

Review of Cloud Storage Techniques

Madhura K^{1*}, Alamelu Mangai J²

Department of Computer Science and Engineering, Presidency University, Bengaluru, India

Corresponding Author: madhura@presidencyuniversity.in

DOI: <https://doi.org/10.26438/ijcse/v7si16.6568> | Available online at: www.ijcseonline.org

Abstract—The term “cloud” is very common term and is used in both biological plus technical fields. During rainy reasons cloud plays a very important role in the life of human beings as it stores all molecules for showering the water and humans can preserve the water for future living. In the same way, in technical fields it offers a list of facilities for the user with respect to data such as data storage, security to data, data sharing, and data preservation for the future, updation and maintenance. Cloud is one such platform which allows the user to store the data and application together at one place and it is made available irrespective of the place of access at any time. Cloud can be used by an individual or a company for many purposes such as storing, retrieving and sharing documents, photos, e-mails, etc. The big question which always runs on every researcher’s mind is the need of cloud. The main problem is to find a place to store the huge growing amount of data. According to a survey done by Peter Lyman and HAL R Varian, world produces approximately two Exabyte’s of data every year which means that each individual contribution is around 250 megabytes. So it is possible to fit this much amount of data in any storage medium like papers, magnetic tapes, disks and drives. This paper gives a brief insight into the existing cloud storage techniques which are used by organizations.

Keywords: IAAS, PAAS, SAAS, GDS-LC, SPMCloud, etc.

I. INTRODUCTION

A cloud offers a virtual view as a single system by aggregating the resources over the network such as soft wares, applications, databases, different servers, etc. and are used in both private and public organizations. “Cloud Computing” is the term which is often coined along with cloud. Cloud computing allows an individual or a company to access these resources on the internet. Cloud provides service to the requested service only and it follows pay-on-demand service. That is only to the demanded service have to pay.

The following sections covers some important aspects of cloud. Section I.A covers key advantages offered by the cloud. In section I.B different varieties of clouds available and its corresponding features are listed. Regarding the different service provided is discussed in the section I.C. The features of cloud based the performance is given in the section I.D. Finally, the section III is covering on the existing storage techniques in cloud based on price, space and security aspects.

A. KEY ADVANTAGES OF CLOUD COMPUTING

1. Reduction in Cost

Cloud services are free from huge initial investment. Payment is based on the operation performed without much investment on hardware.

2. Increased Speed

As clouds are formed with a cluster of computers, it can perform requested operations in a faster manner.

3. Increased Productivity

Because of the large computation capacity, clouds have the increased work rate which in turn improves of the revenue of the company.

4. More storage capacity

Allows organizations to store and access huge amount of data.

5. Automatic Updation

Server of clouds are programmed to have the capacity of automatically update the software versions, without the requirement from cloud service providers.

6. Quality Performance

Since the resources are distributed over different servers and the services are provided in an automated manner, clouds give precise, better quality service in limited amount of time.

7. Security

Gives a persistent storage to the data, so even if an individual loses data can be retrieved back [1].

B. TYPES OF CLOUD

1. Private cloud

A private cloud has a single ownership. That is it belongs to a single person, company or enterprise and it provides services to the owner of cloud only.

2. Public cloud

In public cloud, the resources are distributed over the Internet and can be accessed by an authorized third party. Popular public clouds in use are Amazon's AWS, Google's Google Cloud Platform, Microsoft's Azure, IBM's IBM cloud, etc.

3. Hybrid cloud

It is the combination of more than one cloud which are either private clouds or public clouds or both sometimes, but it is subject to the proprietary of an organization. Hybrid clouds are used to handle clouds bursts where an application from private cloud is transferred into public cloud when the request for capacity increases [2].

C. TYPES OF CLOUD SERVICES

1. Infrastructure as a service (IAAS)

IAAS cloud computing service model, provides resources to build the hardware upon which software can be built. That is, within the infrastructure boundary, it allows cloud users to use their own software and applications to access computing resources like storage, network and servers. Therefore, it allows the user to create their own platform for software and hardware [2]. Best example is a cloud OpenStack service which provides infrastructure by combing all available open source software on the internet [13].

2. Platform as a service (PAAS)

Paas cloud computing service model supplies resources for building the software application and services from Internet. The resources are supplied from a third-party provider for which the user doesn't have to download and install it. The platform for software and hardware is provided by the service provider through its own infrastructure.

3. Software as a service (SAAS)

Here, applications are provided as a service to customers, where customers can access the applications over the internet. Platform and infrastructure both are supplied as a part of service for accessing the cloud based software [2].

D. MAIN FEATURES OF CLOUD IN TERMS OF PERFORMANCE

1. Provides Automated-service

The cloud users expect the resource access and services in an instant manner. To support this clouds work in an automated manner to process and serve the customer requests.

2. Pay on the usage basis

Cloud allows its users to access and use the required resources only and bill can be paid based on the amount of usage.

3. Capacity to Stretch

Cloud gives an illusion to its users that it has enormous amount of computing capacity by providing the demanded resource instantly and manages the server load.

4. Customization

By creating a virtual environment, clouds allow customer to build a personalized environment through remote servers [2].

II. VARIOUS CLOUD STORAGE TECHNIQUES

1. Online data storage with built-in security

This feature has provided a cloud architecture to store data online with built-in security by dividing the data based on finite field polynomial root. This scheme of division is based on two ways. First, the data are randomly chosen from the network and are stored in the cloud servers. Secondly, these partitioned data are retrieved back in order to generate the master copy of data. In this method, cloud uses separate algorithm to perform data partitioning, reconstruction and to handle the complexity in both. The data will be accessible to the one who knows password and location of partitioned data [3].

2. Authentication based on the identity

In cloud computing since the data are distributed among numerous servers, the security of data is at high risk. So along with the service, a cloud must ensure the secured communication too. Till the year 1995, companies were using Secure Socket Layer (SSL) Hand shaking Protocol to secure the unsecured network connection between client and server like the internet. But when the load on SSL increases it fails to secure the communication such as when the networks turns into distributed computing SSL wastes the client and server resources in both computing and communication links. To overcome the drawback of SSL, authentication process which is based on identity came in to picture. It uses a hierarchical model for authentication based on encryption and signatures. Through simulation it has been proved that Identity based authentication method is more efficient and lightweight, especially on client side than SSL scheme. Here, authentication process is composed of encryption based on identity and signature based on identity. That is, when a client sends message to server with session ID, the server converts the message into cypher text and replies back to client with session ID and a unique random number. That is when client is authenticated, the server sends its signature as acknowledgement. From 2015 onwards

companies started using Transport Layer Security Protocol (TLSP) for this purpose [4].

3. Support for dynamic data with public auditing

Cloud clients never own their data physically, therefore it is a big challenge to maintain the integrity support. Technique such as proofs for data possession and retrieval designed to handle this but failed to support dynamic nature of data.

The public auditing scheme, it removes the index range through an index switcher in data storage, thus gives an efficient support for the dynamic nature of data [5] [14].

4. Potential support for auditing by third party

Cloud computing gives immense computing power to consumers mean while it should ensure integrity, security and performance. Since the consumers cannot possess the data physically in cloud, it is not possible to use standard security techniques. The third party auditing (TPA) ensures transparency between cloud service providers and consumers. It is also efficient in terms of monetary and performance as it can run in malicious client also [6] [15].

5. Support for caching storage scheme

This scheme works on increasing the cache hit ratio in terms of latency and cost with the aim of providing an efficient service to cloud users. For this purpose, the cache is partitioned into two areas: one is the high priority latency and other is the low priority price aware regions. This virtual division is done by a Greedy Algorithm based on Dual Size for latency (GDS-LC) normalization. GDS-LC algorithm is highly expandable and it is modified to higher grade to improve the frequency (GDS-LCF) of cache hit ratio. The simulation of GDS-LC and GDS-LCF has been done using Amazon Simple Storage Services (S3) and achieved both the goals of minimizing latency and cost [7].

6. Encrypted Duplication Storage System

This technique is implemented by keying and rekeying system which replaces the already existing key with a new one for the encryption.

The keying system was not efficient during data leakage and security. For redefining the security protection, Rekeying aware encrypted reduplication system (REED) was implemented to protect from data leakage which improved performance, security and dynamic access control of cloud systems [8] [16].

7. Support for File system Backup storage

Cloud servers provides the facility to keep a backup storage of file systems on the Internet. One of the efficient implementation of file system backup storage on cloud is Cumulus which works based on the assumption of thin cloud systems. Rather only storing backup in remote servers,

cumulus collects files from small interfaces also which in turn provides an efficient storage system in terms of portability and cost. The prototype was implemented and proved using Amazon S3 [9].

8. Secure data storage

In the user perspective, a cloud service provider should safeguard the privacy, integrity, availability of data deposited in the cloud systems.

DEPSKY (Dependable and Secure Storage in a Cloud-of-Clouds) system has efficiently implemented all these needs through data encryption, coding and duplication of data in distributed cloud system which made clouds to dump on each other. It achieves this by giving a virtual illusion to the large storage medium to the cloud users [10] [17].

9. Memory based cloud storage

To attain the memory based storage, a chip based cloud storage has been distributed over the network to handle the large scale storage need. To reach this goal, SPMCloud (A ScratchPad Memory based Storage Cloud Embedded on a chip Single Chip) was implemented which proved to be energy saving, scalable to multi-core, reduction in execution time, better archival capacity and increased cache hit ratio [11].

10. Cloud object storage

This is the latest storage techniques used by cloud systems. Object storage is a stateless and persistent storage, where the state of previous data access will not be stored. Because of this, object storage method is a secured storage method until now in cloud systems. For this purpose clouds use REST APIs which is responsible to handle the stateless storage mechanism [12].

III. CONCLUSION

Cloud storage is the biggest step taken by companies to move towards digitization of data. However cloud storage is not a new technique but day by day the storage efficiency gets improved because of the very large set of data generated by an individual and companies. Popular public providers like Amazon, Google allow an individual to safe guard the data on pay on use basis. Even the government of India has initiated its ambitious project called "GI Cloud" called "MeghRaj". This is public cloud service by NIC (National Informatics Center) allows each individual of the country to secure his official documents on cloud without charges up to 1GB. From this it can be concluded that cloud computing helps in accelerating delivery of e-services. This paper has highlighted some of the essential cloud storage techniques which are evolved in past nine years as surveyed from the state of the art research literature.

REFERENCES

- [1]. Anthony V. Velte, Toby J. Velte, Robert Elsenpeter, "Cloud Computing-A Practical Approach", McGraw-Hill Companies, in 2010.
- [2]. Rajkumar Buyya, James Broberg, Andrzej Goscinski, "Cloud Computing-Principles and Paradigms", John Wiley & Sons, in 2011.
- [3]. Abhishek Parakh, Subhash Kak, "Online data storage using implicit security", an Elsevier Article Published in May 2009.
- [4]. Hongwei Li, Yuanshun Dai, Ling Tian and Haomiao Yang, "Identity-Based Authentication for Cloud Computing", Springer-Verlag Berlin Heidelberg 2009.
- [5]. Hao Jin, Ke Zhou, Hong Jiang, "Dynamic and Public Auditing with Fair Arbitration for Cloud Data", an Article in IEEE Transactions on Cloud Computing, February 2016.
- [6]. Swapnali More, Sangita Chaudhari, "Third Party Public Auditing Scheme for Cloud Storage", an Article in Procedia Computer Science (Elsevier), December 2016.
- [7]. Bibing Hou and Feng Chen, "GDS-LC: A Latency and Cost Aware Client Caching scheme for Cloud Storage", ACM Transactions on Storage, November 2017.
- [8]. Chuan Qin, Jingwei Li and Patrick P.C. Lee, "The Design and Implementation of a Rekeying Aware Encrypted Deduplication Storage System", ACM Transactions on Storage, February 2017.
- [9]. Michael Vrable, Stephen Savage, Geoffrey M. Voelker, "Cumulus: File System Backup to Cloud", ACM Transactions on Storage, December 2009.
- [10]. Alysson Bessani Miguel Correia, Bruno Quaresma, Fernando Andre and Paulo Sousa, "DEPSKY: Dependable and Secure Storage in a Cloud-of-Clouds", ACM Transactions on Storage, December 2013.
- [11]. Luis Angel, D. Bathen and Nikhil D. Dutt, "SPM Cloud: Towards the Single-Chip Embedded ScratchPad Memory-Based Storage Cloud", ACM Transactions on Design Automation of electronic system, in June 2014.
- [12]. Marcus Branden Burger, Christian Cachin and Nikola Knezevic, "Don't trust the cloud, verify: Integrity and Consistency for object stores", ACM Transactions on privacy and security, July 2017.
- [13]. Hasan Mohamed Ibou, "OpenStack: The Cloud OS", A Technical Report, October 2013.
- [14]. Hui Tian, Yuxiang Chen, Chin-Chen Chang, Hong Jiang, Yongfeng Huang, Yonghong Chen, Jin Liu, "Dynamic-Hash-Table Based Public Auditing for Secure Cloud Storage", IEEE Transactions on Services Computing, Volume: 10, Issue: 5, PP: 701-714, Sept.-2017.
- [15]. Jia Yu, Kui Ren, Cong Wang, Vijay Varadharajan, "Enabling Cloud Storage Auditing With Key-Exposure Resistance", IEEE Transactions on Information Forensics and Security, Volume: 10, Issue: 6, PP: 1167-1179, June 2015.
- [16]. Bo Mao; Hong Jiang, Suzhen Wu, Lei Tian, "Leveraging Data Deduplication to Improve the Performance of Primary Storage Systems in the Cloud", IEEE Transactions on Computers, Volume: 65, Issue: 6, PP: 1775 – 1788, June, 2016.
- [17]. Lan Zhou, Vijay Varadharajan, K. Gopinath, "A Secure Role-Based Cloud Storage System For Encrypted Patient-Centric Health Records", The Computer Journal (Volume: 59, Issue: 11, Nov. 2016).