

COMPARATIVE ANALYSIS OF DISASTER RECOVERY SOLUTIONS IN CLOUD COMPUTING

Mr Tejas Abhimanyu Patil¹ & Ms. Vaidehi K. Gadewar²

^{1,2}Dept. of MCA, Sterling Institute of Management Studies, Nerul, Navi Mumbai, Maharashtra, India

Abstract: Cloud computing provides accessing of any kind of services dynamically over Internet on demand basis. One of the most significant service that is being provided is storage as a service. Nowadays cloud computing has easily reached on many top lists because of its far-reaching implications in many areas in computing, especially big data which without cloud computing is at best a concept. Cloud customer can store any amount of data into cloud storage results to huge amount of data at the data center. Nowadays, data has been generated in large quantity that requires the data recovery services or techniques. Therefore, there is a requirement for designing an efficient data recovery technique to recover the lost data.

Keywords: Cloud Computing, Data Recovery, Replication, Backup.

INTRODUCTION:

Cloud Computing is a buzzword that has many different meanings to many different people. Cloud computing is based on virtualization that has another perspective for disaster recovery. With virtualization, the entire server, including the operating system, applications, patches and data is encapsulated into a single software bundle or virtual server. One can easily take the entire backup or a copy of a data and host it on a server. Nowadays, data has been generated in large amount that required the data recovery services. This type of computing will generate a large amount of private data on main cloud. Therefore, the necessity of data recovery services is growing day-by-day and it requires a development of an efficient and effective data recovery technique. The main purpose of recovery technique is to help user to collect information from any backup server when the data is lost and unable to provide data to the user.

Today many different techniques have been proposed in the market till date. The main objective of this research paper is to summarize the powerful data recovery techniques that are used in cloud computing domain.

Data Recovery in Cloud Computing: Today data storage in data centre is increasing rapidly day-by-day, hence resulting into large amount of data loss, copying of same data, data breach, etc. To avoid these problems there is immense need of finding of some new solution to avoid this kind of data loss in near future. We need to find out the proper technique to recover the data if the gets deleted or destroyed. Loss of data is very harmful to the organizational level, as it can cause huge financial loss to the company's business. By using some of the data recovery techniques the original data can be recovered. Hence, to recover the loss data, efficient technique is needed to get the loss data back which should be reliable and efficient enough.

PROBLEM DEFINITION

Causes of Data Loss:

A. Natural Disasters

Nowadays many businesses are totally dependent on electronic devices or electronic data there after suffers huge temporary data loss for the company.

This natural disasters like earthquake, tsunami, etc causes the company a major setback and make them responsible or realize the importance of data backup and security. According to the survey, when natural disasters occurs then 2% of the data is lost. The main reasons of occurrence natural disasters are global warming. Due to mundane and nefarious effects one cannot recognize the data loss when disaster occurs.

B. Critical application failure

It is natural, when an application is left unusable for many days then it causes a great damage failure to the organizations. This can cause some critical application failure leading to some important data loss. By using all applications that are stored in cloud may reduce the sudden great damage.

C. Network failure

In common man words cloud server and the clients are connected through internet connection. If the connection fails, then the services which are connected to internet are crashed and data gets lost and also applications which are based on cloud also gets affected.

D. Network trespassing(intrusion)

When an virus is invaded into our system then there are high chances of data loss. By placing unusable applications in that place on a watch list we can prevent occurrence of disaster

E. Hacking or malicious code

Nowadays, hacking is the most common thing happening in any field. Disaster is the most unpredictable thing one cannot say anything about it. So if any kind of disasters occurs in the system that may be inside or outside of the company, the organization can prevent these kind of hacking or malicious code from modifying data then there is a loss of data.

F. System failure

As the name itself suggests, if any kind of infrastructure in an organization fails then whole systems which are connected in that organization will crash. Since in any organization, all the systems are connected to one main server/system. If this system fails, then the entire connection connected to that system gets crash. This will affect the operating systems. The main reason why disaster occurs is due to human beings. Almost 60%-70% of data centers are failed.

Disaster Recovery Solutions:

Some recovery solutions that have been proposed to overcome the issues mention above:

A. Local Backup

The dependency issue is tried to be resolved in. Data can be controlled and complete application can be backed up by arranging the Linux box on customers side. This approach may help to migrate private services to public and vice versa. It might also merge the cloud service providers. By this method, local backup can furnish with various services that service provider replenishes during any disasters or issues.

B. Geographical Redundancy and Backup (GRB)

In traditional model geographical redundancy can also be used. Using this model can be costly and cannot be afforded. Two cloud zones are similar to each other. If any of the zone among both becomes down, then another zone will be on and provide the services

C. Inter-Private Cloud Storage (IPCS)

The proposed approach for storage of cloud data. For business data storage at least three backup locations are necessary according to the Storage Networking Industry Association (SNIA). The Data of the user should be kept or stored in three different locations which are Servers, Local backup server(LBS) and remote backup server (RBS). This model helps in increasing data integration

D. Resource Management

Several hardware and software comprising of hybrid storage and diverse disks are abode in Heterogeneous clouds. The cloud storage hoards complete business data in cloud-based enterprise. Thus data protection, security and recovery play significant roles in these environments. Data that is inscribed into primary host stays at risk unless it directed or processed into the backup host. Hence, use of advanced technology is important for recovering the information in storage clouds. In 2012, has three major solutions for data recovery have been proposed:

- During disasters, use fastest disk technology for replicating the data in danger.
- Changing dirty page threshold: The amount of faulty pages in RAM that have to wait to flush into the disk might get faster (Rudolph, 1990).
- Prediction and alteration of risky devices: Few essential elements like power intake, heat dispersion, carbon credit utilization and importance of data (hoarded on every single disk) can be figured out in definite period of time.

E. Scale Up/Down

At times, conducting functions with greater supremacy can lessen the money loss or even increase the earning during disasters. Cloud service providers were flooded with service requests after a natural disaster took place in an area. In such situations, service providers are compelled to manage the existing users and take care of new requests as well. Service providers are bound to appease the existent users and provide satisfactory service to new customers. During natural calamities (for e.g. Earthquake) the system avails DR scenario scaling up resources for the high-priority services (for e.g., voice communication) and scaling down allocated resources to low-priority service (for e.g., video on-demand).

Data Recovery Technique in Cloud Computing

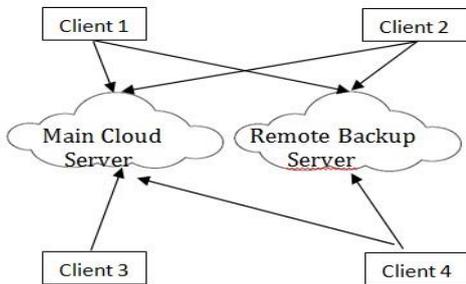


Figure 1 – Data Recovery Backup

Architecture of remote server

The designed architecture includes three different modules such as:

1. Remote Backup Server
2. Main Cloud Server
3. Number of Clients/users

Remote Backup server preserves the replicated copies of main server. It is known as remote repository. The central server, known as the central repository accommodates all users' data. The user uploads the document in main cloud server which is then stored/transferred into the backup server. If the user demands to restore the document from the cloud, it is initially chased in the main cloud server. If it is not available in the main cloud server, then the user goes through the backup server to recover lost data. Sometimes natural calamities or manmade disasters in main servers cause the loss of data or crashes. A recovery technique is essential to retrieve the lost data. This objective can be obtained by using the proposed algorithm adequately. To grant the dependability two or four backup data cloud storages can be used. The Figure 2 explains the system architecture consuming four backup cloud storages.

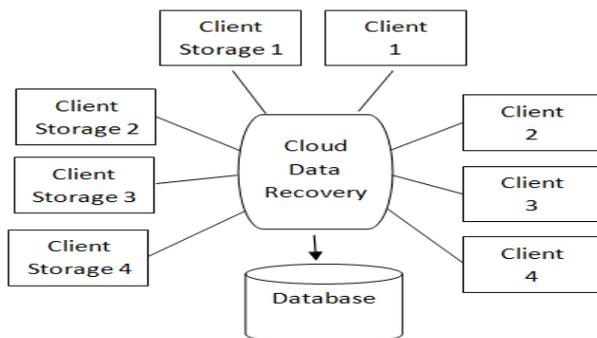


Figure 2 – Data Recovery Technique

SYSTEM ARCHITECTURE

The system architecture illustrates four backup servers. The data of the replicated copies is maintained in more than one server to recover data. When the loss of data occurs at one location then it can be retrieved or get from the other backup maintained in the system.

ANALYSIS AND FINDINGS

Comparison between the Recovery Solutions

A. Local Backup

- **Security**

It may be more appropriate to call this advantage peace of mind. With local backup, which is usually in the form of external hard drives, the data is just as protected as your network. Plus, once it's stored and the drive is disconnected, it's safe from any malicious attacks that affect your infrastructure.

- **Speed**

With on-site backups, speed is not limited by connectivity. Backing up all of your data to an external hard drive usually takes a fraction of the time required by cloud storage. After that initial backup, daily updating can be done in a few minutes.

- **Control**

With a local backup solution, you know exactly where your data is, and you retain control over who can and cannot access your files. Keeping an offline backup locally is a good way to ensure that your information doesn't fall into the wrong hands.

B. Geographical Redundancy and Backup

- **Security**

As more businesses move critical applications and systems to the cloud, the impact a natural disaster or an unexpected outage can have on your business operations is greater today. Even small outages can have long term business impacts. It ensures high availability of business critical systems across multiple locations, mitigating the risk of weather outages.

- **Control**

Businesses can mitigate downtime by replicating applications and data across multiple 'geo-diverse' locations. Also termed as 'geo-replication', the data

that is created or updated in a primary location, is asynchronously replicated to a secondary location so that the same data exists and is readily accessible in both locations.

C. Inter-Private Cloud Storage

- **Control**

Private cloud storage works much like public cloud storage and implements storage virtualization across an organization, providing a centralized storage infrastructure that can only be accessed by the authorized nodes. Private cloud storage operates by installing a data centre, which houses a series of storage clusters that are integrated with a storage virtualization application. Administrative policies and a management console provide access to the different storage nodes and applications within the organization's network. The applications or nodes access the private storage through file access and data retrieving protocols, while the automated storage administrator application allocates storage capacity to them on run time.

D. Resource Management

- **Challenges**

Cloud delivers computing as a utility as it is available to the cloud consumers on demand. It is a simple pay-per-use consumer-provider service model. It contains large number of shared resources. So Resource Management is always a major issue in cloud computing like any other computing paradigm. Due to the availability of finite resources it is very challenging for cloud providers to provide all the requested resources. From the cloud providers perspective cloud resources must be allocated in a fair and efficient manner.

Distinguishing between cloud storage and cloud computing

Cloud computing requires higher processing power than cloud storage. Cloud storage, on the other hand, needs more storage space.

Cloud computing is essentially targeted towards businesses. Cloud storage, on the other hand, is utilized both for professional and personal reasons.

Cloud storage is simply a data storage and sharing medium, while cloud computing gives you the ability to remotely work on and transform data (for example, coding an application remotely).

CONCLUSION:

At present it is very important to all the big or small organization to store large amount of data in the cloud. A thorough literature survey is presented in this paper. The easy way of recovery technique for recovering the deleted data is discussed in detail. The four backup servers concept is used to recover the deleted data. The Proposed method provides the flexibility for the user to recover their lost data from any of the four backup servers.

LIMITATIONS:

In this research we have compare 4 recovery solutions out of 5 which still have to be taken into comparison.

To improve more accuracy in data recovery further work can be done using different solution techniques.

REFERENCES:

- I. Efficient and Reliable Data Recovery Technique in Cloud Computing Published: August 25, 2017 Internet of Things and Cloud Computing
- II. Chintureena Thingom, "A Study on Tools for Cloud Disaster Management", International Journal of Interdisciplinary and Multidisciplinary Studies, 2014.
- III. Jian-hua, Z., & Nan, Z. (2011). Cloud Computing-based Data Storage and Disaster Recovery. IEEE International Conference on Future Computer Science and Education (ICFCSE), (pp. 629-632). <http://dx.doi.org/10.1109/ICFCSE.2011.157>
- IV. Kashiwazaki, H. (2012). Practical uses of cloud computing services in a Japanese university of the arts against aftermath of the 2011 Tohoku earthquake.
- V. Patil, S. R., Shiraguppi, R. M., Jain, B. P., & Eda, S. (2012). Methodology for Usage of Emerging Disk to Ameliorate Hybrid Storage Clouds. IEEE International Conference on Cloud Computing in Emerging Markets (CCEM) (pp.1-5). <http://dx.doi.org/10.1109/CCEM.2012.6354615>
- VI. Patterson, D. A., Gibson, G., & Katz, R. H. (1988). A Case for Redundant Arrays of Inexpensive Disks (RAID). International Conference of Management of Data (SIGMOD) (pp.109-116). Pokharel, M., Lee, S., &
- VII.
- VIII.
- IX.
- X.
- XI.
- XII.