Disaster Recovery(DR)Solutions

In this section, we will discuss some DR solutions which have been proposed to overcome the problems and challenges in cloud-based DR.

Local Backup

A solution for dependency problem has been proposed in (Javaraiah, 2011). A Linux box can be deployed on the side of customers to make control of data and to get backup of both data or even complete application. Local storage can be updated through a secured channel. By this technique, migration between cloud service providers and also migration between public to private, and private to public is possible. In the event of a disaster, local backup can provide the services that were served by the service provider.

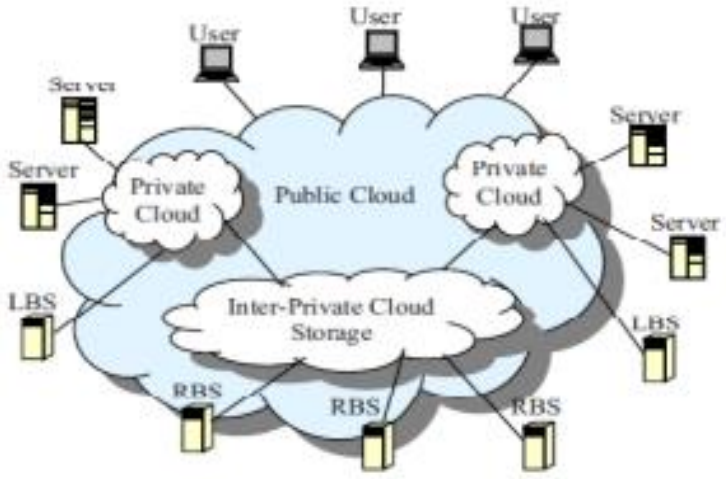
Geographical Redundancy and Backup (GRB)

Although geographical redundancy can be used in traditional model, but it is expensive and unaffordable. In(Pokharel et al., 2010), two cloud zones have a replication of each other. If one zone becomes down, then another zone will be on and provide the services. There is a module that monitors the zones to detect disaster. Primary zone has an active load balancer to request extra resources or even released unused resources. Second zone also has a passive load balancer. Another research (Khan and Tahboub, 2011) has been proposed a method to select optimal locations for multiple backup. The number of places is decided based on the nature of application and priority of services. Distance and bandwidth are two factors to choose the best sites in this method. However, this work neglects some critical factors such as the capacity of mirror sites and the number of node sources which can be hosted in each location.

Inter-Private Cloud Storage (IPCS)

This approach was proposed for cloud data storage (Jian-hua and Nan, 2011). According to Storage Networking Industry Association (SNIA), at least three backup locations are necessary for business data storage. Users' data should be stored in three different geographical locations: Servers, Local backup server (LBS) and remote backup server (RBS). The private clouds are

established for any enterprises consist some servers and an LBS; and also an inter-private cloud storage is created in a public cloud consists the RBSs to be shared between public clouds. This model gives communication ability to backup locations in order to increase data integration.



Resource Management

Heterogeneous clouds consist many different hardware and software such as hybrid storage and diverse disks. In cloud-based enterprises, entire business data are stored in the cloud storage. So, data protection, safety and recovery are critical in these environments. Data in danger is the data which has been processed at the primary host but has not taken place in the backup host yet. So, in the case of disaster, It is necessary to use enhanced technology for data recovery in storage clouds. There are three solutions for data recovery proposed in (Patil et al., 2012):

- Using fastest disk technology in the event of a disaster for replication of data in danger.
- Changing dirty page threshold: The percentage of dirty pages in RAM which have to be waited for flushing to disk might be reduced (Rudolph, 1990).
- Prediction and replacement of risky devices: Some important factors such as power consumption, heat dissipation, carbon

credit utilization and importance of data (stored on each disk) can be calculated in a specific period of time. By these factors, a mathematical equation will be formed to make a replace priority list.

Secure-Distributed Data Backup (SDDB)

An innovative technique has been presented in (Ueno et al., 2010) to protect data in the event of disaster. The data protection technique has six stages:

- First data encryption: Data has to be encrypted after receiving into a data center.
- Spatial scrambling: By a spatial scrambling algorithm, the order of data files is changed.
- Fragmentation, duplication : Data files are divided into some fragments and these fragments are duplicated in terms of service level agreement.
- Second encryption: Fragments are encrypted again with a different key.
- Shuffling & Distribution: In the last stage, fragments are distributed using a shuffling method into unused memory capacities.
- Transferring Metadata to backup server: Metadata including encryption keys, shuffling, fragmentation and distribute information is sent to a supervisory server. If a disaster happens, the supervisory server will gather all information from distributed devices and performs decryption (2nd), sort & merge, inverse spatial scrambling and decryption (1st), respectively.

Pipelined Replication

This replication technique (Wood et al., 2011) aims to gain both the performance of async replication and the consistency of sync replication. In sync replication, processing cannot continue until replication is completely finished at the backup site. Whereas, in async replication, after storing data in the local storage the process can be started. The result can be replied to the client, and then the writes are replicated to the backup site in an epoch. Pipelined replication performs replication and process in parallel as in the following scenario.

- Scenario of usage: The client sends a request to the web server. The web server processes the requests, then sends data to the local database in the primary data center. At this stage, the writes are flushed in the remote backup site, and the process operation can be performed in parallel. However, the reply to the client can be committed only after receiving the Ack from the backup site. Therefore, Pipeline replication facilitates replication procedure, and also guarantees the writes protection.

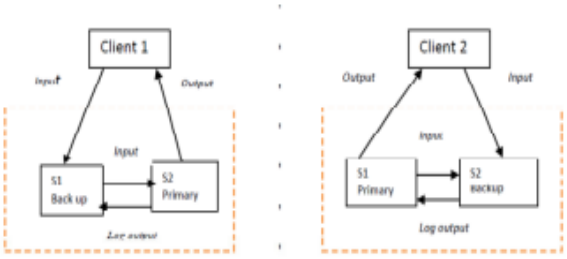
Scale Up/Down

Sometimes, performing functions with high priority can decrease money loss or even increase the revenue in the event of a disaster. Priority of services can be defined by some different features such as service level agreement, and the amount of revenue and urgent needs. After a natural disaster occurs in an area, cloud service providers are faced with flooding service requests. In this case, service providers have to manage their existent users' services and also handle new user requests. Service providers must satisfy existent users and should serve to new customers as much as possible. In (Nakajima et al., 2013), a management engine has been introduced for carrier networks. In case of a large scale natural disaster (like earthquakes), this system uses a DR scenario by scaling up resources for the high-priority services (e.g., voice communication) and scaling down allocated resources to low-priority service (e.g., video on-demand).

Dual-Role Operation

For increasing utilization of recourses, (Aghdaie and Tamir, 2003) introduces a simple technique. As shown in diagram, in this technique each host can operate as the primary host for some applications and can also be the backup host for some other applications. In this architecture, clients send their requests to the backup host first, then the backup host transmits those requests to primary host. After processing, primary host sends a log to the backup and finally reply to the clients. When a failure happens, the primary host becomes unavailable, and backup host has to handle the requests of the failed host. However, this technique cannot guarantee a good service

restoration by itself, because backup site must share the resources between its own requests and redirected requests.



Main challenges and related solutions

Challenges	Solutions	Technique
Dependency	Local backup	• Using a Linux box at the customer premises
Cost	Scale up/down,	• Allocating resources to high priority services.
	Dual-Role operation	• Hiring and running idle physical nodes on the secondary site
Failure prediction and detection	Resource Management , GRB	• Prediction and replacement of risky hardware • Using monitoring unit
Security	SDDB	• Using encryption, scrambling and shuffling techniques
Replication latency	Pipelined replication	• Performing replication and process operations in parallel
Data storage	IPCS	• Using an inter private cloud
Lack of redundancy	GRB, IPCS	• Multiple backup