Service Level Agreement for Distributed Services: A Review

Mohammed Alhamad, Tharam Dillon, Elizabeth Chang

Digital Ecosystems and Business Intelligence Institute (DEBII)
Curtin University of Technology
Perth, Australia
Mohammed.Alhamad@postgrad.curtin.edu.au
Tharam.Dillon@cbs.curtin.edu.au
Elizabeth.Chang@cbs.curtin.edu.au

***Abstract*— Cloud computing has change the strategy of the way of providing distributed services for many business and government agents. Cloud computing delivers a scalable and on demand services for most users in different domains. This new technology brings many challenges to service providers and customers especially for users who already own complicated legacy systems. This paper examines challenges related to the concepts of trust, SLA management, and cloud computing. We focus on SLA definition in cloud computing to achieve the aim of presenting a clear structure of SLA for cloud users and improve the way of building trustworthy relationship between service provider and customer. In this paper, we start with the presenting the importance of cloud computing and the need of SLA for cloud computing. Then, survey of cloud computing architecture is provided. Then, we discuss existing frameworks of service level agreements in different domains such as web services and grid computing. The last part of literature review discusses advantages and limitations of performance measurement models in SOA, distributed systems, grid computing, and cloud services. Finally, we summarize and conclude our work.**

Keywords-trust, cloud computing, SLA

I. INTRODUCTION

The research has been active in cloud computing since late of 2007. Before cloud term there was grid technology. Now, the hot topic of research is cloud and more proposed frameworks and models of various solutions of new technology have started to applied on the cloud architecture. In this section, we survey literature to fine the most proper definition of the term of cloud computing. Also, we review the different architecture frameworks and the common challenges that may be the major problems for providers and customers who are interesting to understand this type of distributed computing.

As shown in Figure 1, the Google trends report shows cloud computing has surpassed grid computing on late 2007.

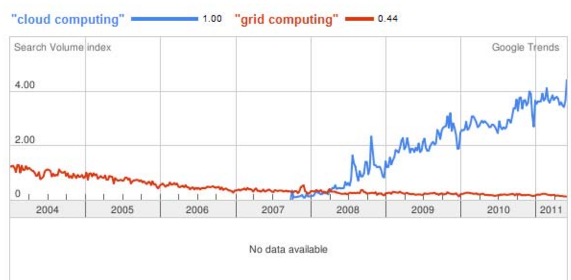

Figure 1. Cloud computing trend

II. DEFINTION

Experts and developers who working to investigate issues and standards related to cloud computing are not from the same background of technology. In the research projects, professionals from grid technology, SOA, business, and other domains of technology and management domains proposed several concepts to define the cloud computing. These definitions of cloud computing still need to be presented in a common standard to cover most technology and other aspects engaged in cloud computing architecture.
In the context of networking and communication, term of "cloud" refer to the metaphor for the common internet concept [1]. The cloud symbol also used to present the meaning of network connection and the way of how the cloud technology is provided by internet infrastructure. "Computing" in the context of cloud domain refers to the

978-0-7695-4612-4/11 $26.00 © 2011 IEEE
DOI 10.1109/DASC.2011.172

technology and applications that are implemented in the cloud datacenters [2].
In [3], Vaquero et al present the lack of common definition of cloud computing. They state that developers and business decision makers confuse of understanding the technology and features of cloud datacenters. So, large budget may be allocated to implement private or even public cloud datacenters. However, these datacenters face several problems when users or public customers want to connect their interfaces of legacy systems with new technology of cloud architecture. Vaquero et al link the challenge of maximizing the revenue of building cloud technology to professionals who involved in distributed services. Because they came from traditional computing domain, they have been confused about the other concepts of distributed services such as grid, and web services. The definition used in [3] which is as follow:

"Clouds are a large pool of easily usable and accessible virtualized resources (such as hardware, development platforms and/or services). These resources can be dynamically re- configured to adjust to a variable load (scale), allowing also for an optimum resource utilization. This pool of resources is typically exploited by a pay-per-use model in which guarantees are offered by the Infrastructure Provider by means of customized SLAs"

Although, this definition presents the main features of cloud computing, important components of cloud architecture included the method of establishing and managing network, applications, and supporting services are not provided in the above proposed definition.

Wang [4] defies cloud computing as:

"*A computing Cloud is a set of network enabled services, providing scalable, QoS guaranteed, normally personalized, inexpensive computing infrastructures on demand, which could be accessed in a simple and pervasive wa".*

Wang presents cloud definition with focusing on the technical aspects of services. Business and functional characteristics are missed in the proposed definition. On other hand, Gruman and Knorr [5] they explain the main technical concepts for cloud services model and define cloud computing from developers point of view. Authors show how the cloud computing architecture takes advantages from the way of implementation of different distributed services, mainly web services and SOA. Two types of cloud services presented with this definition, they define SaaS, and PaaS. Despite of the importance of IaaS as a main component of cloud architecture, they do not give enough discussion about this type of cloud delivery model.

In this research, we adopted and considered the detention provided by U.S. NIST (National Institute of Standards and Technology) [6] that define cloud computing as "Cloud computing is a model for enabling convenient, on demand network access to a share pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management afford or service provider interaction" [6].

Following table conclude the scope and list the main shortcoming of definition discussed above:

Table 1. Conclusion of cloud definitions

Reference	The scope of definition	Missing
Vaquero [3]	Define architecture and service model	Management, supporting, and trust concepts
Wang [4]	Technical concepts	Business and functional characteristics
Gruman [5]	Comparison between cloud computing and web services, and SOA	Definition of IaaS, and DaaS
Mell [6]	Technical features, management, and security concepts	Costing and billing model

III. SERVICE LEVEL AGREEMENTS

I. SLAs for web services

Several specifications for defining SLAs have been proposed for web services. WSLA language [10] introduces a mechanism to help users of web services configuring and controlling their resources in order to meet the service level. Also, the service users can monitor SLA parameters at the run time and report any violation of the service. WSLA, which was developed to describe services in three categories, which are: 1) Parties: in this section, information about service consumers, service providers, and agents are described. 2) SLA parameters: in this section the main parameters which are measurable parameters are presented in two types of metrics. The first is resource metrics, a type of metric used to describe service provider's resources as row information. The second one is composite metrics. This metrics is used to represent the calculation of the combination of information about a service provider's resources. The final section of the WSAL specification is Service Level Objective (SLO). This section is used to specify the obligations and all actions when service consumers or service providers do not comply with the guarantees of services. WSLA provide sufficient level of online monitoring and contracting method but does not define a clear way of which level of service can be considered a violation. WSOL [11] is a service level specification mainly designed to specify different objectives of web services. Defining concepts of service management, cost and other objectives of services can be presenting in WSOL. WSOL is not enough to be used with the objectives of new paradigm of cloud computing.

WS-Agreement [12] is created by Open Grid Forum (OGF) in order to create an official contract between service consumers and service providers. This contract should specify the guarantees, the obligations and penalties in the case of violations. Also, the functional requirements and other specifications of services can be included in the SLA. There are three main sections for WS-Agreement: name, context, and terms. A unique ID and optional names of services are included in the name section. The information about service consumer and service provider, domain of service, and other specifications of service is presented in the context section. Terms of services and guarantees are described with more details in the terms section. These types of online agreements were developed for use with general services. For cloud computing, service consumers lack more specific solutions for SLA to present the main parameters of the visualization environment; at the same time these solutions should be dynamically integrated with the business rules of cloud consumers.

The primary shortcoming of these approaches is that they do not provide dynamic negotiation, and various types of cloud consumers need a different structure of implementation of SLAs to integrate their own business rules with the guarantees that are presented in the targeted SLA.

II. SLAs for grid computing

In the context of grid computing, there are number of proposed specifications which are developed specially to enhance different dimensions related to security and trust for grid services. In [13] a SLA based knowledge domain has been proposed by Sahai to represent the measurable metrics for business relationship between all parties involved in the transaction of grid services. Also, author proposed a framework to evaluate the management proprieties of grid services in the lifecycle. In this work, business metrics and managing evaluation framework are linked to produce an estimated cost model for grid services. In our research, we extend this approach to build general costing model based on the technical and business metrics of cloud domain. The framework proposed in this work lack to the dynamic monitoring technique to help service customers know who is the taken the responsibility when some service level are not provided as they stated in SLA documents. Leff [14] provides a study about the main requirements to define and implement SLAs for grid community. The ontology and details definition of grid computing is provided. Then, scientific discussion id presented about the requirements that can help developers and decision makers deploying a trusted SLAs in grid community. Author implemented basic prototype in order to validate the using of SLAs as a reliable technique when grid service provider and customer need to build trustworthy relationships. The implementation of framework in this study does not consider important aspects of security and trust management in grid computing. Keung [15] Proposed a SLA-based performance prediction tool to analyze the performance of grid services. Keung uses two sources of the information which are the main inputs for the proposed model. The source code information and hardware modeling are used to predict the value of performance metrics for grid services. The model proposed by Keung can be used in other type of distributed computing. But in the cloud environment, this model may not be integrated with dynamic price model of cloud services. It need to be improved with using different metrics for cost parameters to reflect the actual price of cloud services. The system proposed by Padget in [16] considers the response time of applications in the grid systems. The main advantage of the proposed system that it can predicts the CPU time for any node in the grid network before conducting the execution. Padget tested the adaptation SLA model using real experiment on the grid, the prediction system gives values for response time closely to the values when users execute the same application on the grid. Noticed delay recorded for the large size of executed files, the author referee the reason for delay to the external infrastructure such as internet connections. The author also discusses the impact of the time delay caused by external parties to the reputation of service providers in case of using SLA management systems. Although, author provides good method to calculate the response time for grid resources, other metrics such as security and management metrics are missed in this work.

III. SLAs for cloud computing

The context of this research focuses on service level agreement management in cloud communities. In above sections, we present frameworks and models in the literature that mainly designed for managing SLAs in traditional distributed systems. In this section, SLAs and approaches of agreements negotiations in the cloud community are presented.

Valdimir [17] describes the quality of services related to cloud services and different approaches applied to map SLA to the QoS. Services ontology for cloud computing is presented in order to define service capabilities and the cost of service for building general SLAs framework. The proposed framework does not consider all types of cloud services, it is general and it tested on the Amazone EC2 only which need to be considering also other type of cloud providers such as PaaS, DaaS, and SaaS. Our framework in this research considers this issue in the validation phase of research. We conduct the experiments related to the SLA solution chapter in different platform of cloud providers. More details are provided in chapter 6. The framework developed by Hsien [18] focuses on software as a service model of delivery in cloud computing. More details provided on how the services can be integrated to support the concept of stability of cloud community especially for SaaS.

IV. CONCLUSION

In the literature of cloud computing, there is no definition that can be used to provide the clear meaning of all components involved in the architecture of this new paradigm of distributed services. Existing definitions of cloud concepts lack for a comprehensible taxonomy can be easy to understand by professional and business in the cloud community. Service level agreements act the main method to assure the targeted level of quality of cloud services when customers want to move from the traditional infrastructure to the cloud datacenters. However, there is no standard for the implementation of SLAs developed for cloud models. We recommend as a future work to who are interesting about the design and implementation of SLAs in distributed systems to consider following aspects of research issues. First, the cloud community need a clear SLAs structure based on the predefined ontology of cloud computing. Second, the proposed SLAs should be linked to the actual parameters of quality of services and the cost model in order to provide an acceptable framework for cloud customers. Finally, SLAs for cloud computing lacks to reliable tools of simulation and analysis for testing and validation of studied related to SLAs for cloud computing.

REFERENCES

[1] H. Katzan Jr,"On An Ontological View Of Cloud Computing," Journal of Service Science(JSS), vol. 3, 2011.

[2] D. C. Wyld, Moving to the cloud: An introduction to cloud computing in government: IBM Center for the Business of Government, 2009.

[3] L. M. Vaquero, et al., "A break in the clouds: towards a clouddefinition,"ACM SIGCOMM Computer Communication Review, vol. 39, pp. 50-55, 2008.

[4] L. Wang, et al., "Cloud computing: a perspective study," New Generation Computing, vol. 28, pp. 137-146, 2010.

[5] E. Knorr and G. Gruman, "What cloud computing really means," InfoWorld, vol. 7, 2008.

[6] P. Mell and T. Grance, "Draft nist working definition of cloud computing," Referenced on June. 3rd, 2009.

[7] R. Buyya, et al., "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility," Future Generation Computer Systems, vol. 25, pp. 599-616, 2009.

[8] R. Jennings, Cloud Computing with the Windows Azure Platform: Wrox, 2010.

[9] C. Hoefer and G. Karagiannis, "Taxonomy of cloud computing services," 2010.

[10] H. Ludwig, et al., "Web service level agreement (WSLA) language specification," IBM Corporation, 2003.

[11] V. Tosic, WSOL Version 1.2: Carleton University, Dept. of Systems and Computer Engineering, 2004.

[12] A. Andrieux, et al., "Web services agreement specification (WS-Agreement)," 2004.

[13] A. Sahai, et al., "Specifying and monitoring guarantees in commercial grids through SLA," 2003.

[14] A. Leff, et al., "Service-level agreements and commercial grids," IEEE Internet Computing, vol. 7, pp. 44-50, 2003.

[15] H. N. L. C. Keung, et al., "Self-adaptive and self-optimising resource monitoring for dynamic grid environments," 2004, pp. 689-693.

[16] J. Padgett, et al., "Predictive adaptation for service level agreements on the grid," International Journal of Simulation: Systems, Science and Technology, vol. 7, pp. 29–42, 2006.

[17] V. Stantchev and C. Schröpfer, "Negotiating and enforcing qos and slas in grid and cloud computing," Advances in Grid and Pervasive Computing, pp. 25-35, 2009.

[18] C. H. Wen, et al., "A SLA-based dynamically integrating services Saas framework," pp. 306-311.