V5000 Certified Platform Installation and Service Guide

Installing, cabling, configuring and servicing VDURA V5000 Director Nodes, Storage Nodes and JBOD enclosures

Documentation for PanFS 12.0.0

© 2002–2026 VDURA Inc. All rights reserved. VDURA, the VDURA logo, VDURA Data Platform, VeLO, VPOD, DirectFlow, PanView, PanMove, and Multi-Level Erasure Coding (MLEC) are trademarks or registered trademarks of VDURA, Inc. in the U.S. and/or other countries. All other trademarks, registered trademarks, trade names, company names, and service marks are the respective properties of their holders.

Parts of this product are derived from Berkeley Database software, licensed from Sleepycat Software.

Parts of this product are derived from Berkeley BSD systems, licensed from the University of California. The BDB software is also BSD licensed, and carries the following copyrights:

- Copyright © 1990–2009 Oracle. All rights reserved.
- Copyright © 1990, 1993, 1994, 1995 The Regents of the University of California. All rights reserved.
- Copyright © 1995, 1996 The President and Fellows of Harvard University. All rights reserved.
- Copyright © 2000–2005 INRIA, France Telecom

Parts of this product are derived from gd Version 1.8.4, licensed by Boutell.Com, Inc.

Copyright © 1982, 1986, 1991, 1993 The Regents of the University of California. All rights reserved.

© UNIX System Laboratories, Inc.

All or some portions of this file are derived from material licensed to the University of California by American Telephone and Telegraph Co. or Unix System Laboratories, Inc. and are reproduced herein with the permission of UNIX System Laboratories, Inc.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
3. Neither the name of the University nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE REGENTS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE REGENTS OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Disclaimer

VDURA has carefully checked the information in this document and believes it to be accurate. However, VDURA assumes no responsibility for any inaccuracies that this document may contain. In no event will VDURA be liable for direct, indirect, special, exemplary, incidental, or consequential damages resulting from any defect or omission in this document, even if advised of the possibility of such damages.

Any alterations or inconsistent replications of this document cannot be considered accurate and are void. VDURA may not be able to support issues related to altered documents that are not approved by subject matter experts.

Document version 12.0.0-2026-11-12-13:23:43

Introduction 1
About This Guide 1
Document Conventions 2
Glossary 2
Additional Resources 4
Contact VDURA 5
Safety Information 5
Before You Begin 7
Order of Installation 8
Power Procedures 8

1. Storage Installation 10
1.1 V5000 JBOD Installation Requirements 10
1.2 V5000 JBOD Overview 11
1.2.1 Control Panel and Drive LEDs 11
1.2.2 Rear Components 12

1.3 Install the V5000 JBOD Enclosure 13
1.4 V5000 Storage Node Installation Requirements 21
1.5 V5000 Storage Node Overview 21
1.5.1 Control Panel 22
1.5.2 Bezel Power 22
1.5.3 Drive Bay Status LEDs 23
1.5.4 Rear Ports 23
1.5.5 Node Connectivity Configurations 24
1.5.6 NIC Status LEDs 25
1.5.7 Power Supply Status LED 29

1.6 Install the V5000 Storage Node 29
1.7 Cable the Storage Node and JBOD 31
1.8 Configure Storage Node IPMI Support 32

2. Director Node Installation 33
2.1 Installation Requirements 33
2.2 V5000 Director Node Overview 33
2.2.1 Control Panel 34
2.2.2 Bezel Power 35
2.2.3 Drive Bay Status LEDs 35
2.2.4 Rear Ports 35
2.2.5 Node Connectivity Configurations 36

2.2.6 NIC Status LEDs ... 38
2.2.7 Power Supply Status LED ... 42

2.3 Install the V5000 Director Node ... 43
2.4 Connect to the Console ... 44
2.5 Configure Director Node IPMI Support ... 45

3. Software Configuration ... 46
3.1 Initial Network Setup ... 46
3.2 Out of Box Experience ... 48
3.3 Configure PanFS ... 51
3.3.1 PanFS Components ... 51
3.3.2 Configure PanActive Manager ... 53
3.3.3 Upgrade and Downgrade PanFS ... 55

3.4 DirectFlow Client Installation ... 56

4. Configure InfiniBand with Subnet Manager ... 60
4.1 Subnet Manager Requirements ... 60
4.2 OpenSM ... 60

5. IPMI Configuration ... 62
6. Network Overview and Configuration ... 66
6.1 V5000 Network Requirements ... 66
6.1.1 Introduction ... 66
6.1.2 V5000 High-Speed Data Network Connection ... 66
6.1.3 Other V5000 Network Connections ... 67
6.1.4 NTP Server ... 68
6.1.5 InfiniBand Data Network ... 68
6.1.6 Ethernet Data Network ... 69
6.1.7 Data Network Port Connections ... 70

6.2 Logical Networks ... 71
6.3 Networking Modes ... 73
6.4 Configure the 25 GbE Ports ... 74
6.4.1 Configure the 25 GbE Ports Using PanCLI ... 74
6.4.2 Configure the 25 GbE Ports Using PanActive Manager ... 77

7. Product Maintenance Procedures ... 83
7.1 V5000 Director Node Component Replacement ... 83
7.1.1 Replace a Director Node ... 83
7.1.2 Replace an Individual Drive in a Node ... 94

7.1.3 Replace a Power Supply Unit 95

7.2 V5000 Storage Node Component Replacement 96
7.2.1 Replace a Storage Node 96
7.2.2 Replace an Individual Drive in a Node 97
7.2.3 Replace a Power Supply Unit 100

7.3 V5000 JBOD Component Replacement 100
7.3.1 Replace a SAS Cable 101
7.3.2 Replace a Power Supply Unit 101
7.3.3 Replace a Fan 101
7.3.4 Replace a Single Disk 102
7.3.5 Remove a Drive from Its Carrier 105
7.3.6 Replace an IOM 106
7.3.7 Replace the JBOD Enclosure 106

8. System Specifications 107
8.1 V5000 Director Node Server Specifications 107
8.2 V5000 Storage Node Server Specifications 108
8.3 V5000 4U78 JBOD Specifications 109
8.4 V5000 4U108 JBOD Specifications 111

9. Troubleshooting 113
9.1 Collect Logs 113
9.1.1 sysstat all 113
9.1.2 collect_logs 114

9.2 Send Logs 120
9.3 Log Analysis 121
9.3.1 /var/log/net_mon.log 121
9.3.2 /var/log/all.log 127
9.3.3 /var/log/sysstat.log 128

9.4 Diagnostic Utilities 128
9.5 Configure the Maximum Transmission Unit 138

Appendix. Installation and Configuration Checklist 140

Introduction

Welcome to the VDURA® V5000™ Certified Platform Installation and Service Guide. Read this introduction for an overview of the information provided in this guide, an understanding of the documentation conventions used throughout, and overall system safety precautions.

This chapter contains the following sections:

- About This Guide
- Document Conventions
- Glossary
- Additional Resources
- Contact VDURA
- Safety Information
- Before You Begin
- Order of Installation
- Power Procedures

About This Guide

In this guide you will learn how to install and configure VDURA V5000 Certified Platform systems. For information and instructions on legacy systems, refer to older versions of documentation available at support.VDURA.com.

This guide is written for system and network administrators who are familiar with Network Attached Storage (NAS) systems and Storage Area Networks (SANs).

A working knowledge of network protocols, network administration and management, and the Linux®, UNIX®, and/or Windows® operating systems is also required.

Document Conventions

This manual uses the following standard set of typographical conventions to introduce new terms, illustrate screen displays, and describe command syntax.

Convention	Meaning
boldface	Identifiers (names of classes, objects, constants, events, functions, program variables, forms, labels, and reports), environment variables, filenames, table names, column names, icons, menu items, command names, and other similar terms appear in boldface.
`courier`	Information that the product displays and information that you are instructed to enter appear in Courier typeface.

Glossary

The following terms are found throughout this guide.

Term	Explanation
BMC	Baseboard Management Controller
Enclosure	Logical structure to aggregate hypervisor, JBOD, and VPODs
GWIP	Gateway IP; an alias IP address assigned to a Director Node for S3 gateway access
HBA	Host Bus Adapter
HCA	Host Channel Adapter
IB	InfiniBand
I/O	Input/Output
IOM	IO Module
IP	Internet Protocol

Term	Explanation
IPoIB	IP over InfiniBand
JBOD	Just a Bunch of Disks; storage capacity expansion hardware paired with Mixed Fleet Storage Nodes
LAG	Link aggregation group
MConsole	Management Console; the repset Director role that serves PanActive Manager and PanCLI
MLEC	Multi-Level Erasure Coding
Node	Director and storage hardware; listed as VCH-X,Y
OOBE	Out-Of-Box-Experience
President Director	The Director that governs the realm and replicates its decrees and the realm database to the other repset members
Repset	Replication set; the group of Directors (typically three or five) that hold the realm database for redundancy
SM	Subnet Manager
SR-IOV	Single Root I/O Virtualization
VPOD	Virtualized Protected Object Device
VDURA Data Platform	Hypervisor that runs VPODs and hardware monitoring

Additional Resources

Online Help

Online help is available in the PanActive Manager GUI in two ways:

- Entire online help contents: From the Welcome page, click the Help link to view the entire online help contents for PanActive Manager.
- Page-level help: From a particular page in PanActive Manager, click the Help link to view help for that page.

Product Documentation

The full documentation set for the VDURA V5000 system is available in PDF format from the Welcome page of PanActive Manager or from the Content tab of support.VDURA.com.

Consult the following VDURA documentation for additional information regarding your VDURA system:

- VDURA Data Platform User Guide: Describes how to use PanActive Manager to monitor and manage your VDURA system.
- *VDURA PanCLI Reference Guide:* Provides descriptions of the PanCLI commands used to configure and manage your VDURA system.

Release Notes

Release notes supplement the information in this guide and are available at support.VDURA.com or from the PanActive Manager Home page.

Be sure to examine the release notes before you begin the installation and setup process, as they contain important information that can improve your success with the VDURA system, such as fixed issues and known issues/workarounds from previous versions that may still occur in the current release.

Knowledge Base

VDURA customers have access to the articles in the Knowledge Base tab on support.VDURA.com.

Contact VDURA

There are several ways to contact VDURA, depending on your needs.

Internet	**The support.VDURA.com website provides access to submit a service request and track orders online. Support.VDURA.com also offers access to custom software downloads and a robust knowledge base to assist with technical issues.**
Technical Support	Get technical support and file service requests using any of these methods: - Telephone-North America: 1-888-726-2727 - Telephone-International: 1.412.323.6439 - Email: technical-support@vdura.com
File a Service Request	support.VDURA.com

Safety Information

Review the following basic safety information and precautions before assembling the system. Failure to follow these safety guidelines can result in injury or damage to the VDURA system. Mishandling of the VDURA system voids all warranties. Please read and follow safety guidelines and installation instructions carefully.

Equipment

When you install, configure, add, or remove components in a VDURA system:

- Follow all caution and warning instructions in this manual and on the equipment.
- Do not block or cover the openings to the hardware. Do not install the hardware near a radiator or heat register.
- Follow all hardware-dependent environmental requirements.
- All connections, wiring, wire protection, and wire routing must be in accordance with applicable sections and requirements of your national and local electrical authorities.
- Do not make mechanical modifications to the hardware. Any modifications to the hardware will void the warranty.
- Never lift fully loaded VDURA hardware by yourself.
- Ensure that the voltage and frequency of your power source match the voltage and frequency inscribed on the electrical rating label of the equipment.

Electrostatic Discharge

To ensure that your equipment is not damaged from improper handling:

- When you install or perform maintenance tasks, you must wear a grounded electrostatic discharge (ESD) wrist or ankle strap to protect the system hardware against electrostatic discharge. Make sure that the strap makes skin contact prior to handling VDURA hardware.
- Periodically check the resistance value of the antistatic strap to ensure it is functioning properly.
- Store VDURA hardware in its protective packaging until you are ready to install it.

Rack Mounting Basics

Before you begin the rack mount installation process, please review the following mounting information:

- If installed in a closed or multi-unit rack assembly, the operating ambient temperature of the rack environment may be greater than the ambient room temperature. Therefore, install the equipment in an environment compatible with the specified maximum operating temperature. See System Specifications for more information.
- Install the equipment in a rack so that the amount of airflow required for safe operation of the equipment is not compromised.
- A hazardous condition could be created if the equipment is installed in the rack in a way that produces uneven mechanical loading.
- Be careful not to overload circuits when connecting the equipment to the supply circuit. Observe the equipment nameplate ratings when addressing this concern.
- Be sure to maintain reliable grounding of rack-mounted equipment. Pay particular attention to supply connections other than direct connections to the branch circuit, such as use of power strips.

Power Supply and Power Cord Connection

VDURA products are designed to work with single-phase power systems having a grounded neutral conductor. To reduce the risk of electric shock, do not plug VDURA products into any other type of power system. Contact your facilities manager or a qualified electrician if you are not sure what type of power is supplied to your building. To reduce the risk of electric shock, always plug the cord into a grounded power outlet.

Not all power cords have the same current ratings. Extension cords do not have overload protection and are not meant for use with computer systems. Do not use extension cords with your VDURA system.

All supply connections, wiring, wire protection, and wire routing must be made in accordance with applicable sections and requirements of your national and local electrical authorities.

Before You Begin

Verify that any other VDURA systems connected to the same data network are locked. If VDURA systems on the high-speed data network are left unlocked, they may attempt to automatically add any new hardware found to the existing realm which can be very difficult to reverse.

Installing the VDURA Data Platform requires certain information and items. Gathering these in advance will greatly improve your experience. Please make sure to have the following items before beginning the installation process:

- Cables for the VDURA system: Verify that you have the correct power, network, and SAS cables (if applicable) for your hardware.
- Management (IPMI) network: VDURA Data Platform servers have dedicated 1 GbE network ports for IPMI. While operation of the system does not require the management network, it makes remote operation of the system easier.
- Corporate network: Director servers have dedicated 25 GbE network ports for what is often called a "corporate network" that offers additional access to the VDURA Data Platform separate from the data network. While operation of the system does not require the corporate network, it makes remote operation of the system easier and offers additional network storage protocols to clients not on the main data network.
- Data network: The data network connects clients to the VDURA Data Platform. It may also connect additional resources such as DNS, NTP, and gateway servers.
- Data network configuration: The VDURA Data Platform needs information about the data network during installation. A form at the end of this document lists all required details, including:
- An IP range for the VDURA system: The system needs a minimum of thirteen IP addresses per Storage Node and one IP address per Director Node. The first Director needs an initial address within this range, usually the first address.
- Maximum Transmission Unit (MTU) size.
- System DNS name and DNS domain name.
- DNS server IP address or addresses. (Not required, but may make managing and using the system easier.)
- Timezone: e.g. Europe/Amsterdam.
- VDURA Customer ID (supplied by VDURA).
- PanActive Link enabled or disabled: If the system has access to the public Internet, it can directly access software updates and documentation.
- Network Time Protocol (NTP) server: The VDURA Data Platform and clients connected to it must use the same NTP server. The NTP server may connect to the data network or to the corporate network (if used).

> **Note:** These HTTPS features are currently not available; the system cannot download from VDURA. Download software from support.VDURA.com instead.

> **Note:** When configuring NTP, be aware that NTP server time on a client and the VDURA system must be no more than 10 seconds apart, and that tighter synchronization improves performance.

- DirectFlow Client for compute nodes that access the storage system. These are available on the support.VDURA.com site in versions compiled for multiple Linux distributions and processor architectures.

Order of Installation

Under most circumstances, you will want to follow the order below when installing your new VDURA V5000 system.

1. Ensure the user-owned management and data networks are configured correctly.
2. Physical installation of all hardware. Start at the bottom of the first rack and work up.
3. Power on the Director Nodes.
4. Run network configuration from either PanCLI using a keyboard and monitor or via the BMC remote control features.
5. Start the Out Of Box Experience (OOBE) installation on the first Director Node using either PanCLI or PanActive Manager.
6. Power on JBODs if included in the realm.
7. Power on Storage Nodes one at a time.
8. Complete the VDURA installation via PanActive Manager or PanCLI.
9. Update the VDURA system to the latest software version.
10. Configure IPMI, if applicable.
11. Load the DirectFlow Client onto compute nodes.

Power Procedures

IMPORTANT INFORMATION REGARDING SHUTDOWN/RESTART:

Powering on a JBOD must be done by pressing its power button. If you need to power on a JBOD remotely, VDURA can provide instructions for connecting the JBOD BMC.

At Installation

When installing a V5000 system that includes JBODs, each JBOD must be powered on before powering on its attached Storage Node. VDURA recommends powering on all JBODs, then waiting at least 60 seconds before powering on the Storage Nodes.

During Operation

If the realm is shut down from either PanActive Manager or PanCLI, JBODs will not shut down automatically and must be powered off by hand. JBODs can be powered off by holding down the JBOD enclosure power button located on the front of the unit. Storage Nodes should be powered off prior to shutting down the JBODs.

When powering on the realm after a shutdown, JBODs should be powered on before their associated Storage Nodes because the nodes must be able to see their JBOD drives as part of the power-on process. VDURA recommends powering on all JBODs, then waiting at least 60 seconds before powering on the Storage Nodes.

Use PanActive Manager or PanCLI to power down Storage and Director Nodes; do not remove power at the PDU.

1. Storage Installation

This chapter describes installation and configuration of the VDURA V5000 storage appliance.

This chapter contains the following sections:

- V5000 JBOD Installation Requirements
- V5000 JBOD Overview
- Install the V5000 JBOD Enclosure
- V5000 Storage Node Installation Requirements
- V5000 Storage Node Overview
- Install the V5000 Storage Node
- Cable the Storage Node and JBOD
- Configure Storage Node IPMI Support

> **Important:** Be sure to read the Safety Information. Failure to follow these safety guidelines can result in injury or damage to the equipment and may void your warranty.

This chapter assumes you are installing a standard rack configuration of Mixed Fleet Storage Nodes. Begin with the bottom-most JBOD and Storage Node and work your way up the rack. If you are installing All-Flash Storage Nodes, you may skip the steps related to JBOD installation.

1.1 V5000 JBOD Installation Requirements

A V5000 JBOD enclosure can be mounted in any rack that meets these requirements:

- Compliant with the Electronic Industry Alliance (EIA) 310-D specification.
- Cannot have solid or restricted airflow doors.
- Must have 4U available vertical rack space, plus 1U for the Storage Node directly above it.
- Both the 4U78 and 4U108 JBOD require a rack with mounting posts 19 in. (483 mm) apart.
- Both the 4U78 and 4U108 JBOD have a Cable Management Arm (CMA) that adds to the total depth of the enclosures. The 4U78 JBOD total depth with the CMA is 974.7 mm, and the 4U108 JBOD total depth with the CMA is 1273 mm.
- Some racks have rear doors with features that intrude into the space for the systems mounted in the rack. Review the drawings for your rack including the rear doors to determine if there is enough space for the JBOD.

The following tools and equipment are required to install and configure a V5000 JBOD:

- An electrostatic discharge (ESD) preventative wrist or ankle strap. See Safety Information.

- #2 Phillips screwdriver.
- Appropriate cabling to connect to the V5000 Storage Node. Cables will be in the accessory box in the JBOD packaging.
- Power cables included with the JBOD.
- A server lift is strongly recommended.

1.2 V5000 JBOD Overview

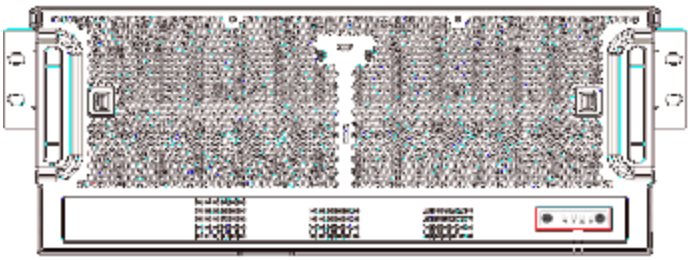
The following descriptions are generally accurate for both the 4U78 and 4U108 V5000 JBOD enclosures. Any exceptions will be noted below.

1.2.1 Control Panel and Drive LEDs

The front of each JBOD contains both internal drive indicator LEDS and a control panel for the enclosure.

Control Panel

Each V5000 JBOD enclosure contains a control panel located at the bottom right corner of the front of the unit. This control panel is used to power the JBOD enclosure on and off and indicate various hardware failures.



V5000 JBOD enclosure, front view; control panel on bottom right with drive LED banks to its left.

The figure above illustrates the 78-drive V5000 JBOD enclosure. The 108-drive JBOD is similar, but with one more LED bank.

The following table explains the items on the JBOD control panel. From left to right:

Item	Description
Mute	- Press: Mute alert tone.
Fan fault LED	- Off: Normal. - Red: Fan failure.

Item	Description
Temperature LED	- Off: Normal. - Red: Overheat condition detected.
Power fail LED	- Off: Normal. - Red: PSU failure.
Power LED	- Blue: Enclosure powered on. - Off: Enclosure powered off.
Power button	- Press: Start system. - Long press: Shut down system.

Drive Indicator LEDs

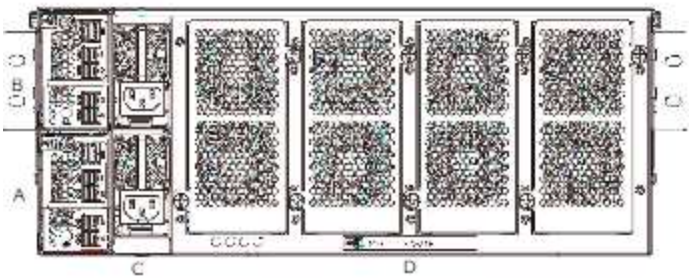
To the left of the control panel, banks of LEDs correspond to the drives inside the JBOD enclosure. Each bank corresponds with a section of drives under part of the top cover of the JBOD. The first drive in the JBOD corresponds with the bottom left-most LED in the left-most bank. Drives increment left to right, bottom to top in each bank, and banks increment left to right.

> **Note:** The first bay is numbered “1”, not “0” as with some other JBODs.

Drive LEDs appear dark blue when a drive is in the slot and functioning normally. A lighter blue or bright purple LED indicates a failed drive in the slot. An unlit LED means no drive is detected in the slot.

1.2.2 Rear Components

The diagram and labels below describe the rear of the V5000 JBOD enclosure.



V5000 JBOD enclosure, rear view.

Label	Description
A	Primary SAS IO Module
B	Secondary SAS IO Module (only installed for some configurations)
C	78-drive JBOD: 2x 1600 W PSUs 108-drive JBOD: 2x 2000 W PSUs
D	8x cooling fans

Each SAS adapter has the following ports (from top to bottom):

Item	Description
BMC LAN port	Labeled “M”
4x Mini SAS HD ports	Labeled “1”, “2”, “3”, and “4” from top to bottom
BMC console port	Labeled “BC”

1.3 Install the V5000 JBOD Enclosure

Preinstallation Information

There are two JBOD options available for the V5000 system: 4U78 and 4U108. These instructions apply to both the 4U78 and 4U108 JBOD enclosures. Be aware that the 4U108 JBOD weighs significantly more than the 4U78 and that the 4U108 needs a 1.2-meter space in front of the rack for service, while the 4U78 needs a 1-meter space.

Before installing the 4U108 enclosure, ensure that the rear of the enclosure does not block the rack power outlets when the enclosure is fully inserted in the rack.

During installation, references to left and right are as seen from the front of the enclosure unless otherwise noted.

When pulled out, each rail is marked on the inside to indicate which end is the rear. Also, the rear of each rail is easily identified by the blue release tab on the arm.

IMPORTANT NOTES REGARDING DRIVES:

- Drives are shipped separately from the JBOD enclosure. DO NOT INSTALL THE DRIVES BEFORE MOUNTING THE ENCLOSURE IN THE RACK.

- Drive location does not matter for a first-time installation. However, if you are moving the VDURA system, or it has been set up at another location before installation at your site, then all drives from a given JBOD must be reinserted into that same JBOD.

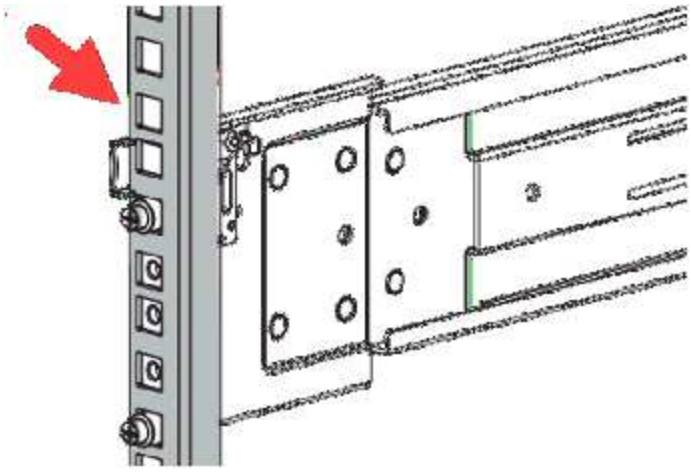
IMPORTANT NOTE REGARDING FRONT BEZEL:

The front bezel is held in place by two small posts and two sliding catches. Consider removing the bezel before working with the JBOD enclosure to prevent damaging it or the enclosure itself.

Step 1: Install Clip Nuts in the Rack

Clip nuts are required for the screws that lock the enclosure into the rack. It is typically easiest to install them before any other parts get in the way.

Insert each clip nut in the bottom hole of the topmost rack unit of the enclosure. For example, if an enclosure is mounted with its bottom at RU 14, count up the three more rack units and install the clip nut in the bottom hole of RU 17. The clip nuts should align with the top hole on each side of the server.



Clip nut location.

Step 2: Install the Outer Rail

1. Pull the front back slightly to find identifying marks on the outside for both front and left/right.
2. The rail system does not have an adjustment at the rear for adjusting the total length. There are no screws to undo to mount the rails in the rack; simply pull the rail out to the correct distance.
3. While one person can install the shorter rail in a rack by themselves, two people will make it easier. For the longer version of the rail for a 1.2-meter rack, two people may be required.
4. Starting with the rail completely folded up, use your thumb to guide the outer part of the rail to about the distance between the front and rear rack posts.
5. Line up the bottom stud with the bottom rack unit in the rack where you want to place the rail.

6. Insert the stud into the bottommost space in the rack unit and then gradually raise the rail until the blue top clip engages.
7. At the front, pull the rail forward and engage the tail studs with the holes in the rack.

The rail kit comes with three types of screws to install the JBOD enclosure. Inside the bag of screws are four baggies. The largest screws are for locking the enclosure against the rack posts. The four smaller screws with triangular heads are used for securing the rails and are matched with a washer to help make sure the screw does not slip through the rack post. The final screws are used for attaching the rail to the chassis.

8. To prepare the screws used to hold the rail against the rack posts, add one washer to each triangular-headed screw. The washer should cup the head of the screw.
9. Insert the screw with the washer into the middle hole of the rail to secure the rail to the rack post.
10. Tighten the screw by hand with a #2 Phillips screwdriver until snug. Make sure that when you finish tightening the screw, the posts on the end of the rail are fully pushed through the rack.
11. Attach another screw and washer in the same middle hole position at the rear of the rail.
12. When you are finished attaching the rail on one side, follow the same process to attach the other rail on the other side.

Step 3: Install the Inner Rail

1. Remove the inner rail before attaching to the chassis.
    - Slide the middle rail out.
    - Slide the inner rail out until it clicks.
    - Release the inner rail by pushing from the outside on the tab and sliding it all the way out.
    - To move the rail back into the rack, push on the tab next to the blue mark to unlock the rail.

To determine correct orientation, the inner rails are stamped L and R to show side of the chassis the inner rail attaches. The clamps to attach the Cable Management Arm (CMA) are on the rear of the inner rails.

2. Align the inner rail with the mushroom studs on the chassis.
3. Fit the rail through all the matching keyholes in the rail.
4. Make sure all holes are lined up and gently push the rail toward the rear. A clip automatically engages to secure the rail in place.
5. Attach three screws on each side to lock in the inner rail to the enclosure. Tighten until snug.
6. When you are finished attaching the rail on one side, follow the same process to attach the other rail on the other side.

Step 4: Engage the Middle Rail

If the enclosure and the rack are not perfectly aligned, and you cannot get the enclosure at the same height as the rails, try to get one side lined up as closely as possible to start and leave the other side a little lower.

During these steps, it may be necessary to slightly lift each side as you join the inner and middle rails.

If the sides of the enclosure are significantly different in alignment, letting it sit could twist the rail too much. Therefore, you may need to use two people to engage both middle rails at the same time.

1. To attach the rail, make sure the enclosure is close to the rack.
2. Pull the middle rail out.
3. There is a ball bearing race on the inside of the middle rail that can move back and forth. Make sure that the race is all the way forward and ready to engage with the rail as you slide the rail over it; otherwise, the enclosure can knock it loose, and the ball bearings will fall out. Best practice is to wrap your thumb around the middle rail to prevent the race from moving.
4. Either slide the inner rail over the middle rail or slide the inner rail inside the middle rail. Note that the end of the middle rail is painted to make it easy to spot that the rail lips at the top and the bottom are fully engaged.
5. While continuing to grip the ball bearing race with your thumb, push everything forward until the middle rail locks.
6. When you are finished attaching the rail on one side, follow the same process to attach the other rail on the other side.
7. At this point, the chassis is firmly in the middle rail.

If it is difficult to extend the middle rail, that may be because the chassis is out of alignment. Try raising the chassis a bit to align it better.

8. Lower the server lift just enough so that you can move the server back and forth freely.
9. The clips on the inner rail are not yet engaged with the middle rail. Gently move the server forward along the rails until these clips are engaged. Make sure that the enclosure locks in the service position.

Step 5: Slide the Enclosure into the Rack

1. Hold the clips on both sides at the same time and push the enclosure in a small amount to make sure it moves freely and then continue to move the enclosure until it's fully inserted into the rack.
2. Slide the enclosure in and out of the rack at least three times to make sure that it is properly installed on the rails.
3. Gently wiggle the enclosure side to side and up and down to make sure that it stays in place. If you push too hard, you will break the rail.

4. Clip nuts are required for the screws that lock the enclosure into the rack. If you have not already done so, install the clip nuts now. See the "Step 1: Install Clip Nuts in the Rack" section at the beginning of these instructions.

Step 6: Insert the Drives

1. Before installing the drives, pull the server all the way out until it locks into position.
2. Make sure a server lift is still positioned underneath the enclosure to catch the enclosure if it falls while installing drives.
3. Open the drive covers by sliding their latches and lifting their front edges.
4. Gently lower the drives into the bays. Due to manufacturing tolerances, some bays may be a tighter fit than others.
5. As you insert each drive, make sure to push it all the way down and continue to push it while you close the latch, since the latch itself will not pull the drive into position; it only locks it in place.
6. Firmly push down while closing the latch to make sure the drive is properly seated. As you push the drive down to close the latch, make sure to push on the black plastic, not the button.
7. When finished, check all the drives to ensure that no drive is higher than the others.
8. Lower the covers.
9. Once all drives are in place and the covers are lowered, slide the enclosure in and out of the rack several times to make sure all is working properly with the rails. With the full weight of the drives, you may notice that the chassis does not move as easily in and out of the enclosure; this is normal.

> **Note:** Due to close rack tolerances, the top JBOD covers may catch on the bottom of the server above it when sliding the JBOD into the rack. If this occurs, best practice is to pull the JBOD back out, then pull the server above it out. While pushing up slightly on the server, push the JBOD into the rack. After the JBOD is in place, push the server in, as well.

Step 7: Install the Cable Management Arm (CMA)

The CMA is attached to the rails at the rear of the JBOD. It holds all cables neatly in place and prevents issues while the unit slides in and out of the rack.

1. Make sure the enclosure is slid all the way in to the rear of the rack.
2. Before installing, the clip on the support arm under the CMA must first be flipped to the opposite direction. Push the silver button to disengage the small lever. Rotate the lever 180 degrees until it snaps into position on the other side.

3. At the rear of the JBOD, each rail has two C-shaped clamps attached to the ends of the inner and middle rails. As viewed from the rear, slide the CMA clip that you rotated in the previous step (with a small blue tab on it) onto the C-shaped clamp located on the middle rail of the right side. The correct clamp protrudes further to the rear than the one on the inner rail.
4. The other side of the CMA has two arms and two attachments. Make sure the CMA is slid shut to its shortest length so these arms fit inside the rack.
5. Slide the outer arm onto the outer (middle rail) C-shaped clamp. Slide the inner arm onto the inner rail C-shaped clamp.
6. The CMA will now open and close with a scissor-like motion as the enclosure is slid in and out of the rack, supporting the cables and preventing tangles and snags.

The distance between the last cable cage on the CMA and the PSUs and IOMs increases a bit as the enclosure slides out. This is important for gauging how much slack to leave in the cables at this end; you do not want to route the cable so tightly that as the enclosure is moved forward, it pulls on the cables and makes it difficult to pull the enclosure all the way out of the rack.

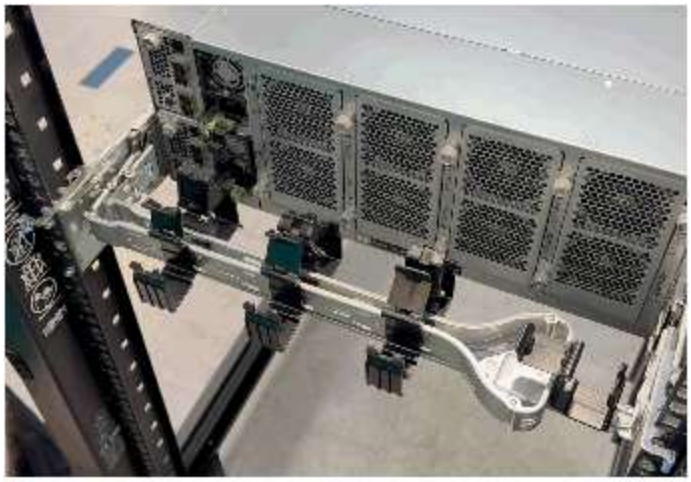


CMA installed on JBOD.

The figure above shows a CMA installed and partially extended. Note the short vertical post on the support arm that holds the CMA up as the JBOD extends; this post points up when the CMA is installed correctly.

Step 8: Connect and Route the Cables

1. Open all the cages or baskets that hold the cables.



CMA with cable cages open.

2. Connect the power cables to the PSUs at the rear of the JBOD.
3. Attach the cable retention strap so that the power cord does not accidentally pull out of the PSU.
    - Pull the tab on the cable retention clip ring to open it.
    - Fit the ring around the cable.
    - Close the cable retention ring clip, but not tightly yet.
    - The clip ring slides on a strap connected to the PSU. There is a second tab on the clip ring that allows the ring to slide on the strap. Move this tab to loosen it, then push the clip ring snugly against the power cord plug.
    - Release the tab and make sure the ring remains snugly in place.
    - Squeeze the ring until it is tightly fitted around the cord. You may need to push and pull the ring catch a little bit to properly squeeze the retention cord.
    - Repeat this process for the other power supply's cord.
4. Lay the cables in the baskets and route the cables through until their entire length is supported. Make sure there is enough slack at the final bend where cords plug into the PSUs and the middle joint of the CMA so that when the enclosure is pulled all the way out, the cables do not want to pull out of the PSUs.
5. Attach the SAS cables. For enclosures with dual-actuator drives, all eight cables will get connected to both IOMs. For other drives, only four cables will get connected to a single IOM.

When you connect the cable to the IOM, you will notice that the cables are stiff and have a sharp bend as they connect to module. This bend is expected and within the bend tolerance of the cables.

6. As with the power cables, route the SAS cables through the baskets, leaving some slack in the middle.

7. Lay the SAS cables above the power cables to keep the baskets neat. With eight cables, the baskets are full and leave little room for cable movement. However, with careful placement and the correct amount of slack, everything will work correctly. Try to lay multiple neat layers of cables within the baskets to give as much room as possible for the cables to slide back and forth.
8. Make sure to give some extra bend radius where the SAS cables leave the last basket and connect to the IOMs, as some of them need to bend upwards to reach their ports.
9. At this point, the provided zip tie may be useful to relieve the strain on the final basket, as the clips on the baskets are only so strong and may pop open. Make sure the zip tie is not so tight that it prevents the cables from moving back and forth.
10. When you have all the cables installed and arranged, close the baskets.



CMA with cable baskets filled and closed.

11. Pull the enclosure out slowly and carefully to test the cabling. Monitor the cables to make sure they are not pulling too hard.

> **Warning:** If the cables do not have enough slack before the CMA or inside the CMA, they can pull out connectors or, in the case of the power cord, the retention strap.

12. If needed, adjust the cable slack at the joints so that the enclosure can be fully extended into the service position, and make sure that the cables are aligned appropriately.
13. Monitor the enclosure again as you push it back into the rack to make sure that the cables fold into their correct positions.

Step 9: Secure the Enclosure to the Rack

1. Use the largest installation screws provided to lock the enclosure in the rack.
2. Insert the screws into the cage nuts at the rear of the rack and tighten until snug.

Step 10: Install the Front Bezel

If you removed the bezel, you can reinstall it after finishing the steps above.

1. There are two pins and two catches on the front of the JBOD enclosure. Match these to their corresponding locations on the rear of the bezel.
2. Push the bezel until it clicks in place.

1.4 V5000 Storage Node Installation Requirements

A V5000 Storage Node can be mounted in any rack that meets these requirements:

- Compliant with the Electronic Industry Alliance (EIA) 310-D specification.
- Cannot have solid or restricted airflow doors.
- Must have 1U available rack space, plus 4U for the JBOD directly below it for a Mixed Fleet Storage Node.
- The V5000 Storage Node requires a rack with mounting posts 19 in. (483 mm) apart and with at least 1000 mm available depth.

The following tools and equipment are required to install and configure a V5000 Storage Node:

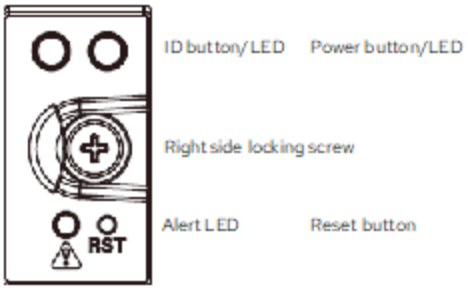
- An electrostatic discharge (ESD) preventative wrist or ankle strap. See Safety Information.
- #2 Phillips screwdriver.
- Power cables included with the V5000 Storage Node.
- Data and management network cables to connect the V5000 Storage Node to the switches are not included and must be provided for the installation.

1.5 V5000 Storage Node Overview

This section applies to installation of both Mixed Fleet Storage Nodes (those with JBODs) and All-Flash Storage Nodes (server-only). All steps are the same for both types of node, except that All-Flash Storage Nodes do not have an SAS HBA and do not need to be installed above a JBOD. Therefore, if you are installing All-Flash Storage Nodes, you may ignore the Cable the Storage Node and JBOD section below.

1.5.1 Control Panel

The front of each Storage Node contains a control panel on the right edge. The control panel is used to power the node on and off, reset it, identify it, and display the status of the node.



Control panel on the right-hand side of the bezel.

The following table explains the items on the control panel for each node.

Item	Description
Power button/LED	- Green: Node is powered on. - Unlit: Node is powered off. - Press to turn power on and off. Note that when off, the node still receives standby power as long as it is plugged in. - If the node loses all power, when power is restored, the node will revert to its last power state, i.e. power back on if it was on prior to power loss or stay off if it was already turned off when power was lost.
ID button/LED	- Blue: Lights up in response to the IPMI ID command, helping you locate a chassis.
Alert LED	- Red: Lights whenever the BMC detects a failure.
Reset button	- Press to reboot the node.

1.5.2 Bezel Power

The front of the node contains a USB port located on the left side. This port is used to power the LED in the bezel and is not visible or usable when the bezel is in place.

1.5.3 Drive Bay Status LEDs

Each drive bay in a Storage Node has three LEDs.

Status LED	Activity Description
Blue Drive Activity Indicator	- Solid: SSD is installed. - Blinking: Drive activity detected. - Off: No drive detected.
Yellow Drive Failure Indicator	- Solid: Drive fault. - Blinking: Drive rebuild. - Off: Normal.
Green Location Indicator	- Solid: Drive location activated. - Off: Normal.

1.5.4 Rear Ports

The rear of the Storage Node contains several ports for connectivity:



V5000 Storage Node rear connections.

Label	Port
A	1 GbE IPMI
B	2x USB 3.2
C	2x 1 GbE RJ45 for VDURA use only
D	2x 10 GbE RJ45 for VDURA use only
E	VGA monitor
Slot 1	See Node Connectivity Configurations below.
Slot 2	See Node Connectivity Configurations below.
Slot 3	See Node Connectivity Configurations below.

Label	Port
OCP 3.0 Slot	See Node Connectivity Configurations below.

1.5.5 Node Connectivity Configurations

Nodes can have either Ethernet or InfiniBand HPC network cards, but not both. Ports will run at 2x200 GbE, but can run at 2x100 GbE, if needed.

Storage Node with Ethernet Option

The following table lists each port type with its device, purpose, and location.

Port Type	Device	Purpose	Location
SAS HBA	4-port SAS PCIe HHHL	JBOD connectivity	Slot 1 or 2
Ethernet	2-port 200 GbE PCIe HHHL	HPC network connectivity	Slot 3
Ethernet	2-port 200 GbE OCP SFF	HPC network connectivity	OCP slot

Storage Node with NDR200 InfiniBand Option

The following table lists each port type with its device, purpose, and location.

Port Type	Device	Purpose	Location
SAS HBA	4-port SAS PCIe HHHL	JBOD connectivity	Slot 1 or 2
InfiniBand	2-port QSFP112 200 GbE / NDR200 PCIe HHHL	HPC network connectivity	Slot 3
InfiniBand	2-port QSFP112 200 GbE / NDR200 PCIe OCP SFF	HPC network connectivity	OCP slot

Storage Node with NDR400 InfiniBand Option

The following table lists each port type with its device, purpose, and location.

Port Type	Device	Purpose	Location
SAS HBA	4-port SAS PCIe HHHL	JBOD connectivity	Slot 1 or 2
InfiniBand	1-port OSFP NDR400 PCIe HHHL	HPC network connectivity	Slot 3
InfiniBand	1-port OSFP NDR400 PCIe OCP	HPC network connectivity	OCP slot

1.5.6 NIC Status LEDs

Storage Nodes have either Ethernet or InfiniBand NICs installed in the rear of the unit. All Mixed Fleet Storage Nodes also contain an SAS card.

Ethernet-Based Systems

PCIe slot 3 contains a Broadcom dual-port 200/400 GbE NIC. These LEDs are located on the outer edges of each port.

Status LED	Activity Description
Activity	- Off: No activity. - Green blinking: Link is up; traffic flowing.
Link	- Off: No link. - Green: 200 GbE or 400 GbE link. - Yellow: 100 GbE link.

The OCP slot contains a slightly different Broadcom dual-port 200/400 GbE NIC. Link LEDs are located on the left side of each port, while the Activity LEDs are on the right.

NDR200 InfiniBand-Based Systems

PCIe slot 3 contains an NVIDIA dual-port NDR200 InfiniBand NIC with two LEDs per port. LED 1 (green/yellow/off) is located closest to each port; LED 2 (green/off) is located to the outside of LED 1.

LED 1	LED 2	Activity Description
- Blinking green at 1 Hz: full port speed - Blinking yellow at 1 Hz: less than full port speed	- Off	- Location mode to identify NIC.
- Blinking yellow at 4 Hz	- On	Error: - I2C access to the networking ports failed. Blinks until error is fixed. - Over-current condition of the networking ports. Blinks until error is fixed.
- Solid green: Full port speed - Blinking yellow: Less than full port speed	- Blinking: This LED blinks green when data is being transferred at full port speed.	- Activity.
- Solid green: Full port speed - Blinking yellow: Less than full port speed	- On: This LED lights solid when the server is powered on and a cable is plugged into the port with the other end of the cable plugged into a functioning port.	- Link is up.
- Solid yellow	- Off	- Physical up.

The OCP slot contains an NVIDIA dual-port NDR200 InfiniBand NIC with two LEDs per port. The LEDs are located between the ports, with LED 0 (green/yellow/off) on the left side of each pair and LED 1 (green/off) on the right.

LED 0	LED 1	Activity Description
- Blinking yellow at 1 Hz	- Off	- Location mode to identify NIC.

LED 0	LED 1	Activity Description
- Blinking yellow at 4 Hz	- On	Error: - I2C access to the networking ports failed. Blinks until error is fixed. - Over-current condition of the networking ports. Blinks until error is fixed.
- Blinking green	- Blinking: This LED blinks green when data is being transferred at full port speed.	- Activity.
- Solid green: Full port speed - Solid yellow: Less than full port speed	- On: This LED lights solid when the server is powered on and a cable is plugged into the port with the other end of the cable plugged into a functioning port.	- Link is up.

NDR400 InfiniBand-Based Systems

PCIe slot 3 contains an NVIDIA single-port NDR400 InfiniBand NIC. LED 1 (green/yellow/off) is located directly to the left of the port; LED 2 (green/off) is located to the left of LED 1.

LED 1	LED 2	Activity Description
- Blinking green at 1 Hz: full port speed - Blinking yellow at 1 Hz: less than full port speed	- Off	- Location mode to identify NIC.
- Blinking yellow at 4 Hz	- On	Error: - I2C access to the networking ports failed. Blinks until error is fixed. - Over-current condition of the networking ports. Blinks until error is fixed.

LED 1	LED 2	Activity Description
- Solid green: Full port speed - Blinking yellow: Less than full port speed	- Blinking: This LED blinks green when data is being transferred at full port speed.	- Activity.
- Solid green: Full port speed - Blinking yellow: Less than full port speed	- On: This LED lights solid when the server is powered on and a cable is plugged into the port with the other end of the cable plugged into a functioning port.	- Link is up.
- Solid yellow	- Off	- Physical up.

The OCP slot contains an NVIDIA single-port NDR400 InfiniBand NIC with two LEDs. LED 0 (green/yellow/off) is on the left; LED 1 (green/off) is on the right.

LED 0	LED 1	Activity Description
- Blinking yellow at 1 Hz	- Off	- Location mode to identify NIC.
- Blinking yellow at 4 Hz	- On	Error: - I2C access to the networking ports failed. Blinks until error is fixed. - Over-current condition of the networking ports. Blinks until error is fixed.
- Blinking green	- Blinking: This LED blinks green when data is being transferred at full port speed.	- Activity.
- Solid green: Full port speed - Solid yellow: Less than full port speed	- On: This LED lights solid when the server is powered on and a cable is plugged into the port with the other end of the cable plugged into a functioning port.	- Link is up.

SAS Card

All Mixed Fleet Storage Nodes above contain a Broadcom SAS card in PCIe slot 1 or 2 to connect to the JBOD. There are two LEDs located on the board that can be seen through the screen.

Status LED	Activity Description
Green	Operating status: - Blinking: Normal operation. - Off: Potential problem.
Yellow	Temperature status: - On: Temperature sensor is over threshold. - Off: Normal operating temperature.

1.5.7 Power Supply Status LED

Each power supply has a status LED.

Status LED	Activity Description
Green	- Solid: Powered and functioning normally. - Off: No power to both PSUs. - Blinking: Standby; AC power connected.
Red	- Solid: Critical event. PSU shutdown: failure, OCP, OVP, fan failure. - Blinking: Non-critical event. PSU continues to operate: high temp, high power, high current, slow fan.
Red and green alternating	- Blinking alternate colors: Warning event. PSU continues to operate: high temp, high power, high current, slow fan.

1.6 Install the V5000 Storage Node

Follow the steps below to install a V5000 Storage Node in a rack.

Preinstallation Information

IMPORTANT NOTE REGARDING FRONT BEZEL:

The front bezel is held in place by two small posts and two sliding catches. Consider removing the bezel before working with the node to prevent damaging it or the enclosure itself.

Step 1: Unpack Storage Node

1. Unpack the Storage Node and rails from the shipping kit. Place the enclosure on a smooth stable surface.
2. Remove any protective film from the top of the enclosure.
3. Verify that all ventilation openings in the enclosure are clear and unobstructed.

Step 2: Install Rails in the Rack

The rail kits require no tools to install; they slide and snap into place.

1. Remove the inner rail from the outer rail. Press the latch inside the rail assembly to allow the inner rail to fully slide out.
2. Attach the inner rail to the Storage Node. Align the holes in the rail to the studs on the side of the node. Once all studs are in their holes, slide the rail to the rear to lock it in place.
3. Attach the outer rail to the rack. Align the front and rear of the outer rail with the correct holes in the rack posts. The rear connector fits over the outside of the rack post.
4. Push front and rear rack studs into the rack post holes until they click in place.
5. Repeat the steps above for the other rail.

Step 3: Install the Enclosure in the Rails

1. Make sure the ball bearing races located in the middle rails are all the way forward.
2. Pull the middle rails all the way out until they lock.
3. Align the rear of the node with the front of the middle rails and push the inner rails into the middle rails.
4. Slide the release latch and push the node into the rack.
5. Secure the node in the rack by inserting the Phillips-head screws on the left and right sides of the front of the node. Tighten until snug.

Step 4: Attach Power Cables

Each Storage Node is equipped with two power supply units (PSUs) to provide redundant power to the system. These PSUs are located on the left of the rear of the node.

Plug one end of each provided power cable into a PSU and the other end into an AC outlet. VDURA recommends that each PSU be plugged into a different circuit or a UPS to protect against power failures.

Step 5: Attach Network Cables

Connect network cables from the high-speed ports to your uplink switches.

Step 6: Install the Front Bezel

If you removed the bezel, reinstall it after finishing the steps above.

1. There are two pins and two catches on the front of the Storage Node. Match these to their corresponding locations on the rear of the bezel.
2. Push the bezel until it clicks in place.

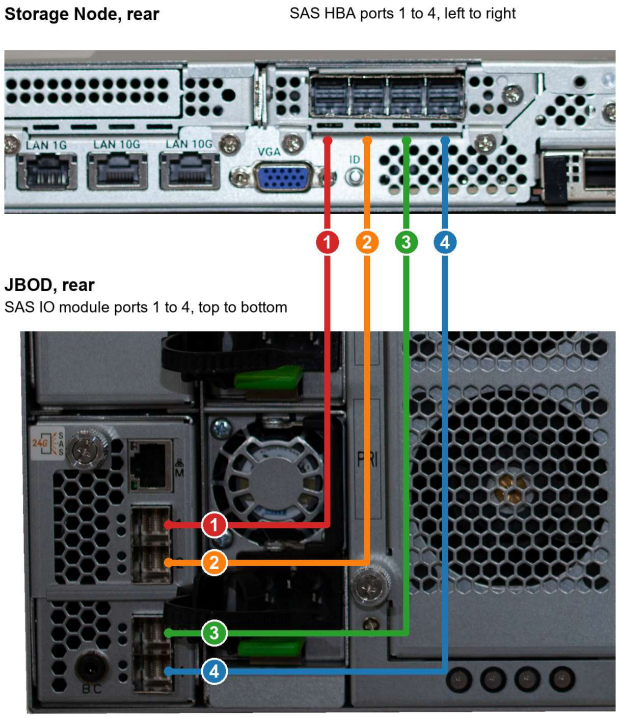
> **Warning:** Due to close tolerances, it is possible to press the power button with the bezel during installation.

1.7 Cable the Storage Node and JBOD

Once you have installed a Mixed Fleet Storage Node/JBOD pair in a rack, you must connect them with SAS cables. JBODs have four cables leading to one HBA on the corresponding Storage Node.

1. If you have not done so already, install the SAS cables on the rear of the JBOD per the instructions in Install the V5000 JBOD Enclosure.
2. Connect the other ends of the SAS cables to the Mixed Fleet Storage Node. Make sure that you are pairing the JBOD to the correct node.
    - Attach the SAS cables such that the top-most cable on the JBOD connects to the left-most port on the Storage Node.
    - Connect the rest of the cables, moving top to bottom on the JBOD and left to right on the Storage Node; i.e. the top-most JBOD cable matches the left-most location on the Storage Node.

The following figure shows the SAS cable connections between a JBOD and a Storage Node.



SAS cabling between a V5000 JBOD and a Storage Node. JBOD SAS IO module ports 1 to 4, from top to bottom, connect to SAS HBA ports 1 to 4, from left to right.

1.8 Configure Storage Node IPMI Support

IPMI is not required for proper realm functionality. To configure IPMI on Storage Nodes, see IPMI Configuration. That chapter also covers the IPMI password rules and how to verify the BMC network settings of a Storage Node.

2. Director Node Installation

This chapter describes installation and configuration of the VDURA V5000 Director Nodes.

This chapter contains the following sections:

- Installation Requirements
- V5000 Director Node Overview
- Install the V5000 Director Node
- Connect to the Console
- Configure Director Node IPMI Support

> **Important:** Be sure to read the Safety Information. Failure to follow these safety guidelines can result in injury or damage to the equipment and may void your warranty.

2.1 Installation Requirements

A V5000 Director Node can be mounted in any rack that meets these requirements:

- Compliant with the Electronic Industry Alliance (EIA) 310-D specification.
- Cannot have solid or restricted airflow doors.
- Must have 1U available rack space.
- The V5000 Director Node requires a rack with mounting posts 19 in. (483 mm) apart and with at least 1000 mm available depth.

The following tools and equipment are required to install and configure a V5000 Director Node:

- An electrostatic discharge (ESD) preventative wrist or ankle strap. See Safety Information.
- #2 Phillips screwdriver.
- Power cables included with the V5000 Director Node.
- Data and management network cables to connect the V5000 Director Node to the switches are not included and must be provided for the installation.

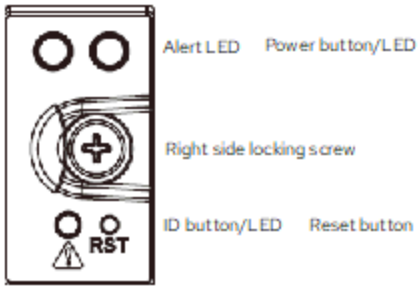
2.2 V5000 Director Node Overview

VDURA V5000 Director Nodes are the control plane of the storage system, handling metadata, data reconstruction, and non-DirectFlow connectivity tasks. Director Nodes can have either Ethernet or InfiniBand HPC network cards, but not both. Depending on the specific hardware configuration, each port may support a speed of 400 Gb, 200 Gb, or 100 Gb.

Note that the 25 GbE network port will be connected to a different network fabric than the HPC network. This port is most often used to connect to a corporate network instead of the compute cluster high-speed data network. Most customers will open PanActive Manager through the corporate port. Customers using S3 connections will likely use this port, as well.

2.2.1 Control Panel

The front of each Director Node contains a control panel on the right edge. The control panel is used to power the node on and off, reset it, identify it, and display the status of the node.



Control panel on the right-hand side of the bezel.

The following table explains the items on the control panel for each node.

Item	Description
Power button/LED	- Green: Node is powered on. - Unlit: Node is powered off. - Press to turn power on and off. Note that when off, the node still receives standby power as long as it is plugged in. - If the node loses all power, when power is restored the node will revert to last power state, i.e. power back on if it was on prior to power loss or stay off if it was already turned off when power was lost.
ID button/LED	- Blue: Lights up in response to the IPMI ID command, helping you locate a chassis.
Alert LED	- Red: Lights whenever the BMC detects a failure.
Reset button	- Press to reboot the node.

2.2.2 Bezel Power

The front of the node contains a USB port located on the left side. This port is used to power the LED in the bezel and is not visible or usable when the bezel is in place.

2.2.3 Drive Bay Status LEDs

Each drive bay in a Director Node has three LEDs.

Status LED	Activity Description
Blue Drive Activity Indicator	- Solid: SSD is installed. - Blinking: Drive activity detected. - Off: No drive detected.
Yellow Drive Failure Indicator	- Solid: Drive fault. - Blinking: Drive rebuild. - Off: Normal.
Green Location Indicator	- Solid: Drive location activated. - Off: Normal.

2.2.4 Rear Ports

The rear of the Director Node contains several ports for connectivity:



V5000 Director Node rear connections.

Label	Port
A	1 GbE IPMI
B	2x USB 3.2
C	2x 1 GbE RJ45 for VDURA use only
D	2x 10 GbE RJ45 for VDURA use only
E	VGA monitor
Slot 1	See Node Connectivity Configurations below.

Label	Port
Slot 2	See Node Connectivity Configurations below.
Slot 3	See Node Connectivity Configurations below.
OCP 3.0 Slot	See Node Connectivity Configurations below.

2.2.5 Node Connectivity Configurations

Director Nodes can have either Ethernet or InfiniBand high-speed data network cards, but not both. Ethernet ports will run at 2x200 GbE, but can run at 2x100 GbE, if needed. This single dual-port card is located in either a PCIe HHHL slot or the OCP slot. For NDR400, the Director will always have two single-port cards.

Director Node with Ethernet Option (Single NIC)

The following table lists each port type with its device, purpose, and location.

Port Type	Device	Purpose	Location
Ethernet	2-port 25/10 GbE PCIe HHHL	Corporate network connectivity	Slot 2
Ethernet	2-port 200 GbE PCIe OCP	HPC network connectivity	OCP slot

Director Node with Ethernet Option (Dual NIC)

The following table lists each port type with its device, purpose, and location.

Port Type	Device	Purpose	Location
Ethernet	2-port 25/10 GbE PCIe HHHL	Corporate network connectivity	Slot 2
Ethernet	2-port QSFP112 200/400 GbE PCIe HHHL	HPC network connectivity	Slot 3

Port Type	Device	Purpose	Location
Ethernet	2-port QSFP112 200/400 GbE PCIe OCP	HPC network connectivity	OCP slot

Director Node with NDR200 InfiniBand Option (Single NIC)

The following table lists each port type with its device, purpose, and location.

Port Type	Device	Purpose	Location
Ethernet	2-port 25/10 GbE PCIe HHHL	Corporate network connectivity	Slot 2
InfiniBand	2-port QSFP112 NDR200 PCIe OCP	HPC network connectivity	OCP slot

Director Node with NDR200 InfiniBand Option (Dual NIC)

The following table lists each port type with its device, purpose, and location.

Port Type	Device	Purpose	Location
Ethernet	2-port 25/10 GbE PCIe HHHL	Corporate network connectivity	Slot 2
InfiniBand	2-port QSFP112 NDR200 PCIe HHHL	HPC network connectivity	Slot 3
InfiniBand	2-port QSFP112 NDR200 PCIe OCP	HPC network connectivity	OCP slot

Director Node with NDR400 InfiniBand or 400 GbE OSFP Option (Dual NIC)

The following table lists each port type with its device, purpose, and location.

Port Type	Device	Purpose	Location
Ethernet	2-port 25/10 GbE PCIe HHHL	Corporate network connectivity	Slot 2

Port Type	Device	Purpose	Location
InfiniBand/Ethernet	1-port OSFP NDR400/400 GbE PCIe HHHL	HPC network connectivity	Slot 3
InfiniBand/Ethernet	1-port OSFP NDR400/400 GbE PCIe OCP	HPC network connectivity	OCP slot

2.2.6 NIC Status LEDs

Each Director Node has Ethernet and/or InfiniBand NICs installed in the rear of the unit. All Director Nodes contain a 25 GbE NIC for corporate network connectivity.

Ethernet-Based Systems

PCIe slot 3 or the OCP slot contains a Broadcom dual-port 200 GbE NIC. These LEDs are located on the outer edges of each port.

Status LED	Activity Description
Activity	- Off: No activity. - Green blinking: Link is up; traffic flowing.
Link	- Off: No link. - Green: 200 GbE or 400 GbE link. - Yellow: 100 GbE link.

NDR200 InfiniBand-Based Systems

PCIe slot 3 contains an NVIDIA dual-port NDR200 InfiniBand NIC with two LEDs per port. LED 1 (green/yellow/off) is located closest to each port; LED 2 (green/off) is located to the outside of LED 1.

LED 1	LED 2	Activity Description
- Blinking green at 1 Hz.: Full port speed. - Blinking yellow at 1 Hz: Less than full port speed.	- Off	- Location mode to identify NIC.

LED 1	LED 2	Activity Description
- Blinking yellow at 4 Hz	- On	Error: - I2C access to the networking ports failed. Blinks until error is fixed. - Over-current condition of the networking ports. Blinks until error is fixed.
- Solid green: Full port speed. - Blinking yellow: Less than full port speed.	- Blinking: This LED blinks green when data is being transferred at full port speed.	- Activity.
- Solid green: Full port speed. - Blinking yellow: Less than full port speed.	- On: This LED lights solid when the server is powered on and a cable is plugged into the port with the other end of the cable plugged into a functioning port.	- Link is up.
- Solid yellow	- Off	- Physical up.

The OCP slot contains an NVIDIA dual-port NDR200 InfiniBand NIC with two LEDs per port. The LEDs are located between the ports, with LED 0 (green/yellow/off) on the left side of each pair and LED 1 (green/off) on the right.

LED 0	LED 1	Activity Description
- Blinking yellow at 1 Hz	- Off	- Location mode to identify NIC.
- Blinking yellow at 4 Hz	- On	Error: - I2C access to the networking ports failed. Blinks until error is fixed. - Over-current condition of the networking ports. Blinks until error is fixed.

LED 0	LED 1	Activity Description
- Blinking green	- Blinking: This LED blinks green when data is being transferred at full port speed.	- Activity.
- Solid green: Full port speed. - Solid yellow: Less than full port speed	- On: This LED lights solid when the server is powered on and a cable is plugged into the port with the other end of the cable plugged into a functioning port.	- Link is up.

NDR400 InfiniBand-Based Systems

PCIe slot 3 contains an NVIDIA single-port NDR400 InfiniBand NIC with two LEDs per port. LED 1 (green/yellow/off) is located closest to each port; LED 2 (green/off) is located to the outside of LED 1.

LED 1	LED 2	Activity Description
- Blinking green at 1 Hz.: Full port speed. - Blinking yellow at 1 Hz: Less than full port speed.	- Off	- Location mode to identify NIC.
- Blinking yellow at 4 Hz	- On	Error: - I2C access to the networking ports failed. Blinks until error is fixed. - Over-current condition of the networking ports. Blinks until error is fixed.
- Solid green: Full port speed. - Blinking yellow: Less than full port speed.	- Blinking: This LED blinks green when data is being transferred at full port speed.	- Activity.

LED 1	LED 2	Activity Description
- Solid green: Full port speed - Blinking yellow: Less than full port speed	- On: This LED lights solid when the server is powered on and a cable is plugged into the port with the other end of the cable plugged into a functioning port.	- Link is up.
- Solid yellow	- Off	- Physical up.

The OCP slot contains an NVIDIA single-port NDR400 InfiniBand NIC with two LEDs per port. The LEDs are located between the ports, with LED 0 (green/yellow/off) on the left side of each pair and LED 1 (green/off) on the right.

LED 0	LED 1	Activity Description
- Blinking yellow at 1 Hz	- Off	- Location mode to identify NIC.
- Blinking yellow at 4 Hz	- On	Error: - I2C access to the networking ports failed. Blinks until error is fixed. - Over-current condition of the networking ports. Blinks until error is fixed.
- Blinking green	- Blinking: This LED blinks green when data is being transferred at full port speed.	- Activity.
- Solid green: Full port speed. - Solid yellow: Less than full port speed.	- On: This LED lights solid when the server is powered on and a cable is plugged into the port with the other end of the cable plugged into a functioning port.	- Link is up.

Corporate Network

All systems above contain a Broadcom dual-port 25 GbE NIC in PCIe slot 2 to connect to the corporate network. These LEDs are located beside each port.

Status LED	Activity Description
Activity	- Off: No activity. - Green blinking: Link is up; traffic flowing.
Link	- Off: No link. - Green: 25 GbE link. - Yellow: 10 GbE or 1 GbE link.

2.2.7 Power Supply Status LED

Each power supply has a status LED.

Status LED	Activity Description
Green	- Solid: Powered and functioning normally. - Off: No power to both PSUs. - Blinking: Standby; AC power connected.
Red	- Solid: Critical event. PSU shutdown: failure, OCP, OVP, fan failure. - Blinking: Non-critical event. PSU continues to operate: high temp, high power, high current, slow fan.
Red and green alternating	- Blinking alternate colors: Warning event. PSU continues to operate: high temp, high power, high current, slow fan.

2.3 Install the V5000 Director Node

Preinstallation Information

IMPORTANT NOTE REGARDING FRONT BEZEL:

The front bezel is held in place by two small posts and two sliding catches. Consider removing the bezel before working with the node to prevent damaging it or the enclosure itself.

Step 1: Unpack Director Node

1. Unpack the Director and rails from the shipping kit. Place the enclosure on a smooth stable surface.
2. Remove any protective film from the top of the enclosure.
3. Verify that all ventilation openings in the enclosure are clear and unobstructed.

Step 2: Install Rails in the Rack

The rail kits require no tools to install; they slide and snap into place.

1. Remove the inner rail from the outer rail. Press the latch inside the rail assembly to allow the inner rail to fully slide out.
2. Attach the inner rail to the Director Node. Align the holes in the rail to the studs on the side of the node. Once all studs are in their holes, slide the rail to the rear to lock it in place.
3. Attach the outer rail to the rack. Align the front and rear of the outer rail with the correct holes in the rack posts. The rear connector fits over the outside of the rack post.
4. Push front and rear rack studs into the rack post holes until they click in place.
5. Repeat the steps above for the other rail.

Step 3: Install the Enclosure in the Rails

1. Make sure the ball bearing races located in the middle rails are all the way forward.
2. Pull the middle rails all the way out until they lock.
3. Align the rear of the node with the front of the middle rails and push the inner rails into the middle rails.
4. Slide the release latch and push the node into the rack.
5. Secure the node in the rack by inserting the Phillips-head screws on the left and right sides of the front of the node. Tighten until snug.

Step 4: Attach Power Cables

Each Director Node is equipped with two power supply units (PSUs) to provide redundant power to the system. These PSUs are located on the left of the rear of the node.

Plug one end of each provided power cable into a PSU and the other end into an AC outlet. VDURA recommends that each PSU be plugged into a different circuit or a UPS to protect against power failures.

Step 5: Attach Network Cables

1. Connect network cables from the high-speed ports to your uplink switches.
2. Connect network cables from the 25 GbE corporate network ports to your uplink switches.

Step 6: Install the Front Bezel

If you removed the bezel, you can reinstall it after finishing the steps above.

1. There are two pins and two catches on the front of the Director Node. Match these to their corresponding locations on the rear of the bezel.
2. Push the bezel until it clicks in place.

2.4 Connect to the Console

Make the following connections to the Director Node to connect to the console.

- 25 Gb Ethernet ports: Connect cc0 and cc1 to infrastructure switches.
- VGA port: Optional port to connect a monitor to manage the node locally.
- USB ports: Optional ports to connect a keyboard and mouse to manage the node locally.
- IPMI port: Optional RJ45 connection to connect to your Ethernet switch for remote IPMI access.
- 1 Gb Ethernet ports: VDURA use only.

Prior to configuring Intelligent Platform Management Interface (IPMI), connect to the Director Node's console via KVM or the admin port.

Connect via KVM

Follow the steps below to connect to the console via KVM.

1. Attach a monitor via the VGA port and a keyboard and mouse via the USB ports.
2. Once connected, refer to Configure Director Node IPMI Support for information on configuring IPMI.

Connect via Admin Port

Follow the steps below to connect to the console via the admin port.

1. Attach an Ethernet cable from a laptop to the 1 Gb admin port 1.
2. Configure that NIC with any IP address in 169.254.0.0/16 except 169.254.12.34.
3. Open your SSH client and connect to the Director Node:

```
ssh admin@169.254.12.34
```

4. Once connected, refer to Configure Director Node IPMI Support for information on configuring IPMI.

2.5 Configure Director Node IPMI Support

IPMI is not required for proper realm functionality. To configure IPMI on Director Nodes, see IPMI Configuration. That chapter also covers the IPMI password rules.

3. Software Configuration

This chapter describes the steps required to configure your VDURA system software once you have installed and cabled the hardware.

This chapter contains the following sections:

- Initial Network Setup
- Out of Box Experience
- Configure PanFS
- DirectFlow Client Installation

3.1 Initial Network Setup

You will likely see the steps below first. If the prompt in step 5 does not appear, proceed to the Out of Box Experience section.

1. Power on the Directors, starting with the one at the bottom. Allow about a minute, then power on the next one above it. Repeat this process until all Directors are powered on. The system will have at least three Director Nodes.
2. Physically connect to the console.
    - Connect via KVM: Attach a monitor via the VGA port and a USB keyboard to the Director Node.
    - Connect via admin port: The admin port is the 1 GbE port furthest to the left. It can be used for installation, but should only be used by a VDURA employee or at a VDURA employee's direction once the realm is running. Attach an Ethernet cable from a laptop to the 1 GbE admin port 1. Configure that NIC with any IP address in 169.254.0.0/16 except 169.254.12.34.
3. Open your SSH client and connect to the Director Node.

```
ssh admin@169.254.12.34
```

4. The system prompts you for a login and password. Log in using the default username and password.

```
login: admin
Password: admin
```

5. The system prompts you to set up the high-speed data network. At the command prompt, enter Yes or press Enter to accept the default value in brackets. For example:

```
pancli -- Initial network setup
Would you like to setup the network now? (recommended) [yes]
```

6. The system prompts you for the IP address, netmask, and default router address.

```
IP Address:  10.1.32.3
Netmask: [255.255.255.0]  255.255.255.0
Default Router Address: [10.1.32.1]  10.1.32.1
```

- IP Address is the IP address of the high-speed data ports of the first Director. The IP address should be the first in the range of IP addresses that you assign to the storage system. VDURA recommends that you assign the VDURA storage cluster its own IP subnet with at least as many available addresses as the maximum predicted number of nodes in the storage system, plus one additional IP address. You also use this IP address to access the Install Wizard.
- Netmask is the subnet mask for the network to which the VDURA system is attached.
- Default Router Address is the name or IP address of the primary routing gateway. The router must be in the same subnet as the VDURA system IP address. If the system does not need routing to reach its clients, leave this field blank.

To configure IP addresses for high-speed data ports, complete steps 5 and 6. These steps do not change the management port IP address.

7. The system displays your settings and prompts you to confirm the values. To confirm the settings, enter Yes or press Enter.

```
Settings chosen:
IP: 10.1.32.3
Netmask: 255.255.255.0
Router: 10.1.32.1
Save settings? [yes]
```

8. Log out of the system.

```
exit
```

At this point, you can unplug the diagnostics port and terminal. However, VDURA recommends that you keep the terminal connected to the serial port while the system comes up for two reasons. First, the next step of the configuration, the Out of Box Experience (OOBE), can be run through the diagnostics/admin port or by connecting to the first Director using the address set during initial network setup. Second, even if you are using the web-based Install Wizard, you can monitor the progress of the system configuration.

After you have configured the first Director, proceed to the next section and use the Install Wizard/OOBE to perform the initial system configuration.

3.2 Out of Box Experience

This section describes the initial Out of Box Experience (OOBE) used to initially power up and configure a VDURA realm.

After the Directors and Storage Nodes are powered on (steps 1 and 2 below) and the network setup is correct, they should establish communication automatically and can be listed with the sysmap nodes command.

- The Node column contains the node identifier.
- The Type column contains the node type: Director, VDURA Data Platform, or VPOD.

> **Note:** VDURA recommends that you run the sysmap nodes command before proceeding to ensure all nodes are ready. You may see clock sync error; while this prevents nodes from joining realm, it will go away after a few minutes. Rerun sysmap nodes to verify.

Below is example output of the sysmap nodes command:

```
[pancli] sysmap nodes
Node Type Serial IP Address State
VCH-1,1 Director 52f200003f4005001 10.101.1.15 Online
VCH-2,1 Director 532200003f4005001 10.101.1.64 Online
VCH-3,1 Director 52f200013f4005001 10.101.1.63 Online
VCH-4,1 VDURA Data Platform 52f200041e4029001 10.101.1.62 Online
VCH-4,2 VPOD 9ec738191d4029001 10.101.1.50 Online
VCH-4,3 VPOD 101e0dc41d4029001 10.101.1.51 Online
VCH-4,4 VPOD b1371a851d4029001 10.101.1.58 Online
VCH-4,5 VPOD bddeb9c11d4029001 10.101.1.49 Online
VCH-4,6 VPOD fc4a659b1d4029001 10.101.1.48 Online
VCH-4,7 VPOD 9400b4811d4029001 10.101.1.45 Online
VCH-4,8 VPOD 68ded6c71d4029002 10.101.1.30 Online
VCH-4,9 VPOD ab784d4a1d4029002 10.101.1.41 Online
VCH-4,10 VPOD 073406511d4029002 10.101.1.34 Online
VCH-4,11 VPOD 0d968c321d4029002 10.101.1.27 Online
VCH-4,12 VPOD 565d71601d4029002 10.101.1.28 Online
VCH-4,13 VPOD e4ed05251d4029002 10.101.1.29 Online
VCH-5,1 VDURA Data Platform 52f200021e4029001 10.101.1.60 Online
VCH-5,2 VPOD d55360781d4029001 10.101.1.42 Online
VCH-5,3 VPOD 90ea9f401d4029001 10.101.1.59 Online
VCH-5,4 VPOD 4bc7cfc71d4029001 10.101.1.46 Online
VCH-5,5 VPOD f6834dae1d4029001 10.101.1.61 Online
VCH-5,6 VPOD 6c058c231d4029001 10.101.1.54 Online
VCH-5,7 VPOD 0cc06cc71d4029001 10.101.1.53 Online
VCH-5,8 VPOD 35f41cec1d4029002 10.101.1.40 Online
VCH-5,9 VPOD 024d00771d4029002 10.101.1.33 Online
VCH-5,10 VPOD 1a5f8c5c1d4029002 10.101.1.32 Online
VCH-5,11 VPOD 991657571d4029002 10.101.1.37 Online
VCH-5,12 VPOD 39e6675a1d4029002 10.101.1.24 Online
VCH-5,13 VPOD 0a433a8e1d4029002 10.101.1.35 Online
VCH-6,1 VDURA Data Platform 52f200031e4029001 10.101.1.65 Online
VCH-6,2 VPOD 69c262fd1d4029001 10.101.1.56 Online
VCH-6,3 VPOD cfdb1fe01d4029001 10.101.1.57 Online
VCH-6,4 VPOD 0da476ff1d4029001 10.101.1.52 Online
VCH-6,5 VPOD 9e68653d1d4029001 10.101.1.55 Online
VCH-6,6 VPOD ec4f4ce21d4029001 10.101.1.44 Online
```

```
VCH-6,7 VPOD 1b5bdc331d4029001 10.101.1.47 Online
VCH-6,8 VPOD b359f4351d4029002 10.101.1.26 Online
VCH-6,9 VPOD cb3793781d4029002 10.101.1.43 Online
VCH-6,10 VPOD fdbedcba1d4029002 10.101.1.38 Online
VCH-6,11 VPOD 40f4702b1d4029002 10.101.1.39 Online
VCH-6,12 VPOD dab33c851d4029002 10.101.1.36 Online
VCH-6,13 VPOD 469f9d4f1d4029002 10.101.1.31 Online
```

The OOBE will automatically create a StorageSet named Set 1 containing / and /home directories, but will assign no VPODs to the StorageSet. You can follow the recommendation to add VPODs to the Set 1 StorageSet in step 5 and either keep the name or change it later. You can also create a new StorageSet and add VPODs to it.

Follow the steps below to complete the OOBE.

1. The Directors should already be powered on from previous steps. If not, power on the Directors, starting with the one at the bottom. Allow about a minute, then power on the next one above it. Repeat this process until all Directors are powered on. The system will have at least three Director Nodes.
2. If the realm includes JBODs, first power on all JBODs and wait at least 60 seconds. Then power on the Storage Nodes one enclosure at a time, starting with the enclosure at the bottom. Allow about a minute, then move to the next higher one.
3. Activate all Directors to bring them into the realm:

```
syscontrol activate all
```

4. Use the syscontrol realmmgr add command to include the second and third Directors in the replication set (repset). You can use the sysmap nodes command to find the identifier for each Director:

```
syscontrol realmmgr add VCH-2,1
syscontrol realmmgr add VCH-3,1
```

The identifiers may be different than the example above.

5. Use the storageset add storage command to add VPODs to the Set 1 StorageSet created during the OOBE process. You can use the sysmap nodes command to find the identifier for each Storage Node, which will begin with VCH. Do not use the identifiers named VDURA Data Platform.

```
storageset addstorage -addspare 0 "Set 1" VCH-4,1
```

3.3 Configure PanFS

This section describes PanFS and how to configure it as part of the installation process.

3.3.1 PanFS Components

PanFS is composed of the following elements:

- PanFS OS
- DirectFlow
- PanFS Realm Management
- PanActive Manager

PanFS Operating System

The PanFS storage operating system (PanFS OS) is the integrated software that manages and directs VDURA systems. It dynamically distributes file activity across Storage Nodes. It also clusters Directors to coordinate file activity, balance system performance, and manage failovers. This distributed cluster-based approach eliminates performance bottlenecks and eases the administration burden for the VDURA storage network.

PanFS is an object-based, parallel file system. Files are divided into chunks that are stored in different objects and accessed in parallel for performance and redundancy. Data objects are containers for data and attributes that describe how the data fits into the overall distributed file system.

Files are striped across objects, and each file has its own map that defines the striping pattern. This allows different objects to have different striping patterns and provides different performance and redundancy characteristics. Because maps are just a small array of object identifiers and associated access capabilities, they are efficient to manage.

For example, small files and directory objects are mirrored so that they are cheap to create and cheap to update with small I/Os often associated with small files and directories. Large files are striped more widely using RAID on a per-file basis to provide redundancy with less capacity overhead than mirroring. Very large files use a two-level striping pattern that optimizes performance in very large storage systems.

The storage map for a file can change dynamically to accommodate changes in the available storage resources. For example:

- If a Storage Node fails, component objects will automatically be reconstructed on alternate VPODs.

- The addition of new storage resources will result in automatic rebalancing of capacity so that all VPODs are loaded uniformly.
- Older hardware can be retired by migrating data away from those VPODs.

All data management is automatic and performed in the background while the system is online.

DirectFlow Client

Similar to massively parallel supercomputer architectures, the DirectFlow Client software supports fully parallel reads and writes from DirectFlow Clients to Storage Nodes without encountering bottlenecks. The DirectFlow software module is a loadable kernel module that transparently adds support for the VDURA file system protocols to Linux compute nodes.

DirectFlow provides high-performance, high-fidelity POSIX semantics for large-scale distributed systems based on a cache consistency protocol with VDURA metadata managers. In addition, there are specialized interfaces that optimize the large shared file workloads found in MPI-IO applications.

PanFS Realm Management

PanFS realm management software logically clusters Directors in an active-active configuration so that all Directors are fully engaged in workload processing.

To enable scalability, realm management allows hundreds of VDURA Directors to cluster together to evenly and effectively distribute the workload. Realm management provides the following advantages:

- Optimizes the initial distribution and ongoing management of files across all hardware elements while minimizing interdependence.
- Adapts to changing client usage patterns and changes in file sizes by dynamically improving system operation.
- Automatically adapts to Director failures by providing a highly available file system.

PanActive Manager

PanActive Manager provides an HTML user interface to manage your VDURA system. This interface makes it easy to securely configure VDURA systems and to display and monitor critical system information. An SNMP interface allows integration of the storage system into enterprise data center management frameworks.

To ensure that problems are detected before they cause a failure, PanActive Manager notifies administrators of impending failures, allowing them to identify and prevent problems before they affect the system. Systems can be monitored for disk, DIMM, and fan failures, as well as temperature and other environmental threats.

For detailed information about the VDURA system and the VDURA PanFS parallel file system, see the *VDURA Data Platform User Guide*.

3.3.2 Configure PanActive Manager

Once VDURA hardware has been installed and cabled, the next step is to configure PanActive Manager to configure and manage storage. For complete information on configuring and managing PanActive Manager, refer to the VDURA Data Platform User Guide.

Prerequisites

The PanActive Manager Install Wizard will ask you for configuration information about your system. Before you begin the configuration process, make sure you have the following information at hand:

- Customer ID provided in your order confirmation. If you do not have your customer ID, call VDURA Technical Support.
- Administrator ID, password, and email address.
- Web proxy host and port, if required. The web proxy server must be configured to handle HTTPS requests if you want PanActive Link to work with your system.
- Simple Mail Transfer Protocol (SMTP) server name or IP address.
- VDURA storage cluster name and IP address.

It is recommended (but not mandatory) that the system name be the DNS name that corresponds to the IP address of the first Director. The system name appears under the /panfs mount point of the DirectFlow file system clients.

The system IP address is the address of the first Director. This address is also used to run the Install Wizard. VDURA recommends that it be the first address of the node address range that you assign to the system.

- Netmask and default router.
- Maximum Ethernet frame size, either 1500 or 9000.
- Node address range: A range of IP addresses that you assign to the VDURA storage cluster. VDURA recommends that you assign the VDURA storage cluster its own IP subnet with at least as many available addresses as the maximum predicted number of nodes in the storage system, plus one additional IP address.
- Domain Name System (DNS) domain name and IP address. If the DNS zone name is not equivalent to system DNS name plus DNS domain name, then the DNS lookups to the DNS zone name fail with NXDOMAIN.
- Network Information Service (NIS) domain name, server, and hostname resolution status.

- Network Time Protocol (NTP) server names or IP addresses. VDURA recommends setting the cluster NTP to the same NTP used by the DirectFlow Clients. A list of NTP servers is available at tf.nist.gov/tf-cgi/servers.cgi. Specify three NTP servers for redundancy. If a node reports clock drift, remove any server that responds inconsistently, or use one reliable server. Servers that are far away on the network are the most likely cause. If no NTP servers are specified, one Director is automatically chosen to operate as a stratum 0 time source for the system.

> **Note:** When configuring NTP, be aware that NTP server time on a client and the VDURA system must be no more than 10 seconds apart, and that closer is better.

- Time zone for the system.

A printable configuration checklist is available in Installation and Configuration Checklist of this guide. A PDF version is also available to download on the first page of the Install Wizard.

Configure the VDURA System

1. Open a web browser.
2. Enter the IP address you assigned to the first Director and press Enter:

```
https://10.1.32.3
```

The first page of the Install Wizard appears.

3. Follow the step-by-step instructions to complete the Install Wizard. To display online help for the Install Wizard, click the controls next to the fields.
4. After you have completed the Install Wizard, click the **Apply Settings** button to save your settings to the system.

The system will automatically configure the first enclosure into a StorageSet and volume and will export the file system for use via DirectFlow. Console messages will indicate when the file system is ready.

For detailed information about configuring and using PanActive Manager, see the VDURA Data Platform User Guide.

You can also configure your system using the VDURA command-line interface. For detailed information about connecting and using PanCLI, see the VDURA PanCLI Reference Guide.

Install DirectFlow Client

For Linux/UNIX users, you can also install the VDURA DirectFlow client software, which enables direct access to the PanFS operating system from supported hosts. Refer to DirectFlow Client Installation.

For complete information on configuring and managing the DirectFlow client, refer to the VDURA Data Platform User Guide.

3.3.3 Upgrade and Downgrade PanFS

Please contact VDURA Technical Support for current information regarding PanFS release compatibility and upgrade/downgrade paths and caveats.

Access a Release

Releases are available only from the Support Portal. As of 12.0.0, the VDURA Data Platform can no longer download upgrade images itself.

Go to the Downloads tab of support.VDURA.com and download the release to a PanFS directory or a machine on your network that your VDURA system can access. Then, in PanActive Manager, navigate to Advanced Tools > Software Management > Upgrade PanFS and use the Upload PanFS Operating System Release section to load the software into your VDURA system. Once uploaded, the software will appear in the Versions Available for Upgrade table. Review the columns that indicate compatibility with your hardware. If an upgrade is not compatible with existing hardware, the column provides a link to a table summary of software releases that are compatible.

For information on upgrading system software with the command-line interface, see the VDURA PanCLI Reference Guide or PanActive Manager online help. Note that remote access via FTP or SCP requires that the remote system be running an FTP or SSH service.

Install a Release

In the PanActive Manager Advanced Tools > Software Management > Upgrade PanFS > Versions Available for Upgrade table, find the release you uploaded. The release shows a Distribute button. Click this button to distribute the upgrade package to all other nodes in the system.

If the distribution is unsuccessful, an error appears on the Upgrade PanFS page, and a new control appears in the Versions Available for Upgrade table. Use this control to redistribute the software either to all nodes or specific ones identified by IP address. Select the appropriate nodes and click Redistribute. You can also go to the Advanced Tools > Software Management > Distribution Status page to view node status and select nodes for redistribution.

Once the image has been successfully distributed, the Distribute button becomes an Install button. Click Install. The system will be taken offline during the upgrade process. When the installation is successful on all nodes, the system will reboot and all nodes will come back online. If the installation fails, the system will revert to the previously running version of the software.

Note that upgrading PanFS may also require upgrading your DirectFlow Clients. Refer to the sections below for instructions on installing and configuring DirectFlow.

Upgrade Time Frame

Upgrade and downgrade times will vary, but in general, the following holds true:

- Distributing the release takes time, but the system is fully online during this step. Since distribution is just a single-click step, there is also practically no administrative time required.
- Distribution time depends on the number of nodes in the system. VDURA uses an exponential distribution method to hasten this process-one node gets the image and sends it to two other nodes, then each of those three nodes send a copy to two additional nodes, then those nine nodes each then send to two additional nodes, etc.
- When the Install button is clicked, the system does some work to validate packages, etc., before it goes offline and reboots. At this point, clients will experience the system hanging. This step will take about ten minutes to complete before coming back online.

3.4 DirectFlow Client Installation

The DirectFlow Client software enables direct access to VDURA storage from supported Linux hosts. This section describes basic installation and configuration of DirectFlow Clients.

1. Download the appropriate RPM or DEB for the client from the Downloads > DirectFlow Client for Linux page of support.VDURA.com. As of 12.0.0, the VDURA Data Platform can no longer download DirectFlow clients itself.
2. Copy the downloaded file to the client.
3. Unmount the realm.

```
umount /panfs
```

4. For RPM installations with a previous version of DirectFlow installed, uninstall the old panfs module before installing the new one, where `[panfs-*]` is the installed panfs module.

```
rpm -e [panfs-*]
```

The rpm -U command is not supported. Using rpm -e will delete all panfs mount lines from the /etc/fstab file, but will backup the file.

5. Run one of the following commands.

For RPM files, where `[panfs*.rpm]` is the name of the RPM file you downloaded:

```
rpm -hiv [panfs*.rpm]
```

For DEB files, where `[path/to/deb]` is the name and location of the DEB file you downloaded:

```
dpkg -i /[path/to/deb]
```

6. Before configuring the DirectFlow client, make sure the client's system is synchronized with the system (e.g., via NTP).
7. If secure boot is enabled on the client, the kernel prompts you to enroll the public PanFS key to the shim security keyring. Secure boot is provided by a BIOS and kernel supporting UEFI 2.3.1 Errata C or higher specification. When the client boots, the shim bootloader must authenticate the operating system for the OS to load. The kernel then only loads kernel modules that it can authenticate using the shim security database. The Machine Owner Key utility allows the root user to manage the public keys for installed kernel modules.

Once DirectFlow is installed, register the DirectFlow key with the following command:

```
# mokutil --import /etc/pki/panfs/panfs-mok.der
```

You will be prompted to enter the client root user password twice. The command will not run without entering the root password. If you get a prompt to run mokutil and cannot run it, DirectFlow will not load. To avoid this issue, do not run config_panfs until you have the client root password.

Once the command is run, the output will indicate that the key was successfully enrolled.

Subsequent DirectFlow installations on this client should not require a public key enrollment, as the public key for any DirectFlow port is the same.

8. Run config_panfs to configure the client.

```
/usr/local/sbin/config_panfs
```

The client must be set to the same NTP as the cluster. Consult the VDURA Data Platform User Guide for any errors encountered while using the script.

9. You will be prompted for the DNS name or IP address of the VDURA system. Typically, you should provide the DNS name unless your site's DNS server is not configured to resolve the system name. If you have configured your VDURA system for realm management with multiple Directors and decide to provide the IP address, enter the IP address of the president Director.

The following is a suggested starting point for a manual configuration of a four-node system:

```
panfs://xxx.xxx.xxx.aaa:global /panfs panfs rw,auto,_netdev,max-async-writepages=4,iscsi-sock-rsize=262144,iscsi-sock-wsize=262144,file-dirty-data-hwm=264241152,sam-resource-level=4096, num-stripes-per-write=112, rmlist=(xxx.xxx.xxx.aaa; xxx.xxx.xxx.bbb; xxx.xxx.xxx.ccc),callback-network-allow= xxx.xxx.xxx.0/24 0 0
```

Regarding DNS/Client Time Slew

PanFS will fail to mount if the time slew between the client and the realm exceeds 10 seconds. Example error message:

```
client:~ # /usr/local/sbin/config_panfs -a -n 192.168.100.10
Enter the DNS/IP address of the PanFS Realm: 192.168.100.10
Started chrony.
PanFS module loaded successfully.
mount.panfs error: mount request rejected because client-realm clock skew exceeds 10 seconds
Error: mount error 1 encountered while running 'mount /panfs'. See mount man page for details.
```

> **Important:** Be aware that there is currently an error in the error message text above. The script error reads "time skew exceeds 10 seconds", when it should read "time slew exceeds 10 seconds".

Corrective Action

No corrective action is required unless clients are unable to mount DirectFlow exports. If clients are unable to mount DirectFlow exports, complete the following steps:

1. Make sure the client and the cluster both use the same NTP server.
2. Force a time sync on the client.
3. Make sure the cluster can ping the NTP server.

4. If the mount still fails, contact VDURA Technical Support. Support may advise rebooting the cluster; be aware that hosts will lose DirectFlow connectivity during the reboot period and may need to restart jobs. After reboot, the storage cluster will synchronize with the reference clock via ntpdate. Ntpdate is a PanFS utility for Directors only.

> **Note:** The ntpdate utility cannot be run while the NTP daemon is running.

For detailed information about installing the DirectFlow client software, see the VDURA Data Platform User Guide.

4. Configure InfiniBand with Subnet Manager

This chapter briefly describes how to use a Subnet Manager tool when configuring InfiniBand installations.

4.1 Subnet Manager Requirements

For InfiniBand networks, the Subnet Manager is a centralized agent responsible for defining and managing an InfiniBand fabric. The Subnet Manager discovers all InfiniBand fabric devices and configures them per the specified subnet configuration. A full overview of the functions of a Subnet Manager are beyond the scope of this document, however, NVIDIA provides their own high-level overview in NVIDIA's Subnet Manager Overview document located at https://docs.nvidia.com/networking/display/mlnxosv3112006/subnet+manager.

The centralized nature of the Subnet Manager mostly avoids the need to individually configure the nodes of the InfiniBand fabric. This does mean that InfiniBand nodes will not appear online until they have been discovered and configured by a valid Subnet Manager.

PanFS places few requirements on the InfiniBand subnet, however, the requirements it does place are:

- All realm nodes must be within same partition.
- The partition used for the realm must enable support for IPoIB along with the corresponding IPoIB broadcast group.
- The Subnet Manager must enable support for Virtualization (virt_enabled 2 for OpenSM). This is because the VPODs are attached to the fabric via SR-IOV virtual functions.

Most InfiniBand switches provide an embedded Subnet Manager. However, in many cases these embedded tools do not support the ability to configure virtual endpoints. In such circumstances, an additional secondary instance of OpenSM may be installed to a node on the InfiniBand subnet to provision the virtual endpoints. Due to OpenSM limitations, it is not possible for a VDURA Data Platform to act as the Subnet Manager.

4.2 OpenSM

OpenSM is a fully-featured Subnet Manager that can be installed on most Linux InfiniBand nodes. It is provided with the NVIDIA DOCA OFED software but is also available in the standard packages offered by many Linux distributions. The full documentation for OpenSM can be found at

https://docs.nvidia.com/networking/display/mlnxofedv24102180lts/nvidia+sm. For the purposes of this installation the procedure for installing and configuring a minimal OpenSM instance to be able to run a realm is covered below.

1. Install the latest NVIDIA DOCA/OFED suite for your Linux distribution using the instructions at https://docs.nvidia.com/doca/sdk/doca+installation+guide+for+linux/index.html.

Alternatively use your Linux distribution's own copy of OpenSM (e.g. yum install opensm, zypper install opensm, or apt install opensm).

2. Create a new partition file in /etc/opensm/partitions.conf:

```
mkdir -p /etc/opensm
touch /etc/opensm/partitions.conf
```

3. Using an editor of your choice, add the following content to /etc/opensm/partitions.conf:

```
Default=0x7fff, ipoib, rate=3, mtu=5, scope=2, defmember=full:
ALL, ALL_SWITCHES=full;
```

4. Enable and start the opensm service:

```
systemctl start opensmd
systemctl enable opensmd
```

5. IPMI Configuration

While IPMI is not required for proper realm functionality, you may wish to use this feature for remote power control, remote sensor monitoring, or the serial-over-LAN console. Users may also consider setting up the BMC network before setting up PanFS.

To use IPMI with your VDURA system, you must first configure it using the ipmi_setup configuration script. Ipmi_setup is a script inside of PanFS and can only run after the realm is configured and online. After connecting a USB keyboard and VGA monitor to the node, use the /pan/bin/ipmi_setup script to configure IPMI.

The ipmi_setup script is used to configure baseboard management controllers (BMCs) for remote IPMI access. This script can either be run manually for each BMC you wish to configure or in batch mode to configure multiple units at once. For all valid commands, the BMC is reset to factory defaults before the desired configuration is applied.

> **Note:** A node may go into Non-reconciled state while being configured.

> **Note:** You can use the PanCLI ipmitool command to further customize IPMI.

> **Important:** IPMI passwords must meet the following criteria:
>
> - Length of 8 to 20 characters
> - Cannot be reverse of username
> - Includes characters from at least three of these character classes: A-Z, a-z, 0-9, special characters

Modes

You can choose from three configuration modes for each BMC.

- Static IP: Use ipmi_setup followed by the BMC IP address, netmask, and gateway IP address. The script will then prompt for a new BMC password.
- DHCP: Use ipmi_setup dhcp to automatically configure the BMC for DHCP. The script will then prompt for a new BMC password.
- Factory reset: Use ipmi_setup factory to reset the BMC to its factory default configuration, including the BMC password.

Argument Styles

You can run the script in two different ways.

- Command-line: Enter arguments via the command line. If needed, the script asks for a password to assign to the BMC.
- Batch: Enter the command-line argument batch followed by the path to a batch file you have previously created and copied to all nodes you are modifying. This file must contain one configuration line for each node on which the script is to be run. Each line must start with a node-id that designates what node the line applies to. The node-id can be either the node's IP address, data serial number, or hardware serial number of the node. After the node-id, the line contains a list of arguments equivalent to the command-line version. For options that require an IPMI password, the password is included in the configuration file in plain text. Since it contains cleartext passwords, the batch file will be deleted after use.

To configure a static IPMI SOL IP address in batch mode, each line must contain the following:

```
[Node Identifier] [IPMI IP address] [IPMI netmask] [IPMI gateway IP address] [IPMI password]
```

To configure IPMI SOL to obtain an IP address via DHCP in batch mode, each line must contain the following:

```
[Node Identifier] dhcp [IPMI password]
```

To reset the IPMI SOL configuration to the factory default in batch mode, each line must contain the following:

```
[Node Identifier] factory
```

Examples

Static IP, command-line: Configure the BMC to use IP address 1.2.3.4, netmask 255.255.255.0, and gateway (router) 1.2.3.255. The script will also prompt for a password to apply to the BMC.

```
ipmi_setup 1.2.3.4 255.255.255.0 1.2.3.255
```

DHCP, command-line: Configure the BMC to obtain IP address, netmask, and gateway (router) via DHCP. This configuration requires an appropriately-configured DHCP server on the broadcast domain. As a final step, the script will prompt for a password to apply to the BMC.

```
ipmi_setup dhcp
```

Factory reset, command-line: Reset the BMC to its factory default configuration.

```
ipmi_setup factory
```

Batch: Configure three nodes using three different modes:

- Node 1: Configure the BMC for static IP using IP address 10.11.12.13 as the node-id.
- Node 2: Configure the BMC for DHCP using data serial number 4a7a2042a1ae36 as the node-id.
- Node 3: Reset the BMC to factory defaults using hardware serial number 42af0001302001001 as the node-id.

Create a batch file containing one line for each node; this example will use /tmp/ipmi_setup.cfg. Note that any type of node_id may be used, and that the desired BMC password is included for each node if the command requires it.

```
10.11.12.13 1.2.3.4 255.255.255.0 1.2.3.255 password1
4a7a2042a1ae36 dhcp password2
42af0001302001001 factory
```

Use scp to copy /tmp/ipmi_setup.cfg to all three nodes:

```
diag_exec " scp root@[director IP address]:/tmp/ipmi_setup.cfg /tmp" all
```

Note that Director IP address above is the address of the node where you created the config file:

```
diag_exec " scp root@192.168.52.10:/tmp/ipmi_setup.cfg /tmp" all
```

After all nodes have the configuration file, run ipmi_setup on all nodes to be configured. To run the command in the background simultaneously on the nodes, use an ampersand (&):

```
diag_exec -timeout 600 "ipmi_setup batch /tmp/ipmi_setup.cfg &" all
```

When run, ipmi_setup will read the batch file, find the specified configuration for each node, and apply it.

Once finished, you should verify that all nodes have been configured successfully. If any node(s) failed, either re-run ipmi_setup with just the failed node(s) or manually run ipmi_setup on the affected node(s).

Verify the Storage Node BMC Network Settings

From the factory, V5000 Storage Nodes are set with IPMI enabled with DHCP. You can verify with the following commands:

```
[pancli] support servicemode yes
[pancli_service] diag_exec -timeout 0 "ipmitool lan print 1"
```

6. Network Overview and Configuration

This chapter describes network requirements for the VDURA V5000 system, as well as the VDURA networking architecture's logical networks and networking modes. It also describes how to configure the 25 GbE ports on the Director Nodes.

This chapter contains the following sections:

- V5000 Network Requirements
- Logical Networks
- Networking Modes
- Configure the 25 GbE Ports

6.1 V5000 Network Requirements

This section covers the following topics.

6.1.1 Introduction

The V5000 Data Platform has specific requirements for the high-speed data network to achieve full functionality and best performance. Note that the requirements change depending on the network type.

6.1.2 V5000 High-Speed Data Network Connection

Customers must define and provide the network environment (switches, cables, and configuration) and VDURA works within that environment. In the following diagrams the customer provides the elements shaded with gray.

The high-speed data network (referred to as the primary network in VDURA documentation) can be either Ethernet or InfiniBand, but not both concurrently.

The realm requires a range of static IP addresses in a single subnet to be assigned to the nodes on the high-speed data network. The subnet may be non-exclusive, meaning it may be shared with clients and other devices. The IP range assigned to the nodes may be non-contiguous, but grouping the IPs as much as possible is recommended for ease of management. The realm is responsible for IP address assignment to its nodes from the supplied range using a proprietary discovery mechanism.

You will need to accommodate the following requirements when determining the range size:

- 1 IP address per Director Node
- 13 or 17 IP addresses per Storage Node depending on the number of NVMe SSDs

- Optionally, 20 extra IP addresses to cover delays in releasing formerly used addresses and for future expansion.

The initial configuration will assign a fixed IP address to the first Director Node. The remaining addresses from the assigned range will be dynamically assigned by this Director to the other nodes.

The realm's dynamic IP address assignment will not conflict with any DHCP services on the high-speed data network. The V5000 realm will ignore DHCP on the high-speed data network. However, it is expected that the static IP address range assigned to the realm has been removed from the address pool of any DHCP servers occupying the realm subnet.

The realm nodes use the default route to contact devices external to the subnet. Static routes may be configured on the data-path network to reach clients and services on other subnets if using the default gateway is not desired.

A V5000 realm does not have a private or back-end network. The realm connects directly to the high-speed data network used by the clients. Director Nodes also have 25 GbE network connections for additional services.

Keep the following points in mind regarding node port bandwidth:

- Director Nodes focus on IOPS instead of bandwidth and require two 200 Gb or 400 Gb ports per node.
- Storage Nodes focus on IOPS and bandwidth and require four 200 Gb ports or two 400 Gb ports per node for a total of 800 Gb available bandwidth.
- All node types configure the physical network ports for failover.

6.1.3 Other V5000 Network Connections

25 GbE ports

Director Nodes have dual 25 GbE ports, known as secondary interfaces, for the following purposes:

- Realm management using PanCLI and PanActive Manager
- S3 gateway services
- Accessing NTP, DNS, and authentication services not available on the high-speed data network
- Connecting to the corporate network

There are several configuration rules regarding the 25 GbE ports:

- The subnet for the 25 GbE links should be separate from the data-path subnet.
- A pool of IP addresses is statically assigned to the secondary network interfaces. One IP address per Director Node should be provided; the addresses do not need to be contiguous.
- Static routes cannot currently be configured for the secondary 25 GbE network.

- It is optional but recommended that the interfaces operate as a failover bond.

1 GbE ports

All node types have 1 GbE ports for out-of-band management using the on-board Baseboard Management Controller (BMC). Addresses may be assigned statically or dynamically (using standard DHCP) as desired.

In the following diagrams, the Storage blocks show two connections representing failover capability but may have up to four physical connections.

6.1.4 NTP Server

The V5000 realm and the clients must synchronize time with the same NTP server, preferably a local NTP server, to maintain coherency between client transactions with shared files.

- DirectFlow running on a client server will not mount the filesystem if the client server time differs from the realm time by more than 10 seconds.
- Variance between client server time and realm time of more than one second will have a negative performance impact.

Director Nodes in the V5000 realm must be able to reach the NTP server through either the high-speed data network or the optional 25 GbE corporate network ports. Storage Nodes synchronize with the Director Nodes and do not need to reach the NTP server directly.

6.1.5 InfiniBand Data Network

For an InfiniBand high-speed data network, the realm must connect to the same high-speed data network as the clients.

- Ports can be split evenly across two InfiniBand switches for redundancy.
- Supported MTU sizes are 1500, 2044, and 4092.
- The clients and realm nodes should communicate through IPoIB.

Director Nodes have dual 25 GbE ports for connections to the corporate network fabric for additional services.

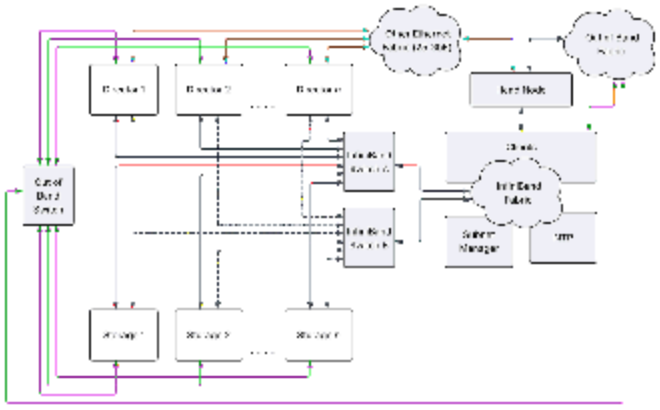
All nodes in the realm have 1 GbE network ports for out-of-band management using a BMC.

Example InfiniBand Data Network

The Head Node in the diagram below is a server that manages the clients and usually has access to all three networks, which are the high-speed data fabric, the Other Ethernet or management fabric, and the Out of Band or IPMI fabric. The Head Node may also replicate services such as NTP and DNS for the

high-speed data fabric.

The diagram below shows the NTP server connected to the InfiniBand fabric, but it could be on a different network available to the Director Nodes, accessed through their 25 GbE ports.



InfiniBand data network.

InfiniBand Subnet Manager Requirements

The InfiniBand Subnet Manager configuration must meet the following requirements:

- Virtualization must be enabled, as Storage Nodes have multiple virtual services sharing the InfiniBand ports using SR-IOV. For OpenSM, this setting is virt_enabled 2.
- All V5000 realm nodes must be in the same partition. The partition must enable support for IPoIB along with the corresponding IPoIB broadcast group.

6.1.6 Ethernet Data Network

For an Ethernet high-speed data network, the realm must connect to the same high-speed data network as the clients.

- The ports can be split evenly across two Ethernet switches for redundancy.
- Supported MTU sizes are 1500 and 9000.

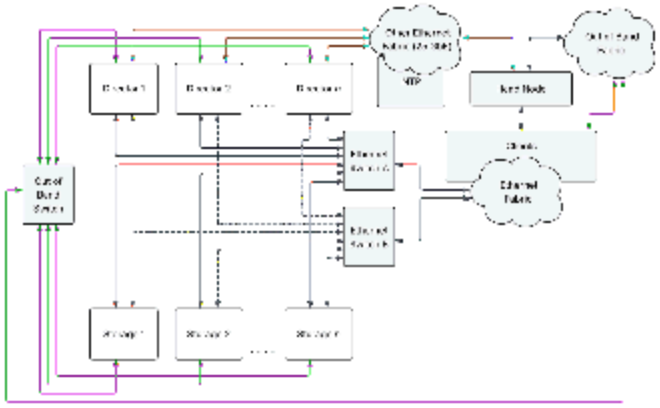
Director Nodes have dual 25 GbE ports for connections to a general network fabric for additional services.

All nodes in the realm have 1 GbE network ports for out-of-band management using a BMC.

Example Ethernet Data Network

The Head Node in the diagram below is a server that manages the clients and usually has access to all three networks, which are the high-speed data fabric, the Other Ethernet or management fabric, and the Out of Band or IPMI fabric. The Head Node may also replicate services such as NTP and DNS for the high-speed data fabric.

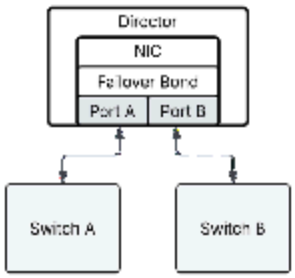
The diagram below shows the NTP server connected to a separate Ethernet fabric available to the Director Nodes accessed through the 25 GbE ports, but it could also be on the high-speed data network available to the Director Nodes accessed through the 25 GbE ports.



Ethernet data network.

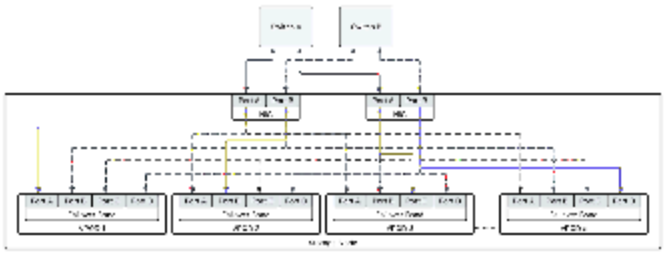
6.1.7 Data Network Port Connections

Director Nodes have two 200 Gb or 400 Gb ports for the data network. These ports are bonded together for active-passive failover. The failover configuration does not require LACP or MLAG; the ports operate as standard connections. The ports can connect to different switches for the best availability.

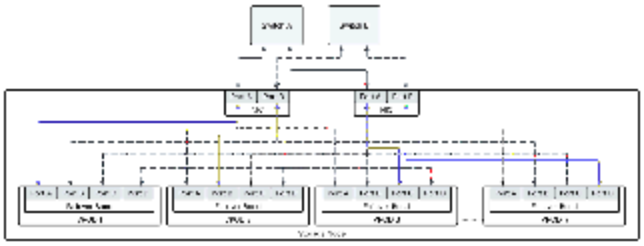


Director Node port connectivity.

Storage Nodes have four 200 Gb or two 400 Gb ports for the data network. Each virtual storage service or VPOD running on the Storage Node bonds the ports together for active-passive failover, but the primary port (shown by blue lines in the diagrams) selected for each VPOD is different to achieve the highest throughput. Ports can connect to different switches for the best availability.



Connectivity with four 200 Gb ports.



Connectivity with two 400 Gb ports.

6.2 Logical Networks

Primary Network

The primary network is on the same physical network or VLAN as the VDURA storage enclosures. The primary data network is accessed through the primary interface (lagg0).

The primary interface may be configured with a DHCP-obtained IP address, a statically configured IP address (for realm managers), and one or more primary gateway IP (GWIP) aliases (for Director Nodes).

Secondary Network

The secondary 25 GbE network is on a separate physical network than the primary network. The secondary network is accessed through the secondary interface (lagg1). For Director Nodes, the secondary interface may be configured with one or more secondary GWIP aliases. Secondary aliases cannot be assigned to a Storage Node, therefore you cannot have a DirectFlow connection over a secondary alias. S3 is supported.

Administrative Network

In addition to the primary and secondary networks, Director Nodes also support a dedicated administrative network on a point-to-point connection to igb0. The administrative network is accessed through a static IP address configured on lagg2.

The administrative network is used by VDURA Service Engineers during initial system configuration and other situations as needed to troubleshoot a Director Node or when a Director Node is unreachable due to misconfiguration. Otherwise, the port should remain disconnected.

Gateway IP Aliases

IP aliases are used to assign a second range of IP addresses to Director Nodes that are distinct from the DHCP and statically assigned ones. IP aliases are used for S3 gateway access to provide high availability to S3 clients, by transparently migrating the GWIP alias from a failing Director Node to a working one.

In situations where S3 is widely used, the use of IP aliases is highly recommended. In this mode, the S3 clients should be configured to connect to an IP alias rather than the physical service address of a Director. This way, if the corresponding Director should fail for some reason, the IP alias will be redirected to another, and the S3 client will see minimal disruption.

Two types of aliases are supported:

- Primary aliases are designed to go on the primary LAG (lagg0) and run on the same network as the rest of the nodes. The DHCP-assigned IP address of each node is often called its service IP address. The primary alias is an IP address that is bound to the same interface as the service IP address. In many cases, the additional range is in the same subnet as the DHCP-assigned addresses. Alternately, if there is a proper gateway, the additional range could be in a different subnet accessed by the same physical network topology.
- Secondary aliases are on a separate physical network than the primary network. With a secondary alias, the lagg1 interface of a Director Node can be given an IP address in the other network.

For the limitation on DirectFlow connections over secondary aliases, see Secondary Network.

Use the following commands to configure primary and secondary aliases:

```
[pancli] netconfig alias add primary [start IP] [end IP] [netmask]
[pancli] netconfig alias add secondary [start IP] [end IP] [netmask]
```

The primary flag indicates that the additional range should share the same physical network as the DHCP-assigned range. The secondary flag indicates that the additional range should share the same physical network as the corporate/IT network.

To validate which Directors have been assigned an IP alias, issue the following command:

```
[pancli] sysmap nodes aliases
```

If you have a Director configured in this manner and it fails, do not physically connect the secondary network until after the replacement has been inserted and is up and running. The new Director will send its startup DHCP request via the best interface it has available; if this request is sent over the external network, it will not get an address.

6.3 Networking Modes

Single Link Mode

Single link mode provides the most basic form of network connectivity. It is appropriate when neither redundancy nor load balancing are required. Select single link mode when a single network uplink or network port is sufficient for your application. The port should be an untagged access port, as VLANs are not supported at this time.

Network Failover Mode

Failover mode provides active/passive networking redundancy. Select failover mode when your application requires active/passive redundancy for high-availability.

The physical ports or uplinks used by the interface must be attached to the same VLAN or physical network in a failover LAG and should reside in the same L2 broadcast domain. The LAG should be an untagged access port, as VLANs are not supported at this time. Also, the upstream switch should not be configured to use link aggregation.

Physical ports, also referred to as links, join the LAG in priority order. One link is active, while the remaining (standby) links are passive. If the active link fails, the LAG will automatically use the next available link, in priority order. The high-availability redundancy benefit is maximized when the physical ports or network uplinks terminate on physically separate peer infrastructure switches.

Active/Active Link Aggregation Mode

In active/active link aggregation mode, the physical ports are bonded together into a single logical port called the link aggregation group (LAG). This provides load balancing for increased throughput in the fault-free case, while also providing fault tolerance in the event a physical port fails to ensure that the bonded interface still operates.

The physical ports or uplinks used by the interface are attached to the same VLAN or physical network in an active/active LAG and should reside in the same L2 broadcast domain. The LAG should be an untagged access port, since VLANs are not currently supported.

There are two types of active/active link aggregation (also referred to as port bonding):

- static link aggregation (Static LAG)
- dynamic link aggregation with Link Aggregation Control Protocol (LACP LAG)

Secondary Network Mode

For a description of the secondary network, secondary GWIP aliases, and the supported protocols, see Secondary Network. To configure the 25 GbE ports, see Configure the 25 GbE Ports.

6.4 Configure the 25 GbE Ports

Director Nodes have dual 25 GbE ports, known as secondary interfaces. You can configure them using PanCLI or PanActive Manager. Both methods do the same two things. They enable the secondary network on each Director Node and define the range of secondary IP addresses that the 25 GbE ports use.

For a description of the secondary network and the configuration rules for the 25 GbE ports, see Other V5000 Network Connections and Secondary Network.

6.4.1 Configure the 25 GbE Ports Using PanCLI

Before following these instructions, ensure that the 25 GbE ports detect a carrier.

Enable the Secondary Network

The 25 GbE ports are the secondary network and must be enabled. The following example enables the secondary network for a Director Node named VCH-1,1. The two ports of the 25 GbE NIC are set to failover mode, so that both ports will use a single IP address, and only one port will operate at a time.

```
netconfig nic node VCH-1,1 secondary link-count 1 mode failover
```

Running the command will generate a series of messages and requires a “yes” answer at the prompt.

```
You are setting the secondary networking mode for node VCH-1,1
to failover. The secondary link-count setting is 1.
For Bravo ASD-100 nodes valid (pri+sec) link-count combinations are:
   4x10GE: (1+0), (2+0), (2+2), (4+0)
   2x40GE: (1+0), (2+0), (1+1)
For Delta ASD-200 nodes valid (pri+sec) link-count combinations are:
   2x25GE: (1+0), (2+0), (1+1)
For ASU nodes valid (pri+sec) link-count combinations are:
   2x10GE: (1+0), (2+0)
   2x25GE: (1+0), (2+0)
The following sequence of actions must be taken to successfully configure
the secondary networking mode.
For target nodes:
 1. Verify that the configuration settings you entered are all correct,
    that the required interfaces and uplinks are all functional,
    and that there are no link-count conflicts.
 2. If you are moving or removing secondary links, take them out of
    service, and disconnect them.
 3. Submit this configuration change.
 4. Configure the appropriate settings at the corresponding external
    switch, as applicable.
 5. Connect each secondary link to its external switch, as applicable.
 6. Disable NIC warnings for links that have been removed
    from service.
 7. Enable NIC warnings for links that have been returned
    to service.

Type "yes" to continue or "no" to quit (yes/no): yes
Successfully configured the Secondary interface networking configuration for the node. Allow a
few minutes for the change to take effect.
```

The next example command verifies that the secondary network was enabled as expected. The output should show “Failover + Failover” as shown.

```
[pancli] netconfig nic node VCH-1,1
Node Name  Node Type     Serial             Nic Mode
VCH-1,1    VCH-5000-D1N  53f30004304001001  (1+1) Failover + Failover (N/A)
```

Define Secondary Aliases

The aliases define what IP addresses can be set for each network type. The secondary alias choices define the 25 GbE port addresses. In the following example, the **netconfig alias add** command adds an alias for the secondary port with the following configuration:

- Starting IP address of the range: 192.168.1.102
- Ending IP address of the range: 192.168.1.103
- Netmask of 255.255.255.0

```
netconfig alias add secondary 192.168.1.102 192.168.1.103 255.255.255.0
```

Define the Secondary Router

If the secondary network sends traffic through a gateway, the following command defines that gateway as a router. In this example, the router is 192.168.1.1.

```
netconfig routers add 192.168.1.1
```

Troubleshooting

If the commands fail, perform the following steps.

1. Confirm that the 25 GbE ports have an active carrier.

```
[pancli] support s yes
Service debug commands are enabled.

 Warning: Service commands should only be used under the
 direction of company service personnel to ensure safe usage.

[pancli_service] ifconfig lagg1
FreeBSD
lagg1: flags=1008843<UP,BROADCAST,RUNNING,SIMPLEX,MULTICAST,LOWER_UP> metric 0 mtu 1500

options=4e507bb<RXCSUM,TXCSUM,VLAN_MTU,VLAN_HWTAGGING,JUMBO_MTU,VLAN_HWCSUM,TSO4,TSO6,LRO,V
LAN_HWFILTER,VLAN_HWTSO,RXCSUM_IPV6,TXCSUM_IPV6,HWSTATS,MEXTPG>
 ether 14:23:f3:2f:9d:10
 hwaddr 00:00:00:00:00:00
 inet 192.168.1.103 netmask 0xffffff00 broadcast 192.168.1.255
 inet 192.168.1.102 netmask 0xffffffff broadcast 192.168.1.102
 laggproto failover lagghash l2,l3,l4
 laggport: bnxt0 flags=1<MASTER>
 laggport: bnxt1 flags=4<ACTIVE>
 groups: lagg
 media: Ethernet autoselect
 status: active
 nd6 options=29<PERFORMNUD,IFDISABLED,AUTO_LINKLOCAL>
```

2. The **status** value near the end of the output should say “active”. If it does not, verify that the cables are plugged in and the network is operational before proceeding.
3. After confirming that the ports have an active carrier, reboot the Director with the PanCLI command **diag_exec {reboot}**.

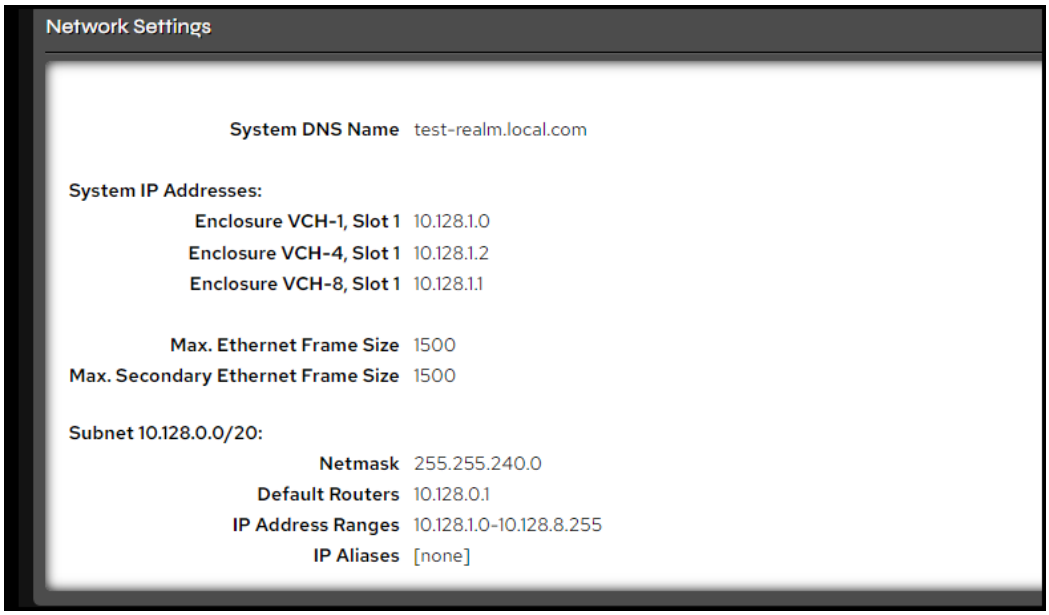
6.4.2 Configure the 25 GbE Ports Using PanActive Manager

This topic covers Enable the Secondary Network and Define Secondary Aliases.

Enable the Secondary Network

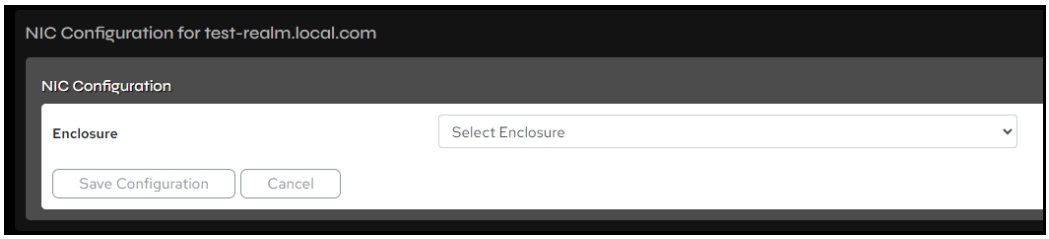
After completing the out-of-the-box experience (OOBE), enable the secondary network as follows:

1. Log in to PanActive Manager and click **Network**. The Network Configuration summary page appears.



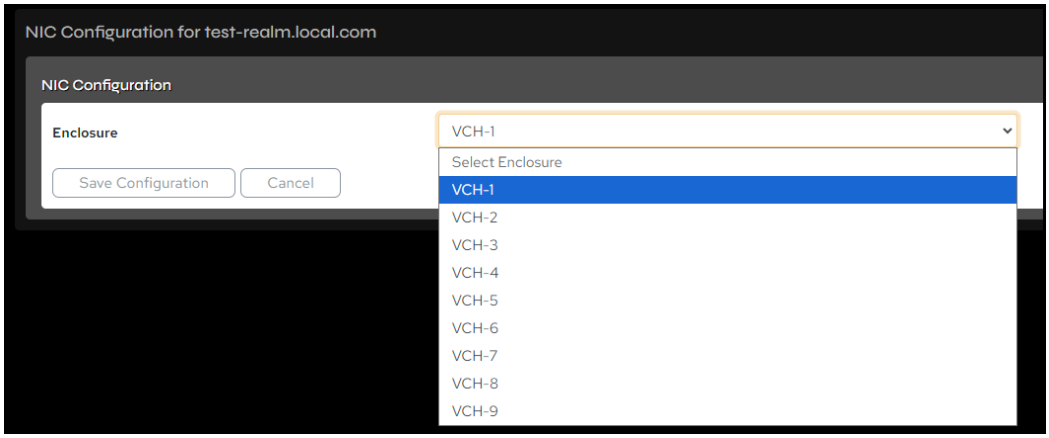
Network Settings on the Network Configuration summary page.

2. Click **NIC Configuration** in the left-hand menu.



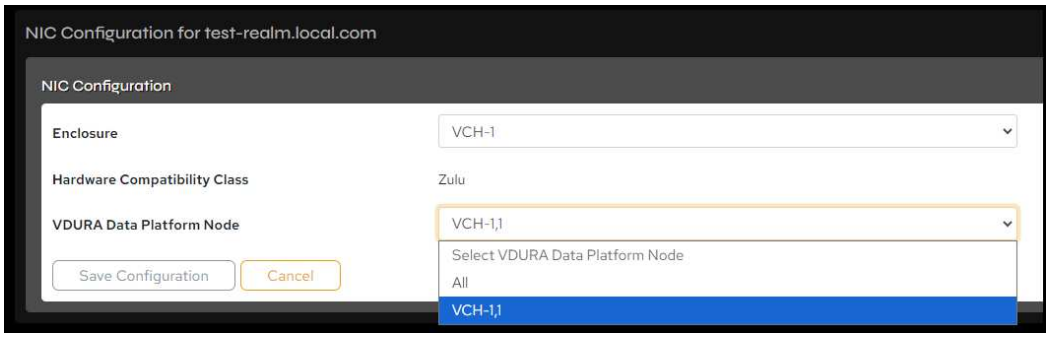
NIC Configuration page.

3. In the **Enclosure** drop-down box, select the name of the Director enclosure to configure. In this example, the Director enclosure is named VCH-1.



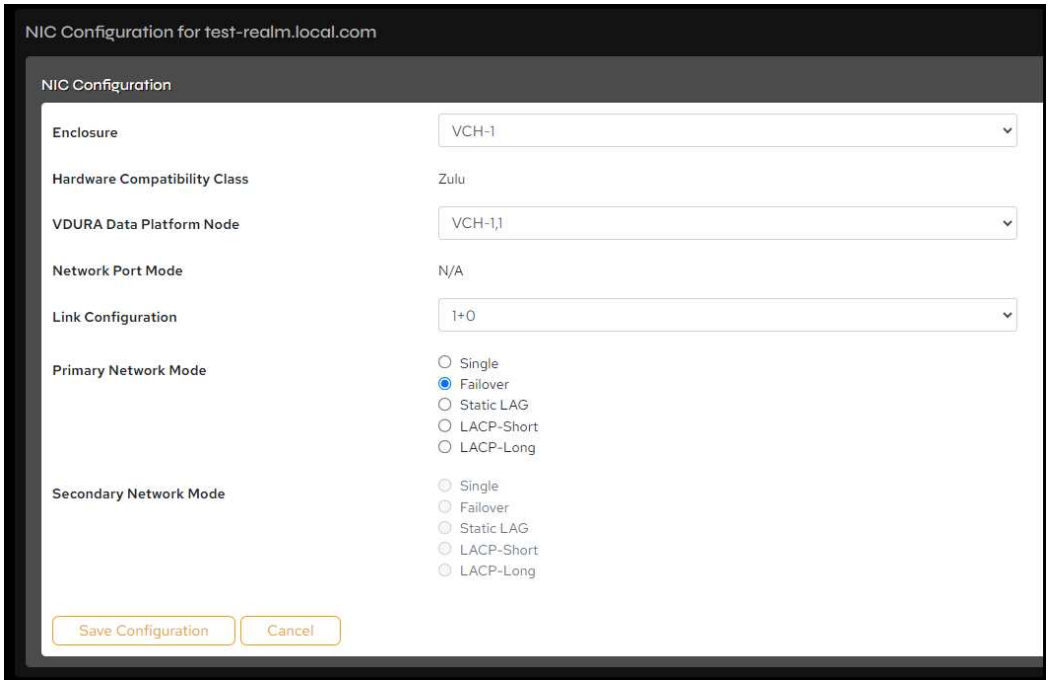
Selecting the enclosure.

4. In the **VDURA Data Platform Node** drop-down box, select the name of the Director Node to configure. In this example, the name is VCH-1,1. To apply the same settings to every Director, select **All**.



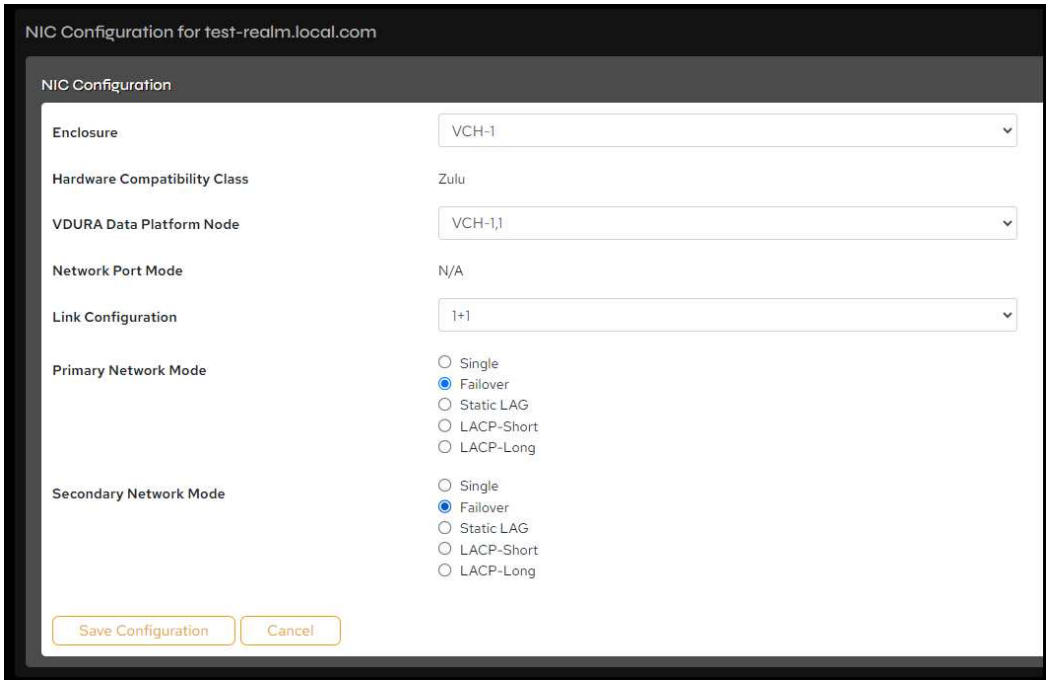
Selecting the node.

5. The **Secondary Network Mode** options control the 25 GbE ports but are unavailable while the **Link Configuration** drop-down box shows the 1+0 option.



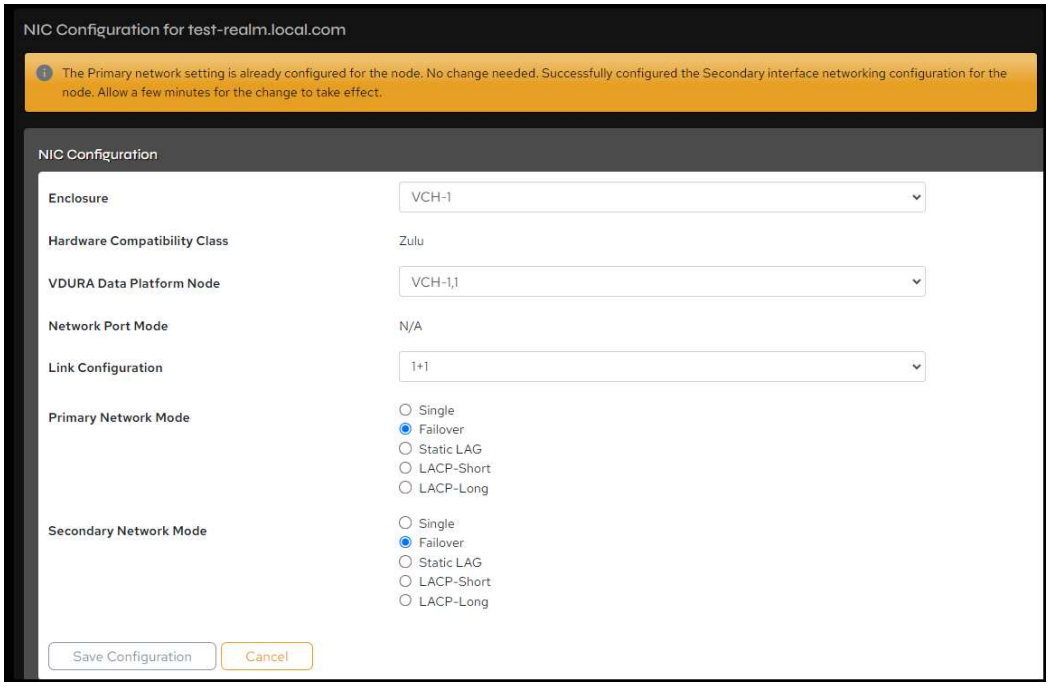
Link Configuration set to 1+0, with the Secondary Network Mode options unavailable.

6. Change **Link Configuration** to 1+1. If it is not already selected, click the **Failover** radio button in the **Secondary Network Mode** options.



Link Configuration set to 1+1 with Secondary Network Mode set to Failover.

7. Click **Save Configuration**. A message confirms that the Secondary interface networking configuration was set and that the change can take a few minutes to take effect.

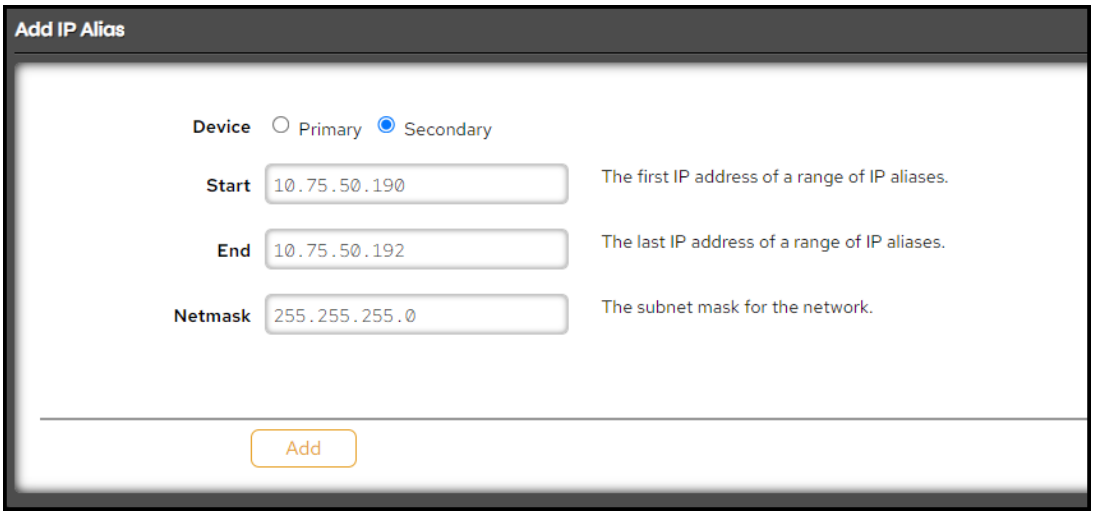


The NIC Configuration page after saving.

Define Secondary Aliases

The aliases define what IP addresses can be set for each network type. The secondary alias choices define the 25 GbE port addresses.

1. In PanActive Manager, click **Network > Network Configuration > IP Aliases**.
2. For **Device**, select **Secondary**.
3. Complete the **Start**, **End**, and **Netmask** fields for the IP range that the 25 GbE ports on the Directors should use. The range cannot overlap the range already assigned to the realm data network addresses. In this example, the range is 10.75.50.190 to 10.75.50.192, and the netmask is 255.255.255.0.
4. Click **Add**.



The Add IP Alias section with a secondary range entered.

7. Product Maintenance Procedures

This chapter describes the various field-replaceable components of your VDURA V5000 system.

This chapter contains the following sections:

- V5000 Director Node Component Replacement
- V5000 Storage Node Component Replacement
- V5000 JBOD Component Replacement

7.1 V5000 Director Node Component Replacement

This section describes how to replace Director Node components and covers the following topics:

- Replace a Director Node
- Replace an Individual Drive in a Node
- Replace a Power Supply Unit

> **Warning:** Use caution when you remove or replace system components; they may be hot to the touch.

7.1.1 Replace a Director Node

Follow the instructions below to replace a single Director Node in an enclosure.

Before You Begin

This section assumes that a replication set (repset) of three or more Directors is configured on the realm. The VDURA realm is governed by the realm president Director. If the realm contains at least three Directors, a three- or five-member repset is typically configured for redundancy. The president Director replicates its decrees and the realm database to the other members of the repset. Other repset member Directors may serve in other roles, such as the Management Console (MConsole) Director or the DHCP server. All repset member Directors also run a DNS service.

Before beginning the steps below, confirm that the Director has actually failed. If necessary, contact VDURA Technical Support for assistance.

In the example below, 10.72.4.12 is the Director that is being replaced. The Director is still online. The output of these commands would be different if 10.72.4.12 were not responding, but the commands would be the same.

Step 1: Remove Target Director from Repset

If the target Director is a member of the repset, it should be removed from the repset before it is replaced.

To determine which Directors are members of the repset, run the syscontrol realmmgr:

```
[pancli] syscontrol realmmgr
NODE ADDRESS ROLE ON/OFFLINE
VCH-1,1 10.72.4.11 RM, MConsole, DNS Online
VCH-2,1 10.72.4.12 RM, DHCP, DNS Online
VCH-3,1 10.72.4.13 RM President, DNS Online
```

If the target Director is not present in this output, move to step 2. Otherwise, continue on below.

> **Note:** If the target Director is the only Director in this output, or if this output shows two repset members, please contact VDURA Technical Support before proceeding. Additional steps will be required to replace the Director. If the target Director is the president, data will typically be unavailable for about a minute while the presidency fails over to another Director in the repset. In rare circumstances, data may be unavailable for longer than that if the new president must restart stuck services. If the target Director is the MConsole Director, the PanActive Manager and PanCLI interfaces will be unresponsive for about a minute while the MConsole service fails over to another repset member Director.

Remove the target Director from the repset with syscontrol realmmgr remove [director IP]:

```
[pancli] syscontrol realmmgr remove 10.72.4.12
/pan/bin/pan_rm_droprm: dropping RM 0x0101c7ef23af0004(RM)
/pan/bin/pan_rm_droprm: dropped RM 0x0101c7ef23af0004(RM).
```

The output may warn that the repset does not have an optimal number of members:

```
Warning: A replication set containing 2 members is not a stable configuration. Please add a
Director to the repset immediately.
```

This is acceptable for a short period of time.

Use syscontrol realmmgr to verify that the target Director is no longer in the repset:

```
[pancli] syscontrol realmmgr
NODE ADDRESS ROLE ON/OFFLINE
VCH-1,1 10.72.4.11 RM, MConsole, DNS Online
VCH-3,1 10.72.4.13 RM President, DNS Online
Warning: A replication set containing 2 members is not a stable configuration. Please add a
Director to the repset immediately.
```

Step 2: Move Volume Metadata Management Services off of Target Director

A Director may be responsible for managing the metadata for one or more volumes. This responsibility is shown by the assignment of a volume service to the Director. Use volservice list to see the volume service assignments.

The following output shows that 10.72.4.12 is the primary manager for volume service ID 0x0400000000000032. This Director is also the backup Director for 0x0400000000000005.

```
[pancli] volservice list
Service Service Primary Primary Backup Backup Volume
ID Status Director IP Director IP Count Volumes
0x0400000000000003(FM) online VCH-1,1 10.72.4.11 VCH-3,1 10.72.4.13 1 /
0x0400000000000004(FM) online VCH-1,1 10.72.4.11 VCH-3,1 10.72.4.13 1 /mdube
0x0400000000000032(FM) online VCH-2,1 10.72.4.12 No backup 1 /esmith
0x0400000000000005(FM) online VCH-3,1 10.72.4.13 VCH-2,1 10.72.4.12 1 /rroy
```

It is usually better to gracefully move the volume services where the target Director is the primary Director to another Director before the target Director is replaced.

To move the volume service, use volservice move [volume service ID] [new primary director IP].

For example, the command below moves the responsibility for volume service 0x0400000000000032 to Director 10.72.4.11. The president Director has also copied the pan_mgr service transaction logs from Director 10.72.4.12 to Director 10.72.4.11. If practical, choose the Director that is the primary Director for

the least number of volume services or whose volume services contain the least active volumes as the new primary Director. If all Directors have about the same volume service load, then you can choose any Director as the new primary Director.

```
[pancli] volservice move 0x0400000000000032 10.72.4.11
Manual move of service id 0x0400000000000032 to 10.72.4.11 completed successfully.
```

If the metadata load on the volume service is high, it may take as long as 10 minutes for this command to complete. The target Director must gracefully close connections with DirectFlow Clients. Do not run the command again if it is slow to complete.

Be aware that if the volume service shows no backup, or if the target Director is not responding or is in distress, moving the volume service may result in the volumes in the service going offline and into mandatory FSRC. Consult with VDURA Technical Support before making the move. For example:

```
[pancli] volservice list
Service Service Primary Primary Backup Backup Volume
ID Status Director IP Director IP Count Volumes
0x0400000000000003(FM) online VCH-1,1 10.72.4.11 VCH-3,1 10.72.4.13 1 /
0x0400000000000004(FM) online VCH-1,1 10.72.4.11 VCH-3,1 10.72.4.13 1 /mdube
0x0400000000000032(FM) online VCH-2,1 10.72.4.12 No backup 1 /esmith
0x0400000000000005(FM) online VCH-3,1 10.72.4.13 VCH-2,1 10.72.4.12 1 /rroy
```

The pan_mgr transaction logs above are not being replicated to a backup Director. If the president Director cannot successfully move the transaction logs from Director 10.72.4.12 to the new primary Director, the volume /esmith will go offline and begin FSRC.

After you have moved the volume service, move the backup assignments with volservice move backup [volume service ID] [new backup director IP]:

```
[pancli] volservice move backup 0x0400000000000005 10.72.4.11
Manual move of backup for service id 0x0400000000000005 to 10.72.4.11 completed successfully.
```

In the example, there is only one choice for the new backup Director. Since 10.72.4.13 is already the primary Director, it cannot be the backup Director.

Verify that the volume service moves completed as expected. In this example, the target Director 10.72.4.12 is no longer the primary Director or backup Director for any volume service:

```
[pancli] volservice list
Service Service Primary Primary Backup Backup Volume
ID Status Director IP Director IP Count Volumes
0x0400000000000003(FM) online VCH-1,1 10.72.4.11 VCH-3,1 10.72.4.13 1 /
0x0400000000000004(FM) online VCH-1,1 10.72.4.11 VCH-3,1 10.72.4.13 1 /mdube
0x0400000000000032(FM) online VCH-1,1 10.72.4.11 VCH-3,1 10.72.4.13 1 /esmith
0x0400000000000005(FM) online VCH-3,1 10.72.4.13 VCH-1,1 10.72.4.11 1 /rroy
```

Step 3: Disable IP Aliases on Director

If IP aliases have been configured for S3 gateway access, see if any have been assigned to this Director using sysmap node [director IP]:

```
[pancli] sysmap node 10.72.4.12
Name: VCH-2,1
State: Online
Type: DB18-48
Services: Enclosure Master
IP Address: 10.72.4.12
Primary Aliases: 10.72.4.55 10.72.4.57
Secondary Aliases: <none>
10GE Secondary Only: disabled
. . .
```

If the Director has either primary or secondary aliases listed, disable them by using syscontrol node [director IP] alias disable:

```
[pancli] syscontrol node 10.72.4.12 alias disable
Successfully disabled alias assignment for node "10.72.4.12".
```

The president Director will move the aliases to other Directors. Use sysmap node [director IP] and sysmap nodes directors aliases to verify that the alias IP addresses have been moved to other Directors:

```
[pancli] sysmap node 10.72.4.12
Name: VCH-2,1
State: Online
Type: DB18-48
Services: Enclosure Master
IP Address: 10.72.4.12
Primary Aliases: disabled
Secondary Aliases: disabled
10GE Secondary Only: disabled
. . .
[pancli] sysmap nodes directors aliases
Node Type Serial IP Address State
VCH-1,1 Director 23a30338300028001 10.72.4.11 Online
Primary Aliases: 10.72.4.57
Secondary Aliases: N/A
VCH-2,1 Director 23a303c5300028001 10.72.4.12 Online
Primary Aliases: N/A
Secondary Aliases: N/A
VCH-3,1 Director 23a3039d300028001 10.72.4.13 Online
Primary Aliases: 10.72.4.55
10.72.4.56
Secondary Aliases: N/A
```

Step 4: Flash Director LED

To physically identify the Director, flash its LED using [pancli] syscontrol node [director IP] led on:

```
[pancli] syscontrol node 10.72.4.12 led on
Now flashing LED for node "10.72.4.12"
```

If this command fails or times out, either skip this step or turn on the LEDs for surrounding nodes to identify the Director to be replaced.

Step 5: Halt the Node

Power off the Director using syscontrol node [director IP] halt:

```
[pancli] syscontrol node 10.72.4.12 halt
```

If the Director is not responding and the command times out, skip this step.

Once the Director becomes non-responsive, the president will move any configured alias IP addresses to a different Director. Clients that connect through the alias IP address should be able to resume data access. Any clients that connect through the Director's non-alias IP address will lose access.

Step 6: Remove Old Director Record from Realm Database

The president automatically removes the record of an OSD that is reconstructed out of a StorageSet and then physically replaced, but does not do this for a Director. Remove the record manually with syscontrol node [director IP] remove:

```
[pancli] syscontrol node 10.72.4.12 remove
```

Step 7: Physically Replace Node

If the Director did not halt in step 5 above, power it off by pressing its front power button. If the Director still does not halt, press and hold its power button for five seconds.

1. Remove the front bezel from the old Director by sliding the catches toward each end and pulling the bezel forward. Consider labeling each network cable and power cable before disconnecting it, to make it easier to attach them to the replacement node.
2. Disconnect all cables from the back of the Director, including the power cables from both power supplies, the network cables, and any console cables. Failure to do this step can result in damaging the NIC or other rear ports.
3. Remove the Phillips-head screws on the left and right sides of the front of the Director that secure it to the rack.
4. Slide the Director forward on its rails until the rails lock in the service position. Support the weight of the node, slide the release latch, and lift the Director out of the middle rails.
5. Remove the inner rails from the old Director by pressing the latch on each rail and sliding the rail toward the front to release it from the studs on the side of the node. Attach the inner rails to the replacement Director by aligning the holes in each rail with the studs on the side of the node, then sliding the rail to the rear to lock it in place.

6. Make sure the ball bearing races in the middle rails are all the way forward and the middle rails are pulled all the way out until they lock. Align the rear of the replacement Director with the front of the middle rails and push the inner rails into the middle rails.
7. Slide the release latch and push the Director into the rack, in the same rack position as the old Director.
8. Secure the Director in the rack by inserting the Phillips-head screws on the left and right sides of the front of the node. Tighten until snug.
9. Reconnect the power cables and the network cables as they were connected to the old Director.
10. Reinstall the front bezel by matching the two pins and two catches on the front of the Director to the corresponding locations on the rear of the bezel, then pushing the bezel until it clicks in place.
11. Power on the Director using the front power button.
12. A replacement node will go through the Not Responding and Clock Sync states before reaching Online.
13. After about 10 minutes, check the status of the Director from a repset member:

```
[pancli] sysmap nodes directors
Node Type Serial IP Address State
VCH-1,1 Director 3c760004301002001 10.72.4.11 Online
VCH-2,1 Director 3c760005301002001 10.72.4.12 Online
VCH-3,1 Director 3c760006301002001 10.72.4.13 Online
```

If the new Director stays in Not Responding or Clock Sync for more than 20 minutes or goes into any other state for more than a few minutes, please contact VDURA Technical Support.

Step 8: Stop LED Flashing

Turn off the identifying LED using syscontrol node [director IP] led off:

```
[pancli] syscontrol node 10.72.4.12 led off
No longer flashing LED for node "10.72.4.12"
```

Step 9: Reassign Old IP Address to Replacement Director

If the old Director was a repset member, and if clients have the IP address of the old Director hard-coded into a mount command, you may want the replacement Director to obtain that same IP address. This procedure may only be used on a Director that is not currently a member of the repset.

Determine which Director currently serves as the DHCP server:

```
[pancli] syscontrol realmmgr
NODE ADDRESS ROLE ON/OFFLINE
VCH-1,1 10.72.4.11 RM, MConsole, DNS Online
VCH-3,1 10.72.4.13 RM President, DHCP, DNS Online
```

Here, the DHCP server is running on the 10.72.4.13 Director. SSH to the DHCP Director and run support servicemode yes and netconfig ipranges ipchange:

```
[pancli] support servicemode yes
[pancli_service] netconfig ipranges ipchange
This command will allow you to change the IP of a node to a new IP provided that the new IP is in the given IP ranges and is not used. Note that this will force a reboot of the node whose IP is getting updated.
```

Enter the IP address that the replacement Director currently has. In this example, that is 10.72.4.36:

```
Enter the current IP address of the node:
```

Enter the IP address that the customer wishes the Director to obtain. In this example, that is 10.72.4.12:

```
Enter the new IP address of the node:
```

The script will restart the DHCP server and then reboot the replacement Director so it can obtain its new IP address:

```
Restarting DHCP server.
Rebooting the remote node.
Node 10.72.4.36 is rebooting
```

Wait for about five minutes. Confirm that the replacement Director has come online and has the desired IP address:

```
[pancli] sysmap nodes directors
Node Type Serial IP Address State
VCH-1,1 Director 23a30338300028001 10.72.4.11 Online
VCH-2,1 Director 23a3033f300028001 10.72.4.12 Online
VCH-3,1 Director 23a3039d300028001 10.72.4.13 Online
```

Step 10: Add Replacement Director to Repset

If the Director that was replaced was a member of the repset, use syscontrol realmmgr add [replacement director IP] to add the replacement Director to the repset. This command may take about a minute to complete.

This example assumes that the replacement Director's IP address has been changed to match the removed Director's IP address:

```
[pancli] syscontrol realmmgr add 10.72.4.12
Added Director: 10.72.4.12 (01c7ef23af0004)
```

Verify that the repset now has either three or five members. If the realm contains less than three Directors, the repset should only have one member:

```
[pancli] syscontrol realmmgr
NODE ADDRESS ROLE ON/OFFLINE
VCH-1,1 10.72.4.11 RM, MConsole, DNS Online
VCH-2,1 10.72.4.12 RM, DNS Online
VCH-3,1 10.72.4.13 RM President, DHCP, DNS Online
```

Step 11: Rebalance Volume Metadata Services

If a volume service was moved in step 2, a Director likely now has too many volume service assignments. The typical way to rebalance the volume service assignments is to move the volume services back to the replacement Director. This example assumes the original volservice move commands from step 2, and that the replacement Director has been reassigned the IP address 10.72.4.12. These commands move the same Primary and Backup assignments as the commands run in step 2.

Use volservice list to view current assignments:

```
[pancli] volservice list
Service Service Primary Primary Backup Backup Volume
ID Status Director IP Director IP Count Volumes
0x0400000000000003(FM) online VCH-1,1 10.72.4.11 VCH-3,1 10.72.4.13 1 /
0x0400000000000004(FM) online VCH-1,1 10.72.4.11 VCH-3,1 10.72.4.13 1 /mdube
0x0400000000000032(FM) online VCH-1,1 10.72.4.11 VCH-3,1 10.72.4.13 1 /esmith
0x0400000000000005(FM) online VCH-3,1 10.72.4.13 VCH-1,1 10.72.4.11 1 /rroy
```

Use volservice move [Service ID] [Primary IP] to move an assignment:

```
[pancli] volservice move 0x0400000000000032 10.72.4.12
Manual move of service id 0x0400000000000032 to 10.72.4.12 completed successfully.
[pancli] volservice move backup 0x0400000000000005 10.72.4.12
Manual move of backup for service id 0x0400000000000005 to 10.72.4.12 completed successfully.
```

Use volservice list again to verify that each Director has a primary and backup assignment:

```
[pancli] volservice list
Service Service Primary Primary Backup Backup Volume
ID Status Director IP Director IP Count Volumes
0x0400000000000003(FM) online VCH-1,1 10.72.4.11 VCH-3,1 10.72.4.13 1 /
0x0400000000000004(FM) online VCH-1,1 10.72.4.11 VCH-3,1 10.72.4.13 1 /mdube
0x0400000000000032(FM) online VCH-2,1 10.72.4.12 VCH-1,1 10.72.4.11 1 /esmith
0x0400000000000005(FM) online VCH-3,1 10.72.4.13 VCH-2,1 10.72.4.12 1 /rroy
```

Step 12: Confirm Director Is Eligible to Receive an Alias

If aliases were disabled for the unhealthy Director in step 3, that configuration will not apply to the replacement Director. The realm president will not move a working alias to the replacement Director immediately, but the replacement Director should receive an IP alias the next time aliases are assigned, if applicable. Confirm the Director is eligible to receive an alias with [pancli] sysmap node [director IP].

```
[pancli] sysmap node 10.72.4.12
Name: VCH-2,1
State: Online
Type: DB18-48
Services: <none>
IP Address: 10.72.4.12
Primary Aliases: <none>
Secondary Aliases: <none>
10GE Secondary Only: disabled
```

7.1.2 Replace an Individual Drive in a Node

Director Nodes contain 12 SSDs located in bays in the front of the node. From the front, the SSDs are numbered from left to right:

1 2 3 4 5 6

7 8 9 10 11 12

Each drive slot has three status LEDs.

Status LED	Activity Description
Blue Drive Activity Indicator	- Solid: SSD is installed - Blinking: Drive activity detected - Off: No drive detected
Yellow Drive Failure Indicator	- Solid: Drive fault - Blinking: Drive rebuild - Off: Normal
Green Location Indicator	- Solid: Drive location activated - Off: Normal

If a drive has failed, the yellow Drive Failure Indicator LED may be lit. PanFS does not have direct control over the failure LEDs for the SSDs. The failure LED only turns on if the BIOS or BMC figures out that the SSD has failed, which may not always occur. Consult with VDURA Technical Support to determine which bay contains the failing SSD.

Contact VDURA Technical Support for detailed instructions on how to proceed. When ready, follow the steps below to physically replace the drive.

> **Caution:** Verifying that an individual drive is the cause of a past or impending failure requires working with VDURA Technical Support. Attempting the steps below without assistance may cause system interruption or even data loss. Confirm the drive number from VDURA Technical Support to positively identify the drive number so you do not accidentally replace the wrong drive in the following step.

1. Remove the bezel from the front of the node by sliding the clips located toward each end and then pulling the bezel forward.
2. Press the ejector button for the failed drive to release its tray handle.
3. Use the tray handle to pull the drive and carrier out of its slot.
4. Insert the replacement drive in its carrier into the empty slot. Push it in and close the tray handle. At this point, the blue power LED will light.

7.1.3 Replace a Power Supply Unit

Director Nodes have two power supplies providing redundant power to the system. If either module fails, the other assumes the system load, allowing the node to continue uninterrupted operation. While the power supplies are hot swappable, one power supply must always be operating, so do not attempt to replace both power supplies at the same time unless you shut down the entire enclosure first.

Each power supply has a status LED.

Status LED	Activity Description
Green	- Solid: Powered and functioning normally - Off: No power to both PSUs - Blinking: Standby; AC connected

Status LED	Activity Description
Red	- Solid: Critical event, PSU shutdown: failure, OCP, OVP, fan failure - Blinking: Non-critical event, PSU continues to operate: high temp, high power, high current, slow fan
Red and green	- Blinking alternate colors: Warning event; power supply continues to operate; high temp, high power, high current, slow fan

If a power supply has failed, the Information LED on the node's control panel will slowly flash red.

To identify a failed power supply, look at the LEDs on the rear of the power supplies. If a power supply has failed, its LED will be red.

1. Disconnect the power cable from the failed power supply.
2. Push the ejector and use the handle to pull the power supply module out of the chassis.
3. Insert a new power supply into the chassis and push until it is fully seated. The LED on the rear of the module will turn blue.
4. Plug the power cable into the module. The power supply LED will turn green.

7.2 V5000 Storage Node Component Replacement

This section describes how to replace Mixed Fleet and All-Flash Storage Node components and covers the following topics:

- Replace a Storage Node
- Replace an Individual Drive in a Node
- Replace a Power Supply Unit

> **Warning:** Use caution when you remove or replace system components; they may be hot to the touch.

7.2.1 Replace a Storage Node

Follow the instructions below to replace a Storage Node.

1. PanActive Manager will alert you to the node that needs to be replaced. For detailed information on setting up alerts, see the VDURA Data Platform User Guide.

2. If the node is in Failed state, the data residing on it will already have been reconstructed to other nodes. Continue to step 3 to physically replace the node.

If the node is in Reconstruction state, please wait until reconstruction is complete.

If the node is still online, drain the node and clear the record before removing it. In PanActive Manager, click the Hardware tab and select the node you want to remove. From the appropriate Node Management page, click Start OSD Drain, then confirm that you want to drain the node. When finished, proceed to step 3.

3. Use the button on the node's control panel to turn off the power.
4. Remove the drives from the front. Remember which slot each drive was in.
5. Disconnect all cables from the back of the node. Failure to do this step can result in damaging the NIC or other rear ports.
6. Remove the node from the enclosure.
7. Reinstall the drives from the original node into the replacement node. Place the drives into the same slot locations as in the old node.
8. Reconnect the network cables as they were connected to the old node and turn the node's power back on.
9. Reinitialize the node to return it to service by issuing the following commands:

```
[pancli] support service yes
[pancli_service] syscontrol node <VCH-#,#> reinit
```

You should see the following output:

```
Reinitializing node "<VCH-#,#>".
```

The reinitialization process takes approximately five minutes. Please consult with VDURA Technical Support if the node has not rejoined the realm after 15 minutes.

10. The VDURA system automatically detects the new node.

7.2.2 Replace an Individual Drive in a Node

If the system has alerted you that an individual Storage Node has failed and is being reconstructed out, or an individual drive in a node is emitting SMART errors, you may be able to replace the faulty drive rather than the entire node to correct the problem and bring the node back into service.

Storage Nodes contain 12 SSDs located in bays in the front of the node. From the front, the SSDs are numbered from left to right:

1 2 3 4 5 6

7 8 9 10 11 12

Each drive slot has three status LEDs.

Status LED	Activity Description
Blue Drive Activity Indicator	- Solid: SSD is installed - Blinking: Drive activity detected - Off: No drive detected
Yellow Drive Failure Indicator	- Solid: Drive fault - Blinking: Drive rebuild - Off: Normal
Green Location Indicator	- Solid: Drive location activated - Off: Normal

1. Drain or reconstruct out the node. If the node in question is still participating in the StorageSet, you will need to drain or reconstruct it out of the StorageSet before proceeding. VDURA Technical Support will let you know which option to use based on the situation.

Physically removing a drive before beginning reconstruction or completing a drain will cause unnecessary system interruption.

If the node is not participating in the StorageSet, proceed to step 2.

In PanActive Manager, click the Hardware tab, then select the node you want to remove. From the appropriate Node Management page, either click Start OSD Drain or Reconstruct depending on the desired action, then confirm. Wait for drain or reconstruction to fully complete, either by the step above or automatically by the system.

2. If a drive has failed, the yellow Drive Failure Indicator LED may already be lit. PanFS does not have direct control over the failure LEDs for the SSDs. The failure LED only turns on if the BIOS or BMC figures out that the SSD has failed, which may not always occur. Consult with VDURA Technical Support to determine which bay contains the failing SSD.

Contact VDURA Technical Support for detailed instructions on how to proceed. When ready, follow the steps below to physically replace the drive.

> **Caution:** Verifying that an individual drive is the cause of a past or impending failure requires working with VDURA Technical Support. Attempting the steps below without assistance may cause system interruption or even data loss. Confirm the drive number from VDURA Technical Support to positively identify the drive number so you do not accidentally replace the wrong drive in the following step.

1. Remove the bezel from the front of the node by sliding the clips located toward each end and then pulling the bezel forward.
2. Press the ejector button for the failed drive to release its tray handle.
3. Use the tray handle to pull the drive and carrier out of its slot.
4. Insert the replacement drive in its carrier into the empty slot. Push it in and close the tray handle. At this point, the blue power LED will light.
5. While logged in as admin on one of the Directors, turn off the location LED on the drive in question with the following PanCLI command:

```
[pancli] syscontrol node <VCH-#,#> drive <#> led normal
```

In the above command, replace <VCH-#,#> with the name of the reconstructed or drained node and <#> with the drive number of the failed or failing drive. This command should yield the following output:

```
Successfully set LED for drive <#> on node <VCH-#,#> to locate.
```

6. Finally, reinitialize the node to return it to service by issuing the following commands:

```
[pancli] support service yes
[pancli_service] syscontrol node <VCH-#,#> reinit
```

You should see the following output:

```
Reinitializing node "<VCH-#,#>".
```

The reinitialization process takes approximately five minutes. Please consult with VDURA Technical Support if the node has not rejoined the realm after 15 minutes.

7.2.3 Replace a Power Supply Unit

Storage Nodes have two power supplies providing redundant power to the system. If either module fails, the other assumes the system load, allowing the node to continue uninterrupted operation. While the power supplies are hot swappable, one power supply must always be operating, so do not attempt to replace both power supplies at the same time unless you shut down the entire enclosure first.

Each power supply has a status LED.

Status LED	Activity Description
Green	- Solid: Powered and functioning normally - Off: No power to both PSUs - Blinking: Standby; AC connected
Red	- Solid: Critical event, PSU shutdown: failure, OCP, OVP, fan failure - Blinking: Non-critical event, PSU continues to operate: high temp, high power, high current, slow fan
Red and green	- Blinking alternate colors: Warning event; power supply continues to operate; high temp, high power, high current, slow fan

If a power supply has failed, the Information LED on the node's control panel will slowly flash red.

To identify a failed power supply, look at the LEDs on the rear of the power supplies. If a power supply has failed, its LED will be red.

1. Disconnect the power cable from the failed power supply.
2. Push the ejector and use the handle to pull the power supply module out of the chassis.
3. Insert a new power supply into the chassis and push until it is fully seated. The LED on the rear of the module will turn blue.
4. Plug the power cable into the module. The power supply LED will turn green.

7.3 V5000 JBOD Component Replacement

This section describes how to replace JBOD components and covers the following topics:

- Replace a SAS Cable
- Replace a Power Supply Unit
- Replace a Fan

- Replace a Single Disk
- Remove a Drive from Its Carrier
- Replace an IOM
- Replace the JBOD Enclosure

You cannot replace an I/O module (IOM), power supply, or fans without first pulling the enclosure partially out of the rack and disengaging the Cable Management Arm (CMA). Depending on your cabling and space available, you may only need to disengage one end of the CMA and let it hang down. It may be easier to also disengage the other side to allow more freedom of movement. You must also detach the appropriate cables before you pull out either a power supply or the bottom IOM.

> **Note:** All items below can be hot-swapped while the JBOD is powered on.

7.3.1 Replace a SAS Cable

1. Disconnect the cable from the rear of the Storage Node.
2. Remove the screws securing the JBOD to the rack, then slide the JBOD forward enough to access its rear.
3. Open the cable baskets.
4. Disconnect the cable from the rear of the JBOD.
5. Connect the new cable to both empty ports.

7.3.2 Replace a Power Supply Unit

1. Remove the screws securing the JBOD to the rack, then slide the JBOD forward enough to access its rear.
2. Unplug the power cable from the PSU.
3. Press the green tab sideways to unlock the PSU.
4. Pull the handle to remove the PSU.
5. Insert the new PSU. Be sure the green tab clicks to lock it in place.
6. Plug the power cable into the new PSU.

7.3.3 Replace a Fan

1. Remove the screws securing the JBOD to the rack, then slide the JBOD forward enough to access its rear.
2. Lift the black tab to unlock the fan unit.
3. Pull the handle to remove the fan unit.
4. Insert the new fan unit. Be sure the black tab clicks to lock it in place.

7.3.4 Replace a Single Disk

If PanActive Manager has alerted you that an individual drive in the JBOD is emitting errors or failing, you may be able to replace the faulty drive and bring the node back into service.

> **Caution:** Verifying that an individual drive is the cause of a past or impending failure requires working with VDURA Technical Support. Attempting the steps below without assistance may cause system interruption or even data loss. Confirm the drive number from VDURA Technical Support to positively identify the drive number so you do not accidentally replace the wrong drive in the following step.

1. Drain or reconstruct out the node. If the node in question is still participating in the StorageSet, you will need to drain or reconstruct it out of the StorageSet before proceeding. VDURA Technical Support will let you know which option to use based on the situation.

Physically removing a drive before beginning reconstruction or completing a drain will cause unnecessary system interruption.

If the node is not participating in the StorageSet, proceed to step 2.

In PanActive Manager, click the Hardware tab, then select the node you want to remove. From the appropriate Node Management page, either click Start OSD Drain or Reconstruct depending on the desired action, then confirm. Wait for drain or reconstruction to fully complete, either by the step above or automatically by the system.

2. Turn on the drive's LED to locate it.

The drives in the 78-drive JBOD enclosure are numbered in the following order, as looking from above while in front of the JBOD:

Rear of JBOD

67 **sd67**	**68** **sd68**	**69** **sd69**	**70** **sd70**	**71** **sd71**	**72** **sd72**	**73** **sd73**	**74** **sd74**	**75** **sd75**	**76** **sd76**	**77** **sd77**	**78** **sd78**			
55 sd55	56 sd56	57 sd57	58 sd58	59 sd59	60 sd60	61 sd61	62 sd62	63 sd63	64 sd64	65 sd65	66 sd66			
43 sd43	44 sd44	45 sd45	46 sd46	47 sd47	48 sd48	49 sd49	50 sd50	51 sd51	52 sd52	53 sd53	54 sd54			
31 sd31	32 sd32	33 sd33	34 sd34	35 sd35	36 sd36	37 sd37	38 sd38	39 sd39	40 sd40	41 sd41	42 sd42			
16 sd16	17 sd17	18 sd18	19 sd19	20 sd20	21 sd21	22 sd22	23 sd23	24 sd24	25 sd25	26 sd26	27 sd27	28 sd28	29 sd29	30 sd30
1 sd1	2 sd2	3 sd3	4 sd4	5 sd5	6 sd6	7 sd7	8 sd8	9 sd9	10 sd10	11 sd11	12 sd12	13 sd13	14 sd14	15 sd15

Front of JBOD

The drives in the 108-drive JBOD enclosure are numbered in the following order, as looking from above while in front of the JBOD:

Rear of JBOD

97 **sd97**	**98** **sd98**	**99** **sd99**	**100** **sd100**	**101** **sd101**	**102** **sd102**	**103** **sd103**	**104** **sd104**	**105** **sd105**	**106** **sd106**	**107** **sd107**	**108** **sd108**
85 sd85	86 sd86	87 sd87	88 sd88	89 sd89	90 sd90	91 sd91	92 sd92	93 sd93	94 sd94	95 sd95	96 sd96
73 sd73	74 sd74	75 sd75	76 sd76	77 sd77	78 sd78	79 sd79	80 sd80	81 sd81	82 sd82	83 sd83	84 sd84
61 sd61	62 sd62	63 sd63	64 sd64	65 sd65	66 sd66	67 sd67	68 sd68	69 sd69	70 sd70	71 sd71	72 sd72
46 sd46	47 sd47	48 sd48	49 sd49	50 sd50	51 sd51	52 sd52	53 sd53	54 sd54	55 sd55	56 sd56	57 sd57	58 sd58	59 sd59	60 sd60
31 sd31	32 sd32	33 sd33	34 sd34	35 sd35	36 sd36	37 sd37	38 sd38	39 sd39	40 sd40	41 sd41	42 sd42	43 sd43	44 sd44	45 sd45
16 sd16	17 sd17	18 sd18	19 sd19	20 sd20	21 sd21	22 sd22	23 sd23	24 sd24	25 sd25	26 sd26	27 sd27	28 sd28	29 sd29	30 sd30
1 sd1	2 sd2	3 sd3	4 sd4	5 sd5	6 sd6	7 sd7	8 sd8	9 sd9	10 sd10	11 sd11	12 sd12	13 sd13	14 sd14	15 sd15

Front of JBOD

Log in as admin on a Director and enter the following PanCLI command:

```
[pancli] syscontrol node <VCH-#,#> drive <#> led locate
```

Replace <VCH-#,#> with the name of the reconstructed or drained node and <#> with the drive number of the failed or failing drive.

This command should yield the following output:

```
Successfully set LED for drive <#> on node <VCH-#,#> to locate.
```

A blinking LED will appear by the drive to be replaced. The LEDs on the front of the JBOD will also indicate which drive slot you have located.

3. Press the button beside the drive to unlock it, then remove it from the enclosure.

> **Caution:** To reduce the risk of shock damage to the drive, remove the drive just enough to make the location LED stop blinking. Wait at least 15 seconds before continuing to remove the drive; this will let the drive spin down completely.

4. Insert a replacement drive in a carrier into the bay.
5. Push the drive in as far as you can, then lock it in place.
6. While logged in as admin on one of the Directors, turn off the location LED on the drive in question with the following PanCLI command:

```
[pancli] syscontrol node <VCH-#,#> drive <#> led normal
```

In the above command, replace <VCH-#,#> with the name of the reconstructed or drained node and <#> with the drive number of the failed or failing drive. This command should yield the following output:

```
Successfully set LED for drive <#> on node <VCH-#,#> to locate.
```

7. Reinitialize the node to return it to service by issuing the following commands:

```
[pancli] support service yes
[pancli_service] syscontrol node <VCH-#,#> reinit
```

You should see the following output:

```
Reinitializing node "<VCH-#,#>".
```

The reinitialization process takes approximately five minutes. Consult with VDURA Technical Support if the node has not rejoined the realm after 15 minutes.

7.3.5 Remove a Drive from Its Carrier

1. The carrier is simply a piece of plastic that wraps around the drive. There are no screws to hold the carrier to the drive.

2. Find the slight plastic ridge in the middle of the carrier.
3. Grip the ridge with your fingers and pull the carrier slightly away from the drive.
4. Gently ease the drive out of the carrier.
5. Lever it up on the side opposite the latch hinge and where the button for the latch is located.
6. To put the drive back into the carrier, reverse the procedure.
7. Line up the side of the drive with the studs built into the inside of the carrier and fit the studs into the screw holes on the drive.
8. You may see that the drive hangs up on this corner, so you might want to do this corner first, then the other corner.

7.3.6 Replace an IOM

1. Disconnect all cables from the unit.
2. Loosen the two screws at the top of the module.
3. Use the handle to remove the module.
4. Insert the replacement module.
5. Tighten the two screws at the top of the module.
6. Reconnect all cables.

7.3.7 Replace the JBOD Enclosure

1. Take the realm or StorageSet offline.
2. Remove the disks. Pay attention to the slot each disk is in. You must place each disk in the same location slot in the new enclosure.

See Replace a Single Disk for tables listing slot and drive numbers.

3. Disconnect all cables from the enclosure. The cables must get connected back into the same ports on the replacement enclosure.
4. Remove the JBOD enclosure from the rack. Leave the PSUs, fans, and IOM in the enclosure; the replacement will have new ones.
5. Slide the new enclosure into the rack.
6. Reconnect all cables.
7. Insert the disks.
8. Bring the realm or StorageSet online.

8. System Specifications

This chapter provides basic operational and environmental specifications for the hardware in this guide. If you have any questions regarding the specifications outlined here, contact your sales representative, call VDURA Technical Support at 1-888-726-2727, or visit the VDURA support site at support.VDURA.com.

This chapter contains the following sections:

- V5000 Director Node Server Specifications
- V5000 Storage Node Server Specifications
- V5000 4U78 JBOD Specifications
- V5000 4U108 JBOD Specifications

8.1 V5000 Director Node Server Specifications

Physical Specifications

The following table lists the physical specifications.

W x D x H	**17.5 x 33 x 1.75 in. - 1 rack unit** **445 x 838 x 44 mm**
Net weight, enclosure plus drives	41 lbs. 18.6 kg

Environmental Specifications

The following table lists the environmental specifications.

Operating temperature	**0-35°C**
Non-operating temperature	-10-60°C
Operating humidity	N/A
Non-operating humidity	5-95% non-condensing

Power Specifications

The following table lists the power specifications.

Power supply output rating	**1300 W**
Input line current	6.25 A @ 208 VAC 50-60 Hz
Power consumption*	776 W @ 208 VAC (typical) 1201 W @ 208 VAC (maximum)
Operating current*	3.73 A @ 208 VAC (typical) 5.77 A @ 208 VAC (maximum)
BTU rating*	2647 BTU/hr @ 208 VAC (typical) 4097 BTU/hr @ 208 VAC (maximum)

- Values derived from example systems in use under normal load. Your actual results may vary based on configuration and load.

8.2 V5000 Storage Node Server Specifications

Physical Specifications

The following table lists the physical specifications.

W x D x H	**17.5 x 33 x 1.75 in. - 1 rack unit** **445 x 838 x 44 mm**
Net weight, enclosure plus drives	38.8 lbs. 17.6 kg

Environmental Specifications

The following table lists the environmental specifications.

Operating temperature	**0-35°C**
Non-operating temperature	-10-60°C

Operating temperature	**0–35°C**
Operating humidity	N/A
Non-operating humidity	5-95% non-condensing

Power Specifications

The following table lists the power specifications.

Power supply output rating	**1300 W**
Input line current	6.25 A @ 208 VAC 50-60 Hz
Power consumption*	722 W @ 208 VAC (typical) 1138 W @ 208 VAC (maximum)
Operating current*	3.47 A @ 208 VAC (typical) 5.47 A @ 208 VAC (maximum)
BTU rating*	2463 BTU/hr @ 208 VAC (typical) 3882 BTU/hr @ 208 VAC (maximum)

- Values derived from example systems in use under normal load. Your actual results may vary based on configuration and load.

8.3 V5000 4U78 JBOD Specifications

Physical Specifications

The following table lists the physical specifications.

W x D x H (without CMA)	**17.5 x 31.9 x 7 in.** **434 x 811 x 176 mm**
W x D x H (with CMA)	17 x 38.4 x 7 in. 434 x 975 x 176 mm
Weight without drives	94 lbs. 42.6 kg

W x D x H (without CMA)	**17.5 x 31.9 x 7 in.** **434 x 811 x 176 mm**
Weight with drives	207.2 lbs. 94.5 kg

Environmental Specifications

The following table lists the environmental specifications.

Operating temperature	**0-35°C**
Non-operating temperature	-20-60°C
Operating humidity	20-80%
Non-operating humidity	10-90%

Power Specifications

The following table lists the power specifications.

Power supply output rating	**1600 W**
Input line current	7.69 A @ 208 VAC 50-60 Hz
Power consumption*	839 W @ 208 VAC (typical) 1155 W @ 208 VAC (maximum)
Operating current*	4.03 A @ 208 VAC (typical) 5.55 A @ 208 VAC (maximum)
BTU rating*	2862 BTU/hr @ 208 VAC (typical) 3940 BTU/hr @ 208 VAC (maximum)

- Values derived from example systems in use under normal load. Your actual results may vary based on configuration and load.

8.4 V5000 4U108 JBOD Specifications

Physical Specifications

The following table lists the physical specifications.

W x D x H (without CMA)	**17 x 41 x 7 in.** **434 x 1050 x 176 mm**
W x D x H (with CMA)	17 x 50 x 7 in. 434 x 1273 x 176 mm
Weight without drives	105 lbs. 47.6 kg
Weight with drives	263.2 lbs. 119.4 kg

Environmental Specifications

The following table lists the environmental specifications.

Operating temperature	**0-35°C**
Non-operating temperature	-20-60°C
Operating humidity	20-80%
Non-operating humidity	10-90%

Power Specifications

The following table lists the power specifications.

Power supply output rating	**2000 W**
Input line current	9.62 A @ 208 VAC 50-60 Hz
Power consumption*	1081 W @ 208 VAC (typical) 1448 W @ 208 VAC (maximum)

Power supply output rating	2000 W
Operating current*	5.20 A @ 208 VAC (typical) 6.96 A @ 208 VAC (maximum)
BTU rating*	3688 BTU/hr @ 208 VAC (typical) 4940 BTU/hr @ 208 VAC (maximum)

- Values derived from example systems in use under normal load. Your actual results may vary based on configuration and load.

9. Troubleshooting

This chapter describes some basic troubleshooting guidelines that may be useful in diagnosing problems in new and existing installations. This is not an exhaustive list and does not replace the need to involve VDURA Technical Support.

This chapter contains the following sections:

- Collect Logs
- Send Logs
- Log Analysis
- Diagnostic Utilities
- Configure the Maximum Transmission Unit

9.1 Collect Logs

This section covers sysstat all and collect_logs.

9.1.1 sysstat all

The sysstat all command provides an overall VDURA system summary. Examples of output include: Director Node CPU utilization, Storage Node disk and CPU utilization, node status, StorageSet information, volume information, network configuration, gateway services, and current warnings and errors.

```
[pancli] sysstat all > /tmp/sysstat_all_{SR-XXXXXXX}
```

You can also get this data by clicking the Download System Summary button on the Status & Reports page on the PanActive Manager web interface. VDURA Technical Support may ask you to save the output as a text file and attach it to a Service Request.

9.1.2 collect_logs

The collect_logs command collects logs commonly used to diagnose issues from specified nodes. VDURA recommends to always collect logs from all repset nodes in addition to any other nodes of interest. In the event of a PanActive Manager outage, the repset option will continue to find the correct nodes to collect from. The host option followed by an IP address should also always work.

```
collect_logs ?[option] [option] ...?
```

Below is a list of all command options.

Option	Description
`age [count] [units]`	Do not collect /var/log/*.[0-9]+ files and corefiles that are more than count units old, where valid units are minute(s), hour(s), day(s), week(s), month(s). Default is one week. If the match option is specified, it defaults to unlimited age. If the count is 0, age is not measured.
`all`	Collect logs from all nodes in the system. This option is only available when collect_logs is run on a Director in the replication set.
`storageset [name]`	Collect logs from all nodes in the named StorageSet. You can use this option multiple times to specify multiple StorageSets. This option is available only when running **collect_logs** on a Director in the replication set.
`container`	Collect logs from all VDURA Data Platforms in the system. This option is only available when collect_logs is run on a Director in the replication set.
`directors`	Collect logs from all Directors in the system. This option is available only when running **collect_logs** on a Director in the replication set.

Option	Description
`enclosure [name]`	Collect logs from all nodes in the named enclosure. You can use this option multiple times to specify multiple enclosures. This option is available only when running **collect_logs** on a Director in the replication set.
`enclosuremaster`	Collect logs from nodes with the enclosuremaster role. One enclosuremaster node in each enclosure controls the enclosure hardware. This option is only available when the command is run on a Director in the replication set.
`fileowner [PanFS file]|[Object ID]`	Collect logs from the nodes that store a portion of the given PanFS file or Object ID, as well as the primary and backup Directors for the file or object's volume. Owners of multiple files can be specified by using this option more than once.
`host [name]`	Collect logs from the node with this name, IP address, or serial number. You can use this option multiple times to specify multiple nodes.
`match [file path pattern]`	Collect only those files whose names match the specified glob-style pattern. Use a glob pattern that captures the full absolute path to the desired files (for example, `/var/log/*http*` for files in `/var/log` that have the substring `http` somewhere in their name). Multiple patterns can be specified by using this option more than once.
`nocores`	Do not collect any corefiles or executables. Stack files are always gathered.
`nodbs`	Do not collect any system databases.
`nodiags`	Do not run any diagnostic tools before collecting logs.

Option	Description
`panfs [dir]`	Copy the output file to the given PanFS directory. The directory will be created if it does not already exist.
`range [ip]-[ip]`	Collect logs from nodes in this range of IP addresses. You can use this option multiple times to specify multiple ranges. This option is available only when running **collect_logs** on a Director in the replication set.
`repset`	Collect logs from all Directors in the replication set. This option is available only when running **collect_logs** on a Director in the replication set.
`scp [user] [host] [dir]`	Scp the output file to the given user, host, and directory arguments. You will be prompted to enter a password. The target directory must exist on that host already.
`storage`	Collect logs from all Storage Nodes in the system. This option is only available when the command is run on a Director in the replication set.
`storagemodules`	Collect logs from all storage modules in the system. This option is only available when collect_logs is run on a Director in the replication set.
`timeout [seconds]`	The timeout for **scp** in seconds. The default is 900 seconds (15 minutes). Use 0 for no timeout.
`upload [user]`	Upload the output file to upload.vdura.com as the given user. You will be prompted to enter a password.

Option	Description
`volumeowner [volume name]`	Collect logs from the Director that owns the named volume. If Volume Services Failover is enabled, also collect logs from the backup Director for the given volume. You can use this option multiple times to specify multiple volume owners. This option is available only when running **collect_logs** on a Director in the replication set.
`waitsavecore`	Wait until savecore is no longer running before collecting logs.

Type help collect_logs [subcommand] to see more details for individual subcommands.

Command Examples

Gather logs for all nodes in VCH-2 and VCH-5, as well as all nodes in the replication set:

```
collect_logs enclosure VCH-2 enclosure VCH-5 repset
```

Gather logs for the enclosure master nodes in VCH-2 and VCH-5, as well as all nodes in the repset:

```
collect_logs enclosuremaster enclosure VCH-2 enclosure VCH-5 repset
```

Gather logs for all Storage Nodes in the StorageSet Set 2, as well as all nodes in the repset:

```
collect_logs storageset "Set 2" repset
```

Gather logs excluding corefiles for all Directors in VCH-4, as well as all nodes in the repset:

```
collect_logs nocores directors enclosure VCH-4 repset
```

Gather logs for both primary and backup Directors for the /home and /build volumes, as well as all Directors in the repset:

```
collect_logs volumeowner home volumeowner build repset
```

Gather logs for the Storage Nodes that cover the file /home/myFile, both the primary and backup Directors for the /home volume, and all Directors in the repset:

```
collect_logs fileowner /home/myFile repset
```

Gather all core/stack files that are less than two days old for the given node:

```
collect_logs host VCH-1,3 match /var/corefiles/* age 2 days
```

By default, collect_logs runs with a timeout value of 900 seconds, or 15 minutes. This timeout is often not adequate for logs from many hosts or for large logs with corefiles. To allow the command to run longer, use a long timeout value such as 3600 (one hour) or 7200 (two hours) for large logs collections.

You can also specify timeout 0 to specify running with no timeout. If timeout 0 is used, collect_logs will run until it either completes or the user ends it with CTRL+C. If you're collecting logs from a host that you think may not respond, consider not using the timeout 0 parameter.

The example below collects files needed to read a corefile; if the nocores option is used, these files will not be collected:

```
[pancli] collect_logs timeout 3600 host {node} waitsavecore age 2 days
```

Note the longer value to prevent a timeout while the command is running.

The example below collects logs for a core where the core sweeper should have already had time to run:

```
[pancli] collect_logs timeout 3600 host {node} age 2 days
```

The age 2 days option instructs the command to only gather the previous two days of data. This covers the date of the crash so you get cores from that date based on error messages telling you what day the crash occurred without extra data from prior to then.

In the following examples, host {node} is used to specify a particular node or nodes by either name or IP address:

```
[pancli] collect_logs host {node} age 12 hours
[pancli] collect_logs host {node} host {node} age 2 days
```

In the following examples, the parameter directly after collect_logs is used to specify a particular group of nodes:

```
[pancli] collect_logs repset age 2 days
[pancli] collect_logs directors age 12 hours
[pancli] collect_logs storage age 2 days
[pancli] collect_logs storageset "Set 1" age 2 days
[pancli] collect_logs enclosure {enclosure} age 2 days
```

In the following examples, the parameter directly after collect_logs is used to specify volume owner, file owner, or enclosuremaster nodes.

```
[pancli] collect_logs volumeowner {volume_name} age 2 days
[pancli] collect_logs volumeowner {volume_name} age 12 hours
[pancli] collect_logs fileowner {/volume_name/path/to/file or object_ID} age 12 hours
[pancli] collect_logs enclosuremaster age 4 days
[pancli] collect_logs enclosuremaster enclosure {enclosure} age 4 days
```

> **Note:** The fileowner option does not work on fenced objects.

For a failed or stuck reconstruction, use the age argument to go back to a time shortly before when reconstruction failed:

```
[pancli] collect_logs directors match /var/log/all* age 12 hours
```

If the volume has snapshots, also get the realm database:

```
[pancli] support servicemode yes
[pancli_service] diag_exec -timeout 1000 "pan_dump_realm" > /tmp/rmdb_{SR-XXXXXXX}
```

Use the match option to collect files that match a specified pattern:

```
[pancli] collect_logs all match /var/corefiles/*stack match /var/crash/*stack age 2 days
[pancli] collect_logs host {node} match /var/log/all* age 8 months
[pancli] collect_logs match /var/log/diag/driveS*.log match /var/log/all.log directors
```

Linux node health check only:

```
[pancli] collect_logs match /pan/var/log/diag/driveS*.log match /pan/var/log/all.log
storagemodules
```

Log collections containing match report an error if the node does not have matching files; this is normal.

If logs fail to collect, ping the node. If the ping succeeds, try to ssh into the host and collect logs locally.

9.2 Send Logs

Once you have collected logs from a server, you may want to send them to VDURA for analysis. Below are scp commands to move log files from a server to your client.

Pull to client from system:

1. Log in to the client.
2. Use the following command to copy a file from a server:

```
scp username@remote:/file/to/get /where/to/put
```

Example:

```
scp admin@serv1.test_lab.vdura.com:/tmp/vol-list ./vol-list
```

Push from system to client:

1. Log in to the VDURA system.
2. Use the following command to copy a file to your client:

```
scp /file/to/send username@remote:/where/to/put
```

Example:

```
scp /var/tmp/vol-list auser@serv1.vdura.com:/fs/home/auser/vol-list
```

9.3 Log Analysis

The VDURA system actively monitors the network and keeps a historical record of events. This information is captured in system logs. If a suspected network problem is occurring, the net_mon.log, all.log, and sysstat.log logs described on the following pages are your best chance to determine root cause.

When looking at logs, be aware that, compared to other VDURA components, the net_mon.log and all.log files for Storage Nodes are located in /pan/var/log/. Furthermore, this folder will be empty until the PanCLI command journal.ctl is used. When run, this command finds all the relevant logs and places them in /pan/var/log.

9.3.1 /var/log/net_mon.log

Every ten minutes, each node in the VDURA system polls vital statistics from the network interfaces and logs them in net_mon.log. Results for ifconfig, netstat -r, and netstat -s are stored here.

ifconfig

Though an administrator should never manipulate the Ethernet interfaces on the node using ifconfig, the output of the command is useful. From this output you can see the state of the link, the effective MTU, and how failover is configured.

The example below is from a Director Node. The laggproto is set to lacp, so cc0 and cc1 are present and active. The two physical interfaces cc0 and cc1 do not have IP interfaces bound to them; they are added to the lagg0 interface to accomplish bonding. The lagg0 laggport information shows which physical interfaces are part of the bond, and the physical interface information shows that the links are active and their speed.

Also notice the igb0 and igb1 ports. VDURA field engineers use igb0, which is part of lagg2, as a service port during installation. If you are using IPMI on the node as this example is, that connection is made on

igb1.

```
igb0: flags=8843<UP,BROADCAST,RUNNING,SIMPLEX,MULTICAST> metric 0 mtu 1500
options=e527bb<RXCSUM,TXCSUM,VLAN_MTU,VLAN_HWTAGGING,JUMBO_MTU,VLAN_HWCSUM,TSO4,TSO6,LRO,WOL_MA
GIC,VLAN_HWFILTER,VLAN_HWTSO,RXCSUM_IPV6,TXCSUM_IPV6>
ether ac:1f:6b:cf:8d:18
media: Ethernet autoselect
status: no carrier
nd6 options=29<PERFORMNUD,IFDISABLED,AUTO_LINKLOCAL>
igb1: flags=8843<UP,BROADCAST,RUNNING,SIMPLEX,MULTICAST> metric 0 mtu 1500
options=e527bb<RXCSUM,TXCSUM,VLAN_MTU,VLAN_HWTAGGING,JUMBO_MTU,VLAN_HWCSUM,TSO4,TSO6,LRO,WOL_MA
GIC,VLAN_HWFILTER,VLAN_HWTSO,RXCSUM_IPV6,TXCSUM_IPV6>
ether a4:bf:01:65:73:c1
media: Ethernet autoselect (1000baseT <full-duplex>)
status: active
nd6 options=29<PERFORMNUD,IFDISABLED,AUTO_LINKLOCAL>
cc0: flags=8843<UP,BROADCAST,RUNNING,SIMPLEX,MULTICAST> metric 0 mtu 9000
options=2ec07bb<RXCSUM,TXCSUM,VLAN_MTU,VLAN_HWTAGGING,JUMBO_MTU,VLAN_HWCSUM,TSO4,TSO6,LRO,VLAN_
HWTSO,LINKSTATE,RXCSUM_IPV6,TXCSUM_IPV6,HWRXTSTMP>
ether 00:07:43:4b:91:80
media: Ethernet autoselect (25GBase-CR <full-duplex>)
status: active
nd6 options=29<PERFORMNUD,IFDISABLED,AUTO_LINKLOCAL>
cc1: flags=8843<UP,BROADCAST,RUNNING,SIMPLEX,MULTICAST> metric 0 mtu 9000
options=2ec07bb<RXCSUM,TXCSUM,VLAN_MTU,VLAN_HWTAGGING,JUMBO_MTU,VLAN_HWCSUM,TSO4,TSO6,LRO,VLAN_
HWTSO,LINKSTATE,RXCSUM_IPV6,TXCSUM_IPV6,HWRXTSTMP>
ether 00:07:43:4b:91:80
hwaddr 00:07:43:4b:91:88
media: Ethernet autoselect (25GBase-CR <full-duplex>)
status: active
nd6 options=29<PERFORMNUD,IFDISABLED,AUTO_LINKLOCAL>
lo0: flags=8049<UP,LOOPBACK,RUNNING,MULTICAST> metric 0 mtu 16384
options=680003<RXCSUM,TXCSUM,LINKSTATE,RXCSUM_IPV6,TXCSUM_IPV6>
inet6 ::1 prefixlen 128
inet6 fe80::1%lo0 prefixlen 64 scopeid 0x5
inet 127.0.0.1 netmask 0xff000000
groups: lo
nd6 options=21<PERFORMNUD,AUTO_LINKLOCAL>
lagg0: flags=8843<UP,BROADCAST,RUNNING,SIMPLEX,MULTICAST> metric 0 mtu 9000
```

```
options=2ec07bb<RXCSUM,TXCSUM,VLAN_MTU,VLAN_HWTAGGING,JUMBO_MTU,VLAN_HWCSUM,TSO4,TSO6,LRO,VLAN_
HWTSO,LINKSTATE,RXCSUM_IPV6,TXCSUM_IPV6,HWRXTSTMP>
ether 00:07:43:4b:91:80
inet 10.65.244.11 netmask 0xffffff00 broadcast 10.65.244.255
laggproto lacp lagghash l2,l3,l4
laggport: cc0 flags=1c<ACTIVE,COLLECTING,DISTRIBUTING>
laggport: cc1 flags=1c<ACTIVE,COLLECTING,DISTRIBUTING>
groups: lagg
media: Ethernet autoselect
status: active
nd6 options=29<PERFORMNUD,IFDISABLED,AUTO_LINKLOCAL>
lagg1: flags=8843<UP,BROADCAST,RUNNING,SIMPLEX,MULTICAST> metric 0 mtu 1500
options=800000<>
ether 00:07:43:4b:91:88
laggproto failover lagghash l2,l3,l4
groups: lagg
media: Ethernet autoselect
status: no carrier
nd6 options=29<PERFORMNUD,IFDISABLED,AUTO_LINKLOCAL>
lagg2: flags=8843<UP,BROADCAST,RUNNING,SIMPLEX,MULTICAST> metric 0 mtu 1500
options=e527bb<RXCSUM,TXCSUM,VLAN_MTU,VLAN_HWTAGGING,JUMBO_MTU,VLAN_HWCSUM,TSO4,TSO6,LRO,WOL_MA
GIC,VLAN_HWFILTER,VLAN_HWTSO,RXCSUM_IPV6,TXCSUM_IPV6>
ether ac:1f:6b:cf:8d:18
inet 169.254.12.34 netmask 0xffff0000 broadcast 169.254.255.255
laggproto failover lagghash l2,l3,l4
laggport: igb0 flags=1<MASTER>
groups: lagg
media: Ethernet autoselect
status: no carrier
nd6 options=29<PERFORMNUD,IFDISABLED,AUTO_LINKLOCAL>
```

netstat -r

The net_mon.log also contains the output of netstat -r. This will show you all of the selected routes active on the current node. This is most useful for routing problems.

```
[pancli] netstat -rn
Routing tables
Internet:
Destination Gateway Flags Refs Use Netif Expire
default 10.17.221.1 UGS 0 1918 lagg0
10.17.221.0/24 link#6 UC 0 0 lagg0
10.17.221.1 00:01:e8:8a:f0:bf UHLW 2 0 lagg0 452
10.17.221.11 00:60:dd:45:bc:4e UHLW 1 15353 lo0
10.17.221.70 00:09:03:01:b1:5b UHLW 1 799 lagg0 599
10.17.221.72 00:09:03:01:ce:58 UHLW 1 793 lagg0 600
...
10.17.221.111 00:09:03:01:c6:db UHLW 1 789 lagg0 599
10.17.221.112 00:09:03:01:b4:81 UHLW 1 778 lagg0 598
10.17.221.113 00:09:03:01:bc:3b UHLW 1 743 lagg0 599
10.17.221.114 00:09:03:01:d2:38 UHLW 1 778 lagg0 599
127.0.0.1 127.0.0.1 UH 0 3488 lo0
255.255.255.255 00:00:00:00:00:00 UHSb 0 134 lagg0
Internet6:
Destination Gateway Flags Netif Expire
::1 ::1 UHL lo0
fe80::%lo0/64 fe80::1%lo0 U lo0
fe80::1%lo0 link#5 UHL lo0
ff01:5::/32 fe80::1%lo0 UC lo0
ff02::%lo0/32 fe80::1%lo0 UC lo0
```

netstat -s

An historical record of tcp statistics can help determine if a failure was network related or not. The netstat -s is the delta from the last record. If you see a value of 100, it is really 100 in the last 10-minute interval. If the difference between the two runs is 0, the value is not listed.

```
tcp:
11944 packets sent
4213 data packets (4683796 bytes)
18 data packets (5273 bytes) retransmitted
17 data packets unnecessarily retransmitted
4690 ack-only packets (1347 delayed)
596 window update packets
2006 control packets
11307 packets received
6808 acks (for 4682649 bytes)
1607 duplicate acks
5467 packets (6040202 bytes) received in-sequence
112 completely duplicate packets (14665 bytes)
596 window update packets
2 out-of-order packets (68 bytes)
97 packets received after close
629 connection requests
944 connection accepts
1573 connections established (including accepts)
1604 connections closed (including 100 drops)
286 connections updated cached RTT on close
286 connections updated cached RTT variance on close
282 connections updated cached ssthresh on close
6808 segments updated rtt (of 5909 attempts)
18 retransmit timeouts
212 keepalive timeouts
204 keepalive probes sent
448 correct ACK header predictions
913 correct data packet header predictions
944 syncache entries added
944 completed
944 cookies sent
```

Look for the following values to flag possible error situations:

- completely duplicate packets: These are packets that were retransmitted by the sender. This can indicate a lost ACK or reordered packets in the write direction.
- out-of-order packets: These are packets that were reordered in the network. This can be caused by non-TCP-flow-based ECMP or LAGs. It can also be caused by small amounts of packet loss in the write direction.
- retransmit timeouts: This is the number of times we gave up on a TCP segment after not receiving an ACK. This typically means a 50ms timeout has been incurred. You should not see these in high numbers. This typically signals excessive packet loss in the read direction.

9.3.2 /var/log/all.log

A number of messages listed in the all.log file have specific bearing on network activities. Following are examples of some of these messages.

This message indicates that the layer 2 stats are incrementing beyond a preset threshold:

```
cpu_mon: cc0: dot3StatsFCSErrors exceed threshold (178 > 1) during previous 10 minutes
```

This message shows the NIC on a node losing link. This could indicate a failure of the NIC hardware on the node:

```
Jul 19 11:31:21 01ab8923af0003 kernel: cc0: link state changed to DOWN
Jul 19 11:31:23 01ab8923af0003 kernel: cc0: link state changed to UP
```

Another key message to look for is the rioe error. This is short for Report IO Error. When a PanFS client is unable to communicate to a Storage Node, it returns a rioe to the metadata manager. These errors are logged and used to isolate network failures.

```
pan_mgr[796]: sm: 0x02008989187f0010(SM): rioe [1269490094:732398995]: client 10.10.0.37:42634,
obj I-xD02008989187f0010-xGb3582637-xU6744b3f20396696d: client_op=OP_WRITE,
vcap_range=0x5a000000:0x2000000, vop_range=0x5a400000:0x400000, handle=0x26088034b920d5fa,
action=0xbf400000
pan_mgr[796]: sm: 0x02008989187f0010(SM): rioe [1269490094:732398995]: comps=0x6008d07187f0007,
op_range=0x3780000:0x80000, cap_range=0x3700000:0x400000, osd_op=OP_WRITE (0x7e9 (RPC: No
response within the timeout period))
```

As shown above, each error produces a stanza of two messages. The first reports the client address and the operation that failed. The second line shows which data components (comps) failed. Each comp ID includes the data serial number of the Storage Node involved.

9.3.3 /var/log/sysstat.log

This log only applies to Director Nodes. Every hour, the mserver node records the output of sysstat storage to disk, as shown in the following partial output:

```
OSD IP CPU Disk Op/s Resp
Address % % msec
east-asu-1,1 10.64.8.61 1 2 28 1.76 ...
east-asu-1,2 10.64.8.228 2 2 31 1.20 ...
east-asu-1,3 10.64.8.93 1 2 29 1.42 ...
east-asu-1,4 10.64.8.88 1 1 29 123.48 ...
east-asu-2,1 10.64.8.90 1 2 30 1.47 ...
east-asu-2,2 10.64.8.226 0 0 12 0.27 ...
east-asu-2,3 10.64.8.27 1 1 30 1.22 ...
east-asu-2,4 10.64.8.212 1 1 21 1.48 ...
east-asu-3,1 10.64.8.203 1 1 30 1.40 ...
east-asu-3,2 10.64.8.6 1 2 25 1.25 ...
east-asu-3,3 10.64.8.201 1 2 28 1.21 ...
east-asu-3,4 10.64.8.4 1 2 28 1.36 ...
test Total 1 1 321 1.27 ...
```

The sixth column is the latency of the node. In many cases, latency can increase due to disk utilization. However, if the disk and CPU utilization are low and the latency is high, the problem is likely network-related. In the example above, east-asu-1,4 10.64.8.88 has high node latency, but the third and fourth column all line up with the other nodes. This indicates that the latency discrepancy is from the network stack. If the fourth column also had an outlier number, it would indicate that the latency is coming from the disk.

9.4 Diagnostic Utilities

There are a number of network diagnostics that can be run on the system from PanCLI. Refer to the VDURA PanCLI Reference Guide for more information on each command.

netperf

Both a netperf server and client are available on every node of the VDURA system. By default, the server is disabled. When enabled, it operates on port 4005 and can be accessed from any standard netperf client:

```
[pancli_service] netperf enable
Successfully enabled netperf server on port 4005.
```

A pre-compiled netperf client is included in the panfs-apps RPM that can be downloaded and installed from the support.VDURA.com portal.

Netperf is a network bandwidth test. It shows the network transfer rate between two end points. It is a memory-to-memory transfer, so there is no disk I/O to affect the results:

```
~$: netperf -p 4005 -H 192.168.1.100
TCP STREAM TEST from 0.0.0.0 (0.0.0.0) port 0 AF_INET to 192.168.1.100 (192.168.1.100) port 0
AF_INET
Recv Send Send
Socket Socket Message Elapsed
Size Size Size Time Throughput
bytes bytes bytes secs. 10^6 bits/sec
524288 16384 16384 10.00 9387.41
```

It is also possible to run the client from a VDURA node:

```
[pancli_service] diag_exec "pan_netperf -p 4005 -H 192.168.1.100"
Tue Oct 27 00:31:10 GMT 2026: Executing "pan_netperf -p 4005 -H 192.168.1.100"...
Tue Oct 27 00:31:20 GMT 2026: Executed "pan_netperf -p 4005 -H 192.168.1.100" successfully:
TCP STREAM TEST from (null) (0.0.0.0) port 0 AF_INET to (null) (192.168.1.100) port 0 AF_INET
Recv Send Send
Socket Socket Message Elapsed
Size Size Size Time Throughput
bytes bytes bytes secs. 10^6bits/sec
524288 524288 524288 10.01 7921.65
```

pan_net_enclosure_nettest

Pan_net_enclosure_nettest is a service mode command for performing basic network connectivity and throughput tests between nodes and between nodes and external hosts.

Network connectivity tests include small and medium pings. Optional tests include a large ping test, ssh, and netperf throughput tests if netperf is enabled on the realm. The large ping test runs by default only if the system MTU is 9000. Test results are output for each node tested and summarized by enclosure.

If pan_net_enclosure_nettest is run with any of the following options, the results are summarized by enclosure in the summary log:

```
--all --encl --directors --storage --dms --sms
```

Results for each node are shown in the detail log.

If pan_net_enclosure_nettest is run either with the –repset option or with IP addresses, the results for those nodes are shown in the summary log.

Pan_net_enclosure_nettest prints the path and names of both logs to the screen when it completes.

The small (56 bytes) and medium (1472 bytes) ping tests are always run and may not be skipped. The large (8972 bytes) ping test is run if the realm is configured for MTU 9000. The large ping test may optionally be skipped via the –nolargeping option.

The netperf test requires that netperf be enabled on the realm. To enable netperf on the realm:

```
[pancli_service] netperf enable
```

Netperf may optionally be disabled once pan_net_enclosure_nettest has completed with the following:

```
[pancli_service] netperf disable
```

Netserver must be installed on any hosts whose IP addresses are supplied on the command line if those hosts are not nodes in the realm. Netserver must be configured to listen on port 4005.

Netperf results for the local enclosure will be unusually high. The test is executed across the enclosure backplane.

Command Examples

Run all tests to all nodes in the realm:

```
[pancli] support servicemode yes
[pancli_service] pan_net_enclosure_nettest --all
```

If netperf is not enabled on the realm, the command stops and asks you to enable it. Enter netperf enable at the pancli_service prompt, or add the –noperf option.

The output below is an example of the console output and summary from the command shown above, run on a realm with three Director Nodes, each in its own enclosure. The ping tests pass if the number of packets sent are received. Large ping shows N/A because the system MTU is not 9000. The ssh test passes if the remote command uptime returns in five seconds. The netperf results are 10^6bits/sec. The netperf results for the enclosures are presented as the lowest, highest, and average result for a Storage (S:) or Director (D:) Node in the enclosure.

```
pan_net_enclosure_nettest is using options: --all

pan_net_enclosure_nettest is running on node 10.128.1.0
   Host              Ping Small  Ping Med    Ping Large  SSH     Netperf
   10.128.1.0        OK          OK          N/A         OK      30849.00
   10.128.1.1        OK          OK          N/A         OK      7902.04
   10.128.1.2        OK          OK          N/A         OK      7888.72

By Enclosure:
   Enclosure         Ping Small  Ping Med    Ping Large  SSH     Netperf (Low/High/Avg)
   VCH-1*            1/1         1/1         N/A         1/1     D: 30849.00/30849.00/30849.00
   VCH-4             1/1         1/1         N/A         1/1     D: 7888.72/7888.72/7888.72
   VCH-8             1/1         1/1         N/A         1/1     D: 7902.04/7902.04/7902.04

*Test results to enclosure VCH-1 may be misleadingly high since
 pan_net_enclosure_nettest was run on 10.128.1.0, which is in enclosure VCH-1.

Tue Oct 27 00:25:53 2026 GMT: Elapsed time: 0 hours, 0 minutes, 56 seconds.
pan_net_enclosure_nettest has completed.
Results are logged to /tmp/pan_net_enclosure_nettest-1793060697-summary.log
Detailed results per node are logged to /tmp/pan_net_enclosure_nettest-1793060697-details.log
```

Test results to enclosure VCH-1 may be misleadingly high since pan_net_enclosure_nettest was run on 10.128.1.0, which is in enclosure VCH-1. Repeat pan_net_enclosure_nettest from a node in another enclosure to see a test through the network for 10.128.1.0.

Run all tests to the repset Directors:

```
[pancli] support servicemode yes
[pancli_service] pan_net_enclosure_nettest --repset
```

The output below is an example from the command, run on the same realm. The node running the command is itself a repset member, so its own results are marked with an asterisk and may be misleadingly high. The results to the repset members are shown both per node and by enclosure.

```
pan_net_enclosure_nettest is using options: --repset

pan_net_enclosure_nettest is running on node 10.128.1.0
   Host              Ping Small  Ping Med    Ping Large  SSH     Netperf
   10.128.1.0        OK          OK          N/A         OK      30157.35
   10.128.1.1        OK          OK          N/A         OK      7917.49
   10.128.1.2        OK          OK          N/A         OK      7912.62

Repset:
   Host              Ping Small  Ping Med    Ping Large  SSH     Netperf
   10.128.1.0*       OK          OK          N/A         OK      30157.35
   10.128.1.2        OK          OK          N/A         OK      7912.62
   10.128.1.1        OK          OK          N/A         OK      7917.49

By Enclosure:
   Enclosure         Ping Small  Ping Med    Ping Large  SSH     Netperf (Low/High/Avg)
   VCH-1*            1/1         1/1         N/A         1/1     D: 30157.35/30157.35/30157.35
   VCH-4             1/1         1/1         N/A         1/1     D: 7912.62/7912.62/7912.62
   VCH-8             1/1         1/1         N/A         1/1     D: 7917.49/7917.49/7917.49

*Test results to node 10.128.1.0 may be misleadingly high since pan_net_enclosure_nettest was
run on 10.128.1.0.
*Test results to enclosure VCH-1 may be misleadingly high since
 pan_net_enclosure_nettest was run on 10.128.1.0, which is in enclosure VCH-1.

Tue Oct 27 00:28:25 2026 GMT: Elapsed time: 0 hours, 0 minutes, 55 seconds.
pan_net_enclosure_nettest has completed.
Results are logged to /tmp/pan_net_enclosure_nettest-1793060850-summary.log
Detailed results per node are logged to /tmp/pan_net_enclosure_nettest-1793060850-details.log
```

Run all tests to all Storage Nodes in the realm, in addition to the client hosts 10.70.15.105 and 10.70.15.127.

```
[pancli] support servicemode yes
[pancli_service] pan_net_enclosure_nettest --storage 10.70.15.105 10.70.15.127
```

Run all tests to all nodes in enclosures VCH-1 and VCH-4 without the netperf test:

```
[pancli] support servicemode yes
[pancli_service] pan_net_enclosure_nettest --encl VCH-1 --encl VCH-4 --noperf
```

Run all tests to all Storage Nodes in StorageSet Set 2:

```
[pancli] support servicemode yes
[pancli_service] pan_net_enclosure_nettest --set 'Set 2'
```

Run a quick connectivity test to find out which enclosures are up:

```
[pancli] support servicemode yes
[pancli_service] pan_net_enclosure_nettest --all --nossh
--noperf --nolargeping --quick
```

Test only node 10.65.23.106. 10.65.23.106 may either be a node in the realm or an external host. No ssh test will be run:

```
[pancli] support servicemode yes
[pancli_service] pan_net_enclosure_nettest 10.65.23.106
```

Below is a list of all command options. At least one option specifying a group of nodes to test is required.

Option	Description
`--help`	Display this usage message.
`--examples`	Display explanation of pan_net_enclosure_nettest use and example command options.

Option	Description
`--all`	Test all nodes in the realm.
`--encl [name]`	Test all nodes in the specified enclosure. Multiple –encl arguments may be used.
`--repset`	Test the repset member Directors.
`--directors`	Test all Director Nodes in the realm.
`--storage`	Test all Storage Nodes in the realm.
`--dms`	Test all Director modules in the realm.
`--sms`	Test all Storage modules in the realm.
`--set`	Test all Storage Nodes in the specified StorageSet. Use single quote marks to enclose a StorageSet name that contains a space, as in 'Set 1'.
`--offline`	Include any nodes that are not marked online in the tests. By default, non-repset member nodes that are not online are excluded. Hosts specified on the command line are assumed to be online.
`--quick`	Run shorter ping and netperf tests. The number of pings is reduced from five to one for each of the small, medium, and large ping tests. The netperf test is reduced from ten seconds to one second.
`--nolargeping`	Skip the large ping test.
`--nossh`	Skip the ssh test. This test is automatically skipped for external hosts.
`--noperf`	Skips the netperf test.

Option	Description
`--noquery`	Do not query the realm to build a list of nodes to test. A space-separated list of node IP addresses to test must be provided. If –noquery is used, all options that specify types of nodes to test are disregarded. The large ping and netperf tests are run without querying the realm to see if netperf is enabled or MTU is set to 9000. The ssh test is not run. This option is useful if the repset is down.

ping

A simple ICMP ping to test connectivity can be issued from PanCLI. By default it sends three 64-byte ICMP pings:

```
[pancli_service] ping compute1
PING panasas1.example.com (172.17.2.234): 56 data bytes
64 bytes from 172.17.2.234: icmp_seq=0 ttl=61 time=1.454 ms
64 bytes from 172.17.2.234: icmp_seq=1 ttl=61 time=0.818 ms
64 bytes from 172.17.2.234: icmp_seq=2 ttl=61 time=0.813 ms
--- compute1.example.com ping statistics ---
3 packets transmitted, 3 packets received, 0.0% packet loss
round-trip min/avg/max/stddev = 0.813/1.028/1.454/0.301 ms
```

Additional ping options (for example, to determine if the MTU is correct) must be executed through diag_exec:

```
[pancli_service] diag_exec "ping -s 9000 -c 3 compute1"
Tue Jul 19 16:44:53 GMT 2015: Executing "ping -s 9000 -c 3 compute1"...
Tue Jul 19 16:44:55 GMT 2015: Executed "ping -s 9000 -c 3 compute1" successfully:
PING compute1.example.com (172.17.2.234): 9000 data bytes
9008 bytes from 172.17.2.234: icmp_seq=0 ttl=61 time=7.729 ms
9008 bytes from 172.17.2.234: icmp_seq=1 ttl=61 time=8.991 ms
9008 bytes from 172.17.2.234: icmp_seq=2 ttl=61 time=8.972 ms
--- compute1.example.com ping statistics ---
3 packets transmitted, 3 packets received, 0.0% packet loss
round-trip min/avg/max/stddev = 7.729/8.564/8.991/0.590 ms
```

tcpdump

Tcpdump is a standard UNIX/Linux utility used to capture raw packets from the network interface. It can be used to diagnose anything from simple connectivity problems to complex network protocol interactions.

Tcpdump is available directly from PanCLI. It can be attached to any interface including the LAG interfaces. By default, it will collect 100,000 packets and exit. That limit can be modified using the -c option.

Below is a simple command to monitor all DHCP traffic on a node:

```
[pancli_service] tcpdump port 9876 or port 9877
```

In the next example, the binary output is saved to a file. When invoked in this way, the system will prompt you to upload the file after the dump is terminated:

```
[pancli_service] tcpdump -w /tmp/trace.bin -s 0 host 172.17.2.234
This tcpdump will stop after collecting 100000 packets.
Use "-c count " option to specify a different packet count.
tcpdump: WARNING: bge0: no IPv4 address assigned
tcpdump: listening on bge0, link-type EN10MB (Ethernet), capture size 65535 bytes
```

traceroute

Traceroute can also be issued directly from PanCLI:

```
[pancli_service] traceroute compute1
traceroute to compute1.example.com (172.17.2.234), 64 hops max, 52 byte packets
1 10.72.3.1 (10.72.3.1) 0.690 ms 0.517 ms 0.519 ms
2 * ROUTER-0.lab.example.com (10.64.0.254) 103.784 ms *
3 compute1.example.com (172.17.2.234) 0.869 ms 0.518 ms 0.513 ms
```

Traceroute is commonly used to determine the actual path a packet will travel in a routed network.

9.5 Configure the Maximum Transmission Unit

Maximum Transmission Unit (MTU) in bridged environments can cause some confusion. The VDURA system supports both 1500- and 9000-byte IP MTUs for Ethernet and 2048- and 4096-byte IP MTUs for InfiniBand. Like MTU settings on hosts, this is the maximum size of the IP packet that will be generated from the host. However, once the packet is sent to the Ethernet interface, it becomes larger due to Ethernet headers encapsulating the packet. If VLAN tagging will be used on the LAN, you must also account for the 4-byte p/q tags.

Let's use a 9000-byte IP packet as an example. To pass a 9000-byte IP packet, the Ethernet infrastructure must be able to pass a larger frame. Many newer switches allow you to specify the exact number of bytes per Ethernet frame. In these cases, a common misconfiguration is to specify 9000. If this is done, 9000-byte IP packets will be lost as they appear larger when encapsulated into Ethernet frames.

If the LAN infrastructure can support the largest MTU that can be generated, it is not necessary to make sure both the clients' and storage servers' MTU match. Part of the TCP connection process involves negotiating the segment size. This allows a 1500-MTU client to properly communicate to a 9000-MTU server. This can be done simultaneously, with a 9000-MTU client in the mix. The key is to make sure the infrastructure in between can handle the largest frame that will pass.

In the case of routed networks, there is the possibility of two LANs with different MTUs. When a packet arrives on a router and its destination is a LAN with a smaller MTU, the IP packet is split to accommodate the destination LAN. Fragmentation of packets can significantly impact performance and should be avoided at all costs.

To make sure the MTU settings are optimal, realms configured with an MTU of 9000 should have all switches configured with an MTU of at least 9018/9022 and routers must have an IP MTU of at least 9000. For realms configured with an MTU of 1500, all switches must have an MTU of at least 1518/1522

and routers must have an IP MTU of at least 1500.

A good test to ensure that MTU is sufficient from end to end is to use the following options on the ping command:

```
ping -M do -s 8972 <IP address>
```

The -M do tells ping to not allow the packets to fragment, and the -s specifies the size of packet to send. If a simple ping succeeds, but this format fails, there is likely some component in the path that is configured with an MTU too low to pass the packet without fragmenting.

To run a jumbo ping from one of the VDURA nodes, use the following command:

```
diag_exec "ping -D -n -s 8972 -c 4 <IP address>"
```

Use the Correct MTU Values

MTU values are commonly misconfigured. Determining the maximum value may be difficult, but the user interface may provide a valid range. Ethernet adds 18 bytes or 22 bytes with an 802.1q VLAN tag, so a switch MTU of at least 9018 or 9022 is required for 9000-byte IP packets. Larger frames might be necessary for support of MPLS, so switches will usually support larger MTU sizes needed to support those larger frames.

The most important issue is that the layer-2 switch MTUs are set higher than the endpoints which are injecting traffic into the network. From the perspective of the VDURA storage system, it is always safe to set the switch MTUs to the largest values supported on the individual switch, even though that might not be exactly the same as the other switches in the same broadcast domain, as long as that value is at least 9018.

If the switch infrastructure is simultaneously being used to transport other traffic which uses larger frame sizes, then the individual switches need to support those MTUs. Note that an MTU mismatch will lead to problems which manifest as either error packets or missing packets, depending on the switch. Store-and-forward architectures will discard the bad frame (incrementing appropriate counters), whereas a cut-through architecture will propagate the bad frame (also incrementing the appropriate counters), ultimately resulting in the end station also incrementing an error packet counter.

Appendix. Installation and Configuration Checklist

An * denotes a required field.

Field	Description	Value
License Agreement		
Customer ID *	Specify the customer number provided by VDURA in your order confirmation. Call 1-888-726-2727 if you do not have your customer ID.	
Administrator Configuration		
Administrator ID	Specify the username for the administrator of the VDURA appliance. The default Administrator ID is admin.	
Administrator Password *	Specify the password for the administrator of the VDURA appliance. VDURA recommends the administrator password contain at least four characters, one digit or special character, and one letter.	

Field	Description	Value
Enable PanActive Link *	Activate VDURA PanActive Link, which provides software updates, up-to-date system information, and error reports. VDURA strongly recommends that you enable PanActive Link. Note: These HTTPS features are currently not available; the system cannot download from VDURA. Download software from support.VDURA.com instead.	
SMTP Server	Specify the name or IP address of the mail server. The system uses Simple Mail Transport Protocol (SMTP) to send mail about problems it encounters. Example: mail.yournetwork.com	
Administrator Email	Specify the email address for the VDURA system administrator, including the full domain name (for example: joe@yournetwork.com).	

Field	Description	Value
Secure Web Proxy Server	Specify the name of the secure web proxy server that connects to the my.VDURA.com support site. This proxy server will be used to make outbound HTTPS requests to VDURA servers if you have chosen to enable PanActive Link. Example: storage.yourhttphost.com	
Secure Web Proxy Port	Specify the secure web proxy port number. Example: 8080	
IPMI Access	Specify a range of IP addresses to be used for the IPMI out-of-band management configuration.	
Name and IP Configuration		
VDURA System DNS Name *	Specify a unique network name for the VDURA system. It is best practice to have this be the DNS name associated with the IP address of the system. Example: West01 Note: If the DNS zone name is not equivalent to “system DNS name + DNS domain name”, then the DNS lookups to the DNS zone name fail with NXDOMAIN.	

Field	Description	Value
System IP Address *	Specify the IP address for the system. This is the IP address you configured for the first Director Node when you first configured your system. Example: 10.10.10.2	
Netmask *	Specifies the subnet mask for the network to which the VDURA appliance is attached. This is the subnet mask that you entered when you configured the first VDURA Director Node. Example: 255.255.255.0	
Default Router	Specify the name or IP address of the primary routing gateway. If the system does not need routing to reach its clients, leave this field blank. This is the default router that you entered when you configured the first VDURA Director Node. Example: 10.10.10.1	
MTU Size *	If your subnet supports jumbo frames (Ethernet Frame Size=9000), specify that size.	

Field	Description	Value
Node Address Range: Start *	Specify the start of a range of IP addresses for the VDURA system. VDURA recommends that you assign the VDURA system its own IP subnet with at least as many available addresses as the maximum predicted number of nodes in the storage system. The start address is the first address in a range of IP addresses that you preassign for the VDURA storage system. By default, the start address is the address that you entered when you configured the first Director Node.	
Node Address Range: End *	Specify the end of a range of IP addresses for the VDURA system. The end address indicates the last address in a block of IP addresses that you preassign for the VDURA system.	
Calculate Number of Nodes	To see the number of IP addresses that can be assigned to VDURA nodes in the range you just specified, click the Calculate Number of Nodes button.	
Network Services Configuration		

Field	Description	Value
DNS Domain Name	Specify the Domain Name Service (DNS) domain name. Example: yournetwork.com	
Primary DNS Server IP Address	Specify the IP address for the DNS server that the system uses for host name lookups. Although the DNS server is not required, its absence will prevent the use of hostnames for NTP, SMTP, NIS, and secure web proxy servers as well as DNS delegation. Example: 10.10.10.1. You can add additional DNS servers in the PanActive Manager.	
NIS Domain Name	Specify the Network Information Service (NIS) name. If your site does not use NIS, leave this field blank. Example: yournetwork.com	
NIS Server Name	Specify an NIS server name or IP address that the system binds to. You can add additional NIS servers in PanActive Manager. Example: nis.yourserver.com.	
NIS Hostname Resolution	Choose to enable or disable NIS hostname resolution.	
NTP Servers *	Specify the names or IP addresses of the NTP servers. A list of NTP servers is available at tf.nist.gov/tf-cgi/servers.cgi.	

Field	Description	Value
Time Zone *	Specify your time zone.	
Storage		
Enable MLEC	Enable or disable Multi-Level Erasure Coding feature for the given StorageSet. This can only be set when a StorageSet is created and cannot be changed after. MLEC is disabled by default.	
Kubernetes (CSI Driver) - optional		
Linux Kernel Versions	If you use the PanFS CSI driver with Kubernetes, list the Linux kernel versions that your Kubernetes nodes run. To collect them, run `kubectl get nodes -o wide`. VDURA uses them to build the DirectFlow Client/KMM images for your kernels.	

User Authentication Services - optional		
Select zero, one, two, or three of the following options for identifying users on the VDURA storage and provide required details.		
NIS		
Enable NIS:	Yes/No	
NIS Domain:		

User Authentication Services - optional		
Domain Controller(s):	IP address or DNS name	
Use NIS for hostname resolution:	Yes/No	
LDAP		
Enable LDAP:	Yes/No	
LDAP Domain:		
LDAP Server:	IP address or DNS name	
Certificate Available?		
LDAP Account Name:	e.g., cn=[username], [baseDN]	
Password for account available for install:	Yes/No	