International Journal of Multidisciplinary Research and Development

Volume: 2, Issue: 6, 300-304
June 2015
www.allsubjectjournal.com
e-ISSN: 2349-4182
p-ISSN: 2349-5979
Impact Factor: 3.762

S. Suganya
M.phil Full Time Research Scholar, Department of Computer Science, Vivekanandha College of Arts and Sciences for Women, Namakkal, Pin Code-637205 TamilNadu, India.

S. Dhanalakshmi
Head of the Department, Assistant Professor, Department of Computer Science and Applications, Vivekanandha College of Arts and Sciences for Women, Namakkal, Pin Code-637205 TamilNadu, India.

Correspondence:
S. Suganya
M.phil Full Time Research Scholar, Department of Computer Science, Vivekanandha College of Arts and Sciences for Women, Namakkal, Pin Code-637205 TamilNadu, India.

Evaluation of disaster recovery in cloud computing

S. Suganya, S. Dhanalakshmi

Abstract
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. Many businesses rely on Disaster Recovery (DR) services to prevent either manmade or natural disasters from causing expensive service disruptions. Unfortunately, current DR services come either at very high cost, or with only weak guarantees about the amount of data lost or time required to restart operation after a failure. In this work, we argue that cloud computing platforms are well suited for offering DR as a service due to their pay-as-you-go pricing model that can lower costs, and their use of automated virtual platforms that can minimize the recovery time after a failure. By using disaster recovery as a service one can handle these disasters and can recover data fast with low cost. As in other techniq Cloud computing, Disaster recovery techniques, Traditional disaster recovery, Disaster recovery as a service.ues DR as a Service doesn't need any initial payment to use it provides pay on use method.

Keywords: Cloud computing, Disaster recovery techniques, Traditional disaster recovery, Disaster recovery as a service.

1. Introduction
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. Rapid development in cloud computing is motivating more industries to use variety of cloud services, for instance near to 61% of UK businesses are relying on some kinds of cloud services. 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.

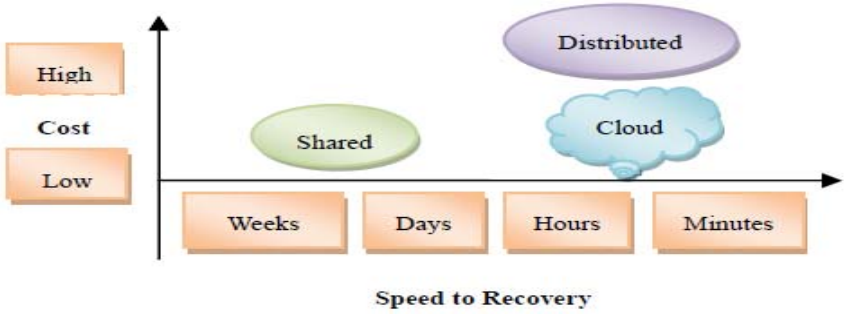

Fig 1: Comparison between traditional and cloud DR models

2. Traditional Disaster Recovery

Traditional disaster recovery was developed by share group which are divided into 6 tiers.

Tier 0: no offsite data that means there is no disaster recovery plan and no saved data. To recover data it may take weeks and it is unsuccessful.

Tier 1: data backup without hotsite that means data is taken backup by offsite not by hotsite. To retrieve the data that is taken backup is time taken process. By not having their own redundant servers it is time taking process to locate and configure appropriate systems.

Tier 2: data backup with hotsite that means organizations maintain data backup as well as hotsite it is the fastest process. By having a hot backup site when disaster occurs we can run applications at stand by servers.

Tier 3: instead of taking backup by physical media it provides an electronic vault so that backup data is network accessible to hot site. As hotsite backup is cost effective it is better to access it by network.

Tier 4: point in time copies means that organization maintains more timely point in time backup of crucial data is network accessible to host site.

Tier 5: transaction integrity means that transactions are consistent between production systems and recovery sites. So, there should be no loss of data.

Traditional disaster recovery offers better RPO's and RTO's. Traditional geographic redundancy is an alternative technique that has data centers having sufficient equipment to store data when backup is made. To assure rapid recovery time objective it is necessary to deploy same type or hardware or software to geo-redundant sites. Virtualization simplifies traditional disaster recovery by relaxing compatibilities requirements by deploying hardware on recovery site. Hardware configuration on recovery site should be equal to primary site to carry the entire traffic load served by impacted site acceptable service quality, reliability and latency.

3. Disaster Recovery as a Service

Disaster recovery as a service is an upcoming service as a nomenclature of cloud computing. It is a low cost service when compared to traditional disaster recovery. It is flexible in replicating physically or virtually. It provides application consistent recovery for some working applications like SQL server. It has pre-built options for virtual recovery environments including security, network connectivity and server failover when continuously replication among servers. When disaster occurs we can take backup and we can run our applications on service provided by disaster recovery until we get backup to primary site. Disaster recovery as a service to replicate critical servers and data centre infrastructure in cloud.

Fig 2: Disaster recovery as a service

Disaster recoveries as a service is free or pay on use offer. When incompatibilities are occurred due to software changes then breaking of DRaaS in cloud may occur.

The architecture of DRaaS is defined by three models.

From Cloud: when the primary application or data is in cloud and backup or recovery site is in private data centre.

In cloud: when both primary site and recovery site are in cloud.

To cloud: when the application is in primary data centre and backup or recovery site is in cloud.

To test the recovery processes sandboxes are used and they test without disrupting running application. It is only accessible to only system administrator. Solutions are pre-packaged services that provide a standard DR Failover to a cloud environment that you can buy on a pay-per-use basis with varying rates based upon your recovery point objective (RPO) and recovery time objective (RTO).

4. DR Requirements

This section discusses the key requirements for an effective DR service. Some of these requirements may be based on business decisions such as the monetary cost of system downtime or data loss, while others are directly tied to application performance and correctness.

Recovery Point Objective (RPO): The RPO of a DR system represents the point in time of the most recent backup prior to any failure. The necessary RPO is generally a business decision—for some applications absolutely no data can be lost (RPO=0), requiring continuous synchronous replication to be used, while for other applications, the acceptable data loss could range from a few seconds to hours or even days.

Recovery Time Objective (RTO): The RTO is an orthogonal business decision that specifies a bound on how long it can take for an application to come back online after a failure occurs. This includes the time to detect the failure, prepare any required servers in the backup site (virtual or physical), initialize the failed application, and perform the network reconfiguration required to reroute requests from the original site to the backup site so the application can be used. Depending on the application type and backup technique, this may involve additional manual steps such as verifying the integrity of state or performing application specific data restore operations, and can require careful scheduling of recovery tasks to be done efficiently [7]. Having a very low RTO can enable business continuity, allowing an application to seamlessly continue operating despite a site wide disaster.

Performance: For a DR service to be useful it must have a minimal impact on the performance of each application being protected under failure-free operation. DR can impact performance either directly such as in the synchronous replication case where an application write will not return until it is committed remotely, or indirectly by simply consuming disk and network bandwidth resources which otherwise the application could use.

Consistency: The DR service must ensure that after a failure occurs the application can be restored to a consistent state. This may require the DR mechanism to be application specific to ensure that all relevant state is properly replicated to the

backup site. In other cases, the DR system may assume that the application will keep a consistent copy of its important state on disk, and use a disk replication scheme to create consistent copies at the backup site.

Geographic Separation: It is important that the primary and backup sites are geographically separated in order to ensure that a single disaster will not impact both sites. This geographic separation adds its own challenges since increased distance leads to higher WAN bandwidth costs and will incur greater network latency. Increased round trip latency directly impacts application response time when using synchronous replication. As round trip delays are limited by the speed of light, synchronous replication is feasible only when the backup site is within 10s of kilometers of the primary. Asynchronous techniques can improve performance over longer distances, but can lead to greater data loss during a disaster. Distance can especially be a challenge in cloud based DR services as a business might have only coarse control over where resources will be physically located.

5. Mechanisms for Cloud DR

While cloud computing platforms already contain many useful features for supporting disaster recovery, there are additional requirements they must meet before they can provide DR as a cloud service.

Network Reconfiguration: For a cloud DR service to provide true business continuity, it must facilitate reconfiguring the network setup for an application after it is brought online in the backup site. We have previously proposed how a cloud infrastructure can be combined with virtual private networks (VPNs) to support this kind of rapid reconfiguration for applications that only communicate within a private business environment. Public Internet facing applications would require additional forms of network reconfiguration through either modifying DNS or updating routes to redirect traffic to the failover site. To support any of these features, cloud platforms need greater coordination with network service providers.

Security & Isolation: The public nature of cloud computing platforms remains a concern for some businesses. In order for an enterprise to be willing to fail over from its private data center to a cloud during a disaster it will require strong guarantees about the privacy of storage, network, and the virtual machine resources it uses. Likewise, clouds must guarantee that the performance of applications running in the cloud will not be impacted by disasters affecting other businesses.

VM Migration & Cloning: Current cloud computing platforms do not support VM migration in or out of the cloud. VM migration or cloning would simplify the failback procedure for moving an application back to its original site after a disaster has been dealt with.

This would also be a useful mechanism for facilitating planned maintenance downtime. The Remus system has demonstrated how a continuous form of VM migration can be used to synchronize both memory and disk state of a virtual machine to a backup server. This could potentially allow for full system DR mechanisms that allow completely transparent failover during a disaster. To support this, clouds must expose additional hypervisor level functionality to their customers, and migration techniques must be optimized for WAN environments.

6. Benefits of the Cloud in DR

Under current pricing schemes, cloud based DR services will not see much benefit when used for applications that require true "hot" standby servers since this can significantly raise the cost during normal operation. However, for applications that can tolerate recovery times on the order of 200 seconds (a typical VM startup time in the EC2 cloud), substantial savings can be found by utilizing low cost servers while replicating state in ordinary conditions and powerful ones only after a disaster occurs.

Cloud DR services may be able to obtain additional economic benefits by multiplexing a single replication server for multiple applications, further lowering the cost of resources under normal operation. For applications with a loose RPO, the cloud can provide even greater benefits by only initiating the replication service a few times a day to create periodic backups.

Cloud computing can facilitate disaster recovery by significantly lowering costs:

- The cloud's pay-as-you go pricing model significantly lowers costs due to the different level of resources required before and during a disaster.
- Cloud resources can quickly be added with fine granularity and have costs that scale smoothly without requiring large upfront investments.
- The cloud platform manages and maintains the DR servers and storage devices, lowering IT costs and reducing the impact of failures at the disaster site.

The benefits of virtualization, while not necessarily specific to cloud platforms, still provide important features for disaster recovery:

- VM startup can be easily automated, lowering recovery times after a disaster.
- Virtualization eliminates hardware dependencies, potentially lowering hardware requirements at the backup site.
- Application agnostic state replication software can be run outside of the VM, treating it as a black box.

These characteristics can simplify the replication and deployment of resources in a cloud DR site, and enable business continuity by reducing recovery times.

7. Disaster Recovery Challenges

In this section we investigate some common challenges of DR in cloud environments.

7.1 *Dependency*

One of the disadvantages of cloud services is that customers do not have control of the system and their data. Data backup is on premises of service providers as well. This issue makes dependency on CSPs for customers (such as organizations) and also loss of data because of disaster will be a concern for customers. Dependency also creates another challenge which is the selection of a trusted service provider.

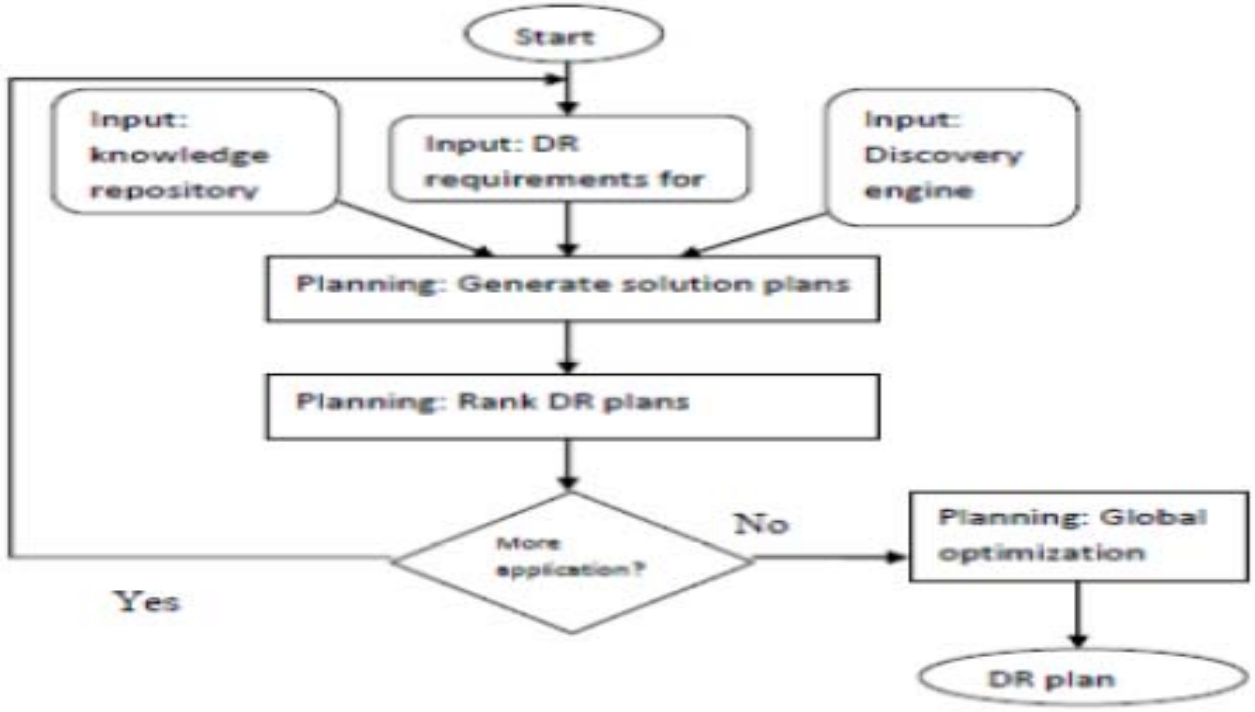

Fig 3: ENDEAVOUR flowchart

7.2 *Cost*

It is obvious that one of the main factors to choose cloud as a DR service is its lower price. So, cloud service providers always seek cheaper ways to provide recovery mechanisms by minimizing different types of cost. The yearly cost of DR systems can be divided in three categories:

- Initializing cost: amortized annual cost
- Ongoing cost: storage cost, data transfer cost and processing cost
- Cost of potential disaster: Cost of recovered disasters and also cost of unrecoverable disasters.

7.3 *Failure Detection*

Failure detection time strongly affects on the system downtime, so it is critical to detect and report a failure as soon as possible for a fast and correct DR. On the other hand, in multiple backup sites there is a major question: How to distinguish between network failure and service disruption.

7.4 *Security*

As mentioned before, DR can be created by nature or can be human-made. Cyber-terrorism attack is one of human-made disasters which can be accomplished for many reasons. In this case, protection and recovery of important data will be a main goal in DR plans beside of system restoration.

7.5 *Replication Latency*

DR mechanisms rely on replication technique to make backups. Current replication techniques are classified into two categories: synchronous and asynchronous (Ji *et al.*, 2003). However, both of them have some benefits and some flaws. Synchronized replication, guarantees very good RPO and RTO, but it is expensive and also can affect on system performance because of large overhead. This issue is more serious in multi-tier web applications, because it can significantly increase Round Trip Time (RRR) between primary and backup site. On the other hand, a backup model adopted with async replication is cheaper and also system suffers low overhead, but the quality of DR Service will be decreased. Therefore, trading off between cost, performance of the system and also replication latency is an undeniable challenge in cloud disaster solutions.

7.6 *Data Storage*

Business database storage is one of the problems of enterprises which can be solved by cloud services. By increasing of cloud usage in business and market, enterprises need to storage huge amount of data on cloud-based storages. Instead of conventional data storage devices, cloud storage service can save money and is also more flexible. The architecture of a cloud storage system includes four layers: physical storage, infrastructure management, and application interface and access layer. In order to satisfy applications and also to guarantee the security of data, computing has to be distributed but storage has to be centralized. Therefore, storage single points of failure and data loss are critical challenges to store data in cloud service providers.

7.7 *Lack of Redundancy*

When a disaster happens, primary site becomes unavailable and secondary site has to be activated. In this case, there is no ability to sync or async replication in a backup site but data and system states only can be stored locally. It is a serious threat to the system. This issue is temporary and will be removed after recovery of the primary site. However, to achieve the best DR solutions, especially in high availability services (such as business data storage), it is better to consider all risky situations.

8. Open Issues and Future Directions

In the last sections, we described the main properties and challenges of DR systems. Then, some related solutions and systems have been introduced. However, some issues still require more effort to reach a worthy level of DR mechanisms in cloud computing. In this section, we introduce some open and related issues in this area:

8.1 *Maximizing Resource Utilization*

Cloud customers pay for DR resources only after a disaster happens. However, these resources must be always available when needed. Since disasters are usually scarce, the revenue of the DR servers is less. Therefore, CSPs need the ways to both increase the utilization and revenue of DR servers and also guarantee DR services, simultaneously.

8.2 *Correlated Failures*

Occurring disaster in a specific area can lead to vast service interruption, and consequently, many customers have to be recovered by CSPs. In this case, it is possible that related servers cannot be able to handle all the customers. So, it can be critical to multiplex customers of the same area in different servers. One major challenge in this case is how to distribute customers between cloud servers to minimize correlated failure risk with respect to required QoS for each server and also cloud SLA (Wood *et al*., 2010).

8.3 *Privacy and Confidentiality*

In the event of disaster, the private data centers of enterprises would be failover by cloud environments. So, one critical issue is that cloud must guarantee confidentiality of data and privacy of resources which are used for DR. On the other hand, the cloud has to guarantee the performance of applications would not be affected by other disasters happened to other enterprises.

8.4 *Failover and Failback Procedure*

Failover and failback procedure are two important stages in DR mechanism. Failover procedure is performed to automatically switch over to a backup site whenever the current active site becomes unavailable. In the event of a disaster, failover procedure excludes failed resources and redirect workloads to a secondary site using a specialized load balancer. Client-transparent procedure and fast IP failover requirements are two main challenges in this issue. On the other hand, after passing disaster, application control has to be reverted to the original site. For this purpose, bidirectional state replication must be supported by the DR mechanism. A portion of data may be lost because of the disaster in the primary site and also new data will be created in the backup site. Therefore, one major challenge is that how to determine new and old data which must be resynchronized to the primary site.

8.5 *Disaster Monitoring*

Since failure tolerance is necessary to deliver expected QoS, it will be essential to determine which processes are operational and which crashed in the cloud systems. In the case of disaster, the sooner failure detection in either primary site or backup site leads to better RTO. So, the challenge is how should the status of cloud be monitored and how a disaster can be detected in its early stages.

8.6 *Resource Scheduling*

The number of cloud-based services are increasing day by day and so has increased the complexity of cloud infrastructures. Hence, resource scheduling is a critical issue in the modern cloud environments. This issue is more crucial for cloud-base DR platforms since they face unpredictable arrival rate and have to consider a variety of catastrophic situations. Building on this, more efficient resource scheduling techniques are needed in order for current DR platforms to be also optimal.

9. Conclusion

As cloud computing is becoming very important in day to day life and every company is based on cloud computing. We have argued that cloud computing platforms are an excellent match for providing disaster recovery services due to their pay-as-you-go pricing model and ability to rapidly bring resources online after a disaster. The flexibility of cloud resources also allows enterprises to make a trade off between data protection and price to an extent not possible when using private resources that must be statically provisioned. In our ongoing work, we are developing Dr. Cloud, a prototype DR system that we can use to understand the potential for using existing cloud platforms to provide DR. This will allow us to better understand what features and optimizations must be included within the cloud platform itself, and to explore the tradeoffs between cost, RPO, and RTO in a cloud DR service.

10. References

1. Albert Greenberg, James Hamilton, David A. Maltz, and Parveen Patel. Cost of a cloud: Research problems in data center networks. In ACM SIGCOMM Computer Communications Review, Feb 2009.
2. Brenda Phillips, “Disaster Recovery”, CRC Press, Mar.2011.
3. Brendan Cully, Geoffrey Lefebvre, Dutch Meyer, Mike Feeley, Norm Hutchinson, and Andrew Warfield. Remus: High availability via asynchronous virtual machine replication. In Proceedings of the Usenix Symposium on Networked System Design and Implementation, 2008.
4. Charlotte Hiatt, “A Primer for Disaster Recovery Planning in an It Environment”, Idea Group Inc (IGI), 2000.
5. Emmanuel Cecchet, Anupam Chanda, Sameh Elnikety, Julie Marguerite, and Willy Zwaenepoel. Performance Comparison of Middleware Architectures for Generating Dynamic Web Content. In 4th ACM/IFIP/USENIX International Middleware Conference, June 2003.
6. Peter Gregory and Philip Jan Rothstein, “IT Disaster Recovery Planning For Dummies”,Wiley Publishing, Inc., Dec.2007.
7. J Peter Bruzzese, “Virtualization and Disaster Recovery”, Realtime Publishers, 2009
8. Rajkumar Buyya, Rajiv Ranjan, and Rodrigo N. Calheiros. InterCloud: Utility-Oriented Federation of Cloud Computing Environments for Scaling of Application Services. In The 10th International Conference on Algorithms and Architectures for Parallel Processing, Busan, Korea, 2010.
9. Regis J. Bates, “Disaster recovery planning: networks, telecommunications, and data communications”, McGraw-Hill, 1992.
10. Zhang Jian-hua, “Cloud Computing-based Data Storage and Disaster Recovery” at Future Computer Science and Education (ICFCSE), 2011 International Conference.
11. Cloud Application Architectures building applications and infrastructure in the cloud by O'RELLY and George Reese.
12. http://www.windstreambusiness.com/data-center-solutions/disaster-recovery/disaster-recovery-as-aservice
13. http://www.windstreambusiness.com/media/299685/draas-large.jpg