# Enabling Data Storage Security in Cloud Computing for Banking Enterprise

S.BIRUNTHA
II ME CSE
Anna University Coimbatore,
Academic Campus
banupriya12317@gmail.com

V.VENKATESA KUMAR,
Lect/CSE,
Anna University Coimbatore,
Academic Campus
mail2venkatesa@yahoo.com

S.PALANISWAMI,
Registrar
Anna university Coimbatore,
Academic Campus
joegct@yahoo.com

**ABSTRACT**

Cloud computing delivers convenient, on-demand access to shared pools of data, applications and hardware over the internet. Cloud computing provides unlimited infrastructure to store and execute customer data and program. As customers we do not need to own the infrastructure, they are merely accessing or renting; they can forego capital expenditure and consume resources as a service, paying instead for what they use. Data can be redundantly store in multiple physical locations. Due to this redundancy the data can be easily modified by unauthorized users which can be stored in the database. This leads to loss of data privacy and security to database. Extensive security and performance analysis shows that the proposed scheme ensures that cyclic redundancy check and time-tested practices and technologies for managing trust relationships in traditional enterprise IT environments can be extended to work effectively in both private and public clouds. Those practices include data encryption, strong authentication and fraud detection, etc.

Keywords: Cloud Computing, On-Demand Access, Cyclic Redundancy Check.

## 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. The increasing network bandwidth and reliable yet flexible network connections make it even possible that users now subscribe high quality services from data and software that reside solely on remote data centers.

Moving data into the cloud offers great convenience to user since they don't have to care about the complexities of direct hardware management. The pioneer of Cloud Computing vendors, Amazon Simple Storage Service (S3) and Amazon Elastic Compute Cloud (EC2) [1] are both well known examples. While these internet–based online services do provide huge amounts of storage space and customizable computing resources, this computing platform shift, however, is eliminating the responsibility of local machines for data maintenance at the same time. As a result, users are at the mercy of their data. Recent downtime of Amazon'sS3 is such an example [2].

From the perspective of data security, which has always been an important aspect of quality of service, Cloud Computing inevitably poses new challenging security threats for number of reasons. Firstly, traditional cryptographic primitives for the purpose of data security protection cannot be directly adopted due to the users' loss control of data under Cloud

Computing.Therefore, verification of correct data storage in the cloud must be conducted without explicit knowledge of the whole data. Considering various kinds of data for each user stored in the cloud and the demand of long term continuous assurance of their data safety, the problem of verifying correctness of data storage in the cloud becomes even more challenging. Secondly, Cloud Computing is not just a third party data warehouse. The data stored in the cloud may be frequently updated by the users, including insertion, deletion, modification, appending, recording, etc. To ensure storage correctness under dynamic data update is hence of paramount importance. However, this dynamic feature also makes traditional integrity insurance techniques futile and entails new solutions. Last but not the least, the deployment of Cloud Computing is powered by data centers running in a simultaneous, cooperated and distributed manner. Individual user's data is redundantly stored and distributed manner. Individual user's data is redundantly stored in multiple physical locations to further reduce the data integrity threats. Therefore, distributed protocols for storage correctness assurance will be of most importance in achieving a robust and secure cloud data storage system in the real world. However, such important area remains to be fully explored in the literature.

Recently the importance of ensuring the remote data integrity has been highlighted by the following research works [3]-[7]. These techniques, while can be useful to ensure the storage correctness without having users possessing data, cannot address all the security threats in cloud data storage, since they are all focusing on single server scenario and most of them do not consider dynamic data operations. As a complementary approach, researchers have also proposed distributed protocols [8]-[10] for ensuring storage correctness across multiple servers to peers. Again, none of these distributed schemes is aware of dynamic data operations. As a result, their applicability in cloud data storage can be drastically limited.

In this paper, we propose an effective and flexible distributed scheme with explicit dynamic data support to ensure the correctness of users' data in the cloud. We rely on cyclic redundancy check in the file distribution preparation to provide redundancies and guarantee the data dependability. This construction drastically reduces the communication and storage overhead as compared to the traditional replication-based file distribution techniques. By utilizing the homomorphism token with distributed verification of redundant data, our scheme achieves the storage correctness assurance as well as data error localization. Whenever data corruption has been detected during the storage correctness verification, our scheme can almost guarantee the simultaneous localization of data errors, i.e., the identification of the misbehaving server(s).

Our contribution can be summarized as the following three aspects:

1) Compared to many of its predecessors, which only provide binary results about the storage state across the distributed servers, the challenge-responses protocol in our work further provides the localization of data error.

2) Unlike most prior works for ensuring remote data integrity, the new scheme supports secure and efficient dynamic operations on data blocks, including: update, delete and append.

3) Extensive security and performance analysis shows that the proposed scheme is highly efficient and resilient against Byzantine failure, malicious data

modification attack, and even server colluding attacks.

## II. PROBLEM STATEMENT

### 2.1System Model

Representative network architecture for cloud data storage is illustrated in fig.1.
Three different network entities can be Identified as follows:

- User: users, who have data to stored in the cloud and rely on the cloud for data computation, consist of both individual consumers and organizations.
- Cloud Service Provider (CSP): a CSP, who has significant resources and expertise in building and managing distributed cloud storage servers, owns and operates live Cloud Computing systems.
- Third Party Auditor (TPA): an optimal TPA, who has expertise and capabilities that users may not have, is trusted to assess and expose risk of cloud storage services on behalf of the users upon request.

In cloud data storage, a user stores his data through a CSP into a set of cloud servers, which are running in a simultaneous, cooperated and distributed manner. Data redundancy can be employed with technique of cyclic redundancy check to further tolerate faults or server crash as user's data grows in size and importance. Thereafter, for application purposes, the user interacts with the cloud servers via CSP to access or retrieve his data. In some cases, the user may need to perform block level operations on his data. The most general forms of these operations we are considering are block update, delete, insert and append.

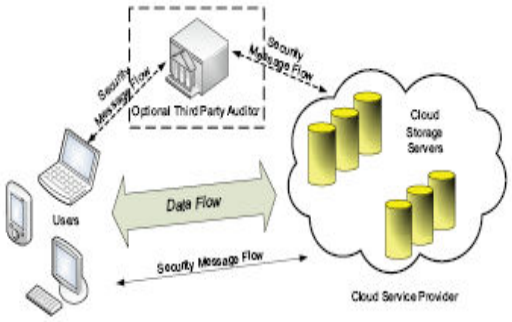


Fig. 1: Cloud data storage architecture

As users no longer possess their data locally, it is of critical importance to assure users that their data are being correctly stored and maintained. That is, users should be equipped with security means so that they can make continuous correctness assurance of their stored data even without the existence of local copies. In case that user do not necessarily have the time, feasibility or resources to monitor their data, they can delegate the tasks to an optional trusted TPA of their respective choices. In our model, we assume that the point-to-point communication channels between each cloud server and the user is authenticated and reliable, which can be achieved in practice with little overhead. Note that we don't address the issue of data privacy in this paper, as in Cloud Computing, data privacy is orthogonal to the problem we study here.

### 2.2 Adversary Model

Security threats faced by cloud data storage can come from two different sources. On the one hand, a CSP can be self-interested, untrusted and possibly malicious. Not only does it desire to move data that has not been or is rarely accessed to a lower tier of storage than agreed for monetary reasons, but it may also attempt to hide a data loss incident due to management errors, Byzantine failures and so on. On the other hand, there may also exist an economically motivated adversary, who has the capability to compromise a number of cloud data storage servers in different time intervals and subsequently is able to modify or delete users' data while remaining undetected by CSP for a certain period. Specifically, we

consider two types of adversary with different levels of capability in this paper:

Weak Adversary: The adversary is interested in corrupting the user's data files stored on individual servers. Once a server is comprised, an adversary can pollute the original data files by modifying or introducing its own fraudulent data files by modifying or introducing its own fraudulent data to prevent the original data from being retrieved by the user.

Strong Adversary: This is the worst case scenario, in which we assume that the adversary can compromise all the storage servers so that it can intentionally modify the data files as long as they are internally consistent. In fact, this is equivalent to the case where all servers are colluding together to hide a data loss or corruption incident.

## III. Best practices for managing trust in private clouds

### 3.1 Set clear policies to define trust and be equipped to enforce them

- In a private cloud, trust relationships are defined and controlled by the organization using the cloud.
- For trust relationships to work, there must be clear, agreed-upon policies for what information is privileged, how that data is managed from fig 2 and how cloud providers will report and validate their performance in enforcing the standards set by the organization.

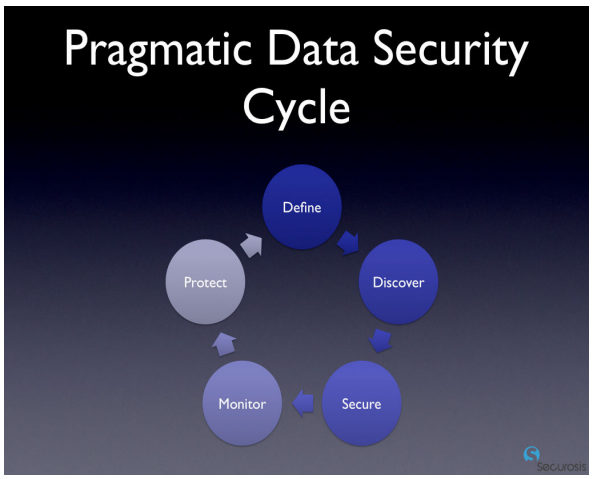


**Fig. 2: Data Security Cycle**

### 3.2 Evaluate whether cloud vendors can deliver on their security claims.

- Because information security is only as strong as its weakest link, it's essential for organizations to evaluate the quality of their cloud vendors. Organizations must aggressively verify whether cloud vendors can deliver upon and validate their security claims

### 3.3 Require transparency into cloud operations to ensure multi-tenancy and data isolation

- In the virtualized environment of the cloud, many different companies, or "tenants," may share the same physical computing, storage and network infrastructure. Cloud providers need to ensure isolation of access - that software, data and services can be safely partitioned within the cloud and that tenants sharing physical facilities cannot tap into their

neighbors' proprietary information and applications.
- Cloud vendors should furnish log files and reports of user activities, and specific performance metrics should be written into managed service agreements and enforced.
- Organizations with private clouds should work with cloud vendors to ensure transferability of security controls.

### 3.4 Preserve segregation of administrator duties

- Cloud administrators need sufficient access to an enterprise's virtual facilities to optimize cloud performance while being prevented from tapping into the proprietary information they're hosting on behalf of their tenants.
- Enterprises should preserve a separation of administrator duties in the cloud. As with physical IT environments, segregating functions within the cloud can provide added security by diffusing control.
- By segregating administrator duties and employing a centralized virtualization management console, organizations can safeguard their private clouds from unauthorized administrator access.

### 3.5 Manage policies for provisioning virtual machines

- To secure their virtual infrastructure, companies using private clouds must be able to oversee how virtual machines are provisioned and managed within their clouds. In particular, managing virtual machine identities is crucial, as they're used for basic administrative functions such as identifying the systems and people with which virtual machines are physically associated and moving software to new host servers.
- Organizations establishing a security posture based on virtual machine identities should know how those identities are created, validated and verified and what precautions their cloud vendors have taken to safeguard those identities.

### 3.6 Employ data encryption and tokenization

- Organizations should encrypt data residing with or accessible to cloud providers to protect proprietary information against unauthorized access, particularly by administrators and other parties within the cloud.
- An additional precaution to secure data residing in clouds is to segregate sensitive data from the users or identities they're associated with.
- Companies also can protect sensitive cardholder information in the cloud through a form of data masking called tokenization.

### 3.7 Adapt federated identity policies backed by strong authentication practices

- In the simplest terms, a federated identity allows a user to access various web sites, enterprise applications and cloud services using a single sign-on.

- Federated identity policies go hand-in-hand with strong authentication policies.
- The federation of identity and authentication policies will eventually become standard practice in the cloud, not just because user convenience is demanded, but also because companies can centralize the access and authentication systems maintained by separate business units.
- Federated identity models, like the strong authentication services that enforce them, are only as strong as their weakest link. Each member of the federation must be trusted to comply with the group's security policies.

## IV. Fraud Protection: Keeping the Bad Guy Out

As rapidly as the cloud is developing, a cyber crime-driven "dark cloud" is growing even more quickly in parallel. The emergence of private clouds represents an opportunity to kick open the doors to the enterprise.

One of the most essential forms of fraud prevention is identity protection: ensuring users actually are who they claim to be. Fraud prevention and identity protection are among the most challenging and fast-changing disciplines within information security.

As cloud security continues to evolve risk-based authentication, which balances security, usability and cost by applying appropriate safeguards based on the risk associated with each activity, will undoubtedly play a major role in preventing fraud within both private and public clouds.

In the coming years, organizations will need to extend their private cloud capabilities in strong authentication and fraud detection to protect against publishing, malware and even intellectual property theft. In building stronger defenses against unauthorized access and online fraud, organizations can borrow from the following fraud prevention practices pioneered by the financial services industry.

### 4.1 Implement strong authentication services

Authentication is often the first line of defense in identity protection. Protecting users poses a challenge, as static passwords are considered too weak. Therefore, many cloud providers are actively seeking to implement a "better than password" authentication technology.

One of the most promising ways to secure online identities in the public cloud is risk-based or adaptive, authentication systems, which intelligently vary authentication processes based on real-time calculations of risk. Risk-based identity protection employs behavior profiling and "invisible," or transparent, authentication processes in which users' requests for cloud services are compared with records of what those users have done in the past. Suspicious activities or patterns that deviate from the norm are automatically challenged.

Risk-based authentication methods are now being broadly deployed in many public clouds, particularly those run by financial institutions.

For advanced risk-based authentication processes, organizations further strengthen user authentication procedures by issuing security tokens to employees. Security token store inalterable, unique identities in protected memory on a small device. The tokens serve as secondary methods of verifying identities after users enter other credentials, such as passwords or PINs.

For enterprises to achieve consistent, guaranteed levels of identity protection, they need to push cloud services providers to deploy identity access and authentication tools that are equal in strength to those used in their enterprise.

### 4.2 Deploy multiple lines of defense to protect against sophisticated malware attacks

The threat posed by increasingly sophisticated malware attacks is a prime example of why layered approaches to identity protection are critical. Enterprises can also participate in external threat protection and intelligence services, such as subscribing to a fraud monitoring network, to minimize the impact of malware attacks.

### 4.3 Map the "Dark Cloud" of cyber crime

In the "Dark Cloud", cyber criminals use an organization's resources to promote their business. The Dark Cloud is the infrastructure fraudsters have built using resources they hijacked from individuals and organizations.

The financial sector and other verticals plagued by publishing and malware often employ cyber crime intelligence services to give them visibility into the Dark Cloud's infrastructure. Such services provide them with the ability to detect mal ware attacking their users, recover credentials stolen from users, shut down infection points and spoofed publishing websites and monitor botnets and command & control mother ships.

## V. Managing Data Compliance in the Cloud

Within cloud environments, the virtualization layer provides an unprecedented degree of visibility into the activity on a system. Hyper visors are exposed to every component and function in the virtual system. This extraordinary degree of visibility means just about every activity involved in providing application services can be monitored and reported for auditing and compliance with relatively little or no extra software instrumentation.

Here are some best practices for dealing with each of these issues and for handling compliance in private clouds:

### 5.1 Monitor cloud vendors for compliance

Organizations deploying private clouds should coordinate with their various cloud providers to ensure the data needed to prove regulatory compliance is fed back into the organization.

### 5.2 Ensure adherence to jurisdictional-specific regulations in borderless clouds

In the cloud, where computing and storage resources are virtualized and can be hosted in several distant locations at once, it's easier for regulated information to "leak" someplace it doesn't belong. Regulatory compliance makes it necessary, in some cases, to manufacture artificial boundaries within borderless clouds.

Organizations needing to ensure compliance with the ever-changing global patchwork of government mandates will benefit from deploying intelligent cloud storage platforms capable of smart provisioning and data loss protection. To illustrate how such platforms work, consider the extreme example of the German Federal Data Protection Act, which essentially prohibits the storage or transfer of German citizen's

personal data outside the jurisdiction of Germany.

In the case of the German Federal Data Protection Act, these "data aware" storage clouds are able to automatically segregate the applicable personal data and store it in German data centers in compliance with the legislation.

## VI. Conclusions

In this paper companies entering the cloud should take steps to ensure they can trust the companies providing them with services, as well as the entities they are transacting with inside the cloud. Enterprises must have the ability to safeguard proprietary information on virtual servers and storage while giving cloud administrators the access and privileges needed to do their jobs [10]. All of cloud issues relate to establishing trust relationships, which form the conceptual foundations for cloud security.

Many of the time-tested practices and technologies for managing trust relationships in traditional enterprise IT environments can be extended to work effectively in both private and public clouds. Those practices include data encryption, strong authentication and fraud detection, etc.

## REFERENCES

[1] Amazon.com, "Amazon Web Services (AWS)," Online at http://aws. amazon.com, 2008.
[2] N. Gohring, "Amazon's S3 down for several hours," Online At http://www.pcworld.com/businesscenter/article/142549/amazons s3 down for several hours.html, 2008.
[3] A. Juels and J. Burton S. Kaliski, "PORs: Proofs of Retrievability for Large Files," *Proc. of CCS '07*, pp. 584–597, 2007.
[4] H. Shacham and B. Waters, "Compact Proofs of Retrievability," *Proc. of Asiacrypt '08*, Dec. 2008.
[5] K. D. Bowers, A. Juels, and A. Oprea, "Proofs of Retrievability: Theory and Implementation," Cryptology ePrint Archive, Report 2008/175, 2008, http://eprint.iacr.org/.
[6] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable Data Possession at Untrusted Stores," *Proc. of CCS '07*, pp. 598–609, 2007.
[7] G. Ateniese, R. D. Pietro, L. V. Mancini, and G. Tsudik, "Scalable and Efficient Provable Data Possession," *Proc. of SecureComm '08*, pp. 1–10, 2008.
[8] T. S. J. Schwarz and E. L. Miller, "Store, Forget, and Check: Using Algebraic Signatures to Check Remotely Administered Storage," *Proc. of ICDCS '06*, pp. 12–12, 2006.
[9] M. Lillibridge, S. Elnikety, A. Birrell, M. Burrows, and M. Isard, "A Cooperative Internet Backup Scheme," *Proc. of the 2003 USENIX Annual Technical Conference (General Track)*, pp. 29–41, 2003.
[10] K. D. Bowers, A. Juels, and A. Oprea, "HAIL: A High-Availability and Integrity Layer for Cloud Storage," Cryptology ePrint Archive, Report 2008/489, 2008, http://eprint.iacr.org/.
[11] L. Carter and M. Wegman, "Universal Hash Functions," *Journal of Computer and System Sciences*, vol. 18, no. 2, pp. 143–154, 1979.
[12] J. Hendricks, G. Ganger, and M. Reiter, "Verifying Distributed Erasurecoded Data," *Proc. 26th ACM Symposium on Principles of Distributed Computing*, pp. 139–146, 2007.
[13] J. S. Plank and Y. Ding, "Note: Correction to the 1997 Tutorial on Reed-Solomon Coding," University of Tennessee, Tech. Rep. CS-03-504, 2003.
[14] Q. Wang, K. Ren, W. Lou, and Y. Zhang, "Dependable and Secure Sensor Data Storage with Dynamic Integrity Assurance," *Proc. of IEEE INFOCOM*, 2009.