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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

Implementation of Block Level File Operations with the Improvement of File Security on Cloud

Keerthana M M

Assistant Professor, Dept. of CSE., ATME College of Engineering, Mysuru, Karnataka, India

ABSTRACT: Today enormous amount of data are generating from individuals and organization in a rapid rate. The data rate is increasing every day and there is a need to maintain them in a right way. The data generated by the organization, business industry or government agency are stored and managed by an external storage provider called cloud storage server (CSS) which have significant data storage and computing power. Users outsource their data to the cloud. Users of cloud cannot physically maintain direct control over their data, this makes data security one of the important concern of using cloud. The communication between CSS and cloud user is controlled and monitored by an optional external agent called Third Party Auditor. First, we are checking the authenticity of the TPA, since it can be malicious. Second, Data on the cloud may change according to the application, in dynamic data storage more overhead is involved where data needs to be updated.

KEYWORDS: cloud computing, datasecurity, authorized auditing, fine-grained dynamic data update.

I. INTRODUCTION

Cloud computing describes the means of delivering any and all information technology components from computing power to computer infrastructure, applications, business processes and collaboration offering IT as a service. Cloud computing is an emerging style of computing where data, applications and resources are provided to users as services over the web. Today technology is changing every day and organizations are expected to adopt changes and transform enterprise IT. Cloud computing provides a shift in the way companies see the IT infrastructure. This technology is driven by the internet and it needs rapid provisioning, scalability, and virtualized environments. The cloud in cloud computing is the set of software, hardware, networks, storage, services. And interfaces that combine to deliver aspects of computing as a service. Information, software and shared resources are given on demand to computers. It allows clients to do things on computer without the need for them to buy and build an IT infrastructure. Enterprises adopt cloud models to improve employer productivity, deploy new products and services faster and reduce operating costs. Today cloud services are delivered in a user friendly manner. The payment model pay as you go elements the need for upfront investment. Clouds are location independent and provide abstract versions of datacenter that are not bound to a specific datacenter. Cloud carries the basic infrastructure characteristics that are helpful to deploy cloud service in a fast and cost effective way. Cloud services today are delivered in a user friendly manner and offered on an unprecedented scale. Clouds will transform the IT industry. They will profoundly affect how we live and how businesses operate. Cloud infrastructure service components include physical infrastructure and management software.

II. RELATED WORK

With the invent of cloud computing, the users can store their data remotely on cloud and can use shared pool of configurable computing resources. Users can outsource their data on to the cloud and they can be free from the data maintenance and data storage. When user outsource huge amount of data on to the cloud, protecting the integrity of the data is a challenging operation. Allowing public auditor for cloud data storage security is important. User has the right to challenge an audit party to check the integrity of the data. External audit party is called Third Party Auditor (TPA).

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

The TPA will help the user in auditing their data. To allow TPA in a secure manner, the TPA should not ask for a copy of user data, TPA should not create vulnerability to user private data. Cloud computing offers flexibility to users. Users will pay as much as they use on cloud they don't have to put up large computers because the operations are handled by the cloud service provider (CSP), user will transfer their data to the CSP and CSP has control over the user data. User should make sure that data is correct on the cloud. Data may be lost due to internal (employee at CSP) or external hackers. Sometimes, CSP might behave unfaithfully. It can hide data loss to protect their reputation or CSP might delete data that is rarely accessed.

[1] C. Wang, Q. Wang, K. Ren and W. Lou " Privacy Preserving Public Auditing for Data Storage Security in Cloud Computing", in Proc. 30st IEEE Conf. on computation and communication.(INFOCOM), 2010, pp.1-9.

To check the integrity of the data it is not practical to simply download the data .TPA can provide the audit report. The TPA should not read the content of the data.Techniques used: we have three participating parties -user, CSP and TPA.KeyGen: It is run by the user to set up the scheme .SigGen: To generate verification metadata users will use this technique. This may consist of Message Authentication Code (MAC), information used for auditing.GenProof: To audit the proof from cloud server this algorithm is run by the TPA.Drawbacks: user computes the MAC of every block. Transfers the file blocks and codes to cloud and shares the key with the TPA. The TPA will demand the random number of blocks and their associated code from the CSP. To verify the correctness of the data TPA will uses the key. Auditor will demand retrieval of the users data which is not privacy preserving.The key can be used only once.The TPA has to keep the state; remembering which key has been used.Communication and computation complexity. This scheme is good for only static data where data does not change in the cloud.[1]

[2] G.Ateniese, R.B. Jhons, R.Curtmola, J.Herring, L.Kissner, Z.Peterson and D. Song "Provable Data Possession at Untrusted Stores", in Proc. 14 ACM Conf. on Computation and Communication. Security (CCS), 2007, pp. 598-609.

The user can verify the data that is stored on untrusted server without retrieving it. The model called provable Data possession (PDP) [2] , generates proofs by sampling the random set of blocks from the server. This will reduce I/O costs. The client will maintain a definite amount of metadata to examine the proof. A small amount of data is transmitted by response protocol which will minimize the network communication. This PDP model for checking the remote data supports huge data sets in various distributed storage systems. We define a model for PDP that provides a proof that a file is stored by the third party. This model is unique which permits the server to access the file in small portions for generating the proof. Other techniques access the entire file. Techniques used: Uses Homomorphic verifiable Tags (HVTs) which are building blocks for the PDP schemes. Block less verification- using HVTs , server can construct a proof that allows client to verify if the server possesses certain file blocks, even when the client does not have access to actual file blocks. Drawbacks: Does not support full data dynamics. [2]

[3] A. Juels and B.S. Kaliski Jr., "PORs: Proof of Retrievability for Large Files", in Proc. 14th ACM Conf. on Computation and Communication. Security (CCS), 2007, pp. 584-597.

A POR is a type of protocol in which a server will give proof to the client that a selected file F is intact, the client can regain all of F from the server with high probability [3]. As PORs incur lower communication complexity than transmission of F itself, they are attractive for building high –assurance remote storage systems. Cloud computing encompasses full spectrum of computing resources, this paper is focusing on archival or backup data, large files subject to infrequent updates. Users may access such files only sporadically. Techniques used: client will encode the file before it is being transmitted in to the storage for archiving. POR enables bandwidth efficient challenge, response protocols to guarantee probabilistically that a file is obtainable at a remote storage provider. This paper uses the technique of spot-checking in the challenge-response protocol to detect the adversarial behavior. In each of the challenge, a subset of file blocks is sampled and the results of its computation over these blocks are returned to the client. The POR model helps in checking of file retrievability of a secret key by a single client in possession, any client are allowed to verify the retrievability of file without the secret key. Dynamic file updates in these protocols are of great challenge. Drawbacks: This scheme is only applicable to static data storage such as archive or library and it does not support full data dynamics. [3]

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

III. PROPOSED METHODOLOGY

To overcome the disadvantages associated with the existing schemes, new scheme was proposed where it offers security and enables dynamic updates. The focus is on small dynamic updates which will benefits the efficiency and scalability of cloud storage server. The existing scheme also aims at providing security in auditing unauthorized requests. This new scheme uses a special hash tree algorithm called Ranked Merkle Hash Tree (RMHT) and uses a flexible data segmentation strategy. To make the scheme more robust and secure security problem is also addressed in supporting the public verifiability. This is achieved by adding authorization among the three participating parties .The client, the CSS and TPA.

The contributions of the proposed scheme are as summarized below:

- Analyzing different types of dynamic fine-grained data update requests.
- To better support fine-grained update requests, BLS signature and RMHT algorithms are used.
- Compared to the existing schemes, the size of file blocks are not restricted. This will offer scalability and also flexibility to the existing schemes.
- To eliminate threats or unauthorized audit challenges from malicious TPA we use a special authorization process. Which will find out whether the TPA is authorized or not.
- Focus on improving the efficiency in verifying file security.

IV. SYSTEM ARCHITECTURE

There are three participating parties:

1. Client
2. Cloud storage server (CSS)
3. Third party auditor (TPA)

Relationships between the three parties are shown in Fig.1.

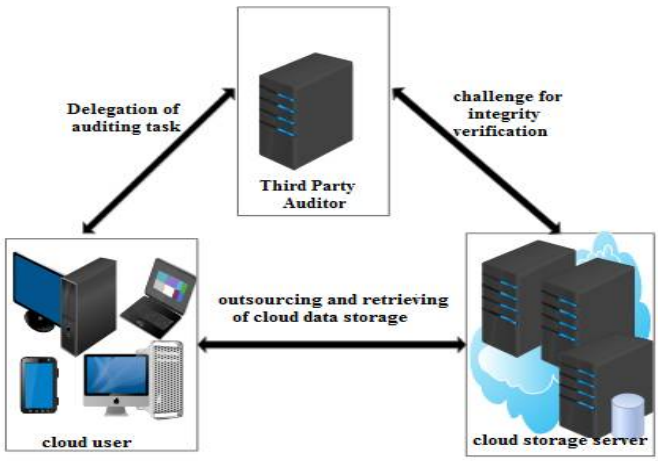

Fig.1: Relationship between Client, CSS and TPA

Third party auditor (TPA) and cloud server are semi-trusted to the client. In order to get proof for certain file blocks any clients can send a challenge message. Distributed Denial of Service Attacks (DDOS) can be sent by a malicious party by sending the challenges from multiple users which will cause extra overheads at the server side and network congestion. Privacy issue can be easily obtained from any adversary from the integrity proof returned by CSS. Adversary can gather statistical information by challenging multiple times from CSS.

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

i. Bilinear Map

Assume a group G is a gap Diffiehelman (GDH) group with prime order P. A bilinear map is constructed as：GXG→GT where GT is a multiplicative cyclic group with prime order. The ‘ e ’ should pose these properties:

- Bilinearity - ∀m,n∈G→e(m^a,m^b)=〖e(m,n)〗^ab
- Non-degenaracy-∀m∈G,m≠0→e(m,n)≠1 and

‘e’ should be efficiently computable.

ii. Ranked Merkle Hash Tree (RMHT)

An extended Merkle Hash Tree (MHT) with ranks is named as RMHT. Similar to the binary tree each node N will have a maximum of two child nodes. Non leaf node will constantly have two child nodes. Information present in one node N is an RMHT T represented as {H,rN}where H is a hash value and rN is the rank of the node. Fig .2 shows RMHT.

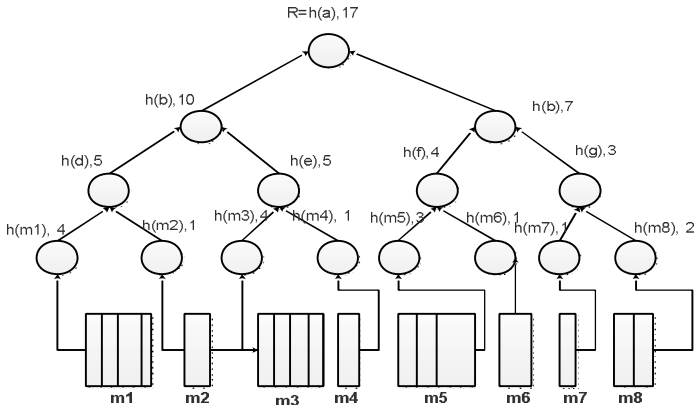

Fig 2: Example of a RMHT

iii. Framework and Definitions

The following are the block level fine-grained update operations. There are six types of operations.

- Partial modifications: A consecutive part of certain blocks needs to be updated.
- Whole block modifications: A whole block is replaced by new set of data.
- Block deletion: A whole block is deleted from the tree structure.
- Block insertion: A whole block needs to be created at the tree structure; contains newly inserted data.
- Block splitting: A part of data is taken out from the block and is inserted next to the new block.

iv. Proposed Scheme

The proposed scheme will describe the roles for each participating parties in the architecture to adopt certain techniques and schemes in the aim of supporting variable-sized data blocks and third party auditing.

Roles of each participating party are described below

Client: Stores data in the cloud and relay on the cloud for data computation consist of both individual consumers and organizations data.Cloud storage server: A cloud service provider, who as resources and expertise in constructing and managing distributed storage servers on cloud, owns and operates the live cloud computing systems.

Third Party Auditor: A TPA has expertise and capabilities that common users may not have. Upon request from the user, a TPA is trusted to asses and exposes risk of Cloud storage services.

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

- **Setup**: The client will generate keying materials via KeyGen and FileProc, and then upload the data to CSS. The client will store a RMHT instead of a MHT as metadata.The client will authorize the TPA by sharing a value sigAUTH. The setup scheme is the first scheme in the proposed system.
- **Verifiable Data updating**: The CSS performs the client's fine-grained update requests via PerformUpdate , then the client runs VerifyUpdate to check whether CSS has performed the updates on both the data blocks and their corresponding authenticator honestly.
- **Challenge, Proof Generation and Verification:** Describes how the integrity of the data stored on CSS is verified by TPA via GenChallenge ,GenProof and Verify.

V. PSEUDOCODE

The pseudo codes for the three participating parties, the client, third party auditor and cloud storage server are given below. All the three modules are performing update operations. Each operations performed by the individual module are given separately in a symmetric manner.

Pseudo code for client which is performing data update in the scheme is given below
Client

1. Send update (block level operation) request.
2. Compute R with {(mi),Ωi}, verify sig with R.
3. If success
4. Compute mi` with {UpdateReq,mi}, then compute Rnew with {H(mi`), Ωi }
5. If Rnew=R'
6. Compute αi' and $sig' = (H(R'))\alpha$

Pseudo code for cloud storage server (CSS) which is performing data update in the scheme is given below
CSS

1. Locate mi
2. Perform the Update.
3. Update mi to mi'
4. Compute R' with $\{mi', \Omega i\}$
5. Update αi to αi' and *sig* to *sig'* accordingly

Pseudo code for Third Party Auditor (TPA) which is performing data update in the scheme is given below
TPA

1. Send challenge request.
2. For $k \in [1, w]$, verify sigAUTH with VID and t.
3. Compute R with $\{H(mi), \Omega i\}$ and verify the following:
   $e(sig,g)=e(H(R),v)$
   $e(\sigma, g) = e(w, v)$
4. If all equations hold.
5. Return TRUE.
6. Else return FALSE.

VI. EXPERIMENTAL RESULTS

Jelastic cloud Deployments

Jelastic acronym for JAVA Elastic is an unlimited platform as a service and container based Infrastructure as a service provider within a single platform that provides high availability of applications to software development clients, enterprisebusiness.Jelastic is a platform as infrastructure provider of java and PHP hosting. Jelastic does not have the

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

restrictions or the code altering requirements. It offers automated vertical scaling and availability from multiple hosting providers.Jelastic cloud offers many benefits. Jelastic runs the codes on standard industry leading software stacks such as Apache Tomcat, glass fish, and wildfly.Jelastic is deployed from multiple data centers worldwide .supported development platforms are java, PHP, Ruby, Python, Node.js, .NET, open shift cartridges etc. Its products are: cloud hosting, Platform as Infrastructure, Private cloud, Public cloud, Hybrid cloud and web hosting.

Steps to install the Jelastic are mentioned below:

Step1: login in to the jelastic environment.
Step2: creating the environment and setting up database and server.
Step3: deploying the war file in to the environmentthe war file being deployed inside the cloud environment.
Step4: Jelastic environment will successfully create.

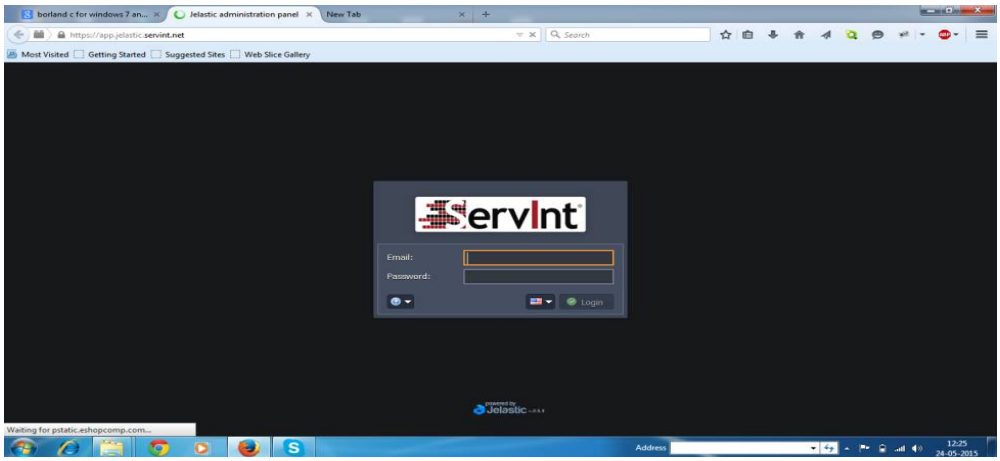

Fig 3: login page of Jelastic

Login in to the jelastic environment as shown in Fig 3.provide valid Email ID and password .After the successful login create the environment and set up database and server. Select tomcat 7 and MySQL 5.7.

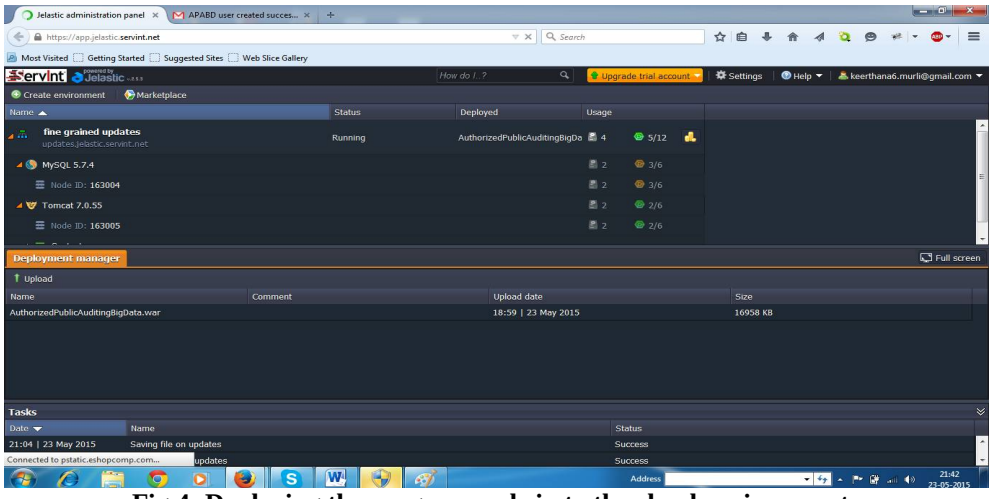

Fig 4: Deploying the program code in to the cloud environment

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

After setting up the database and server we can deploy the war file in to the environment as shown in Fig 4. The war file will be deployed inside the cloud environment.We can perform the file operation on the cloud with file protection .

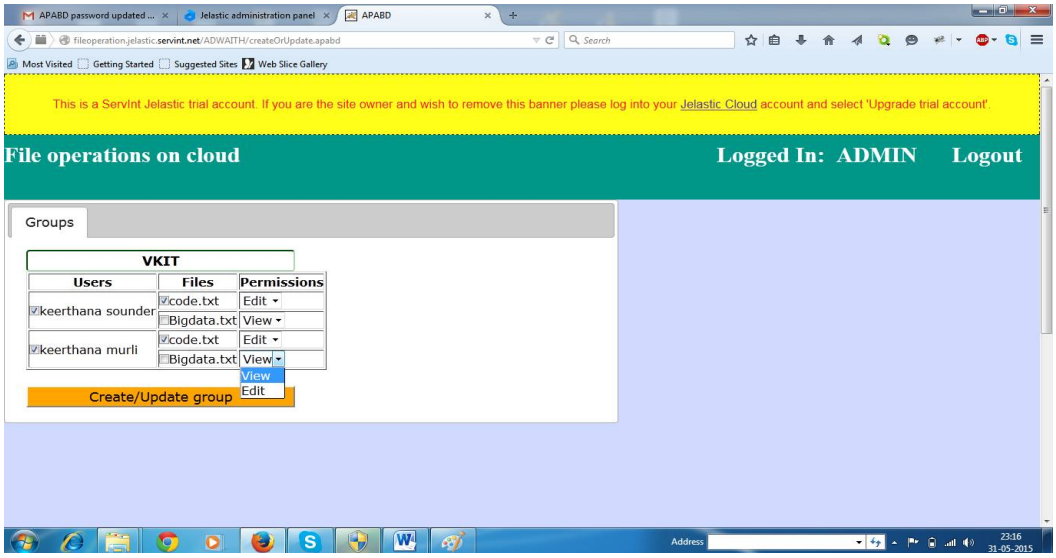

Fig 5: login page for the admin

Login/ sign up page: User should first register. This page takes user information like name, Email ID and phone number.After the successful registration of the user. The user will get unique user ID and password information to his mail then the user can log in obtaining this secret information. User should update the unique password once again to confirm the registration and for user authenticity.

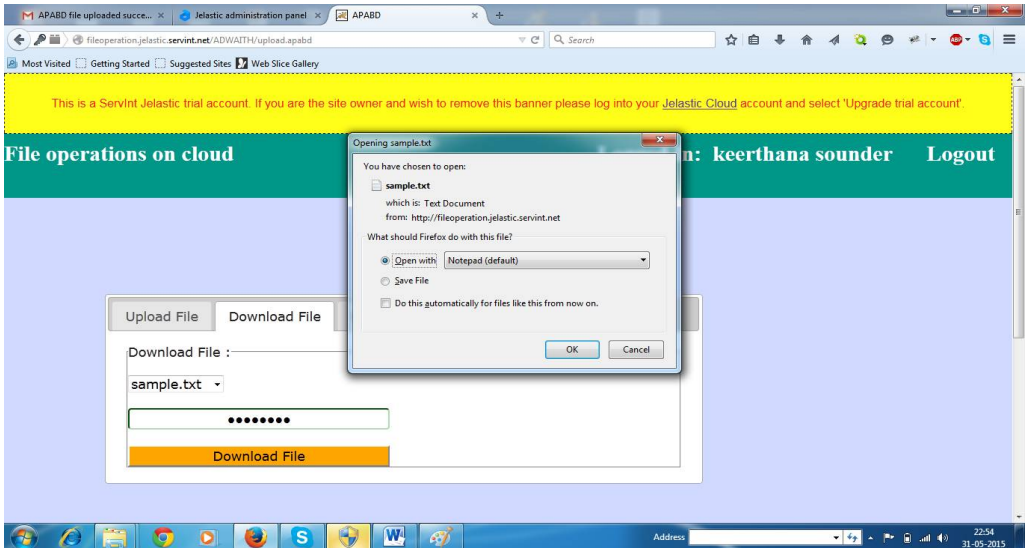

Fig 5: Downloading the file (.txt)

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

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 1, January 2017

Admin can create a group .He has the option of selecting the user and providing permissions to files such as view or edit. User can upload any file (text file). Four operations can be performed a) upload file b) Download file c) update file d) change validity Upload file: user can browse the file to upload, Download: to download user should provide secret key to ensure authenticity, Update: user can modify block of data dynamically. Change validity: user can challenge for validity.

VII. CONCLUSION AND FUTURE WORK

The experimental results show the secure file operations on the cloud and provided authorization to the Third party Auditor. This will enable user to securely store their data in the cloud storage server and user can check the data integrity of the transferred data. In future enhancement improvisation at the server side can be made for efficient data security. Data security is considered as a metric of quality of service. A security aware scheduling scheme can be adopted at the server side protection.

REFERENCES

[1] C. Wang, Q. Wang, K. Ren and W. Lou " Privacy Preserving Public Auditing for Data Storage Security in Cloud Computing", in Proc. 30st IEEE Conf. on comput. andcommun.(INFOCOM), 2010, pp.1-9.
[2] G.Ateniese, R.B. Jhons, R.Curtmola, J.Herring, L.Kissner, Z.Peterson and D. Song "Provable Data Possession at Untrusted Stores", in Proc. 14 ACM Conf. on Comput. andCommun. Security (CCS), 2007, pp. 598-609.
[3] A. Juels and B.S. Kaliski Jr., "PORs: Proof of Retrievability for Large Files", in Proc. 14th ACM Conf. on Comput. and Common. Security (CCS), 2007, pp. 584-597.
[4] H. Shacham and B. Waters," Compact Proofs of Retrievability", in Proc. 14th Int'l Conf. on Theory and Appl. of Crypotl. and Inf. Security (ASIACRYPT), 2008, pp. 90-107.
[5] G.Ateniese, R.D Pietro.L.V.Manchi, and G. Tsudik,ble "Efficient Provable Data possession", in Proc.4th Intl Conf.Security and privacy in communication , 2008, pp. 1-10.

BIOGRAPHY

Keerthana M M is an assistant professor in Department of computer science and engineering from ATME College of engineering Mysuru, Karnataka, India. She received her master degree in computer science and engineering from VKIT, Bengaluru affiliated to VTU University, Belagavi. She received her bachelor degree in computer science and engineering from GECR,Ramanagar , Bengaluru, Karnataka. Her field of interest is on cloud computing.