A Study of Data Storage in Cloud Computing

(Created By): Satish kumar, Karandeep Singh, Vikas, (Guidance): Shubham Pathania, C.K Raina
Cse Department, Adesh Institute of Technology, GharuanMohali(Punjab)

Abstract- Cloud is a collection of computer resources and provides a million of services to its user simultaneously. A Cloud provides a friendly environment to its user and various services. These services are used in Publicly, Private, Hybrid and Community terms. Cloud computing is model of computing that use the internet for sharing of information, software and resources to computer and other devices. Cloud computing provides on demand services to its clients. Data storage is among one of the primary services provided by cloud computing. Cloud service provider hosts the data of data owner on their server and user can access their data from these servers. As data, owners and servers are different identities, the paradigm of data storage brings up many security challenges. An independent mechanism is required to make sure that data is correctly hosted in to the cloud storage server. In this paper, we will discuss the different techniques that are used for secure data storage on cloud.

Keywords- Cloud computing, Data storage, Cloud storage server.

HISTORY

Amazon was the first major organization to modernize its data centres which were utilizing only about 10% of their capacity at any given time (which was commonplace at the time, companies were worried about sudden spikes in capacity needs).Amazon realized that the new cloud computing infrastructure model could allow them to use their existing capacity with much greater efficiency.Meanwhile, Google had become a key player in the Internet commerce marketplace. In 2006 the company launched its Google Docs services, which brought the power of cloud computing and document sharing directly to end users.

I.   INTRODUCTION

Cloud computing is the combination of many pre-existing technologies that have matured at different rates and in different contexts. The goal of cloud computing is to allow users to take benefit from all these technologies. Many organizations are moving into cloud because it allows the users to store their data on clouds and can access at anytime from anywhere. Data breaching is possible in cloud environment, since data from various users and business organizations lie together in cloud. By sending the data to the cloud, the data owners transfer the control of their data to a third person that may raise security problems. Sometimes the Cloud Service Provider(CSP) itself will use/corrupt the data illegally.

The history of cloud computing starts way back in the 1960s, when an "intergalactic computer network" was first suggested, and in recent years the technology has served to shake-up both the enterprise IT and supplier landscape.The term "cloud" is used as a representation of the Internet and other communications systems as well as an abstraction of the underlying infrastructures involved.

EARLY 1990S

Historically, telecommunications companies primarily offered only dedicated, point-to-point data circuits to their users. Beginningin the 1990s, however, they began expanding their offerings to include virtual private network services. This allowed thetelecom companies to provide the same quality of service at a fraction of the cost, as they were able to optimize resource utilization in order to improve the efficiency of their overall bandwidth.

EARLY 2000S

Shortly after Salesforce.com brought this new concept to the world's attention, Amazon.com proved it could outlast the dot-com bubble burst with the introduction of its web-based retail services in 2002.

Security and privacy stands as majorobstacle on cloud computing i.e. preserving confidentiality, integrity and availability of data. A simple solution is to encrypt the data before uploading it onto the cloud. This approach ensures that the data are not visible to external users and cloud administrators.Today World relies on Cloud computing to store their public as well as personal information. Cloud computing, becomes the requirement for every user and provide relevant hardware, software and service according to the need that users put forward.   With the rapid development of the Internet, user's requirement is realized through the Internet, different from changing with the need.
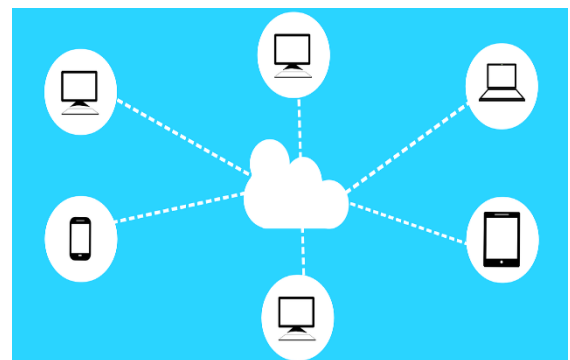
II.   CLOUD STORAGE

Cloud storage is one of the primary use of cloud computing. We can define cloud storage as storage of the data online in

the cloud. A cloud storage system is considered as a distributed data centres, which typically use cloud-computing technologies and offers some kind of interface

There are four main types of cloud storage
- **PRIVATE CLOUD STORAGE**

It is also known as Internal Cloud or on-premises Cloud. It is managed and operated by single organization or a group. It is also known as internal cloud or on-premise cloud, Aprivate cloud provides a limited access to its resources and services to consumers that belong to the same organization that owns the cloud. In Private Cloud Storage the enterprise and cloud storage provider are integrated in the enterprise's data centre. In private cloud storage, the storage provider has infrastructure in the enterprise's data centre that is typically managed by the storage provider. Private cloud storage helps resolve the potential for security and performance concerns while still offering the advantages of cloud storage.

In private cloud storage the data stored in is very safe and secure. Any stranger cannot access easy.The security and control level is highest while using privatenetwork. An example of private cloud storage inyour phones google drives and other drives.
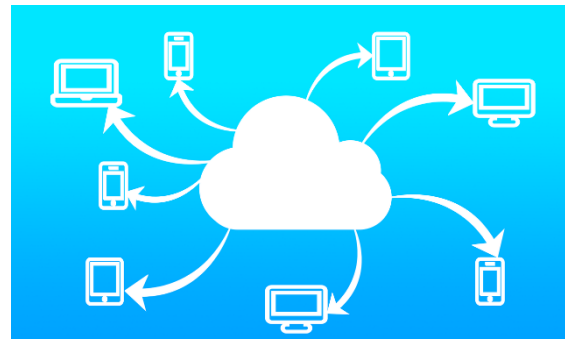
- **PUBLIC CLOUD STORAGE**

It is also known as external cloud or multitenant cloud. It is available and open used by general public. It may be owned and managed by government organization or some combination of them.It is referred as external cloud or multitenant cloud, this model represents an openly accessible cloud environment in this cloud can be accessed by general public. Customer can access resources and pay for the operating resources. Public Cloud can host individual services as well as collection of services.

Public cloud solutions are readily available from Google, Amazon, Microsoft, and others. Public cloud services provide infrastructure and services to the public, and you, or your organization, secure a piece of that infrastructure and network. Resources are shared by hundreds or thousands of people. Gmail and U of I Box are examples of public cloud services. Google reported in April 2017 that it had 1 billion monthly users. While your email account is protected by a password, the hardware on which it is stored is shared by over 1 billion people.

- **HYBRID CLOUD STORAGE**

It is a combination of public and private cloud. Clients can deploy sensitive and private data on private cloud to host less critical information. Hybrid cloud is cloud computing solution for those seeking a balance of accessibility and security. One main advantage of hybrid cloud computing is the ability to put sensitive data on private cloud and keep the bulk of everything else on the public server.

for storing and accessing data. When storing data on cloud, it appearsas if the data is stored in a particular place with specific name.

- **COMMUNITY CLOUD STORAGE**

Community cloud provides cloud computing services to a group of organization or individuals. Community clouds are designed for organization working on joint projects which require a central cloud computing services. It is comparatively more secure than the public cloud but less secure then the private cloud.

It refers to a special purpose cloud environment which is shared and managed by community that shares common concerns (like security, governance, compliance etc).

III.   CLOUD COMPUTING SERVICE MODELS
cloud-computing providers offer their "services" according to different models, of which the three standard models per NIST are Infrastructure as a Service (IaaS), Platform as a Service (PaaS), and Software as a Service (SaaS).

- **Infrastructure as a service (IaaS)**

Infrastructure as a service helps companies to move their physical infrastructure to the cloud with a level of control similar to what they would have in a traditional on-premise data center. IaaS provides the closest resemblance to the in-house data center compared to other services types. Core data center infrastructure components are storage, servers (computing units), the network itself, and management tools for infrastructure maintenance and monitoring. Each of these components has created a separate market niche.

IaaS Providers, such as AWS, supply a virtual server instance and storage, as well as APIs that enables users to migrate workloads to a VM. Users have an allocated storage capacity and can start, stop, access and configure the VM and storage as desired. IaaS providers offer small, medium, large, extra-large and memory- or compute-optimized instance, in addition to customized instances, for various workload needs.

- **Platform as a service (PaaS)**

PaaS (Platform as a Service) is a cloud service model that provides customers with a configurable application platform including a pre-installed software stack. PaaS can be viewed as another abstraction layer above the hardware, operating system, and virtualization stack.

In the PaaS model, cloud providers host development tools on their infrastructures. Users access these tools over the internet using APIs, web portals or gateway software. PaaS is used for general software development, and many PaaS providers host the software after it's developed. Common PaaS providers include salesforce's Force.com, AWS Elastic Beanstalk and google App Engine.

- **Software as a service (SaaS)**

SaaS is a software delivery method that provides access to software and its functions remotely as a web-based service. Instead of paying an upfront fee to purchase and/or licence software, SaaS customers pay a recurring (often monthly or annual) fee to subscribe to the service. In general, user can access the SaaS from any Internet-connected device like mobile or computer from any location, any time day or night. Well Known example of SaaS include Salesforce.com, Microsoft Office 365, Google G suite, Dropbox, Adobe Creative Cloud and others.
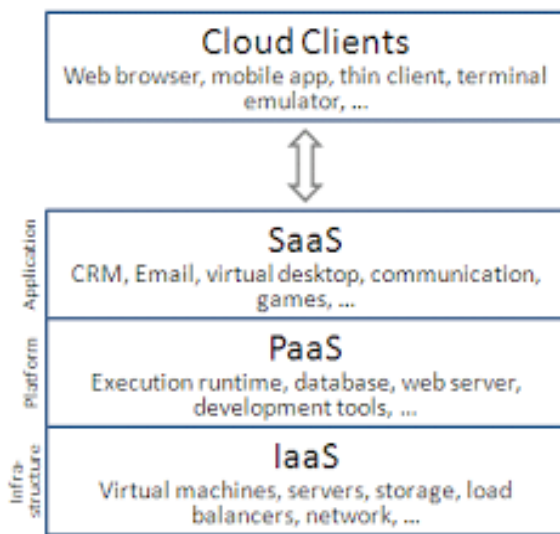
IV. CHARACTERISTICS OF CLOUD COMPUTING

On-Demand self-service; This mean that cloud customers can sign up for, pay for and start using cloud resources very quickly on their own without help from a sales agent.

Broad Network access; Customers access cloud services via the internet.

Measured Services; Customers pay for the amount of resources they use in a given period of time rather than paying for hardware or software upfront.

Rapid elasticity or Expansion; Cloud customers can easily Scale their use of resources up or down as their needs change.

Migration Flexibility;Organizationcan move certain workloads to or from the cloud -or to different cloud platforms- as desired or automatically for better cost savings or to use new services as they emerge.

Hardware and Maintenance: Cloud Computing helps to reduce hardware and maintenance cost because there is no need to be installed any application on user's computer.

V. SECURITY AND PRIVACY

Every enterprise will have its own identity management system to control access to information and computing resources. Cloud providers integrate with customer's identity management system into their own infrastructure using federation or SSD technology or a biometric based identification system. It links the confidential data of the users to their biometrics and stores it in an encrypted fashion.

- **Physical Security**

Cloud service providers physically secure the IT hardware (servers, routers, cables etc.) against unauthorized access, interference, theft, fires, floods etc. and ensure that essential supplies (such as electricity) are sufficiently robust to minimize the possibility of disruption. This is normally achieved by serving cloud applications from 'world-class' (i.e. professionally specified, designed, constructed, managed, monitored and maintained) data centre.

- **Privacy**

Providers ensure that all critical data (credit card numbers, for example) are masked or encrypted and that only authorized users have access to data in its entirety. Moreover, digital identities and credentials must be protected as should any data that the provider collects or produces about customer activity in the cloud.

- **Confidentiality**

Data confidentiality is the property that data contents are not made available or disclosed to illegal users. Only authorized users can access the sensitive data while others, including CSPs, should not gain any information of the data. Meanwhile, data owners expect to fully utilize cloud data services, e.g., data search, data computation, and data sharing, without the leakage of the data contents to CSPs or other adversaries.

- **Access controllability**

Access controllability means that a data owner can perform the selective restriction of access to her or his data outsourced to cloud. Legal users can be authorized by the owner to access the data, while others cannot access it without permissions.

- **Integrity**

Data integrity demands maintaining and assuring the accuracy and completeness of data. A data owner always expects that her or his data in a cloud can be stored correctly and trustworthily. It means that the data should not be illegally tampered, improperly modified, deliberately deleted, or maliciously fabricated. If any undesirable operations corrupt or delete the data, the owner should be able to detect the corruption or loss. Further, when a portion of the outsourced data is corrupted or lost, it can still be retrieved by the data users.

- **Encryption**

Some advanced encryption algorithms which have been applied into cloud computing increase the protection of privacy. In a practice called crypto-shredding the keys can simply be deleted when there is no more use of the data.

Attribute-based encryption(ABE);

Attribute based encryption is a type of public-key encryption in which the secret keyof a user and the ciphertext are dependent upon attributes (e.g. the country in which he lives, or the kind of subscription he has). In such a system, the decryption of a ciphertext is possible only if the set of attributes of the user key matches the attributes of the ciphertext.

Ciphertext-policy ABE(CP-ABE); In the CP-ABE, the encryption controls access strategy. The main research work of CP-ABE is focused on the design of the access structure.

Key-policy ABE(KP-ABE);In the KP-ABE, attribute sets are used to describe the encrypted texts and the private keys are associated to specified policy that users will have.

Searchable encryption(SE);Searchable encryption is a cryptographic system which offer secure search functions over encrypted data. E schemes can be classified into two categories: SE based on secret-key (or symmetric-key) cryptography, and SE based on public-key cryptography. In order to improve search efficiency, symmetric-key SE generally builds keyword indexes to answer user queries.

Fully Homomorphic encryption(FHE);Fully homomorphic encryption allows computations on encrypted data, and also allows computing sum and product for the encrypted data without decryption.

VI.   FUTURE OF CLOUD COMPUTING

Cloud computing is changing businesses in many ways. Whether it is the way they store their data or how they protect their secure information, cloud computing is benefitting all businesses in every sector.

Smart businesses are always looking for the most innovative ways to improve and accomplish their business objectives. When it comes to cloud technology, more and more businesses are realizing the benefits this technology can provide them and are beginning to seek more cloud computing options to conduct their business activities. With the multitude of future technology trends in cloud computing, companies in every sector are benefitted by the opportunities cloud technology offers.

- **Growth in cloud service**

Cloud computing future growth all began when the growth of infrastructure as a service, IaaS, and platform as a service, PaaS, expanded the number of cloud solutions available in public and private sectors. As IaaS and PaaS continue to be used worldwide to achieve diverse goals, we will see these solutions as the most deployed cloud services around the world. Cisco predicts that SaaS, software as a service,

solutions will account for more than 60% of all cloud-based workloads this year. They also predict that PaaS and IaaS solutions will increase throughout 2018. Any business looking to simplify their operations and make services easier to access for customers will most likely move toward cloud services solutions.

- **Increased storage capacity**

A huge aspect affecting the future of cloud computing is the amount of storage cloud computing will offer companies and individuals. This growth is because many businesses are adopting cloud technology as a huge part of doing business. It is predicted that providers will bring more data centres online with larger-capacity storage equipment throughout this year. Cisco estimates the storage capacity of the cloud will double this year alone. With this increased storage, more businesses will be able to store large data sets and perform analytics using cloud computing.Being able to perform analytics on this massive amount of data will allow companies to gain valuable insights into customer behaviour human systems, and strategic financial investments, just to name a few.

- **Enhanced internet quality**

The quality of the internet has been getting immensely better every year since it was created. 2018 is expected to be no different, as the amount of data generated and stored around the world increases. Customers today already expect high-quality, fast-loading services and apps and this expectation will enhance network quality and cloud computing. This high-quality expectation will also lead businesses to upgrade their platforms and services to be more responsive to the needs of their customers. As the quality of the internet is enhanced, IoT and IoE industries will benefit a great deal from the faster network speeds and the ability to receive and deliver data more efficiently in real time.

VII.   CONCLUSION

Cloud computing enables users to store their data in remote storage location. But data security is the major threat in cloud computing. Due to this many organizations are not willing to move into cloud environment. To overcome this, confidentiality, integrity, availability should be encapsulated in a CSP's Service- Level Agreement (SLA) to its customers. Otherwise ensure that any sensitive information is not put into a public cloud and if any it is to be stored in encrypted form. Effective auditing mechanisms also can be used for providing data integrity.

VIII.   REFERENCES

[1]. I.S. Jacobs and C.P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G.T. Rad and H. Suhl, Eds. New York: Academic, 1963, pp. 271-350.
[2]. PARKHILL, D. The Challenge of the Computer Utility. Addison-Wesley Educational Publishers Inc., US, 1966.
[3]. HAMILTON, J. Internet-Scale Service Efficiency. In Large-Scale Distributed Systems and Middleware (LADIS) Workshop (September 2008).
[4]. Above the Clouds: A Berkeley View of Cloud Computing. Michael Armbrust, Armando Fox, Rean Griffith, Anthony D.

Joseph, Randy Katz,AndyKonwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, and MateiZaharia UC Berkeley Reliable Adaptive Distributed Systems Laboratory http://radlab.cs.berkeley.edu/ February 10, 2009

[5]. Architecting for the Cloud: Best Practices, Jinesh Varia , January 2010 [5]

[6]. M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.
http://en.wikipedia.org/wiki/Cloud_computing.
https://www.techfunnel.com/information.../in-cloud-computing/
https://en.wikipedia.org/wiki/Cloud_computing_security
https://www.webopedia.com/TERM/C/cloud_computing.html