

# EMC VNX UNIFIED BEST PRACTICES FOR PERFORMANCE

Applied Best Practices Guide

VNX OE for Block 5.33

VNX OE for File 8.1

## EMC Enterprise & Mid-range Systems Division

### Abstract

This applied best practices guide provides recommended best practices for installing and configuring VNX™ unified systems for best performance.

September, 2014

Copyright © 2014 EMC Corporation. All rights reserved. Published in the USA.

Published September, 2014

EMC believes the information in this publication is accurate of its publication date. The information is subject to change without notice.

The information in this publication is provided as is. EMC Corporation 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 EMC software described in this publication requires an applicable software license.

EMC2, EMC, and the EMC logo are registered trademarks or trademarks of EMC Corporation in the United States and other countries. All other trademarks used herein are the property of their respective owners.

For the most up-to-date regulatory document for your product line, go to the technical documentation and advisories section on EMC Online Support.

## **EMC VNX Unified Best Practices for Performance**

### **Applied Best Practices Guide**

Part Number H10938.5

<b>Chapter 1</b>	<b>System Configuration.....</b>	<b>7</b>
Essential guidelines .....		8
Storage Processor cache .....		8
Physical placement of drives .....		8
Hot sparing .....		9
<b>Chapter 2</b>	<b>Storage Configuration .....</b>	<b>10</b>
General considerations .....		11
Drive type.....		11
RAID level.....		11
Rules of thumb.....		11
Calculating disk IOPS by RAID type.....		12
Classic RAID group considerations.....		13
Drive count .....		13
Large element size .....		13
Drive location selection.....		13
Storage pool considerations.....		14
Storage pool creation.....		14
Virtual LUN type .....		15
Virtual LUN ownership considerations.....		16
Storage tiers .....		16
Considerations for VNX OE for File.....		17
<b>Chapter 3</b>	<b>Data Services.....</b>	<b>19</b>
FAST VP.....		20
General .....		20
Pool capacity utilization .....		20
Data Relocation .....		20
Considerations for VNX OE for File.....		20
FAST Cache .....		21
General considerations .....		21

Enabling FAST Cache on a running system .....	21
Considerations for VNX OE for File .....	22
<b>Chapter 4 VNX OE for Block Specific Considerations .....</b>	<b>23</b>
Availability and connectivity .....	24
Host Connectivity .....	24
Storage group settings .....	24
Host file system alignment .....	25
VNX OE for Block features .....	25
Block replication .....	25
Block compression .....	26
Block deduplication .....	26
VNX OE for Block application tuning .....	29
VMware ESX Server with iSCSI Datastore .....	29
<b>Chapter 5 VNX OE for File Specific Conditions .....</b>	<b>30</b>
File system creation .....	31
Classic RAID group LUNs .....	31
Storage pool LUNs .....	32
VNX OE for File network and protocol considerations .....	32
Interfaces .....	32
NFS .....	32
VNX OE for File features .....	33
SnapSure .....	33
Replicator .....	33
Deduplication and compression .....	33
CAVA .....	33
VNX OE for File application tuning .....	34
Hypervisor / Database over NFS or SMB .....	34
Bandwidth-intensive applications .....	34
<b>Conclusion .....</b>	<b>35</b>

As part of an effort to improve and enhance the performance and capabilities of its product line, EMC from time to time releases revisions of its hardware and software. Therefore, some functions described in this guide may not be supported by all revisions of the software or hardware currently in use. For the most up-to-date information on product features, refer to your product release notes.

If a product does not function properly or does not function as described in this document, please contact your EMC representative.

**Note** This document was accurate as of the time of publication. However, as information is added, new versions of this document may be released to EMC Online Support. Check the website to ensure that you are using the latest version of this document.

## Purpose

The *Applied Best Practices Guide* delivers straightforward guidance to the majority of customers using the storage system in a mixed business environment. The focus is on system performance and maximizing the ease of use of the automated storage features, while avoiding mismatches of technology. Some exception cases are addressed in this guide; however, less commonly encountered edge cases are not covered by general guidelines and are addressed in use-case-specific white papers.

Guidelines can and will be broken, appropriately, owing to differing circumstances or requirements. Guidelines must adapt to:

- Different sensitivities toward data integrity
- Different economic sensitivities
- Different problem sets

The guidelines contain a few “don’ts”:

- DON’T means: Do not do it; there is some pathological behavior.
- AVOID means: All else being equal, it is generally best not to, but it still is okay to do it.

## Audience

This document is intended for EMC customers, partners, and employees who are installing and/or configuring VNX unified systems. Some familiarity with EMC unified storage systems is assumed.

## Related documents

The following documents provide additional, relevant information. Your logon credentials determine your access to these documents. Find all of the documents <http://support.emc.com>. If you do not have access to the following content, contact your EMC representative.

- Introduction to the New EMC VNX Series VNX5200, VNX5400, VNX5600, VNX5800, VNX7600, and VNX8000 – A Detailed Review – White Paper
- VNX MCx Multicore Everything – White Paper
- EMC VNX Multicore FAST Cache VNX5200, VNX5400, VNX5600, VNX5800, VNX7600, and VNX8000 – A Detailed Review – White Paper
- Introduction to EMC VNX Storage Efficiency Technologies VNX5200, VNX5400, VNX5600, VNX5800, VNX7600, and VNX8000 – White Paper
- EMC VNX FAST VP – White Paper
- Virtual Provisioning for the New VNX Series – White Paper
- EMC VNX Deduplication and Compression VNX5200, VNX5400, VNX5600, VNX5800, VNX7600, and VNX8000 – Maximizing Effective Capacity Utilization – White Paper
- EMC VNX Replication Technologies – An Overview – White Paper
- EMC VNX Snapshots – White Paper
- MirrorView KnowledgeBook: Release 30-33 - A Detailed Review – White Paper
- Celerra IP Replicator V2 Best Practices for Performance – Tech Note
- EMC CLARiiON Reserved LUN Pool Configuration Considerations - Best Practices Planning
- EMC SAN Copy – A Detailed Review – White Paper
- Microsoft Exchange 2010: Storage Best Practices and Design Guidance for EMC Storage – White Paper
- EMC Performance for Oracle - EMC VNX, Enterprise Flash Drives, FAST Cache, VMware vSphere – White Paper
- Using EMC VNX Storage with VMware vSphere – TechBook
- EMC Host Connectivity Guides
- Windows/Linux/Solaris/HP-UX/AIX/Tru64/VMware

# Chapter 1 System Configuration

This chapter presents the following topics:

Essential guidelines.....	8
Storage Processor cache .....	8
Physical placement of drives.....	8
Hot sparing .....	9

## Essential guidelines

This guide introduces specific configuration recommendations that enable good performance from a VNX Unified storage system. At the highest level, good performance design follows a few simple rules. The main principles of designing a storage system for performance are:

- Flash First – Utilize flash storage for the active dataset to achieve maximum performance.
- Distribute load over available hardware resources.
- Design for 70 percent utilization (activity level) for hardware resources.
- When utilizing Hard Disk Drives (HDD), AVOID mixing response-time-sensitive I/O with large-block I/O or high-load sequential I/O.
- Maintain latest released VNX OE code versions.

## Storage Processor cache

Storage Processor memory configuration is not required. Memory allocation amounts, cache page size, and cache watermarks, are not configurable parameters.

## Physical placement of drives

When initially placing drives in the array:

- Spread flash drives across all available buses.
- There is no restriction around using or spanning across Bus 0, Enclosure 0.



## Hot sparing

Hot sparing is the process of rebuilding a failed drive's data onto a system selected compatible drive. Beginning with VNX OE for Block 5.33, any unbound drive can be considered for sparing. When planning Hot Spares, consider the following recommendations:

- Plan one additional drive for every 30 provisioned drives
  - Verify count in the CLI or GUI Hot Spare Policy tab in System->Hardware and modify if desired by choosing a Custom Policy
    - **naviseccli hotsparepolicy -list**
- Distribute unbound drives (spares) across available buses.
- Ensure that unbound drives for each drive type are available.
  - SAS Flash / FAST Cache SSD (SLC) must spare for SAS Flash / FAST Cache SSD (SLC).
  - SAS Flash VP / FAST VP SSD (eMLC) must spare for SAS Flash VP / FAST VP SSD (eMLC).
  - SAS must spare for SAS (regardless of rotational speed).
  - NL-SAS must spare for NL-SAS.

Ensure the unbound drive capacity is equal to or larger than the provisioned drives.

## Chapter 2 Storage Configuration

This chapter presents the following topics:

General considerations .....	11
Classic RAID group considerations .....	13
Storage pool considerations .....	14

## General considerations

### Drive type

Match the appropriate drive type to the expected workload:

- SAS Flash / FAST Cache SSD (SLC) for extreme performance; best performance for transactional random workloads; lowest write service time. Required for FAST Cache.
- SAS Flash VP / FAST VP SSD (eMLC) for extreme performance FAST VP tier. These are a higher capacity Flash option; not for use with FAST Cache.
  - Use all of the same SSD technology for the same extreme performance tier.
- SAS for the general performance tier.
- NL-SAS for less active data, well-behaved streaming data, aging data, archive purposes, and backups.

### RAID level

For best performance from the least number of drives, match the appropriate RAID level with the expected workload:

- RAID 1/0 works best for heavy transactional workloads with high (greater than 30 percent) random writes, in a pool or classic RAID group with primarily HDDs.
- RAID 5 works best for medium to high performance, general-purpose and sequential workloads.
- RAID 6 for NL-SAS works best with read-biased workloads such as archiving and backup to disk. It provides additional RAID protection to endure longer rebuild times of large drives.

### Rules of thumb

Disk drives are a critical element of unified performance. Use the rule of thumb information to determine the number of drives to use to support the expected workload.

- These guidelines are a conservative starting point for sizing, not the absolute maximums.

Rules of thumb for drive throughput (IOPS):

- IOPS assumes small block random with good response time.
  - Drives are capable of a sustained IOPS workload based on drive type, as follows:

Throughput	NL-SAS	SAS 10K	SAS 15K	SAS Flash VP (eMLC)	SAS Flash (SLC)
IOPS	90	150	180	3500	5000

- To size IOPS with spinning drives, VNX models scale linearly with additional drives, up to the maximum drive count per model.
- To size for host IOPS, include the RAID overhead as described in the next section [Calculating disk IOPS by RAID type](#).

Rules of thumb for drive bandwidth (MB/s):

- Bandwidth assumes multiple large-block sequential streams.

Bandwidth	NL-SAS	SAS 10K	SAS 15K	Flash (All)
Read MB/s	15	25	30	90
Write MB/s	10	20	25	75

- Parity does not count towards host write bandwidth sizing.
  - For example; a 4+1 RAID group using SAS 15K to support a sequential write workload is sized at 100 MB/s write for the entire RAID group.
- To size bandwidth with HDD as RAID 5 or RAID 6, VNX models scale up to the sweet spot drive count, which varies by model as follows:

Bandwidth	VNX5200	VNX5400	VNX5600	VNX5800	VNX7600	VNX8000
Read Drive Count	100	200	300	450	500	1000
Write Drive Count	100	120	160	200	225	350

- Drive counts include parity.
- Systems can hit maximum bandwidth with fewer drives when the per-drive performance is exceeding the rule of thumb sizing guidelines.

## Calculating disk IOPS by RAID type

Block front-end application workload is translated into a different back-end disk workload based on the RAID type in use.

For reads (no impact of RAID type):

- 1 application read I/O = 1 back-end read I/O

For random writes:

- RAID 1/0 - 1 application write I/O = 2 back-end write I/O
- RAID 5 - 1 application write I/O = 4 back-end disk I/O (2 read I/O + 2 write I/O)
- RAID 6 - 1 application write I/O = 6 back-end disk I/O (3 read I/O + 3 write I/O)

## Classic RAID group considerations

### Drive count

RAID groups can have a maximum count of 16 drives. For parity RAID, the higher count offers higher capacity utilization, but with higher risk to availability.

Sequential optimization occurs when the array can perform full-stripe operations normally associated with large-block I/O. Certain drive counts are more likely to enable this behavior, though these optimizations can also occur with non-preferred drive counts.

With a predominance of large-block sequential operations, the following applies:

- RAID 5 has a preference of 4+1 or 8+1.
- RAID 6 has a preference of 8+2.
- RAID 1/0 has a preference of 4+4.

For predominantly random workloads, use the drive type rule of thumb values and the RAID level IOPS calculations to determine the number of drives needed to meet the expected application workload.

Storage pools have different recommended preferred drive counts as described in the section on [Virtual LUN ownership considerations](#).

### Large element size

When creating a 4+1 RAID 5 RAID group, you can select a 1024 block (512KB) element size.

- Use large element size when the predominant workload is large-block random read activity (such as data warehousing).
  - AVOID using large element size with any other workloads.

### Drive location selection

When selecting the drives to use in a RAID group, drives can be selected either from a DAE or DAEs on the same bus (horizontal positioning), or from multiple DAEs on different buses (vertical positioning). There are now almost no performance or availability advantages to using vertical positioning. Therefore:

- Use the default horizontal positioning method of drive selection when creating RAID groups.
- If a single DAE does not have enough available drives to create the RAID group, selecting drives from a second DAE is acceptable.
  - The SAS buses are fast enough to support drives distributed across enclosures belonging to the same RAID group.

## Storage pool considerations

### Storage pool creation

Create multiple pools to:

- Separate workloads with different I/O profiles.
  - Place predominantly sequential workloads in dedicated pools or RAID groups.
- Dedicate resources when you have specific performance goals.
- Vary pool parameters, such as FAST Cache enabled/disabled
- Minimize failure domains.
  - Although unlikely, loss of a private RAID group in the pool compromises the total capacity of that pool. You can create multiple smaller pools instead of allocating the available capacity in a single pool.

Consider limiting the pool size to the maximum number of drives for initial placement in a pool at creation for a given model.

Model	VNX5200	VNX5400	VNX5600	VNX5800	VNX7600	VNX8000
Maximum drives in a pool at creation	80	80	120	120	120	180
Maximum drives in a single pool	121	246	496	746	996	1496

Tiered storage pools have multiple RAID options per tier for preferred type and drive count.

- Use RAID 5 with a preferred count of 4+1 for the best performance versus capacity balance.
  - Using 8+1 improves capacity utilization at the expense of slightly lower availability.
- Use RAID 6 for NL-SAS tier.
  - Options for preferred drive count are 6+2 and 14+2 where 14+2 provides the highest capacity utilization option for a pool, at the expense of slightly lower availability.
- Consider the following rule of thumb for tier construction:
  - Extreme performance flash tier: 4+1 RAID 5
  - Performance SAS tier: 4+1 or 8+1 RAID 5
  - Capacity NL-SAS tier: 6+2 or 14+2 RAID 6

- AVOID a drive count that results in a small remainder when divided by the preferred drive count specified.
  - Maintain a multiple of the preferred drive count for each tier selected (such as 5, 10, 15, and so on, when using RAID 5 4+1).
- When expanding pools, use a multiple of the preferred modulus drive count for the tier you are expanding
  - Anything other than preferred modulus results in non-standard private RAID Groups that can be less than optimal for balanced performance.
    - For example: 10 drives in a RAID 1/0 configuration, where the preferred modulus is 8, results in a 1+1 mirror that can lead to a hot spot within the tier.
  - An imbalance in the RAID group modulus has less impact on the Flash Tier due to the performance capabilities of individual Flash Drives.

## Virtual LUN type

Block storage pool LUNs can be created as either thick (fully allocated) or thin (virtually provisioned).

- Use thin LUNs when storage efficiency requirements outweigh performance requirements.
  - Thin LUNs maximize ease-of-use and capacity utilization at some cost to maximum performance.
  - A slower response time or fewer IOPS does not always occur; the potential of the drives to deliver IOPS to the host is less.
  - When using thin LUNs, adding a flash tier to the pool can improve performance.
    - Thin LUN metadata can be promoted to the flash tier with FAST VP enabled.
  - DON'T use thin LUNs with VNX OE for File; use only thick LUNs when using pool LUNs with VNX OE for File.
- Use thin LUNs when planning to implement VNX Snapshots, block compression, or block deduplication.
  - Ensure the same storage processor owns all LUNs in one pool if enabling Block Deduplication in the future. Refer to the [Block Deduplication section](#) for details.
- Use thick LUNs for the highest level of pool-based performance.
  - A thick LUN's performance is comparable to the performance of a classic LUN and is better than the performance of a thin LUN.
  - For applications requiring very consistent and predictable performance, use thick pool LUNs or Classic LUNs.

## Virtual LUN ownership considerations

AVOID changing the ownership of a pool LUN after initial creation.

- Use LUN migration to move LUNs to their peer SP if required to change ownership.

## Storage tiers

The following items influence the number of tiers required in a storage pool:

- Performance requirements
- Capacity requirements
- Knowledge of the I/O

The capacity required for each tier depends on expectations for locality of your active data. As a starting point, you can consider capacity per tier of 5 percent flash, 20 percent SAS, and 75 percent NL-SAS. This works on the assumption that less than 25 percent of the used capacity is active and infrequent relocations from the lowest tier occurs.

If you know your active capacity (skew), then you can be more systematic in designing the pool and apportion capacity per tier accordingly.

In summary, follow these guidelines:

- When FAST Cache is available, use a 2-tier pool comprised of SAS and NL-SAS and enable FAST Cache as a cost-effective way of realizing flash performance without the expense of flash dedicated to this pool.
  - You can add a flash tier later if FAST Cache is not capturing your active data.
- For a 3-tier pool, start with 5 percent flash, 20 percent SAS, and 75 percent NL-SAS for capacity per tier if the skew is unknown.
  - You can always expand a tier after initial deployment to effect a change in the capacity distribution.
- Use a 2-tier pool comprised of flash and SAS as an effective way of providing consistently good performance.
  - You can add NL-SAS later if capacity growth and aged data require it.
- AVOID using a flash+NL-SAS 2-tier pool if the skew is unknown.
  - The SAS tier provides a buffer for active data not captured in the flash tier, where you can still achieve modest performance, and quicker promotion to flash when relocations occur.
- Add a flash tier to a pool with thin LUNs so that metadata promotes to flash and overall performance improves.



For each pool LUN:

- Consider dedicated pools or classic RAID groups where there is predominantly sequential activity for specific LUNs.
- DON'T use auto-tier for low-skew random workloads where the active dataset does not fit in the highest tier.
  - Excessive tier relocations that do not benefit the active data can occur.
- AVOID using highest available when the LUN capacity exceeds 90% the highest tier capacity.
  - This can affect the overall efficiency of the highest tier to service active data for LUNs running in auto-tier mode.

## Considerations for VNX OE for File

When using Block storage pools with VNX OE for File, use the following recommendations:

- Create a separate pool for VNX OE for File LUNs.
  - AVOID mixing with Block workloads.
- Pre-provision space in the storage pool
  - Create and assign LUNs to VNX OE for File, so that VNX OE for File has available space for file system creation and extension, snapshots, and so on.
- Use only thick pool LUNs with VNX OE for File.
  - DON'T use thin LUNs with VNX OE for File.
    - If Virtual Provisioning™ is required for VNX OE for File, use a thin-enabled file system on classic or thick LUNs.
  - DON'T use compressed or block deduplicated LUNs with VNX OE for File.
    - If deduplication and compression are required for VNX OE for File, use VNX OE for File Deduplication and Compression.
    - Block Deduplication and Block Compression are not supported with VNX OE for File LUNs.
  - DON'T create VNX SnapShots on LUNs used by VNX OE for File.
    - If snapshots are required for VNX OE for File, use SnapSure.
- Apply the same tiering policies to all LUNs in the storage pool.
- Allow LUNs to complete the prepare process (thick LUN slice allocation) before adding to the File storage group. Use this command to display the status of the prepare:
  - **naviseccli lun -list -opDetails**

## Storage Configuration

When creating pool LUNs:

- Create approximately 1 LUN for every 4 drives in the storage pool.
- Create LUNs in even multiples of 10.
  - $\text{Number of LUNs} = (\text{number of drives in pool divided by } 4), \text{ rounded up to nearest multiple of } 10.$
- Make all LUNs the same size.
- Balance LUN ownership across SPA and SPB.
- Leave 5percent available in the pool for relocations and recovery operations.

## Chapter 3 Data Services

This chapter presents the following topic:

FAST VP	.....	20
FAST Cache	.....	21

## FAST VP

### General

- Use consistent drive technology for a tier within a single pool.
  - Same flash drive technology and drive size for the extreme performance tier.
  - Same SAS RPM and drive size for the performance tier.
  - Same NL-SAS drive size for the capacity tier.

### Pool capacity utilization

- Maintain some unallocated capacity within the pool to help with relocation schedules when using FAST VP.
  - Relocation attempts to reclaim 10 percent free per tier. This space optimizes relocation operations and helps when newly created LUNs want to use the higher tiers.

### Data Relocation

Relocation moves pool LUN data slices across tiers or within the same tier, to move hot data to higher performing drives or to balance underlying drive utilization. Relocation can occur as part of a FAST VP scheduled relocation, or due to pool expansion.

- Enable FAST VP on a pool, even if the pool only contains a single tier, to provide ongoing load balancing of LUNs across available drives based on slice temperature.
- Schedule relocations for off-hours, so that the primary workload does not contend with relocation activity.
- Enable FAST VP on a pool before expanding the pool with additional drives.
  - With FAST VP, slices rebalance according to slice temperature.

### Considerations for VNX OE for File

By default, a VNX OE for File system-defined storage pool is created for every VNX OE for Block storage pool that contains LUNs available to VNX OE for File. (This is a “mapped storage pool.”)

- All LUNs in a given VNX OE for File storage pool must have the same FAST VP tiering policy.
  - Create a user-defined storage pool to separate VNX OE for File LUNs from the same Block storage pool that have different tiering policies.

When using FAST VP with VNX OE for File, use thin enabled file systems for increased benefit from FAST VP multi-tiering.

## FAST Cache

FAST Cache is best for small random I/O where data has skew. The higher locality ensures better FAST Cache benefits.

### General considerations

Use the available flash drives for FAST Cache, which can globally benefit all LUNs in the storage system. Supplement the performance as needed with additional flash drives in storage pool tiers.

- Match the FAST Cache size to the size of active data set.
  - For existing EMC VNX/CX4 customers, EMC Pre-Sales have tools that can help determine active data set size.
- If active dataset size is unknown, size FAST Cache to be 5 percent of your capacity, or make up any shortfall with a flash tier within storage pools.
- Consider the ratio of FAST Cache drives to working drives. Although a small FAST Cache can satisfy a high IOPs requirement, large storage pool configurations distribute I/O across all pool resources. A large pool of HDDs can provide better performance than a few drives of FAST Cache.

Preferred application workloads for FAST Cache include:

- Small-block random I/O applications with high locality
- High frequency of access to the same data
- Systems where current performance is limited by HDD capability, not SP capability

AVOID enabling FAST Cache for LUNs not expected to benefit, such as when:

- The primary workload is sequential.
- The primary workload is large-block I/O.

AVOID enabling FAST Cache for LUNs where the workload is small-block sequential, including:

- Database logs
- Circular logs
- VNX OE for File SavVol (snapshot storage)

### Enabling FAST Cache on a running system

When adding FAST Cache to a running system, enable FAST Cache on a few LUNs at a time, and then wait until those LUNs have equalized in FAST Cache before adding more LUNs.

FAST Cache can improve overall system performance if the current bottleneck is drive-related, but boosting the IOPS results in greater CPU utilization on the SPs. Size systems to ensure the maximum sustained utilization is 70 percent. On an existing system, check the SP CPU utilization of the system, and then proceed as follows:

- Less than 60 percent SP CPU utilization – enable groups of LUNs or one pool at a time; let them equalize in the cache, and ensure that SP CPU utilization is still acceptable before turning on FAST Cache for more LUNs/pools.
- 60-80 percent SP CPU utilization – scale in carefully; enable FAST Cache on one or two LUNs at a time, and verify that SP CPU utilization does not go above 80 percent.
- CPU greater than 80 percent – DON'T activate FAST Cache.

AVOID enabling FAST Cache for a group of LUNs where the aggregate LUN capacity exceeds 20 times the total FAST Cache capacity.

- Enable FAST Cache on a subset of the LUNs first and allow them to equalize before adding the other LUNs.

**Note:** For storage pools, FAST Cache is a pool-wide feature so you have to enable/disable at the pool level (for all LUNs in the pool).

## Considerations for VNX OE for File

AVOID enabling FAST Cache for LUNs containing SavVol (snapshot storage).

- SavVol activity tends to be small block and sequential, characteristics that do not benefit from FAST Cache.

When enabling FAST Cache on a Block storage pool with VNX OE for File LUNs, use a separate storage pool or classic RAID groups for SavVol storage.

## Chapter 4 VNX OE for Block Specific Considerations

This chapter presents the following topics:

Availability and connectivity.....	24
VNX OE for Block features .....	25
VNX OE for Block application tuning .....	29

## Availability and connectivity

The VNX OE for Block storage array offers connectivity to a variety of client operating systems, and with multiple protocols, such as FC, FCoE, and iSCSI. EMC provides connectivity guides with detailed instructions for connecting and provisioning storage with different protocols to the specific host types.

Consult the EMC Host Connectivity Guide for the host type or types that connect to the VNX OE for Block array for any specific configuration options.

### Host Connectivity

Always check the appropriate Host Connectivity Guide for the latest best practices.

- Host connectivity guides cover more detail, especially for a particular operating system. Reference them for host-specific connectivity guidelines.
- Host guides include FC, iSCSI, and FCoE connectivity guidelines, where appropriate.

Balance front-end host port connections across Front End I/O modules, as host port connections affect the preferred CPU core assignment.

- Use the even numbered ports on each Front End I/O module before using the odd numbered ports when the number of Front End port assignments is less than the number of CPU Cores.
- Initially skip the first FC and/or iSCSI port of the array if those ports are configured and utilized as MirrorView connections.
- For the VNX8000, engage the CPU cores from both CPU sockets with Front End traffic.
  - Verify the distribution of Front End I/O modules to CPU sockets.
    - Balance the Front End I/O modules between slots 0-5 and slots 6-10.
    - DON'T remove I/O modules if they are not balanced. Instead, contact EMC support.
  - Balance host port assignments across UltraFlex slots 0-5 and 6-10.

AVOID zoning every host port to every SP port.

### Storage group settings

When registering host HBAs with VNX OE for Block, make sure to set the appropriate failover mode based on the host type. See the Host Connectivity Guides for details.



## Host file system alignment

File system alignment is covered in detail in the Host Connectivity Guides. In general:

- The more recent Linux environments, and the Windows Server 2008 and later, automatically align.
- VNX OE for File automatically aligns when it creates file systems.
- When provisioning LUNs for older Windows and Linux operating systems that use a 63-block header, manually align the host file system needs.

Alignment practices include:

- Use host-based methods to align the file system.
- Align the file system with a 1 MB offset.

## VNX OE for Block features

### Block replication

For MirrorView™:

- AVOID enabling FAST Cache on MirrorView secondary LUNs.
  - MV/S secondary LUNs replicate only writes from the source and serviced well by SP cache.
  - MV/A secondary LUNs replicate writes during updates and incur copy-on-first-write activity. This can incur additional FAST cache promotions that do not lead to performance gain.
- DON'T enable FAST Cache on Write Intent Log (WIL).
  - SP Cache algorithms achieve optimal performance for the WIL I/O pattern.

For SnapView™:

- Use SAS drives for reserve LUN pool (RLP) configurations, with write cache-enabled LUNs.
  - Match the secondary side RLP configuration to the primary side.
- AVOID configuring RLP on the same drives as the primary and secondary LUNs to avoid drive contention.
- DON'T enable FAST Cache on RLP LUNs.
  - RLP exhibits multiple small-block sequential streams that are not suited for FAST Cache.
- DON'T enable FAST Cache on clone private LUNs (CPL).
  - SP Cache algorithms achieve optimal performance for the CPL I/O pattern.

For RecoverPoint:

- DON'T enable FAST Cache for RecoverPoint journal LUNs.
  - Journals exhibit primarily large-block sequential activity not suited for FAST Cache use.

For VNX Snapshots:

- Start with thin LUNs if planning to use VNX snapshots.
  - Thick LUNs eventually convert to thin LUNs once a VNX snapshot is created on; all new writes and overwrites require thin LUNs.
- Plan the deletion of snapshots
  - Whenever possible, schedule the deletion of VNX Snapshots during non-peak hours of operation.
    - If deleting snapshots during peak periods of array activity, lessen the impact by reducing the number of concurrent snapshot deletes.
  - DON'T delete the last snapshot of a thick LUN, if you intend to create another snapshot immediately after deleting the last snapshot.
    - Create the new snapshot before deleting the older snapshot.
    - Deleting the last snapshot of a thick LUN reverses the thin conversion, which then reconverts for the new snapshot.
  - For additional technical information on VNX Snapshots, refer to *EMC VNX Snapshots* at EMC Online Support.

## Block compression

Start with thin LUNs if planning to use Block compression.

- Classic or thick LUNs convert to thin LUNs when compressed.

Manage the processes of Block Compression to minimize impact to other workloads.

- Pause or change the compression rate to Low at the system level when response-time critical applications are running on the storage system.
  - The option exists to pause compression at the system level to avoid compression overhead during spikes in activity.
  - Pausing the compression feature ensures background compression, or associated space reclamation operations, do not impact host I/O.

## Block deduplication

Block deduplication works best in storage environments which include multiple copies of the same data that is read-only and remains unchanged.

Evaluate the I/O profile of the workload and data content to determine if deduplication is an appropriate solution.

- AVOID using Block Deduplication on data that does not have a large amount of duplicated data. The added metadata and code path can impact I/O response time and outweigh any advantage seen by deduplication.
  - In instances where portions of data on a single LUN are a good fit for Block Deduplication while other portions are not, consider placing the data on different LUNs when possible.
- AVOID deploying deduplication into a high write profile environment. This action either causes a cycle of data splitting/data deduplication or data splitting inflicting unnecessary overhead of the first deduplication pass and all future I/O accesses.
  - Block Deduplication is best suited for datasets in which the workload consists of less than 30 percent writes.
    - Each write to a deduplicated 8 KB block causes a block for the new data to be allocated, and updates to the pointers to occur. With a large write workload, this overhead can be substantial.
- AVOID deploying deduplication into a predominantly bandwidth environment.
  - Avoid sequential and large block random (IOs 32 KB and larger) workloads when the target is a Block Deduplication enabled LUN.
    - As Block Deduplication can cause the data to become fragmented, sequential workloads require a large amount of overhead which can impact performance of the Storage Pool and the system.

Use FAST Cache and/or FAST VP with Deduplication

- Optimizes the disk access to the expanded set of metadata required for deduplication.
- Data blocks not previously considered for higher tier movement or promotion can now be hotter when deduplicated.

Start with Deduplicated Enabled thin LUNs if setting up to use Block Deduplication

- After enabling deduplication on an existing LUN, both thick and thin LUNs migrate to a new thin LUN in the Deduplication Container of the pool.
- Start with thin LUNs if planning to use Block Deduplication in the future.
  - Since the thick LUN migrates to a thin LUN when deduplication you enable, this reduces the performance change from thick to thin LUN.

Plan the creation of deduplicated LUNs to maximize and balance performance.

- For best performance, the same SP must own all deduplicated LUNs within the same pool.
  - All deduplicated LUNs in the same pool are placed in a single container and are managed by one SP.
- Balance the deduplication containers between the SPs.
  - If deduplication is enabled on one pool, and the deduplication container is owned by SPA, the next container must be created on SPB.

Manage the processes of Block Deduplication to minimize impact to other workloads.

- Pause or change the Deduplication rate to Low at the system level when response-time critical applications are running on the storage system.
  - The option exists to pause Deduplication at the system level to avoid Deduplication overhead during spikes in activity.
  - Pausing the Deduplication feature ensures background Deduplication, or associated space reclamation operations, do not impact host I/O.
    - Pausing the Deduplication feature also pauses any internal migrations initiated by enabling Deduplication on an existing LUN.
- Separate any data migrations and the Deduplication pass schedule as both can require a large amount of I/O and CPU resources.
  - If amount of raw data is larger than the Deduplication pool, cycle between data migrations and the Deduplication Pass.
  - Migrate classic LUNs to pool LUNs with deduplication already enabled to avoid an additional internal migration.
  - Enabling Deduplication on an existing pool LUN initiates a background LUN Migration at the rate of the Deduplication defined on the pool.
    - Changing the Deduplication rate to Low lowers both the Deduplication Pass and the internal migration reducing resource consumption on the array.
    - To speed up the migration into the deduplication container, pause deduplication on the pool/system and manually migrate to a pool LUN with deduplication already enabled selecting the desired LUN Migration rate (ASAP, High, Medium, or Low).
    - Pool LUN migrations can also initiate any associated space reclamation operations.

## VNX OE for Block application tuning

### VMware ESX Server with iSCSI Datastore

- Disable “Delayed Ack” for iSCSI storage adapters and targets.
  - For further detail, se VMware Knowledge Base article 1002598.

<http://kb.vmware.com/kb/1002598>

## Chapter 5 VNX OE for File Specific Conditions

This chapter presents the following topics:

File system creation .....	31
Classic RAID group LUNs .....	31
Storage pool LUNs .....	32
VNX OE for File network and protocol considerations .....	32
VNX OE for File features.....	33
VNX OE for File application tuning .....	34

## File system creation

Guidelines for creating file systems:

- Balance backup and restore window requirements with capacity requirements when deciding on file system sizes.
- Use Automatic Volume Management (AVM) to create file systems.
  - AVM creates well-designed file system layouts for most situations.
- Manual Volume Management can be used to create file systems with specific attributes for specialized workloads.
  - For metadata-heavy workloads, stripe across an odd number of disk volumes (dVols).
  - For sequential access on striped volumes:
    - AVOID wide striping across a large number of dVols.
    - Match VNX OE for File stripe size to the Block LUN full-stripe width for best sequential write performance.

## Classic RAID group LUNs

When creating LUNs for VNX OE for File from RAID groups, consider the following:

- Preferred RAID group sizes:
  - RAID 1/0: 1+1
  - RAID 5: 4+1, 8+1
  - RAID 6: 8+2, 4+2
- When creating volumes manually:
  - Stripe across 5 LUNs from different RAID groups.
  - Use a stripe size of 262144 (256KB).
  - Balance SP ownership.

## Storage pool LUNs

When creating LUNs for VNX OE for File from Block storage pools, consider the following:

- When creating volumes manually:
  - Stripe across 5 LUNs from the same pool.
  - Use a stripe size of 262144 (256KB).
  - Balance SP ownership.

For more detail, refer to [Storage pool considerations](#) and [Considerations for VNX OE for File](#).

## VNX OE for File network and protocol considerations

### Interfaces

10GbE:

- Best for bandwidth-sensitive applications.
- Configure the first NIC from different UltraFlex I/O modules before moving to the second NIC on the same UltraFlex I/O module.
- Use Jumbo frames.

Trunking and multipathing:

- Use LACP instead of EtherChannel.
- Configure Failsafe Networking (FSN) to avoid a network single point of failure.

### NFS

For bandwidth-sensitive applications:

- Increase the value of param **nfs v3xfersize** to **262144** (256KB).
- Negotiate 256KB NFS transfer size on client with:
  - `mount -o rsize=262144,wsiz=262144`



## VNX OE for File features

### SnapSure

If using SnapSure™ to create user snapshots of the primary file system:

- SnapSure sizing:
  - Size disk layout for PFS to include copy-on-first-write activity.
  - Size layout for SavVol based on expected user load to the snapshot file systems.
  - Include one additional read I/O for PFS, and one additional write I/O for SavVol, for every host write I/O.
- Place SavVol on separate disks when possible.
- DON'T disable SmartSnap traversal.
- AVOID enabling FAST Cache on SavVol.
- DON'T use thin LUNs for SavVol.

### Replicator

- Size disk layout for the primary file system to include copy-on-first-write and replication transfer activity.
- Place the secondary file system on the same drive type as the primary file system.
- It is okay to place SavVol on NL-SAS for replication.
  - If user snapshots are also enabled on the primary file system, then consider the user load from the snapshots to determine whether NL-SAS is still be adequate for SavVol.
- Use 1GbE links for Replicator interconnects when traversing WAN links.
  - 10GbE is typically not necessary for replication across a WAN.
  - When using high-latency networks with Replicator, use a WAN-accelerator to reduce latency.

### Deduplication and compression

If using VNX OE for File deduplication and compression:

- Target deep compression at inactive files.
- DON'T enable CIFS compression on busy Data Movers.
  - CIFS compression occurs in-line with the host write and can delay response time if the compression activity is throttled.

### CAVA

- Always consult with the antivirus vendor for their best practices.
- Ensure CIFS is completely configured, tested, and working before setting up virus checker.
- Ensure the antivirus servers are strictly dedicated for CAVA use only.

- Ensure that the number CIFS threads used are greater than virus checker threads.
- Exclude real-time network scanning of compressed and archive files.
- Set file mask to include only file types recommended for scanning by the antivirus provider.
  - DON'T include \*.\*.
- Disable virus scanning during migrations.

## VNX OE for File application tuning

### Hypervisor / Database over NFS or SMB

DON'T use Direct Writes

- VNX OE for File cached I/O path supports parallel writes, especially beneficial for Transactional NAS.
  - Prior to VNX OE for File 8.x, file systems required the Direct Writes mount option to support parallel writes.
    - Direct Writes mount option (GUI); uncached mount (CLI).

Recommendation is to use cached mount in 8.x

- VNX OE file systems use cached mount by default, which provides buffer cache benefits.
  - Buffer cache ensures that files are 8KB aligned.
  - Allows read cache hits from data mover.

### Bandwidth-intensive applications

Single Data Mover bandwidth has a default nominal maximum of 1600 MB/s (unidirectional); this is due to the 2x 8Gb FC connections from the Data Movers to the Storage Processors.

- Scale VNX OE for File bandwidth by utilizing more active Data Movers (model permitting), or by increasing the number of FC connections per Data Mover from 2x to 4x.

Increasing the FC connectivity to 4x ports requires utilizing the 2x AUX FC ports; as a result, FC-based NDMP backups are not allowed directly from Data Mover. An RPQ is required for this change.

This best practices guide provides configuration and usage recommendations for VNX unified systems in general usage cases.

For detailed discussion of the reasoning or methodology behind these recommendations, or for additional guidance around more specific use cases, see the documents the [related documents](#) section.