Best Practices for Implementing VMware vSphere in a Dell PS Series Storage Environment

Abstract
Dell EMC recommended best practices for configuring VMware vSphere hosts connecting to Dell PS Series storage arrays.

November 2019
## Revisions

<table>
<thead>
<tr>
<th>Date</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>August 2013</td>
<td>Initial Release</td>
</tr>
<tr>
<td>August 2013</td>
<td>Added best practices around CHAP and ACL</td>
</tr>
<tr>
<td>December 2014</td>
<td>Added: ESXi v5.5 EQL FW v7.0 Updated: Delayed ACK / LRO</td>
</tr>
<tr>
<td>December 2015</td>
<td>Added: ESXi v6.0, PS Series FW v8.0, NOOP Timeout, Virtual Volumes</td>
</tr>
<tr>
<td></td>
<td>Moved to appendix: ESX v4.1, ESX v5.0/5.1</td>
</tr>
<tr>
<td></td>
<td>Removed: LRO</td>
</tr>
<tr>
<td>September 2017</td>
<td>Added: ESXi 6.5, EQL FW v9.x</td>
</tr>
<tr>
<td></td>
<td>Removed ESX v4.1, ESX(i) v5.0/5.1</td>
</tr>
<tr>
<td>November 2019</td>
<td>vVols branding update</td>
</tr>
</tbody>
</table>

The information in this publication is provided “as is.” Dell Inc. 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 software described in this publication requires an applicable software license.

© 2013-2019 Dell Inc. or its subsidiaries. All Rights Reserved. Dell, EMC, Dell EMC and other trademarks are trademarks of Dell Inc. or its subsidiaries. Other trademarks may be trademarks of their respective owners.

Dell believes the information in this document is accurate as of its publication date. The information is subject to change without notice.
# Table of contents

Revisions .......................................................................................................................... 2
Table of contents ............................................................................................................... 3
Preface ............................................................................................................................... 4
Executive summary .......................................................................................................... 4
1  Introduction .................................................................................................................... 2
2  Host based connectivity and high availability ............................................................ 3
   2.1 Keep ESXi updated to current revision ................................................................. 3
   2.2 iSCSI subnets and routing recommendations ...................................................... 3
   2.3 Jumbo frames ......................................................................................................... 3
   2.4 CHAP authentication for access control .............................................................. 3
   2.5 Selecting optimal MPIO mode ............................................................................. 4
3  Host based performance ............................................................................................... 7
   3.1 Delayed ACK ......................................................................................................... 7
   3.2 Storage I/O control ............................................................................................... 9
4  Datastore size and queue depth .................................................................................. 10
5  Virtual machine considerations ................................................................................. 11
6  Data Drives .................................................................................................................. 12
   6.1 VMDK on VMFS ..................................................................................................... 12
   6.2 iSCSI in the guest VM .......................................................................................... 14
   6.3 Raw Device Mapped LUNs ............................................................................... 17
7  Conclusion .................................................................................................................... 18
A  Technical support and resources ................................................................................. 19
   A.1 Related documentation ......................................................................................... 19
Preface

Dell™ PS Series arrays optimize resources by automating performance and network load balancing. Additionally, they offer all-inclusive array management software, host software, and free firmware updates.

The information in this guide is intended for VMware® and PS Series storage administrators and installers.

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Model</th>
<th>Software Revision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dell</td>
<td>PS Series storage firmware</td>
<td>9.x</td>
</tr>
<tr>
<td>VMware</td>
<td>VMware vSphere®/ESX™</td>
<td>6.5 Update 1</td>
</tr>
</tbody>
</table>

Executive summary

This technical report is aimed at VMware vSphere and PS Series storage administrators that are configuring a virtual environment. This document covers connectivity and high availability as well as some common performance tuning best practices. In addition, there is a discussion on various methods to implement virtual machine data drive connectivity.
Introduction

1 Introduction

As more organizations are virtualizing datacenters that use both VMware vSphere and PS Series storage, best practices and configuration guides such as this one, help to keep these environments running smoothly. These best practices and recommendations are a combination of various technical report guidelines as well as knowledge provided by the technical support staff at Dell EMC.

Various documents provide extensive detail on how to configure some of the recommended settings. In these instances, instead of adding the same information to this document, the others are referenced.
Host based connectivity and high availability

Several best practices should be implemented when connecting VMware vSphere/ESXi servers to PS Series storage to ensure optimum performance and high availability.

2.1 Keep ESXi updated to current revision

It is important to ensure that the most current ESXi version, also known as the build number, is installed. Using the VMware vSphere Web client, select an individual ESXi server under Hosts and Clusters, and expand the Configuration window to display the version and build number. When using the VMware vSphere Web client, select an individual ESXi server to display the version and build number. VMware Update Manager, part of vSphere vCenter, provides this updating functionality.

2.2 iSCSI subnets and routing recommendations

iSCSI port binding is required for Round Robin and EqualLogic Multipath Extension Module (MEM) MPIO with PS Series storage. Versions of VMware vSphere prior to 6.5 do not support routing when using the iSCSI port binding (described in the VMware KB article 2038869). As a result, the VMkernel ports used to communicate with the PS Series storage must be on the same subnet and able to communicate directly with the SAN and not through a gateway with vSphere versions older than 6.5. Although routing of iSCSI with the use of port binding is supported using vSphere version 6.5 or newer, routed iSCSI environments should only be used when accessing replicas at a remote site for faster verification or restoration of a single virtual machine.

2.3 Jumbo frames

Jumbo frames are not required but are recommended for improved iSCSI performance. In order to support jumbo frames the vmnic, ESXi VMkernel ports, and vSwitch must be configured to use an MTU of 9000. The physical switches must also be capable of using jumbo frames and properly configured for them.

When using iSCSI offload cards, verify the compatibility of the cards and support for jumbo frames before enabling it in the environment.

2.4 CHAP authentication for access control

Manipulating MPIO modules and increasing volume sizes are both great ways to manage volume iSCSI connections. There are other best practices and recommendations that can help control how connections are made to PS Series volumes. CHAP authentication for access control lists is the preferred method of volume access authentication from an ease of administration point of view.

Each PS Series volume ACL (access control list) has a maximum of 16 entries (unrelated to iSCSI connections) that can be CHAP accounts, IQN strings or IP addresses. In keeping with designing a scalable, flexible environment, CHAP allows rapid deployment and discovery of volumes with a single ACL entry. For example, Cluster A could have a single CHAP entry to talk to all of its volumes while Cluster B has a different single CHAP entry to talk to all of its volumes. Shared volumes such as templates or shared data would have both Cluster A CHAP and Cluster B CHAP assigned to it dramatically decreasing the administration of ACL lists.
Note: In a VMware vSphere environment, this allows rapid movement of datastores between clusters. By unregistering the VMs, changing the ACL and rescanning, entire volumes or datastores can be moved between clusters.

Using CHAP helps with scalability, as well as ease of management. When a new host is introduced to a cluster, the administrator only needs to set the CHAP configuration on the iSCSI initiator so the new host can access the cluster storage resources. When a new volume is created, only one ACL entry and an initiator rescan are needed for all the cluster hosts to be able to access the additional storage capacity.

2.5 Selecting optimal MPIO mode

2.5.1 Multipath Extension Module

MEM is a software package that can be installed in VMware vSphere environments to provide intelligent load balancing, multipathing, and failover. With appropriate licensing, VMware allows the installation of vendor specific MPIO software (MEM). These Path Selection Policy (PSP) modules, allow storage vendors to tune MPIO optimally for their hardware. In environments with the appropriate licensing, this is the preferred option for best performance. The MEM will optimize both datastore I/O and Raw Device Mapped (RDM) volumes. In addition to providing intelligent multipathing, the setup.pl script included with MEM can ease the initial iSCSI configuration as well as configure the iSCSI initiator in accordance with Dell EMC’s best practices.

Note: Use of MEM has specific VMware vSphere licensing requirements: vSphere 5.x requires Enterprise or Enterprise + licensing, and vSphere 6.x requires Standard or above licensing.

While the steps may differ between versions, the MEM is supported on all versions of vSphere currently supported by VMware.

For more information on the MEM, including installation and configuration instructions, refer to Configuring and Installing the EqualLogic Multipathing Extension Module for VMware vSphere and PS Series SANs.

2.5.2 VMware Round Robin

If MEM is not an option, set the Path Selection Policy to use VMware Round Robin. This enables all of the physical NICs to participate in load balancing to the datastore. The VMware default path policy (Fixed) will not load balance I/O across all of the physical NICs but will instead use the additional paths only in the event of a failure. This could potentially prevent performance scaling past one NIC.

In order for Round Robin to be configured as an option, the iSCSI connectivity must be properly configured to use more than one physical adapter. In addition, each iSCSI VMkernel port will be bound to only one of these physical adapters as discussed in the Configuring iSCSI connectivity with VMware and Dell PS Series Storage paper. The path selection policy can be changed from Fixed to Round Robin using the VMware vSphere Web client for each datastore. Because this can become unwieldy due to existing and new datastores having to be changed manually, this can be performed using the command line as described in VMware KB article 1017760.

In addition to changing the path selection policy to Round Robin, modify the number of I/Os that have to be sent before switching paths in order to make Round Robin function better. The default behavior is to wait until 1000 I/Os have been sent before switching to the next available path. This may not fully utilize the entire available bandwidth to the SAN when multiple paths are available. Change the number of I/Os from 1000 to 3.
This script sets all PS Series volumes to Round Robin and the IOPS value to 3. This script can be copied into the console or an SSH session to make it easier. This script must be run on each ESXi host accessing PS Series storage.

```bash
esxcli storage nmp satp set --default-psp=VMW_PSP_RR --satp=VMW_SATP_EQL ; for i in `esxcli storage nmp device list | grep EQLOGIC|awk '{print $7}'|sed 's/(//g'|sed 's/)//g'`; do esxcli storage nmp device set -d $i --psp=VMW_PSP_RR ; esxcli storage nmp psp roundrobin deviceconfig set -d $i -I 3 -t iops ; done
```

Use the next command to set the default so that new PS Series volumes inherit the proper setting. To be effective, this new rule requires a reboot of the ESXi host. Until then, if you add more PS Series (EQL) volumes, re-run the script above on all ESXi hosts accessing PS Series storage.

```bash
esxcli storage nmp satp rule add -s "VMW_SATP_EQL" -V "EQLOGIC" -M "100E-00" -P "VMW_PSP_RR" -O "iops=3"
```

To verify that the correct settings are in effect, use:

```bash
esxcli storage nmp device list
```

The output should look similar to the following:

```bash
naa.6090a098703e5059e3e2e483c401f002
Device Display Name: EQLOGIC iSCSI Disk (naa.6090a098703e5059e3e2e483c401f002)
Storage Array Type: VMW_SATP_EQL
Storage Array Type Device Config: SATP VMW_SATP_EQL does not support device configuration.
Path Selection Policy: VMW_PSP_RR
Path Selection Policy Device Config:
{policy=iops,iops=3,bytes=10485760,useANO=0;lastPathIndex=3:
NumIOsPending=0,numBytesPending=0}
```

**Note:** The storage array type VMW_SATP_EQL is used to indicate the PS Series volume and that the Path Selection Policy is VMW_PSP_RR for Round Robin. In addition, the policy=iops,iops=3 verifies that the number of I/Os between path change has been set to three.

### 2.5.3 Virtual Volume related settings

In general, following best practices not specific to VMware® vSphere® Virtual Volumes™ (vVols) is appropriate. However, additional requirements must be adhered to when implementing vVols:

- The VMware vCenter and the Dell EqualLogic Virtual Storage Manager (VSM) virtual appliances:
  - Must be installed on an HA-enabled vSphere cluster
  - Must be installed on a non-vVol shared storage datastore
  - Must not be migrated (Storage vMotion®) to a vVol datastore
- The vSphere cluster HA settings must be modified to set a higher starting priority to vCenter and VSM over other VMs.
- NOOPTTimeout settings (as described in section 2.5.4.2)

Additional information on vVol configuration and options can be found in the [VMware vSphere Virtual Volumes on Dell PS Series](https://www.dell.com) paper.
2.5.4 Default iSCSI login and Noop timeout values

In addition to configuring Round Robin or MEM for multipathing, there is a recommendation for the host iSCSI login and Noop timeout settings. These timeout values might be too low in large environments and prevent paths from reconnecting or potentially cause an all paths down (APD) event. By default, the MEM configuration script attempts to set Login timeout values to the recommended 60 seconds. However, at this time, Noop timeout values must be modified manually.

2.5.4.1 Login timeout

The default login timeout value is 5 seconds. Change this to 60 seconds on each host to allow more time for the iSCSI login process to respond to the host. Timeout is the maximum time that the host will wait to receive an iSCSI login response from the PS Series storage before giving up on the login process and starting over. This is especially important when it comes to controller failover, which occurs during firmware updates on the PS Series storage.

**vSphere Web client:**

1. On the Home screen, click **Hosts and Clusters** and select a host by drilling down into the datacenter.
2. Click the **Configure** tab; in the **Storage** section, click **Storage Adapters** and select the **iSCSI Software Adapter** from the list.
3. Under **Adapter Details**, select **Advanced Options** and click **Edit**.
4. Scroll down to **LoginTimeout**, change the value to 60 and click **OK**.

This can also be done using the CLI with the following command:

```
esxcli iscsi adapter param set --adapter=vmhba\x --key=LoginTimeout --value=60
```

(where \x represents the appropriate vmhba number in the system)

2.5.4.2 Noop timeout

The default Noop timeout value is 10 seconds. Change this to 30 seconds on each host to allow more time to establish non-active paths to the volume are alive. If a response from the volume is not received within the Noop timeout value, the path is marked as dead. While this is a best practice for every install, it is critically important when using vVols with multi-member volumes.

**vSphere Web client:**

1. On the Home screen, click **Hosts and Clusters**, navigate through the datacenter and select a host.
2. In the **Storage** section of the **Configure** tab, click **Storage Adapters** and select the **iSCSI Software Adapter** from the list.
3. Under **Adapter Details**, select **Advanced Options** and click **Edit**.
4. Scroll down to **NoopTimeout**, change the value to 30 and click **OK**.

This can also be done using the CLI with the following command:

```
esxcli iscsi adapter param set --adapter=vmhba\x --key=NoopOutTimeout --value=30
```

(where \x represents the appropriate vmhba number in the system)
3 \hspace{1cm} \textbf{Host based performance}

3.1 \hspace{1cm} \textbf{Delayed ACK}

Delayed ACK is a TCP/IP method of allowing segment acknowledgements to piggyback on each other or other data passed over a connection with the goal of reducing I/O overhead.

![Diagram of No delayed ACK and Delayed ACK](image)

One side effect of delayed ACK is that if the pipeline is not filled, acknowledgement of the data is delayed. In SANHQ, this can be seen as higher latency during lower I/O periods. Latency is measured from the time data is sent to when the acknowledgement is received. With disk I/O, any increase in latency can result in decreased performance. As more customers are moving to 10Gbe and using jumbo frames, this situation shows up more often when monitoring an environment with SANHQ. As a result, Dell EMC recommends delayed ACK be disabled.

\textit{VMware KB article 1002598} has additional information, including instructions for disabling ACK on ESX 3.5 through ESXi 6.x.

\textbf{Note}: The current version of MEM disables Delayed ACK when configured using the Best Practice settings option or flag. If MEM is not installed or was configured without the Best Practices option, Delayed ACK must be disabled manually.

3.1.1 \hspace{1cm} \textbf{Disabling Delayed ACK}

For new installations, using the vSphere Web client:

1. From the Home screen, click \textbf{Hosts and Clusters}, navigate through the datacenter and select a host.
2. In the \textbf{Storage} section of the \textbf{Configure} tab, click \textbf{Storage Adapters} and select the \textbf{iSCSI Software Adapter} from the list.
3. This step differs depending on if you are setting the Delayed ACK property globally for the iSCSI initiator or for an individual discovery address.
• Setting the Delayed ACK property globally for the iSCSI initiator
  a. Under Adapter Details, select Advanced Options and click Edit.
  b. Scroll down to Delayed ACK.
  c. Deselect the checkbox and click OK

• Setting the Delayed ACK property for an individual discovery address
  a. Under Adapter Details, select Targets, Dynamic Discovery.
  b. Select the discovery address to modify and click Advanced.
  d. Scroll down to Delayed ACK.
  e. Deselect the checkbox for DelayedAck in the Inherited column.
  f. Deselect the checkbox for DelayedAck in the Value column and click OK.

  4. Do this for each discovery address that needs to be modified.

For existing connections, using the vSphere Web client:

Disabling Delayed ACK using the above procedure only disables Delayed ACK for new volumes. Any existing volumes must use the following procedure to correct the entries already stored in the iSCSI initiator database.

  1. Put ESXi node into maintenance mode.
  2. To verify Delayed ACK status at the ESXi shell, enter the following command

```
# vmkiscsid --dump-db | grep Delayed
```

iSCSI MASTER Database opened. (0xff889008)
`node.conn[0].iscsi.DelayedAck`='1'
`node.conn[0].iscsi.DelayedAck`='1'
`node.conn[0].iscsi.DelayedAck`='0'
`node.conn[0].iscsi.DelayedAck`='0'
`node.conn[0].iscsi.DelayedAck`='0'
`node.conn[0].iscsi.DelayedAck`='0'

A value of “1” indicates that DelayedACK is enabled, “0” is disabled.

  3. Using the vSphere Web client, click the Configure tab and in the Storage section click Storage Adapters and select the iSCSI Software Adapter from the list.
     (If you are using a HW iSCSI adapter, you must repeat this procedure for each adapter in use.)
  4. Under Adapter Details, select Targets.
  5. In the Static Discovery listings, select the top entry and then hold down Shift while clicking the bottom entry.
  6. Click Remove to delete all the selected entries.
  7. In the Dynamic Discovery listing, remove any PS Series storage entries.
  8. Reboot the node.
9. Return to Configure > Storage > Storage Adapters > iSCSI Software Adapter > Adapter Details > Targets and select Dynamic Discovery again.
10. Add the PS Series storage Discovery address, but do NOT rescan the adapter yet.
11. From the Dynamic Discovery listing, select the discovery address to modify and click Advanced.
12. Scroll down to Delayed ACK.
13. Deselect the checkbox for DelayedAck in the Inherited column.
14. Deselect the checkbox for DelayedAck in the Value column and click OK.
   Do this for each discovery address that needs to be modified.
   This is also a good time to verify that the Login Timeout is set to 60 and NOOP Timeout is set to 30.
15. Now you can rescan the Storage adapter(s) and verify that Delayed ACK is disabled, using the command from Step 2.

3.2 Storage I/O control

Storage I/O Control (SIOC) is a means of ensuring that the excessive storage I/O demands of a particular virtual machine disk (VMDK) do not negatively impact the storage I/O needs of other VMDKs residing on the same datastore.

Because of the various load balancers in the PS Series storage, disable the SIOC feature on PS Series storage based datastores. Otherwise, VMware vCenter might get false information and attempt to load balance at the same time causing unnecessary data movement. For more information about SIOC, refer to Dell PS Series Arrays: Advanced Storage Features in VMware vSphere.
Datastore size and queue depth

While administrators continually try to maintain an optimized data layout and performance, the size of the datastore becomes a question. Because every environment is different, there is no hard-set number. However, you will find recommendations of 10-30 VMs per datastore. A number of factors in this decision include speed of the disks, RAID type and workload intensity of the VMs.

Having one VM for each volume can pose an abundance of administrative overhead, while putting all VMs on a single datastore could cause a performance bottleneck. Determine the most beneficial compromise by monitoring the environment with SANHQ to find volumes that may be underperforming. In addition, monitor the queue depth to see if there are outstanding I/Os to a volume indicating that too many VMs may reside on that datastore.

The current VMware maximum supported datastore is 64TB. The PS Series array can create a maximum volume size of 15TB.
5 Virtual machine considerations

Apply best practices on a virtual machine to allow increased performance. One of these is changing the disk timeout values to 60 seconds. For Windows® guests, the VMware tools typically do this by default. For more information on the disk timeout setting recommendations, refer to the OS Initiator Guide in the PS Series Firmware download section of the support site at support.dell.com/equallogic (this site requires a login).
Data Drives

6 Data Drives

There are four primary ways to connect data drives to VMs in a VMware vSphere and PS Series storage environment, assuming that the C:\ or /root drive of the guest OS is located on a VMFS datastore volume.

- VMDK on VMFS
- iSCSI in the guest (sometimes called direct connect, or storage direct)
- Raw Device Mapping (RDM)
- Virtual Volumes (vVols)

6.1 VMDK on VMFS

A VMDK disk on a VMFS volume is the most commonly used practice for data drives. This scenario involves carving out additional space on a VMFS datastore and assigning it to a virtual machine.

Advantages:

- Easy to configure: Easily add storage to a VM from free space in existing datastores or provision another datastore
- Viewable in vCenter: This keeps the administration overhead light, the data drive location is known and it is easier to integrate vCenter commands, such as cloning, templates and Storage vMotion
- Does not require storage team interaction: As long as the datastore has free space, data drives can be added to a VM without any storage interaction
- Allows for tiering of data drives to different VMFS volumes/pools based on workload: An example of this is the root (OS) drive on the guest can be located on a datastore created on a 10k R50 volume and a database drive on the guest can be located on a datastore created on a 15k R10 volume
- Uses vSphere MPIO to maintain the same performance benefits
- Able to leverage VSM to take hypervisor consistent snapshots of the VM and its associated data

6.1.1 VMDK on VMFS best practices

When configuring a VMDK on VMFS, use the following best practices to maintain performance.

**Using vSphere Web client:**

1. Right-click on the virtual machine where the data drive needs to be added.
2. Select **Edit Settings**.
3. In the **Virtual Hardware** tab, select **SCSI Controller** from the **New Device** drop-down list, and click **Add**. As many as three additional SCSI controllers can be added across which new hard disks can be spread.
4. In the **Virtual Hardware** tab, select **New Hard Disk** from **New Device** drop-down the list, and click **Add**. Do this for each virtual disk to be added to the virtual machine.
5. For VMs that are supported, change the controller type from LSI to a Paravirtual SCSI adapter for better performance.

**Note:** Paravirtual SCSI adapter support and configuration information is located in the [VMware KB article 1010398](https://kb.vmware.com/kb/1010398).

6. Expand the details of the newly listed SCSI controller by clicking >.
7. In **Change Type**, select **VMware Paravirtual** from the list.

8. Expand the details of the newly listed hard disk by clicking >.
9. Select details, including desired capacity, **Disk Provisioning** method, and **Location** (datastore).
10. Change the **Virtual Device Node** to one of the new SCSI controllers.
11. Change the SCSI ID to the next available ID for that controller – shown in the example: **SCSI(1:0)**. When adding multiple hard disk devices, spread them across the up to three added SCSI controllers.
The Dell EMC support staff has seen tremendous performance improvements by making these few changes.

**Note:** Virtualized RDMs and VMDKs prior to ESXi v5.5 are limited to 2TB-512b in size in all versions of ESX.

### 6.2 iSCSI in the guest VM

You can take advantage of iSCSI in the guest VM by using the guest VM iSCSI software initiator. This also allows vMotion and all of the other VMware tools to work.

**Advantages:**

- Isolates the data drives for Data Protection/DR strategies. This technique also allows you to isolate your data drives to a volume that may have different replication and snapshot strategies than the volume that houses the parent VM.
- Can be mounted on physical or other virtual machines. If you have a VM crash or just want to dismount and remount the data, you can easily do this through the iSCSI initiator.
- Uses the same best practices as physical environments. This technique is the same technique used for the physical environment. If you perform a P2V migration of a server that has iSCSI attached volumes, you can continue to run the VM with very little change to the server.
- Allows tiering of data drives to different volumes and pools based on workload. Since the data drive resides directly on a SAN volume, it can perform faster and can be placed on a different SAN tier from the VMs other virtual disks.

**Disadvantages:**

- It is not visible to vCenter. Because the volume is managed from the VM itself, vCenter will not see it in the storage tab and you will not see it connected to the VM when editing the properties. This can cause some additional management overhead.
- Requires storage team intervention. Since it is a brand new volume, you need to create the volume to be seen by the VM not the vSphere environment. This means installing and configuring the iSCSI software initiator, connecting to the SAN with proper pathing, configuring and formatting the volume to be seen by the VM.
- Needs to be considered for DR plans separate from VMs. Because the volume is not seen by vCenter or any of the VMware tools (like Site Recovery Manager), they need to be put into the recovery and protection plans according to their SLAs.

### 6.2.1 iSCSI in the guest VM best practices

Configuring iSCSI in the guest VM is the same as configuring it for a physical server. Install the iSCSI initiator into the guest VM, connect to the SAN and format the presented volume. There are, however, some considerations to think about on the ESX servers:

1. Normally the iSCSI network and Public LAN network are isolated so you first have to make sure that the VM can see the PS Series storage.
2. Verify the VM is configured properly to take advantage of MPIO inside the guest VM if you need it.
3. Modify the vSwitch that is delivering iSCSI traffic to include two iSCSI Guest Port Groups.

In the diagram below, the iSCSI Guest Port Groups are iSCSI Guest 1 and iSCSI Guest 2. These are created on the vSwitch that already communicates on the SAN network.

4. Once the iSCSI Guest Port Groups are created, guarantee traffic across the physical NICs. Monitoring esxtop does not guarantee the traffic across all the physical NICs from a single VM; it has to be forced. Take advantage of the vSwitch NIC Teaming so that each port group is configured to use a single NIC, enabling alternate NICs to be used only for failover.
5. Select one of the iSCSI Guest port groups and click the Edit icon.
6. In the Settings window, click **Teaming and failover** and select **Override** under **Failover order**.

7. Make one of the physical NICs active and set the other to standby (not unused). Do this for the other iSCSI guest port group and use the other adapter. This prevents the guest OS from detecting a link failure. Host MPIO algorithms typically pause I/O for a short time before declaring a path dead. Using a standby prevents this I/O delay. For the guest OS to manage this function, set the alternate NIC to unused.

8. The last thing to do is assign new virtual Network adapters to the VM and attach the iSCSI guest networks to the new adapters.
6.3 **Raw Device Mapped LUNs**

Raw Device Mapped LUNs (RDMs) are used to isolate the data on volumes and retain the view inside vCenter.

**Advantages:**

- Easy to configure: Each RDM is a single volume on the SAN
- Viewable in vCenter under the VMs to which it is attached: The volume is displayed inside vCenter under the VM and enables vCenter integration with software such as SRM
- Enables tiering of data drives to different volumes or pools based on workload: It is possible to isolate the data drive RDMs to another tier of storage for performance
- Uses vSphere MPIO for high bandwidth Apps: Continues to leverage the MEM or Round Robin for multipathing
- Isolates data from OS and enables the use of 3rd party applications which work with RDMs: There are some 3rd party applications that require some form of RDM for either backup or protection or tools

**Disadvantages:**

- No guest integration with Auto Snapshot Manager/Microsoft® Edition: ASM/ME cannot recognize RDM volumes for Smart Copy operations
- No VM integration with VSM Smart Copies: VSM does not recognize RDM volumes for Smart Copy operations for VM snapshots
- Each ESXi server needs connectivity; connection counts need to be evaluated: Since each ESXi server needs to see the RDM volume, there are multiple connections to the RDM even though only one server is using the volume at a time.

Configure RDMs with the same best practices as VMDK on VMFS with the new SCSI adapter and paravirtualized adapter if supported.

6.3.1 **VM data drive formatting options**

Dell EMC has found improvements in performance by using a 64k partition alignment. This is necessary in Windows Operating Systems before Windows 2008, Oracle® Solaris, and Red Hat® Enterprise Linux® guest OSs. Additionally, Microsoft® SQL, Exchange®, and other I/O intensive applications benefit from formatting the filesystem with a 64K cluster size. This format achieves a better alignment of I/Os with the 64K stripe size of the PS Series arrays.
Conclusion

As more environments become virtualized, it is necessary to follow best practices to maintain not only the performance of these environments, but also the high availability. Dell EMC recommends following these practices even if there is no apparent performance issue as it increases the likelihood of the environment being in a known healthy state.
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.

A.1 Related documentation

Dell EMC publications:

- Configuring and Installing the EqualLogic Multipathing Extension Module for VMware vSphere and PS Series SANs
- Configuring iSCSI connectivity with VMware and Dell PS Series Storage
- Dell PS Series Arrays: Advanced Storage Features in VMware vSphere
- VMware vSphere Virtual Volumes on Dell PS Series

VMware publications:

- Considerations for using software iSCSI port binding in ESX/ESXi
  https://kb.vmware.com/kb/2038869
- Changing the default pathing policy for new/existing LUNs
  https://kb.vmware.com/kb/1017760
- ESX/ESXi hosts might experience read or write performance issues with certain storage arrays
  https://kb.vmware.com/kb/1002598
- Configuring disks to use VMware Paravirtual SCSI (PVSCSI) adapters
  https://kb.vmware.com/kb/1010398