

Dell EMC VMAX All Flash and VMAX3: Reliability, Availability, and Serviceability

Abstract

This white paper explains the reliability, availability, and serviceability (RAS) hardware, as well as the software features of Dell EMC™ VMAX All Flash and VMAX3 arrays.

December 2020

Revisions

Date	Description
September 2019	Content and template update
December 2020	PowerMaxOS Q3 2020 release updates, and updates to address frequent questions

Acknowledgements

Author: Michael Bresnahan

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

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

Copyright © 2019 Dell Inc. or its subsidiaries. All Rights Reserved. Dell, EMC, Dell EMC and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be trademarks of their respective owners. [12/5/2020] [Technical White Paper] [H13807.10]

Table of contents

Revisions.....	2
Acknowledgements.....	2
Table of contents	3
Executive summary.....	5
Audience	6
1 Introduction.....	7
2 Active-Active Architecture	8
3 Remote support.....	9
3.1 Supportability through the management module control station.....	9
3.2 Secure Service Credential (SSC), Secured by RSA	10
4 Error detection.....	11
4.1 T10 Data Integrity Field	11
4.2 Data integrity checks	11
4.3 Drive monitoring and correction.....	12
4.4 Physical memory error verification and error correction.....	12
5 Reliable components.....	13
5.1 Component-level redundancy.....	13
5.2 Redundant engine components.....	13
5.3 Channel front-end redundancy.....	19
5.4 SAS back-end redundancy.....	21
5.5 Drive Array Enclosure.....	22
6 Dynamic virtual matrix.....	26
6.1 Internal environmental Ethernet connectivity	26
7 Redundant power subsystem.....	28
7.1 Battery backup unit modules	28
7.2 Vaulting.....	28
7.3 Vault triggers.....	28
7.4 Power-down operation.....	29
7.5 Power-up operation	29
8 Data protection methods	30
8.1 RAID 1 (Mirroring).....	30
8.2 RAID 5	30
8.3 RAID 6	31
8.4 Local RAID.....	31

8.5	Thin provisioning.....	31
8.6	Drive sparing.....	32
8.7	Local replication using TimeFinder.....	35
8.8	Remote replication using SRDF.....	37
8.9	Application I/O serviced by remote array.....	39
8.10	Data at Rest Encryption.....	41
9	Component-level serviceability.....	43
9.1	Flashing rack lights.....	44
10	Non-disruptive upgrades.....	45
10.1	PowerMaxOS and HYPERMAX OS upgrades.....	45
10.2	eNAS upgrades.....	45
10.3	Hardware upgrades.....	45
11	VMAX 250F.....	46
11.1	VMAX 250F V-Brick.....	46
11.2	VMAX 250F back-end.....	47
11.3	VMAX 250F system power.....	47
11.4	VMAX 250F serviceability.....	47
12	VMAX 950F.....	48
12.1	VMAX 950F V-Brick.....	48
12.2	VMAX 950F back-end.....	49
12.3	VMAX 950F system power.....	49
12.4	VMAX 950F serviceability.....	49
13	Unisphere for PowerMax and Solutions Enabler.....	50
13.1	Unisphere for PowerMax system health check.....	50
13.2	Unisphere alerts.....	51
13.3	Solutions Enabler commands.....	52
14	Summary.....	54
A	Resiliency testing.....	55
A.1	Dell EMC internal QE testing.....	55
A.2	On-site proof-of-concept demonstrations.....	56
B	Technical support and resources.....	57
B.1	Related resources.....	57

Executive summary

The IT industry has been growing and changing at a rapid pace, with a large area of focus on cloud computing and cloud-like consumption models. IT organizations are finding that they must be able to cater to both internal and external customers' storage needs, while maintaining alignment with the budget that is provided to them.

Dell EMC™ VMAX™ All Flash and VMAX3™ platforms accommodate all these needs. They provide a simple, service level-based provisioning model that changes the way users consume storage, taking the focus away from the back-end configuration steps and allowing them to concentrate on other key roles.

While simplifying storage consumption is critical, other features also create a powerful platform. Redundant hardware components and intelligent software architecture deliver extreme performance while also providing high availability. This combination provides exceptional reliability, while also leveraging components in new and innovative ways which decreases the total cost of ownership of each system.

Important functionality such as local and remote replication of data used to deliver business continuity must cope with more data than ever before, without impacting production activities. Furthermore, at the end of the day, all these challenges must be met while continually improving economics.

Highlights include:

Service level-based provisioning: The requirements of a storage consumer revolve around capacity and performance. It is important for them to have the right amount of capacity to fulfill their needs as well as the proper performance levels. Service level-based provisioning gives VMAX3 users the opportunity to select a predefined service level, provision storage to it, and get the performance they need, without worrying about back-end configurations. VMAX All Flash systems are architected with a single tier and with the highest possible service level.

Performance: The processing power, memory, and bandwidth deliver high levels of predictable performance. This means providing the performance to meet the IOPS and MB/s needs of the world's most demanding applications. VMAX All Flash and VMAX3 arrays are architected with the ability to absorb I/O bursts due to unpredictable workloads, such as end-of-quarter processing and performance intensive *ad hoc* queries, while continuing to guarantee consistent application-response times.

Information-centric security, built-in: Integrated RSA technology provides industry-leading, information-centric security to secure people, infrastructure, and data. Security features reduce risk and protect information. Organizations can authenticate, authorize, and audit activities on systems and devices.

Always-on availability (HA/RAS): Reliability, availability, and serviceability (RAS) features are crucial for enterprise environments requiring always-on availability. VMAX All Flash and VMAX3 platforms are architected for six-nines (99.9999%) availability. The many redundant features discussed in this document are considered in the calculation of overall system availability. This includes redundancy in the back-end, cache memory, front-end and fabric, as well as the types of RAID protections given to volumes on the back-end. Calculations may also include time to replace failed or failing FRUs (field replaceable units) which also consider customer service levels, replacement rates of the various FRUs, and hot sparing capability in the case of drives.





Eliminate Costly Downtime	Exceed Stringent Replication SLAs (RTO, RPO)	Eliminate Planned Downtime	Ensure 100% Data Integrity
			
<p>Proven 6 Nines of Availability Advanced Fault Isolation, map-out faulty memory DIMMS, mirrored memory no single points of failure</p>	<p>Gold Standard in Multi-Site Replication Proven Disaster Recovery and rapid restart; 2-site, 3-site replication</p>	<p>Non-Disruptive HW and SW Upgrades Continuous IO through parallel microcode NDUs, upgrade HYPERMAX O/S within seconds</p>	<p>T10 DIF Data Coding Single Bit Error Correction, validation checksum through T10 DIFF</p>

Figure 1 Reliability, Availability, and Serviceability highlights

VMAX All Flash and VMAX3 have raised customer expectations for high-end storage in terms of availability, performance, replication, scalability, and management. High-end availability is more than just redundancy; it means non-disruptive operations and upgrades and being “always online.” Delivering high-end performance means handling all workloads, predictable or not, under all conditions. High-end replication means any amount of data anytime and sending it any distance. High-end scalability means more than capacity; it means having the flexibility to handle any service level or application, cost-effectively, no matter how your business changes. And high-end storage management goes beyond monitoring storage arrays; it means managing the service levels the business expects, from provisioning to business continuity. Once again, the standard in high-end storage has been redefined.

Audience

This document is for anyone who needs to understand how the components and technology in VMAX All Flash and VMAX3 systems provide a highly reliable and available platforms. It is also intended for individuals who would like to know more about the serviceability aspects, including Dell EMC customers, sales, and field technical staff.

1 Introduction

Binding service-level agreements (SLAs) commit IT organizations to deliver agreed-to and measurable support metrics for application performance, end-user response time, and system availability. In the event of a component failure, such as a faulty disk drive or power supply, these organizations are required to provide the highest levels of performance, availability, and serviceability – without compromise.

IT executives around the globe have recognized that downtime is not only measured in minutes or hours, but is also calculated as potential revenue loss, missed opportunities, or dissatisfied customers. Any operational impacts resulting from a component failure must be completely transparent to the individual applications and users relying on information availability to drive the business.

Today's mission-critical environments require a high-end storage solution that guarantees uncompromised levels of service, backed by a vendor that will provide quality support and a smooth migration path to new technologies. This solution must include a robust architectural design to withstand potential failures without impacting data availability.

VMAX All Flash and VMAX3 are based on a revolutionary design and include key enhancements that improve the reliability, availability, and serviceability of the new systems – ideal for critical applications and 24x7 environments demanding uninterrupted access to information.

The objective of this technical note is to provide an overview of the reliability, availability, and serviceability (RAS) features architected into VMAX All Flash and VMAX3 storage arrays.

2 Active-Active Architecture

VMAX All Flash and VMAX3 systems aggregate up to sixteen directors with fully shared connectivity, processing, memory, and storage capacity resources. Each pair of highly available directors is contained within a common shell as an engine. The directors operate in a truly symmetric active-active fashion delivering sub-controller fault isolation and component-level redundancy. Each director contains hot-pluggable I/O modules that provide frontend host connectivity, SRDF connectivity, and backend connectivity. Management modules on each director provide environmental monitoring and system management intercommunications to all other directors in the system. Each director also has its own redundant power and cooling subsystems.

Redundant internal fabric I/O modules on each director provide communication interconnection between all directors. This technology connects all directors in the system to provide a powerful form of redundancy and performance by allowing the directors to share resources and act as a single entity. Data behind one director can be accessed by any other director without performance considerations.

The memory modules on each director are collectively distributed as a unified global cache. Global cache/memory is accessible by all directors in the array. When a new write operation is committed to memory, the new data is immediately available to all processors within every engine. While the data is protected in memory, the processors on all directors can work autonomously on the new data to update a mirrored pair, send the update over an SRDF link, update a TimeFinder copy, report the current status of all events to the management software, and handle error detection and correction of a failed component. All tasks can occur simultaneously, without de-staging to the drives and re-staging to a separate region in memory.

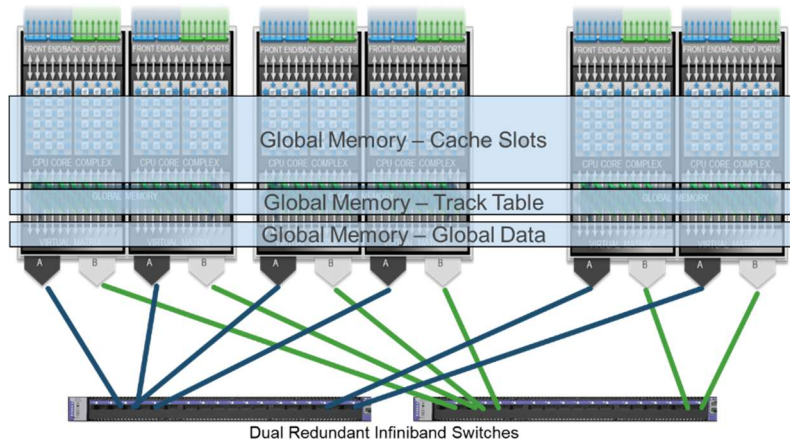


Figure 2 Globally shared resources

3 Remote support

Remote support is an important and integral part of Dell EMC Customer Support. Every VMAX All Flash and VMAX3 system has two integrated Management Module Control Stations (MMCS) that continuously monitor the VMAX environment. The MMCS's can communicate with the Customer Support Center through a network connection to the EMC Secure Remote Support (ESRS) Gateway.

Through the MMCS, the system actively monitors all I/O operations for errors and faults. By tracking these errors during normal operation, the system can recognize patterns of error activity and predict a potential hard failure before it occurs. This proactive error tracking capability can often prevent component failures by fencing off, or removing from service, a suspect component before a failure occurs. The call home capabilities allow the system to automatically notify Customer Support of potential issues before a failure occurs. A Dell EMC Technical Support Engineer handles these calls and can dispatch a local Customer Service Engineer to replace a component without disrupting access to data.

To provide remote support capabilities, the system is configured to call home and alert Customer Support of a potential failure. An authorized Technical Support Engineer can run system diagnostics remotely for further troubleshooting and resolution. Configuring Dell EMC products to allow inbound connectivity also enables Customer Support to proactively connect to the systems to gather needed diagnostic data or to attend to identified issues. The current connect-in support program for the system uses the latest digital key exchange technology for strong authentication, layered application security, and a centralized support infrastructure that places calls through an encrypted tunnel between Customer Support and the MMCS located inside the system.

Before anyone from Customer Support can initiate a connection to a system at the customer site, that person must be individually authenticated and determined to be an appropriate member of the Customer Support team. Field-based personnel who might be known to the customer must still be properly associated with the specific customer's account.

An essential part of the design of the connectivity support program is that the connection to the customer's MMCS must originate from one of several specifically designed Remote Support Networks at Dell EMC. Within each of those Support Centers, the necessary networking and security infrastructure has been built to enable both the call-EMC and call-device functions.

3.1 Supportability through the management module control station

The previous generation of VMAX had a single service processor in each system that was used for fault monitoring, as well as remote connectivity and maintenance provided by Customer Support. In VMAX All Flash and VMAX3, this has been extended to two management module control stations (MMCS) in the first engine of each system (one per director). Each MMCS is powered by the redundant power supplies located in its respective director board.

The MMCS located in director 1 is known as the primary MMCS, and the MMCS located in director 2 is known as the secondary MMCS. The primary MMCS provides all control station functionality when it is operating normally, while the secondary MMCS provides a subset of this functionality. If the primary MMCS fails, the secondary MMCS is put in an elevated secondary state, which allows more functionality for the duration of this state. Both MMCS are connected to the customer network, giving the system the redundant ability to report any errors to Customer Support, as well as allowing Customer Support to connect to the system remotely.

The MMCS is part of the following support and maintenance tasks:

- Troubleshooting by Customer Support
- Component replacement scripts
- Code loads and configuration & installation scripts
- Internal scheduler tasks that monitor the health of the system
- Error collection and logging
- Error reporting and remote connectivity

For more information on MMCS and component-level redundancy, see the [Management module control station \(MMCS\)](#) section of this document.

3.2 Secure Service Credential (SSC), Secured by RSA

The Secure Service Credential technology applies exclusively to service processor activities and not host-initiated actions on array devices. These service credentials describe who is logging in, the capabilities they have, a time frame that the credential is good for, and the auditing of actions the service personnel performed which can be found in the symaudit logs. If these credentials are not validated, the user cannot log in to the MMCS or other internal functions. SSC covers both on-site and remote login.

Some of the security features are transparent to the customer, such as service access authentication and authorization by Customer Support and SC (user ID information) restricted access (MMCS and Customer Support internal functions). Access is definable at a user level, not just at a host level. All user ID information is encrypted for secure storage within the array.

MMCS-based functions honor Solutions Enabler Access Control settings per authenticated user to limit view/control of non-owned devices in shared environments such as SRDF-connected systems.

4 Error detection

A suite of error and integrity checks ensure that data integrity is maintained in the event of a system failure or power outage. The systems are designed with these data integrity features:

- Block CRC error checks
- Data integrity checks
- Drive monitoring and correction
- Physical memory error

4.1 T10 Data Integrity Field

VMAX All Flash and VMAX3 arrays support industry standard T10 Data Integrity Field (DIF) block cyclic redundancy code (CRC) for track formats. For open systems, this enables a host-generated DIF CRC to be stored with user data and used for end-to-end data integrity validation. Additional protections for address and control fault modes provide increased levels of protection against faults. These protections are defined in user-definable blocks supported by the T10 standard. Address and write status information is stored in the extra bytes in the application tag and reference tag portion of the block CRC.

PowerMaxOS further increases data integrity with T10-DIF+ which has additional bits for detecting stale data address faults, control faults and sector signature faults that are not detected in a standard T10-DIF. T10-DIF+ is performed every time data is moved; across the internal fabric, to or from drives, and on the way back to the host on reads.

On the backend, the T10-DIF codes for the expected data is stored and the checksums are verified when the data is read from the host. In addition, a one-byte checksum for each 8 K of data is kept in the track table (not stored with the data) that is used for independent validation of the data against the last version written to the array. This provides protection against situations such as:

- Detecting reads from the wrong block: The data and checksum stored together are fine, but it was from the wrong address. In this case, the additional checksum will not match.
- RAID disagreement: Each data block and the parity block of the RAID group have valid checksums and no errors, but the parity does not match with the data. In this case, each of the data blocks can be validated to determine if a data block or the parity block are stale.

4.2 Data integrity checks

Data integrity is validated at every possible point during the lifetime of the data. From the point at which data enters an array, the data is continuously protected by error detection metadata. This protection metadata is checked by hardware and software mechanisms any time data is moved within the subsystem, allowing the array to provide true end-to-end integrity checking and protection against hardware or software vaults.

The protection metadata is appended to the data stream, and contains information describing the expected data location as well as CRC representation of the actual data contents. The expected values to be found in protection metadata are stored persistently in an area separate from the data stream. The protection metadata is used to validate the logical correctness of data being moved within the array any time the data transitions between protocol chips, internal buffers, internal data fabric endpoints, system cache, and system disks.

4.3 Drive monitoring and correction

Medium defects are monitored by both examining the result of each data transfer and proactively scanning the entire drive during idle time. If a block is determined to be bad, the director:

- Rebuilds the data in physical memory if necessary.
- Remaps the defective block to another area on the drive set aside for this purpose.
- Rewrites the data from physical memory back to the remapped block on the drive.

The director maps around any bad block(s) detected, thereby avoiding defects in the media. The director also keeps track of each bad block detected. If the number of bad blocks exceeds a predefined threshold, the primary MMCS invokes a sparing operation to replace the defective drive and then automatically alerts Customer Support to arrange for corrective action.

4.4 Physical memory error verification and error correction

Correctable single-bit errors report an error code once the single-bit errors reach a predefined threshold. When a multi-bit error occurs, the physical memory segment is fenced-off (removed from service). Data can be retrieved from mirrored memory (if it was unwritten) or from the physical drive. In the unlikely event that physical memory replacement is required, the array notifies Dell EMC support, and a replacement is ordered. The failed FRU is then sent back to Dell EMC for failure analysis.

5 Reliable components

VMAX All Flash and VMAX3 systems use components that have a mean time between failure (MTBF) of several hundred thousand to millions of hours for a minimal component failure rate. A redundant design allows systems to remain online and operational during component repair.

Periodically, the system tests all components. Errors and environmental conditions are reported to the host system as well as to the Dell EMC Customer Support Center.

5.1 Component-level redundancy

All critical components are fully redundant, including director boards, global memory, internal data paths, power supplies, battery backup, and all SAS back-end components. The following is an overview of the redundancy of each of these components.

5.2 Redundant engine components

The engine is a critical building block of VMAX All Flash and VMAX3 systems. It primarily consists of two redundant director boards, which house global memory, front-end connectivity, back-end connectivity, and internal network communications components. Even single-engine configurations are fully redundant.

Table 1 displays the supported V-Brick count for each VMAX All Flash model, and Table 2 displays the supported engine count for each VMAX3 model.

Table 1 Supported V-Brick count per VMAX All Flash model

VMAX All Flash Model	Supported Engine Count
250F	1-2
450F	1-4
850F	1-8
950F*	1-8

Note: More info on the VMAX 950F/950FX can be found on page 32.

Table 2 Supported engine count per VMAX3 model

VMAX3 Model	Supported Engine Count
100K	1-2
200K	1-4
400K	1-8

The following figures display front and rear views of a V-Brick / Engine.



Figure 3 Front view of V-brick / Engine



Figure 4 Rear view of V-Brick / Engine

5.2.1 Redundant director boards

As mentioned above, each engine consists of two director boards. Each director board has a dedicated power and cooling system. Table 3 lists the components within a director, the number of them, and defines their purposes.

Table 3 Director Components

Director Component	Count (per director)	Purpose
Power Supply	2	Provide redundant power to director
Fan	5	Provide cooling for director
Management Module	1	Manage environmental functionality
Flash I/O Module	Up to 4	Enables vault to flash capability
Front-end I/O Module	Up to 4	Provide front-end connectivity to the array; Fibre Channel SCSI, iSCSI, FICON, SRDF, and embedded NAS (eNAS)
Back-end I/O Module	2	Provide back-end connectivity to drives
InfiniBand (IB) module	1	Connects to IB interconnect switches
Memory Module	16	Global memory component

The following figures display front and rear views of director components.

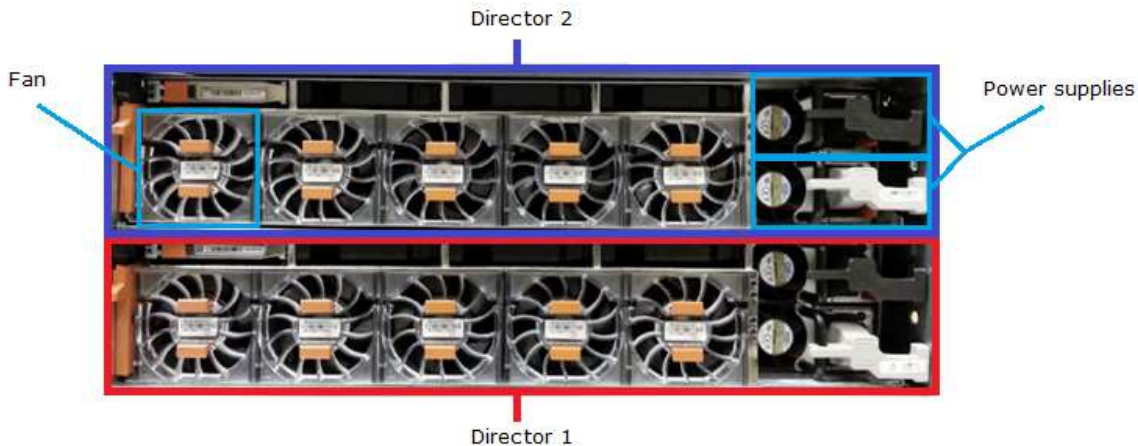


Figure 5 Front view of director components

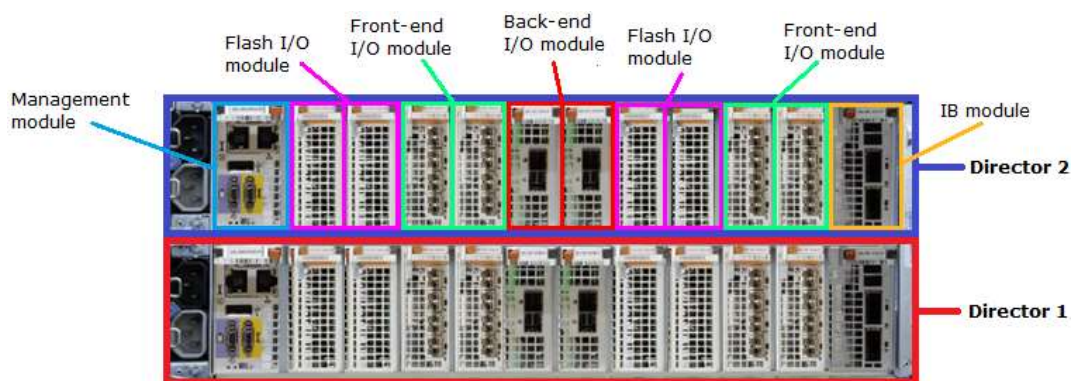


Figure 6 Rear view of director components

5.2.2 Management module control stations and management modules

There are two different types of management modules; Management Module Control Stations (MMCS) and standard Management Modules (MM). The first engine of each system will be deployed with an MMCS in each director. Each subsequent engine (engines 2-8) will be deployed with a management module in each director, in place of an MMCS.

5.2.3 Management module control station (MMCS)

The MMCS combines the management module and control station (service processor) hardware into a single module. It provides environmental monitoring capabilities for power, cooling, and connectivity. Each MMCS monitors one of the system standby power supplies (SPS) through a RS232 connection. Each MMCS is also connected to both internal Ethernet switches within the system as part of the internal communications and environmental control system.

For more information on internal communications and environmental controls, see the [Internal Ethernet connectivity](#) section of this document.

The MMCS also provide support functionality. Each MMCS connects to the customer's local area network (LAN) to allow monitoring of the system, as well as remote connectivity for the Customer Support team. The

primary MMCS is connected to the keyboard/video/mouse (KVM) inside the system and the secondary MMCS has these connection capabilities as a backup option. These can also be connected to an external laptop or KVM source.

For more information on supportability through the MMCS, see the [Supportability through the management module control station \(MMCS\)](#) section of this document.

The MMCS also controls the blue LED bars on the front and back of each bay. These can be used for system identification purposes by Customer Support. For more information, see the [System identification blinking bay LED feature](#) section of this document. The following figure illustrates MMCS connectivity.

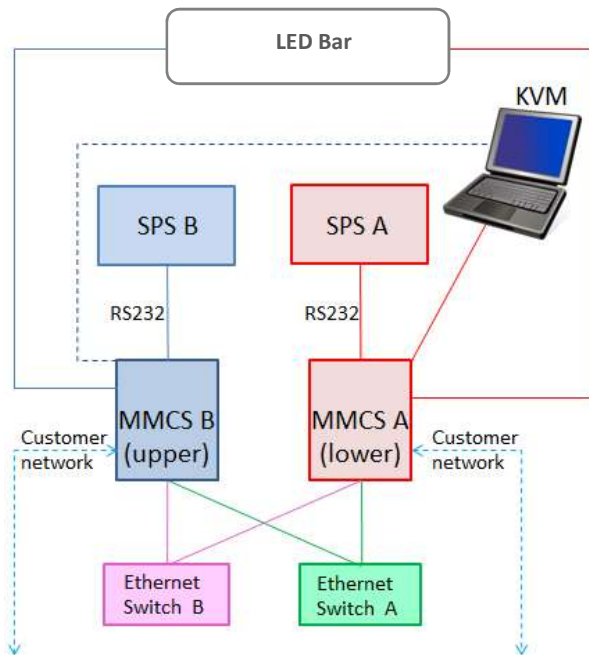


Figure 7 Management module control station connectivity

5.2.4 Management module (MM)

The management module has a subset of features of the MMCS, which does not include the control station functionality. Each management module has an RS232 connection to an SPS. Management module A connects only to Ethernet switch A, and management module B connects only to Ethernet switch B. Like the MMCS, each management module is responsible for monitoring and reporting any environmental issues, such as power, cooling, or connectivity problems. The following figure illustrates MMCS connectivity.

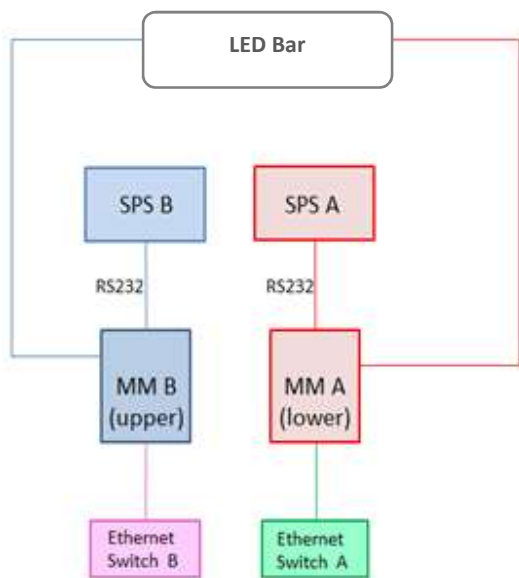


Figure 8 Management module connectivity

5.2.5 Flash I/O module

The flash I/O modules use NVMe technology to safely store data in cache during the vaulting sequence. For more information, see the [Vaulting](#) section of this document.

5.2.6 Front-end I/O module

The front-end I/O modules are used for channel connectivity. There are different types of front-end I/O modules that allow connectivity to various interfaces. These include SAN, FICON, SRDF™, and embedded NAS (eNAS). For more information, see the [Channel front-end redundancy](#) section of this document.

5.2.7 Back-end I/O module

The back-end I/O modules are used to connect the director boards to the back-end of the system, allowing I/O to the system's drives. For more information, see the [SAS back-end redundancy](#) section of this document.

5.2.8 InfiniBand (IB) module

The InfiniBand (IB) modules provide connectivity to the matrix interface board enclosure (MIBE), as part of the Dynamic Virtual Matrix. For more information, see the [Dynamic Virtual Matrix](#) section of this document.

5.2.9 Shared physical memory

Global memory is accessible by any director within the array.

- If an array has a single engine, physical memory pairs are internal to the engine
- If an array has multiple engines, physical memory is paired across engines

Dual-write technology is maintained by the array. In the event of a director or memory failure, the data continues to be available from the redundant copy.

5.2.10 Global memory technology overview

Global memory is a crucial component in the architecture. All read and write operations are transferred to or from global memory. Transfers between the host processor and channel directors can be processed at much greater speeds than transfers involved with physical drives. Complex statistical prefetch algorithms can adjust to proximate conditions on the array. Intelligent algorithms adjust to the workload by constantly monitoring, evaluating and optimizing cache decisions.

Dual-write technology is maintained by the array. Front-end writes are acknowledged when the data is written to mirrored locations in the cache. In the event of a director or memory failure, the data continues to be available from the redundant copy. If an array has a single engine, physical memory mirrored pairs are internal to the engine. Physical memory is paired across engines in multi-engine configurations.

VMAX3 and VMAX All Flash arrays may be configured with up to 2TB of mirrored memory per engine and up to 16TB mirrored per array.

5.2.11 Physical port numbering

Physical port numbering within the engine is used to determine how cables are connected to the front-end, back-end, and fabric I/O modules. The physical port numbering on each director board follows the same rules. The following figure illustrates physical port numbering of VMAX3 Engine and VMAX All Flash 450F and 850F V-Brick.

- I/O module slots are numbered 0-10 from left to right
- I/O modules that have 4 ports are numbered 0-3 from bottom to top
- I/O modules that have 2 ports are numbered 0-1 from bottom to top

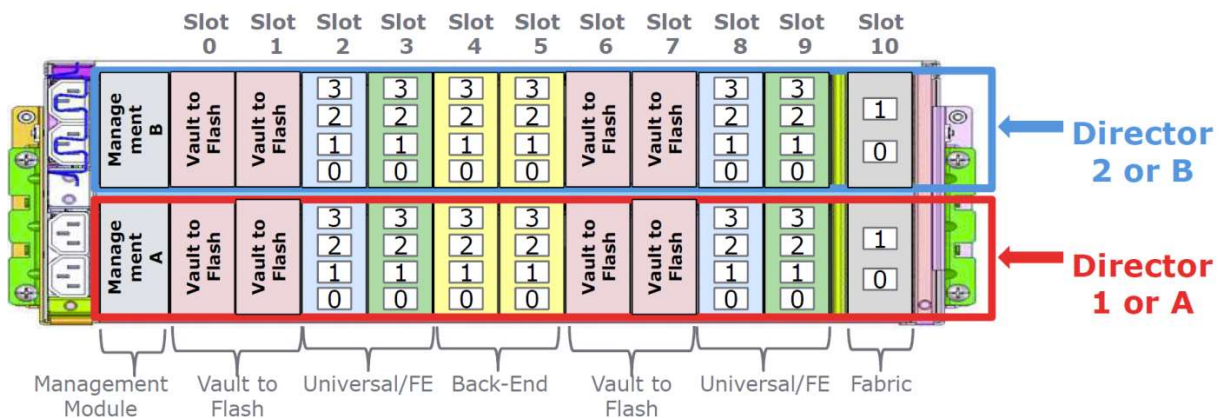


Figure 9 Physical port numbering of VMAX3 Engine and VMAX All Flash 450F and 850F V-Brick

5.2.12 Logical port numbering

Logical port numbering within the engine is used to determine what is connected to the front-end, back-end, and fabric ports. The logical port numbering on each director board follows the same rules. The following figure illustrates logical port numbering of VMAX3 Engine and VMAX All Flash 450F and 850F V-Brick.

- Supports 32 logical ports (ports 0-31)
- Logical ports are numbered from left to right, bottom to top
- Ports 0-3 and 20-23 are internally dedicated to the vault I/O modules

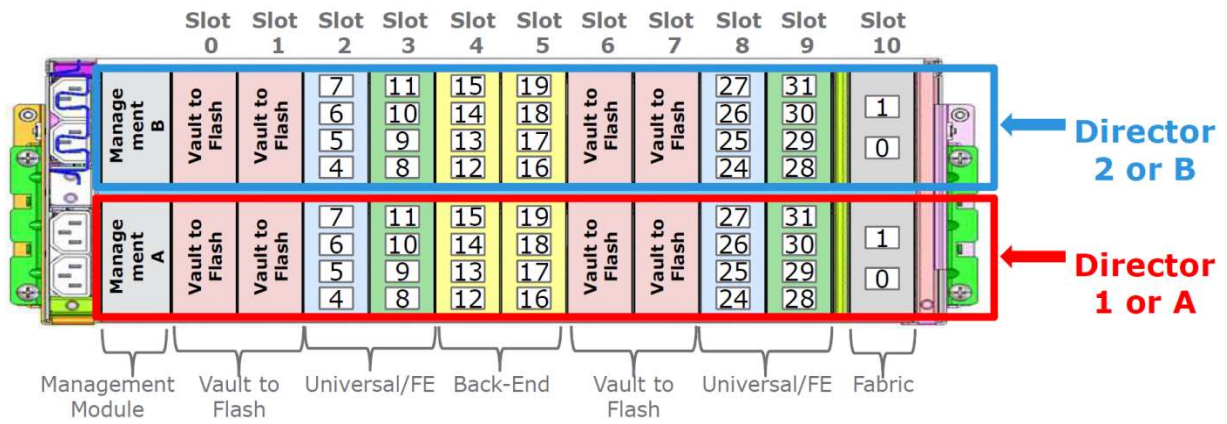


Figure 10 Logical port numbering of VMAX3 Engine and VMAX All Flash 450F and 850F V-Brick

5.3 Channel front-end redundancy

Channel redundancy is provided by configuring multiple connections from the host servers (direct connect) or Fibre Channel switch (SAN connect) to the system. SAN connectivity allows each front-end port on the array to support multiple host attachments, which increases path redundancy and enables storage consolidation across many host platforms. A minimum of two connections per server or SAN to different directors is necessary to provide basic redundancy.

The highest level of protection is delivered by configuring at least 2 or 4 front-end paths in the port group for masking and zones to the host. Single initiator zoning is recommended. Multiple ports from each server should be connected to redundant fabrics.

On the array, distributing the host paths and fabric connections across the physical array components increases fabric redundancy and spreads the load for performance. The relevant array components include directors, front-end I/O modules, and ports.

The following two examples show the options for selecting the array ports for each fabric path. The fabrics can either share I/O modules or be connected to separate I/O modules. If both redundant fabrics connect to the same I/O module, one fabric should be connected to the high ports and the other fabric should be connected to the low ports.

Both are examples of a pair of servers connected through redundant fabrics to a single-engine array with a pair of front-end I/O modules on each director. A multi-engine system has more directors and modules to spread across for redundancy. Each fabric has multiple connections to each director in the array. If a server loses connection to one fabric, it will have access to the array through the remaining fabric. Redundancy can be increased further if additional paths are configured from the servers to the fabrics and/or from the fabrics to the array.

In the first example, each fabric connects to each front-end I/O module. If one of the front-end I/O modules were to fail, each server will have an equal number of paths remaining to each director and to each of the remaining I/O modules. In the event of a director failure, host connectivity remains available through the other director's front-end I/O modules. Fabric A connects to the two high ports on each I/O modules, and Fabric B is connected to the two low ports.

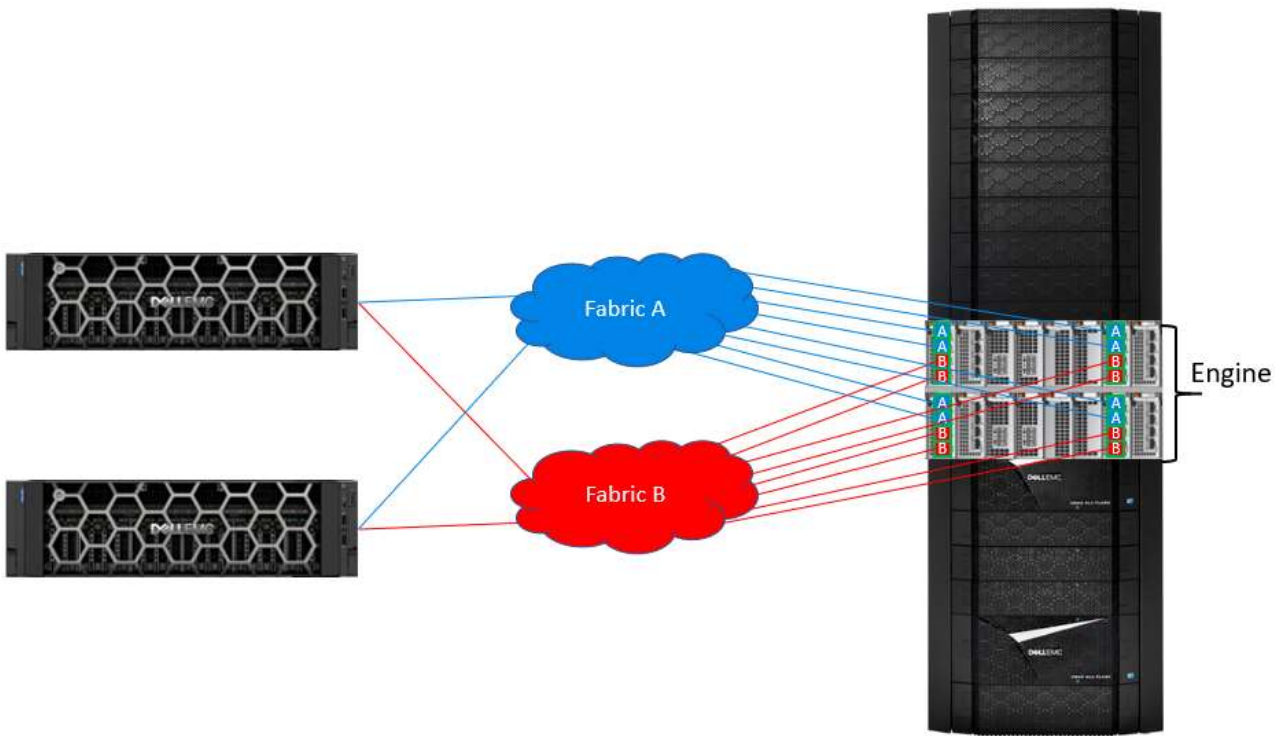


Figure 11 Each fabric connected to each I/O module

In the following example, each front-end I/O module is dedicated to one fabric. A front-end I/O module failure will only affect one fabric. And like the previous example, a director failure will affect 4 paths from each fabric.

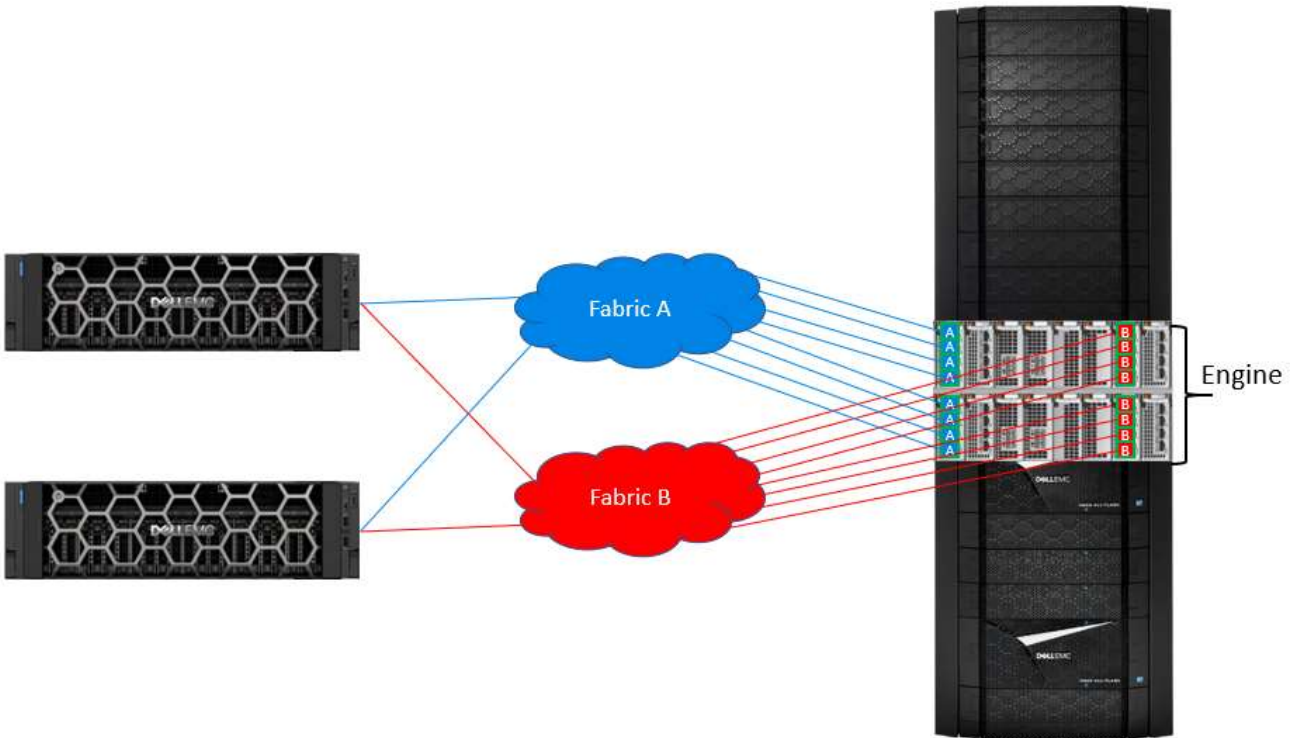


Figure 12 Each I/O module dedicated to one fabric

5.3.1 Dell EMC PowerPath Intelligent Multipathing Software

Dell EMC PowerPath™ is a family of software products that ensures consistent application availability and performance across I/O paths on physical and virtual platforms.

It provides automated path management and tools that enable you to satisfy aggressive service-level agreements without investing in additional infrastructure. PowerPath includes PowerPath Migration Enabler for non-disruptive data migrations and PowerPath Viewer for monitoring and troubleshooting I/O paths.

Dell EMC PowerPath/VE is compatible with VMware® vSphere® and Microsoft® Hyper-V-based virtual environments. It can be used together with Dell EMC PowerPath to perform the following functions in both physical and virtual environments:

- **Standardize Path Management:** Optimize I/O paths in physical and virtual environments (PowerPath/VE) as well as cloud deployments.
- **Optimize Load Balancing:** Adjust I/O paths to dynamically rebalance your application environment for peak performance.
- **Increase Performance:** Leverage your investment in physical and virtual environment by increasing headroom and scalability.
- **Automate Failover/Recovery:** Define failover and recovery rules that route application requests to alternative resources in the event of component failures or user errors.

For more information on PowerPath, refer to the *Dell EMC PowerPath Family Product Guide*.

5.4 SAS back-end redundancy

The system's architecture incorporates a SAS (Serial attached SCSI) back-end design to ensure high performance and full redundancy. SAS is a reliable, high end protocol that uses a connectionless tree structure with unique paths to individual devices. The paths are stored in routing tables which are built during a discovery phase and are used to route I/O to the desired end-point.

The SAS back-end subsystem provides independent redundant paths to the data stored on physical drives. This provides seamless access to information, even in the event of a component failure and/or replacement.

The directors are connected to each DAE through a pair of redundant back-end I/O modules. The back-end I/O modules connect to the DAEs at redundant LCCs. Each connection between a back-end I/O module and an LCC uses a completely independent cable assembly. Within the DAE, each drive has two ports, each of which connects to one of the redundant LCCs.

The dual-initiator feature ensures continuous availability of data in the unlikely event of a drive management hardware failure. Both directors within an engine connect to the same drives using redundant paths. If the sophisticated fencing mechanisms of PowerMaxOS detect a failure of the back-end director, the system can process reads and writes to the drives from the other director within the engine without interruption.

The following is an overview of the redundant components of the back-end subsystem.

5.4.1 Redundant back-end director connectivity

A pair of directors within the same engine is used to access each drive. One director is connected to the primary physical path to the drive, and the other director is connected to the secondary physical path to the drive. Directors are connected through a pair of independent back-end I/O modules and cabling that allow data to be moved back and forth between global memory and the drives. Each director is connected to global memory through primary and secondary paths, to eliminate possible single points of failure. For more information on single points of failure, see the [Dual-initiator feature](#) section of this document.

5.4.2 Redundant cable paths

Each back-end I/O module is connected to its associated link control card (LCC) and drive array enclosure (DAE) chain through a completely independent cable assembly. Each connection carries four paths, providing an increase in reliability and throughput.

5.4.3 Redundant drive paths

Each SAS drive has two ports which are connected via a fully independent path. Each of these ports is connected through separate directors, cables, and LCCs.

5.4.4 Point-to-point back-end

VMAX All Flash and VMAX3 systems use a SAS connectionless, point-to-point network that has an independent relationship with each drive. This relationship between the back-end controller and each drive facilitates analysis of drive health and improves serviceability.

The SAS connectionless, point-to-point network consists of SAS controllers in the back-end I/O modules and SAS expanders in the LCCs (VMAX DAE60 ICM/LCC and VMAX DAE120 LCC). These controlling points of the SAS network contain routing tables necessary to move data from point-A to point-B within the network. The routing tables are built in the discovery phase as the SAS network initializes or in an event of a topology change. For example, a topology change can consist of pull/plug of a drive, or connect/disconnect of a cable, or pull/plug of an LCC.

5.5 Drive Array Enclosure

Two different types of drive array enclosures (DAEs) are available. The VMAX DAE60 accommodates 3.5" drives and 2.5" drives in a 3.5" carrier. The VMAX DAE120 accommodates 2.5" drives. Both types can be mixed together in a single system. The DAE houses the physical drives as well as the LCCs.

5.5.1 DAE components

The DAE60 can hold a maximum of 60 drives and relies on both Inter-connect modules (ICM) and link control cards (LCCs) for providing communication to and from the physical drives to the back-end I/O modules. The ICM contains expanders that interface to the rest of the system, including cables to the back-end I/O modules, as well as expansion DAEs. The LCCs contain two expanders that each interface to the ICM and the drives. LCC A connects to a set of 60 drives and LCC B connects to the second port of the same drives. The two drive expanders in each LCC A and LCC B connecting to 30 drives each give a total of 60 drives maximum connected to the dual ports of each drive for redundancy.

The following figures show top and rear views of the DAE60.



Figure 13 VMAX DAE60 (top)



Figure 14 VMAX DAE60 (rear)

The DAE120 can hold a maximum of 120 drives. There are four drive expanders on each LCC. Each of the drive expanders connects to 30 drives. A total of four drive expanders connecting to 30 drives each gives a total of 120 drives. LCC A and LCC B connect to the dual ports on each drive for redundancy.

The following figures show top and rear views of the DAE120.



Figure 15 VMAX DAE120 (top)



Figure 16 VMAX DAE120 (rear)

5.5.2 Fault zones for DAEs

DAEs are designed to have complete fault zone protection in the event of an unforeseen, power related failure. Both types of DAEs have multiple fault zones which effectively provide four virtual drive enclosures within a single DAE. Figure 17 illustrates the fault zones within the DAE60. Figure 18 illustrates the fault zones within the DAE120.

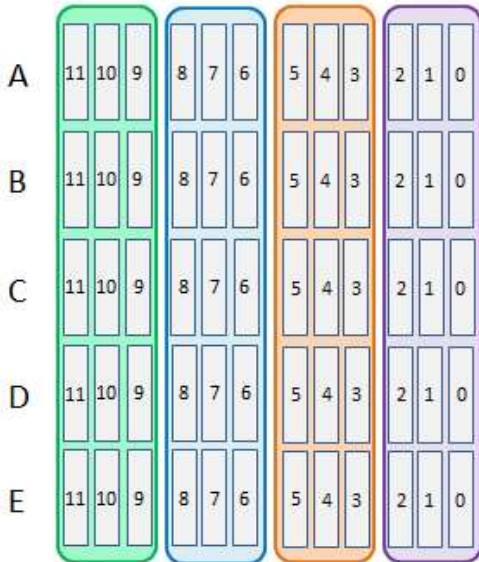


Figure 17 Fault zones for DAE60

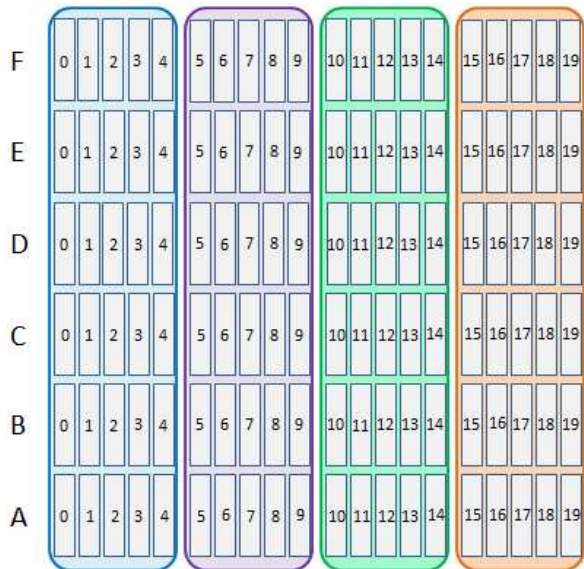


Figure 18 Fault zones for DAE120

5.5.3 Dual-initiator feature

The dual-initiator feature ensures continuous availability of data in the unlikely event of a drive management hardware failure. The dual-initiator does not provide data availability in the event of a drive failure. Drives are configured in RAID groups which provide protection for them. For more information on RAID, see the [Data Protection Methods](#) section of this document.

The dual-initiator feature works by having both directors in a single engine connect to redundant paths in the DAE. A single director can connect to both LCC A and LCC B of the same DAE. The other director in the engine, the dual-initiator, can connect to the other side of both LCC A and LCC B. This cabling scheme provides complete redundancy for availability purposes.

If the sophisticated fencing mechanisms detect a failure on the back-end director, the system can process reads and writes to the drives through a completely independent path from the dual-initiator without interruption.

The following figure illustrates an example of the redundant SAS cabling between a back-end IO modules and Link Control Cards (LCC).

Single engine with 4 DAEs

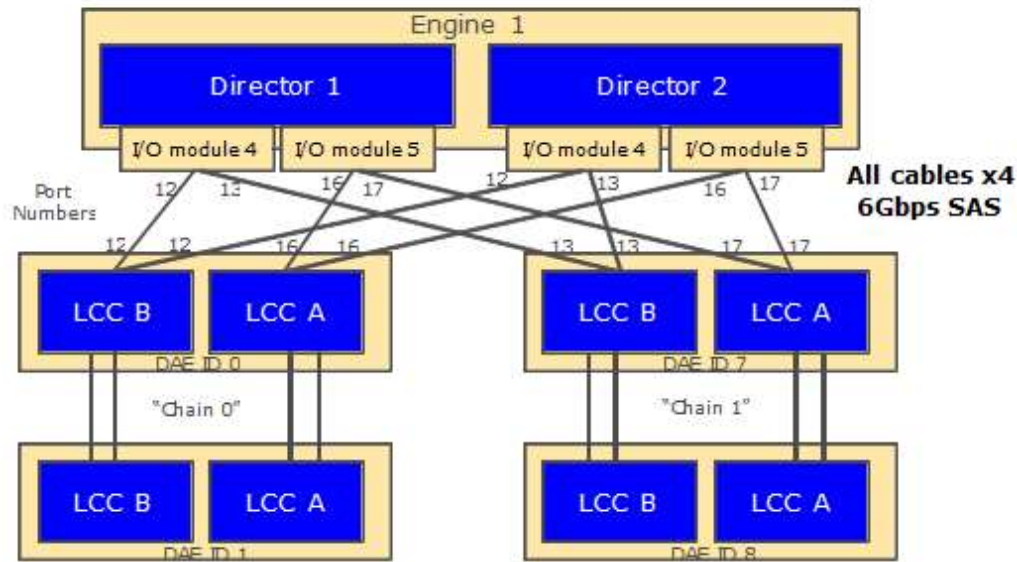


Figure 19 SAS cabling example

6 Dynamic virtual matrix

The dynamic virtual matrix uses InfiniBand (56 Gbps) technology to carry control, metadata, and user data through the system. This technology connects all engines in the system to provide a powerful form of redundancy and performance. This allows all directors to share resources and act as a single entity while communicating.

In any system that has two or more engines, there are two redundant matrix interface board enclosures (MIBE) that connect to the fabric I/O modules of each director board. The purpose of the dynamic virtual matrix is to create a communication interconnection between all directors, and therefore a single engine system does not require a dynamic virtual matrix or MIBEs.

VMAX 450F, 100K and 200K platforms contain two 12-port MIBEs with redundant, non-FRU power supplies, shown in the following figure.

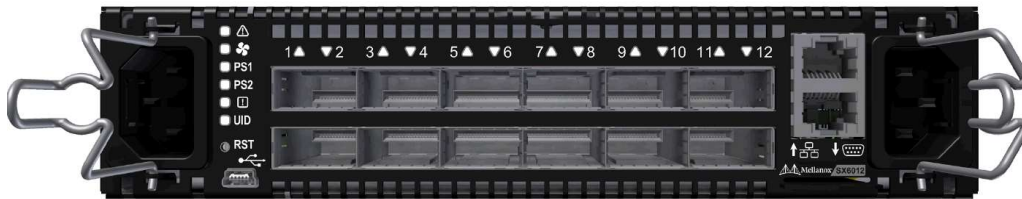


Figure 20 Front view of 12-port MIBE

VMAX 950F, 850F and 400K platform contains two 18 port MIBEs with redundant, hot pluggable power supplies, as well as fans that are replaceable, shown in the following figure.

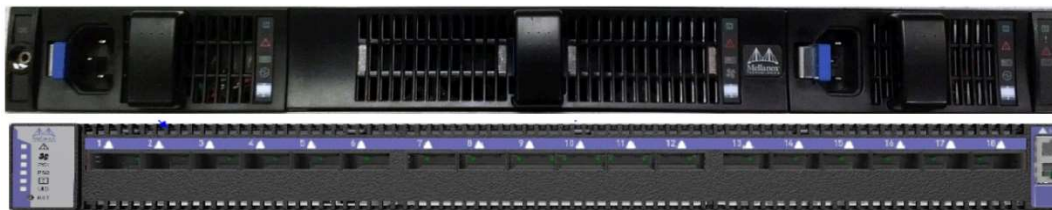


Figure 21 Front (top) and rear (bottom) view of 18-port MIBE

6.1 Internal environmental Ethernet connectivity

Environmental information is carried through two redundant Ethernet switches. Each MMCS connects to both Ethernet switches, and each management module connects to one switch (odd directors to one, even directors to the other). These provide low-level system-wide communications and environmental control for running application software, monitoring, and diagnosing the system from the MMCS. Figure 22 illustrates the internal Ethernet connectivity.

Note: These are separate from the InfiniBand MIBEs that are part of the dynamic virtual matrix.

The internal Ethernet connectivity network monitors and logs environmental events across all critical components and reports any operational problems. Critical components include director boards, global memory, power supplies, power line input modules, fans, and various on/off switches. This network's environmental control capability can monitor each component's local voltages, ensuring optimum power delivery. Temperature of director boards and memory are also continuously monitored. Failing components can be detected and replaced before a failure occurs.

The AC power main is checked for the following:

- AC failures
- Transfer to auxiliary
- DC failures
- Current sharing between DC supplies
- DC output voltage
- Specific notification of overvoltage condition
- Current from each DC supply
- Voltage drops across major connectors

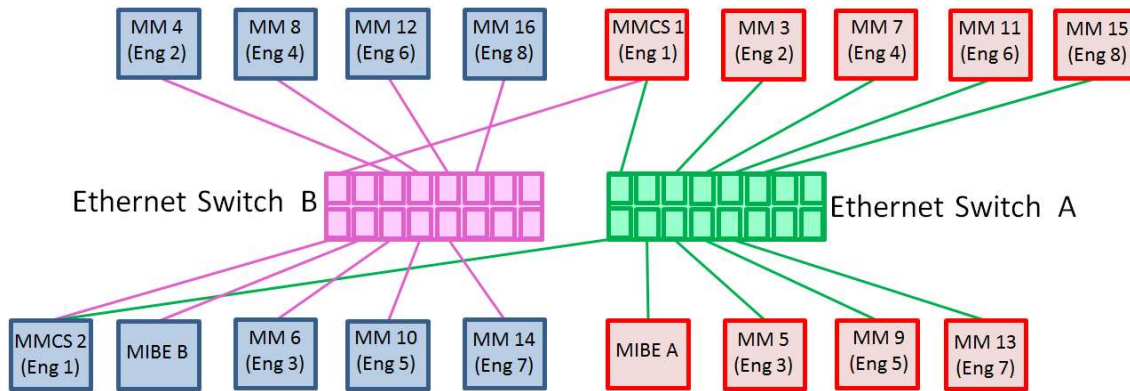


Figure 22 Internal Ethernet connectivity

7 Redundant power subsystem

A modular power subsystem features a redundant architecture that facilitates field replacement of any of its components without any interruption in processing.

The power subsystem has two power zones for redundancy. Each power zone connects to a separate dedicated or isolated AC power line. If AC power fails on one zone, the power subsystem continues to operate through the other power zone. If any single power supply module fails, the remaining power supplies continue to share the load. The fault is reported as an environmental error.

7.1 Battery backup unit modules

Lithium-ion standby power supply (Li-Ion-SPS) modules provide battery backup functionality. These are lighter and have a longer shelf life than lead acid SPS modules. System bays can include up to six Li-Ion-SPS modules. The number and location of SPS modules depends on the type of system bay and the number of engines in the array. The following rules apply to SPS configurations:

- SPS 3A and 3B provide back-up power to the first (odd) engine in that bay. They will also provide back-up power to both matrix interface board enclosures (MIBE) if they are configured in the system.
- SPS 2A and 2B power the second (even) engine in a dual engine system bay and are required on all dual engine system bays.
- SPS 1A and 1B provide back-up power to both MIBEs and the first (odd) engine. They are only required in configurations that have MIBEs, which include dual engine configurations and single engine configurations with two or more bays.

7.2 Vaulting

As cache size has grown, the time required to move all array data to a persistent state has also increased. Vaulting is designed to limit the time needed to power off the system if it needs to switch to a battery supply. Unlike previous platforms, VMAX All Flash and VMAX3 platforms do not vault to back-end drives. Data is now vaulted to dedicated I/O modules, known as flash I/O modules, saving disk space. Vaulting to flash also expedites the vaulting process, and centralizes it to the engine components, removing the need for battery backup to disks, creating an overall denser configuration than previous systems.

7.3 Vault triggers

State changes that require the system to vault are referred to as vault triggers. There are two types of vault triggers: internal availability triggers and external availability triggers.

7.3.1 Internal availability triggers

Internal availability triggers are initiated when global memory data becomes compromised due to component unavailability. Once these components become unavailable, the system triggers the Need to Vault (NTV) state, and vaulting occurs. There are three internal triggers:

1. **Vault flash availability:** The flash I/O modules are used for storage of meta data under normal conditions, as well as storing any data that is being saved during the vaulting process. When the overall available flash space in the flash I/O modules becomes the same size as *N* copy of global memory, the NTV process triggers. This is to ensure that all the data is saved before a potential further loss of vault flash space occurs.

2. **Global memory (GM) availability:** When any of the mirrored director pairs are both unhealthy either logically or environmentally, NTV triggers because of GM unavailability.
3. **Fabric availability:** When both the fabric switches are environmentally unhealthy, NTV triggers because of fabric unavailability.

7.3.2 External availability triggers

External availability triggers are initiated under circumstances when global memory data is not compromised, but it is determined that the system preservation is improved by vaulting. Vaulting in this context is used as a mechanism to stop host activity, facilitate easy recovery or proactively prevent potential data loss. There are two external triggers:

4. **Engine trigger:** When an entire engine fails, the system vaults.
5. **DAE trigger:** If the system has lost access to the whole DAE or DAEs, including dual-initiator failure, and loss of access causes configured RAID members to become non-accessible, the system vaults.

7.4 Power-down operation

When a system is powered down or transitioned to offline, or when environmental conditions trigger a vault situation, a vaulting procedure occurs. During power-down or power loss, the part of global memory that is saved first reaches a consistent image (no more writes). The directors then write the appropriate sections of global memory to the flash I/O modules, saving three copies of the logical data. The battery backup unit (BBU) modules maintain power to the system during the power-down process for up to 5 minutes.

7.5 Power-up operation

During power-up, the data is written back to global memory to restore the system. When the system is powered on the startup program does the following:

- Initializes the hardware and the environmental system
- Restores the global memory from the saved data while checking the integrity of the data. This is accomplished by taking sections from each copy of global memory that was saved during the power-down operation and combining them into a single complete copy of global memory. If there are any data integrity issues in a section of the first copy that was saved, then that section is extracted from the second copy during this process.
- Performs a cleanup, data structure integrity, and reinitialization of needed global memory data structures

At the end of the startup program, the system resumes normal operation when the BBUs are recharged enough to support another vault. If any condition is not safe, the system does not resume operation and calls Customer Support for diagnosis and repair. In this state, Customer Support can communicate with the system and find out the reason for not resuming normal operation.

8 Data protection methods

Although the system has standard features that provide a higher level of data availability than conventional DASDs, the following data protection options ensure an even greater level of data recoverability and availability:

- RAID 1 (Mirroring)
- RAID 5
- RAID 6
- Local RAID
- Thin provisioning
- Drive sparing
- Local replication with TimeFinder
- Remote replication with Symmetrix Remote Data Facility (SRDF)
- Data at Rest Encryption (D@RE)

These data protection options (excluding disk sparing) are configurable on individual physical volumes, so that different levels of protection can be applied to different datasets within the same system.

8.1 RAID 1 (Mirroring)

RAID 1 configurations have higher performance in most applications because each drive has a copy of the data. Accordingly, it is possible for the system to be satisfying two I/O requests simultaneously by sending or receiving one from either copy. Mirrored volumes also have lower response times due to the Dynamic Mirror Service Policy (DMSP), which automatically determines the optimal disk to read to achieve maximum performance. Mirrored volumes also have lower response times since the system can access the data on either of the mirrored drives.

Mirrored configurations also provide higher performance if there is a disk failure since another complete copy of the data is immediately available. Furthermore, since only two disks are used for the mirror, the chance of failure on multiple drives containing the same data is reduced.

8.2 RAID 5

This is an implementation of the industry-standard RAID 5 data protection technique with rotating parity across all members of the RAID 5 set. In the event of a physical drive failure, the missing data is rebuilt by reading the remaining drives in the RAID group and performing XOR calculations.

RAID 5 provides cost-effective data protection against drive failures. While the most demanding environments continue to opt for mirrored storage for maximum performance, RAID 5 configurations offer an extremely attractive alternative for information storage where price is more important than performance.

RAID 5 is available in two configurations:

- RAID 5 (3+1) – Data and parity are striped across 4 drives (3 data, 1 parity)
- RAID 5 (7+1) – Data and parity are striped across 8 drives (7 data, 1 parity)

8.3 RAID 6

Protection schemes such as RAID 1 and RAID 5 can protect a system from a single physical drive failure within a mirrored pair or RAID group. RAID 6 supports the ability to rebuild data if two drives fail within a RAID group.

Dell EMC's implementation of RAID 6 calculates two types of parity. This is important during events when two drives within the same RAID group fail, as it still allows the data in this scenario to be reconstructed. Horizontal parity is identical to RAID 5 parity, which is calculated from the data across all disks in the RAID group. Diagonal parity is calculated on a diagonal subset of data members. For applications without demanding performance needs, RAID 6 provides the highest data availability.

8.4 Local RAID

Local RAID is a new concept that puts all members of a RAID group behind a single engine. This improves back-end performance and drive rebuild time, while still supporting the dual-initiator failover/failback model. Local RAID is implemented in all VMAX All Flash and VMAX3 environments.

The following member distribution rules apply to Local RAID configurations:

- No two members of the same RAID group can be placed on the same disk
- Disks reside in disk groups
- Each disk group supports only one RAID type
- One disk group per disk type is created by default
- Each disk group is provisioned with its own spare disks
- Aside from improved performance, there are other benefits associated with Local RAID, including the following:
 - Elimination of cross-bay cabling for direct/daisy chain DAE cabling
 - Dispersion at the engine/bay level
 - Configuration of new systems or upgrades with any combination of contiguous or dispersed engines/bays
 - Can position engine/bay on either side of system bay #1
 - Most flexible floor planning capability in the industry

8.5 Thin provisioning

Thin Provisioning enables the ability to increase capacity utilization by enabling more storage to be presented to a host than is physically consumed, and by allocating storage only as needed from a shared virtual pool. Thin Provisioning also simplifies storage management by making data layout easier through automated wide striping, and by reducing the steps required to accommodate growth.

Thin Provisioning uses a type of host-accessible device called a virtually provisioned device, also known as a thin device, which does not need to have physical storage completely allocated at the time the devices are created and presented to a host. The physical storage that is used to supply drive space for a virtually provisioned device comes from a shared storage pool, also known as a storage resource pool (SRP). The SRP is comprised of one or more data pools containing internal devices called data devices. These data devices are dedicated to the purpose of providing the actual physical storage used by virtually provisioned devices.

When a write is performed to a portion of the virtually provisioned device, the array allocates a minimum allotment of physical storage from the pool and maps that storage to a region of the virtually provisioned device, including the area targeted by the write.

The storage allocation operations are performed in small units of storage called virtually provisioned device extents. Extents may also be called chunks. The virtually provisioned device extent size is 1 track (128 KB).

When a read is performed on a virtually provisioned device, the data being read is retrieved from the appropriate data device in the SRP to which the virtually provisioned device is bound. Reads directed to an area of a virtually provisioned device that has not been mapped do not trigger allocation operations. Reading an unmapped block returns null data. When more storage is required to service existing or future virtually provisioned devices, data devices can be added to existing virtually provisioned data pools within the SRP.

For more information on Thin Provisioning on VMAX3, refer to *Dell EMC VMAX3 Service Level Provisioning with Fully Automated Storage Tiering (FAST™)* Technical Notes.

8.6 Drive sparing

Drive health is monitored proactively for any indication that they may be trending toward failure. Drive-dependent codes detect and report indications of failing health. For example, conditions such as errors on blocks of NAND media, errors in DRAM buffer media, controller check errors.

When a failing drive is detected, the data on the faulty drive is copied directly to a spare drive. The failing drive is set as read-only while data is copied and written to the spare. The failing drive is made not ready after the spare is synchronized. If the faulty drive stops responding to valid commands prior to spare synchronization, the drive is made not ready and the data is rebuilt onto the spare drive through the remaining RAID members.

The location of the spare drive determines the process that will take place after the faulty drive is replaced. Data will either be copied from the spare to the new drive and the spare drive will become an available spare again after the new drive is synchronized, or the initial spare drive may become the new RAID member and the drive that was replaced will become an available spare.

Direct Sparing automatically replaces a failing drive with a spare drive. Direct Sparing is supported with all protection types, including RAID 6 (14+2).

Two traditional sparing factors, vault drives and specific back-end director within the engine, do not apply to VMAX All Flash and VMAX3 systems. These systems vault to flash I/O modules on the engine rather than vault drives. If needed, spare drives can be dynamically relocated across back-end directors within the same engine by the sparing and replacement operations.

The major factor with Direct Sparing is the power zone within the disk enclosure where the spare drive, failing drive, and other RAID members are located. RAID 1 and RAID 5 are only allowed to have one member per power zone, and RAID 6 can have up to two members per power zone.

Spare drives are designated into three categories: Preferred, Regular, and Non-Preferred. The disk sparing script selects the best available drive by using the rules in Table 4.

Table 4 Drive sparing categories

Spare category	Description	Spare placement behavior		Script handling behavior	
		Same DAE power zone distribution	Keeps high availability configuration rules	Failing drive gets restored to its original positions	Spare Drive Replenishment Process
Preferred Spare (Level 1)	<p>The spare is configured in either the same DAE power zone as the failing drive, or in a different DAE power zone where another RAID member does not already exist.</p> <p>Use of this spare results in a valid RAID configuration.</p> <p>This spare does not create a configuration with multiple RAID members in the same DAE power zone.</p>	Yes	Yes	No	Yes
Regular Spare (Level 2)	<p>The spare is not configured in the same DAE power zone as the failing drive.</p> <p>Use of this spare results in a valid RAID configuration.</p> <p>This creates a legal configuration with multiple RAID members in the same DAE power zone.</p>	No	Yes	Yes	Yes
Non-Preferred Spare (Level 3)	<p>The spare is not configured in the same DAE power zone as the failing drive.</p> <p>Use of this spare results in a configuration that breaks the rules for RAID member/DAE power zone distribution.</p>	No	No	Yes	No

8.6.1 Spare Drive Replenishment Process

The Spare Drive Replenishment Process calls out spare drives for replacement when any drive in the system no longer has a spare available following a direct sparing operation. This allows customers to schedule drive replacement maintenance with Customer Support less frequently, as multiple drives are replaced simultaneously.

For more information on drive sparing, refer to *Drive Sparing in Dell EMC Symmetrix™ VMAX Family Systems White Paper*.

8.6.2 VMAX All Flash and VMAX3 spare drive count

The amount and types of spares required are calculated as follows:

- One spare drive is required for every 50 drives per disk group per engine.
 - This calculation applies to both HDD and EFD drives.
 - The required spares are calculated purely per disk group.
 - There is no minimum spare count per system.

8.6.3 Solutions Enabler commands

Solutions Enabler provides tools to view information related to spare drives.

The `symcfg list -v` output reports total values for Configured Actual Disks, Configured Spare Disks and Available Spare Disks in the system.

The `Number of Configured Actual Disks` field reports only non-spare configured disks, and `Number of Configured Spare Disks` field reports only configured spare disks.

```

Symmetrix ID: 000197800XYZ (Local)
Time Zone   : Eastern Standard Time

Product Model      : VMAX_250F
Symmetrix ID      : 000197800XYZ

Microcode Version (Number) : 5978 (175A0000)

-----< TRUNCATED >-----

Number of Configured Actual Disks : 64
Number of Configured Spare Disks  : 2
Number of Available Spare Disks   : 2

```

Figure 23 `symcfg -sid <sid> list -v`

The `symdisk list -dskgrp_summary -by_engine` reports spare coverage information per Disk Group per Engine. The `-detail` and `-v` options will provide additional information.

The Total and Available spare disk counts for each Disk Group include both spare disks that are in the same Disk Group in the same Engine, as well as shared spare disks in another Disk Group in the same Engine that provide acceptable spare coverage. These shared spares are also included in the total disk count for each Disk Group in each Engine. Therefore, the cumulative values of all Disk Groups in all Engines in this output should not be expected to match the values reported by the `symcfg list -v` command that were described in the previous example.

Total Disk Spare Coverage percentage for a Disk Group is the spare capacity in comparison to usable capacity shown in the output.

Disk			Hyper		Usable Capacity			Spare Coverage			
Grp	Eng	Cnt	Flgs LT	Speed (RPM)	Size (MB)	Disk	Total (%)	Total (MB)	Total Disk (%)	Avail Disk (%)	
1	1	9	IE	0	29063	8	89	14880255	1	12	1 100
2	1	25	IE	0	29063	24	96	44640765	1	4	1 100
2	2	33	IE	0	29063	32	97	59521020	1	3	1 100
Total						64	97	119042040			

Legend:
 Disk (L)ocation:
 I = Internal, X = External, - = N/A
 (T)echnology:
 S = SATA, F = Fibre Channel, E = Enterprise Flash Drive, - = N/A

Figure 24 `symdisk -sid <sid> list -dskgrp_summary -by_engine`

Spare Coverage as reported by the `symdisk list -v` and `symdisk show` commands indicates whether the disk currently has at least one available spare; that is, a spare disk that is not in a failed state or already invoked to another disk.

Symmetrix ID	: 000197800XYZ
Disks Selected	: 66
Director	: DF-1C
Interface	: C
Target ID	: 0
Spindle ID	: 0
-----< TRUNCATED >-----	
Spare Disk	: N/A
Spare Coverage	: True

Figure 25 `symdisk -sid <sid> list -v`

8.7 Local replication using TimeFinder

TimeFinder™ software delivers point-in-time copies of volumes that can be used for backups, decision support, data warehouse refreshes, or any other process that requires parallel access to production data.

Previous VMAX families offer several different TimeFinder offerings, each with their own characteristics and ideal use cases. These offerings also have several similarities, the main one being that they each require a target volume to retain snapshot or clone data.

TimeFinder in HYPERMAX OS 5977 introduced TimeFinder SnapVX which combines the best aspects of the previous TimeFinder offerings, adds some new ease-of-use features, and increases scalability.

SnapVX provides very low impact snapshots and clones for data volumes. SnapVX supports up to 256 snapshots per source volume, which are tracked as versions with less overhead and simple relationship tracking. Users can assign names to identify their snapshots, and they have the option of setting automatic expiration dates on each snapshot.

SnapVX provides the ability to manage consistent point-in-time copies for storage groups with a single operation. Up to 1024 target volumes can be linked per source volume, providing read/write access as pointers or full copies.

TimeFinder also provides compatibility modes for users who rely on their TimeFinder Mirror, Snap, Clone, or VP Snap command scripts. This will allow users to use their existing scripts while learning how to take advantage of the new features of SnapVX.

For more information on TimeFinder SnapVX, refer to *Dell EMC TimeFinder SnapVX Local Replication Technical Notes*.

8.7.1 Snapshot policies

Snapshot Policies, introduced in the PowerMaxOS Q3 2020 release, provide automated scheduling of SnapVX snapshots using a highly available and flexible policy engine that runs internally on the storage array. Snapshot policies can be managed through Dell EMC Unisphere™ for PowerMax, REST API, and Solutions Enabler.

Snapshot policies can be customized with rules that specify when to take snapshots, how many snapshots to take, and how long to keep each snapshot. Compliance requirements can also be specified to send alerts if the rules of a policy are not being met. Applications can be protected by multiple policies with differing schedules and retention parameters according to the requirements of the business. Each policy can protect many applications, even protecting a mix of open systems and mainframe applications.

Snapshot policies provide reliable protection for applications in an automated fashion that requires little to no maintenance by the business. Administrators can manually take snapshots of applications that are protected by snapshot policies to satisfy on-demand requirements. Policy parameters are shown in Figure 26.

The screenshot displays the 'View & Modify Policy' window for a policy named 'DailyDefault'. The window is divided into several sections:

- Properties:**
 - Name: DailyDefault
 - Type: Secure Snapshots
 - Last Execution Time: N/A
 - Description: Every day at 0:00
- Recovery Point Objective (RPO):**
 - Create a snapshot: Daily at 00:00
 - Keep 14 Snapshots Total 14 Snapshots (Max 1024)
- Compliance:**
 - Show as if fewer than N/A snapshots are created
 - if fewer than 10 (71%) snapshots are created

At the bottom right, there are 'CANCEL' and 'MODIFY' buttons. A help icon (?) is located at the bottom left.

Figure 26 Snapshot policy create / modify window

8.8 Remote replication using SRDF

Symmetrix Remote Data Facility (SRDF) solutions provide industry-leading disaster recovery and data mobility solutions. SRDF replicates data between 2, 3 or 4 arrays located in the same room, on the same campus, or thousands of kilometers apart.

- SRDF synchronous (SRDF/S) maintains a real-time copy at arrays located within 200 kilometers. Writes from the production host are acknowledged from the local array when they are written to cache at the remote array.
- SRDF asynchronous (SRDF/A) maintains a dependent-write consistent copy at arrays located at unlimited distances. Writes from the production host are acknowledge immediately by the local array, thus replication has no impact on host performance. Data at the remote array is typically only seconds behind the primary site.

SRDF disaster recovery solutions use “active, remote” mirroring and dependent-write logic to create consistent copies of data. Dependent-write consistency ensures transactional consistency when the applications are restarted at the remote location. You can tailor SRDF to meet various Recovery Point Objectives/Recovery Time Objectives.

Using only SRDF, you can create complete solutions to:

- Create real-time (SRDF/S) or dependent-write-consistent (SRDF/A) copies at 1, 2, or 3 remote arrays.
- Move data quickly over extended distances.
- Provide 3-site disaster recovery with zero data loss recovery, business continuity protection and disaster-restart.

You can integrate SRDF with other Dell EMC products to create complete solutions to:

- Restart operations after a disaster with zero data loss and business continuity protection.
- Restart operations in cluster environments. For example, Microsoft Cluster Server with Microsoft Failover Clusters.
- Monitor and automate restart operations on an alternate local or remote server.
- Automate restart operations in VMware environments.

8.8.1 SRDF/Cascade and SRDF/Star support

SRDF/Cascade configurations use 3-site remote replication with SRDF/A mirroring between sites B and C, delivering additional disaster restart flexibility. The following figure shows an example of an SRDF/Cascade solution.

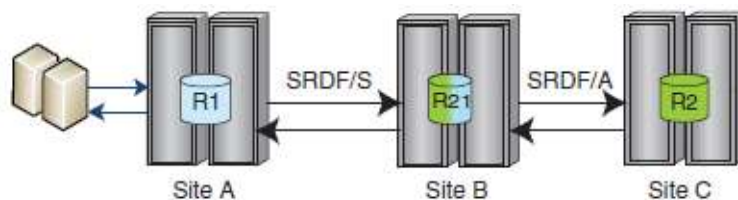


Figure 27 SRDF/Cascade

RDF/Star is commonly used to deliver the highest resiliency in disaster recovery. SRDF/Star is configured with 3-sites enabling resumption of SRDF/A with no data loss between the two remaining sites, providing

continuous remote data mirroring and preserving disaster-restart capabilities. The following figure shows examples of Cascaded and Concurrent SRDF/Star solutions.

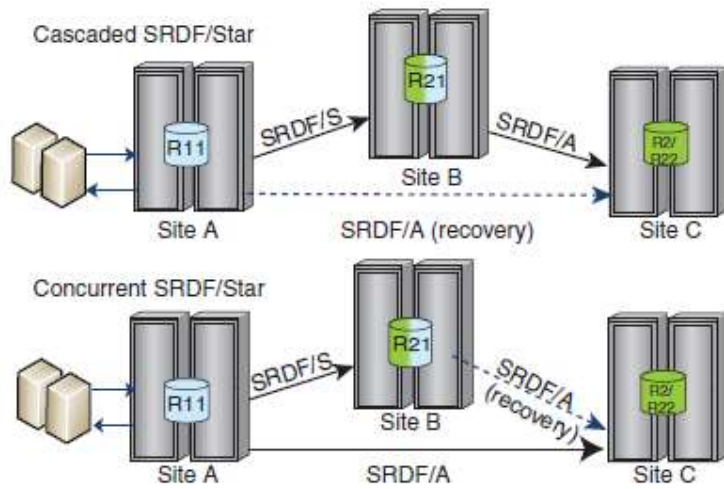


Figure 28 SRDF/Star

8.8.2 SRDF/Metro support

SRDF/Metro™ significantly changes the traditional behavior of SRDF™ Synchronous mode with respect to the remote (R2) device availability to better support host applications in high-availability environments. With SRDF/Metro, the SRDF R2 device is Read/Write accessible to the host and takes on the federated personality of the primary R1 device (geometry, device WWN, etc.). By providing this federated personality on the R2 device, both R1 and R2 devices may then appear as a single virtual device across the two SRDF paired arrays for host presentation. With both the R1 and R2 devices being accessible, the host or hosts (in the case of a cluster) can read and write to both R1 and R2 devices with SRDF/Metro insuring that each copy remains current, consistent, and addressing any write conflicts which may occur between the paired SRDF devices. The following figure shows example SRDF/Metro solutions.



Figure 29 SRDF/Metro

On the left is an SRDF/Metro configuration with a standalone host which has visibility to both arrays (R1 and R2 devices) using multi-pathing software such as PowerPath, to enable parallel reads and writes to each

array. This is enabled by federating the personality of the R1 device to ensure that the paired R2 device appears, through additional paths to host, as a single virtualized device.

On the right is a clustered host environment where each cluster node has dedicated access to an individual array. In either case, writes to the R1 or R2 devices are synchronously copied to its SRDF paired device. Should a conflict occur between writes to paired SRDF/Metro devices, the conflicts are internally resolved to ensure a consistent image between paired SRDF devices is maintained to the individual host or host cluster.

SRDF/Metro may be selected and managed through Solutions Enabler SYMCLI or Unisphere for VMAX 8.1 or greater client software. SRDF/Metro requires a separate SRDF/Metro license to be installed on both arrays to be managed.

8.9 Application I/O serviced by remote array

In the event of a redundant RAID failure on the R1 array, applications local to the R1 site will continue to access data through the remote R2 array. There may be some overhead on response time depending on distance, but the data remains available to the applications even though it cannot be read locally.

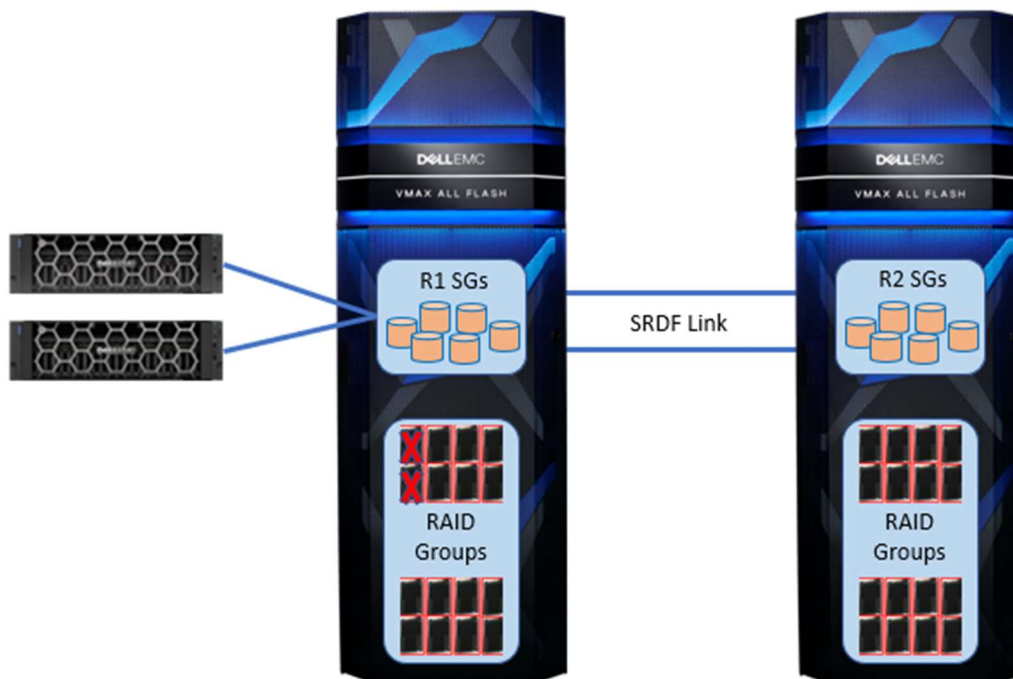


Figure 30 Application I/O serviced by remote array

This core functionality has been an integral piece of SRDF for many years, originally providing the ability for the entire contents of a thick volume to be accessed from a remote array and rebuild data after a drive replacement even before local RAID schemes were implemented. In VMAX AF arrays, which employ virtual provisioning, remote access is not required for an entire TDEV and is instead done on a per-track basis.

The local RAID group will be rebuilt from the remote array after replacement or recovery of the affected drives. An invoked spare drive can also participate in the rebuild in place of a RAID member. If the RAID members had failed at the same time, both can be recovered and rebuilt concurrently. If the second drive failed while the first drive was rebuilding after being replaced, the remote data is used to finish rebuilding the

data to the first drive, and the second drive will be rebuilt locally. Applications can remain online during the rebuild.

This functionality is available with SRDF/Metro, SRDF/S, and SRDF/A (when not in a spillover state). Being able to access data remotely, and rebuild RAID data from a remote array, enable SRDF configurations to take advantage of the efficiency and performance benefits of RAID 5 without incurring the performance penalties of RAID 6 protection. The Mean Time Between Part Replacement (MTBPR) of modern flash drives, and most replacements being proactive, make a dual failure very unlikely.

8.9.1 Dell EMC VMAX All Flash Lab Validation Report, IDC 2017

IDC tested the continuous operations without data loss from a total local RAID group failure with SRDF/Synchronous during overall validation of Dell EMC VMAX All Flash arrays. The following is from the previously published IDC Lab Validation Report *Dell EMC VMAX All Flash: Essential Capabilities for Large Enterprise Mixed Workload Consolidation*:

IDC Opinion: The active volume worked flawlessly during these fault injection tests, both upon initial failure and after re-establishing the failed resource... I/O was not interrupted, no data was lost, and in this case (with a relatively light workload) there was no long-term impact to overall system performance... Throughout the entire fault injection test, the array(s) continued to meet its specified Diamond Service Level.

Exploration of I/O Impact on Dual Local Drive Failure in a RAID 5 Configuration:

- An SRDF/Synchronous configuration was set up using a VMAX 950F and a VMAX 250F where the data was mirrored to both locations; SSDs in each array were separately protected by RAID 5
- An artificial workload was generated to flow continuously against a volume that was mirrored to the two arrays with SRDF/Synchronous; this workload can be characterized as “light” as it was under 20% of each of the array’s rated performance capabilities
- Two SSDs in a single RAID 5 Raid Group in the VMAX 950F were simultaneously failed to create a dual drive failure scenario; the “failure” was validated running the Solutions Enabler command **syndisk –sid <sid> list –failed**. In addition, the failure can generate an Alert in Unisphere and will generate a dial home
- The impact on I/O was observed by continually watching the throughput and latency against the SRDF mirrored volume using IOmeter; upon failure, I/O was not disrupted as I/O to the failed devices continued to be handled by the remote mirror (through reads across the SRDF Link) with no impact to storage latency or throughput during this failure, and throughout the test the volume continued to meet its Diamond Service level
- This is significant because this type of dual failure in a stretched cluster configuration could lead to an outage whereas here the system just continues to access the data using the mirrored volume on the target array
- The “failed” SSDs were turned back on, and with no impact to storage latency or throughput the system performed a background resync (quite short since there was less than 5MB of writes to the mirrored volume during the outage) and returned the volume to a fully synchronized state across the VMAX 950F and the VMAX 250F

8.9.2 PowerMaxOS Q3 2020 release SRDF updates

SRDF/Metro Smart DR: SRDF/Metro and SRDF/A integration provides high-availability disaster recovery solution for SRDF/Metro active/active environments.

Smart DR provides SRDF/Metro with a single asynchronous target R22 volume which may be populated from either the R1 or R2 volume of an SRDF/Metro paired solution. Adding the capability for R1 and R2 to share a single asynchronous R22 volume simplifies setup, maintenance capabilities, system requirements, and reduces the amount of disk space required for a single target system.

Smart DR provides the ability to fail over or fail back to the DR site while retaining the metro environment. Smart DR can be implemented non-disruptively onto existing SRDF/Metro environments.

The following figure depicts SRDF/Metro Smart DR:



Figure 31 SRDF/Metro Smart DR

8.10 Data at Rest Encryption

Data at Rest Encryption (D@RE) protects data confidentiality by adding back-end encryption to the entire array. D@RE provides hardware-based, on-array, back-end encryption. Back-end encryption protects information from unauthorized access when drives are removed from the system.

D@RE provides encryption on the back-end using SAS I/O modules that incorporate XTS-AES 256-bit data-at-rest encryption. These I/O modules encrypt and decrypt data as it is being written to or read from a drive. All configured drives are encrypted, including data drives, spares, and drives with no provisioned volumes.

D@RE incorporates RSA™ Embedded Key Manager for key management. With D@RE, keys are self-managed, and there is no need to replicate keys across volume snapshots or remote sites. RSA Embedded Key Manager provides a separate, unique DEK for each drive in the array, including spare drives.

By securing data on enterprise storage, D@RE ensures that the potential exposure of sensitive data on discarded, misplaced, or stolen media is reduced or eliminated. If the key used to encrypt the data is secured, encrypted data cannot be read. In addition to protecting against threats related to physical removal of media,

media can readily be repurposed by destroying the encryption key used for securing the data previously stored on that media.

D@RE is compatible with all features, allows for encryption of any supported local drive types or volume emulations, and delivers powerful encryption without performance degradation or disruption to existing applications or infrastructure.

D@RE can also be deployed with external key managers using KMIP (Key Management Interoperability Protocol) which will allow for a separation of key management from VMAX3 and VMAX All Flash arrays. KMIP is an industry standard that defines message formats for the manipulation of cryptographic keys on a key management server. External key manager provides support for consolidated key management and allows integration between VMAX3 and VMAX All Flash with an already existing key management infrastructure.

For more information on D@RE, refer to the *Dell EMC VMAX3 and VMAX All Flash Data at Rest Encryption White Paper*.

9 Component-level serviceability

A modular design with a low parts count improves serviceability by allowing nondisruptive component replacements, should a failure occur. This low parts count minimizes the number of failure points.

VMAX All Flash and VMAX3 systems feature nondisruptive replacement of their major components, which can be replaced while the system is powered on, including:

- Engine components:
 - Director boards and memory modules
 - I/O Modules
 - > Fibre Channel (front-end)
 - > iSCSI
 - > FICON
 - > Embedded NAS (eNAS)
 - > SAS (back-end)
 - > Flash (Vault)
 - > SRDF Compression
 - > Inline Compression
 - Management modules/management module control stations
 - InfiniBand (IB) module
 - Power supplies
 - Fans
- Disk Array Enclosure (DAE) components:
 - Link Control Cards (LCC)
 - Power supplies
 - SAS & flash drives
 - Fans
 - System Status Card (SSC)
- Cabinet Components
 - Matrix interface board enclosures (MIBE)
 - Ethernet switches
 - Standby Power Supplies (SPS) - Lithium Ion batteries
 - Power Distribution Units (PDU)

VMAX All Flash and VMAX3 systems provide full component-level redundancy to protect against a component failure and ensure continuous and uninterrupted access to information. This nondisruptive replacement capability allows the Customer Support Engineer to install a new component, initialize it if necessary, and bring it online without:

- Disrupting access to unaffected volumes
- Powering down the unit
- Stopping the operating system
- Taking unaffected channel paths offline
- Taking devices offline (other than the affected device)

9.1 Flashing rack lights

Each VMAX All Flash and VMAX3 bay has an LED bar on the front and rear that perform a valuable function during service activities. Procedures such as installation, component replacement, and cable connectivity verification send a signal to the appropriate LED bar(s) initiating a blinking of the LED, helping to guide service personnel to the correct bay. Additional indicator LEDs inside the bay, along with detailed instructions provided by the service procedure on the MMCS, further guide service personnel to the correct component(s). This is a very valuable feature in today's data centers, especially on dispersed bay systems.

The state of the LED bar on a bay can also be changed by way of the Solutions Enabler command `symcfg set -led` and through Unisphere for VMAX under the System Hardware menu.

Each light bar is connected to both power zones for redundant power. The light bars themselves as well as their cabling can be replaced non-disruptively.

Note: VMAX 250F bays do not have LED bars

10 Non-disruptive upgrades

10.1 PowerMaxOS and HYPERMAX OS upgrades

Interim updates of PowerMaxOS can be performed remotely by the Remote Change Management (RCM) team. These updates provide enhancements to performance algorithms, error recovery and reporting techniques, diagnostics, and PowerMaxOS fixes. They also provide new features and functionality for PowerMaxOS.

During an online PowerMaxOS code load, a member of the RCM team downloads the new PowerMaxOS code to the MMCS. The new PowerMaxOS code loads into the EEPROM areas within the directors and remains idle until requested for a hot load in the control store. The system loads executable PowerMaxOS code within each director hardware resource until all directors are loaded.

Once the executable PowerMaxOS code is loaded, the new code becomes operational in 6 seconds or less through an internal processing operation that is synchronized across all directors.

The system does not require customer action during the upgrade. All directors remain online to the host processor and maintain application access. There is no component downtime, no rolling outage upgrade, no failover or failback processes involved, and switching LUN ownership or trespass is not required. The Fibre Channel port never drops, and the servers never see a logout or login (no fabric RSCN).

This upgrade process, which has been transparent to applications for many years, is continually improved upon and provide the means to perform downgrades in the same, non-disruptive manner.

10.2 eNAS upgrades

Embedded NAS (eNAS) can be added to existing arrays non-disruptively. The Dell EMC upgrade planning process will determine if eNAS can simply be added to an existing configuration or if additional hardware will also be required to provide adequate capacity, cache, and processing power.

10.3 Hardware upgrades

All upgrades to add hardware are non-disruptive, including:

- Engines
- Cache
- I/O modules
- Capacity

10.3.1 Capacity upgrades

No user action is required for the system to begin using newly added capacity. New writes are distributed across the system with some bias towards drives which have the most available capacity, which in this case will be the new drives. Background rebalance activities take place to spread the TDATs on the newly added drives as well as existing data across the compression pools. PowerMaxOS or HYPERMAX OS will bring each compression pool into a reasonable balance transparently to the host applications. Servicing host I/O takes priority over the background activities. Therefore, the total time to achieve a balance will depend on overall system activity as well as total system capacity, used capacity, and amount of newly added capacity.

11 VMAX 250F

The Q3 2016 Service Release of HYPERMAX OS introduced support for the VMAX 250F. The F package includes base software titles. VMAX 250FX ships with the FX package, adding additional software titles to the F bundle. VMAX 250F scales from 1 to 2 V-Bricks, providing a system total of up to 100 drives, 1PBe capacity, 4TBr cache, and 64 front-end ports, all within a single floor tile. A single, fully-configured VMAX 250F will only consume the bottom half of the cabinet. The top half of the cabinet can be used for either a separate VMAX 250F or for non-Dell EMC hardware.

11.1 VMAX 250F V-Brick

Each V-Brick in a VMAX 250F contains two director boards. Each director board has up to four front-end slots with up to four front-end ports that can be used for Fibre Channel, SRDF, iSCSI, or eNAS connectivity.

The following figure illustrates the slots and associated physical ports of a VMAX 250F V-Brick.

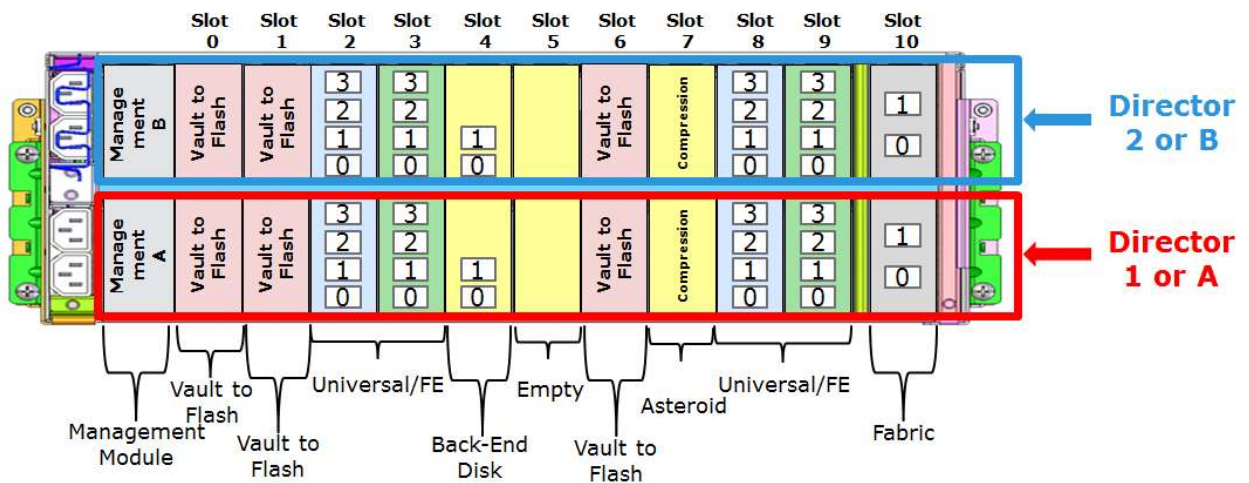


Figure 32 Rear view of VMAX 250F V-Brick with physical port numbering

The following figure illustrates the slots and associated logical ports of a VMAX250F V-Brick.

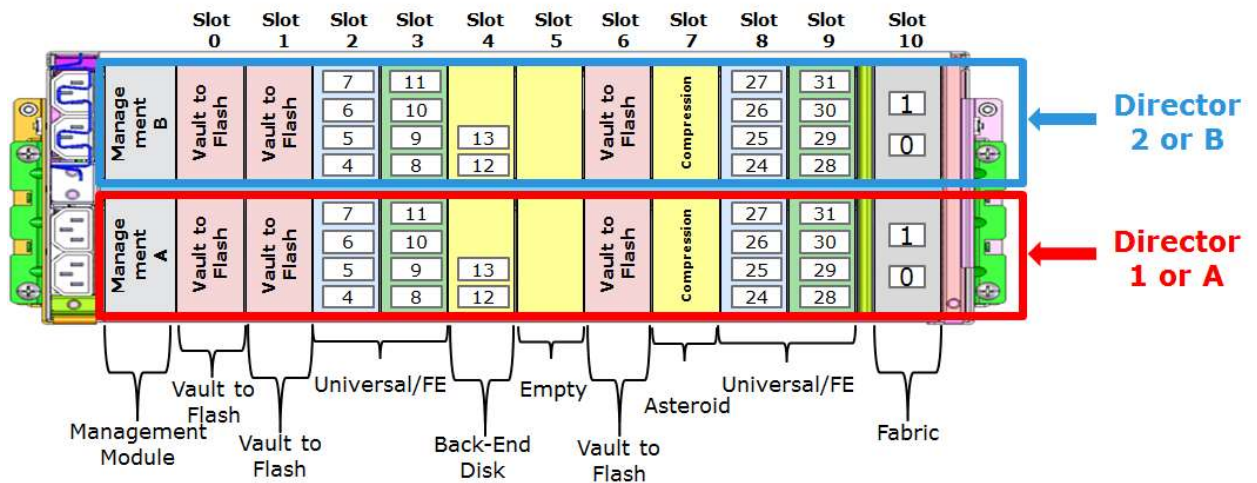


Figure 33 Rear view of VMAX 250F V-Brick with logical port numbering

Each director board also contains an MMCS (Directors 1 and 2) or Management Module (Directors 3 and 4), three Flash I/O Modules, a Compression I/O Module, and a 4-port 12 Gb/s SAS Back-end I/O Module.

VMAX 250F systems do not contain InfiniBand switches. In a two V-Brick system, each director communicates with the directors in the other V-Brick through redundant, directly-connected, 56 Gb/s inter-director links.

Each director also has redundant power and cooling modules.

11.2 VMAX 250F back-end

VMAX 250F systems incorporate a 12 Gb/s SAS back-end. Each V-Brick can have up to two DAEs, with a total of four DAEs per system. Each DAE can hold up to 25 2.5" 12 Gb/s SAS SSD drives.

Each VMAX 250F DAE has a pair of redundant power/cooling modules. Therefore, each DAE does not have internal fault zones like other VMAX All Flash models do, the 250F supports RAID 5 (3+1), RAID 5 (7+1) and RAID 6 (6+2) configurations.

VMAX 250F systems support Direct Sparring and have the same spare drive count requirements as the other VMAX All Flash models: One spare drive is required for every 50 drives per disk group per engine.

11.3 VMAX 250F system power

Like other VMAX All Flash models, each VMAX 250F system has two power zones for redundancy. Each power zone connects to a separate dedicated or isolated AC power line. If AC power fails on one zone, the power subsystem continues to operate through the other power zone.

When two arrays are configured in a single cabinet, they will each have their own power feeds.

Battery backup functionality is provided by a pair of redundant Lithium-ion standby power supply (Li-Ion-SPS) modules for each engine.

11.4 VMAX 250F serviceability

VMAX 250F systems are not configured with a keyboard/video/mouse (KVM). When on-site and needing to access maintenance procedures, Service Personnel can connect a laptop to the dedicated service cables or connect via Remote Connectivity Software. Both methods are secure as they each require Secure Service Credentials (SSC).

12 VMAX 950F

The HYPERMAX OS Q2 2017 Release introduced support for a new VMAX All Flash array, the VMAX 950F. The VMAX 950F scales from 1 to 8 V-Bricks, providing a system total of up to 1920 drives, 4PBe capacity, 16TBr cache, 192 front-end ports (OS/Mixed) or 256 ports (MF). The VMAX 950F can be shipped with either the F package which includes base software titles or the FX package, adding additional software titles to the F bundle.

12.1 VMAX 950F V-Brick

Each V-Brick in a VMAX 950F contains two director boards. Each director board has up to four front-end slots (slot 9 will always be populated with a compression SLIC) with up to four front-end ports that can be used for Fibre Channel, SRDF, iSCSI, FICON or eNAS connectivity.

The following figure illustrates the slots and associated physical ports of a VMAX950F V-Brick.

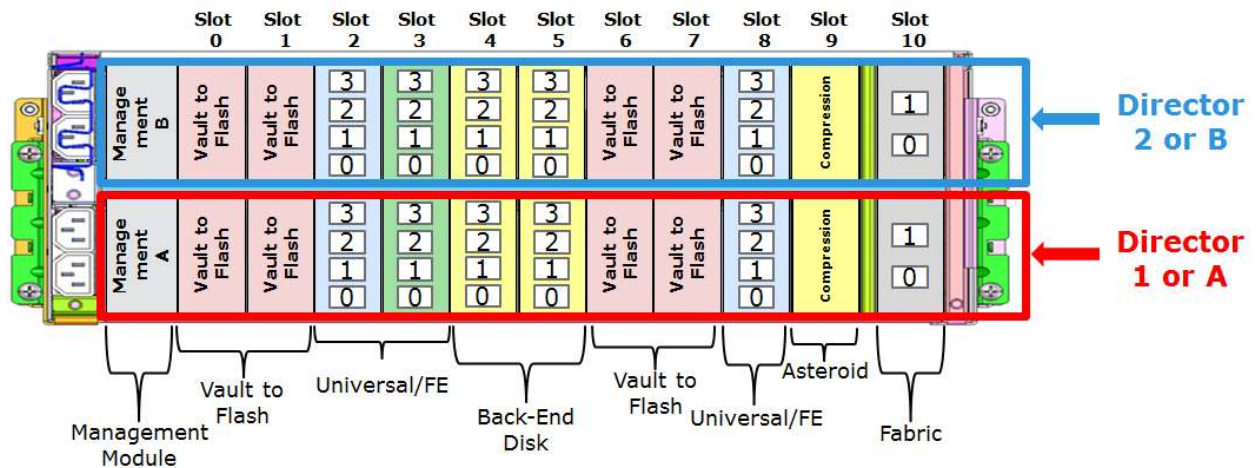


Figure 34 Rear view of VMAX 950F V-Brick with physical port numbering

The following figure illustrates the slots and associated logical ports of a VMAX950F V-Brick.

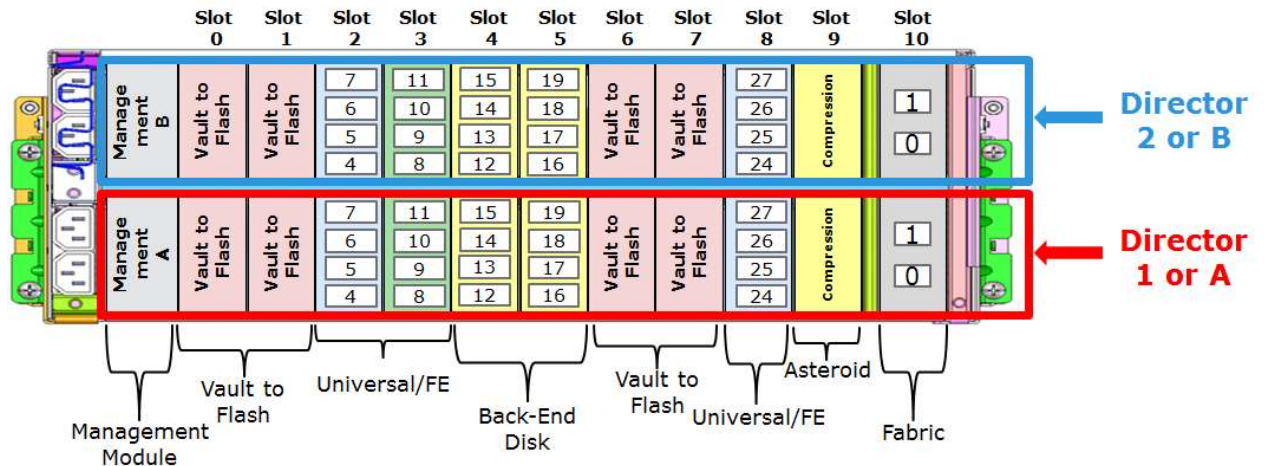


Figure 35 Rear view of VMAX 950F V-Brick with logical port numbering

Each director board also contains an MMCS (Directors 1 and 2) or Management Module (Directors 3-16), four Flash I/O Modules, a Compression I/O Module, a 4-port 6 Gb/s SAS back-end I/O Module and a two port fabric module. The fabric Module is for any system that has two or more engines, there are two redundant matrix interface board enclosures (MIBE) that connect to the fabric I/O modules of each director board. Each director also has redundant power and cooling modules.

12.2 VMAX 950F back-end

VMAX 950F systems incorporate a 6 Gb/s SAS back-end. Each V-Brick can have up to two DAEs, with a total of 16 DAEs per system. Each DAE can hold up to a maximum of 120 2.5" 6 Gb/s SAS SSD drives.

There are four drive expanders on each LCC. Each of the drive expanders connects to 30 drives. A total of four drive expanders connecting to 30 drives each gives a total of 120 drives. LCC A and LCC B connect to the dual ports on each drive for redundancy. The DAE is designed to have complete fault zone protection in the event of an unforeseen, power related failure. The DAE has multiple fault zones which effectively provide four virtual drive enclosures within a single DAE.

VMAX 950F systems support Direct Sparring and have the same spare drive count requirements as the other VMAX All Flash models: One spare drive is required for every 50 drives per disk group per engine.

12.3 VMAX 950F system power

Like other VMAX All Flash models, each VMAX 950F system has two power zones for redundancy. Each power zone connects to a separate dedicated or isolated AC power line. If AC power fails on one zone, the power subsystem continues to operate through the other power zone.

Battery backup functionality is provided by a pair of redundant Lithium-ion standby power supply (Li-Ion-SPS) modules for each engine.

12.4 VMAX 950F serviceability

The 950F systems provide full component-level redundancy to protect against a component failure and ensure continuous and uninterrupted access to information. This nondisruptive replacement capability allows Dell EMC support personnel to install a new component, initialize it if necessary, and bring it online without:

- Disrupting access to unaffected volumes
- Powering down the unit
- Stopping the operating system
- Taking unaffected channel paths offline
- Taking devices offline (other than the affected device)

13 Unisphere for PowerMax and Solutions Enabler

System health and component status can be monitored with Unisphere for PowerMax and Solutions Enabler. The tools provided by the management software are designed to give the end-user a high-level overview of the condition of the array. If these tools report any problems, and the user should contact Dell EMC Customer Support for proper, more in-depth investigation. Users should not try to perform any self-diagnostics or recovery. Dell EMC Technical Support Engineers have access to additional tools that allow for a thorough examination of the system. An investigation may already be underway as the array will send a call home to Dell EMC Customer Support for issues.

13.1 Unisphere for PowerMax system health check

Unisphere for PowerMax has a system health check procedure that interrogates the health of the array hardware. The procedure checks various aspects of the system and reports the results as either pass or fail. The results are reported at a high level with the intent of either telling the user that there are no hardware issues present or that issues were found, and the user should contact Dell EMC Customer Support for further investigation.

The health check procedure is accessed from the System Health.

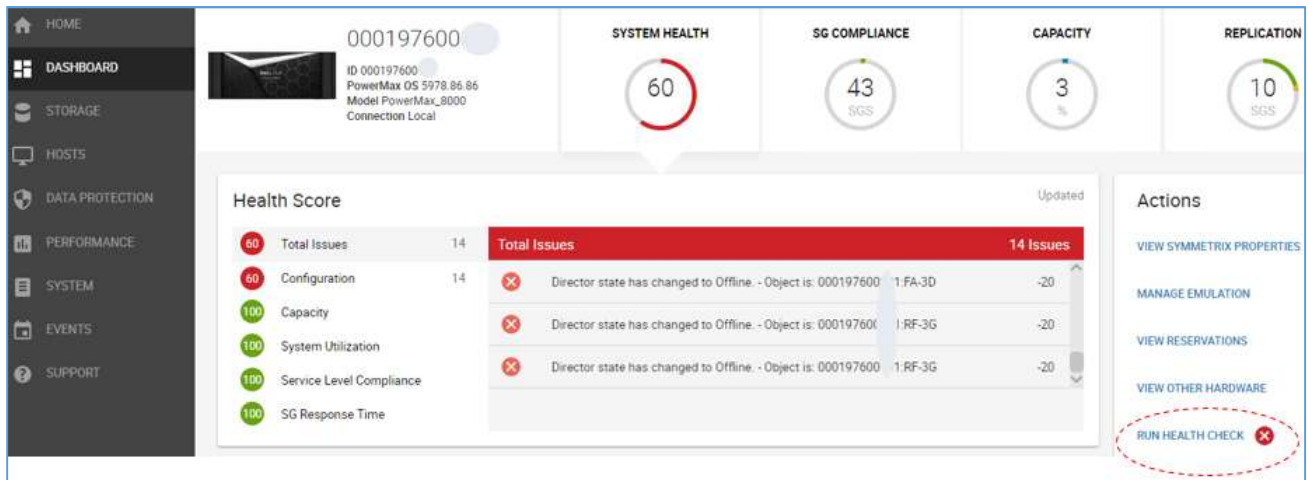


Figure 36 Unisphere System Health Dashboard

The test takes several minutes to complete. When complete, clicking the Run Health Check link displays test results in the format shown in the following figure.

Health Check 000197600	
Time of last run	Fri Dec 01 2017 10:30:58 GMT-0500
Result	✘ FAILED
Name	Status
Vault State Test	✔
Spare Drives Test	✔
Memory Test	✔
Locks Test	✔
Emulations Test	✔
Environmentals Test	✘
Battery Test	✔
General Test	✔
Compression And Dedup Test	✔

Figure 37 Health Check Results

13.2 Unisphere alerts

Unisphere for PowerMax has alerts for component failures. These alerts are optional and not set by default. The intent of these alerts is to inform the user of issues they may be affected by. For example, failure of a front-end I/O module will cause ports to go offline.

Alerts will not be sent for failure of an internal component such as a back-end I/O module because the failure is transparent to the user. The system will call home to alert Dell EMC Customer Support of the failure.

Enable the following alerts for component failures:

- Array Events
- Array Component Events
- Director Status
- Disk Status
- Environmental Alert

See the *Dell EMC Unisphere for PowerMax Installation Guide* for more information.

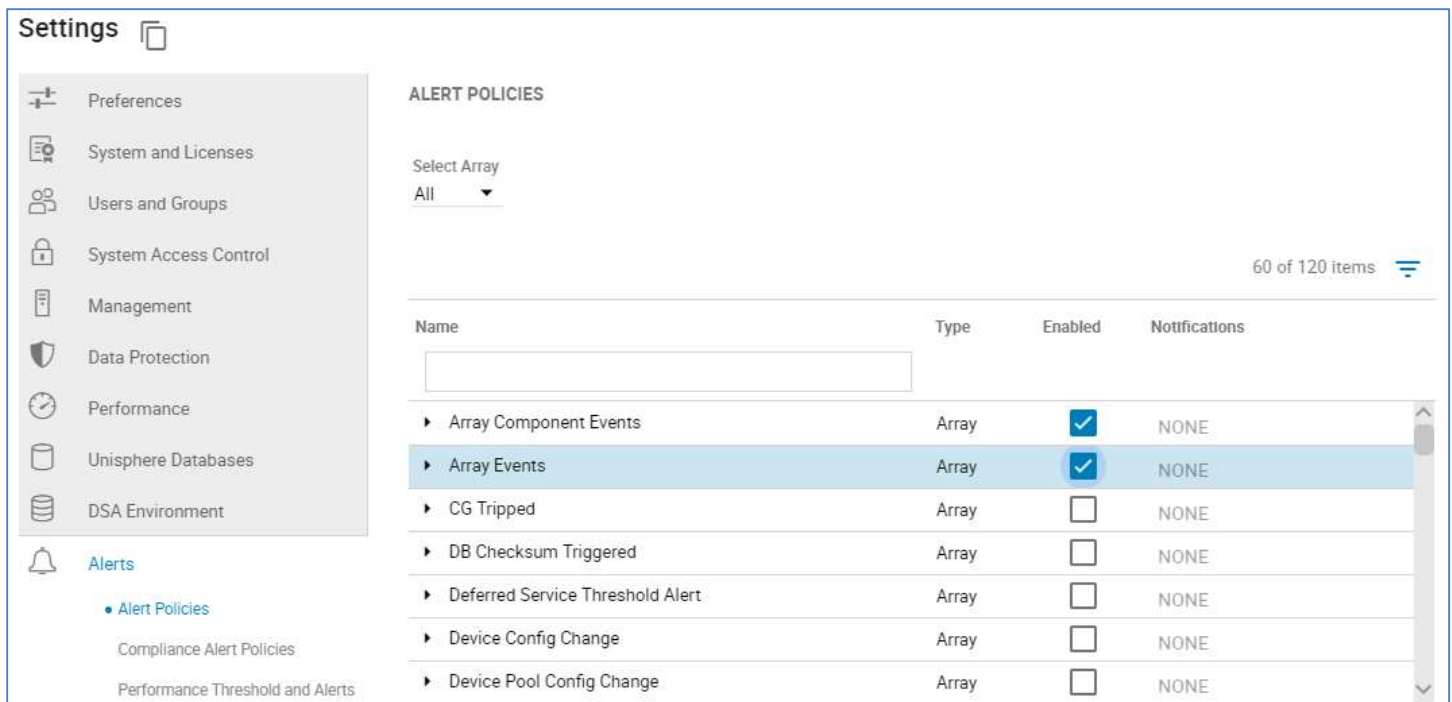


Figure 38 Unisphere for PowerMax Alert Settings

13.3 Solutions Enabler commands

In addition to the spare drive commands mentioned earlier, Solutions Enabler offers the following commands to report component status:

- `symcfg -sid <sid> list -env_data`

```
Symmetrix ID           : 000197800XYZ
Timestamp of Status Data : 08/14/2019 13:11:23

System Bay

    Bay Name           : SB-1
    Number of Standby Power Supplies : 2
    Number of Drive Enclosures       : 1
    Number of Enclosure Slots        : 1
    Number of MIBE Enclosures        : 2

Summary Status of Contained Modules
    All Standby Power Supplies      : Normal
    All Enclosures                  : Normal
    All Link Control Cards          : Normal
    ALL Drive Enclosures Power Supplies: Normal
    All Enclosure Slots             : Normal
    ALL Enclosure Slots Power Supplies : Normal
    All Fans                        : Normal
    All Management Modules          : Normal
    All IO Module Carriers          : Normal
    All Directors                   : Normal
    All MIBE Enclosures             : Normal
    ALL MIBE Enclosures Power Supplies : Normal
```

Figure 39 `symcfg -sid <sid> list -env_data`

- `symcfg -sid <sid> list -env_data -v`

```

Bay Name : SB-1
Bay LED state : Normal
Front Door Bay LED state : Normal
Number of Standby Power Supplies : 2
Number of Drive Enclosures : 1
Number of Enclosure Slots : 1
Number of MIBE Enclosures : 2

Status of Contained Modules
Standby Power Supplies
  SPS-1A (Aggregate) : Normal
  SPS-TRAY-1A : Normal
  -1A : Normal
  SPS-1B (Aggregate) : Normal
  SPS-TRAY-1B : Normal
  -1B : Normal

Drive Enclosure Number : 1
Drive Enclosure State : Normal
LCC-A : Normal
LCC-B : Normal
PS-A : Normal
PS-B : Normal

Enclosure Slot Number : 1
Enclosure Slot State : Normal
MM-1 : Normal
MM-2 : Normal
DIR-1 : Normal
  PS-A : Normal
  PS-B : Normal
  FAN-0 : Normal
  FAN-1 : Normal
  FAN-2 : Normal
  FAN-3 : Normal
  FAN-4 : Normal
  BOOT-DRIVE-0 : Normal
DIR-2 : Normal
  PS-A : Normal
  PS-B : Normal
  FAN-0 : Normal
  FAN-1 : Normal
  FAN-2 : Normal
  FAN-3 : Normal
  FAN-4 : Normal
  BOOT-DRIVE-0 : Normal

MIBE Name : MIBE-A
MIBE State : Normal
PS-A : Normal
PS-B : Normal
CM : Normal

MIBE Name : MIBE-B
MIBE State : Normal
PS-A : Normal
PS-B : Normal
CM : Normal

```

Figure 40 `symcfg -sid <sid> list -env_data -v`

14 Summary

VMAX All Flash and VMAX3 platforms improve upon every aspect of the already extraordinary VMAX product line. This is a system with highly redundant hardware components, creating a remarkably reliable environment that has also been condensed into a configuration that minimizes carbon footprint in the data center and increases total cost of ownership. The introduction of PowerMaxOS further enhances the customer's experience through new technologies such as service level-based provisioning, making storage management easier while also increasing upon availability of data through improvements to concepts such as vaulting, disk sparing, and RAID.

Local and remote replication suites also bring the system to an elevated level of availability, through TimeFinder SnapVX and SRDF, respectively. The serviceability aspects are just as distinguished, making the component replacement process quick and easy.

The key enhancements that improve the reliability, availability, and serviceability of the systems make VMAX All Flash and VMAX3 the ideal choice for critical applications and 24x7 environments demanding uninterrupted access to information.

A Resiliency testing

A.1 Dell EMC internal QE testing

Dell EMC internal QE testing Quality Engineering (QE) teams perform thorough testing of all FRUs. Each FRU is tested multiple times for each code level with very specific pass/fail criteria.

Standard tests perform verification of the GUI-based scripted replacement procedures that are used by EMC field personnel. The tests are designed to verify the replaceability of each FRU without any adverse effects on the rest of the system, and to verify the functionality and ease-of-use of the scripted procedures. These tests are straightforward replacement procedures performed on operational components.

Non-standard tests are also performed on components that have failed either by error injection or hot removal of the component or its power source. These tests also incorporate negative testing by intentionally causing different failure scenarios during the replacement procedure. Please note that removing a drive hot will not cause sparing to invoke. This behavior is optimal as the system knows the device has not gone bad. The correct course of action is to recover the drive rather than go through needless sparing and full rebuild processes.

Negative tests are designed to make sure that the replacement procedure properly detects the error and that the rest of the system is not affected.

Some examples of negative tests are:

- Replacing the wrong component
- Replacing component with an incompatible component
- Replacing component with a faulty component
- Replace component with a new component that has lower code that needs to be upgraded
- Replace component with a new component that has higher code that needs to be downgraded
- Replacing component with the same component and make sure script detects and alerts the user that the same component is being used
- Improperly replace a component (miscabled, unseated, etc)
- Initiate a system vault save (system power loss) operation during a replacement procedure
- Create a RAID group failure in an SRDF R1 array and verify local hosts continue to operate by accessing data from the R2 array
 - Replace the drives and rebuild RAID data from remote R2 array
 - Test with two drives in a RAID 1 or RAID 5 group, and three drives in a RAID 6 group
 - Test with SRDF/Metro, SRDF/S, and SRDF/A (when not in a spillover state)

Both the standard and non-standard tests are performed on all system models and various configurations with customer-like workload running on the array. Tests are also performed repeatedly to verify there are no residual issues left unresolved that could affect subsequent replacements of same or different component(s). Components that are known to fail more frequently in the field, as well as complex component replacements, are typically tested more frequently.

A.2 On-site proof-of-concept demonstrations

Proof-of-Concept (POC) engagements often include requests to demonstrate system resiliency. The requests are aimed at failing a component while the user monitors the behavior of the system. Typically, the main goal is to demonstrate that the system will remain online and available with a component in a failed state. Other goals of the demonstration may include performance monitoring, time to recover the component/array, time to engage support, along with other aspects.

The official Dell EMC process for On-Site POC demonstrations is available to Dell EMC Account Team members through the following link:

<https://inside.dell.com/docs/DOC-224096>

B Technical support and resources

[Dell.com/support](https://www.dell.com/support) is focused on meeting customer needs with proven services and support.

[Storage technical documents and videos](#) provide expertise that helps to ensure customer success on Dell EMC storage platforms.

B.1 Related resources

- Dell EMC VMAX All Flash Product Guide for VMAX 250F, 450F, 850F and 950F
- Dell EMC VMAX3 Family Product Guide
- Dell EMC VMAX3 Service Level Provisioning with Fully Automated Storage Tiering (FAST)
- Dell EMC TimeFinder SnapVX Local Replication Technical Note
- Dell EMC SRDF/Metro Overview and Best Practices Technical Note
- Dell EMC PowerPath Family Product Guide
- Dell EMC VMAX3 and VMAX All Flash Data at Rest Encryption White Paper