Cloud Service Level Agreements –Issues and Development

Isaac Odun-Ayo
Covenant University, Nigeria
isaac.odun-ayo@covenantuniversity.edu.ng

Olasupo Ajayi
University of Lagos Nigeria
olaajayi@unilag.edu.ng

Nicholas Omoregbe
Covenant University, Nigeria
nicholas.omoregbe@covenantuniversity.edu.ng

***Abstract*—Cloud computing is a broad paradigm that has influence across major fields of human endeavour. The unique services it offers makes organisations curious about understanding the cloud and its likely benefits. The cloud offers services such as custom built applications deployed on remote systems and ready to use platforms which reduce the efforts needed to develop and deploy applications for cloud users. In addition to these, there are other services such as storage and infrastructural resources which the cloud also avails to its users. These services are usually provided to users on a pay-per-use bases, thus necessitating the need to have documented agreements in place to ensure a smooth relationship between the providers and the users. These documented agreements are referred to as Service Level Agreements (SLAs). SLAs detail the terms, conditions and service expectation of the users from their service provider in terms of availability, redundancy, uptime, cost and penalties for violations. These ensures users' confidence in the services being offered. In this paper, the state of the art with respect to cloud SLAs is presented. The paper seeks to answer questions related of what the current trends and developments in terms of cloud SLA are and it does so by means of a review of existing literature available. This paper therefore is a survey of cloud SLAs, their issues and developmental challenges. It provides a guide for future research and is expected to benefit prospective cloud users and cloud providers alike.**

Keywords-Cloud computing; Cloud providers; Cloud users; Quality of Service; Service Level Agreements

I. INTRODUCTION

"Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" [1]. Cloud computing is the provision of services by a cloud service provider to facilitate computing, storage and applications development by a user utilizing the Internet. Cloud computing can also be defined as a computing model for permitting omnipresent, suitable and on-demand service access to a common group of configurable computing resources such as networks, services, storage and applications that can be quickly provisioned and released with minimum management involvement [2]. Cloud computing has three basis service types, Software-as-a-Service (SaaS), Platform-as-a-Service (PaaS) and Infrastructure-as-a-Service (IaaS). In SaaS, application are made available over the Internet for cloud users by cloud providers. Users do not need to install and configure applications for usage. Applications can be accessed anywhere and at any time over the Internet. In PaaS, a platform is provided for cloud users to create and deploy own applications using application programming interfaces provided by the cloud provider. In this model, the user has control over the application being developed and deployed. In the IaaS model, infrastructure is provided and controlled by the cloud provider. Cloud users are provided with networks, servers, storage, computing and other resources on-demand and at a cost. The cloud provider controls the entire infrastructure, while the user manages the operating system, applications and storage.

Cloud computing utilizes the concept of virtualization and multi-tenancy to provides service to users. Enterprises can migrate some of their application and utilize services based on suitable agreements. Cloud computing offers four deployment types. The public, private, community and hybrid clouds. The private cloud is owned by an organization. The facilities may be on premise or off – premise and can be managed by a third party. Private clouds are considered more secured. Public clouds are owned by major service providers offering services at a cost to cloud users. Public cloud utilizes large data centres sometimes across several geographical locations. They are considered less secure. Community cloud are owned by several organizations who have a shared common interest. It could be managed by that community or a third party. Hybrid cloud is a combination of either private, community or public cloud. Hybrid cloud leverages on the advantages of the different cloud types.

Service Level Agreement (SLA) is defined as a formal contract between service providers and consumers to guarantee that consumer's service quality expectation can be achieved [3], SLA lifecycle has six steps which are: discover the service providers, define the SLA, establish an agreement, monitor likely SLA violators, terminate SLA and enforce penalties [3]. SLAs embody key elements to achieve full success in cloud computing since they represent the desired guarantees between service providers and customers [4]. SLAs are used to formally describe the functions being offered, the Quality of service (QoS) expected from the provider, responsibilities of both parties and likely penalties [4]. According to [8], SLAs should contain five basic items

which are: set of services being offered by the providers, an unambiguous and future proof terms of service, set of QoS metrics for measuring service delivery levels and a means of monitor these metrics and means of resolving disputes arising from failure to meet the SLA terms. It is important for cloud users to enjoy the resources and service guaranteed by the cloud provider, while the cloud provider should benefit optimally from the resources being provided [6].

According to NIST, a cloud SLA is "a document stating the technical performance promises made by the cloud provider, how disputes are to be discovered and handled and any remedies for performance failures". Clearly, an SLA is essential for a successful relationship between a cloud provider and cloud user. No serious enterprise will want to lose valuable data migrated to the cloud and no viable provider will want to lose patronage because of failed guarantees. Although, cloud users do not have full control over the computing resources made available by the provider, they should enjoy attributes such as quality, accessibility, trustworthiness and performance of these resources when users have transferred some of their business functions to the cloud providers [2].

The purpose of this paper therefore is to discuss cloud computing and SLA. The paper will discuss SLA architecture and the challenges. Thereafter, it examines issues relating to cloud computing and SLA from the perspective of industry. This paper contributes to further understanding of SLA in cloud computing. The rest of the paper is as follows: Section 2 examines related work. Section 3 discusses the SLA architecture and challenges. Section 4 focuses on highlights of SLA in terms of cloud computing based on industry perspective. Section 5 concludes and suggests future work.

II. Related Works

The authors in [3] presented an SLA-Based Admission Control for a Software-as-a-Service Provider in Cloud Computing Environments. In this work, an algorithm determining customer satisfaction for SaaS cloud model was presented. The crux of the work was to develop a means of maximizing profit by saving virtual machine cost through a process formed by conducting extensive analysis. A dynamic data-driven simulation approach for preventing SLA violations in federated cloud environments was presented in [6]. The work presented an architecture for enabling the release of consumer resources without issues in terms of the agreement. This was achieved using multiple cloud service providers being utilized by a consumer. Service Level Agreement in Cloud Computing was presented in [7]; wherein the authors examined service level agreements between consumers and providers. A Web SLA framework was then proposed, implemented and validated. In [8], a framework for negotiating Service Level Agreement of Cloud-based Services was presented. It was a framework that helped decided the most suitable cloud service provider for a cloud user. The authors also highlighted the important roles of cloud brokers for optimum resource utilization by the consumer. A capacity driven utility model for SLA negotiation of cloud services was presented in [9]. The work proposed a dynamic system which ensured that cloud users effectively utilized resources provided by their cloud provider. The authors noted that customers' requirement and utilization of cloud resource were relevant criteria to be considered when drawing up SLAs that are expected to satisfy both parties. The role of governance and other SLA issues in cloud environment were presented in [10]. The work presented key attributes to be considered when drawing up service contracts in cloud computing. The issue of cloud governance was also examine with a view of determining suitable means of managing information with minimum consequences. The authors in [11], presented a pricing strategy for cloud computing. They proposed a hybrid strategy for cloud pricing services between a vendor and consumer. The approach was to use service interruptions as a determinant in the relationship between the cloud users and the provider in terms of services to optimize profit. The issues and challenges faced when providing QoS and SLAs in cloud environments were discussed in [12]. The issues that could arise from the service providers optimizing users' workloads in view of ensuring adequate quality of service were examined. The authors then proposed an architecture for providing SLA between the cloud providers and users. In [13], the authors proposed a model that simultaneously catered for QoS adherence and resource utilization in cloud data centres. The authors proposed an approach which grouped user workloads into categories and then used these categories to prevent SLA violation and also improve resource utilization, thus benefitting both the cloud users and providers. The NIST cloud computing standards roadmap [2] described the role of the cloud broker(s) and identified the broker(s) as a critical aspect of cloud SLA. An Inter Cloud SLA brokering service was presented in [14]. The work discussed inter-cloud environments, which allowed running of applications across different cloud platforms. The paper then proposed a cloud brokerage that allows user find suitable provider(s) for their applications. The authors validated the proposed framework via implementation. Cloud SLA considerations for the government consumer in [4] presented a detailed discussion on SLA. Several aspects of SLA was examined including cloud service brokers. The SLA evaluation model for cloud computing in [5] proposes a new method of developing SLA. As the cloud continues to evolve so also should the SLA, hence the paper presents an optimized SLA that is more suitable for use by both the provider and consumer alike. A formal model for SLA negotiation was proposed in [17]. The authors presented an agent-based multiple round SLA negotiation model. The proposed model was based on Web Service Agreement but targeted at SLA management for Cloud services provided by multiple parties. The model incorporated multiple runs, multiple providers and multiple negotiation round. At each negotiation round a happiness factor was computed. The happiness function was used to represent the level of compromises each party (user agent and CSPs agents) could make and was used to determine the point at which satisfactory SLA could be made. The work assumed that QoS requirements of each party could be prioritized and that

these parties were willing to make compromises on certain QoS requirements.

III. Cloud Computing and Service Level Agreement

A. Service Level Agreement Concepts

1) NIST

The NIST in its publication on cloud computing synopsis and recommendations as described in [16], viewed SLA from one perspective which is from the service provider. According to this report, a typical commercial cloud SLA should consists of:

a. These promises should explicitly contain promises made to the users (concerning service availability, dispute resolution steps, data preservation and legal protection of users' information).
b. The providers service limitations resulting from effect of acts of nature (natural disasters) which are outside the providers' control, service outages and updates. The providers are however required to give reasonable notice to their users.
c. The expectations from cloud users, which includes acceptance of terms and conditions as well as payment for services used.

This NIST's view to SLA is quite rigid and skewed in favour of the Cloud providers. For instance it does not take into consideration the option for users to discuss modification to service agreements with service providers if the default SLA terms do not address all of the users' needs. The report also implores users to be aware that SLAs may be changed at the providers' discretion (though with reasonable advance notice) thus be ready to migrate workloads to alternate providers if the changes are unacceptable. This however is not an easy task, as vendor lock-in and a lack of standard which could allow interoperability between cloud providers is still an open cloud computing issue.

Lastly the NIST's report assumes a blanket approach to SLA agreement for all Cloud services. In [5] however multiple SLAs were presented based on the cloud service model being subscribed. The authors also specified two types of SLA for each service model – the provider SLAs and the user SLAs. These contained the promises and expectations from the service provider and the users respectively.

2) ISO / IEC

International Organization for Standardization and the International Electro-technical Commission (ISO / IEC) defines cloud SLA as service level agreement between a Cloud Service Provider (CSP) and a cloud consumer based on a taxonomy of cloud computing specific terms to set the quality of services delivered. Cloud SLAs have business and technical properties and cover terms regarding the quality of service, security, performance and remedies for failure to meet the terms of the SLA. A CSP can also list within the cloud SLA a set of promises explicitly not made to cloud service consumers. That is, limitations and obligations that cloud service customers need to accept.

3) Cloud Service Council

The Cloud Service Council (CSCC) has defined Cloud Service Agreement (CSA) as having three major parts – a customer agreement, an Acceptable Use Policy (AUP) and an SLA. The CSA describes the overall relationship between the customer and the provider, while the AUP details activities which the provider considers to be illegal or inappropriate. The SLA describes the level of technical performance i.e. availability, serviceability, or performance associated with the service being rendered to the users.

In [11], SLA is described as a document that specifies the terms and conditions between the user and CSP. The SLA indicates minimum performance level that the CSP has to provide, counteractive actions and the consequences in case of breach of the agreement between user and CSP. The user must be clear about security requirements for their assets and all the requirement should be thoroughly agreed upon in the SLA.

Monitoring and enforcing these contractual agreement is often a challenge. This is because the service statistics are usually provider by service provider and the user can not totally rely on these as it might have been tampered with. In most cases both the user and the service provider keep statistical records of service. Unfortunately in instances of conflict between the CSP and the user statistics, evaluation of the statistics and determination of responsibility becomes an issue. Though cloud arbitrators might be employed in cases where numerical metrics are available. However for cases relating to security, it is even more challenging as often times the "deed has already been done" and an audit about the security provided by the CSP is harder to carry out and even agreed upon in the first place; as security is an ever dynamic issue.

B. Service Level Agreement Levels

Some of the significant levels associated with cloud SLAs are discussed:

a. Facility level SLA: At this level, the CSP will deliver on SLA covering the data centre services necessary to maintain the customer owned information and/or applications. These comprise items such as electric power, onsite generator and cooling among others. SLAs would cover high availability, fault tolerance and data replication.
b. Platform Level SLA: This level entails physical servers, virtualization infrastructure, and network related hardware owned by the provider and used by the cloud consumers. SLA at this level would include information about physical security that denies unauthorized access to the building, facility, and resource. Background checks and character analysis of staff should also be done prior to recruitment by the service provider.
c. Operating system level: At this level, providers normally deliver some amount of managed services to the users. This extra service permits the provider to guarantee that the OS is suitably sustained so that it is dependably accessible. SLAs would contains information about security updates, system patches,

confidentiality / encryption, user authorization and audit trails.

d. Application level SLA: This category delivers safety against application level data catastrophes up to and comprising the customer application being executed on the infrastructure provided by the provider. Here, the cloud provider is ensuring the availability, stability and performance of their cloud user software which they are hosting. This is often times difficult to guarantee particularly in IaaS and PaaS wherein the user is solely responsible for the application they put on the cloud.

C. Cloud Provider Service Level Agreement

A typical SLA of a cloud service provider comprises the following:

a. Service Assurance: This is a metric which specifies the service level which a provider must meet over an agreement time period.
b. Service Assurance Time Period: This describes the duration over which a service guarantee should happen. The time period can be a billing month, or as agreed upon by both parties.
c. Service Assurance Granularity: This defines the resource scale on which a provider specifies a service guarantee. For example, the granularity can be as per service, per data centre, per instance, or per transaction basis. Related to time period the service assurance can be inflexible if granularity of service assurance is fine–grained. Service assurance granularity can also be designed as a cumulative of the deliberated resources such as contacts.
d. Service Guarantee: Omissions are instances that are excluded from service guarantee metric calculations. The omission typically include misuse of the system by a customer or any downtime associated with the scheduled maintenance.
e. Service Recognition: This is the account credited to the customer or applied towards upcoming expenditures if the service assurance is not met. The amount can be a comprehensive or restricted recognition of the consumer compensation for the miscalculated service.
f. The service violation measurement and reporting: This describes how and who measures and reports the violation of service assurance respectively. This is usually done by the users but the provider could also take up the role. In some instance an independent service monitoring third party might be responsible for this.

D. Service Level Agreement Benefits and Challenges

1) Service Level Agreement Benefits

It is imperative to have detailed agreement between the consumer and provider to ensure trust and confidence on both sides. The following are some of the benefits of having a cloud SLA according to [10]:

a. It enables strong understanding of the service and accountabilities of all parties.
b. It helps the customer to achieve their view points.
c. It encourages clearness, responsibility and reliability.
d. Provides information on team performance, capabilities and staffing judgment.
e. There is provision of supportive and collective functioning.

2) Service Level Agreement Challenges

Every agreement has some challenges that both parties must contend with. Some of these challenges are as follows:

a. SLAs are hard to accomplish in cloud computing because certain infrastructure and circumstances such as network and force majeure are beyond the control of both the cloud provider and customer, hence difficult to draw up an SLA for.
b. In situations where multi-tenancy is employed, SLAs pertaining to service isolation and high availability might be difficult to for the service provider to guarantee.
c. For cloud SaaS model, SLAs are difficult to accomplish because it is nearly impossible for the service provider to test every possible user software/application with varied system configuration beforehand.
d. It is very difficult to agree on a cloud SLA which covers security. Distributed Denial of Service DDoS, keystroke timing and side-channel attacks have been identified as some of the most common attacks in cloud environment [15] and still remain an open challenge. Due to the dynamic and ever changing nature of these attacks, a service provider can only at best give a generalized security-based SLA and might not even be able to guarantee it.
e. In distributed and multi-cloud application deployment, SLAs are difficult to agree on as the various providers implement varied standards. These standards are often times proprietary and not interoperable hence a single SLA for all parties involved might be difficult to achieve.

IV. Typical Service Level Agreement Offered by Cloud Service Providers

A. Amazon Web Services

AWS [18] is a cloud service provided by Amazon and offers services such as the Elastic Cloud Compute (EC2) and Simple Storage Service (S3). In EC2, a customer can obtain virtual machine instances by the hour or reserve them in advance for an entire year. In addition, EC2 offers spot instances where a customer can bid for compute capacity. The storage service S3 provides mechanism for storing and retrieving data object using put (), get () operating with data size ranging from one byte to five terabytes. Amazon EC2 and S3 service are backed by distinct SLAs [14] and summarized as follows:

a. For Amazon's EC2
   i. SLAs are defined per data centre and on regional basis.

ii. 99.95% regional service availability guarantee with users redirected to other regional after a 5 minutes of waiting.
iii. Users are responsible for monitoring SLA metrics and report same to Amazon.
iv. Dispute resolution and penalties are in terms of service credit rebate for failure of individual instances as a result of region unavailability.

b. For Amazon's S3:
i. A 99.9% storage service request completion rate is guaranteed
ii. Users are responsible for monitoring and reporting service failures.
iii. In terms of dispute resolution and penalties, a service credit rebate of 10% of the customer's bill is paid to the customer if completion rate is below 99.9% per month and 25% if completion rate is less than 99%.

B. Windows Azure

Windows Azure [19] is a PaaS and IaaS cloud service provider that offers Azure compute and Azure storage services. Azure compute comprise three services called roles, namely: web, worker and a VM. A web role provides a web based front end for applications, the worker is used for application deployment, while the VM provides a virtual infrastructure to the user. Azure storage on the other hand is similar to Amazon's S3 storage. The SLAs for these respective services are as summarized:

a. For Azure Compute
i. Azure SLA defines two service guarantees: external network connectivity and uptime which are calculated on monthly basis.
ii. 99.5% and 99.9% service availability rate is guaranteed.
iii. Users are responsible for monitoring and reporting service failures.
iv. Disputes are resolved by offering users a 10% service credit rebate if connectivity and uptime percentage is below 99.5% and 99.9% respectively.

b. For Azure Storage:
i. A 99.9% storage service request completion rate is guaranteed.
ii. Users are responsible for monitoring and reporting service failures.
iii. For disputes and penalties, a service credit rebate of 10% of the customer's bill is paid to the customer if completion rate is below 99.9% per month and 25% if completion rate is less than 99%.

Table 1 shows a concise comparison of cloud storage SLAs of three cloud providers – Amazon, Azure and Rackspace.

TABLE I. COMPARISON OF CLOUD STORAGE SLAS [14]

Service Provider	**Amazon S3**	**Azure Storage**	**Rackspace Cloud Files**
Service guarantee	Completed transactions (with no error response)	Completed transactions (within stipulated time)	Completed transactions, availability
Service granularity guarantee	Per transaction	Per transaction	Per transaction, data center
Service guarantee time period	Billing month	Billing month	Per month
Service credit	10% of bill if < 99.9%, and 25% if < 99%		10% of bill if < 99%, 100% if < 96.5%
Service violation reporting onus	Customer		
Service violation incident reporting	N/A	Within days of incident occurrence	N/A
Service violation claim filing	Within 10 business days following the month in which the incident occurred.	Within one billing month	Within 30 days following unavailability
SLA publish date	October 1, 2007	November 12, 2010	June 23, 2009
Credit applied towards future payments only	Yes	No	

V. CONCLUSION

Cloud computing provides scalable, on demand, elastic, multi-tenant and virtualized services to customers over the Internet through cloud providers. The service types are the SaaS that provides applications, PaaS that provide platform for application development, and IaaS that provides storages and computing infrastructure to users. These services are provided on the basis of service level agreements between the CSP and the user. The SLA specifies the terms of the services provided for a mutually beneficial transaction between both parties. In this paper, a survey of recent developmental trends and issues in cloud SLA negotiations were presented. The paper concluded with a review of SLAs for major CSPs such as Amazon, Microsoft and Rackspace which provide IaaS and PaaS services to clients was then done.

REFERENCES

[1] P. Mell, T. Grance, The NIST Definition of Cloud Computing, NIST Special Publication 800-145, 2011

[2] M. Hogan, F. Liu, A. Sokol, J. Tong, NIST Cloud Computing Standards Roadmap, NIST Cloud Computing Standards Roadmap – Version 1.0, Special Publication 500-291, 2011.

[3] L. Wu, S. K. Garg, R. Buyya, SLA-Based Admission Control for A Software-as-a-Service Provider in Cloud Computing Environments, Journal of Computer and System Sciences 78, 1280-1299, 2012

[4] K. Buck, D. Hanf, D. Harper, Cloud SLA Considerations for the Government Consumer, 2015 The MITRE Corporation.

[5] W. Chenkang, Z. Yonghua, P. Shunhong, The SLA Evaluation Model for Cloud Computing, Intl Conf. on Computer, Networks and Communication Engineering (ICCNCE 2013).

[6] F. Faniyi, R. Bahsoon, G. Theodoropoulos, A Dynamic Data-Driven Simulation Approach for Preventing Service Level Agreement Violations in Cloud Federation", Intl Conf. on Computational Science, ICCS 2012

[7] P. Patel, A. Ranabahu, A. Sheth, Service Level Agreement in Cloud Computing, https://www.researchgate.net/publication/228343067. 2009

[8] R. El-Awadi, A Framework for Negotiating Service Level Agreement of Cloud-based Services, Intl Conf. on Communication, Management and Information Technology (ICCMIT 2015)

[9] N. Ranaldo, E. Zimeo, Capacity –Driven Utility Model for Service Level Agreement Negotiation of Cloud Services, Future Generation Computer Systems 55 (2016) 186–199

[10] M. Cochran, P. Witman, Governance and Service Level Agreement Issues in A Cloud Computing Environment, Journal of Info. Technology Management Volume XXII, Number 2, 2011.

[11] J. Huang, R. Kauffman, D. Ma, Pricing strategy for cloud computing: A damaged services perspective", Decision Support Systems 78 (2015) 80–92

[12] Z. Mahmood (ed.), Cloud Computing, Computer Communications and Networks, DOI 10.1007/978-3-319-10530-7_3 Springer International Publishing Switzerland 2014.

[13] Ajayi, O., Oladeji, F., Uwadia, C., Multi-Class Load Balancing Scheme for QoS and Energy Conservation in Cloud Computing, *West African Journal of Industrial and Academic Research, 17*(1), pp. 28-36. 2016

[14] F. Jrad, J. Tao, A. Streit, SLA Based Service Brokering in Inter cloud Environments, 2nd International Conference on Cloud Computing and Services Science, 2012.

[15] Mahmood, Z. (Ed.). *Cloud Computing: Challenges, Limitations and R&D Solutions*. Springer. 2014.

[16] L. Badger, T. Grance, R. Patt-Corner, J. Voas, Draft cloud computing synopsis and recommendations. *NIST special publication, 800*, 146, 2011.

[17] Baig R., Khan W., Haq I. and Khan I. Agent-based SLA negotiation protocol for cloud computing. Intl. Conf. of Cloud Computing Research and Innovation, CloudAsia2017, p.5.

[18] Amazon Web Services Inc., Products and Services, 2015. Available online at http://aws.amazon.com/products/?nc2=h_ql_sf_keynote/

[19] Zikos M., Microsoft Azure Load Balancing Services, 2014, Available Online at http://azure.microsoft.com/blog/2014/04/08/microsoft-azure-load-balancing-services/