

INFRASTRUCTURE CONSOLIDATION FOR INTERCONNECTED SERVICES IN A SMART CITY USING CLOUD ENVIRONMENT

Jorge F Hernandez¹, Victor M Larios¹, Manuel Avalos¹ and Ignacio Silva-Lepe²

¹Department of Information Systems, CUCEA, UDG Guadalajara, Mexico
jorge.hernandez.mx@ieee.org, victor.m.lariosrosillo@ieee.org,
jmaavalos@mx1.ibm.com

²Thomas J. Watson Research Center, Yorktown Heights,
NY USA Research, New York, USA
isilval@us.ibm.com

ABSTRACT

Sustainability, appropriate use of natural resources and providing a better quality of life for citizens has become a prerequisite to change the traditional concept of a smart city. A smart city needs to use latest generation Information Technologies, IT, and hardware to improve services and data, to offer to create a balanced environment between the ecosystem and inhabitants. This paper analyses the advantages of using a private cloud architecture to share hardware and software resources when it is required. Our case study is Guadalajara, which has seven municipalities and each one monitor's air quality. Each municipality has a set of servers to process information independently and consists of information systems for the transmission and storage of data with other municipalities. We analysed the behaviour of the carbon footprint during the years 1999-2013 and we observed a pattern in each season. Thus our proposal requires municipalities to use a cloud-based solution that allows managing and consolidating infrastructure to minimize maintenance costs and electricity consumption to reduce carbon footprint generated by the city.

KEYWORDS

Smart Cities; Cloud Architectures; Cost Estimation; City Services

1. INTRODUCTION

Improving the services offered by a city and promoting a balance between the environmental sustainability and citizen's quality of life has become an important goal of what we define today as Smart Cities [1]. IT offer a convenient way to connect processes in a city, optimize resources for the benefit of communities and forecast dynamics of the urban environment to better adapt solutions towards the well-being of citizens. However, citizens in smart cities have to deal with the physical and digital dimension.

During the living activities in the urban fabric, inhabitants have a unique identity to access and engage services such as energy, water, communication, and transport, among others. In addition, cities need to offer secure digital platforms for their inhabitants and IT infrastructure becomes vital in terms of communication and processing capabilities and availability. One solution to adapt and scale to the cities services demand is to shift city IT departments to the Cloud Computing paradigm [8].

The Cloud allows grouping various types of hardware and to merge them into a single entity for better and efficient management. Hence, Cloud Computing can work in three categories of services. First, we have the Infrastructure as a Service (IaaS) which provides of virtualization for using hardware resources and this category can offer sensors, storage or processing capacities on demand. Second is the Platform as a Service (PaaS) where users can run Web applications without the complexity of maintaining and running the associated infrastructure, this is critical for e-government service portals. Third, Software as a Service (SaaS) where licenses for critical software in processes such as analytics can be used on demand.

A key aspect of the cloud is the use of virtual machines to achieve its elasticity; a virtual machine is a software application that emulates be a real computer with software and hardware features limited to execute some task. Cloud computing provides a set of principles establishing the rules and principles among suppliers and customers. An important aspect of Cloud is the use of resources based on a “pay as you go” basis, in which the customer must pay for the time that a service, platform or software license is executed/used on a cloud provider.

A service as a process for a Smart City may need hardware, software or a combination of them. Cloud computing proposes benefits of elasticity, resilience, performance, productivity, scalability etc. Hence, this technology offers a better strategy for city governments to manage IT services. This paper is based on the Guadalajara Smart City project selected as IEEE pilot project to share the experience of best practices for smart cities worldwide. Moreover, Guadalajara is not only a city but also a metropolitan area composed of seven interconnected municipalities and we observed and analysed that each municipality has a traditional IT infrastructure consisting of a cluster of servers, routers and intranet access to communicate with other municipalities. In its current state, the data centers on each municipality are isolated infrastructure because they are not interconnected and sharing information and processing capabilities for the metropolitan area.

Our proposal is to consolidate municipal infrastructure by setting up a private cloud for the metropolitan area with the existing infrastructure. The benefits of using a cloud computing architecture allows the acquisition of any hardware configuration (supported by virtualization) in a few minutes. To better understand how Cloud Infrastructure can bring value for Smart Cities, we introduce a Use Case that is based on historical data about pollution in the metropolitan area; an alert system executed in the cloud can inform citizens when bad conditions can expose them to health threats. In Fig 1 we can see the core of the city which is deployed on a typical cloud computing architecture, in which a set of mobile devices or computers are interacting with users continuously to figure out how an alert system for city services can work on the cloud.



Fig. 1 A cloud infrastructure supporting City Services

2. PROBLEM STATEMENT

In the previous section, we discussed the benefits of Cloud Computing in its different layers (IaaS, PaaS, and SaaS) for Smart Cities. We also referred to the case of Guadalajara Smart City looking to shift from traditional IT infrastructure to a cloud computing environment to deal with the city dynamics.

We should mention that an additional important project is the interconnection of all government offices with optical fiber as per the project Connected Mexico, which offers to municipalities the conditions to share their IT infrastructure as a cloud entity. In Mexico, it is possible to process data outside the environment where they occur, i.e., each municipality can analyse data of citizens in another municipality; if and only if the citizen is informed how it will be used his personal data, thanks transparency and access to information law [11].

Current challenges in data centers include identifying the best practices to support a cloud-computing environment. Thus, we propose basic building blocks for this to migrate the traditional IT datacenters to a private cloud as shown in Fig 2. The hardware layer represents the physical resources (routers, computers, switches, hard drives, RAM memory, video cards, etc.) owned by the IT Municipal Datacenters. The second layer is the Virtualization, which enables create virtual machines when a process required it, with its own resources and its own operating system. The third layer has the software tools to complete integrate layers one and two.

For managing virtual machines can use applications such as OpenStack, VMware vSphere, CloudStack, Xen, among others. Government entities usually use open source solutions to minimize licensing costs nonetheless; they could use paid software to manage their virtual machines.

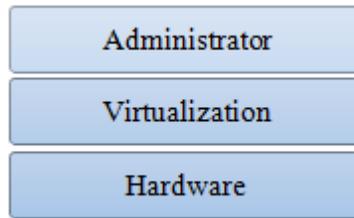


Fig. 2 Layers in a cloud computing environment

We can identify the Administrator as the process of monitoring the behaviour of each virtual machine in the cloud and to perform operations such as: increase or decrease resources hardware or software, delete, or create a new virtual machine. OpenStack Dashboard is an option for this type of module to support a better management. Given the layers in Fig. 2, Fig. 3 shows how the cloud service categories already explained fit into the Smart Cities cycles.

The city deploys sensors and actuators that can be connected to the cloud as an IaaS, offering a global management, security and capacity to scale on demand. In particular, sensors produce data to be curated and stored in the open data city repository. Cloud PaaS is the most indicated to curate and provide storage as well as processing capabilities for analytics. To deal with the complexity of Open Data repositories, specialized software for analytics should be used requiring a SaaS service to use licenses on demand.

One of the key elements of Smart Cities [4] is to break the silos of information among the different government offices offering to share all in a common open data city platform. This action allows avoiding duplicated efforts and investments to understand the city dynamics and provide solutions.

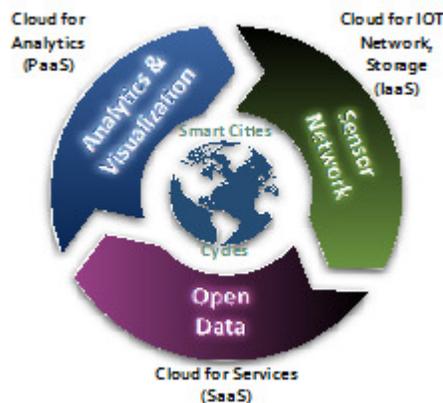


Fig. 3 Smart City Cycles related to the cloud

In order to provide more efficient services, it is possible to correlate information datasets from different indicators. Moreover, a Smart City needs to have a strategy for metrics to understand its performance and where to invest to reach its sustainability.

Metrics for smart cities need to have Key Performance Indicators (KPIs) as well as a subset of indicators for example in Cohen's Wheel there is a section called: Smart Gov, which contains infrastructure, this option could generate alternative metrics such as latency, multitenancy and others. This means that Smart Cities should work with a holistic vision integrating KPIs [5] and indicators to understand city dynamics and decide how the services should adapt. This concept is

based on a systemic approach where a city is a System of Systems or can be modelled and understood as a Complex System.

For this reason, the city should decide how to select KPIs and indicators. Given our work at the IEEE Smart Cities initiative, the model used is shown in Fig. 4 based in the Cohen Wheel. The model proposes five important KPIs related to Smart Economy, Smart Government, Smart People, Smart Living, Smart Mobility and Smart Environment [2][3]. Each KPI has a subset of actions and indicators in a secondary ring. It depends on the amount of sources of information available to feed indicators to be provided by the city, there could be more outer rings, which themselves induce more external rings.

This means that the more the city deploys sensor/actuator networks, the more rings that will appear, resulting in more accurate models to analyse the behaviours and dynamics of the city. That is the reason to have a good cloud strategy in order to scale the KPIs, Indicators and Actions management [6]. Hence, the Smart City project in Guadalajara, following the principles of Metrics based in the Cohen Wheel KPIs, requires an architecture to migrate the metropolitan area of Guadalajara to the cloud. This is the main problem and challenge presented in this paper.

A new issue to introduce is that the metropolitan area of Guadalajara, and for every city that is composed of interconnected municipalities, each one has autonomous infrastructure and budgets. Since all municipalities are interconnected, a challenge is to connect all data centers respecting their autonomy them. A proposed solution is to create a private cloud to support the three types of cloud services. As a use case to create a methodology to estimate the performance and cost of the private cloud integration among the interconnected municipalities, we identified sensors, open data and processing requirements as an example that can be used as reference for all KPIs of the Smart City in Guadalajara.

The sensors are real devices in the city creating datasets of air pollution in various zones of the city. The created datasets are stored in an open web service, and we propose a system that analyses the air pollution data flows to identify harmful pollution levels in zones of the metropolitan area to provide actions for the benefit of citizens (alarms, transport re-routing, etc.). The contribution of this paper is to process and analyse the information produced by an alert system. The system will be fully supported by the Cloud resulting in a consolidation of infrastructure across municipalities in the metropolitan area.

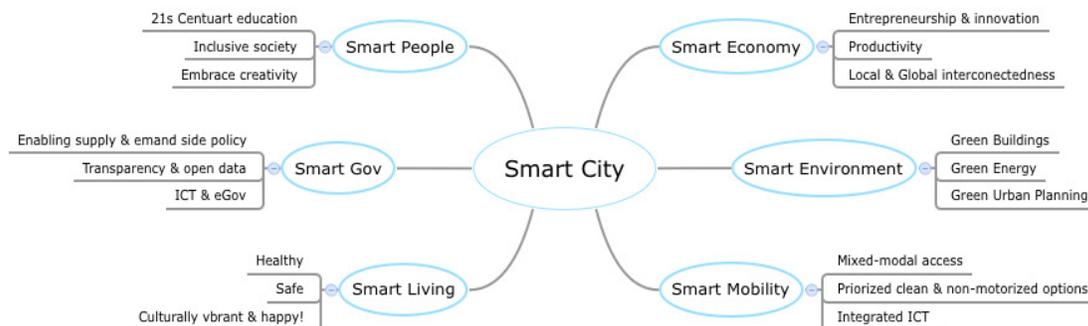


Fig. 4 Representative diagram Cohen's Smart City Wheel

Finally, we propose a methodology to estimate costs of cloud services, which is based in the current municipalities data center infrastructure modelled and extended with a plug in created for the Cloud Simulation Tool Cloudsim [7].

3. METHODOLOGY

We use Java framework to simulate the behaviour of a cloud, CloudSim. CloudSim's goal is to provide a generalized and extensible simulation framework that enables modelling, simulation, and experimentation of emerging Cloud computing infrastructures and application services, allowing its users to focus on specific system design issues that they want to investigate, without getting concerned with the low-level details related to Cloud-based infrastructure and services.

Also, an additional goal of our simulation was to generate an equation to create the actual cost of implementing the service using the cloud, we considered aspects such as payment of electricity, preventive and corrective maintenance, and key and support staff. We decided to group the different environments that are necessities to get the real cost of a specific service. These groups were:

1. *Physical configuration*: it represents the required configuration to execute correctly the service (i.e. hard disk, memory ram, video target, bandwidth, kind of network).
2. *Software configuration*: it refers to the set of programs that the service needs (i.e. operating system, database, file system, simulation programs and parallelism).
3. *Supplies*: items that the provider needs to active the mentioned above services (i.e. electrical power, cables, air conditioning, license fees, space, staff, payment to other providers).

This aggrupation allowed us to understand the elements to be evaluated in each process and we made an equation as follows:

$$C_{2ost}(\mu, \omega) = \sum_{i=1}^{\Psi} (\Omega_i + \phi_i) + \sum_{j=1}^{\mu} \lambda_{ij} + \sum_{k=1}^{\mathbf{Y}} \tau * C_{2ost}(\lambda, \tau) + M(\Omega)$$

where:

ω = a service/process

Ψ = the total amount required resources

Ω = a specific used resource from physical or software configuration

ϕ = fixed cost of used resource finished by the provider

λ = execution time of each resource

μ = execution time needed to complete all the process

\mathbf{Y} = sub process of ω

\mathbf{T} = the total sub process of ω

$M(\Omega)$ = maintenance cost of θ

Using this formula, we can obtain the cost computation of each resource in a specific time. We decided to use a recursive function to recover the used resources of a certain process/service that required a distribution of its job (*parallel tasks*). The primary goal of our formula is used to specify the economic cost in each process to have a log of all the physical resources (hardware) used in a service. Each municipality has its autonomy to decide what kind of software and hardware is needed to accomplish the tasks. With this equation we could also determine the efficiency of a process versus another for each municipality to identify the fastest execution and best low price service for the same service.

4. EXPERIMENTAL FRAMEWORK

Guadalajara is located in western Mexico; it has a current population of about 4,299,000 according to the National Institute of Statistics and Geography (INEGI for its initials in Spanish) [9] in 2008. It is one of the three main cities in Mexico for its economic growth, technological and demographic region.



Fig. 5 Metropolitan area of Guadalajara¹

The city is formed by 9 municipalities: Guadalajara, Zapopan, Tonalá, Tlaquepaque, El Salto, Juanacatlán, Ixtlahuacán de los Membrillos y Tlajomulco de Zuñiga. In Fig. 5, we show the structure the metropolitan area of Guadalajara. Since 1995, the city of Guadalajara has been monitoring air quality 24 hours a day, to make recommendations to care health of its citizens and animals. The Metropolitan Index of Air Quality [10] (IMECA in Spanish) is a Mexican official standard for gauging air quality since 1988. It reports chemicals such as: ozone (O₃), sulfur dioxide (SO₂), nitrogen dioxide (NO₂), carbon monoxide (CO) and particulate less than 10 micrometers (PM₁₀). IMECA is used as a reference for all the states of Mexico to measure environmental pollution. IMECA has a set of ranges for activating notifications, which are shown in the Table 1.

When the range of pollution is above 120 IMECAS, it may cause respiratory problems for some people, mainly children and the elderly. However, if the range is greater than 200 IMECAS a contingency plan is activated: the PRECA (Emergency Response Plan Contingency and

¹ Image taken from: https://commons.wikimedia.org/wiki/File:Mapa_ZMG.svg

Atmospheric Jalisco), which contains various levels, see Table 2. Guadalajara has registered as an open dataset, the information collected about pollution levels from 1996 to 2013, it is in a excel format and files contain data for each metropolitan area and its levels of pollution sensed per day and time (0 - 23 hours).

<i>IMECA</i>	<i>Air quality level</i>
0 – 50	Good
51 – 100	Regular
101 – 150	Bad
151 – 200	Very Bad
>201	Extremely Bad

Table 1. Classification pollution levels

PRECA		
<i>Level</i>	<i>Enabled</i>	<i>Disabled</i>
Pre	Equal or greater 120 IMECAS for 2 consecutive hours	Equal or less than 110 IMECAS for 2 consecutive hours
I	Equal or greater 150 IMECAS for 2 consecutive hours	Equal or less than 140 IMECAS for 2 consecutive hours
II	Equal or greater 200 IMECAS for 2 consecutive hours	Equal or less than 190 IMECAS for 2 consecutive hours
III	Equal or greater 250 IMECAS for 2 consecutive hours	Equal or less than 240 IMECAS for 2 consecutive hours

Table 2. Levels of activation and deactivation of PRECA

Registered contaminants are: O₃, NO, NO₂, NO_x, SO₂, CO, PM₁₀, direction and wind speed and temperature. In our case study, we considered the use of CO as a contaminant, because most Guadalajara inhabitants are in direct contact with it. The CO is an odourless and invisible toxin, thus constant attention is required to monitor its level and to detect when it is higher than environmental health standards in Mexico, to activate contingency plans for the population. The extracted information on the levels of contamination of the ZMG, which is in Excel format, it was migrated to a MySQL database, we normalized the information and we analysed the behaviour of the contamination by zone based on hours, days, months and seasons.

In Figs. 6 to 8, we can observe the behaviour of contamination of three major areas of Guadalajara, where we no increase in pollution levels at the beginning of January, May and December. A decrease in the months of August and September. These factors should be for the rainy season and the reincorporation of students and teachers to their schools. We noted that there is a relationship between the work and school activities with the emission of pollutants, with this premise determined in 3 phases: 1 = low level, 2 = intermediate level and 3 = maximum level. We mapped these levels with an alert system (we did it) to inform to the citizens and to initiate a feedback process to provide recommendations and avoid crossing by this zone.

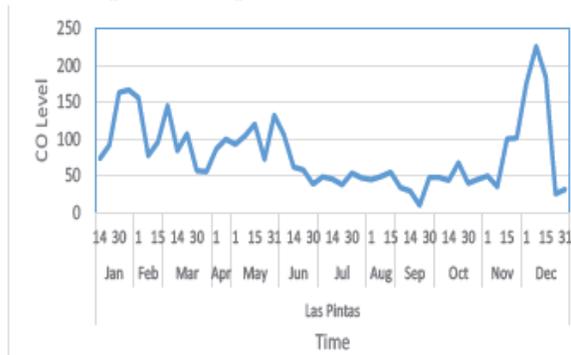


Fig. 6 Co Levels in Las Pintas 2013



Fig. 7 Co Levels in Miravalle 2013



Fig. 8 Co Levels in Tlaquepaque 2013

These levels allow us to identify when a municipality zone needs to share resources among others through the private cloud to manage services. When a level is in phase 1, the municipality may process its information by itself, in phase 2, it could need the assistance of one or more municipalities to improve response time, and the last phase this municipality requests the intervention of others to process the high demand for information to be processed in the shortest possible time. Once the simulation has been performed, we create nine datacenters with hardware features provided by each zone also. The purpose of representing the nine municipalities in a cloud environment, it is to verify that sharing the resources of each infrastructure, improving quality of service to the city, maintenance costs are optimize and the purchase of computer equipment benefits all municipalities and not just one. That is reason, we generated a mechanism to share resources between each area and we built a coordinator agent, which monitors the workload of each area to grant access to other data centers when is necessary.

After we configured a datacenter with the following characteristics: 16 Memory Ram, 4 Tb Hard Drive, Two Xeon X3430 processors and multi-node optical fiber with 100 Mbps transfer rate. This configuration represents the average of resources that a municipality has. The application was executed in two phases: the first, it was using the traditional scheme of each municipality and phase two: it used a cloud-based scheme. In the Fig. 9 we show the workflow, we did for this extraction, interpretation and results obtained.

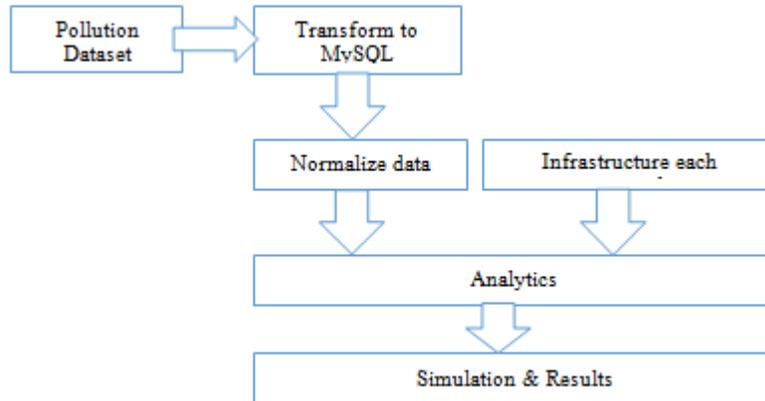


Fig. 9 Workflow process

5. RESULT ANALYSIS

In Fig 10, we show on the left side the number of milliseconds that a service took to execute during a certain time in a traditional scheme, and on the right side the result of our proposal where we suggested use of resources using a private cloud. A considerable decrease is observed and thus electric power consumption was reduced and the cost of use is lower than with other scheme.

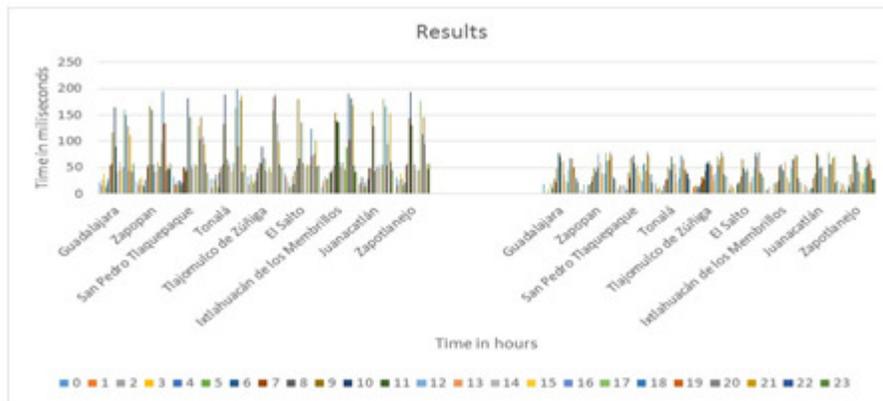


Fig. 10 Traditional IT vs Private Cloud

6. CONCLUDING REMARKS AND PERSPECTIVES

Historical data of pollution levels helped to determine the system of events that should be present in the city to inform and organize services for each zone. For example, when pollution levels are high, systems could report to citizenship through their smartphones of this situation and suggest changes in their routine. Thus begins a process of feedback where the user could request new routes to arrive his destination not passing by contaminated areas.

These kind of services are useful for smart cities. Cloud computing allows enhancing resources and adapt to new hardware equipment that may be present in a data center and to share its resources with other areas. As a future work, we could use docker containers to perform in situ the information that is enclosed within a metropolitan area instead of moving it to another location for processing.

ACKNOWLEDGEMENTS

The work described in this paper was supported by CONACYT through University of Guadalajara in collaboration with Smart Cities Innovation Center.

REFERENCES

- [1] M. Batty, K. W. Axhausen, F. Giannotti, A. Pozdnoukhov, A. Bazzani, M. Wachowicz, G. Ouzounis, and Y. Portugali, "Smart cities of the future," *The European Physical Journal Special Topics*, vol. 214, no. 1, pp. 481–518, Nov. 2012.
- [2] H. Schaffers, N. Komninos, M. Pallot, B. Trousse, M. Nilsson, and A. Oliveira, "Smart Cities and the Future Internet: Towards Cooperation Frameworks for Open Innovation.," *Future Internet Assembly*, vol. 6656, no. 31, pp. 431–446, 2011.
- [3] "Smart city Framework. Guide to establishing strategies for smart cities and communities," BSI British Standards, London.
- [4] N. Komninos, H. Schaffers, and M. Pallot, "Developing a Policy road map for Smart Cities and the future internet," *eChallenges e-2011 ...*, 2011.
- [5] C. Harrison and I. A. Donnelly, "A Theory of Smart Cities," *Proceedings of the 55th Annual Meeting of the ISSS - 2011*, Hull, UK, vol. 55, no. 1, Sep. 2011.
- [6] G. J. Peek and P. Troxler, "City in Transition: Urban Open Innovation Environments as a Radical Innovation," *programm.corp.at*
- [7] Rodrigo N. Calheiros, Rajiv Ranjan, Anton Beloglazov, Cesar A. F. De Rose, and Rajkumar Buyya, "CloudSim: A Toolkit for Modeling and Simulation of Cloud Computing Environments and Evaluation of Resource Provisioning Algorithms, Software: Practice and Experience (SPE)", Volume 41, Number 1, Pages: 23-50, ISSN: 0038-0644, Wiley Press, New York, USA, January, 2011
- [8] Rajkumar Buyya, James Broberg, Andrzej M. Goscinski , "Cloud Computing: Principles and Paradigms", Wiley Editorial, March 2011
- [9] Instituto Nacional de Estadística Geografía e Informática, <http://www.inegi.org.mx/>

[10] Secretaria de Medio Ambiente, <http://siga.jalisco.gob.mx/>

[11] Ley Federal de Transparencia y Acceso a la Información Pública, http://www.diputados.gob.mx/LeyesBiblio/pdf/244_140714.pdf, Junio 2002

AUTHORS

Jorge F. Hernandez, he was born in Guadalajara, Jalisco, in 1975. He received the B.E. degree in computation engineering from the University of Guadalajara, Mexico, in 1998, and the Master in Computer Science in 2000 from Cinvestav. In 2004, he joined the Department of Computer Science, University of Guadalajara, as a teacher. He had published 4 articles in different congress and he had been thesis advisor in bachelor and master level. In 2012, he was admitted in CUCEA to study Postgrade IT and he is working with cloud computing to estimate cost and planning time processing.



Víctor M. Larios, has received his PhD and a DEA (French version of a MS program) in Computer Science at the Technological University of Compiègne, France and a BA in Electronics Engineering at the ITESO University in Guadalajara, Mexico. He works at the University of Guadalajara (UDG) as Professor Researcher and as consultant directed the Guadalajara Ciudad Creativa Digital Science and Technology program during 2013. His research interests are related to distributed systems, visual simulations and smart cities. He is a Senior IEEE member and current chair of the Computer Chapter at the IEEE Guadalajara Section at Region 9. His role in the IEEE-CCD Smart Cities initiative is to lead the working groups.



Manuel Avalos, graduated from Computer Science Engineering from “Universidad Autónoma de Guadalajara (UAG)” back in 1991. In December 2006 Manuel finished his Master Degree on Information Technology by ITESM institution. Manuel is currently a PhD Data Science student at Universidad de Guadalajara campus CUCEA. Manuel joined IBM in 1991 as a Testing Software Engineer and since then Manuel had several Technical, Management and Executive positions in different IBM Divisions, currently he has a Global responsibility for Systems-Storage brand.



Ignacio Silva-Lepe is a Research Staff Member at IBM. His areas of interest include (1) Component Software, designing and building application server infrastructure for distributed components, (2) Distributed Messaging Systems, (3) Advanced Enterprise Middleware, and (4) PaaS Research, designing and building infrastructure for on-boarding and instantiating platform as a service offering onto a compute cloud. Before joining IBM Research, Ignacio was a Member of Technical Staff at Texas Instruments' Corporate Research and Development, subsequently acquired by Raytheon. Prior to that he was a Research Assistant at Northeastern University, where he earned a PhD in Computer Science.

