Redbooks Paper

Jerome Calves

Veritas Storage Foundation for Oracle 9i RAC and DB2 UDB V8 on IBM @server pSeries

This IBM® Redpaper describes our installation and testing of Veritas Storage Foundation for Oracle 9*i* RAC and IBM DB2® Universal Database™ V8 on AIX® 5L™, running on IBM @server® pSeries®. It describes a high availability solution implemented for demonstration purposes at the IBM Benchmark Center in Montpellier, France.

Archived

© Copyright IBM Corp. 2006. All rights reserved.

Introduction

This IBM Redpaper describes a Proof of Concept (PoC) performed in 2004[1] by IBM and Veritas at the Montpellier Benchmark Center. We installed and tested the solution with Oracle 9i Real Application Clusters and IBM DB2 Universal Database v8 on AIX. These products were deployed on IBM @server pSeries. In it we describe:

- The different layers and products in the solution
- The HW architecture of the platform
- The installation of VERITAS Storage Foundation for Oracle and DB2 databases
- High availability testing scenarios

In “Resources” on page 27, we provide URLs for the documents that helped us perform this PoC.

To secure and recover critical data and to ensure continuous business processes, enterprises need to have a robust, available, and manageable infrastructure. Information must be available at all times, and downtime, whether planned or unplanned, is an impossible option.

The system must provide:

- Continuous application services and data access
- Secure universal access to enterprise data
- Application performance that grows with the organization
- A way to contain management costs

Clustering technologies help meet these challenges by coordinating interconnected servers to enhance application and data availability, scalability, and manageability. Clusters are conceptually attractive to users because of their potential to solve some of the most significant problems in computing:

- **Server failure:** If a server or application crashes, another server in the cluster can take over its application, providing nearly continuous service to clients.
- **Network or I/O path failures:** If network or I/O links fail, access to applications and data can continue on alternate paths.
- **Application growth:** If application demands go beyond the ability of a single server, additional servers can be added, and the workload can be shared.
- **Management cost:** As the number of enterprise applications grows, clusters can provide a flexible platform for running multiple applications while allowing the entire cluster to be managed as a single system.

A number of products, such as GPFS and HACMP™, provide solutions with high availability.

Deploying and using *VERITAS Storage Foundation for Databases* on pSeries and AIX is another way to take advantage of High Availability functions on powerful systems.

[1] Original title: *Overview, Installation and High Availability Test Scenario of Veritas Storage Foundation Solution for Oracle 9i RAC and DB2 Universal Database v8 on IBM pSeries and AIX 5L*, by Jerome Calves, IBM Advanced Technical Support – PSSC, IBM Montpellier. Jerome_Calves@fr.ibm.com

Veritas Storage Foundation for Oracle RAC: What it is

Here we explain the characteristics of VERITAS Storage Foundation for Oracle RAC.

Industry-leading high availability technology

VERITAS Storage Foundation for Oracle RAC includes VERITAS Cluster Server technology. VERITAS Cluster Server provides a low latency interconnect to improve the performance and scalability of the Oracle CacheFusion technology. This technology runs on top of low cost, well understood, Ethernet connections; no proprietary interconnects are needed.

VERITAS Storage Foundation for Oracle RAC also includes I/O Fencing technology that prevents database corruption if a cluster split brain occurs. The VERITAS I/O Fencing implementation is robust, protecting against interconnect failures, interconnect time-outs, node hangs, and a variety of other failure mechanisms.

Clustered file system for simplified management

VERITAS Storage Foundation for Oracle RAC provides a cluster file system (CFS) that has been certified by Oracle to work with Oracle RAC (9i and 10g). A CFS simplifies the installation and ongoing management of a RAC database. The VERITAS cluster file system provides excellent I/O performance, so you get the manageability of a file system with no sacrifices in performance. Installation is simple because a single set of Oracle binaries can be installed and used on all nodes. Shared Oracle binaries also simplify the application of Oracle patches. This reduces the repetitive installation and updating work found in an environment without a CFS. Adding new Oracle data files is as simple as creating a new file, not the complicated process of adding a new raw device.

The VERITAS CFS provides excellent I/O performance. All nodes in the cluster can directly read and write to the storage devices; there is no write-shipping being done over the inter-node connections. VERITAS takes advantage of the Oracle Disk Manager API for increased performance and manageability. The VERITAS CFS uses direct I/O, bypassing file system locking and buffering. VERITAS CFS enables ease of management without sacrificing I/O performance.

Flexible storage provisioning to adapt to changing business

VERITAS Storage Foundation for Oracle RAC uses volume management technology to enable physical disks to be managed as logical devices or volumes. The clusterized volume manager included in this solution provides flexible mirroring and striping configurations. Multiple connections from servers to storage devices further improve performance and availability. Online administration allows administrators to reconfigure the storage topology while RAC is up and running.

The VERITAS Storage Foundation 4.0 for Oracle RAC is built on VERITAS Cluster Server (VCS), VERITAS Cluster Volume Manager (CVM), VERITAS Cluster Filesystem (CFS), and VERITAS Storage Foundation, which includes the base VERITAS Volume Manager (VxVM) and file systems (VxFS). The different layers are sorted as follows:

- Oracle 9i RAC
- VCS RAC Extension
- VERITAS Storage Foundation for Oracle RAC
- Cluster File System (CFS)
- Cluster Volume Manager (CVM)
- Hardware

Key terms and definitions

Here we list several key terms for VERITAS Storage Foundation for Oracle RAC.

RAC

Real Application Clusters (RAC) is a parallel database environment that takes advantage of the processing power of multiple, interconnected computers. A cluster contains two or more computers, also known as nodes or servers. In RAC environments, all nodes concurrently run Oracle instances and execute transactions against the same database.

RAC coordinates each node's access to the shared data to provide consistency and integrity. Each node adds its processing power to the cluster. RAC serves as an important component of a robust high availability solution. A properly configured Real Application Clusters environment can tolerate failures with minimal downtime and interruption to users. With clients accessing the same database on multiple nodes, failure of a node does not completely interrupt access, as clients accessing the surviving nodes continue to operate. Clients attached to the failed node simply reconnect to a surviving node and resume access. Recovery after a failure in a RAC environment is far quicker than when using a failover database, because another instance is already up and running. Recovery is simply a matter of applying outstanding redo log entries from the failed node.

Veritas Cluster Server (VCS)

In the Storage Foundation for Oracle RAC environment, VCS functions as a director of operations. It controls start-up and shutdown of the component layers of RAC. In the Storage Foundation for Oracle RAC configuration, the RAC service groups run as parallel service groups. VCS does not attempt to migrate a failed service group, but it can be configured to restart it on failure. VCS also notifies users of any failures. Storage Foundation for Oracle RAC provides specific agents for VCS to operate in a Storage Foundation for Oracle RAC environment, including CVM, CFS and Oracle.

Oracle Disk Manager (ODM)

The Oracle Disk Manager (ODM) is a standard API specified by Oracle for performing database I/O. For example, when Oracle wishes to write, it calls the odm_io function. ODM improves both performance and manageability of the file system. The VERITAS-provided implementation of ODM improves performance by providing direct access for the database to the underlying storage without passing through the actual file system interface. This means the database sees performance that is equivalent to using actual devices. The administrator sees the storage as easy-to-manage file systems, including the support of resizing data files while in use.

Cluster File System (CFS)

The VERITAS Cluster File System is an extension of the industry standard VERITAS Filesystem (VxFS). CFS allows the same file system to be simultaneously mounted on multiple nodes. Unlike other clustered file systems, CFS is a true SAN file system. All I/O takes place over the storage area network.

Cluster Volume Manager (CVM)

Cluster Volume Manager is an extension of VERITAS Volume Manager (VxVM), the industry standard storage virtualization platform. CVM extends the concepts of VxVM across multiple nodes. Each node sees the same logical volume layout, and more importantly, the same state of all volume resources. In a Storage Foundation for Oracle RAC cluster, all storage is managed with standard VxVM commands from one node in the cluster. All other nodes immediately recognize any changes in disk group and volume configuration with no

interaction. CVM supports performance enhancing capabilities such as striping, mirroring, and mirror break-off (snapshot) for off-host backup.

CVM architecture

Cluster Volume Manager uses master-slave architecture. One node in the cluster acts as the configuration master for logical volume management, and all others are slaves. Any node can take over as master if the existing master fails. The CVM master is established on a per-cluster basis. Since CVM is an extension of VxVM, it operates in a very similar fashion. The volume manager configuration daemon, vxconfigd, maintains the configuration of logical volumes. Any changes to volumes are handled by vxconfigd, which updates the operating system at the kernel level when the new volume state is determined. For example, if a mirror of a volume fails, it is detached from the volume and the error is passed to vxconfigd, which then determines the proper course of action, updates the new volume layout, and informs the kernel of a new volume layout.

CVM extends this behavior across multiple nodes using master-slave architecture. Changes to a volume are propagated to the master vxconfigd. The vxconfigd process on the master pushes these changes out to slave vxconfigd processes, each of which in turn updates the local kernel. CVM does not impose any write locking between nodes. Each node is free to update any area of the storage. All data integrity is the responsibility of the upper application. From an application perspective, logical volumes are accessed identically on a stand-alone system as on a CVM system. CVM imposes a “Uniform Shared Storage” model. All systems must be connected to the same disk sets for a given disk group. Any system unable to see the entire set of physical disks as other nodes for a given disk group cannot import the disk group. If a node loses contact with specific disk, it is excluded from participating in the use of that disk. CVM uses Group Atomic Broadcast (GAB) and Low Latency Transport (LLT) for transport of all its configuration data.

Mechanism overview

Storage Foundation for Oracle RAC provides a complete I/O and communications stack to support Oracle RAC. It also provides monitoring and management of instance start-up and shutdown. The following section describes the overall data and communications flow of the Storage Foundation for Oracle RAC stack.

Data stack overview

Figure 1 details the overall data flow from an instance running on a server to the shared storage. The various Oracle processes making up an instance (such as DB Writer, Log Writer, Checkpoint, Archiver, and others) read and write data to storage using the I/O stack shown in the diagram. Oracle communicates through the Oracle Disk Manager (ODM) interface to the VERITAS Cluster File System (CFS), which in turn accesses the storage using VERITAS Cluster Volume Manager (CVM).

Archive

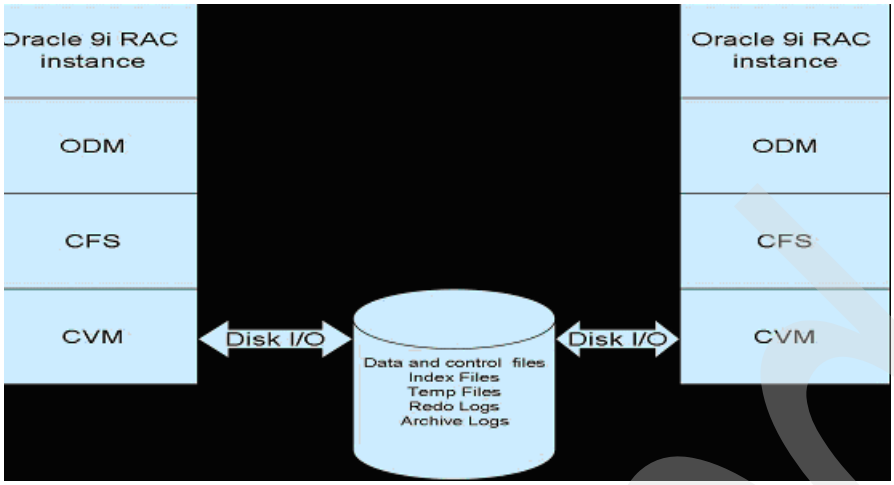

Figure 1 Data stack overview

Communications stack overview

Figure 2 shows the data stack as well as the communications stack. Each of the components in the data stack requires communications with its peer on other systems to function properly. RAC instances must communicate to coordinate protection of data blocks in the database. ODM processes must communicate to coordinate data file protection and access across the cluster. CFS coordinates metadata updates for file systems, and finally CVM must coordinate the status of logical volumes and distribution of volume metadata across the cluster. VERITAS Cluster Server (VCS) starts and stops components in the Storage Foundation for Oracle RAC stack and provides monitoring and notification on failure. VCS must communicate status of its resources on each node in the cluster. For the entire system to work, each layer must reliably communicate.

Figure 2 also shows Low Latency Transport (LLT) and the Group Membership Services/Atomic Broadcast (GAB), which make up the communications package central to the operation of Storage Foundation for Oracle RAC. During an operational steady state, the only significant traffic through LLT and GAB is due to Lock Management and Cache Fusion, while the traffic for the other data is relatively sparse.

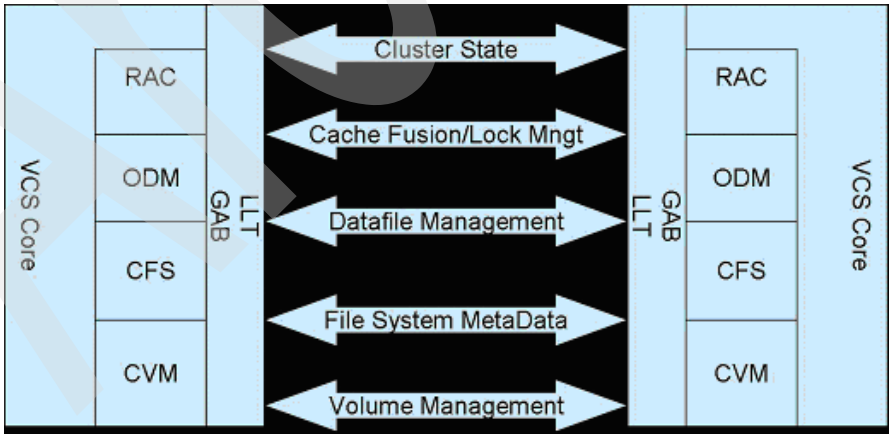

Figure 2 Communications stack overview

Storage Foundation for DB2 overview

VERITAS Storage Foundation for DB2 is an integrated suite of data, storage, and system management technologies that optimize performance, availability, and manageability of DB2 UDB databases. What does this mean for your critical DB2 environments?

Enhanced DB2 performance

VERITAS Quick I/O and Cached Quick I/O provide raw device performance with the ease of file system manageability. Storage Mapping maps database objects to physical disks, enabling performance hot spots to be detected and corrected. Software RAID, Dynamic Multi-pathing, and extent-based allocation increase I/O through put and optimize performance.

High availability for DB2

Storage Checkpoint and Rollback enable faster point-in-time database recovery from operational errors. Using FlashSnap, copy resource-intensive tasks (backups, reporting, etc.) off-host to maintain production database availability. Storage Foundation/HA for DB2 includes VERITAS Cluster Server for automatic database monitoring and failover.

Powerful DB2 storage management

Online management allows database resizing, performance tuning and storage reconfiguration without impacting users. Intelligent Storage Provisioning (ISP) allows policy-based deployment of database storage using standardized templates. Quality of Storage Service (QoSS) segments data and helps ensure that only essential files reside on high-end arrays.

Key terms and definitions

Here is a list of key terms for VERITAS Storage Foundation for DB2.

VERITAS Volume Manager

VERITAS Volume Manager, a key component of VERITAS Storage Foundation for DB2, is a storage virtualization tool and allows the management of physical disks as logical devices, called volumes. A volume is a logical device that appears to data management systems as a physical disk partition device. By allowing volumes to span multiple disks, VERITAS Volume Manager enables the management of virtual storage pools rather than actual physical disks. By using VERITAS Volume Manager as an abstraction layer that virtualizes storage and makes that storage easier to manage, users (as well as operating systems and applications) overcome the physical restrictions imposed by hardware disk devices.

Through its support of RAID, VERITAS Volume Manager protects against disk and hardware failure. Additionally, Volume Manager provides features that enable fault tolerance and fast recovery from disk failure. VERITAS Volume Manager also provides easy-to-use online disk storage management for tasks such as dynamically configuring disk storage while the system is active, and ensuring that data remains available.

Quick I/O and Cached Quick I/O

For performance reasons, database administrators sometimes choose to implement the database on character/raw devices, known in DB2 UDB terms as DMS (database-managed storage) devices. VERITAS File System's **Quick I/O feature** provides performance virtually equal to raw devices, but still maintains File System manageability benefits. DBAs can use Quick I/O files as DMS devices to get the best of both worlds. In addition the **Cached Quick I/O feature** uses the server's extra physical memory as a second-level cache for caching DB2

UDB data blocks. In a 32-bit DB2 UDB environment, this feature can improve DB2 UDB performance significantly beyond that of raw device performance.

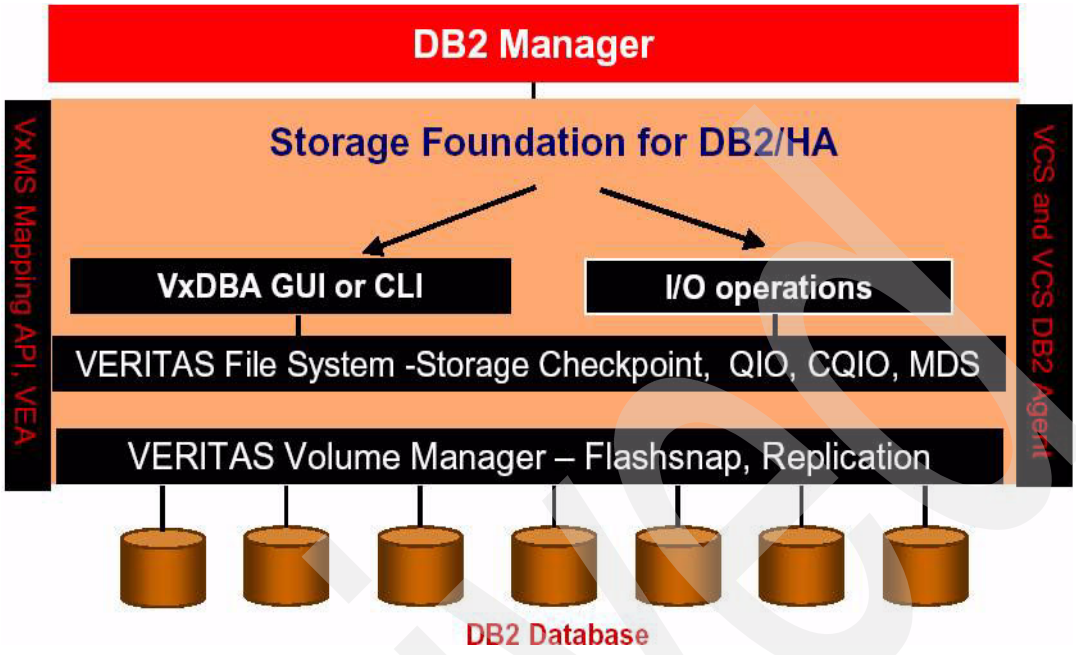

Figure 3 Mechanism overview of Storage Foundation for DB2-HA

Hardware and software platform

Here is the platform for the installation:

System information

- Hardware Platform: 2 X IBM IBM @server pSeries 630
- Operating System and Version: AIX v5.2 ML 3
- Disk space: 140 GB
- Number of CPUs: 4
- Memory: 32 GB

Storage information

- Storage Array Type: ESS 2105-800
- Firmware version: 2.4.0.245
- HBA Firmware Version: 02E01871
- HBA Driver Version: 5.2.0.30
- Fibre Channel Switches Being Used: Direct attachment

Network information

- Private Network Connection: 2 x 1GB Ethernet (used for heartbeat)
- Administration network: 100 MB Ethernet

Archive

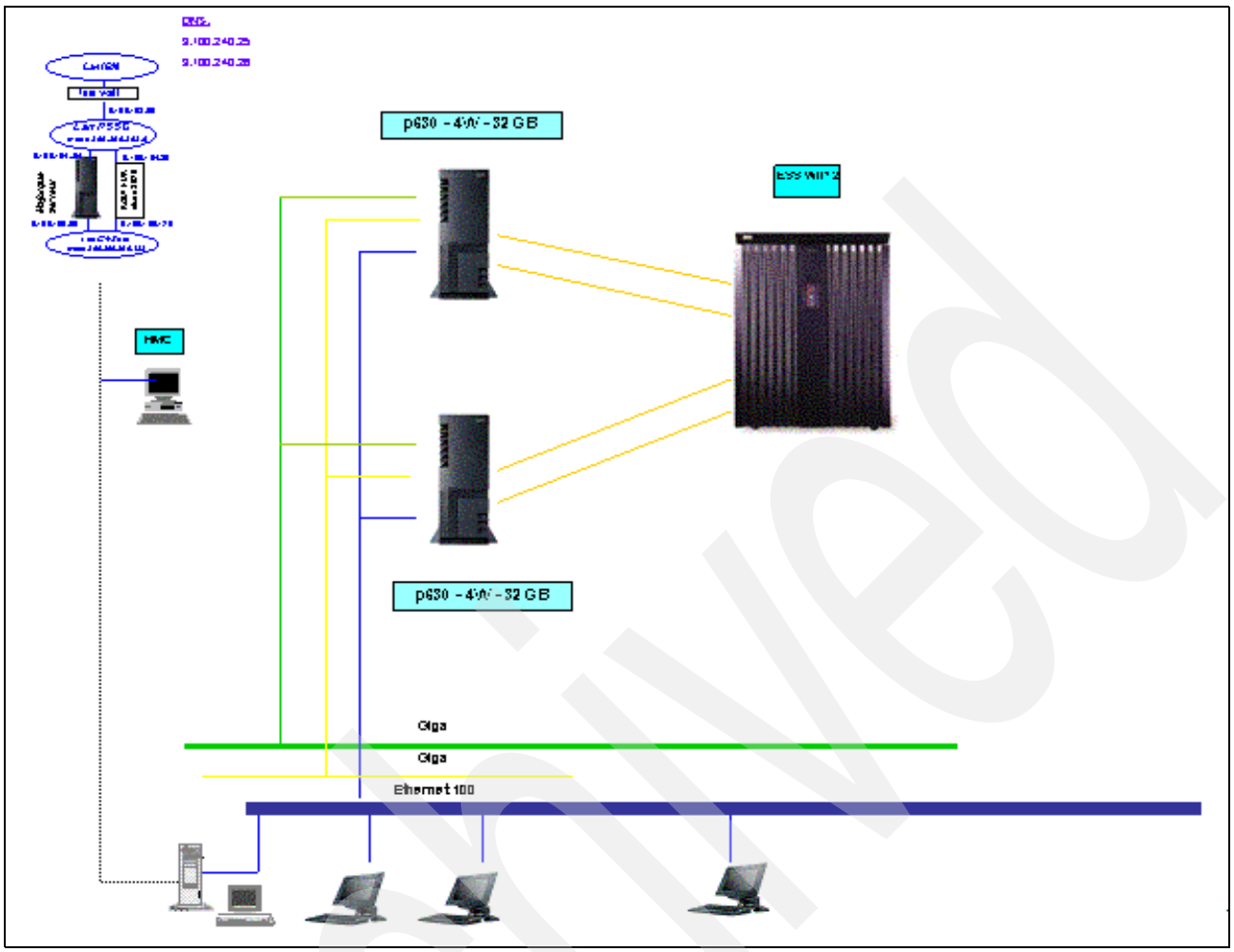

Figure 4 Configuration

Installation

This section describes two installations: VERITAS Storage Foundation for Oracle RAC and VERITAS Storage Foundation for DB2.

Installing Veritas Storage Foundation for Oracle RAC

This section describes the installation of the Storage Foundation for Oracle RAC solution.

- ☐ Cluster Parameters
  - ► Cluster heartbeat 2 x 1 GB links with cross cable
  - ► Cluster heartbeat interface: similar on both servers. en4 and en5 forced to 1000_FullDuplex.
  - ► I/O fencing disks: hdisk5, hdisk6, hdisk7
  - ► Cluster ID: 2
  - ► Cluster node 1: Node A: mcbprod1
  - ► Cluster node 2: Node B: mcbprod2
  - ► Oracle data disk group: oradatadg (shared data)

Archived

- Oracle binary disk group: orabindg (shared binary)
- SRVM shares same disk group as binary
- Files systems size, naming and attributes are same as production environment

☐ Setting the PATH Variable

On each system, add these directories to your PATH environment variable using the following commands:

If you use the Bourne Shell (sh or ksh):

```
$ PATH=/usr/sbin:/sbin:/usr/bin:/usr/lib/vxvm/bin:\
/opt/VRTSvxfs/sbin:/opt/VRTSvcs/bin:/opt/VRTS/bin:\
/opt/VRTSvcs/vxfen/bin:/opt/VRTSvcs/rac/bin:/opt/VRTSob/bin:\
$PATH; export PATH
```

If you use the C Shell (csh or tcsh):

```
% setenv PATH /usr/sbin:/sbin:/usr/bin:/usr/lib/vxvm/bin:\
/opt/VRTSvxfs/sbin:/opt/VRTSvcs/bin:/opt/VRTS/bin:\
/optVRTSvcs/vxfen/bin:/opt/VRTSvcs/rac/bin:/opt/VRTSob/bin:$PATH
```

Note: For root user, do not define paths to a cluster file system in the LIBPATH variable. For example, define $ORACLE_HOME/lib in LIBPATH for user oracle only.The path defined to /opt/VRTSob/bin is optional, required if you install the optional fileset, VERITAS Enterprise Administrator.

☐ Setting the MANPATH Variable

Set the MANPATH variable to enable viewing manual pages:

If you use the Bourne Shell (sh or ksh):

```
$ MANPATH=/usr/share/man:/opt/VRTS/man; export MANPATH
```

If you use the C Shell (csh or tcsh):

```
% setenv MANPATH /usr/share/man:/opt/VRTS/man
```

☐ Obtain license key for SFRAC:

Access this Web site, using the information provided in the License Key Request Form, to generate a license key for SFRAC:

http://vlicense.veritas.com

Ensure all nodes can communicate with each other via RSH.

☐ Perform the following on every node in the cluster:

1. Login as the root user.
2. # echo “+” > ~/.rhosts

☐ Mount the CD.

☐ Launch the installsfrac installer, preferably from mcbprod1 (Node A).

Perform the following from the install host:

1. # cd /cdrom/storage_foundation_for_oracle_rac
2. # ./installsfrac –precheck (to check the system is ready to install or not)
3. # ./installsfrac

Archived

During the installation, be sure to use the ClusterID assigned to your specific cluster.

The following questions, which will be asked by the installer, must be answered exactly as described below:

Do you want to configure SMTP notification? no

Note: When asked whether to start the components now, answer “yes”.

☐ Verify the SFRAC installation.

1. # gabconfig –a

After a successful installation of SFRAC, the “gabconfig –a” output should look similar to:

```
GAB Port Memberships
===============================================================
Port a gen a21901 membership 01
Port h gen a21904 membership 01
Port o gen a21903 membership 01
```

Note: The “cvm” service group will be faulted after installation. This is expected and will not impact this lab. No action is necessary; just proceed to the next section.

☐ Configuring I/O Fencing on hdisk5, hdisk6 and hdisk7

This section will guide you through the configuration of I/O Fencing.

1.Shut down VCS and unload the vxfen driver.

Perform the following on each node of the cluster:

a. **# /opt/VRTSvcs/bin/hastop –local** (or **hastop –all** from one node)

2. Run vxfentsthdw

a. **#/opt/VRTSvcs/vxfen/bin/vxfentsthdw**

3. Create the coordinator disk group.

Perform the following from one node of the cluster.

a. # vxdg init vxfencoorddg hdisk5

b. # vxdg -g vxfencoorddg adddisk hdisk6

c. # vxdg –g vxfencoorddg adddisk hdisk7

d. # vxdg deport vxfencoorddg

e. # vxdg –t import vxfencoorddg

f. # vxdg deport vxfencoorddg

4. Create the /etc/vxfendg file.

Perform the following on each node of the cluster:

a. # echo “vxfencoorddg” > /etc/vxfendg

5. Load the vxfen driver and configure fencing:

Perform the following on each node of the cluster:

a. # /etc/rc.d/rc2.d/*vxfen* start

6. Verify the vxfen configuration. Execute **gabconfig –a** from one node of the cluster and verify that port ‘b’ has proper membership.

7. Add the “UseFence” cluster attribute to the main.cf file.

   Perform the following from the first node (mcbprod1) of the cluster:

   a. Use the vi editor to open the /etc/VRTSvcs/conf/config/main.cf file.

   b. Add the “UseFence=SCSI3” line to the list of cluster attributes in the main.cf file.

   After modifying the main.cf file, it should look similar to the example below.

```
…
cluster rac_test01 (
UserNames = { admin = dqrJqlQn }
Administrators = { admin }
HacliUserLevel = COMMANDROOT
CounterInterval = 5
UseFence = SCSI3
)
…
```

8. Verify the validity of your changes to the main.cf file.

   a. # /opt/VRTSvcs/bin/hacf –verify /etc/VRTSvcs/conf/config

   You successfully modified the main.cf file if no errors were reported by the hacf utility.

9. Start VCS.

   Execute the following on each node of the cluster (starting with the first node):

   a. # /opt/VRTSvcs/bin/hastart

   b. # gabconfig –a

   The “gabconfig –a” output should resemble the following:

```
GAB Port Memberships
===============================================================
Port a gen a21901 membership 01
Port b gen a21907 membership 01
Port f gen a21912 membership 01
Port h gen a2190a membership 01
Port o gen a21903 membership 01
Port v gen a2190e membership 01
Port w gen a21910 membership 01
```

☐ Creating the dba group and the Oracle user:

This section will guide you through creating the dba group and the Oracle user. The UID and GID must be identical on all nodes of the cluster.

1. Create the dba group.

   Perform the following on all nodes of the cluster:

   a. # mkgroup –A id=1002 adms=’root’ dba

2. Create the Oracle user.

   Perform the following on all nodes of the cluster:

   a. # mkuser id=’1001’ pgrp=’dba’ groups=dba home=’/home/oracle’ oracle

   b. # mkdir /oracle

   c. # chown oracle:dba /oracle

☐ Creating a Shared Disk Group and Volume.

This section will guide you through creating a shared disk group and a volume. At least one shared disk group should be used during this lab.

1. Create the shared disk group from the master node.

Determine the master node. Then create the disk group:

a. **`# vxdctl –c mode`** (the output reports MASTER or SLAVE)

Create the disk group:

b. **`# vxdg –st init orabindg hdisk8`** (for Oracle binary)

c. **`# vxdg –st init oradatadg hdisk9-hdisk14`** (for Oracle database)

Create a shared volume from the master node.

d. **`# vxassist –g orabindg make orabinvol 8G`**

e. **`# vxassist –g oradatadg make oradatavol 2G`**

All other volumes can be created using this command syntax.

Grant the Oracle user ownership of the newly created volume.

f. **`# vxedit –g orabindg set user=oracle group=dba mode=755 orabinvol`**

g. **`# vxedit –g oradatadg set user=oracle group=dba mode=755 oradatavol`**

Note: Be sure to create a separate shared raw volume for Oracle. Name the volume srvmvol, and make the volume 200M.

☐ Creating and Mounting the VxFS File System:

This section will guide you through creating and mounting the VxFS file system.

Perform the following from any node of the cluster:

Create:

a. **`/opt/VRTS/bin/mkfs –V vxfs /dev/vx/rdsk/orabindg/orabinvol`**

b. **`/opt/VRTS/bin/mkfs –V vxfs /dev/vx/rdsk/oradatadg/oradatavol`**

All other volumes can be created using this command syntax.

Perform the following on all nodes of the cluster:

a. **`mkdir /oracle /oradata`**

b. **`mount –V vxfs –o cluster /dev/vx/dsk/orabindg/orabinvol /oracle`**

c. **`chown oracle:dba /oracle (only need to do on one node)`**

d. **`mount –V vxfs –o cluster /dev/vx/dsk/oradatadg/oradatavol /oradata`**

e. **`chown oracle:dba /oradata (only need to do on one node)`**

Mount the rest of data filesystems and change ownership to oracle user and group.

☐ Performing Oracle9i pre-installation tasks:

Run the preinstallora script.

```
cd /opt/VRTSvcs/rac/bin/scripts
./preinstallora
```

Ensure the following filesets are installed on all nodes of the cluster:

Archived

```
bos.acct-5.2.0.13
bos.adt.libm-5.2.0.13
bos.perf.tools-5.2.0.14
bos.rte.install-5.2.0.15
bos.rte.tty-5.2.0.14
bos.sysmgt.loginlic-5.2.0.0
perfagent.tools-5.2.0.10
rsct.basic.rte -2.3.1.5
rsct.compat.basic.rte-2.3.1.2
rsct.compat.clients.rte-2.3.1.1
```

- Installing Oracle9i:

  Use the GUI script to install Oracle9i.

  Login as root on the install host and perform the following:

  a. **# export ORACLE_BASE=/oracle**
  b. **# export ORACLE_HOME=/oracle/product/9.2.0.4**
  c. **# export DISPLAY=<hostname>:0.0**
  d. **# cd /path_to_9201/Disk1**
  e. **# ./rootpre.sh**
  f. **# su – oracle**
  g. **$ ./runInstaller**

  Some notes about installing Oracle9i RAC:

  - Do NOT select the option to create a database at installation time. This will be done later.
  - **Cluster Node Selection** screen: Be sure to select all nodes in the cluster.
  - **Available Products** screen: Select *Oracle 9i Database 9.2.0.1.0*

- Applying Oracle9i Patch Set 4:

  This section will guide you through applying the Oracle9i patchset 4.

  From the install host, switch to the Oracle user and install the patchset:

  a. **# su – oracle**
  b. **$ export ORACLE_HOME=/oracle**
  c. **$ export ORACLE_BASE=/oracle/product/9.2.0.4**
  d. **# cd /path_to_9204/Disk1**
  e. **#./runInstaller**

  The first time the patchset installer is launched, select to upgrade the "*Oracle Universal Installer 2.2.0.18.0*". After installation is complete, the installer *must* be closed and relaunched (known Oracle bug). The second time the installer is launched, select to upgrade the "*Oracle9iR2 Patch Set 3 9.2.0.4.0*"(slibclean).

- Using the postinstall or a script to re-link the Oracle9i libraries:

  Create the srvConfig.loc file on all nodes of the cluster:

  a. **# mkdir /var/opt/oracle**
  b. **# echo "srvconfig_loc=/dev/vx/rdsk/orabindg/srvmvol" >**

Archived

```
/var/opt/oracle/srvConfig.loc
```

c. `# chown oracle:dba /var/opt/oracle/srvConfig.loc`

Perform the following from any node of the cluster:

a. `cd /opt/VRTSvcs/rac/bin/scripts`

b. `./postinstallora`

When asked if you want to remove the HACMP classes, select 'y'.

When asked if you want to enter the oracle home:

Enter the <oracle home directory>

When asked to enter the oracle user:

Enter **oracle**

When asked to enter the group:

Enter **dba**

- ☐ Creating the Oracle9i Database:

This section will guide you through the creation of an Oracle9i database.

Perform the following from the first node of the cluster:

a. `# su – oracle`

b. `$ export ORACLE_BASE=/oracle`

c. `$ export ORACLE_HOME=/oracle/product/9.2.0.4`

d. Run your database creation script or dbca

- ☐ Bringing up the Oracle instance on both nodes:

This section will guide you through bringing up the Oracle instance on both nodes.

Perform the following from the first node of the cluster:

a. `# su – oracle`

b. `$ export ORACLE_BASE=/oracle`

c. `$ export ORACLE_HOME=/oracle/product/9.2.0.4`

d. `$ export ORACLE_SID=testdb1`

e. `$ /oracle/product/9.2.0.4/bin/sqlplus '/ as sysdba'`

f. `SQL> shutdown`

g. `SQL> startup pfile='/oracle/product/9.2.0.4/dbs/inittestdb1.ora'`

On the second node of the cluster, perform steps a -g (skip f), but the ORACLE_SID should be set to "testdb2".

Installing Veritas Storage Foundation for DB2

For this installation we describe the prerequisites, environments, fixes, tasks, verifications and more.

Archived

Prerequisite

- Full functionality of VERITAS Storage Foundation for DB2 requires that the DB2 database be built on VERITAS Volume Manager and VERITAS File System (components of VERITAS Storage Foundation for DB2).

Supported platforms in AIX environments

- DB2 Releases
- 8.1 ESE
- 7.2 EE (FixPak 6 or higher)
- AIX Releases
- Versions 5.1 or 5.2
- 32-bit and 64-bit

Confirming the operating system version

- In AIX environments, VERITAS Storage Foundation for DB2 can only be installed on systems running AIX 5L.
- 32 and 64-bit environments are supported on all of these releases.
- Operating system information can be verified at login or by typing: **<uname –a>**

Acquiring operating system fixes

- The required AIX operating system level and fixes are listed in the Storage Foundation 4.0 for DB2 Release Notes and can be obtained from the IBM Web site:

  http://www.ibm.com/servers/eserver/support/pseries/aixfixes.html

- The fix installation process will be covered later.

Obtaining a license key

http://vlicense.veritas.com

Installation tasks

To install Storage Foundation for DB2:

1. Review Release Notes—required patches are listed.
2. Add the required operating system patches.
3. Run the installsfdb2 script.
4. Start VEA (VERITAS Enterprise Administrator) on the server.
5. Install VEA Client (discussed later).

Adding the required operating system fixes

- Check if fixes are already installed.
- If a specific fix file set is installed, the output will be:

```
# instfix -ik IY56497
```

All filesets for IY56497 were found.

- If the patch is not installed, the output will be:

```
# instfix -ik IY56493
```

There was no data for IY56493 in the fix database.

Archived

- To install a new fix:

```
# instfix –T –d <device_path>
```

About the installsfdb2 script

The installsfdb2 script performs the following tasks:

1. Checks for operating system level and fixes.
2. Installs the Storage Foundation for DB2 packages.
3. Prompts for the license key.

Note: The installation script will prompt the user for the DB2 Instance Owner and Group. Please make sure that the DB2 Instance Owner and Group have been created before running the installsfdb2 script.

Running the installsfdb2 script

```
# ./installsfdb2
```

Enter the system names separated by spaces on which to install SFDB2: tarpan
Add name of local machine.
For remote installation, .rhosts has to be configured.

```
Checking OS version on tarpan ............................. AIX 5.2.0
Checking VRTSdb2ed fileset ............................. not installed
Initial system check completed successfully.
Press [Return] to continue:
VERITAS Infrastructure fileset installation:
Installing VERITAS Infrastructure filesets on tarpan:
Checking VRTScpi.rte fileset ........................... not installed
Checking VRTSvlic fileset .............................. not installed
Checking file system space ............... required space is available
Installing VRTScpi 4.0.4001.0 on tarpan ......................... Done
Installing VRTSvlic 3.2.6.1 on tarpan ........................... Done
VERITAS Infrastructure filesets installed successfully.
Press [Return] to continue:
Licensing Verification:
Checking SFDB2 license key on tarpan ..................... not licensed
Enter a SFDB2 license key for tarpan: [?] 3JP6-U3PP-B6I2-S77G-LI98-PCP
Registering VERITAS Storage Foundation for DB2 DEMO key on tarpan
Do you want to enter another license key for tarpan? [y,n,q,?] (n)
SFDB2 licensing completed successfully.
Press [Return] to continue:
installsfdb2 can install the following optional SFDB2 filesets:
VRTSobgui VERITAS Enterprise Administrator
VRTSvmman VERITAS Volume Manager Manual Pages
VRTSvmdoc VERITAS Volume Manager Documentation
VRTSfsman VERITAS File System Manual Pages
VRTSfsdoc VERITAS File System Documentation
VRTSap VERITAS Action Provider
VRTStep VERITAS Task Provider
VRTSd2doc VERITAS Storage Foundation for DB2 Documentation
```

Note: Select one of the following three options:

1. Install all of the optional filesets.
2. Install none of the optional filesets.
3. View fileset descriptions and select optional filesets.

Select the optional filesets to be installed on all systems? [1-3,q,?] (1)

Use the default – install all filesets

```
installsfdb2 will install the following SFDB2 filesets:
VRTSperl.rte VERITAS Perl 5.8.0 Redistribution
VRTSob VERITAS Enterprise Administrator Service
VRTSobgui VERITAS Enterprise Administrator
VRTSveki VERITAS Kernel Interface
VRTSvxvm VERITAS Volume Manager Binaries
VRTSvmman VERITAS Volume Manager Manual Pages
...
VRTSvxmsa VERITAS Mapping Service, Application Libraries
VRTSd2gui VERITAS Storage Foundation Graphical User Interface for DB2
VRTSd2doc VERITAS Storage Foundation for DB2 Documentation
VRTSvail VERITAS Array Provider
Press [Return] to continue:
Checking system installation requirements:
Checking SFDB2 installation requirements on tarpan:
Checking VRTSperl.rte fileset ..... not installed
Checking VRTSob fileset ..... not installed
Checking VRTSobgui fileset ..... not installed
Checking VRTSveki fileset ..... not installed
..
.
Checking file system space ..... required space is available
Checking vxsvc process ..... not running
Checking vxfs driver ..... vxfs check command not installed
Checking veki driver ..... not running
Installation requirement checks completed successfully.
Press [Return] to continue:
Installing Storage Foundation for DB2 4.0.0.0 on tarpan:
Installing VRTSperl 4.0.2.0 on tarpan ..... Done 1 of 24 steps
Installing VRTSob 3.2.514.0 on tarpan ..... Done 2 of 24 steps
Installing VRTSobgui 3.2.514.0 on tarpan ..... Done 3 of 24 steps
Installing VRTSveki 1.0.0.0 on tarpan ..... Done 4 of 24 steps
...
Installing VRTSd2gui 4.0.0.0 on tarpan ..... Done 20 of 24 steps
Installing VRTSd2doc 4.0.0.0 on tarpan ..... Done 21 of 24 steps
Installing VRTSvail 4.2.606.5 on tarpan ..... Done 22 of 24 steps
Adding patch VRTSob 3.2.514.16 on tarpan ..... Done 23 of 24 steps
Adding patch VRTSobgui 3.2.514.16 on tarpan ..... Done 24 of 24 steps
Storage Foundation for DB2 installation completed successfully.
Press [Return] to continue:
```

Output of the installsfdb2 script

The following information is required to set up the permission for VxDBA Utility.

The VxDBA utility for database administration requires permission changes to allow database administrators (DBAs) access to the tools. The default setting at installation time for the /opt/VRTSdb2ed directory allows only the root user access to the directory.

```
Would you like to change permissions to allow a DBA or a group of DBAs access
to the VxDBA tools?
Do you want to add single user access [y,n,q,?] (y)
Enter login account name for DBA user: inst1
Do you want to add group access [y,n,q,?] (y)
```

```
Enter group name for DBA users: db2grp
Specify the user name and group name for the DBA
Specify the user name and group name for the DBA
```

It is optional to configure SFDB2 now.
If you choose to configure SFDB2 later, you can either do so manually, or run the installsfdb2 –configure command.

```
Are you ready to configure SFDB2? [y,n,q,?] (y)
```

installsfdb2 will now ask sets of SFDB2 configuration-related questions. When a [b] is presented after a question, 'b' may be entered to go back to the first question of the configuration set. When a [?] is presented after a question, '?' may be entered for help or additional information about the question. Following each set of questions, the information you have entered will be presented for confirmation. To repeat the set of questions and correct any previous errors, enter 'n' at the confirmation prompt. No configuration changes are made to the systems until all configuration questions are completed and confirmed. Press Return to continue:

```
Do you want to start Storage Foundation for DB2 processes now? [y,n,q] (y)
Evaluating which systems can now be started...
System tarpan is eligible -- can be started.
Preparing to start VxVM on target systems...
```

Note: The vxconfigd daemon will be started, which can take a while, depending upon the hardware configuration.

```
Begin initial start of VxVM on system tarpan
Starting vxconfigd for VxVM ..............................
Succeeded
Done with initial start of VxVM on system tarpan
Done with starting VxVM on target systems...
Press [Return] to continue:
```

Many Volume Manager commands affect the contents or configuration of a disk group. Such commands require that the user specify a disk group. This is accomplished by using the -g option of a command or setting the VXVM_DEFAULTDG environment variable. An alternative to these two methods is to configure the default disk group of a system.

```
Evaluating which systems can now have their default disk group configured...
System tarpan is eligible -- can configure the default diskgroup.
Do you want to set up the default disk group for each system? [y,n,q,?] (y)
Specify a default disk group for system tarpan or type 'l' to display a
listing of existing disk group(s). [?] db2dg
You have decided to specify the default disk group as follows:
Host: tarpan ..................................... Disk group: db2dg
```

Note: If 'nodg' is displayed, then the host will be configured to have no default disk group.

```
Is this correct? [y,n,q] (y)
Setting default diskgroup to db2dg on tarpan ..................... Done
Preparing to start daemons on target system(s)...
Starting vxrelocd on tarpan ................................... Success
Starting vxcached on tarpan ................................... Success
Starting vxconfigbackupd on tarpan ............................ Success
Storage Foundation for DB2 was started successfully.
Press [Return] to continue:
Optional default diskgroupname
Installation of Storage Foundation for DB2 4.0.0.0 has completed successfully.
```

The installation summary is saved at:
/opt/VRTS/install/logs/installsfdb2902111558.summary

The installsfdb2 log is saved at:
/opt/VRTS/install/logs/installsfdb2902111558.log

The installation response file is saved at:
/opt/VRTS/install/logs/installsfdb2902111558.response

When installsfdb2 installs software, some software may be applied rather than committed. It is the responsibility of the system administrator to commit the software, which can be performed later with the -c option of the installp command.

Post-installation tasks

After installation:

- Verify the installation.
- Check and set permissions.
- Add and register a database.
- Add users to the VEA Server Registry.
- Start the VERITAS Enterprise Administrator (VEA) Server.
  VEA is the Storage Foundation for DB2 user interface.
- Install the VEA client.
- Launch the VEA client.

Verifying the installation

To verify that package modules have been installed properly, use this command:

```
# genkex | grep vx
2334088 970 /etc/vx/kernel/dmpjbod
2333680 a00 /etc/vx/kernel/dmpap
2332cd8 9a0 /etc/vx/kernel/dmpaa
23e0460 bfb0 /usr/lib/drivers/vxqio.ext_52
221f668 1470 /usr/lib/drivers/vxportal.ext_52
69bb000 1bdff0 /usr/lib/drivers/vxfs.ext_52
2331810 14c0 /usr/lib/drivers/vxspec
66ec000 2a30d4 /usr/lib/drivers/vxio
23a7f00 38544 /usr/lib/drivers/vxdmp
620b000 2a30d4 /usr/lib/drivers/vxio
22f92c0 38544 /usr/lib/drivers/vxdmp
2220ae0 bfb0 /usr/lib/drivers/vxqio.ext_52
5ef4000 1bdf70 /usr/lib/drivers/vxfs.ext_52
```

Verifying the permissions

Use these commands:

```
# ls -ld /opt/VRTSdb2ed
drwxr-x--- 6 inst1 db2grp 512 Sep 02 11:24 /opt/VRTSdb2ed
```

- This should be the instance owner and group ID as defined during the installfsdb2 script execution.
- Use chmod / chgrp / chown commands to correct the VRTSdb2ed permissions.

Initialize VM storage

Use these commands:

```
# vxdisk list
DEVICE TYPE DISK GROUP STATUS
```

```
Disk_0 auto:none - - online invalid
Disk_1 auto:none - - online invalid
Disk_2 auto:none - - online invalid
Disk_3 auto:none - - online invalid
Disk_4 auto:none - - online invalid
Disk_5 auto:LVM - - LVM
# vxdisk init Disk_0
# vxdisk init Disk_1
# vxdg init db2dg Disk_0 Disk_1
# vxdisk list
DEVICE TYPE DISK GROUP STATUS
Disk_0 auto:cdsdisk Disk_0 db2dg online
Disk_1 auto:cdsdisk Disk_1 db2dg online
Disk_2 auto:none - - online invalid
Disk_3 auto:none - - online invalid
Disk_4 auto:none - - online invalid
Disk_5 auto:LVM - - LVM
```

Create test database

Use these commands:

```
# vxassist [-g db2dg] make gugfs 10G
# mkfs -V vxfs /dev/vx/rdsk/db2dg/gugfs
version 6 layout
20971520 sectors, 10485760 blocks of size 1024,
log size 32768 blocks
largefiles supported
# mkdir /gugfs
# mount -t vxfs /dev/vx/dsk/db2dg/gugfs /gugfs
# chown inst1:db2grp /gugfs
# su - inst1
[YOU HAVE NEW MAIL]
$ db2 create database gugdb on /gugfs
DB20000I The CREATE DATABASE command completed successfully.
```

Update environment and register DB

Use these commands:

```
$ echo $PATH | grep VRTSdb2ed
$ PATH=$PATH:/opt/VRTSdb2ed/bin
$ echo 'PATH=$PATH:/opt/VRTSdb2ed/bin' >> ~/.profile
$ echo $DB2DATABASE
$ DB2DATABASE=gugdb export DB2DATABASE
$ echo 'DB2DATABASE=gugdb export DB2DATABASE' >> ~/.profile
$ db2ed_update -D $DB2DATABASE
```

Add users and start VEA server

Use these commands:

```
# /opt/VRTSd2gui/bin/vxdb2edusr -a user -n inst1
Successfully added 'inst1' into registry
# /opt/VRTSob/bin/vxsvcctrl status
Current state of server: NOT RUNNING
# /opt/VRTSob/bin/vxsvcctrl start
Starting VERITAS Enterprise Administrator Service (vxsvc)
..DBED: Initializing Storage Foundation Provider 4.0 for DB2
```

Archived

```
.........VEA Server has successfully started
```

Installing the VEA client

The VEA client installer for Windows® is located in the ./windows subdirectory of the complete Storage Foundation for DB2 distribution tar archive.

Launching the VEA client

For Windows desktops launch the VEA client from the Programs menu: (**Start**->**Programs**->**VERITAS**->**VERITAS Enterprise Administrator**)

To make the connection:

1. Enter the Hostname.
2. Enter the operating system Username.
3. Enter the operating system Password.
4. Click OK.

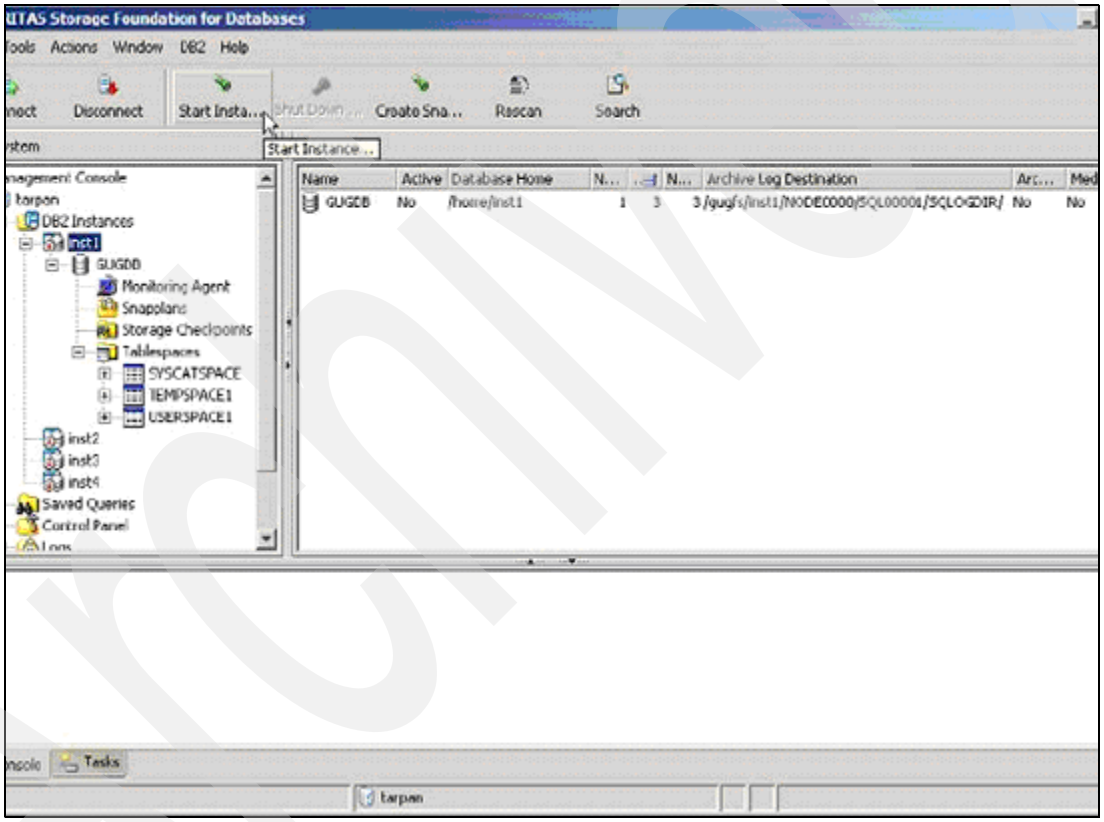

Figure 5 VEA main window

Archived

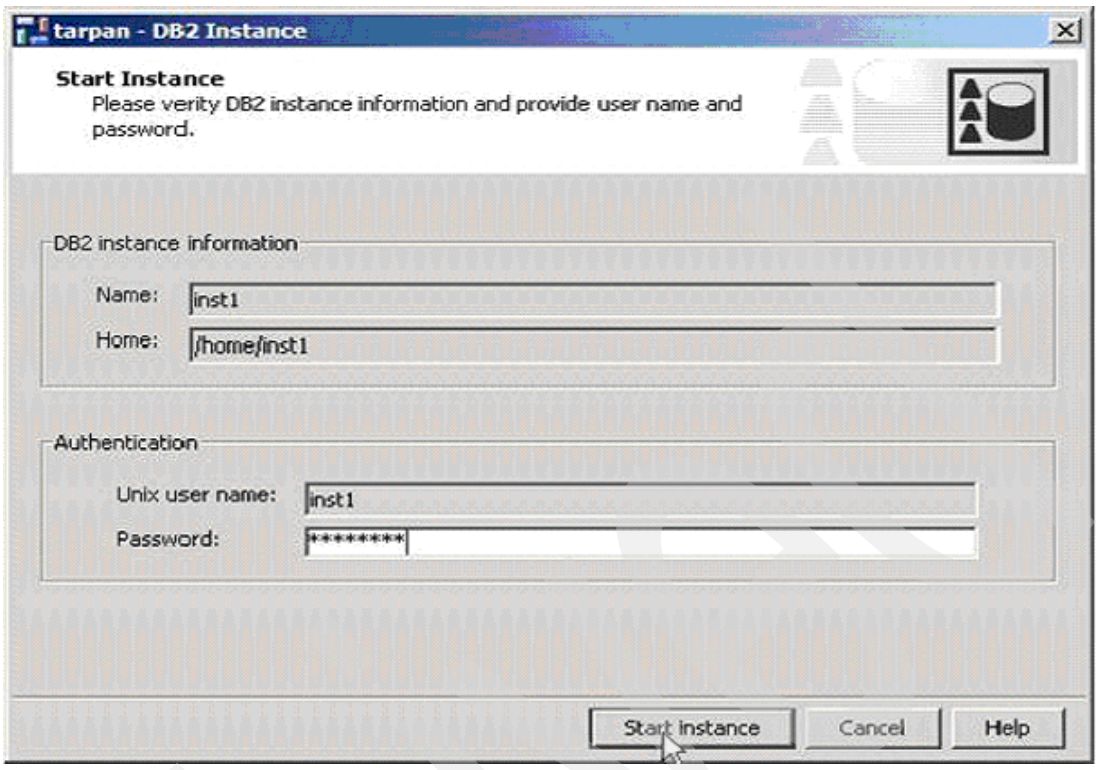

Figure 6 Starting the DB2 Instance Using VEA

The DB2 User Name and DB2 Password correspond to the operating system user name and password authorized to access the DB2 files.

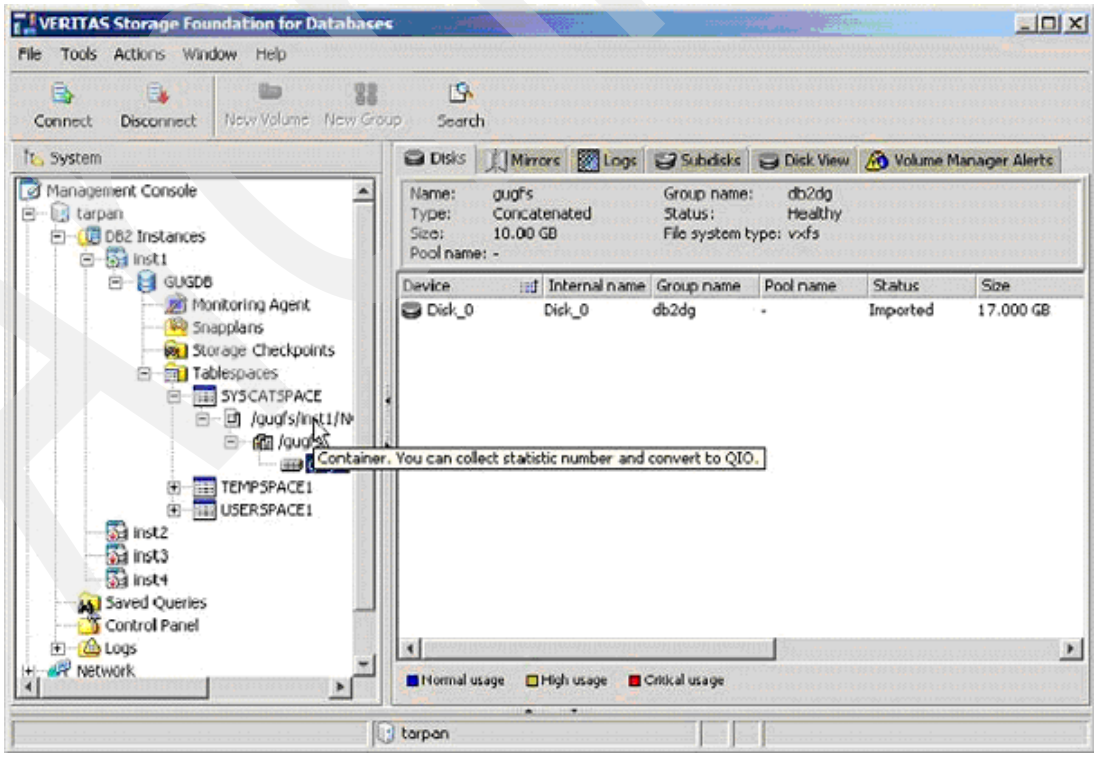

Figure 7 Container View

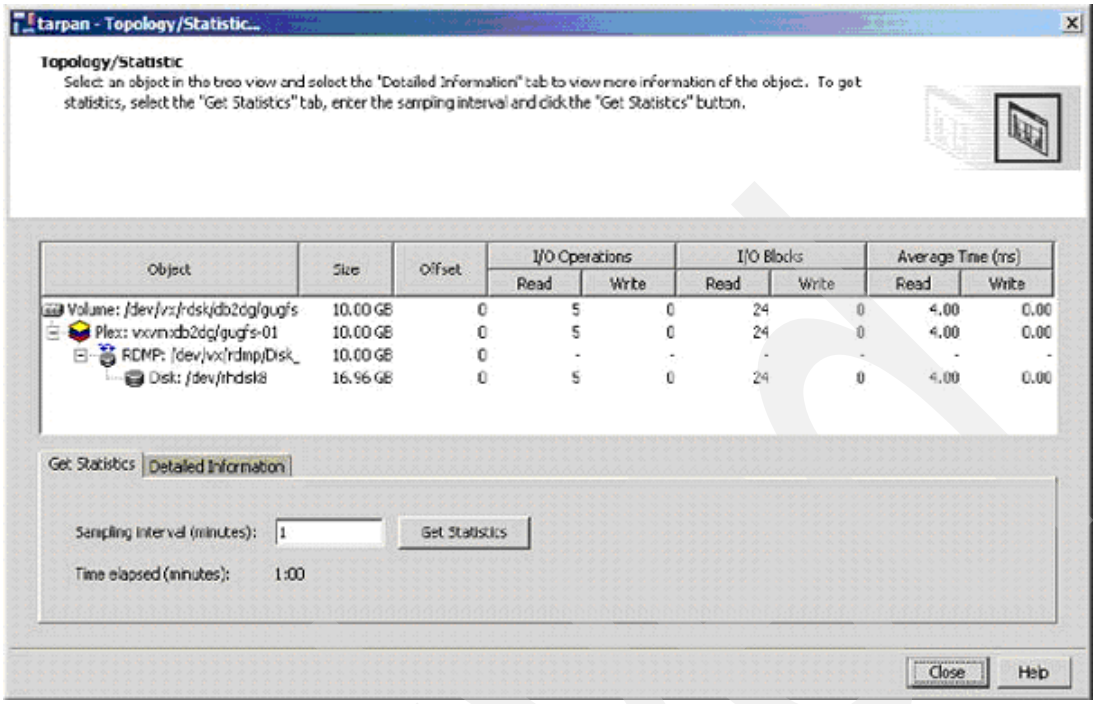

Figure 8 Container Details and Statistics

Test scenarios

Here is additional documentation about our Oracle and DB2 tests.

Oracle RAC

We tested different configurations of failover. A whitepaper on this subject is available at:

http://eval.veritas.com/mktginfo/products/White_Papers/High_Availability/IBM_Paper_on_SFRAC.pdf

Table 1 is a summary of that whitepaper.

Table 1 Test cases

Event	Node A: What Happens?	Node B: What Happens?
Both private networks fail	Node A races for majority of coordinator disks. If Node A wins race for coordinator disks, Node A ejects Node B from the shared disks and continues	Node B races for majority of coordinator disks. If Node B loses race for coordinator disks, Node A ejects Node B removes itself from the cluster.
Both private networks function again after event above	Node A continues to work.	Node B has crashed. It cannot start the database since it is unable to write to the data disks
One private network fails	Node A prints message about an IOFENCE on the console but continues	Node B prints message about an IOFENCE on the console but continues.

Archive

Node A hangs	When Node A is no longer hung, any queued writes to the data disks fail because Node A receives message from GAB about being ejected, it removes itself from the cluster	Node B loses heatbeats with Node A, and races for a majority of coordinator disks. Node B wins race for coordinator disks and ejects Node A from shared data disks.
Nodes A and B and private networks lose power. Coordinator and data disks retain power. Power returns to nodes and they restart, but private networks still have no power.	Nodes A restarts and I/O fencing driver (vxfen) detects Node B is registered with coordinator disks. The driver does not see Node B listed as member of cluster because private networks are down. This causes the I/O fencing device driver to prevent Node A from joining the cluster.	Nodes B restarts and I/O fencing driver (vxfen) detects Node A is registered with coordinator disks. The driver does not see Node A listed as member of cluster because private networks are down. This causes the I/O fencing device driver to prevent Node B from joining the cluster.
Node A crashes while Node B comes up and Node A is still down	Node A is crashed	Node B restarts and detects Node A is registered with the coordinator disks. The driver does not see Node A listed as member of the cluster, so Node B drops out of the cluster and displays a warning message on the console.
The disk array containing two of the three coordinator disks is powered off.	Node A continues to operate as long as no nodes leave the cluster.	Node B continues to operate as long as no nodes leave the cluster.
Node B leaves the cluster and the disk array is still powered off	Node A races for a majority of coordinator disks. Node A fails because only one of the three coordinators disks is available. Node A removes itself from the cluster.	Node B leaves the cluster.

DB2 HA test scenarios overview

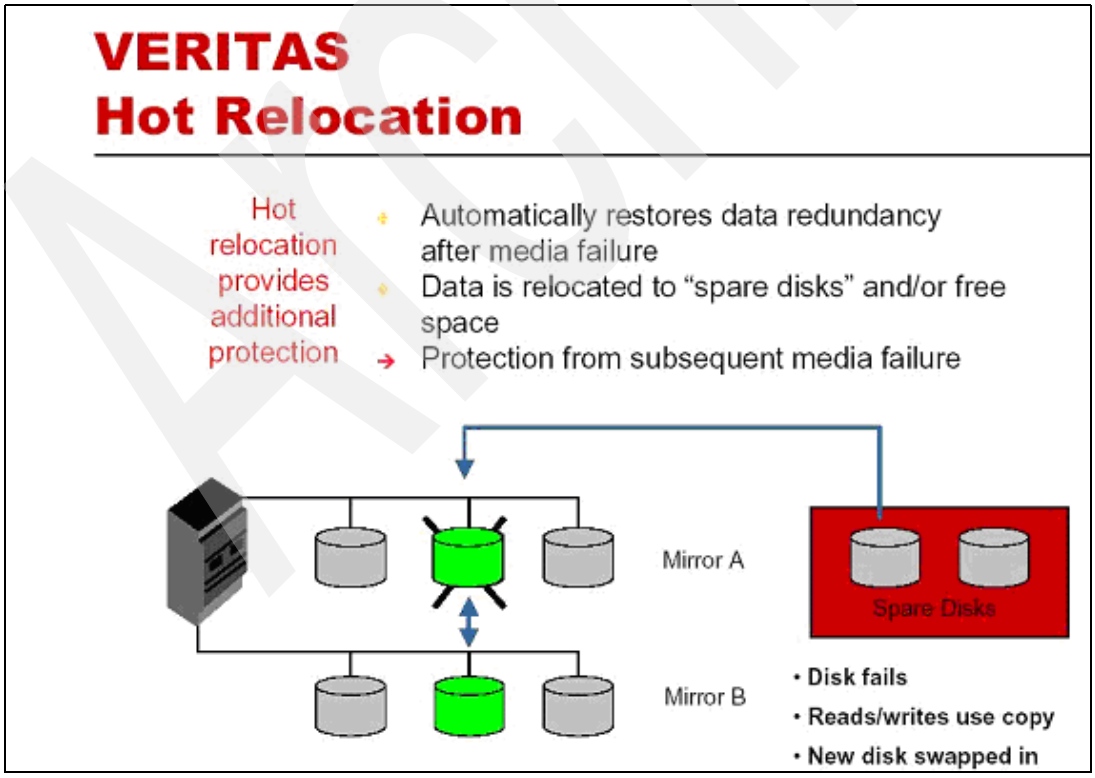

Figure 9 Veritas hot relocation

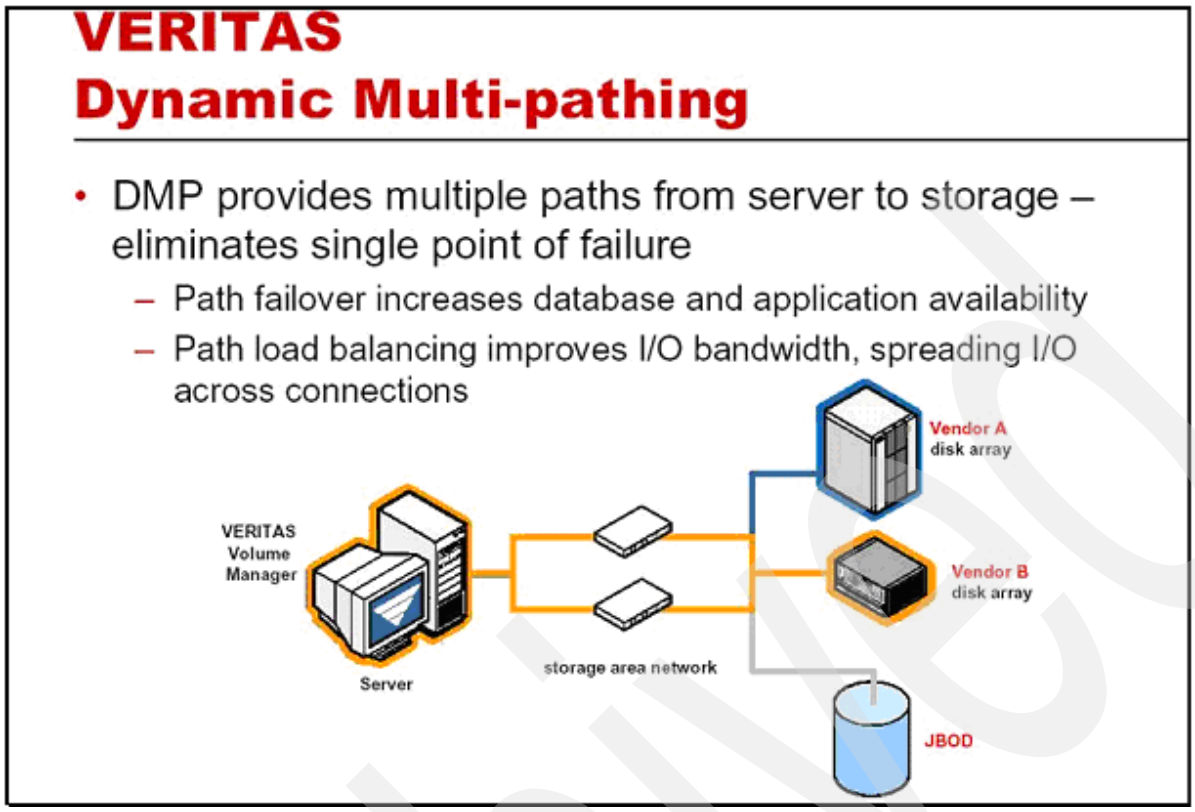

Figure 10 Veritas dynamic multi-pathing

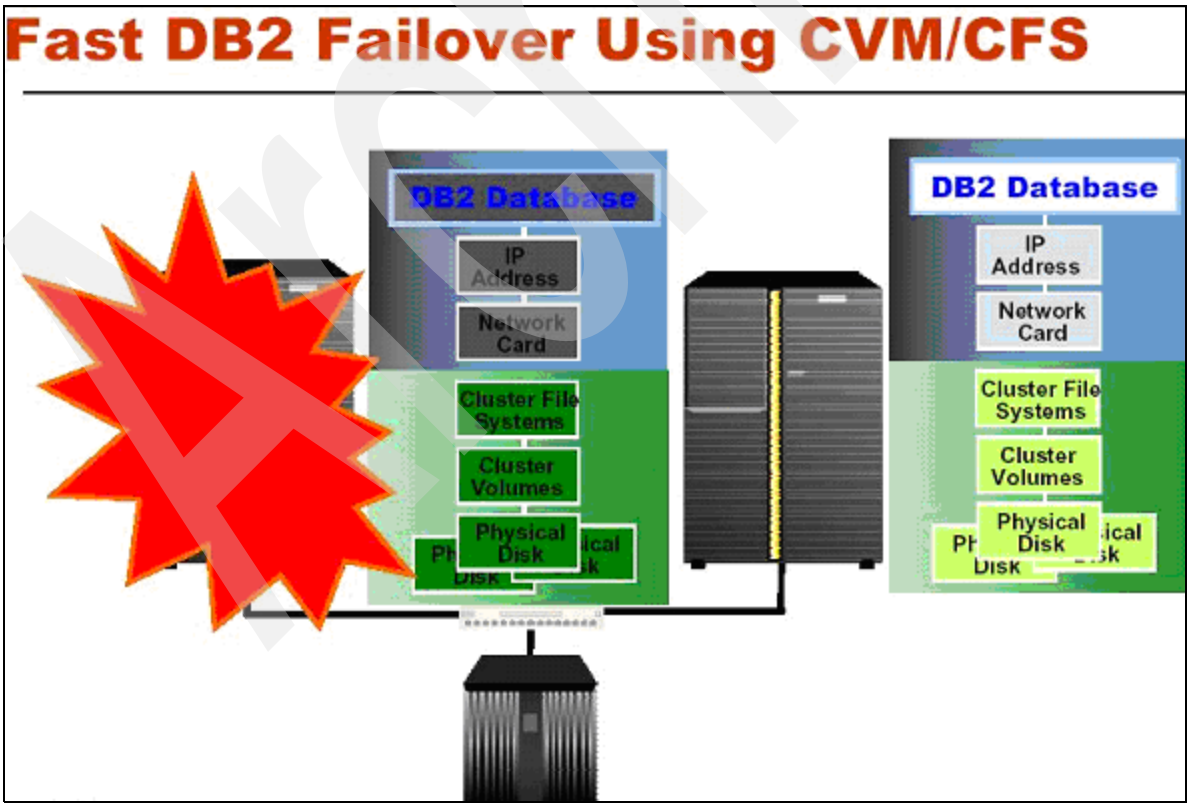

Figure 11 Fast DB2 failover using CVM/CFS

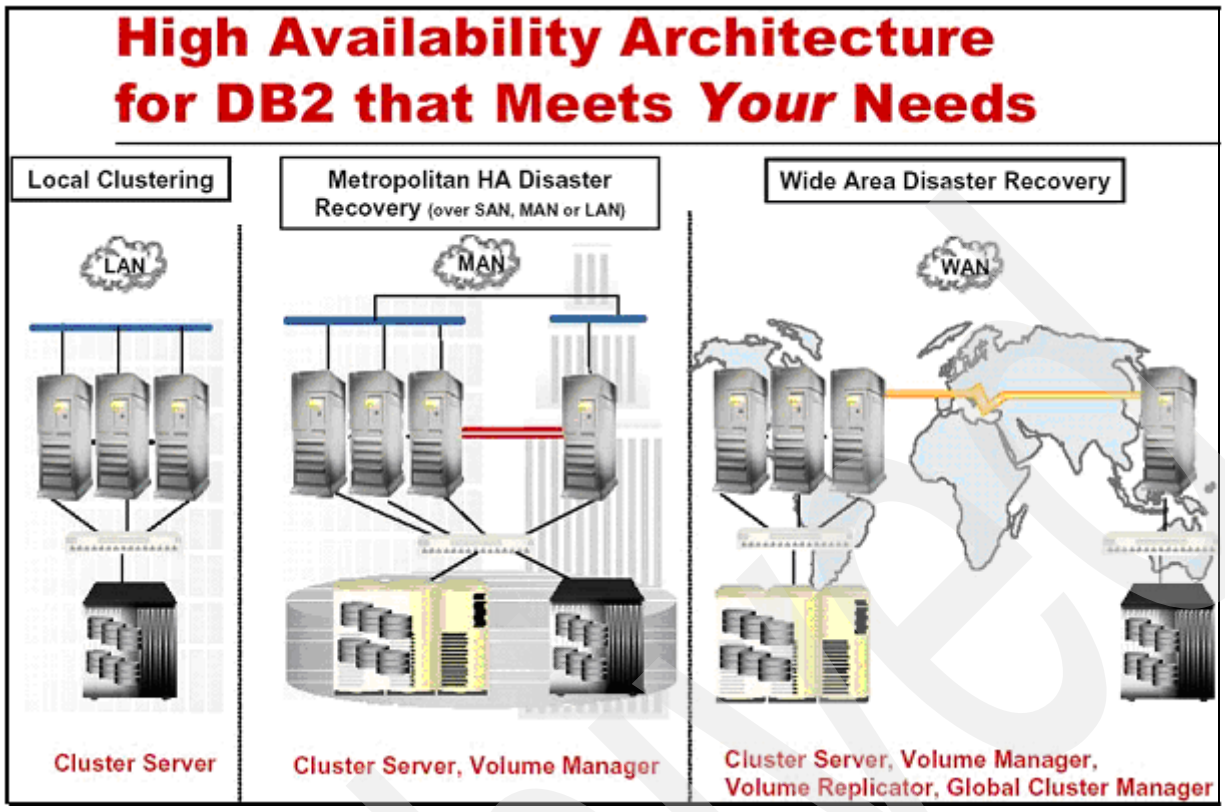

Figure 12 High availability architecture for DB2

Conclusion

This document presents an overview of the Veritas Storage Foundation solution for Oracle 9i RAC and DB2, and their installations and tests. The ease of installation and the efficiency of the high availability mechanism demonstrates the suitability of this solution for companies who want a robust, available, and manageable infrastructure.

Resources

For more information, visit:

http://www.symantec.com/Products/enterprise?c=prodinfo&refId=145
http://www.symantec.com/Products/enterprise?c=prodinfo&refId=208

Archive

Archived

Notices

This information was developed for products and services offered in the U.S.A.

IBM may not offer the products, services, or features discussed in this document in other countries. Consult your local IBM representative for information on the products and services currently available in your area. Any reference to an IBM product, program, or service is not intended to state or imply that only that IBM product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any IBM intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any non-IBM product, program, or service.

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites is at your own risk.

IBM may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:
This information contains sample application programs in source language, which illustrates programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs. You may copy, modify, and distribute these sample programs in any form without payment to IBM for the purposes of developing, using, marketing, or distributing application programs conforming to IBM's application programming interfaces.

© Copyright International Business Machines Corporation 2006. All rights reserved.
Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Send us your comments in one of the following ways:

- Use the online **Contact us** review redbook form found at:
  ibm.com/redbooks
- Send your comments in an email to:
  redbook@us.ibm.com
- Mail your comments to:
  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400 U.S.A.

Trademarks

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

@server®	AIX®	IBM®
@server®	DB2 Universal Database™	Redbooks™
pSeries®	DB2®	Redbooks (logo) ™
AIX 5L™	HACMP™	

The following terms are trademarks of other companies:

Java and all Java-based trademarks are trademarks of Sun Microsystems, Inc. in the United States, other countries, or both.

Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Intel, Intel logo, Intel Inside, Intel Inside logo, Intel Centrino, Intel Centrino logo, Celeron, Intel Xeon, Intel SpeedStep, Itanium, and Pentium are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

UNIX is a registered trademark of The Open Group in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.

Archived