LITERATURE REVIEW: IMPROVING THE QUALITY OF SERVICES IN CLOUD COMPUTING ENVIRONMENT

1Manila Gupta, 2Devendra Singh, 3Bhumika Gupta

1Manila Gupta, PhD Scholar, Department of Computer Science, IFTM University, Lodhipur Rajput, Moradabad.
2Devendra Singh, HOD, Department of Computer Science, IFTM University, Lodhipur Rajput, Moradabad.

ABSTRACT
Cloud computing is an archetypal for conveying data know-how facilities and users get all the resources through internet. It expedite pay-per use value model for computing services. The requirement has suddenly increased for cloud computing services as the companies have move to cloud for their services and the cloud providers necessitate to offer services based on the likely quality requirements. The main task in cloud computing is to manage the quality of services (QoS) i.e. which is the difficulty of assigning resources to the applications to pledge service based on performance, availability and reliability. In this paper, we have presented a survey on the quality of service (QoS) in cloud computing with reference to techniques used, advantages and disadvantages. We combined all the related works of quality of service in cloud computing. The aim of this paper is to present a study of previous works done on QoS methods used in the cloud computing environment.

KEYWORDS - Cloud Computing, Quality of Service (QoS), Service level Agreement (SLA), Software Defined Network (SDN).

1. INTRODUCTION
The emergent tendency in know-how in this present era is cloud computing. It’s a service oriented technology over the internet. It’s a unique kind of a service over the internet as it provides computing as a service. Cloud computing is a kind of work out technology which can be well thought-out as a novel model of computing. Over the internet many cloud users required a variety of services as per their needs. So, there is proper demand and supply ratio but due to limited resources, it is very difficult for cloud providers to provide all the users desired services. From the cloud providers, it’s a perception that cloud resources must be assigned in a coherent manner. And it is difficult to meet cloud users satisfaction and QoS requirements. With the growing development internet is very important in every field of business, research and many other fields, so cloud computing provide the whole lot of services either it is in the form of software, hardware or platform.

Cloud Computing made of two words: Cloud and Computing, where cloud is a bank of services which contain interfaces, hardware, network & software for the end users and Services in the clouds can be software, platform & infrastructure as a service which are delivered to client via use of internet. With the help of Cloud computing any users can access any resources without the need to purchase it, sometimes it provide the Cloud Computing come apart in two words i.e. cloud and computing. In which cloud is a bank of services which contain interfaces, hardware, network, and software for the end users.

A service in the clouds refers as software, platform and infrastructure as a service which are delivered to client via use of internet. It facilitate client to carry out any work on the computer or any other devices which can access internet and perform task with easy installing of any software without buying it, or can pay some amount to get access of that service. The service level agreement signed by the client and vendor before delivering the services to the client. [1].

1.1 Models of cloud computing
Cloud Computing works on SPI model. The framework for cloud computing includes Software-as-a-Service (SaaS), Infrastructure-as-a-Service (IaaS), and Platform-as-a-Service (PaaS) [3]. These are described as under:

Fig 1: Services Provided by a Cloud
a) **Software-as-a-Service (SaaS)**
A Software Distribution model in which applications are hosted by service providers and made available to the clients on the internet, thereby minimizing the need to install and executes the application on the client machine. The key provider who uses SaaS is Amazon Web Service, Salesforce, Google Apps, Facebook and many more.

b) **Platform-as-a-Service (PaaS)**
A computing platform where user does not need to buy or install the hardware and software in the cloud infrastructure. It consists of applications for computing that are required by the consumer. PaaS also provides various resource management functions for on-timer scheduling of processes.

c) **Infrastructure-as-a-Service (IaaS)**
It consists of various resources like storage, network and processing capacity required by the user. The client can regulate the cloud environment as a service and client has to pay only for the duration of time whenever they avail the services. This results in the faster service delivery with low cost. The client can still use the service without the awareness of the storage location. This is also called Hardware-as-a-Service (HaaS).

![Cloud Development Model](image)

**Fig 2:** Cloud Development Model

### 1.2 Cloud Deployment Models

a) **Private cloud:** A private cloud is owned by a singleorganization. Private cloud enables an organization to use cloud computing technology as a means of centralizing access to IT resources in the organization. The actual administration of a private cloud environment may be carried out by internal or outsourced staff. This model has tremendous value from a secure to point of view[2].

b) **Public cloud:** A public cloud is a publicly accessible cloud environment owned by a third party cloud provider. The IT resources on cloud computing are usually provisioned and offered to cloud providers at a cost. The cloud provider is responsible for the creation and on-going maintenance of the public cloud and its IT resources.

c) **Hybrid cloud:** A hybrid cloud is a cloud environment comprised of two or more different cloud deployment models. This deployment model helps business to take advantage of secured applications and data hosting on private cloud, while still enjoying the cost benefit by keeping shared data and applications on the public cloud.

d) **Community cloud:** A community cloud is similar to public cloud except that its access is limited to a specific community of cloud users. The community cloud may be jointly owned by the community members or third-party cloud provider that provisions a public cloud with limited access.

### 2. QUALITY OF SERVICE (QoS) TECHNIQUES

Quality of Service (QoS) indicates the level of performance, reliability, availability obtainable by a application and by the platform or infrastructure that hosts it. QoS is elementary to cloud users, who look forward to providers to deliver the present characteristics, and for cloud providers, who necessitate finding the right trade-off between QoS levels and operational cost [4]. Finding out the finest substitution is not an easy decision dilemma because it rivets Service Level Agreements (SLAs) which spell out QoS targets and economic penalties coupled with SLA violations. Service providers need to act in accordance with SLA contracts which conclude the profits and penalties on the foundation of the accomplishment performance level[5].

Service Level Agreements (SLA’s) are a precursor between the service provider and the customer where SLA violation does something as major restraint. SLA violation is concentrated through mechanisms concerning monitor [6].
In particular, the author proposed a mechanism called DEDCA (Dynamic Enhanced Distributed Channel Access) to deal with the channel access in wireless networks, and a framework that make possible its performance in 802.11-based wireless networks using SDN technology. The key fact of this alternative solution is be in command over the contention window size of the wireless terminals. Thus, an adequate response to dynamic and short-term Quality of Service (QoS) requirements can be accessible to services running on these networks. DEDCA mechanism relies in the lead of use of a scalar parameter called gain. The mathematical model which has endorsed us to attainthis parameter is obtainable and calculates in this paper[7].

In this paper[8], the author spotlight on Quality of Service (QoS) as an high-flying factor make a difference to service provider’s outlook in choosing between a centralized SDN versus a distributed traditional routing network. The proposed prototype of the SDN expresses comparable QoS performance to traditional networks. This paper uncovers ways to diminish latency and suggests improvements in QoS using a practice built OpenFlow controller based on Floodlight. The networking tool iPerf is warned to assess the QoS performance characteristics. The results point towards capital cost savings of 20% and operational cost savings slam to 43% upon relocation to a white box SDN from traditional networks.

The comprehensio of the Quality of Service (QoS) conception happen to understandable with SDN in a convenient way. In this paper [9] center of attention is on the active architectures parameters such as response time, switch capacity and bandwidth isolation and we assesses these parameters here. Open Flow as the best-known SDN standard so far defines a standard protocol for network control. These interpretations of change variability may provide SDN application developer’s insights when realizing QoS concepts in an SDN-based network.

[10] In this survey paper, we aim at building a picture of QoS-motivated narrative in OpenFlow-enabled SDN networks by exhaustively surveying appropriate research studies. We systematize the related studies according to the grouping that are the most important ways in which QoS can benefit from the concept of SDN: Multimedia owe routing mechanisms, inter-domain routing mechanisms, resource reservation mechanisms, queue management and scheduling mechanisms, Quality of Experience (QoE)-aware mechanisms, network monitoring mechanisms, and other QoS-centric mechanisms such as virtualization-based QoS provisioning and QoS policy management etc.

[11] In this paper, a two-step approach is projected for SDN controller selection. First, the controllers are ranked with analytical network process (ANP) according to their qualitative features which persuade the performance of these controllers and then a performance comparison is achieved to verify for the QoS improvement. The most important part of this paper is read-through the applicability of the ANP for controller selection in SDN taking into consideration its features and performance analysis irreel-world Internet and Brita topologies. The recreation results illustrate that the controller work outthrough the proposed approach outperforms the controller selected with activeappraches. The selection of a most favourable controller with ANP results in a diminution of topology discovery time and delay in the normal and traffic load scenario.

Though QoS properties have inward bound constant attention even prior to the development of cloudcomputing, performance and heterogeneity and resource isolation system of cloud platforms have notably intrinsic QoS analysis, prediction and assurance. Thus, several researchers are look into automated QoS management method that can influence the high programmability of hardware and software resources in the cloud [12].

P. C. Hershey et al. [13] proposed a SOS approach to facilitate QoS monitoring, management and response for enterprise systems that convey computing as a service all the way through a cloud environment. Enterprise Monitoring, Management and Response Architecture in Cloud Computing Environments (EMMRA CC) unmitigated previous work to make available structure from which to identify points within the administrativedomains where QoS metrics may be monitored and managed. A realinstance was endow with forapplying the new SOS approach to a real world scenario (viz., distributed denial of service (DDoS)). The approach is very effectual but it was not applied to federated clouds in real time.

M. Salam et al. [14] proposed a QoS-oriented come together cloud computing framework where multiple independent cloud providers can lend a hand effortlessly to make available scalable QoS-assured services. The key elements for smooth the progress of cloud federation used were Cloud Coordinators (CC) and Federation Coordinators (FC). The distinct feature of the projected federation is its QoS-orientation that can activate of-demand supply provisioning across multiple contributor, hence ration to maximize QoS targets and resources usage, reduce SLA violations and augment SLA formalization. However, intricacies services were not assembling using a mixture of services from different cloud providers and no condition was made for distributed denial of service (DDoS) attacks.

W. C. Chu et al. [15] proposed a prescribed model to assist not only the ECC services design and construction through SaaS, PaaS, IaaS but also the concurrent monitoring and self-motivated analysis on the QoS factors for the promises from QoS service providers and the service level agreements (SLA) for multiple ECC consumers. Foundation on the formal model, analysis model and testing model was produce to support automatic testing as well as runtime monitoring to pledge the contentment to thedesires/SLA constraints. This work had some previcinets such as not get used to the features and solutions of IOT into the framework as well as the field experiment.

M. M. Hassan et al. [16] premeditated and experienced the workload of big data by consecutively a group of typical big data jobs on Amazon cloud EC2. They formed a large simulation scenario and evaluate the proposed method with other approaches. Though, the
projected approach was cost effective, performance metrics such as delay, delay variable and throughput were not taken into thoughtfulness.

R. Karim et al. [17] proposed a method to map the users’ QoS requirements of cloud services to the right QoS specifications of SaaS then map them to the best IaaS service that propose the optimal QoS guarantees. The end-to-end QoS values were considered as a result of the mapping. They proposed a set of rules to carry out the mappings process. The QoS provisioning of cloud services was hierarchically modelled using the analytic hierarchy process (AHP) method. The AHP-based model lends a hand to make possible the mapping process thwart the cloud layers and to rank the candidate cloud services for the end users. A case study was used to exemplify and authenticate the solution approach. No presentation of evaluation was done based on real QoS data of cloud services.

S. Lee et al. [18] proposed an architecture that engaged the agent technology to hold the scrutinize of requested Quality of Service (QoS) requirements and service level agreements, to support verification and validation. Moreover, the agent technology enthusiastically analysed resources allocation and deployment. This work’s weedy point was being deficient in of self-learning algorithm to conclude the instant of involuntary allocation of system resources.

L. Bin et al. [19] et al. proposed a new QoS aware service imitation delete approach for disks space and upholding cost saving purpose. Investigational results specific that the DRDS algorithm can set aside disk space and maintenance costs for distributed storage systems while the availability and performance quality of service requirements are guaranteed. However, increased overhead on update and irregularity of data is usually connected with data replication.

P. Zhang et al. [20] obtainable a QoS framework for mobile cloud computing and an adaptive QoS management process to administer QoS pledge in mobile cloud computing environment. Moreover, they presented a QoS management model foundation on fuzzy cognitive map (FCM). No good quality model with appropriate construction was generated.

Y. Xiao et al. [21] accessible a QoS-aware reputation-based QoS provisioning scheme, which can diminish the cost of computing resources, while rewarding the needed QoS metrics. They well thought-out the arithmetic probability of the response time as a matter-of fact metric rather than the typical mean response time. More so, QoS provisioning algorithm was not used to incorporate security and privacy metrics.

M. Xu et al. [22] commence a numerous QoS inhibited scheduling strategy of multi-workflows (MQMW) to tackle the issue of multiple workflows with dissimilar QoS requirements. The projected strategy could schedule several workflows which were taking place at any time though QoS restriction such as availability and reliability were not added to workflows.

**Table 1: Comparison Table for QoS Techniques**

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Technique Used</th>
<th>Simulator Used</th>
<th>Algorithm Used/Parameter</th>
<th>Result Obtained</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1]</td>
<td>Simulation</td>
<td>Cloud analytic simulator</td>
<td>Equally Spread Current Execution (ESCE)</td>
<td>Better result in overall response time and data center processing time.</td>
</tr>
<tr>
<td>[2]</td>
<td>Simulation</td>
<td>CloudSim</td>
<td>Cluster Based Load Balancing Algorithm</td>
<td>Clustering can also be implemented at the client side. We can divide our tasks/cloudlets into different clusters depending upon their task length, cost and priority.</td>
</tr>
<tr>
<td>[3]</td>
<td>Simulation</td>
<td>CloudAnalyst</td>
<td>Round-Robin, Throttled, ESCE</td>
<td>The Throttled algorithm used in load balancer in cloud computing simulations using CloudAnalyst is the average response rate is still within the average range between UB1 and the other.</td>
</tr>
<tr>
<td>[4]</td>
<td>Simulation</td>
<td>CloudSim</td>
<td>Works on CSA gateway.</td>
<td>Results prove that the CSA(Cloud Service Agent) based approach fared better in terms of throughput and service time as compared against the already existing approaches.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Simulation</td>
<td>CloudSim</td>
<td>Proposed an algorithm named Reverse filling LB Algorithm and Compare with Round Robin Algorithm.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Proposed an algorithm that reduces the waiting time of the machines along with the overheads in the system resulting in lesser cost than existing approach. The future work includes modifications in the system leading to greater resource utilization and performance.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>None of the methods has considered the execution time of any undertaking at the runtime. In this manner there is an objective to grow such load balancing method that can enhance the execution of cloud computing alongside most extreme asset usage.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Analysis</td>
<td>Proposed an improved ACO</td>
<td>Ant colony optimization(ACO)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The advantage of the approach lies in the fact that the task of each ant is specialized rather than being general and the task depends on the type of first node that was encountered whether it was overloaded or underloaded.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Simulation</td>
<td>CloudSim</td>
<td>Novel dynamic load balancing algorithm</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Execution analysis of the simulation shows that change of MIPS will effects the response time. Increase in MIPS vs. VM decreases the response time.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Experiment has been conducted by varying the number of nodes in a VM available in a cloud configuration of single data center. The results show the use of other soft computing techniques is needed to be studied for further improvement.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Analysis</td>
<td>Proposed a Load Balancing Algorithm</td>
<td>Round Robin Algorithm</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>By developing new load balancing algorithms and model, the dynamic situations on servers can be easily handled and overloading problem can be avoided.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Simulation</td>
<td>Cloud Analyst</td>
<td>Throttled Algorithm, Round Robin Algorithm, Active Monitoring Load Balancing Algorithm</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Throttled Algorithm proves to be better compared with other two algorithms and gives better response time.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Analysis</td>
<td>Adopting CLBVM(Central Load Balancing Policy For Virtual Machines).</td>
<td>Central Load Balancing Policy</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>By Adopting the CLBVM policy the throughput of the system will increase.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Simulation</td>
<td>Cloud Analyst, CloudSim</td>
<td>Round Robin and Equally Spread Current Execution Load.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The response will become very less with MY Load Balancer .The parameters considers are overall response time and Data Centre Processing time.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Simulation &amp; Analysis</td>
<td>Cloud Analyst Simulation toolkit</td>
<td>Round-Robin and Throttled Algorithm</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>With the modified Throttled Algorithm, if the Number of VM increases it will reduces the response time and processing time of cloud data centers.</td>
<td></td>
</tr>
</tbody>
</table>
3. CONCLUSION
In this review, various Quality of Service methods in cloud computing environment are analysed, we found that till now much research work has been done to improve the QoS in cloud computing architecture. In this paper we studied the techniques which are used till now to enhance the quality and some pros and cons of these methods. Further work on to reduce the limitations and considering those areas where system is requiring improving the QoS (Quality of Service). In Future, more work can be done by considering those parameters and design a new efficient algorithm that helps to achieve better QoS in Cloud Computing environment.

REFERENCES: