ICT for Renewable Energy Integration into Smart Buildings: IoT and Big Data Approach

Mohamed Riduan Abid,
Rachid Lghoul
*Alakhawayn University,
Morocco*

Driss Benhaddou
*University of
Houston
Houston, TX, USA*

Abstract— **SG (Smart Grids) are emerging as a very promising technology with buildings as one of its most critical components. SG are meant to leverage energy-efficiency and promote green usage via efficient renewable energy integration.**

SG pose inherent challenges stemming, mainly, from its inter-disciplinary nature. Besides ICT (Information and Communication Technology), SG involves two other broad disciplines: 1. Power systems and 2. Control systems. Still, ICT is the major component that manifests SG "*Smartness*".

In this work, we present the architecture of a real-world smart micro-grid testbed to be deployed in a university campus. We highlight its main components with an emphasis on the ICT component. We are advocating that SG will utilize Internet of Things (IoT) to source all types of data, and thus making the whole system fall within the realm of Big Data. To handle the huge amount of data that will be collected, we advise the deployment of a private local Cloud using open source platform for HPC (High-Performance-Computing) along with Hadoop/MapReduce as the underlying data storage and processing model.

We deem this project as an ideal blueprint that can be easily adopted for similar real-world smart micro-grid testbeds in Africa and around the world.

Keywords—Internet of Things; Wireless Sensor Networks; Big Data; High Performance Computing; Smart Grids.

I. INTRODUCTION

Smart Grids (SGs) [1] are emerging as a very promising technology meant to cope with the energy efficiency and reliability issues by disseminating relevant information in a real-time mode among the different SG components. Such dissemination requires a reliable network that would connect the different SG components.

Building constitute a central component is SGs. They are the dominant consumers of electrical energy, accounting for more than 40% of total electricity use in the US in 2016 (see table. 1) [2]; and it accounts for more than 70% in third-world countries where transportation and industry contributes with minimal electricity consumption shares.

ICT is a key component in SGs. It is the SG component that is responsible for the "*smartness*" feature. The latter is leveraged via a real-time dissemination of information among the different SGs components, e.g., smart meters at the production site (i.e., renewable sources), sensors (measuring electricity consumption) at the consumer site, actuators (controlling electrical appliances by switching them On/Off), the control plane which decides on whether to injecting renewable into the local Grid or send it to a third party (e.g., providers gird or a storage site), etc. These components are Internet of Things (IoT) that source and sink different types of data and control signals. The interconnection of these IoT components needs a reliable network: the Advanced Metering Infrastructure (AMI) [3], and the disseminated data needs to be reliably stored using NoSQL approach.

	Residential	Commercial	Total	Industry	Transp.
2010	22.5%	18.6%	41.1%	30.8%	28.1%
2011	22.4	18.7	41.1	31.3	27.6
2012	22.5	18.8	41.2	30.5	28.2
2013	22.0	18.7	40.7	30.8	28.5
2014	21.7	18.6	40.4	31.2	28.5
2015	21.5	18.6	40.2	31.4	28.4
2016	21.4	18.6	40.0	31.7	28.3

Table. 1. Buildings Share of U.S. Primary Energy Consumption (percentage)

Indeed, SGs data needs NoSQL approach as it falls in the realm of Big Data [4]. SG data meets the 3 Big Data Vs criteria (Volume, Velocity, and Variety). A major part of SG data will emanate from sensors (in buildings) tracking the electricity consumption as well as environmental sensor and crowd sourced data. With the evolution of IoT (Internet of Things) [5], we will have multiple sensors in each building (regardless of whether it is industrial or residential one) thus generating *voluminous* data. Besides, SG data is continuously produced: it never stops. We will be always/infinitely measuring electricity/energy consumption and who is consuming it, thus meeting the *velocity* criteria. Last, SG data will emanate from different sources (temperature sensors, weather stations, consumers, renewable produces, etc), thus bearing *variable* structure

1

SG (big) data need to be processed using HPC (High Performance Computing) capabilities. In this work, we advocate the implementation of a local private Cloud using open source platform (Openstack [6, 7]) along with the Hadoop distributed system [8] and MapReduce [9, 10] as the programming model.

By deploying a private cloud, we assure the optimization of resources usage. Cloud computing [11] is based on *virtualization* as a cornerstone technology. Virtualization promotes *consolidation*, which consists on grouping/consolidating multiple virtual machines (VMs) [12] instances in the same physical hardware, thus optimizing resources usage. On the other hand, Hadoop/Mapreduce is the ideal storage/programming model dealing with SG big data as the latter cannot be stored in a single physical hardware. Instead, it needs to be distributed (i.e., mapped) among different nodes using a distributed file system, e.g., HDFS (Hadoop Data File System). Once mapped, each node carries a designed sub-task thus reducing the original overhead of the main task.

To further leverage energy-efficiency in smart buildings, we promote "context-awareness" whereby the switching *On/Off* of electrical appliances (e.g., electrical heaters, air-conditioners, water heaters, lightning, etc) is based on the "Context", e.g., mainly temperature, humidity and presence, which are provided by sensors connected to the AMI (Advanced Metering Infrastructure).

In this work, we promote a wireless mesh networking (WMN) architecture for AMI deployment using both Zigbee and Wi-Fi. By adopting WMNs [13], sensors/meters/actuators become easy-to-deploy, i.e., plug-and-play. Besides, the AMI becomes self-healing as well: In case a wireless route failure (a link is down because a sensor is turned *Off*), the network autonomously detects the failure and "heals" it via figuring out of another alternative route.

This paper presents the roadmap for a real-world deployment of a smart microgrid at Alakhawayn University campus in Morocco for research purposes. This can be easily adopted as a blue-print for further similar deployment. We present a general architecture of a smart building deployment in the shade of smart microgird integration, and highlight the relevant ICT components. The rest of the paper is organized as follows: Section 2, presents the general architecture. In section 3, we present the Wireless sensor network architecture. Section 4 details the Big Data Processing Platform. Section 5 overviews the context-awareness component and finally, we conclude in Section 6.

II.   GENERAL ARCHITECTURE

The aim of this work is to develop a holistic testbed platform that integrates smart buildings and renewable energy production/storage. Our proposal for smart building deployment involves four main components, see figure 1:

1. *Renewable*: This consists of PV panels which will be deployed on top of parking roofs.
2. *Storage:* These are fuel cells that store excessive energy and feeds it back to the building when needed.

3. *Control plane:* This decides on whether to forward the produced electricity (from renewable) to the building or forward it to the *storage*. The decision making is based on production and storage levels, and the current /expected consumption levels which are fed by the Big Data Analytics platform. The control plane also instructs the storage to inject electricity into the building when needed. In this work, we are advocating the use of compaqRio controllers.
4. The Big Data Analytics Platform (BDAP): This is the main ICT component. This interfaces with the Wireless Sensors Network (WSN) which gets data from different types of sensors (e.g., rooms' temperatures, consumption, production, weather, etc). This data is processed to provide needed information to the *control plane* for decision making. Besides, the BDAP is connected to the *wireless actuators network* as well. This allows the BDAP to issue commands (e.g., switching *On/Off*) to actuators in order to further boost energy-efficiency via context-awareness.
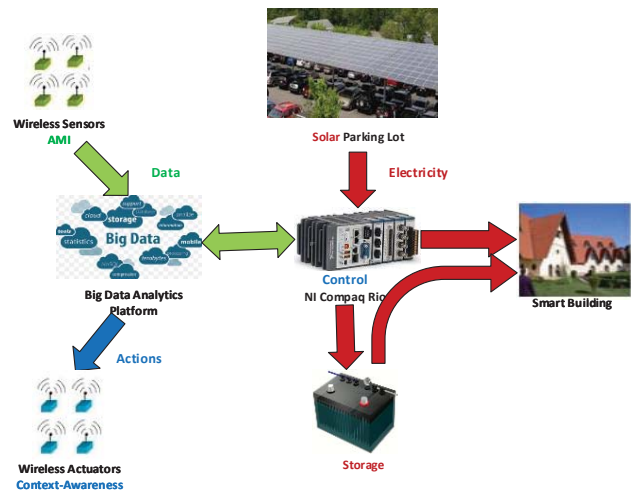

Figure 1. Smart micro-grid testbed General Architecture

The ICT component is made mainly of: 1. Wireless Sensors Network, 2. context-awareness Wireless Actuator Network, and 3. the Big Data Analytics Platform. Next sections presents each of these components.

III.   THE INTERNET OF THINGS

The Internet of Things (IoT) uses Wireless Sensors Network (WSN) [14] as well as other heterogeneous sensors implemented in the CompactRio and are the ICT component responsible for *Data Acquisition*. Its sole role is to sense needed data and forward it to the data processing plane, i.e., the BDAP. To promote ease-of-deployment and self-healing, we are advocating the use of Wireless Mesh Networks (WMN), see figure. 2.

Indeed, due to their easy-to-deploy and self-healing features, WMNs [13] are attracting considerable attention in both academia and industry as well. WMNs are easy-to-deploy as setting a WMN involves minimal wiring and configuration overhead: Placing the WMN nodes and powering them *On* is

2

all that is required to operate a WMN. On the other hand, WMNs are self-healing thanks to the redundancy of wireless links: A failure in a wireless link incites the network to seek an alternative operational link, thus maintaining the network operational. We are deploying the WMN in the floor of a campus building.

With WMNs, each wireless node serves as a data reader and as a router as well, and sensed data is forwarded towards a sink (aka., gateway), thus shaping a tree-like topology with the sink/gateway as the tree root, see figure. 2. The gateway forwards the aggregated data towards the BDAP for processing.

Choosing a suitable tree-based routing algorithm to cater for the "fat-tree" topology whereby the wireless links next to the gateway are witnessing heavier traffic is a must [22]. There are different tree-based routing protocols, e.g., DSR [15], HWMP [16], AODV [17]. We are adopting the AODV routing protocol as this proves ideal for WSN whereby nodes are static, which is the case with smart buildings whereby sensors are fixed in the rooms and halls. In our case, the wireless sensors will sense four types of data:

1. Rooms temperatures
2. Presence in the rooms
3. Weather conditions (temperature, wind, sunny, etc.)
4. Electricity consumption (actual)
5. Electricity production at the renewable site
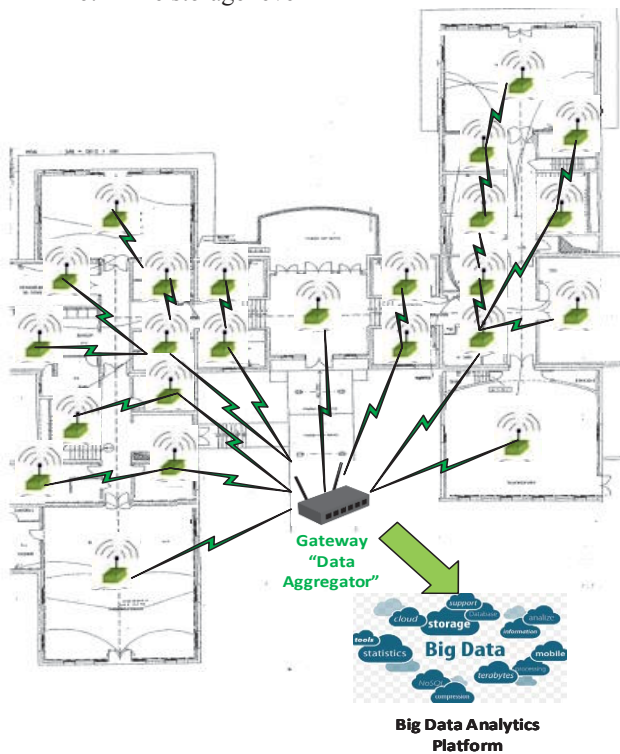6. The storage level

Figure 2. Wireless Mesh Sensor Network for Data Acquisition

IV.   THE BIG DATA ANALYTICS PLATFORM

The Big Data Analytics Platform (BDAP) serves the basis of processing the Data acquired by the sensors and provide needed information to the controller in order to make appropriate control decisions.

At this stage of the work, we did not decide yet on the format of the data as well as the precise information to be requested by the controller. Nevertheless, we present a generic infrastructure of the BDAP that is "*data-format-transparent*".

The BDAP is made of two main components: 1. The BDPA *Back Office*, and 2. The BDAP *Hadoop Cluster*. See figure .3.

The BDAP Back Office is the component that interacts with CompactRio control plane. It stores relevant requests and archives data in an attached database. Mainly, back office implements a middleware that allows interaction between the different components and the underlying hardware, and between the end user and the platform. Via the back office middleware, the end-user can shape and request specific data visualization requests.

The BDPA Hadoop Cluster is the HPC (High-Performance Computing) [18] component of the BDPA. It carries out the distributed storage, and processing, of smart grid big data.

Indeed, since the Data to process in Smart Grids falls within the realm of Big Data, we are adopting a distributed architecture whereby Data is stored in multiple nodes. A single node cannot store Big data. Furthermore, we are advising the adoption of a private Cloud to boost *elasticity* and better usage of the resources.

In fact, there is an inherent, and synergetic, "matching" between SGs and Cloud computing [19]: on one hand, SGs are generating Big Data and are in crucial need of storage and processing power. On the other hand, Cloud computing has been tailored with the main goal of providing compute power (e.g., storage and processing) as a utility. A fact that further eases this "matching" is that Cloud computing masks all the burden of deploying and maintaining the needed IT infrastructure, e.g., recruiting engineers, maintenance, hardware and software purchase, etc.

In this work, we advise the use of an HPCaaS (HPC as a Service) [20] Cloud Computing Platform for SG Big Data processing (For a detailed step-by-step blue-print for how to deploy HPCaaS in a private Cloud refer to [21]).

Our proposed HPCaaS uses Hadoop as the clustering platform to store and process data. This uses HDFS (Hadoop Distributed File System) for data storage and the MapReduce programming model for distributed data processing, and both run on top of a Cloud computing platform: the Openstack platform. The Big Data generated by SGs (mainly generated by the SG sensors and meters) falls in the category of K-V (Key-value) pairs (e.g., sensor id, timestamp, consumption levels, etc).

HDFS manages storage and related issues (e.g., chunks replication, failure recovery), and the Hadoop MapReduce runs the relevant "jobs" on selected chunks of the big data. The proposed architecture is depicted in Fig. 3.
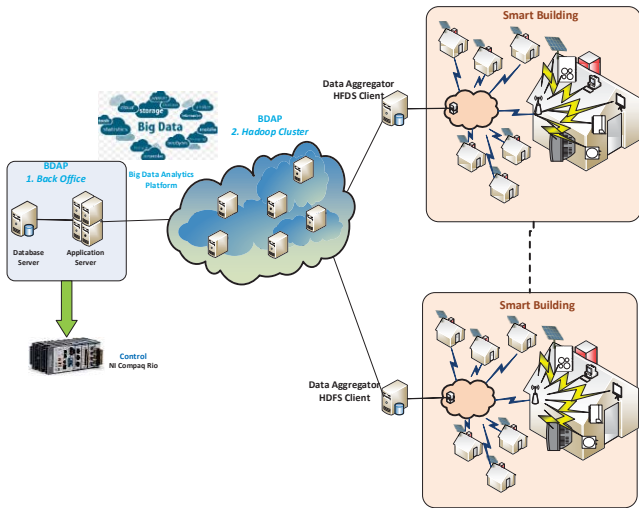
3

Figure. 3 The BDAP Interface

For a scalable deployment of our architecture, each building with has a Data Aggregator that "aggregates" data from the different sensors and forwards to the BDAP where the Hadoop cluster lies. To do so, the Data aggregator will implement and HDFS client that will allow him to communicate with the HDFS server in the cluster via an IPC (Inter-process Communication) protocol, see figure. 4.
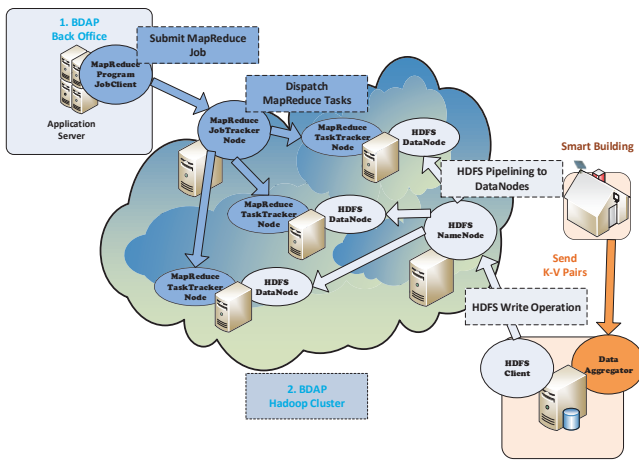

Figure. 4 The Data Flow Architecture

Upon receiving a request (e.g., from an end-user requesting a data Visualization snapshot), the BDAP Back Office *forks a* MapReduce *Job-Client* process and submits a MapReduce *Job*. The latter instantiates a MapReduce *Tracker* in the Hadoop Cluster which "divides/*maps*" the job into multiples tasks (*TaskTrackers*), see figure. 4. Each of these tasktrackers processes runs on a piece of data which is stored in an HDFS Data node.

The HDFS files system divides its data files into multiple chunks (each of size 64 MB) and distribute them among different *DataNodes*. The HDFS *NameNode* is the master process that manages the location of files, and keeps relevant

files metadata and namespace entries. Datanodes are (the real workhorses of the system. They store and retrieve files' blocks/chunks upon request from the master node (i.e., the NameNode).

A key component in collecting data from the sensors sending control back to the actuators is the context-aware middleware. This lies within the BDAP back office.

The primary function of the middleware is to integrate heterogeneous IoT elements without the complexity of the underlying data link layer they bring with them. An IoT element will be need just to be plugged-in without complicated development of new drivers and interfaces that will make the element integrate seamlessly. Heterogeneity, which is inherent into these application as we will have sensors of different types, induces complexity in configuration and integration as well as data collecting and processing. Indeed, various heterogeneous components that will generate diverse data types that will be processed by a middleware system to come out with appropriate information meant for smart grid applications. In addition to configuration, integration, and data collection, the middleware will provide followings functionalities:

1. Collect data and generate contextual information useable by application software
2. Data processing including filtering of redundant data received from sensors and other front end devices
3. Transformation and archival in a database of sensor data into an understandable computer machine format
4. Providing different processing services for energy management applications
5. Providing security services from end to end perspective
6. Provide subscription capabilities for different energy management functions

V. CONTEXT-AWARENESS

Context-awareness is crucial for leveraging energy-efficiency in smart buildings. In this project, we are deploying wireless sensors that track two main "context" aspects: 1. Temperature/Humidity and 2. Motion.

Alakhawayn University in Ifrane (AUI), where the testbed will be deployed, is located in the mid-atlas Moroccan mountains and witnesses a very cold weather season with long periods of snow. During these periods, some of the students (and staff) use electrical heaters in a non-responsible manner, e.g., by setting the heaters' thermostats to the maximum, and thus they will never switch *off* even if none is present in the room. This fosters a non-energy-efficient environment where electrical energy is unnecessarily wasted. In this project, we are deploying a wireless sensor that will communicate the real-time temperatures in rooms/offices. By setting a temperature threshold, the BDAP back office will send commands to the wireless actuators to switch *On/Off* electrical heaters. Besides, we are deploying motion detectors as well: When no motion is detected, for quite large periods (e.g., an hour), a fact reflecting that none is around, the temperature threshold will be decreased to a lower value in order to cope for the situations

4

where students/staff left their rooms for long period. In this context, every electrical heater will be equipped with a wireless actuator that will receive signals from the BDAP over a wireless network.

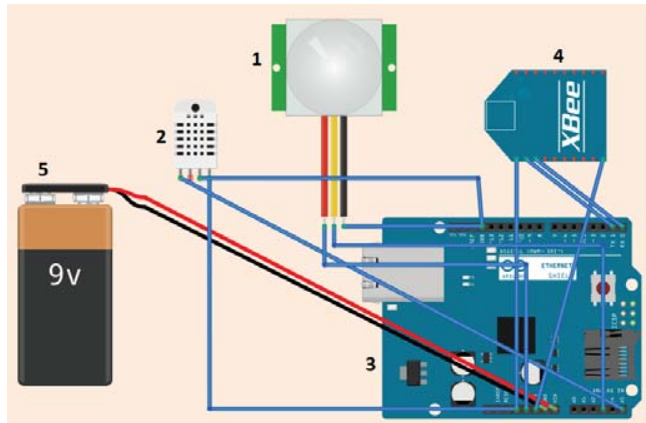To deploy our own sensors, we are using Arduino cards See Figure. 5.

Figure. 5: The Architecture of the Proposed Context-Aware Wireless Sensor (Data Acquisition)

The wireless sensor is made of five elements:

1. The PIR motion detector which operates under 3.3 V DC and outputs a digital signal indicating either the presence or the absence of the motion.
2. The Temperature and Humidity sensor DHT 22 operating under 3.3-5.5 V DC and outputs an analogue signal,
3. The Arduino Uno board with an ATmega328V microcontroller and a 16 Mhz clock speed. This board operates under 7V to 12V,
4. The wireless antenna is the Xbee ZNET 2.5 radio frequency module that operates under 3.3 V. This XBee interface is compatible with the Zigbee protocol (IEEE 802.15.4) and has 20 pins where two of them are reserved for receiving and transmitting data
5. 9V battery.

On the other hand, the wireless actuator is currently deployed using Raspberry PIs. This is made of 5 elements, see figure. 6:

1. 1. The Processing unit (Raspberry Pi) that uses Raspbian as the underlying operating system. This runs a listener process that intercepts incoming request from the BDAP via the WSN in order to put a pin's status either up or down, which corresponds to a switching On/Off command. The latter is forwarded to the electrical appliance via a relay that is connected to one of the pins of the raspberry Pi.
2. The transistor is used to elevate the output voltage of the raspberry Pi's pin which is 3V to 5V. We need this elevation in order to be able to act upon the relay.

3. The relay is the switching point that receives the output voltage being elevated from the Raspberry Pi's output pin.
4. a power source of 240 V is used to power the electrical appliance
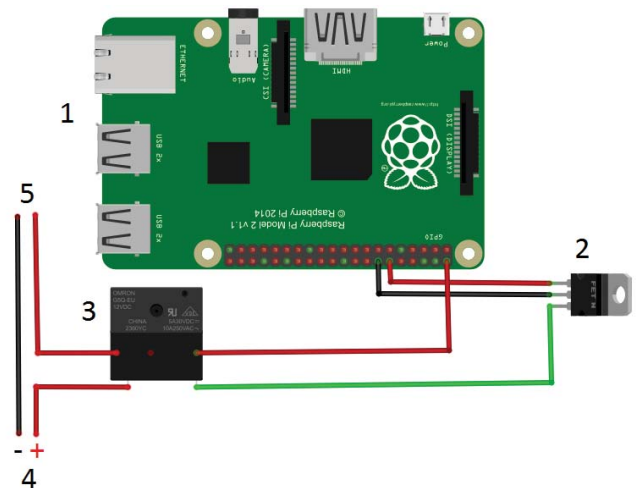5. The electrical entry to the appliance to switch *On/Off.*

Figure. 6: The Architecture of the Proposed Context-Aware Wireless Actuator

VI.   TESTBED DEPLOYMENT

The microgrid testbed will be deployed at one of AUI buildings. See figure 5. The renewable will be integrated via the installation of a Parking Port Solar Farm next to the concerned building. We are intending to deploy the wireless sensor in the base floor of a building following the layout in figure. 2.

Figure. 7: Testbed Smart Building

VII.   CONCLUSION AND FUTURE WORK

In this paper we presented a holistic architecture for deploying a real-world microgrid testbed at AUI university campus. We listed the main components and emphasized upon the ICT component.

5

Since the data emanating from Smart Grids falls within the realm of Big Data, we advocated the use of HPC (High-Performance-Computing) via the deployment of a Hadoop cluster on top of the Openstack Cloud Computing platform, We delineated the wireless network architecture for data dissemination, and highlighted the actual implementation of the sensors to be used.

As a future work, we will further delve into the deployment of the middleware sub-component and interface the BPAP (Big Data Analytics Platform) with the smartgrid *control plane* component.

REFERENCES

[1] X. Fang, S. Misra, G. Xue, D. Yang, "Smart Grid – The New and Improved Power Grid: a survey", IEEE Communications Surveys & Tutorials, Vol. 14, No. 4, 2012

[2] DOE (U.S Department of Energy), "Buildings Energy Data Book, Department of Energy, March 2015."

[3] M. R. Abid, A. Khallaayoun, H. Harroud, R. Lghoul, M. Boulmalf, D. Benhaddou, "A Wireless Mesh Architecture for the Advanced Metering Infrastructure in Residential Smart Grids", IEEE GreenTech, April, 2013

[4] S. Sagiroglu, D. "Big data: A review." IEEE International Conference on Collaboration Technologies and Systems (CTS), 2013

[5] P. Barnaghi, Payam, "Semantics for the Internet of Things: early progress and back to the future." International Journal on Semantic Web and Information Systems (IJSWIS) 8.1 (2012): 1-21.

[6] The Openstack Cloud Software, "Open source software for building private and public clouds"

[7] K. Pepple, "The OpenStack Project" in Deploying Openstack, O'Reilly Media, July 2011.

[8] Apache Hadoop Web site

[9] J. Dean, S. Ghemawat, "Mapreduce: Simplified data processing on large clusters". In In Proceedings of the 6th USENIX OSDI, pages 137–150, 2004.

[10] J. Dean, S. Ghemawat,. "MapReduce: Simplified data processing on large clusters". Communications of the ACM, 51(1):107–113, 2008.

[11] P. Mell, T. Grance, "The NIST definition of cloud computing". NIST Special Publication 800-145, in press.

[12] J. Smith, R. Nair, "The architecture of Virtual Machines", IEEE Computer Journal, Vol. 38, No. 5, 2005

[13] F. Akyildiz and X. Wang, "Wireless Mesh Networks: A Survey," Computer Networks and ISDN Systems, Vol. 47, No 4, pp. 445-487, 2005

[14] F. Akyildiz, W. Su, Y. Sankarasubramaniam, E. Cayirci, "Wireless sensor networks: a survey", The International Journal of Computer and Telecommunications Networking, Vol. 38, No. 4, pp. 393-422, 2002

[15] D. B. Johnson and D. A. Maltz, "Dynamic source routing in ad-hoc wireless networks", Mobile Computing Kluwer Academic Publishers, 1996.

[16] M. Bahr, "Update on the Hybrid Wireless Mesh Protocol of IEEE 802.11s", IEEE Conference on Mobile Adhoc and Sensor Systems, 2007. Computing Kluwer Academic Publishers, 1996.

[17] C. E. Perkins and E. M. Royer, "Ad-hoc on-demand distance vector routing", WMCSA, 1999.

[18] O Achahbar, MR Abid, M Bakhouya, C El Amrani, J Gaber, M Essaaidi, T. Elghazawi, "Approaches for High-Performance Big Data Processing", book chapter, Big Data: Algorithms, Analytics, and Applications, *Chapman and Hall,* 2015.

[19] M. R. Abid, O. Achahbar, T. Rachidi, M. AlAssaf, N. Khalil, D. Benhaddou, "Smart Grids and Cloud Computing: A Synergy", The 8th IADIS International Conference on Information Systems, Madeira, Portugal, March 14-16, 2015

[20] O. Achahbar, M. R. Abid, "The Impact of Virtualization on HPC Clustering in the Cloud", International Journal of Distributed Systems and Technologies (IJDST), Vol. 6, No. 4, pp. 65-81, 2015

[21] O. Achahbar, "The Impact of Virtualization on High Performance Computing Clustering in the Cloud", Thesis

[22] M. R. Abid, S. Biaz, " Interference-aware Routing Metric for IEEE 802.11s Wireless Mesh Networks", International Journal of Mobile Communications (IJMC), Vol. 9, No. 6, pp. 619-641, 2011

6