

Service and Service Assurance Issues in a Cloud Environment

Prachi Deshpande*, S.C.Sharma*, S.K.Peddoju# and Ajith Abraham**

*Department of Applied Science & Engineering,

#Department of Computer Science & Engineering

Indian Institute of Technology Roorkee-India-247667

**Machine Intelligence Research Labs (MIR Labs), WA, USA.

**IT4Innovations - Center of Excellence, VSB -Tech. University of Ostrava, Czech Republic

{psd17dpt,scs60fpt,drpskfec}@iitr.ac.in,
ajith.abraham@ieee.org

Abstract. Cloud security and service assurance is a wide research area with an unrestrained amount of apprehensions, ensuring equipment and stage innovations, to secure information and asset access. In spite of the colossal advantages of Cloud computing paradigm, the security and service concerns have consistently been the center of various Cloud clients and obstruction to its extensive acceptance. The paper reports a meticulous review in the field of Cloud computing with a focus on the security risk assessment and service assurance. This effort will serve as a ready reckoner to the research aspirants to encompass a general thought of the risk factors in security and the service assurance in a Cloud environment.

Keywords: Cloud computing, CSP, Risk, Security, Service assurance.

1 Introduction

Advent of Cloud computing has given rise to a new paradigm in the information processing systems by providing location and resources transparent service to its users. It has dominant features like high degree of scalability, low maintenance, reduced hardware cost, expediency and persistent accessibility, reinforcement and recuperation, environment friendly, scalability and performance, rapid deployment and ease of incorporation, increased storage capability, device multiplicity and location independence etc. Pay-per-usage facility of the Cloud computing had attracted many public and private firms as well as individual users are attracted towards utilizing the various Cloud based services. The national institute of science and technology(NIST) had defined Cloud computing [1] as- 'Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service proadfa, p. 1, 2011.

vider interaction'. Cloud service providers(CSP) have established the obligatory proficiency and infrastructure to deliver the on demand service. Due to the increasing demand, assurance in Cloud service quality and performance is a latent challenge in front of CSPs. Security and allied risk is an important aspect in terms of providing a reasonable service to its clients. Though the services offered by the Cloud paradigm is attractive, the data security and allied risk are of great concern to the Cloud customers and inturn proving to be a big hurdle to its global prevalent adaptation. With growing use of Cloud services, the apprehension about its performance and overall quality has drawn attention of researchers and academicians. As Cloud is operating location transparent, there will be many users with different profile and needs. User's expectation towards services and applications deployed on the Cloud is to have same latency, reliability and availability as in case of traditional hardware configurations. The user of Cloud resources have to pay only for the used services by virtue of a customized service level agreement (SLA). Under such situation ensuring quality service is very tricky. The need and importance of the issue had initiated many efforts in the recent past to focus the problem solving of the service quality in Cloud infrastructure. Fig.1. depicts the basic Cloud structure in terms of user interface.

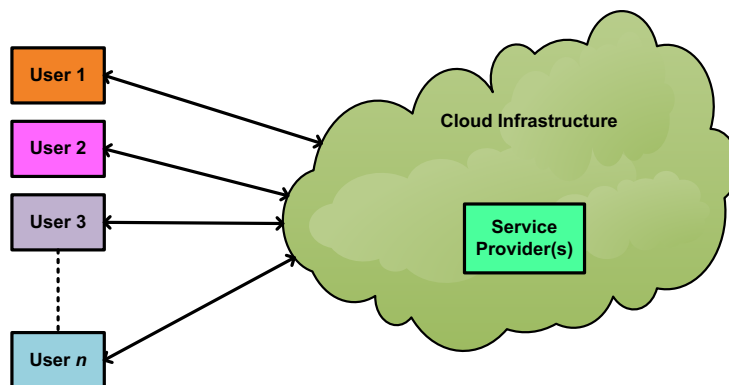


Fig. 1. Service scenario in Cloud computing

End-to- end quality of service, micro-benchmarks and kernels, a neural network based approach, effect virtualization of resources in Cloud scenario, a fuzzy synthetic decision based service performance and SLA as a metric [2-9] are considered as various approach to decide the quality of the service. But, these approaches had been directed towards addressing a solution to a particular application. No effort had been directed towards the fundamental issues in archiving the quality of service in the Cloud environment. Owing to these facts, it becomes necessary to identify and classify the various factors and ingredients on which the quality of service in Cloud computing scenario depends by and large. This paper focuses on the various aspects to have a reliable service with a focus on risk assessment in the field of Cloud computing. Section 2 describes the various security risk aspect associated with the Cloud operation while

section 3 describes the issues in achieving the service assurance in Cloud scenario. The article is concluded in section 4 with its future scope.

2 Risk Assessment in Cloud Scenario

Risk is the possibility of occurrence of an event that can unfavorably influence the attainment of the goals. The nature of risks (i.e. security, reliability, accessibility, and performance) are the same with the Cloud and the non-Cloud technology systems. But, the degree of risk and its profile varies if Cloud solutions are adopted depending on the impact of risk events (residual and natural) associated with the CSP. For the proposed study, we had classified the risk assessment in to two broad categories such as the risk arise due to with the inherent characteristics of the Cloud computing and with the deployment strategy of the Cloud. The various risk associated with the Cloud services and security can be minimized by risk reassignment, risk alleviation, risk approval and risk avoidance.

The risks caused by the CSP can be reassigned. The risks that arises due to use of various Cloud services, such as compliance, specification, and authentication can be alleviated. The risks which are indirect in nature and based on external factors, such as laws has to be accepted. When the risk arises due to the different specifications used by a CSP and the user it can be avoided.

2.1 Security Risk Assessment with Cloud Characteristics

The intent nature of the Cloud opens new paradigm in terms of security risk for a IT network. The emergence of Cloud computing can be seen as a risk episode for some organizations as in terms of becoming obsolete in the future. With acceptance of a third-party-managed Cloud services, the CSP and associated Cloud tenants may influence the organization in various ways. Lack of transparency, data leakage system failure due to over demand of resources, proprietary nature of the solutions provided by a CSP, possibility of security attacks due to the multitenant nature of a CSP's infrastructure are the scenarios which has to be considered as a risk factor in the Cloud. Finally, with adoption of Cloud computing services, an IT organization requires less skilled manpower to run the day to day activity. This may degrade the confidence and morale of the of remaining IT staff members. The risk assessment must be completed before an organization adopts the Cloud solutions. The assessment of a risk will be based on the criteria's such as risk profile, natural and residual risk and likelihood and impact. Risk profile indicates the ability of an organization to cover a specific volume of risk categories. An organization must evaluate the inherent risk and develop its responses and then settle on the residual risk. The ability to decide the possibility and impact of a risk depends on whether the organization has a comprehensive, accurate, and current inventory of risks. After the assessment of the risks in perspective of organizational objectives with Cloud computing, the response to the risk need to be determined. The possible risk can be mitigated by avoidance, reduction, sharing and acceptance of the risk. Table 1 summarizes the different risk and the possible strategy to overcome them.

2.2 Risk Associated with the deployment models of the Cloud

A Cloud can be deployed in different ways depending on the need and applications of the user. It becomes very important to access the risk level and allied issues at each of this models. No studies, till date, have concentrated on this issue. The deployment model is of important issue both for CSP and user point of view as the user must now the topology offered by the CSP to deal with his needs and vice versa. The major risk issues can be categories as Organizational or policy related issues, Technical issues, Legal Issues and miscellaneous issues. Table 2 summarizes the various security risk, its effect associated with the Cloud functioning and possible ways for their alleviation.

Table 1. Risk assessment and its response

Sr.	Classification of the risk	Alleviation policy
1.	Regulation non-compliance.	Monitoring of the external environment
2.	Disclosure noncompliance.	New disclosures in financial reporting
3.	Vendor lock-in.	Preparation of an exit strategy
4.	Reliability, performance, and cyber-attacks.	Incident management
5.	Transparency and relinquishing direct control.	Management oversight and operations monitoring controls
6.	Security, compliance, data leakage, and data jurisdiction.	Data classification policies and processes
7.	Unauthorized Cloud activity.	Cloud policies and controls
8.	Lack of transparency.	Assessments of the CSP control environment

Table 2. Classification of the Risk in Cloud Environment.

Classification	Nomenclature	Severity	Vulnerability	Effect on	Alleviation Policy
Policy related risk	Rk-1:Lock-In	High risk	Lack of standard technologies and solutions	Credibility of CSP, personal data of users, service quality	Risk Reassignment
	Rk-2: Failure of Governance	High risk	Undefined roles, responsibilities, and ownership	Customers trust, employee loyalty and CSP reputation	Risk Reassignment
	Rk-3: Compliance challenges	High risk	Unavailability of audit/certification to the customers	Service limits due to compliance issues	Risk Acceptance
	Rk-4: Loss of business due to cotenant activities	Medium risk	Poor resource isolation	Service delivery and personal data	Risk Avoidance.
	Rk-5: Service termination	Medium risk	Less transparency in terms of use	Service delivery	Risk Avoidance

Technical and security Risk	Rk-6 CSP Acquisition	Medium risk	Inter Cloud application dependency	Intellectual property, personal data	Risk Approval.
	Rk-7: Under or over provisioning of the resources	Medium risk	Inaccurate estimate of resource usage	Access control, authentication and authorization(AAA)	Risk Alleviation
	Rk-8: Segregation failure	High risk	Hypervisor vulnerabilities	QoS due to multi-tenancy	Risk Avoidance.
	Rk-9: Malevolent insider	High risk	Inadequate security procedures	Integrity and availability of all the data	Risk Avoidance.
	Rk-10: Availability of infrastructure	Medium risk	misconfiguration	Real time services	Risk Alleviation
	Rk-11: Snooping of the data	High risk	AAA failure	Security attacks such as sniffing, eavesdropping, man-in-middle, side channel and replay attacks can be dominant	Risk Avoidance
	Rk-12:Data seepage	High risk	AAA and communication encryption failure	transfer of data between CP and user	Risk Reassignment
	Rk-13: ineffective deletion of the data	Medium risk	susceptible media cleansing	Critical data may be lost	Risk Alleviation
	Rk-14:distributed denial of service (DDoS)	High risk	Misconfiguration	CSP management interface	Risk Avoidance.
	Rk-15: loss of encryption keys	High risk	Poor key generation mechanism	Credentials and personal data may be lost	Risk Reassignment
	Legal Risk	Rk-16:Malicious probing or scanning	Medium risk	Internal network probing can occur	User trust.
Rk-17: Compromise service engine		Medium risk	Lack of resource isolation	Service delivery.	Risk Avoidance
Rk-18: Conflict involving customer consolidation measures and Cloud environment		Medium risk	Conflicting SLA clauses and transparency in operation	The roles and responsibilities of CSP and the customers	Risk Reassignment
Rk-19:Subpoena and E-discovery		Medium risk	Lack of resource isolation and transparency in data storage	Users critical data , trust	Risk Approval
Rk-20:Change of Jurisdiction		High risk	information on jurisdictions	Users data held at multiple jurisdiction may put it in a high risk state	Risk Approval
Rk-21:Data protection		High risk	Lack of transparency in location of data storage	Company reputation may be at stake	Risk Avoidance

Miscellaneous Risk	Rk-22:License issues	Medium risk	Lack of transparency in terms of use	Certification and service delivery	Risk Approval
	Rk-23: Network Impairments	High risk	Misconfiguration, system or OS related issues, poor resource isolation	Potentially thousands of customers were affected at the same time	Risk Avoidance
	Rk-24: Network management	High risk	Network congestion, mis-connection and non optimum use	Service latency will be disturbed	Risk Reassignment
	Rk-25:Network traffic modification	Medium risk	No control over vulnerability assessment	Data retrieval	Risk Reassignment
	Rk-26: Privilege escalation	Medium risk	AAA mishap	Users personal data	Risk Avoidance
	Rk-27: Social engineering attacks	Low risk	Lack security awareness and resource isolation	CRP trust and users personal data	Risk Avoidance
	Rk-28: Backup related issues	Medium risk	Lost or stolen backup	Company reputation	Risk Reassignment
	Rk-29: Unauthorized access to the system	Medium risk	Inadequate physical security measures	Company reputation, users trust, sensitive data	Risk Avoidance
	Rk-30: Theft of PCs	Medium risk	Inadequate physical security measures	Company reputation, users trust, sensitive data	Risk Avoidance
	Rk-31: natural disaster	Medium risk		Data storage	Risk Avoidance

Fig.2 shows the classification of the various risk associated with the Cloud when analyzed in the deployment mode. A detailed analysis of risk associated with individual deployment model had been carried out from available literature and listed out in Table 3. in this analysis, a community Cloud is not taken into consideration as the risk level associated with it are more or less the same as that with a private Cloud.

Table 3. Summary of Risk consideration in each deployment model.

Classification	Organizational risk	Technical(Security) risk	Legal risk	Miscellaneous risk
Public Cloud	Rk-1,3,4,5	Rk-7,9,10,14	Rk-19,21	Rk-24,27,31

Private Cloud	Rk-1,3	Rk-8,9,12,13,16	Rk-20,21	Rk-23,25,28,29,30
Hybrid Cloud	Rk-2,3,5,6	Rk-8,9,11,15,16,17	Rk-20,21,22	Rk-23,26
Contribution	[10-12]	[10],[13-21]	[22-28]	[20],[29-31]

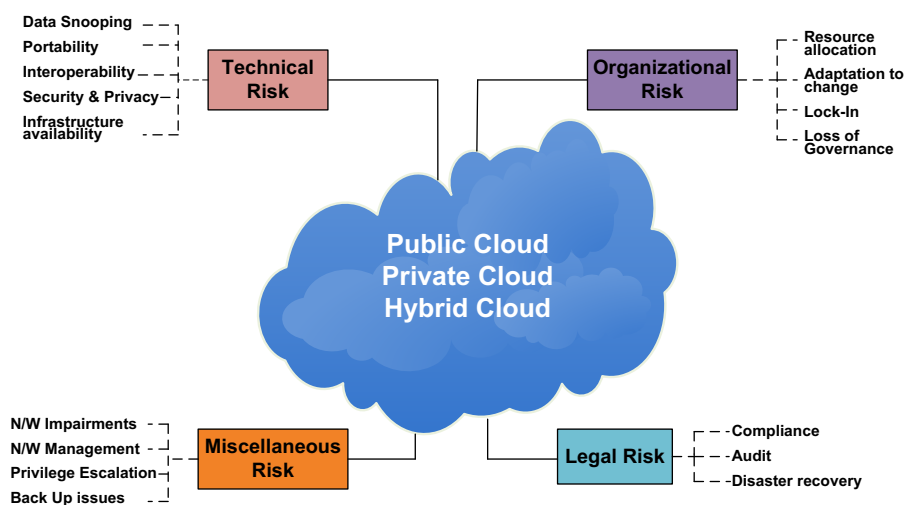


Fig.2 Cloud risk categorization

3. Service Quality Assurance in the Cloud Environment

The Cloud environment have a virtual machine (VM) pool between the user operating system (OS) and the physical resources, providing the required environment for the processing of user requirement. It acts as a bridge between the underlying network resources and the users by means of hypervisors. In general, a VM is a software abstraction of the complete virtualized resource environment to the user. The accuracy and quality of the services by could computing environment depends on the accuracy of the degree of emulation of resources provided by VMs to the guest OS. The Cloud computing environment provides services such as networking, computing and storage to the user's guest OS. The quality of a Cloud based service can be estimated by various cost functions such as availability, retainability, latency, throughput, and reliability. The general frame work for the deployment of services over the Cloud along with the various quality measures is described in Fig.3.

Availability of a service can be defined as the ability to perform its defined function as and when required by the user. Ideally, a service available ‘24x7’ without any constraints is known as the best available service. But practically, due to diverse nature of the Cloud, it may be possible that a particular service may not be available for some interval of time-*the downtime*. Hence, accuracy of a service can be estimated as-

$$Availability = \frac{ERT - DT}{ERT} \quad (1)$$

Where ERT=established response time and DT= down time in seconds.

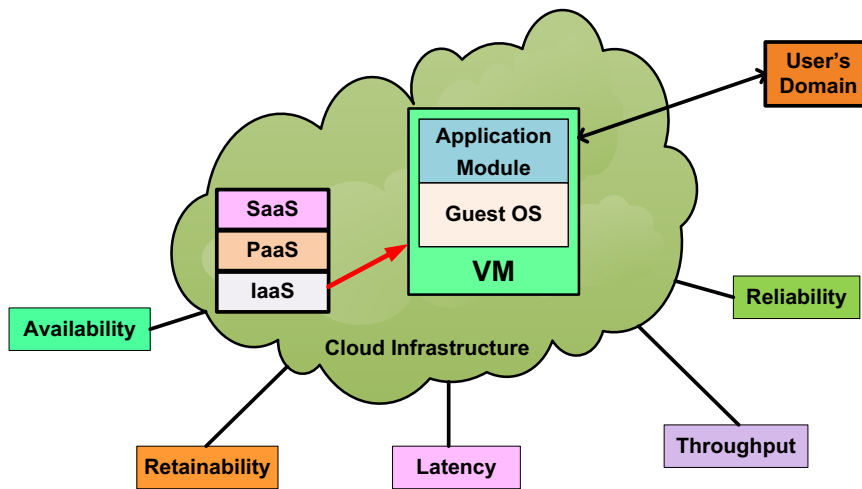


Fig. 3. Quality of Service factors for a Cloud based application

To guarantee the best availability of the service, maximum service interruption tolerance and the down time must be critically set. But for unknown application characteristics, the best way is to set quantitative requirements for key input ingredients.

It is believed that the next generation data communication will be replaced by videos. In such cases, *retainability* is of prime interest to have uninterrupted streaming. A service is said to be retainable if it ends itself due to the users request; not due to any other reasons.

Latency of the service is defined as the time lag between the request of a service and the actual commitment of the service. The service latency is influenced by many factors such as type of disk (storage), the discrepancy in request onset time, congestion in the underlying network, the available bandwidth, network impairments and packet loss, unusual traffic pattern.

The *throughput* of a service is the indicator of its successful operation over the time. Reliability and throughput are very closely related matrices. A service is said to be *reliable* if it is completed itself within the acceptable time frame. This feature should be intact for numerous request of the service by the user.

As the VMs are acting as a pool between users and the Cloud resources, its effective functioning is a critical aspect in the quality of service. Typical VM impairments includes malfunction of VM, mismatch in VM capacity configuration, drift in clock and allied jitters, degraded VM capacity, and failed VM instances. These are evidenced due to amalgamation of one or more operating conditions in the Cloud. The VM malfunction may be an outcome of unambiguous request by the Cloud user or the applications, failure of VM server, hypervisor or host OS and abrupt power failure. Sometimes, the Cloud service provider's infrastructure unable to provide a VM instance to any of the resources for a period of time. This phenomenon is known as mismatch in VM capacity configuration. It can be overcome by using hypervisors in tandem.

Degraded VM capacity arises due to the failure of Cloud infrastructure to allocate the predefined resources so as to function a VM normally. This factor becomes critical during the heavy workload conditions. In strict real time applications the quality of service is affected by the clock drift from its standard time and allied jitters. Seldom, it become difficult for the Cloud service provider to start up the guest OS and configure the VM promptly. There are many strategies to overcome all these drawbacks so as to have best quality of service such as proper work-load distribution, containing the frequent failures of VM, and focus on capacity management. One of the major bottlenecks in achieving a good quality of service is load balancing i.e. load distribution mechanism. It can be achieved by placing a proxy based system (client-server model) or an independent system in the service path.

Fig. 4 shows a generalized proxy based load distributor. In the proxy based load distribution (PLD) approach clients sees the IP address of the load distributor and the servers are obscure to them. The clients sends request to the proxy load distributor. Depending upon the availability it selects a server to work out the client's request. After processing, the selected server returns its response to the load distributor. This response is forwarded to the particular client. Apart from acting as a regulating element, a PLD element keeps the performance information for the selected server.

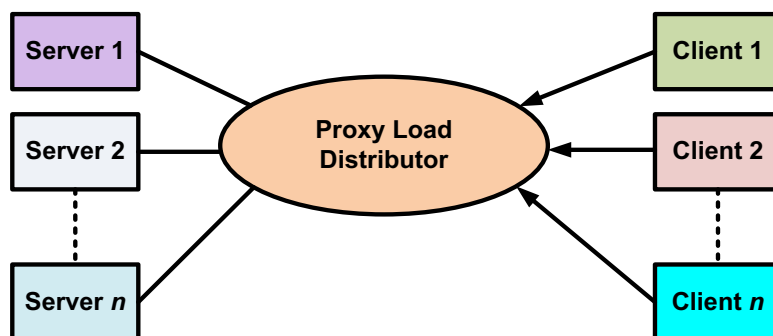


Fig. 4. A proxy load distributor

This information can be used for initiating a server selection decision for the next request. Thus the proxy based approach guarantees the best service for a given instance of request. It is very important to select a particular server for carrying out the

request processing by the clients. The selection of servers can be carried out by using a round robin, static configuration, random choice, status and performance based configuration [32-34]. Fig.5 shows the deployment scheme of the proposed PLD in Cloud. It will act as sub module of IaaS layer in the Cloud. It will be placed in between the host and the client VMs for regulating the traffic load. The host OS is nothing but a particular server available in the Cloud. Depending on the availability of a server, PLD assigns the VMs request to it.

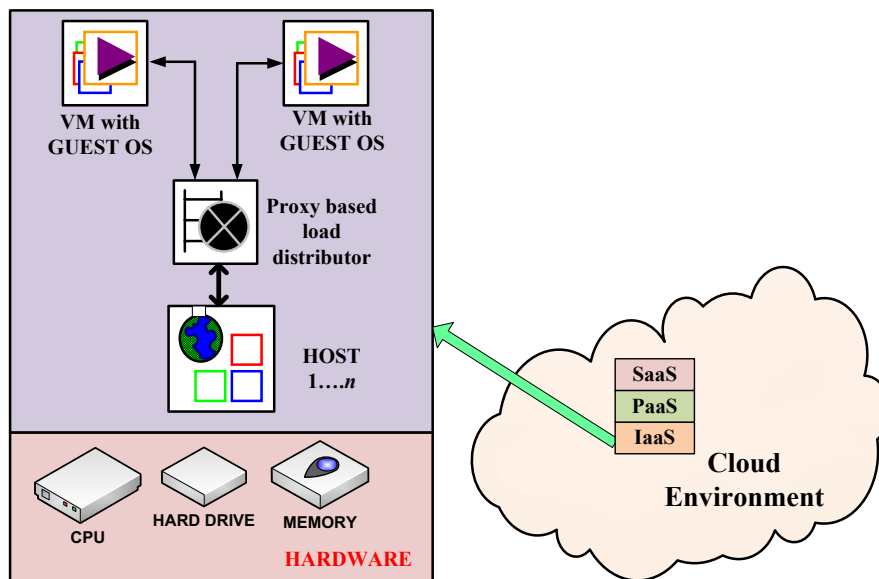


Fig.5. Deployment scheme for the proposed module in the Cloud

In case of *non proxy* load distribution, domain name service (DNS) is the most preferred strategy. In this approach, client sends a domain name to DNS server and in turn, receives an IP address allied with it. Clients transmits request for service using IP address provided by DNS. The process will be repeated till a meaningful response is returned by the DNS. The load distribution can also be carried out by a *static configuration*. In case of static configuration the client has to maintain a configuration file with it. An IP address of a server is clearly written in to it. Client has to send service request to the configured server and receives the response accordingly. The non proxy distribution can also be enforced with any cast or a multicast service. But the non proxy methods are less attractive due to its fixed server selection strategy. The Cloud environment is highly dynamic and demands very rapid responses for a service request. In this scenario a PLD is the best choice.

3.1 Load distribution challenges in Cloud infrastructure

In traditional network, the load distribution is simple as compared to the distributed environment like Cloud computing. This is the result of the fundamental properties of Cloud computing such as-

1. For speedy operation the Cloud is very flexible application pattern. Due to it the work load may get reduced or otherwise at any instance of time. Hence the load distributors must be dynamic enough to cope the situation.
2. Due to higher degree of virtualization, load distributor must distinguish the asymmetric server capability to handle different throughput instances.
3. Pay per use attracts a huge number of service requests from users in small bursts.

With the help of PLD, service availability is guaranteed as it ensures the availability of a server for each request. It also identifies and omits the failed server instances and forwards the load to the active server instances. But it may be possible that the load distributor may become a single point bottleneck in execution of services. Hence it also must be taken into account in decision for availability by the Cloud service providers.

The PLD holds the information of service conditions of the servers during the operation. This information will be useful in allocation of new workload to a particular server. This reduces the latency effect as the workload will be allocated to a server on its ready to process basis. The PLD monitors the resource usage continuously and hold this information for the next instance of service request. Based on this information, it diverts the new service request to the servers with best available resource allocation and thereby guarantees the reliability of service. The continuous monitoring feature of PLD will be helpful in service retainability also as it prevents the allocation of a new service request to an already overloaded server. In case of overload, the PLD moves the service instances to a server with more available resources. This feature will be helpful to achieve a service throughput.

One of the bottlenecks in achieving a high quality of service is the sudden failure of the service resources. Due to high degree of virtualization in Cloud environment, failure becomes a serious concern. Fortunately, due to PLD the effect of resource failure can be dealt effectively as the active node to carryout services (based on SLA) is pre identified and then service instance are initiated.

3.2 The capacity management

Managing the service capacity in the Cloud based infrastructure is another factor of prime interest for the system designer. The reason is, the variable workload situations in the Cloud. The variation in the workload can be due to the factors like increase or decrease in usage, random variation in the service traffic and due to the market conditions. Also the question of capacity management is directly related with the capital investment of the service providers. Traditionally the capacity management has been carried out by several key activities such as workload and allied performance monitoring, demand and resource forecasting, modeling and implementing the capacity

related changes. In case of Cloud computing, rapid elasticity and on-demand service affects the service capacity management by a large extent. It can be mitigated by resource allocation without any manual interference as and when need arises, so as to achieve good service of quality.

3.3 End-to-end service consideration

The Cloud based resource and service utilization by the end user is carried out by their personalized gadgets. The connection with the Cloud infrastructure is mostly in wireless manner (optionally wired manner also). Owing to this fact, user may come across the allied impairments while accessing the Cloud based services. Fig.6 shows the basic end-to-end service (ETES) scenario.

The ETES is primarily affected by the service conditions across the WLAN, the impairments with the VMs or with the Cloud service provider and due to human intervention. Also it is badly affected by the user devices associated at a particular time of instance. The condition becomes worst when one or more of these scenarios operate simultaneously. A careful design and control at individual stage can only mitigate the issues and improves the service quality.

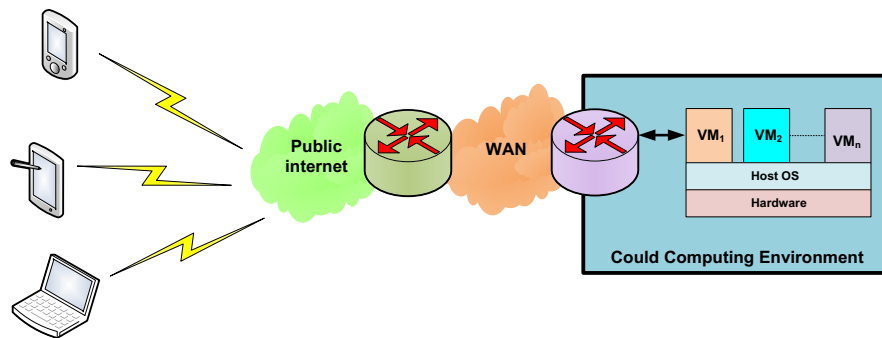


Fig. 6. End-to-end service scenario

3.4 Selection of a correct Cloud service model

As per the NIST, software as a service(SaaS),platform as a service(PaaS) and infrastructure as a service(IaaS) are the three service models associated with a Cloud. Each service model provides a specific level of abstraction to minimize the efforts of users to build and deploy systems. The first level up is IaaS. NIST defines IaaS as: "The capability provided to the consumer is to provision processing, storage, networks and other fundamental computing resources where the consumer is able to deploy and run arbitrary software which can include OS and applications". In short, IaaS provides infrastructure abstraction to the users. Table 4 list out the Cloud service providers.

Table 4. Cloud Service Providers

Sr.	Cloud Service	Vendors
1.	IaaS	Amazon web service, GoGrid, Rackspace
2.	PaaS	Google, Microsoft
3.	SaaS	Amazon EC2 Service

The next level up in the stack is PaaS. PaaS is on the top of IaaS and provides abstraction of many standard applications and provides these functions as a service. NIST defines PaaS as: "The capability provided to the consumer is to deploy onto the Cloud infrastructure consumer created or acquired applications created using programming languages, libraries and tools supported by the providers. The consumer doesn't manage or control the underlying Cloud infrastructure but control over the deployed applications and possibly configuration setting for the application-hosting environment." At the top of the stack is SaaS. It is a complete application delivered as a service to the consumers. The consumers have only to configure some application-specific parameters and manage users. There are many factors that are decisive to select a service model such as technical, financial, strategic, organizational and risk. The aspects like scalability, performance, security, disaster management are some of the aspects which are dealt under technical category while choosing a specific model for Cloud. The total cost of ownership is considered the most when the financial aspect is considered for choice of a specific model. Speed to market is very important strategically for the decision makers. The capability assessment of the organization to provide the required services plays a vital role in selecting a specific Cloud model. Finally the risk factor is decisive in the view point of companies ability to assume the risk in view of downtime, in damaging security breach etc. The risk is also in terms of companies decision to go with a public, private or hybrid Cloud. Generally, security and ownership of data are the major issues in selection of a specific Cloud service and deployment model.

3.4.1. When to use a SaaS

SaaS is the most grown-up of the three Cloud service model. The SaaS providers have total control over the infrastructure, performance, scalability, security of the applications. The SaaS provides offers two way connectivity to their potential users. The first and very common method is a web based connectivity. In such type of connectivity services are accessible via any device that can get connected with the internet. The second way is to have integration of features into the exiting application of the consumers via application programming interface(APIs).

SaaS will be used by a company to outsource all services and features assuming that it is an affordable venture. The company should not engage itself in building these applications. Buying and maintaining these applications are also not advisable it the emergence of SaaS. With SaaS solutions, companies need not to by the costly servers or software and allied man-power to manage them.

3.4.2. When to use PaaS

PaaS is the least mature model of the Cloud application. Many PaaS solutions require that the buyers must use a specific programming language and infrastructure. For small scale or beginner business setup, it may be acceptable but not for a big and complex enterprise. Due to lack of flexibility in terms of programming language and infrastructure, the acceptance of PaaS solutions are not acceptable widely. The second generation PaaS services had tried to overcome these drawbacks and supports multiple languages like Python, Ruby, PHP etc.

PaaS vendor provides a platform shared by many customers. To enforce the reliability of services, PaaS vendor provides many restrictions against heavy loading from an individual customer. These limits are called as throttling. Developers must understand and then design accordingly to the selected platform. Many PaaS service providers protect their consumers from the throttling activity. The developers must account for this in their design. It can be avoided by breaking the big task into smaller ones or continuously trying until got success. But for some applications throttling can create undue delays which in turn affects the quality of the service. In such cases, PaaS is not advisable service model.

3.4.3. When to use IaaS

If certain application requires issues like scalability or performance. to meet these requirements, developers need to manage memory, configurations of database and applications so as to maximize the throughput, specification of data distribution pattern, manipulation of OS, then one should leverage IaaS. When it is not required, then PaaS is the best. Another important aspect is the cost. PaaS is the best in terms of pay-per-use when data chunk is small. But, otherwise it is very expensive. Amazon EC2 is the lead IaaS service provider who has recently reduced the cost of its operations and other IaaS service providers are also following it. Another fact of preferring IaaS over PaaS is the possible risks related to extenuating threat of downtime. When PaaS provider has a shutdown, customer has to wait for the provider to fix the issues to get the service back on the track. But with IaaS, the customer can architect for failure and can build a redundant service over multiple virtual or physical data centers. In recent outage major websites like Reddit, Foursquare were down But many other websites were survived due to cross-zone redundancy.

As we move towards SaaS in the stack, the speed to market increases, reduction in human resource requirement and the operational cost. Unlike as we move towards IaaS, more control of the infrastructure and have a better chance of avoiding from a vendor outage. For startups and green field applications, it is common that entire applications are build in the Cloud. Switching the peaks in the traffic, reduction in the storage cost by shifting the storage in the Cloud, data analysis with Cloud infrastructure and on demand testing environment are the scenarios when the IT companies prefers the Cloud services.

3.5 Security as a service

In achieving a sizable quality of service, security of Cloud based application plays a pivotal role. Generally Cloud operations were maintained and managed by virtual technology. It may be possible that more than one user shares the system environment. In such scenario the security problem exists in terms of segregation and protection of individual customers data from the third party. Any IT infrastructure change or up gradation brings the opportunities as well as the risks with it and same has been true for the Cloud computing also. Due to the inherent nature of the Cloud computing, it has prone to many risk associated with its functioning that had not been experience before. For effective deployment of the Cloud services, these must be addressed in detail. For having agreed end results, the applications must run for what it has been intentioned for. Without proper security mechanism, intruders may get access of the information in the Cloud and in turn, can temper with it. This is where an intrusion detection system (IDS) fits in. In Cloud computing environment, IDS may be host based or network based entity. It is recommended to have a security as a service in IaaS layer of the Cloud. As VMs are responsible to bridge the users with the Cloud resources, IDS must be fit in each VM. Fig. 7 depicts the IDS deployment in the Cloud based environment.

In case of host based IDS, it will scan the system call traces of the OS to detect the abnormal behavior. In case of network based IDS, it has to detect a wide pool of attacks such as DDoS, port scanning, flooding etc. Out of these, port scan attack is very prominent, as it provides the exact information about the working environment and running application processes to the intruder [35].

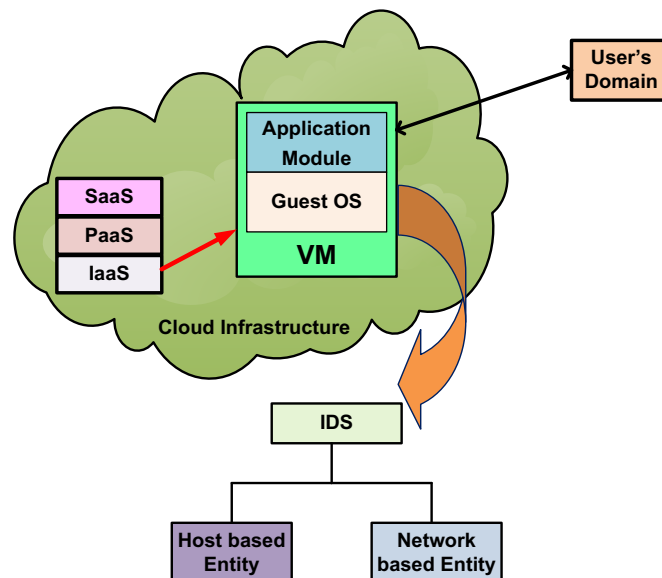
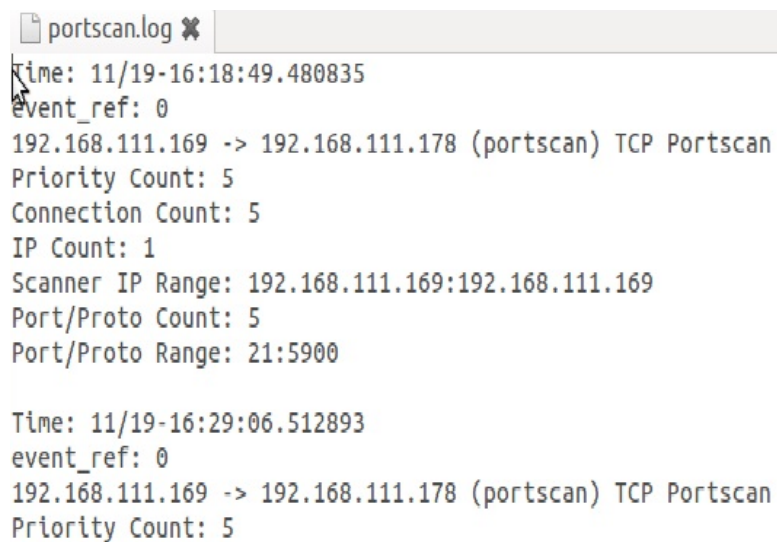


Fig. 7. IDS as a service in Cloud computing paradigm

The behavior of a private Cloud [36] had been evaluated under different security attacks by using a SNORT IDS. Portscan attack provides information of the open ports and the services running on that ports . Using this information attacker can exploit the vulnerabilities and launch the next attack. An attacker can hide its identity by using false IP addresses with TCP port scan. Also, It can even check if the firewall is active on the target system. Portscan attack using TCP port scan was launched on Cloud host and the detection of attack using snort IDS is shown in Fig.8. The Scanning for TCP open ports is performed on target {IP-178} and the same has been detected by the IDS.

Enforcement of security as a service using IDS with Cloud , the attack must be sensed before it gets activated and the source must be identified , port scan log can help in sensing the future attack strategies.

'Flooding' is another destructive security attack. It will force the system to generate false alarms regarding the availability of the service. An IDS will help to eliminate the danger with such attacks.



```
portscan.log ✕
Time: 11/19-16:18:49.480835
event_ref: 0
192.168.111.169 -> 192.168.111.178 (portscan) TCP Portscan
Priority Count: 5
Connection Count: 5
IP Count: 1
Scanner IP Range: 192.168.111.169:192.168.111.169
Port/Proto Count: 5
Port/Proto Range: 21:5900

Time: 11/19-16:29:06.512893
event_ref: 0
192.168.111.169 -> 192.168.111.178 (portscan) TCP Portscan
Priority Count: 5
```

Fig 8. Port scan attack detection with snort.

Fig 9. shows SYN flood attack detection with SNORT IDS. A SYN flood is a form of denial-of-service attack in which an attacker sends a large number of repetitive SYN requests to a target's system in an attempt to consume server resources and make server unavailable to legitimate users.

Security as a service feature will generate an alert as soon as it detects the abnormal behaviors in the execution of an application, data patterns and traffic load. The IDS for providing security as a service must fulfill the quality of service measures like deep packet inspection, system call monitoring, automatic or manual remediation action on intrusion detection, system or application log inspection and VM image repository monitoring. It is possible that flooding can be launched to block the IDS

itself by the attackers. In such cases, the a PLD can be very useful as it will immediately scene the abnormal in rush of service requests. It will redirect them to appropriate servers based on the information stored for previous request(s). This functioning will provide sufficient time for IDS to analyze the malicious activities in the Cloud.

Timestamp	Source IP	Destination IP	Alert Message
11/26-12:08:10.164925	116.134.236.1617471	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 116.134.236.1617471 -> 192.168.111.178:80
11/26-12:08:10.165044	139.169.71.1397481	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 139.169.71.1397481 -> 192.168.111.178:80
11/26-12:08:10.165169	55.82.113.1737491	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 55.82.113.1737491 -> 192.168.111.178:80
11/26-12:08:10.165292	130.65.42.1117501	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 130.65.42.1117501 -> 192.168.111.178:80
11/26-12:08:10.165349	93.130.246.2087511	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 93.130.246.2087511 -> 192.168.111.178:80
11/26-12:08:10.165511	189.114.19.2167521	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 189.114.19.2167521 -> 192.168.111.178:80
11/26-12:08:10.165636	210.204.172.2057531	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 210.204.172.2057531 -> 192.168.111.178:80
11/26-12:08:10.165769	252.233.7.807541	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 252.233.7.807541 -> 192.168.111.178:80
11/26-12:08:10.165877	172.74.119.1657551	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 172.74.119.1657551 -> 192.168.111.178:80
11/26-12:08:10.166003	42.22.240.1737561	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 42.22.240.1737561 -> 192.168.111.178:80
11/26-12:08:10.166110	222.148.238.827571	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 222.148.238.827571 -> 192.168.111.178:80
11/26-12:08:10.166233	80.77.147.587581	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 80.77.147.587581 -> 192.168.111.178:80
11/26-12:08:10.166349	78.174.39.1737591	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 78.174.39.1737591 -> 192.168.111.178:80
11/26-12:08:10.166451	186.205.7.207601	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 186.205.7.207601 -> 192.168.111.178:80
11/26-12:08:10.166583	130.86.224.417611	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 130.86.224.417611 -> 192.168.111.178:80
11/26-12:08:10.166682	109.121.196.1427621	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 109.121.196.1427621 -> 192.168.111.178:80
11/26-12:08:10.166813	1.203.0.867631	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 1.203.0.867631 -> 192.168.111.178:80
11/26-12:08:10.166938	2.243.114.2067641	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 2.243.114.2067641 -> 192.168.111.178:80
11/26-12:08:10.167058	158.191.180.527651	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 158.191.180.527651 -> 192.168.111.178:80
11/26-12:08:10.168312	115.22.14.1377756	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 115.22.14.1377756 -> 192.168.111.178:80
11/26-12:08:10.169113	244.61.204.2467823	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 244.61.204.2467823 -> 192.168.111.178:80
11/26-12:08:10.169751	142.60.94.1627878	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 142.60.94.1627878 -> 192.168.111.178:80
11/26-12:08:10.169871	137.68.169.17888	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 137.68.169.17888 -> 192.168.111.178:80
11/26-12:08:10.170451	178.97.173.1367939	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 178.97.173.1367939 -> 192.168.111.178:80
11/26-12:08:10.170870	169.221.96.697974	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 169.221.96.697974 -> 192.168.111.178:80
11/26-12:08:10.171467	163.95.244.1198026	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 163.95.244.1198026 -> 192.168.111.178:80
11/26-12:08:10.172280	22.114.166.578064	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 22.114.166.578064 -> 192.168.111.178:80
11/26-12:08:10.172729	165.64.48.1428095	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 165.64.48.1428095 -> 192.168.111.178:80
11/26-12:08:10.172927	65.55.152.1268149	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 65.55.152.1268149 -> 192.168.111.178:80
11/26-12:08:10.173337	204.114.92.1528184	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 204.114.92.1528184 -> 192.168.111.178:80
11/26-12:08:10.173679	130.116.57.1198213	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 130.116.57.1198213 -> 192.168.111.178:80
11/26-12:08:10.175159	152.11.244.1918341	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 152.11.244.1918341 -> 192.168.111.178:80
11/26-12:08:10.175594	36.178.74.318376	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 36.178.74.318376 -> 192.168.111.178:80
11/26-12:08:10.176112	123.16.183.2128419	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 123.16.183.2128419 -> 192.168.111.178:80
11/26-12:08:10.176357	158.116.243.803840	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 158.116.243.803840 -> 192.168.111.178:80
11/26-12:08:10.177086	244.133.39.618502	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 244.133.39.618502 -> 192.168.111.178:80
11/26-12:08:10.177391	87.196.216.838528	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 87.196.216.838528 -> 192.168.111.178:80
11/26-12:08:10.177978	240.140.174.2108580	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 240.140.174.2108580 -> 192.168.111.178:80
11/26-12:08:10.178313	244.240.83.1898018	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 244.240.83.1898018 -> 192.168.111.178:80
11/26-12:08:10.178904	48.242.111.588658	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 48.242.111.588658 -> 192.168.111.178:80
11/26-12:08:10.179187	76.136.208.1198083	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 76.136.208.1198083 -> 192.168.111.178:80
11/26-12:08:10.179390	67.165.241.2118718	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 67.165.241.2118718 -> 192.168.111.178:80
11/26-12:08:10.179682	115.185.212.958741	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 115.185.212.958741 -> 192.168.111.178:80
11/26-12:08:10.180202	216.89.286.578768	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 216.89.286.578768 -> 192.168.111.178:80
11/26-12:08:10.180771	205.284.94.1688817	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 205.284.94.1688817 -> 192.168.111.178:80
11/26-12:08:10.181104	96.89.92.2128846	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 96.89.92.2128846 -> 192.168.111.178:80
11/26-12:08:10.181484	116.238.229.1308878	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 116.238.229.1308878 -> 192.168.111.178:80
11/26-12:08:10.181923	67.55.205.2468915	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 67.55.205.2468915 -> 192.168.111.178:80
11/26-12:08:10.182383	142.14.221.1198955	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 142.14.221.1198955 -> 192.168.111.178:80
11/26-12:08:10.182686	204.137.119.1568981	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 204.137.119.1568981 -> 192.168.111.178:80
11/26-12:08:10.183330	227.168.116.1919036	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 227.168.116.1919036 -> 192.168.111.178:80
11/26-12:08:10.183688	280.280.74.579060	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 280.280.74.579060 -> 192.168.111.178:80
11/26-12:08:10.184071	114.202.135.149097	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 114.202.135.149097 -> 192.168.111.178:80
11/26-12:08:10.184685	127.61.14.2199149	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 127.61.14.2199149 -> 192.168.111.178:80
11/26-12:08:10.184822	48.126.178.1339161	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 48.126.178.1339161 -> 192.168.111.178:80
11/26-12:08:10.185259	149.92.16.6089186	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 149.92.16.6089186 -> 192.168.111.178:80
11/26-12:08:10.185797	58.48.48.2049243	192.168.111.178:80	[Classification ID: 0] [Priority ID: 3] [TCP] 58.48.48.2049243 -> 192.168.111.178:80

Fig 9. SYN flood attack detection with snort

4 Conclusions

The security risk and quality of service issue had rapidly acquired a hotspot in Cloud computing research due to the distributed nature of the Cloud. A detailed classification and systematic review has been presented on the issue of various security risk associated with the Cloud service. A possible mitigation approach had been also listed out based on the severity and nature of the risk issue. Also, the fundamental aspects of the service quality in Cloud computing environment has been discussed in the paper. Proper distribution of work load is the key factor to achieve a good quality of service in the Cloud environment. A proxy based load distribution strategy has been discussed in detail for Cloud environment. A focus has been given on the capacity management and security as a service for Cloud applications so as to achieve the hassle free service for all the users. In future, detailed mathematical modeling for the various security risk issues and service quality is to be initiated. This will help to arrive at a

generalized way to alleviate the security risk issues which in turn will enhance the service quality in the Cloud scenario.

References

1. Brown, E.: NIST issues Cloud computing guidelines for managing security and privacy. National Institute of Standards and Technology Special Publication 800-144 (2012).
2. Oberle, K., Cherubini, D., Cucinotta, T.: End-to-End service quality for Cloud applications. *Economics of Grids, Clouds, Systems, and Services*, LNCS, 8193, 228-243 (2013).
3. Ostermann, S., Iosup, A., Yigitbasi, N., Prodan, R., Fahringer, T., and Eperna, D.: A performance analysis of EC2 Cloud computing services for scientific computing. *LNICST*.34, 115-131 (2010).
4. Abraham, A., Thomas, J., and Ghinea, G.: Mining network quality of service for human computer interaction using neural networks. In: 10th International Conference on Human - Computer Interaction.3, 1193-1197 (2003).
5. Youseff, L., Seymour, K., You, H., Dongarra, J., Wolski, R.: The impact of paravirtualized memory hierarchy on linear algebra computational kernels and software. In: Proceedings of the 17th international symposium on high performance distributed computing, ACM, New York. 141-152 (2008).
6. Shangguang, W., Zhipiao, L., Qibo, S., Hua, Z., Fangchun, Y.: Towards an accurate evaluation of quality of Cloud service in service-oriented Cloud computing. *Jr. Inte.Manu.*25(2), 283-291 (2014).
7. Wu, L., Garg, S., Buyya, R.: SLA-based resource allocation for software as a service provider (SaaS) in Cloud computing environments. In: 11th IEEE/ACM Int. Sympo. on Cluster, Cloud and Grid Comp. 195-204 (2011).
8. Goudarzi, H., Ghasemazar, M., Pedram, M.: SLA-based optimization of power and migration cost in Cloud computing. In: 12th IEEE/ACM Int. Sympo. on Cluster, Cloud and Grid Comp. 172-179 (2012).
9. Chhetri, M., Vo, Q., Kowalczyk, R.: Policy-based automation of SLA establishment for Cloud computing services. In: 12th IEEE/ACM Int. Sympo. on Cluster, Cloud and Grid Comp. 164-171 (2012).
10. Chen, D., Zhao, H.: Data Security and Privacy Protection Issues in Cloud Computing. In: *Int. Conf. Comp. Sci. Elect. Eng.* 647-651 (2012).
11. Chou, Y., Oetting, J.: Risk Assessment for Cloud-Based IT Systems. *Int. Jr. of Grid and High Perf. Comput.* 1-13 (2011).
12. Bisong, A., Rahman, S.M.: An Overview of the Security Concerns in Enterprise Cloud Computing. *Int. Jr. Net. Secu. & its Appli.*3(1), 30-45 (2011).
13. Harauz, J., Kauifman, M., Potter, B.: Data Security in the world of Cloud computing. *IEEE Security & Privacy.* 7(4), 61-64 (2009).
14. Subashini, S., Kavitha, V.: A survey on security issues in service delivery models of Cloud computing. *Jr. Net. and Comp. Appli.* 34(1), 1-11 (2011).
15. Takabi, H., Joshi, J., Ahn, G.: Security and Privacy Challenges in Cloud Computing Environments. *IEEE Security and Privacy* 8(6), 24-31 (2010).
16. Hashizume, K., Rosado, D., Medina, E., Fernandez, E.: An analysis of security issues for Cloud computing. *Jr. of Inte. Services and Appli.* 4(5), 1-13 (2013).
17. Ren, K., Wang, C., Wang, Q.: Security Challenges for the Public Cloud. *IEEE Jr. of Internet Comp.*16(1), 69-73 (2012).

18. Ayala, L., Vega, M., Vargas, L.: Emerging Threats, Risk and Attacks in Distributed Systems: Cloud Computing. In: Elleithy, K., Sobh, T. (eds.) *Innovations and Advances in Computer, Information, Systems Sciences, and Engineering*. LNEE,152, 37–52 (2013).
19. Khajeh, A., Sommerville, I., Bogaerts, J., Teregowda, P.: Decision Support Tools for Cloud Migration in the Enterprise. In: *IEEE Int. Conf. Cloud Comp.* 541-548 (2011).
20. Chou, Y., Oetting, J.: Risk Assessment for Cloud-Based IT Systems. *Int. Jr. Grid and High Perf. Comput.* 3(2), 1–13 (2011).
21. Jansen, W., Grance, T.: *Guidelines on Security and Privacy in Cloud Computing*. NIST (2011).
22. Fito, J., Guitart, J.: *Introducing Risk Management into Cloud Computing*. Barcelona Supercomputing Center and Technical University of Catalonia, Barcelona, Spain (2010).
23. Barron, C., Yu, H., Zhan, J.: *Cloud Computing Security Case Studies and Research*. In: *Proc. World Cong. on Engineering*, 2, 1-5 (2013).
24. Julisch, K., Hall, M.: Security and Control in the Cloud. *Info. Security Jr.: A Global Perspective*. 19(6), 299–309 (2010)
25. Dahbur, K., Mohammad, B.: A Survey of Risks, Threats and Vulnerabilities in Cloud Computing. In: *Int Conf. Inte. Semantic Web-Services and Appli.* 1-6 (2011).
26. Cloud Security Alliance CSA: *The Notorious Nine Cloud Computing Threats 2013* (2013)
27. Peiyu, L., Dong, L.: Risk Assessment Model for Information System in Cloud Computing Environment. *Advanced in Control Engineering and Information Science*. 15, 3200-3204 (2011).
28. Iyengar, S., Ganapathy, G., Kumar, M., Abraham, A.: A Multilevel Thrust Filtration Defending Mechanism against DDoS Attacks in Cloud Computing Environment. *Int. Jr. Grid and Utility Comp.* 5(4), 236-248 (2014).
29. Kaufman, L.: Data Security in the World of Cloud Computing. *IEEE Security and Privacy*. 7(4), 61-64 (2009).
30. Che, J., Duan, Y., Zhang, T.: Study on the Security Models and strategies of Cloud Computing. In: *Proc: Int Conf. Power Elect. and Eng. Appli.* 23, 586-593 (2011).
31. Rosado, D., Gomez, R., Mellado, D., Medina, E.: Security Analysis in the Migration to Cloud Environment. *Jr. Future Internet*, 4(2), 469–487 (2012).
32. Kalyanaraman, R.: A rule based static configuration validation technique in an autonomous distributed environment. In: *2nd Int. Conf. on Systems*. 53 (2007).
33. Pengye, X., Gary, S.: Distributed joint optimization of traffic engineering and server selection. In: *Proceedings of 18th International Packet Video Workshop*. 86-93 (2010).
34. Tran, H., Mellouk, A., Perez, J., Hoceni, S., Zeadally, S.: QoE-based server selection for content distribution networks. *IEEE Trans. Comp.* PP(99), 1 (2013).
35. Deshpande, P., Sharma, S., Peddoju, S., Abraham, A.: Distributed port scan attack in Cloud environment. In: *5th Int. Conf. Compu. Aspects of Social Net.* 27-31 (2013).
36. Deshpande, P., Sharma, S., Peddoju, S.: Implementation of a private Cloud: A case study. *Adv. Inte. Syst. and Comp.* 259, 635-647 (2014).