

## Cloud Security: Problems and Solutions



### Engineering

**KEYWORDS :** Cloud computing; cloud security; Internet security, distributed computing

**Dr. Dipti D. Patil**

Associate Professor, Computer Engineering Dept., MIT COE, Pune

**Dr. V. M. Wadhai**

Principal, Computer Engineering Dept, MIT COE, Pune

**Bhagyashri D. Dangewar**

Research Scholar, MIT COE, Pune

### ABSTRACT

*Cloud computing is a network-based environment that focuses on sharing computations or resources. Cloud computing is new computing model which is based on grid computing distributed computing, parallel computing, virtualization technology, utility computing and other computer technologies. In cloud computing data is stored remotely from the customer's location; in fact, it can be stored at any location and due to which security, in particular, is one of the most argued issues in the cloud computing field. Several enterprises look at cloud computing warily due to projected security risks. Single security method cannot solve the cloud computing security problem and many traditional and new technologies and strategies must be used together for protecting the total cloud computing system. This paper describes the cloud computing environment, discusses the security issues arise while handling big data on cloud, and the level of security to be provided at various levels in cloud architecture as solution.*

### I. INTRODUCTION

Cloud computing is one of today's most exciting technologies due to its ability to reduce costs associated with computing while increasing flexibility and scalability for computer processes. During the past few years, cloud computing has grown from being a promising business idea to one of the fastest growing parts of the IT industry. Cloud computing is a network-based environment that focuses on sharing computations or resources. Actually, clouds are internet-based and it tries to disguise complexity for clients. Cloud computing refers to both the applications delivered as services over the internet and the hardware and software in the data centers that provide those services. Cloud providers use virtualization technologies combined with self service abilities for computing resources via network infrastructure. In such environments, several kinds of virtual machines are hosted on the same physical server as infrastructure. Consumers need to pay only for what they use and have not to pay for local resources which they use such as storage or infrastructure. The cloud computing system provides the service for the user and has the character of high scalability and reliability. The resource in the cloud system is transparent for the application and the user does not know the place of the resource. The users can access your applications and data from anywhere.

Resources in cloud systems can be shared among a large number of users. The cloud system could improve its capacity through adding more hardware to deal with the increased load effectively when the work load is growing. Cloud resources are provided as a service on an as needed basis. The cloud itself typically includes large numbers of commodity-grade servers, harnessed to deliver highly scalable and reliable on-demand services. The amount of resources provided in the cloud system for the users is increased when they need more and decrease when they need less. The resource can be the computing, storage or any other specification service. The cloud computing system provides the service for the user and has the character of high scalability and reliability. The resource in the cloud system is transparent for the application and the user does not know the place of the resource. The users can access its applications and data from anywhere. Resources in cloud systems can be shared among a large number of users. The cloud system could improve its capacity through adding more hardware to deal with the increased load effectively when the work load is growing. Cloud resources are provided as a service on an as needed basis. The cloud itself typically includes large numbers of commodity-grade servers, harnessed to deliver highly scalable and reliable on-demand services. The amount of resources provided in the cloud system for the users can be increased when they

need more and decreased when they need less. The resource can be the computing, storage and other specification service. Many companies provide the cloud computing platform such as Google, IBM, Microsoft, Amazon, VMware and EMC [1-7].

### II. CLOUD COMPUTING

In cloud computing, data storage and computing is not carried out on the local computer or server but on the computers distributed over the internet. In this environment the tasks which are implemented on the personal computer and private data center are moved on to the larger computing center which are shared with total user and distributed over the internet. Its applications out of loosely coupled services and one service failure will not disrupt other services. The cloud computing system can be divided into two sections: the front end and the back end. They connect to each other through the internet. The front end is user who use the service provided by the back end which is the cloud section of the system. The computer hardware, software, computing resources and the services which include resource management are shared fully. The services in cloud computing are ubiquitous and they can be accessed from workstations and other devices, such as cell phones. The virtualization has the ability to run multiple operating systems on a single physical system and share the underlying hardware resources. A virtual server can be serviced by one or more hosts, and one host may house more than one virtual server. If the environment is built correctly, virtual servers will not be affected by the loss of a host. Hosts may be removed and introduced almost at will to accommodate maintenance. The virtual servers in the cloud computing system can be scaled out easily and if the administrators check out that the resources supporting a virtual server are being taxed too much in the real environment and they can modify the amount of resources allocated to that virtual server. Cloud computing is developed from many technologies such as parallel computing, distributed computing, grid computing and other computer technologies. The grid computing want to solve the assignment of computing and resource storage and the cloud computing want to share the computing, storage and application resource. The grid computing does not rely on virtualization as much as the cloud computing do and each individual organization maintain full control of their resources. The user need not computing and storage resource and don't provide the application in the cloud computing. The resource and server can be provided by the cloud computing. There are three types of cloud environments: Public, Private, and Hybrid clouds. A public cloud is standard model which providers make several resources, such as applications and storage, available to the public. Public cloud services may be free or paid. In public clouds applications may run externally by large service providers and

offers some benefits over private clouds. Private Cloud refers to internal services of a business that is not available for ordinary people. Essentially Private clouds are a marketing term for an architecture that provides hosted services to particular group of people behind a firewall. Hybrid cloud is an environment that a company provides and controls some resources internally and has some other resources for public use. Combination of private and public clouds is also called Hybrid cloud. In this type, cloud provider has a service that has private cloud part which is only accessible by certified staff and protected by firewalls from outside access and a public cloud environment to which external users can have access to it. The cloud computing is on-demand service and it gives computing capabilities as needed automatically. It can use the service by many machines such as desktop, laptop, PDA and mobile phone. There are three major types of services in the cloud environment: SaaS, PaaS, and IaaS [1]. In the software as a service the consumer uses the provided application and don't manage or control the network, server's storage and the application. It can reduce expenses and is easy to use and access everywhere. It share instance of a software application as a service accessible via internet browser or client based role access and sharing rules. The service provider hosts the software so the user doesn't need to install or manage or buy hardware for it. All they have to do is connect and use it. The examples of SaaS are Flickr, Google Docs, Siri, Amazon and Cloud Drive. In platform as a service the consumer deploys their applications on the cloud computing system and controls their applications but they don't manage servers and storage and delivers a computing platform or solution stack as a service. It shares platform for custom software application configuration, development, testing and deployment. The applications are deployed without the cost and complexity of buying and managing the underlying hardware and software layers and it makes raw hardware made available to the user through the internet but generally includes a specific operating system that is pre-installed and supported by the cloud vendor. The examples of PaaS are Google App Engine, Amazon Web services.

In the infrastructure as a service the consumer get access to the infrastructure to deploy their application and system but they don't manage or control the infrastructure and they control the storage and applications. It shares managed pool of configurable and scalable resources such as network, middleware, database and storage servers. The examples of IaaS are Amazon Elastic Compute Cloud (EC2). The cloud has the elastic character and resource allocation can get bigger or smaller depending on demand. The cloud also has the scalability and the cloud can scale upward for peak demand and downward for lighter demand. The application can get modified while adding users or when application requirements change.

There are many cloud computing systems in the market such as Google, Windows, IBM and Amazon. The Google cloud computing system include GFS (Google File System), MapReduce and Bigtable. The GFS is a distributed file system and which contains one master server and many block servers. The file is segmented into stationary size such as 64Mb file block stored in the block server. The MapReduce is a distributed programming mode and it can decrease the complex of programming in the cloud computing. The MapReduce include Map and Reduce operation and the Map use the Key and Value to create new Key and Value. The Reduce operation merges the same style of Key and Value. The MapReduce is not only programming mode but also efficient parallel task scheduling model. The programmer can provide their own Map function and Reduce function to process data. The Bigtable is a distributed and large scale database management system and the data is stored in the table which is divided into many rows. Many rows make a small tablet stored in the node. The Bigtable depends on the distributed cluster task scheduling, GFS and distributed locker service Chubby. Windows provide the Azure operation system which want to create cloud computing platform for the developer. The developer can make the application on the cloud server, web, PC and data center. The system supports general-purpose computing, rather than a single category of application. Amazon provide the EC2 (Elastic Compute Cloud) and S3 (Simple Storage Ser-

vice). The EC2 can provide many services which running in the virtual machine. The user can choose different virtual machine according to the different requirement and upload it to the S3 and call the machine interface to finish the task. The Hadoop is an open source distributed computing framework provided by the Apache. Many network stations use it to create system such as Amazon, Facebook. The Hadoop cores are MapReduce and HDFS (Hadoop Distributed File System). The MapReduce can make the decomposition of tasks and integration of results. The HDFS is a distributed file system and provide the base support for the storage of file in the storage node. The MapReduce contains job trackers and task trackers. MapReduce [4] is a programming model Google has used successfully in processing big data sets. A map function extracts some intelligence from raw data and a reduce function aggregates the data according to some guidelines output by the map. MapReduce needs a distributed file system and an engine that can distribute, coordinate, monitor and gather the results. The HDFS is a master and slave framework and which contains data nodes and name node. The name node is a center server and manage the name space in the file system. The data node manages the data stored in it. The cloud computing is a large scales distributed computing mode and scale economic driven mode. The large scale is the first character and it can provide more low cost service for the user. The cloud computing uses the abstract entities on the every layer function and provides cloud server for the user. The cloud server is implemented by the virtualization technology. The user can use the cloud server for the cloud system in every place and using of every terminals. The users don't worry about the concrete realization or place of cloud server. The virtualization is the charter of the cloud system and the application don't need the information of hardware platform. The cloud system can make the application in different place or different hardware. The cloud system must provide reliability server for the user and the data in the cloud centre also must be protected. The cloud scale in the cloud system can be extended dynamically and can meet the growth of application and number of users. The application in the cloud also can be extended according to the number of user.

### III.CLOUD SECURITY PROBLEM

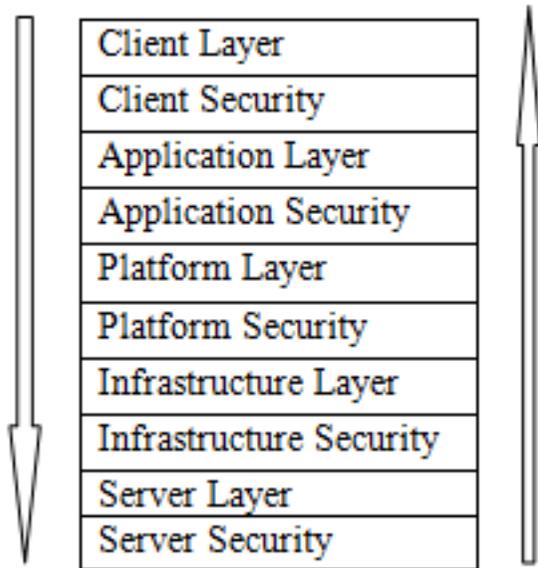
Innately, internet is communication infrastructure for cloud providers that use well-known TCP/IP protocol which users' IP addresses to identify them in the internet. Similar to physical computer in the internet that has IP address, a virtual machine in the Internet has an IP address as well. A malicious user, whether internal or external, like a legal user can find this IP addresses as well. In this case, malicious user can find out which physical servers the victim is using. By implanting a malicious virtual machine at that location attacker can launch an attack [2]. All the users who use same virtual machine as infrastructure, if a hacker steals a virtual machine or take control over it, he will be able to access to all users' data within it. Therefore, the hacker can copy them into his local machine before cloud provider detect that virtual machine is in out of control. The hacker analysing the data may hack valuable data [3]. The cloud system is running in the internet and the security problems in the internet also can be found in the cloud system. The cloud system is not different than the traditional system in the PC and it can meet other special and new security problems. The biggest concerns about cloud computing are security and privacy [5]. The traditional security problems such as security vulnerabilities, virus and hack attack can also make threats to the cloud system and can lead more serious results because of property of cloud computing. The data security audit also can be deployed in the cloud system. Data integrity requires that only authorized users can change the data and Confidentiality means that only authorized users can read data. Cloud computing should provide strong user access control to strengthen the licensing, certification, quarantine and other aspects of data management. In the cloud computing, the cloud provider system has many users in a dynamic response to changing service needs. The users do not know what position the data and do not know which servers are processing the data. Which network is transmitting the data is also unknown because of the flexibility and scalability of cloud system. The user can't make sure that data privacy is

operated by the cloud in a confidential way. The cloud system can deploy the cloud center in different area and the data can be stored in different cloud node. The different area has different law so the security management can meet the law risk. Cloud computing service must be improved in legal protection.

**IV. SECURITY SOLUTIONS**

**A) Level of Security**

Figure 1 shows the security deployment levels in each layer.



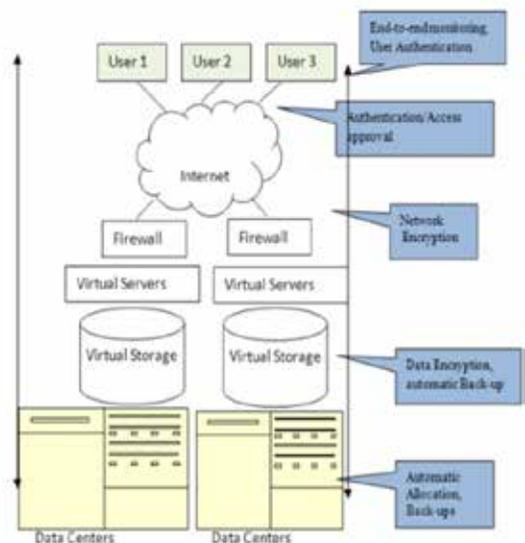
**Figure1. Security Deployment in each Layer**

As illustrated in figure 1, there must be a security deployed in each layer of the cloud computing to ensure that the information coming from client side and the server side is secured and legitimate. In server security requirements point of view, the cloud computing service provider must ensure that there is necessary security for the data centers for both hardware and software they are using. There must be a disaster and recovery plan in case there are unexpected natural disasters. Lacking of preparation may cause a lot of damages that would affect the client's day to day operations. The provider must have back-ups and necessary precautions to handle this kind of problems. Aside from having a physical security, data centers must have strong security Identification mechanism to avoid unauthorized access. Data centers must be secured enough because the core of cloud computing is being maintained inside it. Servers must have its security in form of software like antivirus program, Intrusion Detection and Prevention Systems and other software for maintaining the security, also the username, password and Digital ID for Access Control. Second is the Infrastructure Layer Security, as infrastructure layer provides platform virtualization to the clients/users. The type of the security that is necessary for this layer is internal to the cloud computing provider and the developer of the virtual machines.

Third is the Platform layer, platform layer must be well developed and that is reliable enough to run in any run-time platform. Secure enough that it cannot be used as a tool to launch an attack.

Fourth is the Application layer, the cloud application services deliver software as a service over the internet for eliminating the need to install and run the application on the customer own computers. The cloud computing provider must ensure that these applications are free from bugs and cannot be used as a tool to launch attacks.

Lastly, is the client layer, the cloud client consist of the computer hardware and the computer software that is totally based on the applications of the cloud services. Security needed for this layer is the safety of the hardware and the user access control, the company or enterprise has to ensure the user is legitimate and that is he/she is authorized to access that certain computer. Security overview in cloud computing environment is depicted in figure 2.



**Figure 2: Cloud computing security overview**

As a whole, we can say that if we meet all the security requirements of each layer then we can have secured Cloud Computing System.

**VI. CONCLUSION**

Cloud computing is a network-based environment that focuses on sharing computations or resources. The cloud system is not different than traditional system in the PC and it can meet other special and new security problems. The traditional security problems such as security vulnerabilities, virus and hack attack can also make threats to the cloud system and can lead more serious results because of property of cloud computing. The data security audit also can be deployed in the cloud system. Access control mechanisms are tools to ensure that authorized user can access their data and unauthorized access is prevented towards information systems.

**REFERENCE**

1. S. Roschke, et al., "Intrusion Detection in the Cloud," presented at the Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing, Chengdu, China, 2009. | 2. N. Mead, et al., "Security quality requirements engineering (SQUARE) methodology," Carnegie Mellon Software Engineering Institute. | 3. J.W.Rittinghouse and J.F.Ransome, Cloud Computing: Taylor and Francis Group, LLC, 2010. | 4. Dean, J. and Ghemawat, S. 2008. MapReduce: simplified data processing on large clusters. Communication of ACM 51, 1 (Jan. 2008), 107-113. | 5. Cloud Security Alliance : <http://www.cloudsecurityalliance.org/> | 6. Amazon Elastic Compute Cloud, <http://www.amazon.com/ec2/> | 7. Google App Engine, <http://appengine.google.com/> | 8. Salesforce, <http://www.salesforce.com/platform/> | 9. Microsoft, <http://www.microsoft.com/> | 10. VMware, <http://www.vmware.com/> | 11. IBM Blue Cloud project, | 12. <http://www03.ibm.com/press/us/en/pressrelease/22613.wss> | 13. OpenNEbula Project, | 14. <http://www.opennebula.org/> |