JOURNAL FOR INNOVATIVE DEVELOPMENT IN PHARMACEUTICAL AND TECHNICAL SCIENCE

Volume-2,Issue-4 (Apr-2019)
ISSN (O) :- 2581-6934

A Secure Deduplication and Resource Allocation with SLA

Mrs. Jenefa [1], I.SHARLIN SORNA [2]
[1]Assistant Professor ME , [2]Student ME
[1 2] Computer Science and Technology, Rajas Institute of technology for women.

Abstract— Cloud computing has lately arisen as a significant service to manage applications efficiently over the Internet. As complexity, strength and heterogeneity of cloud environment is rising promptly, it makes cloud systems diffident and insecure. In order to conserve the privacy of consumers, data are habitually accumulated in cloud in an encrypted manner. However, encrypted data introduce new challenges for cloud data deduplication, which becomes crucial for big data storage and processing in cloud. A scheme to deduplicate encrypted data stored in cloud based on ownership challenge and proxy re-encryption had been proposed in Phase I work. It integrates cloud data deduplication with access control. However, SLAs were not considered to improve the QoS. Nowadays various cloud providers offer pay per use cloud services that requires Quality of Service (QoS) management to efficiently monitor and measure the delivered services through Internet of Things (IoT) and thus needs to follow Service Level Agreements (SLAs). To overcome these problems, cloud systems require self-management of services. A SLA-aware autonomic resource management technique called STAR which mainly focuses on reducing SLA violation rate for the efficient delivery of cloud services. The performance of the proposed technique has been evaluated through cloud environment. The tentative results demonstrate that STAR is proficient in dropping SLA violation rate and in enhancing other QoS parameters which upshot the effectual cloud service distribution.

Keywords - Internet of Things (IoT), Service Level Agreements (SLAs), Quality of Service (QoS), data deduplication, SLA-aware autonomic resource management technique (STAR), proxy re-encryption.

1. INTRODUCTION

Cloud offers three types of services such as Infrastructure-as-a-Service (IaaS), Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS) and therefore it requires management of Quality of Service (QoS) to efficiently monitor and measure the delivered services to meet Service Level Agreements (SLAs). In Cloud environment, uncertainty and dispersion of resources encounters problems in efficient management of resources, which is caused due to many reasons such as: i) heterogeneity (due to different type of resources and scheduling techniques), ii) dynamism (detect and fulfill the requirements of application at runtime) and iii) failures (failure of system or resources which leads to performance degradation). However, present cloud computing systems and management techniques are unable to handle above mentioned problems efficiently at runtime. An autonomic system provides a solution to this problem by offering the environment in which applications can be managed efficiently by fulfilling QoS requirements of applications without human involvement. Thus, autonomic cloud system becomes self-managed to overcome the above challenges and to provide reliable, secure and cost efficient services to end users.

Currently, cloud services are provisioned and scheduled according to resources' availability without ensuring the expected performances. The cloud provider should evolve its ecosystem in order to meet QoS requirements of each cloud component. To realize this, there is a need to consider two important aspects which reflect the complexity introduced by the cloud management should be considered: firstly, QoS-aware and secondly autonomic management of cloud services. QoS-aware aspect involves the capacity of a service to be aware of its behavior to ensure the elasticity, high availability, reliability of service, cost, time etc. as mentioned in SLA . Autonomic implies the fact that the service is able to self-manage itself as per its environment

needs. Thus, maximizing cost-effectiveness and resource utilization for applications while ensuring performance and other QoS guarantees, requires leveraging important and extremely challenging tradeoff.

The most important and popular cloud service is data storage service. Cloud users upload personal or confidential data to the data center of a Cloud Service Provider (CSP) and allow it to maintain these data. Since intrusions and attacks towards sensitive data at CSP are not avoidable, it is prudent to assume that CSP cannot be fully trusted by cloud users. Moreover, the loss of control over their own personal data leads to high data security risks, especially data privacy leakages. Due to the rapid development of data mining and other analysis technologies, the privacy issue becomes serious. Hence, a good practice is to only outsource encrypted data to the cloud in order to ensure data security and user privacy. But the same or different users may upload duplicated data in encrypted form to CSP, especially for scenarios where data are shared among many users. Although cloud storage space is huge, data duplication greatly wastes network resources, consumes a lot of energy, and complicates data management. The development of numerous services further makes it urgent to deploy efficient resource management mechanisms. Consequently, deduplication becomes critical for big data storage and processing in the cloud.

Deduplication has proved to achieve high cost savings, e.g., reducing up to 90-95 percent storage needs for backup applications and up to 68 percent in standard file systems. Obviously, the savings, which can be passed back directly or indirectly to cloud users, are significant to the economics of cloud business. How to manage encrypted data storage with deduplication in an efficient way is a practical issue.

The primary aim of this scheme is to develop a SLA-aware autonomic cloud resource management technique called **STAR** ***(S**LA-aware autonomic **T**echnique for **A**llocation of **R**esources)* for effective scheduling of resources which considers SLA violation rate along with other QoS parameters like execution time, cost, latency, reliability and availability.

The objectives are:

- To save cloud storage and preserve the privacy of data holders by proposing a scheme to manage encrypted data storage with deduplication
- An effective approach to verify data ownership and check duplicate storage with secure challenge and big data support
- To integrate cloud data deduplication with data access control in a simple way, reconciling data deduplication and encryption
- To propose an autonomic resource management technique for execution of heterogeneous workloads by considering generic property of self-management
- To optimize the above mentioned QoS parameters
- To reduce SLA violation rate and improve user satisfaction by fulfilling their QoS requirements.

1.2 DATA DEDUPLICATION

In cloud computing, data deduplication is a specialized data compression technique for eliminating duplicate copies of repeating data. The synonymous terms are intelligent compression and single-instance storage. This technique is used to improve storage utilization and can also be applied to network data transfers to reduce the number of bytes that must be sent. In the deduplication process, unique chunks of data, or byte patterns, are identified and stored during a process of analysis. As the analysis continues, other chunks are compared to the stored copy and whenever a match occurs, the redundant chunk is replaced with a small reference that points to the stored chunk. Given that the same byte pattern may occur dozens, hundreds, or even thousands of times, the amount of data that must be stored or transferred can be greatly reduced.

This type of deduplication is different from that performed by standard file-compression tools, such as LZ77 and LZ78(Lampel-Ziv). Whereas these tools identify short repeated substrings inside individual files, the intent of storage-based data deduplication is to inspect large volumes of data and identify large sections – such as entire files or large sections of files – that are identical, in order to store only one copy of it. This copy may be additionally compressed by single-file compression techniques. For example, a typical email system might contain 100 instances of the same 1 MB

ISSN:-2581-6934|www.jidps.com

(megabyte) file attachment. Each time the email platform is backed up, all 100 instances of the attachment are saved, requiring 100 MB storage space. With data deduplication, only one instance of the attachment is actually stored; the subsequent instances are referenced back to the saved copy for deduplication ratio of roughly 100 to 1.

Deduplication has proved to achieve high cost savings., reducing up to 90-95% storage needs for backup applications and up to 68% in standard file systems Obviously, the savings, which can be passed back directly or indirectly to cloud users, are significant to the economics of cloud business. How to manage encrypted data storage with deduplication in an efficient way is a practical issue. However, current industrial deduplication solutions cannot handle encrypted data. Existing solutions for deduplication suffer from brute-force attacks. They cannot flexibly support data access control and revocation at the same time. Most existing solutions cannot ensure reliability, security and privacy with sound performance.

Existing System

Although the existing schemes aim at providing integrity verification for different data storage systems, data dynamics has not been fully addressed. How to achieve a secure and efficient design to seamlessly integrate these two important components for data storage service remains an open challenging task in Cloud storage.

Disadvantages of existing system:

Although the infrastructures under the cloud are much more powerful and reliable than personal computing devices, they are still facing the broad range of both internal and external threats for data integrity.

Second, there do exist various motivations for CSP to behave unfaithfully toward the cloud users regarding their outsourced data status.

In particular, simply downloading all the data for its integrity verification is not a practical solution due to the expensiveness in I/O and transmission cost across the network. Besides, it is often insufficient to detect the data corruption only when accessing the data, as it does not give users correctness assurance for those unaccessed data and might be too late to recover the data loss or damage.

Encryption does not completely solve the problem of protecting data privacy against clous service provider but just reduces it to the complex key management domain. Unauthorized data leakage still remains possible due to the potential exposure of decryption keys.

Proposed System:

proposed the use of the convergent encryption, i.e., deriving keys from the hash of plaintext. Then, Storer et al. pointed out some security problems, and presented a security model for secure data deduplication. However, these two protocols focus on server-side deduplication and do not consider data leakage settings, against malicious users.

Advantage:

1) As a rising subject, cloud storage is playing an increasingly important role in the decision support activity of every walk of life.
2) Get Efficient Item set result based on the deduplication.

IMPLEMENTATION

Implementation is the stage of the project when the theoretical design is turned out into a working system. Thus it can be considered to be the most critical stage in achieving a successful new system and in giving the user, confidence that the new system will work and be effective. The implementation stage involves careful planning, investigation of the existing system and it's constraints on implementation, designing of methods to achieve changeover and evaluation of changeover methods.

Modules

1. User Module.
2. client module.
3. CSP Module.
4. Deduplication module.

Modules Description

User Module

In this module, user should register their details and get the secret key for log in and user can download the clients uploaded files the users are able to access the content stored in the cloud, depending on their access rights which are

ISSN:-2581-6934|www.jidps.com

authorizations granted by the client, like the rights re-store the modified data in the cloud.

Client Module

In this module, a client makes use of provider's resources to store, retrieve and share data with multiple users. A client can be either an individual or an enterprise. client can check the uploaded file he can neglate or upload the file. Client can view the deduplicate file based on this client can delete the unwanted data's.

CSP module:

In this module CSP can view all the user details, client uploads details, client's details. And client's activities. regarding the A Secure Client Side Deduplication Scheme in Cloud Storage Environments

Deduplication module:

In this module, the clients uploaded files can be stored in cloud database. it can be very secure clients can view the file from the database based on the deduplicate factor it can be very secure.

CONCLUSION

Managing encrypted data with deduplication is important and significant in practice for achieving a successful cloud storage service, especially for big data storage. The scheme to manage the encrypted big data in cloud with deduplication based on ownership challenge and PRE. This scheme can flexibly support data update and sharing with deduplication even when the data holders are offline. Encrypted data can be securely accessed because only authorized data holders can obtain the symmetric keys used for data decryption. Extensive performance analysis and test showed that our scheme is secure and efficient under the described security model and very suitable for big data deduplication. The cloud based SLA-aware autonomic resource management technique (STAR) has been proposed for execution of heterogeneous workloads by considering generic property of self-management. The main aim of STAR is to reduce SLA violation rate and improve user satisfaction by fulfilling their QoS requirements.

REFERENCES

[1] G. Ateniese, K. Fu, M. Green, and S. Hohenberger, "Improved proxy re-encryption schemes with applications to secure distributedstorage," ACM Trans. Inform. Syst. Secur., vol. 9, no. 1, pp. 1–
30, 2006, doi:10.1145/1127345.1127346.

[2] M. Bellare, S. Keelveedhi, and T. Ristenpart, "DupLESS: Server aided encryption for deduplicated storage," in Proc. 22nd USENIX Conf. Secur., 2013, pp. 179–194.

[3] M. Bellare, S. Keelveedhi, and T. Ristenpart, "Message-locked encryption and secure deduplication," in Proc. Cryptology—EUROCRYPT, 2013, pp. 296–312, doi:10.1007/978-3-642-38348-9_18.

[4] Buyya, Rajkumar, Saurabh Kumar Garg, and Rodrigo N. Calheiros. "SLA-oriented resource provisioning for cloud computing: Challenges, architecture, and solutions." In Cloud and Service Computing (CSC), 2011 International Conference on, pp. 1-10. IEEE, 2011.

[5 Casalicchio, Emiliano, and Luca Silvestri. "Mechanisms for SLA provisioning in cloud-based service providers." Computer Networks 57, no. 3 (2013): 795-810.

[6] J. R. Douceur, A. Adya, W. J. Bolosky, D. Simon, and M. Theimer, "Reclaiming space from duplicate files in a serverless distributed file system," in Proc. IEEE Int. Conf. Distrib. Comput. Syst., 2002, pp. 617–624, doi:10.1109/ICDCS.2002.1022312.

[7] C. Fan, S. Y. Huang, and W. C. Hsu, "Hybrid data deduplication in cloud environment," in Proc. Int. Conf. Inf. Secur. Intell. Control, 2012, pp. 174–177, doi:10.1109/ISIC.2012.6449734.

[8] Garg, Saurabh Kumar, Adel Nadjaran Toosi, Srinivasa K. Gopalaiyengar, and Rajkumar Buyya. "SLA-based virtual machine management for heterogeneous workloads in a cloud datacenter." Journal of Network and Computer Applications 45 (2014): 108-120.

[9] S. Halevi, D. Harnik, B. Pinkas, and A. Shulman-Peleg, "Proofs of ownership in remote storage systems," in Proc. 18th ACM Conf. Comput. Commun. Secur., 2011, pp. 491–50 doi:10.1145/2046707.2046765.

[10] Morshedlou, Hossein, and Mohammad Reza Meybodi. "Decreasing impact of SLA violations: A proactive resource allocation approach for cloud computing

environments." *IEEE Transactions on Cloud Computing* 2, no. 2 (2014): 156-167.

[11] D. Perttula, B. Warner, and Z. Wilcox-O'Hearn, "Attacks on convergentencryption." (2016). [Online]. Available: http://bit.ly/yQxyvl

[12] J. Pettitt, "Hash of plaintext as key?" (2016). [Online]. Available: http://cypherpunks.venona.com/date/1996/02/msg02013.html

[13] Serrano, Damián, Sara Bouchenak, Yousri Kouki, Frederico Alvares de Oliveira Jr, Thomas Ledoux, Jonathan Lejeune, Julien Sopena, Luciana Arantes, and Pierre Sens. "SLA guarantees for cloud services." Future Generation Computer Systems 54 (2016): 233-246.

[14] J. Simao et al, "Partial Utility-driven Scheduling for Flexible SLA and Pricing Arbitration in Clouds", IEEE T. Cloud Computing, 2014

[15] G. Wallace, et al., "Characteristics of backup workloads in production systems," in Proc. USENIX Conf. File Storage Technol., 2012,pp. 1–16.

[16] Z. O. Wilcox, "Convergent encryption reconsidered," 2011. [Online]. Available: http://www.mailarchive.com/cryptography@metzdowd.com/msg08949.html

[17] T. Y. Wu, J. S. Pan, and C. F. Lin, "Improving accessing efficiency of cloud storage using de-duplication and feedback schemes," IEEE Syst. J., vol. 8, no. 1, pp. 208–218, Mar. 2014, doi:10.1109/JSYST.2013.2256715.

[18] J. W. Yuan and S. C. Yu, "Secure and constant cost public cloud storage auditing with deduplication," in Proc. IEEE Int. Conf. Communic. Netw. Secur., 2013, pp. 145–153, doi:10.1109/CNS.2013.6682702.

[19] Y. Z. Zhou, Y. X. Zhang, H. Liu, N. X. Xiong, and A. V. Vasilakos, "A bare-metal and asymmetric partitioning approach to client virtualization," IEEE Trans. Serv. Comput., vol. 7, no. 1, pp. 40–53, Jan.-Mar. 2014, doi:10.1109/TSC.2012.32.

[20] Zhou, Zhou, Zhigang Hu, and Keqin Li. "Virtual Machine Placement Algorithm for Both Energy-Awareness and SLA Violation Reduction in Cloud Data Centers." Scientific Programming 2016 (2016).