2016 16th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing

Service Level Agreement Assurance between Cloud Services Providers and Cloud Customers

Abdallah Ali Zainelabden A. IBRAHIM, Dzmitry Kliazovich, Pascal Bouvry
University of Luxembourg
Faculty of Science, Technology and Communications
6, rue Richard Coudenhove-Kalergi
Luxembourg City, Luxembourg
Email: {first name, last name}@uni.lu

***Abstract*—Cloud services providers deliver cloud services to cloud customers on pay-per-use model while the quality of the provided services are defined using service level agreements also known as SLAs. Unfortunately, there is no standard mechanism which exists to verify and assure that delivered services satisfy the signed SLA agreement in an automatic way. There is no guarantee in terms of quality. Those applications have many performance metrics. In this doctoral thesis, we propose a framework for SLA assurance, which can be used by both cloud providers and cloud users. Inside the proposed framework, we will define the performance metrics for the different applications. We will assess the applications performance in different testing environment to assure good services quality as mentioned in SLA. The proposed framework will be evaluated through simulations and using testbed experiments. After testing the applications performance by measuring the performance metrics, we will review the time correlations between those metrics.**

Keywords*—*Cloud Computing, Service Level Agreement, Data Centers, Simulation, Quality of Services, Quality of Experience, Applications, Performance, Metrics.

I. Introduction

Services such as Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS) are provided worldwide over the Internet [1] through the new computing paradigm which is called cloud computing [2]. In the cloud computing system, there are many players: cloud service provider, cloud customer (which will be a service provider for end customers), network service provider and end customer. Cloud service provider offers services to the high tier cloud customer. The high tier cloud customer provides cloud services to the end customers through network service providers. Figure 1 shows the players of cloud computing system. In this doctoral thesis, we will assure the SLA compliance between the cloud provider and the end cloud users. SLA is the part of a contract which defines exactly what services a service provider will provide and the required level or standard for those services. This contract is needed to be accepted, negotiated and agreed between the two sides, the cloud customers and the cloud providers. SLA plays an important role in defining the relationship between cloud customers and cloud services providers. Example of SLA is a part of the contract between a business company and its information technology supplier. Cloud data center is the house of cloud computer systems and associated components such as cloud networking infrastructure and cloud storage systems. Cloud data centers host a lot of hardware, which individual components may fail, affecting the performance of the whole system. Once a physical server, network or a virtual machine (VM) failure occurs [3], it can affect degrading the performance of cloud applications or even halt their execution. This could lead to the violation of SLA signed between the cloud services provider and the cloud user. The cloud service provider should pay to the cloud customer a pre-defined penalty for each SLA violation [4].

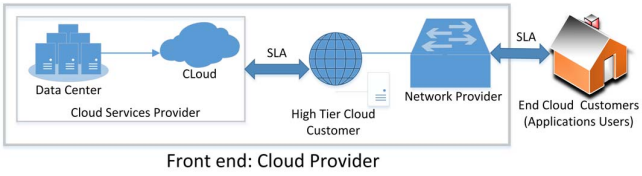

Fig. 1: Cloud services provider and end customers.

SLA guarantees cloud customer a good level of service, giving him confidence that if something goes wrong, the provider will respond quickly. However the problem is that the quality of services level is just written on the SLA document, but usually there is no process to verify that the services are delivered with due quality. In this proposed research work, the quality of services provided to the customers will be assessed, and this allow us to assure what the SLA document cover. SLA covers two main metrics. The first is the uptime metric which guarantees apply and access to services. The second is the response time metric which is the quality and performance of services. In this research work, SLA response times will be assured, in other words, the performance of the services is needed to be assured and a high Quality of Services (QoS) is also needed to be guaranteed all the time as the cloud providers mention in the SLA contract. It is hard for cloud customers to verify on SLA metrics, so we need to evaluate these arguments between cloud providers and cloud customers to provide solutions to cloud providers if failure exist and provide an assurance of SLA QoS to cloud customer if

978-1-5090-2453-7/16 $31.00 © 2016 IEEE
DOI 10.1109/CCGrid.2016.56

services performance really working good all the time even if there are failures in data centers where the cloud services are hosted.

Based on the above, during this proposed work, first, an assessment to the SLA will be provided by assessing the applications performance and evaluating the failures effect on them. Second, an assurance to the SLA between the cloud provider and the cloud customer will be provided by solving and mitigating the failures that appear in the cloud data centers and keep the applications working with good performance. Before that, the performance metrics for the cloud applications will be identified first. This will allow us to measure the performance of the applications. The failures may occur inside cloud data center are needed also to know their impact on the applications performance. Finally, the time correlations between the cloud applications metrics will be explained to know the relations between those metrics.

The rest of this paper is organized as follows. Section II presents the related work and comparison with previous work. Section III presents the research plan to assure SLA. The expected outputs are in Section IV. Section V presents the conclusions.

II. Related Work

In this section, we provide an overview of some of the work related to service level agreement violations and SLA assurance, cloud data centers and their failures and cloud applications performance metrics.

A. Service Level Agreement Assurance

Shyam Wagle in [5] introduced the SLA assured brokering framework which matches the requirements of cloud customers with SLA provided by the cloud services provider. In our research proposal, the SLA provided by the cloud services provider to cloud customer will be assured by measuring the performance of the cloud applications which are hosted on the cloud service provider data center and working on the cloud customer's devices. The author in [5] relied on external cloud auditing services without measuring any quality of services and he ranked the cloud services providers just based on the response time of cloud provider.

But in our work, the ranking will be based on the quality of services which are provided form the cloud provider to the cloud customers.

Arpan Roy *et al.* in [4] introduce the KIM software framework as a cloud controller which aids in minimizing service failures which happen because of the SLA violations. SLA violations are like violation of utilization, availability and response time in the SaaS cloud data centers. They used only the migration as a failure mitigation technique and they focused on the services availability and work load.

But in this proposed work, the focus will be on the services quality and the performance level mentioned on SLA document.

Wood *et al.* in [3] introduced the Sandpiper system for automated mitigation of increasing response time of the host Physical Machines and utilization in a virtualized data center due to workloads. The migration in Sandpiper system only depends on the increasing of utilization threshold and does not depend on SLA failures or violations. Shen *et al.* [6] introduced the CloudScale system that make a proactive forecasting for the upcoming SLA violation, they also did a dynamic resource allocation in addition to the migration of the work load. CloudScale system uses a Markov Chain based state space approach to do the prediction process. Last but not least, Salman in [7] provides a study which states that there are no any cloud providers that offer performance guarantees of computer services or allow the customer to detect the SLA violation and that is why we propose a methodology and framework for SLA assurance in this research work.

B. Cloud Data Center Failures

Kashif Bilal *et al.* in [8] discuss the challenge and trends in the cloud data centers. It provides a review of the cloud data centers architecture and the structure of the cloud data center network. These challenges are representing the failures that may occur inside the cloud data centers. During this proposed work, the resolution of these challenges are needed to assure SLA compliance.

Ranjithprabhu *et al.* in [9] proposed a technique to eliminate the single point of failure in the cloud data centers by elaborating redundancy techniques. They also operate data replication by using data mapping to prevent data losing in cloud computing.

In this work, the redundancy and replication will be used to mitigate the failures may occur inside cloud data centers which host the cloud provider services.

Phillipa Gill *et al.* in [10] answered some questions to understand the network failures in cloud data centers by measuring the impact of the failures then analyzing the measurements and finally discussed the implications of the failures. They also introduced a brief discussion about the cloud data centers network structure. But, they did not provide any resolving ways to those failures so the failures resolution and mitigation are needed and we will provide them during this research work.

C. Cloud Applications

There are many research works related to the cloud computing applications. Ali *et al.* in [1] gave an overview of the cloud computing applications and discussed the characteristics, traits and issues of the cloud computing applications. Puja in [11] discussed the cloud computing and the software as a service applications. The author explained the concept, services provided by cloud computing and different service providers. The authors of [1] and [11] did not introduce any performance metrics for the cloud applications and how to analysis the behavior of those applications.

During this research proposal, the performance metrics for cloud applications will be provided and also the time correlations between those metrics.

III. Research Plan

A. Methodology

Figure 2 shows our methodology for SLA assurance between cloud service providers and cloud customers.

- *Cloud Customers*
  In our work, we classify the cloud customers in different cloud users based on the applications usage which are used by the cloud customers such as web, file, distributed, real time, highly interactive and mobile applications. Due to different application respond differently to failures, the applications performance metrics will be defined in this research work.
- *Applications Performance Metrics*
  It contains the multiple cloud applications performance metrics. Each cloud application has multiple metrics which indicate the performance of this application. For example, response time is one of the performance metrics for web application and end to end delay is one of the performance metrics for real time application. In this proposed work, all performance metrics are measured to indicate the quality of the applications. Quality of applications measurements will indicate the quality of services which are provided by cloud services providers.
- *Cloud Services Providers*
  It contains the multiple data centers with different cloud services resources. The Cloud services are offered by cloud services providers and are provided to the customers according to their needs. The information about the cloud infrastructure like server load, CPU utilization and network metrics such as bandwidth, delay, jitter and error rate are provided by the cloud service provider. These values are mentioned inside the SLA document. They are mentioned as two metrics, the quality of services and the services availability.
- *Network Performance Metrics*
  It contains the network metrics such as the bandwidth, delay time, jitter and error rate. The degradation of network performance metrics can affect the quality of the services provided from cloud services provider to the cloud users. In this proposed work, the degradation in network performance metrics will be used as a failure to study the impact of failures on the applications performance and on the quality of services.
- *SLA Verifier*
  In this proposed work, the quality of services will be assessed and assured by a SLA verifier. The SLA verifier measures the quality of services and the quality of experience based on the applications performance metrics and the metrics of the cloud provider. These measurements are compared with the parameters mentioned in SLA document. If the measurements are verified by the SLA values so, the cloud services provider are providing his services with good quality and as mentioned in SLA document. If the measurements are not verified by SLA values, the cloud services provider are not providing his services with good quality and not as mentioned in SLA document. In this case, SLA verifier will propose solutions to prevent any SLA violations.

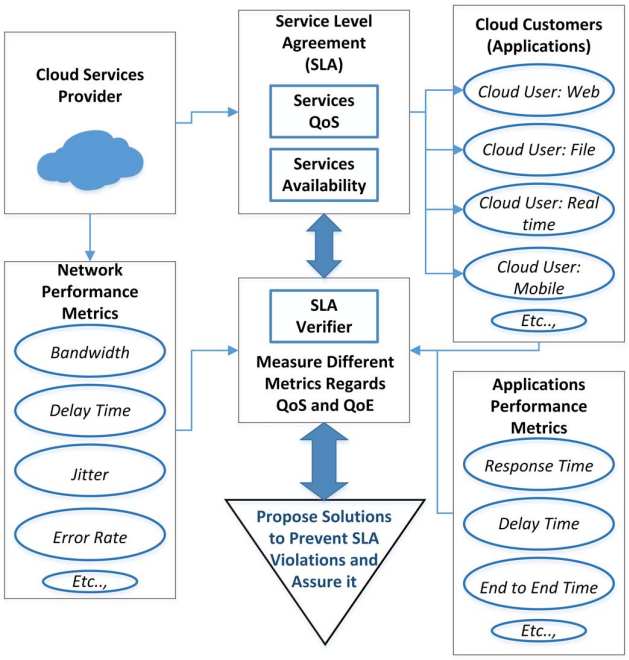

Fig. 2: Our methodology to assure SLA between cloud providers and cloud customers.

Figure 3 shows how the SLA verifier is working in four steps between cloud services providers and cloud end customers. First, SLA verifier introduces failures to the cloud services provider which may occur inside the cloud data centers. Failures such as bandwidth degradation and increasing in delay time and error rate. Second, it measures the applications performance metric to show the impacts of the failures on the applications performance. Because of diversity of applications performance metrics, the measurements will be for each class of application separately. Third, it provides mitigation and resolution techniques to resolve the failures. The mitigation techniques will be proposed to the services provider to use them in case of performance degradation and prevent SLA violations. Finally, SLA verifier measure again the performance metrics and assure the SLA compliance.

B. Testing Models

Testing cloud applications will be performed in two ways. The first consists of testbed experiments which are a real testing scenarios. The second relies on simulation experiments on the green cloud simulator [12]. Using Green cloud simulator, applications performance can be tested in simulating big data center topologies.

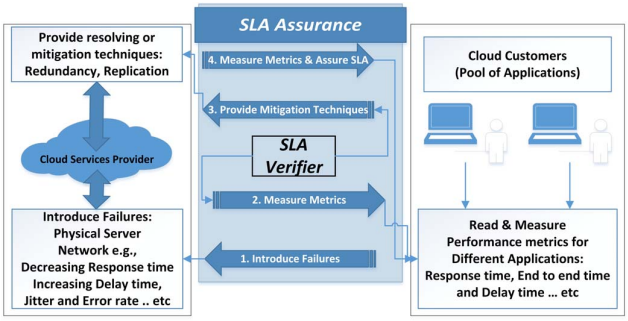

Fig. 3: SLA verifier structure.

IV. Expected Output

The expected output from this proposed research work is a framework which can be used by both cloud providers and cloud users. Inside the framework, the performance metrics for different cloud applications will be defined. The cloud applications performance will be assessed in different testing environment to assure good services quality as mentioned in SLA document.

The time correlations between the applications metrics will be reviewed by using data mining or data analytic algorithms. These correlations will be obtained to know the exact relation between the performance metrics for different cloud applications, and this is the second expected output from this proposed research work.

In the first set of experiments, the framework for SLA assurance in Figure 4 developed with seven steps as training steps and two steps as operational steps. Training steps are used to build a knowledge base for, SLA assurance and the assessment for cloud applications. Operational steps are used to identify the failures and propose a mitigation technique to mitigate the failure and prevent SLA violations.

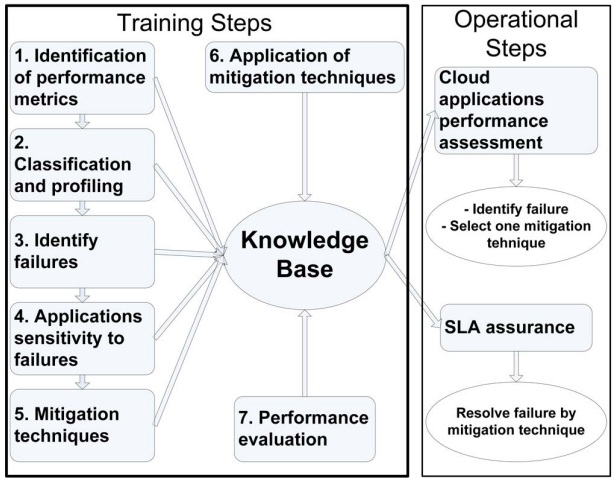

Fig. 4: Our Framework to assure SLA between cloud providers and cloud customers.

V. Conclusions

In this proposal, we presented our conceptual framework and methodology for providing SLA assurance and preventing SLA violations by assessing the cloud applications performance which provided by the cloud services providers to the cloud users. Our framework also assesses the cloud applications performance after identifying their performance metrics. It also provides a mitigation technique to mitigate failures which may occur inside the cloud data centers. During this proposed research work, the time correlation between different cloud applications metrics will be provided.

Acknowledgment

Moreover, the authors would like to acknowledge the funding from National Research Fund (FNR), Luxembourg in the framework of ECO-CLOUD (C12/IS/3977641), Green@Cloud (INTER/CNRS/11/03), and IShOP (POLLUX/ 13/IS/6466384) projects.

References

[1] M. Ali and M. H.Miraz, "Cloud Computing Applications," *Internaional Conference on Cloud Compuing and eGovernance*, 2013.

[2] P. Patel, A. Ranabahu, and A. Sheth, "Service Level Agreement in Cloud Computing," *The Ohio Center of Excellence in Knowledge-Enabled Computing (Kno.e.sis)*, 2009.

[3] T. Wood, P. Shenoy, A. Venkataramani, and M. Yousif, "Sandpiper:Black-box and gray-box resource management for virtual machines," *Computer Networks*, vol. 53, no. 17, p. 2923:2938, 2009.

[4] A. Roy, R. Ganesan, and S. Sarkar, "Keep It Moving: Proactive workload management for reducing SLA violations in large scale SaaS clouds," *IEEE 24th International Symposium on Software Reliability Engineering (ISSRE)*, 2013.

[5] S. S. Wagle, "SLA Assured Brokering (SAB) and CSP Certification in Cloud Computing," *IEEE/ACM 7th International Conference on Utility and Cloud Computing*, 2014.

[6] Z. Shen, S. Subbiah, X. Gu, and J. Wilkes, "Cloudscale: elastic resource scaling for multi-tenant cloud systems," *in Proc. SOCC. ACM*, p. 5:19, 2011.

[7] Salman A. Baset, "Cloud SLAs: Present and Future," *IBM Research, ACM SIGOPS Operating Systems Review*, vol. 46 Issue 2, pp. 57–66, July 2012.

[8] K. Bilal, S. U. R. Malik, and S. U. Khan, "Trends and Challenges in Cloud Datacenters," *IEEE CLOUD COMPUTING SOCIETY*, 2014.

[9] Ranjithprabhu.K and Sasirega.D, "Eliminating single point of failure and data loss in cloud computing," *International Journal of Science and Research (IJSR)*, vol. Volume 3 Issue 4, p. 2319 : 7064, (April 2014).

[10] P. Gill, N. Jain, and N. Nagappan, "Understanding Network Failures in Data Centers: Measurement, Analysis, and Implications," *ACM, SIGCOMM, Toronto, Ontario, Canada*, August 15-19, 2011.

[11] P. Dhar, "Cloud computing and its applications in the world of networking," *International Journal of Computer Science*, vol. 9, issue 1, no. 2, p. 1694:0814, January 2012.

[12] D. Kliazovich, P. Bouvry, and S. U. Khan, "GreenCloud: A Packetlevel Simulator of Energy-aware Cloud Computing Data Centers," *Journal of Supercomputing*, vol. 62, no. 3, p. 1263:1283, 2012.