Cloud Storage Fault Tolerance System using Minimum Storage Regeneration Codes

[1]Ashwini Singh S, [2]Y S Nijagunarya
[1]PG Student, [2]Professor
[1]Computer Science and Engineering,
[1]Siddaganga Institute of Technology, Tumakuru, India

Abstract : With the specific knowledge of application domain, comprehensive analysis and various prevalent techniques are needed for implementing the cloud fault tolerance policies. Here, fault tolerance is the property that enables a system to continue operating properly in the event of the failure of some of its components. Fault tolerance is one of the most imperative issues in the cloud to deliver reliable services. It is difficult to implement due to dynamic service infrastructure, complex configurations and various interdependencies existing in the cloud. In this work, we are using minimum storage regenerating codes which is a class of distributed storage codes, and an important class of optimal regenerating codes which is a class of distributed storage code, and an important class of optimal regenerating codes that minimize the amount of data stored per node and then repair bandwidth. The file can be recovered by downloading the content of any of the cloud storage nodes. Our main aim in this work is to achieve 50 percent fault tolerance using 100 percent extra storage space. And we have implemented three functionalities such as, cryptographic technique for security purpose, a file management technique for managing the file and minimum storage regenerating technique to reconstruct a failed cloud storage nodes in distributed storage systems.

IndexTerms - Cloud Storage, MSR cloud, Fault Tolerance System, Cryptography, AES Algorithm, File Management

I. INTRODUCTION

The statement used to relate a diversity of computing concepts is cloud computing which require a large number of computers connected through a real time transmission network such as the Internet. Cloud computing is a interpretation over a network for distributed computing, that is capability to run a program at the same time on many connected computers. The network-based services commonly refers to a phrase, that runs on one or more ideal machines appears to be provided by hardware and simulated by software. Such virtual server are popular in internet hosting domain, do not physically exist and can scale up and moved around without affecting the organization - feasibly, preferably like a cloud.

Cloud storage furnish an on-demand online managed backup service. However, the single-cloud storage provider raises consumption emission simulation such as single point of failure and customer lock-in. As indicated, a reasonable solution is to streak data across different cloud service providers. By utilizing the variety of multiple clouds, we can upgrade the fault tolerance of cloud storage. While striping data with normal erasure, the codes performs skillfully when some clouds experience lasting a short time temporary failures. There are real-life cases showing that constant failures do occur and are not always predicted. In perspective of this, this exertion focuses on unpredicted permanent cloud failures. When cloud fails imperishably, to maintain fault tolerance and repetition of data is necessary to actuate service.

With a fast flowing of data manufacturing in companies, the storage space demand has become more. This clouds storage leads to its prominence. The data which is stored in virtual disk, is a representation of computer data storage which is Cloud storage. The data centre facility may obtained Cloud storage service, for keeping the data obtained and accessible, these cloud storage service providers are having an obligation. Cloud storage works through a large group of networked computer servers with the process of designing, developing, and deploying on virtualization, providing the person who actually uses a particular product and application with a storage virtual architecture that is climbed according to application requirements.

Figure.1 Cloud Storage System

Private cloud service provides same computing process, flexibility and storage mechanism. To perform operating properly even when one or more fault occurs, the fault tolerance system is a attribute that activates. The system enables a design to continue operating where the fault tolerant operation is intended.

In computer, the term failure is used commonly to describe a reduction in throughput or in response time when it is increased even when partially fails. Minimum storage regenerating codes have attracted much attention and interest on research in last period of years. The important class of minimum storage regenerating is most optimum revitalize codes that reduces the quantity

stored data per node and repair the range of frequencies within a given band. There are distinct types of failures in a cloud platform that are possibly occur and gives the existing fault tolerance techniques. Cloud computing provides various resources in a distributed computing environment as a service on demand to the user. Usually cloud can be classified based on deployment model, such as hybrid cloud, public cloud, community cloud, and private cloud.

In this paper, our proposal is applied to the Hybrid cloud, and the motivation of doing this paper is, here we are using 100 percent storage space and achieving 50 percent fault tolerance to retrieve the data and maintaining cryptographic scheme. The which is prolonged by third person on the storage location, is stored by the users on the cloud. As soon as the data transferred in the cloud, the users may lose data and it cannot be retrieved most of the time due to limited storage space and there will be no much security, so came the motivation to work on this paper.

Varying sizes in businesses field, an outstanding role will be played by cloud computing by providing computing services. But the cloud services failure can have financial loss substantially. To make sure higher trustworthy, fault tolerant is exceptionally to be needed in cloud computing system. The capability of a system to distribute the intended facility is fault tolerance, that when lack of success and errors happen in the system. There are so many expertises in cloud computing for fault tolerance like restart, retry, job migration. In this paper we are using MSR codes technique that is minimum storage regenerating, it is a category of distributed storage codes. For providing the quality of being trustworthy of data and systematic repair of failed nodes in distinctive storage systems, regenerating codes are proposed. By downloading the data stored in any of these n nodes the entire message must be able to reconstruct by the data collector.

In this project there are three functionalities in web server, first is minimum storage regenerating code is a technique were we are going to use it for the fault tolerance purpose. When some storage nodes goes fails also we can able to retrieve the data with the help of the previous presenting node that is called MSR Code (minimum storage regenerating codes). Second is cryptographic that is encryption and decryption, in addition to this we do when once the file is uploaded, the file is divided into blocks for that each block we generate a regeneration code and totally we have eight blocks for a single single file. That eight blocks we are going to encrypt using this cryptographic technique. Third is file management, in the file management which user uploaded which file, who is downloading the file, and whether the storage nodes are active or not active.

There are two primary storage node and two regenerating code storage node we are having, so that it will take the decision of the thing whether from where to download the nodes and merge it and decrypt it and give it to the user, that is called file management.

Existing system focuses on unpredicted lasting cloud failures. When a cloud fails constantly, it is required to operate repair to maintain data repeated unnecessarily and fault tolerance. A repair operation recover data from existing surviving clouds over the network and recreate the lost data in a new cloud. Today's cloud storage contributor charge users for outbound, so moving an huge amount of data across clouds can initiate notable budgetary costs. It is important to lower the repair traffic and hence, the economic cost due to data migration. Existing system need more cloud storage space which in-turn cause more cost.

II. RELATED WORK

In [1] author has proposed the algorithm which efficiently recover the backup during the fraudulent and the cost is also reduced. Further, it says a new log based recovery method to be introduced for retrieving the data from the failure disk. In this algorithm, read only the index of failure file system instead of reading the whole file.

In [2] author has discussed about the strategy which is proposed and uses the pass rate of the computing virtual nodes with the fault using the relay/checkpoint manager technique by applying the reward renewal process theorem. It tries to repair the fault generated before the time limit as a fault tolerance mechanism, with the help of the pass and fail rate analysis obtained from the experimental results as well as comparing the performance of the existing. And further it says they will aim at addressing fault tolerance challenge especially in a high performance large scale environment in a real life scenario.

In [3] author has explained how minimum storage regenerating codes are an important class of optimal regenerating codes that minimize the first amount of data stored per node and then about the repair bandwidth. The result in this paper is an explicit construction of systematic repair codes for all parameters possible values. And it aims to see whether the similar bounds exist for general parameters or not is left for future work. So it is still the questioning of constructing minimum storage regenerating codes that optimally repair parity nodes.

In [4] author has done survey of network codes secure storage (NCSS) which is an arbitrator, here different cloud storage framework are not that much addresses the commitment of today's cloud argumentation cache. NCSS is not as the only source that gives the adaptation to internal failure, FMSR code execution should dispenses with the cryptography interest of capacity all through repair.

III. METHODOLOGY

The methodology in this paper going to use

- Minimum-storage regenerating (MSR) codes technique.
- File upload and download process
- Model-View-Controller (MVC Architecture)
- MD5 algorithm for hash function and for security purpose AES algorithm.

File uploading process takes place in several steps- user has to select the file to be uploaded. Transfer the file to the server. Then the server divides the file into four equal blocks, name the blocks as M N O and P. generate the MAC for all the four blocks, store the MACs in table. Now create another four blocks, M (XOR) O, N (XOR) P, M (XOR) P, (N (XOR) P) (XOR) O). now there are eight blocks, store two blocks in each cloud storage as shown in table.1.

SN1	SN2	SN3	SN4
M	O	M(xor)O	M(xor)P
N	P	N(xor)P	(N(xor)P)(xor)O)

Table.1. SN stands for Storage Node.

File downloading process takes places in several steps- first user has to select the file to be downloaded. Check the status of primary cloud storage (storage node 1 and storage node 2). If it is active, download all the four blocks. Generate the MAC, Retrieve the MAC from the table, Compare the MAC and Display the result. If the result is pass then merge the blocks and form a file. Download the file to the local system. If any one or both the primary cloud storage is inactive, then system has to perform MSR technique as shown in table.2.

Cloud Status				**Block Retrieval Process**			
SN 1	**SN 2**	**SN 3**	**SN 4**	**M**	**N**	**O**	**P**
A	NA	A	NA	M	N	(M+O)+M	(N+P) + N
NA	A	A	NA	(M+O) +O	(N+P) +P	O	P
A	NA	NA	A	M	N	(N+P) +((N+P)+O)	(M +P)+ M
A	NA	A	A	(M+O) +O	(N+P) +P	(N+P) +((N+P)+O)	(M +P)+ M

Table.2. MSR Technique

IV. PROPOSED SYSTEM ARCHITECTURE

System Architecture is relating to a model to a model that describes exactly the behavior, the structure, and more views of a system. The architecture is formal description and of a system representation, that supports reasoning and organized in a way about the behaviors and the structures of the system. The figure 5.1 shows, the system architecture is build in three layers, the top layer is cloud storage layer and we are using four clouds that is storage node 1, storage node 2, storage node 3 and storage node 4. In this four storage nodes the files will be stored. In the second layer there is web server layer and in the web server it has three functionalities which we are using in our work.

First text file will be uploaded in cloud, and the text file size can be of any size like 410 bytes and when the file gets uploaded in the cloud it get encrypted and stored in different storage nodes like storage node1 storage node2 storage node3……storage node n. In our work we are using four storage nodes where the file will be get stored in. The file which is stored in the cloud gets encrypted. Encryption is the process of converting information or data into a code, especially to prevent unauthorized access. There are many algorithms for this encryption and decryption process, and here we are using Advanced Encryption Standards (AES) algorithm, this is one of the most popular and extensively brought up symmetric encryption algorithm which is encountered likely to be nowadays. It is found six times faster, better and stronger than any other algorithm. It is based on permutation and substitution network, which consists of a number of events of linked operation by some of which it involves in replacing inputs by specific outputs that is substitutions and others interestingly involve in shuffling bits around which is a permutations. Then decryption generally is the reverse process of the encryption. It is the decoding process of data which is in secret format and has been encrypted.

AES calculation is mostly done in particular finite field. For example, if there are 16 bytes, then AES operates 4X4 matrix column-major order array of bytes. That is b_0, b_1,……., b_{15}, the key size used in AES specifies the number of transformation and rounds that convert the input called as the plaintext, to the final output called as the ciphertext. With the fixed block size of 128 bits, the file size is of 128, or 192 or 256 bits is specified with block and has a key sizes may be any multiple of 32 bits with a minimum of 128 bits and maximum of 256 bits.

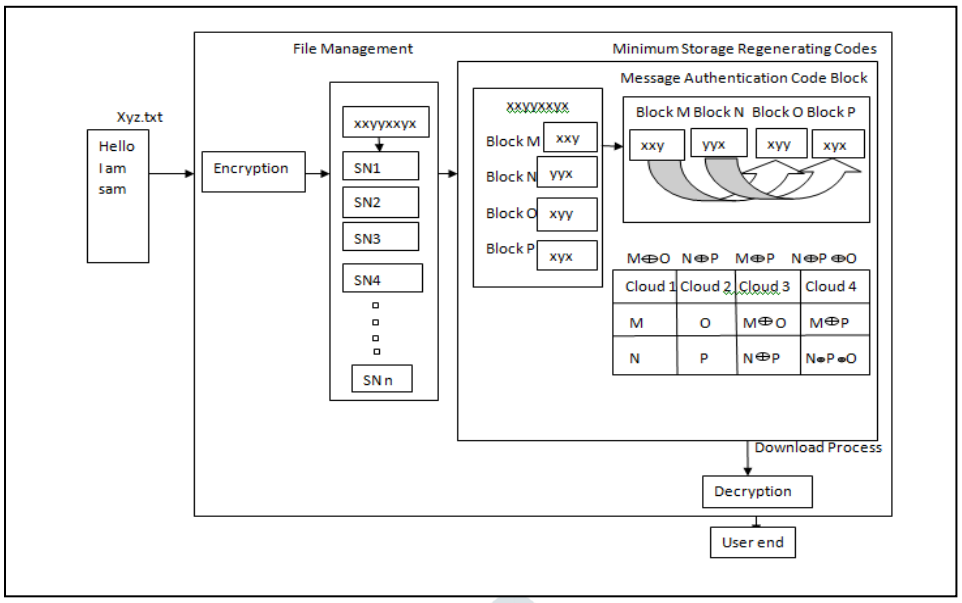


Figure.2. System Architecture

Then the file which is stored in storage nodes that is let us take storage node1 that file will gets divided into four blocks and each block size depends upon file size may be of any multiple of 32 bits. Then for this four block, message authentication code will be generated and message authentication code is symmetric key algorithm to provide message authentication in cryptographic technique. The message authentication code is a encrypted process generated along with a message to protect message authentication. Then the sender will use some known message authentication code algorithm, which inputs the message, the secret key, and then produces a message authentication code value. The algorithm is used MD5 hashing algorithm, and this is one way cryptographic function, that accept a message which is of any length as input and with return to its output of fixed length value used for authenticating the original message. While coming to our work, here the file gets encrypted and get upload in the storage nodes and the uploaded file is divided into blocks and for that block we generate a message authentication code and then have totally eight blocks.

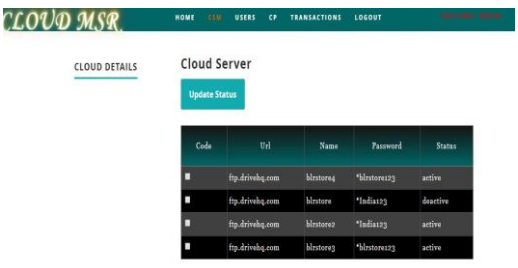
Suppose user has to upload one file called abc.txt, for this we are having four blocks as M, N, O and P and generate message authentication code (MAC) for all the four blocks and will store the MAC in table. Now create another four blocks as M $\oplus$O, N P, M $\oplus$ P, (N$\oplus$P)$\oplus$O) now there are eight blocks, and we store these blocks two each in storage node. Then when user selects the file to download, it will check the status of primary storage node. If it is active, it will first download all the four blocks and generate a MAC and it will retrieve the MAC from the table, compare the MAC and display the result and then get download to the local system.

Then comes the file management, in this file management we can know which user uploaded which file and who is downloading the file and whether the storage nodes are active or not active. Then we are having two primary storage node and two regenerating code storage node, so that it will take the decision from which node to download and merge it and decrypt it and give it to the user that is file management.

V. RESULTS

The final outcome of activity or incident indicated qualitatatively or quantitatively is called result. Production examination is an operational analysis which is a method of examining current performance, and is a set of basic relating to connection between the production quantities.

Figure 1. Admin side Cloud Server Management (CSM)

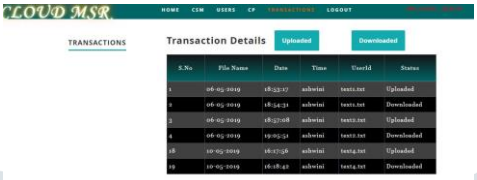


The above screen shots shows the four drive hq cloud which we have taken in this project

Figure 2. Admin side user details

The above screen shot shows the user registration details that is user side registered by giving their details

Figure 3. Admin side Transaction details

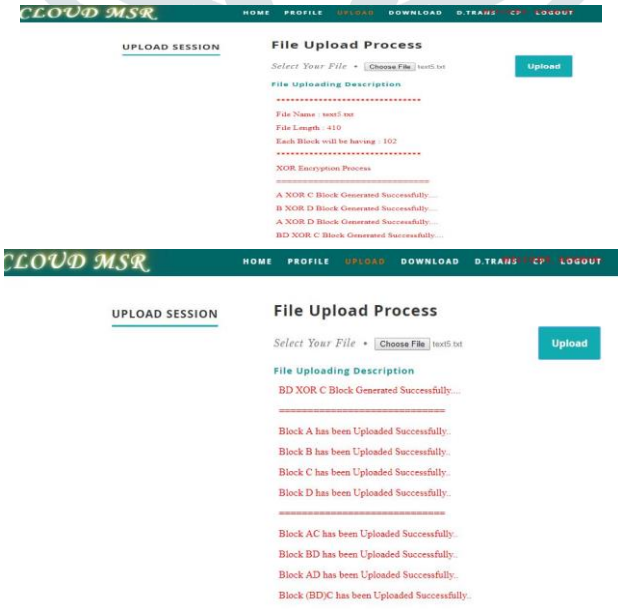


The above screen shot shows the users transaction details of upload and download with date and time

Figure 4. User side Home page

The above screen shot shows the user login profile details that is their user name, email id, phone number.

Figure 5. User side Upload Process

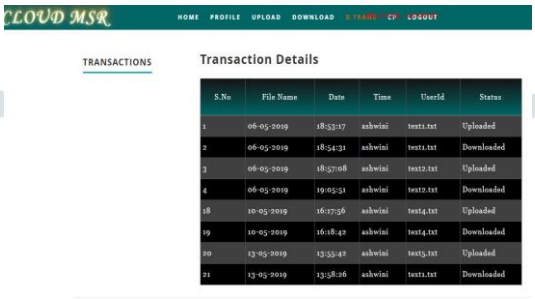


The above session shows the file uploading process, here it first checks the file size and divided into blocks totally eight blocks will be created for one file

Figure 6. User side Download Process

The above screen shot shows the download process, here it shows file download description process and shows in which cloud data is available.

Figure 7. User side Transaction details



The above sessioin shows the user transaction details that is uploaded and downloaded date and time details.

VI. CONCLUSION

Our minimum storage regenerating code implementation maintains double-fault tolerance and has the same storage cost as in traditional data protection schemes, but uses less repair traffic when recovering a single-cloud failure. While this work motivated by and established with multiple-cloud storage in mind, we point out that MSR are prone to failures and network transmission bandwidth is limited. In this case, minimizing repair traffic is important for reducing the overall repair time.

REFERENCES

[1] S.Sangeetha, B.Rasina Begum "Regenerating Failed Cloud With Minimum Storage" International Journal of Innovations in scientific and Engineering Research Vol 2 Issue 4 APR 2015/14.

[2] Bashir Mohammed, Mariam Kiran, Irfan-Ullah Awan and Kabiru .M. Maiyama "Optimising Fault Tolerace in Real-Time Cloud Computing IaaS Environment" 2016 IEEE 4th International Conference on Future Internet of Things and Cloud.

[3] Sreechakra Goparaju, Arman Fazeli, Alexander Vardy "Minimum Storage Regenerating Codes For All Parameters".

[4] Madhu M V, Vani B "A Survey Network Codes Secure Storage in a Cloud-of-clouds" International Journal of Computer Science and Information Technology Research vol. 3, Issue 3, pp☹256-271), Month: July-September 2015.

[5] Suruchi Talwani, Inderveer Chana "Fault Tolerance Techniques for Scientific Applications In Cloud" 2017 2nd International Conference On Telecommunication and Networks.