Backup of real time data and Recovery using cloud computing

[1]Karishma Nadhe, [2]Sushma Somani
SE, VJTI Mumbai, India

***Abstract* - Cloud computing is the technology that is widely used for storing large volume of data in organizations. The important issues are data protection and its confidentiality and to maintain this we propose a system. Storing data at remote location and to restore data in case the data is deleted without requirement of network connectivity are the main objectives of the proposed system. Here we are using Seed Block Algorithm for the achieving our objectives. Seed Block Algorithm focuses on the security concept for the back-up files stored at Remote Server, without using any of the existing encryption techniques. Data include document files as well as real time data like current location of a person or device. Seed block algorithm solves the time related issues. For the purpose of tracking assets, personnel, and/or patients the system implements Real Time Location System (RTLS). The current best use of RTLS is for asset tracking, but importantly whole-hospital deployment as well as centralized control of the system is required, preferably by materials management or biomedical engineering departments.**

***Keywords* - Seed Block Algorithm, Real time data, Remote Server, Real Time Location System**

I. INTRODUCTION

Cloud computing is also called as on-demand computing. Cloud computing provides shared processing resources and data to computer and other devices on demand. Users and enterprises are provided with various capabilities to store and process data in third party data centres by cloud computing and storage solutions. As the service provider can access the data that is in the cloud at any time so cloud computing possess privacy concerns. It could accidentally or deliberately alter or even delete the information. For this reason the information that we store or process in cloud should be done in encrypted format to prevent unauthorized access. For encrypting the data we are making use of Seed Block Algorithm.

Cloud computing is based on virtualization that takes a very different approach to disaster recovery. The entire server which includes the operating system, application and data is encapsulated into a virtual server with virtualization. Within few minutes, this entire virtual server can be copied or backed up to an offsite data centre and spun up on a virtual host. The effective reduction of cost and increase in the performance is due to the added benefit of disaster recovery with cloud computing.

Previously for backup strategies many techniques were proposed but could not fulfil issues like implementation complexity, cost reduction, security and time related issues. Here we are using smart remote backup algorithm, that is, Seed Block Algorithm that meets all the mentioned issues. All the information is collected by it to store at remote location and restore the same if the data is lost or deleted without the need for network connectivity.

In the hospitals, clinical processes can be documented and improved by tracking patients and devices, including identifying routine breakdowns in care provision. Further benefit in this section includes assigning providers to patients automatically, and disciplining clinicians or other staff who are underperforming in their roles. Patients may get lost as they are moved to various departments within a hospital and tracking the patients helps to locate them. Before medical procedures, identities of patients can be verified by keeping track of patients, and it is also useful in providing quicker turnover of beds by improving discharge processes. Considering all the mentioned uses of the systems, increasing efficiency, improving safety and reducing operational costs are the benefits of implementing RTLS.

II. LITERATURE REVIEW

In literature, different algorithms such as PCS[1], HSDRT[2], ERGOT[3], Linux Box[4], Cold/Hot backup strategy[5] are already defined for recent back-up and recovery techniques that have been developed in cloud computing domain. The following review shows that none of these techniques are able to provide best performances under all circumstances such as cost, security, low implementation complexity, redundancy and recovery in short span of time. A survey and comparison of these techniques are given as follows.

PCS is comparatively reliable, simple, easy to use and more convenient for data recovery totally based on parity recovery service but it is unable to control the implementation complexities [1]. Few works have been done for efficient personal data recovery service while lots of effective backup and recovery technologies, including data de-duplication and incremental backup, have been developed for enterprise level data backup service. A plain data backup-based recovery service is not adequate for public service since the privacy protection is a crucial issue for providing a personal data recovery 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 [1].

On the contrary, HSDRT has come out an efficient technique for the movable clients such as laptop but it fails to manage the low cost for the implementation of the recovery and also unable to control the data duplication [2].

Rather, ERGOT combines Distributed Hash Tables (DHTs) and Semantic Overlay Networks (SONs) in distributed infrastructures such as Grids and cloud to enable semantic based service discovery. ERGOT provides efficient way of retrieval of that is based on the semantic analysis but is unable to focus on time and implementation complexity [3].
Similarly, we also found that one technique in addition, Linux Box model is having very simple concept of data back-up and recovery with very low cost. However, in this model protection level is very low [4].
All these techniques tried to cover different issues maintaining the cost of implementation data increases i.e. cold and hot back-up strategy [5] that performs backup and recovery on trigger basis of failure detection.

Table-I describes the advantages and disadvantages of some of these foresaid techniques. The role of a remote data back–up server is very crucial and hot research topic due to the high applicability of backup process in the companies.

Table-1 Comparison between various techniques of Back-up and recovery [7]

Sr no.	Approach	Advantage	Disadvantage
1	HSDRT	Used for Movable clients like laptop, smart phone	Costly Increase redundancy
2	Parity Cloud service	Reliable Privacy Low cost	Implementation Complexity is high
3	ERGOT	Perform exact match retrieval Privacy	Time complexity Implementation complexity
4	Linux box	Simple Low cost for implementation	Required higher bandwidth Privacy Complete server backup at a time
5	Cold/Hot back-up strategy	Triggered only when failure detected	Cost increases as data increases gradually

The most common systems, locate equipment or a person today by employing radiofrequency identification (RFID), Wi-Fi or Wireless Local Area Network (WLAN), ultra-wide band (UWB), infrared (IR), Zig-Bee, Bluetooth, or ultrasound [8]. RFID tags were largely employed in initial systems but the performance was extremely poor as it provided inaccurate information about the location and cumbersome interfaces. Adoption of such systems is obviously low which do not function properly because users are certainly not interested in learning and operating a system where workload is increased.

III. PROPOSED SYSTEM

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

A. *Remote Data Backup Server*

When we talk about Backup server of main cloud, we only think 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 i.e. far away from the main server and having the complete state of the main cloud. The central repository is the term used to refer main cloud and remote repository is used to refer remote backup cloud.

1) Architecture of Remote Data Backup Server

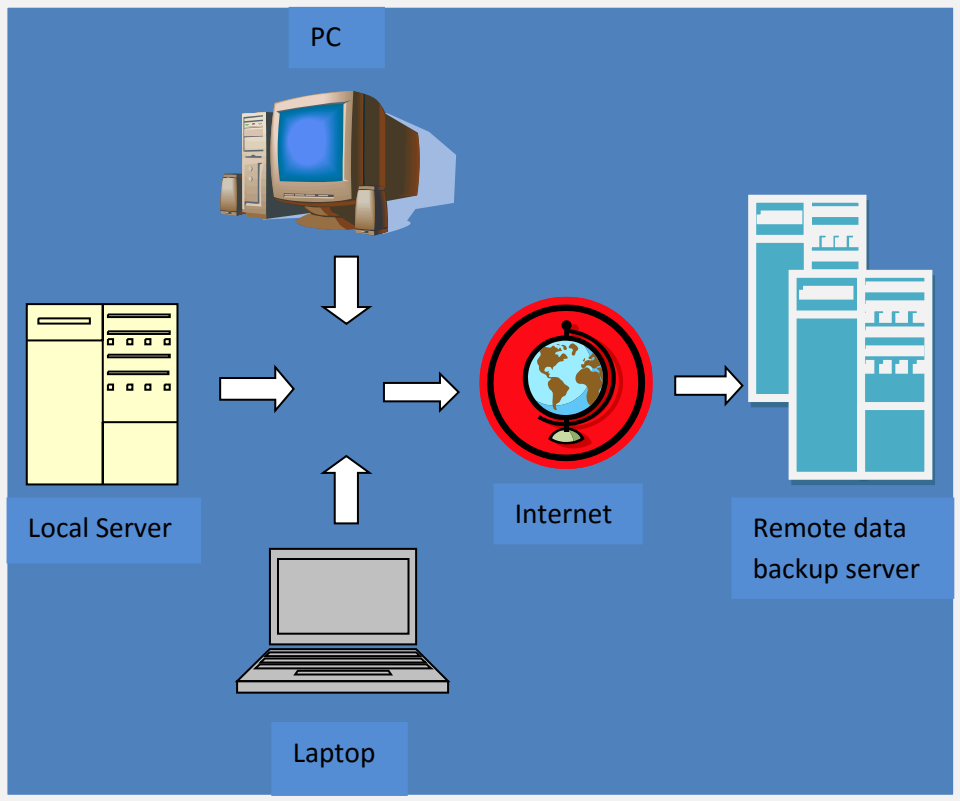

Fig.1 Architecture of Remote data Backup Server

2) Characteristics of Remote Data Backup Server

Following are the characteristics of Remote Data Backup Server:-

1) Data Integrity:

Complete state and the whole structure of the server are concerned in data integrity.

During transmission and reception it verifies the data such that it remains unaltered. It is the measure of the validity and fidelity of the data present in the server.

2) Data security:

The utmost 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/clients either intentionally or unintentionally.

3) Data Confidentiality:

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

4) Trustworthiness:

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

5) Cost efficiency:

For the maximum number of company/clients to avail the advantage of back-up and recovery service the cost of process of data recovery should be efficient.

B. *Smart Remote Data Collection Server*

The main purpose of this algorithm is to focus on simplicity of the back-up and recovery process. In this algorithm it used the concept of Exclusive– OR (XOR) operation of the computing world. For ex: - Suppose there are two data files: X and Y. When we XOR X and Y it produced Z. If we want our X data file back which was destroyed then we are able to get X data file back, it is very easy to get back it with the help of Y and Z data file. The Seed Block Algorithm works to provide the simple Back-up and recovery process. The architecture of this algorithm is shown in Figure 2. Figure 2 consists 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.

3.1 Architecture of Seed Block Algorithm (SBA):

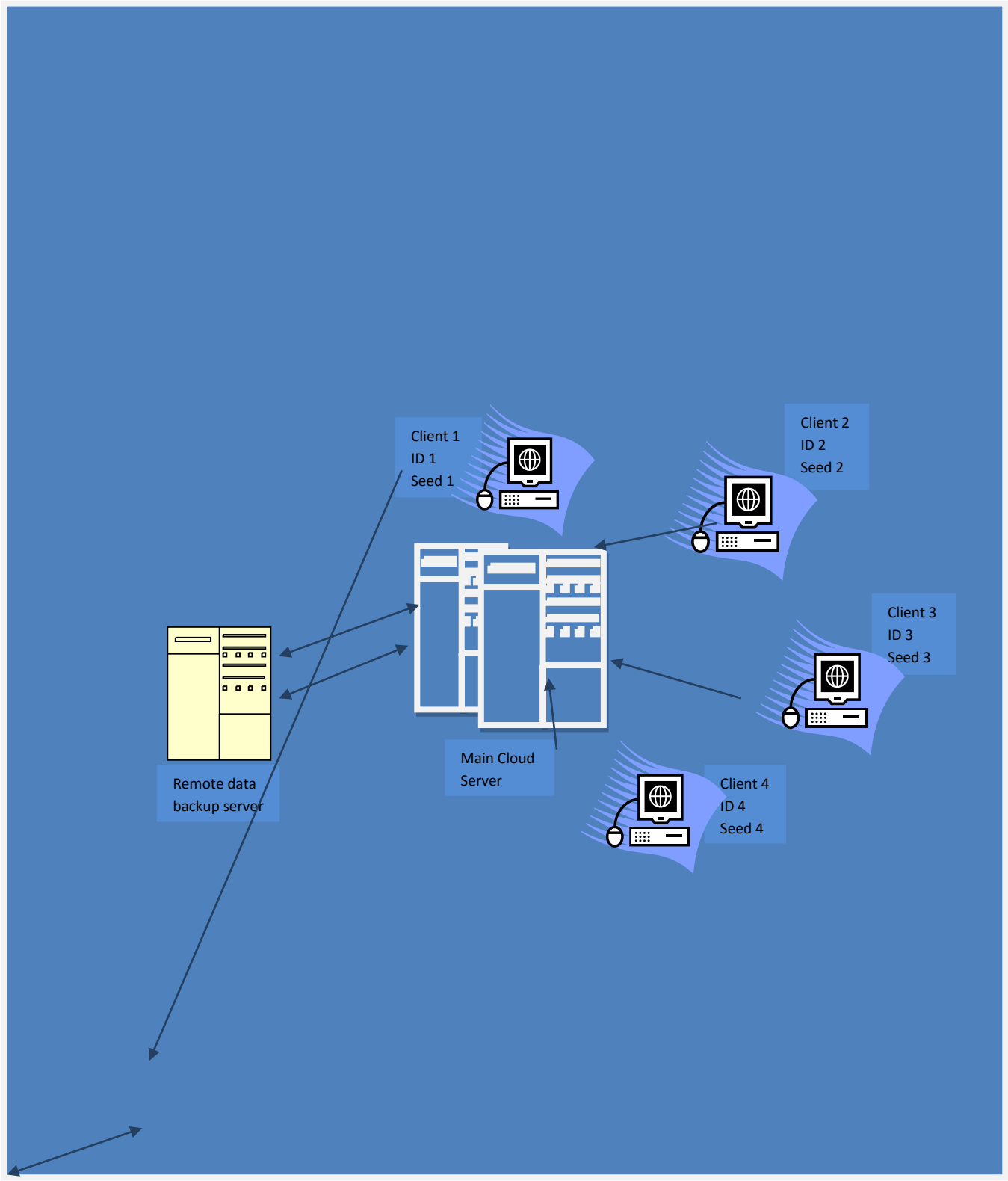

Fig.2: Architecture of seed block algorithm.

3.2 SBA Algorithm [6]:
Initialization: Main Cloud (Mc): Remote Server: (Rs) ;
Clients of Main Cloud (Ci) **;** Files(a1) and (a1');
Seed block: (Si)**;** Random Number**: (r);**
Client's ID: (Client_Idi)
Input: (a1)created by (Ci)**;** **r** is generated at (Mc);
Output: Recovered file a1 after deletion at Mc
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 O
Step 2: Create a seed Block Si for each Ci and Store
Si at Rs
Si= r ⊕ Client_Idi (Repeat step 2 for all clients)
Step 3: If Ci/Admin creates /modifies a a1 and stores
at Mc,
then a1' create as
a1'=a1 ⊕ Si
Step 4: Store a1'at Si
Step 5: If server crashes a1 deleted from Mc,
then, we do EXOR to retrieve the original as a1:
a1=a1'⊕ Si
Step 6: Return a1 to C1**.**
Step 7: END.

C. *RTLS Functionality:*
The Vendors have made impressive claims about the accuracy of RTLS in tracking and locating. In our study, most of the RTLS were RFID based, and most hospitals had implemented their systems for asset tracking. Other systems that were not RFID based included ultrasound-, IR-, UWB-, and ZigBee- based technologies. In hospitals, for patient identification or tracking as well as personnel tracking RTLS is useful. In addition to asset tracking in hospitals, addition of temperature monitoring function to their system is also a useful technique. Newer RTLS offer technological improvements but systems have maintained former or acquired novel drawbacks. The implementation of new systems depicts the biggest improvement with RTLS.
In hospitals, many will be able to operate on existing wired or wireless networks, as "plug-and-play" technologies are also offered that are relatively easy to install. This clearly indicates that it is both faster and less expensive to have systems go live and so it is very beneficial. The clear best purpose to which RTLS has been put is asset tracking. RTLS must be implemented throughout hospital facilities so that tracking coverage is as complete as possible and in this way we can maximize their utility. Along with this, vendors that can guarantee a fair level of accuracy in their tracking systems must be chosen by the hospitals.
Challenges for RTLS are created by lead-shielded rooms to operate within those spaces as well as in contiguous hallways and rooms. Problems with RTLS installations are also created due to access to an electrical supply and the ability to pull cables because the systems depend on strategically placed receivers that will "read" the signals emitted by tags. Newer facilities frequently have interstitial spaces above the ceiling that can be accessed for cable and electrical drops. Many a times, the incident that is observed in hospitals is the disappearance of equipment from hospital facilities. The devices can get lost in patients' bedding and end up damaged or thrown away. This includes small items such as telemetry devices. Hospitals that report loss of bigger equipments like wheelchair due to theft leads to a major contribution to equipment costs. To solve these problems RTLS are designed so that items will not be unknowingly lost or stolen.
To make RTLS a success, a critical element is placing the systems within the domain of a department that is responsible for hospital equipment, such as materials management or biomedical engineering. To make RTLS function well, these departments have the most stake and to make use of the tracking function for inventory control or preventive maintenance they are in the best position. The clinical staffs needs to be educated about the tracking system and permitted to access it should they find it valuable, to minimize sabotage of equipment tags. The hospitals that will assign personnel from materials management to locate and deliver equipment to clinicians will definitely achieve higher rates of use of their RTLS as well as greater satisfaction among all personnel.

IV. CONCLUSION

Due to computerization and availability of data from remote location, vast amount of data is going to be collected on the web servers. It helps in reducing allocation of geographical area required for storing records and also promotes paperless work. Time consumed for searching required documents is less. Every organization prefers computerization as well as remotely accessible web services. Hence data security and protection comes in highest priority so recent developments will be on securing and protecting data collection on web server. There has been tremendous gain with RTLS technology, yet the vendors should further improve the accuracy of their systems' tracking functions and software capabilities. To be effective, implementation of RTLS must go beyond the simple deployment of technology. The RTLS that best matches the specific goals of the hospitals and will provide the best results in their own facilities should be chosen by the hospitals, if their unique material and organizational constraints are given.

V. REFRENCES

[1] 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 IEEETrustCom-11/IEEE. ICESS-11/FCST-11
[2] Yoichiro Ueno, Noriharu Miyaho, Shuichi Suzuki, Muzai Gakuendai, Inzai-shi, Chiba, Kazuo Ichihara, Oriented Architectures" North Carolina State University "Performance Evaluation of a Disaster Recovery System and Practical Network SystemApplications ", Fifth International Conference on Systems Networks Communications.

[3] 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.
[4] 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.
[5] Lili Sun, Jianwei An, Yang Yang, Ming Zeng, 2011, "RecoveryStrategies for Service Composition in Dynamic Network," International Conference on Cloud and Service Computing.
[6] Ms. Mayuri Tidke, Ms. Vijayshree Jadhav, Ms. Sonali Parab, Ms. Shubhrata Patil, Prof. Y. K. Patil, 2015,"Seed Block Algorithm: A smart technique for Data Back-up and Recovery in Cloud Computing", IERJ, Volume 1, Issue 9.
[7] Kalyani Bangale, Karishma Nadhe, Nivedita Gupta, Swati Singh Parihar, Gunjan Mankar, "Smart Remote Health-Care Data Collection Server" IJCSMC, Vol. 3, Issue. 2, 2014.
[8] Jill A. Fisher, Torin Monahan, "Evaluation of real-time location systems in their hospital contexts" ijmi journal, 8I, 2012.