

# HITACHI

## Configuring Hitachi VSP Storage in Red Hat Enterprise Linux Environments

Best Practices Guide

MK-SL-456-00

Hitachi Vantara

© Hitachi Vantara LLC 2026. All Rights Reserved.

# Table of Contents

- Notices and Disclaimer ..... 2**
- About This Document..... 3**
  - Introduction ..... 3
  - Intended Audience..... 3
  - Document Revisions ..... 3
  - Comments..... 3
- Overview ..... 4**
- Hitachi Storage and RHEL Configuration Best Practices ..... 4**
  - Fibre Channel..... 4
  - iSCSI ..... 9
  - FC-NVMe..... 13
  - NVMe/TCP ..... 14
  - Multipathing..... 17
  - DM multipath ALUA configuration for GAD..... 19
  - Expand the filesystem ..... 20
  - Boot from SAN ..... 22
  - Hybrid Cloud ..... 23

## Notices and Disclaimer

© 2026 Hitachi Vantara LLC. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording, or stored in a database or retrieval system for commercial purposes without the express written permission of Hitachi, Ltd., or Hitachi Vantara (collectively, "Hitachi"). Licensee may make copies of the Materials provided that any such copy is: (i) created as an essential step in utilization of the Software as licensed and is used in no other manner; or (ii) used for archival purposes. Licensee may not make any other copies of the Materials. "Materials" mean text, data, photographs, graphics, audio, video and documents.

Hitachi reserves the right to make changes to this Material at any time without notice and assumes no responsibility for its use. The Materials contain the most current information available at the time of publication.

Some of the features described in the Materials might not be currently available. Refer to the most recent product announcement for information about feature and product availability, or contact Hitachi Vantara at [https://support.HitachiVantara.com/en\\_us/contact-us.html](https://support.HitachiVantara.com/en_us/contact-us.html).

**Notice:** Hitachi products and services can be ordered only under the terms and conditions of the applicable Hitachi agreements. The use of Hitachi products is governed by the terms of your agreements with Hitachi Vantara.

By using this software, you agree that you are responsible for:

- 1) Acquiring the relevant consents as may be required under local privacy laws or otherwise from authorized employees and other individuals to access relevant data; and
- 2) Verifying that data continues to be held, retrieved, deleted, or otherwise processed in accordance with relevant laws.

**Notice on Export Controls.** The technical data and technology inherent in this Document may be subject to U.S. export control laws, including the U.S. Export Administration Act and its associated regulations, and may be subject to export or import regulations in other countries. Reader agrees to comply strictly with all such regulations and acknowledges that Reader has the responsibility to obtain licenses to export, re-export, or import the Document and any Compliant Products.

**EXPORT CONTROLS** - Licensee will comply fully with all applicable export laws and regulations of the United States and other countries, and Licensee shall not export, or allow the export or re-export of, the Software, API, or Materials in violation of any such laws or regulations. By downloading or using the Software, API, or Materials, Licensee agrees to the foregoing and represents and warrants that Licensee is not located in, under the control of, or a national or resident of any embargoed or restricted country.

Hitachi is a registered trademark of Hitachi, Ltd., in the United States and other countries.

AIX, AS/400e, DB2, Domino, DS6000, DS8000, Enterprise Storage Server, eServer, FICON, Flash Copy, IBM, Lotus, MVS, OS/390, PowerPC, RS6000, S/390, System z9, System z10, Tivoli, z/OS, z9, z10, z13, z/VM, BCPii™ and z/VSE are registered trademarks or trademarks of International Business Machines Corporation.

Active Directory, ActiveX, Bing, Excel, Hyper-V, Internet Explorer, the Internet Explorer logo, Microsoft, the Microsoft Corporate Logo, MS-DOS, Outlook, PowerPoint, SharePoint, Silverlight, SmartScreen, SQL Server, Visual Basic, Visual C++, Visual Studio, Windows, the Windows logo, Windows Azure, Windows PowerShell, Windows Server, the Windows start button, and Windows Vista are registered trademarks or trademarks of Microsoft Corporation. Microsoft product screen shots are reprinted with permission from Microsoft Corporation.

All other trademarks, service marks, and company names in this document or web site are properties of their respective owners.

**IMPORTANT:** This document can only be used as Hitachi Vantara internal documentation for informational purposes only. This documentation is not meant to be disclosed to customers or discussed without a proper non-disclosure agreement (NDA).

## About This Document

### Introduction

This document describes best practices to configure and manage Hitachi Virtual Storage Platform (VSP) storage in Red Hat Enterprise Linux (RHEL) environments.

### Intended Audience

This document is intended for Hitachi Vantara staff and IT professionals of Hitachi Vantara customers and partners who are responsible for planning and deploying such configurations.

### Document Revisions

Revision Number	Date	Details
v1.0	April 2026	Initial release

### Comments

Send any comments on this document to [Docs-Feedback@hitachivantara.com](mailto:Docs-Feedback@hitachivantara.com). Include the document title, including the revision level, and refer to specific sections and paragraphs whenever possible. All comments become the property of Hitachi Vantara LLC.

Thank you!

## Overview

Hitachi Vantara LLC, a subsidiary of Hitachi Ltd., provides various datacenter infrastructure components to enable IT environments to support the Red Hat Linux ecosystem. This includes mid-range and enterprise storage, converged, and hyperconverged infrastructure, as well as a suite of software and software integrations to enable a robust, automated operational environment.

This document outlines best practices for implementing Hitachi storage with Red Hat Enterprise Linux, including guidance on volume discovery, multipathing configuration, and queue depth management when using Hitachi Storage Virtualization Operating System (SVOS).

These aid in building a Linux environment that provides the performance, scalability, reliability, usability, resilience, and recoverability that is expected when integrated with Hitachi products.

These best practices cover the Hitachi storage products listed in the following table.

Table 1 provides the hardware specifications used in this document.

Hardware	Product
Storage Platforms	Hitachi Virtual Storage Platform One Block High End Hitachi Virtual Storage Platform One Block 20 series Hitachi Virtual Storage Platform 5000 series Hitachi Virtual Storage Platform E series

Table 1: Hardware

Table 2 provides the software specifications used in this document.

Software	Product
Hitachi Storage Software	Hitachi Storage Virtualization Operating System (SVOS): Hitachi RAID Manager (CCI) Hitachi Virtual Storage Platform One Block Administrator Hitachi Device Manager Storage Navigator

Table 2: Software

## Hitachi Storage and RHEL Configuration Best Practices

### Fibre Channel

**Host Configuration:** Configuring a Red Hat Enterprise Linux (RHEL) host for a Fibre Channel (FC) SAN setup with Hitachi Virtual Storage Platform (VSP) requires proper planning of SAN zoning, host parameters, preparing the host bus adapters, validating FC connectivity, enabling multipathing, and ensuring LUN visibility. Correct configuration ensures high availability, optimal performance, and enterprise-grade reliability.

The first step is establishing physical connectivity between the RHEL host and the VSP storage through FC switches. Use dual port Host Bus Adapters (HBAs) on the host for redundancy. Connect each HBA to separate SAN fabrics

(Fabric A and Fabric B). Configure single-initiator zoning on the FC switches (one host HBA WWPN per zone with VSP target ports). This ensures path redundancy and fabric-level fault tolerance.

After physical connections and zoning are complete, the host should verify FC link status. Also verify FC link state, remote WWPN visibility, and adapter details using `sysfs` and vendor utilities.

On RHEL, validate that the correct kernel, FC drivers, and utilities are installed.

After installation, ensure that FC HBAs are detected and identify HBA WWNs. You can use the following commands to retrieve details about the installed HBAs from the OS terminal.

```
# cat /sys/class/fc_host/host*/port_name
# cat /sys/class/fc_host/host*/node_name
# cat /sys/class/fc_host/host*/speed
# cat /sys/class/fc_host/host*/port_state
# cat /sys/class/fc_host/host*/port_type
# cat /sys/class/fc_host/host*/symbolic_name
```

Before attaching storage, verify that all host hardware meets requirements outlined in the Hitachi [Product Compatibility Guide](#). Confirm that the storage system is fully configured and provisioned according to the *Open-Systems Host Attachment Guide for Virtual Storage Platform Family* on the Product Documentation portal ([docs.hitachivantara.com](https://docs.hitachivantara.com)).

The following table lists potential error conditions that can occur during storage system installation on a Linux host, along with instructions for resolving these conditions. If you cannot resolve an error, contact customer support.

Error Condition	Recommended Action
Logical devices are not recognized by the system.	Ensure that the READY indicator lights on the storage system are ON. Verify that LUNs are correctly configured. The LUNs for each target ID must start at 0 and continue sequentially without skipping any numbers.
The file system cannot be created.	Confirm that the correct device name is used with <code>mkfs</code> and that the logical unit is properly connected and partitioned.
The file system is not mounted after reboot.	Ensure that the auto-mount information in the <code>/etc/fstab</code> file is correct.

**HBA port timeout:** In RHEL, the `HBA Port Down timeout` parameter specifies the duration a Linux system waits before terminating a FC connection after losing communication with the port.

These settings are crucial for maintaining stable multipathing and avoiding unnecessary path failovers. To mitigate I/O disruptions, configure the HBA port timeouts according to the vendor's recommendations.

For Emulex, the `lpfc_devloss_tmo` parameter defines the time (in seconds) the driver waits before removing a lost device. The default value is 30.

```
# cat /sys/module/lpfc/parameters/lpfc_devloss_tmo
30
```

For QLogic, the `qlport_down_retry` parameter defines how many times the driver retries communication when a port goes down. The default is 0, meaning driver default retry logic is used. Setting the parameter to -1 enables indefinite retry.

```
# cat /sys/module/qla2xxx/parameters/qlport_down_retry
0
# cat /sys/class/fc_remote_ports/rport-*/dev_loss_tmo
30
```

Additionally, device removal timing is primarily controlled by the SCSI/FC layer parameter

`/sys/class/fc_remote_ports/rport-*/dev_loss_tmo`. The `dev_loss_tmo` parameter controls how long the OS keeps a lost FC device before removing it, helping ensure reliable storage failover and preventing unnecessary device loss during transient SAN disruptions. It controls how many retries occur when a port goes down. When value is 0, timeout is managed by the HBA driver instead of the FC transport layer.

**Host queue depth:** Host queue depth defines the number of I/O operations that a storage device (LUN or disk) can process simultaneously. It determines how many read/write requests the host can issue before waiting for earlier requests to complete. Adjusting queue depth can improve I/O performance, but the optimal value depends on the workload, operating environment, and the specific storage system model.

Queue depth settings are applied per storage port and per LDEV. A LDEV is an individual logical data volume in the storage system. Each LDEV has a unique identifier or “address” within the storage system composed of the logical disk controller (LDKC) number, control unit (CU) number, and LDEV number. The LDEV IDs within a storage system do not change. An LDEV formatted for use by open-system hosts is called a logical unit (LU).

The Queue depth settings table lists the queue depth settings to ensure smooth processing at the ports and best average performance. Other queue depth settings, higher or lower than these values, can provide improved performance for certain workload conditions. See the *Open-Systems Host Attachment Guide for Virtual Storage Platform Family* on the Product Documentation portal ([docs.hitachivantara.com](https://docs.hitachivantara.com)) for more information.

Higher storage port queue depth settings can impact host response times. Exercise caution when using queue depth settings that are higher than the recommended values in the *Open-Systems Host Attachment Guide for Virtual Storage Platform Family* on the Product Documentation portal ([docs.hitachivantara.com](https://docs.hitachivantara.com)).

The current queue depth settings for all LUNs on a RHEL host can be viewed using the following command:

```
# lsscsi -l
[0:0:0:0]    disk      HITACHI  OPEN-V          A001  /dev/sdh
  state=running queue_depth=64 scsi_level=6 type=0 device_blocked=0 timeout=30
[0:0:1:0]    disk      HITACHI  OPEN-V          9301  /dev/sdb
  state=running queue_depth=64 scsi_level=6 type=0 device_blocked=0 timeout=30
[0:0:1:59]   disk      HITACHI  OPEN-V          9301  /dev/sdc
  state=running queue_depth=64 scsi_level=6 type=0 device_blocked=0 timeout=30
#
```

For Emulex HBA, the default Queue Depth value can be checked with the following command.

```
# cat /sys/module/lpfc/parameters/lpfc_lun_queue_depth
64
# cat /sys/module/lpfc/parameters/lpfc_hba_queue_depth
8192
```

The active Queue Depth values at runtime can be obtained with these commands. Each VSP LUN shows a queue depth value of 64.

```
# cat /sys/class/scsi_host/host2/lpfc_lun_queue_depth
64
# cat /sys/class/scsi_host/host2/lpfc_hba_queue_depth
5894
```

To change the Queue Depth value for Emulex HBA, edit the `/etc/modprobe.d/lpfc.conf` file and add the following:

```
options lpfc lpfc_lun_queue_depth=128 (to change LUN queue depth, change 128 to the desired value)
options lpfc lpfc_hba_queue_depth=8192 (to change HBA queue depth, change 8192 to the desired value)
```

The default Queue Depth value for QLogic HBA can be checked with the following command:

```
# cat /sys/module/qla2xxx/parameters/ql2xmaxqdepth
```

To change the Queue Depth value for QLogic HBA, edit `/etc/modprobe.d/qla2xxx.conf` and add the following:

```
options qla2xxx ql2xmaxqdepth=64 (change 64 to the desired value)
```

**Zoning:** Use zoning to enable access control in a SAN environment. Through zoning, a SAN administrator can configure which HBA WWPNs on the RHEL host can communicate to which WWPNs on the Hitachi storage processors. The RHEL host port in the Fibre Channel HBA is referred to as the initiator. The storage processor port in the Hitachi storage system is referred to as the target.

Hitachi generally recommends the following:

- For utmost availability with slightly higher administrative cost, use Single Initiator to Single Target (SI-ST) zoning. Brocade Peer Zoning and Cisco Smart Zoning are supported to reduce the administrative burden.
- Each HBA port should only see one instance of each LUN. This is primarily based on years of experience with fabrics and to avoid potential availability issues where host HBA ports can be overrun leading to performance issues. Error recovery from fabric path issues (transient or otherwise) is faster and impacts hosts less.

**Port Topology:** On the VSP storage system, configure the port topology according to the connectivity type. For SAN environments using Fibre Channel switches, set the port to Fabric ON with Point-to-Point (P-to-P) mode. For direct-attached connections between the HBA and storage port, configure the port to Fabric OFF with Point-to-Point (P-to-P) mode. Correct port configuration ensures stable path discovery and optimized throughput between the Linux host and Hitachi storage.

**Host Group and LU path configuration:** After physically connecting the Linux hosts to the storage system with cables and switches, I/O paths must be configured between the hosts and logical volumes. Logical volumes presented to hosts are called Logical Units (LUs), and the connections between hosts and LUs are known as LU paths.

Before defining LU paths, hosts must be organized into host groups based on the operating systems and the storage port to which they are connected. Each host group can contain only hosts connected to the same storage port. For example, Linux hosts connected to different ports must be placed in separate host groups.

After host groups are created and host WWPN are registered, the logical volumes are mapped to host groups.

The LU address assigned to a host is referred to as the Logical Unit Number (LUN).

Key points:

- A host can access multiple LUs.
- LU paths can be added, modified, or removed without system downtime.
- For high availability, configure alternate LU paths to prevent I/O disruption during hardware failures.
- The recommended multipathing configuration is Single-Initiator Single-Target, where each HBA maintains one path per LU.
- In Fibre Channel environments, up to 2,048 LU paths can be defined per host group and per port. For Bi-directional ports: 1,024 Target function and 1,024 Initiator function can be defined.

Correct host group and LU path configuration ensures controlled access, high availability, and optimal I/O performance.

**Configuring Host mode and Host mode options:** On the Hitachi VSP storage system, configure the Host Mode and Host Mode Options (HMOs) specific to Linux to ensure proper communication and command handling between the host and the storage system. Identify the host mode, review the applicable HMOs, and then apply these settings to the storage system

The host mode and host mode options must be set on the host-facing port before the host is connected to that port. If you change the host mode or HMOs after the host is connected, you must restart the host (server) for the new settings to be recognized. 00 [Standard] is the Host mode for Linux server hosts. HMO best practice recommendations for Linux are 2, 22, 25, and 68. HMO 2, 22, and 25 work by default for VSP One Block 20.

HMO	Function	Host Mode	Description
-----	----------	-----------	-------------

2	VERITAS Database Edition/Advanced Cluster	Any	<p>Purpose: By default, Reservation Conflict is returned for a Test Unit Ready run from the host without the PERSISTENT GROUP RESERVATION key setting.</p> <p>In these conditions, Good Status is expected.</p> <p>When this HMO is enabled, the storage system will switch Test Unit Ready response to Good Status from Reservation Conflict.</p> <p>Use this HMO when any of the following conditions are met:</p> <ul style="list-style-type: none"> <li>• Linux with cluster configuration is used.</li> <li>• Symantec Cluster Server (previously known as Veritas Cluster Server (VCS)) is used.</li> <li>• The response for SPC-3 (Good Status) is required.</li> </ul>
22	Veritas Cluster Server	Any	<p>Purpose: By default, Reservation Conflict is returned to the MODE SENSE command from the host without the PERSISTENT GROUP RESERVATION key setting.</p> <p>In the following condition, GOOD status is expected when a reserved volume receives the MODE SENSE command from a node that is not reserving this volume.</p> <p>When this HMO is enabled, the storage system will switch the MODE SENSE command response to Good status from Reservation Conflict.</p> <p>Use this HMO when any of these conditions are satisfied:</p> <ul style="list-style-type: none"> <li>• Veritas Cluster Server is used.</li> <li>• An OS or middleware that supports SPC-4 is used.</li> <li>• A single-site or global-active device (GAD) dual-site Clustered VMDK configuration is used.</li> </ul>
25	Support SPC-3 behavior on Persistent Reservation	Any	<p>Purpose: By default, Reservation Conflict is returned for PERSISTENT RESERVE OUT (Service Action = REGISTER AND IGNORE EXISTING KEY) command if there is no registered key to be deleted.</p> <p>In the following conditions, Good Status (SPC-3 response) is expected.</p> <p>When this HMO is enabled, the storage system will switch this command response to Good Status from Reservation Conflict.</p> <p>Use this HMO when one of these conditions is satisfied:</p> <ul style="list-style-type: none"> <li>• Symantec Cluster Server (previously known as Veritas Cluster Server (VCS)) is used.</li> <li>• There is no registered key to be deleted when running the PERSISTENT RESERVE OUT command.</li> </ul> <p><b>Note:</b> Host types other than those listed in Special Direction expect the response when the option is set to OFF.</p>

68	Support Page Reclamation for Linux	00 [Standard], 01 [(Deprecated) VMware], or 21 [VMware Extension]	<p>Purpose: When this HMO is enabled, the storage system will change the response so that the Linux OS can issue the <code>WriteSame</code> command to use the Page Reclamation function.</p> <p>Use this HMO when using the Page Reclamation function with a Linux host.</p> <p><b>Notes:</b></p> <ol style="list-style-type: none"> <li>1. The option is applied when Dynamic Provisioning is used by Linux 2.6.33 or higher.</li> <li>2. After setting HMO 68 to ON, perform the operation (such as a server restart) that reissues the <b>INQUIRY</b> command from the host.</li> </ol>
----	------------------------------------	---	---

**LUN security:** To protect mission-critical data in your storage system from illegal access, apply security policies to logical volumes. Use VSP One Block Administrator or Storage Navigator to enable LUN security on ports to safeguard LUs from illegal access.

If LUN security is enabled on ports, host groups affect which host can access which LUs. Hosts can only access the LUs associated with the host group to which the hosts belong. Hosts cannot access LUs associated with other host groups. For example, hosts in the Linux host group cannot access LUs associated with the Windows host group. Also, hosts in the Windows host group cannot access LUs associated with the Linux host group.

**LUN distribution:** The general recommendation is to distribute LUNs and workloads so that each host has 2-8 paths to each LDEV. This prevents workload pressure on a small set of target ports to become a potential performance bottleneck.

You should isolate production and critical systems to dedicated ports to avoid contention from other hosts workloads. However, presenting the same LUN to too many target ports could also introduce additional problems with slower error recovery.

Follow these best practices:

- Each HBA port (HBA) should detect only one instance of each LUN
- The number of paths should typically not exceed the number of HBA ports for better reliability and recovery
- Two to four paths to each LUN provides the optimal performance for most workload environments

**Volume Discovery:** Linux exposes storage and HBA information through the `sysfs` pseudo-filesystem under `/sys`, which provides real-time access to kernel-managed devices. Each HBA port appears as a separate directory under `/sys/class/scsi_host/`.

Folders such as `host0`, `host1`, etc., contain files used to scan for new devices or read adapter parameters. This structure applies to common HBAs such as QLogic and Emulex.

To scan a specific SCSI host without restarting the host, write the wildcard pattern `---` to its scan file:

```
echo "--" > /sys/class/scsi_host/hostX/scan
```

For Fibre Channel HBAs, trigger a Loop Initialization Protocol (LIP) to refresh the FC fabric:

```
echo "1" > /sys/class/fc_host/hostX/issue_lip
```

## iSCSI

Internet SCSI (iSCSI) is a protocol for sending and receiving SCSI commands through an IP network. iSCSI transfers data in block units. Linux distributions provide support for iSCSI and include a built-in software initiator. In RHEL environments, the open-iSCSI package functions as the default initiator. This document focuses specifically on the

default Linux software initiator, open-iSCSI. For advanced configurations or specialized use cases, refer to the respective vendor documentation and support services.

The iSCSI software and related tools are provided through the `iscsi-initiator-utils` package. iSCSI initiator information can be obtained in Red Hat host in the `/etc/iscsi/initiatorname.iscsi` file. In an iSCSI environment, up to 2,048 LU paths can be defined for one iSCSI target, and up to 2,048 LU paths can be defined for one port.

#### iSCSI interface specifications (10 Gbps and 25 Gbps channel boards)

Protocol Layer	Items	Specifications
All	iSCSI target functions	Supported
	iSCSI initiator functions	Supported
	For the 10 Gbps iSCSI copper/optical channel board and 25 Gbps Ethernet channel board: Number of connected hosts (connections).	127 connections per port. Because the load on the iSCSI port increases as the number of connections increases, connect to 127 connections per port or fewer.
Physical layer, MAC layer	Maximum transmission unit (MTU)	1500/4500/9000 bytes
	Jumbo frames	Supported.
	Link aggregation	Not supported.
	VLAN	Supported.  VLANs can be configured in the range of 1 to 4094.
TCP/IP	TCP port number	3260 (default). The TCP port number can be changed in the range of 1 to 65535. Note the following: <ul style="list-style-type: none"> <li>1. When you change the TCP port number, also change the settings of the host to be accessed.</li> <li>2. Verify whether the new number is filtered and disabled by a switch on the path.</li> </ul>
	Window size	Supported.  For 10 Gbps iSCSI channel boards: 64 KB (default), 128 KB, 256 KB, 512 KB, 1024 KB  For 25 Gbps Ethernet channel boards: 64 KB, 128 KB, 256 KB (default), 512 KB, 1024 KB, 2048 KB
iSCSI	Error recovery level	Level 0. Errors are corrected when a host retries. Levels 1 and 2 are not supported.

	Header digest, data digest	Supported.  The header digest and data digest protect headers and data in iSCSI communication from errors. iSCSI ports use these functions based on the host settings. However, if the functions are used, performance degrades (the rate of degradation changes depending on the host capabilities and the content of communication).
	CHAP	Supported.  Up to 256 users per iSCSI port
	Mutual CHAP	Not supported.

If using a 25-Gbps iSCSI channel board does not achieve the level of I/O performance you expected, performance might improve by configuring the MTU, socket buffer size, immediate data, maximum receive data segment length, first burst length, maximum burst length, and header digest parameters in Linux hosts.

In the `sysctl` configuration file, enable TCP window scaling by setting `net.ipv4.tcp_window_scaling=1` which allows the system to use larger TCP window sizes beyond 64 KB. Also, set the `net.core.rmem_max` value to 3145728 or greater. Increasing this value helps optimize iSCSI performance by permitting larger receive buffers, which improves data transfer efficiency on faster networks.

In the iSCSI configuration file, set `node.session.iscsi.ImmediateData = yes`, which allows the server to send write data immediately, which helps improve storage performance and reduce delay.

To process larger data segments efficiently, set a higher value for `MaxRecvDataSegmentLength`, `FirstBurstLength` and `MaxBurstLength` in the iSCSI configuration file.

```
node.conn[0].iscsi.MaxRecvDataSegmentLength = 262144
```

```
node.session.iscsi.FirstBurstLength = 262144
```

```
node.session.iscsi.MaxBurstLength = 262144
```

Set `node.conn[0].iscsi.HeaderDigest = None` in the iSCSI configuration file which disables header digest verification, reduces CPU overhead, and can improve performance when the network is stable and reliable.

### Hitachi Storage iSCSI Guidelines

The following guidelines should be implemented on both ends of a connection, and in some cases require compatibility through all adjoining switches and links.

- The Linux host must have an Ethernet interface, preferably with dedicated ports, to establish reliable communication with the iSCSI ports on the Hitachi VSP storage system. Using a dedicated network interface helps ensure optimal performance and minimizes potential conflicts with other network traffic.
- It is considered best practice to separate iSCSI traffic from regular network traffic using dedicated network ports, switches, and supporting infrastructure. If physical network separation is not possible because of topology constraints, isolate iSCSI traffic using VLAN subnets. Additionally, it is strongly advised to configure iSCSI in a multipath setup to ensure path redundancy and improve fault tolerance.

- Changes to iSCSI port settings usually cause a link down/up event so should be configured with resilience and alternate paths in place within appropriate maintenance periods or quiet times. The default Maximum Transmission Unit (MTU) size is 1500 which is generally compatible across networks. However, increasing the MTU or using jumbo frames (MTU=9000) can provide a performance benefit (less overhead from IP/TCP headers). You must ensure compatibility throughout the end-to-end switches and links when using larger MTU sizes. The MTU frame size or Maximum Segment Size (MSS) payload might need to be adjusted across the environment. The iSCSI ports do not support fragment processing (dividing a packet). When the MTU of a switch is smaller than that of an iSCSI port, packets might be lost, and data cannot be transferred correctly. The MTU value for the switch must be the same as or greater than the MTU value for the iSCSI port. For details of MTU settings and values, see the user documentation for the switch.
- iSCSI performance can be improved by tuning iSCSI parameters such as delayed ACK. By default, delayed ACK is enabled in Hitachi VSP storage. When you use an iSCSI interface between storage systems for replication disable Delayed ACK. In addition, most hosts also benefit from disabling Delayed ACK on adapters with iSCSI traffic (see the operating system documentation). If Delayed ACK is set to Enable, the host might take a long time to recognize the volume used by GAD pairs, UR pairs, or external volumes. For example, 2,048 volumes might take 8 minutes to process. Delayed ACK can be enabled or disabled by editing the iSCSI port in Storage Navigator. or running the `raidcom modify port -delayed_ack_mode disable` command in CCI.
- Do not change the default Selected ACK setting (enabled).
- If the impact of long-distance line delay is significant, consider using devices to optimize or accelerate the WAN. WAN accelerators can provide TCP optimizations, larger window sizes over long distance, or locally acknowledge TCP segments, thus reducing round trip time and even providing optimizations at higher level protocol and application layers.
- Do not mix host and storage replication traffic on the same ports.
- Disable the spanning tree setting on the switch side because it is not compatible with Hitachi Storage iSCSI.
- The host mode option HMO 104 increases the maximum data segment size of an iSCSI PDU. Enabling HMO 104 will change from the default of 64 KB to 2 MB. This can provide performance benefits and reduced latency for GAD and UR solutions. In some environments, enabling HMO 104 improved performance and reduced latency for GAD and UR configurations. To enable HMO 104, you must enable SOM 847 in Storage Navigator. Alternatively, HMO 104 can be set using the RAIDCOM command line without setting SOM 847.

HMO 104 = ON: Login parameters are set as follows during iSCSI path creation.

- a. MaxReceiveDataSegmentLength: 2MB
- b. MaxBurstLength: 2MB

HMO 104 = OFF (default): Login parameters are set as follows during iSCSI path creation.

- a. MaxReceiveDataSegmentLength: 64KB
- b. MaxBurstLength: 64KB

However, as the transferable data size increases, the frequency of data retransmission and packet loss might increase, therefore it is also necessary to prevent that data retransmission and packet loss by tuning the networking.

- The default Maximum Window Size is 64 KB and it can be adjusted up to 1024 KB. Larger values can improve performance using high delay (long distance) and high bandwidth networks. This is a tuning value which usually requires adjustment, and results might vary between environments and applications. Validate and test values before use in a production environment. See the following table lists some approximate calculations (without protocol overhead) demonstrating that window size and round trip time are the important metrics affecting the throughput of a TCP/IP session, and link bandwidth is not a factor.

Bandwidth (Mbps)	Window Size (KiB)	Round Trip Time (ms)	Expected Throughput (MiB/sec)	Actual Throughput (MiB/sec)
1000	64	10	119.2	6.3
1000	256	10	119.2	25.0
1000	64	4	119.2	15.6
1000	256	4	119.2	62.5
1000	1024	10	119.2	100.0

## FC-NVMe

Hitachi Virtual Storage Platform (VSP) storage systems support FC-NVMe–based configurations and solutions with Red Hat Enterprise Linux. For details on supported operating systems and VSP storage models, see the [Hitachi Interoperability Matrix](#).

Red Hat Enterprise Linux provides support for NVMe over Fabrics (NVMeoF), including protocols such as NVMe over RDMA, NVMe over Fibre Channel (NVMe/FC), and NVMe over TCP. This provides high-performance connectivity between hosts and NVMe storage devices over fabric networks using the NVMe protocol. Configuration typically involves defining NVMe subsystems, namespaces, and transport parameters using tools such as `nvme-cli` and `configs`, along with kernel modules such as `nvmet`, `nvme-fabrics`, and `nvme-rdma`.

For NVMe over Fibre Channel, ensure that the Linux host is equipped with an FC HBA that supports NVMe. The recommended zoning configuration is identical to that of a standard Fibre Channel (FC-SCSI) setup.

By default, native NVMe multipathing is enabled in RHEL 9 and is the preferred multipathing method. Verify its status with the following command:

```
# cat /sys/module/nvme_core/parameters/multipath
```

Output options:

- N: Native NVMe multipathing is disabled.
- Y: Native NVMe multipathing is enabled.

If native NVMe multipathing is disabled, enable it by completing the following steps:

- Verify the current running kernel, identify the `vmlinuz` kernel file under the `/boot` directory, and then run the `grubby` command:
 

```
# uname -r
5.14.0-284.11.1.el9_2.x86_64
# grubby --args=nvme_core.multipath=Y --update-kernel /boot/vmlinuz-5.14.0-284.11.1.el9_2.x86_64
```
- Disable Device Mapper multipath if it is enabled.

3. Rebuild `initramfs` by running the following command:  

```
# dracut --force -verbose
```
4. Restart the system.

To set the I/O policy to *round-robin* persistently, create the `udev` rule file `/lib/udev/rules.d/71-nvmeiopolicy-hitachi-SVOS.rules` as follows and restart the system:

```
# Enable round-robin for Hitachi SVOS  
  
ACTION=="add", SUBSYSTEM=="nvme-subsystem", ATTR{model}=="HITACHI SVOS RF-System",  
ATTR{iopolicy}="round-robin"
```

If you change the I/O policy, verify whether *round-robin* is the active I/O policy on NVMe devices as follows:

```
# cat /sys/class/nvme-subsystem/nvme-subsys2/iopolicy  
round-robin
```

Ensure that Device Mapper multipathing is disabled and remove any existing multipath configuration. To configure RHEL FC-NVMe with Hitachi VSP storage, refer to the [RHEL FC-NVMe SAN Boot and NVMe Native Multipathing Implementation Guide](#).

NVMe over Fabrics (NVMe-oF) is an extended protocol that enables communication over a Fibre Channel fabric or an Ethernet fabric with NVMe. In a Fibre Channel environment, NVMe-oF technology supports the NVMe-oF communication protocol on FC-SAN networks using existing Fibre Channel network devices.

To configure a system consisting of a host and a storage system using the NVMe-oF communication protocol, register a logical volume of the storage system as a namespace on the NVM subsystem, and then configure a data I/O route from the host to the logical volume.

Conventional Fibre Channel and iSCSI requires an LU mapping for a port to manage an access route between the host and the logical volume. NVMe-oF, on the other hand, requires the following system components to be configured on the storage system between the host and the logical volume:

- NVM subsystem: A flash memory storage control system that supports the NVMe-oF communication protocol with one or more namespaces, and one or more communication ports (NVM subsystem ports).
- Namespace: A flash memory space formatted into a logical block.
- NVM subsystem port: A Fibre Channel port set to NVMe mode.
- Host identification (host NQN): Host name qualifier.
- Host-namespace path: Access permission to the namespace for each host NQN registered on the NVM subsystem.

Note: There is no concept of Host mode options in an NVMe subsystem.

## NVMe/TCP

NVMe over TCP (NVMe/TCP) is gaining traction because of its ability to deliver high performance and scalability using standard Ethernet networks. When implemented with Linux and VSP One Block 85 and VSP One Block 20 series storage, it provides an efficient platform for supporting critical workloads.

VSP One Block 85 and VSP One Block 20 series storage support NVMe/TCP in a Red Hat Enterprise Linux environment. For more information see the Hitachi Product Compatibility Guide at [https://compatibility.hitachivantara.com/products/interop-matrix?array\\_if=19](https://compatibility.hitachivantara.com/products/interop-matrix?array_if=19).

NVMe queues and data are managed over the standard TCP network protocol, while offering extra features such as data integrity checks (DIGEST) and Transport Layer Security (TLS) for added security. In addition, NVMe/TCP allows smooth and efficient NVMe operations between hosts and controllers over any standard IP network, with great performance and low delay. This means large data centers can use their existing Ethernet networks, including complex switch setups and regular Ethernet adapters.

See the *Configuring NVMe/TCP and Multipathing on RHEL and SUSE Linux with VSP One Block High End Storage NVM Subsystems Implementation Guide* on the Hitachi Vantara company website ([www.hitachivantara.com](http://www.hitachivantara.com)) for more information for details of configuring NVMe/TCP on an RHEL host.

#### NVMe/TCP interface specifications (100 Gbps channel boards)

Protocol Layer	Item	Specifications
All	NVMe/TCP target functions	Supported.
	NVMe/TCP initiator functions	Not supported.
	Path switching	Linux native multipathing.
	Number of connected hosts (subsystems)	255 subsystems per port.  Because the load on the NVMe/TCP port increases as the number of connections increases, connect with 255 connections per port or fewer.
Physical layer, MAC layer	Maximum transmission unit (MTU)	1500/4500/9000 bytes.
	Jumbo frames	Supported.
	Link aggregation	Not Supported.
	VLAN	Supported.
	Network Switch	L2 and L3 switches. 100 Gbps (optical): Conforms to IEEE802.3bm (100GBASE-SR4). Up to 5 stages can be cascaded. If you increase the number of cascaded switches, the number of host I/O delays increases. Keep the number of cascaded switches a minimum.
TCP/IP	TCP port number	For NVMe/TCP: <ul style="list-style-type: none"> <li>○ Discovery controller: 8009 (default)</li> <li>○ I/O controller: 4420 (default)</li> </ul> The TCP port number can be in the range of 1 to 65535. Note the following: <ol style="list-style-type: none"> <li>1. When you change the TCP port number, also change the settings of the host to be accessed.</li> <li>2. Verify whether the new number is filtered and disabled by a switch on the path.</li> <li>3. For NVMe/TCP I/O controllers, the range that can be set is 49152 to 65535</li> </ol>

		(excluding the port number used by the discovery controller).
	Window size	Supported. 64 KB, 128 KB, 256 KB, 512 KB, 1024 KB (default), 2048 KB
NVMe/TCP	Header digest, data digest	Supported. Protects communication headers and data information from errors. NVMe/TCP ports use this function according to the host-side configuration, but performance is degraded when used (the degradation rate varies depending on the host's capabilities and communication content).

See the *Provisioning Guide for VSP One Block 20 Series* on the Product Documentation portal ([docs.hitachivantara.com](https://docs.hitachivantara.com)) for NVMe/TCP port configuration details.

If using a 100-Gbps Ethernet channel board does not achieve the expected level of I/O performance, performance might improve by configuring the following parameters in hosts.

When you configure the parameters, the set values are applied to the I/O of all devices that are connected to the host. If the host is connected to devices other than an NVMe/TCP channel board, make sure that the new values of the parameters do not affect I/O of other devices before setting the values.

In an NVMe architecture, hosts communicate with storage devices through Submission Queues (SQs) and Completion Queues (CQs). SQ Flow Control ensures commands are completed in an orderly manner.

1. **With SQ Flow Control enabled:** The storage system waits for the host to send an acknowledgment (ACK) before considering a read operation complete. This means every read is dependent on the host's response time.
2. **With SQ Flow Control disabled:** The storage system can finalize the read operation immediately after sending the data (C2HData) without waiting for an explicit ACK from the host. This can reduce latency and speeds up overall performance.

For Red Hat Linux, set the following parameters.

Parameter	Value to be set
MTU	9000
SQ flow control	Disable for Red Hat Enterprise Linux.

### Prevent NVMe auto discovery from stopping after 600 seconds

In the event of a path failure, the NVM subsystem tries to reconnect to the failed controller for a specific amount of time, which is managed by the `ctrl-loss-tmo` option of the `nvme connect` command. The default value of `ctrl-loss-tmo` is 600 seconds, and beyond this time the failed path does not connect automatically. Set the `ctrl-loss-tmo` option to `-1` so the NVM Subsystem keeps trying to reconnect for an infinite time period.

To set the value of `ctrl-loss-tmo` to infinite, modify the `/etc/nvme/discovery.conf` file as shown:

```
# cat /etc/nvme/discovery.conf
# Used for extracting default parameters for discovery
#
# Example:
# --transport=<trtype> --traddr=<traddr> --trsvcid=<trsvcid> --host-traddr=<host-traddr> --
host-iface=<host-iface>
--transport=tcp --traddr=192.168.10.30 --host-traddr=192.168.10.45 --trsvcid=8009
```

```
--ctrl-loss-tmo=-1
--transport=tcp --traddr=192.168.20.30 --host-traddr=192.168.20.45 --trsvcid=8009
--ctrl-loss-tmo=-1
```

## Multipathing

Multipathing is a host-level software solution that enables a system such as a Red Hat Linux host to use multiple physical paths between itself and a storage system. While optional, multipathing is highly recommended in production environments or any setup where availability and performance are critical.

In a Fibre Channel/ iSCSI SAN environment, full fault tolerance typically involves the following:

- Two or more Host Bus Adapters (HBAs) on the Linux host.
- Redundant SAN switches.
- Dual storage controllers on the storage system, allowing alternate paths to the disk array.

There are several multipathing software options available, and selecting the most suitable one depends on the specific requirements of the environment. The following are some commonly used solutions:

- Native Linux Multipath (device-mapper-multipath)
- Hitachi Dynamic Link Manager (HDLM)
- Symantec™ Veritas Dynamic Multipathing (VxDMP)

The native Linux multipath solution is bundled with most major Linux distributions and is widely supported. Its availability at no additional cost makes it a popular choice.

In contrast, HDLM and VxDMP offer advanced features tailored to specific storage platforms and provide broader integration with various operating systems beyond Linux. These multipath solutions are preferred in environments requiring enhanced functionality or vendor-specific optimizations. To ensure consistency and avoid conflicts, only one multipathing solution should be active on a host, and the same solution must be deployed across all nodes in a cluster.

### Linux Device Mapper multipath:

Device Mapper (DM) multipath manages multiple I/O paths between server nodes and storage systems. These storage LUN I/O paths are physical SAN connections that include separate cables, switches, and storage controllers. Multipathing aggregates these I/O paths and creates a new device that consists of the combined paths. Multipathing enables load balancing across multiple paths, ensuring optimal resource use. In the event of a path failure, it automatically switches to an active path, improving system availability and reliability.

The DM multipath configuration file is typically not created during a local boot OS installation. The DM multipath configuration file, `multipath.conf`, is located in the `/etc` directory and overwrites the built-in `multipathd` configuration table.

To create an initial configuration file, run the following command:

```
/usr/sbin/mpathconf
```

To create a multipath configuration file, run the following command:

```
mpathconf --enable
```

To enable and start the `multipathd` service, run the following commands:

```
# systemctl enable multipathd
# systemctl start multipathd
```

Verify the `/etc/multipath.conf` file and customize it according to your requirements. To view the default DM multipath parameters, use the following commands:

```
# multipath -t or # multipathd -k'show config'
```

The devices section of the configuration file includes accurate vendor and product identifiers to correctly recognize the storage system:

```
vendor "HITACHI"
product "OPEN-.*"
```

For FC, the recommended value for `no_path_retry` is 10 or greater, and for iSCSI it is 30 or greater.

The following example shows a multipath configuration file for VSP storage systems.

```
# cat /etc/multipath.conf
.....
defaults {
    find_multipaths yes
    user_friendly_names yes
}
blacklist {
}
devices {
    device {
        vendor "HITACHI"
        product "OPEN-.*"
        path_grouping_policy "multibus"
        path_checker "tur"
        features "0"
        no_path_retry 10
        hardware_handler "0"
        prio "const"
        rr_weight "uniform"
    }
}
#
```

Verify the multipath status of the LUNs from the VSP storage system by running the following command:

```
# multipath -ll
mpathb (360060e80282718005080271800000002) dm-4 HITACHI,OPEN-V
size=10G features='1 queue_if_no_path' hwhandler='0' wp=rw
`-+- policy='service-time 0' prio=1 status=active
  |- 5:0:0:1 sdc 8:32 active ready running
  `-- 8:0:0:1 sdn 8:208 active ready running
```

When multipath devices (for example, `/dev/mapper/mpathb`) are available, create partitions and filesystems. Mount and configure a persistent entry in `/etc/fstab`. See the *Configuring Device Mapper Multipath on SUSE Linux for Fibre Channel and iSCSI with VSP One Block High End Storage Implementation Guide* on the Product Documentation portal ([docs.hitachivantara.com](https://docs.hitachivantara.com)) for detailed configuration information.

### Hitachi Dynamic Link Manager

Hitachi Dynamic Link Manager (HDLM) is a multipathing software solution developed by Hitachi Vantara to improve data availability and system resilience by managing paths and failover mechanisms in a heterogeneous storage environment, with Hitachi and non-Hitachi storage systems. By efficiently managing multiple paths, HDLM optimizes connectivity and availability between storage systems and host servers, such as those running Red Hat Linux. HDLM provides high availability and dynamic load balancing for connected storage, ensuring continuous access to data. When integrated with Red Hat Enterprise Linux, HDLM intelligently manages paths between the server and storage infrastructure.

Set the value of the `ErrorPathRetry` parameter for HDLM to 5 or greater.

Set the value of the `ErrorPathDelay` parameter for HDLM to 30 or greater.

**Note:** Set the number of I/O attempts and the I/O attempt latency on a failed path in the relevant parameter. Although it takes 130 seconds to restart during a firmware update of the channel board or a reset when a failure occurs, the hosts and network processing time needs to be taken into account. For this reason, set 5 times for the number of I/O attempts and 30 seconds for the I/O attempt latency on a failed path, so that the time is 150 seconds.

**Procedure:**

1. Set the `ErrorPathRetry` parameter to 5 using the HDLM driver option setting utility (`dlnmsetopt`).

The following command output example shows that HDLM is installed under `/opt`.

```
# /opt/DynamicLinkManager/bin/dlnmsetopt -epr 5
```

**Command output:**

```
KAPL12554-I The utility for setting HDLM driver option has started.
KAPL12555-I The utility for setting HDLM driver option completed normally.
KAPL12558-I Please restart the computer so that the option settings take effect.
```

2. Set the `ErrorPathDelay` parameter to 30 using `dlnmsetopt`.

```
# /opt/DynamicLinkManager/bin/dlnmsetopt -epd 30
```

**Command output:**

```
KAPL12554-I The utility for setting HDLM driver option has started.
KAPL12555-I The utility for setting HDLM driver option completed normally.
KAPL12558-I Please restart the computer so that the option settings take effect.
```

3. Restart the host OS.

The following parameter configurations are recommended to ensure efficient and reliable path management:

```
/usr/DynamicLinkManager/bin/dlnkmgr set -pchk on -intvl 3
/usr/DynamicLinkManager/bin/dlnkmgr set -afb on -intvl 1
/usr/DynamicLinkManager/bin/dlnkmgr set -lb on -lbtype rr
```

See the *Dynamic Link Manager (for Linux®) User Guide* on the Product Documentation portal ([docs.hitachivantara.com](https://docs.hitachivantara.com)) for configuration details..

### DM multipath ALUA configuration for GAD

In an RHEL environment with GAD configured, ALUA is used to ensure that hosts prefer local-site Active-Optimized paths and treat remote-site paths as Active-Non-Optimized.

Enable ALUA mode on GAD volumes when a cross-path configuration exists, allowing the host to manage preferred paths and recognize ALUA path state changes.

Linux device-mapper multipath determines path priorities through the ALUA prioritizer, which assigns higher priority to Active-Optimized paths and lower priority to Non-Optimized paths.

To ensure that ALUA priorities form separate path groups, Linux multipath must use `group_by_prio` rather than `multibus`, which otherwise forces all paths into one group.

To enable ALUA-based path prioritization for Hitachi VSP storage in a stretch cluster GAD configuration, update `/etc/multipath.conf` as follows:

```
defaults {
    user_friendly_names yes
```

```

        find_multipaths on
    }

blacklist {
}
devices {
    device {
        vendor "HITACHI"
        product "OPEN-.*"
        path_grouping_policy "group_by_prio"
        prio "alua"
        path_checker "tur"
        features "0"
        no_path_retry 10
        hardware_handler "1 alua"
        failback "immediate"
        rr_weight "uniform"
    }
}

```

The following are key ALUA parameters:

- The `path_grouping_policy` parameter `group_by_prio` determines how multipath groups paths. It creates one path group per path priority value. The priorities are determined by the `prio` attribute.
- The `prio` parameter `alua` generates path priority based on SCSI-3 ALUA settings reported by the storage system.
- The `hardware_handler` parameter `1 alua` activates ALUA support at the device level.
- The `failback` parameter `immediate` automatically switches back to the optimized path when it becomes available.

Update the `multipath.conf` file with required ALUA parameters and restart the host.

If the setting is changed from non-ALUA to ALUA or vice versa, a host restart is required for ALUA settings to be recognized by the host.

## Expand the filesystem

The following outlines the recommended best practices steps to perform an online LUN expansion when LUNs are presented from a Hitachi Virtual Storage Platform storage system to an RHEL host.

Take a snapshot of the LUN before expanding it. Filesystem resizing activities inherently carry a potential risk of data corruption or loss. It is strongly advised that you create a snapshot of the volume and reliable backups are available before proceeding with these operations. Do not change the LUN ID or host mappings. Do not modify host Group settings.

```

[root@localhost ~]# multipath -ll
mpatha (360060e80341180000091118000000510) dm-1 HITACHI,OPEN-V
size=10G features='1 queue_if_no_path' hwhandler='0' wp=rw
`--+ policy='service-time 0' prio=1 status=active
  |- 0:0:0:1 sdb 8:16 active ready running
  `-- 2:0:0:1 sdd 8:48 active ready running

```

Expand the LUN on the storage system using VSP One Block Administrator or Storage Navigator, depending on the VSP model.

Rescan SCSI devices on the RHEL system.

```
rescan-scsi-bus.sh -resize
```

Scan for a particular storage device.

```
echo 1 > /sys/block/sdX/device/rescan
```

Reload multipath devices.

```
[root@localhost ~]# multipathd -k'resize map 360060e80341180000091118000000510'
ok
```

Verify the new size.

```
[root@localhost ~]# multipath -ll
mpatha (360060e80341180000091118000000510) dm-1 HITACHI,OPEN-V
size=15G features='1 queue_if_no_path' hwhandler='0' wp=rw
`-+- policy='service-time 0' prio=1 status=active
  |- 0:0:0:1 sdb 8:16 active ready running
  `-- 2:0:0:1 sdd 8:48 active ready running
```

Rescan all paths, not just one. Always resize using WWID, not alias names, and ensure all paths show active ready running status.

Resize the partition. If a filesystem is in the partition, resize the partition using the `growpart` command or use `fdisk` manually.

Extend the filesystem online. Use the following for xfs filesystems, and for ext4 use the `resize2fs` command.

```
[root@localhost ~]# df -Th |grep mpath
Filesystem      Type      Size  Used Avail Use% Mounted on
/dev/mapper/mpath1 xfs        10G   228M   9.8G   3% /fs0
/dev/mapper/mpathb1 xfs        10G   228M   9.8G   3% /fs1
```

```
[root@localhost ~]# xfs_growfs /fs0
meta-data=/dev/mapper/mpath1      isize=512    agcount=4, agsize=655296 blks
=                               sectsz=512   attr=2, projid32bit=1
=                               crc=1       finobt=1, sparse=1, rmapbt=1
=                               reflink=1   bigtime=1 inobtcount=1 nnext64=1
=                               exchange=0
data      =                       bsize=4096  blocks=2621184, imaxpct=25
=                               sunit=0    swidth=0 blks
naming    =version 2                   bsize=4096  ascii-ci=0, ftype=1, parent=0
log       =internal log           bsize=4096  blocks=16384, version=2
=                               sectsz=512  sunit=0 blks, lazy-count=1
realtime  =none                    extsz=4096  blocks=0, rtextents=0
data blocks changed from 2621184 to 3931904
```

```
root@localhost ~]# df -Th |grep mpath
Filesystem      Type      Size  Used Avail Use% Mounted on
/dev/mapper/mpath1 xfs        15G   326M   15G   3% /fs0
/dev/mapper/mpathb1 xfs        15G   326M   15G   3% /fs1
```

Validate data integrity:

```
root@localhost ~]# cd /fs0
[root@localhost fs0]# ls
testfile
```

## Boot from SAN

This section describes SAN boot on RHEL systems with DM multipath.

To configure a Hitachi Virtual Storage Platform (VSP) volume as a bootable device on a Linux system, ensure that the target volume is presented to the host with LUN ID 0. This is a prerequisite for the system to recognize and boot from the volume correctly.

Perform following steps to boot from SAN.

1. Install a supported FC HBA in the server and establish connectivity to the FC port of the VSP storage system either directly or through a SAN switch. HBA ports can be connected directly to the FC ports of the VSP storage system or through intermediate SAN FC switches.
2. Configure port topology and LUN security. For configuration information, see the *VSP One Block Storage Provisioning Guide* on the Product Documentation portal (docs.hitachivantara.com).
  - Fabric: ON when connected to a SAN switch.
  - Fabric: OFF when connected directly to an HBA.
  - Connection Type: P-to-P for both direct and fabric connections.
3. Create host groups in the storage ports and assign the same set of LUNs to all host groups. Ensure that the Boot LUN size meets the OS installation requirements.
4. Configure the HBA BIOS. Follow the recommendations provided by the HBA vendor and ensure that boot from SAN is enabled.
5. Verify that the designated SAN Boot LUN from the VSP storage system is detected in the HBA BIOS.
6. Install RHEL OS on a LUN ID 0 provisioned from the VSP storage system.
7. Verify that the storage LUN is detected as a multipath enabled during OS installation.

The DM Multipath configuration file is created during SAN Boot installation as shown:

```
# cat /etc/multipath.conf
defaults {
    find_multipaths yes
    user_friendly_names yes
}
blacklist {
}
#
```

Verify that SAN LUN is detected as a multipath device after OS installation.

```
# multipath -ll
mpatha (360060e8028271800508027180000001d) dm-0 HITACHI,OPEN-V
size=100G features='0' hwhandler='0' wp=rw
`-+- policy='service-time 0' prio=1 status=active
|- 4:0:0:0 sda 8:0 active ready running
`- 7:0:0:0 sdb 8:16 active ready running
#
```

8. Update the `multipath.conf` file for the VSP storage system as shown:

```
# cat /etc/multipath.conf
defaults {
    find_multipaths yes
    user_friendly_names yes
}
blacklist {
}
devices {
```

```

device {
    vendor "HITACHI"
    product "OPEN-.*"
    path_grouping_policy "multibus"
    path_checker "tur"
    features "0"
    no_path_retry 10
    hardware_handler "0"
    prio "const"
    rr_weight "uniform"
}
}

```

Rebuild the `initramfs` file system using the `dracut` command. Restart the server after the DM Multipath configuration is updated and `initramfs` is rebuilt. The following example shows the DM Multipath status:

```

# multipath -ll
mpatha (360060e8028271800508027180000001d) dm-0 HITACHI,OPEN-V
size=100G features='1 queue_if_no_path' hwhandler='0' wp=rw
`-+- policy='service-time 0' prio=1 status=active
  |- 4:0:0:0 sda 8:0 active ready running
  `-- 7:0:0:0 sdb 8:16 active ready running
#

```

See *Configuring Device Mapper Multipath on Red Hat and SUSE Linux with VSP One Block 20 Storage Systems* on the Hitachi Vantara company website ([www.hitachivantara.com](http://www.hitachivantara.com)) for detailed device-mapper configuration details.

## Hybrid Cloud

Hitachi storage systems are supported with Red Hat Enterprise Linux (RHEL) across multiple hybrid cloud environments, including Amazon Web Services (AWS), Google Cloud Platform (GCP), and Microsoft Azure. Supported configurations cover a range of RHEL versions (7.7 to 9.4) and Hitachi Virtual Storage Platform (VSP) models such as VSP 5600, VSP E1090, VSP One Block 20 series, and VSP One Block 85.

These configurations support modern network and storage interfaces such as 10 Gbps/25 Gbps iSCSI and 100 Gbps NVMe over TCP, enabling high-performance and scalable deployment options.

For a complete list of validated and supported configurations, see the Product Compatibility Guide at [https://compatibility.hitachivantara.com/products/hybrid-cloud-interoperability?operating\\_system=45,297,361,382,393,450](https://compatibility.hitachivantara.com/products/hybrid-cloud-interoperability?operating_system=45,297,361,382,393,450).

In hybrid cloud environments, configuration parameters must be set to ensure reliable I/O performance and seamless interoperability.

### Recommendations for iSCSI Configuration:

- MTU: 1500
- Set the SCSI timeout to 60 for iSCSI devices:  
`/sys/block/<device name>/device/timeout = 60`
- Multipath configuration (`/etc/multipath.conf`), set `no_path_retry = 30`
- If the CHB TYPE is 25 Gbps, a fixed speed setting (25 Gbps/10 Gbps) is recommended.
- For Microsoft Azure with RHEL 9.4, set the `OS.RootDeviceScsiTimeout` parameter (a SCSI timeout value) in a `WALinuxAgent` file to 60 seconds. The parameter overwrites the SCSI timeout value set in `/sys/block/sdX/device/timeout`. If the parameter is still set to the default value (300 seconds), an I/O hang can occur if there is a storage hardware failure.

- You need to use an IP switch that complies with the following standards, depending on the channel board that is used.
  - 100 Gbps (Optic): IEEE802.3bm (100GBASE-SR4 QSFP28) compliant
  - 25 Gbps (Optic): IEEE802.3by (25GBASE-SR SFP28) compliant

#### Recommendations for NVMe over TCP Configuration:

- Use NVMe native multipathing. NVMe multipathing is enabled by default in Red Hat Enterprise Linux 9. If it is not enabled, See *Configuring NVMe/TCP with AWS EC2 Instance Running Red Hat Enterprise Linux on VSP One Block 20 NVM Subsystems* on the Hitachi Vantara company website ([www.hitachivantara.com](http://www.hitachivantara.com)) for NVMe native multipathing configuration details.
- Depending on the Linux version, the default timeout value might already be set at the maximum value. For an AWS EC2 instance with Red Hat Linux 9.4 (ami-0c5ebd68eb61ff68d), the default I/O timeout for NVMe devices is set to 4294967295 seconds. Because of the high timeout value, an I/O hang might occur when there is a storage hardware failure in an NVMe/TCP configuration. For uninterrupted operation in the event of a storage hardware failure, you must configure the I/O timeout value to a low value (for example, 60 seconds). The optimal I/O timeout value for NVMe core is 60 seconds according to VSP guidelines.

Set the `nvme_core.io_timeout` parameter (which is an I/O timeout value for NVMe devices) using the following command.

```
# grubby --args=nvme_core.io_timeout=<I/O timeout value> --update-kernel  
/boot/<vmlinuz-file>
```

Verify the I/O time out value:

```
# cat /sys/module/nvme_core/parameters/io_timeout 60
```