Need of SLA Parameters in Cloud Environment. An Evaluation

*Shikha Pandey
Manager-Learning,
Trans-Neuron technologies, Bangalore, India.
Shikha.p@transneuron.com

Ashok Kumar Upadhaya
Member of Technical Staff II,
VM Ware private ltd. India, Bangalore, India.
aupadhaya@vmware.com

Dr.C.K. Jha.
Head of Department (Computer Science),
Banasthali University, India.
ckjha1@gmail.com

Abstract- **Management and establishment of Service Level agreement (SLA) in cloud totally depends upon the various parameters which helps in establishment of SLA between the service providers and users. These parameters are the deciding factors for provisioning of services and in formation of SLA metrics. In this paper, we have analyzed probable parameters that can form or act as a SLA for entire cloud system. These parameters vary from lowest level to highest level of computing stack along with the services offered.**

Keywords: SLA, Cloud Computing, Parameters, Metrics.

I. Introduction:

The model of cloud computing mostly operates in a dynamic environment, therefore in order to provide services to unpredictably diverse set of entities, lightweight and real-time decision supporting functionalities and parameters are required. Requirement of SLA parameters are also for the formation and establishment of the functionalities. Prior to the Cloud and in Traditional IT and Web-based enterprise application deployment, the task of administration was easy, since the single important objective was the performance, but the scenario has changed a lot in the recent years. [3] [4].

A need to establish a Service Level Agreement (SLA) for providers, customers and between provider and customer is there in cloud environment. **SLA is defined as a legal contract between participants to ensure:**

(a) Quality of Service(QoS) requirements are met and if any party violates the SLA terms, the defaulter has to pay penalty according to the clauses defined in SLA.

(b) To indicate the profits, the service provider may obtain if the service is delivered at certain levels,

(c) To indicate the penalty, the service provider has to pay if the agreed-upon performance is not met.

(d) There exist many forms of SLAs with different metrics and measurement methods (e.g., measured at a per-customer level or a per query level).

(e) In other words, Service Level Agreement (SLA) is an agreement negotiated between service users and providers, which defines the metrics, expected QoS, and penalties during service delivery. [2] [3][4][7][8][9]

SLA Model: Front end (Users), Back end (Providers and geographically Distributed servers) and the network are the fundamental cloud components which forms or lays the blue print for Cloud architecture**.** [12]. SLA is between the provider and the consumer depending upon the various factors and parameters. Users request for various services from cloud provider depending upon its need. The provider with the help of broker /auditor fulfills the demand.

For establishment of SLA, the providers depending upon the demand negotiates, allocates and manages the SLA. The entire process is depicted in the below figure.

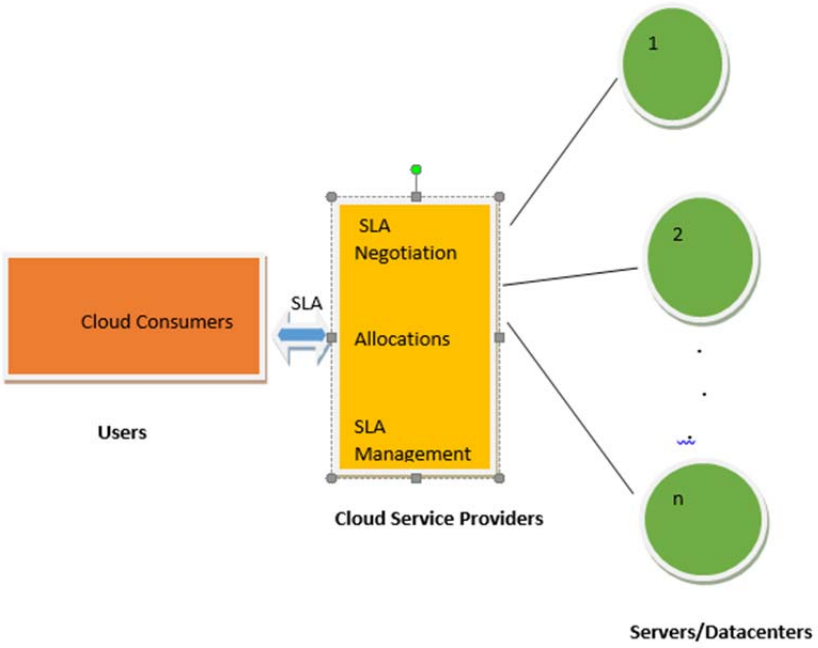

Fig: Service level agreement model.

The model of SLA not only determine the template parameters for providers and users who use different types of cloud service signing contract, but also helps users to select a superior service. The evaluation method based on the SLA model combines assessment from both providers and users. It is obvious that service with high grade has high quality in its own service type [2].

In this paper we will review in detail various SLA parameters that affect cloud environment. The organization of this paper is as follows. Section 2 describes the Measuring strategies; related work is described in Section 3 need for the analyses of parameters is described in Section 4. The parameter and its affect along with the factor is described in Section 5. Finally, concluding remarks are given and possible future work is mentioned in Section 6.

II. Measuring terms

SLA can be measured in terms of: -

- Impact of reservation strategies.
- Impact of QoS parameters.
- Performance Analysis under Uncertainty Future Interest Value.
- Providers
- Customer [2][3][6][7]

III. Survey/Review:

DASTJERD A. et.al have proposed a state-of-the-art solution in order, to automatethe negotiation process for cloud environments with infrastructure as a service as main focus. The negotiation strategy is based on a time-

dependenttactic. For cloud providers, the strategy uniquely considers utilization of resources when generating new offers. The proposed strategy is capable of assessing reliability of offers received from cloud providers. Main focus is on maximizing the profit of cloud provider with respect to deadlinetaking into account reliability metric to discard unreliable offers, and investigating social optimality of time-dependent functions in the negotiation process. [1]

Wu Chenkanget.al in their work have presented an optimized SLA classification model for cloud computing and introduce a corresponding evaluation method for different types of cloud services, which is a basis for both users selecting cloud service and providers scheduling resources. The proposed model not only satisfies the requirements of both cloud service providers and users, but it also makes the evaluation method more fair and reliable. Transformation, inference and grading mechanism has been also included. [2]

The main objective of their work is to minimize cost and SLA violations using resource provisioning strategies. They have worked and proposed thecustomer driven heuristic algorithms which consider customer profiles and provider KPI (key performance index) criteria. i.e. a resource provisioning model for SaaS Clouds considering customer profiles and multiple KPI criteria is proposed. [3]

They have built a novel lightweight data structure called SLA-tree and have also used the SLA metric on query response time. SLA-tree not only encodes the information about the profit situation in the near future, but also is able to efficiently infer the potential profit impact on scheduling, dispatching, capacity planning, and so on which are relevant to cloud computing. A framework is too proposed to efficiently support profit-oriented decisions in database systems in the cloud. [4]

IaaS resources specified by SLAs offer a variety of properties, from which the user can choose the best fit for his/her needs. The main contributions of the work revolve around an SLA with two dimensions, charging model and VM type. It also includes an integer linear program (ILP) to compute the scheduling regarding an SLA with two dimensions. Customers of cloud services choose the VMs profiles (SLAs) offered by the provider, and pay according to how long these VMs are utilized. [5]

A per-service SLA model, by supporting the automatic implementation of cloud Security SLAs tailored to the needs of each customer for specif c service instances is proposed. In particular, the process and the software architecture for per-service SLA implementations detailed. [6]

They are focused on how monitoring of cloud environment take place and it enables how service providers can show, cloud information to customer. Customers inform what kind of services they need and their expectations on the quality of service, which are usually expressed through Service Level Agreements (SLAs). Through monitoring, SPs can show cloud information to customers. [7]

Paul Manuel et.al, have proposed a novel trust management system called QoS model. The work describes how a service level agreement is prepared combining quality of service requirements of user and capabilities of cloud resource provider. It also includes how a cloud resource selection is made by combining its trust and capabilities. [8]

Proposal of an admission control and resource scheduling algorithm is made, the scheduling algorithm not only satisfies QOS requirements of requests as guaranteed in Service Level Agreements (SLAs), but also increases the profit for AaaS (Application as a service) providers by offering a cost effective resource scheduling solution. [9]

They have proposed a CSP ranking model based on service delivery measurements and user experience where the proposed framework is a recommendation system that facilitates selection of the appropriate cloud service provider. Where the recommendations are based on the entire SLA requirements [10]

IV. Need

SLA parameters are required to be established because:

i. Clients and Providers need to establish a Service Level Agreement (SLA) to define the Quality of Service (QoS).
ii. The SLA violation causes penalty.
iii. Existing cloud SLAs focus only on few service terms, completely ignoring all security related aspects.

iv. To satisfy and manage dynamic customer requests and demands.
v. To minimize the total cost and SLA violations for cloud providers.
vi. How to reserve resources by considering the customer profiles and multiple KPI criteria?
vii. To map customer requirements to infrastructure level parameters.
viii. To deal with the infrastructure level heterogeneity (such as different VM types and service initiation time)? [2][3][6]

V. Parameters:

A number of parameters acts as a basis of agreement and governs the entire cloud system. In this section, we take a dig on various parameters from all aspect i.e. from lowest level to highest level of computing stack along with the services offered. Any of these parameters can be grouped to form SLA metrics for entire cloud system. Since cloud is a virtual environment there is number of SLA parameters defined in accordance of it.

S. No.	**Parameter**	**Description**
1	Hard Disk	It is a part of a unit, often called a disk drive that stores and provides relatively quick access to large amounts of data. [1][8]
2	CPU Capacity	It provides the provider with the capacity of CPU. [1][2][8]
3	RAM	A primary storage. [1][8]
4	Resource utilization	Percentage of allocated resources consumed. [1]
5	Resource prices/COST	It includes price with respect to resource and services. The resource includes virtual resource also. [1][3][7]
6	Amounts of requested resources	The number of computing resources in form of services demanded from provider. [1]
7	Availability	Aanywhere and anytime access to services and data is defines as availability. It is holy grail for cloud services as a whole. [1] [2][4]
8	Deadline of requests	It mostly deals with the time period. [1]
9	Initial offer values	The various offers provided by the provider for the services is important condition/parameter in cloud environment. [1]
10	Type of tactic (negotiation parameter)	How to negotiate, and on which factors. It's a parameter generally considered from provider's side. [1]
11	Response time (Service), mean response time, Query response time, Tuning cloud response time	The elapsed time between the demand and the beginning of a response. Mean response time generally governs SLA violations. [2] [3][4][7]
12	Service Initiation Time	The time to initiate a service. of service. in cloud computing it is sum of VM initiation time, application deployment and installation time. [3]

13	Data Transfer Time	The rate at which data is transmitted from one host to another. [3]
14	Delay time	It's also known as a latency. [3]
15	Resource level	A parameter for resources allocation and balancing. [3]
16	Heterogeneity	It is different type of services, resources and characteristics. In Cloud the heterogeneity is with respect to resources, its type, workload and environment. [3]
17	Reservation strategies	It is the different strategy for allocation and scheduling of resources and applications. [3]
18	Throughput	Defined as maximum rate to process anything.2,4,8
19	Security	Security is utmost important parameter in cloud environment. It refers to a broad set of policies, technologies, and controls deployed to protect data, applications, and the associated infrastructure of cloud computing. [4] [2]
20	Timeout/ threshold	Threshold is boundary value measured over the operations result. [4]
21	Slack time	Amount of time a task can be delayed without causing another tasks or an application to be delayed. [4]
22	Level of Redundancy (LoR)	It represents different components and functions. [6]
23	Level of Diversity (LoD)	LOD represents the number of different web server instance. [6]
24	Service characteristics	The features possessed by services to fulfill the demand. [6]
25	Processor load	Capacity of processor. [7]
26	Network usage	Percentage of network utilization. [7]
27	Heterogeneous APIs	Different types of programming interface, needed for provisioning of services. Forms the foundation for IaaS cloud. [7]
28	Energy consumption,	Amount of energy consumed in a process or system. [7]
29	Systems dependability	Measure of a ***system's*** availability, reliability, and its maintainability, along with the maintenance, support and performance. [7]
30	Networking strength	Measured in terms of bandwidth, speed and latency. [2][8]
31	Turnaround efficiency	Measure of efficiency for completion of a process. [8]
32	Data integrity and reliability	Used to compute trust value an important parameter in cloud as a whole. [2][8]
33	Resource configurations	Configuration and allocation of different category of resources. [9]
34	Data processing time	The amount of time required to process an application or service. [9]
35	VM creation time	The time taken for creation of a new VM on the basis of demand. [9]
36	Submission time	It is defined as the submission time for any query or request. [9]
37	Waiting time	Amount of time a process has been waiting for the execution. [**9]**
38	Billing	A general Sla with respect to provider. The parameter is based upon the cost model involved in cloud. [2]
39	Monitoring	To check the progress and review the kind of services over a period of time. [2]
40	Privacy	Protection of information or data. [**2]**
41	Compensation	It is a strategy of reimbursement. [2]
42	Load balance	Process of distributing workloads across multiple computing resources. [2]
43	Recovery	It is a backup and restore strategy. [2]
44	Service Level	A general SLA parameter which governs different level and category

		of services. [2]
45	Migration	Shows the portability of services and virtual instances. [2]
46	Memory Capacity	The parameter is generally used by IaaS providers. Its shows the ability of memory to host or hold the operations. [2]
47	Booting time	Total time to Boot time for VM, IaaS provider. [2]
48	Storage	Size of the external and internal storage of a VM
49	Auto Scaling	Support auto-scaling or not2
50	Virtualization degree	Time to increase/decrease a VM Maximum number of VMs can be deployed on a physical host [2]
51	Manageability	Time proportion of Continuous delivery and the ability to manage request, services, instances etc. [2]
52	Value added service	Service of help and support Description [2]
53	Scalability	Capability to perform and sub stain under varying demand. [2]
54	Application server	Virtual resources that can be allocated to user's application. [2]
55	Browsers	The type of browsers used for accessing PaaS development environment and other applications. [2]
56	Development and deployment environment	Specific mechanism used for deployment and access. [2]
57	Multi terminal support	Includes mobiles, thin and thick client. [2]
58	Open API	Provide API to individual developers. [2]
59	Offline Services	These support offline operations. [2]
60	Usability	Simplification of use of any application. And attractive UI [2]
61	Customization	Adapts to user specific needs. [2]
62	VM Scale	Maximum and minimum amount of Virtual machine that can be used. It also shows the level of scalability. [2]
63	Speed of VM	Parameter describes the speed at which VM operates. [3]
64	VM capacity	The ability of VM to hold and process data. [1]
65	Type of VM	Different types of VM used. [3]
66	Integration	The ability of services to integrate from other platforms. [2]

Some parameters like customization, manageability, browsers etc. which acted as a features or services in traditional IT environment is defined and referred as a SLA parameter in Cloud environment.

VI. Conclusion

The parameters are the deciding factors for provisioning of services along with its use. Therefore, it is desirable to control possible violations before they happen by predicting the various SLA parameters over a period of time. SLA involves constant monitoring and controlling of various SLA parameters**.** As discussed SLA is a contract with service providers and set expectations for the relationship. SLA should be with respect to provider and consumers. In order to develop an effective SLA a list of important criteria needs to be established. Because of these there is a need to know and evaluate the probable parameters which can help in construction of SLA and establish the QOS.

References

[1] AMIR VAHID DASTJERDI and RAJKUMAR BUYYA," An Autonomous Time-dependent SLA Negotiation Strategy for Cloud Computing", The computer journal- Oxford, Vol. 58, No.11, pp3202-3216l, 2015.

[2] Wu Chenkang et.al," The SLA Evaluation Model for Cloud Computing "International Conference on Computer, Networks and Communication Engineering (ICCNCE 2013), pp 331-334,2013.

[3] Wu Li et.al, "SLA-Based Resource Provisioning for HostedSoftware-as-a-Service Applications inCloud Computing Environments"**,** IEEE TRANSACTIONS ON SERVICES COMPUTING, VOL. 7, NO. 3, pp 465-485,2014.

[4] Yun Chi et.al, "SLA-Tree: A Framework for Efficiently Supporting SLA-based Decisions in Cloud Computing", Proceeding EDBT/ICDT '11, 14th International Conference on Extending Database Technology, pp 129-140, ISSBN: 978-1-4503-0528-0,2011

[5] Cristiano Costa et.al, "A Two-Dimensional SLA for Services Scheduling in Multiple IaaS Cloud Providers", International Journal of Distributed Systems and Technologies, 6(4), pp 45-64, 2015

[6] Valentina C et.al, "Per-service Security SLA: A New Model for Security Management in Clouds", IEEE 25th International Conference on Enabling Technologies: Infrastructure for Collaborative Enterprises (WETICE), pp: 83-88,2016
[7] Rodrigues G.et.al, "Monitoring of Cloud Computing Environments: Concepts, Solutions, Trends, and Future Directions", SAC 2016 ACM Proceedings of the 31st Annual ACM Symposium on Applied Computing, pp 378-383, ISBN 978-1-4503-3739-7 ,2016.
[8] Paul Manuel et.al," A trust model of cloud computing based on Quality of Service" This work is supported by Kuwait University, Research Grant No. [WI 07/11]
[9] Zhao Y. et.al., "SLA-Based Resource Scheduling for Big Data Analytics as a Service in Cloud Computing Environments" 44th International Conference on Parallel Processing(ICPP), ISBN: 978-1-4673-7587-0,2015
[10] Wagle S. et.al, "Cloud Service Providers Ranking Based on Service Delivery and Consumer Experience", Cloud Networking (Cloud Net), IEEE 4th International Conference, ISBN: 978-1-4673-9501-4 ,2015
[11] Radha K. et.al, "Service Level Agreements in Cloud Computing and Big Data", International Journal of Electrical and Computer Engineering (IJECE) Vol. 5, No. 1, February 2015, pp. 158-165, ISSN: 2088-8708
[12] https://en.wikipedia.org/wiki/Cloud_computing_architecture.