

Nutanix Best Practices Guide

For
Cigna

Copyright 2018 Nutanix, Inc.

Nutanix, Inc.

1740 Technology Drive, Suite 150

San Jose, CA 95110

All rights reserved. This product is protected by U.S. and international copyright and intellectual property laws.

Nutanix is a trademark of Nutanix, Inc. in the United States and/or other jurisdictions. All other marks and names mentioned herein may be trademarks of their respective companies.

Version History

Date	Ver.	Author	Description	Reviewers
08/29/2018	1	Louis McNair	Cigna	N/A

Table of Contents

Version History	3
VMware vSphere Storage Best Practices	4
Storage Pool Best Practices	4
Capacity Optimization Best Practices	4
Nutanix Controller Virtual Machine Best Practices	4
VMware vSphere Networking Best Practices	5
LLDP and CDP Best Practices	5
Network Resource Pool Share Values Best Practices	5
NIC Teaming and Failover Best Practices	6
VMware Virtual Networking Security Best Practices.....	6
Cluster Management	6
Admission Control.....	6
Configuring HA, DRS, and EVC in vCenter	8
Nutanix vSphere HA, DRS, and other Cluster Settings Checklist	10
Disabling Storage I/O Control on a Container (Datastore in VMware)	11
Node Management	12
Shutting Down a Node in a Cluster (vSphere Web Client)	12
Adding a node to a cluster	12
Starting a Node in a Cluster	13

VMware vSphere Storage Best Practices

Storage Pool Best Practices

- Avoid using storage pool reservations by default.
- Take advantage of thin provisioning to maximize storage capacity utilization
- Where space reservations are required, use eager-zeroed thick VMDKs, which automatically guarantee space reservations in the DSF

Capacity Optimization Best Practices

- Compression
 - Inline - The system compresses data synchronously as it is written to optimize capacity and to maintain high performance for sequential I/O operations. Inline compression only compresses sequential I/O to avoid degrading performance for random write I/O.
 - Post-Process - For random workloads, data writes to the SSD tier uncompressed for high performance. Compression occurs after “cold” data migrates to lower-performance storage tiers. Post-process compression acts only when data and compute resources are available, so it doesn’t affect normal I/O operations.
- Deduplication
 - Disable deduplication for all except full clone VDI VMs. Be sure to increase CVM memory to at least 24 GB.
 - Nutanix does not recommend deduplication for general-purpose server workloads, including business-critical applications.
- Replication Factor Best Practices
 - Use replication factor 3 – Clusters with 32 or more nodes, or as needed to meet availability requirements.

Nutanix Controller Virtual Machine Best Practices

- Resource Sizing Best Practices
 - vCPUs – Use the Nutanix Default for the number of vCPUs.
 - vRAM
 - Minimum is 24 GB.
 - Increase RAM to 32 GB when using deduplication.
 - Memory reservation 100% (all locked).
 - Do not add more than 32 GB of RAM unless you are working with Nutanix support.
- CVM Networking Best Practices
 - General recommendations
 - Most deployments do not need jumbo frames. However, for extreme performance requirements, use jumbo frames where the network can support them end-to-end. Caution: misconfiguring jumbo frames can result in poor performance. If you’re not entirely clear on how to

configure and test end-to-end with your networking switch vendor, avoid this configuration.

- Create a dedicated VLAN for infrastructure that includes the Nutanix CVMs and ESXi hosts.
 - vSphere Distributed Switch
 - Use load-based teaming (LBT).

VMware vSphere Networking Best Practices

LLDP and CDP Best Practices

- Choose the correct type dependent on your switching infrastructure. Use CDP for Cisco-based environments and LLDP for non-Cisco environments.
- Both CDP and LLDP are generally acceptable in most environments and provide maximum operational benefits. Configuring an attribute (listen, advertise, or both) in LLDP allows vSphere and the physical network to openly exchange information. You can also choose not to configure an attribute in LLDP if you don't want this exchange to occur.

Network Resource Pool Share Values Best Practices

Network Resource Pool	Share Value	Physical Adapter Shares
Management traffic	25	Low
vMotion traffic	50	Normal
Fault tolerance traffic	50	Normal
VM traffic	100	High
NFS traffic	100	High
Nutanix CVM traffic (1)	100	High
iSCSI traffic	100	High
vSphere Replication traffic	50	Normal
vSphere Storage Area Network traffic (2)	50	Normal
Virtual SAN traffic (3)	50	Normal

(1) This is a custom network resource pool created manually.

(2) This is exposed in vSphere 5.1, but as it has no function and no impact on NIOC, we can ignore it.

(3) This is exposed in vSphere 5.5, but as Nutanix environments do not use it, it has no impact on NIOC here.

○

NIC Teaming and Failover Best Practices

vSphere Distributed Switch (vDS)	
Load Balancing	Route based on physical NIC load (LBT)
Network Failover Detection	Link status only (1)
Notify Switches	Yes
Failback	No
vSphere Standard Switch (vSwitch)	
Load Balancing	Route based on originating virtual port
Network Failover Detection	Link status only (1)
Notify Switches	Yes
Failback	Yes
(1) Enable link state tracking or the equivalent on physical switches.	

○

VMware Virtual Networking Security Best Practices

- Promiscuous Mode – Reject
- Mac Address Change – Reject
- Forged Transmits – Reject

Cluster Management

Admission Control

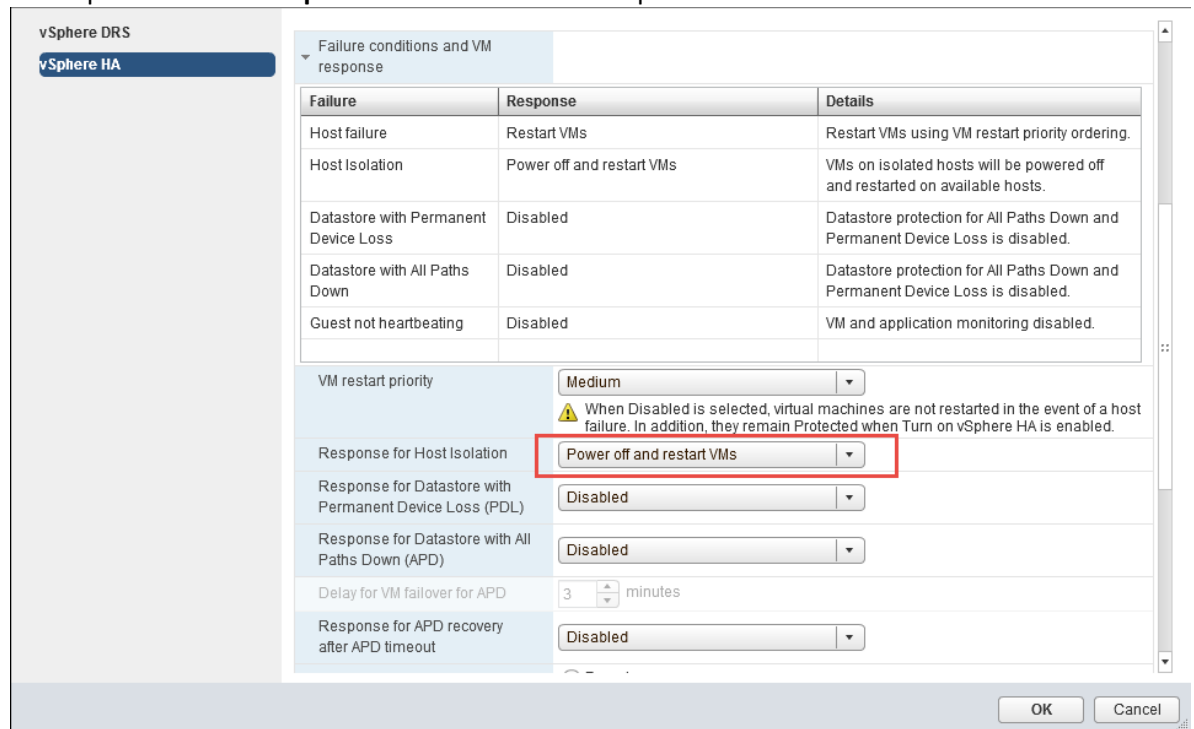
1. If you are using redundancy factor 2 with cluster sizes of up to 16 nodes, you must configure HA admission control setting with the appropriate percentage of CPU/RAM to achieve at least N+1 availability. For cluster sizes larger than 16 nodes, you must configure HA admission control with the appropriate percentage of CPU/RAM to achieve at least N+2 availability.
2. **Table Below - Minimum Reservation Percentage for vSphere HA Admission Control Setting.**
For redundancy factor 2 deployments, the recommended minimum HA admission control setting percentage is marked with single asterisk (*) symbol in the following table. For redundancy factor 2 or redundancy factor 3 deployments configured for multiple non-concurrent node failures to be tolerated, the minimum required HA admission control setting percentage is marked with two asterisks (**) in the following table.
3. Block Awareness
 - a. For deployments of at least three blocks, block awareness automatically ensures data availability when an entire block of up to four nodes configured with redundancy factor 2 can become unavailable.

Nodes	Availability Level			
	N+1	N+2	N+3	N+4
1	N/A	N/A	N/A	N/A
2	N/A	N/A	N/A	N/A
3	33*	N/A	N/A	N/A
4	25*	50	75	N/A
5	20*	40**	60	80
6	18*	33**	50	66
7	15*	29**	43	56
8	13*	25**	38	50
9	11*	23**	33	46
10	10*	20**	30	40
11	9*	18**	27	36
12	8*	17**	25	34
13	8*	15**	23	30
14	7*	14**	21	28
15	7*	13**	20	26
16	6*	13**	19	25

Nodes	Availability Level			
	N+1	N+2	N+3	N+4
17	6	12*	18**	24
18	6	11*	17**	22
19	5	11*	16**	22
20	5	10*	15**	20
21	5	10*	14**	20
22	4	9*	14**	18
23	4	9*	13**	18
24	4	8*	13**	16
25	4	8*	12**	16
26	4	8*	12**	16
27	4	7*	11**	14
28	4	7*	11**	14
29	3	7*	10**	14
30	3	7*	10**	14
31	3	6*	10**	12
32	3	6*	9**	12

Configuring HA, DRS, and EVC in vCenter

1. Log on to vCenter with the vSphere Web Client.
2. Select the Nutanix cluster and go to **Manage > Settings**.
3. If vSphere HA and DRS are not enabled, you can enable them from the **vSphere DRS** and **vSphere HA** tabs.
4. **Note:** It is recommended to configure vSphere HA and DRS even if the customer does not plan to use the features. The settings are preserved within the vSphere cluster configuration, so if the customer later decides to enable the feature, it is pre-configured based on Nutanix best practices.
5. Configure vSphere HA by navigating to **Manage > Settings > vSphere HA** and click **Edit**.
6. Select **Turn on vSphere HA** check box.
 - a. Configure the cluster wide host isolation response settings.
 - b. Open the **Failure conditions and VM response** menu and select **Power off and restart VMs** from the **Response for Host Isolation** drop-down menu.



Failure	Response	Details
Host failure	Restart VMs	Restart VMs using VM restart priority ordering.
Host Isolation	Power off and restart VMs	VMs on isolated hosts will be powered off and restarted on available hosts.
Datastore with Permanent Device Loss	Disabled	Datastore protection for All Paths Down and Permanent Device Loss is disabled.
Datastore with All Paths Down	Disabled	Datastore protection for All Paths Down and Permanent Device Loss is disabled.
Guest not heartbeating	Disabled	VM and application monitoring disabled.

VM restart priority: Medium

Response for Host Isolation: **Power off and restart VMs**

Response for Datastore with Permanent Device Loss (PDL): Disabled

Response for Datastore with All Paths Down (APD): Disabled

Delay for VM failover for APD: 3 minutes

Response for APD recovery after APD timeout: Disabled

- c. Click **OK**.
- d. Configure VM restart priority, host isolation response, and VM monitoring setting for all the Controller VMs.
- e. Go to **Manage > Settings > VM Overrides** and click **Edit**.
- f. Select **Disabled** from the **Response for Host Isolation**, **VM restart priority**, and **VM Monitoring** drop-down menu.

Automation level:	<input type="text" value="--"/>
VM restart priority:	<input type="text" value="Disabled"/>
Response for Host Isolation:	<input type="text" value="Disabled"/>
Response for Datastore with Permanent Device Loss (PDL):	<input type="text" value="Use Cluster Settings"/>
Response for Datastore with All Paths Down (APD):	<input type="text" value="Use Cluster Settings"/>
Delay for VM failover for APD:	<input type="text" value="Use Cluster Settings"/> minutes
Response for APD recovery after APD timeout:	<input type="text" value="Use Cluster Settings"/>
VM Monitoring:	<input type="text" value="Disabled"/>
VM monitoring sensitivity:	<input type="text" value="--"/>

▼ Relevant Cluster Settings

▶ vSphere HA	Expand for details
--------------	------------------------------------

- g. Disable (if enabled) the VM component protection option by clearing the **Protect against Storage Connectivity Loss** check box.
7. Configure datastore monitoring.
 - a. Go to **vSphere HA > Datastore for Heartbeating**.
 - b. Select **Use datastores only from the specified list** and select the Nutanix datastore (NTNX-NFS).
 - c. If the cluster has only one datastore as recommended, click **Advanced Options**, add an **Option** named `das.ignoreInsufficientHbDatastore` with **Value** of `true`, and click **OK**.
8. **Note:** After configuring this setting, you need to clear the **Turn on vSphere HA** check box and wait for all the hosts in the cluster to reconfigure HA, and then again enable HA by selecting the **Turn on vSphere HA** check box.
9. If the cluster does not use vSphere HA, disable it on the **Cluster Features** page. Otherwise, proceed to the next step.
10. Configure DRS by navigating to **Manage > Settings > vSphere DRS** and click **Edit**.
11. Click **Turn On vSphere DRS** check box.
12. In the **DRS Automation** menu, set the default migration threshold of 3 in a fully automated configuration as it is recommended for the Nutanix deployments. This configuration automatically manages data locality in such a way that whenever VMs move writes are always written on one of the replicas locally to maximize the subsequent read performance.
 - a. Configure automation level setting of all the Controller VMs.
 - b. Go to **Manage > Settings > VM Overrides** and click **Edit**.
 - c. Change the **Automation Level** setting of all Controller VMs to **Disabled**.

Automation level:	Disabled	▼
VM restart priority:	--	▼
Response for Host Isolation:	--	▼
Response for Datastore with Permanent Device Loss (PDL):	--	▼
Response for Datastore with All Paths Down (APD):	--	▼
Delay for VM failover for APD:	--	▼ minutes
Response for APD recovery after APD timeout:	--	▼
VM Monitoring:	--	▼
VM monitoring sensitivity:	--	▼

- d. Click **OK**.
 - e. Go to **Manage > Settings > vSphere DRS** and click **Edit**.
 - f. Confirm that **Off** is selected as the default power management for the cluster.
13. If the cluster does not use vSphere DRS, disable it on the **Cluster Features** page.
 14. Click **OK** to close the cluster settings window.
 15. Enable EVC on a cluster.
 - a. Select the cluster in the inventory.
 - b. Shut down all the virtual machines on the hosts with feature sets greater than the EVC mode.
 16. Ensure that the cluster contains hosts with CPUs from only one vendor, either Intel or AMD.
 - a. Click the **Manage** tab, select VMware EVC and click **Edit**.
 - b. Enable EVC for the CPU vendor and feature set appropriate for the hosts in the cluster, and click **OK**.
 - c. Start the virtual machines in the cluster to apply the EVC.
 17. If you try to enable EVC on a cluster with mismatching host feature sets (mixed processor clusters), the lowest common feature set (lowest processor class) is selected. Hence, if VMs are already running on the new host and if you need to enable EVC on the host, you need to first shut down the VMs and then enable EVC.

Nutanix vSphere HA, DRS, and other Cluster Settings Checklist

- Section 1

<input type="checkbox"/>	Enable host monitoring
<input type="checkbox"/>	Enable admission control and use the percentage-based policy with a value based on the number of nodes in the cluster. For more information about settings of percentage of cluster resources reserved as failover spare capacity, vSphere HA Admission Control Settings for Nutanix Environment .
<input type="checkbox"/>	Set the VM Restart Priority of all Controller VMs to Disabled .
<input type="checkbox"/>	Set the Host Isolation Response of the cluster to Power Off .
<input type="checkbox"/>	Set the Host Isolation Response of all Controller VMs to Disabled .
<input type="checkbox"/>	Set the VM Monitoring for all Controller VMs to Disabled .
<input type="checkbox"/>	<p>Enable Datastore Heartbeating by clicking Select only from my preferred datastores and choosing the Nutanix NFS datastore.</p> <p>If the cluster has only one datastore, add an advanced option <code>das.ignoreInsufficientHbDatastore=true</code>.</p>

- Section 2

<input type="checkbox"/>	Set the Automation Level on all Controller VMs to Disabled .
<input type="checkbox"/>	Leave power management disabled (set to Off).

- Section 3

<input type="checkbox"/>	Store VM swapfiles in the same directory as the virtual machine.
<input type="checkbox"/>	Enable EVC in the cluster.

Disabling Storage I/O Control on a Container (Datastore in VMware)

- Risks with not disabling SIOC
 - If SIOC or SIOC in the statistics mode is enabled then storage might become unavailable.
 - If SIOC is enabled and you are using Metro Availability feature, you may face issues with activate and restore operation.
 - If SIOC in the statistics mode is enabled, then this might cause all the hosts to repeatedly create and delete the access and .lck-XXXXXXX files in the .iorm.sf directory in the root directory of the container.
- How to disable SIOC in VMware
 1. Log into the vSphere Web Client.

2. Click **Storage**.
 3. Navigate to the container for your cluster.
 4. Right-click the container and select **Configure Storage I/O Controller**.
The properties for the container is displayed. The **Disable Storage I/O statistics collection** option is unchecked, which means that SIOC is enabled by default.
 5. Select the **Disable Storage I/O statistics collection** option to disable SIOC, and click **OK**.
 6. Select the **Exclude I/O Statistics from SDRS** option, and click **OK**.
- Removing Hidden Files
 1. Log in to the ESXi host by using SSH.
You can log on to any host in the Nutanix cluster if the container is mounted on all the hosts, or log on to the host where the container is mounted.
 2. Go to the container by using the following command.

```
root@esx# cd /vmfs/volumes/container_name
```

 Replace *container_name* with the name of the container.
 3. Remove the contents of the container. For example, to remove the hidden files from ctr1, run the following commands.

```
root@esx# cd /vmfs/volumes/ctr1
root@esx# rm -rf .irom.sf
root@esx# rm -f .iromstats.sf
root@esx# rm -f .lck-XXXXXXXX
```

Changes that you make on one host are applied to all the other hosts in the cluster. Hence, performing this procedure on one host resolves this issue.

This disables SIOC and removes any data related to SIOC, making your container empty. You can now use this empty container to configure metro availability.

Node Management

Shutting Down a Node in a Cluster (vSphere Web Client)

1. Log on to vCenter Server by using vSphere Web Client.
2. If DRS is not enabled, manually migrate all the VMs except the Controller VM another host in the cluster or shut down any VMs other than the Controller VM that you do not want to migrate to another host.
If DRS is enabled skip to the next step.
3. Right-click the host and select **Maintenance Mode > Enter Maintenance Mode**.
4. In the *Confirm Maintenance Mode*, click **OK**.
5. Log on to the Controller VM with SSH and shut down the Controller VM.
(command line) **nutanix@cvm\$ cvm_shutdown -P now**
6. After the Controller VM shuts down, wait for the host to go into maintenance mode.
7. Right-click the host and select **Shut Down**.

Adding a node to a cluster

1. Log into Prism Element of the desired

2. Click on the cog wheel in the upper right corner and select expand cluster
3. Select the host/s that you want to add and click next
4. Download the CSR's for the servers that are going to be added to the cluster
5. Submit the certificates to the team responsible for KMS
6. Once the certificates are received navigate to the cog wheel in the upper right corner and click on expand cluster
7. Select the host/s that you would like to add to the cluster
8. Upload the certificates
9. Proceed with adding the nodes to the cluster

Starting a Node in a Cluster

1. Log on to a running Controller VM in the cluster with SSH.
2. Start the Controller VM.

```
nutanix@cvm$ ~/serviceability/bin/esx-exit-maintenance-mode -s cvm_ip_addr
```

If successful, this command produces no output. If it fails, wait 5 minutes and try again.

```
nutanix@cvm$ ~/serviceability/bin/esx-start-cvm -s cvm_ip_addr
```

Replace *cvm_ip_addr* with the IP address of the Controller VM.

If the Controller VM starts, a message like the following is displayed.

```
INFO esx-start-cvm:67 CVM started successfully. Please verify using ping cvm_ip_addr
```

After starting, the Controller VM restarts once. Wait three to four minutes before you ping the Controller VM.

Alternatively, you can take the ESXi host out of maintenance mode and start the Controller VM using the vSphere Web Client.

3. Verify that all services are up on all Controller VMs.

```
nutanix@cvm$ cluster status
```

If the cluster is running properly, output similar to the following is displayed for each node in the cluster:

```
CVM: 10.1.64.60 Up
Zeus UP [5362, 5391, 5392, 10848, 10977, 10992]
Scavenger UP [6174, 6215, 6216, 6217]
SSLTerminator UP [7705, 7742, 7743, 7744]
SecureFileSync UP [7710, 7761, 7762, 7763]
Medusa UP [8029, 8073, 8074, 8176, 8221]
DynamicRingChanger UP [8324, 8366, 8367, 8426]
Pithos UP [8328, 8399, 8400, 8418]
Hera UP [8347, 8408, 8409, 8410]
Stargate UP [8742, 8771, 8772, 9037, 9045]
InsightsDB UP [8774, 8805, 8806, 8939]
InsightsDataTransfer UP [8785, 8840, 8841, 8886, 8888, 8889, 8890]
```

```

Ergon UP [8814, 8862, 8863, 8864]
Cerebro UP [8850, 8914, 8915, 9288]
Chronos UP [8870, 8975, 8976, 9031]
Curator UP [8885, 8931, 8932, 9243]
Prism UP [3545, 3572, 3573, 3627, 4004, 4076]
CIM UP [8990, 9042, 9043, 9084]
AlertManager UP [9017, 9081, 9082, 9324]
Arithmos UP [9055, 9217, 9218, 9353]
Catalog UP [9110, 9178, 9179, 9180]
Acropolis UP [9201, 9321, 9322, 9323]
Atlas UP [9221, 9316, 9317, 9318]
Uhura UP [9390, 9447, 9448, 9449]
Snmp UP [9418, 9513, 9514, 9516]
SysStatCollector UP [9451, 9510, 9511, 9518]
Tunnel UP [9480, 9543, 9544]
ClusterHealth UP [9521, 9619, 9620, 9947, 9976, 9977, 10301]
Janus UP [9532, 9624, 9625]
NutanixGuestTools UP [9572, 9650, 9651, 9674]
MinervaCVM UP [10174, 10200, 10201, 10202, 10371]
ClusterConfig UP [10205, 10233, 10234, 10236]
APLOSEngine UP [10231, 10261, 10262, 10263]
APLOS UP [10343, 10368, 10369, 10370, 10502, 10503]
Lazan UP [10377, 10402, 10403, 10404]
Orion UP [10409, 10449, 10450, 10474]
Delphi UP [10418, 10466, 10467, 10468]

```

4. Verify storage.

- a. Log on to the ESXi host with SSH.
- b. Rescan for datastores.

```
root@esx# esxcli storage core adapter rescan --all
```

- c. Confirm that cluster VMFS datastores, if any, are available.

```
root@esx# esxcli storage vmfs volumes | awk '{print $5}'
```