

DATA BACKUP AND RECOVERY METHODOLOGY IN CLOUD ENVIRONMENT

Dr.K.Rajasekaran^{1*}, M.C.A., M.Phil., Ph.D., **P.Nisha M.C.A.²**

¹ Associate Professor, P.G. and Research Department of Computer Science
Dhanraj Baid Jain College (Autonomous), Thoraipakkam,
Chennai, India.

E-mail: krs.salem@gmail.com

²M.Phil. Research Scholar, P.G. and Research Department of Computer Science
Dhanraj Baid Jain College (Autonomous), Thoraipakkam,
Chennai, India

E-mail: nishaarasu1997@gmail.com

ABSTRACT

Data loss can happen at any instant, at any scale. It could be as innocent as unwittingly placing your laptop on a magnet and destroying your hard-disk drive, or a bad storm triggering a power surge and taking data stored on an on-premise backup with it. This is where cloud backup comes in—data preserved off-site, securely. Cloud computing is an application-based software is used to store data on remote serves, which can be access through the internet. Cloud computing service providers always preserve multiple copies of the data to mitigate instances of data loss, security threats, data breach, etc., Cloud back and recovery relate the process of creating and storing copies of data that can be used to protect organization against data loss and recovery help to restore the data to the native location or to another location where it can be used in place of the lost or damaged data. This paper explained one of the backup techniques of seed block algorithm. Using seed black algorithm, we can backup and recover the data easily.

Keywords: Cloud Computing, Cloud Backup, Cloud Recovery, Architecture, Seed Block Algorithm

1. INTRODUCTION

Cloud computing is a way of computing that has as central base sharing computing resources instead of having local servers or personal devices to gives a right of entries to all the applications[10]. The term Cloud refers to Internet or Network. In other hand which is present at remote location. Cloud can allow service over network which means, on public network or on private network, i.e., Wide Area Network (WAN), Local Area Network or Virtual Private Network (VPN). Cloud based-application such as web conferencing, e-mail, customer

relationship management (CRM), all run in cloud. Cloud computing provide different services through the internet. These resources include application and tools like servers, networking, software, database and data storage. Cloud computing is the best choice for people and business for a number of reasons including increased productivity, speed and efficiency, cost saving, performance and security. There are the basic platform of cloud computing services and models working behind the scene. It can make the cloud computing feasible, reliable and accessible to end users. The working models for cloud computing is Development Models and Service Models. Deployment Models is the type of access to the cloud that is how the cloud is located. Cloud can have any of the four types of access: Private, public, Hybrid and Community. Public cloud services allow their services over the internet for a fee on the other hand, private cloud service provide services to a certain number of people. Using both public and private services are known as hybrid option. The Community Cloud provides services and system to be accessible by group of organizations. Service Models is reference models on which the Cloud Computing is based. These can be divided into three basic service models Infrastructure as a Service (IaaS), Software as a Service (SaaS) and Platform as a Service (PaaS).

2. CLOUD BACKUP AND RECOVERY

Cloud backup is also known as online backup. Cloud backup is a type of data backup whereby a copy of the data is sent over a secure public network or proprietary to a cloud-based server. Cloud computing services are usually provided by a third-party vendor who sets the service fee based on scalability, reliability, bandwidth, or number of users. Cloud data backup can be set up to run when requested, thus ensuring minimal data loss. The data is then available from various access points and can be divided among multiple cloud users. Cloud Data Recovery is also called as Cloud disaster recovery (cloud DR). Cloud Data Recovery is a backup and restores procedure that includes putting away and keeping up carbon of electronic records in a cloud computing condition as a safety effort. Data backup and recovery can observe a similar progression. The data can alternatively remain within the customer's network and backed up off site to the customer's choice backup location. The data backup and recovery processes can be conducted in the cloud.

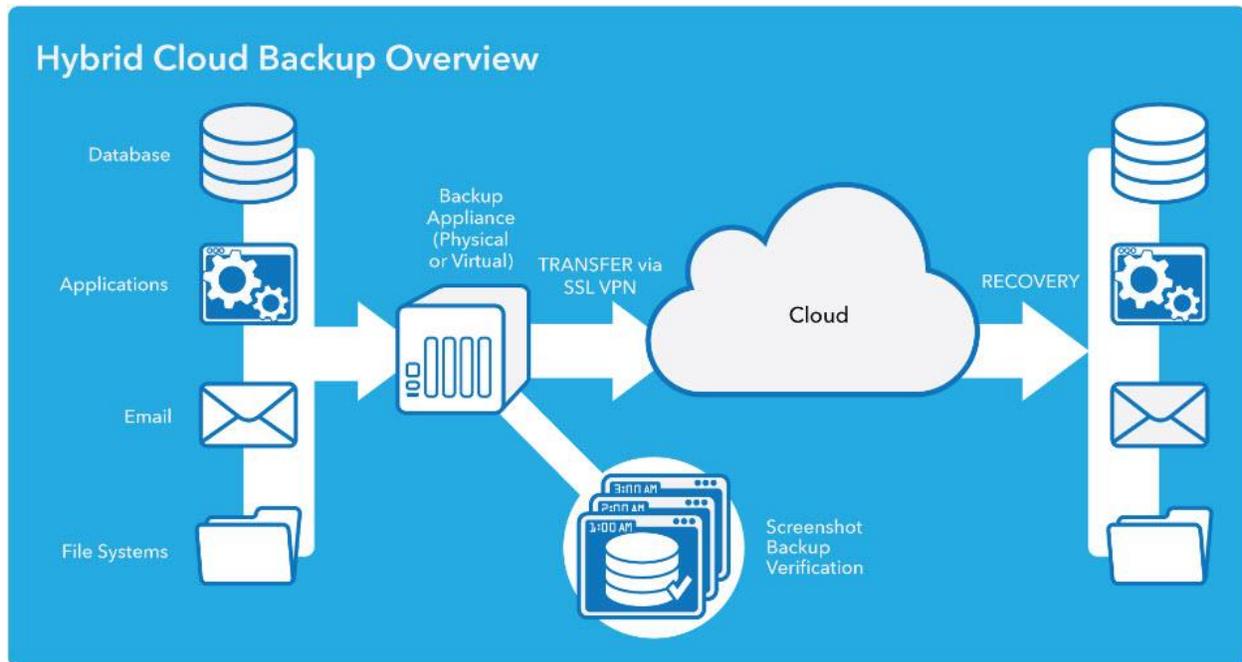


Figure 1. Model of Cloud Backup and Recovery

Cloud Storage

Cloud storage is a model of data storage in which the data can be accessed, conducted, and stored in a remote cloud server via the Internet. Cloud storage is maintained by a cloud storage provider responsible for keeping the user's data information available and accessible at any time. Generally, cloud storage systems share the following characteristics:

1. The cloud storage provider is completely responsible for the back-end support and maintenance of the application.
2. Cloud environments function on a self service basis, meaning that the user or customer can get direct access to cloud-based resources and enjoy the built-in services without involving the service provider.
3. Cloud environments are elastic. Thus, they can be scaled up or down depending on the customer or user needs.
4. Cloud-based resources can be accessed over the Internet at any time and any place.
5. One cloud atmosphere can be shared by multiple users with the help of a multi-tenant model.
6. The cloud storage provider monitors and calculates each user or customer resource use, meaning that you only pay for what you use in a given period of time.

Cloud Storage Types

The following types of cloud storage can be categories:

1. Public cloud storage is essentially a multi-tenant storage environment mostly used for storing unstructured and less sensitive data. Public cloud storage functions as a global data center where computing resources can be preserved and accessed by the general public over the Internet. The peak public cloud storage vendors include Amazon Web Services (AWS), Microsoft Azure, the Google Cloud platform, etc.
2. Private cloud storage is a cloud environment used by one organization exclusively and usually managed via internal resources or by a third-party vendor. Private clouds are designed for management that required full data control, customization, and high-level security. The top most vendors of private cloud storage are VMware, Dell EMC, Hewlett Packard Enterprise (HPE), Open Stack, etc.
3. Hybrid cloud storage represents a combination of public and private cloud storage to form a comprehensive system. In this case, essential data is stored in the private cloud, while less sensitive data is transferred to public cloud storage. To achieve maximum efficiency in a virtual environment, the services of both public and private cloud providers are used.

3. PROPOSED SYSTEM

In proposed system, using cloud server can we recovered data if the system gets physically crash and the data is lost due to some other reason. In today's IT world, it seems like every unique IT headline has something to do with this current technology called "Cloud Server". Cloud computing system, through a variety of computational resources, fundamentally provide access to large amounts of data and interface.

A. Remote Data Backup Server

When we think about Backup server of main cloud, we only talk about the copy of main cloud. The remote location server is termed as Remote Data Backup Server when this Backup server is at remote location that is far away from the main server and having the complete state of the main cloud. The central repository is used to refer main cloud and remote repository is used to refer remote backup cloud.

1) Architecture of Remote Data Backup Server

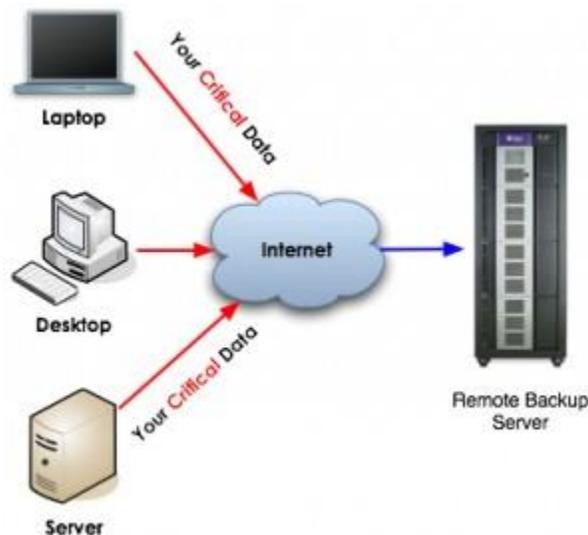


Figure 2. Architecture of Remote Data Server

2) Characteristics of Remote Data Back Server

Data Integrity:

Entire state and the whole structure of the server are concerned in data integrity. During transmission and reception it analysis data such that it remains unaltered. It is the measure of the fidelity and validity of the data present in the server.

Data Confidentiality:

Data Confidentiality is important that client's data files should be kept confidential such that when more than one user are contemporaneous accessing the cloud, then data files that are personal to only specific client must be able to be kept hidden from other clients on the cloud during accessing of file.

Data security:

The maximum priority for the remote server is to give full protection to the client's data. The client's data should not be accessed by the third party or any other users or clients either intentionally or unintentionally.

Cost efficiency:

For the topmost number of company or clients to avail the benefit of back-up and recovery service the cost of process of data recovery should be efficient.

Trustworthiness:

The remote cloud must be believable. Because the user or client stores their private data; therefore the cloud and remote backup cloud must play a trustworthy capacity to protect the private data of user or client.

B. Seed Block Algorithm (SBA) Architecture

The purpose of this algorithm is to focus on simplicity of the back-up and recovery process. This algorithm it used the impression of Exclusive OR (XOR) operation of the computing world. For example: Suppose there are two data files A and B. When we XOR A and B it produced C. If we want our A data file back which was destroyed then we are able to get A data file back, it is very easy to get back it with the help of B and C data file. The Seed Block Algorithm works to allow the simple Back-up and recovery process. First we set a random number in the cloud and different client id for every client. Second, whenever the client id is being listed in the primary cloud then client id and random number is getting EXORed () with each other to generate seed block for the particular client. The generated seed block corresponds to each client is preserve at remote server. When client discover the file in cloud first time, it is stored at the primary cloud. When it is preserve in primary server, the main file of client is being EXORed with the Seed Block of the specific client. And that EXORed file is preserved at the remote server in the form of file. If either unfortunately file in main cloud crashed or file is been deleted mistakenly, then the user will get the original file by EXORing file with the seed block of the corresponding client to produce the original file and return the resulted file that is original file back to the requested client.

Architecture of Seed Block Algorithm (SBA)

SBA Algorithm [6]:

Initialization: Main Cloud (Mc): Remote Server: (Rs) ;

Clients of initial Cloud (Ci) ; Files(a1) and (a1');

Seed block: (Si); Random Number: (r);

Client's ID: (Client_Idi)

Input: (a1) produced by (Ci); r is generated at (Mc);

Output: Recovered file a1 after deletion at Mc

Given: Authenticated clients could allow uploading, downloading and do alternation on its own the files only.

step 1: Generate a random number.

Int r= rand O

step 2: Create a seed Block (SB) Si for each Ci and Store

Si at Rs

Si= r \oplus Client_Idi (Repeat step 2 for all clients)

step 3: If Ci/Admin creates /alter a a1 and stores

at Mc,

then a1' create as

a1'=a1 \oplus Si

step 4: Store a1'at Si

step 5: If server damage a1 deleted from Mc,

then, we do EXOR to rectify the original as a1:

a1=a1' \oplus Si

step 6: Return a1 to C1.

step 7: END

Whenever client creates the file in cloud first time, it is stored at the initial cloud (e.g. 010). It is preserved in main server. The main file of client is being EXORed with the Seed Block of the specific client. And that EXORed file (100) is stored at the remote server in the form of file' (pronounced as File dash). If either unfortunately file in main cloud crashed or damaged or file is been deleted mistakenly, then the user will get the original file by EXORing file' with the seed block of the corresponding client to produce the original file (010) and return the resulted file i.e. original file back to the requested client. It is found that size of original data file stored at main cloud is exactly similar to the size of Back-up file stored at Remote Server.

CONCLUSION

In this paper, study of remote data backup algorithm is also called seed block algorithm is done which helps the user to recover the disaster files from the remote location when the main cloud fails to deliver the files to the client. We have presented detailed design of proposed Seed Block Algorithm (SBA) and also called as smart data backup techniques. Proposed Seed Block algorithm (SBA) is robust in assisting the users to collect data from any remote location in the loss of internet connection and if file deletion occurs due to any reason or unluckily cloud gets destroyed till we can also recover files. The proposed Seed Block Algorithm (SBA) will take minimum time for process recovery so that the problem corresponds to time can be solved.

REFERENCE

- [1] Karishma Nadhe and Sushma Somani, "Backup of real time data and Recovery using cloud computing", *International Journal of Engineering Development and Research*, Vol. 4, 2016.
- [2] Monisha and Venkateshkumar, *Cloud Computing in Data Backup and Data Recovery*, *International Journal for Trend in Scientific Research and Development*, Vol. 2, Sep-Oct 2018.
- [3] R.V. Gandhi and M. Sessaiah, *Data Back-Up and Recovery Techniques for Cloud server Using Seed Block Algorithm*, *International Journal of Engineering Research and Application*, Vol.5, Feb 2015.
- [4] Somesh P Badhel and Vikrant chole, *A Review on Data Back-up Techniques for Cloud Computing*, *International Journal of Computer Science and Mobile Computing*, Vol. 3, Dec 2014.
- [5] K.Lakmi and V. Supriya, *Data Backup and Recovery Techniques in Cloud Computing*, *International Journal of Scientific Research in Computer Science*, Vol. 3, 2018.
- [6] Yashodha Sambrani and Dr. Rajashekarappa, *Efficient Data Backup Mechanism for cloud Computing*, *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 5, July 2016.
- [7] Yogesh Gite and Ankush Pawar, *Efficient Data Backup Technique for Cloud Storage*, *International Journal of Engineering Research in Computer Science and Engineering*, Vol. 5, Mar 2018.
- [8] Mr. Akshay A.Gharat, *Disaster Recovery in cloud Computing*, *International Journal of Advance Research in Computer Engineering and Technology*, Vol. 4, May 2015.
- [9] Mr. G.S. Narke and Mr. A. K. Harijan, *A Smart data technique for cloud computing using seed block algorithm strategy*, *International Research Journal of Engineering and Technology*, Vol.2, Sep 2015.

[10] R. Panner Selvam, Dr.K.Rajasekaran, "A SURVEY ON CLOUD COMPUTING IN BUSINESS DEVELOPMENT", *International Journal of Management, Technology and Engineering*, Volume IX, Issue VI, ISSN NO: 2249-7455, Pp No. 105-116. August 2019.

[11] Mr. A. Srinivas, Y.Seetha Ramayya, *A Study on Cloud Computing Disaster Recovery*, *International Journal of Innovative Research in Computer and Communication Engineering*, Vol. 1, Aug 2013.

[12] Kolipaka Kiran, Janapati Venkata Krishna, *Smart Data Back-up Technique for Cloud Computing using Secure Erasure Coding*, *International Journal of Computer Trends and Technology*, Vol. 16, Oct 2014.

[13] Ms. Ashwini S. Gharde, Ms. Kamini Ghaormare, *Study of Seed Block Algorithm in Cloud Computing Environment*, *IOSR Journal of Computer Engineering (IOSR-JCE)* e-ISSN: 2278-0661,p-ISSN: 2278-8727 PP 37-39 2016.