Maximum Availability via Clustering and DRP in Enterprise Environment

Bakht Shah [1], Adnan Ahmed [2]

[1]MSCS SZABIST
Karachi Pakistan

Abstract: *It is important for organizations to have a standardized strategy or approach for data and applications recovery in case of hardware and/or software failures, human errors and/or natural disasters. The down-time for any planned outages such as maintenance, patching and up-gradation should be minimized and be based on a standard organizational policy. A large number of organizations have the necessary hardware and software resources, but most of them lack any detailed/in-depth disaster recovery plans (DR) and scalability strategies. In this paper, the author puts forward recommendations based on real-time implementations for an effective higher availability (HA), Scalability and disaster recovery strategy. The author also examines and discusses solutions provided by different vendors for higher availability, scalability and disaster recovery.*

Keywords*: High Availability Cluster (HAC), Maximum Availability (MA), High Availability (HA), Real Application Cluster (RAC), Disaster Recovery (DR), planned and unplanned downtime, and modes.*

1. INTRODUCTION

Maximum Availability can be defined as minimization of downtime rather than the complete elimination of downtime. High availability and disaster recovery (DR) have long been like life insurance for businesses. In its simplest form, a high-availability plan encompasses three aspects:

- Resilience
- Recoverability
- Continuous operation

Maximum availability of data means shielding of the business processes from the impact of and availability of data in case of hardware faults, storage failure and disaster occurrence when the data holding servers crashes.

Clusters and disaster recovery setup are generally marketed as the only ways to provide maximum availability and scalability for the applications that run on them.

2. DOWN TIME AND THEIR CAUSES

2.1. What is Downtime

In simple words, we can define the downtime as the time in which the business data is not available to customer/end user for processing due to system crash, site failure and other such reasons *[1]*

2.2. Types of Downtime and Their Impact on Business

Typically there are two causes of downtime:

- Unplanned downtime.
- Planned downtime.

2.2.1. Unplanned Downtime:

If a system or a database is unavailable to customers for certain period of time without informing them in advance, this downtime is called unplanned downtime. This may be due to hardware failure, power failure, operating system failure, storage failure, human error, data corruption or complete site failure.

2.2.2. Planned Downtime

The server or database is unavailable to customers may be due to hardware maintenance, OS maintenance or up-gradation, database patching or up-gradation.

2.2.3. Impact on Business

In both type of downtime the customers can't access business data if there is no implementation of a maximum availability solution for hardware and software.

2.3. Recommendation to Minimize Unplanned Downtime

Figure 2.3.0, shows the root-causes of downtime and recommendations to minimize it.

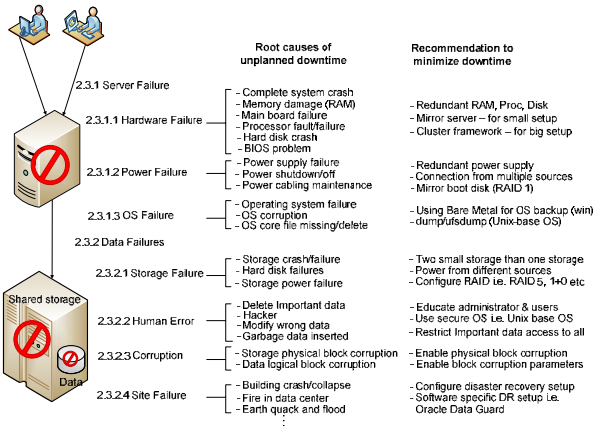


Figure 2.3.0 – Root causes and recommendations to minimize unplanned downtime

The recommendations that mentioned to minimize unplanned downtime in above figure and their pros and cons are discussed in the subsequent section.

2.3.1. Server or System Failure

In this type of failure, the server is completely down or have a critical problem. In this case the data loss depends on whether the data is stored on local hard drives or shared storage. There are several computer failures describe as below

2.3.1.1. Hardware Failure

Hardware Failure may be complete system damage, memory damage (RAM), main board failure, hard disk crash, processor fault or BIOS problem. In such type of failures, the system not function and system recovery depends on the problem.

Recommendation: If a company's business data is too much critical and needs to be operated 24/7 where as they don't have enough budget to implement expensive maximum availability solution then it is recommended to use the Mirror server.

2.3.1.2. Power Failure

All small and high specification servers require power to operate, if there is a single power supply in a server and it is connected with a single power source as well, in case of a power failure the server is not available for operation.

Recommendation: Approximately all Storage types have redundant power supplies but mostly main power is provided from a single source to all power supply, in such a case when there is a power trip or maintenance is required the storage goes down and no one can access application data even when the system is up and running.

2.3.1.3. Operating System Failure

When operating system failure occurs or some core OS file goes missing then generally they manually reinstalls operating system, apply maintenance pack (windows) /packages and patches (Unix-base), apply business application and restore data or database from backup server.

These manual processes takes more time and are required expert sources which can highly affect the business.

Recommendation: The recommended approach should be used to obtain the operating system backup and restore it automatically rather than manually reinstall in case of failure with minimum time.

It is recommended for Microsoft windows (MS 2003 server, XP) to take the whole operating system backup using Bear Metal in which we can completely restore OS to target host automatically(without human interaction).

Dump (ufsdump for Solaris) backup type should be used for Unix-base operating system to take the entire file system backup which includes OS kernel and core OS configuration files.

2.3.2. Data Failure

Author recommends using two storage redundant storage to save data in case of one storage failure. In figure 2.3.1 shown, that storage A failure occurred and nodes transparently connect with storage B to retrieved the same data was stored on Storage A.

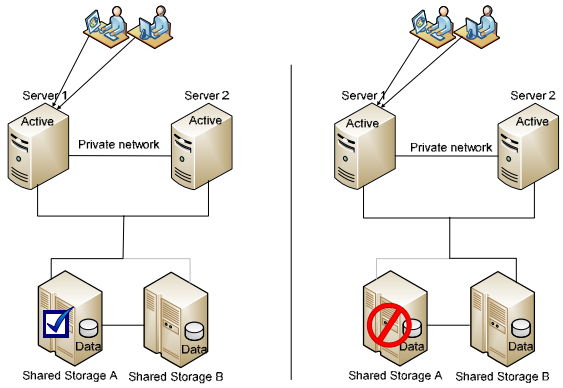


Figure 2.3.1 – Storage redundancy

2.3.2.1. Human Error

We have seen that some times by mistake human being delete business data folder or directory or makes changes in configuration files which may cause data loss.

Recommendation: It is recommended to use Unix-base operating system (UNIX, Linux, AIX, Solaris, HP-UX, HP-tru64, Mac OSX etc) in production environment and set file level security and restriction. Assign exact privileges (read, write and execute) on each directory and file to restrict user access.

2.3.2.2. Corruption

Data is stored physically on hard drive and logically on storage unit. When storage units have any physically issues or data is inserted into application or database in unacceptable order is called physical and logical block corruption.

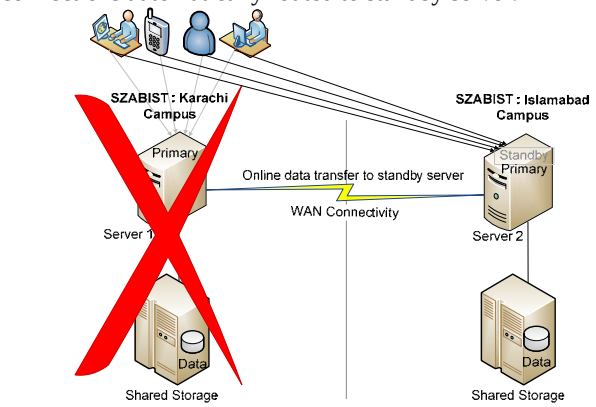
Recommendation: Many hardware and software vendors provide this functionality but unfortunately we do just 'by default' installation. It is recommended to enable the relevant parameters and make the document, which briefly describe how to fix different types of block corruption.

2.3.2.3. Site Failure

It can be defined as complete data centre crash, whole building crash due to earth quack, fire, and flood or due to any other reason.

Recommendation: In today's technologically advanced business environment, all most all of the hardware, applications and database vendors provide the solution for disaster recovery planning.
The world leading database vendor "Oracle Corporation" provides disaster recovery solution for their database known as Oracle Data Guard *[2]*.

Figure 2.3.2 shows primary site crash along with customer's connections automatically routed to standby server.



2.3.2 – Disaster recovery / Oracle Data Guard

2.4. Recommendation to Minimize Planned Downtime

Figure 2.4.0 shows the root causes which required downtime and recommendations to minimize these downtimes.

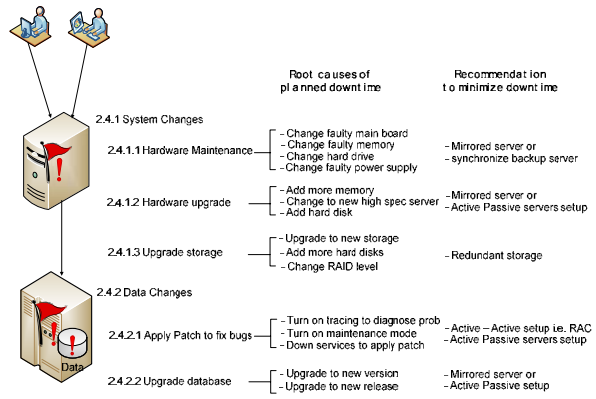


Figure 2.4.0 – Recommendation to minimize planned downtime

The recommendations mentioned to minimize planned downtime in the above figure, the detail for each point have been discussed in subsequent sections.

2.4.1. System Changes

It has always been a requirement to change hardware part or upgrade hardware to new high specification server to provide better performance. This include the following causes

2.4.1.1. Hardware Maintenance

If the application and database runs on a single server and data are stored on local disk (not shared storage) and there is a need to perform system maintenance then planned downtime is required. The system which needs maintenance cannot be put down due to the applications need to be available 24/7.
Recommendation: The business is highly affected whenever planned or unplanned downtime occurs. For such a type of businesses clustering configuration is best solution. If the business is too critical with respect to time but do not have enough budget then mirrored server setup should be used.

2.4.1.2. Hardware Upgrade

Usually hardware upgrades takes place whenever the current system specification reports drop in performance or system do not have the capability to permit new huge load.

Recommendation: For all type of business setup (big, average and small), it is recommended to maintain synchronize backup server of core database and application systems. If you need planned downtime to upgrade or replace the core system while your application don't allow you then provide synchronize backup server to customer temporarily whereas upgrade your core system or server.

2.4.1.3. Upgrade Storage or Changing RAID type

Adding more space (hard drives), changing or upgrading the complete storage or changing the RAID type of existing mount point need planned downtime.

Recommendation: It is recommended to make use of two storages for high redundancy, availability and to minimize both type of downtime.

2.4.2. Data Changes

All software vendors regularly release new version and patches to provide transparent functionalities, new features together with fixing the existing bugs in the software. Same like hardware maintenance, we need downtime to upgrade or apply any patch to fix a bug.

2.4.2.1. Apply Patch to Fix Bugs

Applying a patch is a pre-plan activity and it generally needs planned downtime.

Recommendation: For big setup it is recommended that switch to active-active cluster environment which provides high performance, scalability and high availability. Here the author is considering the leading database vendor "Oracle" solution that provides the above functionalities known as Oracle Real Application Cluster (RAC). In RAC environment you can apply patch on one instance while other instances will operate and customer can access data *[3]*.

2.4.2.2. Upgrade database

If the application and database are running in a cluster environment either in active-active or active-passive setup then first test the upgrade process in Test setup then go ahead with the process on Production setup.

Recommendation: Make sure that you have a cold or consistent backup of all application and database directories before upgrading the software. Try to conduct the pre-testing in an appropriate environment i.e. if the application is currently running in active-active setup then tests the upgrade version in active-active setup (not in single server setup or active-passive setup).

3. CLUSTER TECHNOLOGY

Clustering is a general terminology that describes a group of two or more separate servers operating as a "single unit".

3.1. Maximum Availability via Clustering

Cluster environments provide an environment where, in the case of any single hardware or software failure in the cluster, application services and data are recovered automatically (without human intervention) and quickly "faster than server boot". This is done by taking advantage of the existence of redundant servers in the cluster and redundant server-storage path.

High availability clusters (HAC) improve availability of applications by failing them over or switching them over in a group of systems as opposed to High Performance Clusters which improve performance of applications by allowing them to run on multiple systems simultaneously.

Figure 3.1.0 demonstrates the logical layout of two nodes cluster with redundant switches, disks and private interconnect.

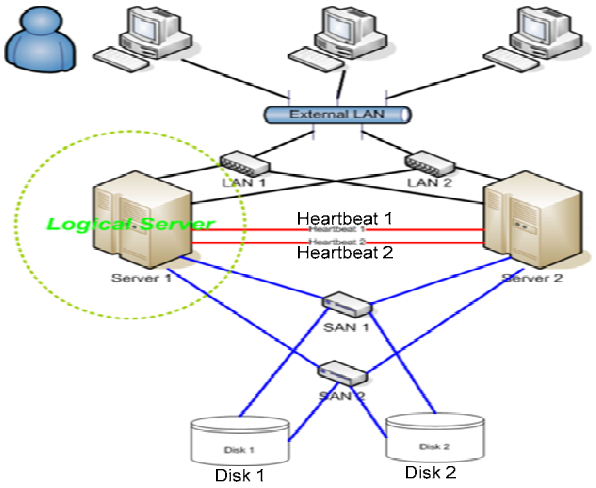


Figure 3.1.0 – Two nodes cluster logical layout

Reference:http://en.wikipedia.org/wiki/High-availability_cluster

3.2. Top Clustering Solutions

There are a number of hardware and software vendors in the market throughout the world who provides clustering solutions. The following are a few of the vendors' solutions.

VERITAS Cluster Server (also known as VCS) is a High-availability cluster software, for Unix, Linux and Microsoft Windows computer systems, created by VERITAS Software (now part of Symantec). It provides application cluster capabilities to systems running databases, file sharing on a network, electronic commerce websites or other applications. VCS is one of the few products in the industry that provides both high availability and disaster recovery across all major operating systems while supporting 40+ major application or replication technologies out of the box.

Other vendors' products include Linux-HA, Red Hat Cluster Suite, HP Service Guard, Sun Cluster, Microsoft Cluster Server (MSCS), IBM Tivoli System Automation for Multi-platforms (SA MP), IBM HACMP and Oracle Real Application Cluster (RAC). Oracle RAC allows multiple computers to run the Oracle RDBMS software simultaneously while accessing a single database, thus providing a clustered database. *[4]*

3.3. Cluster types

The most common size for a High Availability (HA) cluster is two nodes, since that's the minimum required to provide redundancy, but many clusters consist of many more, sometimes dozens of nodes. Such configurations can sometimes be categorized into one of the following models:

- Active/Active

- Active/Passive
- N+1:
- N+M:
- N-to-1:
- N-to-N:

3.4. Further Improving Cluster Availability

In a cluster setup, shared storage is a mandatory component and actual business data is stored on it. In cluster, we have redundant nodes (more than one node), and typically one storage with RAID configuration. When one node completely crashes or fails, another node in the cluster takes over and there is no discontinuity in the business.

But if there is a single shared storage and more than two hard drive fails or complete storage crash occurs then business cannot continue and data may be lost. In this special case the recovery depends on damage and backup strategy.

It is highly recommended to take up-to-date backup on regular intervals and keep them in separate location (another building and city).

In Pakistan, there are a number of big companies running cluster setup but they do not have implementation or strategy for disaster recovery.

Cluster components (nodes & storage) are mostly configured in same data-centre & building. It is good approach to configure remote backup and disaster recovery setup by using the same cluster setup (not a single node) on remote site to avoid any performance and connectivity issue in time of disaster; when the entire customer and internal user connect with standby server.

Figure 3.1.1 shown, four nodes cluster setup, storage redundancy and disaster recovery setup.

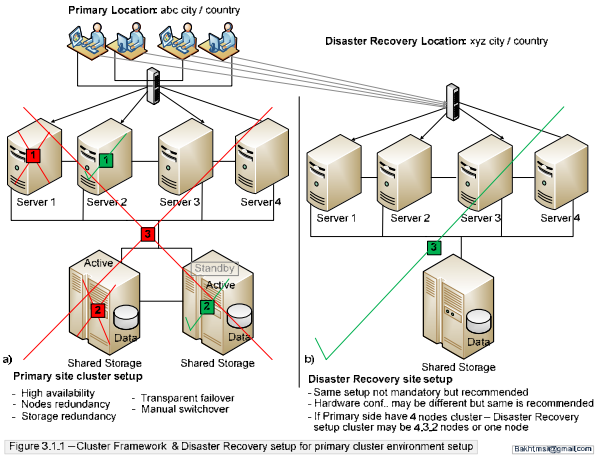


Figure 3.1.1 – Cluster Framework & Disaster Recovery setup for primary cluster environment setup

In the above figure the red line indicates crash, while green line indicates that the new system or storage is available for processing having same data of crash node or storage.

The red line marked with 1 indicates that a node in the cluster become down due to hardware or software failure but there is no impact because other node in the cluster distribute the load and start services for that node.
Second red line (marked with 2), indicating that active storage fail or crash but as there are two redundant storages configured so the standby storage become active and start work to provides data to cluster nodes.

The red line marked with 3 shows that all cluster nodes and storage have crashed due to disaster or building crash but data is safe on remote side and customer connections have been transparently routed to remote side setup.

4. Oracle Data Guard and Network Implication

Oracle Data Guard is Oracle's data protection and disaster recovery solution. Data Guard ensures high availability, data protection, and disaster recovery for enterprise data. It provides a comprehensive set of services that create, maintain, manage, and monitor one or more standby databases to enable production Oracle databases to survive disasters and data corruptions. Data Guard maintains these standby databases as transaction ally consistent copies of the production database.

4.1. Protection modes and variation

There are three Data Guard data protection modes available. One mode can configure at a time. Selection of protection mode totally depends on business requirements for data availability against user demands for response time and performance.

The explanation of each protection mode as below *[5]:*

- Maximum Protection Mode
- Maximum Availability Mode
- Maximum Performance Mode

4.2. Network Link Measurement and Recommendations

One of the most important questions that are how much network bandwidth is required by Data Guard?
The answer is simple. It depends on how busy the production database is *[6]*.

I am recommend that use network tools to measure the exact upload and download bandwidth size of a network link rather than verbally relying.

Measure bandwidth and link performance

It is extremely important to calculate the authentic bandwidth for the following setup:

- Disaster recovery planning
- Real time replication

- Distributed databases
- Streaming
- Remote point in time recovery
- Load balancing

Sometimes, most of the organizations, organize environment by ignoring the actual link bandwidth size which create performance problem later-on when number of user increases.

CONCLUSION

This independent study discussed unlikely events that can and do occur; power outages, disappearing buildings and resources, redundant hardware failures, OS failure or crashes, the number and frequency of storms, and other formerly unthinkable scenarios show that disasters do occur. All of these are root causes of business downtime.

The solutions provided by different hardware and software vendors for maximum availability, scalability and disaster recovery have been discussed and further added recommendations to ensure business availability.

Business processes and technology changes frequently in today's environment, so it is recommended that maintain redundancy implantation on hardware, software, application and database levels and disaster recovery (DR) must change accordingly. DR planning is an ongoing task, not a one-time goal, so capital budgets must include items for maintaining DR sites and equipment.

It is expected that this study would prove helpful in understanding the downtimes and their root causes. The reader would then be able to understand the different methods to minimize these downtimes and improve maximum availability and disaster recovery planning.

The future work in this Independent Study would be to explore all the methods that should be used to improve the business availability and disaster recovery planning which will completely cover up downtime and make sure no data is lost in any critical circumstances.

REFERENCES

[1] Oracle® Database High Availability Overview. "Overview of High Availability.", Part Number: B14210-02, Chap:1, page: 3. Internet: http://database.in2p3.fr/doc/oracle/Oracle_Database_10_Release_2_(10.2)_Documentation/server.102/b14210/overview.htm. July 2006. [April 3, 2009].

[2] Vivian Schupmann ,Oracle Data Guard Concepts and Administration. "Oracle Data Guard." Part Number: B14239-01, Chap: 5, Page: 12. Internet: http://youngcow.net/doc/oracle10g/server.102/b14239/title.htm. June 2005. [April 3, 2009].

[3] Wikipedia. "High-availability cluster." Internet: http://en.wikipedia.org/wiki/High-availability_cluster. March 12, 2009. [April 3, 2009].

[4] Michael Smith, Oracle. "Building a Highly Available and Disaster- Proof Architecture." Page: 19. Internet: http://www.oracle.com/technology/deploy/availability/pdf/oow06/S281208_Smith.pdf. [April 3, 2009].

[5] Vivian Schupmann. "Data Guard: Concepts and Administration." Part Number: B14239-01, Chap: 1, page 7. Internet: http://youngcow.net/doc/oracle10g/server.102/b14239.pdf. June 2005. [April 3, 3009].

[6] Ray Dutcher and Ashish Ray. "Network Bandwidth Implications of Oracle Data Guard." Internet: http://www.oracle.com/technology/deploy/availability/htdocs/dataguardnetwork.htm. [April 3, 2009].