BACKUP AND RECOVERY SYSTEM USING SEED BLOCK ALGORITHM

Kakad Umesh, Kankhar Mahesh ,Mysore Ajay , Nitin Rathee

Abstract
Cloud computing relies on restricting sharing of resources to achieve coherence and economies of scale, similar to a utility (like the electricity grid) over the network. At the setting up of cloud computing is the broader concept of converged infrastructure and services. In cloud computing, to maintain this data efficiently, there is a need of data recovery services. To achieve this we are using seed block algorithm. It collects the information from user and recovered the file in case of file deletion or if cloud gets destroyed. It is time efficient and it also focuses on security concerns. For backup of data we are use remote backup server. In case of data loss or if cloud get destroyed the data will be recovered from the remote backup server.

Keywords-seed block, remote backup server, data recovery services.

1.INTRODUCTION

Cloud computing is more striking a service having a great potential to alter the large part of the IT industry. Cloud computing is the centralized storage for the data and it also provides the online access to various computer services and resources. Cloud computing broadly focuses on maximizing the efficiency of shared resources. Clouds can be classified according to users as private, public or Hybrid. Today, Cloud Computing is itself a gigantic technology which is surpassing all the previous technology of computing (like cluster, distributed etc.) of this competitive and challenging for the IT world. The need of cloud computing is increasing day by day as its advantages overcome the disadvantage of various early computing techniques. The storage on cloud provides online storage where data stored in form of virtualized pool that is usually hosted by third parties.

The hosting company operates large data on large data center and according to the requirements of the customer these data center virtualized the resources and expose them as the storage pools that help user to store files or data objects [1].Simultaneously number of user shares the storage and other resources of cloud; it is possible that other customers can access your data. The human mistake, equipment's fault, network connectivity problem, or any criminal intestinally may put our cloud storage on the risk. And changes into the cloud are also made very fast or frequently it can be called as data dynamics. However, still various successful techniques are lagging behind some critical issues like low cost, implementation complexity, security, time related issues. To overcome these issues, in this paper we are using Seed Block Algorithm (SBA) [1]. The SBA is used to recover the files in case of the file deletion or if the cloud gets destroyed due to any reason.

2. LITERATURE SURVEY

There are many techniques have been proposed HSDRT [2], PCS [3], Linux Box [4] etc. that, discussed the data recovery process.

PCS: - It is reliable, simple, easy to use and more convenient for data recovery totally based on parity recovery service. The probability of recovering data is very high. For the recovery of data, it generates a virtual disk in user system for the backup of data and creates parity groups across that virtual disk, and store parity data of parity group in cloud. It uses the Exclusive–OR ($\oplus$) for creating Parity information. Its limitation is that implementation complexities high.

HSDRT: - It is efficient technique for the movable clients such as laptop, smart phones

etc. But it fails to manage the low cost for the implementation of the recovery and also unable to control the data duplication. This technique makes use of an effective ultra-widely distributed data transfer mechanism and encryption technology. This proposed system follows two sequences one is Backup sequence and second is Recovery sequence. In Backup sequence, it collects the data to be backed-up and in Recovery Sequence, when some disasters take place or periodically, the Supervisory Server (one of the components of the HSDRT) starts the recovery sequence. This model is unable to give a perfect solution for back-up and recovery. Its limitation is costly, increase redundancy.

Linux Box: -This performs data back-up and recovery with very low cost. There is low level of protection. It also makes the process of migration from one cloud service provider to other very easy. This solution eliminates consumer's dependency on the ISP and its associated backup cost. It can do all these at very less cost named as simple Linux box which will sync up the data at block/file level from the cloud service provider to the customer. It incorporates an application on Linux box that will perform backup of the cloud on the local drives. The data transmission will be encrypted and secure. Its limitations are requiring higher bandwidth, complete backup at a time.

Above all techniques are enabling to solve the issues of remote backup. So in our paper SBA overcomes these issues like cost, implementation complexity, etc.

3. REMOTE DATA BACKUP SERVER

In general, when we talk about Backup server of main cloud that is the replica of main cloud. When this Backup server is far away from the main server and having the complete state of the main cloud, then this Backup server is termed as Remote Data Backup Server. The main cloud is called as the central data storage and remote backup cloud is termed as remote repository.

And if the main cloud lost its data under any circumstances due drastic reason or deletion

that is done by mistakenly and then it uses the information from the remote data storage. The purpose of the remote backup is to help user to collect information from any remote location.
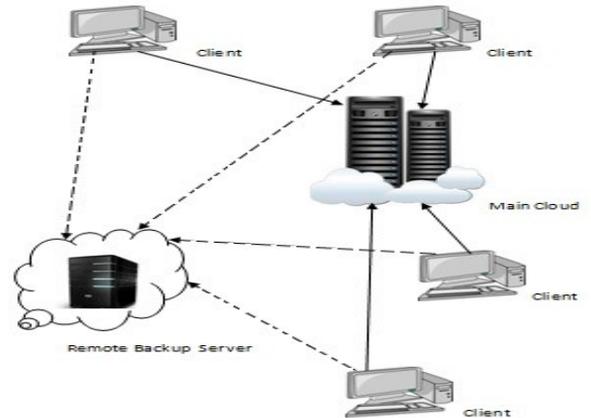
------ Indirect   Direct

Fig.1 Remote data Backup Server and its Architecture

The Remote backup service has the following issues:

1) Data Integrity

Data Integrity is related with complete state and the whole structure of the server. It checks that data such that it remains unaltered during transmission and reception. It is the measure of the validity and correctness of the data present in the server.

2) Data security

Data security is related with the full protection of client's data. It is also the utmost priority for the remote server. And, it should be not able to access by third party or any other users/client's so that it can misuse that data against that client.

3) Data Confidentiality

Data confidentiality related to kept client's data confidential. If no. of users simultaneously accesses the cloud, then the all data files that are personal to a particular client must be able to hide from other clients on the cloud during accessing of file.

4. SEED BLOCK ALGORITHM ARCHITECTURE

Seed block Algorithm focuses on simplicity of the back-up and recovery process. It uses the concept of Exclusive– OR (XOR) operation of the computing world. For ex: - Take two data files: C and D. When the XOR of the C and D it produced Z i.e. $Z = C \oplus D$. If C data file get destroyed and we want our C data file back then we can get C data file back, by using Exclusive– OR (XOR) operation the data file and Z.

Seed Block Algorithm consists of the Main Cloud and its clients and the Remote backup Server. Here, first Random number in the cloud and unique client id for each client. Then, whenever the client id is register in the main cloud; then client id and random number is getting EXORed ($\oplus$) with each other to generate seed block for the particular client. Whenever client creates the file then it is stored at the main cloud. When it is stored at main server, then the main file of client is being EXORed with the Seed Block of the particular client. And that EXORed file is stored at the remote server in the form of file'.

If either due to some reason files in main cloud crashed / damaged, then the user can get the original file back by EXORing file with the seed block of the corresponding client to produce the original file and return the original file back to the requested client.

SEED BLOCK ALGORITHM

Initialization: Cloud: C; Remote Server: S; Clients of Cloud: Cl; Files: f1 and f'1;
Seed block: Sb; Random Number: r; Client's ID: Cl_Idi
Input: f1 created by Cl; r generated at C;
Output: Recovered file f1 after deletion at C
Authenticated clients could allow performing operations such as downloading, uploading and doing modification on their own the files.
Step1: Generate a random number.

int r = rndno ();

step2: create a seed block Sb for each Cl and store Sb at R   Sb =r$\oplus$ Cl_Idi (Repeat Step2 for all clients)

Step3: if Cl / Admin creates/modifies a f1 and stores at C, then f'1 creates as

f'1= f1 $\oplus$ Sb

Step4: store f'1 at S.
Step5: If server crashes or f1 deleted from C,

Then, we do EXOR to retrieve the original f1 as:

f1=f'1$\oplus$ Sb

Step 6: return f1 to Cl.
Step7: END.

5. CONCLUSION

The Seed Block Algorithm is time efficient technique to recover the file. It maintains the data integrity and solves the issues like cost, implementation complexity. SBA also focuses on the security concept for the back-up files stored at remote backup server, without using the existing encryption techniques.

6. REFERENCES

[1]Ms. Kruti Sharma,Prof. Kavita R Singh, 2013"Seed Block Algorithm: A Remote Smart Data Back-up Technique for Cloud Computing"International Conference on Communication Systems and Network Technologies,2013

[2] Yoichiro Ueno, Noriharu Miyaho, Shuichi Suzuki,Muzai Gakuendai,Inzai-shi, Chiba,Kazuo Ichihara, 2010, "Performance Evaluation of aDisaster Recovery System and Practical Network SystemApplications," Fifth International Conference on Systems andNetworks Communications,pp 256-259.

[3] Chi-won Song, Sungmin Park, Dong-wook Kim, Sooyong Kang,2011, "Parity Cloud Service: A Privacy-Protected Personal DataRecovery Service," International Joint Conference of IEEETrustCom-11/IEEE ICESS-11/FCST-11.

[4] Vijaykumar Javaraiah Brocade Advanced Networks andTelecommunication Systems (ANTS), 2011, "Backup for Cloud and Disaster Recovery for Consumers and SMBs," IEEE 5th InternationalConference, 2011.

BIOGRAPHIES

Kakad Umesh Baban is a final year student of the B.E. computer from PVG COE Nasik, Savitribai Phule Pune University.

Kankhar Mahesh Tukaram is a final year student of the B.E. computer from PVG COE Nasik, Savitribai Phule Pune University.

Mysore Ajay Nagaraj is a final year student of the B.E. computer from PVG COE Nasik, Savitribai Phule Pune University.

Nitin Rathee is a final year student of the B.E. computer from PVG COE Nasik, Savitribai Phule Pune University.