Multicloud Based Backup and Disaster Recovery Service

Prof. Nagesh R, Krishna B A, Chandan H R
Department of Information Science and Engineering
SJC Institute of Technology, Chickaballapura, Karnataka, India

Abstract: *In the recently years organizations have started to generate huge amounts of information of different formats. Availability of this huge amount of data collected is required for continuity of business during disasters .It is one of the main necessities of the time. Currently these backup and recovery services run with disadvantages of being very expensive with high restore and backup time. Solution to the above issues is cloud based platforms which will be used for these services. Cloud based platforms can be used for any amount of resources limited or boundless. Scheduling strategy COST-RTO tradeoffs scheduling algorithm is used to find the right server location to back up the replicated data. This algorithm mainly reduces the cost and data recovery time during a disaster. Here 3 backups are made and storage of these replicas of the data is done by this replica scheduler. Methods like incremental backup is being used to reduce backup time, metadata of the backed up data will be generated. This generated meta-data will help to retrieve data faster while restoring .Delivering these solutions as cloud services for faster and lowered cost accessing is the main objective which is fulfilled by the application..*

Keywords: Disaster Recovery

I. INTRODUCTION

Many services in the present generation use IT sector as their main functioning unit. Services such as financial and health care service can be catastrophic if they are inaccessible to the user at a given time or place. There can be huge economic losses or social problems even if there is a short period of downtime or data loss. Catastrophic system failures lead to downtime hence disaster recovery services are used to protect critical data by many important business and public services.

Creating standby systems at geographically different places, periodical backup of data and continuous synchronization of data were some of the popular methods of disaster recovery. These methods prove to be costly like purchasing servers, infrastructure, maintaining facilities, employing human resource and also time-consuming. If any service provider builds a disaster recovery data center of his own he will require huge investment.

In the recent years, Cloud computing has risen to popularity as a new and more efficient service paradigm. These cloud services has risen to popularity because of its „how much you use that much you pay" model. The cost of building and maintaining own data centers is tremendously high when compared to the disaster recovery service model used here. The pay for only the amount of services user has used vastly reduces the cost of disaster recovery.

Data reliability is one of the main services of building the data disaster recovery model. This cloud based model is used for maximizing the data availability. A disaster recovery model which reduces cost is still a challenge. Just like all other computer systems, cloud computing system can also encounter certain risks such as software bugs, hardware faults, network intrusions, human-caused damage, natural disasters, etc. The above mentioned risks may lead to loss of data in many cases which leads to severe damage of clouds main service that is data reliability.

At present there are many data protection strategies like Amazon S3, Google GFS and Apache HDFS have adopted 3-replicas data redundant mechanism in which data is stored in data centers. This leads to a single point failure problem, where an entire data center failure would result in data loss. To avoid this problem, cloud service providers implemented geographically distant places for storage of data to protect their data.

The above method still runs the risks of all data center present in different geographical places running the risk of failures in the same period due to some common cause of disaster across data centers. These data centers can become unreachable due to long distance in geography which can lead to network disasters.

The most advanced cloud services of the present generation have encountered several types of outrages and resulted in many public services being down. Therefore, in this model a solution is found for disaster recovery service where we utilize multiple data centers from different cloud service providers.

In this model for multi-cloud based disaster recovery service, which we call DR-Cloud (DR for disaster recovery). DR-Cloud application will access multiple clouds and use their resources cooperatively. In turn users can access this application by using a simple useable interface. This cloud service provider will lease resources from other cloud service providers in the pay-as-you- go mode. Then based on dynamic scheduling algorithm we can selectively backup the data to the required other cloud service provider. Thus data disaster recovery is no longer limited to the data centers of one cloud service provider. By using an appropriate backup data scheduling strategy disaster recovery service provider can achieve better effect, which is to minimize the cost while promoting the service quality.

- Recovery Point Objective (RPO): Since the disaster has occurred how old the data is when it was last updated or backed-up. The amount of information lost because of the last update this should be as less as possible.
- Recovery Time objective (RTO): The time taken to recover the data once a disaster has been declared or when the application has is back up and running.
- Disaster Recovery Plan (DRP): The recovery [9] plan of the IT system of an organization.

Data center is a centralization of operations and IT devices for organizations and also a place to collect manage and disseminate information. Size of data center is depending on the size of operations for each type of business as well.

II. LITERATURE SURVEY

The below literature survey is made on different topics of traditional backup and recovery tools to the different methods which were used as disaster recovery methods.

2.1 Traditional Backup and Recovery Methods
A. Cold Site Method

In Figure 2.1, primary location of the data[8] is shown. The boxes show different applications, we back them all up on tape and store them. Here the challenge is:

Recovery is a tedious and time consuming.

The data here is backed up infrequently.

It may take many days for the data to be re-assembled.

B. Frugal Warm Site Method

Consider Figure 2.2, expensive recovery location is brought and only the important part of the application is uploaded to that site as it is very expensive and recovery has same problems as above.
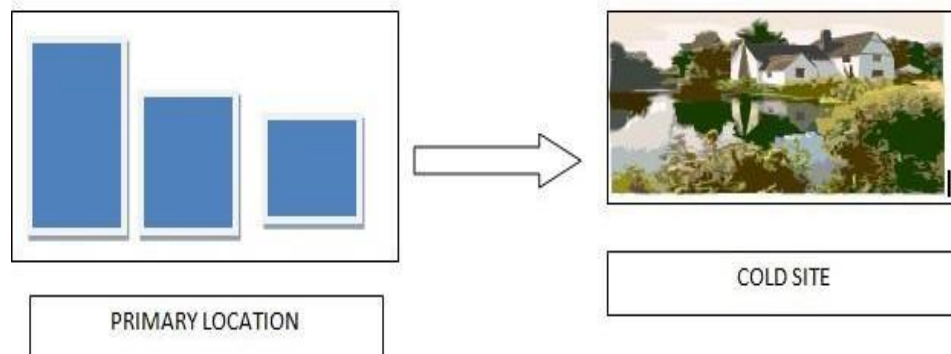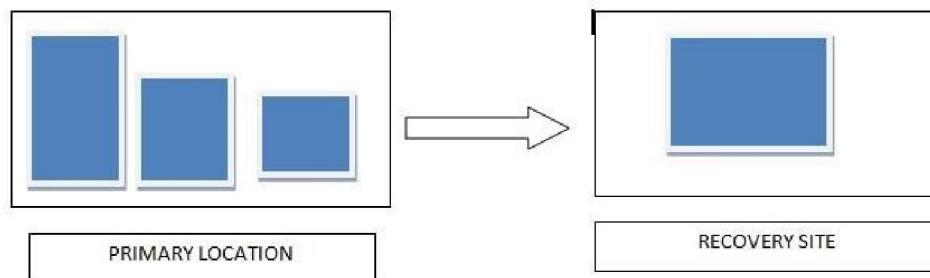

Figure 2.1 Cold Site Method

Impact Factor: 6.252

Figure 2.2: Frugal Warm Site Method

C. Local Hot Site

As shown in Figure 2.3, this method is mainly used by companies that have multiple tiersor educational institutions. In the same campus in another building the backup is kept. Here the main challenge is the same disaster may affect both data centres leading to data losses.

D. Global Hot Site

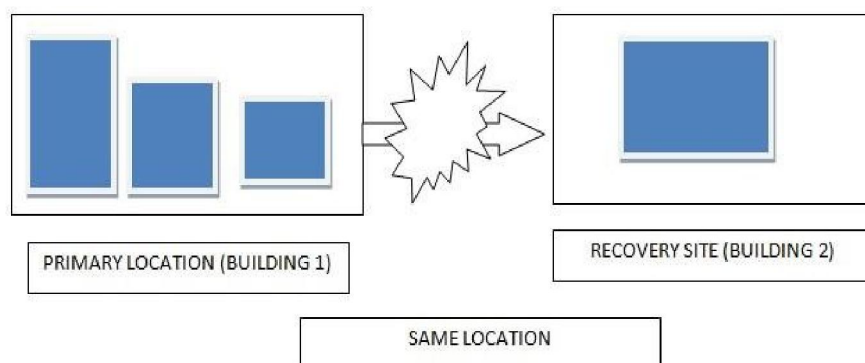Consider Figure 2.4, the data here is stored at far away locations. This is a very expensive method.

Figure 2.3 Local Hot Site

III. PROJECT REQUIREMENTS

3.1 Software Requirements
- Operating System: Windows XP.
- IDE: Eclipse Database: MySQL 5.5.
- Language: JAVA/JEE.
- Cloud creation: Dropbox account.

3.2 Hardware Requirements
- Processor: Pentium IV 2.4 GHz.
- Hard Disk: 5GB recommended.
- Memory: 512 MB of RAM, 1 GB recommended.
- Display, keyboard, mouse: Standard displays, keyboards and 2 button mouse.

3.3 Computing Environments
Operating System: Windows XP

It is a product of Microsoft, windows 7 and above supports hyper-threading where a singleprocessor will act like 2 separate processors for application and the operating system. Multi-cores are also supported here.

IDE: Eclipse Luna version

Eclipse IDE runs on cross platforms with very easy usage for java developers. It containsmany versions like callisto being the oldest to the newest and more popular once like indigo, luna, mars etc.

Database: MySQL 5.5

This is mainly a 2 edition system an open source one and enterprise server. The older versions of MySQL have been used extensively at industry standards. Newer MySQL versions with have features like using view methods, triggers and stored procedures

Server: XAMPP and APACHE

XAMPP[13] server is one of the many other servers WAMP, MAMP, LAMP and SAMPrunning under the license of GNU. This server mainly came during the replacement of MySQL with mariaDB .In project XAMPP is mainly used as a user interface to access MySQL DB for simpler access to user.

Apache being one of the most common and sought after server mainly because the functionalities are made into modules complied and stored to be accessed by user when required easily.

Dropbox

Files are stored just like Google drive in cloud, where we create and an account andgive a Dropbox password .it will open a Dropbox folder which can be easily used for storage or sharing between a groups of people, this folder will be synchronized on all systems where the account is opened.

3.4 Language
JAVA/JEE

Java Platform standard edition and enterprise edition are used in my project.JEE extends the functionalities of the standard edition. It contains packages of JEE5,JEE6 andJEE7 already used as base APIs.

JAVA is mainly object-oriented .Mainly having a functionality of java API"s in thestandard editions is very helpful to the users.

3.5 Functional Requirements
A. Functional Requirements of Incremental-backup Module

Input: Authenticated user asks for backup and presses submit**.**

Processing

- User entered filename; backup method and username are store in metadata.
- The COST-RTO algorithm is run which checks the cloud storage, recovery time and otherdata from the metadata.
- The cloud -data is checked to find out on which cloud user data has to be stored.
- Once selected the cloud-ids of where each replica of the data is backed-up is stored inmetadata for easy retrieval.
- Then display the pop-up "backed-up".

Output: Displays the result with appropriate information "backed-up".

B. Functional Requirements of Recovery Module

Input: Authenticated user asks for recovery of the backed-up data.

Processing

- User entered filename is checked in the metadata.
- The method which was used to back-up the requested file is searched from database.
- The cloud-id where the backup of the requested file is got.
- Using the cloud –id in that where we get the fastest recovery time .this is assessed andretrieval is done.
- Then display the pop-up "restored successfully".

Output: Displays the result with appropriate information "Restored".

C. Non-Functional Requirements
Performance: The proposed project should be very fast and flexible usage for a naïve end user.
Reliability: The system must run and come into action whenever a user requires it or a disaster occursas fast as possible.
Capability: The capability of any function of that project to run smoothly and access required data or other functions.
Maintainability: The systems operation and maintenance shouldn't be of much trouble to the user or the administrator hence approach must be kept as direct as possible to repair if an error occurs. For better usage by the user the updates must be rolled out at required time intervals with new and improved features or fixes for recent bugs.

V. DESIGN

5.1 Architectural Design
Architecture diagrams are used to share the knowledge about the project and seek feedback fromother members of the project. It is the easier way of showing how a project will execute if all conditions are true.

Figures 4.1, 4.2 below are the architectural diagram for the backup and recovery service. In Figure 4.1 The first part of the architecture diagram shows the different components of backup which contains cloud service providers CP1,CP2,…..,CPn, where CP1 is the main cloud and canbe accessed by the user through the cloud interface from which the request buffer accesses the requests got and tells the replica scheduler to run the scheduling algorithm to check different constraints and store that information in the metadata and store the backups to the cloud service provider. In Figure 4.2, the restore takes place easier and faster than backup because of the data entered in the metadata. This data is accessed easily path of backup is found and data is restored.
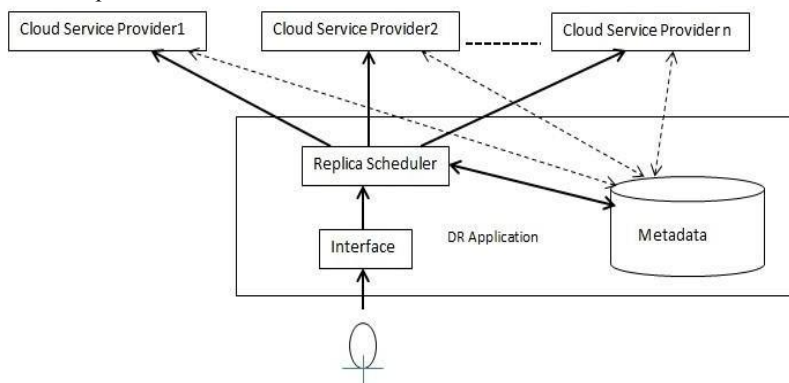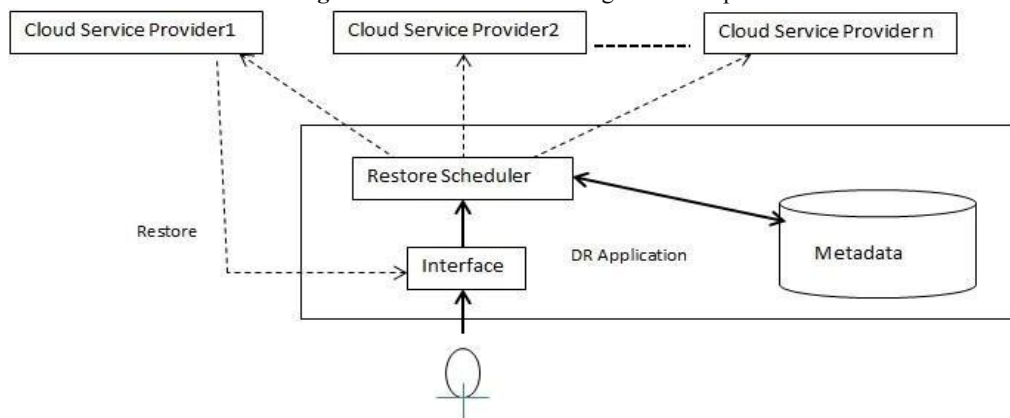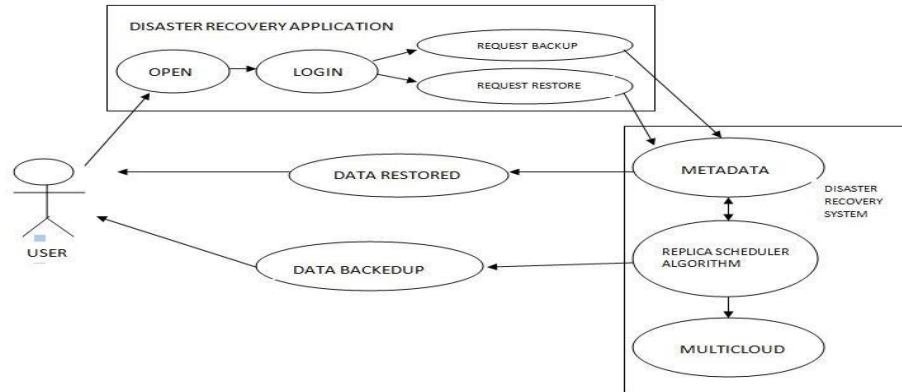

Figure 4.1: Architectural Design of Backup

Figure 4.2: Architectural Design of Restore

5.2 Use Case Diagram

This model is mainly to capture the working of the project when it is running or operating. Compared to other diagrams we look into specific purposes of that module in the system working here.

VI. IMPLEMENTATION

The below is the discussion on implementation made on creating cloud using Dropbox toolalso DR (disaster recovery) service.

6.1 Implementation of cloud using Dropbox

A. Introduction to Dropbox

It is a multi-storage system that can store any kind of files like spreadsheets, photos etc. which will be automatically synchronized to any personal computer or mobiles device that theuser has logged into.

It works by having copies of the files stored on our computer using cloud based storage andsynchronizing them after encrypting the data. Here automatic synchronization takes place which will lead to reduced time and efforts of user.

It has 3 main features which makes Dropbox amazing be anywhere, always safe and simplesharing. Dropbox helps with files being created anywhere and used anywhere and everywhere bythe user. The 2nd point is the data that is saved by user is always safely present in the infinitely large cloud based storage of the cloud. Lastly it is very simple and easy sharing a file with projectteams, friends and family with drop boxes easy to use interface for users.

B. Installation of Dropbox as a Cloud

Step 1: Create a Dropbox account.

Step 2: There are many SDKs [14] available by the Dropbox development environment.

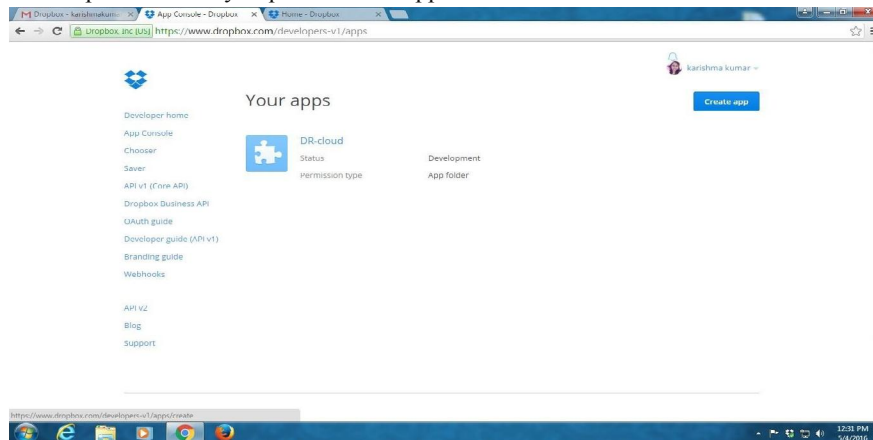Download java SDK import the library required for the application

Figure 5.1 Dropbox Installation-1

179

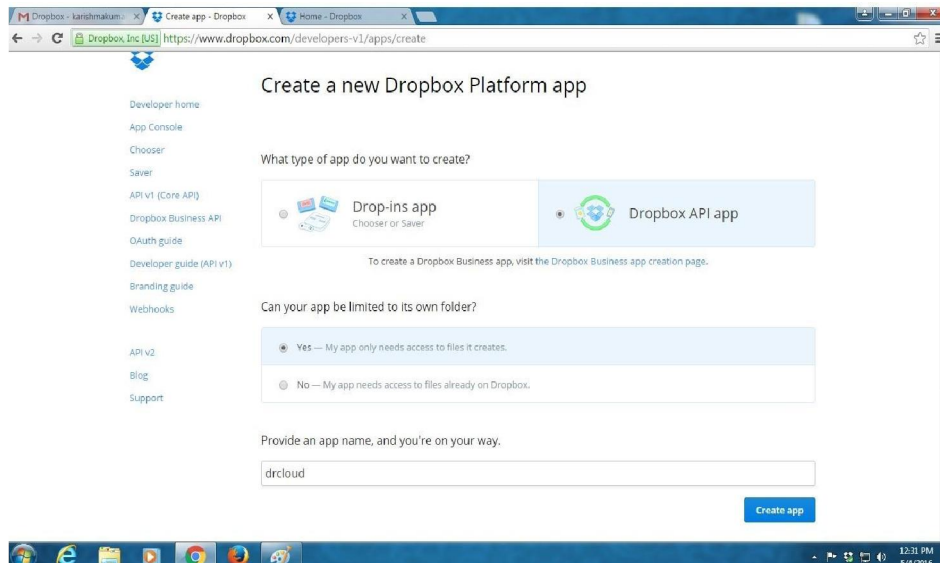Step 3: Connect and register to the App Console where you will need to create a new app.

Figure 5.2 Dropbox Installation-2

Step 4: Once the new app is created from a registered user it is will show APP_KEY andAPP_SECRET. Both will be used in code below to connect the project to the cloud created. Now install java SDK
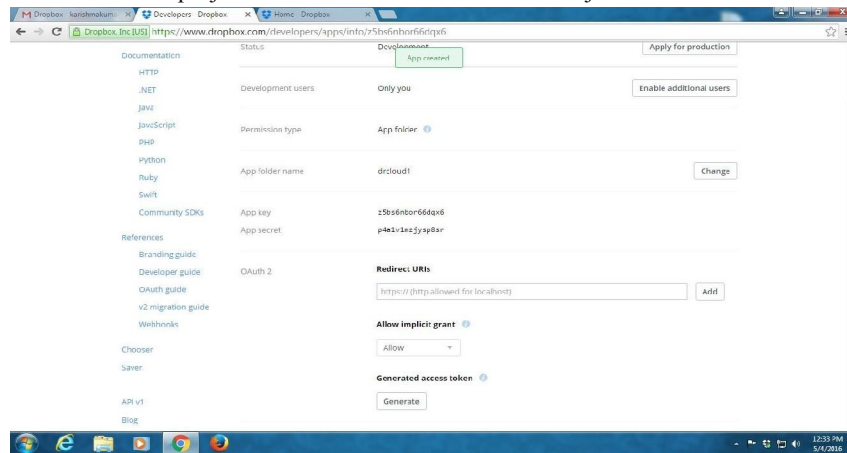

Figure 5.3 Dropbox Installation-3

Step 5: Below code snippet is used to connect the project to the cloud, using APP_KEY andAPP_SECRET.
Pseudo code: Include Dropbox SDK
Import necessary java packagesCreate class {
      Get the app key and app secret from Dropbox developer"swebsite.
      String APP_KEY="insert app key"; String APP_SECRET="insert app secret";
}
Sent to a function DbxAppInfo(app_ key, app _info) as parameters to connect to Dropbox

VII. RESULTS AND ANALYSIS

A brief on implementation results obtained for the project above: having the screenshots of user-interfaces, intermediate steps and final results.

Figure 6.1 shows the ready-made table used in this project at the admin"s side to check the cost and RTO of different multiple- clouds present which will be used by the algorithm to choose the cloud which is best fit for that application.

Impact Factor: 6.252

CLOUD_ID	SPACE	COST(in Rupees)	TIME(sec)
1	10 GB	100	60
2	10 GB	500	20
3	10 GB	300	40

Figure 6.1 Cost-RTO of Different Clouds

Figure 6.2 Backup of the database is made at the different cloud servers located at different locations, as multiple copies of it is done.
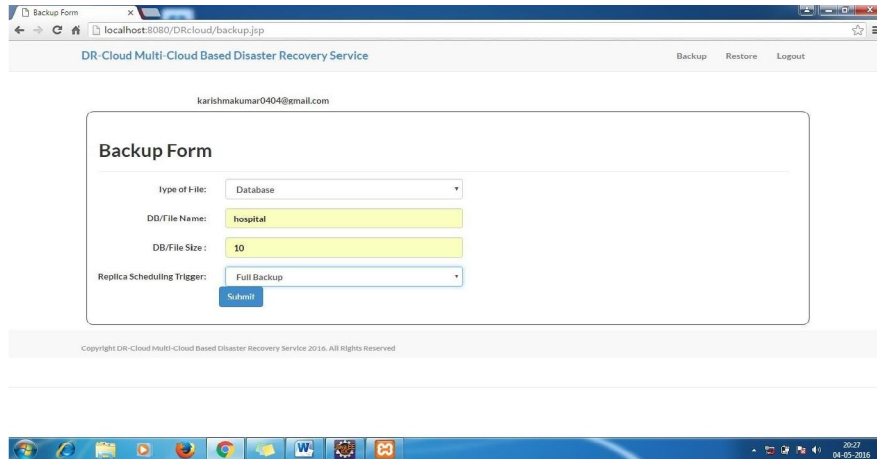

Figure 6.2 Database Backup Form

Figure 6.3 Incremental backup of database which is used by users if we want to backup only recently updated values not the entire data which considerably reduces the bandwidth consumption for storing the replica copies in the cloud.
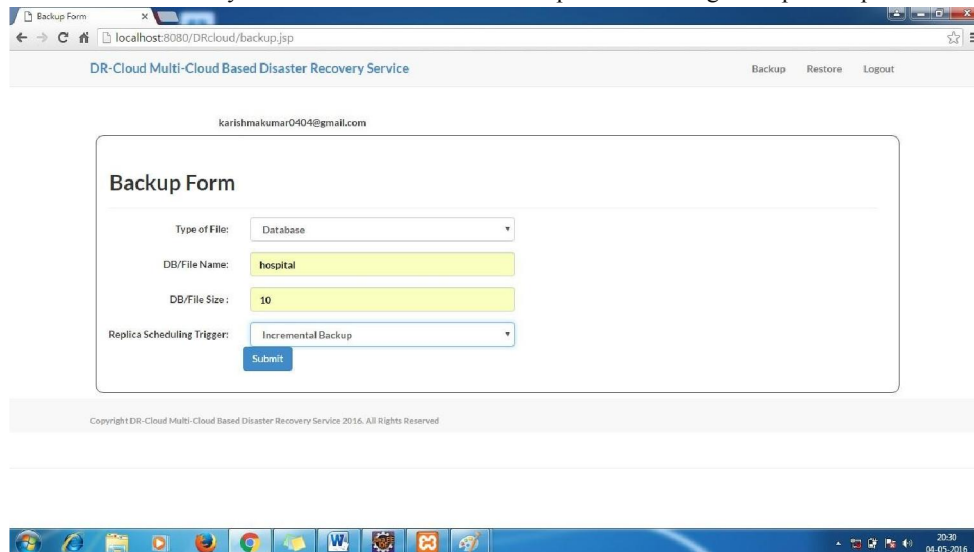

Figure 6.3 Incremental Backup Form

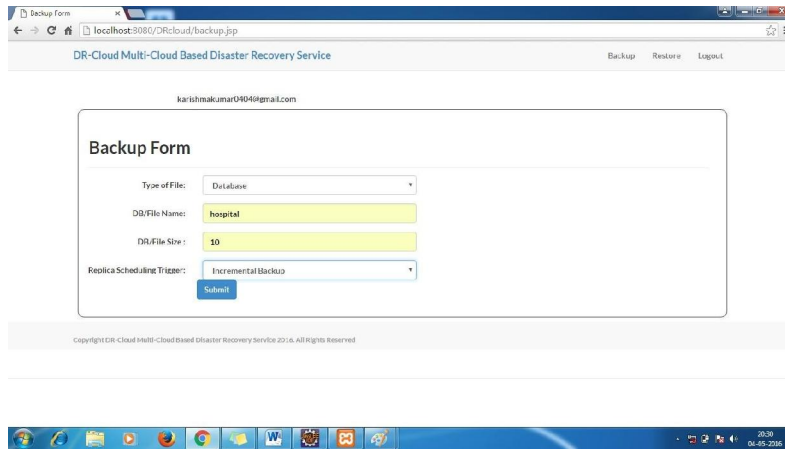Figure 6.4 Backup of any type of files can also be done, with this project.

Figure 6.4 File Backup Form

Figure 6.5 and 6.6, Restore of the database and file is made from the server closest to the disaster site which reduced recovery time and cost.
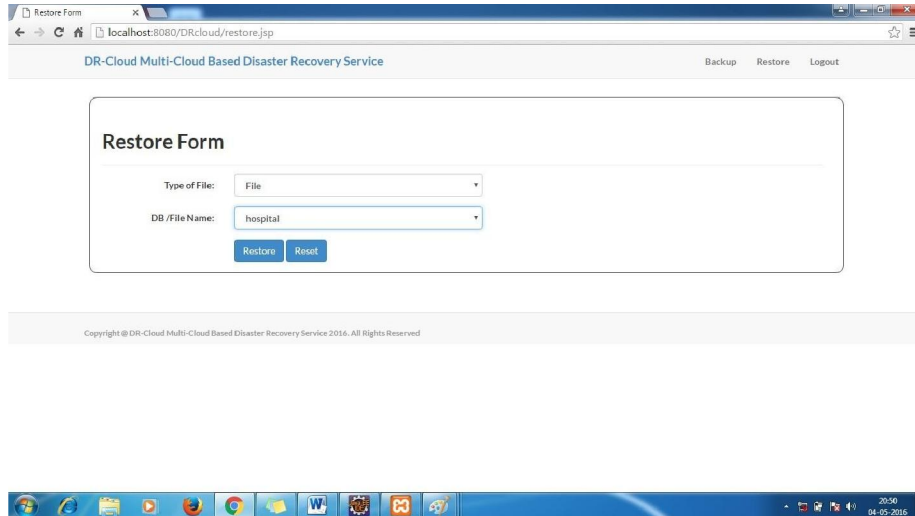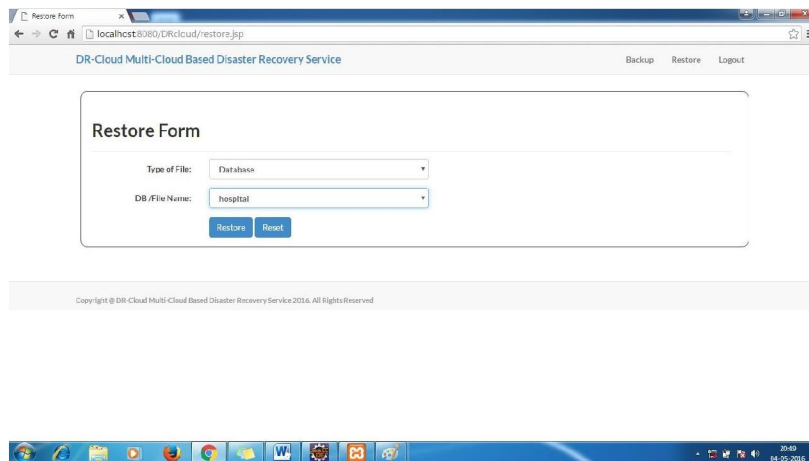

Figure 6.5 Database Restore Form

Figure 6.6 File Restore Form

REFERENCES

[1]. Omar H. Alhazmi, Yashwant K. Malaiya,"Evaluating Disaster Recovery Plans Using the Cloud", 2013 IEEE

[2]. Timothy Wood, Emmanuel Cecchet, K.K. Ramakrishnany,Prashant Shenoy, Jacobus van der Merwey, and Arun Venkataramani,"Disaster Recovery as a Cloud Service: Economic Benefits & Deployment Challenges", AT&T Labs

– Research.

[3]. David Bermbach, Markus Klems ,Stefan Tai , Michael Menzel "MetaStorage: A Federated Cloud Storage System to Manage Consistency-Latency Tradeoffs" 2011 IEEE 4th International Conference on Cloud Computing.

[4]. Zhang Jian-hua and Zhang Nan, "Cloud Computing-based Data Storage and Disaster Recovery",2011 International Conference on Future Computer Science and Education.

[5]. Satoshi Togawa , Kazuhide Kanenishi "Private Cloud Cooperation Framework of e-Learning Environment for Disaster Recovery" , 2013 IEEE International Conference on Systems, Man, and Cybernetics

[6]. Zia Saquib, Veena Tyagi, Shreya Bokare, Shivraj Dongawe, Monika Dwivedi, Jayati Dwivedi, "A New

[7]. Approach to Disaster Recovery as a Service over Cloud for Database system", Computer Networks and Internet Engineering.

[8]. S. Suguna, Dr. A. Suhasini, "Overview of Data backup and Disaster Recovery in Cloud" ISBN No.978-1-47993834-6/14, 2014 IEEE.

[9]. Mike Klien "Disaster Recovery in the Cloud", talks by online expert president of online tech on March 14th.

[10]. Diana Nolting "https://www.bluelock.com/blog/rpo-rto-pto-and-raas-disaster-recovery-explained/", March 1st 2013.

[11]. "The simplest cloud management experience" Courtesy of http://opennebula.org/

[12]. "Openstack installation guide for Ubuntu" Courtesy of http://docs.openstack.org/liberty/install-guide-ubuntu/

[13]. "Introduction to Dropbox" Courtesy of, http://www.commoncraft.com, Nov 16, 2009

[14]. "A brief introduction to XAMPP" Courtesy of http://geekapod.blogspot.in/2011/10/brief-introduction-toxampp.html

[15]. "Using the core API in java" Courtesy of https://www.dropbox.com/developers-v1/core/start/java

[16]. "A hands on introduction for developers" Courtesy of http://dn.codegear.com/article/31863UML Distilled Ch. 3, by M.Fowler