International Journal For Technological Research In Engineering
Volume 3, Issue 8, April-2016 *ISSN (Online): 2347 - 4718*

SURVEY ON DATA DIVISION AND REPLICATION FOR SECURITY AND PERFORMANCE IMPROVEMENT IN CLOUD

Prasanna Goshika[1], Manjunath P.C[2]
[1]M.Tech, [2]Assistant Professor, CSE, Reva Institute of Technology and Management, Bangalore, Karnataka

Abstract: In Cloud Computing, Security is a major concern for distributing data to third-party data control. In the cloud the assailant will steal the information shared by the others. So need to provide a security for data in the cloud. However, data retrieval time also take into account with respect to security. Security and performance will be given to the customer's data by performing the fragment replication. In this method, a file can be divided in to fragments and the fragmented data will be replicated over cloud nodes. A single fragmented data will store in each of nodes, so the attacker won't get proper information regarding a file. And one more thing fragmented data in nodes are separated by definite distance. Even the attacker can't guess the location of the fragments in the nodes, because graph T-coloring method prevents it. No information is revealed to the assailant after applying the data replication. For improving security, each fragment is replicated only once. The results obtained as, data will be in high secure and with little performance.

Keywords: Centrality, Fragmentation, Replication, Security, Performance

I. INTRODUCTION

Resources will be delivered to the customers on demand basis over internet. Performance and security are more important and its gives beneficial to the large scale systems. There are two risks in cloud services for organizations and individual, those are Data loss or leakage. The user, service and infrastructure are the three actors involved in the cloud computing environment. For cloud users, security is the top most concern and where the meaningful data/ information is stored and processed. Unauthorized access to private information and data theft are the major problems of users. The data moving in virtualized and shared environment from cloud storage having a security problem. Resources can be shared among multiple users in cloud. Customer's personal information is leaked due to improper media sanitization. The data must be secured in public cloud. Data retrieval time, availability of data and responding time are the problems in large scale systems and these are dealt with the data replication. Customers are attracted by the services which are provided by providers but at the same time there is no guarantee for the private information of the customer. So here we are introducing the fragment replication for gaining the performance and security. Each file is fragmented as small amount of pieces and those are replicated in proper fashion. The replicated data will store in different locations. So that the fragmented data does not having proper information. Different fragments stored in cloud node will increase the data security. So the attacker doesn't have any idea about the location of fragments in cloud. For improving the security we don't select neighboring nodes for storing the file fragments. Here we are using the T-coloring graph method for node separation. Based on centrality measures nodes are selected for improving the retrieval time. In two ways we can select the nodes for storing the file fragments, (a) for starting position of fragments, nodes are selected based on centrality measures. (b) Nodes will be choose for replication. For improving the performance, here we are using replication techniques.

II. LITERATURE SURVEY

To ensuring high reliability in data centers, distributing the power and resource cooling, storage capacity are purveyed. The operational cost will be reduced by using the replicated data in cloud node. In cloud, resources will be leads to a congestion in service purveying. This will saves the energy consumption. Therefore data replication, which fetches data near to data customer seen as auspicious resolution. Bandwidth and network detain are diminished by the replication. Here replication algorithm is using for gaining the performance as well as energy efficiency. The performance can be measured using availability, response time and datacenter congestion and Failures and loss occurred in data is analyzed by the replication cost model. Replicated data will be supported by the distributed system. Internet plays a vital role in every one's life for accessing and rapid dispersal of data. One possible solution is replicating few of objects at various places is for decreasing network traffic. Then we need to decide where to replicate and what kind of data to replicate, Allocation, consistency and fault tolerance are the three main issues of replication. In this paper, Genetic Replication algorithm (GRA) [2] solves the problems of read/write demands and quality is obtained by comparing it with greedy method (Simple Replication Algorithm). Regrettably, when read/write demands are continuously changing, the static genetic algorithm (GA) is not useful because it involves high running time. To overcome that problem, the author presents an adaptive genetic replication algorithm (AGRA) and it takes as input for current replica distribution and using knowledge we can determines a new one, when changes are happened. More number of requests can be handled in popular web servers for providing good quality services to customers. So that by using replication techniques, it ensure contents of from up-to-date, information can be retrieved in fast, load can be reduced in web-server and it adds reliability of data [3]. Very often accessed data objects are replicated. Data objects are replicated for reducing the average access time in various locations and these are recognized by the customers. For diminishing the transfer of object cost in total here we are using cost model for effective storage space utilization, fault-tolerance and replica-consistency. For data replication problem, A-star, greedy and bin pack and genetic algorithm were used in this. To increasing the performance of system need to determine in which places replicas will be stored and what amount of replicas. Fine grained replication algorithm was used for this purpose. To analyzing the communication cost, evaluate the heuristics, due to transfer of objects under variation of capacity of servers, size of object, permissions of write and read. This paper having the various algorithms that will give assurance of fast or optimal or both types of results. Backup services are provided to customers on pay basis. Public, private and hybrid cloud are the three main cloud storage models. This paper presents [4] a reduce management cost for outsourcing data to third party cloud storage. Although for the outsourcing data need to provide security guarantees. For achieving assured deletion of file and controlling the access based on policy, we need to design and implement the FADE (File assured deletion). In FADE, files are deleted on time basis and can’t recover by others. From cloud, customer can upload and download a file using FADE. Preventing unauthorized access and protecting data which was outsourced, it was developed. FADE acts as laminate system and works with the today’s top most cloud storage services i.e. Amazon S3. For outsourcing data, FADE provides security protection. In Internet, a platform which provides various cost effective services to customers as well as business i.e. cloud. At the same time it gives few problems regarding security, due to this its decreases the usage. The SAP model security issues were described and also the security problems of virtualization, capacity of storage and networking are discussed in this. For this vulnerabilities’ and threats relationship was done. Mobile customers are utilizing the cloud computational and storage services for improving the resource control of mobile devices. And also improves the storage capacity and processing of mobile devices, security and privacy issues are raised by relocation of privileged information on untrusted cloud. In cloud, mobile customer’s data is hidden and which is not shown in to understandable format like dots are using for protecting private data. Once the customer private data is hacked by the assailant then attacker misuses the mobile customer data later on. The mobile customer is unaware of attacker malicious activities. In cloud, to protecting mobile customer’s data light weight security scheme [6] was introduced in this. Depending on exchange of packets in mobile cloud, private information is updated regularly for enhancing the security and reliability. Modern exposure of cloud computing has drastically make changes to everyone’s appreciation of infrastructure architectures development model and software delivery. Security is given by the trusted third party to customers for avoiding the problems faced by it. Security is evaluated by recognizing the definite security measures and avoids the threats are solved in this. To securing the information in cloud in trust basis, a Trusted Third party [7] provides various answers to customers. Safe communication is done in two customers with this of third party. From security aspect a number of threats and problems are introduced from this relocation to the clouds. Security will be constructed on confidence in cloud environment, diminishing protection to a trusted third party. To certify the entities certification scheme was used and it also certifies network devices used in cloud. In private and public organization, cloud infrastructure are continuously used for wide range of computational needs. This paper describe a virtualization environment work for many consumers, at the file system level storage consolidation is precedent because it allows sharing of data, performance optimization. Analyzes the security needs in multiple customers file systems. The intermediate translation layers are needed for networked file access or identity management. This paper introduces a dike authorization architecture [8], which merges native access control with customers name space isolation. Dike authorization architecture demonstrates bounded performance overhead for up to hundred customers. Presented a technique [9] that protecting secure cloud data from integrity and freshness availability of data and customers visibility is offered by the auditing frame work in the cloud. In cloud migration of data is performed by the Irish file system. For file system operations, Irish offers integrity and freshness guarantees for file systems data. In iris file system, no changes occurred in file system operations. The aim is to achieve reducible operational latency, the gateway application is designed and caches file system data and meta data is accessed by tenant’s and it computes integrity and MAC codes on data blocks. For cached data it also maintains the integrity and freshness information. For data confidentiality, the proposed technique is depending on the tenant’s. Authentication scheme was used in Iris file system, Merkle-tree, MAC codes and counters are used for freshness of data. Concurrent and previous file operations are supported by the Merkle-tree. In Auditing framework data retrieval is verified by the Proof of Retrievability scheme in cloud for performing the operations. It helps in identifying the loss of data in cloud. Drive failure resilience problem was solved by the Remote Assessment of Fault Tolerance. Due to data loss providers are not interacting to tenants. High-Availability and Integrity Layer is introduced in this to overcome that problem. In cloud there are so many challenges with respect to performance and security. Security problems are not resolved when foregrounding the cost and performance advantage of cloud. For providing security to complex computer system in large scale computing is also a security problem. Data must be protected in the organization with some policies when the organization will be in cloud. Analyzing the qualitative and quantitative factors to avoid risks. The cloud hook provides a useful data in cloud with outsourced services. The key security issues are categorized as identity management, trust, data protection and availability. Amazon and Google are the providers, those are having the infrastructure for saving the customer’s from assailant.

III. CONCLUSION

Customers are attracted by the services which are provided by providers but at the same time there is no guarantee for the private information of the customer. So here we are introducing the fragment replication for gaining the performance and security. Each file is fragmented as small amount of pieces and those are replicated in proper fashion in cloud in the various nodes. Each node having the single replicated data and fragments can be replicated only once in the node. So the meaningful information is not revealed to the adversary. One more thing nodes are placed in different locations using T-coloring graph. A user can upload and download a file and update the contents of file. An automatic update mechanism needs to develop, that identifies and updates whatever fragments needed. So the future work will be saves the time and resources used in uploading a file, updating the contents of file and download the file again.

REFERENCES

[1] D. Boru, D. Kliazovich, F. Granelli, P. Bouvry, and A. Y. Zomaya, “Energy-efficient data replication in cloud computing datacenters”, In IEEE Globecom Workshops, 2013, pp. 446-451.

[2] T. Loukopoulos and I. Ahmad, “Static and adaptive distributed data replication using genetic algorithms”, Journal of Parallel and Distributed Computing, Vol. 64, No. 11, 2004, pp. 1270-1285.

[3] S. U. Khan, and I. Ahmad, “Comparison and analysis of ten static heuristics-based Internet data replication techniques”, Journal of Parallel and Distributed Computing, Vol. 68, No. 2, 2008, pp.113-136.

[4] Y. Tang, P. P. Lee, J. C. S. Lui, and R. Perlman, “Secure overlay cloud storage with access control and assured deletion”, IEEE Transactions on Dependable and Secure Computing, Vol. 9, No. 6, Nov. 2012, pp. 903-916.

[5] K. Hashizume, D. G. Rosado, E. Fernndez-Medina, and E. B. Fernandez, “An analysis of security issues for cloud computing”, Journal of Internet Services and Applications, Vol. 4, No. 1, 2013, pp. 1-13.

[6] A. N. Khan, M.L. M. Kiah, S. A. Madani, and M. Ali, “Enhanced dynamic credential generation scheme for protection of user identity in mobile-cloud computing”, The Journal of Supercomputing, Vol. 66, No. 3, 2013, pp. 1687-1706.

[7] D. Zissis and D. Lekkas, “Addressing cloud computing security Issues”, Future Generation Computer Systems, Vol. 28, No. 3, 2012, pp. 583-592.

[8] G. Kappes, A. Hatzieleftheriou, and S. V. Anastasiadis, “Dike: Virtualization-aware Access Control for Multitenant Filesystems”, University of Ioannina, Greece, Technical Report No. DCS2013-1, 2013.

[9] A.Juels and A.Opera, “New approaches to security and availability for cloud data”, Communications of the ACM, Vol. 56, No. 2, 2013, pp. 64-73.

[10] W. A. Jansen, “Cloud hooks: Security and privacy issues in cloud computing”, In 44th Hawaii IEEE International Conference on System Sciences (HICSS), 2011, pp. 1-10.