International Journal of Computer Science and Engineering Communications,
Volume.5, Issue.4 (2017): Page.1682-1687
www.ijcsec.com

A Survey on Drops: Division and Replication of Data in Cloud for Optimal Performance and Security

Thulasi Mohan[1] and Shilpa Sudheendran[2]
[1]PG Scholar, [2]Assistant Professor
[1,2]Department of Computer Science and Engineering
[1,2]Royal College of Engineering and Technology, Akkikkavu, Kerala, India.
thulasinair28@gmail.com[1], shilpats@royalcet.ac.in[2]

Abstract: *Features of cloud storage system allow the users to work with the data without any trouble of the resources. As the Private data stored in the cloud may be tampered by hackers or unauthorized users, it needs high security for maintaining the data secrecy. This paper reviews Division and Replication of Data in the Cloud for Optimal Performance and Security (DROPS) that collectively approaches the security and performance issues in cloud. DROPS divide a file into fragments, and replicate the fragmented data over the cloud nodes. Only a single fragment of a particular file is stored on each node. This ensures no meaningful information is revealed to the attacker, even he got success with his attack. Moreover the fragments holding nodes are separated by T-coloring technique to minimize the chance of guessing the fragments locations by the attacker. This system enhances the data security and performance of cloud storage.*

Keywords: *Cloud Computing, DROPS, Fragmentation, Replication, Security, Storage, T-coloring*

I. Introduction

Cloud computing, the delivery of computing services like servers, storage, databases, networking, software, analytics and more over the Internet called as the cloud has dominated the business and single users. Cloud computing boasts several attractive benefits like pay per use, Elasticity, Speed, Productivity, Reliability, Performance and security. New worldview of information facilitating benefit additionally brings new security dangers toward client's information stored on cloud nodes.

Well-known cloud providers have experienced temporary lack of availability lasting at least several hours and loss of personal data. Several companies have built search engines, some websites to communicate with other user, E-mail services, and services to purchase items online that use cloud computing infrastructure. Major issue in information technology is the data security. It becomes particularly serious in the cloud computing environment as the data is located in different places even in all the globe.

The two main factors of user's concerns about the cloud technology is the data security and privacy protection. This survey paper aims to analyze the various unresolved security threats in cloud. It also describes the pros and cons of the existing security strategy. Also introduces the existing issues in cloud computing such as data integrity, data segregation, and security and so on. The disadvantages of the existing system are heavily depends on the user′s employed scheme for data confidentiality, the probable amount of loss in case of data tempering as a result of intrusion or access by other VMs cannot be decreased, do not protect the data files against tempering and has loss due to issues arising from virtualization and multi-tenancy.

The new system propose a Division and Replication of Data in the Cloud for Optimal Performance and Security that judicially fragments user files into pieces and replicates them at strategic locations within the cloud. In order to avoid compromise of all of the data in case of successful attack, DROPS methodology does not store the whole file on a single node. It fragments the file and store the fragments on multiple nodes. The proposed scheme fragments and replicates the data file over cloud nodes. The advantages include in case of a successful attack, no meaningful information is revealed to the attacker, the non-cryptographic nature of the proposed scheme makes it faster to perform the required operations on the data and moreover improved the security of data in cloud and decreased the retrieval time. The system is reliable and more flexible as the user can increase the number of nodes, varying the nodes capacity etc. The remainder of the paper is organized as follows. Section I provides the introduction. Section II provides an overview of the related work in the field. Section III concludes the paper.

II. Related Works

Wayne A. Jansen[1] proposed the paper where the issues of cloud are organized into several general categories: trust, architecture, identity management, software isolation, data protection, and availability. Incidents may involve various types of fraud, sabotage of information resources, and theft of information by current or former employees, contractors, and other parties that have received access to an organization's networks, systems, and data to carry out or facilitate operations. It is not feasible to verify the correct functioning of a subsystem and the effectiveness of security controls as extensively as with an organizational system. Both the client and server side protection is inevitable in cloud storage. It is sufficient to extend the organization's identification and authentication framework as data sensitivity, and privacy of information have increasingly become a concern for organizations, and unauthorized access to information resources in the cloud is a major issue. Data isolation and data location are important in the data protection part in the cloud. The data stored on the cloud nodes should be available for the authorized users with any cost at any time. The migration to a cloud computing environment is in many ways an exercise in risk management and that risks must be carefully balanced against the available safeguards, with the understanding that accountability for security remains with the organization.

G. Kappes, A.Hatzieleftheriou, and S.V.Anastasiadis [2] proposed the paper describing in a virtualization environment, that serves multiple tenants. As it enables data sharing, administration efficiency and performance optimization, storage consolidation at the file system is desirable. Due to intermediate translation layers required for purposes of networked file access or identity management, the scalable deployment of file systems in such environments is challenging. Analyzes the security requirements in multitenant file systems they introduced the Dike authorization architecture that combines native access control with tenant namespace isolation that is backwards compatible to object-based file systems. The advantages of the model are Isolation: Each tenant is free to choose identities for its users. So to prevent collisions, the identity space and access control of different tenants are isolated. Sharing is used to provide flexible access to enable secure file sharing within a tenant or among various tenants. To required achieve the required performance and scalability for enormous numbers of users or files efficient natively support of multi access control is used. Backwards compatibility brings out the architectural characteristics of successful file systems to ensure backward compatibility with existing applications. Manageability provides the maintenance support at the file level that allows the cloud provider to uniformly and flexibly manage the storage resources of different tenants.

Yang Tang, Patrick P.C. Lee, John C.S.Lui, Radia Perlman[3] proposed the paper to reduce data management costs, this paper describes outsource data backups off-site to third-party cloud storage services. We must provide security guarantees for the data outsourced, which is maintained by third parties. FADE is a secure overlay cloud storage system that achieves fine-grained, policy-based access control and file assured deletion. FADE associates the outsourced files with file access policies and it deletes files to make them unrecoverable to anyone upon revocations of file access policies. It is built on a upon a set of cryptographic key operations that are self-maintained by a quorum of key managers. These are independent of third-party clouds to achieve security goals. It can act as an overlay system that works seamlessly upon cloud storage services.

Thanasis Loukopoulos, Ishfaq Ahmad[4] proposed the paper including in large distributed systems such as internet, fast dissemination and access of information has become a norm. To decrease the network traffic, replicating some of the objects at multiple sites is one possible solution. By solving a optimization problem which is NP-complete, it is possible to take decisions of what to replicate where. It is possible to propose a GA to solve the problem when the read/write demands remain static. And to prove the superior solution quality obtained compared to an intuitive greedy method. But the static GA approach involves high running time and is somewhat not useful when read/write demands continuously change. A hybrid GA can tackle such case that takes current replica distribution as its input and computes a new one using knowledge about the network attributes and the changes occurred.

Yves Deswarte, Laurent Blain, Jean-Charles Fahe[5] proposed the paper including intrusion means a large class of attacks, covering not only computer break-ins by external attackers, but also illegitimate use by registered users. The users who have not registered has to pass the authentication and authorization mechanisms. Internal intruders are those who are already registered who tries to read or modify data only possible after the authorization mechanisms. The authorization server has to manage two kinds of access controls. that includes the access to application servers, such as a data processing servers, which manage no specific persistent objects. In the tree structure, leaves directly connected to the root are the descriptors of such servers. : there is no obvious reason to structure these servers as a hierarchy. And the next kind concerns accesses to persistent objects maintained by application servers, such as the Persistent File Server. A method of cutting each file into pages of fixed size is used here. To reach a size equal to a multiple of a page size, the files are padded. In order to improve the speed of access of information, the fragments are of same length. A user does not need to reassemble a whole file if he only needs a single page. Fragmentation is performed at the user side and it has got several advantages. The theft of individual storage media of no avail to the intruder because of the geographical scattering of fragments, even he possesses the cipher key. The ciphers employed can be much simpler to make faster than the conventional ones.

Abdul Nasir Khan, M.L. Mat Kiah, Sajjad A. Madan, Atta ur Rehman Khan, Mazhar Ali[6] proposed the paper to protect mobile users identity with dynamic credentials, a light weight security scheme is proposed. The proposed scheme to keep minimum processing burden the mobile device, it offloads the frequently occurring dynamic credential generation operations on a trusted entity. The credential information is updated frequently on the basis of mobile- cloud packets exchange to enhance security. Trusted entity in the system is mobile device. The cloud servers are assumed to be fully distrusted entities and the mobile users can take precautionary measures to avoid the malicious activities on a mobile device. Manager, the fully trusted party is under the control of client organization. He is capable of handling the entire organizations requests. The manager sends the encrypted Nonce MD and Nonce CSP

to the mobile user and cloud service provider to verify the authenticity of both mobile user and cloud service providers.

Samee Ullah Khan, Ishfaq Ahmad[7] proposed the paper to solve the fine-grained data replication problem over the Internet, compares and analyzes 10 heuristics . The frequently accessed data objects are replicated onto a set of selected sites in fine grained replication so as to minimize the average access time perceived by the end users. The model captures the minimization of the total object transfer cost in the system by using a unified cost model. This leads to effective utilization of storage space, replica consistency, fault-tolerance, and load-balancing. Heuristics includes six A-Star based algorithms, two bin packing algorithms, one greedy and one genetic algorithm. DRPA starts from the root or the start node and the intermediate tree nodes represent the partial solutions, and leaf nodes represent the complete solutions or goals. DRPA-Star always identifies a solution, if there exists one. It always chooses the best solution, if two or more solutions exist. The main purpose of suboptimal assignments is to design algorithms that converge to solution faster. It also overcome the high memory requirements associated with A-Star type algorithms. The users can determine the exact time to invoke an algorithm only after the studying of past user access trends.

Alessandro Mei, Luigi V. Mancini, Sushil Jajodia[8] proposed the paper that describes file allocation by a distributed algorithm for guaranteeing high assurance, availability, scalability in large distributed file system. The files are allocated over multiple servers by using replication and fragmentation schemes. Even in the presence of a successful attack that compromises a subset of the file servers, the file confidentiality and integrity are preserved. As the read-write patterns and the location of the clients in the network change, the file allocation changes. This makes the algorithm adaptive. It provides a higher frequency of writes that guarantees higher confidentiality. Estimating where the fragment is read and written and thereby keeping it close makes minimal cost and increases confidentiality. Propose a dynamic allocation algorithm which, assuming stable read-write frequency patterns and moves fragments between servers in such a way as to converge to a mapping with maximal dynamic assurance. In static assurance, fragments can be encrypted before being stored t improve the security of the system. the encryption must be using a secret key that are stored in the server. In dynamic assurance the fragments can be encrypted with the client's private key before being stored. And moreover if the file fragments are not kept encrypted in the server, the access to a shared file can be simplified. It is important that it must be encrypted before they are send.

Jaydip Sen[9] proposed the paper which discuss regulatory, security and privacy issues in cloud computing. The various threats against information assets that are residing on cloud computing environments. The various types of attackers and their capability of attacking the data stored on cloud. The variety of security risks associated with the cloud, and where relevant considerations of attacks and countermeasures. Examines emerging cloud security risks. Several types of security risks are Integrity, Confidentiality and Availability. Random attacker randomly scan the Internet trying to find vulnerable components. Weak attackers attempt to customize their attacks using available exploit tools. Strong attackers are organized, well-financed and skilled groups of attackers, targeting particular applications and users of the cloud. Organization for the Advancement of Structured Information Standards not-for-profit, international consortium that drives the development, convergence, and adoption of e-business standards with standards for security, e-business, and standardization efforts in the public sector and application-specific markets.

Dr G.Kesavaraj, K.Anitha, R.Divya[10] proposed the paper that evaluate cloud security by identifying unique security requirements and present a viable solution that eliminates these

potential threats. Trust refers to the quality of a person or entity who is accessing the data stored on cloud data centers. Protecting data from unauthorized deletion, modification or fabrication refers to the data integrity. Property of a system being accessible and usable upon demand by an authorized entity is systems availability. It also includes the ability to carry on operations even when some authorities misbehave. The system should also be capable to continue operations even in the possibility of a security breach.

Aiqiang Gao, Luhong Diao[11] proposed the paper where the transaction commits after updating one replica of data with lazy replication. The updates are propagated towards the other replicas after the transaction commits. And these replicas are updated in separate transactions. The replica management component handles the join and quit of a replica site and maintains a whole graph of all replica sites. In immediate-immediate algorithm, when data is submitted to its primary server, the server finds registered subscriber for that data and multicasts to various customers. The method improves throughput of data server. Also reduces the response time. Matching for special cases where database transactions are short, need high throughput.

III. Conclusion

Cloud computing faces security issues like Loss or theft of intellectual property, Loss of control over end user actions, Malware infections that unleash a targeted attack, Contractual breaches with customers or business partners. High security measures are required to protect data within the cloud. DROPS ensure an increased security level by data replication. The data was fragmented and stored on multiple nodes separated by T-coloring. DROPS technology gives high efficiency than full-scale replication techniques. The performance and data retrieval time efficiency got upgraded. The performance can be further enhanced by varying the number of nodes and capacity of nodes.

References

[1] Wayne A. Jansen, *"Cloud Hooks: Security and Privacy Issues in Cloud Computing" Proceedings of the 44th Hawaii International Conference on System Sciences - 2011.*

[2] G. Kappes, A.Hatzieleftheriou, S.V.Anastasiadis, "*Virtualization-aware Access Control for Multitenant Filesystems" 978-1-4799-5671-5/14/$31.00 © 2014 IEEE.*

[3] Yang Tang, Patrick P.C. Lee, John C.S.Lui, Radia Perlman, *"Secure Overlay Cloud Storage with Access Control and Assured Deletion" IEEE transactions on dependable and secure computing, vol.9, No.6, November/December 2012.*

[4] Thanasis Loukopoulos, Ishfaq Ahmad, *"Static and adaptive data replication algorithms for fast information access in large distributed systems" IEEE International conference on Disributed Computing Systems.*

[5] Yves Deswarte Laurent Blain Jean-Charles Fahe, "*Intrusion Tolerance in Distributed Computing Systems" CH2986-8/91/0000/0110$01 .OO@ 1991 IEEE.*

[6**]** Abdul Nasir Khan, M.L. Mat Kiah, Sajjad A. Madan, Atta ur Rehman Khan, Mazhar Ali, " *A Cloud-Manager-Based Re-Encryption Scheme for Mobile Users in Cloud Environment: a Hybrid Approach" Jouranl of Grid Computing, Volume 13 Issue4, December 2015, Pages 651-675.*

[7] Samee Ullah Khan, Ishfaq Ahmad, "*Comparison and analysis of ten static heuristics- based Internet data replication techniques", J. Parallel Distrib. Comput. 68 (2008) 113 – 136.*

[8] Alessandro Mei, Luigi V. Mancini, Sushil Jajodia, "*Secure dynamic fragment and replica allocation in large-scale distributed file systems" IEEE Transactions on Parallel and Distributed Systems,* 14*(9), 885-896.*

[9] Jaydip Sen, "*Security and privacy issues in cloud computing", Innovation Labs, Tata Consultancy Services Ltd., Kolkata, India.*

[10] Dr G.Kesavaraj, K.Anitha, R.Divya, *"Addressing cloud computing security issues", IJIRCCE vol.4 issue 6 June 2016.*

[11] Aiqiang Gao, Luhong Diao, *"Lazy Update Propagation for Data Replication in Cloud Computing" 978-1-4244-9142-1/10/$26.00©2010 IEEE.*