Published by :
http://www.ijert.org

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 01, January-2021

Cloud Storage in Context of Amazon Web Services

Mohd Tajammul
Deptt. of Computer Science
Jamia Millia Islamia
New Delhi, India

Iftikhar Aslam Tayubi
Deptt. of Computer Science
FCITR
King Abdul Aziz University, KSA

Rafat Parveen
Deptt. of Computer Science Jamia Millia Islamia
New Delhi, India

Sumaira Asif
Deptt. of Computer Science
Jamia Hamdard University
New Delhi, India

***Abstract*— Cloud computing is an advancement in the field of computing. It offers customers a robust platform like storage and computation power to use. Many of the organizations are using cloud storage to store their data onto it. Out of these organizations approximately 80% are using Amazon Web Services (AWS) for storage and (or) computation power both because it offers services at economic charges. Moreover, its security features are so updated that their penetration is very difficult. This paper focuses on cloud storage in context of AWS and also proposes an algorithm to authenticate user to protect data from unauthorized access.**

Keywords—Cloud storage; Cloud computing; AWS; Cloud security;

I. INTRODUCTION

This Cloud computing is the best way of using computation and storge from remote location. Nowadays, it has become prior demand of business corporate as it is 24X7 available and accessible from anywhere.

National Institute of Standard Technology (NIST) defines cloud computing as: "Cloud computing is a model for enabling ubiquitous, convenient, on demand network access to a shared pool of configurable computing resources (e.g., Network, Server, storage, database, computation, app etc.) that can be rapidly provisioned and released with minimal management effort or service provider interaction" [4-7].

A. Evolution of cloud computing

To understand any topic (theoretical or practical), or any technology or any computing model, it is very necessary to know the background of that and not only background but also to know how that particular thing was evolved refer **Fig.1**. Moreover, it is intelligence to show the evolution graphically (How the particular thing was developed and sustained in the market according to time). Alvin Toffler (04/10/1928 – 27/06/2016), an American futurist, a businessman and a successful writer in his book "The Third Wave" (Benton 1980) wrote that civilization has evolved in waves. There are many waves of growth of civilization but most important three of them are: Agricultural Scientist, Industrial Age, Information Age. Again the author revealed that each and every wave is composed of large number of sub-waves. We are at the starting of this post-industrial as well as information age. This beginning leads us to what people feel is directly an era of cloud computing. Similarly, Nicholas G. Carr (07/01/1959), an American writer, who has published many books on business, technologies and on culture. In one of his books "The Big Switch", he wrote that information revolution can be considered as similar to an important advancement in industrial era. Nicholas associates the growth of cloud computing in the information age to the growth of electrification in the industrial age. The author argued that it was said in starting that each and every organization needs to provide their own power but it was not so, because organizations were not providing their own power rather, they were just plug in to electrical power supply grid. The author argued that the cloud computing is really an initiation is the same advancement of Information Technology [4].

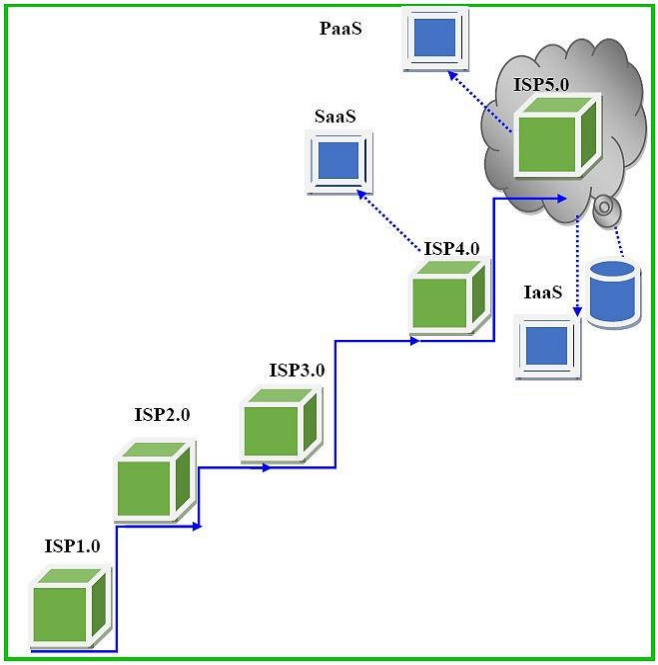

Fig. 1. Evolution of cloud computing

B. Types of Deployment model

Public cloud - Public cloud (refer **Fig.2.**) is available to all for storing data or performing computation. If data to be stored or to be computed is highly secure, it is highly recommended not to go for public cloud in this situation it is better to choose private cloud [4].

Private cloud - This is the cloud which is available only for personal storage or personal computation of a particular organization or particular individual. This type of cloud is recommended if data is highly sensitive or financial [4].

Community cloud - When some organizations have their own cloud for their use, this is an example of community

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 01, January-2021

cloud. This cloud is not owned by the community which uses its services rather its infrastructure is provided by the supplier [4].

Hybrid cloud - This cloud satisfies the customer with heterogeneous requirements. Utilization of the cloud depends upon the nature of data, if it is very sensitive store it on dedicated server and if it is less sensitive upload it on cloud this is recommended because in hybrid cloud there is environment of multitenancy and hence fear of data leakage [5-6].

Cloud computing is an online service provided to the user as per requirement and pay-per-use basis by arranging the available resources in best possible manner in between different users to fulfill their needs. Its dominants an important role in coming generation of Mobile Network and Services that is 5G and CPSC (Cyber Physical and Social Computing). Producing the data within the boundary of organization and storing it outside the boundary of organization (at cloud storage), drastically reduce the burden of storage [4].

Nevertheless security, privacy as well as the trust between both the ends of cloud become the main issue that leverages a great impact on the success of cloud computing and also produce a hurdle in the development of CPSC and 5G. Four main points are here to draw attention [4].

C. Types of Services models

Software-as-a-Service (S-a-a-S) - In this provision dedicated software is given to the user on which customer can perform his operation. For instance, someone wants to start his business and have no money to purchase costly software to run his business, in this situation he can go for S-a-a-S. For instance, NetSuite, SalesForce.com, Microsoft, IBM and Oracle [5].

Platform-as-a-Service (P-a-a-S) - In this provision a dedicated platform is given to the user on which he can develop his applications. Suppose someone wants to start a software company to develop software but have no platform on which to develop that application, in this situation he can go for P-a-a-S. Two famous examples are GAE, Microsoft's Azure [5].

Infrastructure-as-a-Service (I-a-a-S) - Under this provision, infrastructure is given to the user for use. The infrastructure may be of many types like storage like server. More elaborately let us consider an organization having a website with the capacity to handle 100 users at a time. To host these 100 customers, we need a server which can handle the request given by these users of the website and what happened if the server is very costly and out of the budget of the organization, in this case, the organization can go for renting a server from cloud and will pay for time the organization uses its services. Examples are GoGrid, Joyent, Flexiscale and Rackspace [5].

D. Challenges and issues of Cloud Computing [4]

- Issue of Resource Scheduling and Management.
- Issue of Security and Privacy.
- Issue of Power Consumption.
- Issue of Scalability and Elasticity.
- Issue of Portability and Interoperability.
- Issue of Reliability, Availability, and performance.

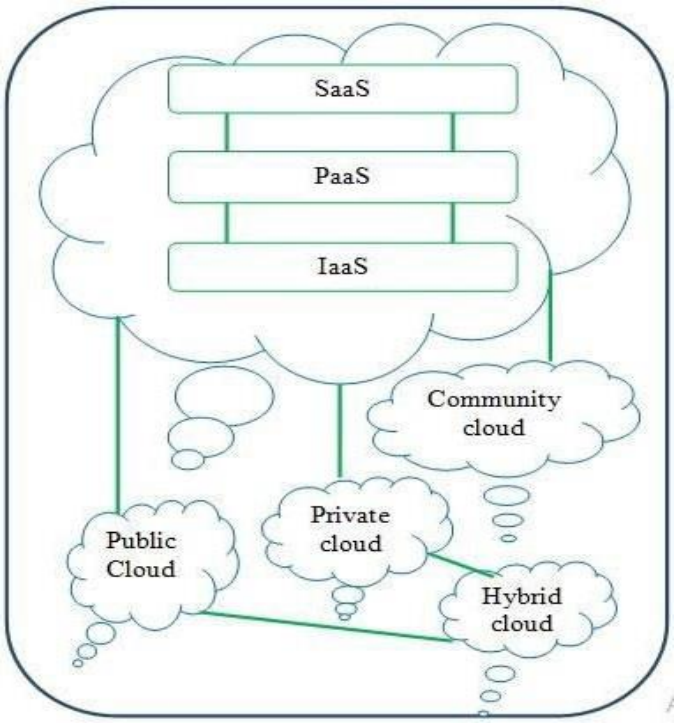

Fig. 2. Cloud computng models

E. Data centres

Thousands of servers are connected in cluster. Cluster of servers are called Data Center connecting all data centers make grid. There are a large number of data centers of AWS around the world and if you are anywhere then you may be maximum 1ooo Miles away from AWS data centers. In a data centers there are 300 to 500 nodes. Some popular data centers are as:

- Yahoo
- Google
- Amazon
- Microsoft MSN
- Facebook
- Twitter
- eBay

AWS data centers are most popular just because of offering variety of methods of charging the customers for example pay per use, region specific prices, term specific prices and monthly billing makes it more versatile [1-3].

Rest of the paper is organized as:

- Section II discusses the literature survey
- Section III discusses the research gaps and research questions
- Section IV discusses AWS storage
- Section V proposes an algorithm
- Section VI concludes the paper with future research directions

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 01, January-2021

II. LITERATRURE SURVEY

V. Chang, Y. Kuo, and M. Ramachandran [12], discussed Cloud Computing Adoption Framework (CCAF). This framework was capable of detecting and subsequently blocking of approximately 99.5 % of the viruses and Trojans. Moreover, the framework was also capable of blocking the SQL injection flaws. Beyond this, the framework was success in blocking of above 85% of hundred continuous attacks in 12 milliseconds. A. M. Talib et al. [13], discussed Multi Agent System (MAS). It was secured framework for cloud data storage based on five main components. Y. Li, K. Gai, L. Qiu, M. Qiu, and H. Zhao [14], discussed intelligent cryptography approach for secure distributed big data storage in cloud computing. This framework was capable of identifying the sensitive and normal data packets. On successful identification the system was dividing sensitive data packets into two parts and uploading them on two separate clouds while normal data packets were directed to upload on just one cloud. Tajammul M., Parveen R., [5-6], discussed big ten Information Security Management System Standards (ISMS) and their effect on cloud computing. Moreover, authors discussed best security standards. Tajammul M., Parveen R., [7], discussed security issues and challenges in cloud computing, authors discussed various methods to resolve those issues. Tajammul M., Parveen R., [8], discussed key generation algorithm coupled with Data Encryption Standards (DES) to secure cloud data storage. This method became successful in key generation and passing this key to DES to encrypt data to maintain confidentiality of the data uploaded on cloud storage. The framework was working in two phases, in first phase it was generating keys and passing to DES and in second phase it was encrypting data on the basis of key produced in first phase. Tajammul M., Parveen R., [9], discussed two pass multidimensional and key generation and encryption algorithm for data storge security to maintain data confidentiality on cloud storage. This framework was also working in two phases, in first phase it was generating keys and in second phase it was encrypting data on the basis of key produced in first phase by substitution and transposition and non-back tracking methods. Tajammul M., Parveen R., [10], discussed algorithm for document integrity testing pre- upload and post-download from Cloud Storage. This framework was designed to focus on data integrity of the stored data on cloud storage. This framework was designed to test the alteration of characters that can takes place in data stored on cloud storage. This framework was detecting the errors that were occurring in data by unauthorized users due to security breaches. Due to availability of large numbers of machines, cloud can be used to crack the security of any of the encryption-decryption algorithm by hit and trial or brute force methods. If any changes occur in such situations in users' data then how a particular user will come to know what changes were made in his or her data for the duration it was residing on cloud space. Keeping in mind the integrity testing framework was designed and developed. Tajammul M., Parveen R., [11], discussed auto encryption algorithm for uploading data on cloud storage. This framework was capable of automatically encrypting-decrypting of data to be uploaded-downloaded on or from cloud storage.

On the basis of literature reviewed, it can be stated that cloud storage is still in its early stage in terms of security because of news coming day to day in the market for the data leakage and security breaches. The research gaps and research question have been identified as:

III. RESEARCH GAPS AND RESEARCH QUESTIONS

On the basis of studying large number of papers on cloud computing and cloud storage it can be stated that it is lacking in terms of security. Many companies have been targeted and attacked in current century. There is great need to focus on security of stored data on the cloud. After authenticating users, the cloud should provide then One Time Password (OTP) to make some changes in their data or it should ask then some questions that the users know only to maintain integrity and confidentiality of the data. Some of the research question have been identified as:

Q1. *How to coupled integrity with confidentiality*

Q2. *How to offer semi-auditing capabilities to cloud user itself.*

Q3. *How to use the concept of two factor authentication.*

Q4. *How to minimize the phishing on cloud.*

IV. AMAZON WEB SERVICES (AWS) STORAGE

AWS offers various types of storage to the users on different charges, customers need to go through them and choose the best fit for them. They are suitable for different-different situations (refer **Fig.3.**).

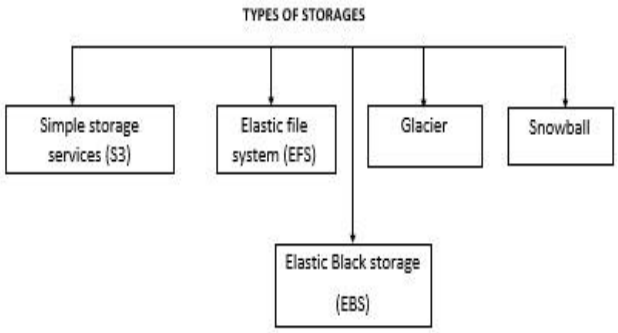

Fig. 3. AWS types of storages

AWS offers a complete Range of cloud Storage Services to support both application land archival compliance requirements (refer **Fig.4.**). Select from objects, file and Block Storage Services as well as Cloud data Migration options to start designing the foundation of your Cloud IT Environment [16-17].

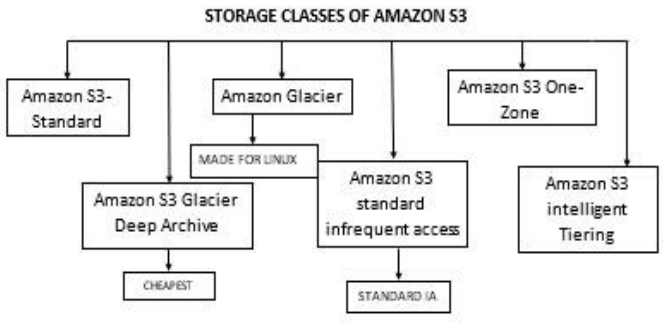

Fig. 4. Classes of AWS S3

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 01, January-2021

A. Amazon S3 Standard [16-17]

- S3 standard offers high durability, availability and performance object storage for frequently accessed data.
- Durability is 99.999999999
- Designed with 99.99% Availability
- Supports Secured Socket Layer (SSL) for data in transit & encryption of data at rest
- The storage cost for the object is fairly high but there is very less charge of accessing the objects.
- Largest object data, that that can be uploaded in a Single PUT is 5GB.

B. Amazon Simple Storage Serivice(S3)-Intelligent Access(IA) [16-17]

- S3-IA is for data that is accessed less frequently, but requires rapid access when needed.
- The storage cost is much cheaper on S3 – standard.
- Durability is 99.999999999
- Resident against events that impact an entire storage
- Availability is 99.9% in year.
- Support SSL for data in transit and encrypt of data at rest.
- Data that is deleted from S3-IA within 30 days will be charged for a full 30 days.

C. Amazon S3 Intelligent Tiering [16-17]

- The S3 Intelligent Tiering Storage class is designed to optimize cost by automatically moving data
- It works by storing objects in two access tiers
- If an object in the infrequent access tiers is accessed, it is automatically moved back to frequent access tier
- There are the retrieval fees when using the S3-Intelligent
- Object less than 128KB cannot move to IA
- Durability is 99.999999999%
- Availability is 99.99%.

D. Amazon One-Zone IA [16-17]

- Data store in single zone
- Ideal for those who want lower cost option of IA-data
- It is good choice for storing secondary backup copies of re-creatable data.
- You can use S3 lifecycle policies
- Durability is 99.999999999%
- Availability is 99.99%
- Because S3 one-zone stores data in single zone, data stored in this storage class will be lost in event of zone destruction.

E. Amazon S3 Glacier [16-17]

- To keep cost low yet suitable for varying needs, S3 Glaciers provides three retrieval options that range from few minutes to few hours
- You can upload objects directly to Glacier or lifecycle policies
- Durability is 99.999999999%
- Data is resilient in the event of an entire zone destruction
- Support SSL for data in transit & encryption data at rest.
- You can retrieve 10GB of your amazon S3 Glacier data per month for free with free tier account

F. Amazon S3 Glacier Deep Archive [16-17]

- Design to retain data for long period. For example, 10 years
- Durability is 99.999999999%
- Ideal alternate to magnetic tape libraries
- Retrieval time within 12 hrs
- Storage cost is up to 75% less than for existing S3-Glaciers storage class
- Availability is 99.9%

Except above storages, two more needs to be focused here: Block storage and object storage.

G. Block Storage [1-3, 16-17]

- Block storage is suitable for transactional database, random read/ write load and structured database storage. Parallel processing is needed.
- Block storage decides the data to be stored in every sized block (data chunks) for instance, a file can be split into every sized block before it is stored. Equally divided into blocks for retrieving the data exactly we design in block storage
- Data block stored in block storage would not contain metadata (data created, data modified, data content types)
- Block storage only keeps the address where the data blocks are stored, it does not care what is in that block just how to retrieve it when require. For example, block storage is EBS.
- Block storage is accessed in instance and through that only it can be access.

H. Object storage [1-3, 16-17]

- In this all the file as whose will be uploaded irrespective if size, as it is 5GB file or 5MB file
- If we divide the parts of object then it will lose its identity. In this case object storage is useful.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 10 Issue 01, January-2021

- In whatever the forms of data, it will be kept as it is, no block will be formed
- Global unique ID must be in AWS, that's why, it's easy to recover from anywhere as it how its own identifier and as a global unique ID
- Object storage can access through internet also
- Objects storage store the file as a whole and does not divide them
- In object storage on object is the file/data itself, Its meta data, Objects global unique ID
- The object global unique ID is a unique identifier the object (can be object name itself and it must be unique such that it can be retrieved disregarding where its physical storage location is.
- Objects storage solution- Dropbox, AWS, S3
- Object storage can be accessed through http or https

V. PROPOSED ALGORITHM

The algorithm will cope up with the existing situation and will try to keep data secure. Moreover, it will reduce the cloud storage ongoing phishing problem.

Input: Data sets

Output: Achieving confidentiality on data sets and mitigating phishing on storage

Step1: *Take data sets into consideration*

Step2: *Hide users' identities and some questions related to genuine users in the data*

Step3: *Ask prestored question to the users trying to manipulate data like one time transaction password implemented in the bank to minimize phishing*

Step4: *Integrate the concept of confidentiality and integrity together to achieve better security results*

Step5: *Check post testing of the data to know if any unauthorized changes occur*

Step6: *Finish*

VI. CONCLUSION AND FUTURE RESEARCH DIRECTIONS

This paper discusses cloud computing and specially cloud storage by taking a particular vender into consideration that is Amazon Web Services (AWS). The paper focused on various types of storages under Simple Storage Service (S3). The paper focuses that the variety of the available storages suits to different situations and based on the situation a customer can go for subscription of the particular type of storage.

Now, the **future research direction or future scope** motivate the researchers to implement the algorithm and compute its efficiency in real environment given under Section V of the paper.

REFERENCES

[1] P. Rajesh, M, Shamresh, and P, Prashant, "Study on Cost Estimation Service Delivery in Cloud Computing Environment", International Journal of Information and Computation Technology, vol. 4, pp299–308, 2014.

[2] K. Cho and H. Bahn, "A Cost Estimation Model for Cloud Services and Applying to PC Laboratory Platforms," pp. 1–13, 2020.

[3] "How AWS Pricing Works," vol. 2018, no. June, 2018. https://d0.awsstatic.com/whitepapers/aws_pricing_overview.pdf

[4] Mohd. Tajammul, Rafat Parveen 2019: Cloud Computing – Introduction to Innovation, pp 188, New Delhi: International Research Publication House.

[5] Tajammul M., Parveen R., "Comparative Study of Big Ten Information Security Management System Standards," International Journal of Engineering Research in Computer Science and Engineering, vol. 5, no. 2, pp. 5–14, 2018.

[6] Tajammul M., Parveen R., "Comparative analysis of big ten ISMS standards and their effect on cloud computing," 2017 International Conference on Computing and Communication Technologies for Smart Nation (IC3TSN), Gurgaon, 2017, pp. 362-367.

[7] Tajammul M., Parveen R., and Shahnawaz M., "Cloud Computing Security Issues and Methods to Resolve : Review," Journal of Basic and Applied Engineering Research, ISSN: 2350-0077 vol. 5, no. 7, pp. 545–550, 2018.

[8] Tajammul M., Parveen R., "Key Generation Algorithm Coupled With DES for Securing Cloud Storage," International Journal of Engineering and Advanced Technology, vol. 8 no. 5, pp. 1452–1458, 2019.

[9] Tajammul M., Parveen R., "Two Pass Multidimensional Key Generation and Encryption Algorithm for Data Storage Security in Cloud Computing", International Journal of Recent Technology in Engineering, no. 2, pp. 4152– 4158, 2019.

[10] Tajammul M., Parveen R., "Algorithm for Document Integrity Testing Pre Upload and Post Download from Cloud Storage", International Journal of Recent Technology in Engineering, no. 2, pp. 973–979, 2019.

[11] Tajammul M., Parveen R., Auto encryption algorithm for uploading data on cloud storage. Int. j. inf. tecnol. (2020). https://doi.org/10.1007/s41870-020- 00441-9 [

[12] 28]P.-C. Chao and H.-M. Sun, "Multi-agent-b

[13] V. Chang, Y. Kuo, and M. Ramachandran, "Cloud computing adoption framework : A security framework for business clouds," vol. 57, pp. 24–41, 2016.

[14] A. M. Talib, R. Atan, R. Abdullah, M. Azrifah, and A. Murad, "Security Framework of Cloud Data Storage Based on Multi Agent System Architecture : Semantic Literature Review," vol. 3, no. 4, pp. 175–186, 2010.

[15] Y. Li, K. Gai, L. Qiu, M. Qiu, and H. Zhao, "Intelligent cryptography approach for secure distributed big data storage in cloud computing," Inf. Sci. (Ny)., vol. 387, pp. 103–115, 2017.

[16] Cloud Object Storage | Store & Retrieve Data Anywhere | Amazon Simple Storage Service (S3)

[17] What is Amazon Simple Storage Service (Amazon S3)? (techtarget.com)