Guidelines for using IBM® TotalStorage® DS Series Family For Oracle® Database Disaster Recovery Using Global Mirror

Author: Betty Mason

Last Updated: October 21, 2005

Version : 1.0

Table of Contents

Product data has been reviewed for accuracy as of the date of initial publication. Product data is subject to change without notice. This information could include technical inaccuracies or typographical errors. IBM may make improvements and/or changes in the product(s) and/or programs(s) at any time without notice.

The performance data contained herein was obtained in a controlled, isolated environment. Actual results that may be obtained in other operating environments may vary significantly. While IBM has reviewed each item for accuracy in a specific situation, there is no guarantee that the same or similar results will be obtained elsewhere.

References in this document to IBM products, programs, or services does not imply that IBM intends to make such products, programs or services available in all countries in which IBM operates or does business. Any reference to an IBM Program Product in this document is not intended to state or imply that only that program product may be used. Any functionally equivalent program, that does not infringe IBM's intellectually property rights, may be used instead. It is the user's responsibility to evaluate and verify the operation of any non-IBM product, program or service.

THE INFORMATION PROVIDED IN THIS DOCUMENT IS DISTRIBUTED "AS IS" WITHOUT ANY WARRANTY, EITHER EXPRESS OR IMPLIED. IBM EXPRESSLY DISCLAIMS ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE OR NONINFRINGEMENT. IBM shall have no responsibility to update this information. IBM products are warranted according to the terms and conditions of the agreements (*e.g.*, IBM Customer Agreement, Statement of Limited Warranty, International Program License Agreement, etc.) under which they are provided. IBM is not responsible for the performance or interoperability of any non-IBM products discussed herein.

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 in connection with this publication 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.

The provision of the information contained herein is not intended to, and does not, grant any right or license under any IBM patents or copyrights. Inquiries regarding patent or copyright licenses should be made, in writing, to:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

Trademarks

Overview

IBM TotalStorage DS Series

The IBM TotalStorage DS8000 series, a member of the IBM TotalStorage DS family

The IBM TotalStorage DS8000 series is composed of disk storage systems designed to provide high capacity and performance using IBM's POWER5™ technology. The DS8000 series has storage capacity options ranging from 1.1 TB up to 192 TB. Using POWER5 technologythe DS8300 Model 9A2 is able to create storage system logical partitions (LPARs) that can be used for completely separate production, test, or other unique storage environments.

The DS8000 Series is intended to be a flexible and extendable disk storage subsystem and is designed to add and adapt to new technologies as they become available. New management tools are also provided with the DS8000 Series, like the DS Storage Manager and the DS Command-Line Interface (CLI), which allow for the management and configuration of the DS8000 series as well as the IBM TotalStorage DS6000 series. The DS8000 series is designed to support high availability environments and offers advanced remote mirror and copy functions for business continuity.

Copy Services

The DS8000 Series copy services are powerful data backup, remote mirroring and recovery functions that can help protect data from unforeseen events. Copy services run on the IBM TotalStorage DS8000 Series and are designed to support a wide range of open systems, pSeries, iSeries, and zSeries environments. Many design characteristics and advanced functions of the DS8000 Series contribute to the protection of data. The DS8000 Series has a number of advanced copy services functions that are part of the IBM TotalStorage Resiliency Family of technologies. (Comparable functions are supported also on the IBM TotalStorage Enterprise Storage Server (ESS) Models 800 and 750 as well as DS6000 series.) Copy services include the following types of functions:
o IBM TotalStorage FlashCopy®, a point-in-time copy function
o Remote mirror and copy functions include:
o IBM TotalStorage Metro Mirror (previously known as Synchronous PPRC)
o IBM TotalStorage Global Copy (previously known as PPRC Extended Distance)
o IBM TotalStorage Global Mirror (previously known as Asynchronous PPRC)

You can manage Copy Services functions through the DS8000 Series's CLI, as well as the GUI-based interface provided by the IBM TotalStorage DS Storage Manager. The DS Storage Manager allows you to set up and manage data-copy functions from any point from which network access is available.

FlashCopy

The FlashCopy feature is designed to provide the ability to create full volume copies of data. When you set up a FlashCopy operation, a relationship is established between source and target volumes, and a bitmap of the source volume is created. Once this relationship and a bitmap are created, the target volume can be accessed as though all the data had been physically copied. While a

relationship between the source and target volume exists, a background process copies the tracks from the source to the target volume.

Remote Mirror and Copy

The DS8000 Series remote mirror and copy feature is a flexible data mirroring technology that allows replication between volumes on two or more disk storage systems. You can also use this feature for data backup and disaster recovery. DS8000 series systems can participate in remote mirror and copy solutions with the ESS Model 750, ESS Model 800, and DS6000 disk systems.

Metro Mirror supports real-time mirroring of logical volumes between two DS8000 series subsystems that can be located up to 300 km from each other. It is a synchronous copy solution where write operations are completed on both copies (local and remote site) before they are considered to be completed.

Global Copy copies data non-synchronously and over longer distances than is possible with Metro Mirror. When operating in Global Copy mode, the source volume sends a periodic, incremental copy of updated tracks to the target volume instead of a constant stream of updates. This is intended to lessen the impact to application writes for source volumes and cause less demand for bandwidth resources, while allowing a more flexible use of the available bandwidth.

Global Mirror provides a long-distance remote copy feature across two sites using asynchronous technology. Global Mirror is designed to provide the following benefits:

- Support for extended distance between the local and remote sites, with the distance typically limited only by the capabilities of the network and the channel extension technology. This ″extended″ distance support enables you to choose your remote site location based on business needs and enables site separation to add protection from localized disasters.
- A consistent and restartable copy of the data at the remote site, created with minimal impact to applications at the local site.

Oracle Storage Compatibility Program

Through Oracle Storage Compatibility Program (OSCP), IBM has demonstrated to both Oracle and IBM customers that its advanced storage technology can be used to protect data in an Oracle database environment. This document describes the procedures that IBM used to test its implementation of remote mirror technologies under OSCP. All scenarios tested and validated are based on the Oracle document *Guidelines for Remote Mirroring Storage Systems for Oracle Database*, which is available at the following URL:

http://otn.oracle.com/deploy/availability/pdf/oscp_remote_mirror_use.pdf

OSCP Remote Mirror Test Kit

As part of OSCP requirements, IBM used the OSCP Remote Mirror Test Kit to test the IBM DS8000's Global Mirror (asynchronous PPRC) function in an Oracle database environment. This Test Kit was used to create and configure Oracle database instances which included an init file,

control files, table spaces and data files for the scenario in which the testing was performed. Two host servers and two IBM DS8000s were used to configure an Oracle Standby Database environment which consists of one production Oracle server and one Standby Oracle server. Each tested scenario emulates certain Failover or Fallback operation in disaster recovery situation that Oracle database environment may implement. A load generator, which simulated transactional stress loads, was used during every scenario testing. We also used an SQL script provided in the Test Kit to check and verify database consistency before and after every database recovery on production and standby site.

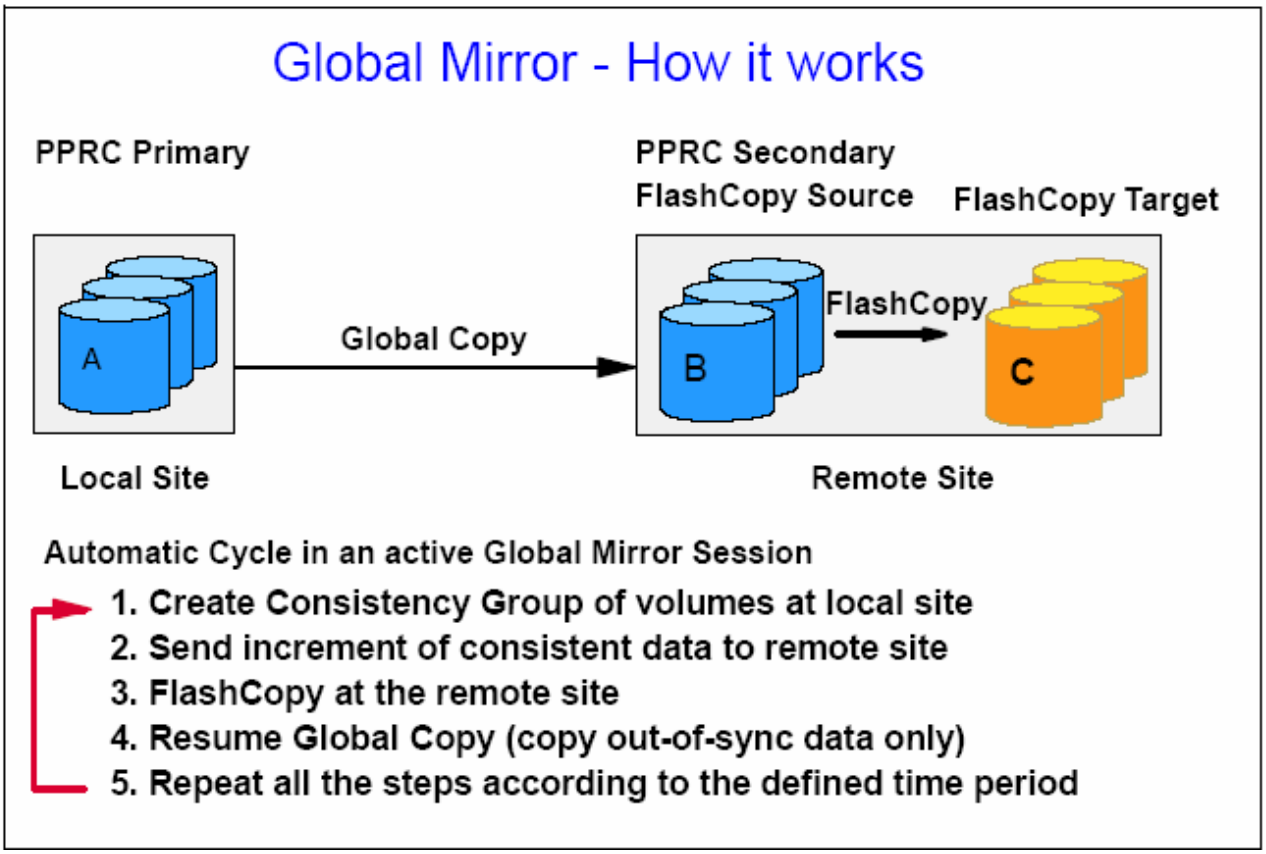
For more information about the OSCP, refer to the following URL:
http://otn.oracle.com/deploy/availability/htdocs/oscpf.html

Overview of Global Mirror (Asynchronous PPRC)

Global Mirror provides a long-distance remote copy feature across two sites using asynchronous technology. With Global Mirror, the data that the host writes to the server enclosure at the local site is asynchronously shadowed to the server enclosure at the remote site. A consistent copy of the data is then automatically maintained on the server enclosure at the remote site.

Global Mirror operations are designed to provide the following benefits:

- Support for virtually unlimited distances between the local and remote sites, with the distance typically limited only by the capabilities of the network and the channel extension technology. This *unlimited* distance enables you to choose your remote site location based on business needs and enables site separation to add protection from localized disasters.
- A consistent and restartable copy of the data at the remote site, created with minimal impact to applications at the local site.
- Data currency where, for many environments, the remote site lags behind the local site typically 3 to 5 seconds, reducing the amount of data exposure in the event of an unplanned outage. The actual lag in data currency that you experience can depend upon a number of factors, including specific workload characteristics and bandwidth between the local and remote sites.
- Dynamic selection of the desired recovery point objective, based upon business requirements and optimization of available bandwidth.
- Session support whereby data consistency at the remote site is internally managed across up to eight storage units that are located across the local and remote sites.
- Efficient synchronization of the local and remote sites with support for failover and failback modes, helping to reduce the time that is required to switch back to the local site after a planned or unplanned outage



How Global Mirror works

The A volumes at the local site are the production volumes; they are used as Global Copy primary volumes. The data from the A volumes is replicated to the B volumes, which are Global Copy secondary volumes. At a certain point in time, a Consistency Group is created using all of the A volumes, even if they are located in different DS boxes. There is typically no application impact because the creation of the Consistency Group is very quick (on the order of milliseconds).

- The copy created with Consistency Group is a power-fail consistent copy, not an application-based consistent copy. When you recover with this copy, you may need recovery operations, such as the **fsck** command in an AIX filesystem.

Once the Consistency Group is created, the application writes can continue updating the A volumes. The increment of the consistent data is sent to the B volumes using the existing Global Copy relationship. Once the data reaches the B volumes, it is copied to the C volumes using the FlashCopy function.

The C volumes should now contain a *consistent* copy of data. Because the B volumes usually contain a *fuzzy* copy of the data from the local site (not when doing the FlashCopy), the C volumes are used to hold the last point-in-time consistent data while the B volumes are being updated by the Global Copy relationship.

- When you implement Global Mirror, you setup the FlashCopy relationship between the B and C volumes with *No Background copy* and *Start Change Recording* options. It means that before the latest data is updated to the B volumes, the last consistent data in the B volume is moved to the C volumes. Therefore, at some time, a part of consistent data is in the B volume, and the other part of consistent data is in the C volume.
- If a disaster occurs during the FlashCopy operation, special procedures are needed to finalize the FlashCopy target.
- In the recovery phase, the consistent copy is created in the B volumes. You need some operations to check and create the consistent copy.
- You need to check the status of the B volumes for the recovery operations. Generally, these check and recovery operations are complicated and difficult with the GUI or CLI in a disaster situation. Therefore, you may want to use some management tools (for example, Global Mirror Utility), or management software (for example, Multiple Device Manager Replication Manager), for Global Mirror to automate this recovery procedure.

The data at the remote site is current to within 3 to 5 seconds of the data last written on the source copy, but this recovery point (RPO) depends on the workload and bandwidth available to the remote site.
Global Mirror automates all of the steps that have to be done manually when using Global Copy.
If you use Global Mirror, you must adhere to the following additional rules: _ You must purchase a FlashCopy function authorization (2244 Model PTC) for the secondary storage unit. If Global Mirror will be used during failback on the secondary storage unit, you must also purchase a FlashCopy function authorization for the primary system.

- Global Mirror and Metro Mirror can support failover and failback operations. A failover operation is the process of temporarily switching production to a backup facility (normally your recovery site) following a planned outage, such as a scheduled maintenance period, or an unplanned outage, such as a disaster. A failback operation is the process of returning production to its original location. These operations use Remote Mirror and Copy functions to help reduce the time that is required to synchronize volumes after the sites are switched during a planned or unplanned outage.

Terminology Overview

In the descriptions of the scenarios, the following terminology is used:

- *Atomic*: To break mirror atomically roughly means to stop propagating writes at a point in time across all mirrored data.

- *Automatic split*: The mode where a system may choose to break the mirror if a write can not succeed on a remote site due to link failure, a crash of the remote system, etc when using synchronous remote mirroring. This mode favors availability rather than reliability.

- *Failover*: The action that transforms a standby database into the production database when there is a failure on the host where the production database is located.

- *Fallback*: The action that constructs a new standby database after the original standby database has been successfully transformed into the production database when a disaster occurs.

- *Master DS*: Using built-in functionality, a Master DS communicates with subordinate DSs to control the creation of Consistency Groups and manage the Global Mirror session. The Master is defined when the Start/Resume Global Mirror Session command is issued to an LSS in the Master DS. The Master DS needs Fibre Channel communication paths to any of the LSSs in each subordinate DS that has volumes as part of the session. This also called *Production Storage.*

- *Primary site*: The main site of the production database. This does not change during failover or fallback.

- *Production database*: The database that handles user requests in a production environment. This may also be called the primary database. The production database may be on either the primary or secondary site. For example, after failover, the standby site may run the primary database.

- *Standby database*: This term refers to the Oracle standby database. This may also change sites. For example, after reverse role, the primary may run standby database.

- *Standby site*: The initial site for the standby database. This may also be called the secondary site. This also does not change during failover.

- *Subordinate DS*: Using built-in functionality, a Subordinate DS communicates with the Master and is controlled by the Master. One of the LSSs of each Subordinate DS needs to have PPRC Fibre Channel communication paths from the Master DS established. It is used to enable the Master to create Consistency Groups using volumes in different DS boxes. This also called *Standby storage.* the local storage system will not acknowledge a write until the local system is told by the remote system that the write has been done persistently on the remote storage system. The write is written persistently on the remote site once it is recorded on disk or NVRAM.

Lab Test Environment

We configured two SUN servers and two IBM DS8000s for our validation testing. The operation system setup, file systems, and Oracle database setup and key database file paths are identical on both host machines.

Table 1. Hosts

Features	Production	Standby
Machine Model	SUN Ultra-SPARC E450	SUN Ultra-SPARC E450
CPU	4 x 296 Hz	4 x 296 Hz
RAM	3GB	2GB
Internal Disk	4 x 9GB	4 x 9GB
I/O adapter	JNI Fibre Channel HBA	JNI Fibre Channel HBA
OS version	Solaris 9	Solaris 9
Oracle version	Oracle 8.1.7 Enterprise Server.	Oracle 8.1.7 Enterprise Server.

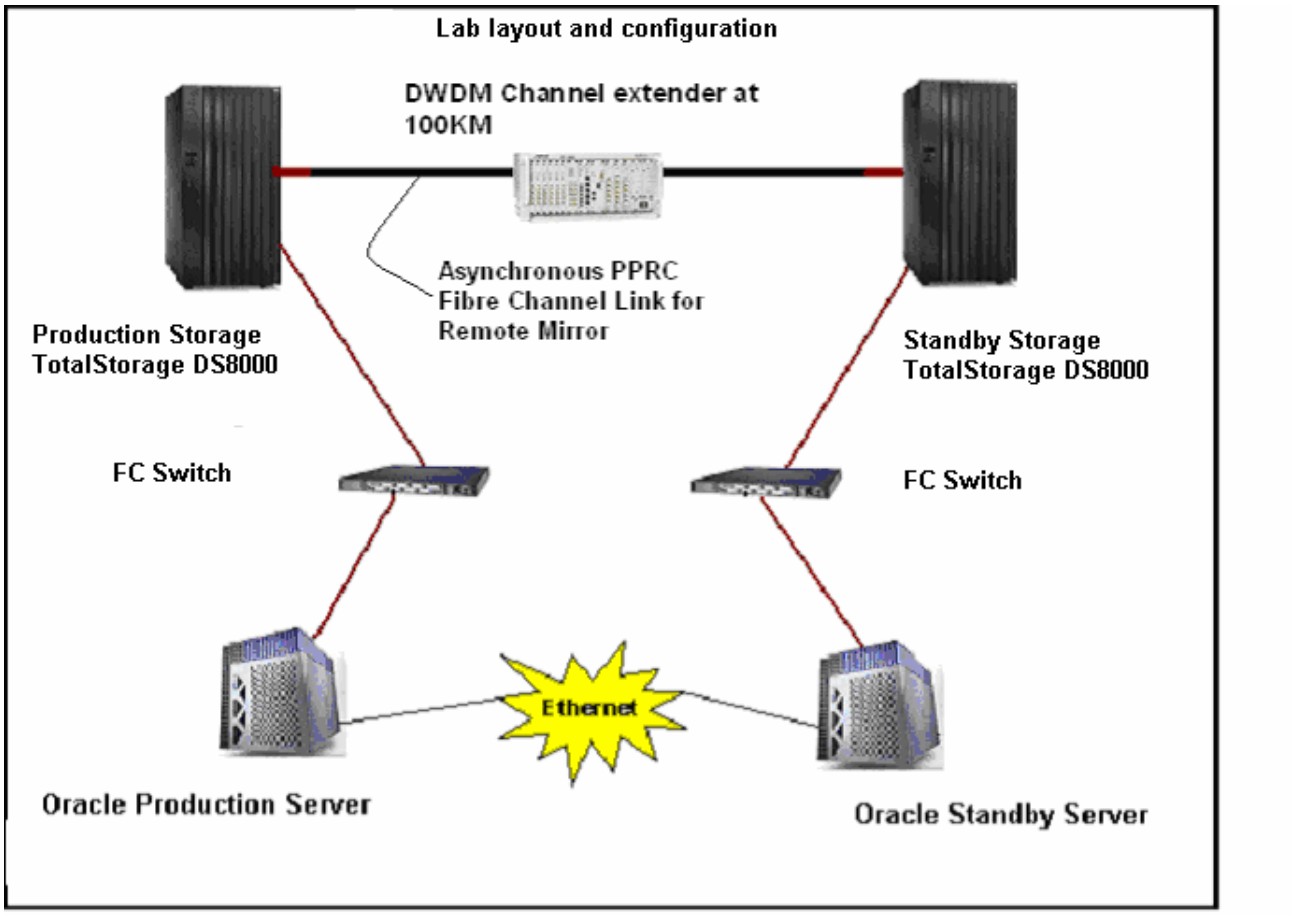
Table 2. Storage Hardware

Features	Primary	Secondary
Storage	IBM DS8000 2107-900	IBM DS8000 2107-900
Fibre Switch	Brocade 2109 F16 16pt FC	Brocade 2109 F16 16pt FC
Network	Ethernet switch	Ethernet switch
DS8000 'name'	oslsq04	oslsq05

Table 3. File System Layout for Both Primary and Secondary Sites

Device Location	Mount Point	Usages
/dev/dsk/c0t1d0s7	/orc	Oracle Home
/dev/dsk/c0t2d0s3	/orcdf	Database files
/dev/dsk/c6t0d0s1	/oscprmm	Redo logs DS8000 oslsq04 - production
/dev/dsk/c5t0d0s3	/oscprmm	Redo logs DS8000 oslsq05 - standby

Lab Configuration



Global Mirror setup steps

First of all identify all PPRC primary and secondary volumes and the corresponding LSSs as well as the FlashCopy target volumes at the remote sites for the Global Mirror session.

Steps below are used to setup our lab test for the OSCP.

1. Establish a direct fibre channel connection between primary storage and secondary storage.
2. Establish PPRC data path from primary storage to secondary storage.
3. Establish PPRC-XD volume Pairs from Primary storage to secondary storage.
4. Using PPRC target volume as a source to establish a FlashCopy relationship to a FlashCopy target. Standby host uses the FlashCopy target volume to bring up database when a failure occurs.
5. Establish PPRC data path from primary DS8000 to secondary DS8000.
6. Establish PPRC-XD volume Pairs from Primary DS8000 to secondary DS8000.
7. Establish FlashCopy relationships for asynchronous PPRC target volume at the remote site. Use the following options: No Background copy, Start Change Recording, Inhibit writes to target and Persistent FlashCopy.
8. Open a session with all LSSs that will participate in the asynchronous PPRC environment.
9. Add all PPRC primary volumes to the session.
10. Start the asynchronous PPRC session using one of the LSSs located in the Master DS8000.

We used DS command line interface (DSCLI) to setup the steps above. The configuration and functional commands were saved in files and reused for the duration of the lab test to ensure consistency of testing. The IBM TotalStorage Management Console may also be used as the focal point for configuration, Copy Services Management, and maintenance for the DS8000.

For a complete DS8000 Global Mirror setup instructions, refer IBM Redbook (SG246452) "*The IBM TotalStorage DS8000 Series: Concepts and Architecture*" or the following web link http://www.redbooks.ibm.com/redbooks/SG246452/wwhelp/wwhimpl/java/html/wwhelp.htm

SUN Ultra-SPARC Host Set-Up

Check the test kit documentation and verify the correct versions of Solaris and Oracle have been correctly installed on the host, log in as root.

One will want to create a location for $ORACLE_HOME\dbs.

1. Select a disk that is not being used to create a new file system that will serve as the location of the ORACLE_HOME\dbs.

2. Now one must create the file system on the host.
   a. Go into the root directory / and create a new directory.
   b. Mount the file system by typing mount /dev/dsk/c0t1d0s7 /newDirName
   c. Type chown oracle newDirName and chgrp oinstall newDirName to make oracle the owner of the newDirName.

3. Now you'll want to make sure the DS Logical Units (LUNS) are set-up correctly and create locations to store the dbs and redo directories.
   a. IBM DS LUNs will appear in the host if you execute "format" command.
   b. Partition the LUNS for the data files and redo log files as mentioned in the above point 2 and 3. They should be named to show that data files will reside in one directory and redo files reside in the other. Similar to Step 3, make sure to change ownership for these directories to oracle and group oinstall.

4. If you type #df, the oracle related directories/disk info will appear:

   On the primary host

/orc	(/dev/dsk/c0t1d0s7):	13617796 blocks	1031877 files
/orcdf	(/dev/dsk/c0t2d0s3):	16518952 blocks	1064439 files
/oscprmm	(/dev/dsk/ c7t1d0s1):	1710910 blocks	473074 files

   On the secondary host

/orc	(/dev/dsk/c0t1d0s7):	13617796 blocks	1031877 files
/orcdf	(/dev/dsk/c0t2d0s3):	16518952 blocks	1064439 files
/oscprmm	(/dev/dsk/c5t0d0s3):	1710910 blocks	473074 files

   The location of our dbs and redo directories is /orcdf and /oscprmm, respectively.

5. Now one must install the test kit so one should be signed in as oracle instead of root.
6. Using the steps described in the Oracle Remote Mirror Test Kit (SDK) Readme file, the SDK is installed exactly as described.

Scenario Guide Overview

Scenarios represent the key Oracle database disaster recovery situation, supported by IBM DS8000 Global Mirror technology, were tested and validated using the Remote Mirror Test Kit. An Oracle database administrator may be able to develop a disaster recovery strategy, based on one of the scenarios or a combination of several scenarios that could best fulfill the disaster recovery requirement of their database environment.

Key tests validated in our OSCP remote mirror validation

- Simple Primary Site to Secondary Site Failover /Atomic Split of Mirror
- Reverse Role Fallback via Database Copy
- Reverse Role Fallback via Restoring Backup
- Direct Fallback via Database Copy
- Direct Fallback via Restoring Backup
- Simple Primary Site to Secondary Site Failover for Managed Recovery with Standby database
- Reverse Role Fallback via Restoring Backup for Managed Recovery with Standby database
- Concurrent Transaction testing
- Write Ordering testing
- Simple Primary Site to Secondary Site Failover for Whole database Mirror
- Simple Secondary Site to Primary Fallback for Whole database Mirror
- Write Ordering for Whole Database Mirror

Each test is described first below with a brief explain of the disaster recovery situation. We then use a table to list the sequence of events and commands for performing the Global Mirror on IBM DS8000 and SQL commands for activating Standby database, performing database Failover or Fallback on the Oracle database environment.

Simple Primary to Secondary Site Failover & Atomic Split of Mirror

Follow the procedures listed below to active the standby database when a disaster, network related or power related, disables the primary site from serving as the production database.

In this scenario, On the primary site, this LUN is mounted as /oscprmm There is a Remote Mirror and Copy relationship which maps the redo directory on the primary site to the redo directory on the secondary site, which is also located at /oscprmm.

The procedure and commands for scenario above is summarized in the table below:

Event Sequence	Primary Site Actions	Secondary Site Actions	IBM DS8000 action
Activate Remote Mirror and Copy Relationship			Start Global Mirror
Setup Standby Database		Issue 'alter database recover automatic standby database;' for recovery mode	
Disaster	Out of commission		Stop Global Mirror Establish FlashCopy pair
Check file system of the target LUN		fsck –y /dev/rdsk/c5t0d0s2	
Mount the file system where the redo logs reside		Mount /dev/dsk/c5t0d0s2 /oscprmm	
Activate standby database (convert into production database)		a. shutdown abort b. startup nomount c. create control file using "create control file" script with noresetlogs option. This script was generated in primary site and saved in the standby site. d. alter database recover automatic database; e. alter database open;	

Reverse Role Fallback via Database Copy

Follow the procedures listed below to use the secondary site as the production database and uses the primary site as the standby database. This is achieved by having the production database fallback to the primary site by copying the datafiles back from the secondary to primary site.

In this scenario, we assume that the mounted LUN on the secondary site, contains the directory storing the redo files, corresponds to the directory /oscprmm. There is a Remote Mirror and Copy relationship which maps the redo directory on the primary site to the redo directory on the secondary site, which is also located at /oscprmm.

The procedure and commands for scenario above is summarized in the table below:

Event Sequence	Primary Site Actions	Secondary Site Actions
Failover to secondary site after disaster	Out of commission	Standby database transformed to production database
Remove error or rebuild primary site	Primary recovered	
Start process to reverse role fallback		a. perform full backup(hot or cold) of current database b. alter database, create standby controlfile; c. alter system archive log current; d. alter database, backup controlfile to trace noresetlogs;
Copy database files and related control files from secondary site to primary site	a. copy datafiles that get backup on secondary site to primary site b. copy archive logs since backup c. copy standby control file d. copy “create control file” script	
Recover database on the primary site	a. startup nomount b. alter database mount standby database; c. alter database recover automatic standby database; d. alter database recover cancel; e. setup sqlnet	
Start sending archive logs from secondary to primary		Start second destination for archive logs on the primary site
Put standby database on primary site in manual recover manual	After database, recover standby database	

Reverse Role Fallback via Restoring Backup

Follow the procedures listed below to use the secondary site as the production database and uses the primary site as the standby database. This is achieved by having the production database fallback to the primary site by restoring the database using backed up datafiles.

In this scenario, we assume that the mounted LUN on the secondary site, contains the directory storing the redo files, corresponds to the directory /oscprmm. There is a Remote Mirror and Copy relationship which maps the redo directory on the primary site to the redo directory on the secondary site, which is also located at /oscprmm.

The procedure and commands for scenario above is summarized in the table below:

Event Sequence	Primary Site Actions	Secondary Site Actions
Failover to secondary site after disaster	Out of commission	Standby database transformed to production database
Remove error or rebuild primary site	Primary recovered	
Start process to reverse-role fallback via restoring backup	Restore backup taken before failover (disaster)	
Prepare files on secondary site		a. alter database, create standby controlfile; b. alter system archive log current; c. alter database, backup controlfile to trace noresetlogs;
Copy control and archive files from secondary site to primary site	a. copy archive logs from secondary site b. copy standby control file c. copy “create control file” script	
Recover database on the primary site	a. startup nomount b. alter database mount standby database; c. alter database recover automatic standby database; d. alter database recover cancel; e. setup sqlnet	
Start sending archive logs from secondary to primary		Start second destination for archive logs on the primary site
Put standby database on primary site in manual recover manual	After database, recover standby database	

Direct FallBack via Database Copy

This approach requires shutting down the production database on the secondary site before performing the fallback to the primary site. It also requires you to setup the standby database on the secondary site again after the successful fallback to primary site

In this scenario, we assume that the mounted LUN on the secondary site, contains the directory storing the redo files, corresponds to the directory /oscprmm. There is a Remote Mirror and Copy relationship which maps the redo directory on the primary site to the redo directory on the secondary site, which is also located at /oscprmm.

The procedure and commands for scenario above is summarized in the table below:

Event Sequence	Primary Site Actions	Secondary Site Actions	IBM DS8000 action
Failover to secondary site after disaster	Out of commission	Standby database transformed to production database	
Remove error or rebuild primary site	Primary recovered		
Start process to fallback direct to primary site via database copy from the secondary to primary site	Remove any leftover database files such as datafiles, old archive logs, redo logs or control files		
Prepare files on secondary site		a. shutdown database b. perform full backup	
Copy datafiles, redo logs, control files and archive logs from secondary site	a. copy datafiles b. copy redo logs c. copy control files d. copy archive logs e. copy init file		
Recover database on the primary site and bring up the production database	a. startup nomount b. alter database mount; c. alter database recover automatic database; d. alter database open;		
Start sending archive logs from secondary to primary		Start second destination for archive logs on primary site	
Activate Remote Mirror and Copy Relationship for mirroring redo logs			Start Global Mirror
Build a completely new standby database on the secondary site		Follow standard procedures to setup standby database	

Direct Fallback via Restoring Backup

This approach requires shutting down the production database on the secondary site before performing the fallback. It also requires you to setup the standby database on the secondary site again after the successful fallback to primary site

In this scenario, we assume that the mounted LUN on the secondary site, which contains the directory storing the redo files, corresponds to the directory /oscprmm. There is a Remote Mirror and Copy relationship which maps the redo directory on the primary site to the redo directory on the secondary site, which is also located at /oscprmm.

The procedure and commands for scenario above is summarized in the table below:

Event Sequence	Primary Site Actions	Secondary Site Actions	IBM DS8000 action
Failover to secondary site after disaster	Out of commission	Standby database transformed to production database	
Remove error or rebuild primary site	Primary recovered		
Start process to fallback direct to primary site via restoring backup	Remove any leftover database files such as datafiles, old archive logs, redo logs or control files		
Restore backed up data and control files	a. restore datafiles b. restore control file created by alter database backup controlfile c. copy archive logs from secondary site	a. shutdown database b. perform full backup	
Recover database on the primary site	a. startup nomount b. alter database mount; c. alter database recover automatic database using backup controlfile d. alter database recover cancel; e. shutdown		
Shutdown production database on secondary		shutdown immediate;	
Copy redo logs or any new archive logs	Copy redo logs and any new archive logs from secondary		
Recover database and bring up production database on primary	a. startup nomount b. alter database mount; c. alter database recover automatic database; d. alter database open;		
Activate Remote Mirror and Copy relationship for mirroring redo logs			Start Global Mirror
Build a completely new standby database on the secondary site		Follow standard procedures to setup standby database	

Simple Site to Site Failover for Managed Recovery

In a Managed Standby Environment, the primary database automatically sends the archive logs to a standby site via the Oracle Net8 connection. When the standby database is configured to run in Managed Recovery mode, Oracle automatically applies the archive logs to the database in the standby site at the time they are received from the primary site. The advantage of the Managed Recovery mode is that the recovery, which activating the standby database, takes much less time since the standby database is being kept current. The drawback is that user errors or data corruptions may be inadvertently propagated to the standby database.

In this scenario, we assume that the mounted source LUNs on the primary site, which contain the directories storing redo files are mounted at /oscprmm. There is a Remote Mirror and Copy relationship which maps the redo directory on the primary site to the redo directory on the secondary site, which is also located at /oscprmm.

The procedure for simple site to site failover scenario is summarized in the table below:

Event Sequence	Primary Site Actions	Secondary Site Actions	IBM DS8000 action
Activate Remote Mirror and Copy Relationship			Start Global Mirror
Setup Standby Database		Issue 'alter database recover managed standby database;' for managed recovery mode	
Disaster	Out of commission		Stop Global Mirror Establish FlashCopy pair
Check file system of the target LUN		fsck –y /dev/rdsk/c5t0d0s2	
Mount the file system where the redo logs reside		Mount /dev/dsk/oscprmm	
Activate standby database (convert into production database)		a. shutdown abort b. startup nomount c. create control file using "create control file" script with noresetlogs option. This script was generated in primary site and saved in the standby site. d. alter database recover automatic database; e. alter database open;	

Reverse Role Fallback via Backup Restore for Managed Recovery with Standby Database

In a managed standby environment, the primary database automatically sends the archive logs to a standby site via the Oracle Net8 connection. When the standby database is configured to run in Managed Recovery mode, Oracle automatically applies the archive logs to the database in the standby site at the time they are received from the primary site. The options and procedures of fallback using Managed Recovery mode are the same as those for manual recovery mode.

In this scenario, we assume that the mounted source LUNs on the primary site, which contain the directories storing redo files are mounted at /oscprmm. There is a Remote Mirror and Copy relationship which maps the redo directory on the primary site to the redo directory on the secondary site, which is also located at /oscprmm.

The procedure and commands for scenario above is summarized in the table below:

Event Sequence	Primary Site Actions	Secondary Site Actions
Failover to secondary site after disaster	Out of commission	Standby database transformed to production database
Remove error or rebuild primary site	Primary recovered	
Start process to reverse role fallback via restoring the backup	Restore backup taken before the failover or disaster	
Prepare files on secondary site		a. alter database, create standby controlfile; b. alter system archive log current; c. alter database, backup controlfile to trace noresetlogs;
Copy database files and related control files from secondary site to primary site	a. copy archive logs from secondary site b. copy standby control file c. copy “create control file” script	
Recover database on the primary site	a. startup nomount b. alter database mount standby database; c. alter database recover automatic standby database; d. alter database recover cancel; e. setup sqlnet	
Start sending archive logs from secondary to primary		Start second destination for archive logs t the primary site
Put standby database on primary site in managed recover manual	After database, recover managed standby database	

Concurrent Transactions test

By performing the following actions, one can verify that when placing concurrent transactions on the databases, the DS's Remote Mirror and Copy (Global Mirror) feature will confirm that the standby database is up to date with the production database up until the time of a disaster.

In this scenario, we assume that the mounted source LUNs on the primary site, which contain the directories storing redo files are mounted at /oscprmm. There is a Remote Mirror and Copy relationship which maps the redo directory on the primary site to the redo directory on the secondary site, which is also located at /oscprmm.

The procedure and commands for scenario above is summarized in the table below:

Event Sequence	Primary Site Actions	Secondary Site Actions	IBM DS8000 action
Activate Remote Mirror and Copy Relationship			Start Global Mirror
Simulate balance transfer between tables		Issue 'alter database recover managed standby database;'	
Disaster	Out of commission		
Check file system of the target LUN		fsck –y /dev/rdsk/c5t0d0s2	
Mount the file system where the redo logs reside		Mount /dev/dsk/oscprmm	
Activate standby database (convert into production database)		a. shutdown abort b. startup nomount c. create control file using "create control file" script with noresetlogs option. This script was generated in primary site saved in the standby site. d. alter database recover automatic database; e. alter database open;	
Data verification		At this point, data should be up to date	

Write Ordering test

By performing the following actions, one can verify that when placing data into the production database by spawning multiple processes, the DS's Remote Mirror and Copy (Global Mirror) feature will confirm that the write ordering is preserved in standby database.

In this scenario, we assume that the mounted source LUNs on the primary site, which contain the directories storing redo files are mounted at /oscprmm. There is a Remote Mirror and Copy relationship which maps the redo directory on the primary site to the redo directory on the secondary site, which is also located at /oscprmm.

The procedure and commands for scenario above is summarized in the table below:

Event Sequence	Primary Site Actions	Secondary Site Actions	IBM DS8000 action
Activate Remote Mirror and Copy Relationship			Start Global Mirror
Generate database write activity	a. drop and recreate test table b. load data into production database		
Disaster	Out of commission		Stop Global Mirror Establish FlashCopy pair
Check file system of the target LUN		fsck –y /dev/rdsk/c5t0d0s2	
Mount the file system where the redo logs reside		Mount /dev/dsk/oscprmm	
Activate standby database (convert into production database)		a. shutdown abort b. startup nomount c. create control file using "create control file" script with noresetlogs option. This script was generated in primary site saved in the standby site. d. alter database recover automatic database; e. alter database open;	
Data verification		At this point, data should be up to date	

Simple Primary Site to Secondary Site Failover in Whole database Mirror

Follow the procedures listed below to allow users to mirror the whole database remotely. This allows for a simpler procedure to follow as well as faster recovery time in case of a disaster.

In this scenario, we assume that the mounted source LUNs on the primary site, which contain the directories storing the datafiles and redo files, is mounted at /oscprmm. There is a Remote Mirror and Copy relationship which maps the redo and datafiles directory on the primary site to the redo and datafile directories on the secondary site, which is also located at /oscprmm.

The procedure and commands for scenario above is summarized in the table below:

Event Sequence	Primary Site Actions	Secondary Site Actions	IBM DS8000 action
Activate Remote Mirror and Copy Relationship			Start Global Mirror
Setup Standby Database	All datafiles, control file, redo logs and archive logs ./oscprmm, which are being mirrored	All datafiles, control file, redo logs and archive logs are in /oscprmm. Database is not running while waiting.	
Disaster	Out of commission		Stop Global Mirror Establish FlashCopy pair
Check file system of the target LUN		fsck –y /dev/rdsk/c5t0d0s2	
Mount the file system where the redo logs reside		Mount /dev/dsk/oscprmm	
Activate standby database (convert into production database)		a. startup database b. verify data in the database that was just brought up	

Simple Secondary Site to Primary Fallback in Whole database Mirror

Role copy information from the secondary site back to the primary site. You have the option of shutting down the secondary site, if downtime can be tolerated, and then copying the entire database. Otherwise to avoid shutting down the database, one can perform a hot full backup and move all relevant files to the primary site.

In this scenario, we assume that the mounted source LUNs on the primary site, which contain the directories storing the datafiles and redo files, is mounted at /oscprmm. There is a Remote Mirror and Copy relationship which maps the redo and datafiles directory on the primary site to the redo and datafile directories on the secondary site, which is also located at /oscprmm.

The procedure and commands for scenario above is summarized in the table below:

Event Sequence	Primary Site Actions	Secondary Site Actions	IBM DS8000 action
Failover to secondary site after disaster	Out of commission	Standby database transformed to production database	
Remove error or rebuild primary site	Primary recovered		
Start process to perform reverse-role fallback		Shutdown database	
Start role reversal		Copying the datafiles, redo logs and control file to the primary site	
Recover database on the primary site and make it the standby database	Startup the database in standby mode	Startup the database	

Write Ordering test for Whole Database Mirror

By performing the following actions, one can verify that when placing data into the production database by spawning multiple processes, the write ordering in preserved in the standby database.

In this scenario, we assume that the mounted source LUNs on the primary site, which contain the directories storing the data files and redo files are mounted at /oscprmm.

The procedure for testing write ordering in whole database mirror is summarized in the table below:

Event Sequence	Primary Site Actions	Secondary Site Actions	IBM DS8000 action
Activate Remote Mirror and Copy Relationship			Start Global Mirror
Setup Standby Database	All datafiles, control file, redo logs and archive logs ./oscprmm, which is being mirrored	All datafiles, control file, redo logs and archive logs /oscprmm. Database is not running while waiting.	
Generate database write activity	a. drop and recreate test table b. load data into the production database		
Disaster	Out of commission		Stop Global Mirror Establish FlashCopy pair
Check file system of the target LUN		fsck –y /dev/rdsk/c5t0d0s2	
Mount the file system where the redo logs reside		Mount /dev/dsk/oscprmm	
Activate standby database by converting into production database		Startup database	
Data verification		At this point, data should be up to date	

Summary

The IBM TotalStorage DS family is designed to provide the means to support continuous data availability with its FlashCopy and Remote Mirror and Copy functions. By successfully testing these Advanced Copy Services using the Oracle Remote Mirror Test Kit, IBM demonstrates that its storage technology is compatible with Oracle's database technology. Oracle database administrators could apply the scenarios described in this usages guide and construct a disaster recovery strategy that helps fulfills the requirements of their database environment.

All the distinct steps described in each scenario can also be threaded together by shell scripts or pertinent commands or tools to build a disaster recovery solution with no or little human intervention for an Oracle database.

References

Warrick, Alluis, Bauer, Blaschek, Fourie, Garay, Knobloch, Laing, O'Sullivan, Rothenwaldt, Sano, Tang, Vandewerdt, Warmuth, Wolf, The IBM TotalStorage DS8000 Series: Concepts and Architecture Copyright IBM Corp. 2005
http://www.redbooks.ibm.com/redbooks/SG246452/wwhelp/wwhimpl/java/html/wwhelp.htm