

DOCUMENT RETRIEVAL IN CLOUD

K.Maheswari, E.Pavithra, S.Keerthana, B.Rahman

¹Assistant Professor, SNS College of Technology, Coimbatore, Tamilnadu, India

^{2,3,4}B.E CSE Student, SNS College of Technology, Coimbatore, Tamilnadu, India

¹mahi.kirubakaran@gmail.com, ²pavibvn95@gmail.com,

³keerthana.surendran69@gmail.com, ⁴rahman82201@gmail.com

Abstract: Cloud computing is used to reduce the expense of the users and the man power they use vastly for maintaining the servers that is mostly needed only for some particular days for important publishes or updates. The main Objective of the System is to review the available document retrieval into cloud. In this proposed System the concept of the cloud and how the document retrieval is in the cloud. The document retrieval consists of these facilities like securing the performance evaluation of the schemes. The documents to be retrieved are consisting of certificates, credit card, PAN card etc., can be retrieved from the cloud. Resources provided by use of the data storage and computation outsourced. Within the unique id, one can access the services and retrieve the data in the cloud. In this paper, we feature the new idea to provide some better document retrieval in cloud. The system that we are going to develop will be available for all people who possess the cloud. For the user whenever the details are needed they can reach the server through an unique id and can access the details in the cloud.

I. INTRODUCTION

Cloud computing is a type of Internet-based computing that provides shared computer processing resources and data to computers and other devices on demand. It is a model for enabling ubiquitous, on-demand access to a shared pool of configurable computing resources (e.g., computer networks, servers,

storage, applications and services), which can be rapidly provisioned and released with minimal management effort. Cloud computing and storage solutions provide users and enterprises with various capabilities to store and process their data in third-party data centers that may be located far from the user—ranging in distance from across a city to across the world. Cloud computing relies on sharing of resources to achieve coherence and economy of scale, similar to a utility (like the electricity grid) over an electricity network.

II. CLOUD COMPUTING WORKS

The goal of the cloud computing is to apply for the traditional supercomputing or the high-performance computing) power, normally used by military and research facilities, to perform tens of trillions of computations per second, in consumer-oriented applications such as financial portfolios, to deliver personalized information, to provide data storage or to power large, immersive online computer games. Cloud computing increases mobility, as you can access your documents from any device in any part of the world

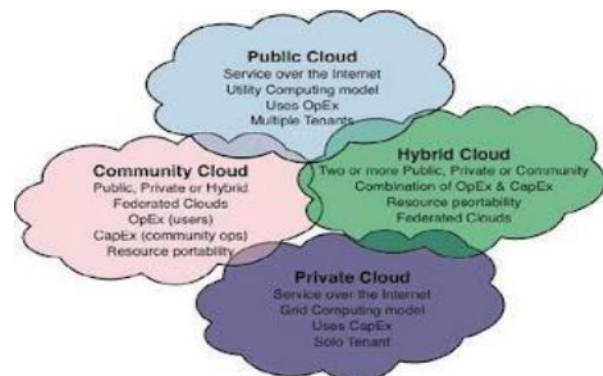


Fig.1.Cloud deployment Model

To do this, cloud computing uses networks of large groups of servers typically running low-cost consumer PC technology with specialized connections to spread data-processing chores across them.

This shared IT infrastructure contains large pools of systems that are linked together.

Often, virtualization techniques are used to maximize the power of cloud computing.

III. INFRASTRUCTURE

- **On-demand self-service** -- A consumer can independently provide computing capabilities, such as compute time, network connectivity and storage are needed automatically without requiring any human interaction with each service's provider
- **Broad network access** --The Capabilities are available over the network and accessed through standard mechanisms that promote use by heterogeneous thin or thick client platforms

generally has no control or knowledge over the exact location of the provided resources, but may be able to specify location at a higher level of abstraction (e.g., country, state, region or datacenter). Examples of computing resources include storage, processing (compute), memory, network bandwidth, and virtual machines

- **Rapid elasticity** -- Capabilities can be rapidly and elastically provisioned, in some cases automatically, to quickly scale out, and rapidly released to quickly scale in. To the consumer, the capabilities available for provisioning often appear to be unlimited and can be purchased in any quantity at any time
- **Measured Service** -- The Cloud systems are automatically control and optimize resource use by leveraging a metering capability at some level of abstraction appropriate to the type of service (e.g., storage, compute, bandwidth, active user accounts, etc.). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilized service.

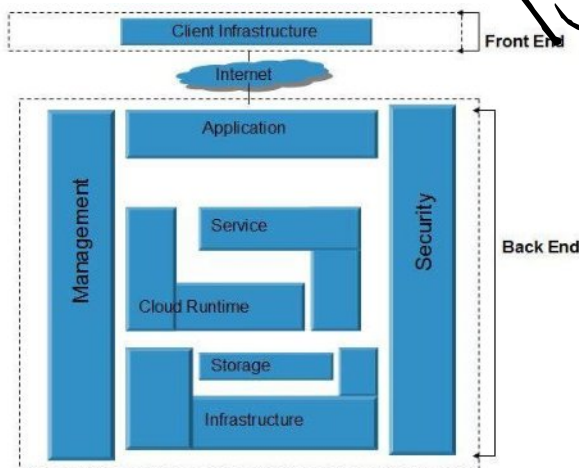


Fig.2.Cloud Infrastructure Architecture

- **Resource pooling** -- The resource pooling is to provide computing resources and are pooled to serve multiple consumers using a multi-tenant model, with different physical and virtual resources dynamically assigned and reassigned according to consumer demand. There is a sense of location independence in that the customer

IV. DOCUMENT RETRIEVAL ARCHITECTURE

Cloud computing is a type of Internet-based computing that provides shared computer processing resources and data to computers and other devices on demand. It is a model for enabling ubiquitous, on-demand access to a shared pool of configurable computing resources (e.g., computer networks, servers, storage, applications and services) which can be rapidly provisioned and released with minimal management effort.

Cloud computing and storage solutions provide users and enterprises with various capabilities to store and process their data in third-party data centers that may be located far from the user—ranging in distance from across a city to across the world. Cloud computing relies on sharing of resources to achieve coherence and economy of scale,

similar to a utility (like the electricity grid) over an electricity network

This system provides various information like people can provide the certificates and other details in the cloud.

If whenever the necessary items are needed the user may not go and search for the necessary documents. Neither can they use the document retrieval in the cloud. Therefore, by accessing a unique id into the website the user can search for the documents. It is been accessed in the form of the private cloud.

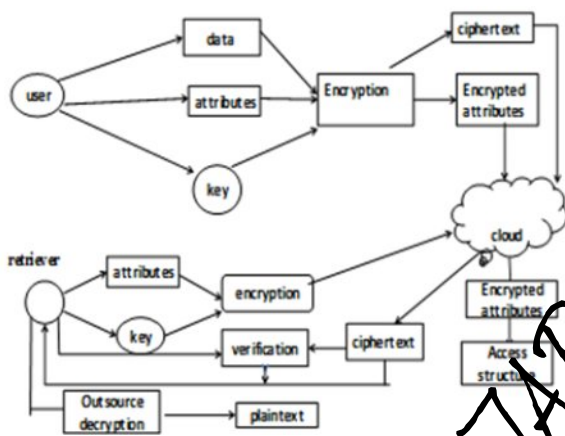


Fig.3. Overview of Document Retrieval System

For retrieving, the documents from cloud user have to create a unique id. With the unique id, the user can reach the server and search for any details that are need to access and so that the server can provide the details to the user.

A.PRIVATE CLOUD

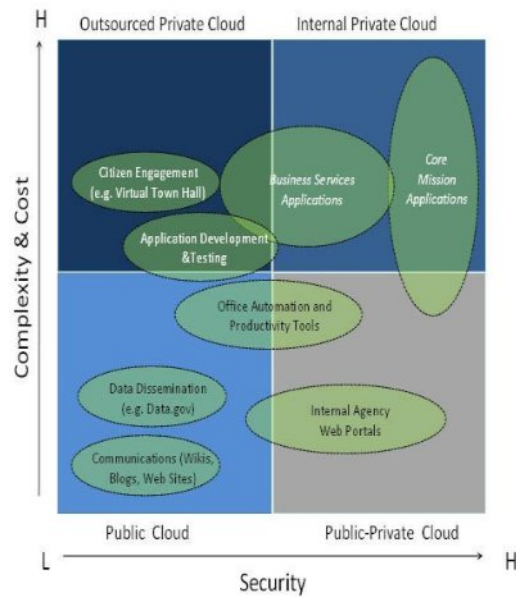


Fig.4. Comparison of Cloud Models

Private cloud is the phrase used to describe a cloud-computing platform that is implemented within the corporate firewall, under the control of the IT department. Private cloud includes the self-service and the scalability, the ability to provision machines, changing computing resources on demand and creating multiple machines for complex computing resources.

B.PUBLIC CLOUD

The cloud infrastructure is made available to the general public or a large industry group and is owned by an organization selling cloud services and the comparison of private and public cloud

C.ENCRYPTION

Encryption is an interesting piece of technology that works by scrambling data so it is unreadable by unintended parties. Encryption in the cloud should embrace the benefits of SaaS applications. As only a small percentage of corporate data in the cloud is sensitive, taking a selective, risk-appropriate approach allows for the necessary level of security without interfering with the native mobile, productive, collaborative benefits of SaaS applications.

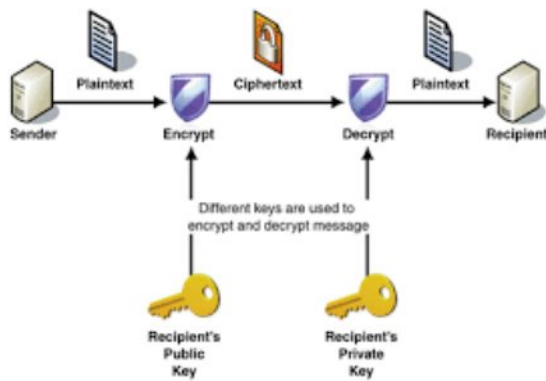


Fig.5.PKI Based Encryption in Cloud

D.HYBRID CLOUD

Hybrid cloud is a composition of two or more clouds (private, community or public) that remain distinct entities but are bound together, offering the benefits of multiple deployment models. Hybrid cloud can also mean the ability to connect collocation, managed and/or dedicated services with cloud resources.

Hybrid cloud service as a cloud computing service that is composed of some combination of private, public and community cloud services, from different service providers.

A hybrid cloud service crosses isolation and provider boundaries so that it can't be simply put in one category of private, public, or community cloud service. It allows one to extend either the capacity or the capability of a cloud service, by aggregation, integration or customization with another cloud service. Varied use cases for hybrid cloud composition exist.

For example, An organization may store sensitive client data in house on a private cloud application, but interconnect that application to a business intelligence application provided on a public cloud as a software service.

This example of hybrid cloud extends the capabilities of the enterprise to deliver a specific business service through the addition of externally available public cloud services.

C.SECURITY AND PRIVACY

Cloud computing poses privacy concerns because the service provider can access the data that is on the cloud at any time. It could accidentally or deliberately alter or even delete information. Many cloud providers can share information with third parties if necessary for purposes of law and order even without a warrant. That is permitted in their privacy policies that users have to agree to before they start using cloud services. Solutions to privacy include policy and legislation as well as end users' choices for how data is stored. Users can encrypt data that is processed or stored within the cloud to prevent unauthorized access.

According to the Cloud Security Alliance, the top three threats in the cloud are "Insecure Interfaces and APIs", "Data Loss & Leakage", and "Hardware Failure" which accounted for 29%, 25% and 10% of all cloud security outages respectively — together these form shared technology vulnerabilities. In a cloud provider platform, being shared by different users there may be a possibility that information belonging to different customers resides on same data server. Therefore, Information leakage may arise by mistake when information for one customer is given to other. Additionally, Eugene Schultz, chief technology officer at Emagined Security, said that hackers are spending substantial time and effort looking for ways to penetrate the cloud. "There are some real Achilles' heels in the cloud infrastructure that are making big holes for the bad guys to get into".

Because data from hundreds or thousands of companies can be stored on large cloud servers, hackers can theoretically gain control of huge stores of information through a single attack — a process he called "hyperjacking".

There is the problem of legal ownership of the data (If a user stores some data in the cloud, can the cloud provider profit from it?). Many Terms of Service agreements are silent on the question of ownership.

Physical control of the computer equipment (private cloud) is more secure than having the equipment off site and under someone else's control (public cloud).

This delivers great incentive to public cloud computing service providers to prioritize building and maintaining strong management of secure services. Some small businesses that do not have expertise in IT security could find that it is more secure for them to use a public cloud.

There is the risk that end users don't understand the issues involved when signing on to a cloud service (persons sometimes don't read the many pages of the terms of service agreement, and just click "Accept" without reading).

This is important now that cloud computing is becoming popular and required for some services to work, for example for an intelligent personal assistant (Apple's Siri or Google Now).

Fundamentally, private cloud is seen as more secure with higher levels of control for the owner, however public cloud is seen to be more flexible and requires less time and money investment from the user.

V. MAIN THEME OF THE PAPER

The current process in the cloud is been time consuming. The main objective for this project is retrieving the documents of personal details like credit card, certificates and pan card so that the user can search and get the details of the required documents easily. The people who had lost their necessary documents and details they can pursue it easily within a time in the cloud. It gives the idea of a how a document can retrieve from cloud on android phones and computers. The platform independent language java and the php is used for the implementation. Along with SQLite and Cloud Database are used as databases for user and server respectively. All these things taken together can change the total concept of retrieving documents easily and effectively from cloud. Cloud computing

relies on sharing of resources to achieve coherence and economy of scale, similar to a utility (like the electricity grid) over an electricity network

VI. CONCLUSION:

The development of this document retrieving helps the people to store and retrieve the documents for the user whenever needed. This decreases their work of searching the documents like certificates, Mark sheet and credit card, PAN card etc., in case of emergency for the particular things.

REFERENCES

- [1] Q Chen ,Q Deng, "The cloud and its key techniques", The Journal of computer applications, 2009.
- [2] F. Dong, AB Song, "Architecture of the cloud" Journal of china institute, 2011.
- [3] Mipro, "The Cloud Security issues" 2009, The proceedings of the 33rd International convention.
- [4] R. Basmadjan, "The Issues of the Cloud computing", 2012 Springer
- [5] Gartner "7 cloud computing security risks" , july 2008.
- [6] Cloud security allience, security guidance for crital areas of focus in cloud computing, December 2009.
- [7] M. Wang and B. Li, "Network coding in live for peer-to-peer streaming," *IEEE Trans. Multimedia*, vol. 9, no. 8, pp. 1554-1567, 2007.
- [8] Fan, Y. Jiang, H. Zhu, and X. Shen, "An efficient privacy-preserving scheme against traffic analysis attacks in network coding," in *Proc. IEEE INFOCOM'09*, Rio de Janeiro, Brazil, Apr. 2009.
- [9] V. Oviya, Dr. S. Kirubakaran, K. Maheswari, "Enhancing Confidential Message Transferring System In A Security Framework" International Journal of

Applied Engineering Research, ISSN 0973-4562 vol. 10 no.29 (2015).

[10]. www.cloudsecurityallience.org

[11]. Dr.S.Kirubakaran,K.Maheswari, “An Improved SIP Protocol in Heterogeneous

Mobile Network for Efficient Communication”, Asian Journal of research in Social Sciences and Humanities, Vol.6.Issue.9,ISSN 2249-7315 Sep 2016.

IJARMET