ISSN 2395-1621

Seed Block Algorithm:A smart technique for Data Back-up and Recovery in Cloud Computing

[#1]Ms. Mayuri Tidke,[#2]Ms. Vijayshree Jadhav, [#3]Ms. Sonali Parab, [#4]Ms. Shubhrata Patil, [#5]Prof. Y. K. Patil

[#1234]B.E Students JSPM's BSIOTR, PUNE

[#5]Assistant Prof. in Computer dept. JSPM's BSIOTR, PUNE

ABSTRACT

In cloud computing the data is stored in the cloud server and stored data is very sensitive. It belongs to different fields like social network and medical science so it is important to store data securely in the main cloud. If in case data gets deleted due to any natural calamity or due to human mistake then it is important to retrieve these data. We proposed SBA a smart technique for data backup and recovery. The data is stored securely on the remote server using the seed block algorithm. Also it checks the user is authenticated person or not. The two main objectives of proposed system are first, to help the user to collect information from any remote location in the absence of network connectivity and second, to recover the files in case if file gets destroyed due to any reason.

Keywords: ***Central Repository; Remote Repository; Parity Cloud Service; Seed Block;***

ARTICLE INFO

Article History

Received:20th October 2015

Received in revised form :

22nd October 2015

Accepted:23rd October, 2015

Published online :

24th October 2015

I. INTRODUCTION

On-demand network access to a share pool of configurable computing service, National Institute of Standard and Technology defines as a model for enabling convenient that can released with minimal management effort or services provider and provisioned rapidly. The cloud computing is also gigantic technology which is surpassing all the previous technology of computing like cluster, grid, etc. To overcome the disadvantages of previous computing techniques, the need of cloud computing is increasing now a days due to its advantages. Cloud computing provide us the online data storage where data stored in form of virtualized pool that is usually hosted by third parties. On large data center the hosting company operates large data 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.

There are number of user share the resources and storage. So that it is possible that other customers can access your data. Our cloud storage may be in risk and danger due to either human errors faulty equipment's, a bug, any criminal intent and network connectivity. And the changes in the cloud may also be made frequently, which is also called as data dynamics. The data dynamics is supported by various operations such as insertion, deletion and block modification. Remote data integrity is needed because, to archiving and taking backup of data, the services are not limited. Because, the data integrity always focus on validity and fidelity of the complete state of the server that takes care of the heavily generated data which remains unchanged during storing at main cloud remote server and transmission. For back-up and recovery services integrity is very important.

In literature many techniques have been proposed, which are HSDRT[1], PCS[2], ERGOT[4], Linux Box [5], Cold/Hot backup strategy [6] etc. These all techniques discussed the data recovery process. However, behind the various successful techniques, there are same critical issue like, implementation complexity, low cost, security and time related issues. To fulfil these issues, in this paper we propose a smart remote data backup algorithm, Seed Block Algorithm (SBA). There are two objectives of our proposed system, first it help the users to collect information from any remote location in the absence of network connectivity and second to recover the files in case of the file deletion or if the cloud gets destroyed due to any reason.

II. LITERATURE REVIEW

Performance Evaluation of a Disaster Recovery System and Practical Network System Applications

In this paper author presents evaluation results for a high security disaster recovery system using distribution and rake technology. In an experimental evaluation, the encryption and spatial scrambling performance and the average response time have been estimated in terms of the data file size. Discussion is also provided on an effective shuffling algorithm to determine the dispersed location sites. Finally, this paper describes a prototype system configuration for several practical network applications, including the hybrid utilization of cloud computing facilities and environments which are already commercialized.

Parity Cloud Service: A Privacy-Protected Personal Data Recovery Service

In this paper author said that, as more and more data are generated in an electronic format, the necessity of data recovery service became larger and the development of more efficient data backup and recovery technology has been an important issue during the past decade. While lots of effective backup and recovery technologies, including data dedeplication and incremental backup, have been developed for enterprise level data backup service, few works have been done for efficient personal data recovery service. Since the privacy protection is a crucial issue for providing a personal data recovery service, a plain data backup-based recovery service is not adequate for public service. Users are not expected to upload their critical data to the internet backup server until they can fully trust the service provider in terms of the privacy protection. In this paper, we propose a novel data recovery service framework on cloud infrastructure, a Parity Cloud Service (PCS) that provides a privacy-protected personal data recovery service. The proposed framework does not require any user data to be uploaded to the server for data recovery. Also the necessary server-side resources for providing the service are within a reasonable bound.

Disaster Recovery Mechanism using Widely Distributed Networking and Secure Metadata Handling Technology

In this paper author said that, an innovative mechanism of a file-backup system concept is discussed. In the proposed file backup mechanism, the combination of the following technologies such as a spatial random scrambling of file data, a subsequent random fragmentation of the file, the corresponding encryption and duplication for each fragmented one by using a stream cipher in each encryption stage, and the corresponding notification of the history data of the used encryption key code sequence which we call metadata in addition can effectively realize the highly secure and prompt file backup system economically only when they are combined at the same time. In case of disaster occurrences in data backup center, the prompt data recovery can be easily and securely achieved by making use of a widely distributed great amount of PCs or cellular phones via several supervisory servers which are secretly diversified and functionally combined mutually. This paper proposes the above mentioned state-of-the art hybrid disaster recovery mechanism and its basic characteristics.

ERGOT: A Semantic-based System for Service Discovery in Distributed Infrastructures

In this paper author said that, the increasing number of available online services demands distributed architectures to promote scalability as well as semantics to enable their precise and efficient retrieval. Two common approaches toward this goal are Semantic Overlay Networks (SONs) and Distributed Hash Tables (DHTs) with semantic extensions. This paper presents ERGOT, a system that combines DHTs and SONs to enable semantic-based service discovery in distributed infrastructures such as Grids and Clouds. ERGOT takes advantage of semantic annotations that enrich service specifications in two ways: (i) services are advertised in the DHT on the basis of their annotations, thus allowing to establish a SON among service providers, (ii) annotations enable semantic-based service matchmaking, using a novel similarity measure between service requests and descriptions. Experimental evaluations confirmed the efficiency of ERGOT in terms of accuracy of search and network traffic.

III.PROPOSED SYSTEM ARCHITECTURE

Many techniques have been proposed for recovery and backup which are HSDRT[1], PCS[2], ERGOT[4], Linux Box[5], Cold/Hot backup strategy[6] etc. Low implementation complexity, low cost, security and time related issues are still challenging in the field of cloud computing. To overcome these issues we propose SBA algorithm and remote data backup server to recover the data.

3.1 REMOTE DATA BACKUP SERVER

We think, Backup server is copy of main cloud. When this Backup server is at far away from the main server and having the complete state of the main cloud, then this remote location server is called as Remote Data Backup Server. The main cloud is called as the central repository and remote backup cloud is called as remote repository.

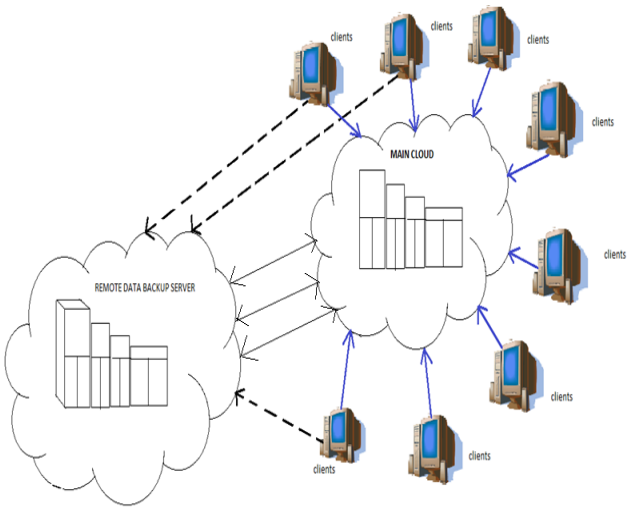

Fig 1: Remote data Backup Server and its Architecture

When the central repository lost its data under any circumstances either of any natural calamity or by human attack or deletion which has been done mistakenly and then it uses the information from the remote repository. The main purpose of the remote backup facility is to help user to collect information from any remote location even if network connectivity is not available or if data not found on main cloud. As shown in Fig-1 If the data is not found on central repository clients are allowed to access the files from remote repository (i.e. indirectly). The Remote backup services should cover the following issues:

1)Data Integrity

Data Integrity is concerned with complete state and the whole structure of the server. It verifies that data such that it does not changed during transmission and reception. It is the measure of the validity and fidelity of the data present in the server.

2) Data security

Giving full protection to the client's data is also the utmost priority for the remote server. And either intentionally or unintentionally, it should be not able to access by third party or any other users/client's.

3) Data Confidentiality

Sometimes client's data files should be kept confidential such that if no. of users simultaneously accessing the cloud, then data files that are personal to only particular client must be able to hide from other clients on the cloud during accessing of file.

4) Trustworthiness

The remote cloud must possess the Trustworthiness characteristic. Because the user/client stores their private data; therefore the cloud and remote backup cloud must play a trustworthy role.

5) Cost efficiency

The cost of process of data recovery should be efficient so that maximum no. of company/clients can take advantage of back-up and recovery service.

3.2 Seed block algorithm(sba) architecture:

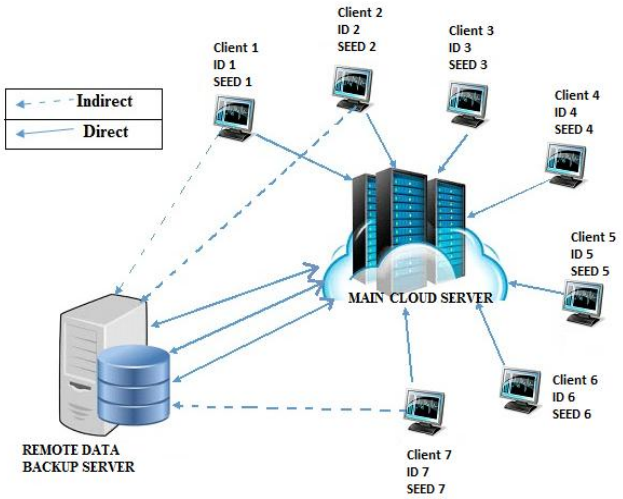

Fig.2: Architecture of seed block algorithm.

The main purpose of this algorithm is focuses on simplicity of the back-up and recovery process. In this algorithm uses the concept of Exclusive– OR (XOR) operation of the computing world. For ex: - Suppose there are two data files: A and B. When we XOR A and B it produced X. If we want our A data file back which was destroyed then we are able to get A data file back, then it is very easy to get back it with the help of B and X data file. The Seed Block Algorithm works to provide the simple Back-up and recovery process. The architecture of this algorithm is shown in Fig.2. Fig.2 consist of the Main Cloud and its clients and the Remote Server. First we set a random number in the cloud and unique client id for every client. Second, whenever the client id is being register in the main 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 stored at remote server.

When client creates the file in cloud first time, it is stored at the main cloud. When it is stored in main server, 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 unfortunately file in main cloud crashed / 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 and return the resulted file i.e. original file back to the requested client.

3.3 SBA Algorithm:

Initialization: Main Cloud:M_C Remote Server: R_S;
Clients of Main Cloud: C_i; Files: a_1 and a_1';
Seed block: S_i; Random Number:r ;
Client's ID: $Client\_Id_i$
Input: a_1 created by C_i; r is generated at M_c;
Output: Recovered file a_1 after deletion at M_c
Given: Authenticated clients could allow uploading, downloading and do modification on its own the files only.
Step 1: Generate a random number.
Int $r = rand\ ()$
Step 2: Create a seed Block S_i for each C_i and Store S_i at R_S
$S_i = r \oplus Client\_Id_i$ (Repeat step 2 for all clients)
Step 3: If $C_i/Admin$ creates /modifies a a_1 and stores at M_c,
then a_1' create as
$a_1' = a_1 \oplus S_i$
Step 4: Store a_1' at S_i
Step 5: If server crashes a_1 deleted from M_c,
then, we do EXOR to retrieve the original as a_1:
$a_1 = a_1' \oplus S_i$
Step 6: Return a_1 to C_1.
Step 7: END.

IV. CONCLUSION

In this paper, we presented detail design of proposed SBA algorithm. Proposed SBA is robust which is very useful for the users to collect information from any remote location without network connectivity and also to recover the files which deleted If the cloud gets destroyed due to any reason, using SBA we can recover all the data. The proposed SBA also focuses on the security concept for the back-up files stored at remote server, without using any of the existing

encryption techniques. It will take minimum time for the recovery process so that the time related issues are also being solved by proposed SBA.

ACKNOWLEDGEMENT

We would like to give our sincere gratitude to our guide Prof. Y. K. Patil who encouraged and guided us throughout this paper.

We especially grateful to seminar Co-ordinator Prof. Bharat Burghate and H.O.D Prof. G. M. Bhandari for their valuable guidance and encouragement. Last but not the least, many thanks and deep regards to Principal Sir for their support and encouragement.

REFRENCES

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

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

[3] Y.Ueno, N.Miyaho, and S.Suzuki, , 2009, "Disaster Recovery Mechanism using Widely Distributed Networking and Secure Metadata Handling Technology", Proceedings of the 4th edition of the UPGRADE-CN workshop, pp. 45-48.

[4] Giuseppe Pirr´o, Paolo Trunfio , Domenico Talia, Paolo Missier and Carole Goble, 2010, "ERGOT: A Semantic-based System for Service Discovery in Distributed Infrastructures," 10th IEEE/ACM International Conference on Cluster, Cloud and Grid Computing.

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

[6] Lili Sun, Jianwei An, Yang Yang, Ming Zeng, 2011, "Recovery Strategies for Service Composition in Dynamic Network," International Conference on Cloud and Service Computing.

[7] Xi Zhou, Junshuai Shi, Yingxiao Xu, Yinsheng Li and Weiwei Sun, 2008, "A backup restoration algorithm of service composition in MANETs," Communication Technology ICCT 11th IEEE International Conference, pp. 588-591.

[8] M. Armbrust et al, "Above the clouds: A berkeley view of cloud computing,"http://www.eecs.berkeley.edu/Pubs/TechRpts/2009//EEC S-2009-28.pdf.

[9] F.BKashani, C.Chen,C.Shahabi.WSPDS, 2004, "Web Services Peerto- Peer Discovery Service ," ICOMP.

[10] Eleni Palkopoulouy, Dominic A. Schupke, Thomas Bauscherty, 2011, "Recovery Time Analysis for the Shared Backup Router Resources (SBRR) Architecture", IEEE ICC.

[11] Lili Sun, Jianwei An, Yang Yang, Ming Zeng, 2011, "Recovery Strategies for Service Composition in Dynamic Network," International Conference on Cloud and Service Computing, pp. 221–226.

[12] P.Demeester et al., 1999, "Resilience in Multilayer Networks," IEEE Communications Magazine, Vol. 37, No. 8, p.70-76.

[13] S. Zhang, X. Chen, and X. Huo, 2010, "Cloud Computing Research and Development Trend," IEEE Second International Conference on Future Networks, pp. 93-97.

[14] T. M. Coughlin and S. L. Linfoot, 2010, "A Novel Taxonomy for Consumer Metadata," IEEE ICCE Conference.

[15] K. Keahey, M. Tsugawa, A. Matsunaga, J. Fortes, 2009, "Sky Computing", IEEE Journal of Internet Computing, vol. 13, pp. 43-51.