International Journal of Electrical, Electronics and Computer Engineering **5**(2): 11-15(2016)

ISSN No. (Online): 2277-2626

Cloud Computing : A Comprehensive View

Priyanka Thakur and Dr. Pawan Thakur

** Assistant Professor, Department of Computer Science,*
KC Group of Institutions, Una, (Himachal Pradesh), INDIA
***Assistant Professor, Department of Master in Computer Applications,*
Govt. P.G. College Dharamshala, Kangra (Himachal Pradesh), INDIA

(Corresponding author: Priyanka Thakur)
(Received 29 June, 2016 Accepted 12 August 2016)
(Published by Research Trend, Website: www.researchtrend.net)

ABSTRACT: Cloud Computing is the technique in which virtual resources and services are offered to the user through the internet. The data centers, known as 'cloud', are used to provide utilities via internet. It's the fastest emerging technology in last few years and has widely influenced both the IT industry and Business World and is used by the major service providers like Google, Microsoft, and Yahoo etc. The rapid growth in the field of Cloud Computing is due to its capabilities like scalability, reduced cost of capital (including hardware and software implementation) and dynamic availability of data and resources, but it also brings with it the major security concerns like the confidentiality and integrity of data. This paper provides a conceptual overview of cloud computing.

Keywords: Cloud Computing, cloud architecture, security issues.

I. INTRODUCTION

Cloud Computing is basically a Distributed Computing that allows dynamic access to the personal data stored by the user on the cloud. It also helps the user to run any application directly from a web based server without installing a suite of software and hardware, thus, making it more economical and user- friendly. According to the official definition by the National Institute of Standards and Technology's (NIST), "Cloud Computing is a model for facilitating convenient and on-demand network access to configurable computing resources like applications, networks, servers, storage and services that are provided to the users at any place and at any time with minimum interaction with service provider [1]. According to Gartner, Cloud Computing is "a technique of computing in which, scalable IT-related capabilities are provided "as a service" via Internet to the multiple users that share the services [2]". The user's perspective of cloud computing is that he is using the various services provided by the company without knowing where his files and data have been kept. He is providing its invaluable information in the hands of a third-party without knowing how they are managing and securing his data. From an organization's perspective, cloud computing ensures quality service to the customers through economical means [2]. In this paper, Section 1 represents the basic concept of Cloud Computing and Section 2 and 3 focuses on the review of cloud computing architecture and framework. Section 4 highlights the security issues involved in Cloud Computing.

II. CLOUD COMPUTING ARCHITECTURE

The cloud computing architecture consists of services and models which make the cloud computing accessible to end users. It allows the user to generate and customize the applications online. The Cloud Computing Architecture is broadly divided into two models: Cloud Service Models and Cloud Deployment Models.

A. Cloud Service Models

There are three service models as follows as shown in Fig. 1:

Software as a service (SaaS). Software as a Service model involves the end users dynamically running the software applications stored in the cloud of the service provider directly on their network [3]. In the traditional system, each user has to install the application on their own servers but with SaaS, a user can run the application remotely thus making it more user-friendly. Though it alleviates the risk of over- provisioning and under- provisioning [16]. In SaaS, the platform responsibilities are managed by the service provider itself [4]. For Example, Yahoo mail is a SaaS where Yahoo is the provider and we are the consumers. Software as a service has four approaches: [8]

(i) Single instance
(ii) multi- instance
(iii) multi-tenant
(iv) Flex tenancy

Platform as a service (PaaS). Platform as a service Model provides the run- time environment for application frameworks and databases as a service. It provides a platform where software can be developed, tested and implemented [3]. Therefore, this model is basically used by developers and deployers. It helps the developers to make the products really quick, easy and cost-effective. For example, Google App Engine helps the users to develop applications and implement it [4]. PaaS reduces the coding complexity by providing a more user- friendly environment for the developer. As these services are platform independent, multiple developers can work on a single project remotely. The basic disadvantage of PaaS is that the developer is not able to use the conventional tools fully [13].

Infrastructure as a service (IaaS). Infrastructure as a service is defined as consuming the physical resources like networks, servers, storage, firewalls, load balancers etc. from the cloud itself [3]. IaaS is based on the concept of virtualization. The user has to pay the usage fee to the service provider, in return for the fully outsourced services, including hardware and software, consumed by it. IaaS provides pay-as-you-go service model, as the organizations only pay for the resources they use in a given period, typically by the hour, week or month, thus increasing the scalability of resources according to the demand of consumers. This pay-as-you-go model eliminates the capital cost of deploying in-house hardware and software [3]. IaaS is suitable for workloads that are temporary or experimental. Its other features include dynamic scaling, virtualization, and policy-based services [14]. However, users should keep a check that he has not been charged for unauthorized services. For example, Amazon Web Services provides "EC2" computing platform, and "S3" storage platform.

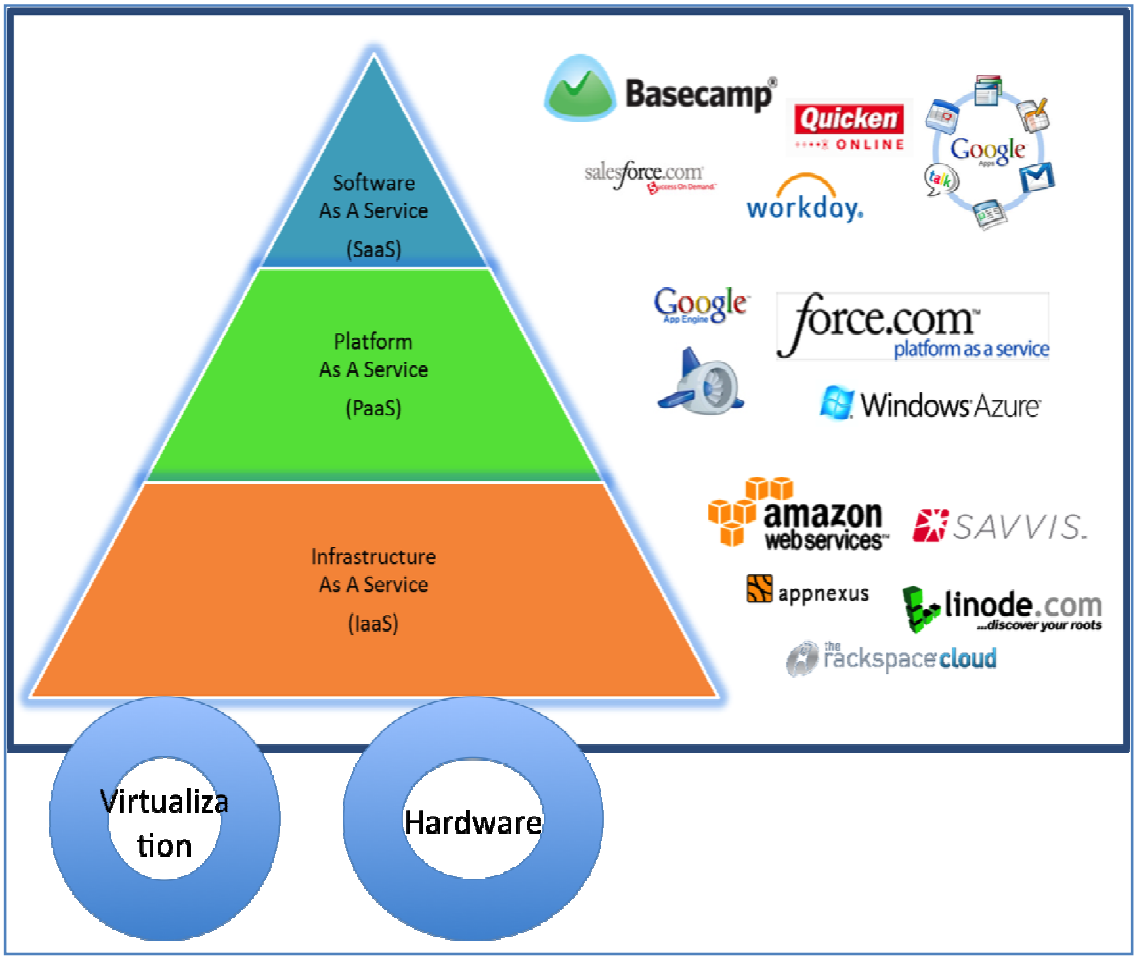

Fig. 1. Types of Service Models.

B. Cloud Deployment Models

A deployment model defines the type of access to the cloud.

Public/External Cloud. In this cloud, the services (resources, network, applications, and storage) are publicly accessible [3]. The services provided by Public cloud may or may not be free to the user and are run on a remote system. The service of a Public cloud can be consumed by several organizations, having similar requirements and which seek to share infrastructure too. These types of clouds are economical to set-up because the application, hardware, and bandwidth costs are covered by the provider [5]. Examples of public clouds are Amazon Elastic Compute Cloud (EC2) and Google AppEngine.

Private/Internal Cloud. This type of model is owned by an organization exclusively. The organization possesses the full control on the cloud whether it is on-premise or managed by the third party [3]. Though such type of cloud doesn't serve the basic purpose of cloud computing as setup and maintenance cost of cloud infrastructure is high. It is favorable for high profit organizations and the decision of owning a private cloud is taken after cost and benefit analysis. For example, VMware has its own private-cloud architecture built on top of VMware vSphere.

Hybrid Cloud. Hybrid cloud computing combines the benefit of both the public and private clouds. An organization can store its valuable and sensitive data on a private cloud simultaneously using the public clouds as and when needed [5]. For instance, during high work load and time constrain projects, public cloud services can be used. This concept is known as Cloud Bursting. It can be used as an off-premise disaster recovery site or business continuity services. The hybrid cloud implementation is based on factors such as data security, compliance requirements, the level of control needed over data, and the applications an organization uses.

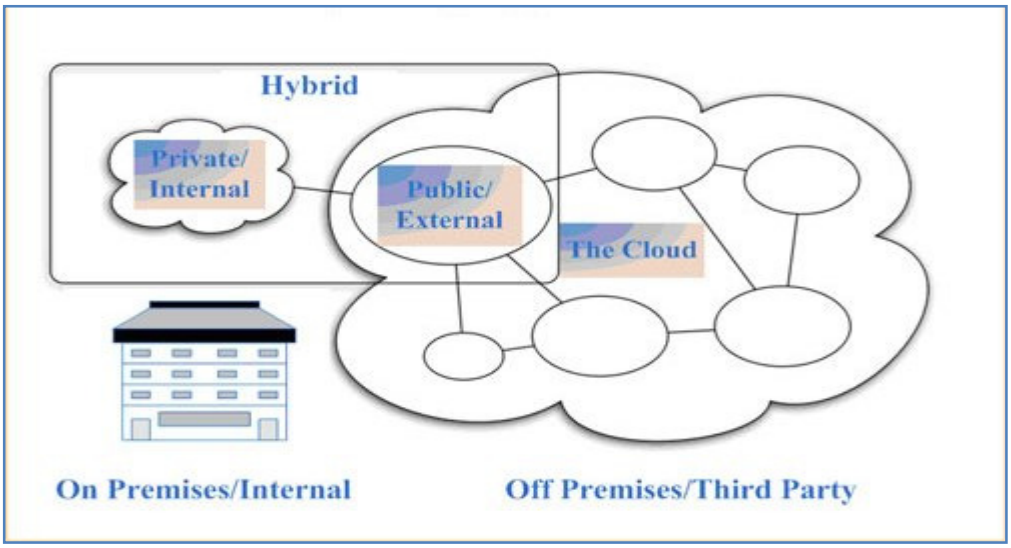

Fig. 2. Types of Deployment Models.

III. CLOUD COMPUTING FRAMEWORK

Once the type of cloud environment has been determined for a specific project or on the consumer's requirement, the cloud computing framework helps the organizations in identifying challenges that should be considered prior to cloud implementation [9]. The most commonly used framework for cloud computing is MapReduce.

MapReduce. Nowadays, millions of people are using the internet for various applications and terabytes of data is generated in a small time interval. This data has to be processed efficiently and accurately. For this, Parallel and Distributed Programming models are used. One such model, MapReduce is proposed by Google. Hadoop is the most popular open source implementation of this model. The main advantage of MapReduce is its efficient scalability. It basically consists of Map() and reduce() procedure. Map() processes a key/ value pair to generate outputs which act as an input for reduce() to give the final result. The two tasks can be represented as:

Map: (k1, v1) → list (k2, v2)
Reduce: (k2, list (v2)) → list (v2)

The intermediate and processed data are kept in HDFS (Hadoop Distributed File System) [7].

Implementation. MapReduce program executes in three stages: map stage, shuffle stage, and reduce stage. (i) **Map stage:** In map stage, the input data is processed. Generally, this input data is stored in the Hadoop file system (HDFS) as a file or a directory. The input file is passed to the mapper function line by line [10]. The mapper processes the data and converts them into several small packets of data, ensuring that redundancy does not exist [15].
(ii) **Reduce stage:** Reduce stage is the combination of the Shuffle stage and the Reduce stage. The data packets formed in the Map stage are then redistributed such that all data belonging to one key are located on the same node. Each group of data that comes from the Mapper is processed in parallel by the Reducer [10]. After processing, the whole new set of output is stored in the HDFS.

Execution Overview. The execution process is divided into three steps. In the first step, data is filtered or processed and the output is stored in a temporary storage. It also removes the redundant data. In the second step, data is shuffled or redistributed on the output keys so that identical data is kept in one place. In the third step, all the key values are processed in parallel to give the final output.
MapReduce uses an easy way to manage the faults. The master node keeps a check on every mapper and reducer by sending signals/message at times. If they don't respond in a fixed interval of time, the respective node is considered to be botched and the task is reassigned to some other node from the scratch in case of a mapper as the data is stored on a local file system [7]. But in the case of a reducer, if the execution of the task is completed, the output can be fetched anytime from the global file system as shown in Fig. 3.

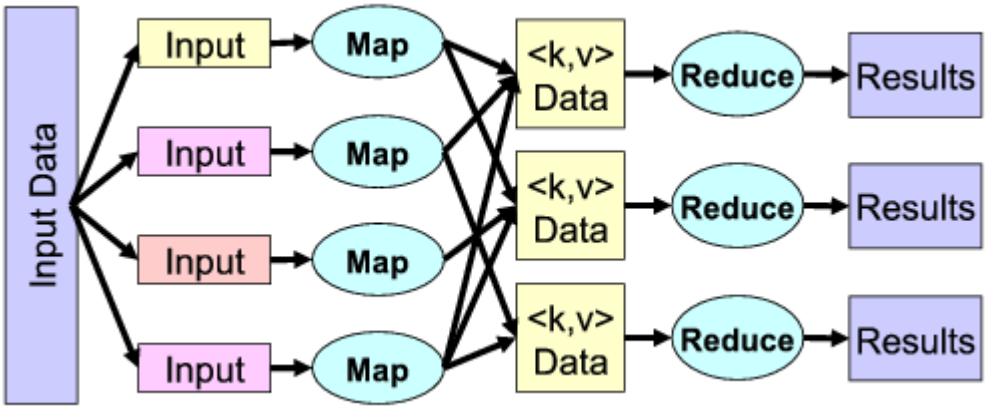

Fig. 3. Map Reduce Framework.

IV. CLOUD COMPUTING SECURITY ISSUES

Data Confidentiality and Integrity. Privacy of data is one of the major issues in cloud computing as all the invaluable and private data of the client is handled by a third party. Also, in cloud computing, virtual resources are shared between clients and public networks are used, thus the data is exposed and becomes more prone to the cyber attacks and threat to the privacy of data increases. It is also possible that if the user has deleted a particular data, it may not actually get deleted because more than one copy of the same data is stored which are not available at the time of deletion. The risk of intrusion is highest in cloud computing due to multi-tenancy [2]. A hacker can add a masked code at the application level and if this code gets executed, then it can corrupt or steal the data [2]. The service should be intelligent enough to segregate the data from different users [6]. To ensure the confidentiality of data, encryption techniques are used by the service providers. Hashing is the most common technique used for encryption. The data is encrypted and the key resides with the user. The security risks are high during the data transfer through communication networks [6]. Therefore, security checks are applied to prevent breaches. This involves the use of various encryption techniques such as Secure Sockets Layer (SSL) and Transport Layer Security (TLS) [2]. Moreover, access control techniques should be used by service providers to check the authorization and authentication of the user.
Locality and Availability of resources. Transferring data between client and service providers is a complex task as different users and service providers have different software and hardware architectures.

Moreover, as the services are provided by the same cloud to different clients, therefore, the quality of service depends on various factors like availability of resources, data transfer speed etc. Horizontal and vertical scaling is used to improve the efficiency of cloud services during multi-tenancy [3]. The locality of data is of utmost importance in enterprise architecture [2]. The Service provider must be capable of providing reliability to the consumer. The service providers have to make sure that all the files are regularly backed up so that data can be recovered under any circumstances. Data availability means that if the accidents such as hard disk damage or network failures occur, then to what extent that user's data can be recovered and how the users would verify their data by technically rather than depending solely on the guarantee of the cloud service provider alone.

The issue of storing data on the remote servers located throughout the world is a serious concern of clients because the cloud vendors are administered by the local laws and, therefore, the cloud clients should be aware of those laws. Moreover, the service provider should ensure the confidentiality, integrity, and security of the data [12]. The cloud provider should discuss all such concerns with the client and build a trust relationship in this connection. The cloud vendor should explain the jurisdiction of local laws to the clients and ensures the security of the data.

V. CONCLUSION

Though cloud computing has numerous advantages but the major threat of security prevents many organizations to adopt it [2]. Though many measures have been taken by the IT researchers and professionals to make cloud computing a secure technique, but still a lot of improvement is needed. Even most of the frameworks come with the burgers; they are naïve and not very effective. Mutual understanding between the service provider and user is necessary for providing better services. Virtualization technology provides good support to the aim of cloud computing like higher resource utilization, elasticity, handling IT cost, as well as cloud computing has various flexible services of deployment model which is also one of the major issues of adopting this computer paradigm [3]. Virtualization concept has open shared nature which is responsible for violation of security policies and laws as well as degrades the computing reputation and performance [17]. We can say that need for further work on various security mechanisms is required in order to provide reliable and transparent services that can be trusted by the user.

REFERENCES

[1]. NIST Special Publication 800-145 The NIST Definition of Cloud Computing by Peter Mell ,Timothy Grance.

[2]. Dr. S. Mehfuz, Dr. G. Sahoo, Rashmi, International Journal on Cloud Computing: Services and Architecture (IJCCSA), Vol. **3**, No.4, August 2013, Securing Software as Service Model of Cloud Computing: Issues and Solutions.

[3]. Suruchee V. Nandgaonkar, Prof. A.B Raut, *International Journal of Computer Science and Mobile Computing,* Vol. **3** Issue.4, April- 2014, A Comprehensive study on Cloud Computing

[4]. Cloud Computing Available at: https://en.wikipedia.org/wiki/Cloud_computing

[5]. Types of Clouds available at: http://www.asigra.com/blog/cloud-types-private-public-and-hybrid.

[6]. Monjour Ahmed and Mohammad Ashraf Hossain, *International Journal of Network Security & Its Applications (IJNSA),* Vol. **6**, No.1, January 2014, Cloud Computing and Security issues in the Cloud.

[7]. Sara Dadizadeh, Amit Goyal, A survey on Cloud Computing.

[8]. Cloud Computing Architecture at: https://en.wikipedia.org/wiki/Cloud_computing_architecture

[9]. Hadoop Model Available at: https://hadoop.apache.org

[10]. Map Reduce Model Available at: https://en.wikipedia.org/wiki/MapReduce

[11]. Shucheng Yu, Wenjing Lou, and Kui Ren, Data Security in Cloud Computing.

[12]. Raj Kumar, International Journal of Advanced Research in Computer Science and Software Engineering," 2015, Research on Cloud Computing Security Threats using Data transmission"

[13]. Types of Service Model Available at: https://www.linkedin.com/pulse/3-service-4-deployment-models-cloud-computing-sankar-somepalle

[14]. Types of Deployment Model available at: http://www.itinfo.am/eng/cloud-computing/

[15]. MapReduce Framework Available at: http://alumni.cs.ucr.edu/~jdou/misco/

[16]. Yi-Ju Chiang, Yen-Chieh Ouyang and Ching-Hsien Hsu, An Optimal Cost-Efficient Resource Provisioning for Multi-servers Cloud Computing, available at: http://ieeexplore.ieee.org/document/6820997/?reload=true&arnumber=6820997.

[17]. Thakur Pawan and Thakur Sikma (2013), "*Cloud Computing",* Satya Prakashan New Delhi.