Providing Secure Data Storage on Cloud

Garima Singh[1], Akansha Chitransh[2]

[1]Computer Science & Engineering Department, V.N.I.T Nagpur

[2]Computer Science & Engineering Department, G. L. Bajaj Institute of Technology & Management

Abstract*—Cloud computing provides us a means by which we can access the applications as utilities, over the Internet. It allows us to create, configure, and customize applications online. Cloud computing allows to computer infrastructure namely hard disk, development platform, database, computing power or complete software applications as services rather than a product available "on-demand" basis. Storage as a Service (StaaS) business model of cloud computing allows cloud service provider to rent space in their storage infrastructure to a smaller company or individuals. Storage as a Service (StaaS) allows users to store their data at remote disks and access them anytime from anywhere as long as they have access to the cloud. Storage as a Service is generally seen as a good alternative for a small or mid-sized business that lacks the capital budget and/or technical personnel to implement and maintain their own storage infrastructure. This data outsourcing service poses various challenges with respect to the outsourced data including data privacy protection, high availability, reliability, performance, replication and data consistency. This paper focusses on a framework which securly stores and manages the data stored on cloud.

***Index Terms*- Cloud computing, Data confidentiality, Data integrity, Data auditing.**

I. INTRODUCTION

Cloud computing is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable com- puting resources namely networks, servers, storage, applications, and services these resources can be rapidly provisioned and released with minimal management effort [?], technologies that are helpful to enable cloud computing are virtualization, multite- nancy, web services, distributed computing, utility computing and system automation. Virtualization allows to share single physical instance of an application or resource among multiple client-organizations. Multitenancy enables virtual isolation across a large pool of users, and cloud's users can use and customize the application as they individually have its own instance running. Cloud resources can be accessed by the users through web services.

Cloud computing has the following characteristics:

1) Shared resource pooling: The providers computing re- sources are pooled together to serve cloud's users using multiple- tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand.

2) Elastic resources: Cloud services can scale up or down quickly and easily to meet user's demand. Cloud service provider can smoothly add or remove users, software features, and other resources.

3) Pay for use: Cloud computing resource usage can be measured, by this metered service consumer only pays for what he used, the more consumer utilize cloud resources the higher the bill. The amount of resources that consumer may use can be monitored and controlled from both consumer side and cloud providers side which provides transparency.

4) Broad network access: Cloud services are available over

the network and accessed through thin or thick client platforms such as mobile phones, laptops and PDAs.

5) On-demand self- service: On-demand self-service allows

users to obtain, configure and deploy cloud services themselves, users are able to provision cloud computing resources without requiring human interaction, mostly done though a web-based self-service portal (management console).

6) Cost Effectiveness: Resource sharing improves utilization of physical resources and thus reduces the associated cost.

II. CLOUD SERVICE DELIVERY MODELS

In cloud computing the services provided by cloud provider is classified into three main categories Software as a service, Platform as a service, and Infrastructure as a service, as shown in Fig. 1, Each service delivery model consumes the services provided by the layer beneath.

- Software-as-a-Service (SaaS)

SaaS allows the consumer to access and use cloud provider's software applications including video-on-demand, email ser- vices, financial, business applications, office assistance that is hosted, deployed, and managed by the cloud provider. consumers have limited control over the applications and restriction on how to use and intract with the applications. Consumer can access these services through thin client like web browser. Salesforce.com, Nimsoft Monitor are the examples of SaaS service.

- Platform-as-a-Service (PaaS)

PaaS allows the service consumer to define, develop, configure, deploy, manage, and monitor cloud applications. Cloud service provider offers a development environment to application developers. Developers develop applications and offer those services through the cloud service provider's platform. PaaS allows consumers to deploy and control applications and their hosting environment configurations, consumers do not have direct control over the underlying cloud infrastructure. Google App Engine, Microsoft Azure, SQL Azure are the examples of PaaS service.

- Infrastructure-as-a-Service (IaaS) It outsources infrastruc- ture capabilities including data storage resources, computa- tional power, hardware, servers and networking components based on demand. The cloud service provider hostess the equipment and is responsible for running and maintaining them. Consumers are not given direct access to resources but have the ability to select and configure resources as required based on their needs.

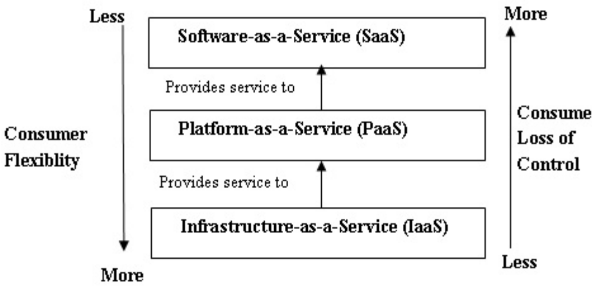


Fig. 1: Cloud service delivery models

Cloud services are available only for authorized users. Use of services is governed by contracts, which specify the responsi- bilities of the cloud service provider and the responsibilities of the customer. Cloud services can be deployed in different ways, depending on the organizational structure and the provisioning location. Four cloud computing deployment models are:

- Private Cloud: Private cloud is a cloud infrastructure operated only for a single organization. Mutually trusted consumers provides more flexibility and highest trust level. Private cloud is managed internally or by a third-party and may exist on premise or off premise. It is not available for general public. As shown in Fig. 2, private cloud services are only for the customers of organization A, a Customer from any other organization can not use the services provided by the private cloud.

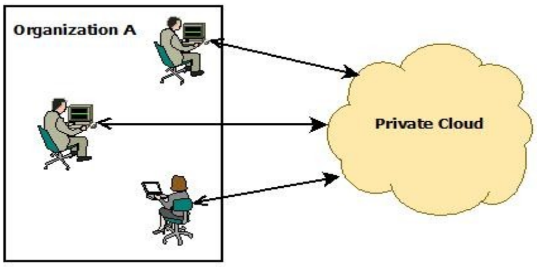


Fig. 2: Private Cloud

- Public Cloud: Public cloud allows multiple public organi- zations or consumers to purchase cloud provider's services such as applications and storage, using the same shared infrastructure. A public cloud is hosted, operated, and managed by a third-party vendor from one or more data centers. The services are provided to the public based on the service level agreement between the provider and the consumer. The lack of a trust model between the cloud providers and consumers is the main obstacle for this model. There are various public cloud service providers like Amazon Elastic Compute Cloud (EC2), IBM's Blue Cloud, Google AppEngine and Windows Azure Services Platform. Public cloud is shown in Fig 3, in which Customer A, Customer B, Customer C etc are from general public these can be any organization, or individuals.

Community Cloud: The community cloud infrastructure is shared by several organizations that supports a specific community or who may have strategic relationships or shared concerns (e.g. mission, jurisdiction, security requirements, policy, and compliance considerations). It may be managed by the internally or a third party and may exist on premise or off premise. Community cloud shown in Fig. 4, provides services to organization A, organization B and organization C, these three organizations have some shared concern they can use the resources provided by community cloud, any other organization can't use services provided by community cloud.

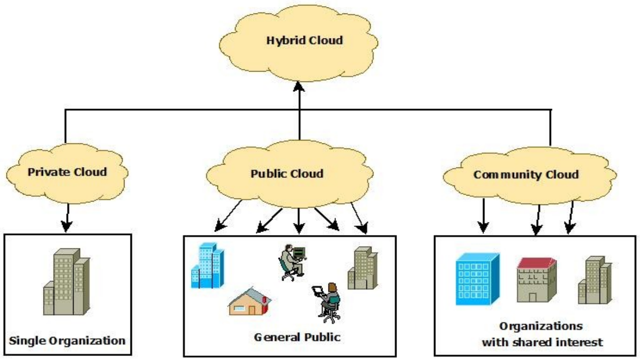


Fig. 5: Hybrid Cloud

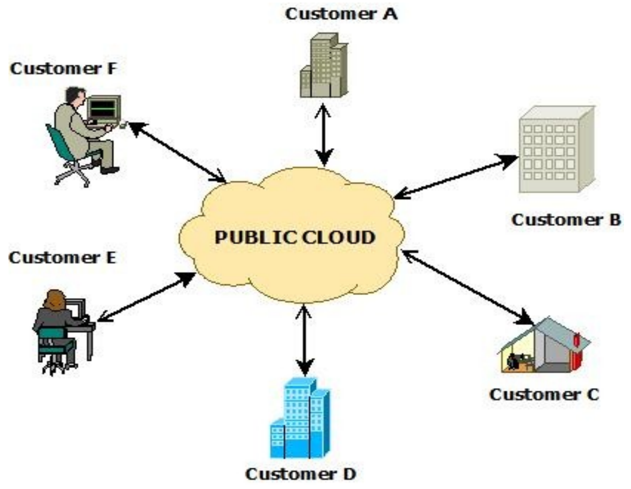


Fig. 3: Public Cloud

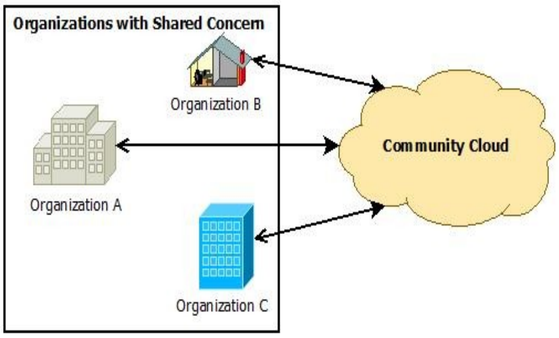


Fig. 4: Community Cloud

I. Hybrid Cloud: A cloud that is a composition of two or more types of clouds (private, community, or public) is called hybrid cloud. After composing these entities remain unique but are bound together by standardized or proprietary technology that enables data and application portability (e.g., cloud bursting for load-balancing between clouds), as shown in Fig. 5.

III. MOTIVATION

Cloud provides the facility to use resources as per the payment. There is no need to worry about maintenance of machine and management of data. A user can get required computing power on demand. Cloud offers businesses the opportunity to do things faster and better:

- *Scalability*: Data and application resources can be quickly provisioned whenever and wherever you need them.
- *Availability*: The cloud provider ensures that your resources remain continuously available and always secure.
- *Less Maintenance*: Hardware, applications and bandwidth are managed by the provider.
- *Expert Service*: At Expedient, cloud computing services are continuously monitored and maintained by onsite staff of expert data center technicians.

IV. STORAGE-AS-A-SERVICE

By cloud's storage service, data owner stores sensitive data like personal health records, photo albums, tax documents, financial transactions, emails and other files on cloud [1]. There are several benefits of cloud data storage service, like relief from the burden of storage management, universal data access with independent geographical locations, avoidance of capital expenditure on hardware, software, personal maintenance, and so on [2].

The moving of data to the cloud means that the responsibility of data security becomes shared between the cloud provider and customer. The extent to which the data is secure is now limited to the security controls and policies applied by both the cloud consumer and cloud service provider. There are various security issues with the cloud. Most basic data outsourcing security issues are [3] [4]:

- *Authentication*: The process of identifying an individual, which ensures that the individual is who he claims to be.
- *Authorization*: The process of giving individuals access to system objects based on their identity.
- *Confidentiality*: It refers to keep data private. Metadata must also be safeguarded along with internal secrets and sensitive personal data, because metadata and transactional data can also leak important details about IT enterprises or individuals.
- *Integrity*: It is a degree of confidence that the data in the cloud is what it was supposed to be there, and is protected against alteration without authorization.

Security of the data stored on the cloud is one of the major issue, which can hamper the popularity of the cloud. Although infrastructure of the cloud is highly powerful and reliable than personal computing devices, but data loss can happen in any infrastructure. Despite of so much benefits of cloud computing, if security and privacy measures of data will not be primary concern, then, it will make cloud computing an insecure and unreliable technology.

V. RELATED WORK

A lot of work has already been done for securing data stored on cloud, and still new researches are undergoing to secure data stored on cloud. In [5], [6], authors developed an approach to secure data, wherein the cloud stores encrypted data, and the organization (third party) stores decryption keys. The clients fetch the two and decrypt the data locally. In this scheme, the client has an overhead to decrypt the data, and a third party involvement is required. When users put their data on cloud, data integrity protection is challenging. In [7], author proposed a framework to check integrity of data that also involves a third party auditor to perform data auditing on cloud. This whole scheme depends upon a third party auditor (TPA).

VI. SECURITY ISSUES

I. Framework For Enhancing Data Storage Security On Cloud

A. System Assumptions

There are two system assumptions:

- File transfer activity from data owner to cloud controller and from cloud controller to data owner needs to be fully secure.

- Cloud controller is an honest entity.

B. Definition of a System Model

We consider a system for cloud storage as shown in Fig . 6, which involves data owners, the cloud controller, and the storage servers. The owners create data and host their data on the cloud. The cloud controller authenticates and authorizes the incoming data requests, stores owner's data and provides the data access to the owner. The storage servers store owner's data.

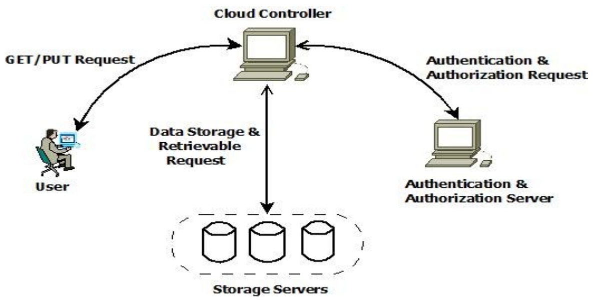

Cloud Controller
GET/PUT Request
Authentication & Authorization Request
Data Storage & Retrievable Request
User
Authentication & Authorization Server
Storage Servers

Fig. 6: System Model

II. Network Layout

Network layout for private cloud is shown in Fig. 7, cloud controller node has ip address 172.31.134.15, authentication and authorization server has ip address 172.31.134.13, there are three data storage servers in different zones, first data storage node is placed in zone1 and it has ip address 172.31.134.11, second data storage node is placed in zone2 and it has ip address 172.31.134.12 and third data storage node is placed in zone3 and has ip address 172.31.134.22. Storage nodes are connected with each other through a private switch over the local area network (LAN).

III. Framework For Enhancing Confidentiality Of Data On Cloud

To enhance confidentiality of data stored on cloud, the cloud controller has to maintain abstract information about the data.

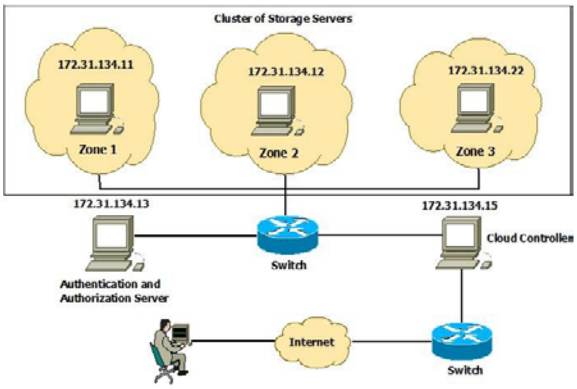


Fig. 7: Cloud's Network Layout

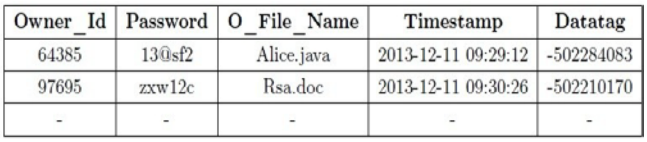
Cloud controller stores abstract information of the data into a database known as Abstract Info.

A. Abstract Info Database

For the organization of data on cloud, the cloud controller stores following information into the AbstractInfo database:

- Identity of the data owner (Owner_Id) who is storing his data on cloud servers.
- Password of the data owner.
- Original data name (O_Data_Name) by which it is stored on the owner's system.
- Timestamp at which data has been received by the server.
- Datatag corresponding to the received data, datatag will be a unique value and will help to identify data of any particular data owner on cloud, as shown in Table 1.

Table 1: AbstractInfo Database



Owner_Id	Password	O_File_Name	Timestamp	Datatag
64385	13@sf2	Alice.java	2013-12-11 09:29:12	-502284083
97695	zxw12c	Rsa.doc	2013-12-11 09:30:26	-502210170
-	-	-	-	-

B. Cloud Data Storage Framework

To store data on cloud, data owner will send his data on cloud. On cloud, the cloud controller will authenticate the data storage request and will check authorization of the request. If the request is authenticated and have the rights to store data on cloud, then the cloud controller will store abstract information (Owner_Id, Password, O_Data _Name, Timestamp, Datatag) of the data into Abstract Info database. Now cloud controller will encrypt this data. By encryption, with the content of data, data name will be converted to the corresponding datatag name and format of the data will also be converted. Cloud controller will send this encrypted data to the storage servers. Finally, the cloud controller will send confirmation of data storage back to the data owner.

By doing encryption, content of the data and metadata (data name, data type) will be converted into a non-understandable form by this encryption an eavesdropper will not be able to get any kind of information about the data. By following this procedure, data owner receives data storage confirmation message from the cloud, so data owner is sure about storage of data on cloud, as shown in Fig. 8.

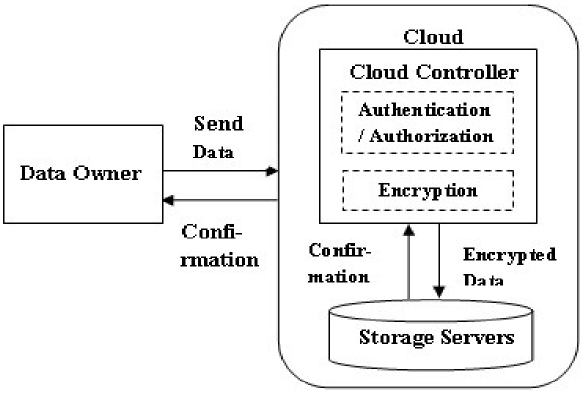


Fig. 8: Data Storage Framework

C. Cloud Data Retrieval Framework

To retrieve data from cloud, the data owner has to send data request to the cloud. On cloud, the cloud controller finds datatag name corresponding to the requested data, and sends data request corresponding to the datatag name to the storage servers, the storage server will send back the data to cloud controller, where its decryption will be done and data with original name and format will be sent back to the data owner securely, as shown in Fig. 9.

IV. Framework For Cloud Data Integrity Check

By data-as-a-service model, the data is stored on a remote location, the data owner does not have any control over his the Abstract Info database, which will help to check integrity of the data stored on cloud. Now the cloud controller will send this encrypted data to the storage servers and will send back confirmation of data storage to the data owner.

C. Auditing framework

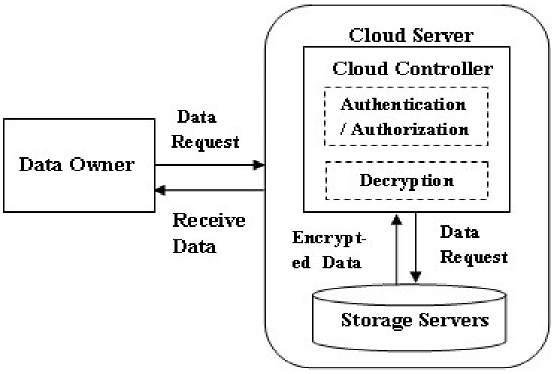


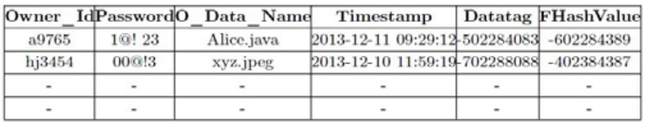
Fig. 9: Data Retrieval Framework

data. On cloud, checking data integrity will be a big challenge for the data owner. Cloud has to provide an auditing scheme to check integrity of the data stored on cloud.

A. Abstract Info Database

To perform data auditing on cloud, Abstract Info database has to be modified. A new field FHash Value is included into the database that stores hash value of the encrypted data, as shown in Table 2.

Table 2: AbstractInfo Database



Owner_Id	Password	O_Data_Name	Timestamp	Datatag	FHashValue
a9765	1@! 23	Alice.java	2013-12-11 09:29:12	-502284083	-602284389
hj3454	00@!3	xyz.jpeg	2013-12-10 11:59:19	-702288088	-402384387
-	-	-	-	-	-
-	-	-	-	-	-

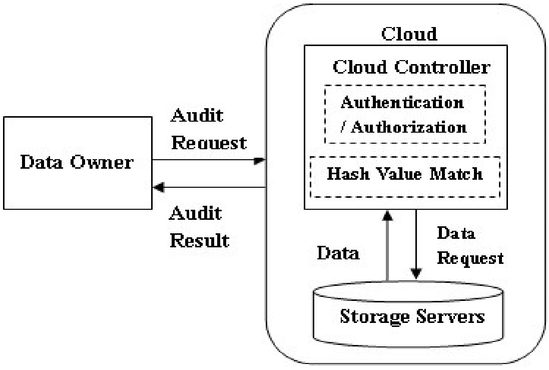


Fig. 10: Data Framework

To perform data auditing, data owner will send auditing request for his data to the cloud. On cloud, the cloud controller will retrieve owner's data back from the storage servers and will calculate hash value corresponding to the owner's data. Cloud controller will match this calculated hash value with the already stored hash value, and sends the audited result back to the data owner. By this audited result, the owner will be able to verify that weather the integrity of his data on cloud is maintained or not, as shown in Fig. 10.

B. Cloud Data Storage Framework

To store data on cloud, the data owner will send his data on cloud. On cloud, the cloud controller will authenticate the data storage request and will check authorization of the request. If the request is authenticated and have the rights to store data on cloud then cloud controller will store abstract in- formation (Owner_Id, Password, O_Data _Name, Timestamp, Datatag, F Hash Value) of the data. The cloud controller will encrypt this data. By encryption, with the content of data, data name will be converted to the corresponding datatag name and format of the data will also be converted. The cloud controller stores hash value (F Hash Value) of the data into Fig. 10: Framework for Integrity Check

VII. CONCLUSION

In this paper, the proposed framework is able to provide authentication, authorization, data privacy protection and in- tegrity checking. In this framework, there is no requirement of any third party organization for auditing or key management functions. File owner does not have any overhead of perform- ing encryption and decryption. All the load of auditing has been moved to the cloud server, which greatly improves the auditing performance.

REFERENCES

[1] K. Ren, C. Wang, and Q. Wang, "Toward secure and effective data utilization in public cloud," *IEEE Network*, no. 12, pp. 69–73, 2012.

[2] K. Ren, C. Wang, J. Li, and W. Lou, "Toward publicly auditable secure cloud data storage service," *IEEE Network*, no. 10, pp. 19–24, 2010.

[3] K. Ren, C. Wang, and Q. Wang, "Security challenges for public cloud," *IEEE Computer Society*, no. 12, pp. 69–72, 2012.

[4] K. S. O., I. F., and A. O., "Cloud computing security issues and challenges,"

[5] K. P. N. Puttaswamy, C.Kruegel, and B. Zhao, "Silverline: Toward data confidentiality in storage-intensive cloud applications," *ACM*, 2011.

[6] K. Hashizume, D. G. Rosado, E. Fernndez-Medina, and E. B. Fernandez, "An analysis of security issues for cloud computing," *Internet Services and Applications*, 2013.

[7] K. Yang, "An efficient and secure dynamic auditing protocol for data stor- age in cloud computing," *IEEE Transaction on Parallel and distributed systems*, vol. 24, no. 9, pp. 1717–1726, 2013.