Sun StorEdge™ Availability Suite Software — Improving Data Replication Over a Highly Latent Network

A Best Practice

Sun Microsystems, Inc.
4150 Network Circle
Santa Clara, CA 95054 U.S.A.
650-960-1300

Part No. 816-7180-10
August 2002, Revision A

Send comments about this document to: docfeedback@sun.com

Please
Recycle

Contents

Improving Data Replication Performance Over A Highly Latent Network

Businesses today often implement data replication as part of a disaster recovery and business continuance strategy. Remote mirroring can be used to create multiple data copies to help enable increased information access. While data replication can provide many opportunities for enhancing business operations, network bandwidth constraints and latencies can introduce additional latencies to storage activity. In turn, these delays can impede application performance in ways that can be hard to diagnose. This best practice document provides procedures that address the issue of improving data replication performance over a highly latent network.

The Sun StorEdge™ Availability Suite Software includes the remote mirror software and the point-in-time copy software. These two software packages are designed and integrated to help enable improved application performance by offloading volume replication activity to a copy of the application's volumes.

The point-in-time copy software is a point-in-time snapshot facility that runs in a Solaris™ Operating Environment (Solaris OE). A point-in-time copy snapshot, also called a point-in-time copy, is an instantly-available, time-fixed, replicated view of a momentarily-quiesced volume. The point-in-time copy software is intended for use as part of a backup or secondary application processing plan to help provide nearly instantaneous availability of replicated storage.

The remote mirror software is a remote replication facility that runs in a Solaris OE. Remote replication is a continuously updated, replicated view of a constantly changing volume. The remote mirror software is primarily used as part of a disaster recovery and business continuance plan. It provides a means with which to create redundant storage of critical information across physically separate sites.

Decoupling storage activity of the application from the remote replication of the storage means network latencies introduced during replication are no longer propagated back to the application. This helps to enable improved application

efficiency while helping to meet requirements for providing redundant storage to a remote site across a bandwidth-constrained network. Using the point-in-time copy software and the remote mirror software can also help increase data availability, help to provide new opportunities for information sharing, and provide capabilities for hot backup while continuing online activity.

Network Latency

The single greatest issue in any remote replication scenario is network latencies introduced by speed of light, network contention, and limited bandwidth. These latencies can severely impact overall application performance. Consider that today's disk storage systems can routinely deliver 10,000 or more input/output operations per second (iops) at latencies less than 6ms over multiple 2Gbps channels. Networks have far less bandwidth over greater distances with much higher costs associated.

Consider the standard wide area network (WAN) connections in TABLE 1. When evaluating how to implement a remote replication environment, bandwidth versus cost is a significant factor in how to configure and implement an application. When you use this to calculate the possible iops for typical links (TABLE 2), it is obvious that synchronous writes for most applications will suffer significant performance issues. In fact, because of the potential overrun of queues, asynchronous writes will have the same issues.

TABLE 1 Link Bandwidth and Cost

Link	Bandwidth	Est. Cost/mo
T-1	1.544Mbps	$750 - $1200
T-3	43.232Mbps	$12,000 - $24,000
OC-3	155Mbps	$32,000 - $49,000
OC-12	6.22 Mbps	
OC-48	2.5 Gbps	
OC-192	9.6 Gbps	

TABLE 2 Bandwidth Constrained Latency

Link	Bandwidth	Throughput	8iopsKB	300GB init
T-1	1.544mbps	192KB/sec	24	434hrs
OC-3	155Gbps	19MB/sec	2400	4.3hrs

The concerns increase when you consider the impact of the speed of light, switching latencies, and network contention. TABLE 3 describes theoretical latency for round trips at the speed of light as compared to a test ping across a live WAN during off peak hours. If you consider a link of any bandwidth between NYC and San

Francisco, the theoretical serialized (per process or volume) synchronous write maximum is 33 iops. If you consider actual tested latency and include the write at the other side, latency is increased and throughput is reduced further.

TABLE 3 Round-trip Theoretical Versus Real WAN Latencies

From NYC to	Distance	Latency	Real test
San Francisco	2565mi	28ms	95ms
Denver	1627mi	18ms	64ms
Atlanta	825mi	9ms	126ms
Chicago	776mi	8ms	98ms
Boston	205mi	2ms	14ms
New Jersey	50mi	<1ms	47ms

The following procedures gives one solution for dealing with replicating data. The objective is to maximize and manage production performance while ensuring data consistency at the remote site.

Method

Network delays during replication can impede application performance in these two situations:

- When applications wait on acknowledgement of write completion before continuing

  Synchronous replication writes need to complete and be acknowledged on the remote host as well as the primary host before the primary host can continue, which can have a significant impact on transaction latency.

- When access to a disk block is requested while the block is being replicated

  Synchronous replication reserves the disk block until replication is acknowledged. Therefore, any application writes to the disk block are delayed while the block is being replicated. Asynchronous replication reserves the disk block until a write queue entry can be allocated. Therefore, any application writes to the disk block are delayed only until a write queue entry can be allocated. Write queue entries remain allocated while the block is being replicated, which can cause a moderately write-intensive application to use all available write queue entries. When this occurs, asynchronous replication transitions in and out of synchronous replication causing some application writes to be delayed.

This best practice describes a technique to offload the replication source from the application volume(s) to help improve the application's timely access to storage. Snapshots taken at regular intervals with the point-in-time copy software provide the source for volume replication to the remote host. Replicating from a snapshot of the application volume(s) helps alleviate disk contention so the application can run more efficiently. The remote mirror software replicates only the changes that occur between snapshots. The use of snapshots helps avoid over-replication, which can happen with I/O hot spots that are updated repeatedly over very short periods of time.

Snapshots are also performed on the remote host for increased data availability and information access. After data is replicated to the remote host, a snapshot is taken so that the point-in-time copy is available while more recent changes are being replicated. The snapshot on the remote host should be available as a read-only copy to maintain the availability of a consistent point-in-time copy in the event of a failure of the local host. The point-in-time copy software supports multiple shadows of a

single master volume, such that more than one snapshot of the replicated volume can be made. Additional snapshots can be used for purposes such as backup or secondary application processing.

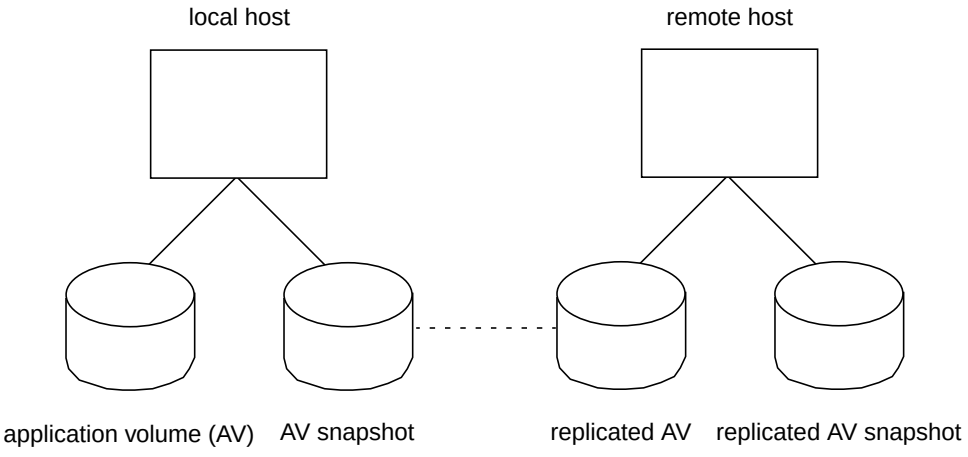

FIGURE 1

Figure 1: The point-in-time copy software is used to create a snapshot of the application volume (AV). The AV snapshot is then replicated to the remote host. On the remote host, a snapshot is taken of the replicated AV to maintain a consistent point-in-time snapshot while application and replication activities continue.

Note – The method discussed in this best practice is designed for non-Sun Cluster configurations. See 816-5127, the *Sun Cluster 3.0 and Sun StorEdge Software Integration Guide*, for information on using the Sun StorEdge Availability Suite Software with Sun Cluster configurations.

Planning and Setup

Before starting replication activity:

- Define appropriately-sized point-in-time copy volume sets and remote mirror volume sets.
- Evaluate methods for performing full volume copies to the remote site.
- Assess network bandwidth and latency, volume set size, and channel bandwidth to see if network initialization is feasible.
- Determine whether the network bandwidth is shared or dedicated.
- Estimate the expected quantity of new and changed data within an update interval to specify the required type of WAN link.
- Calculate an estimated time to copy using new and changed data and available bandwidth. There can be significant protocol overhead involved, and it is best to run a test remote mirror copy to see how to account for that overhead.

When the primary and secondary volumes are consistent, a replication cycle can be initiated.

Point-in-time Copy Volume Sets

Point-in-time copy volume sets consist of a master volume, a shadow volume, a bitmap volume, and an optional overflow volume. A bitmap is used to track changes in the master and shadow volumes.

A volume set can be enabled in several configurations. An independent shadow volume is a full volume copy of the master. A dependent shadow volume relies on the master volume for all unmodified data blocks (blocks are copied to the dependent shadow upon write request to the master). A dependent shadow volume is used when a separate, independent copy of the data is not required. When a dependent shadow volume is smaller than the master volume, it is said to be compact. A compact dependent shadow volume is used when the amount of change that occurs to a master volume is a significant percentage less than the size of the volume. Overflow volumes can be designated for times when the amount of modified data exceeds the space allocated to the compact dependent volume.

Use of an independent shadow volume on the local host can provide additional application efficiency since it does not rely on the master volume after its copy is complete. An independent shadow volume is required at the remote site to maintain a consistent point-in-time copy in the event of a failure during replication. In this case, the volume being updated at the time of failure would be considered

inconsistent, and the independent shadow volume would be used for recovery purposes. See 816-4313, the *Sun StorEdge Availability Suite 3.1 Point-in-time Copy Software Administration and Operation Guide* for further information.

Remote Mirror Volume Sets

A remote mirror volume set consists of a primary volume residing on a local host and a secondary volume residing on a remote host. The volume set also includes a bitmap volume on each host to track write operations and differences between the volumes. During normal replication operations, the primary volume is the point-in-time copy shadow volume on the local host, and the secondary volume is the point-in-time copy master volume on the remote host.

Required Volumes

To replicate a single application volume, the master volume, between a local host and a remote host requires that the following volumes be correctly sized and allocated on their respective hosts. After being created, they need to be configured into the point-in-time copy software and the remote mirror software as outlined in the remainder of this best practice.

The following volumes need to be created on the local host:

- Point-in-time copy master volume
- Point-in-time copy shadow volume = remote mirror primary volume
- Point-in-time copy bitmap
- Remote mirror bitmap

The following volumes need to be created on the remote host:

- Point-in-time copy master volume = remote mirror secondary volume
- Point-in-time copy shadow volume
- Point-in-time copy bitmap volume
- Remote mirror bitmap volume

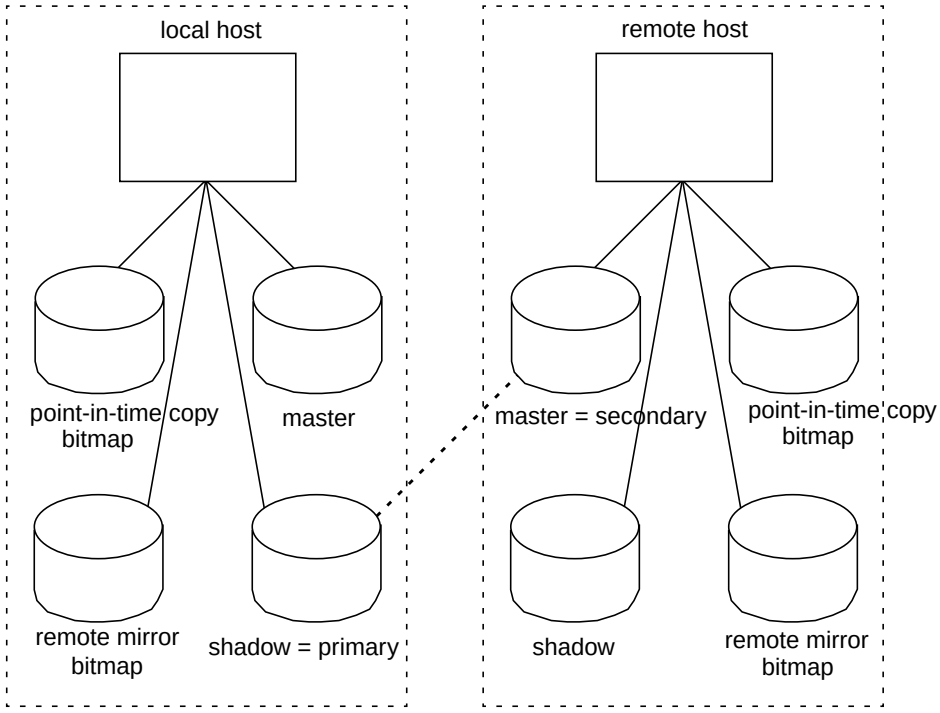

FIGURE 2

Figure 2: Separate point-in-time copy volume sets are defined on both the local and remote hosts. The point-in-time copy shadow volume on the local host serves as the primary volume for the remote mirror volume set. The point-in-time copy master volume on the remote host serves as the secondary volume for the remote mirror volume set. A single remote mirror volume set has a bitmap volume on the local host and another bitmap volume on the remote host.

Volume Size

Where an independent shadow is used, the shadow volume must be equal in size to the master. For use with the remote mirror software for a disaster recovery application, it is recommended you use an independent shadow volume, but not required. If a compact dependent shadow volume is used, the percentage of change of the master volume between snapshots needs to be understood to determine the sizing requirements. The point-in-time copy bitmap volume is sized according to the master volume and the type of volume set.

This best practice is based on the fact that a fully replicated point-in-time copy of the local host's point-in-time copy master volume is desired on the remote host. Therefore the remote host's secondary volume must be the same size as the local host's point-in-time copy master volume. This sizing of the point-in-time copy shadow volumes, on either the local or remote host, can be sized as either independent, dependent, or compact dependent shadows. Multiple shadow volumes of either the local host's point-in-time copy master volume or the remote host's secondary volume can be incorporated.

For independent and fully dependent volume sets:

- 8 Kbyte per 1 Gbyte of master volume size (rounded up to the nearest whole Gbyte), plus an additional 24 Kbyte for overhead.

  For example, to shadow a 3 Gbyte master volume, the bitmap size must be (3 * 8 Kbyte) + 24 Kbyte, or 48 Kbyte in size.

For compact dependent shadow volume sets:

- 256 Kbyte per 1 Gbyte of master volume size (rounded up to the nearest whole Gbyte), plus 8 Kbyte per Gbyte of master volume (rounded up to the nearest whole Gbyte), plus an additional 24 Kbyte for overhead.
- For example, to shadow a 3 Gbyte master volume, the bitmap size must be (3 * 256 Kbyte) + (3 * 8 Kbyte) + 24 Kbyte, or 816 Kbyte in size.

If the source of replication is an independent shadow, the primary and the secondary volumes must be the same size. If the source of replication is a compact dependent shadow, the secondary volume must be sized equal to the point-in-time copy master volume on the local host because an initial full copy of the volume needs to be replicated to the secondary.

For remote mirror bitmap volumes:

- 1 Kbyte + 4 Kbyte per Gbyte of device storage space.

  For example, a 2 Gbyte data device requires a bitmap size of 9 Kbyte. Round up any partial Gbyte counts to the next Gbyte.

All bitmap volumes should be mirrored and the volumes should be mounted at boot. Do not use memory based bitmaps.

Determining Initial Full Synchronization Method

Consideration must be given to the method used to synchronize the primary and secondary volumes. To determine feasibility of performing a full volume copy via the network, divide the volume size by the actual transfer rate of the network, over the time period available, to arrive at the length of time required for the operation. For example copying a 300Gbyte volume across a T1 network (not including network overhead) 434 hours. For a situation like this, alternative methods are in order. One alternative would be to co-locate the secondary server for initial synchronization, then ship it the secondary site. A third and probably more easily accomplished method is initial synchronization of a volume set by sending physical copies of the local application volume devices to the remote site for installation.

To resume application activity while the volume is not yet restored on the remote site, the remote mirror software can be placed in logging mode so changes to the data are tracked. After the secondary volume is in place, an update to the point-in-time copy on the local host followed by a resynchronization from the primary volume to the secondary volume captures those changes made during initialization of the secondary volume.

Performing an Initial Full Copy via Network

A series of steps are required to synchronize the secondary volume with the primary volume before the replication cycle can be implemented. First, the remote mirror volume sets on both the local and remote hosts are enabled, putting the remote mirror software in logging mode. Then the point-in-time copy volume sets are added so that the remote mirror software can track changes in the snapshot volume. Next, the application is momentarily quiesced while the point-in-time copy volume set on the local host is enabled, and then immediately resumed. The full copy is then issued from the local host. When the copy is complete, the point-in-time copy volume set on the remote host is enabled to establish a physically separate copy of the replicated data. When the copy to the snapshot on the remote host is complete, the shadow volume on the local host is updated with changes that have occurred since the initial point-in-time copy. Those changes are then replicated to the secondary volume on the remote host and the remote shadow is updated as well.

Performing an Initial Full Copy via Physical Media

Instead of performing the full volume copy, you can use third party backup software to create a volume-level tape backup, or use the `dd` utility to create physical disk copies. Send the tapes or disks to the remote site for installation.

Make sure the backup copy is a physical copy (using `dd`(1M) command) and not a logical copy (for example, one made using `tar` or `cpio`). The copies must have identical blocks, not just identical files.

Replication Cycle

The objective of the replication cycle is to take point-in-time snapshots of the master volume on the local host at regular intervals, replicate the changes between snapshots to the secondary volume, and update the snapshot on the remote host with the changes. An interval time between updates needs to be determined. Factors to consider in determining how often to update snapshots include the percentage of changed data for a given window of time and the potential risk of data loss. The snapshot on the remote host can be as much as an interval behind the snapshot on the local host. Also, the application continues to update the master volume on the local host while replication and remote copying take place. This means the potential for data loss approaches the amount of data changed during two update intervals. Ideally, windows are timed as small as possible to minimize the risk of data loss.

Caution – Do not automate the replication cycle in such a way that when the local host crashes and reboots, a potentially inconsistent volume is automatically replicated, invalidating volumes at the remote host.

Failure and Recovery

Upon failure, the first critical step is to determine where the most recent consistent volume exists. The data services log, `/var/opt/SUNWesm/ds.log` reflects the most recent successfully completed operation.

- If the local host crashed during synchronization with the remote host, then it is possible the secondary volume is inconsistent. If the local volumes are unavailable, the most recent point-in-time copy snapshot would be on the remote shadow volume.
- If the local host crashed after replication successfully completed, but before the point-in-time copy update to the remote shadow completed, then the most recent consistent volume on the remote host would be the master volume.
- If the point-in-time copy to the remote shadow volume did complete before the crash, then the shadow volume would be where the most recent consistent volume exists.

The recovery methods in this document cover each of these scenarios.

When the local host becomes available for synchronization with the remote host, the application can continue to run with access to the shadow volume, while a full copy is achieved by replicating from the remote master volume to the local master volume. Point-in-time copies to the remote master are ceased for the duration of the replication, whether through a network or when the physical copy is performed. After the local master volume is consistent with the remote master volume, a full copy to the local shadow is performed. Then the application is quiesced so the remote master can be updated with changes that occurred to the remote shadow during replication. After the changes are propagated to the local shadow, the replication cycle can be initiated.

If it was previously determined that a full copy over the network is not feasible, refer to the procedure, “Performing an Initial Full Copy via Physical Media” on page 12, for an alternate recovery method.

Procedures

Note – You must already have installed the Sun StorEdge Availability Suite Software on both the local and remote hosts before beginning the following procedure. For help with installing the software, see 816-4413, the *Sun StorEdge Availability Suite 3.1 Remote Mirror Software Installation Guide;* and 816-4312, the *Sun StorEdge Availability Suite Point-in-time Copy Software Installation Guide*.

This section includes the following procedures:

▼ To Perform an Initial Full Copy via Network

To establish the initial copy you need to first establish an initial point-in-time copy.

1. **Quiesce the application.**
2. **On the local host, enable the point-in-time copy volume set.**

```
# iiadm -e {ind | dep} master_vol shadow_vol bitmap_vol
```

3. **Resume the application.**

   If you wish to view the copy in progress:

```
# iiadm -i volume_set
```

4. **On the local host, enable the remote mirror volume set where** `phost` **is the local host,** `pdev` **is the application volume (AV) snapshot,** `pbitmap` **is the local remote mirror bitmap,** `shost` **is the remote host,** `sdev` **is the replicated AV, and** `sbitmap` **is the remote host's remote mirror bitmap. See the** `sndradm` **man page for more information on a volume set.**

```
# sndradm -e phost pdev pbitmap shost sdev sbitmap ip {sync | async}
```

   For example:

```
# sndradm -e localhost /dev/vx/rdsk/data_dg/vol01 \
/dev/vx/rdsk/data_dg/vol01_rm_bmp \
remote_host /dev/vx/rdsk/data_dg/vol01_snap \
/dev/vx/rdsk/data_dg/vol01_rm_bmp ip async
```

5. **On the remote host, enable the remote mirror volume set.**

```
# sndradm -e phost pdev pbitmap shost sdev sbitmap ip {sync | async}
```

   On each host, view the enabled remote mirror set:

```
# sndrstat

A P /vx/rdsk/data_dg/vol01_snap => remote_host:vx/rdsk/data_dg/vol01 100.00 LOG
```

Also on each host, view the remote mirror volume set as follows:

```
# sndradm -i

local_host /dev/vx/rdsk/data_dg/vol01_snap /dev/vx/rdsk/data_dg/vol01_rm_bmp

remote_host /dev/vx/rdsk/data_dg/vol01 /dev/vx/rdsk/data_dg/vol01_rm_bmp ip

async
```

6. **On the local host, initiate the full copy and wait for completion.**

```
# sndradm -m RM-set
# sndradm -w RM-set
```

7. **On the remote host, enable the independent point-in-time copy and wait for completion.**

```
# iiadm -e ind master_vol  shadow_vol  bitmap_vol
```

8. **Quiesce the application**

9. **On the local host, put the remote mirror software in logging mode.**

```
# sndradm -l RM-set
Put SNDR into logging mode? (Y/N) [N]:Y
```

Confirm that the RM-set is in logging mode:

```
# sndrstat
  A P /vx/rdsk/data_dg/vol01 => remote_host:vx/rdsk/data_dg/vol01 0.00 LOG
```

To print information about the status of the remote mirror volume set:

```
# sndradm -p
```

10. **On the local host, update the shadow volume from the master volume.**

    The point-in-time copy volume set is referred to by the shadow volume name.

```
# iiadm -u -p s volume-set
```

11. **Resume the application.**

12. **On the local host, replicate to the secondary volume changes made to primary volume, and wait for the synchronization to complete.**

```
# sndradm -u RM-set
# sndradm -w RM-set
```

13. **From the local host, put the RM-set in logging mode.**

```
# sndradm -l RM-set
```

14. **On the remote host, update the shadow with changes replicated to the secondary volume.**

```
# iiadm -u -p s volume-set
```

    Both the local and secondary volumes are now consistent. The application can be resumed and the replication cycle can be initiated.

▼ To Perform an Initial Full Copy via Copy to Physical Media

1. **Quiesce the application.**
2. **On the local host, enable the point-in-time copy volume set.**

```
# iiadm -e {ind | dep} master_vol shadow_vol bitmap_vol
```

3. **Resume application.**
4. **Perform the backup of the application volumes to tape. Make sure you are using a block level copy mechanism such as** `dd` **or** `ufsdump`**.**

   `dd` example:

```
# dd if=/dev/rdsk/c1t0d3s2 ibs=1024 of=/dev/rmt/0 obs=1024
```

5. **Send the physical media to the remote site and install as the remote master volume.**
6. **On the remote host, enable the point-in-time copy volume set and wait for the copy to complete.**

```
# iiadm -e ind master_vol shadow_vol bitmap_vol
```

7. **On both hosts, enable the remote mirror volume set where** *phost* **is the local host,** *pdev* **is the application volume (AV) snapshot,** *pbitmap* **is the local remote mirror bitmap,** *shost* **is the remote host,** *sdev* **is the replicated AV, and** *sbitmap* **is the remote remote mirror bitmap. Notice the use of the** `sndradm -E` **option to indicate that the primary and secondary volumes are consistent. See the** `sndradm` **man page for more information on a volume set.**

```
# sndradm -E phost pdev pbitmap shost sdev sbitmap ip {sync | async}
```

   For example:

```
# sndradm -E localhost /dev/vx/rdsk/data_dg/vol01 \
/dev/vx/rdsk/data_dg/vol01_rm_bmp \
remote_host /dev/vx/rdsk/data_dg/vol01_snap \
/dev/vx/rdsk/data_dg/vol01_rm_bmp ip async
```

On each host, view the enabled RM-set:

```
# sndrstat

A P /vx/rdsk/data_dg/vol01_snap => remote_host:vx/rdsk/data_dg/vol01 100.00 LOG
```

Also on each host, you can view the RM-set as follows:

```
# sndradm -i

local_host /dev/vx/rdsk/data_dg/vol01_snap /dev/vx/rdsk/data_dg/vol01_rm_bmp

remote_host /dev/vx/rdsk/data_dg/vol01 /dev/vx/rdsk/data_dg/vol01_rm_bmp ip

async
```

8. **Quiesce the application.**
9. **On the local host, update the shadow with changes that occurred during the replication of the application volume to the secondary volume.**

   ```
   # iiadm -u s volume-set
   ```

10. **Resume the application.**
11. **On the local host, replicate the changes from the primary volume to the secondary volume, and wait for completion. Upon completion, place the RM-set in logging mode.**

    ```
    # sndradm -u RM-set
    # sndradm -w RM-set
    # sndradm -l RM-set
    ```

12. **On the remote host, update the remote shadow.**

    ```
    # iiadm -u -p s volume-set
    ```

13. **Begin the replication cycle.**

▼ To Replicate

After performing the initial full copy, ongoing replication can be accomplished using the following steps. This replication cycle can be run at regular intervals to continuously update the most recent copy at the remote site.

1. **On the local host, put the remote mirror volume set in logging mode if it is not already.**

```
# sndradm -l RM-set
```

2. **Quiesce the application**
3. **On the local host, update the shadow volume from the master volume.**

```
# iiadm -u -p s volume-set
```

4. **Resume application.**
5. **On the local host, replicate to the secondary volume changes made to primary volume and wait for the synchronization to complete.**

```
# sndradm -u RM-set
# sndradm -w RM-set
```

6. **On the local host, put the volume set in logging mode.**

```
# sndradm -l RM-set
```

7. **On the remote host, update the shadow with changes replicated to the secondary volume.**

```
# iiadm -u -p s volume-set
```

8. **Repeat steps Step 2 through Step 7 at the desired interval.**

▼ To Recover via Network

1. **Review** `/var/opt/SUNWesm/ds.log` **to determine where the most recent consistent data copy exists.**

Note – Before making both the remote master and shadow volumes write-accessible, it is suggested that a backup be performed to preserve the most recent consistent volume.

---

2. **If most recent consistent data copy is on the remote master volume, take a point-in-time copy to update the remote shadow.**

```
# iiadm -u s volume-set
```

Or, if the most recent consistent data copy is on the remote shadow volume, update the master from the shadow volume.

```
# iiadm -u m volume-set
```

3. **If desired, resume application services on the remote shadow.**

4. **While the local host is unavailable, periodically quiesce the application to update the remote master from the remote shadow.**

```
# iiadm -u m volume-set
```

5. **From both hosts, disable the original RM-set that enabled replication from the local to the remote site.**

```
# sndradm -d RM-set
```

To view that the set is gone, use the print option for `sndradm`.

```
# sndradm -p
```

6. **When the local system is ready to receive the data from the remote site, on both hosts, add a new RM-set to enable replication from the remote master volume to the local master volume.**

```
# sndradm -e remote_host remote_master remote_bitmap local_host local_master local_bitmap ip async
```

7. **From the remote host, initiate the full copy from the remote host to the local host, and wait for replication to complete. Upon completion, place the RM-set in logging mode.**

```
# sndradm -m local_host:local_master
# sndradm -w local_host:local_master
# sndradm -l local_host:local_master
```

Note – After the full copy has been initiated, the point-in-time copy software updates to the remote master should cease until progressing to Step 10 below.

---

8. **Confirm that the original point-in-time copy volume set exists on the local host.**

```
# iiadm -i volume-set
```

If the set exists, do a full copy to the shadow volume.

```
# iiadm -c s master shadow bitmap
```

If the set does not exist, enable the set.

```
# iiadm -e {ind | dep} master_vol shadow_vol bitmap_vol
```

9. **Quiesce the application.**

10. **Update the remote master with changes to the shadow that occurred during the full copy to the local host.**

```
# iiadm -u m volume-set
```

11. **From the remote host, replicate the changes to the local master volume.**

```
# sndradm -u local_host:local_master
# sndradm -w local_host:local_master
```

12. **When replication is complete, update the local shadow.**

```
# iiadm -u s volume-set
```

13. **On both hosts, disable the RM-set that was used for replicating data back to the local host.**

```
# sndradm -d local_host:local_master
```

14. **On both hosts, add the original RM-set that enables replication from the local shadow volume to the remote master volume. Here, the -E option is used to reflect that the primary and secondary volumes are consistent when the set is enabled.**

```
# sndradm -E local_host local_shadow local_bitmap remote_host remote_master remote_bitmap ip {sync
| async}
```

15. **Resume the application on the local master.**

16. **Initiate the replication cycle.**

▼ To Recover via Copy to Physical Media

1. **Review** `/var/opt/SUNWesm/ds.log` **to determine where the most recent consistent data copy exists.**

Note – Before making both the remote master and shadow volumes write-accessible, it is suggested that a backup be performed to preserve the most recent consistent volume.

---

2. **If most recent consistent data copy is on the remote master volume, take a point-in-time copy to update the remote shadow.**

```
# iiadm -u s volume-set
```

Or, if the most recent consistent data copy is on the remote shadow volume, update the master from the shadow volume.

```
# iiadm -u m volume-set
```

3. **Perform the copy of the most recent consistent volume to the physical device using third party backup software or the** `dd` **utility.**

`dd` example:

```
# dd if=/dev/rdsk/c1t0d3s2 ibs=1024 of=/dev/rdsk/c3t0d1s2 obs=1024
```

4. **If desired, resume application services on the remote shadow.**

However, to preserve consistency between the primary and secondary volumes, point-in-time copy updates cannot be done until after the RM-set is enabled on both hosts.

5. **While the local host is unavailable, periodically quiesce the application to update the remote master from the remote shadow.**

```
# iiadm -u m volume-set
```

6. **From both hosts, disable the original RM-set that enabled replication from the local to the remote site.**

```
# sndradm -d RM-set
```

View that the set is no longer enabled using `sndrstat`.

```
# sndrstat RM-set
```

7. **Send the physical media to the local site and restore the volume as the local application volume when the local host is available.**
8. **Confirm that the original point-in-time copy volume set exists on the local host.**

```
# iiadm -i volume-set
```

If the set exists, do a full copy to the shadow volume.

```
# iiadm -c s master shadow bitmap
```

If the set does not exist, enable the set.

```
# iiadm -e {ind | dep} master_vol shadow_vol bitmap_vol
```

9. **Reenable the original RM-set where the primary volume is the local shadow volume and the secondary volume is the remote master volume.**

Notice the `-E` option to indicate the primary and secondary volumes are consistent.

```
# sndradm -E local_host local_shadow local_bitmap remote_host remote_master remote_bitmap ip {sync | async}
```

10. **Quiesce the application and put the RM-set in logging mode.**

```
# sndradm -l RM-set
```

11. **On the remote host, update the remote master with changes to the shadow that occurred during the full copy to the local host.**

```
# iiadm -u m volume-set
```

12. **On the local host, do a reverse update to replicate the changes from the secondary to the primary volume.**

```
# sndradm -u -r remote_host:remote_master
# sndradm -w remote_host:remote_master
```

13. **When replication is complete, update the local master from the local shadow.**

```
# iiadm -u m volume-set
```

14. **Resume the application on the local master.**

15. **Initiate the replication cycle.**

Sun Microsystems, Inc.
4150 Network Circle
Santa Clara, CA 95054 USA
650-960-1300 Fax 650-969-9131

For U.S. sales office locations, call: 800-821-4643

In other countries, call:

Corporate headquarters:	650-960-1300
Intercontinental sales:	650-688-9000