Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 4, Issue. 10, October 2015, pg.388 – 395

SURVEY ARTICLE

A SURVEY ON CLOUD COMPUTING SYSTEM USING PROVABLE MULTICOPY WITH DYNAMICALLY FRAGMENTED DATA

[1]R.THENMOZHI, [2]I.JANANI
PG STUDENT, ASSISTANT PROFESSOR
[1,2]NPR COLLEGE OF ENGINEERING AND TECHNOLOGY, TAMILNADU, INDIA
[1] then.2044@gmail.com; [2] janani.npr@gmail.com

Abstract— Distributed data to a remote Cloud Service Provider is a flourishing trend for infinite customers and institution relieve the difficulty of local data storage and maintenance. Moreover the customer gets confidence on the data mirroring provided by the CSP. The structure used for data mirroring can be different based on the data; more facsimile(copies) are demanded for derogatory data which cannot quickly be reproduced This critical data should mirrored on multiple servers across multiple data centers. Before mirroring CSP split the data for improve the efficiency. The system aims the desire riskless and efficient protocol that conquers the limitations of the previous model. It will address the problem of creating multiple facsimile of data over the cloud server, and also review all these facsimile to verify the truthfulness.The Message Authentication Code based algorithm should be used for identifying the integrity and also provide security based on the Elliptic Curve Cryptography encryption and decryption schemes.

Keywords— Cloud Service Provider (CSP); Data mirroring; Elliptic Curve Cryptography Algorithm (ECC); Message Authentication Code (MAC)

1. INTRODUCTION

Cloud computing technology helps to access the data from anywhere in the world. It is a rental based service depends on the user demand. User can able to access the information in an efficient manner. Cloud Computing enables the application software and databases to the centralized large data centers. Management of the data and its services may not be trustworthy. This will create the problem of assuring the integrity of data storage in Cloud.

The considerable attention of Cloud Computing paradigm is due to a number of key advantages which make it a challenging research area in both academia and industry. Cost effective and avoid capital expenditure on hardware, software, and services. Moreover, It can

© 2015, IJCSMC All Rights Reserved

provides low management overhead and immediate access to abroad range of applications. Cloud environment can be used to reduce the maintenance cost. It is flexible to scale up and down IT capacity over time to business needs. CC offers more mobility where customers can access information wherever they are, rather than having to remain at their desks. CC allows organizations to store more data on remote servers. Organizations will no longer be worried about constant server updates and other computing issues; they will be free to concentrate on innovations. Storing data remotely allows users to access the data from various geographic locations making it more convenient to them. Also, some organizations may create large data files that must be achieved for many years but are rarely accessed, and so there is no need to store such files on the local storage of the organizations. According to a recent survey, IT outsourcing has grown by a staggering 79% as companies seek to reduce costs and focus on their core competencies. Data security and integrity protection in Cloud Computing is a challenging task for the researchers. Data security can be achieved by encrypting sensitive data before outsourcing to remote servers. As such, it is a crucial demand of customers to have strong evidence that the cloud servers still possess their data and it is not being tempered with or partially deleted over time, especially because the internal operation details of the CSP may not be known by cloud customers. The completeness and correctness of customers' data in the cloud is being put at risk due to the following reasons. First, the CSP goal is to make a profit and maintain a reputation. Second, a greedy CSP might delete some of the data or might not store all data in fast storage required by the contract with certain customers, i.e., place it on CDs or other offline media and thus using less fast storage. Third, the cloud infrastructures are subject to wide range of internal and external security threats.

Cloud is boon to computing world but its adoption in IT sector is lack behind due to many issues. The main objective is related to data security which may include the concepts like confidentiality, integrity, data leakage, availability, data access. Cloud computing provides services on-demand-services. Cloud computing delivers four types of basic deployment models: Public Cloud, Private Cloud, Hybrid Cloud, Community Cloud. Three delivery models: SaaS (Software as a Service), PaaS (Platform as a Service), IaaS (Infrastructure as a Service). Cloud computing comes with many benefits like Flexibility, increases collaboration, reduces cost and may more.. Data security is one of the major concerns today due to which the organizations not fully take up this technology. The users do not know where the data is being residing after uploading to cloud and even they do not know who is handling their data. Hence the concerns come into mind that whether their data is confidential. Integrity of the data is lost if the alteration is done by any unauthorized person.

© 2015, IJCSMC All Rights Reserved

If the data is not highly available when the user wants to access the data then it may lead to degradation to the business

The foundations of cloud computing lie in the outsourcing of computing tasks to the third party. It entails the security risks in terms of confidentiality, integrity and availability of data and service. Remote data integrity checking is a primitive to address the issues. For the general case, when the client stores his data on multi-cloud servers, the distributed storage and integrity checking are indispensable. On the other hand, the integrity checking protocol must be efficient in order to make it suitable for capacity-limited end devices.

Cloud storage has been gaining tremendous popularity among individuals and corporations because of its low maintenance cost and on-demand services for the clients. To improve the availability and the reliability of critical data, storing multiple replicas on multiple servers is a commonly used strategy. Currently, several provable data possession (PDP) protocols for multiple replicas of dynamic data have been proposed to ensure the integrity of outsourced multi-copy data, but the efficiency of these protocols on verifying multiple replicas one by one is not satisfactory. In this paper, we propose a provable multiple replication data possession protocol with full dynamics, named MR-DPDP. In MR-DPDP, we utilize a novel authenticated data structure called Merkle hash tree with rank to support both full dynamic data updates and efficient integrity verification. In addition, our construction with RSA signature can support both variable-sized file blocks and public verification..

A number of computing concepts and technologies are combined in Cloud Computing to satisfy the computing needs of users, it provides common business applications online through web browsers, while their data and software's are stored on the servers. This is an approach that is used to maximize the scope or step up capabilities robustly without investing in new infrastructure, sustenance new personnel or licensing new software. It provides tremendous storage for data and rapid computing to customers over the internet. There is fear between the data owner's especially in large organizations that their data possibly misuse by the cloud provider without their knowledge. Data security of the users can be ensured by using the concept of virtual private networks, firewalls, and by enforcing other security policies within its own circumferences.

2. LITERATURE REVIEW

2.1. Title: Efficient Provable Data Possession for Hybrid Clouds (2010)

Provable data possession is a technique for ensuring the integrity of data in outsourcing storage service. In this paper, we propose a cooperative provable data possession

© 2015, IJCSMC All Rights Reserved

scheme in hybrid clouds to support scalability of service and data migration, in which we consider the existence of multiple cloud service providers to cooperatively store and maintain the clients' data. In this work, we focus on the construction of PDP scheme for hybrid clouds, supporting privacy protection and dynamic scalability. We first provide an effective construction of Cooperative Provable Data Possession (CPDP) using Homomorphic Verifiable Responses (HVR) and Hash Index Hierarchy (HIH). This construction uses homomorphic property, such that the responses of the client's challenge computed from multiple CSPs can be combined into a single response as the final result of hybrid clouds. More importantly, a new hash index hierarchy is proposed for the clients to seamlessly store and manage the resources in hybrid clouds. The experimental results also validate the effectiveness of our construction. The code was written in C++ and the experiments were run on an Intel Core 2 processor with 2.16 GHz. All cryptographic operations utilize the QT and bilinear cryptographic library. We propose a fragment structure of CPDP scheme based on following characters:

1) A file is split into n×s sectors and each block (sectors) corresponds to a tag, so that the storage of signature tags can be reduced with the order of s;

2) The verifier can check the integrity of a file by random sampling approach, which is a matter of the utmost importance for large or huge files; and

3) This structure re-lies on homomorphic properties to aggregate the data and tags into a constant size response, which minimizes network communication overheads.

Here we consider the existence of multiple CSPs to cooperatively store and maintain the clients' data, and a publicly verifiable PDP is used to verify the integrity and availability of their stored data in CSPs. The clients are allowed to dynamically access and update their data for various applications, and the verification process of PDP is seamlessly performed for the clients in hybrid clouds. Hence, it is a challenging problem to design a PDP scheme for supporting dynamic scalability. It is used to integrate multiple responses from the different CSPs in cooperative PDP scheme. The total of communication over-heads is not significantly increased.

2.2. Title: On Verifying Dynamic Multiple Data Copies over Cloud Servers (2011)

Currently, many individuals and organizations outsource their data to remote cloud service providers (CSPs) seeking to reduce the maintenance cost and the burden of large local data storage. The CSP offers paid storage space on its infrastructure to store customers' data. Replicating data on multiple servers across multiple data centers achieves a higher level of

scalability, availability, and durability. The more copies the CSP is asked to store, the more fees the customers are charged. Therefore, customers need to be strongly convinced that the CSP is storing all data copies that are agreed upon in the service contract, and the data-update requests issued by the customers have been correctly executed on all remotely stored copies. In this paper we propose two dynamic multi-copy provable data possession schemes that achieve two main goals:

i) They prevent the CSP from cheating and using less storage by maintaining fewer copies, and

ii) They support dynamic behaviour of data copies over cloud servers via operations such as block modification, insertion, deletion, and append.

We prove the security of the proposed schemes against colluding servers. Encrypting sensitive data before outsourcing to remote servers can handle the confidentiality issue. One fundamental advantage of using CC is pay-as-you-go pricing model, where customers pay only according to their usage of the services. This approach suffers from a severe drawback; the communication complexity is linear with the queried data size which is impractical especially when the available bandwidth is limited.

2.3. Title: PORs: Proofs of Retrievability for Large Files (2014)

In this paper, we define and explore proofs of retrievability (PORs). A POR scheme enables an archive or back-up service (prover) to produce a concise proof that a user (verifier) can retrieve a target file F, that is, that the archive retains and reliably transmits file data sufficient for the user to recover F in its entirety. The goal of a POR is to accomplish these checks without users having to download the files themselves. We explore POR protocols here in which the communication costs, number of memory accesses for the prover, and storage requirements of the user (verifier) are small parameters essentially independent of the length of F. In addition to proposing new, practical POR constructions, we explore implementation considerations and optimizations that b ear on previously explored, related schemes. In a POR, unlike a POK, neither the prover nor the verifier need actually have knowledge of F. PORs give rise to a new and unusual security definition whose formulation is another contribution of our work. We view PORs as an important tool for semi-trusted online archives This MAC-based approach is quite efficient in terms of file-expansion overhead, computational costs, and bandwidth. It has an important drawback, though: It does not permit the prover to return a digest of its responses, i.e., to hash or XOR them together.

2.4. Title: Proving Possession and Retrievability within a Cloud Environment: A Comparative Survey (2014)

In this paper we surveyed two core integrity proving schemes in detail along with different methods used for data integrity in both the schemes. Suppose that a CSP offers to store n copies of an owner's file on n different servers to prevent simultaneous failure of all copies. Thus, the data owner needs a strong evidence to ensure that the CSP is actually storing no less than n copies, all these copies are complete and correct, and the owner is not paying for a service that he does not get.

The PB-PMDP scheme supports public verifiability. Generating unique differentiable copies of the data file is the core to design a multi-copy provable data possession scheme. Identical data copies enable the CSP to simply deceive the owner by storing only one copy and pretending that it stores multiple copies. Using a simple yet efficient way, the proposed scheme generates distinct copies utilizing the diffusion property of any secure encryption scheme. There will be an unpredictable complete change in the cipher text, if there is a single bit change in the plaintext. The interaction between the authorized users and the CSP is considered through this methodology of generating distinct copies, where the former can decrypt and access a file copy received from the CSP without recognizing the copy index. Homomorphic linear authenticators (HLAs) are basic building blocks in the proposed scheme. This is a data correctness scheme which involves the encryption of the few bits of data per data block thus reducing the computational overhead on the clients. It's based on the fact that high probability of security can be achieved by encrypting fewer bits instead of encrypting the whole data. The client storage overhead is also minimized as it does not store any data with it and it reduces bandwidth requirements. Hence this scheme suits well for small memory devices and low power devices. The freshness of the response computation by the server is guaranteed by the fact that a challenge is never reused before reboot of the server. Although the protocol does not require exponentiation of the entire file, a local copy of the fingerprints whose size is linear in the number of file blocks must be stored on the verifier side.

2.5. Title: Provable Possession and Replication of Data over Cloud Servers (2010)

In this paper we address this challenging issue and propose Efficient Multi-Copy Provable Data Possession (EMC-PDP) protocols. We prove the security of our protocols against colluding servers. Through extensive performance analysis and experimental results,

© 2015, IJCSMC All Rights Reserved

we demonstrate the efficiency of our protocols. The only advantage of the MR-PDP scheme is the reduction of the tag generation cost which is done only once during the life time of the outsourced storage system. Unfortunately, this contribution resulted in many various limitations that have been explained earlier, and one of these critical resulting limitations is the inability of the authorized users to access the file copies for the opaqueness nature of the CSP.

2.6. Title: Security of Data Dynamics in Cloud Computing (2012)

This brings the problem of ensuring the integrity of data storage in Cloud. Cloud computing share distributed resources via network in the open environment thus it makes security problem. we consider the task of allowing a third party auditor (TPA), on behalf of the cloud client, to verify the integrity of the dynamic data stored in the cloud. It is one of the most frequently used operations in cloud data storage. Public auditability also allows clients to delegate the integrity verification tasks to TPA while they themselves can be unreliable or not be able to commit necessary computation resources performing continuous verifications.

2.7. Title: Efficient Dynamic Provable Possession of Remote Data via Balanced Update Trees (2013)

In this work we develop a novel and efficient scheme, computation and communication over-head of which is an order of magnitude lower than those of other state-of-the-art schemes. Our solution has a number of new features such as a natural support for operations on ranges of blocks, and revision control. The advantages come at the cost of requiring the client to maintain a data structure of modest, but non-constant size. A disadvantage of Merkle hash tree based solutions is that the trees becomes unbalanced after a series of insert and delete requests.

3. CONCLUSIONS

The research area of cloud computing is not only academic, but has also notice the considerable attention from industry due to a number of key advantages: cost effectiveness, low management overhead, immediate access to a wide range of applications, flexibility to scale up and down information technology capacity, and mobility where customers can access information wherever they are, rather than having to remain at their desks. Currently, we are living in a digital world, where a large amount of sensitive data (e.g., personal information, electronic health records, and financial data) is generated by various

© 2015, IJCSMC All Rights Reserved

organizations. Managing such huge amount of data locally at the organization's end is problematic and costly due to the requirements of high storage capacity and qualified personnel. Therefore, cloud service providers (CSPs) offer Storage-as-a-Service as a paid facility to reduce the maintenance cost and mitigate the burden of large local data storage. Through this facility, data owners are enabled to outsource their data to be stored over cloud servers.

The proposed MB-PMDDP scheme based on Secure Hash Algorithm and Elliptic Curve Cryptography encryption algorithm approach outperforms the TB-PMDDP approach derived from a dynamic single-copy PDP models. The TB-PMDDP leads to high storage overhead on the remote servers and high computations on both the CSP and the verifier sides. This scheme significantly reduces the computation time and makes it more practical for applications where a large number of verifiers are connected to the CSP causing a huge computation overhead on the servers.

REFERENCES

[1] Y. Zhu, H. Wang, Z. Hu, G.-J. Ahn, H. Hu, and S. S. Yau, "Efficient provable data possession for hybrid clouds," 2010.
[2] A. F. Barsoum and M. A. Hasan. "On verifying dynamic multiple data copies over cloud servers," 2011.
[3] A. Juels and B. S. Kaliski, Jr., "Pors: Proofs of retrievability for large files," 2007.
[4] Kochumol Abraham, Win Mathew John, "Proving Possession and Retrievability within a Cloud Environment: A Comparative Survey", 2014.
[5] A. F. Barsoum and M. A. Hasan. "Provable possession and replication of data over cloud servers," 2010.
[6] M.Yugandhar, D. Subhramanya Sharma. "Security of Data Dynamics in Cloud Computing" 2012.
[7] Yihua Zhang and Marina Blanton," Efficient Dynamic Provable Possession of Remote Data via Balanced Update Trees", 2013.

© 2015, IJCSMC All Rights Reserved