

Cloud Computing Environment for Secure Data Storage – An Overview

Anupama P.L.¹, Bismi M²

PG Scholar, Computer Science & Engineering, LBSITW, Thiruvananthapuram, India^{1,2}

Abstract: Cloud computing is an innovative computing paradigm that offers amazing advantages in economical aspects such as reducing time, improved performance, flexible computing capabilities, and advantageous computing power. But the security of the data in cloud is under risk. We can explore the cloud computing for so many applications as an individual or as a group work. This paper gives the basics of cloud computing with its history, architecture, features along with pros and cons. The paper also gives two ideas for secure data storage in cloud environment. It helps to store data without confidentiality leakage and also helps to retrieve only accurate files.

Keywords: Cloud Computing, Security, Storage, User data.

I. INTRODUCTION

Cloud computing is an evolving term that describes the development of many existing technologies and approaches to computing into something different. The NIST defined cloud computing as “Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort of service provider interaction”[1].

In the traditional desktop computing, we run copies of software programs in our system. The documents that we created are stored in the same system. i.e. PC centric. With cloud computing, the software programs are not run from our computer rather stored on servers and accessed via the Internet. i.e., service centric. The advantage is that if our system crashes, the software is still available. The same thing will happen in the case of the data that stored in the cloud. Anyone with permission can access and modify the cloud data. “Cloud” consists of huge number of computers and servers, linked and accessible via internet [2].

II. HISTORY

A. Client/Server Computing

Servers are the huge mainframe computers where the software applications, data, and all the control reside. Users connected to the server via a computer terminal, called a workstation or client. It is similar to a centralized storage and applications. i.e., all the control rested in the server. There is a bottleneck problem may arise on server side.

B. Peer-to-Peer Computing

Peer-to-Peer computing is defines as the network architecture in which each computer has equal capabilities and responsibilities. i.e., more than one computer acts as servers.

C. Distributed Computing

Distributed computing [3] uses the resources that most of the PC does not used. Here more computing power is

utilized. The software is installed in the remote machine and processes the activities during idle time of PC. This sparse time processing is periodically uploaded to the network of distributed computing and concatenate for the results. Here multiple computers are networked together. First internet based distributed computing project was distributed.net.

D. Collaborating Computing

Collaborative means working as a group. i.e., multiple users work simultaneously on the same computer based project. It support multiple user video conferencing, file sharing etc.

E. Grid Computing

Grid computing [4] combines more computer resources from various regions to get a main objective. Here different computers in the same network can work together for a single job. Thus act like a large supercomputing type resources.

F. Utility Computing

Utility computing is a pay- per-use service model. According to the customer’s requirements the service provider makes the computer resources available.

III. ARCHITECTURE

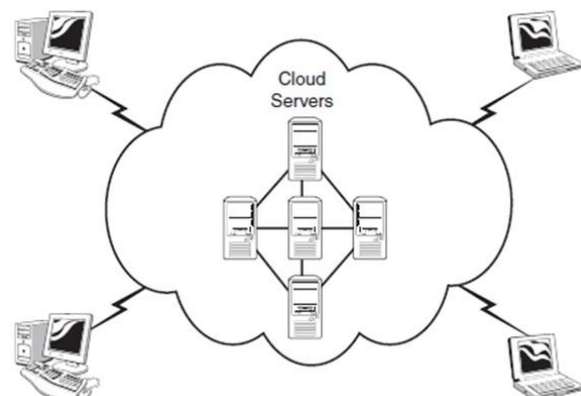


Figure 1: Cloud Computing Architecture

Cloud Computing is the most popular one to store the data in geographical environment with infinite computing resources and access the data anytime and anywhere. Hence it provides greater availability, scalability, and reliability to the users. The following Figure 1 shows the architecture of cloud computing.

III.A.FEATURES AND CHARACTERISTICS

The following are the characteristics of cloud computing. [5]



Figure 2: Characteristics of Cloud Computing

- **Broad Network Access:** - The capabilities are available over the network and accessed through standard mechanisms (mobile phones, laptops and PDAs) as well as other traditional or cloud based software services.
- **On-demand Self Service:** - A consumer can unilaterally provision computing capabilities as needed automatically.
- **Measured Service:** - Cloud computing resources are provided to users on a pay-per-use model. Resource usage can be monitored, controlled and reported. It provides transparency for both the provider and consumer of the service.
- **Rapid Elasticity:** - Cloud resources can be rapidly scaled up or down based on demand. Capabilities are rapidly and elastically provisioned. These capabilities appear to be unlimited and can be purchased in any quantity at any time.
- **Resource Pooling:** - The computing and storage resources provided by cloud service providers are pooled to serve multiple users using multi-tenancy.
- **Multi-tenancy:** - It is the feature of being capable of running multiple instances under the same shared platform. Each instance can be accessed by one or more users, called tenants, while sharing a common platform.

III.B.SERVICE MODELS

Following are the services (SPI) provided by cloud [5]

A. Software as a Service (SaaS)

This provides a service to the user by offering different software to the different user over internet. A distinct instance of service which runs in the cloud, here one or more user can utilize the service. E.g., Gmail

B. Platform as a Service (PaaS)

It provides a storage mechanism for the various applications and consumptions. User can have an independency to build their personal applications that provides infrastructure for the user. E.g., Google app engine.

C. Infrastructure as a Service (IaaS)

In this mechanism cloud computing provides a service over the internet, hardware and software in data centers as

a services. The data centre of hardware and software is called as Cloud. E.g., Amazon EC2.

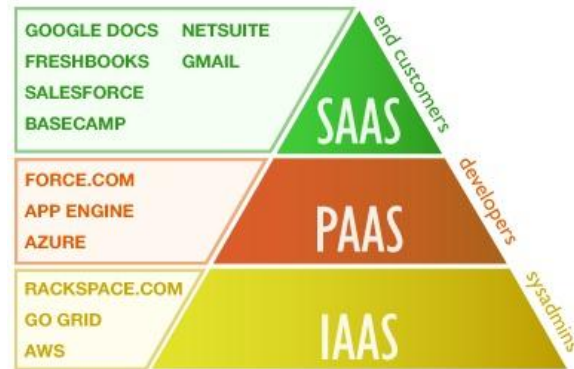


Figure 3: Service Models

III.C. DEPLOYMENT MODELS

Cloud deployment models [6] [7] are shown in figure 4

A. Public Cloud

All the resources of the cloud are publically available to the users over internet. It is less secure than the other cloud models, because all the data is directly placed in the cloud. A public cloud is basically the Internet. Service providers use the Internet to make resources, such as applications (also known as Software-as-a-service) and storage, available to the general public, or on a public cloud. Examples of public clouds include Amazon Elastic Compute Cloud (EC2), IBM's Blue Cloud, Sun Cloud, Google App Engine and Windows Azure Services Platform.

For users, these types of clouds will provide the best economies of scale, are inexpensive to set-up because hardware, application and bandwidth costs are covered by the provider. It's a pay-per-usage model and the only costs incurred are based on the capacity that is used. There are some limitations, however; the public cloud may not be the right fit for every organization. The model can limit configuration, security, and SLA specificity, making it less-than-ideal for services using sensitive data that is subject to compliance regulations.

B. Private Cloud

Due to security and confidential problems in public cloud, private cloud came into picture. Private cloud is set within an organization i.e. resources can be used only by that particular organization. Private clouds are data center architectures owned by a single company that provides flexibility, scalability, provisioning and automation and monitoring. The goal of a private cloud is not sell "as-a-service" offerings to external customers but instead to gain the benefits of cloud architecture without giving up the control of maintaining your own data center.

Private clouds can be expensive with typically modest economies of scale. This is usually not an option for the average Small-to-Medium sized business and is most typically put to use by large enterprises. Private clouds are driven by concerns around security and compliance, and keeping assets within the firewall.

C. Hybrid Cloud

It is combination of public and private cloud. An organization can have a private cloud in a public cloud, is a best example of hybrid cloud.

By using a Hybrid approach, companies can maintain control of an internally managed private cloud while relying on the public cloud as needed. For instance during peak periods individual applications, or portions of applications can be migrated to the Public Cloud. This will also be beneficial during predictable outages: hurricane warnings, scheduled maintenance windows, rolling brown/blackouts.

The ability to maintain an off-premise disaster recovery site for most organizations is impossible due to cost. While there are lower cost solutions and alternatives the lower down the spectrum an organization gets, the capability to recover data quickly reduces. Cloud based Disaster Recovery (DR)/ Business Continuity (BC) services allow organizations to contract failover out to a Managed Services Provider that maintains multi-tenant infrastructure for DR/BC, and specializes in getting business back online quickly.

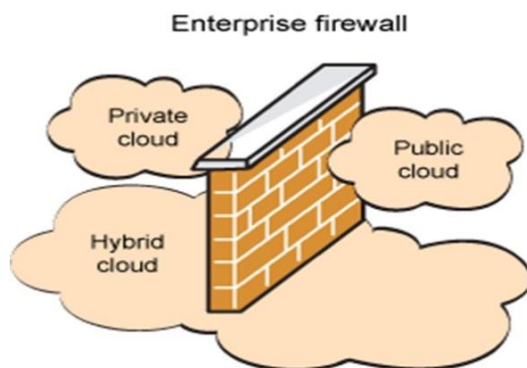


Figure 4: Cloud Deployment Models

D. Community Cloud

The group of organization of same community or group owns a common cloud. This cloud is maintained by the third trusted party.

III. D. PROS AND CONS

Here discussed about the advantages and disadvantages of cloud computing [8].

A. Advantages

- Lower Software Cost: - User does not need to purchase separate software packages for separate computers in the organization.
- Instant Software Updates: - The software updates are done automatically.
- Unlimited Storage Capacity: - Cloud offers limitless storage capacity for their users.
- Increased Computing Power: - User gets the power of the entire cloud.
- Easier Group Collaboration: - Multiple users can easily collaborate on their files and documents.
- Improved Performance: - The users in the cloud computing boot up faster and run faster.

- Available the Latest Version: - The cloud always hosts the latest version of the documents, not fear about the outdated version.

B. Disadvantages

4.2.1 Can be Slow

Because the traffic of the network the application in cloud system can be slower than desktop applications

4.2.2 Security

The users have worry about the security and privacy problems from the cloud provider. In some cases cloud provider provides a security to the front-end resources only and failed to provide a security to the backend resources, so the attackers may hack the data easily from the back-end resources. Hence malicious user could compromise the data integrity and confidentiality. Leakage details of data might be in the users cloud resources and the cloud provider is the responsible for this. Thus users commonly think of using encryption for preventing unauthorized access of data.

The major security issues [9] are shown in Figure 5.

- **Integrity:** - It makes sure that data held in a system is in proper representation and not be modified by unauthorized person.
- **Availability:-** The data processing resources are unavailable for malicious purpose.
- **Confidentiality:** - The data is not disclosed to unauthorized person and secure data access and transfer.



Figure 5: Cloud Security Issues

A. Confidentiality in Cloud Environment

Today, encryption of personal and enterprise data is strongly recommended for security. Cloud environments are shared with many tenants. Thus confidential data hosted in a cloud must be protected using a combination of access control policies and encryption. Encrypting data on disk or in a live production database has value, as it can protect against a malicious cloud service provider or a malicious co-tenant as well as against some types of application abuse. For long-term archival storage, some customers encrypt their own data and then send it as cipher text to a cloud data storage vendor. The customer then controls and holds the cryptographic keys and decrypts the data, if necessary, back on their own premises.

Loss of cryptographic keys may cause a serious issue. [10] introduce a new scheme that combines encryption of data with secure file access mechanisms and also describes

multiple decryption keys for supporting confidentiality of user data. [10] also claim that it reduces the internal information leakage shown in figure 6.

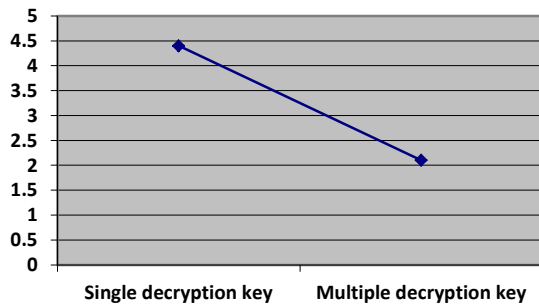


Figure 6: Comparison of Information Leakage with single and multiple keys

Encrypting data is common within IaaS environments, using a variety of provider and third party tools. Encrypting data within PaaS environments is generally more complex, requiring instrumentation of provider offerings or special customization. Encrypting data within SaaS environments is a feature cloud customers cannot implement directly, and need to request from their providers.

B. Multiple Keyword Ranked Search (MRSE)

The cloud provider encrypts the confidential data and stores it in the cloud so that only authenticated users can access the data. Keyword search is the commonly used method for efficient and effective search service over encrypted outsourced data. Searchable symmetric encryption methods are used to search matching keywords in the files. Existing works support only single or Boolean keywords in the search request; which gives results with low efficiency. So a new scheme is proposed which allows multiple keywords in search request along with synonym based search[11]. Co-ordinate matching is the similarity measure used to find relevance of a document to the query keyword. Resulting files can be retrieved according to total relevance score. Proposed scheme will give more accurate results than that of existing systems. But the time taken will be more in the MRSE scheme. Graphs showing time taken and accuracy in new and existing systems are shown below.

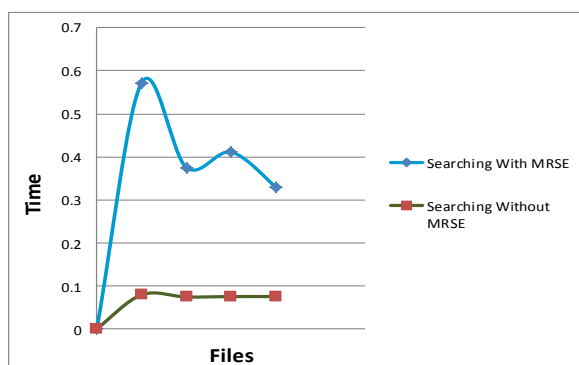


Figure 7: Comparison of time taken for file retrieval in two cases.

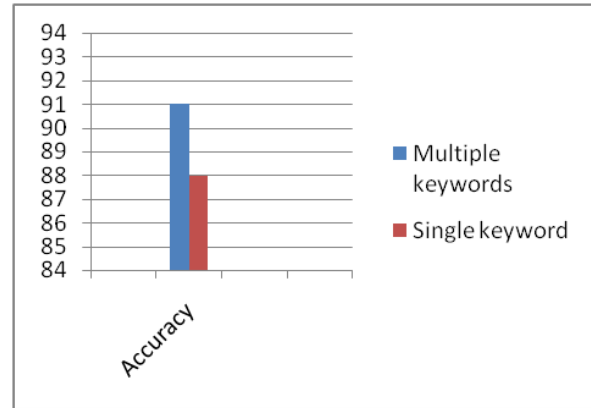


Figure 8: Accuracy obtained while retrieving files with single and multiple keywords.

IV. COMPANIES IN THE CLOUD

Some companies offered cloud services are [12]

A. Google: - Google offers a powerful collection of web based applications served through cloud architecture. Cloud base word processing – Google Docs, presentation software – Google Presentation, email – Gmail, Calendar – Google Calendar etc.

B. Microsoft: - It offers Windows Live suite of web based applications.

C. Amazon: - It has its Elastic Compute Cloud (EC2), a web service which provides cloud based resizable computing capacity for application developers.

D. IBM: - It provides open source workload scheduling software called Hadoop, based on the MapReduce software used by Google in its offering.

V. APPLICATIONS

Following are some applications [12] using the features of cloud computing.

A. For the Family

Computing in the cloud can help a family to communicate and collaborate. It brings the family members closer together.

- Centralized Email Communication (Gmail, Yahoo etc.)
- Collaborating with Schedules (Google Calendar Yahoo Calendar etc.)
- Collaborating on Grocery List (Google Docs)
- Collaborating on To-Do Lists (Ta-da List)
- Collaborating on Contact List (My Events)
- Sharing Family Photos

B. For the Community

Any time any group of people in the community can communicate and collaborate.

- Collaborating on Schedules – Sports team schedules, School Schedules, Event Schedules etc.
- Collaborating on Group projects and Events
- Collaborating on Budgets
- Collaborating on Event Marketing

C. For the Corporation

The cost savings and productivity enhancement are done through the cloud. Companies can do more with limited budgets.

- Managing Schedules (AppointmentQuest)
- Managing Contact List (BigContacts)
- Managing Projects (AceProjects)
- Collaborating on Reports (Google Docs)
- Collaborating on Marketing Materials
- Collaborating on Presentations
- Presentation on the Road (WebEx)
- Collaborating on Budgets (HostBudgets)

VI. CLOUD STORAGE

Cloud storage [13] is a model of data storage where the digital data is stored in logical pools, the physical storage spans multiple servers (and often locations), and the physical environment is typically owned and managed by a hosting company. These cloud storage providers are responsible for keeping the data available and accessible. People and organizations buy or lease storage capacity from the providers to store user, organization, or application data. Cloud storage is based on highly virtualized infrastructure and is like broader cloud computing in terms of accessible interfaces, near-instant elasticity and scalability, multi-tenancy, and metered resources.



Figure 9: Cloud Storage

Cloud storage is made up of many distributed resources, but acts as one – often referred to as federated storage clouds. It is highly fault tolerant through redundancy and distribution of data. It is efficient and cost effective. i.e., through cloud storage, companies need only pay for the storage and service they actually use. Cloud storage provides users with immediate access to a broad range of resources and applications hosted in the infrastructure of another organization via a web service interface. The service provider handles the storage capacity so the user need not worry about the capacity and capability of storage. It also provides mechanism for creating, accessing and updating the outsourced data. The examples of cloud storage are Amazon S3, Microsoft Azure, etc.

In cloud environment, the sensitive data of data owners are stored in the cloud storage and can be accessed from anywhere, everywhere and at any time. To protect data privacy, some cryptographic techniques like encryption can be introduced [14]. Thus the sensitive data is encrypted before uploading to the cloud storage. Mass

storage and low expense provided by the cloud storage invites more and more enterprises and organizations to store their private data in cloud with effective security. A virtual private storage service is designed by taking the advantages of both public and private clouds. A public cloud provides scalable and dynamic storage and provides availability and reliability of data where as private clouds provide security and privacy for the data. Thus the virtual private storage service based on cryptographic techniques achieves both the security of a private cloud and functionality and cost savings of a public cloud. Other advantages of cryptographic cloud storage are the control of data is in the hands of customer and security properties are taken from cryptography.

VII. CONCLUSION

Cloud computing is an innovative technology to store and share the data publicly with the assurance of data security. The various features and characteristics of cloud described above have proved its significance. Also cloud has got numerous applications in every point of human life. However, this new computing paradigm exhibits serious privacy challenges on users' data stored on remote servers which belong to a different trust domain. Here comes the need for data encryption in cloud. i.e., for data security, cryptographic techniques are used and thus the cloud storage is modified as cryptographic cloud storage. Two schemes for secure data storage in cloud premises is also mentioned along with graphs. The awesome applications of cloud inspire more and more companies to offer cloud facility in future.

ACKNOWLEDGMENT

We are extremely thankful to our guides Smt. Sumithra M.D, Mr. Anver S.R and Mr. Sulphikar A, Associate Professors of Computer Science and Engineering, LBSITW for their valuable guidance and timely suggestions for this paper work and also to all the authors of the papers which were referred, to get enough information to write this paper.

REFERENCES

- [1] P. Mell and T. Grance, "Draft NIST working definition of cloud computing", Online at <http://csrc.nist.gov/groups/SNS/cloud-computing/index.html>, Referenced on 3rd, 2009
- [2] https://en.m.wikipedia.org/wiki/Cloud_computing
- [3] <http://study.com/academy/lesson/distributed-computing-principles-environments.html>
- [4] <https://www.techopedia.com/definition/87/grid-computing>
- [5] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Zonwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "Above the clouds: A Berkeley view of cloud computing," University of California, Berkeley, Tech. Rep. UCBECS-2009-28, Feb 2009.
- [6] Mohiuddin Ahmed, Abu Sina Md. Raju Chowdhury, Mustaq Ahmed, "An Advanced Survey on Cloud Computing and State-of-the-art Research Issues," International Journal of Computer Science Issues, Vol.9, No.1 pp. 201-207, 2012.
- [7] <http://www.asigra.com/blog/cloud-types-private-public-and-hybrid>
- [8] <http://www.levelcloud.net/why-levelcloud/cloud-education-center/advantages-and-disadvantages-of-cloud-computing>
- [9] Kuyoro S.O, Ibikunle F, Awodele O, "Cloud Computing Security Issues and Challenges," International Journal on Computer Networks, pp. 478-481, 2010.



- [10] Anupama P.L., Sumithra M.D., Anver S.R., “Supporting Confidentiality in Cloud Databases and Role Based Access”, International Journal For Advance Research In Engineering And Technology, Volume 3, Issue IX, Sep 2015 ISSN 2320-6802
- [11] Bismi M, Sulphikar A, “Improving Accuracy with Ontology based Secured Search in Encrypted Cloud Data”, International Journal For Advance Research In Engineering And Technology, Volume 3, Issue IX, Sep 2015 ISSN 2320-6802
- [12] Michael Miller, “Cloud Computing Web-Based Applications That Change the Way You Work and Collaborate Online”, e-book.
- [13] Cong Wang, KuiRen, Jin Li, “Toward publicly Auditable Secure cloud data storage Service,” IEEE Network, August 2010.
- [14] S. Kamara and K. Lauter, “Cryptographic cloud storage”, in RLCPS, January 2010, LNCS, Springer, Heidelberg.

BIOGRAPHIES



Anupama P.L. has received her B-Tech degree from CUSAT University (MG College of Engineering, Trivandrum) in 2012 and now she is pursuing her M-Tech degree from University of Kerala (LBS Institute of Technology for Women, Trivandrum). She has published

two international journals. Her research interests include Cloud computing and Grid computing.



Bismi M. has received her B-Tech degree from University of Kerala (LBS Institute of Technology for Women, Trivandrum) in 2012 and now she is pursuing her M-Tech degree also from University of Kerala (LBS Institute of Technology for Women, Trivandrum).

She has published two international journals. Her research interests include Data mining and Cloud computing.