Nutanix Tech Note

Configuration Best Practices for Nutanix Storage with VMware vSphere

Nutanix Virtual Computing Platform is engineered from the ground up to provide enterprise-grade availability for critical virtual machines and data. This Tech Note describes Nutanix storage concepts and how they should be configured in a VMware vSphere environment to ensure optimal performance and availability.

Table of Contents

Executive Summary

The Nutanix Virtual Computing Platform is a highly resilient converged compute and storage platform, designed for supporting virtual environments such as VMware vSphere. The Nutanix architecture runs a storage controller in a VM, called the Nutanix Controller VM (CVM). This VM is run on every Nutanix server node in a Nutanix cluster to form a highly distributed, shared-nothing converged infrastructure. All CVMs actively work together to aggregate storage resources into a single global pool that can be leveraged by user virtual machines running on the Nutanix server nodes. The storage resources are managed by the Nutanix Distributed File System (NDFS) to ensure that data and system integrity are preserved in the event of node, disk, application, or hypervisor software failure. NDFS also delivers data protection and high availability functionality that keeps critical data and VMs protected.

Because of this highly distributed, shared-nothing architecture, the Virtual Computing Platform is not constrained by a limited number of storage controllers. This is a huge advantage compared to traditional shared storage platforms designed with two or four storage controllers. As a result, architecting and scaling a solution with Nutanix to service all datacenter requirements is much less complex.

This document is intended to help the reader understand core storage configuration best practices for a Nutanix cluster running with VMware vSphere. Implementing the following best practices will enable Nutanix customers to get the most out of their storage and virtualization investments.

Nutanix Overview

The Nutanix Distributed File System takes local SSD and HDD storage resources and aggregates them into one single global pool of resources called a storage pool. From this storage pool, the administrator can create several containers, which are presented to the hypervisor and used to host virtual machines. Each container can have a different set of compression, deduplication, and RF policies applied.

Storage Pools

A *storage pool* on the Nutanix Virtual Computing Platform is a group of physical disks from one or more tiers. When creating a storage pool for use in a VMware vSphere environment, it is generally best to have a single storage pool, which contains all storage available in the cluster. This configuration allows you to use all available resources in the cluster, and lets Nutanix optimally place data across the nodes, SSD, and HDD. This configuration supports the majority of use cases. The following table contains the recommendations for setting up storage pools for vSphere.

Recommendations for Storage Pools	
Nodes	Create one Nutanix Cluster with all nodes
SSD Drives	Include all drives (SSD and SATA) in a single storage pool. This is done automatically with NOS 3.0 or later.
SATA Drives	
Data tiering (ILM) threshold	Default (75%)

Containers

On top of a storage pool, an administrator creates one or more Nutanix *containers*. A *container* is presented to VMware ESXi hosts as a NFS datastore. This NFS datastore is attached to every ESXi host within the Nutanix cluster, and can be serviced by any CVM. NFS datastores are used to host VMware virtual machines and their associated files. The number of VMs per datastore is only limited by the available capacity in the Nutanix cluster and by VMware vSphere configuration maximums.

Recommendation for Containers (Datastore)	
General	1. Ensure sufficient capacity available within the cluster to tolerate node unavailability or maintenance. 2. Create one container per configuration policy as required. Example container configuration policies include: Resiliency Factor (RF); Deduplication On/Off; and Compression On/Off (In-Line or Post-Process).
Virtual Machines per Container	Up to 2048. This is the maximum number of virtual machines per datastore in a vSphere HA cluster, as of ESXi 5.5.

Thin Provisioning

All Nutanix containers are thin provisioned by default; this is a feature of NDFS. Thin provisioning is a widely accepted technology that has been proven over time by multiple storage vendors, including VMware. As containers are presented by default as NFS datastores to VMware vSphere hosts, all VMs will also be thin provisioned by default. This results in dramatically improved storage capacity utilization without the traditional performance impact.

Thick provisioning on a VMDK level is available if required for the limited use cases such as fault tolerance (FT) or highly demanding database and I/O workloads. Thick provisioning can be accomplished by creating Eager Zero Thick VMDKs. Eager Zero Thick VMDKs will automatically guarantee space reservations within NDFS.

Similar to compute-level reservations in VMware vSphere, reservations for the Virtual Computing Platform allow an administrator to configure an amount of storage capacity from the underlying storage pool, which is guaranteed to the Nutanix container. The purpose of this is to enable a minimum level of available storage capacity to a Nutanix container, regardless of the utilization level of the storage pool. As with reservations in VMware vSphere, reservations at the infrastructure layer should be applied as an exception rather than by default.

The following are general recommendations; however Nutanix suggests all customers carefully consider the advantages and disadvantages of reservations for their specific applications:

Recommendation for Reservations per Container (Datastore)	
General	1. Avoid using reservations by default. 2. Leverage thin provisioning to maximize storage capacity utilization. 3. Where space reservations are required, use Eager Zero Thick VMDKs, which automatically guarantee space reservations in NDFS.

Capacity Optimization

The Virtual Computing Platform provides capacity optimization features that improve storage utilization and performance. Two of these key features are compression and deduplication. The following sections provide key considerations for using these features.

Compression

The system currently provides two types of compression policies, described in the following table:

Type of Compression	
In-line	Data is compressed as it is written (synchronously) to optimize capacity and to maintain high performance for sequential I/O operations. In-line compression checks the type of I/O and will only compress sequential I/O to avoid degrading performance for random write I/O.
Post-Process	For random workloads, data is written uncompressed to the SSD tier for high performance. Compression occurs after “cold” data is migrated to lower-performance storage tiers. Post-process compression acts only when data and compute resources are available, so normal I/O operations are unaffected. Post process compression is generally suited to random write I/O to avoid performance impacts during the write operation.

Nutanix recommends all customers careful consider the advantages and disadvantages of compression for their specific applications. For further information on compression, refer to the Nutanix Compression Benefits Tech Note.

Deduplication

The next feature for capacity optimization delivered by the Virtual Computing Platform is the software-driven Elastic Deduplication Engine. It increases the effective capacity in the disk tier, as well as the utilization of the performance tiers (RAM and flash), by eliminating duplicate data. This substantially increases performance for de-duplicable workloads due to larger effective cache sizes in the performance tier.

Deduplication savings will vary greatly depending on workload and data types, but will generally provide the largest benefit for common data sets, such as operating system data. Where a Nutanix cluster is not capacity constrained, deduplication can be disabled to optimize write performance.

The following are general recommendations, but Nutanix recommends all customers careful consider the advantages and disadvantages of deduplication for their specific applications. For further information on deduplication on the Nutanix Platform, see the Elastic Deduplication Engine Tech Note.

Recommendation for deduplication per container (datastore)	
Container(s) hosting business-critical applications	Enable. Be sure to increase CVM memory to 24 GB or more. Consider disabling the option 'fingerprint on write' if the measured deduplication rates are less than 1.1:1 or if the Nutanix CVMs are becoming CPU-constrained.
Container(s) hosting VDI	
Container(s) hosting general server workloads	
Container(s) hosting Big Data	

NDFS Data Protection

The Nutanix Distributed File System (NDFS) uses synchronous replication of writes between multiple nodes to provide data protection. The replication or resiliency factor (RF) is used to control the number of nodes data should be written to in the Nutanix cluster.

For example, when a Nutanix container is configured to use a policy of RF=2, each write is written to the local node's SSD and synchronously written to another remote Nutanix node's SSD. RF=2 provides excellent data protection, while still providing high performance write I/O. Subsequent writes can be stripped across any remote nodes of the Nutanix cluster, NDFS is not merely mirroring data to a remote node. This feature is key to maintaining high performance across a large distributed system.

Recommendation for Replication Factor	
Use Replication Factor 2 (RF2) - Default	1. Suitable for all workloads types 2. High/Medium/Low performance requirements

Nutanix CVM Sizing

The following are general recommendations for Nutanix CVM sizing. Nutanix suggests all customers carefully consider the requirements for their specific environments.

Recommendation for CVM Sizing	
vCPUs	• Use the Nutanix default for number of vCPUs • CPU reservation of 50% of the VMs total physical cores speed
vRAM	• Default (16GB) • Add an additional 8GB when enabling deduplication • Add an additional 8GB for nodes servicing business critical applications (to extend caching capabilities) • Add additional memory if the host memory is idle and to increase the size of the Nutanix cache • Memory reservation 100% (All locked)
VM Options	• For nodes servicing latency-sensitive applications, set the latency sensitivity to "High" (vSphere 5.5 or later)

CVM Networking

The following are general recommendations for Nutanix CVM networking. Nutanix suggests all customers evaluate requirements for their specific environments. For more information relating to VMware vSphere networking, refer to the VMware vSphere Networking with Nutanix Tech Note.

Recommendation for CVM Networking	
General	• Use Jumbo Frames where supported end-to-end by the network for maximum performance. Caution: Misconfiguration of Jumbo Frames can result in poor performance. If unclear on how to configure and test end-to-end with your networking switch vendor, avoid this configuration. • Create a dedicated VLAN for Infrastructure, including Nutanix CVM, IPMI, and ESXi hosts
When using Virtual Standard Switch (VSS)	• Use Jumbo Frames where supported end-to-end by the network for maximum performance. Caution: Misconfiguration of Jumbo Frames can result in poor performance. If unclear on how to configure and test end-to-end with your networking switch vendor, avoid this configuration.

When using Virtual Distributed Switch (VDS)	• Use network I/O control • Use load-based teaming

Multiple vSphere Clusters with Nutanix

A single Virtual Computing Platform can efficiently service multiple VMware vSphere clusters. There are several different reasons to deploy multiple vSphere clusters on a single Nutanix Virtual Computing Platform, including high availability rules, unique performance needs, and growth and virtualization-driven workload/management isolation. Multiple vSphere clusters can be used for a variety of roles, including:

1. Management clusters
2. DMZ
3. vCloud or vCAC resource clusters
4. Dedicated customer clusters (in multi-tenant / service provider environments)
5. Business-critical applications
6. Virtual desktop infrastructure (VDI)
7. When and where they are restricted by application-level licensing

For example, a 12-node Virtual Computing Platform could be deployed in any combination including:

1. One single 12-node VMware vSphere cluster for all virtualized workloads
2. Two six-node VMware vSphere clusters, with one being used for business-critical applications and the other for VDI
3. Three 4-node VMware vSphere clusters, with each one being assigned to different lines of businesses
4. Six 2-node VMware vSphere clusters, with each running certain applications for licensing reasons

When deciding on vSphere cluster design, it is important to take into account the specific workloads, business requirements, and resource demands. Having all vSphere clusters accessing a traditional separate shared SAN/NAS is a high risk, especially in mixed workload environments. However, using the Nutanix scale out, shared-nothing architecture mitigates or eliminates these risks associated with traditional shared storage:

Traditional Storage	Nutanix Solution
Noisy neighbor issue	• Natively eliminating noisy neighbor issues by leveraging CVMs across the entire Virtual Computing Platform, each

	with its own flash and capacity tiers of storage. • Not being limited by LUN or HBA queue depth of SAN. • All Nutanix controllers concurrently serving I/O to all datastores. • Allowing capacity reservations on datastores hosting critical VMs.
Large single fault domain	• The Nutanix solution is designed to eliminate any single point of failure using a shared-nothing architecture and self-healing mechanisms. • The larger the Virtual Computing Platform, the faster the cluster can self-heal from a failure such as SSD, HDD, or an entire node.
Finite performance and capacity limits	• NDFS currently does not have a limit to the maximum number of nodes. • NDFS enables deployments to scale performance linearly – there is no overhead associated with scaling. • Multiple generations of nodes can be incorporated into the same system, eliminating the need for forklift upgrades. • Any scaling changes to the system, including adding additional nodes, requires no downtime. • The performance storage tier based on SSDs is pooled and shared across the entire system, enabling not only low latency reads, but also fast writes. • Similarly, the capacity tier is pooled and shared, allowing VMs to overcommit without impact to utilization.
Operational and Human Error	• The VM-centric nature of operations enables automation of common tasks and eliminates the need to map VMs to storage. • Non-disruptive and automated software upgrades minimize the chance of operational issues during updates. CVMs are updated sequentially after an automated health check of the newly upgraded node and Virtual Computing Platform.

Handling Isolation Requirements

For environments requiring isolation in a shared virtualization environment, the Virtual Computing Platform is the ideal fit to vSphere functionality. The following table details the common isolation requirements and solutions that meet these isolation requirements when using the shared Virtual Computing Platform:

Isolation Requirement	Solutions
Networking	• Use separate ports on a NIC connected to separate networks, which can be leveraged by VMware vSphere (VLAN, VxLAN, NSX) for network isolation and redundancy. o For example, on a NX-305 node, use a set of ports across two NICs (1 x 10GbE as the active and 1x 1GbE as the standby) to connect to network A, and the other set of ports on the same set of NICs (1 x 10GbE as the active and 1x 1GbE as the standby) to connect to network B for full redundancy.
Storage	Depending on the need: • Create one or more containers serviced by the same storage pool within a single Virtual Computing Platform. • Create a dedicated storage pool and containers for isolation all within a single Virtual Computing Platform.
Compute	• Create one or more separate vSphere Clusters on a single Virtual Computing Platform.

Conclusion

The Nutanix Virtual Computing Platform is a highly resilient converged compute and storage platform designed for supporting virtual environments such as VMware vSphere. Understanding fundamental Nutanix storage configuration features and recommendations is key to designing a scalable and high performing solution, which meets customer requirements. Leveraging the best practices outlined in this document will enable Nutanix and VMware customers to get the most out of their storage, compute, and virtualization investments.

Further Reading

For more information relating to Nutanix Storage configurations for VMware vSphere or to review other Nutanix Tech Notes, including information on data protection and disaster recovery, capacity optimization, and performance measurements, please visit the Nutanix website at: http://www.nutanix.com/resources/