

Redistribution of Load in Cloud Using Improved Distributed Load Balancing Algorithm with Security

R.Rajavadivu, A.Bazila Banu

Post Graduate, Dept of Information Technology, Velammal college of Engineering and Technology, Madurai,
Tamil Nadu, India

Assistant Professor, Dept of Information Technology, Velammal college of Engineering and Technology,
Madurai, Tamil Nadu, India

Abstract—In cloud computing security plays an important role. The nodes in the cloud simultaneously serve computing and storage functions. To provide security the file is partitioned into a number of chunks and are allocated in distinct. Before partitioning, the data is encrypted to provide more security. The encrypted data is then portioned and are stored in various nodes. The files are splitted dynamically on the basis of number of systems present in the cloud. In a cloud computing environment, nodes may be upgraded, replaced, and added in the system. Files in the node can be dynamically created and deleted. This results in load imbalance in a distributed file system; that is, the file chunks are not distributed as uniformly as possible among the nodes. In order to balance the load in the cloud load balancing algorithm is used. Load balancing is done by moving the chunks of the file from one system to the neighbouring system. If a node is deleted the files in the particular node is moved and stored in a balanced manner among the neighbouring nodes. VM ware offers an cloud environment that implement transferring data's like as cloud.

Keywords— Load balance, cloud, security, storage, VMware

I. INTRODUCTION

Cloud computing is evolving and considered next generation architecture for computing. Cloud computing

plays a prominent role in this modern era due its compelling benefits and services. Cloud Storage system plays its part too in making cloud usage as Storage as Service. Cloud storage system is a collection of cloud servers which provide continuous access. Data storage is the one which grabs a tremendous part in technology of world. Typically cloud computing is a combination of computing recourses accessible via internet. Historically the client or organizations store data in data centers with firewall and other security techniques used to protect data against intrudes to access the data. In a cloud computing environment, the service content offered by service providers can be adjusted according to the needs of the user. One technique could be encrypting the data on client side before storing it in cloud storage.

Load balancing in cloud computing systems is really a challenge now. Always a distributed solution is required. Because it is not always practically feasible or cost efficient to maintain one or more idle services just as to fulfil the required demands. Jobs can't be assigned to appropriate servers and clients individually for efficient load balancing as cloud is a very complex structure and components are present throughout a wide spread area. Load balancing is a process of reassigning the total load to the individual nodes of the collective system to make resource utilization effective and to improve the response time of the job, simultaneously removing a condition in which some of the nodes are over loaded while some others are under loaded. A load balancing algorithm which is dynamic in nature does not consider the previous state or behaviour of the system, that is, it depends on the present behaviour of the system. The important things to consider while developing such

algorithm are : estimation of load, comparison of load, stability of different system, performance of system, interaction between the nodes, nature of work to be transferred, selecting of nodes and many other ones. This load considered can be in terms of CPU load, amount of memory used, delay or Network load. In order to distribute the necessary tasks, load balancers go through a series of steps. First, the load balancer will query the available servers to ensure their availability. The load balancer pings a server, and if the expected response occurs, it will be included in the available list. If the server fails to respond, it will not be used until another test is performed and it returns with the appropriate response. Load balancing software is very flexible in this environment, as the administrator can quickly tweak the system to ensure it is checking servers appropriately and accurately. Distributing the load between the active servers can be done in several different ways. The load balancer may use a round-robin method, where each server is used in turn. It can also use a weighted round robin system, where servers are assigned traffic based on their configured capabilities.

II. MOTIVATION

Data security in cloud is one of the main issue in cloud computing. The data's in the cloud has no security. In order to provide security, the data's are split into many chunks and stored over various nodes in the network. While storing a single file in various nodes by splitting it into many chunks a load imbalance problem occur. The load can be a memory, CPU capacity, network or delay load. It is always required to share work load among the various nodes of the distributed system to improve the resource utilization and for better performance of the system. Load balancing aims to optimize resource use, maximize throughput, minimize response time, and avoid overload of any one of the resources. Using multiple components with load balancing instead of a single component may increase reliability through redundancy. Load balancing provides a means to efficiently and consistently distribute the processing load for clients across multiple servers rather than having each server respond to every client. This can help to avoid the situation where nodes are either heavily loaded or under loaded in the network. Load balancing is the process of ensuring the evenly distribution of work load on the pool of system node or processor so that without disturbing, the running task is completed. All types of files can be made into chunks. The number of chunks increases as the size of the file increases. Number of chunks is not an issue here. Any number of chunks can be stored in the cloud among the nodes present in a balanced manner. The goals of load balancing are:

- To improve the performance substantially
- To have a backup plan in case the system fails even partially
- To maintain the system stability
- To accommodate future modification in the system

III. LITERATURE SURVEY

[1] proposed a binary tree structure that is used to partition the simulation region into sub-domains. The characteristics of this fast adaptive balancing method are to be adjusted the workload between the processors from local areas to global areas. According to the difference of Workload, the arrangements of the cells are obtained. But the main workload [2] proposed an algorithm named honeybee behaviour inspired load balancing algorithm. Here in this session well load balance across the virtual machines for maximizing the throughput. al.[3] proposed a dynamic file migration load balancing algorithm based on distributed architecture. Considered the large file system there were various problems like dynamic file migration, algorithm based only on centralized system etc. So these problems are to be avoided by the introduction of the algorithm called self acting load balancing algorithm (SALB). [4] Proposed an efficient cell selection scheme and two heat diffusion based algorithm called global and local diffusion. Considered the distributed virtual environments there were various number of users and the load accessing by the concurrent users can cause problem. This can be avoided by this algorithm. [5] addressed the concept of overlay networks for the interconnection of machines that makes the backbone of an online environment. the proposed network that makes better feasibility and load balancing to the dynamic virtual environments.

IV. PROBLEM DEFINITION AND LIMITATIONS

The problem here is to providing security to the data stored in cloud. The main method used for ensuring data security in the cloud is by encryption. Encryption seems like the perfect solution for ensuring data security; however, it has some drawbacks. Encryption takes considerably more computational power, and this is multiplied by several factors in the case of databases. There are several approaches developed to handle data encryption, each having its own compromises and downsides, some provide better security mechanisms, and some focus on facilitating more operations to the customers. Some methods have been developed that serve as alternatives to encryption. These methods are generally faster than encryption. The encrypted data's are made into chunks and are stored in cloud. Data Splitting is another way to provide data security. The idea is to split the data over multiple hosts that cannot communicate with each other; only the owner who can access both hosts can collect and combine the separate datasets to recreate the original.

Load balancing is the issue in cloud computing. The load in the network is balanced among the nodes in the network. The file that is to be stored in the cloud is encrypted for security purpose. The encrypted file is made into chunks and is stored evenly in various nodes. When a new file is uploaded or deleted in a node load imbalance problem occur. The load imbalance problem

may arise even due to the failure of the node. We can overcome this problem by load rebalancing algorithm. When an imbalance in the load occur the load rebalancing algorithm balances the load by moving the load from the failure node to the neighbour nodes. The server node which handles the load balancing among the nodes in the network has the back up of all the files that are stored in the nodes. If any of the files are deleted in a particular node due to the failure of that node, then the main server which has the back up of the deleted files splits it and stores it to the neighbour nodes in the balanced manner. The load rebalancing algorithm used here considerably outperforms the prior distributed algorithm in terms of load imbalance factor, movement cost, and algorithmic overhead.

V.ARCHITECTURE

In this architecture a centralized load balancer is used to split the file into chunks in order to store the data in various nodes. The data to be stored in the cloud is encrypted before storage for more security. The encryption is done by the key generated at the client side. Then the encrypted data is made into chunks and stored in various nodes. When the server control performs operations on data like deletion or updation load imbalance problem occurs. This problem can be solved by the rebalancing algorithm which balances the load in the cloud after the above operations performed. As the centralized node has the backup of the deleted files it can store the deleted files in the other nodes even in the case of node failure.

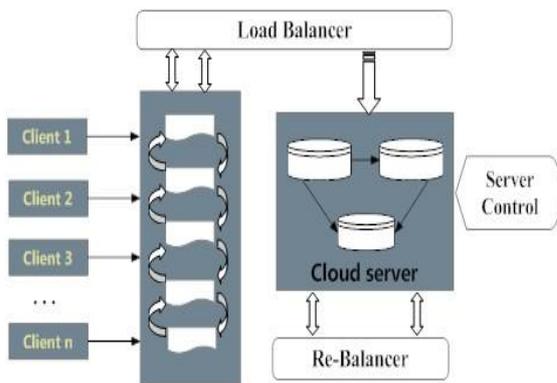


Fig.1 Architecture

VI.METHODOLOGIES

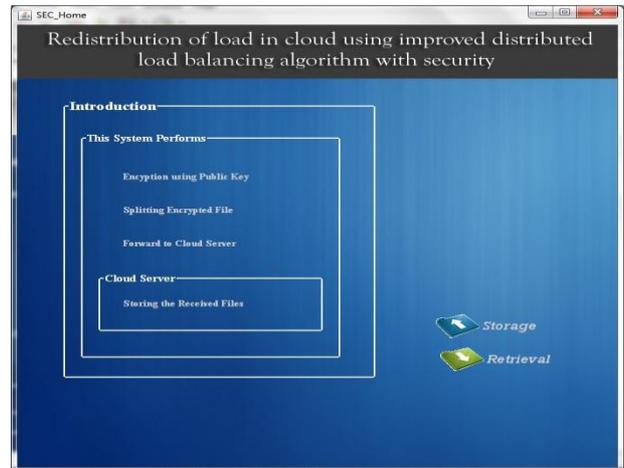


Fig.2 Home page

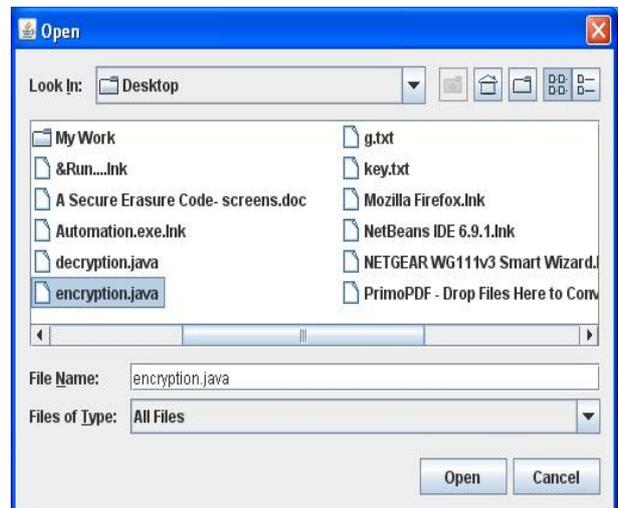


Fig.3 File selection

A.Encryption of data

The data that is to be stored in cloud is not secure. In order to provide security to data, the data is stored in the encrypted form in the nodes. The file that is to be uploaded in the cloud is selected by the client. The encryption process is performed over the data with the key provided. The encrypted file is made in chunks and stored in various nodes.

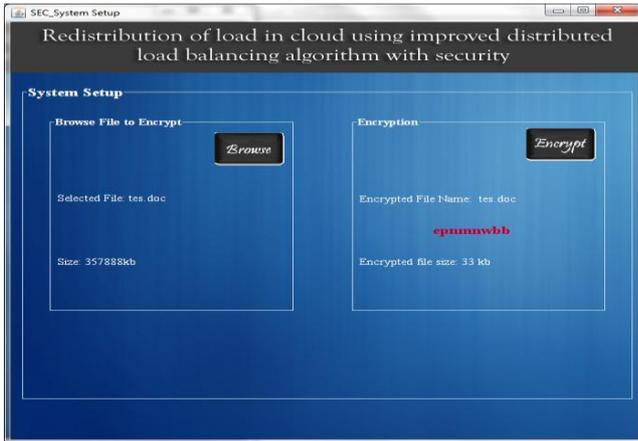


Fig.3 data encryption

B. Splitting the data

The encrypted file is partitioned into a number of chunks as shown in figure 3 is allocated in distinct nodes. The load of a node is typically proportional to the number of file chunks the node possesses. Because the files in a cloud can be arbitrarily created, deleted, and appended, and nodes can be upgraded, replaced and added in the file system. The chunks of files are allocated uniformly among the nodes such that no node manages an excessive number of chunks.

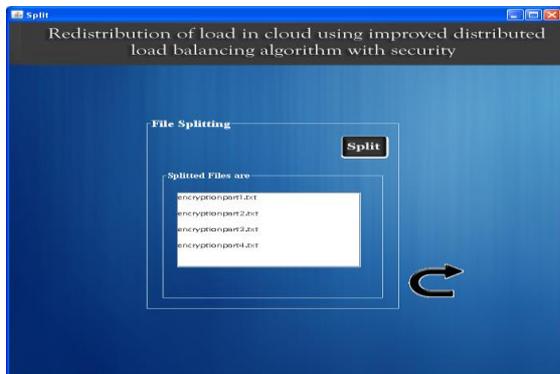


Fig.4 data splitting

C. sending data to cloud

As cloud is a centralized storage the data's have no security. So the encrypted file is made into chunks to provide more security to store the data in the cloud. The splitted files are stored in the cloud and can be accessed from anywhere whenever needed. Thus storing a single file in various nodes has more security when compared to the file that is stored in a single node.

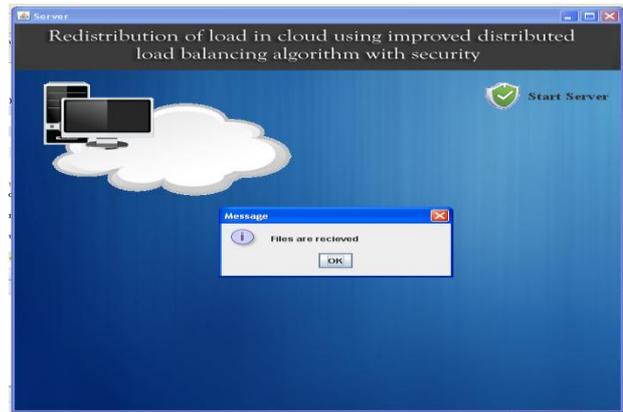


Fig.5 sending data to cloud

VII CONCLUSION

The data are stored in the cloud in the secure manner. The security for the data is provided by the encryption and data splitting techniques. The file that is to be uploaded in the cloud is encrypted by the key generated at the client side. The encrypted data is splitted into chunks to provide more security to the data. The chunks of the files are stored in various nodes present in the network.

VIII REMAINING WORK

In future rebalancing the data's among the nodes is done. The data's are splitted and stored in the balanced manner in the cloud environment among the nodes present in it. When a node is deleted or added in the cloud environment imbalance problem occur which can be solved by using the load rebalancing algorithm which reduces the loss of data

ACKNOWLEDGEMENT

I would like to thank our head of the department Mr.S.Raj Pandian and my guide Mrs.A.Bazila Banu for motivating me in doing this projects and thanks to all my friends and the web sources.

REFERENCES

[1] Dongliang Zhang, Changjun Jiang,Shu Li, "A fast adaptive load balancing method for parallel particle-based simulations", Simulation Modelling Practice and Theory 17 (2009) 1032–1042.
 [2] Dhinesh Babu L.D, P. VenkataKrishna, "Honey bee behaviour inspired load balancing of tasks in cloud computing environments", Applied Soft Computing 13 (2013) 2292–2303.
 [3] Bin Dong, Xiuqiao Li, Qimeng Wu, Limin Xiao, Li Ruan, "A dynamic and adaptive load balancing strategy for parallel file system with large-scale I/O servers", J. Parallel Distribution Computing. 72 (2012) 1254–1268.
 [4] Yunhua Deng, Rynson W.H. Lau, "Heat diffusion based dynamic load balancing for distributed virtual environments", in: Proceedings of the 17th ACM Symposium on Virtual Reality Software and Technology, ACM, 2010, pp. 203–210.
 [5] Markus Esch, Eric Tobias, "Decentralized scale-free network construction and load balancing in Massive Multiuser Virtual

- Environments”, in: Collaborative Computing: Networking, Applications and Worksharing, Collaborate Com, 2010, 6th International Conference on, IEEE, 2010, pp. 1–10.
- [6]Qian Wang ,Cong Wang ,Kui Ren ,Wenjing Lou ,Jin Li, "Enabling Public Auditability and Data Dynamics for Storage Security in Cloud Computing", IEEE Transactions on Parallel and Distributed Systems, vol. 22,no. 5, may 2011.
- [7]Jing-Jang Hwang, Hung-Kai Chuang,Yi-Chang Hsu, Chien-Hsing Wu, "A Business Model for Cloud Computing Based on a Separate Encryption and Decryption Service," Proceedings of the 2011 International Conference on Information Science and Application, April 2011.
- [8]Haozheng Ren, Yihua Lan, Chao Yin School of Computer Engineering Huaihai Institute of Technology,Lianyungang, China "The Load Balancing Algorithm in Cloud Computing Environment" . 2012 2nd International Conference on Computer Science and Network Technology.
- [9]D. Karger and M. Ruhl, "Simple Efficient Load Balancing Algorithms for Peer-to-Peer Systems," in Proc. 16th ACM Symp. Parallel Algorithms and Architectures (SPAA'04), June 2004, pp. 36–43
- [10]Gaochao Xu, Junjie Pang, and Xiaodong Fu" A Load Balancing Model Based on Cloud Partitioning for the Public Cloud" TSINGHUA SCIENCE AND TECHNOLOGY ISSN 1007-0214 04/12 pp34-39 Volume 18, Number 1, February 2013
- [11] D. Karger and M. Ruhl, "Simple Efficient Load Balancing Algorithms for Peer-to-Peer Systems," in Proc. 16th ACM Symp. Parallel Algorithms and Architectures (SPAA'04), June 2004, pp. 36–43
- [12]Amandeep Kaur Sidhu, Supriya Kinger "Analysis of Load Balancing Techniques in Cloud Computing" International Journal of Computers & Technology Volume 4 No. 2, March-April, 2013, ISSN 2277-3061
- [13]Nidhi Jain Kansal and Inderveer Chana "Existing load balancing techniques in cloud computing: a systematic review" Journal of Information Systems and Communication ISSN: 0976-8742, E-ISSN: 0976-8750, Volume 3, Issue 1, 2012, pp- 87-91.
- [14] Tejinder Sharma, Vijay Kumar Banga " Efficient and Enhanced Algorithm in Cloud Computing" International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-1, March 2013.
- [15] S. Surana, B. Godfrey, K. Lakshminarayanan, R. Karp, and I.Stoica, "Load Balancing in Dynamic Structured P2P Systems,"Performance Evaluation,vol.63,no.6,pp.217-240,Mar.2006.
- [16] VMware, <http://www.vmware.com/>, 2012.
- [17] P. Ganesan, M. Bawa, and H. Garcia-Molina, "Online Balancing of Range-Partitioned Data with Applications to Peer-to-Peer Systems,"Proc. 13th Int'l Conf. Very Large Data Bases (VLDB '04), pp. 444-455,Sept.2004.

