

Various Effective Performance Methods In Backup And Data Recovery Techniques For A Cloud Based Environment

¹R.Viswanathan, ²T.Edison, ³J.Vaishnavi, ⁴Dr.K.Kalaivani.

^{1,2,3,4}Professor at school of computer science in Vels University Chennai

rvnathan06@gmail.com, edison.gis@gmail.com, vaishnavi13jay@gmail.com, kalaivani.se@velsuniv.ac.in

Abstract— Cloud computing provides various kinds of services to its users. Storage-as-a-service is one of them, it is a service provided by cloud infrastructure in which large amount of electronic data is stored in cloud. In remote location on cloud we must always assure that our valuable data must be safe and it can be available in any time. In situations like natural calamities or any hardware crash or any unintended crossing out our data might be no longer to stay available. To recover data in cloud environment we are using data backup techniques for efficient use of data. This paper provides a review on various backup techniques used for Cloud computing platform.

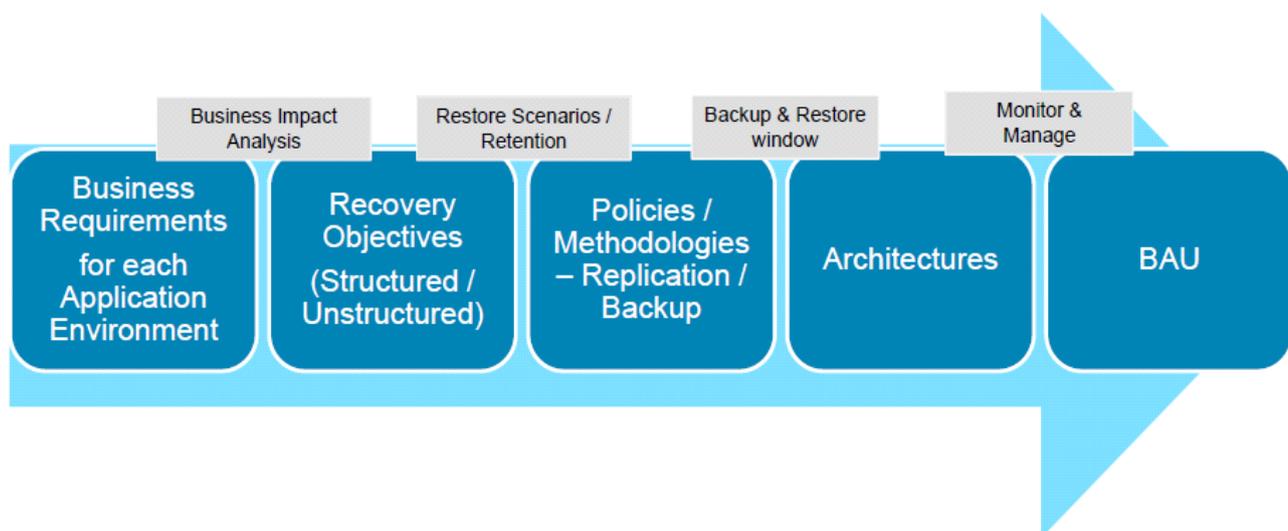
Keywords— Cloud Computing, Load Balancing, Quality of Service (QOS) Backup; Recovery; Incremental backup; Differential backup; Progressive incremental backup; RTO; RPO

1. INTRODUCTION

Cloud computing is an established prototype for on-demand computing services to a wide range of clients, as well as enterprises, software developers and researchers. Infrastructure Providers (IPs) manage the base infrastructure including servers, storage and network connectivity, and typically present this infrastructure as Virtual Machines (VMs). Other providers rent these resources and resell value-added services as Platform as a Service (PaaS) or Software as a Service (SaaS). Amazon web services assets to analyze compliance levels and security. It also enables an administrator to troubleshoot why a resource stopped working properly. Amazon Web Services resources have been analyzed according to best practices into four categories:

- Cost optimization
- Security
- Fault tolerance
- Performance improvement.

With the advent of server technology, the concept of data sharing and application usage has been affected the most. Companies have started to adopt the concept of cloud and have brought a revolutionary change in the way to make business decisions or provide business solution. There has been a significant increase in the data generation, about 1.7 megabytes of data is generated every second for every human on the planet. To suffice the needs of data processing and storage, service providers implement high end dedicated hardware and data centers. With increase in data the need of securing and backing up data has also risen. Data backup is one of the least preferred but the most essential part for an IT firm. If a firm does not invest on data backup, chances are high that the firm won't be able to recover its own or customer's data and thus will suffer huge loss. Each cloud deployment model has a backup database and backs up the data according to the services it provides and the essentiality of the data that has to be backed up.



2. TYPES OF BACKUPS

• **Incremental:** Full backup is preferably initiated on a lean business day or on an off peak time. On the consecutive days only the data that has been modified is

backed up. Each day the data that has been modified on the earlier day is backed up. It is suitable for business whose RTO(recovery time objective) is high.

• **Differential:** Full backups are performed on a lean business day while on the consecutive days the data that

has been modified is backed up. On every 3rd day all the data that has been backed up in the consecutive days is consolidated together. It is suitable for the business who have less RPO and RTO and can spend more on storage.

- **Progressive Incremental:** It is similar to incremental backups but here on the 6th day all the data that has been backed up till now is consolidated and a full backup set is prepared. It is a flat file system backup option.

Recovery

Recovery means to recover the copy of original data that has been backed up on some remote backup server, after a failure is encountered., Recovery is the 2nd phase of the data disaster mitigation process, after backing up data retrieval of data becomes a major challenge. To keep a business running moderate amount of data(Essential data) is necessary, without it the business will collapse. Recovery ensures the running of a business even after a disaster has occurred.

The recovery of any backup is done according to its RTO(recovery time objective), if RTO is very less than the availability of the hardware should be high.

The essential data of the business should be backed up first so that the business keeps running after that all the remnant data is recovered.

Reliability

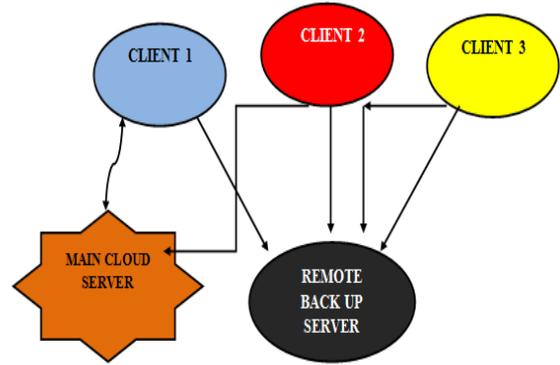
The remote cloud must possess the reliability characteristics. Because in cloud computing the main cloud stores the complete data and each client is dependent on the main cloud for each and every little amount of data; therefore the cloud and remote backup cloud must play a trustworthy role. That means, both the server must be able to provide the data to the client immediately whenever they required either from main cloud or remote server.

Relocation Of Server

For data recovery there must be relocation of server to the cloud. The Relocation of server means to transfer main server’s data to another server; however the new of location is unknown to the client. The clients get the data in same way as before without any intimation of relocation of main server, such that it provides the location transparency of relocated server to the clients and other third party while data is been shifted to remote server.

3.PROPOSED DATA RECOVERY TECHNIQUE IN CLOUD COMPUTING

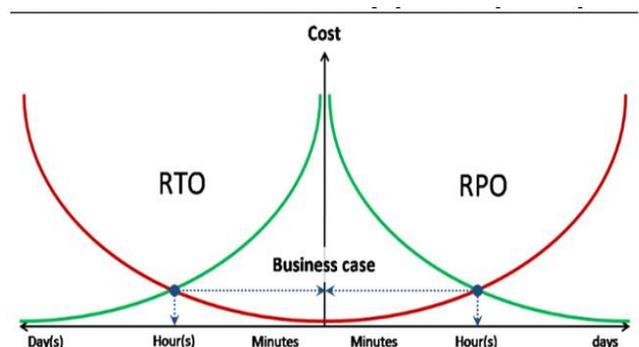
1. Remote Backup Server 2. Main Cloud Server 3. Number of Clients/users. Remote Backup server maintains the replicated copies of main server and is called as remote repository. The main server is called as central repository it stores all the user data. The user uploads the file to main cloud server; the main cloud server stores all the data in backup server. If user wants to retrieve the file from cloud then file is searched in main cloud server firstly, if the data is not present in main server then it is checks in backup



Implementation

In cloud the type of backup implemented is generally a use case of 2 or more types combined, to ensure a safe and complete restoration of original data. Some IT firms also provide DRaaS- Disaster recovery as a Service, that is the client can pay a 3rd party vendor to backup and secure his data. In DRaaS the customer agrees to SLA of 3rd party vendor for his data backup and security services all the upkeep and maintenance of the data centre is done by the provider, the client does not has to worry about anything. Other methods are use of mirroring and geo-redundancy where a backup server is used to store the copy of original data. The mirroring servers are decided on the basis of the location of client to provide him the fastest recovery time objective, with the lowest latency. Backup can also be done on distributed data servers rather than having all the backup under single roof and the recovery can be initiated and with the help of a wide bandwidth network data can be recovered parallel at a more quicker pace.

Cloud creators also provide backup on the basis of services they provide, say a SaaS provider will focus on user data while IaaS will spend least in any kind of user data backup and will provide user the backup of allotted hardware for their VM.A PaaS provider will always backup the complete snapshot of the ongoing development of the software. Different providers might use different approaches to deal with data backup problem. In general backups are made on the most cheapest and durable storage media, Tapes are one of the best options to store or archive data and retrieve it when required. Magnetic tapes are used in either NAS architecture or SAN architecture making it possible for a quick backup and recovery.



Recovery

Recovery means to recover the copy of original data that has been backed up on some remote backup server, after a failure is encountered. Recovery is the 2nd phase of the data disaster mitigation process, after backing up data retrieval of data becomes a major challenge. To keep a business running moderate amount of data(Essential data) is necessary, without it the business will collapse. Recovery ensures the running of a business even after a disaster has occurred. The recovery of any backup is done according to its RTO(recovery time objective), if RTO is very less than the availability of the hardware should be high. The essential data of the business should be backed up first so that the business keeps running after that all the remnant data is recovered.

4.CONCLUSION

A thorough literature survey is presented in this paper. The efficient recovery technique for recovering the deleted data is discussed in detail. Result and analysis section shows that the proposed way to handle the cloud server efficient and reliable. The four backup servers concept is used to recover the deleted data. Proposed method provides the flexibility for the user to recover their data from any server among four backup servers. The time related issues are being solved by proposed RTO such that it will take minimum time for the recovery process. Many reputed companies can store the confidential data without much risk, Security is also being increased.

5. REFERENCES

- [1] The Disaster Recovery Handbook: A Step-by-Step Plan to Ensure Business Continuity and Protect Vital Operations, Facilities, and Assets by Michael Wallace, 2010.
- [2] Hadoop Backup and Recovery Solutions by Gaurav Barot, 2015.
- [3] Chi-won Song Sungmin Park Dong-wook Kim Sooyong Kang "Parity Cloud Service: A Privacy-Protected Personal Data Recovery Service" International Joint Conference of IEEE Trust Com-11/IEEE ICES-11/FCST-11 2011.
- [4] Vijaykumar Javaraiah Brocade Advanced Networks and Telecommunication Systems (ANTS) "Backup for Cloud and Disaster Recovery for Consumers and SMBs" IEEE 5th International Conference 2011.
- [5] Andrea Bianco Jorge Finochietto Luca Giraud Marco Modesti Fabio Neri Network Planning for Disaster Recovery IEEE September 2009.
- [6] Glen Robinson Ianni Vamvadelis Attila Narin Using Amazon Web Services for Disaster Recovery <http://media.amazonwebservices.com/AWS-Disaster-Recovery.pdf> January 2012.
- [7] T. Wood E Cecchet K. K. Ramakrishnan P. Shenoy J. van der Merwe and A. Venkataramani "Disaster recovery as a cloud service: economic benefits & deployment challenges" Proc. 2nd USENIX Conference on Hot topics in cloud computing (Hot Cloud'10) Berkeley CA USA 2010 pp. 8-8.
- [8] N Vedashree KC Praveen Kumar G Anilkumar Data Recovery in Cloud Environment Using Seed Block Algorithm.
- [9] Kruti Sharma Kavita R Singh "Online Data Back-up and Disaster Recovery Techniques in Cloud Computing A Review" in IEEE Paper Nagpur (M.S) 441 110 India:Computer Science Engineering YCCE 2012.
- [10] R. V. Gandhi M Seshaiyah A. Srinivas C. ReddiNeelima "Data Back-Up and Recovery Techniques for Cloud Server Using Seed Block Algorithm" IEEE paper February 2015.
- [11] Vijay kumar Javaraiah "Backup for Cloud and Disaster Recovery for Consumers and SMBs" IEEE 5th International Conference 2011.
- [12] S. Zhang, X. Chen, and X. Huo, 2010, "Cloud Computing Research and Development Trend," IEEE Second International Conference on Future Networks, pp. 93-97.
- [13] M.Armbrust et al, "Above the clouds: A berkeley view of cloud computing," <http://www.eecs.berkeley.edu/Pubs/TechRpts/2009//EECS-2009-28.pdf>.