IASTED

# THE 6th INTERNATIONAL COMPETITION of Ph.D. STUDENTS

## *2nd Place*

## *Mr. Dimitris KARADIMAS (Greece)*

*The 2007 IASTED International Conference on*
***Web-Based Education (WBE-2007)***
*March 14 – 16, 2007*
*Le Majestic Centre De Congres, Chamonix, France*

*March 15, 2007*

*Prof. Vladimir Uskov*
***WBE-2007 General Chair***

IASTED

# THE 3rd INTERNATIONAL COMPETITION OF NON-COMMERCIAL SOFTWARE FOR WEB-BASED EDUCATION

## *3rd Place*

## *Mr. Dimitris KARADIMAS (Greece)*

*The 2007 IASTED International Conference on*
***Web-Based Education (WBE-2007)***
*March 14 – 16, 2007*
*Le Majestic Centre De Congres, Chamonix, France*

*March 15, 2007*

*Prof. Vladimir Uskov*
WBE-2007 General Chair

# On The Study of a Digitally Calibrated R-2R Ladder Architecture

D. S. Karadimas, K. A. Efstathiou
Dept. of Electrical Engineering & Computer Technology, University of Patras, Greece
karadimas@apel.ee.upatras.gr

*Abstract*—**This paper presents a study on a digitally calibrated DAC, based on a strictly R-2R topology that is able to derive high resolution – high performance DACs, in terms of INL and DNL. It has been proven by simulations that the performance of the conventional R-2R DAC can be optimized, regardless of resistors' tolerance and the DAC resolution.**

## I. Introduction

A Digital to Analog Converter (DAC) is a significant building block that can be found almost in every modern system playing significant role to its overall performance. The quality of the DAC basically depends on the tolerance of the technology used for its implementation and is characterized by the Integral Non-Linearity (INL) and Differential Non-Linearity (DNL) errors as well as by other, second order parameters, such as the Spurious Free Dynamic Range (SFDR), the Signal-to-Noise Ration (SNR) and the conversion time.

High performance, in terms of linearity and high resolution, DACs are implemented using $\Delta\Sigma$ techniques [1], [2] or dynamically calibrated current sources [3], at the cost of slow conversion speed.

High conversion speed can be achieved by using either thermometer code [4]-[6] or ladder DACs [7], [8]. The first architecture has much better linearity at the expense of area and power dissipation, while the second one is characterized by convenient design, low cost and area efficiency.

The combination of the above two architectures derives a hybrid DAC [9], [10] that compromises the linearity and the area tradeoffs regarding with the resolution.

However, high resolution linear DACs can not be achieved by using the above techniques, unless expensive manufactory procedures are employed (laser trimming) [11].

This paper proposes an extension of the conventional R-2R architecture, invoking branches for digital calibration, able to be implemented using all MOS ladders as it is proposed by [12]. The resulted architecture is a strictly based on the R-2R, area efficient, topology that, theoretically, is able to implement high performance, high resolution DACs, in terms of INL and DNL. Additionally, simulations of the proposed architecture and empirical formulas that enable us to forecast the performance of the DAC as a function of the technology tolerance and resolution, are presented.

## II. Properties Of The Conventional R-2R Ladder

A study and simulation results of the R-2R ladder are presented in this paragraph and its performance as a function of resistors' tolerance and the resolution is derived, assuming that there is no resistance difference at the switches.

A conventional $N$ -bits R-2R Ladder is depicted in Fig. 1.

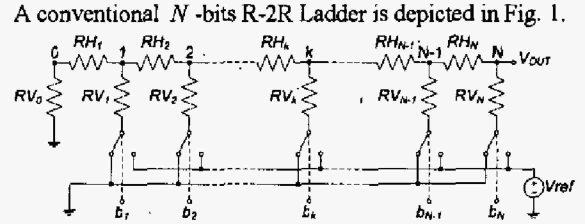



Fig. 1. A conventional $N$ -bits Ladder topology.

We denote $RL_k$ the resistance at the left of the node-$k$ as:

$$RL_k = \begin{cases} RV_0, & k=0 \\ RV_k // (RH_k + RL_{k-1}), & k = 1,2,\ldots,N. \end{cases} \quad (1)$$

Assuming that only the $k$ -bit is active, the ladder's output voltage will be given as follows:

$$V_{OUT}(k) = Vref \cdot Q_k, \quad (2)$$

where:

$$Q_k = \frac{RH_k + RL_{k-1}}{RV_k + (RH_k + RL_{k-1})} \times \prod_{m=k+1}^{N} \frac{RV_m}{RV_m + (RH_m + RL_{m-1})}. \quad (3)$$

Subsequently the output voltage of the ladder, for each possible digital input word $(D)$, due to the contribution of all the bits $(b_1, b_2, \ldots, b_N)$, equals to:

$$V_{OUT}(D) = \sum_{k=1}^{N} b_k V_{OUT}(k), \; D = \sum_{k=1}^{N} b_k 2^{k-1}. \quad (4)$$

Thus, the Thevenin equivalent of a ladder is an ideal voltage source $V_{OUT}(D)$ in series with a resistor $RL_N \cong R$.

Taking into consideration (4) we can calculate the INL using:

$$\text{INL} = \max\left|\frac{V_{OUT}(D) - V_{OUT}(0)}{Vref} - \frac{D}{2^N}\right|, \quad (5)$$

expressed in LSBits, where $0 < D < 2^N - 1$, while the DNL is given by:

$$\mathrm{DNL} = \max\left|\frac{V_{OUT}(D+1) - V_{OUT}(D)}{Vref}2^{N-1} - 1\right|, \quad (6)$$

expressed in LSBits as well, where $0 < D < 2^N - 2$.

Applying the above formulas we simulated the ladder performance for a range of tolerances and resolutions. After statistical processing on the simulated results, the behavior of the INL's mean value and deviation, as a function of resistors' tolerance $T$ and resolution $N$ was derived and are shown in Fig. 2 and 3.

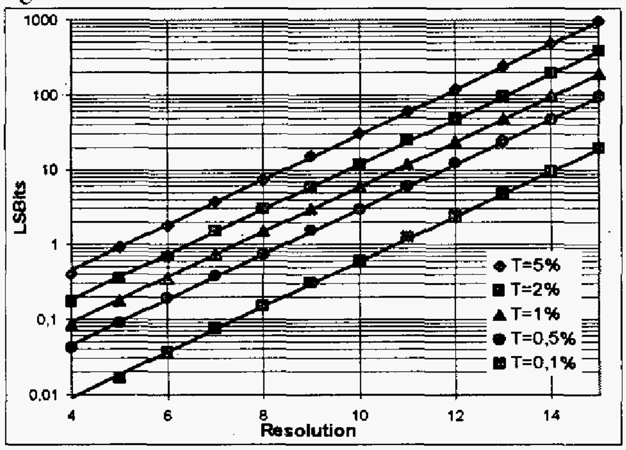



Fig. 2. Mean INL value as a function of resolution for several resistor's tolerances.

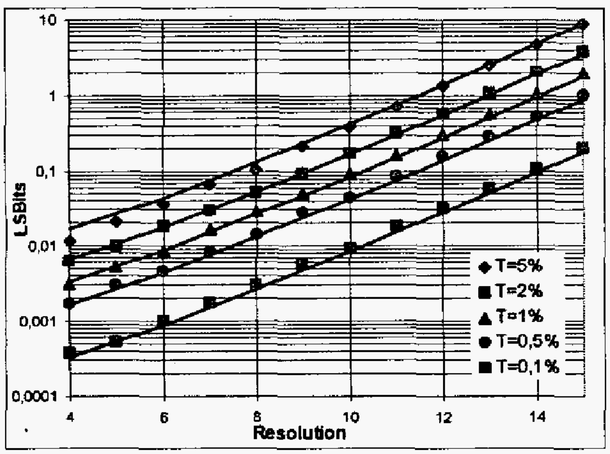



Fig. 3. Deviation of INL as a function of resolution for several resistor's tolerances.

Additionally, we found out that the empirical formulas (7) and (8) fit on the statistical results for the INL mean value and deviation:

$$\overline{\mathrm{INL}}(T,N) = 0{,}577 \cdot T \cdot 2^N \qquad (\mathrm{LSBits}), \quad (7)$$

$$\sigma_{\mathrm{INL}}(T,N) = 0{,}082 \cdot T \cdot \frac{2^N}{N} \qquad (\mathrm{LSBits}). \quad (8)$$

Markers in Fig. 2 depict the INL simulated mean value, while lines depict the graph of (7). Similarly, markers and lines in Fig. 3 depict the INL's deviation according to simulation results and (8), respectively. It can be seen that empirical formulas fit well on the simulated results.

Simulations regarding the DNL exhibited identical statistical results and subsequently, resulted in identical empirical formulas. Equations (9) and (10) are the corresponding formulas for the DNL's mean value and deviation, respectively.

$$\overline{\mathrm{DNL}}(T,N) = 0{,}639 \cdot T \cdot 2^N \qquad (\mathrm{LSBits}), \quad (9)$$

$$\sigma_{\mathrm{DNL}}(T,N) = 0{,}101 \cdot T \cdot \frac{2^N}{N} \qquad (\mathrm{LSBits}). \quad (10)$$

The above empirical formulas could be used to estimate the performance of a ladder, taking into account resistors' mismatch and resolution.

## III. Architecture Of The Digitally Calibrated DAC

Fig. 4 depicts an $NR$-bits DAC of the proposed architecture. It can be seen that an $NR$-bits ladder identical to this in Fig. 1 is employed, where $k$ $(k = 1,2,\ldots,NR)$ vertical resistor $RV_k$ has been substituted by an $R$-valued $R_AV_k$ resistor in series with an $NC$-bits ladder, named as $k$-calibration ladder, of an also $R$-valued output resistance, $R_BV_k$.

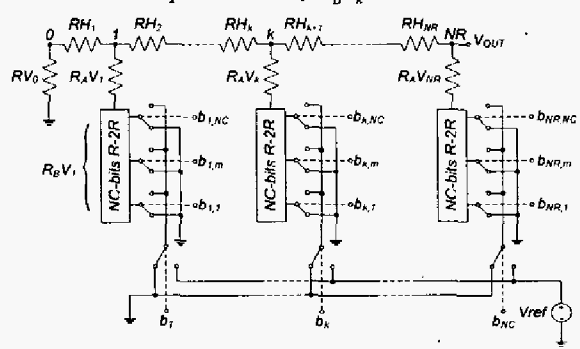



Fig. 4. Architecture of the proposed digitally calibrated Ladder.

In more details, $R_AV_k$ is half the $RV_k$ of the conventional ladder, while $R_BV_k$ represents the output resistance of the $k$-calibration ladder, which is theoretically equal to $R$, as it has been proved in (1). Thus, the functional behavior of the proposed calibrated ladder remains the same with the conventional one, as depicted in Fig. 5.

This topology provides the advantage of trimming the reference voltage $VC_k$ for each resolution bit $k$, by assigning a different digital word (Digital Calibration word – $DC_k$), to the corresponding calibration ladder. Consequently, the contribution of each resolution bit at the output voltage can be adjusted so as to cancel resistors mismatch; meaning that an accurate adjustment of the reference voltage is able even to eliminate the non-linearity behavior.

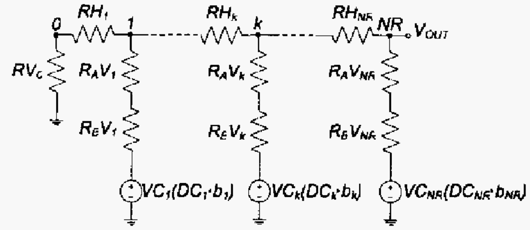



Fig. 5. Thevenin equivalent of the proposed architecture.

Since the reference voltage of each bit can be trimmed with a resolution of $NC$ -bits, the length of each calibration ladder should apparently be selected so as to derive the desired linearity.

## IV. Calibration Principles

The goal of a calibration procedure, regarding the proposed architecture, is the estimation of the $DC_k$ calibration word that forces $k$ -calibration branch to contribute a voltage at the output of the ladder, as close as possible to the ideal. Assuming that $Vmax$ is the maximum output voltage of the DAC, the contribution $V_{OUT}(k)$ of the $k$ -branch, applying (2), should be as close as possible to the value assigned by (11),

$$V_{OUT}(k) = VC_k(DC_k) \cdot Q_k = 2^{k-NR-1} \cdot Vmax. \quad (11)$$

In other words the normalized voltage contribution $V_{NORM}(k)$ of each bit should ideally be the same,

$$V_{NORM}(k) = 2^{NR+1-k} \cdot V_{OUT}(k) = Vmax. \quad (12)$$

Since the MSBit plays the most significant role to the DAC performance, the selection of $Vmax$ should exactly fit to a normalized voltage that MSBit is able to derive applying the calibration word $DC_{NR}$.

$$Vmax = 2 \cdot V_{OUT}(NR) = V_{NORM}(NR) \quad (13)$$

Moreover, the normalized voltage contributions of the rest bits should closely approximate $Vmax$. Thus, $Vmax$ should be less or equal to the minimum normalized voltage contribution regarding all the resolution bits,

$$Vmax \leq \min_{k=1,\dots,NC} \{V_{NORM}(k)\}. \quad (14)$$

$Vmax$ derived by (13) and (14) is the reference voltage of the calibrated ladder in terms of the conventional ladder, since this will be the maximum voltage at the output of the calibrated ladder. $DC_{NR}$ is the calibration word for the calibration branch at the MSBit of the ladder. The calibration words for the least significant calibration branches should be selected so as to comply with

$$VCn_k(DC_k) \cong VC_{NR}(DC_{NR}). \quad (15)$$

Following the above guidelines we simulated the DAC performance of the proposed architecture for various resolutions and for various resolutions of the calibration branches.

## V. Statistical Processing Of The Simulation Results

Several simulations, using formulas described in II and calibration guidelines issued in IV, were applied on this DAC architecture for a range of tolerances, DAC resolutions and calibration resolutions. All simulations were issued using the same length for all the calibration branches. The statistical processing of the simulation results indicates that increasing the resolution of the calibration branches, the linearity of the DAC, in terms of INL and DNL, follows an exponential improvement. Similarly to Fig. 2, Fig. 6 depicts the simulated mean INL value for a tolerance $T = 0.01$ (1%) and for various resolutions of the calibration branches.

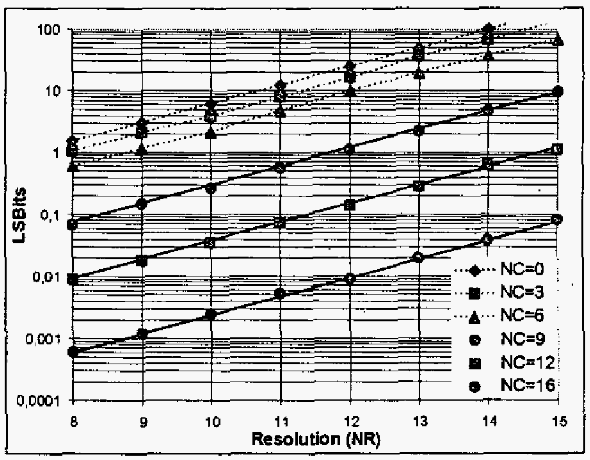



Fig. 6. Mean INL value as a function of DAC resolution (NR) for several calibration resolutions and resistor's tolerance equal to 1%.

Fig. 7 depicts the mean value of INL as a function of the resolution $NC$ of the calibration branches for a tolerance $T = 0.01$ (1%). Apparently, the exponential improvement of the INL is observed for a calibration resolution above 5-bits, whereas below 5-bits the improvement becomes negligible, thus resulting to the break-points depicted in this figure. The cause that generates these break-points is that the calibration algorithm is unable to trim the reference voltages, when the calibration resolution is not adequate, regarding the resistor tolerance. Thus, all the calibration words are forced to be $2^{NC} - 1$.

Simulations proved that the position of a break-point, thus the minimum length required for the calibration ladders in order to be able to improve the DAC linearity, depends on the tolerance of the resistors employed. Empirical inequality (16) defines the minimum number of bits $NC$, required for the calibration ladders, as a function of the tolerance $T$.

$$NC \geq -1.68 - \log_2 T \qquad (LSBits) \quad (16)$$

Evaluating (16) using $T = 0.01$, as is considered in Fig. 7, we derive that $NC$ should be greater than 4.96 bits.

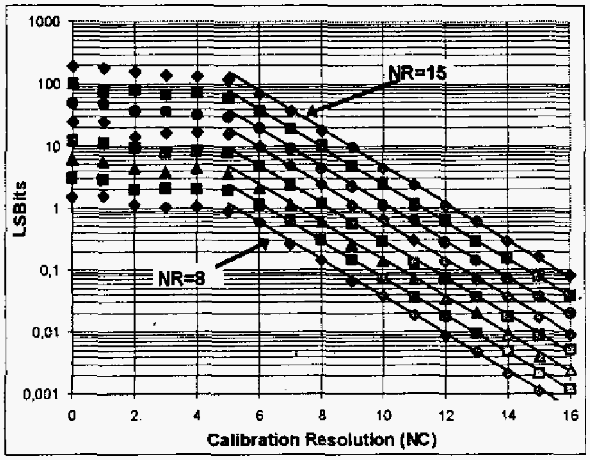



Fig. 7. Mean INL value as a function of calibration resolution for several DAC resolutions (NR)and resistor's tolerance equal to 1%.

The empirical formulas (17) and (18) fit on the statistical results for the INL and DNL mean values, assuming that $NC$ complies with inequality (16).

$$\overline{\mathrm{INL}}(NR,NC,T) = 1.476 \cdot T \cdot 2^{NR-NC} \text{ (LSBits)} \quad (17)$$

$$\overline{\mathrm{DNL}}(NR,NC,T) = 1.566 \cdot T \cdot 2^{NR-NC} \text{ (LSBits)} \quad (18)$$

Thick lines in Fig. 6 and 7 are the plot of (17), while markers represent the statistical values derived by the simulations.

Finally, we can conclude that the linearity of the proposed architecture could also be derived by (7) and (9), on the condition that tolerance $T$ equals to an equivalent tolerance $T_e$

$$T_e = 2.56 \cdot T \cdot 2^{-NC}. \quad (19)$$

## VI. Further Considerations

The proposed architecture demands a significantly reduced area for its implementation, since the total number of components required, is proportional to the square of the resolution length, while the number of components required by a thermometer code DAC is exponentially proportional to its resolution.

The use of the same resolution for all the calibration branches is not obligatory. It has been found by simulations that the least significant bits are not so sensitive to the selection of the calibration word. Thus, less resolution for the LS-branches may be employed without affecting the overall performance. This will also lead to an additional reduction of the component count.

The presented architecture concerns voltage mode DACs; however it can also be applied, as it is, to current mode DACs of similar architecture.

Moreover, the adopted calibration procedure could be further extended in order to provide even better performance. A weighted least square algorithm could be employed for selecting the appropriate calibration words, thus deriving the best matching of the normalized voltages $V_{OUT}(k)$.

Finally, it is of great interest to study the behavior of the proposed DAC when taking into account the switch resistor difference. It is expected that this difference will degrade the performance of the proposed DAC. However, it will still be able to come up with much better INL and DNL characteristics in contrast with those of the conventional one.

## VII. Conclusion

In this paper we presented an area efficient DAC architecture strictly based on the R-2R ladder topology attaining at the same time high performance, high resolution and low cost. Additionally, empirical formulas were derived for estimating the conventional and the digitally calibrated ladders' performance as a function of the technology's tolerance and the desired resolution. Implementation of the presented architecture in a chip and measurements to qualify the real performance of the DAC, in terms of accuracy and speed, is of further work.

## Acknowledgements

The authors acknowledge the support of the "Karatheodoris" research grant awarded by the Research Committee of the University of Patras. Additionally, authors would like to thank Professor G. D. Papadopoulos for his useful comments and suggestions on the issues discussed in this paper.

## References

[1] I. Fujimori and Tetsuro Sugimoto, "A 1.5 V, 4.1 mW dual-channel audio delta-sigma D/A converter," IEEE J. Solid State Circuits, vol. 33, Dec. 1998.

[2] T. Shui, R. Schreier and F. Hudson, "Mismatch shaping for a current-mode multibit delta-sigma DAC," IEEE J. Solid State Circuits, vol. 34, Mar. 1999.

[3] M. P. Tiilikainen, "A 14-bit 1.8-V 20-mW 1-mm² CMOS DAC," IEEE J. Solid State Circuits, vol. 36, Jul. 2001.

[4] C. Lin and K. Bult "A 10-b, 500-msample/s CMOS DAC in 0.6 mm²," IEEE J. Solid State Circuits, vol. 33, Dec. 1998.

[5] A. Bosch, M. Borremans, M. Steyaert and W. Sansen, "A 10-bit 1-Gsample/s Nyquist current-steering CMOS D/A converter," IEEE J. Solid State Circuits, vol. 36, Mar. 2001.

[6] J. Bastos, A. M. Marques, M. Steyaert and W. Sansen, "A 12-bit intrinsic accuracy high-speed CMOS DAC," IEEE J. Solid State Circuits, vol. 33, Dec. 1998.

[7] L. Wang, Y. Fukatsu and K. Watanabe, "Characterization of current-mode CMOS R-2R ladder digital-to-analog converters," IEEE Trans. Instr. and Meas., vol. 50, Dec. 2001.

[8] M. P. Kennedy, "On the robustness of R-2R ladder DAC's," IEEE Trans. Circuits & Syst. II, vol. 47, Feb. 2000.

[9] Y. Cong and R. L. Geiger, "A 1.5 V 14-bit 100ms/s self-calibrated DAC," IEEE J. Solid State Circuits, vol. 38, Dec. 2003.

[10] A. R. Bugeja and B. Song, "A self-trimming 14-b 100-MS/s CMOS DAC," IEEE J. Solid State Circuits, vol. 35, Dec. 2000.

[11] "R/2R ladder networks," Application Note AFD006, International Resistive Company, Inc., 1998.

[12] K. Bult and J. Geelen, "An inherently linear and compact MOST-only current division technique," IEEE J. Solid State Circuits, vol. 27, Dec. 1992.

# A Digitally Calibrated R-2R Ladder Architecture for High Performance Digital-to-Analog Converters

D. S. Karadimas, D.N. Mavridis, K.A. Efstathiou
Dept. of Electrical Engineering & Computer Technology
University of Patras, Greece
{karadimas, mavridis, efstathiou}@apel.ee.upatras.gr

***Abstract*—The paper presents a new model for R-2R ladder-based digital-to-analog converters, developed in terms of resistors' tolerance. The widely used R-2R ladders are easy to be realized as an integrated circuit, while their performance is basically limited due to the resistors' mismatch. However, it is possible to achieve higher linearity, in terms of Integral and Differential Non-Linearity errors, if calibration is employed. In this paper, the conventional R-2R ladder architecture is briefly discussed, and then a new digitally calibrated architecture is presented with its calibration algorithm, based on the resistors' tolerance formulation. It has been proven by simulations that the performance of an R-2R-based digital-to-analog converter can be optimized regardless of the resistors' tolerance and the required resolution.**

## I. INTRODUCTION

The performance of a digital-to-analog converter (DAC) plays significant role to almost every modern system design. One of the most popular methods for converting a digital word into a quantized analog signal is based on R-2R ladder networks. These ladders have the significant advantage of being easily integrated onto a small silicon area, as they require a small device count, and therefore constitute key components in mixed analog and digital ASICs.

Nowadays digital-to-analog conversion demands high performance in terms of linearity, resolution and speed conversion, which are not easy to be straightforwardly implemented. Although R-2R ladder-based DACs and especially current-output ladders can inherently operate as high-speed current dividers [1], they are not able to meet the linearity and resolution requirements.

ΔΣ techniques [2] or dynamically calibrated current sources [3] are mainly employed to provide high performance, in terms of linearity and high resolution, but at the cost of slow conversion speed. On the other hand, high conversion speed with high performance can be achieved using either thermometer code DACs [4], at the expense of area and power dissipation, or laser-trimmed R-2R ladder-based DACs, using expensive manufactory procedures [5]. Another compensating method includes hybrid DACs [6] or self-calibrated R-2R ladders [7].

The motivation of the work is to increase the linearity, in terms of Integral/Differential Non-Linearity (INL/DNL) errors and resolution of the conventional R-2R ladder, concluding in a high performance, easily integrated DAC architecture. The proposed architecture extends the conventional ladder with R-2R branches for digital calibration and concludes in a strictly R-2R-based, area efficient topology, able to implement high performance, high resolution DACs.

## II. PRELIMINARIES

Fig. 1 shows the conventional R-2R ladder DAC. $V_{ref}$ is the reference voltage and $b_k$ (k=1…N) is the N-bit digital word to be converted into the analog voltage.

Assuming that all resistors are perfectly matched and the switches are ideal, the output of the ladder will be

$$V_{out} = \left(b_1 2^{-N} + \ldots + b_N 2^{-1}\right) V_{ref} = V_{ref} \sum_{k=1}^{N} b_k 2^{k-N-1}, \quad (1)$$

since the impedance seen from the k-node to the left is 2R and the reference voltage switched by $b_k$ is attenuated by $2^{-k}$ to reach the output. The contribution of the k-bit to the output voltage is

$$V_{out}(k) = V_{ref} Q_k, \ k = 1..N, \quad (2)$$

where

$$Q_k = \frac{RH_k + RL_{k-1}}{RV_k + \left(RH_k + RL_{k-1}\right)} \prod_{m=k+1}^{N} \frac{RV_m}{RV_m + \left(RH_m + RL_{m-1}\right)} \quad (3)$$

and

$$RL_k = \begin{cases} RV_0, & k=0 \\ RV_k // \left(RH_k + RL_{k-1}\right), & k=1\ldots N. \end{cases} \quad (4)$$

Now the output of the ladder will be given by

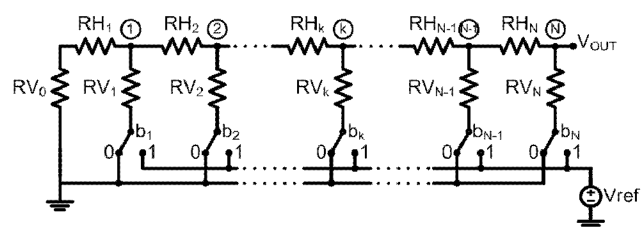



Figure 1. Conventional R-2R ladder DAC.

$$V_{out}=\sum_{k=1}^{N} V_{out}(k)=\sum_{k=1}^{N} Q_k V_{ref}\,. \qquad (5)$$

Simulations for a range of resistors' tolerance (T) and resolutions (N), followed by statistical analysis based on the above formulation of an R-2R ladder network, proved that the INL's and DNL's mean value can be approximated by

$$\overline{\mathrm{INL}}(\mathrm{T,N})=0.577\,\mathrm{T}\,2^{\mathrm{N}}\,(\mathrm{LSBits}) \qquad (6)$$

$$\overline{\mathrm{DNL}}(\mathrm{T,N})=0.639\,\mathrm{T}\,2^{\mathrm{N}}\,(\mathrm{LSBits})\,, \qquad (7)$$

while their deviation has been expressed by similar formulas. These empirical formulas enable the forecast of a ladder's linearity given the resistors' tolerance and resolution.

## III. Architecture Of The Digitally Calibrated R-2R Ladder

Equation (5) points out that the trimming of the reference voltage for each bit is able to equalize the error of the $Q_k$ factors, introduced by the resistors' mismatch, in such a way that the voltage contribution of each bit to the output voltage is as close as possible to the optimum value.

Based on the above ascertainment, the proposed architecture has been designed so as to enable the trimming of the reference voltage for each resolution bit. Fig. 2 depicts an NR-bits ladder of the proposed architecture. A conventional NR-bits ladder is employed, where each of the NR-vertical resistors $RV_k$ (k=1…NR) has been replaced by an in series combination of an R-valued resistor $R_AV_k$ and an NC-bits ladder, named as k-calibration ladder.

$R_AV_k$ is R-valued, while $R_BV_k$ represents the output resistance of the k-calibration ladder, which is also R-valued (4). Thus, the proposed architecture retains the functional principle of the voltage/current division from the k-node to the left.

The Thevenin equivalent of the proposed architecture, shown in Fig. 3, indicates that the reference voltage $VC_k$, for each resolution bit k, can be trimmed by assigning a different Digital Calibration word ($DC_k$) to the corresponding calibration ladder. The proper selection of each $DC_k$-word is

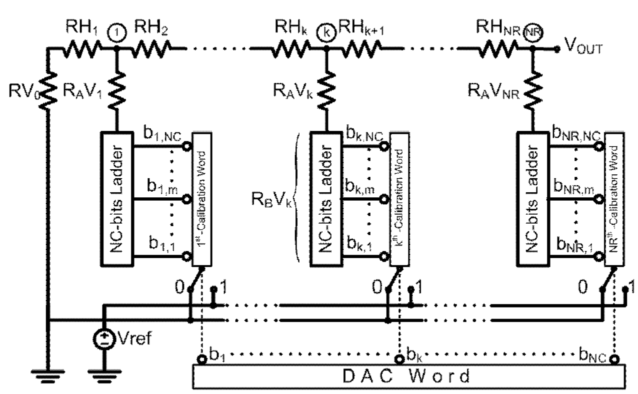



Figure 2. Architecture of the proposed digitally calibrated ladder.

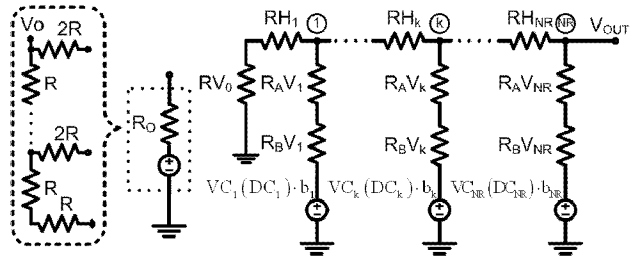



Figure 3. Thevenin equivalent of the proposed architecture.

able to adjust the contribution of the corresponding resolution bit at the output voltage of the ladder; so as to cancel non-linearity effects of resistors' mismatch. Thus, a one-time accurate adjustment of the reference voltage, for each resolution bit, can eliminate permanently the inherently non-linear behavior of the R-2R ladder.

Since the reference voltage of each resolution bit can be trimmed at $2^{NC}$ different voltage values, the selection of the length NC for each calibration ladder is a major issue that can be encountered by the designer. Many policies could be applied; stepwise, linear or even logarithmic lowering of the calibration length for the low order resolution bits are possible solutions for even more area effective designs. The calibration technique, described in the next section, is independent of the length of each calibration branch. However, all simulations were performed assuming identical NC for all the calibration branches.

## IV. Calibration

The goal of a calibration procedure, taking into account the proposed architecture, is the estimation of the $DC_k$ calibration word that forces k-calibration ladder to contribute a voltage at the output, as close as possible to the ideal. Assuming that $V_{Full\ Scale}$ is the maximum output voltage of the ladder and applying (2), the contribution $V_{OUT}(k)$ of the k-branch (k=1…NR) should be as close as possible to the value

$$V_{OUT}(k)=VC_k(DC_k)Q_k \approx 2^{k\text{-}NR\text{-}1}V_{Full\ Scale}\,. \qquad (8)$$

Obviously, the normalized voltage contribution $V_{NORM}(k)$ for all resolutions bits should ideally have the same value

$$V_{NORM}(k)=2^{NR+1-k}V_{OUT}(k) \approx V_{Full\ Scale}\,. \qquad (9)$$

Since the MSBit plays the most significant role to the ladder linearity, the selection of $V_{Full\ Scale}$ should fit exactly to a normalized voltage that MSBit is able to produce, applying the calibration word $DC_{NR}$.

$$V_{Full\ Scale} \equiv 2V_{OUT}(NR) \equiv V_{NORM}(NR)\,. \qquad (10)$$

Additionally, the normalized reference voltages that the rest of the resolution bits (except for the MSBit) can produce should closely approximate $V_{Full\ Scale}$. Thus, the possible $V_{Full\ Scale}$ from (10) should also be less or equal to the minimum normalized reference voltage, taking into account all the resolution bits, apart from the MSBit (k=1…NR-1),

$$V_{\text{Full Scale}} \leq \min\left\{2^{NR+1-k} VC_k \left(2^{NC} - 1\right)\right\}. \quad (11)$$

$V_{\text{Full Scale}}$, derived by (10) and (11), is the reference voltage of the calibrated ladder in terms of a conventional ladder, since this will be the maximum voltage at its output. $V_{\text{Full Scale}}$ will not be equal to Vcc; yet this is not a great disadvantage, as DAC design emphasizes more on its linearity rather than on its output gain, which can be externally adjusted.

Taking into account the resulted $V_{\text{Full Scale}}$ and applying k=NR in (8), the calibration word for the most significant calibration ladder ($DC_{NR}$) is determined, while all the other calibration words (k=1…NR-1) should be selected using the criterion

$$\left|2^{NR-k} VC_k\left(DC_k\right) Q_k - V_{OUT}\left(NR\right)\right| \rightarrow 0. \quad (12)$$

The above criterion constitutes the basic requirement for the discussed calibration algorithm, as the trimming of the reference voltages, managed by the calibration ladders, targets at the accomplishment of the above criterion, so as the remained error from the NR-1 voltage differences results in extensively decreased non-linearity effects.

## V. Error Considerations Of The Calibration Algorithm

The presented calibration algorithm takes into account only the resistors' tolerance. Thus, the analysis is limited to the basis of the resistors' tolerance and resolution as well. Apparently, the formulation of a ladder is a more difficult task, when more error sources are involved. An exhaustive, but complete, analysis of the ladder network should also take into consideration the non-linear behavior of the switches and the $R_{ON}$ effect of a MOST-only implementation as well.

Nevertheless, the adopted calibration procedure could be further extended, if weighted least square root algorithm would be employed for selecting the proper calibration words, thus deriving a further more better matching of the normalized voltage outputs than (12) dictates.

Finally, in practice, a measurement device should be utilized in order to determine the best resolution bit voltage contribution at the output of the ladder. Since every measurement device has finite accuracy, another error source is added. However, simulation results, presented in the next section, prove that the proposed architecture is able to come up with improved characteristics, in terms of INL, DNL and resolution, over against the conventional ladder.

## VI. Simulation Results

Several simulations for the digitally calibrated ladder architecture were performed using formulas described in II and the calibration algorithm discussed in IV, for a range of resistors' tolerance (T=0.1%, 0.2%, …,2.0%), resolutions (NR=8, 9, …, 15) and calibration lengths (NC=1, 2, …, 16).

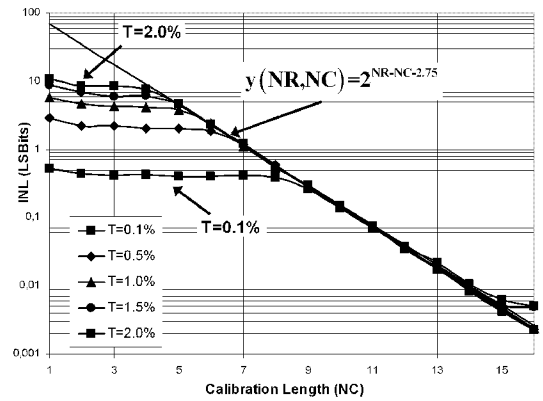



Figure 4. Simulated mean INL value for 10-bits resolution at various resistors' tolerance.

The above simulations were transacted assuming the same length for all the calibration ladders.

Fig. 4 depicts the simulated mean INL value as a function of the calibration length (NC) for various resistors' tolerance (T), assuming 10-bits resolution (NR). A particular analysis on the simulation results implies that the proposed ladder's behavior is twofold:

- Calibration ladders are not inherently able to offer a satisfying reference voltage trimming, when their length is not adequate. Thus, the employed calibration algorithm forces all calibration words to the certain value of $2^{NC}$-1 and consequently ladder's linearity is only affected by the resistors' tolerance; which is the case of the conventional ladder. Any negligible improvement may take place only because of the more complex resistor network that the proposed architecture utilizes.
- When the calibration length is adequate, calibration ladders achieve an exponentially improved linearity due to the reference voltage trimming. It is remarkable that, in this case, ladder's linearity depends only on the calibration length, independently of the resistors' tolerance; at least in the range of the performed simulations. Diagonal line ($2^{NR-NC-2.75}$) in Fig. 4 fits on the INL exponential improvement due to the corresponding calibration length.

This remarkable fact can be intuitively explained taking

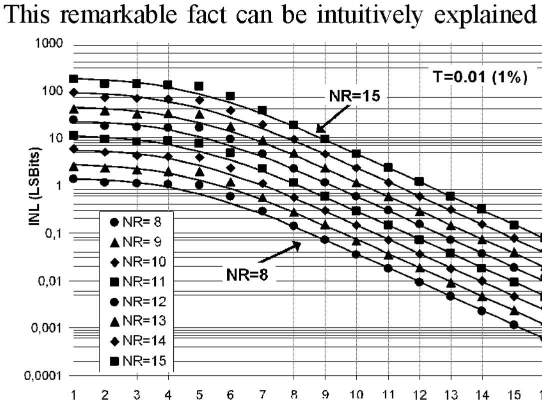



Figure 5. Simulated INL mean value for resistors' tolerance T=0.01 (1%) for various resolutions.

into consideration that the linear increment of the length for each calibration ladder constitutes an exponentially increased possibility to accomplish the criterion asserted by (12). This results in an exponentially improved linearity of the proposed ladder, since an exponentially improved matching between the normalized voltages of each resolution bit and the MSBit occurs.

Further statistical analysis of the above simulation results indicates that the predictable exponential improvement in ladder's linearity is forced when the length of the calibration ladders complies with

$$NC \geq 7.887\text{-}NR\text{-}\log_2(T), \qquad (13)$$

while the behavior of the proposed ladder, regarding the INL and DNL mean value, can be compactly approximated by

$$INL(NR,T,NC) = \frac{0.577\,T\,2^{NR}}{1+0.577\,T\,2^{NC+2.75}}, \qquad (14)$$

$$DNL(NR,T,NC) = \frac{0.639\,T\,2^{NR}}{1+0.639\,T\,2^{NC+2.75}}. \qquad (15)$$

Lines in Fig. 5 represent the plot of (14) for various resolutions NR, while markers stand for the corresponding simulated mean INL value. Apparently, (14) is able to predict with high precision the linearity of the digitally calibrated ladder. However, it should be emphasized that (14) and (15) comply with the results within the range of the performed simulations. It is of further work to confirm or modify these empirical formulas for a wider range of resistors' tolerance and calibration lengths.

## VII. Characterization Of The Digitally Calibrated Ladder

An area comparison, in relation with common ladder topologies, via the required component count, can be performed as depicted in Fig. 6. The proposed architecture requires a quite suppressed total component count in relation with the popular R-String DAC for implementing high resolution DACs, while at the same time an improved overall performance is attained.

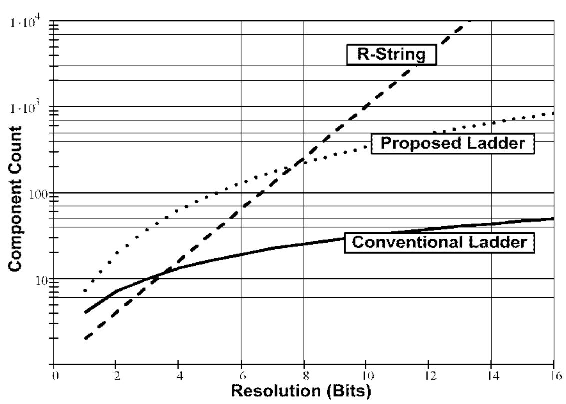



Figure 6. Component count comparison for 3 DAC topologies.

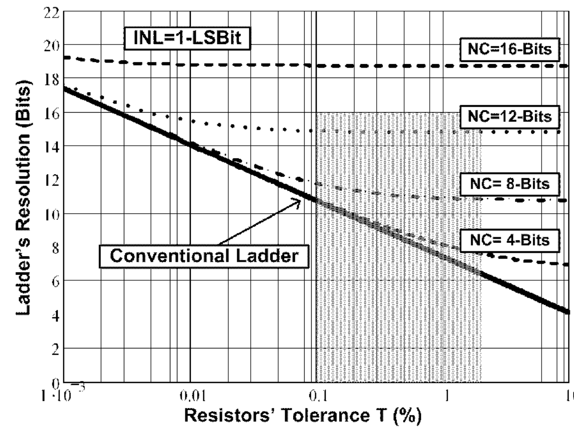



Figure 7. Conventional and the proposed ladder's performance.

Additionally, the performance of the proposed and the conventional ladder can be straightly correlated with the resolution capabilities, on demand of the crucial value of 1-LSBit INL, as depicted in Fig. 7. Thick line in Fig. 7 depicts the inherent performance limitation of the conventional ladder, while the discontinuous lines present the effect of the calibration technique for various calibration lengths. Shadowed area indicates the simulation range.

## VIII. Conclusions

The proposed, strictly based on the R-2R topology, area efficient ladder architecture enables the design of high resolution DACs, attaining at the same time high linearity, regardless of the resistors' tolerance at a quite rational area expense. Empirical formulas were derived for estimating the digitally calibrated ladder's performance, as a function of the desired resolution and the calibration length. Additionally, an algorithm for the required one-time calibration has been presented in details.

## Acknowledgment

The paper was supported by the "Karatheodoris" research grant, awarded by the Research Committee of the University of Patras.

## References

[1] B. Vargha, I. Zoltan, "Calibration algorithm for current-output R-2R ladders," IEEE Trans. Instr. and Meas., vol. 50, Oct. 2001.

[2] T. Shui, R. Schreier and F. Hudson, "Mismatch shaping for a current-mode multibit delta-sigma DAC," IEEE J. Solid State Circuits, vol. 34, Mar. 1999.

[3] M. P. Tiilikainen, "A 14-bit 1.8-V 20-mW 1-mm$^2$ CMOS DAC," IEEE J. Solid State Circuits, vol. 36, Jul. 2001.

[4] J. Bastos, A. M. Marques, M. Steyaert and W. Sansen, "A 12-bit intrinsic accuracy high-speed CMOS DAC," IEEE J. Solid State Circuits, vol. 33, Dec. 1998.

[5] "R/2R ladder networks," Application Note AFD006, International Resistive Company, Inc., 1998.

[6] Y. Cong and R. L. Geiger, "A 1.5 V 14-bit 100ms/s self-calibrated DAC," IEEE J. Solid State Circuits, vol. 38, Dec. 2003.

[7] M. P. Kennedy, "On the robustness of R-2R ladder DAC's," IEEE Trans. Circuits & Syst. II, vol. 47, Feb. 2000.

# Linear Range Extension of a Phase-Frequency-Detector with Saturated Output

Michail Papamichail, Dimitrios Karadimas, Konstantinos Efstathiou and George Papadopoulos
Department of Electrical and Computer Engineering, University of Patras
Rion Campus, 26500, Patras, Greece
mpapamicha@apel.ee.upatras.gr

*Abstract*—**A Phase Frequency Detector (PFD) with wide linear operating range and the ability to saturate beyond that range is presented in this paper. Its use in phase locked loops (PLLs) implies elimination of cycle slipping, which results in an analytically predictable transient behavior and a possible improvement in acquisition speed, depending on the PLL design. Simulation results confirm these advantages. Moreover, a detailed study on the applicability of this PFD is conducted in this paper.**

## I. INTRODUCTION

The Phase Detector (PD) constitutes a critical component of Phase Locked Loops (PLLs). The PD having prevailed in PLL applications is the Phase Frequency Detector (PFD). However, limitations in the width of its linear operating range, which occasionally have impact on the acquisition speed of the PLL, have encouraged many designers to propose PDs that present superior properties. The accumulative PD proposed in [8], for example, achieves wider linear operating range at the cost of a significant circuitry increase. Most other techniques aim to improve acquisition speed and employ nonlinear PDs, or adaptive PDs that control the gain or the bandwidth of the loop [9,10,11,12]. In [9] a more detailed survey of these techniques can be found.

This paper proposes a PD based on the conventional PFD, modified so as to obtain wide linear operating range, while at the same time eliminating the cycle slipping effect. Employing the proposed PFD in a PLL improves the acquisition speed in wide transients, without affecting its operation in steady state.

Section II of this paper studies the limitations of the conventional PFD. Section III presents the proposed PFD architecture and its operation. Section IV studies the proposed PFD properties and applications and, finally, Section V summarizes the results.

## II. CONVENTIONAL PFD - LIMITATIONS

Conventional PFD (C-PFD) is a sequential logic circuit which detects the phase difference of two signals, while at the same time provides a frequency sensitive signal aiding acquisition when the loop is out of lock. A simple implementation of PFD and its state transition diagram (STD) are shown in Figure 1 (a) and (b).

The digital outputs of the PFD are converted into analog quantity by using a charge pump, and are then passed through the PLL's loop filter that drives the VCO. The transfer characteristic of C-PFD is illustrated in Figure 1 (c).

It can be seen that the range of linear operation is ideally $\pm 2\pi$. For greater phase difference between the input signals PFD behaves nonlinearly, a fact that deteriorates the acquisition process of the loop. Figure 2 depicts a simulation example of the acquisition process of a PLL, where the phase difference of the two input signals exceeds $2\pi$. As it can be seen, at time $\omega_n t_{2\pi}=1.111$ a cycle slipping occurs and the response of the loop diverges from the expected.

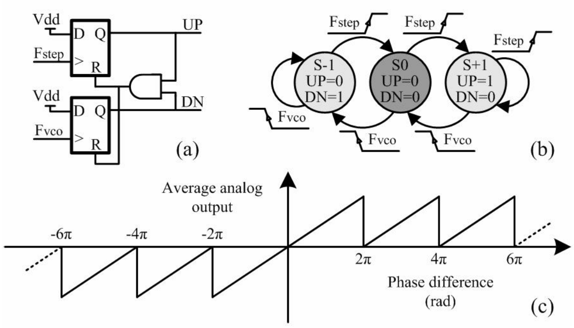



Figure 1. Conventional PFD. (a) Circuit implementation, (b) State Transition Diagram, (c) Transfer characteristic.

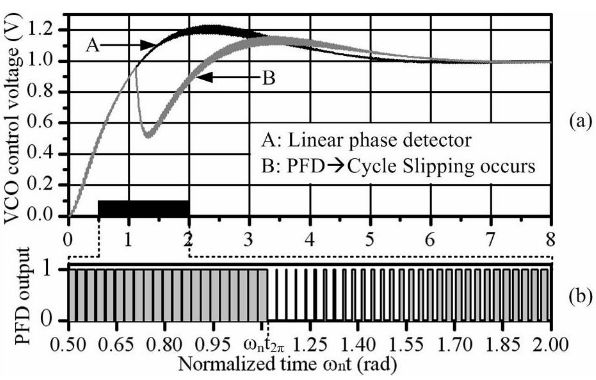



Figure 2. Transient behavior of a PLL upon a frequecy step. (a) VCO control voltage (Lowpass filtered), (b) PFD output (zoomed): at $\omega_n t_{2\pi}$ the duty cycle erroneously resets to zero.

Cycle slips in general impose a more sluggish transient behaviour to the PLL, which cannot be analytically predicted with linear techniques. The implications become more severe as the frequency step increases or the natural frequency of the loop ($\omega_n$) decreases or the damping factor ($\zeta$) decreases.

## III. THE PROPOSED PD ARCHITECTURE

As is established in section II, cycle slipping of C-PFD deteriorates the acquisition process of the PLL. A PD with a transfer characteristic that saturates for phase errors beyond $\pm 2\pi$, as the one depicted in Figure 3 (a), could improve the acquisition speed by eliminating cycle slips. When phase error exceeds $\pm 2\pi$, such a PD, the saturated PFD (S-PFD), drives the loop filter with the maximum output, thus forces VCO to converge to the proper frequency with the maximum speed. Figure 3 (b) illustrates the state transition diagram (STD) of S-PFD. However, in the case of S-PFD the transient behaviour of the PLL cannot still be expressed in a closed form.

It is apparent that a linear transfer characteristic exceeding the $\pm 2\pi$ limitation would further increase the acquisition speed in wide frequency steps, making at the same time the behavior of the loop predictable. Figure 4 (a) illustrates the transfer characteristic of such a phase detector that is able to operate linearly in a range of $\pm 2n\pi$, while exhibiting saturation beyond this range. The STD of this Wide-range Saturated PFD (WS-PFD) is shown in Figure 4 (b).

The operation of the STDs depicted in Figure 3 (b) and Figure 4 (b), is quite simple and is based on accumulating UP/DN levels according to the count of the edges of the Fstep/Fvco signals, i.e. on each edge of Fstep/Fvco the count of UP/DN increases by one or the count of DN/UP decreases by one, depending on the current state of the STD. Notation UP=0 means that all UP outputs equal zero, while notation $UP^n=1$ means UP1=UP2=…=UPn=1.

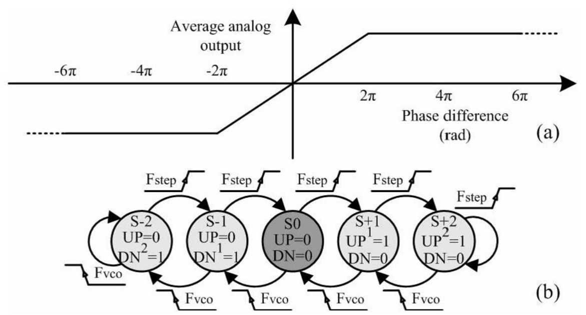



Figure 3. Saturated-PFD. (a) Transfer characteristic, (b) State Transition Diagram.

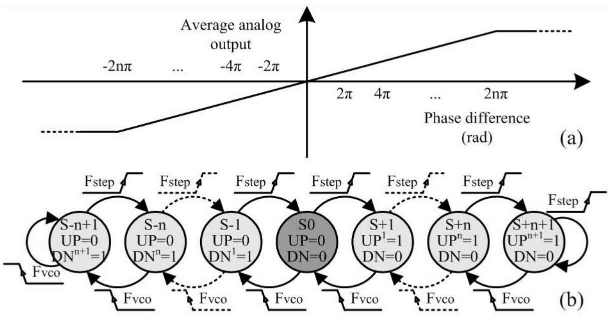



Figure 4. n-level Wide-range Saturated-PFD. (a) Transfer characteristic, (b) State Transition Diagram.

The implementation of the above STDs can be simplified by using the two cells depicted in Figure 5 (a) and (b), while the block diagrams of S-PFD and n-level WS-PFD are depicted in Figure 5 (c) and (d) respectively. The charge pump of the proposed PD is identical to the C-PFD's one. Each UP/DN output drives an equally sized pmos/nmos current source that sources/sinks a predefined amount of current. It should be denoted that WS-PFD's (S-PFD's) outputs UPn+1, DNn+1 (UP2, DN2) are ignored in order to obtain saturation. Additionally, it should be emphasized that in the phase locked condition, or generally for phase differences less than $\pm 2\pi$, the active states are S0, S+1, S-1 and the operation of S-PFD and WS-PFD is identical to that of C-PFD. Therefore, the power consumption of the proposed PFDs can be considered almost the same as the C-PFD's one, since the additional circuitry is inactive during the locked state. Moreover, the dead zone issue appears only during the locked state and thus, design techniques for the C-PFD that handle this problem can be applied to the proposed PFDs as well.

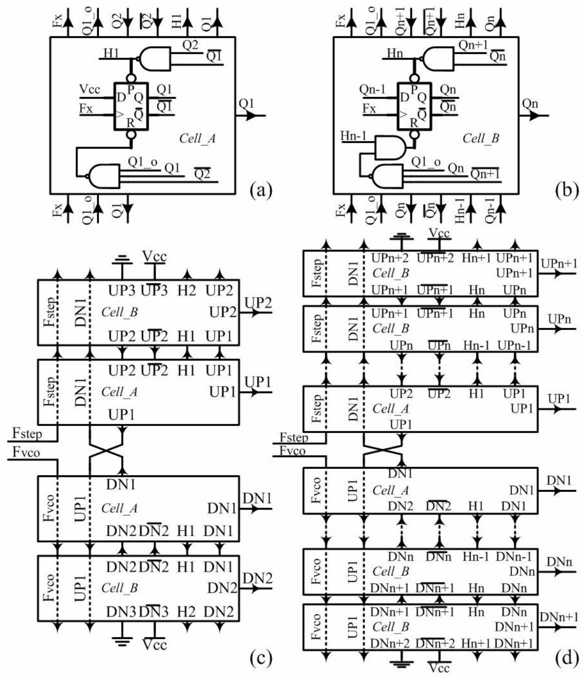



Figure 5. Circuit implementation of the proposed phase detectors. (a) Cell_A, (b) Cell_B, (c) S-PFD, (d) n-level WS-PFD.

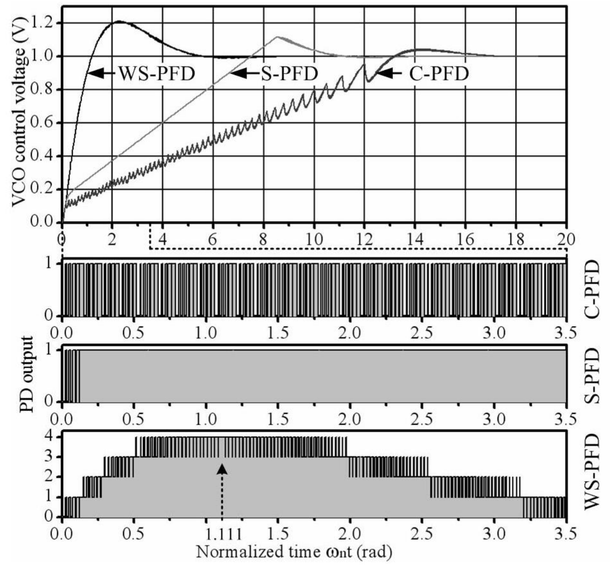



Figure 6. Transient response of a PLL for three different PDs. (The response of the PDs is zoomed)

The transient response of a PLL incorporating C,S-PFDs and a 4 level WS-PFD, all with identical $K_d$ gains, was simulated using Matlab. A type-II, second-order PLL with $\omega_n = 2\pi 1000$ rad/sec and $\zeta = 0.707$ was tested. Figure 6 shows the transient response for a frequency step input of $\Delta f$=55.123kHz ($f_{vco\_free\_run}$=200kHz, $f_{step}$=255.123kHz) for the three types of PFDs. It is clear that no cycle slipping occurs when S-PFD is used as the phase detector, resulting to a 57% faster settling than the C-PFD's one. Regarding to WS-PFD, which does not saturate, it is clear that it presents a typical type II second order step response.

Since the transfer function, and thus $\omega_n$ and $\zeta$, are the same for all PLLs, there is no obvious penalty of phase noise. Indeed, the phase noise power spectral density $S_\varphi(f)$, as is given in [3], is the same for PLLs with the same transfer function and the same VCO gain $K_O$, i.e. the noise coming from the VCO, the reference, the loop filter and the feedback divider remains the same. In [4], the noise contribution due to the phase detector is also considered, and is found to increase for lower PD gain. Unless it is not possible to use a greater value for the full scale current than the one used in C-PFD, the noise due to the PD will not change either.

## IV. Applications of the Proposed PFDs

Incorporation of S-PFD or WS-PFD in a PLL experiencing frequency steps is of advantage when the transient behavior of the loop suffers from cycle slipping. In order to find out which PLL designs could benefit from the proposed phase detector architecture we have to investigate the behavior of the maximum phase error developed during the acquisition process upon the width of the frequency step and upon physical quantities of the loop such as the natural frequency ($\omega_n$) and the damping factor ($\zeta$). For all those PLL applications where the maximum phase error exceeds $\pm 2\pi$ during frequency steps, use of the proposed phase detector seems to be beneficial.

For the purposes of the analysis, a second order PLL having the transfer function in (1) will be considered.

$$H(s) = \frac{2\zeta\omega_n s + \omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (1)$$

This is the transfer function of PLLs with voltage output PD that use active PI filter [1, 2] and of PLLs with current output PD that use first order passive RC filter [1]. It is also approximately the transfer function of most practical high gain second order PLLs with voltage output PD [2]. The error transfer function is shown in (2).

$$H_e(s) = \frac{s^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (2)$$

For a frequency step $\Delta\omega$ applied to the input at $t=0$, the phase error signal is given in Laplace form in (3) [2].

$$\Theta_e(s) = \frac{s^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \cdot \frac{\Delta\omega}{s^2} \quad (3)$$

Applying the inverse Laplace transform to (3) we derive the phase error signal in time domain, given in (4) for $\zeta \neq 1$.

$$\vartheta_e(t) = \Delta\omega \cdot \frac{e^{-\zeta\omega_n t}}{\omega_n\sqrt{1-\zeta^2}} \cdot \sin(\omega_n\sqrt{1-\zeta^2}\,t) \quad (4)$$

The time where the maximum phase error occurs can be found by solving the equation $d\theta_e(t)/dt=0$ and is given in (5). This time depends only on $\omega_n$ and $\zeta$.

$$t_{\max}(\omega_n,\zeta) = \frac{\arctan\left(\frac{\sqrt{1-\zeta^2}}{\zeta}\right)}{\omega_n\sqrt{1-\zeta^2}} \quad (5)$$

Evaluating $\theta_e(t)$ at $t_{max}$, the maximum phase error $\theta_{e\_MAX}$ as a function of $\Delta\omega$, $\omega_n$ and $\zeta$, is derived:

$$\vartheta_{e\_\mathrm{MAX}}(\Delta\omega,\omega_n,\zeta) = \frac{\Delta\omega}{\omega_n} \cdot \frac{e^{\frac{\arctan\left(\frac{-\sqrt{1-\zeta^2}}{\zeta}\right)}{\frac{\sqrt{1-\zeta^2}}{\zeta}}}}{\sqrt{1-\zeta^2}} \cdot \sin\left(\arctan\left(\frac{\sqrt{1-\zeta^2}}{\zeta}\right)\right) \quad (6)$$

Which can be simplified to:

$$\vartheta_{e\_\mathrm{MAX}}(\Delta\omega,\omega_n,\zeta) = \frac{\Delta\omega}{\omega_n} \cdot C(\zeta) \quad (7)$$

Or it can be approximated to:

$$\vartheta_{e\_\mathrm{MAX}}(\Delta\omega,\omega_n,\zeta) \cong \frac{\Delta\omega}{\omega_n} \cdot \frac{1}{1.767\cdot\zeta + 0.960} \quad \text{for } \zeta \text{ in } (0,1.5] \quad (8)$$

Figure 7 illustrates the PLL's cycle slipping free operating areas for several levels of WS-PFDs. The area under any curve constitutes the locus of the operating points of the PLL in which it is safe to use a phase detector with a $\pm 2n\pi$ rad linear range without experiencing cycle slipping.

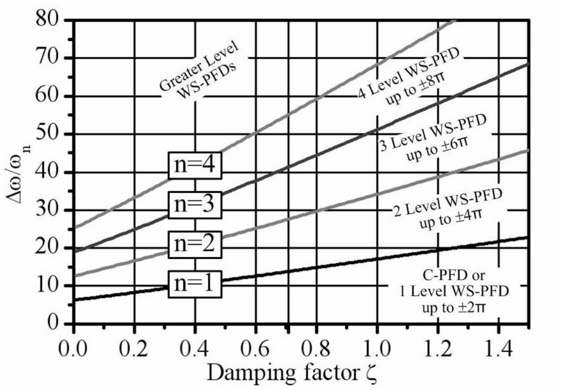



Figure 7. Cycle slipping free operational areas of a PLL with various PFDs.

In order to verify the results, we consider the PLL simulation in Figure 6, where $\Delta\omega/\omega_n=55.123$ (actually just over 55.123) and $\zeta=0.707$. According to Figure 7, a 4 level WS-PFD is marginally adequate. Indeed, in the simulation in Figure 6, we can see that at $\omega_n t$=1.111 the phase detector output saturates for one cycle (pointed by an arrow) because the phase difference marginally exceeds $8\pi$. Furthermore, for $\zeta=0.707$ (5) gives normalized time that equals 1.111.

In the case of an Integer N frequency synthesizer, the feedback loop incorporates a frequency divider with divisor values in the region $[N_{min}, N_{max}]$. The maximum frequency difference in the phase detector inputs is observed during the $F_{max}$ to $F_{min}$ transition and is equal to $(F_{max} - F_{min})/N_{min}$. Substituting this difference in (7) and setting $\theta_{e\_MAX}=2n\pi$ we derive (9).

$$\frac{2\cdot\pi\cdot\frac{F_{\max}-F_{\min}}{N_{\min}}}{2\cdot\pi\cdot f_n}=\frac{2\cdot n\cdot\pi}{C(\zeta)} \tag{9}$$

Taking into account that $f_{step}=F_{min}/N_{min}$ we derive (10) which is plotted in Figure 8 for $\omega_n/\omega_{step}=1/100$ and various values of $n$.

$$\frac{F_{\max}}{F_{\min}}(\zeta)=\frac{2\cdot n\cdot\pi}{C(\zeta)}\cdot\frac{f_n}{f_{step}}+1 \tag{10}$$

Every curve denotes the maximum $F_{max}/F_{min}$ ratio of a synthesizer incorporating an n-level WS-PFD that ensures cycle slipping free operation of the loop for a particular value of $\omega_n/\omega_{step}$ and for various values of $\zeta$. Figure 9 compares the maximum permissible $F_{max}/F_{min}$ of the C-PFD and of WS-PFDs for loops with $\zeta=0.707$ and $\omega_n/\omega_{step}$ in the range [1/200, 1/20].

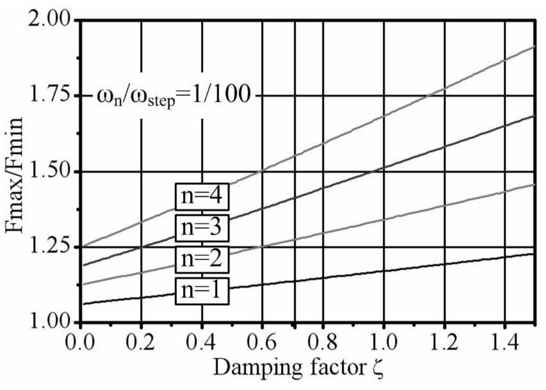



Figure 8. Fmax over Fmin ratio of a synthesizer with various PFDs for cycle slipping free operation.

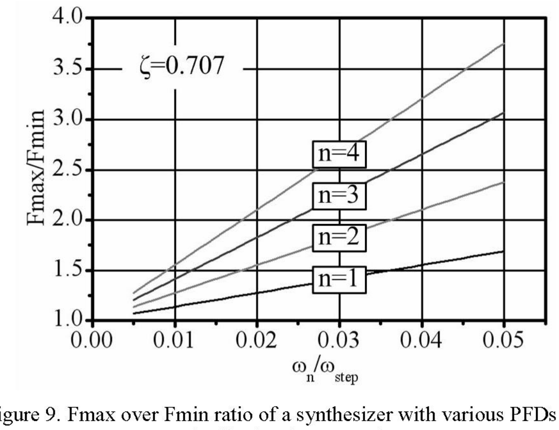



Figure 9. Fmax over Fmin ratio of a synthesizer with various PFDs for cycle slipping free operation.

As a design rule, the ratio $\omega_n/\omega_{step}$ should not exceed 1/10 [5, 6, 7] to avoid stability issues due to the discrete nature of the PFD. In conventional applications it is selected as 1/100, or even smaller [6, 7]. For loops with $\omega_n/\omega_{step}=1/100$, $\zeta=0.707$ and $n=1$ (C-PFD) the $F_{max}/F_{min}$ ratio should not exceed 1.13781, or else cycle slipping will occur. It is evident that PLLs used in applications where wide bandwidth coverage is required, could benefit significantly by incorporating a WS-PFD.

## V. Conclusion

A phase detector with a wide linear operating range, exceeding the $\pm 2\pi$ limitation, and a saturation mechanism was presented in this paper. The so called WS-PFD, unlike the conventional PFD, can provide to a PLL all the advantages that linear theory predicts, i.e. cycle slipping free operation and a predefined settling time that does not depend on the frequency step applied to the PLL input. Applications involving wide frequency coverage or multiband operation, could undoubtedly take advantage by incorporating a WS-PFD in the PLL, which can replace C-PFD without the need of redesigning the loop. All the above, in conjunction with the relatively low circuit complexity and the anticipated low power consumption, render WS-PFD a rather competitive phase detector. As future work, the case of a nonlinear transfer characteristic with gradually increasing gain, yet with saturation, could be investigated.

## References


[1] V. F. Kroupa, "Phase Lock Loops and Frequency Synthesis," John Wiley & Sons, 2003.

[2] R. E. Best, "Phase-Locked Loops: Theory, Design and Applications, Fifth Edition," McGraw-Hill, 2003.

[3] D. R. Stephens, "Phase-Locked Loops for wireless communications: Digital, Analog and Optical Implementations, Second Edition," Kluwer Academic Publishers, 2002.

[4] K. Kundert, "Predicting the Phase Noise and Jitter of PLL-Based Frequency Synthesizers," www.designers-guide.com.

[5] B. Razavi, "RF Microelectronics," Prentice Hall, 1998.

[6] F. M. Gardner, "Charge-Pump Phase-Locked Loops," IEEE Trans. Comm., Vol. COM-28, pp.1849-1858, November 1980.

[7] D. Banerjee, "PLL Performance, Simulation, and Design, 3rd Edition," 2003.

[8] P. V. Brennan, "Performance of phase-locked loop frequency synthesizer using accumulative phase detector," IEE Proc. – Circuits Devices Syst., Vol. 143, No. 5, pp. 249-254, October 1996.

[9] G. T. Roh, Y. H. Lee, and B. Kim, "Optimum Phase-Acquisition Technique for Charge-Pump PLL," IEEE transactions on Circuits and Systems-II: Analog and Digital Signal Processing, Vol. 44, No. 9, pp. 729-740, September 1997.

[10] C. Y. Yang and S. I. Liu, "Fast-Switching Frequency Synthesizer with a Discriminator-Aided Phase Detector," IEEE Journal of Solid-State Circuits, Vol. 35, NO. 10, pp. 1445-1452, October 2000.

[11] I. C. Hwang, S. H. Song, and S. W. Kim, "A Digitally Controlled Phase-Locked Loop With a Digital Phase-Frequency Detector for Fast Acquisition," IEEE Journal of Solid-State Circuits, Vol. 36, No. 10, pp. 1574-1581, October 2001.

[12] J. Lee, and B. Kim, "A Low-Noise Fast-Lock Phase-Locked Loop with Adaptive Bandwidth Control," IEEE Journal of Solid-State Circuits, Vol. 35, No. 8, pp. 1137-1145, August 2000.

# Design, Implementation and Evaluation of a Remote Laboratory System for Electrical Engineering Courses

Dimitris Karadimas and Kostas Efstathiou
Dept. of Electrical Engineering and Computer Technology, University of Patras, Greece
*{karadimas, efstathiou}@apel.ee.upatras.gr*

## Abstract

*This paper describes an Internet-based laboratory, named Remote Monitored and Controlled Laboratory (RMCLab) developed at University of Patras for electrical engineering laboratory education. The key feature of this remote laboratory is the utilization of real experiments, in terms of instrumentation and lab circuits, rather than simulation or virtual reality environment. RMCLab is able to provide affordably and simultaneously its services to many users through the Internet. A wide variety of lab experiments are supported, including not only fixed circuits but also students' custom circuit designs; thus contributing to students' authentic understanding of theory subjects. RMCLab can be accessed via the web through http://www.apel.ee.upatras.gr/rmclab.*

## I. Introduction

In almost any engineering department, laboratory experience is an integral part of the educational process. In fact, electrical engineering laboratory education is mandatory and demands an amount of material resources (instrumentation and hardware), which are not always affordable. As a result, the possibility of personal experimental training, for each student in large classes, would be extremely suppressed, leading anywise to a somehow reduced level of education.

The rapid progress of Internet and computer technology, along with its increasing popularity, enables the development of remote laboratories for supporting distance laboratory courses, where the experiments can be accessed, monitored and controlled remotely [1]. This new interpretation of the experimental training procedure offers to the students the opportunity to interact with the laboratory at any time, while at the same time reduces the experiment cost per student and also extends the capabilities of the entire experimental framework [2]. Moreover, remote laboratories can offer high level experimental training if they are able to realize, support and interact with real lab experiments, rather than simulations or simple presentation of the reality.

Many e-learning software systems can be found that enable distance laboratory education via online courses and simulated virtual lab environments [3]. These software systems integrate many of the desired functions, such as presentation of the course theory, communication support, through email, chat rooms and forums, collaboration through discussion pages and self-evaluation questions. Despite the fact that nowadays simulators can accurately estimate the circuits' performance, the employment and utilization of real lab experiments ensure the measurements' reliability while at the same time increase the educational value.

This paper presents the specifications and the structure of an integrated educational platform that enables the instant remote access to real lab experiments, with the use of real hardware and instrumentation. This platform, named Remote Monitored and Controlled Laboratory (RMCLab), is able to provide educational services to a great number of students for a wide range of real electrical engineering experiments, either pre-configured or customizable, at a very low total implementation cost. RMCLab is already in use since March 2004 at the Dept. of Electrical Engineering and Computer Technology of University of Patras, Greece, where it was developed and implemented.

## II. The RMCLab System Design Issues

The basic purpose of the developed platform is to provide high quality lab education in electrical engineering courses to a great number of students. The design of such a remote laboratory for real-time Internet-based experiments [4]-[5] should consider all aspects of the system, including communication, instrumentation and hardware control.

The RMCLab system has been designed so as to integrate all potentials of the physical laboratory to a user friendly interface, among other subsystems, such as lab assessment and scheduling, user administration, instrument operation and hardware management. Since a remote laboratory has two user types, the student and the instructor, which have completely different requirements from the system, two different sets of specifications have been assigned, in order to make RMCLab not only usable and acceptable but also necessary, for both types of user.

The primary service that RMCLab should provide to students is the possibility to study on the lab subjects, by accomplishing their lab-courses at any time and anywhere. For this reason RMCLab's basic specification is defined as the ability to serve simultaneously and at real time, 24-hours per day, 7-days per week, any potential user. Significant role to the acceptance of a remote educational tool plays also the usability, the functionality and the time response of its software components; however these parameters are easily supplied by nowadays software development tools.

An integrated remote laboratory platform should moreover reinforce the instructor's faculties, regarding the lab experiments' setup and the evaluation of students' knowledge on the lab subjects, by providing an effortless, and with the minimum possible human intervention, management of any educational procedure. Therefore, any kind of assessment functionality should be supported; assignment of the required measurements and instruments' settings, as well as multi-type questions. Finally, the ability of the instructor to monitor and interactively advise the students during a remote lab experiment is of great educational value; whereas administrative and automated secretarial procedures integrate the RMCLab system by offering usability and functionality regarding both types of user.

## III. Architecture of the RMCLab System

The RMCLab system has been developed based on the conventional client-server architecture as depicted in Fig.1, and consists of the following basic entities: client, instructor client, application server, resource server and lab infrastructure, including the instrumentation and the hardware modules.

The server side of the architecture employs two sub-servers; the Resource Server (RS) and the Application Server (AS). Resource server manages and operates hardware and instrumentation resources, providing to application server an abstract communication language that enables access to lab infrastructure. Application server undertakes all the communication tasks between clients and the physical remote laboratory (resource server and lab instrumentation), as clients are not able to directly access the resource server.

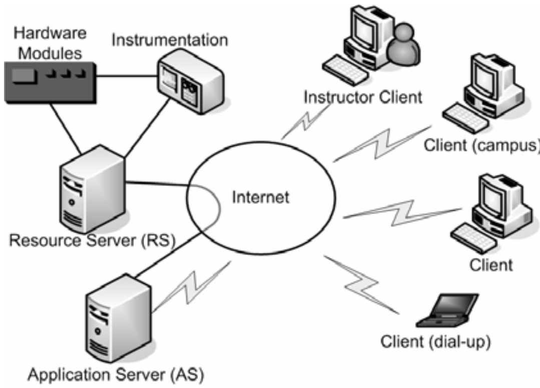



**Fig. 1. RMCLab system architecture.**

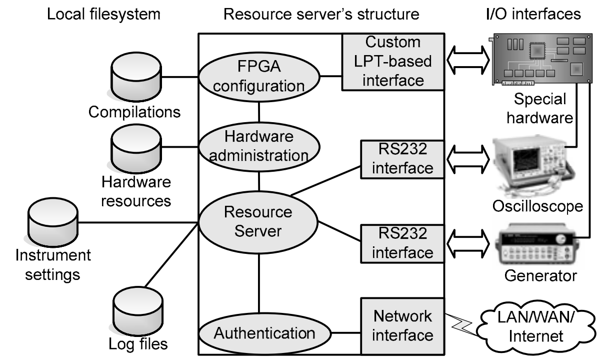



**Fig. 2. Software and Hardware modules of the resource server.**

This topology simplifies the architecture of the server-side and expands platform's capabilities as it facilitates the development and customization of a resource server, while at the same time enables the utilization of its shared resources by many application servers. Additionally, both instructor and student are able to transparently access physical resources of the remote laboratory, increasing system's flexibility and expandability.

The real measurement laboratory is based on a low cost and easy implementation, while it is realized around the resource server. Resource server is equipped with suitable, custom interface toward the signals of the lab experiment (both digital and analog experiments), via a custom bus based in LPT, and the instruments, via the RS232 interface, as depicted in Fig.2. Standard or other interfaces, like PCI, USB, etc, may be also supported by the resource server.

Multiple types of circuits can be hosted in the platform's resource server; standard, pre-configured or re-configurable analog, digital or mixed circuits. For this reason, resource server is outfitted with a motherboard that is able to host up to 64-cards, where each of them is incorporated with an FPGA and extra auxiliary modules, as depicted in Fig.3.

Each card employs also a PLD, which is responsible for the card addressing and the configuration of the FPGA. Each of these cards can host 8-different analog, digital or mixed independent arbitrary circuits, since the FPGA is segmented into 8-sectors, each of them corresponding to a specific lab experiment. The internal operation of the FPGA is controlled by a register file (TABLE I) which is employed within it.

As each sector of the FPGA can host either a specific multi-mode lab experiment or a user's custom circuit, the mode register and two auxiliary registers control its operation mode and behavior. For example, a single sector could implement both synchronous and asynchronous counters, being realized as two different lab experiments. Sector register points to the active sector, on which measurements are performed. Finally, when a measurement is carried out, two more registers, Probe1

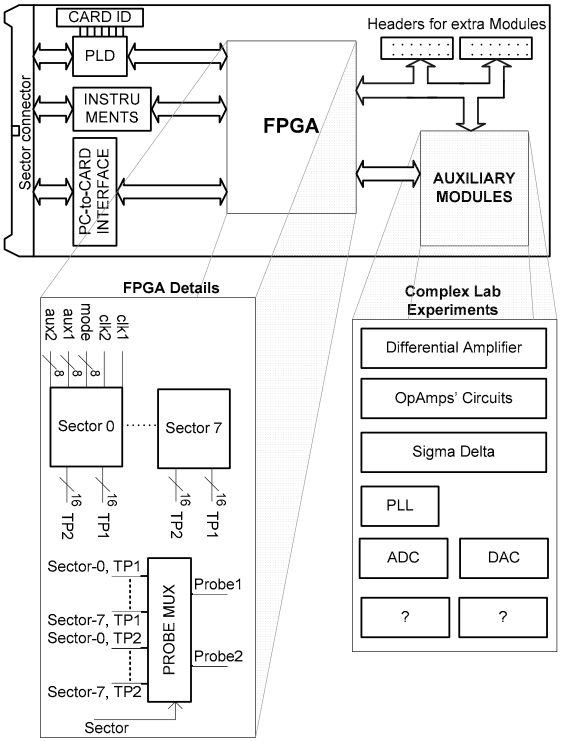



**Fig. 3. Hardware architecture.**

and Probe2, assign the active nodes of the active sector, on which the two probes of the oscilloscope become physically connected through cross-point switches. Moreover, each card may be equipped with additional on-board or external circuitry, in order to implement a wide range of more complex electronic circuits, which is an offline procedure.

When a client raises a measurement request, application server logs and routes this request to the proper resource server. Afterwards, resource server has to accomplish multiple tasks, as the authentication of the request and the lab infrastructure (hardware and instrumentation) setup, so as to be ready for the requested operations. A hardware administration module is responsible for the aforementioned task, giving first priority to the servicing of the request. This may lead to a real-time online re-configuration of one's card FPGA, so as to implement the requested circuit, or to the removal of an unutilized sector's circuit in order to provide a free sector to the system. As soon as the hardware is configured, the measurement is performed and the acquired data are transmitted back to the client again via the application server. The above procedures serve each

**TABLE I. FPGA register file.**

| Name | Address | Width (bits) | Operation |
|---|---|---|---|
| Sector | 0 | 3 | Select the active sector |
| Probe1 | 1 | 4-6 | Select the active nodes of Oscilloscope's Ch-A |
| Probe2 | 2 | 4-6 | Select the active nodes of Oscilloscope's Ch-B |
| Aux1 | 3 | 8 | Auxiliary register 1 |
| Aux2 | 4 | 8 | Auxiliary register 2 |
| Mode | 5 | 8 | Sector's operation mode |

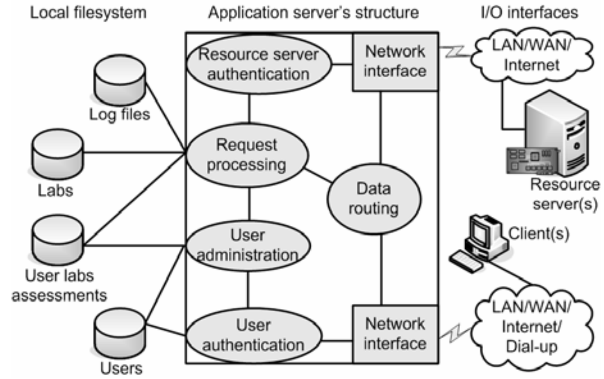



**Fig. 4. Software modules of the application Server.**

possible request in a FIFO priority so as to time-share the physical resources to all available requests.

Apart from the dataflow control and the routing procedure between client(s) and the resource server(s), application server is also responsible for the logging, assessment and the evaluation of the students' actions. These log files can be offline presented or actively reproduced by the instructor in order to illustrate the students' skills and level of knowledge for each lab experiment. Although some assessment procedures can be automatically performed by the software of the application server, like multiple choice questions and measurements, there are also different types of assessments, self-assigned questions or evaluation on the range and the cohesion of the student's actions, which demand the offline human intervention for accurate and fair evaluation of the students. The above presented characteristics and functionalities of the RMCLab's application server define its architecture, as depicted in Fig.4.

The client part of the RMCLab's system is designed so as to meet the requirements provided by the server side. Thus, client embeds interfaces, unified as scenario interface, for supporting the remote monitor and control of lab infrastructure, like full-functional and user friendly interfaces for lab instrumentation (function generator, oscilloscope) and lab hardware, as is depicted in Fig.5.

In more details scenario interface provides to the student graphic information, related with the real hardware of the corresponding lab experiment.

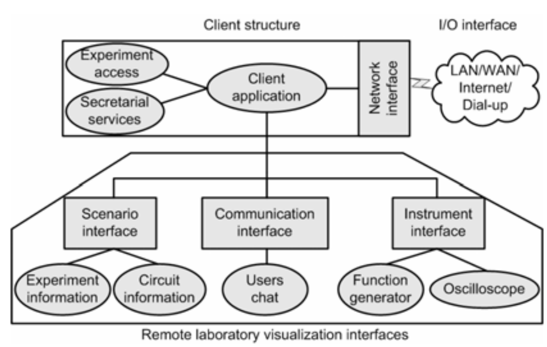



**Fig. 5. Software modules of the Client.**


0-7695-2632-2/06 $20.00 © 2006 IEEE

Additionally, it enables the student to control circuit parameters of the lab experiment (variable pots, caps, etc) and monitor any node of the circuit by setting the probes of the oscilloscope and the function generator on it. Moreover, scenario interface provides the necessary information, regarding the technical and theoretical aspects of the experiment, among with the prerequisites for the students' evaluation, such as multiple choices, measurement or self-assigned questions; all defined by the instructor.

Finally, RMCLab's platform embeds an identical to the students' one client interface, called as instructor client, which incorporates the ability to replicate, monitor and control any online student's environment to the instructor's side. This feature gives the instructor the ability to closely observe the actions and efficiently tutor any online student, concluding to a 'near-to-real' lab environment.

## IV. Advanced RMCLab System's Properties

The architecture described in the previous section allows several functional scenarios for an advanced utilization of RMCLab's system.

The most important benefit of the RMCLab's system architecture is that provides to the students the feasibility to design their own custom circuits and test/measure them under real hardware and real instrumentation. Students can offline design almost any circuit using a separate software package, which can be MAX+plus II or Quartus, both offered at no-cost to the students from Altera. Using one of the specific software packages, students can design their circuits following a reduced set of rules and confirm its proper operation. Once the design is verified at the client-side, it can be easily uploaded to the server-side and after a few seconds students will be able to perform any measurement on their custom design, which is now implemented on real hardware by using real instrumentation.

RMCLab system's architecture is able to support even more functionalities. Assume that a resource server, physically located somewhere in the world, supports and shares through the Internet the hardware and instrumentation required for 'Lab A', as depicted in Fig.6. This shared lab infrastructure can be utilized by one

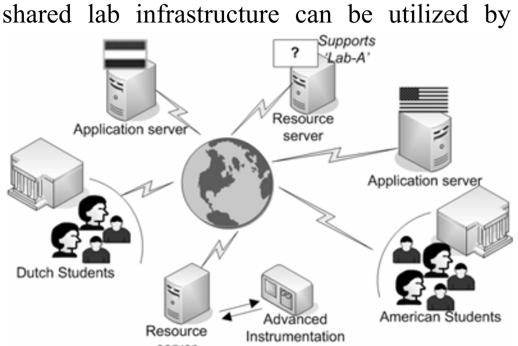



**Fig. 6. Advanced RMCLab system utilization.**

application server in Nederland, where the instructor has developed the required educational material regarding 'Lab A' for his students. Thus, Dutch students are enabled to attend the lab course regarding 'Lab A' in their native language, while the Dutch instructor will be able to take the advantages of this service. At the same time a second application server, located in USA, may also utilize the same lab infrastructure of the resource server supporting 'Lab A', while American instructor can prepare the corresponding required educational material in the native language of his students, enabling also them to attend the specific lab course.

Obviously, each instructor has the opportunity to mark and review his students' performance according to his own educational and pedagogical criteria, as the set of the assessments rules for each lab is defined in the RMCLab's application server, which is available and accessed by the instructor, while resource server transparently executes the measurement requests.

The prospects of the RMCLab system may hopefully expand world-wide, as the above scenario can be further extended if one adds more resource servers. Each resource server can be expert and focused on a specific subject, incorporating the appropriate hardware and instrumentation. Instructors all over the world may take the advantage of using such laboratory resources and develop educational material in their local application servers, so as to offer advanced experimental training to their students, without any requirement for the development and maintenance of any expensive lab infrastructure.

RMCLab's advanced utilization modes are not limited within the above examples. The real-time use of real hardware and real instrumentation can significantly contribute to the educational procedure, since it enables the instructor to prepare 'Active Lessons' and present in details the operation of a circuit or a system under real-world circumstances.

## V. Realized RMCLab System

The architecture described in Section III has been implemented at the University of Patras. Current configuration is a cost effective implementation, which employs a single PC with a Celeron 2.4GHz processor, 396MB RAM for both the resource and the application server of the RMCLab's system. This PC, running Windows 2000 Server, is permanently connected to the campus LAN and through this same LAN to Internet. An Agilent 54622D mixed signal oscilloscope, controlled via RS232@56kbps and an Agilent 33120A function generator, controlled via RS232@19.2kbps are physically connected to this PC.

Hardware interface is implemented based on a custom, low-cost bus, through LPT in EPP mode. Each card of the hardware infrastructure contains an Altera FPGA of the

**TABLE II. RMCLab's time response characteristics.**

| Property | Average Delay (sec) |
|---|---|
| Hardware setup and measurement time | 3 |
| Compilation time in the server side | 10 |
| Hardware re-configuration time | 5 |
| Measurement delay from client side using PSTN line @56kbps | 5 |

FLEX8K series and other components required for the implementation of the experiments. Currently, all cards are hosted by a small motherboard outfitted with 5-cards, out of 64-cards that hardware structure is able to support, for lab experiments and a dedicated card acting as the power supplier for the hardware infrastructure. These cards have been mounted to a custom-designed box, which also provides the required connectivity to the PC and the instrumentation.

Client software minimum requirements include a PC-based computer running at least Windows '98SE, screen resolution 1024x768 (or higher), DirectX runtime libraries (version 8.1 or higher) and a conventional Internet connection (PSTN or better).

RMCLab's system provides its educational services since March 2004 for the Dept. of Electrical Engineering and Computer Technology of University of Patras, Greece, regarding Analog and Digital Circuits lab experiments. Analog lab experiments include 2-stage, feedback and cascode/folded-cascode amplifiers, whereas digital lab experiments include a wide variety of counters, adders and accumulators. The technical characteristics demonstrated by the up-to-now use of the RMCLab system consist of the average timings/delays, presented in TABLE II.

In particularly, during the second semester of academic year 2004-'05, RMCLab provided to the 3rd year's class (about 90-groups of 3-students each) of the aforementioned department, educational services regarding three laboratory experiments. The first experiment regarded an introductory exercise, aiming at the students' familiarization with the use of RMCLab platform, while the other two regarded synchronous, asynchronous, binary, decimal and programmable counters. For these two obligatory lab-exercises 1264-accesses were logged, of 473h 5m 6s total duration. During this period, up-to-8 simultaneous requests have been raised to the RMCLab resource server, without importing any extra delay to the users' requests servicing. Fig.7 gives a notion of RMCLab usage for this period. Additionally, 17200-measurements were logged on the RMCLab's instrumentation, where 1666 of them regarded the introductory exercise, and 4458, 11076 measurements regarded the second and third obligatory exercise, respectively. For the first obligatory exercise the performed measurements were 4.35-times more than the total number required for completing the exercise, while for the second one this was reduced to 3.97, despite the fact that this exercise was significantly more demanding for the students.

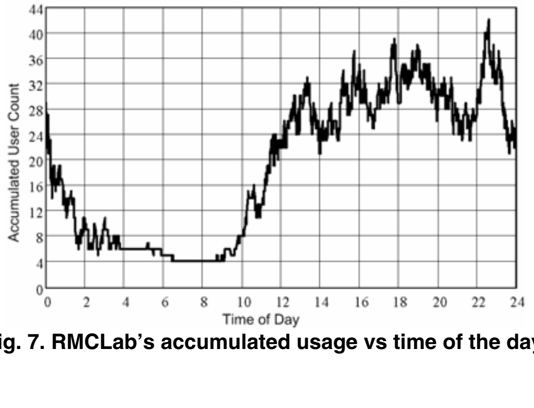



**Fig. 7. RMCLab's accumulated usage vs time of the day**

## Conclusion

RMCLab's system is able to provide a wide range of high educational services in a great number of students. It increases the productivity of the students by enabling them to have access to the lab infrastructure at non-working hours, while at the same time affects significantly their psychological mood regarding the level of the offered education by their institute.

The structure of RMCLab enables sharing of hardware and instrumentation resources, thus makes possible the extensive exploitation of an expensive lab infrastructure, facilitating the wide spread of remote real lab experiments, which are indisputably valuable for engineers' education. Additionally, hardware re-configurability permits the remote implementation and measurement of electronic circuits, providing further more a high-valued educational service.

The concentrated use of RMCLab system during 5-academical semesters, for the courses of Analog and Digital Integrated Circuits, consisting of classes of about 300-students per class, has definitely proved the high value of this educational tool, for the students and for the instructors as well.

## References


[1] Ferrero A., Salicone S., Bonora C., Parmigiani M., (2003), "ReMLab: A Java-Based Remote, Didactic Measurement Laboratory", *IEEE Trans. on Instr. and Meas.*, June 2003, pp. 710-715.

[2] A. Bagnasco and A. Scapolla, "A Grid of Remote Laboratory for Teaching Electronics", *2nd International LeGE-WG Workshop on e-Learning and Grid Technologies: a fundamental challenge for Europe*, Paris, March 2003.

[3] J. Palop and M. Teruel, "Virtual Work Bench for Electronic Instrumentation Teaching", *IEEE Trans. on Education*, Feb. 2000, pp. 15-18.

[4] T. Fjeldly and M. Shur, "Lab on the Web: Running Real Electronics Experiments Via the Internet", *Wiley-IEEE Press*, Oct. 2003

[5] C. Ko, B. Chen and J. Chen, "Creating Web-based Laboratories (Advanced Information and Knowledge Processing)", Springer, Oct. 2004.

# An Analog PLL-based Technique for VCO Phase Noise Reduction

*D. Mavridis, D.Karadimas, M. Papamichail, K.Efstathiou and G.Papadopoulos*
Department of Electrical and Computer Engineering, University of Patras
Rion Campus, 26500, Patras, Greece
Email: {mavridis, karadimas, mpapamicha, efstathiou, papadopoulos}@ee.upatras.gr

**Abstract—A PLL-based technique, able to reduce phase noise and non linearity of voltage controlled oscillators, is presented. According to this technique, the phase of a VCO can be sampled using an analog circuit (phase to voltage converter). The resulting signal is filtered and fed back to the VCO, decreasing its phase noise. Based on the phase/jitter properties extracted from transistor level analysis, a voltage domain behavioral model of the system was simulated and significant phase noise reduction was confirmed.**

## I. INTRODUCTION

Voltage-controlled oscillators (VCOs) constitute a critical component in modern communication systems. The specifications of wireless standards require tight performance for local oscillators, which cannot be met without several design tradeoffs. Phase noise is one of the most significant constraints, as there is a growing demand for higher number of channels and for better exploitation of the available bandwidth. Phase noise reduction comes at the cost of increased power consumption which limits the operating time of battery supplied wireless devices [1].

There are two main categories of oscillators, harmonic and relaxation. The first consist of an inductance-capacitance (LC) resonating circuit and a negative resistance active circuit used to compensate for the resonator losses [2]. Extensive studies of their noise mechanisms have led to optimized designs [3]-[6]. The used passive elements however, pose several constraints. The quality factor of on-chip spiral inductors is low and their dimensions are relatively large compared to transistors, resulting in increased power consumption and chip areas. Moreover, varactors, which give tunability to the circuit, are not able to provide wide tuning range.

Oscillators of the second category charge and discharge periodically a capacitor to produce the oscillating signal. Ring oscillators are the most representative topology of this category [7]. The delay of the inverters and the number of stages determine the oscillation frequency. Their disadvantage is the noise performance which is very poor for today's demanding applications in Gigahertz frequencies [6,8]. Nevertheless, their small chip area, the reduced power consumption and the wide tuning range [9,10] constitute attractive characteristics to designers.

This paper presents a technique able to reduce significantly the phase noise of any VCO type. According to this technique, a Frequency to Current Converter (FCC) is employed, embedding the frequency noise properties of the VCO output. The FCC's output current is integrated and thus, the output voltage of the integrator holds information regarding the VCO's phase noise. The integrator's output voltage is filtered and fed to the VCO's input. Since the control variable of this negative feedback loop is the phase, it can be considered as a Phase Locked Loop (PLL).

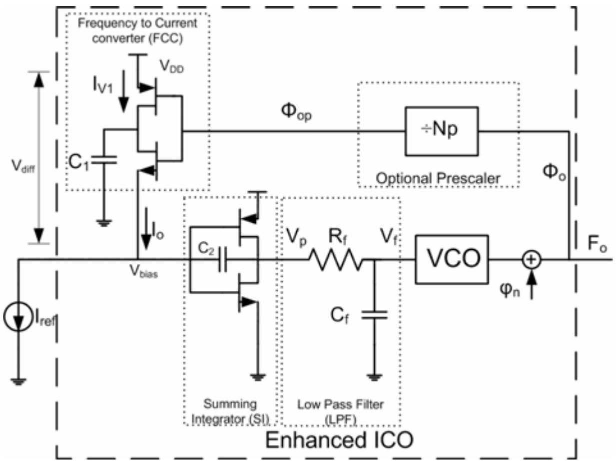


Figure 1. The proposed Current Controlled Oscillator.

The resulting closed loop system acts as a CCO (Current Controlled Oscillator) or VCO and has improved characteristics in terms of phase noise and linearity, employing a few analog elements. It is meaningful to apply this technique to a ring VCO since its main drawback, the phase noise, will drop significantly without affecting the rest of its properties.

The paper is organized as follows. Section II describes the operation of the system, its mathematical analysis and the noise characteristics. Section III discusses the simulation setup employed for the confirmation of the theoretical operation. In section IV, the simulation results are shown. Finally, design aspects and the conclusions are considered.

## II. DESCRIPTION OF THE SYSTEM

Fig. 1 depicts the proposed system, which employs the noisy VCO with gain $K_V$ (Hertz/Volt). The output frequency of the VCO, which can be optionally scaled down by a prescaler of ratio $N_p$, drives a simplified Frequency to Current Converter (FCC) that consists of an inverter with a small capacitance $C_1$ at its output. $C_1$ is charged at $V_{DD}$ and discharged at $V_{bias}$ on every cycle of the prescaler's output. Thus, the current $I_o$ acquires the frequency properties of the output of the prescaler and consequently of the VCO. The average value of $I_o$ is given by:

$$I_o(t) = \frac{F_o(t)}{N_p} V_{diff} \cdot C_1 \tag{1}$$

where $V_{diff}=V_{DD}-V_{bias}$ and $F_o$ is the VCO's output frequency.

The difference of $I_o$ and the input current $I_{ref}$ is integrated by the simplified Summing Integrator (SI) providing at its output a voltage that carries the properties of the integral of the VCO's output frequency, thus the properties of VCO's phase, as shown in the following expression:

$$V_P(t) = -\frac{1}{C_2}\int\left(\frac{F_o(t)}{N_p}V_{diff}\cdot C_1 - I_{ref}\right)dt \Rightarrow$$

$$V_P(t) = -\frac{C_1\cdot V_{diff}}{C_2\cdot N_p}\cdot\frac{\Phi_o(t)}{2\pi} + \int\frac{I_{ref}}{C_2}\cdot dt \quad (2)$$

It should be emphasized that the input voltage of the SI, $V_{bias}$ remains constant since SI is a current summing integrator with negative feedback.

SI's output voltage is filtered by a simple low pass filter with $R_f C_f = T_f$, so as to reject any high frequency components and drives the VCO's input. Thus, a negative feedback, proportional to the VCO's phase feeds its input. The system is at steady state, when FCC's mean output current, given in (1), equals to the input current $I_{ref}$. This occurs when the output frequency is:

$$F_o = I_{ref}\frac{N_p}{C_1\cdot V_{diff}} = I_{ref}\cdot K_I \quad (3)$$

It becomes apparent that this system is a Current Controlled Oscillator (CCO) with a constant gain $K_I$ (Hz/A) which does not depend on the gain $K_V$ of the employed VCO. The system can also be analyzed in the frequency domain by rewriting (2), as:

$$V_p(s) = \frac{I_{ref}(s)}{sC_2} - \frac{C_1\cdot V_{diff}}{C_2\cdot N_p}\frac{\Phi_o(s)}{2\pi} \quad (4)$$

At the output of the VCO we have:

$$\Phi_o(s) = \frac{2\pi K_V}{s}\frac{1}{1+sT_f}V_p(s) \quad (5)$$

Substituting $V_p(s)$ in (5), we obtain the phase to current transfer function:

$$P(s) = \frac{\Phi_o(s)}{I_{ref}(s)} = \frac{2\pi}{s}\cdot\frac{N_p}{V_{diff}\cdot C_1}\cdot\frac{\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2} \quad (6)$$

where:

$$\omega_n = \sqrt{\frac{V_{diff}\cdot K_V\cdot C_1}{N_p\cdot T_f\cdot C_2}} \quad (7)$$

and:

$$\zeta = \frac{1}{2\cdot\omega_n\cdot T_f} = \frac{1}{2\cdot\omega_n\cdot \mathrm{R_F}\cdot \mathrm{C_F}} \quad (8)$$

Consequently, the transfer function for the frequency is:

$$G(s) = \frac{F_o(s)}{I_{ref}(s)} = \frac{N_p}{V_{diff}\cdot C_1}\cdot\frac{\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2} \quad (9)$$

Equation (9) confirms that the derived CCO has a gain $K_I$ calculated in (3) and acts as a second order low pass filter with natural frequency $\omega_n$ and damping factor $\zeta$.

The derived constant gain should be considered as a significant improvement since it contributes to the linear behavior of the system that employs this type of CCO. Additionally, its intrinsic low pass behavior is rather desirable in most of the real world applications. The natural frequency $\omega_n$ and the damping factor $\zeta$ can be adjusted so as to comply with the specifications and the requirements of the application where this CCO will be employed.

The key characteristic of the circuit is its ability to reduce significantly the in-band phase noise of the output frequency. The transfer function of the system with respect to the modeled phase noise $\Phi_n$ of the employed VCO is:

$$R(s) = \frac{\Phi_o(s)}{\Phi_n(s)} = \frac{s^2+2\zeta\omega_n s}{s^2+2\zeta\omega_n s+\omega_n^2} \quad (10)$$

implying that the VCO's phase noise is filtered by a second order high pass filter at $\omega_n$ and with damping factor $\zeta$. Therefore, a phase noise reduction up to 40 dB per decade can be achieved in-band, for frequencies lower that the $\omega_n$.

## III. Simulation Setup

The circuit described above has been modeled and simulated in order to validate the theoretical results. The model was built using Verilog-A [14], an industry standard modeling language and the simulations were performed in Cadence environment. The process of predicting the jitter of the CCO system, with voltage domain models involves the following steps [11]:

- Use of SpectreRF [12] to predict the phase noise of the individual blocks that make up the CCO. These blocks are simulated in transistor level.
- Converting the phase noise of the individual blocks to jitter.
- Building high-level behavioral models for each block that exhibit jitter. This is established by dithering the time at which events occur. In this case the events are the zero crossings of the output signal.
- Assembling the blocks of the CCO into a general model.
- Simulating the CCO and sampling the zero crossings of the output signal.

After specifying the phase noise of each block in the second step of the previous process, the modeling requires to convert it to jitter. This last term describes an undesired perturbation or uncertainty in the timing of events. There are several metrics of jitter. In the case of autonomous systems such as the VCO, the period jitter is used which is defined as the standard deviation of the length of a single period. Following the analysis in [11], the rms value of period jitter is given by:

$$J = \sqrt{cT} \tag{11}$$

where c is a quantity given in seconds and T is the oscillation period.

Oscillator phase noise is a variation in the phase of the oscillator as it proceeds along its limit cycle. It is the description of the jitter in the frequency domain. The most common metric of phase noise is the normalized single sideband power spectral density of the output voltage at a frequency offset Δf from the carrier, $\mathcal{L}(\Delta f)$, given in decibels below the carrier per Hertz (dbc/Hz) [6]. In [11], it is proven that:

$$\mathcal{L}(\Delta f) = \frac{c \cdot f_0^2}{\Delta f^2} \tag{12}$$

where $f_0$ is the oscillation frequency and $\Delta f$ the frequency offset from the carrier.

Defining the phase noise specifications of the VCO, it is easy to extract the jitter information, using (11) and (12). Supposing an oscillator, operating at a frequency of $f_0$ =1GHz with $\mathcal{L}(\Delta f)$ =-140dbc/Hz at 20MHz offset, the rms period jitter equals 63 femtoseconds (fs).

Simulation results were derived by transient analysis for hundreds of microseconds and required several hours to complete since the simulator needs to capture the details caused by the jitter.

However this time is considered short, compared to the time required for the transistor level simulation, since Verilog-A [14] modeling is simple, accurate and time effective. The collected samples of the output signal zero crossings are post-processed in Matlab [15] in order to extract the closed-loop phase noise.

## IV. Simulation Experiments

Simulations were performed for three setups of the CCO, varying the loop natural frequency and the damping factor. The results are taken for $f_n$=20MHz with ζ=0.5 and 0.7 and for $f_n$=40MHz with ζ=0.5. The desired performance of the circuit and the parameters of the employed components are listed in Tables I and II respectively. The capacitor at the output of the FCC is fixed to 200fF and the filter resistance to 500 Ohms.

One should notice that the selected natural frequency for the loop is high in contrast to classic PLL topologies where $\omega_n$ is in the range of KHz and is usually dictated by the reference frequency [13]. This is attributed to the fact that the applied frequency to the SI circuit, that samples the VCO's phase noise, can be equal even to its output frequency.

TABLE I. CCO Characteristics

| Design parameters for the CCO | |
|---|---|
| ***Operating Frequency*** | 1 GHz |
| ***VCO gain*** | 1.2 GHz/V |
| ***Supply Voltage*** | 3.3V |
| ***Ring VCO phase noise*** | -142 dbc/Hz@20MHz |
| ***Natural frequency*** | 20 MHz, 40MHz |
| ***Prescaler ratio*** | 4@$f_n$=20 MHz,2@$f_n$=40 MHz |

TABLE II. Element Values

| Passive elements | | | |
|---|---|---|---|
| ***System Variables ($f_n$, ζ)*** | 20 MHz,0.5 | 20 MHz,0.7 | 40 MHz, 0.5 |
| ***Capacitor $C_1$ (F to I)*** | 200fF | | |
| ***Capacitor $C_2$ (SI)*** | 1.58pF | 2.2pF | 1.58 pF |
| ***Capacitor $C_f$(Filter)*** | 15.9pF | 11.4pF | 7.95 pF |
| ***Resistance $R_f$*** | 500 Ohm | | |

Figure 2 depicts simulation results for the VCO phase noise and the derived phase noise of the enhanced CCO for the values of ζ and $\omega_n$ mentioned above. The open-loop VCO presents a phase noise of -142dbc/Hz at 20MHz frequency offset. The simulation confirms the conclusions of the theory. The phase noise of the open loop VCO is high while the phase noise of the closed loop is kept flat inside the band. It is apparent that while the employed open loop VCO cannot be used for high performance applications, the derived CCO has significant better phase noise performance and can be compared directly to state of the art VCOs. [1,2,4].

Taking for example the GSM application, the specifications for the phase noise are listed in Table III. The numbers in parentheses indicate the difference between the devices' phase noise and the GSM standard. A VCO that has a phase noise of -142dbc/Hz at 20MHz offset, performs far away from the specifications. However, applying the proposed technique for a loop natural frequency of 20MHz for example, the phase noise level at 20MHz is maintained throughout the band below 20MHz, satisfying the demands of the specific standard. For $f_n$=40 MHz the closed-loop apparently performs better since the phase noise level at the specific frequency offset is lower.

TABLE III. Phase Noise Values

| Phase noise (dBc/Hz) | | | | |
|---|---|---|---|---|
| ***Offset frequency (Hz)*** | ***GSM*** | ***VCO*** | ***CCO($f_n$=20 MHz)*** | ***CCO($f_n$=40MHz)*** |
| 100K | -112 | -96 (+16) | -142 (-30) | -150 (-38) |
| 600K | -126 | -112 (+14) | -142 (-16) | -150 (-24) |
| 3M | -142 | -125 (+15) | -142 (0) | -150 (-8) |

## V. Design Considerations

The designer of the proposed CCO should take into account some important considerations concerning design and technical parameters of the system. First of all, the loop natural frequency $\omega_n$ should not exceed one tenth of the phase sampling frequency. This poses an upper limit in the selection of the loop frequency. The lower limit for $\omega_n$ is imposed by the VCO's phase noise at this frequency, which actually will be the highest phase noise level for the system. For example, for the specific VCO and the specific standard (GSM) the minimum value of $f_n$ is 20 MHz.

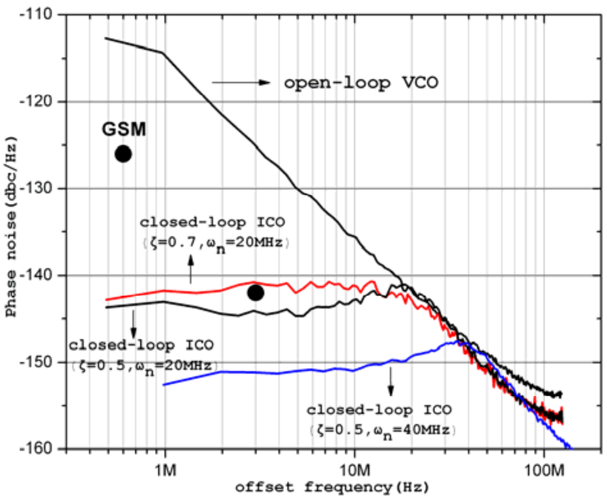



Figure 2. Phase noise of the closed-loop CCO at the output of the VCO when only the VCO exhibits jitter (closed-loop) for different values of damping factor and loop natural frequency versus the phase noise of the open loop VCO.

Moreover, it should be noted that the CCO gain $K_I$ given in (3) is a very large quantity. As a result, attention should be paid to the design of the driving current source as it directly impacts the output phase noise below frequency $\omega_n$. Alternatively, the use of a constant low noise current source will assign a non-zero free run frequency for the CCO, reducing at the same time the sensitivity of the system.

Additionally, Eqn. (3) implies that any fluctuations in $V_{diff}$ will affect $I_o$ and thus phase information. Therefore, the phase noise of the circuitry depicted in Fig. 1 suffers from $V_{dd}$ fluctuation. However, a comprehensive circuit design, not presented here for sake of simplicity, could provide high immunity to this type of fluctuations.

## VI. CONCLUSION

A PLL-based technique was presented, able to reduce by many decibels the phase noise of any VCO type. The implementation of the proposed technique requires a small number of additional elements and negligible power consumption overhead and, thus, it can be easily integrated. Moreover, it provides enhanced voltage to frequency linearity while retaining useful VCO characteristics, such as tuning range. However, the proposed system requires careful design regarding the noise of the low frequency components such as the input source and the integrator.

The proposed technique can be applied to today's high frequency demanding applications and especially in frequency synthesizers relaxing their design tradeoffs.

## REFERENCES


[1] Emad Hegazi, Jacob Rael, Asad Abidi, "The Designer's Guide to High-Purity Oscillators", Springer, 2004

[2] A. Hajimiri and T. H. Lee, "Design issues in CMOS differential LC oscillators," IEEE Journal of Solid-State Circuits, vol. 34, pp. 717 - 724, May 1999.

[3] T. H. Lee and A. Hajimiri, "Oscillator phase noise: A tutorial," IEEE Journal of Solid-State Circuits, vol. 35, pp. 326 - 336, March 2000.

[4] D. Ham and A. Hajimiri, "Concepts and methods in optimization of integrated LC VCOs," IEEE Journal of Solid-State Circuits, vol. 36, pp. 896 - 909, June 2001.

[5] E. Hegazi, H. Sjöland, and A. A. Abidi, "A filtering technique to lower LC oscillator phase noise," IEEE Journal of Solid-State Circuits, vol. 36, pp. 1921 - 1930, December 2001.

[6] A. Hajimiri and T. H. Lee, "The Design of Low Noise Oscillators," Kluwer Academic Publishers, 1999.

[7] B. Razavi, "Design of Analog CMOS Integrated Circuits", McGraw-Hill, 2001

[8] A. Hajimiri, S. Limotyrakis, and T. H. Lee, "Jitter and phase noise in ring oscillators," IEEE Journal of Solid-State Circuits, vol. 34, pp. 790 - 804, June 1999.

[9] T.Miyazaki, M.Hashimoto, H.Onodera, "A Performance Comparison of PLLs for Clock generation Using Ring Oscillator VCO and LC Oscillator in a Digital CMOS Process", Proceedings of the 2004 Asia and South Pacific Design Automation Conference (ASP-DAC'04)

[10] C. Park and B. Kim, "A low-noise, 900-MHz VCO in 0.6-µm CMOS," IEEE Journal of Solid-State Circuits, vol. 34, pp. 586 - 591, May 1999.

[11] K. Kundert, "Predicting the Phase Noise and Jitter of PLL-Based Frequency synthesizers," Available from www.designers-guide.com

[12] Cadence Design Systems, SpectreRF simulation option, www.cadence.org/datasheets/spectrerf.html.

[13] K. Efstathiou, G. Papadopoulos. "An Enhanced Frequency Synthesizer using an Analog Dual Input Accumulator", Proceedings of the Sixth IEEE International Conference on Electronics, Circuits and Systems, Vol. 1, pp.17-20, Paphos, Cyprus, Sept. 1999

[14] Cadence®Verilog®-A Language Reference, January 2004

[15] Matlab, www.mathworks.com

# AN INTEGRATED PLATFORM, IMPLEMENTING REAL, REMOTE LAB-EXPERIMENTS FOR ELECTRICAL ENGINEERING COURSES

Dimitris Karadimas [1], Kostas Efstathiou [2]
[1] Ph.D. Student, [2] Lecturer
[1,2] Dept. of Electrical and Computer Engineering, University Campus, Rion, Patras
Greece
{karadimas, efstathiou}@apel.ee.upatras.gr

**ABSTRACT**
Nowadays, the design of a remote, real instrumentation and measurement laboratory requires flexible hardware platforms. This paper describes an Internet-based laboratory, named Remote Monitored and Controlled Laboratory (RMCLab) developed at University of Patras, Greece, for electrical engineering experiments. The key feature of this remote laboratory is the utilization of real experiments, in terms of instrumentation and under-measurement circuits, rather than simulation or virtual reality environment. RMCLab's hardware infrastructure contains multiple reconfigurable sub-systems (FPGAs), which can be enhanced by almost any analog expansion module. The main characteristics of this system include the versatility of the hardware resources, due to the dynamic reconfiguration potentiality and the low cost of the hardware components. Moreover, this system enables its users to test, in real time, their own custom circuit designs. The paper concludes with a specific example regarding an elementary circuit in digital electronics and a short statistical review of the RMCLab educational usage. RMCLab can be accessed via the web through http://www.apel.ee.upatras.gr/rmclab.

**KEY WORDS**
Client-server architecture, remote laboratory, distributed instrumentation and resources

## 1. Introduction

The rapid progress of internet and computer technology along with its increasing popularity enables the development of remote laboratories, where the experiments can be remotely accessed, monitored and controlled [1]-[2]. This new interpretation of the measurement process offers to anyone the opportunity to interact with the laboratory at any time, reducing at the same time the experiment cost per user and extending the capabilities of the entire experimental framework.

Moreover, remote laboratories can offer high-level experimental training and experience, when they are able to realize, support and interact with real experiments, rather than present simulation results or simple depiction of reality. Additionally, expensive, often dedicated experiments, of modern, cut-edge technology can be shared worldwide, contributing thus to a high-valued remote laboratory framework [3]-[4].

There can be found many internet-enabled software systems that afford distance laboratories via simulated, virtual environments [6]-[7]. These software systems often integrate many of the desired functionalities, especially from the user's side, such as accompaniments to the experiment documentation, communication support and collaboration among their users. Despite the fact that modern simulators can accurately estimate circuits' performance, the employment and utilization of real circuits and real instrumentation, for electrical engineering laboratories, ensures the measurements' reliability, while at the same time increases the educative value of the remote laboratory.

This paper presents the specifications and the basic structure of an integrated remote laboratory platform that enables the instant remote access to real experiments, employing real hardware and real instrumentation. This platform, named Remote Monitored and Controlled Laboratory (RMCLab), is able to provide high-level services to a great number of users for a wide-range of real electrical engineering experiments; either pre-configured, reconfigurable or customizable, at a very low hardware infrastructure cost. RMCLab is already in use since March 2004, at the Dept. of Electrical and Computer Engineering of University of Patras, Greece, where it was developed and implemented.

## 2. Proposed Approach

The basic purpose of the developed platform is to provide high-quality lab-training in electrical engineering subjects to a wide range of users. The design of such a remote laboratory for real-time, internet-based experiments, should consider all aspects of the system, including communication and data flow, as well as instrumentation and hardware control [8]-[9]. The RMCLab system has been designed so as to integrate all potentials of a physical laboratory to a simple user interface, among with other sub-systems, such as lab-administration, instrument operation and hardware management.

The primary service that RMCLab platform should provide to its users is the possibility to study on the experiment subjects, by accomplishing their measurements at any time and anywhere. For this reason, RMCLab's basic specification is defined as the ability to serve at any time, simultaneously and at real time, any potential user. On the other hand, an integrated remote laboratory platform should reinforce lab-administrator's faculties, regarding the experiment setup, hardware and instrumentation control, users' management and also lab-maintenance. RMCLab offers also many kinds of assessment functionalities for the lab-experiments, such as the assignment of several different evaluation criteria (measurements, instrument settings and multi-type questions), so as the whole platform can be configured as an advanced tool for high-level educational services, an aspect that characterizes our usage case and also the initial motivation.

## 3. Architecture

The RMCLab system has been developed based on the conventional client-server architecture, expanded in the server side, as depicted in Fig. 1, and consists of the following basic entities: client, instructor client (IC), application server (AS), resource server (RS) and lab-infrastructure, including the real instrumentation and all the hardware modules.

### 3.1 Network Topology

The server-side of the proposed architecture employs at least two sub-servers; the resource server and the application server, which can also be replicated in a more complex network topology. Resource server manages and operates hardware and instrumentation resources, providing to application server an abstract layer for communication that enables access to lab-infrastructure.

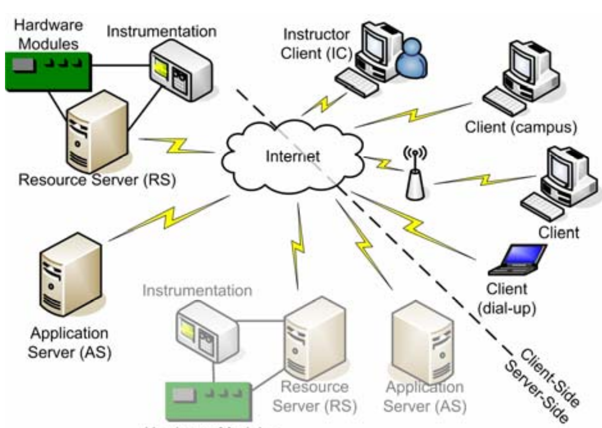



**Fig. 1. RMCLab system overview.**

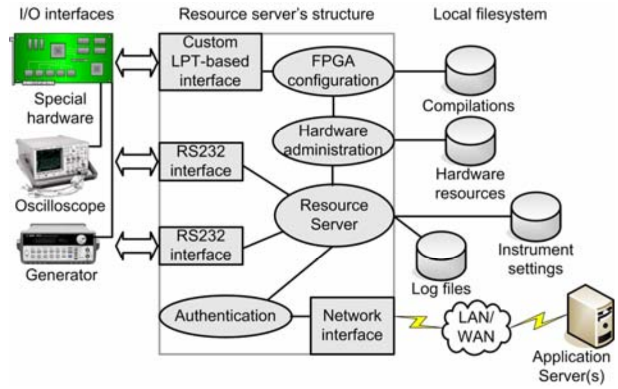



**Fig. 2. Hardware, instrumentation and resource server's software modules.**

Application server undertakes the data-flow control task between clients and the physical remote laboratory, realized by the resource server and the lab-instrumentation, since clients are not able to directly communicate with any of the resource servers. The communication between application server(s) and the resource server(s) or the client(s) is based on a custom, abstract language that integrates all potential tasks of a conventional, physical laboratory, while however it can be modified according to future or different requirements.

This topology simplifies the architecture of the server-side and expands platform's capabilities, as it facilitates the robust development and customization of a resource server. Moreover, it enables many application servers to utilize the components shared by the resource server(s), and as a result, a wide range of users are able to transparently access, via the application server(s), these shared resources. Additionally, application server grants to its clients transparently access to the physical resources of the physical laboratory, thus increasing system's robustness, flexibility and expandability.

### 3.2 Hardware, Instrumentation and Resource Server

The real measurement laboratory is based on a low-cost and easy implementation, while it is realized around the resource server. Resource server is equipped with suitable, interfaces toward the signals of lab-experiments (both digital and analog experiments), via a custom, LPT-based bus, and the instruments, via the RS232 interface, as depicted in Fig. 2. Standard or other commercial interfaces, like PCI, USB, etc, may be also supported by the resource server.

Multiple types (standard, programmable, pre-configured or re-configurable) of analog, digital or mixed circuits can be hosted in the platform's resource server(s). For this reason RMCLab's hardware if outfitted with a motherboard that is able to host up to 64-cards, where each of them is incorporated with an FPGA and extra auxiliary circuitry required for the lab-circuits, as depicted in Fig. 3.

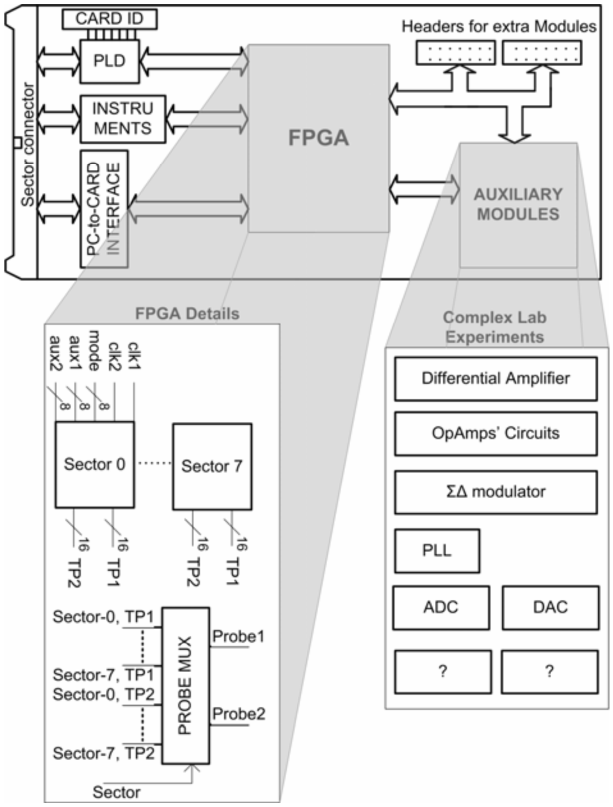



**Fig. 3. Hardware, instrumentation and resource server's software modules.**

Each card employs also a PLD, which is responsible for the card addressing and the configuration of the FPGA. Each of these cards can host 8-different analog, digital or mixed independent arbitrary circuits, since the FPGA is segmented into 8-sectors, each of them corresponding to a specific lab-experiment. The internal operation of the FPGA is controlled by a register file (Table II) which is employed within it.

As each sector of the FPGA can host either a specific multi-mode lab-experiment or a user's custom circuit, the mode register and two auxiliary registers control its operation mode and behavior. For example, a single sector could implement both synchronous and asynchronous counters, being realized as two different lab-experiments, while the selection is performed by the value of the mode register. Sector register points to the active sector, on which measurements are performed. Finally, when a measurement is carried out, two more registers, Probe1 and Probe2, assign the active nodes of the active sector, on which the two probes of the oscilloscope become physically connected through cross-point switches. Moreover, each card may be offline equipped with

| Name | Address | Width (bits) | Operation |
|---|---|---|---|
| Sector | 0 | 3 | Select the active sector |
| Probe1 | 1 | 4-6 | Select the active nodes of Oscilloscope's Ch-A |
| Probe2 | 2 | 4-6 | Select the active nodes of Oscilloscope's Ch-B |
| Aux1 | 3 | 8 | Auxiliary register 1 |
| Aux2 | 4 | 8 | Auxiliary register 2 |
| Mode | 5 | 8 | Sector's operation mode |

**Table I.FPGA register file.**

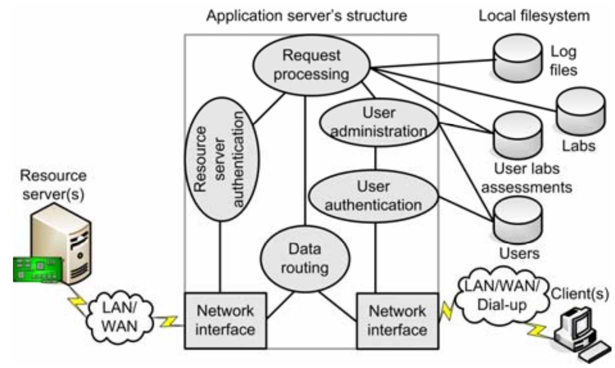



**Fig. 4. Software modules of the application server.**

additional on-board or external circuitry, in order to implement a wide range of more complex electronic circuits, including PLL, D/A or A/D, ΣΔ Modulators, etc.

When a client raises a measurement request, application server logs and routes this request to the proper resource server, which is able to support the under-measurement circuit. Afterwards, resource server has to accomplish multiple tasks, such as the authentication of the request and the lab-infrastructure (hardware and instrumentation) setup, so as to be prepared for the requested operations. A hardware administration module, within the resource server, is responsible for the aforementioned task, giving first priority to the completion of the request. This may lead to a real-time, online re-configuration of one's card FPGA, so as to implement the requested circuit, or even to the removal of an unutilized sector's circuit. As soon as the hardware is configured, the measurement is performed and the acquired data are transmitted back to the specific client, again via the application server. The above procedure has been designed so as to time-share the physical resources, in a FIFO priority, to all available requests.

### 3.3 Application server

Apart from the dataflow control and the routing procedure between client(s) and the resource server(s), application server is also responsible for the authentication and logging, as well as for the assessment and evaluation of its clients' actions, if educative usage is intended. The above presented characteristics and functionalities of the application server define its architecture, as depicted in Fig. 4. Additionally, each one of the application servers of the RMCLab platform needs to be offline aware of the resource servers, that are able to communicate with, and their list of supported circuits, which are dynamically acquired upon each successful transaction with one of the resource servers.

### 3.4 Client and Instructor Client

The client part of RMCLab's system has been designed so as to meet the requirements provided by the server-side. Thus, client embeds a specific interface, named as scenario interface, for supporting the remote monitor and

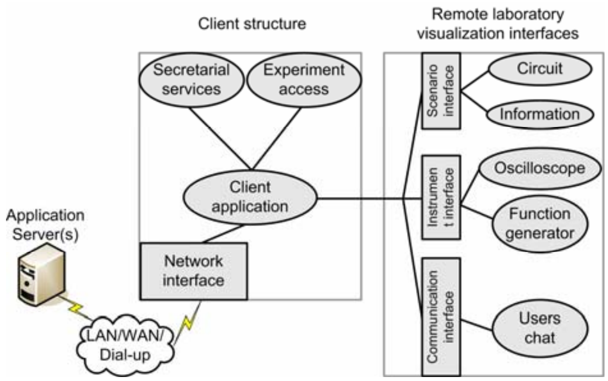



**Fig. 5. Software modules of the client.**

control of lab-infrastructure, and other full-functional and user friendly interfaces for lab-instrumentation (function generator, oscilloscope, frequency analyzer), as depicted in Fig. 5.

In more details, scenario interface provides a user with graphic information, related with the under study circuit. Additionally, grants to the user the control of circuit parameters (variable pots, caps, etc) and also the monitoring of any active node of the circuit, by selecting and setting the probes of the oscilloscope and the function generator on the circuit. Moreover, extra documentation, regarding the technical and theoretical aspects of the experiment, which can also be separated into multiple steps, can be presented to the user, via the scenario interface.

Finally, RMCLab's platform embeds an identical to the user's one client module, named as instructor client, which offers to a supervisor/instructor of the experiment the ability to replicate, monitor and control any online user's lab-environment. This feature is focused on the educational aspect of RMCLab platform, as it provides an instructor with the potentiality to observe closely and tutor efficiently the actions of any online users, concluding to a 'near-to-real' lab environment.

## 4. Advanced Properties of the RMCLab

The most important benefit of the RMCLab's system is that it provides its users with the feasibility to design their own custom circuits and test/measure them under real hardware and real instrumentation. RMCLab users can offline design almost any circuit using a separate software package (MAX+plus II or Quartus, both offered at no-cost for academic institutes from Altera). Using one of the aforementioned specific software packages, one can design his circuits following a reduced set of rules and confirm by simulation its proper operation. Once the design is verified at the client-side, it can be easily uploaded to the server-side and after a while (<15-seconds) he will be able to perform any measurement on his custom design, which is now implemented on real hardware, by employing real instrumentation.

The network architecture of RMCLab platform enables the world-wide distribution of resources, in terms of lab-experiments, by utilizing multiple application servers in a single network topology. Thus, instructors all over the world can take the advantages of employing a running lab-experiment and present it in their native language and personal point of view. Obviously, each supervisor has the opportunity to review his users' performance by his own criteria, according to the assessments rules for each experiment, that are defined in the RMCLab's application server, which is available and accessed by the supervisor, as resource server transparently executes the measurement requests.

RMCLab's advanced utilization modes are not limited within the above example. The real-time use of real hardware and real instrumentation can significantly contribute to the educational procedure, since it enables an instructor to prepare 'Active Lessons' and present in details, during a class, the operation of a circuit or a system under real world circumstances, while at the same time can be utilized as a mean of demonstration for expensive products.

## 5. The Realized RMCLab System

### 5.1 Technical Characteristics

The architecture described in section 3 has been implemented at the University of Patras, Greece. Current configuration is a cost effective implementation, which consists of a single PC, with a Celeron 2.4GHz processor and 394MB RAM, for both the resource and the application server of our running RMCLab system. This PC, running Windows 2000 Server, is permanently connected to the campus LAN and also to an Agilent 54622D mixed signal oscilloscope and an Agilent 33120A function generator, both controlled via RS232 @ 57600 and 19.2 baud-rate, respectively. The PC-interface with the hardware modules is implemented based on a custom, low-cost bus, through LPT in EPP mode. Each card of the hardware infrastructure contains an Altera FPGA of the FLEX8K series and also other components required for the implementation of the experimental circuits. The aforementioned infrastructure provides fast enough access and response (<5-secs per measurement) to the client requests, as summarized in Table II.

### 5.2 A Simple Educational Paradigm

Fig. 6 illustrates the measurement result, regarding the CLK and LOAD signals of an Early Decoding, Count Down, 4-bit Decimal Counter. The top part in Fig. 6

| Property | Average Delay (sec) |
|---|---|
| Hardware setup and measurement time | 3 |
| Compilation time of a custom | 10 |
| Hardware re-configuration time | 5 |
| Measurement delay from client side using PSTN line @56kbps | <5 |

**Table II.RMCLab time response characterictics.**

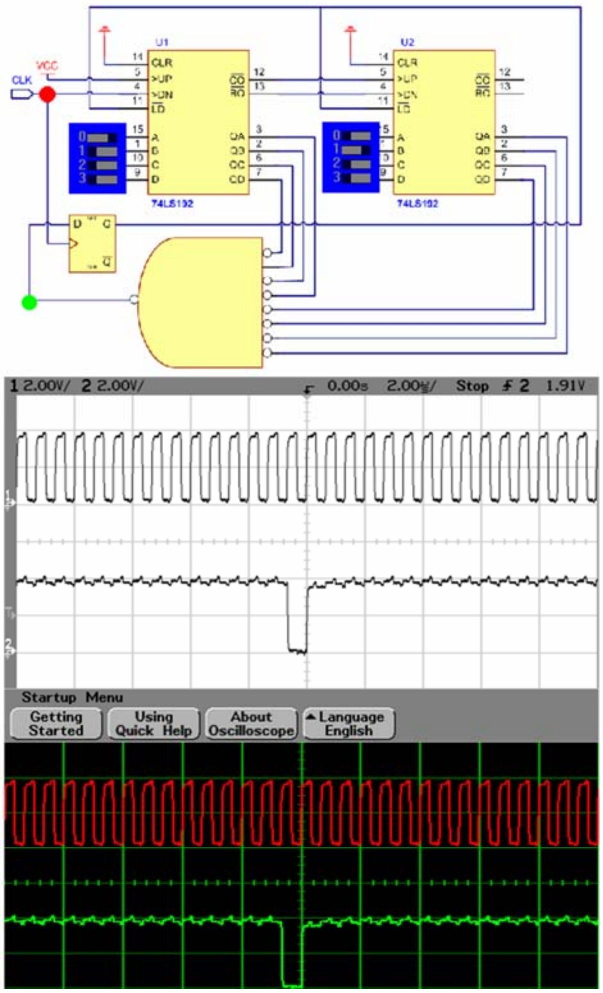



**Fig. 6. Educational paradigm of RMCLab usage.**

depicts the circuit information available to RMCLab users for the specific circuit, while the middle and bottom part depict the measurement representation of the real oscilloscope and the remote interface, respectively. Obviously, RMCLab system is able to take full advantages of the real hardware and real instrumentation utilization, providing measurements, regarding a wide variety of signals' properties, and the full control of both instruments (oscilloscope and function generator).

### 5.3 Educational Utilization

RMCLab system provides its educational services since March 2004 for the Dept. of Electrical and Computer Engineering of University of Patras, Greece, supporting classes of approximately 300-students, in two core-lessons, regarding Analog and Digital Electronics. Analog lab-experiments include 2-stage, feedback and cascade/folded-cascade amplifiers, whereas digital experiments include oscillators, multi-vibrators and a wide variety of counters, adders and accumulators.

A class of the Dept. of Electrical and Computer Engineering of our University has about 300-students. Students are grouped in teams of 3-to-4 persons per team,

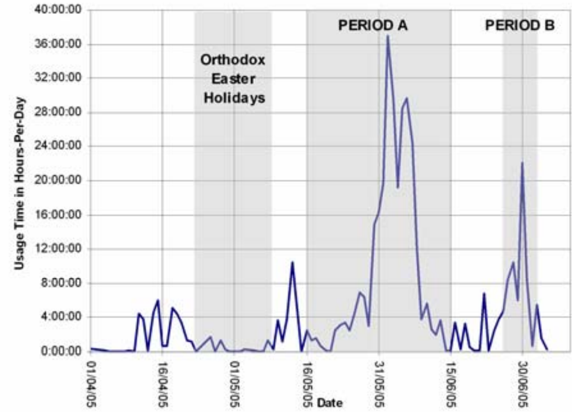



**Fig. 7. RMCLab's cumulative utilization during the second semester of academic year 2004-'05.**

in order to perform 6-obligatory lab-experiments per semester. For convenience, the same grouping was retained for accessing RMCLab services. Thus, for the second semester of academic year 2004-'05, RMCLab has to offer its services to 80-teams, for 2-obligatory lab-experiments. Four more obligatory lab-experiments had also to be carried out in the conventional way. The two RMCLab-based experiments referred to binary/decimal synchronous counters and special purpose programmable up/down counters, using the 74192 chip.

Students' obligations regarding the RMCLab-based experiments were announced at the beginning of April '05. As depicted in Fig. 7, they had to carry out these experiments in PERIOD A, while the re-examination took place in PERIOD B. It should be noticed that 6-teams did not fulfill their obligations to these labs.

During the same academic semester, RMCLab has been extensively employed for its services; thus 17200-measurements were logged on the platform's instrumentation, where 1666 of them regarded an introductory exercise, and 4458, 11076 measurements regarded the first and the second obligatory exercise, respectively. For the first obligatory exercise the performed measurements were 4.35-times more than the

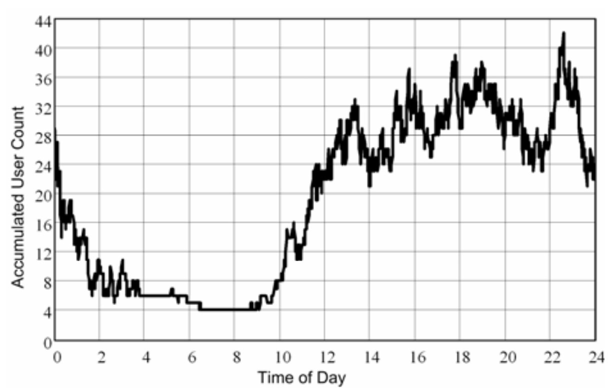



**Fig. 8. RMCLab's cumulative user count vs. time of the day for the second semester of academic year 2004-'05.**

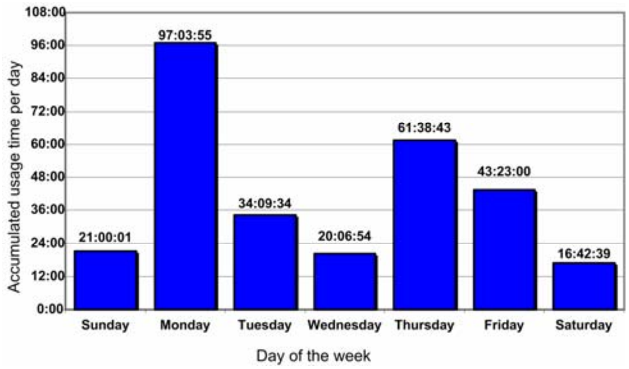



**Fig. 9. RMCLab's cumulative usage time vs. day of the week for the second semester of academic year 2004-'05.**

total number required for completing the exercise, while for the second one this ratio was reduced to 3.97, despite the fact that this exercise was significantly more demanding for the students. Students employed these oscillographs in their two lab-reports, which were of 25 and 55-pages in average, respectively.

During the same period, up-to 8-simultaneous requests have been raised to the RMCLab resource server, without importing any extra delay to the users' requests servicing. The 74-active teams fulfilled their lab obligations in 383h 30m 20s, thus, approximately 2.6-hours per lab-experiment. Fig. 8 and Fig. 9 depict the cumulative user count and the cumulative usage time of the RMCLab services versus the time of the day and the day of the week, respectively, for the aforementioned period.

Commenting on the statistics, derived by the use of RMCLab services, it can be pointed out that students have to acquire 4-measurements in average, in order to obtain the desired results; concluding that they are not experienced enough with the instrumentation and the RMCLab client application. On the other hand, RMCLab makes possible the easy, instant acquisition of multiple measurements. Finally, Fig. 8 and Fig. 9 prove that services like RMCLab enable students to take over their obligations in a reasonable time while exploiting efficiently their wide spread working hours.

## 6. Conclusion

The structure of RMCLab enables distribution of hardware and instrumentation resources, moreover makes possible the extensive exploitation of expensive lab-infrastructure, facilitating the wide spread of remote, real lab-experiments, which are indisputably valuable for engineers' education. Additionally, hardware re-configurability enables the remote implementation and measurement of electronic circuits, providing yet a high-valued educational service, while at the same time affects significantly students' psychological mood regarding the level of the offered education by their institute.

The intensive usage of the RMCLab system, during 5-academical semesters, for the courses of Analog and Digital Integrated Electronics, consisting of classes of about 300-students, has definitely proved the high value, for the students and for the instructors as well, of this educational tool, placing the proposed platform in a high position among related platforms, in terms of exploitation, capabilities and service quality.

It is anticipated that the proposed architecture guidelines along with the success of the RMCLab platform will motivate educational community to co operate, so as to develop an integrated World-Wide-Lab environment.

## Acknowledgements

Authors would like to thank Prof. G.D. Papadopoulos, S.A. Koubias and N.M.Avouris for their useful suggestions and their support to the promotion of the RMCLab platform. Development of several lab-experiments was supported by the Third Framework of E.P.E.A.E.K. II.

## References

[1] G. Mason, A handheld data acquisition system for use in an undergraduate data acquisition course, *IEEE Trans. on Education, 45*(4), 2002, 388-393.
[2] J. Sebastian, D. Garcia & F. Sanchez, Remote-access education based on image acquisition and processing through the internet, *IEEE Trans. on Education, 46*(1), 2003, 142-148.
[3] P. Lin & M. Lin, Design and implementation of an internet-based virtual lab system for eLearning support, *Proc. 5$^{th}$ IEEE International Conf. on Advanced Learning Technologies*, Kaohsiung, Taiwan, 2005, 295-296.
[4] A. Bagnasco & A. Scapolla, A grid of remote laboratory for teaching electronics, *2nd International LeGE-WG Workshop on e-Learning and Grid Technologies: a fundamental challenge for Europe*, Paris, 2003.
[5] F. Pianegiani, D. Macii & P. Carbone, An open distributed measurement system based on an abstract client-server architecture, *IEEE Trans. on Instr. and Meas., 52*(3), 2003, 686-692.
[6] J. Palop & M. Teruel, Virtual work bench for electronic instrumentation teaching, *IEEE Trans. on Education, 43*(1), 2000, pp. 15-18.
[7] C. Daly & J. Horgan, An automated learning system for java programming, *IEEE Trans. on Education, 47*(1), 2004, 10-17.
[8] T. Fjeldly & M. Shur, *Lab on the web: Running real electronics experiments via the internet* (Wiley-IEEE Press, 2003).
[9] C. Ko, B. Chen & J. Chen, *Creating web-based laboratories* (Advanced Information and Knowledge Processing, Springer, 2004).

Instrumentation and Measurement
Technology Conference – IMTC 2007
Warsaw, Poland, May 1-3, 2007

# An R-2R Ladder-Based Architecture for High Linearity DACs

Kostas A. Efstathiou and Dimitris S. Karadimas
Department of Electrical and Computer Engineering, University of Patras
Applied Electronics Laboratory, Rio Campus, 26500 Patras, Rio Campus, Greece
Phone: +30-2610-996-425, Fax: +30-2610-996-818, email: *efstathiou@ece.upatras.gr*

***Abstract*** *– Digital to Analog Conversion performance is mainly characterized by its resolution, linearity and speed. Additional implementation characteristics include area and power dissipation. This paper presents a DAC architecture based on the conventional R-2R ladder topology that is able to derive a high-resolution, high-linearity and high-speed DAC, while requiring reasonable operating power and implementation area in a standard CMOS technology. Simulation results, derived both from numerical and circuit-level simulations, point out that the proposed architecture is able to derive any desirable resolution and linearity, without requiring low mismatch technologies or expensive trimming procedures.*

***Keywords*** *– digital to analog conversion, resolution, linearity, settling time.*

## I. INTRODUCTION

Nowadays signal processing typical requirements are increased, resulting to employment of digital rather than analog processing, since digital processing has better potentials. Consequently, signal conversion is mandatory, since more often than not the input and the output signals are analog. Signal conversion requirements continuously increase, in terms of conversion-speed, resolution, accuracy and power dissipation. Current technology for Digital-to-Analog Conversion (DAC) demonstrates techniques that achieve significant improvements regarding the aforementioned characteristics. However, several trade-offs prevent the design and implementation of a DAC, able to satisfy all the requirements.

DAC's linearity and resolution state-of-the art prescribe techniques such as ΔΣ conversion [1] and dynamically calibrated current sources [2], which however have poor conversion speed. On the other hand, ladder-based DACs [3] can offer high conversion speed, small area and power consumption, but they fail to comply with increased resolution and linearity requirements. Techniques able to approach high performance on speed, linearity and resolution, include thermometer code DACs [4], at the expense of area and power dissipation, or laser-trimmed [5] R-2R ladder-based DACs, using expensive manufactory procedures. The combination of the aforementioned techniques, derive hybrid DACs [6] or self-calibrated R-2R ladders [7], thus counterbalancing the strict requirements of the digital-to-analog conversion procedure.

This paper presents a DAC architecture based on the conventional ladder, able to fulfill tough requirements regarding resolution, linearity, conversion speed and power dissipation. Section II presents the proposed DAC architecture, section III presents circuit and numerical simulation results and finally, section IV proposes an innovative method for fast and accurate calibration of the proposed DAC.

## II. ARCHITECTURE

Fig. 1 depicts the conventional current-mode R-2R ladder proposed by [8]. In this design, FETs labeled as $QH_X$ represent the R resistances of the ladder, while the 2R branches are composed by the in series combination of FETs $QV_{X,A}$ and $QV_{X,B}$ or $QV_{X,C}$. All FETs are assumed of the same size, while active FETs operate at the linear region. The voltage difference Vref-Vbias is assumed small enough and thus body effect does not affect significantly the mismatching of the FETs' $R_{ON}$. However, the technology mismatch, along with the small difference in the voltage $V_{SB}$ of the FETs, will indeed derive a mismatch on the FETs' $R_{ON}$, which bounds the linearity of the DAC.

The architecture proposed by [9] is based on the conventional R-2R ladder and is depicted in Fig. 2. In this architecture all FETs related to the R-2R topology are denoted as Q and are of the same size, while they operate at the linear region and have an $R_{ON}$ resistance.

Each FET labeled as $QV_{X,B/C}$ in Fig. 1 has been replaced in Fig. 2 by an *NC*-bits ladder, acting as a bit calibration ladder, that is identical to the one depicted in Fig. 1, and a current mirror that mirrors the output current of each branch to the DAC's output according to the *NC*-bits digital input word of the specific bit calibration ladder. This replacement does not affect the R-2R topology benefits, since the input resistance of each calibration branch equals to one's FET $R_{ON}$. This modified topology enables the trimming of the output current of each calibration branch, according to the specific calibration word. Therefore, the modified topology offers the capability to trim each branch's current close-enough to the

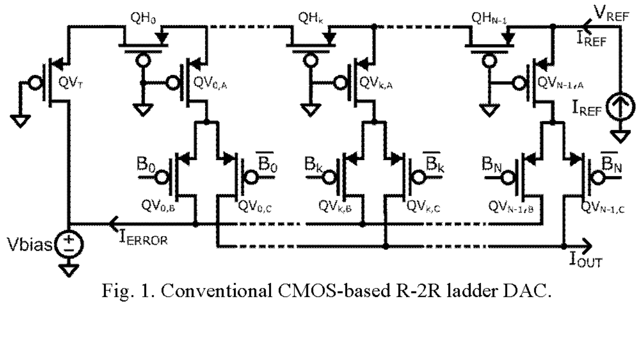



Fig. 1. Conventional CMOS-based R-2R ladder DAC.

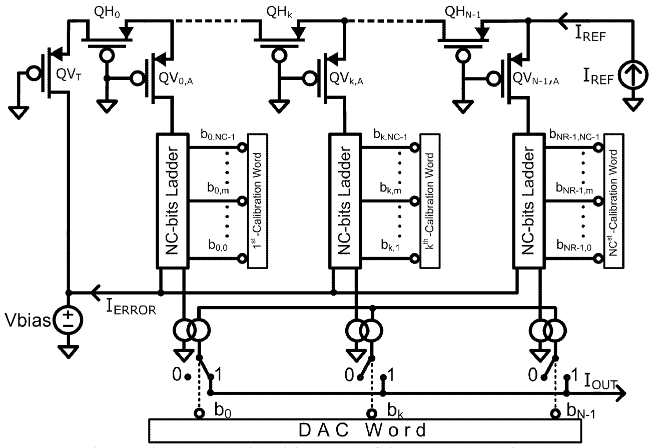


Fig. 2. The proposed architecture of the current mode DAC

optimum one. Apparently, the resolution of each calibration branch (*NC*) plays a major role, since it assigns the accuracy of the trimming procedure.

Fig. 3 denotes that the proposed ladder architecture can be synthesized by combining *NR* similar blocks, assuming that calibration words for each branch are settled using a simple serial interface. Each one of these blocks consists of the bit-calibration ladder and the DAC backbone (FETs $QH_X$, $QV_{X,A}$), as well as the required circuitry for current-cascading and the interface for assigning the calibration word. Digital signals Din, Dout, SCLK and LOAD are responsible for the assignment of the calibration word. Analog signals Cin and Cout are employed for the current-cascading, while $CI_{OUT}$ is the output current of each block. Finally, the digital signal $b_k$ represents the corresponding bit of the DAC's digital input word.

Each k-block of Fig. 3 (k=0…*NR*-1), depicted in more details in Fig. 4, consists of a single element of the DAC backbone ($QH_k$ and $QV_{k,A}$). Each calibration branch is composed by *NC*-similar blocks, one for each calibration bit, and a current mirror (FETs M1 to M4), that mirrors the branch's output current to the output node, while the error current is let to the ground (through FETs E1 and E2) and FETs S1 and S2 switch the bit's current to the ladder's output.

Finally, each block $CBit_m$ (m=0…*NC*-1),, representing a calibration bit, depicted in Fig. 4, consists of one bit calibration register and a single bit calibration ladder, as depicted in Fig. 5. Calibration register consists of two D-FF's and enables the serial-input/ parallel-load of the DAC's calibration words. The single bit calibration ladder is identical to the cell of the ladder as depicted in Fig. 1.

As soon as the calibration words have been assigned, all FETs, labeled as Q, M and E operate at a steady-state, regardless of the DAC's digital input. Only FETs S1 and S2, of each calibration branch, alternate their operating conditions according to the DAC word. This implies that most of the circuitry operates in constant bias conditions, as only a small portion of it has to switch, thus contributing to high speed performance. Since current, in most FETs, is constant, the power consumption of the circuit will be approximately close to $2\times VDD\times I_{REF}$.

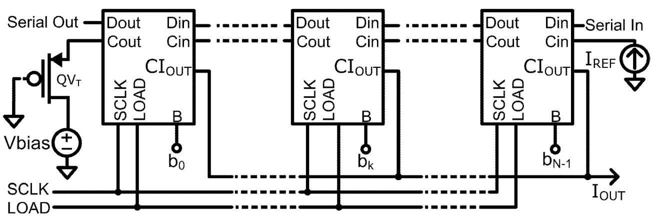


Fig. 3. Block diagram of the proposed DAC circuitry.

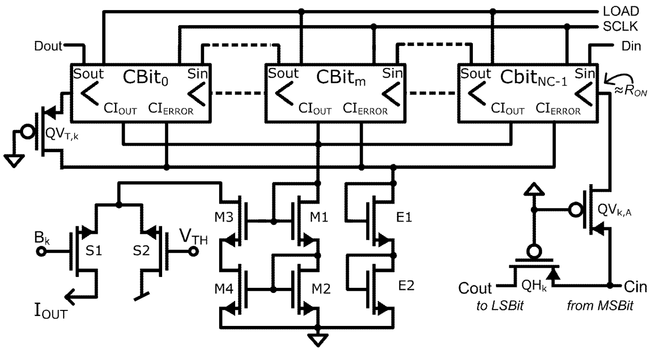


Fig. 4. Internal architecture of a ladder's block.

## III. SIMULATION RESULTS

The proposed architecture has been simulated both numerically, employing Matlab for studying the DAC's behavior over a wide range of cases regarding the resolution and the resistors' mismatch, and at circuit level, employing Tanner circuit simulator on an 8-bits DAC with 8-bits calibration branches, implemented using a standard CMOS technology.

### A. *Numerical Simulation*

Numerical simulations, covering a wide range of resolutions (*NR*=4-16) and calibration lengths (*NC*=1-16), for a wide range of components' mismatches ($m_e$=0.1%-10%) have been performed using Matlab.

Fig. 6 presents the estimated mean linearity (INL) of a ladder employing components of constant mismatch ($m_e$=0.01 or 1%), for a range of resolutions (*NR*=8-15) and calibration lengths (*NC*=1-16). Apparently, the more calibration length is increased, the more the DAC's linearity gets improved, since the current trimming capabilities are exponentially increased.

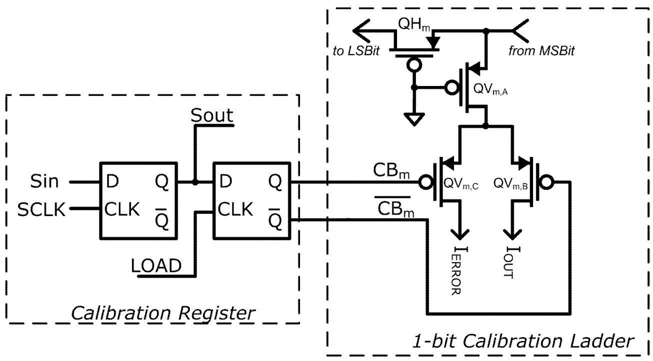


Fig. 5. Internal architecture of a single bit of a calibration ladder.

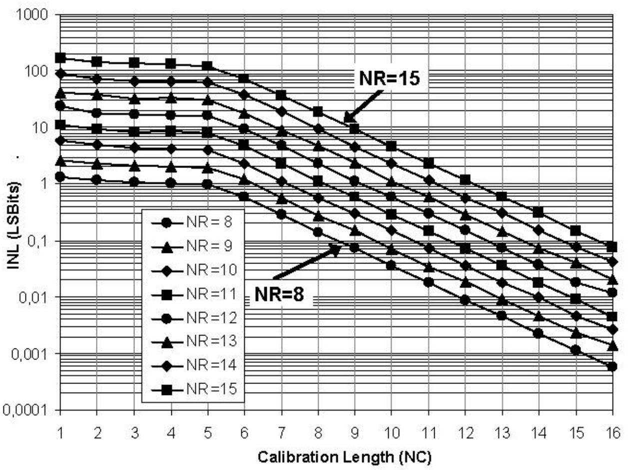



Fig. 6. Simulated INL mean value for mismatch me=0.01 (1%) for various resolutions.

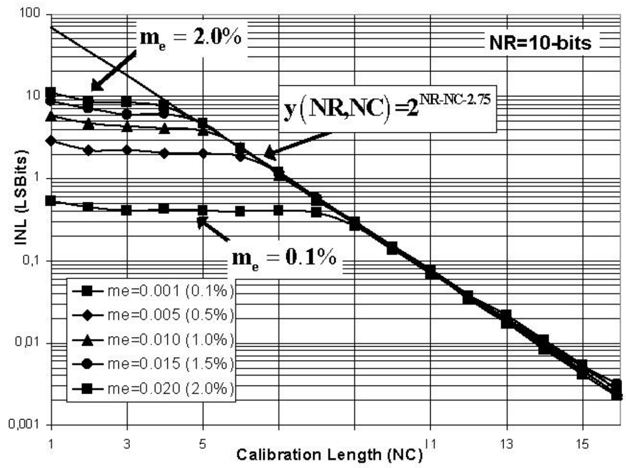



Fig. 7. Simulated INL mean value for 10-bits resolution at various mismatches.

However, for small values of calibration length ($NC$<5 for $m_e$=1%) no linearity improvement can be considered, which is reasonable, since calibration branches do not have the adequate resolution in order to trim the current contribution of each bit.

Fig. 7 presents the estimated mean INL for a certain ladder's resolution ($NR$=10), employing components of a wide range of mismatches ($m_e$=0.1%-2%) and for a wide range of calibration lengths ($NC$=1-16). The pattern of these simulated results leads to a twofold approach, regarding the calibration length $NC$:

- When calibration length $NC$ is small, the proposed architecture does not improve DAC's linearity, which has a certain value that depends only on the DAC's resolution $NR$ and the components' mismatch $m_e$.
- When calibration length $NC$ is adequate, the ladder's linearity depends on the DAC's resolution $NR$ and the calibration length $NC$ while does not depend on the components' mismatch $m_e$.

This indicates that the proposed architecture is able to provide high-resolution, high-linearity ladders, with out requiring advanced technology, in terms of components' mismatch. Equation (1) is the empirical formula that fits on the horizontal lines of Fig. 6 and Fig. 7, thus in the region where the proposed architecture cannot achieve linearity improvement. Therefore, (1) gives the estimated mean INL value of the conventional ladder as a function of its resolution, $NR$, and the mismatch, $m_e$, of the employed components,

$$\overline{\mathrm{INL}}(m_e, NR) = 0.577 \cdot m_e \cdot 2^{NR} (LSBits) \quad (1)$$

Equation (2) is the empirical formula that fits on the diagonal line of Fig. 7 and derives the estimated mean INL value of the proposed ladder, as a function of the resolution and the calibration length $NC$,

$$\overline{\mathrm{INL}}(NR, NC) = 2^{NR-NC-2.25} (LSBits) \quad (2)$$

According to the simulation results, the proposed ladder architecture is able to derive a linear DAC of any resolution, regardless of the technology mismatch. Equation (2) implies that designing a ladder with calibration length equal to the resolution length minus two ($NR$-2) will produce a DAC with mean INL value less than one. Further statistical analysis on the simulation results pointed out that when $NC$ equals to $NR$-1, the maximum derived INL would always be less than 1-LSBit, regardless the technology's mismatch. Apparently, this is a significant conclusion, since it implies that the proposed architecture is able to implement high-resolution, high-linearity and high-speed D/A converters employing a simple methodology on a standard CMOS technology.

*B. Circuit simulation*

An 8-bits DAC, with 8-bits calibration branches, has been simulated using the AMS08 technology and the Tanner tool. During the simulation, zero mismatch error for the FETs' size and ideal connections between them has been assumed. Despite this, circuit simulation derives an INL of 5-LSBits if employing the word 255 (0xFF) as the calibration word for all branches, thus concluding to an un-calibrated DAC.

Calibration procedure requires the adjustment of the calibration words, so as each bit to contribute with a current value as close as possible to the ideal one. For each calibration word of the MSBit, a set of calibration words for the rest bits of the DAC that derive the minimum INL can be found.

As Fig. 8 denotes, there are many sets of calibration words, one set for each calibration word of the MSBranch; however, one set would provide the best results. Employing the aforementioned calibration algorithm, the best INL (0.054-LSBits) achieved by setting the value 237 (0xED) as the calibration word for the MSBit. However, it should be emphasized that, as it can be seen in Fig. 8, for any calibration word of the MSBit, in the range 255-192 (0xFF-0xC0), the worst INL achieved was less than 0.5-LSBits. Additionally it should be emphasized that according

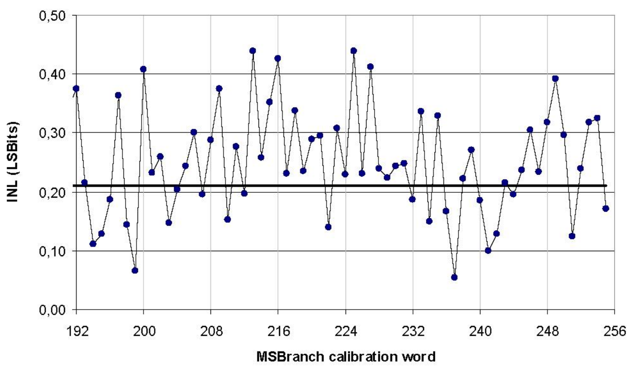



Fig. 8. INL achieved for a range of calibration words for the MSBranch.

to the empirical formula (2), derived by the numerical simulations, the estimated mean INL for this DAC is $2^{-2.25} \approx 0.21$, which is in accordance to the achieved INL value.

The current switching time for any possible transition did not exceeded the 20ns, a time that can be considered as more than satisfactory for the employed technology.

## IV. CALIBRATION

Calibration of the proposed ladder demands the estimation of the proper calibration word, for each branch, so as to equalize their normalized currents. The first step is to calculate the normalized current for each branch and for each calibration word. The second step requires finding for each possible calibration word, of the MSB branch, the nearest normalized currents of the rest branches, and calculating the error matching, by applying the least square algorithm. Finally, the best linearity will be achieved by applying the set of calibration words that provides the minimum error. However, if applying any other set of calibration words derived by the calibration algorithm, a DAC with significantly improved linearity characteristics will be obtained in comparison to the conventional R-2R ladder.

In real world, the calibration of a ladder, based on the proposed architecture, cannot be easily performed by following the aforementioned calibration procedure, applied for the simulation, since it is not easy to accurate measure all the currents.

An elegant calibration method, based on the properties of an alternative phase detector, designed towards to PLL-based frequency synthesis is presented in [10], and can be employed to facilitate and accelerate the calibration of the proposed ladder. This phase detector, named as “Dual Input Phase Accumulator” (DIPA), derives at its output the phase difference of two non-equal frequencies $F_1$ and $F_2$ by applying a normalization technique. DIPA consists of a digital portion that implements the required function and a DAC for generating the analog output corresponding to the normalized phase difference. Additional to the two inputs for the frequencies $F_1$ and $F_2$, DIPA employs two more digital inputs $W_1$ and $W_2$ where the phase normalization factors are applied. DIPA’s proper operation requires applying on $W_1$ a value proportional to the frequency applied on input $F_2$ and consequently, a value proportional to the frequency $F_1$ should be applied at input $W_2$.

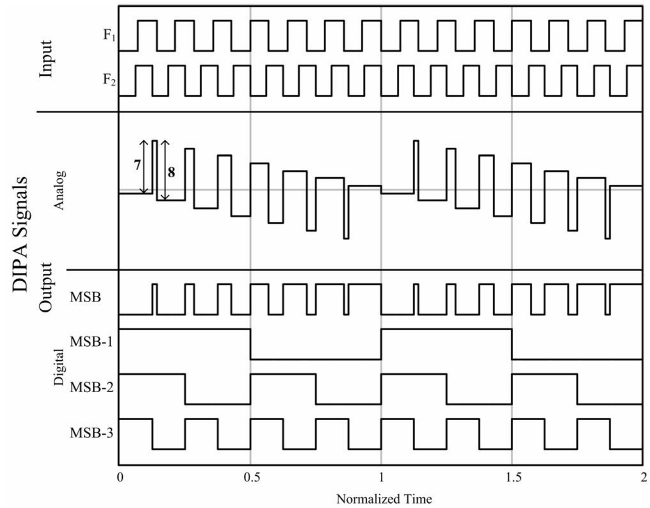



Fig. 9. DIPA’s input/output analog/digital signals.

An example that demonstrates the operation of a 4-bits DIPA (*NR*=4) is depicted in Fig. 9, assuming that two frequencies with normalized values $F_1=7$ and $F_2=8$ are applied on its two inputs. On each falling edge of $F_1$ the analog output of the DIPA is decreased by $W_1=8$, while on each edge of $F_2$ is increased by $W_2=7$. Thus, DIPA’s output mean value remains constant, indicating that the normalized phase difference of the two input frequencies is constant. It has been proven in [10] that DIPA’s AC output current in this example can be expressed in Fourier series as shown in (3),

$$I_{DIPA}(t)=\frac{W_2 I_{FS}}{2^{NR}\pi}\sum_{i=1}^{\infty}\frac{1}{i}\sin(2\pi i F_2 t)-\frac{W_1 I_{FS}}{2^{NR}\pi}\sum_{i=1}^{\infty}\frac{1}{i}\sin(2\pi i F_1 t), \quad (3)$$

where $I_{FS}$ is the full scale output current of the DAC.

Apparently, DIPA’s output signal consists of harmonics of the two input frequencies on the condition that the employed DAC is ideal. Observing, in Fig. 9, the digital waveforms that drive DIPA’s DAC, it is easy to conclude that a mismatch error between the current contributions of MSB and MSB-k will superimpose to the ideal output, a square waveform with a frequency $2^{k-1}$-times the beat frequency of $F_1$ and $F_2$ ($F_b=|F_2-F_1|$) and an amplitude Ierr(k) proportional to the aforementioned mismatch error:

$$Ierr(k)=I_{MSB-k}-\frac{I_{MSB}}{2^k} \quad (4)$$

and is expressed in time domain employing Fourier series as:

$$Ierr(k,t)=\frac{2}{\pi}\left(I_{MSB-k}-\frac{I_{MSB}}{2^k}\right)\sum_{i=1,3,5...}\frac{1}{i}\sin\left(2^k\pi i F_b t\right). \quad (5)$$

Apparently, in the case of employing a non-ideal DAC, DIPA’s output will include harmonics of the beat frequency, thus low frequencies. It is noticeable that, in the example depicted in Fig. 9, the amplitude of 1st, 2nd and 4th harmonic of the beat frequency that will appear at DIPA's output will be directly proportional to the mismatch error between the analog contribution of the MSB and MSB-1, MSB-2 and MSB-3 respectively. By measuring the amplitude of the aforementioned harmonics of the beat frequency, the mismatch errors of the corresponding bits could be easily calculated by multiplying these amplitudes with $\pi/2$. The phase of these harmonic point out the sign of the amplitude, thus indicate if the current contribution of a bit is more or less the ideal value.

For calibrating MSB-1, 2 and MSB-3 bits of a DAC it is required to minimize the corresponding errors, thus to set the calibration words that minimize the 1st, 2nd and 4th harmonic of the beat frequency.

A simple strategy for calibrating rest least significant bits of a multi-bits DAC, assumes as MSB the already calibrated MSB-k (k=1…*NR*-1) and requires shifting DIPA's 4-bits digital output bus down on the least significant, uncalibrated bits of DAC's input bus. Then the same calibration procedure is applied, assuming MSBits already calibrated. This simplified calibration procedure requires reduced infrastructure in terms of hardware (4-bits DIPA) and instrumentation (low frequency resolution spectrum analyzer), however may not conclude to the best calibration, since reference current is not unique.

In order to achieve the best calibration results for a multi-bits DAC, also a multi-bits DIPA should be employed so as to drive simultaneously all the DAC bits. Assuming that an *NR*-bits DAC has to be calibrated also an *NR*-bits DIPA should be employed. The ratio of the two input frequencies $F_1/F_2$ should be assigned to be $(2^{NR-1}-1)/(2^{NR-1})$ and phase-normalizing input words to the DIPA should be $W_1=2^{NR-1}$ and $W_2=2^{NR-1}-1$ respectively. In this case, calibration procedure requires minimizing the amplitudes of *NR*-1 specific harmonics of the beat frequency $F_b$ located at frequencies $2^{k-1}F_b$ ($1 \le k \le NR-1$). Based on (5), a formula that assigns DIPA's harmonic contents due to the DAC's non-linearity can be found:

$$\mathrm{Ierr}(t)=\sum_{k=1}^{NR-1}\left(\frac{2^k I_{MSB-k}-I_{MSB}}{2^{k-1}\pi}\sum_{i=1,3,5\ldots}\frac{1}{i}\sin\left(2^k\pi i F_b t\right)\right) \qquad (6)$$

According to the above formula DAC's non-linearity generates harmonics of the beat frequency $F_b$. Further investigation on this formula discloses that bandwidth below frequencies $F_1$ and $F_2$ is filled-up with harmonics of $F_b$. Each one of these harmonics is generated due to a single bit mismatch, however the first and consequently the strongest harmonic of MSB-k bit is located at frequency $2^{k-1}F_b$ ($1 \le k \le NR-1$). This is a significant property of the DIPA since it facilitates the accurate measurement of the mismatch contribution of each bit of the DAC, and consequently its accurate, fast and easy calibration. DIPA based calibration procedure has been confirmed by numerical simulations using Matlab.

In system calibration of the proposed architecture DAC is also an option. Since the performance of a system is affected by the linearity of the employed DAC, it can be calibrated using as criterion the improvement of the system's performance.

## V. CONCLUSION

A current-mode, MOST-only, R-2R ladder-based, digitally calibrated DAC architecture has been presented in this paper. The architecture and circuitry of the DAC has been discussed in details along with the properties that affect the linearity, the conversion speed and the power dissipation. The implementation of the proposed DAC was based on standard CMOS technology and requires no special treatment for the analog components regarding the mismatch error, concluding in a low cost design. Simulation results have certified the capabilities of the proposed DAC architecture, in achieving high linearity. Moreover, empirical formulas have been derived, interpreting the design properties of the ladder as a function of its desired linearity. Finally, an innovative method for accurately measuring the contribution of each DAC's bit to the overall DAC's non-linearity has been proposed, aiming to simplify and accelerate the calibration procedure of the proposed ladder.

## REFERENCES


[1] T. Shui, R. Schreier and F. Hudson, "Mismatch shaping for a current-mode multibit delta-sigma DAC," IEEE J. Solid State Circuits, vol. 34, Mar. 1999, pp. 331-338

[2] M. P. Tiilikainen, "A 14-bit 1.8-V 20-mW 1-mm2 CMOS DAC," IEEE J. Solid State Circuits, vol. 36, Jul. 2001, pp. 1144-1147.

[3] M. P. Kennedy, "On the robustness of R-2R ladder DAC's," IEEE Trans. Circuits & Syst. II, vol. 47, Feb. 2000, pp. 109-116.

[4] M. Vadipour, "Gradient error cancellation and quadratic error reduction in unary and binary D/A converters," IEEE Trans. Circuits & Syst. II, vol. 50, Dec. 2003, pp. 1002-1007.

[5] "R/2R ladder networks," Application Note AFD006, International Resistive Company, Inc., 1998.

[6] Y. Cong and R. L. Geiger, "A 1.5 V 14-bit 100ms/s self-calibrated DAC," IEEE J. Solid State Circuits, vol. 38, Dec. 2003, pp. 2051-2060.

[7] B. Vargha, I. Zoltan, "Calibration algorithm for current-output R-2R ladders," IEEE Trans. Instr. and Meas., vol. 50, Oct. 2001, pp. 1216-1220.

[8] K. Bult and J. Geelen, "An inherently linear and compact MOST-only current division technique," IEEE J. Solid State Circuits, vol. 27, Dec. 1992, pp. 1730-1735.

[9] D.S. Karadimas, D.N. Mavridis, and K.A. Efstathiou, "A Digitally Calibrated R-2R Ladder Architecture for High Performance Digital-to-Analog Converters", IEEE International Symposium on Circuits and Systems, Kos, Greece, May 2006.

[10] K. Efstathiou, G. Papadopoulos, Implementation of a High Speed Frequency Synthesizer employing a Dual Input Phase Accumulator, International Journal of Electronics, pp. 43-56, No. 1 Vol. 87, January 2000

Instrumentation and Measurement
Technology Conference – IMTC 2007
Warsaw, Poland, May 1-3, 2007

# A Remote Electrical Engineering Laboratory based on Re-Configurable Hardware

K. Efstathiou, D. Karadimas and K. Zafeiropoulos
Department of Electrical and Computer Engineering, University of Patras
Rion Campus, 26500, Greece
Phone: +30-2610-996-425, Fax: +30-2610-996-818, Email: *efstathiou@ee.upatras.gr*

***Abstract*** *– This paper describes an Internet-based laboratory, named Remote Monitored and Controlled Laboratory (RMCLab) developed at University of Patras, Greece, for electrical engineering experiments. The key feature of this remote laboratory is the utilization of real experiments, in terms of instrumentation and under measurement circuits, rather than simulation or virtual reality environment. RMCLab's hardware infrastructure contains multiple reconfigurable sub-systems (FPGAs), which can be enhanced by almost any analog expansion module. The main characteristics of this system include the versatility of the hardware resources, due to the dynamic reconfiguration potentiality and the low cost of the hardware components. Moreover, this system enables its users to test, in real time, their own custom circuit designs. The paper concludes with a specific example regarding an elementary circuit in digital electronics. RMCLab can be accessed via the web through http://www.apel.ee.upatras.gr/rmclab.*

***Keywords*** *– Client-server architecture, remote laboratory, distributed instrumentation and resources.*

## I. INTRODUCTION

The exponential growth of computer and internet technology enables the development of complex, hybrid systems such as remote laboratories where experiments can be remotely accessed, monitored and controlled [1]-[3]. This new interpretation of the measurement process offers to anyone the opportunity to interact with the laboratory at any time, reducing at the same time the experiment cost per user and extending the capabilities of the entire experimental framework.

Paradigms of using these advanced facilities apply either for educational purposes [4]-[5] or for products' advertisement. Remote laboratories can offer high-level experimental training and experience, when they are able to realize, support and interact with real experiments, rather than present simulation results or simple depiction of reality. Additionally, expensive, often dedicated experiments, of modern, cut-edge technology can be shared worldwide, contributing thus to a high-valued remote laboratory framework.

Many internet-enabled software systems that afford distance laboratories via simulated, virtual environments can be found in the web [6]-[7]. These software systems often integrate many of the desired functionalities, especially from the user's side, such as accompaniments to the experiment documentation, communication support and collaboration among their users. Although modern simulators can accurately estimate circuits' performance, the employment and utilization of real circuits and real instrumentation, for electrical engineering laboratories [8], ensures the measurements' reliability, while at the same time increases the educative value of the remote laboratory.

Remote laboratories offering access to real lab-experiments and real instrumentation also exist, however the majority of them cannot share their resources simultaneously to many users, thus they fail to serve and support large classes of several hundreds of students.

This paper presents the specifications and the basic structure of an integrated remote laboratory platform that enables the instant remote access to real lab-experiments, employing real hardware and real instrumentation. This platform, named Remote Monitored and Controlled Laboratory (RMCLab), is able to provide high-level services to a great number of users for a wide-range of real electrical engineering experiments; either pre configured, reconfigurable or customizable, at a very low hardware infrastructure cost. RMCLab offers its services since March 2004, at the Dept. of Electrical and Computer Engineering of University of Patras, Greece, where it was developed and implemented.

## II. PROPOSED APPROACH

The basic purpose of the developed platform is to provide high-quality lab-training in electrical engineering subjects to students, all over the world. The design of such a remote laboratory for real-time, internet-based lab-experiments, should consider all aspects of the system, including communication and data flow, as well as instrumentation and hardware control [9]-[10]. RMCLab system has been designed so as to integrate all potentials of a physical laboratory to a simple user interface, among with other sub-systems, such as lab-administration, instrument operation and hardware management.

The primary service that RMCLab platform should provide to its users is the possibility to study on the lab-experiment subjects, by accomplishing their measurements at any time and from anywhere. For this reason, RMCLab's basic specification is defined as the ability to serve at any time, simultaneously and at real time, any

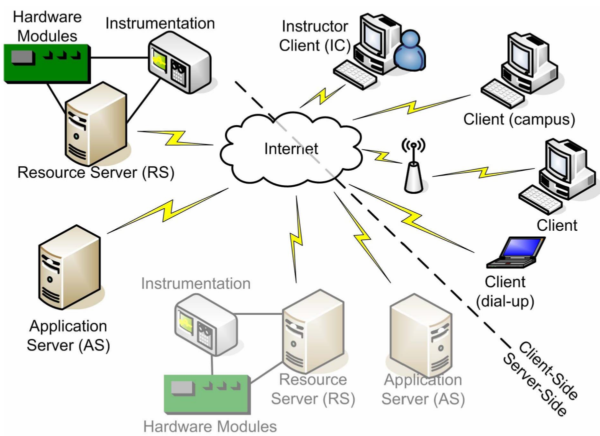



Fig. 1. RMCLab system overview.

potential user for any available lab-experiment. On the other hand, an integrated remote laboratory platform should reinforce lab-administrator's tasks and responsibilities, regarding the experiment setup, hardware and instrumentation control, users' management and also lab-maintenance. RMCLab offers also many kinds of assessment functionalities for the students' lab-skills, regarding the lab-experiments, such as the assignment of several different evaluation criteria (measurements, instrument settings and multi-type questions, etc), so as the whole platform can be configured as an advanced tool for automated, high level educational services, an aspect that characterizes the offered educational activities and also our initial motivation.

## III. ARCHITECTURE

RMCLab system has been developed based on the conventional client-server architecture, expanded in the server-side, as depicted in Fig. 1, and consists of the following basic entities: client, instructor-client (IC), application server (AS), resource server (RS) and lab-infrastructure, including the real instrumentation and all the hardware modules.

### *A. Network topology*

The server-side of the proposed architecture employs at least two sub-servers; the resource server and the application server. This structure could also be replicated in a more complex network topology. Resource server manages and operates hardware and instrumentation resources, providing to application server an abstract layer for communication that enables access to lab-infrastructure.

Application server undertakes the data flow control task between clients and the physical remote laboratory, realized by the resource server and the lab-instrumentation. The intermediary role of the application server is mandatory, since clients should not be to directly communicate with any of the resource servers. The communication between application server(s) and the resource server(s) or the client(s) is based on a custom, abstract language that integrates all potential tasks of a conventional, physical laboratory.

This topology simplifies the architecture of the server side and expands platform's capabilities, as it facilitates the robust development and customization of a resource server. Moreover, it enables many application servers to utilize the components shared by the resource server(s). As a result, users all over the world are able to transparently access, via the application server(s), these shared resources. Additionally, application server grants to its clients transparently access to the real resources of the physical laboratory, thus increasing system's robustness, flexibility and expandability.

### *B. Hardware, Instrumentation and Resource server*

The real measurement laboratory is based on a low cost and easy implementation, while it is realized around the resource server. Resource server is equipped with suitable, interfaces toward the signals of lab-experiments (both digital and analog experiments), via a custom, LPT based bus, and the instruments, via the RS232 interface or the GPIB bus, as depicted in Fig. 2.

Multiple types (standard, programmable, pre configured or re-configurable) of analog, digital or mixed circuits can be hosted in the platform's resource server(s). For this reason RMCLab's hardware is outfitted with a motherboard that is able to host up to 64-cards, where each of them incorporates an FPGA and extra auxiliary circuitry required for implementing the lab-circuits, as depicted in Fig. 3.

Each card employs also a PLD, which is responsible for the card addressing and the configuration of the FPGA. Each of these cards can host 8-different analog, digital or mixed independent arbitrary circuits, since the FPGA is segmented into 8-sectors, each of them corresponding to a specific lab-experiment. The internal operation of the FPGA is controlled by a register file (Table 1) which is employed within it.

As each sector of the FPGA can host either a specific multi-mode lab-experiment or a user's custom circuit, the mode register and two auxiliary registers control its operation mode and behavior, while sector register points to the active sector, on which measurements are performed. Therefore, a single sector could implement alike lab-circuits, which can be

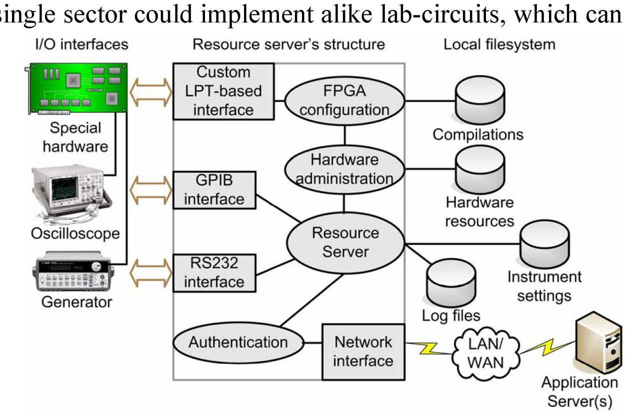



Fig. 2. Software and hardware modules of the Resource server.

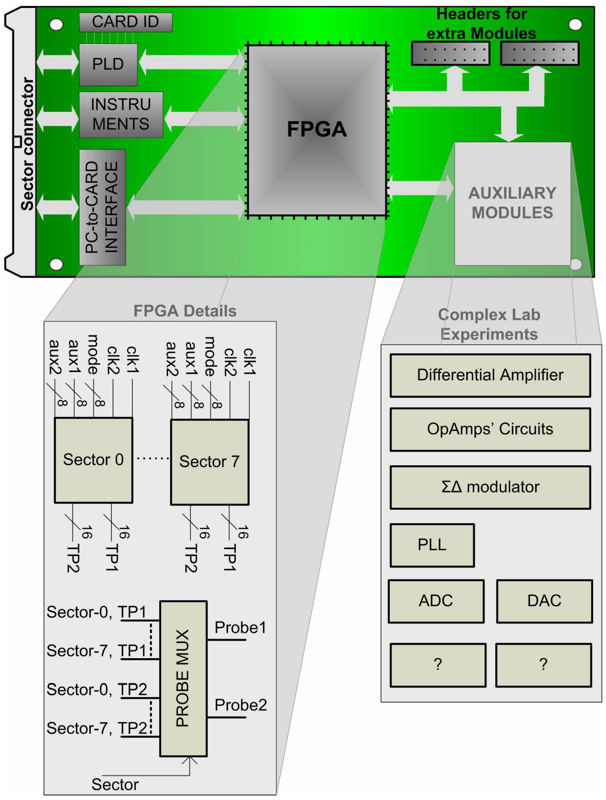



Fig. 3. Hardware architecture.

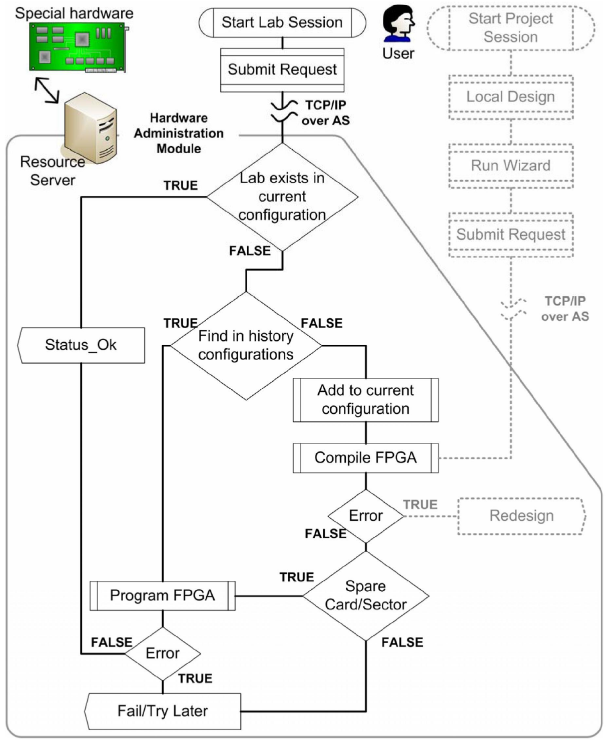



Fig. 4. Hardware administration module flowchart.

externally presented as different lab-experiments, while the selection of the operation is performed by the value of the mode register. For example, in our case, both synchronous and asynchronous digital counters are implemented in the same sector, but are presented as two different lab-experiments. Moreover, when a measurement is carried out, two extra registers, Probe1 register and Probe2 register, assign the active nodes of the active sector, on which the two probes of the oscilloscope become physically connected through cross point switches. Finally, each card may be offline equipped with additional on board or external circuitry, in order to implement a wide range of more complex electronic circuits, including PLL-based Frequency Synthesizers, several types of D/A or A/D Converters, ΣΔ Modulators, etc.

The aforementioned hardware architecture characteristics in combination with the network topology complexity necessitate an elegant and efficient management of the hardware resources and the measurement requests. A hardware administration module, depicted in Fig. 4, within the resource server, undertakes this management role. A measurement transactions starts when a client raises a measurement request, regarding either a pre-configured lab-experiment or a user's custom circuit. The later is discussed in details in Section IV. After the request is raised to the corresponding application server, by a client, is logged and forwarded, via the same application server, to the proper resource server, which supports the under measurement circuit. Afterwards, resource server has to accomplish multiple tasks, such as the authentication of the request and the lab-infrastructure (hardware and instrumentation) setup, so as to be prepared for the requested operations. This may lead to a real-time, online re-configuration of one's card FPGA, so as to implement the requested circuit, or even to the removal of an unutilized sector's circuit, if an empty sector cannot be found. As soon as the hardware is configured, the measurement is performed and the acquired data are transmitted back to the specific client, again via the application server. The above procedure has been designed so as to time-share the lab-infrastructure, in a FIFO priority, to all available requests.

Table 1. FPGA register file

| Name | Address | Width (bits) | Operation |
|---|---|---|---|
| Sector | 0 | 3 | Select the active sector |
| Probe1 | 1 | 4-6 | Select the active nodes of Oscilloscope's Ch-A |
| Probe2 | 2 | 4-6 | Select the active nodes of Oscilloscope's Ch-B |
| Aux1 | 3 | 8 | Auxiliary register 1 |
| Aux2 | 4 | 8 | Auxiliary register 2 |
| Mode | 5 | 8 | Sector's operation mode |

### *C. Application server*

Apart from the dataflow control and the routing procedure between client(s) and the resource server(s), application server is also responsible for the authentication and logging, as well as for the assessment and evaluation of its clients' actions, when educative usage is intended. The above presented characteristics and functionalities of the application server define its architecture, as depicted in Fig. 5. Additionally, each one of the application servers of the RMCLab platform needs to be offline aware of the resource

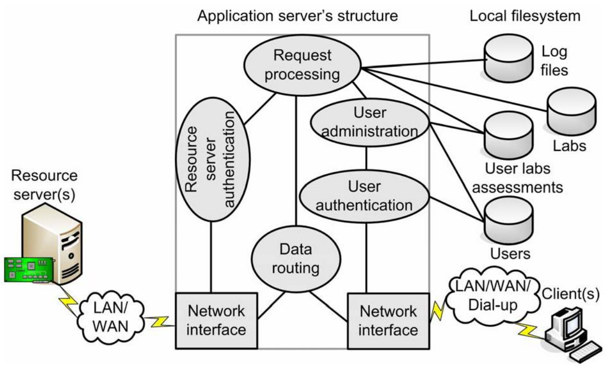



Fig. 5. Software modules of the application server.

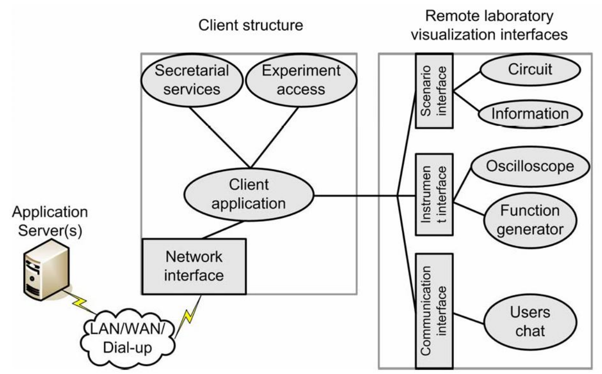



Fig. 6. Software modules of the client.

servers, that are able to communicate with, and their list of supported circuits, which are dynamically acquired upon each successful transaction with one of the resource servers.

Application server comprises also an advanced tool for the development and maintenance of a laboratory class that is available to the administrator of the laboratory, as can it be unrelated to the location of the physical laboratory, realized by the resource server, the hardware and the instrumentation. The development and maintenance of a laboratory class has been merged into a single database system, which contains all the required data for the design and assignment of a lab-exercise. Additionally, the same database system includes data regarding the students. One lab-exercise may consist of several active-images, which correspond to the real lab-circuits. For each lab-circuit, test-points and active-elements (switches and variable components) can be assigned. Hardware properties required for the assignment of the test-points and the active-elements are specified in the custom abstract language, used by the RMCLab system. On the other hand, a lab-exercise is separated in several steps, where each step may contain information, regarding the theoretical and practical aspects, measurements, multiple choice questions and a free-text question. Moreover, assessment rules may be provided in each lab-step. Along with the lab-exercises' data, students' data, regarding their personal information and assessments are stored in the same database system.

The layered networking of the RMCLab system permits each laboratory administrator to present a lab-experiment, running at a specified resource server, according to his personal educational aspects, regardless of the physical location of the real hardware of the lab-experiment.

### *D. Client*

The client-side of RMCLab's system has been designed so as to comply with the demands of a potential user. Thus, client module embeds a specific interface, named as 'scenario interface', for supporting the remote monitor and control of lab-infrastructure, and other full functional and user friendly interfaces for lab-instrumentation (function generator, oscilloscope, etc), as depicted in Fig. 6.

In more details, scenario interface provides a user with graphic information, related with the under study circuit. Additionally, grants to the user the control of circuit parameters (variable pots, caps, etc) and also the monitoring of any active node of the circuit, by selecting and setting the probes of the oscilloscope and the function generator on the circuit. Moreover, extra documentation, regarding the technical and theoretical aspects of the experiment, which can also be separated into multiple steps, can be presented to the user, via the scenario interface.

### *E. Communication module & Instructor-client*

In order to meet the basic requirements of the collaborative interactive e-Learning, a communication interface has been incorporated into the RMCLab system. This communication interface consists of a simple chat module enabling the collaboration and the information exchange, during a remote lab-experiment. The communication interface has been integrated into the RMCLab system by request of its early users, while it is under development the expansion of the chat module with voice and video capabilities.

Finally, RMCLab's platform embeds an identical to the user's client module, named as instructor-client, which offers to a supervisor/instructor of the experiment the ability to replicate, monitor and control any online user's lab-environment. This feature is focused on the educational aspect of RMCLab platform, as it provides an instructor with the ability to closely observe and efficiently tutor the actions of any online users, concluding to a 'near-to-real' lab-environment.

### *F. Architecture overview*

The described hardware architecture is suitable for developing circuits of low-to-medium complexity, at a low-cost. For accessing the properties of this specific hardware, a software driver has been developed and embedded in the resource server application. Apparently, the platform is able to employ and control any hardware, under the condition that the corresponding software driver enables its access. Thus, even the use of complex or commercially available products is possible.

Additionally, the abstract language used to communicate RMCLab entities each other, can be modified according to

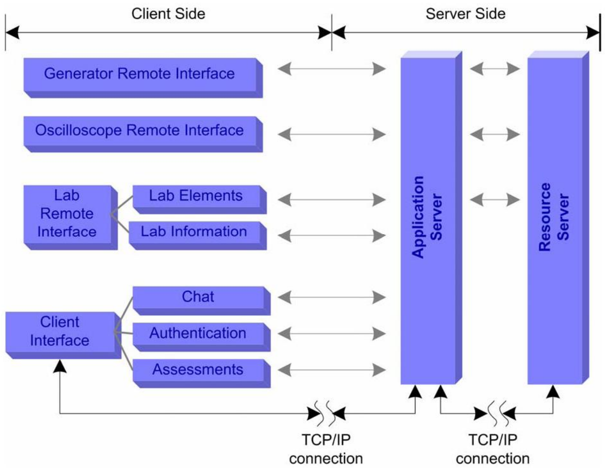



Fig. 7. RMCLab system's dataflow overview.

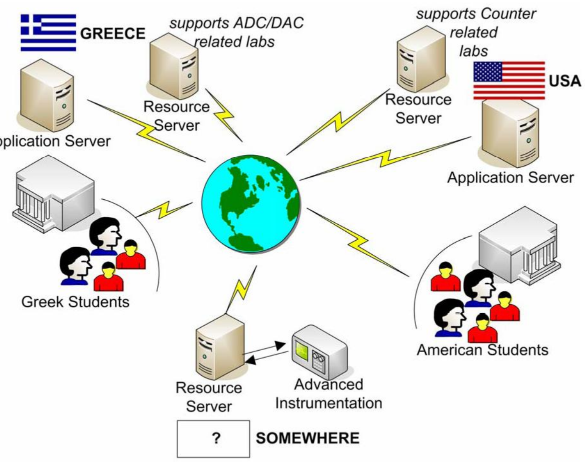



Fig. 8. Advanced RMCLab utilization.

future or different requirements. Fig. 7 depicts the dataflow diagram among the RMCLab entities and focuses on the correspondence between them.

## IV. ADVANCED PROPERTIES OF THE RMCLAB

Conventional lab-education is based on the study of pre-defined lab-experiments. RMCLab system provides an outstanding benefit to its users, thus the feasibility to design and test/measure their own custom circuits under real hardware and real instrumentation. RMCLab users can offline design almost any circuit using separate software package (MAX+plus II or Quartus, both offered at no-cost for academic institutes from Altera). Using one of the aforementioned specific software packages, one can design his own circuits following a reduced set of rules and confirm by simulation its proper operation. Once the design is verified at the client-side, it can be easily uploaded to the server-side and after a while (<15 seconds) he will be able to perform any measurement on his custom design, which is now implemented on real hardware, by employing real instrumentation. The aforementioned procedure is supported by the hardware administration module of the resource server, as depicted in the dotted part of Fig. 4.

The network architecture of RMCLab platform enables the world-wide distribution of resources, in terms of lab-experiments, by utilizing multiple application servers in a single network topology, as depicted in Fig. 8. Thus, instructors all over the world can take the advantages of employing a running lab-experiment and present it in their native language and personal educational point of view. Obviously, each supervisor has the opportunity to review his users' performance by his own criteria, according to the assessments rules for each experiment, that are defined in the RMCLab's application server, which is available and accessed by the supervisor, as resource server transparently executes the measurement requests.

The prospects of the RMCLab system may hopefully expand world-wide, as the above scenario can be further extended if one adds more resource servers, as depicted in Fig. 8. Each resource server can be expert and focused on a specific subject, incorporating the appropriate hardware and instrumentation. Instructors all over the world may take the advantage of using such laboratory resources and develop educational material in their local application servers, in their native language, according to their own educational criteria, so as to offer advanced experimental training to their students, without any requirement for the development and maintenance of any expensive lab-infrastructure.

RMCLab's advanced utilization modes are not limited within the above example. The real-time use of real hardware and real instrumentation can significantly contribute to the educational procedure, since it enables an instructor to prepare 'Active Lessons' and present in details, during a class, the operation of a circuit or a system under real world circumstances, while at the same time can be utilized as a mean of demonstration for expensive products.

## V. THE REALIZED RMCLAB SYSTEM

### *A. Technical characteristics*

The architecture described in section III has been implemented at the University of Patras, Greece. Current configuration is a cost effective implementation, which consists of a single PC, with an Intel Hyper-Threading 2.6GHz processor and 1024MB RAM, embedding both the resource and the application server of our running RMCLab system. This PC, running Windows 2003 Server, is permanently connected to the campus LAN and also to an Agilent 54622D mixed signal oscilloscope and an Agilent 33120A function generator. Oscilloscope is connected with the PC via a high-speed GPIB interface, while function generator is controlled via RS232 @19.2kbps. The PC-interface with the hardware modules is implemented based on a custom, low-cost bus, through LPT in EPP mode. Each card of the hardware infrastructure contains an Altera FPGA of the FLEX8K series and also other components required for the implementation of the experimental circuits.

Table 2. RMCLab time response characteristics.

| Property | Average Delay (sec) |
|---|---|
| Hardware setup and measurement time | 3 |
| Compilation time of a custom circuit | 10 |
| Hardware re-configuration time | 5 |
| Measurement delay from client side using PSTN line @56kbps | <5 |

The aforementioned infrastructure provides fast enough access and response (<3-secs per measurement) to the client requests, as summarized in Table 2.

*B. A simple educational paradigm*

Fig. 9 illustrates the measurement result, regarding the CLK and LOAD signals of an Early Decoding, Count Down, 4-bit Decimal Counter. The top part in Fig. 9 depicts the circuit information available to RMCLab users for the specific circuit, while the middle and bottom part depict the measurement representation of the real oscilloscope and the remote interface, respectively. Obviously, RMCLab system is able to take full advantages of the real hardware and real instrumentation utilization, providing measurements, regarding a wide variety of signals' properties, and the full

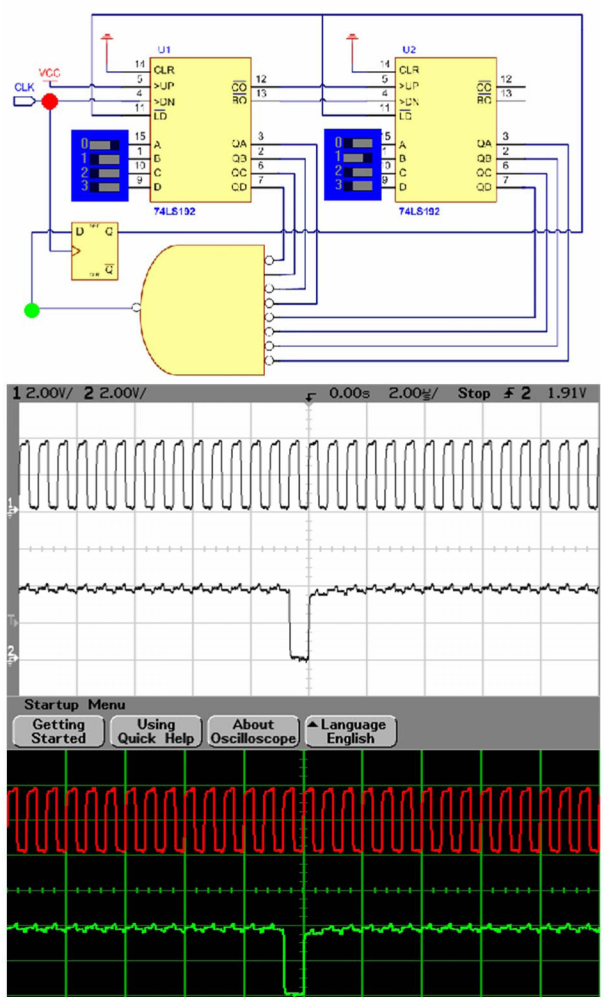



Fig. 9. Educational paradigm of RMCLab usage.

control of both instruments (oscilloscope and function generator).

## VI. CONCLUSION

RMCLab platform is able to provide a wide range of high educational services in a great number of students. It increases the productivity of the students by enabling them to have access to the lab-infrastructure at non working hours, while at the same time affects significantly their psychological mood regarding the level of the offered education by their institute.

Moreover, RMCLab accomplishes its services employing a single PC and a single set of hardware and instrumentation, thus pointing out that is able to provide high-quality lab-education at low-cost, without time consuming human insteraction.

The structure of RMCLab enables sharing of hardware and instrumentation resources, thus makes possible the extensive exploitation of an expensive lab-infrastructure, facilitating the wide spread of remote real lab-experiments, which are indisputably valuable for engineers' education. Additionally, hardware re-configurability permits the remote implementation and measurement of electronic circuits, providing further more a high-valued educational service.

It is anticipated that the proposed architecture guidelines along with the success of the RMCLab platform will motivate educational community to cooperate, so as to develop an integrated World Wide Lab environment.

## REFERENCES


[1] Harry Boldt, "Didaktisches Design medialer Lernangebote im Umfeld ingenieurwissenschaftlicher Laborpraktika: Didactical Design of media learning offers in the surrounding field of engineer-scientific laboratory practical courses", University Duisburg-Essen, Campus Duisburg, Aachen 2005

[2] G. Mason, "A handheld data acquisition system for use in an undergraduate data acquisition course", *IEEE Trans. on Education*, vol. 45, No4, pp. 388-393, November, 2002.

[3] J. Sebastian, D. Garcia & F. Sanchez, "Remote-access education based on image acquisition and processing through the internet", *IEEE Trans. on Education*, vol. 46, No1, pp. 142-148, February, 2003.

[4] P. Lin & M. Lin, "Design and implementation of an internet-based virtual lab system for eLearning support", *Proc. 5th IEEE International Conf. on Advanced Learning Technologies*, Kaohsiung, Taiwan, 2005, pp. 295-296.

[5] A. Bagnasco & A. Scapolla, "A grid of remote laboratory for teaching electronics", *2nd International LeGE-WG Workshop on e-Learning and Grid Technologies: a fundamental challenge for Europe*, Paris, 2003.

[6] J. Palop & M. Teruel, "Virtual work bench for electronic instrumentation teaching", *IEEE Trans. on Education*, vol. 43, No1, pp. 15-18, February, 2000.

[7] C. Daly & J. Horgan, "An automated learning system for java programming", *IEEE Trans. on Education*, vol. 47, No1, pp. 10-17, February, 2004.

[8] Dimitris Karadimas and Kostas Efstathiou, "Design, Implementation and Evaluation of a Remote Laboratory System for Electrical Engineering Courses", Proc. 6th IEEE International Conf. on Advanced Learning Technologies, Kerkrade, The Netherlands, 2006, pp. 408-412.

[9] T. Fjeldly & M. Shur, "Lab on the web: Running real electronics experiments via the internet", Wiley-IEEE Press, 2003.

[10] C. Ko, B. Chen & J. Chen, "Creating web-based laboratories", Advanced Information and Knowledge Processing, Springer, 2004.

# A MOST-Only R-2R Ladder-Based Architecture for High Linearity DACs

Dimitris Karadimas, Michail Papamichail and Kostas Efstathiou
Dept. of Electrical Engineering & Computer Technology, University of Patras, Greece
karadimas@apel.ee.upatras.gr

***Abstract*— The paper presents a MOST-Only, digitally calibrated DAC architecture, based on the R-2R ladder topology. The proposed DAC architecture employs circuitry that enables the fine trimming of each bit's current contribution at the DAC's output, thus concluding in a high linear DAC architecture. The architecture of the proposed DAC is discussed in details along with simulation results that confirm its high linearity performance. The proposed DAC topology maintains the conventional ladder's performance in speed and power dissipation without requiring large area for its implementation.**

## I. Introduction

Nowadays technology has realized the data processing of almost any system in the digital domain. However, in most cases, systems have to exchange data with the real world thus, data conversion from analog-to-digital and vice versa is mandatory. Significant role to the overall performance of a modern system plays the data converters' characteristics, regarding linearity, resolution, speed and power consumption. Especially digital-to-analog converters (DAC) exist in many systems, however their performance cannot be easily optimized and uniquely determined, since designers have to face trade-offs regarding the performance criteria.

Several digital-to-analog conversion techniques have been presented improving the performance of a DAC. However, none of these techniques is able to satisfy all the requirements. For example, ΔΣ techniques [1] or dynamically calibrated current sources [2] have improved characteristics, in terms of resolution and linearity, while their performance lack of speed. Similarly, ladder DACs [3], [4] can offer high speed operation, small area and low power consumption, but they cannot conform to high resolution and high linearity requirements. On the other hand, high conversion speed along with high linearity can be achieved using either thermometer code DACs [5]-[7], at the expense of area and power dissipation, or laser-trimmed [8] R-2R ladder-based DACs, using expensive manufactory procedures. The most common solution in order to compensate the above requirements is to employ hybrid DACs [9], [10] or self calibrated R-2R ladders [11], thus counterbalancing the strict requirements of the digital-to-analog conversion procedure.

A recently proposed technique [12], which is based on a voltage-mode ladder DAC, improves exponentially its performance, regarding the resolution and linearity, without affecting speed, area and power consumption. Based on this technique, new results regarding a current-mode, MOST only, digitally calibrated R-2R ladder architecture, its design specifications and performance characteristics, are presented in this paper.

This work has been motivated by the increased performance, in terms of speed and linearity, required by several applications. The proposed DAC architecture is meant to be employed with a high performance frequency synthesizer, which is based on the Dual Input Phase Accumulator technique [13]. However, the proposed DAC technique can be employed by any application that requires high linearity, along with high speed and low power, data conversion, assuming any practical resolution.

## II. Architecture

Fig.1 depicts the current-mode version of the R-2R ladder proposed by [14]. In a conventional design the voltage difference Vref-Vbias is assumed small enough and thus body effect does not affect significantly the mismatching of the FETs' $R_{ON}$. All FETs are assumed of the same size, while active FETs operate at the linear region.

The architecture proposed by [12] is based on the voltage-mode R-2R ladder, while the current-mode, MOST only version is depicted in Fig.2. In this architecture all FETs related to the R-2R topology are denoted as Q and are of the same size, while they operate at the linear region and have an $R_{ON}$ resistance.

As it can be seen in Fig.2, each switching pair of FETs has been replaced by an NC-bits CMOS current-mode ladder, identical to the one depicted in Fig.1, thus introducing a calibration branch. This replacement does not affect the R-2R topology, since the input resistance of each calibration branch is equivalent with one's FET $R_{ON}$. The modified topology enables the trimming of the output current of each calibration branch, according to the specific calibration word. Therefore, it is able to trim each branch's current close-enough to the optimum one. Apparently, the resolution of each calibration branch (NC) plays

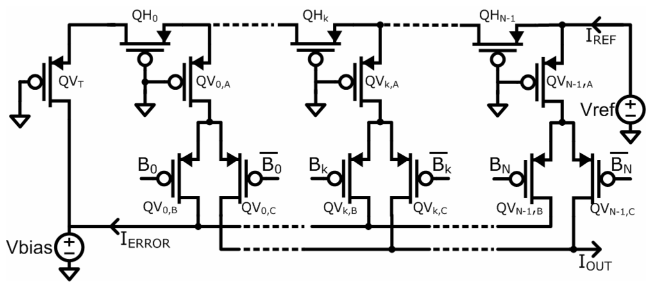


Figure 1. Conventional CMOS-based R-2R ladder DAC.


1-4244-2420-7/08/$20.00 

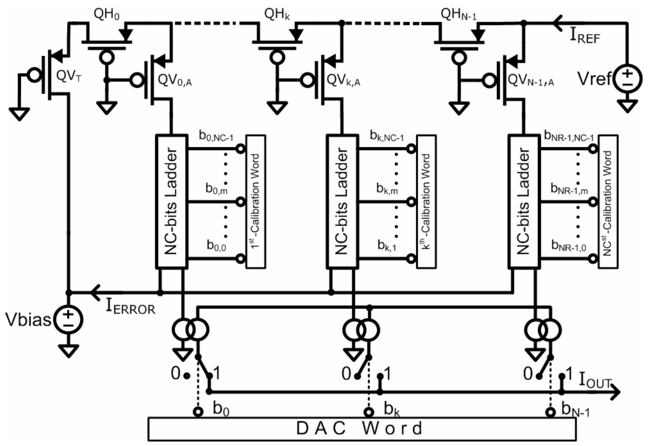



Figure 2. The proposed architecture of the current-mode DAC.

a major role, since it assigns the accuracy of the trimming procedure. Finally, the output currents of each calibration branch are mirrored and summarized at the output of the DAC ($I_{OUT}$).

Fig.3 indicates that the proposed DAC circuitry consists of a number of identical blocks, one for each DAC bit. Each one of these blocks consists of the corresponding bit's calibration ladder as well as the required circuitry, for current-cascading and the assignment of the calibration word. Digital signals Din, Dout, SCLK and LOAD are responsible for the assignment of the calibration word for each bit. Analog signals Cin and Cout are employed for the current-cascading, while $CI_{OUT}$ is the output current of each block. Finally, the digital signal $b_k$ represents the corresponding bit of the DAC's digital input word.

Each block of Fig.3, depicted in more details in Fig.4, consists of a single element of the DAC backbone ($QH_k$ and $QV_{k,A}$), the calibration branch, composed by NC-similar blocks, one for each calibration bit, and a current mirror (M1 to M6), that mirrors the proper amount of current to the output node, while the error current is let to the ground (through E1 to E3). E1 to E3 and M1 to M6 are of the same size and should be large enough, in comparison with FETs denoted as Q, since they have to operate at the saturation region and preserve the minimum voltage drop across.

Each block, representing a calibration bit in Fig.4, consists of the calibration register and a single bit calibration ladder, as depicted in Fig.5. Calibration register consists of two D-FF's and enables the serial-input/parallel-load of the DAC's calibration words. The single bit calibration ladder is identical to the cell of the ladder depicted in Fig.1.

As soon as the calibration words have been assigned, all FETs, denoted as Q, operate at a steady-state, regardless of the

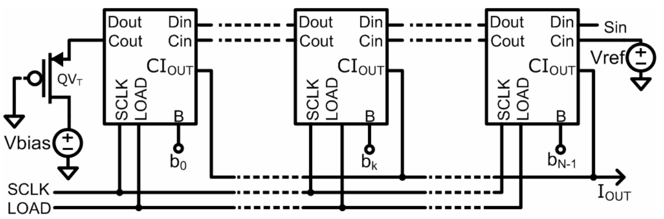



Figure 3. Block diagram of the proposed DAC circuitry.

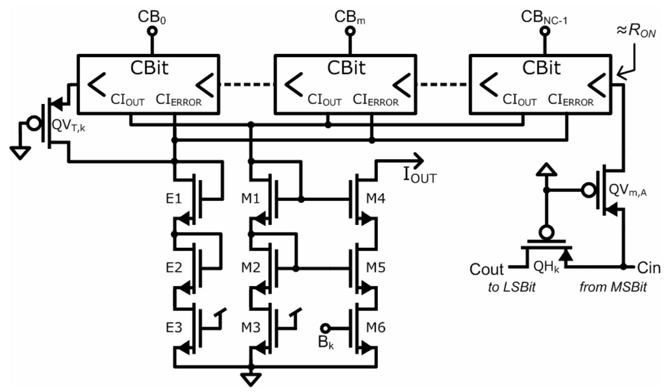



Figure 4. Internal architecture of a ladder's block.

DAC's digital input. Only FETs M4 to M6, of each calibration branch, alternate their operating conditions according to the DAC word. This implies that most of the circuitry operates in constant conditions, as only a small portion of it has to switch; thus contributing to high speed performance. Since current, in most FETs, is constant, the power consumption of the circuit will be approximately equal to Vref×$I_{REF}$. Depending on the output current of the DAC, extra power consumption, dissipated on FETs M4 to M6, may be accounted.

## III. Calibration Technique and Simulation Results

An 8-bit DAC, with 8-bit calibration branches, has been simulated using the AMS08 technology and the TANNER tool. During the simulation, no mismatch error in the FETs and ideal connections between them have been assumed.

Fig.6 depicts the simulated INL of the proposed DAC, when the word 255 (0xFF) is employed as calibration word, for each branch. Despite the fact that no error sources (FET mismatch, wiring resistance) have been taken into consideration, the result can not be satisfactory (INL=2.4LSBits), since body effect of FET's, denoted as Q, introduces mismatches on their $R_{ON}$.

Apparently, the selection of the proper set of calibration words may lead to the improvement of DAC's linearity. The criterion for selecting the proper set of calibration words is based on the optimization of DAC's Signal-to-Noise Ratio (SNR), assigned as

$$\mathrm{SNR} = 10\log_{10}\left[\left(\frac{2^N - 1}{2\sqrt{2}}\right)^2 \Big/ \mathrm{Noise}_{RMS}^2\right], \qquad (1)$$

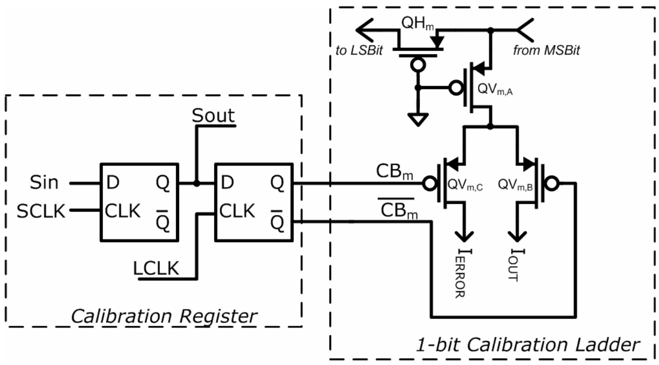



Figure 5. Internal architecture of a single bit of a calibration ladder.

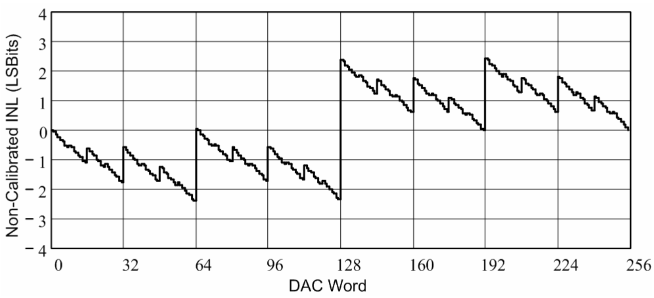



Figure 6. Simulated INL of the non-calibrated 8-bit ladder.

where $\text{Noise}_{RMS}^2$ is the RMS power of the quantization noise in $\text{LSBits}^2$ along with the noise introduced by DAC's non-linear effects, as shown in (2),

$$\text{Noise}_{RMS}^2 = \frac{1}{2^N - 1}\sum_{k=0}^{2^N-1}\int_{-1/2}^{1/2}\left(x + \text{INL}_k\right)^2 dx,\ \text{LSBits}^2. \qquad (2)$$

Eq. (2) denotes that the Noise RMS power can be directly connected with the INL produced for each DAC's word, as

$$\text{Noise}_{RMS}^2 = \frac{1}{12} + \frac{1}{2^N - 1}\sum_{k=0}^{2^N-1}\text{INL}_k^2,\ \text{LSBits}^2, \qquad (3)$$

where

$$\text{INL}_k = \left(2^N - 1\right)\frac{\left[I_{OUT(k)} - I_{OUT(0)}\right]}{I_{OUT\left(2^N-1\right)} - I_{OUT(0)}} - k,\ \text{LSBits}, \qquad (4)$$

assuming that $0 \le k \le 2^N - 1$. Combining (1)-(4) SNR can be simplified as follows,

$$\text{SNR} = 10\log_{10}\left[\frac{\left(\frac{2^N - 1}{2\sqrt{2}}\right)^2}{\frac{1}{12} + \frac{1}{2^N - 1}\sum_{k=0}^{2^N-1}\text{INL}_k^2}\right]. \qquad (5)$$

Eq. (5) denotes that in order to maximize SNR, the INL's

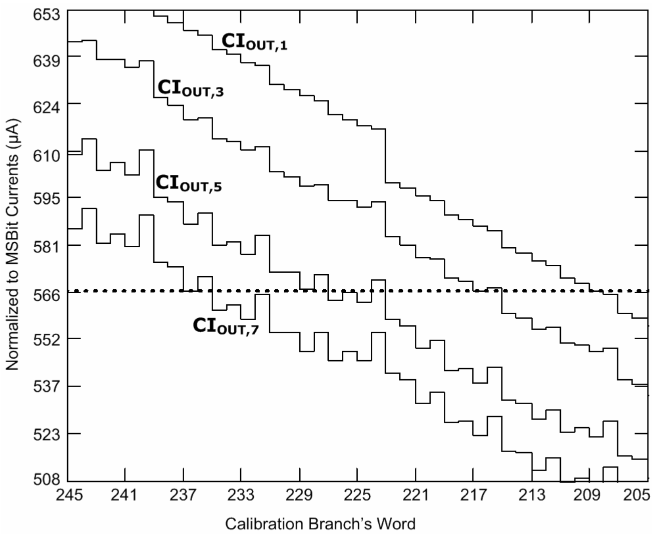



Figure 7. Normalized output currents of the calibration branches.

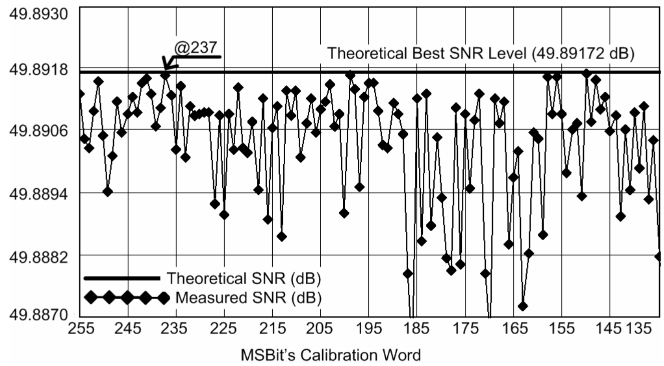



Figure 8. DAC's SNR as a function of the MSBit's calibration word.

noise power should be minimized, thus the least square algorithm should be applied for the corresponding INL, for each possible set of calibration words. Applying the above calibration algorithm, the output current of each calibration ladder has been measured. Fig.7 depicts, for means of clarity, only four out of eight, normalized to the MSBit, DAC bit currents, for a small range of calibration words (245-205).

Processing on these data has proved that, for each calibration word of the MSBit a specific SNR can be achieved, as depicted in Fig.8. TABLE I lists the 10-best SNR values achieved and the corresponding sets of calibration words. The selection of the proper set of calibration words can be easily performed, if the maximum current requirements are taken into account. After the calibration, the impact on the SNR of this selection is obviously negligible, as depicted in Fig.8. Therefore, the second row of TABLE I has been selected, while this selection is represented in Fig.7 as a dotted line. The derived INL, after applying the above calibration procedure, is depicted in Fig. 9.

Apparently, the INL of the calibrated DAC is significantly improved in comparison with the one depicted in Fig.6, as the maximum INL value has become less than 0.045 LSBits, while the corresponding SNR (49.891675 dB) can be considered as equal to the theoretical value (49.891716 dB). The goal of the presented architecture and its calibration procedure is that the optimization of the SNR is based on the DAC bits' currents trimming, according to calibration resolution (NC), regardless of FETs' $R_{ON}$ mismatching; thus there always exist a 'most-proper' set of calibration words that leads to high linearity.

## IV. Speed, Area and Power Characteristics

Speed performance of the proposed architecture can be considered identical to the one of the conventional current-mode R-2R ladder, as the current switching circuit, depicted in Fig.4,

TABLE I. Calibration Words for 10-Best SNR Values.

| SNR (dB) | Calibration Word | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Bit7 | Bit6 | Bit5 | Bit4 | Bit3 | Bit2 | Bit1 | Bit0 |
| 49.891716 | *Theoretical best SNR level* | | | | | | | |
| 49.891686 | 150 | 148 | 147 | 144 | 143 | 139 | 139 | 143 |
| 49.891675 | **237** | **233** | **226** | **222** | **217** | **210** | **209** | **213** |
| 49.891668 | 199 | 194 | 191 | 189 | 184 | 181 | 180 | 188 |
| 49.891645 | 156 | 155 | 153 | 151 | 149 | 144 | 144 | 149 |
| 49.891627 | 158 | 156 | 155 | 152 | 151 | 146 | 147 | 151 |
| 49.891619 | 241 | 236 | 235 | 226 | 222 | 216 | 215 | 218 |
| 49.891561 | 148 | 147 | 145 | 143 | 141 | 138 | 136 | 142 |
| 49.891549 | 251 | 247 | 243 | 236 | 231 | 224 | 223 | 227 |
| 49.891528 | 194 | 191 | 188 | 186 | 183 | 178 | 177 | 185 |
| 49.891520 | 195 | 190 | 189 | 184 | 182 | 177 | 176 | 184 |

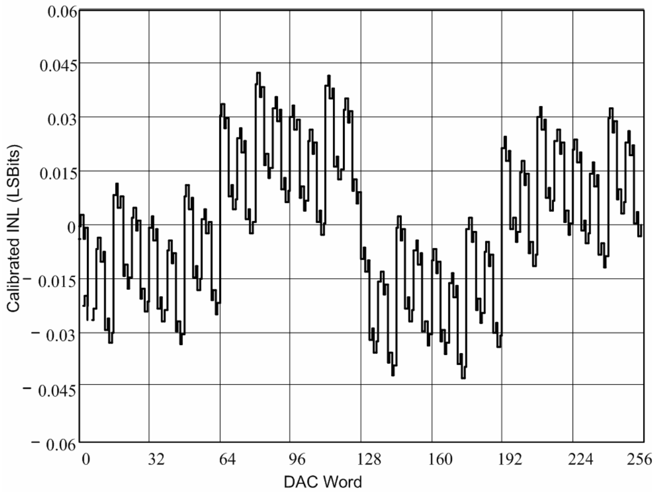



Figure 9. INL of the calibrated DAC.

employs a current cascode topology (FETs M4 to M6). Additionally, it should be emphasized that during the DAC operation all devices are in a steady state except the aforementioned FETs, which assign the output current.

The total device count of the proposed architecture can be compared with the one of advanced DAC architectures. Fig.10 compares the minimum device count required by a thermometer code or R-string DAC with the corresponding device count of the proposed architecture, assuming that NC=N.

The power consumption of the proposed architecture can be considered as at most double of the conventional current-mode R-2R ladder DAC. Taking into account Fig.1 and Fig.2 it can be found that the power consumed by a conventional ladder equals to $Vref \times I_{REF}$ while the proposed architecture requires this power for the biasing and additionally sinks the output current, thus increasing the total power consumption by $Vref \times I_{OUT}$. However, the maximum power consumption cannot exceed the $2 \times Vref \times I_{REF}$ value. Finally, it should be noticed that the extra digital logic, depicted in Fig.5 as calibration register, does not switch during the operation of the DAC, since it stores constant calibration words and thus does not contribute to the power dissipation.

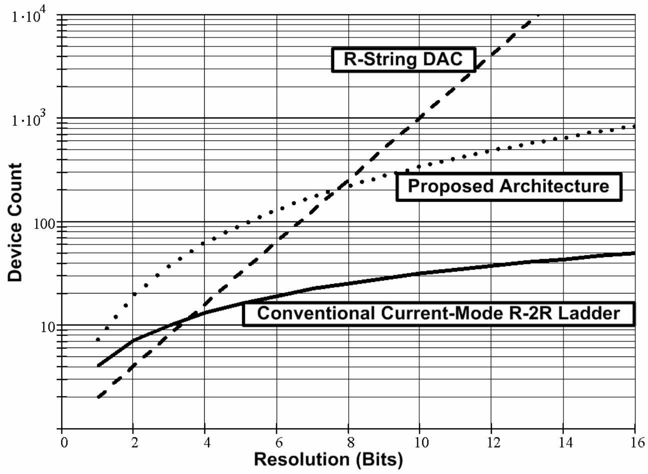



Figure 10. Device count comparison for 3-DAC architectures.

## CONCLUSION

A current-mode, MOST-only, R-2R ladder-based, digitally calibrated DAC architecture has been presented in this paper. The architecture and circuitry of the DAC has been discussed in details along with a calibration algorithm, which is focused on obtaining the best SNR. The current trimming capabilities, derived by the proposed architecture, assign an inherent linear ladder-based DAC, as impacts of FETs' mismatching or any other error source's non-idealities are diminished.

Simulation results have certified the capabilities, of the proposed DAC architecture, in achieving high linearity. The proposed architecture minimizes the current and voltage switching, thus implying high speed along with low power characteristics. The implementation of the proposed DAC was based on standard CMOS technology and requires no special treatment for analog components, concluding in a low cost design.

## ACKNOWLEDGMENT

This work was supported by the "Karatheodoris" research grant, awarded by the Research Committee of the University of Patras, Greece.

## REFERENCES

[1] T. Shui, R. Schreier and F. Hudson, "Mismatch shaping for a current-mode multibit delta-sigma DAC," IEEE J. Solid State Circuits, vol. 34, Mar. 1999, pp. 331-338

[2] M. P. Tiilikainen, "A 14-bit 1.8-V 20-mW 1-mm2 CMOS DAC," IEEE J. Solid State Circuits, vol. 36, Jul. 2001, pp. 1144-1147.

[3] M. P. Kennedy, "On the robustness of R-2R ladder DAC's," IEEE Trans. Circuits & Syst. II, vol. 47, Feb. 2000, pp. 109-116.

[4] L. Wang, Y. Fukatsu, and Kenzo Watanabe, "Characterization of Current-Mode CMOS R-2R Ladder Digital-to-Analog Converters," IEEE Trans. Instr. and Meas., vol. 50, Dec. 2001, pp.1781-1786.

[5] C. Lin and K. Bult "A 10-b, 500-msample/s CMOS DAC in 0.6 $mm^2$," IEEE J. Solid State Circuits, vol. 33, Dec. 1998.

[6] J. Bastos, A. M. Marques, M. Steyaert and W. Sansen, "A 12-bit intrinsic accuracy high-speed CMOS DAC," IEEE J. Solid State Circuits, vol. 33, Dec. 1998.

[7] M. Vadipour, "Gradient error cancellation and quadratic error reduction in unary and binary D/A converters," IEEE Trans. Circuits & Syst. II, vol. 50, Dec. 2003, pp. 1002-1007.

[8] "R/2R ladder networks," Application Note AFD006, International Resistive Company, Inc., 1998.

[9] Y. Cong and R. L. Geiger, "A 1.5 V 14-bit 100ms/s self-calibrated DAC," IEEE J. Solid State Circuits, vol. 38, Dec. 2003, pp. 2051-2060.

[10] A. R. Bugeja and B. Song, "A self-trimming 14-b 100-MS/s CMOS DAC," IEEE J. Solid State Circuits, vol. 35, Dec. 2000.

[11] B. Vargha, I. Zoltan, "Calibration algorithm for current-output R-2R ladders," IEEE Trans. Instr. and Meas., vol. 50, Oct. 2001, pp. 1216-1220.

[12] D. S. Karadimas and K.A. Efstathiou, "On The Study of a Digitally Calibrated R-2R Ladder Architecture", IEEE International Symposium on Signals, Circuits and Systems, vol. 1, pp. 1335-1338, Iasi, Romania, July 2005.

[13] K. A. Efstathiou, and G. D. Papadopoulos, "An Enhanced Frequency Synthesizer using an Analog Dual Input Phase Accumulator", Proceedings of the Sixth IEEE International Conference on Electronics, Circuits and Systems, vol. 1, pp.17-20, Paphos, Cyprus, Sept. 1999.

[14] K. Bult and J. Geelen, "An inherently linear and compact MOST-only current division technique," IEEE J. Solid State Circuits, vol. 27, Dec. 1992, pp. 1730-1735.

# Web Services based Collaborative B2B e-commerce in the Oil Industry

J. Gialelis[1], D. Karadimas[1], P. Chondros[1], A. Kalogeras[2], S. Koubias[1] D.Serpanos[2]

[1]Applied Electronics Laboratory
Electrical and Computer Engineering Dept.
University of Patras Campus
26500 Rio Patras, Greece
{gialelis, koubias}@ece.upatras.gr
{karadimas, pchondros}@apel.ee.upatras.gr

[2]Industrial Systems Institute
Science Park of Patras Building
Stadiou Str, 26506 Platani Patras, Greece
{kalogeras, serpanos}@isi.gr

**Abstract.** Collaborative B2B e-commerce provides enterprises with the necessary level of flexibility and efficiency to retain competitiveness under the increasingly turbulent business environment. XML-based frameworks may be utilized by enterprises to integrate enterprise level applications to provide B2B e-commerce without, though seamlessly integrating inter-enterprise processes with intra-enterprise processes to provide collaborative B2B e-commerce. This paper presents an innovative approach towards collaborative B2B e-commerce by utilizing the RosettaNet emerging standard in combination with Service Oriented Architecture and semantically enriched information in order to seamlessly integrate the inter-enterprise (public) and intra-enterprise (private) processes. An implementation of a collaborative B2B e-commerce business model in the oil industry is given as a use case example.

**Keywords:** Collaborative B2B e-commerce, Service Oriented Architecture, RosettaNet semantics.

## 1. Introduction

Economic pressure and turbulences in the global business environment make it increasingly more difficult for manufacturing enterprises to make predictable long-term provisions, as they radically change the way that organizations operate and interoperate with other enterprises in order to accomplish a business goal. In this context, electronic Business-to-Business interactions (B2B e-commerce) provide the necessary backbone for organizations to interact with their suppliers and customers in order to respond more quickly to changes. As a result, business competition is gradually becoming more value chain to value chain rather than enterprise-to-enterprise.

E-Business frameworks represent the state of the art in e-commerce; since framework-based B2B e-commerce effectively addresses the seamless linking of inter-enterprise (public) processes involving Enterprise layer's applications to those of business partners. These frameworks provide the necessary layers to achieve interoperability in e-commerce, since their interactions involve various business processes and business components, such as executable applications, systems, and their associated information [1].

In order to support the interchange of information among "public" modules, several standards and languages have been developed. EDI (Electronic Data Interchange), EDIINT (EDI over the Internet), ebXML (e-business eXtended Markup Language) and BizTalk are some generic standards addressing vertical as well as horizontal integration issues. Furthermore, standards such as RosettaNet, IOTP (Internet Open Trading Protocol), ICE (Information and Content Exchange) and OBI (Open Buying on the Internet) deal more with vertical market-oriented integration issues.

However, current framework-based developments consider inter-enterprise processes, involving only Enterprise layer applications, information, etc; without taking into consideration the intra-enterprise processes involving applications, information residing into the other enterprise layers such as Plant and Shop Floor layers thus, not assuring agility, flexibility and autonomy of the interacting partners [2]. The work presented in this paper, intents to overcome this drawback by proposing a set of technologies imposed over the components of a presented e-business frameworks' holistic architecture which extends the functionality of current frameworks in order to incorporate all enterprise layers into B2B e-commerce interactions thus elevating an innovative collaborative business model which seamlessly integrates the inter-enterprise (public) with the intra-enterprise (private) processes [3]. More specifically, section 2 describes the collaborative B2B e-commerce, section 3 depicts the structure of e-business frameworks and describes the proposed set of technologies imposed over the e-business frameworks' holistic architecture and section 4 presents our implementation of a collaborative B2B e-commerce business model in the oil industry according to the proposed approach.

## 2. Collaborative B2B e-commerce

Collaborative B2B e-commerce differs from basic B2B commerce, since collaborative commerce goes beyond on-line document exchanges, indicating that organizations adopt B2B networks to establish new collaboration mechanisms with channel partners. Thus, enterprises should establish collaborative B2B processes with cooperating manufacturing enterprises, in order to best exploit and gain the most out of the adoption of electronic commerce networks. Collaborative B2B processes are considered to be any

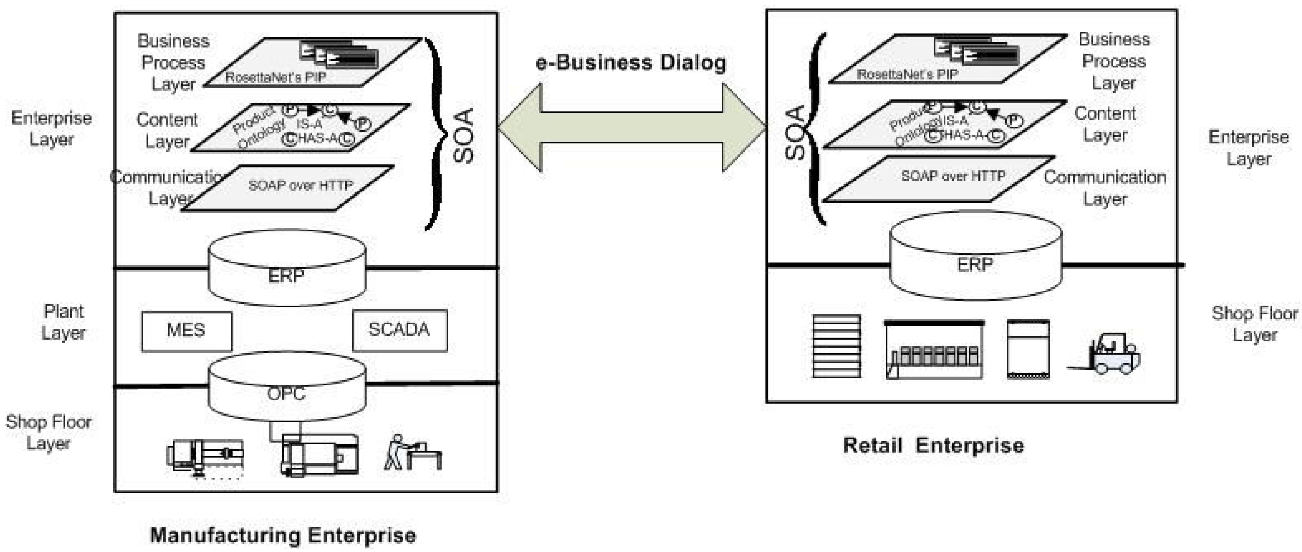



Fig. 1. Proposed e-business frameworks' holistic architecture.

end-to-end business processes, which results from the seamless integration of intra-enterprise processes with inter-enterprise processes.

A designated example of a collaborative B2B process is the Continuous Replenishment Planning (CRP) model on top of which one of the most efficient inventory management functions may be developed. The key characteristics of CRP are the sharing of real-time inventory data by retailers with manufacturers and continuous replenishment of retailers' inventory by manufacturers. This function, called continuous replenishment planning, is defined as the practice of partnering between distribution channel members that changes the traditional replenishment process from a warehouse/distributor centric process handling purchase orders, to an end-user/retailer centric replenishment process involving the manufacturing enterprise in the overall process utilizing actual and forecast data.

Even though, the CRP collaboration model has the potential to alter significantly production-scheduling methods and inventory management practices so that enterprises enjoy the above-mentioned benefits, survey results do not keep up with these expectations [4]. This drawback is caused, mainly, due to CRP deployment methodologies. Such methodologies are based on point-to-point architectures and on the utilization of proprietary solutions/technologies which present very limited tolerance to system/applications alterations and minimum scalability to value-network expansion.

In order to overcome these drawbacks and achieve both integration of heterogeneous systems/applications, at any layer in the enterprises they may reside in and unlimited expansion of the value-network, this paper proposes the deployment of specific set of technologies over both ERP (Enterprise Resource Planning) and OPC (Manufacturing Execution Systems and Supervisory Control And Data Acquisition) systems. These technologies are based on Internet ubiquity and emerging standards. Therefore, Service oriented Architecture (SoA) elements, RosettaNet's ontology and standard processes for the B2B transactions are utilized, in order to fully implement the CRP collaboration model. Moreover, SOAP is utilized as the transport message protocol in order to satisfy communication needs, while ontologies are used in order to standardize content semantics and WSDL (Web Service Description Language) is used in order to expose systems' functionality as Web Services. Figure 1 depicts the e-business frameworks' holistic architecture, based on the above mentioned principles which are further described in the next section.

## 3. e-Business Framework Holistic Architecture and Components Technologies

Several contemporary B2B interaction frameworks are based on the XML notation. They aim at overcoming some of the limitations of traditional electronic data interchange standards. XML-based frameworks for B2B transactions follow a hierarchical model, which comprises three layers: the Communication, the Content and the Business Process layer [5]. The Communication layer provides the appropriate means of communication for exchanging the required messages among the involved systems/applications, while the Content layer provides the necessary tools (languages and conceptual models) so that the involved entities agree upon the semantics of contents and the types of the exchanged documents. Finally, the Business Process layer is concerned with the conversational interactions among services. The following three sub-sections present the proposed imposed technology solutions for the three aforementioned layers.

### 3.1 Communication Layer

For the communication needs, the SOA paradigm is utilized by employing the use of Web Services (WSs) technology in order to open up enterprise systems, plant systems and production systems; thus making available their functionalities. Therefore, the Simple Object Access Protocol (SOAP) and the Web Services Description Language (WSDL) are deployed at the communication layer. Universal Description Discovery & Integration (UDDI) is not integrated in our used case implementation.

### 3.2 Content Layer

The content layer provides the suitable means to describe and organize exchanged information in such a way that it can be understood and used by all the collaborating partners. Content interactions require that the involved systems understand document content semantic. Our work is based upon the development of an elaborated petroleum product ontology in order to achieve a uniform product terminology. Furthermore, existing work in different standards, such as RosettaNet Technical Dictionary (RNTD), is utilized wherever required.

### 3.3 Business Process Layer

Our work emphasizes on the RosettaNet framework in order to address business process layer needs. RosettaNet focuses on three key areas of standardization to automate B2B interactions. First of all, the vocabulary needs to be aligned. Secondly, the business process governing the interchange of the Business Messages themselves must be harmonized and specified. Finally, the way in which Business Messages are wrapped and transported should also be specified.

Next section describes in details an implementation example of a use case scenario of the proposed approach depicting a real world Continuous Replenishment Planning process application in the oil industry which involves gas stations or other enterprises as end users/retailers and a refinery as the manufacturing enterprise.

## 4. An illustrative Use Case

A use case scenario is presented in order to show a generic infrastructure of an online Oil-gas ordering Content Management System (OilCMS). The implementation presented in this section focuses on the specification and deployment of intra-enterprise (private) processes as well as the integration with the inter-enterprise (public) processes. However, there are some key characteristics of such an online system: (a) should be able to connect to any other available ERP system of the enterprise in order to update product and pricing information and (b) should be able to support enterprises with different products and pricing policies.

### 4.1 Scenario Description

Oil ordering and delivering is a demanding procedure, in terms of catching-up with the deadline dates, managing the paperwork and generally taking care of all the rest non-automated processes inside the Distributor Enterprise. Three entities are defined as follows: (a) the customer/user which makes the order(s), (b) the Distributor Enterprise which collects the orders and has the duty to deliver the order to the customers, accordingly and (c) the Supplier Enterprise which supplies the distributor with the appropriate quantities of the ordered products. The OilCMS automates this end-to-end online ordering infrastructure based on intra/inter-enterprise Web Services. Moreover, it is able to keep track and publish the different order states during the whole process.

An operational diagram of the implemented system is depicted in Fig. 2. The distributor enterprise exposes several Web Services identifiable by the UDDI protocol. These Web Services are implemented by the OilCMS software and are installed locally at the Distributor Enterprise's server. Figure 2 shows the three basic web services:

- `GetQuote()` – The customer can access information about available products, pricing and discounts policies of the Distributor Enterprise through this Web Service. The product catalogue is updated according to the enterprise's policy.
- `Order()` – The order Web Service allows the user to make an online order. The order is stored at the Distributor Enterprise local database and can be managed by the administrator through the OilCMS software.
- `TrackStatus()` – The different states of the order throughout the entire processing path are published through this Web Service.

Major role in the above process plays the intra-enterprise procedure which is implemented by the OilCMS software. Each OilCMS of a distributor company is connected to the local database of the corresponding enterprise. This database stores data relative to:

- User data – contains personal and technical information (i.e. a customer may receive the products by sheep, by land or both) about each user account.
- Product data – Contains information about available product types and costs, respectively.
- Order data – All the data associated with an order are stored in this structure. Some of these data could be order delivery date, customer's address, total cost, cost per product, information about pricing policies per user and discounts policies, etc.

All aforementioned data structures can be accessed by the OilCMS administrators. The OilCMS software is designed to support five different types of users that have special rights, defined as follows:

- **Unconfirmed Users** – An unconfirmed user has already made a registration request but has not yet been confirmed by the OilCMS administrators. Only when confirmed, he will be able to access the system.
- **Confirmed Users** – A confirmed user is able to access product and pricing information, place an

order and check his order status. Moreover, he can edit his account data in case of changes. The customer is represented by this type of user.

- **Banned Users** – A confirmed user can be banned by the administrators for several reasons (did not paid on time, etc.). A banned user cannot access the enterprise's system.
- **Administrators** – An administrator can manage users' accounts, product types and prices, update catalogs and ban or confirm potential customers/users. An administrator has also the duty to call the inter-enterprise service by uploading the orders to the supplier enterprise.
- **Super-Administrator** – The main difference between an administrator and the super-administrator, is that the latter has the privileges to manage administrators' accounts. The super-administrator is assumed unique in the proposed implementation.

### 4.2 The OilCMS Architecture Implementation

The OilCMS architecture has been designed according to object-oriented approach principles and implemented using PHP programming. Along with the basic classes, which are depicted in Fig. 3, come two auxiliary classes and several library functions. A configuration script is also provided and allows the

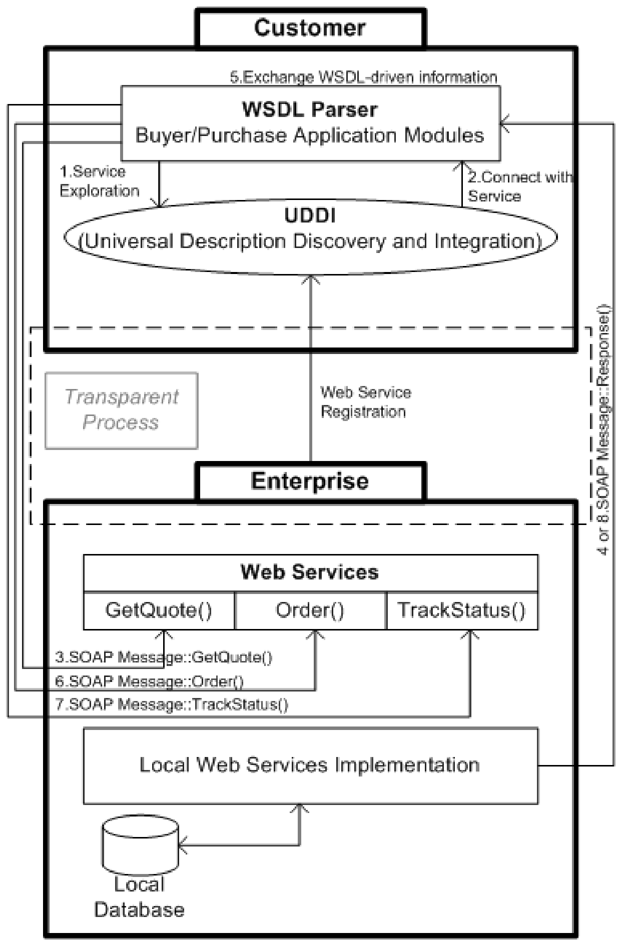



Fig. 2. Operational diagram of the online order management system based on web services.

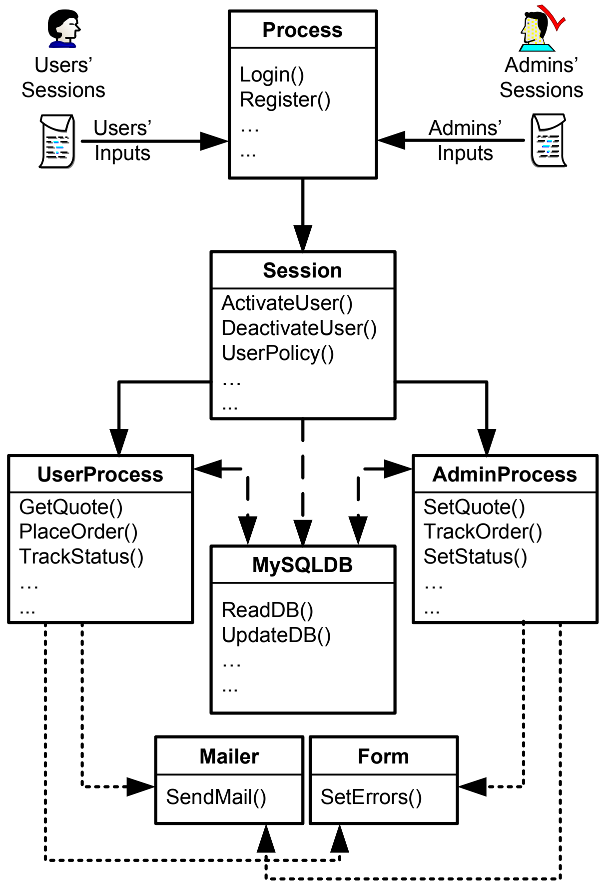



Fig. 3. Class Collaboration Diagram within the Distributor Enterprise Web Service (OilCMS).

installation of the software to be platform and technology independent. The system can support many users and administrators at the same time. Following there is a short presentation of the implemented classes, which are illustrated in Fig. 3.

- **Process** – manages the actions performed by a potential user.
- **Session** – keeps track of the OilCMS active users. Session is active until the user logs out from the system. It stores any information related to the online user such us user level, time of being active, last time the user entered the system, etc. One session object is generated each time a user enters the system.
- **UserProcess** – implements the navigation to the internal procedures of the OilCMS for confirmed users.
- **AdminProcess** – implements the navigation to the internal procedures of the OilCMS for the administrators and the super-administrator.
- **MySQLDB** – implements all methods employed to access the Distributor Enterprise local database.
- **Mailer** – auxiliary class used to send informative e-mails. A mail server needs to be properly configured on the Distributor Enterprise's side.
- **Form** – auxiliary class that manages all the errors caused by mistake on the user interface.

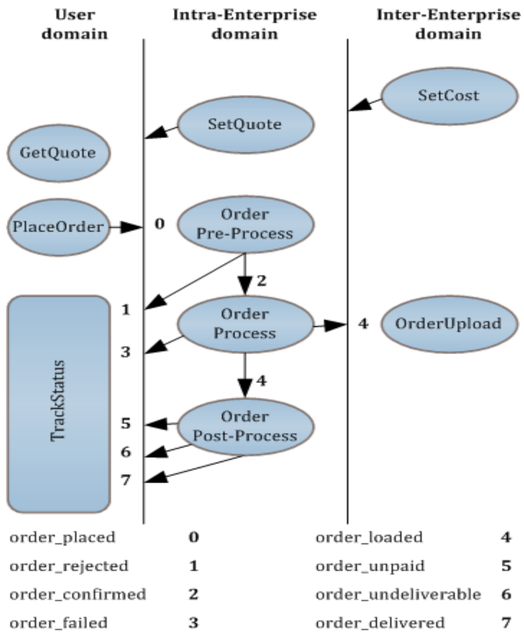



Fig. 4. Order state diagram versus implementation domains.

Many different security aspects have been taken into account so that the final architecture will be robust and tight enough in order to meet all the requirements from security point of view. In particular, every user is uniquely combined with his username and password, while the last one is strongly encrypted by employing hash functions. A second security layer is provided by employing the aforementioned logic of five different levels of users. The system is able to impart the appropriate privileges of a user level to each designated user, as all basic principles of object-oriented design have been followed. Finally, Session object is responsible for the security implementation for the entire Web Service as it grants users with the appropriate privileges. Figure 4 depicts the transactions that take place between the user, the Distributor Enterprise and the Manufacturing Enterprise (Supplier). Each of the three order sub-processes sets accordingly the order status, based on inputs provided by any of the administrators.

The order status can be accessed by the user through the `TrackStatus()` Web Service. The designed algorithm provides a real-time representation of the ordering status throughout the whole processing path.

### 4.3 Service Expandability

As already mentioned, the proposed replenishment planning process, aims to automate the intra-enterprise procedures by utilizing a Web Service. The basic components of this process include the Web Service implementation itself, the administration and management policy, the configuration and a local database, as depicted in one of the two cells at the left part of Fig. 5.

The standard scenario provides a potential buyer with the capability to employ the service offered by the Distributor Enterprise, while this service is able to transparently communicate with the corresponding application of the Supplier Enterprise. The Supplier Enterprise's application could be also implemented as a

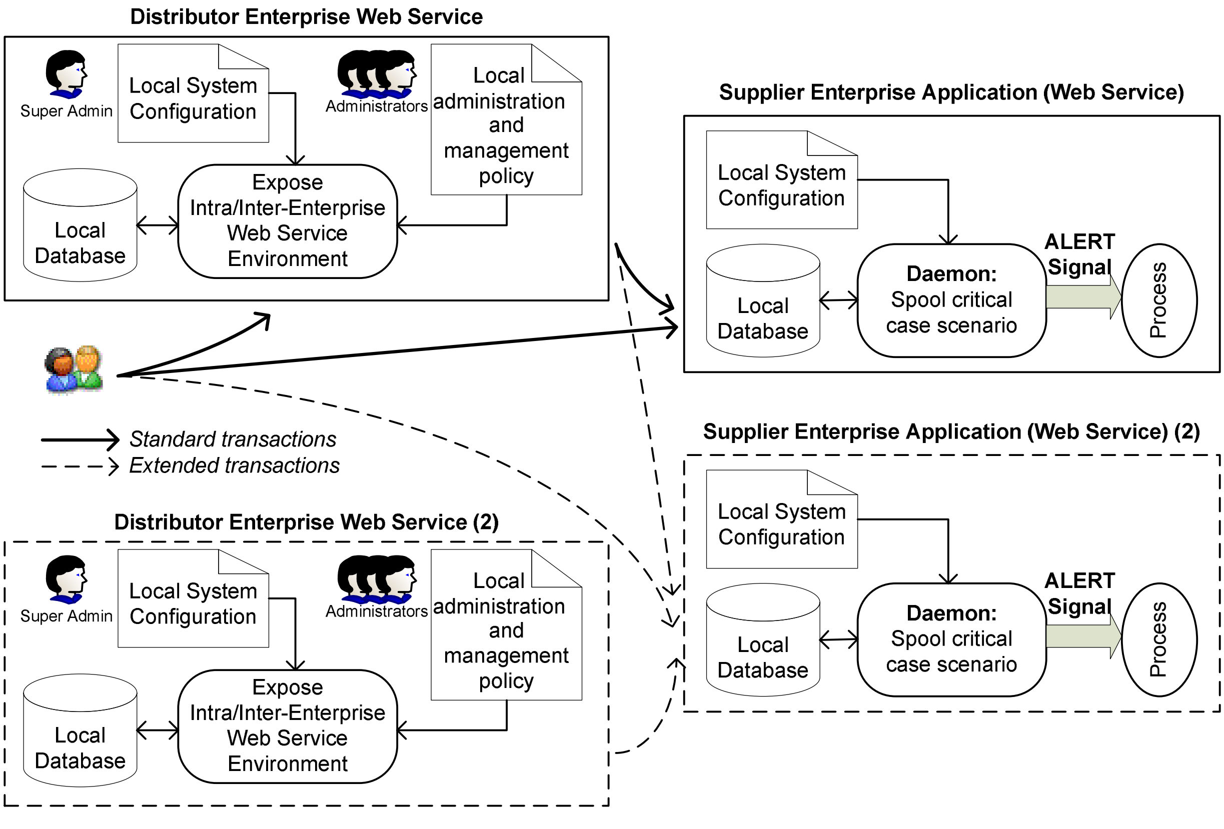



Fig. 5. Proposed Service Expansion Scenario.

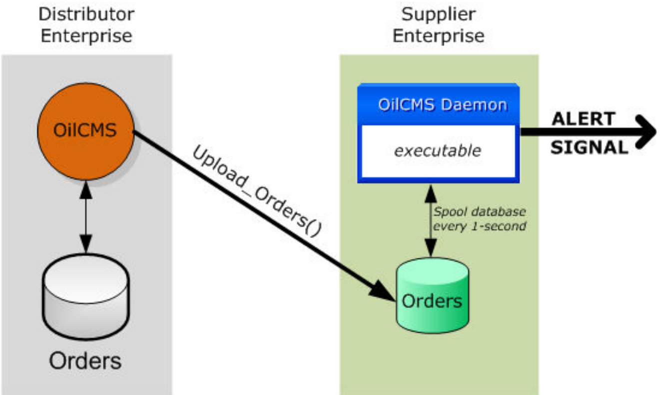


Fig. 6. OilCMS daemon process in details.

separate Web Service, which would be able to be exploited by other Distributor Enterprises and/or individual buyers, if they are granted with this privilege. Distributor Enterprise's transactions with the Supplier Enterprise's application or Web Service are secure and can be operated only by the administrators group of the last mentioned Enterprise. Transactions from the Distributor Enterprises Web Services are uploaded and finally stored in a local database next to the Supplier Enterprise application or Web Service.

The overall system architecture is able to provide many extra benefits, apart from the intra-enterprise automation of the ordering procedures within the Distributor Enterprise. Thus, the proposed implementation is suitable for also contributing to the automation of the production procedures within the Supplier Enterprise. The local database of the Supplier Enterprise could be implemented either as an extension of the possible existing ERP of the company or as a standalone database. In the first case, the use of an intermediate program, acting as a wrapper, in order to enable communication and data exchanging between the two databases is mandatory, while in the second case the local database of the Supplier remains very similar to the one of the Distributor Enterprise. Thus, the proposed replenishment and planning process does not repeal the utilization of any existing ERP within the Distributor or the Supplier Enterprise.

Moreover, in both cases, a special standalone process, could be deployed and utilized in order to take the advantages of the aforementioned local database of the Supplier Enterprise. Figure 6 depicts the idea of how this process is utilized, called oilcmsdaemon, while the main key functions of the daemon have been presented in Fig. 5. The major role of this daemon process is to spool the database at certain time intervals and make critical decisions, according to specific configuration scenarios that are embedded in the Supplier Enterprise application or Web Service. This process is able to produce a signal which can be appropriately driven and managed by the Supplier Enterprise Manufacturing Execution Systems or other systems, related with the production process chain.

Finally, Fig. 5 depicts that the proposed replenishment and planning process is well structured so as the Web Services of the Distributor and the

Fig. 7. User Interface from Distributor Enterprise Web Service.

Supplier Enterprise could be replicated in a more complex network of interactive Web Services.

### 4.4 OilCMS Implementation Details

This section provides the implementation details regarding the aforementioned use case scenario and concludes with a specific section of the exposed Distributor Enterprise Web Service, implemented for the need of use case scenario.

The Distributor Enterprise Web Service, named as OilCMS, has been developed with PHP/MySQL employing the Simple Object Access Protocol (SOAP), while its communication interface is based on the Web Services Description language (WSDL) technology. The OilCMS software is installed at the Distributor Enterprise's side and is designed to work on a PHP/MySQL/Apache server. Moreover, the installation of a suitable ODBC driver is required for the daemon process implementation at the Supplier Enterprise's side. Figure 7 depicts a screenshot from the user interface of Distributor Enterprise Web Service that presents the developed product ontology.

Finally, Fig. 8 depicts a segment from the final WSDL document that defines the implemented Distributor Enterprise Web Service. In particular, the presented segment refers to a specific method of the Web Service, named as CreateProduct(), while this method is part of a sub-service of the entire Web Service, named as AdminProcess().

## 5. Conclusions

This paper presents an innovative approach towards collaborative B2B e-commerce by utilizing the RosettaNet emerging standard in combination with Service Oriented Architecture and semantically enriched information in order to seamlessly integrate the inter-enterprise (public) and intra-enterprise (private) processes. An implementation of a special case of collaborative B2B e-commerce namely the Continuous Replenishment Planning process in the oil industry is given as a use case example.

The work presented in this paper has been partially supported by the General Secretariat for Research and Technology (GSRT) of the Hellenic Ministry of Development under the framework "Competitiveness" through the project No68, Measure 4.3, Action 4.3.6.2g

```xml
<?xml version='1.0' encoding='UTF-8'?>

<definitions name="Adminprocess" targetNamespace="urn:Adminprocess"
xmlns:typens="urn:Adminprocess" xmlns:xsd="http://www.w3.org/2001/XMLSchema"
xmlns:soap="http://schemas.xmlsoap.org/wsdl/soap/"
xmlns:soapenc="http://schemas.xmlsoap.org/soap/encoding/"
xmlns:wsdl="http://schemas.xmlsoap.org/wsdl/" xmlns="http://schemas.xmlsoap.org/wsdl/">

<!-- types -->
        <types>
                <xsd:schema xmlns="http://www.w3.org/2001/XMLSchema"
targetNamespace="urn:Adminprocess">
                        <xsd:complexType name="Adminprocess">
                                <xsd:all/>
                        </xsd:complexType>
                </xsd:schema>
        </types>

<!-- message -->
        <message name="CreateProduct"/>
        <message name="CreateProductResponse"/>

<!-- portType -->
        <portType name="AdminprocessPortType">
                <operation name="CreateProduct">
                        <input message="typens:CreateProduct"/>
                        <output message="typens:CreateProductResponse"/>
                </operation>
        </portType>

<!-- binding -->
        <binding name="AdminprocessBinding" type="typens:AdminprocessPortType">
                <soap:binding style="rpc"
transport="http://schemas.xmlsoap.org/soap/http"/>
                <operation name="CreateProduct">
                        <soap:operation soapAction="urn:AdminprocessAction"/>
                        <input>
                                <soap:body namespace="urn:Adminprocess" use="encoded"
encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"/>
                        </input>
                        <output>
                                <soap:body namespace="urn:Adminprocess" use="encoded"
encodingStyle="http://schemas.xmlsoap.org/soap/encoding/"/>
                        </output>
                </operation>
        </binding>

<!-- service -->
        <service name="AdminprocessService">
                <port name="AdminprocessPort" binding="typens:AdminprocessBinding">
                        <soap:address location=""/>
                </port>
        </service>
</definitions>
```

Fig. 8. Part of WSDL document of the implemented use case scenario.

[6] and by the Network of Excellence "INTERMEDIA", [7].

## References


1. S. Shim et al., "Business-to-Business E-Commerce Frameworks",IEEE Computer, pp.40-47, October 2000
2. T. Kim et al, "A modeling Framework for Agile and Interoperable Virtual Enterprises", Computers in Industry 57, pp. 204-217, 2006
3. J. Gialelis, A.P. Kalogeras, A. Kaklis, and S. Koubias, "Collaborative Continuous Replenishment Planning Process Implementation Utilizing Web Services, Product Ontologies and Workflow Management System" Proceedings of the IEEE International Conference on Emerging Technologies and Factory Automation, ETFA2006 proceedings, pp. 574 – 581.
4. S. Lee et al., "Business Value of B2B Electronic Commerce", Electronic Commerce Research and Applications 2, pp. 350-361, 2003
5. B. Medjahed et al., "Business-to-Business Interactions: issues and enabling technologies", The VLDB Journal 12, pp. 59-85, 2003
6. http://www.gsrt.gr
7. European Project INTERMEDIA, "Interactive Media with Personal Networked Devices", http://intermedia.miralab.unige FP6 – IST- 38419.

# RMCLAB - PRIMJER KORIŠĆENJA DALJINSKE LABORATORIJE U NASTAVI ELEKTRONIKE NA UNIVERZITETU CRNE GORE


Bojan Radović*, Zoran Jakšić*, Radovan Stojanović* , Kostas Efstathiou**, Dimitris Karadimas**, Budimir Lutovac* i Nedeljko Lekić*

*Elektrotehnički Fakultet, Univerzitet Crne Gore, stox@cg.ac.yu
**University of Patras, Department of Electrical Engineering and Computer Technology, Greece



**Sadržaj -** Rad opisuje Internet-baziranu laboratoriju, nazvanu RMCLab, razvijenu na Univerzitetu u Patrasu, Grčka i adaptiranu za potrebe Univerziteta Crne Gore. Glavna odlika ove laboratorije je rad sa realnim kolima i instrumentima nasuprot simulacija i virtuelnog okruženja. Hardverska infrastruktura RMCLaba sastoji se od više rekonfiguarabilnih FPGA sistema koji prate svaki analogni i digitalni modul koji predstavlja odredjenu vježbu. Na ovaj način se postiže dinamička rekonfiguracija mjernih mjesta i položaja komponenti. Takodje se može testirati u realnom vremenu i sopstveni dizajn za generalnu namjenu. U radu se prezentira i konkretni primjer kola iz digitalne elektronike. Saradnja na ovom projektu je ostvarena u okviru Tempus projekta CD_40017_2005.


## 1. UVOD

Veliki broj univerziteta i obrazovnih institucija je suočen sa problemom obezbijedjivanja adekvatnih uslova za rad velikog broja studenata. Problem je još izraženiji kada je riječ o laboratorijskim vježbama gdje je potrebno organizovati istovremeni rad u više grupa. Savremena laboratorijska oprema je još uvijek prilično skupocena, naročito u nerazvijenim i zemljama u razvoju, tako da se proces vježbanja ograničava na mali broj instrumenata koji se dijele na veliki broj grupa.

Omogućavanje pristupa stvarnoj laboratorijskoj opremi preko Interneta bi u velikoj mjeri rješilo ovaj problem. U tom pravcu je aktuelno nekoliko pokušaja, [1] i [2]. Jedan od njih prestavlja RMCLab razvijena na Univerzitetu u Patrasu i koja obezbedjuje implementaciju laboratorijskih vježbi upotrebom realnih instrumenata u bilo koje vrijeme i sa bilo kojeg mjesta [3], [4] i [5]. Korišćenjem RMCLaba broj potrebnih instrumenata se drastično smanjuje na samo jedan laboratorijski set kojeg može dijeliti nekoliko stotina studenata, vježbajući više različitih vježbi.

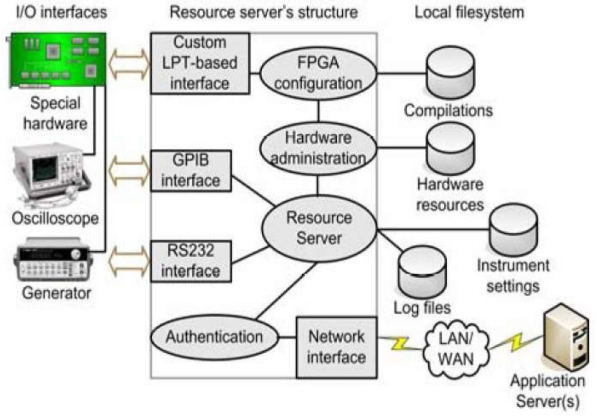
U okviru Tempus projekta CD_40017_2005 čiji je koordinator Elektrotehnički fakultet Univerziteta Crne Gore jedna od aktivnosti je i upotreba sistema daljinskog učenja za potrebe laboratorijskog treninga studenata na odsjeku elektronike. U tom smislu na postojeću strukturu RMCLaba su dodate vježbe koje se odradjuju na našem fakultetu. Na taj način je omogućeno da čitav set laboratorijskih vježbe iz analogne i digitalne elektronike bude efikasno implementiran preko Interneta.

Ovaj rad u osnovnim crtama opisuje arhitekturu i način korišćenja RMCLaba. Takodje, kroz opis načina implementacije proste laboratorijske vježbe iz digitalne elektronike teži se demonstrirati njegova svrsishodnost i efikasnost.

## 2. STRUKTURA RMC LABORATORIJE

RMCLab predstavlja moćan alat za daljinsko upravljanje i kontolisanje laboratorijskih mjernih instrumenata. Na jednom serveru preko standardnih komunikacionih portova (RS-232, EPP) povezani su mjerni instrumenti i hardver neophodan za realizaciju lab. vježbi, slika 1. Mjerni instrumenti imaju sposobnost komunikacije sa računarom. Na taj način server prikuplja podatke sa više mjernih tačaka istovremeno. Tako akumulirane informacije se prosliјeđuju korisniku (klijentu) na njegov zahtjev i u odgovarajućoj formi. Na serveru se pogoni specijalno razvijeni softver koji kontroliše rad pojedinih instumenata i namjenskog hardvera. Softver je zamišljen da u što većoj mjeri približi korisniku rad sa realnim instrumentima. U tom cilju grafički interfejs koji vidi korisnik sadrži sve funkcije za kontrolu i upravljanje realnim instrumentima.

Na ovaj način omogućeno je da se jedan set mjernih instrumenata istovremeno nalazi na raspolaganju većem broju korisnika, što je jedna od glavnih prednosti RMCLaba. RMCLab ne zahtijeva velike brzine prenosa podataka između klijenta i servera, što omogućava korisnicima i sa sporijom Internet vezom korišćenje resursa laboratorije Univerziteta u Patrasu za svoje potrebe. Studenti odsjeka za primjenjenu elektroniku ovdašnjeg fakulteta uspješno koriste RMCLab od 2004. godine na dva akademska kursa.



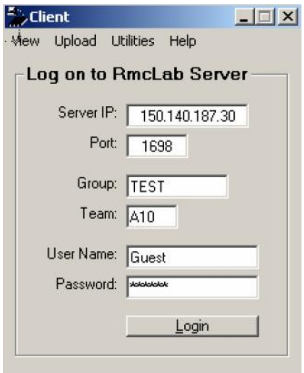

*Slika 1. Osnovna struktura RMCLaba*

RMCLab ima modularnu strukturu koja omogućava laku nadogradnju i izmjene postojećeg hardvera. Svaka laboratorijska vježba realizovana je na posebnoj štampanoj ploči i ima odgovarajuću grafičku predstavu u programskom interfejsu. To omogućava lake izmjene sadržaja laboratorijskih vježbi (nadogradnju postojećih, uklanjanje i dodavanje novih). RMCLab trenutno od mjernih instrumenata podržava rad sa osciloskopom, generatorom funkcija i izvorima napajanja, dok buduće verzije predviđaju upotrebu analizatora spektra, logičkog analizatora, frekvencijskog i digitalnog sintisajzera. Trenutno, platformu podržavaju jedan PC sa 2.6GHz procesorom i 1024MB RAMa, Agilentov osciloscop 54622D (2 analogna i 16 digitalnih ulaza) i Agilentov signal generator 33120A. Kao I/O jedinica koristi se ploča bazirana na Alterinom FPGAu iz FLEX8K familije.

Za što vjerodostojnije testiranje svaka štampana ploča tj. svaka laboratorijska vježba posjeduje više testnih tačaka sa kojih je moguće izvršiti mjerenja (snimiti talasne oblike signala ili izmjeriti neku od karakterističnih vrijednosti). Od ukupnog broja testnih tačaka, korisnik maksimalno može izabrati dvije u jednom vremenskom trenutku.

## 3. UPOTREBA RMCLaba

U svrhu korišćenja RMCLaba potrebno je instalirati besplatni klijent program koji se može downloadovati sa sajta laboratorije za primjenjenu elektroniku http://www.apel.ee.upatras.gr/rmclab/ . Pri pokretanju programa korisniku će se prikazati prozor kao na slici 2.




*Slika 2. Client log on interface.*

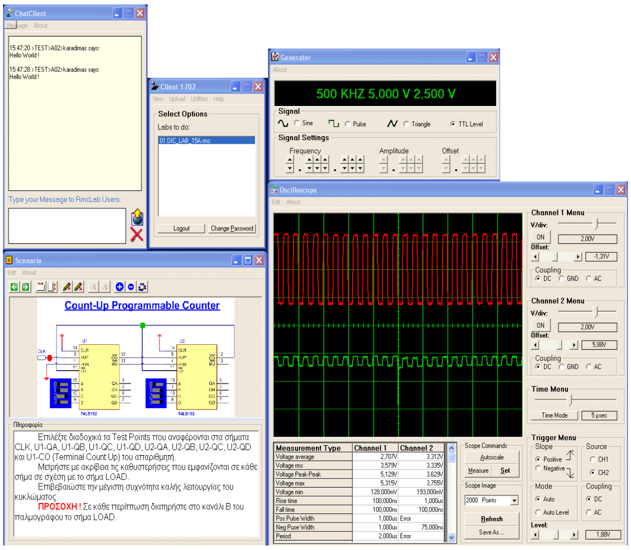
Da bi pristupio RMCLabu korisnik u naznačena polja ukucava odgovarajući nalog i šifru. Za potrebe Univerziteta Crne Gore kreirana je posebna grupa i 10 korisničkih naloga tako da studenti ovog Univerziteta mogu ravnopravno da koriste resurse kao njihovi kolege na Univerizitetu u Patrasu.

Nakon što se klient loguje, pred njim će se naći izbor od sedam laboratorijskih vježbi od koje ovaj bira jednu. Tri od ponuđenih sedam vježbi namijenjene su za osnovni kurs analogne elektronike, dok su ostale 4 namjenjene kursu digitalne elektronike. Naravno kako je to prethodno rečeno, modularnost RMCLaba omogućava njegovu stalnu nadogradnju pa će neka od novijih verzija ponuditi veći izbor laboratorijskih vježbi. Važno je napomenuti da su laboratorijske vježbe koje su do sada realizovene prilagođene jezički i programski potrebama studenata crnogorskog univerziteta.

Kada se pristupi jednoj od ponuđenih laboratorijskih vježbi, prikaže se glavni interface prikazan na slici 3. Svaka vježba sastoji se iz više koraka koji su prikazani u prozoru *Scenario*. Svaki korak sadrži određeni zadatak. U datom prozoru prikazana je elektronska šema kola koje se testira. Na posebno označenim tačkama kola moguće je postaviti sonde osciloskopa kako bi se izvršila potrebna mjerenja. Tekst ispod šeme kola daje studentu sve potrebne instrukcije u cilju implementacije vježbe. Ulazni pinovi svih kola (*clk* za digitalna kola i *Vul* za analogna i sl.) povezani su sa generatorom funkcija. Podešavanje ovog signala vrši se u prozoru generatora funkcija *Generator.* Mjerenja karakterističnih veličina posmatranih signala, kao i prikazivanje njihovih talasnih oblika kontrolište se preko ponuđenih komandi u prozoru *Oscilloscope*.

Važno je napomenuti da rezultati mjerenja dobijeni na ovaj način nijesu proizvod simulacije, koja obično ne odražava pravo stanje i ne može biti totalno vjerodostojna.

Od studenta se zahtjeva da sačuva rezultate mjerenja u izvještaju koji se po završetku vježbe šalje predavaču na razmatranje i bodovanje. Novija verzija RMCLaba će omogućiti automatsko ocjenjivanje odrađenih vježbi čime će se izuzeti faktor subjektivnosti predavača, što predstavlja jedan od ciljeva u savremenom sistemu obrazovanja.




*Slika 3. Glavni interface RMCLaba.*

Vremenske performanse sistema su testirane u odnosu na skup kašnjenja, Tabela 1. (i) Td1 - kašnjenje prouzrokovano zadavanjem parametara mjernog hardvera (amlitude, vremenske baze itd.) i na osnovu njih izvršenog mjerenja, (ii) Td2- kašnjenje za rekonfiguraciju FPGA hardvera, (iii) Td3 – kašnjenje za kompalaciju šeme laboratorijske vježbe, (iv) Td4 – kašnjenje komunikacije, slanja paarmetara i prijema podataka o izmjereneim veličinama. Kašnjene Td1 direktno utiču na brzinu dijeljenja (sharing) resursa za implementaciju vježbe, a testiranja su pokazala da je ono obično <3 sec.

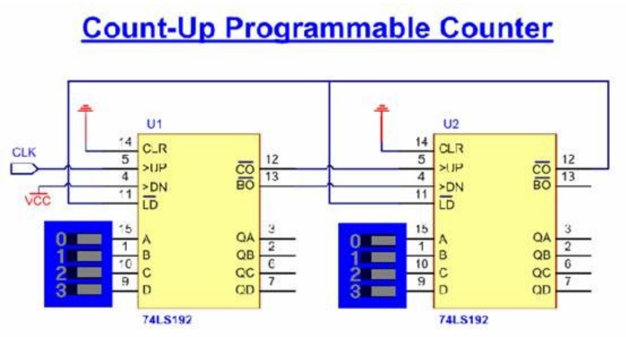
*Tabela 1: Vremenska kašnjenja pri korišćenju RMCLab.*

| Tip kašnjenja | Iznos (sec.) |
|---|---|
| Td1 | 3 |
| Td2 | 5 |
| Td3 | 10 |
| Td4 | <5 |

## 4. PRIMJER

Na narednoj slici dat je prikaz jednostavne laboratorijske vježbe iz digitalne elektronike. Radi se o 8bitnom «count-up» brojaču realizovanom pomoću dva TTL 4bitna brojača 74LS192. U prvom koraku, studentima se daje teorijska osnova vježbe, princip rada komponenti i kola u cjelini.

U sledećem koraku od studenta se traži da izvrši odgovarajuća podešavanja generatora funkcija, koji je povezan na *clk* ulaz kola. Da bi kolo ispravno funkcionisalo, izlaz generatora funkcija treba podesiti na TTL nivo (5V peak-to-peak, 2.5V srednja vrijednost) frekvencije 1.2 MHz.



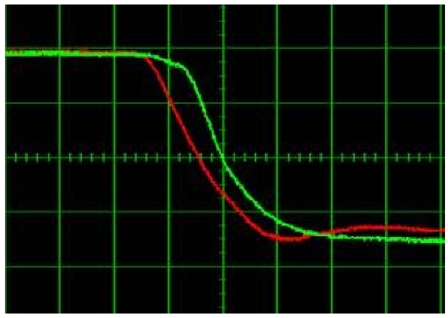

*Slika 4. Šema kola 8bitnog brojača.*

Kada se izvrše sva podešavanja generatora funkcija treba pristupiti mjerenjima. Sonde osciloskopa koje se označavaju crvenom i zelenom bojom na grafiku kola, potrebno je postaviti na neku od željenih mjernih tačaka. Kanal A osciloskopa se postavlja na *clk* ulaz, a kanal B mjenjamo u zavisnosti od zadatka vježbe.

Snimanje karakterističnih signala se postiže na sledeći način. Klikom na dugme *Autoscale* a zatim i na dugme *Refresh*, u prozoru osciloskopa dobija se prikaz talasnih oblika signala. Koristeći opciju *Save as* talasni oblici se mogu snimiti kao slika u jpg formatu, koju je kasnije potrebno uključiti u izvještaj.

U sledećem koraku od studenta se zahtjeva da izmjeri karakteristične veličine signala (amplituda, period, rise time, fall time). Klikom na dugme *Measure*, tabela u prozoru osciloskopa sa karakterističnim veličinama će se ažurirati.

Naprednijú opciju predstavlja mjerenje kašnjenja između pojedinih tačaka kola, slika 5. Na taj način se dokazuje realnost mjerenja u odnosu na simulaciju. Mjerenje kašnjenja zahtjeva ručno podešavanje osciloskopa. Pod tim se podrazumjeva podešavanje vremenske baze i ofseta svakog pojednačnog kanala. U prozoru osciloskopa se nalaze jednostavne komande kojima se podešavaju ovi parametri. Nakon ovih podešavanja potrebno je ažurirati prikaz osciloskopa, što se postiže ponovnim pritiskom na dugme *Refresh*, poslije čega se može očitati vremenska razlika između signala.



*Slika 5. Mjerenje kašnjenja, vremenska baza 50ns.*

U tekstu vježbe studentu se daju sugestije kako da izvrši mjerenja na što precizniji način, kao i upozorenja na koja treba da obrati pažnju. Pored toga, na kraju vježbe dat je set pitanja vezanih za laboratorijsku vježbu.

Izvršena mjerenja kao i odgovori na pitanja sumiraju se u posebnom dokumentu koji se šalje predavaču na evaluaciju, čime se završava laboratorijska vježba.

## 5. ZAKLJUČAK

Izvodjenje laboratorijskih vježbi na daljinu pokazalo se dosta efikasnim i korisnim. U prvom planu se misli na finansijsku uštedu, skraćeno vrijeme rada i mnogostruko uvećane kapacitete (broj studenata koji istovremeno vježbaju).

Još jedna, možda ne takouočljiva, ali veoma bitna prednost korišćanja RMCLaba ogleda se u rezultatima koji su studenti postigli prilikom izvođenja laboratorisjkih vežbi na ovakav način. Ocijenili su kao veliku prednost što mogu da izvode vježbe od kuće pa su i postignuti rezultati bili znatno bolji.

Dodatna pogodnost RMCLaba ogleda se u njegovoj modularnosti što omogućava njegovu stalnu nadogradnju. Ova prednost se naročito može zapaziti ukoliko je potrebno izvršiti testiranje nekog posebno dizajniranog elektronskog kola. Specijalno dizajnirana ploča se lako povezuje na RMCLab pa na ovaj način sistem dobija na značaju kada je naučno-istraživački rad u pitanju.

Nastavnici i saradnici sa Univerziteta Crne Gore su učestovali u adaptaciji RMCLaba za potrebe korišćenja u svom nastavnom procesu, a studenti smjera elektronika su odradili ciklus laboratorijskih vježbi upotrebljavajući ovaj alat.

## 6. ZAHVALNOST

**CD_40017_2005** Projekat je finansiran od strane Evropske komisije u okviru Tempus programa. Autori se najsrdačnije uzahvaljuju na pruženoj pomoći i podršci.

Education and Culture
TEMPUS

## 7. LITERATURA

[1] F. A. Candelas Herías*, C. A. Jara Bravo, F. Torres Medina, Flexible virtual and remote laboratory for teaching Robotics, Current Developments in Technology-Assisted Education (2006), pp. 1959-1963.

[2] Kin Yeung, Jie Huang, Development of a remote-access laboratory: a dc motor control experiment, Computers in Industry , Volume 52, Issue 3, pp. 305-311

[3] Dimitris Karadimas, Kostas Efstathiou: Design, Implementation and Evaluation of a Remote Laboratory System for Electrical Engineering Courses. ICALT 2006: Kerkrade, The Netherlands

[4] Dimitris Karadimas and Kostas Efstathiou Dept. of Electrical and Computer Engineering, University of Patras, Greece: *An Integrated Educational Platform Implementing Real, Remote Lab-Experiments for Electrical Engineering Courses.* JOURNAL OF COMPUTERS, VOL. 2, NO. 2, APRIL 2007, pp. 37-44.

[5] http://www.apel.ee.upatras.gr/

**Abstract** – Rad opisuje Internet-baziranu laboratoriju, nazvanu RMCLab, razvijenu na Univerzitetu u Patrasu, Grčka i adaptiranu za potrebe Univerziteta Crne Gore. Glavna odlika ove daljinske laboratorije je rad sa realnim kolima i instrumentima nasuprot simulacija i virtuelnog okruženja. Hardverska infrastruktura RMCLaba sastoji se od više rekonfiguarabilnih FPGA sistema koji prate svaki analogni i digitalni modul koji predstavlja odredjenu vježbu. Na ovaj način se postiže dinamička rekonfiguracija mjernih mjesta i položaja komponenti. Takodje, se može testirati u realnom vremenu i sopstveni dizajn za generalne primjene. U radu se prezentira i konkretni primjer kola iz digitalne elektronike. Saradnja na ovom projektu je ostvarena u okviru Tempus projekta CD_40017_2005.

**RMCLAB – AN EXAMPLE OF THE USING REMOTE LABORATORY FOR ELECTRONICS COURSES AT UNIVERSITY OF MONTENEGRO**

Bojan Radović, Zoran Jakšić, Radovan Stojanović, Kostas Efstathiou, Dimitris Karadimas, , Budimir Lutovac, Nedeljko Lekic

# Collaborative B2B e-Commerce Utilizing Semantic Web-Services and Smart Agents

J. Gialelis1,2, D. Karadimas1,2, P. Chondros1,2, D. Serpanos1,2
[1] Industrial Systems Institute/RC Athena
Patras Scientific Park Building
Stadiou St., GR-26504 Platani Patras, Greece
{gialelis, pchondros, karadimas, serpanos}@isi.gr

[2] Dept. of Electrical and Computer Engineering
University of Patras, GR-26504 Patras, Greece

*Abstract*

Collaborative B2B e-commerce provides enterprises with the necessary level of flexibility and efficiency in order to retain competitiveness under the increasingly turbulent business environment. Web-Services are often utilized by enterprises in order to integrate high and low level enterprise applications, thus providing a collaborative B2B e-commerce environment without affecting inter- and intra-enterprise processes. Nevertheless, the above paradigm should be enhanced in order to comply with demanding B2B e-commerce environments, such as the oil/gas domain, where restricted specifications should be met. This paper presents an integrated approach towards collaborative B2B e-commerce for demanding industrial domains, in terms of precision and reliability, enriched with smart agents for on-demand dynamic adaptation of existing Web-Services. The proposed approach utilizes the RosettaNet emerging standard, in combination with Service Oriented Architecture and semantically enriched information in order to seamlessly integrate the inter and intra-enterprise processes. Finally, the proposed approach is presented as a study case in the oil/gas industry domain.

## 1. Introduction

Economic pressure and turbulences in the global business environment introduce more and more demands for manufacturing enterprises in order to make predictable long-term provisions, as these factors impose radically changes in the way that enterprises operate with other enterprises in order to accomplish a business goal. In this context, Business-to-Business electronic transactions (B2B e-commerce) outfit the enterprises with the necessary tools for real-time interaction with their suppliers, dealers and customers. As a result, competition among businesses turns gradually from an enterprise-to-enterprise matter to a value chain-to-value chain one.

E-Business frameworks represent the state of the art in e-commerce; since framework-based B2B e-commerce effectively addresses the seamless linking of inter-enterprise (public) processes involving Enterprise layer's applications to those of business partners. These frameworks provide the necessary layers to achieve interoperability in e-commerce, since their interactions involve various business processes and business components, such as executable applications, systems, and their associated information [1].

In order to support the interchange of information among "public" modules, several standards and languages have been developed. EDI (Electronic Data Interchange), EDIINT (EDI over the Internet), ebXML (e-business eXtended Markup Language) and BizTalk are some generic standards addressing vertical as well as horizontal integration issues. Furthermore, standards such as RosettaNet, IOTP (Internet Open Trading Protocol), ICE (Information and Content Exchange) and OBI (Open Buying on the Internet) deal more with vertical market-oriented integration issues.

However, current framework-based developments consider inter-enterprise processes, involving only Enterprise layer applications, information, etc; without taking into consideration the intra-enterprise processes involving applications, information residing into the other enterprise layers such as Plant and Shop Floor layers thus, not assuring agility, flexibility and autonomy of the interacting partners [2].

The work presented in this paper is focused in the demanding and critical industry domain of oil/gas and intents not only to overcome the aforementioned drawback but also to enhance existing framework architectures with smart agents for on-demand, dynamic adaptation of existing Web-Services, thus managing to incorporate critical industry-related requirements, i.e. precision and reliability, within the B2B framework. Additionally, the proposed architecture extends the functionality of current frameworks in order to incorporate all enterprise layers into B2B e-commerce interactions, thus elevating an innovative collaborative business model which seamlessly integrates the inter-enterprise (public) with the intra-enterprise (private) processes.

The paper is organized as follows: Sections 2 and 3 describes the collaborative B2B e commerce and the Semantic Web-Services and Smart Agents, respectively, while Sections 4 and 5 describe the proposed set of technologies imposed over the e business frameworks' holistic architecture and the implementation of a collaborative B2B e-commerce business model for the oil/gas industry, according to the proposed approach.

## 2 Collaborative B2B e-commerce

Collaborative B2B e-commerce differs from basic B2B e-commerce, since collaborative commerce goes beyond on-line document exchanges, indicating that organizations adopt B2B networks to establish new collaboration mechanisms with channel partners. Thus, enterprises should establish collaborative B2B processes with cooperating manufacturing enterprises, in order to optimize exploitation and gain the most out of the adoption of electronic commerce networks. Collaborative B2B processes are considered to be any end-to-end business processes, which results from the seamless integration of intra-enterprise processes with inter-enterprise processes.

A designated example of a collaborative B2B process is the Continuous Replenishment Planning (CRP) model, which can constitute the base for the development of one of the most efficient inventory management function. The key characteristics of CRP are the sharing of real time inventory data by retailers with manufacturers and continuous replenishment of retailers' inventory by manufacturers. This function, called continuous replenishment planning, is defined as the practice of partnering between distribution channel members that changes the traditional replenishment process from a warehouse/distributor centric process handling purchase orders, to an end-user/retailer centric replenishment process involving the manufacturing enterprise in the overall process utilizing actual and forecast data.

Even though, the CRP collaboration model has the potential to alter significantly production-scheduling methods and inventory management practices, so that enterprises enjoy the above-mentioned benefits, survey results do not keep up with these expectations [3]. This drawback is caused, mainly, due to CRP deployment methodologies. Such methodologies are based on point-to-point architectures and on the utilization of proprietary solutions/technologies which present very limited tolerance to system/applications alterations and minimum scalability to value-network expansion. In order to overcome these drawbacks and achieve both integration of heterogeneous systems/applications, at any layer in the enterprises they may reside in and unlimited expansion of the value-network, this paper proposes the deployment of specific set of technologies over ERP (Enterprise Resource Planning), MESS (Manufacturing Execution SystemS), SCADA (Supervisory Control And Data Acquisition) and OPC (OLE for Process Control) systems, as depicted in Fig. 1. These technologies are based on Internet ubiquity and emerging standards.

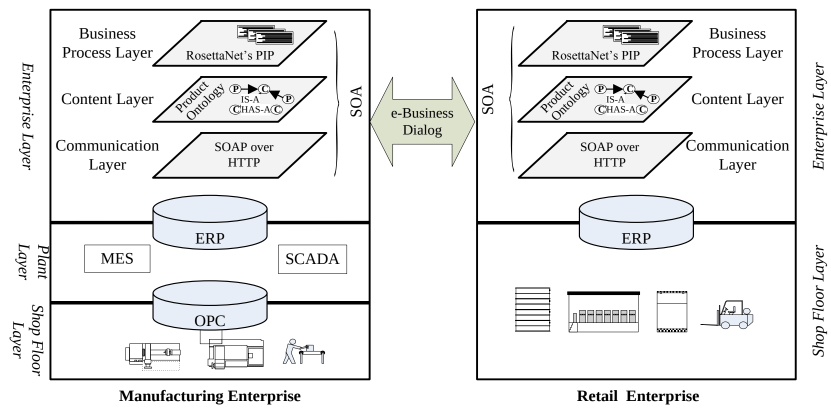



**Fig. 1.** e-Business framework's holistic architecture.

Therefore, Service oriented Architecture (SoA) elements, RosettaNet's ontology and standard processes for the B2B transactions are utilized, in order to fully implement the CRP collaboration model. Moreover, SOAP is utilized as the transport message protocol in order to satisfy communication needs, while ontology is used in order to standardize content semantics and WSDL (Web Service Description Language) is used in order to expose systems' functionality as Web Services. The above mentioned principles of the e-business frameworks' holistic architecture, for a manufacturing and a retail enterprise are further described in the next section.

## 3 e-Business Framework and Components' Technologies

Several contemporary B2B interaction frameworks are based on the XML notation. They aim at overcoming some of the limitations of traditional electronic data interchange standards. XML-based frameworks for B2B transactions follow a hierarchical model, which comprises three layers: the Communication, the Content and the Business Process layer [4]. The Communication layer provides the appropriate means of communication for exchanging the required messages among the involved systems/applications, while the Content layer provides the necessary tools (languages and conceptual models) so that the involved entities agree upon the semantics of contents and

the types of the exchanged documents. Finally, the Business Process layer is concerned with the conversational interactions among services. The following three sub-sections present the proposed imposed technology solutions for the three aforementioned layers.

### *3.1 Communication Layer*

For the communication needs, the SOA paradigm is utilized by employing the use of Web-Services technology in order to open up enterprise systems, plant systems and production systems; thus making available their functionalities. Therefore, the Simple Object Access Protocol (SOAP) and the Web-Services Description Language (WSDL) are deployed at the communication layer.

### *3.2 Content Layer*

The content layer provides the suitable means to describe and organize exchanged information in such a way that it can be understood and used by all the collaborating partners. Content interactions require that the involved systems understand document content semantic. Our work is based upon the development of elaborated petroleum product ontology in order to achieve a uniform product terminology. Furthermore, existing work in different standards, such as RosettaNet Technical Dictionary (RNTD), is utilized wherever required.

### *3.3 Business Process Layer*

Our work, at this layer, emphasizes on the RosettaNet framework in order to address business process layer needs. RosettaNet focuses on three key areas of standardization to automate B2B interactions. First of all, the vocabulary needs to be aligned. Secondly, the business process governing the interchange of the Business Messages themselves must be harmonized and specified. Finally, the way in which Business Messages are wrapped and transported should also be specified.

Next section presents the principles of semantic Web-Services and also the mechanism to enhance the Web-Services –based e-Business dialog with the use of smart agents for the automatic synthesis of semantic information from both higher and lower enterprise layers, in order to provide an integrated e-Business framework compatible with the demanding requirements of reliability, security that appear in an enterprise within the oil/gas industry domain.

## 4 Semantic Web-Services and Smart Agents

Communication between Web-Services -based systems is done through the use of the SOAP and the Web Service Description Language (WSDL). WSDL specifies a way to describe the functionalities of Web-Services and how and where to invoke them. However, the WSDL W3C Recommendation does not include semantics in the definition of a Web-Service. Thus, two services can have similar description but different meaning, or they can have different descriptions but similar meaning. Resolving these ambiguities in Web-Service definitions is not only an important but also a necessary step toward automating the discovery, the composition and the deployment of Web-Services. Being able to discover, compose and deploy new Web-Services, based on existing Web-Services, will give a more dynamic, adaptive and accurate grid of services, able to be employed in critical environments.

Semantic tools provide capabilities for automatic discovery of topics and concepts. These tools may also extract the meaning of the information provided, categorize, correlate and map information from different sources. By using ontologies and semantic technologies it is possible to automatically discover and deploy semantic annotated Web-Services and employ their functionality. Moreover, semantic described Web-Services allow the automatic combination of services in order to create new services, and in this sense Web-Services will grow together with the data on the web. The sequence of Web-Service requests may be orchestrated, and then responses may be gathered together into composites that deliver a more comprehensive view of the web.

The aim is to automatically discover, connect to and deploy Web-Services. A client should be able to do semantic search for Web-Services and upon discovery automatically connect to and deploy the selected Web-Service(s). This automation requires that all the processes that can be performed by a Web-Service are described. Processes should then be categorized on the basis of what they do. To enable efficient Web-Service discovery, the services need to be registered somewhere, since the publication of their description on the service provider’s home page is not adequate. Several solutions for Web-Service registry and discovery exist, like the Universal Description, Discovery and Integration specification (UDDI). These are primarily key-word based and human interaction is needed to discover and connect to a Web-Service. Therefore, the automatic discovery and deployment requires the use of a semantic search engine that describes Web-Services functionalities. Then, service providers could register their services at the search engine repository by adding the semantic description of their services, while clients could thereafter search for Web-Services by description, since the description of Web-Services is functional-based, providing a semantic description of the operations performed by the service.

To fulfill more advanced tasks for an industrial environment a matching engine is needed. This engine must provide discovery, negotiation, filtering, choreography and reasoning, based on the population of the descriptions of different services; which services perform which operations, what are the expected inputs, outputs, preconditions and effects. Thus, three essential types of knowledge about a Web-Service needed to be provided: what the service does/provides, how it is used/works and how to interact with it.

The Web Ontology Language (OWL) descriptions, collected in a search engine, enable semantic search for operations and proper Web-Services to perform the desired operations. The semantic description enables queries for Web-Services based on their expected input and/or output, since all the concepts in the queries are related to ISO definitions. Based on the Web-Service descriptions, smart agents can reason over the descriptions and orchestrate new services. The use of Business Process Execution Language (BPEL) enables different parties to construct their own smart agents with reasoning abilities as desired.

The goal regarding a demanding industrial environment is to provide reliable and secure Web-Services. For the accomplishment of this task a smart agent as a matching engine should consider not only the public Web-Services, within a public UDDI registry, but also private/protected services exposed by the OPC layer of an enterprise. Thus, a smart agent takes over the composition of new services, based on information located both on high level layers, such as the ERP, and on low level layers, such as the OPC. The smart agent is not only sophisticated enough to reason over the descriptions of the services and orchestrate new services but also is capable to access low-level layers and incorporate their reliable and protected from public view information to the new services, thus providing reliable semantic Web-Services, without endanger the security of the low level information.

Fig. **2** depicts a generic example of the aforementioned functionality. For this example we assume three different semantic Web-Services A, B and C, where the first two are connected to the ERP layer and the third is connected to the OPC layer. Both services A and B could be public and registered within a public UDDI registry, while service C is private/protected since it exposes low level information. A smart agent is then capable of reasoning over more complex information, ABC in our example, acquired by an external enterprise, search for Web-Services that might provide this information, access the low level and protected service C, combine the information and return it, as illustrated in

Fig. **2**. After the smart agent is queried for an ABC information, interprets the ABC element and appreciates that it consists of several elements that can be acquired from several Web-Services. Smart agent can authenticate itself and gain access to the low level services, if required. Then, smart agent requests the information, collects it and responds to the query with an ABC element.

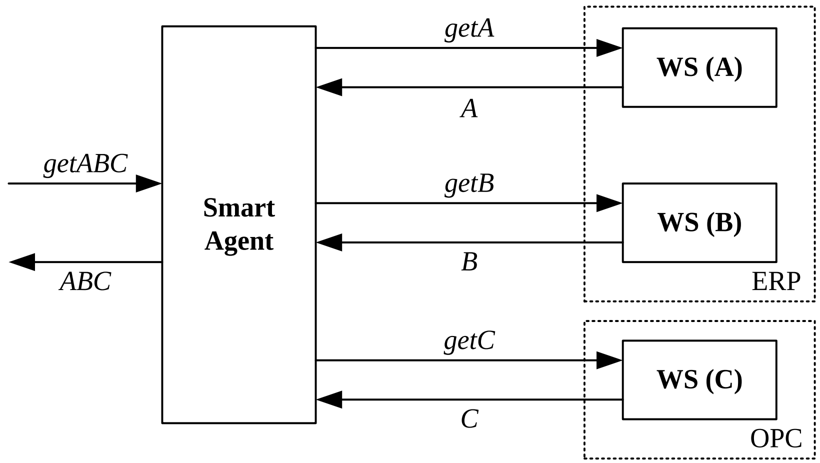



**Fig. 2.** Smart Agent – collect, organize and compose

The aforementioned mechanism provides semantic Web-Services with advanced functionality, high reliability and accuracy, since it enhances high level information with low level elements, without affecting the security and privacy of the low level enterprise layers. The deployment of this mechanism could enforce an industry enterprise to optimize its manufacturing production and forecast the production levels, which are both two of the most significant factors that influence the performance of enterprises within the oil/gas industry domain.

Next section describes in details an implementation example of a use case scenario of the proposed approach depicting a real world Continuous Replenishment Planning process application in the oil/gas industry which involves gas stations or other enterprises as end users/retailers and a refinery as the manufacturing enterprise.

## 5 B2B Framework Architecture for Oil/Gas Industry

The architecture presented in this section aims at the integration of high-level inter-enterprise (public) processes with the low-level intra-enterprise (private) processes by utilizing the mechanism of smart agents, described in the previous section. However, there are some key characteristics of such an online

system/framework: (a) ability to connect to any available ERP system of the enterprise in order to update product and pricing information, (b) ability to

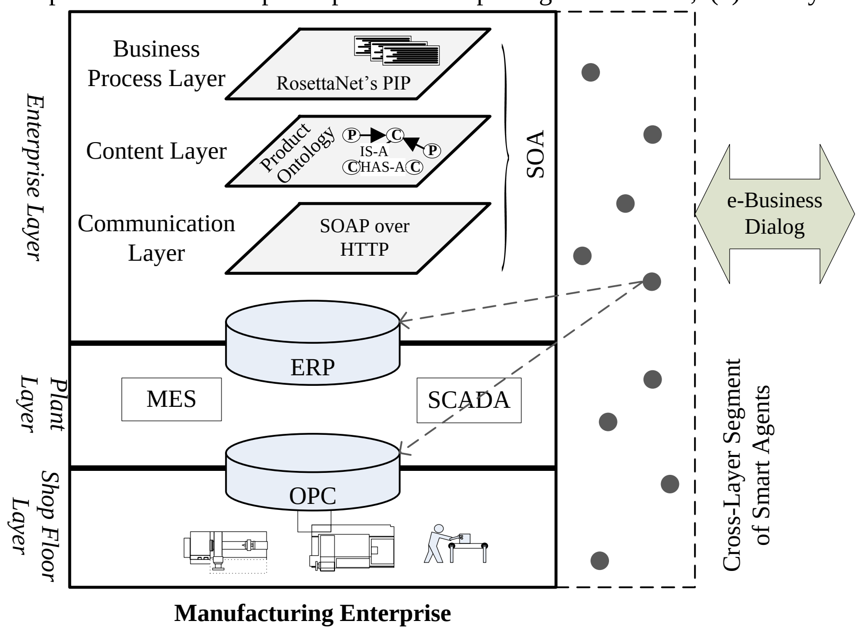



**Fig. 3**. Proposed B2B framework architecture.

support enterprises with different products and pricing policies and finally (c) ability to preserve privacy and security when accessing low-level information among with reliability and accuracy of the final provided semantically enriched information.

Fig. **3** illustrates the incorporation of the smart agent mechanism within the Manufacturing Enterprise layers, as presented in Section 3 and depicted in Fig. 1. Since the goal of the smart agents is to combine information from both higher and lower layers, so as to comprise reliable and accurate Web-Services, they have to be vertically expanded through the different enterprise layers, leading to vertical cross-layer segment. Smart agents provide the ability to securely access the sensitive lower layers and are also hereafter responsible for the e-Business dialog, occurring between Enterprises or between an Enterprise and its Customers, where the first mentioned case is the more demanding for the oil/gas industry.

### 5.1 Use Case Scenario

Oil/gas ordering and delivering is a demanding procedure, in terms of catching-up with the deadline dates, managing the paperwork and generally taking care of all the non-automated processes, inside an Enterprise, between a Manufacturing and a Retail Enterprise and finally between a Customer and an Enterprise. Thus, three entities have to be defined as follows: (a) the Customer/User which makes the order(s), (b) the Retail Enterprise (playing a distributor role) which collects the orders and has the duty to deliver the order to the customers and (c) the Manufacturing Enterprise (playing a supplier role) which supplies the distributor with the appropriate quantities of the ordered products. The proposed framework automates this end-to-end online ordering infrastructure, based on semantic Web-Services together with the smart agents' mechanism. Apparently, it is also able to keep track and provide, through Web-Services, the different order states during the whole process, which is of major significance for the logistics part regarding the specific use case.

An operational diagram of the implemented use case scenario is depicted in Fig. 4. The distributor enterprise exposes several Web-Services via a segment of smart agents that communicate with high and low layers within the enterprise. The aforementioned Web-Services are identifiable either by the UDDI protocol or with the use of OWL-S ontology for semantic Web-Services. Fig. 4 illustrates the functionality of the smart agent mechanism for composing Web-Services enriched with information acquired from both higher and lower layers of the enterprise.

Fig. 5 depicts a screenshot from the user interface of Distributor Enterprise Web-Service that presents the developed product ontology. Furthermore, the process terminology that has been employed accrues from RossettaNet technical dictionary.

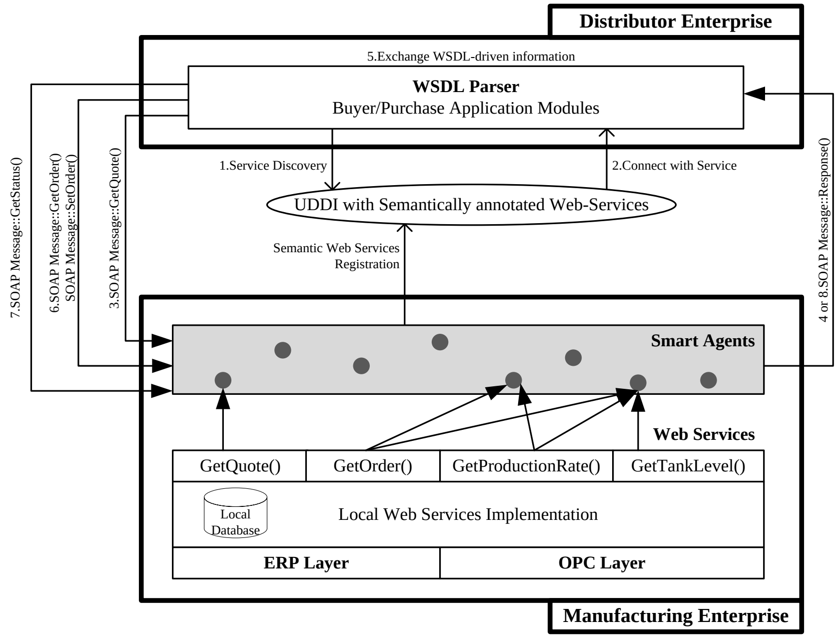



**Fig. 4**. Operational diagram of the implemented use case scenario.


**Fig. 5**. The developed product ontology for the specific use case scenario.

## 6 Conclusions

This paper presents an innovative approach towards collaborative B2B e-commerce by utilizing RosettaNet standard in combination with Service Oriented Architecture and semantically enriched information in order to seamlessly integrate the inter-enterprise (public) and intra-enterprise (private) processes. An implementation of a special case of collaborative B2B e-commerce namely Continuous Replenishment Planning process in the oil/gas industry is given as a use case scenario. Also, the proposed B2B framework architecture incorporates semantic Web-Services and a smart agent mechanism that is able to provide reliable and accurate information.

Although the proposed framework architecture has been initiated and also tested for the logistics automation between the entities of a Supplier Enterprise, a Manufacturing Enterprise and a Customer, it could be also applied for the end-to-end support of the oil/gas industry domain; i.e. from the oil-well to the oil/gas station, since it manages to provide semantically enriched low level information in a secure, reliable and accurate way.

The work presented in this paper has been partially supported by the General Secretariat for Research and Technology (GSRT) [5] of the Hellenic Ministry of Development under the framework "Competitiveness" through the project No68, Measure 4.3, Action 4.3.6.2g and Network of Excellence (NoE) Intermedia [6].

### References

1. Simon S.Y. Shim, Vishnu S. Pendyala, Meera Sundaram, Jerry Z. Gao: Business-to-Business E-Commerce Frameworks. IEEE Computer, vol. 33, no.10, pp. 40--47, (October 2000)
2. Kim Tae-Young Kim, Sunjae Lee, Kwangsoo Kim, Cheol-Han Kim: A modeling framework for agile and interoperable virtual enterprises. Computers in Industry, vol. 57, issue 3, pp. 204--217. Elsevier Science Publishers (April 2006)
3. S. Lee et al.: Business Value of B2B Electronic Commerce. Electronic Commerce Research and Applications 2, pp. 350--361 (2003)
4. B. Medjahed et al.: Business-to-Business Interactions: issues and enabling technologies. The VLDB Journal 12, pp. 59--85 (2003)
5. General Secretariat for Research and Technology, http://www.gsrt.gr
6. European Project INTERMEDIA: Interactive Media with Personal Networked Devices, FP6–IST- 38419, http://intermedia.miralab.unige

John Gialelis received his B.S.E.E degree from Fairleigh Dickinson University, NJ, USA, his M.S.E.E degree from Polytechnic University, NY, USA and his Ph. D from the University of Patras, Greece. He is currently in the Department of Electrical and Computer Engineering at the University of Patras and a senior member of the Industrial Systems Institute. He has published over 40 papers in international journals and conference proceedings. His main research interests include enterprise integration, embedded systems in critical applications, B2B architectures.

Dimitris Karadimas received his B.Sc. in electrical and computer engineering from the University of Patras, Greece. He is currently a Ph. D candidate in the Department of Electrical and Computer Engineering at the University of Patras and a member of the Industrial Systems Institute. He has published over 10 papers in international journals and conference proceedings. His main research interests include embedded systems in critical applications and B2B architectures.

Petros Chondros received his B.Sc. in electrical and computer engineering from the University of Patras, Greece. He is currently a Ph. D candidate in the Department of Electrical and Computer Engineering at the University of Patras and a member of the Industrial Systems Institute. His main research interests include embedded systems in critical applications and B2B architectures.

Dimitrios Serpanos is the Director of ISI/RC Athena, Chairman of the Governing Board of the University of Western Greece and a Professor of Electrical and Computer Engineering at the University of Patras, Greece. He holds a PhD and an MA in Computer Science from Princeton University and a Diploma in Computer Engineering and Informatics from the University of Patras. Professor Serpanos works in the area of computer architecture with emphasis on embedded systems, network systems and security systems. He has published extensively in scientific journals and conferences and he has co-authored and co-edited several books in the above technical areas.

# Identifying Chronic Disease Complications Utilizing State of the Art Data Fusion Methodologies and Signal Processing Algorithms

John Gialelis, Petros Chondros, Dimitrios Karadimas, Sofia Dima, and Dimitrios Serpanos

Industrial Systems Institute/RC Athena,
PSP Bld, Stadiou St., GR26504 Platani Patras, Greece
{gialelis,serpanos}@isi.gr,
{pchondros,karadimas,sdima}@ece.upatras.gr

**Abstract.** In this paper a methodology for identifying patient's chronic disease complications is proposed. This methodology consists of two steps: a. application of wavelet algorithms on ECG signal in order to extract specific features and b. fusion of the extracted information from the ECG signal with information from other sensors (i.e., body temperature, environment temperature, sweating index, etc.) in order to assess the health state of a monitoring patient. Therefore, the objective of this methodology is to derive semantically enriched information by discovering abnormalities at one hand detect associations and inter-dependencies among the signals at the other hand and finally highlight patterns and provide configuration rulesets for an intelligent local rule engine. The added value of the semantic enrichment process refers to the discovery of specific features and meaningful information with respect to the personalized needs of each patient.

**Keywords:** patient monitoring, algorithms, data fusion.

## 1 Introduction

Heart disease is the most important cause of death in many countries. Thus an automated solution of pervasive heart monitoring is required in order to take care of senior chronic heart patients. The electrocardiogram is an important signal for providing information about functional status of the heart. It shows how fast the heart is beating, whether the heart beat is steady or irregular and the strength and timing of electrical signals as they pass through the heart. The processing of the ECG signal in order to extract specific features of the P wave, the QRS complex, the T wave and ST segment, which in conjunction with other physiological parameters such as blood pressure and sweating index, is of great importance in the detection of cardiac anomalies. Furthermore, it is essential to acquire physiological signals from any type of environment - clinical, domestic, rural and urban - accurately and properly classify them so physicians or medical experts are able to correctly evaluate the patient's

K.S. Nikita et al. (Eds.): MobiHealth 2011, LNICST 83, pp. 256–263, 2012.

health status and perform the appropriate actions. Therefore it is essential to follow an approach using various signals and features in order to measure the same underlying clinical phenomenon: the gradually worsening condition of a chronic patient. The aim is to improve the quality and robustness of the indicator, thus improving sensitivity without causing false alarms. The final goal is to come up with a system that measures a number of parameters from easily applicable low-cost sensors, and simultaneously forms a powerful diagnostic tool. The envisaged high level architecture of such a system is depicted in Fig. 1.

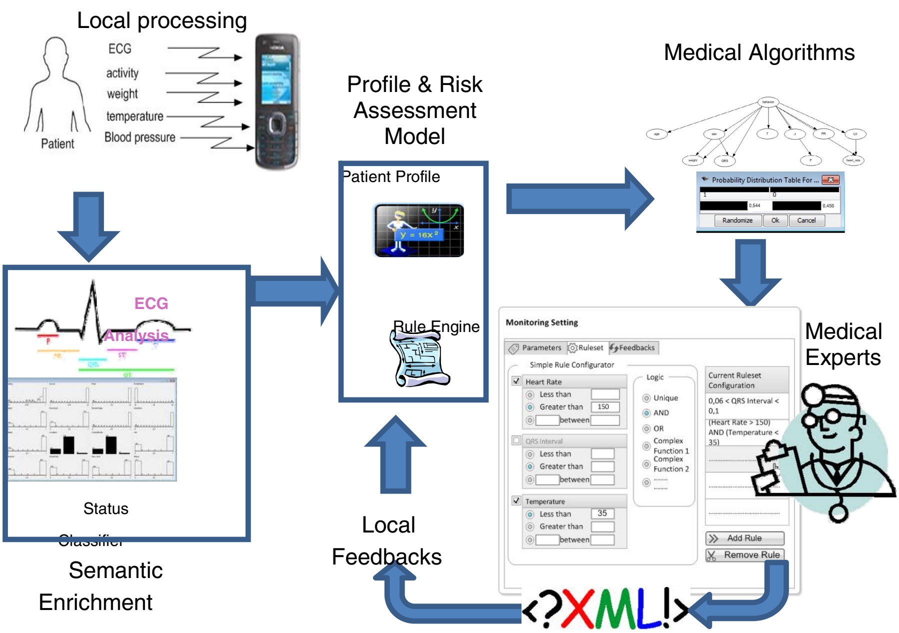



**Fig. 1.**The envisaged high level architecture of the proposed system

## 2 Signal Processing Algorithms for ECG Features Extraction

In the clinical domain of patients with chronic heart disease physicians can evaluate patients' health status based on ECG signal analysis. Furthermore, the extraction of ECG features, as they are depicted in Fig. 2, enables the medical experts not only to evaluate heart problems, such as arrhythmia and cardiovascular diseases, but we argue that additional features and meaningful information with respect to the personalized needs of a patient could be discovered.

ECG signals are gathered through appropriate sensors. The measurement of an ECG signal always imposes noise and artifacts within the frequency band of interest.

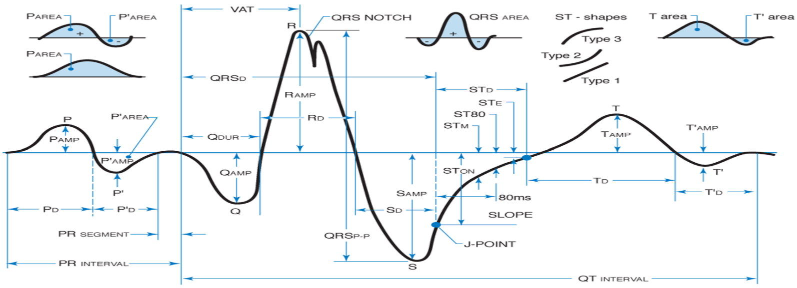



**Fig. 2.** ECG waveform features

Therefore suitable methods and algorithms have been developed in order to derive the useful information from the noisy ECG signals. Specifically, a preprocessing stage removing noise and/or artifacts from the raw ECG results the starting point of ECG analysis. Wavelet based algorithms can provide efficient localization in both time and frequency [1] thus constitute the most efficient and commonly used approach. The detection of fiducial points [2] which are relevant and strongly correlated to several heart diseases [3] is performed through a parameter extraction stage, as depicted in Fig. 3. This task becomes even more challenging due to the presence of artifacts and time-varying morphology. The key point is to provide a description of the ECG signal in the time-scale domain allowing the representation of the temporal features of a signal at different resolutions.

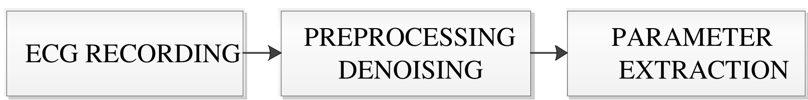



**Fig. 3.** ECG signal processing for parameter extraction

Furthermore the basic steps of the ECG signal processing procedure, starting with the QRS complex detection, are depicted in Fig. 4. With respect to the QRS complex, Q and S peaks occur before and after R peak, respectively. R peak is detected on the first scale of Wavelet Analysis. Multi resolution analysis of Wavelet Transform on time-domain ECG is utilized at different Wavelet scales, and the modulus-maxima and zero-crossing approach has been exploited at each resolution level. At the second resolution level, QRS onset and offset (J point) can be identified by using modulus-maxima on the wavelet coefficients. Since wavelet analysis preserves the time domain information of the original signal, the location of these characteristic points (Q, R and S) can be easily obtained from the wavelet domain. Finally, T wave analysis is achieved by considering details at lower frequencies. At the fifth resolution level of wavelet analysis, T wave is detectable using the modulus-maxima based approach on the wavelet coefficients. Appropriate time windows and amplitude thresholds are required in order to get rid of irregular points and misdetections. This is essential so as to provide a robust ECG signal processing methodology.

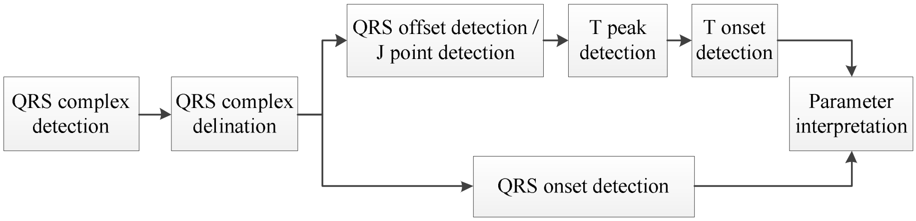



**Fig. 4.** ECG signal processing procedure

## 3 Data Fusion Model

An appropriate model describing a fusion process of data obtained from body area sensors, EHR (Electronic Health Records) and/or other historical data has been utilized. Based on these raw data, patient's characteristics and expert's knowledge, a mathematical framework is proposed to extract knowledge. The framework's basic goal is to determine the patient's health status and potentially to be used as a medical decision support tool.

Specifically, medical data, as they may come from different sources, are in different format and types. Typical examples are ECG, medical images, historical data and also habits, such as smoking, exercising etc. Due to the unique nature of these data and in order to extract knowledge, data mining approach is needed as any patterns found should be capable of human interpretation. Moreover relationships or patterns that are extracted may not be commonly accepted or conform to current medical knowledge. Results may indicate an association between the physiological parameters and the class (medical condition of the patient) but there is no implication of cause and effect. The interpretation of these associations is fully up to the experts, so additional visualization and statistical processing is needed to present results in a human interpretable format. Moreover, as the evaluation of the results fully depends on the experts, the proposed methodology can only serve as a decision support tool regarding the prognosis and the diagnosis of the monitoring patient.

There are two main approaches during the design process: supervised learning and unsupervised learning. The basic requirement of classification of patient's health status using supervised learning algorithms is that the dataset must be annotated. The most commonly used algorithms are C4.5 (decision tree approach), Multilayer Perceptron and Naïve Bayes. Unsupervised learning's goal is to determine how a set of data is structured. Clustering and blind source separations are typical unsupervised learning techniques.

Having WEKA [4] available, it was decided to use both supervised and unsupervised machine learning techniques. Besides the typical classification problems, there is an additional need to discover any associations between attributes (physiological parameters). The resulting relations and rules may enlighten hidden interdependencies and provide the experts with meaningful information regarding the patient's health status.

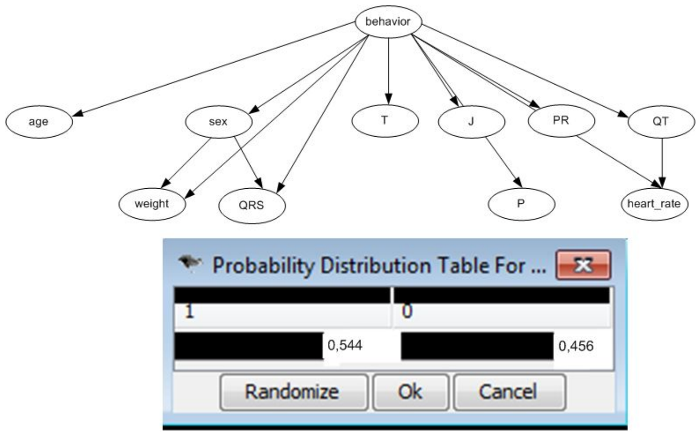



**Fig. 5.** Screenshot of the Weka tool depicting the developing classifier

In Fig. 5, a screenshot of the Weka tool is depicted. Specifically, the classification process is implemented using NaiveBayes algorithm with 3-parent configuration. The class 'behavior' represents the health status of the patient that can be either normal or abnormal, while the selected attributes that form the annotated dataset are general personal data (age, sex, weight), physiological monitoring parameters (body temperature), as well as extracted ECG features (QRS duration, J point, P wave, Q-T duration, Heart Rate, P-P duration). The training of the model ended up to the structure depicted and indicates some kind of relationship between P wave and P-P duration as well as QT duration and heart rate regarding the ECG extracted features. During testing and run of different algorithms and approaches, it was made clear that results were strongly depended on the dataset we were using, meaning that the algorithms' performance is highly correlated to the size and the quality (missed values) of the used dataset. Therefore, the above findings may trigger the expert for further investigation of the proposed associations outside the scope of the machine learning approach.

## 4 Methodology for Patient's Personalized Monitoring

The aforementioned data fusion techniques aim to support the doctors and the medical experts in general and also to potentially contribute in discovering additional personalized medical features that will provide information about patients' health status. Under this scope, personalized monitoring of patients could be conclusively accomplished as a workflow mechanism, and lead the doctors to a decision driven by patient's unique parameters as depicted in Fig. 6.

As previously described, the applied data fusion models regularly investigate sensor data along with ECG extracted features for irregularities, associations and inter-dependencies so as to detect abnormal patterns. Once an irregularity or an

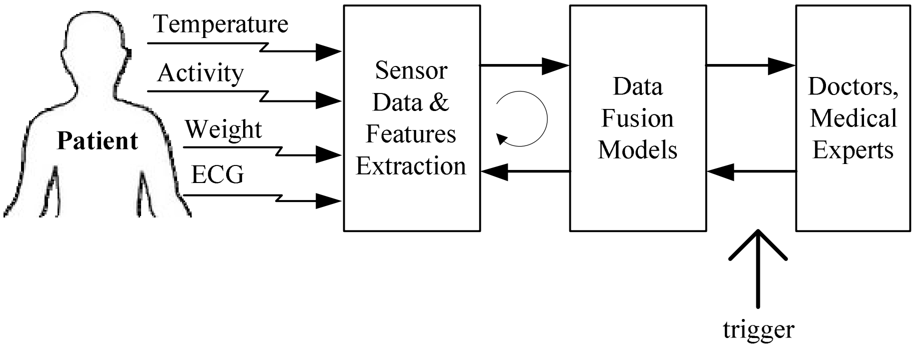



**Fig. 6.** Workflow mechanism exploiting data fusion assets for conducting doctors personalized monitoring

abnormal pattern is discovered, a trigger signal is provided to the medical expert. After this, the medical expert is able to re-assess the set of rules for the specific patient through a simple set of user interfaces (UIs) as depicted in Fig. 7.

The first step for the re-assessment involves the adding or the subtracting of physiological parameters to be monitored depending on the findings of the data fusion model, designated as “Machine Learning (ML) Findings”.

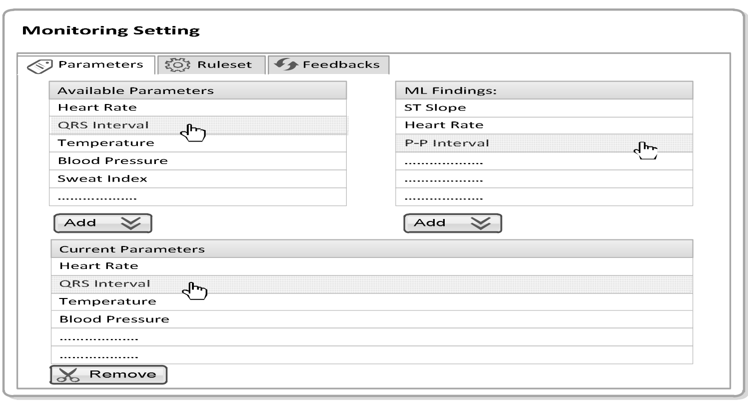



**Fig. 7.** Adding physiological parameters to be monitored

At a second step, the definition of a set of rules, according to the doctor’s expertise and experience, complements patient’s treatment towards personalization. The personalization concept lies on the definition of different physiological parameters ranges that correspond to the normal state for each patient. The key feature of personalization is associated with the ability to create rules with the selected physiological parameters as presented in Fig. 8. So far the proposed methodology will provide the doctors and medical experts a ruleset configuration covering almost any

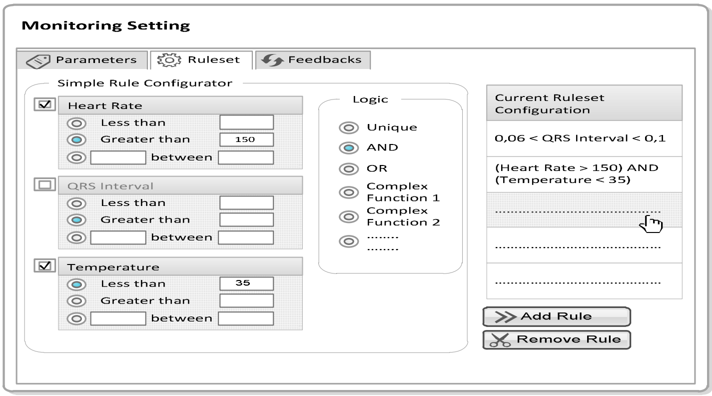



**Fig. 8.** Creation of simple rulesets

potential irregularity for a given patient, since this ruleset will be constructed based not only on doctors' aspect but also on identified abnormalities issued by the "Machine Learning Findings" as described previously.

Eventually, the constructed ruleset configuration, derived in an interoperable XML-based format such as PMML v4.0 Standard [5], is forwarded to an appropriate, operated on patient's portable device, application that monitors and records his health status. Moreover, doctors and medical experts are also able to predefine a set of feedback actions that will be executed from that application in case of a hit in any of the rules within the defined ruleset, as depicted in Fig. 9.

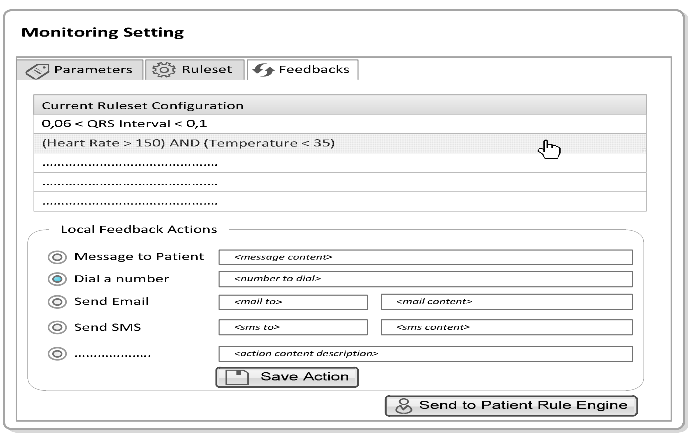



**Fig. 9.** Defining feedback actions and exporting configuration rulesets

## 5 Conclusion

In this paper high level system architecture for identifying patient's chronic disease complications is depicted. Furthermore, a methodology was proposed. This methodology which consists of two tasks: application of algorithms on primary sensor readings (e.g. ECG) in order to derive semantically enriched information and a data fusion model that uses the above extracted information as well as data from other sensors (body temperature, environment temperature, sweating index, etc.) in order to assess the health state of a chronic heart disease person. Our goal is to investigate data irregularities, associations and inter-dependencies as well as to detect abnormal patterns and provide interpretation of continuous data which can support an intelligent rule engine machine efficiently.

**Acknowledgment.** The presented work is taken place in the framework of the ARTEMIS JU Project CHIRON (ARTEMIS - 2009-1 100228 - http://www.chiron-project.eu).

## References

1. Li, C., Zheng, C., Tai, C.: Detection of ECG characteristic points using wavelet transforms. IEEE Transactions on Biomedical Engineering 42(1), 21–28 (1995)
2. Martinez, J.P., Almeida, R., Olmos, S., Rocha, A.P., Laguna, P.: A wavelet-based ECG delineator: evaluation on standard databases. IEEE Transactions on Biomedical Engineering 51(4), 570–581 (2004)
3. Choi, S.: Detection of valvular heart disorders using wavelet packet decomposition and support vector machine. Expert Systems with Applications 35(4), 1679–1687 (2008)
4. WEKA, http://www.cs.waikato.ac.nz/ml/weka/
5. PMML v4.0 Standard, http://www.dmg.org/v4-0-1/GeneralStructure.html

# Archetype-based solution to tele-monitor patients with chronic diseases

Juan Mario Rodríguez[1], Carlos Cavero Barca[1], John Gialelis[2], Petros Chondros[2], Dimitris Karadimas[2], Paolo Emilio Puddu[3], Kevin Keene[4] and Jan-Marc Verlinden[4]

[1]ATOS, ATOS Research and Innovation (ARI), Spain
[2] Industrial Systems Institute/RC Athena, Platani Patras, Greece
[3]Sapienza University of Rome, Department of Cardiological Sciences, Rome, Italy
[4]ZorgGemak BV, Voorschoten, Netherlands

*Abstract*— **Health tele-monitoring systems are aimed to improve chronic diseases treatment and reduce cost of care delivery. Behind this innovative and promising philosophy in the care of people with chronic diseases several benefits can be found: hospitalizations may be reduced, patients' quality of life improved and clinical outcomes considered more complete. As tele-monitoring includes measuring and collecting health information about individual patients, the evolving concept of Electronic Health Record (EHR) is crucial. Getting a shareable and universally accessible EHR is a challenge whose importance is contemplated by organizations that establish and manage standards. In the context of the CHIRON project[1], Congestive Heart Failure (CHF) patients are enrolled in an observational study to be tele-monitored by experienced doctors. Technical solutions have been designed to deal with EHR desirable features and visual requirements for remote visualization and study. An EN13606/openEHR-kernel is the core component dealing with multisource patient data making up a complete EHR system assuring semantic interoperability. EN13606 [1] and OpenEHR [2] follow the two-level modelling approach describing specifications of a reference model and archetypes to store, retrieve, exchange and manage health data in EHRs. Dynamic components to access, visualise and insert data into patients' records are shown through a doctor-friendly user interface. It has been called Slim MEST (Medical Expert Support Tool) and combines high flexibility and adaptability as it is built upon the same archetypes defined in openEHR-kernel. Functionality is extended by means of an ECG signal viewer application, which proofs versatility of data collected by sensors used in the tele-monitoring process while offers the clinicians a practical tool for their diagnosis.**

*Keywords*—**EHR, archetypes, tele-monitoring, parameter, openEHR.**

## I. INTRODUCTION

The use of communication technology to monitor patients and their health status is a focal point for improving chronic disease management. Hospitalizations can be reduced, patients' self-care is enhanced and outcomes are meant to be more valuable. Risk identification, particularly short- and very short-term, and multiple interactions among old and newly investigated risk indexes [3] may be an effective tool for patient-centric tele-monitoring [4] with the potential of fostering the continuum of care [5] in patients with chronic diseases.

To detect diseases in an early stage, before symptoms have occurred, the CHIRON project investigates patients with Congestive Heart Failure (CHF) both at home and in ambulatory conditions. Newly envisaged risk functions are applied to go beyond what has been done recently [6].

Provision of Electronic Health Record (EHR), as well as demographic information and data integration, are assessed following openEHR open standard specification **[**1**]**, which relies on archetypes as the standardized methodology to get reusability and scalability. Patient data is obtained from heterogeneous sources can be stored using the openEHR reference model [7].

The project equips the patients with unobtrusive sensors to collect multiple physiological parameters such as ECG, skin temperature, sweating index and activity patterns, as long as means to store other parameters, getting a complete list of 67 parameters which were found relevant for CHF patients, after a survey between cardiologists experts. Raw data collected by the sensors is made available for visualisation purposes by means of the ECG Viewer which access directly to the raw data. Only significant values are incorporated to the EHR, such as periodical averages. The tele-monitoring process generates raw data but only averages are copied to the EHR in order to give the doctors only the relevant CHF parameters. The tool also shows other EHR-based information coming from the Hospital Information System (HIS) such as medications, interventions, etc.. Slim MEST tool has been conceived to take advantage of the openEHR flexible dual model concept which separates out the reference model and the clinical knowledge (archetypes) modelled by doctors [9]. It encapsulates archetypes, data storage and retrieval and offers a doctor oriented interface which gives a complete overview of parameters trends and the health profile of the patient by accessing the EHR. Figure 1 shows components involved in the solution and their communication flows.

[1] www.chiron-project.eu

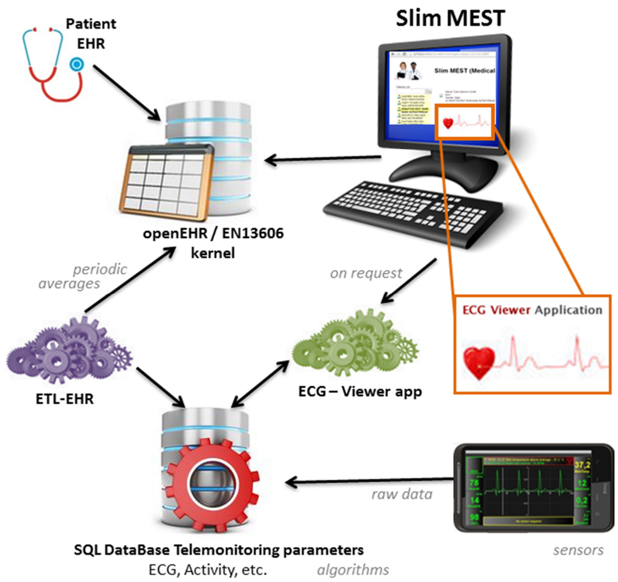



Fig. 1 - Components diagram.

The achievements of the medical objectives are assessed through a six months observational study started at April. Although number of patients and people involved is limited, it attempts to provide a proof of concept and useful conclusions about the potential benefits of this exposed solution.

## II. EHR INTEROPERABILITY

### A. Archetypes

To store (medical) parameters for the CHIRON project, the reference model EN13606 was used, and more specific the openEHR reference model. The use of EN13606 in eHealth systems is a proven solution for standardization [10]. The key concept of EN13606/openEHR is the so-called archetypes to model the medical knowledge providing interoperability, standardisation and computability regarding electronic health records. For every medical parameter archetypes were searched for in the openEHR clinical knowledge manager [11] which is a repository of archetypes already created and validated by community members are stored and shared. For the project these archetypes were clustered to the specific needs. Concepts were added for, by example, NYHA classification, daily activity, daily weight change and CHF etiology. More generable used parameters already existed in the openEHR knowledge manager, for example archetypes about ECG and lab tests.

After the creation of these archetypes they were approved by the project's cardiologist, which validated the workflow medically.

### B. Transforming raw data into EN13606 information

Because of the observational study, the raw sensor data had to be stored correctly in the EHR to be accessible for the doctors. Storing raw data in the archetype model is not useful for medical use in the context of electronic health records. The EHR consists mainly of summaries of observations and evaluations. Therefore a technical translation was done. For example to store the relevant information for an ECG, instead of storing every single point, a calculation was done with an algorithm to extract the useful parameters from the ECG concerning CHF.

## III. SLIM MEST

### A. openEHR configurator

Clinical concepts modelled as archetypes in openEHR are written in Archetype Definition Language (ADL) [7]. Due to the path/value nature of archetypes, a component to configure the clinical concepts and the corresponding archetypes was found useful and consequently developed.

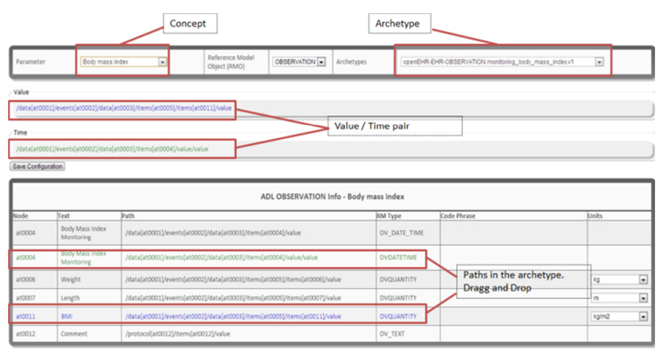



Fig. 2 - openEHR configurator.

This openEHR-configurator simplifies the associations between the parameters needed by the doctors and the archetypes stored in the openEHR kernel. Scalability and flexibility are key factors to allow different archetypes in the future, so the Graphical User Interface for this separate component is built dynamically according the one that is being considered. Figure 2 shows an example of concepts and archetypes.

Some initiatives have already been deployed with full archetype-based development [9-14] demonstrating at a vendor level the lack of dynamicity to constraint the archetype nodes, so the stakeholders continued to use their own software interfaces linking them to the dual model solution. Archetypes are defined as maximum datasets, therefore there is a need of selecting the information to be shown.

OpenEHR-configurator plays an important role in the automation of EHR accesses. It sets the properties which allow the correct communication between the client application and the openEHR-kernel. As a result, heterogeneous medical parameters can be treated uniformly through its API and client applications maintenance gets deeply simplified.

### B. Graphical User Interface

Interface for medical experts is provided through Slim MEST. This tool is a web based application that allows ubiquitous secure access to doctors, where they can visualise, update and analyse the data of their patients. The application accesses the EHR, located in openEHR-kernel, remotely and shows up data collected in the observation period plus some data coming from the EHR.

Slim MEST relies on archetypes that are included in the EHR. Following this strategy makes the tool specially robust and flexible to the variations introduced in the openEHR-kernel it accesses to. This allows also a simpler adaptation to another supposed openEHR system in any other context. A communication layer, configured by openEHR-configurator (Section III. A), makes remote calls independent from the tool and relates archetypes in the EHR with concepts that will let the GUI being built dynamically.

The Slim MEST provides the following features:

- *Visualise parameters health status*: the last patient health status (last sensor measurements and heterogeneous EHR values, as medications or previous diagnosis) is shown.

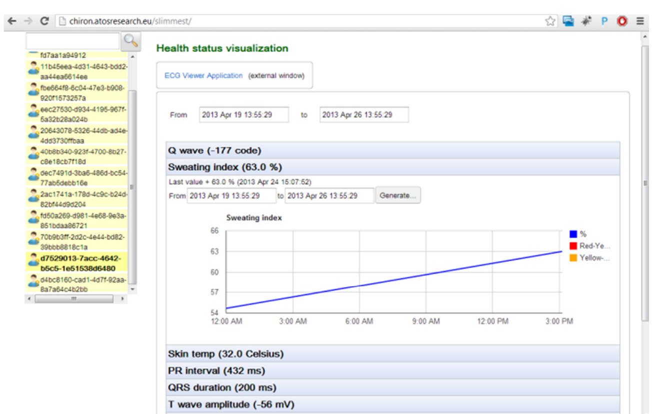



Fig. 3 - Slim MEST: Visualisation.

A self-expandable table is provided showing parameter name beside its last value and allows generating graphs between certain period of time of the numeric parameters illustrating the variation together with the stipulated thresholds. Diseases and medications are displayed indicating start and end of diagnoses and treatment. A printable report is available with this information.

- *Insert parameters health status*: this option brings out a paginated list of all available parameters in the system and allows introducing new measured or calculated values. As a reference, last value stored is shown beside the insertion area, accompanied by the date when the measurement was taken. Doctors guidelines have been widely taken into account in the design of the data update mechanism. As a result, new values are stored with the current date and time. A printable document that summarizes the actions taken can be obtained.

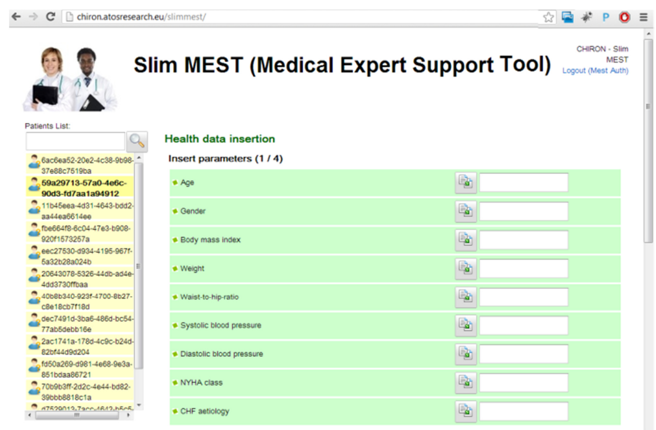



Fig. 4 - Slim MEST: Insertion

## IV. ECG VIEWER

As described in the previous section the overall objective of the proposed solution is the effective management of the patient data coming from various sources (demographic, monitored and clinical data) coming from the EHR; but the medical experts need to analyse the raw data collected during the observational study protocol.

Under this concept the present research also includes an ECG (Electrocardiogram) Viewer application aiming at real electrocardiogram signal visualization. The ECG Viewer allows the doctors to visualise a specific part of the real ECG signal when such a demand rises. This means that when medical experts receive a critical alarm for a specific patient from the MEST they are able to have access in the original ECG signal, in an appropriately visualized depiction, for that specific patient. The ECG Viewer has been set to provide a 30-secs, 60-secs or 90-secs ECG signal after the time of interest has been selected, at which the critical situation has been observed in the Slim MEST. The expert is able then to move back and forth in time, observing the ECG signal and make up his conclusions based on his knowledge. The ECG Viewer provides a user-friendly, easy and responsive environment with unlimited time-axis zoom capabilities, signal coordinates tracking and even exportation of the ECG signal in various formats.

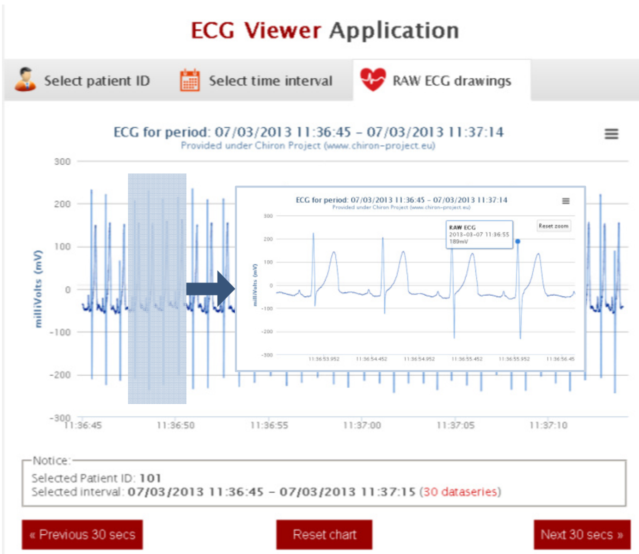



Fig. 5 - ECG Viewer and zoom feature demonstration.

ECG Viewer application has been developed to be presented within an Internet browser environment. Since the ECG signal is being sampled with a 500Hz frequency, the viewer has been set to provide a maximum of 90-secs visualisation. As the ECG signal length increases, it becomes more demanding to be visualised within a browser, due to the significant large amount of data that the both modules have to parse. The visualisation engine is based on jQuery framework and its extensions, thus providing maximum compatibility.

## V. CONCLUSIONS

Health is a topic that concerns everybody. The conjunction between health care and technology has always been a source of hope to improve quality of life. This paper studies in depth an EHR-based tele-monitoring implantation for CHF patients and presents the satisfactory solution that has been reached through a collaborative work. EN13606 / openEHR model provides many benefits as an EHR standard and creates an effective synergy with tools designed to ease the doctor participation and intervention in the care of the patients with chronic diseases converting the archetypes in the basis of the overall system. What makes this approach especially significant is that the tele-monitored patients and doctors get benefit from this progress due to its dynamicity and simplicity. Additional parameters do not require new releases; just archetype modelling and concepts assignment.

Future success in eHealth, tele-monitoring and standardisation resides in collaboration, intercommunication and investigation that becomes exciting when day by day idealistic forecasts become closer.

## Acknowledgment

This research was carried out in the CHIRON project (Cyclic and person-centric Health management: Integrated appRoach for hOme, mobile and clinical eNvironments) co-funded by the ARTEMIS Joint Undertaking (grant agreement #2009-1-100228) and by national authorities.

## References

1. CEN/TC251-ISO/TC215 (2010) Electronic Healthcare Record (EHR) Communication. Parts 1: Reference Model, Part 2: Archetype Model, Part 3: Reference Archetypes and Term lists, Part 4:Security and Part 5: Interface Specification.
2. Beale T., Heard S. (2008) Architecture overview. [Online] openEHR Foundation. [Cited: 24th April 2013] www.openehr.org/releases/1.0.2/architecture/overview.pdf.
3. Puddu PE, Morgan JM, Torromeo C et al. (2012) A clinical observational study in the Chiron Project: rationale and expected results. In Proceedings of ICOST, 2012. pp 74-82
4. Chaudhry SI, Mattera JA, Curtis JP et al. (2010) Telemonitoring in patients with heart failure. N Engl J Med 363: 2301-2309
5. Bonfiglio S. (2012) Fostering a continuum of care. In Proceedings of ICOST, 2012., pp 35-41
6. Klersy C, De Silvestri A, Gabutti G et al. (2009) A meta-analysis of remote monitoring of heart failure patients. J Am Coll Cardiol (2009) 54: 1683-1694
7. Beale T., Heard S., Kalra D., Lloyd D. (2007) The openEHR Reference Model - EHR Information Model. [Online] openEHR foundation. [Cited: 24th April 2013] www.openehr.org/releases/1.0.1/architecture/rm/ehr_im.pdf.
8. Beale, T (2013) Archetype Definition Language ADL 1.5, openEHR Foundation. [Online] [Cited 24th April 2013]: http://www.openehr.org/releases/trunk/architecture/am/adl1.5.pdf.
9. Beale T. (2002) Archetypes constraint-based domain models for future-proof in-formation systems. s.l. : OOPSLA-2002 Workshop on Behavioural Semantics, 2002.
10. Roberta Gazzarata, Jan-Marc Verlinden et al. (2012) The integration of e-health into the clinical workflow – Electronic Health Record and standardization efforts. Proceedings of ICOST, 2012. pp 107-115.
11. Clinical Knowledge Manager. (2010) [Online] OpenEHR Foundation. [Cited: 24th April 2013] http://www.openehr.org/ckm/.
12. Stroetman K. A., Artmann J., Stroetman V. N. et al. (2011) European countries on their journey towards national eHealth infrastructures. s.l. : eHealth Strategies Report,
13. Chen, R., G. Klein, et al. (2009) Archetype-based conversion of EHR content models: pilot experience with a regional EHR system. s.l. : BMC Medical Informatics and Decision Making, 9(1): 33..
14. Bernstein, K., I. Tvede, et al. (2009) Can openEHR archetypes be used in a national context? The Danish archetype proof-of-concept project. s.l. : Stud Health Technol Inform. 150: 147-151.

ACEEE

*Proc. of World Cong. on Multimedia and Computer Science*

# An ONS-based Architecture for Resolving RFID-enabled Objects in Collaborative Environments

J Gialelis, D Karadimas, P Chondros, E Polytarchos
Industrial Systems Institute/RC Athena, PlataniPatras, Greece
Athens University of Economics and Business, Athens, Greece
[1]gialelis@isi.gr, [2]karadimas@ieee.org, [3]chondros@isi.gr, [4]e.polytarchos@gmail.com

***Abstract:*** **Collaborative B2B and B2C environments provide enterprises with the necessary level of flexibility and efficiency in order to retain competitiveness under the increasingly turbulent e-business area. Web-Services are utilized by enterprises in order to integrate high and low level enterprise applications, thus providing a collaborative e-business environment without affecting inter- and intra-enterprise processes. Nevertheless, the above paradigm should be enhanced in order to comply with the Web-of-Things concept. This paper describes a sustainable approach towards the above requirement by employing ONS-based services able to provide targeted information regarding RFID-enabled physical objects (i.e. products) that are handled in a B2B/B2C collaborative environment.**

***Keywords:*** **collaborative b2b, supply chain collaboration, web services, RFID, ONS**

## I. Introduction

Business-to-Business and Business-to-Customer electronic transactions (B2B/B2C e-commerce) outfit the enterprises with the necessary tools for real-time interaction with their suppliers, dealers and customers. As a result, competition among businesses turns gradually from an enterprise-to-enterprise matter to a value chain-to-value chain one.

E-Business frameworks represent the state of the art in e-commerce; since framework-based B2B e-commerce effectively addresses the seamless linking of inter-enterprise (public) processes involving Enterprise layer's applications to those of business partners. These frameworks provide the necessary layers to achieve interoperability in e-commerce, since their interactions involve various business processes and business components, such as executable applications, systems, and their associated information [1].

In order to support the interchange of information among "public" modules, several standards and languages have been developed. EDI (Electronic Data Interchange), EDIINT (EDI over the Internet), ebXML (e-business eXtended Markup Language) and BizTalk are some generic standards addressing vertical as well as horizontal integration issues. Furthermore, standards such as RosettaNet, IOTP (Internet Open Trading Protocol), ICE (Information and Content Exchange) and OBI (Open Buying on the Internet) deal more with vertical market-oriented integration issues [2].

However, current framework-based developments consider inter- as well as intra-enterprise processes, involving information at the level of enterprise layer applications, without taking into consideration the tracking information of a physical object (i.e. a product) through its lifecycle.The work presented in this paper focuses in the challenging concept of an RFID-enabled value chain from the manufacturing industry up to the point of sales, by introducing an architecture that utilizes components of the EPC global network, such

*DOI: 02.WCMCS.2013.1.17*

as the. Object Naming Services (ONS) and the EPC Information Services (EPCIS), in order to support the Internet of Things concept. Our current add on enhances the architecture of e-business frameworks in order to involve all enterprise layers into e-commerce interactions, thus introducing an innovative collaborative business model which seamlessly integrates the inter-enterprise (public) with the intra-enterprise (private) processes.

## II. Collaborative E-Business Environments

Continuous Replenishment Planning (CRP) model is an example of a collaborative B2B model which can constitute the base for the development of an efficient inventory management function. Even though, CRP model has the potential to alter significantly production-scheduling methods and inventory management practices, so that enterprises enjoy the above-mentioned benefits, survey results do not keep up with these expectations [3]. This drawback is caused, mainly, due to CRP deployment methodologies. Such methodologies are based on point-to-point architectures and on the utilization of proprietary solutions/technologies which present very limited tolerance to system/applications alterations and minimum scalability to value network expansion.

In order to fully overcome these drawbacks and achieve both integration of heterogeneous systems and applications, at any layer in the enterprises they may reside and unlimited expansion of the value-network, this paper proposes an enhancement of an approach described in [4] and [5]. Fig. 1 depicts the corresponding architecture. Our approach was based on Internet ubiquity and the XML standard. In a nutshell, a hierarchical model was utilized based on [6], which comprises three layers: the Communication, the Content and the Business Process layer.

The technologies per layer are as follows. For the communication needs the SOA paradigm (WSDL. SOAP, etc.)was utilized by employing the use of Web-Services technology in order to open up enterprise systems, plant systems and production systems; thus making available their functionalities. For the Content layer which provides the means to describe and organize exchanged information in such a way that it can be understood and used by all the collaborating partners, product ontologies in order to achieve a uniform product terminology were utilized. For the Business Process layer which is concerned with the conversational interactions among services, RosettaNet Partner Interface Processes (PIPs) was the paradigm that was adopted.

Therefore, our proposed enhancement comprises the introduction of an additional layer called Object Naming Services (ONS) layer which resides above the business process layer and it is responsible for resolving the Electronic Product Code (EPC) of RFID tags to Internet addresses which correspond to specific services thus, fully supporting the expansion of the value-network.

## III. Resolving EPC of RFID-Enabled Physical Objects

In order to support an RFID-enabled value chain from the manufacturing industry [7]to the point of sales and beyond in a global scale, the proposed architecture utilizes the standards and specifications concerning the Electronic Product Code (EPC) published by the GS1 [8].

### *A. The EPC global*

The GS1 EPC global is a suite of standards and specifications that leverage the RFID technology in order to enable visibility and collaboration in the value-chain on an item level. An organizational entity that supports these standards will be able to manage, keep track, provide value added services and collaborate in a global scale with other entities that support the EPC global stack, for every EPC tagged item that has been detected within its premises.

- Tag data & tag data translation. These standards define EPC tag data, the different schemes of encoding and their translation to various forms of presentation.
- Tag protocol. These standards define the physical and logical requirements for the EPC RFID systems (tags and interrogators).
- Low level reader protocol. The low level reader protocol defines the interface between the RFID interrogators and their clients.
- Discovery configuration and initialization & reader management. These standards define the management interface between the RFID interrogators and Access controllers that includes the

discovery of other components within the network, the configuration of the interrogators and the monitoring of their status.

- Application level events. The application level events standard specifies the filtering and data collection interface that provides tag data acquired from a interrogators.
- EPC information services. This standard provides a standard interface to provide services related with the EPC tag data.
- Object naming service. The Object Naming Service (ONS) is a crucial component of the EPC global framework, essential for the discovery of services related with an EPC tagged object. [9]. Concisely, it describes a mechanism, which utilizes the DNS infrastructure to translate the EPC encoded on the RFID tag of a product to a collection of services (EPC information services) and their corresponding addresses. The EPCs are structured using any one of the various numbering schemes available in the Tag Data Standard (TDS) [10]. Most of the schemes have the general structure depicted in Fig. 2. The ONS uses the EPC Manager Number (company prefix) and the Object Class components of an EPC to search through the DNS for appropriate services and it returns the list of services along their corresponding addresses. These services are in essence links to their formalized description (ServiceType XML) that can enable a machine to automatically determine how to interact with the service located at that address [9].
- Discovery services. These standards are still in development, but they will enable trading partners in the value chain and other entities to share resources, knowledge and data related to EPC tagged objects. Concisely, they will provide a standardized interface for an interested party to learn which EPCIS servers provide information for an EPC tagged object.
- Certificate profile. This standard specifies the digital certificates used to securely identify RFID interrogators.

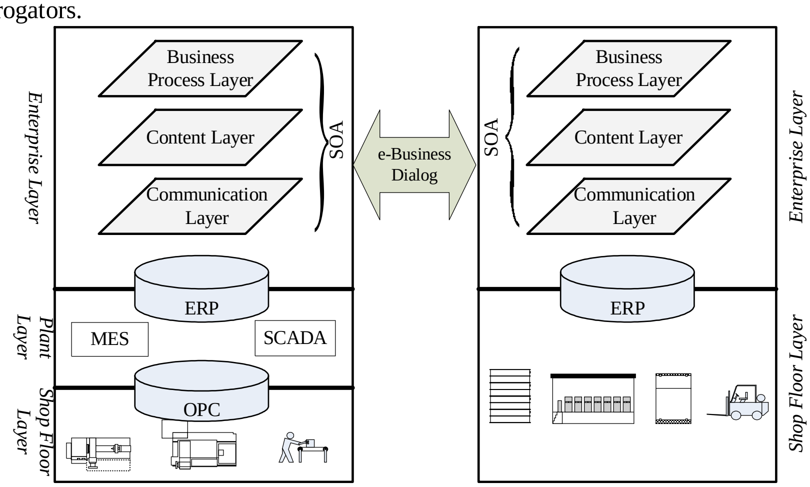



Figure 1. e-Business holistic framework.

## III. Enhanced Architecture Framework

The proposed enhancement of the existing e-Business Framework aims to support as many of the EPC global standards as possible; in order to provide the ability to resolve single physical objects to URIs and to track related information regarding the entire lifecycle of the representation for all involved enterprises in the value chain, i.e. manufacturer, logistics, retailand end-user (customer), together with the realization of an ONS-based web service available in the cloud.

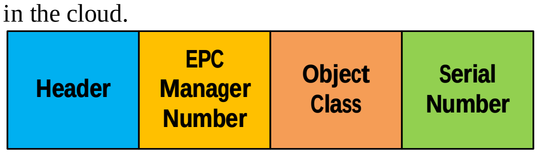



Figure 2. Common structure of TDS schemes.

Fig. 3illustrates the incorporation of such an ONS service layer at the top of each existing CRP model. The proposed ONS service layer consists of a collection of ONS-based web services that are able to serve information from the internal of the CRP model breakdown, using the global EPC notation, without affecting the existing e-Business dialogs among the enterprises of the value chain. Each enterprise's ONS service layer is responsible for providing per object, both public and private information, using the global EPC notation.
The private web services satisfy per sector- needs for real-time, synchronous physical objects tracking. These needs have various orientations, depending on the corresponding node of the chain. The public web services address needs of purchasers, regarding single product's lifecycle information

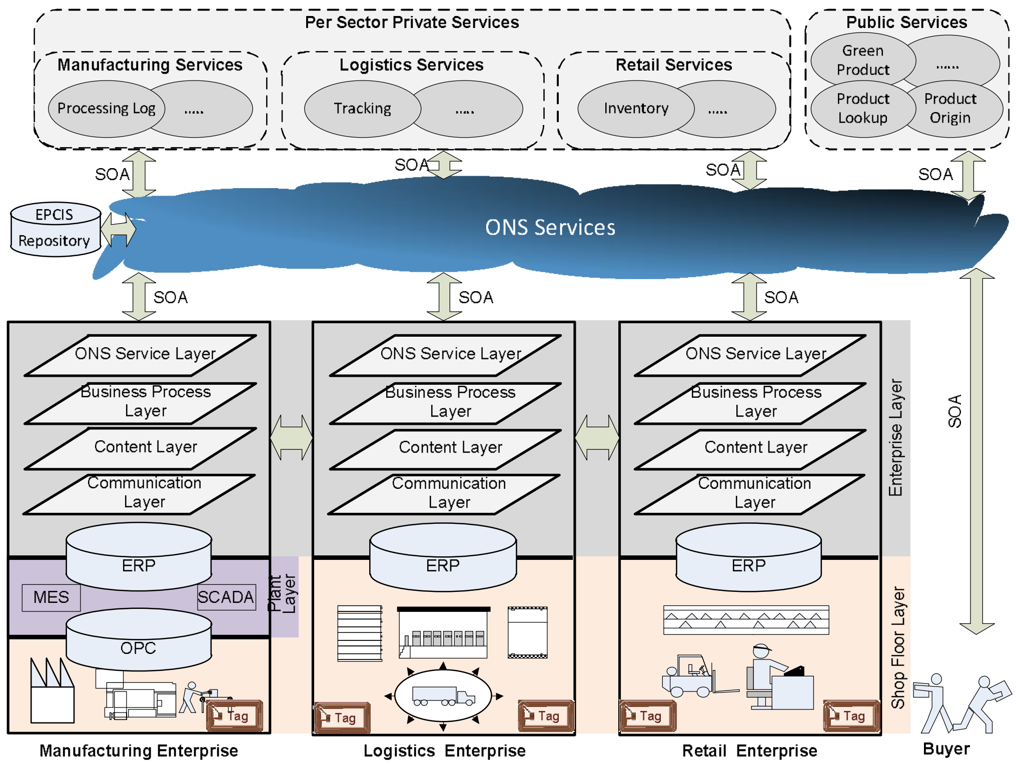



Figure 3. Proposed architecture framework

The integration of an RFID subsystem, into an existing architecture framework raised two main issues. The first regards the capturing of tags' information as well as the identification of the tags themselves. The RFID reader scaling, range and reading angle are of major importance since it is incorporated in an industrial environment. The second issue regards the employment and integration of the protocol stack into the CRP model. In the following section, the ONS service layer is described.

### *A. The ONS service layer*

Since the main focus of the architecture presented in this paper is the collaboration within B2B/B2C environments, one of the key motivators for integrating an EPC RFID system in the architecture would be collaboration itself, which, within the EPC global framework, is mainly represented by the Object Naming Service, that provides interested parties with the ability to resolve EPC tagged objects to addresses of arbitraryservices, but with a standardized interface.
Adhering to the well-defined and widely supported SOA paradigm of SOAP Web-Services, we have defined services that provide the ONS functionality to EPC-agnostic web service clients. Thus, if a client wants to discover the available services related to an EPC tagged item, both the required by the ONS standard NAPTR DNS query and a SOAP web service call will be available.
The architecture also embraces another key component of the EPC global, the EPCIS [11], in order to provide value added services that include traceability (i.e. a history of the item that includes the locations

where it has been detected and the corresponding business events) and a mapping service that acts as a translator of legacy bar-code based application dependent codes to EPCs.

Additionally, in order to simplify the development of the various components of the architecture, the ONS service layer has been designed so as to provide web services that provide a uniform interface to the traceability service.Thereby, whenever the system gets input from an RFID reader and uses the ONS service layer to detect the network address of an EPCIS server, a corresponding entry is added in the traceability service that contains:

- the EPC oftheitem,
- thelocationofthedetection (whichalsoindicatesthebusinessevent),
- theservicerequested,
- thedateoftherequest.

The next section presents a small subset of possible services that have been employed as private or public services, illustrating the usage and large scale capabilities of the proposed architecture.

*B. Use case scenario*

The main use case scenario illustrated in this paragraph utilizes the enhanced architecture framework as described previously, premises that discrete product manufacturing enterprises are involved in the chain.

Additionally, the enterprise of primal production within the chain should own a company prefix (Global Trade Item Number - GTIN) so as their products are not only RFID-capable but also the corresponding EPC tags have global visibility and uniqueness. The above requirements among with the proposed architecture framework constitute the integrated, collaborative environment that is able to deliver significant benefits to all involved parts in the chain, i.e. the enterprises and the consumers.

An integral substitution of the existing e-Business dialog in a stable chain is not an option in large scale enterprises, since huge investments and efforts have been achieved for such an accomplishment. The accomplished benefits from such an integrated environment for a contemporary B2B and B2C trading scheme, aim to complement the existing e-Business dialog of collaborative B2B environments so as to include ONS-obtained functionality, thus the single object resolving capability. However, many critical or just informational services could be deployed on such an architectural framework for a B2B/B2C environment.

The critical under-tracking information within a manufacturing enterprise is the processing log and tracking of a single product. This is more meaningful if the manufacturing enterprise fabricates large objects, following a build-to-order approach.

Logistics and retail enterprises could benefit the most from such an e-Business framework since the ability of a single object tracking is of great importance for them. Assuming that all cooperating partners of a logistic and a retail enterprise follow the proposed architecture, then all involving parts are capable to track a single product not only within their facilities but worldwide.

Finally, consumers are able to track provided information from a single product's lifecycle, assuming that the points of sales are equipped with RFID readers. The real-time and synchronous tracking of all relevant information regarding a product will provide next generation informational services regarding the consumable products. These services may include product energy characterization, product designation of origin, the obvious information about manufacturing and packaging dates, even information regarding the intermediate steps of manufacturing or supply chain.

## V. ARCHITECTURE'S HOLISTIC SERVICE LAYER

The implementation of the proposed service layer as an enhanced architecture scheme for collaborative environments is based on two major aspects: the XML structures that are handled by the ONS service layer of each CRP Model and also the operations that this layer is capable of, which are both described in this section.

*A. Structures used by the ONS service layer*

The developed ONS service layer uses the XML structure described below in order to represent the pairs of information services and their respective internet addresses (Pairs):

- Pair

  - Service (The definition of the service, according to [9])
  - Servers (The URI of the server that provides the service that is contained in the Pair e.g. "https://location.randomserver.org/locate")

The following structure represents the Pairs that are related with an EPC:

- EUSSN (the name is an acronym derived from EPC, URI, (Service-Server)*N) contains the fields:
  - EPC (an EPC writtenas a string according to [10])
  - URI (The URI that results from the conversion of the EPC according to[10])
  - Location (A string that describes the location where the detection took place)
  - Time Stamp (The time the item was detected, in UTC)
  - A number of Pairs

The structure below is used in the operations that require authentication and authorization in order to function. It is comprised by a user, an encrypted password and a number (0 or more) of Pairs that are related with an EPC (EUSSN).

- UEUSSN (User + EUSSN)
  - User (a username)
  - Pass (an encrypted password)
  - EUSSN (zero or more)

The Location Time structure is used to specify the location and the date that an item was detected

- Location Time:
  - Description (A string that describes the location of the detection e.g. "storehouse A")
  - Timestamp (The detection time - UTC)

Finally, the structure below represents a list of the locations where the item has been detected:

- Locations:
  - A number of LocationTime entries

*B. Service Operations*

Based on the previously described structure, the implemented ONS service layer supports the operations described in this section: EPC query, EPC service Query, EPC2URI, EPC addition, EPC deletion, User Addition, User Deletion and Location Query.
Below, the detailed description of the operations can be found.

*EPC query:*

The EPC query returns the Pairs located for the specified EPC or URI (hereby mentioned collectively as EPC), included in an EUSSN structure.
This, EPC2URI and EPC service Query services do not require authentication. The EPC query operation accepts a single argument:

- EUSSN (A EUSSN node that does not contain any Pairs. In it, either the EPC or the URI or both should be mentioned (if both are mentioned; only the URI will be considered). If a location is not mentioned, the location is not stored into the tracking service. If the Time Stamp field exists, it is used for the detection time in the tracking service. If that is omitted, the query arrival time is used.)

The operation returns a EUSSN node that contains:

- The EPC for which the query was for. If the request did not contain an EPC, but only a URI, the EPC field will contain the corresponding translated value of the URI, according to[10]
- The URI representation of the EPC. Conversely with the EPC, if a URI hasn't been contained in the request, the URI field of the response will contain the result of the EPC to URI conversion.
- The previous location where the item was detected, along with the date of that detection

- A number of service-server pairs that will contain the pairs of service tags and server addresses for the EPC-URI contained in the EUSSN

*EPC Service Query*

The EPC service query returns the Pairs located for the specified EPC or URI (hereby mentioned collectively as EPC) and service, included in a EUSSN structure.

This operation does not require authentication. The operation accepts a single argument:

- EUSSN (An EUSSN node. In it, either the EPC or the URI or both should be mentioned (if both are mentioned, only the URI is considered) and a Pair that should contain only the desired service name. Optionally, the query can also include a Location and a TimeStamp.)

The operation returns a EUSSN node that contains:

- The EPC for which the query was for. If the request did not contain an EPC, but only a URI, the EPC field will contain the correspondingtranslatedvalueofthe URI, accordingto[10]
- The URI representation of the EPC. Conversely with the EPC, if a URI hasn't been contained in the request, the URI field of the response will contain the result of the EPC to URI conversion.
- The previous location where the item was detected, along with the date of that detection.
- A number of service-server pairs that will contain the pairs of service tags and server addresses for the EPC-URI contained in the EUSSN that matched the service in the request.

*EPC 2URI*

The EPC to URI returnsthe URI ofthetranslationaccordingto[10]ofthe EPC provided. This operation does not require authentication. The operation accepts a single argument:

- EPC (The EPC that should be translated.)

The operation returns a string that contains the resulting URI. If the EPC could not be translated, a zero string is returned.

*EPC addition*

This action adds a Pair for the specified EPC (the pair and the EPC are provided by an EUSSN structure).

If the EPC isn't already contained in the system, a new entry for that EPC will be created. A valid user with sufficient permissions is required. Along with this operation the user must pass his encrypted password, in order for the server to verify whether the user is authorized to add a service-server pair for an EPC. The required arguments are:

- EPC additionRequest (A UEUSSN node that contains the credentials of the user and an EUSSN node that must contain the EPC and the Pair that should be added in the ONS. The EUSSN must contain only one pair.)

This operation returns an integer. If the new service was successfully added to the ONS, the operation returns 0, else if the Pair already exists 1 is returned. If the Pair could not be added -1 is returned. If the user doesn't have sufficient permissions to add an EPC, 2 is returned.

*EPC deletion*

The provided Pair is removed for the specified EPC (the data are provided using an EUSSN structure).

If no Service-server pair has been provided, all the records for the in the local ONS server will be deleted.

A valid user with sufficient permissions is required. Along with this operation the user must pass his encrypted password, in order for the server to verify whether the user is authorized to delete a service-server pair for an EPC. The required arguments are:

- EPC deletion Request (A UEUSSN node that contains the credentials of the user and an EUSSN node that must contain the EPC and the service-server pair that should be removed from the ONS. The EUSSN must contain only one pair.)

This operation returns an integer. If the service has been successfully removed from the ONS, the operation returns 0, else if the service-server pair doesn't exist 1 is returned. If the service-server pair could not be deleted -1 is returned. If the user isn't permitted to delete an EPC, code 2 is returned.

*User Addition*

This operation creates a user with certain permissions. It must be executed by an authorized user. The required arguments are:

- User (A UEUSSN node that must contain the credentials of a user with the appropriate user management rights. No EPCs or Pairs should be contained.)
- Added User (A UEUSSN node with the credentials of the user that will be created. No EPCs or Pairs should be contained.)
- Permissions
    - 3bits: *add EPC* (If the first bit is set - i.e. Permissions Value | 1 -, the user will be able to add EPCs), *delete EPC* (If the second bit is set - i.e. Permissions Value | 2 -, the user will be able to delete EPCs) and *user management* (If the third bit is set - i.e. Permissions Value | 4 -, the user will be able to add or delete users.)

The operation returns an integer: 0 for success, 1 if the user already exists, 2 for insufficient permissions, -1 for general failure.

*User Deletion*

This operation removes a user. It must be executed by an authorized user. The required arguments are:

- User (A UEUSSN node that must contain the credentials of a user with the appropriate user management rights. No EPCs or Pairs should be contained.)
- Deleted User (The username of the user that will be deleted. This must be different than the user that requested the deletion.)

The operation returns an integer: 0 for success, 1 if the user doesn't exist, 2 for insufficient permissions, -1 for general failure.

*Location Query*

The Location Query service returns the locations where the EPC provided as an argument has been detected.

This service does not require a valid user and requires a single argument, the EUSSN structure that will provide the EPC of the item:

EUSSN (An EUSSN node that does not contain any service-server pairs. In it, either the EPC or the URI or both should be mentioned - if both are mentioned; only the URI is considered.)

The Location Query call returns a list of the locations where the item has been detected, using a Locations structure.

## VI. Experimental Implementation

The experimental implementation of the ONS service layer employs the bind 9.8.4 DNS server for the ONS instance, as specified in [9] and the Tomcat Java 6.0.35 applet container to deploy the web services, supported by the PostgreSQL 9.1 database and the OpenJDK Java platform. The whole ONS service layer has been included in a virtual machine running the debian 7.1 linux distribution, in order to streamline installation and ease deployment.

The validation and verification of the functionality of the system included the creation two such ONS service nodes in different physical servers. Each instance contained a user that could add and delete EPCs, as described in section V. Then, their web services were used to add a set of services (location tracking and legacy code mapping) for a number of dummy EPCs. Finally, the system was queried for these services both through the web service interface and via DNS queries.

## VII. Conclusion

This paper presents an innovative approach towards collaborative e-business environments by utilizing Rosetta Net standard, Service Oriented Architecture and semantically enriched information combined with Object Name Services in order to seamlessly integrate the inter-enterprise (public) and intra-enterprise (private) processes in an ONS-based perspective.

The main focus of the architecture presented in this paper is the interoperable integration of B2B/B2C environments, which constitutes one of the key motivators for involving an EPC RFID system in the architecture. The integration itself, which, within the EPC global framework, is mainly achieved by utilizingthe Object Naming Service, provides collaborating parties with the ability to resolve EPC tagged objects to addresses of arbitrary services, but with a standardized manner.

## Acknowledgment

The work presented in this paper has been partially supported by the General Secretariat for Research and Technology (GSRT) [12] of the Hellenic Ministry of Development in the framework of project SELIDA - contract # 09SYN-72-646 which runs under the "Cooperation", 2009 Call.

## References

[1] Simon S.Y. Shim, Vishnu S. Pendyala, Meera Sundaram, Jerry Z. Gao: Business-to-Business E-Commerce Frameworks. IEEE Computer, vol. 33, no.10, pp. 40--47, (October 2000)

[2] Kim Tae-Young Kim, Sunjae Lee, Kwangsoo Kim, Cheol-Han Kim: A modeling framework for agile and interoperable virtual enterprises. Computers in Industry, vol. 57, issue 3, pp. 204--217. Elsevier Science Publishers (April 2006)

[3] S. Lee et al.: Business Value of B2B Electronic Commerce. Electronic Commerce Research and Applications 2, pp. 350--361 (2003)

[4] J. Gialelis, A. Kalogeras, A. Kaklis, S. Koubias: RosettaNet-based implementation for a collaborative continuous replenishment planning model utilizing Web services and Ontologies. IEEE Emerging Technologies and Factory Automation, ETFA 2005.

[5] J. Gialelis, A. Kalogeras, A. Kaklis, S. Koubias: Collaborative Continuous Replenishment Planning Process Implementation Utilizing Web Services, Product Ontologies and Workflow Management Systems. IEEE Emerging Technologies and Factory Automation ETFA 2006.

[6] B. Medjahed et al.: Business-to-Business Interactions: issues and enabling technologies. The VLDB Journal 12, pp. 59--85 (2003)

[7] Dongmyung Lee and Jinwoo Park (2010). RFID-enabled Supply Chain Traceability: Existing Methods, Applications and Challenges, Sustainable Radio Frequency Identification Solutions, Cristina Turcu (Ed.), ISBN: 978-953-7619-74-9

[8] GS1, EPC global network specifications, http://www.gs1.org/gsmp/kc/epcglobal/

[9] GS1, Object Name Service (ONS) Version 2.0, Ratified Standard, Issue 1, 2012/12/20

[10] GS1, EPC Tag Data Standard 1.6, Ratified Standard, September, 2011

[11] GS1, EPC Information Services (EPCIS) Version 1.0.1 Specification, 2007/09/21

[12] General Secretariat for Research and Technology, http://www.gsrt.gr

# END-TO-END SYSTEM FOR THE RELATIONSHIP MODELLING OF PHYSICAL EXERCISE WITH SLEEP QUALITY

**ABSTRACT**

In this paper we are presenting the ongoing work of an end-to-end system that we claim it will improve the quality of sleep in people with sleep disorders through physical exercise, in a convenient and friendly for the user way. In the proposed system we incorporate various sensing devices that measure physiological parameters that the literature manifests to be strongly related with exercise and sleeping activity. The data are collected by a mobile device which handles the heterogeneity among the devices and provides secure and reliable transmission to a back-end system where the data are processed and permanently stored. Upon conclusion of the ongoing work, the system will provide the necessary modeling and relation among the exercise and sleep that will benefit people with sleep disorders.

**KEYWORDS**

Sleep disorder, body area network, artificial intelligence, wearable devices

## 1. INTRODUCTION

Chronic insomnia is a sleep disorder characterized by long-term difficulties with initiating and maintaining sleep, waking up too early, or nonrestorative sleep. It is often associated with occupational and psychosocial impairments such as daytime fatigue, mood disturbances, cognitive deficits, and poor quality of life. Physical activity and exercise has been proposed in the literature as a non-drug treatment alternative with low cost and easy access. Exercise is assumed to affect the subsequent night of sleep through alternations of physiologic systems. The tool that has dominated as the gold standard for recording the sleep activity and assessing the quality of sleep so far is polysomnography. The polysomnogram records several body functions including brain activity, eye movement, oxygen and carbon dioxide blood levels, heart rate rhythmbody muscle activity, air flow through the nose, etc. However, despite the detailed recording that this tool provides it is quite resource demanding as it needs to be performed in-lab by specialized personnel. At the same time it becomes inconvenient for the patient as he needs to spend one – or more – nights sleeping in the lab.

At the moment, application algorithms aiming at the support of the experts in improving the sleep quality of insomniac and healthy individuals aren't used in large scales. However, an application constructing an exercise schedule suitable for the age, gender, BMI, physical condition of an individual, arranging time and duration of the exercise, reminding simple advice such as the appropriate time for exercise or the last meal, could be very useful to this direction.

Therefore an automated solution that addresses the relation of physical exercise and sleep disorders in a non-invasive and cost-effective way may address the sleep disorders symptoms and consequently improve the user's quality of life. From this perspective in this work in progress we describe an end – to – end system for recording parameters that evaluate the exercise and sleep activity where algorithms could be implemented that manifest how the physical exercise is related to the sleep quality.

## 2. RELATED WORK AND STATE OF THE ART

Research studies present evidence that exercise or physical activity can have a beneficial effect on sleep quality. Systematic reviews and meta-analyses have shown that many aerobic or resistance exercises can act as a supported intervention for insomnia and improve quality of sleep in middle-aged and elderly with sleep complaints [Kline CE et al][ Yang PY et al]. In general, the authors of these studies do not describe what type of physical activity is most prevalent or the intensity and duration of exercise. The generalization of the results is also limited by the small sample size or the fact the many subjects are good or young sleepers as well as the type of the exercise which may be unsuitable for the elderly [Montgomery P. & Dennis J]. Ohayon et al, determined normative sleep data sleep and daily activities for over 1000 participants aged 60 years or older. He found that loss of autonomy in daily activities is associated with sleep factors including sleep duration (long vs short), wake-up time (early vs late) and bedtime (early vs late). The sleep data were collected through telephone interviews and evaluated based on the Sleep-EVAL System.

The methodology for assessing sleep quality is often based on self-evaluated reports employing questionnaires, the most popular being the Pittsburg Sleep Quality Index (PSQI), a widely used measure suitable for large-scale epidemiologic investigations [Buysse et al]. The alternative dominant methodology is polysomnography (PSG), a multi-parametric study where body functions are monitored including brain (EEG), eye movements (EOG), muscle activity (EMG) and heart rhythm (ECG).

Passos et al, studied the effects of three different types of exercise (moderate aerobic/walking, intense aerobic/running, moderate resistance/weight training) on sleep quality. Both exercise and sleep data were obtained in controlled environment. PSG data demonstrated changes in sleep onset latency (reduction by 55%) and in total wake time (reduction 30%). However, on the polysomnogram inspection, it is necessary to stick many electrodes and sensors and the subjects need to be in a laboratory setting in order to ensure reliability of results. The study in [T. Igasaki, et al] researched ways of identifying sleep quality based on hypnogram which can be obtained through an EEG. They employed Artificial Neural Network technology which resulted in 71% accuracy.

Despite the growing consensus that regular exercise can improve sleep it remains unclear how this relationship is modeled. Work in [Buman MP et al] suggests that physically inactive older adults with poor initial sleep quality are most likely to receive sleep benefits from moderate-intensity physical activity. The authors observed that a 12-month intervention of moderate aerobic exercise had antidepressant effects that improved sleep quality. Their results showed that factors like exercise intensity and even the time of the day the workout is taking place are important in modeling the relationship between exercise and sleep quality. Moreover, recent studies suggest the potential of reciprocal relationships between exercise and sleep. Specifically sleep quality also predicted subsequent exercise behavior [Dzierzewski JM et al]. Similar results observed in [Baron, K.G et al], where in case of the 16 volunteers slept poorly, their workout the next day were significantly shorter. On the other hand the participant’s insomnia improved but only after the duration of 16 weeks of the intervention.

The above mentioned studies suggest that exercise and physical activity is strongly related to sleep quality. The proposed system will follow a hybrid methodology regarding the assessment of sleep quality, i.e both self-reported and objectively measured based on EEG alone. The main benefit is that the technology employed, can be used from the user/patient in his living environment without the requirement of being physically present in laboratory.

## 3. SYSTEM ARCHITECTURE

In this section we present the technological solution and the architecture design that allows the processing and integration of data obtained from the various medical sensors, in order to offer the appropriate services within a comprehensive integrated system.

The design of our system was driven by the requirements the application introduces, such as:

- Flexibility and Interoperability
- Security of the data, privacy and effective access policies
- Service-level and communication-level interoperability.
- User friendly non-invasive and richness of the experience for the users

A holistic view of the system’s architecture is presented in figure 1, while the system’s three sub-layers that it is consisted of and their functionalities are listed in Table 1. The various types of vital parameters that we needed to record in order to perform an accurate assessment of the exercise intensity and the sleep quality required respective sensor boards in the bottom layer of the wearable devices. The products of the two sub-layers (Wearable and mobile devices) that fit best to the application’s requirements and design’s principles are listed in table 2.

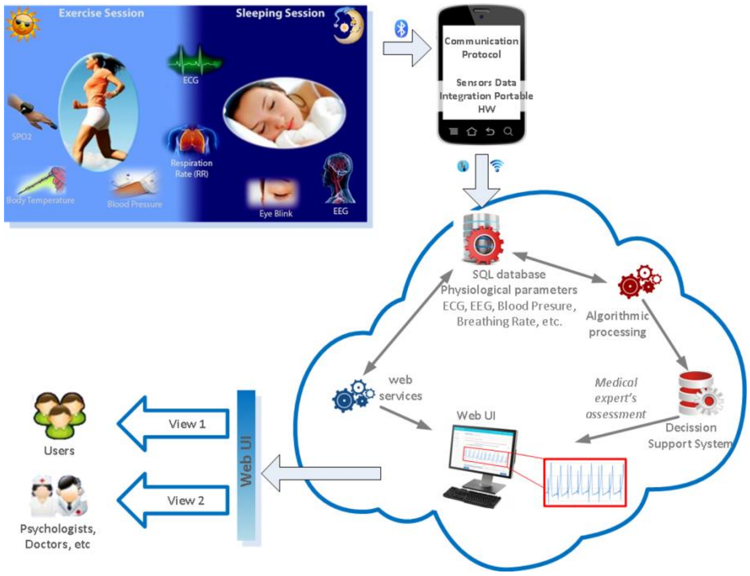



Figure 1: System Architecture

Table 1 . Systems sub-layers description

| **Sub Layer** | **Functionalities** |
|---|---|
| Wearable Devices | Data gathering<br>Signal processing algorithms for signal denoising, artifact removal and feature extraction<br>Data transmission (Bluetooth) |
| Mobile device | Data collection processing and storage (at mobile device level)<br>Data transmission to the Backend data repository (dynamically employ physical interfaces such as Bluetooth, Wi-Fi or Broadband communication) with JSON based messages<br>Communication security provided by OpenSSL. |
| Backend System | Data repository and Information Interchange<br>Algorithmic analysis and medical knowledge extraction<br>Incorporates a short health profile for each anonymous user handled by the system, connected with user's mobile device<br>Provides medical experts the capability to have a quick reference to relevant medical metrics and user's habits in order to deploy personalized training scenarios for each user.<br>User Interfaces |

Table 2 Sensor Boards

| **Sensor Board** | **Application Scenario** | **Features** |
|---|---|---|
| Zephyr BioHarness | Sleep/Exercise Sensor Board | Encapsulates 2 conductive ECG sensor pads and 1 internal breathing sensor location.<br>Battery Lifetime: 26 hours<br>Bluetooth connectivity |
| Nonin - 4100 pulse oximiter +<br>Nonin - 8000AA reusable SpO2 sensor (Articulated Internal Spring finger clip) | Exercise Sensor Board | Finger clip Sensor<br>Battery Lifetime: 120 hours of continuous operation with new batteries<br>Bluetooth Connectivity |
| A&D Medical - UA-767BT Automatic Blood Pressure monitor | Exercise Sensor Board | Internal Storage Capabilities storing up to 40 measuments.<br>Bluetooth Connectivity |
| MyndPlay BrainBandXL | Sleep Sensor Board | Head strap with two dry sensors attached, one ear clip as a reference point<br>Neurosky proprietary algorithms.<br>Bluetooth Connectivity |

| | | |
|---|---|---|
| Samsung Galaxy Nexus | Sleep/Exercise Mobile Device | OS: Android Jelly Bean<br>Bluetooth, 3G, WiFi connectivity<br>Internal Storage: 16GB enough for temporary storing days of records. |

## 4. DISCUSSION

The proposed architecture is the basis for modeling the relationship between physical activity and sleep quality. The system's functionality supports non-invasive monitoring of exercise and sleep of older people in their living environment allowing the medical expert to provide personalized intervention and feedbacks.

Biosignals, such as the ECG and EEG are undergoing algorithmic processing in frequency and time domain in order to extract features such as breathing rate, heart pulse rate, QRS complex duration from ECG and additionally non-REM brainwaves such as alpha, theta, delta, and gamma extracted from EEG with frequency domain analysis utilizing Fast Fourier Transforms. These algorithms are integrated in the backend-system.

Furthermore, current work comprises the extension of the functionality of the proposed system by employing artificial intelligence algorithms in the back-end towards medical knowledge extraction. The focus is on detecting semantically enriched events from EEG processing, which can be correlated with the monitored physiological parameters and questionnaires. Under this context, a predictive analytics approach is being followed to address specific classification problems, i.e. important EEG events, towards a rule-based end-to-end system revealing the dynamics of exercise and sleep quality.

The system's architecture is being tested in real environment by employing elderly people with sleep complaints. Preliminary results confirm system's reliability and operational capacity.

## ACKNOWLEDGEMENT

This work is being executed in the framework of a specific program funded by ERANET/LEADERA.

## REFERENCES

Baron, K.G et al, 2013.Exercise to improve sleep in insomnia: exploration of the bidirectional effects. Journal of Clinical Sleep Medicine, 15;9(8):819-24.

Buman MP et al, 2012. Moderators and mediators of exercise-induced objective sleep improvements in midlife and older adults with sleep complaints. Health Psychology, 30(5):579-87

Buysse DJ et al,1989. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 28:193-213.

Dzierzewski JM et al, 2014. Exercise and sleep in community-dwelling older adults: evidence for a reciprocal relationship. *Journal of Sleep Research*,.23(1):61–8

Kline CE et al, 2011. The effect of exercise training on obstructive sleep apnea and sleep quality: a randomized controlled trial. Sleep ,34(12):1631–1640.

Montgomery P. & Dennis J, 2004. A systematic review of non-pharmacological therapies for sleep problems in later life. *Sleep Medicine Reviews*, 8: 47–62.

Ohayon MM & Vecchierini MF,2005. Normative sleep data, cognitive function and daily living activities in older adults in the community. Sleep, 28(8):981–989

Passos GS et al, 2010. Effect of acute physical exercise on patients with chronic primary insomnia. Journal of Clinical Sleep Medicine, 6(3):270-5.

T. Igasaki, et al.,2012. Identification of sleeping quality by EEGs in humans. Proceedings of International Conference on Complex Medical Engineering, pp. 535–539, 2012.

Yang PY et al, 2012. Exercise training improves sleep quality in middle-aged and older adults with sleep problems: a systematic review. J Physiother, 58(3):157-163.

# Information System Framework Architecture for Organization Agnostic Logistics Utilizing Standardized IoT Technologies

Dimitris Karadimas
Industrial Systems Institute/RC Athena, Platani Patras, Greece
Email: karadimas@ieee.org

Elias Polytarchos
Athens University of Economics and Business, Athens, Greece
Email: e.polytarchos@gmail.com

Kyriakos Stefanidis
Library & Information Center, University of Patras, Greece
Email: stefanidis@ece.upatras.gr

John Gialelis
Industrial Systems Institute/RC Athena, Platani Patras, Greece
Email: gialelis@isi.gr

**Abstract—** Logistics or supply-chain services provide enterprises and organizations with the necessary level of flexibility and efficiency in order to retain competitiveness under the increasingly turbulent e-business area. Web-Services are utilized by organizations in order to integrate high and low level applications, thus providing a collaborative environment without affecting inter- and intra-enterprise processes. Nevertheless, the above context should be enhanced in order to comply with the Web-of-Things concept. This paper describes a sustainable approach towards the above requirement by employing ONS based services able to provide targeted information regarding RFID-enabled physical objects that are handled in an organization agnostic collaborative environment.

## I. Introduction

RADIO Frequency Identification (RFID) technology has already delivered revolutionary aspects in various areas such as logistics (supply chains), e-health management and materials identification / traceability. RFID technology itself allows an object's identification with effectiveness and efficiency. However traceability of an object calls for a robust and reliable system operating seamlessly over its entire lifecycle. Such a traceability system has to be implemented so as: a) its data model allows unique object identification and scalable, often big-data, databases, b) its underlying framework supports interoperability and c) its mechanism is capable to achieve end-to-end tracing providing full history information.

Despite RFID technology's nature in tracking, there are several challenges that need to be addressed. Since an RFID tag can be read from a quite long distance without requiring line-of-sight, it is possible that collisions may occur whereas also multiple tags could be read simultaneously. Therefore, there is no guarantee that a single tag will be consecutively detected on consecutive scans. Moreover, the use of RFID may constitute a serious threat for the information privacy, as it could be easily facilitated to espionage or unauthorized requests.

Based on the previously described situation there is a real need for an underlying framework able not only to support and complement the tracing functionality offered by the RFID technology but also to take into consideration the relevant tracking information of a physical object through its entire lifecycle in a secure and effectively protected way.

The work presented in this paper focuses in the challenging concept of an RFID-enabled, organization unaware, logistics management, by introducing an architecture that utilizes components of the Electronic Product Code (EPC) global network, such as the Object Naming Services (ONS) and the EPC Information Services (EPCIS), in order to support the Internet of Things concept. Our implementation is capable of enhancing the architecture of e-business frameworks, thus introducing an innovative collaborative business model which seamlessly integrates the inter-enterprise (public) with the intra-enterprise (private) processes.

The implemented architecture is demonstrated by the presentation of a case study in documents (books, papers, etc.) tracking and management in an academic library environment. Despite the fact that such an environment has quite lot variations from an e-business logistics environment, it has the potential to illustrate (and, even, simulate) the key-points of the presented architecture when not a standalone but a whole network of academic libraries are taken into account.

## II. Current Status

Traceability is defined as the ability to trace the history, application or location of an entity, by means of recorded identifications. It also may be defined in general as the ability to trace and follow any product through all stages of production, processing and distribution. Traceability itself can be divided into three types:

- Back traceability (supplier traceability)

This work was financially supported by the General Secretariat for Research and Technology (GSRT) [16] of the Hellenic Ministry of Development in the framework of project SELIDA – contract #09SYN-72-646 which runs under the "Cooperation", 2009 Call.

- Internal traceability (process traceability)
- Forward traceability (end-user traceability)

Having the end-to-end traceability encompasses all three types of traceability and since traceability is defined over every stages of a value-chain, several researchers have pointed out various elements that should be taken into account.

Traceability systems store information and show the path of a particular object of interest along the whole value-chain from the supplier/producer to the retailer/distributor and eventually to the consumer/end-user. Throughout this process, secure, reliable and automatic object identification is crucial to provide effective and efficient tracing.

Barcode technology, in the past, has been used for the identification of items. However, in order to meet the traceability requirements imposed by the governments, a new technology that allows automated recording of information was needed. This need has been partially fulfilled by the revolutionary developments regarding the RFID technology.

Many logistic services have already integrated RFID identification technology into their services and products but these solutions most often implement a custom or proprietary communication flow. This means that it is quite difficult to come up with a generic approach against these solutions.

On the other hand, providing traceability services apart from trading logistics services i.e. physical documents interchange is even more demanding since existing services (e.g. Xerox Docushare, Papyrus, etc.) refer only to digitized documents management. Nowadays, document interchange between organizations, authorities and citizens is realized via the well-known courier shipping services (i.e. FedEx, etc.), but these services are almost always built into proprietary protocols while a gap often occurs when different services, even of the same type, in inter-continental transactions are involved.

Especially for book tracking (i.e. lending libraries, etc.) libraries and Inter-Librarian Loan (ILL) services in general employ standard interchange formats, exploited via web services, in order to share repositories and establish collaboration among them. However, a global standard elaborating libraries worldwide does not exist.

## III. Architecture of the Deployed Information System

In the context of this section we are going to concisely present the architecture of the Document Tracking System (DoTS), which has been designed in order to tackle the issues mentioned in the previous section.

Section A concisely describes the technologies employed whereas section B presents a brief description of the entire architecture and its basic components.

### A. Technologies and specifications employed

The following technologies and specifications have been utilized in the context of the system: **RFID**, used to uniquely identify physical objects and **EPC**, providing the underlying framework that the system takes advantage of in order to offer standardized tracking services.

#### a) RFID

The RFID (Radio Frequency IDentification) is a well-documented [1]-[6] and widely adopted [7], [8] technology, that provides the ability to uniquely identify objects tagged with RFID tags using special readers [9]. The main goal of the architecture is to be able to track documents on a potentially global scale. RFID provides significant advantages over other automatic ID technologies (specifically the widely applied bar and QR codes), as RFID tags:

- can be detected in bulk
- don't need to be aligned with the reader (line of sight) in order to be read
- can be detected from a greater distance
- have a larger data capacity
- are less susceptible to damage

These outweigh the benefits of the bar and QR code tags (which are less expensive and quite ubiquitous compared to RFID tags) for the application on important documents and can enable the introduction of innovative services, such as real time traceability and theft prevention. Additionally, RFID is intrinsic in the framework of global standards published by the GS1 that concern the EPC, which have been exploited in order to provide a method to globally provide tracking information services.

#### b) EPCglobal

The GS1 EPC global is a suite of standards and specifications that leverage the RFID technology in order to globally enable visibility and collaboration on an item level. These standards comprise the framework depicted in Fig. 1.

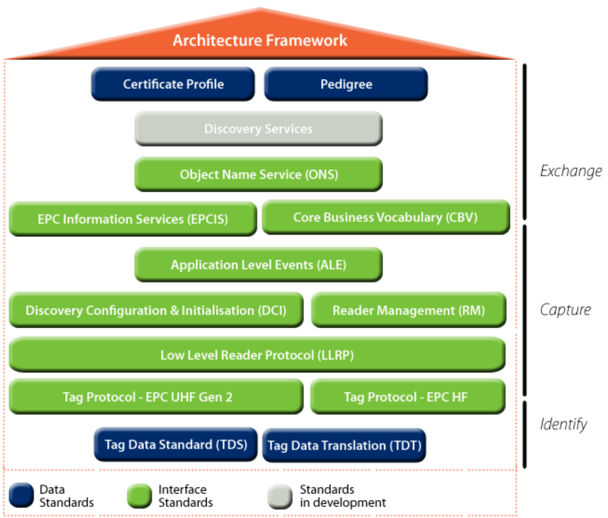



Fig. 1 GS1, EPCglobal framework standards, 2014 [6]

Information about these standards can be found in the GS1 website [5]. In the context of the proposed system, most of the EPC related GS1 standards have been utilized.

c) Object Name Service Specification

The discovery and tracking service of physical documents that has been implemented exploits the Object Name Service (ONS) 2.0.1 [5] and the EPC Information Services (EPCIS) 1.0.1 [2], in order to enable the mapping of EPC tagged documents to addresses of arbitrary, but with a standardized interface, object management services (OMS).

## B. Framework Architecture

The proposed framework aims to support as many of the EPC global standards as possible; in order to provide the ability to map single physical objects to URIs and to track related information regarding the entire lifecycle of the representation for all involved organizations in the value chain, together with the realization of ONS-based web services available in the cloud.

The proposed architecture is value chain agnostic pertaining:

- common logistics value-chain (i.e. manufacturer, logistics service, retail and end-user/customer)
- physical documents inter-change value-chain between public authorities or organizations of the public sector and citizens or companies of the private sector
- objects inter-change value-chain in demanding cases such as insurance organizations, shipping companies, courier companies, etc.

Fig. 2 illustrates the incorporation of the aforementioned technologies in the proposed framework architecture.

The proposed ONS service layer consists of a collection of ONS-based web services that are able to provide information from the organization's internal hierarchy model breakdown, using the global EPC notation, without affecting the existing processes of the value chain. Each organization's ONS service layer is responsible for providing per object, both public and private information, using the global EPC notation.

The private web services satisfy per sector- needs for real-time, synchronous physical objects tracking. These needs have various orientations, depending on the corresponding node of the chain. The public web services address needs of common operations, regarding single object's lifecycle information, such as location, history etc.

The integration of the RFID subsystem, into the architecture framework raises two main issues. The first regards the capturing of tags' information as well as the identification of the tags themselves. The RFID reader scaling, range and reading angle are of major importance since it is incorporated in a versatile environment (many different types of organizations performing totally different internal processes). The second issue regards security and privacy issues that are raised when the identified object's data should be classified. Security and privacy issues are presented in detail in next paragraph.

a) RFID Middleware

The RFID middleware is responsible for receiving, analyzing processing and propagating the data collected by the RFID readers to the Information System that supports the business processes. The middleware hides the complexity of the actual RFID infrastructure and only provides business events. On the other hand, the middleware is oblivious of how the data it provides gets handled afterwards.

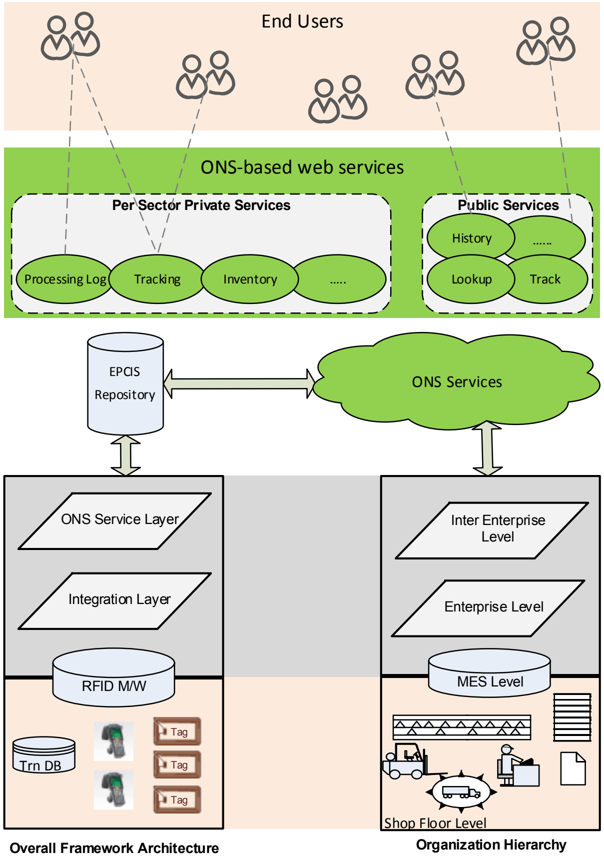



Fig. 2. Proposed framework architecture.

Specifically, the RFID middleware provides facilities for:

1. EPC Allocation
2. Device Management and Monitoring
3. Data Collection and Integration
4. Data Structure and Data Association
5. Data Filtering and Data Routing
6. Line Coordination and Process Control
7. Legend and Graphics Creation
8. Visibility and Reporting
9. Track and Trace Applications

b) ONS Resolver

The purpose of the ONS Resolver is to provide secure access to the ONS infrastructure, so that its clients would not only be able to query for the OMSs related to EPCs (which is the de facto use case of the ONS), but also introduce new or delete any existing OMSs for the objects. This has been accomplished by creating a SOAP web service layer that functions on top of the ONS, which provides authenticated and authorized users with the capability to query the whole

ONS infrastructure and discover the OMSs for the given EPCs, to add or delete OMSs or to add or delete users of the ONS Resolver, depending on their permissions.

In addition to the secure access to the ONS infrastructure, the ONS Resolver acts as an authorization server [10] for the relevant OMSs. This way, whenever a user uses the ONS Resolver to get the address of an OMS and authenticates successfully, an access token will be returned along with the result of the query, which, if used in the subsequent interaction between the user and the OMS, it can provide privileged access to the service. The authorization procedure that has been developed is depicted in Fig. 3.

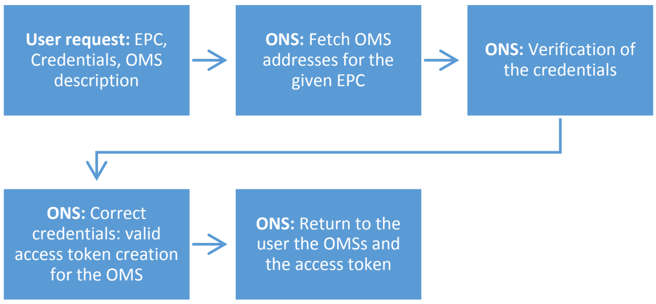



Fig. 3. ONS authorization procedure.

c) Object Management Services

The object management services (OMS) provide management, tracking and other value added services for the EPC tagged objects. The ONS Resolver maps the object management services to the objects according to their owner and type (for example the OMS for objects tagged using an SGTIN EPC is going to be determined by their GTIN; i.e. the OMS for the EPCs with the same GTIN is common [5]) and they should be implemented according to the EPCIS specification [2].

The most common case of the aforementioned context is the presentation of all historical data relevant to a single physical object. The resulting historical data can contain information regarding object transactions within locations controlled by the owner of the item or other organizations, as long as all the implicated parties implement the described framework; this was rendered possible through the utilization of the ONS Service Layer. The data of the objects stored by the different parties adhere to the same API and, as a result, one can select the exact kind of these data that are to be presented. For example historical data could describe an object's transaction only with dates on interaction with an RFID M/W, or with geographic interpreted location of the interaction, or even with more meta-data of the object like its purpose of transportation, original owner, final destination and whether it is classified or not. In an expanded version of multiple organizations running the presented framework, any tagged object (document, package, suitcase, etc.) could be easily and reliable recognized in the entire framework's context.

Finally, the arbitrary nature of the OMS themselves, even though they are implemented with standardized formation, enables each organization to level-up objects' related information according to the specific organization needs and requirements.

d) Security and Privacy aspects

Since the proposed architecture is based on web services, our first goal is to identify and classify those services in regards to their security requirements. Even before that, we can safely assume that all web services can and should implement TLS as a standard form of encrypting the data channel in use (usually the Internet).

Based on the aforementioned architecture diagram, we can easily identify that there are public and non-public facing web services. In regards to the non-public facing web services, the ones that reside within the internals of the proposed architecture stack, we can exploit the useful fact that both the clients and the servers of this part of the architecture are known and can be controlled in regards to their implementation of the security mechanisms. Therefore it is safe to assume that the use of client certificates is a feasible security mechanism. Client certificates are a very robust way of handling secure and, in conjunction with TLS, encrypted authentication and authorization and the issues with scalability and deployment that are usually encountered in more general scenarios are not applicable to our proposed architecture. As for the public facing web services, we follow the industry standard of API keys due to the fact that although we name those services "public facing", in reality those services will be accessed not by casual end users but by the information systems of the organizations that will employ that services of our proposed architecture. A case study of such a deployment is described in the next chapter.

Moreover, all the web services should follow a standard of secure design that, although already a common practice in popular web services around the web, we will briefly describe below.

As mentioned above, all services should be authenticated over an encrypted communication channel. Messages should be digitally signed, as well as encrypted, to provide privacy and tamper-proofing when the message travels through intermediary nodes route to the final destination even within different organizations that implement the same proposed architectural stack. The usage of the access token (a unique ID or nonce, a cryptographically unique value) generated by the ONS Resolver within every request, will, obviously, provide protection against unauthorized usage and it will also aid to the detection message replay and man in the middle attacks. HTTP methods should be valid for each API key and associated resource collection and method by white-listing allowable methods. Any request for exposed resources should be protected against CSRF and insecure direct object references should be avoided.

# IV. Case Study on Physical Documents Tracking in a Library Environment

In order to evaluate the applicability of the proposed architecture on a real environment we deployed it on the existing Integrated Library System (ILS) that is being used in the central university library of Patras, Greece. The ILS that is being used by the institution is the well-known open source ILS named KOHA [11].

As with all ILS, KOHA supports a variety of workflows and services to accommodate the needs of the institution. Our proposed scheme focuses on a handful of those services and augments them with additional features. This is generally done by adding, in a transparent way, the additional UI elements and background processes that are needed for our scheme to work.

The specifics on how this is done will be presented in the following section, while first we will discuss briefly on the exact services of the KOHA ILS that our scheme aims to augment.

## A. Supported services

In its initial design, our scheme aims to provide additional functionality on the core services of an ILS. Those services are:

- Check Out
- Check In
- New Record
- Delete Record

There are also a number of tracking services that our scheme aims for and those are:

- History
- Location
- Search

To elaborate, when the check-out or check-in services are called in KOHA, an additional call is made on the SELIDA (the name of the developed framework from the related project) middleware that updates the status of the affected documents on the SELIDA database. The details on how those calls are made will be described in the next section. Similar functionality can be seen on the entire core and tracking services that are applicable in our scheme as described in the previous chapters.

## B. Integration layer

In order to provide the added functionality to the existing KOHA services we designed and implemented an integration layer that seamlessly handles all the extra work along with the usual service workflow.

The primary reason to provide a seamless layer instead of changing the actual services (i.e. the source code) is the fact that every integrated system that is actually in a production environment needs the benefits of a continuous and stable update process that is offered by the systems development team along and implemented by the organization's administrator. Adding extra functionality in the form of changes to the system's code would require continuous maintenance of this part of the system on par with the normal updates of the KOHA ILS.

To overcome this obstacle, and given to the fact that KOHA is a web based application as most of the modern ILS, we designed the integration layer so that it is injected upon page load as a JavaScript file on the pages that we are interested in. On our specific case study, we used the "mod_substitute" directive on the Apache web server that was serving the KOHA web pages. Each time a module/page of interest is requested by the server (i.e. check-out), the web server adds a <script> tag that loads all the additional functionality in the form of a JavaScript file.

This layer, in the form of a JavaScript application module, adds the required UI elements for our proposed scheme to work and handles all the web service requests that are necessary.

As an example, we will showcase the check-out process. When the user navigates to the check-out KOHA module by requesting the module's URL, the apache server injects the selida.js file which is the integration layer code (so called SELIDA module). SELIDA module starts executing upon page load and adds the button "Scan" next to the button "Checkout". When the user presses the button "Scan", a web service request is launched from the SELIDA module which starts up the RFID reader via the SELIDA middleware services. The results (per example the barcodes and titles of the documents that the reader identified) are communicated back to the SELIDA module which in turn shows a pop-up window to the user indicating those results. After this procedure, the normal check-out workflow resumes by sending the required POST requests to the KOHA web server (just like when the user presses the "Checkout" button after adding the barcode). When the check-out process ends and the web server responds with the next web page as per KOHA workflow, the SELIDA module sends a second web service request to the middleware indicating that the check-out is complete so that the rest of SELIDA architecture continues with its own check-out workflow presented in the previous chapters.

As we can see on this example, the extra steps that are needed for our scheme to work are completely transparent from the ILS itself. KOHA as an application retains its original functionality and workflow while the user benefits from the added services that our SELIDA module provides.

Fig. 5 illustrates the framework's activities during the first object identification phase, previously described.

## C. Object Management Services

Since RESTful [12] web services are becoming the most common developers' choice for implementing API's and data retrieval services SELIDA's services have been designed following the RESTful architectural style and JavaScript Object Notation (JSON) [13] for data exchanging. Although these technologies do not comply with a specific standard, as other technologies do, i.e. SOAP, XML web services, when combined together they constitute a lightweight data-interchange mechanism that is easy for humans to read and

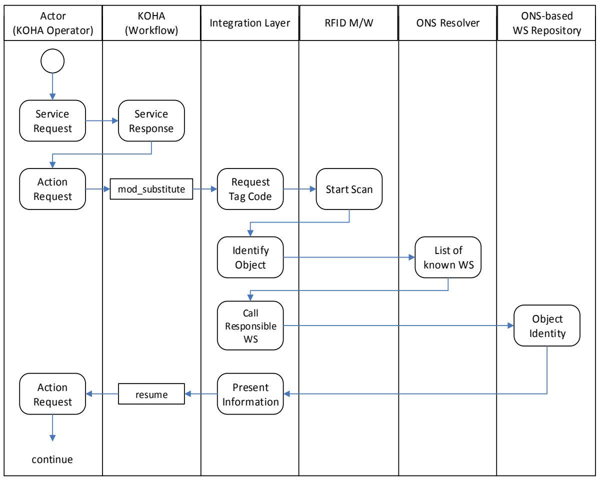



Fig. 5. Framework's activities during an object identification phase.

write and also easy for machines to parse and generate, thus appearing to be the easiest and most comprehensive messaging way among web services. Moreover, nowadays numerous open standards-like specifications (i.e. JSONDoc [14], Swagger [15], etc.) have been presented towards describing, producing, consuming, and visualizing RESTful web services so as RESTful web services can become a complete framework.

Currently, based on the deployment of the proposed architecture on the existing Integrated Library System (ILS) that is being used in the central university library of Patras, Greece, three types of management services (history, location


```
{
    "EPCList": [
    {
        "EPC": "1111;2222;3333"
    },
    {
        "EPC": "4444;5555;6666"
    },
    {
        "EPC": "7777;8888;9999"
    }
    ]
}
```

```
{
    "ErrID": 0,
    "EPCList": [
    {
        "EPC": "1111;2222;3333",
        "ItemCode": "abc",
        "Title": "Title of document",
        "Status":"Self",
        "Owner":"LIS UPATRAS",
        "Location":
        {
            "Lat":"38.28923",
            "Lon":"21.785369"
        }
    },
    {
        "EPC": "4444;5555;6666",
        "ItemCode": "def",
        "Title": "Title of document",
        "Status":"Loaned",
        "Owner":"LIS UPATRAS",
        "Location":
        {
            "Lat":"38.28923",
            "Lon":"21.785369"
        }
    },
    {
        "EPC": "4444;5555;6666",
        "ItemCode": "ghi",
        "Title": "Title of document",
        "Status":"Self",
        "Owner":"LIS AUEB",
        "Location":
        {
            "Lat":"37.9940902",
            "Lon":"23.7324801"
        }
    },
    ]
}
```


**Request** **Response**

Fig. 4. An example of JSON request and response data structure for status OMS.

and search) have been implemented, as described in paragraph A.

As for an example, the JSON schemes for the request and response of the search service is presented in Fig. 4.

As depicted in Fig. 4 request to the status OMS may contain multiple EPC tags at a time, while the implemented system response contains an indicative error code (0 in case of no error) and relevant data for recognized books. In the illustrated example the first two EPCs belong to books owned by the central university library of Patras, Greece (LIS UPATRAS) while the third belongs to a book owned by the university library of Athens University of Economics and Business, Athens, Greece (LIS AUEB).

## V. Benefits and Scalability

Although the described framework has been deployed in an ILL (Inter-Librarian Loan) environment its initial inspiration and design has been originated in general logistics or warehouse inventory stocktaking environments. Based on this origin the whole framework has been designed and deployed so as to offer organization agnostic information, supported by ONS-based web-services that are integrated at the top of each organization's hierarchy stack, as illustrated in Fig. 2.

The proposed architecture offers a set of versatile characteristics due to the combination of reliability and uniqueness, induced by RFID technology along with the ONS perspective implementation throughout the architecture that derives interoperable information exchange. These characteristics could constitute the basis for the employment of such a framework so as to derive into a generic Internet-Of-Things service platform able to handle information and processes of any value-chain type, including and not limited at food supply chain, luggage handling, physical document interchange, etc.

This prospect could be further substantiated by the nature of the ONS, which is actually a mechanism that leverages Domain Name System (DNS) to discover information about an object and its related services from the EPC code. Conclusively, the presented architecture could be scaled in the same way as internet does, since internet is based on DNS and the presented architecture is based on ONS, inheriting DNS scalability capabilities.

## VI. Conclusion

In spite of the promising nature of RFID technology, numerous applications in the actual logistics field area have not been reported. Only a few pilot studies as well as experimental tests have proved that RFID would be a successful tool to enable supply chain traceability. The reasons why companies are yet reluctant to have confidence in adopting the technology to gain their product visibility may be attributed to the several challenges such as lack of standards, immaturity of RFID, and privacy issues.

This paper presents a ubiquitous approach towards a collaborative logistics environment and concludes with a case study implementation involving physical documents tracking in an academic library. The main focus of the proposed methodology and implemented architecture addresses the issue of empowering the whole framework with a standard specification for objects tracking services, thus the integration of the ONS-perspective in the architecture. The integration itself, which, within the EPC global framework is mainly achieved by utilizing the Object Naming Service, enables involved organizations to act agnostically of their entities and provides them with the ability to resolve EPC tagged objects to arbitrary services, but with a standardized manner.

However, security and privacy issues should be further investigated as future work, apart from the issues already covered though the implementations of the framework's web services, so as the presented framework will be promising enough for evolutionizing the way currently exploited for tracking items in the supply chain.

## References


[1] ISO. (2010). ISO RFID Standards: A Complete List. Retrieved from http://rfid.net/basics/186-iso-rfid-standards-a-complete-list

[2] GS1. (2007). EPC Information Services (EPCIS) Version 1.0.1 Specification. Retrieved from http://www.gs1.org/gsmp/kc/epcglobal/epcis/epcis_1_0_1-standard-20070921.pdf

[3] GS1. (2010). EPCglobal: Specifications. Retrieved from http://www.gs1.org/epcglobal/standards/specs

[4] GS1. (2010). EPCglobal: Standards. Retrieved from http://www.gs1.org/gsmp/kc/epcglobal

[5] GS1. (2013). Object Name Service (ONS). Retrieved from http://www.gs1.org/gsmp/kc/epcglobal/ons/ons_2_0_1-standard-20130131.pdf

[6] GS1. (2014). EPCglobal framework standards. Retrieved from http://www.gs1.org/gsmp/kc/epcglobal

[7] Hossain, M. (2012). A Comparison of Voluntary and Mandatory Adoption of Radio Frequency Identification (RFID) Technology in Organizations. Retrieved from http://www.wbiconpro.com/421-Hossain.pdf

[8] RFIDjournal. (2013). Survey Shows Half of U.S. Retailers Have Already Adopted Item-Level RFID. Retrieved from http://www.rfidjournal.com/articles/view?9168

[9] Wikipedia. (2007). Radio-frequency identification. Retrieved from http://en.wikipedia.org/wiki/Radio-frequency_identification

[10] IETF. (2012). RFC 6749: The OAuth 2.0 Authorization Framework. Retrieved from http://tools.ietf.org/html/rfc6749

[11] Official Website of Koha Library Software. Retrieved from http://koha-community.org/

[12] Wikipedia (2014). Representational state transfer. Retrieved from http://en.wikipedia.org/wiki/Representational_state_transfer

[13] JSON (JavaScript Object Notation). Retrieved from http://json.org

[14] JSONDoc. Retrieved from http://jsondoc.org

[15] Swagger. Retrieved from https://helloreverb.com/developers/swagger

[16] General Secretariat for Research and Technology, http://www.gsrt.gr

# Dynamic Cargo Routing on-the-Go: The Case of Urban Solid Waste Collection

Artemios G. Voyiatzis, John Gialelis, and Dimitrios Karadimas
Industrial Systems Institute
Athena Research and Innovation Center in ICT and Knowledge Technologies
PSP building, Stadiou Str., Platani Patras, GR-265404, Greece
Email: {bogart,gialelis,karadimas@isi.gr}

***Abstract*—Efficient collection of urban solid waste is a growing problem for cities and urban areas. Optimizing service routes based on actual waste bin fill level information can result in significant cost savings, reduced emissions, and improved city image. Fill level sensor installation and networking at an urban scale poses unique challenges. We describe Dynacargo, an innovative smart city application that leverages RFID technology in the waste bins and delay-tolerant networking (DTN) in roaming vehicles acting as mobile sinks. Dynacargo collects and transmits waste bin status information to the backend systems utilizing already existing network infrastructures and single-hop communications thus, reducing associated costs and time for installation, operation, and management of the necessary networks and ICT systems.**

## I. Introduction

There are six major categories of waste: industrial (manufacturing); construction and demolition; mining and quarrying, agricultural and forestry; radioactive; and urban solid waste. While the formers involve only a handful of collection points (e.g., plants), predictable waste production, and long filling periods, urban solid waste involves numerous waste bins that exhibit significant filling variations and must be emptied many times per week or even per day.

A poor or inappropriate urban solid waste collection system can introduce or amplify problems affecting the urban environment, the authorities, and the citizens. Main sources of problems and debate include: time of collection; plan and frequency of collection; waste bin placement; number and capacity of waste bins; garbage truck routes; and age of the fleet. On the other hand, an optimized collection system allows reductions in traffic congestion, noise levels, emissions, and costs, while improving public image and general sense of duty for public cleanliness [1].

Urban solid waste bins are still emptied *by experience* and *a good guess approach*, although the benefits of utilizing *actual status information*, like fill level, frequency of use, and environmental conditions are evident [2], [3], [4]. Significant reductions can be achieved not only by optimizing the service routes but also when a garbage truck avoids to idle as to allow service personnel to check if a waste bin should be emptied or not. Idling and low-speed moving -even when unloaded- can account to more than 50% of a truck's emissions during a service route.

The collection of actual status information about the waste bins of a whole city or municipality requires the deployment and operation of sensors at *urban scale*: hundreds or even thousands of sparsely-distributed sensors installed in waste bins in populated areas. As an example, about 20,000 waste bins are installed throughout the city of Berlin in Germany. At this scale, network connectivity and formation problems become apparent. The communication cost for using cellular networks becomes prohibitively large; the short-range multi-hop networks can be unmanageable and unreliable; and a very large number of long-range gateways (e.g., WiFi) must be installed and operated to cover the whole area of a city.

In this paper, we present Dynacargo, an innovative system that requires minimal infrastructure costs and handles urban-scale waste bin status information collection. Short-range sensors are installed in the waste bins. Vehicles passing by the bins collect status information, acting as mobile sinks. While roaming around the city, the vehicles explore connectivity opportunities, like through WiFi hotspots, and, using delay-tolerant networking (DTN), they send the collected information to the waste collection system for further processing.

The rest of the paper is organized as follows. In Section II, we present requirements and challenges for integrating field information in an urban solid waste management system, while in Section III, we revisit existing approaches in the literature. Section IV presents the Dynacargo approach, while Section V discusses the mobile sink system design and the challenges faced. Finally, Section VI concludes our paper and presents future directions of work.

## II. Requirements and Challenges

The implementation of an efficient urban solid waste[1] management system for a city is a challenging task. It requires not only the optimal use diminishing city budgets but also the increased satisfaction of the city visitors and citizens for the quality of the experience they receive.

Earlier attempts to solve the waste collection problem i.e., the garbage truck fleet routing problem, did not consider any variations in the waste bin fill levels [5]. The collection and the integration of information regarding the actual status of waste bins in the field can assist city planners into devising waste collection plans that serve both aforementioned aims.

In the next, we consider the requirements and challenges for collecting waste bin information from the field at an

[1] For the sake of simplicity, we shall use the terms "waste" and "urban solid waste" interchangeably hereon.

urban scale. The areas of attention include the integration of sensors in waste bins, the accuracy in measuring the fill level, any additional information that must be collected, and the transmission of the information to the waste management backend systems.

### A. Urban Solid Waste Bins

The *waste bin identification* for inventory management is the most elementary information processed by a waste management system. Advances in RFID technology allow the use of RFID tags for fast bin identification. A patrol service can be realized by driving near the waste bin equipped with an RFID reader. The driver or a backend system can be notified of missing bins and initiate necessary investigation. It can also be used to postpone routing a garbage truck to the specific area, until a bin is installed again. This can result in significant cost savings.

The *fill level* is an indicator for emptying a bin. Once filled, the citizens cannot use it, they may be annoyed, and they may even leave their waste next to the bins. This results in emptying delays since the waste must be handpicked by service personnel while the truck idles, in a bad city image, and in possible environmental and hygiene threats. On the other hand, frequently servicing near-empty bins is not an optimal use of the available resources.

The *fill level estimation* of a solid waste bin is not an easy task. Unlike liquid or granular materials, randomly-placed solid objects of different shapes and dimensions severely hamper the estimation accuracy. Only a handful of works are published on *sensing solid waste level* or fill, mostly focused on specific waste types. Examples include paper [6]; corrugated board and cardboard [2]; recyclable materials (Tecmic Ecogest); moisture content for combustion in a furnace [7], optical sensing [8], and camera image processing for urban solid waste [9], [4], [10]. The environmental conditions around or inside the bin may affect significantly a sensor's operation. Example conditions include ambient light, humidity, moisture, and dust. This can lead to unreliable measurements and need for frequent sensor cleaning.

It may be beneficial to collect additional information regarding a waste bin, such as *loading events* (time and frequency) and *weight*. This information could be used to plan better service (reduce the time that the waste stays in the bin) and to optimize routes. Garbage trucks can compress the waste once loaded on truck and reduce its volume. Trucks are weight-saturated in contrast to waste bins that are volume-saturated. Weight limits must be obeyed for safety and truck-health reasons. On the other hand, citizens observe the waste volume in a bin, which also affects the image of the city. Thus, a good combined estimator of volume and weight can be beneficial. However, we notice that it is simpler to record service time and waste weight during the loading process of the waste in the garbage truck, rather than integrate such sensors in the bin.

Service delays can occur for a wide range of reasons. *Environment sensors* (e.g., air quality) could provide additional information on the exact conditions so as to assess the accompanying risks and received citizen complaints.

Waste bins are *manufactured of different materials*, such as plastic (PVC), metal (aluminum), or a combination of them. The variety and the specific materials must be considered in order to ensure proper operation, sensing, and communication of electronic devices. The size and shape of waste bins is not standardized thus, sensing systems may need to be adapted to a specific design and shape.

The *economic benefits* of the so-called "smart bins" are still not well-received by many decision makers. In many cases, it requires a large upfront investment to replace the existing waste bins. This cost cannot be fully justified. In response, waste bin manufactures have not introduced waste bins with intelligence integrated in their frames, due to the increased costs involved and little market interest. As a result, smart bins are nowadays mostly implemented as a computing system *retrofitted* into existing bins. This may increase costs and create significant reliability issues in the operation of the system, since it is exposed to a severely harsh environment, including high temperatures; chemical reactions; and violent shaking.

A large number of bins is needed for servicing a city or an urban environment. These bins are often *placed unattended in public spaces*. Thus, they may be vandalized and require frequent replenishment. Furthermore, a computing system, either integrated or retrofitted in existing designs, can be an attractive target for theft due to the valuable components it may contain (e.g., advanced electronics, metals, and batteries).

### B. Actual Status Information Collection

The collection of actual status information from each and every waste bin at an urban scale poses significant challenges. A simple solution is to equip each bin with a sensor node and a long-range communication unit, such as a cellular modem, and directly transmit information to the backend systems for further processing. Clearly, this option is not economically viable at urban scale due to telecommunication costs. Furthermore, these nodes will need some source of battery and frequent recharge, either automatically (e.g., via a photovoltaic panel) or manually (e.g., via frequent visits by service personnel). Also, radio coverage is not guaranteed throughout a city terrain. Tests must be performed at each bin location and iit may be necessary to move the bins to new locations.

A second approach is to take advantage of any available city network infrastructure, such as public WiFi hotspots, and use them as gateways to connect the sensor nodes with the backend systems. This cuts the long-range telecommunication costs but still an urban-scale network of sensor nodes must be installed and tested for connectivity with the gateways. In case there is no sufficient coverage at some areas, the authorities must install and maintain new infrastructure, incurring additional costs for procurement and network management at both the gateway and the sensor node levels.

A different line of thinking is to take advantage of communication units that roam around the city and collect information. These units act as mobile gateways or sinks for sensor nodes equipped only with low-power wireless communication units. The *garbage trucks* are a profound option, since they are already in the field. In case the trucks are already equipped with a fleet management system (FMS) unit, it may be possible to enhance the unit and include an interface for receiving and transmitting all the information collected from the field unit (i.e., the waste bin or a nearby unit) since the last visit. In this setting, it is necessary to adapt the on-truck software and hardware, as for example to integrate an antenna and a network protocol stack for communicating with the field unit. Furthermore, it is required to either extend the backend FMS system to include processing of collected information or to implement interfaces for exporting to a waste management system. The garbage truck option does not come without limitations. The truck collects information about the already visited waste bins and not about the ones way ahead. This is more useful for collecting historical information and cannot be used to dynamically adapt the service plan.

In many cases, *city or public service vehicles* roam around the city. Examples include municipality patrol service; public transportation like buses, trams, and trains; post office and courier services. These vehicles could be used as mobile sinks and transmit collected information either in real-time through a cellular network (if they already have Internet access for business operations) or once they connect to a WiFi hotspot, either while on the go or when they return to their depot after completion of service. Such vehicles can cover large parts of the city and provide fresh information. The large periods of disconnections and the opportunistic nature of connectivity episodes with both the sensor nodes and the Internet gateways poses significant challenges for application robustness.

The *citizen involvement* can dramatically increase the coverage, accuracy, and timeliness of information. Citizens' participation and contribution can transform their role from passive receivers of the waste collection service into active "*prosumers*" that both *produce* information *for* the service and *consume* information *of* the service with the aim to improve the quality of the received service. In contrast to the aforementioned approaches, the citizen involvement leads to manual rather than automated information collection. How to successfully motivate and reward participation remains an open discussion in the literature and heavily depends on the business models of the service provider. Due to this, we focus our discussion in the following sections on the automated collection systems.

## III. LITERATURE REVIEW

### A. Fill Level Sensing Prototypes and Architectures

A prototype waste bin system was constructed for the city of Pudong, Shangai, PR China [4]. The system consists of two parts. The first part is installed in the bottom of the waste bin, occupying about 20% of its storage capacity. This part contains the battery and the GPRS communication system. The second part is installed in the lid of the bin and contains a photo camera, a PC104 with 12 VDC supply, three LEDs and temperature and humidity sensors. The system takes photos of the inside of the bin (LEDs are used as flash to enhance the contrast). The photos are sent through GPRS to a central server for processing and bin fill estimation.

A less bulky system was constructed in Malaysia [10]. The waste bins are equipped only with a passive RFID tag with a range of 6 meters. The garbage trucks are equipped with an RFID reader, a GPS receiver, a low-resolution camera, and a GPRS connection. Once the truck stops nearby the waste bin, the RFID reader reads the tag of the bin and activates the camera to take a photo of the surrounding area $(3 - 6\mathrm{m}^2)$ before and after waste collection. The photos are then sent to a central server that uses image processing and estimates the volume of waste inside and nearby the bin.

Fill level estimation can be realized with sensors based on a modulated infrared beam detected by a photodiode. The accuracy of the measurement can be influenced by transparent objects, the reflection of light on object surfaces, the ambient light, and the dirt over emitter and detector surfaces. One solution is to use multiple line-of-view sensors in parallel and a majority decision system [8]. However, this results in increase of cost per bin. Another approach is to use hall effect sensors and count the times a waste bin is opened or moved. Under some (strong) assumptions and simplifications, the count can be used as a proxy for the amount of waste inside the bin [8].

The SEA project designed a smart bin prototype using an ultrasonic sensor and IDEA's ArgosD (TelosB) sensor nodes running a custom TinyOS 2.1.1 application. The smart bin connects with gateways that are based on the New/Linux OS. SEA proposes a three-layer architecture for information collection [11]. In the lower layer, short-range *single-hop* communication through IEEE 802.15.4 is realized to transmit fill levels. In the middle layer, gateways with both short-range and long-range modules (GSM/GPRS) are used to collect information from the sensor node and send them through IPv4 to the backend. Finally, in the upper layer, servers receive, store, and process the collected information.

A similar architectural approach is followed in [12]. There, two sensors are used: an ultrasonic for fill level and a load cell for weight. Sensed values are transmitted to nearby gateways installed in light poles. The design allows detecting removal of a bin through the lack of communication. An interesting idea explored in the paper is the need to promote the citizen participation in the smart waste process. A prototype application for Google Android smartphones is reported, through which citizens locate the closest bins that are not filled and they can associate their activity (loading of waste) with bin filling through scanning a QR code placed on the bin.

A three-layer approach is followed also in [13]. There, multiple sensors are used in the lower layer: ultrasonic for fill level, load cell for weight, temperature and humidity, hall effect and accelerometer for detecting bin cover open events. Sensed values and operational parameters, such as bin identity, date, time, and battery power level, are collected and transmitted when a cover opening is sensed. This way, energy-efficient, real-time fill level reporting can be achieved.

The EU FP7 OUTSMART project designed a mesh wireless sensor network for Berlin, Germany [14]. The project uses an ultrasonic sensor for fill level estimation and a wireless sensor network in mesh topology based on IEEE 802.15.4 for connecting nodes with gateways through *multiple* short-range hops. From the gateways, the information can reach the backend systems for further processing.

The EU FP7 Future Cities project[2] develops an urban-scale living lab in the city of Porto in Portugal. In the context of the project, the Municipality of Porto is developing an

[2] http://futurecitiesproject.eu/

innovative data collection system for monitoring fill level of garbage containers. Functional pilot tests will be run using buses as “data mules”. Both sensor-node-to-bus and bus-to-cloud communications are based on IEEE 802.11 (WiFi). The latter utilize roadside units (RSUs) and WiFi hotspots already installed in the city of Porto[3].

The EU FP7 Straightsol project[4] streamlines charity collection (e.g., clothes and books) from donation banks installed in public spaces and retail shops. The project used an infrared sensor for fill level estimation (at 20% reported accuracy) and transmitted the information twice per day through GSM for scheduling next day's collection. A small-scale actual pilot included 37 donation banks, 50 retail shops, and 5 vans and resulted in an estimated 5% revenue gain [15].

### B. Commercial Products and Services

Commercial products and services for estimating waste bin fill levels are also available. Most offerings use ultrasonic sensors, as they can provide more reliable measurements. Examples include Agora Energy (France), Enevo (Finland), Libelium (Spain), IGS Research (Spain), Trimble (USA), Urbiotica (Spain), and SmartBin (Ireland).

Communication with the waste bins can be classified in two main categories. The first is proprietary protocols transmitting at the license-free ISM bands (like 868 MHz or 2.4 GHz). These require nearby Internet gateways from the same company (Agora Energy uses Sigfox Ultra-Narrow Band, Urbiotica uses U-Sense and U-Box). The second is cellular protocols based on GSM/GPRS/3G (Enevo and SmartBin). Libelium offers systems that can be configured to work with both gateway-based short-range protocols (IEEE 802.15.4, ZigBee, WiFi) and direct, long-range protocols (GPRS and 3G).

Sending as little as 1 message per day, all these approaches claim a battery life of 5-10 years, even when using cellular communication.

### C. Mixing RFID and WSN

Wireless sensor networks (WSN) and radiofrequency identification (RFID) are pervasive environments. The integration of these two pervasive computing environments results in reliable, energy-efficient, survivable, and cost-effective solutions for various applications. From an energy consumption point of view, in an ideal setting, the passive RFID tags use energy supplied by RFID readers, receive a query, power up the sensor, collect the measurement, and send back a reply. This is a battery-free setup that requires zero maintenance.

The Wireless Identification and Sensing Platform (WISP) is a first approach combining RFID tags and sensors [16]. A 1-bit accelerometer (sensor) is integrated in an Alien Technologies ALL-9250 EPC-compliant UHF RFID tag for a total cost of US$0.90 per unit. In comparison, battery-powered communication systems cost more than US$5.0 per unit. The sensed quantity is transmitted as part of the tag identity in a range of one meter. This is useful for tracking small objects in a house but cannot be utilized in an open-space, harsh environment, as the reader must be placed close to the tag. An improved WISP version is presented in [17]. The WISP now uses the whole 64 bits of the address space for transmitting sensed values. The platform is implemented in PCB rather than as an integrated circuit and uses two antennas: one for power reception and one for communication. The EPC transmission protocol is implemented in software for the MSP430 microprocessor. This approach departs from the traditional thinking of RFID tags as integrated circuits. A large storage capacitor (e.g., $10\mu F$)) is used to collect power. Once enough power is stored, a burst of activity occurs.

From an architecture point of view, there are several options combining sensors, WSNs, and RFID [18]. The key difference is how the communication with the backend systems is realized and the sensing requirements. A first option is to integrate sensors into the RFID tags (circuits) creating “sensor tags” [19]. These tags require an RFID reader to collect the sensed values. Due to the low power available and the integration capabilities, such devices usually integrate simple environmental sensors, such as temperature, light, humidity, and contact (magnets). A second option is to integrate an RFID tag within a WSN node. A combined device acting as a WSN gateway and an RFID reader is used to identify the nodes through the RFID tag and collect the measurements through WSN technologies. Yet another option is to use a WSN as to detect unusual activity and then trigger an independent (power-hungry) RFID reader to detect presence of specific tags or to limit search area for tags. The scenario could also work on the opposite direction; an RFID reader continuously monitors an inventory. Once a tag is reported missing, a WSN is awakened to collect field evidence (e.g., sound and images). An extended analysis of available company products and academic designs on integrating RFID tags and WSN for various application domains is provided in [20], [21].

## IV. The Dynacargo Approach

The Dynacargo project considers urban solid waste collection as a case example of dynamic cargo routing on the go. This is an extension of the capacitated vehicle routing problem, where a fleet of vehicles (garbage trucks) with limited carrying capacity of goods (waste) service a number of customers (waste bins) and the aim is to minimize the total route cost. The Dynacargo approach introduces two important aspects: a) the exact waste bin fill levels may be unknown at the start of the service (*dynamic cargo*) and b) the optimization algorithm must be capable to compute on-the-fly a new service path (*dynamic routing*), while trucks are *on the go* and new fill level information becomes available.

The Dynacargo approach proposes a modular architecture consisting of four layers:

1) Field units, for sensing and storing information about waste bin usage and status.
2) Information collection units, for collecting and transmitting information to the backend systems.
3) Backend systems, for computing optimal routes dynamically based on available information and distributing related information.
4) Information consumption units, for accessing and utilizing information produced in the system. These include systems installed in garbage trucks for getting

[3]http://futurecities.up.pt/site/buses-used-as-date-mules-at-porto/
[4]http://www.straightsol.eu/

route updates, web portals, and mobile applications for citizens to get informed about service frequency, fill level, and notifications for service outages.

The most critical part for the successful operation of Dynacargo is the process design for collecting status information from thousands of bins. At this urban scale of bins, network formation and routing between nodes becomes unmanageably complex and devices with constrained sources of energy become very quickly points of failure and cause network partitions and node isolations.

The approach of Dynacargo is to utilize vehicles that roam around the city and network infrastructure that is already deployed. The system architecture is depicted in Figure 1. The vehicles are equipped with information collection units that act as mobile sinks for the sensors. The vehicles exploit communication opportunities via public WiFi hotspots already installed throughout the city or once they return back to their depots. If the vehicle has already Internet connectivity (e.g., in the case of a bus offering Internet access to its passengers through a cellular connection), this can also be utilized. Once an Internet connection is available, the vehicles send to the backend systems the status information they collected. Since there is no need to transfer the bin status information in real-time, the deferred delivery of information is acceptable.

The selection of appropriate technologies is a crucial factor for the successful deployment and operation of the system. Regarding the sensor nodes, it is a design choice to use only single-hop links between sensor nodes and mobile sinks. This choice offers significant benefits over multi-hop communications. Single-hop reduces *energy consumption* for constrained nodes that need not turn on their radio to receive and process packets by other nodes. The short range of node transmissions allows *spatial reuse* throughout the whole city. *Routing overheads*, in terms of protocol processing, energy consumption, and network traffic are avoided. There is *no single point of network failure*, as multiple vehicles cover the city. The design is characterized by *scalability and simplicity*, as new nodes can be added to the network with minimal effort and without affecting network capacity.

All communications in the system are opportunistic in nature. From the sensor node point of view, it is by and large unknown when one of the free-roaming vehicles will reach its vicinity and establish a communication link. Also, it is unknown how long the contact will last, since we do not assume that the vehicle stops moving at each and every bin. From the mobile sink point of view, it is again by and large unknown when and for long the free-roaming vehicle will be within the range of a WiFi hotspot and thus, if and how much information it will be able send to the backend systems.

In this environment of operation, there are challenges to overcome and tradeoffs to explore for an efficient function of the collection system. Examples include the range of operation for the sensor nodes (longer distances versus power consumption and interferences), detection of contact opportunities (always-on versus power-off periods for energy efficiency and missed communication opportunities), transmission costs (rich information versus power consumption), disruption handling (transmission of long packet sizes and fragmentation issues), and protocol stack complexity (protocols with large headers and rich functionalities versus simple packet structures and required functions).

In principle, we opt for a lean, modular, and scalable system and network architecture that can support the urban scale of devices that must be integrated. As requirements may emerge in the future, the modular approach allows system components and technologies to be replaced with ones providing richer functionality. In the next section, we describe the system architecture and design for the mobile sinks. The mobile sinks are the critical chaining component that supports the whole operation of the system.

## V. Mobile Sink System Design

The Dynacargo information collection system has an adaptable, modular, and configurable design that allows optimizing its operation for multiple scenarios. The system is built using open-source software. It is based on a small-form-factor computer, consuming about 10 W of power (maximum) and having two USB 2.0 ports, built-in WiFi and Bluetooth, and one Ethernet port. It is supported by GNU/Linux and Google Android operating systems running from a microSD card.

### A. Field Information Collection

The waste bins are equipped with ultrasonic sensors and active RFID tags. The active RFID tags transmit periodically their (waste bin) identity and a sensed value regarding its fill level. This option allows the minimal possible information to be transmitted periodically. All tags are configured to operate in beacon mode, broadcasting the information every few seconds in a range of about 100 meters. Active RFID tag transmission technology is much lighter in complexity and in coping with harsh environments compared to that of WSNs.

The mobile sink connects with an active RFID reader installed in the vehicle through Bluetooth or one of the USB ports. As the vehicle roams around the city, the RFID reader reads the tags and collects the information to the sink. Given the short "packet" size (less than 128 bits), the tag range (100 meters), and a realistic average speed of 40 km/h in a city, the tag must beacon at least every nine seconds to be read by the passing vehicle. At this rate, its battery life will be depleted in about 5 years, which is rather satisfactory. Increasing the tag range or allowing a more stochastic operation, as in the case that an area is covered by multiple vehicles, this constraint can be further relaxed.

The use of RFID technology allows reading multiple tags simultaneously (group of nearby bins), without collisions and re-transmissions, as it would be the case of a wireless sensor network. We note however that the mobile sink is not bound to RFID technology. If more status information must be collected per bin, such as multiple fill level measurements or measurements of different parameters, sensor nodes can be used, like ones based on IEEE 802.15.4 protocol. In this case, a USB port of the system can be used to plug an appropriate node and antenna. In this case, there are issues to be studied regarding to the detection of contact opportunities by nodes in an energy-efficient manner and the transmission of information [22]. In many cases, waste bins are installed as groups (rows) of bins or in proximity (e.g., opposite sides of the road). Further studies are necessary towards exploring energy-efficient policies for

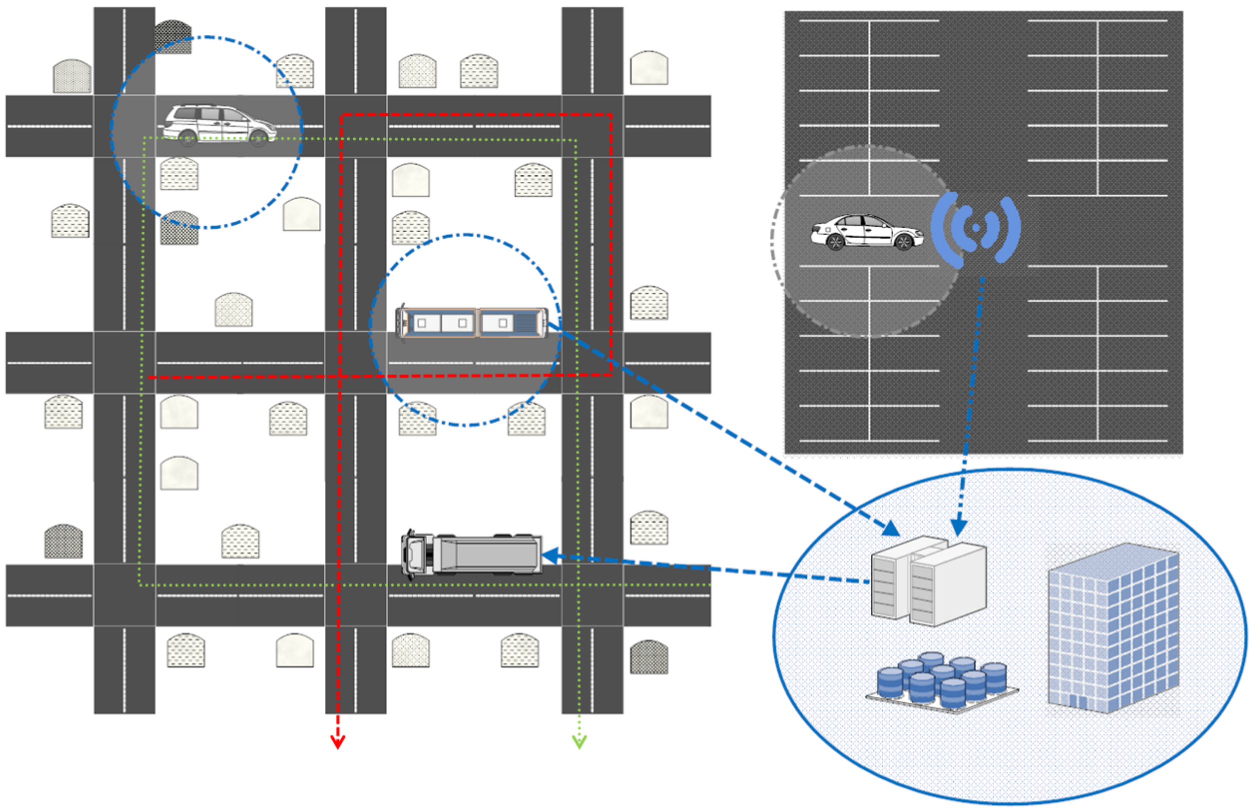


Fig. 1. The Dynacargo approach. Initial service plan (green line) optimized (red line) based on collected bin status information.

combined detection-transmission opportunities with low-duty cycles in such an environment: the mobile sinks may need to swap between beaconing and receiving information so as to allow detection of nearby bins that were not identified in first contact.

Initial simulations with Contiki-based sensor nodes, ContikiMAC, Rime protocol stack, and a mobile sink indicate that the node can detect contact opportunity and transmit information to the passing by vehicles even with a 1% radio duty cycle. This translates to significant energy savings that can lead to battery lifecycles comparable to those of RFID tags. Introducing more complex network protocols, protocol layers and stacks, or application-level logic at this stage of communication must be carefully considered and evaluated, given the episodic nature of communications [23]. Longer packets or complex processing can cause communication disruption and packet corruption, as the vehicle moves away from the bins. This translates in lost opportunities to collect viable information from the field and significant energy consumption for the sensor nodes.

### *B. Backend Communication*

An application running on the mobile sink is responsible for receiving the information from the field units. As explained in the previous, we assume that the vehicle has access to the Internet once it returns to its depot or sporadically through any public WiFi hotspots available in the city. Handling the network issues arising in such an opportunistic environment at the application layer can be troublesome. Our approach is to build the mobile sink with a delay/disruption-tolerant networking (DTN) architecture. A DTN architecture allows extending network connectivity fighting large delays or periods of disconnection, where traditional Internet protocols cannot cope with [24], [25]. In the case of Dynacargo, a DTN protocol stack that places the Bundle Protocol (BP) on top of the transport layer is implemented.

The DTN stack is utilized as follows. The application collects status information from the field and is configured to send them to the backend system through the Internet. There are different options in realizing the application-to-stack communication. One is to develop the application using the DTN API (application programming interface) instead of the traditional network sockets API. Another one is to introduce a transparent DTN proxy/tunnel with one end lying in the mobile sink and the other end in the backend system. Yet another option is to use file interfaces and utilize already-available DTN-enabled applications for sending the files to the other end. The exact stack implementation (e.g., DTN2, IBR-DTN, ION) and tool choice is more a matter of personal taste and familiarity with the DTN technology.

The important aspect of a DTN approach is that the collection application succeeds in sending “immediately” the status information to the network, even when there is no network connection available. Then, the DTN is responsible for handling and restoring networking connectivity, coping with disruptions and bundle (packet) retransmissions, and temporarily storing the information in long-term storage, until a network connection is available. The DTN approach allows to totally liberate the application from management tasks: it needs neither to handle connectivity issues (the network is always on) nor storage (the information bundles will be stored by DTN is necessary). The Linux operating system handles automatic connection with open WiFi hotspots and on top of that the DTN stack transmits to the backend system any pending bundles for further processing.

### C. Concept Validation and Functional Tests

A prototype mobile sink is now implemented. The sink interfaces an active RFID reader and an application is developed that collects the readings and transmits to the backend systems through a DTN tunnel. The DTN functionality is implemented using the DTN2 software, customized and optimized for fitting the constrained resources available. Another instance of the DTN2 software runs on the backend systems in order to receive the bundles sent by the mobile sink. The storage in the sink is realized using a USB flash drive occupying the second available port of the system. This is necessary, as the microSD card will wore out very quickly if small files are written continuously to it.

The prototype system is tested in a controlled environment with emulated network disruptions. A system acting as a WiFi access point exposes a wireless link for random period of times. Upon detecting a WiFi link opportunity, the prototype connects through it to a public server (DTN endpoint) and unloads any collected information so far. A process running on the prototype creates at random point of time set of bundles, effectively emulating the collection of information from waste bins.

## VI. Conclusions and Future Directions

The increasing population concentration in cities and urban areas creates the need for improved waste collection processes. Optimized use of available resources can be achieved by devising service plans based on actual fill level information. At an urban scale, the collection of status information from sensors installed in waste bins poses significant challenges in network setup and operation, both technical and financial. We presented Dynacargo, an innovative smart city application for urban solid waste collection. Dynacargo exploits already available network infrastructure, such as public WiFi hotspots, delay-tolerant networking (DTN) technology, and single-hop RFID-based communications in order to reduce costs, simplify system operation, and support scaling at urban level. We also describe a system design and the developed prototype of the mobile sinks carried on vehicles roaming around the city.

The proposed Dynacargo system will be pilot-tested in the Municipality of Nafpaktia in Greece. The Municipality has an area of 870.38 $\mathrm{km}^2$ and a population of 27,800 people (in 2011 census). Eleven municipal trucks are used for waste collection from the bins. There are about 2,100 solid waste bins installed in about 140 settlements, villages, and towns in the area, including urban, suburban, and rural areas. The pilot tests will allow gaining insights on areas of improvement and experiment with vehicle movement and network connectivity issues in a city terrain.

## Acknowledgment

This work was financially supported by GSRT under grant 11SYN_10_456, project "Dynamic Cargo Routing on-the-go" (Dynacargo).

## References

[1] M. Faccio, A. Persona, and G. Zanin, "Waste collection multi objective model with real time traceability data," *Waste Management*, vol. 31, pp. 2391–2405, 2011.

[2] O. Johansson, "The effect of dynamic scheduling and routing in a solid waste management system," *Waste Management*, vol. 26, pp. 875–885, 2006.

[3] S. Suzuki, M. Saito, Y. Matsufuji, and K. Yamamoto, "Waste amount analysis based on "sphere of users" for waste collection point by mutual exploitation of GPS and GIS," in *Proceedings of Korea-Japan Special Symposium on Technology and R&D Trends on Inorganic Waste Recycling*, 2005, pp. 122–124.

[4] F. Vicentini, A. Giusti, A. Rovetta, X. Fan, Q. He, M. Zhu, and B. Liu, "Sensorized waste collection container for content estimation and collection optimization," *Waste Management*, vol. 29, pp. 1467–1472, 2009.

[5] M. Ghose, A. Dikshit, and S. Sharma, "A GIS based transportation model for solid waste disposal - a case study on Asansol municipality," *Waste Management*, vol. 26, no. 11, pp. 1287–1293, 2006.

[6] F. Reverter, M. Gasulla, and R. Pallà-Areny, "Capacitive level sensing for solid waste collection," in *Proceedings of IEEE Sensors 2003*, vol. 1. IEEE, 2003, pp. 7–11.

[7] A. Fuchs, H. Zangl, G. Holler, and G. Brasseur, "Design and analysis of a capacitive moisture sensor for municipal solid waste," *Measurement Science and Technology*, vol. 19, no. 2, p. 025201, 2008.

[8] P. Leitão and M. Marques, "Reliability of a line-of-view sensor for recycling point waste containers," 2009, available on http://www.iswa.org/uploads/tx_iswaknowledgebase/13-291paper_long.pdf.

[9] A. Rovetta, F. Xiumin, F. Vicentini, Z. Minghua, A. Giusti, and H. Qichang, "Early detection and evaluation of waste through sensorized containers for a collection monitoring application," *Waste Management*, vol. 29, no. 12, pp. 2939–2949, 2009.

[10] M. Hannan, M. Arebey, R. A. Begum, and H. Basri, "An automated solid waste bin level detection system using a gray level aura matrix," *Waste Management*, vol. 32, pp. 2229–2238, 2012.

[11] S. Longhi, D. Marzioni, E. Alidori, G. Di Buo, M. Prist, M. Grisostomi, and M. Pirro, "Solid waste management architecture using wireless sensor network technology," in *New Technologies, Mobility and Security (NTMS), 2012 5th International Conference on*. IEEE, 2012, pp. 1–5.

[12] V. Catania and D. Ventura, "An approch for monitoring and smart planning of urban solid waste management using Smart-M3 platform," in *Proceedings of the 15th Conference of Open Innovations Association FRUCT*. ITMO University Publisher House, 2014.

[13] M. Mamun, A. Al, M. Hannan, A. Hussain, and H. Basri, "Wireless sensor network prototype for solid waste bin monitoring with energy efficient sensing algorithm," in *Computational Science and Engineering (CSE), 2013 IEEE 16th International Conference on*. IEEE, 2013, pp. 382–387.

[14] D. Cassaniti, "A multihop 6LoWPAN wireless sensor network for waste management," Master's thesis, Universit degli Studi di Padova, Italy, 2011.

[15] F. McLeod, G. Erdogan, T. Cherrett, T. Bektas, N. Davies, D. Shingleton, C. Speed, J. Dickinson, and S. Norgate, "Improving collection efficiency through remote monitoring of charity assets," *Waste management*, vol. 34, no. 2, pp. 273–280, 2014.

[16] M. Philipose, J. R. Smith, B. Jiang, A. Mamishev, S. Roy, and K. Sundara-Rajan, "Battery-free wireless identification and sensing," *Pervasive Computing, IEEE*, vol. 4, no. 1, pp. 37–45, 2005.

[17] A. P. Sample, D. J. Yeager, P. S. Powledge, A. V. Mamishev, and J. R. Smith, "Design of an RFID-based battery-free programmable sensing platform," *Instrumentation and Measurement, IEEE Transactions on*, vol. 57, no. 11, pp. 2608–2615, 2008.

[18] L. Zhang and Z. Wang, "Integration of RFID into wireless sensor networks: Architectures, opportunities and challenging problems," in *Grid and Cooperative Computing Workshops, 2006. GCCW'06. Fifth International Conference on*. IEEE, 2006, pp. 463–469.

[19] N. Cho, S.-J. Song, S. Kim, S. Kim, and H.-J. Yoo, "A 5.1-$\mu$w UHF RFID tag chip integrated with sensors for wireless environmental monitoring," in *Solid-State Circuits Conference, 2005. ESSCIRC 2005. Proceedings of the 31st European*. IEEE, 2005, pp. 279–282.

[20] H. Liu, M. Bolic, A. Nayak, and I. Stojmenovic, "Taxonomy and challenges of the integration of RFID and wireless sensor networks," *Network, IEEE*, vol. 22, no. 6, pp. 26–35, 2008.

[21] A. Mitrokotsa and C. Douligeris, "Integrated RFID and sensor networks: Architectures and applications," in *RFID and Sensor Networks: Architectures, Protocols, Security and Integrations*. Auerbach Publications, CRC Press: Boca Raton, FL, 2009, pp. 511–535.

[22] G. Anastasi, M. Conti, and M. Di Francesco, "Data collection in sensor networks with data mules: An integrated simulation analysis," in *Computers and Communications, 2008. ISCC 2008. IEEE Symposium on*, 2008, pp. 1096–1102.

[23] E. Borgia, G. Anastasi, and M. Conti, "Energy efficient and reliable data delivery in urban sensing applications: A performance analysis," *Computer Networks*, vol. 57, no. 17, pp. 3389–3409, 2013.

[24] A. Voyiatzis, "A survey of delay- and disruption-tolerant networking applications," *Journal of Internet Engineering*, vol. 5, no. 1, pp. 331–344, 2012.

[25] ——, "The quest for a killer application for opportunistic delay-tolerant networks," in *Advances in delay-tolerant networks (DTNs): Architecture and Enhanced Performance*, J. Rodrigues, Ed. Woodhead Publishing, 2014, pp. 251–267.

# Architecture and Implementation Issues, Towards a Dynamic Waste Collection Management System

George Asimakopoulos, Sotiris Christodoulou, Andreas Gizas, Vassilios Triantafillou, Giannis Tzimas
Computer & Inform. Engineering Dept.
Technical Educational Inst. of Western Greece
Antirion, Greece
{gasimakop@gmail.com, sxristod@teimes.gr, gizas@ceid.upatras.gr, triantaf@teimes.gr, tzimas@cti.gr}

John Gialelis, Artemios G. Voyiatzis, Dimitris Karadimas, Andreas Papalambrou
Industrial Systems Institute, "Athena" RIC in ICT and Knowledge Technologies, Patras Science Park, Stadiou Str.,Platani, PATRAS, Greece, GR-26504
{gialelis@isi.gr, bogart@isi.gr, karadimas@isi.gr, papalambrou@isi.gr}

## ABSTRACT
Dynacargo is an ongoing research project that introduces a breakthrough approach for cargo management systems, as it places the hauled cargos in the center of a haulage information management system, instead of the vehicle. Dynacargo attempts to manage both distribution and collection processes, providing an integrated approach. This paper presents the Dynacargo architectural modules and interrelations between them, as well as the research issues and development progress of some selected modules. In the context of Dynacargo project, a set of durable, low cost RFID tags are placed on waste bins in order to produce crucial data that is fed via diverse communication channels into the cargo management system. Besides feeding the management system with raw data from waste bins, data mining techniques are used on archival data, in order to predict current waste bins fill status. Moreover easy-to-use mobile and web applications will be developed to encourage citizens to participate and become active information producers and consumers. Dynacargo project overall aim is to develop a near real-time monitoring system that monitors and transmits waste bins' fill level, in order to dynamically manage the waste collection more efficiently by minimizing distances covered by refuse vehicles, relying on efficient routing algorithms.

## Categories and Subject Descriptors
**C.0 GENERAL:** System architectures, **H.2.8 Database Applications:** Data Mining, **G.2.2 Graph Theory:** Graph algorithms

## General Terms
Algorithms, Design, Experimentation.

## Keywords
Urban solid waste collection; intelligent transportation systems; data mining; graph routing algorithms.


*WWW '15 Companion,* May 18–22, 2015, Florence, Italy.
ACM 978-1-4503-3473-0/15/05.
http://dx.doi.org/10.1145/2740908.2742134

## 1. INTRODUCTION
Nafpaktia is a typical example of how solid waste collection is managed throughout the country, as waste collection occurs on regular time intervals by following fixed routes. The waste collection scheme and all related decisions are drawn based solely on personnel experience.

This "rule of thumb" approach sometimes displays an unacceptable result, as it was found that in some cases overfilled waste bins were uncollected for some days, while at the same time other unfilled bins were collected anyway. These inaccuracies result in citizen dissatisfaction and a noticeable cost increase. In the Municipality of Athens[8] it has been estimated that the 60% to 80% of the total cost of waste collection, transportation and disposal is spent during collection. The main factor for reducing the cost is to minimize the distance and duration of the routes [6]. Johansson [7] proved that if the fill level of bins was taken into account and waste collection adapts accordingly, it could reduce the cost of waste collection up to 20%. Dynacargo project (Dynamic Cargo Routing on-the-Go) aims at developing a near real-time monitoring system that monitors and transmits waste bins' fill level, in order to make waste collection more efficient by cost reduction which is accomplished by minimizing distances covered by refuse vehicles.

## 2. RELATED WORK & INNOVATION OF DYNACARGO
At the moment, several approaches exist that tackle issues regarding waste collection management and especially information collection at the point of waste disposal [2,3,4, 5,7,9,10]. However, none of these solutions can be applied as-is in the Dynacargo case. Most of them are expensive and they do not match the Dynacargo functional requirements, while most of them deal with the problem of collecting recyclable waste, which can be collected at less frequent intervals.

Although Dynacargo aims at automating the waste collection process as existing approaches do, it also exhibits major differences when compared against such solutions. The differences arise mainly due to the Dynacargo architecture, which builds on the concept of a generic multipurpose cargo-based dynamic vehicle routing system that takes into account dynamically changing cargo that needs to be collected from disperse points. This implies that Dynacargo does not only serve as a data collection system that transmits data from waste bins to a central server, but utilizes this data in order to achieve optimized vehicle routing in a dynamic manner during the waste collection process via decision making system.

A major breakthrough that Dynacargo introduces is the utilization of a diverse set of data transmission techniques in order to ensure that data is sent from points of collection to the system in the most efficient manner. Rather than relying on GSM, which induces fixed telecom costs, Dynacargo adapts Delay-Tolerant Networking concepts in order to transmit data from disperse collection points to the system. In order to achieve this, a set of existing public commuters serve as data hosts that transport data as they execute unaltered standard procedures.

## 3. NAFPAKTIA AS CASE STUDY

The functional requirements of Dynacargo resulted from the analysis of the solid waste collection and management system of Municipality of Nafpaktia, Greece, which expands in an area of 870,38 $km^2$ and displays a population of 27.800 citizens. It has geographic and demographic peculiarities that make it ideal as a pilot Municipality for Dynacargo. It includes a coastal city with narrow busy streets in the historic center, coastal towns and villages whose population increases substantially during the summer months, but also mountain remote villages (up to 120 km. away) with a few dozen residents. The network coverage and frequency of public bus routes are very diverse among the areas of the municipality. Finally, it supports the transshipment of waste from a small to a big refuse truck. An as-is analysis was performed initially in order to formally document the current waste management process. A set of indicative waste collecting routes were selected in a way that ensures the incorporation of peculiarities and modeled on Google Earth (Figure 1). The analysis included available waste collection historical data and information regarding related processes.

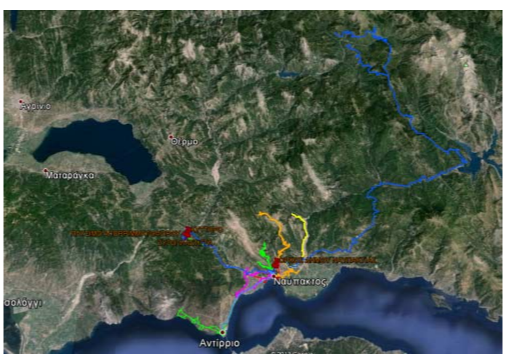



**Figure 1. Indicative waste collection routes for Nafpaktia.**

## 4. USERS & USE-CASE MODELS

The main user groups realized throughout the analysis are: waste collection related municipal services, system administrators, truck drivers, bins' data collectors and individual citizens.

In order to reassure the accurate Dynacargo operation when used as a prediction system, information regarding waste bin fill level must be acquired at brief intervals from as many as possible waste bins. In order to achieve this goal, existing organized transportation systems are utilized such as public bus services, postal office vehicles, taxis, municipal police vehicles, along with active social entities, as these vehicles traverse the area of interest regularly. A Data Collector can be either a vehicle driver (e.g., bus, taxi, postman car, etc.) equipped with the Dynacargo equipment, or anyone else who may be involved in data collection (e.g., postmen), who may use any means of transportation such a car, a bike, or by foot. Moreover, citizens can improve the operation of the system by reporting fill-level through a mobile application, reporting estimations of produced waste volume on unforeseen events and by checking online the bin fill levels near their residence so as to discard their waste on nearby unfilled bins.

After collecting and analyzing the users' functional requirements, we have constructed several application scenarios that are divided into two categories: *Full-scale Application Scenarios* and *Pilot Scenarios* (scenarios that will be demonstrated during the project). Based on the full-scale scenarios, use-case models defining use cases for each subsystem were constructed, including the vast majority of functional requirements along with some non-functional requirements. We have constructed six major use-case models, modeling different usage cases of the system: Bins' Data Collector, Refuse Truck, Central System, Optimal Routes Calculation, Citizen's SmartPhone App and Citizen's Information Portal. These use cases and more detailed description of the functional requirements of Dynacargo, are described in a previous research work [1].

## 5. ARCHITECTURE

Based on the functional requirements and the use case models, the most significant Dynacargo structural components were pinpointed and designed and furthermore decomposed into subsystems and smaller functional units. Operations, roles and significant logical relationships with other subsystems were defined for each subsystem.

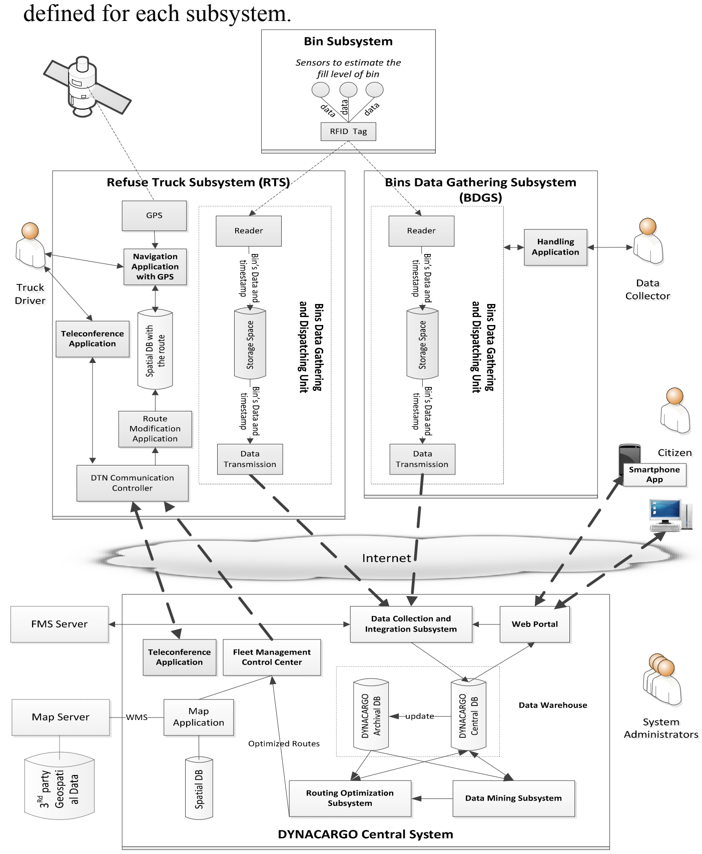



**Figure 2. Dynacargo Architecture.**

Dynacargo specifications were derived after thorough analysis of the waste collecting process as it is implemented by Nafpaktia municipality. However, the Dynacargo architecture is designed in a manner that allows its utilization as a generic cargo-based routing system along with the inherited ability to be adapted to any other municipality regardless of specific waste collecting process characteristics. The Dynacargo architecture is presented in figure 2 and in more detail in [1]. Following, we shortly describe its main subsystems and architectural parts.

**Bin Subsystem:** Special equipment (sensors) that estimates the waste bin fill level and collects, stores, and transmits bins data.

**Bins Data Gathering and Dispatching Unit:** A unit that communicates through a short range network with the Bin Subsystem and dispatches the collected information (Bin's Data) to the Central System, through a long range network. It can be installed at passing vehicles and consists of three components:

- *"Reader"*, which refers to the communication interface with the bin subsystem.
- *"Storage space"* which temporarily stores the data until transferred to the Central System.
- *"Data transmission"* to the Central System, through Delay Tolerant Networking (DTN) technology.

**Bins Data Gathering Subsystem (BDGS):** The BDGS is the subsystem carried by the Data Collector. It consists of the Bins Data Gathering and Dispatching Unit, combined with a very simple application for handling it. The BDGS should be portable and able to work with a battery.

**Refuse Truck Subsystem (RTS):** This subsystem will be installed on each refuse truck. The main part of the RTS is the Bins Data Gathering and Dispatching Unit. Additionally the RTS will be equipped with a GPS unit and a camera and will incorporate the appropriate software to provide the following functionalities:

- *GPS Navigation Application:* It is a classical navigation application through GPS, which will provide navigation guidance to the truck driver, and instructions about which bins should be collected.
- *Teleconference Application:* Teleconference communication of driver with the Central System, especially for reporting emergency situations.

The RTS will be powered by the refuse truck. The **communication** with the Central System will be based on DTN technology, which will transparently select the most suitable communication method, considering criteria like cost or speed.

**Citizen Application:** An application for mobile devices (smartphone, tablet), through which the citizen can choose any bin and report data for it (fill level, photo, comment, etc.).

**Central System:** It is the back-end system of Dynacargo. Its main part is the Data Warehouse storing all historic data, bins' data, vehicles' data and everything needed for calculating the best routes based on the current load of bins and some restrictions (Bin Collection Settings) specified by the system administrators. Moreover, Data Warehouse will store any information derived from the other subsystems, particularly from the Data Mining Subsystem and the Routing Optimization Subsystem.

Apart the above databases, the Central System will include:

**Fleet Management Control Center:** receives the optimized routes from the Routing Optimization Subsystem, two hours before the start of the routes and on the fly (if the routes have to be modified), and forwards them to the appropriate RTS. All spatial information for the corresponding route will be stored locally in a Spatial RTS Database. The Fleet Management Control Center can send route modifications to any refuse truck, informing properly the Navigation Application, if such need arises and provided that the refuse truck has network coverage.

**The Map Application:** A cartographic JavaScript API that will provide all the required functionality for creating rich-web applications based on geographic and descriptive data.

**Data Collection and Integration Subsystem:** This subsystem will collect bins data from the RTS, the BDGS and citizens' applications, and integrate them into the central DBs. Furthermore, it will be able to integrate data from other sources, if needed, like an independent FMS server that can operate in parallel with the Dynacargo system. The architecture design of Dynacargo does not require the existence of a classical FMS.

**Data Mining Subsystem:** Mainly used to estimate fullness of bins when we do not have the available information updated or is not fairly recent. It generally seeks to apply fusion techniques from multiple sources to produce more semantically rich data, thereby obtaining information on a higher level of abstraction.

**Routing Optimization Subsystem:** The role of this subsystem is the dynamic route planning before the starting of truck routes, and the on-the-fly modification of routes, either due to exceptional events (accident), or if the new data collected during the routes impose such changes. The routes will be calculated two hours ahead, taking into account various data that will exist in the central data warehouse system.

**Citizens Web Portal:** A web site, through which citizens can be informed about the current completeness of bins or report bins' data (fill level, photography, etc.) in the Central System, by selecting a bin on a map. The purpose of the Citizen Web Portal and the smartphone app are to motivate people to participate in sustainable waste management.

In the three following sections we further discuss the current design and implementation issues of three important subsystems.

## 6. BIN SUBSYSTEM

The bin subsystem consists of three building blocks: the fill sensing unit, the active RFID tag for data transmission, and the protective enclosure. The fill level estimation of urban solid waste bins appears to be a challenging a task. The irregular shape and the variety of the materials require advanced sensing approaches. The harsh environmental conditions (e.g., humidity, temperature, and dust) can significantly affect the sensor measurement accuracy and reliability.

A comparison of various solutions including infrared proximity sensors, optical sensors and ultrasonic sensors, indicated that ultrasonic sensors appear to be the most suitable solution for the purpose of the presented architecture taking into account the aforementioned conditions. The ultrasonic sensors are advantageous in providing ranging measurements independently of the contained objects, thus making possible the corresponding translation into fill level measurements. Since the resolution of the ultrasonic sensors is only centimeters (less than an inch) the selected solution can offer fine-grained accuracy. The ultrasonic sensors should be mounted in the bin lid, exposing only a small part of the sensor body. Since the sensors will operate unattended in the field, low power consumption models that also offer IP-67 protection rating are to be used. An analog output is provided for connecting the sensing units to the data transmission unit of each bin, thus the active RFID tag.

Active RFID tags have been selected as the data aggregation and transmission unit for the bin subsystem. The selected tags can operate with a standard 3.3V battery providing a lifetime of more than 5 years (that equals to millions of beacon transmissions). The employed active tags offer extra I/O pins for communicating with external devices, thus the ultrasonic sensing units. This setup allows powering the sensors and the RFID tag from the same source, either tag's battery or an external power source. The tag itself supports various operating modes, including standard beacons at programmable time intervals, sleep, wake up at regular intervals and also wake up at external trigger. The combination of the above operating modes allows the extension of the entire bin subsystem energy lifetime, since minimal power consumption occurs when the tag operates in sleep mode and the sensor is not powered up. When the tag is awake, it powers up the sensors and temporarily stores their values in its internal memory.

In the Dynacargo scenarios the tag wakes up at predefined times of the day, depending on the location of the waste bin installation and on the desired measurement frequency defined by the Municipality operators.

The Dynacargo operation does not assume availability of a fixed network infrastructure that reaches all the installed bins. At an *urban scale* of operation, this is a quite realistic assumption, since *thousands* of sensors are *sparsely* deployed in a complex city and suburban terrain.

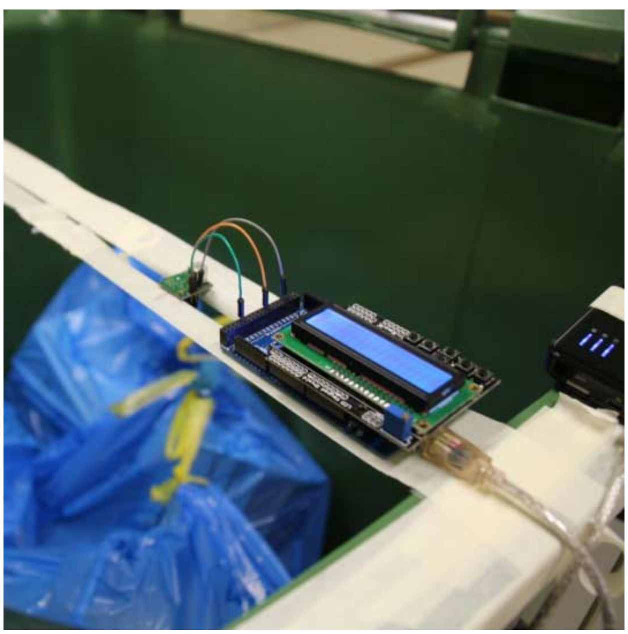


**Figure 3. Experimental setup with sensors and an Arduino board for control**

System installation and maintenance, ensuring radio coverage, and retaining network formation can become an unmanageable task. The Dynacargo project opts for low-range, point-to-point communications based on RFID technology so as to cope with these issues. Vehicles roaming around the city and equipped with readers collect the information from the bins. In order to cope with increased telecommunication costs and infrastructure upgrades, these mobile sinks defer transmissions until an Internet connection becomes available. The Delay Tolerant Networking paradigm ensures that the information is retained in the bins, until a mobile sink passes nearby and then in the sinks, until an Internet connection is available.

# 7. DATA MINING & PREDICTION ANALYSIS SUBSYSTEM

The data mining and prediction analysis subsystem is based on the data warehouse infrastructure, from which the data are extracted, so that the various data mining scenarios can be realized. The data mining process is quite exploratory, as the parametrization of the applied methods and their analysis in order to select the more efficient ones, is a process quite empirical. The research team tried to execute different scenarios in order to analyze the behavior of existing data mining methods and provide the rules that arise wherever is possible.

**Data mining algorithms study and results visualization**

The procedure followed in order to study and fine tune the data mining algorithms within the scope of the project, as well as the visualization of the results, is discussed below. The whole subsystem is realized in a Microsoft SQL Server RDBMS, utilizing the relevant Analysis Server (Microsoft Analysis Services), where a Mining project has been created in the supported programming environment, linking the central database with the Analysis Server.

At the first stage, a data mining structure was constructed based on a single table, merging the data from the various RDBMS tables. When creating the mining structure the variables included in the analysis were chosen. Furthermore, each variable type was defined in order to be effectively handled by the data mining algorithm. For example, some variables are continuous space such as dates, some are distinct and some must be discretized for the sake of the analysis at specific intervals. Also, a variable must be defined as the key which imposes distinctness between different records.

In the scope of the subsystem a clustering model was defined, as well as variants of it with different parameters, aiming to reveal the general distribution range of variables in the entire range of data. From the different clusters we got from the analysis, we tried to define which variable(s), make the distinction between them and whether it makes sense to separate the data into clusters.

In parallel with the above, a study was performed using other data mining tools, because not all methods were fully covered by the selected configuration. The same data were imported in WEKA which provides a richer set of data mining methods, and also additional algorithms to some of the methods. With this configuration we attempted to find the logical expressions that link data together. Also, a classification took place according to the distance between the waste bins and investigated how this affects the total cost.

Besides the above, at the Analysis Server a cube was constructed in order to visualize the data. The cube has as many dimensions (columns) as the data we want to use. Queries on these data are performed utilizing the PivotViewer control, a Silverlight web browser plug-in. PivotViewer was used to implement the main querying interface, since it leverages Deep Zoom which is the fastest, smoothest, zooming technology on the Web. As a result, it displays full, high resolution content without long loading times, while the animations and natural transitions provide context and prevent users from feeling overwhelmed by large quantities of information. The PivotViewer enables users to interact with thousands of objects at once, and sort and browse data in a way that helps them see trends and quickly find what they are looking for. In PivotViewer, variables filters can be applied depending on the range of each one. Filters can be applied simultaneously to many variables. Besides filtering, the data can be sorted on any variable without the need of a filter. For example, to show the fullness of the waste bins for a specific week of a month in some specific time range, the appropriate filters can be selected and sort the data by date.

# 8. ROUTING OPTIMISATION

In this section we describe the static and dynamic data we decided to store and calculate for each bin, we present how we construct the graph models, we highlight the routing particularities of Dynacargo and match them to known families of routing problems that present similar characteristics and technical limitations.

## 8.1 Static and Dynamic Bins' Data

For each bin we store the following static data: longitude, latitude, altitude, bin type, bin capacity. Moreover, we identify some dynamic data for each bin that can be provided either by system administrators or other subsystems, such as Bins Data Gathering Subsystem or the Data Mining Subsystem:

- Bin fill-level that it is collected many times daily from several collectors.
- Bin fill-level predicted value at the time of routes, provided by the Data Mining Subsystem.
- Actual bin fill-level collected by the refuse truck. This value is compared to the predicted value and used for more accurate future predictions.
- Maximum time that a bin cannot be serviced, depending on the seasonal period.
- Specific days and / or hours that the bin must be serviced, depending on the seasonal period.
- Bin priority (1 to 3). Based on other dynamic data and the seasonal period, the priority to service a bin will be: 1 (do not service), 2 (desirable but not necessary to service it), 3 (must be serviced). The algorithm that calculates the priority for each bin can be parameterized by system administrators.

## 8.2 Graph Model

Because most non-urban municipalities in Greece contain geographically scattered villages or towns, our model will group nearby bins into bin clusters. For each pair of bins within a cluster the actual distance between them is calculated considering traffic restrictions, in both directions. Moreover, even if the distance of a bin pair in both directions is the same, the travel costs may not be. For each bin cluster, we set one or more entry / exit points. The entry / exit points of all clusters, the landfill location and vehicle parking spaces are modeled as the nodes of another graph. For each pair of nodes in this graph, the actual distance between them is calculated considering traffic restrictions, in both directions.

The calculation of actual distance between two geographic points is carried out by using the QGIS tool and the Open Street Map via the online routing api that provides. We are based on open standards and the commonly used WGS84 coordinate system, thus this procedure could be also performed by using the routing API of Google or Bing Maps.

## 8.3 Dynacargo Routing Problem

2100 bins and 2 bin types. 11 refuse trucks, 6 kinds of trucks. Some trucks are servicing only one bin type. Only one small truck can service bins in the historic center of Nafpaktos, because of narrow roads. This small truck is also used to collect other small bins, some of which are located at long distances, mainly due to lower fuel consumption. This small truck does not go to the landfill, but tranships its waste cargo to a larger refuse truck.

All routes should begin and end in a truck parking, and in our system there can be more than one parking areas and selected parking areas for each refuse truck. This information is provided for each refuse truck, as well as which trucks are available every day. Each truck route (except the route of the small truck) must pass from the landfill once and then move directly to a parking area without servicing other bins.

## 8.4 Our Algorithmic Approach

Currently, we are designing our routing algorithms in order to implement and evaluate them. In this section we outline our approach. At first we are going to solve the subproblem of the small truck which will service all small bins with priority 3, that are located in the historical center of Nafpaktos and are the most distant ones. Depending on the available capacity of the truck, assign to its route some extra small bins of priority 2, starting from those that are closer to the initial route.

Afterwards, reset the load of all small bins that belong to the small truck route, and solve the complete problem with the other available trucks, where the small truck appears as a separate node with its load. The exact location of this node should be on the route from the last bin serviced by the small truck and its depot. The optimal position of this node (which is the location of the waste transship to a bigger truck) will be calculated by the algorithm. The algorithm should synchronize the routes of the two trucks so as to minimize the waiting time of the truck that arrive first at the transship location. The complete problem will be solved in two levels: within each cluster and inter-cluster.

In overall, the main routing problem is defined as finding a set of optimal routes (lowest total km and fuel consumption) for a subset of the available trucks, that begin at parking areas, go through each bin of priority 1 at least once, and end up in landfills and then in a parking area, satisfying the capacity constraints of the available trucks and the maximum time of any route. It should also be taken into account the issue of transhipment of the small truck. Depending on the available capacity of each truck, assign to its route some extra bins of priority 2, starting from those that are closer to the initial route. At this point, we will examine ways to optimize the selection of best bins of priority 2 (or even 1), relying on the prediction of these bins fill-level for the next days.

## 8.5 Assigning to known problem families

This specific problem is a variant of the Capacitated Vehicle Routing Problem (CVRP) where the approach methods are grouped into the following main categories [11]: Branch-and-Bound, Branch-and-Cut (hybrid methods between Branch-and-Bound and Cutting Plane methods), Set-covering based algorithms, Heuristics and Metaheuristics.

Furthermore, Dynacargo routing problem can also be approached as an Orienteering Problem (OP). The OP is a combination of vertex selection and determining the shortest Hamiltonian path between the selected vertices. As a consequence, the OP can be seen as a combination between the Knapsack Problem and the

Travelling Salesperson Problem (TSP). The OP's goal is to maximize the total score collected, while the TSP tries to minimize the travel time or distance. Furthermore, not all vertices have to be visited in the OP. Determining the shortest path between the selected vertices will be helpful to visit as many vertices as possible in the available time. The main categories of OP are the following: [12]: Orienteering problem (OP), Team Orienteering Problem (TOP), Orienteering Problem with Time Windows (OPTW), Team Orienteering Problem with Time Windows (TOPTW).

For the implementation of our algorithms, we will be based on the library Or-Tools of Google, possibly in combination with other tools. The aim of this library, unlike other constraints programming libraries, is not to provide a complete set of constraint-based algorithms solutions. Conversely it provides a tool for initial analysis and programming on which we can build solutions for more specific problems.

## 9. CONCLUSIONS

Intelligent transportation systems constitute key components for ecologically sustainable development in urban spaces. Dynacargo project aims at developing a cargo-centric transport management system and demonstrates it in the case of urban solid waste collection management. Dynacargo extends and expands existing fleet management system functionality in two directions. The first direction is to fuse into the monitoring and decision support process near real-time waste related information (fill level of waste bins) before refuse truck visits a collection point. Alternatively, if this information is impossible or not justified on a cost-benefit basis, historical information is utilized in order to predict waste bins fill status, using data mining technics. The second direction is to encourage end users (i.e. citizens) to participate and become active information producers and consumers. Dynacargo will utilize low-cost, durable units as RFID tags, explore alternative network protocols like DTN, will be 4G-ready and will utilize dedicated dynamic routing algorithms in order to minimize telecommunication and hardware costs. In this paper, we outline the system functional requirements derived from the needs of the main user groups, illustrate some use case models and present the major architectural parts of Dynacargo, their decomposition into subsystems and smaller modules. Now we are in the phase of deciding the implementation details of each module and integrate them in a compact system.

## 10. ACKNOWLEDGMENTS

This work has been financially supported by the Greek General Secretariat of Research and Technology and European Union under the project "11SYN_10_456: Dyanacargo".

## 11. REFERENCES

[1] Asimakopoulos G., Christodoulou S., Gizas A., Triantafillou V., Tzimas G. and Vasilopoulos N. (2014). Functional specifications and architecture of a dynamic routing system for urban waste collection using real-time information. IADIS 13th International Conference on WWW/Internet, 25-27 Oct 2014, Porto, Portugal, pp. 203-210.

[2] Al Mamun M.A., Hannan M.A., Hussain A., Basri, H. (2013). Wireless Sensor Network Prototype for Solid Waste Bin Monitoring with Energy Efficient Sensing Algorithm. *IEEE 16th International Conference on Computational Science and Engineering*.

[3] Arebey M., Hannan M A, Basri H., Begum R A, Huda A. (2010). Solid Waste Monitoring System Integration based on RFID, GPS and Camera. *Proceedings of IEEE Intelligent and Advanced Systems (ICIAS)*

[4] Faccio M, Persona A and Zanin G. (2011). Waste collection multi objective model with real time traceability data. *In Waste Management, Elsevier,* Vol. 12, No. 2,pp 391-405.

[5] Gomes T., Brito N., Mendes J., Cabral J. and Tavares A. (2012). WECO: A wireless platform for monitoring recycling point spots. *Proceedings of 16th IEEE Mediterranean Electrotechnical Conference (MELECON)*.

[6] Gukhool O., Bokhoree C., Mohee R. and Jamnejad G. (2008). Solid Waste Collection Optimization using a Route Time Approach: A case study for an urban area of Mauritius, *23rd Int. Conf. on Solid Waste Technology and Management, Philadelphia PA, p501*.

[7] Johansson Ola (2006). The effect of dynamic scheduling and routing in a solid waste management system. *In Waste Management,* Vol. 26, No. 8,pp 875 – 885.

[8] Municipality of Athens. (2003) Estimation, Evaluation and Planning Of Actions for Municipal Solid Waste Services During Olympic Games 2004. Municipality of Athens.

[9] Nuortio T, Kytöjoki J, Niska H and Bräysy O (2006). Improved route planning and scheduling of waste collection and transport. *In Expert Sys Appl*, Vol. 30, No. 2,p 223-232.

[10] Roshan I. and Akshai M. (2013). SVASTHA: An Effective Solid Waste Management System For Thiruvalla Municipality in Android OS. *Proc. of IEEE Global Humanitarian Technology Conf. South Asia Satellite*.

[11] Toth P. and Vigo D. (2001). The vehicle routing problem. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.

[12] Vansteenwegen P., Souffriau W., Oudheusden D. V. (2011). The orienteering problem: A survey. European Journal of Operational Research, Vol. 209, Issue 1, pp. 1-10.

# A Comparative Study between SVM and Fuzzy Inference System for the Automatic Prediction of Sleep Stages and the Assessment of Sleep Quality

Ch. Panagiotou(*), I. Samaras, J. Gialelis, P. Chondros, D. Karadimas
Electr. & Comput. Eng. Dept. University of Patras, Greece
(*)Corresponding author email: cpanag@ece.upatras.gr

**Abstract—this paper compares two supervised learning algorithms for predicting the sleep stages based on the human brain activity. The first step of the presented work regards feature extraction from real human electroencephalography (EEG) data together with its corresponding sleep stages that are utilized for training a support vector machine (SVM), and a fuzzy inference system (FIS) algorithm. Then, the trained algorithms are used to predict the sleep stages of real human patients. Extended comparison results are demonstrated which indicate that both classifiers could be utilized as a basis for an unobtrusive sleep quality assessment.**

***Keywords—EEG, sleep stages, SVM, FIS***

## I. INTRODUCTION

The relevance of sleep abnormalities with chronic diseases and inflammatory conditions such as depression, heart disease, obesity, diabetes, stroke and arthritis has been already manifested in the literature [1-3]. Currently, there is a significant number of people suffering from sleep disorders, like insomnia, narcolepsy, sleep apnea, etc. The constantly increasing number of people facing sleep abnormalities intensifies the strong relation between sleep quality and quality of life [4-5].

Polysomnography test has been during the past years the dominant tool for sleep quality monitoring and assessment. However, this test requires specialized equipment and can be performed only within a certified lab. Thus, patients cannot repeat this test routinely and in their convenience (i.e. at their home environment). The appropriate, non–obtrusive way for assessing the sleep quality refers to subjective metrics and methods such as the Pittsburgh questionnaire (PSQI)[6].

Sleep is a dynamic phenomenon which is characterized by individual sleep stages. These sleep stages alter during person's rest sessions and contribute towards his sleep cycle formation. The two main sleep stages are the Random Eye Movement stage (REM) and the non – REM stage (NREM). More precisely, NREM can be further distinguished in 4 stages. Stages 1 and 2 are denoted as light sleep phase while stages 3 and 4 as deep sleep or slow waves phase. These stages and the related brain wave frequencies are presented in Table I. The transitions from stage to stage and the duration of each stage during a person's sleep are the main markers for sleep quality assessment.

According to literature quantitative analysis of sleep electroencephalography (EEG) data can provide valuable additional information in sleep research [7].These data recordings are considered a reliable method of assessing a person's sleep stages. However, recent evolution in artificial intelligence has encouraged new efforts on the detection of sleep stages and finally the assessment of sleep quality through machine learning algorithms. Therefore, clinicians and researchers' effort has nowadays been focused on analyzing and extracting enriched features that could feed classifiers, in order to produce efficient and accurate models for the identification of each person's sleep cycle.

TABLE I. SLEEP STAGES AND BRAIN WAVES

| Stage | Frequency (Hz) | Brain Wave |
|---|---|---|
| Awake | 13-30 | Beta |
| 1 | 8-13 | Alpha |
| 2 | 4-13 | Spindle and theta waves |
| 3 | 2-4 | Spindle and Delta waves |
| 4 | 0.3-2 | Delta waves |
| REM | 13-30 | |

This paper is organized as follows. Section II demonstrates the related work and Section III analyzes the methodology employed in this study. The experimental results obtained by applying two different machine learning algorithms are presented in Section V while Section VI concludes the paper.

## II. RELATED WORK

A performance comparison among popular classifiers for the detection of sleep stages is presented is presented in [8]. More precisely, SVM ensemble and Random forest has been tested on ten healthy subjects. In this study the random forest algorithm fed with spectral linear features has outperformed SVM.

An artificial neural network approach reaching 76% performance of identifying stages 1,2,3,4, REM and wake is presented in [9]. The reformation of stages in three larger groups such as (wake), (stage 1, stage 2, REM), (stage 3 stage 4) increased the performance by 82%.

The authors of [10] present another comparison of sleep stage classification by testing the performance of k – Nearest

Neighbor (kNN), Quadratic Discrimminant Analysis and SVM. In these experiments the SVM achieved the most accurate classification by identifying correctly the 73.1% of the stages on healthy subjects and 76% on subject with obstructive apnea.

Finally in [11] an SVM classifier is applied on the proposed features that derive from detrended fluctuation analysis on the ECG (MIT – BIH polysomnography database), achieving a classification rate of 80%.

However the clinical golden standard so far has been the manual scoring from medical experts while the applications and devices that estimate sleep quality indices, based on actigraphy, have not been proven yet as reliable enough to produce accurate and significant outcomes.

In the current work we present the main guidelines for a sleep stage prediction system used as a primary screening and unobtrusive tool for sleep quality assessment. For this purpose, we have compared two dominant machine learning algorithms, namely the support vector machines (SVM) and the fuzzy inference systems (FIS). Both of them present significant advantages and tradeoffs. While many prediction techniques have been reviewed side – by – side for the sleep stage classification problem, a straightforward comparison among FIS and SVM on the same basis has not been observed. A feasibility study concerning the reliable sleep stage assessment using these two algorithms is the main aim of the current study.

## III. METHODOLOGY

### A. Data

The extraction of reliable and accurate models based on data mining and machine learning techniques requires numerous datasets. However, data mining and machine learning in the healthcare domain lacks of data availability. On that context public databases, that make available medical data, try to address this issue. For our study the MIT-BIH Posylomnographic Database [12] which is available from Physionet [13] has been employed. This database contains physiological signals that were recorded during the sleep session of 16 subjects. In total the database consists of 18 records with over 80 hours of polysomnographic recordings. The recorded signals are the electrocardiogram (ECG), EEG, Electroocculogram (EOG) and respiration rate. Since our effort focuses on the development of a non - obtrusive system, the usage of a single physiological signal was a demanding specification. Thus, the EEG signal was considered as the main signal that manifests the sleep activity. EEG records have been recorded with a sampling rate of 250Hz and have been annotated by medical experts every 7500 samples (30 seconds).

### B. Feature extraction

The brain activity is captured in the EEG signals as voltage alterations that hardly exceed a threshold of 100μV. The low amplitudes of an EEG signal are prone to increased signal – to – noise ratio. EEG is highly affected by surrounding signals such as body movement, eye blinking, ECG and the power line inference. Therefore a preprocessing step is required for the extraction of these artifacts from the EEG. On the employed EEG signals a band pass filter with cutoff frequencies at 0.3Hz and 40Hz has been applied.

In order to extract meaningful and semantic information from the EEG signal a further processing stage was conducted. This stage produced features that have been extracted from the frequency and the time domain. Since the sleep stages are strongly related with the brain waves (presented in Table I) the respective frequency bands and their power spectrum have been extracted with a 512 samples Hanning window and 50% overlap. Some further statistic calculations have been applied for the extraction of the frequency, with the higher power on each epoch, and the median frequency spectral power.

Statistical and time domain analysis has been proven to extract useful characteristics that expose important patterns of the brain activity. Based on a thorough study of previous works [14], [15] we extracted the time domain features that has been proved to present the higher degree of correlation with sleep activity. This processing resulted in the extraction of the respective following features:

- Hjorth Mobility & Complecity
- Kurtosis & Skewness
- Interquartile range
- Maximum, minimum, mean and range
- Variance standard deviation
- Shannon Entropy
- Zero Crossing Points (strongly related with presence of spindles)
- Mean absolute and median absolute deviation

Finally our datasets have been completed with the auto regressive filter coefficients extracted from the 6$^{th}$ order auto regression analysis of the signals. All the processing and feature extraction has been applied on non – overlapping EEG epochs of 30 seconds duration.

### C. Classification Strategy

For the final sleep stage prediction, we have employed classification techniques from the machine learning domain, as it has been already mentioned. Especially we focused on a performance comparison between two well established and popular techniques of supervised learning, namely the SVM and the FIS.

The cornerstone of every machine learning classification approach are the data that feed and train the classifiers. However, medical data present very low availability for the researchers. The datasets constructed from the MIT – BIH sleep database offer a sufficient quantity but the classes are not uniformly distributed so as to construct an ideal and unbiased dataset. In total, the dataset consisted of 10181 instances (each instance refer to on 30 seconds EEG epoch) with the following distribution of classes:

TABLE II. CLASSES DISTRIBUTION

| Groups | Sleep Stages | # Stages | # Groups |
|---|---|---|---|
| Wake | Wake | 3120 | 3120 |
| Light Sleep | REM | 700 | 6397 |
| | Stage 1 | 1814 | |
| | Stage 2 | 3883 | |
| Deep Sleep | Stage 3 | 483 | 664 |
| | Stage 4 | 181 | |
| Total: | | 10181 | 10181 |

In order to overcome such problems that cause overfitting issues we used the k-fold Cross Validation technique for evaluating the classifiers. In k-fold cross validation the training set is randomly divided into K disjoint sets of equal size with similar class distribution in every set. Then the classifier is trained with the respective k-th training set while its performance is evaluated with the respective test set that was held out. Finally the estimated performance metric is the average of the values obtained from the k folds. A second approach we followed in order to increase as possible the number of instances from each class was to group them in three classes by joining the classes with common characteristics (clinical, qualitative and quantitative). For example Stage 3 and 4 both share the presence of dominant slow waves, while in stages 1 and 2 are occur theta waves.

The classifiers' performance have been evaluated from the study of the respective confusion matrices derived from the experiments along with the accuracy metric, defined as: $accuracy = \frac{correctly\_identified\_records}{total\_records}$. Further popular metrics used for the evaluation of the classification performance of multiclass classifiers are recall (rec) and precision (prec) defined as: $rec = \frac{TP}{TP+FN}$, $prec = \frac{TP}{FP+TP}$, where TP: True Positive, TN: True Negative and FP: False Positive (respective metrics for binary classification are sensitivity and specificity)[16].

*1) FIS*

A fuzzy rule-based expert system contains fuzzy rules in its knowledge base and derives conclusions as outputs from the user inputs and the fuzzy reasoning process. All these features constitute a fuzzy inference system (FIS) [17].

In this study, the learning algorithm introduced in [17] was used in order to automatically derive the membership functions and the fuzzy IF/THEN rules from the real EEG data together with its corresponding sleep stages.

In particular, initially the subtractive clustering algorithm [17] was utilized for separating the training EEG data together with its corresponding sleep stages into clusters. This algorithm does not involve any iterative nonlinear optimization, and therefore is robust and fast. The following value was defined for each training instance $P_s = \sum_{j=1}^{50} e^{-a\|L_s - L_j\|^2}$ where $s \in [1,50]$, $a$ is a positive constant (here $a = 0.5$), and $L_s$ denotes the multidimensional real numerical data of the $s$th training instance. Then, the procedure described in [5] has been utilized. The constructed FIS parameters are listed in Table II.

*2) SVM*

SVM algorithms have offered great impact on the evolvement and the application of machine learning in general. SVM consider the data points as vectors in a high dimensional space and tries to estimate the optimal hyperplane that separates the data in the respective classes. This properties describe SVM as a binary linear classifier. The linearity however can be overridden through the adoption of the kernel methods instead of vectors. Kernel methods maps the data points of the training sets to hyperplanes that may offer better separation among the classes of the data.

TABLE II. PARAMETERS OF THE FIS.

| Parameter | Value |
|---|---|
| AND method | Algebraic product |
| OR method | Probabilistic OR |
| Implication method | Algebraic product |
| Aggregation method | Max |
| Type of membership functions | Gaussian |
| Fuzzy inference method | Sugeno |
| Defuzzifier | Weighted average |

SVM present high degree of generalization performance in many problems but they add also significant computational complexity during the training phase. This drawback has been partially addressed in [18] with the Sequential Minimal Optimization that we employ in our study along with the polynomial kernel $K(x, y) = (x^T + c)^d$ where x and y are vectors of the feature space and c a constant and d a positive integer.

## IV. EXPERIMENTAL RESULTS

As already described our evaluation strategy was based on the k – fold Cross Validation (k = 10) on a multiclass dataset with 3 classes (Wake), (S1, S2, REM) and (S4, S5). The classifiers were tested initially on the whole dataset with the 10181 instances. The confusion matrix for the FIS and SVM classifiers are depicted on Table III and Table IV respectively.

TABLE III. FIS CONFUSION MATRIX WITH UNBALANCED DATASET

| Classified as: | W | (S1,S2,REM) | (S3,S4) | Total |
|---|---|---|---|---|
| W | **2160** | 940 | 20 | 3120 |
| (S1,S2,REM) | 1076 | **5304** | 7 | 6397 |
| (S3,S4) | 30 | 264 | **370** | 664 |
| **Total** | 2980 | 6774 | 427 | 10181 |
| **Performance Metrics** | rec=69% prec=72% | rec=83% prec=78% | rec=56% prec=87% | Acc= **77%** |

TABLE IV. SVM CONFUSION MATRIX WITH UNBALANCED DATASET

| Classified as: | W | (S1,S2,REM) | (S3,S4) | Total |
|---|---|---|---|---|
| W | **2147** | 966 | 7 | 3120 |
| (S1,S2,REM) | 487 | **5775** | 135 | 6397 |
| (S3,S4) | 7 | 233 | **424** | 664 |
| **Total** | 2641 | 6974 | 566 | 10181 |
| **Performance Metrics** | rec=69% prec=81% | rec=90% prec=83% | rec=64% prec=75% | Acc=**82%** |

From this first round of experiments we observe that we get a classification accuracy of 82% for the SVM classifier while the FIS reaches the 74%. From a more thorough study of the two confusion matrices we observe how the unbalanced training dataset is highlighted on the classification stats for each class. The FIS identifies the (S1, S2, REM) class with 83% success rate while the SVM with 90% (respective recall metrics), exceeding both their average hit rate (accuracy). Significant poor performance is also observed in the identification of deep sleep stages. This is obviously attributed to the biased dataset, since the instances labeled as (S1, S2, and REM) and (S3, S4) are the 62% and 6.5% respectively of all instances.

The next experiments were conducted with balanced datasets in order to study how the distribution of classes in the training data affect the performance of the classifiers. The class containing the less instances is the deep sleep stage (S3, S4) which consists of 664 instances. In order to construct the balanced dataset we selected randomly 664 instances from the two other sets of instances labeled as (W) and (S1, S2, REM) respectively. The new dataset now consists of a total number of 3*664 = 1992 instances. The tradeoff for building the balanced dataset is that now we utilize only the 20% of the available instances. The experiments have been repeated multiple times after selecting randomly 664 instances from the classes with a surplus of instances. The respective statistic results from the classification processes are presented on Table V and Table VI for FIS and SVM respectively.

TABLE V. FIS CONFUSION MATRIX WITH BALANCED DATASET

| Classified as: | W | (S1,S2,REM) | (S3,S4) | Total |
|---|---|---|---|---|
| W | **500** | 159 | 5 | 664 |
| (S1,S2,REM) | 145 | **450** | 69 | 664 |
| (S3,S4) | 11 | 80 | **573** | 664 |
| **Total** | 632 | 682 | 678 | 1992 |
| **Performance Metrics** | rec=75% prec=79% | rec=68% prec=66% | rec=86% prec=85% | **Acc**=78% |

TABLE VI. SVM CONFUSION MATRIX WITH UNBALANCED DATASET

| Classified as: | W | (S1,S2,REM) | (S3,S4) | Total |
|---|---|---|---|---|
| W | **541** | 118 | 5 | 664 |
| (S1,S2,REM) | 80 | **515** | 69 | 664 |
| (S3,S4) | 11 | 49 | **604** | 664 |
| **Total** | 632 | 682 | 678 | 1992 |
| **Performance Metrics** | rec=81% prec=86% | rec=78% prec=76% | rec=91% prec=89% | **Acc**=83% |

None of the two classifiers present significant improvement on the measured accuracy after the completion of the second configuration of the experiments. However we observe as expected a more balanced performance on the identification on each particular class. Particularly now the deep sleep stage is identified with accuracy 86% and 91% for the FIS and SVM respectively. This can be attributed to the fact that sleep stages 3 and 4 that form the deep sleep stage have characteristics that do not overlap with any of the other classes (i.e. the Delta waves). In contradiction the early sleep stage 1 which is assigned to the light sleep stage is harder to de identified from REM and wake.

Both classifiers achieved satisfactory results but they present potentials for further improvement. This improvement could be achieved from the enhancement of the dataset with new features either from the dimensionality reduction of the dataset through sophisticated feature selection algorithms. The SVM achieved significant accuracy over 80% but the FIS gave us a 78% with less computational complexity.

## V. CONCLUSION

In this paper, two well-known machine learning algorithms, namely SVM and the FIS, have been used for the prediction of the human sleep stages. These algorithms have been trained with real human EEG data together with corresponding sleep stages. The trained algorithms have been assessed in cases of predicting the sleep stages of real human patients. This assessment has shown that the SVM verified the expectations for better performance over the FIS, but both techniques can deliver sufficient accuracy. The quality of life for people suffering from chronic diseases and sleep disorders could be benefitted by tools that monitor and assess their sleep.

## ACKNOWLEDGMENT

This work was financially supported by the Operational Program DEPIN 2007-2013 of Western Greece Region under the LEADERA framework, project “Healthy Sleep and Exercise Analysis Tool (hSEAT).

## REFERENCES

[1] Von Ruesten, Anne, et al. "Association of sleep duration with chronic diseases in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam study." PLoS One 7.1 (2012): e30972.

[2] Foley, Daniel, et al. "Sleep disturbances and chronic disease in older adults: results of the 2003 National Sleep Foundation Sleep in America Survey." Journal of psychosomatic research 56.5 (2004): 497-502.

[3] Ranjbaran, Z., et al. "The relevance of sleep abnormalities to chronic inflammatory conditions." Inflammation Research 56.2 (2007): 51-57.

[4] Albuquerque, Rachel et al.. "Sleep quality implicates in life quality: An analysis about children who suffered burns." Burns 40.4 (2014): 774-775.

[5] Do Valle, C. Ribeiro, et al. "Quality of life and sleep disorders in elderly." Sleep Medicine 14 (2013): e291.

[6] Shahid, Azmeh, et al. "Pittsburgh Sleep Quality Index (PSQI)." STOP, THAT and One Hundred Other Sleep Scales.Springer New York,2012. 279-283.

[7] P. Anderer, et al., "Artifact processing in computerized analysis of sleep EEG - a review", *Neuropsychobiology*, vol.40,no.3,Sep 1999,pp.150-157.

[8] Radha, Mustafa, et al. "Comparison of feature and classifier algorithms for online automatic sleep staging based on a single EEG signal." Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE. IEEE, 2014.

[9] Ronzhina, Marina, et al. "Sleep scoring using artificial neural networks." *Sleep Medicine Reviews* 16.3 (2012): 251-263.

[10] Yılmaz, Bülent, et al. "Sleep stage and obstructive apneaic epoch classification using single-lead ECG." *Biomedical engineering online* 9 (2010): 39.

[11] Adnane, Mouradet al. "Sleep–wake stages classification and sleep efficiency estimation using single-lead electrocardiogram." *Expert Systems with Applications* 39.1 (2012): 1401-1413.

[12] Lee, Jong-Min, et al. "Detrended fluctuation analysis of EEG in sleep apnea using MIT/BIH polysomnography data." Computers in Biology and Medicine 32.1 (2002): 37-47.

[13] Goldberger, Ary L., et al. "Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals." Circulation 101.23 (2000): e215-e220.

[14] Mporas, Iosif, et al. "Online seizure detection from eeg and ecg signals for monitoring of epileptic patients." Artificial Intelligence: Methods and Applications. Springer International Publishing, 2014. 442-447.

[15] Koley, B., and D. Dey. "An ensemble system for automatic sleep stage classification using single channel EEG signal." Computers in biology and medicine 42.12 (2012): 1186-1195.

[16] Sokolova, Marina, and Guy Lapalme. "A systematic analysis of performance measures for classification tasks." Information Processing & Management 45.4 (2009): 427-437.

[17] T.-P. Hong, and C.-Y. Lee, “Induction of fuzzy rules and membership functions from training examples,” Fuzzy Sets and Syst., vol. 84, no. 1, Nov. 1996, pp. 33-47..

[18] S. L. Chiu, “Fuzzy model identification based on cluster estimation,” J. Intelligent and Fuzzy Syst., vol. 2, no. 3, 1994, pp. 267-278.

Platt, John. "Fast training of support vector machines using sequential minimal optimization." Advances in kernel methods—support vector learning 3 (1999).

4th Mediterranean Conference on Embedded Computing MECO - 2015 Budva, Montenegro

# Sleep Monitoring Classification Strategy for an Unobtrusive EEG System

J. Gialelis(*), C. Panagiotou, D. Karadimas, I. Samaras, P. Chondros, D. Serpanos, S. Koubias
Electrical & Computer Engineering Department, University of Patras, Greece
(*)Corresponding author email: gialelis@ece.upatras.gr

*Abstract*—**the advances in the wearable devices and Artificial Intelligence domains highlight the need for ICT systems that aim in the improvement of human's quality of life. In this paper we present the sleeping tracking component of an activity and sleeping tracking system. We present the sleep quality assessment based on EEG processing and support vector machines with sequential minimal optimization classifiers (SVM-SMO). The performance of the system demonstrated by respective experiments (accuracy: 83% and kappa coeff: 72%) exhibits significant prospects for the assessment of sleep quality and the further validation through an evaluation study.**

***Keywords: sleep stages; EEG; SVM;***

## I. INTRODUCTION

The evolution observed nowadays in the domain of pervasive health and wellbeing assisted from the maturation of Internet of Things (IoT) technologies has led to the hyper growth of the market regarding wearable devices and applications related to health and fitness. According to [1], the respective market is expected to increase rapidly from US$750 million in 2012 to US$5.8 billion in 2018.

Nevertheless significant work has been published so far, regarding systems employing data mining techniques that are tailored on the detection of specific pathological anomalies (heart diseases, epilepsy, etc.) [2]. On the other hand the multi-modal acquisition of data and context from wearable devices gains also interest and attraction. These data along with respective algorithms are expected to help researchers to extract meaningful patterns that will help them to monitor and detect human's daily routines that jeopardize his health and deteriorates quality of life in short or long term.

The significance of sleep quality is strongly connected with quality of life since it directly affects mental health [3]. At the same time, literature highlights the coexistence of sleep disorders with chronic diseases as well. Therefore, tools for the sleep quality assessment and interventions for its improvement could significantly contribute towards a qualitative life. Pilot studies have manifested the strong relation among sleep quality and physical exercise [4].

These systems however need to address many contradictory challenges. First of all, in order to achieve a high degree of adoption by the end users and hence commercial success, they need to operate unobtrusively and seamlessly without distracting the users. At the same time, they need to collect data and produce knowledge that could have medical meaning for doctors and users.

In that context this paper presents the design and development of a sleep and activity tracking platform that intends to deliver unobtrusiveness and a more fine grained sleep and activity analysis. This paper extends the primary work presented in [5] by focusing on the sleep stage identification component integration as well as system's architecture design and implementation along with some preliminary results.

Section II presents a brief description of the background, regarding sleep staging and the most popular approaches followed for similar systems. Section III decomposes the system in its components while in Section IV we present the intelligence core of the proposed system that is responsible for the sleep stage identification and sleep quality assessment. The paper concludes in Section V and summarizes the progress made so far along with the future steps to follow.

## II. BACKGROUND & RELATED WORK

The sleeping session is a dynamic process that consist of sleep stages and is described by the transition from one to another along with the total time spent on each stage. Namely, the sleep stages are discriminated in Wakefulness (W), stage 1 (N1), stage 2 (N2) (these two stages are characterized as light sleep), stage 3 (N3) (characterized as deep sleep) and Random Eye Movement (REM).

According to literature, the golden standard for the accurate sleep stage classification is polysomnography (PSG) along with the R&K rules scoring system as restructured in 2007 by the American Academy of Sleep Medicine (AASM) [6]. The guidelines for this particular methodology require the use of at least 6 EEG channels, 2 electromyography (EMG) leads attached to the chin and 2 electrocculography (EOG) leads that track eye movements. Obviously, such a setup is cumbersome for the person that needs to track his sleep as it can be feasibly applied only in lab environments, with experts' assistance and not on as a daily routine.

The devices that flood the market recently, such as Fitbit Flex [7], utilize an alternate, yet less obtrusive method, the actigraphy. Actigraphy is applied through a wearable wrist band which integrates accelerometer, capable of identifying body movements [8]. While these devices can produce results, related to the detection of sleep/wake, close to PSG, no further details about the sleep cycle can be extracted so. Therefore, the adoption of such techniques in medical practice have been poor and only supplementary to the rest of the methods.

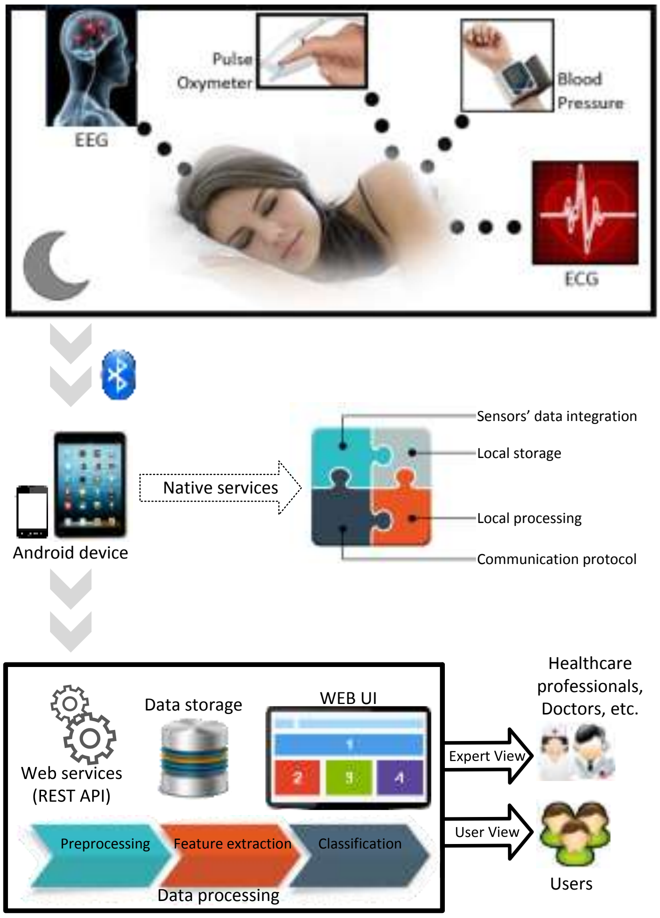



Figure 1. Sleep tracking component

Authors in [9] present a unobtrusive approach by utilizing only a smartphone's sensors (microphone and inertial measurement unit) for the detection of body movements, snoring, heartbeat and respiration while authors in [10] use a smartphone for detecting sleep patterns by studying smartphone's usage patterns (e.g., the time and length of smartphone usage) and ambient observations (e.g., prolonged silence and darkness).

EEG processing for sleep stage identification takes advantage of the recent progress in machine learning, so as to employ robust and accurate sleep stage classifiers. Authors in [11] present a feature extraction methodology and a relevance vector machine (RVM) for automatic sleep staging classification. Although they achieve a total accuracy of 76.7% they fail to detect sufficiently N1 stage (23.2%). Artificial neural networks are proposed in [12] for sleep scoring and they also achieve a significant accuracy of 81.55%. In [13] two popular classifiers are tested on how they perform on a specific feature set for online automatic sleep staging. Namely Random Forest (RF) and Support Vector Machine (SVM) were compared and the RF outperformed for the specific configuration.

## III. SYSTEM DESCRIPTION

Regardless the new approaches presented in the recent literature for unobtrusive sleep monitoring, the most prominent tool to provide all the necessary information regarding the brain activity during sleep remains the single channel EEG. The ecosystem of the presented system is based on three main pillars: the wearable devices, the smartphone and the backend system. The sleep tracking component of the system architecture is presented in Figure 1.

### A. Wearable Devices

Wearable devices are the cornerstone of all related systems for health monitoring and ambient assisted living (AAL) systems, as they constitute the raw data source. In our use case we decided to use components of the self (COTS) that employ widely adopted and popular wireless standards, such as Bluetooth or Wi-Fi, that incorporate built-in flexible and interoperable communication interfaces while at the same time the same time the devices address the needs for comfort and unobtrusiveness.

In that context, the EEG source is based on the MyndPlay BrainBandXL which is a commercial EEG recording product, based on the Neurosky platform, with two dry sensors and an ear clip as a reference point. The Surface Mount Device (SMD) board is attached on a headband made of a soft and comfortable fabric. The wireless connectivity offered by the MyndPlay platform is based on Bluetooth.

However, the modalities supported by the proposed platform are not restricted to the EEG. The activity recording depending on the user profile and his comfort zone may varies from a simple step counter to fitness and medical chest straps for ECG recording. At the moment the presented system supports 2 activity recording modes. The most unobtrusive and simple is the step counter provided by the smartphone, while for a more sophisticated heart activity recording the Zephyr BioHarness is employed. Zephyr BioHarness encapsulates 2 conductive ECG sensor pads and 1 internal breathing sensor. The wireless connectivity remains the Bluetooth.

Finally, the proposed system also integrates the A&D Medical – UA-767BT automatic blood pressure monitor along with the NONIN 4100 pulse oximeter as complementary metrics for the completeness of user's health profile. Both devices provide Bluetooth connectivity.

### B. Smartphone Functionalities

The smartphone platform could be any Android device with Bluetooth and Wi-Fi connectivity. The smartphone provides a twofold contribution to the system.

The first and primary role of the smartphone is derived by its capability for providing interconnection between the wearable devices and the backend system, thus playing the role of a gateway. Communication drivers have also been implemented and deployed at the application level for the interaction with each wearable device. The required connectivity is implemented as individual android services that run in the background. As long as a paired wearable device of the user is detected, in the vicinity of the smartphone's wireless interfaces, a connection is established and the smartphone receives either data streams from EEG and ECG sensors or single values from the pulse oximeter and blood pressure monitor devices. All received values are temporary stored in the smartphones internal storage until they are transmitted to the backend system for further processing and storage.

The second functionality that the smartphone provides is derived by the native services that any Android device implements. For example, step counter values are also gathered and provide some primary indication of the person's activity while the GPS sensor can enhance the activity record with speed and distance covered during running.

### C. Backend System

The backend part of the proposed system is responsible for acquiring and storing data that have been collected by the wearable devices during each user's sessions and transmitted via the Android services deployed within the smartphone. Apart from simple data gathering, the backend system also incorporates some data processing components for high level information creation and storage. Overall, the backend system consists of four main components as described below.

- Web services provider for data gathering. This component is mainly responsible for gathering, handling and further processing of transmitted data from the smartphone services. Since continuous recording of EEG and ECG signals can easily derive huge amount of raw data a special treatment has been applied as a custom communication protocol between the backend system and the smartphone's services towards minimization of data volumes that have to be transmitted. Such a technique results in major advantages in processing and storage resources in backend as well as in mobile pillar of the proposed system. Finally the component is responsible for feeding the visualization component (Web UI) through proper web interfaces.
- Data processing. The software modules responsible of the processing of the biosignals and the physiological data along with the respective classifiers are integrated in the backend system. An extensive description of the processing conducted on the EEG for the sleep stage classification is presented on Chapter IV.
- Data storage. Raw data together with extracted high level information from data processing component are stored in backend's database which is deployed under a typical MySQL server. On top of the database engine an ORM framework, with a tiny processing and memory footprint, is responsible for abstracting the actual database records, thus enabling large scale employment.
- Web application (Web UI) for system's stakeholders. The backend system is finally responsible for providing useful information (such as advices, tips, etc.) to all involved stakeholders. The presentation view of such information is responsive to the stakeholder type. For example typical user's view includes his activity history in a timeline basis and some advices/tips that are specified by his corresponding doctors or health care team. At the same time doctors and health care professional have full access to either visualize a user's raw data or become aware of features extracted from raw data (i.e. sleep stages from EEG, HR-artifacts from ECG, etc.). Doctors' view has the ability to intelligently adapt the relevant information from simple diagrams to cohesive dashboards in order to enable the doctor to get the most with the less possible effort and interaction.

## IV. SLEEP QUALITY ASSESSMENT

In this paper we focus on how the proposed system handles the sleep related biosignals and specifically the EEG which is a non-stationary, nonlinear and noisy signal. EEG is an error-prone signal since it coexists with stronger biosignals that interfere and introduce noise as artifacts. These artifacts originate mainly from ECG, EMG and eye blinks. Further noise is introduced by the power line inference and the body movements. The methodology followed for the identification of sleep stages consists of three sequential steps: signal preprocessing (artificacts/noise elimination), feature extraction (frequency & time domain analysis) and final the classification.

To eliminate this noise as much as possible without eliminating valuable sleep frequencies we apply a 2$^{nd}$ order butterworth filter with low and high cutoff frequencies set at 0.3 and 40Hz respectively. Since the EEG is a non-stationary signal, it is processed in epochs of 30 seconds where stationarity is assumed.

Subsequently, the filtered signal feeds the feature extraction algorithms so as to extract the set of parameters that describe the brain activity during the sleep cycle. For that purpose we follow a frequency and time domain analysis and we calculate the most dominant features that are documented in the literature for sleep stage classification.

The sleep stages described in Section II are manifested in the EEG as frequency components. Therefore, we perform a respective spectral analysis in order to obtain the EEG power density spectrum based on Welch's method [14].

Further features obtained from time domain analysis include the estimation of mean, max and min values as long as standard deviation and variance. The Zero Crossing Points value is another promising feature that manifests the presence of spindles that occur mainly in Stage 2. The set of the dominant features for the sleep stage identification is completed with the estimation of Hjorth Mobility & Complexity, Kurtosis, Skewness and Interquartile range along with the auto regressive filter coefficients occurred from the 6$^{th}$ order auto regression analysis of the signal [15].

The classification scheme implemented in our system is based on Support Vector Machines with the Sequential Minimal Optimization (SMO). Our prototype classification model has been built on data obtained from Physionet [16] and particularly the MIT-BIH polysomnography sleep database [17]. This database consists of 18 records obtained from 16 subjects with the method of polysomnography. From this database we utilized the EEG records which where sampled with a frequency of 250Hz. All the records are segmented in epochs of 30s duration and annotated based on the initial sleep scoring scheme (4 non-REM stages) by medical experts.

The method followed for the performance assessment of our classifier was based on the 10-fold Cross Validation. This method splits the dataset in 10 independent and disjoint test sets with almost equal class distribution. The classifier is trained with the remaining data left from the i-th test set and tested on the i-th test set. The experiments are repeated 10 times for each i-th test/train dataset for i=1:10. The performance metrics of the classifier occur from the average of the 10 repetitions.

Our analysis was based mainly on two statistical measures, the total accuracy and Cohen's kappa coefficient. As accuracy is defined as the ratio of the correct predictions across all classes divided by the total number of instances. For more coherent interpretation of the classification results we employ the kappa coefficient as an indicator of the relative improvement of the classifier over the random predictor.

The first experiment was conducted with all the instances of the database as input to the classifier. Our initial aim was to identify the three main sleep stages, namely wakefulness, light sleep (stages 1, 2 and REM) and deep sleep (stages 3, 4). This resulted in unbalanced dataset with 10181 instances. The results are depicted in Table I.

TABLE I. UNBALANCED DATASET - SVM CONFUSION MATRIX

| Classified as: | W | (S1,S2,REM) | (S3,S4) | Total |
|---|---|---|---|---|
| W | **2147** | 966 | 7 | 3120 |
| (S1,S2, REM) | 487 | **5775** | 135 | 6397 |
| (S3,S4) | 7 | 233 | **424** | 664 |
| **Total** | 2641 | 6974 | 566 | 10181 |
| | **TP rate = 69%** | **TP rate = 90%** | **TP rate = 63%** | **Acc=8346/10181 =82% Kappa = 63%** |

While the model occurred from this dataset performed with an accuracy of 82% it doesn't identify every class with the same accuracy. Respective True Positive (TP) rates present that our model is biased towards the light sleep stage (TP rate 90%). This is also indicated from the kappa coefficient value which reaches the 63%. In order to improve the predictability of our model we fed the classifier with a balanced dataset. The new dataset consists of the 664 instances of the deep sleep class and 664 instances randomly chosen from wake and light sleep classes respectively. The performance our classifier achieved is presented on Table II.

TABLE II. BALANCED DATASET – SVM CONFUSION MATRIX

| Classified as: | W | (S1,S2,REM) | (S3,S4) | Total |
|---|---|---|---|---|
| W | **541** | 118 | 5 | 664 |
| (S1,S2, REM) | 80 | **515** | 69 | 664 |
| (S3,S4) | 11 | 49 | **604** | 664 |
| **Total** | 632 | 682 | 678 | **1992** |
| | **TP rate = 82%** | **TP rate = 78%** | **TP rate = 91%** | **Acc = 83.33% Kappa = 72%** |

The strategy followed in the second experiment achieved a slight increase in the accuracy (83.33%) but significant improvement in the prediction capabilities. The kappa coefficient value reached 72% while the TP rates where more equally balanced among the three classes. The TP rate of deep sleep outperformed with a TP rate of 91% which was expected as the deep sleep stages are manifested more clearly in the EEG.

## V. CONCLUSIONS

In this paper we presented the main architectural and implementation approaches followed for the development of an activity and sleep tracking system. The main focus of the particular paper was the sleep tracking component of the presented system. The EEG processing and classification strategy which we followed, tested on data provided by open databases (MIT – BIH polysomnography), achieved significant performance on the prediction of sleep stages therefore we are confident that it will be validated through a respective evaluation study.

## ACKNOWLEDGMENT

This work was financially supported by the Operational Program DEPIN 2007-2013 of Western Greece Region under the LEADERA framework, project "Healthy Sleep and Exercise Analysis Tool (hSEAT).

## REFERENCES

[1] Wei, J., "How Wearables Intersect with the Cloud and the Internet of Things: Considerations for the developers of wearables.," *Consumer Electronics Magazine, IEEE* , vol.3, no.3, pp.53,56, July 2014

[2] Shukla, D. P., Shamsher Bahadur Patel, and Ashish Kumar Sen. "A literature review in health informatics using data mining techniques." *Int. J. Softw. Hardware Res. Eng. IJOURNALS* (2014).

[3] Zeitlhofer, J., et al. "Sleep and quality of life in the Austrian population." *Acta Neurologica Scandinavica* 102.4 (2000): 249-257.

[4] Wang, Xuewen, and Shawn D. Youngstedt. "Sleep quality improved following a single session of moderate-intensity aerobic exercise in older women: Results from a pilot study." *Journal of Sport and Health Science* 3.4 (2014): 338-342

[5] J. Gialelis, C. Panagiotou, P. Chondros, D. Karadimas and E. Batzi. "End-to-End System for the Relationship Modeling of Physical Exercise with Sleep Quality." *Proceeding of the International Conferences on ICT, Society and Human Beings (2014): 367 – 372*.

[6] Rosenberg, Richard S., and Steven Van Hout. "The American Academy of Sleep Medicine inter-scorer reliability program: sleep stage scoring." *J Clin Sleep Med* 9.1 (2013): 81-87.

[7] Fitbit Flex wristband, Available at: http://www.fitbit.com/store

[8] Bhagat, Y.A.; Byunghun Choi; Do Yoon Kim; Jaegeol Cho; Black, R.; Foster, G.H.; Insoo Kim, "Clinical validation of a wrist actigraphy mobile health device for sleep efficiency analysis,"*Healthcare Innovation Conference (HIC), 2014 IEEE* , vol., no., pp.56,59, 8-10 Oct. 2014.

[9] Kalkbrenner, C.; Stark, P.; Kouemou, G.; Algorri, M.-E.; Brucher, R., "Sleep monitoring using body sounds and motion tracking," *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE* , vol., no., pp.6941,6944, 26-30 Aug. 2014.

[10] Zhenyu Chen; Mu Lin; Fanglin Chen; Lane, N.D.; Cardone, G.; Rui Wang; Tianxing Li; Yiqiang Chen; Choudhury, T.; Campbell, A.T., "Unobtrusive sleep monitoring using smartphones,"*Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2013 7th International Conference on* , vol., no., pp.145,152, 5-8 May 2013

[11] Huang, Chih-Sheng, et al. "Knowledge-based identification of sleep stages based on two forehead electroencephalogram channels." *Frontiers in neuroscience* 8 (2014).

[12] Ronzhina, Marina, et al. "Sleep scoring using artificial neural networks." *Sleep Medicine Reviews* 16.3 (2012): 251-263.

[13] Radha, Mustafa, et al. "Comparison of feature and classifier algorithms for online automatic sleep staging based on a single EEG signal." *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*. IEEE, 2014.

[14] Welch, P.D. "The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms", *IEEE Transactions on Audio Electroacoustics, AU-15 (1967)*: 70–73.

[15] Mporas, Iosif, et al. "Online seizure detection from eeg and ecg signals for monitoring of epileptic patients." *Artificial Intelligence: Methods and Applications. Springer International Publishing, (2014)*: 442-447.

[16] Goldberger, Ary L., et al. "Physiobank, physiotoolkit, and physionet components of a new research resource for complex physiologic signals." Circulation 101.23 (2000): e215-e220.

[17] Lee, Jong-Min, et al. "Detrended fluctuation analysis of EEG in sleep apnea using MIT/BIH polysomnography data." Computers in Biology and Medicine 32.1 (2002): 37-47.

# Towards a Dynamic Waste Collection Management System using Real-time and Forecasted Data

George Asimakopoulos,
Sotiris Christodoulou, Andreas Gizas,
Vassilios Triantafillou,
Giannis Tzimas, Emmanouil Viennas
Computer & Inform. Engineering Dept.
Technical Educational Inst. of Western Greece
Antirion, Greece
{gasimakop@gmail.com, sxristod@teimes.gr,
gizas@ceid.upatras.gr, triantaf@teimes.gr,
tzimas@cti.gr, viennas@gmail.com}

John Gialelis, Dimitris Karadimas,
Andreas Papalambrou
Industrial Systems Institute, "Athena" RIC in ICT and Knowledge Technologies, Patras Science Park, Stadiou Str.,Platani, PATRAS, Greece, GR-26504
{gialelis@isi.gr, karadimas@isi.gr,
papalambrou@isi.gr}

## ABSTRACT
Dynacargo, a research project under implementation, aims to change supply chain management, as the hauled goods are at the center of attention instead of vehicles, as in traditional management approaches. At the same time Dynacargo manages both delivery and collection cycles. Dynacargo aims at developing a near real-time monitoring system that monitors and transmits waste bins' fill level, which is used to dynamically manage the waste collection process by introducing distance minimization, relying on efficient routing algorithms. This paper deals with research issues that have emerged along with developments regarding specific architectural modules. In general, Dynacargo places a set of durable, low cost sensors and RFID tags on waste bins. These tags store the fill-level estimated by the sensors, which is passed through diverse communication channels and ends to a central cargo information management system. Along with this real time data harvesting, data mining techniques are utilized on historical data collected prior to Dynagargo implementation, in order to predict future waste bins fill rates. Mobile and web applications are developed in order to harvest relevant data from citizens.

## Categories and Subject Descriptors
**C.0 GENERAL:** System architectures, **H.2.8 Database Applications:** Data Mining, **G.2.2 Graph Theory:** Graph algorithms

## General Terms
Algorithms, Design, Experimentation.


16th EANN workshops, September 25 - 28, 2015, Rhodes Island, Greece
© 2015 ACM. ISBN 978-1-4503-3580-5/15/09…$15.00
DOI: http://dx.doi.org/10.1145/2797143.2797154

## Keywords
Urban solid waste collection; intelligent transportation systems; data mining; graph routing algorithms.

## 1. INTRODUCTION
Nafpaktia municipality is a typical case of solid waste management throughout Hellas, as waste collection is based on standard time intervals and according to fixed vehicle routes. Decision making is solely empirical, which leads to biased decisions that do not take into account real needs based on data.

This has led to results far from optimal; as it is not uncommon to have overfilled waste bins were uncollected for some days, while at the same time other unfilled bins were collected, resulting to an unsatisfied local society along with increased cost. Regarding Athens[8], it has been estimated that the 60% to 80% of the total cost of waste collection, transportation and disposal is spent during collection. The main cost reduction efforts should deal with distance and duration minimization of vehicle routes [6]. Johansson [7] proved that if the fill level of bins was taken into account and waste collection adapts accordingly, it could reduce the cost of waste collection up to 20%. Under the light of these findings, Dynacargo (Dynamic Cargo Routing on-the-Go) attempts to build a near real-time monitoring system that monitors and transmits waste bins' fill level, in order to make waste collection more efficient by cost reduction which is accomplished by minimizing distances covered by refuse vehicles.

## 2. RELATED WORK & INNOVATION OF DYNACARGO
There are various approaches available in order to harvest data from points of waste collection [2,3,4, 5,7,9,10]. Although they introduce interesting approaches, they cannot be used as-is in Dynacargo. This is because they introduce major costs, they do not meet Dynacargo functional requirements and are mostly customized for recyclable waste which can be managed with longer collection cycles compared to domestic waste.

This implies that Dynacargo, despite aiming at waste collection automation as existing approaches do, displays major differentiations against these approaches. At the heart of these differentiations lies Dynacargo architecture which originates from a generic multipurpose cargo-based dynamic vehicle routing approach which copes with cargo changes and collection from disperse points of concentration.

Dynacargo moves forward from sole data collection from waste concentration points, as it utilizes such data in real time in order to optimize vehicle routes during the waste collection process execution.

Another point at which Dynacargo differentiates from existing approaches is the utilization of a diverse set of data transmission techniques that is incorporated to establish data transmission from points of collection to the system in near real time. Besides GSM which induces fixed telecom costs, Delay-Tolerant Networking concepts are adapted for data transmission from the disperse concentration points to the central information system. Dynacargo DTN is based on a set of existing public commuters that are utilized as data hosts that carry data as they execute unaltered standard procedures.

## 3. NAFPAKTIA AS CASE STUDY

Dynacargo functional requirements evolved from the analysis of the solid waste collection and management system of Municipality of Nafpaktia, Greece, which expands in an area of 870,38 km$^2$ and displays a population of 27.800 citizens. Nafpaktia displays geographic and demographic aspects that sum to an ideal pilot environment for Dynacargo. It is composed from a coastal city with narrow busy streets in the historic center, coastal towns and villages with population varying following tourist seasonality, along with remote villages (up to 120 km. away) with a few dozen residents located in mountainous areas. Telecom network coverage along with public transportation frequency display significant diversion between different Municipality areas. Another diversion is that Nafpaktia displays waste transshipment from small to big refuse vehicles. Requirement analysis initiated with an as-is analysis in order to formally document the current waste management process. Several indicative waste collecting routes were selected and modeled on Google Earth (Figure 1), after an assessment of peculiarities each route presents. The analysis included available waste collection historical data and information regarding related processes.

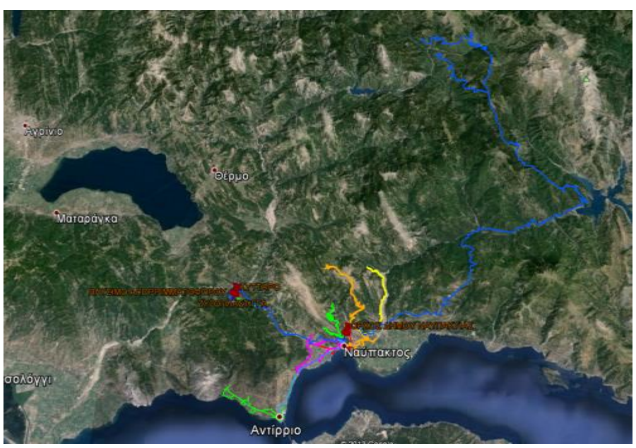



**Figure 1. Indicative waste collection routes for Nafpaktia.**

## 4. USERS & USE-CASE MODELS

Dynacargo identifies a set of target user groups which include waste collection related municipal services, system administrators, truck drivers, bins' data collectors and individual citizens.

In order to reassure the accurate Dynacargo waste level predictions, waste bin fill level data is acquired at short time intervals from as many as possible waste collection points. Data collection is realized through diverse data harvesting tools take advantage of existing transportation systems such as public bus services, postal office vehicles, taxis, municipal police vehicles, along with active social entities as these vehicles traverse the area of interest in an irregular but frequent manner. The Data Collector role can be served by either a vehicle driver (e.g., bus, taxi, postman car, etc.) equipped with the Dynacargo equipment, or anyone else who may be involved in data collection (e.g., postmen) who may use any means of transportation such a car, a bike, or by foot. Citizens improve data harvesting by reporting fill-level with a mobile application, reporting estimations of produced waste volume on unforeseen events and by checking online the bin fill levels near their residence so as to discard their waste on nearby unfilled bins.

Based on users' functional requirements, several application scenarios have been set out, which fall into two categories based on time of implementation: *Full-scale Application Scenarios* and *Pilot Scenarios* (scenarios that will be demonstrated during the project). Full-scale scenarios define use-case models for each subsystem, including the vast majority of functional requirements along with several non-functional requirements. Six major use-case models were set out modeling different usage cases of the system: Bins' Data Collector, Refuse Truck, Central System, Optimal Routes Calculation, Citizen's SmartPhone App and Citizen's Information Portal. These use cases and more detailed description of the functional requirements of Dynacargo, are described in a previous research work [1].

## 5. ARCHITECTURE

Functional requirements and use case models led to the definition and design of the major Dynacargo structural components, followed by decomposition to subsystems and smaller functional units. Operations, roles and significant logical relationships with other subsystems were defined for each subsystem.

Dynacargo specifications were derived based on exhausting analysis of the waste collecting process as it is performed by Nafpaktia municipality. However, Dynacargo architecture is designed as a generic cargo-based routing system along with the inherited adaptability to any other municipality regardless of specific waste collecting process characteristics. The Dynacargo architecture is presented in figure 2 and in more detail in [1]. Following, we shortly describe its main subsystems and architectural parts.

**Bin Subsystem:** Sensors estimate waste bin fill level and collect, store, and transmit relevant data.

**Bins Data Gathering and Dispatching Unit:** This unit communicates through a short range network with the Bin Subsystem and routes the collected information to the Central System, through a long range network. It can be installed at passing vehicles and is built of three components:

- *"Reader"*, which refers to the communication interface with the bin subsystem.
- *"Storage space"* which temporarily stores the data until transferred to the Central System.
- *"Data transmission"* to the Central System, through Delay Tolerant Networking (DTN) technology.

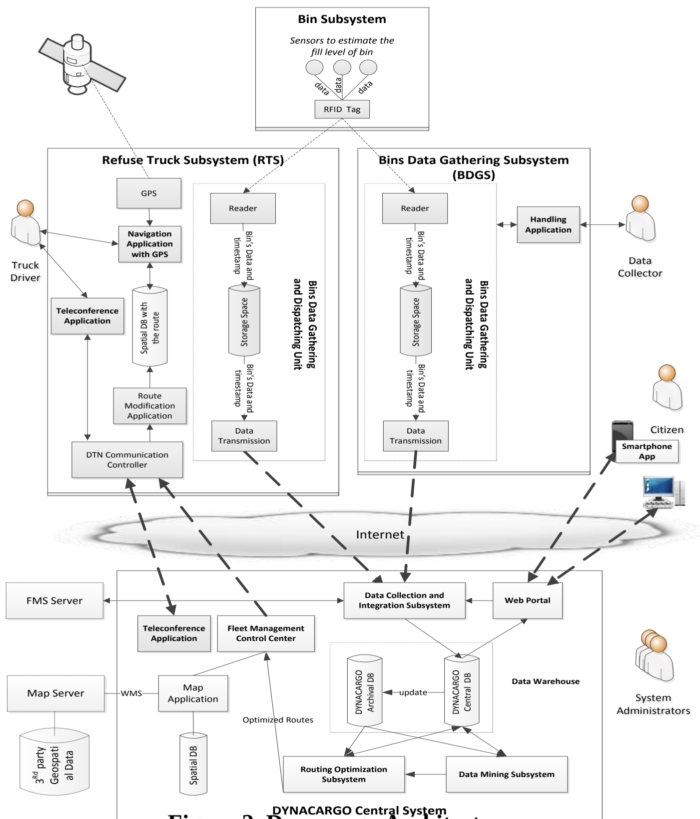



**Figure 2. Dynacargo Architecture.**

**Bins Data Gathering Subsystem (BDGS):** The BDGS is the subsystem carried by the Data Collector. It consists of the Bins Data Gathering and Dispatching Unit, combined with a very simple application for handling it. The BDGS should be portable and able to work with a battery.

**Refuse Truck Subsystem (RTS):** This subsystem will be installed on each refuse truck. The main part of the RTS is the Bins Data Gathering and Dispatching Unit. Additionally the RTS will be equipped with a GPS unit and a camera and will incorporate the appropriate software to provide the following functionalities:

- *GPS Navigation Application:* It is a classical navigation application through GPS, which will provide navigation guidance to the truck driver, and instructions about which bins should be collected.
- *Teleconference Application:* Teleconference communication of driver with the Central System, especially for reporting emergency situations.

The RTS will be powered by the refuse truck. The **communication** with the Central System will be based on DTN technology, which will transparently select the most suitable communication method, considering criteria like cost or speed.

**Citizen Application:** An application for mobile devices (smartphone, tablet), through which the citizen can choose any bin and report data for it (fill level, photo, comment, etc.).

**Central System:** It is the back-end system of Dynacargo. Its main part is the Data Warehouse storing all historic data, bins' data, vehicles' data and everything needed for calculating the best routes based on the current load of bins and some restrictions (Bin Collection Settings) specified by the system administrators. Moreover, Data Warehouse will store any information derived from the other subsystems, particularly from the Data Mining Subsystem and the Routing Optimization Subsystem.

Apart the above databases, the Central System will include:

**Fleet Management Control Center:** receives the optimized routes from the Routing Optimization Subsystem, two hours before the start of the routes and on the fly (if the routes have to be modified), and forwards them to the appropriate RTS. All spatial information for the corresponding route will be stored locally in a Spatial RTS Database. The Fleet Management Control Center can send route modifications to any refuse truck, informing properly the Navigation Application, if such need arises and provided that the refuse truck has network coverage.

**The Map Application:** A cartographic JavaScript API that will provide all the required functionality for creating rich-web applications based on geographic and descriptive data.

**Data Collection and Integration Subsystem:** This subsystem will collect bins data from the RTS, the BDGS and citizens' applications, and integrate them into the central DBs. Furthermore, it will be able to integrate data from other sources, if needed, like an independent FMS server that can operate in parallel with the Dynacargo system. The architecture design of Dynacargo does not require the existence of a classical FMS.

**Data Mining Subsystem:** Mainly used to estimate fullness of bins when we do not have the available information updated or is not fairly recent. It generally seeks to apply fusion techniques from multiple sources to produce more semantically rich data, thereby obtaining information on a higher level of abstraction.

**Routing Optimization Subsystem:** The role of this subsystem is the dynamic route planning before the starting of truck routes, and the on-the-fly modification of routes, either due to exceptional events (accident), or if the new data collected during the routes impose such changes. The routes will be calculated two hours ahead, taking into account various data that will exist in the central data warehouse system.

**Citizens Web Portal:** A web site, through which citizens can be informed about the current completeness of bins or report bins' data (fill level, photography, etc.) in the Central System, by selecting a bin on a map. The purpose of the Citizen Web Portal and the smartphone app are to motivate people to participate in sustainable waste management.

In the three following sections we further discuss the current design and implementation issues of three important subsystems.

# 6. BIN SUBSYSTEM AND FILL LEVEL ESTIMATION

The fill level estimation of urban solid waste bins appears to be a challenging a task. The irregular shape and the variety of the materials require versatile sensing approaches. The harsh environmental conditions (e.g., humidity, temperature, and dust) can significantly affect the sensor measurement accuracy and reliability.

## 6.1 The Physical bin system

The proposed smart waste-bin system is envisaged to be mounted on the top lid of a waste-bin and it consists of the sensing units, an active RFID tag for data aggregation and transmission, as well as a protective enclosure for the sensors and the RFID tag that may optionally include an external battery source. On the other side, the waste-bin system interacts with an active RFID reader which could be optionally accompanied by a bare-bone-pc, i.e. a Raspberry Pi for enhanced capabilities. The aforementioned modules are described in the following paragraphs starting from the waste-bin which constitutes the main physical component and moving towards the receiver side.

### *6.1.1 Mobile Garbage Bin*

The physical system is based on the standard type of a mobile garbage bin (MGB), also known as a "wheelie bin", used in the municipality of Nafpaktia, Greece, where the pilot of the Dynacargo project will take place. The bin has a main compartment whose internal dimensions are approximately 110x90x90 cm, consisting of slight curved-shaped walls. It has a lid-opening mechanism and is made of durable hard plastic. The lid of the bin was chosen as the point of placement of the sensors because the main compartment is subject to harsher conditions due to the presence of the waste as well as the washing procedure of the bins.

### *6.1.2 Sensing Unit*

A comparison of various solutions including infrared proximity sensors, optical sensors and ultrasonic sensors indicated that ultrasonic sensors are the most suitable solution for the purpose of the presented architecture taking into account the harsh environmental conditions (e.g., humidity, temperature, and dust) that can significantly affect the sensor measurement accuracy and reliability. The ultrasonic sensors are advantageous in providing ranging measurements independently of the contained objects, thus making possible the corresponding translation into fill level measurements. For the purpose of this task, the ultrasonic sensors of Maxbotix were selected which provide a variety of detection patterns, accuracy and durability. Since the resolution of the ultrasonic sensors is only a few centimeters (about an inch), the selected solution can offer fine-grained accuracy for the purpose of the application. The ultrasonic sensors should be mounted in the bin lid, exposing only a small part of the sensor body. Since the sensors will operate unattended in the field, low power consumption models that also offer IP-67 protection rating can be used. An analog output is provided for the transmission of the measurements. The experimental placement of sensors on the top of the bin can be found in Figure 3.

**Figure 3. Experiment setup for verification.**

The measurements of the sensors essentially provide the distance to the nearest object of efficient size. Since waste in the form of irregular-shaped objects can occupy the bin space in various arrangements, there is typically some difference between the highest bin level and the actual fill percentage. Following experimentation with placement and number of sensors it was concluded that a single sensor placed in the center of the lid was providing good results but over-estimating the fill level due to the typically higher level of waste in the center of the bin. As a result, various combinations of two sensors were examined. It was concluded that the optimal results were provided by placement of two sensors in a way that their areas of detection would not overlap.

### *6.1.3 Active RFID Tag*

Active RFID tags have been selected as the data aggregation and transmission unit for the bin subsystem. The selected tags (ZT-50-mini Tag by TagSense) can operate with standard 3V voltage input providing long lifetime that equals to millions of beacon transmissions. The employed active tags offer extra I/O pins for communicating with external devices, in our case the ultrasonic sensing units. This setup allows powering the sensors and the RFID tag from the same source, either tag's battery or an external power source.

In the Dynacargo scenarios the tag wakes up at predefined times of the day, depending on the location of the waste bin installation and on the desired measurement frequency defined by the Municipality operators.

The Dynacargo operation does not assume availability of a fixed network infrastructure that reaches all the installed bins, especially those at distant villages. Moreover, at an urban scale of operation, this is a quite realistic assumption, since thousands of sensors are sparsely deployed in a complex city and suburban terrain.

## 6.2 Fill Level Estimation

### *6.2.1 Computer model of the system*

The waste-bin model that was developed is based on the use of a three dimensional matrix representing the entire waste-bin's internal volume. This model was developed using MATLAB because it offers powerful matrix handling tools. The ultrasonic sensors were also modeled as three dimensional matrices. These matrices were of the same size as the bin matrix with zero values for the areas of the bin which fall outside the detection range of the sensor and with unit values for the areas inside the detection area. One of the most crucial parameter for the simulation was waste modelling, mostly concerning its generation. The representation of waste was made using the bin matrix by setting 1 in the elementary volumes where waste was present and 0 otherwise. More challenging was the method to generate waste in a way that would be equivalent to the waste of a real bin. In absence of an existing model in the bibliography concerning how waste of irregular shape and size gradually fills a bin, it was chosen to perform extensive random scenarios of bin filling that would cover all possible combinations. Then, the less likely situations can be filtered out by statistical analysis of the derived combinations.

### *6.2.2 Simulation*

The bin system simulation in MATLAB took place via the following steps: i) Parameters are set (sensors types, bin sizes etc.); ii) The matrices of the bin and the sensors are created; iii) Waste is generated through a random process based on input parameters; iv) The waste surface is calculated; v) The sensor values are calculated based on the smallest distance between the sensor and the waste surface; vi) The waste volume is calculated; vii) The fill percentage is calculated; and finally viii) A 3D plot of the waste inside the bin is created. The results of interest for the purposes of this work are the actual fill percentage and the sensor measurements. Figure 4 depicts the simulated end results for a 30k-run simulation and an elementary base-edge length of 20cm for the area of responsibility of a single sensor. In this figure, the vertical axis represents the actual fill-level percentage of the area calculated each time based on the waste-bin and the sensor models, while horizontal axis represents the actual value based on the sensor model. It is noticeable that selecting other values for the elementary base-edge length did not alter the shape of the scattered points, but only the amount of spreading of the scatters for medium to low sensor values. This is expected taking into account the actual physical problem.

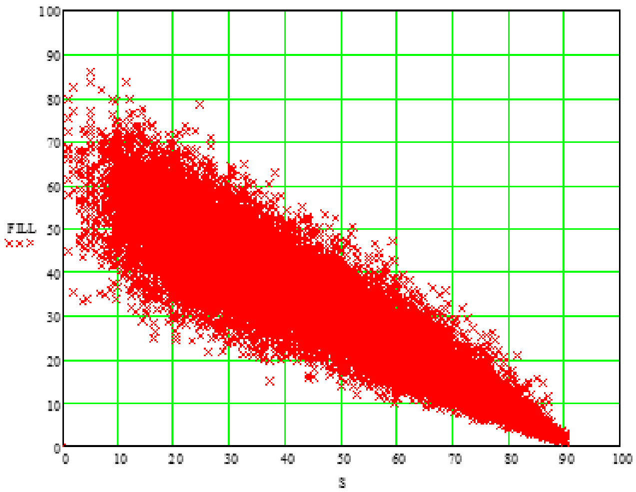



**Figure 4. Simulated end results for a 30k-run simulation and an elementary base-edge length of 20cm for the area of responsibility of a single sensor**

The raw results derived from a sufficient number of simulations for a variety of models' parameters selection can be further statistically processed. The selected totally random method for filling a bin followed in the simulations provided different fill-levels for a range of identical sensor values. These fill-levels have been grouped so as to extract the corresponding mean value and deviation for each group, thus treat the bin filling with elementary wastes as a Gaussian process.

## 7. BINS' DATA GATHERING AND DISPATCHING UNIT

The Bins' Data Gathering and Dispatching Unit (or simply Collector Unit) consists of an active RFID reader and a small-form-factor computer (currently a Raspberry Pi), consuming about 10 W of power (maximum) and having two USB 2.0 ports, built-in WiFi and Bluetooth, and one Ethernet port that is supported by GNU/Linux operating system running from a microSD card. The selected active RFID reader (ZR-USB RFID Reader by TagSence) is inherently designed to communicate with the TagSense ZT-50 active tag. The ZR-USB active reader communicates with the active tags using a variable packet length protocol which is designed to conserve power on the tag. The air interface protocol layer is based on the underlying IEEE 802.15.4 industry standard, which is best-known as the physical layer for Zigbee.

The waste bins are equipped with the Field Unit and the active RFID tags transmit periodically their identity and a sensed value regarding bin's fill level. This option allows the minimal possible information to be transmitted periodically. All tags are configured to operate in beacon mode, broadcasting the information every few seconds in a range of about 100 meters. Active RFID tag transmission technology is much lighter in complexity and in coping with harsh environments compared to that of WSNs making the architecture robust and resistant.

The Collector Unit could be held by any approved personnel or installed in a vehicle of existing organized transportation systems such as public bus services, postal office vehicles, taxis, municipal police vehicles, etc. in order to unobtrusively collect the required information. As the vehicle roams around the city, the RFID reader reads the tags and collects and stores the information to the storage area. Given the short "packet" size (less than 128 bits), the tag range (100 meters), and a realistic average speed of 40 km/h in a city, the tag beacon interval is extracted so as the passing vehicle is able to complete all necessary transactions with the smart-bin. The use of RFID technology allows reading multiple tags simultaneously (group of nearby bins), without collisions and re-transmissions, as it would be the case of a wireless sensor network.

In order to cope with increased telecommunication costs and infrastructure upgrades, the Collector Unit defers transmissions to the Central System until an Internet connection becomes available. The Delay Tolerant Networking paradigm ensures that the information is retained in the bins, until a Collector Unit passes nearby and then in the Collector Units, until an Internet connection is available.

## 8. DATA MINING & PREDICTION ANALYSIS SUBSYSTEM

The data mining and prediction analysis subsystem is based on the data warehouse infrastructure, from which the data are extracted, so that the various data mining scenarios can be realized. The data mining process can be quite challenging and time consuming since the efficient parametrization of the chosen variables requires experience on behalf of the scientists and is mostly empirical. The research team tried to execute different scenarios in order to analyze the behavior of existing data mining methods on this dataset and to investigate the existence of association rules and possibly to make interesting conclusions.

### 8.1 Data mining algorithms study

Below we present the process that was followed in order to study and adjust the data mining algorithms within the scope of the project, as well as the visualization of the results. The whole subsystem is realized in a Microsoft SQL Server RDBMS, utilizing the relevant Analysis Server (Microsoft Analysis Services), where a Mining project has been created in the supported programming environment, linking the central database with the Analysis Server.

At the first stage, a data mining structure was constructed based on a single table, merging the data from the various RDBMS tables. While creating the data mining structure, the variables that took part in the analysis process where chosen. Furthermore, each variable type was carefully defined in order to be effectively handled by the data mining algorithm. For example, some variables are continuous such as dates, some are distinct and some must be discretized for the sake of the analysis at specific intervals. Also, one of the variables was defined as the key in order to ensure distinctness between different records.

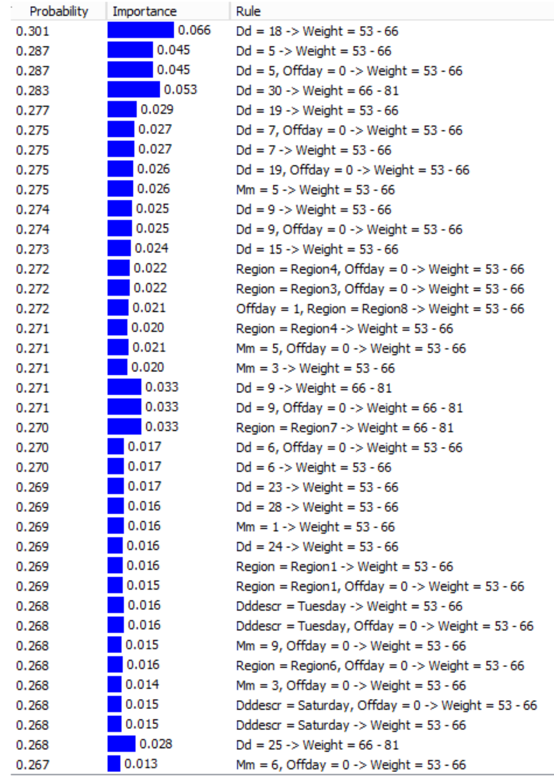
In the scope of the subsystem a clustering model was defined, as well as variations of it with different parameters, aiming to reveal the general distribution range of variables in the entire range of data. From the different clusters we got from the analysis, we tried to define which variable(s) have a distinct role and can serve as representatives and whether it makes sense to separate the data into clusters.

In parallel with the above, a study was performed using other data mining tools, because not all methods were fully covered by the selected configuration. The same data were imported in WEKA which provides a richer set of data mining methods, and also additional algorithms to some of the methods. Using this configuration, we tried to find a combination of variables that could link the data and provide some insight. Also, a classification process took place according to the distance between the waste bins and we investigated how this affects the total cost.

## 8.2 Results visualization

Besides the above, using the Analysis Server, we constructed a cube in order to visualize the data. The cube has as many dimensions (columns) as the data we want to use. Queries on these data are performed utilizing the PivotViewer control, a Silverlight web browser plug-in. We made use of PivotViewer in order to implement the main interface for user queries. The choice was made based on the fact that PivotViewer leverages Deep Zoom which is the fastest, smoothest, zooming technology on the Web. As a result, it displays full, high resolution content without long loading times. Moreover, the animations and natural transitions provide context in a user-friendly way and prevent users from feeling overwhelmed by large quantities of information. PivotViewer enables users to interact with thousands of objects at once while sorting and browsing data, making it possible to track trends easily and find the information they are looking for. In PivotViewer, the user can apply filters on variables based on the respective variable range. Filters can also be applied simultaneously to many variables. Apart from filtering, the data can be sorted based on any variable without using filters. For example, to show the fullness of the waste bins for a specific week of a month in some specific time range, the appropriate filters can be selected and sort the data by date.

In figure 5, we present a small sample of results for the experiments we conducted. The model that was chosen made use of association rules. The concept was to predict the weight percentage of the waste bins based on the day, month and region where the waste bin is located, as well as the respective weight percentage of the waste bin that were given as input.



| Probability | Importance | Rule |
|---|---|---|
| 0.301 | 0.066 | Dd = 18 -> Weight = 53 - 66 |
| 0.287 | 0.045 | Dd = 5 -> Weight = 53 - 66 |
| 0.287 | 0.045 | Dd = 5, Offday = 0 -> Weight = 53 - 66 |
| 0.283 | 0.053 | Dd = 30 -> Weight = 66 - 81 |
| 0.277 | 0.029 | Dd = 19 -> Weight = 53 - 66 |
| 0.275 | 0.027 | Dd = 7, Offday = 0 -> Weight = 53 - 66 |
| 0.275 | 0.027 | Dd = 7 -> Weight = 53 - 66 |
| 0.275 | 0.026 | Dd = 19, Offday = 0 -> Weight = 53 - 66 |
| 0.275 | 0.026 | Mm = 5 -> Weight = 53 - 66 |
| 0.274 | 0.025 | Dd = 9 -> Weight = 53 - 66 |
| 0.274 | 0.025 | Dd = 9, Offday = 0 -> Weight = 53 - 66 |
| 0.273 | 0.024 | Dd = 15 -> Weight = 53 - 66 |
| 0.272 | 0.022 | Region = Region4, Offday = 0 -> Weight = 53 - 66 |
| 0.272 | 0.022 | Region = Region3, Offday = 0 -> Weight = 53 - 66 |
| 0.272 | 0.021 | Offday = 1, Region = Region8 -> Weight = 53 - 66 |
| 0.271 | 0.020 | Region = Region4 -> Weight = 53 - 66 |
| 0.271 | 0.021 | Mm = 5, Offday = 0 -> Weight = 53 - 66 |
| 0.271 | 0.020 | Mm = 3 -> Weight = 53 - 66 |
| 0.271 | 0.033 | Dd = 9 -> Weight = 66 - 81 |
| 0.271 | 0.033 | Dd = 9, Offday = 0 -> Weight = 66 - 81 |
| 0.270 | 0.033 | Region = Region7 -> Weight = 66 - 81 |
| 0.270 | 0.017 | Dd = 6, Offday = 0 -> Weight = 53 - 66 |
| 0.270 | 0.017 | Dd = 6 -> Weight = 53 - 66 |
| 0.269 | 0.017 | Dd = 23 -> Weight = 53 - 66 |
| 0.269 | 0.016 | Dd = 28 -> Weight = 53 - 66 |
| 0.269 | 0.016 | Mm = 1 -> Weight = 53 - 66 |
| 0.269 | 0.016 | Dd = 24 -> Weight = 53 - 66 |
| 0.269 | 0.016 | Region = Region1 -> Weight = 53 - 66 |
| 0.269 | 0.015 | Region = Region1, Offday = 0 -> Weight = 53 - 66 |
| 0.268 | 0.016 | Dddescr = Tuesday -> Weight = 53 - 66 |
| 0.268 | 0.016 | Dddescr = Tuesday, Offday = 0 -> Weight = 53 - 66 |
| 0.268 | 0.015 | Mm = 9, Offday = 0 -> Weight = 53 - 66 |
| 0.268 | 0.016 | Region = Region6, Offday = 0 -> Weight = 53 - 66 |
| 0.268 | 0.014 | Mm = 3, Offday = 0 -> Weight = 53 - 66 |
| 0.268 | 0.015 | Dddescr = Saturday, Offday = 0 -> Weight = 53 - 66 |
| 0.268 | 0.015 | Dddescr = Saturday -> Weight = 53 - 66 |
| 0.268 | 0.028 | Dd = 25 -> Weight = 66 - 81 |
| 0.267 | 0.013 | Mm = 6, Offday = 0 -> Weight = 53 - 66 |

**Figure 5. Prediction of waste bin weight percentage based on association rules.**

# 9. ROUTING OPTIMISATION SUBSYSTEM

In this section we describe the static and dynamic data we store and calculate for each bin, we describe the construction methodology used for the graph models, we highlight the routing particularities of Dynacargo, we outline our algorithmic approach and discuss known families of routing problems that present similar characteristics and technical limitations.

## 9.1 Static and Dynamic Bins' Data

At first, we classified all bins according to the type of their geographic region, in order to be able to analyze historic data (plus future data) and facilitate administrators to configure the routing subsystem. The categories we chose are general and can be applied to any municipality. They are outlined in Table 1.

**Table 1. Generic Region Types**

| Region Type ID | General description | Regions of Nafpaktia |
|---|---|---|
| 1 | Historic Center | Historic center of Nafpaktos city (around the castle) |
| 2 | City | City of Nafpaktos (densely populated areas with blocks of flats) |
| 3 | Suburb | Suburb area of Nafpaktos city (mainly houses and small apartment buildings) |

| | | |
|---|---|---|
| 4 | Town | Antirrion |
| 5 | Village – Rural area | Routes to Panw Neokastro, Paleochoraki and St Ioanni monastery |
| 6 | Remote area | Route of Apodotia. |

The following **static data are stored for each bin**:

- id
- longitude,
- latitude,
- altitude,
- bin type (small or large and its capacity in lt)
- Region Type ID (1 to 6 according to the above table)
- id of fixed route that it belongs
- number in the order of collection within the fixed route that belongs
- Average daily volume growth rate (Winter)
- Average daily volume growth rate (Summer)

Additionally, we store the following **dynamic data for each bin** that are provided either by system administrators or other subsystems, such as Bins Data Gathering Subsystem or the Data Mining Subsystem:

**Dynamic data by administrator**

- **Day Limit:** Maximum time period (in days) that a bin cannot be serviced, depending on the seasonal period and the location of the bin.
- **Upper Limit (% volume):** If the bin fill-level exceeds this limit at the collection time, then it should be collected.
- **Lower limit (% volume):** If the bin fill-level not exceeds this limit at the collection time, then it should not be collected.
- **Waste average weight (kg) per liter** per region type.

The dynamic data as estimated for the municipality of Nafpaktia are shown in table 2. These values can be modified by the administrator whenever necessary.

**Table 2. Administrator's Dynamic Data for Nafpaktia**

| Region Type ID | Waste average weight (kg) per lt | Lower limit (% volume) | Upper Limit (% volume) | Day Limit (winter) | Day Limit (summer) |
|---|---|---|---|---|---|
| 1 | 0.10 | 20 | 60 | 3 | 2 |
| 2 | 0.10 | 20 | 60 | 3 | 2 |
| 3 | 0.09 | 20 | 70 | 5 | 3 |
| 4 | 0.08 | 20 | 70 | 5 | 3 |
| 5 | 0.07 | 20 | 75 | 7 | 4 |
| 6 | 0.06 | 20 | 80 | 7 | 4 |

**Dynamic data by collectors**

- ID Collector
- Timestamp (date and time of collection)
- Bin fill-level percentage as collected several times daily by various collectors passing by the bin, carrying a Bins Data Gathering Subsystem.

**Dynamic data by DataMining Subsystem**

- Timestamp of prediction (date and time of the next scheduled routes)
- Bin fill-level percentage forecasted value for the above timestamp.
- Waste weight based on the forecasted volume. It is calculated by using the waste average weight (kg) per liter per region type.

**Dynamic data by refuse truck**

- Actual Bin fill-level percentage collected by the refuse truck. This value is compared to the forecasted value, in order to improve the accuracy of future predictions.

## 9.2 Refuse Tracks' Data

For each refuse truck the administrators provide the following data:

- Registration number
- Capacity
- Average fuel consumption
- Daily availability
- Which bins can service (which bin types and/or which geographical regions)
- Parking area that it is currently parked

## 9.3 Graph Model

Most non-urban municipalities in Greece include geographically scattered villages or towns, thus in our model we group of a village or a town into bin clusters. For each pair of bins within a cluster the actual distance between them is calculated considering traffic restrictions, in both directions. Even if the distance of a bin pair in both directions is the same, the travel costs may not be, in case of an uphill route. For each bin cluster, we set one or more entry / exit points, usually 1, 2 or 3 for a typical village or town. By using the entry / exit points of all clusters, the landfill location and the vehicle parking space, we construct another graph. For each pair of nodes in this graph, the actual distance between them is calculated considering traffic restrictions, in both directions.

The actual distance between two geographic points is calculated by using the QGIS tool and the Open Street Map platform via the online routing API that provides. We are based on open standards and the commonly used WGS84 coordinate system, thus this procedure could be also performed by using the routing API of Google or Bing Maps.

## 9.4 Dynacargo Routing Problem

2100 bins and 2 bin types. 11 refuse trucks, 6 kinds of trucks. Some trucks are servicing only one bin type. Only one small truck can service bins in the historic center of Nafpaktos, because of narrow roads. This small truck is also used to collect other small bins, some of which are located at long distances, mainly due to lower fuel consumption. This small truck does not go to the landfill, but tranships its waste cargo to a larger refuse truck.

All routes should begin and end in a single truck parking area. Each truck route (except the route of the small truck) must pass from the landfill once and then move directly to a parking area without servicing other bins.

## 9.5 Our Algorithmic Approach

Currently, we are designing our routing algorithms in order to implement and evaluate them. In this section we outline our approach.

**Step1: Calculate the bins' priority**

Based on bins dynamic data and the seasonal period, the priority to service a bin is calculated: 1 (do not service), 2 (desirable but not necessary to service it), 3 (must be serviced). The algorithm that calculates the priority for each bin can be parameterized by system administrators and it is described below :

- If we have reached the day limit since the last collection of the bin, then Priority =3
- If forecasted value of the bin fill-level >= Upper Limit, then Priority =3
- If Lower Limit <= forecasted value of the bin fill-level < Upper Limit, then Priority =2
- If forecasted value of the bin fill-level < Lower Limit, then Priority =1

**Step2: Solve the subproblem of the small truck**

- Find the best route that service all small bins with priority 3 including the small bins in the historical center of Nafpaktos (no other truck can service them) and the most distant ones (because of lower fuel consumption).
- If there is available capacity in the small vehicle, chose some extra small bins with priority 2 to include in the route of the small vehicle, preferring:
    - The bins that cause the smaller presentence increase on the initial route.
    - The bins that the next day will probably have priority 3 (ie. if current %volume + daily volume growth rate >= upper limit %volume or if tomorrow the bin will reach the collection day limit).
- Reset the load of all small bins that belong to the small truck route. Assign the total load of the small truck route to the last bin in its route, and set priority 3 for this bin. We consider this bin as the location of the waste transship to a bigger truck. The algorithm should synchronize the routes of the two trucks so as to minimize the waiting time of the truck that arrive first at the transship location.

**Step3: Solve the total routing problem**

The complete problem will be solved in two levels: within each cluster and in inter-cluster level.

In overall, the main routing problem is defined as finding a set of optimal routes (lowest total km and fuel consumption) for a subset of the available trucks, that begin from the parking area, service every bin of priority 3 (including the small truck node), and end up in landfill and then in the parking area, satisfying the capacity constraints of the available trucks. Depending on the available capacity of each truck, we could assign to its route some extra bins of priority 2, starting from those that cause the less percentage increase to the length of the initial route and/or those that will probably have priority 3 the next day.

**Step4: Dynamic routing**

- **At bin level:** The final routes will be enriched by those bins with priority 2 or 1, that are located on the route, but the algorithm did not choose. These bins will be collected by the refuse truck, only if their fill-level exceeds the upper volume limit of each bin.
- **At route level:** as the bins are collected by the refuse truck, the truck system calculates the total difference between the actual wastes in bins with the predicted values. If the collected waste is more than the predicted one, then start removing bins of priority 2, from the end of the route.

## 9.6 Assigning to known problem families

This specific problem is a variant of the Capacitated Vehicle Routing Problem (CVRP) where the approach methods are grouped into the following main categories [11]: Branch-and-Bound, Branch-and-Cut (hybrid methods between Branch-and-Bound and Cutting Plane methods), Set-covering based algorithms, Heuristics and Metaheuristics.

Furthermore, Dynacargo routing problem can also be approached as an Orienteering Problem (OP). The OP is a combination of vertex selection and determining the shortest Hamiltonian path between the selected vertices. The OP's goal is to maximize the total score collected, while the TSP tries to minimize the travel time or distance. Furthermore, not all vertices have to be visited in the OP. The main categories of OP are the following: [12]: Orienteering problem (OP), Team Orienteering Problem (TOP), Orienteering Problem with Time Windows (OPTW), Team Orienteering Problem with Time Windows (TOPTW).

For the implementation of our algorithms, we will be based on the library Or-Tools of Google, possibly in combination with other tools.

# 10. CONCLUSIONS

Sustainable growth, in regards to urban areas, requires intelligent waste collection management. Dynacargo serves this need by developing a cargo-centric waste transport management system and by implementing a fully functional instance regarding domestic waste collection in a real life large scale scenario.

This is achieved by expanding traditional fleet management functionality in two manners. One breakthrough that Dynacargo utilizes is to stream near real-time waste related information (fill level of waste bins) into the monitoring and decision support process, prior to collection from waste concentration points. If this information is not available or not efficient from a cost-benefit perspective, it is substituted by historical data in order to predict waste bins status. The second breakthrough that Dynacargo introduces is active citizen involvement, by turning them into active information producers and consumers. Dynacargo utilizes low-cost durable RFID tags, along with alternative network protocols such as DTN, is 4G-ready and utilizes dedicated dynamic routing algorithms in order to minimize telecommunication and hardware costs.

This paper presents the scope of the Dynacargo project, the architectural modules and the interrelations between them, as well as the research issues and development progress of some significant modules, namely the Bin Subsystem, the Data Mining Subsystem and the Routing Subsystem.

## 11. ACKNOWLEDGMENTS

This work has been financially supported by the Greek General Secretariat of Research and Technology and European Union under the project “11SYN_10_456: Dyanacargo”.

## 12. REFERENCES

[1] Asimakopoulos G., Christodoulou S., Gizas A., Triantafillou V., Tzimas G. and Vasilopoulos N. (2014). Functional specifications and architecture of a dynamic routing system for urban waste collection using real-time information. IADIS 13th International Conference on WWW/Internet, 25-27 Oct 2014, Porto, Portugal, pp. 203-210.

[2] Al Mamun M.A., Hannan M.A., Hussain A., Basri, H. (2013). Wireless Sensor Network Prototype for Solid Waste Bin Monitoring with Energy Efficient Sensing Algorithm. *IEEE 16th International Conference on Computational Science and Engineering*.

[3] Arebey M., Hannan M A, Basri H., Begum R A, Huda A. (2010). Solid Waste Monitoring System Integration based on RFID, GPS and Camera. *Proceedings of IEEE Intelligent and Advanced Systems (ICIAS)*

[4] Faccio M, Persona A and Zanin G. (2011). Waste collection multi objective model with real time traceability data. *In Waste Management, Elsevier,* Vol. 12, No. 2,pp 391-405.

[5] Gomes T., Brito N., Mendes J., Cabral J. and Tavares A. (2012). WECO: A wireless platform for monitoring recycling point spots. *Proceedings of 16th IEEE Mediterranean Electrotechnical Conference (MELECON)*.

[6] Gukhool O., Bokhoree C., Mohee R. and Jamnejad G. (2008). Solid Waste Collection Optimization using a Route Time Approach: A case study for an urban area of Mauritius, *23rd Int. Conf. on Solid Waste Technology and Management, Philadelphia PA, p501.*

[7] Johansson Ola (2006). The effect of dynamic scheduling and routing in a solid waste management system. *In Waste Management,* Vol. 26, No. 8,pp 875 – 885.

[8] Municipality of Athens. (2003) Estimation, Evaluation and Planning Of Actions for Municipal Solid Waste Services During Olympic Games 2004. Municipality of Athens.

[9] Nuortio T, Kytöjoki J, Niska H and Bräysy O (2006). Improved route planning and scheduling of waste collection and transport. *In Expert Sys Appl*, Vol. 30, No. 2,p 223-232.

[10] Roshan I. and Akshai M. (2013). SVASTHA: An Effective Solid Waste Management System For Thiruvalla Municipality in Android OS. *Proc. of IEEE Global Humanitarian Technology Conf. South Asia Satellite*.

[11] Toth P. and Vigo D. (2001). The vehicle routing problem. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA.

[12] Vansteenwegen P., Souffriau W., Oudheusden D. V. (2011). The orienteering problem: A survey. European Journal of Operational Research, Vol. 209, Issue 1, pp. 1-10.

# A Versatile Scalable Smart Waste-bin System based on Resource-limited Embedded Devices

A. Papalambrou, D. Karadimas, J. Gialelis, A. G. Voyiatzis
Industrial Systems Institute, "Athena" Research Innovation Center in Information, Communication and Knowledge Technologies
Patras Science Park, Stadiou Str., Platani, Patras, Greece
{papalambrou, karadimas, gialelis, bogart}@isi.gr

***Abstract*— This work presents the architecture, modelling, simulation, and physical implementation of a versatile, scalable system for use in common-type waste-bins that can perform and transmit accurate fill-level estimates while consuming minimal power and consisting of low-cost embedded components. The sensing units are based on ultrasonic sensors that provide ranging information which is translated to fill-level estimations based on extensive simulations in MATLAB and physical experiments. At the heart of the proposed implementation lies RFID technology with active RFID tags retrieving information and controlling the sensors and RFID readers receiving and interpreting information. Statistical processing of the simulation in combination with physical experiments and field tests verified that the system works accurately and efficiently with a tiny data-load fingerprint.**

***Keywords—urban solid waste; ultrasonic sensors; waste-bin fill-level estimation; active RFID tag; smart-cities; sustainability***

## I. INTRODUCTION

Waste in general appears in many different forms such as agricultural, biomedical, chemical, electronic, mineral, organic/inorganic, radioactive, and urban/municipal; etc. All of the above waste forms, apart from urban, are characterized by specific collection points, uniform and predictable production, and equal, usually long, filling periods. Urban waste on the other hand involves numerous waste bins that exhibit significant filling variations (over time or location) and diverse requirements for emptying, from sporadic (a few times within a week) to very frequent (several times a day).

A poor or inappropriate urban waste collection may introduce or amplify problems affecting the urban environment, the authorities, or even the citizens. One of the major problem-raising source is the remote fill-level estimation, since such a waste management could potentially lead to great improvements and/or optimizations in the overall process, like: a) waste-bins' number, capacity and placement; b) garbage trucks' routes that could in turn lead to reductions in traffic congestion, noise levels, emissions, and operating costs; c) public image and general sense of duty for public sanitation.

In some cases, urban waste is separated, according to its source material, to respective waste bins types such as glass; books; textiles; oils or even silos. Despite that during the past years many commercial solutions [1], [2] have appeared dealing with the fill-level estimation of such waste bins containers, the fill-level estimation of solid-waste bins is a challenging task. The detection of the fill-level for solid-waste urban bins presents many difficulties due to the various irregularities of the bin filling process, such as the irregular shape and the variety of the included materials. Furthermore, the physical experimentation with this process is difficult as the number of experiments needed to reach safe statistical conclusions is very large. More challenges exist for the economical and energy efficient data aggregation from a large number of bins while the harsh environmental conditions (e.g., humidity, temperature, and dust) can significantly affect the sensor measurement accuracy and reliability. These challenges are being dealt within this work, which is part of the Dynacargo project [3].

The paper is organized as follows: Section II presents the scope of the Dynacargo project and revises existing approaches. Section III is dedicated to the description of the proposed architecture and its components while Sections IV and V demonstrate the modelling and statistical processing towards fill-level estimation, and the verification method followed along with power characteristics of the deployed system. Finally, the paper concludes with Sections VI and VII justifying the system's possibilities regarding scalability and future actions, respectively.

## II. OVERVIEW OF DYNACARGO PROJECT AND RELATED WORK

### *A. Overview of Dynacargo Project*

Dynacargo is an ongoing research project that introduces a breakthrough approach for cargo management systems, as it places the hauled cargos in the center of a haulage information management system, instead of the vehicle. Dynacargo attempts to manage both distribution and collection processes, providing an integrated approach. In order for the Dynacargo project team to achieve their goal, a fill-level monitoring system is placed on waste-bins in order to produce crucial data that is fed via diverse communication channels into the cargo management system. Besides feeding the management system with raw data from waste bins, data mining techniques are used in order to predict current waste-bins fill-status and easy-to-use mobile and web applications will be developed to encourage citizens to participate and become active information producers and consumers.

The Dynacargo project overall aim is to develop a near real-time monitoring system that monitors and transmits waste-bins' fill-level, in order to dynamically manage the waste collection more efficiently by minimizing distances covered by refuse

978-1-4673-7929-8/15/$31.00 ©2015 IEEE

vehicles, relying on efficient routing algorithms, as described in [4].

In this work, we present the part of the Dyncargo project that is related to the waste-bin fill-level monitoring and data communication subsystem.

### *B. Related Work*

Currently, several research approaches exist tackling issues related with waste collection and management and especially information collection at the point of waste disposal [5][3]-[11]. However, most of them are expensive in terms of total required equipment, while some of them deal with the problem of collecting only recyclable waste, which can be collected at less frequent intervals and have a more stable production frequency. Moreover, a few fill-level sensing prototypes and architectures can be found in the literature, which are shortly described hereby.

A prototype waste-bin system was constructed for the city of Pudong, Shangai, PR China [12] and in Malaysia [13]. These systems are very similar to each other since they both employ GPRS communication to send waste-bins images to a central server for processing and bin fill-level estimation. Another approach for waste-bin fill-level estimation can be realized with sensors based on a modulated infrared beam detected by a photodiode. As described in [14], multiple line-of-view sensors in parallel and a majority decision system can be used in order to count the times a waste-bin is opened or moved and finally correlate the result with the amount of waste inside the bin. Apart from the assumptions and simplifications in the above correlation, the accuracy of the measurement can be influenced by transparent objects, the reflection of light on object surfaces, the ambient light, and the dirt over emitter and detector surfaces.

The SEA project designed a smart bin prototype using an ultrasonic sensor and IDEA's ArgosD (TelosB) sensor nodes running a custom TinyOS application. The smart bin connects with gateways that are based on the New/Linux OS, realizing a three-layer architecture for information collection [15]. A similar architectural approach is followed in [16], where two sensors (an ultrasonic for fill-level and a load cell for weight) are used to transmit their sensed values to nearby gateways installed in light poles. A three-tier approach is followed also in [5], where multiple sensors are used in the lower tier: ultrasonic for fill level, load cell for weight, temperature and humidity, Hall Effect and accelerometer for detecting bin cover open events. The sensed values and operational parameters, such as bin identity, date, time, and battery power level, are collected and transmitted when a cover opening is sensed, thus achieving energy-efficient, real-time fill-level reporting.

The EU FP7 OUTSMART project designed a mesh wireless sensor network for Berlin, Germany [17] employing an ultrasonic sensor for fill-level estimation and a wireless sensor network based on IEEE 802.15.4 for connecting nodes with gateways through multiple short-range hops. The EU FP7 Future Cities project [18] developed an urban-scale living lab in the city of Porto in Portugal. In the context of the project, the Municipality of Porto has developed an innovative data collection system for monitoring fill-level of garbage containers.

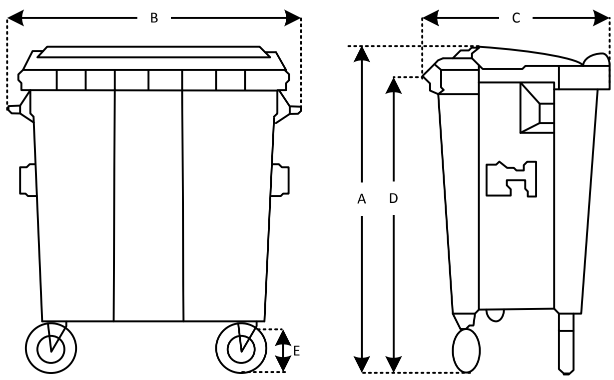
The EU FP7 Straightsol project [19] streamlines charity collection (e.g., clothes and books) from donation banks installed in public spaces and retail shops by using an infrared sensor for fill-level estimation (at 20% reported accuracy) and transmitting the information twice per day through GSM for scheduling next day's collection. A small-scale actual pilot included 37 donation banks, 50 retail shops, and 5 vans and resulted in an estimated 5% revenue gain [20].

## III. Smart Waste-Bin System Architecture

Our proposed smart waste-bin system is envisaged to be mounted on the top lid of a waste-bin and it consists of the sensing units, an active RFID (Radio-Frequency IDentification) tag for data aggregation and transmission, as well as a protective enclosure for the sensors, and the RFID tag that may optionally include an external battery source. On the other side, the waste-bin system interacts with an active RFID reader which could be optionally accompanied by a bare-bone-pc, i.e. a Raspberry Pi for enhanced capabilities. The aforementioned modules are described in the following paragraphs starting from the waste-bin which constitutes the main physical component and moving towards the receiver side.

### *A. Mobile Garbage Bin*

The physical system is based on the standard type of a mobile garbage bin, also known as a "wheelie bin", used in the municipality of Nafpaktia, Greece, where the pilot of the Dyncargo project will take place. The bin has a main compartment whose internal dimensions are approximately 110x90x90 cm, consisting of slight curved-shaped walls. It has a lid-opening mechanism and is made of durable hard plastic, as depicted in Fig. 1. The lid of the bin was chosen as the point of placement of the sensors because the main compartment is subject to harsher conditions due to the presence of the waste as well as the washing procedure of the bins.




| | | |
|---|---|---|
| Volume (lt) | | 1148 |
| Maximum Load (kg) | | 530 |
| Total Height (mm) | A | 1460 |
| Total Width (mm) | B | 1360 |
| Total Depth (mm) | C | 1050 |
| Upper Edge Comb (mm) | D | 1230 |
| Wheel diameter (mm) | E | 200 |

Fig. 1. The mobile garbage bin.

### *B. Sensing Unit*

A literature-based comparison in research and industrial efforts of various solutions including infrared proximity sensors, optical sensors and ultrasonic sensors, indicated that ultrasonic sensors are the most suitable solution for the purpose of the presented architecture taking into account the harsh environmental conditions (e.g., humidity, temperature, and dust) that can significantly affect the sensor measurement accuracy and reliability. The ultrasonic sensors are advantageous in providing ranging measurements independently of the contained objects, thus making possible the corresponding translation into fill level measurements. Moreover, ultrasonic sensors are most suitable to our application because they can be placed on the lid and thus avoid the harsh conditions (contact with waste, washing procedure etc.) of the main compartment. Finally, the chosen ultrasonic sensors can have a beam with a wide field of view and therefore the whole bin can be sensed with a limited number of sensors, thus reducing cabling and interconnection needs.

For the purpose of this task, the ultrasonic sensors of Maxbotix were selected which provide a variety of detection patterns, accuracy and durability. The wide selection of available detection patterns in the same family means that this setup can be expanded and modified for bins of various sizes. For the standard type of bin used in this pilot, the MB1040 LV-MaxSonar-EZ4 [21] was selected whose beam dimensions fit well with the depth and width of the bin as depicted in Fig. 2.

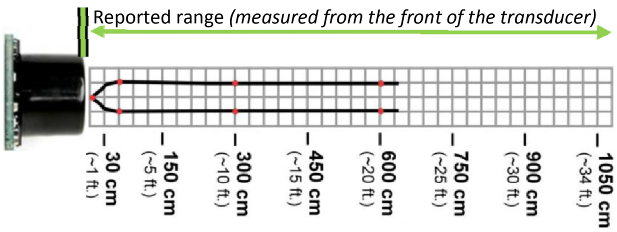



Fig. 2. MB1040 LV-MaxSonar-EZ4 beam characteristics drawn to a 1:95 scale.

Since the resolution of the ultrasonic sensors is only a few centimeters (about an inch), the selected solution can offer fine-grained accuracy for the purpose of the application. The ultrasonic sensors should be mounted in the bin lid, exposing only a small part of the sensor body. Since the sensors will operate unattended in the field, low power consumption models that also offer IP-67 protection rating can be used. An analog output is provided for the transmission of the measurements.

The measurement of the sensors essentially provide the distance to the nearest object of efficient size. Since waste in the form of irregular-shaped objects can occupy the bin space in various arrangements, there is typically some difference between the highest bin level and the actual fill percentage. Accuracy of the fill-level estimation can be increased by providing more sensing points which will decrease the influence of irregular arrangements as these will only be limited to the sensing area of each sensing point. Following experimentation with placement and number of sensors it was concluded that a single sensor placed in the center of the lid was providing good results but over-estimating the fill level due to the typically higher level of waste in the center of the bin. As a result, various combinations of two sensors were examined. It was concluded that more accurate results were provided by placement of two sensors in a way that their areas of detection would not overlap. This prevented problematic arrangements from affecting more than one sensor, increased the resolution of the measurement for each independent area and also simplified the task of estimating the fill level. Having two sensing points with independent areas of detection allows for the fill estimates of each area to be summed so as to provide the total fill estimate of the bin. The exact placement of the sensor is limited by the distance to the bin walls which must be outside the beam or they will provide a false detection. The exact number of sensors to be used depends on the size of the bin and the beam pattern of the sensors. Typically, using more narrow-beam sensors will increase resolution and accuracy while using fewer wide-beam sensors will reduce cost but increase measurement ambiguity. The final number of sensors for our application was based on compromising the above parameters with cost limitations.

### C. Active RFID Tag

Active RFID tags have been selected as the data aggregation and transmission unit for the bin subsystem. The selected tags (ZT-50-mini Tag from TagSense [22]) can operate with standard 3V voltage input providing increased lifetime that equals to millions of beacon transmissions.

The employed active tags offer extra I/O pins for communicating with external devices (as depicted in Fig. 3), thus the ultrasonic sensing units. This setup allows powering the sensors and the RFID tag from the same source, either tag's battery or an external power source. The tag itself supports various operating modes, including standard beacons at programmable time intervals, sleep, wake up at regular intervals and also wake up at external trigger.

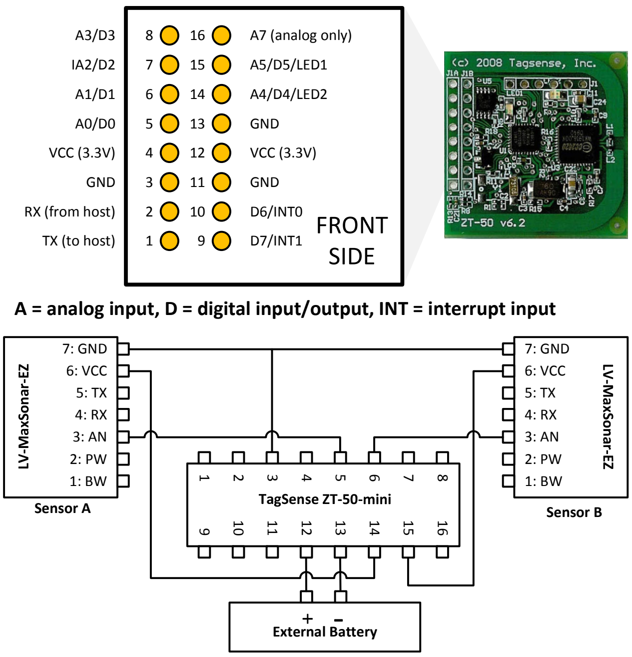



Fig. 3. ZT-50-mini photograph of the front side with a description of the 2-rows header connector and the circuit diagram of the hardware setup.

The combination of the above operating modes allows the extension of the entire bin subsystem's energy lifetime, since minimal power consumption occurs when the tag operates in sleep mode and the sensor is not powered up. When the tag is awake, it powers up the sensors and temporarily stores their values in its internal memory.

In the Dynacargo scenarios, the tag wakes up at predefined times of the day, depending on the location of the waste bin installation and on the desired measurement frequency defined by the Municipality operators. The Dynacargo operation does not assume availability of a fixed network infrastructure that reaches all the installed bins. At a wide area scale of operation, this is a quite realistic assumption, since thousands of sensors are sparsely deployed in a complex city and suburban terrain. System installation and maintenance, ensuring radio coverage, and retaining network formation can become an unmanageable task. The Dynacargo project opts for low-range, point-to-point communications based on RFID technology so as to cope with these issues. Vehicles roaming around the city, equipped with readers, collect the information from the bins. In order to cope with the increased telecommunication costs and infrastructure upgrades, these mobile sinks defer transmissions until an Internet connection becomes available.

### D. Active RFID Reader

The selected active RFID reader (ZR-USB RFID Reader from TagSence [22]) is inherently designed to communicate with the TagSense ZT-50 active tag. The ZR-USB active reader communicates with the active tags using a variable packet length protocol which is designed to conserve power on the tag. The air interface protocol layer is based on the underlying IEEE 802.15.4 industry standard, which is best-known as the physical layer for Zigbee. However, the full Zigbee protocol is not well-suited for most RFID applications since it contains a great deal of overhead and requires a larger program memory on the tag to support routing tables and multi-hop/mesh capability.

In order to support general RFID applications, and conserve power, the TagSense active RFID tags use a star topology, where all tags communicate directly to the reader, and the reader sends control commands to the tags. The TagSense Active RFID tag protocol is also "tag-talks-first" or TTF, which works very well for ad-hoc networks, where tags are continuously entering or leaving the network. Since the reader needs to quickly process all the packets that are being received from multiple tags, the reader does a minimum amount of processing on the tag data and passes it on to the host.

### E. Integrated system

The proposed system has an adaptable, modular and configurable design that allows optimizing its operation for multiple scenarios. In more detail, the integrated system is able to glue the "smart" and "cyber-physical" characteristics with a waste-bin containing any type of wastes consists of two main modules, as depicted in Fig.4:

- the *Field Unit*, that is mounted onto an mobile garbage bin and consists of an active RFID tag, a couple of ultrasonic sensors and optionally an external power battery, as described in previous paragraphs,
- the *Mobile Sink* consists of the active RFID reader, described in previous paragraph and a small-form-factor computer (currently a Raspberry Pi), consuming about 10 W of power (maximum) and having two USB 2.0 ports, built-in WiFi and Bluetooth, and one Ethernet port that is supported by GNU/Linux operating system running from a microSD card.

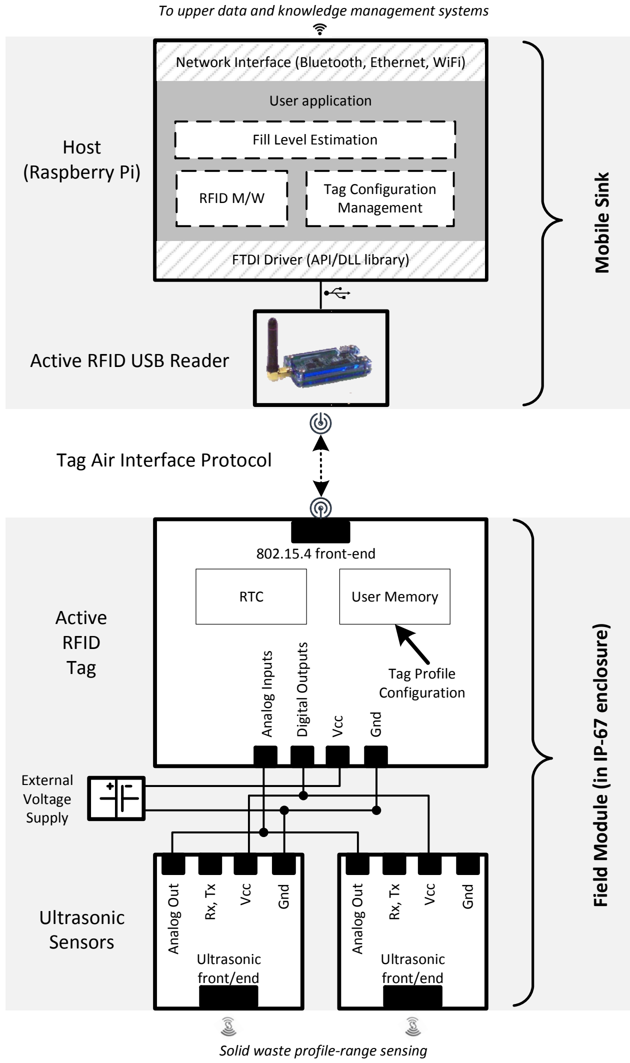



Fig. 4. Proposed smart-bin system architecture.

The waste bins are equipped with the *Field Unit* and the active RFID tags transmit periodically their identity and a sensed value regarding its fill level. This option allows the minimal possible information to be transmitted periodically. All tags are configured to operate in beacon mode, broadcasting the information every few seconds in a range of about 100 meters. Active RFID tag transmission technology is much lighter in complexity and in coping with harsh environments compared to that of WSNs making the architecture robust and resistant.

The *Mobile Sink* could be held by any approved personnel or installed in a vehicle of existing organized transportation systems such as public bus services, postal office vehicles, taxis, municipal police vehicles, etc. in order to unobtrusively collect the required information. As the vehicle roams around the city, the RFID reader reads the tags and collects the information to the sink. Given the short "packet" size (less than 128 bits), the tag range (100 meters), and a realistic average speed of 40 km/h in a city, the tag beacon interval is extracted so as the passing vehicle completes all necessary transactions with the smart-bin. The use of RFID technology allows reading multiple tags simultaneously (group of nearby bins), without collisions and re-transmissions, as it would be the case of a wireless sensor network.

Finally the application running on the mobile sink is responsible for receiving/caching the information from the field units and, when needed, forward this information to upper data and knowledge management systems. The main tasks of this application are: a) tag communication and configuration management; b) tag frame disassembling; c) fill level estimation based on detected range values from sensors; d) creation of an XML-structured packet containing all relevant per tag/waste-bin information (tag ID, estimated fill level, battery level, timestamp of measurement, etc.) that will be stored or forwarded via the internet to appropriate, high-level systems when this is feasible.

### F. Application Scenarios

The system is easily customizable for various application scenarios. Parameters that determine the nature of the scenario can be related to the waste-bin and include its location (urban, suburban or countryside) which correlates to frequency of measurements and possibly the desired accuracy but also the time period (seasonal population changes etc.). Other parameters can be related to the *Mobile Sink* such as which type of vehicle carries it, its power availability and frequency of measurements. Finally, even more parameters of an adopted scenario are related to design choices of the whole system such as desired lifetime of battery before servicing and cost. Various application scenarios were tested for efficacy and power consumption and for the purposes of Dyncargo project the prevalent scenario is as follows.

- The ultrasonic sensors are powered via the RFID tag power output pins in order to be able to control when they operate and thus save power.
- Two sensors are used to estimate the fill level of each bin as they provide the best cost to accuracy ratio.
- The active RFID reader mandates the powering on and off of the sensors when the tags are in range using the minimum of five transmissions. This prevents unnecessary measurements if a reader/tag interaction is not imminent.
- The active RFID tag is programmed to transmit at an interval which will be sufficient for five transmissions to take place while in range with a reader.

## IV. WASTE-BIN FILL LEVEL ESTIMATION

This section presents bin, waste and sensor modelling, measurement simulation and data statistical processing that was necessary to be performed in order to extract an accurate fill-level estimation pattern of waste-bins that contain any type of materials and incorporate this pattern in the Mobile Sink. The modelling was based on parameters of the actual physical system.

### A. Waste-Bin Model

The waste-bin model that was developed is based on the use of a three dimensional matrix representing the entire waste-bin's internal volume. This model was developed using MATLAB because it offers powerful matrix handling tools. The main matrix of the model, the bin matrix, is a matrix whose dimensions are equal to the bin dimensions in centimeters. This allows both an accuracy that is equivalent (or better) than the physical sensors will provide but is also lightweight enough for multiple random simulations. Each value of the bin matrix represents the state of a cubic centimeter found inside the bin. By default, the bin matrix is a zero matrix which corresponds to the empty state of the bin. The size of the bin is parametric and any bin size or shape can be used simply by changing the constants.

### B. Sensor Model

The ultrasonic sensors were also modeled as three dimensional matrices. The matrices used were of the same size as the bin matrix with zero values for the areas of the bin which fall outside the detection range of the sensor (Fig. 2) and with unit values for the areas inside the detection area. In order for the correct values of the matrix to be selected, it was necessary to translate the sensor detection diagrams. This was performed manually since sensor detection patterns are only provided schematically by the manufacturer. First, the schematics were converted to parametric distance equations and then the distance equations were used to fill in the values of the sensor matrix which correspond to the detection areas. The number and placement of the sensors, their detection range and their dead range are all parameters that can be changed in the simulation. This allows for a combination of sensor parameters to be simulated as well as different sensor types.

### C. Waste Model

One of the most crucial parameters for the simulation was waste modelling, mostly concerning its generation. The representation of waste was made using the bin matrix by setting 1 in the elementary volumes where waste was present and 0 otherwise. More challenging was the method to generate waste in a way that would be of equivalence to the actual physical problem. In absence of an existing model in the bibliography concerning how waste of irregular shape and size gradually fills a bin, it was chosen to perform extensive random scenarios of bin filling that would cover all possible combinations. Then, the less likely situations can be filtered out by statistical analysis of the derived combinations.

Waste generation in the simulation was performed so that the following two outcomes are delivered:

- the upper surface of the waste, in order to be used for sensor measurement simulation
- the actual volume of the waste, in order to be used for bin fill percentage.

These two parameters don't necessarily relate in a well-defined way since sensors cannot detect possible voids underneath the waste surface.

Finally, the simulation was parameter-controlled based on the following input:

- *Elementary Waste Base-Edge*. This defines the base-edge length of the elementary cubic waste generated. After experimenting with various values, it was derived that an elementary base-edge length of 10 to 20 cm provides both reliable results and is similar to the physical problem.
- *Maximum Waste Height*. This defines the maximum height the generated waste could reach. This parameter was placed in order to give control over the waste generation during the first iteration (filling the lower segments of the bin), thus better representing the physical process.

### D. Simulation Steps and Results

The complete simulation process took place via the following steps: i) Parameters are set (sensors types, bin sizes etc.); ii) The matrices of the bin and the sensors are created; iii) Waste is generated through a random process based on input parameters; iv) The waste surface is calculated; v) The sensor values are calculated based on the smallest distance between the sensor and the waste surface.; vi) The waste volume is calculated; vii) The fill percentage is calculated; and finally viii) A 3D plot of the waste inside the bin is created.

The final results of interest for the purposes of this work are the actual fill percentage and the sensor measurements. Figure. 5 depicts the simulated final results for a 30k-run simulation using an elementary base-edge length of 20cm for the area of responsibility of a single sensor. In this figure, the vertical axis represents the actual fill-level percentage of the area calculated each time based on the waste-bin and the sensor models, while horizontal axis represents the actual distance detected from the sensor-value based on the specific sensor model.

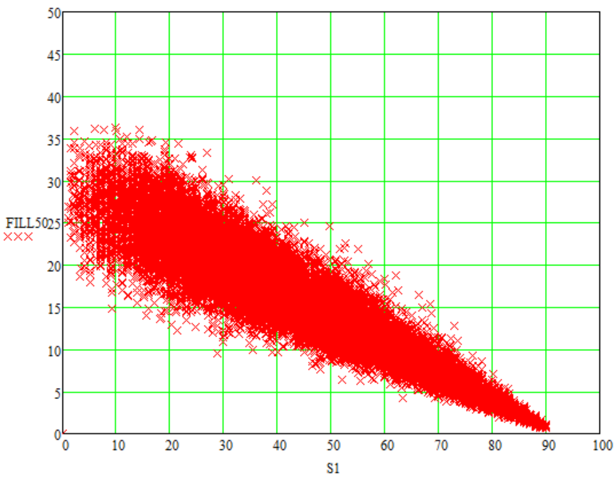



Fig. 5. Scattered data regarding a 30k-runs simulation of single sensor's area of responsibility random fills based on elementary wastes of 20cm base-edge.

It is noticeable that selecting other values for the elementary base-edge length did not alter the shape of the scattered points, but only the amount of spreading of the scatters for medium to low sensor values. This is expected taking into account the actual physical problem.

### E. Statistical Analysis and Fill-Level Estimation Model

The raw results derived from a sufficient number of simulations for a variety of models' parameters selection have been further statistically processed. The selected totally random method for filling a bin followed in the simulations provided different fill-levels for a range of identical sensor values. These fill-levels have been grouped so as to extract the corresponding mean value and deviation for each group, thus treat the bin filling with elementary wastes as a Gaussian process. Finally, Fig. 6 presents the elaborated results regarding the fill-level percentage and deviation versus sensor detected waste range in common axis, whereas the dotted lines represents the corresponding linear regression, which is expressed by the following equation:

$$f(x) = 30.219 - 0.327x \quad (1)$$

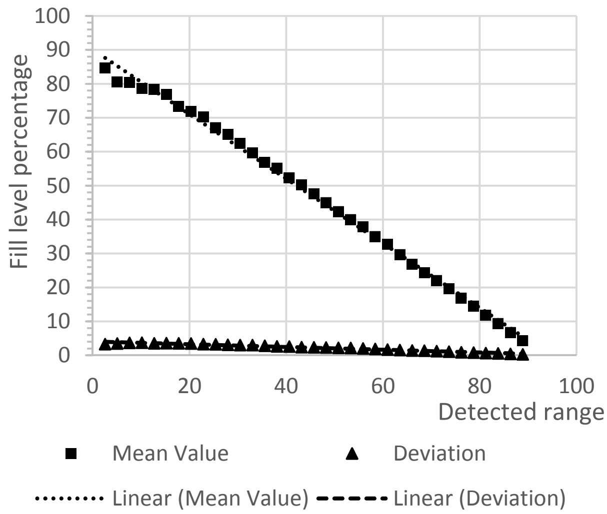



Fig. 6. Fill-level percentage and deviation versus detected waste range.

## V. Verification Measurements and Power Consumption

### A. Experiment Setup for Verification

The verification of the simulations has been performed using the standard waste-bin type mentioned in a previous section. The ultrasonic sensors were mounted on a strip running along the top edge of the main waste-bin compartment. The ultrasonic sensors were initially connected to an Arduino Mega prototyping board with an LCD display which allowed for easy setting up of the experiment and initial measurements, as depicted in Fig. 7.

At the main stage, the experiment was performed using the RFID tags and readers which allows for verification of the complete data chain. The aim of the verification was not to perform an exhaustive series of experiments, which only a simulation can do, but to verify carefully selected scenarios of various fill-levels and waste placements which would indicate that the simulation results correspond to the real world scenario. An experiment which included gradual filling of the bin using cardboards (whose dimension were measured to estimate waste

Fig. 7. Experiment measurement setup for verification.

volume inserted) showed that the algorithm was estimating fill level with better than 90% accuracy for this specific scenario.

In order to interpret the received measurements, it was necessary to construct the appropriate equations leading from the sensor measurement to the reported valued at the RFID reader. The ultrasonic sensor outputs its distance (expressed in inches) measurements as analog voltage divided to 512 distinct levels ranging from 0V to Vcc, therefore it is necessary to also take into account the current voltage of the power supply. The RFID tag digitizes the analog voltage input using 10-bits and transmits a hexadecimal value which is then decoded into decimal centimeters. The chain of equations leading to the interpretation of the measurements is as follows:

$$TagValue = \frac{TagData \cdot BatteryLevel}{1024}(Volts) \tag{2}$$

$$Range = \frac{TagValue \cdot 2.54 \cdot 512}{1024}(cm) \tag{3}$$

thus deriving the final formula:

$$Range = \frac{2.54}{2} TagData(cm) \tag{4}$$

The chosen sensor model for the physical system was made so that it is possible to use two sensors, each measuring one half of the bin without overlapping their areas of detection. This allowed better verification of the measurements as well as the ability to provide a fill level estimate which would be the sum of the two estimates for the half-bin.

### B. *Power Consumption of the Field Module*

As described in Section III, the proposed system consists of the *Field Module*, mounted onto physical waste-bins and the *Mobile Sink*, mounted onto vehicles with approved personnel. Therefore, it is of great importance to calculate the power consumption of the *Field Unit*, since it is going to be deployed in large scale and of course operating under field conditions.

The power consumption of the *Field Module* is equal to the consumption of the sensors and the RFID tags. The sensors do not include any programmable functions so they must be powered via the RFID tag which can be programmed and controlled extensively. Considering the range of voltages the sensors (2.5V to 5.5V) and the tag (2.2V to 3V) can receive, it was chosen to use 3V as the voltage input which can be readily provided with ordinary rechargeable batteries.

Sensors consume 2mA [21] when powered and no power when idle. Assuming k to be the daily number of tag/reader sessions and d the tag transmission interval, the daily power consumptions of the 2-sensors, employed in our prototype, can be derived as:

$$P_{sensor} = 2mA \cdot t_h = 2 \cdot \frac{2 \cdot k \cdot d}{3600} mAh \tag{5}$$

The power consumed by the active RFID tag depends on the number of transmissions since the tag consume virtually no power when idle. According to information provided from the tag specifications [22] and assuming d to be the transmission interval, the daily power consumption of the tag can be derived as:

$$P_{tag} = \frac{86400 \sec}{d \sec / trans} \cdot 4.18 \cdot 10^{-4} \frac{mAh}{trans} \tag{6}$$

However, tags can be programmed to be in sleep mode for part of the day. Assuming $f_{alarm}$ to be the percentage of the day that the tag is powered, the total daily consumption of the system can be derived as:

$$P_{daily} = P_{tag} \cdot f_{alarm} + P_{sensor} \tag{7}$$

For a typical scenario of 3-sec transmission interval, 10-daily tag/reader sessions and 50% sleep time we get an average daily consumption of 6.05 mAh which would allow for continuous operation for 200 days using a typical Lithium battery of 1200 mAh assuming no recharging. Finally, since the *Field Unit* is meant to operate at the open field, a solar power module could significantly extend battery and the system's lifetime.

## VI. System Scalability

A significant advantage of the proposed approach is that the system is not limited to a specific type or size of waste-bin but can be scaled regarding any of its components including waste-bin sizes and shapes, numbers of sensors and information data flow. This is due to the fact that the architectural components have discrete roles and functional independence. Scalability was considered in both modelling and simulation as well as physical component selection. The scalability of the system is described in detail per scalable component in the next paragraphs.

### A. *Scaling for Waste-Bin Geometry and Sensor Type*

The scalability of waste-bin size and shape is made possible based on the following design choices:

- Each sensor has its own independent area of responsibility. The way the system was designed, each sensor is responsible for a specific area of the waste-bin and there is no overlap between areas of various sensors.
- The chosen ultrasonic sensor comes in multiple versions of beam range and width. The type of sensor that was chosen is very versatile because a wide range of models exist with different characteristics concerning their beam width, detection range and resolution. All these models provide the same basic functionalities and logic of measurement.

However, the use of versatile programmable active RFID tags, which provides multiple analog inputs and digital I/O's, allows the co-existence of other sensor types that can trigger events or assist in measurement. Types of sensors that can be of use may include magnetic latches to monitor waste-bin lid status as well as temperature and humidity sensors. As the active RFID tag can sample multiple inputs and collectively transmit them to the reader, addition of various sensors can take place by slightly altering the tag profile and RFID M/W software, regarding tag frame disassembling; without affecting the architecture.

### B. *Scaling for Information Data Flow*

The information data flow across the system is based on RFID technology. Even though the active RFID tag can control a number of sensors, in large applications where more sensors are needed additional RFID tags can be used for the extra sensors. The RFID tags and readers automatically exchange information whenever they are in range. This means that the addition of new waste-bins to the system or even the addition of extra RFID tags per waste-bin, could be realized without affecting the architecture and not even raising the need for reprogramming the tags or the reader. In such a case the collection of the data would be again realized automatically; however it would be of the related backend software's responsibility to correlate the new additions with their physical interpretations.

## VII. CONCLUSIONS AND FUTURE DIRECTIONS

The presented system has been developed and implemented as part of the Dynacargo project, and is part of an innovative smart city framework for urban solid waste collection. Our proposed architecture exploits RFID communication in order to reduce costs, simplify system operation and support scaling at urban level. The fill-level estimation method using ultrasonic sensors provides accurate results both in the simulation and the physical experiments. We also described the system design and the implemented prototype of the presented architecture among with measurements verification that gave very promising results.

The proposed system will be thoroughly tested during the Dynacargo project's pilot that will take place in the Municipality of Nafpaktia in Greece, during 2015 in the forthcoming months. The Municipality has an area of 870.38 $km^2$ and a population of 27,800 people (in 2011 census). Eleven municipal trucks are used for waste collection from the bins. There are about 2,100 solid waste bins installed in about 140 settlements, villages, and towns in the area, including urban, suburban, and rural areas. The pilot tests will allow gaining insights on areas of improvement and experiment with vehicle movement issues in a city terrain.

## ACKNOWLEDGMENT

This work has been financially supported by the Greek General Secretariat of Research and Technology [23] and European Union under the project "11SYN_10_456: Dyanacargo" [3].

## REFERENCES

[1] "UBi: Ultrasonic sensor." [Online]. Available: www.smartbin.com.

[2] "U-Dump: Urban waste collection sensor." [Online]. Available: www.urbiotica.com.

[3] "Dynacargo Project." [Online]. Available: http://dynacargo.cied.teiwest.gr.

[4] G. Asimakopoulos, S. Christodoulou, A. Gizas, V. Triantafillou, G. Tzimas, J. Gialelis, A. Voyiatzis, D. Karadimas, and A. Papalambrou, "Architecture and Implementation Issues, Towards a Dynamic Waste Collection Management System," in Proceedings of the 24th International Conference on World Wide Web Companion (WWW '15 Companion), International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 2015, pp. 1383-1388.

[5] M. A. Al Mamun, M. A. Hannan, A. Hussain, and H. Basri, "Wireless Sensor Network Prototype for Solid Waste Bin Monitoring with Energy Efficient Sensing Algorithm," in 2013 IEEE 16th International Conference on Computational Science and Engineering, 2013, pp. 382–387.

[6] M. Arebey, M. A. Hannan, H. Basri, R. A. Begum, and H. Abdullah, "Solid waste monitoring system integration based on RFID, GPS and camera," in 2010 International Conference on Intelligent and Advanced Systems, ICIAS 2010, 2010.

[7] M. Faccio, A. Persona, and G. Zanin, "Waste collection multi objective model with real time traceability data," Waste Manag., vol. 31, no. 12, pp. 2391–2405, 2011.

[8] T. Gomes, N. Brito, J. Mendes, J. Cabral, and A. Tavares, "WECO: A wireless platform for monitoring recycling point spots," in Proceedings of the Mediterranean Electrotechnical Conference - MELECON, 2012, pp. 468–472.

[9] O. M. Johansson, "The effect of dynamic scheduling and routing in a solid waste management system," Waste Manag., vol. 26, no. 8, pp. 875–885, 2006.

[10] R. Issac and M. Akshai, "SVASTHA: An effective solid waste management system for Thiruvalla Municipality in Android OS," in c2013 IEEE Global Humanitarian Technology Conference: South Asia Satellite, GHTC-SAS 2013, 2013, pp. 254–259.

[11] P. Reis, R. Pitarma, C. Goncalves, and F. Caetano, "Intelligent system for valorizing solid urban waste," in 2014 9th Iberian Conference on Information Systems and Technologies (CISTI), 2014, pp. 1–4.

[12] F. Vicentini, A. Giusti, A. Rovetta, X. Fan, Q. He, M. Zhu, and B. Liu, "Sensorized waste collection container for content estimation and collection optimization," Waste Manag., vol. 29, no. 5, pp. 1467–1472, 2009.

[13] M. A. Hannan, M. Arebey, R. A. Begum, and H. Basri, "An automated solid waste bin level detection system using a gray level aura matrix," Waste Manag., vol. 32, no. 12, pp. 2229–2238, 2012.

[14] P. O. Leitão, M. B. Marques, P. Dep, and I. Porto, "Reliability Of A Line-of-view Sensor For Recycling Point Waste Containers," in ISWA/APESB2009 World Congress, 2009.

[15] S. Longhi, D. Marzioni, E. Alidori, G. Di Buo, M. Prist, M. Grisostomi, and M. Pirro, "Solid Waste Management Architecture using Wireless Sensor Network technology," 2012.

[16] V. Catania and D. Ventura, "An Approach for Monitoring and Smart Planning of Urban Solid Waste Management Using Smart-M3 Platform," in 15th Conference of Open Innovations Association FRUCT, 2014, pp. 24–31.

[17] D. Cassaniti, "A Multi-Hop 6LoWPAN Wireless Sensor Network for Waste Management Optimization," 2011.

[18] "EU FP7 Project: 'Future Cities.'" [Online]. Available: http://futurecitiesproject.eu.

[19] "EU FP7 Project: 'Straightsol.'" [Online]. Available: http://www.straightsol.eu.

[20] F. McLeod, G. Erdogan, T. Cherrett, T. Bektas, N. Davies, D. Shingleton, C. Speed, J. Dickinson, and S. Norgate, "Improving collection efficiency through remote monitoring of charity assets," Waste Manag., vol. 34, no. 2, pp. 273–280, 2014.

[21] "MB1040 LV-MaxSonar-EZ4 Datasheet." [Online]. Available: www.maxbotix.com.

[22] "ZT-50-mini Active RFID Tag and ZR-USB Active RFID Reader." [Online]. Available: www.tagsense.com.

[23] "General Secretariat for Research and Technology." [Online]. Available: www.gsrt.gr.

# An integrated node for Smart-City applications based on active RFID tags; Use case on waste-bins

Dimitris Karadimas, Andreas Papalambrou, John Gialelis and Stavros Koubias
Industrial Systems Institute / "Athena" Research Innovation Center and Department of Electrical and Computer Engineering, University of Patras, Patras, Greece
{karadimas, papalambrou, gialelis}@isi.gr

*Abstract*— **This work presents the concept and methodology as well as the architecture and physical implementation of an integrated node for smart-city applications. The presented integrated node lies on active RFID technology whereas the use case illustrated, with results from a small-scale verification of the presented node, refers to common-type waste-bins. The sensing units deployed for the use case are ultrasonic sensors that provide ranging information which is translated to fill-level estimations; however the use of a versatile active RFID tag within the node is able to afford multiple sensors for a variety of smart-city applications. The most important benefits of the presented node are power minimization, utilization of low-cost components and accurate fill-level estimation with a tiny data-load fingerprint, regarding the specific use case on waste-bins, whereas the node has to be deployed on public means of transportation or similar standard route vehicles within an urban or suburban context.**

***Keywords***— ***internet of things; urban solid waste; ultrasonic sensors; waste-bin fill-level estimation; active RFID tag; smart-city***

## I. INTRODUCTION

A smart city is an urban development vision to integrate multiple information and communication technology (ICT) solutions in order to manage a city's assets. City's assets may include, but not limited to, local departments information systems, schools, libraries, transportation systems, hospitals, power plants, law enforcement, vehicle traffic, waste management and other community services. The goal of building a smart city is to improve quality of life by using technology to improve the efficiency of services and meet residents' needs. Nowadays, ICT developments and especially Internet of Things (IoT) concept and approach allows city officials to interact directly with the community and the city infrastructure by being informed, in real time, for what is happening within the city context or regarding its assets, how the city is evolving, and finally to enable development of new (or enhancement of existing) policies for obtaining a better quality of life.

The aforementioned concept is being realized through the use of real-time systems and sensors, where (a) data are collected from citizens and objects (things), then (b) processed in real-time and finally (c) the gathered information and related extracted knowledge are becoming the keys to tackling inefficiency [1].

In this context, urban/suburban waste management involves numerous waste bins that exhibit significant filling variations (over days and seasons or location) and diverse requirements for emptying, from sporadic (a few times within a week) to very frequent (several times a day). On the other hand other waste forms (i.e. agricultural, biomedical, chemical, electronic, mineral, organic/inorganic, and radioactive, etc.) are characterized by specific collection points, uniform and predictable production, and equal, usually long, filling periods. The detection of the fill-level for urban solid-waste-bins presents many difficulties due to the various irregularities of the waste-bin filling process, such as the irregular shape and the variety of the included materials. More challenges exist for the economical and energy efficient data aggregation from a large number of bins, as the harsh environmental conditions (e.g., humidity, temperature, and dust) can significantly affect the sensor measurement accuracy and reliability, while on the other hand these conditions constitute parameters that one should also take into account for a holistic waste management process. These challenges are being dealt within this work, which is part of the Dynacargo project [2].

## II. RELATED WORK AND SIMILAR PRODUCTS

Currently, several research approaches exist tackling issues related with waste collection and management and especially information collection at the point of waste disposal [3]-[9]. However, most of them are expensive in terms of total required equipment, focusing only on waste management, apart from the smart city vision, while some of them deal with the problem of collecting only recyclable waste, which can be collected at less frequent intervals and have a more stable production frequency. Moreover, a few fill-level sensing prototypes and architectures can be found in the literature [10]-[18].

In some cases, urban waste is separated, according to its source material, to respective waste bins types such as glass; books; textiles; oils or even silos. Currently many commercial solutions [19]-[22] or open-source initiatives [23] have appeared dealing with the fill-level estimation and monitoring of such waste bins containers and the corresponding efficient routing of waste vehicles. However, an integrated solution encompassing smart city vision is not yet available on one hand, while on the other hand all above solutions are based either on existing networking infrastructure that should exist within a city (i.e. WiFi or ZigBee, etc.) or on standalone 3G or GSM networking capability, which in turn is a costly solution for a long-term establishment.

978-1-5090-1314-2/16/$31.00 ©2016 IEEE

Finally, the above research approaches or commercial solutions are not able to tackle with regions that are outside an urban environment [24] (i.e. suburban, isolated or rural areas), where smart city vision is rather difficult to be accomplished within the next few years. Based on the above, the presented integrated node promises to be a versatile and scalable solution that is based on resource-limited embedded devices, able to transfer smart-city vision even outside the context of a smart-city, at a low operational cost and with accuracy as described in the next sections.

## III. Concept and Methodology of the Integrated Node

### A. *Concept of the RFID-based integrated node*

An integrated smart city node should meet some basic requirements regarding its ability to be used in multiple smart city scenarios and provide affordable, reliable and versatile functions including sensor connectivity, energy efficiency and communication. The most important requirements that form the concept of the integrated smart city node based on RFID technology are the following:

- Flexibility of the hardware node. While it is possible to implement a node that performs a single task (such as waste bin monitoring which is the main scenario of this work), an integrated smart city approach requires the ability to implement multiple smart city features. This is due to cost considerations (a single task sometimes cannot justify the investment) but also due to the fact that each smart city application can benefit from information mining performed by other tasks. For example, traffic information from nodes monitoring car traffic can assist the waste collection route optimization combined with waste bin level monitoring. Other smart city applications that can be integrated include environmental monitoring, parking spot monitoring and more. In order to be future-proof and support additional smart city applications, the hardware node should provide a number of inputs of various types (digital, analog etc.) and be able to simultaneously control and transmit information regarding all of them. All these requirements are met with the choice of the active RFID node which provides analog inputs and digital input/output pins with the ability to transmit all sensor information simultaneously within a single beacon.
- Energy considerations. Energy is one of the scarcest resources when deploying free wireless nodes that are not permanently wired to a power or network connection. While it is possible for some applications to permanently wire nodes, many applications such as waste bin monitoring require them to be wireless (i.e. to allow bin collection, bin replacement etc.). As a result, power consumption of the node and sensors should be minimal and managed in an intelligent way so as to prolong battery lifetime. Our RFID-based node provides significant power saving since transmission only takes place at specified intervals and is of minimal power. Moreover, timing functions that are built in the active RFID chip can turn sensors on and off; thus only consuming power when a measurement needs to be taken.
- Communication methods. Various communication options exist that allow city-wide networks to exchange data such as WiFi, GSM, Zigbee and more. Some options (like GSM) make use of existing infrastructure while other, such as Zigbee, require the deployment of specific purpose metropolitan networks. All options exhibit some disadvantages such as high energy consumption, communication provider costs, infrastructure deployment costs, communication range and more. Moreover, such networks should survive connectivity disruptions and should show resilience to the availability and quality of available links. RFID communication does not rely on any communication infrastructure, does not consume a lot of energy and its range needs only be sufficient for the mobile sinks to receive the signal.

The above considerations combined with a desired ubiquitous networking approach led us to the choice of RFID technology combined with Delay Tolerant Networking (DTN) [25] employment for the networking infrastructure of the integrated smart city node.

### B. *Components of the RFID-based integrated node*

The active RFID tag constitutes the basis of the data aggregation and transmission unit for the smart city integrated node. The selected tags (ZT-50-mini Tag from TagSense [26]) can operate with standard 3V voltage input providing increased lifetime that equals to millions of beacon transmissions. The employed active tags offer extra I/O pins for communicating with external devices such as the ultrasonic sensing units used in the waste monitoring scenario. This setup allows powering the sensors and the RFID tag from the same source, either tag's battery or an external power source. The tag itself supports various operating modes, including standard beacons at programmable time intervals, sleep, wake up at regular intervals and also wake up at external trigger.

The proposed system has an adaptable, modular and configurable design that allows optimizing its operation for multiple scenarios including single sensor, multiple sensor and combinations. The basic units for the RFID-based data exchange are the following:

- the *Field Unit*, that is mounted onto an mobile garbage bin and consists of an active RFID tag, the sensors (two ultrasonic sensors in IP-67 enclosure for the waste monitoring scenario), an external battery and the required cabling.
- the *Mobile Sink* consists of the active RFID reader and a Raspberry Pi, consuming less than 10 W of power and running a version of GNU/Linux operating system from a microSD card and the corresponding RFID middleware.

The aforementioned units are presented in detail in the following figure.

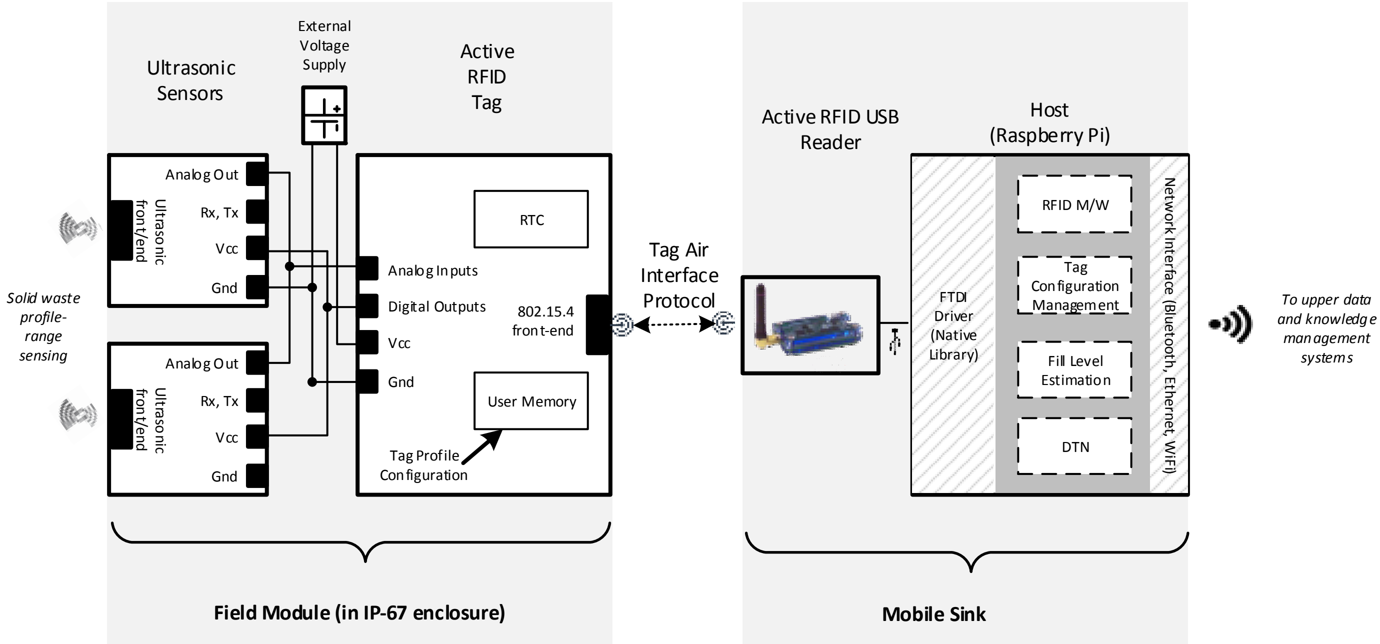



Fig. 1. Details of the RFID-based integrated node.

### *C. Waste-bin fill-level monitoring methodology*

The above presented architectural description of the proposed system indicates that it is easily customizable for the waste-bin fill level monitoring application. Parameters that determine the nature of the scenario can be related to the sensing needs and include its location (urban, suburban or countryside) which correlates to frequency of measurements and possibly the desired accuracy but also the time period (seasonal population changes etc.). Other parameters can be related to the Mobile Sink such as which type of vehicle carries it, its power availability and frequency of measurements. Finally, even more parameters of the adopted scenario are related to design choices of the whole system such as desired lifetime of battery before servicing and cost. Various application scenarios were tested for efficacy and power consumption and the prevalent scenario is as follows.

- The sensors are powered via the RFID tag's digital output pins in order to be able to control when they operate and thus save power.
- The active RFID reader mandates the powering on and off of the sensors when the tags are in range using the minimum of three transmissions. This prevents unnecessary measurements if a reader/tag interaction is not imminent.
- The active RFID tag is programmed to transmit at an interval which will be sufficient for three transmissions to occur while in range with a Mobile Sink with an RFID reader mounted on a moving vehicle.

Bin, waste and sensor modelling, measurement simulation as well as data statistical processing were necessary to be performed in order to extract an accurate fill-level estimation pattern of waste-bins that contain any type of materials and incorporate this pattern in the Mobile Sink. The modelling was based on parameters of the actual physical system. The waste-bin model that was developed is based on the use of a three dimensional matrix representing the entire waste-bin's internal volume. This model was developed using MATLAB because it offers powerful matrix handling tools. Each value of the bin matrix represents the state of a cubic centimeter found inside the bin. By default, the bin matrix is a zero matrix which corresponds to the empty state of the bin. The ultrasonic sensors [27] were also modeled as three dimensional matrices. The matrices used were of the same size as the bin matrix with zero values for the areas of the bin which fall outside the detection range of the sensor and with unit values for the areas inside the detection area, based on the sensor detection diagrams. One of the most crucial parameters for the simulation was waste modelling, mostly concerning its generation. The representation of waste was made using the bin matrix by setting 1s in the elementary volumes where waste was present and 0s otherwise. More challenging was the method to generate waste in a way that would be of equivalence to the actual physical problem. Therefore, extensive random scenarios for bin filling process simulation have been selected, in order to cover all possible combinations. Then, the less likely situations have been filtered out by statistical analysis of the derived results.

The simulation process produced pairs of the actual fill percentage and the sensor measurements using an elementary waste object of base-edge length of 20cm for the area of responsibility of a single sensor. The raw results derived from a sufficient number of simulations (30k-runs) for a variety of models' parameters selection have been further statistically processed. The selected, totally random, method for simulating filling process of a waste-bin followed in the simulations provided different fill-levels for a range of identical sensor values. These fill-levels have been grouped so as to extract the corresponding mean value and deviation for each group, thus treat the bin filling with elementary wastes as a Gaussian process.

## IV. Deployment of the Integrated Node

### A. Physical system deployment

The integrated node has to be deployed at the necessary locations for each application. In the waste monitoring scenario, the waste-bin node is envisaged to be mounted on the top lid of a waste-bin, as depicted in Fig. 2. The bin has a main compartment whose internal dimensions are approximately 110x90x90 cm, consisting of slight curved-shaped walls. It has a lid-opening mechanism and is made of durable hard plastic. The node (Field Unit) consists of the sensing unit, the active RFID tag for data aggregation and transmission, as well as an external battery source. The sensing unit is based on ultrasonic sensors that are advantageous in providing ranging measurements independently of the contained objects, thus making possible the corresponding translation into fill level measurements. For the purpose of this task, the ultrasonic sensors of Maxbotix were selected which provide a variety of detection patterns, accuracy and durability. The measurement of the sensors essentially provides the distance to the nearest object of efficient size. Since waste in the form of irregular-shaped objects can occupy the bin space in various arrangements, there is typically some difference between the highest bin level and the actual fill percentage. Accuracy of the fill-level estimation can be increased by providing more sensing points which will decrease the influence of irregular arrangements as these will only be limited to the sensing area of each sensing point. Following experimentation with placement and number of sensors it was concluded that a single sensor placed in the center of the lid was providing good results but over-estimating the fill level due to the typically higher level of waste in the center of the bin. As a result, various combinations of two sensors were examined and two sensors at specific positions were chosen for the final system.

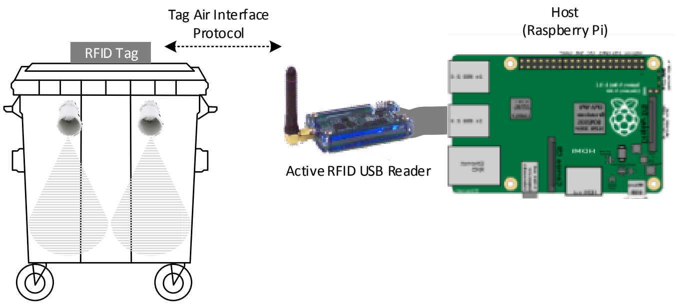



Fig. 2. Deployment of the physical system.

### B. Data flow

First-level data flow through the system regards the sensor measurement transmission through the RFID field unit to the mobile sink that were described in section III. The waste bins are equipped with the Field Unit and the active RFID tags transmit periodically the sensed value among other RFID-related attributes (tag's identity, battery level, RSSI, etc.). This option allows the minimal possible information to be transmitted periodically. All tags are configured to operate in beacon mode, broadcasting the information every few seconds in a range of about 100 meters.

The Mobile Sink could be held by any approved personnel or installed in a vehicle of existing organized transportation systems such as public bus services, postal office vehicles, taxis, municipal police vehicles, etc. in order to unobtrusively collect the required information. As the vehicle roams around the city, the RFID reader reads the tags and collects the information to the sink. Given the short "packet" size (less than 128 bits), the tag range (100 meters), and a realistic average speed of 40 km/h in a city, the tag beacon interval is extracted so as the passing vehicle completes all necessary transactions with the waste-bin. The use of RFID technology allows reading multiple tags simultaneously without collisions and re-transmissions, as it would be the case of a wireless sensor network.

Second-level data flow in the system takes place between the mobile sink and the backend system. For this communication and in order to account for varying availabilities of communication (rural areas, dead zones, temporary loss of communication), DTN technology was used. The DTN2 [30] implementation of the Bundle protocol was deployed in the mobile sink as well the backend in order to provide an effective and reliable way of getting the data to the backend. Figure 3 provides some details on data flow and software components of the mobile-sink.

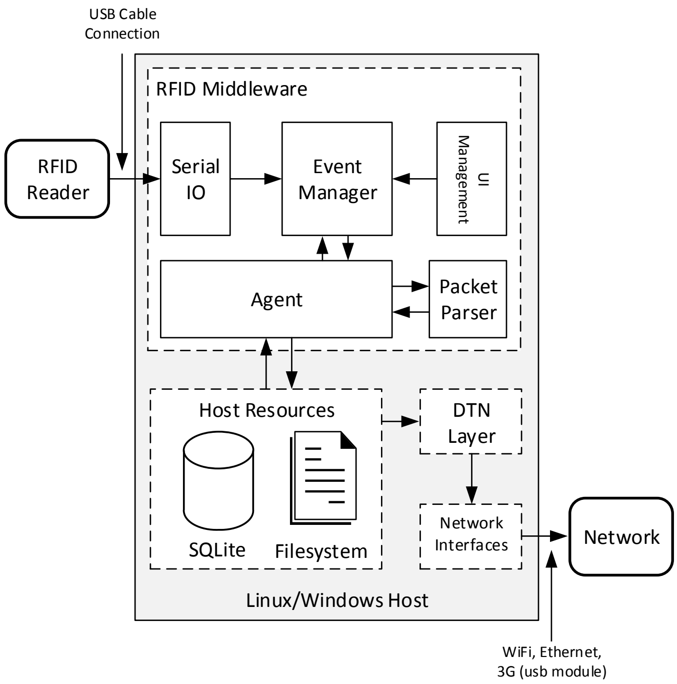



Fig. 3. Software component of the mobile-sink.

### C. Scalability

A significant advantage of the proposed approach is that the system is not limited to a specific type or size of waste-bin but can be scaled regarding any of its components including waste-bin sizes and shapes, numbers of sensors and information data flow. This is due to the fact that the architectural components have discrete roles and functional independence. Scalability was considered in both modelling and simulation as well as physical component selection. The scalability of waste-bin size and shape is made possible based on the design choices that each sensor has its own independent area of responsibility as well because ultrasonic sensors come in multiple versions of beam range and width. The use of versatile programmable active RFID tags, which provides multiple analog inputs and digital I/O's, allows the co-existence of other sensor types that can trigger events or assist in measurement. Even though the active RFID tag can control a number of sensors, in large applications, where more sensors are needed, additional RFID tags can be used for the extra sensors. The RFID tags and readers automatically exchange information whenever they are in range. This means that the addition of new waste-bins to the system or even the addition of extra RFID tags per waste-bin, could be realized without affecting the architecture and not even raising the need for reprogramming the tags or the reader. In such a case the collection of the data would be again realized automatically; however it would be of the related backend software's responsibility to correlate the new additions with their physical interpretations.

## V. Verification and Small-Scale Demonstration

The integrated node, as described in sections III and IV, has been at first deployed in a physical verification at a laboratory environment. The test environment included, apart from the integrated node, a waste-bin of 1,148 liters volume (identical with the one mostly used) and a set of cardboard boxes of known pre-calculated volumes. Figure 4 depicts the volume distribution of the employed experiment items, which have been selected so as to cover a volume range from a few up to some dozens of liters, in order to fit the closest possible with real world's circumstances.

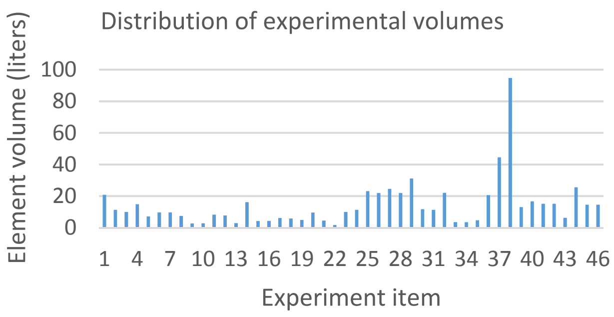



Fig. 4. Volume distribution of employed experimental cardborad boxes.

Subsequently, a large set of experiments has been performed by simulating the real world's waste-bin fill process with dropping a random set of cardboard boxes in the waste-bin. In each round, the volume of the used cardboard boxes was known in advance and also a measurement was acquired using the integrated node.

Figure 5 depicts the per cent error between the two aforementioned volumes for a sample set of 30-experiments. This error is considered to be the maximum possible, since the cardboard boxes haven't been placed within the waste-bin in any specific order, but fully randomly. However the maximum per cent error observed is approximately 12%, which is considered to be significantly low, taking into account that the accurate volume value of a waste-bin's contents does not provide more information to a designated authority than the fill level status approximation does. A few experiments with uniform placement of the cardboard boxes have been also performed in order to verify sensors and measurements accuracy, which concluded to a less than 3% error.

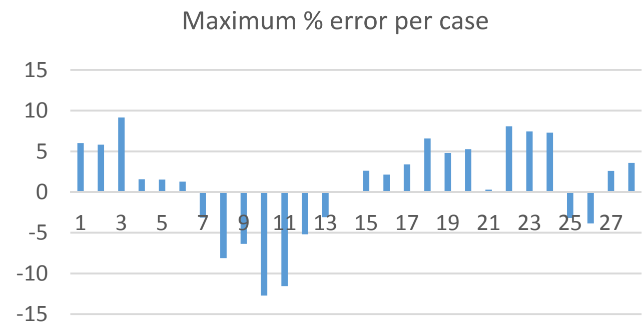



Fig. 5. Physical experiments and corresponding estimation errors.

Finally, a small-scale end-to-end verification of the integrated node has been performed in University of Patras campus, by deploying a simple scenario of three waste-bins and a moving vehicle, as depicted in following figure.

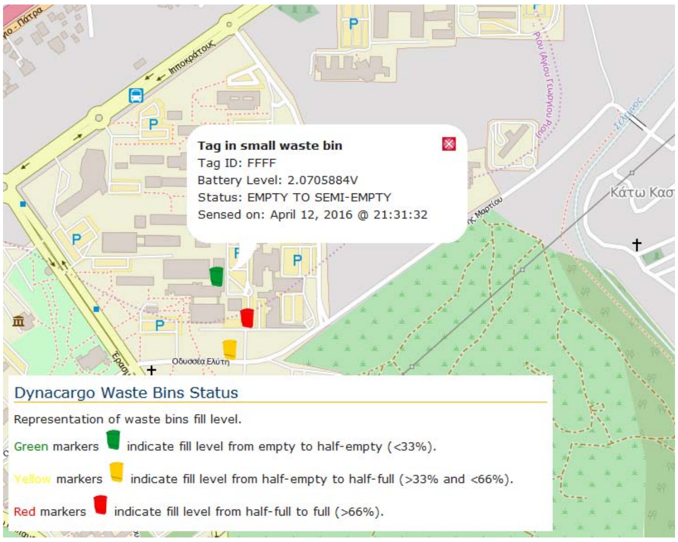



Fig. 6. Waste-bins actual position in a small-scale verification scenario.

The verification scenario consisted of 3-waste-bins, each one equipped with a set of ultrasound sensors and an active RFID tag, where the mobile sink node (the active RFID reader mounted on the Raspberry Pi) was mounted on a moving vehicle speeding up to 60km/h (which is assumed as high speed for loaded waste-trucks). Data from active tags has been successfully received by the mobile sink for various speeds of the moving vehicle. RFID middleware transformed the acquired measurements into JSON structures that were afterwards pushed to a central database via DTN feature, which ensured information transmission at any circumstances, i.e. with or without an actual wireless network connection to be present, thus validating the DTN scenario for isolated rural areas. Finally, DTN was responsible for data uploading to the central database via REST web services, whereas the information stored in the central database is afterwards presented in a simple environment, as depicted in Fig. 6, based on Open Street Map (OSM) [28].

The JSON structure that is pushed in the central database contains raw data from the sensors as well as processed data, i.e. fill level estimation that is produced within the RFID middleware. The following figure depicts the actual notation of the aforementioned structure which is explained in TABLE I.


```
{
    "RID": "FFFF",
    "data": [
        {
            "TID": "0005",
            "BL": "2.95",
            "RSSI": "E7",
            "TS": 1429617039,
            "EFL": "60.21",
            "SNSR": [
                {
                    "ID": "A",
                    "STS": "1429617035",
                    "EHFL": "75.23",
                    "DCM": "15.2"
                },
                {
                    "ID": "B",
                    "STS": "1429617043",
                    "EHFL": "33.44",
                    "DCM": "27.34"
                }
            ]
        }
    ]
}
```


Fig. 7. Sample data snapshot from the demonstration.

TABLE I. JSON STRUCTURE NOTATION

| ***Node*** | ***Description*** |
|---|---|
| RID | Active RFID reader ID (hex value) |
| data | Payload for multiple nodes (array) |
| TID | Active RFID tag ID (hex value) |
| BL | Battery level of active RFID tag (volts) |
| RSSI | RFID communication signal strength (hex value) |
| TS | Timestamp of RFID middleware measurement event |
| EFL | Estimated fill level percentage, according to RFID middleware |
| SNSR | Sensors raw data (array)[a] |
| ID | Sensor identification tag |
| STS | Timestamp pf sensor measurement |
| EHFL | Estimated half fill level (per sensor estimation) |
| DCM | Distance measured from the sensor (cm) |

a. Multiple sensors of other types can also be supported within the same integrated node

Data collected in the central database can be further elaborated (i.e. combined with corresponding historical data) via advanced data processing techniques in order to provide input to other information systems, i.e. schedulers of the waste-trucks timetables, etc.

## VI. CONCLUSIONS

The presented system has been designed and implemented in the framework of the national funded 11SYN_10_456 “Dynacargo” project, promoting an innovative smart city framework for urban solid waste collection. The proposed architecture exploits RFID communication in order to reduce costs and enhance robustness, simplifies system operation and fully supports scaling at urban level while promoting social involvement as it should be in solid waste management. The fill-level estimation method using ultrasonic sensors provided accurate results both in the simulation and the physical experiments. We also described the system design and the implemented prototype of the presented architecture along with experimental measurements and a real-world small-scale verification with very promising results. Future work comprises the extension of this idea to a more granular level, i.e., block of flats, involving all types of waste including recyclable.


## ACKNOWLEDGMENT

This work has been financially supported by the Greek General Secretariat of Research and Technology [29] and European Union under the project “11SYN_10_456: Dynacargo” [2].


## REFERENCES


[1] Sam Musa. "Smart City Roadmap". January 2016. [Online]. Available: www.academia.edu/21181336/Smart_City_Roadmap

[2] "Dynacargo Project." [Online]. Available: dynacargo.cied.teiwest.gr

[3] M. A. Al Mamun, M. A. Hannan, A. Hussain, and H. Basri, "Wireless Sensor Network Prototype for Solid Waste Bin Monitoring with Energy Efficient Sensing Algorithm," in 2013 IEEE 16th International Conference on Computational Science and Engineering, 2013, pp. 382-387.

[4] M. Arebey, M. A. Hannan, H. Basri, R. A. Begum, and H. Abdullah, "Solid waste monitoring system integration based on RFID, GPS and camera," in 2010 International Conference on Intelligent and Advanced Systems, ICIAS 2010, 2010.

[5] M. Faccio, A. Persona, and G. Zanin, "Waste collection multi objective model with real time traceability data," Waste Manag., vol. 31, no. 12, pp. 2391-2405, 2011.

[6] T. Gomes, N. Brito, J. Mendes, J. Cabral, and A. Tavares, "WECO: A wireless platform for monitoring recycling point spots," in Proceedings of the Mediterranean Electrotechnical Conference - MELECON, 2012, pp. 468-472.

[7] O. M. Johansson, "The effect of dynamic scheduling and routing in a solid waste management system," Waste Manag., vol. 26, no. 8, pp. 875-885, 2006.

[8] R. Issac and M. Akshai, "SVASTHA: An effective solid waste management system for Thiruvalla Municipality in Android OS," in c2013 IEEE Global Humanitarian Technology Conference: South Asia Satellite, GHTC-SAS 2013, 2013, pp. 254-259.

[9] P. Reis, R. Pitarma, C. Goncalves, and F. Caetano, "Intelligent system for valorizing solid urban waste," in 2014 9th Iberian Conference on Information Systems and Technologies (CISTI), 2014, pp. 1-4.

[10] F. Vicentini, A. Giusti, A. Rovetta, X. Fan, Q. He, M. Zhu, and B. Liu, "Sensorized waste collection container for content estimation and collection optimization," Waste Manag., vol. 29, no. 5, pp. 1467-1472, 2009.

[11] M. A. Hannan, M. Arebey, R. A. Begum, and H. Basri, "An automated solid waste bin level detection system using a gray level aura matrix," Waste Manag., vol. 32, no. 12, pp. 2229-2238, 2012.

[12] P. O. Leit?o, M. B. Marques, P. Dep, and I. Porto, "Reliability Of A Line-of-view Sensor For Recycling Point Waste Containers," in ISWA/APESB2009 World Congress, 2009.

[13] S. Longhi, D. Marzioni, E. Alidori, G. Di Buo, M. Prist, M. Grisostomi, and M. Pirro, "Solid Waste Management Architecture using Wireless Sensor Network technology," 2012.

[14] V. Catania and D. Ventura, "An Approach for Monitoring and Smart Planning of Urban Solid Waste Management Using Smart-M3 Platform," in 15th Conference of Open Innovations Association FRUCT, 2014, pp. 24-31.

[15] D. Cassaniti, "A Multi-Hop 6LoWPAN Wireless Sensor Network for Waste Management Optimization," 2011.

[16] "EU FP7 Project: 'Future Cities.'" [Online]. Available: www.futurecitiesproject.eu.

[17] "EU FP7 Project: 'Straightsol.'" [Online]. Available: www.straightsol.eu.

[18] F. McLeod, G. Erdogan, T. Cherrett, T. Bektas, N. Davies, D. Shingleton, C. Speed, J. Dickinson, and S. Norgate, "Improving collection efficiency through remote monitoring of charity assets," Waste Manag., vol. 34, no. 2, pp. 273-280, 2014.

[19] "UBi: Ultrasonic sensor." [Online]. Available: www.smartbin.com.

[20] "U-Dump: Urban waste collection sensor." [Online]. Available: www.urbiotica.com.

[21] Libelium Waspmote for Smart Cities. [Online]. Available: www.libelium.com

[22] Enevo. [Online]. Available: www.enevo.com

[23] Open Source Software Society, "The smart garbage bin." [Online]. Available: github.com/ellak-monades-aristeias/Kadoi-aporrimatwn

[24] George Asimakopoulos, Sotiris Christodoulou, Andreas Gizas, Vassilios Triantafillou, Giannis Tzimas, Emmanouil Viennas John Gialelis, Dimitris Karadimas, Andreas Papalambrou, "Towards a Dynamic Waste Collection Management System using Real-time and Forecasted Data", 2nd Workshop on Innovative European Policies and Applied Measures for Developing Smart Cities (IPMSC2015) - 16th International Conference on Engineering Applications of Neural Networks (EANN 2015), Island of Rhodes, Greece, 25-28 September 2015

[25] Delay Tolerant Networking. [Online]. Available: en.wikipedia.org/wiki/Delay-tolerant_networking

[26] "ZT-50-mini Active RFID Tag and ZR-USB Active RFID Reader." [Online]. Available: www.tagsense.com.

[27] "MB1040 LV-MaxSonar-EZ4 Datasheet." [Online]. Available: www.maxbotix.com.

[28] Open Street Map. [Online]. Available: www.openstreetmap.org

[29] "General Secretariat for Research and Technology." [Online]. Available: www.gsrt.gr.

[30] DTN2 implementation, [Online] https://sites.google.com/site/dtnresgroup/home/code/dtn2documentation

*Proceedings of 7th Transport Research Arena TRA 2018, April 16-19, 2018, Vienna, Austria*

# Overview of people localization systems for safe evacuation of large passenger ships

Antonis Kalis a*, Ioannis Panaretou b, Dimitris Karadimas b, Charalampos Kostopoulos b, Renato Campi c, Corinne Kassapoglou-Faist d, Hans Cristian Juul e, Zacharias Siokouros f, Konstantinos Sfakianakis g, Petros Ganos b, Lorenza Alfieri h, Jesus Mediavilla Varas i, Naglaa El Agroudy j, George Georgiades k, Tasos Kounoudes a

[a]*Signal GeneriX LTD, Grigori Afxentiou 23c, 4003 Limassol, Cyprus*
[b] *OPTIONSNET IT & Consulting Services, Kanakari 73-75 Str., 26221 Patras, Greece*
[c]*Canepa & Campi SRL, Via Antonio Gramsci, 4, 16010 Manesseno GE, Italy*
[d]*CSEM Centere Suisse D'electronique et de Microtechnique SA, Rue Jaquet-Droz 1, 2002 Neuchâtel, Switzerland*
[e]*Autronica Fire and Security AS, Bromstadveien 59, 7047 Trondheim, Norway*
[f]*Maritime Institute of Eastern Mediterranean, 150A Franklin & Roosevelt & Omonoias Avenue, 3045 Limassol, Cyprus*
[g]*RCL Cruises Ltd, The Heights Weybridge, KT13 0NY Surrey, United Kingdom*
[h]*Safe Marine SRL, Via Corpo Volont. di Liberazione, 14, 34075 San Canzian d'Isonzo (GO), Italy*
[i]*Lloyd's Register EMEA, Burgess Road, Southampton S016 7QF, United Kingdom*
[j]*Technische Universitaet Dresden, Helmholtzstr. 10, 01069 Dresden, Germany*
[k]*G.G. Dedalos Technology Services LTD, Yiangou Michaelide 26, Nicosia 1048, Cyprus*

**Abstract**

Maritime disasters in recent years are a stark reminder of the imperative need for timely and effective evacuation of large passenger ships during emergency. Tragedies at sea, notably the 2014 South Korean ferry Sewol and the 2012 Costa Concordia incidents, have magnified the urgent need for improvements in the mustering, evacuation, and abandoning procedures, and have led to a series of new global safety initiatives and measures. Driven by this need, several technologies and systems for people localization have been considered, studied and demonstrated over the past decade, which would enable tracking of passengers and crew either on-board, in case of an emergency in order to improve mustering, evacuation, and abandoning procedures, or overboard, after ship abandoning. In this work, we present an overview of these initiatives, evaluating various aspects of such systems, their advantages and disadvantages. Key factors that are considered in our study have been retrieved after extensive analysis of end-user data, acquired over the past three years in the framework of the Lynceus2Market EU funded project (H2020), which has brought together European global players in the field aiming at implementing the first market replication of these technologies and products.

*Keywords:* evacuation procedure; safety localization system; emergency monitoring; decision support.

* Corresponding author. Tel.: +357-258-700-72 fax: +357-258-700-76
*E-mail address: akal@signalgenerix.com*

**Nomenclature**

| | |
|---|---|
| API | Application Programming Interface |
| BLE | Bluetooth Low Energy |
| CCTV | Closed Circuit Television |
| CLIA | Cruise Lines International Association |
| DSS | Decision Support System |
| ECC | European Cruise Council |
| EU | European Union |
| IMO | International Maritime Organization |
| IoT | Internet of Things |
| ISM | Industrial Scientific Medical |
| L2M | Lynceus2Market |
| RF | Radio Frequency |
| RFID | Radio Frequency Identification Device |
| RSSI | Received Signal Strength Indicator |
| RTLS | Real Time Localization System |
| SOLAS | Safety Of Life At Sea |
| SWOT | Strengths Weaknesses Opportunities Threats |
| TRL | Technology Readiness Level |
| UAV | Unmanned Aerial Vehicle |
| UWB | Ultra Wideband |
| WSN | Wireless Sensor Network |

## 1. Introduction

Maritime disasters in recent years are a stark reminder of the imperative need for timely and effective evacuation of large passenger ships during emergency. Tragedies at sea, notably the 2014 South Korean ferry Sewol and the 2012 Costa Concordia incidents, have magnified the urgent need for improvements in the mustering, evacuation, and abandoning procedures, and have led to a series of new global safety initiatives and measures. Shortly after the ship went aground off the Italian coast with the loss of 32 lives, CLIA and ECC launched a Global Cruise Industry Operational Safety Review. One of the issues identified in the review was the difficulty faced by both crew and rescue bodies when trying to locate people during the evacuation and abandonment of a ship; a problem that often adds to the number of fatalities in an incident. As passenger ships grow larger, this only makes the issue ever more urgent.

Driven by this need, several technologies and systems for people localization have been considered, studied and demonstrated over the past decade, which would enable tracking of passengers and crew either on-board in case of an emergency in order to improve mustering, evacuation, and abandoning procedures, or overboard, after ship abandoning. In summary, these technologies can be divided into two main technologies. Those that operate by localizing people directly, through visual or biometric methods, and those that locate people indirectly, through devices that people carry. In the former category, systems include the use of CCTV cameras, or biometric sensors, while in the latter, systems use devices ranging from mobile phones to passive RFIDs, BLE or sub-GHz transceivers.

In this work, we present an overview of these initiatives, evaluating various aspects of such systems, their advantages and disadvantages. Key factors that are considered in our study have been retrieved after extensive analysis of end-user data, acquired over the past three years in the framework of the Lynceus2Market EU funded project (H2020), which has brought together European global players in the field, such as cruise ship owners, operators, ship builders, maritime equipment manufacturers, a classification society, industry associations and important technology organizations with the aim to implement the first market replication of these technologies and products.

Based on this information, we therefore present a comparative study that evaluates diverse system characteristics, while at the same time taking into consideration installation and maintenance costs, and retrofitting abilities. Our goal is to provide an overview of available systems, highlighting both the current technologies and their evaluation criteria, in order to inform on the developments of the basic challenges that have to be considered for building systems for people localization that will create significant impact by saving passenger lives during maritime accidents.

## 2. Review on technologies for large passengers' ships incidents management

Both the size and the capacity of cruise ships have dramatically increased during the last years, creating new concerns in respect to safety and evacuation. Considering the "post Concordia" safety review recommendation, IMO agreed on draft amendments to SOLAS Reg II-2/13 "means of escape" to extend the requirements for evacuation analysis for all passenger ships and not just for the Ro-Ro passenger ships. Such evacuation analysis, able to evaluate escape routes, is required in an early design stage, and should be able to identify and eliminate congestion possibilities during an evacuation and ship abandonment operation. The IMO under SOLAS Chapter III Regulation 21.1.4 requires that all survival craft shall be capable of being launched with their full complement of persons within a period of 30 minutes from the time the abandon ship signal is given. However, for cruise-ships this regulation stipulates that the thirty-minute timeframe only starts when all the passengers have been mustered with lifejackets donned. As one can see, the process of evacuating a large passenger ship is very complex because it involves the management of large numbers of people on a complex moving platform, while people often have very little knowledge about this procedure. Following these developments, several candidate technologies have been considered for the task. This section presents a review on existing technologies for incidents management in large passenger ships, including evacuation procedure, passenger localization, emergency monitoring and decision support. The existing technologies can be divided into two main technology blocks as these are described in the following paragraphs.

### *2.1. User Transparent Localization*

In this paper, we describe as "User Transparent Localization" technologies, those which operate by localizing people directly, without the need for them to carry any devices. This is a significant advantage both in terms of operational costs, and user acceptance. Such systems have already been considered for many on-shore applications, and there are a few examples of maritime use. We have identified the following two main categories of such systems, described in the following paragraphs:

#### *2.1.1. Cameras*

The main concept here is the use of visual sensors, like low-cost CCTV cameras to cover the whole area where people can be located. This is not a new idea, since it has been used for decades in many surveillance systems (Black & Ellis (2006), Black & Ellis et al. (2002)). Recent improvement in this system concept is that people localization is done automatically, using dedicated image processing software, capable of identifying individuals in an area. The technologies behind these systems is described in several published works. For a comprehensive overview, in J.R. Martínez-de Dios et al. (2011) the authors describe various sensor fusion approaches for detection, localization and tracking of mobile objects using camera-based WSN. Most of the work in this field is focusing towards reducing the cost of required cameras, and on the type of image processing algorithms that would enable robust and precise localization of multiple targets.

However, regardless of the technological advancements and the advantages that this type of approach presents for people localization onboard a ship, there are quite a few unresolved issues involved. The less significant of them are relevant to the current state of this technology. It could be argued for example that the reported state of the art (Black & Ellis (2006), Black & Ellis at al. (2002)) can cope with small numbers or small densities of people, which is a scenario irrelevant to large passenger vessels evacuation. Moreover, visual capturing would be challenged when considering a wide variety of emergency scenarios, where low visibility would be an issue due to smoke, fire, or other causes. However, the technical barriers could be considered as secondary, since larger densities of cameras and larger servers could cope with these issues in the future. The primary issues involved with this type of solutions are related to privacy, and user acceptance. The former is a significant issue since there are large areas onboard that cannot be covered due to privacy issues, user acceptance and current legislation. It would not be accepted for example to install cameras in cabins, restrooms, certain recreation areas, or even certain charter cruises.

*2.1.2. Human Activity Sensors*

This is a relatively new and promising field of research, where new sensing devices can detect the presence of human activity. Ranging from low-cost, low-complexity motion sensors, to smart floors and respiration sensors, these sensing devices seem to have the answer to resolving both privacy issues and localization transparency. However, current solutions still lack in maturity. Reported motion sensing devices using infrared or ultrasonic sensors (L. Wang et al. (2010), P. Wojtczuk (2011)) can detect presence, but are rather unsuitable for localizing and identifying large number of people on-board, both in small and large open areas. More promising technologies include UWB respiration sensors (Yuan Liu et al. (2015)) and smart floors (Contigiani et al. (2014)) which can detect presence using low-cost force sensors. However, none of the aforementioned technologies is yet robust enough nor economically viable for deployment in large passenger vessels.

*2.2. User Assisted Localization*

With this term, we describe systems that do not detect people directly, but rather localize devices that we expect people to carry, using radio frequency devices. Although this is clear disadvantage with respect to the previous direct localization methods, user assisted localization is the norm for current onboard localization systems, due to technology maturity level and cost efficiency. Moreover, such systems can be used for onboard as well as for overboard localization, which makes them more appealing for maritime applications. Their development is based on the emergence of several innovative technologies like wireless sensor networks, sensor and actuator systems, intelligent data processing and fusion techniques, innovative decision support processes, low powered microelectronics and digital signal processing algorithms. Recently, a number of such systems have been developed for the task, which are usually evaluated against the following metrics:

- Cost: Both the cost of infrastructure and of each device is critical, due to the fact that in large passenger vessels, very large areas have to be covered, and very large numbers of people have to be equipped with wearable devices.
- Accuracy: The accuracy requirements of a localization system depend on the application needs. For some areas (e.g. cabins, rooms) knowing the number of people present would be sufficient. In open areas and open decks on the other hand, an accuracy in the order of a few meters would be required.
- Bandwidth: The available bandwidth for RF localization largely dictates the accuracy of the system. However, bandwidth in wireless systems is a scarce resource, which must be contended for by both localization and communication systems in the area.

User assisted localization systems are often faced with the need to make trade-offs between these three main characteristics in order to satisfy the requirements of a given application. In the following paragraphs, the most popular technologies are reviewed.

*2.2.1. Passive RFID*

Passive RFID devices are devices that have no internal power supply and rely solely on RF energy transferred from the reader to the tag. Although this method is highly cost effective, since the cost of each tag can be reduced significantly, the bandwidth and range of these passive devices is significantly lower than their active counterparts examined below. This has a significant impact on their use and applicability in localization scenarios onboard large passenger vessels. Since the distance from the reader is limited from a few centimetres to a meter, the density of high-cost readers should be significant if we were to achieve acceptable accuracies. Moreover, in the presence of high densities of people, RFID tags have been reported to present lower robustness. Therefore, the use of passive RFIDs is usually confined to act as complementary to other technologies (e.g. Ekahau's - http://www.ekahau.com) or for other onboard uses like giving access to cabins or performing contactless payments onboard.

*2.2.2. BLE*

BLE technology is a version of the popular Bluetooth technology, customized for very low power consumption. It is an active technology, requiring internal power supply for its tags, and as such, has a higher cost than any passive RFID device. It operates at the 2.4GHz ISM band, which is can be used globally, and provides a significant bandwidth for localization. However, this band is prone to high external as well as co-channel interference, since it is already widely used by many popular systems like WiFi, other Bluetooth devices, Zigbee, etc.

#### *2.2.3. Sub-GHz Radio*

This technology uses a number of RF bands below 1GHz, for example the 433MHz ISM band, or the 868/916MHz bands. As is the case with BLE, sub-GHz systems use active tags. Moreover, the frequency bands used have lower bandwidth, and also face different restrictions of use in different parts of the world. Their main advantage is that they are not highly populated by other systems, but are rather dedicated to IoT applications, as the one considered in this work.

#### *2.2.4. Comparison of Different technologies*

The following table summarizes some of the main features of the aforementioned technologies.

Table 1. Comparison of different tag based solutions

| Competitive Systems | Passive RFID | BLE | Sub-GHz | Comments |
|---|---|---|---|---|
| Frequency | RFID | 2.4GHz | 433MHz, 868/916MHz | |
| Battery Lifetime | N/A | Up to 2 years | Up to 2 years | Depending on use |
| Range | Up to 1m | Up to 5m | Up to 10m | Indoors |
| Tag Cost | Very low (<1$) | Low (<10$) | Low (<10$) | |
| Reader Cost | Very high (500 – 2000 $) | Medium (<200 $) | Medium (<200) | |

Although a thorough evaluation of each technology largely depends on the application, we are able to give some general pointers regarding the applicability of each technology to large passenger vessel evacuation. The passive RFID provides a low-cost tag solution. However, its small range, and the high cost of the readers limits its use to mustering and boarding applications, as this will be shown in the next section. On the other hand, its small range does not make it a suitable candidate for overboard localization applications, compared to active RFID solutions. The latter, namely RFID and sub-GHz solutions have comparable performances in terms of battery lifetime, tag and reader cost. The cost of installation however could be a bit higher for BLE systems, due to their smaller range. The main difference of these systems lies in the use of frequency bands. This has the following consequences. BLE's use of the 2.4GHz makes it prone to interference, particularly in crowded areas, where a large number of devices using the same frequency band is present. On the other hand, sub-GHz systems should be capable of switching frequencies from 868MHz to 916MHz and back, depending on the geographical location of the ship.

## 3. Review on systems for large passengers' ships incidents management

In this section, we present an analysis of existing solutions for onboard and overboard localization, that use any of the technologies presented in chapter two, as well as the L2M equivalent sub-systems. The competitor products shortly presented in the following paragraphs are organized into two main categories, i.e. i) On-board Localization and Evacuation Support systems, ii) Overboard localization systems and electronics.

### *3.1. On-board Localization & Evacuation Support Systems*

#### *3.1.1. AiRISTA Flow - Ekahau Real-Time Location System*

AiRISTA Flow enables a Real-Time Locating System (RTLS), using RFID tags. It offers three types of RFID Wi-Fi tags: passive, active, and semi-passive. AiRISTA Flow gives customers the option of RFID, Bluetooth Low Energy (BLE), GPS, or IR technology. Currently it offers two RTLS platforms. Based on the hardware components and applications, AiRISTA Flow Universal Visibility Software (UVS) or Ekahau Vision software provides an intelligent view into the enterprise for all of the RTLS applications in place. With either AiRISTA Flow UVS or Ekahau Vision the executives, managers, and end users get visibility into the location, condition, and status of assets, people, and workflows plus reporting that can be delivered automatically.

*3.1.2. Extronics - AeroScout MobileView Software*

AeroScout MobileView provides a simple, yet powerful means of tracking the location, status and condition of assets and people. The web-based software platform has a full range of applications with visualization, reporting, management and automated alerting options, as well as the ability to deliver context-aware visibility data to a variety of third-party applications. MobileView unifies asset visibility data from many sources, such as Extronics Advance‘s Wi-Fi Tags, Wi-Fi clients, Passive RFID, GPS, third-party devices and sensors, and turns this data into automated business processes – delivering a scalable, enterprise-proven software solution.

*3.1.3. Sea Hawk Surveillance – Real-time safety solution*

The mission of Sea Hawk Surveillance is to provide a real-time safety solution to monitor passengers and the cruise lines personnel with a reusable waterproof RFID wristband device. This wristband device contains state-of-the-art technology that provides a 24/7 monitoring of passengers. It also identifies all persons in the vicinity of any incident and stores all historical information during the cruise. The wristband is an intelligently designed wristband for localization of personnel using RFID technology. The innovative micrometre integrated circuits allow for the wristband's compact, convenient size and ultra-low power consumption. This wristband is integrated within soft, durable material and a rigid waterproof housing. If the wristband is broken or opened without approval or permission, it immediately will send an alarm signal.

*3.1.4. AutoCrew - Automated Crew Management System*

AutoCrew or “Automated Crew Management System (ACMS)” is an application designed to fulfil a variety of personnel localization and reporting functions within a variety of industrial work sites, and possibly multiple sites working together. The AutoCrew is an IT based solution that provides multiple management stations and unlimited number of readers of RFID tags for real time situational awareness and readiness assessment. The AutoCrew solution helps mustering and boarding, by scanning the tags of crew and passengers at predefined checkpoints.

*3.1.5. Nanotron technologies*

Nanotron Technologies provides high-throughput real time location systems with wireless transmission of vital data based on WSN technology. These systems are built around their embedded location platform based on chirp spread spectrum modulation, which they have patented for commercial wireless data transmission. The company proposes collaborative localization (relative positioning of the nodes) using time-of-flight (TOF) ranging and a patented technique to avoid node synchronization and fixed localization (localization with respect to fixed points or anchors) using time difference of arrival (TDOA). The latter requires time synchronization of the anchors. This solution may have higher accuracy. However, it has higher energy consumption and uses the already crowded 2.4 GHz frequency band.

*3.1.6. SIREVA project*

The focus of the SIREVA project is on optimizing the workflow and communication of the emergency on-board organization by providing technical support for the crew member's key functions such as counting and identifying passengers at their assembly stations. The system developed by the SIREVA project is based on localizing wearable Bluetooth Low Energy (BLE) tags within all areas that are equipped with the respective transceivers. Depending on the technical and financial efforts made, the accuracy for locating people varies between an area of e.g. a vertical fire zone on a certain deck and less than one meter. Variable distances between transceivers are used to reflect different accuracies needed: the borders of a muster station must be detected very precisely, while an accuracy of a few meters is enough to ensure that e.g. public spaces are completely evacuated.

*3.1.7. Lynceus2Market On-board Localization System*

The Lynceus2Market project addresses the challenge of localizing people onboard a large passenger vessel as well as overboard in case of an emergency, through delivering a revolutionary operational system for safe evacuation based on the integration of several innovative people localization technologies and evacuation and safety related systems. The Lynceus2Market on Board system consists of:

- Localizable life jackets that can provide passenger location in real-time during emergency

- A network of gateways that act as the backbone of the localization system, and can also accommodate several fire and safety sensors.
- Innovative localizable bracelets able to send activity data to the emergency management team
- Low cost fire and flooding escalation monitoring sensor nodes
- Novel mustering handheld devices for automatic identification and counting of passengers during evacuation
- Smart localizable cabin key cards
- Intelligent decision support software able to fuse all localization, activity and disaster escalation data to provide an integrated real-time visualization, passenger counting and evacuation decision support

Lynceus2Market bases its on-board localisation on low-cost, ultra-low-power wireless sensor network technology, using signal strength as a perception of distance (RSSI). Measurement of RSSI is inherent to RF communication (a built-in function in the node hardware) and therefore an attractive method. An intelligent decision support software is able to fuse all localization behavioural and disaster escalation data so as to provide real-time visualization, passenger identification and evacuation decision support. It provides real-time information regarding people and crew location, disaster escalation, passenger behavioural information and mustering information.

*3.2. Overboard localization systems and electronics*

Over board localization systems are responsible for localizing people at sea. All these systems rely on the assumption that people overboard are wearing lifejackets equipped with electronic equipment used for radio localization. The latter can be either life-jacket based, or system based, and utilize one of the following technologies:

a. GPS reception. GPS is excellent for providing a localization accuracy of 10 meters or less at open spaces, where view to GPS satellites is unobstructed. This information is available at the GPS receiver on the lifejacket, but is not available at system level. Therefore, GPS based solutions need to be complemented with a communication system to transmit their calculated position to the search and rescue agency. The cost of this combination is usually high, and therefore GPS based systems target the high end of the market.
b. Radio Beacon transmission. In this setup, life-jackets are equipped with electronic equipment that can either send out a radio beacon, or can respond to radio signals transmitted from search and rescue agencies. This solution is more cost effective than the GPS solution, but is limited in range by the range of the radio equipment.

In the following paragraphs, we present some key overboard localization systems.

*3.2.1. AiRISTA outdoor location tracking*

AiRISTA uses the same RSSI algorithmic engine together with GPS for overboard localization. The RSSI Engine would require a Wi-Fi infrastructure, while GPS reading would require a communication link to transmit tag readings to a central search and rescue control center.

*3.2.2. Extronics - Outdoor Equipment Tracking*

Extronics aims to optimize outdoor workflows with greater visibility of open-air work areas. Its solution incorporates an active RFID tag either affixed to a user, while the use of a battery suggests that this is an active RFID system. The active RFID tags send location data in regular, short bursts over an infrastructure network, which means that localization functions are executed on the tags, and would therefore need a number of anchor points for localization and for communicating their information to the backbone.

*3.2.3. Sea Hawk Surveillance - Overboard*

The product locates personnel using RFID technology through designed wristband. Sea Hawk Surveillance monitoring program receives the data from the wristband via satellite indicating that a passenger is overboard. The ships control bridge is also immediately notified. Sea Hawk NOC immediately starts localizing the person in the water and relaying the information to the ship's control. The ship's bridge is also receiving the data from the satellite.

### *3.2.4. Deutsche Telekom M2M (machine-to-machine) Localization*

The product aims to enable machines, sensors and vehicles to communicate with each other or with a central data processing platform. People and objects are localized by GPS and locations are relayed via the mobile network to a control center. Small GPS trackers that fit into a handbag or a jacket pocket relay the wearer's exact position on demand to a control center.

### *3.2.5. Lynceus2Market Overboard Localization System*

The L2M overboard Localisation system consists of two main components; An innovative shore or ship-launched Unmanned Aerial Vehicle for localizing people in the sea in short time and assisting search and rescue operations and a low-cost rescue-boat mounted radar for people localization in the vicinity of the boat. The UAV carries GPS, radio-location, processing and communication modules. Life-jackets are equipped with localizable active reflector tags. The localisation algorithm takes full advantage of two diverse frequency bands in order to achieve both large range and high accuracy.
In the next paragraph the results of a comparative analysis performed in the context of the L2M project, indicating the advantages of the L2M onboard and overboard systems compared to their competitors.

## 4. Comparison analysis of people localization systems

This section presents a comparative analysis of the existing competitive products compared to the Lynceus2Market on-board and overboard systems Table 2 summarizes the major characteristics of the solutions presented in terms of the employed Onboard/Overboard technology and the capability of integration with native or third-party emergency monitoring systems.

Table 2. Comparative analysis of existing solutions for localization in large passengers ships

| Competitive Systems | Onboard Technology | Overboard Technology | Emergency Monitoring System Integration | Extra features |
|---|---|---|---|---|
| AiRISTA Flow | RFID, BLE, GPS | GPS | Yes | Wide range of accompanying applications |
| AutoCrew | RFID | n/a | No | n/a |
| Deutsche Telekom M2M | n/a | GPS | No | n/a |
| Extronics | RFID, WiFi | GPS | No | Wide range of accompanying applications |
| Lynceus2Market | RFID/sub-GHz | Sub-GHz | Yes | Passengers health status information |
| Nanotron | 2.4GHz active RFID | n/a | No | No incident management |
| Sea Hawk Surveillance | Active RFID | Active RFID | No | n/a |
| SIREVA Project | BLE | n/a | No | No incident management |

## 5. Conclusion

In this paper, we presented a comparative study that evaluates several technological solutions and existing systems for localizing people both onboard large passengers' vessels and overboard. Our goal was to provide an overview of available systems, highlighting both the current technologies and their individual characteristics, to inform on the developments of the basic challenges that have to be considered for building systems for people localization that will create significant impact by saving passenger lives during maritime accidents.
Our study has shown that if we consider the technological and privacy barriers as well as the critical metrics that influence user acceptance (e.g, frequency, lifetime, range and cost) as well as the capability to perform well both onboard and overboard, the Lynceus2Market provides an excellent solution with respect to other existing systems, effectively overcoming the basic challenges that have to be considered for building systems for people localization that will create significant impact by saving passenger lives during maritime accidents.

**Acknowledgements**

Work presented in this paper is co-funded by the European Union under Horizon 2020 programme framework, Project H2020-MG-4.2-2014: LYNCEUS2MARKET, “An innovative people localisation system for safe evacuation of large passenger ships”.

## 6. References

AiRISTA Flow – Ekahau Real-Time Location System. Online: https://www.airistaflow.com

AutoCrew – Automated Crew Management System. Online: http://www.marinerescuetechnologies.com

Black, J. & Ellis, T. (2006). Multi Camera Image Tracking. Image and Vision Computing, Vol.24, No.11, (November 2006), pp. 1256-1267, ISSN 0262-8856

Contigiani, Marco & Frontoni, Emanuele & Mancini, Adriano & Gatto, Andrea. (2014). Indoor people localization and tracking using an energy harvesting smart floor. 1-5. 10.1109/MESA.2014.6935623.

Extronics – AeroScout MobileView Software. Online: http://www.extronics.com

J. Black, T. Ellis, and P. Rosin. Multi view image surveillance and tracking. Proceedings of Workshop on Motion and Video Computing, 2002.

J.R. Martínez-de Dios, A. Jiménez-González and A. Ollero, Localization and Tracking Using Camera-Based Wireless Sensor Networks, Sensor Fusion - Foundation and Applications, Openscience, 2011, pp 41-60

Kongsberg Information & Safety Systems. Online: http://www.km.kongsberg.com

L.Wang, W.Gong, L.He, H.Xiao and Y.Huang, “Human motion recognition using pyroelectric infrared signal”, Journal of Optoelectronics. Laser, Vol. 21, No. 3, 2010, pp. 440-443.

Martec – Safety Monitoring and Control System. Online: http://www.martec.it

P.Wojtczuk, A.Armitage, T.D.Binnie and T.Chamberlain, “PIR sensor array for hand motion recognition”, SENSORDEVICES 2011, The Second International Conference on Sensor Device Technologies and Applications, Nice/Saint Laurent du Var, France, 2011,pp. 99-102.

Project Lynceus2Market. Online: http://www.lynceus-project.eu

Sea Hawk Surveillance – Real-time safety solution. Online: http://seahawksurveillance.com

Yuan Liu, Shiyou Wu, Jie Chen, Guangyou Fang, and Hejun Yin, “Human Respiration Localization Method Using UWB Linear Antenna Array,” Journal of Sensors, vol. 2015, Article ID 601926, 11 pages, 2015. doi:10.1155/2015/601926

# An Integrated Low Cost IoT Node based on Discrete Components for Customized Smart Applications; Use case on Precision Agriculture

John Gialelis
Electrical and Computer Engineering Department, University of Patras, Greece
gialelis@ece.upatras.gr

Gerasimos Theodorou
Electrical and Computer Engineering Department, University of Patras, Greece
makosemail@gmail.com

Maria Fokaeos
Electrical and Computer Engineering Department, University of Patras, Greece
fokaeosm@gmail.com

Dimitrios Karadimas
Electrical and Computer Engineering Department, University of Patras, Patras, Greece
karadimas@ieee.org

***Abstract*— This work depicts the concept and methodology as well as the architecture and physical implementation of a low-cost Internet of Things (IoT) node for customized smart-applications. The presented node efficiently integrates state-of-the-art discrete electronic components able to support a variety of smart applications. The node comprises an ARM CortexM based microcontroller designed for low power wireless applications, a power management unit and several sensors for which special adapters have been implemented. The proposed IoT node is being validated in vineyard fields, in the framework of precision agriculture practice, with the aim of collecting critical environmental parameters, to be used as input to algorithmic models for the detection and subsequently prevention of related agricultural diseases.**

***Keywords*— *internet of things, disrete electronic components, Customized Smart Applications***

## I. Introduction

The Internet started as a global network with services but through its revolution over time has become an unprecedented ecosystem with interconnected objects / devices characterized by advanced sensing and actuating capabilities, distributed processing power and storage capacity whilst their size is getting smaller and smaller. All these technology advances are expanding the Internet's boundaries as well as changing the nature and the opportunities offered by the devices. This network of interconnected devices not only collects information from the environment and interacts with the physical world, but it also uses existing Internet standards to provide diverse services regarding information and communication. The use of intelligently connected devices and systems to leverage data gathered by embedded sensors and actuators define the "Internet of Things" (IoT). For the consumers, IoT has the potential to deliver solutions that dramatically improve energy consumption, security, health, ageing and many other aspects of daily life while for the enterprises IoT can support solutions that improve decision-making and productivity in manufacturing, in retail, in agriculture and other sectors.

In this work, we introduce a low-cost Internet of Things (IoT) node ideal for smart-application domains, i.e, smart cities, smart hospitals, smart agriculture, etc. The presented node is built around a state-of-the-art microcontroller which is based on a low power Cortex™ M3 processor from ARM®, a high-performance RF and mixed-signal system, a high speed 12-bit ADC, nonvolatile Flash/EE memory and SRAM. A power supply system based on an integrated boost regulator that converts dc power from a PV cell charges a Lithium-ion storage element for unceasing operation. Moreover, the proposed node provides I2C, SPI and UART connectors for connecting various sensors and Serial Wire Debug (SWD) for system debugging or firmware download. Furthermore, the operation of the proposed IoT node is demonstrated in the precision agriculture by aggregating customized environmental data aiming to feed algorithmic models for the detection as well as the prevention of vineyard fields diseases.

The rest of the paper is organized as follows. Section II briefly describes the background regarding open IoT platforms / protocols and related work about their utilization. Section III depicts our approach and decomposes our proposed system in its components while Section IV reveals a demonstration of the proposed system. Section V summarizes the progress made so far along with the future steps to follow.

## II. Related Work and Similar Products

### *A. Open Hardware Platforms*

The open source initiative is capable of supporting the free use of open source-licenses. It is through this open source movement that the popularity of open-hardware has been triggered and consequently made available with different options for open-hardware microcontrollers-based platforms. The question of which platform would be ideal to use, depends on the requirements of the system. These requirements comprise power consumption, cost and type of connectivity among others. A description of some of the most notable open-hardware platforms and their features follows:

- Arduino Genuino 101: The module contains two tiny cores, an x86 (Quark) and a 32-bit ARC architecture core, both clocked at 32MHz. The Intel toolchain compiles Arduino sketches optimally across both cores to accomplish the most demanding tasks. The Real-Time Operating Systems (RTOS) and framework developed by Intel is open sourced. The 101 comes with 14 digital I/O pins (of which 4 can be used as PWM outputs), 6 analog inputs, a USB connector for serial communication and sketch upload, a power jack,

978-1-7281-1740-9/19/$31.00 ©2019 IEEE

an ICSP header with SPI signals and I2C dedicated pins. The board operating voltage and I/O is 3.3V but all pins are protected against 5V overvoltage. Finally, the 101 can be programmed with the Arduino Software (IDE) and the board is preprogrammed with the RTOS, which handles USB connection and allows to upload new code without using an external hardware programmer [1].

- LinkIt ONE: The hardware core for LinkIt ONE development platform is MediaTek MT2502A and its processor core is 32-bit ARM7EJ-S, clocked at 260 MHz. The MediaTek MT2502 is a feature-rich and extremely powerful single-chip solution for high-end GSM/GPRS (2G) capabilities. Furthermore, it comes with pin-out similar to Arduino UNO, including digital and analog I/O, PWM, I2C, SPI and UART. It also features a highly integrated Bluetooth transceiver which is fully compliant with Bluetooth specification v4.0. This chipset also works with Wi-Fi (MT5931) and GNSS (MT3332) chips, offering high performance and low power consumption for Wearable and IoT devices [2].
- ESP8266EX: The ESP8266EX microcontroller integrates a Tensilica L106 32-bit RISC processor, which achieves extra-low power consumption and reaches a maximum clock speed of 160 MHz. The Real-Time Operating System (RTOS) and Wi-Fi stack allow about 80% of the processing power to be available for user application programming and development. ESP8266EX achieves low power consumption with a combination of several proprietary technologies. The power-saving architecture features three modes of operation: active mode, sleep mode and deep sleep mode. This allows battery-powered designs to run longer [3].
- Seeed Studio BeagleBone Green: BeagleBone Green is a low-power single-board computer with an AM3358 ARM Cortex-A8 processor running at 1 GHz and 512 MB of DDR3 RAM. It features also a 3D graphics accelerator, a NEON floating-point accelerator and 2 PRU 32-bit microcontrollers and GPIO comes with 65 digital and 7 analog I/O. As it concerns connectivity, it includes an USB client for power and communications, an USB host, Ethernet, 2-46 pin headers and I3C and UART as Grove connectors. BeagleBone has also a power-saving mode, in a sense that it provides the ability to let processor put board into a sleep mode to save power. Unlike RTOS, BeagleBone Green run on Linux and Android, that is necessary for applications of higher end IoT nodes [4].

### *B. IoT Protocols*

The IoT system can function and transfer information only when the devices are safely connected wired or wireless to a communication network. This connection is established by choosing the appropriate type IoT protocol based on our system's specific requirements, like the bandwidth and power consumption [5]. The two major M2M (Machine to Machine)/IoT protocols are defined below:

- Constraint Application Protocol (CoAP) is a specialized web transfer protocol created for use within constrained nodes, like low power sensors and actuators, and constrained networks in IoT. It enables those nodes to communicate with other constrained nodes over the Internet. More specifically, the CoAP runs on devices that support UDP protocol, in which client and server communicate through connectionless datagrams. Due to its compatibility with the HTTP, its strongest potential use would be among other protocols. Finally, in a similar way of the HTTP, the CoAP has the ability to carry different types of payloads and to integrate any data format of the user's choice [6].
- Message Queuing Telemetry Transport (MQTT) is one of the most commonly used protocols in the IoT projects and like other internet protocols, it is based on the clients and a server. Contrary to others, the MQTT makes it easier to implement a software, transmit data at quicker pace, but it also minimizes the data packets and consumes less energy. This protocol functions on top of the TCP/IP and it includes a subscriber, a publisher and a broker. The publisher collects the data and sends it to subscribers. Consequently, the broker tests publishers and subscribers, check their authorization. It's through the real time interaction between the devices on a wide area network that makes the MQTT an ideal protocol for IoT applications [7].

Both the MQTT and the CoAP are designed as a long-term vision that will enable them to get used in a lightweight environment. Both work well with low power and network constrained devices. Thus, the choice depends on the application use case. If an M2M network needs to be created and the messages would need to be sent from one node to other multiple interested nodes, then the best choice would be to use an MQTT protocol. If on the other hand, an M2M network has to be created, in which commands need to be circulated among nodes, then the CoAP will be most appropriate choice.

Our proposed IoT node is a microcontroller-based device (MCU) following a modular lean design, both at the hardware components level and at the behavioral level. In this way it is intended to achieve, within a limited budget, a transversal robust structure able to satisfy as many as possible aggregation and transfer requirements thus practically accommodating many smart application use cases. Based on our system's specific requirements, such as bandwidth and power consumption, MQTT protocol is adopted.

## III. IoT Node Architecture

Our proposed IoT node architecture proposal follows a modular design both at hardware level (physical constituent modules) as well as the behavioral level (operational conceptual modules). This section is divided into three subsections. The first subsection presents how the IoT node is implemented at the hardware level, the second subsection reveals the firmware configuration and how the IoT node operates at the software level, while the third subsection details the system performance and the energy consumption aspects.

### *A. Physical Constituent Modules*

At the hardware level, the physical constituents are grouped in three different function blocks as described below:

- Microcontroller and Radio Function Block

The proposed node is utilizing Analog Devices ADuCRF101 microcontroller. The ADuCRF101 is a fully integrated single chip data acquisition solution designed for low power wireless applications. It features a 12-bit ADC, a low power Cortex™ M3 processor from ARM®, a 431 MHz to 464 MHz and 862 MHz to 928 MHz RF transceiver (ADF702x radio on chip), and 128KBFlash/EE memory packaged in a 9 mm × 9 mm LFCSP.

- Power Supply Function Block

The power supply system is built around Analog Devices ADP5090 chip. The ADP5090 is an integrated boost regulator that converts DC power from PV cells or TEGs. The device charges storage elements such as rechargeable Li-Ion batteries, thin-film batteries, super capacitors, and conventional capacitors, and powers up small electronic devices and battery-free systems. The ADP5090 provides efficient conversion of the harvested limited power from 16 µW to 200 mW. A 4x4cm 2V solar panel has been used with 60mA peak output. The photovoltaic panel charges a 3,7V - 650mAh Li-ion battery. The output from the solar panel and the battery drives a ADP190 (linear voltage regulator) which finally provides a 3.3V stable output.

- Input / Output Function Block

The node provides connectors for various sensors, such as I2C interface, SPI, UART, Interrupt pins, A / D interface and SWD for system debugging or firmware download.

- Sensors and corresponding adapters

Through the use of the appropriate interface a wide range of sensors can be connected to the IoT node. Currently, there is an extensive list of sensors for which the corresponding adapters have already been implemented as shown in Table 1. More adapters for additional sensors are envisaged to be developed eventually.

*Table 1List of Integrated Sensors*

| Sensor | Scope | Interface | Product |
|---|---|---|---|
| 12-Bit Digital Temp. (±2°C accuracy) | Node's Battery Temp. | I2C | ADT75 |
| Humidity & Temp. | Ambient H&T | I2C | SHT21 |
| Humidity & Temp. & Pressure | Ambient H&T&P | I2C | BME280 |
| Rain Collector | Rain amount | Interrupt | Davis6463 |
| Wetness | Leaf Dew & Precipitation | A/D | Davis6420 |
| Moisture | Soil Moisture | A/D | Davis6440 |

## B. Operational Conceptual Modules

Analog Devices offers the RapidNet IP and the AD6LoWPAN stack. The RapidNet IP is a star or ex-star topology self-healing wireless communication protocol stack focused on providing a high level of scalability, extreme ease of use, and small code footprint. The AD6LoWPAN is an alternative mesh topology Analog Device's 6LoWPAN solution. Even though both stacks easily could be placed on our node, a more simplistic and very stable stack was built with the aim of minimizing energy consumption. In a typical configuration there are many broadcast node stations (end nodes) and a single center node station. The topology of the network is star or extended star since a node could be used as a repeater. No mesh logic is used. The end node is in a hibernate state. Every 10 minutes wakes up, collects the measurements from the sensors and emits them at a pre-agreed speed and frequency channel without storing them locally. Each end node has a unique 16bit ID. The end node does not expect any acknowledgement from the receiving node (center node) assuming that the measurements have reached the destination. The end node, with no hardware changes, is also used as a center node but with the appropriate firmware installed. The center node is continuously in reception mode at the pre-agreed speed and frequency channel. When the center node receives a packet of measurements with the correct CRC it forwards it directly to its serial port which is connected either with a PC or any other device with a full TCP / IP stack. i.e., Raspberry Pi. Eventually, the measurements are pushed to a back end through the Internet.

A center node can serve concurrently multiple end nodes. The probability of simultaneous packet transmission from the end nodes is negligible. In the unfortunate situation of a collision, the packets of measurements are lost but a small shift on the end nodes' clocks ensures that such a collision will not occur during the next broadcasting session.

Figure 1 depicts the final view of our IoT node pcb and its components.

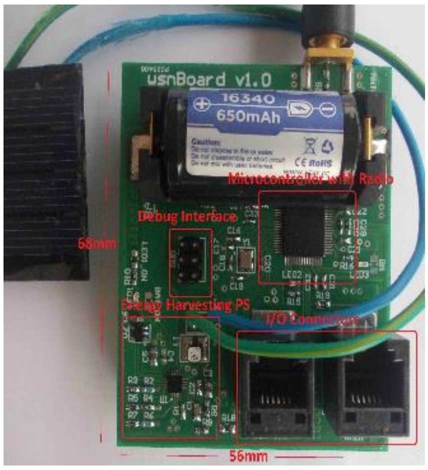



*Figure 1 IoT node pcb and components*

## C. System Performance

- Firmware

No embedded operating system is used. The bare metal firmware code is in C language and its total size, along with radio and sensors control libraries, is approximately 10Kbytes.

- Coverage Distance

The maximum, transmission distance is a function of antenna power, frequency and transmission speed. For 868MHz radio frequency, 1Kbps transmission speed and 10dbm antenna power, a 12.5Km line of sight has been achieved. For a 433MHz frequency the distance is doubled. The transmitter and receiver use antennas of half wave 1.2dBi peak gain. In case the field topology requires longer distances, then one or more IoT nodes will be used as repeaters (with minor modifications to the firmware).

- Energy Consumption

Hibernate mode is the dominant mode for the IoT node since the total consumption in this state is 6µA. During broadcast, for output power of 10dbm in the antenna, the consumption is 32mA. Fifteen (15) levels of output power are available, while the transmission speed can be selected from a range of 1Kbps to 300Kbps. Therefore, for a 20bytes transmission packet every ten (10) minutes with maximum antenna power and 1Kbps transmission rate, the average power consumption of the IoT node is approximately 22μA and its 650 mAh battery could endure for approximately 3.5 years. Node's battery will be recharged through the photovoltaic panel which has the ability to fully charge the battery within 20 hours of solar radiation.

- Cost Aspects

The below mentioned material costs are valid for prototype development. For larger scale reproduction, prices are significantly lower. More specifically, the cost of passive (L, R, C) is 6€, the cost for the ICs is around 11€, for the connectors 8€, for the battery 5€, for the photovoltaic panel 1€, for the antenna 6€ and 3€ for the IP65 casing. All the above sum up to a cost of 40€ per IoT node. Figure 2 depicts the final prototype.

*Figure 2 The Prototype IoT Node*

## IV. VERIFICATION AND SMALL-SCALE DEMONSTRATION

To demonstrate the operation ability of the proposed system the application domain of precision agriculture is chosen due to numerous sensors being used enabling the farmers to closely monitor the field and take precaution actions- thus substantially increasing the yield and quality and lower production costs. In this case, the system is used to gather the needed data from a vineyard pilot. The IoT prototype node, as described in section III has been reproduced in five (5) copies which are deployed at the pilot site. A representative example is shown in Figure 3.

*Figure 3 The IoT Node Integrated with Sensors*

The physical installation of the IoT nodes and the sensors followed a sensing cell approach as Figure 4 depicts. Each sensing cell covers 25m radius and the minimum distance between two nodes is 50m. In the particular use-case, in each sensing cell we gather various set of data. For example, in sensing cell 1 the following parameters are collected: soil moisture, air temperature & humidity and atmospheric pressure. In sensing cell 2 the following parameters are collected: leaf wetness, air temperature & humidity, atmospheric pressure. In sensing cell 3 the following parameters are collected: rain amount, air temperature & humidity, atmospheric pressure. All the above parameter data have been gathered uninterruptedly in a backend system for approximately five (5) months.

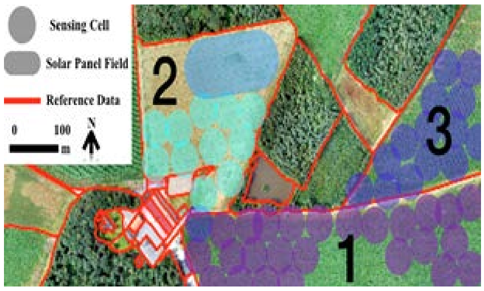



*Figure 4 Topology of the Application Use Case*

## V. CONCLUSION

In this paper, an integrated low cost IoT Node based on discrete components is presented that efficiently collects, and processes field data. It is designed to be robust, modular and easily adopted to any application domain. Its uniqueness lies in the way it integrates features as well as the integrity it exposes, and the new business model it supports. All its architectural as well as operational aspects are depicted, and the use-case example demonstrates the overall functionality of the proposed IoT node. Future actions include the conduction of exhaustive system audits in the field to assess the unobstructed connectivity of the nodes. The presented IoT node has been designed and implemented in the framework of the national project "AgrIoT", promoting a customized precision agriculture framework for environmental data collection.

## ACKNOWLEDGMENT

This research has been co-financed by the European Union and Greek national funds, the Regional Operational Program "Western Greece 2014-2020", under the Call "Regional research and innovation strategies for smart specialization (RIS3) in Microelectronics and Advanced Materials" (project: 5021449 entitled "Intelligent Services Based on the Internet of Things to Support Agriculture "AgrIoT")

## REFERENCES

[1] Official Arduino Website. Available online: https://store.arduino.cc/genuino-101

[2] Official LinkIt Website. Available online: https://labs.mediatek.com/en/platform/linkit-one

[3] https://www.espressif.com/en/products/hardware/esp8266ex/overview)

[4] https://www.seeedstudio.com/SeeedStudio-BeagleBone-Green-p-2504.html

[5] Naik, N. Choice of effective messaging protocols for IoT systems: MQTT, CoAP, AMQP and HTTP. Proceedings of the 2017 IEEE International Systems Engineering Symposium (ISSE)

[6] Official CoAP Website. http://coap.technology/

[7] Official MQTT Website. http://mqtt.org/

# Data Simulation as a Service: The I3T Paradigm on the IoT

D. Karadimas(*), C. Panagiotou, J. Gialelis, C. Chousiadas, N. Panagis, S. Koubias
Industrial Systems Institute, ATHENA Research and Innovation Centre, Patras, Greece
Applied Electronics Laboratory, Electrical and Computer Engineering dpt., University of Patras, Greece

*Abstract*—**This paper describes the architecture of a new cloud-based tool that focuses on the simulation of data generation from CPS/IoT devices and services. The paper defines the problem statement that the tool addresses and positions itself with respect to the current state of the art. The architectural overview of the tool and the adopted technologies are presented in detail. Finally, the status of the work in progress is presented along with the forthcoming steps.**

***Keywords-IoT, CPS, cloud, simulated data***

## I. Introduction

The evolution of embedded networked devices has driven the rise of the disruptive technologies of cyberphysical systems (CPS) and Internet of Things (IoT). In parallel, artificial intelligence and machine learning, took advantage of the massive data generation produced by devices sensing and acting on people and their environment, and presented remarkable progress in various application domains (e.g. healthcare, ambient sense & control, industry, etc.) [1].

The intelligence that is introduced in computational systems and services is strongly related and depended on the availability of data. Prior the deployment of an Artificial Intelligence (AI) algorithm in production environment, data scientists require a large volume of data in order to train, test and validate their models [2]. While, the number of devices and services that produce a variety of modalities is increasing day by day, access on these data is not always feasible due to various factors. In many cases, access on these data is not permitted due to privacy regulations. In other cases, the integration with live systems and the usage of their Application Programming Interface (API) to gather the required data demand significant efforts that cost both money and time. Moreover, while some tools that produce data based on existing models do exist, they fail to scale. Consequently, simulating the large volume of data that the numerous CPS and services generate nowadays, in the context of IoT, is not yet efficient. The impact of this lack of modern data simulation tools, reflect on the accuracy and the feasibility of applications that are based on AI and machine learning techniques.

The scope of this paper is to present the architectural design of a web-based, publicly available data simulation tool that addresses the need for "data generation as a service".

In that context, this paper identifies the need for designing and developing tools that address the need for the ever-increasing trend for data. In Chapter 2, problem definition and the contribution of this paper with respect to the literature is manifested. The overall architecture and operational model of the tool is described in detail in Chapter 3. The tool is evaluated through a power consumption web application that is presented in Chapter 4. Finally, the paper concludes with the current status of the work Chapter 5.

## II. Motivation

### A. Problem Definition

As briefly described in the introductory section of this paper, nowadays, the ever-creasing invasion of AI and machine learning in everyday consumer electronics and web services demand for large volume data consumption. Modern intelligent computational systems owe their evolution to the increase of the generated data coming from cyber-physical systems (CPS), IoT devices and services. However, researchers working in various application domains do not have always this high availability of data. On the other hand, artificial intelligence algorithms cannot succeed their cause unless they are trained and/or applied to large volumes of data. This paper has been motivated by the need for services that provide the necessary tools to generate data of variable volume based on predefined data generation models that the tool is provided by the user.

### B. Background

The typical approach for data scientist is to search for public available data repositories that publish appropriate datasets. Such popular repositories are Physionet for medical data [3] and Kaggle, provided by Google, offering data for a wide range of domains. The common path is to handle large csv files with the obtained data and write handlers that feed machine learning libraries written in popular languages such as Python, R, Mathlab, etc. Other approaches try to develop or utilize existing a series of models that use the physical and dynamical knowledge on the domain as well as computational intelligence [4].

## III. The I3T Approach

In the context of I3T project (Innovative Application of Industrial Internet of Things (IIoT) in Smart Environments) [5], a novel architecture has been designed for the development of a

978-1-7281-1740-9/19/$31.00 ©2019 IEEE

data simulation tool in order to facilitate researchers, scientists and engineers to feed their web services, and AI systems with generated data, that simulate the behavior of modern CPSs. The high-level architectural overview of the tool is depicted on *Figure 1*. The requirements that were defined in order to address the scope of the proposed tool comprise mainly of scalability, technology agnostic, application agnostic, open connectivity.

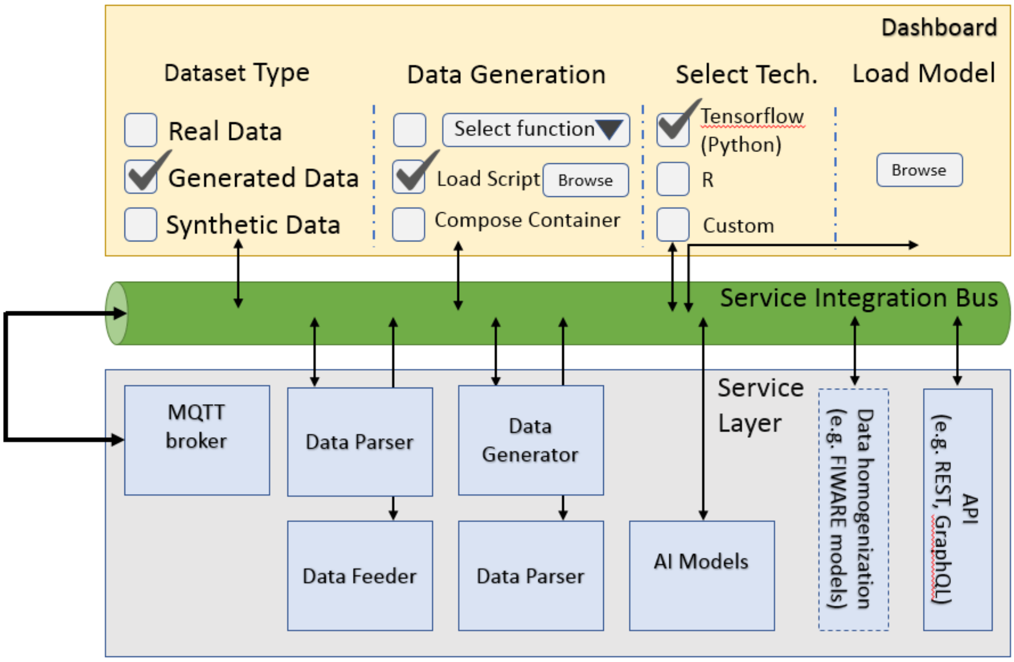



*Figure 1: Tool Architecture*

*Scalability*: The tool should be able to generate large volumes of data to address applications with similar characteristics (e.g. smart-cities). Therefore, the designed architecture should scale seamlessly to address the required throughput of data generation.

*Technology agnostic:* As mentioned earlier, data processing methods and AI techniques are implemented in several technologies. Apparently, data experts need to reuse their expertise without spending time on getting familiar with new programming environments.

*Application agnostic:* The focus of this work was to present an architecture as generic as it gets that tackles a large variety of application domains and not an application specific solution.

*Open standards:* Finally, in order to facilitate the integration with 3rd party services and infrastructure, the tool should support popular integration patterns based on open technologies.

In that context, the I3T-DaaS (Device as a Service) tool was designed on the principles of microservices architecture [7]. Microservices have gained significant interest lately due to its inherent characteristics such as: high maintainability/testing, loosely coupled services, improved fault isolation, highly distributed environment.

To serve the principles of microservices architecture the docker framework was adopted as the container platform that hosts all the required services [8]. All the individual services of the tool have been isolated in independent docker containers each one dedicated to deliver the functionality of the service it implements. The core of these services comprises of:

*1)* Dataset generation and feed: As already described, the main contribution of this tool is the ability to generate data with various configurations. The first basic functionality of this service is to give the ability to the user to load files with data and parse them in order to feed the other services such as AI prediction models, APIs, or other 3rd party services. When real data are not available, the user is able to generate simulated data with a variety of data distributions (e.g., normal, Poisson, polynomial, uniform, etc.) and configurations (volume, dimensions, frequency, etc.). For more sophisticated data generation, the user can pass to the service its own scripts in R or Python and generate data based on custom models. Another option that the tool provides is the creation of synthetic data, where simulated data are generated based on existing datasets with real data. Apparently, the user can enhance the performance of the service, by providing its own models and adapt to the characteristics of its application. All this functionality is also depicted in *Figure* 2 and is provided to the user through the web interface of the Dashboard.

*2)* Data persistence: An additional but optional service, enables the local storage of the data (whether coming from real dataset or models). This service is fed by the generator service and stores the generated data in databases optimized for timeseries (e.g. InfluxDB [9]), which are ideal for the sensor data of IoT applications.

*3)* Data homogenization comes in the picture in order to provide a common representation format for the data that the tool generates. Through the adoption of an open standard, the tool can integrate with IoT infrastructures with less implementation efforts and facilitate the optimization of existing services. The homogenization service implements the FIWARE data models to represent the generated data along with the virtual IoT/CSP devices that generate the data [10].

*4)* API is another key service that enables the extraction of the data (whether real of simulated) using popular http technologies such as the REST web services paradigm. Given the volume of data that could be stored by the tool and made available through the API, query languages such as GraphQL are supported on top of the REST API in order to enable clients to ask and get exactly what they need [11].

*5)* Another core service that links all the services into a single communication bus is the MQTT [12] broker (as implemented by RabbitMQ [13]). The communication between the I3T-DaaS dashboard and the rest of its services is achieved through message passing as defined by the popular MQTT protocol.

*6)* Finally, the Dashboard service, is responsible for the configuration of each service and the orchestration of the data flow through an intuitive web interface. The user through the dashboard can define the data generation strategy, to decide whether the data will be stored locally, to enable/disable AI services and expose the data through the API.

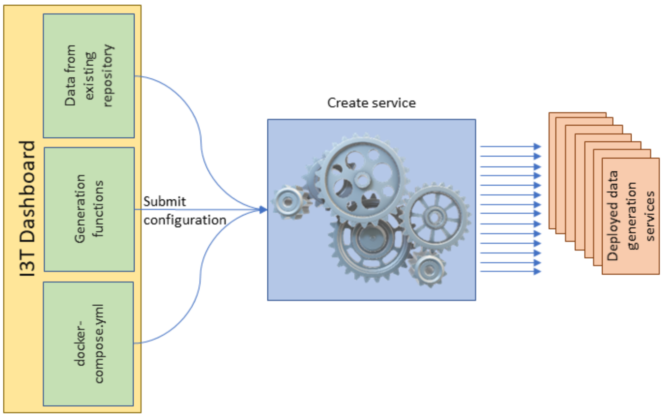



*Figure 2: Data generation service*

The provided functionality is given as a constellation of docker services where the user can deploy to its own infrastructure and enhance with his custom models and datasets. Based on the volume of load and the required resources the user can scale its deployment though the tools that the docker technology provides. As an example, if the API service of the I3T-DaaS is under high load to serve requests coming from external services, the user can increase the number of the API instances with a single command (e.g. *docker service scale i3t_api=5*). These instances can be distributed to several docker hosts that form a docker swarm without any additional effort from the user to orchestrate and configure these services.

## IV. Evaluation

For the evaluation of the tool, we developed an end-to-end application capable of monitoring the power consumption of a building through a self-intuitive web interface and providing short-term predictions based on the current usage of the domestic appliances. The I3T-DaaS tool had a 2-fold contribution. Initially it was used to support the training and testing process of the AI model. Then, it was used to simulate electric and electronic devices of the building by exposing their consumption.

Particularly, the datasets that were used regard a household's measurements of electric power consumption, taken with one-minute sampling rate over a period of almost 4 years, containing 2.075.259 measurements gathered in 47 months. The data attribute comprises of:

- Time
- Global active power
- Global reactive power
- Voltage
- Global intensity

The data generator service was assigned with the task of data preprocessing and resampled the data from a sampling rate of 1-sample/minute to 30-samples/minute. For that purpose, the python pandas library was utilized [14]. The purpose of this resampling was to reduce the dataset size and predict forthcoming consumption with a smaller sequence of previous values. The data were normalized through the minmaxscaler in order to facilitate the training of our model. Finally, using the Tensorflow [15] tool, the Long Short Term Memory (LSTM) artificial recurrent neural network architecture [16] was trained on these data in order to predict feature power consumption of the devices.

Without any further optimizations, the model presents the performance as described in Table 1, while plots of the predicted values against the ground truth are given in Figure 3.

TABLE 1: MODEL PERFORMANCE

| | |
|---|---|
| Loss | 0.040 |
| Mean Absolute Error (MAE) | 0.14 |
| Root Mean Square Error (RMSE) | 0.78 |

The generated data and its predictions have been consumed by the an "Online Platform for the Control of Domestic Appliances". The platform is mainly targeted to electrical installation service companies. It aims to enable them to provide their customers with a user friendly and efficient experience for controlling their domestic appliances and monitoring the ambient conditions and the residential security. It also gives them the ability to have a user-friendly graphical view of all the networked devices within a space. The status and the controls of each of the devices are placed in a top down view that fully reflects the actual design of the residence. In that sense, the user needs minimal effort to get familiar with the interface since all that he needs is to virtually navigate to his own residence and interact with the avatars of the electric/electronic devices that are installed and integrated to the platform.

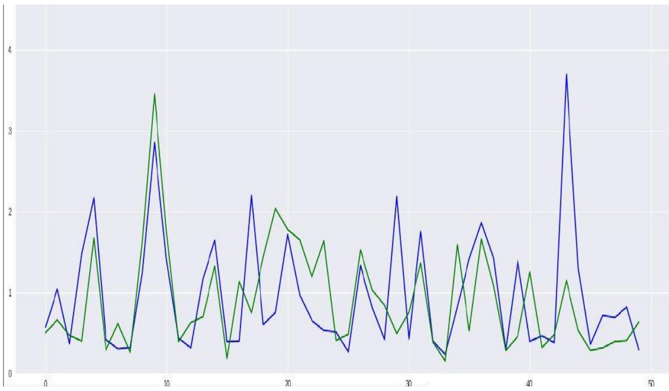


Figure 3: Prediction


Figure 4: Power Consumption graph (history and prediction)

The monitored data (power consumption of domestic appliances) is a mixture of real data coming from devices installed in the lab and simulated data generated by the I3T-DaaS tool and exposed through its REST API. The graphical interface presents the real time values of the consumption of each one, along with the available controls (on/off). Finally, the user is offered a provision of future consumption based on the current patterns given by the trained LSTM model. A view of that interface is given in Figure 4.

## Conclusions

The scope of this use case was not to present an optimized prediction tool for domestic power consumption but to highlight the role of the I3T – DaaS in such workflows. The services of the tool could be deployed in local environment where the user could take advantage of its hardware resources (both CPU and GPU). The tool source code will be available when completed as an open source project through a github repository while, a live web version will be also available for researchers to submit and run their project.

## Acknowledgment

This work was undertaken under the “Innovative Application of Industrial Internet of Things (IIoT) in Smart Environments (I3T)” project funded by National Strategic Reference Framework (2014 – 2020) for the Industrial Systems Institute, ATHENA Research Center.

## References

[1] A. Ghosh, D. Chakraborty and A. Law, "Artificial intelligence in Internet of things," in CAAI Transactions on Intelligence Technology, vol. 3, no. 4, pp. 208-218, 12 2018.

[2] Witten, Ian H., et al. Data Mining: Practical machine learning tools and techniques. Morgan Kaufmann, 2016.

[3] Goldberger, Ary L., et al. "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals." Circulation 101.23 (2000): e215-e220.

[4] Haupt, Sue Ellen, and Branko Kosović. "Variable generation power forecasting as a big data problem." IEEE Transactions on Sustainable Energy 8.2 (2017): 725-732.

[5] https://i3t.isi.gr/en/

[6] https://www.kaggle.com/

[7] J. Thönes, "Microservices," in IEEE Software, vol. 32, no. 1, pp. 116-116, Jan.-Feb. 2015.

[8] https://www.docker.com/

[9] https://www.influxdata.com/

[10] FIWARE http://www. fi-ware. Org

[11] GRAPHQL https://graphql.org/

[12] MQTT http://mqtt.org/

[13] RabbitMQ https://www.rabbitmq.com/

[14] Pandas https://pandas.pydata.org/

[15] Tensorflow https://www.tensorflow.org/

[16] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink and J. Schmidhuber, "LSTM: A Search Space Odyssey," in IEEE Transactions on Neural Networks and Learning Systems, vol. 28, no. 10, pp. 2222-2232, Oct. 2017.

# An Integrated Educational Platform Implementing Real, Remote Lab-Experiments for Electrical Engineering Courses

Dimitris Karadimas and Kostas Efstathiou
Dept. of Electrical and Computer Engineering, University of Patras, Greece
Email: {karadimas, efstathiou}@ee.upatras.gr

*Abstract*—**This paper describes an Internet-based laboratory, named Remote Monitored and Controlled Laboratory (RMCLab) developed at University of Patras, Greece, for electrical engineering experiments. The key feature of this remote laboratory is the utilization of real experiments, in terms of instrumentation and under measurement circuits, rather than simulation or virtual reality environment. RMCLab's hardware infrastructure contains multiple reconfigurable sub-systems (FPGAs), which can be enhanced by almost any analog expansion module. The main characteristics of this system include the versatility of the hardware resources, due to the dynamic reconfiguration potentiality and the low cost of the hardware components. Moreover, this system enables its users to test, in real time, their own custom circuit designs. The paper concludes with a specific example regarding an elementary circuit in digital electronics and a short statistical review of the RMCLab educational usage. RMCLab can be accessed via the web through http://www.apel.ee.upatras.gr/rmclab.**

*Index Terms* — **Client-server architecture, remote laboratory, distributed instrumentation and resources**

## I. Introduction

The exponential growth of computer and internet technology enables the development of complex, hybrid systems such as remote laboratories where experiments can be remotely accessed, monitored and controlled [1]-[4]. This new interpretation of the measurement process offers to anyone the opportunity to interact with the laboratory at any time, reducing at the same time the experiment cost per user and extending the capabilities of the entire experimental framework.

Paradigms of using these advanced facilities apply either for educational purposes [5]-[6] or for products' advertisement. Remote laboratories can offer high-level experimental training and experience, when they are able to realize, support and interact with real experiments, rather than present simulation results or simple depiction of reality. Additionally, expensive, often dedicated experiments, of modern, cut-edge technology can be shared worldwide, contributing thus to a high-valued remote laboratory framework.

Many internet-enabled software systems that afford distance laboratories via simulated, virtual environments can be found in the web [7]-[8]. These software systems often integrate many of the desired functionalities, especially from the user's side, such as accompaniments to the experiment documentation, communication support and collaboration among their users. Although modern simulators can accurately estimate circuits' performance, the employment and utilization of real circuits and real instrumentation, for electrical engineering laboratories [9], ensures the measurements' reliability, while at the same time increases the educative value of the remote laboratory.

Remote laboratories offering access to real lab-experiments and real instrumentation also exist, however the majority of them cannot share their resources simultaneously to many users, thus they fail to serve and support large classes of several hundreds of students.

This paper presents the specifications and the basic structure of an integrated remote laboratory platform that enables the instant remote access to real lab-experiments, employing real hardware and real instrumentation. This platform, named Remote Monitored and Controlled Laboratory (RMCLab), is able to provide high-level services to a great number of users for a wide-range of real electrical engineering experiments; either pre-configured, reconfigurable or customizable, at a very low hardware infrastructure cost. RMCLab offers its services since March 2004, at the Dept. of Electrical and Computer Engineering of University of Patras, Greece, where it was developed and implemented.

## II. Proposed Approach

The basic purpose of the developed platform is to provide high-quality lab-training in electrical engineering subjects to students, all over the world. The design of such a remote laboratory for real-time, internet-based lab-experiments, should consider all aspects of the system, including communication and data flow, as well as instrumentation and hardware control [10]-[11]. RMCLab system has been designed so as to integrate all potentials of a physical laboratory to a simple user interface, among with other sub-systems, such as lab-administration, instrument operation and hardware management.

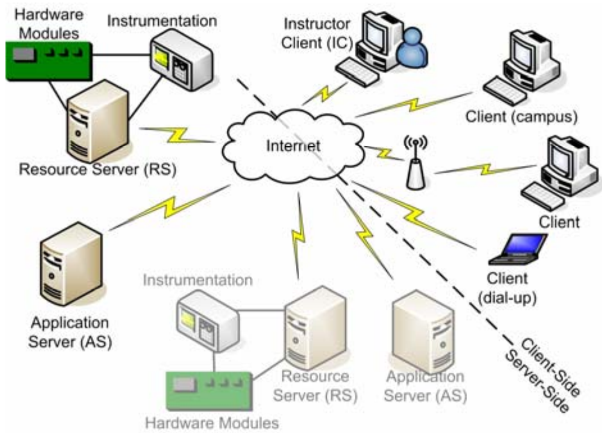


Figure. 1. RMCLab system overview.

The primary service that RMCLab platform should provide to its users is the possibility to study on the lab-experiment subjects, by accomplishing their measurements at any time and from anywhere. For this reason, RMCLab's basic specification is defined as the ability to serve at any time, simultaneously and at real time, any potential user for any available lab-experiment. On the other hand, an integrated remote laboratory platform should reinforce lab-administrator's tasks and responsibilities, regarding the experiment setup, hardware and instrumentation control, users' management and also lab-maintenance. RMCLab offers also many kinds of assessment functionalities for the students' lab-skills, regarding the lab-experiments, such as the assignment of several different evaluation criteria (measurements, instrument settings and multi-type questions, etc), so as the whole platform can be configured as an advanced tool for automated, high-level educational services, an aspect that characterizes the offered educational activities and also our initial motivation.

## III. Architecture

RMCLab system has been developed based on the conventional client-server architecture, expanded in the server-side, as depicted in Fig. 1, and consists of the following basic entities: client, instructor-client (IC), application server (AS), resource server (RS) and lab-infrastructure, including the real instrumentation and all the hardware modules.

### A. Network Topology

The server-side of the proposed architecture employs at least two sub-servers; the resource server and the application server. This structure could also be replicated in a more complex network topology. Resource server manages and operates hardware and instrumentation resources, providing to application server an abstract layer for communication that enables access to lab-infrastructure.

Application server undertakes the data flow control task between clients and the physical remote laboratory, realized by the resource server and the lab-instrumentation. The intermediary role of the application server is mandatory, since clients should not be to directly communicate with any of the resource servers. The communication between application server(s) and the resource server(s) or the client(s) is based on a custom, abstract language that integrates all potential tasks of a conventional, physical laboratory.

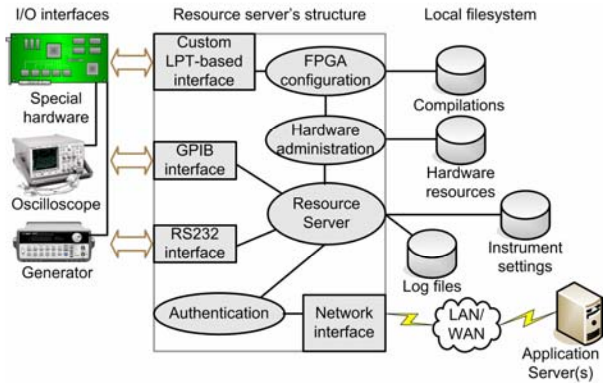


Figure. 2. Software and hardware modules of the Resource server.

This topology simplifies the architecture of the server-side and expands platform's capabilities, as it facilitates the robust development and customization of a resource server. Moreover, it enables many application servers to utilize the components shared by the resource server(s). As a result, users all over the world are able to transparently access, via the application server(s), these shared resources. Additionally, application server grants to its clients transparently access to the real resources of the physical laboratory, thus increasing system's robustness, flexibility and expandability.

### B. Hardware, Instrumentation and Resource Server

The real measurement laboratory is based on a low-cost and easy implementation, while it is realized around the resource server. Resource server is equipped with suitable, interfaces toward the signals of lab-experiments (both digital and analog experiments), via a custom, LPT based bus, and the instruments, via the RS232 interface or the GPIB bus, as depicted in Fig. 2.

Multiple types (standard, programmable, pre-configured or re-configurable) of analog, digital or mixed circuits can be hosted in the platform's resource server(s). For this reason RMCLab's hardware is outfitted with a motherboard that is able to host up to 64-cards, where each of them incorporates an FPGA and extra auxiliary circuitry required for implementing the lab-circuits, as depicted in Fig. 3.

Each card employs also a PLD, which is responsible for the card addressing and the configuration of the FPGA. Each of these cards can host 8-different analog, digital or mixed independent arbitrary circuits, since the FPGA is segmented into 8-sectors, each of them corresponding to a specific lab-experiment. The internal operation of the FPGA is controlled by a register file (Table I) which is employed within it.

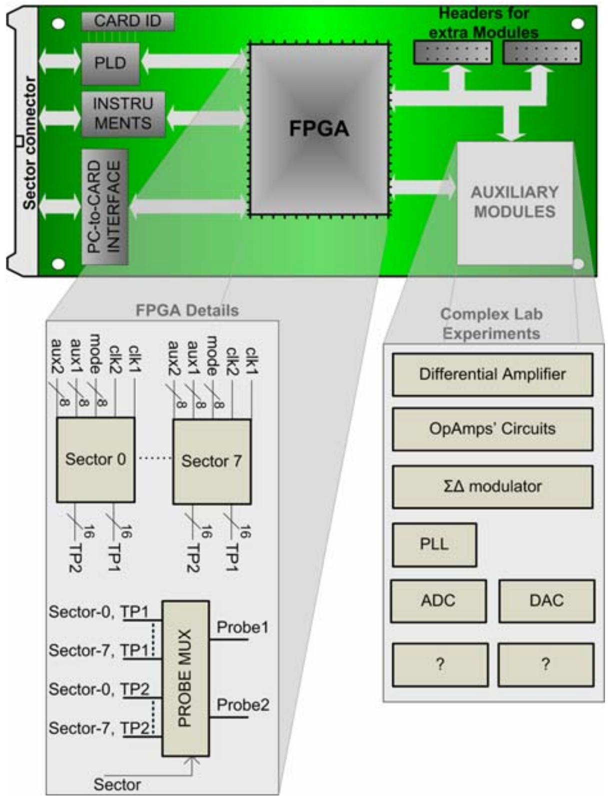


Figure. 3. Hardware architecture.

As each sector of the FPGA can host either a specific multi-mode lab-experiment or a user's custom circuit, the mode register and two auxiliary registers control its operation mode and behavior, while sector register points to the active sector, on which measurements are performed. Therefore, a single sector could implement alike lab-circuits, which can be externally presented as different lab-experiments, while the selection of the operation is performed by the value of the mode register. For example, in our case, both synchronous and asynchronous digital counters are implemented in the same sector, but are presented as two different lab-experiments. Moreover, when a measurement is carried out, two extra registers, Probe1 register and Probe2 register, assign the active nodes of the active sector, on which the two probes of the oscilloscope become physically connected through cross-point switches. Finally, each card may be offline equipped with additional on board or external circuitry, in order to implement a wide range of more complex electronic circuits, including PLL-based Frequency Synthesizers, several types of D/A or A/D Converters, ΣΔ Modulators, etc.

The aforementioned hardware architecture characteristics in combination with the network topology

TABLE I. FPGA REGISTER FILE

| Name | Address | Width (bits) | Operation |
|---|---|---|---|
| Sector | 0 | 3 | Select the active sector |
| Probe1 | 1 | 4-6 | Select the active nodes of Oscilloscope's Ch-A |
| Probe2 | 2 | 4-6 | Select the active nodes of Oscilloscope's Ch-B |
| Aux1 | 3 | 8 | Auxiliary register 1 |
| Aux2 | 4 | 8 | Auxiliary register 2 |
| Mode | 5 | 8 | Sector's operation mode |

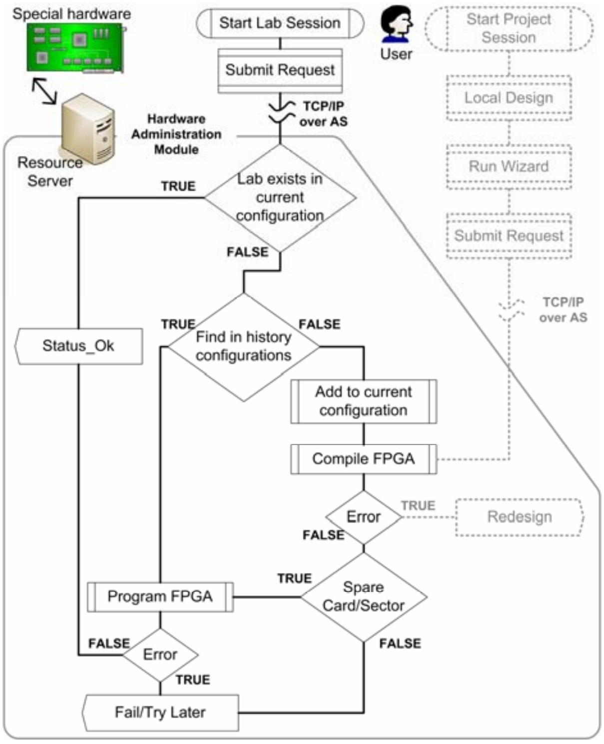


Figure. 4. Hardware administration module flowchart.

complexity necessitate an elegant and efficient management of the hardware resources and the measurement requests. A hardware administration module, depicted in Fig. 4, within the resource server, undertakes this management role. A measurement transactions starts when a client raises a measurement request, regarding either a pre-configured lab-experiment or a user's custom circuit. The later is discussed in details in Section IV. After the request is raised to the corresponding application server, by a client, is logged and forwarded, via the same application server, to the proper resource server, which supports the under-measurement circuit. Afterwards, resource server has to accomplish multiple tasks, such as the authentication of the request and the lab-infrastructure (hardware and instrumentation) setup, so as to be prepared for the requested operations. This may lead to a real-time, online re-configuration of one's card FPGA, so as to implement the requested circuit, or even to the removal of an unutilized sector's circuit, if an empty sector cannot be found. As soon as the hardware is configured, the measurement is performed and the

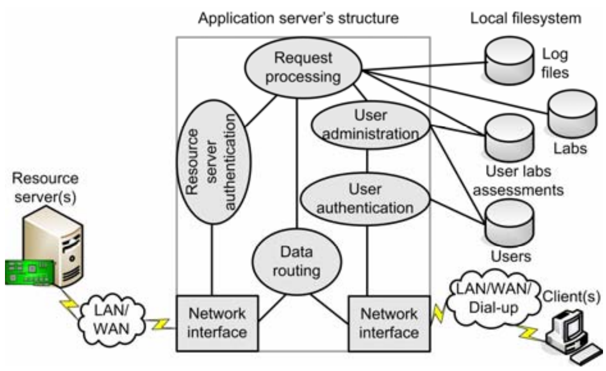


Figure. 5. Software modules of the application server.

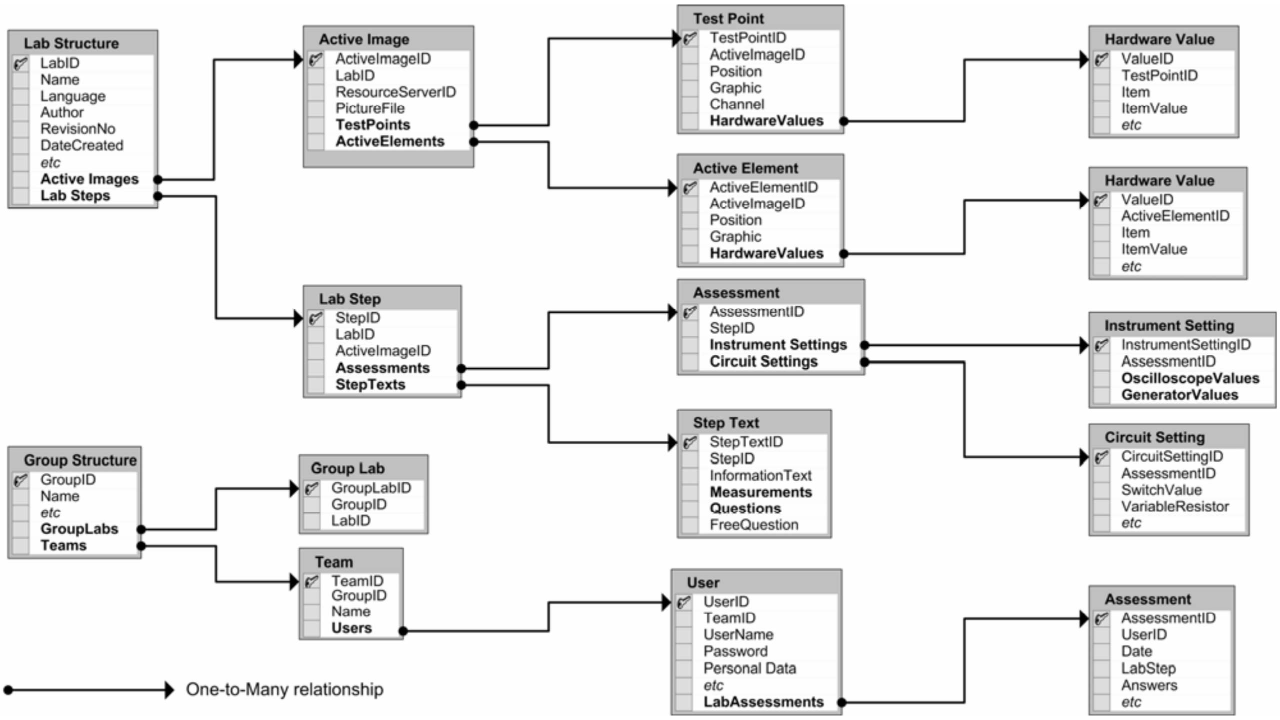


Figure. 6. Labs' and Users' data structure overview.

acquired data are transmitted back to the specific client, again via the application server. The above procedure has been designed so as to time-share the lab-infrastructure, in a FIFO priority, to all available requests.

### *C. Application Server*

Apart from the dataflow control and the routing procedure between client(s) and the resource server(s), application server is also responsible for the authentication and logging, as well as for the assessment and evaluation of its clients' actions, when educative usage is intended. The above presented characteristics and functionalities of the application server define its architecture, as depicted in Fig. 5. Additionally, each one of the application servers of the RMCLab platform needs to be offline aware of the resource servers, that are able to communicate with, and their list of supported circuits, which are dynamically acquired upon each successful transaction with one of the resource servers.

Application server comprises also an advanced tool for the development and maintenance of a laboratory class that is available to the administrator of the laboratory, as can it be unrelated to the location of the physical laboratory, realized by the resource server, the hardware and the instrumentation. The development and maintenance of a laboratory class has been merged into a single database system, depicted in Fig. 6, which contains all the required data for the design and assignment of a lab-exercise. Additionally, the same database system includes data regarding the students. One lab-exercise may consist of several active-images, which correspond to the real lab-circuits. For each lab-circuit, test-points and active-elements (switches and variable components) can be assigned. Hardware properties required for the assignment of the test-points and the active-elements are specified in the custom abstract language, used by the RMCLab system. On the other hand, a lab-exercise is separated in several steps, where each step may contain information, regarding the theoretical and practical aspects, measurements, multiple choice questions and a free-text question. Moreover, assessment rules may be provided in each lab-step. Along with the lab-exercises' data, students' data, regarding their personal information and assessments are stored in the same database system.

The layered networking of the RMCLab system permits each laboratory administrator to present a lab-experiment, running at a specified resource server, according to his personal educational aspects, regardless of the physical location of the real hardware of the lab-experiment.

### *D. Client*

The client-side of RMCLab's system has been designed so as to comply with the demands of a potential user. Thus, client module embeds a specific interface, named as 'scenario interface', for supporting the remote monitor and control of lab-infrastructure, and other full-functional and user friendly interfaces for lab-instrumentation (function generator, oscilloscope,

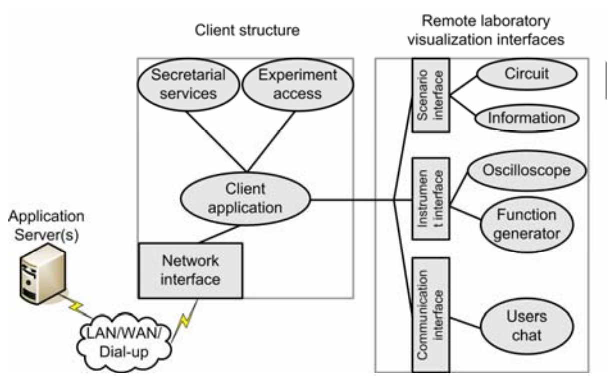


Figure. 7. Software modules of the client.

etc), as depicted in Fig. 7.

In more details, scenario interface provides a user with graphic information, related with the under study circuit. Additionally, grants to the user the control of circuit parameters (variable pots, caps, etc) and also the monitoring of any active node of the circuit, by selecting and setting the probes of the oscilloscope and the function generator on the circuit. Moreover, extra documentation, regarding the technical and theoretical aspects of the experiment, which can also be separated into multiple steps, can be presented to the user, via the scenario interface.

### *E. Communication Module & Instructor-Client*

In order to meet the basic requirements of the collaborative interactive e-Learning, a communication interface has been incorporated into the RMCLab system. This communication interface consists of a simple chat module enabling the collaboration and the information exchange, during a remote lab-experiment. The communication interface has been integrated into the RMCLab system by request of its early users, while it is under development the expansion of the chat module with voice and video capabilities.

Finally, RMCLab's platform embeds an identical to the user's client module, named as instructor-client, which offers to a supervisor/instructor of the experiment the ability to replicate, monitor and control any online user's lab-environment. This feature is focused on the educational aspect of RMCLab platform, as it provides an instructor with the ability to closely observe and efficiently tutor the actions of any online users, concluding to a 'near-to-real' lab-environment.

### *F. Architecture Overview*

The described hardware architecture is suitable for developing circuits of low-to-medium complexity, at a low-cost. For accessing the properties of this specific hardware, a software driver has been developed and embedded in the resource server application. Apparently, the platform is able to employ and control any hardware, under the condition that the corresponding software driver enables its access. Thus, even the use of complex or commercially available products is possible.

Additionally, the abstract language used to communicate RMCLab entities each other, can be modified according to future or different requirements. Fig. 8 depicts the dataflow diagram among the RMCLab entities and focuses on the correspondence between them.

## IV. Advanced Properties of the RMCLab

Conventional lab-education is based on the study of pre-defined lab-experiments. RMCLab system provides an outstanding benefit to its users, thus the feasibility to design and test/measure their own custom circuits under real hardware and real instrumentation. RMCLab users can offline design almost any circuit using separate software package (MAX+plus II or Quartus, both offered at no-cost for academic institutes from Altera). Using one of the aforementioned specific software packages, one can design his own circuits following a reduced set of rules and confirm by simulation its proper operation. Once the design is verified at the client-side, it can be easily uploaded to the server-side and after a while (<15 seconds) he will be able to perform any measurement on his custom design, which is now implemented on real hardware, by employing real instrumentation. The aforementioned procedure is supported by the hardware administration module of the resource server, as depicted in the dotted part of Fig. 4.

The network architecture of RMCLab platform enables the world-wide distribution of resources, in terms of lab-experiments, by utilizing multiple application servers in a single network topology, as depicted in Fig. 9. Thus, instructors all over the world can take the advantages of employing a running lab-experiment and present it in their native language and personal educational point of view. Obviously, each supervisor has the opportunity to review his users' performance by his own criteria, according to the assessments rules for each experiment, that are defined in the RMCLab's application server, which is available and accessed by the supervisor, as resource server transparently executes the measurement requests.

The prospects of the RMCLab system may hopefully expand world-wide, as the above scenario can be further extended if one adds more resource servers, as depicted

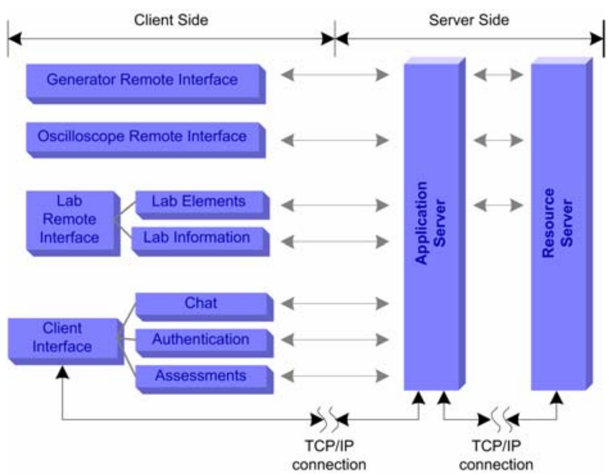


Figure. 8. RMCLab system's dataflow overview.

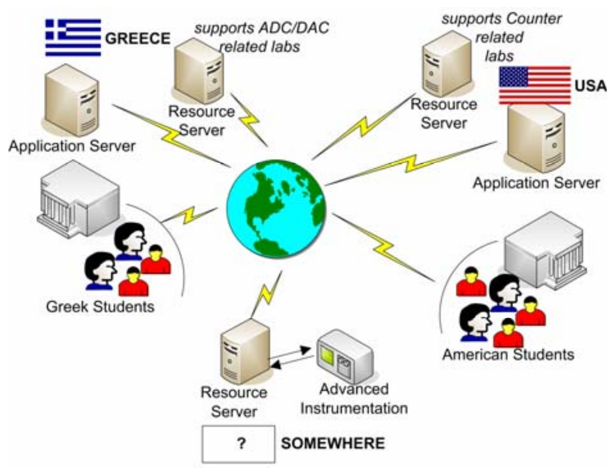


Figure. 9. Advanced RMCLab utilization.

in Fig. 9. Each resource server can be expert and focused on a specific subject, incorporating the appropriate hardware and instrumentation. Instructors all over the world may take the advantage of using such laboratory resources and develop educational material in their local application servers, in their native language, according to their own educational criteria, so as to offer advanced experimental training to their students, without any requirement for the development and maintenance of any expensive lab-infrastructure.

RMCLab's advanced utilization modes are not limited within the above example. The real-time use of real hardware and real instrumentation can significantly contribute to the educational procedure, since it enables an instructor to prepare 'Active Lessons' and present in details, during a class, the operation of a circuit or a system under real world circumstances, while at the same time can be utilized as a mean of demonstration for expensive products.

## V. THE REALIZED RMCLAB SYSTEM

### A. *Technical Characteristics*

The architecture described in section III has been implemented at the University of Patras, Greece. Current configuration is a cost effective implementation, which consists of a single PC, with an Intel Hyper-Threading 2.6GHz processor and 1024MB RAM, embedding both the resource and the application server of our running RMCLab system. This PC, running Windows 2003 Server, is permanently connected to the campus LAN and also to an Agilent 54622D mixed signal oscilloscope and an Agilent 33120A function generator. Oscilloscope is connected with the PC via a high-speed GPIB interface, while function generator is controlled via RS232 @19.2kbps. The PC-interface with the hardware modules is implemented based on a custom, low-cost bus, through LPT in EPP mode. Each card of the hardware infrastructure contains an Altera FPGA of the FLEX8K series and also other components required for the implementation of the experimental circuits. The aforementioned infrastructure provides fast enough access and response (<3-secs per measurement) to the client requests, as summarized in Table II.

### B. *A Simple Educational Paradigm*

Fig. 10 illustrates the measurement result, regarding the CLK and LOAD signals of an Early Decoding, Count Down, 4-bit Decimal Counter. The top part in Fig. 10 depicts the circuit information available to RMCLab users for the specific circuit, while the middle and bottom part depict the measurement representation of the real

TABLE II. RMCLAB TIME RESPONSE CHARACTERISTICS.

| Property | Average Delay (sec) |
|---|---|
| Hardware setup and measurement time | 3 |
| Compilation time of a custom circuit | 10 |
| Hardware re-configuration time | 5 |
| Measurement delay from client side using PSTN line @56kbps | <5 |

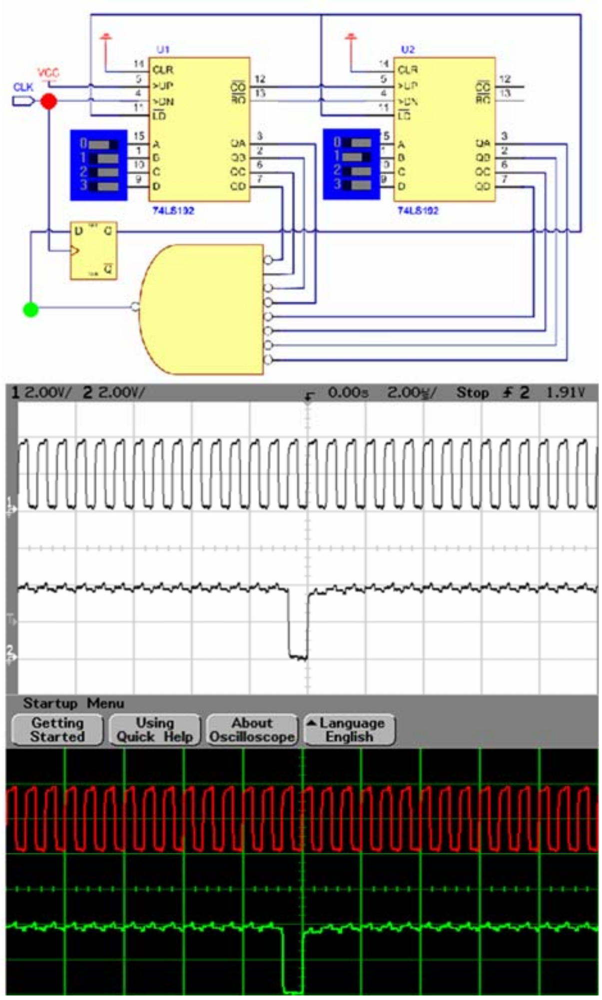


Figure. 10. Educational paradigm of RMCLab usage.

oscilloscope and the remote interface, respectively. Obviously, RMCLab system is able to take full advantages of the real hardware and real instrumentation utilization, providing measurements, regarding a wide variety of signals' properties, and the full control of both instruments (oscilloscope and function generator).

### C. *Educational Utilization*

RMCLab system provides its educational services since March 2004 to the Dept. of Electrical and Computer Engineering of University of Patras, Greece, supporting classes of approximately 300-students, in two core lessons, regarding Analog and Digital Electronics. Analog lab-experiments include 2-stage feedback amplifiers and cascade/folded-cascade amplifiers, whereas digital experiments include a wide variety of topologies regarding counters, adders and accumulators.

A class of the Dept. of Electrical and Computer Engineering of our University has about 300-students. Students are grouped in teams of 3-to-4 persons per team, in order to perform 6-obligatory lab-experiments per semester. For convenience, the same grouping was retained for accessing RMCLab services. Thus, for the second semester of academic year 2004-'05, RMCLab has to offer its services to 80-teams, for 2-obligatory lab-experiments. Five more obligatory lab-experiments were also carried out in the conventional way. The two

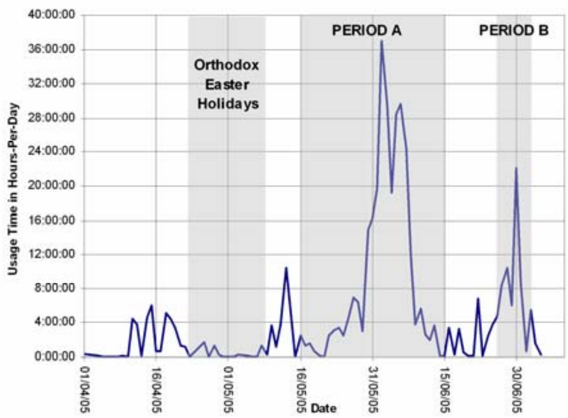



Figure. 11. RMCLab's cumulative utilization during the second semester of academic year 2004-'05.

RMCLab-based experiments referred to binary/decimal synchronous counters and special purpose programmable up/down counters, using the 74192 chip.

Students' obligations regarding the RMCLab-based experiments were announced the second week of April '05. Fig. 11 depicts the RMCLab's cumulative utilization, considering that students had to carry out these experiments in PERIOD A, while the re-examination took place in PERIOD B.

During the aforementioned period, RMCLab has been extensively employed for its services; thus 17200 measurements were logged on the platform's instrumentation, where 1666 of them regarded an introductory exercise, and 4458, 11076 measurements regarded the first and the second obligatory exercise, respectively. For the first obligatory exercise the performed measurements were 4.35-times more than the total number required for completing the exercise by an expert, while for the second one this ratio was reduced to 3.97, despite the fact that this exercise was significantly more demanding for the students. This implies that students easily acquired experience on the use of the RMCLab platform.

During the same period, up-to 8-simultaneous requests have been raised to the RMCLab resource server, without importing any extra delay to the users' requests servicing. The 74-active teams fulfilled their lab obligations in 383h 30m 20s, thus, approximately 2.6-hours per lab-experiment. Fig. 12 and Fig. 13 depict the cumulative

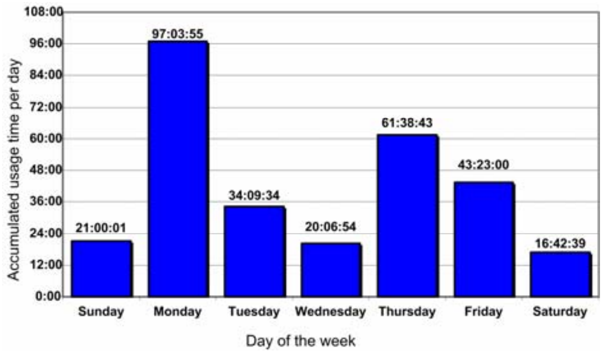



Figure. 12. RMCLab's cumulative usage time vs. day of the week for the second semester of academic year 2004-'05.

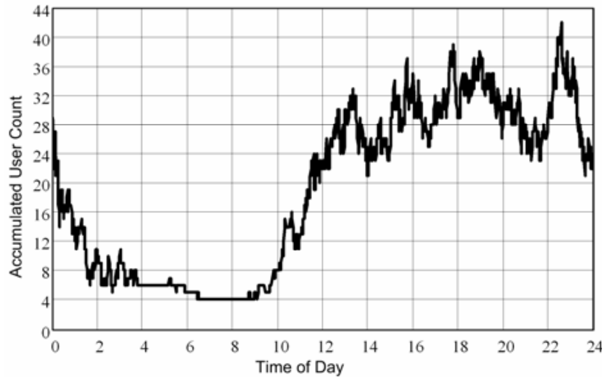



Figure. 13. RMCLab's cumulative user count vs. time of the day for during the second semester of academic year 2004-'05.

user count and the cumulative usage time of the RMCLab services versus the time of the day and the day of the week, respectively, for the aforementioned period. Apparently, services like RMCLab enable students to take over their obligations in a reasonable time while exploiting efficiently their wide spread working hours. Additionally, RMCLab platform carried out extensively and fair the assessments of the students' performance, regarding their lab skills, thus providing a valuable service for the instructor.

## VI. Conclusion

RMCLab platform is able to provide a wide range of high educational services in a great number of students. It increases the productivity of the students by enabling them to have access to the lab-infrastructure at non-working hours, while at the same time affects significantly their psychological mood regarding the level of the offered education by their institute.

Moreover, RMCLab accomplishes its services employing a single PC and a single set of hardware and instrumentation, thus pointing out that is able to provide high-quality lab-education at low-cost, without time-consuming human insteraction.

The structure of RMCLab enables sharing of hardware and instrumentation resources, thus makes possible the extensive exploitation of an expensive lab-infrastructure, facilitating the wide spread of remote real lab-experiments, which are indisputably valuable for engineers' education. Additionally, hardware re-configurability permits the remote implementation and measurement of electronic circuits, providing further more a high-valued educational service.

The concentrated use of RMCLab system during 6-academical semesters, for the courses of Analog and Digital Integrated Circuits, consisting of classes of about 300-students per class, has definitely proved the high value of this educational tool, for the students and for the instructors as well.

It is anticipated that the proposed architecture guidelines along with the success of the RMCLab platform will motivate educational community to cooperate, so as to develop an integrated World-Wide-Lab environment.

## References

[1] Harry Boldt, “Didaktisches Design medialer Lernangebote im Umfeld ingenieurwissenschaftlicher Laborpraktika: Didactical Design of media learning offers in the surrounding field of engineer-scientific laboratory practical courses”, University Duisburg-Essen, Campus Duisburg, Aachen 2005

[2] G. Mason, “A handheld data acquisition system for use in an undergraduate data acquisition course”, *IEEE Trans. on Education*, vol. 45, No4, pp. 388-393, November, 2002.

[3] J. Sebastian, D. Garcia & F. Sanchez, “Remote-access education based on image acquisition and processing through the internet”, *IEEE Trans. on Education*, vol. 46, No1, pp. 142-148, February, 2003.

[4] F. Pianegiani, D. Macii & P. Carbone, “An open distributed measurement system based on an abstract client-server architecture”, *IEEE Trans. on Instr. and Meas.*, vol. 52, No3, pp. 686-692, June, 2003.

[5] P. Lin & M. Lin, “Design and implementation of an internet-based virtual lab system for eLearning support”, *Proc. 5th IEEE International Conf. on Advanced Learning Technologies*, Kaohsiung, Taiwan, 2005, pp. 295-296.

[6] A. Bagnasco & A. Scapolla, “A grid of remote laboratory for teaching electronics”, *2nd International LeGE-WG Workshop on e-Learning and Grid Technologies: a fundamental challenge for Europe*, Paris, 2003.

[7] J. Palop & M. Teruel, “Virtual work bench for electronic instrumentation teaching”, *IEEE Trans. on Education*, vol. 43, No1, pp. 15-18, February, 2000.

[8] C. Daly & J. Horgan, “An automated learning system for java programming”, *IEEE Trans. on Education*, vol. 47, No1, pp. 10-17, February, 2004.

[9] Dimitris Karadimas and Kostas Efstathiou, “Design, Implementation and Evaluation of a Remote Laboratory System for Electrical Engineering Courses”, Proc. 6th IEEE International Conf. on Advanced Learning Technologies, Kerkrade, The Netherlands, 2006, pp. 408-412.

[10] T. Fjeldly & M. Shur, “Lab on the web: Running real electronics experiments via the internet”, Wiley-IEEE Press, 2003.

[11] C. Ko, B. Chen & J. Chen, “Creating web-based laboratories”, Advanced Information and Knowledge Processing, Springer, 2004.



**Dimitris S. Karadimas** was born in Patras, Greece in 1980. He received his Diploma on Electrical and Computer Engineering in 2003 at the Dept. of Electrical and Computer Engineering of University of Patras, Greece.

Since then he is a Ph.D. student in the aforementioned Dept., where his research area includes data converters, frequency synthesis techniques and remote laboratory education.

Mr. Karadimas is a member of IEEE.

**Kostas A. Efstathiou** was born in Volos, Greece in 1958. He received his Diploma on Electrical Engineering in 1981, and his PhD on the same field in 1996, both at the Dept. of Electrical and Computer Engineering of University of Patras, Greece.

He is with the aforementioned Dept. since 1981 as a Research Associate and as Faculty since 1998. He has published several conference and journal papers, regarding industrial electronics and networking, frequency synthesis and synthesizers, data converters and lab education.

Dr. Efstathiou is a member of IEEE and of the Technical Chamber of Greece.

# A WEB-BASED, DISTRIBUTED, REAL EDUCATIONAL LABORATORY FOR ELECTRICAL ENGINEERING COURSES

E. Kostas* and K. Dimitris*

**Abstract**

This paper presents an Internet-based laboratory, named as Remote Monitored and Controlled Laboratory (RMCLab), aiming to provide high-quality lab training in electrical engineering subjects to students all over the world. The key feature of this remote laboratory is the utilization of real experiments, by employing real instrumentation and real circuits, rather than simulation or virtual reality environment. RMCLab's hardware infrastructure contains multiple reconfigurable sub-systems (FPGAs), which can be enhanced by many analog expansion modules. The main characteristics of this system include the versatility of the hardware resources, due to the dynamic reconfiguration potentiality, as well as the low cost of the hardware components. Moreover, this system enables its users to test, in real time, their own custom circuit designs. This paper provides information regarding the architecture of RMCLab and concludes with a specific example regarding the implementation of an elementary circuit in digital electronics and a short statistical review of the RMCLab educational usage.

**Key Words**

Client–server architecture, remote laboratory, distributed instrumentation and resources

## 1. Introduction

The exponential growth of computer and Internet technology enables the development of complex and hybrid systems, in terms of software and hardware composition, such as remote laboratories where experiments can be remotely accessed, monitored and controlled [1]. This new interpretation of the measurement process offers to anyone the opportunity to interact with the laboratory at any time, at the same time reducing the cost of experiments per user, while the use of modern technology extends the capabilities of the entire experimental framework.

Paradigms of using these advanced facilities apply either for educational purposes [2] or for products' advertisement. Remote laboratories can offer high-level experimental training and experience, provided that they are able to realize, support and interact with real experiments, rather than present simulation results or simple depiction of reality. Additionally, expensive, often dedicated experiments, of modern, cut-edge technology can be shared worldwide, contributing thus to a high-valued remote laboratory framework.

Many Internet-enabled software systems that employ distance laboratories via simulated, virtual environments can be found in the web [3]. These software systems often integrate many of the desired functionalities, especially from the user's side, such as accompaniments to the experiment documentation, communication support and collaboration among their users. Although modern simulators can accurately estimate circuits' performance, the employment and utilization of real circuits and real instrumentation, for electrical engineering laboratories [4], ensures the measurements' reliability, while at the same time increases the educational value of the remote laboratory and affects positively the psychological mood of the user.

Remote laboratories offering access to real lab experiments and real instrumentation also exist, however the majority of them cannot share their resources simultaneously to many users, thus they fail to serve and support large classes of several hundreds of students.

This paper presents the specifications and the basic structure of an integrated remote laboratory platform that enables the instant remote access to real lab experiments, employing real hardware and real instrumentation [5]. This platform, named Remote Monitored and Controlled Laboratory (RMCLab), is able to provide high-level services to a large number of users for a wide-range of real electrical engineering specific experiments. These experiments can be pre-configured, reconfigurable or customizable, at a very low hardware infrastructure cost.

## 2. Proposed Approach

The basic purpose of the developed platform is to provide high-quality lab training in electrical engineering subjects to students all over the world. The design of such a remote laboratory for real-time, Internet-based lab experiments, should consider all aspects of the system, including communication and data flow, as well as instrumentation and

* Electrical & Computer Engineering Department, University of Patras, Greece; e-mail: efstathiou@ece.upatras.gr, karadimas@ieee.org

*(paper no. 208-0928)*

hardware control [6]. RMCLab has been designed so as to integrate all the capabilities of a physical laboratory to a simple but efficient user interface, along with other sub-systems, that perform lab administration, instrument operation and hardware management.

The primary service that RMCLab should provide to its users is the capability to study on the lab experiment subjects, by affordably accomplishing measurements at any time and from anywhere. For this reason, RMCLab's basic specification is defined as the ability to serve at any time, simultaneously and at real time, any potential user for any available lab experiment. On the other hand, an integrated remote laboratory platform should reinforce the lab administrator's tasks and responsibilities, regarding the experiment setup, hardware and instrumentation control, users' management and also lab maintenance. RMCLab offers also many kinds of assessment functionalities for the students' lab skills, regarding the lab experiments, such as the assignment of several different evaluation criteria (measurements, instrument settings and multi-type questions, etc.), so as the whole platform can be configured as an advanced tool for automated, high-level educational services, an aspect that characterizes the offered educational activities and also our initial motivation.

## 3. Architecture

RMCLab has been developed based on the conventional client–server architecture, expanded in the server-side, as depicted in Fig. 1. It consists of the following basic entities: client, instructor-client (IC), application server (AS), resource server (RS) and lab infrastructure, which includes the real instrumentation and all the hardware modules.

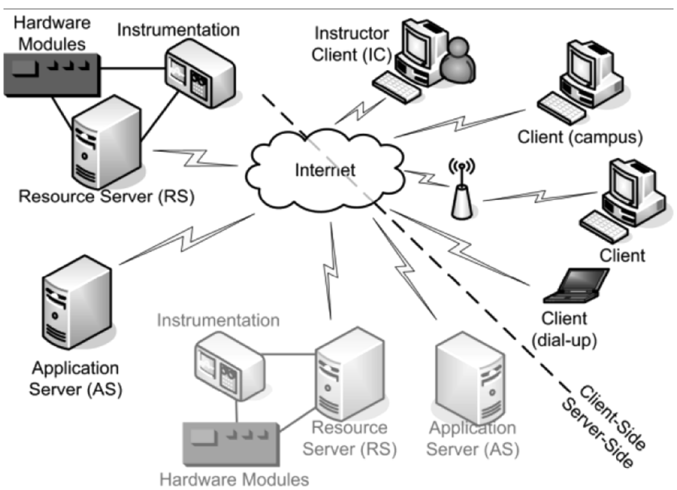



Figure 1. RMCLab system overview.

### 3.1 Network Topology

The server-side of the proposed architecture employs at least two sub-servers: the resource server and the application server. This structure could also be replicated in a more complex network topology. The resource server manages and operates hardware and instrumentation resources, providing to the application server an abstract layer for communication that enables access to the lab infrastructure.

The application server undertakes the data flow control task between clients and the physical remote laboratory, which consists of the resource server and the lab instrumentation. The intermediary role of the application server is mandatory, as system abstraction dictates that clients should not directly communicate with any of the resource servers. The communication between the application server(s) and the resource server(s) or the client(s) is based on a custom, abstract language that integrates all potential tasks of a conventional, physical laboratory.

This topology simplifies the architecture of the server side and expands the platform's capabilities, as it facilitates the robust development and customization of the resource server. Moreover, it enables many application servers to utilize the physical resources (such as lab instrumentation and lab circuits) shared by the resource server(s). As a result, users all over the world are able to transparently access these shared resources via the application server(s). Additionally, the application server grants to its clients transparent access to the real resources of the physical laboratory, thus increasing system's robustness, flexibility and expandability.

### 3.2 Hardware, Instrumentation and Resource Server

The measurement laboratory is based on low cost and low complexity hardware, while it is realized around the resource server. The resource server is equipped with suitable interfaces for the signals of lab experiments (both digital and analog experiments), via a custom LPT-based bus, and the instruments, via the RS232 interface or the General Purpose Interface Bus (GPIB), as depicted in Fig. 2.

Multiple types (standard, programmable, preconfigured or reconfigurable) of analog, digital or mixed circuits can be hosted on the platform's resource server(s). For this reason, RMCLab's hardware is outfitted with a motherboard that is able to host up to 64 cards, where each of them incorporates an FPGA and extra auxiliary circuitry required for implementing the lab circuits, as depicted in Fig. 3.

Each card employs also a PLD, which undertakes card addressing and FPGA configuration. Each of these cards can host eight different analog, digital or mixed (analog and digital) independent arbitrary circuits; as the FPGA is segmented into eight sectors, each of them corresponding to a specific lab experiment. The number of sectors or circuits that a single FPGA can host has been selected taking into account the average complexity of electrical engineering lab experiments and the spare slots of the FPGA. The internal operation of the FPGA is controlled by a register file (Table 1) which is employed within it.

Each sector of the FPGA can host either a specific multi-mode lab experiment or a user's custom circuit. The

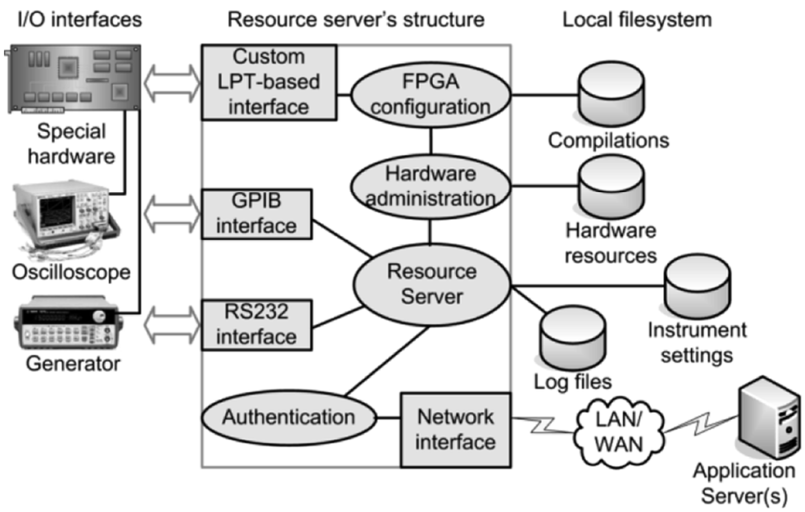



Figure 2. Software and hardware modules of the resource server.

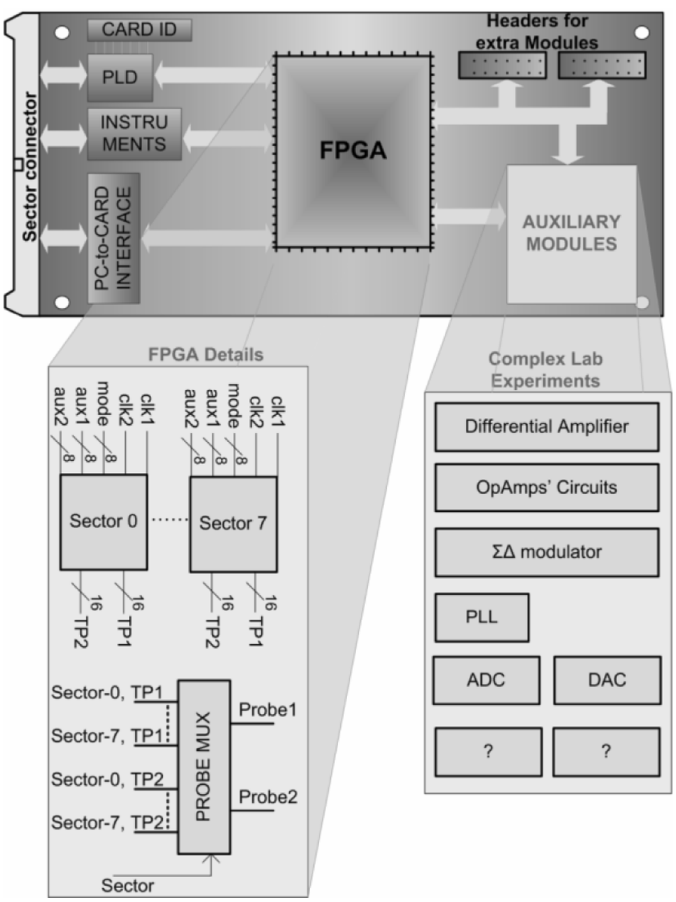



Figure 3. Hardware architecture.

mode register controls the operation mode of a multi-mode lab experiment (e.g., a counter can be either decimal or binary) and the two auxiliary registers allow the user to control the behaviour of lab experiment (e.g., allows the user to assign the modulo of a programmable counter). Finally, the sector register points to the active sector, on which measurements are performed. Therefore, a single sector could implement alike lab circuits, which can be externally presented as different lab experiments, while the selection of the operation is performed by the value of the mode register. For example, in our case, both synchronous and asynchronous digital counters are implemented in the same sector, but they are presented as two different lab experiments. Moreover, when a measurement is carried out, two extra registers, registers Probe1 and Probe2, assign the active nodes of the active sector on which the two probes of the oscilloscope become physically connected through cross point switches. Finally, each card may be offline equipped with additional on board or external circuitry to implement a wide range of more complex electronic circuits, including several types of PLL-based Frequency Synthesizers, several types of D/A or A/D Converters, ΣΔ Modulators, etc.

Table 1
FPGA Register File

| Name | Address | Width (Bits) | Operation |
|---|---|---|---|
| Sector | 0 | 3 | Select the active sector |
| Probe1 | 1 | 4–6 | Select the active nodes of oscilloscope's Ch-A |
| Probe2 | 2 | 4–6 | Select the active nodes of oscilloscope's Ch-B |
| Aux1 | 3 | 8 | Auxiliary register 1 |
| Aux2 | 4 | 8 | Auxiliary register 2 |
| Mode | 5 | 8 | Sector's operation mode |

The aforementioned hardware architecture characteristics in combination with the network topology complexity necessitate an elegant and efficient management of the hardware resources and the measurement requests. A hardware administration module, depicted in Fig. 4, within the resource server, undertakes this management role. A

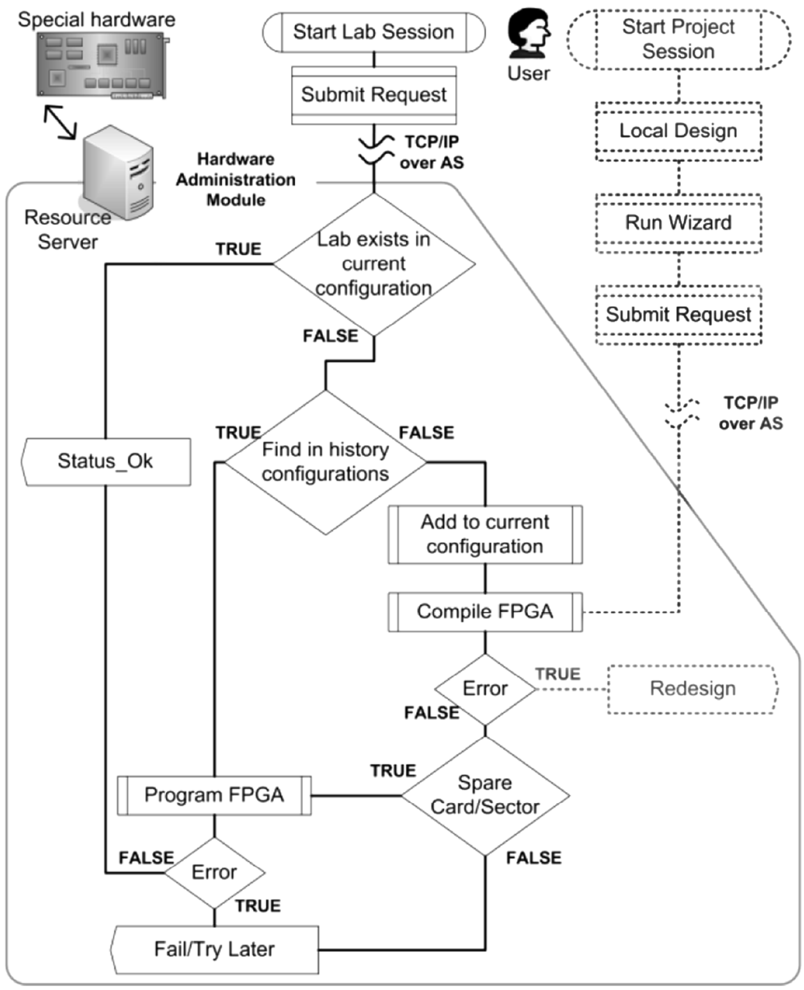



Figure 4. Hardware administration module flowchart.

measurement transaction starts when a client raises a measurement request regarding either a preconfigured lab experiment or a user's custom circuit. The later is discussed in detail in Section 4. After the request is raised by a client to the corresponding application server, it is logged and forwarded via the same application server to the proper resource server which supports the under measurement circuit. Afterwards, the resource server has to accomplish multiple tasks, such as the authentication of the request and the lab infrastructure (hardware and instrumentation) setup, so as to be prepared for the requested operations. This may lead to a real-time, online reconfiguration of the FPGA in a card, so as to implement the requested circuit, or even to the removal of an unutilized sector's circuit, if an empty sector cannot be found. As soon as the hardware is configured, the measurement is performed and the acquired data are transmitted back to the specific client, again via the application server. The above procedure has been designed so as to time-share the lab infrastructure, in a FIFO priority, to all available requests.

### 3.3 Application Server

Apart from the dataflow control and the routing procedure between the client(s) and the resource server(s), the application server is also responsible for the authentication and logging, as well as for the assessment and evaluation of its clients' actions, when educational usage is intended. The above presented characteristics and functionalities of the application server define its architecture, as depicted in Fig. 5. Additionally, each one of the application servers of RMCLab needs to be offline aware of the resource servers, that is able to communicate with, and their list of supported circuits, which are dynamically acquired upon each successful transaction with one of the resource servers.

The application server constitutes also an advanced tool for the development and maintenance of a laboratory class that is available to the administrator of the laboratory, as it can be unrelated to the location of the physical laboratory, realized by the resource server, the hardware and the instrumentation. The development and maintenance of a laboratory class has been merged into a single database system, as depicted in Fig. 6, which contains all the required data for the design and assignment of a lab exercise. Additionally, the same database system includes data regarding the students. One lab exercise may consist of several active images, which correspond to the real lab circuits. For each lab circuit, test-points and active elements (switches and variable components) can be assigned. Hardware properties required for the assignment

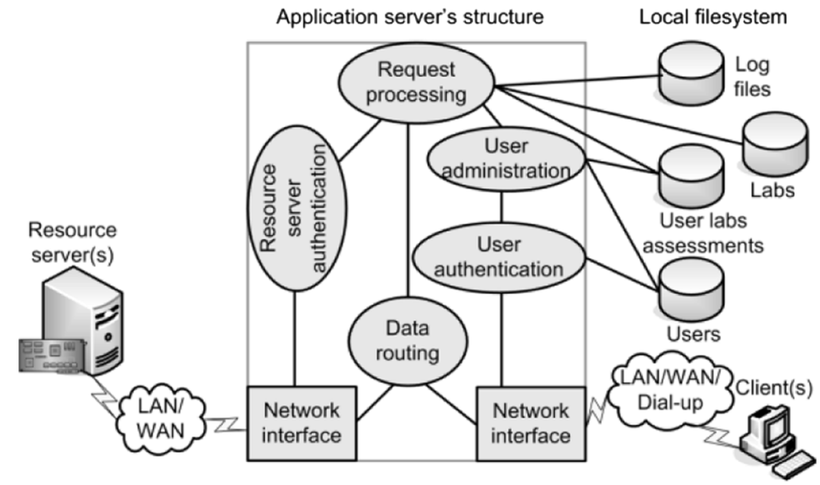



Figure 5. Software modules of the application server.

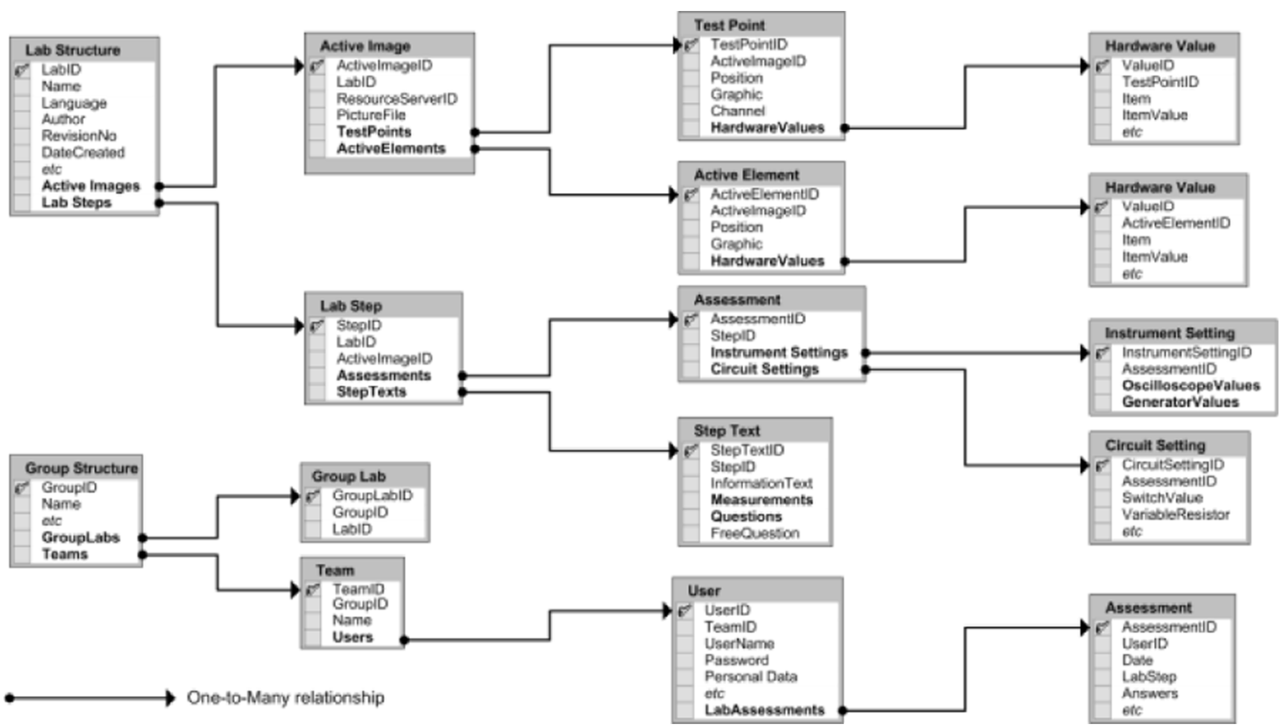



Figure 6. Labs' and users' data structure overview.

of the test-points and the active elements are specified in the custom abstract language, used by the RMCLab. On the other hand, a lab exercise is divided into several steps, where each step may contain information, regarding the theoretical and practical aspects, requests for measurements, multiple choice questions and a free-text question. Moreover, the assessment rules may be provided in each step of a specific lab. Along with the lab experiment's data, students' data regarding their personal information and their assessment are stored in the same database system.

The layered networking of RMCLab permits each laboratory administrator to present a lab experiment, running at a specified resource server, according to his personal educational aspects, regardless of the physical location of the hardware of the lab experiment.

### 3.4 Client

The client-side of RMCLab has been designed so as to comply with the demands of a potential user. Thus, the client module embeds a specific interface, named "scenario interface", for supporting the remote monitor and control of lab infrastructure, and other full functional and user friendly interfaces for lab instrumentation (function generator, oscilloscope, etc.), as depicted in Fig. 7.

In more detail, the scenario interface provides a user with graphic information, related to the under study circuit. Additionally, the interface grants to the user the control of circuit parameters (variable pots, caps, etc.) and also the monitoring of any active node of the circuit, by selecting and setting the probes of the oscilloscope and the

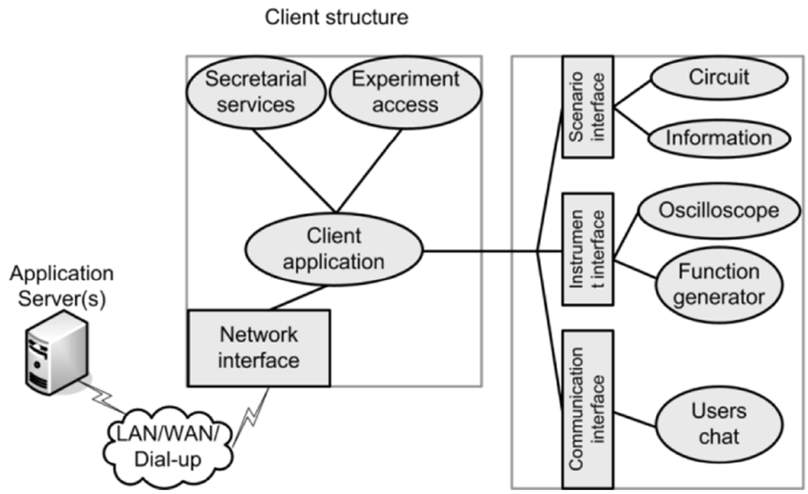



Figure 7. Software modules of the client.

function generator on the circuit. Moreover, extra documentation, regarding the technical and theoretical aspects of the experiment, which can also be separated into multiple steps, can be presented to the user, via the scenario interface.

### 3.5 Communication Module and Instructor Client

To meet the basic requirements of the collaborative interactive e-Learning, a communication interface has been incorporated into the RMCLab. This communication interface consists of a simple chat module enabling the collaboration and the information exchange during a remote lab experiment. The communication interface has been integrated into RMCLab by request of its early users, while the expansion of the chat module with voice and video capabilities is under development.

Finally, RMCLab platform embeds an identical to the user's client module, named the instructor-client. This module offers to a supervisor/instructor of the experiment the ability to replicate, monitor and control any online user's lab environment. This feature is focused on the educational aspect of RMCLab, as it provides an instructor with the ability to closely observe and efficiently tutor the actions of any online users.

### 3.6 Architecture Overview

The described hardware architecture is suitable for developing circuits of low to medium complexity, at a low cost. For accessing the properties of this specific hardware, a software driver has been developed and embedded in the resource server application. Apparently, the platform is able to employ and control any hardware, under the condition that the corresponding software driver enables its access. Thus, even the use of complex or commercially available products is possible.

Additionally, the abstract language used by RMCLab entities to communicate with each other can be modified according to future or different requirements. Fig. 8 depicts the dataflow diagram among the RMCLab entities and focuses on the correspondence between them.

## 4. Advanced Properties of the RMCLab

Conventional lab education is based on the study of predefined lab experiments. RMCLab provides an outstanding benefit to its users, i.e., the feasibility to design and test/measure their own custom circuits under real hardware and real instrumentation. RMCLab users can offline design almost any circuit using a separate software package (MAX+plus II or Quartus, both offered at no-cost for academic institutes from Altera [7]). Using one of the aforementioned specific software packages, one can design his own circuits following a reduced set of rules and confirm by simulation its proper operation. Once the design is verified at the client-side, it can be easily uploaded to the server-side and after a while ($<15\,\mathrm{s}$) the user will be able to perform any measurement on his custom design, which is now implemented on real hardware, by employing real instrumentation. The aforementioned procedure is supported by the hardware administration module of the resource server, as depicted in the dotted part of Fig. 4.

The network architecture of RMCLab enables the world-wide distribution of resources, in terms of lab experiments, by utilizing multiple application servers in a single network topology, as depicted in Fig. 9. Thus, instructors all over the world can take the advantages of employing an existing lab experiment and present it in their native language and personal educational point of view. Obviously, each supervisor has the opportunity to review his users' performance by his own criteria, according to the assessments rules for each experiment, that are defined in the RMCLab's application server, which is available and accessed by the supervisor, as the resource server transparently executes the measurement requests.

The prospects of the RMCLab may hopefully expand world-wide, as the above scenario can be further extended

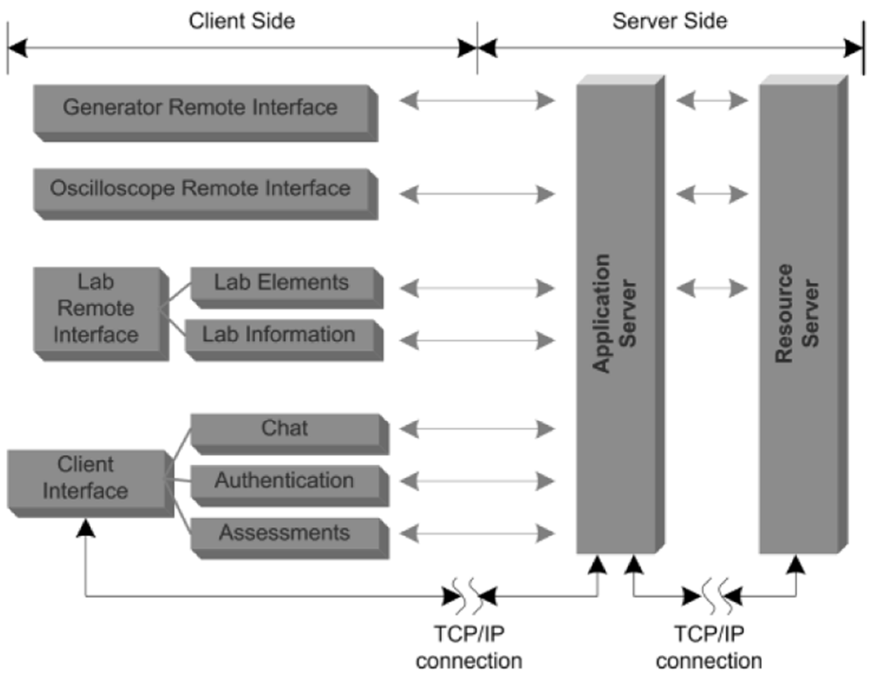



Figure 8. RMCLab's dataflow overview.

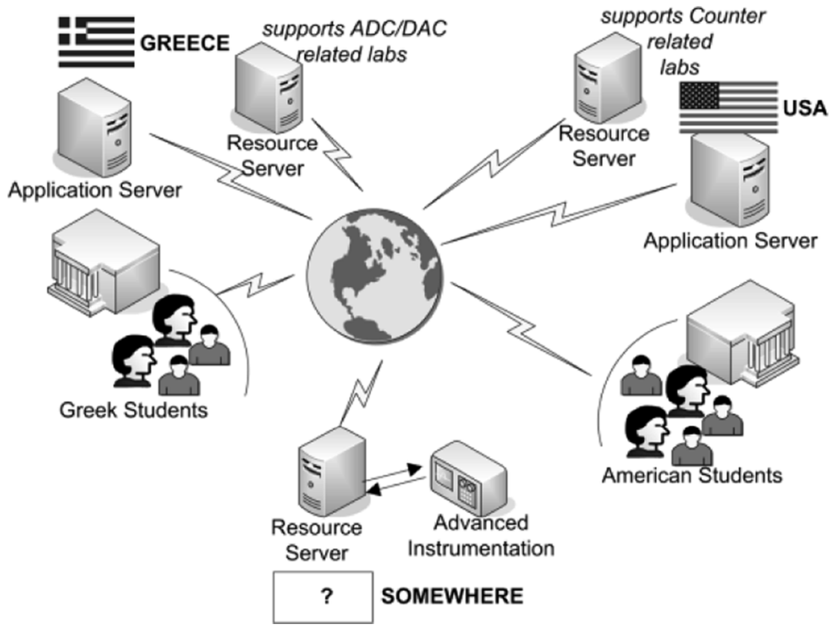



Figure 9. Advanced RMCLab utilization.

if one adds more resource servers, as depicted in Fig. 9. Each resource server can be focused on a specific subject, incorporating the appropriate hardware and instrumentation. Instructors all over the world may take the advantage of using such laboratory resources and develop educational material in their local application servers, in their native language, according to their own educational criteria, so as to offer advanced experimental training to their students, without any requirement for the development and maintenance of any expensive lab infrastructure.

RMCLab's advanced utilization modes are not limited within the above example. The real-time use of real hardware and real instrumentation can significantly contribute to the educational procedure, because it enables an instructor to prepare "Active Lessons" and present in detail, during a class, the operation of a circuit or a system under real world circumstances, while at the same time it can be utilized as a means of demonstration for expensive products.

## 5. The Realized RMCLab System

The architecture described in Section 3 has been implemented at the University of Patras, Greece. The current configuration is a cost effective implementation, consisting of a single PC, with an Intel Hyper-Threading 2.6 GHz processor and 1024 MB RAM, embedding both the resource and the application server of our running RMCLab. This PC, running Windows 2003 Server, is permanently connected to the campus LAN and also to an Agilent 54622D mixed signal oscilloscope and an Agilent 33120A function generator. The oscilloscope is connected with the PC via a high-speed GPIB interface, while the function generator is controlled via RS232 at 19.2 kbps. The PC interface with the hardware modules is implemented based on a custom, low-cost bus, through LPT in EPP mode. Each card of the hardware infrastructure contains an Altera FPGA of the FLEX8K series and also other components required for the implementation of the experimental circuits. The aforementioned infrastructure provides fast enough access and response (<3 s per measurement) to the client requests, as summarized in Table 2.

Table 2
RMCLab's Time Response Characteristics

| Property | Average Delay (s) |
|---|---|
| Hardware setup and measurement time | 3 |
| Compilation time of a custom circuit | 10 |
| Hardware reconfiguration time | 5 |
| Measurement delay from client side using PSTN line at 56 kbps | <5 |

### 5.1 A Simple Educational Paradigm

Fig. 10 illustrates how a lab circuit is presented to students via RMCLab and also the information that is available to the students for the specific circuit. The specific lab circuit regards an Early Decoding, Count Down, 4-bit Decimal Programmable Counter. It should be mentioned that actually this circuit does not contain only a static image, but some "active elements", such as the test-points and two switches, as depicted in Fig. 10. The student is able to interact with these active elements, e.g., he could change the value of the switches or select a different test-point to measure.

Fig. 11 depicts the result of the measurement, as presented in RMCLab, regarding the CLK and LOAD signals of the 4-bit counter, while Fig. 12 depicts the same measurement if acquired directly from the oscilloscope. Obviously, RMCLab is able to take full advantage of the real hardware and real instrumentation utilization, providing measurements, regarding a wide variety of signals' properties, and the full control of both instruments (oscilloscope and function generator).

Figure 11. The measurement result presented in the RMCLab's remote interface.

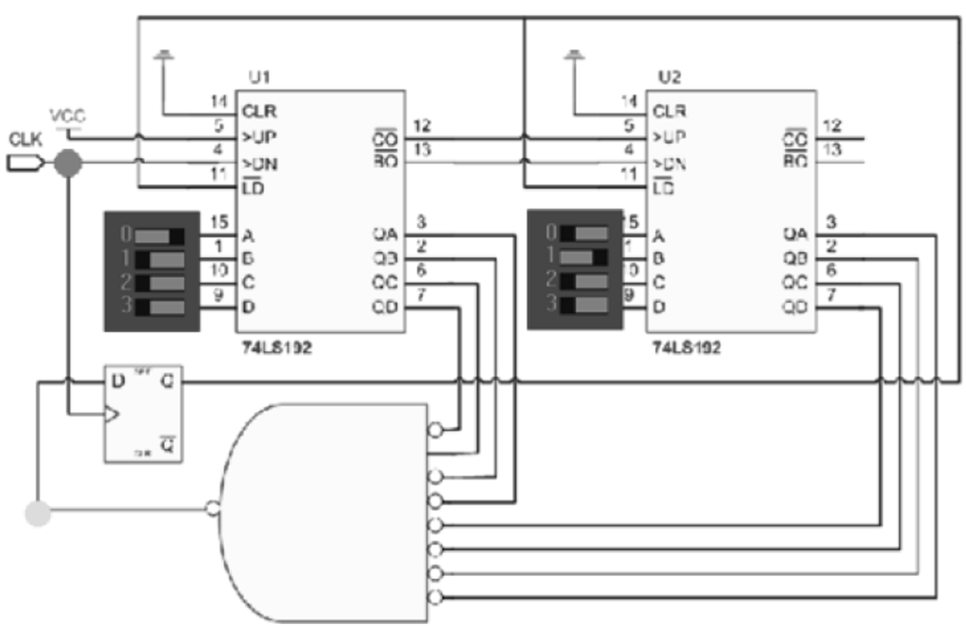



Figure 10. The lab circuit presented via RMCLab, for an Early Decoding, Count Down, 4-bit Decimal Counter.

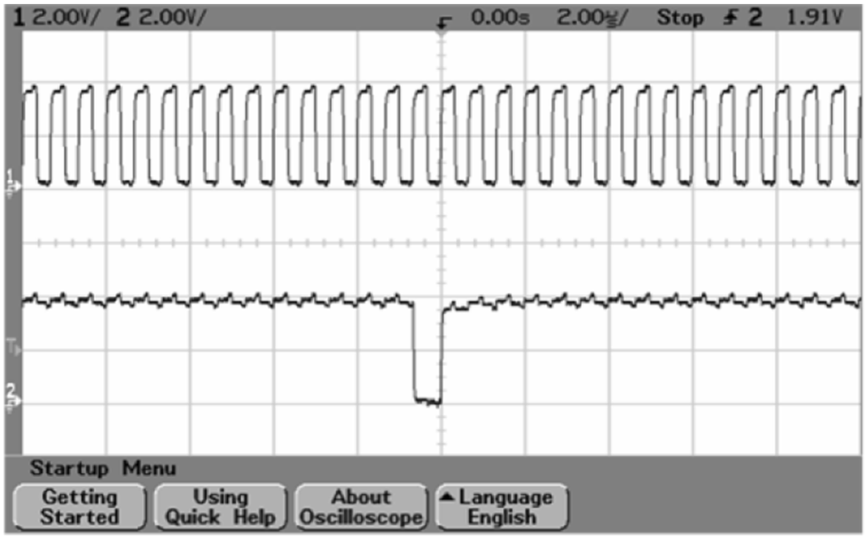



Figure 12. The measurement result acquired directly from the oscilloscope.

Figure 13. RMCLab's cumulative users count versus time of the day for the first and second semester of 2006–07.

### 5.2 Educational Utilization

RMCLab provides its educational services since March 2004 to the Department of Electrical and Computer Engineering of University of Patras, Greece, supporting large classes, in two core lessons, regarding Analog and Digital Electronics. Analog lab experiments include two-stage feedback amplifiers and cascade/folded-cascade amplifiers, whereas digital experiments include a wide variety of topologies regarding counters, adders and accumulators.

Figs. 13 and 14 depict the cumulative user count versus time of the day and the cumulative usage time versus day of week for a single RMCLab-based lab experiment that students had to carry out during the first and second semester of 2006–07 for the lesson Analog and Digital Integrated Circuits, respectively. It is noticeable that RMCLab server has logged 6022 oscillographs and 495 total accesses for the lab experiment of the first semester for all the active teams participated in this lab while 6404 oscillographs and 336 total accesses have been logged for the lab experiment of the second semester, despite the fact that this exercise was significantly more demanding for the students. This implies that students easily acquired experience on the use of the RMCLab platform.

Figure 14. RMCLab's cumulative usage count versus day of the week for the first and second semester of 2006–07.

## 6. Conclusion

RMCLab is able to provide a wide range of high educational services to a great number of students. It increases the productivity of the students by enabling them to have access to the lab infrastructure at non-working hours, while at the same time affects significantly their psychological mood regarding the level of the offered education by their institute.

Moreover, RMCLab accomplishes its services employing a single PC and a single set of hardware and instrumentation, thus pointing out that is able to provide high-quality lab education at low cost, without time consuming human interaction. The structure of RMCLab enables the sharing of hardware and instrumentation resources, thus making possible the extensive exploitation of an expensive lab infrastructure, facilitating the wide spread of remote real lab experiments, which are indisputably valuable for engineers' education. Additionally, hardware reconfigurability permits the remote implementation and measurement of electronic circuits, providing furthermore a high-valued educational service. Finally, services like RMCLab enable students to fulfil their obligations in a reasonable time while exploiting efficiently their widespread working hours.

It is anticipated that the proposed architecture guidelines along with the success of the RMCLab platform will motivate the educational community to cooperate so as to develop an integrated World Wide Lab environment.

## References


[1] H. Boldt, *Didaktisches Design medialer Lernangebote im Umfeld ingenieurwissenschaftlicher Laborpraktika: Didactical Design of media learning offers in the surrounding field of engineer-scientific laboratory practical courses* (Campus Duisburg, Aachen, University Duisburg-Essen, 2005).

[2] A. Bagnasco & A. Scapolla, A grid of remote laboratory for teaching electronics, *2nd International LeGE-WG Workshop on e-Learning and Grid Technologies: A fundamental challenge for Europe*, Paris, 2003.

[3] J. Palop & M. Teruel, Virtual work bench for electronic instrumentation teaching, *IEEE Transactions on Education*, *43* (1), February 2000, 15–18.

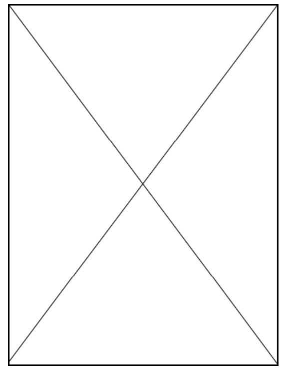
[4] D. Karadimas & K. Efstathiou, An integrated platform, implementing real, remote lab experiments for electrical engineering courses, *Proc. 6th IASTED International Conf. on Web-based Education*, Chamonix, France, 2007.

[5] Online: http://www.apel.ee.upatras.gr/rmclab/

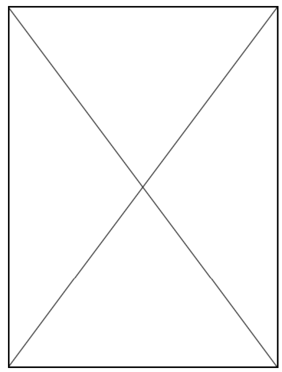
[6] T. Fjeldly & M. Shur, *Lab on the web: Running real electronics experiments via the internet* (Wiley-IEEE Press, 2003).

[7] Online: http://www.altera.com/education/univ/software/unv-software.html


## Biographies



*Efstathiou Konstantinos* received his diploma and Ph.D. degree in electrical engineering from the University of Patras, Greece, in 1981 and 1997, respectively. Since 1981 he is with the Department of Electrical and Computer Engineering at the University of Patras and he is involved in several National and European R&D projects. His interests are in the fields of mixed VLSI design, indirect frequency synthesis, digital signal processing, industrial networking and remote lab education practices and methodology. He has published over thirty conference and journal papers and holds one patent.



*Karadimas Dimitris* was born in Patras, Greece, in 1980. He received his diploma in electrical and computer engineering from the Department of Electrical and Computer Engineering, University of Patras, Greece, in 2003. Since 2003 he is a Ph.D. student in the aforementioned department. His thesis involves mixed analog and digital design, focused on new frequency synthesis techniques. In parallel he is working on a project, under the 3rd Framework of EPEAEK II for the Department of Electrical and Computer Engineering at the University of Patras, Greece, regarding the improvement of the under-graduate curriculum of the aforementioned department, where he has designed and developed a Distance Laboratory Training Platform, named as Remote Monitored and Controlled Laboratory – RMCLab.

Number 99 October 2014

# ERCIM NEWS

www.ercim.eu

## Special theme:

# Software Quality

## 25 years ERCIM: Challenges for ICST

### Also in this issue:

*Joint ERCIM Actions:*
ERCIM 25 Years Celebration

*Keynote:*
The Future of ICT: Blended Life
*by Willem Jonker, CEO EIT ICT Labs*

*Research and Innovation:*
Learning from Neuroscience to Improve Internet Security

*ERCIM News is the magazine of ERCIM. Published quarterly, it reports on joint actions of the ERCIM partners, and aims to reflect the contribution made by ERCIM to the European Community in Information Technology and Applied Mathematics. Through short articles and news items, it provides a forum for the exchange of information between the institutes and also with the wider scientific community. This issue has a circulation of about 6,000 printed copies and is also available online.*

*ERCIM News is published by ERCIM EEIG*
*BP 93, F-06902 Sophia Antipolis Cedex, France*
*Tel: +33 4 9238 5010, E-mail: contact@ercim.eu*
*Director: Jérôme Chailloux*
*ISSN 0926-4981*



***Contributions***
*Contributions should be submitted to the local editor of your country*





***ERCIM News online edition***
*The online edition is published at http://ercim-news.ercim.eu/*

***Subscription***
*Subscribe to ERCIM News by sending an email to en-subscriptions@ercim.eu or by filling out the form at the ERCIM News website: http://ercim-news.ercim.eu/*

***Next issue***
*January 2015, Special theme: Scientific Data Sharing*

*Cover image: an Intel processor wafer.*
*Photo: Intel Corporation.*

## Contents

## RESEARCH AND INNOVATION

This section features news about research activities and innovative developments from European research institutes



## EVENTS, IN BRIEF

### Announcements



### In Brief

comes gained from the relocation rules currently being deployed by most ambulance providers. Currently we are in the field trial phase and are setting up a number of newly developed DAM algorithms, in collaboration with a number of ASPs in the Netherlands. These activities are part of the project 'From Reactive to Proactive Planning of Ambulance Services', partly funded by the Dutch agency Stichting Technologie & Wetenschap.

**Links:**
http://repro.project.cwi.nl
Dutch movie: http://www.wetenschap24.nl/programmas/de-kennis-van-nu-tv/onderwerpen/2014/april/wiskunde-redt-mensenlevens.html

**References:**
[1] M. S. Maxwell et al.: "Approximate dynamic programming for ambulance redeployment", INFORMS Journal on Computing 22 (2) (2010) 266-281
[2] C.J. Jagtenberg, S. Bhulai, R.D. van der Mei: "A polynomial-time method for real-time ambulance redeployment", submitted.
[3] T.C. van Barneveld, S. Bhulai, R.D. van der Mei: "A heuristic method for minimizing average response times in dynamic ambulance management", submitted.

**Please contact:**
Rob van der Mei
CWI, The Netherlands
E-mail: mei@cwi.nl

# An IoT-based Information System Framework towards Organization Agnostic Logistics: The Library Case

by John Gialelis and Dimitrios Karadimas

***SELIDA, a printed materials management system that uses radio frequency identification (RFID), complies with the Web-of-Things concept. It does this by employing object naming based services that are able to provide targeted information regarding RFID-enabled physical objects that are handled in an organization agnostic collaborative environment.***

Radio Frequency Identification (RFID) technology has already revolutionised areas such as logistics (i.e., supply chains), e-health management and the identification and traceability of materials. The challenging concept of RFID-enabled logistics management and information systems is that they use components of the Electronic Product Code (EPC) global network, such as Object Naming Services (ONS) and the EPC Information Services (EPCIS) in order to support the Internet of Things concept.

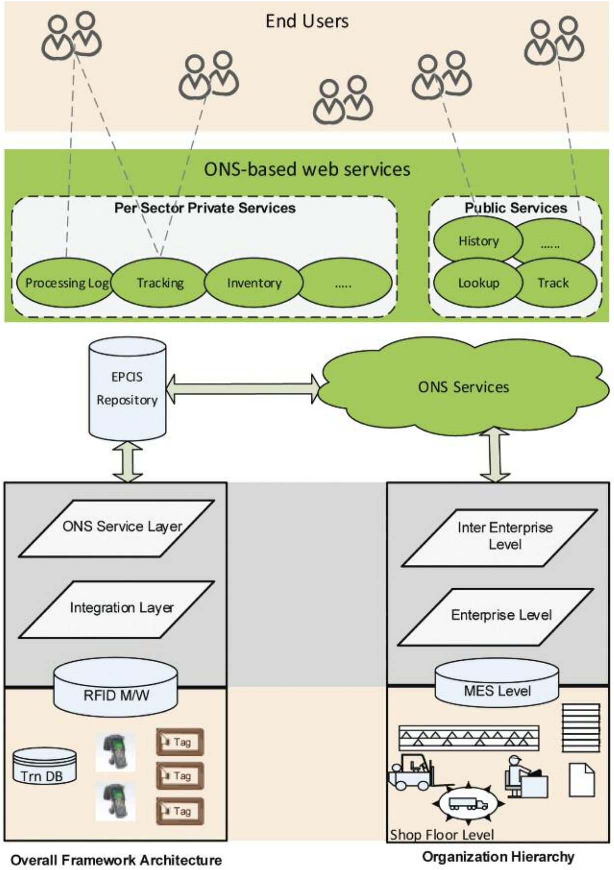



*Figure 1: The architectural framework of SELIDA.*

SELIDA is a joint research project between the Industrial System Institute, the University of Patras (Library and Information Center), the Athens University of Economics and Business, Ergologic S.A and Orasys ID S.A. This project introduces an architectural framework that aims to support as many of the EPC global standards as possible (Figure 1). The project's main goal is the ability to map single physical objects to URIs in order to provide, to all involved organizations in the value chain, various information related to these objects (tracking, status, etc). This is mainly achieved by SELIDA's architectural framework which is able to support as many of the EPC global standards as possible (Figure 1) along with the realization of ONS-based web services available in the cloud. This architectural framework is a value-chain agnostic which relates to:
- the common logistics value-chain;
- the physical documents inter-change value-chain; and
- in demanding cases, the objects inter-change value-chain.

The discovery and tracking service of physical documents that has been implemented exploits both ONS 1.0.1 and EPCIS 1.0.1, in order to allow EPC tagged documents to be mapped to the addresses of arbitrary object management services (OMS), albeit ones with a standardised interface.

The main constituents of the architectural framework are:
- The RFID middleware which is responsible for receiving, analysing processing and propagating the data collected by the RFID readers to the information system which supports the business processes.
- The Integration Layer which seamlessly integrates the EPC related functions to the existing services workflow.

While the existing legacy systems could be altered, such a layer is preferable because of the reliability offered by shop floor legacy systems in general.
- The ONS Resolver which provides secure access to the ONS infrastructure so that its clients can not only query the OMSs related to EPCs (which is the de facto use case for the ONS) but also introduce new OMSs or delete any existing OMSs for the objects.
- The OMS which provides management, tracking and other value added services for the EPC tagged objects. The ONS Resolver maps the OMS to the objects, according to their owner and type, and they should be implemented according to the EPCIS specification (see link below).

The SELIDA architecture has been integrated into KOHA, the existing Integrated Library System used in the University of Patras Library and Information Center. As with all integrated library systems, KOHA supports a variety of workflows and services that accommodate the needs of the Center. The SELIDA scheme focuses on a handful of those services and augments them with additional features. This is generally done by adding, in a transparent way, the additional user interface elements and background processes that are needed for the scheme to work. In order to provide the added EPC functionality to the existing KOHA operations, the integration layer was designed and implemented to seamlessly handle all the extra work, along with the existing service workflow. The SELIDA scheme provides additional functionality to services such as Check Out, Check In, New Record and Delete Record. There are also a number of tracking services that our scheme aims to enhance; these are History, Location and Search/Identify.

The implemented architecture focuses on addressing the issue of empowering the whole framework with a standard specification for object tracking services by utilising an ONS. Thus, the organisations involved are able to act agnostically of their entities, providing them with the ability to resolve EPC tagged objects to arbitrary services in a standardised manner.

**Links:**
KOHA: www.koha.org
ISO RFID Standards: http://rfid.net/basics/186-iso-rfid-standards-a-complete-list
Survey: http://www.rfidjournal.com/articles/view?9168
EPCglobal Object Name Service (ONS) 1.0.1: http://www.gs1.org/gsmp/kc/epcglobal/ons/ons_1_0_1-standard-20080529.pdf
EPCglobal framework standards: http://www.gs1.org/gsmp/kc/epcglobal

**Reference:**
[1] J. Gialelis, et al.: An ONS-based Architecture for Resolving RFID-enabled Objects in Collaborative Environments", IEEE World Congress on Multimedia and Computer Science, WCMCS 2013

**Please contact:**
John Gialelis or Dimitrios Karadimas
Industrial Systems Institute, Patras, Greece
E-mail address: {gialelis,karadimas}@isi.gr

# Lost Container Detection System

by Massimo Cossentino, Marco Bordin and Patrizia Ribino

***Each year thousands of shipping containers fail to arrive at their destinations and the estimated damage arising from this issue is considerable. In the past, a database of lost containers was established but the difficult problem of identifying them in huge parking areas was entrusted to so-called container hunters. We propose a system (and related methods) that aims to automatically retrieve lost containers inside a logistic area using a set of sensors that are placed on cranes working within that area.***

*Figure 1: A common shipping area (by courtesy of Business Korea).*

The Lost Container Detection System (LostCoDeS) [1] is an ICT solution created for avoiding the costly loss of containers inside large storage areas, such as logistic districts or shipping areas. In these kinds of storage areas (Figure 1), several thousand of containers are moved and stacked in dedicated zones (named cells) daily. Nowadays, the position of each stacked container is stored in a specific database that facilitates the later retrieval of this location information. As the movement and management of containers involves many different workers (e.g., crane operators, dockers, administrative personnel, etc.), communication difficulties or simply human distraction can cause the erroneous positioning of containers and/or the incorrect updating of location databases. In large areas that store thousands of containers, such errors often result in containers becoming lost and thus, result in the ensuing difficulties associated with finding them.

At present, to the best of our knowledge, there are no automatic solutions available that are capable of solving this particular problem without the pervasive use of tracking devices. Most of the proposed solutions in the literature address container traceability during transport (either to their destinations or inside logistic districts) by using on-board tracking devices [2] or continuously monitoring the containers with ubiquitous sensors [3], only to name a few.

# Εξ Αποστάσεως Εργαστηριακή Εκπαίδευση Ηλεκτρονικών και Ηλεκτροτεχνίας

Κώστας ΕΥΣΤΑΘΙΟΥ, Δημήτρης ΚΑΡΑΔΗΜΑΣ
Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών
Πανεπιστήμιο Πατρών
efstathiou@ee.upatras.gr, karadimas@apel.ee.upatras.gr

**ΠΕΡΙΛΗΨΗ**

Στην εργασία αυτή παρουσιάζεται ένα ολοκληρωμένο σύστημα εξ Αποστάσεως Εργαστηριακής Εκπαίδευσης, το RMCLab (Remote Monitored & Controlled Laboratory). Το RMCLab καλύπτει τις απαιτήσεις ενός συμβατικού Εργαστηρίου ηλεκτρονικής και ηλεκτροτεχνίας. και δίνει τη δυνατότητα ταυτόχρονης και πραγματικού χρόνου εξυπηρέτησης μεγάλου αριθμού εκπαιδευομένων, ενώ απαιτεί ελάχιστους πόρους σε πάγιες υποδομές και σε προσωπικό, παρέχοντας υψηλού επιπέδου Εργαστηριακή Εκπαίδευση. Το RMCLab χρησιμοποιείται ήδη στα μαθήματα ηλεκτρονικών του τμήματος Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών του Πανεπιστημίου Πατρών και είναι διαθέσιμο μέσω του ιστού στην διεύθυνση: http://www.apel.ee.upatras.gr/rmclab.

**ABSTRACT**

This paper presents RMCLab (Remote Monitored & Controlled Laboratory), which is an integrated system for remote laboratory training. RMCLab accomplishes all requirements that a conventional laboratory of electronics and electrical technology introduces and provides the capability of instant and real time access to multiple users; while at the same time, demands minimum resources of equipment and personnel, providing high level laboratory training. RMCLab is already in use for the curriculum of electronics of the Electrical Engineering and Computer Technology Department at the University of Patras and is also available via the Internet in http://www.apel.ee.upatras.gr/rmclab.

## 1 ΕΙΣΑΓΩΓΗ

Η εργαστηριακή εκπαίδευση στις θετικές επιστήμες έχει καθιερωθεί ως υποχρεωτική και είναι αναμφισβήτητα απαραίτητη διότι δίνει τη δυνατότητα της πρακτικής εφαρμογής της θεωρίας, που είναι άλλωστε και το ζητούμενο από κάθε εκπαιδευόμενο. Όμως, η εργαστηριακή εκπαίδευση είναι σαφώς πιο απαιτητική από την θεωρητική, διότι απαιτεί όχι μόνο ακριβή υποδομή αλλά και πολλαπλάσιο χώρο, όπως και χρόνο διδακτικού προσωπικού, ανά εκπαιδευόμενο.

Η ραγδαία αύξηση του αριθμού των φοιτητών που παρατηρείται τη τελευταία πενταετία, σε συνδυασμό με την αναντίστοιχη αύξηση των εργαστηριακών υποδομών, χώρων και διδακτικού προσωπικού, οδηγεί αναπόφευκτα στη μείωση της ποιότητας της παρεχόμενης Εργαστηριακής Εκπαίδευσης (ΕΕ).

Η εξ Αποστάσεως Εργαστηριακή Εκπαίδευση (εξΑΕΕ) ίσως δεν είναι δυνατή για όλους τους τομείς εκπαίδευσης, όμως, αυτή που αφορά σε εργαστήρια ηλεκτροτεχνίας και ηλεκτρονικής διαπιστώνεται ότι είναι εφικτή. Στον παγκόσμιο ιστό είναι δυνατόν να βρεθούν πολλές λειτουργικές προτάσεις προς την κατεύθυνση αυτή ([Online-1]–[Online-16]), που είναι ενδεικτικές των δυνατοτήτων που υπάρχουν, χωρίς όμως καμία από αυτές να καλύπτει πλήρως τις απαιτήσεις ενός εργαστηριακού μαθήματος ηλεκτρονικών και ηλεκτροτεχνίας.

Στην εργασία αυτή, παρουσιάζονται οι προδιαγραφές και οι κανόνες που πρέπει να διέπουν ένα σύστημα παροχής εξΑΕΕ που ειδικεύεται σε θέματα ηλεκτρονικής και ηλεκτροτεχνίας, έτσι ώστε να δίνει τη δυνατότητα παροχής ποιοτικής ΕΕ σε μεγάλους αριθμούς φοιτητών, μειώνοντας ταυτόχρονα το απαιτούμενο κόστος υποδομών και διδακτικού προσωπικού.

Στο πρώτο τμήμα της εργασίας, παρουσιάζεται η δομή και περιγράφεται η λειτουργικότητα ενός συμβατικού εργαστηρίου ηλεκτρονικής, επισημαίνοντας τις σημαντικότερες διαδικασίες του. Επιπλέον, εντοπίζονται οι αδυναμίες που σχετίζονται με την αύξηση του αριθμού των φοιτητών και την ανεπάρκεια των υποδομών.

Στο δεύτερο τμήμα της εργασίας, παρουσιάζονται οι βασικές οντότητες από τις οποίες αποτελείται το προτεινόμενο σύστημα εξΑΕΕ, αναπτύσσεται η λειτουργικότητά τους, παρατίθενται οι ανάγκες της ΕΕ που καλύπτει και τέλος συγκρίνονται οι επιδόσεις των δύο μορφών εργαστηριακής εκπαίδευσης. Επιπλέον, παρουσιάζονται στατιστικά στοιχεία από την εφαρμογή ενός συστήματος εξΑΕΕ σε δύο εργαστηριακά μαθήματα (Αναλογικά και Ψηφιακά Ηλεκτρονικά) στο Τμήμα Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών του Πανεπιστημίου Πατρών.

Το τελευταίο τμήμα της εργασίας, εστιάζει στις επιπλέον δυνατότητες που μπορεί να προσφέρει η προτεινόμενη εξΑΕΕ και διερευνώνται νέοι τρόποι ανάπτυξης και διάδοσης της επιστημονικής εργαστηριακής εμπειρίας.

## 2 ΣΥΜΒΑΤΙΚΗ ΕΡΓΑΣΤΗΡΙΑΚΗ ΕΚΠΑΙΔΕΥΣΗ

Οι βασικοί στόχοι της ΕΕ ηλεκτρονικής και ηλεκτροτεχνίας είναι η εμπέδωση της θεωρίας και η εξοικείωση του εκπαιδευόμενου με τα υλικά, τις διατάξεις και τα όργανα που αφορούν στο αντικείμενο.

Η συμβολή της ΕΕ ηλεκτρονικής και ηλεκτροτεχνίας στην εμπέδωση της θεωρίας είναι ιδιαίτερα σημαντική διότι ο εκπαιδευόμενος έχει τη δυνατότητα να μελετήσει στην πράξη την λειτουργία και την συμπεριφορά των θεωρητικών κυκλωμάτων. Επιπλέον, του δίδεται η δυνατότητα να εντρυφήσει σε τεχνικές λεπτομέρειες που είναι σημαντικές για τη λειτουργία του κυκλώματος στην πράξη, εμβαθύνοντας με τον τρόπο αυτό ακόμα περισσότερο στην θεωρία.

Η ΕΕ στα ηλεκτρικά και ηλεκτρονικά κυκλώματα απαιτεί τη χρήση πολύπλοκων οργάνων όπως είναι ο παλμογράφος και η γεννήτρια συχνοτήτων καθώς και την εξοικείωση με τις διάφορες μορφές των ηλεκτρονικών στοιχείων (αντιστάσεις, πυκνωτές, ολοκληρωμένα κυκλώματα). Επιπλέον, η κατασκευή, η ρύθμιση και η αποσφαλμάτωση ενός πραγματικού κυκλώματος είναι από τις βασικές και σπουδαιότερες διαδικασίες της ΕΕ. Συνεπώς, η ΕΕ συμβάλλει σημαντικά στην απόκτηση πολύτιμης εμπειρίας.

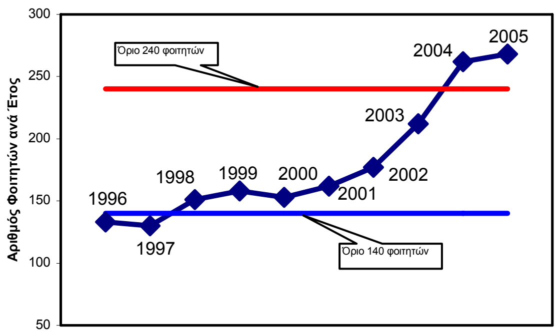
Οι παραπάνω διαδικασίες, που συμβάλλουν στην απόκτηση εμπειρίας, απαιτούν σημαντική προσπάθεια και χρόνο από τους εκπαιδευόμενους, άλλωστε η εμπειρία δεν διδάσκεται αλλά αποκτιέται. Η συμβολή του εκπαιδευτικού προσωπικού στη λύση των πρακτικών προβλημάτων που παρουσιάζονται κατά τη διάρκεια της διεξαγωγής μίας εργαστηριακής άσκησης, συμβάλλει στη μείωση του χρόνου αυτού, αλλά ταυτόχρονα μειώνει την προσλαμβανόμενη, από τον εκπαιδευόμενο, εμπειρία. Η συμμετοχή και η επίβλεψη του εκπαιδευτικού προσωπικού κατά την διεξαγωγή των εργαστηριακών ασκήσεων καθίσταται όμως απαραίτητη, διότι διαφυλάσσει την διαδικασία της αξιολόγησης των εκπαιδευομένων και εγγυάται την σωστή διεξαγωγή της άσκησης και εμπέδωση της θεωρίας.

Η επίτευξη των παραπάνω στόχων εξαρτάται σημαντικά από την αρμονική αναλογία των υλικοτεχνικών υποδομών, του αριθμού του εκπαιδευτικού προσωπικού και φυσικά από τον αριθμό των εκπαιδευομένων.

Έχει αποδειχτεί ότι η συνεργασία δύο–τριών φοιτητών, κατά τη διεξαγωγή μιας εργαστηριακής άσκησης, έχει θετικά αποτελέσματα ως προς τη κατανόηση του αντικειμένου, μειώνοντας ταυτόχρονα τον απαιτούμενο χρόνο. Για το λόγο αυτό, οι φοιτητές χωρίζονται σε ομάδες, συνήθως τριών (3) ατόμων, που πρέπει να συνεργαστούν για τη διεξαγωγή των εργαστηριακών ασκήσεων ενός εξαμηνιαίου μαθήματος (12-13 διδακτικές εβδομάδες). Με δεδομένο ότι ένα εργαστήριο ηλεκτρονικής μπορεί να διαθέτει 10–15 πάγκους εργασίας, εξοπλισμένους με τα κατάλληλα όργανα (παλμογράφο, γεννήτρια συχνοτήτων, τροφοδοτικά κλπ), η κάθε βάρδια, που συνήθως διαρκεί τρεις με τέσσερεις ώρες, μπορεί να εξυπηρετήσει 30–45 φοιτητές. Συνεπώς, διεκπεραιώνοντας 4 βάρδιες ανά εβδομάδα, είναι δυνατό να παρέχεται εργαστηριακή εκπαίδευση σε 120–160 εκπαιδευόμενους, εβδομαδιαίως.

Έχοντας σαν παράδειγμα τα τμήματα των Ηλεκτρολόγων Μηχανικών, στα οποία εισάγονται πάνω από 250 φοιτητές ανά έτος, θα πρέπει, με δεδομένες τις παραπάνω παραδοχές, είτε η κάθε εργαστηριακή άσκηση να διεκπεραιώνεται μέσα σε χρονικό διάστημα δύο εβδομάδων, είτε να αυξηθεί ανάλογα ο αριθμός των εκπαιδευόμενων ανά ομάδα. Και οι δύο αυτές μέθοδοι εξυπηρέτησης μεγάλου αριθμού φοιτητών, οδηγούν αναπόφευκτα σε μείωση της ποιότητας της παρεχόμενης εργαστηριακής εκπαίδευσης, ενώ είναι προφανές, αλλά τις περισσότερες φορές ανέφικτο, ότι η σωστή αντιμετώπιση του προβλήματος είναι ο διπλασιασμός των υλικοτεχνικών υποδομών, καθώς και του εκπαιδευτικού προσωπικού.

Το παραπάνω πρόβλημα εμφανίστηκε την τελευταία τριετία στο εργαστήριο Ηλεκτρονικών Εφαρμογών του Τμήματος Ηλεκτρολόγων Μηχανικών & Τεχνολογίας Υπολογιστών του Πανεπιστημίου Πατρών. Στην Εικόνα 1 παρουσιάζεται ο αριθμός των εκπαιδευόμενων ανά έτος στο εν λόγω τμήμα, κατά την διάρκεια της τελευταίας δεκαετίας.



Εικόνα 1. Κατανομή των εκπαιδευόμενων.

Με στόχο την ανάκτηση της ποιότητας της παρεχόμενης εργαστηριακής εκπαίδευσης, μελετήθηκε, σχεδιάστηκε και αναπτύχθηκε στο παραπάνω εργαστήριο μια εναλλακτική μέθοδος ΕΕ, η οποία, χρησιμοποιώντας τις δυνατότητες της σύγχρονης τεχνολογίας και του Internet, επιτρέπει τον εξ αποστάσεως έλεγχο των οργάνων και κυκλωμάτων ενός πραγματικού εργαστηρίου. Η μέθοδος αυτή εφαρμόστηκε δοκιμαστικά για τρία εξάμηνα σε υποχρεωτικά εργαστηριακά μαθήματα (τάξεις πλέον των 250 φοιτητών), που αφορούν σε αναλογικά και ψηφιακά ηλεκτρονικά.

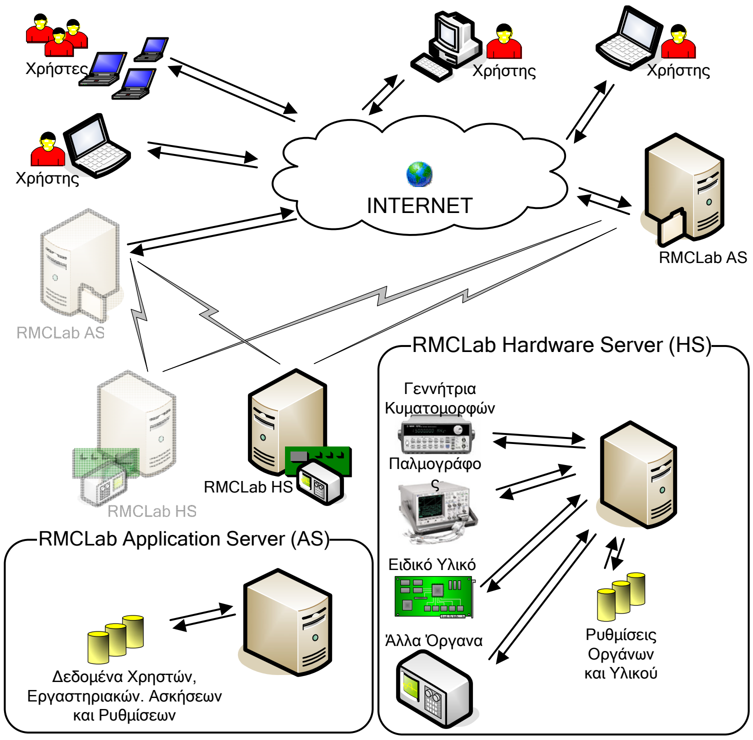
## 3 εξΑΕΕ ΗΛΕΚΤΡΟΝΙΚΩΝ

Για την παροχή της εξΑΕΕ αναπτύχθηκε ένα ιδιαίτερα πολύπλοκο, αλλά καλά δομημένο σύστημα, που ονομάζεται RMCLab (Remote Monitored & Controlled Laboratory). Το RMCLab συνδυάζει τις λειτουργίες εργαστηριακών οργάνων (παλμογράφος, γεννήτρια συχνοτήτων) και ειδικά σχεδιασμένου υλικού, που υλοποιεί τα εργαστηριακά κυκλώματα, με λειτουργίες λογισμικού, έτσι ώστε να παρέχεται μέσω του Διαδικτύου ένα ολοκληρωμένο εργαστηριακό περιβάλλον στον εκπαιδευόμενο, καθώς επίσης και πολλοί βαθμοί ελευθερίας ανάπτυξης και διαχείρισης εργαστηριακών διαδικασιών στον εκπαιδευτή.

Προκειμένου να ικανοποιήσουμε τους βασικούς στόχους της συμβατικής ΕΕ, το RMCLab εξοπλίστηκε με τα παρακάτω κύρια χαρακτηριστικά και δυνατότητες:

- δυνατότητα παρατήρησης και πλήρους ελέγχου πραγματικών κυκλωματικών διατάξεων και εργαστηριακών οργάνων,

- δυνατότητα ταυτόχρονης προσπέλασης πολλών χρηστών σε μία ή περισσότερες εργαστηριακές ασκήσεις, σε πραγματικό χρόνο,
- δυνατότητα αναδιάταξης των κυκλωμάτων των εργαστηριακών ασκήσεων, σε πραγματικό χρόνο,
- δυνατότητα αυτόματης και αξιόπιστης αξιολόγησης των εκπαιδευομένων.



Εικόνα 2. Αρχιτεκτονική του RMCLab.

Στο σημείο αυτό πρέπει να τονιστεί ότι το RMCLab χρησιμοποιεί **μόνο** ένα παλμογράφο, **μόνο** μία γεννήτρια κυματομορφών και ειδικό υλικό στο οποίο υλοποιούνται οι εργαστηριακές ασκήσεις (Εικόνα 2). Η παραπάνω υποδομή έχει τη δυνατότητα να εξυπηρετήσει ταυτόχρονα και σε πραγματικό χρόνο πολλούς χρήστες, δίνοντας σε κάθε έναν από αυτούς την ‘αίσθηση’ της αποκλειστικής χρήσης του εξοπλισμού. Επίσης, το RMCLab παρέχει τις εξΑΕΕ υπηρεσίες υλοποιώντας δύο ανεξάρτητους σταθμούς εξυπηρέτησης, έναν Application Server (AS), που αφορά αποκλειστικά στα δεδομένα που το RMCLab διαχειρίζεται (στοιχεία χρηστών, εργαστηριακών ασκήσεων, κτλ) και έναν Hardware Server (HS), που σχετίζεται με τους πραγματικούς πόρους (εργαστηριακά όργανα, ειδικό υλικό εργαστηριακών ασκήσεων).

Ο βασικός παράγοντας επιτυχίας ενός συστήματος εξΑΕΕ, είναι προφανώς η ευχρηστία και η αποδοχή του από τους χρήστες, ήτοι τους εκπαιδευτές και τους εκπαιδευόμενους. Για το λόγο αυτό, οι προδιαγραφές και οι λειτουργικότητες του RMCLab καθορίστηκαν ανθρωποκεντρικά, με βάση τις ανάγκες των χρηστών του. Το RMCLab χωρίζεται σε δύο βασικές οντότητες, αυτή που αφορά στον εκπαιδευτή και αυτή που αφορά στον εκπαιδευόμενο. Στις ακόλουθες παραγράφους παρουσιάζονται οι δυνατότητες και τα βασικά χαρακτηριστικά των δύο αυτών οντοτήτων.

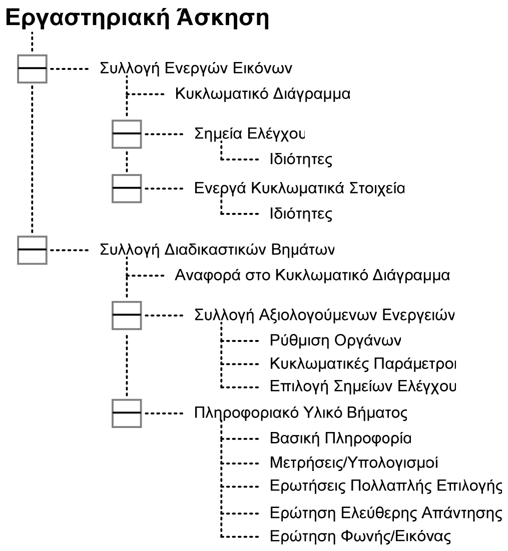
### 3.1 ΔΙΕΠΑΦΗ ΕΚΠΑΙΔΕΥΤΗ-RMCLab

Οι δραστηριότητες που αναπτύσσει ο εκπαιδευτής για την ΕΕ είναι πολλαπλές και συνεπώς η διεπαφή του εκπαιδευτή με το σύστημα εξΑΕΕ θα πρέπει να παρέχει τη δυνατότητα της απρόσκοπτης και σύντομης διεκπαιρέωσής τους. Οι δραστηριότητες αυτές χωρίστηκαν, στα πλαίσια του RMCLab, σε τρεις βασικές κατηγορίες, ήτοι την ανάπτυξη των εργαστηριακών ασκήσεων, τον συντονισμό του εργαστηριακού έργου και την ενεργή εποπτεία των εκπαιδευομένων.

#### 2.1.1 ΑΝΑΠΤΥΞΗ ΕΡΓΑΣΤΗΡΙΑΚΩΝ ΑΣΚΗΣΕΩΝ

Η διαδικασία ανάπτυξης εργαστηριακών ασκήσεων σε ένα σύστημα εξΑΕΕ, όπως είναι και το RMCLab, ενσωματώνει, εκτός από τη συγγραφή των έντυπων σημειώσεων και την προσαρμογή του ειδικού υλικού (κυκλώματα των ασκήσεων), έτσι ώστε να δίνει στους εκπαιδευόμενους τη δυνατότητα πρόσβασης αλλά και ελέγχου του υλικού. Επιπλέον, ο εκ των προτέρων προσεκτικός σχεδιασμός, η μελέτη, η οργάνωση καθώς και ο καθορισμός των κριτηρίων αξιολόγησης σε κάθε εργαστηριακή άσκηση, συμβάλλουν σημαντικά στην ποιότητα της παρεχόμενης εργαστηριακής εκπαίδευσης μέσω ενός συστήματος εξΑΕΕ.

Με βάση τα παραπάνω, μια εργαστηριακή άσκηση στο RMCLab οργανώθηκε όπως δείχνεται στην Εικόνα 3 και αποτελείται από μια συλλογή «ενεργών εικόνων» που αντιπροσωπεύουν τα πραγματικά κυκλώματα της άσκησης και από μία «συλλογή διαδικαστικών βημάτων» που αντιπροσωπεύουν τη ροή της εργαστηριακής άσκησης.



Εικόνα 3. Δομή Εργαστηριακής Άσκησης.

Μία «ενεργή εικόνα» έχει σα στόχο να δώσει τη δυνατότητα στον εκπαιδευόμενο να αλληλεπιδράσει με τη κυκλωματική διάταξη που μελετά, έτσι όπως θα έκανε και στο πραγματικό εργαστήριο. Δηλαδή, σε μια ενεργή εικόνα ο εκπαιδευτής καθορίζει το σύνολο, το είδος, αλλά και το επίπεδο της αλληλεπίδρασης του εκπαιδευόμενου με το πραγματικό κύκλωμα. Έτσι, ο εκπαιδευτής είναι δυνατό να ορίσει σε μία ενεργή εικόνα τα σημεία ελέγχου του κυκλώματος (κόμβοι) και τα στοιχεία (διακόπτες, ποτενσιόμετρα, μεταβλητοί πυκνωτές, κλπ.) των οποίων τις τιμές ο εκπαιδευόμενος θα μπορεί να ρυθμίσει για τον έλεγχο της συμπεριφοράς του κυκλώματος. Οι δυνατότητες αυτές συνηγορούν σε μεγάλο βαθμό στην αυξημένη ποιότητα εργαστηριακής εκπαίδευσης που το RMCLab μπορεί να παρέχει.

Επιπλέον, ο εκπαιδευτής μπορεί να ορίσει σε κάθε εργαστηριακή άσκηση το πλήθος των διαδικαστικών βημάτων που ο κάθε εκπαιδευόμενος θα πρέπει να εκτελέσει, ενώ σε κάθε τέτοιο βήμα ορίζει και το είδος της πληροφορίας που θα διαθέσει στους εκπαιδευόμενους. Συγκεκριμένα, κάθε βήμα μιας εργαστηριακής άσκησης είναι μια παραμετρική οντότητα που περιλαμβάνει το σύνολο πληροφορίας που διατίθεται στους εκπαιδευόμενους. Στο πληροφοριακό αυτό σύνολο συμπεριλαμβάνονται τόσο βασικές πληροφορίες για την ομαλή διεξαγωγή της άσκησης, όπως είναι η σχετική θεωρία και τεχνικές λεπτομέρειες που αφορούν στο χειρισμό του κυκλώματος, όσο και πληροφορίες για τις μετρήσεις που πρέπει να πραγματοποιηθούν. Οι πληροφορίες αυτές συμβάλλουν στην απρόσκοπτη ενασχόληση του εκπαιδευόμενου με το ουσιαστικό περιεχόμενο της άσκησης, απομπλέκοντας τον από οποιαδήποτε μορφή αναζήτησης εξωτερικής πληροφορίας. Επιπρόσθετα, ο εκπαιδευτής είναι σε θέση να συμπεριλάβει σε κάθε βήμα ένα σύνολο από στοιχεία αξιολόγησης, τα οποία μπορεί να είναι ένας συνδυασμός από τα παρακάτω στοιχεία:

- μετρήσεις, ή/και υπολογισμούς που χαρακτηρίζουν τη συμπεριφορά του πραγματικού κυκλώματος,
- ερωτήσεις πολλαπλών επιλογών,
- ερωτήσεις που πρέπει να απαντηθούν με γραπτό ηλεκτρονικό κείμενο,
- χειρισμούς των εργαστηριακών οργάνων και ρυθμίσεις των παραμέτρων του κυκλώματος.

Ειδικά το γεγονός ότι ο εκπαιδευτής μπορεί επιλεκτικά να καθορίσει σε κάθε βήμα της άσκησης ένα σύνολο από συνδυασμούς χειρισμών των εργαστηριακών οργάνων και ρυθμίσεων των παραμέτρων του κυκλώματος, δίνει τη δυνατότητα του πλήρους ελέγχου των διαδικασιών, που ακολούθησε ο εκπαιδευόμενος για την διεξαγωγή της άσκησης, και συνεπώς, της ουσιαστικής και αξιόπιστης αξιολόγησής του. Αυτό ισοδυναμεί πρακτικά στην αντιστοίχηση ενός εκπαιδευτή ανά εκπαιδευόμενο, σε ένα συμβατικό ΕΕ.

### 2.1.2 ΣΥΝΤΟΝΙΣΜΟΣ ΕΡΓΑΣΤΗΡΙΑΚΟΥ ΕΡΓΟΥ

Το RMCLab ενσωματώνει σημαντικές υπηρεσίες γραμματειακής υποστήριξης που συνοπτικά είναι:

- αποστολή/παράδοση εργαστηριακών αναφορών,
- οργάνωση και διαχείριση των στοιχείων των εκπαιδευομένων,
- ανάθεση εργαστηριακών ασκήσεων ανά εκπαιδευόμενο,
- τήρηση στατιστικών χρήσης,
- βαθμολόγηση των επιδόσεων των εκπαιδευόμενων.

Η ανάπτυξη των εργαστηριακών ασκήσεων και ο συντονισμός του εργαστηριακού έργου ενσωματώνονται σε ένα λογισμικό (Εικόνα 4), στο οποίο έχει πρόσβαση ο υπεύθυνος του εργαστηριακού έργου.

Το RMCLab έχει επιπλέον εξοπλίσει τον εκπαιδευτή με κατάλληλο λογισμικό για την ενεργή εποπτεία και παροχή βοήθειας προς τους εκπαιδευόμενους, σε πραγματικό χρόνο. Το λογισμικό αυτό είναι πανομοιότυπο με την διεπαφή του εκπαιδευόμενου, που περιγράφεται παρακάτω, αλλά επιπλέον, προσφέρει την δυνατότητα στον εκπαιδευτή να εποπτεύσει, να παρέμβει επικουρικά και να αξιολογήσει οποιονδήποτε εκπαιδευόμενο είναι online.

## 3.2 ΔΙΕΠΑΦΗ ΕΚΠΑΙΔΕΥΟΜΕΝΟΥ-RMCLab

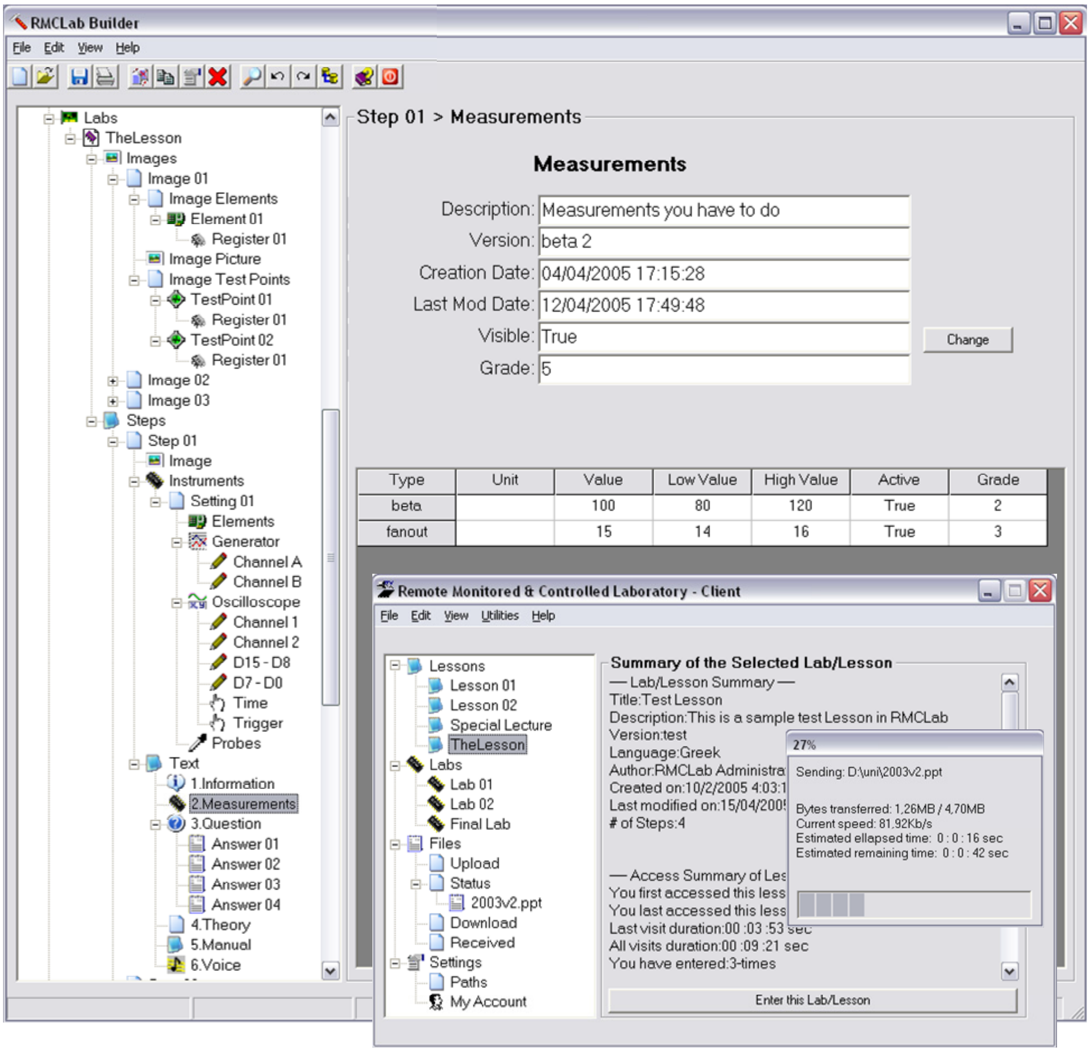
Ο εκπαιδευόμενος θα πρέπει να αντιμετωπίσει ένα φιλικό εργαστηριακό περιβάλλον, παρόμοιο με αυτό το οποίο αντιμετωπίζει στο πραγματικό εργαστήριο, δηλαδή, θα πρέπει να έχει τη δυνατότητα της χρήσης και ρύθμισης των εργαστηριακών οργάνων (παλμογράφος, γεννήτρια συχνοτήτων), της πρόσβασης σε οποιοδήποτε κόμβο του κυκλώματος, καθώς και της ρύθμισης των διαφόρων παραμέτρων αυτού. Επιπλέον, θα πρέπει να παρέχονται με φιλικό τρόπο στον εκπαιδευόμενο οι απαραίτητες πληροφορίες, που αφορούν στην διεκπεραίωση της εργαστηριακής άσκησης, όπως είναι η σχετική θεωρία και τυχόν χρήσιμες συμβουλές για την χρήση των οργάνων και τον χειρισμό του κυκλώματος. Ακόμα, η ύπαρξη της δυνατότητας επικοινωνίας με τους online χρήστες του συστήματος κρίνεται ουσιαστική, διότι έτσι καλύπτεται το κενό που δημιουργεί η έλλειψη της φυσικής αλληλεπίδρασης και επικοινωνίας τόσο μεταξύ εκπαιδευτή-εκπαιδευόμενου όσο και μεταξύ των εκπαιδευομένων. Τέλος, η παροχή υπηρεσιών και διευκολύνσεων γραμματειακής υποστήριξης προς τον εκπαιδευόμενο συμβάλλει θετικά στην αποδοχή της εξΑΕΕ.

Για την εξυπηρέτηση των παραπάνω αναγκών του εκπαιδευόμενου σχεδιάστηκε και υλοποιήθηκε ένα λογισμικό (Εικόνα 4), το οποίο αποτελεί ουσιαστικά τον «Client» του συστήματος εξΑΕΕ και δίνει στον εκπαιδευόμενο τα ακόλουθα εργαλεία λογισμικού.

- **Εισαγωγική διεπαφή χρήστη-εκπαιδευόμενου.** Το τμήμα αυτό της διεπαφής δίνει την δυνατότητα στον εκπαιδευόμενο να διεκπεραιώσει διαδικασίες γραμματειακής υποστήριξης, όπως παράδοση/παραλαβή αναφορών, ενημέρωση των στοιχείων του, να επικοινωνήσει με τους άλλους χρήστες του συστήματος (chat) και να ενημερωθεί για τις εργαστηριακές του υποχρεώσεις και τα διαθέσιμα «ενεργά» μαθήματα.
- **Κεντρική εργαστηριακή διεπαφή.** Η διεπαφή αυτή παρέχει στον εκπαιδευόμενο την τοπολογία του κυκλώματος, δίνοντάς του πρόσβαση σε όλους τους κόμβους του και τη δυνατότητα ρύθμισης των παραμέτρων των στοιχείων του (μεταβλητές αντιστάσεις, διακόπτες, κλπ). Επιπλέον, παρέχει τις πληροφορίες που αφορούν στη θεωρία, στη ροή

της διαδικασίας και στο χειρισμό των οργάνων, καθώς επίσης και στο υλικό αξιολόγησης του εκπαιδευόμενου, όπως επεξεργασία μετρήσεων, ερωτήσεις πολλαπλών επιλογών, καθώς και τα υπόλοιπα στοιχεία που καθορίζει ο εκπαιδευτής και αναφέρθηκαν σε προηγούμενη παράγραφο.

- **Διεπαφή παλμογράφου.** Η διεπαφή αυτή δίνει τη δυνατότητα του πλήρους χειρισμού ενός πραγματικού παλμογράφου. Ειδικά στο RMCLab, η διεπαφή αυτή προσφέρει στον εκπαιδευόμενο τον έλεγχο ενός παλμογράφου Agilent 54622D, που χρησιμοποιείται για την διεξαγωγή των μετρήσεων.
- **Διεπαφή γεννήτριας κυματομορφών.** Η διεπαφή αυτή δίνει τη δυνατότητα του πλήρους χειρισμού μιας πραγματικής γεννήτριας κυματομορφών. Ειδικά στο RMCLab, η διεπαφή αυτή προσφέρει στον εκπαιδευόμενο τον έλεγχο μιας γεννήτριας κυματομορφών Agilent 33120A, που χρησιμοποιείται για την παραγωγή των απαιτούμενων σημάτων εισόδου των κυκλωμάτων.




Εικόνα 4. Στιγμιότυπο από τις διεπαφές Εκπαιδευτή και Εκπαιδευόμενου.

## 4 ΣΤΑΤΙΣΤΙΚΑ ΧΡΗΣΗΣ

Κατά τη διάρκεια της χρήσης του RMCLab διαπιστώθηκε ότι μπορεί απρόσκοπτα να παρέχει εξΑΕΕ σε μεγάλο αριθμό εκπαιδευομένων. Ενδεικτικά αναφέρουμε, ότι σε διάστημα ενός μήνα εντατικής χρήσης του (25/5-25/6 2004), διεκπεραιώθηκαν δύο εργαστηριακές ασκήσεις υποχρεωτικού εργαστηριακού μαθήματος (περίπου 220 φοιτητές) και διακινήθηκαν 129 εργαστηριακές αναφορές. Μέσα στο διάστημα αυτό, η εκμετάλλευση των πόρων του RMCLab από τους χρήστες του δεν ξεπέρασε το 20% των διαθέσιμων, αφού οι ώρες πραγματικής χρήσης του συστήματος ήταν λιγότερες από 150. Ο μέγιστος αριθμός των χρηστών που το σύστημα κλήθηκε να εξυπηρετήσει ταυτόχρονα, δεν ξεπέρασε τους 10, ενώ είναι ενδιαφέρον να τονιστεί ότι η χρήση του συστήματος κορυφώνονταν τις πρώτες πρωινές ώρες (1-3 π.μ.). Με βάση τα παραπάνω, εκτιμάται ότι κατά το διάστημα αυτό, το RMCLab θα μπορούσε να εξυπηρετήσει μέχρι και πέντε φορές περισσότερους χρήστες (περίπου 1000 φοιτητές), χωρίς να παρουσιάσει προβλήματα λόγω υπερβολικού φόρτου.

Ταυτόχρονα διεξήχθει μια έρευνα σε ένα περιορισμένο αριθμό φοιτητών, που είχε σα στόχο τη καταγραφή της αποδοχής και των εντυπώσεων των χρηστών του RMCLab. Η έρευνα αυτή έδειξε ότι το 80% των φοιτητών κρίνει θετικά την πλατφόρμα εργαστηριακής εκπαίδευσης RMCLab, ενώ μόνο ένα μικρό ποσοστό (6%) την απορρίπτει. Το 75% των φοιτητών θεωρεί ότι το σπουδαιότερο πλεονέκτημα του RMCLab είναι το γεγονός ότι του δίνει τη δυνατότητα να εκτελέσει τις εργαστηριακές του υποχρεώσεις «όποτε θέλει, χωρίς χρονικό και χωρικό περιορισμό». Ακόμα, περίπου το 90% των φοιτητών θεωρεί σημαντική την ύπαρξη της δυνατότητας εξΑΕΕ στην τριτοβάθμια εκπαίδευση.

Παρόλα αυτά, πρέπει να τονιστεί ότι το 43% των ερωτηθέντων φοιτητών θεωρεί πως απαιτείται περισσότερος χρόνος για την εκτέλεση μιας εργαστηριακής άσκησης στο RMCLab, απ' ότι σε ένα συμβατικό εργαστήριο. Αυτό οφείλεται στο γεγονός, ότι είχε γίνει εκ των προτέρων γνωστό, πως τόσο η διεξαγωγή όσο και η παράκαμψη ή η αποφυγή μέρους της διαδικασίας των ασκήσεων στο RMCLab, επιτηρείται και αξιολογείται αυτόματα. Αντίθετα, στη συμβατική ΕΕ, κάθε εργαστηριακή άσκηση διαρκεί τρεις ώρες, χρόνος που δεν είναι ικανός για την ενδελεχή αξιολόγηση όλων των εκπαιδευομένων και συχνά ούτε για την πλήρη διεξαγωγή της άσκησης.

## 5 ΕΠΙΠΛΕΟΝ ΔΥΝΑΤΟΤΗΤΕΣ ΚΑΙ ΧΡΗΣΕΙΣ

Η δομή και η οργάνωση του RMCLab επιτρέπει, εκτός από την παροχή βασικής εργαστηριακής εκπαίδευσης, και την ενσωμάτωση νέων, προηγμένων εκπαιδευτικών και επιστημονικών υπηρεσιών.

Η δομή του RMCLab επιτρέπει την εύκολη ανάπτυξη μεγάλου αριθμού εργαστηριακών ασκήσεων. Επειδή ο αριθμός των εργαστηριακών ασκήσεων, που μπορούν να ζητηθούν από τους εκπαιδευόμενους, είναι σαφώς περιορισμένος, το επιπλέον αυτό εκπαιδευτικό υλικό μπορεί να δοθεί στους εκπαιδευόμενους με την μορφή «ενεργών-συνεργατικών» διαλέξεων. Στις διαλέξεις αυτές είναι δυνατό να παρουσιαστεί η θεωρία και η λειτουργία του προς μελέτη κυκλώματος, αποφεύγοντας την απαίτηση στοιχείων (αξιολόγηση, αναφορές), που αφορούν στις εργαστηριακές ασκήσεις. Αυτή η μορφή εκπαίδευσης δίνει τη δυνατότητα στον εκπαιδευτή να παρουσιάσει τη λειτουργία ακόμα και πολύπλοκων κυκλωματικών διατάξεων, διευρύνοντας έτσι τους διδακτικούς ορίζοντες και την ποιότητα της παρεχόμενης εκπαίδευσης.

Ο εκπαιδευόμενος χρησιμοποιώντας τις δυνατότητες αναδιάταξης του ειδικού εργαστηριακού υλικού, είναι σε θέση να υλοποιήσει εξ αποστάσεως σύνθετες κυκλωματικές διατάξεις, που αφορούν είτε σε εξαμηνιαίες (project), είτε σε διπλωματικές, είτε ακόμη και σε πρότυπες ερευνητικές εργασίες, και να πιστοποιήσει την ορθή τους λειτουργία, όπως θα έκανε και σε ένα πραγματικό εργαστήριο. Ακόμα ένα πλεονέκτημα που προκύπτει από τα παραπάνω, είναι το γεγονός ότι η εργασία αυτή μπορεί να παραμείνει διαθέσιμη προς παρουσίαση στο διαδίκτυο.

Η αρχιτεκτονική του RMCLab (Εικόνα 4) επιτρέπει την δημιουργία πλεγμάτων (grids) ([Bagnasco & Scapolla, 2003]) από σταθμούς εξυπηρέτησης, κάτι που δίνει τη δυνατότητα της κατανομής ανεξάρτητων πόρων (όργανα και υλικό) σε διαφορετικές γεωγραφικές θέσεις, επιτρέποντας όμως την ενιαία πρόσβαση σε αυτούς, από οποιαδήποτε σημείο του πλανήτη.

Οι δυνατότητες του RMCLab δεν περιορίζονται μόνο σε εκπαιδευτικές διαδικασίες, αλλά μπορούν να επεκταθούν στην προώθηση ηλεκτρονικών προϊόντων με την online επίδειξη των δυνατοτήτων τους από τους κατασκευαστές.

## 6 ΣΥΜΠΕΡΑΣΜΑΤΑ

Στην εργασία αυτή παρουσιάστηκαν οι βασικές αρχές ενός ολοκληρωμένου συστήματος διαχείρισης και διεξαγωγής εργαστηριακών ασκήσεων ηλεκτρικών και ηλεκτρονικών κυκλωμάτων, του RMCLab.

Η χρήση του RMCLab επιτρέπει τη παροχή ολοκληρωμένης, υψηλής ποιότητας εργαστηριακής εκπαίδευσης, σε μεγάλους αριθμούς εκπαιδευόμενων, έχοντας πολύ μικρές απαιτήσεις τόσο σε εργαστηριακό εξοπλισμό και υποδομή, όσο και σε επιστημονικό προσωπικό, σημαντικά μικρότερες από αυτές που απαιτεί η συμβατική ΕΕ. Συνεπώς, οι επιπλέον πόροι που θα αναλώνονταν για την συμβατική ΕΕ, μπορούν τώρα να δαπανηθούν πιο εποικοδομητικά για την αναβάθμιση της ποσότητας και της ποιότητας της προσφερόμενης ΕΕ.

Η εξΑΕΕ που παρέχει το RMCLab πλεονεκτεί της συμβατικής ΕΕ τόσο στην ποιότητα όσο και στην ποσότητα της παρεχόμενης υπηρεσίας, αλλά μειονεκτεί ως προς την προσλαμβανόμενη από τους εκπαιδευόμενους εμπειρία. Το μειονέκτημα αυτό μπορεί να υπερκεραστεί συμπληρώνοντας την εξΑΕΕ με μια «κατ' οίκον» εξαμηνιαία εργασία (project), που θα απαιτεί την σχεδίαση και κατασκευή υλικού, άμεσα συνδεόμενου με το αντικείμενο του μαθήματος. Όμως, ακόμα και αυτή η εργασία είναι δυνατό να εκπονηθεί μέσω του RMCLab.

Η δοκιμαστική εφαρμογή της εξΑΕΕ σε τρία υποχρεωτικά εργαστηριακά μαθήματα έδειξε ότι γίνεται θετικά αποδεκτή από τους εκπαιδευομένους διότι τους απομπλέκει από τυπικές διαδικασίες (υποχρεωτική φυσική παρουσία σε συγκεκριμένο χώρο και χρόνο), παρέχοντας ταυτόχρονα άνεση χρόνου και καλύτερο περιβάλλον για τη διεκπεραίωση των ασκήσεων. Επιπλέον, διαπιστώθηκε ότι η χρήση της σύγχρονης τεχνολογίας και η παροχή συστημάτων εξΑΕΕ, όπως το RMCLab, συμβάλει σημαντικά στη τόνωση της αίσθησης και της εκτίμησης των φοιτητών για την ποιότητα των σπουδών τους καθώς και στην αύξηση της αναγνωρισιμότητας του εκπαιδευτικού ιδρύματος που την παρέχει.


## ΕΥΧΑΡΙΣΤΙΕΣ

Οι συγγραφείς της εργασίας αυτής αισθάνονται την ανάγκη να ευχαριστήσουν τον καθηγητή και διευθυντή του εργαστηρίου Ηλεκτρονικών Εφαρμογών του τμήματος Ηλεκτρολόγων Μηχανικών και Τεχνολογίας Υπολογιστών του Πανεπιστημίου Πατρών, κ. Γ. Παπαδόπουλο, καθώς και τον καθηγητή κ. Στ. Κουμπιά για την αμέριστη συμπαράσταση και οικονομοτεχνική υποστήριξη που προσέφεραν κατά την διαδικασία ανάπτυξης του RMCLab.


## ΑΝΑΦΟΡΕΣ


[Online-1]. *'A Java-Based Remote, Measurement Laboratory - ReMLab'*, Politecnico di Milano. Διαθέσιμο: http://sermis.etec.polimi.it.

[Online-2]. *'Automated Internet Measurement Laboratory - AIM-Lab',* Rensselaer Polytechnic Institute. Διαθέσιμο: http://nina.ecse.rpi.edu/shur/remote.

[Online-3]. *'Distance Lab'*, Blekinge Institute of Technology. Διαθέσιμο: http://www.its.bth.se/distancelab/english.

[Online-4]. *'Information for WebLab at MIT'*, Massachusetts Institute of Technology. Διαθέσιμο: http://www-mtl.mit.edu/users/alamo/weblab/index.html.

[Online-5]. *'Interactive Systems Laboratory - ISL'*, University of Illinois. Διαθέσιμο: http://iel.isl.uic.edu.
[Online-6]. *'Internet Based Remote Control System - IBRCS'*, Advanced IT Applications and Consultancy. Διαθέσιμο: http://aitac.np.edu.sg/ibrcs.
[Online-7]. *'Internet Lab - ILab', Chalmers University of Technology'*. Διαθέσιμο: http://www.ic.chalmers.se/ilab.
[Online-8]. *'Internet Remote Experimentation - VLAB'*, National University of Singapore. Διαθέσιμο: http://vlab.ee.nus.edu.sg/vlab/index.html.
[Online-9]. '*Internet Shared Instrumentation Laboratory - ISILab'*, University of Genoa. Διαθέσιμο: http://isilab-esng.dibe.unige.it.
[Online-10]. *'Lab-on-Web - LoW'*, University Graduate Center. Διαθέσιμο: http://www.lab-on-web.com/NET/WebApplication/LoW.
[Online-11]. '*Next Generation Lab - NGL - Analog CMOS Course'*, Norwegian University of Science And Technology. Διαθέσιμο: http://www.ngl.fysel.ntnu.no/pages/index.aspx.
[Online-12]. *'Remote experimentation over the Internet'*, Swiss Federal Institue of Technology. Διαθέσιμο: http://iawww.epfl.ch/Staff/Christophe.Salzmann/MS_HTML/ChMS.html.
[Online-13]. *'RETWINE'*, Universidad Autonoma de Madrid. Διαθέσιμο: http://www.retwine.net.
[Online-14]. *'Virtual Laboratory - VirtLab'*, John Hopkins University. Διαθέσιμο: http://www.jhu.edu/%7Evirtlab/virtlab.html.
[Online-15]. *'Web-based Educational framework for Analysis, Visualization, and Experimentation - WEAVE'*, WEAVE Project under DUKE University. Διαθέσιμο: http://weaveserve.egr.duke.edu:8080/weave/index.jsp.
[Online-16]. *'WebLab'*, Massachusetts Institute of Technology. Διαθέσιμο: http://weblab.mit.edu.
Bagnasco A., Scapolla A., (2003), *'A Grid of Remote Laboratory for Teaching Electronics'*, 2nd International LeGE-WG Workshop on e-Learning and Grid Technologies: a fundamental challenge for Europe, March.
Dixon W., Dawson D., Costic B., Queiroz M., (2002), *'A MATLAB-Based Control Systems Laboratory Experience for Undergraduate Students: Toward Standardization and Shared Resources'*, in IEEE Transactions on Education, August 45, pp. 218-226.
Ferrero A., Salicone S., Bonora C., Parmigiani M., (2003), *'ReMLab: A Java-Based Remote, Didactic Measurement Laboratory'*, in IEEE Transactions on Instrumentation and Measurement, June 52, pp. 710-715.
Fjeldly T., Shur M., (2003), *'Lab on the Web: Running Real Electronics Experiments Via the Internet'*, Publisher, John Wiley & Sons Inc.
Ko C., Chen B., Chen J., (2004), *'Creating Web-based Laboratories (Advanced Information and Knowledge Processing)'*, Publisher, Springer.
Mason G., (2002), *'A Handheld Data Acquisition System for Use in an Undergraduate Data Acquisition Course'*, IEEE Transactions on Education, November 45, pp. 388-393.
Notash L., Garriock M., (2003), *'International Virtual Design Studio (IVDS) Program: Overview, Challenges, and Students' Perspectives'*, in IEEE Transactions on Education, August 46, pp. 345-349.
Palop J., Teruel M., (2000), *'Virtual Work Bench for Electronic Instrumentation Teaching'*, in IEEE Transactions on Education, February 43, pp. 15-18.
Ringwood J., Galvin G., (2002), *'Computer-Aided Learning in Artificial Neural Networks'*, in IEEE Transactions on Education, November 45, pp. 380-387.
Sebastian J., Garcia D., Sanchez F., (2003), *'Remote-Access Education Based on Image Acquisition and Processing Through the Internet'*, in IEEE Transactions on Education, February 46, pp. 142-148.