International Journal of Research in Advent Technology, Vol.2, No.5, May 2014

E-ISSN: 2321-9637

Virtualizing Disaster Recovery Management Based On Cloud Computing

Ms.Shital V. Bahale [1] , Prof. Dr.Sunil Gupta [2].

M.E.(II Year)- Department of Computer Science and Engg. P.R.M.I.T.R, Badnera-Amravati [1],

Assistant Professor- Department of Computer Science and Engg. P.R.M.I.T.R, Badnera-Amravati [2]

Email: shitalbahale@rediffmail.com[1] , sunilguptacse@gmail.com[2]

Abstract - Almost from the beginning of widespread adoption of computers, organizations realized that disaster recovery was a necessary component of their information technology plans. Business data had to be backed up, and key processes like order entry, billing, payroll and procurement needed to continue even if an organization's data center was disabled due to a disaster. Growing reliance on crucial computer systems means those even short periods of downtime can result in significant financial loss, or in some cases even put human lives at risk. Many business and government services utilize Disaster Recovery (DR) systems to minimize the downtime incurred by catastrophic system failures.

Cloud computing provides the third leg of a disaster recovery plan that is essential for business continuity. Cloud-based storage services take advantage of Internet access to deliver reliable, low-cost online storage, helping you to bounce back from a full-scale data center disaster for less than the cost of a dedicated online storage solution.

Virtualization is the means of ushering in a new, productive era of cloud computing, driven by this need for cost management and increased agility. Virtualization can also provide the basic building blocks for your cloud environment to enhance agility and flexibility. This paper delineate how virtualization of cloud computing can be used to address the concerns resulting in improved computer infrastructure that can easily be restored following a natural disaster ,reduced expenses, improved scalability, better performance and is easier to manage.

Index Terms **-** Disaster Recovery requirements; Dedicated and shared DR models; Virtualization; Cloud based DR mechanisms;

1. INTRODUCTION

A key challenge in providing DR services is to support Business Continuity (BC), allowing applications to rapidly come back online after a failure occurs. By minimizing the recovery time and the data lost due to disaster, a DR service can also provide BC, but typically at high cost. 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[1,2].

A typical DR service works by replicating application state between two data center's if the primary data center becomes unavailable, then the backup site can take over and will activate a new copy of the application using the most recently replicated data[8].

Virtualization is the foundation for an agile, scalable cloud and the first practical step for building cloud infrastructure [11]. Virtualization abstracts and isolates the underlying hardware as virtual machines (VMs) in their own runtime environment and with multiple VMs for computing, storage, and networking resources in a single hosting environment. These virtualized resources are critical for managing data, moving it into and out of the cloud, and running applications with high utilization and high availability [14].

Virtualization is managed by a host server running a hypervisor software, firmware, or hardware that creates and runs VMs[17]. The VMs are referred to as guest machines [6,9].

Virtualization also provides several key capabilities for cloud computing, including resource sharing, VM isolation, and load balancing. In a cloud environment, these capabilities enable scalability, high utilization of pooled resources, rapid provisioning, workload isolation, and increased uptime.

In this paper we explore how virtualized cloud platforms can be used to provide low cost DR solutions that are better at enabling Business Continuity. In the first section this paper discusses data recovery requirements , second section explores traditional approaches to disaster recovery then in third section describes data recovery mechanisms and fourth section describes cloud computing mechanisms for data recovery lastly it concludes with how organizations can use cloud computing to help plan for both the mundane interruptions to service—cut power lines, server hardware failures and security breaches—as well as more-infrequent disasters.

2. DATA RECOVERY 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.

2.1 Recovery point objective (RPO)

The RPO of a DR system represents the point in time of the most recent backup prior to any failure.

2.2 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.[7]. Having a very low RTO can enable business continuity, allowing an application to seamlessly continue operating despite a site wide disaster.

2.3 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.

2.4 Consistency

The DR service must ensure that after a failure occurs the application can be restored to a consistent state.

2.5 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. 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.

3. TRADITIONAL DISASTER RECOVERY APPROACHES

In traditional disaster recovery models—dedicated and shared— organizations are forced to make the trade-off between cost and speed to recovery.

3.1 Dedicated disaster recovery model

In a dedicated model, the infrastructure is dedicated to a single organization. This type of disaster recovery can offer a faster time to recovery compared to other traditional models because the IT infrastructure is duplicated at the disaster recovery site and is ready to be called upon in the event of a disaster. Although this model can reduce RTO because the hardware and software are preconfigured, it does not eliminate all delays. The process is still dependent on receiving a current data image, which involves transporting physical tapes and a data restoration process. This approach is also costly because the hardware sits idle when not being used for disaster recovery. As illustrated in Figure 1, data restoration can take up to 72 hours including the tape retrieval, travel and loading process[3].

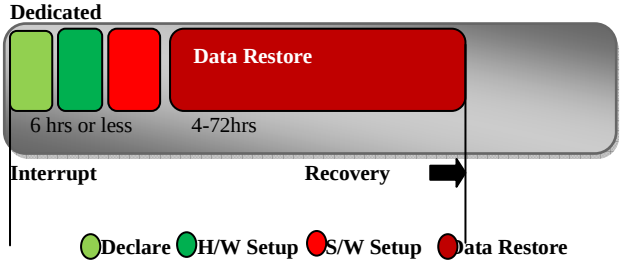

Figure 1. Time To Recovery using a Dedicated Infrastructure

3.2 Shared disaster recovery model

In a shared disaster recovery model, the infrastructure is shared among multiple organizations. Shared disaster recovery is designed to be more cost effective because the off-site backup infrastructure is shared among multiple organizations. After a disaster is declared, the hardware, operating system and application software at the disaster site must be configured from the ground up to match the IT site that has declared a disaster, and this process can take hours or even days. In addition, the data restoration process must be completed as shown in Figure 2, resulting in an average of 48 to 72 hours to recovery[3].

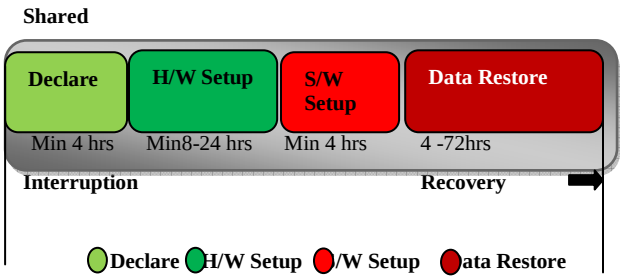

Figure 2. Time To Recovery using a Shared Infrastructure

With dedicated and shared disaster recovery models, organizations have traditionally been forced to make tradeoffs between cost and speed. As the pressure to achieve continuous availability and reduce costs continues to increase, organizations can no longer accept tradeoffs. Any downtime reflects directly on their brand image, and customers view any interruption of key applications such as e-commerce, online banking and customer self-service as being unacceptable. As a result, the cost of a minute of downtime may be thousands of dollars.

4. DR MECHANISMS

Disaster Recovery is primarily a form of long distance state replication combined with the ability to start up applications at the backup site after a failure is detected. Backup mechanisms operating at the file system or disk layer replicate all or a portion of the file system tree to the remote site without requiring specific application knowledge [6].

The use of virtualization makes it possible to not only transparently replicate the complete disk, but also the memory context of a virtual machine, allowing it to seamlessly resume operation after a failure however, such techniques are typically designed only for LAN environments due to significant bandwidth and latency requirements [4, 9].

DR services fall under one of the following categories:

4.1 Hot Backup Site

A hot backup site typically provides a set of mirrored stand-by servers that are always available to run the application once a disaster occurs, providing minimal RTO and RPO. Hot standbys typically use synchronous replication to prevent any data loss due to a disaster. This form of backup is the most expensive since fully powered servers must be available at all times to run the application, plus extra licensing fees may apply for some applications.

4.2 Warm Backup Site

A warm backup site may keep state up to date with either synchronous or asynchronous replication schemes depending on the necessary RPO. Standby servers to run the application after failure are available, but are only kept in a “warm” state where it may take minutes to bring them online. This slows recovery, but also reduces cost.

4.3 Cold Backup Site

In a cold backup site, data is often only replicated on a periodic basis, leading to an RPO of hours or days. In addition, servers to run the application after failure are not readily available, and there may be a delay of hours or days as hardware is brought out of storage or repurposed from test and development systems, resulting in a high RTO. It can be difficult to support business continuity with cold backup sites, but they are a very low cost option for applications that do not require strong protection or availability guarantees.

5. MECHANISMS FOR CLOUD DISASTER RECOVERY

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.

5.1 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[10]. 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.

5.2 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[12,13].

5.3 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[4][16].

Cloud computing offers an attractive alternative to traditional disaster recovery. "The Cloud" is inherently a shared infrastructure a pooled set of resources with the infrastructure cost distributed across everyone who contracts for the cloud service. This shared nature makes cloud an ideal model for disaster recovery. Even when we broaden the definition of disaster recovery to include more mundane service interruptions, the need for disaster recovery resources is sporadic. Since all of the organizations relying on the cloud for backup and recovery are very unlikely to need the infrastructure at the same time, costs can be reduced and the cloud can speed recovery time[5].

Because the server images and data are continuously replicated, recovery time can be reduced dramatically to less than an hour, and, in many cases, to minutes—or even seconds. However, the costs are more consistent with shared recovery.

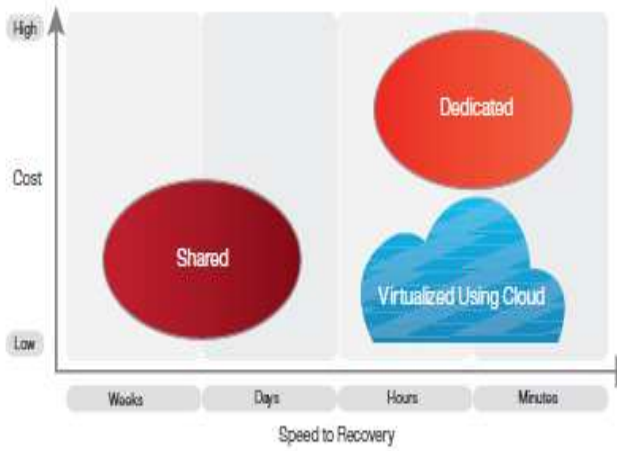

Figure 3. Cloud based approach to disaster recovery

Cloud computing based on virtualization, takes a very different approach to disaster recovery. With virtualization, the entire server, including the operating system, applications, patches and data is encapsulated into a single software bundle or virtual server[15]. This entire virtual server can be copied or backed up to an offsite data center and spun up on a virtual host in a matter of minutes.

Since the virtual server is hardware independent, the operating system, applications, patches and data can be safely and accurately transferred from one data center to a second data center without the burden of reloading each component of the server. This can dramatically reduce recovery times compared to conventional (non-virtualized) disaster recovery approaches where servers need to be loaded with the OS and application software and patched to the last configuration used in production before the data can be restored.

The cloud shifts the disaster recovery trade-off curve to the left, as shown below. With cloud computing (as represented by the red arrow), disaster recovery becomes much more cost-effective with significantly faster recovery times.

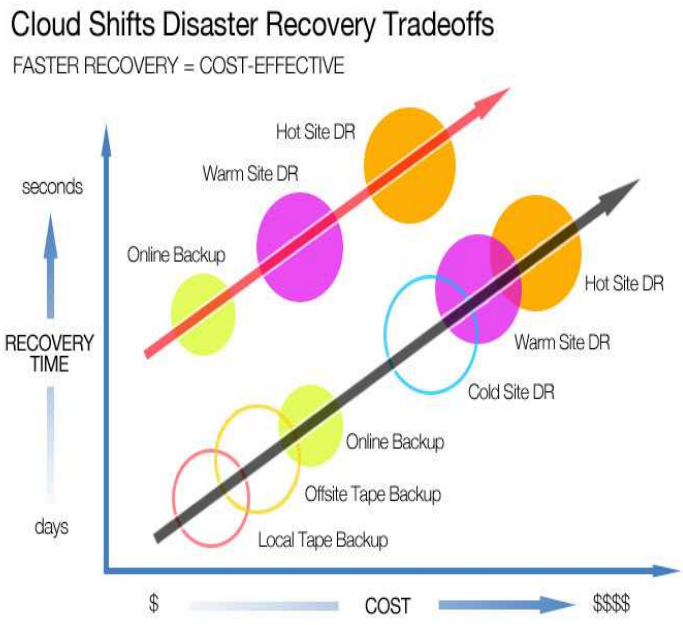

Figure 4. Cloud Disaster Recovery Trade-offs

The cloud makes cold site disaster recovery antiquated. With cloud computing, warm site disaster recovery becomes a very cost-effective option where backups of critical servers can be spun up in minutes on a shared or private cloud host platform.

One of the most exciting capabilities of disaster recovery in the cloud is the ability to deliver multi-site availability. SAN replication not only provides rapid failover to the disaster recovery site, but also the capability to return to the production site when the DR test or disaster event is over.

One of the added benefits of disaster recovery with cloud computing is the ability to finely tune the costs and performance for the DR platform. Applications and servers that are deemed less critical in a disaster can be tuned down with less resources, while assuring that the most critical applications get the resources they need to keep the business running through the disaster.

6. CONCLUSION

With pay-as-you-go pricing and the ability to scale up as conditions change, cloud computing can help organizations meet the expectations of today's frenetic, fast paced environment where IT demands continue to increase but budgets do not. Virtualization also eliminates hardware dependencies, potentially lowering hardware requirements at the backup site.

By coordinating disaster recovery and data back-up, data loss can be reduced and reliability of data integrity improved. In future focus is on using fault tolerant server hardware within virtualized cloud environments to reduce management complexity to sustain the high service levels.

Virtualizing disaster recovery start up can also be automated for lowering recovery times after a disaster.

REFERENCES

[1] 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.

[2] 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.

[3]Virtualizing disaster recovery using cloud computing,IBM Global Technology Services Thought Leadership White Paper January 2013.

[4] 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.

[5] Albert Greenberg, James Hamilton, David A. Maltz, and Parveen Patel.Cost of a cloud: Research problems in data center networks. In ACM SIGCOMMComputer Communications Review, Feb 2009.

[6] Kimberley Keeton, Cipriano Santos, Dirk Beyer, Jeffrey Chase, and John Wilkes. Designing for Disasters. Conference On File And Storage Technologies,2004.

[7] Kimberly Keeton, Dirk Beyer, Ernesto Brau, Arif Merchant, Cipriano Santos,and Alex Zhang. On the road to recovery:restoring data after disasters.European Conference on Computer Systems, 40(4), 2006.

[8] Tirthankar Lahiri, Amit Ganesh, Ron Weiss, and Ashok Joshi. Fast-Start:quick fault recovery in oracle. ACM SIGMOD Record, 30(2), 2001.

[9] Vmware high availability. http://www.vmware.com/products/high-availability/.

[10] T. Wood, A. Gerber, K. Ramakrishnan, J. Van der Merwe, and P. Shenoy.The case for enterprise ready virtual private clouds. In Proceedings of the Usenix Workshop on Hot Topicsin Cloud Computing (HotCloud), San Diego, CA, June 2009.

[11] Marston S., Li Z., Bandyopadhyay S., Zhang J. , Ghalsasi A. (2010) „Cloud computing - The business perspective" Decision Support Systems [online] 51 (2011) 176–189 .

[12] Golden; B. (2009), 'Capex vs. Opex: Most People Miss the Point About Cloud Economics'.

[13] Fellows, W. (2009), 'The State of Play: Grid, Utility, Cloud' - available at http://old.ogfeurope.eu/uploads/Industry%20Expert%20Group/FELLOWS_CloudscapeJan09-WF.pdf

[14] Marc Malizia White Paper On Virtualization+Cloud Equals Perfect Storm For Disaster Recovery Services **Version1.1,** 20 March 2013.

[15] White paper on Server Virtualization and Cloud Computing by Stratus Technologies November, 2011.

[16] 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.

[17] Buyya R., Broberg J., Goscinski A.M. (2011) Cloud Computing: Principles and Paradigms. New York: John Wiley & Sons.