RESEARCH ARTICLE OPEN ACCESS

A Survey on data security models in cloud computing

Noman Mazher[1] Imran Ashraf[2]
[1]*Department of Information Technology University of Gujrat, Gujrat Pakistan*
[2]Lecturer Faculty of CS & IT, University of Gujrat, Gujrat Pakistan

Abstract:
The world of data computing is becoming larger and complex day by day. Could computing is the most popular model for supporting large and complex data. Organizations are moving toward cloud computing for getting benefit of its cost reduction and elasticity features. However cloud computing has potential riks and vulnerabilities. One of major hurdle in moving to cloud computing is its security and privacy concerns. As in cloud computing environment data is out of user possession this leads to great risk of data integrity, data confidentiality and data vulnerability etc. A number of security models have been developed to cope with these security threats. Our research aims at investigating security models that were developed for securing data during whole lifecycle of cloud computing.
Keywords—*Cloud computing , data security , security model*

I. INTRODUCTION

The world of computation has becoming larger and complex day by day. Could computing is most popular model for supporting large and complex data. Organizations are moving toward cloud computing for getting benefit of its cost reduction and elasticity features. Yet cloud computing have potential disadvantages and threats. One of the major hurdles in moving to cloud computing is its security and privacy concerns. As in cloud computing environment data is out of user possession this increases the risk of data integrity, data confidentiality etc. To reduce these risks researchers have purposed many security models. In our research we will investigate security models that were developed for securing data during whole lifecycle of cloud computing**.**

The data life cycle is divided into seven stages that are Data generation , data transfer , use , share , storage , archival and destruction (Chen e Zhao, 2012) .

In proceeding section we will introduce different security models purposed for different security threats.

II. RELATED WORKS

Cloud computing collects all the computing resources and manages them automatically through software. The historical data and present data are integrated to make the collected information more accurate. In this way cloud computing provides more intelligent service to the users. The users are not bothered about how to buy a server or solution. Instead they can buy the computing resources on the internet according to their need. However cloud computing emerge as promising technology in order to provide the services remotely. But there are many security issues in cloud computing. For example in February, 2010, the Amazon network host service, S3 (Simple Storage Service) was broken down for 4 hours. This made people think about the security of cloud computing again. Since Amazon provides S3, it has attracted a lot of entrepreneur on Web 2.0 put their website on the data center of Amazon to save a large hardware investment (Zhang *et al.*, 2010). So the service of cloud computing is not stable and believable. Security is still a major concern in cloud computing and one of the reason that's why cloud computing is still not admitted by the users.

To improve security of cloud many models have been purposed. User would like to know which security scheme they should implement for securing our data. Comparison of these security schemes have also been done by many researchers such as Farzad Sabahi discuss many security problems to cloud computing against these problems. (Sabahi, 2011) .
In this section we will introduce different security models which were purposed by different researchers. We will also tell about each security model that in which layer it works.

A. Private Virtual Infrastructure (PVI):

This model is proposed by F.John Krothein .(Author Guidelines for 8 - krautheim.pdf, 2013) In this model he addresses risks associated with data. As in cloud computing data is out of control of organization so privacy and security of data is the major concern of the organizations. PVI ensures security of data while transferring data from client to service provider and vice versa. Main purpose of PVI is to secure data during its **transfer stage.** In this

purposed architecture data security is ensured with the help of two layers: PVI layer and cloud fabric layer.

PVI layer:

Datacenter of PVI is controlled by information owner. Each client is responsible of security of data through firewall, intrusion detection system and other monitoring and logging system to ensure that data is confidential.

Cloud fabric layer:

Cloud datacenter or fabrics are controlled by cloud service provider. This layer is responsible of maintaining physical and logical security of data. They ensure the security through security tools.

Another tool Locator Botprovider is also used. Locator Botprovider provides details of all activities by monitoring the cloud security. He claims that locator Botprovider will also ensure security during **destruction stage** of data. This architecture is general and does not restrict itself to any specific XaaS (IaaS, PaaS, and SaaS) service. It is a suggested architecture and there is no implementation detail provided.

B. Privacy-preservation public Auditing:

This model is purposed by Cong-wang et...al. (Wang *et al.*, 2010)This architecture is suggested typically for securing data of cloud during during **storage stage.** In prototype of this model he takes three entities 1) cloud user 2) cloud server 3) cloud service provider.

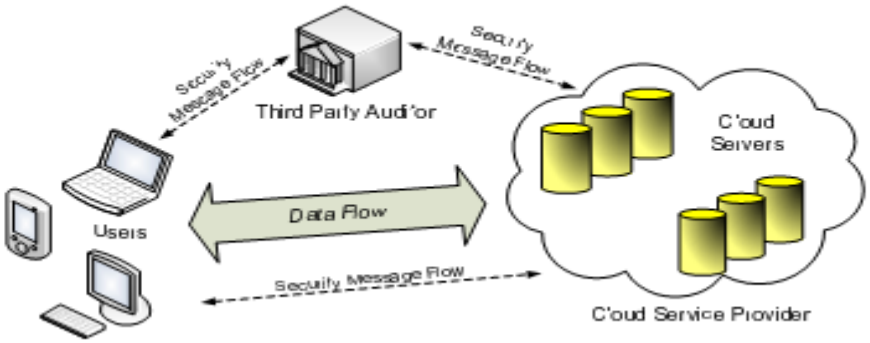

Fig 1: cloud data storage service (Wang *et al.*, 2010)

Third Party Auditor (TPA) is pivotal point in this architecture as he has expertise and experience of auditing data. TPA can send regular audit reports of data. From those report user can evaluate that either his data is going securely on cloud or not. To show the performance of the proposed model a mathematical notational proof is provided as well. However no implementation detail in any domain is provided. This architecture can be used for all XaaS (IaaS, PaaS, and SaaS) services.

C. Cloud data storage security scheme:

This scheme is developed by Syed Azahad and Mr.Srinivas Rao(Azhad e Rao). In this scheme their main focus is on **data storage** security. This scheme is developed with **SaaS** in mind. This scheme addresses the correctness of data in the cloud environment. In this scheme they use homomorphic token with distributed verification. They claim to achieve integration of correctness of storage data and error localization of data. They also give security analysis and the results of their analysis with the help of graphs.

D. Hypbrex (Hybrid Execution) model:

Steven.Y.Ko et.al purposed that model (Ko, Jeon e Morales, 2011). The HybrEx model provides a seamless way for an organization to utilize their own infrastructure for sensitive, private data and computation, while integrating public clouds for no sensitive, public data and computation. In this paper the purposed model Hybprex is also implemented in a specific environment called Mapreduce. Mapreduce is a programming model used to execute large dataset with parallel algorithm.

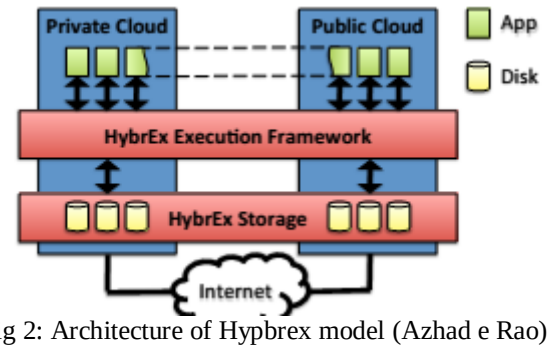

Fig 2: Architecture of Hypbrex model (Azhad e Rao)

Hypbrex model typically worsk on **storage stage** of data. For ensuring integrity of data on public cloud he purposed to maintain hashes of public data in private cloud. With the help of these hashing algorithms hybprex can validate integrity of both public and private data

E. . Airavat :

Indraji roy et al purposed that model for security and privacy for Mapreduce .(Untitled - roy.pdf, 2013) Mapreduce is a programming model used to execute large dataset with parallel algorithm. Airavat guarantees to bound information leakage from any individual data that is computed over mapreduce..

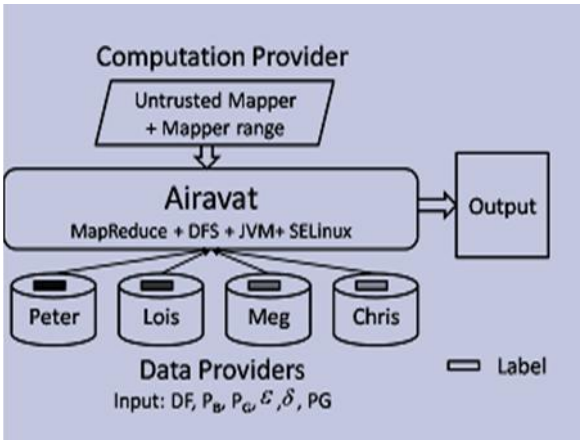

Fig 3 : Architecture of mapreduce (Roy *et al.*, 2010)

In this paper they implemented Airavat on mapreduce environment and gave the result with the help of graphs .Airavat provides security on **data storage** stage.

F. Integrated conceptual digital forensic framework

This paper is written by Ben Martenie (An integrated conceptual digital forensic framework for cloud computing - r_6_130217100243.pdf, 2013) . In this paper he integrates two framework of digital forensics. These two frameworks are McKemmish (1999) and NIST (Kent et al., 2006) . This framework is analyzed with help of survey and questioner. These frameworks are work for data security in **data transfer** stage.

G. Insider Threat remedies:

William R.ClayComb discussed a series of security threats which are associated to cloud computing in general and data security in specific. A general data security threat is insider threat.(Claycomb e Nicoll, 2012) Insider threat is commonly known as "Rouge Administrator" threat. According to definition of CERT insider threat is *"current or former employee, contractor, or other business partner who has or had authorized access to an organizations network, system or data and intentionally exceeded or misused that access in a manner that negatively affected the confidentiality, integrity, or availability of the organizations information or information systems."*

They gave the following suggestions to prevent our data from insider threat:

- Enforce strict supply chain management and conduct a comprehensive supplier assessment.
- Specify human resource requirements as part of legal contracts.
- Require transparency into overall information security and management practices, as well as compliance reporting.
- Determine security breach notification processes.

Insider threat or "rogue administrator" threat can be categorized on **data storage** stage of cloud computing.

H. POR and PDP:

Jeff WhiteWorth discussed two issue of SaaS storage model (Proving-Retrievability-and-Possession-within-the-Cloud-Storage-Model.pdf, 2013).

1) POR (Prove of Retrievability)

First issue is Prove of Retrievability. This means that user should have any proof that data stored on Cloud storage device can be retrieved without any error or loss, when needed. Solution provided for POR is to use hash function on data. So data is sent in encrypted state. When original data owner or *prover* wants to retrieve data from cloud service provider or *verifier* he can ensure that data which are retrieved are same as sent for storing .

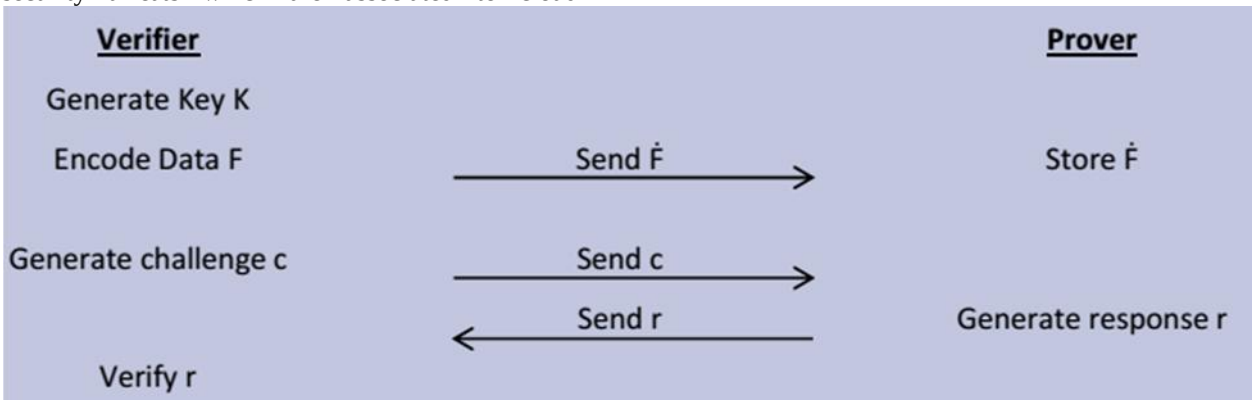

Fig 4: Overview of POR process (Proving-Retrievability-and-Possession-within-the-Cloud-Storage-Model.pdf, 2013)

2) PDP (Proving data possession)

Provable data possession (PDP) refers to the ability of a client to verify that data stored with a server still possesses the data without retrieving it. The building block of the proposed PDP scheme utilizes

Homomorphic Verifiable Tags (HVTs) that are, in the most basic concept, metadata stored with the file that provides unforgeable data tags used for verification. These building blocks are utilized in the four main polynomial-time algorithms used in the PDP scheme: KeyGen, TagBlock, GenProof and CheckProof. The first two algorithms are used in the scheme setup to produce public and secret keys as well as HVTs. GenProof is used by the server to generate a proof in response to the challenge sent by the client and CheckProof is used by the client to verify the proof generated by the server.

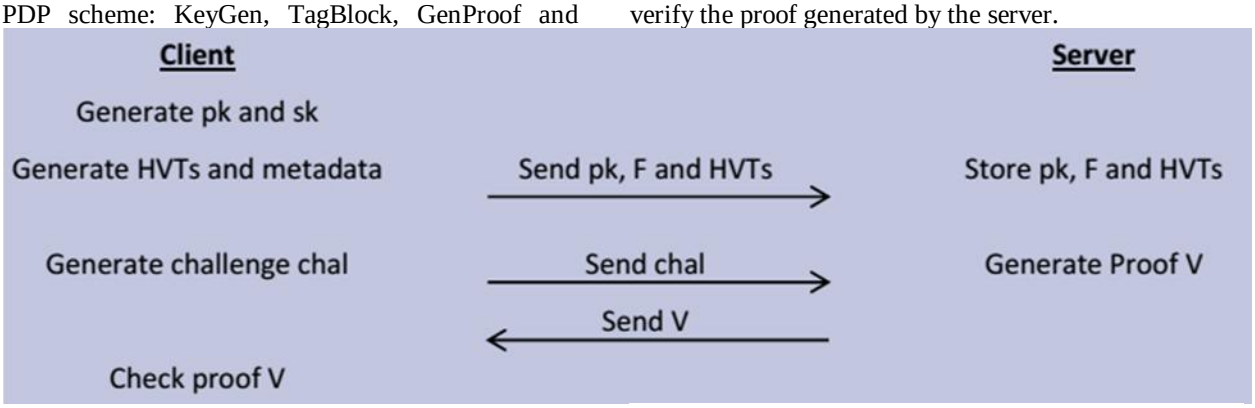

Fig 5: PDP overview (Proving-Retrievability-and-Possession-within-the-Cloud-Storage-Model.pdf, 2013)

Both of these approaches address data security at **storage stage**.

I. Transparent Table encryption Technique:

Ms. Ankita Parkash et al purposed an oracle storage based technique for ensuring data security in cloud computing (Baheti e Patil, 2013). This technique ensures the reliability of data that is stored in cloud environment.

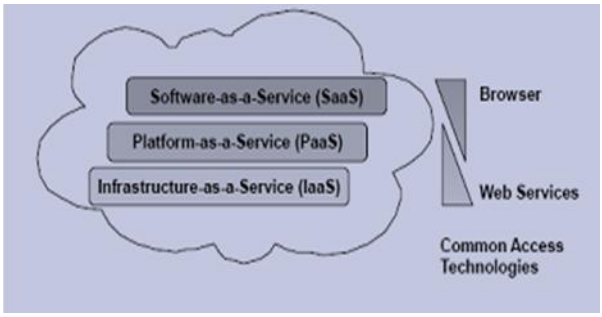

Fig 6: Working diagram of proposed system (Baheti e Patil, 2013)

In this approach they purposed a novel approach of storing data. As user data is stored on any particular cloud so we divide that particular cloud into four groups or region R0, R1, R3 and R3. When user stores his file on cloud his file is partitioned in four parts. A segment file is maintained for each file.

TABLE I. SEGMENT TABLE

Seg. No.	Base	Bounds	rw
0			
1			
2			
3			

Fig 7: segment table (Baheti e Patil, 2013)

Here Base represents the starting address of segment in memory. The bound represents end address of the segment. Rw field represents read and write access of that segment. This all work will be done on cloud administrator side. Cloud administrator encrypts entries of segment table. When user wants to retrieve their data cloud administrator will decrypt that segment table entries and with the help of base and bound address gets user data and sends it to the user. So in this way user data can be secured in cloud environment. This technique also addresses security threats at **data storage stage**.

III. CONCLUSION

In this research we analyzed different security models that were purposed for different stages of data in cloud computing. It presents a comprehensive report on different security models which has been purposed so far.

For securing data on different stages of data lifecycle different security models can be implemented within different perspectives. Confidentiality is an immense concern of data security. To ensure users confidentiality Hypbrex model can be implemented as it is developed by taking confidentiality in mind specifically. While in the environment where information leakage is more important Airavat is the best suit. If data will have to store for long time their Proof of reliability and proof of data possession is very important. So model of OR

and PDP scheme presented is best suited in scenario where data will have to store for a long time. Reliability of stored data is very important for user. Transport Table Encryption Scheme is best suited for ensuring reliability of data. Cloud computing security scheme can be used where integration and correctness of data is important. Transfer of data during client to server can be secured using PVI model.

References:

[1] http://www.ccbc.ir/files_site/files/r_6_130217100243.pdf >.
[2] Author Guidelines for 8 - krautheim.pdf. 2013. Disponível em: < http://static.usenix.org/event/hotcloud09/tech/full_papers/krautheim.pdf >.
[3] AZHAD, S.; RAO, M. S. Ensuring Data Storage Security in Cloud Computing.
[4] BAHETI, A. P.; PATIL, S. V. Cloud Security Based On Oracle Storage Using Transparent Table Encryption Technique. **IJCER,** v. 2, n. 2, p. 132-136, 2013. ISSN 2278-5795.
[5] CHEN, D.; ZHAO, H. Data security and privacy protection issues in cloud computing. Computer Science and Electronics Engineering (ICCSEE), 2012 International Conference on, 2012, IEEE. p.647-651.
[6] CLAYCOMB, W. R.; NICOLL, A. Insider Threats to Cloud Computing: Directions for New Research Challenges. Computer Software and Applications Conference (COMPSAC), 2012 IEEE 36th Annual, 2012, IEEE. p.387-394.
[7] KO, S. Y.; JEON, K.; MORALES, R. The hybrex model for confidentiality and privacy in cloud computing. Proceedings of the 2011 conference on Hot topics in Cloud Computing. USENIX Association, Portland, OR, 2011.
[8] Proving-Retrievability-and-Possession-within-the-Cloud-Storage-Model.pdf. 2013. Disponível em: < http://www.whitworthtech.com/sysadmin/wp-content/uploads/2012/07/Proving-Retrievability-and-Possession-within-the-Cloud-Storage-Model.pdf >.
[9] ROY, I. et al. Airavat: Security and Privacy for MapReduce. NSDI, 2010. p.297-312.
[10] SABAHI, F. Cloud computing security threats and responses. Communication Software and Networks (ICCSN), 2011 IEEE 3rd International Conference on, 2011, IEEE. p.245-249.
[11] Untitled - roy.pdf. 2013. Disponível em: < https://www.usenix.org/legacy/event/nsdi10/tech/full_papers/roy.pdf >.
[12] WANG, C. et al. Privacy-preserving public auditing for data storage security in cloud computing. INFOCOM, 2010 Proceedings IEEE, 2010, IEEE. p.1-9.
[13] ZHANG, S. et al. Cloud computing research and development trend. Future Networks, 2010. ICFN'10. Second International Conference on, 2010, IEEE. p.93-97.