A Comprehensive Survey on Storage Techniques in Cloud Computing

Amol S. Choure
Department of computer Science & engineering
SGGS IE&T, Nanded-431606 (M. S.), India.

S. M. Bansode
Department of computer Science & engineering
SGGS IE&T, Nanded-431606 (M. S.), India.

ABSTRACT

It has been great development in cloud computing since past few years. It offers different kinds of services for example Software as a Service (SaaS), Platform as a Service (PaaS), Infrastructure as Service (IaaS). Cloud computing enables the user and organizations to store their data remotely and enjoy good quality applications on the demand without having any burden associated with local hardware resources and software managements but it possesses a new security risk towards correctness of data stored at cloud. There has been different techniques which in turn provide correctness of data, for example Merkle Hash Tree (MHT), Distributed erasure-coded data & flexible distributed storage integrity auditing mechanism. This work is based on the survey of different techniques of cloud storage with their benefits, disadvantages and security challenges. This particular study allows you to discover long term research areas and techniques for bettering the current downsides.

General Terms

Cloud computing, Data storage, Security, Privacy.

Keywords

Cloud Computing, Data Security, Storage Techniques, Integrity, Survey, Authentication.

1. INTRODUCTION

Cloud services are available on-demand and often bought on a “pay-as-you go” or subscription basis. Exactly where the hardware and software located and how it works doesn’t matter to the user he only needs to have a desktop with internet connection. Most importantly, the service you use is provided by someone else and managed on your behalf. Why people are moving towards cloud computing because it allows user to access data and resources from any geographical location at any time also it has numerous advantages such as reduced infrastructure costs, scalability, no maintenance and need to pay only for what we use. The innovator of cloud computing is amazon simple storage service (AS3). The different kind of services being offered in cloud computing are Software as a Service (SaaS), Platform as a Service (PaaS), Infrastructure as Service (IaaS). Examples of services includes web hosting, web based emails, google documents and google app engine etc. Deployment models of cloud computing includes public cloud, private cloud, hybrid cloud and community cloud. Though there are many advantages of cloud computing, security concern has become the biggest obstacle in adoption of cloud as data is completely under the control of cloud service provider (CSP), so lack of control. This paper discuses different kinds of storage techniques with their advantages and drawbacks.

2. STORAGE TECHNIQUES IN CLOUD COMPUTING

Different kind of techniques for secure cloud storage are described below.

2.1 Identity-Based Authentication

In Cloud Computing, there is security risk because resources and services are distributed across many clients. Hence, authentication of users along with services is a substantial matter for the trust and security of the cloud computing. Secure socket layer (SSL)Authentication Protocol, once connected in distributed computing, will turn out to be complicated to the point that clients will experience an overburdened point both in communication and computation. This paper [1], described identity-based hierarchical model for cloud computing and its corresponding encryption and signature schemes. Encryption and Signature schemes are projected to accomplish security in cloud communication. It has been observed that authentication protocol based on identity is more efficient and weightless than SSL authentication protocol

2.2 Implicit Security for Online Data Storage

Providing security to data which is stored on distributed servers is of major concern in cloud computing. The conventional (explicit) way to deal with securing data is to store are likely to be broken later on. As a result, explicit security architecture will be inadequate for several applications. In this technique of online data storage using implicit security [2], stored data is divided into two or more pieces and put away at arbitrarily chosen places on the network. This information of exactly where the pieces are kept is known and back it up on a single server and permit admittance upon the utilization of passwords that are expected to be changed repeatedly. At the same time, there is an inclination among clients to keep passwords basic and notable, which leads to brute force attacks. In addition considering that the files on the internet will be aged, keys offering adequate encryption currently to the proprietor of the data only. Access to these pieces relies not only upon the knowledge of a password but also on the information of where the pieces are stored. The partition of data is performed in such a manner that the knowledge of every last one of pieces is obliged to reproduce the data and that none of the individual pieces reveals any valuable information.

2.3 Public Auditing with Complete Data Dynamics support

In the previous work of ensuring remote data integrity there is deficiency of provision of public auditability and dynamic data operation. This approach [3] first of all discovers the possible security threats and difficulties associated with previous work and then constructed a sophisticated public auditing scheme with a protocol that gives provision of dynamic data operation. To achieve efficient data dynamics current proof of storage model is improved by operating the classic Merkle Hash Tree construction for block tag verification. For managing the several auditing task efficiently the technique of bilinear aggregate signature is used. Thus TPA will be capable of performing multiple auditing tasks at a time in multiuser setting. This scheme is extremely proficient and provably secure.

2.4 Efficient Third Party Auditing

Cloud computing is a technology which allows the user to store their data at remote location and gives access to different services so that user can develop his own application by using cloud platform and can store related data at cloud by using cloud infrastructure though cloud service providers enables us to do so but security is of main concern thus user is not confident about whether the data stored by him possess integrity or not so to check integrity of data an efficient auditing mechanism is presented in [4] .Author has used the technique of symmetric key cryptography which enable TPA to perform auditing without the local copy of users data and thus reduces computation and transmission overheads also integrating encryption with hashing prevents TPA from learning knowledge of data stored during auditing.

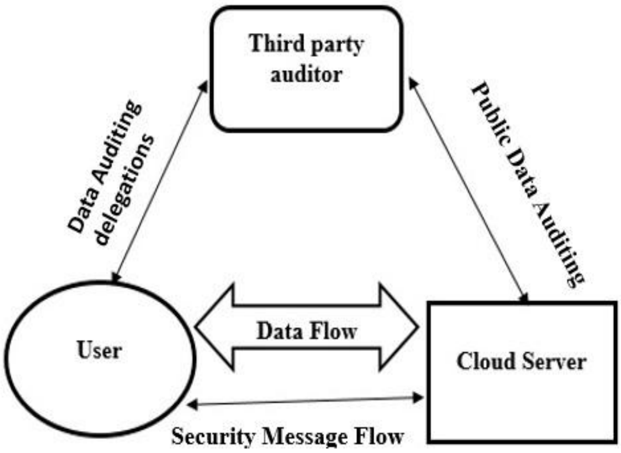

Fig 1: Architecture of third party auditing scheme.

This scheme achieves batch auditing multiple auditing requests from different users are handled also this scheme supports dynamic data operations such as insert, delete, update. The architecture of efficient third party auditing protocol is as depicted in Fig 1.

2.5 Effective and Secure Storage Protocol

Cloud service provider offers large storage space with low cost, so it reduces burden of local data storage but when data is outsourced user loses control of their data which brings a new security risk toward integrity and confidentiality so to address these issues a new security protocol is described here [5], which is based on Sobol Sequence and Elliptic Curve Cryptography. This scheme permits third party auditor to periodically confirm integrity of the data kept at cloud without recovering original data. It generates probabilistic proofs of integrity by challenging random sets of blocks from the server, which significantly reduces the communication and I/O costs. The challenge response protocol exchanges small amount of data on network also it maintains confidentiality and does not reveals any useful information to malicious parties.

2.6 Ways of Dynamically Storing Data to Cloud

As in cloud computing user do not possess copy of outsourced data, data storage is not trustable and also it is not an easy task to preserve all data securely specially when there is high demand, to consider and solve these problems a novel protocol system is described in [6], which uses data reading protocol algorithm to verify data honesty. Service provider gives provision of checking data security with the help of automatic data reading algorithm. For recovering data in future automatic data reading algorithm is used.

2.7 Storage Security of Data

Since in the case of cloud computing resources tend to be distributed all over web widely, which produces extreme troubles in providing data security. Sending data on web will be unsafe because of the invader assault. So data encryption represents an essential purpose in cloud environment. The proposed strategy [7] contains some crucial security benefits that are supplied to cloud framework. A system structure is made up of three data backups for recuperation of data. These backups situated in distant area from primary server. This technique utilized SHA Hash algorithm for encryption, SFSPL algorithm for splitting files and GZIP algorithm for compression. In this way, a guaranteed cross platform is proposed for distributed computing.

2.8 Towards Secure and Dependable Storage

Cloud computing storage service allows user to store their data at cloud without worrying about maintenance but there are some issues concerning to correctness of data as the user doesn’t possess local copy of data. Proposed system [8] permits clients inspecting the cloud data storage. This mechanism uses homomorphic token with reed-Solomon erasure correcting code technique, which promises the correctness assurance and also identifies which server is misbehaving. This design also extended to support block level dynamic operations. If users do not have sufficient resources and time available for processing data then user can delegate this task to TPA. Thus this technique allows user to store data at remote place securely and supports dynamic operations such as insert, update & delete.

2.9 Optimal Cloud Storage

An efficient scheme to accomplish storage service optimality with resource provider, consumer’s lifecycle is explored in [9]. Recommended scheme gives definition of system used for storage, optimality of storage and administrator architecture, which is used for storage and is aware of optimality. Architecture consists of three components data processor that processes data before it is sent to the cloud, data verifier that checks whether the data in cloud has been tampered and token generator that generates the token which helps the storage providers to retrieve segments of consumer data.

2.10 Method to Access and Store Small Files

Hadoop distributed file system (HDFS) is developed for supporting internet services. Inherent HDFS does not do well for small size files with large numbers and this is the main issue to be focused. In this work [10] several reasons are analyzed associated with small file problem such as Name Node of HDFS suffered with heavy load imposed by large numbers of small files, correlations among small files are not taken into consideration for data placement and no optimization mechanism, for example prefetching, is provided to improve I/O performance. Based on correlation feature files are divided into three types structurally related files, logically related files and independent files for structurally-related small files, file merging and prefetching scheme is applied and for managing logically-related small files, file grouping and prefetching scheme is used.

2.11 File Storage security Maintenance

In this work [11] to ensure the data security a framework is introduced, which usages distributed scheme. Proposed framework includes a master server and an arrangement of slave servers. There is no direct connection in the middle of clients and slave servers. Master server is responsible for preparing the clients appeals and at slave server chunking operation is completed. Chunking operation is responsible for storing duplicates of records to give backup of data for recovery of document in future. Clients can likewise perform powerful and dynamic data operations. Client's document is kept at main server as tokens and records were chunked on slave servers for file recovery. Subsequently proposed scheme accomplished storage integrity and accessibility of data for that token generation and merging algorithms were used.

2.12 Privacy Preserving Public Auditing

Outsourcing of data ultimately relinquishes the control of data from user and the lot of data is in control of the cloud server. The cloud server is succeeded by cloud service provider which is a different administrative entity, so ensuring the data integrity is of prime importance. The work presented [12] studies the problems of ensuring data storage correctness and proposes an effective and secure scheme to address these issues. A third party auditor (TPA) is introduced securely, who on behalf of users request will periodically verify integrity of the data stored on cloud server. There will not be any online burden on user and security of data will be maintained as the data will not be shared directly with the third party auditor. A homomorphic encryption scheme is used to encrypt the data by using Elliptic curve Cryptography (ECC) which will be shared with the TPA. ECC provides efficient and secure solutions for the cloud storage servers. It leads to fast computation time, reducing in processing power, save the storage and bandwidth. The results can be further extended to enable the third party auditor to do multiple auditing tasks.

3. CONCLUSION

Cloud Computing is a web based developing computing model, which permits the clients to access data and resources from any geographical location at any time on subscription basis. Despite the fact that Cloud gives advantages to clients, security and privacy of data stored in cloud are still major issues in cloud computing. Cloud computing is a more useful and profitable than the prior conventional storage frameworks particularly in adaptability, cost diminishment, portability and functionality requirements. This paper introduced an overview on secure storage techniques in Cloud Computing. Initially a few storage strategies that give security to data in cloud have been discussed and furthermore highlighted the need for future research on storage methods to provide vastly improved security and accountability.

4. REFERENCES

[1] Li H, Dai Y and H. Yang, "Identity-Based Authentication for Cloud Computing", M. G. Jaatun, G. Zhao, and C. Rong (Eds.): Cloud Computing, Lecture Notes in Computer Science, 2009.

[2] Parakh A, and Kak S, "Online data storage using implicit security", Information Sciences, 2009.

[3] Wang Q, Wang C, Kui Ren and Wenjing Lou, "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing", IEEE Transactions on Parallel and Distributed Systems, 2011.

[4] Prasanthi, C. Balasubramanian, S. Kimsukha Selvi, K. Kala , "An Efficient Auditing Protocol for Secure Data Storage in Cloud Computing", Proceedings of the World Congress on Engineering 2014.

[5] Subramanian R , Kumar S P, "An efficient and secure protocol for ensuring data storage security in Cloud Computing", International Journal of Computer Science Issues, 2011.

[6] Dinesh C, "Data Integrity and Dynamic Storage Way in Cloud Computing", Distributed, Parallel, and Cluster Computing, 2011.

[7] Sajithabanu S, Raj E G P, "Data Storage Security in Cloud", International Journal of Computer Science and Technology, 2011.

[8] Wang C, Wang Q, Kui Ren, Ning Cao and Wenjing Lou, "Toward Secure and Dependable Storage Services in Cloud Computing", IEEE Transactions on Services Computing, 2012.

[9] Spillner J, Müller J and Schill A, "Creating optimal cloud storage systems", Future Generation Computer Systems, 2012.

[10] Zheng Q, Dong B et al. (2012), "An optimized approach for storing and accessing small files on cloud storage", Journal of Network and Computer Applications, 2012.

[11] Deshmukh P M, Gughane A S et al, "Maintaining File Storage Security in Cloud Computing", International Journal of Emerging Technology and Advanced Engineering, 2012.

[12] Cong Wang, Sherman S.M.Chow, Qian Wang, Kui Ren, and Wenjing Lou, "Privacy Preserving Public Auditing for Secure Cloud Storage", IEEE, February 2013.

[13] Pradnya B. Godhankar, Deepak Gupta, "Review of Cloud Storage Security and Cloud Computing Challenges", International Journal of Computer Science and Information Technologies, 2014.

[14] Kishan Lathkar, Ambulgekar H. P, "Public Auditability and Privacy preserving in Cloud Storage'', Journal of Information Security Research, 2015