# Towards Secure and Dependable Storage Services in Cloud Computing

Sagar B Patil,
*Computer Science*
*Shivaji University*

Pooja A Vhatkar,
*Computer Science*
*Shivaji University*

Jaya Gajwani,
*Computer Science,*
*Shivaji University*

***Abstract**: A cloud storage system, consisting of a collection of storage servers, provides long-term storage services over the Internet. Storing data in a third party's cloud system causes serious concern over data confidentiality. In this paper, we present a secure private cloud for cloud services. We deal with user anonymous access to cloud services and shared storage servers. Our solution offers anonymous authentication. This means that users' personal attributes (personal details, social details, valid registration) can be proven without revealing users' identity. Thus, users can use services without any threat of profiling their behavior. We analyze current privacy preserving solutions for cloud services and outline our solution based on advanced encryption cryptographic components. Data loss is another concerning issue in cloud computing. Our solutions to this is providing data backup and restore facility for the users in private cloud. This paper tries to address challenges towards private cloud. Our method fully integrates data uploading, encrypting, data backup and restore.*

***Keywords**: Cloud Computing, Authentication, Encryption, Cryptography, Security, Secure Storage System, Data Backup and Restore*

## I. Introduction

The term Cloud refers to a Network or Internet. In other words, we can say that Cloud is something, which is present at remote location. Cloud can provide services over network, i.e. on public networks or on private network. Cloud Computing has begun to emerge as a hotspot in both industry and academia; It represents a new business model and computing paradigm, which enables on demand provisioning of computational and storage resources. Cloud computing is the delivery of computing services over the Internet. Cloud services allow individuals and businesses to use software and hardware that are managed by third parties at remote locations. The cloud computing model allows access to information and computer resources from anywhere that a network connection is available. Cloud computing provides a shared pool of resources, including data storage space, networks, computer processing power, and specialized corporate and user applications.

### *A .Cloud Characteristics and Security Challenges*

The Cloud Security Alliance has summarized five essential characteristics [4] that illustrate the relation to, and differences from, traditional computing paradigm.

- **On-demand self-service** – A cloud customer may unilaterally obtain computing capabilities, like the usage of various servers and network storage, as on demand, without interacting with the cloud provider.
- **Broad network access** – Services are delivered across the Internet via a standard mechanism that allows customers to access the services through heterogeneous thin or thick client tools (e.g., PCs, mobile phones, and PDAs).
- **Resource pooling** – The cloud provider employs a multitenant model to serve multiple customers by pooling computing resources, which are different physical and virtual resources dynamically assigned or reassigned. According to customer demand. Examples of resources include storage, processing, memory, network bandwidth, and virtual machines.
- **Rapid elasticity** – Capabilities may be rapidly and elastically provisioned in order to quickly scale out or rapidly released to quickly scale in. From customers' point of view, the available capabilities should appear to be unlimited and have the ability to be purchased in any quantity at any time.
- **Measured service** – The service purchased by customers can be quantified and measured. For both the provider and customers, resource usage will be monitored, controlled, metered, and reported.

### *B. Motivation*

The cloud acts as a big black box, nothing inside the cloud is visible to the clients. Clients have no idea or control over what happens inside a cloud. Even if the cloud provider is honest, it can have malicious system admin who can tamper with the VMs and violate confidentiality and integrity. Clouds are still subject to traditional data confidentiality, integrity, availability, and privacy issues, plus some additional attacks. Providers generally assert that they are not responsible for the impacts of security breaches or for security in general, i.e. unauthorized modification or disclosure of consumer data, or service interruptions caused by malicious activity. To overcome this security issues we will be providing encryption to the data by using more secure algorithm i.e. MD5/SHA [1]. In July 2008, Amazon's S3cloud storage service went down for the second time that year. A lot of applications were hosted by company and all those servers could not be accessed until techs could fix the problem.

Some applications were down for 8 hours. If you have sensitive or proprietary information, your IT security group may simply mandate that you not store it on someone else machine. In public cloud you can create your own back up cloud program using Amazon web service and S3 tools command line program then going to create script that sync files to our Amazon S3 buckets allowing us to create our own backup cloud program. This solution is implemented in public cloud, hence we are trying to implement data backup and restore facility in private cloud.

The aim of this paper is to present the current state of security in the cloud computing, and highlight possible security issues and vulnerabilities associated with cloud infrastructures. Research is based on an overview of existing literature based on security issues and data loss in Private Cloud. In this paper, our main focus is to investigate the safety mechanisms in the open source cloud-platform OwnCloud. The paper is organized as follows. In Section 2 a brief review of cloud computing architectures is made, including essential characteristics, deployment and service models. General cloud computing security issues are discussed in Section 3, while Section 4 focuses on features of OwnCloud, where section 5 focuses on algorithm and data backup and restore, Conclusions and future work are presented in Section 6.

## II. Cloud Architecture

Cloud Computing architecture consists of two components "the front end " and "the back end ".The front end of the cloud computing system comprises the client's device(or it may be computer network) and some applications are needed for accessing the cloud computing system. Back end refers to the cloud itself which may encompass various computer machines, data storage systems and servers. Group of these clouds make a whole cloud computing system. The whole system is administered via a central server that is also used for monitoring client is demand and traffic ensuring smooth functioning of the system. A special type of software called "Middleware" is used to allow computers that are connected on the network to communicate with each other .Cloud Computing systems also must have a copy of all its clients' data to restore the service which may arise due to a device breakdown. Making copy of data is called redundancy and cloud computing service providers provide data redundancy.

### *A. Cloud computing infrastructure models*

There are many considerations for cloud computing architects to make when moving from a standard enterprise application deployment model to one based on cloud computing. There are public and private clouds that offer complementary benefits, there are three basic service models to consider, and there is the value of open APIs versus proprietary ones. Deployment models define the type of access to the cloud. Cloud can have any of the four types of access:
Public Cloud, Private Cloud, Community Cloud, Hybrid Cloud.

*1) Public clouds*: Public clouds are run by third parties, and applications from different customers are likely to be mixed together on the cloud's servers, storage systems, and networks (Figure 1). Public clouds are most often hosted away from customer premises, and they provide a way to reduce customer risk and cost by providing a flexible, even temporary extension to enterprise infrastructure. If a public cloud is implemented with performance, security, and data locality in mind, the existence of other applications running in the cloud should be transparent to both cloud architects and end users. Indeed, one of the benefits of public clouds is that they can be much larger than a company's private cloud might be, offering the ability to scale up and down on demand, and shifting infrastructure risks from the enterprise to the cloud provider, if even just temporarily

Figure 1. A public cloud provides services to multiple customers, and is typically deployed at a colocation facility.

Portions of a public cloud can be carved out for the exclusive use of a single client, creating a virtual private datacenter. Rather than being limited to deploying virtual machine images in a public cloud, a virtual private datacenter gives customers greater visibility into its infrastructure [5]. Now customers can manipulate not just virtual machine images, but also servers, storage systems, network devices, and Network topology. Creating a virtual private datacenter with all components located in the same facility helps to lessen the issue of data locality because bandwidth is abundant and typically free when connecting resources within the same facility.

*2) Private clouds:* Private clouds are built for the exclusive use of one client, providing the utmost control over data, security, and quality of service (Figure 2). The company owns the infrastructure and has control over how applications are deployed on it. Private clouds may be deployed in an enterprise datacenter, and they also may be deployed at a colocation facility. Private clouds can be built and managed by a company's own IT organization or by a cloud provider. In this "hosted private" model, a company such as Sun can install, configure, and operate the infrastructure to support a private cloud within a company's enterprise datacenter. This model gives companies a high level of control over the use of cloud resources while bringing in the expertise needed to establish and operate the environment.

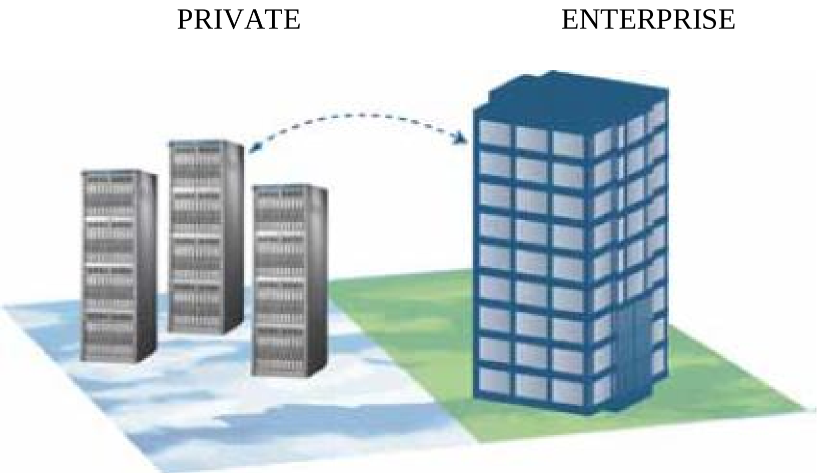

Figure 2. Private clouds may be hosted at a colocation facility or in an enterprise datacenter. They may be supported by the company, by a cloud provider, or by a third party such as an outsourcing firm.

*3) Hybrid clouds:* Hybrid clouds combine both public and private cloud models (Figure 3) [5]. They can help to provide on-demand, externally provisioned scale. The ability to augment a private cloud with the resources of a public cloud can be used to maintain service levels in the face of rapid workload fluctuations.

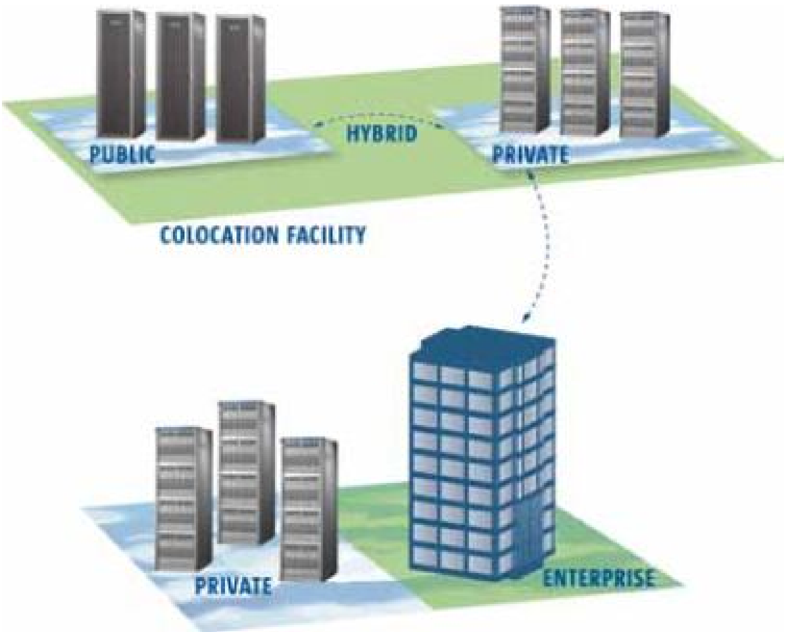

Figure 3. Hybrid clouds combine both public and private cloud models, and they can be particularly effective when both types of cloud are located in the same facility.

This is most often seen with the use of storage clouds to support Web 2.0 applications. A hybrid cloud also can be used to handle planned workload spikes. Sometimes called "surge computing," a public cloud can be used to perform periodic tasks that can be deployed easily on a public cloud. Hybrid clouds introduce the complexity of determining how to distribute applications across both a public and private cloud. Among the issues that need to be considered is the relationship between data and processing resources. If the data is small, or the application is stateless, a hybrid cloud can be much more successful than if large amounts of data must be transferred into a public cloud for a small amount of processing.

### *B .Service Models*

There are certain services and models working behind the scene making the cloud computing feasible and accessible to end users.
Infrastructure as a Service (IAAS), Platform as a Service (PAAS), Software as a Service (SAAS).

*1) Infrastructure as a Service (IaaS):* Infrastructure as a Service is a provision model in which an organization outsources the equipment used to support operations, including storage, hardware, servers and networking components. The service provider owns the equipment and is responsible for housing, running and maintaining it. The client typically pays on a per-use basis.
Characteristics and components of IaaS include:
Utility computing service and billing model.
Automation of administrative tasks.
Dynamic scaling.
Desktop virtualization.
Policy-based services.
Internet connectivity.

*2) Platform as a Service (PaaS):* Platform as a Service (PaaS) is a way to rent hardware, operating systems, storage and network capacity over the Internet. The service delivery model allows the customer to rent virtualized servers and associated services for running existing applications or developing and testing new ones.

Platform as a Service (PaaS) is an outgrowth of Software as a Service (SaaS), a software distribution model in which hosted software applications are made available to customers over the Internet. PaaS has several advantages for developers. With PaaS, operating system features can be changed and upgraded frequently. Geographically distributed development teams can work together on software development projects. Services can be obtained from diverse sources that cross international boundaries. Initial and ongoing costs can be reduced by the use of infrastructure services from a single vendor rather than maintaining multiple hardware facilities that often perform duplicate functions or suffer from incompatibility problems. Overall expenses can also be minimized by unification of programming development efforts.

*3) Software as a Service (SaaS):* Software as a Service (SaaS) is a software distribution model in which applications are hosted by a vendor or service provider and made available to customers over a network, typically the Internet [4]. SaaS is becoming an increasingly prevalent delivery model as underlying technologies that support Web services and service-oriented architecture (SOA) mature and new developmental approaches, such as Ajax, become popular. Meanwhile, broadband service has become increasingly available to support user access from more areas around the world. SaaS is closely related to the ASP (application service provider) and on demand computing software delivery models. IDC identifies two slightly different delivery models for SaaS. The hosted application management (hosted AM) model is similar to ASP: a provider hosts commercially available software for customers and delivers it over the Web. In the software on demand model, the provider gives customers network-based access to a single copy of an application created specifically for SaaS distribution.

Characteristics of the SaaS model include:
Easier administration.
Automatic updates and patch management.
Compatibility: All users will have the same version of software.
Easier collaboration, for the same reason global accessibility.

Figure. 4 depicts the general architecture of a cloud platform, which is also called cloud stack [4]. Building upon hardware facilities (usually supported by modern data centers), cloud services may be offered in various forms from the bottom layer to top layer. In the cloud stack, each layer represents one service model. Infrastructure-as-a-Service (IaaS) is offered in the bottom layer, where resources are aggregated and managed physically (e.g., Emulab) or virtually (e.g., Amazon EC2), and services are delivered in forms of storage (e.g., GoogleFS), network (e.g., Openflow), or computational capability (e.g., Hadoop MapReduce). The middle layer delivers Platform-as a-Service (PaaS), in which services are provided as an environment for programming (e.g., Django) or software execution (e.g., Google App Engine). Software as a Service (SaaS) locates in the top layer, in which a cloud provider further confines client flexibility by merely offering software applications as a service.

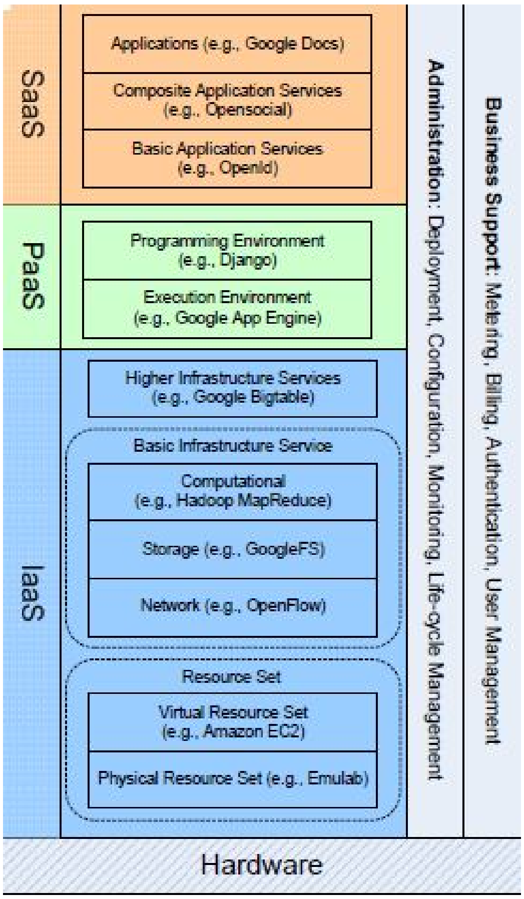
SaaS
Applications (e.g., Google Docs)
Composite Application Services (e.g., Opensocial)
Basic Application Services (e.g., OpenId)
PaaS
Programming Environment (e.g., Django)
Execution Environment (e.g., Google App Engine)
IaaS
Higher Infrastructure Services (e.g., Google Bigtable)
Basic Infrastructure Service
Computational (e.g., Hadoop MapReduce)
Storage (e.g., GoogleFS)
Network (e.g., OpenFlow)
Resource Set
Virtual Resource Set (e.g., Amazon EC2)
Physical Resource Set (e.g., Emulab)
Administration: Deployment, Configuration, Monitoring, Life-cycle Management
Business Support: Metering, Billing, Authentication, User Management
Hardware

Figure. 4. Architecture of Cloud Computing

Apart from the service provisioning, the cloud provider maintains a suite of management tools and facilities (e.g., service instance life-cycle management, metering and billing, dynamic configuration) in order to manage a large cloud system.

## III. Issues

Cloud got many issues when it comes to security especially on Data loss, Data integrity, Data theft, and Privacy and security.

**Data Loss**

Data loss is a very serious problem in Cloud computing. If the vendor closes due to financial or legal problems there will be a loss of data for the customers.

**Data Integrity**

When a data is on a cloud anyone from any location can access those data's from the cloud. Cloud does not differentiate between a sensitive data from a common data thus enabling anyone to access those sensitive data's.

**Data Theft**

Most of the cloud Vendors instead of acquiring a server tries to lease a server from other service providers because they are cost affective and flexible for operation.

**Distributed Storage Systems**

At the early years, the Network-Attached Storage (NAS) and the Network File System (NFS) provide extra storage devices over the network such that a user can access the storage devices via network connection. Afterward, many improvements on scalability, robustness, efficiency, and security were proposed.

## IV. Own Cloud

To implement private cloud, Own Cloud [6] platform is used. Own Cloud is an open source platform. Own Cloud gives you universal access to your files through a web interface or WebDAV. It also provides a platform to easily view and synchronize your contacts, calendars and bookmarks across all your devices and enables basic editing right on the web. Own Cloud is extendable via a simple but powerful API for applications and plugins.

**Features of Own Cloud are-**

Access your data.
Synchronize your data.
Share your data.
Versioning
.Encryption.

Figure 5.OwnCloud

As shown in (figure 6) the main focus of ownCloud is storing and sharing documents, calendars and contacts in a secure way, which also embraces open standards. There are two main reasons why you or your group may want to use own Cloud rather than a similar service like Google Docs or Dropbox.The first is that it offers more security for the files that you store, because you store them on your own server and can also encrypt them. The second reason is that by using ownCloud, you are supporting a more decentralized and less monitored internet. If you are uncomfortable about the level of data analysis and surveillance involved in signing up for corporate internet services, then ownCloud provides a very usable alternative.

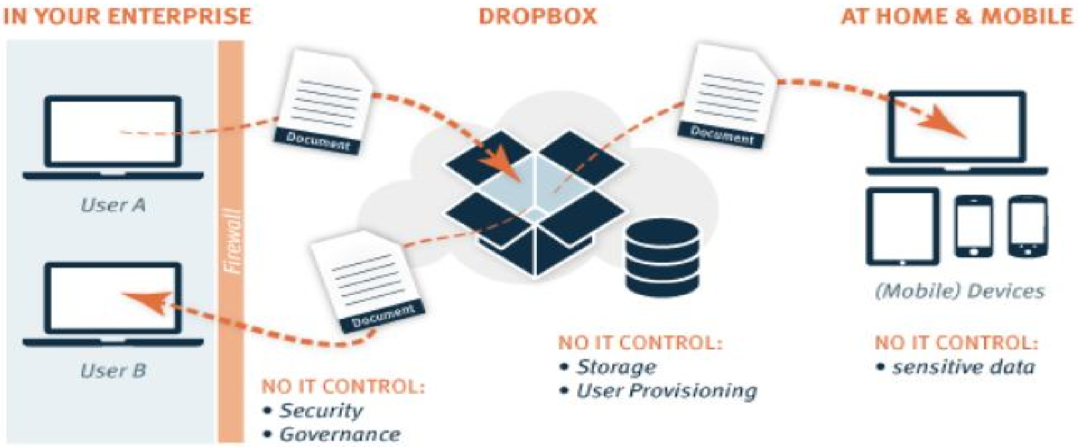

Figure 6. OwnCloud-overview

## V. Cryptographic Algorithm and Data Backup

The first thing you must look into is the security measures that your cloud provider already has in place. These vary from provider to provider and among the various types of clouds. What encryption methods do the providers have in place? What methods of protection do they have in place for the actual hardware that your data will be stored on? Will they have backups of my data?

### *A. Encryption while storing the data on cloud*

Storing data in a third party's cloud system causes serious concern on data confidentiality. In order to provide strong confidentiality for messages in storage servers, a user can encrypt [2] messages by a cryptographic method before applying an erasure code method to encode and store messages. Our system is highly distributed where storage servers independently encode and forward messages and key servers independently perform partial decryption. Our system meets the requirements that storage servers independently perform encoding and encryption and key servers independently perform partial decryption [3]. The author [1] explores the cost of common cryptographic primitives (AES, MD5, SHA-1, RSA, DSA, and ECDSA) and their viability for cloud security purposes.

### *1) How Algorithm [7] Works?*

To ensure data access security, only authorize users are allowed to access data in the cloud. Therefore, to access data new user will have to register by cloud service provider [sh3].The data owner computes a message digest using D5 for every file belonging to the data set available with it. In this algorithm128 bit MD5 hash is used over any other data like SHA-1

(160bit) for data integrity because symmetric key is using for digest (Figure 7). This cryptographic strength ensures data confidentiality and integrity between data owner and user.MD5- (Message-Digest algorithm 5), a widely used cryptographic hash function with a 128-bit hash value, processes a variable-length message into a fixed-length output of 128 bits. The input message is broken up into chunks of 512-bit blocks. The message is padded so that its length is divisible by 512.

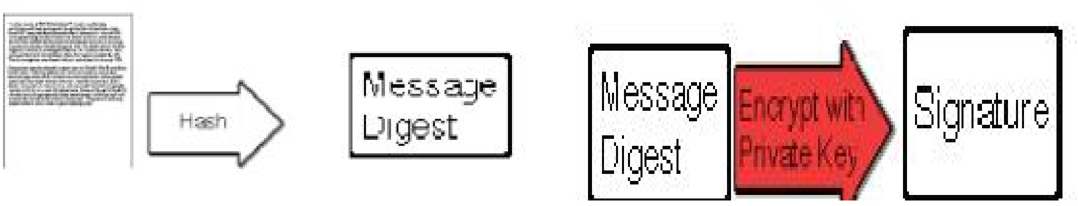

Figure 7.Working of MD5

**MD5 hash algorithm consist of 5 steps:**
Step 1. Append Padding Bits
Step 2. Append Length
Step 3. Initialize MD Buffer
Step 4. Process Message in 16-Word Blocks
Step 5. Output

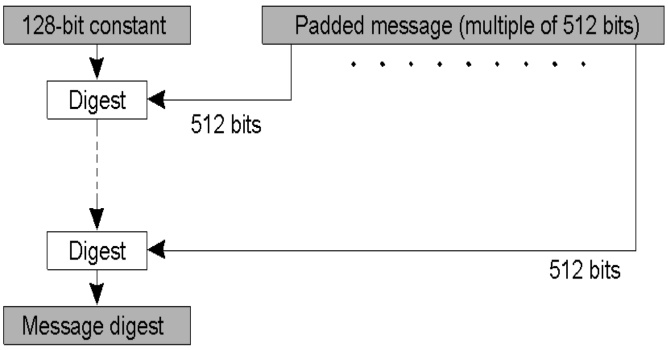

Figure 8.Steps included in MD5

*2) SHA (SECURE HASH ALGORITHM):*

Various types of SHA:-

- ● SHA 0
- ● SHA 1
- ● SHA 256
- ● SHA 512

The Secure Hash Algorithm is one of a number of cryptographic hash functions [8]. There are currently three generations of Secure Hash Algorithm:

- ● SH`A-1 is the original 160-bit hash function. Resembling the earlier MD5 algorithm.
- ● SHA-2 is a family of two similar hash functions, with different block sizes, known asSHA-256 and SHA-512. They differ in the word size; SHA-256 uses 32-bit words whereSHA-512 uses 64-bit words.
- ● SHA-3 is a future hash function standard still in development.

**Basic Properties**

- ● Output Size: - 160(bits)
- ● Internal State Size: 160(bits)
- ● Block Size: 512(bits)
- ● Maximum Message Size 264 − 1
- ● Length Size 64(bits)
- ● Word Size 32(bits)
- ● Collision Attacks No (251 attack)
- ● Rounds 80

ALGORITHM:

There are 80 rounds in SHA Algorithm. The hash value generated by the SHA hash function. The SHA algorithm uses 5 state variables, each of which is a 32 bit integer (an unsigned long on most systems). These variables are sliced and diced and are (eventually) the message digest. The variables are initialized as follows:

- h0 = 0x67452301
- h1 = 0xEFCDAB89
- h2 = 0x98BADCFE
- h3 = 0x10325476
- h4 = 0xC3D2E1F0

### *B. Remote backup and restore function*

Amazon S3 buckets allowing us to create our own backup cloud program. This solution is implemented in public cloud; hence we will be implement data backup and restore facility in private cloud. Proposed system will provide data backup and restore facility in private cloud in confidential form.When we talk about Backup server of main cloud, we only think about the copy of main cloud. When this Backup server is at remote location (i.e. far away from the main server) and having the complete state of the main cloud, then this remote location server is termed as Remote Data Backup Server [9]. The main cloud is termed as the central repository and remote backup cloud is termed as remote repository.

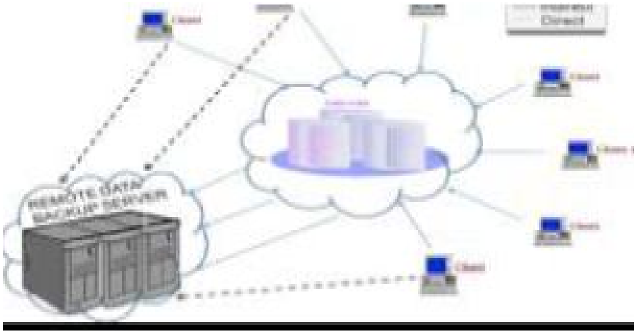

Figure.9 Remote data Backup Server and its Architecture

And if the central repository lost its data under any circumstances either of any natural calamity (for ex - earthquake, flood, fire etc.) or by human attack or deletion that has been done mistakenly and then it uses the information from the remote repository [9]. The main objective of the remote backup facility is to help user to collect information from any remote location even if network connectivity is not available or if data not found on main cloud. As shown in Figure-9 clients are allowed to access the files from remote repository if the data is not found on central repository (i.e. indirectly).

## VI. Conclusion

To summarize, cloud provides many options for everyday computer user as well as large and small business. It opens up the world of computing broader range of uses and increase the ease of used by giving access through internet connection. However, with this increase ease also come drawback. If you are considering using cloud, be certain that you identify what information you will be putting out in cloud, who will have access the information and what you will you need to make sure it is protected. In future, we will extend our research by deploying an application in own cloud and also by providing security and backup to justify our concepts of security for cloud computing.

## References

[1] Y. Chen and R. Sion, "*On securing untrusted clouds with cryptography*," in Proceedings of the 9th annual ACM workshop on Privacy in the electronic society. ACM, 2010, pp. 109–114.

[2]Lukas Malina and Jan Hajny, "*Efficient security solution for privacy-preserving cloud services*", in Proceedings of the 6TH INTERNATIONAL CONFERENCE ON TELECOMMUNICATIONS SIGNAL PROCESSING YEAR 2013.

[3] Hsiao-Ying Lin, Member, IEEE, and Wen-Guey Tzeng, Member, IEEE,"*A Secure Erasure Code-Based cloud storage system with secure data forwarding*" , JUNE 2012.

[4] K.S.Suresh and Prof K.V.Prasad," *Security Issues and Security Algorithms in Cloud Computing*", Volume 2, Issue 10, October 2012.

[5]Sun Microsystems Inc. ,"*Introduction to Cloud Computing Architecture*", 1st Edition, June 2009

[6] http://owncloud.org.

[7] Sunil Sanka, Chittaranjan Hota, Muttukrishnan Rajarajan, "*Secure Data Access in Cloud Computing*", Computer Science and Information Systems Group, Birla Institute of Technology and Science-Pilani Hyderabad Campus, Shameerpet, Hyderabad, INDIA. Information Security Group, School of Engineering and Mathematical Sciences City University, Northampton Square, LONDON, EC1V0HB r.muttukrishnan@city.ac.uk.

[8] Hashing Algorithms/binarywarriors@gmail.com.

[9] Ms. Kruti Sharma and Prof. Kavita R Singh," *Seed Block Algorithm: A Remote Smart Data Back-up Technique for Cloud Computing*", 2013 International Conference on Communication Systems and Network Technologies.