

SLA in Cloud Computing Architectures: A Comprehensive Study

Eman Aljournah¹, Fajer Al-Mousawi², Imtiaz Ahmad³, Maha Al-Shammri⁴ and Zahraa Al-Jady⁵

^{1,2,3,4,5}Kuwait University, Department of computer Engineering,
Kuwait

¹Eman-aljumaa@hotmail.com, ²eng.fajer163@gmail.com,
³Imtiaz.ahmad@ku.edu.kw, ⁴engmahaalshammri@gmail.com,
⁵z.aljady@gmail.com

Abstract

Cloud computing, the model for providing on-demand access to a pool of shared resources with minimum provider interference, is emerging as a substitute to common IT infrastructure. As increasing numbers of cloud consumers dispatch their workloads to cloud providers, Service Level Agreement (SLA) between consumers and providers becomes of paramount importance to guarantee that service quality is preserved at satisfactory levels regardless of the dynamic nature of the cloud environment. SLA contains an explanation of the agreed service, parameters of the level of service, the guarantees regarding the Quality of Service, arrangements and cures for all cases of violations. In this paper, we provide a study about the general structure of SLA, its components, the management processes (in particular SLA monitoring), SLA lifecycle, and pricing. Then, we explore the importance of SLA for cloud computing services related to both the cloud user and the service provider including the differences between SLA for cloud services and other Web services. Finally, we present a comparison between the current major cloud computing service providers in terms of SLA's.

Keywords: Cloud Computing, Service Level Agreement, SLA Management, SLA Pricing

1. Introduction

Cloud computing is a new type of IT mediums which supplies dynamic, elastic and scalable resources virtualization. It is explained by NIST [1] as a pattern for allowing global, appropriate and on-request network entrance to a mutual group of computing resources such as; networks, servers, storage, applications, and services that could be quickly provided and issued with least administration work or service supplier dealings. It has three main service models:

- Software as a Service (SaaS)

By using this model, the client can exploit and enter the application (like an email or Dropbox etc.) that works on the supplier cloud groundwork without having authorization to make changes on any of its elements like networks, servers, OS, or storage [1].

- Platform as a Service (PaaS)

By using this model, the client can operate in the groundwork of the cloud applications created by programming languages, libraries, or services, supplied through the service supplier. The client cannot do changes on any of the main groundwork elements like networks, servers, OS, or storage. But, the user has control on the operating applications and their settings [1].

- Infrastructure as a Service (IaaS)

Using this model, the client can enter the hardware and software resources. He does not own the authorization to the administration of the original cloud groundwork, but only can control the OS, storage, and operating applications (*e.g.*, Amazon EC2) [1].

Cloud computing has emerged as a new frame for transporting effective computing services [2]. In Clouds, services are presented when requested and organizations can access their services and applications anyplace in the realm whenever they want. As clients give their works to cloud suppliers, SLA between users and suppliers becomes of great importance to guarantee that service superiority is preserved at satisfactory levels regardless of dynamicity. Because of the dynamic nature of the cloud environment, uninterrupted controlling on Quality of Service (QoS) characteristics is essential to apply SLAs. Also, other aspects like confidence (on the cloud supplier) arise into thought, mainly for projects clients that might outsource its important information. This compound kind of the cloud confirms advanced ways of handling SLAs.

Service Level Agreement between the cloud provider and the cloud customer is examined in this paper to attract attention to its importance. We start with a general overview of Service Level Agreement (SLA) explaining the main advantages it delivers such as: improving customer acceptance level, enhancing relationship between the parties, and enhancing service quality (QoS: Quality of services).

SLA components and the main information it should include are also mentioned along with the major SLA categories, the measurable and un-measurable qualities that can be determined in an SLA, and the major SLA parameters and metrics. In addition, we explain the management process for SLA, and the importance of SLA monitoring.

SLA lifecycle, pricing, violations, and penalties are also discussed. Then, Web Service Level Agreement (WSLA) is defined through its components, users, objectives and its architecture.

SLA for cloud computing services is examined by away in which we define the cloud user, provider and the importance of SLA for them. Also, the main differences between SLA in Cloud services and other Web services are introduced along with the proposed cloud SLA metrics that are mentioned based on the researches done in this field.

We conduct a comparison between the major cloud providers' SLAs. In addition, the cloud SLA based on WSLA is clarified through examining its operations, functions, classifications, main metrics and architecture.

The rest of this paper is organized as follows: Section 2 presents a general overview of Service Level Agreement (SLA). The lifecycle of SLA is described in Section 3. Section 4 contains the management processes for SLA concentrating on SLA monitoring in particular. Section 5 discusses SLA for cloud computing services in details with a comparison between major cloud providers' SLAs. We conclude this paper in Section 6.

2. A General Overview of Service Level Agreement (SLA)

The SLA main concepts which are going to be discussed in this section are shown in Figure 1.

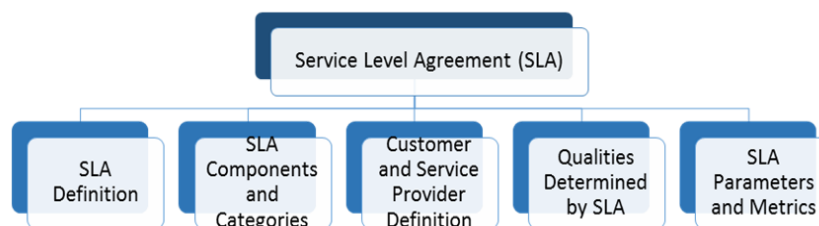


Figure 1. SLA Main Concepts

2.1. SLA Definition

Researchers defined “Service Level Agreement” as a format that contains an explanation of the agreed service, parameters of the level of service, the guarantees regarding the Quality of Service, and arrangements for all cases of violations. The SLA is very significant as a contract that is held between the provider of the service and another party who could be one of following; consumer of the service, broker negotiator, or monitoring negotiator. The key concept of SLA is to provide a clear description of the official agreements about service expressions such as performance, availability and billing....., etc. It is essential that the SLA should contain the duties and the activities that will be done in the occurrence of any contravention. So this contract determines, commonly in quantifiable terms, the services that the service provider will provide and what sanction the Service Provider will pay if the dedicated objectives cannot be met. SLA has been applied in a multiplicity of domains that are related to IT such as: Web Services, Networking, Internet, and Data Centre Management; the existing descriptions differ from domain to domain.

The most summarized SLA contains generic and technical terms such as; business endings, pricing strategy, and attributes of the resources needed to operate the service [3]. As stated by Sun Microsystems Internet Data Center Group's report [4], an appropriate SLA puts borders and prospects for service supplying and delivers the following advantages:

- Improved customer acceptance level: An obviously explained SLA enlarges the customer approval level, as it supports providers to concentrate on the customer needs and confirms that the work is placed on the right way.

- Enhanced relationship between the parties: An obvious SLA specifies the remuneration and retribution policies of the service provided. The consumer can examine services according to Service Level Objectives (SLO) determined in the SLA. Moreover, the specific contract assists parties to solve disagreements without difficulty [5-9].

- Enhanced Service Quality (QoS: Quality of Services): Each element in an SLA suits a Key Performance Indicators (KPI) that determines the customer service among an inner institute by checking whether or not these indicators match the Service Level Objectives (SLOs) of the agreed contract between customers and service providers. A contract is a means of clarifying the difference in needs for both the provider and the user. An SLO always contains level services, specific period, and a specific value as a target to achieve. These actual performance values are compared with the stated ones in the contract to make performance reports used for evaluation [32].The key for reaching the desired level of quality is to select the quality indicators in the early stages to use them later in determining violations. SLA templates and mechanisms can be described based on their performance, availability, and fault tolerance. Templates are customized by prioritizing one metric at a time, performance or fault tolerance. For instance, the fast SLA template prioritizes performance, while the safe SLA template prioritizes fault tolerance. On the other hand, the classic and standard SLA templates prioritize neither performance nor fault tolerance [33].

- Performance: is defined by the response time QoS metric which refers to the maximum time that the request treatment can take from the time of storage, and the arrival of the response. The performance QoS assurance mechanism is shown in sequence diagram in Figure 2. When a customer has a deal or contract for the service provider, then the provider translates and determines the response time depending on its requirements. After that, the service provider gets service resources from the infrastructure during whole duration of contract. The requirements are used to decide and establish the service instance. Thus, the customer can send requests to the appropriate instance within the

given response time. Then as a response, the instance treats the received requests according to the requirements. Thus, by booking resources and shaping the instances to suit the right resource requirements, the performance is guaranteed [34].

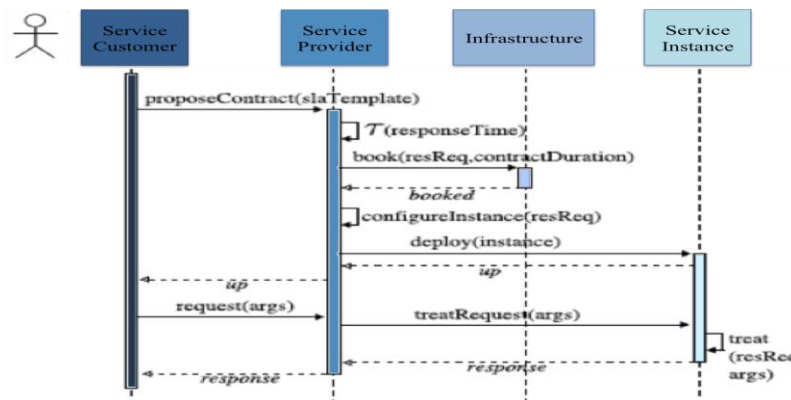


Figure 2. Sequence Diagram That Shows the Performance QoS Assurance Mechanism [46]

- Fault tolerance: is defined by the reliability of the system which refers to the degree of dependability. Fault-tolerance QoS assurance mechanism is mainly used by the service provider to catch and resist faults or crashes during the treatment. These faults can be classified as either job failures or job delays. An unsuccessful job execution can be classified as job failure. To determine a job delay, some provider's information needs to be known such as execution time of the job and request response time. Finally, by solving the delayed jobs during the request treatment, fault tolerance is guaranteed [33].

2.2. SLA Components

An ideal SLA has the following components:

- Purpose – mentions why SLA is formed.
- Parties – mentions the parties included in the SLA and their jobs.
- Validity Period- states the time period covered by the SLA. This is defined by both the start time and the final time of the period.
 - Scope – describes the services mentioned in the SLA; SLA structure should illustrate the service so that the consumer can simply recognize the services procedure.
 - Restrictions – states the essential steps to be done in order to supply the required service levels.
 - Service-level objectives – the service levels that are approved by the customer and the providers. It contains a group of service level indicators such as; availability, performance, and reliability. Each part of the service level, like availability will have a target level to complete. Service Level objectives have day-time restrictions related to them to describe their validity.
 - Service-level indicators – those indicators are used to measure these levels of service.
 - Penalties – describes what is to be done when the provider cannot achieve the goals in the SLA. If the SLA is taken with an external provider, there should be a choice of concluding the contract.
 - Optional services – services that are not ordinarily needed by the customer, but might be needed as exclusion.
 - Exclusions – states what is not included in the SLA.

- Administration – defines the procedures formed in the SLA to achieve and measure its goals [8].

SLAs have been utilized for a long period in IT fields to determine the demands of the clients of IT services. An SLA specified the anticipations of the service client and provider. It is frequent for providers to transport services at variable levels of quality depending on the cost of the service. An SLA is precious for assisting all actors to recognize the tradeoffs between cost, plan, and quality. Same as any kind of contract, an SLA cannot assure that all commitments will be maintained, but it describes what will take place if those commitments are not met. Guaranteeing the superiority of services supplied over the internet is a large challenge, because the internet is dynamic. Some of these challenges are:

- Low performance of typical protocols.
- Security cases.
- Infrastructure malfunctions.

“An SLA cannot guarantee that you will get the service it describes, any more than a warranty can guarantee that your car will never break down. In particular, an SLA cannot make a good service out of a bad one. At the same time, an SLA can avoid the risk of choosing a bad service [10]”. A “good quality” service is one that mitigates the requirements of the client that include goodness and appropriateness. The way to design SLAs is to supply sufficient data or metrics for a client to preselect services depending on the preferred stage of superiority. Usually, SLAs are stated in basic content, using forms or toolkits. Providers design their systems in a way that measurements are gathered and then matched to the metrics determined in the SLA.

There are three major SLA categories:

1. Basic – an SLA with well-organized metrics that are calculated and/or confirmed. The gathering of these metrics is usually completed physically.
2. Medium – a multi-stage superiority depending on the cost of the service. The goal is to equalize the stages of superiority and cost.
3. Advanced – dynamic distribution of resources to achieve requirements [3].

2.3. Customer and Service Provider Definitions

The word Service Provider refers to firms who offer business such as communications and/or data services. Service suppliers may run networks, or they may combine the services of other suppliers to provide the service to their users. The Service supplier could be a worker, a transporter, an Internet Service Provider (ISP) or an Application Service Provider (ASP). The word Customer refers to firms or organizations that employ the services of telecommunication supplied by a Service Provider. The client could be a transporter, an ISP, a project or an end client [6]. The SLA is a lawful layout authenticated in the way that services will be offered as well as supplying a structure for service charges. Service providers exploit this base to enhance their usage of infrastructure to achieve signed conditions of services. Service consumers exploit the SLA to reach the stage of quality of service they require and to keep appropriate business models for a long period [6].

2.4. Qualities That Can Be Determined in an SLA

Identifying any quality in an SLA can be done, as long as all actors recognize how to measure or capture its accomplishment. There are two groups of qualities that can be identified in SLAs: measurable and un-measurable qualities. Measurable qualities could be calculated by metrics; for example, the percentage of time a system is presented. Un-measurable qualities can be calculated from a specified estimation; for example, the complexity of automating the cost of an alternating service.

Measurable Qualities:

- Accuracy cares about the service error rate. Identification of the average number of errors over a given time period is possible.
- Availability is defined by the ratio of service working time. Another definition can be stated for availability which is the time when the service is usable compared to the total time [35]. This aspect is comparing mean time to failure of services. Thus, it affects the reputation of the cloud by increasing or decreasing the confidence of the user. Availability is usually calculated by the probability that the system will be equipped when required. It is likely to identify:
 - The system's reply when a malfunction appears.
 - The time taken to identify a failure.
 - The period taken to get better from a malfunction.
 - If error managing is exploited to cover malfunctions.
 - The downtime needed to apply improvements.
- Capacity is the number of simultaneous demands that can be held by the service in a specified time range. It is probable to specify the maximum number of simultaneous demands that can be held by a service in a set bulk of time.
 - Cost cares about each service demand cost. It is likely to identify:
 - The cost per demand.
 - The cost based on the volume of the information.
 - Cost distinctions associated to peak employment times.
 - Latency cares about the largest quantity of time between the appearance of a demand and the finishing point of that demand.
 - Provisioning-related time (*e.g.*, the time taken for a new customer's account to be prepared)
 - Reliable messaging cares about the assurance of message transmission. It is likely to identify:
 - How message transmission is assured.
 - If the service assists transmitting messages in the appropriate order.
 - Scalability cares about the eligibility of the service to enlarge the number of successful processes accomplished over a specified time range. It is likely to identify the maximum number of processes.

Un-measurable Qualities:

- Interoperability cares about the eligibility of a group of communicating units to share precise information and function on it according to an agreed-upon prepared semantics [38]. It is likely to identify the standards assisted by the service and to validate them at runtime. Important difficulties still require to be solved to accomplish semantic interoperability at runtime.
 - Modifiability cares about how frequently a service is possible to modify. It is likely to identify how frequently the service's:
 - Interface modifies.
 - Implementation modifies.
 - Security cares about the system's facility to prevent not permitted usage, while supplying legal clients with entrance to the service. Security is also specified as a system supplying non-rejection, privacy, honesty, guarantee, and auditing. It is likely to identify the procedures for:
 - Trusting services or clients.
 - Permitting services or clients.
 - Encrypting the information [11].

2.5. SLA Parameters and Metrics

Metrics are exploited in monitoring procedure, software procedure enhancement, business policy employment, and mainly any area where information has to be gathered to confirm whether objectives are being achieved. Past researches specifies that generating metrics is a hard job that is why it usually covers rules for generating proper metrics. SLA parameters are identified by a group of metrics. These metrics decide the amounts that need to be collected to confirm whether the SLA parameters are being achieved. The following strategies are suggested to describe SLA parameters in a way simplifies SLA management. Assumed that SLAs are commonly declared and negotiated, these strategies are applicable to both providers and clients. Thus the parameters must be:

- Reasonable; they must encourage contributory actors to perform in a way that is commonly useful. For example, an SLA parameter that identifies cost depending on the amount of services appealed might weaken service plan in the concern of making the most of revenue.
- Achievable; metrics that are outside the monitoring of both actors would not be involved. For example, because of the unexpected kind of the communication of internet, there will be no need to contain metrics for the entire latency of the operation that goes through the internet.
- Enforceable; this superiority is particularly significant for the service client. SLAs must define and cover indication of how SLA parameters are controlled and applied by the provider. A complete SLA monitoring resolution must determine equipment and measurement on the client and the provider end. Clients can't guess that the provider will inform them when an SLA contravention happened.
- Quantifiable; metrics must be measurable and permit for measurements. Measurements must not be very hard or very expensive to gather. For example, if gathering the metrics exploits important resources, it might not deserve to spend any effort.
- Target; any SLA parameter that has many explanations perhaps desires to be reviewed. For example, an SLA parameter can state that the provider will inform clients when dealings are set but state nothing regarding the period of those announcements. This parameter might guide to altered explanations of suitable latency. It might be suitable to state advanced level metrics that are associated to SLA parameters. For example, gathering performance information from a group of servers can offer convenient information for load equilibrium and resource managing. As described in Figure 3, conducting data associated to SLAs occurs at different stages of a spread organization and can be addressed as follows:

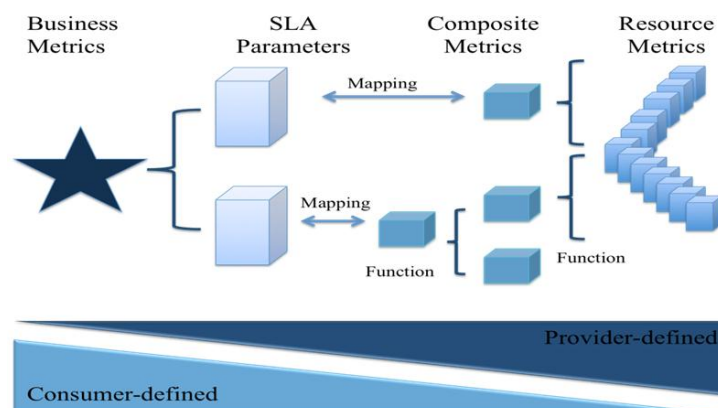


Figure 3. Aggregating Business Metrics, SLA Parameters and Metrics through Different Organizations [41, 45]

- Resource metrics are restored immediately from the organized resources during at the supplier's stage like (routers, servers, and applications). To incorporate a resource easily in a web service medium, WSLA utilizes the idea of Measurement Directive. For each metric shown in an SLA, a Measurement Directive is determined, which should include the orders and other required data to get the metric from the organized resource agency.
- Composite Metrics are made by merging many resource metrics depending on a particular algorithm. This is commonly made by the service supplier's area. It is supposed that the composite metrics either determined in the SLA by using a function (a form explaining the input metrics and the mathematical processes to assemble them) or detected by the service supplier through using a well-stated interface.
- SLA parameters place the metrics existed from a service supplier into the perspective of a particular client and are for that reason the center fraction of an SLA.
- Business Metrics associate SLA parameters to economic expressions related to the service client [9, 11].

3. SLA Life Cycle

SLA has six main stages to be completed. These stages are as follows; development of both service and SLA templates, discovery and negotiation of an SLA, service provisioning and deployment, execution of the service, assessment and corrective actions during execution, and both termination and decommission of the Service. The SLA lifecycle was described by the Tele Management Forum [12] as shown in Figure 4.



Figure 4. SLA Life Cycle [12]

3.1. Development of Service and SLA Templates

This stage includes the identification of customer requirements and needs, the network capabilities, the identification of the suitable service features and parameters, service's levels, service executional environment, and the implementation of the standard of SLA templates [11].

3.2. Discovery and Negotiation of an SLA

Discovery stage consists of; the negotiation of an SLA with the consumer to select the values of SLA parameters related to specific services, the costs gained from the service customer after signing the SLA, the costs incurred by the service provider when the SLA is violated, the definition and at last periodicity of the reports associated with service to be delivered to the service customer.

There was a study presented in conceptual SLA framework for cloud computing which mentioned that there are many methods of initiating the compromise procedure in an online medium [13, 14]. The probable compromise scenarios related to cloud computing were described. The first situation is direct compromise between the cloud user and the cloud supplier. The service supplier can make a single outline and state all SLA sections like duration of convention, payment, time of reply. When the SLA form is set, cloud users can look at the SLA conditions and react by signing the SLA, compromising again or ending the compromise. Direct compromising is a familiar technique exploited by many of cloud suppliers. The second situation is compromise by a confident agent, who is, an agent with knowledge in choosing the cloud suppliers and stating the serious

parameters for the SLA. This can be a main issue when a business needs to spotlight on the centre business actions. A number of actions should be specified to the outside agents who carry out the compromise in elastic and consistent steps. They could begin with the examination of business both procedures and objectives and accomplish the compromise by controlling all or some of the SLA parameters. Also, the confident agent can make use of other agents to complete some actions such as service detection and controlling of the performance.

In the third situation many agents are employed to perform the same kind of compromise. As there are four various kinds of cloud services, a cloud user could sign a convention with four various agents (IaaS agent, PaaS agent, and SaaS) which obtain the job of stating SLA parameters and finish the compromise procedure. This kind of compromise can be good if the cloud user needs more than one kind of cloud service.

3.3. Service Provisioning and Deployment

This stage include the service's resource provisioning, where the service is enabled and prepared for the service shopper consumption, configuration of the network which might be to achieve specific requirements in the service, or to support the service network overall, and service activation. Service provisioning and deployment stage may need the reconfiguration of the service resources to support the executional stage which will lead to a successful achievement of the SLA parameters [12].

3.4. Execution of the Service

This stage is the actual test of the service. It consists of three main phases , The first is service execution and monitoring, Then the real time of reporting and at last the validation of QOS which refers to the quality of service. The final phase of this stage is SLA violation processing [11].

3.5. Assessment and Corrective Actions during Execution

SLA assessment stage consists of two parts, the assessment with the individual customer, and the overall service assessment. The SLA assessment of the customer includes reviewing the Quality of Customer Service (QoS), customer gratification, achieving the possible enhancements, and altering requirements are examined for each SLA. Overall service assessment for major activities are readjusting of service goals, service operations modifying, defining the support problems of the service, and finally establishing different service levels.

3.6. Termination and Decommission Of the Service

Termination and Decommission of the Service stage in charge with the termination of the service. This termination may be a result of different reasons; it might be an issue in the contract, expiration, or violation. The decommissioning of discontinued services can cause termination to the SLA [12].

4. SLA Management

Conditions of an SLA are guaranteed by modifying the systems to provide suitable KPI's and KQI's (Key performance indicators and key quality indicators) at the necessary sample rate. In the design stage, it is important to make sure that the measurement process does not add extra load to the system (added traffic overhead, more processing power). Furthermore, the SLA should be continually monitored to help create management reports and also assure that remedial actions can be taken. These management reports, as for the SLA itself, should be concise and clear. The report is generated by the provider of the service or the client, or a third party trusted by both of them [15].

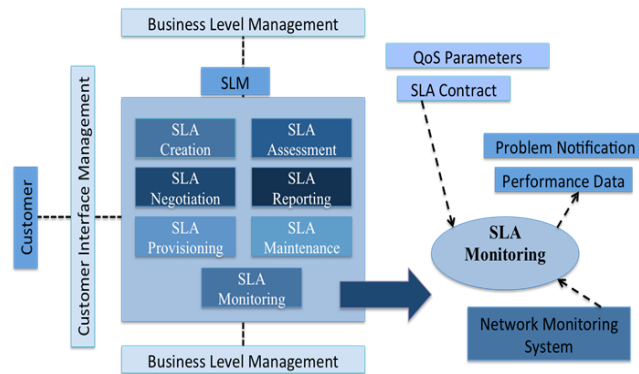


Figure 5. SLM and SLA Monitoring [43]

SLA management is the integrated process of managing various SLAs from start to assessment. SLA management can be categorized into three groups; business level management, service level management, and network level management. The SLA service level management consists of several functions starting with SLA creation, negotiation, provisioning, monitoring, maintenance, reporting and assessment which are shown in Figure 5.

As monitor ability is one important SLA requirement. This specific function needs to be further looked into.

Monitor ability denotes that the service provider and the client can observe and manage the behavior of the service related to the SLA, or employ a trusted third party to do so. Without this requirement, it would be impossible for a party to state that there is an SLA violation. Therefore its terms may be overlooked by the service provider [16]. The problem faced when monitoring compliance with unanimous performance metrics is a big challenge for SLA engineers. The SLA must be designed to guarantee high monitor ability, and decrease the probability of low compliance.

The network level management consists mainly of network monitoring. Network monitoring is the process of the value of the network performance metrics (NPM) by different network monitoring tools and techniques. There are three known methods to monitor a network:

1-The first method is the active monitoring which is traditionally used to measure loss, connectivity and delay. Active monitoring sends extra traffic between machines after setting up those test machines where measurements is to be taken to obtain the current status of the network. Active monitoring uses simple and easy tools, such as ping and trace-route. The system load is very low with active monitoring because the quantity of generated traffic is small compared with that of other methods. However, the generated test packets may be lost due to their low priority which makes it difficult to obtain the exact network status sometimes.

2- The second method is the passive monitoring which relies on capturing the packets to obtain the current network status. This is why passive network is ideal to measure NPMs (network performance metrics) like utilization and throughput.

3- The third and the last method to measure the status of the network is by using SNMP agents. Although this method is practical and simple, it is limited to measure the throughput and the functionality of NPMs [17].

As easy as it might seem, obtaining different NPMs using the described network monitoring methods above, it rather challenges applying the values obtained directly to QoS parameters. SLAs are constructed in terms of QoS while actual measurements are NPMs. This is why a mapping mechanism is needed and NPMs must be defined before deciding on the QoS parameters in an SLA. The mapping between QoS parameters and the measured NPMs depends mainly on the type of the provided service. It can be

complicated and its outcome presentation form should be clear and understandable to the user (in QoS terms not NPMs). More on this topic can be found in [18].

Table 1. Network Monitoring Mechanism and Related Projects

Monitoring method	Mechanism	Related projects	NPMs measured
Active monitoring	-Generate test traffic periodically or on-demand. - Measure performance of test packet or response.	-NIMI -Surveyor -NLANR AMP -PingER -Skitter	Loss, Delay and Connectivity
Passive monitoring	-Capture the traffic by mirroring or splitting. -Analyze the captured packets.	-CoralReef -WAND -WebTrafMon -NLANR PMA	Utilization and Throughput
SNMP	-Using existing SNMP agents.	-Internet2 -MAWI	Functionality and Throughput

Similarly, service monitoring is performed as part of the service level management. Through service monitoring, data related to performance is reclaimed from service resources for each one of the promised services. These reclaimed instances are then collated and integrated to form KQI for both service resource and product KQI as shown in Figure 6.

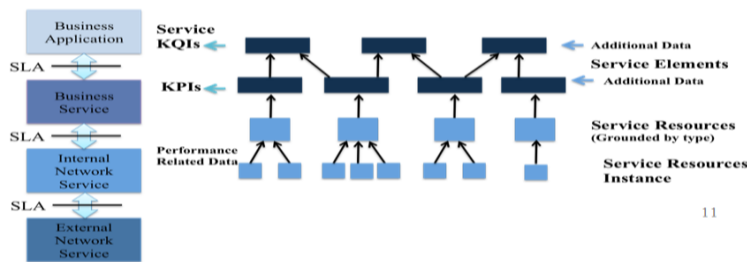


Figure 6. Relationship between Service Resources, KQI, and KPI [42]

Another aspect of the service management process is SLA reporting. It is obvious now that the generated management reports are to be seen by more than one functional group that may include a senior management, SLA engineers, a financial group for handling both charging and billing, and at last end users. This is why the output format of these reports should be appropriate and understandable by all audience. New reporting tools are used and it is applicable to allow users to develop reports on their own.

5. SLA in Cloud Computing

5.1. Definition of Cloud Consumer and Cloud Provider

The cloud client is the person with an interest or concern for the cloud computing service. A cloud client stands for an individual or institute that exploits the service from a cloud supplier. A cloud client looks over the service index from a cloud supplier, demands the proper service, and constitutes service conventions with the cloud supplier.. The cloud user can be asked to pay for the service supplied, and to organize expenses consequently. Cloud users require SLAs to identify the technical performance demands

satisfied by a cloud supplier. SLAs can accomplish conditions about the superiority of service, security, and cures for the facing malfunctions. A cloud supplier could also state in the SLAs a group of guarantees that are not prepared to users clearly, i.e. restrictions, and duties that cloud users have to approve on. A cloud user can select a cloud supplier with preferable pricing and more complimentary conditions. Normally, a cloud supplier's pricing strategy and SLAs are non-discussable, except if the user looks forward to intensive employment and can be able to discuss for superior convention.

Relying on the services demanded, the actions and employment situations can be diverse over cloud users. Figure 7 shows several cloud services accessible to a cloud client.

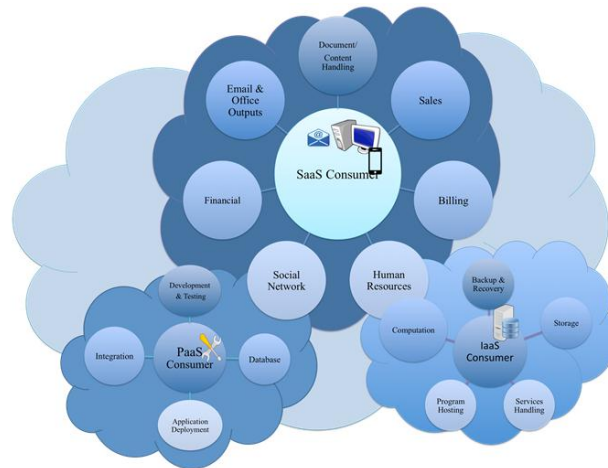


Figure 7. Several Cloud Services Accessible to a Cloud Client [44]

The clients of SaaS might be corporations that offer their participants with entrance to software applications, end clients who immediately exploit software applications, or software application directors who constitute applications for the clients. SaaS expenses can be paid according to the number of the end clients, the usage time, the network bandwidth spent, the quantity of information kept or the period of keeping information.

Cloud clients of PaaS can exploit the instruments and the resources supplied by cloud suppliers to progress, examine, install and administer the applications presented in a cloud medium. PaaS clients can be application designers who develop and accomplish application software. Also, they can be application examiner who execute and examine applications in cloud-based locations. They can be application publishers who distribute applications through the cloud, or can be application managers who constitute and control applications. PaaS expenses can be paid based on, operation, database space, network resources used by the PaaS application, or the period of the platform convention.

Clients of IaaS have an entrance to virtual computing machines, network storing space, network groundwork elements, and other essential resources on which they can install and operate random software. The clients of IaaS can be system designers or system managers who are concerned in making, running, organizing and controlling services for IT groundwork processes. IaaS users are provided with the abilities to enter these resources, and are paid depending on the quantity or time period of the resources used like; CPU hours consumed by virtual computing machines, capacity, network bandwidth used, and quantity of IP addresses utilized for particular periods. Cloud users want an SLA before delivering their groundwork of cloud information stations to have confidence about the resources supplied and to have the facility to get the preferred level of efficiency.

The cloud supplier is an individual or an institute that is accountable for providing accessible service to concerned actors. A Cloud Supplier develops and administers the computing groundwork needed for supplying the services, operates the cloud software

that supplies the services, and makes procedure to transport the services to the Cloud Clients through network entrance.

For SaaS, the cloud supplier installs, constitutes, preserves and improves the process of the software applications on the cloud groundwork so that the services are provided at the estimated service levels to cloud clients. The supplier of SaaS considers many of the tasks in handling and monitoring the applications and the groundwork, while the cloud users have partial managerial monitoring of the applications.

For PaaS, the Cloud supplier organizes the computing groundwork for the platform and operates the software that supplies the elements of the platform like; software implementation stack, databases, and other elements. Moreover, the PaaS Cloud supplier usually provisions the improvement, organization and administration procedure of the PaaS cloud user by supplying instruments like; integrated development environments (IDEs), improvement form of cloud software, software development kits (SDKs), distribution and organization instruments. The PaaS Cloud user has monitoring on the applications and probably some of the introducing locations settings, but has no or restricted entrance to the groundwork likes the network, servers, operating systems (OS), or storage.

For IaaS, the Cloud supplier obtains the tangible resources of the service, such as the servers, networks, storage and hosting groundwork. The Cloud supplier operates the software needed to have computing resources available to the IaaS cloud user through a group of service interfaces and computing resource ideas such as virtual computers and virtual network interfaces. The IaaS Cloud users exploit these resources, such as a virtual machine for their essential computing requirements suitable for SaaS and PaaS Cloud users. An IaaS Cloud user has entrance to more essential shapes of computing resources and has more monitoring on the software elements inside the application, OS and network. The IaaS Cloud supplier has control on the tangible hardware and software that makes the providing of these groundwork services probable. Some main fields accomplished by a Cloud Provider are presented in Figure 8. Cloud suppliers want an SLA to state the confidence and superiority of services they supply to customers along with a settled structure for prices and duties [7, 19].

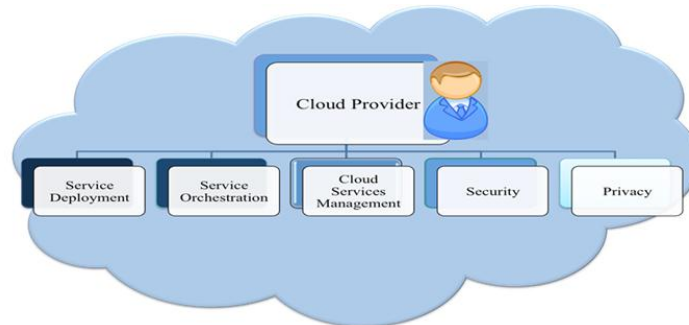


Figure 8. Some Main Fields Accomplished by a Cloud Provider [44]

Cloud computing that supplies less cost and gives price-as-you-use resources is quickly earning activity as a substitution to the conventional IT Groundwork. As users come up to exploit cloud computing, the superiority and consistency of the services come to be significant parts. But the requirements of the service clients differ meaningfully, so equilibrium has to be done through the negotiation procedure. At the conclusion of the negotiation procedure, supplier and user reach to a covenant (SLA). This SLA assists as the basis for the predictable level of service between the user and the supplier. The QoS characteristics that are commonly portion of an SLA alters frequently, so the parameters have to be carefully controlled and observed [5, 20].

5.2. The Difference between SLA in Cloud Services and Other Web Services

In dynamic mediums like Clouds, a number of challenges must be considered: automatic compromise and dynamic SLA organization depending on environmental modification. They are the mainly significant examples. A research conducted by [5] has mentioned the variances between SLAs used in cloud computing and in the other web services:

- QoS Parameters: many web services emphasis on parameters like response time, SLA contravention degree for the job, reliability, availability, stages of client variation and charge of service. In Cloud computing, additional QoS parameters require to be included, (*e.g.*, energy associated QoS, Security associated QoS, Privacy associated QoS, and trust associated QoS). Over 20 QoS parameters are mentioned by the SMI (Service Management Index) group to be exploited.
- Automation: The entire procedure of SLA discussion, providing service, transfer and controlling has to be automated for a very high dynamic and accessible service employment.
- Allocation: SLA oriented resource distribution in Cloud computing is probable contrastive from allocation in web services, because in Clouds, resources are assigned and spread worldwide without centric guide. So the plan and the construction for SLA resource distribution in those environments are contrastive.

On the other hand, most of the requirements are still the same for both kinds of services such as: Availability, Scalability, Security, and privacy, *etc.* Moreover, a clear technique for cost computation, because users consume cloud computing wish to give price as they consume, so yearly or monthly billing durations are not appropriate for cloud computing. A cost computation for resource booking technique is not a single technique for each kind of cloud service. For example, the storage service can be paid according to the time and volume of the client's information. Instead, cloud CRM might be paid according to the number of customers.

5.3. Proposed Cloud SLA Metrics

The study on SLA and QoS metrics has been done by lots of investigators in business and service-oriented construction like e-commerce and web services. Nevertheless, SLA metrics in these fields are not appropriate for cloud computing because the kind and form of resources supplied and sent is various. So, new SLA samples are needed to provide elastic technique for making conventions between users and suppliers. So, a theoretical SLA Framework for Cloud Computing has been presented in. [7] also dynamic SLA metrics for various sets of cloud clients have been identified.

In their planned structure, the SLA parameters are determined by metrics. These metrics state how service parameters can be calculated. Also determines estimations of quantifiable parameters. The planned SLA metrics for cloud computing examine the four kinds of cloud services which are (SaaS, PaaS, IaaS, and Storage as a Service). For every branch of the SLA they state the mainly significant parameters that users can utilize to make a consistent form of compromise with the service supplier. Their research spotlights on the explanation of these parameters.

5.3.1. SLA Metrics for IaaS

Firms such as amazon.com supply infrastructure as a service. Many clients don not know clearly which significant parameter must be declared in the hardware side of the SLA. The study mentioned the mainly significant parameters for clients who are concerned in utilizing cloud as an infrastructure service as shown in Table 2.

Table 1. SLA Metrics for IAAS

Parameter	Description
CPU capacity	CPU speed for VM (Virtual Machine)
Memory size	Cash memory size for VM
Boot time	Time for MV to be ready for use
Storage	Storage size of data for short or long term of contract
Scale up	Maximum of VMs for one user
Scale down	Minimum number of VMs for one user
Scale up time	Time to increase a specific number of VMs
Scale down time	Time to decrease a specific number of VMs
Auto scaling	Boolean value for auto scaling feature
Max number can be configured on physical server	Maximum number of VMs that can be run on individual server
Availability	Uptime of service in specific time
Response time	Time to complete and receive the process

5.3.2. SLA Metrics for PaaS

In platform as a service case, developers who exploit PaaS do not need to install instruments or organize hardware to do the developing jobs. For SLA metrics associated to PaaS, the study illustrates the key parameters that can be utilized as an essential principle when developers wish for compromising with PaaS suppliers as shown in Table 3.

Table 3. SLA Metrics for PaaS

Parameter	Description
Integration	Integration with e-services and other platforms.
Scalability	Degree of use with a large number of online users
Pay as you go billing	Charging based on resources or time of service
Environments of deployment	Supporting offline and cloud systems
Browsers	Firefox, Explorer, etc.
Number of developers	How many developers can access to the platform

5.3.3. SLA Metrics for SaaS

Superior examples of SaaS are mail, calendar and social web sites supplied by Google, Yahoo and Microsoft. The study shows the familiar metrics and parameters for SaaS as an illustration of metrics for this kind of cloud service as shown in Table 4.

Table 4. SLA Metrics for SaaS

Parameter	Description
Reliability	Ability to keep operating in most cases
Usability	Easy built-in user interfaces
Scalability	Used with individual or large organizations
Availability	Uptime of software for users in specific time
Customizability	Flexible to use with different types of users

5.3.4. SLA Metrics for Storage as a Service

Online clients enter their information from diverse places. Some time ago, online storage suppliers were not able to preserve a huge amount of information because there was not enough area in storage disks, network, and information supervision systems. Now, information storage service suppliers like S3 by amazon.com build up big numbers of storage hardware. Also they can handle and provide millions of clients powerfully with their technique of information delivering and guaranteeing that information are suitable for diverse kinds of applications. The parameters for information storage service metrics are fundamental necessities to compromise with storage suppliers as shown in Table 5.

Table 5. SLA Metrics for Storage as a Service

Parameter	Description
Geographic location	Available zones in which data are stored
Scalability	Ability to increase or decrease storage space
Storage space	Quantity of units of data storage
Storage billing	How the cost of storage is calculated
Security	Cryptography for storage, transferring data, authentication, and authorization
Privacy	How the data will be stored and transferred
Backup	How and where images of data are stored
Recovery	Ability to recover data in disasters or failures
System throughput	Amount of data that can be retrieved from system in a specific unit of time
Transferring bandwidth	The capacity of communication channels
Data life cycle management	Managing data in data centers, and using network infrastructure

5.3.5. SLA General Terms

There are common metrics that can be stated for SLA with any kind of cloud clients. The study mentioned the significant parameters as an illustration when making the essential SLA convention between cloud computing clients and suppliers as shown in Table 6.

Table 6. SLA General Terms

Parameter	Description
Monitoring	Who do the monitoring and what method of monitoring
Billing	Cost of service and how can be calculated
Security	Issues like cryptography, authentication, and authorization are main requirements for cloud users
Networking	The number of IPs, throughput, and load balancing
Privacy	How the data will be stored and transferred
Support service	Cloud providers should clearly define the methods of help and support
Local and international policies	The policy standards that providers follow

5.4. Comparison between SLAS of the Existing Cloud Service Providers

In this section, we explore how different cloud providers implement SLA. The characteristics chosen for the sake of comparison are selected based on similarities in attributes in the cloud SLAs we examined. The comparison outcomes can be found in Table 7. Furthermore, there is a number of steps developed by the cloud standards customer council that presents a series of ten steps for the consumer of the cloud service to evaluate and base its negotiation with the cloud vender based on. The steps are explained briefly below:

1- To understand the roles and responsibilities: AUPs (acceptable use policies) are what cloud consumer is mainly concerned about. Reviewing them thoroughly and carefully allows the consumer to understand exactly what their roles and responsibilities are along with the cloud providers' roles and responsibilities.

2- Evaluate business level policies: When reviewing SLA the consumer should consider major policy issues because the SLA policies, the business strategy and policy are somewhat dependent.

3- Understand Service and Deployment Model Differences: Service and deployment models are defined in [21]. This step is to make sure that the consumer understand what is the service model of the cloud (SaaS, PaaS, or IaaS), what are its characteristics? What are its objectives and KPI's? Furthermore, to understand the deployment model of the cloud presented in the service agreement (private, public, community, or hybrid). It is critical that the consumers understand the differences between those models to select the best to suit their requirements.

4- Identify Critical Performance Objectives: Four key components are considered in this step; service commitment, credits, credit process, and exclusions. Further explanation can be found in [22].

5- Evaluate Security & Privacy Requirements: Security and Privacy assurances should be obvious, distinct, and in clearly stated documents. Consideration should be taken for the consumer's data privacy.

6- Identify Service Management Requirements: Consumers should follow reasonable steps to guarantee that the provider is managing the level of service properly.

7- Prepare for Service Failure Management: Considering the offerings of public clouds, consumers must keep in mind the possible impact of service failure on their business operations.

8- Understand the Disaster Recovery Plan: The consumer should plan mainly for cases of disasters because the precautions taken by the cloud provider may not be sufficient to ensure the consumer satisfaction.

9- Define an Effective Management Process: Usually consumers expect good management from the cloud provider for any problem they might encounter. That is not the case actually. Concurrent cloud SLA's does not contain delivery of consumer-provider management process.

10- Understand the Exit Process: Every cloud SLA should contain an exit clause. An exit clause describes in details how the exit process is to be handled, what the provider and consumer has to do on contract termination.

Table 7. Comparison between SLAS of the Existing Cloud Service Providers

SLA characteristic / cloud provider	Amazon EC2 [48]	Microsoft Azure Storage [49]	Rackspace Cloud Servers [50]	Dell Boomi [51]	Google Cloud Storage [52]		
Type of cloud service	IaaS	PaaS	IaaS	SaaS	PaaS		
Service Provider Discovery	Manual discovery	Manual discovery	Manual discovery	Manual discovery	Manual discovery		
Service Availability	<99.95%	<99.9%	100% excluding scheduled maintenance periods down time	<99.9% Service is down for 1 minute once a week for scheduled maintenance	>= 99.9%		
SLA Outlining	Predefined terms and QoS parameters	Predefined terms and QoS parameters	Predefined terms	Predefined terms and QoS parameters	Predefined terms and QoS parameters		
Agreement Establishment	SLA document provisioned by the provider	SLA document provisioned by the provider	SLA document provisioned by the provider	SLA document provisioned by the provider	SLA document provisioned by the provider		
Service Management & Monitoring	Third party monitoring systems can be used under the terms of Amazon's AWS Agreements	Management services are delivered by the provider. Consumer can use third party monitoring systems	Provider offers proactive infrastructure monitoring, operating sys maintenance and patching, application maintenance	Provider offers system upgrades and scheduled maintenance and emergency maintenance	Third party monitoring systems can be used		
SLA Violation Credits (provider penalty)	Monthly Uptime %	Credits	Monthly Uptime %	Credits	5% for each 30 minutes network or infrastructure downtime and 5% for each additional hour up to 100% of the fees	Monthly Uptime %	Credits
	99.0% – < 99.95%	10%	<99.9%	10%	Users are not entitled to a credit if they are in breach of their services agreement with Rackspace	99.0% – < 99.9%	10%
	<99%	30%	<99%	25%		95.0% – < 99.0%	25%
				< 95.0%		50%	
SLA Exclusion	The service commitment does not apply to any unavailability, suspension or termination or performance issues that are due to the reasons stated in the SLA	SLA commitment doesn't apply in case any of the performance & availability issues stated in the SLA Exclusion section happens	Users are not entitled to a credit if they are in breach of their services agreement with Rackspace	Customer shall not receive any credits under this SLA in connection with any failure or deficiency of Service Availability caused by or associated with cases stated in it	The SLA does not apply to few exceptional cases stated in the document		

5.5. Proposed Cloud SLA Based on WSLA

A Web Service level Agreement (WSLA) is a reference for SLA monitoring between the provider and the user of web services that called deployment process. It gives a direct clarification for the, obligations, performance metrics, required performance levels and the corrective actions against violations that are associated with web service application. All these demands are checked according to an agreed contract for the level services of a web service. This includes the process of measuring the metrics and the algorithms used. Also WSLA can control the process of monitoring and managing the web service. This process may include all agreed obligations and interactions of multi-parties that involved

and then translated into system level arrangement. IBM has developed the WSLA in 2001. It uses XML language and deals with several technologies. This language includes many types of operations, special functions, and derived functions [23].

5.5.1. WSLA Users

WSLA users are classified into three types:

- Authors: They use WSLA to create their agreements.
- Implementers of deployment tool: They use WSLA to explain third party responsibilities and documents
- Authors of domain specific extensions of the language.

The language is used in defining the formal design of SLA in order to implement the web service system for the provider. Also, it is used to monitor and manage the agreed qualities for the user. The language should be designed to accommodate the changes to capture services development using agreed parameters. It is used to define, measure, and judge the performance metrics stated in the agreed contract. The language drives new types of syntactical elements from the high level ones in order to help the SLA authors to clarify their main concepts of their Web Service SLA. To enhance customer's confidence, the SLA will analyze all required information for the third party, so there is no need for accessing the full SLA.

5.5.2. WSLA Components

WSLA Components based on [11, 20]:

- Parties' characterization and their ways to follow the agreed contract and deal with other parties including provider, user and the third party. Web services are needed to be implemented in order to treat the behavior demanded by SLA. For example, a party for measuring performance is responsible for creating a web service to aggregate the alerts of other parties. Each party should provide his own information such as contact data, its organization, and description of implemented interface of web services. The main parties of any web service are the user and the service provider. In most of the cases there is a third party which works as assistant party. Contractual parties are classified into:
 - Signatory party (including their identification ,and their interface definition)
 - Supporting party (including their identification, interface definition and sponsors attributes)
- Defining SLA parameters or SLA metrics descriptions which clarify the service, its operations, how they are used for measuring the system and for whom. This ensures the understanding between the parties. Parameters should be specified for each operation separately. Each parameter has its own name, type, and unit. It refers to one metric. This metric can be used to define a function or measurements directive. While a function is used to derive the formula of computing the measurement algorithm and refers to specific schedule. This schedule refers to start, end, and time intervals of executing functions. A Directive Measurement is used to return the agreed metrics from the source.
- Service Level Objective: Shows the obligations followed in case of not meeting the contract goals or parameters. Also, the reactions in case of violating the service level objectives should be warranted. Obligations can be used to clarify the guarantees and constraints. The objectives can be defined as the summary of all elements of SLA parameters. For each operation there is a specified penalty. There are two types of obligations:
 - Service Level Objectives: Defines SLA parameters and preserves a particular situation of the system in a particular duration.
 - Action guarantees: Carrying out the related violation by the signatory party.
- The price: Stated for using the web services and is a clarification of penalties in case of violating the objectives.

5.5.3. WSLA Objectives

The Web Service Level Language has some goals to be pursued [24].

- The language needs to be adaptable, official, and flexible to suit all kinds of SLAs in order to narrow the variability in defining the quality of service parameters and metrics. And then, to be able to fit changes through time.
- WSLA should be compatible with electronic commerce and business systems. This will be suitable for taking care of advertisements, negotiation, order process, and payment systems for most business companies.
- Entrusting the third parties in monitoring the process to deal with some functions and responsibilities that both user and provider do not want to handle. In addition, third parties are playing an important role in supporting unit in measurement, evaluation, violation penalties, or computation.
- Distributing SLAs, which means the evaluating process of the web service will be divided and assigned to multiple SLAs providers. While all the chosen parties should be involved in a single SLA document with all agreed missions. Hence, each party will receive only the part of the document that is related to his responsibility. Having multiple parties will result in a complex SLA system because of security and privacy issues for the various parties; the provider and the user.

5.5.4. WSLA Architecture

Based on the findings of [20], the cloud is dynamic and its needs are also dynamic. Thus, the SLA has to follow these changes. So, the architecture of WSLA is specified as followed:

- Measurement services, which is classified as the third party to measure the agreed parameters of the web service provider resources. This third party is involved in security issues. SLA violations mostly occur in transition stages.
- A condition evaluation service, which gets the output from measurement unit and evaluates the service level objectives, then it, identifies the required violations.
- Management services, which take the appropriate actions in case of violations.

In Figure 9, the cloud provider and the cloud consumer are taking part in the negotiation process. Also, they are accepting a contract of parameters by distributing SLA Document. There is a control in the process of distributing this document so that not all parties will get the document for privacy issues.

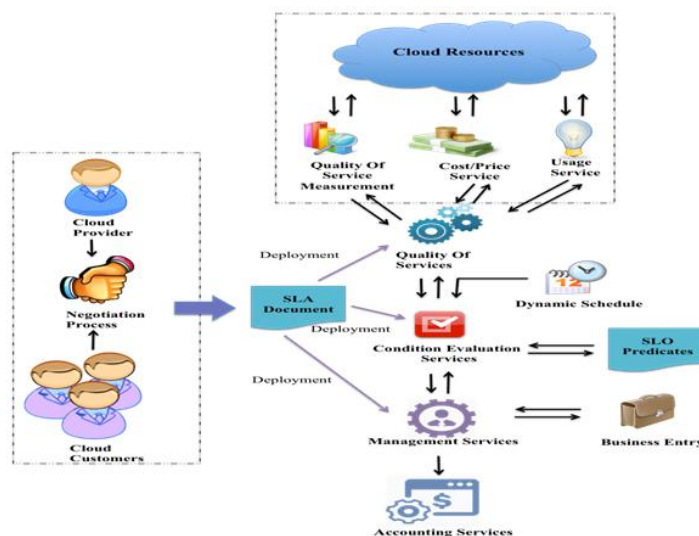


Figure 9. WSLA Architecture [47]

5.6. SLA Pricing

The service-level agreement between the consumer and the service provider states the lawful relationship between the two parties. This agreement protects the rights of both sides under any situation. SLA provides fairly knowledge about the service resources and features such as the quality of service and the price of providing the service [25].

The SLA Management part is the associated part of several resources, one of these resources is pricing. Pricing strategies are ways to determine the service price based on the service's demand and the service's equipment. Service providers use their rule to implement a smart pricing mechanism which will increase their profit. There are various charging strategies to set the service prices. For example, services price can be calculated built on the request peak/peak-off delivery time, service demand, service availability, service supply and charging rates even if it's static or dynamic. Most of the cloud computing providers use one of the three basic models for pricing which are; bid price model, static pricing, and dynamic pricing. Providers seem to prefer the dynamic pricing because they have to increase the service's price constantly, to maximize their incomes [26].

5.6.1. Dynamic Pricing: Dynamic pricing means the continuous altering of the service price. The price changes continuously based on the service supply and service demand. If the service demand increased or the service supply decreased, the price will rise up. And if the service demand decreased or the service supply increased, the price will go down [27]. Dynamic pricing has an influence on the SLA negotiation between the consumers and the providers because the price might change during the negotiation process. But, it is significant to understand that when an SLA has been approved, the price of that service level agreement must be fixed for the SLA lifetime rest. The formatting of an SLA contract and the particular price setting affect that specific provider-consumer collaboration only. The price might change if the same provider interacts with different consumers.

Price is calculated by certain functions that might be simple or complex depending on the variety and the quantity of its parameters. Simple functions will depend on few parameters. However, the complex functions will depend on many parameters. Those parameters can be measurements for the internal or the external state of the provider. Instances on internal state parameters are the service current, service loads, and historical data. External state parameters show the actual condition of the marketplace, but they are difficult to measure [28].

A critical parameter in service charging functions relates to the usage of the current service. The availability of resources is difficult to guarantee for future service supply and service demand. That's why most of the biggest cloud providers such as Amazon EC2, Microsoft Azure, Dell Boomi, and Google Cloud state at their SLAs that the service availability is 99.9%.

Another main parameter in calculating the service price is the risk. In some situations the service level agreement in cloud has fixed deadlines or extreme obligation so the negotiators from both sides should find a solution to cover the liability. So, insurance premium must be involved in the price.

The base cost of service such as both hardware and software purchasing, storing cost, and maintaining cost is another major price parameter. Some service providers may face business's problems in their start so they offer their services in lower costs than the base cost. But that can't last in the long term because successful business model must be gainful.

5.6.2. Price Architecture SLA Negotiating: Based on The Next GRID Project [29], The BREIN Project [30,31] SLAs negotiating price architecture consists of seven main functions which are; SLA template repository, resource capabilities, resource availability,

business objectives, dynamic pricing component, SLA negotiator and the signer as shown in Figure 10. They are in details below;

- SLA template repository: The provider sends non-obligated SLA templates to the consumer as a procedure to announce the offering services by the cloud.
- Resource capabilities: Resource Capabilities presents the documents and the data that concerned with the capabilities of the service's resources.
- Resource availability: This function supplies the up-to date data about the system latest status, containing existent load, predicted request and upcoming reservations.
- Business objectives: This component is more related to the provider of the service. It is the Logical clarification of his business preferences, performance, behavior and management...., etc.
- Dynamic pricing component: Dynamic pricing component calculates the service price based on the previous functions.

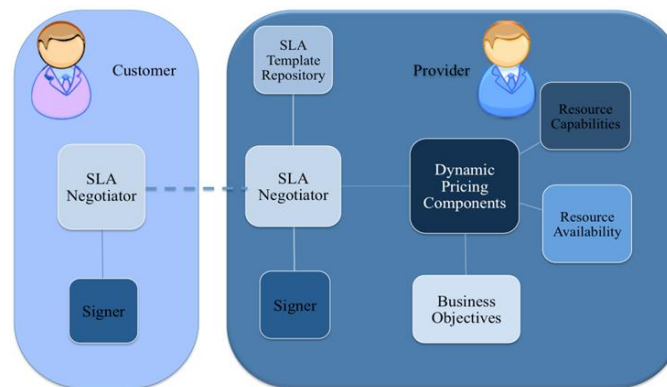


Figure 10. SLA Negotiating Price Architecture [40]

- SLA negotiator: This function is the core function that allows the consumer and the service provider to communicate and negotiate the SLA. The protocol of the negotiation process describes the messages which are sending to consumer's negotiators by the provider's negotiators and vice versa. Those messages might consist of quotes requests, the current quotes, deals, discount, and at last both approved and unapproved notifications.

- Signer: After the final agreement on the service's price, all parties including the provider and the consumers should present their approval digitally signing.

The architecture seems only a specific communication between a consumer and a provider. But in an actual system, several negotiations are rolling at the same period. Every single negotiation might be separate from the other negotiations [31].

7. Conclusion and Future Work

In this work, we examine the SLA for cloud computing services in depth, discussing the difference between web service SLA and cloud service SLA. Also, we discussed the requirements of cloud SLA, proposed metrics for SLA in clouds, and finally providing a comparison among the current major cloud service providers SLAs. As previously indicated, there is a very genuine need for a robust approach to handle SLAs in cloud computing. One important observation we made in the perspective of clouds is lack of standards. This is especially significant when we try to apply monitoring through various clouds. Although, it is applicable to accommodate several types of cloud interfaces by using a middleware, a universal set of metrics that aim at monitoring different cloud services is not yet implemented. Attempts to standardize a model for SLA for the clouds is really exist and we emphasize the significance of such efforts. For future work,

attempts will be made to contribute these efforts for standardizing SLA models for clouds by proposing further metrics for both SLA monitoring and standardized SLA monitoring framework. SLA pricing models needs further investigation as well.

References

- [1] P. Mell and T. Grance, "The NIST Definition of Cloud Computing", <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>, (2011) September.
- [2] R. Buyya, C. H. Yeo, S. Venugopal, J. Broberg and I. Brandic, "Cloud computing and emerging IT platforms, Vision, hype, and reality for delivering computing as the 5th utility", Future Generation Computer Systems, <http://www.buyya.com/papers/Cloud-FGCS2009.pdf>, (2009) June.
- [3] C. S. Yeo DeAssuncao MD, Y. J. Sulistio, A. Venugopal, S. Placek and M. Buyya, "Utility computing on Global Grids", http://www.buyya.com/papers/HandbookCN_Utility_Grids.pdf, (2006).
- [4] "Sun Microsystems, Service Level Agreement in the Data Center", (2010) March. <http://www.sun.com/blueprints>
- [5] L. Wu and R. Buyya, "Service Level Agreement (SLA) in Utility Computing Systems", The University of Melbourne. Australia. <http://arxiv.org/ftp/arxiv/papers/1010/1010.2881.pdf>.
- [6] E. Marilly, O. Martinot, S. Betgé-Brezetz and G. Delègue, "Requirements for Service Level Agreement Management", France, http://www-rp.lip6.fr/adanets/PublicDoc/Papers/IPOM2002_SLA_corrected_modifencoursv2.pdf.
- [7] M. Alhamad, T. Dillon and E. Chang, "Conceptual SLA Framework for Cloud Computing", Australia, http://ieeexplore.ieee.org/xpl/login.jsp?tp=&arnumber=5610586&url=http%3A%2F%2Fieeexplore.ieee.org%2Fxppls%2Fabs_all.jsp%3Farnumber%3D5610586.
- [8] A. Sahai, S. Graupner, V. Machiraju and V. Moorsel, "Specifying and Monitoring Guarantees in Commercial Grids through SLA", <http://www.hpl.hp.com/techreports/2002/HPL-2002-324.pdf>.
- [9] A. Keller and H. Ludwig, "The WSLA Framework: Specifying and Monitoring Service Level Agreements for Web Services", http://clip.dia.fi.upm.es/Projects/S-CUBE/papers/keller03:wsla_framework.pdf, (2003).
- [10] P. Allen, "Service Level Agreements", http://www.cbdiforum.com/report_summary.php3?page=/secure/interact/2006-12/service_level_agreements.php&area=silverH, (2003) .
- [11] P. Bianco, G. A. Lewis and P. Merson, "Service Level Agreements in Service-Oriented Architecture Environments", <http://www.sei.cmu.edu/reports/08tn021.pdf>, (2008) September.
- [12] "SLA Management Handbook, Concepts and Principles, TeleManagement Forum", http://www.afutt.org/Qostic/qostic1/SLA-DI-USG-TMF-060091-SLA_TMForum.pdf, vol. 2, (2005).
- [13] A. Pichot, "Dynamic SLA-negotiation based on WS Agreement", <http://cui.unige.ch/~dimarzo/courses/services/Lecture6.pdf>, (2008).
- [14] P. Rubach and M. Sobolewski, "Dynamic SLA Negotiation in Autonomic Federated Environments", <http://sorcsoft.pl/publications/papers/2009/OTM-2009.pdf>, (2009).
- [15] "The Open Group, SLA Management Handbook", Enterprise Perspective, http://www.afutt.org/Qostic/qostic1/SLA-DI-USG-TMF-060091-SLA_TMForum.pdf, vol. 4, (2004) October.
- [16] J. Kene, A. Skene, J. Crampton and W. Emmerich, "The Monitorability of Service-Level Agreements for Application-Service Provision", Proceeding WOSP '07 Proceedings of the 6th international workshop on Software and performance, New York, (2007), pp. 3-14.
- [17] S. Waldbusse, "Remote Network Monitoring Management Information Base", IETF RFC1757, (1995), February.
- [18] H. J. Lee, M. S. Kim, J. W. Hong, "Mapping between qos parameters and network performance metrics for sla monitoring", Asia-Pacific Network Operations and Management Symposium (APNOMS) Conference. 2002. <http://dpm.postech.ac.kr/papers/APNOMS/02/>
- [19] L. J. Tong, J. Mao, R. Bohn, J. Messina, L. Badger and D. Leaf, "Nist cloud computing reference architecture", NIST special publication, vol. 500, (2011), pp. 292.
- [20] P. Patel, A. H. Ranabahu and A. B. Sheth, "Service Level Agreement in Cloud Computing. Wright State University", <http://corescholar.libraries.wright.edu/cgi/viewcontent.cgi?article=1077&context=knoesis>, (2009).
- [21] "National Institute for Standards and Technology: NIST Cloud Computing Reference Architecture", www.nist.gov/customer/get_pdf.cfm?pub_id=909505, (2011).
- [22] M. Greer, "The Practical Guide to Cloud Service Level Agreements. Cloud Standard Customer Council", (CSCC), <http://www.cloudstandardscustomerCouncil.org/04102012.html>, (2012) April.

- [23] H. Ludwig, A. Keller, A. Dan, R. P. King and R. Franck, "Web Service Level Agreement (WSLA) Language Specification", <http://www.research.ibm.com/people/a/akeller/Data/WSLASpecV1-20030128.pdf>, (2003).
- [24] H. Ludwig and A. Keller, "The WSLA Framework: Specifying and Monitoring Service Level Agreements for Web Services", IBM Research Division. Thomas J. Watson Research Center, (2002) May.
- [25] H. Goudarzi, M. Ghasemazar and M. Pedram, "SLA-based Optimization of Power and Migration Cost in Cloud Computing", University of Southern California, Department of Electrical Engineering, http://sportlab.usc.edu/~hadi/index_files/SLA-based%20Optimization%20of%20Power%20and%20Migration%20Cost%20in%20Cloud%20Computing.pdf
- [26] A. Anandasivam and M. Premm, "Bid Price Control And Dynamic Pricing In Clouds", www.is2.lse.ac.uk/asp/aspecis/20090089.pdf.
- [27] V. Kamra, K. Sonawane and P. Alappanavar, "Cloud Computing And Its Pricing Schemes. India", <http://www.enggjournals.com/ijcse/doc/IJCSE12-04-04-127.pdf>.
- [28] M. Mihailescu and Y. M. Teo, "Dynamic Resource Pricing on Federated Clouds", <http://www.comp.nus.edu.sg/~teoy/pub/10/ccgrid2010-teoy.pdf>, (2010) May.
- [29] "The NextGRID Project", Priority IST-2002-2.3.2.8. <http://www.nextgrid.org/>.
- [30] "The BREIN Project", Activity IST-2005-2.5.4. <http://www.gridsforbusiness.eu/>.
- [31] P. Hasselmeyer, B. Koller and I. Kotsiopoulos, "Negotiating SLAs with Dynamic Pricing Policies", <http://www.hasselmeyer.com/pdf/socinside07.pdf>.
- [32] S. L. Frey, C. Uthje and C. Reich, "Key Performance Indicators for Cloud Computing SLAs", Furtwangen University, (2013).
- [33] R. N. Freire, P. Cunha and G. R. Justo, "An Approach for Reasoning and Refining Non-Functional Requirements", Journal of the Brazilian Computer Society, 10.1007/BF03192354, vol. 10, (2004), pp. 62–84.
- [34] A. L. Freitas, N. Parlavantzas and J. L. Pazat, "An Integrated Approach for Specifying and Enforcing SLAs for Cloud Services", Université Européenne de Bretagne, IEEE Fifth International Conference on Cloud Computing, (2012).
- [35] S. Frey, C. L. Uthje and C. Reich, "Key Performance Indicators for Cloud Computing SLAs", Furtwangen University, (2013).
- [36] A. Sahai, A. Durante and V. Machiraju, "Towards Automated SLA Management for Web Services", (2012) July.
- [37] B. I. Music, P. Wieder, R. Yahyapour and S. Dustdar, "Service Mediation and Negotiation Bootstrapping as First Achievements Towards Self-adaptable Grid and Cloud Services", In Grids and Service-Oriented Architectures for Service Level Agreements, New York, USA.
- [38] L. Lisa Brownsword, J. Carney, D. Fisher, D. Lewis, G. Meyers, C. Morris, Edwin, J. R. Patrick, J. Smith and L. Wrage, "Current Perspectives on Interoperability", <http://www.sei.cmu.edu/pub/documents/04.reports/pdf/04tr009.pdf>, (2004).
- [39] D. C. Verma, "Service Level Agreementson IP Networks", (2004).
- [40] P. Hasselmeyer, B. Koller and I. Kotsiopoulos, "Negotiating SLAs with Dynamic Pricing Policies", <http://www.hasselmeyer.com/pdf/socinside07.pdf>
- [41] H. Ludwig and A. Keller, "Defining and Monitoring Service Level Agreements for Dynamic e-Business", IBM Watson REasearch Centre, PA USA, http://www.cs.uoi.gr/~pvassil/downloads/WebServices/QoS/Ludwig/lisa2002_slides.pdf, (2002) April.
- [42] "SLA Management Handbook, Concepts and Principles", TeleManagement Forum, http://www.afutt.org/Qostic/qostic1/SLA-DI-USG-TMF-060091-SLA_TMForum.pdf, vol. 2, (2005).
- [43] H. J. Lee, M. S. Kim and J. W. Hong, "Mapping between qos parameters and network performance metrics for sla monitoring", Asia-Pacific Network Operations and Management Symposium (APNOMS) Conference, <http://dpm.postech.ac.kr/papers/APNOMS/02/>, (2002).
- [44] L. J. Tong, J. Mao, R. Bohn, J. Messina, L. Badger and D. Leaf, "Nist cloud computing reference architecture", NIST special publication, vol. 500, (2011), pp. 292.
- [45] H. Ludwig and A. Keller, "The WSLA Framework, Specifying and Monitoring Service Level Agreements for Web Services", http://clip.dia.fi.upm.es/Projects/S-CUBE/papers/keller03:wsla_framework.pdf, (2003).
- [46] A. L. Freitas, N. Parlavantzas and J. L. Pazat, "An Integrated Approach for Specifying and Enforcing SLAs for Cloud Services", Université Européenne de Bretagne, IEEE Fifth International Conference on Cloud Computing, (2012).

- [47] P. Patel, A. H. Ranabahu and A. B. Sheth, "Service Level Agreement in Cloud Computing. Wright State University", <http://corescholar.libraries.wright.edu/cgi/viewcontent.cgi?article=1077&context=knoesis>, **(2009)**.
- [48] "Amazon Web Services, Inc. Amazon EC2 Service Level Agreement", <http://aws.amazon.com/ec2/sla/>, **(2013)** June.
- [49] "Microsoft Azure, Service Level Agreement for Microsoft Azure", <http://www.microsoft.com/en-us/download/details.aspx?id=44584>, **(2014)** December.
- [50] "Rackspace Inc. Cloud Big Data Platform SLA", <http://www.rackspace.com/information/legal/cloud/sla>, **(2011)** February.
- [51] "Dell Boomi, Service Level Agreement", <http://www.boomi.com/sla>, **(2014)**.
- [52] "Google Cloud Platform, Google Cloud Storage, Google Prediction API, and Google BigQuery SLA", <https://cloud.google.com/storage/sla>, **(2014)**.
- [53] M. M. Q. Zhou and Y. C. Wang, "Systematic Analysis Of Public Cloud Service Level Agreements And Related Business Values", Services Computing (SCC). IEEE International Conference, **(2013)**, June-July.

