Disaster Recovery in Cloud Computing: A Survey

PROBLEM DEFINITION:

Disaster recovery is a persistent problem in IT platforms. This problem is more crucial in cloud computing, because Cloud Service Providers (CSPs) have to provide the services to their customers even if the data center is down, due to a disaster. In the past few years, researchers have shown interest to disaster recovery using cloud computing, and a considerable amount of literature has been published in this area. However, to the best of our knowledge, there is a lack of precise survey for detailed analysis of cloud-based disaster recovery. To fill this gap, this paper provides an extensive survey of disaster recovery concepts and research in the cloud environments. We present different taxonomy of disaster recovery mechanisms, main challenges and proposed solutions. We also describe the cloud-based disaster recovery platforms and identify open issues related to disaster recovery

DESIGN THINKING:

Cloud computing becomes more popular in large-scale computing day by day due to its ability to share globally distributed resources. Users can access to cloud-based services through Internet around the world. The biggest IT companies are developing their data centers in the five continents to support different cloud services. The total value of the global cloud computing services market revenues is expected to reach about $241 billion by the end of 2020 (Reid et al., 2011). Rapid development in cloud computing is motivating more industries to use variety of cloud services (Arean, 2013), for instance near to 61% of UK businesses are relying on some kinds of cloud services (White paper, 2013). However, many security challenges have been raised, such as risk management, trust and recovery mechanisms which should be taken into account to provide business continuity and better user satisfaction. Disasters, either manmade or natural, can lead to expensive service disruption. Two different disaster recovery(DR) models can be used to prevent failure in a network or CSPs : Traditional and cloud-based service models. Traditional model can be used as either dedicated infrastructure or shared approach. Based on speed and cost, customers can choose the appropriate model. In dedicated approach, an infrastructure is assigned to one customer, so both cost and speed is high. On the other hand, in the shared model (we can also call it distributed approach) an infrastructure is assigned to more multiple users. This approach decreases both cost and speed of recovery. As shown in Figure 1, cloud computing is a way to gain both dedicated and shared model benefits. It can serve DR with low cost and high speed.

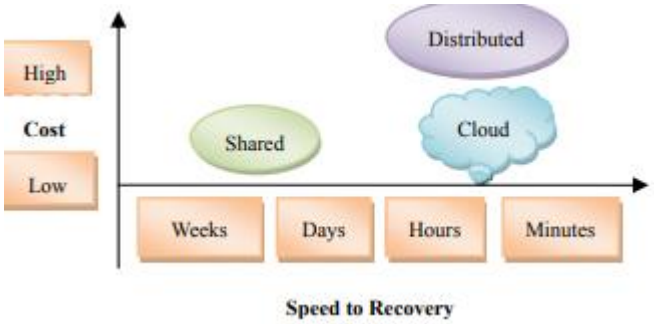

Table 1 shows a comparison between these three DR categories in terms of different features. Cloud computing decreases data synchronization between primary and backup site, minimizes different kinds of cost while increases independency between users' infrastructure and their DR systems

Table 1. Disaster recovery models (Alhazmi and Malaiya, 2013)

DR model	Data synchronization	Independency	Initial Cost	Ongoing cost	Cost of potential disasters
Dedicated	High	Low	High	Depends	High
Distributed	Medium	High	Medium	Depends	High
Cloud	Low	High	Low	Depends	Low

According to IBM research (IBM white paper, 2012), only 50% of disasters in IBM are because of weather and the rest are because of other causes. For instance, such as cut power lines, server hardware failures and security breaches. Hence, DR is not only a mechanism for natural events, but also for all severe disruptions in cloud systems. Organizations and businesses can use DR services which are served by cloud service providers. Using these services, data protection and service continuity are guaranteed for customers at different levels. Table 2 shows different DR services offered by IBM. In addition, one critical issue in DR mechanisms is that how can cloud providers tolerate disaster to prevent data lost and service disruption of their own data, infrastructure and services. In this paper we investigate both challenges and solutions for DR mechanism in cloud provider's point of view. For enterprises, the main goal of DR is business continuity which means resuming back services online after a disruption. Recovery time objective (RTO) and Recovery Point Objective (RPO) are two important parameters which all the recovery mechanisms try to improve. By minimizing RTO and RPO business continuity can be achieved. RTO is the time duration between disruption till restoration of service, and RPO denotes the amount of data lost after a disaster. Failover delays consist of 5 steps depending on the level of backup (Alhazmi and Malaiya, 2013): S1: Hardware setup S2: OS initiation time S3: Application initiation time S4: Data/process state restoration time S5: IP switching time Therefore, RPO and RTO can be defined as:

$$RPO \propto \frac{1}{\mathrm{Fb}}$$

Where Fb is Frequency of backup.

$$TO = fraction\ of\ RPO + \sum_{\mathrm{j\ min}}^{\mathrm{S5}} \mathrm{T_j}$$

Table 2. IBM different DR service level

IBM SmartCloud recovery service level	Recovery time	Description
Gold	1 minute	For mission-critical applications
Silver	30 minutes	For rapid recovery
Bronze	6 to 24 hours	Assisted failover and failback

The rest of this paper is organized as follows: In the section 2 cloud computing has been introduced briefly. In the section 3 we discuss cloud-based DR in detail. In section 4 and section 5 we investigate main challenges in DR mechanisms and some proposed solutions, respectively. It is followed by section 6 discussing somecloud-based DR systems will be introduced. In the section 7, the Open issues have been investigated. Finally, the paper ends with the proposed overall DR procedure and conclusion.