Compaq

Solution Guide

May 2002

Part Number:
16MV-0502A-WWEN
Edition 1.0

Prepared by:
Global SAP Solutions and
Enterprise Storage Group
Compaq Computer Corporation

Contents

StorageWorks DataSafe for mySAP.com (Tru64)

Abstract: This document gives an overview of a Compaq solution for adding disaster tolerant capabilities to a Tru64 Cluster environment using SANworks Data Replication Manager for mySAP.com.

SAP-ICCEngineering@compaq.com

Notice

The information in this publication is subject to change without notice and is provided “AS IS” WITHOUT WARRANTY OF ANY KIND. THE ENTIRE RISK ARISING OUT OF THE USE OF THIS INFORMATION REMAINS WITH RECIPIENT. IN NO EVENT SHALL COMPAQ BE LIABLE FOR ANY DIRECT, CONSEQUENTIAL, INCIDENTAL, SPECIAL, PUNITIVE, OR OTHER DAMAGES WHATSOEVER (INCLUDING, WITHOUT LIMITATION, DAMAGES FOR LOSS OF BUSINESS PROFITS, BUSINESS INTERRUPTION, OR LOSS OF BUSINESS INFORMATION), EVEN IF COMPAQ HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

The limited warranties for Compaq products are exclusively set forth in the documentation accompanying such products. Nothing herein should be construed as constituting a further or additional warranty.

This publication does not constitute an endorsement of the product or products that were tested. The configuration or configurations tested or described may or may not be the only available solution. This test is not a determination of product quality or correctness, nor does it ensure compliance with any federal, state or local requirements.

Compaq, NonStop, Deskpro, Compaq Insight Manager, Systempro, Systempro/LT, ProLiant, ROMPaq, QVision, SmartStart, NetFlex, QuickFind, PaqFax, and Prosignia are registered with the United States Patent and Trademark Office.

ActiveAnswers, Netelligent, Systempro/XL, SoftPaq, Fastart, QuickBlank, QuickLock are trademarks and/or service marks of Compaq Computer Corporation.

Microsoft, Windows and Windows NT are trademarks and/or registered trademarks of Microsoft Corporation.

The following are trademarks or registered trademarks of SAP AG; ABAP/4, InterSAP, RIVA, R/2, R/3, R/3 Retail, SAP (Word), SAPaccess, SAPfile, SAPfind, SAPmail, SAPoffice, SAPscript, SAPtime, SAPtronic, SAP-EDI, SAP EarlyWatch, SAP ArchiveLink, SAP Business Workflow, and ALE/WEB. The SAP logo and all other SAP products, services, logos, or brand names included herein are also trademarks or registered trademarks of SAP AG.

Intel, Pentium and Xeon are trademarks and/or registered trademarks of Intel Corporation.

Oracle is a registered trademark of Oracle Corporation.

Other product names mentioned herein may be trademarks and/or registered trademarks of their respective companies.

Solution Guide prepared by Global SAP Solutions and Enterprise Storage Group

Edition 2.03 (May 2002)
Part Number: 16MV-0502A-WWEN

Introduction

Overview

When data security and availability are critical to the success of their businesses, mySAP.com customers require a computing solution that protects their information systems from disasters, such as power outages, earthquakes, fires, floods, or acts of vandalism. The effects of a disaster range from temporary loss of availability to outright physical destruction of a facility and its assets. In the event of such a disaster, the mySAP.com system must allow customers to shift their information processing activities to another site as quickly as possible. Procedures for disaster recovery must therefore be predictable, well-defined, and immune to human error.

Site replication is a method of achieving disaster tolerance in a mySAP.com environment. Disaster tolerance is characterized by a short recovery time and avoidance of data loss. In a disaster-tolerant system based on this approach, redundant, active servers and client interconnects are located at geographically separated sites. As mySAP.com applications produce data, this data is copied by a replication system whose function is to maintain consistent replicas of the data at each site. Should the system at one site suffer a disaster, mySAP.com instances that were running at the now disabled site can be failed over to a surviving site that has the resources to support them. The process of failing over a mySAP.com application to the target node involves making the application's replicated data accessible, and starting instances on the target node to restore application availability.

Compaq Links

For further information on mySAP.com, please refer to the following Compaq publications:

- Compaq AlphaServer and Tru64 UNIX Performance Guide for mySAP.com

Enterprise Storage Group:

- http://www.compaq.com/storage/san_index.html.
- http://www.compaq.com/products/sanworks/drm/index.html
- http://www.compaq.com/storage/whitepapers.html#soft

Tru64 Unix Group:

- http://www.tru64unix.compaq.com
- http://css-ww.inet.cpqcorp.net/dt_campus.htm

SAP Links

Please refer to the following SAP publications for further technical information on mySAP.com:

- SAP OSS Note 516316 ***StorageWorks DataSafe for mySAP.com (Tru64)***

Solution Description

This document describes a method for configuring a disaster-tolerant system distributed over distant computer sites by combining the Compaq SANworks Data Replication Manager (DRM) Solution Kit with TruCluster Server technology. The combination of these two products has been qualified by Compaq as a Tru64 UNIX Campus-Wide DT Cluster configuration. The actual configurations are published in the *Tru64 UNIX V5 Campus-Wide Disaster Tolerant Cluster Fact Sheet.* In a stretched cluster using DRM, some member systems reside at one site and the others reside at a different site. A mySAP.com application can run the database instance on the initiator site and the corresponding central instance on the target site. All I/O occurs on the storage subsystem on the initiator site under non-disaster conditions. The DRM has exclusive access to storage at the target site, to which it replicates the I/O performed on the initiator site's storage. If a significant failure occurs at the initiator site, data processing can be resumed at the target site where the data is intact.

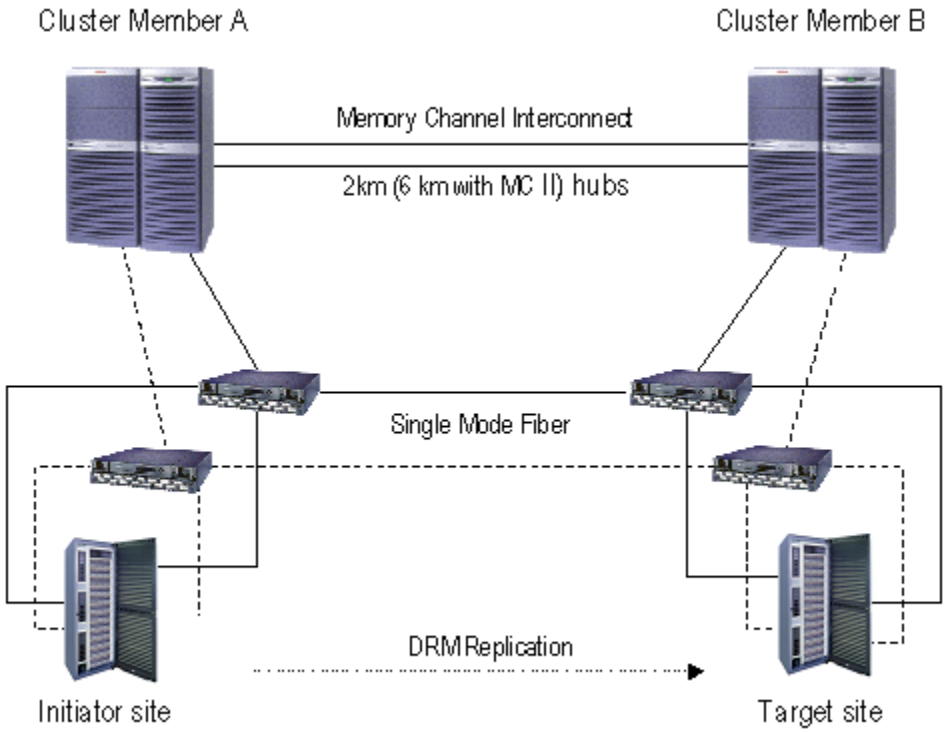

Figure 1 Campus-wide DT Cluster

The StorageWorks DataSafe solution takes advantage of the best features of both the DRM Solution Kit and TruCluster Server technology. Cluster members can span distances across a commercial or college campus or across a small metropolitan area. Data replication hardware ensures correct and consistent mirroring across sites, while TruCluster Server's single-system image (SSI) application and storage management features allow you to manage all cluster members, regardless of whether they are at the local or remote site. These capabilities save time during normal system administration and recovery procedures, and they also eliminate many opportunities for human error in communicating and dispatching operations at different sites. Although storage failover across sites is a manual process, cluster alias and Cluster Application

Availability (CAA) services automatically restart mySAP.com applications at the target site when the systems are rebooted after failover is complete.

Background

As customer applications and 24x7 access to data become business-critical, requirements for high-availability solutions with no single point of failure increase. Customers' ability to continue application processing and maintain data access in the event of a catastrophic disaster becomes critical to their business operations. Disaster-tolerant solutions provide high levels of availability with rapid data access recovery, no single point of failure, and continued data processing after the loss of one or more components of a configuration.

Please refer to the Glossary for terms and expressions used in the following sections.

Data Replication Manager (DRM)

The HSG80 ACS Version 8.6-4P Data Replication Manager Design Guide - Application Notes, describes DRM as a controller-based data replication software solution for disaster tolerance and data movement. DRM currently works with HSG80-based storage systems and allows all data to be mirrored between storage elements in two different storage arrays that can be in separate geographical locations. Each I/O write access is sent to both storage locations, and reads occur only at the local storage location. DRM copies data online and in real time to remote locations via a local or extended storage area network (SAN). For more information about DRM functionality, refer to the *Compaq* Features and Benefits of HSG80 ACS 8.6P Data Replication Manager – White Paper.

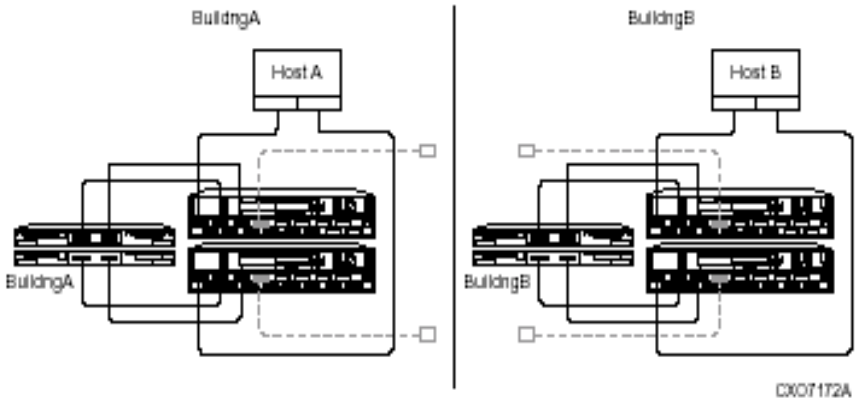

Figure 2 Basic DRM configuration

Tru64 and TruCluster Server

The *High Availability in the Compaq Tru64 UNIX Environment* white paper describes the terminology as well as features and benefits of Tru64 and TruCluster servers delivering high-availability features on AlphaServers today. Some of the major features are Single System Image (SSI) management with a cluster-wide file system (CFS), Cluster Application Availability (CAA) and simplified system management.

DRM and the Oracle Storage Compatibility Program

As part of Oracle's Storage Compatibility Program (OSCP), Oracle has created a test suite that tests remote mirroring technologies to ensure their compatibility with Oracle databases. The self test suite is provided for qualified vendors. Compaq chose to implement these tests using SANworks Data Replication Manager. As a member of OSCP, Compaq has successfully completed all test requirements stated in Oracle's remote mirroring test suite. The results were submitted to Oracle for verification and approved for entry in the program.

Replicating Oracle Databases

The Compaq white paper Oracle Databases Replication and Solutions highlights various concepts concerning Oracle database replication scenarios and gives a comparison table for different replication scenarios.

Configure a Stretched Cluster with DRM

Requirements

The current *Campus-wide Disaster Tolerant Cluster Application Note* specifies the configuration details for the combination of TruCluster and DRM as follows:

System

- Supported AlphaServers are: ES and GS series
- Compaq *Tru64* UNIX Version 5.1A + PK1 (latest patch kit)
- Redundant network, memory channel and Fibre Channel interconnects are required between the two sites

Cluster

- One or more cluster members at each of the two sites (maximum of eight nodes between the two sites)
- Compaq TruCluster™ Server, version 5.1A for Compaq Tru64 UNIX, version 5.1A
- Synchronous replication of cluster and data file systems from the initiator site to the target site (no quorum disk synchronous replication)

Storage

- HSG80 Array Controller minimum ACS Version 8.6P-4
- One or more DRM storage subsystems at each of two sites (maximum of four storage subsystems on each site)
- Each HSG80 configured with 512MB of memory
- Maximum of 12 remote copy sets per HSG80 is supported

Distance

- A maximum of 6 km between the two sites with memory channel hubs in the middle
- 2 km in a memory channel virtual hub configuration

Restrictions

- One quorum disk can be physically configured on each site. During normal operation only the initiator's site quorum disk is active and visible to the cluster. During a site failover, the target site's quorum will be enabled through Selective Storage Presentation (SSP) on the HSG80.
- Logical Storage Manager (LSM) is not supported at this time
- When memory channel is set up in standard hub mode, hubs must be located in separate buildings in order to prevent the loss of the two hubs at the same time. Loss of both hubs will induce loss of communications between the hosts and then cluster loss. If no hub is available, no cluster member boot is possible.

Replication Scenarios for Oracle with DRM

In a DRM environment, there are two major configuration options for replicating the Oracle database synchronously to the target site with no data loss.

The COMPAQ white paper Using Data Replication Manager with Oracle8i Under Tru64 and the Oracle Storage Compatibility Testing - Remote Mirroring White Paper give several considerations for mirroring the entire Oracle database or only the redo log information. As mySAP.com applications are based entirely on the underlying database, these suggestions are also valid in an SAP environment.

Replicating the Entire Oracle Database

- In this configuration, all volumes that contain either Oracle data files, online redo log files or control files are configured equally at both sites and linked to each other via a remote copy set on the HSG80 CLI level. All remote copy sets must be synchronous and in the same association set to be treated as a single entity. All members of an association set will be served by one HSG80 controller at one point in time. This ensures that the target site receives only consistent information.
- Depending on the customer's backup strategy, the Oracle archived redo log files do not need to be in the association set or replicated in this scenario at all, because archived redo log information is not necessary in case of a disaster failover, when all database files are replicated.
- There is a maximum of 12 remote copy sets per HSG80 pair. Bearing in mind that a minimum of one remote copy set is used per cluster member boot disk and one for the cluster root, there will be 9 remote copy sets left for the Oracle SAP database in a two-node cluster configuration. This can be a serious restriction.
- An association set cannot span multiple HSG80 pairs. This means that replicating the whole SAP Oracle Database is currently limited to one HSG80 pair. Compaq and Oracle are in the process of testing Oracle configurations that overcome this limit.

- A configuration with two HSG80 pairs is possible. The first HSG80 pair contains all Oracle data files, the online redo logs and the control files, providing capacity of 10 x 6 member Raid 0+1 sets and forming the database association set. The other HSG80 pair contains Oracle and SAP executables, SAPReorg, SAPTrace, SAPBackup, SAPCheck, SAPArch, SAPStat and the remote copy sets for the cluster system environment.
- A real advantage of mirroring the entire database is that it is a much simpler solution to manage, because it does not require the maintenance of a second database at the standby site.
- A failover to the target site in case of a disaster (DRM unplanned failover) is faster when mirroring the entire database, because recovery is similar to a standard instance recovery for the database after a site failover. Fallback is another matter. In this scenario the entire database has to be renormalized (copied back to the primary site) before the customer can fall back.
- In terms of cost, replicating the whole database in a campus-wide environment is less expensive as only a DRM license is required and there is no need for an additional Oracle license on the target site.

Replicating Oracle Redo Log Information Only

- Here, the Oracle standby database mechanism is used only to replicate Oracle redo log information to the target site via DRM to achieve a disaster tolerant state for the SAP Oracle database. Using the Oracle standby database mechanism without DRM is a common approach at SAP customer sites today, who accept that the latest transactional updates in the Oracle database are lost in the event of a disaster at the primary site. The setup of an Oracle standby database is integrated in the SAP BRBACKUP utility.
- In this scenario all LUN's/storagesets that contain control files, online redo log files and archived redo log files have to be in synchronous remote copy sets.
- With ACS version 8.6, no server access to the remote copy sets on the target site is allowed. Therefore the archived redo log information has to be copied over to the target site and has to be applied regularly. There are several options available to achieve this. One option is to use a TCP based service such as FTP or NFS. If the initiator node and the target node are members of a Tru64 cluster, the archived redo logs can be accessed via the common cluster file system operation and can be directly applied.
  The archived redo log files must still be replicated via DRM. In the event of a disaster, there is no guarantee that the latest archived redo log has been completely copied to the target site before the whole site is lost. As a result, it may be that the DRM replicated online redo log files containing the latest transactional updates could not be applied.
- One advantage of redo log shipping is that transactional updates can be applied to the target database with a timed delay. If the primary database information is destroyed by human error, the target standby database is protected from this kind of error being propagated immediately.
- In general, an Oracle standby database can also run in a read-only mode, allowing the remote machine to be used as a query-only database for reporting and consistency checks. As mySAP.com applications tend to write to the database during startup, this feature has only limited value in this environment.
- A further benefit of replicating log information via DRM is that the SAP Oracle database can include multiple controller HSG80 pairs.

- Comparing the two replication scenarios, replicating only redo log information requires less bandwidth between the two sites. This is not that important in a campus environment where the customer is more flexible to increase bandwidth for moderate costs compared to renting additional bandwidth from a telecommunications company.

Managing DRM Failover and Failback

An essential part of a DRM based solution is the mechanism, for managing a planned/unplanned failover or failback operation in the event of a disaster or maintenance operations. A Compaq-supported utility is the HSG Scripting Tool Kit (HSTK), providing automated failover and failback for DRM. The scripts require a system from which the CLI commands for DRM operations are sent to the HSG80. Two communication configurations are possible, either out-of-band (maintenance port of the HSG80 controller) via terminal server or in-band with either Fibre Channel (command scripter V1.0A) via (command console) LUN or an agent (SANappliance or SWCC). The system running the scripts can be a member of a production cluster or a dedicated server with at least one HBA in case of in-band communication.

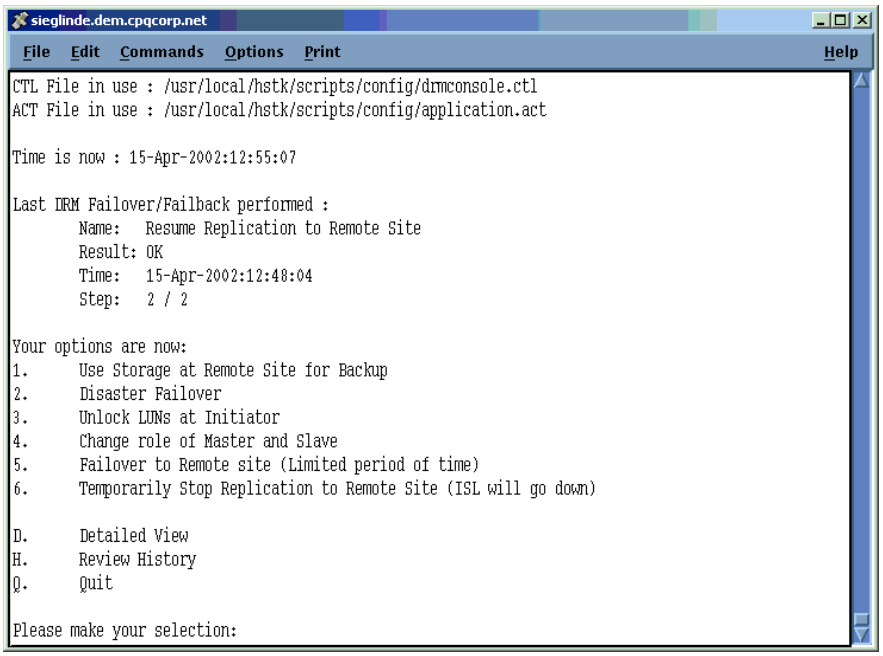

Figure 3 drmconsole Script

In a DRM environment, after a site failover the data is already available at the target site. Failback moves data operations back to the initiator after the initiator site has been brought back online. The drmconsole script in Figure 3 shows the available options for managing a DRM site failover. Option 1 in drmconsole suspends remote mirroring for a volume and enables access to that volume on the target site, for example for a backup server. Using this functionality is beyond the scope of this paper. Please see the *Guide of Operations For Data Replication Manager and Clone and Snapshot scripts* for further information on the supported Perl scripts.

Solution Verification

To verify this solution a system has been set up and configured as described in **Appendix A.** The software versions described are minimum requirement versions for running the solution in a supported environment.

Verification Configuration

The test configuration used for the solution verification is a standard TruCluster setup within the gray box (Figure 4). One cluster member is at the initiator storage site and the second cluster member at the target storage site. The memory channel is configured as a virtual hub (2 km maximum distance between the two cluster members without using a physical hub device). Under normal operating conditions, the shared storage for the cluster system disks and the SAP environment runs at the initiator site. Node A runs the SAP database and node B the central instance service. A LUN/unit D1 is replicated via DRM to D1' to the target site. A unit D2 could be local to the initiator and D3 to the target site.

Please see Appendix A for setup details and software versions.

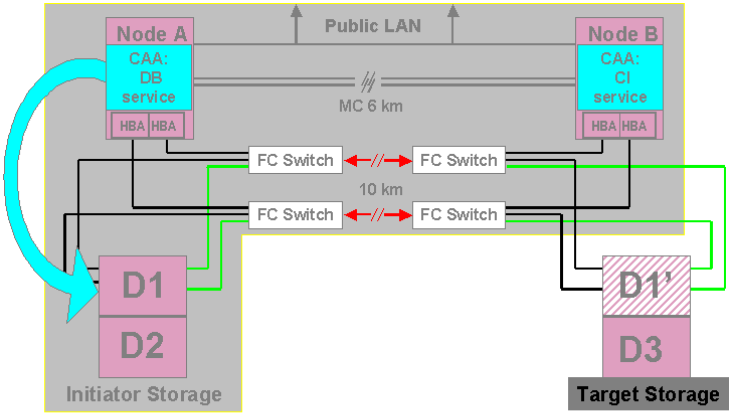

Figure 4 Solution Configuration under Normal Operating Conditions

As referred to in the *Campus-wide DT Cluster Application Notes* the two-node configuration is a special configuration, but very common at SAP Alpha sites. This is because in the event of a node A failure, the database service would automatically be started on Node B and access the shared storage on the initiator site while DRM replication continues. This could result in an overload of the inter-switch links. There are various configuration options for this situation:

- Have more than one ISL per fabric and use static routes on the FC switches or use the 2Gbit FC switch products and the licensed trunking feature, once the 2Gbit products are supported for DRM.
- Fail over the complete site in case of a failure of node A
- Have a second cluster member at the initiator site to serve the shared storage via memory channel

Solution Verification Workload

To verify the functionality and performance of DRM in a TruCluster environment for mySAP.com, an ABAP program is started via transaction SE38 in the SAP frontend. This ABAP inserts a specified number of 200 Byte records into 5 Oracle tables containing unique indices into the USER1 table space of an R/3 standard database. This scenario simulates the behavior of a generic R/3 batch job. The tables will be deleted and recreated after each run to ensure equal conditions for different runs. The size of the configured SAP/ORACLE database has no direct impact on the workload.

To verify the solution in terms of functionality and overhead of DRM, while not focusing on high-water benchmarking of a specific type of server hardware, the ABAP parameters have been adjusted as follows to make sure that neither the servers nor the network will become a bottleneck in the verification scenario:

The default workload specifies an insert/update of **1.5 million records** via 3 SAP D+W processes. A commit occurs every 1000 records. During this workload **5 x 100MB** archived transaction log files are generated. The ABAP provides wall clock time for the whole run (transaction response time), as well as inserted records per second.

The ABAP program is completed in **83 seconds**, and provides the 100% baseline for a non-DRM scenario with all units/LUN's on **one** HSG80 with the exception of the Oracle achieved redo logs. TheDRM overhead will be calculated using this baseline in the two DRM scenarios.

For Tru64 the collect utility measured system status on the Alpha servers in terms of CPU, IO, memory, network usage.

On the SAN switches the portshowperf utility is used to monitor the switch ports.

The vtdpy display status utility is used for non-DRM scenarios to monitor the behavior of the HSG80 controller in terms of idle time.

Replicating the Entire Oracle Database

As discussed above, in this scenario the whole R/3 directory structure and all Oracle database files, with the exception of the achieved redo log files, plus the cluster member boot disks and the cluster file systems (/, /usr, /var) are replicated via DRM to the target site. All remote copy sets are SYNCHRONOUS and in FAILSAFE mode. All database related remote copy sets belong to the same association set.

The ABAP program running the write intensive workload completes after 102 seconds, having generated 5 x 100MB achieved redo log information. After deleting all remote copy sets to run without DRM, the same workload is completed after 83 seconds. This means that the DRM overhead under a heavily write-intensive workload is within the range of 22% compared to a non-DRM scenario as shown in Figure 5.

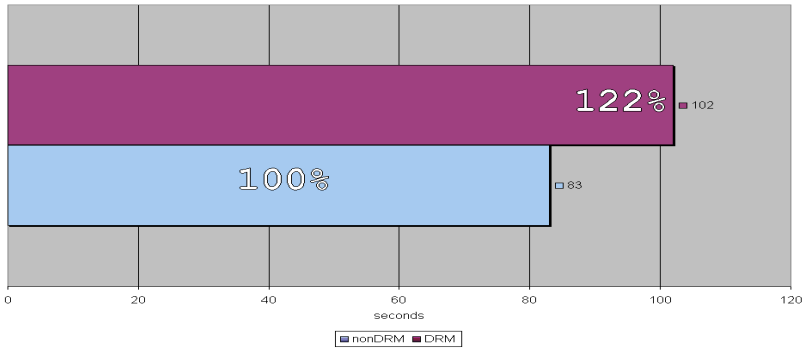

Figure 5 Job Completion Comparison Non-DRM versus DRM Replicating the Entire db

While running the job *without* DRM, the vtdpy performance utility of the HSG80 controller reported 35-40% idle time, which means that this workload puts a considerable load on the HSG80 controller as a controller is supposed to be saturated when around 25% idle time is reported via vtdpy. The collect utility on the database server reported a constant CPU idle time within the range of 40% and a total throughput of 25-28 MB/s.

As a rule of thumb this means that a customer who is already running TruCluster V5 and is considering implementing DRM and experiences an average range of 40% idle time or more on the HSG80 controller, will see about 20% performance decrease in the worst case. Considering a conservative average read/write ratio of 7:3, there will be an overhead of DRM within a range of less than 10%.

Replicating Oracle Redo Log Information Only

When replicating only the LUN's/storagesets containing Oracle redo log and control file information, there is less data to be transferred to the target site. This is reflected in the job completion time of the ABAP program shown in Figure 6.

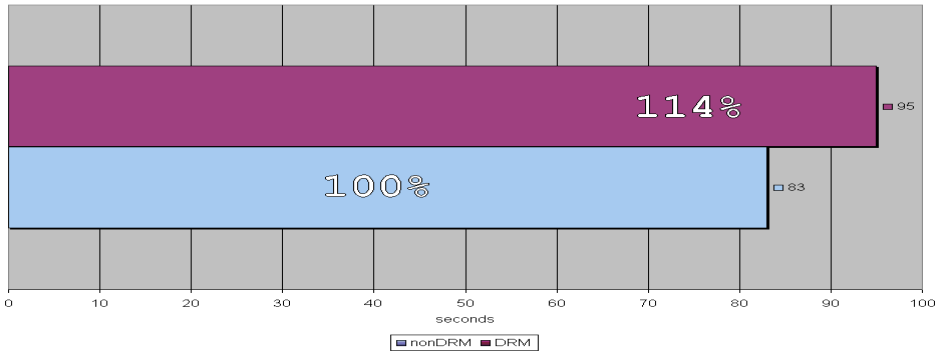

Figure 6 Job Completion Comparison Non-DRM versus DRM Replicating Redo Log Information Only

The write-intensive workload needs 14% more time to complete when the redo log information is replicated, compared to a situation with no DRM active. In a mixed read/write environment, the overhead caused by DRM in this scenario is within the range of 5%.

Please note that the measured DRM overhead in the TruCluster environment is caused only by the additional I/Os for the controllers. The distance between the primary and target controller (up to 2km or 6km) in this environment has hardly no influence on the response time. Please see the *Compaq SANworks Data Replication Manager Inter-site Link Performance Analyzer* white paper on how to calculate the impact of distance for worst-case I/O scenarios for DRM.

Failover and Failback Operations

For various reasons it is necessary to use the service that DRM provides and fail over to the target site. Table 1 lists possible failover situations and the recommended actions in a specific situation. If a type of failure requires a site failover, it is important to verify that all components at the target site are operational before a failover is initiated. It might be preferable in some situations to fix a single component within an acceptable timeframe and continue processing, rather than performing a complete failover

Table 1 Possible Failover Situations

Type of Failure	Recommended action when Error_Mode = Failsafe
Total initiator site loss	Manual Intervention to fail over data processing to target site
Loss of initiator site fabric	Manual Intervention to fail over data processing to target site
Loss of initiator controller pair	Manual Intervention to fail over data processing to target site
Loss of all inter site links	Decide on which side should continue processing
Total target site loss	Manually continue processing at initiator site
Loss of target fabric	Manually continue at initiator site
Loss of target controller pair	Manually continue processing at initiator and target sites
Loss of single initiator controller	Failover not necessary
Loss of both initiator switches	Manual intervention to fail over to the target site and restart of processing at both sites
Loss of a single initiator switch	Failover not necessary
Extended power outage at the initiator site	Manual Intervention to fail over data processing to target site
Loss of a host bus adapter	Failover not necessary
Loss of single disk in redundant storage	Failover not necessary
Loss of single storageset	Failover not necessary
Loss of single host of cluster	Failover not necessary

Unplanned Site Failover (Disaster)

In the event of a series of failures at the Initiator site, this might result in a total loss of access to the storage on this site (Figure 7). This leads to a TruCluster halt situation. A human decision has to be made to initiate a site failover to the target site.

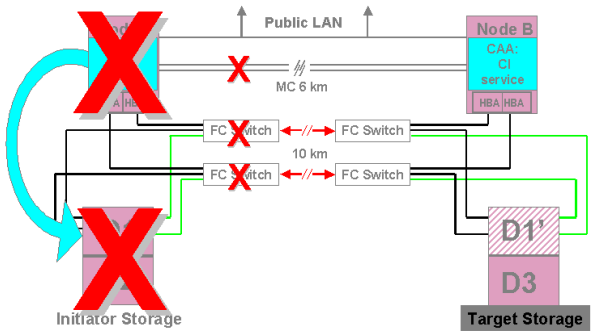

Figure 7 Disaster Situation

The following steps have to be taken to complete an unplanned Site Failover.

1. Run the **drmconsole** script on the management node (see Figure 3) and choose option number **2** (disaster failover). The steps for the failover can be completed within a ten minute time frame.
2. Reboot the cluster member at the target site:

   >>> **boot -fl "ia"**
3. Adjust the quorum expected votes for the cluster when prompted during the boot process:

   'vmunix':**vmunix clubase:cluster_expected_votes=1[Return]**
4. When the system has booted, adjust cluster votes and membership depending on when you expect other cluster members to join as outlined in chapter 4 of the TruCluster Server *Cluster Administration* manual.
   http://www.tru64unix.compaq.com/docs/cluster_doc/cluster_51A/HTML/ARHGYDTE/TITLE.HTM

   Delete the old information about the quorum disk from the initiator site and enable the quorum disk at the target site, if the two cluster nodes do not have the same device special names:

   # clu_quorum -f -d remove

   # clu_quorum -d add <dskXX> 1

5. SAP Oracle database recovery

   a) If the entire database has been replicated, the defined CAA SAP database service will automatically start the Oracle database instance. During the database startup, Oracle will automatically perform an instance recovery. The time it takes until the database is available depends on the number of open transactions during the disaster situation.

   b) If only redo log information has been replicated the following steps must be taken:

   Start up the standby database in recovery mode in order to apply archived redo logs that were not successfully transferred before the disaster occurred.

   SVRMGR> STARTUP NOMOUNT ;
   SVRMGR> ALTER SYSTEM MOUNT STANDBY DATABASE ;
   SVRMGR> RECOVER STANDBY DATABASE;

   Shut down the standby database after all archived redo logs have been applied.

   Apply the replicated online redo log information to the standby database using the latest replicated control file from the initiator site. This is achieved by either specifying a PFILE qualifier in the STARTUP command with a prepared init<SID>.ora for the standby and the production database, or by copying the latest initiator control files over the existing ones or using logical links.

   SVRMGR> STARTUP MOUNT <SID> ;
   SVRMGR> RECOVER DATABASE ;
   SVRMGR> ALTER DATABASE OPEN ;

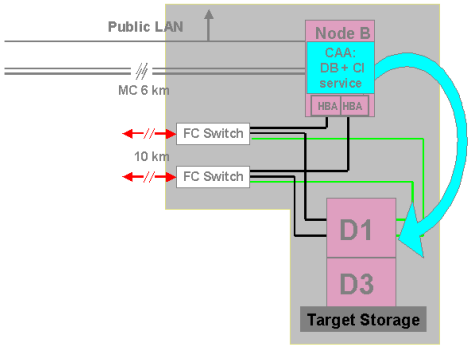

Figure 8 After a Site Failover

After the site failover has been completed, the CAA services for the SAP database and the central instance are running on the target site as shown in Figure 8. The total failover time in the verification scenario is less than 20 minutes.

Failback Procedure

After a planned or an unplanned site failure has occurred, a DRM configuration still has no single point of failure, but is no longer in a disaster tolerant state. To achieve this status again, the necessary actions depend on the customer's disaster plan and strategy, the type of disaster that had occurred, and the replication scenario the customer is using.

To failback the verification configuration to the original primary site, the following steps are taken after an unplanned site failover.

In the scenario replicating the entire database:

- Start the **drmconsole** script (Figure 3) on the management node and choose the appropriate option. The script is aware of a preceding site failover and offers only valid actions.
- In a *first step* of the script, the controller at the initiator site is restarted and prepared for the renormalization of the remote copy sets for the Oracle database and the cluster LUN's. This step can run while SAP services are up and running.
- In a *second step* of the script, the renormalization process is started. This will impact the performance of the running SAP environment as shown in Figure 9 .
- *Before the third step* of the script can start, the SAP service and the cluster needs to be shutdown, because this step will disable access to the units at the target site and failback the remote copy sets to the initiator site.
- *After the third step* of the script is completed the cluster can be rebooted and is in the same state as before the disaster had occurred.

In the scenario for replicating redo log information simply:

- Establish a standby database on the node at the primary site on the primary site storage. The Oracle Storage Compatibility Testing - Remote Mirroring White Paper suggests several ways to achieve this:
  Reverse role via database copy and via restoring backup
  Reverse role via recovery
  Direct fallback via DB Copy
  Direct fallback via restoring backup
- Run the **drmconsole** script on the management node to failback the remote copy sets for this scenario as described above. In the *second step* of the script the performance of the running SAP service will hardly be affected, as only the LUN containing redo log information is normalized.
- Before the SAP CAA services can be started after the cluster reboot, the standby database has to be activated on the primary site in the same way as during a site failover.

Path Failure and Normalization Process

At both sites, a DRM configuration has No Single Point Of Failure in the I/O path from the server to the data on disk. There are at least two paths in two distinct fabrics to ensure that an unplanned site failover (disaster) can only occur if a series of failures occur. To test this functionality in a mySAP.com environment, a path failure was simulated by powering off one FC switch while the ABAP program was running. The path failure was acknowledged after 30 seconds and completed after 80 seconds (Figure 9). This resulted in a job completion time of 187 seconds compared to the 102 seconds, which is the job completion time for replicating the entire database without error conditions.

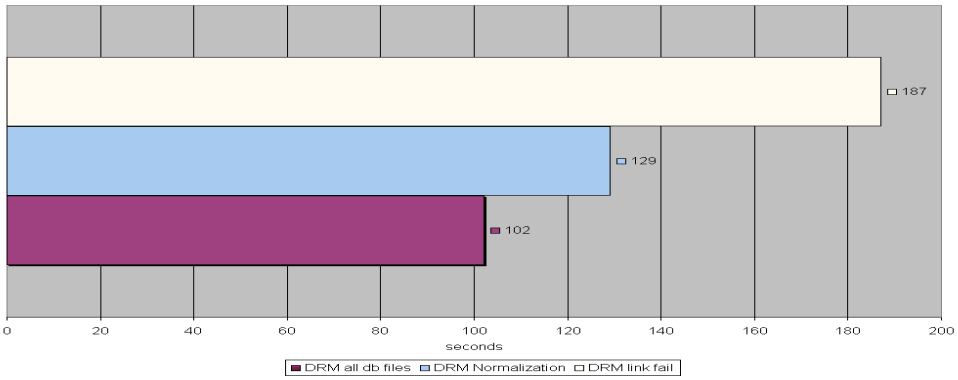

Figure 9 Path Failure and Normalization Process

The normalization process in a DRM configuration is the full copy of a LUN between the initiator and the target site controllers with a size of 64KB I/O. This must happen when a remote copy set is created or once the two sites are out of synchronization. In the verification configuration, the portperfshow utility for the FC switches reported 8 MB/s for a two-member mirror set and 15 MB/s for a six-member raid 0+1 set. Normalizing five remote copy sets in parallel on one HSG80 controller never exceeded 30MB/s.

Running the write-intensive ABAP program while a full normalization was in process, the job completion time increased by about 30% to 129 seconds (Figure 9). This means that it might not be acceptable for a customer to renormalize after a disaster to prepare a failback while the SAP service is available.

General Recommendations

The StorageWorks DataSafe solution for mySAP.com enhances the TruCluster high-availability features with the disaster tolerant capabilities of DRM, maintaining an acceptable DRM overhead in terms of performance.

Replicating the whole SAP Oracle database is a robust, managed solution for SAP customers that do not currently need more than one/two HSG80 pairs. This solution provides almost full FC speed over the TruCluster distance and a failover/recovery time at a remote computing site measured in minutes. This scenario allows existing non-DRM TruCluster V5 customers a straightforward implementation of DRM with careful planning.

Replicating only SAP Oracle Redo log information via DRM using the Oracle Standby Database mechanism provides DT functionality up to the latest transactional update as well. This scenario overcomes the limitation of one/two HSG80 pairs for the total configuration. In addition, this scenario requires less bandwidth and allows database changes to be propagated with a timed delay at the target site to protect the standby database from human error. The tradeoff for this scenario is the additional management effort required to maintain the standby database and the Oracle database expertise necessary in case of a site failover.

Replication beyond the Campus-wide Distance

For SAP customers who need to span a distance beyond the campus-wide distance, Tru64 supports synchronous replication up to 70 km. In this case, an SAP configuration can have a TruCluster configuration at the initiator site replicating the SAP Oracle database or the redo log information to a target site with a single system or a second cluster. This configuration provides a highly available mySAP.com application with a disaster-tolerant database. In a typical SAP environment with an SAP test/development system in addition to the production system, the test/development system is placed at the target site and becomes the production system in the event of a disaster.

Customers, who run more than one mySAP.com application have the option of active/active bi-directional replication when they have a production system at each computing site. In this case, each site is an initiator and target for different applications with at least two HSG80 controller pairs per site. With additional servers and storage, the bi-directional DRM configuration can scale up to 20 switches in each fabric, supporting, for example, 96 servers and 8 arrays at each site. All intersite technologies, such as DRM over ATM, DRM over direct fiber, DRM over IP, and DRM over WDM, support bi-directional use of DRM.

Depending on the mySAP.com application and the customer's workload an evaluation must be carried out according to the distance to determine the interlink technology required to maintain performance and consistency. The *Data Replication Manager HSG80 ACS Version 8.6-4P Configuration Guide* (ftp://ftp.compaq.com/pub/products/SANworks/techdoc/drm/AA-RPHZD-TE.pdf) provides the configuration options that are currently supported.

Appendix A - Description and Setup of the Verification Configuration

Table 2 Hardware

	Initiator Site	#	Target Site	#
Server	ES40	1	ES40	1
	CPU (440 MHz)	4	CPU (600 MHz)	2
	2 GB memory		1 GB memory	
	KGPSA-BC	2	KGPSA-BC	2
	MC II	1	MC II	1
Storage	HSG80 pair	1	HSG80 pair	1
	10K RPM disk drives	24	10K RPM disk drives	24
FC Infrastructure	SAN Switch/16	1	SAN Switch/16	1
	SAN Switch/8	1	SAN Switch/8	1
SAN mgmt node	DS10			1
Client (SAP)	PL1850R			1
Network Infrastructure	The servers and the client are connected via a 10/100 NICs			

Table 3 Software

Software	Version
Tru64 and TruCluster	5.1A and IPK1
Array Controller Version ACS	8.6-4 P
Fabric OS	2.6
SAP R/3	4.6D
Oracle	8.1.7
Cluster Scripts for SAP	V009
Command Scripter	1.0A
DRM Failover Scripts	1.6

Table 4 SANswitch Port Allocation

	Port	0	1	2	3	4	5	6	7
Initiator	TopFabric	PRI TOP1 (HSG80)	PRITOP2 (HSG80)		ES40-INI-HBA1				ISL
	BottomFabric	PRI BOT1 (HSG80)	PRIBOT2 (HSG80)		ES40-INI-HBA2				ISL
Target	TopFabric	TARTOP1 (HSG80)	TARTOP2 (HSG80)		ES40TAR-HBA1		DS20HBA1		ISL
	BottomFabric	TARBOT1 (HSG80)	TARBOT2 (HSG80)		ES40TAR-HBA2				ISL

Table 5 Identical Storage Configuration at Initiator and Target Site

	SCSI 1	SCSI 2	SCSI 3	SCSI 4	SCSI 5	SCSI 6
Position 0	RCS11		RCS12			
	D11 BOOTWOG		D12 BOOTSIEG		D21 CLQUORUM	
	DISK10000 18.2GB/10K	DISK20000 18.2GB/10K	DISK3000 18.2GB/10K	DISK40000 18.2GB/10K	DISK50000 18.2GB/10K	DISK60000 18.2GB/10K
Position 1						
			RCS43		RCS22	
	D42 DRMLOG		D43 SAPARLOG		D22 CLSSYSTEM	
	DISK10100 18.2GB/10K	DISK20100 18.2GB/10K	DISK30100 18.2GB/10K	DISK40100 18.2GB/10K	DISK50100 18.2GB/10K	DISK60100 18.2GB/10K
Position 2						
	RCS44					
	D44					
	SAPLLOG2					
	SAPMIRL1		SAPMIRL2		SAPMIRL3	
	DISK10200 18.2GB/10K	DISK20200 18.2GB/10K	DISK30200 18.2GB/10K	DISK40200 18.2GB/10K	DISK50200 18.2GB/10K	DISK60200 18.2GB/10K
Position 3						
	RCS41					
	D41					
	SAPSTR1					
	SAPMIR1		SAPMIR2		SAPMIR3	
	DISK10300 18.2GB/10K	DISK20300 18.2GB/10K	DISK30300 18.2GB/10K	DISK40300 18.2GB/10K	DISK50300 18.2GB/10K	DISK60300 18.2GB/10K

Table 6 Database Configuration

Name	Drives	RAID	LUN	Description
SapArch	2 x 18.2-GB	1+0	D43	Archived redo logs
OrigLogA	6 x 18.2-GB	1+0	D44	Redo log set A
OrigLogB				Redo log set B
MirrLogA				Mirrored redo log set A
MirrLogB				Mirrored redo log set B
Oracle	6 x 18.2-GB	1+0	D41	Oracle exec
SapReorg				Reorg, backup, check, stat and trace
SapData1				DB data area 1

Oracle Parameter (init<SID>.ora)

```
Parameter
Sort_area_retained_size = 0
Sort_area_size = 2097152
Open_cursors = 750
dml_locks = 1250
enqueue_resources = 4000
db_file_multiblock_read_count = 8
transactions_per_rollback_segment = 20
log_archive_start = true
log_checkpoint_interval = 3000000000
log_checkpoint_timeout = 0
Processes = 100
sessions = 45
timed_statistics = true
db_name = T5O
db_block_size = 8192
db_files = 254
Optimiser_mode = choose
Optimiser_index_cost_adj = 10
Optimiser_search_limit = 3
background_dump_dest = /oracle/T5O/saptrace/background
user_dump_dest       = /oracle/T5O/saptrace/usertrace
core_dump_dest       = /oracle/T5O/saptrace/background
log_archive_dest     = /oracle/T5O/saparch/T5Oarch
Remote_os_authent = true
distributed_transactions = 0
replication_dependency_tracking = FALSE
transaction_auditing = FALSE
control_file_record_keep_time = 30
IFILE = /oracle/T5O/dbs/init_806.ora
IFILE = /oracle/T5O/dbs/initora.addon
Shared_pool_size = 382520524
Shared_pool_reserved_size = 38252052
db_block_buffers = 50000
control_files = (
/oracle/T5O/sapdata1/cntrl/cntrlT5O.dbf,/oracle/T5O/ori
glogA/cntrl/cntrlT5O.dbf,/oracle/T5O/saparch/cntrl/cntr
lT5O.dbf )
MAX_ROLLBACK_SEGMENTS = 400
rollback_segments =
(PRS_0,PRS_1,PRS_2,PRS_3,PRS_4,PRS_5,PRS_6,PRS_7,PRS_8,
PRS_9,PRS_10,PRS_11,PRS_12,PRS_13,PRS_14,PRS_15,PRS_16,
PRS_17,PRS_18,PRS_19)
#log_archive_buffer_size      127
#log_archive_buffers      4
log_buffer = 1126400
```

Set up Zoning in a DRM Configuration

The *Data Replication Manager HSG80 ACS Version 8.6-4P Configuration Guide* explains the way zoning on the FC switches is set up in a DRM environment. Figure 10 shows the three zones of the top fabric plus the management zone through the telnet interface of the initiator top FC switch:

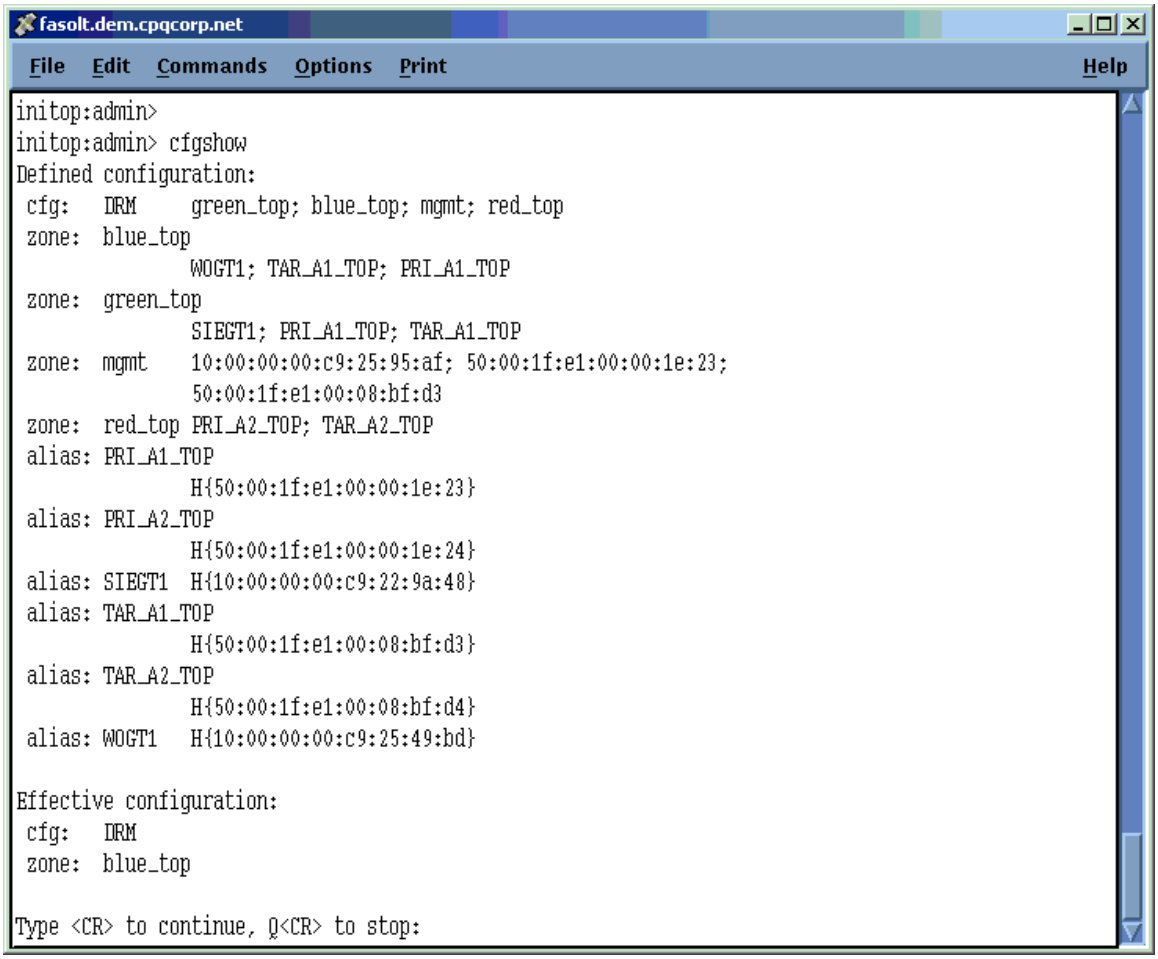

Figure 10 Zoning DRM

Set up the Tru64 Boot Device in the SRM Console

To boot an AlphaServer via FC Host Bus Adapter, the related environment variables have to be set via the World Wide ID Manager utility . The WWID MGR users' manual (http://cybrary.inet.cpqcorp.net/ARCHIVE/PUBS/USERS/WWIDMGR_V11.pdf) specifies how to do this. Figure 11 shows the setup for the verification environment:

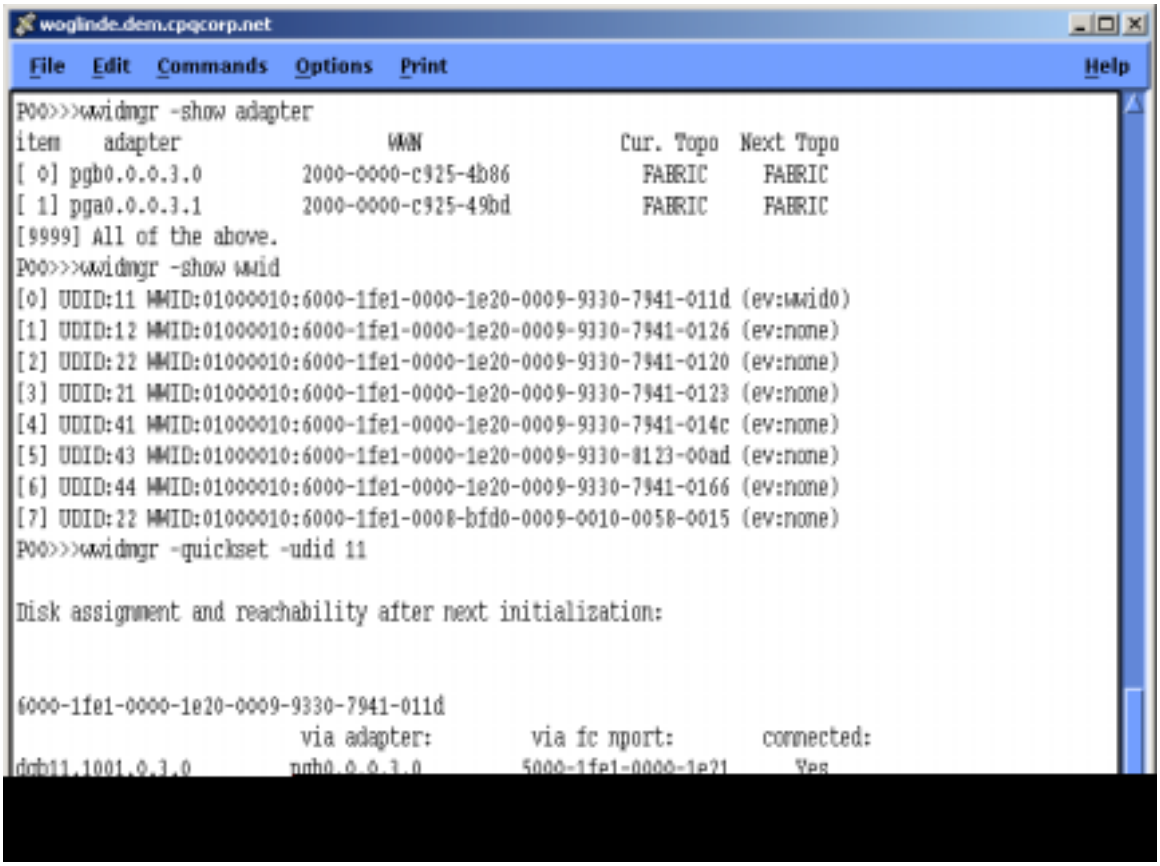

Figure 11 Set the FC Boot Device with wwidmgr in the SRM Console

Configuring Storage

- Install the cable initiator and target site storage, FC switches and target site storage as described in *Data Replication Manager HSG80 ACS Version 8.6-4P Configuration Guide* (ftp://ftp.compaq.com/pub/products/SANworks/techdoc/drm/AA-RPHZD-TE.pdf)

- Set up initiator site and target site HSG80 storage the same way at the CLI level:

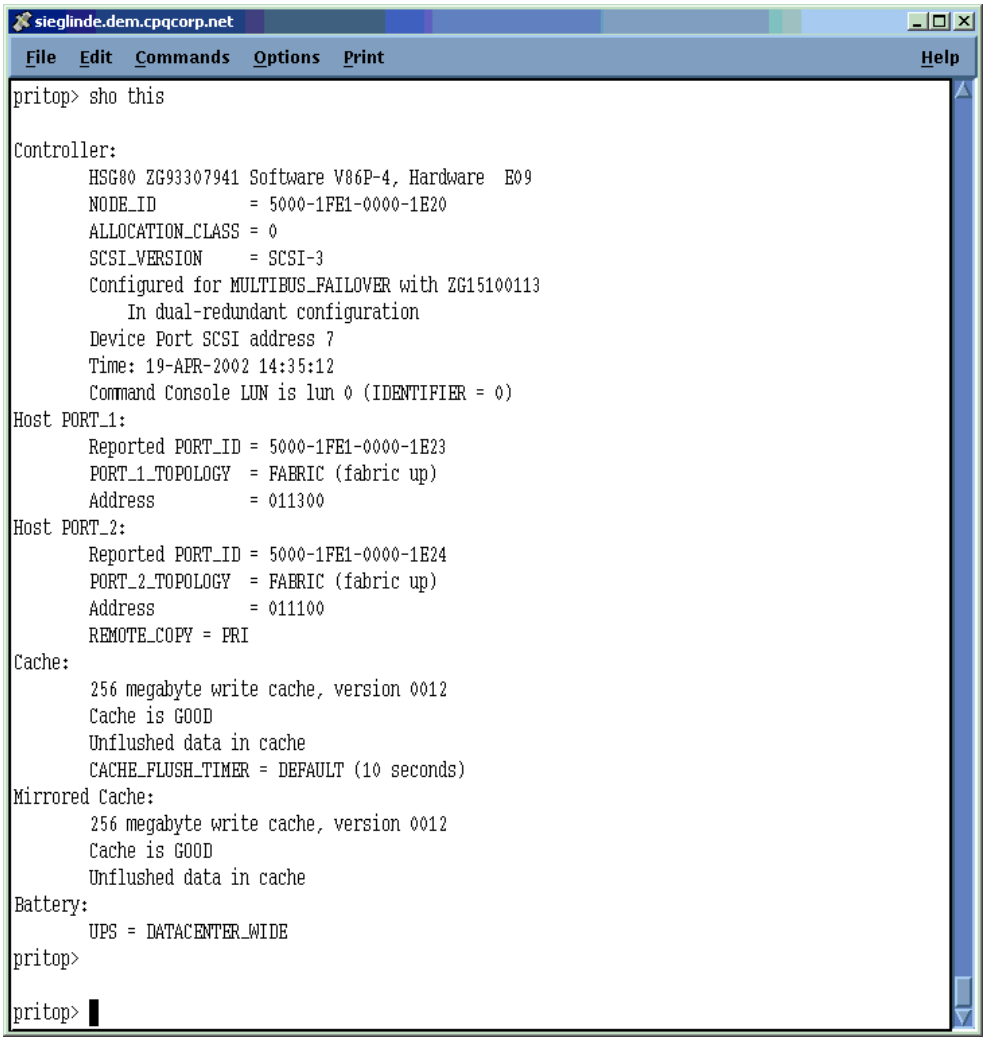

Figure 12 HSG80 Initiator Site

- Managing the HSG80 connection table

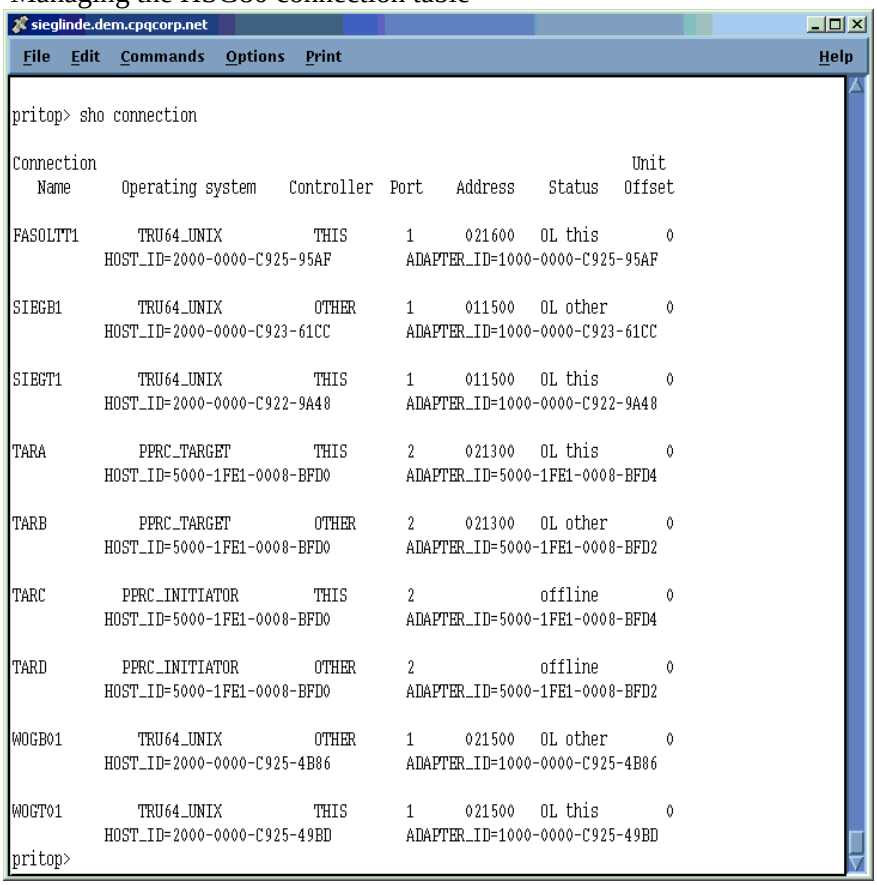

Figure 13 Initiator Site Connection Table

- Set up remote copy sets

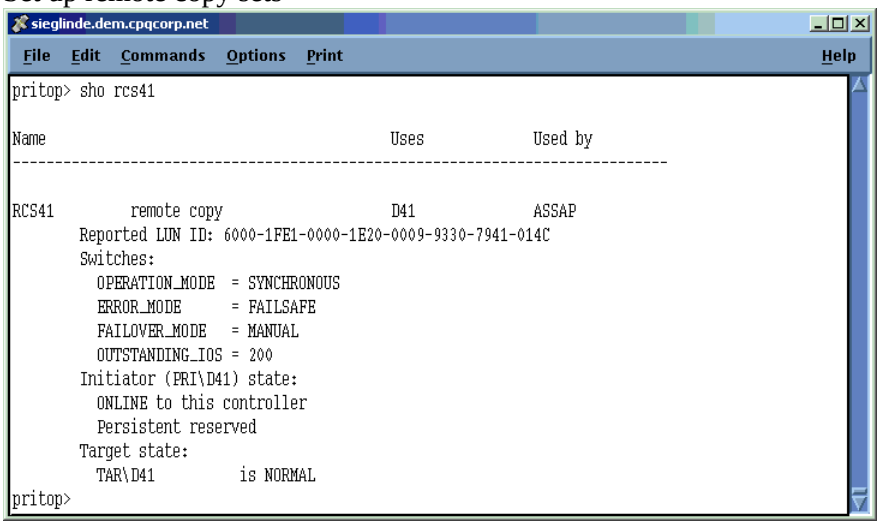

Figure 14 Remote Copy Set

- Set up the association set for the SAP Oracle database

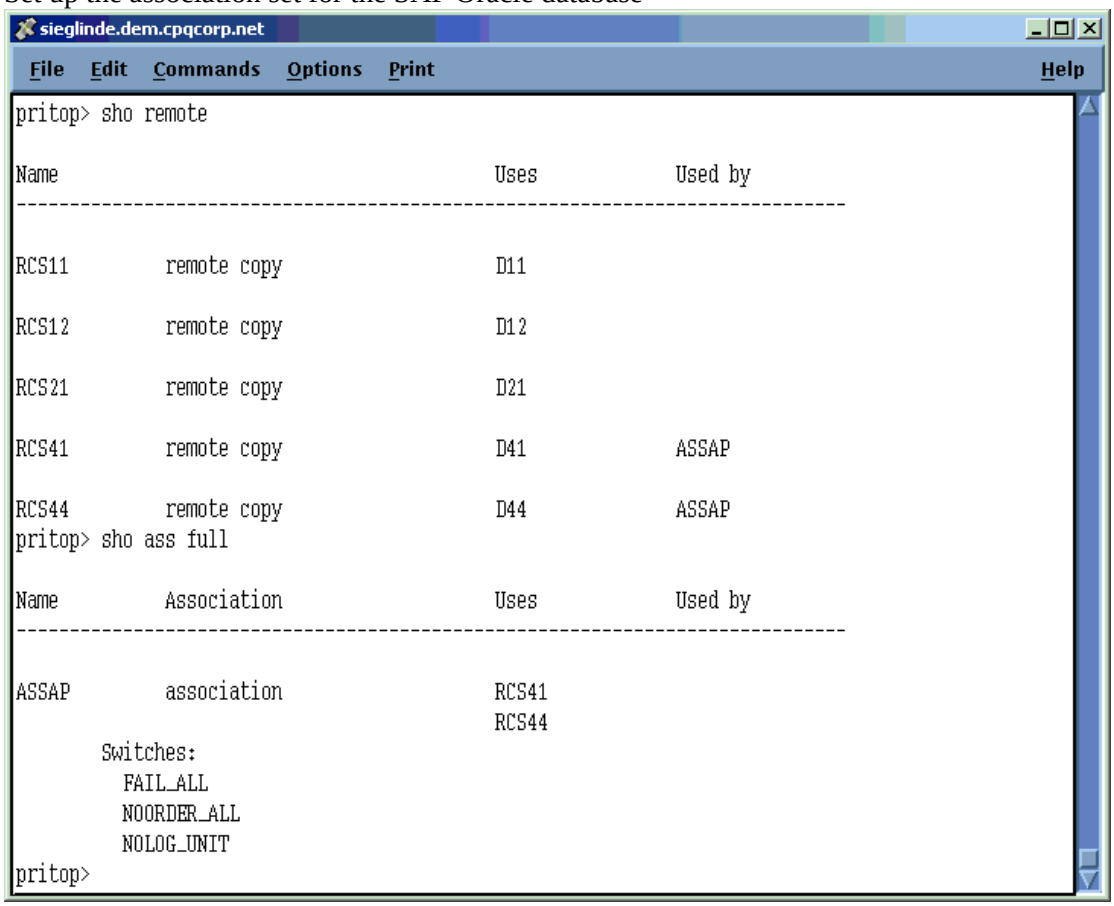

Figure 15 Association Set for the Replicated Database

Set up the Cluster Scripts for mySAP.com

After installing the operating system and TruCluster, the setup of the cluster scripts for mySAP.com is shown in Figure 16.

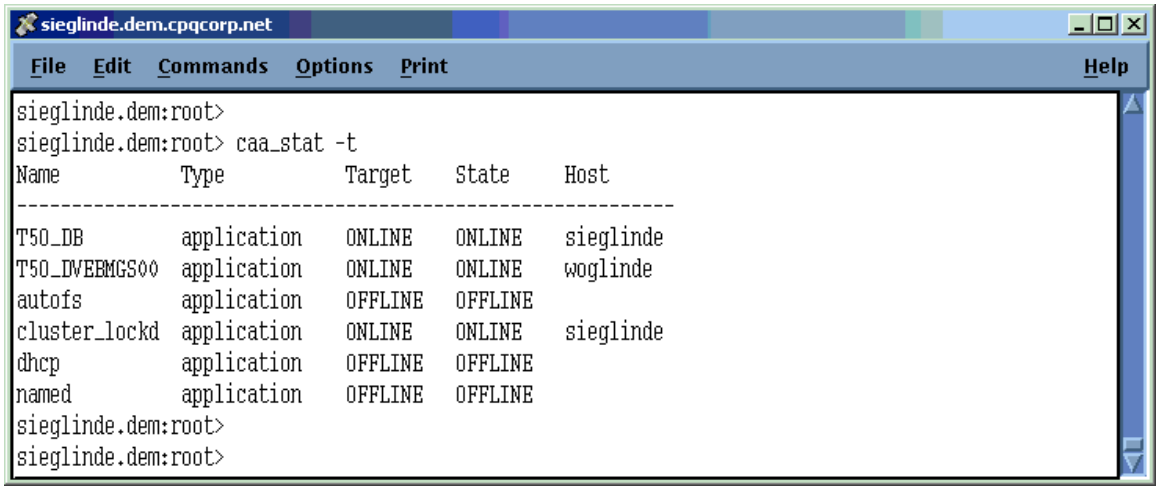

Figure 16 CAA Setup for the Database Service (T50_DB) and the Central Instance (T5O_DVEBMGS00)

Set up the HSG Scripting Tool Kit (HSTK)

1. Install Command Scripter V1.0 and upgrade to V1.0A . Verify that Command Scripter can communicate with the HSG80.
   ```
   # cmdscript –fscp0
   ```
2. Install the HSTK
   ```
   # perl install.pl
   ```
3. Set the environment variable CLONE_HOME
   ```
   # export CLONE_HOME =/usr/local/hstk/scripts
   ```
4. Create the configuration file of the target site controller
   ```
   # /usr/local/hstk/scripts/sh/generate_cfg.sh com=cs tar scp1
   ```
5. Create the configuration file of the initiator site controller
   ```
   # /usr/local/hstk/scripts/sh/generate_cfg.sh com=cs pri scp0
   ```
6. Adjust the access rights in target site configuration file in preparation for a site failover
7. Run `drmconsole` to verify the HSTK works (Figure 3).

Appendix B - Related Documents

- Guide of Operations For Data Replication Manager and Clone and Snapshot scripts (HSTK) http://storage.jgo.cpqcorp.net/perl-scripts/V1.60/
- SANworks Data Replication Manager HSG80 ACS Version 8.6-1P Scripting User Guide by Compaq, Part Number EK-DRMSC-OA.
- SANworks Data Replication Manager HSG80 ACS Version 8.6-4P Failover/Failback Procedures Guide by Compaq, Part Number AA-RPJOC-TE
- SANworks Data Replication Manager HSG80 ACS Version 8.6-4P Configuration Guide by Compaq, Part Number AA-RHPZB-TE
- SANworks Release Notes - Data Replication Manager HSG80 ACS Version 8.6-4P by Compaq,
- Guidelines for Testing Remote Mirroring Storage Systems for Oracle Database, Bill Lee and Cynthia Yip
- Guidelines for Using Remote Mirroring Storage Systems for Oracle Database, Bill Lee
- Oracle Storage Compatibility Testing - Remote Mirroring White Paper
- Using Data Replication Manager with Oracle8i Under Tru64
- Disaster Tolerance white paper: Disaster Tolerance - ftp://ftp.compaq.com/pub/solutions/customsystems/disaster_tolerance_business_continuity.pdf
- High Availability in the Compaq Tru64 UNIX environment http://www.tru64unix.compaq.com/unixhawp.pdf
- Campus-Wide DT Cluster Application Note
- Compaq SANworks Data Replication Manager Inter-site Link Performance Analyzer white paper ftp://ftp.compaq.com/pub/products/SANworks/drm/142G-0100A-WWEN.pdf
- HSG80 ACS Version 8.6-4P Data Replication Manager Design Guide - Application Notes
- Features and Benefits of HSG80 ACS 8.6P Data Replication Manager - White Paper

Appendix C - Glossary

array controller software *(ACS*): Software that is contained on a removable PCMCIA program card that provides the operating environment for the array controller.

association set: A group of remote copy sets that share common attributes. Members of an association set can be configured to transition to the same state at the same time. An association set:

• Shares the same log unit

• Has its host access removed from all members when one member fails

• Keeps I/O order across all members

CLI commands available are ADD ASSOCIATIONS and SET ASSOCIATIONS.

asynchronous mode: A mode of operation of the remote copy set whereby the write operation provides command completion to the host after the data is safe on the initiating controller, and prior to the completion of the target command. Asynchronous mode can provide greater performance and faster response time, but the data on all members at any time cannot be assumed to be identical. *See also* synchronous mode.

availability: The percentage of time a functional unit is operational during a given interval of time. The interval of time and "operational" are defined by the requirements of the user.

block: A stream of data stored on disk or tape media that is transferred and error-checked as a unit. In a disk drive, a block is also called a *sector* (the smallest collection of consecutive bytes addressable on a disk drive). In integrated storage elements, a block contains 512 bytes of data, error codes, flags, and the block address header.

business continuity: This is the broadest term that covers all aspects of keeping your business *in* business including recovery, planning, information technology, environmental, and crisis situations. The concept of business continuity is gaining wide acceptance throughout the business world.

cache: A fast, temporary storage buffer in a controller or computer

cache memory: A portion of high-speed memory used as an intermediary between a data user and a larger amount of storage. The objective of designing cache into a system is to improve performance by placing the most frequently used data in the highest performance memory.

cascaded switch: As applied to the Data Replication Manager, a cascaded switch is one where its output is connected to the input of another switch, which then may in turn be connected to another switch or host or controller.

CLI: Command Line Interface. The CLI is the configuration interface that operates the controller software.

clone: A utility that physically duplicates data on any unpartitioned single-disk unit, stripeset, mirrorset, or striped mirrorset.

connection: As applied to the Data Replication Manager, a connection between two end Fibre Channel ports. An example is the connection between a Host Bus Adapter (by way of the Fibre Channel Switches) and the HSG80 controller. CLI commands available are ADD CONNECTIONS, SET *connection name*. *See also* link.

controller: A hardware device that uses software to facilitate communications between a host and one or more storage devices organized in an array. The HS-series StorageWorks family of controllers are all array controllers.

controller failover: The process that takes place when one controller in a dual-redundant configuration assumes the workload of a failed companion controller. Failover continues until the failed controller is repaired or replaced. The CLI command is SITE_FAILOVER. *See also* failback, dual-redundant configuration, *and* planned failover.

copying member: In a mirrorset, a copying member is a container introduced to the mirrorset after the mirrorset has already been in use. None of the blocks can be guaranteed to be the same as other members of the mirrorset. Therefore the *copying* member is made the same by copying all the data from a *normal* member. This is in contrast to *normalization*, where all blocks written since creation are known to be the same. When all of the blocks on the copying member are the same as those on the normal member, the copying member becomes a normal member. Until it becomes a normal member, the copying member contains undefined data and is not useful for any purpose.

data integrity: The assurance that the data you receive is exactly what was sent you and that it stays that way until deliberately modified. Hardware problems, power failures, disk crashes, or software errors can threaten data integrity.

disaster tolerance: 1. As applied to high availability, the ability to maintain data integrity in operations following a catastrophic event to a computing site. 2. As applied to DRM, disaster tolerance provides the ability for rapid recovery of user data from a remote location when a significant event or a disaster occurs at the primary computing site. *See also* remote copy sets.

dual-redundant configuration: A storage subsystem configuration consisting of two active controllers operating as a single controller. If one controller fails, the other controller assumes control of the failing controller's devices. *See also* controller failover, site failover *and* failback.

fabric: A network of Fibre Channel switches or hubs and other devices.

failback: The process of restoring data access to the newly-restored controller in a dual-redundant controller configuration. The failback method (full copy or fast-failback) is determined by the enabling of the Logging or Failsafe switches, the selected mode of operation (synchronous or asynchronous), and whether the failover is planned or unplanned. *See also* controller failover, site failover, *and* dual-redundant configuration.

failover: *See* controller failover *and* site failover.

failsafe locked: The failsafe error mode can be enabled by the user to fail any write I/O whenever the target is inaccessible or the initiator unit fails. When either of these conditions occurs, the remote copy set goes into the inoperative (offline) state and the failsafe error mode is "failsafe locked." The CLI command SET *remote-copy-set-name* ERROR_MODE=FAILSAFE enables this error mode.

fast-failback: The synchronization of the initiator site with the target during a planned failover of the initiator subsystem. Write operations are logged to the target site write history log and, during the fast-failback, the initiator site is updated from the write history log. *See also* mini-merge, unplanned failover, planned failover, *and* write history logging.

Fibre Channel: An ANSI standard name given to a low-level protocol for a type of serial transmission. The Fibre Channel specifications define the physical link, the low level protocol, and all other pertinent characteristics.

frame: The basic unit of communication using Fibre Channel protocol. Each frame consists of a payload encapsulated in control information. The initiator breaks up the exchange into one or more sequences, which in turn are broken into one or more frames. The responder recombines the frames into sequences and exchanges. *See also* initiator.

ISL: Intersite link or Inter switch link. The abbreviation is context sensitive. *See also* multiple intersite links.

Initiator: 1. A SCSI device that requests an I/O process to be performed by another SCSI device, namely, the SCSI target. The controller is the initiator on the device bus. 2. For subsystems using the disaster tolerance Data Replication Manager solution, the initiator is the site that is the primary source of information. In the event of a system outage, the database would be recovered from the target system. *See also* target.

latency: The amount of time required for a transmission to reach its destination.

link: A connection between two adjacent Fibre Channel ports, consisting of a transmit fiber and a receive fiber. An example is the connection between the Fibre Channel switch port and the HSG80 controller. *See* also connection.

logical unit: A physical or virtual device addressable through a target ID number. The logical unit numbers (LUN's) use their target's bus connection to communicate on the SCSI bus. *See* also LUN.

Logical Unit Number: *See* LUN.

LOG_UNIT: A CLI command switch that (when enabled) assigns a single, dedicated log unit for a specific association set. The association set members must all be in the NORMAL error mode (not failsafe).

LUN: Logical Unit Number. A value that identifies a specific logical unit belonging to a SCSI target ID number. A number associated with a physical device unit during a task's I/O operations. Each task in the system must establish its own correspondence between logical unit numbers and physical devices.

Mean Time Between Failure *(MTBF)*: a statistically derived length of time a user may reasonably expect a component, device, or system to work between two incapacitating failures.

mini-merge: As applied to the Data Replication Manager, the data transfers to be made whenever a target becomes inaccessible. This occurs when both links or both target controllers have gone down. The transfers that would have been made are instead logged into the association set's assigned log unit to wait until the remote copy set subsystem comes back online. *See* fast-failback, write history logging.

mirroring: The act of creating an exact copy or image of data.

mirrorset: 1. A group of storage devices organized as duplicate copies of each other. Mirrorsets provide the highest level of data availability at the highest cost. Another name for RAID 1. Also called *mirrored units* or *mirrored virtual disks*.

2. Two or more physical disks configured to present one highly reliable virtual unit to the host.

3. A virtual disk drive consisting of multiple physical disk drives, each of which contains a complete and independent copy of the entire virtual disk's data.

multiple intersite links

Each intersite link (ILS) is a fiber link between two switches. As applied to Data Replication Manager, increasing bandwidth between switches is handled by adding additional connections between the switches, to a maximum of two connections.

mission critical: A term applied to information systems upon which the success of an organization depends and the loss of which results in unacceptable operational or financial harm.

normal member: A mirrorset member that, block-for-block, contains exactly the same data as that on the other members within the mirrorset. Read requests from the host are always satisfied by normal members.

normalizing: A state in which, block-for-block, data written by the host to a mirrorset member is consistent with the data on other normal and normalizing members. The normalizing state exists only after a mirrorset is initialized. Therefore, no customer data is on the mirrorset.

normalizing member: A mirrorset member whose contents are the same as all other normal and normalizing members for data that has been written since the mirrorset was created or since lost cache data was cleared. A normalizing member is created by a normal member when either all of the normal members fail or all of the normal members are removed from the mirrorset. *See also* copying member

other controller: The controller in a dual-redundant pair that is not connected to the controller serving your current CLI session with a local terminal. *See also* this controller *and* local terminal.

planned failover: As applied to the Data Replication Manager, an orderly shutdown of the controllers for installation of new hardware, updating the software, and so on. The host applications are quiesced and all write operations permitted to complete before the shutdown. The controllers must be in synchronous operation mode before starting a planned failover. *See also* synchronous mode *and* unplanned failover.

reliability: A measure of how dependable a component or system is once it is in use. Availability might be considered the sum of reliability and data integrity.

remote copy sets: A feature that allows data to be copied (mirrored) from the originating site (initiator) to a remote site (target). The result is a mirror copy of the data (remote copy set) at two disparate sites. Used in disaster tolerant (DT) applications such as the Data Replication Manager. CLI commands available are ADD REMOTE_COPY_SETS, SET *remote-copy-set-name*, SET *controller* REMOTE_COPY.

remote copy set metadata: Data that describes the remote copy set membership and state. To assist with site failover, this metadata is located in the mirrored write-back cache on the controller where each member resides. Backup copies of the metadata reside in the controller NVRAM at each site. Only the initiator modifies the metadata and ensures all copies are subsequently updated.

site failover: The process that takes place when storage processing is moved from one pair of controllers to another. All processing is shifted to the target (remote) site. This is possible because all data generated at the initiator site has been replicated at the target site, in readiness for such a situation.

storage array: An integrated set of storage devices. Storage arrays can be manipulated as one unit with a single command.

storage unit: The generic term for storagesets, single-disk units, and all other storage devices that are installed in a subsystem and accessed by the host. A storage unit can be any entity that is capable of storing data, whether it is a physical device or a group of physical devices.

storageset: 1. A group of devices configured with RAID techniques to operate as a single container. 2. Any collection of containers, such as stripesets, mirrorsets, striped mirrorsets, JBODs, and RAIDsets.

surviving controller: The controller in a dual-redundant configuration pair that serves its companion's devices when the companion controller fails.

synchronous mode: A mode of operation of the remote copy set whereby the data is written simultaneously to the cache of the initiator subsystem and the cache of the target subsystem. The I/O completion status is not sent until all members of the remote copy set are updated. *See also* asynchronous mode.

target: A SCSI device that performs an operation requested by another SCSI device, namely the SCSI initiator. The target number is determined by the device's address on its SCSI bus. For subsystems using the disaster-tolerant Data Replication Manager solution, data processing occurs at the initiator site and the data is replicated or mirrored to the target site. In the event of a system outage, the database is recovered from the target system. *See also* initiator.

this controller: The controller that is serving the current CLI session through a local or remote terminal. *See also* other controller.

unit: A container made accessible to a host. A unit may be created from a single disk drive or tape drive. A unit may also be created from a more complex container, such as a RAID set. The controller supports a maximum of eight units on each target.

unplanned failover: As applied to the Data Replication Manager, recovery from an unplanned outage of the controllers. This may occur when the site communication is lost or it may be due to some other failure whereby remote copy sets cannot be implemented. The controllers do not perform an orderly shutdown. *See also* planned failover.

write history logging: As applied to the Data Replication Manager, the use of a log unit to log a history of write commands and data from the host. Write history logging is used for mini-merge and fast-failback. *See* mini-merge *and* fast-failback.

zoning: As applied to the Data Replication Manager, an optional, licensed feature of the SilkWorm switch that allows a finer segmentation of Storage Area Networks (SANs) by allowing ports or WWN addresses to confine access to devices that are in a common zone.