Dell Technologies

Technical White Paper

Dell EMC PowerStore: Best Practices Guide

Abstract

This document provides best practices for installing and configuring Dell EMC™ PowerStore™ for optimal performance and availability.

June 2021

H18241.3

Revisions

Date	Description
April 2020	Initial release: PowerStoreOS 1.0
June 2020	Terminology updates
September 2020	Terminology updates
June 2021	PowerStoreOS 2.0 updates

Acknowledgments

Author: Stephen Wright

Table of contents

Executive summary

This white paper provides best practices guidance for using Dell EMC™ PowerStore™ in a mixed-business environment. It focuses on optimizing system performance and availability, and maximizing usability of the automated storage features.

These guidelines are intended to cover most use cases. They are recommended by Dell Technologies™ but are not strictly required. Some exception cases are addressed in this guide. Less-common edge cases are not covered by these general guidelines and are addressed in use-case-specific white papers.

For questions about the applicability of these guidelines in a specific environment, contact your Dell Technologies representative to discuss the recommendations.

Audience

This document is intended for IT administrators, storage architects, partners, and Dell Technologies™ employees. This audience also includes any individuals who may evaluate, acquire, manage, operate, or design a Dell EMC networked storage environment using PowerStore systems.

1 Introduction

This document introduces specific configuration recommendations that enable optimal performance from Dell EMC PowerStore.

1.1 PowerStore overview

PowerStore achieves new levels of operational simplicity and agility. It uses a container-based microservices architecture, advanced storage technologies, and integrated machine learning to unlock the power of your data. PowerStore is a versatile platform with a performance-centric design that delivers multidimensional scale, always-on data reduction, and support for next-generation media.

PowerStore brings the simplicity of public cloud to on-premises infrastructure, streamlining operations with an integrated machine-learning engine and seamless automation. It also offers predictive analytics to easily monitor, analyze, and troubleshoot the environment. PowerStore is highly adaptable, providing the flexibility to host specialized workloads directly on the appliance and modernize infrastructure without disruption. It also offers investment protection through flexible payment solutions and data-in-place upgrades.

1.2 Terminology

The following terms are used with PowerStore.

Appliance: Term used for solution containing a base enclosure and any attached expansion shelves. The size of an appliance could be only the base enclosure or the base enclosure plus expansion shelves.

Base enclosure: Used to reference the enclosure containing both nodes (node A and node B) and 25 NVMe drive slots.

Cluster: One or more appliances in a single grouping and management interface. Clusters are expandable by adding more appliances to the existing cluster, up to the allowed amount for a cluster.

Embedded module: Connectivity card in the PowerStore node that provides ports for Ethernet connections, and various service and management ports.

Expansion enclosure: Enclosures that can be attached to a base enclosure to provide additional storage in the form of SAS drives.

Fibre Channel: A protocol used to perform NVMe or SCSI commands over a Fibre Channel (FC) network.

File system: A storage resource that can be accessed through file-sharing protocols such as SMB or NFS.

Internet SCSI (iSCSI): Provides a mechanism for accessing block-level data storage over network connections.

I/O module: Optional connectivity cards that provide additional Fibre Channel or Ethernet ports.

IOPS: I/Os per second, a measure of transactional performance for small-block workloads.

MBPS: Megabytes per second, a measure of bandwidth performance for large-block workloads.

Network-attached storage (NAS) server: A virtualized network-attached storage server that uses the SMB, NFS, or FTP/SFTP protocols to catalog, organize, and transfer files within file system shares and exports. A

NAS server, the basis for multi-tenancy, must be created before creating file-level storage resources. NAS servers are responsible for the configuration parameters on the set of file systems that it serves.

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

Node: A storage node that provides the processing resources for performing storage operations and servicing I/O between storage and hosts.

PowerStore Command Line Interface (PSTCLI): An interface that allows a user to perform tasks on the storage system by typing commands instead of using the user interface (UI).

PowerStore T model: Container-based storage system that is running on purpose-built hardware. This storage system supports unified (block and file) workloads, or block optimized workloads.

PowerStore X model: Container-based storage system that is running inside a virtual machine that is deployed on a VMware hypervisor. In addition to the block optimized workloads that this storage system offers, it also allows users to deploy applications directly on the array, through AppsON functionality.

Server Message Block (SMB): An access protocol that allows remote file data access from clients to hosts on a network. This is typically used in Microsoft® Windows® environments.

Snapshot: A point-in-time view of data stored on a storage resource. A user can recover files from a snapshot, restore a storage resource from a snapshot, or provide access to a host.

Thin clone: A read/write copy of a volume, volume group, file system, or snapshot that shares blocks with the parent resource.

Virtual Volumes (vVols): A VMware® storage framework which allows VM data to be stored on individual VMware vSphere® Virtual Volumes™ (vVols). This allows for data services to be applied at a VM-granularity level while using Storage Policy Based Management (SPBM).

Volume: A block-level storage device that can be shared out using a protocol such as iSCSI or Fibre Channel.

2 Hardware considerations

At the highest level, design for optimal performance follows a few simple rules. The main principles of designing a PowerStore system for performance are as follows:

- Distribute workloads across available resources.
- Simplify the configuration.
- Design for resilience.
- Maintain the latest-released Dell EMC PowerStoreOS version.

Hardware components are the foundation of any storage system. This section discusses some key hardware differences between PowerStore models that help determine performance, and also explains how different configuration options can result in different performance from the same hardware.

2.1 PowerStore deployment modes

The PowerStore platform consists of six different models, from the PowerStore 500 model through the PowerStore 9000 model. All models use a common base enclosure and I/O modules. The models differ by CPU core count and speed, memory size, and number of NVMe NVRAM drives. These hardware differences give each model a unique performance profile.

Besides the hardware differences between the models, PowerStore can be installed in one of three different deployment modes. Each deployment mode has different capabilities, as detailed in Table 1. Choose the deployment mode that provides the required capabilities.

Table 1 PowerStore configurations

Deployment mode	External block access	External file access	AppsON functionality
PowerStore T model: Unified	✓	✓	X
PowerStore T model: Block optimized	✓	X	X
PowerStore X model	✓	X	✓

Note that the PowerStore 500 is only available as a T model (either unified or block optimized).

The PowerStore system has different performance characteristics depending on deployment mode.

2.1.1 PowerStore T models

PowerStore T models run the PowerStoreOS directly on the hardware. PowerStore T models can be installed in a unified configuration which provides file and block access, or in a block optimized configuration that provides only block access.

2.1.1.1 PowerStore T models (unified)

PowerStore T models with unified modes can provide access to block and file storage resources simultaneously. This is the default deployment mode.

2.1.1.2 PowerStore T models (block optimized)

If the PowerStore T model appliance will not be used for file access, it can be installed in block optimized mode, which disables the NAS capabilities. This mode can increase the amount of block workload that the system can provide, since it can devote the additional CPU and memory that is no longer needed for file.

2.1.2 PowerStore X models

PowerStore X models run the PowerStoreOS as a virtual machine on ESXi hypervisor. This configuration allows the PowerStore X model appliance to service external host I/O, and run guest VMs directly on the PowerStore hardware. PowerStore X models reserve a portion of the CPU and memory to be used for hosting user VMs. Therefore, fewer resources are available for serving external storage. The relative performance for storage from a PowerStore X model is expected to be less than the same PowerStore T model.

2.1.3 Relative performance expectations

This section gives a relative comparison of the performance potential of the different PowerStore models when serving external workload. Performance scales based on the specific hardware complement of the model, and is also impacted by the configuration type.

In general, the IOPS capability of the PowerStore models scales linearly from PowerStore 500 up to 9000 models. As mentioned previously, deployment mode also impacts performance capability. A PowerStore T model in block optimized mode can deliver more block IOPS than the same model in unified mode. A PowerStore X model has less capability for block IOPS, since some of the compute resources are reserved for running VMs.

With the exception of PowerStore 500, PowerStore systems use NVMe NVRAM drives to provide persistent storage for cached write data. PowerStore 1000 and 3000 model arrays have two NVRAM drives per system, while PowerStore 5000, 7000, and 9000 model arrays have four NVRAM drives per system. The extra drives mean that these systems can provide higher MBPS for large-block write workloads.

2.1.4 PowerStore cluster

PowerStore systems can be clustered. A PowerStore cluster combines multiple PowerStore appliances into a single grouping that is managed as a single storage system. A PowerStore cluster delivers aggregate performance from all appliances in the cluster, but a single volume is serviced by only one appliance at any given time. While not required, it is recommended that all appliances in a cluster be of the same model and have similar physical capacities to provide consistent performance across the cluster. A cluster must contain only Power Store T appliances, or only PowerStore X appliances; they cannot be mixed within a single cluster.

Volumes can be migrated between appliances in a cluster. It is recommended that any host that is connected to a PowerStore cluster has equivalent connectivity to all appliances in the cluster. All appliances in a cluster should be physically located in the same data center, and must be connected to the same LAN.

Clustering is applicable to block storage resources only. While a PowerStore T model in Unified mode can serve as the cluster's primary appliance, the file resources cannot migrate to a different appliance. When deploying multiple appliances for file, plan to have multiple clusters.

2.2 Drive configuration

PowerStore can be configured with NVMe solid state device (SSD) or NVMe storage class memory (SCM) drives for user data. SSD-based systems (except for PowerStore 500) can also be expanded with SAS SSD drives to increase the amount of available storage capacity. It is recommended that all drives within a PowerStore system be the same size, which can maximize the usable capacity from each drive.

PowerStore Dynamic Resiliency Engine (DRE) is used to manage the drives in the system. All drives are automatically used to provide storage capacity. DRE groups the drives into resiliency sets to protect against drive failure. User configuration of the drives is not necessary, and dedicated hot spare drives are not required in PowerStore. Spare space for rebuilds is automatically distributed across all drives within each resiliency set. This configuration provides better resource utilization, and enables faster rebuilds in case of a drive failure.

At initial installation of the PowerStore system, DRE can be configured with either single or double drive failure tolerance. To provide the greatest usable capacity from the same number of drives, it is recommended to initially install PowerStore with a minimum of ten drives for single drive failure tolerance, or nineteen drives for double drive failure tolerance.

2.2.1 SCM drives

PowerStore can utilize SCM drives either by having only SCM drives installed in the system, or, with PowerStoreOS 2.0, by mixing SCM and SSD drives.

Systems with all SCM drives are recommended for small-block workloads that require the absolute lowest latencies. A system with all SCM drives will place both data and metadata on the SCM drives.

Systems with mixed SSD and SCM drives will use the SCM drives for metadata acceleration; the SCM drives will store metadata for faster lookups. This can reduce latency on read operations in systems with large physical capacities.

3 Network considerations

External hosts send and receive data from PowerStore through Fibre Channel, Ethernet, or both networks. These networks play a large role in determining the performance potential of PowerStore. This section discusses considerations for the external network, and for the PowerStore network ports.

Host networking recommendations are covered in section 6.

3.1 General network performance and high availability

It is recommended to use redundant switch hardware between the PowerStore system and external clients. For more details about configuring a redundant network, see the document *Dell EMC PowerStore Host Configuration Guide* on the PowerStore Info Hub.

3.1.1 Fibre Channel fabrics

For Fibre Channel connectivity, configure dual redundant fabrics, with each PowerStore node and each external host having connectivity on each of the fabrics. Minimize the number of hops between host and PowerStore.

For performance, load balancing, and redundancy, each host should have at least two paths to each PowerStore node (four ports per PowerStore appliance).

3.1.2 Ethernet networks

For Ethernet connectivity, use multiple switches that are connected with Virtual Link Trunking interconnect (VLTi) and Link Aggregation Control Protocol (LACP) or equivalent technologies. Each PowerStore node should have connectivity to all linked switches.

For PowerStore T models, the first two ports of the embedded module 4-port card on each PowerStore node are bonded together within the PowerStoreOS. For highest performance and availability from these ports, it is recommended to also configure link aggregation across the corresponding switch ports. LACP is not applicable to PowerStore X models.

3.2 PowerStore front-end ports

PowerStore supports Ethernet connectivity through ports on the embedded module, and on optional I/O modules. PowerStore supports Fibre Channel connectivity through ports on optional I/O modules.

3.2.1 PowerStore Fibre Channel ports

PowerStore Fibre Channel ports support speeds for 32 Gb/s, 16 Gb/s, 8 Gb/s, and 4 Gb/s. This speed depends on the SFP used and the switchport or HBA that is connected. Higher speeds allow for greater MBPS and IOPS capabilities, so it is recommended to use the highest speed supported by the environment.

Fibre Channel ports are available on I/O modules that are inserted into I/O module slots on the nodes. The Fibre Channel I/O module is 16-lane PCIe Gen3. I/O module slot 0 is also 16-lane, while I/O module slot 1 is 8-lane. When a Fibre Channel I/O module is being installed, it is recommended to always use I/O module slot 0 first. If Fibre Channel I/O modules are installed in both I/O module slots, it is recommended to cable the ports in I/O module slot 0 first, due to the PCIe difference. The PCIe lanes in I/O module slot 1 are only a limiting factor for total MBPS, and only when all four ports on the Fibre Channel I/O module are operating at 32 Gb/s.

3.2.1.1 NVMe over Fibre Channel

The NVMe over Fibre Channel protocol provides connectivity using the same Fibre Channel ports, but can decrease the transport latency between PowerStore and the host. Note that all parts of the network, including switches and HBAs, must also support the NVMe over Fibre Channel protocol, and must run at 32Gb/s.

3.2.2 PowerStore Ethernet ports

PowerStore optical Ethernet ports support speeds up to 25 Gb/s, based on the SFP that is used. Copper Ethernet ports support speeds of up to 10 Gb/s. Higher speeds allow for greater MBPS and IOPS capabilities, so it is recommended to use the highest speed supported by your environment.

Jumbo frames (MTU 9000) is recommended for increased network efficiency. Jumbo frames must be supported on all parts of the network between PowerStore and the host.

3.2.2.1 Ethernet ports for iSCSI

Map additional Ethernet ports for iSCSI to increase system MBPS capabilities. Enable Jumbo frames for iSCSI by setting the Cluster MTU to 9000, and setting the storage network MTU to 9000.

The embedded module 4-port card and the optional network I/O modules are 8-lane PCIe Gen3. When more than two 25 GbE ports are used, these cards are oversubscribed for MBPS. To maximize MBPS scaling in the system, consider cabling and mapping the first two ports of all cards in the system first. Then, cable and map other ports as needed.

When PowerStore T models that are in unified mode are used for both iSCSI and file access, it is recommended to reserve the bonded ports for only file access. It is also recommended to log in host iSCSI initiators to iSCSI targets on the other mapped ports.

3.2.2.2 Ethernet ports for NAS

PowerStore NAS servers create their network interfaces on the two bonded ports on the embedded module 4-port card. For highest performance and availability from these ports, it is recommended to configure link aggregation across the corresponding switch ports. Enable Jumbo frames for NAS by setting the cluster MTU to 9000.

If the PowerStore is also providing block access through iSCSI, it is recommended to reserve the bonded ports for file access only, and map other physical network ports for iSCSI; logout iSCSI hosts from the initiators on the bonded ports. If both file and replication are used, tag other ports for replication to avoid contention with file traffic on the bonded ports.

4 PowerStore storage resources

PowerStore allows access to storage through block, file, or both protocols. This section provides recommendations for the different types of storage resources.

4.1 Block storage resources

Block storage resources are accessed through Fibre Channel, Fibre Channel over NVMe, or iSCSI protocols. A host should only access a block resource using one of these protocols. It is not recommended for the same host to access the same block resource using more than one protocol.

4.1.1 Appliance balance for block workloads

PowerStore block storage resources are accessed using ALUA active/optimized or active/non-optimized paths between the host and the two nodes within the PowerStore appliance. I/O is normally sent on an active/optimized path. PowerStore automatically chooses one of the nodes for the active/optimized path when the volume is mapped to the host to maintain a balanced workload across the nodes. This characteristic is called **node affinity**, and can be viewed and modified with PSTCLI or REST. These changes take effect immediately, and are nondisruptive if the host is correctly configured for multipathing.

4.1.2 Performance policy

All block storage resources in a PowerStore system have a defined performance policy. By default, this policy is set to Medium. The performance policy does not have any impact on system behavior unless some volumes have been set to Low Performance Policy, and other volumes are set to Medium or High. During times of system resource contention, PowerStore devotes fewer compute resources to volumes with Low Performance Policy. Reserve the Low policy for volumes that have less-critical performance needs.

4.2 File storage resources

File storage resources are accessed through NAS protocols, such as NFS and SMB. A NAS server can provide access to a file system using all NAS protocols simultaneously, if configured for multiprotocol access.

4.2.1 Appliance balance for file workloads

A single NAS server uses compute resources from only one node of the appliance. It is recommended to create at least two NAS servers (one on each node) so that resources from both nodes contribute to file performance.

NAS servers can be manually moved from one node to the other. This action can be done to balance the workload if one node is busier than the other. All file systems that are served by a given NAS server move with the NAS server to the other node.

5 PowerStore features and layered applications

This section discusses the features and layered applications available with PowerStoreOS.

5.1 Data reduction

PowerStore provides data-reduction capabilities such as zero-detect, compression, and deduplication. Data reduction is integrated into the PowerStore architecture, and is always active. During periods of high write activity, PowerStore may defer the deduplication of data, and devote those resources to servicing the client workload. During periods of low activity, PowerStore will use excess resources to re-examine any data written during these periods for duplicates, in order to regain any space savings that were not initially realized.

5.2 Snapshots and thin clones

All storage resources in PowerStore are thinly provisioned and space efficient, including snapshots and thin clones. Creation of a snapshot or thin clone requires only a quick duplication of pointers. After this action, they behave as independent storage resources and do not impact performance of the source resource.

5.3 AppsON functionality for PowerStore X models

When PowerStore X models are used with AppsON (hosting VMs on PowerStore), other configuration settings are recommended to provide optimal performance. These configurations include creating additional internal iSCSI targets, increasing internal iSCSI queue depths, and enabling Jumbo frames. These changes can be applied as part of the Initial Configuration Wizard, or manually. These changes should be applied before provisioning any storage resources. For detailed configuration steps, see the KnowledgeBase article HOW17288 or the document *Dell EMC PowerStore: Virtualization Integration* on Dell.com/StorageResources.

5.4 Cluster migrations

Block storage resources which are migrating between appliances in a cluster may see impacted performance from the migration activity. It is recommended to run migrations at a time when the resource is less busy.

5.5 PowerStoreOS upgrades

New versions of the PowerStoreOS are applied using a nondisruptive upgrade process. Since half of the system hardware resources are unavailable during parts of the upgrade, it is recommended to perform upgrades during planned maintenance windows. Alternately, perform upgrades when the system is less busy to minimize the impact to clients.

6 External host considerations

This section highlights host configuration changes that may be necessary to efficiently access PowerStore volumes. For details about the appropriate settings for a host type, see the *Dell EMC PowerStore Host Configuration Guide* on the PowerStore Info Hub.

6.1 Host configuration

Since PowerStore is a new class of storage, host operating systems may not natively recognize PowerStore volumes and apply the appropriate settings. For optimal performance, check that the appropriate configuration changes have been applied to all hosts that are connected to a PowerStore. The *Dell EMC PowerStore Host Configuration Guide* has recommendations for the following:

- MPIO settings: Path checker and timeout values
- iSCSI settings: Timeout and queue depth values; disabling delayed ACK
- Fibre Channel settings: Queue depth values
- Network settings: Jumbo frames and flow control
- Unmap operations
- VMware ESXi™ claim rules

For other recommended configurations for VMware ESXi and vSphere, see the document *Dell EMC PowerStore: Virtualization Integration* on Dell.com/StorageResources.

6.2 Host file systems

When a host is attached to a PowerStore block volume, the host can use this volume as a raw device, or it can create a local file system on the volume first. When a local file system is being created, it is recommended to disable SCSI unmap. When PowerStore creates a volume, all space is already unmapped; the host-based unmap is redundant, and generates unnecessary load on PowerStore. See the document *Dell EMC PowerStore Host Configuration Guide* on the PowerStore Info Hub for commands to disable unmap on your host operating system.

When creating a local file system, it is recommended to use a file system block size (allocation unit) of 4 KB, or a larger size that is an even multiple of 4 KB.

It is typically not necessary to perform alignment when creating a local file system. If alignment is performed, it is recommended to use an offset of 1 MB.

6.3 VMware

PowerStore is tightly integrated with VMware applications.

For other recommended configurations for VMware ESXi and vSphere, see the document *Dell EMC PowerStore: Virtualization Integration* on Dell.com/StorageResources.

6.4 Application considerations

PowerStore is well integrated with the most widely used enterprise applications. For best practice recommendations for specific applications, see the solutions-focused white papers available on Dell.com/StorageResources.

7 Conclusion

This best practices guide provides configuration and usage recommendations for PowerStore in general use cases. For a detailed discussion of the reasoning or methodology behind these recommendations, or for additional guidance around more specific use cases, see the documents that are listed in appendix A, or contact your Dell Technologies representative.

A Technical support and resources

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.

The PowerStore Info Hub provides detailed documentation on how to install, configure, and manage Dell EMC PowerStore systems.