Debug and RAS Guide for NVIDIA Data Center Products

Application Note

Document History

DA-11437-001_v13

Version	Date	Authors	Description of Change
01	October 31, 2023	DP, SM	Initial release – Preliminary information
02	February 21, 2024	DP, SM	> Added error code to Table 1 > Added “SXID Error Codes” chapter > Added error codes to Table 3 > Updated “Partner Diagnostics Error Codes” chapter > Updated “Extended Utility Diagnostic Error Codes” chapter > Updated “Unable to Reach Maximum GPU Clocks” section
03	May 13, 2024	DP, SM	> Updated Table 4 with error codes > Added Section “NVIDIA Grace Partner Diagnostics Error Code to Action Mapping” > Added Section “Troubleshooting Steps”
04	May 20, 2024	DP, SM	Updated “Next Steps” in Table 4
05	May 29, 2024	DP, SM	> Updated error codes in Table 3 > Updated actions in “Troubleshooting Actions for CPU Diagnostic Error Codes”
06	July 9, 2024	DP, SM	Added networking platforms
07	October 29, 2024	DP, AB, SM	> Added Section “NVIDIA GB200 NVL Related NVLink MODS Error Code to Action Mapping” > Added “NVLink5 and Link Quality Monitoring” chapter
08	December 2, 2024	AB, SM	> Updated PCIe Pass or Fail criteria > Updated Chapter 10 “NVLink5 Diagnostics and Link Quality Monitoring” > Added the following sections: • “Ongoing NVLink5 Monitoring Validations” • “GB200 NVL Systems Maintenance Flow Overview” • “GB200 NVL Replacement Steps” • “NVLink5 Data Center Acceptance Validation”
09	January 24, 2025	AB, KP, SM	Updated sections on GPU EUD and GPU FD error codes to action mapping

Version	Date	Authors	Description of Change
10	February 18, 2025	AB, MA, GL, SM	> Updated Table 1-2: Related Documents > Updated Chapter 3 SXID Error Codes > Chapter 4: Updated sections on L10 Compute Diagnostics and L11 Partner Diagnostics Error Code to Action Mapping > Updated Section 5.2: CPU EUD Related Error Code to Action Mapping > Updated Chapter 10: NVLink5 Diagnostics and Link Quality Monitoring > Updated Section 10.3: FEC Histogram Bins • Updated Section 10.7: Thresholds and Pass or Fail Criteria • Added Section 10.13 NVLink5 Debug GB200 NVL NVLink Mapping Tools > Added Section 14.2: NVDebug for Datacenter Products
11	April 18, 2025	AB, SM	> Added Section 4.6 “NVIDIA CPU Related Common Error Code to Action Mapping” > Updated Section 4.8 “NVIDIA CPU Related Error Codes to Action Mapping” > Updated Chapter 10 “NVLink5 Diagnostics and Link Quality Monitoring” > Updated troubleshooting for Ethernet in Table 20-8
12	June 11, 2025	ZH, AB, SM	> Updated Chapter 10 "NVLink5 Diagnostics and Link Quality Monitoring" > Updated Section 4.2 "L10 Compute Field Diagnostics Error Code to Action Mapping" > Updated Section 4.3 "L11 Rack Field Diagnostics Error Code to Action Mapping" > Updated Section 4.6 "NVIDIA CPU Related Common Error Code to Action Mapping" > Added Chapter 22 "Telemetry in Multi-node NVLink Systems > Document restructuring to differentiate debug for Switch devices in HGX-8 GPU Systems and NVSwitches in Multi-node NVLink Systems
13	September 10, 2025	AB, SM	> Removed sections applicable to "MultiNode NVLink Troubleshooting" > Added references to "Multi Node NVLink Troubleshooting Guide" > Added Error Codes to GPU FD, L10 FD, L11 FD and CPU FD Error Codes to Action Mapping Tables. > Updated Chapter 2 "XID Error Codes" > Updated Chapter 3 "SXID Error Codes" > Updated Section 20.5 "Link Status Opcodes"

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

This debug and reliability, availability, and serviceability (RAS) guide for data center products is intended to assist partners and customers in analyzing and triaging data center product-related hardware and software issues.

This application note describes different error scenarios that may be encountered when using the NVIDIA data center graphics processing units (GPUs), NVIDIA® NVSwitch™ devices, NVIDIA MGX™, NVIDIA DGX™ and NVIDIA HGX™ baseboard products. They include XID, SXID, Field Diagnostics, NVIDIA® NVLink™, and PCIe related errors.

XID message is an error message from the NVIDIA driver that is printed to the operating system's (OS) kernel log or event log. XID messages indicate that a general GPU error occurred. The messages can be indicative of a hardware problem, an NVIDIA software problem, or a user application problem.

SXID message is an error message from the NVIDIA driver for NVSwitch related error conditions.

For detailed information on XID and SXID messages, refer to the "GPU Management and Deployment" website at: https://docs.nvidia.com/deploy/xid-errors/index.html

This application note also focuses on different networking platform products such as NVIDIA® BlueField®-3 (BF-3) and NVIDIA® CONNECTX®-7 (CX-7), along with shared interfaces across all platforms including PCIe, NVLink, InfiniBand, and Ethernet. Every section presents relevant commands and debugging procedures to address various error scenarios.

Note: For troubleshooting guidance related to GB NVL Multi Node NVLink Systems, refer to the *Multi Node NVLink Troubleshooting Guide* (NVOnline: 1138824).

1.1 Terminology

The following table lists the key terms and their descriptions used within this application note.

Table 1-1. Terminology

Term	Description
AER	Advanced Error Reporting
AN	Auto Negotiation
BDF	Bus Device Function
BMC	Baseboard Management Controller
CEC	Consumer Electronics Control
CPO	Co Package Optics
DDR	Double Data Rate memory
DLL	Data Link Layer
ELS	External Laser Source
ERoT	External Root-of-Trust
EUD	Extended Utility Diagnostic
FEC	Forward Error Correction
FW	Firmware
LACC	Linear Active Copper Cable
MDIO	Management Data Input/Output
NIC	Network Interface Card
NMI	Non-Maskable Interrupt
OOB	Out-of-band
OS	Operating System
PCIe	PCI Express
PLL signals	Phase-Locked Loop Signals
PLR	Physical Layer Retransmission
PSID	Parameter Set Identification
RAS	Reliability, Availability, and Serviceability
RMA	Returned Material
RMA	Return Merchandise Authorization
TLP	Transmission line pulse
VL	Virtual Lanes
WJH	What-Just-Happened
NVSwitch Tray	A physical tray containing the NVSwitch ASICs to which the GPU NVLink devices connect.
Compute Tray	A physical tray that is inserted into a slot in the rack and has CPUs and GPUs.
FM	Fabric Manager - NVLink network control plane service is provided by the FM service.

Term	Description
GFM	Global Fabric Manager An instance of FM with a specific set of features enabled. There is one GFM per NVLink domain (cluster).
NVLSM	NVLink Subnet Manager A service that originates from NVIDIA InfiniBand switches and has the modifications to effectively manage NVSwitches.
Access NVLink	An NVLink between a GPU and an NVSwitch
Trunk NVLink	An NVLink between NVSwitches
NVOS	NVIDIA Networking OS, which was previously known as MLNX-OS. NVOS is used as the Switch OS for L1 NVSwitch Trays.
XID	The XID message is an error report from the NVIDIA driver that is printed to the operating system's kernel log or event log. XID messages indicate that a general GPU error occurred.

1.2 Related Documents

The following table lists the documents that are relevant to this application note.

Table 1-2. Related Documents

NVOnline ID	Document Title or Link
1113192	Collect eMMC Lifetime Tool
1115699	XID and SXID Catalog for NVIDIA Data Center Products
1092300	NVIDIA Data Center Products Telemetry Catalog
1120466	Partner Diagnostics User Guide
1091416	Data Center Partner Diagnostics Playbook
1123468	NVIDIA GB200 NVL Service Flow User Guide
1114436	NVIDIA NVOS User Manual
1001187	SMBus Post Box Interface (SMBPBI) For GPUs
1093241	NVIDIA Hopper HGX 8-GPU Telemetry Playbook
1106532	NVOnline Partner Portal Guidelines
1117886	GB200 NVL72 Rack Level Integration and Partner Guidelines
1118938	GB200 NVL72 Rack System Build and Assembly Instruction Guide
1116117	NVIDIA Server RAS Catalog
1115302	NVIDIA Grace CPER Catalog
1109504	NVIDIA Debug Tool for Datacenter Products (NVDebug)
1124813	GB200 NVL NVLink Mapping Tool
1126922	GB200 NVL Mapper Diagnostic Visualization Tool
1138824	Multi Node NVLink Troubleshooting Guide

NVOnline ID	Document Title or Link
Not applicable	NVIDIA Cumulus Linux User Guide
Not applicable	NVIDIA Firmware Tools (MFT) Documentation
Not applicable	NVIDIA Adapter Cards Firmware Release Notes

Chapter 2. XID Error Codes

When the GPU driver is run, it displays any XID errors in the dmesg log. GPU error events cause the driver to display XID error codes and GPU information in the kernel log.

Error codes provide insight into the cause of the failure and could be used to infer steps required to address the failure.

XID error codes can also appear within Virtual Machines (VMs). XID reported within VMs can be caused by configuration issues in the VM creation and does not necessarily refer to real hardware faults.

For the latest GPU XID error code to action mapping, refer to the *NVIDIA Server RAS Catalog* (NVOnline: 1116117).

Chapter 3. SXID Error Codes

When an NVSwitch (based on NVLink4 or older generations of NVLink) port generates a runtime error, the error report from NVSwitch will be logged with the SXID error code.

Note: SXID error codes are not applicable to NVLink5 NVSwitch devices.

Error codes provide insight into the cause of the failure and could be used to infer steps required to address the failure.

For more information, refer to the *NVIDIA Server RAS Catalog* (NVOnline: 1116117).

Chapter 4. Partner Diagnostics Error Codes and Actions

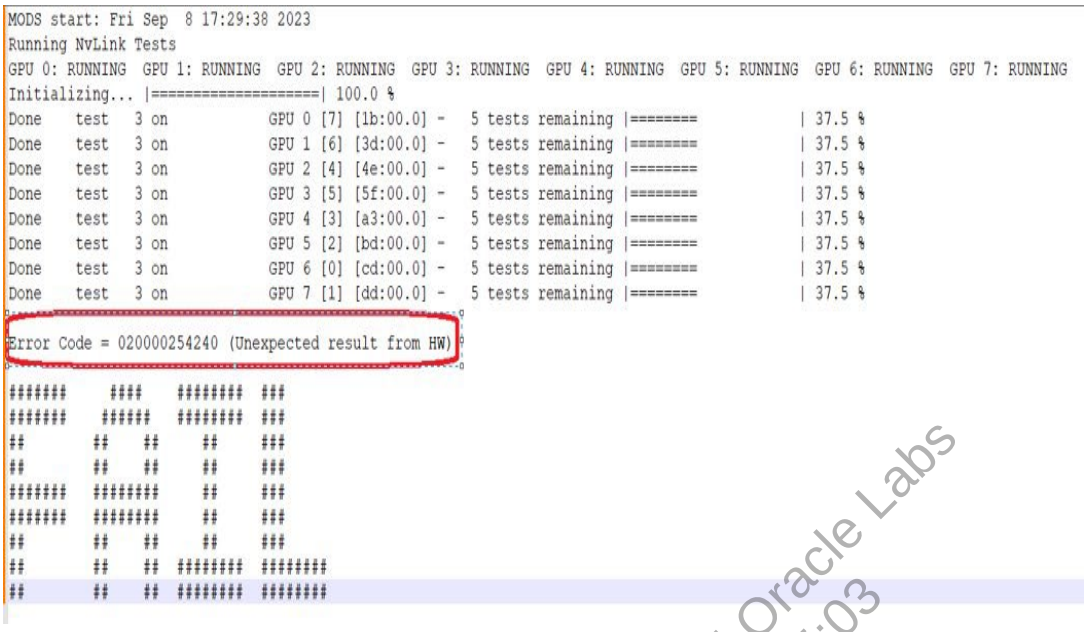
When the NVIDIA Partner Diagnostics is run, it will show the results of the run on the console and create a tarball containing logs from the run. The console output will be saved in a file named ‘run.log’ inside the tarball. Faulty systems will cause the diagnostic to fail, and it will display ‘FAIL’ along with error codes and error messages for the failing tests on the console and at the end of the run.log file.

Error codes provide insight into the cause of the failure and could be used to infer steps required to address the failure. Refer to the following sections to learn more about the common measures to address the failure per error code.

- Section 4.1 “GPU Related Error Code to Action Mapping”
- Section 4.2 “L10 Compute Field Diagnostics Error Code to Action Mapping”
- Section 4.3 “L11 Rack Field Diagnostics Error Code to Action Mapping”
- Section 4.4 “HGX Baseboard Diagnostics Error Code to Action Mapping”
- Section 4.7 “CPU Field Diagnostics Error Code to Action Mapping”

Note: Error Codes to Action Mapping tables presented in this Chapter include a dedicated column that summarizes the Data Center Actions. Steps itemized in this column may be different from the details under the Next Steps column. In some cases, the Next Steps column assumes the Integrator who can perform low level device/component triaging and debug, which may not be possible in the Data Center.

Figure 4-1. GPU Field Diagnostics Error Code



4.1 GPU Field Diagnostics Error Code to Action Mapping

For “GPU Field Diagnostic” related failures, look up the error code in Table 4-2 and follow the “Next Steps” column to address the failing system. The “Data Center Action Category” corresponding to each error code is also included with the corresponding action description in the following table.

Table 4-1. Data Center Action to Data Center Description Mapping – GPU FD

Data Center Action Category	Data Center Description
N/A	Not applicable.
PROD_FIT	System is production ready if the error is not seen again, for e.g. after re-running the test with latest FW and latest Diag.
PROD_FIT.RESET_DEVICE	Reset failing device aka hot reset failing device. The system is then production ready.
PROD_FIT.RESET_TRAY	Reset tray with the failing device aka hot reset failing device. The system is then production ready.
RECOVERY	Running additional tools can get the system back in production.
RECOVERY.RESET_BM	Restart bare metal aka cold reboot the system. Running additional tools can get the system back in production.
RECOVERY.RESET_TRAY	Restart tray aka cold reboot the entire tray. Running additional tools can get the system back in production.

Data Center Action Category	Data Center Description
RECOVERY.RESET_SW	Warm reboot the system aka reload latest supported installed firmware / driver. Running additional tools can get the system back in production.
RECOVERY.RUN_INFOROM	Run InfoROM recovery tool. Running additional tools can get the system back in production. For Blackwell generation of products, please refer to *InfoROM Cleansing for RMA - User Guide* (NVOnline: 1120261).
REPORT_NV_BUG	Report Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached.
RETURN_FOR_FURTHER_TRIAGE	Take the failing tray or board out of production and return the failing tray or board to the integrator for further triage and possible RMA.
RMA_DEVICE	For Hopper products, take the failing device out of production and start RMA qualification for failing device. For Blackwell products, take the failing board out of production and start RMA qualification for failing device. Note that RMA decisions must involve the integrator. Datacenter actions may involve returning the component to the integrator and may possibly translate to RMA following further testing and investigation done by the integrator.
TRIAGE_BOARD	Take the failing board out of production and follow next action steps for that error code. If qualified for RMA, RMA the board. To RMA the board, return the board to the integrator for further triage.
TRIAGE_DEVICE	For Hopper products, take the failing device out of production and follow next action steps for that error code. If qualified for RMA, RMA the device. For Blackwell products, take the failing board out of production and follow next action steps for that error code. If qualified for RMA, RMA the device.
TRIAGE_TRAY	Take the tray with the failing device out of production and follow next action steps for that error code. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The guide for RMA of the networking component(s) is RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).
ISOLATE_TRAY	Isolate the failing tray with the reported information. If it is unreported which tray is failing, swap and test each Compute and Switch Tray to confirm what component the problem follows.

Table 4-2. GPU Diagnostics Error Code to Action Mapping

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
MODS-xxxxxxxxx002	software error	An unexpected software error has occurred.	1. Re-test with the latest GPU FD if not tested with the latest version. 2. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx008	bad parameter passed to function	Software error / bad parameter passed to software function.	1. Re-test with the latest GPU FD if not tested with the latest version. 2. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx021	script failed to execute	Script failed to execute.	1. Re-test with the latest GPU FD if not tested with the latest version. 2. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx041	cannot hook interrupt	A software process or GPU state is preventing hooking interrupts. -poll_interrupts option can be used in case the system does not support MSI or MSI-X IRQs.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU FD if not tested with the latest GPU FD. 3. Test with latest GPU FD by passing 'poll_interrupts' command line option if the same error code is seen after the run. 4. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	1. PROD_FIT.RESET_DEVICE
MODS-xxxxxxxxx060	NVRM invalid param struct		1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU FD if not tested with the latest GPU FD. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	1. REPORT_NV_BUG 2. PROD_FIT.RESET_DEVICE
MODS-xxxxxxxxx077	timeout error	Tests failed with SW timeout issue and needs further investigation.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not tested with the latest version. 3. File Software NVBug to debug root	1. TRIAGE_DEVICE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	
MODS-xxxxxxxxx083	CRC/Checksum miscompare	Computation has returned an incorrect answer.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not tested with the latest version. 3. Start RMA Qualification if the same error code is seen after the GPU FD run.	1. RETURN_FOR_ FURTHER_TRIAGE
MODS-xxxxxxxxx097	unexpected device interrupts	Diag received an unexpected interrupt that it considers an error.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not tested with the latest version. 3. Start RMA Qualification if the same error code is seen after the GPU FD run.	1. RETURN_FOR_ FURTHER_TRIAGE
MODS-xxxxxxxxx124	Invalid InfoROM	InfoROM corruption has occurred. Recover InfoROM by running InfoROM recovery tool.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Run InfoROM recovery tool. For Blackwell generation of products, please refer to *InfoROM Cleansing for RMA - User Guide* (NVOnline: 1120261). 3. Re-test with the latest GPU FD if not tested with the latest version. 4. Start RMA Qualification if the same error code is seen after the GPU FD run.	1. TRIAGE_DEVICE
MODS-xxxxxxxxx127	Vbios Certificate Error	An issue with the VBIOS certificate.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU FD if not tested with the latest GPU FD. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	1. RECOVERY.RESET_SW 2. PROD_FIT
MODS-xxxxxxxxx139	Acceptable temperature limits exceeded, or the thermal sensor is broken or miscalibrated	Check cooling systems and engage with NVIDIA. Failing part qualifies for RMA only if the problem is isolated to NVIDIA hardware.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU FD if not tested with the latest GPU FD. 3. Start RMA qualification if the same error code is seen after the GPU FD run and only if the cooling solution is part of NVIDIA hardware.	1. TRIAGE_DEVICE
MODS-xxxxxxxxx140	NvLink bus error	NVLink is down or NVLink errors occurred during NVLink testing.	1. Reseat the failing GPU. 2. Ensure GPU devices are detected on the system using lspci utility. 3. Re-test with the latest GPU FD if not tested with the latest version.	1. RETURN_FOR_ FURTHER_TRIAGE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			4. Start RMA qualification if the same error code is seen after the GPU FD run.	
MODS-xxxxxxxxx143	PCI Express bus error	PCIe link errors cause PCIe tests to fail.	1. Reseat the failing GPU. 2. Ensure GPU devices are detected on the system using lspci utility. 3. Re-test with the latest GPU FD if not tested with the latest version. 4. Start RMA qualification if the same error code is seen after the GPU FD run.	1. RETURN_FOR_ FURTHER_TRIAGE
MODS-xxxxxxxxx144	CUDA error	One of the CUDA tests is failing. File Software NV bug and attach dmesg, diagnostic error logs, NV bug report tool log, to debug root cause.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not tested with the latest version. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	1. TRIAGE_DEVICE
MODS-xxxxxxxxx167	GFW boot reported a failure	GPU firmware failed to initialize	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Reseat failing GPU. 3. Perform cold reboot i.e., full system shutdown and manually reboot again. 4. Ensure GPU and NVSwitches are detected on the system using lspci utility. 5. Re-test with latest GPU FD if not tested with the latest GPU FD. 6. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the GPU FD run.	1. REPORT_NV_BUG 2. RECOVERY.RESET_BM 3. PROD_FIT
MODS-xxxxxxxxx194	bad memory	Computation has returned an incorrect answer.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not tested with the latest version. 3. Start RMA qualification if the same error code is seen after the GPU FD run.	1. RETURN_FOR_ FURTHER_TRIAGE
MODS-xxxxxxxxx220	pci device not found	Failure with PCIe device. Must isolate the failure to GPU or baseboard.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Reseat the failing GPU. 3. Ensure GPU and NVSwitch devices are detected on the system using lspci utility. 4. Perform cold reboot i.e., full system shutdown and manually reboot again. 5. Re-test with the latest GPU FD if not tested with the latest version.	1. RECOVERY.RESET_BM 2. PROD_FIT

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			6. Start RMA qualification if the same error code is seen after the GPU FD run.	
MODS-xxxxxxxxx229	hardware was not initialized	Something prevented successful hardware initialization.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU FD. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	1. TRIAGE_DEVICE
MODS-xxxxxxxxx240	Unexpected result from hardware	If running GPU FD on a single GPU, make sure the 'skip_nvlink' parameter is used. If it is, then there is a possible setup issue, confirm that NVLink topology is connected as expected by running: nvidia-smi nvlink -s If NVLinks are reported as inactive, debug NVLink issue.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not tested with the latest version. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	1. TRIAGE_DEVICE
MODS-xxxxxxxxx272	Read parameter differs from expected	Re-test with the latest GPU Field Diagnostics. (v22.06-12 and later)	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not tested with the latest version. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen, after the GPU FD run.	1. TRIAGE_DEVICE
MODS-xxxxxxxxx276	Hardware reports wrong status	Hardware failure has occurred.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not tested with the latest version. 3. Start RMA qualification if the same error code is seen after the GPU FD run.	1. RETURN_FOR_ FURTHER_TRIAGE
MODS-xxxxxxxxx280	Power is above specified limit	Power is too high	1. Re-test with latest GPU FD if not already tested with the latest GPU FD. 2. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
MODS-xxxxxxxxx288	Power is below specified limit	Power is too low	1. Re-test with latest GPU FD if not already tested with the latest GPU FD. 2. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx311	NvLink discovered topology does not match required topology	NVLink topology does not match expected topology.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Ensure GPU devices are detected on the system using lspci utility. 3. Perform cold reboot i.e., full system shutdown and manually reboot again. 4. Reseat failing GPU if reboot did not resolve the issue. 5. Re-install NVIDIA release driver if rebooting and reseating the GPU did not resolve the issue. 6. Re-test with latest GPU FD if not tested with the latest GPU FD. 7. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	1. TRIAGE_DEVICE
MODS-xxxxxxxxx317	ECC detected an uncorrectable error in FB	Uncorrectable error has occurred, reset for row remapping to take place	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Power cycle GPU to trigger row remapping. 3. Re-test with the latest GPU FD if not tested with the latest version. 4. Start RMA qualification if the same error code is seen after the GPU FD run.	1. TRIAGE_DEVICE
MODS-xxxxxxxxx319	ECC detected an uncorrectable error in L2	An uncorrectable error has occurred in the L2 cache, start RMA qualification for baseboard including the exact L2.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not tested with the latest version. 3. Start RMA qualification if the same error code is seen after the GPU FD run.	1. TRIAGE_DEVICE
MODS-xxxxxxxxx320	ECC detected a correctable error in SM, threshold exceeded	Rate of correctable errors in the SM have exceeded threshold, start RMA qualification for baseboard including the exact SXMs.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not tested with the latest version. 3. Start RMA Qualification if the same error code is seen after the GPU FD run.	1. TRIAGE_DEVICE
MODS-xxxxxxxxx321	ECC detected an uncorrectable error in SM	An uncorrectable error has occurred in the SM during testing, start	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not	1. TRIAGE_DEVICE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
		RMA qualification for baseboard including the exact SXMs.	tested with the latest version. 3. Start RMA qualification if the same error code is seen after the GPU FD run.	
MODS-xxxxxxxx332	Cannot pulse fast enough to accurately make a waveform	GPU is taking too long to either: (a) Complete the GEMM workload, or (b) Report that it completed the GEMM workload.	1. Reset the tray and retest. 2. File Software NVBug to debug root cause with dmesg, diagnostic error log, and NVIDIA Bug report tool log attached if the same error is seen after the GPU FD run.	1. RECOVERY.RESET_BM 2. REPORT_NV_BUG 3. PROD_FIT
MODS-xxxxxxxxx341	Buffer mismatch	Start RMA qualification for baseboard including the exact SXMs.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not tested with the latest version. 3. Start RMA qualification if the same error code is seen after the GPU FD run.	1. RETURN_FOR_ FURTHER_TRIAGE
MODS-xxxxxxxxx351	memory not strapped correctly	Need access to logs or failing system to debug root cause.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU FD if not tested with the latest GPU FD. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	1. TRIAGE_DEVICE
MODS-xxxxxxxxx363	Row remapping failed	A row remapping failure has occurred, start RMA qualification.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not tested with the latest version. 3. Start RMA qualification if the same error code is seen after the GPU FD run.	1. RETURN_FOR_ FURTHER_TRIAGE
MODS-xxxxxxxxx539	NVRM generic falcon error	Ensure the NV driver is fully unloaded, and other internal processes do not try to load the driver back while running GPU FD.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Ensure NVIDIA driver is fully unloaded. 3. Re-test with latest GPU FD if not tested with the latest GPU FD. 4. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT.RESET_DEVICE
MODS-xxxxxxxxx541	NVRM detected memory error	Need access to logs or failing system to debug root cause	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not tested with the latest version. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached,	1. TRIAGE_DEVICE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			if the same error code is seen after the GPU FD run.	
MODS-xxxxxxxxx542	NVRM VBIOS invalid or rejected	It could be due to a varied NVRM VBIOS validation failure. File Software NVBug and attach dmesg, diagnostic error logs, NV bug report tool log, to debug root cause.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU FD if not tested with the latest GPU FD. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	1. REPORT_NV_BUG 2. PROD_FIT.RESET_DEVICE
MODS-xxxxxxxxx582	gnu stress test found pixel miscompares	Computation has returned an incorrect answer.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not tested with the latest version. 3. Start RMA qualification if the same error code is seen after the GPU FD run.	1. RETURN_FOR_ FURTHER_TRIAGE
MODS-xxxxxxxxx599	fan does not seem to cool the chip	Check cooling systems and engage with NVIDIA. Failing part qualifies for RMA only if the problem is isolated to NVIDIA hardware.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU FD if not tested with the latest GPU FD. 3. Start RMA qualification if the same error code is seen after the run and only if the cooling solution is part of NVIDIA hardware.	1. TRIAGE_DEVICE
MODS-xxxxxxxxx609	Interrupt request mechanism does not work	-poll_interrupts option can be used in case the system does not support legacy interrupts, MSI or MSI-X IRQs.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU FD if not tested with the latest GPU FD. 3. Test with latest GPU FD by passing '-poll_interrupts' command line option if the same error code is seen after the GPU FD run. 4. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	1. PROD_FIT.RESET_DEVICE
MODS-xxxxxxxxx614	Extra golden code miscompare	One of the GPU Field Diagnostics tests failed the consistency check on expected values.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not tested with the latest version. 3. Start RMA qualification if the same error code is seen after the GPU FD run.	1. RETURN_FOR_ FURTHER_TRIAGE
MODS-xxxxxxxxx679	NVRM invalid argument	Need access to logs or failing system to debug root cause if error is seen after	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Run InfoROM recovery tool. For Blackwell generation of products, please refer to *InfoROM Cleansing for RMA* -	1. REPORT_NV_BUG 2. PROD_FIT.RESET_DEVICE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
		running InfoROM recovery tool.	*User Guide* (NVOnline: 1120261). 3. Re-test with the latest GPU FD if not tested with the latest version. 4. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the GPU FD run.	
MODS-xxxxxxxxx735	Failed to perform hot reset	Need to restore PCI Config space to resolve the issue.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Perform hot reboot, i.e. reboot the system to restore PCI config space. 3. Perform cold reboot i.e., full system shutdown and manually reboot again if hot reboot does not resolve the issue. 4. Re-test with latest GPU FD if not tested with the latest GPU FD. 5. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	1. RECOVERY.RESET_BM 2. PROD_FIT
MODS-xxxxxxxxx779	Voltage value out of range	GPU Field Diagnostics bug has been fixed.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Check the input voltage to the GPU and the system power supply. 3. Re-test with the latest GPU FD if not tested with the latest version. 4. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	1. REPORT_NV_BUG 2. PROD_FIT.RESET_DEVICE
MODS-xxxxxxxxx818	Mods detected an assertion failure	A software error has occurred. Need access to logs or failing system to debug root cause of the generic failure.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with the latest GPU FD if not tested with the latest version. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU FD run.	1. REPORT_NV_BUG 2. PROD_FIT

4.2 L10 Compute Field Diagnostics Error Code to Action Mapping

For “L10 Compute Field Diagnostic” related failures, look up the error code in Table 4-3 and follow the “Next Steps” column to address the failing system. The “Data Center Action Category” corresponding to each error code is also included with the corresponding action description in the following table.

The L10 Compute Diag is used to ensure the proper functioning and reliability of various hardware components and configurations of a GB NVL Compute Tray. This includes checking running environment and infoROM, validating NVIDIA® NVLink®-C2C, PCIe, and NVLink™ connectivity, and testing CPU and GPU performance. It tests thermal, compute, and memory performance.

Table 4-3. L10 Diagnostics Error Code to Action Mapping

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
MODS-xxxxxxxxx001	MODS exited with a specific status	MODS received a signal to exit	File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the L10 Compute Diag run.	1. REPORT_NV_BUG
MODS-xxxxxxxxx002	software error	An unexpected software error has occurred.	1. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx005	Bad command line argument	Bad command line argument which could be an issue in the Diag tool or a user error in the Diag command line.	1. Check Diag command line. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. If still unresolved, file Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx008	bad parameter passed to function	Software error / bad parameter passed to software function.	1. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx014	Low Bandwidth	Bandwidth measured during test run is lower than thresholds. Could be	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute	1. ISOLATE_TRAY 2. TRIAGE_TRAY 3. REPORT_NV_BUG

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
		an issue with RM driver or NVLink connectivity.	Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, NVOS tech support from master NVSwitch and NVIDIA Bug report tool log if the same error code is seen after the run.	
MODS-xxxxxxxxx015	C2C bus error	Potential issue in C2C bus.	1. If the same error code is seen after the run with latest L10 Diag, return the board to the integrator for further triage. 2. Isolate the failing GPU board(s) with reported PCIe ID and test with GPU FD. 3. If the same error code is seen after the run with latest Diag, RMA the board or the NBU unit. To RMA the board, return the board to the integrator for further triage. The guide for RMA the NBU unit is RMA Checklists and Support Upgrade Procedures (https://docs.nvidia.com/networking/display/rmachecklistrev).	1. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx021	script failed to execute	Script failed to execute.	1. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx024	Process signal received	MODS receives a signal and stops.	1. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx041	cannot hook interrupt	A software process or GPU state is preventing hooking interrupts.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI id and test with GPU FD. 4. File Software NVBug to debug root cause with dmesg, diagnostic error logs,	1. PROD_FIT.RESET_TRAY

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			and NVIDIA Bug report tool log attached if the same error code is seen after the run.	
MODS-xxxxxxxxx043	Diag crashed, timeout, or did not respond	Device stopped responding, crashed, or timed out.	1. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 2. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. PROD_FIT.RESET_TRAY 2. REPORT_NV_BUG
MODS-xxxxxxxxx060	NVRM invalid param struct		1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT.RESET_TRAY
MODS-xxxxxxxxx077	timeout error	Tests failed with SW timeout issue and needs further investigation.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing unit. To isolate GPU board(s), test the failing GPU board(s) with reported PCI id with GPU FD. To isolate networking component(s), swap suspect networking component(s) with known good parts and retest with the latest L10 Compute Diag. 4. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run. If qualify, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	1. TRIAGE_TRAY
MODS-xxxxxxxxx083	CRC/Checksum miscompare	Computation has returned an incorrect answer.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD.	1. RETURN_FOR_FURTHER_TRIAGE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			4. If the same error code is seen after the run, return the board to the integrator for further triage.	
MODS-xxxxxxxxx097	unexpected device interrupts	Diag received an unexpected interrupt that it considers an error.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI id and test with GPU FD. 4. If the same error code is seen after the run, return the board to the integrator for further triage.	1. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx124	Invalid InfoROM	InfoROM corruption has occurred. Recover InfoROM by running InfoROM recovery tool.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Run InfoROM recovery tool. For Blackwell generation of products, please refer to *InfoROM Cleansing for RMA - User Guide* (NVOnline: 1120261). 3. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 4. If the same error code is seen after the run, return the board to the integrator for further triage.	1.TRIAGE_TRAY
MODS-xxxxxxxxx127	Vbios Certificate Error	An issue with the VBIOS certificate.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. RECOVERY.RESET_SW 2. PROD_FIT
MODS-xxxxxxxxx139	Acceptable temperature limits exceeded, or the thermal sensor is broken or miscalibrated	Check cooling systems and engage with NVIDIA. Failing part qualifies for RMA only if the problem is isolated to NVIDIA hardware.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing unit. To isolate GPU board(s), test the failing GPU board(s) with reported PCI ID with GPU FD. To isolate networking component(s), swap suspect networking component(s) with known good parts and retest with the latest L10 Compute Diag 4. If the same error code is seen after the run, return the board to the integrator for further triage. This should be RMA'ed only if the cooling solution is part of NVIDIA	1. TRIAGE_TRAY

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			hardware. If qualified, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	
MODS-xxxxxxxxx140	NvLink bus error	NVLink is down or NVLink errors occurred during NVLink testing.	1. Reseat the tray containing the failing GPU. 2. Ensure GPU and NVSwitches are detected on the system using lspci utility. 3. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 4. Isolate the failing unit. To isolate GPU board(s), test the failing GPU board(s) with reported PCI id with GPU FD. To isolate the networking component(s), swap suspect networking component(s) with known good parts and retest with the latest L10 Compute Diag. 5. If the same error code is seen after the run, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	1. TRIAGE_TRAY
MODS-xxxxxxxxx143	PCI Express bus error	PCIe link errors causes PCIe tests to fail.	1. Reseat the tray containing the failing GPU. 2. Ensure GPU and NVSwitches are detected on the system using lspci utility. 3. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 4. Isolate the failing unit. To isolate GPU board(s), test the failing GPU board(s) with reported PCI id with GPU FD. To isolate networking component(s), swap suspect networking component(s) with known good parts and retest with the latest L10 Compute Diag. 5. If the same error code is seen after the run, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking	1. TRIAGE_TRAY

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	
MODS-xxxxxxxxx144	CUDA error	One of the CUDA tests is failing.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run. 5. If the same error code is seen after the run, return the board to the integrator for further triage.	1. TRIAGE_TRAY
MODS-xxxxxxxxx145	culnit failed	Use updated version of diag. Diag releases later than 1.1 should no longer see this error code. The resolution for this error code is the same as that for 688.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx160	Resource in use	Resource is reserved by another thread or test.	1. Perform cold reboot i.e., full system shutdown and manually reboot again. 2. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the reboot.	1. PROD_FIT.RESET_TRAY 2. REPORT_NV_BUG
MODS-xxxxxxxxx167	GFW boot reported a failure	GPU firmware failed to initialize	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Reseat the tray containing the failing GPU. 3. Perform cold reboot i.e., full system shutdown and manually reboot again. 4. Ensure GPU and NVSwitches are detected on the system using lspci utility. 5. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 6. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 7. File Software NVBug to debug root cause with dmesg, diagnostic error logs,	1. REPORT_NV_BUG 2. RECOVERY.RESET_BM 3. PROD_FIT

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			and NVIDIA Bug report tool log attached if the same error code is seen.	
MODS-xxxxxxxxx194	bad memory	Computation has returned an incorrect answer.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 4. If the same error code is seen after the run, return the board to the integrator for further triage.	1. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx216	Could not find specified device	Diag detected inactive NVLinks across GPUs. If seen on all NVLinks on a GPU, could be a GPU specific problem.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. If failure seen in NVLink related test (NVLink connectivity issue), isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. If failure seen in non-NVLink related test, it could be a PCIe interface problem. Check if GPUs are enumerated in PCIe tree via lscpi. Reset Compute Tray to confirm if issue persists. 5. Perform cold reboot i.e., full system shutdown and manually reboot again. If issue persists, isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 6. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code persists, and the failure is not isolated to a particular tray.	(NVLink) 1. ISOLATE_TRAY 2. TRIAGE_TRAY 3. REPORT_NV_BUG (GPU) 1. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx220	pci device not found	Failure with PCIe device. Need to isolate the failure to GPU, GB200 module or tray.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Reseat the tray containing the failing GPU. 3. Ensure GPU and NVSwitches are detected on the system using lspci utility. 4. Perform cold reboot i.e., full system shutdown and manually reboot again. 5. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 6. Isolate the failing unit. To isolate GPU board(s), test the failing GPU board(s)	1. RECOVERY.RESET_BM 2. PROD_FIT

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			with reported PCI id with GPU FD. To isolate the networking component(s), swap suspect networking component(s) with known good parts and retest with the latest L10 Compute Diag. 7. If the same error code is seen after the run, return the board to the integrator for further triage.	
MODS-xxxxxxxxx229	hardware was not initialized	Something prevented successful hardware initialization.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Check all the necessary power cables are attached to the tray. 3. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 4. Isolate the failing unit. To isolate GPU board(s), test the failing GPU board(s) with reported PCI ID with GPU FD. To isolate the networking component(s), swap suspect networking component(s) with known good parts and retest with the latest L10 Compute Diag. 5. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	1. TRIAGE_TRAY
MODS-xxxxxxxxx233	bad NVIDIA chip	IST detects bad NVIDIA chip.	1. If the same error code is seen after the run with latest L10 Compute Diag, file an NVBug. 2. RMA the board or the networking component. The guide for RMA the networking component is RMA Checklists and Support Upgrade Procedures (https://docs.nvidia.com/networking/display/rmachecklistrev).	1. REPORT_NV_BUG 2. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx240	Unexpected result from HW	If running GPU Field Diagnostic on single GPU, make sure 'skip_nvlink' parameter is used. If it is, then there is possible setup issue, confirm nvlink topology is connected as	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing unit. To isolate GPU board(s), test the failing GPU board(s)	1. TRIAGE_TRAY

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
		expected by running: nvidia-smi nvlink -s If NVLinks are reported as inactive, debug NVLink issue per debug tips document.	with reported PCI id with GPU FD. To isolate networking component(s), swap suspect networking component(s) with known good parts and retest with the latest L10 Compute Diag. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	
MODS-xxxxxxxxx272	Read parameter differs from expected	Re-test with latest L10 Compute Diagnostic.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	1. TRIAGE_TRAY
MODS-xxxxxxxxx276	HW reports wrong status	HW failure has occurred.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing unit. To isolate GPU board(s), test the failing GPU board(s) with reported PCI ID with GPU FD. To isolate networking component(s), swap suspect networking component(s) with known good parts and retest with the latest L10 Compute Diag. 4. If the same error code is seen after the run, RMA the board or the networking component. To RMA the board, return the	1. TRIAGE_TRAY

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	
MODS-xxxxxxxxx301	FSP returned non-zero error code	If the plain-text log shows MNOC_LOC message, then this error is fixed in GB NVL FW 1.3 release. Otherwise, this may imply the FW hit an error.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI id and test with GPU FD. 4. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx311	NvLink discovered topology does not match required topology	NVLink topology does not match expected topology.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Ensure GPU and NVSwitches are detected on the system using lspci utility. 3. Perform cold reboot i.e., full system shutdown and manually reboot again. 4. Reseat the tray containing the failing GPU if reboot did not resolve the issue. 5. Re-install NVIDIA release driver if rebooting and reseating did not resolve the issue. 6. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 7. Isolate the failing unit. To isolate GPU board(s), test the failing GPU board(s) with reported PCI id with GPU FD. To isolate networking component(s), swap suspect networking component(s) with known good parts and retest with the latest L10 Compute Diag. 8. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	1. TRIAGE_TRAY

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
MODS-xxxxxxxxx316	ECC detected a correctable error in FB, threshold exceeded	N/A	N/A	N/A
MODS-xxxxxxxxx317	ECC detected an uncorrectable error in FB	Uncorrectable error has occurred, reset for row remapping to take place.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Power cycle to trigger row remapping. 3. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 4. Isolate the failing GPU board(s) with reported PCI id and test with GPU FD. 5. If the same error code is seen after the run, return the board to the integrator for further triage.	1. TRIAGE_TRAY
MODS-xxxxxxxxx318	ECC detected a correctable error in L2, threshold exceeded	N/A	N/A	N/A
MODS-xxxxxxxxx319	ECC detected an uncorrectable error in L2	An uncorrectable error has occurred in the L2 cache, start RMA qualification for the GB200 module including the failing GPU.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 4. If the same error code is seen after the run, return the board to the integrator for further triage.	1. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxx171321	ECC detected an uncorrectable error in SM	N/A	N/A	N/A
MODS-xxxxxxxxx320	ECC detected a correctable error in SM, threshold exceeded	Rate of correctable errors in SM have exceeded threshold, start RMA qualification for the GB200 module including the failing GPU.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 4. If the same error code is seen after the run, return the board to the integrator for further triage.	1. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx321	ECC detected an uncorrectable error in SM	An uncorrectable error has occurred in the SM during testing, start RMA qualification for the GB200 module including the failing GPU.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD.	1. RETURN_FOR_FURTHER_TRIAGE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			4. If the same error code is seen after the run, return the board to the integrator for further triage.	
MODS-xxxxxxxxx332	Cannot pulse fast enough to accurately make a waveform	GPU is taking too long to either: (a) Complete the GEMM workload, or (b) Report that it completed the GEMM workload.	1. Reset the tray and retest. 2. File Software NVBug to debug root cause with dmesg, diagnostic error log, and NVIDIA Bug report tool log attached if the same error is seen after the L10 Compute Diag run.	1. RECOVERY.RESET_BM 2. REPORT_NV_BUG 3. PROD_FIT
MODS-xxxxxxxxx341	Buffer mismatch	Start RMA qualification for the GB200 module including the failing GPU.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 4. If the same error code is seen after the run, return the board to the integrator for further triage.	1. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx351	memory not strapped correctly	Memory not strapped correctly.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. TRIAGE_TRAY
MODS-xxxxxxxxx363	Row remapping failed	A row remapping failure has occurred, start RMA qualification.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 4. If the same error code is seen after the run, return the board to the integrator for further triage.	1. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx539	NVRM Generic falcon error	Ensure NV driver is fully unloaded, and other internal processes don't try to load driver back while running GPU Field Diagnostic.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Ensure GPU are detected on the system using lspci utility. Ensure loopback or switches are connected. 3. Perform cold reboot i.e., full system shutdown and manually reboot again. 4. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute	1. REPORT_NV_BUG 2. PROD_FIT.RESET_TRAY

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			Diag. 5. Isolate the failing GPU board(s) with reported PCI id and test with GPU FD. 6. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	
MODS-xxxxxxxxx541	NVRM Detected memory error	Memory issue detected by NVRM.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing unit. To isolate GPU board(s), test the failing GPU board(s) with reported PCI id with GPU FD. To isolate networking component(s), swap suspect networking component(s) with known good parts and retest with the latest L10 Compute Diag. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	1. TRIAGE_TRAY
MODS-xxxxxxxxx542	NVRM VBIOS invalid or rejected	It could be due to a varied NVRM VBIOS validation failure. File Software NV bug and attach dmesg, diagnostic error logs, NV bug report tool log attached to debug root cause.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT.RESET_TRAY
MODS-xxxxxxxxx582	gnu stress test found pixel miscompares	Computation has returned an incorrect answer.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 4. If the same error code is seen after the	1. RETURN_FOR_FURTHER_TRIAGE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			run, return the board to the integrator for further triage.	
MODS-xxxxxxxxx597	NVRM not supported	Need to first check whether this error code happens in an NVLink-related test or a non-NVLink related test. Need to check FW version.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. If failure seen in NVLink related test (NVLink connectivity issue), isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. If failure seen in non-NVLink related test, it could be a PCIe interface problem. Check if GPUs are enumerated in PCIe tree via lscpi. Reset Compute Tray to confirm if issue persists. If issue persists, isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 4. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	(NVLink) 1. ISOLATE_TRAY 2. TRIAGE_TRAY 3. REPORT_NV_BUG (GPU) 1. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx599	fan does not seem to cool the chip	Check cooling systems and engage with NVIDIA. Failing part qualifies for RMA only if the problem is isolated to NVIDIA hardware.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. If the same error code is seen after the run, return the board to the integrator for further triage. This should be RMA'ed only if the cooling solution is part of NVIDIA hardware. If qualify, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	1. TRIAGE_TRAY
MODS-xxxxxxxxx603	NVRM Timeout	Driver timeout possibly because system is in a bad state.	1. Perform cold reboot i.e., full system shutdown followed by a manual reboot. 2. Re-test with latest L10 Compute Diag if not already tested with the latest L10 Compute Diag. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. RECOVERY.RESET_BM 2. REPORT_NV_BUG

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
MODS-xxxxxxxxx609	Interrupt request mechanism does not work	A problem with the interrupt request mechanism.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. PROD_FIT.RESET_TRAY
MODS-xxxxxxxxx614	Extra golden code miscompare	One of the GPU Field Diagnostic tests failed the consistency check on expected values.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 4. If the same error code is seen after the run, return the board to the integrator for further triage.	1. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx665	NVRM invalid parameter	Usually indicative of a SW/FW issue. Does not point to a system HW problem.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not already tested with the latest L10 Compute Diag. 4. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx679	NVRM invalid argument	If error is seen after running InfoROM recovery tool, file Software NV bug and attach dmesg, diagnostic error logs, NV bug report tool log attached to debug root cause.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Run InfoROM recovery tool. For Blackwell generation of products, please refer to *InfoROM Cleansing for RMA - User Guide* (NVOnline: 1120261). 3. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 4. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 5. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT.RESET_TRAY
MODS-xxxxxxxxx682	NVRM invalid data	GPU driver returns invalid data.	1. If the same error code is seen after a rerun with latest L10 Compute Diag, return the board to the integrator for further triage.	1. RETURN_FOR_FURTHER_TRIAGE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
MODS-xxxxxxxxx688	NVRM Invalid State or Config	NVRM reports an invalid state or configuration of an NVLink. One or more NVLinks is in an invalid state. Could be an issue with the Global Fabric Manager on master NVSwitch.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Check GFM health on master NVSwitch by running 'nv show cluster apps running' NVOS CLI command. 3. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 4. If failure seen in NVLink related test (NVLink connectivity issue), isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 5. If failure seen in non-NVLink related test, it could be a PCIe interface problem. Check if GPUs are enumerated in PCIe tree via lscpi. Reset Compute Tray to confirm if issue persists. 6. Perform cold reboot i.e., full system shutdown and manually reboot again. If issue persists, isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 7. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, NVOS tech support from master NVSwitch and NVIDIA Bug report tool log if the same error code persists, and the failure is not isolated to a particular tray.	(NVLink) 1. ISOLATE_TRAY 2. TRIAGE_TRAY 3. REPORT_NV_BUG (GPU) 1. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx735	Failed to perform hot reset	An issue with hot resetting the device via Diag. File Software NV bug and attach dmesg, diagnostic error logs, NV bug report tool log attached to debug root cause.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Perform hot reboot, i.e. reboot the system to restore PCI config space. 3. Perform cold reboot i.e., full system shutdown and manually reboot again if hot reboot does not resolve the issue. 4. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 5. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 6. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. RECOVERY.RESET_BM 2. PROD_FIT
MODS-xxxxxxxxx779	Voltage value out of range	Voltage value is out of range.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Check the input voltage to the GPU and the system power supply. 3. Re-test with latest L10 Compute Diag if	1. REPORT_NV_BUG 2. PROD_FIT.RESET_TRAY

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			not tested with the latest L10 Compute Diag. 4. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 5. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	
MODS-xxxxxxxxx818	Mods detected an assertion failure	A software error has occurred.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx936	TLW Error	An error in the TLW (Transaction Layer Wrapper) module on an NVLink port.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, NVOS tech support (for Switch Tray) and NVIDIA Bug report tool log if the same error code is seen after the run.	1. ISOLATE_TRAY 2. TRIAGE_TRAY 3. REPORT_NV_BUG
MODS-xxxxxxxxx937	NVLINK: TREX Error	NVLink transmission error in the TREX component.	1. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx999	Segfault		1. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG
DGX-xxxxxxxxx001	MODS exited with status 1	Software issue	1. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute	1. REPORT_NV_BUG 2. PROD_FIT

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	
DGX-xxxxxxxxx002	timeout	An unexpected watchdog timeout during the Diag test run.	1. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
DGX-xxxxxxxxx007	script failed to execute	Script failed to execute. Software issue.	1. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
DGX-xxxxxxxxx009	System setup error	A setup error in system.	1. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
DGX-xxxxxxxxx086	Inventory item mismatch	Mismatch in an inventory (e.g. firmware version) check.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
DGX-xxxxxxxxx111	Infiniband property mismatch	Infiniband check fails.	1. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
DGX-xxxxxxxxx184	Repair action	A repair action failed.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs	1. REPORT_NV_BUG 2. PROD_FIT

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			and NVIDIA Bug report tool log if the same error code is seen after the run.	
DGX-xxxxxxxxx189	GPU health check failed	The system might be in a bad state.	1. Perform cold reboot i.e., full system shutdown followed by a manual reboot. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. RECOVERY.RESET_BM 2. REPORT_NV_BUG
DGX-xxxxxxxxx190	Environment check failed	Preconditions of running diagnostics are not met	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI id and test with GPU FD. 4. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT

4.3 L11 Rack Field Diagnostics Error Code to Action Mapping

For “L11 Rack Field Diagnostics” related failures, look up the error code in Table 4-4 and follow the “Next Steps” column to address the failing system. The “Data Center Action Category” corresponding to each error code is also included with the corresponding action description in the following table.

The L11 Rack Diag is intended to verify correct assembly of the rack. This includes verifying the NVLink connections between the Compute and Switch Trays and running cross-node traffic tests. The Diag also includes pulsed workload and a steady-state thermal workload test.

Table 4-4. L11 Diagnostics Error Code to Action Mapping

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
MODS-xxxxxxxxx001	MODS exited with a specific status	MODS received a signal to exit	File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the L11 Rack Diag run.	1. REPORT_NV_BUG
MODS-xxxxxxxxx002	software error	An unexpected software error has occurred.	1. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack	1. REPORT_NV_BUG 2. PROD_FIT

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	
MODS-xxxxxxxxx005	Bad command line argument	Bad command line argument which could be an issue in the Diag tool or a user error in the Diag command line.	1. Check Diag command line. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. If still unresolved, file Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx008	bad parameter passed to function	Software error / bad parameter passed to software function.	1. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx014	Low Bandwidth	Bandwidth measured during test run is lower than thresholds. Could be an issue with RM driver or NVLink connectivity.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, NVOS tech support from master NVSwitch and NVIDIA Bug report tool log if the same error code is seen after the run.	1. ISOLATE_TRAY 2. TRIAGE_TRAY 3. REPORT_NV_BUG
MODS-xxxxxxxxx015	C2C bus error	Possible issue in C2C bus	1. If the same error code is seen after the run with latest L11 Rack Diag, return the board to the integrator for further triage. 2. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The guide for RMA the networking component is RMA Checklists and Support Upgrade	1. RETURN_FOR_FURTHER_TRIAGE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			Procedures (https://docs.nvidia.com/networking/display/rmachecklistrev).	
MODS-xxxxxxxxx021	script failed to execute	Script failed to execute.	1. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx024	Process signal received	MODS receives a signal and stops.	1. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx041	cannot hook interrupt	A software process or GPU state is preventing hooking interrupts.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. PROD_FIT.RESET_RACK
MODS-xxxxxxxxx043	Diag crashed, timeout, or did not respond	Device stopped responding, crashed, or timed out.	1. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 2. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. PROD_FIT.RESET_TRAY 2. REPORT_NV_BUG
MODS-xxxxxxxxx044	MODS test failed on a remote node	Test fails on another node and lead current node exits. Triage should focus on the remote node.	1. Isolate the failing tray. If it is a Compute tray, test with L10 Compute Diag. If it is a switch tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 2. File Software NVBug to debug root cause with dmesg, diagnostic error	1. ISOLATE_TRAY 2. TRIAGE_TRAY 3. REPORT_NV_BUG

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	
MODS-xxxxxxxxx060	NVRM invalid param struct		1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT.RESET_TRAY
MODS-xxxxxxxxx077	timeout error	Tests failed with SW timeout issue and needs further investigation.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	1. ISOLATE_TRAY 2. TRIAGE_TRAY
MODS-xxxxxxxxx083	CRC/Checksum miscompare	Computation has returned an incorrect answer.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. If the same error code is seen after	1. ISOLATE_TRAY 2. RETURN_FOR_FURTHER_TRIAGE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			the run, return the board to the integrator for further triage.	
MODS-xxxxxxxxx097	unexpected device interrupts	Diag received an unexpected interrupt that it considers an error.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. If the same error code is seen after the run, return the board to the integrator for further triage.	1. ISOLATE_TRAY 2. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx124	Invalid InfoROM	N/A	N/A	N/A
MODS-xxxxxxxxx127	Vbios Certificate Error	An issue with the VBIOS certificate.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. RECOVERY.RESET_SW 2. PROD_FIT
MODS-xxxxxxxxx139	Acceptable temperature limits exceeded, or the thermal sensor is broken or miscalibrated	Check cooling systems and engage with NVIDIA. Failing part qualifies for RMA only if the problem is isolated to NVIDIA hardware.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. If the same error code is seen after the run, return the board to the integrator for further triage. This should be RMA'ed only if the cooling solution is part of NVIDIA hardware. If qualify, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is	1. ISOLATE_TRAY 2. TRIAGE_TRAY

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	
MODS-xxxxxxxxx140	NvLink bus error	NVLink is down or NVLink errors occurred during NVLink testing.	1. Reseat the tray containing the failing GPU. 2. Ensure GPU and NVSwitches are detected on the system using lspci utility. 3. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 4. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 5. If the same error code is seen after the run, return the board to the integrator for further triage. If qualify, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	1. ISOLATE_TRAY 2. TRIAGE_TRAY
MODS-xxxxxxxxx143	PCI Express bus error	PCIe link errors causes PCIe tests to fail.	1. Reseat the tray containing the failing GPU. 2. Ensure GPU and NVSwitches are detected on the system using lspci utility. 3. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 4. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 5. If the same error code is seen after the run, return the board to the integrator for further triage. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support	1. ISOLATE_TRAY 2. TRIAGE_TRAY

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	
MODS-xxxxxxxxx144	CUDA error	One of the CUDA tests is failing. File Software NV bug and attach dmesg, diagnostic error logs, NV bug report tool log attached to debug root cause.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run. Start RMA qualification if the same error code is seen after the run.	1. ISOLATE_TRAY 2. TRIAGE_TRAY
MODS-xxxxxxxxx145	culnit failed	Use updated version of diag. Diags later than release 1.1 should no longer see this error code. The resolution for this error code is the same as that for 688.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx160	Resource in use	Resource is reserved by another thread or test.	1. Perform cold reboot i.e., full system shutdown and manually reboot again. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the reboot.	1. PROD_FIT.RESET_TRAY 2. REPORT_NV_BUG
MODS-xxxxxxxxx167	GFW boot reported a failure	GPU firmware failed to initialize.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Reseat the tray containing the failing GPU. 3. Perform cold reboot i.e., full system shutdown and manually reboot again. 4. Ensure GPU and NVSwitches are detected on the system using lspci utility. 5. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack	1. REPORT_NV_BUG 2. RECOVERY.RESET_BM 3. PROD_FIT

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			Diag. 6. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 7. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen.	
MODS-xxxxxxxxx194	bad memory	Computation has returned an incorrect answer.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. If the same error code is seen after the run, return the board to the integrator for further triage.	1. ISOLATE_TRAY 2. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx216	Could not find specified device	Diag detected inactive NVLinks across GPUs. If seen on all NVLinks on a GPU, could be a GPU specific problem.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. If failure seen in NVLink related test (NVLink connectivity issue), isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. If failure seen in non-NVLink related test, it could be a PCIe interface problem. Check if GPUs are enumerated in PCIe tree via lscpi. Reset Compute Tray to confirm if issue persists. 5. Perform cold reboot i.e., full system shutdown and manually reboot again. If issue persists, isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 6. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if	(NVLink) 1. ISOLATE_TRAY 2. TRIAGE_TRAY 3. REPORT_NV_BUG (GPU) 1. RETURN_FOR_FURTHER_TRIAGE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			the same error code persists, and the failure is not isolated to a particular tray.	
MODS-xxxxxxxxx220	pci device not found	Failure with PCIe device. Need to isolate the failure to GPU, GB200 module or tray.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Reseat the tray containing the failing GPU. 3. Ensure GPU and NVSwitches are detected on the system using lspci utility. 4. Perform cold reboot i.e., full system shutdown and manually reboot again. 5. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 6. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 7. If the same error code is seen after the run, return the board to the integrator for further triage.	1. RECOVERY.RESET_BM 2. PROD_FIT
MODS-xxxxxxxxx229	hardware was not initialized	Something prevented successful hardware initialization.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Check all the necessary power cables are attached to the tray. 3. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 4. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 5. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	1. ISOLATE_TRAY 2. TRIAGE_TRAY

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
MODS-xxxxxxxxx233	bad NVIDIA chip	IST detects bad NVIDIA chip.	1. If the same error code is seen after the run with latest L11 Rack Diag, return the board to the integrator for further triage. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The guide for RMA the networking component is RMA Checklists and Support Upgrade Procedures (https://docs.nvidia.com/networking/display/rmachecklistrev).	1. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx240	Unexpected result from HW	If running GPU Field Diagnostic on single GPU, make sure 'skip_nvlink' parameter is used. If it is, then there is possible setup issue, confirm NVLink topology is connected as expected by running: nvidia-smi nvlink -s If NVLinks are reported as inactive, debug NVLink issue per debug tips document.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	1. ISOLATE_TRAY 2. TRIAGE_TRAY
MODS-xxxxxxxxx272	Read parameter differs from expected	Re-test with latest L11 Rack Diagnostics.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The	1. ISOLATE_TRAY 2. TRIAGE_TRAY

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	
MODS-xxxxxxxxx276	HW reports wrong status	HW failure has occurred.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. If the same error code is seen after the run, return the board to the integrator for further triage. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	1. ISOLATE_TRAY 2. TRIAGE_TRAY
MODS-xxxxxxxxx301	FSP returned non-zero error code	If the plain-text log shows MNOC_LOC message, then this error is fixed in GB NVL FW 1.3 release. Otherwise, this may imply the FW hit an error.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L10 Compute Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing GPU board(s) with reported PCI id and test with GPU FD. 4. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx311	NvLink discovered topology does not match required topology	NVLink topology does not match expected topology.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Ensure GPU and NVSwitches are detected on the system using lspci utility. Ensure the topology file is correct based on the user guide. 3. Perform cold reboot i.e., full system shutdown and manually reboot again. 4. Reseat the tray containing the failing	1. ISOLATE_TRAY 2. TRIAGE_TRAY

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			GPU if reboot did not resolve the issue. 5. Re-install NVIDIA release driver if rebooting and reseating did not resolve the issue. 6. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 7. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 8. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	
MODS-xxxxxxxxx316	ECC detected a correctable error in FB, threshold exceeded	N/A	N/A	N/A
MODS-xxxxxxxxx317	ECC detected an uncorrectable error in FB	Uncorrectable error has occurred, reset for row remapping to take place.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Power cycle to trigger row remapping. 3. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 4. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 5. If the same error code is seen after the run, return the board to the integrator for further triage.	1. ISOLATE_TRAY 2. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx318	ECC detected a correctable error in L2, threshold exceeded	N/A	N/A	N/A

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
MODS-xxxxxxxxx319	ECC detected an uncorrectable error in L2	An uncorrectable error has occurred in the L2 cache, start RMA qualification for the GB200 module including the failing GPU.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. If the same error code is seen after the run, return the board to the integrator for further triage.	1. ISOLATE_TRAY 2. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxx171321	ECC detected an uncorrectable error in SM	N/A	N/A	N/A
MODS-xxxxxxxxx320	ECC detected a correctable error in SM, threshold exceeded	Rate of correctable errors in SM have exceeded threshold, start RMA qualification for the GB200 module including the failing GPU.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. If the same error code is seen after the run, return the board to the integrator for further triage.	1. ISOLATE_TRAY 2. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx321	ECC detected an uncorrectable error in SM	An uncorrectable error has occurred in the SM during testing, start RMA qualification for the GB200 module including the failing GPU.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. If the same error code is seen after the run, return the board to the integrator for further triage.	1. ISOLATE_TRAY 2. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxx332	Cannot pulse fast enough to accurately make a waveform	GPU is taking too long to either: (a) Complete the GEMM workload, or (b) Report that it	1. Reset the tray and retest. 2. File Software NVBug to debug root cause with dmesg, diagnostic error log, and NVIDIA Bug report tool log	1. RECOVERY.RESET_BM 2. REPORT_NV_BUG 3. PROD_FIT

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
		completed the GEMM workload.	attached if the same error is seen after the GPU FD run.	
MODS-xxxxxxxxx341	Buffer mismatch	Start RMA qualification for the GB200 module including the failing GPU.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. If the same error code is seen after the run, return the board to the integrator for further triage.	1. ISOLATE_TRAY 2. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx351	memory not strapped correctly	Memory not strapped correctly.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. ISOLATE_TRAY 2. TRIAGE_TRAY
MODS-xxxxxxxxx363	Row remapping failed	A row remapping failure has occurred, start RMA qualification.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. If the same error code is seen after the run, return the board to the integrator for further triage.	1. ISOLATE_TRAY 2. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx539	NVRM generic falcon error	Ensure NV driver is fully unloaded, and other internal processes don't try to load driver back	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Ensure GPU and NVSwitches are	1. REPORT_NV_BUG 2. PROD_FIT.RESET_TRAY

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
		while running GPU Field Diagnostic.	detected on the system using lspci utility. 3. Perform cold reboot i.e., full system shutdown and manually reboot again. 4. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 5. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 6. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	
MODS-xxxxxxxxx541	NVRM Detected memory error	Memory issue detected by NVRM.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document. (https://docs.nvidia.com/networking/display/rmachecklistrev).	1. ISOLATE_TRAY 2. TRIAGE_TRAY
MODS-xxxxxxxxx542	NVRM VBIOS invalid or rejected	It could be due to a varied NVRM VBIOS validation failure.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide	1. REPORT_NV_BUG 2. PROD_FIT.RESET_TRAY

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			Section 4.3. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	
MODS-xxxxxxxxx582	gnu stress test found pixel miscompares	N/A	N/A	N/A
MODS-xxxxxxxxx597	NVRM not supported	Need to first check whether this error code happens in an NVLink-related test or a non-NVLink related test. Need to check FW version.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. If failure seen in NVLink related test (NVLink connectivity issue), isolate the failing tray. If it is a Compute Tray, test with L11 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. If failure seen in non-NVLink related test, it could be a PCIe interface problem. Check if GPUs are enumerated in PCIe tree via lscpi. Reset Compute Tray to confirm if issue persists. If issue persists, isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 5. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	(NVLink) 1. ISOLATE_TRAY 2. TRIAGE_TRAY 3. REPORT_NV_BUG (GPU) 1. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx599	fan does not seem to cool the chip	Check cooling systems and engage with NVIDIA. Failing part qualifies for RMA only if the problem is isolated to NVIDIA hardware.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Start RMA qualification if the same error code is seen after the run and only if the cooling solution is part of NVIDIA hardware. If qualified for RMA, RMA the board or the networking component. To RMA the board, return the board to the integrator for further triage. The RMA guide for networking components is the RMA Checklists and Support Upgrade Procedures document.	1. ISOLATE_TRAY 2. TRIAGE_TRAY

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			(https://docs.nvidia.com/networking/display/rmachecklistrev).	
MODS-xxxxxxxxx603	NVRM Timeout	Driver timeout possibly because system is in a bad state.	1. Perform cold reboot i.e., full system shutdown followed by a manual reboot. 2. Re-test with latest L11 Rack Diag if not already tested with the latest L11 Rack Diag. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. RECOVERY.RESET_BM 2. REPORT_NV_BUG
MODS-xxxxxxxxx609	Interrupt request mechanism does not work	A problem with the interrupt request mechanism.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. PROD_FIT.RESET_RACK
MODS-xxxxxxxxx614	Extra golden code miscompare	One of the GPU Field Diagnostic tests failed the consistency check on expected values.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. If the same error code is seen after the run, return the board to the integrator for further triage.	1. ISOLATE_TRAY 2. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx665	NVRM invalid parameter	Usually indicative of a SW/FW issue. Does not point to a system HW problem.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not already tested with the latest L11 Rack Diag. 4. File Software NVBug to debug root cause with dmesg, diagnostic error	1. REPORT_NV_BUG 2. PROD_FIT

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	
MODS-xxxxxxxxx679	NVRM invalid argument	Invalid argument passed to NVRM. If error is seen after running InfoROM recovery tool, file Software NV bug and attach dmesg, diagnostic error logs, NV bug report tool log attached to debug root cause.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Run InfoROM recovery tool. For Blackwell generation of products, please refer to *InfoROM Cleansing for RMA - User Guide* (NVOnline: 1120261). 3. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT.RESET_TRAY
MODS-xxxxxxxxx682	NVRM invalid data	GPU driver returns invalid data.	1. If the same error code is seen after a rerun with latest L11 Rack Diag, return the board to the integrator for further triage.	1. RETURN_FOR_FURTHER_TRIAGE
MODS-xxxxxxxxx688	NVRM Invalid State or Config	NVRM reports an invalid state or configuration of an NVLink. One or more NVLinks is in an invalid state. Could be an issue with the Global Fabric Manager on master NVSwitch.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Check GFM health on master NVSwitch by running 'nv show cluster apps running' NVOS CLI command. 3. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 4. If failure seen in NVLink related test (NVLink connectivity issue), isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 5. If failure seen in non-NVLink related test, it could be a PCIe interface problem. Check if GPUs are enumerated in PCIe tree via lscpi. Reset Compute Tray to confirm if issue persists. 6. Perform cold reboot i.e., full system shutdown and manually reboot again. If issue persists, isolate the failing GPU board(s) with reported PCI ID and test with GPU FD. 7. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if	(NVLink) 1. ISOLATE_TRAY 2. TRIAGE_TRAY 3. REPORT_NV_BUG (GPU) 1. RETURN_FOR_ FURTHER_TRIAGE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			the same error code persists, and the failure is not isolated to a particular tray.	
MODS-xxxxxxxxx735	Failed to perform hot reset	Hot reset did not work.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Perform hot reboot, i.e. reboot the system to restore PCI config space. 3. Perform cold reboot i.e., full system shutdown and manually reboot again if hot reboot does not resolve the issue. 4. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 5. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 6. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. RECOVERY.RESET_BM 2. PROD_FIT
MODS-xxxxxxxxx779	Voltage value out of range	N/A	N/A	N/A
MODS-xxxxxxxxx818	Mods detected an assertion failure	A software error has occurred.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx936	TLW Error	An error in the TLW (Transaction Layer Wrapper) module on an NVLink port.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L10 Compute Diag. 3. Isolate the failing tray. If it is a Compute Tray, test with L10 Compute	1. ISOLATE_TRAY 2. TRIAGE_TRAY 3. REPORT_NV_BUG

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			Diag. If it is a Switch Tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, NVOS tech support (for Switch Tray) and NVIDIA Bug report tool log if the same error code is seen after the run.	
DGX-xxxxxxxxx001	MODS exited with status 1	Software issue	1. Re-test with latest L11 Rack Diag if not tested with the latest L11 Compute Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
DGX-xxxxxxxxx002	timeout	An unexpected watchdog timeout during the Diag test run.	1. Re-test with latest L11 Rack Diag if not tested with the latest L11 Compute Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs and NVIDIA Bug report tool log if the same error code is seen after the run.	1. PROD_FIT.RESET_TRAY 2. REPORT_NV_BUG
DGX-xxxxxxxxx007	script failed to execute	Script failed to execute. Software issue.	1. Re-test with latest L11 Rack Diag if not tested with the latest L11 Compute Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
DGX-xxxxxxxxx009	System setup error	A setup error in system.	1. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
DGX-xxxxxxxxx086	Inventory item mismatch	Mismatch in an inventory (e.g. firmware version) check.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if	1. REPORT_NV_BUG 2. PROD_FIT

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			the same error code is seen after the run.	
DGX-xxxxxxxxx184	Repair action	A repair action failed.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs and NVIDIA Bug report tool log if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
DGX-xxxxxxxxx189	GPU health check failed	The system might be in a bad state.	1. Perform cold reboot i.e., full system shutdown followed by a manual reboot. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Compute Diag. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. RECOVERY.RESET_BM 2. REPORT_NV_BUG
DGX-xxxxxxxxx190	Environment check failed	Preconditions of running diagnostics are not met	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest L11 Rack Diag if not tested with the latest L11 Rack Diag. 3. Isolate the failing tray. If it is a compute tray, test with L10 Compute Diag. If it is a switch tray, follow NVIDIA GB200 NVL Service Flow User Guide Section 4.3. 4. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. ISOLATE_TRAY 2. TRIAGE_TRAY 3. REPORT_NV_BUG

4.4 HGX Baseboard Diagnostics Error Code to Action Mapping

For NVIDIA HGX baseboard related failures, look up the error code in Table 4-5 and follow the ‘Next Steps’ column to address the failing system.

Table 4-5. HGX Baseboard Diagnostics Error Code to Action Mapping

Error Code	Test	Error Message	Explanation	Next Step
DGX-000000000110	inventory	'VBIOS_version': not found. expected 'Y'	GPU property mismatch	Refer to “Missing GPU Property” section.
DGX-000000000110	inventory	'VBIOS_version': found 'Y' expected 'Z'	GPU property mismatch	Update GPU firmware.
DGX-000000000110	inventory	'VendorID': not found. expected 'Y'	GPU property mismatch	Refer to “Missing GPU Property” section.
DGX-000000000110	inventory	'VendorID': found 'Y' expected 'Z'	GPU property mismatch	Refer to “Missing GPU” section.
DGX-000000000110	inventory	'DeviceID': not found. expected 'Y'	GPU property mismatch	Refer to “Missing GPU Property” section.
DGX-000000000110	inventory	'DeviceID': found 'Y' expected 'Z'	GPU property mismatch	Refer to “Missing GPU” section.
DGX-000000000110	inventory	'SSVendorID': not found. expected 'Y'	GPU property mismatch	Refer to “Missing GPU Property” section.
DGX-000000000110	inventory	'SSVendorID': found 'Y' expected 'Z'	GPU property mismatch	Update GPU firmware.
DGX-000000000110	inventory	'SSDeviceID': not found. expected 'Y'	GPU property mismatch	Refer to “Missing GPU Property” section.
DGX-000000000110	inventory	'SSDeviceID': found 'Y' expected 'Z'	GPU property mismatch	Update GPU firmware.
DGX-000000000110	inventory	'X': not found. expected 'Y'	GPU property mismatch	Refer to “Missing GPU Property” section.
DGX-000000000110	inventory	'X': found 'Y' expected 'Z'	GPU property mismatch	Update GPU firmware.
DGX-000000000108	inventory	'VBIOS_version': not found. expected 'Y'	NVSwitch property mismatch	Refer to “Missing NVSwitch Property” section.
DGX-000000000108	inventory	'VBIOS_version': found 'Y' expected 'Z'	NVSwitch property mismatch	Update NVSwitch firmware.
DGX-000000000108	inventory	'VendorID': not found. expected 'Y'	NVSwitch property mismatch	Refer to “Missing NVSwitch Property” section.
DGX-000000000108	inventory	'VendorID': found 'Y' expected 'Z'	NVSwitch property mismatch	Refer to “Missing NVSwitch” section.
DGX-000000000108	inventory	'DeviceID': not found. expected 'Y'	NVSwitch property mismatch	Refer to “Missing NVSwitch Property” section.
DGX-000000000108	inventory	'DeviceID': found 'Y' expected 'Z'	NVSwitch property mismatch	Refer to “Missing NVSwitch” section.
DGX-000000000108	inventory	'SSVendorID': not found. expected 'Y'	NVSwitch property mismatch	Refer to “Missing NVSwitch Property” section.

Error Code	Test	Error Message	Explanation	Next Step
DGX-000000000108	inventory	'SSVendorID': found 'Y' expected 'Z'	NVSwitch property mismatch	Update NVSwitch firmware
DGX-000000000108	inventory	'SSDeviceID': not found. expected 'Y'	NVSwitch property mismatch	Refer to “Missing NVSwitch Property” section.
DGX-000000000108	inventory	'SSDeviceID': found 'Y' expected 'Z'	NVSwitch property mismatch	Update NVSwitch firmware.
DGX-000000000108	inventory	'X': not found. expected 'Y'	NVSwitch property mismatch	Refer to “Missing NVSwitch Property” section.
DGX-000000000108	inventory	'X': found 'Y' expected 'Z'	NVSwitch property mismatch	Update NVSwitch firmware.
DGX-000000000086	inventory	'GPUNum': found 'Y' expected 'Z'	An unexpected number of GPUs found.	Refer to “Missing GPU” section.
DGX-000000000086	inventory	'NVSWITCHNum': found 'Y' expected 'Z'	An unexpected number of NVSwitches found.	Refer to “Missing NVSwitch” section.
DGX-000000000086	inventory	'HMC_FW_version': not found. expected 'Y'	Unable to fetch HMC Firmware version	Refer to “Custom BMC Commands” section.
DGX-000000000086	inventory	'HMC_FW_version': found 'Y' expected 'Z'	Unexpected HMC Firmware version found	Update HMC Firmware.
DGX-000000000086	inventory	'HMC_FW_boot_complete': not found. expected 'Y'	Unable to fetch HMC Boot Complete Status	Refer to “Custom BMC Commands” section.
DGX-000000000086	inventory	'HMC_FW_boot_complete': found 'Y' expected 'Z'	Unexpected HMC Boot Complete Status	File NVBug for NVIDIA to review for RMA candidacy qualification and refer to “Providing Partner Diagnostic Logs” section.
DGX-000000000086	inventory	'X': not found. expected 'Y'	Unable to fetch attribute 'X'.	File NVBug for NVIDIA to review for RMA candidacy qualification and refer to “Providing Partner Diagnostic Logs” section.
DGX-000000000086	inventory	'X': found 'Y' expected 'Z'	Inventory item mismatch	Update firmware.

Error Code	Test	Error Message	Explanation	Next Step
DGX-xxxxxxxxx001	any	FieldDiag exited with status X	Software Error	Retest with the latest Partner Diagnostics. If this issue still occurs, file NVBug for NVIDIA to review for RMA candidacy qualification. Refer to “Providing Partner Diagnostic Logs” section.
DGX-xxxxxxxxxxx	any	any	-	File NVBug for NVIDIA to review for RMA candidacy qualification and refer to “Providing Partner Diagnostic Logs” section.
MODS-xxxxxxxx665	any	any	Incorrect PCI lane width	File NVBug for NVIDIA to review for RMA candidacy qualification and refer to ““Providing Partner Diagnostic Logs” section.
MODS-xxxxxxxx220	any	any	PCI device not found	Refer to “Missing GPU” and “Missing NVSwitch Property” sections.
MODS-xxxxxxxx143	any	any	PCI Express Bus Error	File NVBug for NVIDIA to review for RMA candidacy qualification and refer to “Providing Partner Diagnostic Logs” section.
MODS-xxxxxxxx140	any	any	NVLink Bus Error	Retest with the latest Partner Diagnostics. If it still fails, Start RMA qualification for baseboard.
MODS-xxxxx248272	any	any	Unexpected Result	Retest with the latest Partner Diagnostics. If this issue still occurs, file NVBug for NVIDIA to review for RMA candidacy qualification. Refer to “Providing Partner Diagnostic Logs” section.
MODS-xxxxxxxx609	any	any	GPU/NVSwitch Interrupt request mechanism does not work	Retest with the latest Partner Diagnostics. If issue persists, refer to “GPU Error Codes” section.

Error Code	Test	Error Message	Explanation	Next Step
MODS-xxxxxxxxxxxx	any	any	GPU/NVSWITCH diag failure	Refer to “GPU Error Codes” section.

4.5 Troubleshooting Actions for Partner Diagnostics Error Codes

This section describes troubleshooting the Partner Diagnostics error codes.

4.5.1 Missing GPU Property

If the 'GPUNum' check passed, then proceed to Section 4.5.6 “Providing Partner Diagnostic Logs.” Otherwise, proceed to Section 4.5.2 “Missing GPU.”

4.5.2 Missing GPU

Confirm all GPUs are present with lspci. If so, then proceed to Section 4.5.6 “Providing Partner Diagnostic Logs.” If not, try power-cycling the system. If the devices are still not present, try swapping the missing GPU module, reseating the baseboard, and updating firmware. If the issue persists, start RMA qualification.

4.5.3 Missing NVSwitch Property

If 'NVSWITCHNum' check passed, then proceed to Section 4.5.6 “Providing Partner Diagnostic Logs.” Otherwise, proceed to Section 4.5.4 “Missing NVSwitch.”

4.5.4 Missing NVSwitch

Confirm that all NVSwitch are present with lspci. If so, then proceed to Section 4.5.6 “Providing Partner Diagnostic Logs.” If not, try power-cycling the system. If the device is still not present, try reseating the baseboard and updating the firmware. If the issue persists, start RMA qualification.

4.5.5 Custom BMC Commands

This optional test requires system-specific commands to communicate with the BMC. The *NVIDIA Baseboard Field Diagnostics Software Guide* has instructions on how to provide these commands to the diag. If this BMC functionality is unavailable, the test can be skipped by adding "–-no_bmc" to the diag command line.

4.5.6 Providing Partner Diagnostic Logs

The diagnostics generates a tar ball of the log file called logs-<yyyymmdd-hhnnss>.tgz and places this file in the /logs Baseboard Partner Diagnostics folder.

The log file name is based on the test date and timestamp in the following format:

- logs/logs-yyyymmdd-hhnnss
- where:
- mm = Month
- dd = Day
- yy = Year
- hh = Hours
- nn = Minutes
- ss = Seconds

You can find the plain text logs of the Partner Diagnostics run in run.log. Additional details on the contents of the logs can be found in the *NVIDIA Baseboard Field Diagnostics Software Guide*.

4.6 CPU Field Diagnostics Common Error Code to Action Mapping

For NVIDIA Grace CPU Field Diagnostic failures, look up the error code in Table 4-6 and follow the ‘Next Steps’ section to address the failing system.

Figure 4-2. Grace CPU Partner Diagnostics Error Log

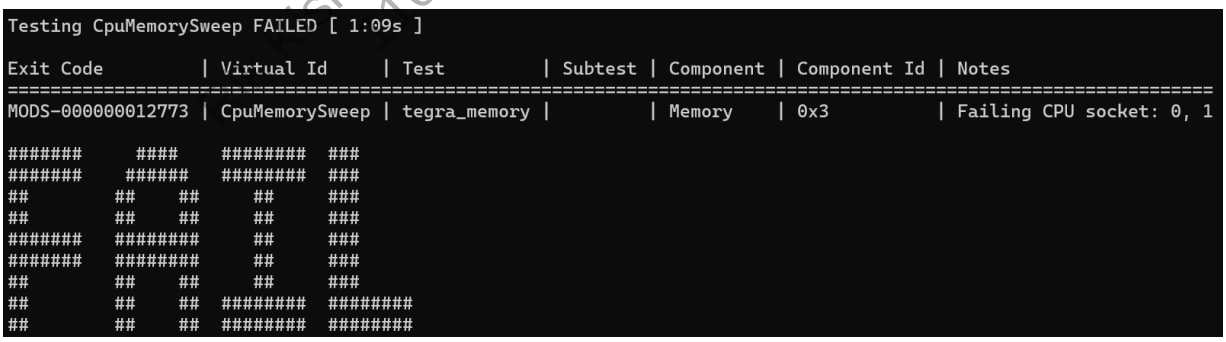


Table 4-6. Grace CPU Field Diagnostics Common Error Code to Action Mapping

Error Code	Error Message	Explanation	Next Action	Data Center Action Category
MODS-xxxxxxxxx002	Software error	Generic SW Error. Possible environment issue.	1. Check firmware versions against the latest NVIDIA released Firmware versions. 2. Check if the Diag tool was run with sudo. 3. Check if MODS kernel driver is installed. 4. If the issue persists, report Software NVBug to debug root cause with dmesg, diagnostic error logs and NVIDIA Bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx224	Feature is not supported in the hardware	The tests are running on a non-TH500 product. This error code should not be seen on a TH500 product.	Check the test HW platform and confirm if it is a TH500 based product.	TRIAGE_DEVICE
MODS-xxxxxxxxx612	Invalid value for Tegra configuration variable(s)	Invalid test ID or configuration parameter provided.	1. Check the test command line used and verify the correct test ID. 2. If the issue persists report Software NVBug to debug root cause with dmesg, diagnostic error logs and NVIDIA Bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx774	Tegra Test Fail	The tested IP failed unexpectedly.	1. Refer to individual tests for further actions. 2. If no action is suggested for individual tests, then report Software NVBug to debug root cause with dmesg, diagnostic error logs and NVIDIA Bug report tool logs attached.	REPORT_NV_BUG
MODS-xxxxxxxxx077	Timeout Error	Tests failed with SW timeout issue and needs further investigation.	Report Software NVBug to debug root cause with dmesg, diagnostic error logs and NVIDIA Bug report tool logs attached.	REPORT_NV_BUG
For any other Error Codes		Unexpected test failure	Report Software NVBug to debug root cause with dmesg, diagnostic error logs and NVIDIA Bug report tool logs attached.	REPORT_NV_BUG

4.7 CPU Field Diagnostics Error Code to Action Mapping

For NVIDIA Grace™ CPU related failures while running Partner Diagnostics, look up the error code in the following table and follow the 'Next Actions' column to address the failing system.

Table 4-7. CPU Field Diagnostics Error Code to Action Mapping

Test	Error Code	Error Message	Explanation	Next Action	Data Center Action Category
TegraCpu	MODS-0000000xx087 (xx could be 01, 02, 03 or 13)	Failed while writing a file	SW issue while writing a file.	1. Perform a power cycle and retest. 2. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool.	1. REPORT_NV_BUG 2. PROD_FIT. RESET_DEVICE
	MODS-0000000xx128 (xx could be 01, 02, 03 or 13)	Invalid input	Input to the test is missing, does not match the expected format or parameter validation failed.	1. Retest with the latest Diag. 2. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool.	1. REPORT_NV_BUG 2. PROD_FIT
	MODS-0000000xx008 (xx could be 01, 02, 03 or 13)	Bad parameter passed to function	Bad parameter passed to software function.	1. Retest with the latest Diag. 2. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool.	1. REPORT_NV_BUG 2. PROD_FIT
	MODS-0000000xx010 (xx could be 01, 02, 03 or 13)	Cannot open file	Cannot open specified file in the test.	1. Retest with the latest Diag. 2. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool.	1. REPORT_NV_BUG 2. PROD_FIT
	MODS-0000000xx236 (xx could be 01, 02, 03 or 13)	The expected value and the destination memory value do not match	Data mismatch occurring during the test.	Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE
	MODS-0000000xx149 (xx could be 01, 02, 03 or 13)	Specific test was requested to run but skipped	The specified test was not executed because it was either already in progress or skipped due to other conditions.	1. Check if the same test was launched already and kill it. 2. If the test was not launched previously, perform a power cycle and run the test again. 3. If the issue persists, report software NVBug to debug the root cause with dmesg,	1. REPORT_NV_BUG 2. PROD_FIT. RESET_DEVICE

Test	Error Code	Error Message	Explanation	Next Action	Data Center Action Category
				diagnostic error logs and NVIDIA bug report tool.	
	MODS-0000000xx160 (xx could be 01, 02, 03 or 13)	Resource is reserved by another thread or test	Resource is reserved by another thread or test	1. Perform a power cycle and run the test again. 2. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool	1. REPORT_NV_BUG 2. PROD_FIT. RESET_DEVICE
	MODS-0000000xx077 (xx could be 01, 02, 03 or 13)	Timeout error	Unexpected test timeout or crash	Report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool logs attached.	REPORT_NV_BUG
	MODS-0000000xx012 (xx could be 01, 02, 03 or 13)	Failed while reading a file	SW issue while reading a file.	1. Perform a power cycle and retest. 2. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool.	1. REPORT_NV_BUG 2. PROD_FIT. RESET_DEVICE
	MODS-0000000xx547 (xx could be 01, 02, 03 or 13)	Invalid CPU frequency measured	Invalid CPU Frequency measured during the test.	Report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool logs attached.	REPORT_NV_BUG
	MODS-0000000xx164 (xx could be 01, 02, 03 or 13)	APEI Error detected	Detected APEI Error on Dmesg.	HW error detected. Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE
	MODS-0000000xx039 (xx could be 01, 02, 03 or 13)	Unable to set element	Unable to set the CPU voltage to the specified value. Possible hardware issue with voltage control.	For C1 down design, refer to "CPU Isolation" action. For Grace products, refer to Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE
	MODS-0000000xx009 (xx could be 01, 02, 03 or 13)	Can not allocate memory	Error happens during allocating memory.	1. Refer to the "Free Memory" action. 2. If issue persists, report Software NVBug to debug root cause with dmesg, Diagnostic error logs and NVIDIA bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT
	MODS-0000000xx139 (xx could be 01, 02, 03 or 13)	Thermal Sensor Error	CPU internal temperature limits exceeded acceptable range.	1. Verify the cooling solution is working properly and re-test. Sensors could be broken. 2. If issue persists, start RMA qualification. Refer to "RMA Actions".	TRIAGE_DEVICE

Test	Error Code	Error Message	Explanation	Next Action	Data Center Action Category
	MODS-000000013774	Tegra test fail	CPU test failure during execution.	Report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool logs attached.	REPORT_NV_BUG
	MODS-000000013011	File does not exist	CPU related system files not found in /sys/devices/system/cpu/ path.	1. Enable the kernel configurations related to CPU as per NVIDIA recommendations in "grace-patch-config-guide". 2. Reboot the device and retest. 3. If issue persists, report Software NVBug to debug root cause with dmesg, Diagnostic error logs and NVIDIA bug report tool logs attached.	1. REPORT_NV_BUG 2. RECOVERY. RESET_BM
TegraMemory	MODS-0000000xx236 (xx could be 06 or 12)	The expected value and the destination memory value do not match	Data mismatch occurring during the test.	Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE
	MODS-0000000xx216 (xx could be 06 or 12)	Could not find the specified device	Possible issue with NUMA configuration of the device.	1. Check NUMA kernel configurations in "grace-patch-config-guide". 2. Reboot the device and retest. 3. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool logs attached.	1. REPORT_NV_BUG 2. RECOVERY. RESET_BM
	MODS-0000000xx164 (xx could be 06 or 12)	APEI error	Detected an APEI error in dmesg.	HW error detected. Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE
	MODS-0000000xx009 (xx could be 06 or 12)	Can not allocate memory	Error happens during allocating memory.	1. Refer to the "Free Memory" action. 2. If issue persists, report Software NVBug to debug root cause with dmesg, Diagnostic error logs and NVIDIA bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT
	MODS-0000000xx240 (xx could be 06 or 12)	Unexpected result from HW	Detected HW Error in dmesg.	HW error detected. Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE
	MODS-0000000xx773	Data mismatch	Data written and read back from the memory mismatch.	Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE

Test	Error Code	Error Message	Explanation	Next Action	Data Center Action Category
Clink	MODS-000000005272	Read parameter differs from expected	Clink EOM measurements are below the threshold.	Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE
	MODS-000000005229	Hardware was not initialized	Clink driver is not initialized.	1. Install the recommended clink driver and retest. 2. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT
	MODS-000000005198	NVLink minion error	Issue in transactions with the Clink driver.	1. Reboot the device and retest. 2. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool logs attached.	1. REPORT_NV_BUG 2. RECOVERY. RESET_BM
	MODS-000000005160	Resource is reserved by another thread or test.	Resource is reserved by another thread or test.	1. Perform a power cycle and run the test again. 2. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool	1. REPORT_NV_BUG 2. PROD_FIT. RESET_DEVICE
	MODS-000000005774	Tegra test fail	Clink error flags report error. One or more hardware error flags were detected on the Clink links during the test.	Collect and share the Partner Field Diag logs with NVIDIA. Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE
	MODS-000000005276	HW status error	Clink hardware error counter is greater than 0, or the hardware reports an abnormal status.	Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE
	MODS-000000005773	Data mismatch	Data written and read back from the memory mismatch.	Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE
	MODS-000000005140	NVLink BER bus error	Bit error rate observed on the link is above the acceptable threshold.	Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE
	MODS-000000005014	Bandwidth outside the range	The link bandwidth is below the acceptable threshold.	1. Perform a power cycle. 2. If issue persists, report Software NVBug to debug root cause with dmesg, diagnostic error logs and NVIDIA Bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT. RESET_DEVICE

Test	Error Code	Error Message	Explanation	Next Action	Data Center Action Category
	MODS-000000005009	Can not allocate memory	Error happens during allocating memory.	1. Refer to the "Free Memory" action. 2. If issue persists, report Software NVBug to debug root cause with dmesg, diagnostic error logs and NVIDIA Bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT
C2C	MODS-000000008128	Invalid input	Input to the test is missing , does not match the expected format or parameter validation failed.	1. Retest with latest Diag. 2. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT
	MODS-000000008003	Function is not supported	Secure partition not available.	1. Refer to "Secure Partition" section and try again. 2. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool logs attached.	REPORT_NV_BUG
	MODS-000000008229	Initialization failure	Issue with C2C SW initialization	1. Retest with the latest Diag. 2. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool logs attached.	REPORT_NV_BUG
	MODS-000000005014	Bandwidth outside the range	The link bandwidth is below the acceptable threshold.	1. Perform a power cycle. 2. If issue persists, report Software NVBug to debug root cause with dmesg, diagnostic error logs and NVIDIA Bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT. RESET_DEVICE
	MODS-000000008774	Tegra test fail	Sending FFA message / command failed. Secure partition is unavailable.	1. Refer to "Secure Partition" action. 2. Install the correct version of arm_ffa module and retest. 3. If issue persists, report Software NVBug to debug root cause with dmesg, diagnostic error logs and NVIDIA Bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT
	MODS-000000008276	HW status error	C2C HW hardware error counter is greater than 0,	Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE

Test	Error Code	Error Message	Explanation	Next Action	Data Center Action Category
			or if hardware reports an abnormal status.		
	MODS-000000008773	Data mismatch	Data written and read back from the memory mismatch.	Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE
	MODS-000000008009	Can not allocate memory	Error happens during allocating memory.	1. Refer to the "Free Memory" action. 2. If issue persists, report Software NVBug to debug root cause with dmesg, diagnostic error logs and NVIDIA Bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT
	MODS-000000008284	Unsupported hardware configuration	Secure partition not available.	1. Refer to "Secure Partition" action and try again. 2. If issue persists, report Software NVBug to debug root cause with dmesg, diagnostic error logs and NVIDIA Bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT
PCIe	MODS-000000004014	Bandwidth outside the range	The link bandwidth is below the acceptable threshold.	Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE
	MODS-000000004143	PCIe bus error	AER is above the acceptable threshold.	Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE
	MODS-000000004276	HW reports wrong status	Issue with PCIe link health.	Refer to "RMA Actions".	RETURN_FOR_ FURTHER_TRIAGE
	MODS-000000004229	Hardware was not initialized	Required hardware modules were not initialized.	1. Check if the following kernel configuration setting is enabled and retest: CONFIG_TEGRA_IVC=y 2. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT
	MODS-000000017002	Software error	Invalid test configuration	1. Perform a power cycle and run the test again. 2. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool.	1. REPORT_NV_BUG 2. PROD_FIT. RESET_DEVICE
	MODS-000000017244	Feature is not supported in the hardware	Fail to find the specified PHY lanes.	1. Check if PCI devices are correctly connected to the specified PHY lanes and run the test again.	1. REPORT_NV_BUG 2. PROD_FIT. RESET_DEVICE

Test	Error Code	Error Message	Explanation	Next Action	Data Center Action Category
				2. Perform a power cycle and run the test again. 3. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool.	
	MODS-000000017272	Read parameter differs from expected	EOM validation failed	Report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool logs attached.	REPORT_NV_BUG
	MODS-000000004773	Data mismatch	Data written and read back from the memory mismatch.	Refer to "RMA Actions".	RETURN_FOR_FURTHER_TRIAGE
	MODS-000000004284	Unsupported hardware configuration	Secure partition not available.	1. Refer to "Secure Partition" action and try again. 2. If issue persists , report Software NVBug to debug root cause with dmesg, diagnostic error logs and NVIDIA Bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT
	MODS-000000004009	Can not allocate memory	Error happens during allocating memory.	1. Refer to the "Free Memory" action. 2. If issue persists , report Software NVBug to debug root cause with dmesg, diagnostic error logs and NVIDIA Bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT
	MODS-000000004426	Network error	Randomly selected network port is being used by another process.	1. Rerun the test again. Since the port selected is random, the test should be able to pick a free port. 2. If issue persists , report Software NVBug to debug root cause with dmesg, diagnostic error logs and NVIDIA Bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT
	MODS-000000004143	PCIe bus error	AER is above the acceptable threshold.	Refer to "RMA Actions".	RETURN_FOR_FURTHER_TRIAGE
	MODS-000000004014	Bandwidth outside the range	The link bandwidth is below the acceptable threshold.	Refer to "RMA Actions".	RETURN_FOR_FURTHER_TRIAGE

Test	Error Code	Error Message	Explanation	Next Action	Data Center Action Category
	MODS-000000004272	Unexpected result error	Physical lane EOM status reported is not in the acceptable threshold.	Refer to "RMA Actions".	RETURN_FOR_FURTHER_TRIAGE
Cper	MODS-000000019001	Exit	CPER test command returns non-zero exit code.	Report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool logs attached.	1. REPORT_NV_BUG 2. PROD_FIT
	MODS-000000019010	Cannot open file	Ruleset filename not provided.	Verify the ruleset file is provided and correctly located.	1. REPORT_NV_BUG 2. PROD_FIT
	MODS-000000019458	File operation was interrupted by a signal	The Cper-collect process is terminated by a signal.	1. Retest with the latest Diag. 2. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool logs attached.	REPORT_NV_BUG
	MODS-000000019483	Unknown file error	The exit status of the process is an unknown state that is neither a normal exit nor a signal termination.	1. Retest with the latest Diag. 2. If the issue persists, report software NVBug to debug the root cause with dmesg, diagnostic error logs and NVIDIA bug report tool logs attached.	REPORT_NV_BUG

4.8 Diagnostic Error Codes

The following actions help troubleshoot diagnostic error codes.

4.8.1 Free Memory

1. Check for available free memory and use the 'free' command to get the amount of free memory.
2. If observed low free memory available,
   a. Use 'htop' command to check if there are any high memory consuming processes that can be killed before kicking off the test again.

4.8.2 Secure Partition

1. To enable MODS secure partition, reboot the module.
2. During reboot go into the ‘Setup’ mode.
3. Navigate through Device Manager
   a. NVIDIA Configuration
      i. Grace configuration
         (1) MODS Secure Partition: If checkbox is not checked, then enable it.
         (2) Reboot the module for the MODS Secure Partition to take effect.

4.8.3 RMA Actions

1. Check firmware versions against the latest NVIDIA released firmware versions.
2. Retest with the latest version of the Grace Partner Diagnostics (Field) if not using the latest version.
3. Start RMA qualification if the same error is seen after the previous steps and reruns.

4.8.4 CPU Isolation

1. Further isolation is needed to identify if the error is caused by the CPU or other system components.
2. If isolated to a faulty CPU, refer to “RMA Actions.”

4.8.5 Providing Partner Diagnostic Logs

The diagnostics generates a tarball of the log file called logs-<yyyymmdd-hhnnss>.tgz and places this file in the /logs Partner Diagnostics folder. The log file name is based on the test date and timestamp in the following format: logs/logs-yyyymmdd-hhnnss

where:

- mm = Month
- dd = Day
- yy = Year
- hh = Hours
- nn = Minutes
- ss = Seconds

You can find the plain text logs of the partner diagnostic run in run.log. Additional details on the contents of the logs can be found in *the NVIDIA Baseboard Field Diagnostic Software Guide*.

Chapter 5. Extended Utility Diagnostic Error Codes

When the Extended Utility Diagnostic (EUD) is run, it will show the results of the run on the console and output it to the corresponding .log file. Faulty systems will cause the diagnostic to fail, and it will display 'FAIL' along with error codes and error messages for the failing tests on the console and at the end of the .log file. Extended Utility Diagnostic can be run in two different variants: Opportunistic EUD and GPU EUD Field Diagnostics.

Error codes provide insight into the cause of the failure and could be used to infer steps required to address the failure. Refer to Section 5.1 "GPU EUD Error Code to Action Mapping," to learn more about the common measures to address the failure per error code seen after running the specific variant of EUD.

The "GPU EUD Field Diagnostics" provides exhaustive coverage of the hardware and may qualify faulty parts for RMA. It runs for roughly 10 hours, regressing all the available power states, and extensively tests compute, memory, high-speed I/O, and performance of the system. GPU EUD Field Diagnostics should be used for exhaustive coverage, that is when the hardware is intermittently failing and must repair or qualify faulty hardware for RMA.

The Opportunistic EUD runs for roughly 30 minutes and targets a small set of tests. It can be run in between workloads to check the system's health. The Opportunistic EUD will catch failures related to silent data corruption.

Look up the error code in Table 5-2 and follow the "Next Steps" column to address the failing system.

5.1 GPU EUD Error Code to Action Mapping

Look up the error code in Table 5-2 and follow the "Next Steps" column to address the failing system. The GPU EUD will catch failures related to silent data corruption and identify parts for the RMA process. The Data Center Action Category corresponding to each error code is also included with the corresponding action description in Table 5-1.

Table 5-1. Data Center Action to Data Center Description Mapping – GPU EUD

Data Center Action Category	Data Center Action Description
N/A	Not applicable.
PROD_FIT	System is production ready if the error is not seen again.
PROD_FIT.RESET_DEVICE	Reset failing device aka hot reset failing device. The system is then production ready.
RECOVERY	Running additional tools can get the system back in production.
RECOVERY.RESET_BM	Restart bare metal aka cold reboot the system. Running additional tools can get the system back in production.
RECOVERY.RESET_SW	Warm reboot the system aka reload latest supported installed firmware / driver. Running additional tools can get the system back in production.
RECOVERY.RUN_INFOROM	Run InfoROM recovery tool. Running additional tools can get the system back in production.
REPORT_NV_BUG	Report Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached.
RMA_BOARD	Take the failing board out of production and start RMA qualification for failing board.
RETURN_FOR_FURTHER_TRIAGE	Take the failing tray or board out of production and return the failing tray or board to the integrator for further triage and possible RMA.
RMA_DEVICE	For NVIDIA Hopper™ products, take the failing device out of production and start RMA qualification for failing device. For Blackwell products, take the failing board out of production and start RMA qualification for failing device. Note that RMA decisions must involve the integrator. Datacenter actions may involve returning the component to the integrator and may possibly translate to RMA following further testing and investigation done by the integrator.
TRIAGE_BOARD	Take the failing board out of production and follow next action steps for that error code. If qualified for RMA, RMA the board. To RMA the board, return the board to the integrator for further triage.
TRIAGE_DEVICE	For NVIDIA Hopper products, take the failing device out of production and follow next action steps for that error code. If qualified for RMA, RMA the device. For Blackwell products, take the failing board out of production and follow next action steps for that error code. If qualified for RMA, RMA the device

Table 5-2. GPU EUD Error Code to Action Mapping

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
xxxxxxxxx002	software error	An unexpected software error has occurred.	1. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after running GPU EUD.	1. REPORT_NV_BUG 2. PROD_FIT
xxxxxxxxx008	Bad parameter passed to function	Software error / bad parameter passed to software function.	1. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log if the same error code is seen after running GPU EUD.	1. REPORT_NV_BUG 2. PROD_FIT
xxxxxxxxx021	script failed to execute	File Software NVBug and attach dmesg, diagnostic error logs, NVBug report tool log, to debug and root cause.	1. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 2. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log, if the same error code is seen after running GPU EUD.	1. REPORT_NV_BUG 2. PROD_FIT
xxxxxxxxx060	NVRM invalid param struct	This could be caused by varied NVRM assertion failure. File Software NVBug and attach dmesg, diagnostic error logs, NVBug report tool log, to debug and root cause.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU EUD run.	1. REPORT_NV_BUG 2. PROD_FIT.RESET_DEVICE
xxxxxxxxx077	timeout error	Follow the triage flow to diagnose/repair root cause.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest GPU EUD.	1. TRIAGE_DEVICE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			3. Follow Triage Flow (NVOnline: 1091416) if the same error code is seen after the GPU EUD run.	
xxxxxxxxx083	CRC/Checksum miscompare	Computation has returned an incorrect answer.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. Start RMA qualification if the same error code is seen after the GPU EUD run.	1. RETURN_FOR_ FURTHER_TRIAGE
xxxxxxxxx097	Unexpected device interrupts	Diag received an unexpected interrupt that it considers an error. It could be a MMU fault during a CUDA test, which is indicative of a GPU hardware issue.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. Start RMA qualification if the same error code is seen after the GPU EUD run.	1. RETURN_FOR_ FURTHER_TRIAGE
xxxxxxxxx124	Invalid InfoROM	InfoROM corruption has occurred. Recover InfoROM by running InfoROM recovery tool.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Run InfoROM recovery tool. For Blackwell generation of products, please refer to *InfoROM Cleansing for RMA - User Guide* (NVOnline: 1120261). 3. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 4. Start RMA qualification if the same error code is seen after the GPU EUD run.	1. TRIAGE_DEVICE
xxxxxxxxx127	Vbios Certificate Error	An issue with the VBIOS certificate.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest version. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU EUD run.	1. RECOVERY.RESET_SW 2. PROD_FIT

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
xxxxxxxxx139	Acceptable temperature limits exceeded or the thermal sensor is broken or miscalibrated	Possible issue with cooling systems. Check cooling systems and engage with NVIDIA. Failing part qualifies for RMA only if the problem is isolated to NVIDIA hardware.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest version. 3. Start RMA qualification if the same error code is seen after the run and only if the cooling solution is part of NVIDIA hardware.	1. TRIAGE_DEVICE
xxxxxxxxx140	NvLink bus error	NVLink is down or NVLink errors occurred during NVLink testing. Follow to triage flow if same error code is seen after running latest GPU EUD.	1. Reseat the failing GPU. 2. Ensure GPU and NVSwitches are detected on the system using lspci utility. 3. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 4. Follow RMA Triage Flow (NVOnline: 1091416) if the same error code is seen after the GPU EUD run.	1. TRIAGE_DEVICE
xxxxxxxxx143	PCI Express bus error	PCIe link errors cause PCIe tests to fail.	1. Reseat the failing GPU 2. Ensure GPU and NVSwitches are detected on the system using lspci utility. 3. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 4. Follow RMA Triage Flow (NVOnline: 1091416) if the same error code is seen after the GPU EUD run.	1. TRIAGE_DEVICE
xxxxxxxxx144	CUDA error	One of the CUDA tests is failing. File Software NVBug and attach dmesg, diagnostic error logs, NVBug report tool log, to debug and root cause.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU EUD run.	1. TRIAGE_DEVICE
xxxxxxxxx194	bad memory	Computation has returned an incorrect answer.	1. Check firmware versions against the latest NVIDIA release firmware versions.	1. RETURN_FOR_ FURTHER_TRIAGE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. Start RMA qualification if the same error code is seen after the GPU EUD run.	
xxxxxxxxx220	PCIe device not found	Failure with PCIe device. Need to isolate the failure to GPU or baseboard.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Reseat failing GPU. 3. Ensure GPU and NVSwitches are detected on the system using lspci utility. 4. Perform cold reboot i.e., full system shutdown and manually reboot again. 5. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 6. Start RMA qualification if the same error code is seen after the GPU EUD run.	1. RECOVERY.RESET_BM 2. PROD_FIT
xxxxxxxxx229	Hardware was not initialized	Something prevented successful hardware initialization.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. Follow Triage Flow (NVOnline: 1091416) if the same error code is seen after the GPU EUD run.	1. TRIAGE_DEVICE
xxxxxxxxx240	Unexpected result from HW	If running GPU EUD on a single GPU, make sure the 'skip_nvlink' parameter is used. If it is, then there is a possible setup issue, confirm the NVLink topology is connected as expected by running: nvidia-smi nvlink -s If NVLink connections are reported as inactive, debug NVLink issue.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. Follow RMA Triage Flow (NVOnline: 1091416) if the same error code is seen after the GPU EUD run.	1. TRIAGE_DEVICE
xxxxxxxxx272	Read parameter	File Software NV bug and attach dmesg, diagnostic error logs,	1. Check firmware versions against the latest NVIDIA release firmware versions.	1. TRIAGE_DEVICE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
	differs from expected	NV bug report tool log, to debug root cause.	2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the run.	
xxxxxxxxx276	HW reports wrong status	Hardware failure has occurred.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. Start RMA qualification if the same error code is seen after the run.	1. RETURN_FOR_ FURTHER_TRIAGE
MODS-xxxxxxxxx280	Power is above specified limit	Power is too high	1. Re-test with latest GPU EUD if not already tested with the latest GPU EUD. 2. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
MODS-xxxxxxxxx288	Power is below specified limit	Power is too low	1. Re-test with latest GPU EUD if not already tested with the latest GPU EUD. 2. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached if the same error code is seen after the run.	1. REPORT_NV_BUG 2. PROD_FIT
xxxxxxxxx311	NVLink discovered topology does not match required topology	NVLink topology does not match expected topology.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Ensure GPU and NVSwitches are detected on the system using lspci utility. 3. Perform cold reboot i.e., full system shutdown and manually reboot again. 4. Reseat failing GPU if reboot did not resolve the issue.	1. TRIAGE_DEVICE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			5. Re-install NVIDIA release driver if rebooting and reseating the GPU did not resolve the issue. 6. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 7. Run HGX diagnostic mentioned in the Triage Flow (NVOnline: 1091416) if the same error code is seen after the GPU EUD run.	
xxxxxxxxx317	ECC detected an uncorrectable error in FB	Uncorrectable error seen in the frame buffer, reset the GPU for row remapping to take place, follow the triage flow if same failure signature is seen after retesting with latest GPU EUD.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Power cycle the GPU to trigger row remapping. 3. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 4. Follow RMA Triage Flow (NVOnline: 1091416) if the same error code is seen after the GPU EUD run.	1. TRIAGE_DEVICE
xxxxxxxxx319	ECC detected an uncorrectable error in L2	Uncorrectable error seen in the L2 cache, reset for row remapping to take place, then retest with latest GPU FD.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Power cycle GPU to trigger row remapping. 3. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 4. Start RMA qualification if the same error code is seen after the GPU EUD run.	1. TRIAGE_DEVICE
xxxxxx171321	ECC detected an uncorrectable error in SM	Uncorrectable error seen in the SM while running T171, reset the GPU for row remapping to take place, follow the triage flow if same failure signature is seen after retesting with latest GPU EUD.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Power cycle GPU to trigger row remapping. 3. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 4. Start RMA qualification if the same error code is seen after the GPU EUD run.	1. RETURN_FOR_ FURTHER_TRIAGE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
xxxxxxxxx320	ECC detected a correctable error in SM, threshold exceeded	Rate of correctable errors seen in the SM have exceeded threshold. Follow the RMA triage flow if same failure signature is seen after re-testing with latest GPU EUD.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Power cycle GPU to trigger row remapping. 3. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 4. Follow Triage Flow (NVOnline: 1091416) if the same error code is seen after the run.	1. TRIAGE_DEVICE
xxxxxxxxx321	ECC detected an uncorrectable error in SM	Uncorrectable error seen in the SM. Follow the RMA triage flow if same failure signature is seen after retesting with latest GPU EUD.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Power cycle GPU to trigger row remapping. 3. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 4. Follow Triage Flow (NVOnline: 1091416) if the same error code is seen after the run.	1. TRIAGE_DEVICE
xxxxxxxxx341	Buffer mismatch	Start RMA qualification for baseboard including the exact SXMs.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD . 3. Start RMA qualification if the same error code is seen after the run.	1. RETURN_FOR_ FURTHER_TRIAGE
xxxxxxxxx363	Row remapping failed	A row remapping failure has occurred.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. Start RMA qualification if the same error code is seen after the GPU EUD run.	1. RETURN_FOR_ FURTHER_TRIAGE
xxxxxxxxx539	NVRM generic falcon error	This could be caused by various NVRM generic failures. File Software NVBug and attach dmesg, diagnostic error logs and NVBug report tool log, to debug and root cause.	1. Check firmware versions against the latest NVIDIA release firmware versions.2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA	1. REPORT_NV_BUG 2. PROD_FIT.RESET_DEVICE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
			Bug report tool log, if the same error code is seen after running GPU EUD.	
xxxxxxxxx541	NVRM detected memory error	This could be caused by various NVRM memory failures. File Software NVBug and attach dmesg, diagnostic error logs and NVBug report tool log, to debug and root cause.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log, if the same error code is seen after running GPU EUD.	1. TRIAGE_DEVICE
xxxxxxxxx542	NVRM VBIOS invalid or rejected	NVRM VBIOS validation failure. File Software NVBug and attach dmesg, diagnostic error logs, NV bug report tool log attached to debug root cause.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU EUD run.	1. REPORT_NV_BUG 2. PROD_FIT.RESET_DEVICE
xxxxxxxxx582	GPU stress test found pixel miscompares	Computation has returned an incorrect answer.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. Start RMA qualification if the same error code is seen after the run.	1. RETURN_FOR_ FURTHER_TRIAGE
xxxxxxxxx599	Fan does not seem to cool the chip	Possibly an issue with cooling systems. Failing part qualifies for RMA only if the problem is isolated to NVIDIA hardware.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. Start RMA qualification if the same error code is seen after the run and only if the cooling solution is part of NVIDIA hardware.	1. TRIAGE_DEVICE

Error Code	Error Message	Explanation	Next Steps	Data Center Action Category
xxxxxxxxx614	Extra golden code miscompare	One of the EUD tests failed the consistency check on expected values.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. Start RMA qualification if the same error code is seen after the GPU EUD run.	1. RETURN_FOR_ FURTHER_TRIAGE
xxxxxxxxx679	NVRM invalid argument	This could be caused by various NVRM assertion failures. File Software NVBug and attach dmesg, diagnostic error logs and NVBug report tool log, to debug and root cause.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU EUD run.	1. REPORT_NV_BUG 2. PROD_FIT. RESET_DEVICE
xxxxxxxxx779	Voltage value out of range	The input voltage supply to the GPU is out of specification. Please check the system power supply and try again.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Check the input voltage to the GPU and the system power supply. 3. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 4. File Software NVBug to debug root cause with dmesg, diagnostic error logs, and NVIDIA Bug report tool log attached, if the same error code is seen after the GPU EUD run.	1. REPORT_NV_BUG 2. PROD_FIT. RESET_DEVICE
xxxxxxxxx818	MODS detected an assertion failure	This could be the result of unexpected chip configuration, programming error, hardware fault or failure. File Software NVBug and attach dmesg, diagnostic error logs, NVBug report tool log to debug and root cause.	1. Check firmware versions against the latest NVIDIA release firmware versions. 2. Re-test with latest GPU EUD if not tested with the latest GPU EUD. 3. File Software NVBug to debug and root cause with dmesg, diagnostic tool logs, and NVIDIA Bug report tool log attached, if the same error code is seen after running GPU EUD.	1. REPORT_NV_BUG 2. PROD_FIT

5.2 CPU EUD Related Error Code to Action Mapping

For CPU EUD related failures, look up the error code in Table 5-3 and follow the **Next Steps** column to address the failing system.

Table 5-3. CPU EUD Related Error Code to Action Mapping

Error Code	Error Message	Explanation	Next Steps
xxxxxxxxx002	Software Error	Unexpected software errors	Re-test with the latest version of EUD. If it still fails, file a Software NVBug with log files.
xxxxxxxxx005	Bad Command Line Argument	Input argument is not supported, or the test suite name is incorrect	Check the supported argument and try again. It supports the following test suite names: - RmaFull - Opportunistic - Trickle
xxxxxxxxx008	Bad Parameter	[2,3,4,8,9,11] Invalid parameter passed to function	[2,3,4,8,9,11] Re-test with the latest version of CPU EUD. If it still fails, file a Software NVBug with log files.
xxxxxxxxx009	Cannot Allocate Memory	[9] Failed to allocate memory	[9] Ensure free memory size is sufficient for test.
xxxxxxxxx012	File Read Error	[1] Tarball file for test is damaged	[1] Re-download the EUD package and try again.
xxxxxxxxx023	Could Not Compile File	[1] Failed to compile file	[1] Check if gcc and md5sum are installed on the system, or check if CPU hardware is damaged.
xxxxxxxxx077	Timeout Error	Test timeout	Collect and share the logs with NVIDIA.
xxxxxxxxx084	Cannot Parse File	[11] Config file parse error	[11] Re-test with the latest version of EUD. If it still fails, file a Software NVBug with log files.
xxxxxxxxx087	File Write Error	[11] Failed while writing config file	[11] Check free disk space and ensure file can be written.
xxxxxxxxx128	Invalid Input	[5] Incorrect input file causes exec pipe error [7] Failed to read from lspci command [9] Failed to get pattern data or test mode	[5,7,9] Check the test HW platform and re-test with the latest version of EUD.
xxxxxxxxx130	Invalid Test Input	[11] Attempting to map inaccessible Cpu coreId	[11] Re-test with the latest version of EUD. If it still fails, file a Software NVBug with log files.

Error Code	Error Message	Explanation	Next Steps
xxxxxxxxx139	Thermal Sensor Error	[11] Acceptable temperature limits exceeded the threshold, or the thermal sensor is broken or is not calibrated	[11] Check the fan or thermal sensor status.
xxxxxxxxx149	Specific Test Not Run	[1-8] The test is already running, so skip this request is skipped	[1-8] Ensure only one EUD instance is running.
xxxxxxxxx150	No Test Run	No tests were run	Assign valid test IDs to run the tests.
xxxxxxxxx160	Resource In Use	[1,4,6,8] No thread is created for the test	[1,4,6,8] Test again after the hardware resource is available.
xxxxxxxxx164	APEI Error	APEI error detected	Check the APEI info from dmesg for more details.
xxxxxxxxx216	Device Not Found	[4,6,8] Cannot find available memory node. Try running test 8 on the system with single NUMA node	[4,6,8] Check the system configuration for NUMA.
xxxxxxxxx220	PCI Device Not Found	[7] Cannot find PCIe device	[7] Check PCIe device installation state.
xxxxxxxxx236	Memory to Memory Result Not Match	[4,6,8] The expected value and the destination memory value do not match	[4,8] Check if DRAM hardware is damaged. [6] Check if CPU caches (L1/L2/L3) is damaged.
xxxxxxxxx229	Hardware Was Not Initialized	[11] Init function has not been called before testing	[11] Re-test with the latest version of EUD. If it still fails, file a Software NVBug with log files.
xxxxxxxxx240	Hardware Error	Memory ECC errors are detected	Check if hardware is damaged.
xxxxxxxxx276	Hardware Status Error	[7] Mismatch between PCIe host & device [10] PCIe link retraining failed	[7] Check PCIe device installation state. [10] Check PCIe device hardware status.
xxxxxxxxx510	Test Cannot Run	The tests are running on a non-TH500 product	Check the test HW platform.
xxxxxxxxx612	Tegra Cfg Invalid Value	[11] Invalid value for configuration variables	[11] Re-test with the latest version of EUD. If it still fails, file a Software NVBug with log files.
xxxxxxxxx709	Unsupported Device	Test is not supported on this system	Check the test HW platform.
xxxxxxxxx719	Parameter Mismatch	[5] Expected system configurations do not match	[5] Check the system configuration.

Error Code	Error Message	Explanation	Next Steps
xxxxxxxxx773	Data Mismatch	[1] MD5 values are different between CPU cores [5] Expected system configurations do not match [9] Pattern data and input data are different	[1] Check if CPU hardware is damaged. [5] Check the system configuration. [9] Check if DRAM hardware is damaged.
xxxxxxxxx774	Tegra Test Failure	[11] Decompressed compiling data failed	[11] Check if tarball is installed on the system or check if CPU hardware is damaged.

5.2.1 List of Tests

The following list corresponds to the tests mentioned in Table 5-3.

- [1] Compile Test
- [2] CPU Stress Test
- [3] CPU Stress Test with Pulse Loading
- [4] Basic Memory Stress Test
- [5] Inventory Test
- [6] CPU Cache Stress Test
- [7] PCIe Link Cap Status Check Test
- [8] Numa Memory Stress Test
- [9] Full Memory Sweep Test
- [10] PCIe Link Retraining Test
- [11] CPU DVFS Workload Test

Chapter 6. ECC and Error Containment

This chapter describes the different ECC and error containment scenarios. Debug steps are provided to triage reported ECC-related issues. If applicable, in-band and out-of-band telemetry were identified to assist partners to triage the associated error scenario.

Internal data corruption in DRAM (or bit flips) can crash or cause data corruption in workloads. The bit flips can be caused by electrical or magnetic interference inside a computer system. At a higher altitude, there is also a higher possibility of bit flips through neutron ray.

With Error Correction Code (ECC) memory enabled, this can correct single-bit errors and flag double bit errors. This is in exchange for a small amount of memory reserved for ECC purposes.

ECC error containment in NVIDIA data center GPUs attributes errors to specific GPU processes and instances (MIG) and prevents errors from impacting other processes or instances. With row remapping, a hardware-based process to replace a faulty memory row with a spare one. No holes in physical address space are visible to any software.

The following table lists different ECC error sources, GPU state, and recommended action that should be taken by partners.
For further information on NVIDIA GPU Memory Error Management, visit https://docs.nvidia.com/deploy/a100-gpu-mem-error-mgmt/index.html.

Table 6-1. ECC Error Containment

Error Source	GPU State	Flag	Recommended Action
Multi-bit errors	Dynamically offline affected memory. GPU is fully functional	Row remapping pending bit asserted.	Continue running jobs on the GPU / launching new jobs
Multi-bit errors affecting entire GPU	Entire GPU impacted; no new CUDA jobs can be launched	Reset pending	Immediate GPU reset
Multi-bit errors affecting selected MIG partition	One MIG partition is impacted and at least 1/7 GPU is not functional. Other MIG partitions continue running workloads	Drain and reset	Drain work on other GPU partitions and reset the GPU

Chapter 7. Thermal Issues

The following sections describe the different GPU thermal issue scenarios. Debug steps are provided to triage reported thermal issues. If applicable, in-band and out-of-band telemetry were identified to assist partners to triage the associated error scenario. Refer to Section 12.5 “Thermal Issue” in Chapter 12 “Debug and Triage.”

7.1 Thermal Hardware Slowdown

Thermal failsafe controller will initiate thermal hardware slowdown when the GPU temperature has reached the thermal slowdown temperature (T_{LIMIT} is equal or less than -2). The GPU clock is cut in half (2x hardware slowdown) in hardware to prevent thermal runaway. Hardware slowdown (0x08) reason code will be generated when the NVIDIA GPU Boost™ clock optimization is unable to contain the GPU temperature to within the maximum operating temperature and triggers a 2x hardware slowdown to the GPU clock.

7.2 Thermal Hardware Shutdown and OVERT Event

When a GPU over-temperature event occurs, a corresponding **THERM_OVERT** signal will be asserted to alert the user of the over-temperature condition. The **THERM_OVERT** is a signal driven by the onboard FPGA to notify BMC of the thermal event. System BMC should check which device has exceeded its maximum temperature by using “SMBPBI Opcode B0H – Get OVERT Info” and log the information. Once the OVERT reason has been logged, system BMC is expected to shut off power to the baseboard within 1 second of the interrupt assertion to avoid component damage.

7.3 Thermal Issue Caused by TIM and Thermal Solutions for Partner Cooled SKUs

When a partner cooled GPU over temperature event occurs, partners are advised to inspect and review their TIM and thermal solutions installation. Partners are advised to review their die pressure and TIM specifications for the specific GPU products. For NVIDIA HGX products, partners can use the “HGX Thermal Solutions Checklist” to verify the NVIDIA specifications against their TIM and thermal solutions.

Chapter 8. System Power Checks

The following sections describe the different GPU power issue scenarios. Debug steps are provided to triage reported power issues. If applicable, in-band and out-of-band telemetry were identified to assist partners to triage the associated error scenario.

8.1 Unable to Reach Maximum TGP

If the GPU is unable to reach the maximum TGP defined for the product, users can use the NVSMI utility to check both the enforced power limit for the GPU:

```
nvidia-smi -q -d POWER
```

The enforced power limit for the GPU should match the default power limit. If the enforced power limit is lower than the default power limit, the GPU is considered power capped and therefore will not be allowed to reach the maximum TGP. Users can reset the enforced power limit to the default power limit by using the NVSMI utility:

```
nvidia-smi -pl [power limit]
```


8.2 Unable to Reach Maximum GPU Clocks

If the GPU is unable to reach the maximum GPU clocks defined for the product, users can use the NVSMI utility to check both the application clock limits for the GPU:

```
nvidia-smi -q -d CLOCK
```

If the maximum clocks are lower than the maximum clocks allowed for the product, users can reset the clock by using the NVSMI utility:

```
nvidia-smi -rgc
```

Furthermore, if Programmable EDP was reduced by user, this may also limit the maximum clocks that the product is able to reach. To check and to reset the current

Programmable EDP setting, user can either access through the SMBPBI or NVFlash utility. In addition, on the NVIDIA Hopper HGX 8-GPU, Redfish can also be used to get and to set the EDPp scaling factor.

8.2.1 Redfish for Programmable EDPp

The Redfish for Programmable EDPp can be used to get and to set the Programmable EDP set point. This command allows users to query and control the programmable EDPp thresholds by specifying a percentage scaling value. This can be used to program an EDPp limit without necessarily lowering the current TGP limit. IP address and exact Redfish may be different from the following examples.

```
# To read back current value “EDPpPercent” curl -X GET
http://192.168.31.1/redfish/v1/Systems/HGX_Baseboard_0/Processors/GPU_SXM_1/EnvironmentMetrics
```

```
# To set programmable EDP to 100% curl -X PATCH
http://192.168.31.1/redfish/v1/Systems/HGX_Baseboard_0/Processors/GPU_SXM_1/EnvironmentMetrics -d
'{ "Oem": { "Nvidia": {"EDPpPercent": { "SetPoint": 100} } } }'
```

8.2.2 SMBPBI for Programmable EDPp

The Programmable EDPp can be set using an OOB channel SMBPBI protocol. This command allows users to query and control the programmable EDPp thresholds by specifying a percentage scaling value. This can be used to program an EDPp limit without necessarily lowering the current TGP limit.

Table 8-1. SMBPBI for Programmable EDPp Commands

Specification	Value
Opcode	10h – Submit/Poll Asynchronous Request
Arg1	11h – Get/Set Programmable EDPp scaling factor
Arg2	Refer to the NVIDIA SMBus Post-Box interface for GPUs NVOnline: 1001187

8.2.3 NVFlash for Programmable EDPp

NVFlash can be used to change the programmable EDPp scaling factor using the following set of commands. The following example shows how to configure power policy limits for GPU0. The change through NVFlash is persistent across reboots.

To change the programmable EDPp scaling factor, progMultRated setting for the TGP power policy must be specified as a percentage in decimal format.

To add a power policy limit in milliwatts:

```
/> ./nvflash --addpp TGP progMultRated <value>

Where <value> is,
    value = 0.9 for 10% reduction in EDP
    value = 0.7 for 30% reduction in EDP
```

To delete a power policy limit for programmable EDPp:

```
/> ./nvflash --delpp TGP progMultRated
```

8.3 Checking Power Brake

Power brake is used to reduce peak and continuous power draw during emergency situations by reducing NVSwitch bandwidth to 1/4 and GPU clock to 25% of normal clock. Power brake can be asserted by using an external signal **PWR_BRAKE_N** or through the baseboard FPGA. If the GPU or NVSwitch is unable to reach the target clock or bandwidth, partners should confirm if Power Brake status is asserted using SMBPBI Opcode 0xF5 “Get Power Brake State” by reviewing the clock throttle reason 0x80 “HW Power Brake Slowdown.” Or in-band with nvidia-smi -q for the “HW Power Brake Slowdown” field.

Chapter 9. PCI Express Hardware Error and GPU Falling off the Bus Scenarios

The following table shows the different “PCIe Hardware Error” scenarios. Available in-band and out-of-band telemetry can be used to correlate the PCIe error report with an error scenario. Depending on the specific error scenario, the SMBPBI server may not be operational or may be partially available.

Table 9-1. PCIe Error Scenarios

Error Scenario	SMBPBI Server State	In-Band Telemetry	Out-of-Band Telemetry
PCIe Surprise Link Down (SLD) without Downstream Port Containment (DPC) This case covers multiple error scenarios related to SXM VR/Cap fault that causes the GPU to shut down	Server not operational	Link status of DSP of PCIe Switch connected to GPU Link status of the PCIe Retimer connected to the GPU PCIe AER message for SLD in dmesg log XID 79 message in dmesg log	Link status of the PCIe Retimer connected to the GPU (Instructions in Section 14.1 “Checking PCIe Retimer Link Status Corresponding to SXM”) PCIe SLD event logged in BMC
PCIe Link Down due to Downstream Port Containment (if enabled)	Server not operational	Link status of DSP of PCIe Switch connected to GPU Link status of the PCIe Retimer connected to the GPU PCIe DPC message in dmesg log XID 79 message in dmesg log	Link status of the PCIe Retimer connected to the GPU (Instructions in Section 14.1 “Checking PCIe Retimer Link Status Corresponding to SXM”) PCIe DPC event logged in BMC System Event Log (SEL) GPU PCIe Correctable and Uncorrectable Error Count logged in BMC System Event Log (if applicable)
GPU shutdown due to THERM_OVERT	Server not operational	Link status of DSP of PCIe Switch connected to GPU Link status of the PCIe Retimer connected to the GPU XID 79 message in dmesg log	Link status of the PCIe Retimer connected to the GPU (Instructions in Section 14.1 “Checking PCIe Retimer Link Status Corresponding to SXM”)

Error Scenario	SMBPBI Server State	In-Band Telemetry	Out-of-Band Telemetry
			THERM_OVERT event in BMC System Event Log (SEL) SXM THERM_OVERT status (Instructions in Section 14.1 “Checking PCIe Retimer Link Status Corresponding to SXM”)
GPU shutdown for unknown reason	Server not operational	Link status of DSP of PCIe Switch connected to GPU Link status of the PCIe Retimer connected to the GPU XID 79 message in dmesg log	Read SXM presence, power good and reset status (Instructions in Section 14.2 “Checking for SXM PWR_GOOD Status”) Link status of the PCIe Retimer connected to the GPU (Instructions in Section 14.1 “Checking PCIe Retimer Link Status Corresponding to SXM”) FPGA register table dump (Instructions in Section 14.5 “Downloading FPGA Register Table”) ERoT logs (Instructions in Section 14.6 “Downloading ERoT Logs”) PCIe Retimer LTSSM logs (Instructions in Section 14.7 “Downloading PCIe Retimer LTSSM Log”)
GPU VR issue	Server not operational	Link status of DSP of PCIe Switch connected to GPU Link status of the PCIe Retimer connected to the GPU XID 79 message in dmesg log (if MPS issue was triggered at runtime)	HMC Event for GPU Power Good state change/fault (Instructions in Section 14.2 “Checking for SXM PWR_GOOD Status”) Host BMC System Event Log (SEL) with I2C1_ALERT triggered and pointing to abnormal PWR_GOOD state change (This alert will not be cleared until power good is achieved). Link status of the PCIe Retimer connected to the GPU (Instructions in Section 14.1 “Checking PCIe Retimer Link Status Corresponding to SXM”). Check for SXM PWR_GOOD status (Instructions in Section 14.2 “Checking for SXM PWR_GOOD Status”)
GPU in recovery mode	Server not operational	PCIe sub-system device ID changes when the GPU goes into recovery mode DEVICE_ID in recovery mode is 0x2300	Link status of the PCIe Retimer connected to the GPU (Instructions in Section 14.1 “Checking PCIe Retimer Link Status Corresponding to SXM”)

Error Scenario	SMBPBI Server State	In-Band Telemetry	Out-of-Band Telemetry
		FLR, MSI-X, and SR-IOV capabilities are disabled and not advertised in the PCIe Capabilities Register	SPDM Index 43 provides information regarding to AP Boot Status, which can be used to determine AP state.
ERoT fault	Server not operational	Host OS will not be able to discover any GPU and no XID will be reported	Check for ERoT Fatal Error (Instructions in Section 14.4 “Checking for ERoT Fatal Error Status”) Check for SXM PWR_GOOD status (Instructions in Section 14.2 “Checking for SXM PWR_GOOD Status”) Link status of the PCIe Retimer connected to the GPU (Instructions in Section 14.1 “Checking PCIe Retimer Link Status Corresponding to SXM”)
ERoT keeping GPU in reset state	Server not operational	Host OS will not be able to discovery any GPU and no XID will be reported	Check for ERoT Fatal Error (Instructions in Section 14.4 “Checking for ERoT Fatal Error Status”) SPDM Index 43 provides information regarding to AP Boot Status, which can be used to determine AP state. ERoT logs (Instructions in Section 14.6 “Downloading ERoT Logs”) FPGA register table dump (Instructions in the “Downloading FPGA Register Table” section)
Retimer failures	SMBPBI may be operational	If the failure happens at runtime: XID 79 message in dmesg log PCIe AER message in dmesg log If the failure happens at boot time: Host OS will not be able to discover any GPU and no XID will be reported nvidia-bug-report to capture host OS state	PCIe Retimer LTSSM logs (Instructions in Section 14.7 “Downloading PCIe Retimer LTSSM Log”) GPU PCIe Correctable and Uncorrectable Error Count logged in BMC System Event Log (if applicable)

Chapter 10. NVLink Error Scenarios

The following table shows the different “NVLink Hardware Error” scenarios. Available in-band and out-of-band telemetry can be used to correlate the XID error report with an error scenario. “NVLink Error” can be caused by NVLink initialization faults, incorrect software configuration, unexpected internal hardware faults, and poor signal quality on NVLink.

For further information on NVLink health monitoring, refer to the *NVIDIA NVLink Health Monitoring Application Note* (NVOnline: 1087258).

Note: Content within this chapter is mainly applicable to NVIDIA systems using NVLink4 and older generations of NVLink.

Table 10-1. NVLink Error Debug and Triage Scenarios

Error Scenario	In-Band Telemetry	Out-of-Band Telemetry
NVLink Initialization Fault	“NVLink: Failed to train link # to remote PCI:B.D.F” message in dmesg log. NVLink state using NVSMI/DCGM Inspect kernel logs for any system errors before NVLink initialization failure.	NVLink state through SMBPBI
Incorrect Software Configuration	“NVLink MINION: link # interrupts disabled due to fatal MINON error” message in dmesg log. NVLink state using NVSMI/DCGM Inspect kernel logs for any system errors before XID 74 error message.	NVLink state through SMBPBI
Unexpected Internal Hardware Fault	“NVLink: fatal error detect on link #” message in dmesg log. NVLink state using NVSMI/DCGM NVLink Flit CRC, Replay, Recovery, and CRC error counters using NVSMI/DCGM.	NVLink state through SMBPBI NVLink Flit CRC, Replay, Recovery and CRC error counters via SMBPBI.

Error Scenario	In-Band Telemetry	Out-of-Band Telemetry
	Inspect kernel logs for any system errors before SXID and XID error messages.	
Poor Signal Quality	SXID and XID error message in dmesg log. NVLink Flit CRC, Replay, Recovery, and error counters using NVSMI/DCGM NVLink state using NVSMI/DCGM	NVLink state through SMBPBI NVLink Flit CRC, Replay, Recovery and CRC error counters via SMBPBI.
NVLink ECC Correction	Drivers provide ECC correction counts across each NVLink lane. ECC is a transparent BER correction scheme that has no impact on performance. The information contained within these registers are for informational purposes only, and high counts within these registers do not suggest a link issue, as any data that escapes ECC will get flagged as FLIT CRC. NVLink devices that exhibit an excessively high rate of ECC correction will trigger an SXID 20009 in dmesg. When encountering this non-fatal SXID, partner is advised to run additional diagnostic tests (GPU Field Diagnostic) to determine if there is an underlying hardware problem.	NVLink state through SMBPBI NVLink Flit CRC, Replay, Recovery and Data CRC error counters via SMBPBI.

Chapter 11. Checking NVLink State with In-Band and Out-of-Band Tools

This chapter provides partners and customers instructions on how to obtain the in-band telemetry using NVSMI and DCGM commands. The NVLink telemetry can be used to assist in the analysis and triage of XID and SXID error scenarios described in the previous chapter.

Note: Content within this chapter is mainly applicable to NVIDIA systems using NVLink4 and older generations of NVLink.

11.1 Checking for NVLink State

The following NVSMI and DCGM command can be used to check the NVLink state status. The NVLink status information can be used for further analysis of NVLink related error scenarios.

11.1.1 NVSMI

1. User can run the following command to check the NVLink state.
   ```
   nvidia-smi nvlink -s
   ```
2. Checking return status:
   ```
   GPU<#>: NVIDIA Graphics Device (UUID: <GPU # UUID>)
   Link<#>: <Link speed GB/s>
   ```

If NVLink is up, the maximum link speed will be reported.

If the link is down, the link speed will be reported as inactive.

11.1.2 DCGM

1. User can run the following command to check the NVLink state.

```
dcgmi nvlink --link-status
```

2. Checking return status:

```
+----------------------+
|  NvLink Link Status  |
+----------------------+
GPUs:
    gpuId <X>:
        <U or D for each link>
Key: Up=U, Down=D, Disabled=X, Not Supported=_
```

11.1.3 SMBPBI for NVLink State Status

The following SMBPBI opcode can be used to check the NVLink state status. The NVLink status information can be used for further analysis of NVLink related error scenarios.

```
Query NVLink Information:
Opcode = 0x1A
Arg1 = 0x01

Return:
Bitmap of the first 32 NVLink states, where each bit encodes the state of the corresponding NVLink.
0 – NVLink is down
1 – NVLink is up
```

Opcode 0x1A in the command corresponds to the SMBPBI Opcode 1Ah “QUERY NVLINK INFORMATION.” Full details of this SMBPBI Opcode 1Ah can be referenced from the *SMBus Post Box Interface (SMBPBI) For GPUs Application Note* (NVOnline: 1001187).

11.1.4 Checking for NVLink Error Counters

The following NVSMI and DCGM commands can be used to check the NVLink error counters, including:

- NVLink Replay Error status
- Recovery Error status
- Flit CRC Error status

The NVLink error counters information can be used for further analysis of NVLink related error scenarios.

11.1.5 NVSMI

User can run the following command to check the NVLink error counters.

```
nvidia-smi nvlink -e
```

11.1.6 NVML API

User can the follow NVML API to check the NVLink error counters.

```
nvmlDeviceGetNvLinkErrorCounter
```

11.1.7 SMBPBI for NVLink Error Counters

The following SMBPBI opcode can be used to check the NVLink error counters. The NVLink error counters information can be used for further analysis of NVLink related error scenarios.

```
Query NVLink Information:
Opcode = 0x1A
Arg1 =
0x03 (NVLink Replay Error Count)
0x04 (NVLink Recovery Error Count)
0x05 (NVLink Flit CRC Error Count)

Arg2 = NVLink Index (NVLink Index = 0xFF will provide aggregate error count of all links)

Return:
Error count for the provided NVLink Index in Arg2
```

Opcode 0x1A in the command corresponds to the SMBPBI Opcode 1Ah “QUERY NVLINK INFORMATION.” Full details of this SMBPBI Opcode 1Ah can be referenced from the *SMBus Post Box Interface (SMBPBI) For GPUs Application Note* (NVOnline: 1001187).

Chapter 12. Debug and Triage

The following sections describe the triage checklist for baseboard, GPU, and NVSwitch errors. In addition, this triage checklist provides further guidance on additional telemetry and debug logs, which must be captured for firmware update, PCIe enumeration and booting, thermal, NVQual, and Field Diag issues.

12.1 Baseboard Debug

When gathering and reporting an NVIDIA HGX baseboard issue, customers and partners are required to follow the triage checklist and provide relevant debug information as follows (in addition to the standard issue reporting and logging guidelines):

- Summary of the issue and the number of devices impacted (GPU, NVSwitch, FPGA, ERoT, and HMC).
- Describe whether the issue is isolated to a specific device and slot.
- Provide information on the number of baseboards and components affected.
- Can the system or device recover after a power cycle or reset.
- Provide the last known state of the system before failure (for example, was the system running a specific workload or application).
- Provides all firmware versions being used (GPU, NVSwitch, FPGA, ERoT, and HMC).
- List all software versions being used (example: NVIDIA GPU driver, fabric manager, NVIDIA® CUDA®, and DCGM).
- Provides relevant debug logs from Field Diag.
- Software and firmware versions installed on the platform, including FRU information.
- Detailed steps on reproduction of the hardware error or issue.
- Frequency of the production and information regarding whether the issue reported is an isolated issue or affecting multiple systems.
- Hardware sample version (for example. ES, QS, PS, or other sample type).

12.2 Out-of-Band Debug

The following list provides detailed out-of-band telemetry to be captured when baseboard, GPU, or NVSwitch error occurs. Customers and partners are required to provide relevant debug information as follows (in addition to the standard issue reporting and logging guidelines):

- Confirm if out-of-band communication is functional for baseboard, GPU, NVSwitch, and FPGA.
- Capture GPU SMBPBI Opcodes
  - 02H – Get Temperature
  - F1H – Get Power Supply Status
  - F3H – Get PCIe Reset Status
- Capture FPGA and Baseboard Devices SMBPBI Opcodes
  - 02H – Get Temperature
  - 04H – Get Power
  - B0H – Get OVERT Information
  - B1H – Get I2C-1 GPU Interrupt Information
  - B2H – Get I2C-2 Onboard Device Interrupt Information
  - B3H – Get HSC Device Alert Information
  - B5H – Get Clock Output Enable Status
  - BBH – Get GPU Presence and Power Status
  - F1H – Get Power Supply Status
  - F2H – Assert/De-assert Device PCIe Fundamental Reset State
  - F3H – Get PCIe Reset Status
- Capture NVSwitch SMBPBI Opcodes
  - 02H – Get Temperature
  - 18H – Query Miscellaneous NVSwitch State Flags
  - 21H – Query PCIe Link Status and Error Counts

12.3 Firmware Update Issue

When gathering and reporting a firmware update issue, customers and partners are required to follow the triage checklist and provide relevant debug information as follows (in addition to the standard issue reporting and logging guidelines):

- Ensure firmware write-protection and BMC polling are disabled.
- Provide NVFlash, HMC, or ERoT firmware update logs.

12.4 PCIe Enumeration and Non-Booting Issue

When gathering and reporting a PCIe enumeration and non-booting issue, customers and partners are required to follow the triage checklist and provide relevant debug information as follows (in addition to the standard issue reporting and logging guidelines):

- Confirm ERoT has successfully booted and provide relevant ERoT logs if there is an ERoT boot failure.
- Provide partner's power-up sequence measurements and verify that power-on sequencing steps are being adhered.
- Verify all power rail status and reference clock status.
- Verify PCIe reset and enumeration sequencing steps are being adhered.
- Confirm any component and connector damage to the baseboard and SXMs.
- Verify PCIe link status in upstream retimer components and PCIe switches.

12.5 Thermal Issue

When gathering and reporting a thermal issue, customers and partners are required to follow the triage checklist and provide relevant debug information as follows (in addition to the standard issue reporting and logging guidelines):

- Review GPU clock event reason, if any. For a complete list of "Clock Event Reason Codes and Description," refer to the *NVIDIA GPU Boost Overview for Data Center Products* (NVOnline: 1100183).
- 0x08 HW Slowdown
- 0x40 HW Thermal Shutdown
- 0x80 HW Power Brake Slowdown
- 0x20 Clock Optimized for Thermal (previously noted as SW Thermal Slowdown).
- Provide details on which components were impacted by the thermal issue.
- Provide and confirm if OVERT is triggered (SMBPBI Opcode B0H).
- Confirm and inspect if heat sink and cooling solutions are installed, seated, and operated correctly. Particularly if "HW Slowdown," "Thermal Shutdown," "Power Brake," or "SW Thermal Slowdown" events have occurred.
- Confirm and review thermal interface material (TIM) used and installation steps for partner cooled product SKUs.
- Provide details on the operating environment at the time of failure.

Chapter 13. Issue Reporting Guidelines

The following sections describe the standard issue reporting and telemetry logging guidelines. When customers and partners report an issue, the following platform information and telemetry logging information described in this section should be provided to NVIDIA.

13.1 Hardware and Software Platform Information

When gathering and reporting a hardware issue, customers and partners are required to provide the following information regarding the platform:

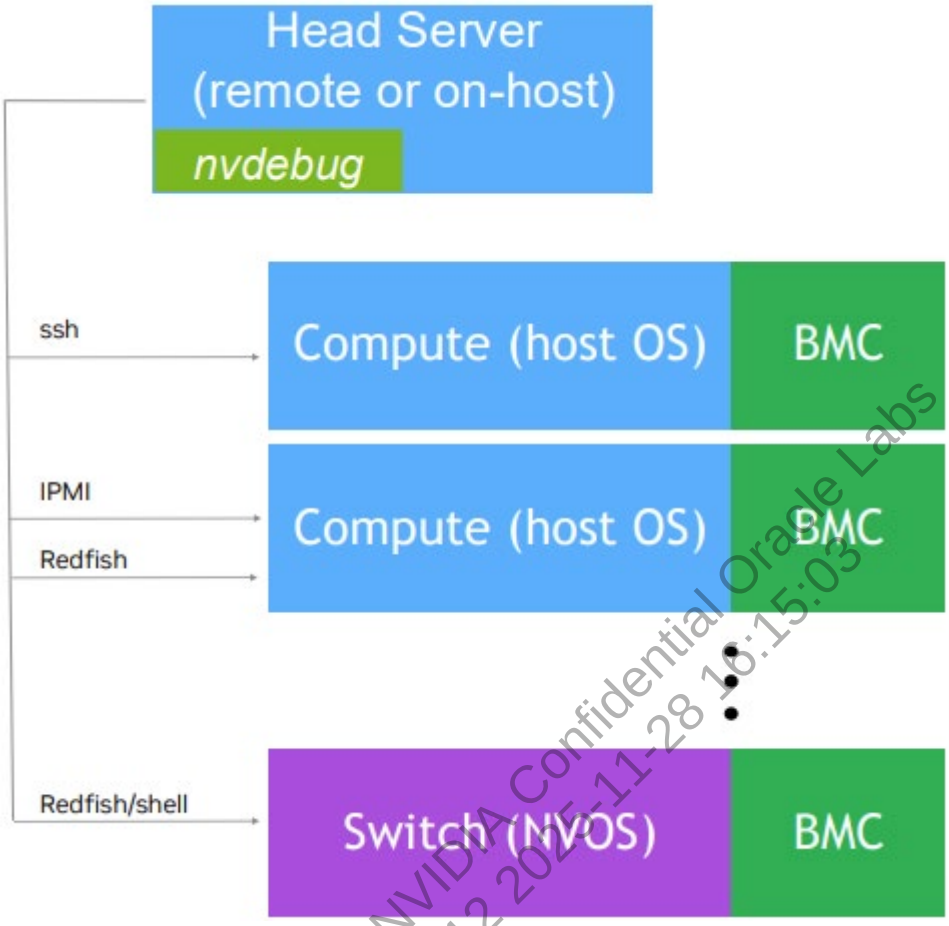
- Software and firmware versions installed on the platform, including FRU information.
- Detailed steps on reproduction of the hardware error or issue.
- Frequency of the production and information regarding whether the issue reported is an isolated issue or affecting multiple systems.
- Hardware sample version (for example. ES, QS, PS, or other sample type).

13.2 NVDebug for Datacenter Products

The NVIDIA NVDebug tool can be deployed on management or head servers in GB NVL systems and other NVIDIA data center products. In GB NVL systems, the tool implements a wrapper to aggregate compute, NVSwitch, and BMC logs across racks in the Data Center. It provides an easy-to-use, customizable, and scalable approach to collecting necessary logs across multiple racks and complex array of subcomponents. In cases where a focused investigation is needed on a subset of nodes in GB NVL or other NVIDIA rack systems, the tool can be used to retrieve the necessary logs from the same nodes of interest.

Customers are strongly encouraged to collect NVDebug logs and attaching the same to NVBugs while reaching out to NVIDIA for support.

Figure 13-1 NVDebug Tool in GB NVL Systems Overview



For further information, please refer to NVOnline: 1109504 NVIDIA Debug Tool for Datacenter Products (NVDebug). The *NVDebug Tool User's Guide* is included in the package under the same posting.

13.2.1 Collecting Rack-level Logs in GB NVL Systems

The NVDebug tool can be used to collect logs from all the trays of the GB200 NVL rack. The tool supports and performs the full rack-level log collection simultaneously from each Compute Node and NVSwitch node.

To collect the full GB200 NVL rack log collection, please use the following steps:

1. Create a separate Compute Node config.yaml and separate NVSwitch node nvswitch_config.yaml file from NVDebug package.
2. Update the PLATFORM parameter in the Compute Node config to arm64 and in the Switch Tray config to NVSwitch.
3. Make other modifications as needed based on the *NVDebug Tool User's Guide*.

Here is an example of the config guide for the Switch Node (nvswitch_config.yaml).

```
# Server Platform Type
# NVDebug will run tests for the listed platform
# The accepted values are HGX-HMC, DGX, arm64, x86_64 and NVSwitch
PLATFORM: "NVSwitch"
```

Here is an example of the config file for the Compute Node (compute_config.yaml).

```
# Server Platform Type
# NVDebug will run tests for the listed platform
# The accepted values are HGX-HMC, DGX, arm64, x86_64 and NVSwitch
PLATFORM: "arm64"

# Server Baseboard Type
TargetBaseboard: "gb200 nvl"
```

4. Update and create a DUT object for each compute and Switch node in the dut_config.yaml file with DUT specific IP and access credentials.
5. For the compute nodes, set the NodeType to Compute and the ConfigFileToUse as the Compute Tray config file.
6. For the NVSwitch node, set the NodeType to SwitchTray and ConfigFileToUse as the Switch Tray config.

   Here is an example with two compute nodes and one switch node.

```
DUT_Defaults: &dut_defaults
  ipmi_cipher: "-C17"
  RF_DEFAULT_PREFIX: "/redfish/v1"
  RF_AUTH: true
  IP_NETWORK: 'ipv4'

gb200-compute-tray-1:
  <<: *dut_defaults
  NodeType: "Compute"
  BMC_IP: "xxxx"
  BMC_USERNAME: "xxxx"
  BMC_PASSWORD: "xxxx"
  HOST_IP: "xxxx"
  HOST_USERNAME: "xxxx"
  HOST_PASSWORD: "xxxx"
  ConfigFileToUse: "compute_config.yaml"

gb200-compute-tray-2:
  <<: *dut_defaults
  NodeType: "Compute"
  BMC_IP: "xxxx"
  BMC_USERNAME: "xxxx"
  BMC_PASSWORD: "xxxx"
  HOST_IP: "xxxx"
  HOST_USERNAME: "xxxx"
  HOST_PASSWORD: "xxxx"
  ConfigFileToUse: "compute_config.yaml"

gb200-switch-tray-1:
  <<: *dut_defaults
  NodeType: "SwitchTray"
  HOST_IP: "xxxx"
  HOST_USERNAME: "xxxx"
  HOST_PASSWORD: "xxxx"
  ConfigFileToUse: "nvswitch_config.yaml"
```

7. Run NVDebug without specifying the credentials or the platform because this information will be picked up from the config files.

```
$ ./nvdebug
Multiple DUTs found. Ignoring CLI DUT details
Log directory created at /tmp/nvdebug_logs_02_10_2024_14_33_05
Starting collection for DUT gb200-compute-tray-1
Starting collection for DUT gb200-compute-tray-2
Starting collection for DUT gb200-switch-tray-1
gb200-switch-tray-1: [14:33:06] Unknown platform. User provided platform type: NVSwitch
gb200-switch-tray-1: [14:33:06] BMC IP: None
Log collection has started for gb200-switch-tray-1
gb200-compute-tray-2: [14:33:07] Identified system as Model: P3809, Partno: 699-13809-0404-
600, Serialno:1583124820340
gb200-compute-tray-2: [14:33:07] User provided platform type: arm64
gb200-compute-tray-2: [14:33:07] BMC IP: XXXX
Log collection has started for gb200-compute-tray-2
gb200-compute-tray-1: [14:33:08] Identified system as Model: P3809, Partno: 699-13809-0404-
600, Serialno:1583124820254
gb200-compute-tray-1: [14:33:08] User provided platform type: arm64
gb200-compute-tray-1: [14:33:08] BMC IP: XXXX
Log collection has started for gb200-compute-tray-1
gb200-switch-tray-1: [14:37:11] Log collection is now complete
gb200-switch-tray-1: [14:37:11] Log collection took 4m 5.34s
gb200-compute-tray-1: [14:40:19] Log collection is now complete
gb200-compute-tray-1: [14:40:19] Log collection took 7m 11.21s
DUT gb200-compute-tray-1 completed.
gb200-compute-tray-2: [14:40:20] Log collection is now complete
gb200-compute-tray-2: [14:40:20] Log collection took 7m 12.77s
DUT gb200-compute-tray-2 completed.
DUT gb200-switch-tray-1 completed.
```

A log zip file is created at `/tmp/nvdebug_logs_02_10_2024_14_33_05.zip`. The log output will have a subfolder for each DUT that was specified in the `dut_config` file.

13.3 System Event Logs and Telemetry Logging

In addition to providing NVIDIA the hardware and software platform information when an issue occurs, customers and partners are required to following the set of standard telemetry and logs to NVIDIA:

- Host OS logs
  - lspci logs
  - Kernel dmesg syslog
  - Fabric manager log

    - BMC system event log
- nvidia-bugs-report and nvidia-smi -q logs
- Field Diag logs
- Device dumps through Redfish or I2C (FPGA, ERoT, and HMC).
- Device dumps through Redfish (Retimer).
- Log collection through Redfish or I2C (Self-test, LTSSM, HW Info, FW Inventory, and Health Check).
- Baseboard telemetry logging through SMBPBI.
    - Power good and voltage regulators status
    - Reference clock status
    - SXM presence status
    - Device interrupts
    - Thermal alerts
- GPU telemetry logging through SMBPBI.
    - PCIe recovery counter
    - PCIe correctable and uncorrectable error counters
    - PCIe NAK received and sent counter
    - PCIe equalization parameters
- NVSwitch telemetry logging through SMBPBI.
    - PCIe recovery counter
    - PCIe correctable and uncorrectable error counters
    - PCIe NAK received and sent counter
    - PCIe equalization parameters
- Retimer telemetry logging through SMBPBI.
- PCIe equalization parameters

For best practices on reporting issue through NVOnline, refer to *NVOnline Partner Portal Guidelines* (NVOnline: 1106532).

Chapter 14. Out-of-Band Telemetry through Redfish

This chapter provides partners and customers with instructions on how to obtain out-of-band telemetry using HMC and Redfish commands. The out-of-band telemetry can assist in the analysis and triage of XID and SXID error scenarios described in the previous section.

14.1 Checking PCIe Retimer Link Status Corresponding to SXM

The following Redfish command can be used to obtain the PCIe Retimer Link Status from each of the eight SXMs from the HMC:

```
curl
http://192.168.31.1/redfish/v1/Fabrics/HGX_PCIeRetimerTopology_[0-7]/Switches/PCIeRetimer_[0-7]/Ports/DOWN_0
```

The output would indicate if the PCIe Link Status, for example, an ActiveWidth=0 and CurrentSpeedGbps=0 would indicate the PCIe Link Down status:

```
“ActiveWidth”:0
“CurrentSpeedGbps”:0
```

Figure 14-1 shows a full example output of the PCIe Retimer Link Status from HMC. The example shows that the current retimer link status is operational and running at PCIe x16 at Gen5 speed.

Figure 14-1. Output of PCIe Retimer Link Status through Redfish Example

```
~ # curl http://192.168.31.1/redfish/v1/Fabrics/HGX_PCIeRetimerTopology_0/Switches/PCIeRetimer_0/Ports/DOWN_0
{
  "@odata.id": "/redfish/v1/Fabrics/HGX_PCIeRetimerTopology_0/Switches/PCIeRetimer_0/Ports/DOWN_0",
  "@odata.type": "#Port.v1_4_0.Port",
  "ActiveWidth": 16,
  "CurrentSpeedGbps": 32.0,
  "Id": "DOWN_0",
  "Links": {
    "AssociatedEndpoints": [
      {
        "@odata.id": "/redfish/v1/Fabrics/HGX_PCIeRetimerTopology_0/Endpoints/GPU_SXM_1"
      }
    ]
  },
  "Metrics": {
    "@odata.id": "/redfish/v1/Fabrics/HGX_PCIeRetimerTopology_0/Switches/PCIeRetimer_0/Ports/DOWN_0/Metrics"
  },
  "Name": "DOWN_0 Resource",
  "Oem": {
    "Nvidia": {
      "@odata.type": "#NvidiaPort.v1_0_0.NvidiaPort"
    }
  },
  "PortProtocol": "PCIe",
  "PortType": "DownstreamPort",
  "Status": {
    "Conditions": [],
    "Health": "OK",
    "HealthRollup": "OK",
    "State": "Enabled"
  }
}~ #
```

14.2 Checking for SXM PWR_GOOD Status

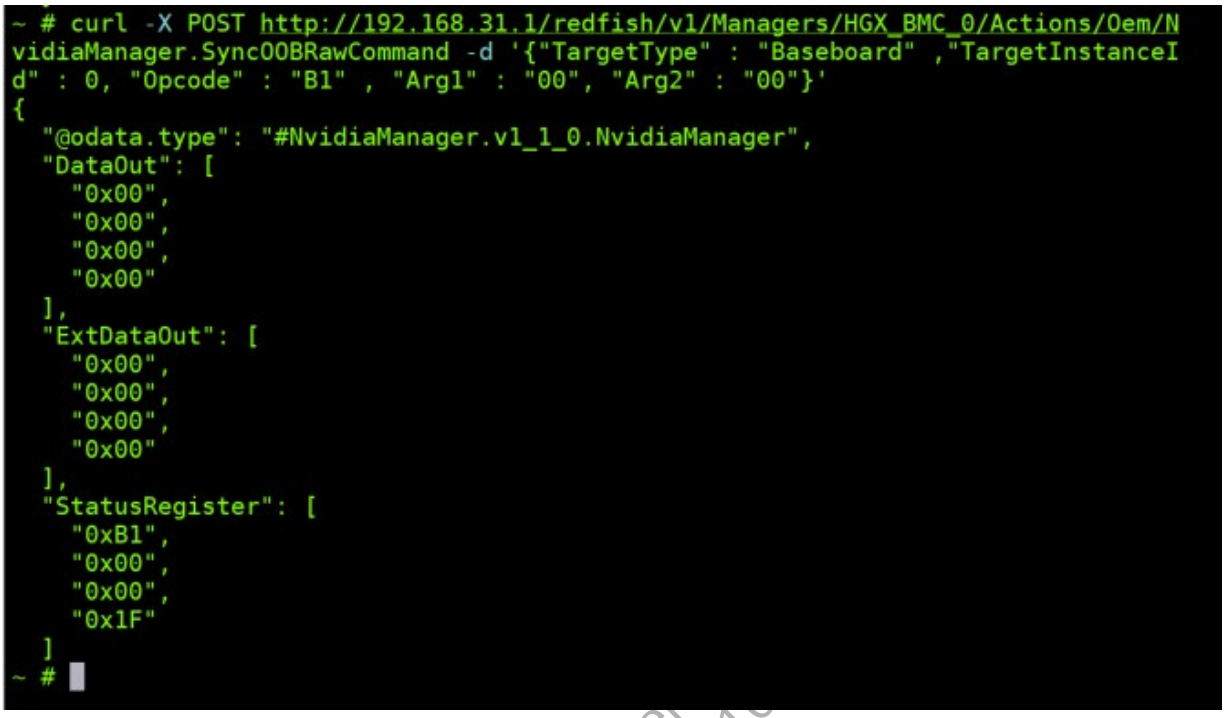
The following Redfish command can be used to obtain the **SXM PWR_GOOD** status from the HMC:

```
curl -X POST
http://192.168.31.1/redfish/v1/Managers/HGX_BMC_0/Actions/Oem/NvidiaManager.SyncOOBRawCommand -d
'{"TargetType" : "Baseboard" ,"TargetInstanceId" : 0, "Opcode" : "B1" , "Arg1" : "00", "Arg2" :
"00"}'
```


Opcode 0xB1 in the command corresponds to the SMBPBI Opcode B1h “GET GPU/EROT INTERRUPT INFO.” Full details of this SMBPBI Opcode B1h can be referenced from the *SMBus Post Box Interface (SMBPBI) For NVIDIA Baseboards Application Note* (NVOnline: 1030060).

Figure 14-2 shows an example output from the Redfish command. Bit 8 through 15 in the “DataOut” output represents the **PWR_GOOD** status for SXMs 1 through 8. A value of 1b indicates that there is an issue with the **PWR_GOOD** status, and a value of 0b indicates that **PWR_GOOD** status is normal.

Figure 14-2. Output of SXM PWR_GOOD Status through Redfish Example



14.3 Checking for SXM THERM_OVERT Status

The following Redfish command can be used to obtain the SXM **THERM_OVERT** status from the HMC:

```
curl -X POST
http://192.168.31.1/redfish/v1/Managers/HGX_BMC_0/Actions/Oem/NvidiaManager.SyncOOBRawCommand -d
'{"TargetType" : "Baseboard" ,"TargetInstanceId" : 0, "Opcode" : "B0" , "Arg1" : "00", "Arg2" :
"00"}'
```

Opcode 0xB0 in the command corresponds to the SMBPBI Opcode B0h “GET OVERT INFO.” Full details of this SMBPBI Opcode B0h can be referenced from the *SMBus Post Box Interface (SMBPBI) For NVIDIA Baseboards Application Note* (NVOnline: 1030060).

Bit 0 through 7 in the “DataOut” field of the Redfish command response represent the **THERM_OVERT** status for SXMs 1 through 8. A value of 1b indicates that the corresponding SXM has an over-temperature issue.

14.4 Checking for ERoT Fatal Error Status

The following Redfish command can be used to obtain the ERoT Fatal Error status from the HMC:

```
curl -X POST
http://192.168.31.1/redfish/v1/Managers/HGX_BMC_0/Actions/Oem/NvidiaManager.SyncOOBRawCommand -d
'{"TargetType" : "Baseboard" ,"TargetInstanceId" : 0, "Opcode" : "B1" , "Arg1" : "80", "Arg2" :
"00"}'
```

Opcode 0xB1 in the command corresponds to the SMBPBI Opcode B1h “GET GPU/EROT INTERRUPT INFO.” Full details of this SMBPBI Opcode B1h can be referenced from the *SMBus Post Box Interface (SMBPBI) For NVIDIA Baseboards Application Note* (NVOnline: 1030060).

The following figure shows an example output from the Redfish command. Bit 0 through 7 in the “DataOut” output represents the ERoT Error status for SXMs 1 through 8. A value of 1b indicates that there is an ERoT Fatal Error status, and a value of 0b indicates that ERoT status is normal.

Figure 14-3. Output of SXM ERoT Fatal Error Status through Redfish Example

```
~ # curl -X POST http://192.168.31.1/redfish/v1/Managers/HGX_BMC_0/Actions/Oem/N
vidiaManager.SyncOOBRawCommand -d '{"TargetType" : "Baseboard" ,"TargetInstanceI
d" : 0, "Opcode" : "B1" , "Arg1" : "00", "Arg2" : "00"}'
{
  "@odata.type": "#NvidiaManager.v1_1_0.NvidiaManager",
  "DataOut": [
    "0x00",
    "0x00",
    "0x00",
    "0x00"
  ],
  "ExtDataOut": [
    "0x00",
    "0x00",
    "0x00",
    "0x00"
  ],
  "StatusRegister": [
    "0xB1",
    "0x00",
    "0x00",
    "0x1F"
  ]
~ #
```

14.5 Downloading FPGA Register Table

The following Redfish command can be used to download the FPGA register table dump using HMC. This FPGA register table dump can be provided to NVIDIA for further analysis of error scenario.

1. Initiate the FPGA register dump collection.

```
curl -X POST
http://${HMC_IP}/redfish/v1/Systems/HGX_Baseboard_0/LogServices/Dump/Actions/LogService.Collect
DiagnosticData/ -d '{"DiagnosticDataType":"OEM",
"OEMDiagnosticDataType":"DiagnosticType=FPGA"}'
```

2. Check the task completion progress.

```
curl -X GET http://${HMC_IP}/redfish/v1/TaskService/Tasks/{task_id}
```

3. Download the FPGA register dump after TaskState becomes “Complete.”

```
curl -X GET
http://${HMC_IP}/redfish/v1/Systems/HGX_Baseboard_0/LogServices/Dump/Entries/1/attachment --
output /tmp/fpga-report.tar.xz
```

14.6 Downloading ERoT Logs

The following Redfish command can be used to download the ERoT logs using HMC. These ERoT logs can be provided to NVIDIA for further analysis of error scenario.

1. Initiate the ERoT dump collection.

```
curl -X POST
http://${HMC_IP}/redfish/v1/Systems/HGX_Baseboard_0/LogServices/Dump/Actions/LogService.Collect
DiagnosticData/ -d '{"DiagnosticDataType":"OEM", "OEMDiagnosticDataType":"DiagnosticType=EROT"}
```

2. Check the task completion progress.

```
curl -X GET http://${HMC_IP}/redfish/v1/TaskService/Tasks/{task_id}
```

3. Download the ERoT logs after TaskState becomes “Complete.”

```
curl -X GET
http://${HMC_IP}/redfish/v1/Systems/HGX_Baseboard_0/LogServices/Dump/Entries/{entry_id}/attachm
ent --output erot_dump_1.tar.xz
```

14.7 Downloading PCIe Retimer LTSSM Log

The following Redfish command can be used to download the PCIe Retimer LTSSM log using HMC. This PCIe Retimer LTSSM log can be provided to NVIDIA for further analysis of error scenario.

1. Initiate the PCIe Retimer LTSSM log dump collection.

```
curl -X POST
http://${HMC_IP}/redfish/v1/Systems/HGX_Baseboard_0/LogServices/Dump/Actions/LogService.Collect
DiagnosticData/ -d '{"DiagnosticDataType":"OEM",
"OEMDiagnosticDataType":"DiagnosticType=RetLTSSM"}'
```

2. Check the task completion progress.

```
curl -X GET http://${HMC_IP}/redfish/v1/TaskService/Tasks/{task_id}
```

3. Download the PCIe Retimer LTSSM log after TaskState becomes “Complete.”

```
curl -X GET
http://${HMC_IP}/redfish/v1/Systems/HGX_Baseboard_0/LogServices/Dump/Entries/1/attachment --
output /tmp/retimer.tar.xz
```

Chapter 15. Reliability, Availability, and Serviceability

15.1 Error Counters and Thresholds

This section provides guidance and error counters that can be used to assess the health of the NVIDIA GPU system. When a particular error threshold is exceeded, NVIDIA recommends partners review any XID and SXID messages that have been generated by following the *Server RAS Catalog* (NVOnline: 1116117). NVIDIA GPU and NVIDIA HGX Field Diagnostics tools can be used to determine if the system is eligible for RMA.

For NVIDIA Hopper HGX baseboard products, additional information on telemetry data to capture, in-band, and out-of-band APIs available to retrieve health metrics can be referenced from the *NVIDIA Hopper HGX 8-GPU Telemetry Playbook* (NVOnline: 1093241)

Table 15-1. Error Counters and Thresholds

Error Counter	Threshold	Next Steps
GPU PCIe Fatal Error (GPU-PCI-ERR-CTR-FATAL)	> 0	Review any XID and SXID error messages using the “XID error to Action Mapping” in the previous section. GPU and HGX Field Diagnostics is recommended if this error is persistent and reproducible
GPU PCIe Non-Fatal Error (GPU-PCI-ERR-CTR-NON-FATAL)	> 0	Review any XID and SXID error messages using the “XID error to Action Mapping” in the previous section. GPU and HGX Field Diagnostics is recommended if this error is persistent and reproducible
GPU PCIe Unsupported Request (GPU-PCI-ERR-CTR-UNSUPP-REQ)	> 0	Review system topology, external system queries and bus scans. Review any XID and SXID error messages using the “XID error to Action Mapping” in the previous section. GPU and HGX Field Diagnostics is recommended if this error is persistent and reproducible
GPU PCIe Recovery (GPU-PCI-ERR-CTR-RECOVERY)	TxRecovery > 2	Review any XID and SXID error messages using the “XID error to Action Mapping” in the previous section. GPU and HGX Field Diagnostics is recommended if this error is persistent and reproducible

Error Counter	Threshold	Next Steps
GPU Row Remapping Failure (GPU-ROW-REMAP-FAILED)	Binary 1 = Failed	GPU and HGX Field Diagnostics is recommended if this error is persistent and reproducible
GPU ECC Uncorrectable Error (ECC-ERR-CTR-UCE)	>1	Review any XID and SXID error messages using the “XID error to Action Mapping” in the previous section. GPU and HGX Field Diagnostics is recommended if this error is persistent and reproducible
GPU ECC Correctable Error (ECC-ERR-CTR-CE)	No limit	Review any XID and SXID error messages using the “XID error to Action Mapping” in the previous section. GPU and HGX Field Diagnostics is recommended if this error is persistent and reproducible
NVSwitch PCIe Correctable Error (NVSWITCH-PCI-ERR-CTR-CORR)	> 19 per 10 minutes	Review any XID and SXID error messages using the “XID error to Action Mapping” in the previous section. GPU and HGX Field Diagnostics is recommended if this error is persistent and reproducible
NVSwitch PCIe Fatal Error (NVSWITCH-PCI-ERR-CTR-FATAL)	> 0	Review any XID and SXID error messages using the “XID error to Action Mapping” in the previous section. GPU and HGX Field Diagnostics is recommended if this error is persistent and reproducible
NVSwitch PCIe Non-Fatal Error (NVSWITCH-PCI-ERR-CTR-NON-FATAL)	> 0	Review any XID and SXID error messages using the “XID error to Action Mapping” in the previous section. GPU and HGX Field Diagnostics is recommended if this error is persistent and reproducible
NVSwitch PCIe Unsupported Request (NVSWTICH-PCI-ERR-CTR-UNSUPP-REQ)	> 0	Review any XID and SXID error messages using the “XID error to Action Mapping” in the previous section. GPU and HGX Field Diagnostics is recommended if this error is persistent and reproducible
NVSwitch PCIe Recovery (NVSWTICH-PCI-ERR-CTR-RECOVERY)	TxRecovery: > 2	Review any XID and SXID error messages using the “XID error to Action Mapping” in the previous section. GPU and HGX Field Diagnostics is recommended if this error is persistent and reproducible
NVSwitch ECC Correctable Error (NVSWTICH-ECC-ERR-CTR-CE)	No Limit	Review any XID and SXID error messages using the “XID error to Action Mapping” in the previous section. GPU and HGX Field Diagnostics is recommended if this error is persistent and reproducible

Chapter 16. ConnectX-7 Platforms

This chapter lists the different monitoring measures and debug steps relevant to the ConnectX-7 platforms.

16.1 ConnectX-7 Thermal Issues

The following subsections are the different thermal issue scenarios encountered in ConnectX-7 (CX-7). Debugging steps are listed to address reported thermal issues. Where applicable, in-band and out-of-band telemetry have been identified to aid partners in diagnosing the associated error scenarios.

16.1.1 Thermal Warnings

ConnectX-7 firmware initiates thermal warnings, triggering a hardware slowdown when the ASIC reaches the thermal warning threshold. This is indicated by the assertion of a corresponding **THERM_warning** signal, alerting the user to the over-temperature condition.

16.1.2 Thermal Hardware Shutdown and OVERT Event

Upon reaching the thermal shutdown threshold, the ASIC asserts a **THERM_SHUTDN** signal to notify the user of the over-temperature condition. A message is sent to the driver, prompting the firmware to initiate the shutdown process for the ASIC.

16.2 ConnectX-7 Debug Steps

Following are the CX-7 debug steps for the data center products.

16.2.1 ASIC Thermal Warning Message

Perform the following debug steps when the following ASIC thermal warning message is received by dmsg:

```
mlx5_core 0000:41:00.1: temp_warn:171:High temperature on sensors with bit set 2 8000 0000 0000
0000
```

Debug Steps:

1. To read and verify the high temperature, run:
   ```
   mget_temp -d [device name]
   ```
2. Perform a visual inspection.
3. Ensure that the airflow direction aligns with the product specifications.
4. Verify that fans are operating at the expected performance level.
5. Replace the faulty NIC with a functioning one from another slot in the server and observe if the issue persists with the NIC.
6. If the problem persists in the same server slot, investigate the airflow within that slot.
7. Replace the cards between servers; if the failed card functions properly in another server, investigate the airflow within the original failing server. If the issue occurs wherever the NIC is installed, initiate an RMA process.

16.2.2 Transceiver Thermal Warning Message

Perform the following debug steps when the following a transceiver thermal warning message is received by dmesg:

```
mlx5_core 0000:33:00.0: Port module event[error]: module 0, Cable error, High Temperature
mlx5_core 0000:41:00.1: temp_warn:171:High temperature on sensors with bit set 2 8000 0000 0000
0000
```

Debug Steps:

1. Run:
   ```
   mlxlink -d [device]  -m
   ```
2. Verify that fans are operating at the anticipated performance level.
3. It is possible that a defective transceiver is causing the issue. Replace the transceiver and check if issue is resolved.

16.3 ConnectX-7 Power

If the power of the transceiver exceeds the supported power of the port, the link will be disabled.

Run the following command to identify the reason:

```
mlxlink -d [device]
```

Ensure that the transceiver being used is supported. For a list of validated and supported cables, refer to: NVIDIA Adapter Cards Firmware Release Notes

Chapter 17. BlueField-3 Platforms

This chapter highlights the advanced functionalities of the NVIDIA BlueField-3 Data Processing Unit (DPU) and Super Network Interface Cards (Super NIC), focusing on Reliability, Availability, and Serviceability (RAS) aspects. The BlueField-3 platforms integrate a suite of RAS features aimed at improving system reliability, minimizing downtime, and facilitating maintenance tasks. The following commands are used for monitoring and debugging critical BlueField-3 (BF-3) components.

Note: The BMC, eMMC, NVMe SSD, ERoT, and 1G OOB interfaces are relevant to BlueField-3 platforms only when the Arm cores are activated.

17.1 OS Dump

The sysinfo.snapshot tool is designed to capture a snapshot of all configuration details and relevant information concerning both the server and the NVIDIA boards. It conducts a thorough scan of the system, providing information on the current settings of the operating system, networking, and hardware components.

Run the tool on both the server host OS and the BlueField Arm OS to check for driver errors, assertions, PCIe, and network errors.

```
# first, ssh to the Host server OS

Sysinfo.snapshot.py

# Second, ssh to the ARM OS

Sysinfo.snapshot.py

# Output example
localhost:~# sysinfo-snapshot.py
Sysinfo-snapshot is still in process...please wait till completed successfully
Gathering the information may take a while, especially in large networks
Your patience is appreciated
------------------------------------------------------------

Running sysinfo-snapshot has ended successfully!
```

```
Temporary destination directory is /tmp/
Out file name is /tmp/sysinfo-snapshot-v3.7.6-bu-fae4-bf3-20240312-103303.tgz

/tmp/sysinfo-snapshot-v3.7.6-bu-fae4-bf3-20240312-103303.tgz:
amber_info
bu-fae4-bf3-20240312-103303.html
cables
commands_txt_output
devlink
dmesg
dmidecode
ecn
err_messages:
        dummy_functions         - contains all not found commands
        dummy_paths             - contains all not existing internal files (/paths)
        dummy_external_paths    - contains all not existing external files (/paths)
etc_udev_rulesd                 - contains all files under /etc/udev/rules.d
ethtool_S                       - contains all files which are generated from invoking ethtool -S
<interface>
firmware                        - contains all firmware files (mst dump files and commands outputs)
journal
lib_udev_rulesd                 - contains all files under /lib/udev/rules.d
libvma.conf
lshw
pcie_files
performance_tuning_analyze.html
pkglist
show_irq_affinity_all
sr_iov.html
trace
var_log_dmesg
var_log_syslog
```

17.2 DDR

The following table lists the BlueField-3 platforms DDR.

Table 17-1. DDR

Action	Command
Collect the DDR temperature	# ssh to the ARM OS modprobe mlxbf-ptm cd /sys/kernel/debug/mlxbf-ptm/monitors/status/ cat ddr_temp
Check the RShim for DDR errors through the Arm OS	# From the ARM OS: bfrshlog
Check the RShim for DDR errors through the Host server OS	# From the Host OS: echo "DISPLAY_LEVEL 2" > /dev/rshim0/misc cat /dev/rshim0/misc

17.3 BMC

The following table lists the BlueField-3 platforms BMC.

Table 17-2. BMC

Action	Command
Get BMC version	#Using redfish curl -k -u root:'<password>' -X GET https://<bmc_ip>/redfish/v1/UpdateService/FirmwareInventory/BMC_Firmware | jq -r ' .Version' #From BMC CLI cat /etc/os-release
Get BMC diagnostic data	# Using redfish # Start diagnostic data task curl -k -u root:'<password>' -H 'Content-Type: application/json' -X POST -d '{"DiagnosticDataType":"Manager"}' https://<BMC IP>/redfish/v1/Managers/Bluefield_BMC/LogServices/Dump/Actions/LogService.CollectDiagnosticData #check the <Id> and then Monitor unit generation progress curl -k -u root:'<password>' -H 'Content-Type: application/json' -X GET https://<BMC IP>/redfish/v1/TaskService/Tasks/<Id> # Once "TaskState": "Completed",

Action	Command
	# Check the "Payload":"Location: /redfish/v1/Managers/Bluefield_BMC/LogServices/Dump/Entries/<Entry Id>" # Download the BMC logs curl -k -u root:'<password>' -H 'Content-Type: application/json' -X GET https://<BMC IP>/redfish/v1/Managers/Bluefield_BMC/LogServices/Dump/Entries/<Entry Id>/attachment --output <output file>
Collect dmesg	# From the BMC CLI dmesg > <output file>
Detect the I2C devices connected to the BMC	# From the BMC CLI i2cdetect -y <i2c bus number>

17.4 USB

The following table lists the BlueField-3 platforms USB.

Table 17-3. USB

Action	Command
Check USB interface	# From the BMC CLI, verify 22dc device exists lsusb

17.5 eMMC

The following table lists the BlueField-3 platforms eMMC.

Table 17-4. eMMC

Action	Command
Collect eMMC lifecycle logs 1. Download the "Collect eMMC Lifetime Tool" via NVOnline #1113192. 2. Compile the file. 3. Run the tool.	# ssh to the ARM OS # Compile Gcc eMMC_LTM_demo.c -o eMMC_LTM # Run the tool ./eMMC_LTM

17.6 NVMe SDD

The following table lists the BlueField-3 platforms NVMe SSD.

Table 17-5. NVMe SSD

Action	Command

Run data monitoring using the smartontools	# ssh to the ARM OS
	
	`apt install smartontools`
	
	`smartctl --all /dev/nvme0`
	
	# Output example
	`root@bu-fae4-bf3-1:~# smartctl --all /dev/nvme0`
	`smartctl 7.2 2020-12-30 r5155 [aarch64-linux-5.15.0-1032-bluefield] (local build)`
	`Copyright (C) 2002-20, Bruce Allen, Christian Franke, www.smartmontools.org`
	
	`=== START OF INFORMATION SECTION ===`
	`Model Number: MTFDHBL128TDP`
	`Serial Number: 22443C22C214`
	`Firmware Version: MU05`
	`PCI Vendor/Subsystem ID: 0x1344`
	`IEEE OUI Identifier: 0x00a075`
	`Total NVM Capacity: 128,035,676,160 [128 GB]`
	`Unallocated NVM Capacity: 0`
	`Controller ID: 0`
	`NVMe Version: 1.3`
	`Number of Namespaces: 4`
	`Local Time is: Tue Mar 12 12:25:30 2024 UTC`
	`Firmware Updates (0x12): 1 Slot, no Reset required`
	`Optional Admin Commands (0x001f): Security Format Frmw_DL NS_Mngmt Self_Test`
	`Optional NVM Commands (0x0055): Comp DS_Mngmt Sav/Sel_Feat Timestmp`
	`Log Page Attributes (0x06): Cmd_Eff_Lg Ext_Get_Lg`
	`Maximum Data Transfer Size: 64 Pages`
	`Warning Comp. Temp. Threshold: 115 Celsius`
	`Critical Comp. Temp. Threshold: 123 Celsius`

Action	Command
	Supported Power States St Op Max Active Idle RL RT WL WT Ent_Lat Ex_Lat 0 + 6.00W 300.00W - 0 0 0 0 500 500 1 + 6.00W - - 1 1 1 1 500 500 2 + 6.00W - - 2 2 2 2 500 500 3 - 0.2270W - - 3 3 3 3 5000 10000 4 - 0.0130W - - 4 4 4 4 10000 50000 === START OF SMART DATA SECTION === SMART overall-health self-assessment test result: PASSED SMART/Health Information (NVMe Log 0x02) Critical Warning: 0x00 Temperature: 34 Celsius Available Spare: 100% Available Spare Threshold: 10% Percentage Used: 11% Data Units Read: 19,174,769 [9.81 TB] Data Units Written: 16,561,766 [8.47 TB] Host Read Commands: 2,106,598,041 Host Write Commands: 1,896,538,679 Controller Busy Time: 8,555 Power Cycles: 109 Power On Hours: 4,413 Unsafe Shutdowns: 85 Media and Data Integrity Errors: 0 Error Information Log Entries: 0 Warning Comp. Temperature Time: 0 Critical Comp. Temperature Time: 0 Temperature Sensor 1: 34 Celsius Error Information (NVMe Log 0x01, 16 of 64 entries) No Errors Logged

17.7 Power and Thermal

The following table lists the BlueField-3 platforms power and thermal.

Table 17-6. BF-3 Power and Thermal

Action	Command
Collect the BlueField-3 SoC temperature:	# ssh to the ARM OS mget_temp -d <mst dev>
Check for thermal/power throttling events of the Arm temperature:	# ssh to the ARM OS Modprobe mlxbf-ptm cd /ys/kernel/debug/mlxbf-ptm/monitors/status/ cat thermal_throttling_event_count cat power_throttling_event_count # ARM core power (PCIe goldfingers) cat vr0_power # NIC and misc power (ATX) cat vr1_power

17.8 ERoT

The following table lists the BlueField-3 platforms ERoT.

Table 17-7. ERoT

Action	Command
Collect CEC firmware version	# ssh to the Host OS curl -k -u root:'<password>' -X GET https://<bmc_ip>/redfish/v1/UpdateService/FirmwareInventory/Bluefield_FW_ERoT | jq -r '.Version'

17.9 1G OOB Interface

The following table lists the BlueField-3 platforms 1G OOB interface.

Table 17-8. 1G OOB Interface

Action	Command
Collect MDIO data	# ssh to the ARM OS # Clone tool from https://github.com/wkz/phytool.git and run make to build it # Read MDIO standard register in the PHY using clause 22 convention ./phytool read IFACE/0x3/reg> # example: ./phytool read oob_net0:/0x3/3

Chapter 18. NVIDIA Quantum-3 Platforms

This chapter applies to Q3200 switches only.

The NVIDIA Firmware Tools (MFT) package is a set of firmware management and debug tools for NVIDIA devices. This section describes MFT features, tools content and configuration. For more information on the MFT tool, refer to: NVIDIA Firmware Tools (MFT) Documentation.

Note: MFT tools must only be used for debug purposes on HGX 8-GPU platforms. For Multi-node NVLink Switches, NVOS CLI must be used.

18.1 Link Monitoring

- Checks the status of the links in normal operation. This feature ensures comprehensive monitoring, analysis, and troubleshooting capabilities across different interfaces, enhancing the overall reliability and serviceability of the system.
- **AmBER Tool Usage:** This tool utilizes the "mlxlink" command with the "--amber_collect" option to gather various chip parameters and readings. It enables the retrieval of various ASIC configurations, connectivity readings, and other relevant metrics.

```
mlxlink -d <device> --amber_collect <output_file.csv>
```

- **Output Report:** The tool generates an output file in CSV format containing collected data.
- **Detection of Problems:** The report generated includes columns for Raw, Effective, and Symbol Bit Error Rate (BER) to facilitate the detection of issues on specific ports.
- **IB XDR KPI Values:** Provides Key Performance Indicator (KPI) values for UPHY (Unified PHY). Refer to Table 18-1 for IB.

Table 18-1. IB XDR KPI Values

Link Type	FEC Interleaving (514,516, 30)	PLR	CRC Protected	External BER	Internal BER
IB (1X XDR) over short optics/copper	2 way interleave	With PLR	Yes	1E-25[2]	1E-26[2]
IB (1X XDR) over long optics	4-way interleave[3]	No PLR	Yes	1E-19	1E-20

Notes:

[1]The FEC interleaving scheme implemented on the physical line Power-On-Reset (POR) exclusively employs Symbol interleaving. Bit interleaving is reserved only for GearBox applications.

[2]The targeted BER for InfiniBand (IB) of 1e-19 externally and 1e-20 internally is a sufficient request.

[3]A preferred configuration involves a two-way interleave setup to mitigate Packet Loss Ratio (PLR), subject to characterization results. A 4-way interleave configuration lacks L1 capability.

If errors are detected on a link, perform a deeper, dedicated test using the PRBS test in test mode, as described in the following section.

18.2 PRBS in Test Mode - mlxlink

The PRBS test serves to assess the quality of the link.

Note: Execute the following commands on both ends of the test to function properly.

- Activate PRBS mode, run:
  ```
  mlxlink -d --test_mode EN
  ```
  - The default speed per lane is 25G and PRBS31.
- Adjust default speeds using flags: --tx_rate and --rx_rate.
- For GB NVL systems with DC-coupled cable cartridges, use the –dc_cpl_allow option with mlxlink while adjusting the default speeds.
- Tune the receiver, run:
  ```
  mlxlink -d device_name --test_mode TU
  ```
- Check results, run:
  ```
  mlxlink -d <device_name> -c
  ```
- Ensure successful tuning and that all lanes are locked.
- Deactivate PRBS mode, run:
  ```
  mlxlink -d <device> --test_mode DS
  ```

18.3 Loopback Test

If peer connectivity for PRBS testing is not available, a loopback test can detect issues with the board or module.

Activate loopback mode using the following command:

```
mlxlink <device_name> -l <loopback_mode>
```

Where <loopback_mode> options include:

- NO (no loopback)
- RM (phy remote Rx-to-Tx loopback)
- PH (internal Phy Tx-to-Rx loopback)
- EX (external loopback, requires connector)
- LL (link layer local loopback)

Select the appropriate loopback mode based on your requirements.

18.4 Sideband Management

MST devices serve as hardware interfaces for accessing devices. Each interface is represented by a Linux file located at /dev/mst.

- **PCI (configuration cycles):** This mode is the most used. Example: /dev/mst/mt4123_pciconfl
- **PCI (direct memory access):** Requires access to PCI memory. Not compatible with Live-Fish mode (NO firmware on device). Example: /dev/mst/mt4123_pci_cr0
- Utilizing an MST device through I2C with USB involves physical bus access facilitated by the MTUSB USB to I2C adapter. Example: /dev/mst/mtusb-1
- **Remote Access:** Remote MST devices transmit access requests over a TCP/IP socket to the destination device. Example: fit-l-vrt-80:23108,@dev@mst@mt4123_pciconf1

In-Band Access: Devices can be accessed in-band through the InfiniBand fabric.

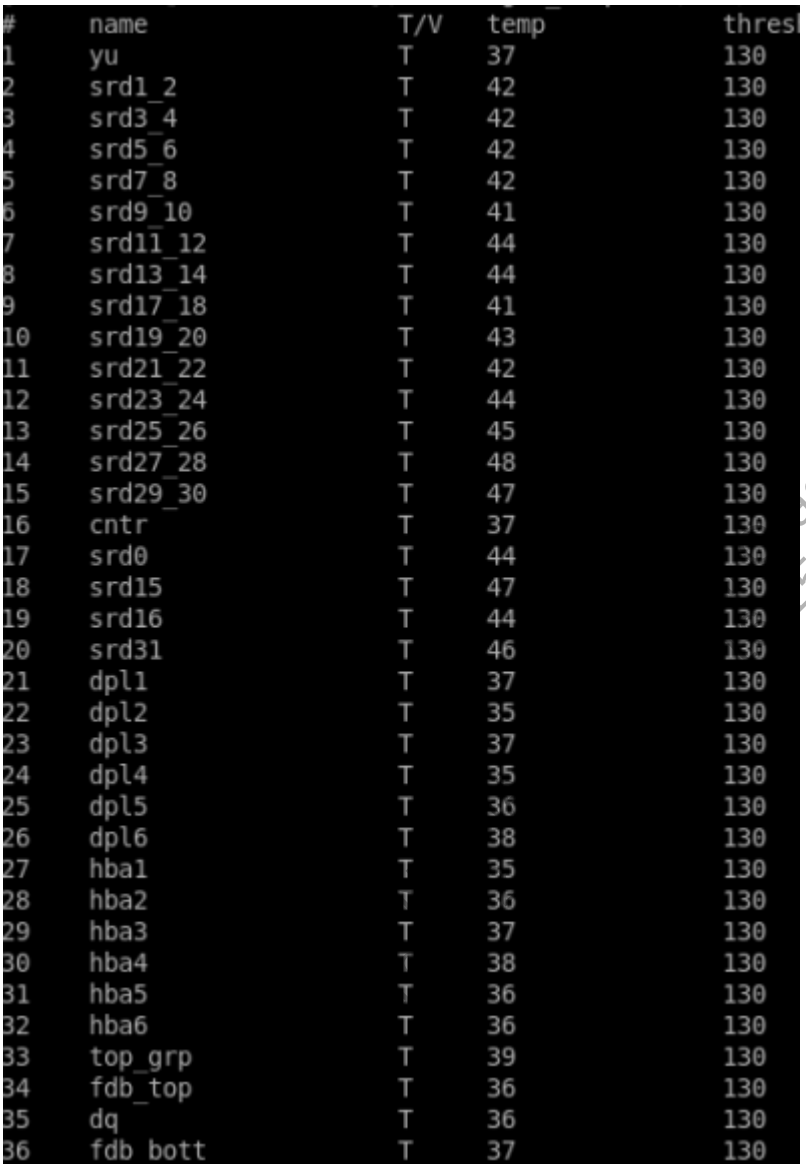
18.5 Quantum-3 Thermal

For remote temperature readings from the ASIC on HGX-8 platforms, run the following command:

```
mget_temp -d <device> -v
```

Example Output:

```
sudo mget_temp -d /dev/mst/mt54004_pciconf0
41
```



#	name	T/V	temp	thresh
1	yu	T	37	130
2	srd1_2	T	42	130
3	srd3_4	T	42	130
4	srd5_6	T	42	130
5	srd7_8	T	42	130
6	srd9_10	T	41	130
7	srd11_12	T	44	130
8	srd13_14	T	44	130
9	srd17_18	T	41	130
10	srd19_20	T	43	130
11	srd21_22	T	42	130
12	srd23_24	T	44	130
13	srd25_26	T	45	130
14	srd27_28	T	48	130
15	srd29_30	T	47	130
16	cntr	T	37	130
17	srd0	T	44	130
18	srd15	T	47	130
19	srd16	T	44	130
20	srd31	T	46	130
21	dpl1	T	37	130
22	dpl2	T	35	130
23	dpl3	T	37	130
24	dpl4	T	35	130
25	dpl5	T	36	130
26	dpl6	T	38	130
27	hba1	T	35	130
28	hba2	T	36	130
29	hba3	T	37	130
30	hba4	T	38	130
31	hba5	T	36	130
32	hba6	T	36	130
33	top_grp	T	39	130
34	fdb_top	T	36	130
35	dq	T	36	130
36	fdb_bott	T	37	130

From the switch terminal, run the command:

```
show temperature
```

This command displays both ASIC and board temperatures.

The maximum permitted threshold for NVIDIA Quantum-3 IC is 105°C on junction (T_j). The IC shuts down at 110°C.

18.6 Quantum-3 Power

Q3200 Switch Output Example:

```
admin@croc-94:~$ nv show platform environment voltage
Name                             Actual (V)  Maximum (V)  Minimum (V)  State
-------------------------------  ----------  -----------  -----------  -----
PMIC-1-12V-VDD-ASIC1-In-1        12.25       16.00        0.00         ok
PMIC-1-ASIC1-VDD-Out-1           0.70        1.00         0.50         ok
PMIC-2-12V-HVDD-DVDD-ASIC1-In-1  12.22       16.00        0.00         ok
PMIC-2-ASIC1-DVDD-PL0-Out-2      0.84        1.00         0.59         ok
PMIC-2-ASIC1-HVDD-PL0-Out-1      1.20        1.40         1.00         ok
PMIC-3-12V-HVDD-DVDD-ASIC1-In-1  12.31       16.00        0.00         ok
PMIC-3-ASIC1-DVDD-PL1-Out-2      0.84        1.00         0.59         ok
PMIC-3-ASIC1-HVDD-PL1-Out-1      1.20        1.40         1.00         ok
PMIC-4-12V-VDD-ASIC2-In-1        12.31       16.00        0.00         ok
PMIC-4-ASIC2-VDD-Out-1           0.71        1.00         0.50         ok
PMIC-5-12V-HVDD-DVDD-ASIC2-In-1  12.25       16.00        0.00         ok
PMIC-5-ASIC2-DVDD-PL0-Out-2      0.83        1.00         0.59         ok
PMIC-5-ASIC2-HVDD-PL0-Out-1      1.20        1.40         1.00         ok
PMIC-6-12V-HVDD-DVDD-ASIC2-In-1  12.31       16.00        0.00         ok
PMIC-6-ASIC2-DVDD-PL1-Out-2      0.84        1.00         0.59         ok
PMIC-6-ASIC2-HVDD-PL1-Out-1      1.20        1.40         1.00         ok
PMIC-7-12V-MAIN-In-1             12.25       16.00                     ok
PMIC-7-CEX-VDD-Out-1             1.03        1.46         0.66         ok
PSU-1-12V-Out                    12.27       12.95        11.00        ok
PSU-2-12V-Out                    12.25       12.95        11.00        ok
PSU-3-12V-Out                    12.27       12.95        11.00        ok
PSU-4-12V-Out                    12.27       12.95        11.00        ok
```


18.7 IBDiagnet Tool

The ibdiagnet tool can be used to provide power supply data.

In the ibdiagnet2.db_csv file, search for "START_POWER_SENSORS" to find: NodeGuid, SensorIndex, SensorName, Voltage, Current

18.8 Security

To get the life cycle secure firmware status, run:

```
flint -d <device> -qq q full

admin@juliet-68:~$ sudo flint -d /dev/mst/mt54004_pciconf0 -qq q full
Image type:            FS5
FW Version:            35.2014.0402
FW Release Date:       15.5.2024
Part Number:           920-9K36F-00MV-JS0_QPN_Ax
```

```
Description:           QPN for NVIDIA Quantum-3; NVLink Tray Solution; 1U; with 72 NVLink5 Internal
Ports and 72 NVLink5 External Ports
Product Version:       35.2014.0402
Rom Info:              type=UEFI version=skipped cpu=skipped
                       type=PXE  version=skipped devid=skipped cpu=skipped
                       type=NVMe version=skipped devid=skipped cpu=skipped
Description:           UID                GuidsNumber
Base GUID:             9c63c0030072b213        8
Orig Base GUID:        N/A                     8
Base MAC:              9c63c072b213            8
Orig Base MAC:         N/A                     8
Image VSD:             N/A
Device VSD:            N/A
PSID:                  MT_0000001185
Security Attributes:   N/A
Default Update Method: fw_ctrl
Life cycle:            PRODUCTION
Secure Boot Capable:   Disabled
Encryption:            Disabled
```

Chapter 19. Troubleshooting Networking Interfaces

19.1 PCIe Interface

PCIe is used in any system for communication between different modules. Network adapters need to communicate with the CPU and memory (among other modules). This means that to process network traffic, the different devices communicating via the PCIe should be well configured. When connecting the network adapter to the PCIe interface, automatic negotiation occurs to establish the highest supported capabilities between the network adapter and the CPU.

19.1.1 PCIe Troubleshooting

To gain deeper insights into the PCIe issue, address the following:

- **Description of the Issue:** Provide details about the server, device, network topology, device PSID, and any security functionalities in place.
- **Reproduction Steps:** Outline the steps to replicate the issue and specify the frequency of occurrence. If it involves a cluster, indicate the number of affected nodes.
- **Recent Changes:** Identify any alterations made before the issue arose.
- **Special Components:** Determine if there are any unique risers, switches, cables, or repeaters in the setup.
- **Compatibility Testing:** Verify if the issue persists when using another similar NVIDIA product.

For initial PCIe link debugging, use the mlxlink tool. For more information, refer to the "mlxlink Utility" documented under "Debug Utilities" in the latest NVIDIA Firmware Tools (MFT) Documentation.

To determine the PCIe link speed and width, use the following flag: `--port_type PCIE`.

Output Example:

```
PCIe Operational (Enabled) Info
-------------------------------
Depth, pcie index, node     : [Depth, pcie index, node]
Link Speed Active (Enabled) : [Freq – Gen]
```

```
Link Width Active (Enabled) : [Width]
To determine the physical PCIe counters, use the following flag: --port_type PCIE –c flag
```

Output Example:

```
Management PCIe Timers Counters Info
------------------------------------
dl down                         : [link down counter]

Management PCIe Performance Counters Info
-----------------------------------------
RX Errors                       : [Rx Errors]
TX Errors                       : [Tx Errors]
CRC Error dllp                  : [CRC Errors dllp]
CRC Error tlp                   : [CRC Errors tlp]
RX Errors: indicate the number of transitions to recovery required due to framing errors and CRC
(dlp and tlp) errors.
TX Errors: indicate the number of transitions to recovery required due to EIEOS and TS errors.
CRC Error dllp: indicate CRC error in Data Link Layer Packets.
CRC Error tlp: indicate CRC error in Transaction Layer Packet.
```

19.1.2 PCIe Useful Metrics

To assess the PCIe link margin through the mlxlink utility, run:

```
mlxlink -d [device] --port_type PCIE --margin
```

19.1.3 Pass or Fail Criteria

The following tables list the pass or fail criteria for PCIe Gen 3.0, 4.0, and 5.0.

Table 19-1. PCIe Gen 3.0

Last FOM	Criteria
Last FOM < 70	Fail
70 < Last FOM < 99	Marginal
100 < Last FOM	Pass

Table 19-2. PCIe Gen 4.0

Last FOM	Criteria
Last FOM < 70	Fail
70 < Last FOM < 99	Marginal
100 < Last FOM	Pass

Table 19-3. PCIe Gen 5.0

Last FOM	Criteria
Last FOM < 49	Fail
50 < Last FOM < 69	Marginal
70 < Last FOM	Pass

19.1.4 Error Detection

The following table lists the common PCIe related issues.

Table 19-4. PCIe Related Issues

Error Type	Debugging Steps
PCIe link down	Check PERST and clock PLL signals
Speed degradation	> Check the expected card and server speed > Replace riser card
Width degradation	> Check the expected card and server width > Replace PCIe slot
Server is hanging or in Non-Maskable Interrupt (NMI)	> Capture three MST dumps using the mtusb utility > Replicate the scenario in Livefish mode
Firmware assertion error	Attempt to improve or cancel the completion timeout
AER	Review error type
The "PCIe" term appears in dmesg	Capture sysinfo or three MST dumps
Performance issue	Tune MaxReadReq
Signal integrity	Check the capacitors on the card

19.1.5 Error Reporting and Logging

Review the Advanced Error Reporting (AER) by running the `lspci -vvv` command.

Table 19-5. Error Reporting and Logging

Error Type	Description	Unit
ABORT	Cancels the current pending error, if exists.	Not applicable
BAD_DLLP_LCRC	Flips a bit in the LCRC of the next “error_duration” DLLPs that are transmitted through the port.	Packets
BAD_TLP_LCRC	Flips a bit in the LCRC of the next “error_duration” TLPs that are transmitted through the port. The packets are VDM TLPs that are sent by the port to the destination BDF - “dbdf”.	Packets

Error Type	Description	Unit
BAD_TLP_ECRC	Flips a bit in the ECRC of the next “error_duration” TLPs that are transmitted through the port. The packets are VDM TLPs that are sent by the port to the destination BDF - “dbdf”.	Packets
ERR_MSG	Sends an error signaling message to the RC.	Packets
MALFORMED_ TLP	Sends an “error_duration” PM_ACTIVE_STATE_NACK message to the destination BDF - “dbdf” with TC=1 instead of 0.	Packets
POISONED_TLP	Sends an “error_duration” VDMs with data to the destination BDF - “dbdf” with EP = 1.	Packets
UNEXPECTED_CPL	Sends “error_duration” completions to the destination BDF - “dbdf” with 0xff tag.	Packets
ACS_VIOLATION	Sends “error_duration” VDMs to the destination BDF - “dbdf” with source_bdf=0.	Packets
SURPRISE_LINK_DOWN	Sets a port state to DETECT.	Not applicable
RECEIVER_ERROR	Sends a clock instead of data for “error_duration” usecs. A value of 0 in ‘error_duration’ means that this error must be toggled by the firmware as fast as possible.	uSec

19.2 Ethernet and InfiniBand Interfaces

Ethernet and InfiniBand (IB) represent advanced networking technologies recognized for their high performance, primarily applied in high-performance computing (HPC) environments and data centers.

19.2.1 Ethernet and InfiniBand Troubleshooting

Troubleshooting the InfiniBand and Ethernet links is divided into different scenarios, outlined in Table 19-6.

19.2.2 Link is Down

Table 19-6 lists the troubleshooting for ethernet and InfiniBand when the link is down.

Table 19-6. Link is Down

Interface	Troubleshooting
Ethernet	If the link is inactive and has either never been active or was active previously but became inactive, verify the following elements: > Ensure that the link is properly connected on both ends. > Confirm that the link is configured with the same protocol on both ends: • If auto-negotiation is enabled, ensure it is activated on both sides. • Disable Auto-Negotiation and manually set the speed and FEC at both ends of the link. > Check the output of the mlxlink tool, which is part of the MFT package. For more information, refer to the "mlxlink Utility" documented under "Debug Utilities" in the latest: NVIDIA Firmware Tools (MFT) Documentation • Execute the mlxlink tool with the -m flag. For an example, refer to: Utility Output Example > If possible, conduct a loopback test to determine if the link will activate. This involves connecting the cable and module to two ports on the same device. • IMPORTANT: Be cautious as this setup may trigger a loop; ensure the ports used are NOT within the same L2 segment. • If the link becomes active during the loopback test on both ends, it indicates a potential interoperability issue between the link partners. Verify this scenario using a certified cable by referring to the network device product brief and data sheet.
InfiniBand	In this situation, it is important to verify the following components: > If the interface is in the "INIT" state, ensure that OpenSM is operational in the network and has a physical path (via switches) to the affected devices. > Ensure that the link is properly connected on both ends. > Check the output of the mlxlink tool, which is part of the MFT package. For more information, refer to the "mlxlink Utility" documented under "Debug Utilities" in the latest

19.2.3 Link is Flapping

> Reseat the cable. If this resolves the issue, refer to Error Detection to monitor the link and possibly replacing the cable, as it may degrade in quality over time.

> If the interface continues to flap, try relocating the cable to a different port to determine if the issue persists with the cable or the port. If the issue persists with the cable, consider replacing it with a verified spare. For optic connections, ensure the optical connectors are clean and retry the link before considering replacing the fiber itself. If the problem recurs, it is likely that one of the transceivers is failing.

> Use the "What-Just-Happened" feature on compatible switch systems to examine Layer 1 link down reasons. An example for Cumulus Linux switch OS can be found here, and similar features are available on SONiC and Onyx platforms.

19.2.4 Ethernet and InfiniBand Useful Metrics

To better understand the mlxlink utility output and for useful link debug matrics, refer to: Understanding mlxlink Utility Output.

19.2.5 Ethernet and InfiniBand Error Detection

The following table lists the error detection for ethernet and InfiniBand interfaces.

Table 19-7. Ethernet and InfiniBand Error Detection

Protocol	References
Ethernet	Use the What Just Happened (WJH) tool which provides real time visibility into network problems. For more information, refer to What-Just-Happened in the latest NVIDIA Cumulus Linux User Guide. Use SNMP to monitor link errors and transceiver DDMI values. For more information, refer to the "Simple Network Management Protocol - SNMP" in the latest NVIDIA Cumulus Linux User Guide
InfiniBand	Refer to the NVIDIA InfiniBand Cluster Operation and Maintenance Guide.

19.2.6 Ethernet and InfiniBand Error Reporting and Logging

The following table lists the error reporting and logging for ethernet and InfiniBand interfaces.

Table 19-8. Ethernet and InfiniBand Error Reporting and Logging

Protocol	References
Ethernet	For Cumulus Linux, refer to Troubleshooting Network Interfaces
InfiniBand	Refer to the "Supported Counters and Events" section, refer to the NVIDIA UFM Enterprise User Manual here

Chapter 20. Understanding the mlxlink Utility Output

The NVIDIA Firmware Tools (MFT) package is a set of firmware management and debug tools for NVIDIA devices. This section describes MFT features, tools, and configuration. For more information on the MFT tool, refer to the NVIDIA Firmware Tools (MFT) Documentation.

The mlxlink tool is used to check and debug link status and related issues. The output includes information regarding the following aspects, each described as follows with an output example.

- Operational Information
- Supported Information
- Troubleshooting Information
- Physical Counters and BER Information
- Module Information

20.1 Operational Information

The following table lists the details of the operational status of the module and parameters.

Table 20-1. Operational Information

Parameter	Description
State	Admin status of the port. - "Active": link is active - "Polling": system is in the process of establishing a link
Physical State	Represents the current state of the Layer 1 machine during the link-up process. - "LinkUp": Indicates successful establishment of the link - "ETH_AN_FSM_ENABLE": Indicates the initiation of the link establishment process - "ETH_AN_FSM_ABILITY_DETECT": Signifies recognition of a link partner and efforts to establish optimal link parameters

Parameter	Description
	> "ETH_AN_FSM_AN_GOOD_CHECK": Marks the conclusion of link establishment with agreement on optimal parameters, signaling the imminent "Link Up" status
Speed	Link speed
Width	PCIe width
FEC	Indicated the FEC mechanism
Loopback Mode	Indicates if Loopback mode is activated
Auto Negotiation	Indicates if the auto-negotiation feature is activated

Output Example:

```
Operational Info
----------------
State                           : Active
Physical state                  : LinkUp
Speed                           : 100G
Width                           : 4x
FEC                             : Standard RS-FEC - RS(528,514)
Loopback Mode                   : No Loopback
Auto Negotiation                : ON
```


20.2 Supported Information

The following table provides details on the enabled link and supported cable speeds.

Table 20-2. Supported Information

Parameter	Description
Enabled Link Speed	The speed the interface is configured for, both Onyx and Cumulus will automatically set this to the highest rate the port supports.
Supported Cable Speed	The speeds the module declares it supports based on the EEPROM.

Output Example:

```
Supported Info
--------------
Enabled Link Speed (Ext.)       : 0x00000200 (100G_4X)
Supported Cable Speed (Ext.)    : 0x000002f2 (100G_4X,50G_2X,40G,25G,10G,1G)
```

20.3 Troubleshooting Information

This section offers information on the link's status. The most important row is the "Recommendation" row. Refer to Section 20.5 “Link Status Opcodes” for descriptions of the various link status OpCodes.

Output Example:

```
Status Opcode                       : 0
Group Opcode                        : N/A
Recommendation                      : No issue was observed.
```

20.4 Physical Counters and BER Information

This section provides information on the module, with the key focus being on the "Rx/Tx Power Current" row seen in the below output example.

In "R/Tx Power Current [dBm]: 2,0,0,2 [-6..8].”

- 2,0,0,2 represent the current power in dBm scale (1dBm = ~1 milliwatt), with each value corresponding to a lane. In the case of a QSFP module, there are 4 lanes, whereas an SFP module would have only 1 lane.
- [-6..8] represent the scale of acceptable power values. If the current power presented in 2,0,0,2 is below the lowest scale value:
- For Rx, it indicates a lack of signal reception from the peer, often accompanied by a recommendation of "No partner detected."
- For Tx, it suggests that no signal is being transmitted to the peer (link is disabled).

If the Rx value exceeds the highest scale value, it implies that the customer should use a longer fiber cable or an attenuator to decrease the signal strength.

It is unlikely for the Tx value to surpass the highest scale value; if encountered, escalation is recommended.

Note: The power section is not accessible for DAC cables or optic devices lacking Digital Diagnostic and Monitoring (DDMI) support. In such cases, these values are derived from the module EEPROM.

Output Example:

```
Module Info
-----------
Identifier                          : QSFP28
Compliance                          : 100G AOC (Active Optical Cable) or 25GAUI C2M AOC with FEC
Cable Technology                    : 850 nm VCSEL
Cable Type                          : Active cable (active copper / optics)
OUI                                 : Other
Vendor Name                         : O-NET
```

```
Vendor Part Number              : 1AT-3Q4M01XX-12A
Vendor Serial Number            : 2QA-0050444
Rev                             : 03
Wavelength [nm]                 : 850
Transfer Distance [m]           : 1
Attenuation (5g,7g,12g) [dB]    : N/A
FW Version                      : N/A
Digital Diagnostic Monitoring   : Yes
Power Class                     : 2.5 W max
CDR RX                          : ON,ON,ON,ON
CDR TX                          : ON,ON,ON,ON
LOS Alarm                       : N/A
Temperature [C]                 : 36 [-5..75]
Voltage [mV]                    : 3352.3 [2970..3630]
Bias Current [mA]               : 6.392,6.166,6.240,6.240 [2..10]
Rx Power Current [dBm]          : 2,0,0,2 [-6..8]
Tx Power Current [dBm]          : 4,2,2,2 [-6..8]
```


20.5 Link Status Opcodes

The following table lists the various link status Opcodes, along with their descriptions and recommended mitigation methods.

Table 20-3. Link Status Opcodes

Status Opcode	Description	Recommended Mitigation
PHY FW indication (0 - 1023):		
0 – No issue observed		Wait 5 seconds and check again. If the message continues, please check peer side. Additionally, review the mlxlink -m flag output and to see if anything is suspicious (for example with transceivers. Rx power is -40dbm then need to check fiber and peer side; example 2: Tx power -40dbm.
1 – Port is closed by command	Port is shut down by system.	Need to check who sent the command to close the port and reopen it.
2 – AN failure	Both sides did not agree on speed/FEC or DME is missing.	Debug steps: 1) Check Tx power and Rx power from both sides. a. Low Tx power: Check transceiver issue b. Low Rx power: Check Tx power from peer side+ clean fiber and transceiver (both ends). 2) Check both sides are configured correctly. a. Same speeds and FECs or that AN is fully enabled.

Status Opcode	Description	Recommended Mitigation
9 – Logical mismatch between link partners	Did not acquire block lock	1) Check both sides are configured correctly. a. Same speeds and FECs or that AN is fully enabled. 2) If the issue repeats, please collect data from both sides and report to NVIDIA.
10 – Logical mismatch between link partners	Did not acquire AM lock (NO FEC)	1) Check both sides are configured correctly. a. Same speeds and FECs or that AN is fully enabled. 2) If the issue repeats, please collect data from both sides and report to NVIDIA.
11 – Logical mismatch between link partners	Did not get align status. AN is done but the signal is not locked. Very rare event.	1) Check both sides are configured correctly. a. Same speeds and FECs or that AN is fully enabled. 2) If the issue repeats, please collect data from both sides and report to NVIDIA.
12 – Logical mismatch between link partners	FC FEC is not locked.	1) Check both sides are configured correctly. a. Same speeds and FECs or that AN is fully enabled. 2) If the issue repeats, please collect data from both sides and report to NVIDIA.
13 – Logical mismatch between link partners	RS FEC is not locked.	1) Check both sides are configured correctly. a. Same speeds and FECs or that AN is fully enabled. 2) If the issue repeats, please collect data from both sides and report to NVIDIA.
14 – Remote fault received		Wait 5 seconds and check again. If the message persists, please check peer side.
15 – Bad signal integrity	Low Raw BER. Please ensure to have it running for minimum duration before re-checking.	The link is up, but with low BER it is indication for monitoring and review. There are two different steps to debug- with Symbol BER and clean Symbol BER: 1) Wait to test again after some time and see it is still reoccurring. 2) Read the mlxlink -c flag. 3) If the Symbol BER is clear of errors, then we just need to monitor the port and see if the Raw BER and Effective BER are stable over time. 4) If there are Symbol BERs then check Low Rx power (mlxlink -m) : Check Tx power

Status Opcode	Description	Recommended Mitigation
		from peer side + clean fiber and transceiver (both ends) 5) Collect SNR (electrical and optical) from both sides: a. Available in mlxlink -m and other tools 6) In case the link stays with low BER, test for 125 minutes.
16 – Cable compliance code mismatch (protocol mismatch between cable and port)		1) Need to see the port speed is configured as expected with the cable. 2) Need to see if the cable is the right one for the port if it is as expected.
20 – Stamping of non-NVIDIA Cables/Modules		Replace the cable with an NVIDIA cable.
22 – Internal error		Not relevant for the field.
30 – Port is closed, no backplane enabled speed over backplane channel		1) Check the port is configured correctly: a. Same speeds, width and FECs or that AN is fully enabled.
31 – Port is closed, no passive protocol enabled over passive copper channel		1) Check the port is configured correctly: a. Same speeds, width and FECs or that AN is fully enabled.
32 – Port is closed, no active protocol enabled overactive channel		1) Check the port is configured correctly: a. Same speeds, width and FECs or that AN is fully enabled.
33 – Port width does not match the port speed enabled		1) Check the port is configured correctly: a. Same speeds, width and FECs or that AN is fully enabled.
34 – Local Speed degradation		The link is up, but with lower speed than expected. Steps: 1) Wait to test again after some time 2) Cleaning the fiber from both sides + the transceivers (can be delayed being a final step before escalating) 3) Look at the Tx power and Rx Power a. Low Tx power: Check transceiver issue b. Low Rx power: Check Tx power from peer side+ clean fiber and transceiver (both ends) 4) Collect SNR (electrical and optical) from both sides. a. Available in mlxlink -m and other tools. 5) In case the link stays with low BER, test with PRBS. a. See the steps in the mlxlink help flag or in the attached excel.

Status Opcode	Description	Recommended Mitigation
		b. In case specific lane doesn't lock, it might be transceiver, cable, fiber issue.
35 – Remote Speed degradation		Review remote side status.
36 – No Partner detected during force mode.		Debug steps: 1) Check Tx power and Rx power from both sides. a. Low Tx power: Check transceiver issue b. Low Rx power: Check Tx power from peer side+ clean fiber and transceiver (both ends) 2) Check both sides are configured correctly: a. Same speeds and FECs or that AN is fully enabled.
37 – Partial link indication during force mode.		Debug steps: 1) Check Tx power and Rx power from both sides. a. Low Tx power: Check transceiver issue b. Low Rx power: Check Tx power from peer side+ clean fiber and transceiver (both ends) 2) Check both sides are configured correctly: a. Same speeds and FECs or that AN is fully enabled.
38 – AN failure	FEC mismatch during override	1) Check both sides are configured correctly. a. Same speeds and FECs or that AN is fully enabled.
39 – AN failure	No HCD	1) Check both sides are configured correctly. a. Same speeds and FECs or that AN is fully enabled.
42 – Bad SI, cable is configured to non-optimal rate		1) Check the port is configured correctly: a. Same speeds, width and FECs or that AN is fully enabled.
51– HST speed mismatch		1) Check both sides are configured correctly. a. Same speeds and FECs or that AN is fully enabled. 2) It seems to me that for more than that, collect data and report to NVIDIA.
52 – Eq Fail	Equalization failure	Collect logs. Toggle the port.

Status Opcode	Description	Recommended Mitigation
		Reset the firmware.
53 – Link failure due to MCB at link up		Wait for 10 seconds, and if the message is reread then share information from both sides and toggle the link.
54 – PLR didn't get Rx good non sync cell		Wait for 10 seconds, and if the message is reread then share information from both sides and toggle the link.
55 – PSI fatal error		Wait for 10 seconds, and if the message is reread then share information from both sides and toggle the link.
57 – Signal not detected	Power in the Serdes is not detected.	1) Wait to test again after some time. 2) Look at the Tx power and Rx Power. a. Low Tx power: Check transceiver issue b. Low Rx power: Check Tx power from peer side+ clean fiber and transceiver (both ends). 3) In case the link stays with low BER, test with PRBS. a. See the steps in the mlxlink help flag or in the attached excel.
59 – Did not get module configuration done		1) Wait to test again after some time. 2) Look at the Tx power and Rx Power. a. Low Tx power: Check transceiver issue b. Low Rx power: Check Tx power from peer side+ clean fiber and transceiver (both ends). 3) In case the link stays with low BER, test with PRBS. a. See the steps in the mlxlink help flag or in the attached excel. 4) Collet mlxlink and mstdumps and share with NVIDIA. a. In case of successful PRBS results: need to debug the FW. b. In case of low BER PRBS results: need to debug the Serdes. c. In case specific lane doesn't lock, well it gets in interesting and might be transceiver, NIC, FW or Serdes.
128 – Troubleshooting in process		Wait 3 seconds and run the command again.
1023 – Info not available		Wait for 10 seconds, and if the message is reread then share information from both sides and run power cycle.

Status Opcode	Description	Recommended Mitigation
1024 – Cable is unplugged	No physical transceiver detected on cage.	Plug transceiver. Please confirm that no one executed command simulating unplugged transceiver.
1025 – Long Range for non-Mellanox cable/module		Should not be seen. Please collect logs in failed state and contact NVIDIA for support.
1026 – Bus stuck (I2C Data or clock shorted)	Received failure on the I2C EEPROM communication line.	Transceiver reset (Disable/enable), if the issue continues, please collect information and data, and then run power cycle.
1027 – Bad/unsupported EEPROM	Failed to read EEPROM from transceiver or transceiver id is not recognized.	Test with another approved transceiver. Id the issue continues, please collect data and share.
1028 – Part number list	Transceiver is not permitted by vendor list.	Replace the cable with cable from the supported list.
1029 – Unsupported cable.	SFP transceiver is not supported.	Replace the cable with cable from the supported list.
1030 – Module temperature shutdown	Transceiver temperature exceeded allowed threshold.	Check the cable temperature and cool the environment if it is indeed too hot.
1031 – Shorted cable	Receive over current on the transceiver.	Bad transceiver, please test with a different transceiver.
1032 – Power budget exceeded	Board power budget have exceeded.	Review supported power by the transceiver and board INI.
1033 – Management forced down the port	Module shutdown by server command.	Review the server commands.
1034 – Module is disabled by command	Transceiver admin status is disabled.	Enable admin status.
1035 – System Power is Exceeded therefore the module is powered off	Power drawn from the system is above supply, which resulted in power throttling in module.	Check system power input. Could be a system platform issue.
1036 - Module's PMD type is not enabled (see PMTPS)	Transceiver type not supported.	Replace transceiver.
1040 – pcie system power slot Exceeded	PCIe device required more power than what the slot if capable of.	Hardware issue on host system. Only applicable to NICs. Not applicable to NVIDIA Switch devices.
1042 – Module state machine fault	Module CMIS state machine returned fault for state transition.	Reset transceiver and replace if problem repeats.
1043 – Module's stamping speed degeneration		Replace the cable with an NVIDIA cable.
1044 – Module's stamping speed degeneration	HDR speed is not supported	Replace the cable with an NVIDIA cable.
1045 – Module's stamping speed degeneration	EDR speed is not supported	Replace the cable with an NVIDIA cable.

Status Opcode	Description	Recommended Mitigation
1046 – Module's stamping speed degeneration	FDR10 speed is not supported	Replace the cable with an NVIDIA cable.
1047 – Modules DataPath FSM fault	Failed to configure speed (application) by transceiver	Wait for 10 seconds, and if the message is reread.
1048 – Modules DataPath FSM fault	Failed to activate speed (application) by transceiver	Wait for 10 seconds, and if the message is reread then share information from both sides and run power cycle.
1050 – Module Boot Error	Module failed to boot. No response with I2C bus.	Reset transceiver and replace if problem repeats.
1051 – Module Boot Error	Module failed to boot. No response with I2C bus.	Reset transceiver and replace if problem repeats.
1052 – Module Boot Error	Module failed to boot. No response with I2C bus.	Reset transceiver and replace if problem repeats.
1053 – Module Boot Error	Module failed to boot. No response with I2C bus.	Reset transceiver and replace if problem repeats.
1054 – Module Forced to Low Power by command	Module was set to low power by command (EEPROM write or GPIO).	Check byte 26, bit 6, page 0x0. Ensure it is not set.
CPO only (1055 – 1059)		
1055 – ELS laser fiber is contaminated	Applicable to systems with CPO only.	
1056 – ELS laser fiber	Applicable to systems with CPO only.	
1057 – ELS unplugged	Applicable to systems with CPO only.	
1058 – ELS fatal indication	Applicable to systems with CPO only.	
1059 – ELS temperature error	Applicable to systems with CPO only.	
Core/Driver (2048 - 3071)		
2048 – MPR Violation (Under 64 bytes between two starts)		Wait for 10 seconds, and if the message is reread then share information from both sides and run power cycle.

Notice

This document is provided for information purposes only and shall not be regarded as a warranty of a certain functionality, condition, or quality of a product. NVIDIA Corporation (“NVIDIA”) makes no representations or warranties, expressed or implied, as to the accuracy or completeness of the information contained in this document and assumes no responsibility for any errors contained herein. NVIDIA shall have no liability for the consequences or use of such information or for any infringement of patents or other rights of third parties that may result from its use. This document is not a commitment to develop, release, or deliver any Material (defined below), code, or functionality.

NVIDIA reserves the right to make corrections, modifications, enhancements, improvements, and any other changes to this document, at any time without notice.

Customer should obtain the latest relevant information before placing orders and should verify that such information is current and complete.

NVIDIA products are sold subject to the NVIDIA standard terms and conditions of sale supplied at the time of order acknowledgement, unless otherwise agreed in an individual sales agreement signed by authorized representatives of NVIDIA and customer (“Terms of Sale”). NVIDIA hereby expressly objects to applying any customer general terms and conditions with regards to the purchase of the NVIDIA product referenced in this document. No contractual obligations are formed either directly or indirectly by this document.

NVIDIA products are not designed, authorized, or warranted to be suitable for use in medical, military, aircraft, space, or life support equipment, nor in applications where failure or malfunction of the NVIDIA product can reasonably be expected to result in personal injury, death, or property or environmental damage. NVIDIA accepts no liability for inclusion and/or use of NVIDIA products in such equipment or applications and therefore such inclusion and/or use is at customer’s own risk.

NVIDIA makes no representation or warranty that products based on this document will be suitable for any specified use. Testing of all parameters of each product is not necessarily performed by NVIDIA. It is customer’s sole responsibility to evaluate and determine the applicability of any information contained in this document, ensure the product is suitable and fit for the application planned by customer, and perform the necessary testing for the application in order to avoid a default of the application or the product. Weaknesses in customer’s product designs may affect the quality and reliability of the NVIDIA product and may result in additional or different conditions and/or requirements beyond those contained in this document. NVIDIA accepts no liability related to any default, damage, costs, or problem which may be based on or attributable to: (i) the use of the NVIDIA product in any manner that is contrary to this document or (ii) customer product designs.

No license, either expressed or implied, is granted under any NVIDIA patent right, copyright, or other NVIDIA intellectual property right under this document. Information published by NVIDIA regarding third-party products or services does not constitute a license from NVIDIA to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property rights of the third party, or a license from NVIDIA under the patents or other intellectual property rights of NVIDIA.

Reproduction of information in this document is permissible only if approved in advance by NVIDIA in writing, reproduced without alteration and in full compliance with all applicable export laws and regulations, and accompanied by all associated conditions, limitations, and notices.

THIS DOCUMENT AND ALL NVIDIA DESIGN SPECIFICATIONS, REFERENCE BOARDS, FILES, DRAWINGS, DIAGNOSTICS, LISTS, AND OTHER DOCUMENTS (TOGETHER AND SEPARATELY, “MATERIALS”) ARE BEING PROVIDED “AS IS.” NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE MATERIALS, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE. TO THE EXTENT NOT PROHIBITED BY LAW, IN NO EVENT WILL NVIDIA BE LIABLE FOR ANY DAMAGES, INCLUDING WITHOUT LIMITATION ANY DIRECT, INDIRECT, SPECIAL, INCIDENTAL, PUNITIVE, OR CONSEQUENTIAL DAMAGES, HOWEVER CAUSED AND REGARDLESS OF THE THEORY OF LIABILITY, ARISING OUT OF ANY USE OF THIS DOCUMENT, EVEN IF NVIDIA HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. Notwithstanding any damages that customer might incur for any reason whatsoever, NVIDIA’s aggregate and cumulative liability towards customer for the products described herein shall be limited in accordance with the Terms of Sale for the product.

Trademarks

NVIDIA, the NVIDIA logo, BlueField, ConnectX, CUDA, NVIDIA GPU Boost, NVIDIA Grace, NVIDIA HGX, NVIDIA Hopper, NVLink, NVLink-C2C NVSwitch, and Tegra are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright

NVIDIA Corporation | 2788 San Tomas Expressway, Santa Clara, CA 95051
http://www.nvidia.com