



DELL EMC EMBEDDED NAS TECHNICAL OVERVIEW

Technical Note

ABSTRACT

This technical note explains the Embedded NAS (eNAS) that is available for Dell EMC PowerMax, VMAX All Flash, and VMAX3 arrays.

May, 2018

TABLE OF CONTENTS

EXECUTIVE SUMMARY	4
Audience.....	4
INTRODUCTION	4
POWERMAX AND VMAX HYPERVISOR	5
Hypervisor CPU Core Allocation – Multicore Emulation	5
Hypervisor Memory Allocation	6
Hypervisor Storage Allocation – Cut-Through Device (CTD).....	6
Hypervisor Network Connectivity	6
EMBEDDED NAS	7
eNAS Components	7
eNAS with Service Levels	9
eNAS with FAST.X.....	9
File System Logging.....	10
Common Log File Systems.....	10
Split Log File Systems	10
Migrating to Split Log File Systems	11
MANAGING ENAS	11
Provision Storage for File	14
Configuring Link and Launch.....	15
Enabling Alerts	17
EMBEDDED NAS FILE AUTO RECOVERY	18
File Auto Recovery	18
Failover.....	19
Cleaning a File Auto Recovery Session	19
Reversing Replication.....	20
Adding Storage (Refresh Pairs).....	21
FAR Health Check	21
Preserve ReplicatorV2 Sessions for FAR	23
File Auto Recovery Manager	24
CONCLUSION.....	25
References.....	25

EXECUTIVE SUMMARY

Organizations around the globe need IT infrastructure that can deliver instant access to the huge volumes of data intrinsic to traditional transaction processing/data warehousing and to a new generation of applications built around the world of social, mobile, and big data. Dell EMC® is redefining Data Center Cloud Platforms to build the bridge between these two worlds to form the next generation of Hybrid Cloud.

[Dell EMC PowerMax](#) family, VMAX® All Flash and VMAX3™ storage arrays support embedded file data services that enable customers to consolidate islands of block and file storage, simplify management, and reduce deployment costs by up to 33%. Embedded NAS (eNAS) uses the hypervisor provided in PowerMaxOS and HYPERMAX OS to create and run a set of virtual machines on VMAX controllers. These virtual machines host two major elements of eNAS: software Data Movers and Control Stations and are distributed based on the mirrored pair architecture of PowerMax, VMAX All Flash, and VMAX3 to evenly consume resources for both performance and capacity. Common eNAS use cases include running Oracle® on NFS, VMware® on NFS, Microsoft® SQL on SMB 3.0, home directories, file sharing, and Windows server consolidation.

AUDIENCE

This document is for anyone who needs to understand Embedded NAS (eNAS) along with the components and technology in PowerMax, VMAX All Flash and VMAX3 arrays to provide a unified platform.

INTRODUCTION

PowerMax, VMAX All Flash and VMAX3 arrays support the industry's first open storage and hypervisor converged operating system, HYPERMAX OS 5977. It combines industry-leading high availability, I/O management, quality of service, data integrity validation, storage tiering, and data security with an open application platform.

HYPERMAX OS 5977 features the first real-time, non-disruptive storage hypervisor that manages and protects embedded services by extending VMAX high availability to services that traditionally would have run external to the array. It also provides direct access to hardware resources to maximize performance.

VMAX All Flash and VMAX3 arrays are based on the world's first and only Dynamic Virtual Matrix™ that delivers agility and efficiency at scale. Hundreds of CPU cores are pooled and allocated on-demand to meet the performance requirements for dynamic workloads.

POWERMAX AND VMAX HYPERVISOR

PowerMaxOS 5978 and HYPERMAX OS 5977 runs on the Dynamic Virtual Matrix leveraging its scale out flexibility of cores, cache, and host interfaces. The embedded storage hypervisor reduces external hardware and networking requirements, delivers high levels of availability, and dramatically reduces latency. Hypervisor upgrades are performed non-disruptively.

Within the PowerMax and VMAX Hypervisor, virtual machines (VMs) provide the host platform that includes CPU processing, memory, network interface card (NIC), ports, data storage by using a Cut-Through device (CTD), and external network through the Management Module Control Station (MMCS). VMs run within the front end FA emulation of the PowerMax and VMAX.

Figure 1 shows the primary components of the VMAX Hypervisor.

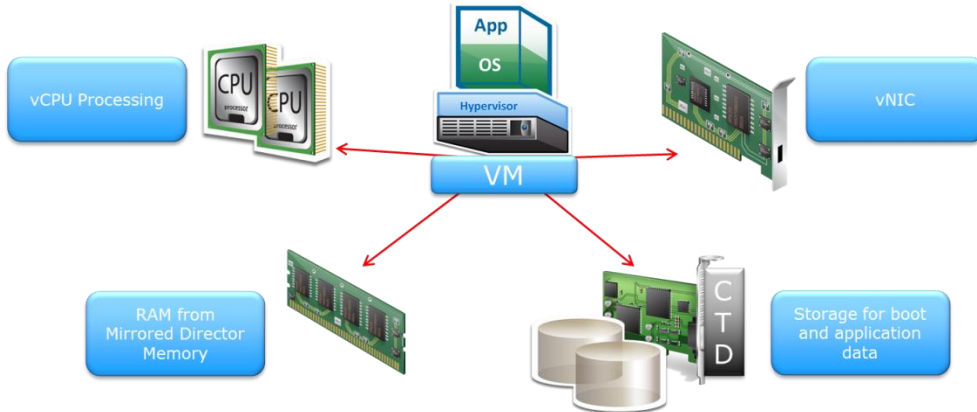


Figure 1. Hypervisor Concepts – Virtual Machines

HYPERVISOR CPU CORE ALLOCATION – MULTICORE EMULATION

Using the Multi-Core emulation capability in PowerMax, VMAX All Flash, and VMAX3 arrays, the CPU processing is provided using CPU cores from the FA emulation. The cores are pooled for front-end, back-end, and for PowerMaxOS or HYPERMAX OS functions as shown in Figure 2. All of the CPU cores on the director will work on I/O from all of the ports. This helps ensure PowerMax and VMAX directors' ports are always balanced.

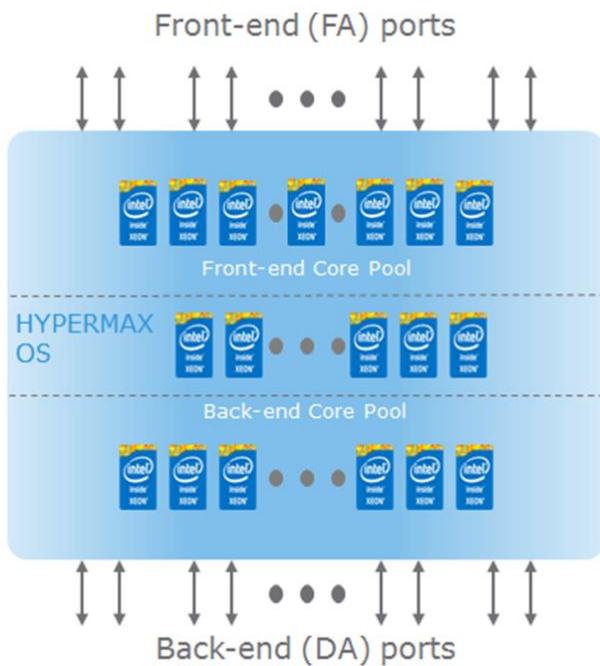


Figure 2. VMAX Multi-Core Emulation

HYPERVISOR MEMORY ALLOCATION

Memory is allocated to the hypervisor from the director cache during the initial setup as shown in Figure 3. This memory is then allocated to each Virtual Machine (VM) on that director for the purpose of embedded applications. The amount of memory allocated to a VM is dependent on the type of application, for example Control Station or Data Mover.

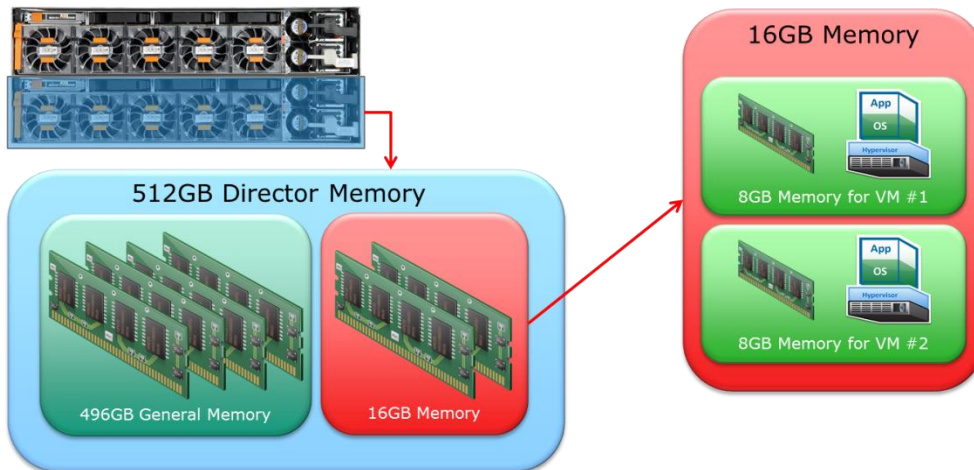


Figure 3. HYPERMAX OS 5977 Hypervisor Memory Allocation

HYPERVISOR STORAGE ALLOCATION – CUT-THROUGH DEVICE (CTD)

Data storage for both the boot and application data is provided using a Cut-Through Device (CTD) as shown in Figure 4 which acts like an HBA that accesses LUNs in the PowerMax, VMAX All Flash, and VMAX3 array. The CTD has two components to enable access to the LUNs through an FA port. The first is the CTD Server thread. This runs on the FA emulation and communicates with the CTD Client in the embedded operating system. The second is the CTD Client Driver. The CTD Client Driver is embedded in the host operating system and communicates with the CTD server running on the FA emulation. An operating system running in a VM must have the CTD client driver installed to access the LUNs.

Embedded application ports are virtual ports specifically provided for use by the VMs that contain the applications, such as Embedded NAS. They are addressed as ports 32-63 per director FA emulation. The virtual ports are provided to avoid contention with physical connectivity. As with physical ports, LUNs can be provisioned to the virtual ports.

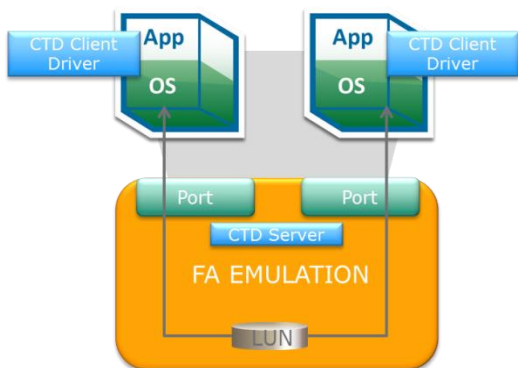


Figure 4. Cut-Through Device

HYPERVISOR NETWORK CONNECTIVITY

Network connectivity for the VMs is provided by a virtual NIC (vNIC). The vNIC is connected to the internal network providing communication to PowerMaxOS, HYPERMAX OS 5977 and other VM instances. The VM management external network connectivity is provided through a PowerMaxOS and HYPERMAX OS component called the network address translation (NAT) Gateway which is part of the Infrastructure Manager (IM) emulation. The NAT Gateway provides translation services between external and internal IP addresses and utilizes a separate network connection on each of the two Management Module Control Stations (MMCS).

EMBEDDED NAS

Embedded NAS (eNAS) extends the value of PowerMax, VMAX All Flash, and VMAX3 to file storage by enabling the customer to leverage vital enterprise features including Service Level based provisioning, Host I/O limits, and FAST technologies for both block and file storage. eNAS is a multi-controller, traditional/transactional NAS solution designed for customers requiring hyper consolidation for block storage (the traditional VMAX use case) combined with moderate file storage in mission-critical environments. Common eNAS use cases include running Oracle® on NFS, VMware® on NFS, Microsoft® SQL on SMB 3.0, home directories, file sharing, and Windows server consolidation.

eNAS uses the hypervisor provided in PowerMaxOS and HYPERMAX OS to create and run a set of virtual machines on VMAX controllers. These virtual machines host the two major elements of eNAS: software data movers and control stations. The VMs are distributed based on the mirrored pair architecture of PowerMax, VMAX All Flash, and VMAX3 to evenly consume resources for both performance and capacity.

eNAS provides equivalent NAS capabilities to those found on the VNX2 File Operating Environment (OE).

ENAS COMPONENTS

Figure 5 shows how eNAS fits within the PowerMax and VMAX3 and VMAX All Flash architecture.

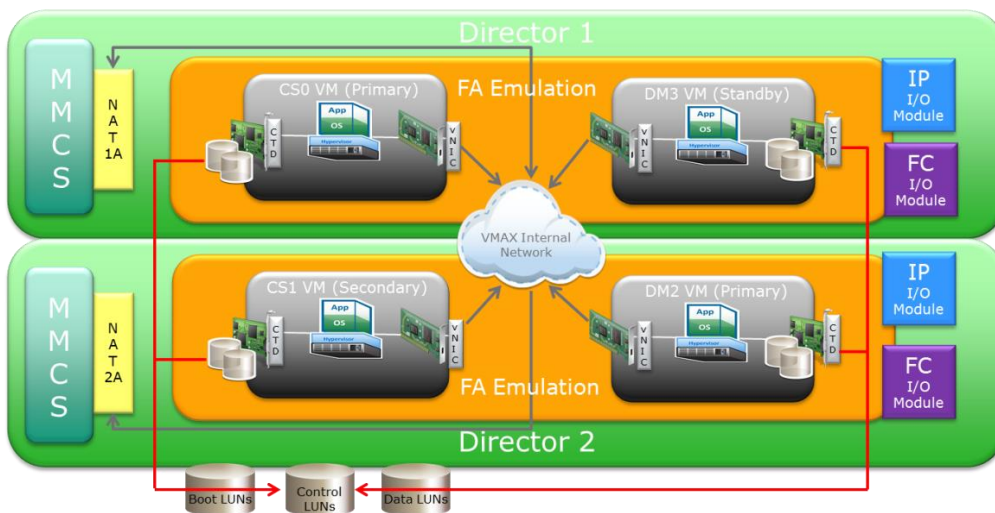


Figure 5. Embedded NAS Architecture

Management Module Control Station (MMCS): Provides environmental monitoring capabilities for power, cooling, and connectivity. Each MMCS has two network connections that connect to the customer's local area network (LAN). One to allow monitoring of the system, as well as remote connectivity for the Dell EMC Customer Support team, and the second for use by the Network Address Translation (NAT) Gateway.

Network Address Translation (NAT) Gateway: Provides translation services between external and internal IP addresses and utilizes a separate network connection on each of the two MMCS.

Control Station (CS): Provides management functions to the file-side components referred to as Data Movers. The control station is responsible for Data Mover failover. There is a primary CS which is configured with a matching secondary CS to ensure redundancy.

Software Data Mover (DM): Accesses data from the backend and provides host access using the I/O Modules that supports the NFS, CIFS, and pNFS protocols.

Data Mover I/O options: Using PCI Pass-through, the following I/O modules are available for the NAS Services on the Data Mover:

- 2-port 10 Gb BaseT Ethernet copper module
- 2-port 10 Gb Ethernet optical module ^a

- 4-port 1 Gb BaseT Ethernet copper module ^b
- 4-port 8Gbps Fibre Channel module for NDMP Back-up to tape use. There is a maximum of 1 per DM.

Note: A minimum of 1 Ethernet I/O module is required per Data Mover.

- The standard I/O module for the VMAX All Flash is the 2-port 10 Gb Ethernet optical module. Other I/O modules are available when ordered separately.
- The 4-port 1Gb BaseT Ethernet copper module is not available on the VMAX All Flash 250F/FX, 950F/FX and the PowerMax family.

Within the PowerMax, VMAX All Flash, and VMAX3 families the eNAS specifications depend on the array model, as Table 1 shows.

Table 1. Embedded NAS Model Comparison

		VMAX3			VMAX All Flash				PowerMax	
Components		VMAX 100K	VMAX 200K	VMAX 400K	VMAX 250F/FX	VMAX 450F/FX	VMAX 850F/FX	VMAX 950F/FX	PowerMax 2000	PowerMax 8000
Control Stations (2)	Shared Logical Cores	2	2	2	2	2	2	4 ^d	4	4 ^d
	Memory (GB)	8	8	8	8	8	8	8	8	8
Data Movers (2/4/6 ^a /8 ^a)	Logical Cores	8	10/20	16/32 /48 ^a /64 ^a	12/24	10/20	16/32 /48 ^b /64 ^b	16/32 /48 ^b /64 ^b	12/24	16/32 /48 ^b /64 ^b
	Memory (GB)	12	48/96	48/96 /144 ^a /192 ^a	48/96	48/96	48/96 /144 ^b /192 ^b	48/96 /144 ^b /192 ^b	48/96	48/96 /144 ^b /192 ^b
	I/O Modules (Max)	4	6/12	6/12 /18 ^a /24 ^a	6/12	6/12	6/12 /18 ^b /24 ^b	6/12 /18 ^b /24 ^b	6/12	6/12 /18 ^b /24 ^b
Max Software Data Movers		2	4	8 ^a	4	4	8 ^b	8 ^b	4	8 ^b
Max Capacity per Data Mover		256TB	512TB ^c	512TB ^c	512TB	512TB ^c	512TB ^c	512TB ^c	512TB	512TB
Max eNAS capacity per VMAX		256TB	1.5PB ^c	3.5PB ^c	1.5PB	1.5PB ^c	3.5PB ^c	3.5PB ^c	1.5PB	3.5PB

- Support for more than four Data Movers in the VMAX 400K requires HYPERMAX OS 5977.813.785 or later.
- The PowerMax 8000, VMAX All Flash 850F/FX and 950F/FX can be configured via Sizer with a maximum of 4 Data Movers. However, via RPQ that can be increased to either 6 or 8 Data Movers.
- Each VMAX 100K Data Mover can support up to 256TB of usable capacity. Starting with HYPERMAX OS 5977.813.785 or later VMAX 200K, 400K, 250F, 450F, 850F, 950F each Data Mover can support up to 512TB of usable capacity.
- The PowerMax 8000 and VMAX All Flash 950F/FX has dedicated Control Station cores.

ENAS WITH SERVICE LEVELS

eNAS supports Service Level features in both PowerMaxOS and HYPERMAX OS. Using Service Levels with eNAS provides the ability to ensure the highest priority NAS storage pool's response times are not impacted by lower priority applications or storage pool IOPS. All storage groups that are system managed are represented as mapped pools on eNAS. A storage group is considered system managed if it has an explicitly set Service Level or Storage Resource Pool (SRP) value. Devices outside of system managed control are discovered as Default Service Level (DSL) devices and associated to the system defined pool (symm_dsl).

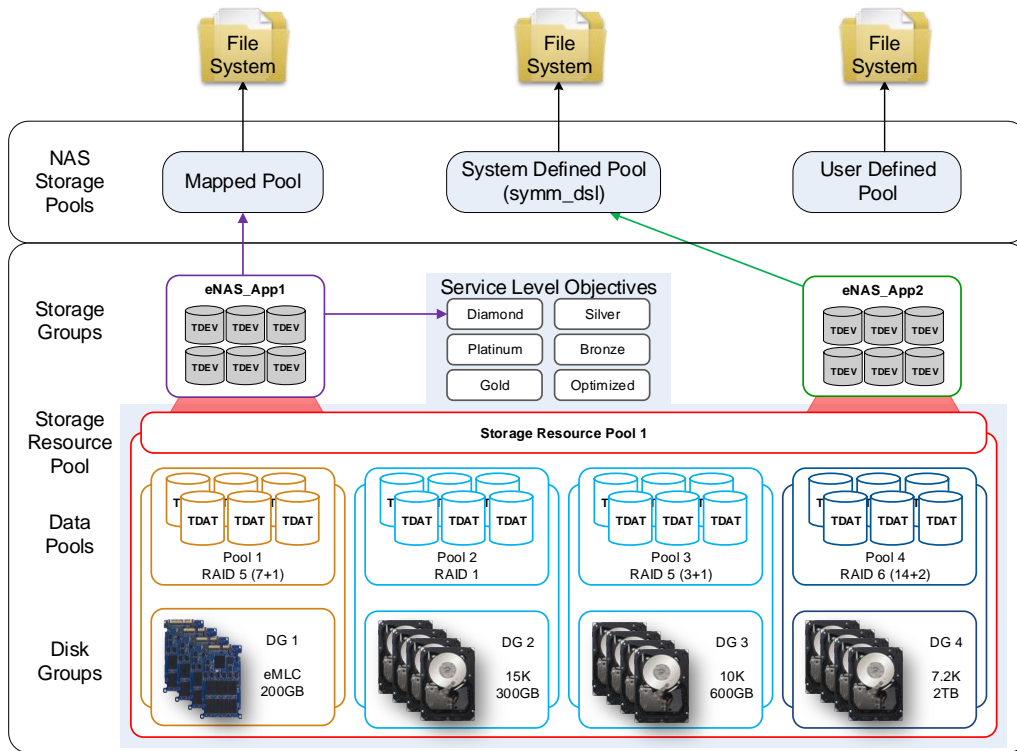


Figure 6. Overview of Disks consumed from VMAX3 on eNAS

Figure 6 is an overview of how a file system created on eNAS is related to underlying disk groups on a VMAX3. File systems can be created on a NAS storage pool or a NAS meta volume. There are three types of NAS Storage Pools: mapped pool, system-defined pool, and user-defined storage pool. All system managed storage groups are represented as mapped pools. Mapped Pools are the recommended type of storage pool.

On VMAX3 systems, if the storage group devices are created out of multiple disk groups, the technology type is “mixed”. For single disk groups, the disk technology type is the physical disk type (EFD, SAS). PowerMax and VMAX All Flash systems have the disk type EFD.

Additional information about PowerMax and VMAX Service Levels is available in the Dell EMC Service Levels for PowerMaxOS document.

ENAS WITH FAST.X

FAST.X for VMAX3 evolves and extends storage tiering and service-level management to other storage platforms and to the cloud. FAST.X integration with XtremIO extends service level management and zero data loss SRDF replication to Dell EMC XtremIO.

If the Storage Group is associated to an SRP that only has external disk groups in it, the system marks the device as “FTS” disktype. If the SRP has both internal and external disk groups, the devices are marked as “MIXED” disktype.

Note: eNAS FAST.X support requires HYPERMAX OS 5977.691.684 or later.

FILE SYSTEM LOGGING

File system logging is used to track metadata transactions on the File System. The log is replayed during a recovery process to ensure all tracked updates are committed. There are two types of logs: a Common Log and a Split Log. Split log file systems are the default, starting with version HYPERMAX OS 5977.691.684, eNAS File OE 8.1.7

Note: File systems can have an IP Replication session only if they have the same type of logging. For example: Common log file systems can only have a replication session with a common log file system, not a split log file system.

COMMON LOG FILE SYSTEMS

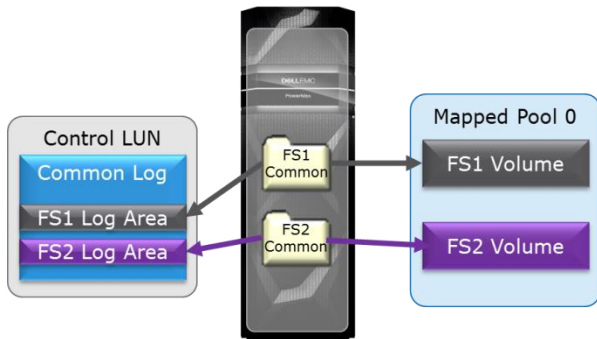


Figure 7. Common Log File Systems

Common Log file systems share a log mounted on the Data Mover. Log space is allocated to each Data Mover from a File Control LUN. Common log file systems are created with the CLI only, using the command:

```
nas_fs -create <volname> size=<size> pool=<pool> log_type=common
```

There is no option to create a common log file system using EMC Unisphere.

SPLIT LOG FILE SYSTEMS

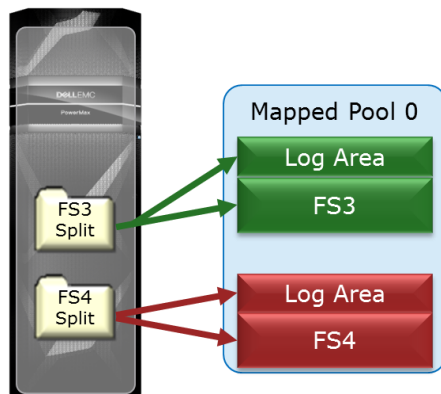


Figure 8. Split Log File Systems

Split Log file systems utilize a log area that is unique to each file system. The space for the log is allocated from the file system volume, and is outside the reported usable space for the file system. Split log file systems are the default with HYPERMAX OS 5977.691.684, and are created by using CLI or EMC Unisphere. The split log file systems are required for the eNAS File Automated Recovery solution. Common log files are not available with this feature.

MIGRATING TO SPLIT LOG FILE SYSTEMS

The migration of data from a common log file system type to a split log file system requires a split log file system type to be created. Leveraging an external host, both file systems would be mounted and the following utility examples could be used to copy the data:

- rsync – for linux / NFS environments
- EMCopy / lgdup – for CIFS/SMB environments

Dell EMC Cloud Tiering Appliance (CTA) is an option to migrate the files from the source common log file system to a split log file system.

Note: More information on CTA is in the white paper: *Dell EMC VNX with Cloud Tiering Appliance*.

MANAGING ENAS

Unisphere for PowerMax and Unisphere for VMAX 8.0.1 and higher includes a file dashboard as the interface to create and provision storage to the eNAS Data Movers which facilitate the file I/O for end customers. Unisphere has additional functionality that provides faster troubleshooting and provisioning for administrators who are supporting file storage. Using the File Dashboard, shown in Figure 9, the user can:

- View Capacity details
- View and manage block and file assets
- Display the mapping of file systems to storage groups
- Allow the provisioning of storage for file systems
- View Data Mover health status
- View File Storage alerts
- Use Unisphere Link and Launch to EMC Unisphere for VNX

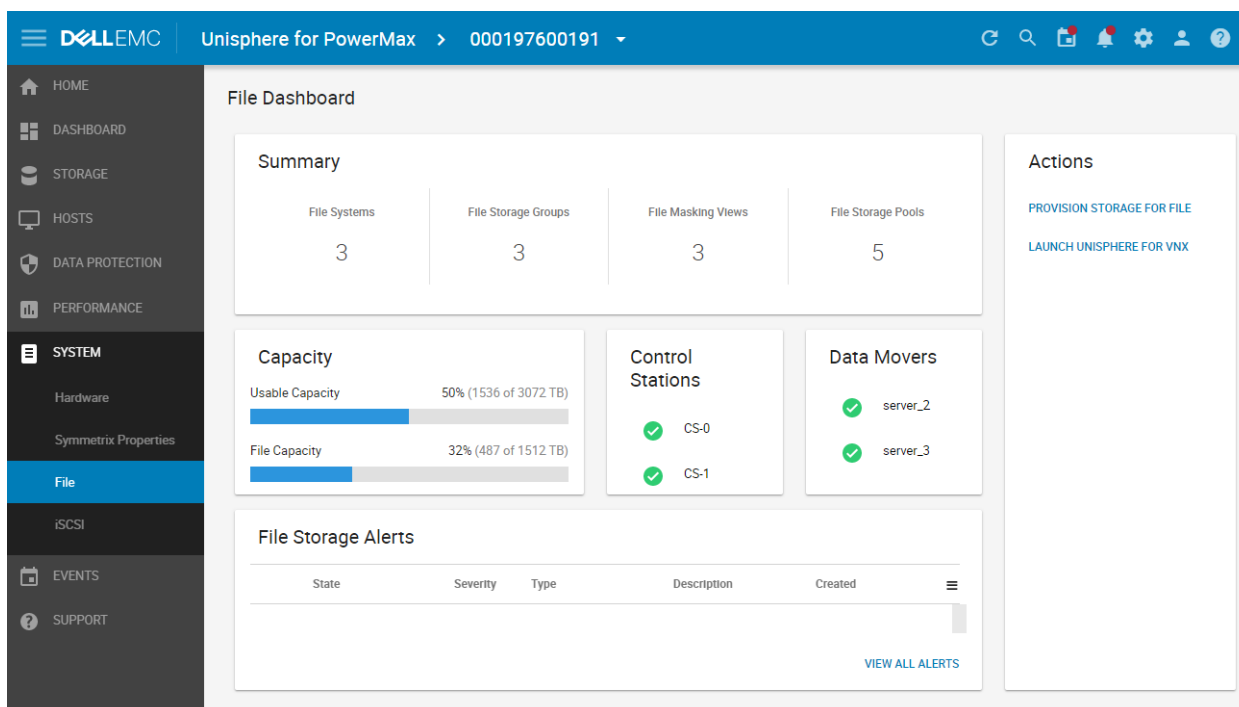


Figure 9. Unisphere for PowerMax - File Dashboard

The capacity panes in the File Dashboard, shown in Figure 10, provide a summary of the total usable and File capacity of the File Storage Groups.

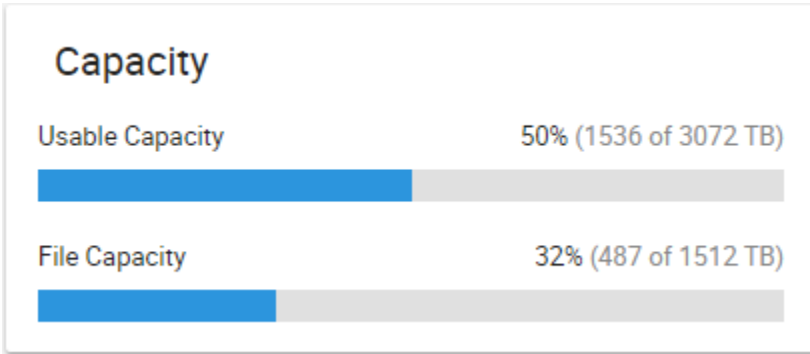


Figure 10. File Dashboard Capacity Details

The summary pane of the File Storage Dashboard shows the number of File Systems, File Storage Groups, File Masking Views, and File Storage Pools (see Figure 11).

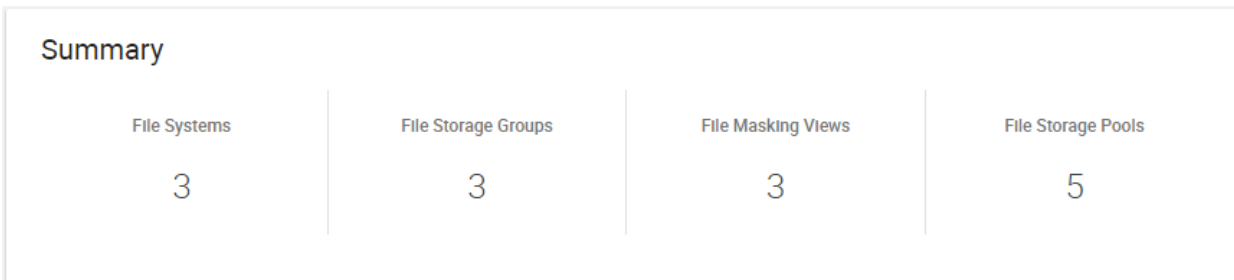


Figure 11. File Dashboard Summary

Each of the summary items provides a link to detailed information. Figure 12 shows an example of the file systems, the associated storage group, pool, and allocated capacity.

The screenshot shows the Dell EMC interface for 'Unisphere for PowerMax'. The breadcrumb path is 'APM00112400946 > File Systems'. There are 3 items in the list. The table below shows the details for each file system.

File System Name	Storage Group(s)	Storage Pool	Allocated %	Capacity (GB)
<input type="checkbox"/> App1_DB	eNAS_App1, eNAS_App1_DB	eNAS_App1_DB	46.8	504.13
<input type="checkbox"/> App1_Log	eNAS_App1, eNAS_App1_Log	eNAS_App1_Log	100	252.03
<input type="checkbox"/> App2_DB	eNAS_App2, eNAS_App2_DB	eNAS_App2_DB	0	756.22

Figure 12. File Systems Summary

The File Storage Groups details in Figure 13, allow you to view and modify the storage groups associated with eNAS.

Name	Compliance	SRP	Service Level	Capacity (GB)	Emulation	Compression Ratio	Masking Views	# File Systems	# Storage Pools
EMBEDDED...	●	NONE	Diamond	105.82	FBA	—	1	0	0
▼ eNAS_App1	—	NONE	NONE	1,536.02	FBA	—	1	2	2
□ eNAS_App...	✓	SRP_1	Diamond	1,024.02	FBA	1.1:1	1		
□ eNAS_App...	✓	SRP_1	Gold	512.01	FBA	1.0:1	1		
▼ eNAS_App2	—	NONE	NONE	1,536.02	FBA	—	1	1	2
□ eNAS_App...	✓	SRP_1	Diamond	1,024.02	FBA	16.0:1	1		
□ eNAS_App...	✓	SRP_1	Silver	512.01	FBA	16.0:1	1		

Figure 13. File Storage Group Summary

Detailed NAS active management (such as share management, replication, and quotas) is performed within EMC Unisphere. EMC Unisphere, shown in Figure 14, is pre-installed and runs on the eNAS Control Stations. The link between Unisphere for PowerMax and EMC Unisphere is based on context sensitive Link and Launch.

This is the same management interface that is used for VNX OE File 8.1. Additional information is in the VNX OE for File 8.1 documentation available on <http://support.emc.com>

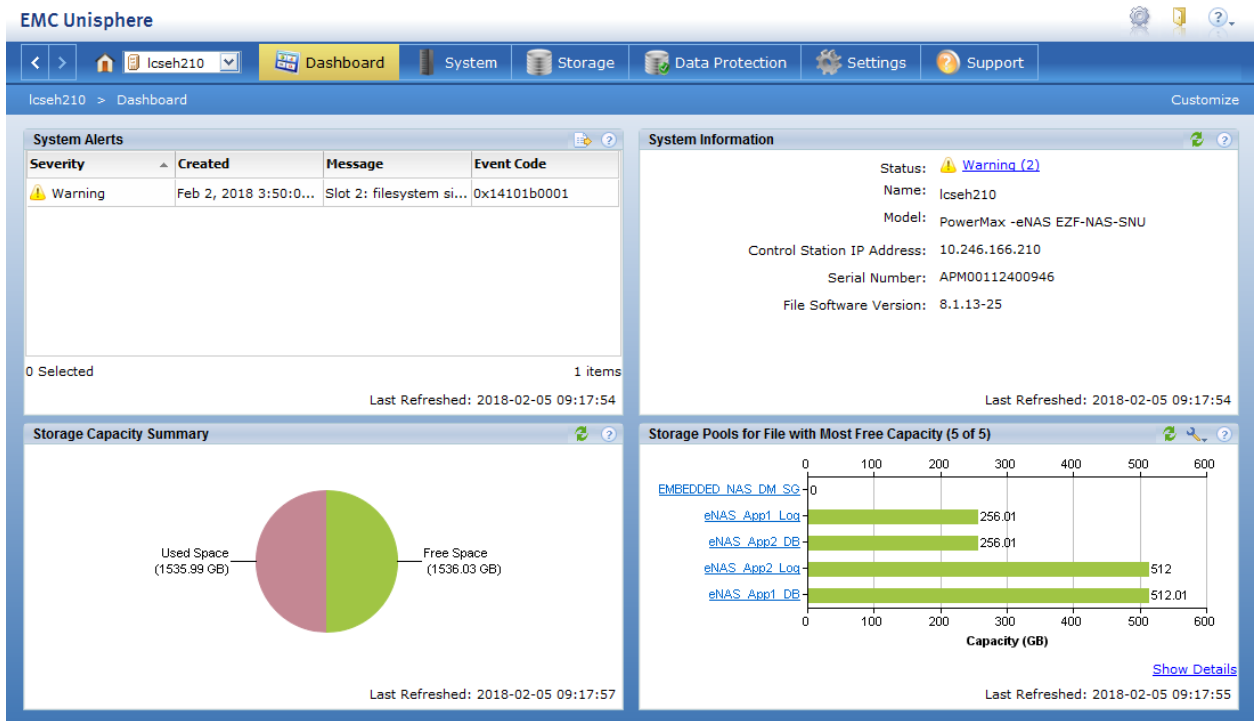


Figure 14. EMC Unisphere

PROVISION STORAGE FOR FILE

Creating or expanding storage for file has been simplified and automated with Unisphere for PowerMax and eNAS. Using the **Provision Storage for File** wizard, see Figure 15, Service Levels, Compression, and Dynamic Host I/O Limits can be assigned to storage groups at the time of provisioning. After any change to the eNAS storage, an optional automatic and intelligent discovery of devices occurs, called Diskmark. Diskmark maps managed devices to system mapped pools and non-system managed devices to the system defined pool called symm_dsl. Diskmark also updates the properties for pools and file systems for changes that happen after the Storage Group is created, such as a change in the Service Level.

Provision Storage For File

1 Storage Group | 2 Summary

Storage Group Name * eNAS_App1 | Storage Resource Pool SRP_1

Name	Service Level	Volumes	Volume Capacity
eNAS_App1_1	Diamond	8	256 GB
eNAS_App1_2	Silver	8	128 GB

Total Capacity **3072 GB** | Total Service Levels **2**

Enable Compression (2 of 2)

? | CANCEL | NEXT

Figure 15. Provision Storage for File

CONFIGURING LINK AND LAUNCH

In order to use the context sensitive links in EMC Unisphere, you first register the Link and Launch Client. To create the client registration in Unisphere for PowerMax, select **Settings**, and select **Management > Link and Launch**.

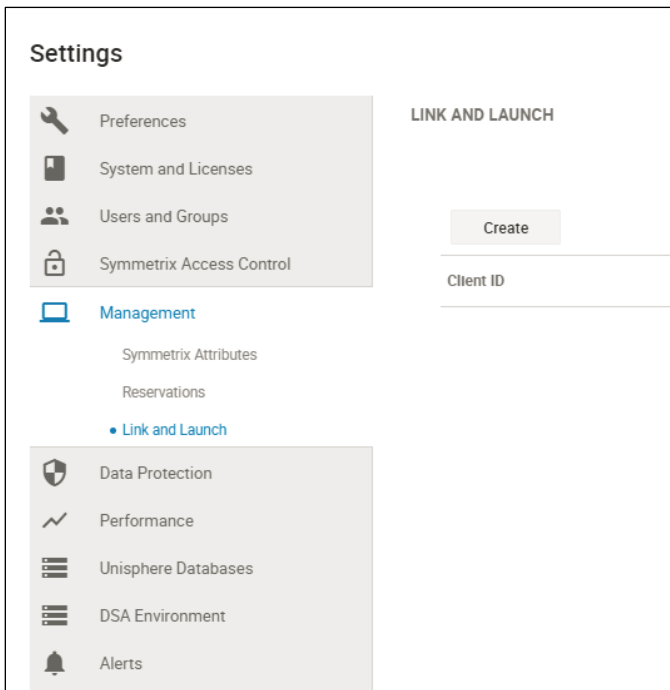


Figure 16. Unisphere for PowerMax - Link and Launch

Click the **Create** button to open the **Register Launch Client** dialog box. Provide a unique Client ID and password to be associated with the client. This information is be entered in EMC Unisphere to complete the registration.

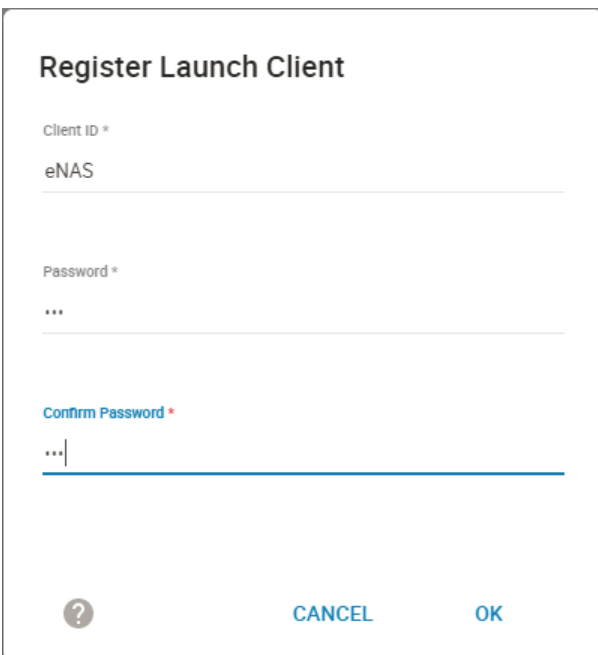


Figure 17. Unisphere for PowerMax - Register Launch Client

Once the launch client is registered in Unisphere for PowerMax, open EMC Unisphere. From the **Storage** pane, in the Configure Unisphere for PowerMax task list, click **Register Unisphere for PowerMax**.

In the Configure Unisphere PowerMax Credentials window opens, the Symmetrix ID field is populated with the PowerMax or VMAX serial number.

Select from the drop-down the Unisphere for PowerMax Host. The port and version fields are populated.

Fill in the Username, Client ID, and Client Password that were setup in Unisphere for PowerMax.

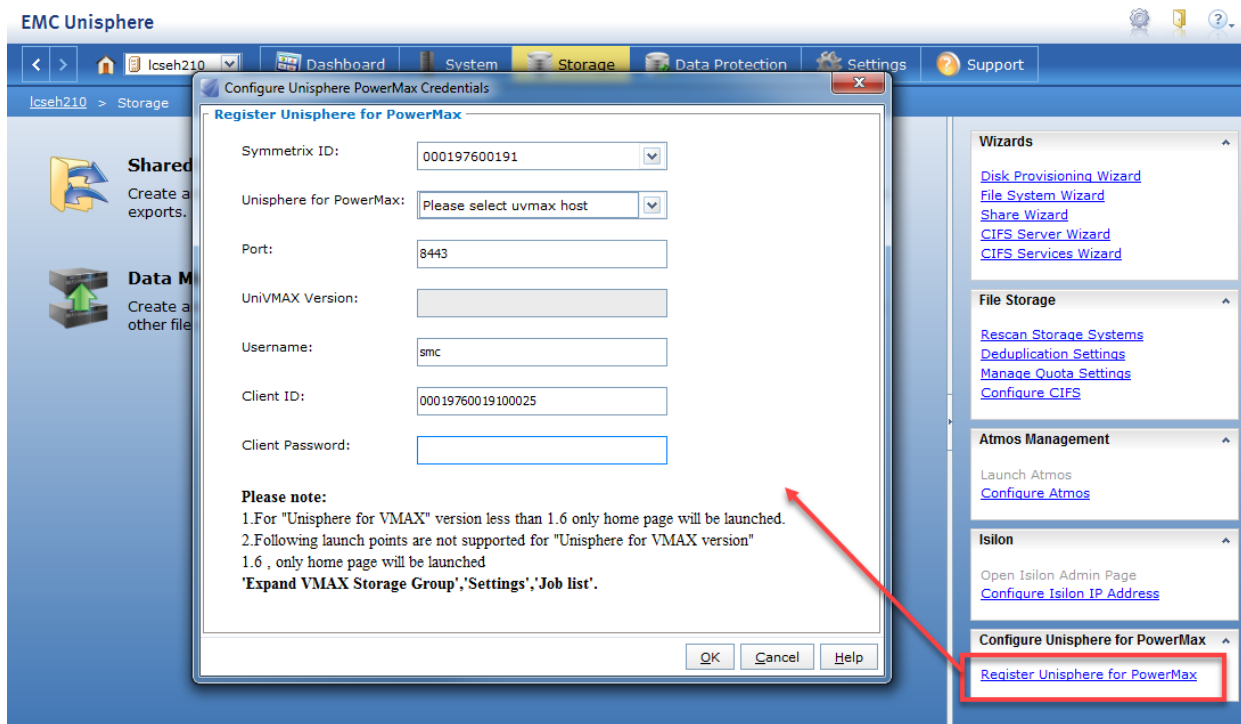


Figure 18. Register Unisphere for VMAX

Once Link and Launch is registered, from EMC Unisphere the context sensitive links can be used to view the Dashboard, Alerts, Performance, and Expand PowerMax and VMAX Storage Groups from Unisphere for PowerMax.

ENABLING ALERTS

The alerts that appear on the File Dashboard in Unisphere for PowerMax by default are disabled. This is to minimize the number of alerts storage administrators initially receive. You can enable the alerts in the Alert Policies settings (see Figure 19). To access the Alert Policies in Unisphere for PowerMax, select **Settings**, and select **Alerts > Alert Policies**.

Settings

- Preferences
- System and Licenses
- Users and Groups
- Symmetrix Access Control
- Management
- Data Protection
- Performance
- Unisphere Databases
- DSA Environment

Alerts

- Alert Policies**
- Compliance Alert Policies
- Performance Threshold and Alerts
- Symmetrix Threshold and Alerts
- Notifications

ALERT POLICIES

Select Array
All

Name	Type	Enabled	Notifications
000197800130	Array	<input type="checkbox"/>	NONE
▼ Storage Pool Projection Alert	File	<input checked="" type="checkbox"/>	NONE
000197600191	File	<input checked="" type="checkbox"/>	NONE
000197800130	File	<input checked="" type="checkbox"/>	NONE
▼ Storage Pool Projection Almost Full Alert	File	<input checked="" type="checkbox"/>	NONE
000197600191	File	<input checked="" type="checkbox"/>	NONE
000197800130	File	<input checked="" type="checkbox"/>	NONE
▼ Storage Pool Projection Full Alert	File	<input checked="" type="checkbox"/>	NONE
000197600191	File	<input checked="" type="checkbox"/>	NONE
000197800130	File	<input checked="" type="checkbox"/>	NONE
▼ Storage Pool Usage Available Alert	File	<input checked="" type="checkbox"/>	NONE
000197600191	File	<input checked="" type="checkbox"/>	NONE

CANCEL APPLY CLOSE

Figure 19. Enable File Dashboard Alerts

EMBEDDED NAS FILE AUTO RECOVERY

FILE AUTO RECOVERY

Introduced in HYPERMAX OS 5977.691.684, the File Auto Recovery (FAR) feature provides the ability to use the eNAS CLI to manually failover or move a virtual Data Mover (VDM) from a source eNAS system to a destination eNAS system. The failover or move uses block-level Symmetrix Remote Data Facility (SRDF) synchronous replication, so the failover or move involves zero data loss in the event of the failover/move being unplanned. This feature consolidates VDMs, file systems, file system checkpoints and schedules, CIFS servers, networking, and VDM configurations into their own separate pools. This feature works for a true disaster recovery (DR) where the source becomes unavailable. For DR support in the event of an unplanned failover, an option is provided to recover and clean up the source system to make it ready as a future destination.

Users can manually initiate failover and reverse operations using the EMC File Auto Recovery Manager (FARM) product. FARM provides the ability to automatically failover a selected sync-replicated VDM on a source eNAS system to a destination eNAS system. FARM can be used to set up monitoring of configured sync-replicated VDMs and to trigger automatic failovers based on Data Mover, File System, Control Station, or IP network unavailability that would cause the NAS client to lose access to data.

File Auto Recovery (FAR) can be used in the case of disaster recovery, maintenance, and load balancing operations. FAR allows active VDMs from either the local or remote sites to be protected as shown in Figure 20.

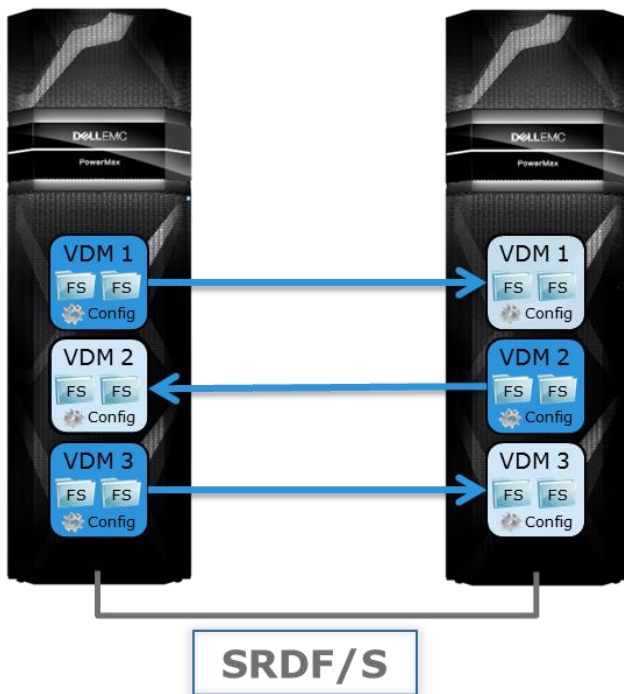


Figure 20. File Auto Recovery

File Auto Recovery has the following requirements:

- PowerMaxOS or HYPERMAX OS 5977.691.684 (eNAS File OE 8.1.7 or later)
- SRDF configured and enabled
- VDM enabled with “syncreplicable” option
- VDM with Split Log File System
- Checkpoints (SavVol) must be on the same pool as the production file system

FAILOVER

In the event of a primary site failure due to unplanned environmental catastrophes or any system failures, the external NFS/CIFS clients that are using the file system(s) served by the VDM(s) active on the primary site experience a Data Unavailability (DU) situation. In such scenarios, the File Storage Administrator can run the replication failover option to quickly mitigate the DU scenario.

Meet the following pre-requisites before running the failover:

- The primary site has one or more of these issues:
 - The PowerMax or VMAX itself failed due to power fluctuations, bad hardware or a catastrophic environmental disaster.
 - The Embedded NAS Data Mover(s) are not responding.
 - The network interface on the Data Mover(s) is malfunctioning due to failure of the hardware IO Modules.
 - Any unplanned disasters that would stop the NFS/CIFS clients from accessing their file systems served by the syncreplicable VDM.
- Run failover on the secondary site only.

Using the following command to failover the replication session:

```
nas_syncrep -failover <session_name>
```

Figure 21 shows the failover operation of a VDM replication session named HR_App1_session1.

Depending on the configuration and load on the syncreplicable source and destination VDM, the failover operation takes some time to successfully complete.

```
[nasadmin@eNAS_CS0 ~]$ nas_syncrep -failover HR_App1_session1

WARNING: You have just issued the nas_syncrep -failover command. Verify whether the
peer system or any of its file storage resources are accessible. If they are, then
you should issue the nas_syncrep -reverse command instead. Running the nas_syncrep -
failover command while the peer system is still accessible could result in Data
Unavailability or Data Loss. Are you sure you want to proceed? [yes or no] yes

Now doing precondition check... done: 53 s
Now doing health check... done: 0 s
Now cleaning local... done: 0 s
Now importing sync replica of NAS database... done: 111s
Now creating VDM... done: 4 s
Now importing VDM settings... done: 0 s
Now mounting exported FS(s)/checkpoint(s)... done: 5 s
Now loading VDM... done: 2 s
Now turning up local network interface(s)... done: 1 s
Service outage end: 176s
Now mounting unexported FS(s)/checkpoint(s)... done: 0 s
Now importing schedule(s)... done: 0 s
Elapsed time: 176s

done
```

Figure 21. VDM Session Failover

CLEANING A FILE AUTO RECOVERY SESSION

When the failover of a VDM sync replication session successfully completes, the primary site is not clean (unless the auto-clean operation as part of the failover succeeds) with all the stale NAS objects like the syncreplicable VDM, filesystem(s), checkpoint(s), and checkpoint schedule(s).

The `--clean --all` operation has the following pre-requisites:

- VDM sync replication session is in a failed over state

- The original primary site has been recovered
- The RDF link status between the local and remote sites are online

Run the following `-clean` command on the primary site only:

```
nas_syncprep -Clean -all
```

Figure 22 shows the output of running the clean operation.

```
[nasadmin@eNAS_CS0 ~]$ nas_syncprep -Clean -all

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

Now cleaning session HR_App1_session1 (may take several minutes)...
done
Now starting session HR_App1_session1...                already started, skipped
```

Figure 22. Clean VDM Session

REVERSING REPLICATION

If the primary site is going to be unavailable due to planned maintenance, the Storage Administrator can run the VDM sync replication -reverse operation. This reduces the overall impact of existing file system IO and ensures the original primary is cleaned up as part of the same operation.

The `-reverse` operation has the following pre-requisites:

- Reverse requires both sites to be operational and the VDM sync replication service session and data session is in an `in_sync` state.

Run the reverse operation on the secondary site only.

```
[nasadmin@eNAS_CS0 ~]$ nas_syncprep -reverse HR_App1_session1

WARNING: There will be a period of Data Unavailability during the reverse operation,
and, after the reverse operation, the VDM/FS(s)/checkpoint(s) protected by the sync
replication session will be reversed to the local site. Are you sure you want to
proceed? [yes or no] yes

Now doing precondition check...                done: 68 s
Now doing health check...                      done: 24 s
Now cleaning local...                          done: 8 s

                                Service outage start.....
Now turning down remote network interface(s)... done: 4 s
Now switching the session (may take several minutes)... done: 106s
Now importing sync replica of NAS database...   done: 94 s
Now creating VDM...                          done: 6 s
Now importing VDM settings...                 done: 0 s
Now turning up local network interface(s)...   done: 1 s
                                Service outage end: 211s

Now importing schedule(s)...                  done: 0 s
Now unloading remote VDM/FS(s)/checkpoint(s)... done: 2 s
Now cleaning remote...                        done: 11 s
                                Elapsed time: 324s

done
```

Figure 23. VDM Session Reverse

ADDING STORAGE (REFRESH PAIRS)

To accommodate future file system storage needs on an existing replication session, the File Auto Recovery feature has an option to add new disk volumes and refresh the RDF session without impacting the existing RDF session(s).

To extend the storage, first add new devices to local and remote storage groups using Unisphere or CLI. Then run the `nas_syncrep -refresh_pairs` command as shown in Figure 24.

```
[nasadmin@eNAS_CS0 ~]$ nas_syncrep -refresh_pairs HR_App1_session1
Now validating local bonded pool... done
Now validating remote bonded pool... done
Now validating RDF pairs in the mapped pools... done
Now extending the RDF session... done
done
```

Figure 24. Adding Storage to VDM Session

FAR HEALTH CHECK

Included in PowerMaxOS and HYPERMAX OS 5977.1125.1125 Release (File OE 8.1.12) or later is a FAR Health Check option to the existing `nas_syncrep` command. The FAR Health Check analyzes FAR sessions for conditions and scenarios that can cause a failover or reverse operation to fail.

It is recommended to periodically perform a FAR Health Check to monitor the health of VDM sync sessions. Using cron jobs, the FAR Health Check can be automated. This will enable you to take proactive measures to help minimize the chances that an unplanned failover or reverse operation will fail.

Table 2 lists the conditions and scenarios analyzed by a FAR Health Check.

Table 2. FAR Health Checks

Check	Description
SRDF group state	Checks if the relevant SRDF group exists and if it is in the correct state. The SRDF groups and the directors should be in the Online state for this check to pass.
SRDF session(s) status	Checks if the SRDF session is in the correct state. It checks for the existence and validity of the device groups associated with the VDM Sync sessions from the PowerMax side.
VDM Sync session(s) disktype(s)	Checks for the disks to be of the correct type. All the disks involved in a VDM Sync session need to be of the same disk type.
File system has 128K free space	Checks if the File systems mounted on VDM have at least 128K free space. During failover or reverse, if file system to be mounted does not have at least 128K free space, the file system could be marked corrupt due to GID corruption.
Network Configuration Check	Checks if VDM interfaces are present on the local and remote sites.
Check SRDF link status	Checks if the remote SRDF link is pingable, using the <code>SymCli</code> command.
Data Mover status	Checks if the Data Movers on the local and remote sites have reason codes four or five. A sample <code>getreason</code> output : 10 - slot_0 primary control station 11 - slot_1 secondary control station 5 - slot_2 contacted 5 - slot_3 contacted

Check	Description
eNAS, Solutions Enabler, PowerMaxOS version check	Checks if the local and the remote have the same or in-family versions of eNAS, Solutions Enabler and PowerMaxOS.
Filesystem ID consistency	Checks if the local and remote file system IDs overlap.
Pool to SRDF session mapping	Checks if all the devices from mapped pool are added to the SRDF session.
Director ports online/offline	Checks for the availability of PowerMax or VMAX directors and ports related to eNAS.
Equivalent Data Services	Checks if the local and remote have equivalent data services like SLO (service level objective), SRP and Host I/O Limits.

The success of FAR Health Check does not guarantee that failover or reverse operations will succeed. The states of various entities involved in sessions, such as interfaces, and Data Movers can change over time, and could result in the failure of the failover and reverse operations. The health check is intended to be used for proactive monitoring of the VDM sync sessions.

Use the following command to run a FAR Health Check:

```
nas_syncrep -health_check
```

Figure 25 shows the health check operation.

```
[nasadmin@eNAS_CS0 ~]$ nas_syncrep -health_check

Health check starting ...
Initializing ...
Check No.      | Check Name                                     | Message ID | Status | Brief Description
Check ( 1/12 ) | Check SRDF Group State                       | 34906964006 | PASS  | SRDF Group online.
Check ( 2/12 ) | Check SRDF session(s) status                 | 34906964010 | PASS  | SRDF session in sync.
Check ( 3/12 ) | Check VDM Sync session(s) disktype(s)       | 34906964014 | PASS  | VDM Sync session
disktypes in correct state.
Check ( 4/12 ) | File system has 128K free space               | 34906964018 | PASS  | File system has at
least 1MB free space.
Check ( 5/12 ) | Network Configuration Check                  | 34906963984 | PASS  | Interfaces correct.
Check ( 6/12 ) | Check SRDF link status                       | 34906963989 | PASS  | Remote SRDF link is
pingable.
Check ( 7/12 ) | Data Mover status                            | 34906963992 | PASS  | Data Movers are in the
correct state.
Check ( 8/12 ) | eNAS, SE, Enginuity version check            | 34906963993 | PASS  | eNAS, Solutions Enabler
and microcode version check passed.
Check ( 9/12 ) | Check for filesystem ID consistency           | 34906963997 | PASS  | No conflict in file
system IDs.
Check ( 10/12 ) | Pool to SRDF session mapping                 | 34906963981 | PASS  | VDM Sync session(s)
disk type OK.
Check ( 11/12 ) | Check for director ports online/offline      | 34906963999 | PASS  | VDM Sync session's
directors and ports online.
Check ( 12/12 ) | Check for Equivalent Data Services           | 34906964026 | PASS  | Data services (Service
level and Host IO Limits) match.
Health check complete.
Check /nas/log/nas_syncrep.log for more details.
```

Figure 25. FAR Health Check

Preserve ReplicatorV2 Sessions for FAR

With synchronous replication enabled between two systems, it is possible to add asynchronous replication to a third system using ReplicatorV2 (RepV2). This allows the third system to be located further away and enables it to be used as a backup and recovery solution.

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

Figure 26 shows the coexistence of FAR and file system Replicator sessions. There are three PowerMax or VMAX systems with eNAS: System A, System B, and System C. A single VDM with several file systems is synchronously replicated from System A to System B using FAR. There is also an asynchronous Replicator session created from System A to System C for each file system on this VDM. This allows the file systems on the VDM to be simultaneously protected by synchronous and asynchronous replication.

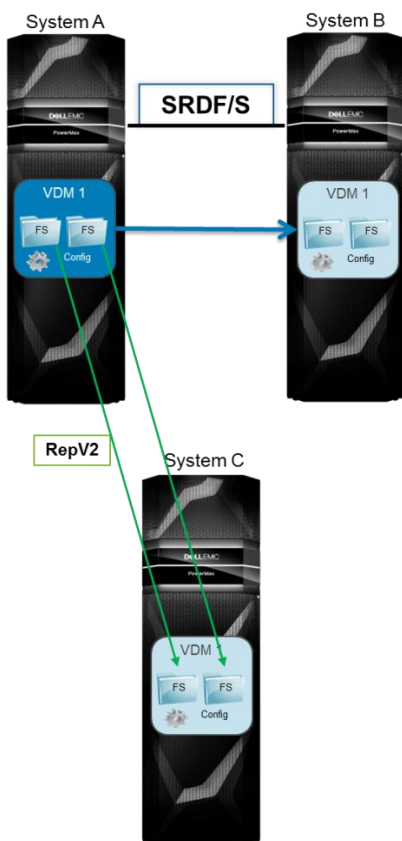


Figure 26. File Auto Recovery with Preserve ReplicatorV2

The Preserve ReplicatorV2 Sessions feature has the following requirements:

- Three PowerMax, VMAX All Flash or VMAX3 systems with eNAS
 - Two eNAS systems used for FAR (Systems A and B)
 - One eNAS system used as a RepV2 destination (System C)
 - PowerMaxOS or HYPERMAX OS 5977.1125.1125 Release (eNAS File OE 8.1.12) or later
- The FAR service enabled
- The systems have the RepV2 license enabled

For more information about the configuration and management of FAR and Preserve ReplicatorV2 sessions, refer to the *Dell EMC PowerMax eNAS File Auto Recovery with SRDF/S* documentation.

FILE AUTO RECOVERY MANAGER

Dell EMC File Auto Recovery Manager (FARM) is an optional Windows utility that can manage FAR sessions. In addition to the manually initiated operations, FARM's automated failover is provided by a background service that you can manage through a GUI.

FARM can monitor replication sessions from one source eNAS system to one target eNAS system. If there are active VDMs on both sides, a second Windows host with FARM will need to be setup to monitor the sessions going in the other direction.

Figure 27 below shows FARM monitoring active VDMs on the source eNAS system.

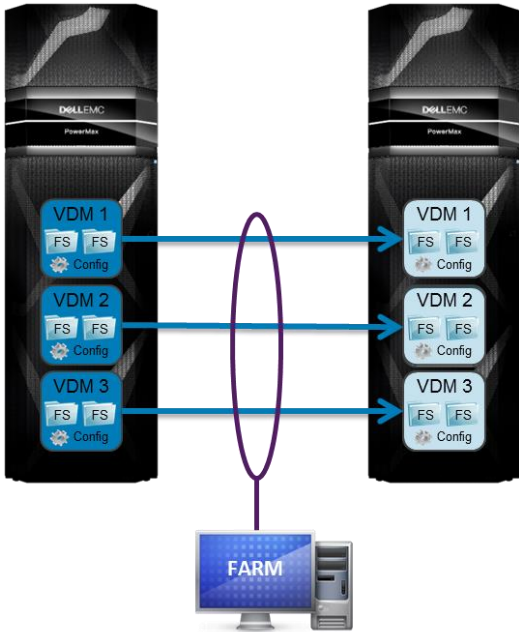


Figure 27. File Auto Recovery Manager

Using the FARM GUI, manual failover, reverse, and restore operations can be run as shown in Figure 28.

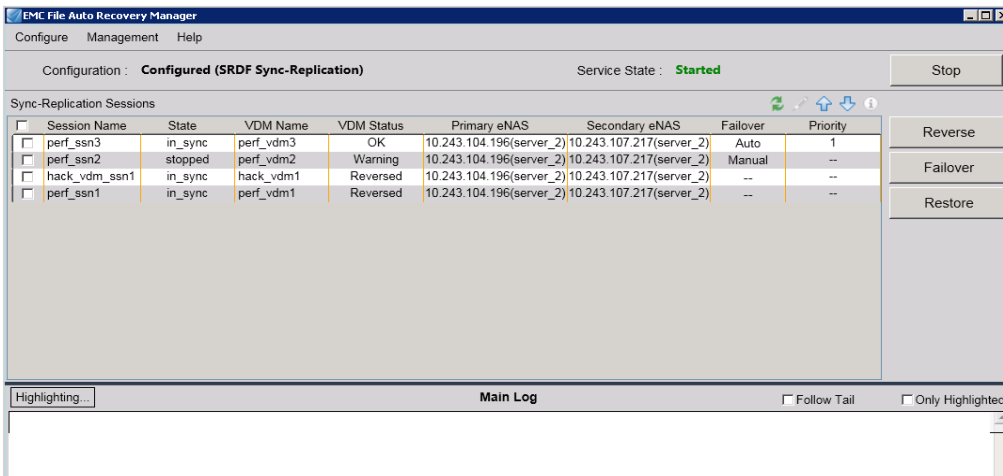


Figure 28. File Auto Recovery Manager Operations

For more information about the FARM product, refer to the *Dell EMC PowerMax Embedded NAS File Auto Recovery Manager Product Guide*.

CONCLUSION

[PowerMax](#) and VMAX unified storage introduces embedded file data services that enable customers to consolidate islands of block and file storage, simplify management, and reduce deployment costs by up to 33%. eNAS uses the hypervisor provided in PowerMaxOS and HYPERMAX OS 5977 to create and run a set of virtual machines on PowerMax and VMAX controllers. These virtual machines host two major elements of eNAS: software data movers and control stations and are distributed based on the mirrored pair architecture of PowerMax, VMAX All Flash and VMAX3 to evenly consume resources for both performance and capacity.

eNAS extends the value of PowerMax, VMAX All Flash, and VMAX3 to file storage by enabling customers to leverage vital enterprise features including Service Level Provisioning, Compression, and Host I/O limits technologies for both block and file storage. eNAS is a multi-controller, transaction NAS solution designed for customers requiring hyper consolidation for block storage (the traditional PowerMax and VMAX use case) combined with moderate file storage in mission-critical environments.

PowerMax, VMAX All Flash, and VMAX3 systems with eNAS can now achieve even higher availability by using eNAS File Auto Recovery. This solution allows customers to replicate at a more granular level of storage using VDMs in order to protect a specific set of file systems. Using the optional File Auto Recovery Manager, automatic failovers can be accomplished reducing Recovery Time Objective (RTO).

REFERENCES

Reference information and product information can be found at www.dell EMC.com including:

Dell EMC PowerMax eNAS CLI Reference Guide

Dell EMC PowerMax eNAS Quick Start Guide

Dell EMC PowerMax eNAS File Auto Recovery with SRDF/s

Dell EMC PowerMax eNAS Release Notes

Dell EMC PowerMax Family Product Guide

Dell EMC Service Levels for PowerMaxOS

Dell EMC Unisphere for PowerMax Installation Guide

Dell EMC Unisphere for PowerMax Release Notes

Dell EMC VMAX3 & VMAX All Flash eNAS Best Practices

VNX OE for File 8.1 Documentation

Dell EMC VMAX3 Family with HYPERMAX OS Product Guide

Dell EMC VMAX All Flash Product Guide

Dell EMC VMAX3 Service Level Provisioning with Fully Automated Storage Tiering (FAST)

HYPERMAX OS 5977 for VMAX3 Family and VMAX All Flash Release Notes