Intel® Virtual RAID on CPU (Intel® VROC) for X11 Serverboards

Linux User's Guide

Revision 1.0e

The information in this USER'S GUIDE has been carefully reviewed and is believed to be accurate. The vendor assumes no responsibility for any inaccuracies that may be contained in this document, makes no commitment to update or to keep current the information in this manual, or to notify any person organization of the updates. Please Note: For the most up-to-date version of this manual, please see our web site at www.supermicro.com.

Super Micro Computer, Inc. ("Supermicro") reserves the right to make changes to the product described in this manual at any time and without notice. This product, including software, if any, and documentation may not, in whole or in part, be copied, photocopied, reproduced, translated or reduced to any medium or machine without prior written consent.

DISCLAIMER OF WARRANTY ON SOFTWARE AND MATERIALS. You expressly acknowledge and agree that use of the Software and Materials is at your sole risk. FURTHERMORE, SUPER MICRO COMPUTER INC. DOES NOT WARRANT OR MAKE ANY REPRESENTATIONS REGARDING THE USE OR THE RESULTS OF THE USE OF THE SOFTWARE OR MATERIALS IN TERMS OF THEIR CORRECTNESS, ACCURACY, RELIABILITY, OR OTHERWISE. NO ORAL OR WRITTEN INFORMATION OR ADVICE GIVEN BY SUPER MICRO COMPUTER INC. OR SUPER MICRO COMPUTER INC. AUTHORIZED REPRESENTATIVE SHALL CREATE A WARRANTY OR IN ANY WAY INCREASE THE SCOPE OF THIS WARRANTY. SHOULD THE SOFTWARE AND/OR MATERIALS PROVE DEFECTIVE, YOU (AND NOT SUPER MICRO COMPUTER INC. OR A SUPER MICRO COMPUTER INC. AUTHORIZED REPRESENTATIVE) ASSUME THE ENTIRE COST OF ALL NECESSARY SERVICE, REPAIR, OR CORRECTION.

LIMITATION OF LIABILITY. UNDER NO CIRCUMSTANCES INCLUDING NEGLIGENCE, SHALL SUPER MICRO COMPUTER INC. BE LIABLE FOR ANY INCIDENTAL, SPECIAL, OR CONSEQUENTIAL DAMAGES THAT RESULT FROM THE USE OR INABILITY TO USE THE SOFTWARE OR MATERIALS, EVEN IF SUPER MICRO COMPUTER INC. OR A SUPER MICRO COMPUTER INC. AUTHORIZED REPRESENTATIVE HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Any disputes arising between manufacturer and customer shall be governed by the laws of Santa Clara County in the State of California, USA. The State of California, County of Santa Clara shall be the exclusive venue for the resolution of any such disputes. Super Micro's total liability for all claims will not exceed the price paid for the hardware product.

Manual Revision 1.0e
Release Date: 12/15/2020

Document Revision History

Date	Revision	Description
12/19/2019	1.0	Initial document.
1/5/2019	1.0a	In section 3, replaced the term “strip size” with “stripe size”.
1/22/2019	1.0b	Reworded definition for “strip”.
5/16/2019	1.0c	Updated specifications in the following tables: Supported SSDS, Supported OS, Supported Hardware Configurations, and Hardware Components. Updated the supported OS hyperlink. Added the following sections: Additional Settings for the Use of M.2 NVMe SSD and Hybrid RAID. Added Figure 2-1. Combined three sections in chapter 4.
6/27/2019	1.0d	Added sections 2.2 and 6.1.1. Updated the second table in section 4.4. Updated section 4.4.
12/15/2020	1.0e	Added VROC Integrated Caching configurations.

1 Preface

About this Manual

This manual is written for Supermicro field application engineers, Supermicro systems engineers, and IT specialists wanting to learn how to properly configure Intel® Virtual RAID on CPU (VROC) and use and manage NVMe drives supported in Supermicro servers.

To use this guide, you need familiarity with Intel® X86 Architecture, the system UEFI BIOS configuration, the Linux operating system, and basic NVMe commands.

Notes

For your system to work properly, please follow the links below to download the user's manuals:

- Supermicro product manuals: http://www.supermicro.com/support/manuals/
- Product safety info: http://www.supermicro.com/about/policies/safety_information.cfm

If you have any questions, please contact our support team at: support@supermicro.com.

This manual may be periodically updated without notice. Please check with the Supermicro website for possible updates to the manual revision level.

Contents

2 Introduction

Intel® Virtual RAID on CPU (VROC) enables Redundant Array of Independent Disks (RAID) volume management with the Intel® Xeon® Scalable Processor family. Intel® VROC leverages this architecture to enable non-volatile memory express (NVMe) RAID, connect via a peripheral component interconnect express (PCIe) connection, and directly manage on the CPU. Intel® VROC provides compelling RAID performance that unleashes the full potential of NVMe drives, while eliminating the need for a discrete hardware RAID controller card. This reliable hybrid RAID solution is easy to scale with flexible drive configurations. Intel® VROC is compatible with RAID 0, RAID 1, RAID 5, and RAID 10.

This guide describes the procedures for configuring Intel® VROC and the use of NVMe drives on Supermicro X11 Serverboards.

View the latest information and product updates at https://www.supermicro.com/products/accessories/addon/AOC-VROCxxxMOD.cfm.

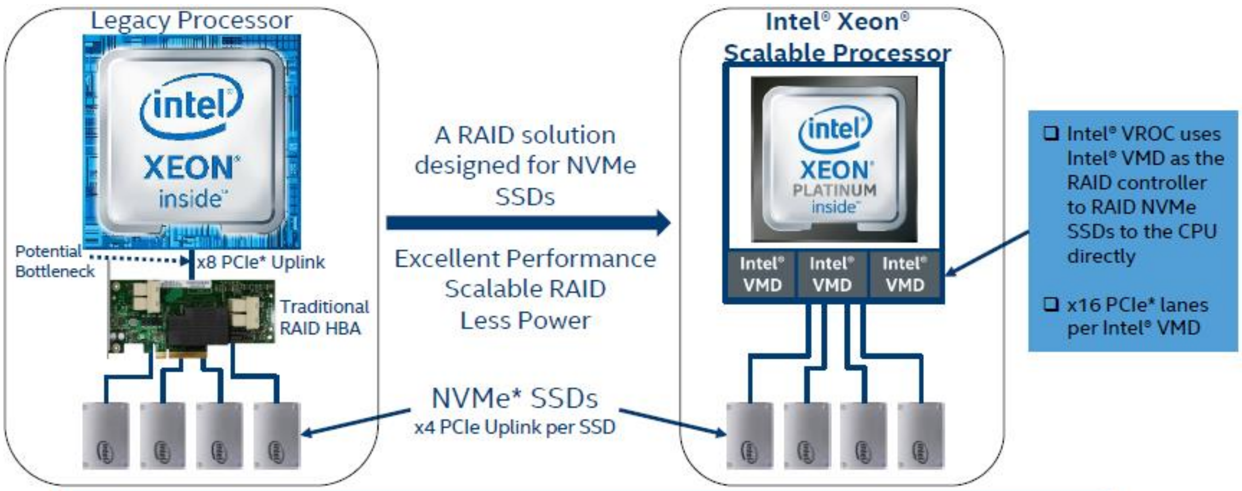


Figure 2-1. Intel® VROC

2.1 Terms

The table below displays the terms used throughout this guide.

Term	Description
API	Application Programming Interface
BIOS	Basic Input/Output System
CPU	Central Processing Unit
Data Striping	The method of writing data in sequential order so that the data is stored evenly across all drives within the RAID array.
GB	Gigabyte
GUI	Graphical User Interface
HBA	Host Bus Adapter
HDD	Hard Disk Drive
Intel® VMD	Intel® Volume Management Device
Intel® VROC	Intel® Virtual RAID on CPU

Term	Description
JD	Journaling Drive
KB	Kilobyte
Matrix	A Matrix occurs when there is more than one array within a container.
MB	Megabyte
NVMe	Non-Volatile Memory Express
OS	Operating System
Parity	Parity in a RAID detects and rebuilds missing information in a failed drive.
PCIe	Peripheral Component Interconnect Express
POST	Power-on Self-Test
PPL	Partial Parity Logging
Pre-OS	An option during the UEFI BIOS stage to configure Intel® VROC and Intel® RSTe RAID
RAID	Redundant Array of Independent Disks
RAID 0 (Disk Striping)	The data in the RAID volume is striped across all drives in the RAID array, which improves read/write performance. However, no data redundancy is created. If one drive fails, all data is lost and irrecoverable.
RAID 1 (Disk Mirroring)	The data in the RAID volume is duplicated (mirrored) across all drives in the RAID array, which creates data redundancy and fault tolerance. If one drive fails, data is still usable.
RAID 5 (Disk Striping with Parity)	The data in the RAID volume and parity are striped across all drives in the RAID array. If one drive fails, data can be recovered from the remaining drives.
RAID 10 (Disk Striping and Mirroring or 1+0)	Consists of two sets of drives; the data in the RAID volume is striped across one set of drives while the other set of drives are mirrored for fault tolerance.
RAID Recovery	A process in which data is recovered from a RAID array that has failed.
RAID Size	The total amount of drive space allocated for the drives in the RAID array.
RAID Volume	The total amount of usable space amongst the drives in the RAID array.
RSTe	Rapid Storage Technology Enterprise
RWH	RAID Write Hole
SAS	Serial Attached SCSI
SCSI	Small Computer System Interface
Spare Drive	A backup drive that is used in a RAID recovery.
SSD	Solid-State Drive
Strip Size	A block size of data forming a stripe that is evenly distributed across the physical drives within a RAID array.
Stripe Size	The size of interleaved data segments that are written across multiple drives, excluding parity drives.
TB	Terabyte
UEFI	Unified Extensible Firmware Interface. A system setting in the BIOS.
VMD	Volume Management Device

2.2 Technical References

Below is a list of technical references available to view:

- FAQ
- Product Brief
- RAID Performance Analysis
- Software and Product Updates
- Support Information
- Supported Configurations

3 Requirements and Restrictions

- **Intel® VROC is only available when the system is configured for UEFI boot mode.**
- Intel® VROC is not compatible with secure boot. This feature must be disabled.
- When creating a bootable OS RAID, physical drives in the array must reside on the same CPU and VMD on that CPU.
- Spanning drives when creating RAID devices is not recommended due to performance issues, even though it is supported.
- The rpm patch in "RSTe_VROC_RHEL7.4_5.4_2018.3.30" package must be installed. This enables the mdadm command to see Intel® VROC RAID. It is available here: https://downloadcenter.intel.com/download/28158/Intel-Virtual-RAID-on-CPU-Intel-VROC-and-Intel-RapidStorage-Technology-enterprise-Intel-RSTe-Driver-for-Linux.
- To enable Intel® VROC, insert an Intel® VROC hardware key on the motherboard, and enable the appropriate processor's Virtual Management Devices in the UEFI BIOS setup.

Note: A hardware key is a license that enables RAID functionalities.

4 Supported Features

4.1 Hybrid RAID

Intel® VROC is a hybrid RAID product. The table below displays the differences between hardware RAID, software RAID, and Intel® VROC.

RAID Features	**Hardware RAID**	**Intel® VROC**	**Software RAID**	**Description**
Error Handling Isolation	✓	✓	X	Intel® VMD isolates SSD error/event handling from OS to reduce system crashes or reboots due to errors
Reliable Data Storage	✓	✓	X	Protects RAID 5 data even if power loss occurs while volume is degraded
Boot Support	✓	✓	X	Provides protection to system volume
Management	✓	✓	X	UEFI, GUI, CLI, RESTful, remote web
Dedicated I/O Processor for RAID	✓	1	X	1 = Uses powerful Intel® Xeon CPU to RAID the unprecedented fast NVMe SSDs
Protected Write Back Cache	✓	2	X	2 = Uses the power loss to protect write back cache inside SSDs
Easily Upgraded	X	✓	✓	Software update vs. new hardware purchase
Less Hardware Required	X	✓	✓	HBA, cable, and battery backup unit is not needed. Saves power and precious PCIe lanes for others.

4.2 Solid-State Drives

The table below lists all Intel® VROC supported NVMe SSD products. To view the latest supported SSDs, visit https://www.intel.com/content/www/us/en/support/articles/000030310/memory-and-storage/ssd-software.html.

Note: Intel® RSTe has been replaced by Intel® VROC.

Brand	**Model**
Intel®	All Intel® SSDs for Data Center with NVMe, such as: • Intel® SSD DC P3100 Series • Intel® SSD DC P3500 Series • Intel® SSD DC P3520 Series • Intel® SSD DC P3600 Series • Intel® SSD DC P3608 Series

Brand	Model
	• Intel® SSD DC P3700 Series • Intel® SSD DC P4101 Series • Intel® SSD DC P4500 Series • Intel® SSD DC P4501 Series • Intel® SSD DC P4510 Series • Intel® SSD DC P4511 Series • Intel® SSD DC P4600 Series • Intel® SSD DC P4610 Series • Intel® Optane™ SSD DC P4800X Series • Intel® Optane™ SSD DC P4801X Series All Intel® Professional NVMe SSDs: • Intel® SSD Pro 7600p Series • Intel® SSD Pro 6000p Series Other selected Intel® SSD Series: • Intel® Optane SSD 900P Series • Intel® Optane SSD 905P Series
Samsung®	• SM951 • SM961 • PM953 • PM961 • PM963 • PM983
Toshiba®	• XG3
Micron®	• 9100 Series • 9200 Series
Lenovo®	• Astani
Huawei®	• ES3600P
Western Digital®	• SN720

4.3 Operating Systems

The table below lists the supported operating systems (OS). To view the latest supported OS, visit https://www.intel.com/content/www/us/en/support/articles/000030310/memory-and-storage/ssd-software.html?wapkw=vroc+operation+systems.

RedHat Enterprise Linux (RHEL)	SUSE Linux Enterprise (SLES)
RHEL 7.5	SLES 12 SP3
RHEL 7.6	SLES 12 SP4
RHEL 7.8	SLES 15
RHEL 8.0	

4.4 Hardware Components

Intel® VROC is compatible with all processors in the Intel® Xeon Scalable Processor family.

The Intel® VROC hardware key is a license key that detects the Intel® VROC SKU and activates the Intel® VROC function accordingly. The key must be plugged into the Supermicro motherboard.

		Pass-Thru	Standard SKU	Premium SKU	Intel® SSD SKU
NVMe RAID	CPU attached NVMe	✓	✓	✓	✓
	3rd party SSD support	✓	✓	✓	-
	Hot-Plug/Surprise Removal	✓	✓	✓	✓
	LED management	✓	✓	✓	✓
	Web UI management	✓	✓	✓	✓
	Bootable RAID	-	✓	✓	✓
	RAID 0/1/10	-	✓	✓	✓
	RAID 5	-	-	✓	✓
	RAID Write Hole closed (BBU replacement)	-	N/A	✓	✓

Figure 4-1 is an example of an Intel® VROC key being plugged into a JRK1 connector. The table on the following page lists the key options.

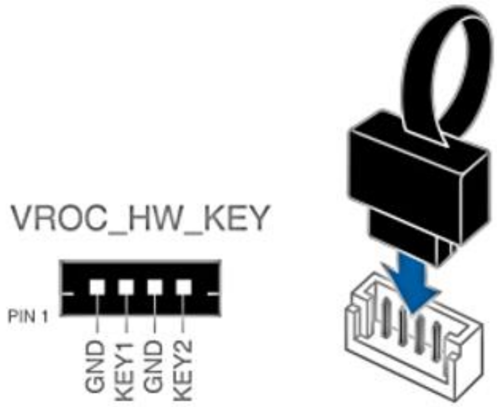


Figure 4-1. Intel® VROC RAID Key and Motherboard Connector JRK1

The table below lists the supported hardware configurations on the CPU.

SSD Totals	Platform Considerations
Up to 4 Direct Attached SSDs per Intel® VMD controller	Up to 2 levels of switches
Up to 24 SSDs per Intel® VMD controller when using switches	
Up to 48 SSDs per platform when using switches	Data volumes can span Intel® VMD controllers and CPUs and boot volumes cannot span Intel® VMD controllers

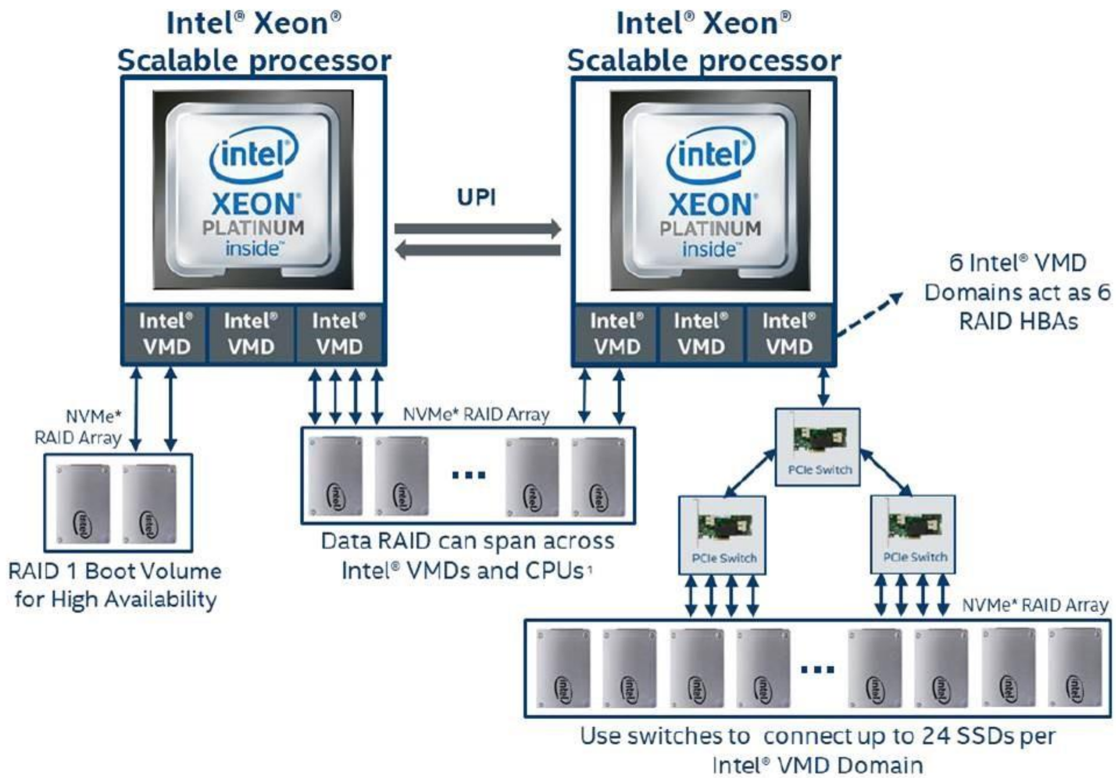


Figure 4-2 Supported Hardware Configurations

4.5 RAID Types

The table below lists the supported RAIDs.

Note: The stripe sizes are 4K, 8K, 16K, 32K, 64K, and 128K.

RAID	Number of Drives Supported	Description
0	Minimum of 2 and maximum of 24	Data is striped across all drives in the RAID array, but no data redundancy is created. RAID 0 is **not** recommended for critical systems.
1	Minimum and maximum of 2	Data is duplicated (mirrored) from one drive to another creating data redundancy and fault tolerance.
5	Minimum of 3 and maximum of 24	Data and parity are striped across all drives in the RAID array.
10	Minimum and maximum of 4	Combines the read performance of RAID 0 with the fault-tolerance of RAID 1.

4.6 RAID Write Hole

A RAID Write Hole (RWH) occurs when a drive failure and power failure happens simultaneously or close to each other during the write. The events of a drive failure and power failure are connected and can lead to irrecoverable data or silent data corruption. This is caused by the lack of atomicity of write operations across the disks in the RAID.

Note: RAID Write Holes are known to affect RAID 5 arrays; however, other array types can be affected.

5 Configuring Intel® VROC RAID Settings

The Supermicro X11 platform supports Intel® VROC. Intel® VROC is a RAID solution, which integrates with Intel® Volume Management Device (Intel® VMD), for NVMe SSDs. A single processor supports up to 12 NVMe SSDs and six RAID arrays. A dual processor supports up to 24 NVMe SSDs and 12 RAID arrays.

5.1 Enabling Intel® VMD Controllers and VMD Mode

Follow the instructions below to enable Intel® VMD Controllers and VMD mode.

Note: Images displayed in this manual are for illustration only. Your UEFI BIOS screens might look different from those shown in this manual.

1. Select the option that will allow access to the UEFI BIOS setup menu during the system boot. For Supermicro UEFI BIOS, this option is typically performed by pressing the <F2> key continuously, and then pressing the <Del> key when prompted to do so.
2. Select the **Advanced** tab.
3. Use the arrow keys and press <Enter> to select **Chipset Configuration** -> **North Bridge** -> **IIO Configuration** -> **Intel® VMD technology** -> **Intel® VMD for Volume Management Device on CPU** to access the tab items.
4. Select each VMD on the CPU and enable VMD mode for the NVMe device (Figure 5-1).

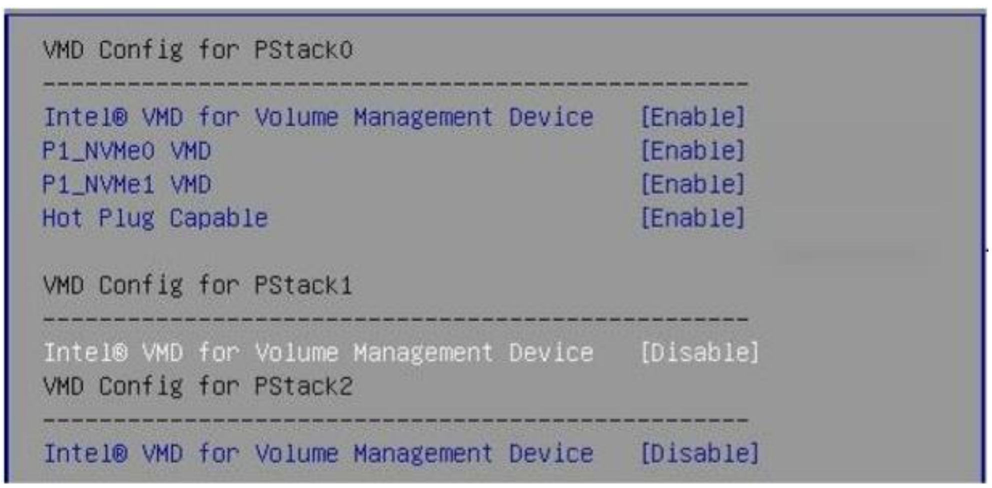


Figure 5-1. Enabled VMD mode for NVME devices

5. Select the desired PStack# to **Enable** or **Disable** the corresponding Intel® VMD controller (Figure 5-2).

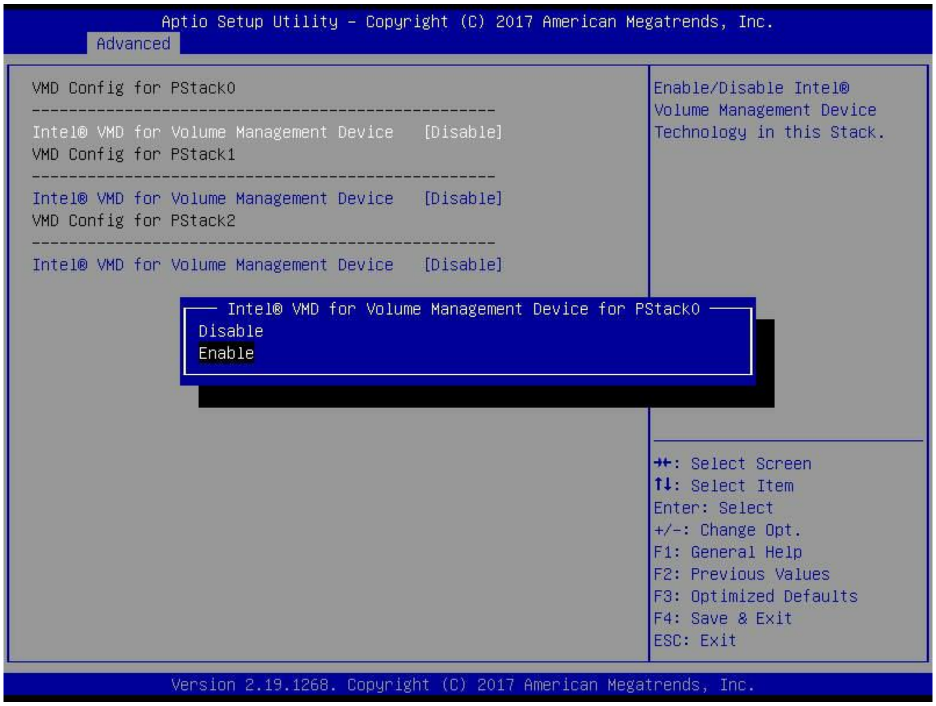


Figure 5-2. Enable Intel® VMD for Volume Management Device for Pstack0

6. Select the desired PCIe slot to **Enable** or **Disable** Intel® VMD functionality according to the current hardware configuration being used. Hot Plug Capability can also be **Enabled** or **Disabled** (Figure 5-3).

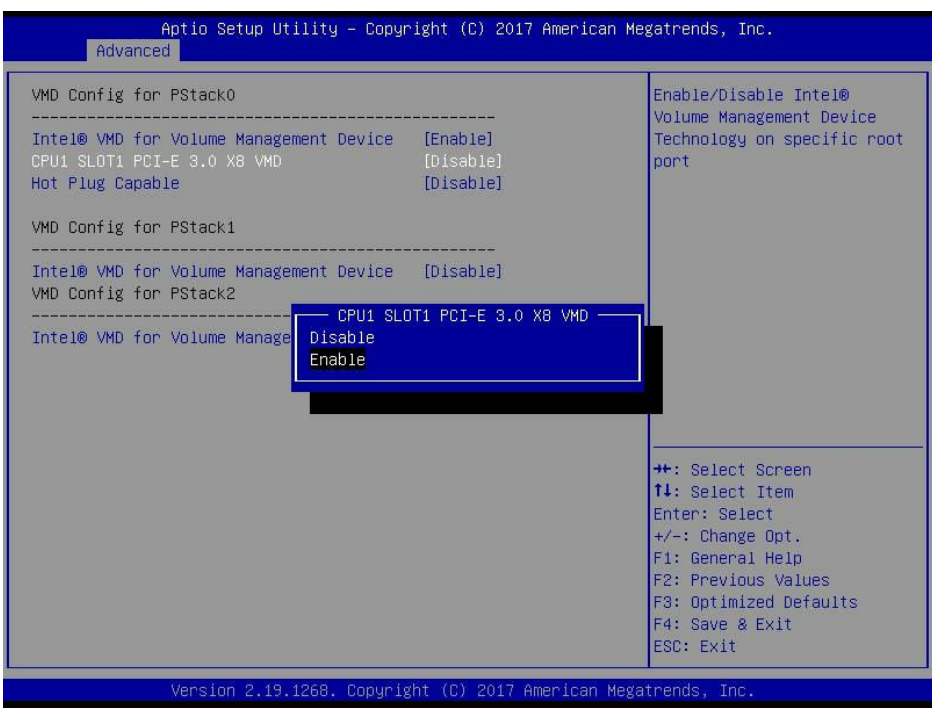


Figure 5-3. Enabling VMD Functionality per Slot

7. Repeat steps 8-9 for each PStack# on each CPUX to be enabled or disabled. In this example, we **enabled** *CPU1 Slot1* (Figure 5-4) and *CPU2 Slot5* (Figure 5-5) (our 4x U.2 form factor SSDs) as well as *CPU1 M.2 C-1* and *CPU1 M2. C-2* (our 2x M.2 form factor SSDs).

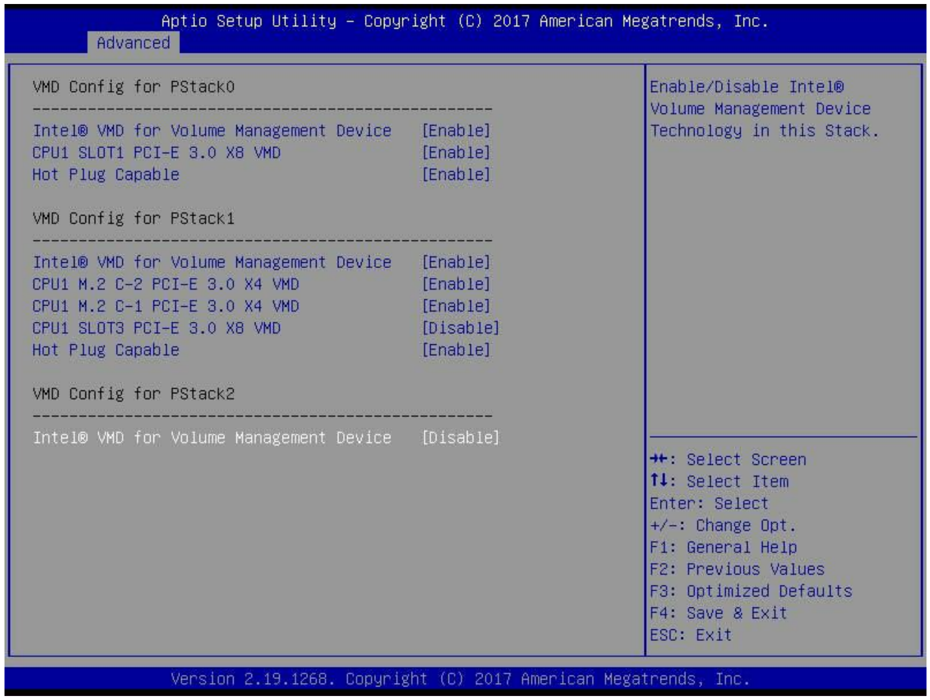


Figure 5-4. Enabled CPU1 example

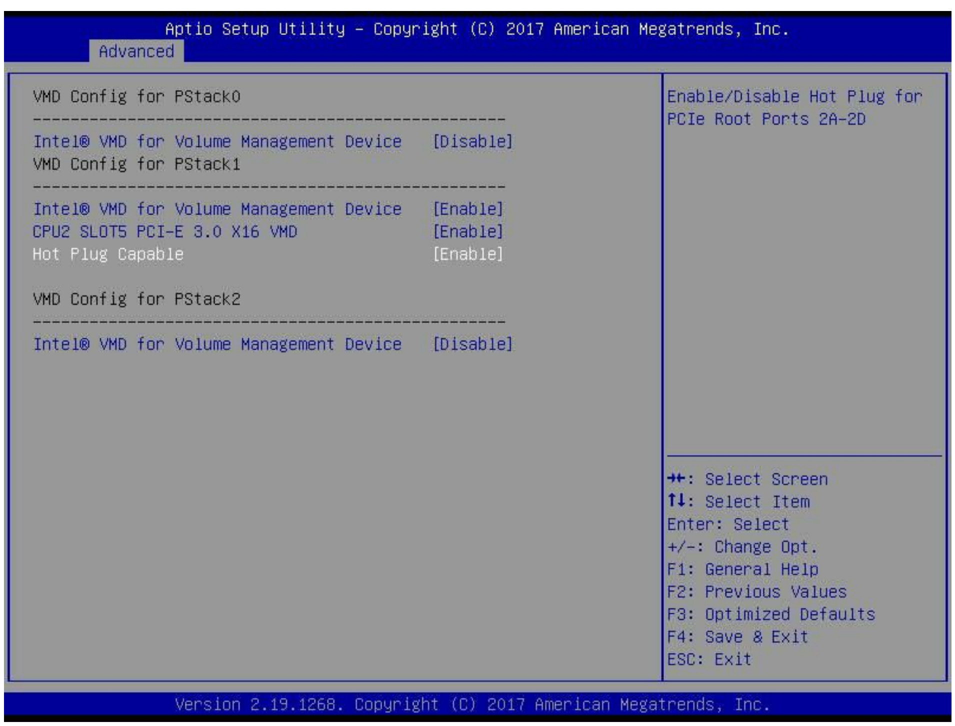


Figure 5-5. Enabled CPU2 example

8. Save the changes and enter the UEFI BIOS again (Figure 5-6).

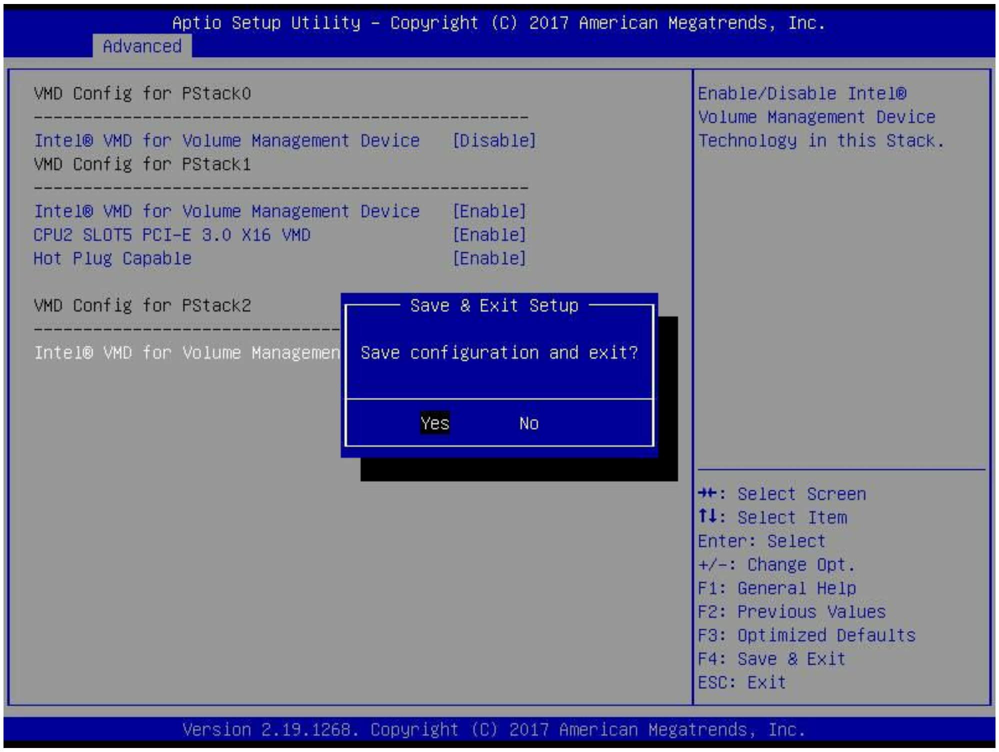


Figure 5-6. Save Configuration and Reboot

9. Enter the UEFI BIOS again and select the Advanced tab. The system is now ready to create a RAID volume.

Note: Disabling the VMD controller without first deleting the associated existing RAID volume can lead to unexpected behavior. This action is strongly **not** recommended.

Note: The effects of physically changing or swapping a CPU on the VMD controller enablement has not yet been thoroughly tested or documented.

5.2 Additional Settings for the Use of M.2 NVMe SSD

Depending on the system, motherboard, and BIOS version, the following BIOS settings may be necessary for the proper operation of M.2 NVMe drives:

- Set the CPU IOU settings to x4x4x4x4 PCIe bifurcation. This option may be found under **BIOS > Advanced > Chipset Configuration > North Bridge > IIO Configuration > CPU Configuration > IOU Setting > x4x4x4 x4.**
- Set sthe NVMe Firmware Source to AMI Native Support. This option may be found under **BIOS Setup > Advanced > PCIe/PCI/PnP Configuration > NVMe Firmware Source > AMI Native Support**.

Note: Refer to the applicable system or motherboard user manual.

5.3 Creating a RAID Volume

Follow the instructions below to create a RAID volume.

Note: Images displayed in this manual are for illustration only. Your UEFI BIOS screens might look different from those shown in this manual.

Note: Use the arrow keys to make a selection.

1. Select the option that will allow the user to access the UEFI BIOS setup menu during the system boot. For Supermicro UEFI BIOS, this option is typically performed by pressing the <F2> key continuously, and then pressing the <Del> key when prompted to do so.
2. Select the **Advanced** tab.
3. Select **Intel® Virtual RAID on CPU** and press <Enter> (Figure 5-7).

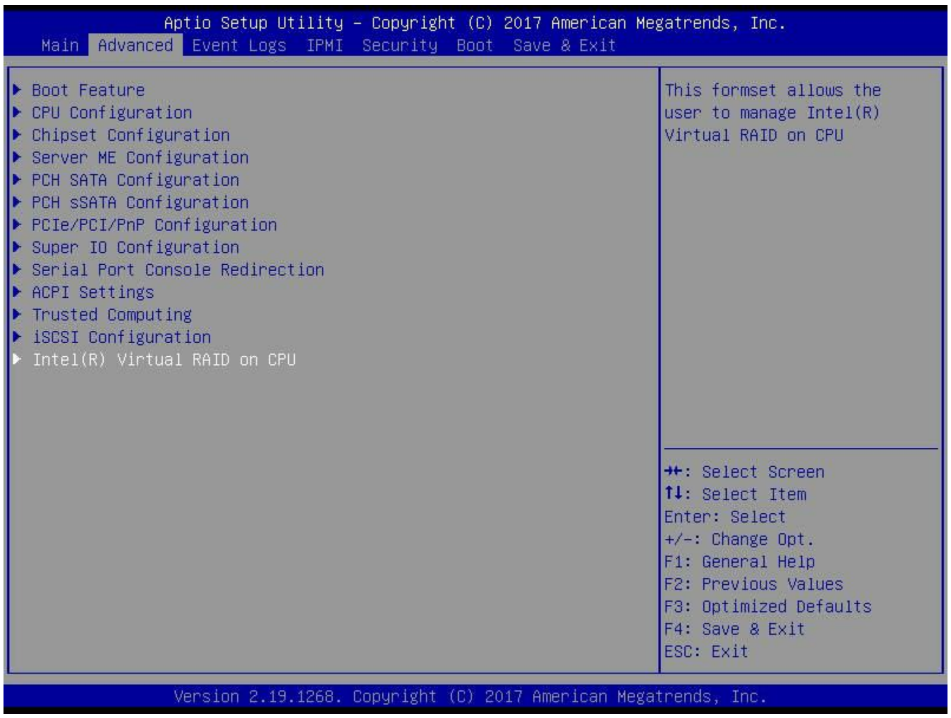


Figure 5-7. Intel® Virtual RAID on CPU (with VMD Functionality Enabled) option on Advanced tab

4. Select **All Intel® VMD Controllers** and press <Enter> (Figure 5-8).

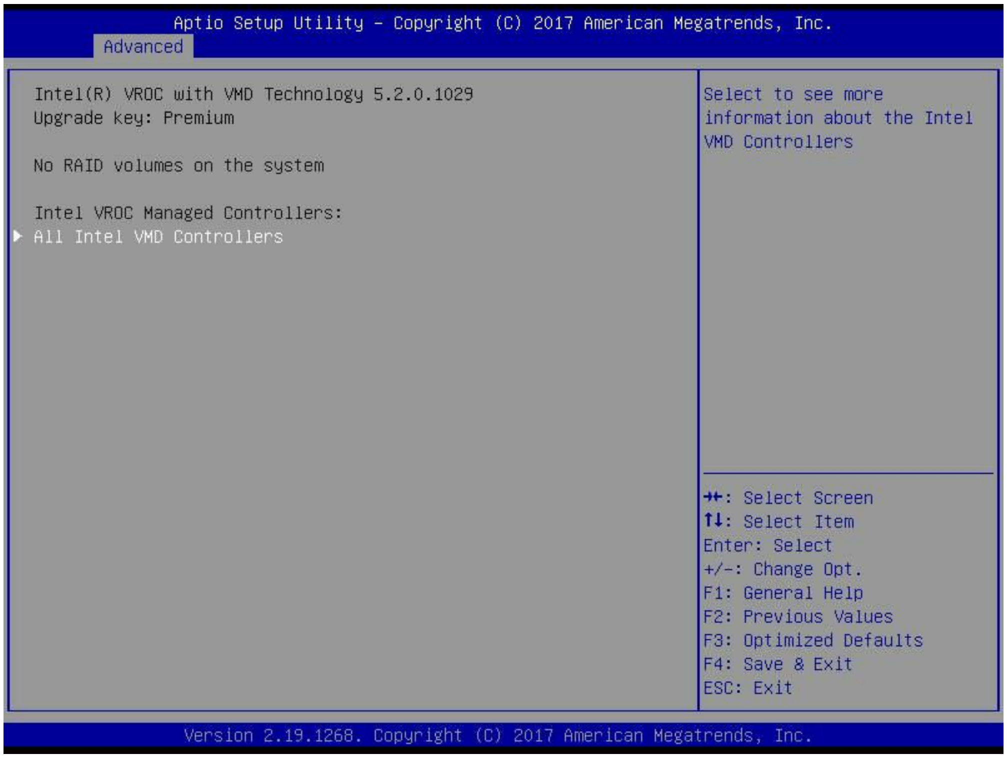


Figure 5-8. All Intel® VMD Controllers on Advanced tab

5. Select **Create RAID Volume** and press <Enter>.
6. Select **Name** and press <Enter> to rename the volume.
7. Select **RAID Level** and press <Enter> to scroll through the available RAID values and make a selection.
8. If the intended RAID volume will consist of member disks attached to different VMD controllers, select **Enable RAID spanned over VMD Controller** and press <Spacebar>[1] to toggle the selection.
9. Highlight each intended member disk one by one and press <Spacebar> to toggle the selection.
10. Select **Strip Size** and press <Enter> to scroll through the available values and make a selection.
    Note:
11. If a RAID 1 was selected, the strip size will be automatically set to 128KB and this option will not be available.

[1]Alternatively, press <Enter> and select the **X** value.

12. Select **Capacity (MB)** and use the keyboard to adjust the RAID size (Figure 5-9).

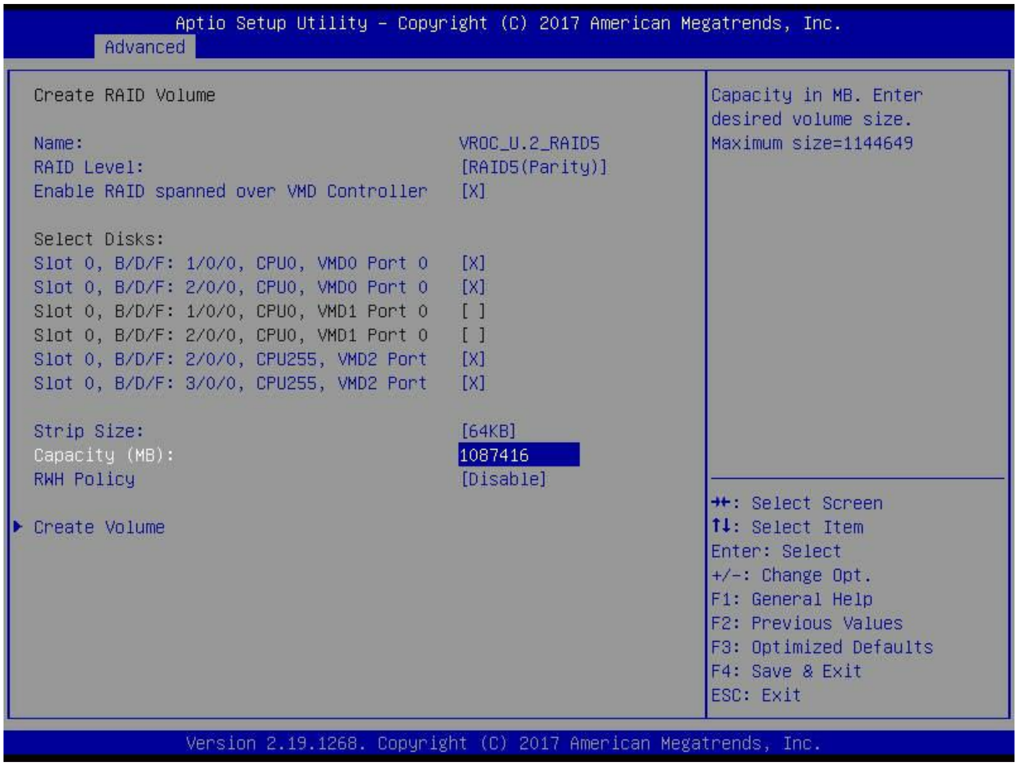


Figure 5-9. Default Capacity is 95% of Maximum Size

Note: The default size is calculated based on the selected disks and RAID level, and then taking 95% of the maximum value due to the Disk Coercion feature.[2] The rightmost sub-window displays the maximum value.

13. If a RAID 5 was selected, select **RWH Policy** and press <Enter> to scroll through the available RWH policies and make a selection.

[2] This feature will analyze the physical disks and will automatically adjust (round down) the capacity of the disk(s) to 95% of the smallest physical disk.

14. Select **Create Volume** and press <Enter>. If the RAID volume spans multiple VMD controllers (step 8), then the following warning message will appear: "You have selected NVMe drives that are connected to multiple VMD controllers. Press 'y' to create, 'n' to discard" (Figure 5-10).

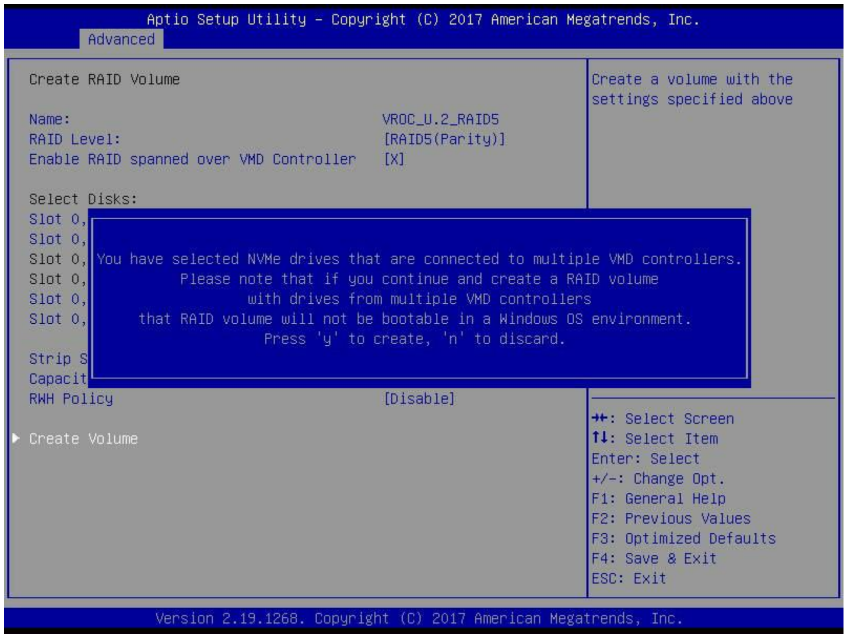


Figure 5-10. Warning Message for Spanned VMD Volume

15. The RAID volume configuration can be verified by looking at the Intel® Virtual RAID on the CPU screen. If the RAID volume was created successfully, it will be listed under the heading of Intel® VROC Managed Volumes (Figure 5-11).

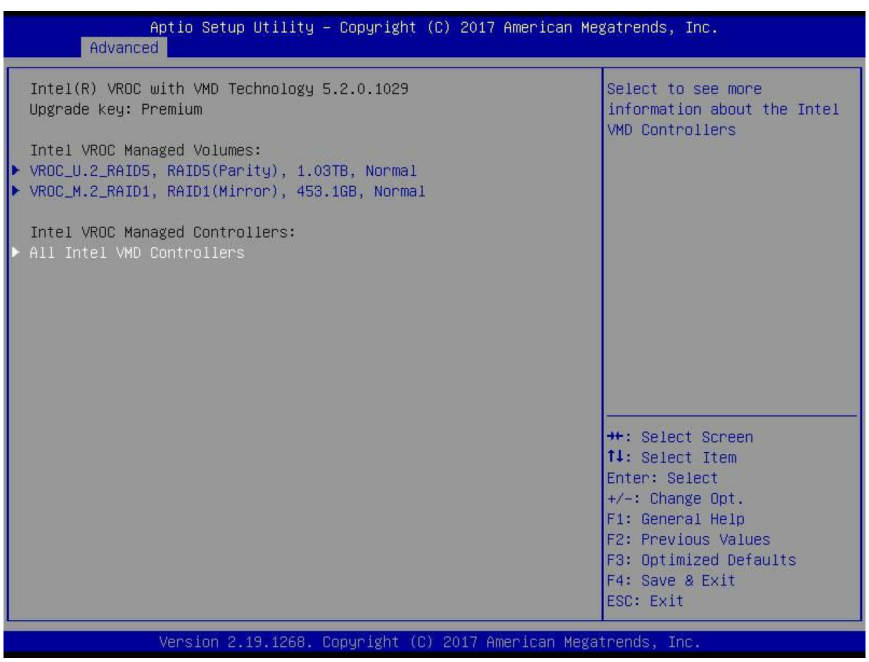


Figure 5-11. Intel® VROC Summary

5.4 RAID Volume Information

Follow the step below to view the RAID volume information.

Note: Images displayed in this manual are for illustration only. Your UEFI BIOS screens might look different from those shown in this manual.

Note: Use the arrow keys to make a selection.

1. Select and press <Enter> on the desired RAID volume on the Advanced tab. The following screen will appear (Figure 5-12).

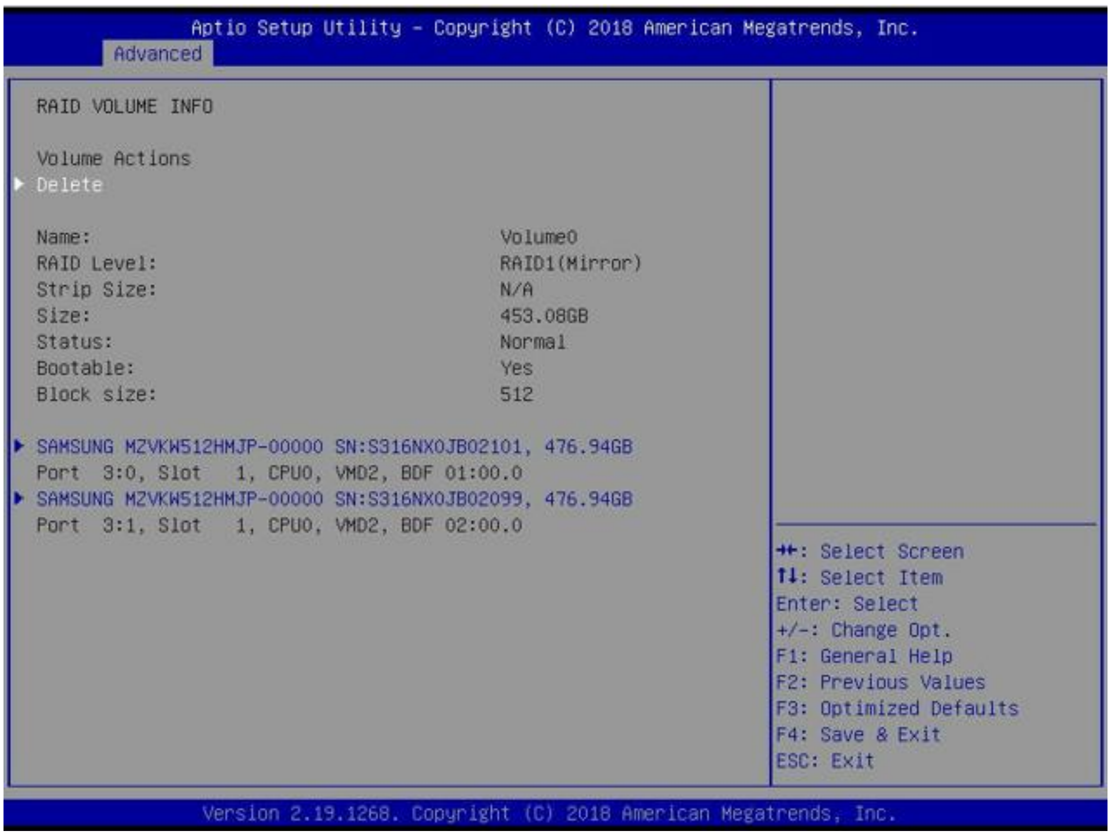


Figure 5-12. RAID Volume Info screen

5.5 Status Indications

An LED indicator on the drive carrier shows the RAID status of the drive. The table below lists the different LED statuses.

Drive Carrier Status LED Indicator	
Status	State (amber)
Normal function	Off
Locate	4 Hz blink
Fault	Solid on
Rebuild	1 Hz blink

6 Intel® VROC Setup

6.1 Modes of Operation and Upgrade Key

Intel® VROC supports three modes of operation: Pass-Through, Standard, and Premium, see table below. Enabling the Intel® VROC Standard SKU or the VROC Premium SKU requires the purchase and installation of a separate Intel® VROC RAID upgrade hardware key.

Supermicro Part Number	Item Description	Hardware Key	RAID Support	Label Color
N/A	Intel® VROC Pass Through No key (default)	No hardware key needed	No RAID support	N/A
AOC-VROCINTMOD	Intel® VROC Intel® SSD Only Upgrade module Intel® SSD only	Intel® SSD Only hardware key needed	RAID 0/1/5/10 support	Light Green
AOC-VROCSTNMOD	Intel® VROC Standard Upgrade module standard	Standard hardware key needed	RAID 0/1/10 support	Black
AOC-VROCPREMOD	Intel® VROC Premium Upgrade module premium	Premium hardware key needed	RAID 0/1/5/10 support	Red

6.1.1 Intel® VROC 90-Day Trial

A 90-day trail is included with the Intel® VROC package and does not require an Intel® VROC Upgrade key. The trial begins when the first Intel® VROC RAID volume is created. The trial will conclude if an Intel® VROC Upgrade key is inserted prior to the 90-day expiration.

Caution: Inserting the upgrade key during the trial may cause existing RAID volumes to not be seen and may be in an unknown state.

Note: Do **not** mix NVMe vendors while using the trial version as unexpected behavior may occur.

6.2 Intel® VROC RAID Volumes on a Single VMD Domain

The Intel® VMD is an integrated PCIe endpoint within the CPU root complex and is classified as a RAID controller. Each Intel® Xeon Scalable Processor CPU provides 48 PCIe lanes, which are subdivided into three domains of 16 PCIe lanes each controlled by a separate VMD. An Intel® VMD can be turned on/off at x4 lane granularity and supports either PCIe switch devices or NVMe SSD devices (PCIe add-in cards, M.2 form factor, or U.2 form factor).

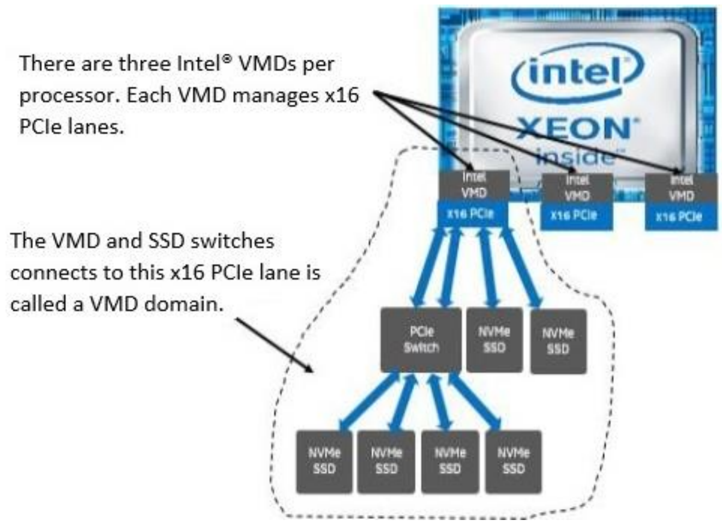


Figure 6-1. Intel® Volume Management Device Overview

Note: A single Intel® VMD supported processor supplies 48 PCIe lanes and contains three Intel® VMD controllers (domains). Refer to Figure 6-1.

6.3 Intel® VROC RAID Volumes across Multiple VMD Domains

Intel® VROC allows DATA RAID volumes on NVMe disks that span across multiple Intel® VMD domains (Figure 6-2).

Note: Although supported, spanning across CPUs is generally **not** recommended as this configuration may incur a performance penalty.

Note: Intel® VROC does not support installing an OS onto an Intel® VROC RAID volume that spans across multiple Intel® VMD domains.

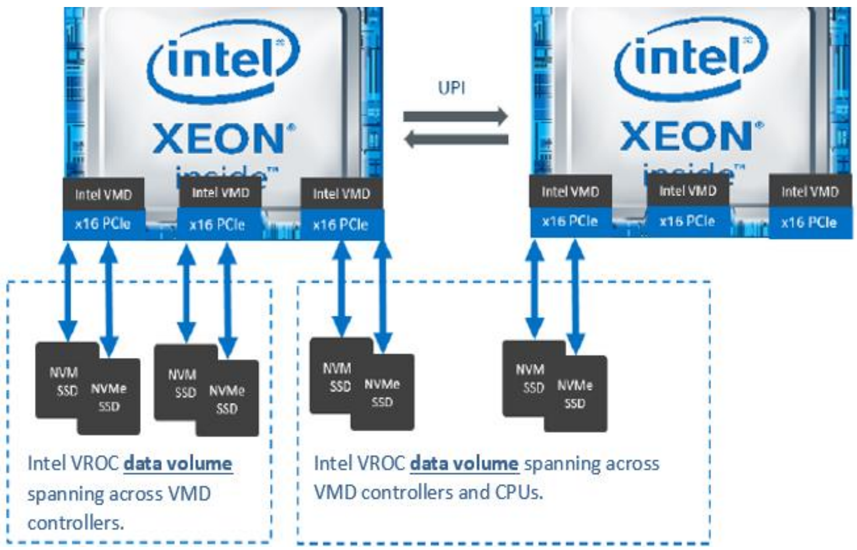


Figure 6-2. Intel® VROC Data Volume Spanning Across Multiple VMD Domains

6.4 Intel® VROC Bootable RAID Volume with HII

Verify that the following has been done before reviewing the instructions:

- Intel® VMD is enabled in UEFI BIOS.
- There is a sufficient number of drives with the appropriate size to create the RAID volume.
- The drives are properly connected to the system.
- All drives can be seen in UEFI BIOS.

Note: Images displayed in this manual are for illustration only. Your UEFI BIOS screens might look different from those shown in this manual.

Note: Use the arrow keys to make a selection.

1. Select the option that will allow access to the UEFI BIOS setup menu during the system boot. For Supermicro UEFI BIOS, this option is typically performed by pressing the <F2> key continuously, and then pressing the <Del> key when prompted to do so.
2. Select **Advanced** and press <Enter>.
3. Select **PCI Configuration** and press <Enter>.
4. Select **UEFI Option ROM Control** and press <Enter>.
5. Select **Intel® Virtual RAID on CPU** under Storage Controllers and press <Enter>.
6. Select **All Intel® VMD Controllers** and press <Enter>.
7. Select **Create RAID Volume** and press <Enter>.
8. Modify the volume name by pressing <Enter>, or keep the default name by navigating to the next item.
9. Select **RAID Level** and press <Enter>. The available RAID settings are shown. Select a RAID level and press <Enter> to set the value.
10. **Enable RAID spanned over VMD Controllers**: Highlight the empty brackets and pressing <Enter>. Set the status as **enabled**: Navigate to the X and press <Enter>.
11. Select the drives to be within the array: Highlight the empty bracket next to the desired drive, press <Enter> and highlight the X. Press <Enter> to finalize the selection. Repeat step until all required drives are added.

Note: You may alter the **Strip Size** of the array, except for RAID 1. The default Strip Size is 64k.

Note: You may alter the **Capacity** if the total capacity value is less than 95%. Highlight the current value, press <Enter>, enter the preferred value in megabytes, and press <Enter>.

12. Select **Create Volume** and press <Enter>.
13. Save the changes and reboot into the UEFI BIOS. For Supermicro UEFI BIOS, this option is typically <F4> to save, followed by <Ctrl>+<Alt>+<Del> to reboot the system.

6.5Intel® VROC Integrated Caching

Intel® VROC Integrated Caching supports configuring Intel® Optane™ SSDs as a cache level storage with RAID protection[3]. Intel® VROC IC is included as a software package in the Linux operating system.

6.5.1 Configure BIOS Settings

Set up the BIOS settings for Intel® VROC IC by enabling the VMD controllers and optionally creating RAID volumes. See Chapter 5 Configuring Intel® VROC RAID Settings. If creating RAID volumes in the BIOS, skip section 6.5.4 below.

6.5.2 Install RHEL 7.8

Download RHEL 7.8 from the Redhat website https://access.redhat.com/downloads

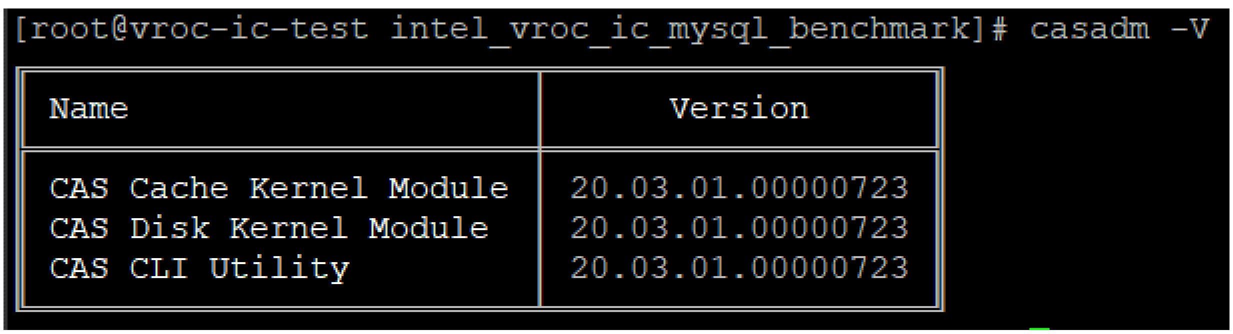
6.5.3 Install Intel® VROC Packages

Install the cache acceleration software.

```
sudo rpm -i open-cas-linux-modules_k3.10.0_1127.e17-20.03.1.0292-master.x86_64
sudo modprobe cas_cache
sudo rpm –i open-cas-linux-20.03.1.0292-master.x86_64.rpm
```

Confirm installation.

```
sudo casadm –V
```



```
[root@vroc-ic-test intel_vroc_ic_mysql_benchmark]# casadm -V
```

Name	Version
CAS Cache Kernel Module	20.03.01.00000723
CAS Disk Kernel Module	20.03.01.00000723
CAS CLI Utility	20.03.01.00000723

Figure 6-3. Cache Acceleration Software Installation

6.5.4 Configure Intel® VROC IC with Quick Start Script or Command Line

6.5.4.1 Quick Start Script

Edit the RAID and cache parameters in the quick_start.sh script.

Run the quick_start.sh.

```
chmod +x ./quick_start.sh
sudo ./quick_start.sh
```

Related Information Links

[3] https://www.intel.com/content/www/us/en/support/articles/000057962/memory-and-storage.html [intel.com]

6.5.4.2 Command Line

Alternatively, use mdadm and casadm to configure RAID and cache parameters.

1. Create Core RAID Volume:

```
mdadm -C /dev/md/imsm0 /dev/nvme[0-3]n1 –n4 –e imsm
mdadm –C /dev/md/md1 /dev/md/imsm0 –l5 –n4 –c16
```

2. Create Cache RAID Volume:

```
mdadm -C /dev/md/imsm1 /dev/nvme[5-6]n1 –n2 –e imsm
mdadm –C /dev/md/md2 /dev/md/imsm1 –l1 –n2
```

3. Configure the cache:

```
casadm -S –d /dev/md/md2 –c wo –x 4
casadm –A –i 1 –d /dev/md/md1
```

4. List all the cache instances and core devices:

```
casadm -L
```

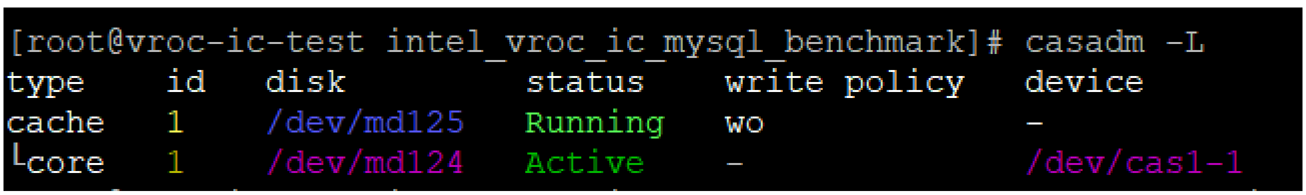


Figure 6-4. Cache Instances and Core Devices

7 Linux

7.1 mdadm Overview

mdadm is a command utility line used to create and manage RAID volumes for Linux operating systems.

7.1.1 Help Documentation

The mdadm help page

```
mdadm <command --help
```

where *<command>* can be several options, including create, assemble, build, stop, etc.

7.1.2 Basic Usage

The table below lists the available command parameters.

Command Parameter	Description
-c	Specifies the strip size in kilobytes.
-l	Specifies the RAID level (i.e., 0, 1, 5, 10).
-n	Number of devices to be used in a RAID volume.
-x	Number of spare devices in the initial RAID array.
-z	Specifies the size (in kilobytes) of space dedicated on each disk to the RAID volume. This must be a multiple of the chunk size. A suffix of ‘M’ or ‘G’ can be given to indicate megabytes or gigabytes, respectively (this also applies to the –c parameter).

```
mdadm --create<newDevice><devicesToUse> --raid-devices=<numberActiveDisks> --
level=<raidLevel>
```

Most mdadm commands have an associated shorthand term. For example, the above usage can be shortened to:

```
mdadm –c <newDevice> <devicesToUse> -n <numberActiveDisks> l <raidLevel>
```

```
mdadm --detail<mdDevice>
mdadm --stop <mdDevice>
mdadm --zero-superblock <device>
```

Note: For the remainder of this document, all examples will assume the creation of a RAID 1 device using */dev/nvmeXn1* and */dev/nvmeXn1.*

7.2 OS Installation for Linux

The instructions below should be followed after the creation of an Intel® VROC RAID volume.

Note: Steps 2-5 below may be unnecessary for RHEL 7.4 as Intel® VROC drivers may already be included.

1. Extract/unzip the contents of the rste-5.1_PV_rhel7.3.iso file to a USB device and attach it to the system.
    a. The name of the USB device **must** be "RSTE".
    b. The USB device will preferably be clean and formatted in FAT32.
2. Begin installing RHEL 7.3 and when prompted, go to the "Install Red Hat Enterprise Linux 7.3" option (Figure 7-1) and press <e> to enter edit mode (Figure 7-2).

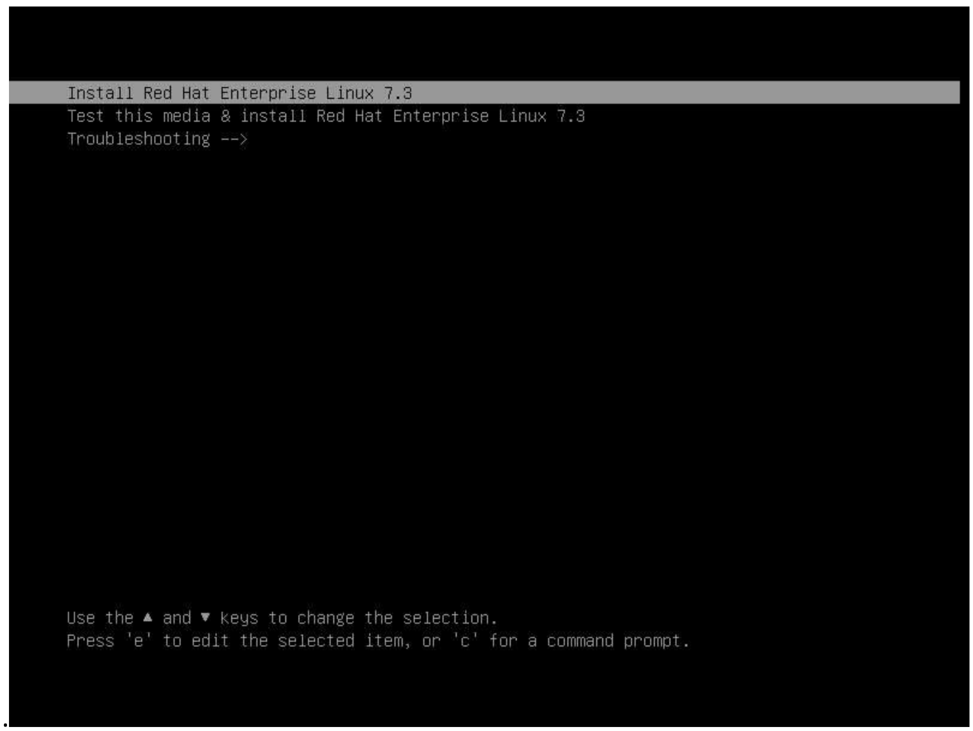


Figure 7-1. Install Red Hat Enterprise Linux 7.3 option

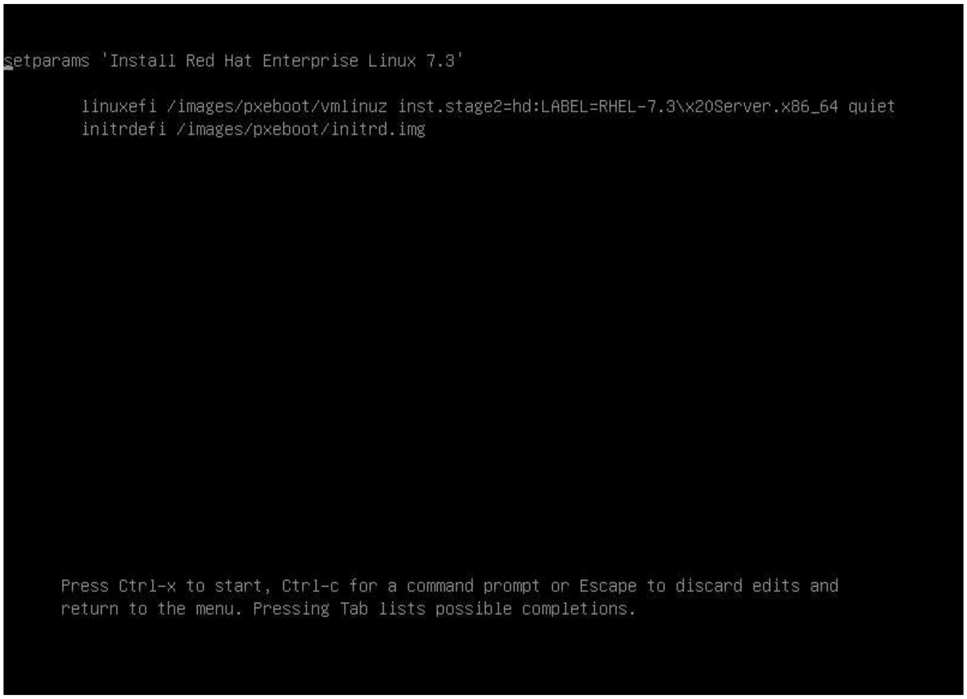


Figure 7-2. Edit Mode

3. Append the following without quotes after the word “quiet” on the third line: “inst.updates=LABEL=RSTE” (Figure 7-3).

Figure 7-3. Append “inst.updates=LABEL=RSTE” to Text

4. Press <Ctrl>+<X> to continue the installation process (Figure 7-4).

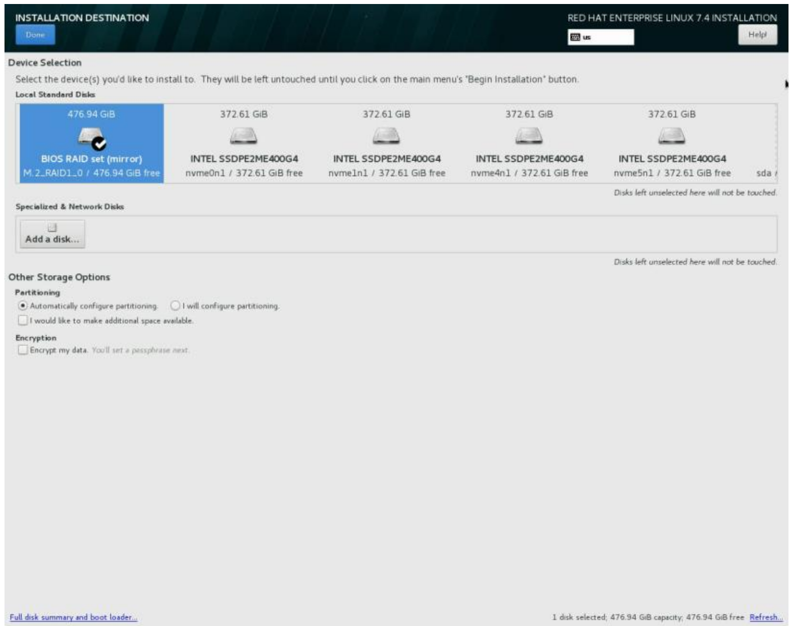


Figure 7-4. Continue the Installation Process

5. When configuring the installation destination, it is highly recommended for the **ext4** file system to be used instead of the default **xfs** file system (Figure 7-5).

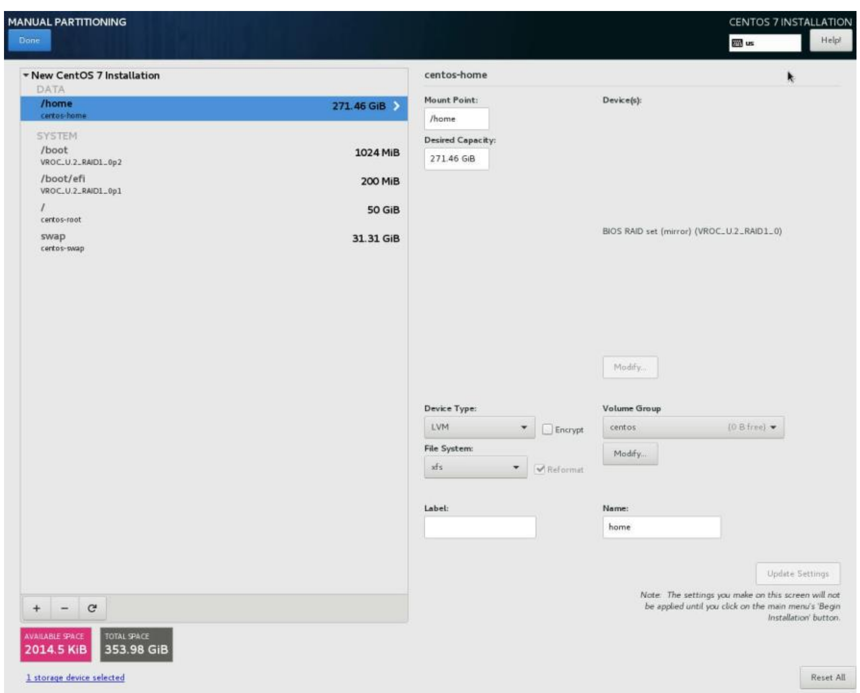


Figure 7-5. Change File Types from xfs to ext4

6. Proceed with the Linux installation as normal.

7.3 Creating a RAID Volume with mdadm

Follow the instructions below to create a RAID volume with mdadm.

Note: In some Linux installation media environments it is possible to switch from the GUI to the shell environment via <Ctrl>+ <Alt>+ <F2>. Return to the GUI via <Ctrl>+<Alt>+<F6>.

Note: The commands below each figure should be used to perform the associated step.

Warning! Back up all important data before performing the steps below. Existing data on the selected drives will be permanently deleted after creation of RAID array.

1. Erase any previous RAID metadata. Incorrect or bad metadata can cause RAID volumes to be assembled incorrectly. This operation does not attempt to wipe existing user data, but will delete the corresponding array of a current member disk.

```
[anaconda root@localhost /]#
[anaconda root@localhost /]# mdadm --zero-superblock /dev/nvme4n1
mdadm: Unrecognised md component device - /dev/nvme4n1
[anaconda root@localhost /]#
```

2. An Intel® IMSM metadata container **must** first be created with the -e flag. In this example, a 2-disk RAID container consisting of drives *nvme4n1* and *nvme5n1* is assigned the designation */dev/md0*. The wildcard expression */dev/nvme[4-5]n1* can be used to specify the range of drives, although individual expressions can be used.

```
[anaconda root@localhost /]#
[anaconda root@localhost /]# mdadm -C /dev/md0 /dev/nvme[4-5]n1 -n 2 -e imsm
mdadm: container /dev/md0 prepared .
[anaconda root@localhost /]#
```

Note: Creating an Intel® RAID requires that the *<newDevice>* name be */dev/md/<variable>*, the *metadata* flag be set to *imsm*, and the *level* flag be set to *container* (or leave the *level* flag out entirely).

3. Next, the RAID volume itself is created using the IMSM container. The following parameters may be used in conjunction to give finer control for the creation of the RAID volume.

```
[anaconda root@localhost /]#
[anaconda root@localhost /]# mdadm -C /dev/md/Volume0 /dev/md0 -n 2 -l 1
mdadm: array /dev/md/Volume0 started .
[anaconda root@localhost /]#
```

4. Create a file system on the RAID volume.

```
[anaconda root@localhost /]#
[anaconda root@localhost /]# mkfs.ext4 /dev/md/Volume0
mke2fs 1.41.9 (28-Dec-2013)
Discarding device blocks: done
Filesystem label=
OS type: Linux
Block size=4096 (log=2)
Fragment size=4096 (log=2)
Stride=0 blocks, Stripe width=0 blocks
24420352 inodes, 97677056 blocks
4883852 blocks (5.00%) reserved for the super user
First data block =0
Maximum filesystem blocks=2246049792
2981 block groups
32768 blocks per group, 32768 fragments per group
8192 inodes per group
Superblock backups stored on blocks:
      32768, 98304, 163840, 229376, 294912, 819200, 884736, 1605632, 2654208,
      4096000, 7962624, 11239424, 20480000, 23887872, 71663616, 78675968

Allocating group tables: done
Writing inode tables: done
Creating journal (32768) blocks: done
Writing superblocks and filesystem accounting information: done

[anaconda root@localhost /]#
```

5. Afterwards, the RAID volume can be mounted to the location of choice.

```
# mount/dev/md/Volume0 /mnt/<mountpoint>
```

7.4 Adding a Hot Spare

Adding a spare drive allows for immediate reconstruction of the RAID volume if a device is detected. If this occurs, mdadm will mark the failed device as “bad” and begin reconstruction with the first available spare drive. Spare drives stay in idle during normal operations. When using mdadm with imsm metadata, the spare drive added to a container is dedicated to that specific container.

The command below adds a spare drive to the designated container */dev/md0*.

```
# mdadm -a /dev/md0 /dev/<nvmeXn1>
```

7.5 RSTe and Intel® VROC

The RSTe driver and usage is currently beyond the scope of this document.

A useful command to display the status of md RAID devices is: # cat /proc/mdstat

7.6 Intel® VROC RAID Characteristics

Under Linux, individual RAID member disks are not masked by the OS and remain visible to all system operations. Thus, there is an increased risk of inadvertently destructive behavior with administrative commands such as fdisk, parted, dd, etc. This characteristic is shared with the standard Intel® RSTe RAID over SATA/SAS drives.

8 Removing RAID Volumes

8.1 Removing RAID Volumes with mdadm

Follow the instructions below to remove RAID volumes with mdadm.

Note: The commands below should be used to perform the associated step.

1. cat /proc/mdstat

```
[anaconda root@localhost /]#
[anaconda root@localhost /]# cat /proc/mdstat
Personalities : [raid0] [raid1] [raid6] [raid5] [raid4] [raid10] [linear]
md126 : active (auto-read-only) raid1 sda[1] sdb[0]
      74240000 blocks super external :md/127/0 [2/2] [UU]


md127 : inactive sdb[1](S) sda[0](S)

      6306 blocks super external:imsm

unused devices: <none>
[anaconda root@localhost /]#
```

2. # mdadm --stop *<mdDevice>*

```
[anaconda root@localhost /]#
[anaconda root@localhost /]# mdadm --stop /dev/md126 /dev/md127
mdadm: stopped /dev/md126
mdadm: stopped /dev/md127
[anaconda root@localhost /]#
```

3. # mdadm --zero-superblock *<devices>*

```
[anaconda root@localhost /]#
[anaconda root@localhost /]# mdadm --zero-superblock /dev/sd[a-b]
[anaconda root@localhost /]#
```

8.2 Removing RAID Volumes with HII

The process to remove a RAID volume is similar to creating the volume. Follow the instructions below to remove RAID volumes with HII.
Note: Use the arrow keys to make a selection.

1. Select the option that will allow access to the UEFI BIOS setup menu during the system boot. For Supermicro UEFI BIOS, this option is typically performed by pressing the <F2> key continuously, and then pressing the <Del> key when prompted to do so.
2. Select **Advanced** and press <Enter>.
3. Select **PCI Configuration** and press <Enter>.
4. Select **UEFI Option ROM Control** and press <Enter>.
5. Select **Intel® Virtual RAID on CPU** under Storage Controllers and press <Enter>.
6. Select the RAID to be removed below **Intel® VROC Managed Volumes** and press <Enter>.
7. Select **Delete** and press <Enter>.
8. Select **Yes** and press <Enter>.

8.3 Removing an Array within a Matrix

A Matrix is when there is more than one array within a container. Follow the instructions below to remove an array within a Matrix.

1. Stop the array:
```
# mdadm --stop /dev/md{number}
```

2. Remove the volume metadata:
```
# mdadm --kill-subarray={index} /dev/md{container_number}
```

Note: A ‘0’ or ‘1’ should appear in place of “{index}” representing the first or second volume.

8.4 Removing a Failed Drive with mdadm

Only container based RAID volumes can be removed. Use the following command to remove a failed drive:

```
# mdadm --remove /dev/md0 /dev/<nvmeXn1>
```

Then, power off the system before removing the physical NVMe SSD from the system.

8.4.1 Removing a Failed Drive Example

Below is an example of removing a failed drive from a RAID 5 volume with IMSM metadata. In the example, “md127” is the IMSM container and “md126” is the RAID 5 volume, and the RAID 5 container has the following drives: “/dev/nvme0n1”, “/dev/nvme1n1”, “/dev/nvme2n1”, and “/dev/nvme3n1”.

```
Personalities : [raid0] [raid1] [raid5] [raid10]
md127 : active (read-only) raid5 nvme0n1[3] nvme1n1[2] nvme2n1[1] nvme3n1[0]
                  39999488 blocks super external:/md0/0 level 5, 512k chunk,
algorithm 0 [5/5]
[UUUUU]

md0 : inactive -
            11285 blocks super external:imsm

unused devices: <none>
```

The following is displayed in /proc/mdstat when a drive fails (in this case, /dev/nvme2n1).

```
Personalities : [raid0] [raid1] [raid5]
md127 : active raid5 nvme0n1[3] nvme3n1[0]
                  39999488 blocks super external:/md0/0 level 6, 512k chunk,
algorithm 0 [5/4]

md0 : inactive - nvme2n1[1](S)
            1045 blocks super external:imsm

unused devices: <none>
```

To remove the failed drive, use the following command:

```
# mdadm --remove /dev/md0 /dev/nvme2n1
```

or use the following command to remove the failed drive from the container:

```
# mdadm /dev/md/Volume0 --fail detached --remove detached
```

After the failed drive is removed, the system **must** be powered off before replacing the NVMe SSD.

9 RAID Volume Recovery

Recovery allows for redundant RAID volumes to be rebuilt if a drive fails. RAID volumes can be recovered for RAID levels 1, 5, and 10. In most cases, drives can be recovered if only one drive fails.

For a RAID 10, recovery is possible if two of four drives fail and the two failed drives are not of the same mirrored pairs. Recovery is not possible if the two failed drives belong to the same mirror.

9.1 Rebuilding RAID Volumes

RAID volumes automatically rebuild if there are spare drives available in the container. Spare drives can also be manually added to commence the rebuild process. The RAID volume is still operational and runs in degraded mode.

9.2 Auto-Rebuild

Auto-rebuild occurs when a drive fails and the RAID volume is automatically rebuilt. Auto-build can occur in the following three scenarios:

1. If a drive fails and is replaced by a new drive.
2. If a drive fails and there is at least one spare drive in the container.
3. If there is a spare drive in one of two containers, mdadm is running in monitor mode, and the necessary policy is configured in mdadm.conf file. The spare drive will be automatically moved from one container to another when the RAID volume is degraded and requires a spare drive for rebuild.

10 Hot Swap Drives

Intel® VMD enables hot-plug and hot-unplug for NVMe SSDs, whether from Intel® or other manufacturers. Under vSphere ESXi, several steps are necessary to avoid potential stability issues[4].

10.1 Hot-unplug

1. Prevent devices from being re-detected during rescan:

```
esxcli storage core claiming autoclaim --enable=false
```

2. Unmount the VMFS volumes on the devices[5].
3. Detach the device[6].
4. Physically remove the device.

10.2 Hot-plug

Physically install the device.

ESXi will automatically discover NVMe SSDs, but a manual scan may be required in some cases.

Related Information Links

4 https://kb.vmware.com/s/article/2151404

5 https://docs.vmware.com/en/VMware-vSphere/6.5/com.vmware.vsphere.storage.doc/GUID-1B56EF97-F60E-4F21-82A7-8F2A7294604D.html

6 https://docs.vmware.com/en/VMware-vSphere/6.5/com.vmware.vsphere.storage.doc/GUID-F2E75F67-740B-4406-9F0C-A2D99A698F2A.html

Contacting Supermicro

Headquarters

Address:	Super Micro Computer, Inc. 980 Rock Ave. San Jose, CA 95131 U.S.A.
Tel:	+1 (408) 503-8000
Fax:	+1 (408) 503-8008
Email:	marketing@supermicro.com (General Information) support@supermicro.com (Technical Support)
Website:	www.supermicro.com

Europe

Address:	Super Micro Computer B.V. Het Sterrenbeeld 28, 5215 ML ‘s-Hertogenbosch, The Netherlands
Tel:	+31 (0) 73-6400390
Fax:	+31 (0) 73-6416525
Email:	sales@supermicro.nl (General Information) support@supermicro.nl (Technical Support) rma@supermicro.nl (Customer Support)
Website:	www.supermicro.nl

Asia-Pacific

Address:	Super Micro Computer, Inc. 3F, No. 150, Jian 1st Rd. Zhonghe Dist., New Taipei City 235 Taiwan (R.O.C.)
Tel:	+866-(2) 8226-3990
Fax:	+866-(2) 8226-3992
Email:	support@supermicro.com.tw
Website:	www.supermicro.com.tw