

SANRAD Application Note: Designing your GDR Solution

APP-021-01

Important Notice

The configurations described or tested in this application note are not the only available solution. This application note is not intended nor may it be construed as an endorsement of any product(s) tested.

This application note provides no warranty of any kind. SANRAD limited warranties for SANRAD products are stated in separate documentation accompanying each product. No damages or remedy of any kind shall be recoverable by any party for any claim or loss in any way arising or alleged to arise from or as a result of this application note, whether considered alone or in addition to any other claim.

Copyright SANRAD 2007

All rights reserved. The copyright and all intellectual property rights in this article belong to SANRAD. It is strictly forbidden to copy, duplicate or otherwise use this article or any part thereof in any way shape or form without the prior written consent of SANRAD.

Table of Contents

Introduction	1
GDR Architecture	1
Primary and Secondary Storage	2
Allocating Disk Space for the Journal	2
Allocating Disk Space for Snapshots (for asynchronous replication)	2
Allocating Disk Space for Replicated Volumes.....	3
Network Architecture	3
Replication.....	3
Deciding which Type of Replication to use, Synchronous or Asynchronous?.....	3
How often to Replicate when using Asynchronous Replication	3
Initial Data Synchronization	4
GDR Deployment.....	5
Creating a new GDR setup using two new V-Switches	6
Adding GDR to existing V-Switches already being used in production.....	7
Pre-production Tests.....	7
Verifying Replication.....	7
Synchronous Replication	7
Asynchronous Replication	7
Testing Planned Failover.....	9
Testing Disaster Failover (optional).....	9
Final Implementation.....	9
Monitoring	10
Asynchronous Replication.....	10
Synchronous Replication Status	11
Determining PIT Size	12
Appendix A – Sample Design Information Template.....	14

Figures

Figure 1. GDR Deployment Flowchart	5
Figure 2. Pending Pits	8
Figure 10. Last Successful Replication	8
Figure 11. Monitoring Asynchronous Replication.....	10
Figure 12. Monitoring Synchronous Replication.....	11
Figure 13. Monitoring Pits for GDR Pair.....	12
Figure 14. Monitoring Pits for Consistency Group.....	13
Figure 15. Monitoring Size of Pits being Replicated.....	13

Tables

Table 1. Sample Design Template.....	14
Table 2. GDR Design Template.....	15

Introduction

Designing and implementing an efficient, reliable and well-tuned Global Data Replication & Recovery (GDR) System requires careful planning. Different aspects of the system design must be considered. Before implementation, it is crucial to verify that the design will meet your Service Level Agreement (SLA) requirements. Thorough testing before implementation and consistent monitoring afterwards will help prevent major problems by detecting bottlenecks and other changes in condition. This guide will help you understand the critical design considerations for designing and implementing a successful GDR solution including testing and monitoring.

Note:

This paper assumes that you are familiar with SANRAD's V-Switch StoragePro Graphical User Interface. SANRAD recommends reading the V-Switch StoragePro Graphical User Interface User Manual as well as V-Switch Global Data Replication & Recovery User Manual.

GDR Architecture

This section describes the important considerations affecting your GDR solution's design.

Notes:

1. *Once a GDR pair is defined you can't make any changes to your GDR parameters. It's very important to define your requirements carefully before implementation (e.g. replication interval & journal size).*
2. *SANRAD recommends writing down your GDR solution's information in a simple table to facilitate configuration. A sample table is provided in [Appendix A – Sample Design Information Template](#), page 14.*

Primary and Secondary Storage

GDR requires you to allocate storage space for the journal, snapshots and replicated volumes.

Allocating Disk Space for the Journal

Note:

Allocating disk space for the journal is only relevant when using Asynchronous replication. The allocated Journal space must be defined for both V-Switches (one in each site). Once replication has started, Journal size cannot be reduced, so it's very important to determine the size for the Journal in advance.

- SANRAD recommends defining the journal on a separate LUN (RAID set) than the one used for the production volume.
- When deciding on how much space to allocate for the Journal, you must consider the following factors:
 1. The anticipated average amount of IO changes (e.g. MB/hour or GB/day) which will be replicated.
 2. The space needed to accumulate IO Changes that can not be replicated due to network connectivity issues with the remote site or due to a planned failover mode.
 3. The space to accumulate IO changes before the remote site is online (refer to *Determining PIT Size, page 12*).

For example:

The following example is a guideline for estimating how much space to allocate for the journal based on estimated IO changes and time before system is online.

We have a system with:

- | | |
|---|--------------------|
| • <i>Estimated average amount of IO changes:</i> | <i>250 MB/hour</i> |
| • <i>We want to replicate the data every 60 minutes:</i> | <i>1 time/hour</i> |
| • <i>We anticipate network downtime of 8 hours:</i> | <i>8 hours</i> |
| • <i>We anticipate 2 days before the remote site is online:</i> | <i>48 hours</i> |

$$\{IO\ changes/hour \times downtime\} + \{IO\ changes/hour \times time\ until\ remote\ is\ online\} = Journal\ size\ in\ GB$$

$$\{250\ MB \times 8\} + \{250\ MB \times 48\} = \mathbf{14\ GB}$$

Allocating Disk Space for Snapshots (for asynchronous replication)

- Disk space for the snapshots must be allocated. The disk space can be created on a subdisk on the same physical disk used for the production volume or on a separate physical disk.
- SANRAD recommends that the size allocated for the snapshots is 20% of the production volume size.

For example:

If the size of the production volume to be replicated is 100GB then snapshot size should be 20GB.

- | | |
|--|---------------|
| • <i>Size of the production volume to be replicated:</i> | <i>100 MB</i> |
| • <i>Snapshot should be 20% of production volume:</i> | <i>20%</i> |

$$100\ GB \times 0.20 = \mathbf{20\ GB}$$

Allocating Disk Space for Replicated Volumes

- Disk space allocated on both the local and remote sites must be exactly the same size.
- If the replicated volumes will be mirrored or striped, allocate the necessary space accordingly.

For example:

If you want to replicate a volume that is 100 GB, you must define a remote volume of exactly 100GB.

Network Architecture

- SANRAD recommends having a separate subnet for the management IP ports and a "stretch" subnet for the Portal IPs. If "stretch" subnet is not possible, Gateway Routing IP addresses must be defined for both the local and remote sites on the V-Switches.
- At least one IP address for the iSCSI portals must be defined for each V-Switch.
- Verify IP connectivity to the remote site from within the local V-Switch. Make sure to configure the Firewall (and NAT if needed) accordingly.
- When using MPIO, you must define at least two IP addresses (on different subnets) for the iSCSI portals on the local site. For the remote site, this is optional.
- Make sure StoragePro GUI has access to both V-Switch from both sites.

Replication

How often you replicate your data and which type of replication you select depends on your system design.

Deciding which Type of Replication to use, Synchronous or Asynchronous?

Data can be replicated asynchronously or synchronously. The type of replication depends mainly on the WAN bandwidth capacity and latency between the local and remote site.

- Synchronous replication is recommended with WAN bandwidths greater than 1Gb.
- Asynchronous replication is recommended with WAN bandwidths at least 1.5Mb.

Other important factors for consideration are:

- Is the WAN line is being used for other applications and if so how much of the bandwidth capacity is being used for these applications?
- If the distance between the local site and the remote site is more then 5 miles it may require using asynchronous and not synchronous replication even when the line between the two sites is 1Gb (or higher).

How often to Replicate when using Asynchronous Replication

With asynchronous replication the main issue is how often to replicate the changes. Factors influencing the decision on how often to replicate are:

- The bandwidth between the local and remote sites
- The expected amount of IO changes (MB/hour, GB/day)
- The time it will take to replicate the changes
SANRAD GDR Replication speed is:
For T1 (1.5Mb) line: approximately 140KB/s
For T3 (45Mb) line: approximately 320KB/s

Note:

Under normal usage, the configuration should support the replication of one PIT completely transferred before starting to replicate the next PIT.

For example:

1. *How often should you replicate when the expected IO changes are 250 MB per hour using a T3 WAN (45 Mb/s)?*
 - $250 \text{ MB} / 6 \text{ B/s} (=45\text{Mb/s}) = 42 \text{ seconds}$
Since it should take less than a minute to replicate 250 MB with a T3 line, you can replicate every 1 minute. Actual replication speed may be lower than the MAX potential.
2. *How often should you replicate when the expected IO changes are 250 MB per hour using a T1 WAN (1.5 Mbs)?*
 - $250 \text{ MB} / 190 \text{ KB/s} (=1.5\text{Mb/s}) = 1347 \text{ seconds} = 23 \text{ minutes}$
Since it will take 23 minutes to replicate 250 MB with a T1 line, it will be safe to replicate every 25 minutes. Actual replication speed may be lower than the MAX potential.

Initial Data Synchronization

- When the volume in the local site already has data on it, the volume on the remote site needs to be fully synchronized with that of the local site before starting replication.

Note:

When creating your GDR pair you have the option of running an initial synchronization. This tool is supplied in order to replicate a primary volume which already contains data. If you are building a new GDR pair with new storage, then there is no need to run an initial synchronization. The GDR is built and considered "in synch" from the start.

The time it takes to complete an initial synchronization depends on the speed of the link between your sites and the size of the volumes in the GDR pair. It does not depend on the amount of actual data on the volume, as the process is block-based, and requires the synchronization of all blocks on the volume regardless of which blocks actually contain data.

In the case where you create large volumes, but only a relatively small amount of data at the time of initial synch, you might consider migrating the data off the primary volume before creating the GDR pair. The pair could then be created with no initial synchronization, (which might otherwise take a long time to complete.) Once created, copying the data back to the primary volume will only require the synchronization of that data via the regular replication process, and not of the entire volume.

- Since initial data synchronization over small bandwidth lines may take a while (hours or even days), it is recommended to first connect the remote site hardware (remote V-Switch and disk array) to the same LAN in the local site and complete the initial data synchronization, start the replication and then move the remote site hardware to the remote site physical location after initial sync has completed. While the remote site hardware is being physically moved, the local journal will accumulate the IO changes until the remote site is online at which point, all these changes will be replicated to the remote site. (Refer to [Allocating Disk Space for the Journal](#)).
- If you are planning to migrate data from another volume to the replicated volume in the local site, SANRAD recommends migrating the data before creating the GDR pair. This will avoid problems with the Journal volume not being adequately large.

GDR Deployment

Once you have designed the GDR system and collected all of the required information, it's time to configure your GDR setup according to your design criteria.

There are two possible setup scenarios:

Creating a new GDR setup using two new V-Switches

Adding GDR to existing V-Switches already being used in production

The Flowchart in *Figure 1* summarizes the recommended steps for configuring GDR.

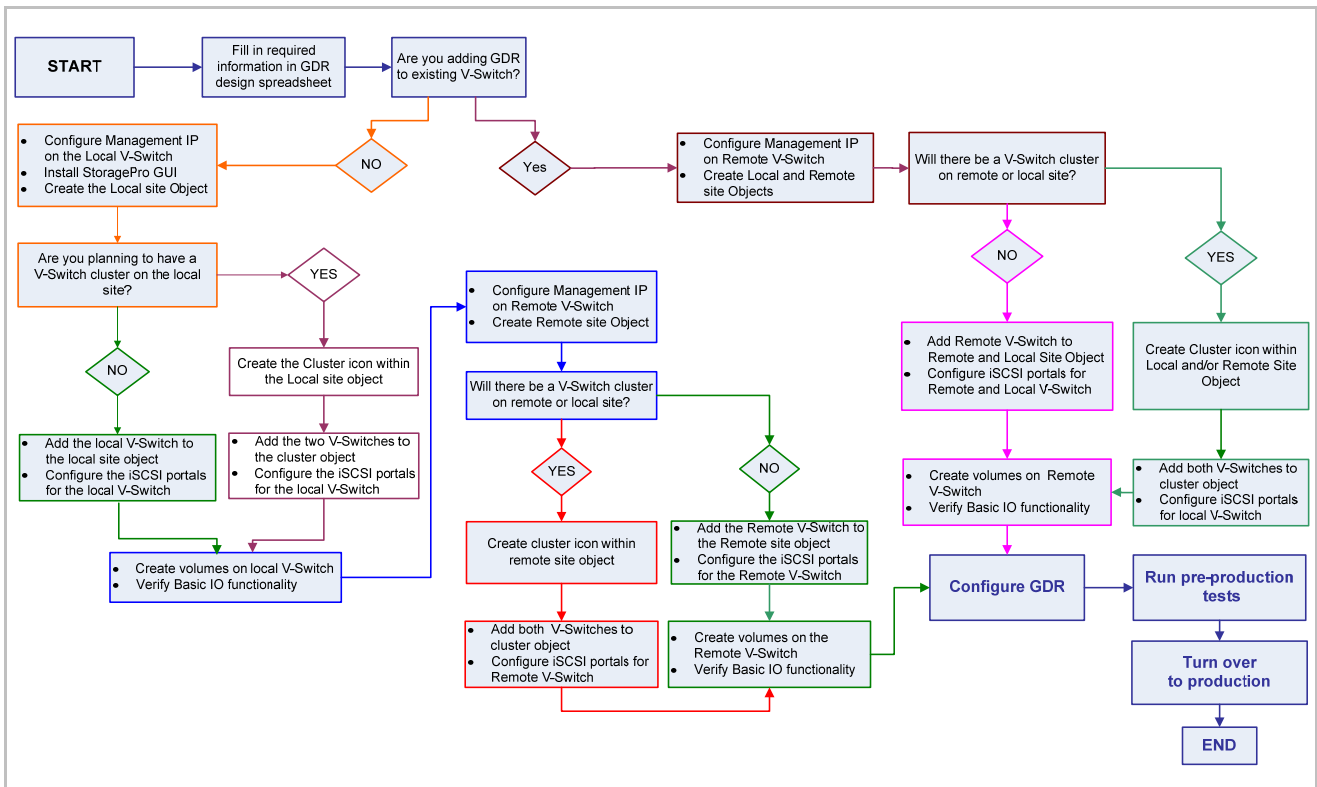


Figure 1. GDR Deployment Flowchart

Creating a new GDR setup using two new V-Switches

Note:

SANRAD recommends temporarily connecting both the local and remote hardware (V-Switch and storage) to the local LAN (local site) while performing the following steps and tests.

1. Configure the management IP for the local site V-Switch.
2. Install SANRAD StoragePro GUI and start it.
3. Using StoragePro GUI create the local site object.
4. Add V-switch to site or, if using a cluster, create the cluster object in the site and add both V-Switches to cluster.
5. Configure the iSCSI portals for the local V-Switch.
6. If using different subnets (one for the local site and a different one for the local site), configure the routes.
7. Make sure the firewall is configured to allow iSCSI traffic
8. From within the StoragePro GUI click the local V-Switch icon and then click the storage icon.
9. Verify that the local V-Switch can see the LUNs exposed by the disk array properly.
10. Configure the volume(s) based on your design (for example: simple, mirror or stripe).
11. Connect from a host via the iSCSI initiator to the volumes on the local V-Switch.
12. Verify that all basic volume operations are working properly from the host's perspective (write, read, copy, delete).

Note:

SANRAD recommends testing the application which will be used in production with the iSCSI volumes exposed by the local V-Switch.

13. If the DR design includes a V-Switch cluster on the local site, test the cluster functionality of failover and fallback.
14. Once these tests are completed, close the applications and disconnect the iSCSI initiator from the iSCSI volumes.
15. Configure the management IP for the remote V-Switch.
16. Using StoragePro GUI create the remote site object.
17. Add V-switch to site or, if using a cluster, create the cluster object in the site and add both V-Switches to cluster.
18. Configure the iSCSI portals for the remote V-Switch.
19. Click the storage icon and verify that the remote V-Switch can see the LUNs exposed by the disk array properly.
20. You are now ready to configure your GDR solution. From StoragePro, launch the GDR pair wizard by selecting **GDR** followed by **Create Pairs...** The wizard guides you through the implementation of your GDR solution in a straightforward manner. Before launching the wizard, make sure that you have filled in the information in the sample form provided in [Table 2. GDR Design Template](#).
21. Proceed with the steps described in [Pre-production Tests](#).

Adding GDR to existing V-Switches already being used in production

Note:

SANRAD recommends temporarily connecting both the local and remote hardware (V-Switch and storage) to the local LAN (local site) while performing the following steps and tests.

You should plan a downtime in which the application using the iSCSI volumes exposed by the current V-Switch will be offline.

1. Configure the management IP for the remote V-Switch.
2. Install SANRAD StoragePro GUI and start it.
3. Using StoragePro GUI create the remote site object.
4. Add V-switch to site or, if using a cluster, create the cluster object in the site and add both V-Switches to cluster.
5. Configure the iSCSI portals for the remote V-Switch.
6. From within the StoragePro GUI click the remote V-Switch icon and then click the storage icon.
7. Verify that the remote V-Switch can see the LUNs exposed by the disk array properly.
8. Refer to the GDR manual and follow the instructions to create the DR setup according to your design requirements.
9. Proceed with the steps described in, [page 7](#).

Pre-production Tests

Before implementing GDR, SANRAD recommends performing tests to verify functionality and ensure that your GDR design is viable. These tests should be conducted while users are not connected to the iSCSI volumes and after the GDR system setup configuration has been completed (including starting the replication and finishing the initial data synchronization).

Verifying Replication

Synchronous Replication

If your GDR setup includes synchronous replication, test it by performing planned failover and fallback described on page 9.

Asynchronous Replication

Verify functionality by following these steps:

1. Using StoragePro click the GDR icon
2. Check the “Pending PIT” column (see Figure 2). Under normal circumstances you should see either:
 - a. “Zero” pending PITs with increasing total data (in MB) indicating that the journal is accumulating the IO changes for the next PIT.
 - b. “One” pending PIT with increasing and decreasing total data (in MB) indicating that there is one PIT currently being replicated simultaneously while the journal is accumulating the IO changes for the next PIT. If you see any number higher than “One” this usually indicates that there is a problem with connectivity between the sites or with the replication interval (amount of data changes is higher than what the WAN line can transfer within the interval range).

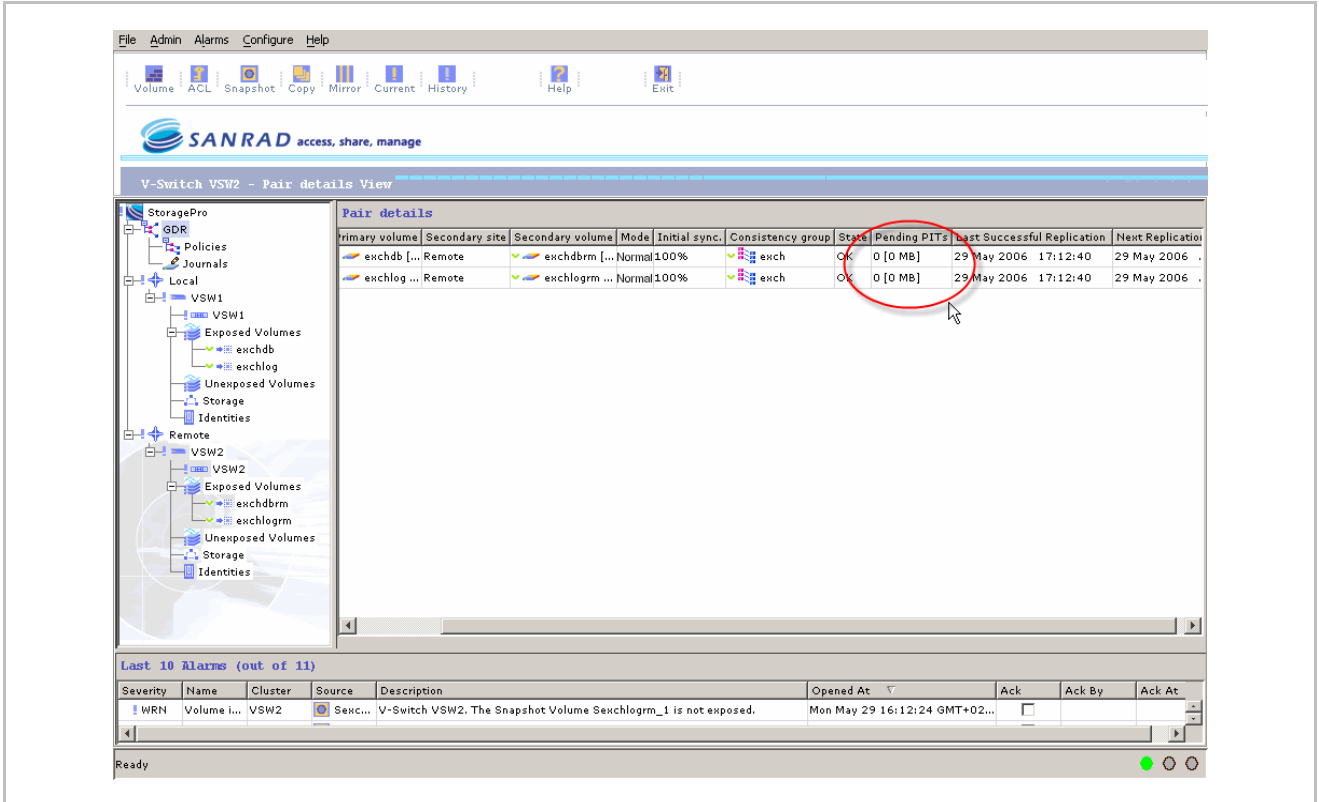


Figure 2. Pending Pits

3. Check when the last successful replication was (see Figure 3). Compare it to the next replication column and to the replication policy.

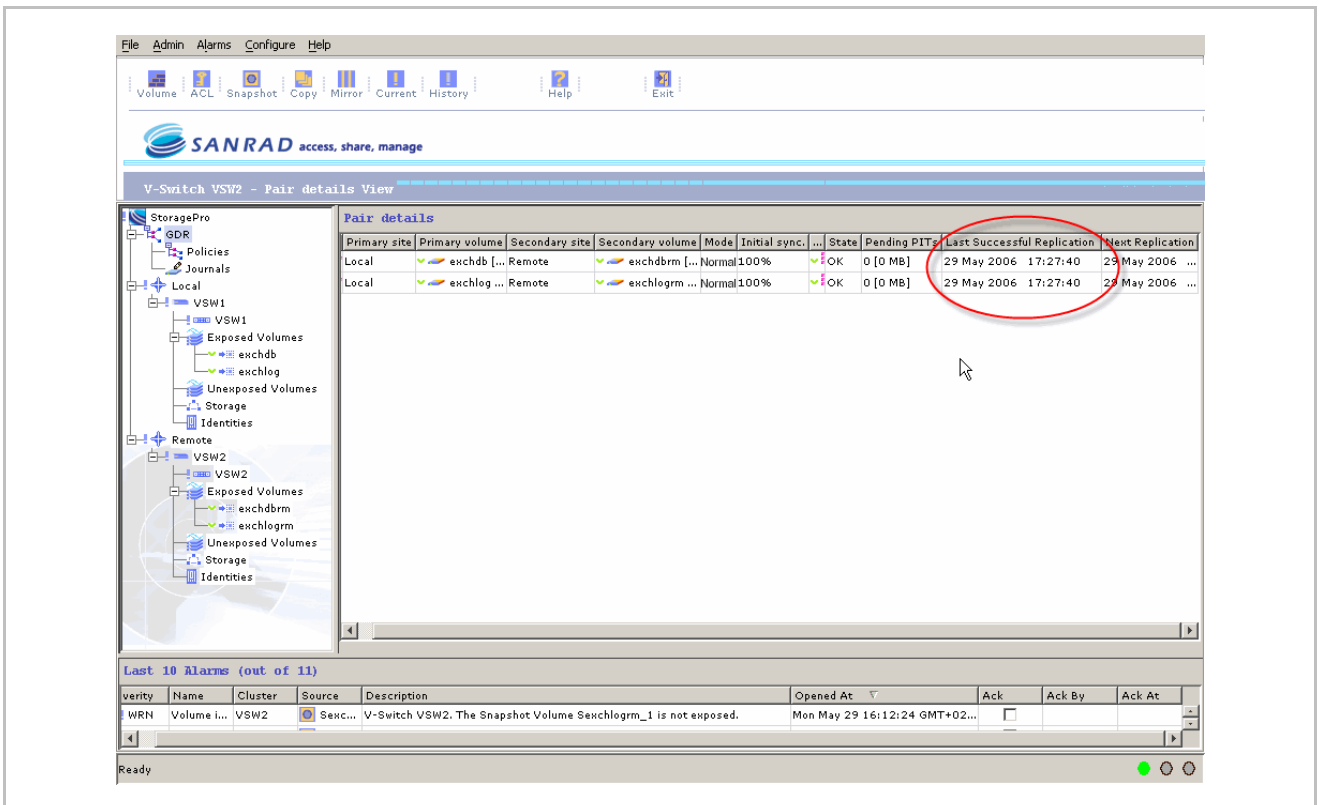


Figure 3. Last Successful Replication

Testing Planned Failover

1. Open Notepad, type some text and save it onto one of the iSCSI volumes that is being replicated.
2. Close all applications using the iSCSI volume.
3. Disconnect all iSCSI initiators connected to that volume.
4. Follow the instructions on how to perform Planned Failover in the GDR user manual or in the Online Help.
5. Verify that the hosts can access the iSCSI volume which is now exposed by the remote V-Switch.
6. Find the Notepad file created in the step 1, open it and verify its data integrity.
7. Disconnect all iSCSI initiators connected to that volume.
8. Follow the instructions on how to perform Fallback in the GDR user manual or in the Online Help.
9. Verify that the hosts can access the iSCSI volume which is exposed by the local V-Switch.
10. Restart all applications using the iSCSI volume.

Testing Disaster Failover (optional)

Warning:

This is a destructive test. The local site configuration will be deleted during the Failover process and will be recovered only after the Disaster Recovery process has been completed.

1. Open Notepad, type some text and save it onto one of the iSCSI volumes that is being replicated.
2. Disconnect the power cables from the local V-Switch.
3. Follow the instructions on how to perform Disaster Failover in the GDR user manual or in the Online Help.
4. Verify that the hosts can access the iSCSI volume which is now exposed by the remote V-Switch.
5. Find the Notepad file created in the step 1, open it and verify its data integrity.
6. Follow the instructions on how to perform Disaster Fallback in the GDR user manual or in the Online Help.
7. Verify that the hosts can access the iSCSI volume which is exposed by the local V-Switch.
8. Restart all applications using the iSCSI volume.

Final Implementation

Once all the tests have been completed, your GDR system is ready to be implemented.

1. Start replication and allow the initial data synchronization to complete.
2. If the remote hardware was connected to the local LAN in the local site, disconnect it from the local LAN and send it to the remote site location and bring it online.
3. Follow the instructions in the GDR user manual or in the Online Help to update the configuration.
4. Start the applications.
5. Allow the users to connect to the iSCSI volumes.

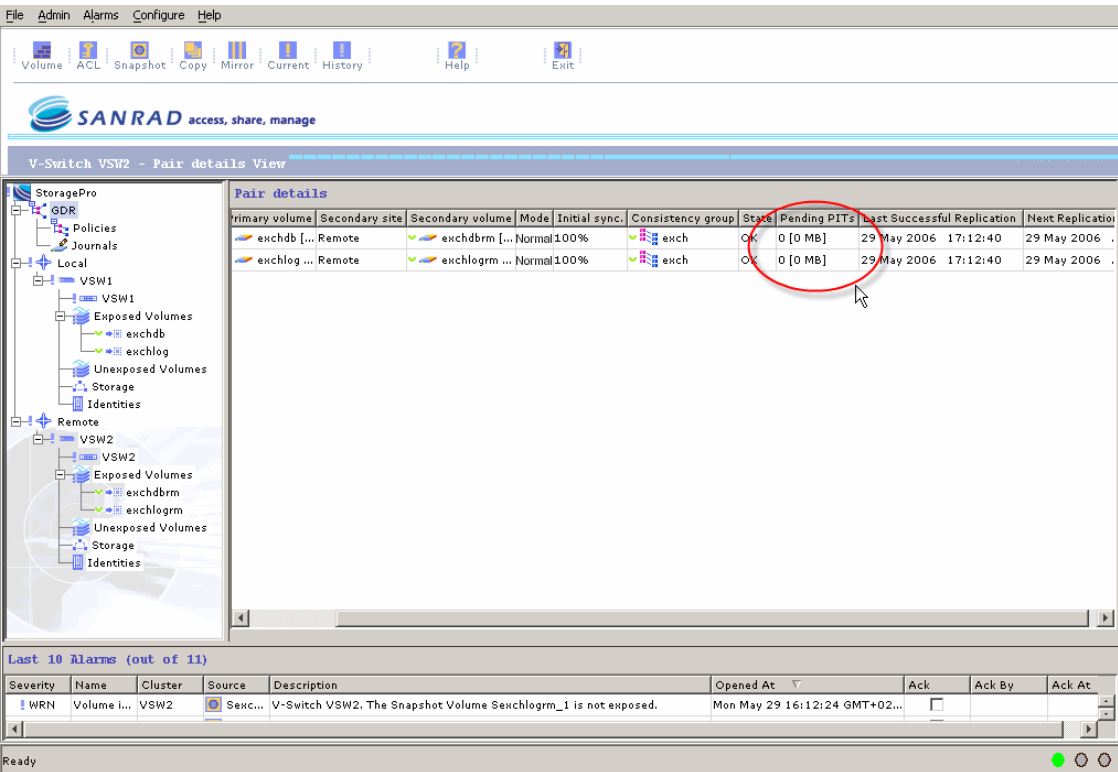
Monitoring

Monitoring your GDR system on a daily basis will help in detecting and preventing replication problems.

Asynchronous Replication

Monitor the status of asynchronous replication as follows:

1. Using StoragePro click the GDR icon
2. Check the “Pending PIT” column (see Figure 4). Under normal circumstances you should see either:
 - a. “Zero” pending PITs with increasing total data (in MB) indicating that the journal is accumulating the IO changes for the next PIT.
 - b. “One” pending PIT with increasing and decreasing total data (in MB) indicating that there is one PIT currently being replicated simultaneously while the journal is accumulating the IO changes for the next PIT. If you see any number higher than “One” this usually indicates that there is a problem with connectivity between the sites or with the replication interval (amount of data changes is higher than what the WAN line can transfer within the interval range).



Primary volume	Secondary site	Secondary volume	Mode	Initial sync.	Consistency group	State	Pending PITs	Last Successful Replication	Next Replication
exchdb [...]	Remote	exchdbrm [...]	Normal	100%	exch	OK	0 [0 MB]	29 May 2006 17:12:40	29 May 2006 .
exchlog [...]	Remote	exchlogrm [...]	Normal	100%	exch	OK	0 [0 MB]	29 May 2006 17:12:40	29 May 2006 .

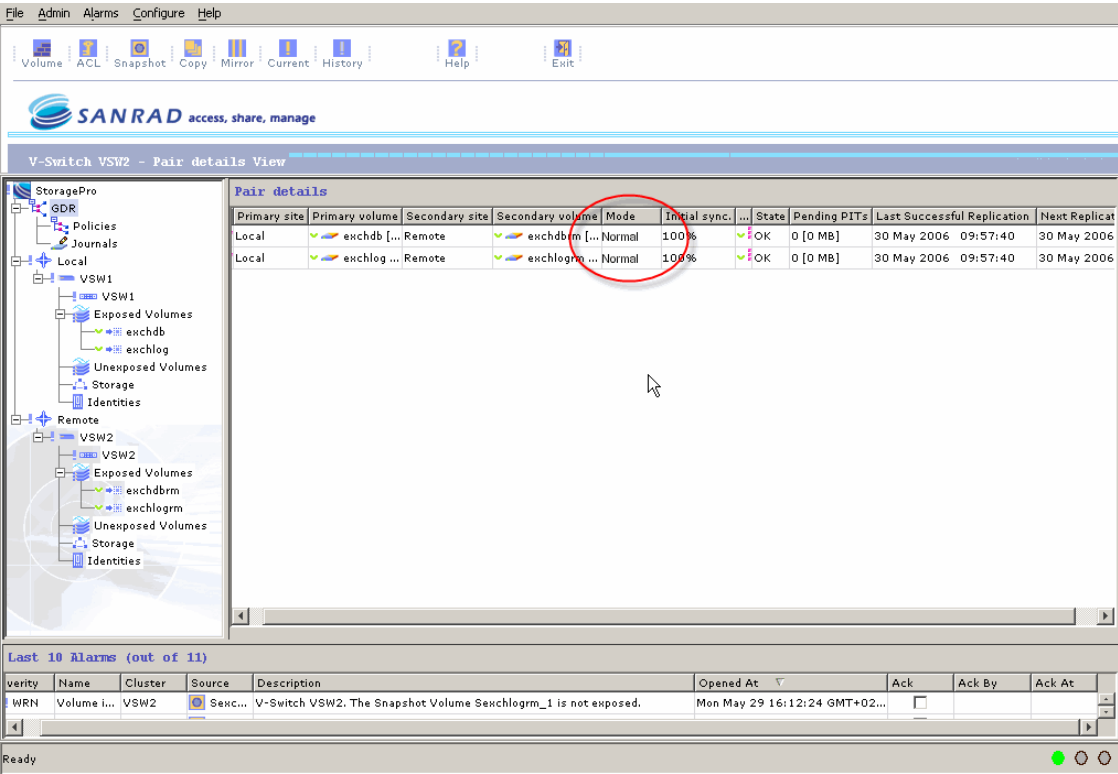
Severity	Name	Cluster	Source	Description	Opened At	Ack	Ack By	Ack At
WRN	Volume i...	VSW2	Sexc...	V-Switch VSW2. The Snapshot Volume Sexchlogrm_1 is not exposed.	Mon May 29 16:12:24 GMT+02...	<input type="checkbox"/>		

Figure 4. Monitoring Asynchronous Replication

Synchronous Replication Status

Monitor the status of synchronous replication as follows:

1. Using StoragePro click the GDR icon.
2. Verify that the status indicated in the Mode column is “Normal” and not “Need Sync” (See Figure 5). “Need Sync” indicates that the GDR system could not replicate the IO to the remote V-Switch.



The screenshot shows the SANRAD StoragePro interface. On the left is a tree view of the storage configuration. The main area displays 'Pair details' for V-Switch VSW2. Below this is a table of replication pairs.

Primary site	Primary volume	Secondary site	Secondary volume	Mode	Initial sync. ...	State	Pending PITs	Last Successful Replication	Next Replicat
Local	exchdb [...]	Remote	exchdbm [...]	Normal	100%	OK	0 [0 MB]	30 May 2006 09:57:40	30 May 2006
Local	exchlog ...	Remote	exchlogm ...	Normal	100%	OK	0 [0 MB]	30 May 2006 09:57:40	30 May 2006

Below the table is a section for 'Last 10 Alarms (out of 11)'. One alarm is visible:

verity	Name	Cluster	Source	Description	Opened At	Ack	Ack By	Ack At
WRN	Volume i...	VSW2	Sexc...	V-Switch VSW2. The Snapshot Volume Sexchlogm_1 is not exposed.	Mon May 29 16:12:24 GMT+02...	<input type="checkbox"/>		

Figure 5. Monitoring Synchronous Replication

Determining PIT Size

For asynchronous replication, how often we replicate the IO changes is influenced by several factors, one being the amount of IO changes (measured in MB). The amount of IO changes may change over time. This can cause a bottleneck in which one PIT did not complete its replication and another PIT is already waiting for replication.

To monitor the PIT size for a GDR pair:

1. Using StoragePro click the GDR icon.
2. Click the GDR pair, right click and choose “PIT Management” (see Figure 6).
3. In case of a consistency group, under the GDR object click Policies, in the right panel right click the Consistency Group followed by Show PIT (see Figure 7).

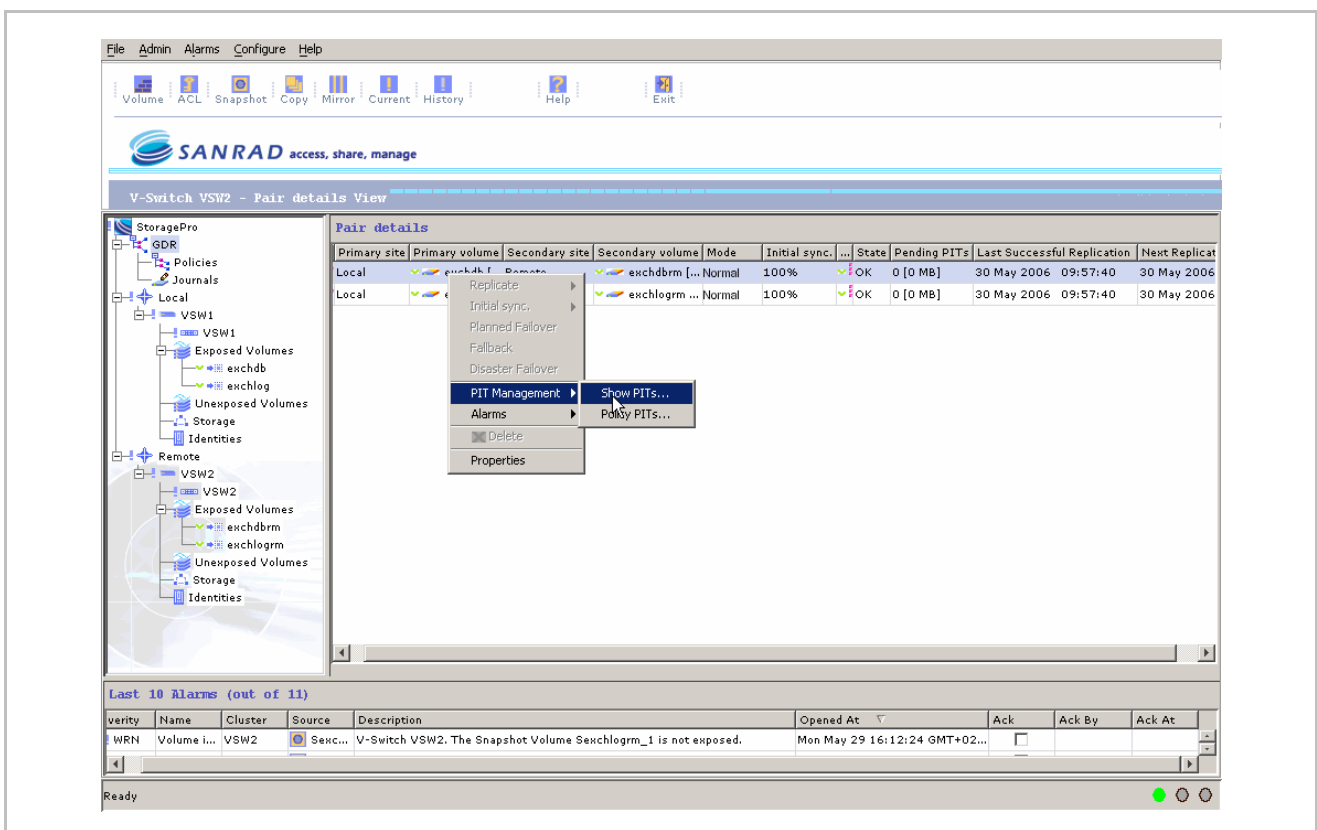


Figure 6. Monitoring Pits for GDR Pair

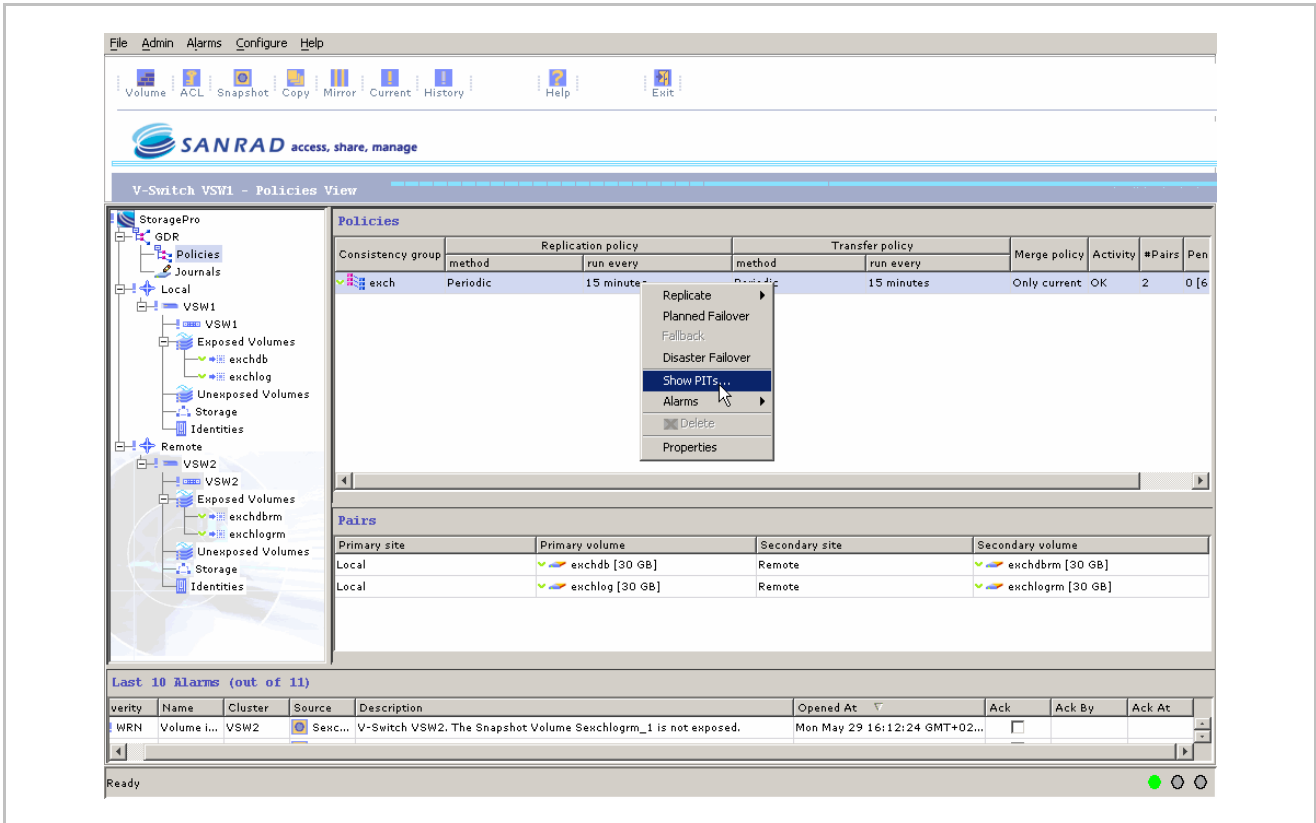


Figure 7. Monitoring Pits for Consistency Group

4. Check the number under “Pending Transfer [MB]” next to a PIT that is being replicated state (Figure 8).
5. Verify that this number is what you are expecting. If the total is higher than anticipated, consider revising the replication interval.

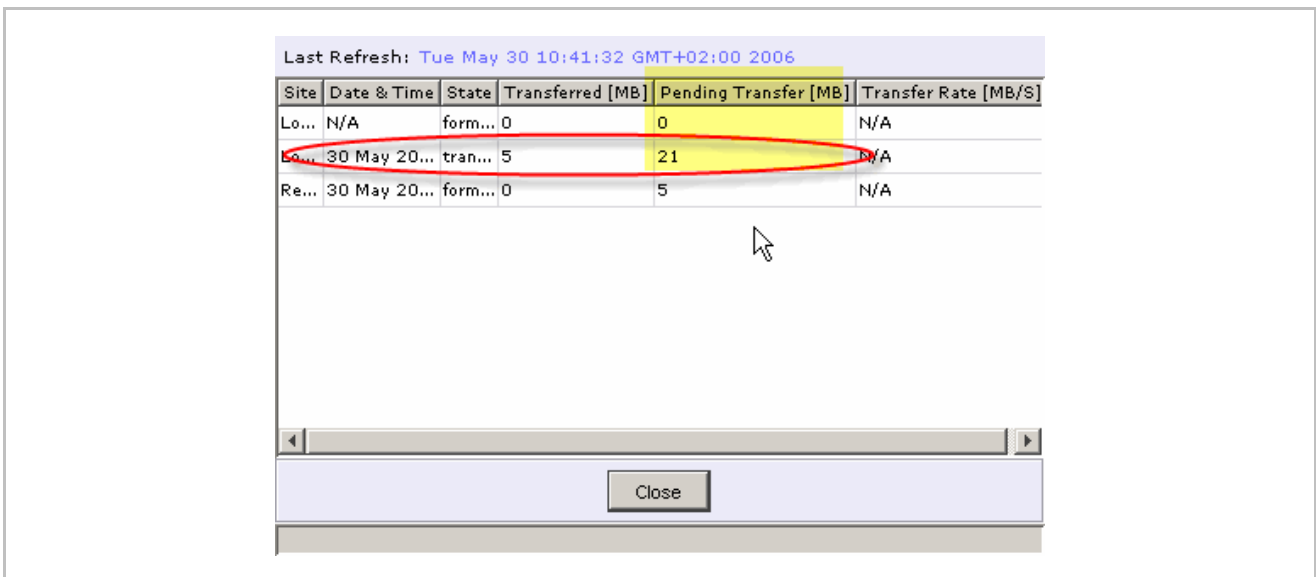


Figure 8. Monitoring Size of Pits being Replicated

Appendix A – Sample Design Information Template

Table 1 shows a sample design information template. Table 2 is blank and SANRAD recommends filling in a sample template before starting to configure your GDR application.

Table 1. Sample Design Template

General Info			
Amount of Data Changes		250MB/Hour	
WAN Line Bandwidth		T3: 45Mb	
Site Information			
		<u>Primary Site</u>	<u>Remote Site</u>
Site Name		Office	Warehouse
V-Switch name		VSooffice	VSwarehouse
Mgnt. IP Address		192.168.10.101	10.11.12.101
iSCSI Portal IP Address		192.168.10.110	10.11.12.110
Target name		office.data.targ	warehouse.data.targ
GDR Volume		office.data.vol	warehouse.data.vol
Resource for Volume	100 GB	office.data.vol.sub	warehouse.data.vol.sub
Journal Volume		office.journal	warehouse.journal
Resource for Journal	50 GB	office.journal.sub	warehouse.journal.sub
Snapshot Volume		office.snap	warehouse.snap
Resource for Snapshot	20 GB	office.snap.sub	warehouse.snap.sub
Routing information			
Gateway IP		192.168.10.1	10.11.12.1

Table 2. GDR Design Template

General Info			
Amount of Data Changes			
WAN Line Bandwidth			
Site Information			
		<u>Primary Site</u>	<u>Remote Site</u>
Site Name			
V-Switch name			
Mgmt. IP Address			
iSCSI Portal IP Address			
Target name			
GDR Volume			
Resource for Volume	GB		
Journal Volume			
Resource for Journal	GB		
Snapshot Volume			
Resource for Snapshot	GB		
Routing information			
Gateway IP			