WHITEPAPER

DELL EMC UNITY: METROSYNC

A Detailed Review

Abstract

This white paper explains the usage of MetroSync for Dell EMC Unity™, a synchronous disaster recovery solution for file resources. The paper outlines the available commands and configurations available when replicating file data using this feature as well as utilizing advanced features. This feature is available on Dell EMC Unity OE version 4.4 and later.

February, 2019

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 source system 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 to the destination system when critical issues are detected.

This white paper provides a comprehensive overview of MetroSync for Dell EMC Unity™, a synchronous disaster recovery solution for file resources. This solution is designed to meet the above requirements.

AUDIENCE

This white paper is intended for IT planners, storage architects, system administrators, partners, Dell EMC employees and any others involved in evaluating, acquiring, managing, operating, or designing a MetroSync environment using Dell EMC Unity 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.

Converged Network Adapter (CNA) – A physical port on Dell EMC Unity systems that can be configured as Fibre Channel (FC), Optical Ethernet, Copper Ethernet, or TwinAx. This configuration is done at the factory and selected during the system ordering process. There are 2x CNA ports on each storage processor for physical Dell EMC Unity systems.

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

NAS Server – A Dell EMC Unity storage server that uses the SMB, NFS, or FTP/SFTP protocols to catalog, organize, and transfer files within designated file system shares. A NAS Server, the basis for multi-tenancy, must be created before you can create file-level storage resources such as file systems or VMware file datastores.

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.

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.

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

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 destination systems before an acknowledgement is sent back to the host.

Storage Pool – A collection of physical drives organized in a logical grouping for use on Dell EMC Unity systems for storage resource provisioning.

Storage Processor (SP) – A storage node that provides the processing resources for performing storage operations as well as servicing I/O between storage and hosts.

Unisphere – A web-based Dell 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.

Unisphere Command Line Interface (UEMCLI) – An interface that allows a user to perform tasks on the storage system by typing commands instead of using the graphical user interface.

INTRODUCTION

MetroSync for Dell EMC Unity is a disaster recovery (DR) solution for file resources, also known as File Synchronous Replication, which leverages a synchronous connection to create a zero data loss replication solution. This feature is available on Dell EMC Unity systems running OE version 4.4 or later. MetroSync allows for replication of a NAS Server along with all of its contents as well as file systems, association of file systems to snapshot schedules, snapshots, SMB servers, exports, interfaces, and so on. It can be configured in either a uni-directional configuration where the source NAS Servers are constrained to one system, or a bi-directional configuration where each system has its own set of source NAS Servers. NAS Servers can be moved or failed over from one system to another for load balancing, maintenance, or disaster recovery. Synchronous replication has a distance limitation based on latency between systems. This limitation is generally 60 miles or 100 kilometers between sites. To support MetroSync, the link latency must be less than 10 milliseconds.

With synchronous replication enabled between two systems, it is also possible to add asynchronous replication to a third system by using coexisting synchronous/asynchronous feature. This allows the third system to be located further away and enables it to be used as a backup and recovery solution. When NAS Servers are moved or failed over between the synchronously replicated systems, the asynchronous sessions to the third system can be preserved. Common base snapshots are replicated along with the NAS Server, which removes the requirement for a full synchronization after a failover. The asynchronous sessions can be incrementally updated and restarted on the new source system following a failover where the NAS Server is now active.

Along with file synchronous replication sessions, read-only snapshots are replicated to the destination which ensures consistent snapshots on both sites. Snapshot replication is automatically enabled for all resources that have file synchronous replication enabled which could include File Systems and VMware NFS Datastores. Both scheduled snapshots and user created snapshots are replicated. Another feature of MetroSync is snapshot schedule replication which gives the ability to replicate snapshot schedules as well as apply the replicated snapshot schedules to synchronous file resources. This allows a snapshot schedule to remain intact upon a failover for a particular resource thereby reducing storage administration and saving time.

In OE 4.5 or later, an application called MetroSync Manager can be utilized which enables the ability for automatic failover in the event of a unplanned outage or disaster. MetroSync Manager serves as an offsite witness to monitor the system statuses of two sites participating in file synchronous replication and will initiate a cabinet level unplanned failover if it detects a critical failure at one of the sites. This gives the benefit of reducing downtime in the event of a disaster so business operations can continue unhindered from the destination site. MetroSync Manager is an optional tool and is not required to utilize other MetroSync capabilities.

Some of the typical use cases for MetroSync 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

MetroSync for Dell EMC Unity systems includes multiple advanced features that make it an enterprise ready solution. The advanced features will be covered in separate sections. See below for the full list. Click on the associated hyperlinks to jump directly to that particular section.

- Asynchronous Replication to a 3rd Site
- MetroSync Manager

METROSYNC PRE-REQUISITES

REQUIREMENTS

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

- Two physical Dell EMC Unity systems
  - Dell EMC Unity OE version 4.4 or later
    - To utilize MetroSync Manager, OE version 4.5 or later is required
  - Fibre Channel (FC) connectivity (direct connect or via switch) of the Synchronous Replication Ports between the two systems
  - Both systems' time needs to be within 5 minutes of each other
  - The two Dell EMC Unity systems do not need to be the same model, but it is highly recommended to pair system models with the same default NFS transfer sizes*
- (Optional) Third Dell EMC Unity system (physical or Dell EMC UnityVSA) – For asynchronous replication to a 3rd site
  - Dell EMC Unity OE version 4.4 or later
  - **Note**: Take into account system limitations when sizing out environments (i.e. Number of supported asynchronous replication sessions)
- Recommended: NTP configured on both systems to ensure times are in sync

* If you're using NFSv3 or NFSv4, ensure both systems are configured with the same `nfs.v3xfersize`. See Table 1 for a list of models and their default NFS transfer sizes. If these values do not match, this parameter must be updated. You should update the system that has the larger value to match the system with the smaller value. This change requires a reboot of both storage processors. For instructions on how to modify this parameter, reference KB article 521956.

Table 1. Default NFS Transfer Sizes

System Model	Default Transfer Size
Dell EMC Unity 300(F)	65535
Dell EMC Unity 350F, 400(F), 500(F)	131072
Dell EMC Unity 450F, 550F, 600(F), 650F	262144

METROSYNC CONFIGURATION

In this section, we will go over technical details for each setup required to successfully setup MetroSync on Dell EMC Unity systems. This feature can be configured via Unisphere GUI, CLI, or REST API. For this document, all steps will be covered using Unisphere examples unless otherwise specified. Below is a high-level summary of the steps we will be covering in this section:

1. Configure Sync Replication Management Interfaces on both SPs for each system
2. Establish Fibre Channel (FC) connection (direct connect or via switch) of the Synchronous Replication Ports both systems
3. Configure Synchronous Replication Connection between both systems
4. Create synchronous replication sessions

In the below example, we start with two physical Dell EMC Unity systems which have not been configured for MetroSync yet, as seen in Figure 1. In this example, assume that both systems have been properly installed, licensed, and have the same NTP server configured. Also, there is an Ethernet switch and FC switch which will be referenced in later examples. Lastly, management connectivity (seen in yellow in Figure 1) is already configured which allows access via Unisphere.

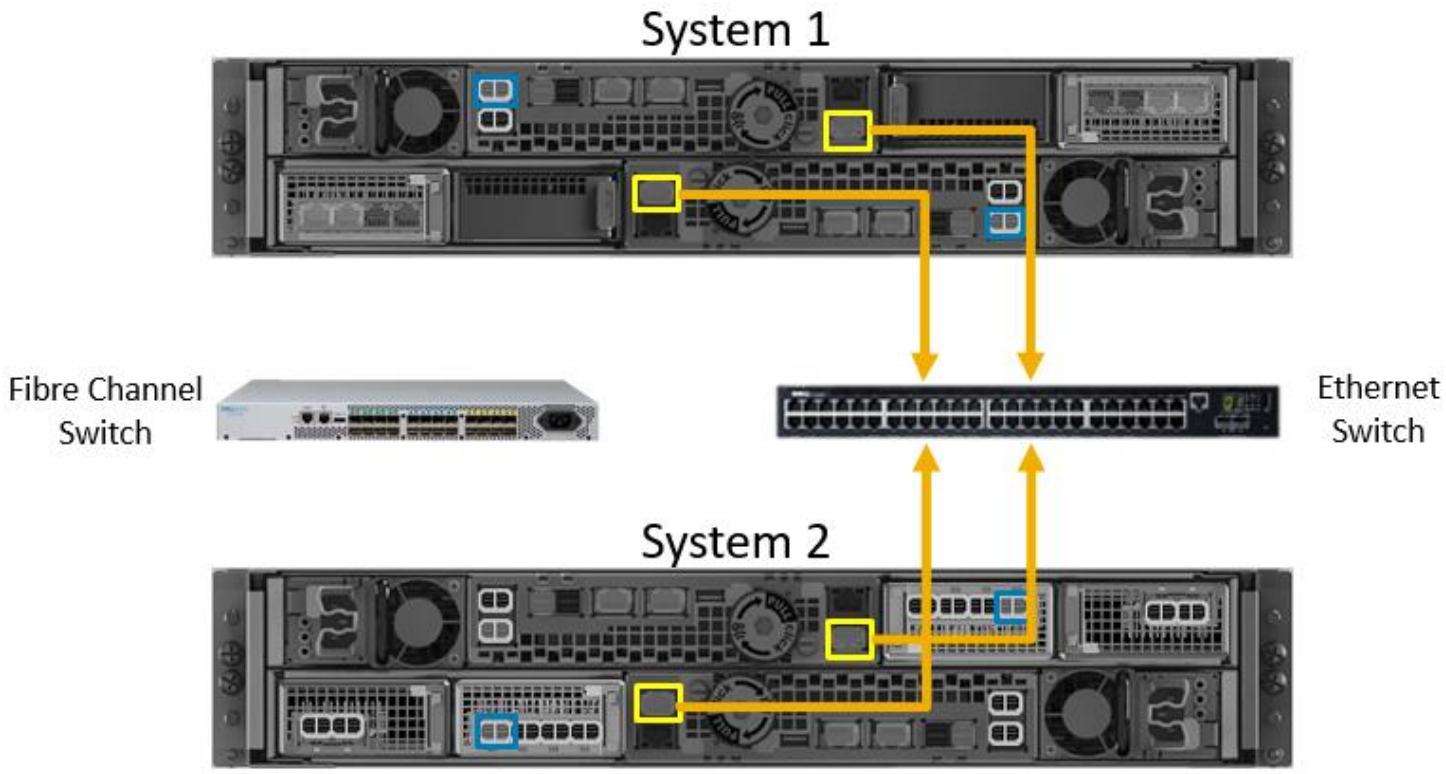


Figure 1 - Example Setup (with Ethernet connected)

CONFIGURE THE SYNC REPLICATION MANAGEMENT INTERFACES

In order to use MetroSync, there needs to be a secure communication path for replication commands between the two systems. Therefore, you must configure sync replication management interfaces on each SP for both systems. The interfaces are configured on the existing management port connections, which are already used for Unisphere, so no additional physical cables are required.

To configure sync replication management interfaces via Unisphere, navigate to the **Interfaces** page and click on the **Add** button as seen in Figure 2.

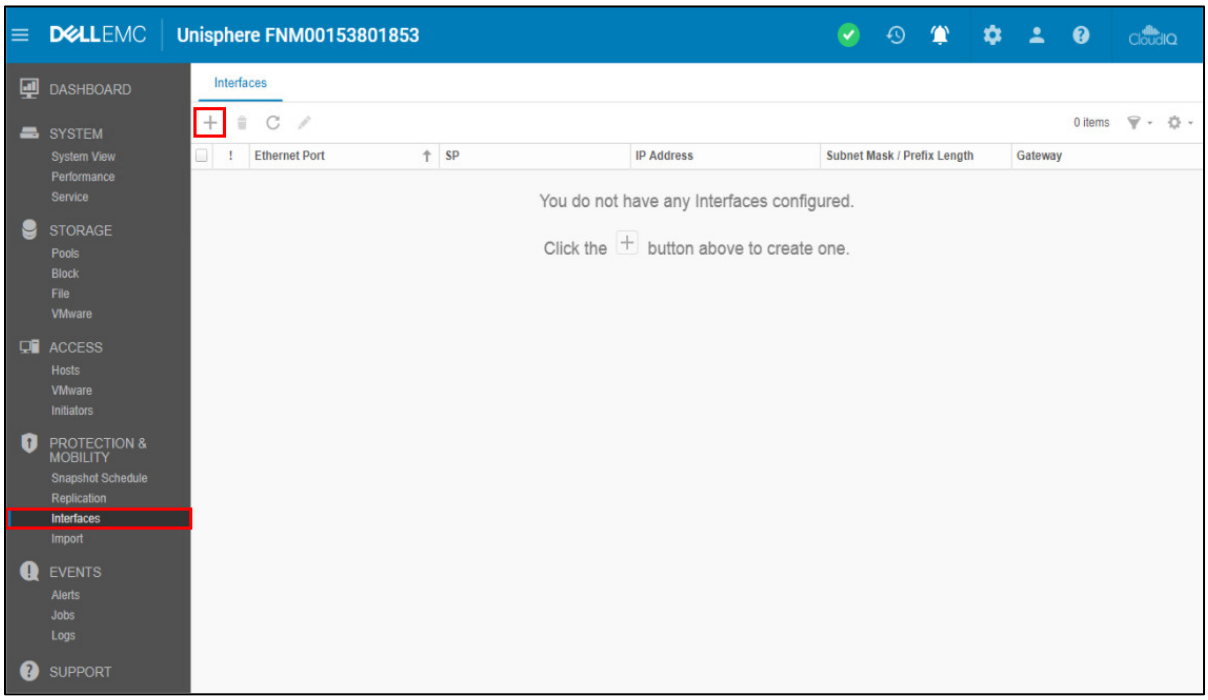


Figure 2 - Interfaces page

In the resulting window, make sure the Ethernet Port selected is “Sync Replication Management Port” and then specify the network information for each SP’s interface including IP address, subnet mask, and gateway (if necessary) as seen in Figure 3. Note that the sync replication management interfaces do not need to be on the same subnet as your management network, but they must be able to communicate over the network with the second system’s sync replication management interfaces. For example, you could utilize 192.168.1.x IP addresses for the sync replication management interfaces as seen in Figure 3. After successful configuration on the first system, repeat this step on the second system.

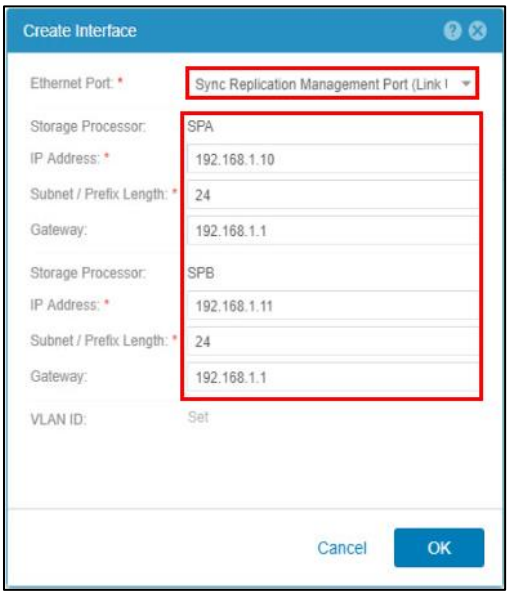


Figure 3 - Create Interface - Sync Replication Management Port

Once your sync replication management interfaces have been setup successfully, you can see the interfaces on the **Interfaces** page as seen in Figure 4. Another way to verify successful configuration is by running a simple UEMCLI show command as seen in Figure 5.

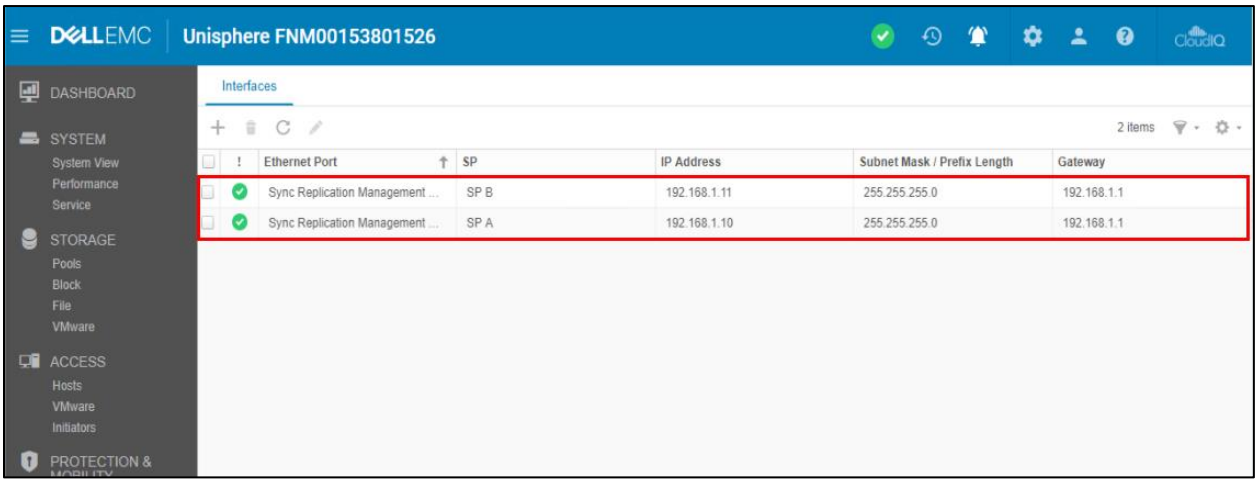


Figure 4 - Interfaces page with configured sync replication management interfaces

```
C:\>uemcli -d <ip_address> -u username -p password /net/if show -detail
2:     ID                = if_4
       Type              = replication
       Port              = spa_srm
       IP address        = 192.168.1.10
3:     ID                = if_3
       Type              = replication
       Port              = spb_srm
       IP address        = 192.168.1.11
```

Figure 5 - UEMCLI /net/if show command example

ESTABLISH FIBRE CHANNEL CONNECTION

For MetroSync on Dell EMC Unity systems, data is synchronously replicated over the first Fibre Channel (FC) port on each Storage Processor (SP). If the CNA ports on your system are configured as FC, then CNA port 4 on each SP are the designated synchronous replication ports. If the CNA ports on your system are not configured for FC, then the first port on the first I/O module on each SP configured for FC are the designated synchronous replication ports. You can identify the synchronous replication ports on the system by either checking the **System View** page in Unisphere as seen in Figure 6, or by running a simple UEMCLI command as seen in Figure 5 which is a sample output with only relevant information displayed.

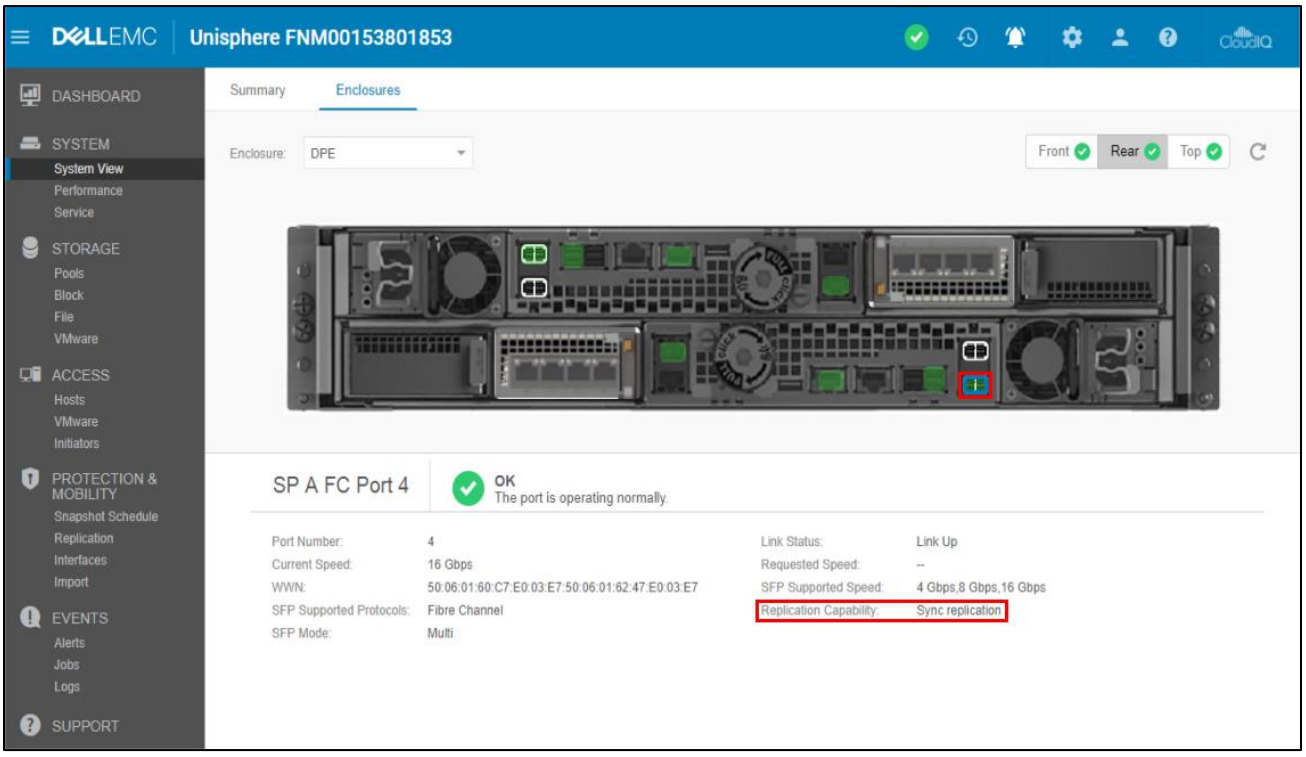


Figure 6 - System View page (Synchronous replication port)

```
C:\>uemcli -d <ip_address> -u username -p password /net/port/fc show -
detail
1:     ID                 = spa_fc4
       Replication capability  = Sync replication
4:     ID                 = spb_fc4
       Replication capability  = Sync replication
```

Figure 7 - UEMCLI /net/port/fc show command example

Once the synchronous replication ports are identified on the system, they need to be cabled with FC cables either to a switch or directly to the second system's synchronous replication ports. An example can be seen in Figure 8 with all synchronous replication ports connected to a switch. If connecting via a FC switch, make sure to zone the first system's SPA FC initiator to the second system's SPA FC initiator in one zone and create a separate zone for the SPB FC initiators. On a similar note, if you choose to directly connect the two systems, you must cable SPA to SPA and SPB to SPB. Hosts can also be zoned to the synchronous replication ports, but it is recommended to dedicate the synchronous replication ports for only synchronous replication traffic.

In the example in Figure 6, "System 1" has FC configured on the CNA ports so FC Port 4 is the synchronous replication port for that system while "System 2" has FC configured only on the second I/O module so I/O Module 1 FC Port 0 is the synchronous replication port on that system. Therefore, those identified ports on each SP have been cabled to the same FC switch in the example and zoned accordingly. For best practices, it is recommended to connect SPA FC ports to one switch and the SPB FC ports to a different switch for redundancy.

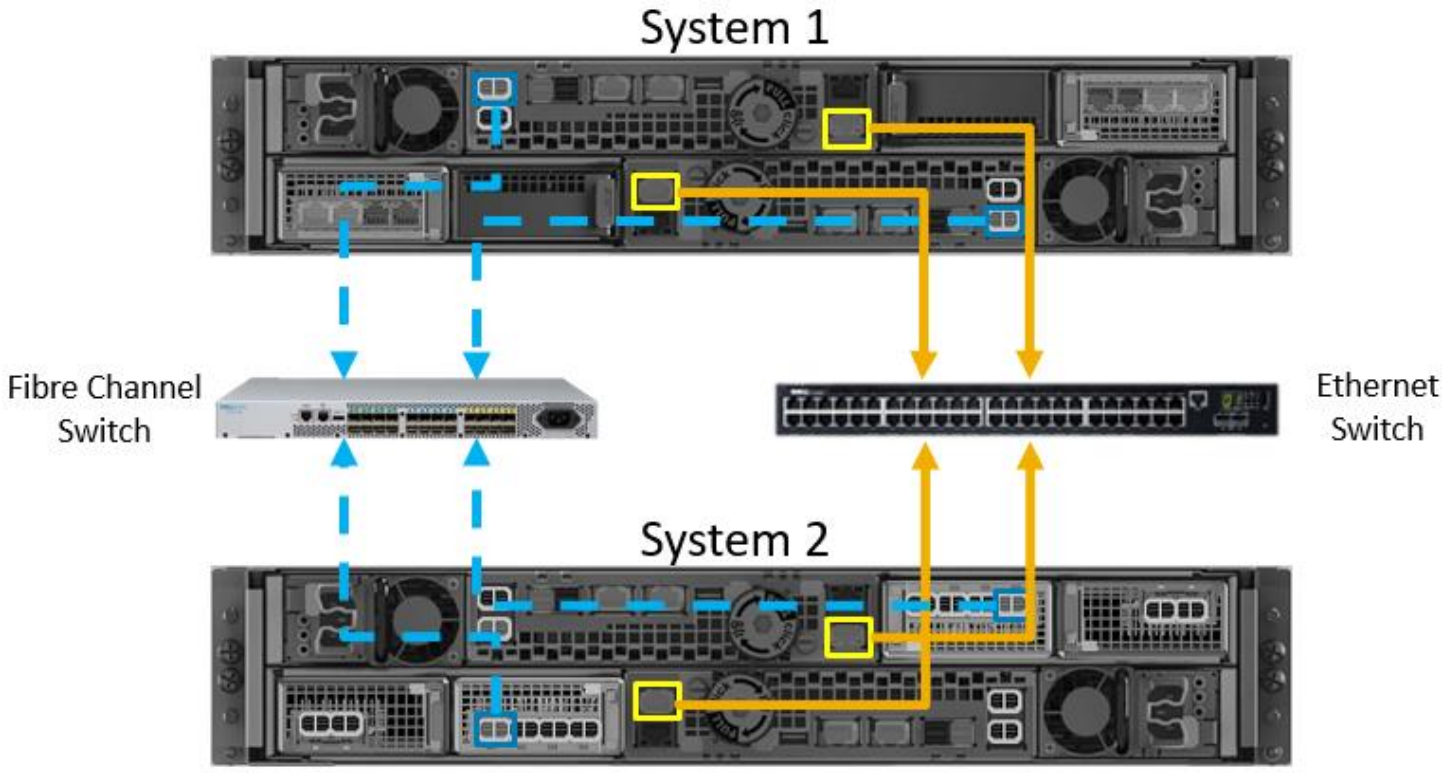


Figure 8 - Example setup with Ethernet and FC connected

CONFIGURE REPLICATION SYSTEM CONNECTION

When configuring synchronous replication, a trusted link must be created between two systems before any replication sessions can be made. A replication connection is a logical link that is created between systems that will participate in replication. The replication connection establishes a link for management and the data path between a pair of systems. After the sync replication management interfaces are configured and FC ports are connected/zoned properly, you can configure the replication system connection between the systems. Once configured, all synchronous replication sessions (include synchronous block replication sessions) utilize the same connection to transport data to another system.

To configure the replication system connection in Unisphere, navigate to the **Replication** page, click on the **Connections** tab and then click the **Add** button as seen in Figure 9.

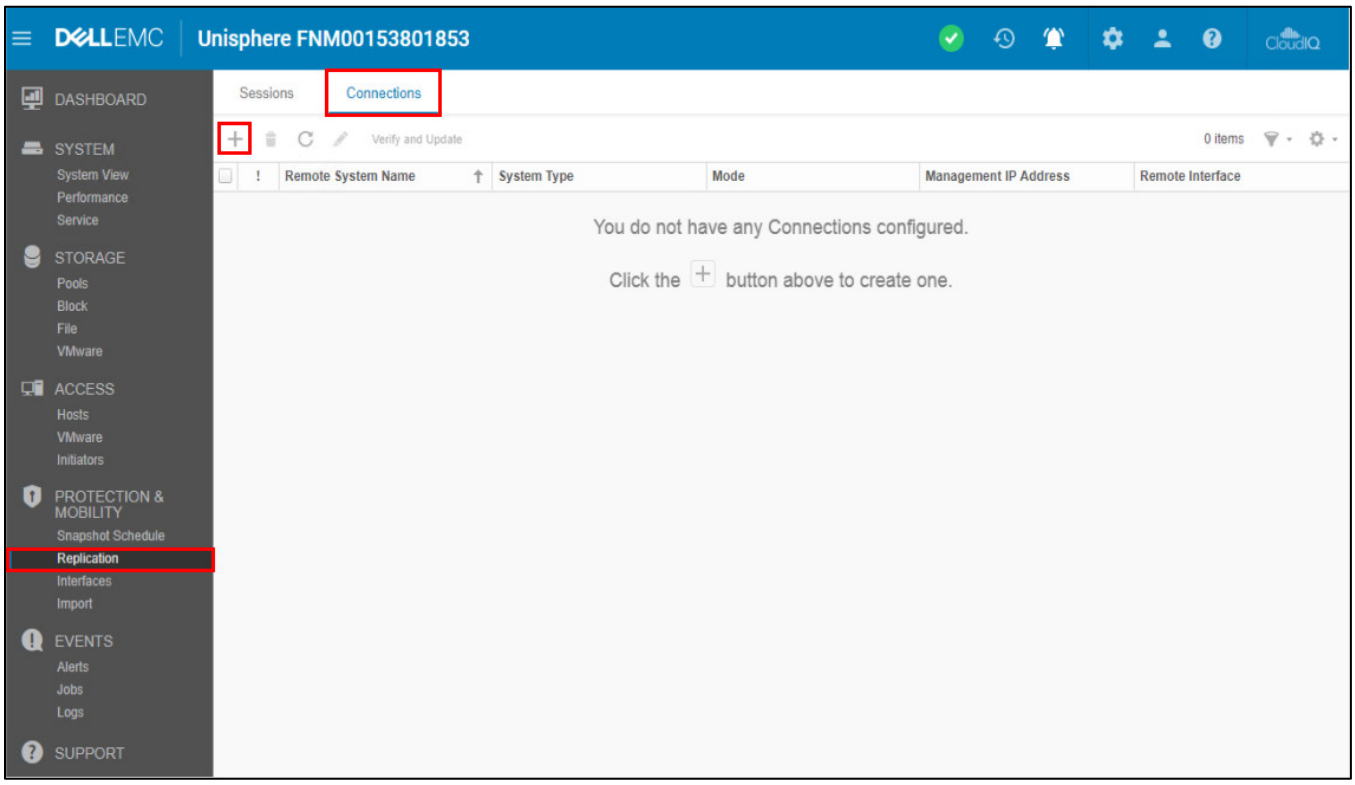


Figure 9 - Replication Page - New replication system connection

In the resulting “Create Replication Connection” window (see Figure 10), enter the management IP address of your second system, the credentials of an administrator account, and the admin password for the local system. For the Connection Mode, choose “Synchronous” if you only want to configure synchronous replication sessions between the two systems, or choose “Both” if you want to configure both asynchronous and synchronous replication sessions between the systems. Note that you must have associated asynchronous replication interfaces configured if you’d like to configure “Both” for the replication connection mode successfully. Choosing “Asynchronous” for the connection mode does not allow you to configure synchronous replication sessions. In the example, as seen in Figure 10, the mode is chosen to be “Synchronous”.

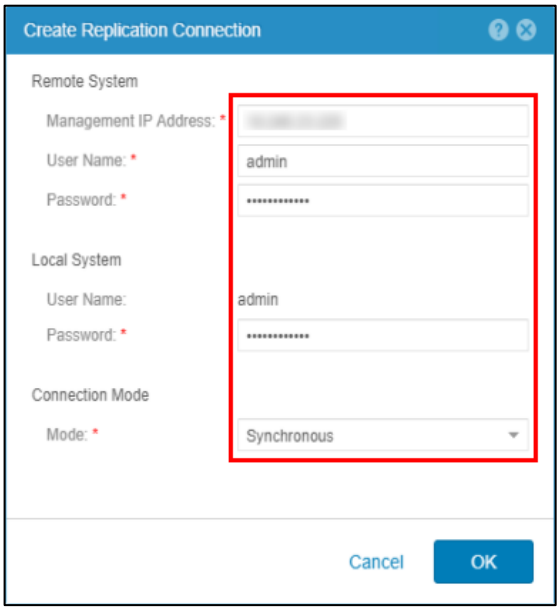


Figure 10 - Create Replication Connection

The replication connection creates a bi-directional communication path between the two systems. The system automatically configures the connection on the peer system and once complete, it becomes visible on both systems. Note that you can only create one synchronous replication connection on the system, meaning that synchronously replicated systems are always configured in pairs. It is not possible to configure multiple systems as the target for multiple synchronous replication sessions. This is different than asynchronous replication, which allows for multiple asynchronous connections per system.

Once a replication connection is successfully made, you'll notice there is an available "Verify and Update" button when selecting a given replication connection as seen in Figure 11. This operation can be used to test a replication connection to the peer system and update the replication connection information if changes have been made (i.e. synchronous replication management port IP address change). The operation can also be used to re-establish a replication connection to the peer system after a power down or outage scenario.

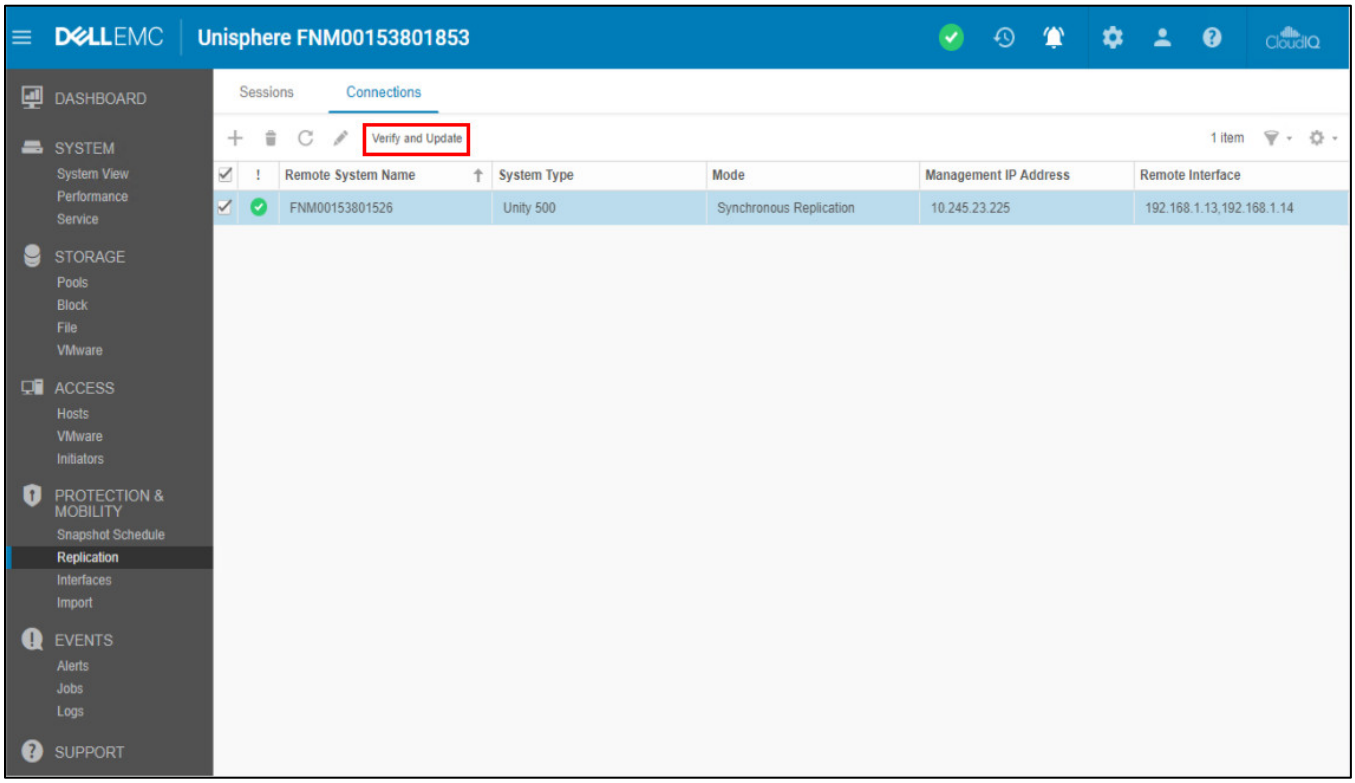


Figure 11 - Verify and Update

CONFIGURE SYNCHRONOUS REPLICATION SESSIONS

A synchronous replication session utilizes a configured replication connection to transfer data from a source resource to a destination resource. Replicating locally to the same system for synchronous replication sessions is not supported. If a synchronous replication session is created on an existing resource using Unisphere, a storage resource of the same size and type is created on the peer system automatically and a full synchronization occurs between the source and newly created resource. If replication is configured during the creation of a brand new resource, then the initial synchronization is quick as no user data needs to be copied over the link.

When configuring replication, you must first replicate the NAS Server before replicating the underlying file systems or NFS datastores. When utilizing the GUI to configure replication, the system automatically configures replication for all related resources when configuring replication for a NAS Server, saving the administrator time and effort. You can always delete replication on a per resource basis after the fact. Note that if an existing asynchronous replication session already exists and you'd like to use synchronous replication instead to the same system, you would need to first delete the existing asynchronous replication session and then setup synchronous replication, which requires a full synchronization.

During a full synchronization, incoming new host writes to the source are automatically replicated to the destination if the data block has already been copied to the destination as part of the synchronization process. If the data has yet to be synchronized, then only the source resource acknowledges the write and the data is copied at a later point once the synchronization process gets to that data block.

This document does not go over the exact steps of configuring replication. Please see the *Dell EMC Unity Family Configuring Replication* technical guide for step-by-step configuration instructions.

The following steps outline a write operation to a storage resource with a synchronous replication session configured. In this example, assume the initial synchronization is complete.

1. A write I/O is sent to a storage resource on the production system.
2. System cache on the production system accepts the write I/O.
3. The production system replicates the data to the peer system.
4. The peer system accepts the data into system cache.
5. The peer system responds to the Production System and acknowledges the write has been saved.
6. The production system acknowledges the host that the data has been accepted and saved on the system.

Replication sessions configured with synchronous replication can have different synchronization states which describe the current state of user data on the destination storage resource. These states can be useful for troubleshooting. For the latest information on the synchronization state of a replication session, refer to the status on the source system since the destination system status may not update immediately. Below are the available synchronization states and their associated descriptions.

- **Consistent** – The data on the peer system is consistent from the host perspective and I/O has recently occurred. The destination resource is the latest or point in time copy of the source resource.
- **In Sync** – The data on the peer system is consistent from the host perspective. The destination resource is an exact copy of the source resource.
- **Syncing** – The data on the peer system is getting updated from the source system. This can be a full synchronization or just the pending changes from the source system.
- **Out of Sync** – The state of the data on the destination resource cannot be determined by the replication session. The update to the peer system may be incomplete.
- **Inconsistent** – The state is only reported when a replication session is failed over. It appears when the synchronization state of the session was not “In Sync” or “Consistent” prior to failover. In this case, it is recommended that you check for the consistency of the destination storage resource.

SUPPORTED STORAGE RESOURCES

When configuring replication on Dell EMC Unity systems, the source and destination storage resource must be of the same type and size. MetroSync is supported on the following storage resources:

- NAS Servers
- File Systems
- VMware NFS Datastores

Once synchronous replication is configured, any host writes to File Systems or NFS Datastores are first copied to the peer system over the replication connection before the host is acknowledged, thereby ensuring a zero RPO disaster recovery capability when failover is needed. Most configuration changes made to a source NAS Server are also propagated to the peer system. For example, adding additional production IP interfaces or applying different sharing protocols makes the same change immediately to the destination. For file systems and NFS datastore settings, you can change the size (extend or shrink) of the source storage resource which makes the same change to the destination storage resource automatically. Data reduction is not propagated to the destination automatically and you can enable/disable on a per resource basis, if the configuration supports data reduction. Also, during creation of a replication session in the GUI, you may choose for the destination resource to have data reduction enabled even if the source resource does not have data reduction enabled.

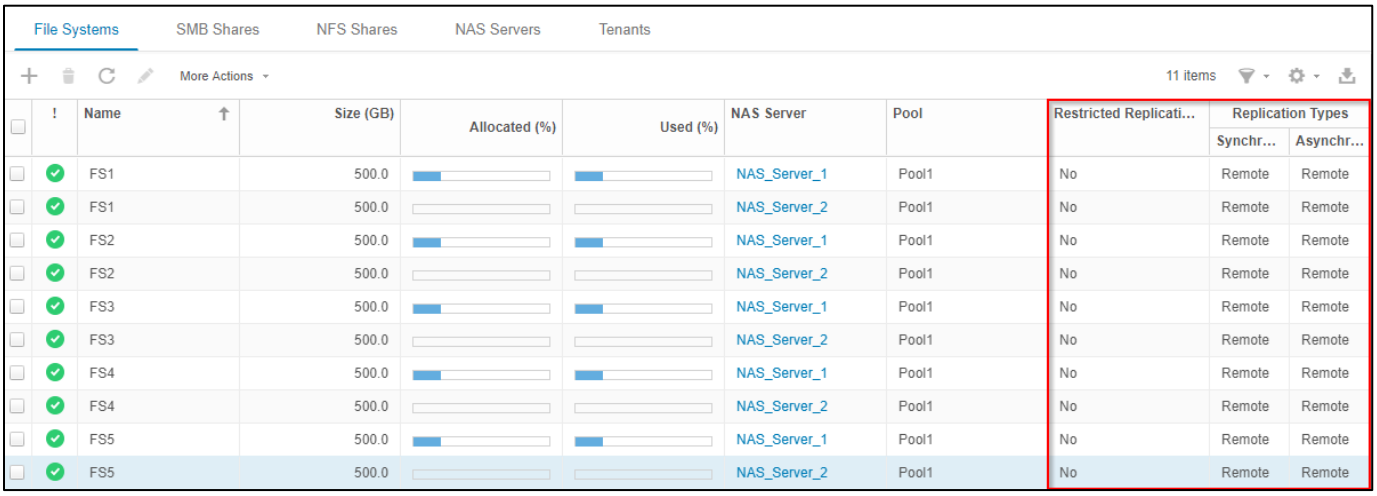
REPLICATION ROLES

At least two storage resources are required for replication:

- A source storage resource which will be replicated
- A destination storage resource which is copied to from the source and is not directly host accessible

When a replication session is created in Unisphere, the destination storage resource is automatically created with the session. Upon creation, the destination resource is marked as a destination image. This means host access is restricted on the destination storage resource. To view the data contained in a destination storage resource, you may take a Snapshot of the resource and view the data through various methods like Backup and DR interfaces or Proxy NAS Servers. For more information and details on those methods, see the *Dell EMC Unity: DR Access and Testing* white paper on Dell EMC Online Support.

In Unisphere, you can easily determine which storage resources are replicated and which are designated as destination resources from any of the related storage resource pages such as the **File Systems** tab on the **File** page. While in this tab, select the gear icon, and hover over the Columns option to view the available columns that can be viewed. Select the check boxes for "**Replication Types**" and "**Restricted Replication Access**". An example of the page with optional columns selected can be seen in Figure 12.



!	Name	Size (GB)	Allocated (%)	Used (%)	NAS Server	Pool	Restricted Replicati...	Replication Types	
								Synchr...	Asynchr...
✓	FS1	500.0			NAS_Server_1	Pool1	No	Remote	Remote
✓	FS1	500.0			NAS_Server_2	Pool1	No	Remote	Remote
✓	FS2	500.0			NAS_Server_1	Pool1	No	Remote	Remote
✓	FS2	500.0			NAS_Server_2	Pool1	No	Remote	Remote
✓	FS3	500.0			NAS_Server_1	Pool1	No	Remote	Remote
✓	FS3	500.0			NAS_Server_2	Pool1	No	Remote	Remote
✓	FS4	500.0			NAS_Server_1	Pool1	No	Remote	Remote
✓	FS4	500.0			NAS_Server_2	Pool1	No	Remote	Remote
✓	FS5	500.0			NAS_Server_1	Pool1	No	Remote	Remote
✓	FS5	500.0			NAS_Server_2	Pool1	No	Remote	Remote

Figure 12 - File Systems Page - Optional Columns

Replication Types displays what type(s) of replication the storage resource is participating in, whether it is None, Remote, or Local. The Restricted Replication Access column displays "**Yes**" if the storage resource is labeled as a replication destination resource. If a replication session is deleted, the storage resources themselves are not deleted. For the destination resource, the replication destination designation must be edited manually via Unisphere CLI before the resource is allowed to be accessed from hosts and receive I/O. For example, to remove the replication destination designation setting for a file system, use the `uemcli /stor/prov/fs { -id <value> | -name <value> } set -replDest no` command. For more information on removing the replication destination designation setting on other storage resource types, consult the *Dell EMC Unity Unisphere Command Line Interface User Guide* on Dell EMC Online Support.

REPLICATION OPERATIONS

Once a synchronous replication session is created, a number of operations are available to modify the state of the replication sessions from the Unisphere GUI. Not all operations are available at all times as some depend on the current particular state of the session. Also, certain operations perform differently depending on which system they are issued from, source or destination. Only one replication operation can be issued and be running per session at any particular point in time.

Failover

When issuing a failover operation, the destination resource becomes the production resource and is available for host Read/Write operations while the original source is no longer available for host access. The effects of issuing the failover operation depends on which system the failover was initiated from.

Planned Failover

A failover issued from the source system is also referred to as a planned failover. It is highly recommended to quiesce I/O to the source resource first before running a planned failover. After running a planned failover, the original destination resource becomes host accessible and the direction of replication reverses to the original source. When this occurs, the original destination resource starts replicating all new writes it receives to the original source resource automatically. Issuing a planned failover from the source is recommended when testing a site failover to ensure the DR configuration is working properly. To return the session back to the original setup before the failover, first quiesce I/O to the source

resource, then issue a planned failover from the source (original destination). This makes the original source resource the source again and reverses replication to replicate back to the original destination.

Unplanned Failover

A failover issued from the destination system is also referred to as an unplanned failover and is usually only done in disaster situations where the source is no longer available and/or not recoverable. An unplanned failover operation assumes an actual disaster has occurred on the source system and the destination needs to be brought up in read/write mode as production meaning it is now host accessible and can accept new writes. Note that the replication role remains "**Destination**" for the new production resource and "**Source**" for the original source system, but since replication is no longer occurring, this information is not relevant. You can view the details of the replication session to verify whether the resource is accepting host I/O locally.

If the source system becomes temporarily unavailable, such as due to a power outage, an unplanned failover can be initiated to enable data access during the outage. However, once power is restored and the original source system boots back up, this could lead to a duplicate IP address on the network. MetroSync includes a duplicate IP avoidance mechanism to prevent against this situation. While the original source system is booting up, it communicates with the peer system over the synchronous replication management interface and checks the status of the replication sessions. If the replication sessions are failed over, it keeps the interfaces on the original source system in an offline state to avoid creating a duplicate IP scenario. Note that this mechanism may not work in certain situations such as if failover is initiated while the source system is still available or if the synchronous replication management network is unavailable.

After an unplanned failover has occurred, the replication session is effectively broken and a full synchronization is required to restart replication. If the original source system is recovered and is available again after an unplanned failover is completed, users have the option to run a *Failback* or *Resume* operation from the new production resource (original destination). See the related sections below for details for those operations.

If an unplanned failover tries to be performed while the source system is still available, a warning message appears prior to applying the operation stating that the network status of the source system is seen as "**OK**" and recommends the user to failover from the source storage resource instead as seen in Figure 13.

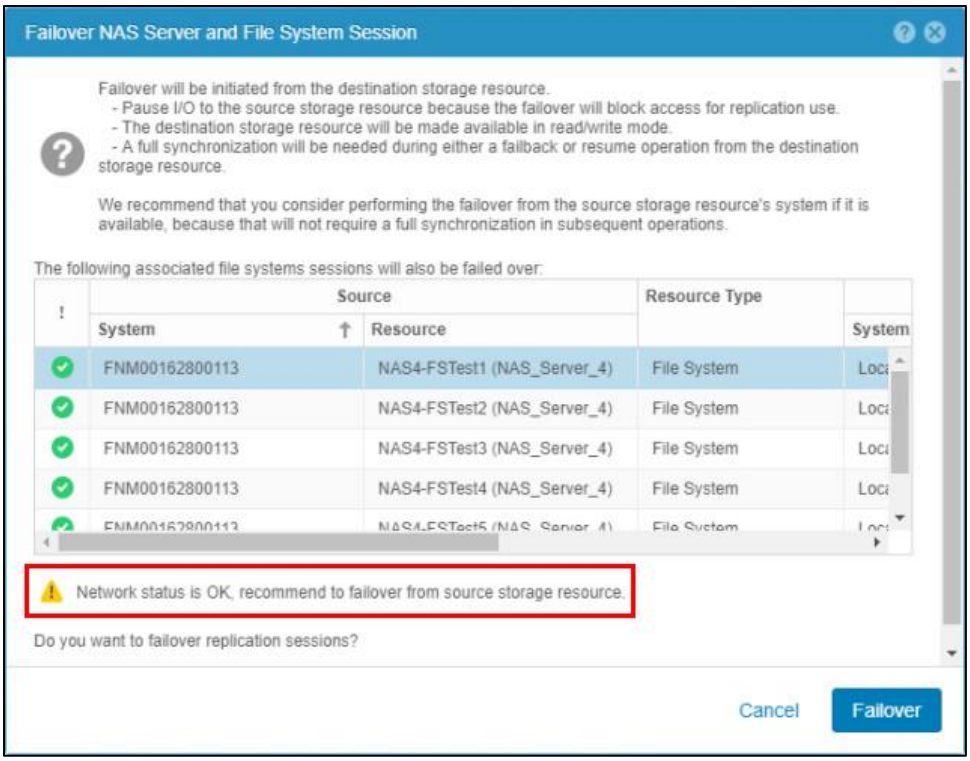


Figure 13 - Failover operation from destination NAS Server (with source system still available)

Failback

On Dell EMC Unity systems, a failback operation is only issued after an unplanned failover operation. Since an unplanned failover operation stops replication, running a failback first initiates a full synchronization of the production resource. Once the synchronization is complete, the system then switches the roles of replication to make the original source and

destination reprise their roles again. From the standpoint of the original setup before an unplanned failover, the source is actually the original source, so a failback is essentially resetting back to the original replication paradigm before the unplanned failover occurred.

Pause

A pause operation is only issued on the replication session from the source system. Once a pause operation is completed, all writes to the storage resource are only saved on the source system before the host is acknowledged. All changes on the source storage resource are tracked while the replication session is paused so that those changes can be propagated to the destination once a *Resume* operation is performed. Pausing a replication session may be done for a number reasons including the need to power off the peer system for planned maintenance, a configuration change on the network on either system, or maybe to physically move the system from one data center to another. In certain circumstances, configuring replication and synchronizing data between systems can be done within the same site, then the peer system is later moved to its final destination site. Another use case for pause is for system software upgrades. It is recommended to pause all replication sessions including file synchronous replication sessions prior to an upgrade to mitigate potential issues during the upgrade process.

Resume

A resume operation is performed on the production system either after an *Unplanned Failover* or *Pause* operation is completed.

After an unplanned failover, running a resume operation initiates a full synchronization to the original source system before the unplanned failover operation occurred and then continues replicating new writes in the same direction. This operation effectively reverses the direction of replication from the standpoint of the original setup prior to the unplanned failover. A user can later run a planned failover from the new source system to switch the roles of replication to get back to the original setup and original replication paradigm.

After a pause operation, running a resume simply copies over any outstanding writes since when the pause operation was initiated and then continues syncing new writes per normal replication operations.

Delete

Deleting replication sessions for synchronous file sessions must be first be performed on underlying file systems and NFS datastores before the replication session for the associated NAS Server can be deleted.

Deleting a replication session can be issued on the source system or destination system, but it is recommended that the operation be issued on the source when the source is available. When the configuration is in a healthy state and a delete operation is issued on the source system, the replication session is deleted from both the source and destination system. If the delete operation is issued while the destination system cannot be reached, the session needs to be deleted from the destination system manually. If the delete operation is issued from the destination system, the source replication session is left configured and must be deleted manually. Once a replication session is deleted, a full synchronization needs to occur if replication is reconfigured.

Note that the storage resource, source or destination, is not deleted when the associated replication session is deleted. If a synchronous replication session is deleted before the initial synchronization completes, the destination file systems returns an unrecoverable error. In this case, the destination file system cannot be used and should be deleted.

A delete operation can also be issued for a replication connection. A replication connection can only be deleted after all related replication sessions utilizing the connection have been deleted.

Group Operations

Group operations at a NAS Server level are supported for file synchronous replication sessions. Group operations allow for a given replication operation to be propagated from the NAS Server to its underlying File Systems automatically which saves management time for system administrators. Group operations are available for the following operations: Failover, Failback, Pause, and Resume. All other operations (Create, Delete, and Modify) are individual commands.

Do not perform a group operation at both sides of a replication session at the same time. This action is not prohibited by the storage system, however, a group performed at the same time at both sides of a replication session can cause the session to enter an unhealthy state. Although a group operation looks like one operation, each File System is replicated individually. Group operations skips File System replication sessions that are in a paused, error, or non-replicated state. If any of the individual File System replication sessions fails during a group operation, you can resolve the issue and then select the individual File System to complete the necessary replication operation.

DATA PROTECTION MECHANISMS

Synchronous replication has mechanisms to resync data differences between the source and/or destination resources in the event of replication disruption thereby preventing the need for full synchronization. The fracture log protects primarily against loss of communication with the destination resource. The write intent log protects primarily against interruptions to the source resource. Both of these structures exist to enable partial synchronizations in the event of interruptions to the source or destination resources.

Fracture log

The fracture log is a bitmap held in the memory of the storage processor that owns the source resource. It indicates which physical areas of the source have been updated since communication was interrupted with the destination.

The fracture log is automatically invoked when the destination resource of a replication session is lost for any reason and becomes out of sync. The replication session is out of sync (no longer replicating) if the destination is not available, or it can be administratively paused through Unisphere or UEMCLI. Dell EMC Unity sets a replication session as out of sync if an outstanding I/O to the destintation is not acknowledged within 25 seconds. While in a state of out of sync, the source pings the destination every 20 seconds to determine if communication has been restored.

The fracture log tracks changes on the source resource for as long as the destination resource is unreachable. It is a bitmap that represents areas of the source resource with regions called extents. The amount of data represented by an extent depends on the size of the data resource. Since the fracture log is a bitmap and tracks changed areas of the source resource, it is not possible to run out of fracture log capacity.

When the destination resource returns to service, it must be synchronized with the source. This is accomplished by reading those areas of the source addressed by the fracture log and writing them to the destination resource. This activity occurs in parallel with any writes coming into the source and replicated to the destination. Bits in the fracture log are cleared once the area of the source marked by an extent is copied to the destination. This ability to perform a partial synchronization can result in significant time savings. It may be necessary, depending on the length of the outage and the amount of write activity, to resynchronize the entire dataset.

By default, the fracture log is stored in memory. Therefore, it would be possible for a full resynchronization to be required if a destination resource is out of sync and an interruption in service occurs on the source SP. To protect against such scenarios, the write intent log is used.

Write intent log

The write intent log is a record stored in persistent memory (disk) on the storage system on which the source resource resides. During normal operation, the write intent log tracks in-flight writes to both the source and destination resources in a sync replication relationship. Much like the fracture log, the write intent log is a bitmap composed of extents indicating where data is written. The write intent log is always active, but the fracture log is only enabled when the replication session is out of sync.

When in use, Dell EMC Unity makes an entry in the write intent log of its intent to update the source and destination resources at a particular location, then proceeds with the attempted update. After both images respond that data has been written (i.e. written to write cache), the system clears previous write intent log entries. For performance reasons, the write intent log is not cleared immediately following the acknowledgement from the source and destination resources. It will be cleared while subsequent write intent log operations are performed.

In a recovery situation, the write intent log can be used to determine which extents must be synchronized from the source storage system to the destination system. For instance, if a single SP becomes unavailable (for example during a reboot or failure), there may be in-flight writes that were sent to the destination, but not acknowledged before the outage. These writes will remain marked in the write intent log. Then server software trespasses the resource to the peer SP. The remaining SP directly accesses the unavailable SP's write intent log and recovers the recent modification history. The SP then resends the data marked by the extents in the write intent log. This allows for recovery using only a partial resynchronization, rather than a full resynchronization because it ensures that any writes in process at the time of the failure are acknowledged by the destination resource. If the entire array becomes unavailable, then the write intent log is used to facilitate a partial resynchronization from source to destination, once the source array is recovered.

SUPPORTED REPLICATION CONFIGURATIONS

At the system-level, Dell EMC Unity systems support both uni-directional and bi-directional replication topologies. For uni-directional, this means you can have all replicated source resources on one system while all destination resources are on the other system. When using bi-directional replication topology, you have some production resources on one system with

other production resources on the other system, while using each other as the destination. Bi-directional is typically utilized when production I/O needs to be spread across multiple systems or locations.

With the introduction of MetroSync in OE version 4.4, multiple different file replication topologies are now supported, giving users flexibility in configuring data protection/backup schemes for file resources. The following topologies can be configured on a given file storage resource (Note this is specific to file resources and does not apply to block resources):

- One asynchronous replication session to another system or locally within the same system
- One synchronous replication session to another system
    - Local replication within the same system is not supported for synchronous replication
- One synchronous replication session to 2nd system and one asynchronous replication session to 3rd system
    - The asynchronous replication session must be configured to different system than the system used for synchronous replication
    - Local replication within the same system is not supported with the dual replication configuration

This means a single file storage resource can be replicated up to two different systems (one synchronously and one asynchronously), which provides additional backup and expands data protection use cases for Dell EMC Unity systems. Cascading replication is not supported.

For more information on configuring dual replication, see the *Asynchronous Replication to a 3rd Site* section of this document.

SNAPSHOT REPLICATION

As a part of the MetroSync solution for Dell EMC Unity, snapshot replication provides the ability to replicate read-only file snapshots to the peer system to ensure consistent snapshots on both sites. Read-write snapshots are not replicated. Snapshot replication is automatically enabled for all resources that have file synchronous replication enabled which includes File Systems and VMware NFS Datastores. Both scheduled snapshots and user created snapshots are automatically replicated. This means anytime a read-only snapshot is created for the source resource, whether by snapshot schedule (local or synchronously replicated) or manually by a user, the snapshot is automatically replicated to the destination with the same properties as the source snapshot such as retention policy and snapshot name. The destination snapshot can later be modified to be a different retention policy than the source as necessary. Note that changes to the source snapshot including modification of the retention policy, or deleting the snapshot, automatically updates the destination snapshot, even if it was previously modified.

When the synchronous replication session is not in an "**In Sync**" or "**Consistent**" state, snapshots are not replicated. These snapshots cannot be replicated later, even if the session returns to a replicated state. Taking a manual user snapshot from the destination system is supported, but this does not replicate the snapshot back to the source system.

Storage resources that are actively participating in a synchronous replication session cannot be restored to a snapshot. In order to perform a snapshot restore, the file system synchronous replication session must first be deleted. Deleting the replication session does not delete the source file system or snapshots. Once the restore is complete, the file system synchronous replication session can be created. Note that this requires a full synchronization. An alternative to this is to access the snapshot and perform a manual restore of the files or folders to the production file system.

SNAPSHOT SCHEDULE REPLICATION

One of the advanced features for the MetroSync solution is snapshot schedule replication. This feature allows users to replicate their snapshot schedules to the destination to ensure consistent schedules on both sites as well as apply the replicated snapshot schedules to file synchronous replication resources. Snapshot schedule replication utilizes the same replication management connection as synchronous replication. Making any changes to a replicated snapshot schedule, whether on the source or destination system, automatically makes the same change to the peer system.

Applying a replicated snapshot schedule is only allowed on synchronously file replicated resources. By applying a replicated snapshot schedule to an associated resource, that resource automatically applies the same schedule to the destination. Note that the destination system's snapshot schedule is not active for synchronously file replicated resources. Only after a failover occurs does the destination system's snapshot schedule become active and continue snapshot operations. Therefore, only the source snapshot schedule is ever active for synchronous file replicated resources, but due to the snapshot replication feature, all scheduled snapshots are automatically replicated to the destination anyways. In lieu of a replicated snapshot schedule, a local snapshot schedule could be applied to the source system, but that schedule is not replicated to the destination system, meaning no schedule would be set after a failover operation.

Therefore, it is recommended for users to utilize synchronously replicated snapshot schedules for synchronously file replicated resources.

To create a new synchronously replicated snapshot schedule, navigate to the **Snapshot Schedule** page under **Protection & Mobility** in Unisphere. After clicking the **Add** button, the corresponding pop-up window, as seen in Figure 14, provides the available options for scheduling as well as a checkbox stating "**Synchronize snapshot schedule to remote system**". Clicking this checkbox and creating the schedule automatically replicates the same schedule to the peer system given that the associated replication connection is in a healthy state. A similar dialog can be found if creating a new schedule during the file system creation workflow.

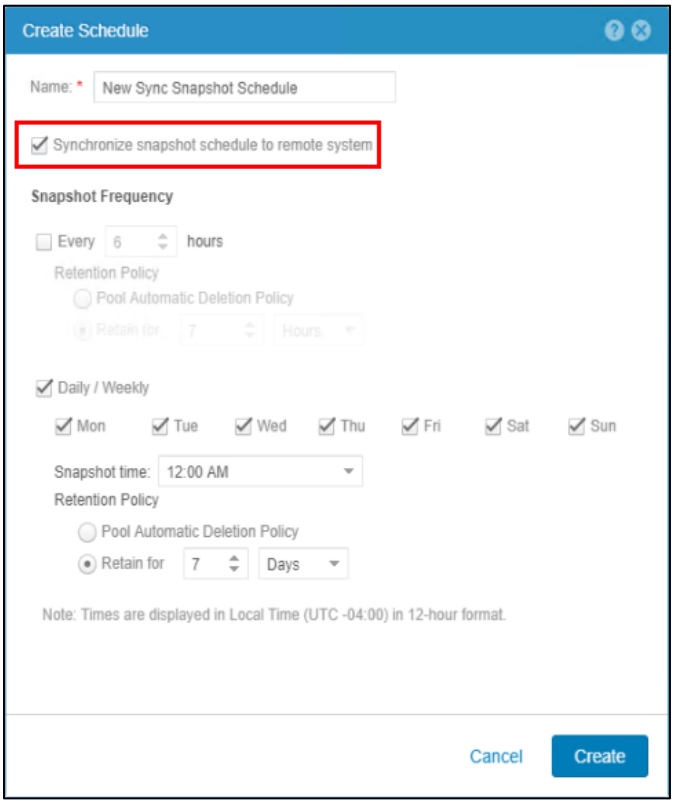


Figure 14 - Create Schedule window (Synchronize schedule option)

When creating a new file system on a synchronously replicated NAS Server, Unisphere only allows users to choose a synchronously replicated snapshot schedule during the **Snapshots** step of the wizard as seen in Figure 15. This is only a restriction during the file system creation workflow and can be modified afterwards in the properties of the file system.

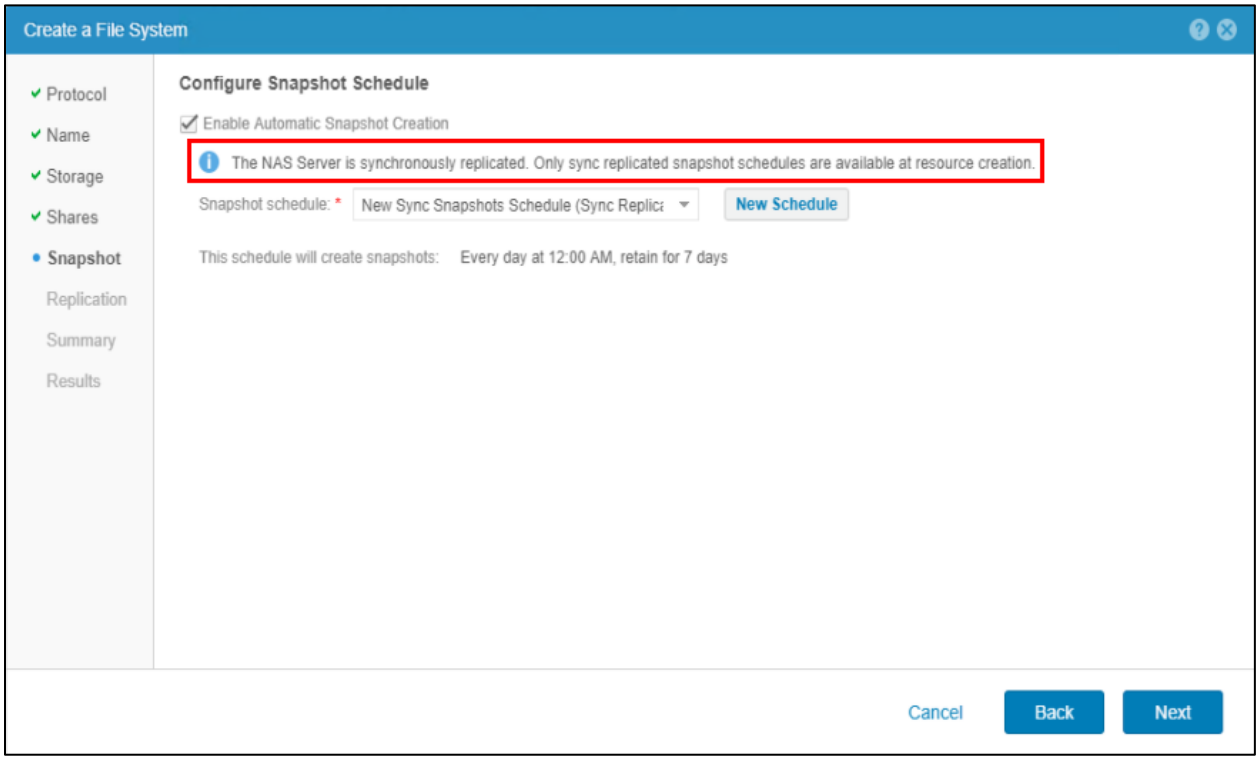


Figure 15 - Create a File System Wizard - Snapshots step (Only synchronous replication schedules selected during creation)

ASYNCHRONOUS REPLICATION TO A 3RD SITE

Another advanced feature of the MetroSync solution is asynchronous replication to a 3rd site. This feature allows users to create an asynchronous replication session on top of an existing synchronous replication session to a 3rd system separate from the first two systems. Also, if a user has an existing asynchronous replication session today, then they may configure a new synchronous replication session without breaking any existing settings. Configuring this additional replication session provides a backup/DR solution in case a disaster happens to the two source systems simultaneously. Note that this feature is limited to one synchronous replication session and one asynchronous replication session for a given source storage resource and is only applicable to file resources. Also, the asynchronous replication session must be to a system running OE version 4.4 or later and not a local loopback to the same system. Note that the 3rd system for asynchronous replication destination is not required to be a physical system and can also be a Dell EMC UnityVSA system.

To illustrate an example setup, let's say there are three sites: Site A, Site B, and Site C. Site A and Site B are our metro distance sites where a synchronous replication connection is created between them and Site C is a third site that is located farther away and utilizes asynchronous replication connections back to the main sites. Once those connections are setup successfully, you can create the replication sessions (both synchronous and asynchronous) for the NAS Server and associated file resources from the source site, in this case Site A, to their respective sites as depicted in Figure 16. Note that a separate NAS Server could be configured across Site A and Site B and then use the same system for the 3rd site (Site C) or different system for the 3rd site for the asynchronous replication session (i.e. Site D).

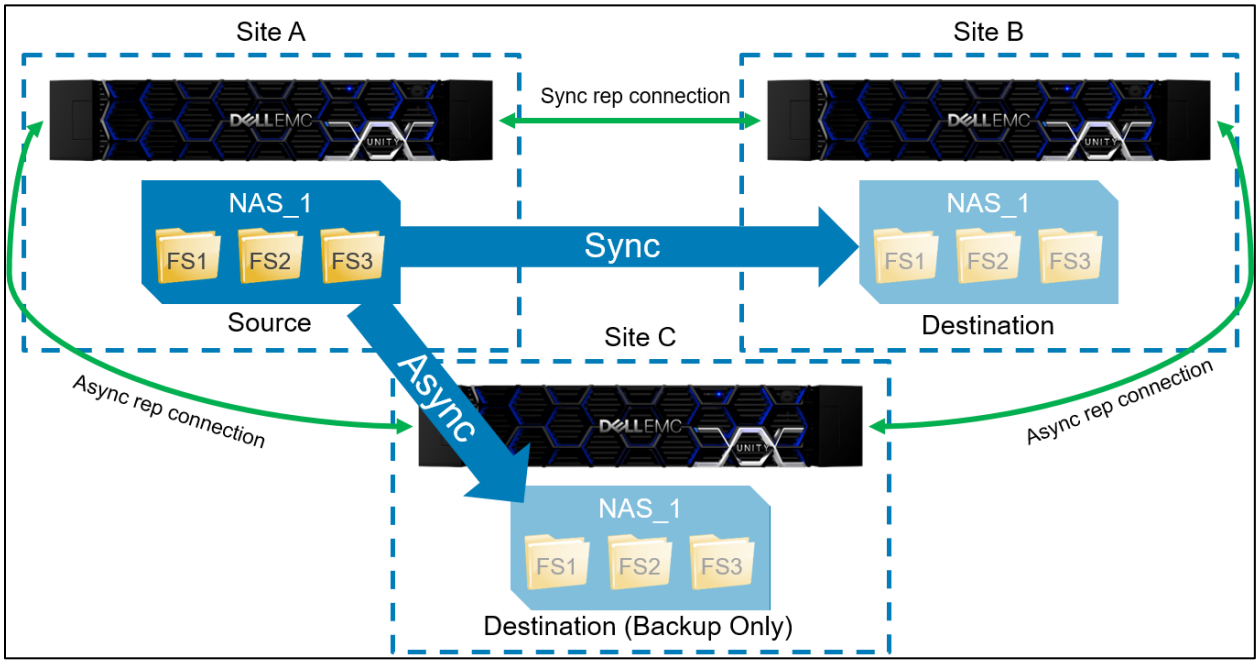


Figure 16 - Illustration for synchronous and asynchronous replication sessions

Create Options

The following sections show some example creation workflows and provides some recommendations for setting up both asynchronous and synchronous replications sessions. For a step-by-step guide on setting up replication sessions, reference the *Configuring Replication Guide* on Dell EMC Online Support.

In order to configure Asynchronous Replication to a 3rd Site, the following Replication Connections must be configured:

- Site A ⇔ Site B – Synchronous or Both
    - Select Both if you also plan on creating regular asynchronous replication sessions between these two systems
- Site A ⇔ Site C – Asynchronous
- Site B ⇔ Site C – Asynchronous

It is important to ensure all of these connections are in place prior to configuring any asynchronous replication sessions. Otherwise, the Preserve operation, used to restart the asynchronous replication sessions, may not work properly. For example, if the connection between Site B ⇔ Site C is not configured, the asynchronous replication sessions cannot be restarted on Site B in the event of a failover.

As can be seen in Figure 16, all replication connections should be set up between each of the sites as a part of initial setup. Once those connections are setup properly, then you can start creating related replication sessions. When creating a brand new NAS Server on OE version 4.4 code or later, the replication step has two separate replication options that are independently selectable. An example can be seen in Figure 17 which shows both synchronous replication and asynchronous replication selected during NAS Server creation workflow on the Replication step.

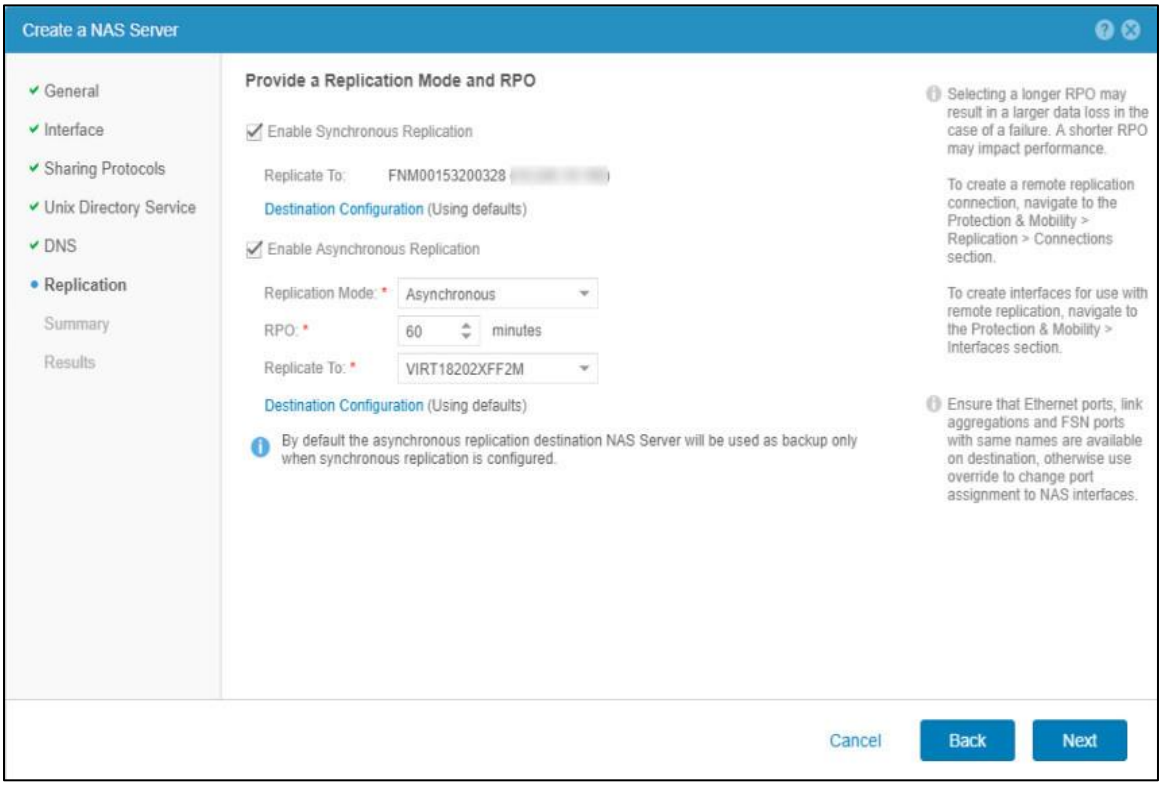


Figure 17 - Create NAS Server workflow - Replication step

If you have an existing NAS Server and file systems that are already being either asynchronously or synchronously replicated, you can setup the other replication type without disrupting the current replication configuration. Like before, once you have the necessary replication connections setup between the three separate sites, you can create a new replication session in addition to the existing replication session. An example can be seen in Figure 18 and Figure 19 where there was an existing synchronous replication session and then a new asynchronous replication session is added. A similar workflow can be done with an existing asynchronous replication session. Note that Unisphere automatically detects which replication session can be added and won't allow you to add a replication session of the same type as the existing session.

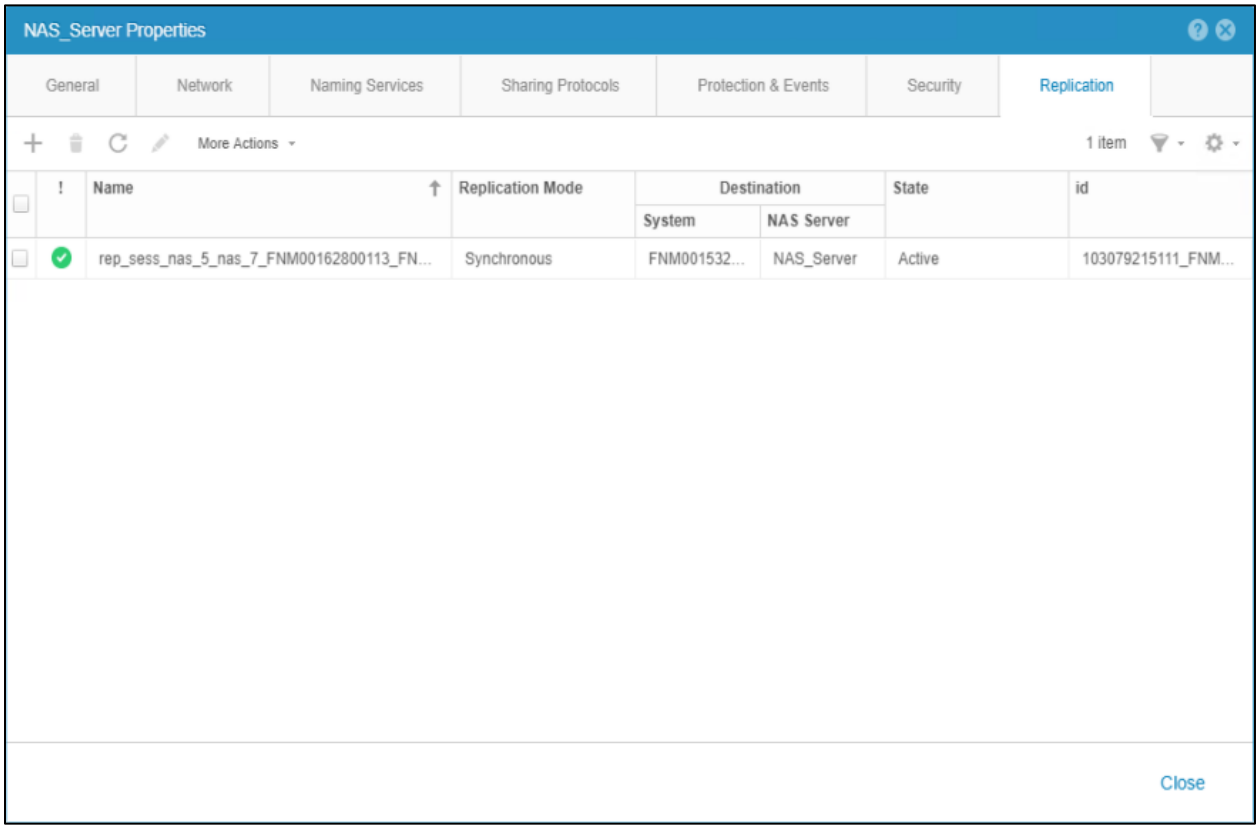


Figure 18 - NAS Server with existing synchronous replication session

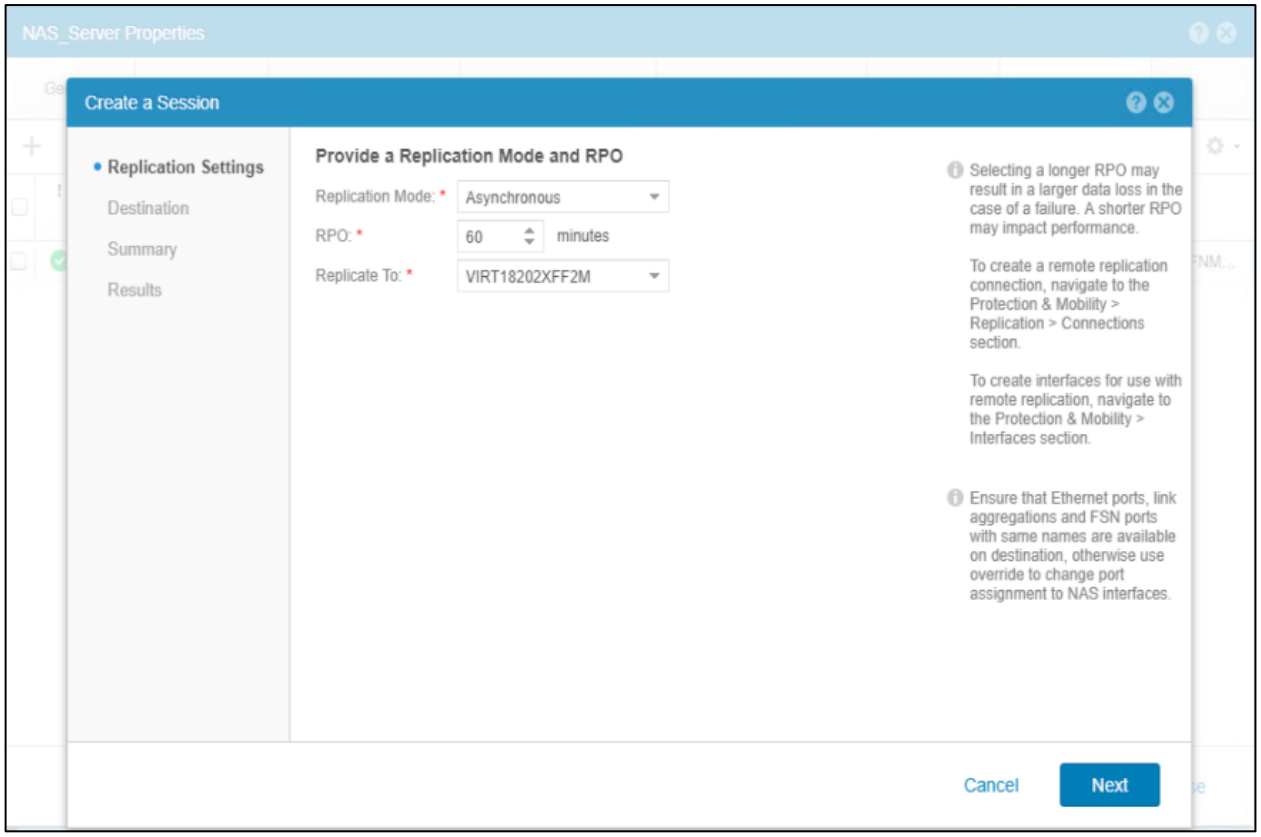


Figure 19 - Creating new asynchronous replication session on top of existing synchronous session

Preserve Operation

Starting with the example in Figure 20, in the event that the NAS Server and file systems on Site A need to failover to Site B, the asynchronous replication sessions from Site A to Site C are no longer active and is seen as "**Hibernated**". In this case, asynchronous sessions should be restarted from Site B to Site C to continue backup operations by using a *Preserve* operation. Note that this Preserve operation does not require a full synchronization of data since the internal common base snapshots are replicated along with the synchronous replication session between Site A and Site B. Therefore, the system can utilize these snapshots as a common base to synchronize any changes since the last refresh (based on the RPO). An example illustration of the setup after running a preserve command can be seen in Figure 20. Notice that Site B is now the source and is now asynchronously replicating to Site C while Site A to Site C's sessions have been "**Hibernated**".

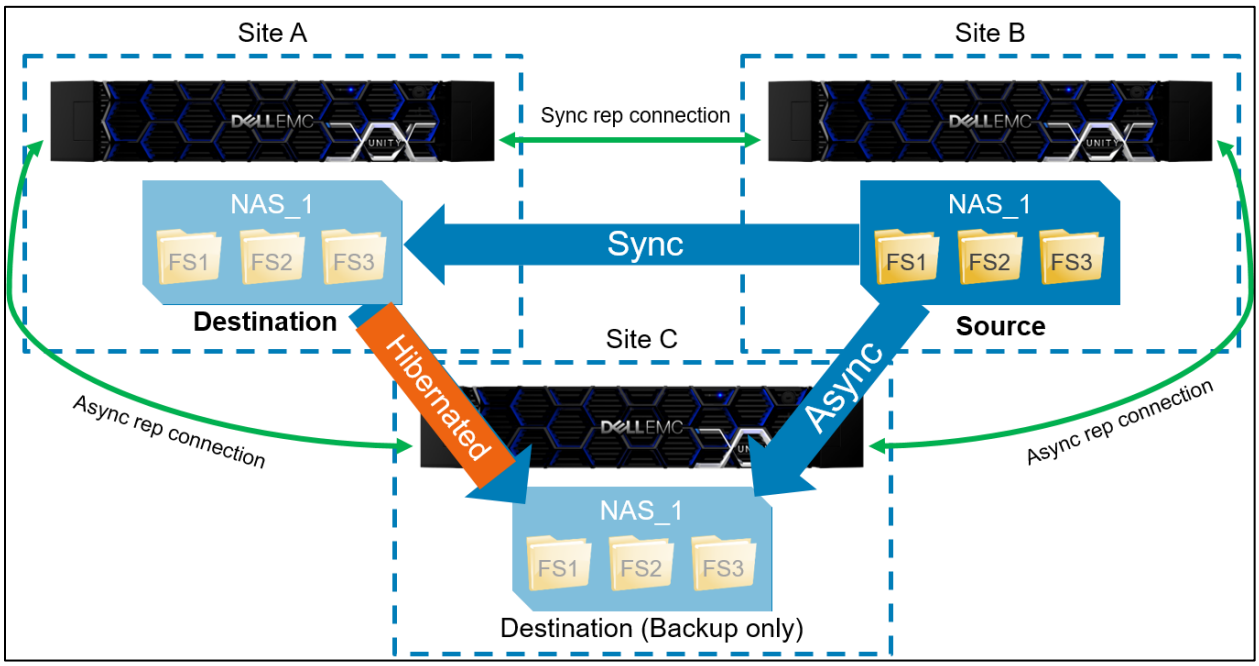


Figure 20 - Example illustration of synchronous and asynchronous sessions after failover and preserve operations

After a failover operation has completed successfully, a preserve operation can be run. First, navigate to the new source NAS Server properties page and on the **Replication** tab, select the synchronous replication session and choose **More Actions**. You'll notice in the **More Actions** dropdown menu that there is a replication operation called "**Preserve async replication**" as can be seen in Figure 21. This initiates the preserve operation and restarts asynchronous replication operations from the new source site. The preserve operation is always initiated on the synchronous replication session on the current source system.

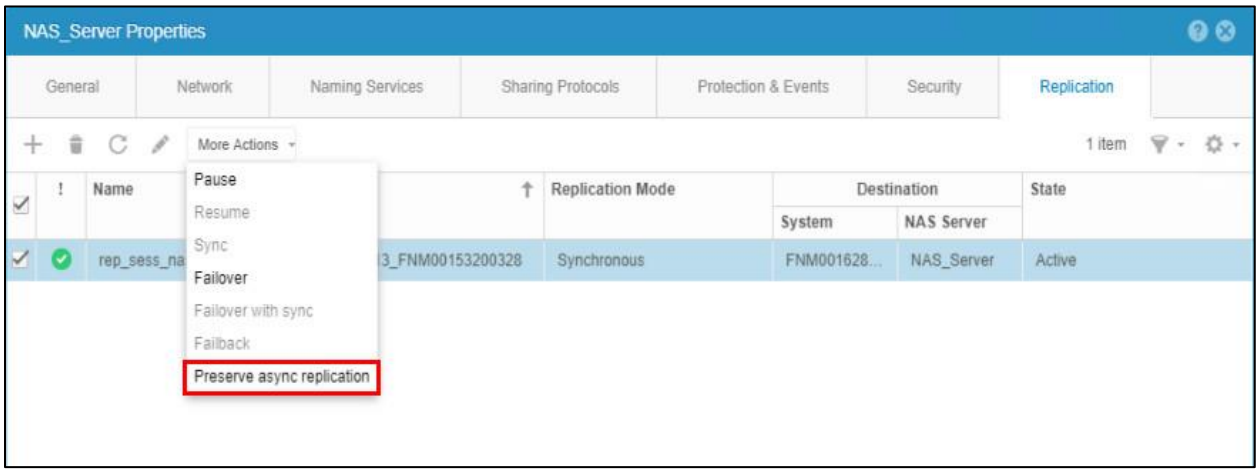


Figure 21 - Preserve asynchronous replication operation in NAS Server Replication tab (Source site)

After failover operation, you'll notice on the destination site that the asynchronous replication session is still seen, but is now in a state of "**Hibernated**" as seen in Figure 22. Deleting this hibernated session is supported if you'd like to clean up the view to only show active replication sessions, but it is not necessary and won't affect ongoing replication operations. If you do delete the hibernated session, an error message appears stating only the local session was deleted and the remote session could still exist. This is expected behavior. Also, if the hibernated session is deleted prior to preserving the session on the new source (Site B), then Site C's asynchronous sessions goes into an error state. Again, this is expected behavior and running a preserve operation on Site B changes the session states back into a healthy state.

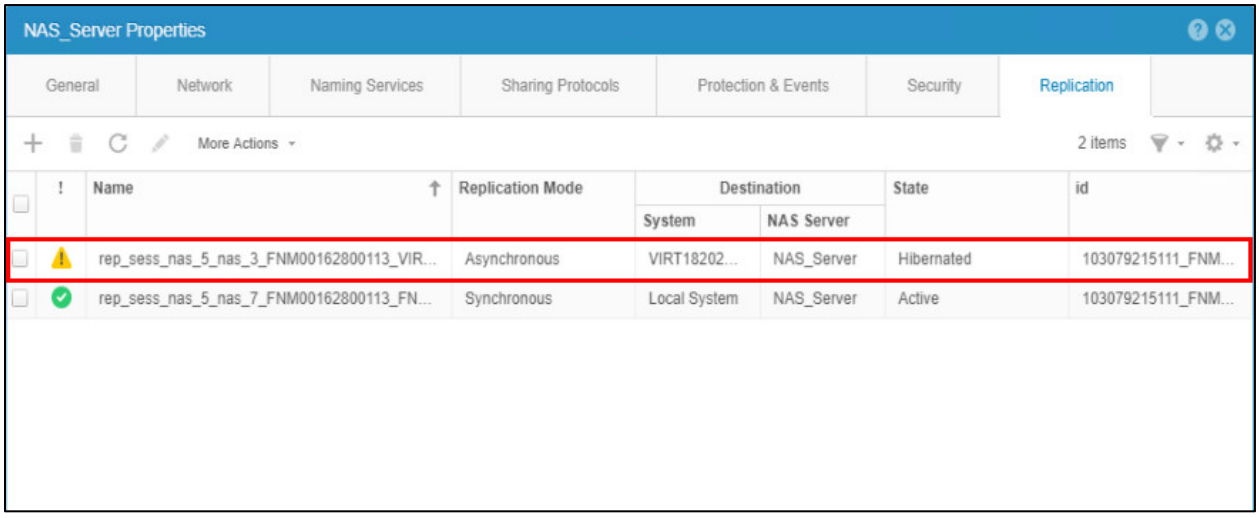


Figure 22 - NAS Server Replication tab with hibernated asynchronous replication session (Destination site)

Backup Only Flag

The "**Backup Only**" flag is a new setting for NAS Servers in OE version 4.4 code which disables the ability to failover to that particular NAS Server. This means running a "**Failover**" command on the source site (Site A) can only failover to the DR system (Site B) and never to the third system (Site C) under normal circumstances. Users could remove the Backup Only flag manually and a run a failover operation from Site C, but this operation should only ever be done in the case of a multiple site disaster scenario such as both Site A and Site B are unavailable and/or non-recoverable. This is because failing over to Site C is an *Unplanned Failover* which would break all current replication operations and require a full synchronization for subsequent replication setup. Therefore, under normal operations, Site C is recommended solely for backup purposes. The "**Backup Only**" flag can be seen in the **General** tab for the properties of a destination NAS Server as seen in Figure 23. Note that this flag must be enabled on an existing asynchronous replication destination if you want to create a new synchronous replication session in addition to it.

If you do failover to Site C and then later want to reinstate the original replication setup prior to the unplanned failover, you'll need to do the following steps:

- Delete all related old replication sessions including Site A to Site B sessions
- Delete existing resources on Site A and Site B
- Create an asynchronous replication session from Site C to Site A
- Once the full synchronization is complete, run a planned failover from Site C to Site A
- Change Site C's NAS Server to "**Backup Only**"
- Create synchronous replication session from Site A to Site B

Note that if you are utilizing UEMCLI, you can skip deleting existing resources on Site A and Site B and setup replication sessions to the existing resources. This initiates a full synchronization regardless.

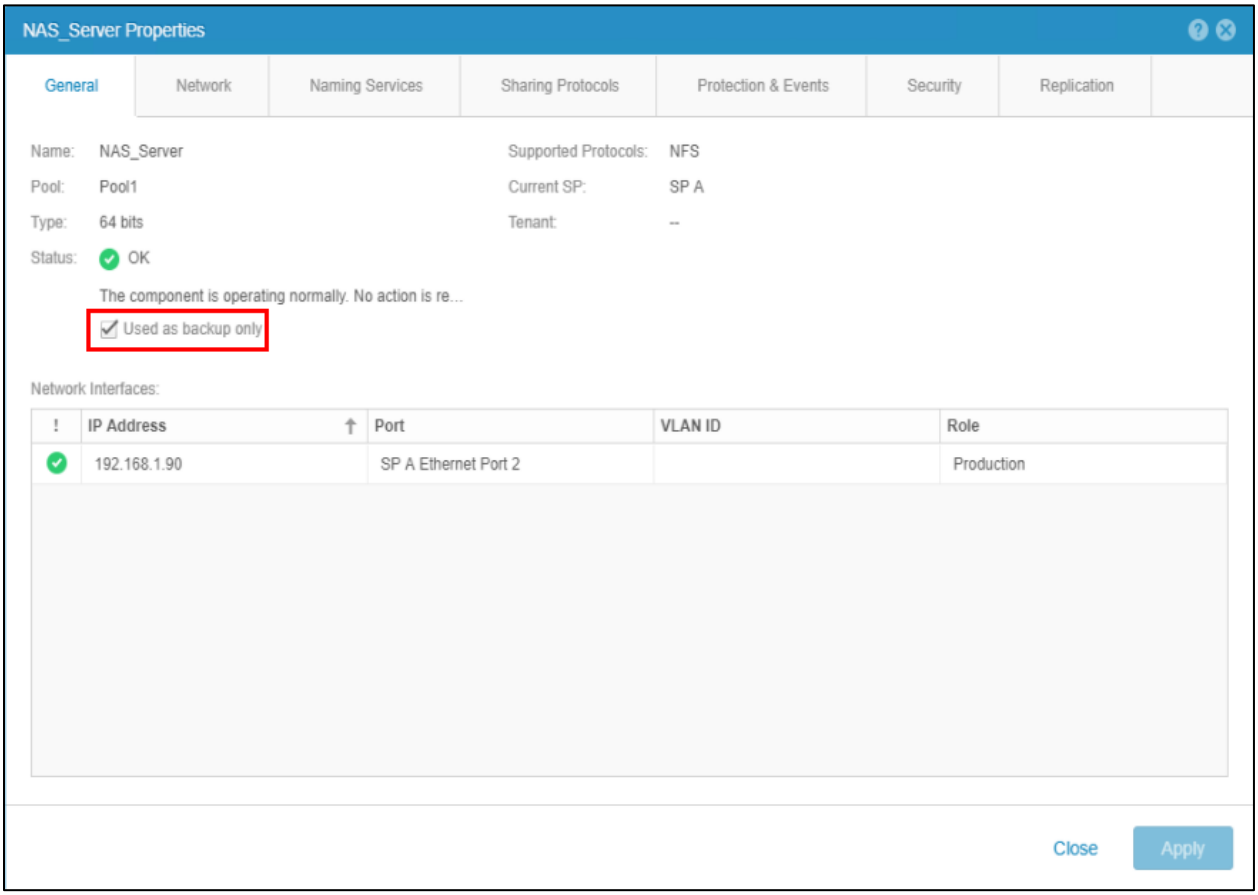


Figure 23 - Destination NAS Server with Backup Only flag enabled

CABINET LEVEL FAILOVER

Another advanced feature as a part of the overall MetroSync solution for Dell EMC Unity is cabinet level failover. This feature allows users to quickly bring up all synchronous file resources on the DR system in the event of an unplanned outage or disaster for the source site. Cabinet level failover is a single UEMCLI command that initiates a failover of all synchronous file resources from the destination system. This command creates separate job threads for each NAS Server that is currently designated as "**Destination**" and brings it up as the source. This ensures that any issues with specific NAS Servers does not affect the other NAS Servers and related resources from coming online on the new source site. For other replicated resources (i.e. asynchronous file and block sessions, and synchronous block sessions), those sessions must be individually failed over and are not included in the cabinet level failover operation. Note that cabinet level failover is specifically for unplanned failover scenarios where the source site is unavailable and/or non-recoverable. Once a cabinet level failover is performed, any subsequent replication operations (i.e. Resume or Failback) requires a full synchronization of data.

Cabinet level failover is a UEMCLI and REST API only feature. It cannot be initiated from the Unisphere GUI. To run a cabinet level failover, either open up an SSH session to your Dell EMC Unity system or install the Unisphere CLI application on a supported host and open up a command prompt session and then run the following command:

```
uemcli /remote/sys –id <value> failover [-force]
```

- Replace <value> with CLI ID of the associated synchronous replication system connection
- The CLI ID can be found in Unisphere: **Replication** page > **Connections** tab, enable optional column "**CLI ID**" for the table, and find the corresponding synchronous replication connection

The above command should be run on the destination system and only when the source system is unavailable and/or non-recoverable. When not utilizing the optional `[-force]` flag option, the system first checks whether the source site is still accessible via the management network and return an error if it detects that the source is still available and not run the command. This check tries to ensure that any unnecessary unplanned failovers are prevented and the error suggests to run planned failovers from the source site instead.

Using the optional `[-force]` flag option as a part of the command ignores all network checks and continues with the cabinet level failover regardless of the source side status. If a cabinet failover is forced while the source site is still available, the duplicate IP avoidance mechanism (also known as split brain) initiates to ensure that only the new source site has host access and prevents host access to the original source.

After an unplanned failover, the NAS servers and file systems on the original source system must be updated to reflect the new status. If there is a large number of NAS servers and file systems, this change may take several minutes to complete. During this period, resume and failback operations of the synchronous replication sessions will not work. It is recommended to wait for all of the updates to complete before running a resume or failback operation. There is no impact to data access while this update is occurring.

For more information on the command itself and to see an example, see the *Dell EMC Unity Family Configuring Replication* Guide on Dell EMC Online Support.

METROSYNC MANAGER

MetroSync Manager (MSM) is a standalone Windows application that can be configured to monitor the system statuses of two systems (“Site A” and “Site B”) participating in file synchronous replication. MSM is available starting with OE 4.5. This optional tool enables the ability to allow automatic failover in the event of a critical failure, for example, an entire site going offline due to power outage or entire network outage. Without MSM, users would need to manually initiate the cabinet level unplanned failover command, as mentioned in the Cabinet Level Failover section. MSM utilizes the same cabinet level failover feature, but does not require a user manually run it and instead automatically initiates the failover if it senses a critical failure. The overall benefit of this is the reduction of overall downtime in the event of a disaster to ensure production resources can continue accessing data without issue from the destination site. MSM can monitor a one-way configuration with one site replicating exclusively to another site or can monitor both systems in a bi-directional configuration with some source objects replicating in one direction and other source objects replicating in the opposite direction, see Figure 24 as an example.

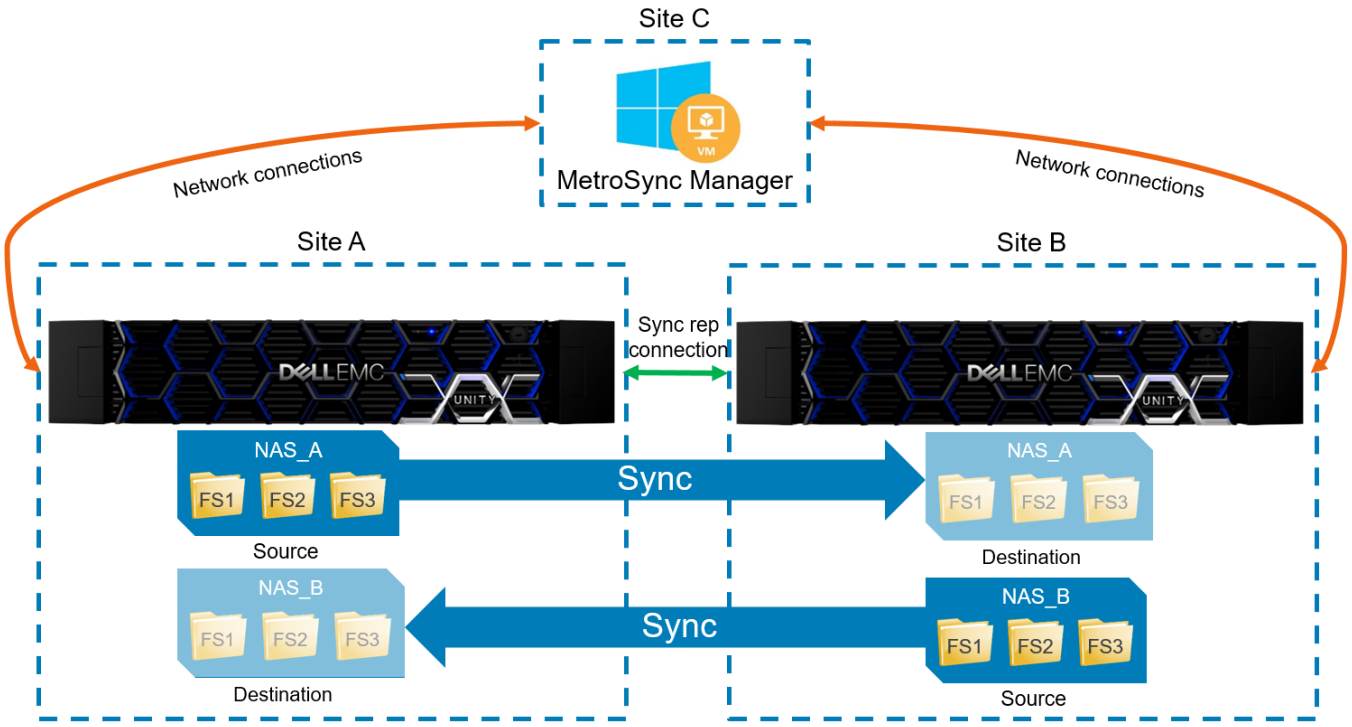


Figure 24 - MetroSync Manager Example Setup (Bi-Directional Configuration)

MetroSync Manager Configuration

To install the MSM application, the following requirements are needed for the Windows host:

- At least 4 cores CPU
- At least 2GB of memory
- At least 4GB of storage for log space

- Windows version: Windows 10, Windows Server 2012 R2, or Windows Server 2016
- Windows host must have .NET Framework 4.6.1 or later installed

MSM initiates a automatic failover in two critical cases which are as follows:

- Site outage
    - Power outage – Entire system unavailable
    - Network outage – Management network (and data storage network)* are unavailable
        - *Applicable only when there are configured “Production IP List” IP addresses
- Pool offline
    - Source pool is unavailable

When configuring MSM to monitor two Dell EMC Unity systems participating in file synchronous replication, there are four main pages that should be filled out depending on the datacenter environment and available resources:

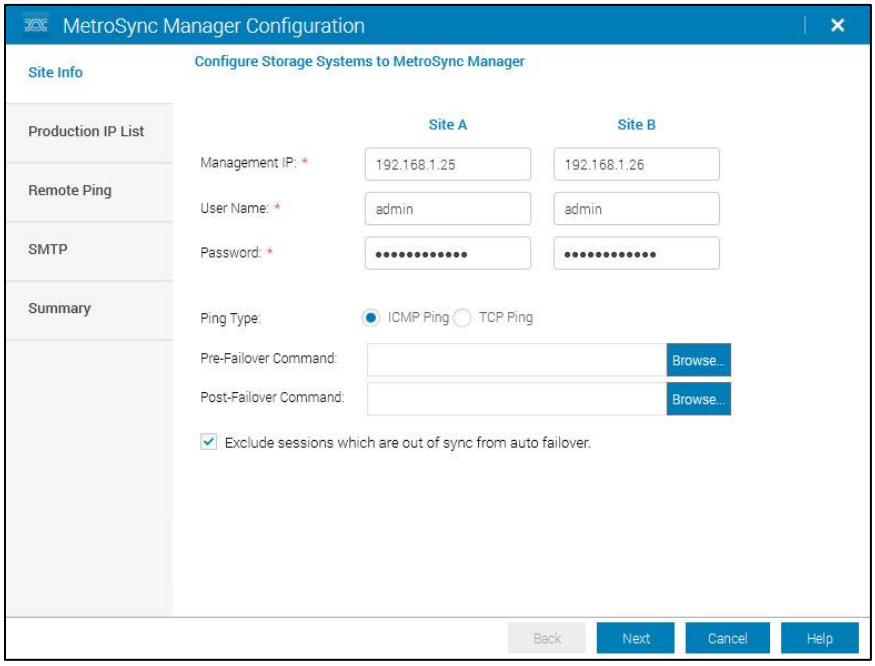


Figure 25 - MSM Configuration (Site Info)

- Site Info (See Figure 25 for an example)
    - Configure management IP address, administrator user credentials for both sites (“Site A” and “Site B”)
    - Choose ping type (ICMP or TCP)
        - ICMP ping is most common and usually allowed in datacenter management networks.
        - In cases where ICMP ping is specifically blocked by configured network firewalls, TCP ping is available to be chosen which usually is not blocked by network firewalls.
            - When choosing TCP ping, a TCP port must be specified. The port chosen should be an open port such as 443 for HTTPS.
        - Both ping types are valid and do not change the functionality or behavior of MSM.
    - Configure pre-failover and post-failover scripts (Optional)
        - In the event of an automatic failover, the pre-failover script is run prior to the failover command being sent. Note that if the script returns a non-zero value, then MSM will not initiate the cabinet failover.
        - In the event of an automatic failover, the post-failover script is run after the failover command completes regardless if the failover is successful or not.

        - These scripts are useful in doing additional checks in a datacenter environment as well as doing any needed automation in the event of a failover.
    - Choose whether to exclude out of sync sessions in the event of a failover
        - This option allows the user to exclude sessions that are not in a fully synced and healthy state (i.e. paused, initial syncing, non-recoverable, etc.) upon failover meaning they are not failed over in the event of a critical failure.
        - If this option is not checked then all file synchronous replication sessions will be failed over regardless of current health/sync status.

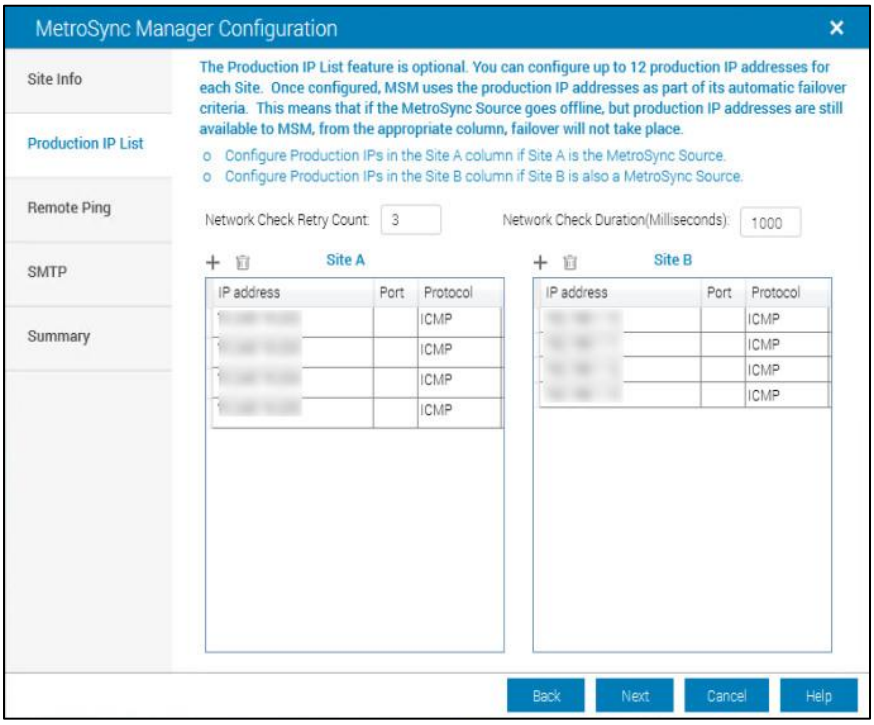


Figure 26 - MetroSync Manager Configuration (Production IP List)

- Production IP List (See Figure 26 for an example)
    - Configure IP addresses that are serving as production purposes (Optional)
        - These IP addresses serve as a redundant check in the event that management interfaces are all seen as unavailable as MSM. Therefore, it is recommended to configure production IP addresses to prevent unnecessary failover whereby the management network is not available, but the data storage network is still available. Note that putting Production IP addresses in the wrong Site could result in automatic failover being disabled upon critical failure.
            - Configure Production IPs in the Site A column when Site A is a MetroSync Source.
            - Configure Production IPs in the Site B column when Site B is a MetroSync Source.
        - If a NAS Server is failed over manually to the peer site, the Production IP List will need to be manually updated if the corresponding IP address was configured in the list..
        - Recommended that configured IP addresses should be production NAS Server interfaces that are visible by MSM.
            - Note that any production IP addresses configured that remain pingable in the event of an outage will prevent automatic failover from occurring.
        - Up to 12 IP addresses can be configured per site.
            - Choose ICMP or TCP ping for each IP address configured.

        - If TCP ping is chosen, a TCP port must be specified. The port chosen should be an open port such as port 445 for SMB or port 111 for NFS.
    - Configure Network Check Retry Count
        - Sets the max number of times MSM will retry the configured production IP addresses before confirming ping unavailability. Default count is set to 3.
    - Configure Network Check Duration
        - Sets the amount of time between each network check for configured production IP addresses. Default time is set to 1000 milliseconds (1 second).

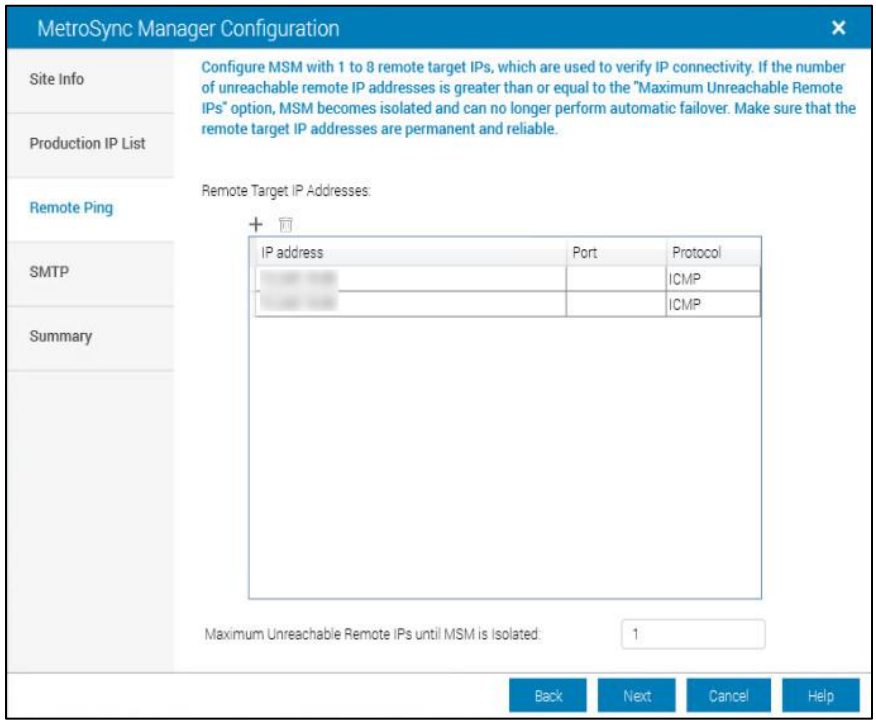


Figure 27 - MetroSync Manager Configuration (Remote Ping)

- Remote Ping (See Figure 27 for an example)
    - Configure IP addresses used by MSM to verify IP connectivity and to determine whether it is isolated from the datacenter network or not
        - These IP addresses should not be addresses on the Dell EMC Unity systems. They should be permanent and reliable IP addresses that should always remain available (i.e. Gateway IPs, DNS Server IPs, etc.). At least one IP address must be configured to start the MSM monitoring service. Up to 8 IP addresses can be configured. If MSM determines it is isolated from the network, then MSM believes it is unsafe to perform automatic failovers and automatic failover protection is disabled.
        - Choose ICMP or TCP ping for each IP address configured.
    - Configure the number of “Maximum Unreachable Remote IPs until MSM is isolated”
        - The number configured here determines how many Remote Ping IP addresses can be unavailable before MSM determines it is isolated from the network and disables automatic failover protection. For example, if the number is set to 2, then if 2 configured Remote Ping IP addresses are seen as unavailable, then MSM will disable automatic failover protection. The default and minimum number is set to 1. The maximum number is limited by the number of configured Remote Ping IP addresses.

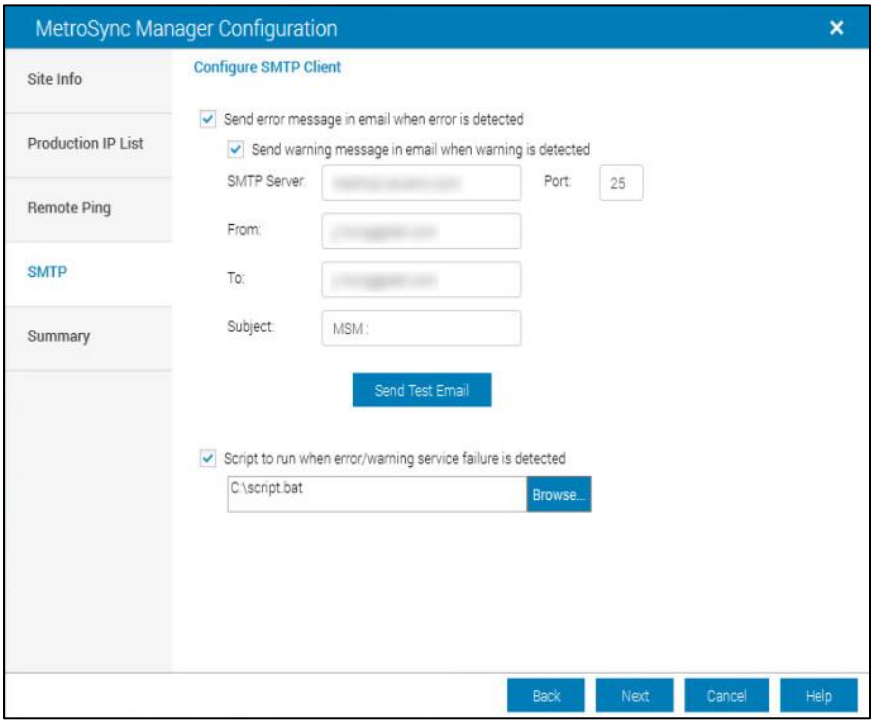


Figure 28 - MetroSync Manager Configuration (SMTP)

- SMTP (See Figure 28 for an example)
  - In the event of a warning or failure, configure SMTP to send an email to a user specified email address
    - Configure SMTP Server (DNS name or IP address), Port number, From address, To address, Subject header
  - In the event of a warning or failure, a script can be configured to run (optional)
    - The script is only run when an email needs to be sent

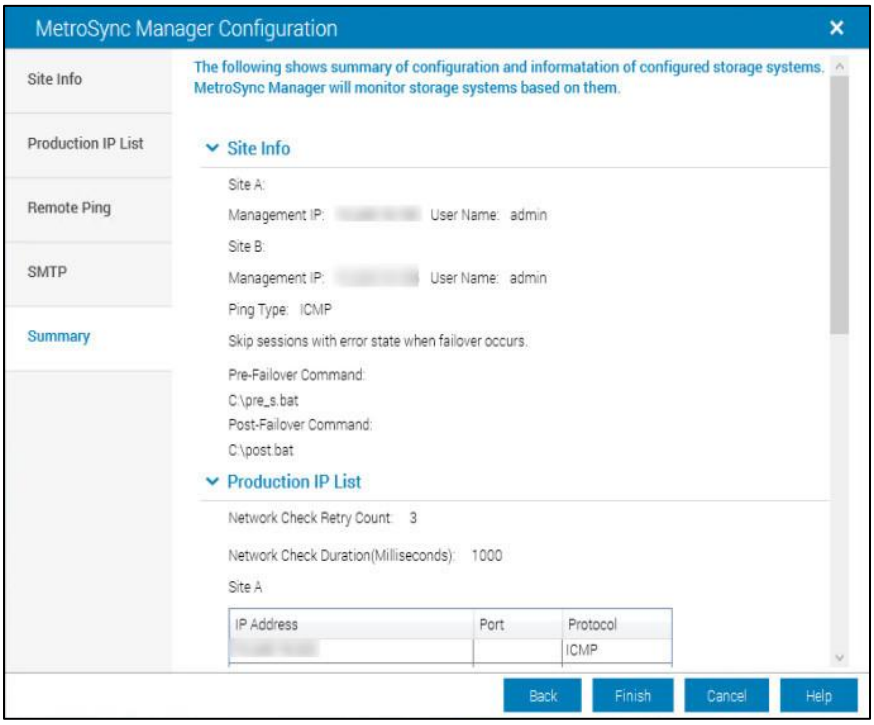


Figure 29 - MetroSync Manager Configuration (Summary)

- Summary (See Figure 29 for an example)
    - Displays a summary of the MetroSync Manager configuration for storage systems participating in file synchronous replication. Each page heading is able to be collapsed for easier viewing purposes.

MetroSync Manager Operations

After the setup of MSM is complete and initial network checks are successful, the main MSM window enables the “Start” button. After clicking “Start”, MSM starts its monitoring operations and continuously checks the following conditions:

- Site A and Site B sites are reachable
- Storage pools on Site A and Site B are online
- MetroSync sessions between Site A and Site B are active
- MSM host is connected to the network

Once properly configured and the monitoring service has started successfully, MSM can detect two types of critical failures: site/network outage and pool offline/failure. In the event of a site/network outage, MSM first checks the management interfaces including Unisphere management interface and sync replication management interfaces. If all management interfaces are no longer available, then MSM checks any Production IP addresses, if configured, for the corresponding site. For example, if Site B’s management network goes down, then MSM checks any Production IP addresses configured for Site B. If the production IP addresses are also unavailable, then MSM assumes a critical issue has occurred at that site and initiates a cabinet level unplanned failover to the peer system. If “Exclude sessions which are out of sync” option is checked then any out of sync sessions are not failed over as part of the process.

In the event of a pool offline/failure issue, MSM detects the issue via REST API queries and initiates a cabinet level unplanned failover to the peer system. Any replication sessions that are associated with the pool failure are also failed over even if the “Exclude sessions which are out of sync” option is checked.

MSM only initiates automatic failover in the above use cases whereby the entire system is suddenly unavailable (i.e. total network outage or power outage) or a pool suddenly fails or goes offline. MSM does not protect against cascading failures or manual shutdown/reboots. Below is a list of some of the scenarios where MSM will not initiate an automatic failover:

- Management network goes down, and data network goes down at a later time (assuming corresponding Production IP addresses are configured)
- Data network goes down and management network remains available
- Graceful shutdown of a storage system
- Reboot both SPs of a storage system

Manual failover operations and preserve operations are not supported via MSM. All replication operations remain available via Unisphere, REST API, and UEMCLI.

MetroSync Manager Management

While MSM monitoring service is running, there are two check boxes for the Main Log window that can be selected: “Follow Tail" and “Only Highlight”. When checking the “Follow Tail” checkbox, the Main Log window automatically shows the latest messages continuously as MSM does its checks, as seen in Figure 30. This helps in not needing to scroll down the log window each time a new set of log messages appear. To disable automatic scrolling, uncheck the “Follow Tail” checkbox. The “Only Highlight” checkbox allows the user to only show Warning, Error, and Critical messages in the Main Log window which show up in red and orange colors, by default. This helps in troubleshooting cases to easily identify what errors were detected by MSM.

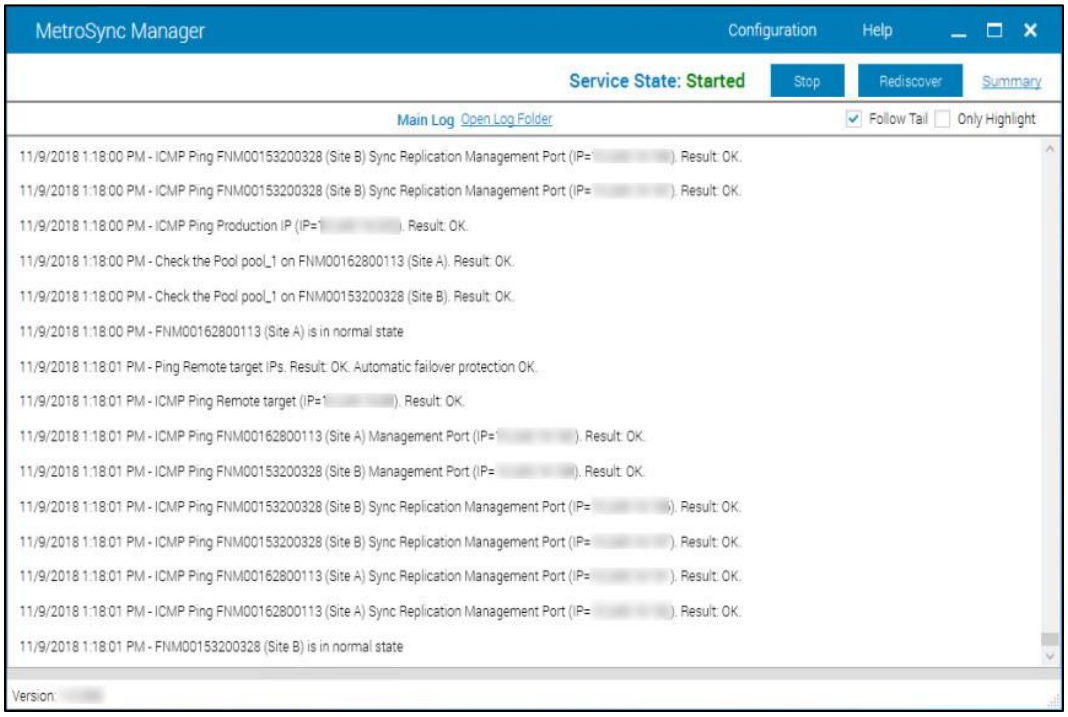


Figure 30 - MetroSync Manager Main Window (Follow-Tail Enabled)

To change the color of Warning, Error, and Critical messages or to highlight specific messages with certain character strings (i.e. name of specific replication session), MSM provides a Log Configuration window which can be found in the Configuration menu. In the Log Configuration Window, there are three default highlight patterns as seen in Figure 31. The three default patterns cannot be deleted or modified, but additional customized highlight patterns can be added/deleted as needed. Customized highlight patterns can be any alphabetical string including spaces and is not case sensitive (i.e. “icmp ping”). Note a particular log item in the GUI can only have one highlight color and highlight patterns are applied in the order configured in the Log Configuration list from top down.

Figure 31 - Log Configuration Page (Default Values)

If any changes are made to the site configurations (i.e. configured additional file system replication sessions or a user credentials for either system are modified), then a “Rediscover” operation needs to be run for MSM to be aware of the new resources/configuration for proper monitoring purposes. The “Rediscover” operation prompts the user to stop the monitoring service if it was started which subsequently needs to be manually restarted after the operation is initated and completes successfully. Note that if a site configuration update occurs whereby a NAS Server(s) are manually failed over to the peer site, the Production IP List page should be updated in the Site Configuration menu prior to the “Rediscover” process being initated.

In the event of system maintenance or non-disruptive upgrade, the MSM monitoring service should be stopped prior to entering the maintenance period. This is to prevent false positives from occurring and unnecessary failovers. To stop the MSM monitoring service, click on the “Stop” button and the Service State will change to “Stopped” in a red color, as seen in Figure 32.

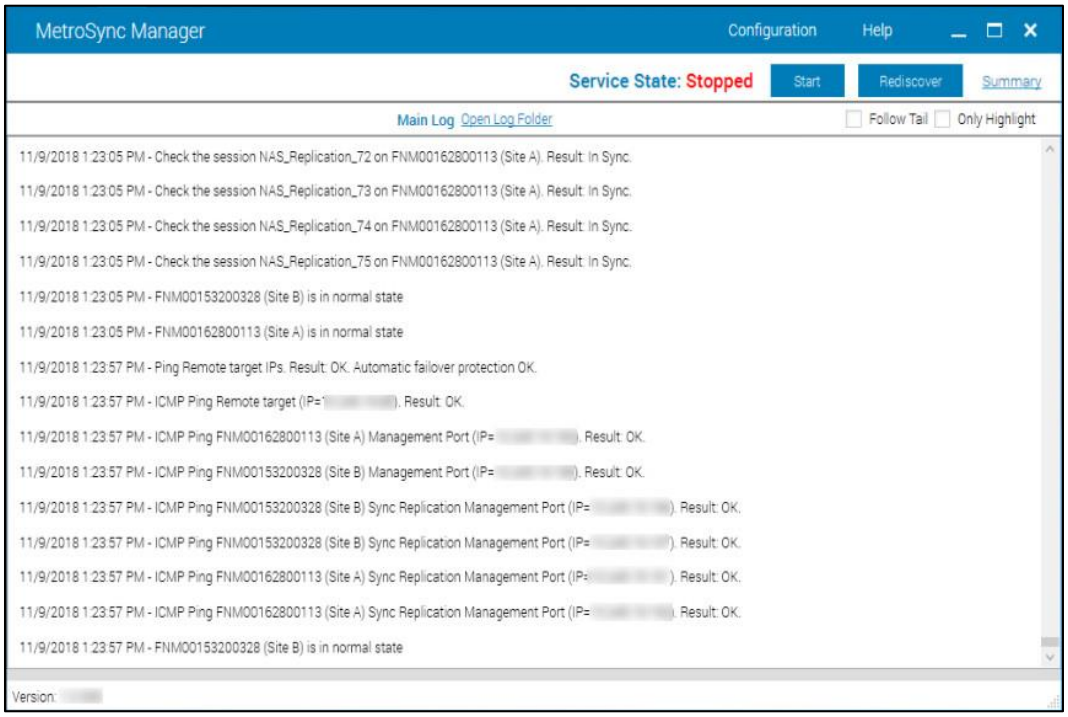


Figure 32 - MetroSync Manager Service Stopped

MetroSync Manager Help

The Help menu is located at the top of the main MSM window which has quick links to the Online Help page, Support Material Collection, and About. The Help page, as seen in Figure 33, provides a user guide to describe the usage of MSM and associated descriptions. The Support Material Collection link captures the latest MSM related logs and zips them up into a log bundle which can be saved and provided to Support for troubleshooting and support purposes. The “Open Log Folder” link in the Main Log window can also be utilized for troubleshooting purposes. The About page displays the MSM Software License Agreement.

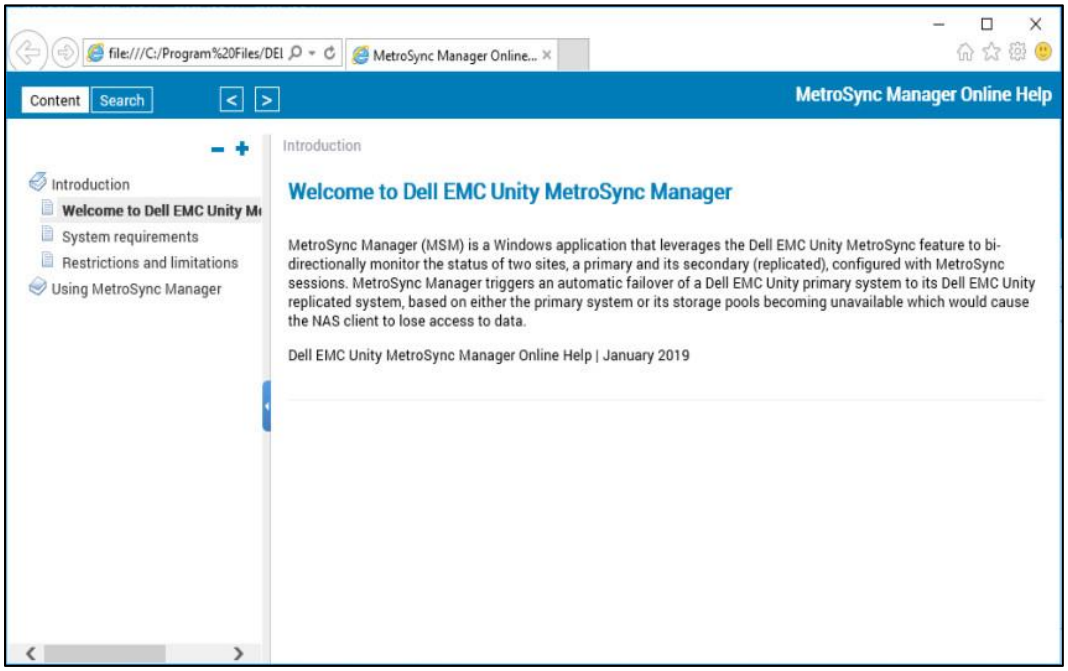


Figure 33 - MetroSync Manager Help Page

UNISPHERE MANAGEMENT

Creating and managing replication in Unisphere is easy and intuitive. All replication operations, including configuring of synchronous replication management interfaces, replication connections, and replication sessions can be performed in the Unisphere GUI. With the help of easy-to-follow wizards, replication can be configured quickly by IT generalists and advanced users alike. Replication can also be configured using Unisphere CLI or REST API. For more information on configuring and managing replication using Unisphere CLI, refer to the *Unisphere Command Line Interface Guide* which can be found on Dell EMC Online Support. For more information on REST API, refer to the *Unisphere Management REST API Programmer's Guide* which can be found on Dell EMC Online Support as well.

The beginning sections of this document already go over how to configure your environment and successfully setup replication in Unisphere. Therefore, this section focuses specifically on viewing and managing existing replication sessions.

Viewing and Managing Replication Sessions

All replication sessions for a particular Dell EMC Unity system can be viewed from the **Replication** page. This includes synchronous and asynchronous replication sessions. To navigate to this page, click **Replication** under **Protection & Mobility** in Unisphere. Figure 34 shows an example of the **Sessions** tab with multiple replication sessions created on the system. In this example, multiple NAS Servers and File Systems are being synchronously replicated to another system. From this window, you can easily see information regarding each session. The following is a list of information displayed on this screen:

- The Replication Session **Name**
- The current **State**
- The **Source**, which includes the source system and the source storage resource
- The **Resource Type**
- The **Destination**, which includes the destination system name and the destination storage resource
- The Replication Session **ID** (Column hidden by default)

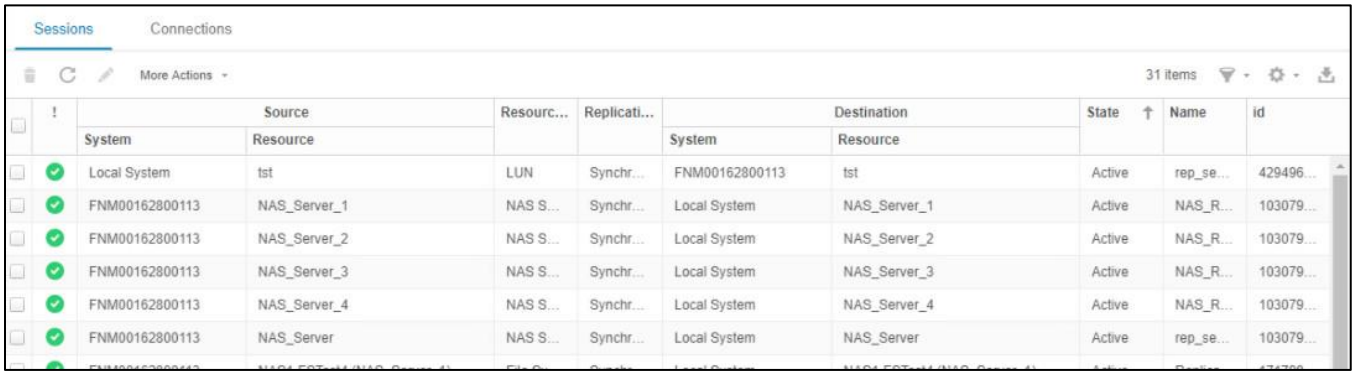


Figure 34 - Replication Page - Sessions Tab

From the **Sessions** tab, you can also issue replication operation commands on the available sessions. After selecting a specific replication session in the list, you can select the **More Actions** dropdown menu to view the available replication operations depending on the current state. In Figure 35 below, you can see that only **Pause** and **Failover** are valid options based on the currently selected resource's replication session state.

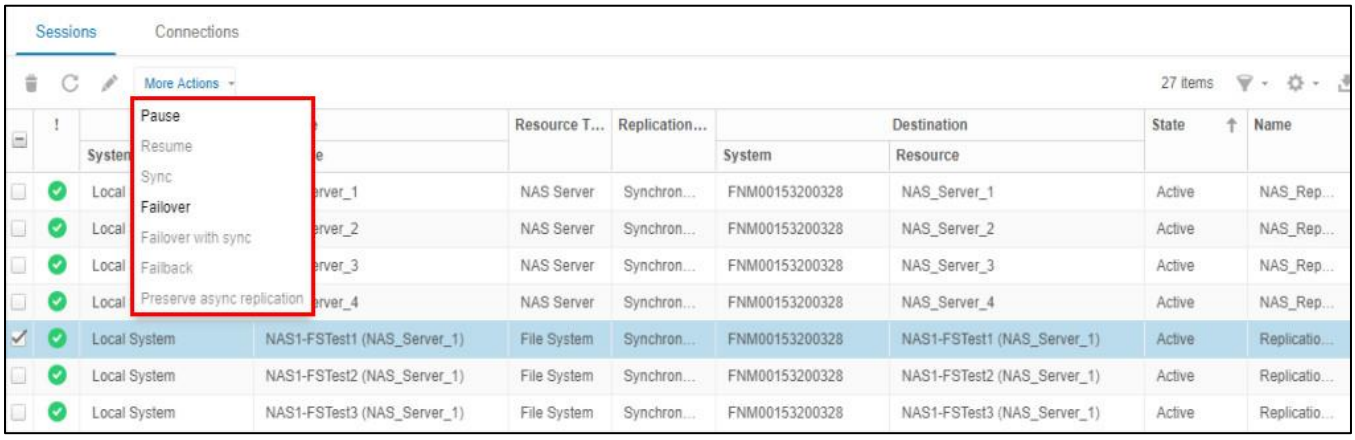


Figure 35 - Replication Page - Session Tab - More Actions

After a synchronous replication session is configured on a resource, you can view information about the session in the resource's properties window. From Unisphere, select the storage resource in question and click **Edit** or double-click the name of the storage resource. From the properties window, view the **Replication** tab, select the associated replication session and click the **Edit** button or double-click the session item to see more information about the session. An example of this session details pane is shown in Figure 36. On this tab, you can view the following information:

- The replication **Session Name**
- The replication **Mode**
- The **Sync State**

Also shown is a pictorial representation of the replication session. The picture shows which storage resource is available for I/O, which direction the data is replicating in, its current state, the destination system name, IP address, and the destination resource name. As the state of the replication session changes, this picture updates to reflect the new state.

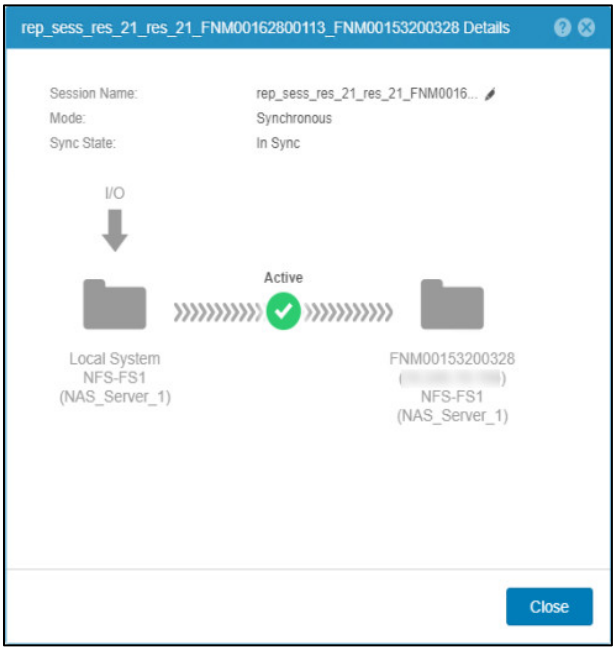


Figure 36 - Replication session details window

MetroSync also includes the ability to do group operations at a NAS Server level. This means running an operation for a particular NAS Server initiates the same operation for all of its associated storage resources at the same time. Group operations are available for the following operations: Failover, Failback, Pause, and Resume. Figure 37 shows an example of executing a group Failover for a NAS Server which automatically includes all of its associated file systems.

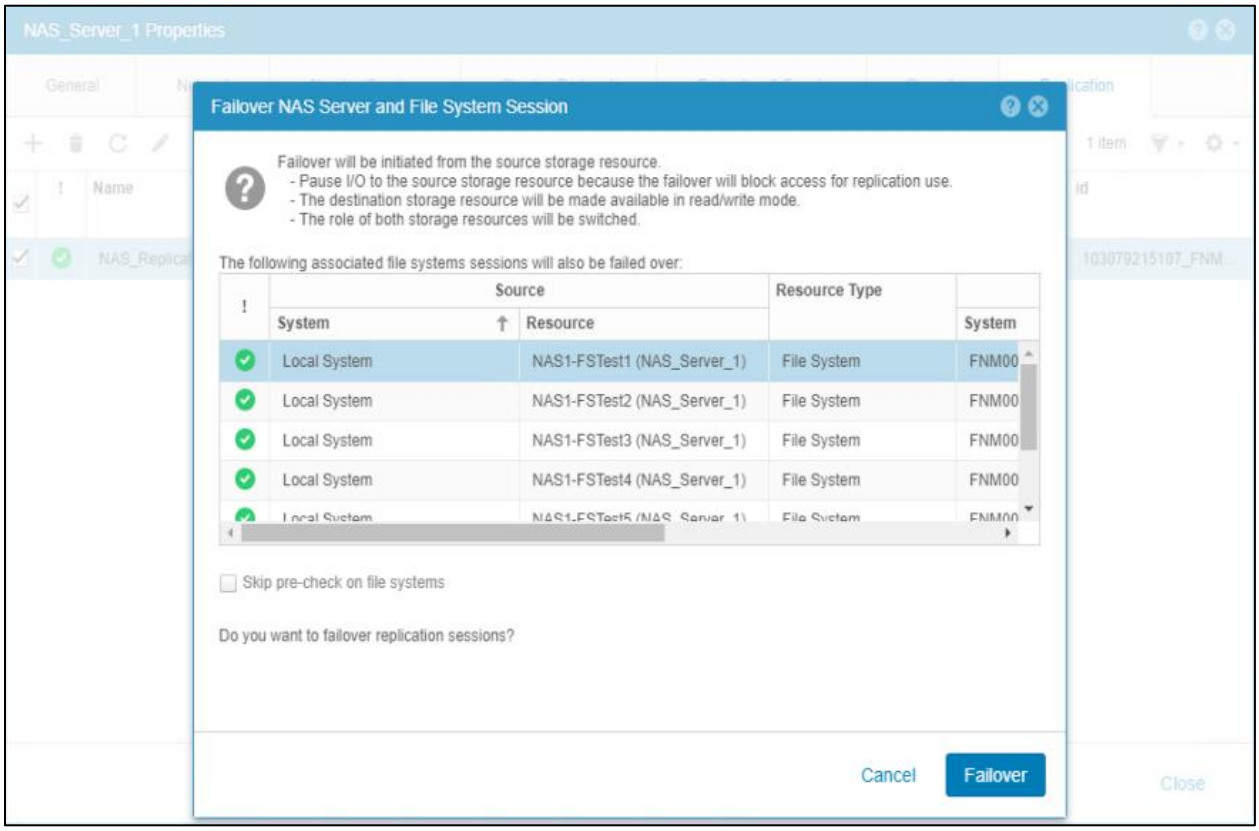


Figure 37 - Group failover operation for NAS Server and its underlying File Systems

Override Network Addresses for File Replication

If you have file replication configured to a destination site on a different physical network, to ensure minimal downtime before running a failover, ensure you modify the destination NAS Server properties to include an override address for the associated network interface. Figure 38 and Figure 39 show an example of the destination NAS Server with the Production Interface being overridden to a different IP address. Note that a different Ethernet Port can also be selected in this process. Note that a similar process can be done for other services as well such as DNS, NIS, and LDAP.

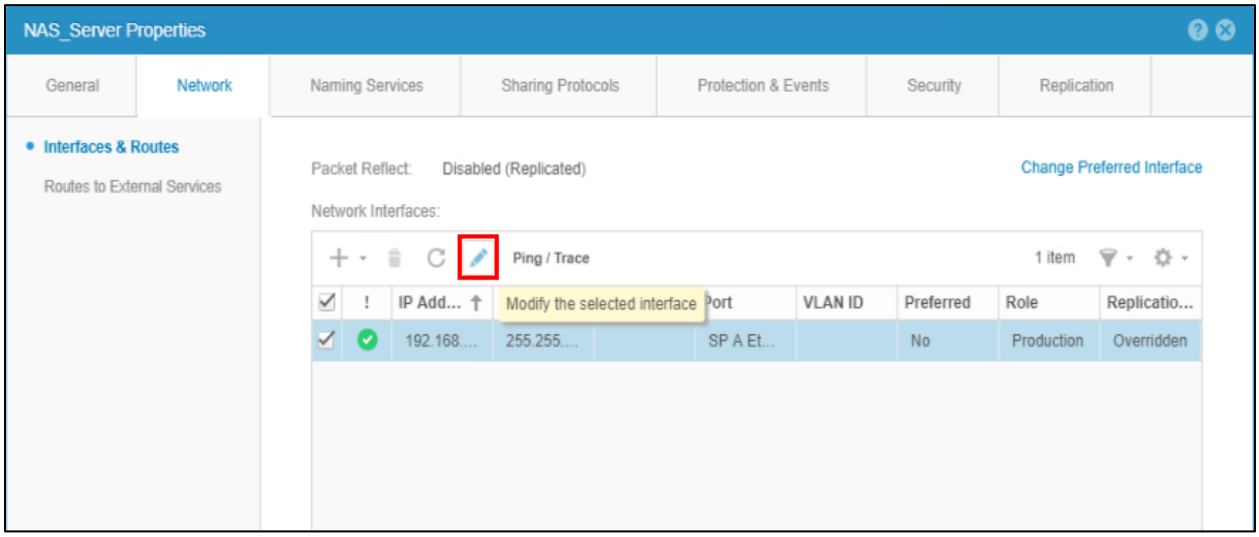


Figure 38 - Network tab - Modify production interface (Destination NAS Server)

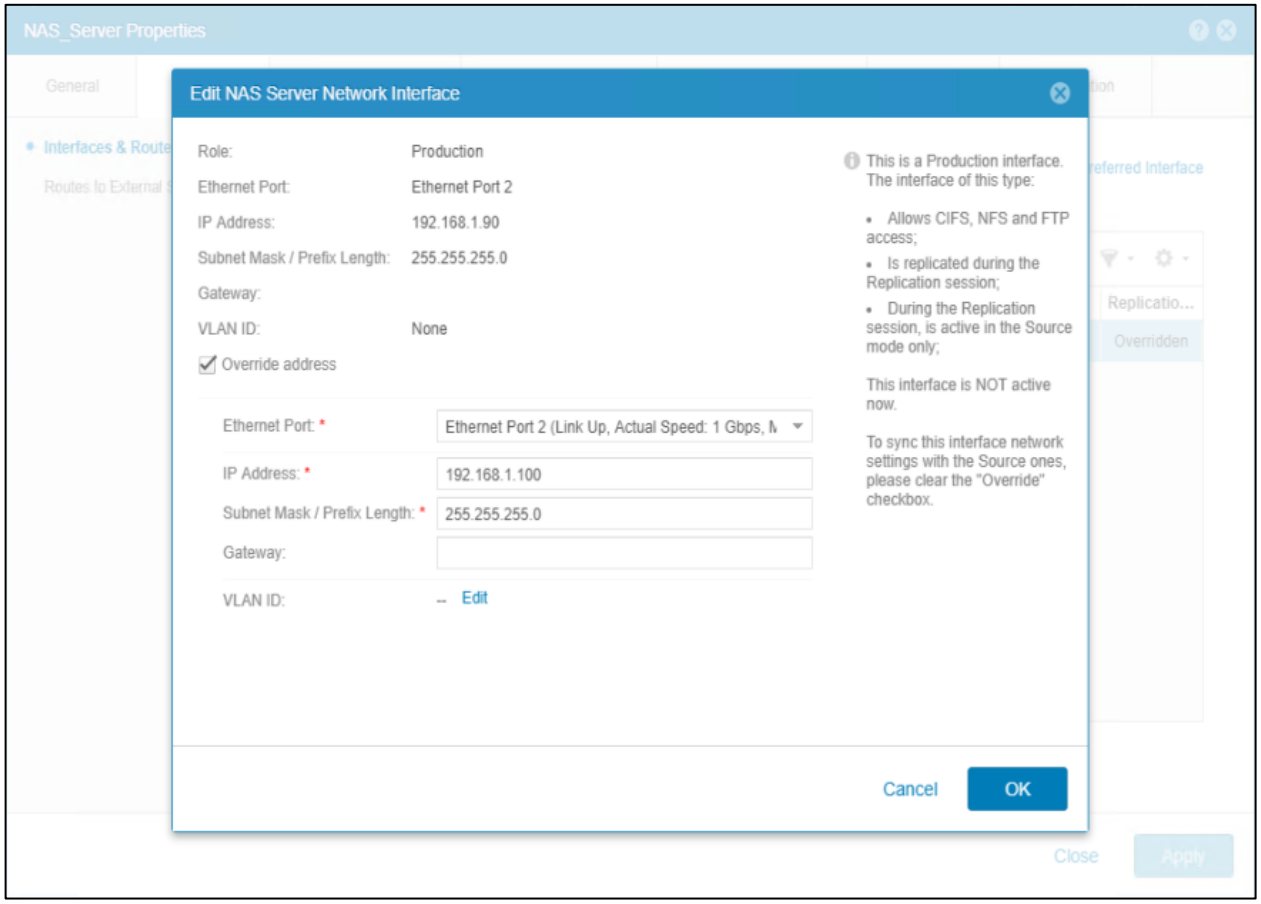


Figure 39 - Edit NAS Server Network Interface Window (Destination NAS Server)

Enable NDMP on NAS Server

NDMP (Network Data Management Protocol) support allows for data backup to NDMP backup clients which could be a physical tape library or virtual tape library. NDMP can also work with replication for additional data protection. In OE version 4.4, Dell EMC Unity supports both 3-way NDMP backup over Ethernet and 2-way NDMP backup over FC. If replication is enabled on a NAS Server and additional data protection is needed, NDMP can be enabled through the **Protection & Events** tab in the properties page of a NAS Server, as seen in Figure 40. NDMP can be enabled on any NAS Server as a part of a MetroSync solution, but note that in OE version 4.4.0, replication should be paused if running an NDMP backup from a source or destination NAS Server to ensure no issues during backup operations. In OE 4.4.1 and later, this limitation is not an issue and both async and sync replication can remain enabled during NDMP backup operations.

For more information on NDMP setup and configuration, see the *Dell EMC Unity: NAS Capabilities* white paper on Dell EMC Online Support.

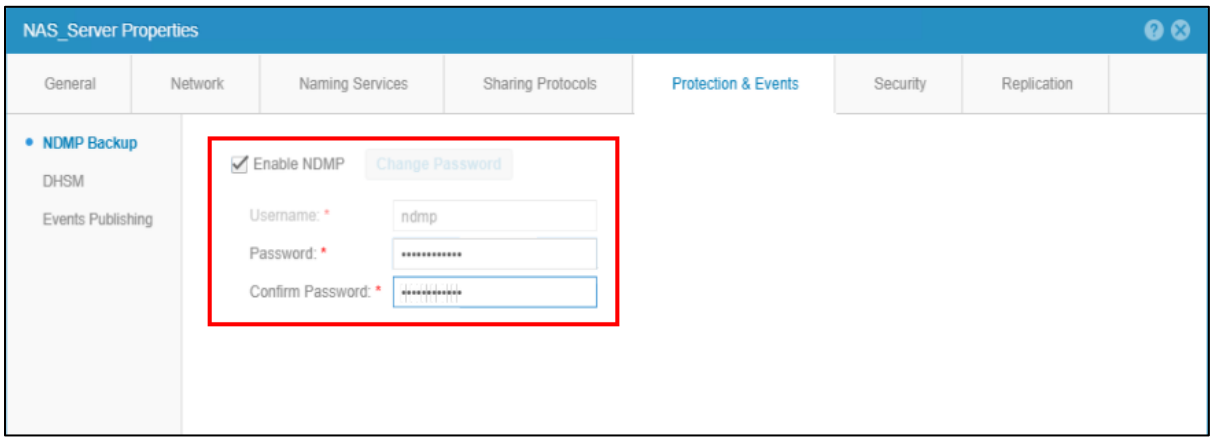


Figure 40 - Enable NDMP on Destination NAS Server (Protection & Events Tab)

Upgrades

Prior to initiating an upgrade of the Dell EMC Unity OE, it is needed to pause all ongoing replication sessions. The Pre-Upgrade Health Check (PUHC) checks for any active replication sessions and generates a warning if any are found. Once the upgrade completes, you can resume each replication session. The pause and resume operations should be initiated from the source system for each replication session. Note that if you're running Dell EMC Unity OE version 4.2 or later, pausing and resuming the NAS server replication session also pauses and resumes all of its associated file system replication sessions automatically.

CONCLUSION

The 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 source 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.

MetroSync also supports asynchronous replication to a third system as a backup and recovery solution. When NAS Servers and file systems/NFS datastores are failed over between the synchronous systems, the asynchronous sessions to the third system can be preserved. The asynchronous sessions can be recreated and incrementally updated on the new system where the NAS Server is active thereby ensuring continuation of backup services after failover.

MetroSync also supports cabinet level failover which allows users to quickly bring up the destination system as source in the event of an unplanned outage or disaster. This minimizes the downtime and data unavailability of resources and provides a rapid recovery method to continue critical business operations.

Since the OE 4.5 release, Dell EMC Unity supports MetroSync Manager as part of the overall MetroSync solution which allows for automatic cabinet level failover in the event of a disaster or critical failure. This provides the additional benefit of minimizing downtime for such disaster scenarios providing smoother disaster recovery.

REFERENCES

The following references can be found on Dell EMC Online Support:

- Dell EMC Unity: Cloud Tiering Appliance (CTA)
- Dell EMC Unity: Data Reduction
- Dell EMC Unity: DR Access and Testing
- Dell EMC Unity: Dynamic Pools
- Dell EMC Unity: Introduction to the Platform
- Dell EMC Unity: Unisphere Overview
- Dell EMC Unity: NAS Capabilities
- Dell EMC Unity: Snapshots and Thin Clones
- Dell EMC UnityVSA

APPENDIX A: REPLICATION MAXIMUMS

The following table outlines the Dell EMC Unity replication maximums.

	Dell EMC Unity 600/600F/650F	Dell EMC Unity 500/500F/550F	Dell EMC Unity 400/400F/450F	Dell EMC Unity 300/300F/350F	Dell EMC UnityVSA
Max Replication Sessions (Synchronous + Asynchronous)	2000	1500	1000	1000	16
Max Replication Sessions (Synchronous Replication)	2000	1000	750	500	N/A
Max Target Systems (Synchronous Replication)	1	1	1	1	N/A
Max Consistency Group Replication Sessions (Synchronous Replication)	128	64	64	64	N/A
Max LUNs per Replicated Consistency Group (Synchronous Replication)	32	32	32	32	N/A
Max Concurrent Initial Syncs (Synchronous Replication)	32	28	24	16	N/A
Max Concurrent Replication Sessions (Asynchronous Replication)	256	256	256	256	8
Max Target Systems (Asynchronous Replication)	16	16	16	16	16
Max LUNs per Replicated Consistency Group (Asynchronous Replication)	75	75	75	75	50
Max Concurrent Initial Syncs (Asynchronous Replication)	32	32	32	32	4
Max replicated NAS Servers (Asynchronous Replication)	256	128	128	90	4

The table above outlines a number of system maximums when using synchronous and asynchronous replication. The maximum replication sessions includes all replication sessions on the system, which includes both synchronous and

asynchronous replication sessions, local or remote. The replication destination storage resources count towards the system maximums, even though they are not host accessible when acting as a destination image. In Dell EMC Unity, only one replication connection used for synchronous replication, or synchronous and asynchronous replication, can be created. This also means that only 1 pair of systems can replicate synchronously to each other.

APPENDIX B: REPLICATION SUPPORT ACROSS PLATFORMS

The following table outlines replication support across the various Dell EMC midrange platforms.

Source	Destination	Block Replication		File Replication		RecoverPoint (Block)
		Synchronous	Asynchronous	Synchronous	Asynchronous	
Dell EMC Unity	Dell EMC Unity	✓	✓	✓**	✓	✓
Dell EMC Unity	Dell EMC UnityVSA	×	✓	×	✓	×
Dell EMC Unity	VNXe3200	×	✓*	×	×	✓
Dell EMC Unity	VNX1/VNX2	×	×	×	×	✓
Dell EMC Unity	VNXe1600	×	✓	×	×	×
Dell EMC UnityVSA	Dell EMC UnityVSA	×	✓	×	✓	×
Dell EMC UnityVSA	Dell EMC Unity	×	✓	×	✓	×
Dell EMC UnityVSA	VNXe3200	×	✓*	×	×	×
Dell EMC UnityVSA	VNX1/VNX2	×	×	×	×	×
Dell EMC UnityVSA	VNXe1600	×	✓	×	×	×
VNX1/VNX2	Dell EMC Unity	×	×	×	×	✓

*Min VNXe3200 OE v3.1.5.6801782
**Min Dell EMC OE version 4.4

Footnote:
✓– Supported
×– Not Supported