A Novel and Reliable Multi Data Hosting Model Over Cloud

Vellanki Sree Divya[1] ,Ch. Kodandaramu[2]
[1]M.Tech Scholar,[2]Associate Professor
[1,2]Dept of Computer Science and Engineering, Avanthi Institute of Engineering and Technology

Abstract: *We propose an efficient model of secure data storage or hosting model over cloud. Millions of data components uploaded to server every day, duplicate copy of data components reduce the space of cloud drive. Cloud acts as resource area for data owners and various cloud service providers. In this model we reduce the duplication of the data components without violating privacy and privileges or access permissions of the while sharing between multiple data owner and two-level cloud service storage. The second indirect cloud service maintains the replication of the original uploaded files.Our protocol improves the confidentiality through secure hosting with efficient authentication technique and encoding model.Our proposed model improves the performance and maintains data confidentiality than the traditional approaches.*

I.INTRODUCTION

In the recent years of online or cloud data hosting services , data hosting increased in a rapid rate such as Amazon S3, Windows Azure, Google Cloud Storage, AliyunOSS [1], and so forth. Major impressive factors of the cloud data hosting services are data reliability, scalability and low cost, most of the services works charge based on end user usage rather than fixed costs. Most of the cloud services work as data storage, Infrastructure, back end support and virtual machines[2].

Service providers may choose different types of architectures based on their requirements, if domains are specific financial, they take more care of data confidentiality and authentication while store and retrieval of data from the server.

They also design different system architectures and apply various techniques to make their services competitive. Such system diversity leads to observable performance variations across cloud vendors. End users are choosing cloud service provider based on their features or requirements, in online market various providers provides various service with cheap prices. Thegeneral status quo is that customers usually put their data into a single cloud and then simply trust to luck[3][4].

This is subject to the so-called “vendor lock-in risk”, because customerswould be confronted with a dilemma if they want to switch to other cloud venders. The vendor lock-in risk first lies in that data migration inevitably generates considerable expense. For example, moving 100 TB of data from Amazon S3 (California datacenter) to Aliyun OSS (Beijing datacenter) would consume as much as 12,300 (US) dollars. Besides, the vendor lock-in risk makes customers suffer from price adjustments of cloud vendors which are not uncommon. For example, the fluctuation of electricity bills in a region will affect the prices of cloud services in this region[5]. We notice that giant cloud vendors like Windows Azure and Google Cloud Storage have been adjusting their pricing terms[6].

II.RELATED WORK

Cloud service providers pricing is based on usage of end user, consumption of services and usage of bandwidth which requires to access the data. Service providers may choose different types of architectures based on their requirements, if domains are specific financial, they take more care of data confidentiality and authentication while store and retrieval of data from the server. They also design different system architectures and apply various techniques to make their services competitive[7][8].

Major challenges involved in cloud services are data reliability, data integrity and authentication. Cloud maintain replica, so data unavailability or data lost chances are less compared to regular servers. End users consume the services but do not have idea of where it stores and how it stores, it maintains the various auditing protocols internally for efficient data integrity and authentication of services between end users and service provider[9].

A cloud storage service provider should base its pricing on how much storage capacity a business has used, how much bandwidth was used to access its data, and the value-added services performed in the cloud such as security. Unfortunately, all the CSPS are not functioning in equal manners‟. Data storage paradigm in “Cloud” brings about many challenging design issues because of which the overall performance of the system get affected. Most of the biggest concerns with cloud data storage are: Data integrity verification at un-trusted serversFor example, the storage service provider, which experiences Byzantine failures occasionally, may decide to hide the data errors from the clients for the benefit of their own. What is more serious is that for saving money and storage space the service provider might neglect to keep or deliberately delete rarely accessed data files which belong to an ordinary client[10].

We first assign each cloud a value di which is calculated based on four factors (i.e., availability, storage, bandwidth, and operation prices) to indicate the preference of a cloud. We choose the most preferred n clouds, and then heuristically exchange the cloud in the preferred set with the cloud in the complementary set to search better solution. This is similar to the idea of Kernighan-Lin heuristic algorithm, which is applied to effectively partition graphs to minimize the sum of the costs on all edges cut.

IV. PROPOSED WORK

We propose an empirical and HEURISTIC data hosting or storage model. Storage switching models makes a replica when end user uploads documents to main server which helps while data lost. Our proxy-based implementation reduces the additional overhead on the server. Data components can be segmented in to number of chunks and encoded and uploaded o the server. We are proposing an empirical model of data deduplication technique over cloud for elimination redundant components and private cloud takes care of authentication mechanism, it obliviously reduces the additional overhead on cloud. Usually data components over cloud are encrypted and apply signatures over encoded blocks, so while uploading new components it needs to compare with same format. This proposed model reduces the redundancy of data over cloud and reduces additional overhead while authentication of users.

Data Hosting/Upload Implementation:

In our method data owner apply signature generation method on each blocks of the data and creates the hash code and encrypts the content with Triple DES algorithm and uploads in to the server. Data Components are divided into $m_1,m_2….m_n$& generates random tag key set($t_1,t_2…..t_n$) . Every individual block can be encrypted with tag keys and then it forward the file meta data details and key to the third-party auditor (verifier). There the auditor process same signature generation method and generates signature on the blocks and then verifies the both signatures if any block code is not matched that sends alert message to the data owner, then the administrator can forward only the revised information instead of total content then the user can browse the information which is given by the cloud service provider.

Step by Step Process for protocol Implementation:

Step1: Data proprietor fragments Data component D into n blocks (m1,m2… .mn).

Step2: Generates an arbitrary tag key set T (t1,t2… ..tn) to encrypt the piece with triple DES calculation and discovers signatures on encrypted blocks for authentication

Step3 : Generates irregular difficulties RA,RB and computes hash value of xor amongst RA and RB.

x := hash (RA XOR RB)

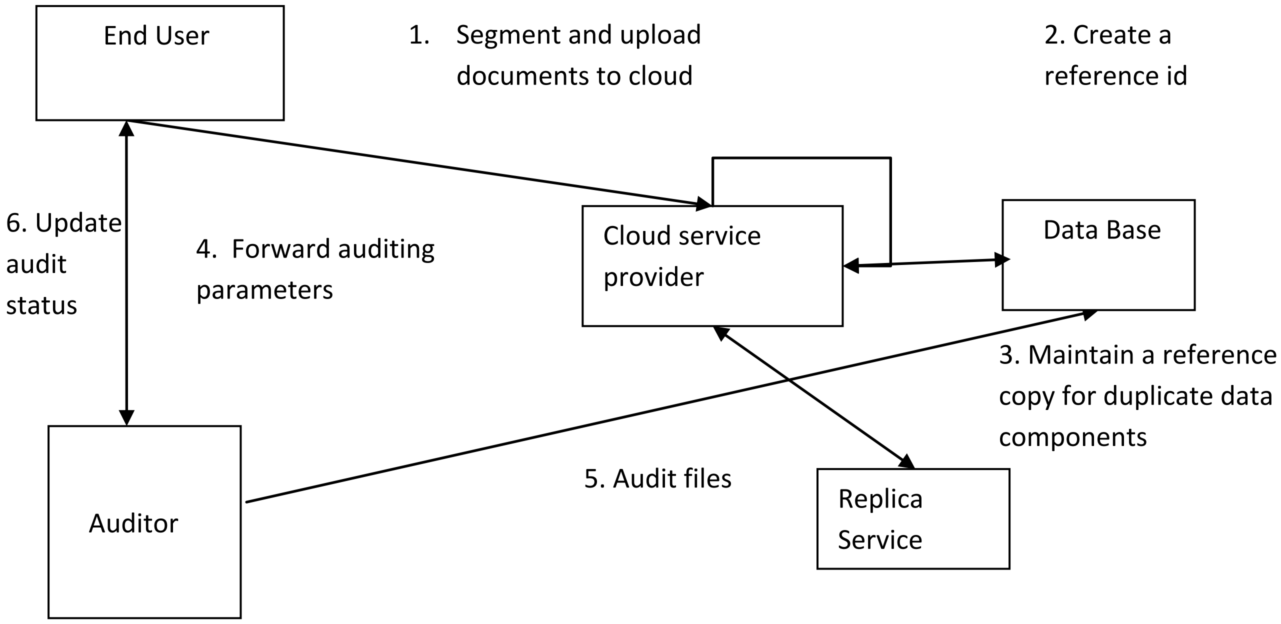
Step4 : Forward Data component, Tag key set and RB to specialist co-op and meta data and authentication parameters (Minfo RA,T (t1,t2 … . Tn)) to Auditor

Step5 : data proprietor Checks authentication by re-computing hash code with reviewer RA.

Step6 : Auditor again isolates D in ti number of blocks at server end, encrypts and applies same mark and analyzes signatures of comparing blocks

Step7 : Monitoring Status can be sent t Data proprietor through smtp usage

Step8: Auditor refreshes Data component status to the Data proprietor and updates the square if adulterated.


End User
1. Segment and upload documents to cloud
2. Create a reference id
6. Update audit status
4. Forward auditing parameters
Cloud service provider
Data Base
3. Maintain a reference copy for duplicate data components
5. Audit files
Auditor
Replica Service

Replica Switching Model:

Whenever it receives a request from end user, central service provider directly uploads to cloud server and followed by replica in to second server with switching model. If data lost or unavailable in in central server, request internally forwards to replica server and makes a copy to central server and returns documents to requested end user.

Proxy implementation improves the performance by reducing the additional overhead on cloud service. It is a virtual service, handles the requests returned by service provider and sends response to service provider. It improves the performance of the server, minimizes the cost of service access.

V. CONCLUSION

We have been concluding our current research work with efficient data hosting or storage technique with replica-based server. Which maintains a copy of hosted documents and gets to end user when central service is unavailable or lost? Proxy will improve the performance. Data components can be securely divided into chunks followed by encryption and signature generation .Our proposed model gies more efficient results than traditional models.

We can enhance our research work by enhancing the random key generation approach instead of generating with sample random number process, Even though AES algorithm provides the optimal security than the other algorithms, time complexity is more .

References:

[1] S. Liu, X. Huang, H. Fu, and G. Yang, "Understanding data characteristics and access patterns in a cloud storage system," in Proc. 13th IEEE/ACM Int. Symp. Cluster, Cloud, Grid Comput., 2013, pp. 327–334.

[2] P. Wendell, J. W. Jiang, M. J. Freedman, and J. Rexford, "Donar: Decentralized server selection for cloud services," in Proc. ACM SIGCOMM Conf., 2010, pp. 231–242.

[3] H. H. Liu, Y. Wang, Y. R. Yang, H. Wang, and C. Tian, "Optimizing cost and performance for content multihoming," in Proc. ACM SIGCOMM Conf. Appl., Technol., Archit., Protocols Comput. Commun., 2012, pp. 371–382.

[4] T. G. Papaioannou, N. Bonvin, and K. Aberer, "Scalia: An adaptive scheme for efficient multi-cloud storage," in Proc. Int. Conf. High Perform. Comput.,Netw., Storage, Anal., 2012, p. 20.

[5] J. S. Plank, "Erasure codes for storage systems: A brief primer," Usenix Mag., vol. 38, no. 6, pp. 44–50, 2013.

[6] J. S. Plank, K. M. Greenan, and E. L. Miller, "Screaming fast galois field arithmetic using Intel SIMD instructions," in Proc. 11th USENIX Conf. File Storage Technol., 2013, pp. 299–306.

[7] H. Weatherspoon and J. D. Kubiatowicz, "Erasure coding vs. replication: A quantitative comparison," in Proc. 1st Int. Workshop Peer-to-Peer Syst., 2002, pp. 328–338.

[8] R. Rodrigues and B. Liskov, "High availability in DHTs: Erasure coding vs. replication," in Proc. 4th Int. Conf. Peer-to-Peer Syst., 2005, pp. 226–239.

[9] I. M. Bomze, M. Budinich, P. M. Pardalos, and M. Pelillo, "The maximum clique problem," in Handbook of Combinatorial Optimization. New York, NY, USA: Springer, 1999, pp. 1–74.

[10] B. W. Kernighan and S. Lin, “An efficient heuristic procedure for partitioning graphs,” Bell Syst. Tech. J., vol. 49, no. 2, pp. 291–307, 1970

BIOGRAPHIES

Vellanki SreeDivya is an M.Tech Scholar studying in Dept Of Computer Science And Engineering in Avanthi Institute Of Engineering And Technology. her interests in cloud computing.

CH.Kodanda Ramu completed Msc and MTech. He is working as Associate professor in Avanthi Institute of Engineering And Technology.