Conceptual Service Level Agreement Mechanism to Minimize the SLA Violation with SLA Negotiation Process in Cloud Computing Environment

Surendra Kumar  Narander Kumar*

Babasaheb Bhimrao Ambedkar University (A Central University), Lucknow, INDIA

*Corresponding Author: nk_iet@yahoo.co.in, kumar.surendra1989@gmail.com

ORCID ID: https://orcid.org/0000-0003-4633-1879, https://orcid.org/0000-0003-1718-8102

Received 16/5/2020, Accepted 31/3/2021, Published 20/6/2021

Abstract:
Online service is used to be as Pay-Per-Use in Cloud computing. Service user need not be in a long time contract with cloud service providers. Service level agreements (SLAs) are understandings marked between a cloud service provider and others, for example, a service user, intermediary operator, or observing operators. Since cloud computing is an ongoing technology giving numerous services to basic business applications and adaptable systems to manage online agreements are significant. SLA maintains the quality-of-service to the cloud user. If service provider fails to maintain the required service SLA is considered to be SLA violated. The main aim is to minimize the SLA violations for maintain the QoS of their cloud users. In this research article, a toolbox is proposed to help the procedure of exchanging of a SLA with the service providers that will enable the cloud client in indicating service quality demands and an algorithm as well as Negotiation model is also proposed to negotiate the request with the service providers to produce a better agreement between service provider and cloud service consumer. Subsequently, the discussed framework can reduce SLA violations as well as negotiation disappointments and have expanded cost-adequacy. Moreover, the suggested SLA toolkit is additionally productive to clients so clients can secure a sensible value repayment for diminished QoS or conceding time. This research shows the assurance level in the cloud service providers can be kept up by as yet conveying the services with no interruption from the client's perspective.

Keywords: Cloud Computing, QoS, Service Level Agreement, SLA Negotiation Model, SLA Toolkit.

Introduction:
A significant component that gives some level of confirmation to the two clients and service providers of these cloud resources is the Service Level Agreement which characterize the extent of use and arrangement of resources. Cloud service user need a SLA before they move their framework to cloud server, to give sureness with respect to the resources gave and the capacity to arrive at the ideal degree of efficiency. Cloud service providers need a SLA to characterize the trust and Quality of service (QoS) they give to clients just as a concurred system for cost and charges. The exploration on SLA and QoS measurements has been considered by researchers, for example, web based applications and e-commerce. In any case, SLA measurements in these advancements are not appropriate for cloud computing as the quality and resource type being given and sends is different. Thus, new SLA models are as yet required to give adaptable technique to negotiate and the marking of electronic agreements among customers and service providers.

Cloud computing focuses on the ubiquitous, advantageous, on-request server access to a shared processing resources in cloud computing such as servers, online applications and cloud services) which can be quickly provisioned, sent to insignificant balance exertion or cloud service providers (1). The resources are given for Cloud Services Provider (CSP). CSPs are conveying its cloud services by cloud deployment models like, Infrastructure-as-a-Service (IaaS), Software-as-a-Service (SaaS) and Platform-as-a-Service (PaaS). CSPs are utilizing the SLA to characterize the quality of services and execution that will be conveyed to service user. Specifically, this characterizes status and attributes for delivering services as per SLAs. Services are delivered to clients should pay for per-use for each utilization.
framework and maintain the QoS ought to be related to the SLA guaranteed parameters (2).

Basically, consider the circumstance wherein an employment is basic to the accomplishment of the business. On a primary level, the customer might have demanded a penalty in the first SLA that is equivalent to the estimation of customer's business, as remuneration for the failure because of the disappointment of that business demanding work. However, this is typically unreasonable, since such a penalty can be bigger even than the complete resources of the service providers. Thus, the customer will never be completely redressed, and the supplier faces lost future customers because of degrade in reputation as the quantity of violated positions aggregates. Therefore, the supplier and the customer would typically like to renegotiate utilizing another SLA with another related time (i.e., an expansion past the main deadline), as opposed to take care of a precarious penalty and acknowledge the wiping out of a business demanding work. The new SLA will for the most part remember a discount by the supplier for the initially settled upon cost, as a concession by the supplier for escape the precarious penalty conceived by the first SLA violation.

Key challenges for this research is as follows,

- Firstly, the main problem deals with agreements for major parameters which are modified by SLA violation to provide real-time results to minimize the more violation as well as mannered process for timely consideration.
- Secondly, to manage the regular occurring violations in cloud computing environment and a most effective mechanism is much needed to deal with to captured service performance to send a response if any problem occurred.

Main objectives of this research are as follows:

- Designing a VM management architecture framework model for resource management architecture with significant proportions of the application and SLA analysis.
- The introduction of SLA life cycle which shows the creation of SLA and controlling the operations as well as tells when the SLA negotiations take place and role of cloud service providers.
- Designing a resources allocation decision is the basic leadership of allotment fractions for given to each VM and significant decision variable in our resource assignment issue as well as SLA framework model.
- Designing a model for cloud service user will have the information required to choose the most reasonable cloud service providers with the most significant level of affirmation of covering every single required SLA parameters.

To improve the collaboration with associated objects to convey new sort of services to SLA policies for cloud service providers and cloud service users. This condition requires provoking procedures to offer secure services with a Quality of Service (QoS) ensure so as to fulfill clients' desires. A framework empowering cloud services conveyance with QoS ensures is presented in this paper. The proposed framework depends on a particular SLA life cycle just as various Service Level Agreements (SLAs) and co-operations so as to accomplish a worldwide cloud to minimize SLA violation.

The major contributions of this research are as follows,

1) Proposing an SLA renegotiation model for cloud system to target the customer as well as service provider desired for an optimal renegotiation.
2) Proposing the SLA framework to understand the working and cloud service client will have the information required to choose the most reasonable cloud service providers with the most significant level of affirmation of covering every single required SLA parameters.
3) To minimize the violation a proposed mechanism is introduced for service providers.

Different sections are as follow: Section I, contains the overall introduction between Service Level Agreements to cloud service providers and cloud service users. Section II, contains the related works done previously. In Section III, an algorithm and framework is proposed for cloud service providers to SLA negotiation and minimize the SLA violation. Section IV contains the results validations and testing the results by the algorithm. Section V contains the conclusion of the given algorithm and a proposed framework.

Related Work:

They categories the SLA issues and analyze the SLA issues. Results are basically focused on academically or limited industrial setting intend of real industrial environments and the lack of tools as well as evolution perspectives supports in SLA managements and focused on assessments like qualitative and quantitatively of QoS in SLA (3,4). In the cloud computing SLA-aware deployment are take place and analyze workload with case study. The technique depends on the models and controls the tools as well as runtime. So it fitted in the combining the software design with DevOps (5).
To secure the cloud application moving target defuse approach is used. So it can exchange, admissible application configuration to confuse the hackers and invalidate the surveillance exertion while safeguarding the application (6). Service provider’s deals with the various issue like based stock (penalty) and review period, they affect the cost with multiple service users. So a service provider prefers to deal with more service users to maintain the demands and large performance review is helpful under the penalty contracts (7-9).

Autonomous SLA Negotiation framework, adaptable agents as well as multi agent frameworks based on the suitable tools for self detection of declines and controlling for the operation as well as services in the cloud computing (10). To remove the redundancy and network externality is analyzed for the cloud service providers with resource assurance by optimal contract design in cloud computing. This model characteristic is analyzed for optimal costing and resource allocation decisions as well as made a contract of cloud service providers which shows that both service providers and service users compensation ratio is unfulfilled (11).

SLA is used as a rule for service user and service providers to provisioned service observed with maximum profits for both and detailed review of the existing SLA assurance issues in the cloud computing (12). Analysis of various Meta heuristics techniques is to optimize the SLA violation cost in cloud computing and algorithm gives the better handling to multifarious constraints. To improve the SLA penalty cost and constraints have optimized by mathematically to QoS which gives the feasible domain to maximize the profits of service providers (13).

Dynamic SLA’s Negotiation system consistent with dynamically which is overcome with mathematically as well as fuzzy models in the cloud computing and defines the preconditions, take decisions to provide the complete access without hang-up SLA at the runtime (14). To assure the QoS a context aware framework is used for SLA and defines the contextual parameters for cloud service users. Using the reasoning’s they predict SLA violations selections as well as implement it and getting proactive balancing between cloud services and SLA parameters (15).

Reputation-based Trust Management (RTM) defines the performance matrix and measures it for assigned cloud services. Various techniques and user experiences matrices compute to check the quality of cloud services and validity of the services (16). Cloud based framework for a real time and SLA renegotiation model is designed. Various situations are dependent on parameters for renegotiation. So multi offer and real time techniques are used to detect and predict SLA violations for renegotiation (17).

To maximize the utility values and the execution rate without negotiate and simulate with efficient strategies which show the highest utility rate and execution rate using learning mechanism without any negotiation conflicts for cloud service providers (18). A negotiation approach is designed to increase the benefits for service providers, cloud service users and simulation place to evaluate the performance of negotiation techniques for speed and negotiating utility (19).

Basically SLA based cloud framework is used in resource allocation which manages the workload and location of the datacenters in cloud computing. They design a cloud framework which performs the auto negotiation and a workload- and location-aware resource allocation (WLARA), implement the cloud testbed framework and compare with existing approaches (20). Resource manager and VM reallocation have some special components and various techniques are used. ETC matrixes with PSO algorithm are basically introduced for SLA priority clustering and simulate with the help of CloudSim which gives the efficient reallocation results in cloud computing (21). SLA violation detection model used Markov decision process theory to avoid the violation in cloud computing. Cloud service provides needs some specific requirements from service user which are recognized by the model and evaluate the credibility of service providers as well as monitor the events that cause for violations (22). Decrease the latency in term of cloud computing by using the fog computing and research shows many advantages like, scalability, privacy, resource processing, security and low latency (23). Fog computing increases the performance of tasks using machine learning applications to get significant results (24).

Related works show very well, researchers have managed the inconsistent conduct of the supplier's resources and the client's prerequisites by checking the request status during the run-time. Proposed mechanism is introduced to the clients prior to beginning the SLA instatement. Requests are viewed as the foundation of setting up and arranging the proposed framework where they are replied by the client before the start of the exchange cycle and afterward, they are utilized to help boundaries to system in the SLA arrangement.

**Proposed Materials and Methods:**

**SLA Based Framework Design**

The QoS is a basic approach in cloud environment where redistributing is utilized. The
significance of services as an approach to determine, its utilization on data and security for such SLA. The SLA is a piece of the contract among cloud service providers and client, and depicts the ideal QoS. A SLA alone doesn't ensure that the predetermined characteristics are met, so it characterizes the vital observing instruments, the importance and characterizes disciplines and pays if conditions are not met.

In cloud computing, SLA use adopts a various techniques in comparison to that of cloud service providers. The agreements will represent to both clients and service providers desires. Likely, commitments might be indicated for the two gatherings. The extent of data contained in the agreements is likewise separated. It is pivotal to point that, for SLA, service measurement observing is as significant as their determination. For this reason, measurements are applied to evaluate consistence with the ideal characteristics of services. The manner in which these measurements are estimated relies upon the kind of services and quality key features that one needs to evaluate.

Figure 1. Framework Architecture of VM management for SLA in Server.

The framework model expected for virtual machine managements in a single machine given in the Fig. 1. The significant part in this model is management of the resource allocations module which is to deal with all request of workloads task and to support these workload requests by acknowledging framework attributes, properties of their applications, VM accessibility and SLA guarantee for increasing the income of the cloud service providers. Resource management architecture is designed with significant proportions of the application and SLA analysis. This analysis is refreshed by application or SLA development in cloud computing. Figure 2 shows a SLA life cycle for cloud service provider to maintaining the service request of user.
Working of SLA based Proposed Mechanism:

In this systematic framework architecture, resource allocation is constrained by parameter of SLA as well as some other parameters of framework. It is evaluated analysis determined from parameters of SLA applications request of workloads plays a decision-making role. Cloud computing goes for services proficiency is particularly related to SLA parameters, thus the SLA execution has identified with the VM's storage support services by fulfilling the service reaction time determined in the SLA. Parameters of SLA, focusing on throughput (T^<sup>T</sup>) of the service, reaction time threshold (RT^<sup>T</sup>), probability PB(), and the framework parameters are total number of VMs (TV) made by the virtualization layer, the used value (u_i) of each VM and the service time (ST) of the application on a physical server. Among these qualities, the used values are observed as the most extreme value given by the cloud service providers and the throughput value is the max throughput or throughput threshold (T^<sup>TH</sup>).

Resource management part gets a estimates AR_i (arrival rate for requests) from the mentioned requested workload segment for every application during the upcoming controller interim. On the off chance that there happen a few deviations in arrival rate from the evaluated progression, at that point a optimization of load is started and the deviations are enhanced with the optimization. Here, AR_i is the arrival request rate over the thought about controller intervals. From these showed up demands, some might be dismissed as a result of resources confinements thus the real appearance rate gets lesser than AR_i. Among the prepared arrangement of requests, some may violate the concurred reaction time thus they are not taken for the estimation of r AT_i (actual throughput).

If there should be an occurrence of fixed controller intervals, the occasions which are extensively less than the controller interval could deceive the allotment, for example, mass quantum of request (with less term), originating from certain applications can stop the assignments of resources to some different class of uses due to lack in resources accessibility and this will prompt substantial penalties to the service providers. For limiting this undesired impact, the mentioned requested workload component part gives the estimated probability (EP_i) of a class of requests having higher rate of arrivals for the upcoming controller intervals. The estimated parameter EP_i is to show conviction level to guarantee greatest benefit to the service providers for VM_i, and it very well may be skirted, to consider outstanding workload changes, by set the value 1.

To continue with the framework model, it is expected that the application started from a client is a one of a kind substance and is submitted to specific VM. Application is starting when client submit to VM_i which is adjusted in a ST_i (mean service time), and the usage value to maximum point of limit of VM_i is ul_i. In this model, each VM_i is qualified for an ensured portion of accessible physical server and the average service time is measure by taking fn_i, the division of service time given to VM_i as ST_i/fn_i. Additionally, resources allocation decision is the basic leadership of allotment fractions fn_i (i = 1, 2 . . . V) given to each VM_i. Along these lines, fn_i is seen as the significant decision variable in our resource assignment issue and SLA framework model.
Proposed SLA Negotiation Model:

Figure 3. SLA Negotiation for cloud service provider and customer relationships.

Figure 3 shows SLA negotiation process for cloud service provider and customer relationships. Some conditions are satisfied by the negotiation process. If a customer needs to change their current preference then customer has to change the weightage of every parameter in next cycle. When CSP needs to change preference for considering next upcoming maintenance or limited resources may affect the services requests. So CSP will negotiate and modify weightage. CSP and customers are adjusting the weightage for prior negotiation. If CSP predict the SLA violation then SLA renegotiation takes place and weightage of parameters are automatically restructured.

Result and Discussion:
Step.1: Prior to the discussion with the service providers, the user ought to distinguish all the necessary SLA parameters utilizing the proposed SLA parameters recognizable form.

Step.2: During the beginning of negotiation procedure, the user needs to examine each and every chosen SLA parameters with service providers to characterize degree of service needed, degree of consistence for every cloud service providers. At the point when the service providers consent to a given needs of client, relating SLA parameter ought to be selected as "Yes", else "No" is picked.
**Explanation:** The parameter of data location is needed for the service client X. Subsequently, service providers will perceive that it is important for cloud service clients to know where the information is located. On the off chance that service providers consents to give such data, "Yes" is picked. Then again, when service providers will not show the location, at that point "No" is picked as well as cloud service client ought to write in the Notes segment whether cloud service providers make some other arrangement, such as locating that the information could be put away in a various locations however not the accurate locations. These notes will be utilized in the following stage.

**Step.3:** After gathering with picked cloud service providers the client needs to analyze the degree of consistence for every cloud service providers by confirming which cloud service providers offer clings much more to the client's necessities or contains appropriate other options.

**NOTE:** This analysis can be effectively acted in a simple structure by doling out numeric estimation of "1" for "Yes", an estimation of "0" for "No", a reflecting weight scale significance of parameter (i.e. 0 to 5), evaluating average weighted for each parameter in the given form. If there is service providers with a similar level of consistence the Note segment ought to be surveyed for any additional data that can recognize them.

**Step.4:** Finally, utilizing all recently given data, cloud service client will have the information required to choose the most reasonable cloud service providers with the most significant level of affirmation of covering every single required SLA parameters.

![Image of SLA working flow diagram](image)

**Figure 4. SLA working flow diagram for cloud service provider and service users.**

Figure 4 shows SLA working flow diagram for cloud service and service users. The majority of the related work accepts that once established a SLA can't be changed and when violated the service is consequently discard, without first permitting the cloud service provider and client the choice of renegotiation. The service provider takes care of a high penalty and loses notoriety, while the purchaser may have a request discard. A SLA renegotiation model and life cycle are designed to deal with proactively handles such SLA violations. The analyzing model makes conceivable a better situation (guaranteeing congruity of cloud service and lessening SLA penalty costs). The dynamic procedures depend on a utility capacity for the supplier and planning data produced by the rescheduling alternative picked before renegotiation. The presented SLA renegotiation model produces to a critical advance to empower a change of the SLA boundaries over the span of services. Furthermore, by utilizing the methods presented, the trust level in the service providers can be kept up by as yet providing the services with no interruption from the client's perspective. In the extended works, incorporate finishing the situation where violation is examined during service
providers and consolidating service cost count into the cloud environment.

In this work, Cloudsim is utilized to test the SLA toolkit and recreate the SLA cycle model as well as dissect the quality of the service. Cloudsim is an extensible that bears the cost of fundamental classes for characterizing clients, resource, virtual machines (VM), server, cloudlets and numerous different methodologies for the administration of various components of the framework. The major objectives of CloudSim are: it takes minimum time or exertion is for actualizing the distributed computing conditions. Simulation components cover different client characterized arrangements for assigning host to VMs, more adaptable to appoint various centers to more virtualized administrations. SLA toolkit is made out of five fundamental attributes (self assistance in On-request, network accessibility, Resource scheduling, Rapid versatility, calculated services), to target three model Saas, PaaS, IaaS as well as various cloud deployment models.

A. Hypothesis:

- Proposed a toolkit or a conventional model, which is fit for assessing a specific business or private client, regardless of whether he ought to decide on a cloud IaaS services.
- The cloud service supplier and client ought to and could work together, to formalize the best SLAs negotiation between them.

This research is focused on pro-active as well as continuous renegotiation to keep a specific degree of trust, particularly on the cloud service provider (CSP) side. To help the pro-active methodology, the CSP should evaluate the services conveyance in an opportune way to estimate the chance of any violation. SLA violation grouping is acquainted with improving the expectation. At the point when SLA violation is anticipated, the satisfactory estimations of QoS will be re-define simultaneously in the renegotiation cycle. Client will at that point pick the most appropriate one. From client's perspective, this choice is seemingly better to be taken as opposed to ending the service.

The limitation is if the missing matrix founded so penalties are made by SLA and under this situation customer takes over advantage of them. Service negotiation techniques have some various systems to accomplish an understanding between the cloud service providers and clients. Despite the fact that negotiation can resolve the contention with no total data given among the providers, the best provider offers counter-offers procedure is a time taken process in arriving at the arrangement. So, this negotiation process is a success-lose negotiation that is less reasonable for renegotiation. Offer validity based on negotiation gives a mutually beneficial stage and can produce multi-offer simultaneously, which makes this reasonable for renegotiation.

Conclusion:

Service level agreement (SLA) is a mechanism planned to help in characterizing the quality of given service to cloud clients and distinguish mentioned demands for service providers. It contains numerous perspectives that cloud service users need to talk about during the negotiation procedure with service providers. For scaling parameter, SaaS suppliers need to lease resources from IaaS suppliers to run their profoundly scalable applications. To increase their benefit and to fulfill customers, a SaaS supplier utilizes a SLA toolkit model that effectively allocates customer solicitations to leased resources. Since the cloud is exceptionally powerful, sudden occasions may happen that adjust the initially chosen parameter and lead to SLA violated. A SLA negotiation model is proposed to verify the resource requirements for every application and accept the system. The proposed model is ready to manage the resource provisioning according to the demands. Hosted applications and assignments of VMs on the physical machines (PMs) are also considered. In additionally, the proposed SLA toolkit is very profitable to include other phase of the SLA Lifecycle like SLA creations and SLA controlling in cloud computing environment. This research shows the assurance level in the cloud service providers can be kept up by as yet conveying the services with no interruption from the client's perspective. The proposed SLA toolkit benefits the client demands from its interlinked cloud datacenters and has an SLA Negotiation mechanism to reduce the cost for penalty and increase consumer comfort by minimizing SLA violations. Future perspective will aim on the various types of SLA parameters and presents a renegotiation framework to provide a significant process to make changes in SLA parameter during performing cloud services operations as well as CPS can be maintained by providing services without any interruptions from client side. Additionally, the estimated service delivery and incorporation cost into current system are also computed.

Abbreviations:
Cloud service Providers (CSP), Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS), Physical Machines (PMs), Quality of Service (QoS), Service Level Agreement (SLA), Software-
as-a-Service (SaaS), Virtual Machine (VM), workload- and location-aware resource allocation (WLARA).

Acknowledgment:
Declared none.

Authors' declaration:
- Conflicts of Interest: None.
- We hereby confirm that all the Figures and Tables in the manuscript are mine ours. Besides, the Figures and images, which are not mine ours, have been given the permission for re-publication attached with the manuscript.
- Ethical Clearance: The research paper (project) was approved by the local ethical committee in University of Babasaheb Bhimrao Ambedkar.

References:
الخلاصة:

تُستخدم الخدمة عبر الإنترنت لبديل الدفع لكل استخدام في الحوسبة السحابية. لا يحتاج مستخدم الخدمة إلى عقد طويل مع مزودي الخدمة السحابية. اتفاقية مستوى الخدمة (SLAs) هي إتفاهمات تم تحديدها بين مزودي الخدمة السحابية وغيرهم، على سبيل المثال، مستخدم الخدمة أو المشغل الوسيط أو المشغلين المراقبين. نظرًا لأن الحوسبة السحابية هي تقنية متعددة، تقدم العديد من الخدمات لكي تكون متاحة، وتتراوح من الخدمات العامة إلى الخاصة. إذا فشل مزود الخدمة في الحفاظ على الخدمة المطلوبة، فإن اتفاقية مستوى الخدمة تعتبر متفق عليها للاستمرار في اتفاقية مستوى الخدمة. الهدف الرئيسي هو تقليل انتهاكات اتفاقية مستوى الخدمة (SLA) للحفاظ على جودة الخدمة لمستخدم السحابة. في هذه المقالة البحثية، اقترحنا صندوق أدوات للمساعدة في إجراء تبادل اتفاقية مستوى الخدمة مع مزودي الخدمة والذي يمكن العميل السحابي من الإشراء إلى متطلبات جودة الخدمة. واقترح خوارزمية بالإضافة إلى نموذج التفاوض من أجل التفاوض على الطلبات، مع خدمة متقدمة لخدمات صغيرة لإنتاج اتفاقية أفضل بين مقدم الخدمة ومستهلك الخدمة السحابية. وبالتالي، يمكن للطيار تقليل انتهاكات اتفاقية مستوى الخدمة وكذلك توجيه الأمل في المفاوضات. ويعتبر نماذج كفاية التكلفة علامة على ذلك، فإن مجموعة أدوات اتفاقية مستوى الخدمة المتقدمة منتجة بشكل إضافي للعملاء حتى يمكن العملاء من تأمين سداد قيمة معقولة مقابل تقديم جودة الخدمة وأوقات التنازل. يوضح هذا البحث أن يمكن الحفاظ على مستوى الضمان في موفري الخدمات السحابية من خلال نقل الخدمات دون انقطاع من منظور العملاء.

الكلمات المفتاحية: الحوسبة السحابية ، جودة الخدمة ، اتفاقية مستوى الخدمة ، نموذج تفاوض اتفاقية مستوى الخدمة ، مجموعة أدوات اتفاقية مستوى الخدمة.