Configuring SR-IOV and SIOV in VMware vSphere on Lenovo ThinkSystem Servers

Planning / Implementation

With the rapid development of cloud computing and virtualization technology, the computing environment for businesses has started to use virtual machines to improve resource utilization and flexibility. It creates the need for efficient I/O virtualization technologies. Traditional approaches to I/O virtualization usually rely on the software layer to manage the sharing of I/O devices, which can lead to low performance and high latency. Single Root I/O Virtualization and Sharing (SR-IOV) and Intel Scalable I/O Virtualization (SIOV) are two approaches to address these challenges by hardware-assisted I/O virtualization.

Single Root I/O Virtualization and Sharing (SR-IOV) is an optional feature of the PCIe standard that enables hardware-assisted I/O virtualization, allowing physical I/O devices to be shared across multiple VMs or containers. An SR-IOV capable device has a Physical Function (PF) and multiple Virtual Functions (VFs). Device sharing is achieved by assigning VFs to different VMs or containers. The PF is responsible for managing and configuring VFs by the driver running in the hypervisor.

Intel Scalable I/O Virtualization (SIOV) is also an approach to hardware-assisted I/O virtualization developed by Intel. SIOV enables high scalability and high-performance sharing of I/O devices across VMs or containers by further granulating I/O devices.

The difference between SR-IOV and SIOV

A Single Root I/O Virtualization and Sharing (SR-IOV) device comprises a Physical Function (PF) and multiple Virtual Functions (VFs). As illustrated in the following figure, the PF manages and configures the device, while VFs provide I/O functionality without configuration or management capabilities. The number of VFs of SR-IOV devices is limited and the allocated resources are fixed. It is preset at device design time; users can only enable a specific number of VFs within this predefined range.

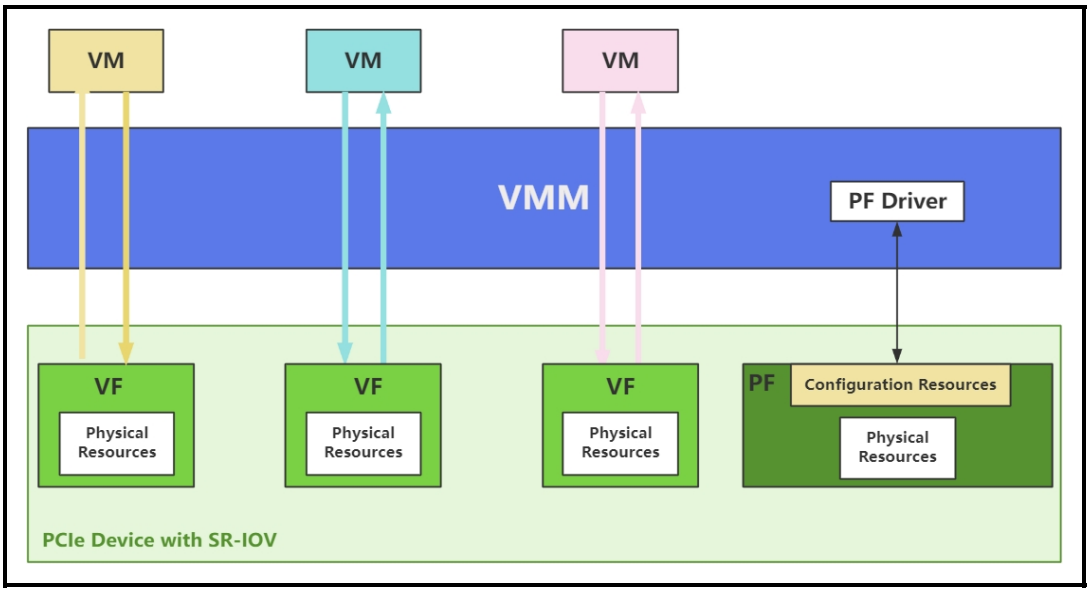

Figure 1. Single Root I/O Virtualization

Intel Scalable I/O Virtualization (SIOV) is a new hardware-assisted IO virtualization solution, the structure is shown in the figure below. The resources of SIOV devices are divided into finer granularities, with the type and number of backend resources grouped to compose an Assignable Device Interface (ADI). An ADI refers to the set of device backend resources that are allocated, configured, and organized as an isolated unit, forming the unit of device sharing. ADI composing a virtual device (VDEV) is assigned to the VM. Each VDEV may be backed by one or more ADIs. A VDEV may be composed of a static number of ADIs that are pre-allocated at the time of VDEV instantiation or composed dynamically in response to guest driver requests to allocate/free resources.

SR-IOV is a mature standard that is widely supported, and easy to configure and use. While SIOV offers higher flexibility and scalability, SIOV is relatively new and less standardized.

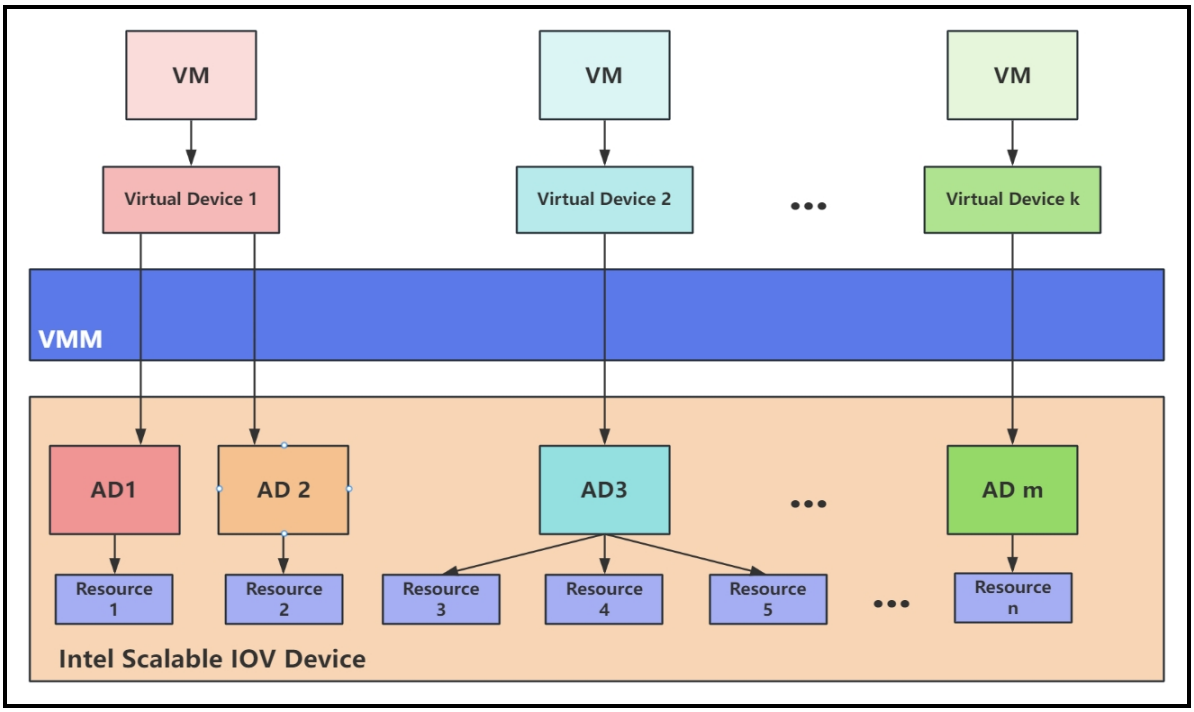

Figure 2. Intel Scalable I/O Virtualization

All in all, SR-IOV and SIOV have their own benefits and limitations, and their differences are summarized in the table below.

Table 1. Comparing SR-IOV and SIOV

Feature	SR-IOV	SIOV
Granularity	Limited to predefined Virtual Functions (VFs).	Offers finer granularity through Assignable Device Interfaces (ADIs).
Scalability	Fixed number of VFs determined during device design.	High scalability with dynamic ADI allocation.
Flexibility	Lacks flexibility; VFs are static.	Highly flexible; ADIs can be allocated/reallocated dynamically.
Configuration Ease	Widely supported and easier to configure.	Requires more complex setup and newer hardware/software.
Adoption Status	Mature and standardized.	Emerging and less standardized.
Use Cases	Common in traditional virtualization scenarios.	Ideal for AI, machine learning, and dynamic workloads.

Implementing SR-IOV

This section describes the steps to configure and enable SR-IOV in VMware ESXi 8.0 U3 on the Lenovo ThinkSystem Server. In our lab tests, we are using the Broadcom 57454 10GBASE-T 4-port PCIe Ethernet Adapter as an example. This adapter has its own SR-IOV switch in the BIOS, so it it is well-suited as an example.

The test configuration of ThinkSystem SR650 V3 is listed in the following table.

Table 2. SR-IOV test configuration

Component	Configuration
Server	ThinkSystem SR650 V3 Server
CPU	2x Intel Xeon Platinum 8468 processors
Memory	2x DDR5 4800MHz 16GB DIMMs
HDD	4.0TB SATA HDD
Host	ESXi 8.0 U3 Custom Image for Lenovo ThinkSystem
Guest OS	RHEL 9.3
SR-IOV Device	Broadcom 57454 10GBASE-T 4-port PCIe Ethernet Adapter

Prerequisites for SR-IOV

To use SR-IOV, the system must meet the following requirements:

- Supported System: Intel VT for Directed I/O (VT-d) or AMD Virtualization (AMD-V).
- The device must be supported SR-IOV by ESXi: To check supported devices, see the VMware Compatibility Guide and select SR-IOV under Features as a search option.
- The vSphere ESXi host supports a limited number of VFs and VMDirectPath PCI/PCIe devices. Refer to the configuration maximum guide for more information: https://kb.vmware.com/s/article/1003497.

Enabling SR-IOV in UEFI

The process to enable SR-IOV in UEFI is as follows.

1. Power on the Lenovo ThinkSystem server, then press F1 to enter System Setup.
2. For the Intel platform: Enable Intel VT-d and SR-IOV, and reboot. Select **System Settings > Devices and I/O Ports**, ensure that **Intel VT for Directed I/O (VT-d)** and **SRIOV** options are enabled as shown in the following figure.

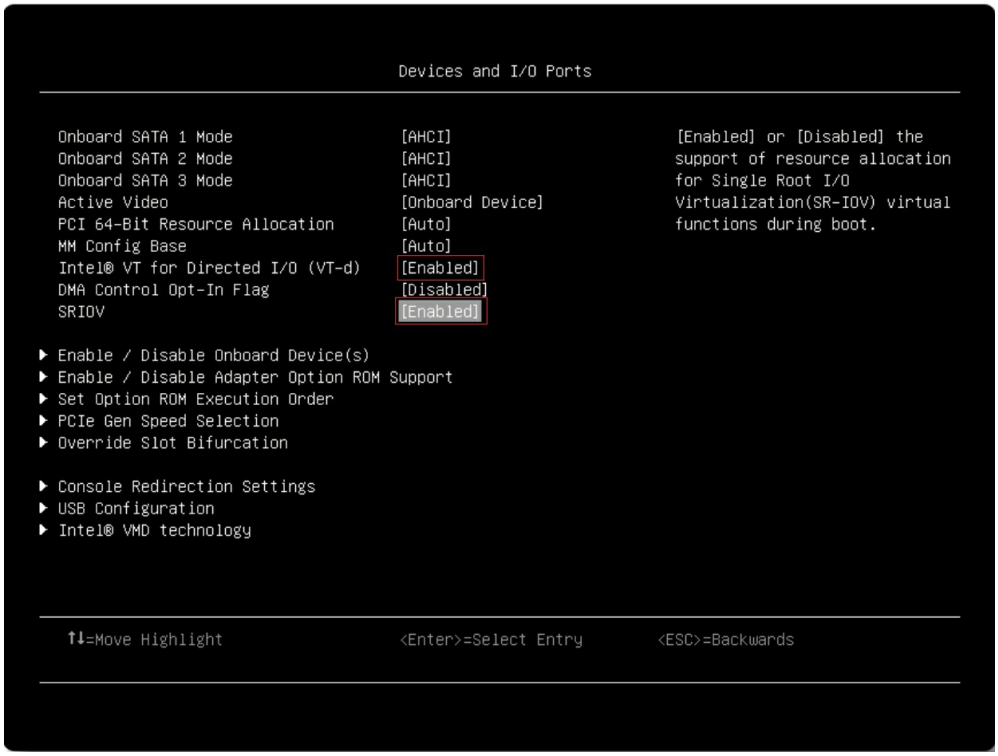

Figure 3. Enabling VT-d and SRIOV in UEFI

3. For the AMD platform: Enable SVM Mode, SR-IOV, IOMMU, and reboot. Select **System Settings > Devices and I/O Ports** and ensure that **IOMMU** and **SRIOV** options are enabled as shown in the following figure.

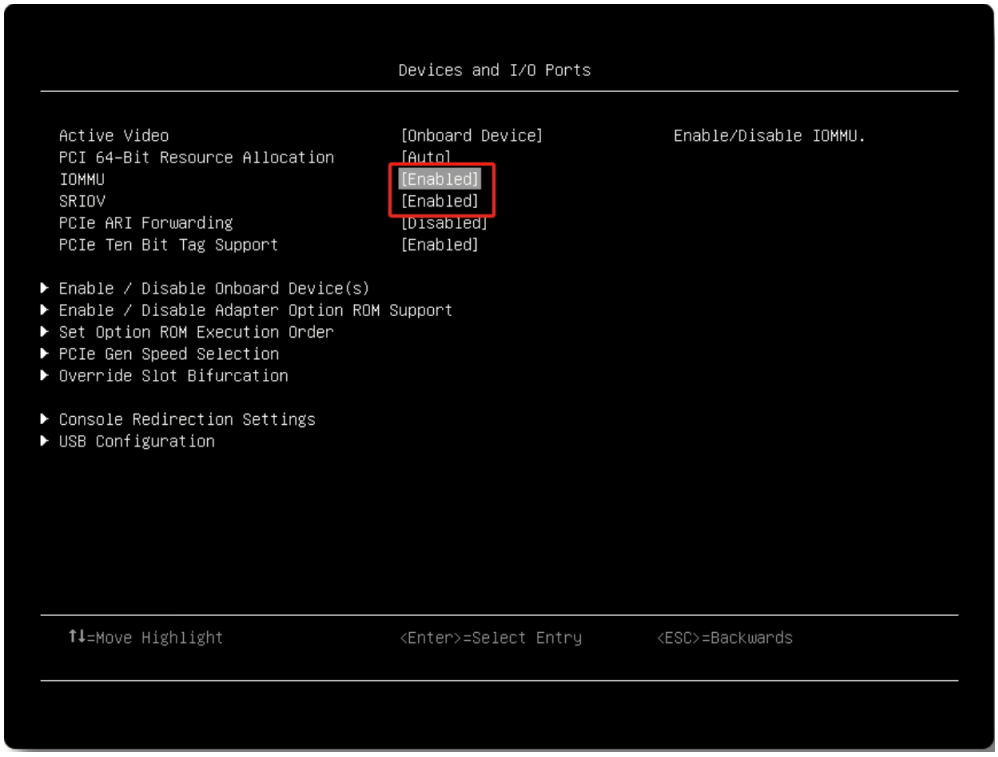

Figure 4. Enabling IOMMU and SRIOV in UEFI

4. Select **System Settings** > **Processors** and ensure that the **SVM Mode** option is enabled as shown in the following figure.

Figure 5. Enabling SVM Mode in UEFI

5. Some devices have a separate SR-IOV option in UEFI, make sure it is enabled. For example, the device

Broadcom 57454 10GBASE-T 4-Port PCIe Ethernet Adapter. You can enter in **System Settings** > **Network** > **Broadcom 57454 10GBASE-T 4-Port PCIe Ethernet Adapter** > **Device Configuration Menu** to find its SR-IOV Settings, and ensure it is enabled.

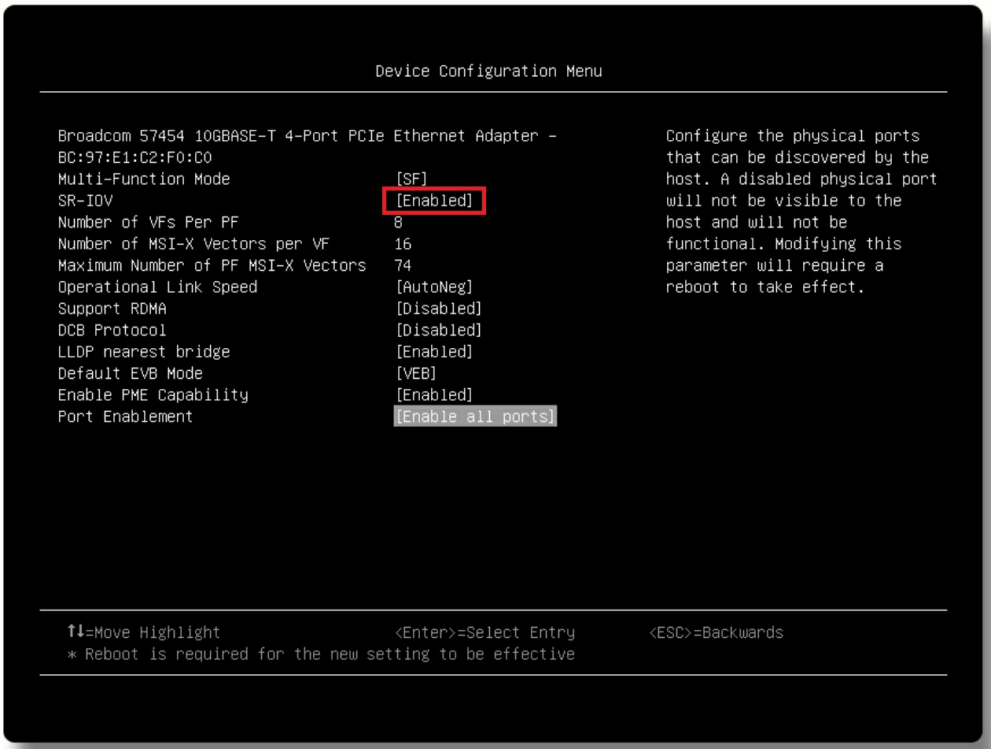

Figure 6. Enabling SR-IOV in the adapter settings in UEFI

6. Save settings and reboot.

Configuring SR-IOV in ESXi

After enabling SR-IOV in UEFI Settings, boot into ESXi and enable SR-IOV according to the following steps.

1. Log in to the vSphere client.
2. Select **Manage** > **Hardware** > **PCI Devices** from the left navigation menu. Select the SR-IOV device, the SR-IOV item should be "**Disabled**" and select "**Configure SR-IOV**". If the item is "Not Capable", please check the configuration of the system and UEFI settings.

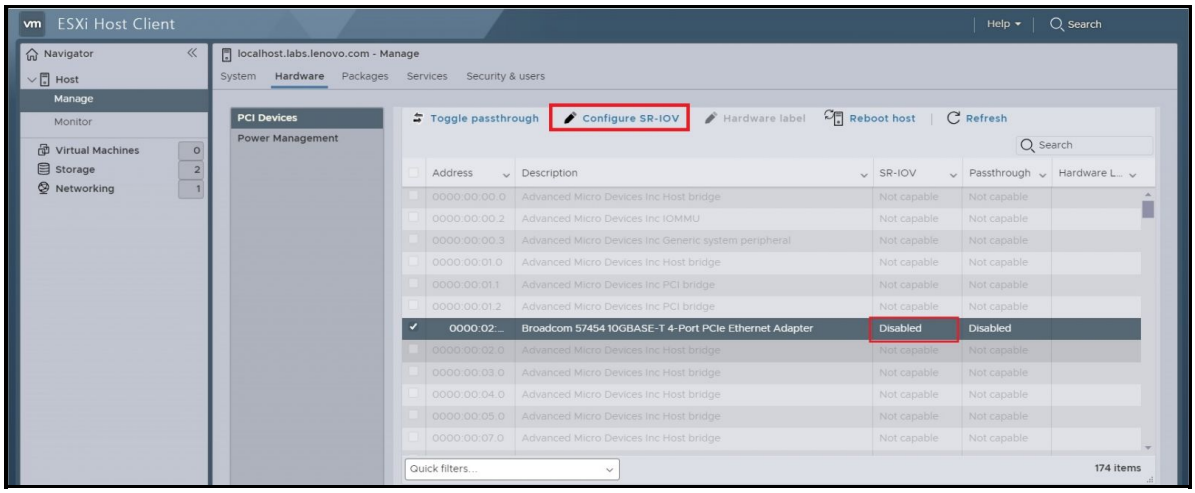

Figure 7. Configure SR-IOV in ESXi client

3. As shown in the figure below, set the **Enabled** option to **Yes** and enter the desired number of virtual functions between 1 and the maximum value shown in the window.

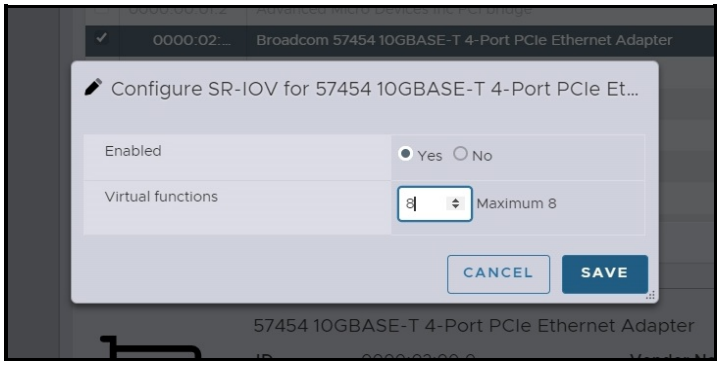

Figure 8. Enabling SR-IOV and set the number of VFs

4. Save the Settings and reboot for the configuration to take effect.
5. Navigate to the PCI Device menu again to ensure the VF devices are on the list, as shown below. The number of VF devices is equal to the number set in Step 3. The function of each VF device is identical, with no priority order; The only difference is their PCI device ID, which you can use to distinguish them.

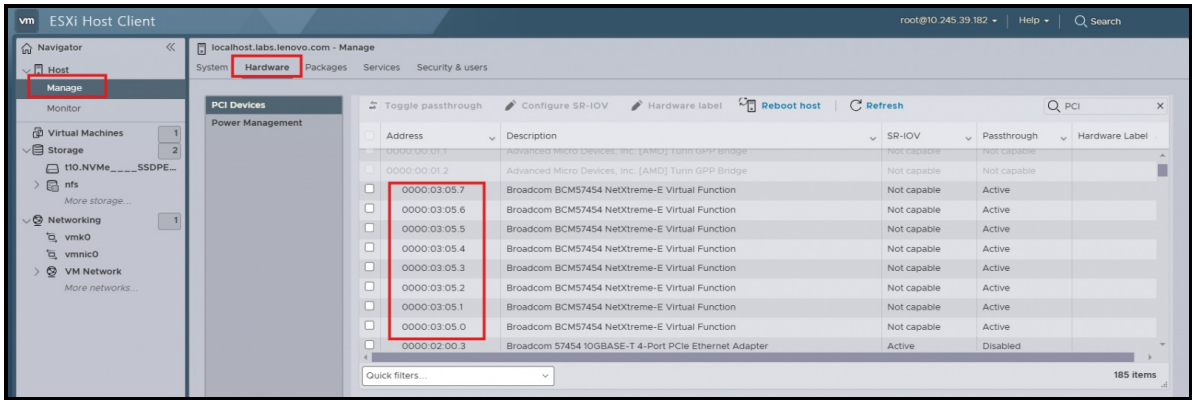

Figure 9. Verify SR-IOV is active

6. Passthrough the VF to Virtual Machine (VM) as shown below. Choose a VM, click the **Edit** > **Add other device** > **PCI device**.

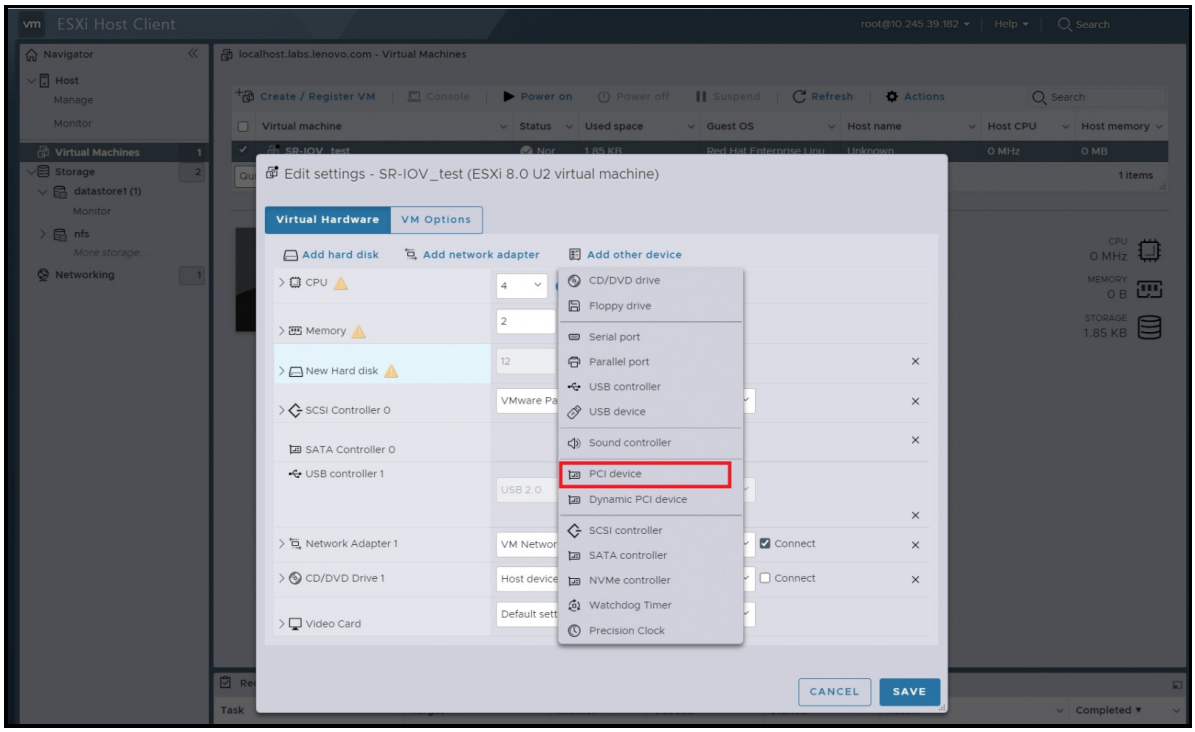

Figure 10. Add a PCI device

7. You can see the VF devices that just appeared in the PCI device list under the New PCI device option as shown below. Determine the VF device to use based on the name and PCI device ID. In our lab environment, the SR-IOV device is the BCM57454 and we want to select VF0, so we select **BCM57454 NexXtreme-E Virtual Function-000:03.02.0** as highlighted in the figure. Select it and save the settings.

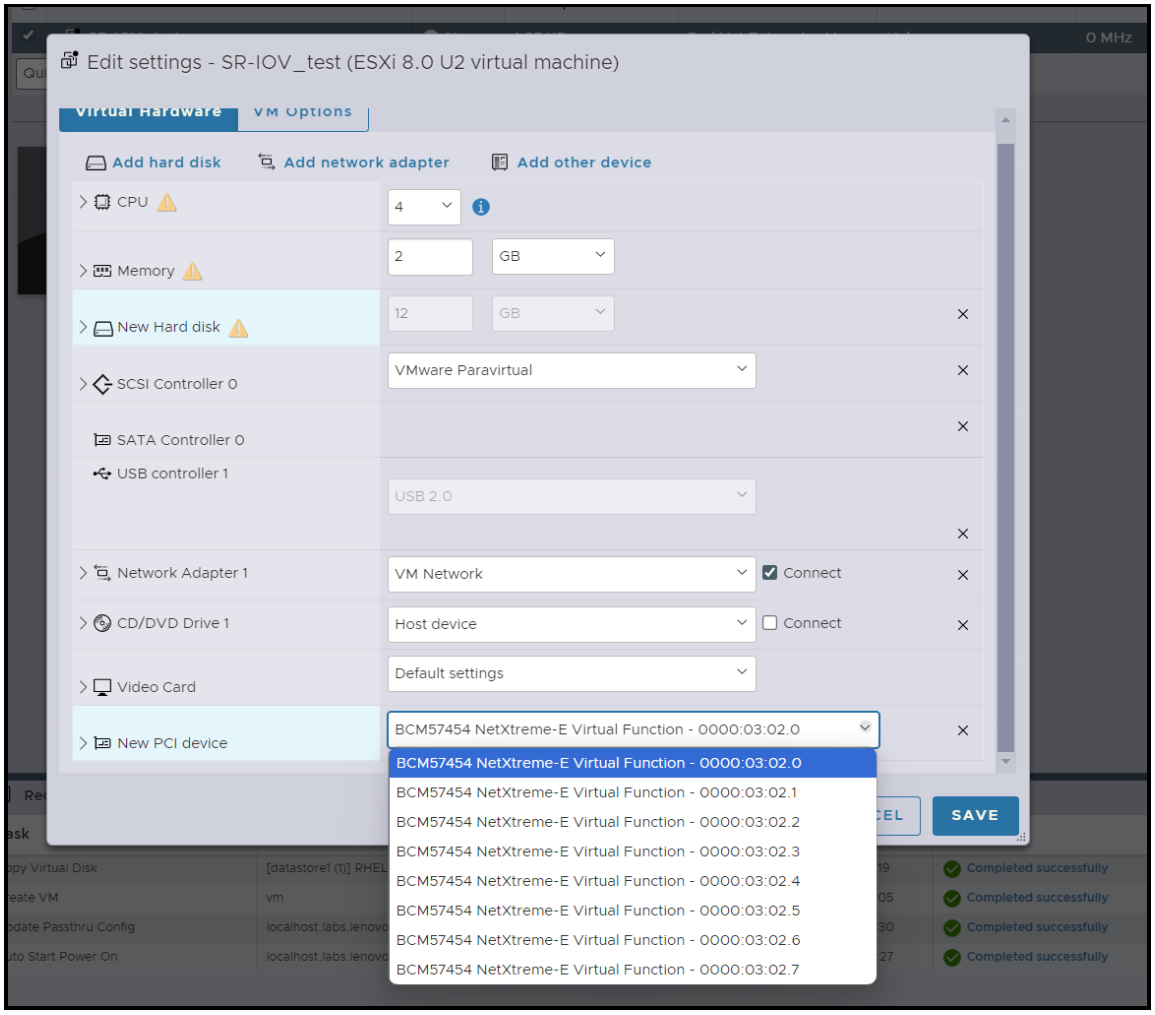

Figure 11. Select the VF device

8. Power on the VM and boot into the guest OS. You can see the VF in the PCI device list using the `lspci` command.

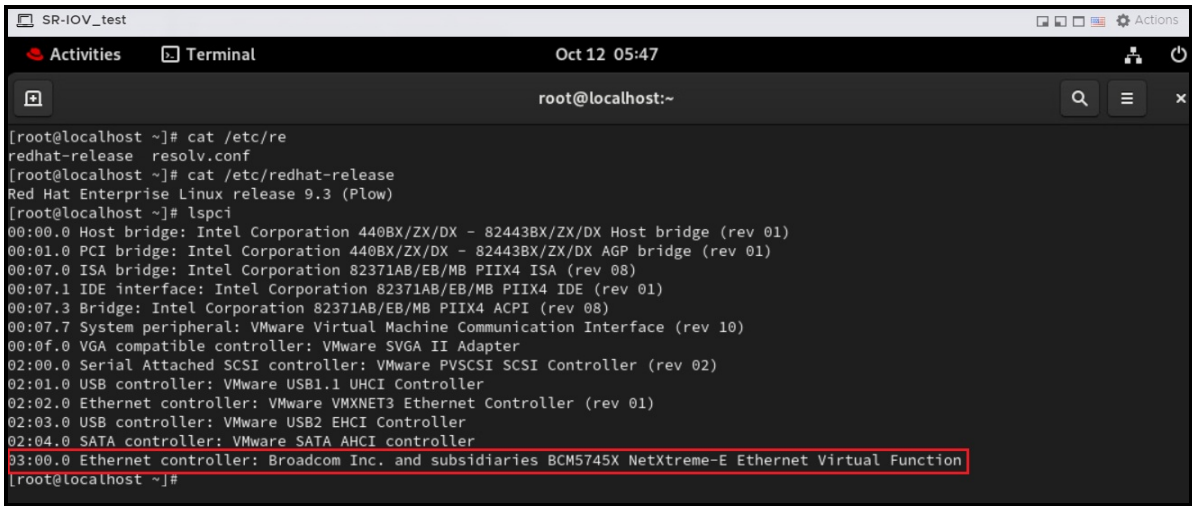

Figure 12. Verifying the VF device

Implementing SIOV

This section describes the steps to configure and enable SIOV and passthrough VDEV to the virtual machine in VMware ESXi 8.0 U2 on the server.

In our lab tests, we are using the Intel QuickAssist Technology (Intel QAT) accelerator as an example. Intel QAT is PCIe device that provides cryptographic and compression acceleration. It enables sharing of a Physical Function across multiple guest Virtual Machines (VM) using SIOV.

The test configuration of ThinkSystem SR650 V3 is listed in the following table.

Table 3. SIOV test configuration

Component	Configuration
Server	ThinkSystem SR650 V3 Server
CPU	2x Intel Xeon Platinum 8592+ processors
Memory	16x DDR5 4800MHz 16GB DIMMs
HDD	4.0TB SATA HDD
Host	ESXi 8.0 U2 Custom Image for Lenovo ThinkSystem
Guest OS	Ubuntu 22.4
SIOV Device	Intel QuickAssist Technology (Intel QAT) accelerator

Prerequisites for SIOV

ESXi has these limitations when using SIOV:

- The device must be supported SIOV by ESXi: Check with the adapter vendor's web site to confirm that the device supports SIOV in ESXi.
- The vSphere ESXi host supports a limited number of PCI passthrough devices. Refer to the configuration maximum guide for more information: https://kb.vmware.com/s/article/1003497.
- The virtual IOMMU option for VM is not compatible with VDEVs on ESXi.
- vSphere Host Client doesn't support SIOV setting. You need VMware vCenter to complete the configuration of SIOV devices.

Enabling SIOV in UEFI

To enable SIOV in UEFI, follow the same steps to enable SR-IOV as described in Enabling SR-IOV in UEFI. The steps remain consistent across all devices, except Step 5. Step 5 depends on the specific device.

1. Please skip Step 5 if the device does not have its own SR-IOV switch in BIOS.
2. The configuration entry is device-specific if the device has its own SR-IOV switch in BIOS.

Consult your device manufacturer for information on whether the device has its own SR-IOV switch and how to configure it.

Configuring SIOV in ESXi

The following steps build on the assumption that Intel QAT devices are already available on ESXi. For more information on Intel QAT device driver selection and installation, see the Lenovo Press paper, Implementing the Intel DLB and Intel QAT Accelerators on ThinkSystem Servers Running VMware ESXi.

The steps are as follows:

1. Log in to the VMware vCenter via vSphere Client.

   Note: vSphere Host Client doesn't support VDEVs assignment to VM via VM settings.

2. Locate the target VM in the inventory. Make sure the VM is powered down.
3. In the center pane, click **Edit Settings** to edit VM configuration. A pop-up window with the VM settings will appear.
4. All memory for the VM must be reserved. Expand **Memory** and check that Reservation section is reporting the following (see the figure below):
   `All VM memory is reserved for this VM.`

   If you do not see this message, check the **Reserve all guest memory (All locked)** checkbox.

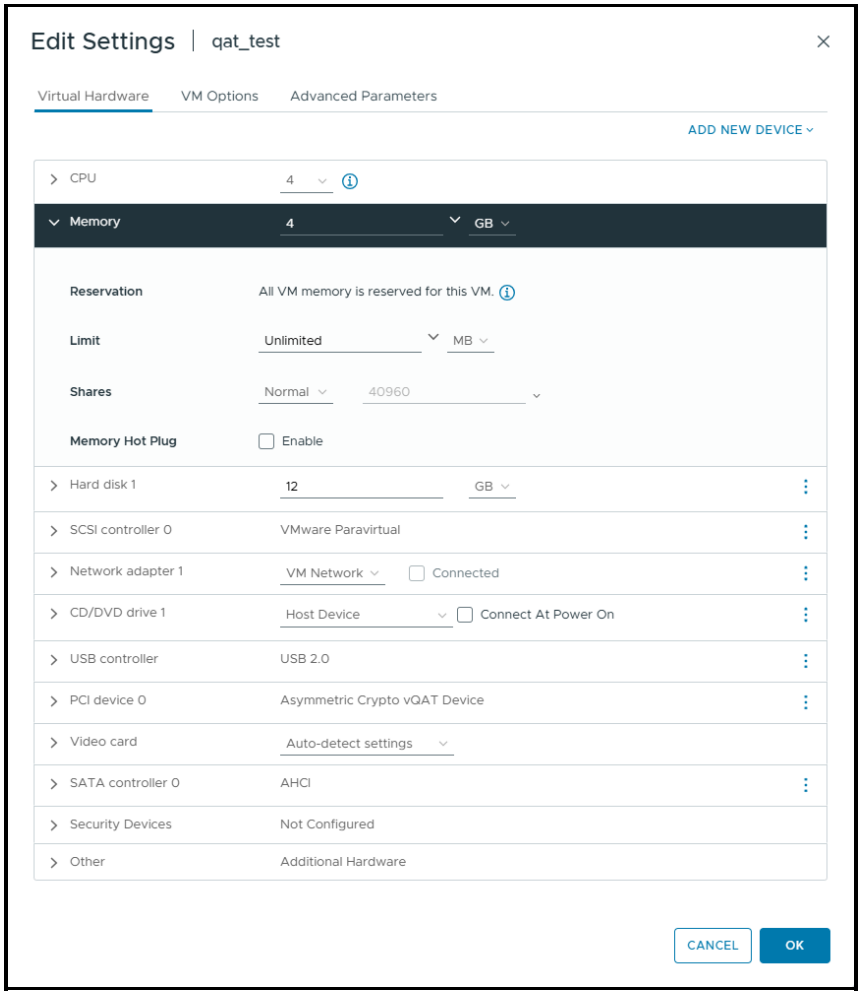

Figure 13. Edit VM settings

5. Click the **Add New Device** dropdown menu and choose **PCI Device** as shown below.

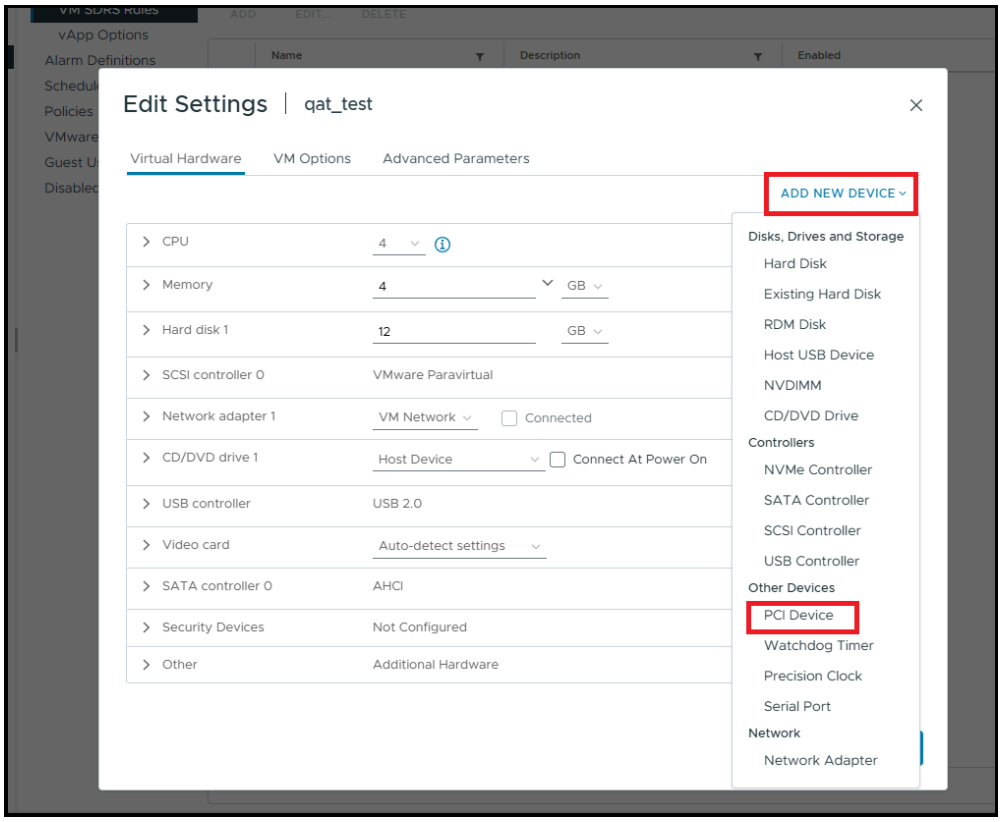

Figure 14. Add New Device menu

6. The available VDEV devices are listed as shown below. Select the VDEV device that belongs to your SIOV device by name and apply to continue. Each SIOV device has its own VDEV name, which depends on the manufacturer of the device. If you do not know, consult the manufacturer of the SIOV device. In our lab configuration, the SIOV Device is Intel QAT and its VDEV device name is **Asymmetric Crypto vQAT Device**.

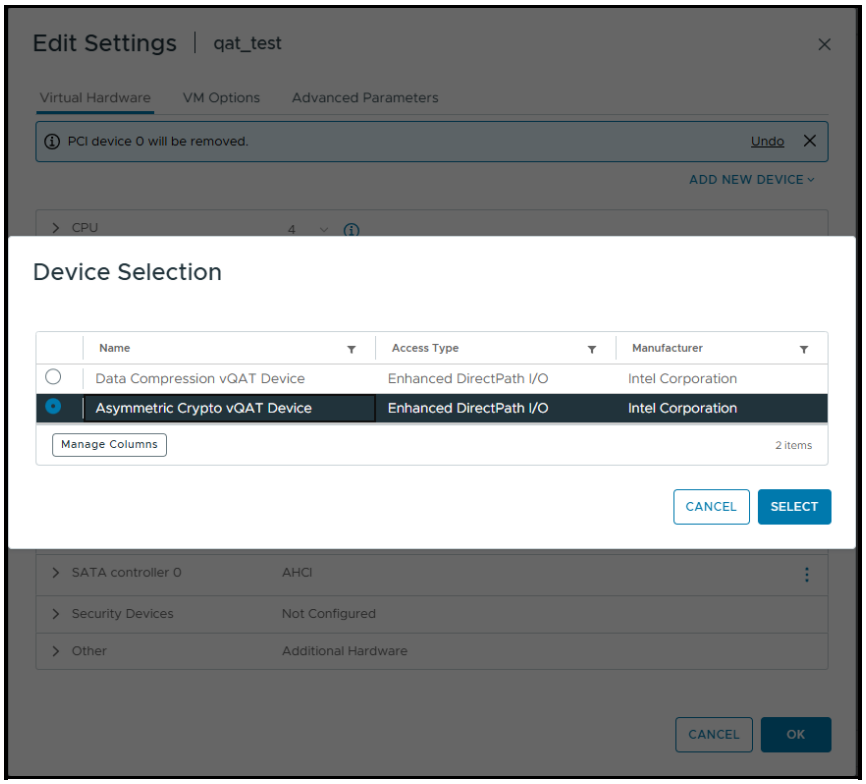

Figure 15. Select the VDEV device

7. Click **OK** to save the VM configuration and power on the VM.
8. Log in to the guest OS.
9. Use the `lspci` command to view PCI devices including VDEV devices. The VDEV PCI device name and PCI device ID are specified by the device manufacturer. If you do not know, please consult your SIOV device manufacturer. In this case, the PCI Device identification for VDEV is **03:00.00: Intel Corporation Device 0da5**.

```
root@ubuntu22:/QAT# lspci
00:00.0 Host bridge: Intel Corporation 440BX/ZX/DX - 82443BX/ZX/DX Host bridge (rev 01)
00:01.0 PCI bridge: Intel Corporation 440BX/ZX/DX - 82443BX/ZX/DX AGP bridge (rev 01)
00:07.0 ISA bridge: Intel Corporation 82371AB/EB/MB PIIX4 ISA (rev 08)
00:07.1 IDE interface: Intel Corporation 82371AB/EB/MB PIIX4 IDE (rev 01)
00:07.3 Bridge: Intel Corporation 82371AB/EB/MB PIIX4 ACPI (rev 08)
00:07.7 System peripheral: VMware Virtual Machine Communication Interface (rev 10)
00:0f.0 VGA compatible controller: VMware SVGA II Adapter
02:00.0 Serial Attached SCSI controller: VMware PVSCSI SCSI Controller (rev 02)
02:01.0 USB controller: VMware USB1.1 UHCI Controller
02:02.0 Ethernet controller: VMware VMXNET3 Ethernet Controller (rev 01)
02:03.0 USB controller: VMware USB2 EHCI Controller
02:04.0 SATA controller: VMware SATA AHCI controller
03:00.0 Co-processor: Intel Corporation Device 0da5 (rev 02)
```

Figure 16. Verifying the virtual device is active

References

For more information, see these resources:

- Configure a Virtual Machine to Use SR-IOV
  https://docs.vmware.com/en/VMware-vSphere/8.0/vsphere-networking/GUID-CC021803-30EA-444D-BCBE-618E0D836B9F.html
- Intel QAT downloads
  https://www.intel.com/content/www/us/en/developer/topic-technology/open/quick-assist-technology/overview.html

Author

Skyler Zhang is a VMware Engineer at the Lenovo Infrastructure Soiutions Group, based in Beijing, China. She joined the OS team in 2022.

Thanks to the following specialists for their contributions and suggestions:

- Gary Cudak, Lenovo OS Architect
- Alpus Chen, Lenovo VMware Engineer
- Chengcheng Peng, Lenovo VMware Engineer
- David Hsia, Lenovo VMware Engineer
- David Watts, Lenovo Press

Notices

Lenovo may not offer the products, services, or features discussed in this document in all countries. Consult your local Lenovo representative for information on the products and services currently available in your area. Any reference to a Lenovo product, program, or service is not intended to state or imply that only that Lenovo product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any Lenovo intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any other product, program, or service. Lenovo may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

Lenovo (United States), Inc.
8001 Development Drive
Morrisville, NC 27560
U.S.A.
Attention: Lenovo Director of Licensing

LENOVO PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. Lenovo may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

The products described in this document are not intended for use in implantation or other life support applications where malfunction may result in injury or death to persons. The information contained in this document does not affect or change Lenovo product specifications or warranties. Nothing in this document shall operate as an express or implied license or indemnity under the intellectual property rights of Lenovo or third parties. All information contained in this document was obtained in specific environments and is presented as an illustration. The result obtained in other operating environments may vary. Lenovo may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

Any references in this publication to non-Lenovo Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this Lenovo product, and use of those Web sites is at your own risk. Any performance data contained herein was determined in a controlled environment. Therefore, the result obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

This document, LP2166, was created or updated on February 28, 2025.

Send us your comments in one of the following ways:

- Use the online Contact us review form found at:
  https://lenovopress.lenovo.com/LP2166
- Send your comments in an e-mail to:
  comments@lenovopress.com

This document is available online at https://lenovopress.lenovo.com/LP2166.

Trademarks

Lenovo and the Lenovo logo are trademarks or registered trademarks of Lenovo in the United States, other countries, or both. A current list of Lenovo trademarks is available on the Web at https://www.lenovo.com/us/en/legal/copytrade/.

The following terms are trademarks of Lenovo in the United States, other countries, or both:
Lenovo®
ThinkSystem®

The following terms are trademarks of other companies:

AMD, AMD Virtualization™, and AMD-V™ are trademarks of Advanced Micro Devices, Inc.

Intel® and Xeon® are trademarks of Intel Corporation or its subsidiaries.

Other company, product, or service names may be trademarks or service marks of others.