EMC²

VDM METROSYNC FOR VNX2

A Detailed Review

ABSTRACT

This white paper introduces the architecture and functionality of the VDM MetroSync feature for VNX2. It also discusses VDM MetroSync Manager for automated failover and the ability to preserve asynchronous replication sessions after a reverse or failover.

July, 2018

To learn more about how EMC products, services, and solutions can help solve your business and IT challenges, contact your local representative or authorized reseller, visit www.emc.com, or explore and compare products in the EMC Store

EMC believes the information in this publication is accurate as of its publication date. The information is subject to change without notice.

The information in this publication is provided "as is." EMC Corporation makes no representations or warranties of any kind with respect to the information in this publication, and specifically disclaims implied warranties of merchantability or fitness for a particular purpose.

Use, copying, and distribution of any EMC software described in this publication requires an applicable software license.

For the most up-to-date listing of EMC product names, see EMC Corporation Trademarks on EMC.com.

Part Number H14695.4

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Being able to access data is a critical component in the daily operation and function of many organizations. Implementing a replication solution enables data centers to avoid disruptions in business operations by providing a disaster recovery (DR) plan and additional redundancy.

The demand for continuous data availability is higher than ever before. IT organizations are seeking synchronous replication solutions that provide zero data loss in the event the primary site becomes unavailable. They must be able to recover from a disaster quickly and efficiently, in order to bring their business back online as soon as possible. They expect minimal downtime by automating failover to the secondary site when critical issues are detected.

This white paper provides a comprehensive overview of VDM MetroSync, VDM MetroSync Manager, and Preserve Replicator Sessions, which are designed to meet these requirements.

AUDIENCE

This white paper is intended for IT planners, storage architects, administrators, partners, EMC employees and any others involved in evaluating, acquiring, managing, operating, or designing an EMC VDM MetroSync environment using VNX2 storage systems.

TERMINOLOGY

Bandwidth – The amount of data that can be transferred in a given period of time. Bandwidth is usually represented in bytes per second (Bps) or MB/s.

Common Internet File System (CIFS) – An access protocol that allows data access from Windows/Linux hosts located on a network. Also known as Server Message Block (SMB).

Consistency Group – A set of LUNs that are grouped together and managed as a single entity to ensure write-order consistency.

Control Station – Hardware that provides management functions to the VNX File components of the system.

Data Mover – A Data Mover is a component that runs its own operating system. It retrieves data from a storage device and makes it available to a network client by using a NAS protocol.

Fibre Channel Protocol – Transfer protocol used to communicate Internet Protocol (IP) and Small Computer Systems Interface (SCSI) commands over a Fibre Channel network.

iSCSI Protocol – The iSCSI (Internet Small Computer System Interface) protocol provides a mechanism for accessing block-level data storage over network connections. The iSCSI protocol is based on a network-standard client/server model with iSCSI initiators (hosts) acting as storage clients and iSCSI targets acting as storage servers.

MirrorView/S – A synchronous replication product that mirrors data in real time between local and remote storage systems.

NAS Pool – A collection of Block disk volumes organized in a logical grouping for use on VNX File.

Network Attached Storage (NAS) – File-based storage for a wide range of clients and applications that access storage over IP connectivity.

Network File System (NFS) – An access protocol that allows data access from Linux/UNIX hosts located on a network.

Primary Image – The LUN that contains production data and the contents of which are replicated to the secondary image.

Recovery Point Objective (RPO) – RPO is a defined period of time in which data can be lost but still allow an organization to continue operations. For example, if an organization determined that it could handle an RPO of 30 minutes, the business would be able to experience a disaster, lose 30 minutes of data, and still be able to perform operations normally.

Recovery Time Objective (RTO) – RTO is the duration of time within which a business process must be restored after a disaster. For example, an RTO of 1 hour means that in case of a disaster, the data and business process needs to be restored in 1 hour.

Replicator - An IP-based asynchronous replication feature that is capable of replicating File Systems and Virtual Data Movers (VDMs). Also known as RepV2.

Round Trip Time (RTT) – RTT is the length of time it takes for a signal to be sent.

Secondary Image – A LUN that contains a mirror of the primary image LUN. Also referred to as the secondary. This LUN must reside on a different storage system than the primary image.

SnapView Clone – VNX Block feature that allows you to take a full copy of a source LUN.

Synchronous Replication – A replication mode in which the host initiates a write to the system at the local site. The data must be successfully stored in both the local and remote sites before an acknowledgement is sent back to the host.

Standby Data Mover – A Data Mover held in a reserved state, waiting to assume the state and roles of a failed active partner.

Throughput – The rate at which data is transmitted in a given amount of time and is usually represented in IOPs.

Unisphere – A web-based EMC management interface for creating storage resources, and configuring and scheduling protection for stored data. Unisphere is also used for managing and monitoring other storage operations.

Virtual Data Mover (VDM) – A logical container that holds the data needed to support one or more CIFS servers, NFS servers, or both CIFS and NFS servers, and their file systems. Each VDM has access only to the file systems mounted to that VDM, providing a logical isolation between physical Data Movers and other VDMs on the VNX system.

INTRODUCTION

VDM MetroSync is a Disaster Recovery (DR) solution for VNX2 File which leverages a MirrorView/S replication session to create a zero data loss replication solution at a Virtual Data Mover (VDM) granularity. It allows for replication of a VDM along with all of its contents including file systems, checkpoints, checkpoint schedules, CIFS servers, exports, interfaces, and so on. It can be configured in either an active/passive configuration where the active VDMs are constrained to one site, or an active/active configuration where each site has its own set of active VDMs. VDMs can be moved or failed over from one system to another as needed.

VDM MetroSync Manager is optional software that can be installed on a Windows server which works with VDM MetroSync. It provides a GUI interface to display VDM MetroSync session information and run operations to move, failover, or restore VDMs. It also has the ability to continuously monitor sessions and automatically initiate failover when issues are detected.

With synchronous replication enabled between two systems, it is also possible to add asynchronous replication to a third system by using Replicator. This allows the third system to be located further away and enables it to be used as a backup and recovery solution. When VDMs are moved or failed over between the VDM MetroSync systems, the Replicator sessions to the third system are preserved. Since the Replicator checkpoints are replicated along with the VDM, a common base checkpoint is available which removes the requirement for a full synchronization after failover. The Replicator sessions can be incrementally updated and restarted on the new system where the VDM is active.

Compared to the cabinet-level DR solution, VDM MetroSync provides several advantages:

- Finer granularity - Allows for the failover of a single VDM as opposed to the entire system.
- Failover times - Quicker failover times because the entire cabinet does not need to be failed over.
- More efficient Data Mover usage - No idle local standby Data Movers are required.
- Automated failover - Provides automatic failover capabilities by using VDM MetroSync Manager.

Some of the typical use cases for this feature include:

- Disaster Recovery (DR):
    - Power outages
    - Network outages
    - Human error (accidental reboot, cable pull, and so on)
    - Environmental (flood, storm, fire, and so on)
- Data Mobility
    - Maintenance
    - Load balancing
    - Upgrades

VDM METROSYNC CONFIGURATION

REQUIREMENTS

In order to leverage the VDM MetroSync feature, the following requirements must be met:

- Two VNX2 systems
    - VNX OE for File 8.1.9.155 or later and Block 05.33.009.5.155 or later
- MirrorView/S and SnapView Clone enablers installed

- All file systems on the VDM must be Split-Log
- ICMP and port 443 for communication between systems
- NTP to ensure Control Stations and Data Movers times are in sync

The two VNX2 systems can be any model and they do not need to be the same. The disk configuration on both systems also do not need to be the same as long as enough capacity is available to store the replicated data. If performance and scalability is important, both systems can be configured have matching configurations.

VDM MetroSync only replicates the data that is inside of the VDM. Any features or functionality that operate at the physical Data Mover or cabinet level are not included. This includes services such as DNS, NIS, NTP, local password/group files, Usermapper client, FTP/SFTP, LDAP, HTTP, CEPP, CAVA, server parameters, netgroups, nsswitch, hosts, CIFS service, and so on. These services can be migrated by using the `migrate_system_conf` command. This command should be run prior to configuring VDM MetroSync and again any time the configuration changes. Note that this command does not migrate the routing table, including the default route, so that should be configured after an interface has been created on the target system.

CONFIGURATION

In this section, we will build out a configuration to review the technical details of each component. This feature requires Unisphere and VNX File CLI. A high-level summary of the steps required are:

1. Establish a control path between the Control Stations on both systems.
2. Enable MirrorView connection.
3. Configure and allocate Write Intent Logs (WILs) and Clone Private LUNs (CPLs) on both systems.
4. Enable the VDM MetroSync Service.
5. Provision storage for VNX File and configure a user-defined NAS Pool on both systems.
6. Create a VDM using the `syncreplicable` option.
7. Create a VDM MetroSync replication session.
8. Create VNX File resources.

More details about the configuration requirements and procedure can be found in the *Using VDM MetroSync with VDM MetroSync Manager for Disaster Recovery* document.

First, let us start with two independent VNX2 systems that have not been configured for VDM MetroSync yet, as shown in Figure 1. Note that each system has its own NAS DB (NAS Database) which holds information about its own NAS configuration, such as VDMs, file systems, checkpoints, interfaces, disks, and so on.

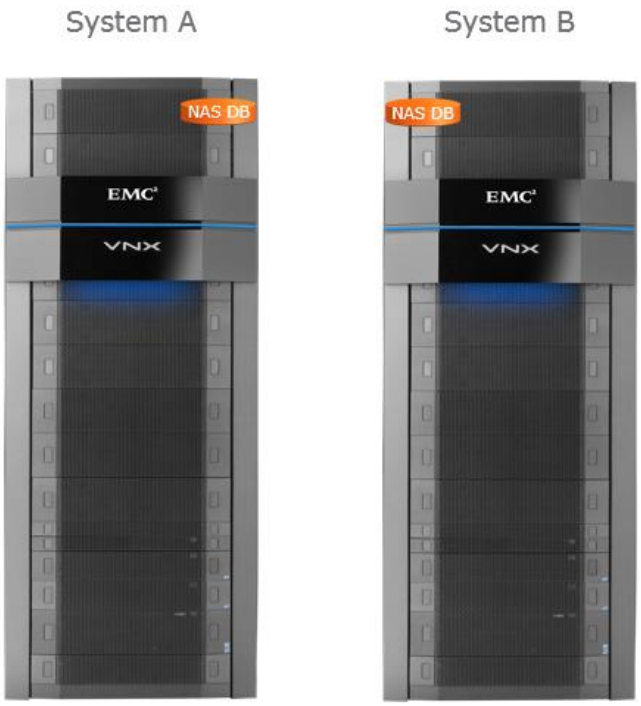

Figure 1. Two Independent VNX2 Systems

ESTABLISH CONTROL PATH

In order to start using VDM MetroSync, there needs to be a control path between the Control Stations of both systems. This allows the two systems to establish a secure communication channel to issue commands to the other system.

If a system currently is, or has plans to be, a dual Control Station (CS) configuration, you must first create a Control Station IP alias by using the `nas_config` command. This is an IP address that will always be active on the primary CS, regardless if it is running on CS0 or CS1. Without this, management of the VDM MetroSync sessions will become degraded if the CS fails over to CS1. For a single CS configuration, this step is optional. If this is skipped and a secondary CS is added later, the IP address of CS0 must be used for the alias and a new IP address must be assigned to CS0.

Run the `nas_cel` command to create the relationship to the peer system's CS. For a dual CS configuration, the alias IP must be used in the command. Also, a passphrase needs to be specified and must match on both systems. This command must be run on both systems to allow for communication in both directions. Ensure port 443 is open and the time on each Control Station is within 10 minutes of the other.

ENABLE MIRRORVIEW

The first Fibre Channel and iSCSI port on each SP are the designated MirrorView (MV) ports. You can establish the MirrorView connection using either Fibre Channel or iSCSI over IP. This port must be connected to the corresponding MV ports on the peer system. Note that SPA must be connected to SPA on the peer system while SPB must be connected to SPB on the peer system. Once these connections are in place, check Unisphere's **Hosts → Initiators** page to ensure that the MV ports display as Registered and Logged In.

Both VNX systems must also be added to the same Unisphere Domain. This can be done from the System List page in Unisphere. When adding a new system to the domain, input an SP IP address along with the sysadmin credentials for that system. Confirm the new system now displays in the System List.

After the MV ports are connected and both systems have been added to the Unisphere Domain, the MV connection can be enabled. This can be done in the Data Protection tab in Unisphere. Once the connection has been enabled, MV is activated and can be used to replicate data, as shown in Figure 2.

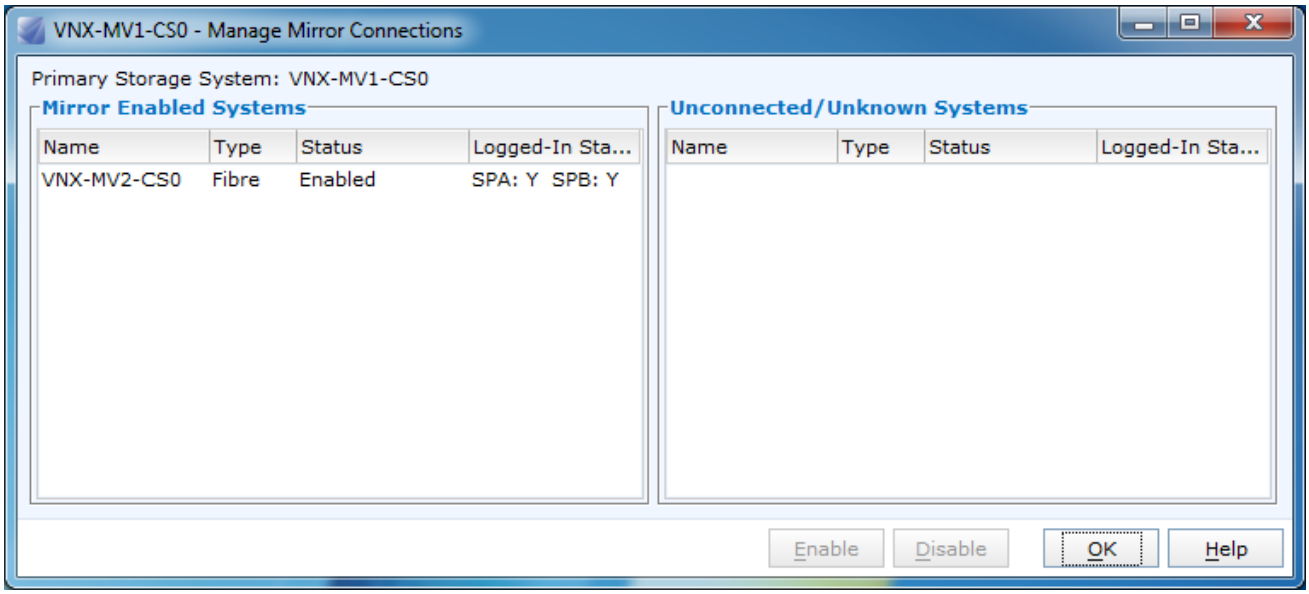

Figure 2. MirrorView Connection

CONFIGURE WRITE INTENT LOGS & CLONE PRIVATE LUNS

The VDM MetroSync feature requires enabling Write Intent Logs (WILs) for MirrorView and Clone Private LUNs (CPLs) for SnapView Clone on each array.

The WIL tracks in-flight writes to both the primary and secondary images in a mirror relationship. It is a bitmap that is composed of extents indicating where data is written. The WIL consists of two LUNs, one assigned to each SP, and each LUN services all the mirrors owned by that SP that have the WIL enabled. LUNs that will be used for the WIL must be 128 MB in size and created on RAID Group-based LUNs. The same RAID Group can also be used for CPLs and the NAS DB mirrors. FAST Cache should not be enabled on WILs. Once these LUNs are created, they can be allocated as WILs in the Data Protection tab in Unisphere.

CPLs record information that identifies data chunks that are modified if a clone is fractured. The log records the location of the changed data chunks, but no actual data is written to the CPL. Using CPLs reduces the time needed to synchronize or reverse-synchronize a clone and its source LUN because the software copies only modified chunks. CPLs must also be created on RAID Group-based LUNs that are 2 GB in size. FAST Cache should not be enabled on CPLs. Once these LUNs are created, they can also be allocated as CPLs in the Data Protection tab in Unisphere.

ENABLE VDM METROSYNC SERVICE

To begin using VDM MetroSync, the VDM MetroSync service must first be enabled by using the `nas_cel -syncrep` command. This is a new option that was added for this feature. The command syntax to enable the service is:

```
nas_cel -syncrep -enable {<cel_name>|id=<cel_id>}
-local_fsidrange <from>,<to> -remote_fsidrange <from>,<to>
[-local_storage {raid_group=<rg_id> | block_pool=<pool_id>}]
[-remote_storage {raid_group=<rg_id> | block_pool=<pool_id>}]
```

This is an example of this command:

```
nas_cel -syncrep -enable id=1 -local_fsidrange 5000,18000 -remote_fsidrange 18001,31000 -local_storage
raid_group=0 -remote_storage raid_group=0
Now doing precondition check...                                  done
Now saving FSID range [18001,31000] on remote system...          done
Now saving FSID range [5000,18000] on local system...            done
Now adding remote storage info to local system...                done
Now creating sync replication mirror LUNs on local system...    done
Now creating sync replication mirror LUNs on remote system...   done
```

```
Now creating Mirrors and Clones (may take several minutes)...      done
Now waiting for Mirrors and Clones to finish initial copy...       done
Now adding NBS access to local server server_2...                  done
Now adding NBS access to local server server_3...                  done
Now creating mountpoint for sync replica of NAS database...        done
Now mounting sync replica of NAS database...                       done
Now configuring and rebooting secondary CS...                      done
Now enabling sync replication service on remote system...          done
done
```

During this process, the system configures the following settings on both systems:

FSID (FILE SYSTEM ID) RANGES

Configures an FSID range to be used on System A and a different range for System B. This helps ensure there are no FSID conflicts when VDMs are reversed or failed over. These ranges are user configurable and are specified in the command to enable the service. The valid values are 1-32767.

By default, a minimum of 8192 FSIDs for each system is required to minimize the chance of running out of FSIDs. For advanced users who would like to configure a smaller range for each system, add an entry for `fsidrange:<range>:` to `/nas/site/nas_param`. For example, if you would like to configure a system with 5000 FSIDs reserved, add `fsidrange:5000:` to `/nas/site/nas_param`. If you are unsure, keep the default of 8192.

If your systems already have resources created, prior to configuring FSID ranges for each system, confirm which FSIDs are already in use. For VDMs already under VDM MetroSync sessions on the system, use the `nas_checkup` command to identify any potential FSID conflicts.

NAS DB MIRRORS (LUN 8)

The local NAS DB on each system needs to be synchronously replicated to the peer system. This ensures the peer system's VDM configuration is available when a reverse or failover is initiated. This mirror is automatically created by the system on the Pool or RAID Group that is specified in the `-local_storage`/`-remote_storage` switches. The initial synchronization is also started automatically. Note that if these LUNs already exist, the `-local_storage`/`-remote_storage` switches can be omitted.

NAS DB MIRROR CLONES (LUN 9)

The system also automatically creates a clone of LUN 8 on the Pool or RAID Group that is specified in the `-local_storage`/`-remote_storage` switches. This is required because the data on LUN 8 cannot be read while it is synchronizing. LUN 9 synchronizes with LUN 8 until a reverse or failover operation occurs. At that point, LUN 9 fractures from LUN 8, and the system reads the necessary VDM configuration from the peer system. Once the VDM configuration is read, LUN 9 returns to a synchronizing state with LUN 8.

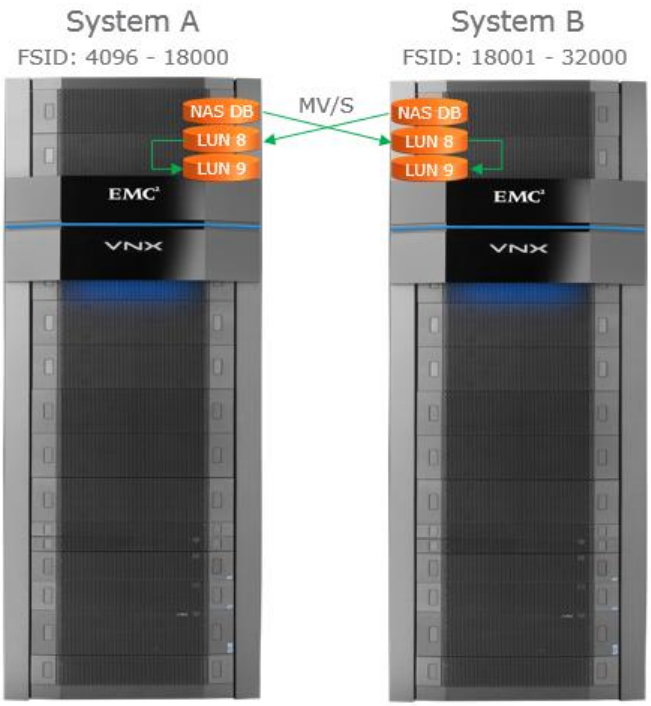

Figure 3. VDM Synchronous Replication Service Enabled

Figure 3 shows a pair of systems that have the Synchronous Replication service enabled. It includes the FSID ranges that are set, the NAS DB Mirror (LUN 8), and the NAS DB Mirrors Clone (LUN 9) on each system.

PROVISION STORAGE RESOURCES

When provisioning LUNs for File, follow the recommendations referenced in the *VNX2 Unified Best Practices for Performance* white paper to achieve the best performance and availability. The LUNs must be added to the ~filestorage Storage Group and then a rescan must be run in order to make them available as disk volumes for VNX File. After provisioning LUNs on the source system, the same size and number of LUNs must also be provisioned to VNX File on the peer system. In order to achieve the same performance, you should build the LUNs using the same disk and RAID types on both systems.

VDM MetroSync requires the use of user-defined NAS Pools. When LUNs are added to the ~filestorage Storage Group and a rescan is performed, they become unused disk volumes. A user-defined NAS Pool can be created by running the `nas_pool` command and adding one or more unused disk volumes. Note that each user-defined NAS Pool can only support a single synchronously replicated VDM. This feature requires the VDM and all of its resources (file systems, checkpoints, and so on) to be stored on disk volumes that are on a single user-defined NAS Pool. The NAS Pool should be sized appropriately to store all of the VDM's resources. Once a VDM MetroSync session has been started, the NAS Pool can be expanded but not shrunk without requiring a full synchronization.

On the source system, a single VDM can be created from the user-defined NAS Pool by using the `syncreplicable=yes` option in the `nas_server` command. An existing VDM can also have the `syncreplicable` option enabled or disabled by running the `nas_server` command. This option makes the VDM eligible to have a VDM MetroSync session created on it. Only one `syncreplicable` VDM is allowed for each user-defined NAS pool. Once the VDM is available, file systems, checkpoints, schedules, CIFS servers, interfaces, exports, and other resources that reside on the VDM can also be created normally. As these resources are being provisioned, the synchronously replicated NAS DB is being updated with the latest VDM configuration.

File systems that are created on the user-defined NAS Pool can only be mounted to the VDM. Note that file systems must be of the split-log type, which is the default setting starting with VNX2 File OE 8.1.6.96. You can check the log type of a file system or VDM by using the `nas_fs` command. If `log_type=split` is not displayed for the file system, a host-based migration can be used to migrate the data to a new split-log file system.

Note that unlike the cabinet-level DR solution, VDM MetroSync does not require any dedicated local or remote standby Data Movers. VDMs and their associated resources can be created on any Data Mover and moved to the other system on an as-needed basis. Ensure each pair of source and destination Data Movers have enough resources to host all of the VDMs in case of failover.

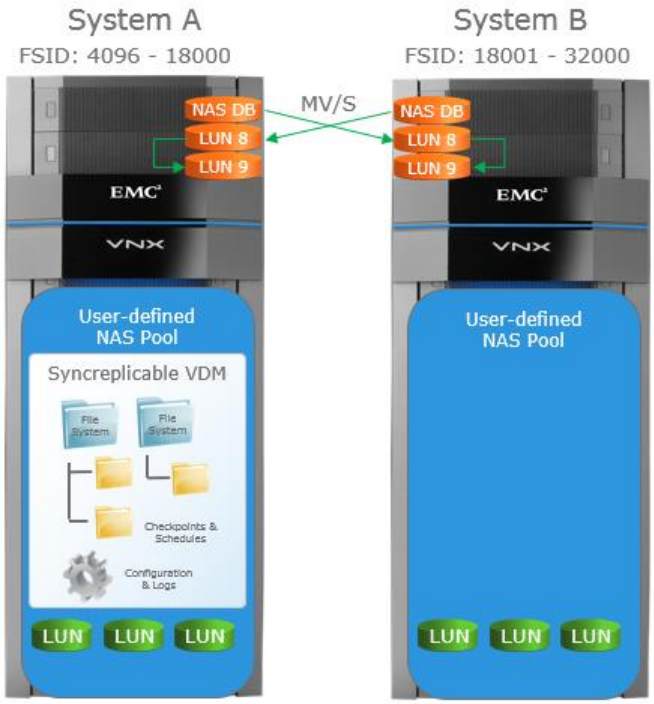

Figure 4. Storage and VDM Resources Provisioned

Figure 4 shows a pair of systems that have the Synchronous Replication Service enabled and all of the storage and resources provisioned. This includes the LUNs, user-defined NAS Pool, `syncreplicable` VDM, file systems, checkpoints, and schedules.

CONFIGURE CLIENT TRANSPARENCY

Client transparency during reverse or failover operations can be achieved if the outage does not exceed the timeout of the application. The clients and arrays must also be properly configured in order for this to work.

Ensuring both systems have the same configuration is crucial when trying to maintain client transparency. It is important to note that VDM MetroSync only replicates the data that is inside of the VDM. Any features or functionality that operate at the physical Data Mover or cabinet level are not included. This includes services such as DNS, NIS, NTP, local password/group files, Usermapper client, FTP/SFTP, LDAP, HTTP, CEPP, CAVA, server parameters, netgroups, nsswitch, hosts, CIFS, and so on. These services can be migrated by using the `migrate_system_conf` command. This command should be run prior to configuring VDM MetroSync and again any time the configuration changes. Note that this command does not migrate the routing table, including the default route, so that should be configured after an interface has been created on the target system.

Attaching an interface to the VDM provides the ability for the interface to follow the VDM during reverse or failover operations. For interfaces that are already attached when the VDM MetroSync session is created, the system automatically creates identical interfaces on the peer system in a *down* state. When a reverse or failover operation occurs, the system automatically brings up the interface on the peer system. Note that for any interfaces that are attached after a VDM MetroSync session has already been created, the user must manually create an identical interface on the peer system in a *down* state by using the `server_ifconfig` command. If you are using different IP addresses at each site, create an interface on the peer system using the same interface name but the new IP address. Once the interfaces are created on the target system, ensure the default gateway is also added since it is not included in the `migrate_system_conf` command.

CIFS

For CIFS, the SMB3 Continuous Availability (CA) feature allows Windows clients to persistently access CIFS shares without the loss of the session state. You can leverage this feature when planning storage availability for business-critical applications such as Microsoft SQL Server, IIS, and Hyper-V. In order to leverage SMB3 CA, the file system must be mounted by using the server_mount command with the `smbca` mount option. It must also be exported by using the server_export command with the `type=CA` option.

If necessary, the CIFS parameter `smb2.maxCaTimeout` can be modified. This parameter has a default value of 360 seconds but can be configured between 0 to 600 seconds. This setting should be configured to match the timeout of the application.

Using the `server_cifs` command, ensure that the CIFS service is started and the I18N mode matches on both systems. If the service has not been started, use the `server_setup` command to start the CIFS service on both systems. If the I18N mode does not match, update the internationalization mode on one or both sites so they are the same. Enabling Unicode on both systems is recommended.

NFS

Ensure the interface that is used for the NFS mount is attached to the VDM. This allows NFS access to the file systems on the VDM and also provides the ability for the interface to follow the VDM during reverse or failover operations. For clients, use the hard mount option. Since the FSID remains unchanged, the client should not see any stale file handles.

VDM METROSYNC OPERATIONS

The new `nas_syncrep` command can be used to manage the VDM MetroSync sessions. The syntax for this command is:

```
$ nas_syncrep
          -list
        | -info { -all | <name> | id=<id> } [-verbose]
        | -create <name>
             -vdm <vdm_name>
             -remote_system <cel_name>
             -remote_pool <pool_name>
             -remote_mover <mover_name>
             [-network_devices <local_device_name>:<remote_device_name>[,...]]
        | -start { -all | <name> | id=<id> }
        | -delete { <name> | id=<id> }
        | -reverse { <name> | id=<id> }
        | -failover { <name> | id=<id> }
        | -Clean { -all | <name> | id=<id> }
```

Table 1 lists these commands and their descriptions.

Table 1. VDM MetroSync Commands

Command	Description
list	Lists the name and status of the VDM MetroSync sessions on the system.
info	Provides details about the specified VDM MetroSync sessions.
create	Creates a new VDM MetroSync session.
start	Restarts a VDM MetroSync session that has been stopped.
delete	Deletes a VDM MetroSync session.
reverse	Gracefully moves a VDM to the other system and reverses the direction of replication. This is the preferred method to move a VDM and should be used any time the source system is still available.
failover	Forcefully fails over the VDM to the other system. This command should only be used in the event that the source VNX2 system is not reachable.
Clean	The Clean command is related to failover. This is run on the original source system to clean up residual objects in the NAS DB that were left over from the failover operation.

CREATE

At this point, a VDM MetroSync session can be created for each VDM that will be replicated by using the `nas_syncrep` command. The syntax to create a new VDM MetroSync session is:

```
nas_syncrep -create <name> -vdm <vdm_name>
-remote_system <cel_name> -remote_pool <pool_name> -remote_mover <mover_name>
[-network_devices <local_device_name>:<remote_device_name>[,...]]
```

This is an example of this command:

```
nas_syncrep -create VDM1_syncrep -vdm VDM1 -remote_system VNX-2 -remote_pool VDM1_Pool -remote_mover server_2 -
network_devices cxg-1-0:fxg-1-0
Now validating params...                                    done
Now marking remote pool as standby pool...                  done
Now creating CG...                                          done
Now creating remote network interface(s)...                 done
Now updating local disk type...                             done
Now updating remote disk type...                            done
Now generating session entry...                             done
done
```

Running this command:

- Creates a VDM MetroSync session for the VDM between the local and remote system.
- Designates the remote user-defined NAS Pool as the standby pool.
- Creates the MirrorView Consistency Group (CG).
- Adds all the LUNs from the user-defined NAS Pool in to the MirrorView CG.
- For any attached interfaces, creates identical interfaces on the remote system by using the network device mapping.
- Starts the initial synchronization of the LUNs.

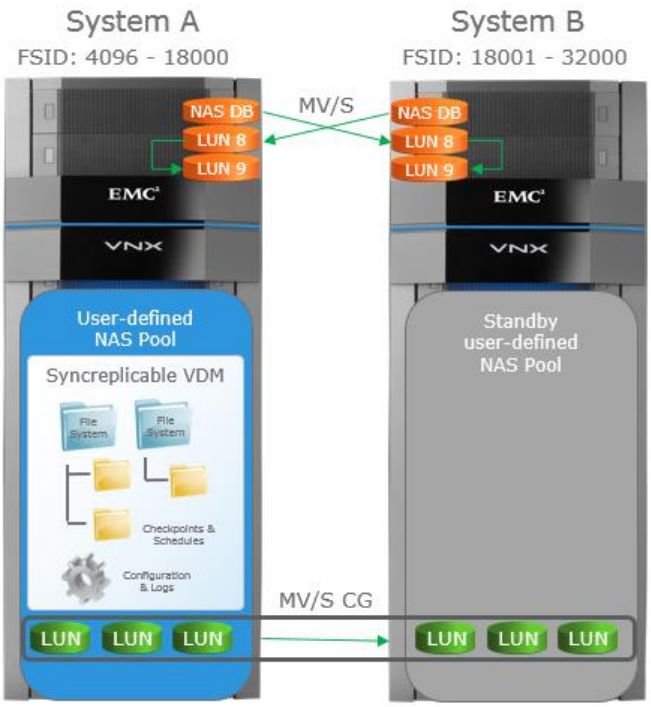

Figure 5. VDM Synchronous Replication Session

Figure 5 shows a pair of systems that has a VDM Synchronous Replication session created. You can check the status of the initial synchronization by running `the nas_syncrep -info -verbose` command. Once the initial synchronization is complete, the VDM is protected.

REVERSE

The reverse operation is used to gracefully move a VDM to the other system and reverse the direction of replication. This is the preferred method to move a VDM and should be used any time the source system is still available. After running this command, there will be a period of data unavailability until the VDM is brought online on the other system. Reverse can be used for situations such as load balancing, maintenance, or upgrades.

Prior to running the reverse command, use the `nas_syncrep` command to confirm the session has a status of `in_sync`. Since this is a pull operation, this command must be executed on the destination system. After running this command, there will be a period of data unavailability until the VDM is brought online on the other system. It is crucial to confirm that the configuration matches on both systems prior to running this command. Refer to the "Configure Client Transparency" section for more details.

This is an example of this command:

```
nas_syncrep -reverse VDM1_Session

WARNING: You have just issued the nas_syncrep -reverse command. There will be a period of Data Unavailability
during the reverse operation. After the reverse operation, the VDM/FS(s)/checkpoint(s) protected by the sync
replication session will be reversed to the local site. Are you sure you want to proceed?  [yes or no] yes
Now doing precondition check...                          done: 7  s
Now doing health check...                                done: 11 s
Now cleaning local...                                    done: 1  s

                                 Service outage start......
Now turning down remote network interface(s)...          done: 3  s
Now unloading remote VDM/FS(s)/checkpoint(s)...          done: 3  s
Now switching the session (may take several minutes)...  done: 13 s
Now importing sync replica of NAS database...            done: 9  s
Now creating VDM...                                      done: 6  s
Now importing VDM settings...                            done: 0  s
Now mounting exported FS(s)/checkpoint(s)...             done: 4  s
```

```
Now loading VDM...                                      done: 8  s
Now turning up local network interface(s)...            done: 2  s
                                  Service outage end: 48 s

Now mounting unexported FS(s)/checkpoint(s)...          done: 0  s
Now importing schedule(s)...                            done: 0  s
Now unloading remote VDM/FS(s)/checkpoint(s)...         done: 15 s
Now cleaning remote...                                  done: 8  s
                                       Elapsed time: 90 s
done
```

The reverse command:

- Disables the source interfaces on the source system.
- Unmounts the file systems and checkpoints, and unloads the VDM.
- Promotes the MirrorView Consistency Group.
- Reads the configuration from the NAS DB.
- Rebuilds all the objects for the VDM.
- Mounts the exported file systems and checkpoints, and loads the VDM.
- Enables the destination network interfaces.
- Mounts unexported file systems and checkpoints.
- Rebuilds the checkpoint schedules.

This operation gracefully unloads the VDM and its associated resources from the source system and brings them online on the destination system. The time needed to complete this command depends on the number of resources, such as file systems and checkpoints, on the VDM. While the command is running, a detailed output of the operation and timing is printed to the screen. The client outage begins when the network interfaces are brought down and ends once the interfaces are brought up on the other system. The total time between these two events is measured and printed as the service outage. To minimize the service outage time, resources that are in use are prioritized and brought online first. Non-critical resources such as unexported, temporarily unmounted, temporarily unloaded objects and checkpoint schedules are brought online afterwards. Once the command completes, the elapsed time it took for the command to run is printed.

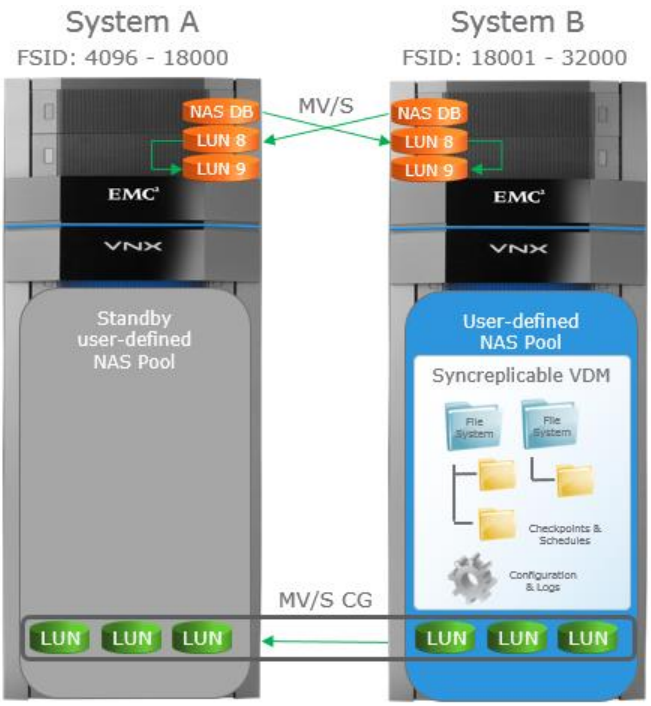

Figure 6. VDM MetroSync Reverse

Figure 6 shows the VDM MetroSync session status after a reverse operation is complete. System B becomes the active system for the VDM and clients can access data from it. The direction of replication now runs in the opposite direction and System A becomes the standby. If you would like to move the VDM back to System A, execute a reverse command on System A.

Since the interface is brought down as part of the reverse process, reversing the last VDM on a Data Mover might remove all available interfaces from that Data Mover. This prevents the Data Mover from communicating with the environment such as NTP servers, DNS servers, and so on. It is recommended to leave at least one interface available on each physical Data Mover at all times on both systems. This can be accomplished by adding an interface to each primary Data Mover on both sites but leaving it unattached to any VDMs. The additional interface should be on the same subnet as the one that is on the VDMs.

FAILOVER

The failover operation forcefully fails over the VDM to the other system. Since this is not a graceful procedure, this command should only be used in the event that the source VNX2 system is not accessible. If there is no failure that results in the source system being unavailable, a reverse must be issued instead. Failover can be used for situations such as power outages, network outages, or environmental issues such as a fire.

To the

This command assumes the source site is unavailable and should only be used in those situations. Prior to running the failover command, confirm whether the source site is truly unavailable to clients. Depending on the type of outage, the VDM MetroSync session has a status of either `in_sync` or `stopped`. If the source site becomes completely unavailable, the VDM MetroSync session status changes to `stopped` since replication is unavailable. However, if the source site is only partially down, such as a VLAN outage, the MirrorView link can continue to replicate and the VDM MetroSync session continues to be `in_sync`.

If a failover is initiated while the source system is still accessible, its backend LUNs are changed to read-only mode. The VDM configuration remains in the source system's NAS DB and it cannot be gracefully cleaned up. However, the failover process sends a command to attempt to turn down any interfaces that are still *up* on the source system to prevent an IP address conflict. Because this requires the source Control Station and Data Mover to be accessible, depending on the failure scenario, it may not be possible. Due to the combination of the read-only LUNs and the residual VDM configuration in the NAS DB, the Data Mover is highly likely to panic if there are I/O failures. This will impact client access to any resources that are hosted on that Data Mover.

Note that if a failover is issued while the source system is still online and the session is not in sync, data loss can occur because there are changes that have not been replicated. If a failover is issued while the source SPs are unavailable, such as a power outage,

a force promote is required to bring the destination MirrorView Consistency Group online. This means a full synchronization is required if the original source site is recovered. In these situations, the following warning is displayed:

```
WARNING: The MirrorView/S consistency group is not in a proper state for failover activation. If you continue
while the consistency group is in this state and force the failover, a full synchronization of the source from
the destination will automatically occur during the restore to reconstruct the group and mirrors, which is time-
consuming. Are you sure you want to proceed?  [yes or no]
```

Since this is a pull operation, this command must be executed on the destination system. Prior to running this command, it is crucial to confirm that the configuration matches on both systems. Refer to the "Configure Client Transparency" section for more details.

This is an example of this command:

```
nas_syncrep -failover VDM1_Session

WARNING: You have just issued the nas_syncrep -failover command. Verify whether the peer system or any of its
file storage resources are accessible. If they are, you should issue the nas_syncrep -reverse command instead.
Running the nas_syncrep -failover command while the peer system is still accessible could result in Data Loss if
the session is not in sync. Are you sure you want to proceed?  [yes or no] yes
Now doing precondition check...                          done: 7  s
Now doing health check...                                done: 2  s
Now cleaning local...                                    done: 1  s
Now switching the session (may take several minutes)...  done: 18 s
Now importing sync replica of NAS database...            done: 8  s
Now creating VDM...                                      done: 5  s
Now importing VDM settings...                            done: 0  s
Now mounting exported FS(s)/checkpoint(s)...             done: 6  s
Now loading VDM...                                       done: 6  s
Now turning up local network interface(s)...             done: 1  s
                                    Service outage end: 54 s

Now mounting unexported FS(s)/checkpoint(s)...           done: 0  s
Now importing schedule(s)...                             done: 0  s
                                         Elapsed time: 55 s
done
```

The failover command:

- Fails over the MirrorView Consistency Group.
- Reads the configuration from the NAS DB.
- Rebuilds all the objects for the VDM.
- Loads the VDM and mounts exported file systems and checkpoints.
- Enables the destination network interfaces.
- Mounts unexported file systems and checkpoints.
- Rebuilds the checkpoint schedules.

This operation reads the VDM configuration from the NAS DB and brings the VDM, along with its associated resources, online on the destination system. The time needed to complete this operation depends on the number of resources, such as file systems and checkpoints, on the VDM. While the command is running, a detailed output of the operation and timing is printed to the screen. The service outage is the elapsed time between when the command is issued and when the interfaces are brought online. To minimize the service outage time, non-critical resources such as unexported, temporarily unmounted, temporarily unloaded objects, and checkpoint schedules are brought online afterwards. Once the command completes, the elapsed time it took for the command to run is printed.

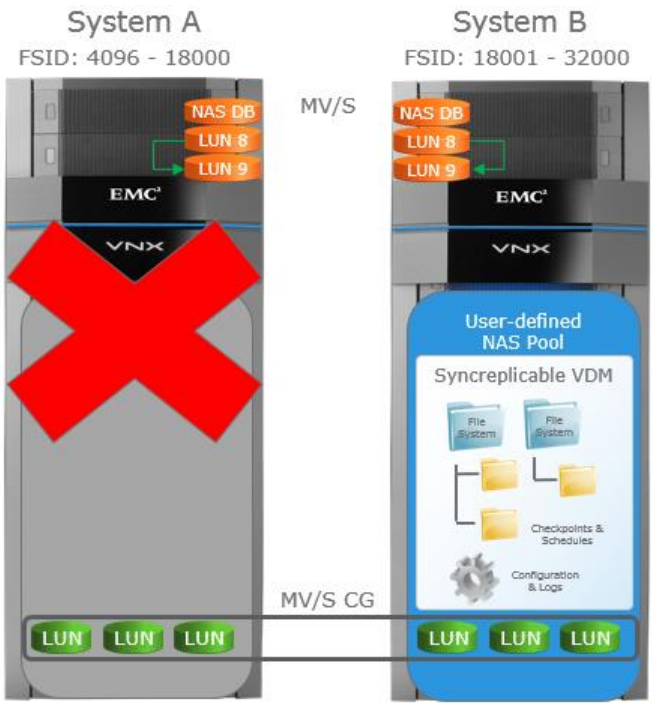

Figure 7. VDM MetroSync Failover

Figure 7 shows the VDM MetroSync session status after a failure on System A and a failover operation is complete. System B becomes the active system for the VDM and clients can access data from it. Note that replication between the two systems is stopped since System A is unavailable.

CLEAN

The Clean command is related to failover. This is run on the original source system to clean up residual objects in the NAS DB that were left over from the failover operation. If the primary system is brought offline temporarily, such as for a power outage, and the session is failed over, it is important to ensure the system remains offline. This can be accomplished by using an intelligent power breaker switch which holds the system down after power loss. If the original source system powers back up automatically, this could lead to a situation where both systems are active and duplicate IP address conflicts cause disruptions for users. When the system is ready to be brought back online, the Data Mover's network cables should be removed prior to powering it on. Once the system is powered up, a Clean operation needs to be run to bring down the interfaces, remove the residual resources, and restart the replication. After the Clean operation has been run, it is safe to reconnect the Data Mover's network cables.

This is an example of this command:

```
nas_syncrep -Clean VDM1_Session

WARNING: You have just issued the nas_syncrep -Clean command. This will result in a reboot of the original
source Data Mover that the VDM was failed over from. Verify whether or not you have working
VDM(s)/FS(s)/checkpoint(s) on this Data Mover and plan for this reboot accordingly. Running the nas_syncrep -
Clean command while you have working VDM(s)/FS(s)/checkpoint(s) on this Data Mover will result in Data
Unavailability during the reboot. Are you sure you want to proceed? [yes or no] yes
Now cleaning session VDM1_Session (may take several minutes)...     done
Now rebooting Data Mover server_2 ...                               done
Now starting session VDM1_Session...                                done
done
```

The Clean command does the following:

- Removes residual VDM configuration information from the NAS DB.
- Reboots the Data Mover, if necessary.
- Restarts replication, if stopped.

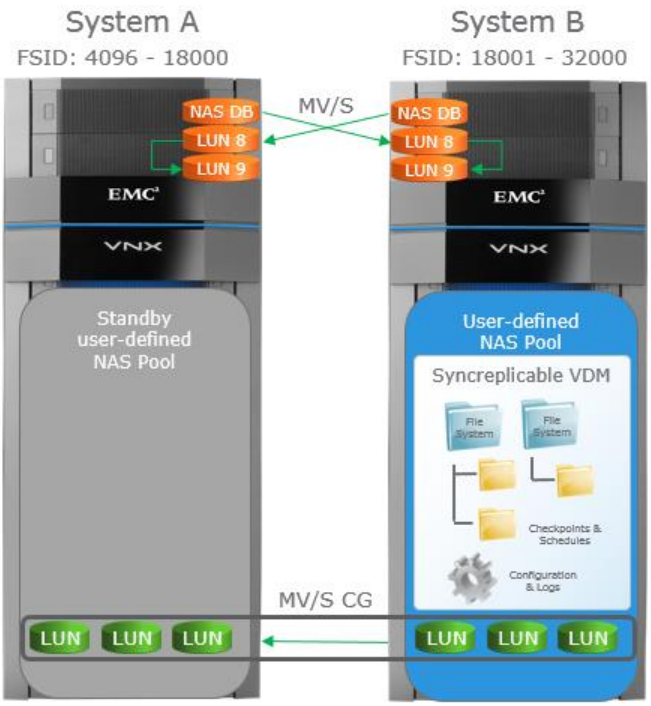

Figure 8. VDM MetroSync Clean

Figure 8 shows the VDM MetroSync session status after the Clean operation is complete on System A. System B remains active and replication is restarted towards System A. The NAS DBs on both systems are also synchronized.

The Clean command also enables the ability to reverse the VDM MetroSync session back to the original source system. Until this is complete, any attempts to run a reverse operation on this VDM will return an error. In situations where the source site is already cleaned, running the Clean command again will detect this and take no action. Once the Clean operation is complete, if the Data Mover's network cables were removed, they can be reconnected. Then, a reverse operation can be issued to gracefully bring the VDM back online on the original source system.

EXPAND VDM METROSYNC NAS POOL

If a NAS Pool that is a part of a VDM MetroSync session runs low on space, the NAS Pool can be extended by adding new LUNs. Note that VNX File does not support extending the size of existing LUNs. Adding new LUNs to a NAS Pool that is used for VDM MetroSync is possible without the need for a full synchronization. When this procedure is followed, only the newly-added LUNs that are used for the extension need to be synchronized. This allows the NAS Pool to be extended without any disruption.

Prior to starting this procedure, confirm the following:

- Ensure the session is in sync and not failed over by using the `nas_syncrep` command.
    - If a failover has occurred, the original source system must be cleaned prior to starting the NAS Pool expansion.
- Check how much free space and which disk volumes are currently in the NAS Pool by using the `nas_pool` command.
- Confirm the size of the disk volumes that are currently in the NAS Pool by using the `nas_disk` command.

When provisioning these LUNs, follow the recommendations referenced in the *VNX2 Unified Best Practices for Performance* white paper to achieve the best performance and availability. It is highly recommended to create LUNs that are the same size as the existing LUNs. The LUNs must be added to the ~filestorage Storage Group and then a rescan must be run in order to make them available as disk volumes for VNX File. After provisioning LUNs on the source system, the same number and size LUNs must also be provisioned to VNX File on the peer system. In order to achieve the same performance, you should build the LUNs using the same disk and RAID types on both systems.

Run the `nas_disk` command to ensure you can see the newly-added disk volumes and note their names on both systems. The disk volume names will be used in the `syncrep_modify_mirror` command, which has the following syntax:

```
/nas/sbin/syncrep/syncrep_modify_mirror -create -session <session_name>
-local_disk_volumes <volume_name>[,<volume_name>,...]
-remote_disk_volumes <volume_name>[,<volume_name>,...]
```

This is an example of this command:

```
/nas/sbin/syncrep/syncrep_modify_mirror -create -session VDM1_Session -local_disk_volumes d17,d18 -
remote_disk_volumes d17,d18

Now creating mirror(s)...                    done
Now waiting for mirror(s) to complete initial copy...       done
Done
```

The `syncrep_modify_mirror` command must be run on the source system. This command creates a remote mirror for each local disk volume and also starts the initial synchronization. The command will output the initial synchronization status while it is running. This command can be aborted by pressing CTRL+C without interrupting the initial synchronization, which continues to run in the background. The initial synchronization status can be checked again by repeating the same command. While the initial synchronization is occurring, the VDM remains protected, and reverse and failover operations can be issued normally. If a failover occurs, ensure a clean operation is issued prior to starting the next step.

Once the initial synchronization is complete, the `syncrep_modify_cg` command can be run to add the new LUNs in to the session's existing MirrorView Consistency Group. The syntax for the `syncrep_modify_cg` command is:

```
/nas/sbin/syncrep/syncrep_modify_cg -addmirror -session <session_name>
-local_disk_volumes <volume_name>[,<volume_name>,...]
-remote_disk_volumes <volume_name>[,<volume_name>,...]
```

This is an example of this command:

```
/nas/sbin/syncrep/syncrep_modify_cg -addmirror -session VDM1_Session -local_disk_volumes d17,d18
-remote_disk_volumes d17,d18
```

The `syncrep_modify_cg` command must also be run on the source system. Once the LUNs are in the Consistency Group, the NAS Pool can be extended. The `-skip_session_check` switch in the `nas_pool` command must be used for the extension. The syntax for the `nas_pool` command is:

```
nas_pool -xtend <name> -volumes <volume_name> [,<volume_name>,...] -skip_session_check
```

This is an example of this command:

```
nas_pool -xtend VDM1_Pool -volumes d17,d18 -skip_session_check
```

Expanding the NAS Pool using the `-skip_session_check` switch must only be used after the mirrors have been synchronized and added to the MirrorView Consistency Group. The destination side needs to be extended first. Once the destination has been extended, the source can be extended using the same command. This finalizes the extension procedure and allows the newly-added disk volumes to be used.

ROLLBACK VDM METROSYNC NAS POOL EXPANSION

If any issues are encountered during VDM MetroSync NAS Pool extension procedure, it is possible to rollback each command. Using the rollback commands allows you to remove the newly-added LUNs from the NAS Pool, remove the LUNs from the MirrorView Consistency Group, and deletes the primary/mirror relationship. It is important to note that this procedure is not designed to be used to shrink a NAS Pool. It should be used only to undo a NAS Pool expansion before any resources are built on the disk volumes. To undo each step, these commands are run in reverse order compared to the expansion.

The `nas_pool` command is used to shrink the NAS Pool, also using the `-skip_session_check` switch. This is an example of this command:

```
nas_pool -shrink VDM1_Pool -volumes d17,d18 -skip_session_check
```

If both sides have been extended, shrink the source side first and then shrink the destination side. Otherwise, only the destination side needs to be shrunk.

The next step is to use the `syncrep_modify_cg` command to remove the LUNs from the MirrorView Consistency Group. This is an example of this command:

```
/nas/sbin/syncrep/syncrep_modify_cg -removemirror -session VDM1_Session -local_disk_volumes d17,d18 -
remote_disk_volumes d17,d18
```

Once the mirrors have been removed from the CG, the primary LUN and mirror relationship can be deleted by running the `syncrep_modify_mirror` command:

```
/nas/sbin/syncrep/syncrep_modify_mirror -delete -session VDM1_Session -local_disk_volumes d17,d18 -
remote_disk_volumes d17,d18
```

Once this command completes, the NAS Pool expansion has been completely rolled back. These disk volumes will be marked as unused and can be added to another NAS Pool or deleted.

LIMITATIONS

- VDM MetroSync is not supported on systems where Cabinet-level File DR solutions are in use.
- Common log file systems are not supported. Only split-log VDMs and file systems are supported in a VDM MetroSync session. To transfer a common log file system to a split-log file system, use a host-based migration.
- Shrinking a user-defined NAS pool is not allowed if there is an active VDM MetroSync session. You must delete the session, shrink the user-defined NAS pool, and then recreate the session.
- NDMP backups may be stopped when a VDM reverse or failover occurs because the full mount paths of the file systems will be changed. To work around this, a symbolic link can be created to point to the proper path on the system.
- When performing OE upgrades, upgrade the target system prior to upgrading the primary system.
- For each system pair, there is a maximum of 64 VDM MetroSync sessions.
- Since MirrorView is the underlying replication technology, the MirrorView limits cannot be exceeded. The limits for each system model are shown in Table 2.

Table 2. MirrorView Limits

Name	VNX5200	VNX5400	VNX5600	VNX5800	VNX7600	VNX8000
Max Number of Mirrors	128	128	128	256	512	1024
Max Number of Consistency Groups	64	64	64	64	64	64
Max Number of Members per CG	32	32	32	32	64	64

TROUBLESHOOTING

For troubleshooting and support purposes, a dedicated log is created for VDM MetroSync. This log is located in `/nas/log/nas_syncrep.log`.

VDM METROSYNC MANAGER

VDM MetroSync Manager is optional software that can be installed on a Windows server which works with VDM MetroSync. It provides a GUI interface that provides VDM MetroSync session information and allows for running operations such as reverse, failover, and restore. Note that the initial configuration of the VDM MetroSync sessions must be done by CLI. VDM MetroSync Manager can only be used to manage sessions that are already running. It also has the ability to continuously monitor sessions and automatically initiate failover when issues are detected.

Prior to installing VDM MetroSync Manager, it is important to consider the VDM configuration. The software allows you to designate a primary and target site and only monitors and manages the VDMs that are active on the designated primary site. In an active/passive array configuration, VDM MetroSync Manager should be installed at the remote site because this allows it to trigger a failover if there is a disaster at the primary site.

If there are active VDMs at both sites, an instance of VDM MetroSync Manager must be installed at each site. Each VDM MetroSync Manager instance should designate the remote site as the primary. When a session is either reversed or failed over, it is no longer monitored by that instance of VDM MetroSync Manager. In the other instance of VDM MetroSync Manager, a discover operation needs to be run in order to see and start monitoring the new session.

CONFIGURATION

The VDM MetroSync Manager server has the following requirements:

- 2 CPUs
- 2 GB RAM
- 2GB of space
- Dedicated Windows Server 2008 R2, Windows Server 2012, or Windows 7 32-bit or 64-bit client
- ICMP protocol must be allowed

As part of the VDM MetroSync Manager installation process, a reboot of the server is required. After rebooting, launching VDM MetroSync Manager for the first time opens the configuration wizard. This wizard allows you to:

- Configure VNX2 systems
    - Provide primary site IP address and credentials. VDM MetroSync Manager automatically discovers the target site. If the target site credentials are different, they must be provided here as well.
- Configure automatic failover operations:
    - Set a VDM Network to "Failed" when all VDM network interfaces fail or any VDM network interface fails.
    - If VDM MetroSync Manager failover fails, set the VDM failover retry count. The default is 3.
    - VDM MetroSync Manager sends a ping request to the specified Remote Target IP Address to determine if it has been isolated from the network. If it does not receive a response, it assumes the host has been isolated and other failures may be invalid. In this situation, VDM MetroSync will not initiate an automatic failover. The Remote Target IP should be a highly available network address such as a gateway.
- Configure SMTP email alerts (optional)
- Configure pre-failover and post-failover scripts (optional)
- Discover the system configuration and provide a detailed summary

Figure 9 shows an example of the summary page in the VDM MetroSync Manager configuration wizard.

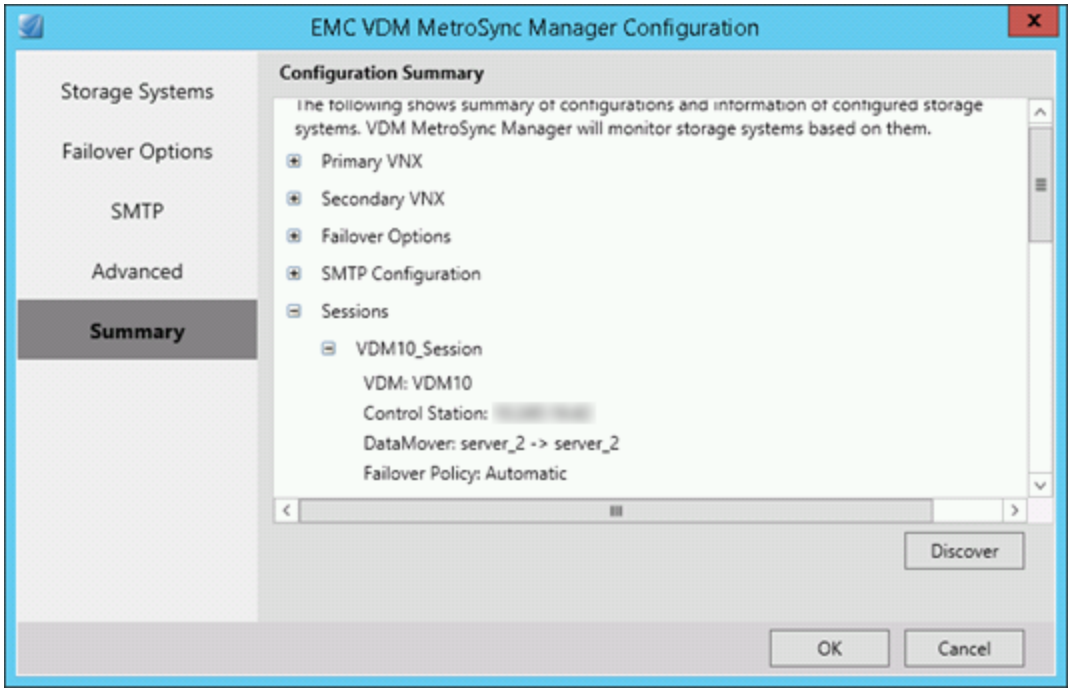

Figure 9. VDM MetroSync Manager Summary

VDM METROSYNC MANAGER SESSIONS

Once the configuration wizard has been completed, the main VDM MetroSync Manager is displayed as shown in Figure 10.

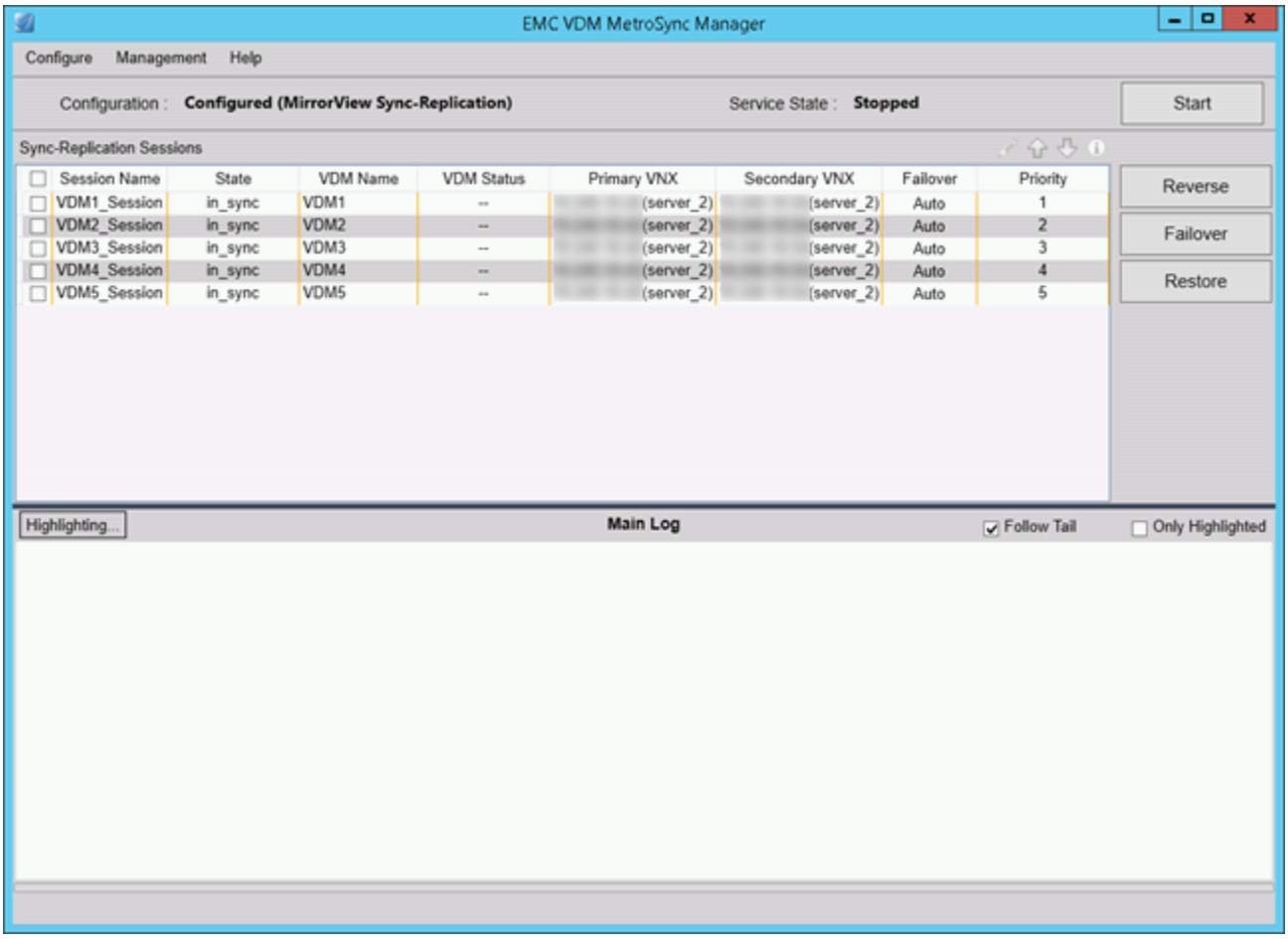

Figure 10. VDM MetroSync Manager Main Window

The VDM MetroSync Manager main window allows you to view a list of VDM MetroSync sessions and additional details such as status, direction, and failover settings. Each VDM MetroSync session can be configured by selecting the session and clicking the pencil icon or double-clicking on the session. This launches the configuration window as shown in Figure 11.

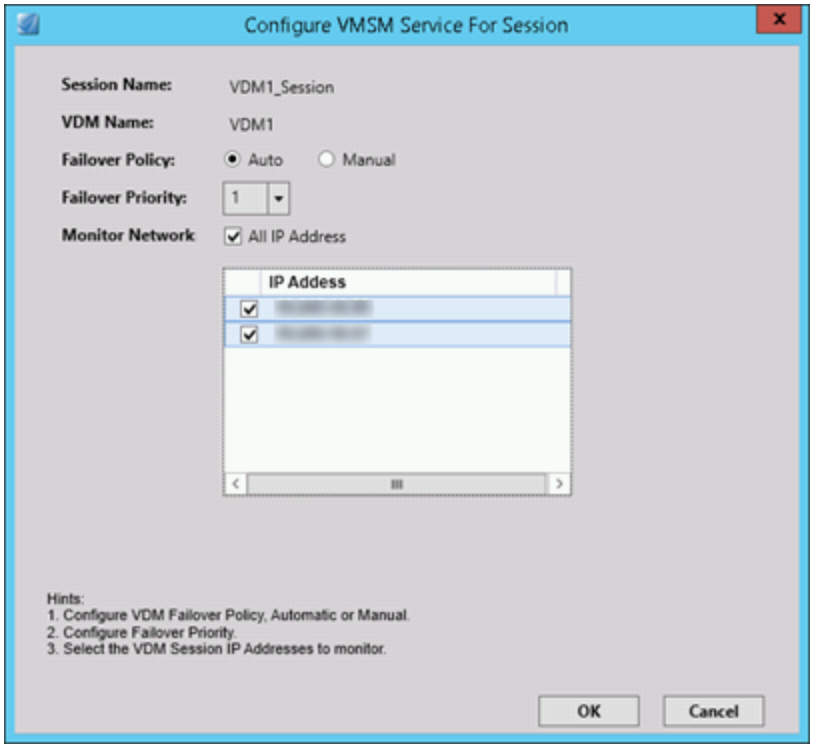

Figure 11. VDM Configuration Window

Each VDM session can be individually configured for either Manual or Automatic failover. When Automatic failover is enabled, the VDM is monitored and will automatically fail over to the secondary VNX when issues are detected. The failover priority can be edited here or on the main window using the up and down arrows. This determines the order that VDMs are failed over when multiple failures are detected. Also, choose whether all, some, or none of the attached network interfaces should be monitored for ping connectivity. When Manual failover is selected, the user has to initiate the failover, and none of these additional options need to be configured.

MANUAL OPERATIONS

On the main window, you can select one or more VDM MetroSync sessions and run operations such as Reverse, Failover, or Restore. The reverse and failover operations behave the same as the CLI. Failover should only be used in situations where the source system is unavailable. In cases where the source is still available, a reverse must be used instead. Restore is only available for sessions that are either reversed or failed over to the secondary VNX. This cleans the session, if necessary, and reverses it back to the primary VNX.

After a session is reversed or failed over, it is no longer monitored by VDM MetroSync Manager. The priorities of the remaining VDMs that are still active on the primary VNX are shifted up. When the VDM is reversed back to the primary system, it reclaims its original priority and the remaining VDMs are shifted down. A popup is displayed to inform the user of the new priorities whenever they change as shown in Figure 12.

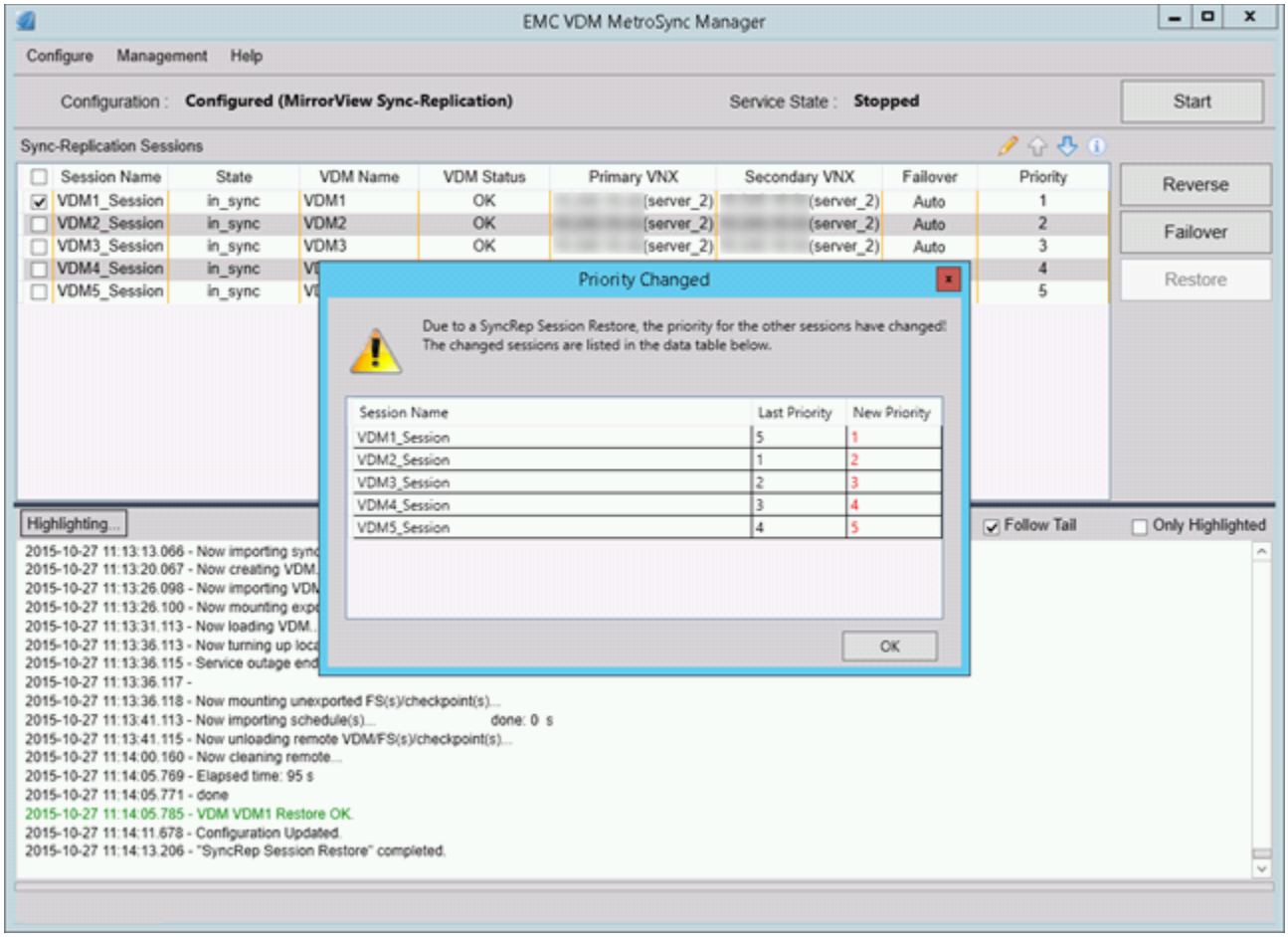

Figure 12. VDM MetroSync Manager Reverse

While these operations are running, the real-time output will be displayed in the Main Log window at the bottom of the VDM MetroSync Manager main window. By default, a copy of this log can also be found at C:\Program Files\EMC\VMSM\log\vmsm.main.log or in the Log Collection. "Follow Tail" is enabled by default which automatically scrolls the window to continuously show the latest log output. "Only Highlighted" can be checked to only show important events that are highlighted red or green, by default. The default colors can be changed by clicking the "Highlighting" button.

VDM METROSYNC MANAGER SERVICE

Once the sessions have been configured, the VDM MetroSync Manager service can be started. This service continuously monitors the environment for issues that could disrupt data access for users. The components that are monitored include:

- Data Movers
- VDM interfaces
- File systems
- LUNs

While the service is running, the results of each component check is displayed in the log window. Note that manual operations cannot be run while the service is running. If a manual operation is initiated while the service is running, the user will be prompted to stop the service first. Once the operation is complete, it is important to start the service again because VDM MetroSync Manager does not take any automatic actions while the service is stopped. When the service is started, it is displayed in green as shown in Figure 13.

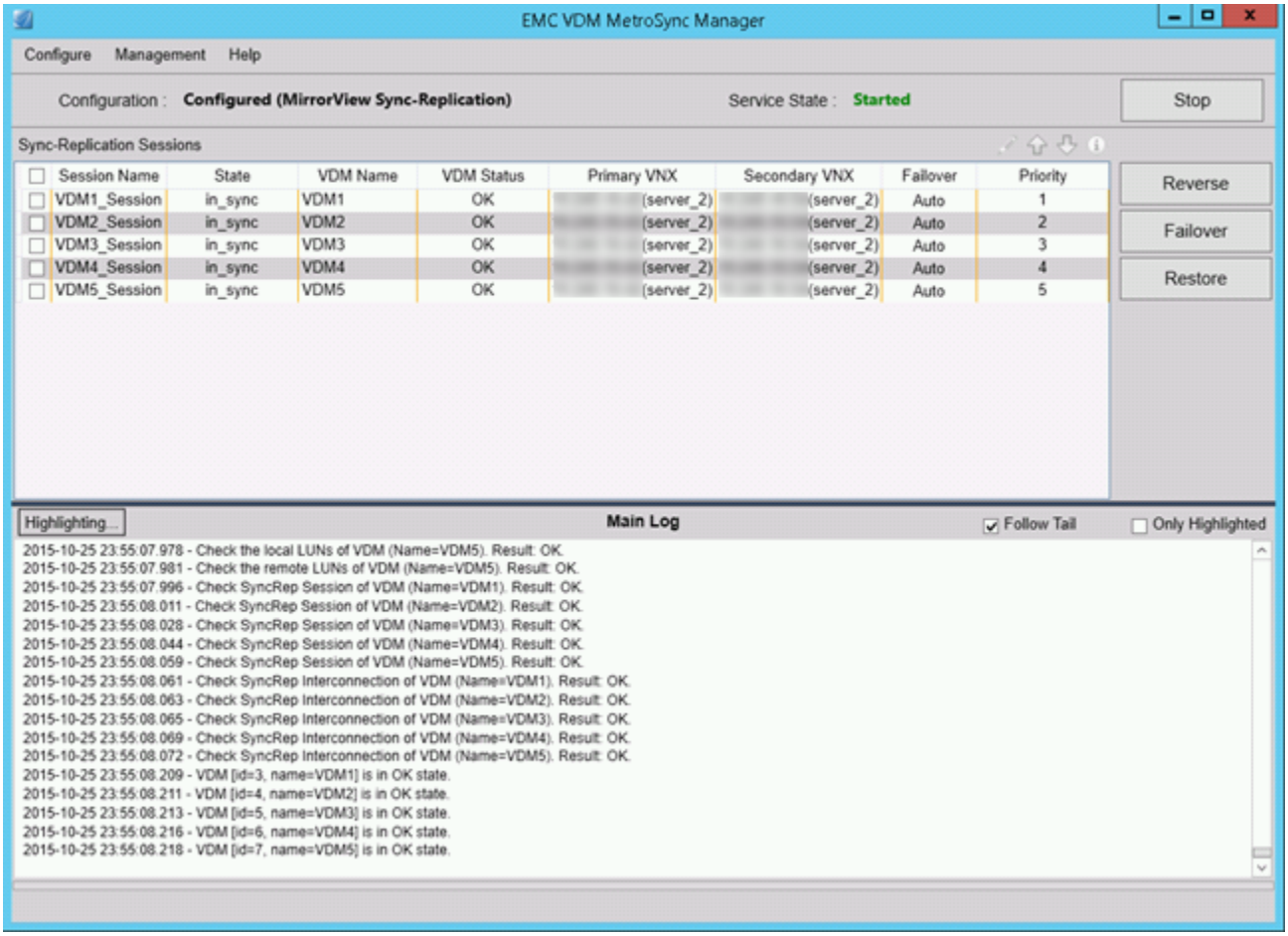

Figure 13. VDM MetroSync Manager Service

Table 3 lists the checks that are monitored by the VDM MetroSync Manager service. Note that some checks result in warnings but others result in failures. VDM MetroSync Manager only initiates a failover when a failure is detected.

Table 3. VDM MetroSync Manager Service Checks

Event	Warnings	Failures	Description
Remote Ping	✓		Check remote IP ping
Primary Site CS0	✓		Check /nbsnas mount on primary site CS0
Primary Site CS1	✓		Check /nbsnas mount on primary site CS1
Secondary Site CS0	✓		Check /nbsnas mount on secondary site CS0
Secondary Site CS1	✓		Check /nbsnas mount on secondary site CS1
Primary Site SPA	✓		Check SPA status of primary site
Primary Site SPB	✓		Check SPB status of primary site
Secondary Site SPA	✓		Check SPA status of secondary site
Secondary Site SPB	✓		Check SPB status of secondary site
Primary Site Data Mover	✓	✓	Check Data Mover status of primary site
Secondary Site Data Mover	✓		Check Data Mover status of secondary site
VDM Interface Check	✓	✓	Ping interfaces bound on the VDM
VDM File System Check	✓	✓	Access file systems mounted on VDM
VDM Local LUN Status Check	✓	✓	Check status of local LUNs used by VDM (ERROR/DEGRADED)
VDM Remote LUN Status Check	✓		Check status of remote LUNs used by VDM (ERROR/DEGRADED)

Event	Warnings	Failures	Description
Synchronous Replication Interconnect Check	✓		Check status of interconnect of synchronous replication session
Synchronous Replication Session Check	✓		Check status of synchronous replication session

If a failure is detected while the service is running, VDMs are failed over starting with the VDM with the highest priority. The log output will display the failure detection, failover progress, and failover completion. Also, if SMTP alerts are configured, email notifications are sent to notify the administrator about the latest status. The automatic failover process is shown in Figure 14.

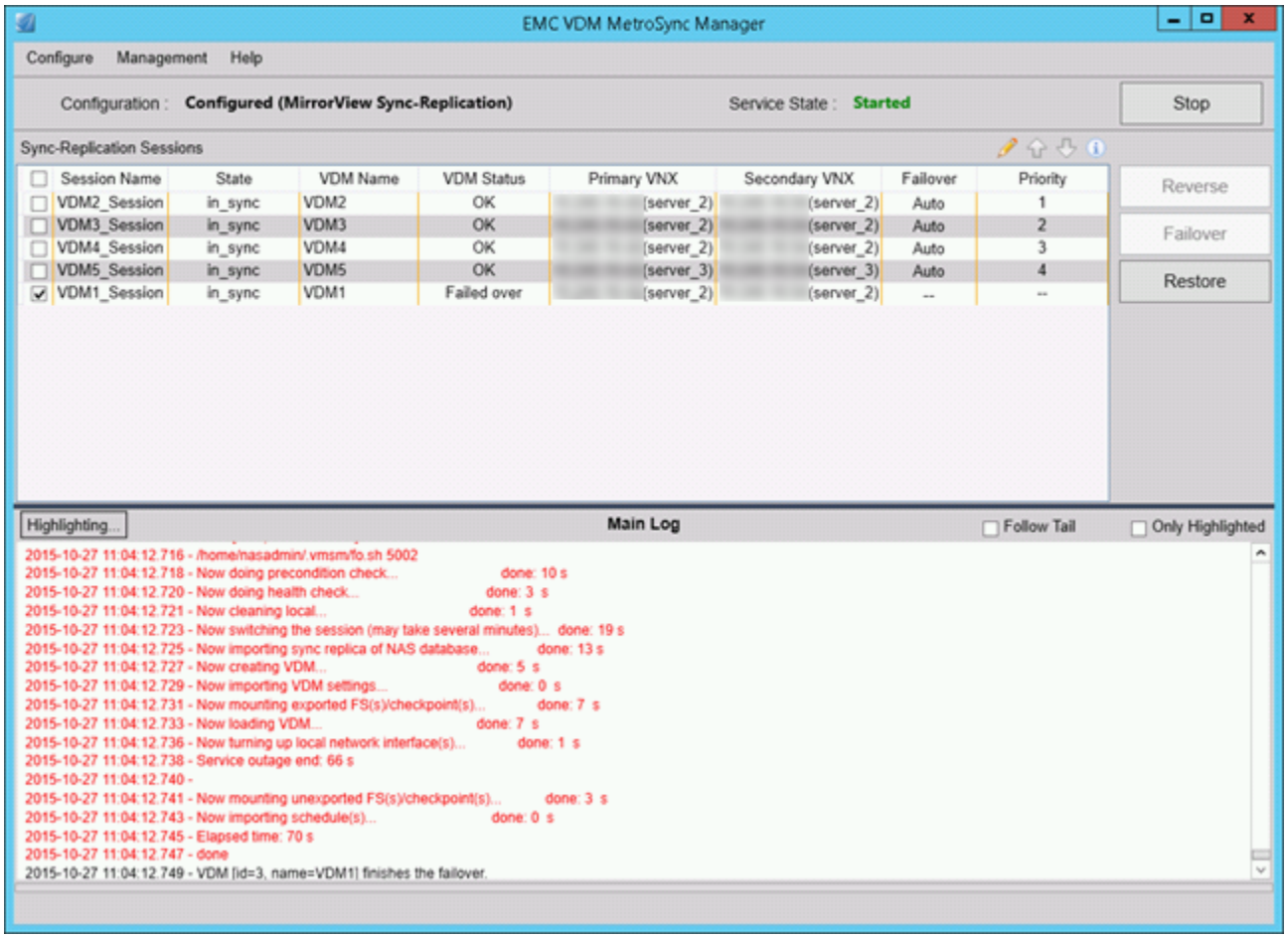

Figure 14. VDM MetroSync Manager Automatic Failover

If a failure is detected and there is a local standby Data Mover configured, VDM MetroSync Manager waits for the local Data Mover failover to complete to see if that resolves the issue. If the issue is resolved, then no further action is taken and the monitoring continues normally. If issues are still detected, then failover to the secondary VNX is initiated. This delay is avoided when the primary system is configured without any standby Data Movers.

After the failover is complete, the VDMs will be running on the secondary VNX. After the failure on the primary system has been resolved, it can be brought back online. If the system was powered off, the Data Mover network cables should be disconnected before powering it on, to avoid an IP address conflict. Afterwards, a restore operation can be run on the failed over sessions. This cleans the sessions on the primary VNX and runs a reverse to bring the VDMs back to the primary VNX. After the clean operation is complete, but before the reverse operation is started, you must reconnect the Data Mover's network cables. Note that the VDM MetroSync Manager service must be stopped in order to run the restore. Once the operation is complete, it is important to start the service again because VDM MetroSync Manager does not take any automatic actions while the service is stopped.

If VDM MetroSync Manager is used for automated failover, it should be installed as a VM and protected by a High Availability (HA) or Fault Tolerance (FT) service. VDM MetroSync Manager cannot initiate failover if the VM is offline.

Prior to upgrading the VNX for File OE, you must shut down the VDM MetroSync Manager service. This prevents VDM MetroSync Manager from initiating a failover when the Data Movers are rebooted.

CONFIGURE

The Configure menu, located at the top of the VDM MetroSync Manager main window, is used to edit the current configuration or to reset the VDM MetroSync Manager configuration. For configuration changes such as IP addresses, credentials, SMTP, and failover scripts, the configuration wizard is launched and edits can be made. After configuration changes are made, VDM MetroSync Manager will run a discover operation on the environment to retrieve the latest changes. There is also a reset configuration option which removes all of the user configuration so it can be set up again. Note the configuration cannot be changed while the VDM MetroSync Manager service is running.

MANAGEMENT

The Management menu, located at the top of the VDM MetroSync Manager main window, has options for Discover and Check Status. When changes are made to the VDM MetroSync environment, a discover operation must be run to rescan the environment for the changes. Changes include creating or deleting a VDM MetroSync session, attaching a new interface, running a manual operation using CLI, and so on. The Check Status button runs the same checks that the VDM MetroSync Manager service runs when it is enabled. Running this check manually initiates a single check of each component and reports the results in the Main Log window. Even if failures are detected during this command, because the VDM MetroSync Manager service is not running, no action is taken.

HELP

The Help menu, located at the top of the VDM MetroSync Manager main window, has links to the Help page, Log Collection, and About. The Help page launches the user guide for VDM MetroSync Manager which describes how to use the software. The Log Collection button captures all of the VDM MetroSync Manager related logs and zips them up in a log bundle. The user will be prompted to save the zip file to a location. This is useful for troubleshooting and support purposes. The About page displays the EMC VDM MetroSync Manager Software License Agreement.

PRESERVE REPLICATOR SESSIONS

OVERVIEW

With synchronous replication enabled between two systems, it is also possible to add asynchronous replication to a third system using Replicator. This allows the third system to be located further away and enables it to be used as a backup and recovery solution.

After a reverse or failover of a VDM MetroSync session, the Replicator sessions on the VDM become unavailable. This feature provides an automated method to recreate the Replicator sessions on the VDM without requiring a full copy. Since the Replicator checkpoints are replicated along with the VDM, a common base checkpoint is available. Using the common base checkpoint, these sessions can be recreated and incrementally updated on the new system where the VDM is active.

CONFIGURATION

The Preserve Replicator Sessions feature has the following requirements:

- Three VNX2 systems
    - Two VNX2 systems used for VDM MetroSync (Systems A and B)
    - One VNX2 system used as a Replicator destination (System C)
    - VNX OE for File 8.1.9.155 or later and Block 05.33.009.5.155 or later

ESTABLISH CONTROL STATION INTERCONNECTS

In order to create Replicator sessions between two systems, there needs to be a control path between the Control Station(s) of both systems. This allows the two systems to establish a secure communication channel to issue commands to the other system.

If a system currently is or has plans to be a dual Control Station (CS) configuration, you must first create a Control Station IP alias by using the `nas_config` command. This is an IP address that will always be active on the primary CS, regardless of whether it is running on CS0 or CS1. Without this, management of the Replicator sessions will become degraded in the event the CS fails over to CS1. For a single CS configuration, this step is optional. If this is skipped and a secondary CS is added later, the IP address of CS0 must be used for the alias and a new IP address must be assigned to CS0.

Run the `nas_cel` command to create the following CS relationships:

- System A → System C
- System B → System C
- System C → System A
- System C → System B

The CS interconnect between System A → System B should already be available for the VDM MetroSync sessions. For a dual CS configuration, the alias IP must be used in the command. Also, a passphrase needs to be specified and must match on both systems. This command must be run on both systems to allow for communication in both directions. Ensure port 443 is open and the time on the Control Stations are within 10 minutes of each other.

ESTABLISH DATA MOVER INTERCONNECTS

A Data Mover interconnect defines the data communication path between a pair of Data Movers for Replicator. First, ensure the control path between the two systems is established by using the `nas_cel` command. Then, run the `nas_cel –interconnect` command to establish a Data Mover interconnect to the peer system. This command configures the local and destination Data Movers and interfaces for the interconnect. It also has the ability to configure a bandwidth schedule for all Replicator sessions that are using this interconnect. Use the `nas_cel -interconnect` command to create the following Data Mover interconnects:

- System A → System C
- System B → System C
- System C → System A
- System C → System B

CREATE REPLICATOR SESSIONS

Once the Control Station and Data Mover interconnects are in place, Replicator sessions can be created for the file systems that are on the VDM. Note that Replicator sessions of the VDM can also be created but they will not be preserved upon reverse or failover. The Replicator sessions should be created on the system where the VDM is active. This means System A or B can be the source, depending on which system is the primary for VDM MetroSync, but the target is always System C.

By using the `nas_replicate` command, create Replicator sessions for the file systems that are on the VDM. When the session is created, the system starts the initial synchronization of the file system.

Figure 15 shows the coexistence of VDM MetroSync and file system Replicator sessions. There are three VNX2 Systems: System A, System B, and System C. A single VDM with several file systems is synchronously replicated from System A to System B using VDM MetroSync. There is also an asynchronous Replicator session created from System A to System C for each file system on this VDM. This allows the file systems on the VDM to be simultaneously protected by synchronous and asynchronous replication.

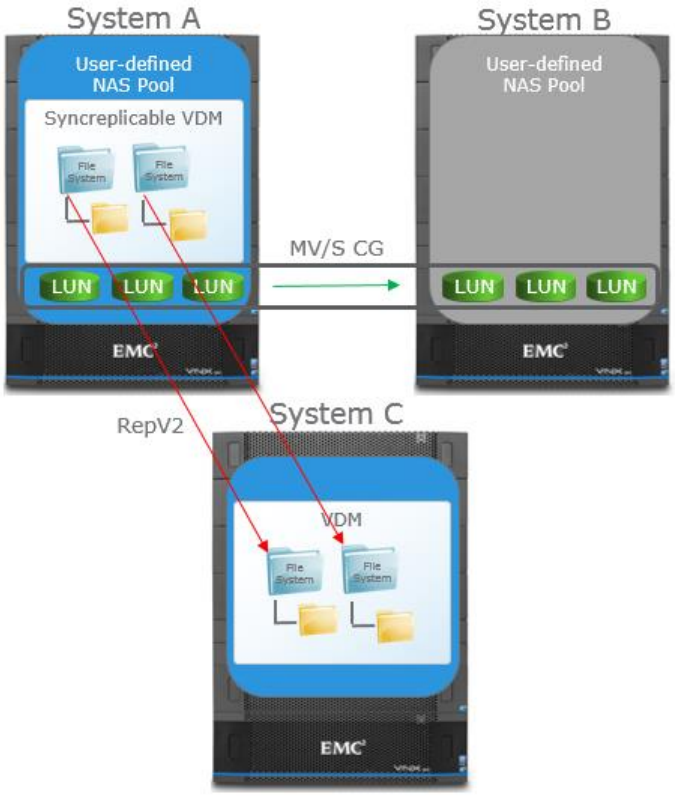

Figure 15. Coexistence of VDM MetroSync and Replicator Sessions

OPERATIONS

The `nas_syncrep_rr` command is a new command that can be used to manage the Replicator sessions after a VDM MetroSync reverse or failover. The syntax for this command is:

```
nas_syncrep_rr

          -list

       | -config { -enable | -disable | -info }

       | -restore { -all | -vdm { <vdm_name> | id=<vdm_id> } }

       | -free_intermediate_data

       | -Clean -server <server_name>
```

Table 4 lists these commands and their descriptions.

Table 4. Preserve Replicator Sessions Commands

Command	Description
list	Lists the name and state of the Replicator sessions that can be restored.
config	Configures the Preserve Replicator Sessions service.
restore	Restores the Replicator sessions on the system.
free_intermediate_data	Frees the intermediate Replicator internal checkpoints after restore is complete.
Clean	The Clean command is related to VDM MetroSync failover. This command is run on the original source system to clean up residual objects in the NAS DB. It also cleans up all broken Replicator sessions on the specified Data Mover which may remain after the failover operation.

CONFIG

Before this feature can be used, the Preserve Replicator service must be enabled by using the `config` command. This is an example of this command:

```
nas_syncrep_rr -config -enable

Note:'Preserve RepV2 sessions for VDM Sync MV Replication' only supports Mirror View environment. If you're in
VDM Sync RP/SRDF environment, please disable this feature immediately.

Enabling "Preserve RepV2 for SyncRep" on remote array
Enable Successfully

Enabling "Preserve RepV2 for SyncRep" on local array
Enabled Successfully
```

Once the service is enabled, VDM MetroSync and Replicator sessions can coexist. The user is now allowed to create Replicator sessions on file systems that are mounted on `syncreplicable` VDMs. This also enables the user to convert a normal VDM with existing Replicator sessions on it to a `syncreplicable` VDM. Note that Replicator sessions of the VDM are not supported.

The `nas_syncrep_rr -config -disable` command can also be used to disable this feature. Prior to running the `disable` command, all Replicator sessions on `syncreplicable` VDMs must first be deleted. If any Replicator sessions on `syncreplicable` VDMs still exist, the `disable` command fails.

LIST

After the Replicator sessions have been configured and synchronized, a VDM MetroSync reverse operation is initiated from System B. This brings the VDM online on system B and reverses the direction of the synchronous replication. However, this also breaks the asynchronous replication session since they are pointing from System A to System C. Figure 16 shows the results of a reverse operation on the VDM MetroSync session from System A to System B.

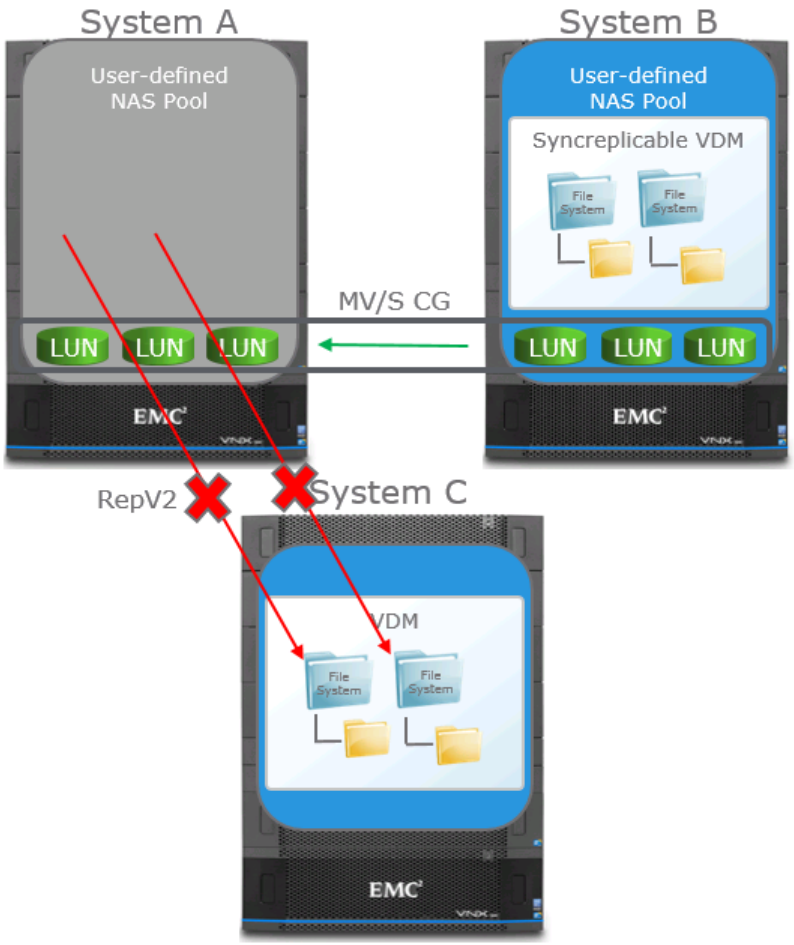

Figure 16. VDM MetroSync Reversed with Replicator Sessions

The list operation displays a list of broken Replicator sessions and details including the Session Name, VDM name, Local and Remote File System/VDM, CMU, Restore State, and Session Type.

This is an example of this command:

```
nas_syncrep_rr -list

Session Name    VDM Name        Local FS/VDM    Remote FS/VDM   CMU                 Restore State   Session Type

FS_Rep1          VDM1            5020            5020           APM001429131072007   ToBeRestored    Remote

FS_Rep2          VDM1            5021            5021           APM001429131072007   ToBeRestored    Remote

VDM_Rep1         VDM1            VDM1            VDM1           APM001429131072007   N/A             Remote
```

This command only displays the broken Replicator sessions on `syncreplicable` VDMs. If the VDM does not have the `syncreplicable` flag, the Replicator sessions cannot be restored. Also, this command checks whether the Control Station interconnects are available. If the interconnects are not available, the Replicator sessions cannot be restored. Also, any Replicator sessions that are already restored are not listed.

The Restore State has five possible values:

- ToBeRestored
- OnGoing
- Timeout

- Failed
- N/A

The Replicator sessions with a state of ToBeRestored can be preserved using the restore command. The Replicator sessions that have a state of OnGoing indicate that they are in the process of being restored.

If the restore fails due to an error, the state results in Timeout or Failed. Replicator sessions that have a state of Timeout or Failed can be restored by running the restore command again once the error is resolved.

In order to be restored, the file system Replicator must be in an OK state prior to the VDM MetroSync failover or reverse operation. Replicator sessions that have a state of stopped, switched over, or failed over cannot be restored. These sessions will have a Restore State of N/A and they will not be preserved. Also, VDM-level, local, and loopback Replicator sessions cannot be restored.

RESTORE

To restore the Replicator sessions, run the restore command on System B. The restore command runs the following operations:

1. Collects a list of the broken Replicator sessions along with their associated internal checkpoints. There are two internal checkpoints on System A (which are also available on System B when you use VDM MetroSync) and another two on System C.

2. Deletes the broken Replicator sessions between System A and System C, while keeping the internal checkpoints.

3. Reconstructs the Replicator sessions from System B to System C using the internal checkpoints. Since these checkpoints act as a common base, only the deltas need to be copied to restart the Replicator session.

4. After the Replicator session is created successfully, the system deletes the previous internal checkpoints.

This is an example of this command:

```
nas_syncrep_rr -restore -all

Info 26316963879: Command result: The nas syncrep rr command may take a long time to complete, please avoid
performing any syncrep operations for related vdm(s) during restore, use nas_task to check task {id = 13853}
status. Follow up by running "/nas/sbin/syncrep/RestoreRepv2/nas_syncrep_rr -free_intermediate_data" command to
clean all the intermediate data after restore finish.

OK
```

Because this process may take a while to complete, this command starts a background task which restores all of the Replicator sessions on this system. You can monitor the status of the task by using the `nas_task` command:

```
nas_task -i 13853
Task Id                = 13853
Celerra Network Server = VNX-MV2
Task State             = Succeeded
Movers                 =
Description            = VDM1_FS_RepV2,Succeeded,Reconstruct Repv2
id=458_APM00140916758_2007_116_APM00142913107_2007 successfully

Originator             = nasadmin@cli.localhost
Start Time             = Wed Oct 28 18:38:35 EDT 2015
End Time               = Wed Oct 28 18:40:27 EDT 2015
Schedule               = n/a
Response Statuses      = Info 26316963879: Command result: The nas_syncrep_rr command may take a long time to
complete, please avoid performing any syncrep operations for related vdm(s) during restore, use nas task to
check task {id = 13853} status. Follow up by running "/nas/sbin/syncrep/RestoreRepv2/nas_syncrep_rr -
free_intermediate_data" command to clean all the intermediate data after restore finish.
OK
```

The `nas_task` output provides detailed information for each Replicator session that is being restored. If any Replicator sessions fail to restore, additional details are provided here. Once the background task is complete, all of the Replicator sessions have been incrementally synchronized and are running normally from System B → System C, as shown in Figure 17.

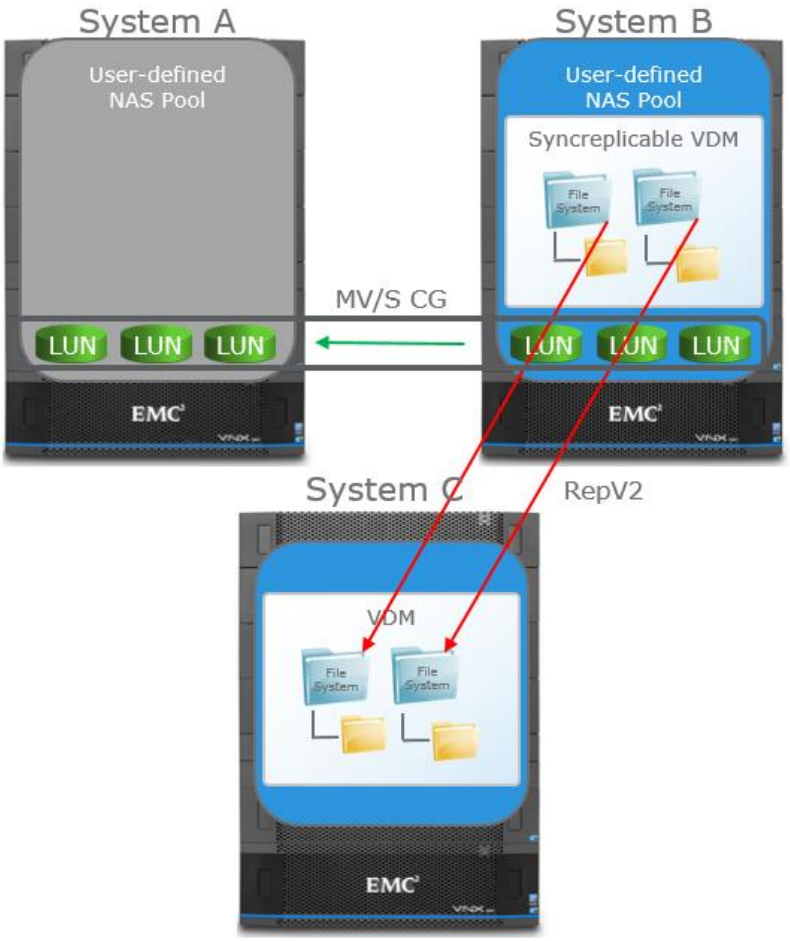

Figure 17. Preserved Replicator Sessions

Note that if you plan to reverse the VDM MetroSync session back to System A, the `nas_syncrep_rr -restore` command does not need to be run on System B prior to running the reverse. Once the VDM MetroSync session is active on System A again, the `nas_syncrep_rr -restore` command can be run directly on System A.

FREE INTERMEDIATE DATA

After the restore is complete, the free intermediate data command must be run to clean up the intermediate data that is left over after the restore process. This removes residual information from the broken Replicator sessions. This is an example of this command:

```
nas_syncrep_rr -free_intermediate_data

        Clean renamed internal ckpts for successfully restored RepV2 sessions:
        Clean Count: 12      Success: 12      Fail: 0
Free_intermediate_data Done
```

Once this command is complete, both the VDM MetroSync and Replicator sessions should be running normally. If the VDM is reversed or failed over at a later time, this process must be run again to restore the Replicator sessions on the new system.

CLEAN

The Clean command is related to VDM MetroSync failover. This command is run on the original source system to clean up residual objects in the NAS DB. It also cleans up all broken Replicator sessions on the specified Data Mover which may remain after the failover operation. Running this command is recommended after running the `nas_syncrep –Clean` command. The command requires root privileges to run.

This is an example of this command:

```
nas_syncrep_rr -Clean -server server_2
```

```
Now start cleaning up DART BDBs as well as all broken repV2 sessions.

Deleting server_2 repV2 session info.                          done
```

COMMON SCENARIOS

Here are some common scenarios for the Preserve Replicator Sessions feature.

VDM METROSYNC REVERSE

The Replicator sessions can be preserved after the VDM MetroSync session reverse. A high-level summary of the steps are:

1. Reverse the VDM MetroSync session to System B.
2. Restore the Replicator sessions to System B.
3. Free intermediate data on System B.

VDM METROSYNC FAILOVER

The Replicator sessions can be preserved after the VDM MetroSync session failover. A high-level summary of the steps are:

1. Failover the VDM MetroSync session to System B.
2. Clean the VDM MetroSync session on System A.
3. Clean the Replicator sessions on System A.
4. Restore the Replicator sessions to System B.
5. Free intermediate data on System B.

LIMITATIONS

- Local and loopback, one-to-many, and cascading Replicator sessions are not supported.
- Replicator sessions between the VDM MetroSync systems are not supported.
- Replicator sessions with states other than "OK", such as initial copy, switched over, or failed over, cannot be preserved.
- While a restore operation is running, no VDM MetroSync operations should be initiated.

CONCLUSION

The VDM MetroSync feature is a zero data loss replication solution designed for customers that demand continuous data availability. It can be used for load balancing and maintaining availability during scheduled maintenance events such as upgrades. In the event the primary site becomes unavailable, it allows an organization to recover from a disaster quickly and efficiently, in order to bring their business back online as soon as possible.

It is able to minimize downtime by automating failover to the secondary site when critical issues are detected using VDM MetroSync Manager. It provides a GUI interface to display VDM MetroSync session information and run operations to move, failover, or restore VDMs.

VDM MetroSync also supports asynchronous replication to a third system as a backup and recovery solution. When VDMs are moved or failed over between the VDM MetroSync systems, the asynchronous sessions to the third system are preserved. The asynchronous sessions can be recreated and incrementally updated on the new system where the VDM is active.

REFERENCES

Specific information related to the features and functionality described in this document are included in the following documents and white papers.

- Documentation
    - Using VDM MetroSync with VDM MetroSync Manager for Disaster Recovery
    - Configuring Virtual Data Movers on VNX
    - EMC VNX Command Line Interface Reference for File
    - Configuring and Managing CIFS on VNX
    - Configuring NFS on VNX
    - Using VNX SnapSure
    - Using VNX Replicator
    - Using MirrorView/Synchronous with VNX for File for Disaster Recovery
    - Parameters Guide for VNX for File
- White Papers
    - Virtual Data Movers on EMC VNX
    - EMC VNX Replication Technologies
    - EMC VNX2 Unified Best Practices for Performance
    - EMC MirrorView Knowledgebook Releases 30 – 33

The complete set of EMC VNX2 customer publications is available on the EMC Online Support website at https://support.emc.com. After logging in to the website, click the Support by Product page, to locate information for the specific product or feature required.