VERITAS Volume Replicator™

Interoperability in Mixed-Vendor Environments

Toby Creek | July 2003 | TR 3271

Abstract

VERITAS Volume Replicator™ is a product that integrates with Volume Manager™ to enable data to be replicated over any TCP/IP protocol network. This paper will examine issues around VVR configuration and integration with Network Appliance® storage systems and third-party storage. The paper will provide configuration recommendations and best-practices suggestions while detailing the implementation of the software.

Network Appliance® Inc.

1

Table of Contents

Network Appliance® Inc.

1) List of Figures and Tables

2) Purpose and Scope

This document will detail implementation of VERITAS Volume Replicator™ on a Solaris™ host. The products covered will include:

Configuring VERITAS Volume Manager™ (VxVM)

Configuring VERITAS Volume Replicator™ (VVR)

Setup and configuration of Network Appliance storage will be discussed only as they relate to the example configurations used in this document. References will be provided in the last section of this paper to other sources that will cover these tasks in detail.

3) Requirements and Assumptions

For the information contained in this report to be useful to the reader, several assumptions are made:

The reader has at least intermediate Solaris administration skills and has access to the root login or superuser privileges through another mechanism.

The reader has at least basic knowledge of VERITAS Volume Manager™ and has access to the administrative commands.

The reader has at least basic knowledge of Network Appliance filer administration and has administrative access to the filer via the command-line interface as well as FilerView® software.

The filer and host have the required licenses necessary to perform the activities outlined in this document.

The target system has the required protocol interconnects to perform the activities outlined in this document. At a minimum, both Ethernet and Fibre Channel interfaces should exist in the filer and the host, and connectivity should be established and verified.

The hosts must have Internet protocol connectivity available between them. Name services should be configured to simplify management.

This paper is not intended to be a definitive implementation guide. There are many factors that are not taken into account in this paper. Significant expertise may be required to solve logistical problems when the system is designed and implemented.

For the purposes of this report, filer administration is performed via the filer command line for clarity. FilerView can also be used.

4) Infrastructure

The specific infrastructure pieces used in this paper are described below:

The primary host is a SPARC-based host running Solaris 8 and VERITAS Volume Manager™ 3.5 with a VVR license installed. Two Fibre Channel host bus adapters (HBAs) are used to connect to the primary storage

The secondary host is a SPARC-based host running Solaris 8 and VERITAS Volume Manager™ 3.5 with a VVR license installed. A single Fibre Channel HBA is used to connect to the secondary storage.

Both hosts have a VxVM rootdg that created using local disk. The rootdg cannot be used as the VVR Replicated Disk Set (RDS). It is recommended that the rootdg exist on disks that are local to the system. Encapsulating and mirroring the disks containing the operating system is a recommended method of utilizing the rootdg.

This document assumes that the storage systems are already configured. References will be provided for additional documentation specific to Network Appliance storage configuration. For other vendors, consult the vendors' documentation.
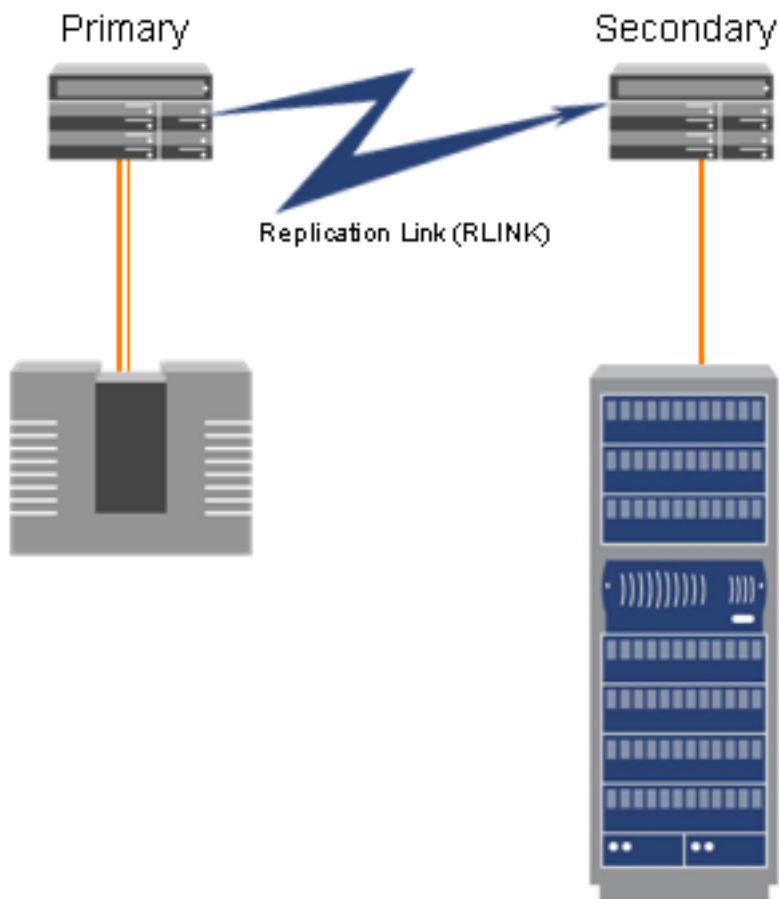

Figure 4-1) Basic Infrastructure Diagram
The diagram shows the primary host connected to the primary storage subsystem via redundant paths. The replication link is used to send data to the secondary, which stores the data on its attached subsystem. Since there is not an availability requirement on the secondary, that host only has a single path to its storage.

There are many possible configurations for VVR and its storage subsystems. The actual configuration will be dictated by technical constraints and business requirements. This paper will not cover all of the possible configurations; only the most basic will be discussed to reveal possible issues.

5) VERITAS Volume Manager™ Configuration with Network Appliance Storage

This section will detail the necessary configuration within VERITAS Volume Manager™ to prepare the host infrastructure.

5.1) Installing the Network Appliance Array Storage Library

The first task in configuring VERITAS Volume Manager™ is to install the array support libraries (ASL) for VxVM. Network Appliance provides an ASL for VxVM 3.5 and later on Solaris 8 and later. The ASL needs to be installed only on hosts that are connected to Network Appliance storage systems. If another vendor's storage is used, consult the documentation from that vendor to see if an ASL is available. The `vxddladm` command is used to describe the Network Appliance storage system capabilities to the Device Discovery Layer (DDL) of VxVM.

If you are using VxVM 3.5 or later, perform the following procedure:

1. To ensure compatibility, review the requirements for the ASL on the NOW™ (NetApp on the Web) Web site at http://now.netapp.com/NOW/download/tools/vasl. Download the `NTAPasl.pkg` file to the hosts running VxVM. We suggest you download and read the installation guide.

2. Install the package using the pkgadd utility.

    ```
    # pkgadd –d NTAPasl.pkg
    ```

    Follow the prompts according to the ASL installation documentation to complete the package installation.

3. Use the `vxdctl` command to refresh the VxVM configuration database.

    ```
    # vxdctl enable
    ```

4. Verify that the installation and detection of the filer were successful.

    ```
    # vxdmpadm listenclosure all
    ENCLR_NAME        ENCLR_TYPE      ENCLR_SNO            STATUS
    ===========================================================
    OTHER_DISKS       OTHER_DISKS     OTHER_DISKS          CONNECTED
    Disk              Disk            DISKS                DISCONNECTED
    Netapp0           Netapp          50:a9:80:00:02:00:8c CONNECTED
    ```

If the ASL is not available for your configuration, the array will be configured as "just abunch of disks" (JBOD). VxVM will have no specifications for the features of the Network Appliance array. Use the following method to describe the array where a Network Appliance ASL is not available:

1. Use the `vxddladm` command to add the array description.

```
# vxddladm addjbod vid=NETAPP
```

2. Verify that the array description was added correctly.

```
# vxddladm listjbod
VID      PID              Opcode   Page Code   Page Offset    SNO length
=======================================================================
SEAGATE  ALL PIDs            18        -1          36            12
SUN      SESS01              18        -1          36            12
NETAPP   ALL PIDs            18        -1          36            12
```

5.2) Adding LUNs to VERITAS Volume Manager™

LUNs that are to be placed under Volume Manager™ control need to have a Solaris volume label before VERITAS can recognize them. This requires using the `format` command. Systems with many LUNs connected can use the prtvtoc and fmthard programs to automate the process, though these commands are not discussed here. For more information, see the Solaris manual pages fmthard(1M) and prtvtoc(1M).

In the following example, a LUN on a Network Appliance filer is given a default volume label:

1. Run the `format` command as the root user and select the Network Appliance LUN.

```
# format
Searching for disks...done

AVAILABLE DISK SELECTIONS:
       0. c0t0d0 <FUJITSU-MAJ3182MC-3702 cyl 12435 alt 2 hd 5 sec 574>
          /pci@17,4000/scsi@3/sd@0,0
       1. c0t1d0 <FUJITSU-MAJ3182MC-3702 cyl 12435 alt 2 hd 5 sec 574>
          /pci@17,4000/scsi@3/sd@1,0
       2. c1t1d0 <NETAPP-LUN-0.2 cyl 5118 alt 2 hd 16 sec 512>
          /pci@15,2000/fibre-channel@1/sd@1,0
       3. c1t1d1 <NETAPP-LUN-0.2 cyl 5118 alt 2 hd 16 sec 512>
          /pci@15,2000/fibre-channel@1/sd@1,1
       4. c3t1d0 <NETAPP-LUN-0.2 cyl 5118 alt 2 hd 16 sec 512>
          /pci@15,4000/fibre-channel@5/sd@1,0
       5. c3t1d1 <NETAPP-LUN-0.2 cyl 5118 alt 2 hd 16 sec 512>
          /pci@15,4000/fibre-channel@5/sd@1,1
Specify disk (enter its number): 2
selecting c1t1d0
Disk is not labeled.  Label it now? y
```

In the above output, two LUNs are presented on two different paths. Each device need only be labeled once. Accepting the default label is the least time-consuming method. VxVM will write its own label when the LUNs are initialized.

2. Repeat until all LUNs have a valid Solaris volume label.

3. Prompt Volume Manager™ to refresh its configuration database.

```
# vxdctl enable
```

4. Verify that the disks are presented correctly to Volume Manager™.

```
# vxdisk list
DEVICE       TYPE      DISK        GROUP       STATUS
c0t0d0s2     sliced    -           -           offline
c0t1d0s2     sliced    c0t1d0s2    rootdg      online
c1t1d0s2     sliced    -           -           error
c1t1d1s2     sliced    -           -           error
c3t1d0s2     sliced    -           -           error
c3t1d1s2     sliced    -           -           error
```

5. If multiple paths are used, ensure that VERITAS Dynamic Multipathing (VxDMP) has recognized the device paths.

```
# vxdisk list c1t1d0s2
Device:    c1t1d0s2
devicetag: c1t1d0
type:      sliced
flags:     online error private autoconfig
errno:     Disk is not usable
Multipathing information:
numpaths:   2
c1t1d0s2        state=enabled
c3t1d0s2        state=enabled
```

6. Initialize the LUNs and place them under Volume Manager™ control. In this step, the c1 parameter on the command line instructs VxVM to initialize all disks on that controller. Only one path of a multipathed device should be specified.

```
# vxdiskadd c1

FORMAT MENU:

Add or initialize disks
Menu: VolumeManager/Disk/AddDisks

  Here are the disks selected.  Output format: [Device_Name]

  c1t1d0 c1t1d1

Continue operation? [y,n,q,?] (default: y) y

  You can choose to add these disks to an existing disk group, a
  new disk group, or you can leave these disks available for use
```

```
       by future add or replacement operations. To create a new disk
       group, select a disk group name that does not yet exist. To
       leave the disks available for future use, specify a disk group
       name of "none".

   Which disk group [<group>,none,list,q,?] (default: rootdg) appdg

      There is no active disk group named appdg.

   Create a new group named appdg? [y,n,q,?] (default: y) y

   Use default disk names for these disks? [y,n,q,?] (default: y) y

   Add disks as spare disks for appdg? [y,n,q,?] (default: n) n

   Exclude disks from hot-relocation use? [y,n,q,?] (default: n) n

      A new disk group will be created named appdg and the selected disks
      will be added to the disk group with default disk names.

      c1t1d0 c1t1d1

   Continue with operation? [y,n,q,?] (default: y) y

      Initializing device c1t1d0.

   Use a default private region length for this disk?
   [y,n,q,?] (default: y) y

      Initializing device c1t1d1.

   Use a default private region length for this disk?
   [y,n,q,?] (default: y) y

      Creating a new disk group named appdg containing the disk
      device c1t1d0 with the name appdg01.

      Adding disk device c1t1d1 to disk group appdg with disk
      name appdg02.

   Goodbye.
```

7. Repeat these steps as required on the secondary host(s).

5.3) Volume Manager™ Configuration Requirements for Volume Replicator™

The configuration of volume replication with VVR depends on the configuration of the underlying VxVM disk group. The reader should be aware of several requirements.

At least two LUNs are required in the volume group. The data change map (DCM) is automatically created and used by VVR to record volume changes when the storage replication log (SRL) overflows or is taken offline. Two DCM mirrors are created and cannot exist on the same LUN or physical disk.

As much as possible, the names for the replicated data set (RDS) volume group on the primary and secondary hosts should match. If they do not, additional configuration must be performed, which is outside the scope of this document.

The target disk group on the secondary must be the same size or larger than the source disk group on the primary.

A VxVM disk group can be expanded only by adding additional LUNs or physical disks to it. When using a Network Appliance storage system, it is recommended that a uniform LUN size be selected such that additional increments can be of the same size.

The ability to resize LUNs, presented from Network Appliance filers, cannot be used to expand disks that are initialized by Volume Manager™ without destroying and recreating the volume. Attempting to do so will, at best, result in wasted space. In the worst case, the LUN may be corrupted.

It is strongly recommended that each RDS get its own VxVM disk group if multiple application data sets are replicated from the primary to the secondary.

Since VVR uses VxVM volumes, it is not dependent on the underlying physical disk structure, such as RAID levels and controller types. It is possible to use a different or less robust storage system at the secondary sites to control costs.

6) VERITAS Volume Replicator™ Configuration

This section will discuss the configuration of VERITAS Volume Replicator™ (VVR). Ensure that all requirements outlined are met before proceeding with VVR configuration.

6.1) Replicated Volume Group Configuration

The replicated volume group (RVG) is a set of volumes in a Volume Manager™ disk group that is replicated between the primary and secondary hosts. A storage replication log (SRL) is associated with the RVG and records all data changes. The SRL will also buffer changed data for transmission in case of a loss of contact with the secondary host. A data change map (DCM) is a bit map used to record changes should the SRL overflow. While not as granular as the SRL, this facility avoids having to resynchronize the entire RVG when the connection with the secondary is reestablished. The SRL is a circular log and must be sized appropriately such that the oldest entries are not overwritten before the log wraps. Sizing of the SRL requires detailed knowledge of the data change rate and other factors. Calculation of the proper SRL size is outside the scope of this document.

To create an RVG, follow the steps outlined below:

1. Use the `vxassist` command to create the required volumes on all hosts.

```
# vxassist –g appdg –b make appvol01 10g
# vxassist –g appdg –b make appvol02 10g
```

These commands create two 10GB volumes in the "appdg" disk group.

2. Create the SRL to be associated with the RVG on all hosts.

```
# vxassist –g appdg –b make appsrl 1g
```

This command creates a volume named "appsrl" in the appdg disk group.

3. On the primary host only, create the RVG using the `vradmin` command.

```
# vradmin –g appdg create pri apprvg appvol01 appsrl
```

This command will automatically create DCM mirrors for each data volume in the disk group unless the "–nodcm" switch is specified. DCM overflow protection is the default and is highly recommended.

4. Confirm the configuration of the RVG using the `vxprint` command.

```
# vxprint
Disk group: rootdg

TY NAME          ASSOC       KSTATE    LENGTH    PLOFFS    STATE     TUTIL0
PUTIL0
dg rootdg        rootdg      -         -         -         -         -         -

dm disk01        c0t1d0s2    -         35679840 -         -         -         -

Disk group: appdg

TY NAME          ASSOC       KSTATE    LENGTH    PLOFFS    STATE     TUTIL0
PUTIL0
dg appdg         appdg       -         -         -         -         -         -

dm appdg01       c1t1d0s2    -         41902080 -         -         -         -
dm appdg02       c1t1d1s2    -         41902080 -         -         -         -

rv apprvg        -           ENABLED   -         -         ACTIVE    -         -
v  appvol01      apprvg      ENABLED   20971520 -         ACTIVE    -         -
pl appvol01-01   appvol01    ENABLED   20971520 -         ACTIVE    -         -
sd appdg01-01    appvol01-01 ENABLED   20971520 0         -         -         -
pl appvol01-02   appvol01    ENABLED   LOGONLY  -         ACTIVE    -         -
sd appdg01-02    appvol01-02 ENABLED   64        LOG       -         -         -
pl appvol01-03   appvol01    ENABLED   LOGONLY  -         ACTIVE    -         -
sd appdg02-02    appvol01-03 ENABLED   64        LOG       -         -         -
v  appvol02      apprvg      ENABLED   20971520 -         ACTIVE    -         -
pl appvol02-01   appvol02    ENABLED   20971520 -         ACTIVE    -         -
sd appdg02-01    appvol02-01 ENABLED   20971520 0         -         -         -
pl appvol02-02   appvol02    ENABLED   LOGONLY  -         ACTIVE    -         -
sd appdg01-02    appvol02-02 ENABLED   64        LOG       -         -         -
```

Network Appliance® Inc.

```
pl appvol02-03  appvol02     ENABLED  LOGONLY  -       ACTIVE  -       -
sd appdg02-02   appvol02-03  ENABLED  64       LOG     -       -       -
v  appsrl       apprvg       ENABLED  2097152  SRL     ACTIVE  -       -
pl appsrl-01    appsrl       ENABLED  2097152  -       ACTIVE  -       -
sd appdg02-01   appsrl-01    ENABLED  2097152  0       -       -       -
```

5. On the primary host, get the disk group ID to place in the `/etc/vx/vras/.rdg` file on all secondary hosts. Optionally, a "+" may be used in the file to indicate that any disk groups can be replication targets.

    ```
    # vxprint –l appdg
    952738334.1025.sybok
    ```

    In this example, sybok is the name of the primary replication host.

6. On the secondary host, start "vradmind".

    ```
    # vradmind
    ```

7. On the primary host, use the `vradmin` command to designate secondary hosts for the RVG.

    ```
    # vradmin –g appdg addsec apprvg sybok sarak
    ```

    In this example, sarak is the name of the secondary replication host.

8. Start replication from the primary to the secondary.

    ```
    # vradmin –g appdg –a startrep sybok sarak
    Message from Primary:
    vxvm:vxrlink: WARNING: Attaching rlink to non-empty rvg. Autosync will be
    performed.
    vxvm:vxrlink: INFO: Secondary data volumes detected with rvg apprvg as parent:
    vxvm:vxrlink: INFO: appvol01:     len=20971520 primary_datavol=appvol01
    vxvm:vxrlink: INFO: Autosync operation has started
    ```

Once the synchronization is complete, changed data blocks will automatically be replicated from the primary host to the secondary host(s).

6.2) First-Time Synchronization

It is possible that the infrastructure and application requirements are such that special methods may be employed to perform the initial synchronization before deploying the storage. These instances may include:

A large volume of data and relatively small available replication bandwidth.

Availability or performance requirements that cause a resource contention issue at the primary site.

Storage management teams structured such that detailed interaction is not available, such as using an outsourced secondary provider.

In these cases, it may be necessary to use special methods to perform the first-time synchronization of the storage to avoid these issues. These methods include:

> Colocating the hosts and storage and using a high-bandwidth interconnect to perform the initial synchronization.

> Using a block-level backup and restore mechanism to make an identical copy of a secondary. This media can then be sent to the secondary location, restored, and synchronized.

Each of these methods will require that the SRL be used to buffer changed data until the secondary can be brought online. A larger SRL may be needed during this time and resized once the secondary is fully deployed.

6.3) Changing the Synchronization Mode

VVR has the option of designating synchronous operation for replication. When synchronous replication is used, writes to volumes in the RVG are not acknowledged back to the application until writes are confirmed on both the primary and all secondary hosts. This guarantees that there will be no data loss for operations that were completed on these volumes should the primary host go offline. Application performance can be seriously impacted by the latency induced by synchronous replication. It is recommended that all parties concerned with performance of the system and business requirements thoroughly examine the impact of synchronous replication before enabling it.

To enable synchronous replication, the following command is used:

```
# vradmin –g appdg set apprvg sarak synchronous=override
```

To disable synchronous replication, the same command is used with different arguments:

```
# vradmin –g appdg set apprvg sarak synchronous=off
```

6.4) Additional VERITAS Volume Replicator™ Configuration

VVR has an extensive array of options that are beyond the scope of this report. These options are covered in detail in the documentation provided with VxVM and VVR. References to these documents are provided at the end of this paper.

6.5) Feature Comparison: Network Appliance® SnapMirror® and VERITAS Volume Replicator™

Network Appliance SnapMirror and products enable volume-level replication between two Network Appliance storage systems over networks running the Internet protocol. This section is intended to highlight some of the features of SnapMirror and compare them to VVR's feature set.

Criteria	Network Appliance SnapMirror	Network Appliance SyncMirror™*	VERITAS Volume Replicator™
Replication Type	Asynchronous BatchUpdate (minimum 1 minute)	Synchronous	Asynchronous or Synchronous
Host Dependency	Host Independent	Host Independent	Host Dependent
Licensing	Per Filer	Per Filer	Per Server (requires VxVM and VVR license)

Transport	TCP/IP	Fibre Channel *	TCP/IP
Distance Limitation	Unlimited	500 Meters 30 kM (MetroCluster)	Unlimited
Environments	Homogenous NetApp	Homogenous NetApp	Heterogeneous

Table 6-1: Feature Comparison

* Network Appliance will be releasing an IP-based synchronous mirror product in the near future.

7) References

VERITAS Volume Manager™ 3.5 Installation Guide

VERITAS Volume Manager™ 3.5 Administrator's Guide

VERITAS Volume Replicator™ 3.5 Installation Guide

VERITAS Volume Replicator™ 3.5 Administrator's Guide

Network Appliance product documentation can be found at http://now.netapp.com/ (login required).

8) Conclusions

Network Appliance storage systems using block access transports integrate readily with the VERITAS Volume Manager™ and Volume Replicator™ products. The online and nearline storage arrays offered by Network Appliance provide cost-effective and scalable business continuance solutions. The NearStore™ product line is particularly well suited for the role as the replication target.

9) Caveats

Performance of the replication is dependent on many factors, the most influential being the bandwidth of the network connection between the two hosts. The SRL size is related to the data change rate and the required bandwidth to support that rate. Sizing the SRL requires specific knowledge of the application and infrastructure characteristics. A large amount of research and planning is required to size the SRL to avoid replication problems. Consultation with VERITAS engineers may be required to ensure a successful deployment.

NetApp has not tested this procedure with all of the combinations of hardware and software options available. There may be significant differences in your configuration that will alter the procedures necessary to accomplish the objectives outlined in this paper. If you find that any of these procedures do not work in your environment, please contact the author immediately.