A Secure Data Sharing Technique using Attribute based Encryption Algorithm in Cloud Storage Server

[1]Rajendrababu Nagavarapu

[1]M.Tech Scholar, Department of Computer Science and System Engineering,

Andhra University College of Engineering (A), Visakhapatnam, Andhra Pradesh, India.

Abstract: Data sharing is a challenging issue in public cloud storage systems with proper security. In existing system, data integrity auditing scheme that realizes data sharing with sensitive information hiding. However this provide only the security for the sensitive data where it has poor accessible structure during data sharing. In this paper, attribute based encryption technique is used for data sharing technique with the file permission technique. Ciphertext-Policy Attribute-Based Encryption (CP-ABE) has been adopted as a promising technique to provide flexible, fine-grained and secure data access control for cloud storage. CP-ABE scheme is adopted in a large-scale cloud storage system. This improves the flexible and fast access of the shared data from the cloud

Index Terms: Ciphertext policy, attribute based encryption, data sharing, public key generator.

INTRODUCTION

Cloud computing, along with big data, is the biggest buzz of tech world these days. Just like Internet and Web took the world by storm in 1990s and early 2000s, and smart phones have shaped the new world order in communication in last decade or so, cloud computing is also expected to revolutionize the way in which businesses would be conducted and services would be provided to potential consumers. Almost all major tech giants like Microsoft, Google, Amazon and Apple provide cloud services to their consumers and even to other major businesses – for instance Netflix uses Amazon Web Services for hosting their streaming services. The cloud industry grew by 16.5% last year and is expected to rise to $204B by the end of 2016 according to tech research company Gartner Inc. But despite all its hype and usage, the concept of cloud computing is pretty elusive and its definition quite vague. In simplistic terms, cloud provides remote computing and storage services from a pool of shared resources to its consumers.

NIST, in its 16th and final Definition of Cloud Computing, codenamed SP (Special Publication) 800- 145, has highlighted five key characteristics that every cloud technology should incorporate. They are:

1. On-demand Self-Service: The consumer should be able to change the provisioned computing capabilities like number of cloud clusters and online storage unilaterally, without the intervention of human service provider.

2. Broad Network Access: Cloud services should be easily available through standard Internet mechanisms on all kinds of devices like mobiles, desktops, laptops, workstations etc.

3. Resource Pooling: It must be able to serve multiple consumers concurrently in location-independent way from same physical resources which are separated on logical level in a secure manner.

4. Rapid Elasticity: Resources should be provisioned and released on demand, and at any point of time, the consumer should have exactly the amount of resources he needs for his product. In essence, consumer should be able to scale up and down the resources, remove or add users, provision for more machines or storage in a seamless manner, and to him the resources should seem to be infinite, any amount of which can be provisioned at any point of time.

5. Measured Service: Cloud services should follow the pay-as-you-go pricing model. All consumption and usage of cloud resources should be monitored, logged and reported to consumer accordingly, and controlled from both sides under some agreement. A user should only be charged for what he used and also if there are limits on usage per user, it is the responsibility of service provider that such limits are not breached under normal circumstances.

RELATED WORK

Cloud storage services[11] such as Dropbox, Google Drive, and Microsoft One Drive provide users with a convenient and reliable way to store and share data from anywhere, on any device, and at any time. The cornerstone of these services is the data synchronization (sync) operation which automatically maps the changes in users' local file systems to the cloud via a series of network communications in a timely manner. If not designed properly, however, the tremendous amount of data sync traffic can potentially cause (financial) pains to both service providers and users.

This paper addresses a simple yet critical question: Is the current data sync traffic of cloud storage services efficiently used? We first define a novel metric named TUE to quantify the Traffic Usage Efficiency of data synchronization. Based on both real-world

traces and comprehensive experiments, we study and characterize the TUE of six widely used cloud storage services. Our results demonstrate that a considerable portion of the data sync traffic is in a sense wasteful, and can be effectively avoided or significantly reduced via carefully designed data sync mechanisms. All in all, our study of TUE of cloud storage services not only provides guidance for service providers to develop more efficient, traffic economic services, but also helps users pick appropriate services that best fit their needs and budgets.

As tools for personal storage [12], file synchronization and data sharing, cloud storage services such as Dropbox have quickly gained popularity. These services provide users with ubiquitous, reliable data storage that can be automatically synced across multiple devices, and also shared among a group of users. To minimize the network overhead, cloud storage services employ binary diff, data compression, and other mechanisms when transferring updates among users. However, despite these optimizations, we observe that in the presence of frequent, short updates to user data, the network traffic generated by cloud storage services often exhibits pathological inefficiencies. Through comprehensive measurements and detailed analysis, we demonstrate that many cloud storage applications generate session maintenance traffic that far exceeds the useful update traffic. We refer to this behavior as the traffic overuse problem. To address this problem, we propose the update-batched delayed synchronization (UDS) mechanism. Acting as a middleware between the user's file storage system and a cloud storage application, UDS batches updates from clients to significantly reduce the overhead caused by session maintenance traffic, while preserving the rapid file synchronization that users expect from cloud storage services. Furthermore, we extend UDS with a backwards compatible Linux kernel modification that further improves the performance of cloud storage applications by reducing the CPU usage.

While many public cloud providers[13] offer pay-as-you-go computing, their varying approaches to infrastructure, virtualization, and software services lead to a problem of plenty. To help customers pick a cloud that fits their needs, we develop CloudCmp, a systematic comparator of the performance and cost of cloud providers. CloudCmp measures the elastic computing, persistent storage, and networking services offered by a cloud along metrics that directly reflect their impact on the performance of customer applications. CloudCmp strives to ensure fairness, representativeness, and compliance of these measurements while limiting measurement cost. Applying CloudCmp to four cloud providers that together account for most of the cloud customers today, we find that their offered services vary widely in performance and costs, underscoring the need for thoughtful provider selection. From case studies on three representative cloud applications, we show that CloudCmp can guide customers in selecting the best-performing provider for their applications.

The increasing popularity of cloud storage services has lead companies that handle critical data to think about using these services for their storage needs. Medical record databases, power system historical information and financial data are some examples of critical data that could be moved to the cloud. However, the reliability and security of data stored in the cloud still remain major concerns. In this paper we present DEPSKY, a system that improves the availability, integrity and confidentiality of information stored in the cloud through the encryption, encoding and replication of the data on diverse clouds that form a cloud-of-clouds. We deployed our system using four commercial clouds and used PlanetLab to run clients accessing the service from different countries. We observed that our protocols improved the perceived availability and, in most cases, the access latency when compared with cloud providers individually. Moreover, the monetary costs of using DEPSKY on this scenario is twice the cost of using a single cloud, which is optimal and seems to be a reasonable cost, given the benefits.

PROBLEM DEFINITION

We investigate how to achieve data sharing with sensitive information hiding in remote data integrity auditing, and propose a new concept called identity-based shared data integrity auditing with sensitive information hiding for secure cloud storage. In such a scheme, the sensitive information can be protected and the other information can be published. It makes the file stored in the cloud able to be shared and used by others on the condition that the sensitive information is protected, while the remote data integrity auditing is still able to be efficiently executed.

DISADVANTAGE

- Increase computational time on storing the data.
- Increase computational resource on processing.
- Poor and direct accessing structure on documents from the cloud.

PROPOSED SYSTEM

In this paper, attribute based encryption technique is used for data sharing technique with the file permission technique. Ciphertext-Policy Attribute-Based Encryption (CP-ABE) has been adopted as a promising technique to provide flexible, fine-grained and secure data access control for cloud storage. CP-ABE scheme is adopted in a large-scale cloud storage system. This improves the flexible and fast access of the shared data from the cloud.

ADVANTAGES

- Reduce computational time on storing the data.
- Reduced computational resource on processing.
- Improved accessing structure.
- Flexible and fast data sharing process.

LITERATURE REVIEW

Cloud service providers determine the access control mechanisms for data on the cloud. Access control is a procedure that restricts, denies or allows access to system. In the cloud, data security is crucial to protect against inside attack, denial of service attack and collision attack. Traditionally different expressive access control policies are used to protect data stored locally and data stored remotely. The approaches include Discretionary Access Control (DAC), Mandatory Access Control (MAC), Role-Based Access Control (RBAC), and Attribute-Based Access Control (ABAC). In DAC, users are given complete control over resources on the basis of user identity. The use of DAC is not feasible when the size of the network and the number of users increase or when data is distributed acrdifferent servers. In RBAC, access is based on particular roles and varies depending on the user. A role is assigned to different tasks, for example members of staff have different roles.

RBAC is not feasible because all entities have the right to access and large groups would have same type of access. ABAC considers attributes based on user requests including names and value pairs and are associated with actions, users, subjects, objects, contexts and policies. ABAC is more flexible, secure, and scalable and works in a hierarchical fashion. ABAC solved the RBAC problem of assigning privileges to a user. However, such access control schemes or the use of a server as a reference monitor cannot be applied in cloud environments because clouds have plenty of resources, lot of dynamic users and flexible construction because every autonomous system has its own security policy. As networks grow and the number of users increases, a more complex structure must be created to improve the performance and reliability of stored data. The data are replicated across several locations and stored in distributive fashion across many servers. This creates a lack of confidentiality and security. The only method for protecting sensitive data across multiple sites is to encrypt the data before uploading to the server. Data stored on the cloud must be protected through different mechanisms. One of the vital techniques is public key encryption. In the traditional public key infrastructure, the data owner encrypts the data with the data user public key, before uploading it to the cloud. When a data user sends a request to access data on the cloud, the cloud decrypts the cipher text with the private key. There are two major disadvantages with this technique. First, for encryption, the data owner must obtain the data user public key before uploading. Second, because the same plaintext is used with different public keys, the storage overhead becomes excessive.

ATTRIBUTE BASED ENCRYPTION: The Sahai-Waters (ABE) cryptosystem as implemented in this paper is specifically detailed. Attribute-centered Encryption can be considered as a generalization of identity-headquartered Encryption (IBE). In IBE a person identification is a string which is similar to “bobsmith@yahoo.com". A celebration in the method can encrypt a message to this designated person with handiest the competencies of the recipient identity and the procedure public parameters. In exact the encryption algorithm ought no longer to have entry to a separate public key certificate of the recipient. In Attribute-established Encryption, user identification consists of a collection, S, of strings which serve as descriptive attributes of the user. For illustration, a person identity could include attributes describing their university, department and job function. A get together in the system can then specify an extra set of attributes for the receiver to decrypt a message if his identity S has at the least single set of attributes with the set S0.
The user needs to get related with each the parties before getting the set of keys. The role of KGC is to authenticate customers along with the distribution of the set of attribute keys in order that user is ready to generate secret key via combining the important thing accessories received from the each authorities.

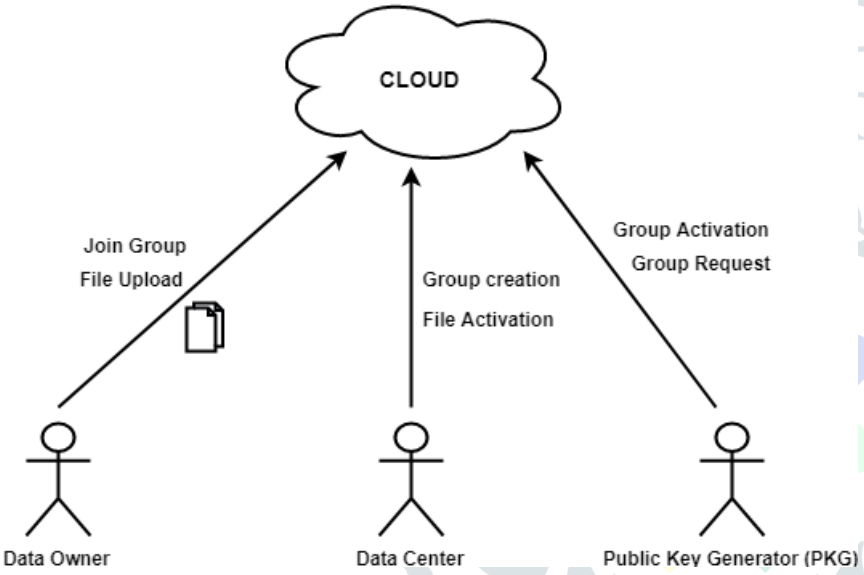
SECURITY BASED ON ABE FOR DATA SHARING: In proposed a mediated Cipher text-policy Attribute-head quartered Encryption (CP-ABE) which extends CP-ABE with instantaneous attribute revocation. Additionally, they reveal how user can observe the proposed CP-ABE scheme to soundly manage private wellbeing records (PHRs).
In offered the inspiration of Fuzzy identification founded Encryption, which enables the error-tolerance between the identification of secret key and the general public key used for encrypting a cipher text. Two realistic functions of Fuzzy IBE of encryption using biometrics and attribute-based encryption and also provided the development of a Fuzzy IBE scheme that makes use of set overlap as the gap metric between identities. Ultimately the proved scheme is the Selective identity mannequin with the aid of decreasing an assumption that can be viewed as a modified variant of the Bilinear Decisional Diffie Hellman assumption. Confidential information is shared and stored on websites on the web, on these websites will need to encrypt data stored. Encrypt knowledge to a problem is that it selectively best a rough-grained level will also be shared (i.e., supply your private key to yet another occasion). The attribute Key-satisfactory-grained policy-founded encryption (KP-Abe) is to share encrypted data to advance a new cryptosystem. Elements of

our cryptosystem texts are represented and secret keys which might be capable to decrypt cipher strength texts customers are associated with access control constructions.

In provided procedure for Cipher textual content-policy Attribute situated Encryption. The approach enables for a new sort of encrypted access control the place person exclusive keys are distinct by using a suite of attributes and a party encrypting information can specify a policy over these attributes specifying which users are competent to decrypt.

In presented a dispensed KP-ABE scheme that solves the key escrow quandary in a multi authority system. In this process, all (disjoint) attribute authorities are participating in the important thing new release protocol in a distributed way such that they cannot pool their data and link multiple attribute units belonging to the identical consumer. One drawback of this kind of entirely distributed process is the performance degradation. Due to the fact that there is no centralized authority with master secret understanding, all attribute authorities will have to communicate with the other authorities in the procedure for the reason that the grasp key is a centralized authority with understanding, procedure presenting all officers to generate secret key a person to be in contact with different officers. M to generate a user's secret key.

ARCHITECTURE



There are three four entity in the architecture, they are data owner, data center, and public key generator. The Data owner request to join in the group and Public Key Generator (PKG) allows permission for joining into the group, then the data owner can upload the file. Data center will create the group and activates the file for further process. Public Key generator will activate the group and handles the group request.

Data Owner: When the client wants to send the information and desires to add it into the external data storage center for data sharing or for cost saving. A knowledge proprietor is liable for outlining (attribute founded) access coverage and implementing it on its possessing data by means of encrypting the data beneath the coverage before distributing it. Data owner should get key from key generator and encrypt the file. Encryption is the process of converting the data into cipher text that can't be simply understood with the aid of unauthorized individuals

Data Storing Centre: This component provides the information to the service provider. It is accountable for controlling the accesses from external customers to the storing knowledge and supplying corresponding contents offerings. The data storing center is a different key authority that generates personalized consumer key with the Key Generation Center and revokes attribute staff keys to legitimate customers per every attribute which might be used to enforce a high-quality-grained consumer entry control. Knowledge storage facilities presents offsite file and tape storage.

User: This is an entity who wishes to access the information. If a consumer possesses a suite of attributes enjoyable the access coverage of the encrypted information outlined by way of the data proprietor and it is not revoked in any of the attribute businesses, then user has to decrypt the cipher text to obtain the information

Key Generation: The key authority center generates public and secret parameters for CP-ABE. This will be used for revoking and updating the user key attributes. It provides various access rights to individual customers with respect the attributes. Key iteration is the system of generating keys for cryptography.

The structure of the Attribute based data sharing is shown in Figure.3. The entities involved are administrator and customers which stand as UI for the approach. Key generation Centre (KGC) is a key authority that generates public and secret parameters for CP-ABE. Data storage center is the component that presents a knowledge sharing carrier. The information storing center is one more key authority that generates customized consumer key with the KGC and revoked attribute staff keys to legitimate customers per each attribute that are used to provide a fine-grained user access control. The client who owns information and desires to add it into the external data storage center for data sharing or for cost saving. A knowledge proprietor is responsible for defining entry policy and also it possesses data by encrypting the information below the policy earlier than distributing it. Consumer is an entity who wishes to access the information

CONCLUSION

The most challenging issue in public cloud storage systems are due to proper security. Maintaining the integrity and auditing the documents simultaneously are the difficult task to do. However, this provide only the security for the data sharing technique and results in poor accessible structure. In this paper, Attribute based encryption technique is used for data sharing with file permission approach is proposed. Ciphertext-Policy Attribute based Encryption (CP-ABE) has been adopted as a challenging technique to provide flexible, fine-grained and secure data access control for cloud storage. This improves the flexible and fast access of the data shared from the cloud storage server.

REFERENCES

[1] K. Ren, C. Wang, and Q. Wang, "Security challenges for the public cloud", IEEE Internet Computing, vol. 16, no. 1, pp. 69–73, Jan 2012.

[2] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song, "Provable data possession at untrusted stores", in Proceedings of the 14th ACM Conference on Computer and Communications Security, ser. CCS '07, 2007, pp. 598–609.

[3] A. Juels and B. S. Kaliski, "Pors: Proofs of retrievability for large files", in Proceedings of the 14th ACM Conference on Computer and Communications Security, ser. CCS '07, 2007, pp. 584–597.

[4] H. Shacham and B. Waters, "Compact proofs of retrievability", J. Cryptology, vol. 26, no. 3, pp. 442–483, Jul. 2013.

[5] C. Wang, S. S. M. Chow, Q. Wang, K. Ren, and W. Lou, "Privacy-preserving public auditing for secure cloud storage", IEEE Transactions on Computers, vol. 62, no. 2, pp. 362–375, 2013.

[6] S. G. Worku, C. Xu, J. Zhao, and X. He, "Secure and efficient privacy-preserving public auditing scheme for cloud storage", Comput. Electr. Eng., vol. 40, no. 5, pp. 1703–1713, Jul. 2014.

[7] C. Guan, K. Ren, F. Zhang, F. Kerschbaum, and J. Yu, "Symmetric-key based proofs of retrievability supporting public verification", in Computer Security – ESORICS 2015. Cham: Springer International Publishing, 2015, pp. 203–223.

[8] W. Shen, J. Yu, H. Xia, H. Zhang, X. Lu, and R. Hao, "Light-weight and privacy-preserving secure cloud auditing scheme for group users via the third party medium", Journal of Network and Computer Applications, vol. 82, pp. 56–64, 2017.

[9] J. Sun and Y. Fang, "Cross-domain data sharing in distributed electronic health record systems", IEEE Transactions on Parallel and Distributed Systems, vol. 21, no. 6, pp. 754–764, June 2010.

[10] G. Ateniese, R. D. Pietro, L. V. Mancini, and G. Tsudik, "Scalable and efficient provable data possession", in Proceedings of the 4th international conference on Security and privacy in communication netowrks, 2008, pp. 1–10.

[11] Z. Li, C. Jin, T. Xu, C. Wilson, Y. Liu, L. Cheng, Y. Liu, Y. Dai, and Z.-L. Zhang, "Towards Network-level Efficiency for Cloud Storage Services", in IMC. ACM, 2014.

[12] Z. Li, C. Wilson, Z. Jiang, Y. Liu, B. Y. Zhao, C. Jin, Z.-L. Zhang, and Y. Dai, "Efficient Batched Synchronization in Dropbox-like Cloud Storage Services", in Middleware. ACM/IFIP/USENIX, 2013.

[13] A. Li, X. Yang, S. Kandula, and M. Zhang, "CloudCmp: Comparing Public Cloud Providers", in IMC. ACM, 2010.

[14] Bessani, M. Correia, B. Quaresma, F. Andre, and P. Sousa, "DepSky: Dependable and Secure Storage in a Cloud-of-Clouds", in EuroSys. ACM, 2013.