

PowerStore Technischer Deep Dive

September 23, 2020

DELLTechnologies

Hardware Overview

Terminology (Hardware) - PowerStore

PowerStore Term	Definition	Previous Terms
Node	Component within the base enclosure that contains processors and memory. Each appliance consists of 2 nodes.	Storage Processor (SP), Controller
Base Enclosure	Used to reference the enclosure containing both Nodes (Node A and Node B) and 25x NVMe drive slots	Disk Processor Enclosure (DPE), Array
Expansion Enclosure	Enclosures that can be attached to a base enclosure to provide additional storage in the form of SAS drives	Disk Array Enclosure (DAE), JBOD
Appliance	Term used for solution containing a base enclosure and any attached expansion enclosures. The size of an appliance could be just the base enclosure or the base enclosure plus expansion enclosures.	N/A
Cluster	Multiple appliances in a single grouping. Clusters can consist of one appliance or more. Clusters are expandable by adding more appliances (up to 4).	N/A
Embedded Module	Module embedded to each node providing mgmt, host front-end connectivity, SAS expansion ports, and selectable 4-port card	N/A
4-Port Card	Card for each node that provides 4 ports for mgmt. and front-end ports. Selectable in 25GbE and 10GbE Base-T options.	N/A

Overview Models

PowerStore™

- Overall family brand that encompasses all models is called “PowerStore”
- “PowerStore T models” refers to SAN/NAS only deployment mode of PowerStore
- “PowerStore X models” refers to SAN and built-in ESX hypervisor deployment mode of PowerStore
- From marketing perspective, it is acceptable to omit “T” and “X” when referring to a particular system configuration
 - i.e. “PowerStore 5000 models have these system limits...”

PowerStore™



PowerStore T models

- PowerStore 1000T
- PowerStore 3000T
- PowerStore 5000T
- PowerStore 7000T
- PowerStore 9000T

PowerStore X models

- PowerStore 1000X
- PowerStore 3000X
- PowerStore 5000X
- PowerStore 7000X
- PowerStore 9000X

Overview

PowerStore Models

	PowerStore 1000	PowerStore 3000	PowerStore 5000	PowerStore 7000	PowerStore 9000
CPU per Appliance	4x 8C @ 1.8G Intel Xeon	4x 12C @ 2.1G Intel Xeon	4x 16C @ 2.1G Intel Xeon	4x 20C @ 2.4G Intel Xeon	4x 28C @ 2.1G Intel Xeon
Memory per Appliance	384GB	768GB	1152GB	1536GB	2560GB
NVRAM drives per Appliance	2		4		
Max Storage Drives per Appliance	96		96		
Supported Drives	NVMe SCM, NVMe SSD, SAS SSD				
Embedded Ports	4x 25/10/1 GbE Optical or 4x 10/1 GbE BaseT				
Support I/O Modules (2 slots per node)	4x 32/16/8 Gb FC, 4x 25/10/1 GbE Optical (PowerStore T only), 4x 10/1 GbE BaseT (PowerStore T only)				
Supported Expansion Shelves	2.5" 25-Drive SAS SSD				

Chassis

Base Enclosure (Front)

- 25x Drive Slots supports:

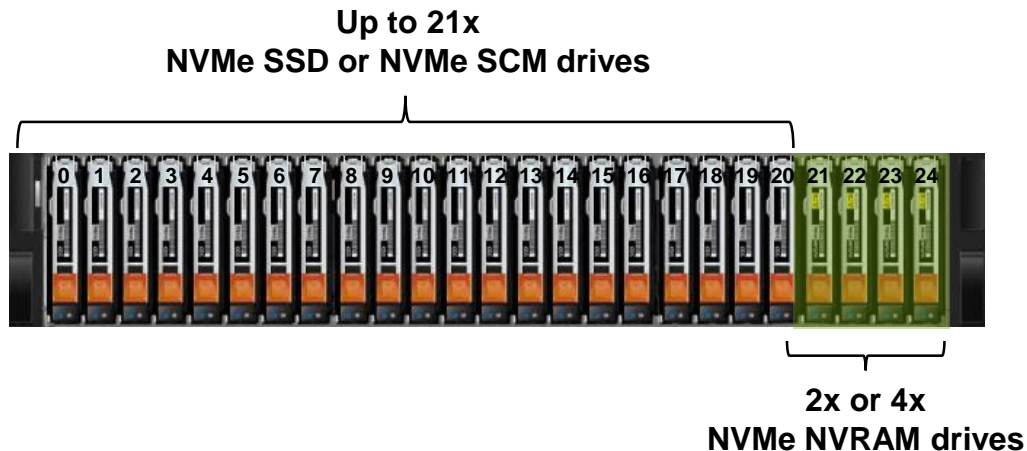
- NVMe SSD or
- NVMe SCM

- Last 4x slots support:

- NVMe NVRAM used for write caching
- Support either 2 or 4 NVRAM drives depending on model
- Four slots will always be reserved for the NVRAM drives even when the model uses 2 NVRAM drives

- 2U in height (rack units), 31.2” depth, ~92 lbs fully loaded

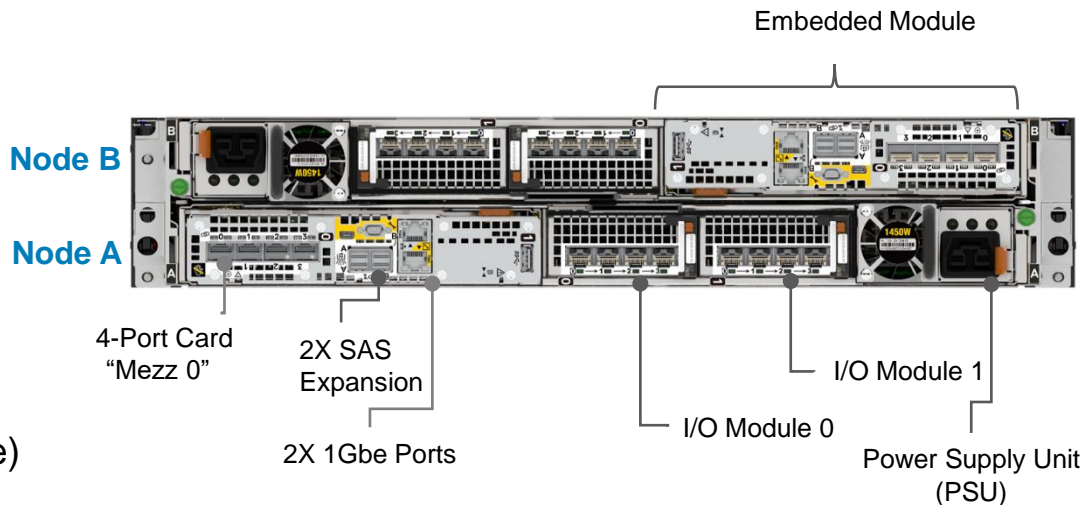
- Two personnel required to install the system



Chassis

Base Enclosure (Rear View)

- 2 Nodes in inverted arrangement
- Each Node has Embedded Module with:
 - 1x 4-Port Card Slot – Mezz 0
 - Embedded 2x1GbE Port (Mgmt/Service)
 - Embedded SAS Expansion Ports (x2)
- Every appliance must contain a 4-Port Mezz 0
 - The 4-Port Card is used for connections such as cluster interconnect and management of the appliance
 - Customers will have 2 types of 4-Port Mezz cards to select from: Optical or Copper
- Each node has 2x I/O Module slots for configuring optional additional frontend ports
- Each node has a power supply unit (PSU) which can power both nodes if needed



PowerStore Drive Support

Summary

- All 25 drive slots support NVMe drives, SAS SSD drives are not supported in any slot on base enclosure
- SAS SSD drives are only supported in attached expansion enclosures



Drive Type	Location	Purpose
NVMe NVRAM	Slots 21-24 in base enclosure	Cache
NVMe SCM	Slots 0-20 in base enclosure	User Data/Meta Data
NVMe SSD	Slots 0-20 in base enclosure	User Data/Meta Data
SAS SSD	Slots 0-24 in expansion enclosure	User Data/Meta Data

PowerStore Drive Support

NVMe SCM – System Configuration



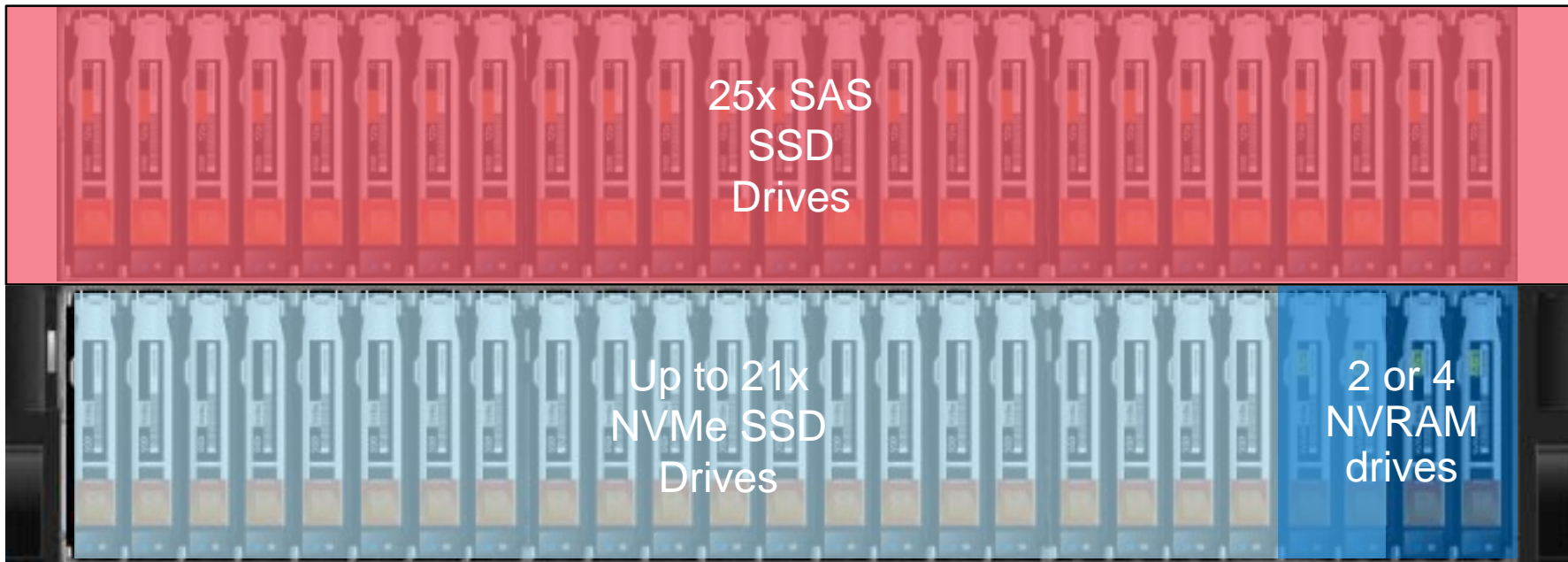
PowerStore Drive Support

NVMe SSD – System configuration (no expansion enclosures)



PowerStore Drive Support

NVMe SSD – System configuration (With expansion enclosure)



PowerStore Drive Support

Storage Type	Usage / Purpose	GB	2.5" Base Enclosure (25 slot)	2.5" Expansion Enclosure (25 slot)
SAS SSD	User Data/Metadata	1,920		✓
	User Data/Metadata	3,840		✓
	User Data/Metadata	7,680		✓
NVMe SSD	User Data/Metadata	1,920	✓	
	User Data/Metadata	3,840	✓	
	User Data/Metadata	7,680	✓	
	User Data/Metadata	15,360	✓	
NVMe SCM	User Data/Metadata	375	✓	
	User Data/Metadata	750	✓	
NVMe NVRAM	Cache	8	✓	

- All drive offerings are encrypted (SEDs)
- All drive offerings are FIPS certified except for NVMe NVRAM drives

PowerStore T Model Overview

Overview

PowerStore T Model

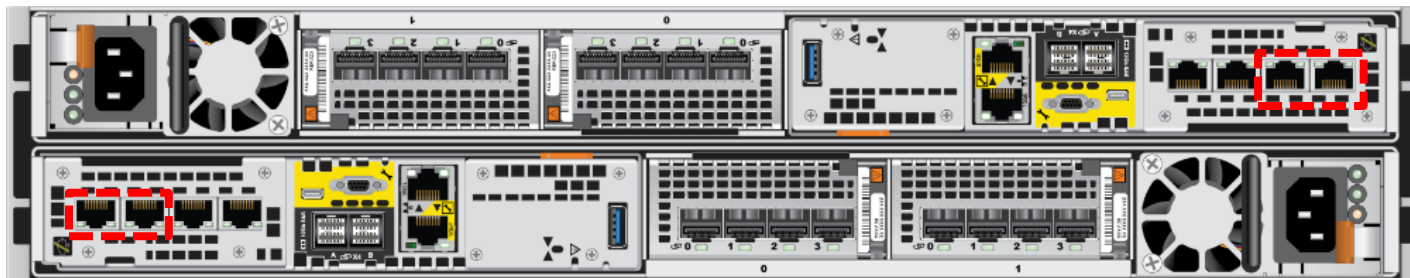
- PowerStoreOS installed directly on purpose built hardware
 - 2U2N
 - All NVMe Base Enclosure
 - Dual-socket Intel Xeon architecture
- Unified Storage array
 - SAN (FC/iSCSI)
 - NAS (NFS/SMB/FTP/SFTP)
 - vVol (FC/iSCSI)
- Active-Active architecture
 - Each node has access to the same storage
 - Active-optimized/Active-unoptimized front end connectivity



Networking

System Bond

- PowerStore T models contain a system bond by default
- Ports 0 & 1 of 4-Port Card are automatically bonded together in LACP mode
- This bond is essential to the networking configuration and cannot be removed
- System bond can provide high availability to cluster data and metadata traffic
- Link aggregation configuration on the switch side is optional
 - If switches are not configured to support LACP, Ports 0 & 1 will operate in Active/Passive mode

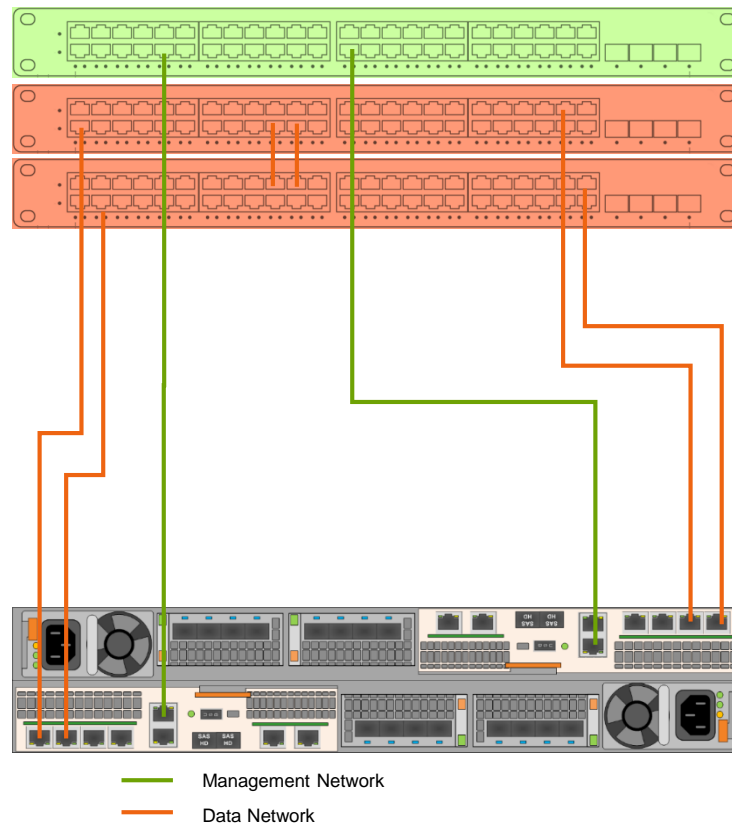


Deployment

Cabling

- Management Network
 - 1x Out of Band (OOB) Management switch
 - 2x OOB Management is supported for HA
 - Onboard 1GbE Ports
- Data Network
 - 2x Top of Rack (ToR) Ethernet Switches
 - Bonded 4-Port Card Ports 0 & 1
 - Layer 2 interconnect
- Diagram depicts minimal required cabling
 - Cluster will alert on not having dual redundant switches

Example Config

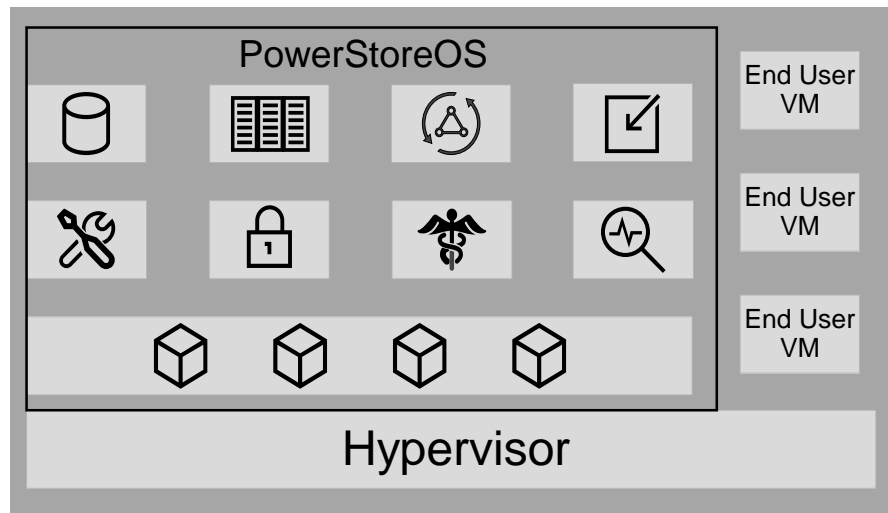


PowerStore X Model Overview

Overview

PowerStore X Model

- ESXi runs directly on purpose built hardware
 - 2U2N
 - All NVMe Base Enclosure
 - Dual-socket Intel Xeon architecture
- PowerStoreOS runs in a virtual machine
 - Referred to as the Controller VM
- Capabilities:
 - SAN (FC/iSCSI)
 - vVol (FC/iSCSI)
 - Embedded Applications (Virtual Machines)
 - Currently, clustering of multiple X model appliances is not supported
- Active-Active architecture
 - Each node has access to the same storage
 - Active-optimized/Active-unoptimized front end connectivity



Overview

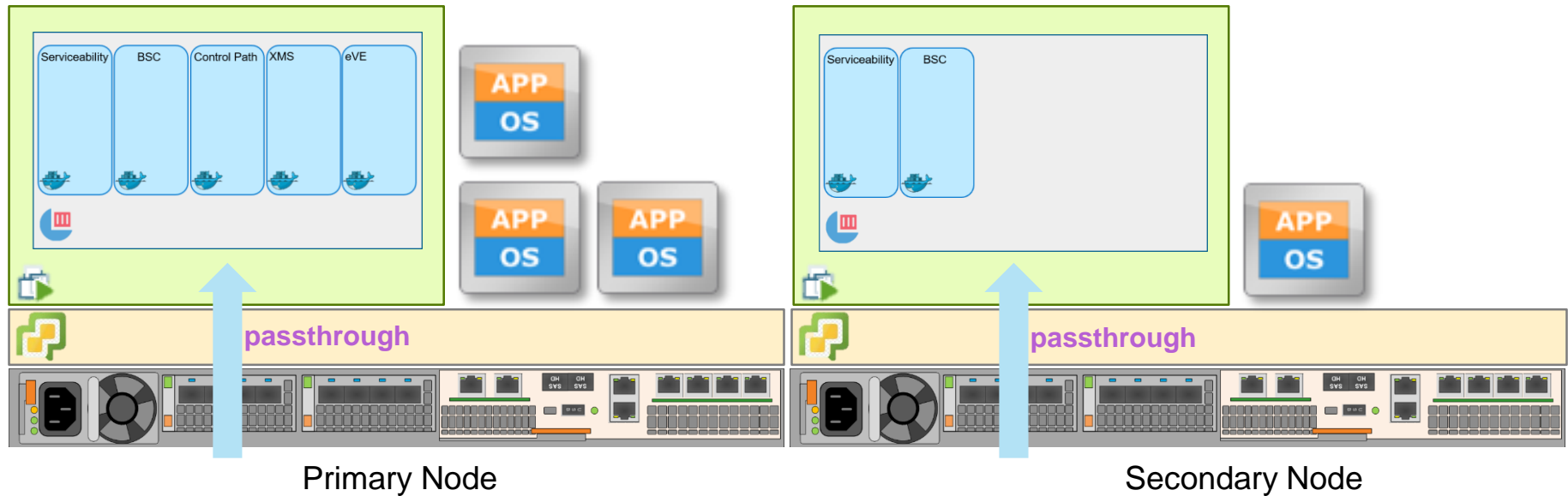
ESXi



- VMware ESXi 6.7 U2 installed directly onto each node
- PowerStore X model appliance contains ESXi Cluster of two ESXi hosts
 - 1 ESXi host per node
- Requires existing vCenter and license for deployment
 - VMware vSphere Enterprise Plus license
 - Customer can purchase license or use existing
- VMware components are automatically configured
 - ESXi Cluster
 - vSphere HA
 - Distributed Virtual Switch
 - Etc.
- Customer Virtual Machines will leverage PowerStore storage and data services

Overview

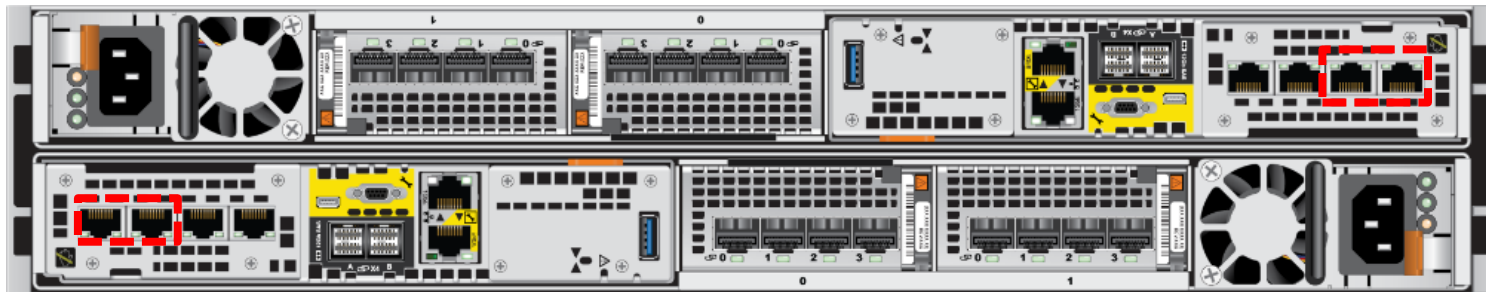
Diagram



Networking

Teaming

- PowerStore X implements various port teaming configurations
 - This provides high availability for network traffic
- PowerStore X requires 4-Port Card Ports 0 & 1 cabled at a minimum
 - ESXi networking is configured to support cabling all 4 ports on the 4-Port Card
- The use of these ports by specific networks will be detailed

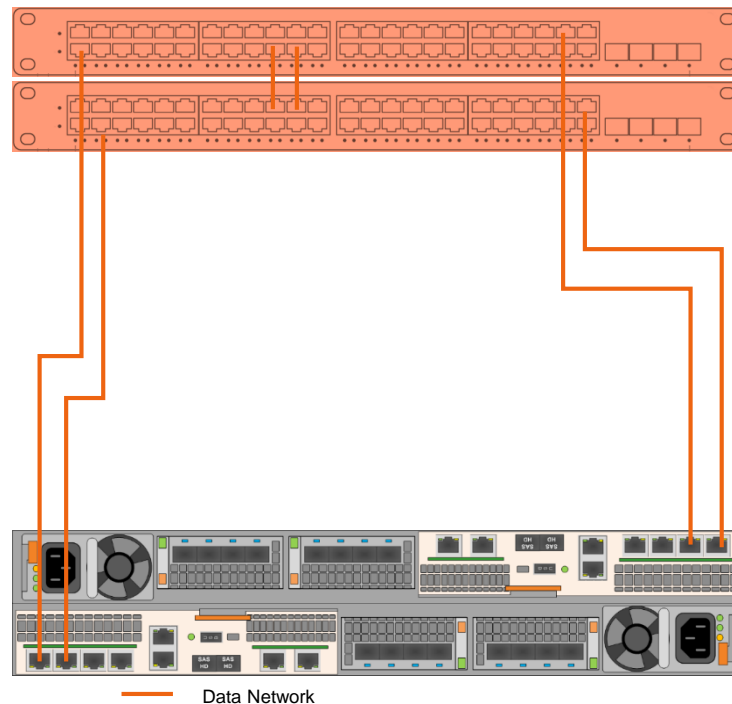


Deployment

Cabling

- Data Network
 - 2x Top of Rack (ToR) Ethernet Switches
 - Bonded 4-Port Card Ports 0 & 1
 - Layer 2 interconnect
- Diagram depicts minimal required cabling
 - Cluster will alert on not having dual redundant switches

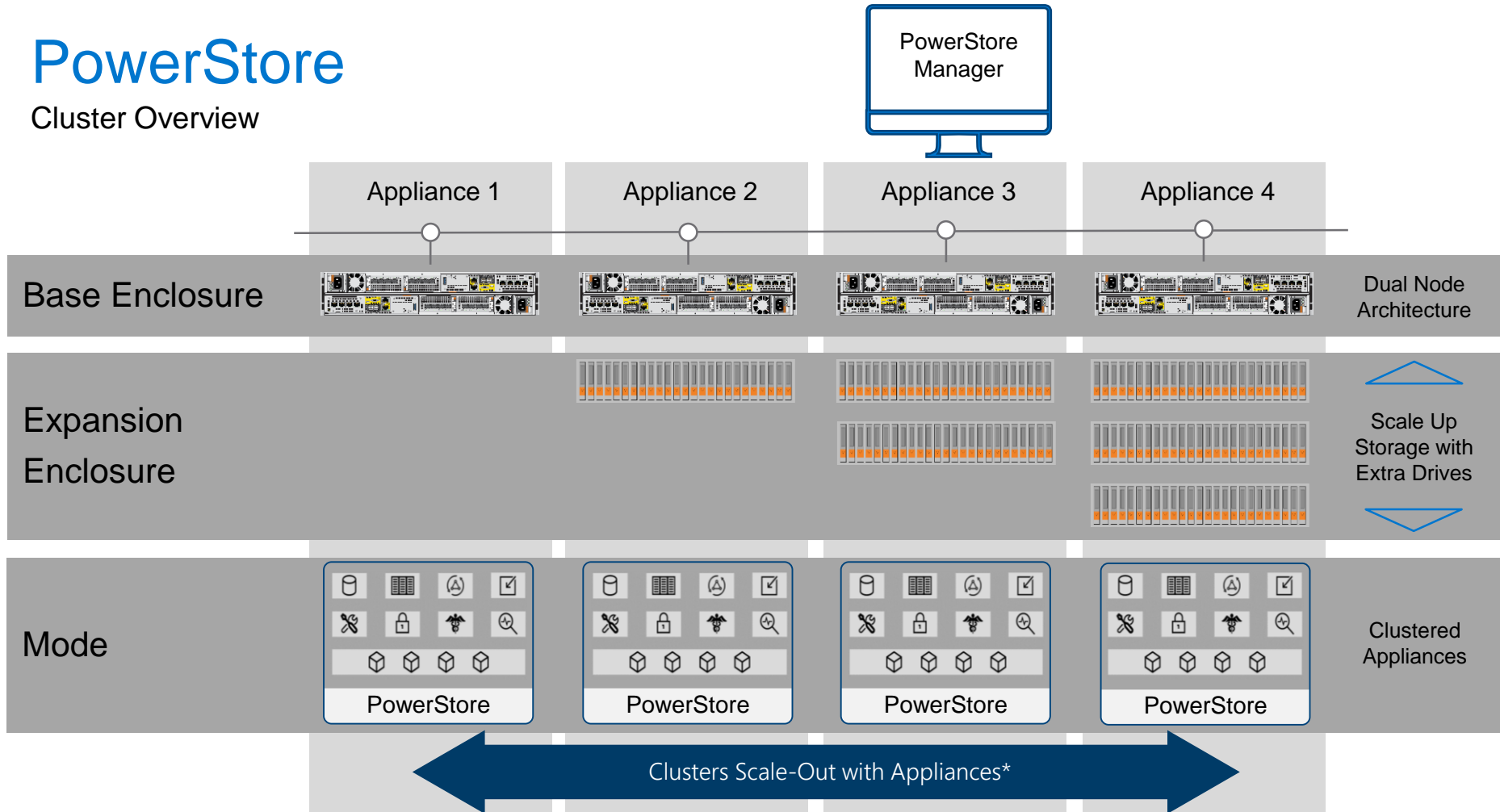
Example Config



PowerStore Clustering

PowerStore

Cluster Overview



*Clustering not supported for PowerStore X appliances in PowerStore v1

Clustering

Details

- Currently, you can only cluster PowerStore T Model appliances
 - PowerStore X Model appliances do not support clustering
 - You can mix models within the series (e.g, PowerStore 1000 & PowerStore 3000)
- Maximum size of 4 appliances

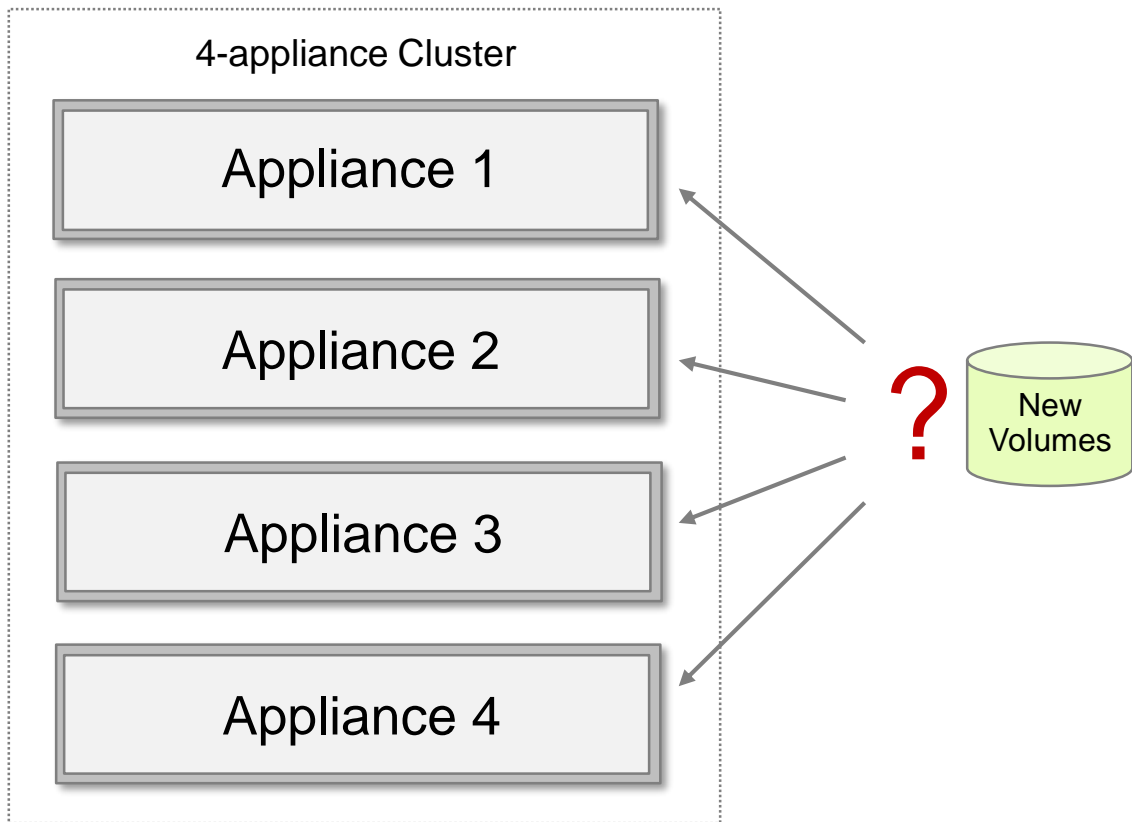
<input type="checkbox"/>	Name	Model	Service Tag	Alerts	IP Address	Total Capacity	Used Capacity
<input type="checkbox"/>	Appliance A	PowerStore 1000	DemoDST		10.245.60.1	3.0 MB	2.1 KB
<input type="checkbox"/>	Appliance B	PowerStore 2000	DemoDST		10.245.60.0	81.1 MB	22.5 KB

Resource Balancing

Overview

What is Resource Balancer?

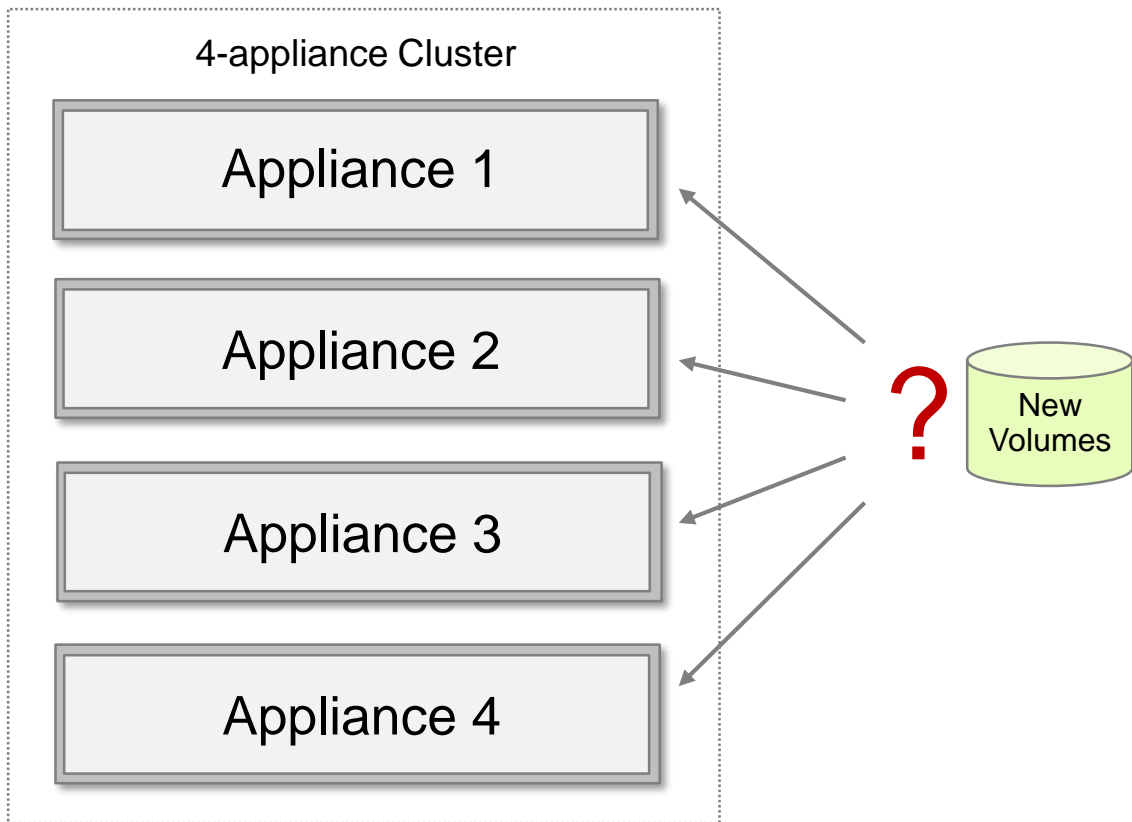
- The ability of PowerStore to use analytics to balance storage resources (volumes)
 - Which node to assign a new volume to on an appliance
 - Which appliance to assign a new volume to in a multi-appliance cluster



Overview

What is Resource Balancer?

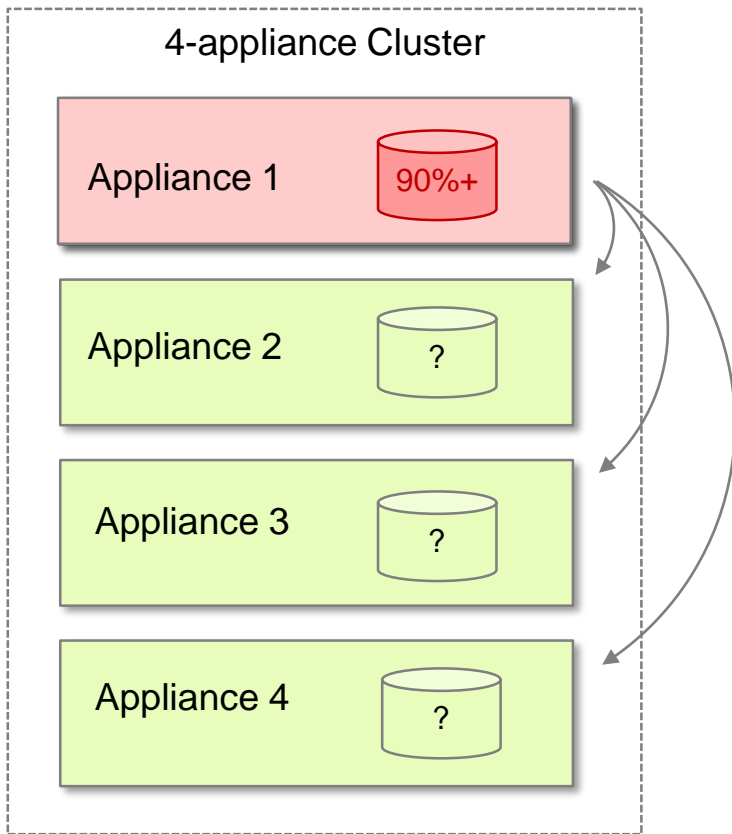
- Appliance assignments determined by:
 - Current storage space utilization
 - › Storage trends and forecasts
 - System limits
 - › Max volumes per appliance / volume group
 - Appliance status and health
 - › Offline, failures, read-only (100% full)
 - Performance metrics are not considered
 - Resource Balancer does not proactively or automatically move existing volumes from one appliance to another



Overview

What is Resource Balancing?

- The ability to migrate storage resources between appliances in a cluster
 - Manual migration
 - Assisted migration
- Leverage capacity monitoring, forecasts, and alerts with suggested remediation options

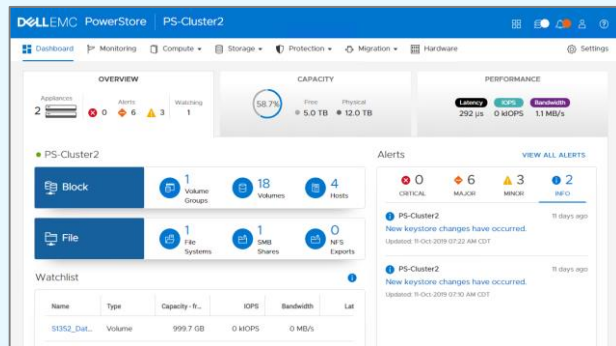


Management

PowerStore Manager GUI Overview

What is PowerStore Manager?

- HTML5-based GUI for managing PowerStore clusters
- No Java required
- Supported with popular browsers: Chrome, Firefox, Internet Explorer, Edge, Safari
- Simple, modern, intuitive management and monitoring interface
- No client install required
- Connect by HTTPS
 - HTTP or HTTP redirect to HTTPS is not supported



PowerStore CLI

- Standalone Client
 - Translates CLI into REST API calls in the background
- CLI still used frequently for administration
 - Easy to use, OS commands can be used to filter
- Allows batch jobs without programming skills
- Available for Windows and Linux
 - RPM for SLES 12, RHEL 7.x, Oracle Linux 6.5, 7
 - Windows 7,8,10 / Server 2012, 2012R2, 2016
 - Requires administrative privileges for installation
- Objects almost identical to REST-API

REST API

- The REST API allows you to interact with PowerStore Management functionality, including:
 - System settings and monitoring
 - Host and remote system connections
 - Network settings
 - Storage management
 - Data protection
 - Support configuration
- Presents a single, consistent interface to manage
- SSL encryption allows secure connection from client to system

SwaggerUI

Swagger UI – https://<PowerStoreClusterIP>/swaggerui

The screenshot shows the Swagger UI interface for the Storage Cluster API V1.0. At the top, there is a blue header with the Dell EMC logo on the left and a dropdown menu labeled "Select a definition" with "Storage Cluster API V1.0" selected. Below the header, the main content area displays the API title "Storage cluster API" with a version indicator "1.0" in a small circle. Underneath the title, there is a description: "[Base URL: 10.245.11.11/api/rest]" and a link to the JSON definition: "https://10.245.11.11:443/api/rest/openapi_1_0.json". A sub-description reads "Storage cluster REST API definition." followed by a link to the "End User License Agreement". Below this, there is a "Schemes" section with a dropdown menu currently set to "HTTPS". A search bar labeled "Filter by tag" is positioned below the schemes. At the bottom, an "alert" section is visible, stating "Use these resource types to manage events and alerts in the system." with a downward arrow icon.

Data Path

PowerStore Data Path

Dynamic RAID – RAID Resiliency Sets

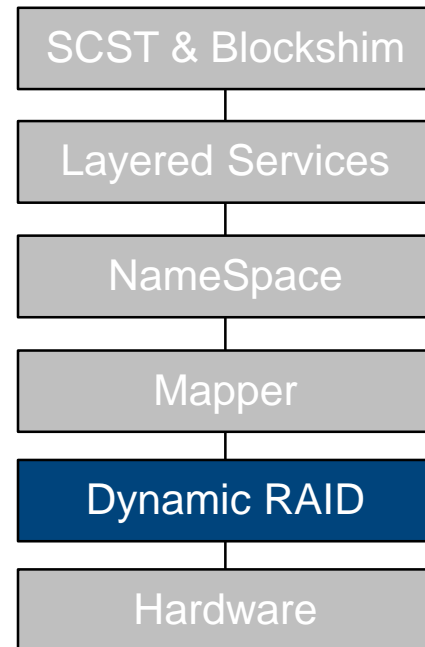
- The Solution

- Up to 25 Drives Per RRS
- May Span Across Enclosures
- 1 Spare Per 25 Drives
 - Spare space is distributed across all drives

- RAID Geometry

- Set on Initialization
- Automatic selection to maximize usable capacity

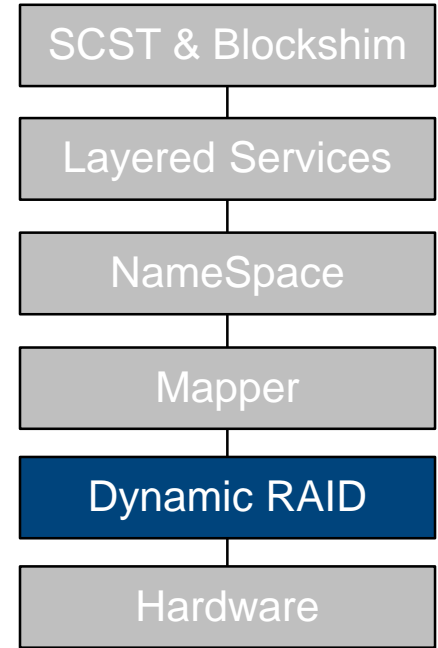
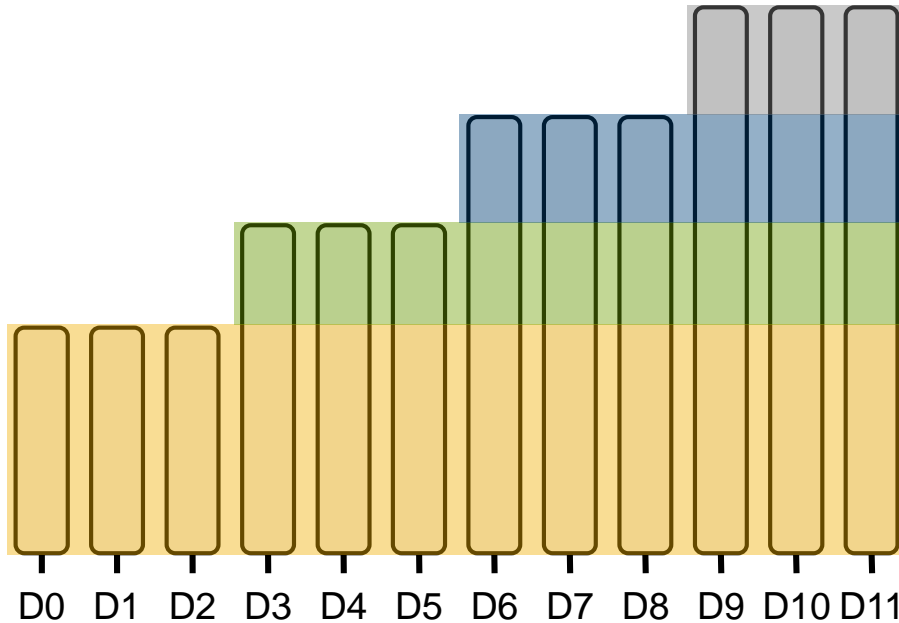
Initial Disk Count Requirement	Geometry	Maximum Drives
6	4+1	25
10	8+1	25



PowerStore Data Path

Dynamic RAID – Drive Types

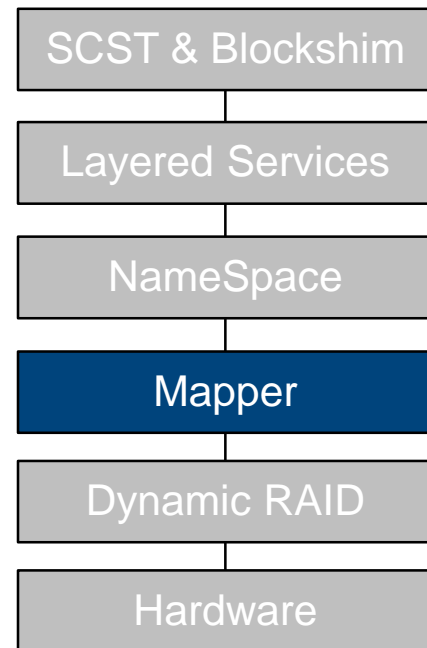
- Mixing Drive Types



PowerStore Data Path

Dedupe and Compression

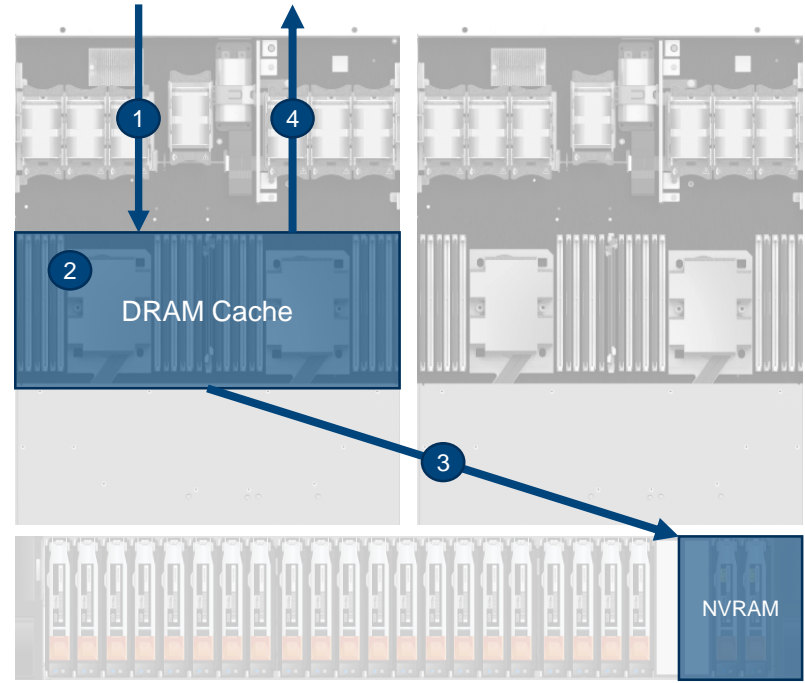
- **Main Concepts**
 - 4:1 data reduction ratio guarantee
 - 4K granularity
 - HASH > Read & Compare
- **Hardware Compression Offloaded**
 - Lewisburg Chip
 - Quick-Assist Technology
- **Checksum**
 - Stored with Metadata



Data Path Write Path

I/O Flow

1. The system receives a Write operation
2. The write request is stored in DRAM Cache
3. The data is saved in NVRAM Write Cache.
4. The system acknowledges the host.

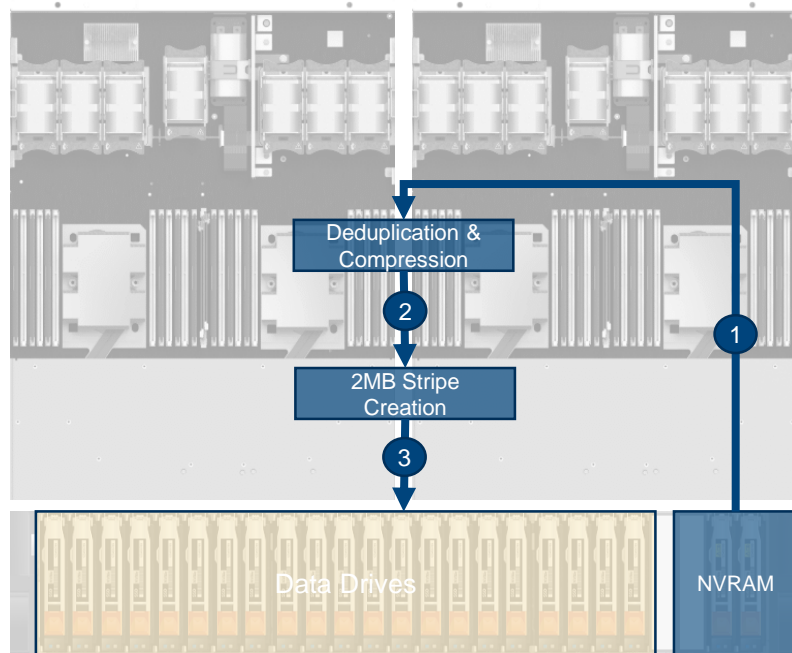


Data Path Write Path – “Late Bind”

I/O Flow

After the Host has been acknowledged, the system still processes the Write request in the background

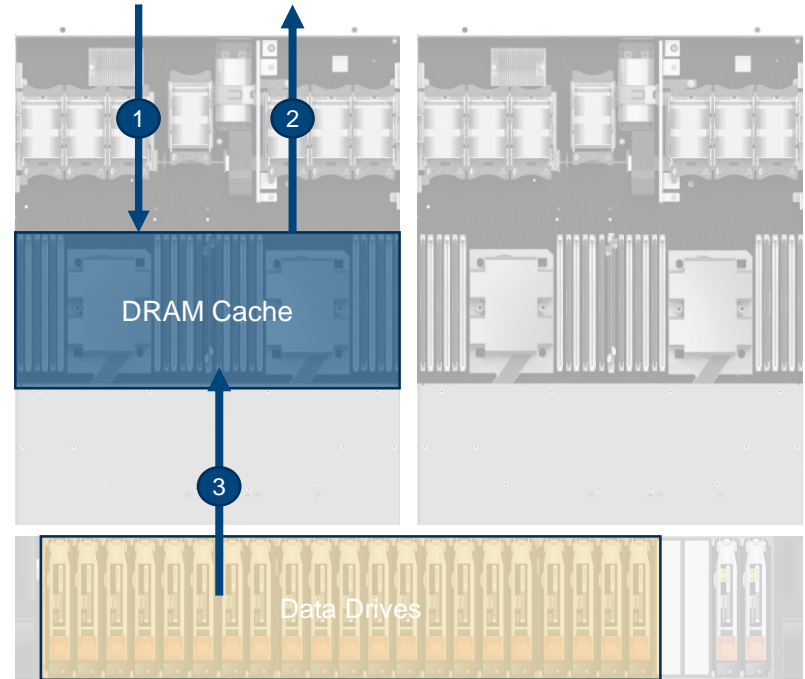
1. Write request is deduplicated and compressed
2. Continue to fill a 2MB stripe of compressed and deduplicated Write requests
3. Once the 2MB stripe is full, destage to drives



Data Path Read Path

I/O Flow

1. The system receives a Read operation
2. If the data is in the (DRAM) cache, fulfill the Read request
3. If the data is not in the (DRAM) cache, copy the data from the drives



File

PowerStore File Overview

- PowerStore File enables clients to access data over file protocols:
 - Server Message Block (SMB)
 - Network File System (NFS)
 - File Transfer Protocol (FTP)
 - SSH File Transfer Protocol (SFTP)
- Currently File Replication is not available
- File is only available on PowerStore T model appliances
 - Currently PowerStore X model appliances do not support File

Block Optimized or Unified

- PowerStore T model appliances can be configured as Block Optimized or Unified (block and file)
 - Selection determines resource allocation on the appliance
 - PowerStore X model appliances do not have this option as they do not support NAS
- Must be decided during initial configuration
 - Select Unified if there is a chance that you will need file functionality
 - No option to convert between Block Optimized and Unified afterwards
- PowerStore Sizer will have performance details for each configuration
 - Block-optimized systems have slightly higher block IOPS potential
- If Unified is selected, NAS installation is performed automatically after cluster creation completes

NAS Server High Availability

- In the event of a PowerStore node failure, NAS Servers automatically failover from one NAS node to the other
 - Failover generally completes within 30 seconds to avoid host timeouts
 - NAS Servers are automatically moved to the peer node during NDU
 - Failback is a manual process
- New NAS Servers are automatically assigned round-robin across nodes
 - All file systems associated with the NAS Server are served by the NAS Server's current node
- Current Node
 - The node that the NAS server is currently running on
 - Changing this moves the NAS server to run on a different node
- Preferred Node
 - The node that the NAS server should be ideally running on
 - Acts as a marker that is based on the round-robin algorithm

Protection Policies

Protection Policies

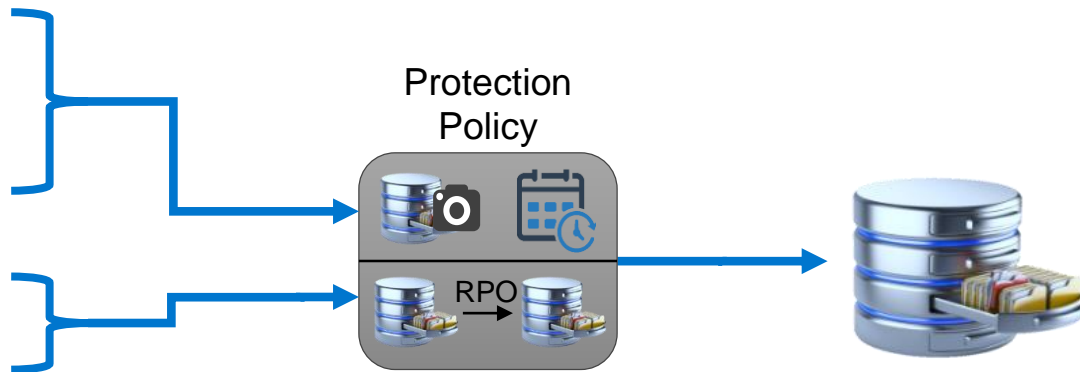
Overview

- A Protection Policy is a set of user defined rules used to establish local or remote data protection across storage resources
 - Users do not configure snapshot schedules or replication sessions on a resource, but rather assign a Protection Policy to it
- A Protection Policy consists of rules which define what level of protection to apply
- When a Protection Policy is assigned to a resource:
 - The Snapshot Rule is automatically applied
 - Replication is automatically configured

Protection Policies

Overview - Example

- Snapshot Rules:
 - Hourly snapshots
 - Daily snapshots
 - Weekly snapshots
- Replication Rules:
 - Asynchronous Replication
 - 1 hour RPO



Snapshots

Snapshots

Overview

- Snapshots are crash-consistent point-in-time copies of data that are stored in supported storage resources
 - Snapshots are the local data protection solution within PowerStore
 - Note: A crash-consistent snapshot is not application consistent. Application consistency requires I/O to be quiesced at the host, host buffers flushed, and a snapshot taken.
- Snapshots are not full copies of the original data
 - Pointer based, redirect on write technology
 - You should not rely on snapshots for mirrors, disaster recovery, etc.
- Snapshots consume overall system storage capacity as changes to the source are made
 - Ensure that the system has enough capacity to accommodate snapshots

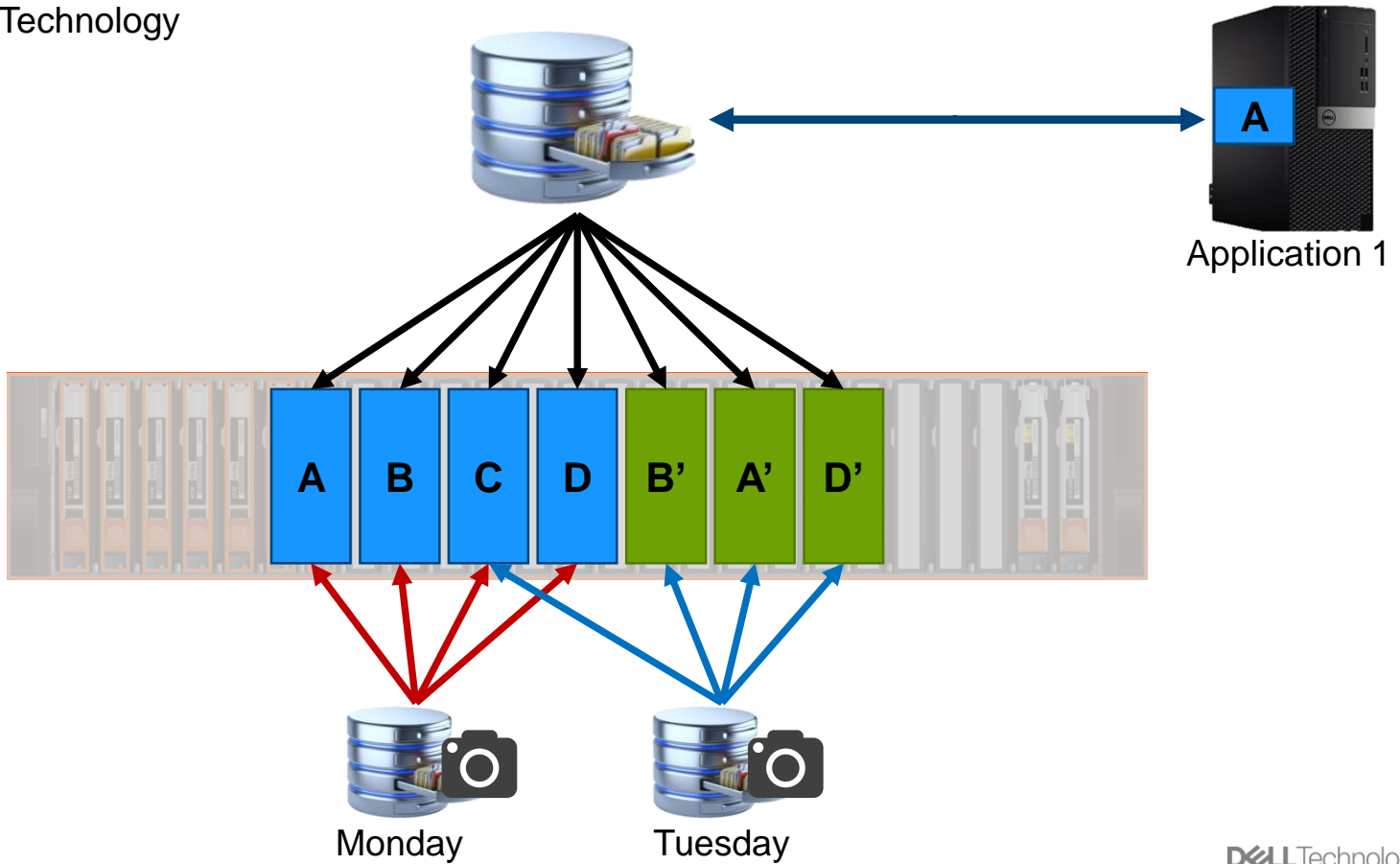
Snapshots

Overview (Continued)

- Snapshots are supported on volumes, volume groups, virtual machines, file systems, and thin clones
- Snapshots can be created:
 - Manually
 - Automatically via a user created Snapshot Rule within the system
 - To have snapshots automatically created via a predefined schedule on a storage resource, a Protection Policy containing a Snapshot Rule must be added to the storage resource
 - Automatically by a feature within the system
 - Such as replication
 - Hidden from the user

Snapshots

Redirect on Write Technology



Block Remote Replication

Overview

Use Cases - Remote Data Protection

- **Redundancy**
 - Increase data redundancy and fault tolerance levels
 - Failover to a secondary site
- **Compliance**
 - Require additional footprint.
 - for instance insurance companies, bank, and government sites
- **Migration**
 - Migrate data between storage systems
 - Tech Refresh

Replication Concepts

Supported Replication with PowerStore T / PowerStore X

- Supports PowerStore T model Arrays and PowerStore X model Arrays – also mixed
- RPO based Asynchronous Remote Replication
 - The acceptable amount of data, measured in units of time, which may be lost due to a failure
 - With RPO the also frequency of synchronization between Source and Destination is defined
 - Factor of time, not amount of changed data
 - PowerStore supports fixed RPO values between 5min and 24hr
- Replication is supported for following Storage Resources
 - Volumes
 - Volume Groups
 - Thin Clones
- IP based only
- Policy based
- Leverage PowerStore Snapshot technology

Replication Topologies

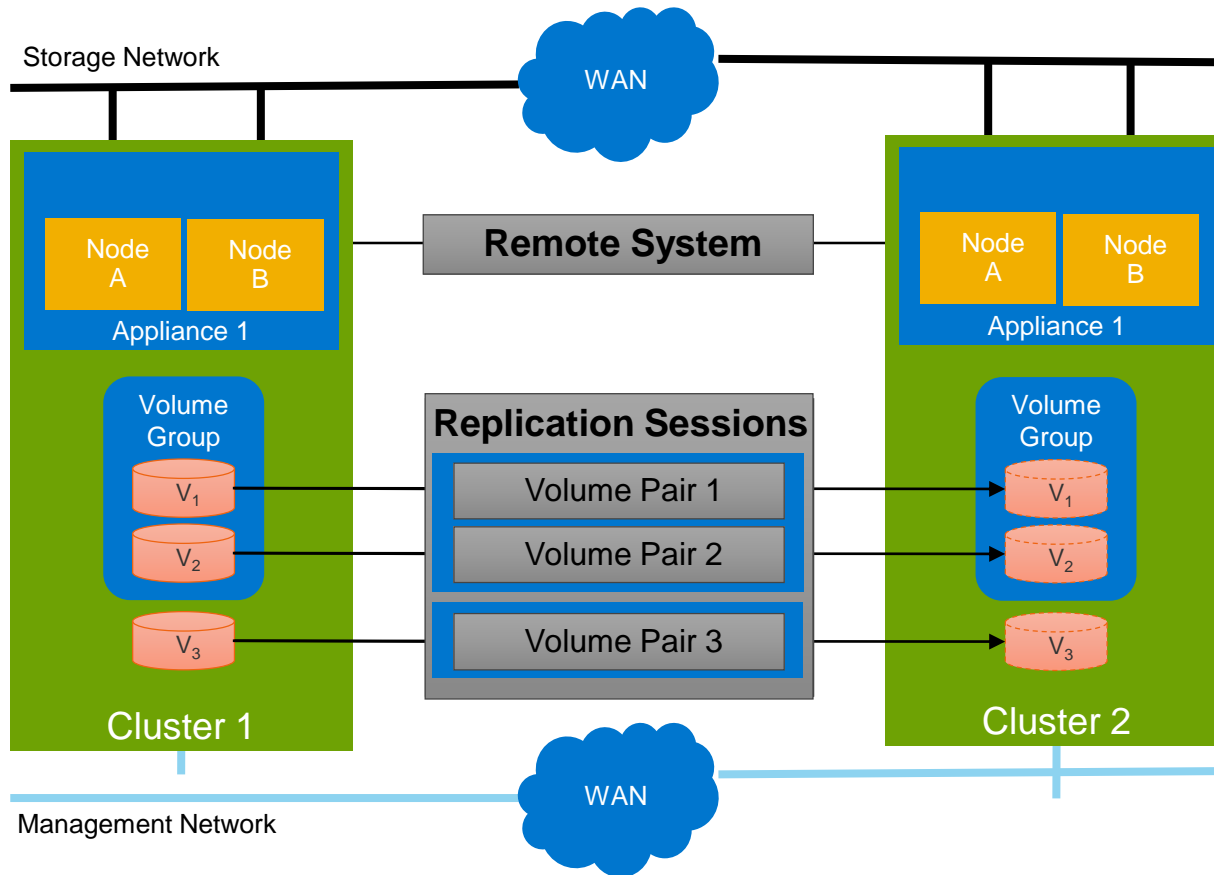
Overview

Supported Replication Topologies with PowerStore platform

- **One-to-One (1:1)**
 - Uni Directional
 - One or multiple Volumes are replicated from one cluster to one other cluster
 - Bi-Directional
 - Various Volumes are replicated in different directions between two clusters
- **One-to-many (1:n)**
 - Different Volumes are replicating to multiple destination clusters
 - No Support for fan-out or cascaded/chain replication
- **Many-to-One (n:1)**
 - Individual Volumes from different cluster are replicating to a single cluster

Remote Replication

Diagram



DELLTechnologies