

# Efficient Multi-Cloud Storage System

Tushar Wable, Mahesh Gavhane, Hemant Jadhav, Gaurav Deokar and Prof. Sanjay Agrawal

**Abstract**— Now a day's cloud computing plays a major role in IT Industries and Personal use. It is not used for business activities but it is now also used for educational purpose. Cloud computing is just internet based computing, you can say it is another property of internet. Customers plug into "cloud" and access applications, services which is priced and on demand. One of the prominent service offered in cloud computing is the cloud data storage, in that, subscribers do not have to store their data on their own servers, where instead their data will be stored on the Cloud Service Provider's servers. In cloud computing, subscribers have to pay the service providers for this storage service. This service does not only provides flexibility and scalability for the data-storage, it also provides customers with benefit of paying only for the amount of data they need to store for a particular period of time, without any concerns for efficient storage mechanism and maintainability issues with large amount of data storage.

**Index Terms**— Cloud Service Provider, Extensible Markup Language, Linear Programming, Quality of Service, Secure Hash Algorithm, Simple Object Access Protocol .

## 1 INTRODUCTION

CLOUD computing is the computing a resources use for hardware and software that are delivered as a service over a network. It is an abstraction for the complex infrastructure. Cloud computing have a key characteristics like a Agility, Application programming interface (API), Cost, Device and Location Independence, Virtualization, Multi-tenancy, Centralization, Reliability, Scalability and Elasticity, Performance, Security, Maintenance etc. And it shares its characteristics with Autonomic Computing, Client-Server Model, Grid Computing, Mainframe Computer, Utility Computing, Peer-to-Peer, Cloud Gaming etc. [8],[9].

integrity and cloud privacy are the most major issues that need to be consider in Cloud Computing. Even if they are standard infrastructure to ensure cloud users that they provide better data privacy and availability and other services.

In cloud computing, subscribers have to pay the service providers for this storage service. This service does not only provides flexibility and scalability for the data storage, it also provide customers with the benefit of paying only for the amount of data they need to store for a particular period of time, without any concerns for efficient storage mechanisms and maintainability issues with large amounts of data storage.

In addition to these benefits, customers can easily access their data from any geographical region where the Cloud Service Provider's network or Internet can be accessed. Since Cloud Service Providers (CSP) are separate market entities, data integrity and privacy are the most critical issues that need to be addressed in cloud computing. Even though the Cloud Service Providers have standard regulations and powerful infrastructure to ensure customer's data privacy and provide a better availability, the reports of privacy breach and service outage have been apparent in last few years. Also the political influence might become an issue with the availability of services.

## 2 SYSTEM REVIEW

Customer's stored data at cloud service providers is vulnerable to various threats. By referring previous studies [1],[2],[3], we discussed in detail that a Cloud Service Provider can be a victim to Denial of Service Attacks (DOS) or its variants. In our proposed work, we consider two types of existed threat models, which will affect the data availability and data security with respect to cloud data storage. These threat models are Single Service Provider and Colluding of Service Providers.

### 2.1 Single Service Providers

In this work we observed that [2],[3], from a customer's point of view, depend upon a solo SP for his outsourced data is not very promising. In this threat model, if the single point of fail-

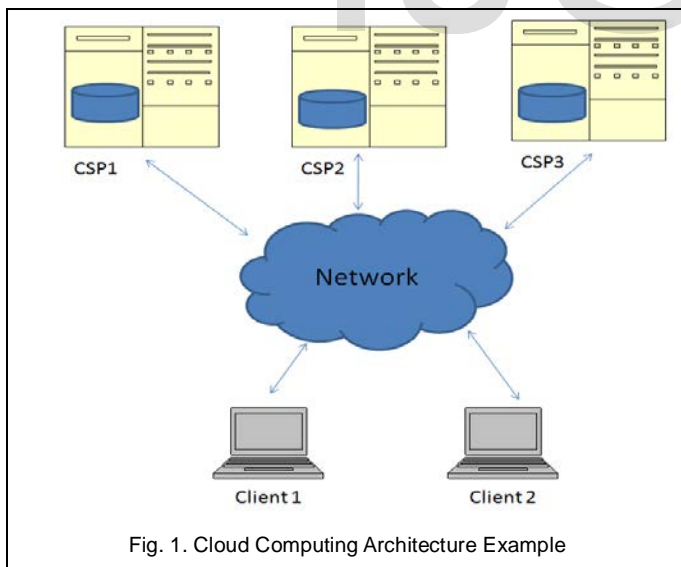


Fig. 1. Cloud Computing Architecture Example

Cloud Computing is simply a rate server. A huge amount of data being retrieved from geographically distributed data sources, and non-localized data handling requirements, creates such a change in technological as well as business model. One of the prominent services offered in cloud computing is the cloud data storage, in that, subscribers do not have to store their data on their own servers, where instead their data will be stored on the cloud service provider's servers. So the Cloud Service Providers (CSPs) are the separate market entities, data

ure then that could occur if a server at the cloud service provider failed or crashed, which make it hard for the customer to retrieve their stored data from a server. Availability of data is also an important issue which could be affected, if the Cloud Service Provider (CSP) runs out of business. Such worries are no more hypothetical issues. Therefore, a cloud service customer can not entirely depend upon a solo Cloud Service Provider to ensure the storage of their vital data.

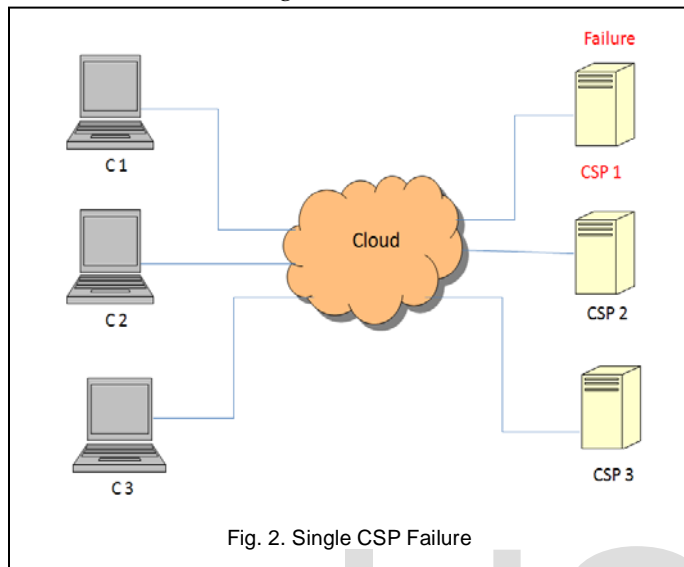


Fig. 2. Single CSP Failure

A Fig. 2. Single CSP Failure shows a customer C1, C2, C3 are cloud service users and they store their complete data or any kind of information on Cloud Service Providers CSP1, CSP2, CSP3 respectively. If a CSP1 is failed or bankrupt due to security attack or it's system failure or any other reason to cause failed, then customer C1 can loss his or her complete data and can't retrieved his or her outsourced data. Therefore this system model contains a problem of availability of CSPs. So customer can't entirely depend upon a solo or single Cloud Service Provider. Here is one solution that user store their data on multiple server by using distributing of data concept.

## 2.2 Colluding Of Service Providers

And now second existed threat model, in which the Cloud Service Providers might collude together to reconstruct and access the user stored data [1]. By providing the idea for distributing the data among two storage clouds such that, an adversary can't retrieve the contents of the data without having access to both the storage clouds [4]. Relying entirely upon a couple of service providers for the storage and retrieval of data might not be secured against colluding service providers. Such an attack scenario is entirely passive, because the cloud user can't detect that his information has been collectively retrieved from the service providers without his consent.

A Fig. 3. Colluding of Cloud Service Providers shows a customer store his or her data on one or two CSPs by distributing. A customer C1 divide his or her data into D1 and D2 part and store it on CSP1 and CSP2 respectively. Now here the problem is that if CSP1 and CSP2 are might collude with each other and exchange the parts of data that customer store on their server. If CSP1 and CSP2 are reconstruct whole data without being detected by the user, then CSPs will misuse of customer data.

Therefore, factors like trust and data privacies are need to be consider here.

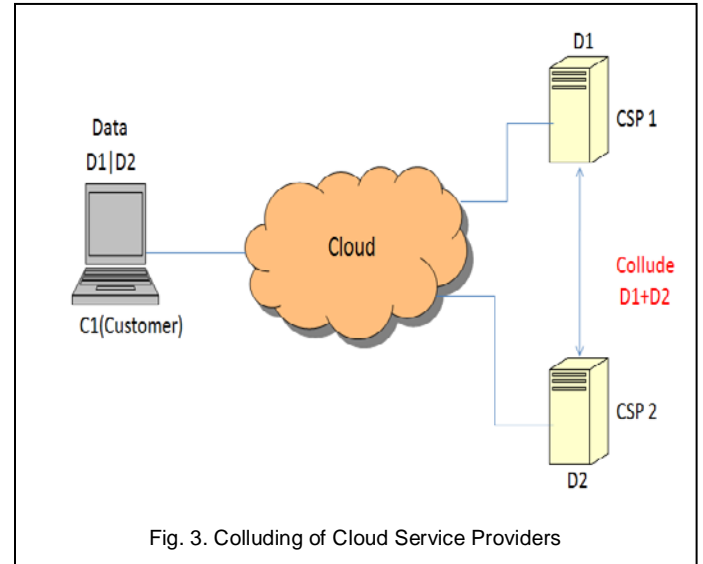


Fig. 3. Colluding of Cloud Service Providers

Here is solution is that customer will trust or sure on Third Party Auditor (TPA) [26]. So that users may resort to an independent Third Party Auditor (TPA) to audit their outsourced data whenever needed. The TPA has expert and capabilities that users don't have. TPA can periodically check the integrity of all the data stored on the cloud. It provides a much more easier and affordable way for the customer to ensure their storage correctness in the cloud.

## 2.3 Limitations of Existing Systems

1. This service does not only provide flexibility and Scalability for the data storage.
2. Data losses accrued.
3. Do not use cryptography technology so less security.
4. High cost for cloud maintains process.

## 3 PROBLEM STATEMENT

Efficient Multi-Cloud Storage System (EMCSS) model in Cloud Computing which holds an economical distribution. EMCSS provide customers with data availability as well as secure storage. And better decision for customers according to their available budgets and amount of data they need to store particular period of time.

## 4 SYSTEM DESIGN

### 4.1 System Architecture

The system design section is basically design on the bases of Infrastructure as a service (IaaS), Platform as a service (PaaS), Software as a service (SaaS), which is most useful services of Cloud Computing [6],[7]. Using software as a service, users also rent application software and databases. And the cloud providers manage the infrastructure and platforms on which the applications run.

We store the data among different number of CSPs, that's by using the idea of distributing [4],[5], we not only solve the problem of availability of CSPs, but also solve the issue of security for data storage. Here we consider the storage services for cloud data storage between two entities, cloud users (U) and Cloud Service Providers (CSP). The cloud storage service is generally priced on two factors, how much data is to be stored on the cloud servers and for how long the data is to be stored. According to that the cost of CSP's are decide for storing data on CSP.

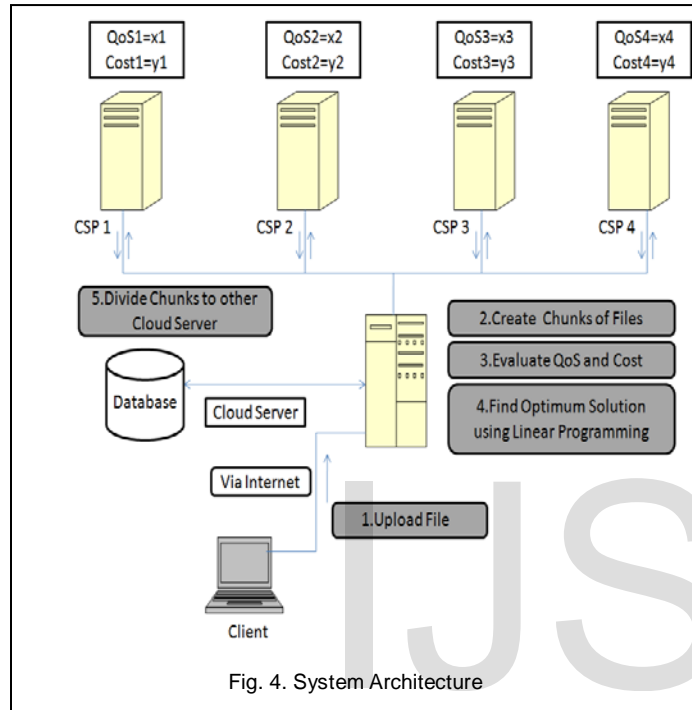


Fig. 4. System Architecture

A Fig. 4. System Architecture shows that a customer store their data using Third Party Auditor (TPA) on different Cloud Service Providers. The TPA has own Cloud Server and Database to store customers all information like customer authentication and what authorities are a customer have, what kind of data store on different CSPs, how long time data store on CSPs and other privacy information related to client. And TPA also contains the all information related with different CSPs like a Quality of Service provided by CSP, availability of CSPs for storing, outsourcing and retrieving a data from CSPs etc. and other status related with different CSPs. So, TPA is one of best solution for customer to trust for storing data on multiple CSPs as efficiency of security, privacy, availability of CSPs and affordable budget.

We consider  $p$  number of Cloud Service Providers (CSPs), each available Cloud Service Provider is associated with a Quality of Service (QS) factor, along with its cost of providing storage service per unit of stored data ( $C$ ). Every CSP has a different level of Quality of Service (QS) offered as well as a different cost associated with it. Hence, the cloud user can store his data on multiple SPs according to the required level of security and their affordable budgets. System application is contain some following modules:

1. Client Side Application Module
2. Server Side Application Module

3. Storage and Database Handler Module
4. Distribution Module etc.

**Client Side Application Design:** User or a client application contains a GUI Application. User must have to login first. The authentication is done by using Secure Hashing Algorithm (SHA). After successfully login the user request to upload a file through Java I/O technology. Similarly user download file using Java I/O technology from the Cloud Server. While knowing the status of number of Cloud Service Providers (CSPs), a verification of data is done by user, the data which is uploaded and downloaded to from different CSPs by connecting to Cloud Storage.

The communication among the user application and server is done by using Simple Object Access Protocol/Extensible Markup Language (SOAP/XML) [11],[13]. SOAP is a protocol specification and use for exchanging structured information in the implementation of Web Services for a computer networks. It is depends or use the Extensible Markup Language (XML) for its message format, and usually relies on other Application Layer protocols, like Hypertext Transfer Protocol (HTTP) and Simple Mail Transfer Protocol (SMTP), for message negotiation and transmission.

**Server Side Application Design:** The availability and security problems with data storage are solved in server side application design, the availability of Cloud Service Providers and Security [1], for store data. Some security issues are also handles through a client application design section.

On the server side, server also communicate and give response to client through SOAP/XML technology by using (GlassFish ) Webserver. Server accepts a data and split a data into a chunks or blocks. After splitting a data server apply Linear Programming (LP) concept to evaluate Quality of Service (QS) and Cost of each CSPs, then retrieving optimum solution using Linear Programming. And finally divide the chunks of data on different Cloud Service Providers.

#### 4.2 Step to Store Data On CSPs

1. User authentication is done for login purpose.
2. User request for upload data via internet.
3. User gets information from cloud server to know their status.
4. Application verifies data which is uploaded by user.
5. Create chunks of files for uploaded data.
6. Evaluate QS and cost of CSP.
7. Find optimum solution using linear programming.
8. Divide chunks to other cloud server.
9. Finally user successfully store data on different CSPs.

### 5 TECHNICAL SPECIFICATIONS

#### 5.1 Mathematical Model

We are used here a Linear programming concept [15],[16],[17]. It can be applied to various fields of study. It is used in business and economics, but can also be utilized for some engineering problems. In the Industries section that use linear programming models include transportation, energy, telecommunications, and manufacturing. It has proved useful in modeling diverse types of problems in planning, routing, schedul-

ing, assignment, and design. Linear Programming is an algorithmic and computational tool for captures the first order effects of various system parameters that should be optimized, and it is essential to the engineering optimization. It has been widely used in various engineering disciplines that analyze and optimize real world systems, such as - flow control, power management of data centers, packet routing etc.

By using a Linear programming (LP, or linear optimization) is a mathematical method for determining a better way to achieve the best outcome (such as maximum profit or lowest cost) in a given mathematical model. Linear programming is a technique for the optimization of a linear objective function [19]. It is a subject to linear equality and linear inequality constraints. The following is the mathematical equation for overall system application:

$$S = \{F, M, C, CC, CL, QS\} \quad (1)$$

Where,

S= System

F= Set of File {f1, f2, f3, ..., fn}

M= Mapping Function

C= Set of Cloud Service Provider {CSPi ∈ C}

CC= Cost of Cloud Service Providers

CL= Storage Limit of CSP

QS= Quality of Service

$$\{CSP1, CSP2, \dots, CSPn\} = M \{F, C, CC, CL, QS\} \quad (2)$$

Where M is calculate by Linear Programming

The objective of a system is,

Minimize [total cost (CC)] and Maximize [QS]

Or

Maximize [QS – total cost (CC)]

## 5.2 Advantages

1. Without any concerns for efficient storage mechanisms and maintainability issues with large amounts of data storage.
2. Cloud data storage also redefines the security issues targeted on customer's outsourced data.
3. Using cryptography technology for data base security.
4. Less cost and cost based on client requirements.
5. Easy to maintains large databases with security.
6. Avoid database losses.

## 6 CONCLUSION

This system application is try to show the importance of Data Storage Availability and Data Storage Security Management in Multi-Storage Cloud Computing. This proposed work evaluates secure data outsourcing for widely applicable large-scale systems using Linear Programming (LP). By dividing and distributing customer data, our application is show its ability of providing a customer with a secured storage under their affordable budget.

## REFERENCES

- [1] J. Du, W. Wei, X. Gu, T. Yu, "RunTest: assuring integrity of dataflow processing in cloud computing infrastructures", *In Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security (ASIACCS '10)*, ACM, New York, NY, USA, 293-304.
- [2] N. Gruschka and M. Jensen, "Attack surfaces: A taxonomy for attacks on cloud services", in *Cloud Computing (CLOUD), 2010 IEEE 3rd International Conference on*, pp. 5–10, July 2010.
- [3] M. Jensen, J. Schwenk, N. Gruschka, and L. L. Iacono, "On Technical Security Issues in Cloud Computing", in *IEEE International Conference on Cloud Computing, (CLOUD II 2009)*, pp. 109–116, Bangalore, India, September 2009.
- [4] P. F. Oliveira, L. Lima, T. T. V. Vinhoza, J. Barros, M. M'edard, "Trust-ed storage over untrusted networks", *IEEE GLOBECOM 2010*, Miami, FL, USA.
- [5] A. Shamir, "How to share a secret", *Commun. ACM* 22, 11(November 1979).
- [6] "The NIST Definition of Cloud Computing" (<http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>). National Institute of Science and Technology. Retrieved 24 July 2011.
- [7] Voorsluys, William, James; Broberg, Buyya, Rajkumar (February 2011). "Introduction to Cloud Computing" ([http://media.johnwiley.com.au/product\\_data/excerpt/90/04708879/0470887990-180.pdf](http://media.johnwiley.com.au/product_data/excerpt/90/04708879/0470887990-180.pdf)). In R. Buyya, J. Broberg, A.Goscinski. *Cloud Computing: Principles and Paradigms*. New York, USA: Wiley Press. pp. 1–44. ISBN 978-0-470-88799-8.
- [8] "Cloud Computing Portal and Scholarly Resources on Academic Room". Online at <http://www.academicroom.com/physical-sciences/information-technology/cloud-computing>.
- [9] P. Mell and T. Grance, "Draft NIST working definition of cloud computing", Master's thesis, June 2009. Online at <http://csrc.nist.gov/groups/SNS/cloudcomputing/index.html>.
- [10] W. Itani, A. Kayssi, and A. Chehab, "Privacy as a Service: Privacy-Aware Data Storage and Processing in Cloud Computing Architectures", *Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing*, December 2009.
- [11] W3C XML homepage. Online at <http://www.w3.org/XML/>
- [12] XML 1.0 Specification. Online at <http://www.w3.org/TR/REC-xml>.
- [13] W3C SOAP page. Online at <http://www.w3.org/TR/soap/>
- [14] SOAP Version 1.2 specification. Online at <http://www.w3.org/TR/soap12/>
- [15] Guidance On Formulating LP Problems. Online at <http://people.brunel.ac.uk/~masfjb/jeb/or/lp.html>.
- [16] The Linear Programming FAQ. Online at <http://lpsolve.sourceforge.net/4.0/LinearProgrammingFAQ.htm>.
- [17] 2009 Linear Programming Software Survey (Online at <http://www.lionhrtpub.com/orms/surveys/LP/LP-survey.html>) - *OR/MS Today*.
- [18] Benchmarks For Optimization Software. Online at <http://plato.asu.edu/bench.html>.
- [19] "A Modern Language for Mathematical Programming". Online at <http://www.ampl.com>.
- [20] Amazon.com, "Amazon s3 availability event: July 20, 2008", 2008. Online at <http://status.aws.amazon.com/s3-20080720.html>.
- [21] M. Arrington, "Gmail Disaster: Reports of mass email deletions", December 2006. Online at <http://www.techcrunch.com/2006/12/28/gmail-disasterreports-ofmass-email-deletions/>.
- [22] A. Cavoukian, "Privacy in clouds", *Identity in the Information Society*, December 2008.
- [23] R. Gellman, "Privacy in the clouds: Risks to privacy and confidentiality from cloud computing", *Prepared for the World Privacy Forum*, Feb 2009. Online at

<http://www.worldprivacyforum.org/pdf/WPF-Cloud-Privacy-Report.pdf>.

- [24] The Official Google Blog, "A new approach to China: an update", March 2010. Online at <http://googleblog.blogspot.com/2010/03/new-approach-to-china-update.html>.
- [25] W. Itani, A. Kayssi, and A. Chehab, "Privacy as a Service: Privacy-Aware Data Storage and Processing in Cloud Computing Architectures", *Eighth IEEE International Conference on Dependable, Autonomic and Secure Computing*, December 2009.
- [26] C. Wang, Sherman S.-M. Chow, Q. Wang, K. Ren, W. Lou, "Privacy preserving public auditing for secure cloud storage", in *InfoCom2010, IEEE*, March 2010.
- [27] B. Krebs, "Payment Processor Breach May Be Largest Ever", January 2009. Online at [http://voices.washingtonpost.com/securityfix/2009/01/payment\\_processor\\_breach\\_may\\_b.html](http://voices.washingtonpost.com/securityfix/2009/01/payment_processor_breach_may_b.html).
- [28] S. H. Shin and K. Kobara, "Towards secure cloud storage", *Demo for CloudCom2010*, December 2010.

IJSER