An Efficient and Secure Protocol for Ensuring Data Integrity in Multi Cloud Environment

Shravan S Kamath,
BE, Department of Computer Science,
B T L Institute of Technology and Management,
Bengaluru, India.

Deepak Sai V
BE, Department of Computer Science,
Gopalan College of Engineering and Management,
Bengaluru, India

Sharath S V
BE, Department of Computer Science,
B T L Institute of Technology and Management,
Bengaluru, India.

***Abstract*- Cloud computing is an internet based computing which enables obtaining resources (hardware, software, platform and services) from the internet on a scalable basis. . Many users place their data in the cloud, so correctness of data integrity and security are prime concerns. Security issue is one of the major concern in cloud computing. To maintain our data in cloud, it may not be fully trustworthy because client doesn't have copy of all stored data and hence maintaining the integrity of the data is an important and integrated aspect of cloud computing. In this paper we propose a framework that would provide integrity of data of multiple users through Third Party Auditor (TPA) and proposes different set of algorithms based on the sensitivity of the data. In the proposed framework concept of multi cloud platform has been used to provide best cost optimization for various requirements of user. Finally we have implemented different algorithms for the various platforms in the proposed framework. In the obtained results, we can find that the proposed framework is easy to implement and provides better performance by using the traditional algorithms of network security over the various issues of cloud computing.**

Key-Words- AES, Bcrypt, Cloud computing, data integrity, RSA, security, data management in cloud, Data integrity, TPA.

Terminologies-

- MAS - Multi agent system. It is a technique where several agents communicate with each other.
- PDP - Prove able data possession. It checks that a file which consists of a collection of n blocks is retained by the outsourced storage site.
- QOS - Quality of service. It defines the degree to which a provided activity promotes customer satisfaction.

I.INTRODUCTION

Cloud Computing is the use of Internet for the tasks performed on the local machine, with the hardware and software demands maintained elsewhere. It represents a different way to architect and remotely manage computing resources.

The only thing the user's computer needs to run is the cloud computing systems' interface software, which can be as simple as a Web browser and the cloud's network takes care of the rest. Cloud computing is being driven by many which includes Google, Amazon and Yahoo as well as traditional vendors including IBM, Intel and Microsoft. The data should be available in the cloud for it to be accessed. There are four main types of cloud storage:

1) Mobile cloud storage: Mobile cloud storage stores the individual's data in the cloud and provides access to the data from anywhere.

2) Public cloud storages: There is no connection between the enterprise and storage service provider and the cloud resources are stored separately from the enterprise's data center.

3) Private cloud storage: In private cloud storage, the storage provider has infrastructure in the enterprise's data center that is typically managed by the storage provider.

4) Hybrid cloud storage: It is a combination of public and private cloud storage where some critical data resides in the enterprise's private cloud while other data is stored and accessible from a public cloud storage provider [18].

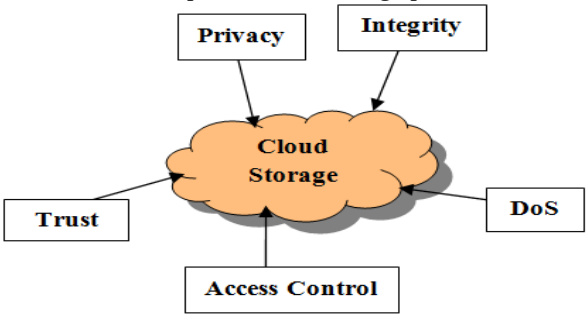

Fig. 1 Challenging areas to focus in cloud computing

As multiple users / parties access information on cloud, the integrity and privacy of the information stored is at risk. Cloud provides distribution of data over computers. When data is sent by the user to be processed in the cloud; the control of the data is given to a remote party that may not address security concerns of the user. As a user has no physical access to the data, he is unaware about the location of his data and is not sure whether the integrity of his data is maintained or compromised in cloud. It is important to ensure that the information being processed on cloud is secure and no tampering of information is done when previously unknown parties may be present. A framework is proposed to provide data integrity using TPA to guarantee the various users that their data is unaltered.

The rest of the paper is classified as follows: Section 2 discusses the literature review, Section 3 describes the proposed model to provide integrity of data, Section 4 describes the algorithm used, Section 5 shows the analysis of the algorithms used and Section 6 presents the conclusion for the work done till now.

II. LITERATURE SURVEY

Cloud computing is an emerging trend in the field of technology. There are various issues related to cloud computing, major ones being the security and integrity of data. Many frameworks have been designed and many algorithms have been proposed to resolve such issues.

Nirmala et al. [1] proposed a new scheme to resolve integrity problem by introducing user authenticator to audit and check the integrity of data. Their research focused on providing solutions to all issues of cloud computing and to develop a model that would provide secure cloud infrastructure which would help to adopt the cloud as and when required.

Raju et al. [2] introduced a protocol for integrity checking of cloud storage that would provide integrity protection of user information. This protocol supports public verifiability and is evidenced to be secure against associate un-trusted server. It's additionally non-public against third-party verifiers. Attas and Batrafi proposed an integrity checking model over cloud with help of TPA using DSA algorithm.

To resolve the problem of privacy in the clouds Metri and Sarote [3] introduced a threat model. Juels et al. [4] described a formal "proof of retrievability" (POR) model for ensuring the remote data integrity. It uses error correction code in order to check that the data is correct or not. Talib et al. [5] proposed layer architecture based on MAS architecture having two main layers:

- cloud resource layer [cloud server side]
- MAS architecture layer [cloud client side].

The MAS architecture has two agents namely Cloud service provider agent [CSPA] and Cloud data integrity backup agent [CDIBA]. This layered architecture collectively was called "Cloud Zone". A prototype of proposed "Cloud Zone" would be designed using Prometheus Methodology and implemented using the Java Agent Development Framework Security (JADE-S). Ateniese et al. [6] considered public audit ability in their defined "provable data possession" model to verify if the client's data is stored at untrusted server. Homomorphic Verifiable Tags are used for data auditing. The model samples random sets of blocks from the server and generates probabilistic proofs of possession which reduces I/O costs. The client verifies the proof by maintaining a constant amount of metadata. The transmission of a small, constant amount of data is done by the response protocol which minimizes network Communication.

Bhosale et al. [7] provides a 3 dimensional framework along with digital signature and RSA algorithm where the user will upload the data over cloud based on the various security levels. Protection ring 1 will provide high level of security, ring 2 will provide less security and ring 3 will provide least level of security. Security of cloud is enhanced by using this framework with RSA and DSA algorithm combination.

III. PROPOSED NEW SYSTEM

This section defines the proposed framework to provide data integrity in multi cloud system. Proposed framework is shown in Figure 3 and has following three mai0n roles:

A. Client (User): It is a network entity that stores data on the cloud server and relies on it for the maintenances and storage of the data.

B. Cloud Service Provider (CSP): It is the cloud server that provides significant storage space, resources and maintenance for user data. In the block diagram, two more blocks are present, Storage Server and the Backup server.

C. *Third Party Auditor (TPA)*: TPA is an entity that has knowledge and expertise that client does not possess. It is responsible for data integrity verification and works on behalf of the client.

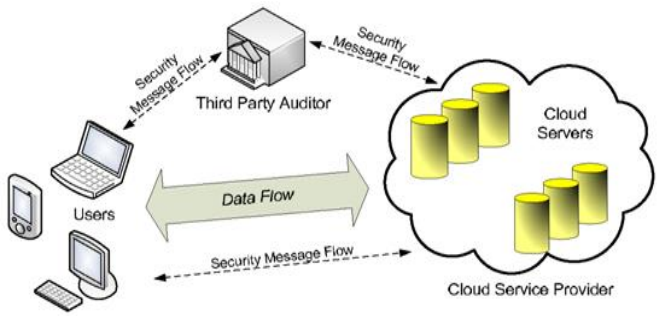

Fig. 2 System design model

There are many cloud service providers and each of them provides different storage plan along with different QoS parameters so it becomes a tough task for users to keep moving their data from one cloud to another based on QoS and cost optimization. In the proposed model concept of multi cloud is used to provide best cost optimization for various requirements of user. To give a clear design of the model, use of connectors are labelled as a, b and c.

Depending on the type of data to be stored on various clouds, there are three main platforms in the model namely:

A. Platinum- Sensitive data will be stored here like data related to transactions of atm, bank account information etc. The data will be stored on private cloud.

B. Gold- Data related to simple login on any page like facebook, ebooking and email login is stored. The level of security needed is not that high. Security only on password is required.

C. Silver- Data related to only simple browsing of sites, uploading of images, downloading of files like downloading of music files or images is stored. The level of security needed is the least.

Also we use encryption algorithm for protecting the integrity and authenticity of the given data.

V. PROPOSED ALGORITHM

In this section we have implemented various algorithms like RSA algorithm, Bcrypt algorithm, and AES algorithms to implement the proposed framework.

In module m1 RSA algorithm is used to provide integrity of data because for storing sensitive information on cloud, hashing algorithms are used. RSA is based on the difficulty of factoring large numbers. There are various advantages of RSA due to which it is preferred over DSA.

- DSA can only be used for authentication while RSA can be used for both authentication and to encrypt a message.
- Faster at encrypting than DES.

RSA: This is an Internet encryption and authentication system that uses an algorithm developed in 1977 by Ron Rivest, Adi Shamir, and Leonard Adleman. The RSA algorithm is the most commonly used encryption. Till now it is the only algorithm used for private and public key generation and encryption.

In module m2 Bcrypt algorithm is used for hashing the passwords. A password hashing algorithm should preferably be slow in order to prevent brute force attacks; it should have features which actually decrease the feasibility of a distributed brute force attack on the hashes. Bcrypt algorithm is derived from the Blowfish block cipher which uses look up tables that are initiated in memory to generate the hash.

In module m3 AES algorithm is used to provide security on the data stored. AES is asymmetric encryption algorithm in which to encrypt the message sender uses public key of receiver and its private key is used by receiver to decrypt the message.

- AES is preferred over DES algorithm as it is more secure.
- AES data encryption is mathematically more efficient and elegant cryptographic algorithm.
- Block size of DES is small compared to AES

AES: (Advanced Encryption Standard), is the new encryption standard recommended by NIST to replace DES. Brute force attack is the only effective attack known against it, in which the attacker tries to test all the characters combinations to unlock the encryption. Both AES and DES are block ciphers. It has variable key length of 128, 192, or 256 bits; default 256. It encrypts data blocks of 128 bits in 10, 12 and 14 round depending on the key size. AES encryption is fast and flexible.

Algorithm:
/* Variables used for:
User => u,
Platinum => p,
Gold => g,
Silver => s */

Begin
If *u chooses p*
Then *call module m1*
If *u chooses g*
Then *call module m2*
If *u chooses s*
Then *all module m3*
End

Module m1:
Begin

1. *User's data to be stored on cloud*
2. *Data is encrypted using RSA*
3. *Data is verified by CSP using RSA*
4. **If** *data is valid*

Go To Module T
Else
Corrupted data
End

Module m2:
Begin

1. *Data stored on cloud*
2. *Data is encrypted*
3. *Data is verified by CSP using Bcrypt algorithm.*

4. **If** *data is valid*

Go To Module T

Else

Intrusion on data

End

Module m3:

Begin

1. Data is stored on cloud

2. Data is encrypted

3. Verification of data is done by CSP using AES

4. **If** *data is valid*

Go To Module T

Else

Invalid data

End

Module T:

Begin

1. Check the data stored.

2. **If** *proof = direct then*

Report = direct access

Else

Return {1, 0}

1: if integrity of data is verified as correct

0: if integrity of data verified is incorrect

End

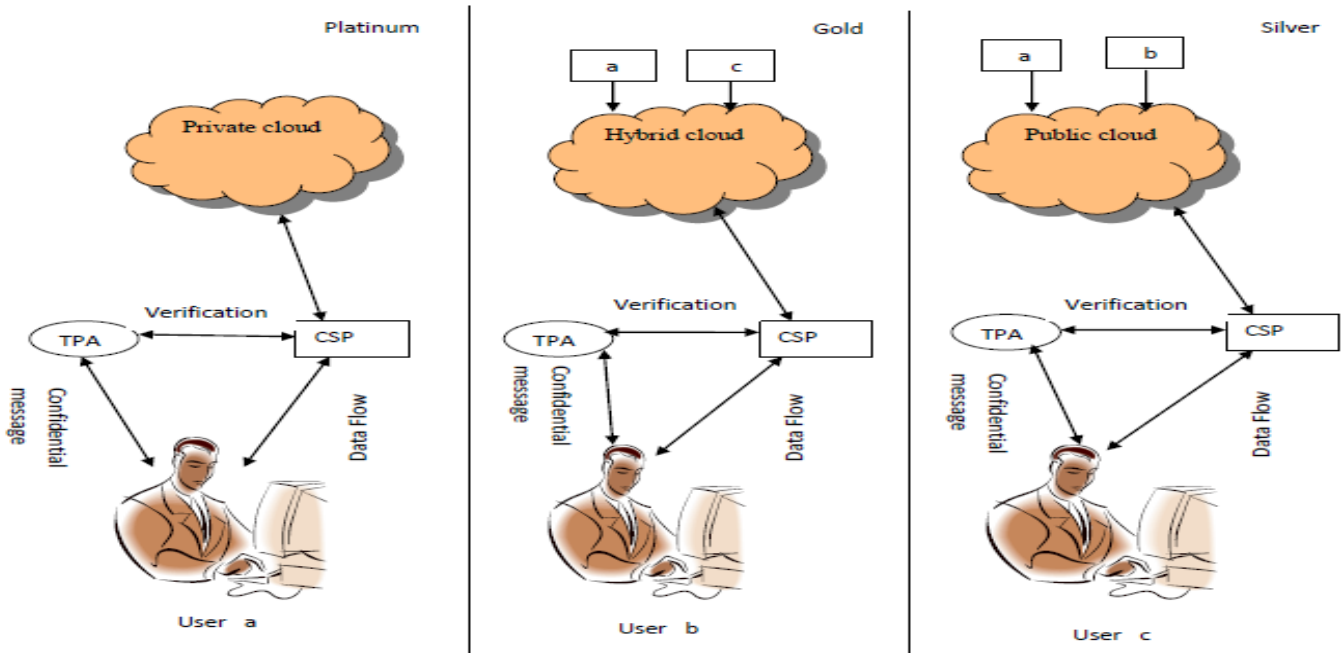

Fig. 3 The proposed system

V. RESULT ANALYSIS

To estimate the output of the programs, matlab tool is used.

A. RSA algorithm analysis:

RSA algorithm is tested for integer numbers ranging from a single digit message length to 16-digit message length. The execution time t is in seconds. The execution time depends on the values of p and q which are prime numbers. Different values of p and q are taken and depending on these values graph between message length and time are plotted as shown in Figure 4(a) and Figure 4(b).

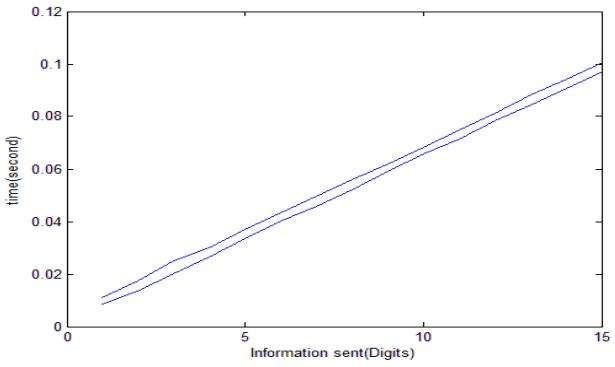

Fig. 4a Message length vs. Time for p=3 and q=7

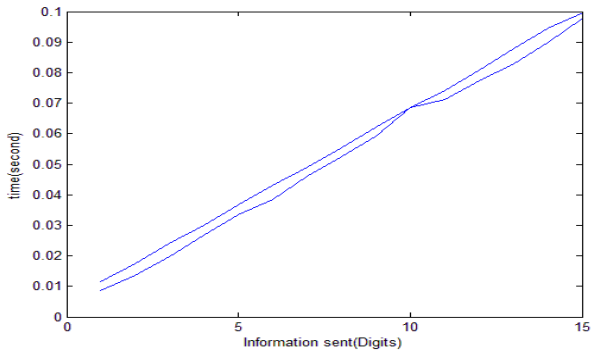

Fig. 4b Message length vs. Time for p=23 and q=17

B. AES algorithm analysis:

AES is used here to provide integrity to data while simple browsing of internet. Plain text is encrypted to hexa decimal format. The change in graph depends on the value of plain text. By using combination of alphabets and digits in a text the time taken increases. A graph between information sent and time is plotted which can be seen in the Figure 5.

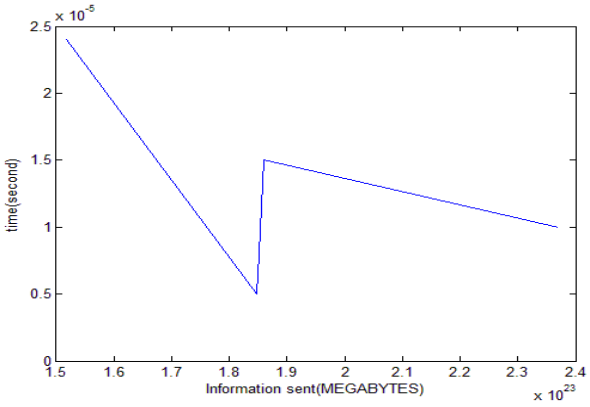

Fig. 5 AES Algorithm

VI. ACTIVITIES OF THE CLOUD

A. Devoted Property Assertion

All resources have high-quality assurance in its multi cloud servers with the help of strong platform. The mentioned possessions of RAM, CPU, and Bandwidth capacity in network produce sufficient atmospheres for its concerned usage. Thus the activity of could server can‘t ever opposite to the current world while considering its level of integrity and security even though there is some little bit of conflict that we want yet to recover from its already existing level.

B. Excess Data Storage

All cloud server storage resources are managed by high-performance and high-availability storage area network. Many cloud solutions operate on local disks from the host system, which means any computing or storage failure can result in down time and potential data loss. As cloud servers are autonomous, if there happens any server crack in stored data, these can be protected against internal and external attacks.

C. Accurate Choice of CSP

To get excellent service from multiple servers, good service providers are important to be considered and selected. So much care must be taken in this respect so that the CSP itself can be elastic with clients in order to get accessed with all places (anywhere and anytime).

VII. CONCLUSION AND FUTURE WORK

In this paper proposed framework resolve the issue of integrity of user data with better performance using the traditional algorithms of network security. Cost is also optimized using multi cloud concept and different platforms for various categories of the users. In our future work we will implement the hybrid algorithms using this framework. We will also explore the new opportunities for security in multi cloud environment.

ACKNOWLEDGEMENT

Authors would like to acknowledge the effort of internet experts who gave us vast amount of knowledge when going through a research on the multi cloud environment. We also are thankful to our computer science department head and faculty members in motivating us to carry out this paper effectively.

REFERENCES

[1] Nirmala V., Sivanandhan R.K., and Lakshmi R.S. “*Data confidentiality and Integrity Verification using Authenticator scheme in cloud,”* Proc. of 2013 International Conference on Green High Performance Computing,

[2] Raju et al. “*Data Integrity using Encryption in Cloud Computing,”* Journal of Global Research in Computer Science, vol. 4, no. 5, pp. 40-43, May 2013.

[3] Metri P. and Sarote G., “*Privacy Issues and Challenges in Cloud computing*,” International Journal of Advanced Engineering Sciences and Technologies, vol. 5, no. 1, pp. 5-6, 2011

[4] A. Juels and B. S. Kaliski *“PORs: Proofs of retrievability for large files*,” Cryptology ePrint archive, June 2007. Report 2007/243.

[5] A. M. Talib, R. Atan, and R. Abdullah, “*CloudZone: Towards an Integrity Layer of Cloud Data Storage Based on Multi Agent System Architecture*,” 2011 IEEE Conference on Open Systems (ICOS2011), September 25 - 28

[6] Giuseppe Ateniese, Randal Burns, Reza Curtmola, Joseph Herring, Lea Kissner, Zachary Peterson, and Dawn Song “*Provable Data Possession at Untrusted Stores*,” Proc. of the 14th ACM conference on computer and communications security,

[7] Pradeep Bhosale, Priyanka Deshmukh, Girish Dimbar, and Ashwini Deshpande, “*Enhancing Data Security in Cloud Computing Using 3D Framework & Digital Signature with Encryption*,”

[8] Framework to Improve Data Integrity in Multi Cloud, Anandita Singh Thakur, P. K. Gupta, International Journal of Computer Applicati