ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

Third Party Authorized Auditing of Dynamic Big Data on Cloud with Fine Grained Upgrades

Shubham M. Burghate, Akshay C. Dighe, Rahul S. Ghyar, Akash G. Chourange, Prof. Yogesh Thorat

B.E Student Dept. of Computer Engineering, Dr.D.Y.Patil, School of Engineering, Pune, India

B.E Student Dept. of Computer Engineering, Dr.D.Y.Patil, School of Engineering, Pune, India

B.E Student Dept. of Computer Engineering, Dr.D.Y.Patil, School of Engineering, Pune, India

B.E Student Dept. of Computer Engineering, Dr. D.Y.Patil, School of Engineering, Pune, India

Assistant Professor, Dept. of Computer Engineering, Dr. D.Y.Patil, School of Engineering, Pune, India

ABSTRACT: Cloud computing is a method of providing a set of shared computing resources that includes applications, computing, storage, networking, development, and deployment platforms as well as business processes. In cloud computing, data while transferring as well as storage, and get data back when it is needed there is no assurance about the security of that data stored and also it is not changed by the cloud or TPA. Therefore Security and control over data remain to play a significant role in plans for cloud computing initiatives. Existing research work earlier grant data integrity to be verified, but still there are different drawbacks, firstly basic and mostly needed authorization/authentication process is not present in between Cloud Service provider and TPA. Techniques like authentication and encryption are important, our system can provides those things for security issues. Second: in recent research POR protocol in which the verifier stores only a single cryptographic key, POR can only capable of detection of file corruption or loss, and not prevention. Maintaining the storages can be a tough task and second it requires high resource costs for the implementation. This paper, Propose a formal analysis method called full grained updates. It includes the well-organized searching for downloading the uploaded file and also emphases on designing the auditing protocol to advance the server-side protection for the efficient data confidentiality and data availability.

KEYWORDS: Cloud computing, data security, Big Data, provable data control, authorized auditing, fine-grained dynamic data update, TPA (Third Party Authenticator).

I. INTRODUCTION

Cloud service providers provide an enterprise-class infrastructure that offers a scalable, secure and reliable environment for users, at a much lower cost due to the sharing nature of resources. It is best way for users to use cloud storage services to share data with others in a cluster, as data sharing becomes a standard feature in most cloud storage scenarios.

Compared to traditional systems, scalability and elasticity are key advantages of cloud. As such, efficiency in supporting dynamic data is of great importance.

When you move to the cloud, time spent maintaining hardware and upgrading software is significantly reduced—eliminating problems with it. Now your IT team can focus on advancing your organization's technology, more rather than being a repair service. Plus, you will have more time to spend improving business operations and beginning agile creativity. Instead of spending more and more portions of your investment budget on servers used for email storage and loads, you can think strategically and support business managers in a much more agile fashion, quickly responding to their needs [8]

But many enterprises are dragging their feet, worried about security. Definitely, sharing data over immense networks outside the business's perimeter does raise big questions: Who has access to your data? How can you establish trust among users and the services with which they interact? How can you control it? Data security is one of

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

the major concerns in the acceptance of cloud computing [8][9].Required process of Authentication/Authorization is not present in between the cloud service provider and auditor. i.e. anyone who is willing to challenge the cloud service provider to acquire the integrity of a certain file, thereby puts the quality of so termed 'auditing-as-a-service' at risk. The recent research work that was carried out on BLS signature can support updates full dynamic data on constant/fixed size of data blocks, this support is only towards fixed size blocks as the basic unit which is Coarse-grained updates. Due to which every small update would cause re-computation and updating of the authenticator for a whole file block, which results in overheads like higher usage of storage space and communication overheads. In this Project, I would permit a formal analysis for all possible types of fine-grained updates and bring out a system that can fully provision authorized auditing and fine grain update requests.

II. RELATED WORK

Cloud computing promises incredible benefits in terms of speed, flexibility, and cost [9].**RalphC.Merkle**[1] proposed new scheme "Digital signature based on a conventional encryption function such as DES is described which is more secure as the basic function. Previous work of Ralph C.Merkle in 1982 "Secrecy, Authentication, and public key systems these all rely on conventional encryption functions like one-way-function. But no single among them are much succeed in providing the convenience of the system based on the more complex mathematical problem. **Suthan and Kesavaraja**[4] proposed scheme "Granule based File Storage System with Secure Transparent Availability" by Suthan and Kesavaraja [4] suffering from drawbacks such as it only provide a security until the intruders get hold of the file. "Digital signature based on conventional encryption function"[1] by Ralph C. Merkle which is symmetric encryption technique"Dynamic Big Data Storage on Cloud with Efficient Verifiable Fine-grained Updates using Secure Erasure Code-Based Cloud Storage System"[8] which uses the concept of proxy servers but it's limited to certain file types. "Proofs of retrievability for Large Files"[2] by Ari Juels provide checks for regular retrievability but it can only protect the static archived files. In any of this application public audibility and variable sized data block are not supported by default. Compared to conventional systems, scalability and flexibility are key advantages of cloud [5], [6], Cloud users may also want to split big datasets into smaller datasets and store them on different physical servers for reliability, privacy-preserving or efficient processing purposes. Among the most vital problems linked to the cloud is data security/privacy [1], [8]. It has been one of the most recurrently raised concerns. There is a lot of work trying to enhance cloud data security/privacy with technological approaches on CSP side.The idea of **POR** and its first model was proposed by Jules et al. [2]. But that system can only be applied to static data storage such as records or files. As such, good organization in behind dynamic data is of great importance. Protection and privacy protection on dynamic data has been considered extensively in the past [3], [7]. The drawback in their scheme can only be functional to static data storage such as archive.

III. PROPOSED ALGORITHM

A. *Design Considerations:*

To achieve fine-grained data updates different operations are carried out. Block level operation in fine-grained data updates contains 5 types of operations.

- Partial Modification (PM):- It is ease now to update a consecutive part of a certain block.
- Whole Block Modification (M):- And also the whole block is replaced by new set of data.
- Block Deletion (D):- The whole block is deleted from the tree structure.
- Block Insertion (I):- A whole block is inserted to carry new set of data
- Block Splitting (SP):- Some part of the existing block is taken out and the new block is created to be added to the tree structure.

B. *Description of the Proposed Algorithm:*

The aim of the proposed algorithm is to update a certain portion of the data block client has to adopt the PM process i.e partial modification. This involves the following steps:

Step 1: client compose an Update Request and send to CSS and CSS run the Perform Update (UpdateRequest,F) algorithm.

Step 2: CSS send the P update to client and client to run the Verify Update (Pk; P update) algorithm.

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

Step 3:if the algorithm returns true value it updates the old value with new ones otherwise if it returns false then client send the Update request again.

IV. PSEUDO CODE

Step 1: Challenge, Verification and Proof generation.

This is the last step in verifying TPA's authenticity. In this phase, TPA has to show that it is the genuine one who is challenging the CSS for data integrity checking. TPA runs the GenChallenge() algorithm with private key and signature as parameters. Then a challenge message is generated with TPA's new ID selected randomly from the set of total blocks. This VID is encrypted with CSS's public key. After this process, TPA can send challenges to CSS.

Step 2: When CSS get the challenges it will run another algorithm to verify the signature, VID and client's public key.

If the algorithm returns a true,

Then CSS will send a proof "p" to TPA and TPA will run the algorithm

Verify (pk, challenge, p)

Else

If the algorithm run by CSS returns false,

End: the request is rejected.

Step 3: For TPA authorization, a signature scheme is chosen which can be forged by malicious TPAs. No malicious TPA cannot make the CSS respond to its challenge which contains an integrity proof for a subset of existing file in CSS.

Step 4: Fine-Grained Data Updates

We can define the fine-grained update request as the set of 3 variables,

1) Starting address of the update in the file

2) The length of the file and

3) The new message to be inserted into the File

Condition: For this update, we have to guarantee that the new information added is not exceeding the maximum block size after updating.

Step 5: Prepare for Authorization:

The client asks (her/his choice of) TPA for its ID VID (for security, VID is used for authorization only). TPA will then return its ID, encrypted with the client's public key.

Step 6: The CSS will construct an RMHT for verification which should be identical to the tree spawned at the client side.

Step 7: go to step 3.

Step 8: End.

V. SYSTEM ARCHITECTURE

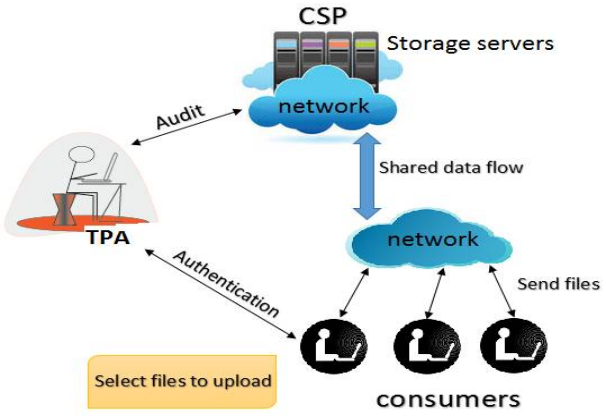

Fig 1. System architecture

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

To enable the TPA efficiently and securely verify shared data of users, our scheme should be designed to achieve following properties:
Public Auditing: The third party auditor is able to verify the integrity of shared data without retrieving the entire data. It informs the user when any unauthorized entity tries to steal his data on the cloud. **Correctness:** The third party auditor is able to correctly detect whether there is any corrupted block in shared data. **Enforceability:** Only a user which is valid can generate valid verification information on shared data. Also, **Identity Privacy:** During auditing, the TPA cannot distinguish the identity of the signer on each block in shared data.

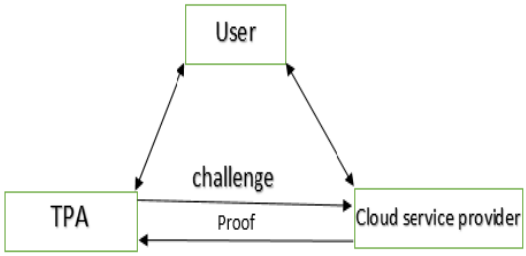

Fig 2 : Verification Of TPA

Participating parties: As Shown in fig.1 **Consumers**: There are different types of consumers. A cloud services consumer might be a specific, or a small business team. Sectors in large companies can be cloud services consumers. The IT department can use cloud services also in addition to existing data center services, or to provide specific cloud-based applications like customer relationship management to the sales team. **Service providers**: Cloud service providers are enterprises that offer packaged facilities to consumers. Many different types of providers range from those who offer services to entities and those that assist a broad set of elements. Many service providers focus on certain markets or certain types of workloads so they can optimize their offerings inexpensively.
TPA: Fig.3.TPA, in cloud computing environment, data while transferring as well as storage, and get data back when it is needed there is no assurance about the security of that data stored and also it is not changed by the cloud or TPA. Even a company that provides cloud services to consumers may use third-party cloud services to supplement their ability.

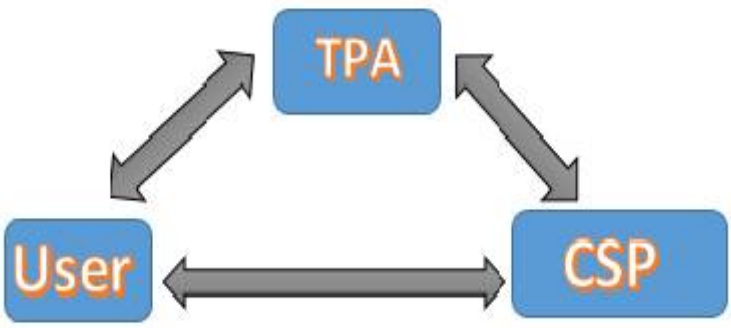

Fig.3 TPA in cloud enviornment

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

VI. SIMULATION RESULTS

AES Algorithm[Advanced **E**ncryption **S**tandard (AES)] is a symmetric key block cipher published by the NIST in December 2001. AESencrypts and decrypts a data block of 128 bits. The key size can be 128, 192, 256 bits.
The number of round: 10 rounds for 128 bits
12 rounds for 192 bits
14 rounds for 256 bits

Verifiable Fine-Grained Data Operations: To update a certain portion of the data block customer has to implement the PM process i.e partial modification. This involves the following steps:
First, client composes an Update Request and send to CSS and CSS run the PerformUpdate (UpdateRequest,F) algorithm Secondly CSS send the Pupdate to client and client run the VerifyUpdate (Pk,; Pupdate) algorithm .
Finally, if the algorithm returns true value it updates the old value with new ones otherwise if it returns false then client send the Update request again.

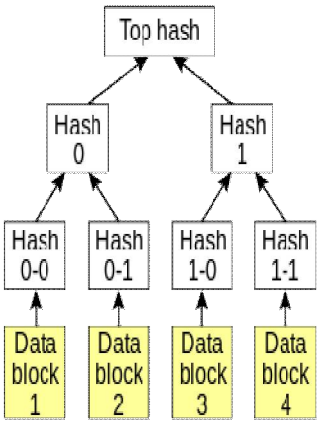

Fig.4 MHT

As shown in fig.4 MHT:
Files data is divided into equal data blocks. In advance storing the data blocks on cloud servers, user pre-computes the verification tokens. These tokens are used to check the integrity of data stored on cloud servers. Also these tokens are used to locate the cloud server on which data has been changed by the attacker. Before data division and dispersing file consumer generates tokens on individual data blocks. When consumer wants to check the correctness of the data, he sends the file identifier to the cloud servers. User may send challenge on particular data block also. Upon receiving challenge token, each cloud server computes the token on the data blocks, and sends them to client.

Merkle Hash Tree:
The Merkle hash tree, the root hash beside with the total size of the file and the piece size are now the only information in the system that needs to come since a trustworthy source. A consumer that has only the root hash of a file can check any piece as follows. It heads computes the hash of the piece it received. Merkle hash tree block that can support full-grained update requests.
Specially, the server.

- Replaces the chunk
- Replaces outputs and
- Replaces the hash functions.

The usage of MHT, the block size of the child nodes is equally divided and the user can update their chosen block, Which they want sinceall the block size are same.

Partition Algorithm

- Load the Input file and size.
- Check size of file

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

If file size is invalid then declare as Invalid size.
Else
Count size = S
Split file into partitions with extension and index value.
Return files.

Merging Algorithm

- Collect all decrypted file partitions
- Check file status
  If (file!) then File is missing.
  Else

Count the index value
Merge files.
Return file.

VII. CONCLUSION AND FUTURE WORK

We have implemented a formal analysis on possible types of fine-grained data updates and proposed a scheme that can fully support authorized auditing and fine-grained update requests. Based on the contributions of this paper on better data auditing, we plan to additional investigate the next step on how to develop other server-side protection methods for efficient data security with effective data confidentiality and availability. Also, we also plan to investigate audit ability-aware data scheduling in cloud computing. For providing more security we are using TPA (Third party authenticator). The proposed applications support a large number of regular small updates such as applications in social media and business transactions.

ACKNOWLEGEMENT

We would like to thank Dr.D.Y.Patil School of Engineering for providing us with all the required amenities. We would thank our guide Prof. Yogesh Thorat sir for giving us all the help and guidance we needed. We are also grateful to Prof. S. S. Das, Head of Computer Engineering Department, DYPSOE, Lohegaon, Pune for their indispensable support, suggestions and motivation during the entire course of the project we would like to thank Dr.D.Y.Patil School of Engineering for providing us with all the required amenities. We would thank our guide Prof. Yogesh Thorat sir for giving us all the help and guidance we needed. We are also grateful to Prof. S. S. Das, Head of Computer Engineering Department, DYPSOE, Lohegaon, Pune for their indispensable support, suggestions and motivation during the entire course of the project.

REFERENCES

[1] Digital signature based on conventional encryption function" by Ralph C. Merkle in 1998.
[2] Proofs of retrievability for Large Files" by Ari Juels and Burton S. KaliskiJr in 2007.
[3] Ensuring Data Storage Security in Cloud Computing" by Cong Wang, Qian Wang, Kui Ren and Wenjing Lou in 2012
[4] Granule based File Storage System with Secure Transparent Availability " bySuthan and Kesavaraja in 2011.
[5] Cloud Computing and Emerging IT Platforms: Vision, Hype, Reality for Delivering Computing as the 5th Utility,'' Future Generation Computation System," by. R. Buyya, C.S. Yeo, S. Venugopal, J. Broberg, and I. Brandic in June 2009.
[6] A View of Cloud Computing,'' by M.Armbrust, A.Fox,R.Griffith, A.D.Joseph, R.Katz,A.Konwinski, G.Lee, D.Patterson, A.Rabkin, I.Stoica, and M.ZahariaCommun in April 2010.
[7] Scalable and Efficient Provable Data Possession," by G. Ateniese, R.D. Pietro, L.V. Mancini, and G. Tsudik, in 2008.
[8] The Secure Cloud: Best Practices for Cloud Adoption (Semantec)
[9] 10 Myths about Moving to the Cloud (Microsoft Office 365).
[10]Secure Distribution of File on Cloud Niyamat I. Ujloomwale, RanjanaBadreDept. of Computer, MIT Academy of Engineering, Alandi, SavitriPhule Pune University, Pune, India

ISSN(Online) : 2320-9801
ISSN (Print) : 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

BIOGRAPHY

Shubham M. Burghate, Akshay C. Dighe, Rahul S. Ghyar, Akash G. Chaurange,we student in computer engineering,Dr.D.Y.Patil School Of Engineering,SavitribaiPhule University. We are now pursuing bachelor in engineering (BE) degree -2016 from SPPU Pune.

Mr. Yogesh. A. Thorat Assistant Professor B.Tech (Computer Science & Engg.), M.E. (Computer Engg.)

Selected Publications:
1. “Optimization of Univesity Course timetabling using Genetic Algorithm: A Survey”. In 2011 International IEEE Conference, Kanyakumari, India.
2. Paper Titled “Review on Multimode Resource Constrained Project Scheduling Problem” International Journal of Computer Science & Engineering Technology (IJCSET), VOLUME 3 ISSUE 5, May 2012 - IJCSET (ISSN: 2229-3345)
3.Paper Titled “REVIEW ON AUDITING OF DYNAMIC BIGDATA ON CLOUD WITH FINED GRAINED UPGRADES” Multidisciplinary Journal of Research in Engineering and Technology, Volume 2, Issue 4, Pg.737-742
Our research interest is cloud computing emerging technology (cloud storage).