



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

## Verification of Authorized Auditing Accuracy Using Genchallenge Method

Revatthy Krishnamurthy<sup>1</sup>, K.P. Kaliyamurthie<sup>2</sup>

<sup>1</sup>M.Tech Scholar, Department of Computer Science & Engineering, Bharath University, Chennai, India

<sup>2</sup>Professor, Department of Computer Science & Engineering, Bharath University, Chennai, India.

**ABSTRACT:** Cloud computing provides various elastic and scalable IT services and users can reduce the huge capital investments in their own IT infrastructure. Cloud computing falls into three broad categories such as software as a service(SaaS), infrastructure as a service(IaaS) and platform as a service(PaaS). Third party auditor (TPA) allows to ensure the correctness of data on behalf of the cloud client and to verify the integrity of the data stored in the cloud. In this paper, we provide one of the TPA algorithms such as Genchallenge for accuracy of an authorized auditing and verification. The results demonstrate that our scheme can offer an enhanced security and flexibility and also significantly lower overhead for big data applications.

**KEYWORDS:** Cloud computing, Big data, Third party Auditor, authorized auditing.

### I. INTRODUCTION

Cloud computing is an internet based computing which enables sharing of services. It is a large group of interconnected computers and it provides services as a software, an infrastructure and a platform [1]. Cloud is a provider and provides resources by the service provider [2]. The main advantage of cloud is cost savings and the disadvantage is security. Although current development of cloud computing is rapid, but some hesitations to use cloud is still exist. In cloud computing, data security / privacy is one of the major concerns in the adoption [3] [4].

To ensure the correctness and the integrity of data stored in the cloud, the third party auditor (TPA) is used. There are two categories of auditing such as private and public. Private auditing can achieve higher scheme efficiency and the public auditing allows anyone, not just the client but its storage keeping no private informations [5]. The TPA will audit data and help data owners to make secure that their data are safe in the cloud and management the data will be easy and less burdening to data owner.

The Merkle Hash Tree (MHT) has been studied from two decades. RMHT is an extended Merkle hash tree with ranks. It is similar to binary tree and each node has maximum of two child nodes. Keygen, FilePreproc, Challenge, verify, Genproof, Performupdate and verifyupdate are some of the algorithms which are supported the public auditing scheme in the aim of supporting variable - size data blocks and authorized third - party auditing. In this paper, Genchallenge algorithm is described and how the integrity of data will be stored on CSS and which is verified by TPA. For testing the audit process, weather report data set is taken and the output results obtained are recorded. In a remote verification scheme, the cloud storage server (CSS) cannot provide a valid integrity proof of a given proportion of data to a verifier unless all this data is intact. It is especially recommended that data auditing is to be conducted on a regular basis for the users who have high-level security demands over their data.

The organization of the paper is as follows. Section II presents the review of related work. Genchallenge algorithm is explained in section III and an experimental results discussed in section IV and section V concludes the paper.

### II. RELATED WORK

Cloud computing provides reliable services to built on virtualized computer and storage technologies [7]. Cluster is a type of parallel and distributed system i.e a collection of inter-connected stand-alone computers working together as a single integrated computing resource [8][9]. Clouds are commodity computers and network attached storage. The size / scalability is 100s to 1000s computers in the cloud. The node operating system of cloud is a hypervisors (VM) on which multiple operating systems run. In cloud both centralized and decentralized inter-operability web services are SOHP and REST. Potential for building third party or value added solutions in cloud is high potential and can create



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

new services by dynamically provisioning of compute, storage and applications services and offer as their own isolated or composite cloud services to users. Internetworking in cloud is high potential and third party solution providers can loosely tie together and service. The third party auditor (TPA) introduced to ensure that data integrity and same the cloud user's computation resources as well as online burden and it is of critical importance to enable public auditing service for cloud data storage. Public auditability has been proposed ensuring remotely stored data integrity under different system and security models [10] [11] [12] [13].

The public key based on homomorphic linear authenticator which enables TPA to perform the auditing without demanding the local copy of the data and thus drastically reduces the communication and computation overhead as compared to the straightforward data auditing approaches.

### Formal Analysis

Aim of our proposed scheme is supporting variable - sized data blocks, authorized third - party auditing. Our scheme is described in three parts.

1. Setup: Initially the client generates keying materials via keygen and fileproc, then upload the data to CSS. The previous schemes such as Merkle Hash Tree (MHT) is used as metadata. But in this proposed scheme, metadata is stored in Ranked Merkle Hash Tree (RMHT). The client will authorize the Third party - Auditor (TPA) by sharing a value  $sig_{AUTH}$ .

2. Verifiable Data Updating: CSS performs the clients request and then the client update to check whether CSS has performed on both the data blocks and their corresponding authenticators (used for auditing) honestly.

3. Challenge, proof generation and verification: Describes how the integrity of the data stored on CSS is verified by TPA via Genchallenge, Genproof and verify. We now describe our scheme in detail as follows:

### Setup:

The phase is similar to the existing BLS based schemes except for the segmentation of file blocks. Let

$e: G \times G \rightarrow G_T$  be a bilinear map where

$G$  is a GDH group supported by  $Z_p^2$ .  $H: (0,1)^* \rightarrow G$  is a collision - resistant hash function, and  $h$  is another cryptographic hashfunction. In the proposed scheme, TPA runs the Genchallenge algorithm. In our setting, TPA must show CSS that it is indeed authorized by the file owner before it can challenge a certain file.

## III. GENCHALLENGE ALGORITHM

Genchallenge (Acc, pk,  $sig_{AUTH}$ ): According to the accuracy required in this auditing, TPA will decide to verify  $c$  out of the total  $l$  blocks. Then, a challenge message

$Chal = (S_{sig_{AUTH}}, (VID)_{PK_{CSS}}, (i, v_i)_{i \in [1, l]})$  is generated where VID is TPA's ID,  $l$  is a randomly selected  $[1, l]$  with  $C$  elements and  $v_i \in Z_p$  are randomly chosen coefficients. VID is encrypted with the CSS's public key  $PK_{CSS}$  so that CSS can later decrypt  $(VID)_{PK_{CSS}}$  with the corresponding secret key. TPA then sends Chal to CSS.

### 2. After receiving chal, CSS will run the following algorithm:

Gen Proof (Pk, F,  $sig_{AUTH}$ ,  $\phi$ , chal) : Let  $w = \max (s_i)_{i \in [1, l]}$ . CSS will first verify  $sig_{AUTH}$  with AUTH,  $t$ , VID and the client's public key  $spk$ , and reject if it fails. CSS will compute  $\mu_k = \sum_{i \in [1, l]} v_i m_{ik}$ ,  $k \in [1, w]$  and  $\sigma = \prod_{i \in [1, w]} (H(mi), \Pi_{i \in [1, l]} Dsig)$ , then output  $p$ . During the computation of  $\mu_k$  and let  $m_{ik} = 0$  if  $k > s_i$ . After execution of this algorithm, CSS sends  $P$  to TPA.

## IV. EXPERIMENTAL RESULTS

This experiment is conducted using Cloudera - a cloud computing operating system and LINUX customized operating system are used. Red Hot Linux 18.2 version is used for server and workstation applications. Program coding is implemented in Java 7.0 version and mysql5.4 version and upon virtualized data centers, hadoop facilitates the Map Reduce programming model and distributed file systems. Moreover, we installed virtual Machine 9.0 version with 64-bit software. The output results are shown below.



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 4, April 2015

```
Applications Places System cloudera@localhost:~/Desktop/Demo
File Edit View Search Terminal Help
User from : 127.0.0.1
Enter Username:
demo
Enter Password:
Confirm Password:
Enter eMail:
demo@gmail.com
User Registration - Success

Welcome to HadoopApp
Choose Option to perform:
Login - 1 | Registration - 2 | Exit - 0

1
Enter Username:
demo
Enter Password:
Login Success

Choose Operation: Data Process - 1 | Exit - 0
1
Enter number of inputs :
2
Enter Input0:
audit-input-1
Enter Input1:
audit-input-2
Enter Output:
audit-output
15/03/11 02:03:28 WARN mapred.JobClient: Use GenericOptionsParser for parsing the arguments. Applications should implement Tool for the same.
15/03/11 02:03:31 INFO input.FileInputFormat: Total input paths to process : 2
15/03/11 02:03:37 INFO mapred.JobClient: Running job: job_201503110017_0004
15/03/11 02:03:38 INFO mapred.JobClient: map 0% reduce 0%
```

```
Applications Places System cloudera@localhost:~/Desktop/Demo
File Edit View Search Terminal Help
[cloudera@localhost Demo] cd /home/cloudera/Desktop/Demo/
[cloudera@localhost Demo] ls
data1.csv execution_steps.txt hadoop-csv-jar-with-dependencies.jar Hadoop File system run command- mysql_steps.txt-
data2.csv execution_steps.txt- Hadoop File system run command mysql_steps.txt
[cloudera@localhost Demo] ls -ltr
total 25652
-rw-rw-r-- 1 cloudera cloudera 388 Dec 6 05:12 Hadoop File system run command-
-rwxr-xr-x 1 cloudera cloudera 124619 Jan 3 23:52 data1.csv
-rw-rw-r-- 1 cloudera cloudera 388 Jan 24 06:35 Hadoop File system run command
-rw-rw-r-- 1 cloudera cloudera 348 Feb 15 04:17 mysql_steps.txt-
-rw-rw-r-- 1 cloudera cloudera 354 Feb 15 04:17 mysql_steps.txt
-rwxr-xr-x 1 cloudera cloudera 122271 Mar 1 00:02 data2.csv
-rwxr-xr-x 1 cloudera cloudera 25989847 Mar 1 01:42 hadoop-csv-jar-with-dependencies.jar
-rw-rw-r-- 1 cloudera cloudera 427 Mar 11 01:27 execution_steps.txt-
-rw-rw-r-- 1 cloudera cloudera 427 Mar 11 01:27 execution_steps.txt
[cloudera@localhost Demo]
[cloudera@localhost Demo]
[cloudera@localhost Demo] hadoop fs -copyFromLocal data1.csv audit-input-1
[cloudera@localhost Demo]
[cloudera@localhost Demo]
[cloudera@localhost Demo] hadoop fs -copyFromLocal data2.csv audit-input-2
[cloudera@localhost Demo]
```

## V. CONCLUSION

In this paper, we have provided a formal analysis and the proposed scheme that can fully support authorized auditing. Based on our scheme, we have also proposed a modification that can dramatically reduce communication overheads. Theoretical analysis and experimental results have demonstrated that our scheme can offer not only enhanced security and flexibility, but also significantly lower overheads for big data applications.



# International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 3, Issue 4, April 2015**

## REFERENCES

- [1].Customer Presentation on Amazon Simmid Australia, Syney, 2012, accened on March 25, 2013. (Online) Available:  
<http://sed.amazon.com/apac/awssummit.au/>
- [2].P. Mell and T. Grance, "Draft Nist working definition of cloud Computing".
- [3].J. Yao, S.Chan, S. Nepal, D. Levy, and J. Zic, "Trust store : Making Amazon S# Trustworthy with services Composition, in Proc. 10th IEEE/ACM Intsymposium on cluster, cloud and Grid Computing (CCGRID)", 2010, PP.600 - 605.
- [4].D.Zissis and D.Lekkas, "Addressing cloud computing security Issue", Future Gen.Compert. Sust., Vol.28, no.3, PP.583-592, March.2011
- [5].G.Ateniese et al., - Provable Data Possession at instructed stores, I Proc. ACM CCS.07, Oct.2007, PP.598-609.
- [6].S. Napal, S. Chen, J.Yao, and D. Tilakanathn, "Diaas : Data Integrity as a service in the cloud," in Procx. 4th international conference on cloud computing, 2011, P.P.308-315.
- [7].A. Weiss, Computing in the clouds, networker 11(4) (2007) 16-25.
- [8].G.P.P. Fisher, in search of Clusters, 2<sup>nd</sup> ed., Prentice Hall, Upper Saddle River, USA 1998.
- [9].Ans :R.Buyya (Ed), High Performance Cluster computing : Architectures and systems, Vol.1, Prentice Hall, Upper Saddle River, USA 1999.
- [10].G.Ateniese, R. Burns, R. Curtmola, J. Herriney, L., Kissor, Z. Petersa, and D.Song, "Provable data possession at entrusted stores," in Proc.of CCS'07,Alexandina, VA, October 2007, PP.598-609.
- [11].Q.Wang, C.Wang, J.Li.K.Cen and W.Lon, "Gnabling Public verifiability and data dynamics for storage secuity in cloud computing," in Proc.ofESDRICS'09, Volume 5789 od CNCS. Spinger - verlag, Sep.2009, PP.355- 370.
- [12].A. Juels and J. Burton S. Kal;iski, "Pros: Proofs of irretrievability for large files," in Proc.of CCS' of, PP.584-597.
- [13].H. Shacham and B. Waters, "Compact proofs of irretrievability," in Proc. of Asia crept 2005, vol.5350, Da 2008, PP.90-107.