

A Generic Framework for Modeling and Simulation of Cloud Computing Services

Maria Salama
British University in Egypt
Cairo, Egypt

Ahmed Shawish
Ain Shams University
Cairo, Egypt

Amir Zeid
American University of Kuwait
Kuwait City, Kuwait

ABSTRACT

Cloud Computing is a paradigm in which tasks are assigned to a combination of computing resources, software and services accessed over a network following the pay-as-you-go financial model. It has been also described as on-demand computing. With the continuous increases of cloud service providers, it becomes crucial to develop a simulation tool to reflect the properties of such complex environment to help clients selecting the appropriate providers. Available Cloud-based tools are designed for cloud architectures and resources scheduling, not the problem of provider selection. Even the Grid-based tools who share many features with the Cloud cannot cope with such problem due to the novel characteristics and services of the Cloud. This paper provides a new simulation tool that reflects the nature of clouds embedding all its aspects, as well as its QoS parameters. Such tool is designed to simulate any framework or solution for service provider selection problem. The proposed simulation tool is validated by running a framework developed for service provider selection problem based on QoS and utility functions. The paper also reviews various mathematical approaches that have been used to model cloud services, where most of them are formulations of cloud services that aim to optimize its quality of service, performance or energy efficiency under given constraints.

General Terms

Cloud Computing, Simulation

Keywords

Cloud Computing, Simulation, Quality of Service

1. INTRODUCTION

Cloud Computing is a paradigm in which tasks are assigned to a combination of computing resources, software and services accessed over a network. This network of servers and connections is collectively known as "the cloud." Computing at the scale of the cloud allows users to access supercomputer-level power. Using a thin client or other access point, like smart phones or laptop, users can reach into the cloud for resources as they need them. For this reason, Cloud Computing has also been described as "on-demand computing". Standing at different level of hardware and software stack levels (IaaS, PaaS, and SaaS), this form of service provides several benefits to consumers; no cost for purchasing, free of maintenance, accessibility through Internet, and high availability. Computing services are offered on-demand according to an auto-scaling paradigm following the pay-as-you-go financial model. Such flexible key tenets lead the Cloud Computing to become the next wave of technological revolution that will be able to provide IT-services required by business.

With the emergence of the cloud services in the academic framework, scientific researches, governmental services and enterprises, many service providers rather than the famous ones like Microsoft, Amazon, Google, and IBM started to rise up providing various services with very competitive prices. The clients on the other side are daily facing a huge amount of service offers that simply put them in a confusing selection to fulfill their real needs with an appropriate price and best quality of services.

The available tools that can be logically used in such dilemma cannot solve this problem as it should be due to many factors. From one side, Cloud-based tools are designed for cloud architectures and resources scheduling, not the problem of provider selection. Even though, those tools are a combination of mathematical equations that do not embed all QoS parameters and do not represent a comprehensive solution since regular customers cannot use mathematical equations in selecting providers. From the other side, the remaining available tools were originally designed to cope with the Grid that may share many aspects and features with the Cloud. However, they are still unable to cover the gap between the two infrastructures. For instance the Grid technology targets to build and run only distributed processes and applications accessible via the Internet; while Cloud have multiple service models (i.e., IaaS, PaaS and SaaS) with the purpose of enabling dynamic resources allocation for running web services. Also, Grid Computing encapsulates application functionality and information resources only; while each cloud service model encapsulates certain computing resources. For example, IaaS offers virtualized hardware and storage on top of which customers can build their infrastructure; encapsulating computational resources; such as memory, CPU, storage and operating systems. SaaS provides customers with various kinds of applications that are accessible anytime and from anywhere. These issues reflect that clouds are computing platforms offering services from a wider scope and range. Thus, due to the difference in nature between the two technologies, applying Grid Computing tools in Cloud Computing will lead to incorrect definition of the quality parameters for Cloud Computing.

It becomes crucial to design a simulation tool that is able to reflect the properties of the Cloud complex environment to help clients selecting the appropriate providers. In this paper, we provide a new simulation tool that reflects the nature of clouds embedding all its aspects as well as its QoS parameters. Such tool is designed to simulate any framework or solution for service provider selection problem. The proposed simulation tool is validated by running a framework developed for service provider selection problem based on QoS and utility functions. The paper also reviews various mathematical approaches that have been used to model cloud services, where most of them are formulations of cloud

services that aim to optimize its quality of service, performance or energy efficiency under given constraints.

The rest of this paper is organized as follows. Section 2, introduces the background and related work on the available cloud-based simulation tools and mathematical modeling of the service provider selection problem. In section 3, we present the proposed simulation framework. In section 4, we introduce related work - about a QoS model for Cloud Computing and a mathematical model for service provider selection problem - which will be used in the conducted simulation studies. In Section 5, we conduct simulation studies to evaluate the efficiency of the proposed simulation tool. Finally, we conclude and reflect our future work in Section 6.

2. BACKGROUND

In this section, we discuss the available simulation tools developed for Cloud Computing. We also discuss the background on various mathematical approaches that have been used to model cloud services, where most of them are formulations of cloud services that aim to optimize its quality of service, performance or energy efficiency under given constraints.

As for the simulation tools for Cloud Computing, there are several frameworks developed; such as CloudSim, CloudAnalyst, NetworkCloudSim, MDCSim and iCanCloud. In [1], authors have proposed CloudSim, a simulation framework that enables seamless modeling, simulation, and experimentation of emerging Cloud Computing infrastructures and application services. CloudSim could be used to investigate specific system design issues, development of new algorithms, methods, and protocols. It is a self-contained platform for modeling data centers, service brokers, scheduling, and allocations policies. Another proposed framework named CloudAnalyst [2], supports visual modeling and simulation of large-scale applications that are deployed on Cloud Infrastructures. CloudAnalyst allows description of application workloads, including information of geographic location of users generating traffic and location of data centers, number of users and data centers, and number of resources in each data center. Using this information, CloudAnalyst generates information about response time of requests, processing time of requests, to determine the best strategy for allocation of resources among available data centers, strategies for selecting data centers to serve specific requests, and costs related to such operations. There is also a simulation framework named NetworkCloudSim [3], which supports modeling of real Cloud data centers and generalized applications such as HPC, e-commerce and workflows. It mainly addressed modeling a data center's internal network and applications. Another cloud simulator is MDCSim [4], which has been designed with an emphasis on multi-tier data centers. It can analyze a cluster-based data centre with detailed implementation of each individual tier. It has been configured into three layers, including a communication layer, a kernel layer and user-level layer, for modeling the different aspects of a cloud, and can estimate the throughput, response times, and power consumption. The latter is approximated using linear functions of the server utilization, which in turn is calculated based on the number of nodes, number of requests and average execution time of requests. Nunez et al. [5, 6] have developed the iCanCloud simulator. It is based on, which is a software simulation framework for large storage networks. iCanCloud can predict the trade-off between costs and performance of a particular application in a specific hardware in order to inform the users about the costs

involved. They have focused in particular on Amazon-like policies which charge users in a pay-as-you-go manner. iCanCloud has a full graphical user interface from which experiments can be designed and run, but existing software systems can only be modelled manually. It also allows parallel execution of one experiment over several machines. It is notable that existing simulation tools are designed for clouds architecture, resources scheduling, and performance evaluation; not for simulating problems related to QoS; such as service selection and service evaluation.

On the other hand, mathematical modeling and analysis used for providing an understanding of the interdependencies involved in cloud computing. It is particularly suitable for identifying optimal values and predicting behavior. With respect to performance evaluation, response time and throughput factors have been mathematically modeled in [7], where authors focused on the response time from the perspective of the customer who is more inclined to request a statistical bound on its response time than an average response time. That research is concerned with a percentile of the response time; that is the time to execute a service request is less than a pre-defined value with a certain percentage of time. Authors in [8] have developed an elasticity model for cloud instances, taking the assumption that each resource type (CPU, memory, network bandwidth, etc.) can be allocated in units and that the user is aware of the allocated resources and the relevant QoS metrics for their requests, as in the case of Amazon CloudWatch. The model combines the cost of provisioned but unutilized resources and the performance degradation cost due to under provisioning. Others [9] have tackled the optimization problems of virtual machine provisioning and placement. For the former, they have developed a global utility function representing power saving and SLA satisfaction, while they have approached the latter as a multiple knapsack problem with capacity constraints on the physical hosts. They have also described how their mathematical models can be used with a middleware framework managing the provisioning and placement of virtual machines on an actual cloud system. Such form of mathematical modeling doesn't represent a comprehensive solution for problems based on QoS factors; since it doesn't embed all QoS parameters. Also, Cloud consumers, not IT specialists, find it hard to run the mathematical solutions to solve their problems; such as service selection.

With the simulation tools developed for designed for clouds architecture, resources scheduling, and performance evaluation, and the mathematical modeling of quality of service; arises the deep need for a simulation tool that reflects the nature of Cloud Computing services and embeds all QoS aspects related to clouds. Such tool is required to simulate any framework or solution related cloud services based on QoS factors.

3. SIMULATION FRAMEWORK

In this paper, we provide a comprehensive dynamic simulation framework that reflects the nature of the cloud services, embed all quality parameters related to cloud services that could be dynamically used to run any simulation problem related to cloud services.

The simulation framework is designed and developed with the purpose of simulating cloud computing services, their QoS and modeling related problems that are mathematically formulated; such as optimum provider selection. The framework will be opted to run models related to QoS in an easy and expedient way and helps decision makers to set their

preferences and get out analytical results for their problems related to cloud services.

3.1 Functionalities

The simulation framework is programmatically implemented to perform the following functionalities. The system fulfills the following functional requirements:

- A list of quality metrics, from which the user will select his preferences, is carried out.
- The addition of more quality dimensions and parameters dynamically is supported.
- The system is designed to adopt criteria for solving user problems; such as optimum service selection; and dynamically runs the simulation. Criteria may include mathematical formulations.
- The decision maker is allowed to set his preferences for the criteria. This should be done by choosing the quality metrics and set his preferences to run any service related problem.
- All service parameters should be submitted with their quality parameters.
- Framework is capable to run simulate services to run solutions for the imposed problem.
- The system should display the optimum solution that yields the maximum utility value as a final output.
- Application should also display a graphical representation for the final calculations, as well as intermediate results.
- The system should allow the decision makers to repeat their experience under various parameters.

3.2 Simulating Cloud Services

The cloud services are simulated in terms of their quality parameters and imposed problem for simulation is settled by its criteria, parameters and mathematical formulations. For instance, in the optimum service provider selection problem, candidate providers data are represented by the quality parameters required for the selection process to take place of candidate providers. The selection problem is settled by the selection criteria; selecting the quality parameters that will take place in the selection process, their constraints, and their weights.

The simulation is designed to output the following:

- Intermediate results of the imposed simulation problem.
- Graphical representation of the final values.
- Optimal result.

3.3 Design and Implementation

Figure 1 shows the system design. QoS is represented in terms of Quality Dimensions and Quality Attributes. QoS attributes (QA) are simple sets of quality parameters; such as response time, throughput. For each QoS attribute a constraint value - minimum or maximum - is specified expressing the consumer's requirement regarding the corresponding quality attribute. Also, a range of actual values, from which the consumer will pick up the constraint value, is previously specified. Second, a further set of QoS dimensions (QD) is defined for better categorization for quality parameters.

A set of measurable parameters coincide with each QD, reflecting in the use of QoS in any simulation related to cloud services and providers. With the aim of providing cloud consumers with a flexible framework useful for running any simulation related to Cloud services, the proposed QoS attributes and dimensions can be applied on different Cloud Computing services: SaaS, PaaS or IaaS. The flexibility supported by the model to include various QoS attributes and

dimensions make it suitable for different kind of services. For instance, factors such as reusability, availability and scalability of the software will be applied in case of SaaS, while in case of IaaS factors like performance and availability of resources such as processors memory, storage and networks will be considered. As a matter of fact, characterization of the services facilitates the determination of the appropriate QoS attributes and dimensions for the imposed problem. It is also worth to note that Quality Attributes and Dimensions can be extended by addition or removal according to the user QoS requirements. In addition, quality parameters may also be extended within one Quality Dimension. Moreover, those dimensions and their parameters are weighted upon consumer preferences.

4. RELATED WORK

In this section, we present a comprehensive QoS model for Cloud Computing [10] and a mathematical model for service provider selection problem [11]. The broad QoS model combines all the quality of service related parameters of both service providers and consumers for different Cloud platforms - IaaS, PaaS and SaaS. The service selection model is a mathematical model based on utility functions addressing the problem of cloud service provider selection, to satisfy the utility-maximization consumer need.

4.1 Cloud Computing QoS Model

The QoS model describes the concrete level of quality requirements and defines QoS properties for cloud services. The defined QoS properties are a set of domain-independent quality concepts and can be widely used by various cloud computing applications. Each quality factor is described in light of a defined ontology, where each one is defined as a QoS property with a specific metric providing the necessary information of how to measure, compare and assess its value. QoS properties are classified into two main categories: technical factors and business factors. Technical factors include performance, availability, reliability and others; while business factors are mainly concerned with manageability, transparency, economics and others. A general overview of the QoS framework is depicted in fig. 2.

4.2 Utility-Based Model for Cloud Computing Service Provider Selection

The service selection model is a mathematical model addressing the problem of cloud service provider selection, to satisfy the utility-maximization consumer need. This approach uses utility functions in a QoS management framework with the goals to meet the minimum requirements and maximize the total utility for the consumer. By computing the utility value (eq. 1) associated to each provider, the optimal provider will be the one with the highest utility value.

$$f(U_i) = \sum_{j=1}^m (\sum_{l=1}^{q_j} P_{ijl} * w_{jl}) * W_j \quad (1)$$

where the calculation of the utility value u_{ij} is addressed as follows. Any quality dimension j , where $j \in \{1, \dots, m\}$, is defined by a group of measurable parameters P_{jl} , where $l \in \{1, \dots, q_j\}$, and q_j is the size of this group. For example; a quality dimension like the performance is defined in terms of measurable parameters like response time, worst-case execution time, and throughput. Each of these parameter P_{jl} is further associated with a weight w_{jl} , where for $w_{jl} \in [0, 1]$, and $\sum_{l=1}^{q_j} w_{jl} = 1$. Varying P_{jl} , different parameters can be considered for one QoS dimension.

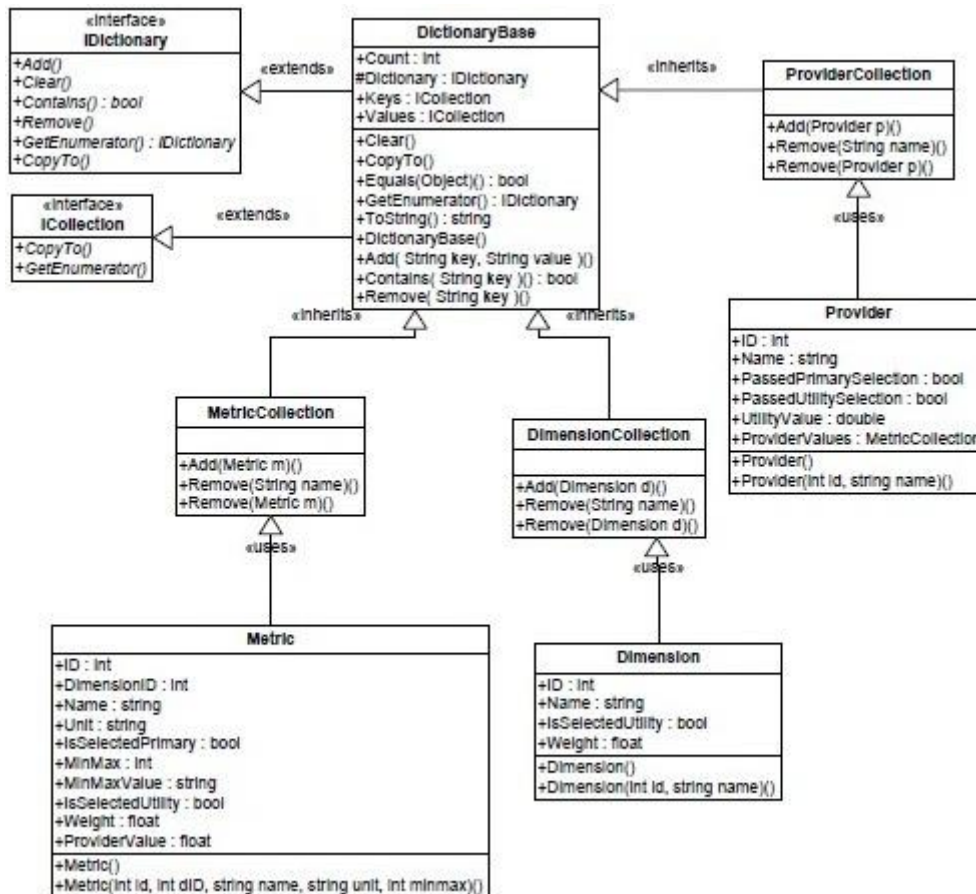


Fig 1: Framework Design

5. VERIFICATION AND RESULTS

In this section, the developed simulation tool is validated by running the utility-based model for cloud computing service provider selection.

Simulation data used in the experimental study has been published in [12]. This simulation setting has been established for evaluation under a more realistic environment. Consider a set of 20 software services $P_i, i = 1..20$, say, for instance, 20 storage service providers, with 20 different quality attributes, as illustrated in Table 1.

For the primary selection phase, Table 2 shows the list of the considered quality attributes, and their upper or lower bounds. These values are assigned by the consumer based on his own interest, within a range of values, previously specified, in order to make it easier for the consumer.

With respect to the utility value calculation, Table 3 holds the set of QoS dimensions and their assigned weights, along with the measurable parameters for each dimension.

Using the proposed simulation framework, a primary selection is firstly performed using the criteria settled by the customer. Thus, the providers $P_2, P_5, P_8, P_9, P_{10}, P_{14}, P_{15}, P_{16}$ and P_{17} will pass to the next phase of selection. The framework will then apply the next selection phase; dominance selection, where none of those providers will be rejected as none of them dominates one another in all QAs. The utility calculation is then performed. Next, the framework will calculate the utility values for each QD. The utility function is then calculated out

of the utility values of all quality dimensions times their corresponding weights using Eqn. 1. The calculated utility values for providers that passed to the second phase are illustrated in Figure 3. Hence, the simulation framework will easily output the provider is P_9 as the optimum one since it yields maximum utility value (79.21). Implementing the framework, a screen shot of the detailed results form appears in Figure 4.



Fig 2: Cloud Computing QoS Model

Table 1. Table captions should be placed above the table

Provider	Fees	Uptime	TBF	TTR	ATM
0	74.0	97.99	165	14	0
1	98.0	97.99	184	25	86
2	81.0	99.99	133	43	44
3	71.0	98.99	151	47	83
4	98.0	97.99	105	11	71
5	75.0	99.99	177	26	43
6	76.0	97.99	182	42	28
7	100.0	99.99	156	44	92
8	77.0	99.99	116	16	46
9	81.0	99.99	188	32	82
10	74.0	99.99	123	26	89
11	75.0	98.99	111	48	71
12	98.0	97.99	186	17	63
13	81.0	97.99	142	34	66
14	81.0	99.99	180	42	65
15	73.0	98.99	120	44	73
16	80.0	99.99	149	29	10
17	81.0	99.99	176	20	6
18	88.0	97.99	102	30	6
19	92.0	97.99	185	14	41

Table 2. Simulation Parameters – Quality Attributes, Weights and Boundaries Values

QoS Attribute	Range	Constraint
Time to Recovery (TTR)	15 – 50	Max. 45
Uptime	97.00 – 99.99	Min. 98%
Service fees	71.0 – 100.0	Max. 85

Table 3. Simulation Parameters – Quality Dimensions, Parameters and Weight Values

QoS Dimension	Weight (W_j)	QoS Parameters	weight (w_j)
1. Dependability	0.40	1.1 Uptime	0.60
		1.2 Time to Recovery (TTR)	0.40
2. Performance	0.25	2.1 Time Between Failures (TBF)	1.0
3. Reputation	0.10	3.1 Arrival Time To Market (ATTM)	1.0
4. Service Value	0.25	4.1 service fees to be paid	1.0

The simulation scenario demonstrated the effective use of the proposed simulation framework. The framework succeeded to simulate a complex scenario for the service provider selection problem; incorporating quality of service factors and mathematical formulations. It is also worth to note that the proposed framework has easily adopted the complications of the consumer preferences towards the quality metrics, presented in the specified weights, had affected the performed calculations and hence the obtained results.

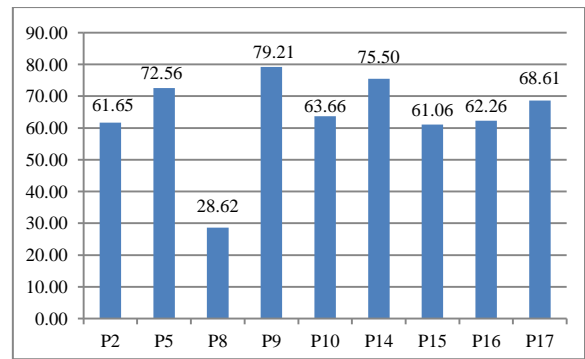


Figure 3. Simulation Results (Utility functions)

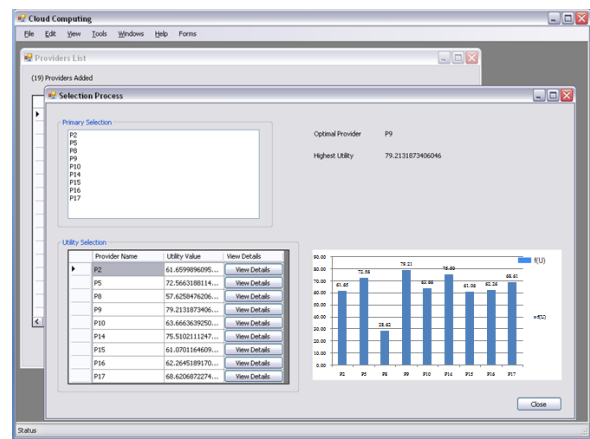


Figure 4. Detailed results form

6. CONCLUSION

This paper proposed a new simulation tool that considers all of the Cloud's properties, features, and services to help clients selecting the appropriate Cloud service providers; embedding all the aspects of clouds, as well as its QoS parameters. The conducted simulation studies reports the ability of the proposed tool to simulate any framework or solution related to the imposed selection problem. The proposed simulation tool has been validated by running a framework developed for the selection problem based on QoS and utility functions. The paper has also reviewed the available cloud-based simulation tools and various mathematical approaches of the service provider selection problem.

Future work will focus on further developments of the simulation tool to cover other approaches for quality assurance and monitoring. Approaches for quality assurance will be settled in parallel while the service is running to monitor and assure the quality level.

7. ACKNOWLEDGMENTS

This work was supported by Professor Mohamed Kouta, Arab Academy for Science, Technology & Maritime Transport.

8. REFERENCES

- [1] R. Calheiros, R. Ranjan, A. Beloglazov, C. De Rose, and R. Buyya, "CloudSim: A Toolkit for Modeling and Simulation of Cloud Computing Environments and Evaluation of Resource Provisioning Algorithms," in *Software: Practice and Experience*, Vol. 41, No. 1, pp. 23-50, Wiley Press, New York, USA, January 2011.
- [2] B. Wickremasinghe, R. Calheiros, and R. Buyya, "CloudAnalyst: A CloudSim-based Visual Modeller for Analysing Cloud Computing Environments and Applications," in *Proc. of 24th IEEE International Conference on Advanced Information Networking and Applications (AINA 2010)*, Perth, Australia, April 2010.
- [3] S. Garg and R. Buyya, "NetworkCloudSim: Modelling Parallel Applications in Cloud Simulations," in *Proc. of 4th IEEE/ACM International Conference on Utility and Cloud Computing (UCC 2011)*, Melbourne, Australia, December 2011.
- [4] S. Lim, B. Sharma, G. Nam, E. Kim, C. Das, "Mdcsim: a multi-tier data center simulation, platform," in *Proc. of IEEE Int. Conference on Cluster Computing and Workshops*, IEEE, New Orleans, Louisiana, USA, 2009.
- [5] A. Nunez, J. Vazquez-Poletti, A. Caminero, G. Castane, J. Carretero, I. Llorente, "iCanCloud: a flexible and scalable cloud infrastructure simulator," in *Journal of Grid Computing*, vol. 10, issue 1, 2012, pp. 185–209.
- [6] A. Nunez, J. Vazquez-Poletti, A. Caminero, G. Castane, J. Carretero, I. Llorente, "Design of a new cloud computing simulation platform," in *Proc. Int. conference on Computational science and its applications (ICCSA'11)*, Santander, Spain, 2011, pp. 582–593.
- [7] K. Xiong, and H. Perros, "Service Performance and Analysis in Cloud Computing," in *Proc. of 2009 Congress on Services – I (SERVICES 09)*, IEEE Computer Society Washington, 2009, pp. 693 - 700.
- [8] S. Islam, K. Lee, A. Fekete, A. Liu, "How a consumer can measure elasticity for cloud platforms," in *Proc. of 3rd Joint WOSP/SIPEW International Conference on Performance Engineering (ICPE '12)*, Boston, Massachusetts, USA, 2012, pp. 85–96.
- [9] H. Van, F. Tran, J. Menaud, "Performance and power management for cloud infrastructures," in *Proc. of 3rd International Conference on Cloud Computing*, IEEE, Miami, FL, USA, 2010, pp. 329–336.
- [10] Maria Salama, Ahmed Shawish, and Mohamed Kouta, "Generic Ontology-based QoS Model for Cloud Computing," in *Proc. of International Conference on Computer and Information Technology (ICCIT 2011)*, Amsterdam, The Netherlands, July 2011.
- [11] Maria Salama, Ahmed Shawish, Amir Zeid and Mohamed Kouta, "Integrated QoS Utility-Based Model for Cloud Computing Service Provider Selection," in *Proc. of 36th Annual IEEE Computer Software and Applications Conference (COMPSAC 2012)*, Izmir, Turkey, July 2012.
- [12] N. Limam and R. Boutaba, "Assessing Software Service Quality and Trustworthiness at Selection Time" in *IEEE Transactions on Software Engineering*, vol. 36, issue 4, July 2010, pp. 559 - 574.