International Journal of Scientific Research & Engineering Trends
Volume 5, Issue 1, Jan-Feb-2019, ISSN (Online): 2395-566X

Privacy Preserving Ranked Keyword Search Over Cloud Computing

M.Tech. Scholar Ashwini S
Department of Computer Science and Engineering
S.J.C. Institute of Technology
Chikkaballapur, India
ashi.tulips@gmail.com

Prof. Dr. T N Anitha
Department of Computer Science and Engineering
S.J.C. Institute of Technology
Chikkaballapur, India
anithareddytn72@gmail.com

***Abstract* - Cloud computing is one of the emerging technologies in today's world (trend).Cloud computing has envisioned as the next generation architecture of IT enterprise. The flip side to this coin is that cloud storage emerges the security issues of confidentiality, data integrity and data availability. The concept of Third Party Auditor(TPA) is to eliminate the involvement of client through the auditing of whether his data stored in the cloud are indeed intact, which can be important in achieving economies of scale for cloud computing. The task of TPA, on behalf of cloud client is to verify the integrity of the dynamic data stored in cloud. The motive of this paper is to provide data security of cloud in cloud computing using digital signature and elliptic curve cryptography. The Provable data possession scheme is implemented to support the dynamic operation on data. An improvement over the conventional technique is done by allowing the user to search their files in the encrypted database with the help of Ranked keyword search.**

***Keywords* – Cloud computing, Data security, Elliptic curve cryptography, searching, and ranked keyword search.**

I. INTRODUCTION

Several trends are opening up the era of Cloud Computing, which is an Internet-based development and use of computer technology. The ever cheaper and more powerful processors, together with the "software as a service" (SaaS) computing architecture, are transforming data centers into pools of computing service on a huge scale. Meanwhile, the increasing network bandwidth and reliable yet flexible network connections make it even possible that clients can now subscribe high quality services from data and software that reside solely on remote data centers. Although envisioned as a promising service platform for the Internet, this new data storage paradigm in "Cloud" brings about many challenging design issues which have profound influence on the security and performance of the overall system. Cloud computing has tremendous advantages such as

- Ubiquitous access
- High reliability
- Resilience
- Scalability
- Cost-effective.

The fundamental service models in cloud computing are Platform as a Service(PaaS),Infrastructure as a Service(IaaS),Software as a Service(SaaS).Numerous methods have been suggested to solve the problem of data integrity. The third party auditor detects the data on behalf of the client. The elliptic curve cryptosystem security scheme is proposed to clinch the data integrity on the remote server. This strategy uses the notion of provable data possession (PDP) to forge the data operations dynamic. The data can be audited by the client itself without being downloaded every time. Homomorphism encryption scheme is executed on the signature level security. This enables the client to do their operations on the cipher text without decrypting it. Client encrypts the data and provides the public key to the TPA, who carries out the operations on its behalf.

II. RELATED WORK

In paper [1] by G. Ateniese proposed a system based on symmetric cryptography, which is not suitable for public verification i.e. third party verification. The number of verifications will be static and storage server has to access more blocks per query and this may become increasingly expensive.

In[2] by Y. Deswarte proposed system is well adapted for software distribution, presents many drawbacks for other applications: (i)It is not directly applicable to web services: the replies to http requests are generally not a simple file content, and even when it is the case, the integrity checks would have to be integrated in browsers, with all the complexity associated with PKI management.

A hacker could still replace the current copies of the files with obsolete copies with their original signatures. It would not solve the remote server management problem: the administrator would still have to retrieve the contents of all the files to check their integrity.

Erway at el. [3] were the first to come up with the architecture for dynamic provable data possession. They expanded the model discussed in [2] to support the dynamic data operations. Jules and Kaiski et al explained the basic idea behind PoR [4]. The model ensured two aspects of data storage on the cloud, that of possession and irretrievability.

Chen et al. [6] have described a scheme of mutual authentication for cloud based on Elliptic curve cryptography (ECC). Its major focus was on ensuring the identity of the user before accessing the cloud contents. The scheme was efficient in increasing the level of security on the cloud, but contributed in increasing the overhead at the server. To remove this, the concepts of TPA and homomorphism authentication were implemented on the cloud.

Verraju et al. [7] implemented the ECC on the cloud. They considered that communication is taking place only between two public clouds without any user in the scenario. The major focus was on the encryption and decryption of data stored.

Dr. Rao Mikkilineni et al [8] and Robert Lolita et al [11] explains secure storage has already captured the attention of many cloud providers, offering a higher level of protection for their customer's data. We think that more advanced techniques such as searchable encryption and secure outsourced computation will become popular in the near future, opening the doors of the Cloud to customers with higher security requirements. The storage services that do not trust on the storage servers are relevant to our work. The data is secured at the client side before it is send to the servers.

This technique secures the data stored at the service provider's end and provides intimation to the data owner when there are any security breaches. All commercial cloud providers offer to their customers at least one cloud storage service: Amazon' S3 and Simple DB, Microsoft Windows Azure Storage services, Google App Engine Data store, just to cite the most popular. All these products are very powerful for their scalability and storage capacity, but their security mechanisms 4 Isaac Agudo, David Nuñez, Gabriele

III. ELLIPTIC CURVE IN CRYPTOGRAPHY

ECC is a public key cryptographic system based on the certain mathematical problems. An elliptic curve is given by the equation in the form of

$$y^2 = x^3 + ax + b$$

Where,

$4a^3 + 27b2 \neq 0$

Considering an elliptic curve in equation 1.

$$E: y^2 = x^3 + x + 1 \qquad (1)$$

If *p*1 and *p*2 are on E, where

$p1 = (x1, y1)$,
$p2 = (x2, y2)$,
$p3 = (x3, y3)$ and
$p1 \neq p2$

as shown in Fig.1, then from the definition we have

$$p3 = p1 + p2$$

Every user has the set of the public and private key, which are used for encryption and decryption respectively. It is an extension of the Diffie -Hellman key agreement scheme. The ECCSS is a variant of the Digital Signature Algorithm which uses elliptic curve cryptography.

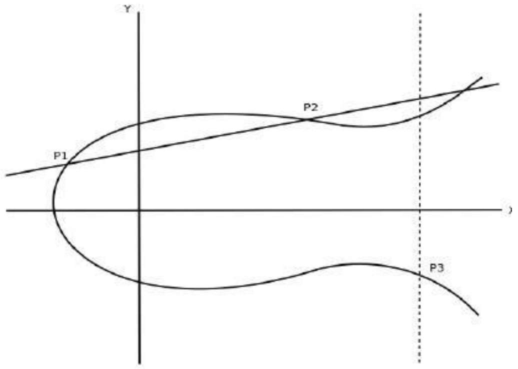

Fig. 1. Elliptic curve

IV. SYSTEM ARCHITECTURE

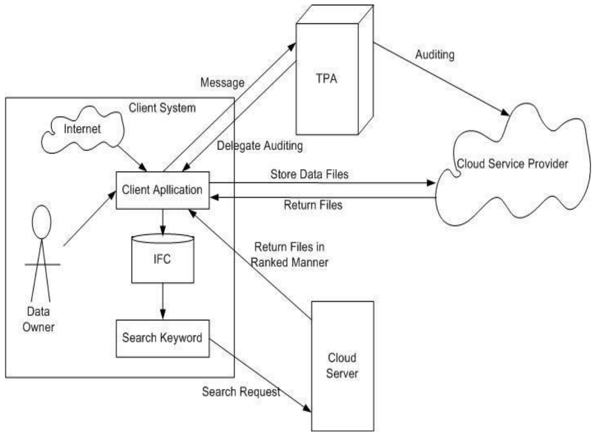

Fig. 2 System Architecture.

V. METHODOLOGY

An ongoing review is done to the data stored at server using ECC based storage security system. In this stratagem PoR mechanism is used. The entire technique comprises of a challenge and response process. The integrity of any block data can be instigated by the client and in return server has to produce a response proof. The proof sent by the server is examined and accuracy of data is certified. ECCSS consists of two phases i.e., setup phase and integrity phase.

VI. SETUP PHASE

In this stage, the file F to be stored is broken into η data blocks by the client= {m1,m2,m3…mn}.Reed-Solomon code is used to encode the data blocks so that it is unaffected by errors slightly. The public key p and private key x are generated using ECC algorithm. The whole process is carried out by the following steps:

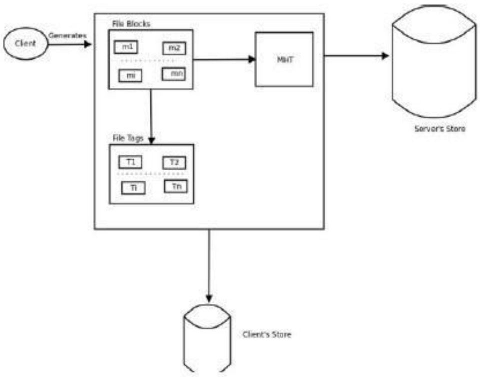

Fig. 3. Processing of file blocks

Firstly, a random number k, such that $1 \leq k \leq (n-1)$ is choosen by the client where n refers to the order of a point Q selected on the elliptic curve. Next the signature tag is created for each data block using the secure key and hash algorithm (SHA-1).

The second step is to generate a collective signature set ϕ={Ti},In third step, for each data block a Merkle Hash Tree(MHT) based on the values from the hash algorithm. In the fourth step, the constructed root of MHT is signed using the secret key. Finally client describes it to the server. After advertising it to the server the client deletes from its local storage. The client conveys the public key to the TPA to examine in the remote files.

VII. ALGORITHMS

The ECCSS comprises of four algorithms. Generation of public key and private key for the signature is done through Key-generation (Key gen) algorithm. Signature for each block is generated using Signature-generation (Sig Gen) algorithm. The Proof-generation (Proof Gen) algorithm generates the proof for the challenge send. The final one Verify-proof (Ver Proof) which is used to verify the proof generated.

1. Key Gen

Input: None

Output: public key p, private key and point Q.

- The client will run the algorithm which will generate the public key p and the private key x.
- A point Q is chosen on the elliptic curve E (K), where K is a finite field.
- It selects a pseudo random number such that $1 \leq x \leq (n - 1)$.
- Point p = Q.
- ECC key pair is (p,x), where p is the public key, x is the private key.

2. SigGen

Input: File block F, secret key , chosen point Q.

Output: Signature set φ.

- A random number k is generated such that $1 \leq k \leq (n-1)$
- The hash value of blocks are calculated ei = H (mi).
- Computes point kQ = (x1, y1).
- Computes r = x1(mod n). If r = 0, go to step 1.
- Next computes s = k−1 (ei + xr) (mod n), if s = 0, go to step 1.
- The signature for a particular block s = (r, s).
- The block tag generated for a block is Ti-(e,kmi) for i ϵn.
- The signatures grouped together as φ.

3. ProofGen

Input: Subset of file blocks mi, co-efficient ai.

Output: Proof P.

- The server generates T, M and AAI for the client to generate MHT.
- The proof P contains {T, M, {H (mi), Ωi} s1≤i≤sc , sigx(H(R)}

4. Verproof

Input: Proof P.

Output: Boolean value TRUE, FALSE.

- The verifier validates the proof by generating the MHT using AAI.
- If step 1 is TRUE, then the further verification is done, else it yields FALSE.
- Initially verification of (r, s) is done over the interval [1, n-1].
- Next it computes e = H (mi).
- Then computes w = s−1 (mod n).
- Next u1 = ew(mod n) and u2 = rw(mod n) are computed.
- Then computes X= u1Q + u2P.
- If X = 0, then S is rejected. Otherwise v = x1(mod n) is computed.
- The proof is accepted only if v = r.

VIII.EXAMPLES USED

1. Client Application

The client application is in charge for all the cryptographic operations done on IFC. Using the RSA key pair the user logs into the application as described in the settings file. This file includes the address of the compute cloud server instance, idea of user log in and storage access key. Numbers of functions are performed depending on the access rights of the user as the user logs in. A number of file system commands can be executed by the client application. Some of them are:

1.1 Upload - to store the IFC on the storage service this command is used.

1.2 Download - Files can be retrieved from the storage service given a filename with the help of this command. Remove - to delete an object from the storage service given a filename.

1.3 List - a list of items within the storage service 3 can be reacquired.

1.4 Search - search capability is taken and a request is sent to server. Then the server replies with a list of files that matches the search.

2. Server Application

The server application acts as a medium between the client application and storage service and executes in the compute cloud. The clients are verified using this application and a secure connection is established between the client and the storage service. While responding to search queries bulk processing is involved otherwise the received requests are simply forwarded to storage service.

IX. CONCLUSION

This paper proposes a secure technique where the data of the user can be stored on a public cloud and the user have access rights on his own hands. Through the ranked manner, the user has the capability to find their data and overhead will be saved in the encrypted database. Efficient security is obtained and signature is created using ECC algorithm. By this integrity and confidentiality of this data secured. In future we can incorporate the system in dynamic real world application to have more effective data storage on the cloud.

Acknowledgment

In the emerging era, Security overhead must be kept low with respect to performance and storage. When the system is busy in executing the cryptographic operations the user need not wait for long time. Large security overhead is not required because the data is stored on the storage provider as it incurs higher costs. The result of this paper shows that it consumes a low overhead on the files. As the file size increases the overhead decreases and becomes constant

REFERENCES

1. G. Ateniese, R. D. Pietro, L. V. Mancini, G. Tsudik “Scalable and Efficient Provable Data Possession”, Fourth International Conf. Security and Privacy in Comm. Networks (Secure Comm 08), 2008.
2. Y. Deswarte, J. Quisquater, and A. Saidane “Remote integrity checking”, Conference on Integrity and Internal Control in Information Systems, November 2003.
3. C. Erway, A. Kupcu, C. Papamanthou, R. Tamassiav, “Dynamic Provable Data Possession”, 16th ACM Conf. Computer and Comm. Security, 2009.
4. A. Juels and B. S. Kaliski Jr. “Pors: Proofs of Irretrievability for Large Files”, 14th ACM Conf. Computer and Comm. Security, pp. 584 - 597, 2007.
5. Qian Wang, Cong Wang, Kui Ren, Wenjing Lou, “Enabling Public Audit ability and Data Dynamics for Storage Security in Cloud Computing”, IEEE Transactions on Parallel and Distributed Systems, May 2011.
6. Tien-Ho Chen, Hsiu-lien Yeh, Wei-Kuan Shih, “An Advanced ECC Dynamic ID-Based Remote Mutual Authentication Scheme. for Cloud Computing”, Fifth FTRA International Conference on Multimedia and Ubiquitous Engineering, 2011
7. Veerraju Gampala, Srilakshmi Inuganti, Satish Muppidi, “Data Security in Cloud Computing with Elliptic Curve Cryptography”, International Journal of Soft Computing and Engineering, July 2012.
8. Dr. Rao Mikkilineni and Vijay Sarathy Kawa Objects, Inc. Los Altos, CA .Cloud Computing and the Lessons from the Past. , 18th IEEE International Workshops 2009.
9. Tharam Dillon and Chen Wu and Elizabeth Chang .Cloud Computing: Issues and Challenges., 24th IEEE International Conference on Advanced Information Networking and Applications,2010
10. Wayne A. Jansen, NIST .Cloud Hooks: Security and Privacy Issues in Cloud Computing., Proceedings of the 44th Hawaii International Conference on System Sciences – 2011.