ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

Enhancing Security in Multicloud by Reallocation of Data

V.S.Harish kumar[1], S.Hariharan[2], Y.Nishant kumar[3], N.V.S.Natteshan[4]

Student, Department of Computer Science and Engineering, Kingston College of Engineering, Vellore, India[1,2,3]

Assistant Professor, Department of Computer Science and Engineering, Kingston College of Engineering, Vellore, India[4]

ABSTRACT: In Cloud services, Data storage and data retrieval play a major role to enhance these process we are providing a secure solution. We provide a MultiCloud environment to store data. Multicloud environment is a gateway to multiple number of cloud storages. The data is split into several blocks and storage in the multi cloud environment.Everytime user access the data, Data will be dynamically reallocated into different blocks. Thus providing a secure data storage. Proposed work focuses on security and corruption detection of data. For this purpose we developed own system which co-actively work with parallel data processing approach. We also utilized concept of MD5 to tag files with document identity number. Hence if any part of file missing system can recover data using anti-parallel resource algorithm.

KEYWORDS: Cloud Computing,MultiCloud,Dynamic allocation

I. INTRODUCTION

Cloud computing can be useful to organizations and industries for conducting their core business activities in effective ways since the management and monitoring task for data cores is minimized. Businesses can also save on power costs as the resources needs are reduced. People may consider if cloud computing is such a great thing then why most of the businesses are not going for it and as per the research the reason for this problem is security. The third party is tangled called CSP (Cloud Service Provider) to whom businesses have to provide their files with delicate data. This paper survey's recent research associated to security of single and multi-cloud comes up with possible solutions for conservation of security. However, multi-cloud computing is relatively new concept, biggest security aspects in cloud computing basically are data intrusion, data integrity and service availability are handled in much better way in multicloud than single cloud computing .This project work encourages the use of multi-cloud architecture. By organizing IT infrastructure and services over the network, an organization can acquire these resources on as-needed base and avoid the capital costs of software and hardware. Fundamental Characteristics of Cloud Computing

1. Availability: cloud amenities can possibly gain access to over the network through the usage of standardized mechanism that hold up different users, like mobile phones, tablets, mainframes and work stations.
2. Flexibility: The user probably will make a judgment on the use of computing amenities such as server time and network storage, based on of their current needs, with no more communication with dissimilar service providers in cloud.
3. Utilising multiple resources: In addition to classical virtualization, cloud computing uses in adding the capabilities of mechanization of services and multi-tenancy of users at shared information resources. Common use of the same technical resources is the vital feature of cloud

Computing

4. Elasticity: The client may simply rise or fall the capacities afforded using the current requirements. The capacities are limitless for the user

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

II. LITERATURE REVIEW

Cloud computing is giving numerous infrastructure to store and work client's data a. customers need not to own the facility and system, they are just accessing or making use of the infrastructure; they can escape from funding a huge amount of capital and utilize resources as a service, paying instead for what they use. Cloud storage can save users lots of money over local storage units [1]. When user uses the cloud, user doesn't know exactly where your data is hosted, what country it will be stored in? Data should be stored and processed only in specific jurisdictions as define by user. Provider should also make a contractual commitment to obey local privacy requirement on behalf of their customers. Data-centered policies that are generated when a user provides personal or sensitive information, that information travels throughout its lifetime to ensure that the information is used only in accordance with the policy [5]. Data that is generated during running of program on instances is all customer data and therefore provider should not perform backups [7, 8]. Cong Wang investigated the problem of data security [2] in cloud data storage, which is essentially a distributed storage system but doesn't provide recovery facility. The processing frameworks which are currently used have been designed for static, homogenous cluster setup and disregard the particular nature of cloud [3]. Fast and secure message authentication in cryptography [4] provides splitting of file only supports for encryption and decryption of file in cloud computing. Proofs of retrievability protocol represents the proof of data stored in server is intact and retrievable but not support for file updates, as well as publicly verifiable PORs .To improve security and trust we also have included continuous auditing of data by which we can audit the security of data every now and then and give recourses a very high chance of safety [10].

III. BACKGROUND

A. CLOUD COMPUTING

Cloud computing is a general term for the delivery of hosted services over the internet. Cloud computing enables companies to consume a compute resource, such as a virtual machine (VMs), storage or an application, as a utility -- just like electricity -- rather than having to build and maintain computing infrastructures in house. Cloud computing boasts several attractive benefits for businesses and end users. Three of the main benefits of cloud computing are:

Self-service provisioning: End users can spin up compute resources for almost any type of workload on demand. This eliminates the traditional need for IT administrators to provision and manage compute resources.

Elasticity: Companies can scale up as computing needs increase and scale down again as demands decrease. This eliminates the need for massive investments in local infrastructure which may or may not remain active.

Pay per use: Compute resources are measured at a granular level, allowing users to pay only for the resources and workloads they use.

B. MULTI CLOUD

Multicloud is the use of multiple cloud computing services in a single heterogeneous architecture.

There are a number of reasons for deploying a multicloud architecture, including reducing reliance on any single vendor, increasing flexibility through choice, and mitigating against disasters. It is similar to the use of best-of-breed applications from multiple developers on a personal computer, rather than the defaults offered by the operating system vendor. It is recognition of the fact that no one provider can be everything for everyone. It differs from hybrid cloud in that it refers to multiple cloud services rather than multiple deployment modes (public, private, and legacy).

Various issues also present themselves in a multicloud environment. Security and governance is more complicated, and more "moving parts" may create resiliency issues. Selection of the right cloud products and services can also present a challenge, and users may suffer from the paradox of choice.

C. DYNAMIC REALLOCATION

Dynamic reallocation is a method that allocates data in different location that of previous location. Dynamic reallocation enables security in multicloud data storage.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

In early methods multicloud environments did not have dynamic reallocation this flaw led to multiple failures in the previous models. So by analysing technologies and methods we arrive at scheduling data into blocks that are frequently altered.

D. CONTINUOUS AUDITING

Continuous auditing is defined as a methodology that enables independent auditors to provide written assurance on a subject matter, using a series of auditors' reports issued virtually simultaneously with, or a short period of time after, the occurrence of events underlying the subject matter [10]

IV. OVERALL APPROACH

To provide parallel data allocation and data security we have stored data on multiple cloud nodes and in case if any data/part of file is missing or virus infected then we can able to recover data/resources. For recovery purpose we have implemented self-derived algorithm which collectively work with MD5, dynamic data allocation approach and logical addresses of file. We aim to recover data 100% which is not possible with existing systems. In our proposed system user can register with cloud service provider. Registered client have to login to upload file which he/she want to store on cloud. Here client expects higher level of security for his/her resources. In this system server receives resources uploaded by client. Server reviews received resource to filter banned data/documentation. The main server receives resources uploaded by client. Consider user have upload "test.txt" file proposed algorithm encrypts the file and splits into number of data chunks. These splitted data chunks parallelly processed to store on auxiliary cloud server-1 and cloud server-2.This system provides more security for data and there will not be any direct access to user to auxiliary server's data. We have implemented MD5 algorithm to check all files are stored as it is or any files are modified due to impact of any mishap. In case of data corruption or missing file part, we can recover original data from cloud server by using MD5 hash value.

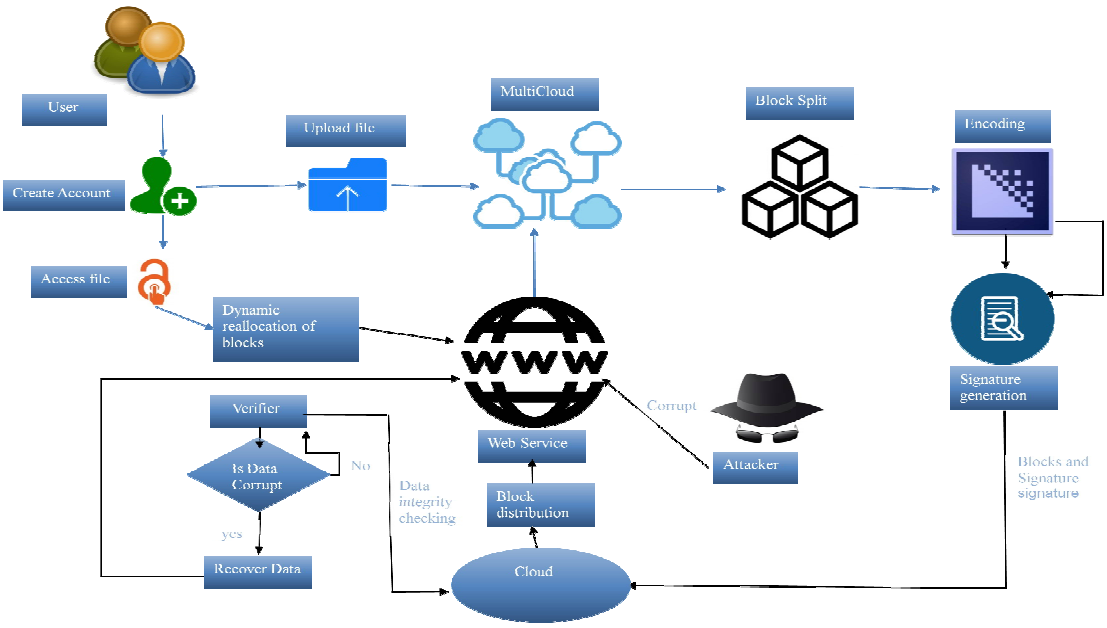


Fig 1: Overall Architecture

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

A. CLIENT REGISTRATION

Client registration and login window has been developed. By which new user can avail system facility. In this module existing client can log-in to upload his/her data file. If user is registered then he/she gets ID & password. At the time of login user has to provide correct ID & password. Then system checks the user ID & password if both are correct then user logged into the system. Unauthorized person cannot be able to login into the system.

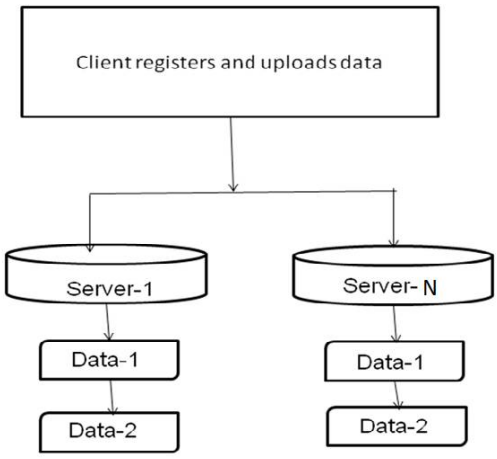


Fig 2:Block splitup

B. FILE ENCRYPTION AND UPLOAD

In this module, when user is registered he/she uploads file on main server and the main server encrypts the file and stored on server. Base64 algorithm is implemented for encryption of the data .The main server receives files uploaded by client. Then file is splitted into number of data chunks. The Base64 algorithm will encrypt the splitted file parts. These data will be parallelly processed to store on nodes.

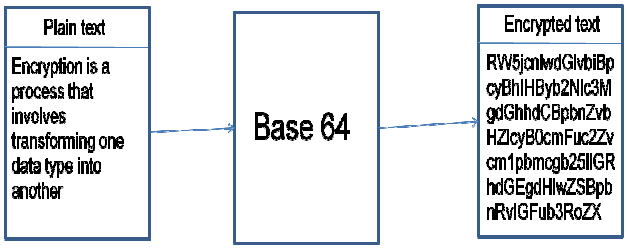


Fig 3: Encryption

C. FILE DECRYPTION AND DOWNLOAD

User can download the files from server. The main server decrypts the file using Base64 algorithm and gives to the user. The Base64 is also implemented for decryption of the data. This will provide more security for data and there will not be any direct access of user to server's data.

D. DYNAMIC DATA ALLOCATION

In this module, we have created virtual nodes for allocating different servers for storage. In our project, we have created nodes for storing the files. Suppose file is encrypted and stored on different nodes. When we are trying to download file then file parts are collected from various nodes. We have presented a system that uses virtualization

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

technology to store data resources (file) dynamically and support green computing by optimizing the number of servers in use. We have combined different types of workloads (storage) nicely and improved the overall utilization of server resources. We have developed a set of heuristics that prevent overload in the system effectively while saving energy used.

E. FILE INTEGRITY DETECTION AND RECOVERY

In this module, we have implemented MD5 algorithm. User's file is splitted and encrypted using Base64 algorithm. That file parts are stored on different nodes. By using MD5, we have calculated hash value and stored on database. When user wants to download his/her file then again hash value of current file is calculated and verified with old hash value. If both hash values are matched then user gets his original file. If hash values are not same then we can say that file parts are corrupted or infected. When we know that file is corrupted then we can recover this file.

V. RESULTS

A. BLOCK SPLITTING ALGORITHM

	File size	Fragment /Block
File1	50kb	25kb
File 2	50mb	5mb
File3	500mb	50mb
File 4	1Gb	150mb

Table 1: Block size Description

In the above table, we have shown the number of divisions that the multi cloud makes based on the blocks that are allocated and the original source file size. The number of blocks depends on the file size as specified in the above table. We use a block splitting algorithm to split the source file into number of chunks/parts.

B.DATA RECOVERY

In this phase, we have analysed the percentage efficiency of data recovery by comparing the duplicate copy of the source file and the recovered data. We have recovered the data using data recovery techniques involving sign verification and hash functions. The result of the study is found to be a 100% match between the recovered data and original source data.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

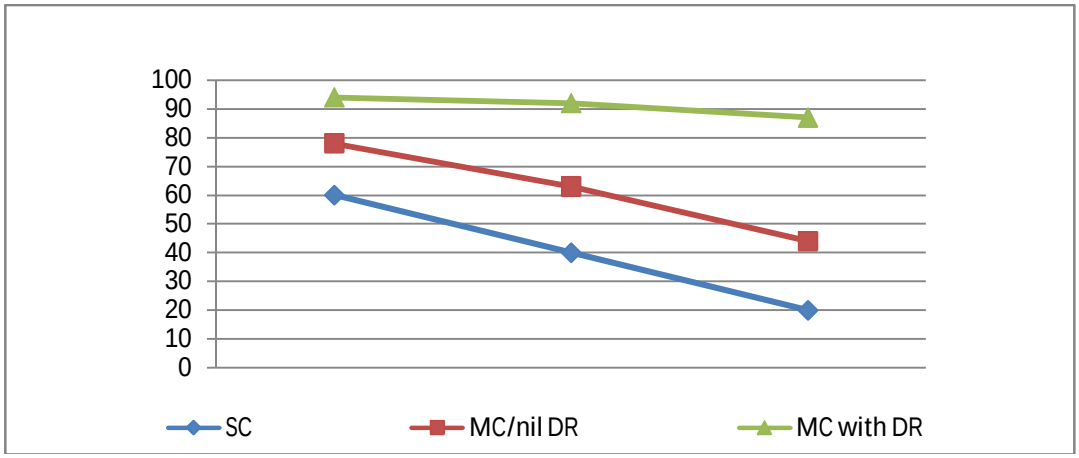
International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

C.STABILITY OF DATA IN MULTICLOUD



Graph 1:stability vs intrusion

Xaxis-Stability of system in percentage
Yaxis-Intrusions to the system
SC* -Single server environment
MC/nil DR -multicloud without Dynamic Reallocation
MC with DR -Multicloud with dynamic Reallocation

In this graph we have plotted stability of the system vs. number of intrusion. Three simulations with single server environment, multicloud environment without dynamic reallocation and multicloud with dynamic reallocation were done. By the analysis of the simulations and study, we found that our proposed dynamic allocation technique has a higher stability.

VI. CONCLUSION

We constructed a system with multiple numbers of servers to store data which dynamically rearranged data chunks on every access of the user and also in any case of intrusion ultimately leading to better and improved cloud security. Thus we conclude the paper by stating that our paper provides hack proof data integrity checking in a multicloud environment.

REFERENCES

[1] .Amazon.com. Amazon simple storage service (Amazon S3), 2016.
[2] Cong Wang, Qian Wang, Kui Ren and Wenjing Lou "Ensuring Data Storage Security in Cloud Computing", IEEE 2015.
[3] Daniel Warneke and Odej Kao "Dynamic resource allocation for efficient parallel data processing in cloud",IEEE 2015.
[4] D. Warneke and O. Kao, "Nephele: Efficient Parallel Data Processing in the Cloud," Proc. Second Workshop ManyTask Computing on Grids and Supercomputers (MTAGS '09), pp. 1-10, 2016.
[5] G. Ateniese, R. Burns, R. Curtmola, J. Herring, L. Kissner, Z. Peterson, and D. Song. "Provable data possession at untrusted stores" in 14th ACM CCS, pages 598–609, 2015.
[6] J. Black, S. Halevi, H. Krawczyk, T. Krovetz, and P. Rogaway. "UMAC: Fast and secure message authentication" in CRYPTO, volume 1666 of LNCS, pages 216–233, 1999.
[7] K. D. Bowers, A. Juels, and A Oprea " A high-availability and integrity layer for cloud storage,",2008. IACR ePrint manuscript 2008/489.
[8] K. D. Bowers, A. Juels, and A Oprea "Proofs of retrievability: Theory and implementation", 2008. IACR ePrint manuscript 2008/175.
[9] D. Nurmi, R. Wolski, C. Grzegorczyk, G. Obertelli, S. Soman, L.Youseff, and D.Zagorodnov, "Eucalyptus: A Technical Report on an Elastic Utility Computing Architecture Linking Your Programs to Useful Systems," technical report, Univ. of California, Santa Barbara, 2008.
[10] CICA/AICPA, "Continuous auditing", 1999.