NVIDIA UFM-SDN Appliance User Manual v4.14.0

Exported on Nov/27/2023 05:27 PM
https://docs.nvidia.com/networking/x/wBOgC

Table of Contents

About This Document 27
Software Download 28
Document Revision History 29
Release Notes 30
- Key Features 30
- Changes and New Features 31
  - Unsupported Functionalities/Features 33
- Installation Notes 34
  - Supported Devices 34
    - Supported NVIDIA Externally Managed Switches 34
    - Supported NVIDIA Internally Managed Switches 34
  - System Requirements 36
    - UFM GUI Client Requirements 36
    - MFT Package Version 36
    - UFM SM Version 37
    - UFM NVIDIA SHARP Software Version 37
- Bug Fixes in This Release 37
- Known Issues in This Release 38
- Changes and New Features History 39

Bug Fixes History 44
Known Issues History 58
Overview 69
Scale-Out Your Fabric with Unified Fabric Manager 69
UFM Benefits 69
Central Console for Fabric Management 69
In-Depth Fabric Visibility and Control 69
Advanced Traffic Analysis 70
Enables Multiple Isolated Application Environments on a Shared Fabric 70
Service-Oriented Automatic Resource Provisioning 70
Quick Resolution of Fabric Problems 70
Seamless Failover Handling 71
Open Architecture 71
Main Functionality Modules 71
Fabric Dashboard 71
Fabric Design and Configuration 71
Fabric Discovery and Physical View 72
Central Device Management 72
Monitoring 72
Configuration 72
Logging 72

High Availability 73
InfiniBand Fabric Managed by UFM 73
UFM Software Architecture 76
Graphical User Interface 78
Client Tier API 78
Client Tier SDK Tools 78
UFM Server 78
Subnet Manager 78
NVIDIA Scalable Hierarchical Aggregation and Reduction Protocol (SHARP)™ Aggregation Manager 78
Performance Manager 79
Device Manager 79
UFM Switch Agent 79
Communication Protocols 79
UFM Installation and Initial Configuration 80
Additional Configuration (Optional) 80
General Settings in gv.cfg 80
Enabling SHARP Aggregation Manager 80
Running UFM in Monitoring Mode 81
Enabling Predefined Groups 81
Enabling Multi-NIC Host Grouping 81
Defining Node Description Black-List 82

Running UFM Over IPv6 Network Protocol 82
Adding SM Plugin (e.g. lossymgr) to event_plugin_name Option 82
Multi-port SM 83
Configuring UDP Buffer 84
Virtualization 84
Static SM LID 84
Configuring Log Rotation 85
Configuration Examples in gv.cfg 85
SM Trap Handler Configuration 88
CPU Affinity on UFM 88
Quality of Service (QoS) Support 89
UFM Failover to Another Port 90
Delegating Authentication to a Proxy 91
Configuring Partial Switch ASIC Failure Events 92
Enabling Network Fast Recovery 93
Disabling Rest Roles Access Control 94
Setting up Telemetry in UFM 94
Enabling UFM Telemetry 95
Changing UFM Telemetry Default Configuration 95
Supporting Generic Counters Parsing and Display 96
Supporting Multiple Telemetry Instances Fetch 97

Secondary Telemetry 97
Installing License 100
Historical Telemetry Collection in UFM 100
Storage Considerations 100
UFM-SDN Appliance Installation and Initial Configuration 101
Configuring UFM-SDN Appliance for the First Time 101
Connecting UFM-SDN Appliance via Serial Port 101
Connecting UFM-SDN Appliance via Remote Management 102
Configuring UFM-SDN Appliance Using CLI Wizard 103
Obtaining the License 108
Initial UFM-SDN Application Configuration 111
Changing UFM Running Mode 111
Running the UFM Server in Monitoring Mode 112
Modifying UFM Configuration Files 114
Configuring General Settings in gv.cfg 116
Configuring Fabric Interface 116
Running UFM in Monitoring Mode 116
Enabling Predefined Groups 117
Enabling Multi-NIC Host Grouping 117
Running UFM SM Only (UFM HA with Additional SMs) 117
Running UFM Over IPv6 Network Protocol 118

Adding SM Plug-in to event_plugin_name Option 118
Enabling SHARP Aggregation Manager 118
Multi-port SM 119
Setting UDP Buffer 119
Virtualization 120
Static SM LID 120
Maximum Live Telemetry Sessions 120
Configuring Log Rotation 121
Configuration Examples in gv.cfg 121
SM Trap Handler Configuration 125
Configuring Two Appliances to Run in High Availability 125
Prerequisites for UFM High Availability (HA) Installation 125
Configuring UFM-SDN Appliance with External UFM-SMs 127
Configuring External UFM Subnet Manager 128
CPU Affinity on UFM-SDN Appliance 130
Quality of Service (QoS) Support 130
UFM Failover to Another Port 131
Starting UFM 133
User Authentication 133
Licensing 133
Launching a UFM Web UI Session 137

Installing UFM-SDN Appliance License 138
Setting Up Telemetry in UFM-SDN Appliance 139
Installing and Running UFM-SDN Agent Software 139
UFM Agent Installation Prerequisites 140
Installing UFM Agent Software 140
UFMA Installation Parameters 141
Running UFM Agent Software 142
InfiniBand Bonding on Managed Host 142
UFM SNMP Agent 142
UFM Agent Upgrade 143
Upgrading UFM-SDN Appliance Software 143
High Availability Out-Of-Service Upgrade 150
High Availability In-Service Upgrade 151
UFM-SDN Appliance Common Use Cases 153
Restoring UFM Application Data on the UFM®-SDN Appliance 153
Restoring Data on Standalone Configuration 153
Restoring Data on High Availability Configuration 154
Backing Up the UFM Database 155
Backing Up the UFM Configuration 155
Generating a Debug Dump File 156
Creating Users 156

Creating Local User 156
Linking Users to LDAP Server 157
Updating HCA Firmware 157
Recovering Primary and Secondary Servers Failure in Standby-Standby Mode 158
Getting Familiar with UFM's Data Model 160
Overview of Data Model 160
UFM Model Basics 160
Physical Model 160
UFM Web UI 162
UFM Web UI Main Navigation Buttons 162
Fabric Dashboard 164
Dashboard Views and Panel Management 165
Dashboard Timeline Snapshots 167
Dashboard Panels 167
Top N Servers/Switches by Rx or Tx Bandwidth 168
Top N Congested Servers/Switches by Rx/Tx Bandwidth 174
Top N Utilized PKeys 179
Top N Alarmed Servers/Switches 181
Inventory Summary 186
Fabric Utilization 188
Recent Activities 190

Traffic Map 192
Network Traffic Map 193
Levels Network Traffic Map 195
Port Group Traffic Map 199
Traffic Map Bar Chart 200
Events History 201
Network Map 204
Network Map Components 206
Selecting Map Elements 207
Map Information and Settings 208
Map View Tab 213
Map Zoom In Tab 214
Map Layouts 216
Information View Tab 219
Link Analysis 221
Topology Compare 229
Properties Tab 232
Network Map Elements Actions 236
Supported Actions for Internally Managed Switches 236
Supported Actions for Externally Managed Switches 238
Supported Actions for Hosts 239

Managed Elements 240
Devices Window 241
Mark Device as Unhealthy 245
Mark Device as Healthy 250
Upgrading Software and Firmware for Hosts and Externally Managed Switches 252
Device Information Tabs 263
Ports Window 285
Physical Grade and Eye Opening Information 289
Auto-isolation of High-BER Ports 290
Virtual Ports Window 292
Unhealthy Ports Window 295
Unhealthy Port Connectivity Filter 298
Health Policy Management 299
Cables Window 301
Groups Window 301
Inventory Window 306
PKeys Window 307
Creating New PKey 307
PKey Members Tab 311
PKey QoS Tab 311
PKey Actions 312

Support Pkey with Virtual Ports 313
HCAs Window 316
Events & Alarms 317
Device Status Events 319
Link Status Events 321
Telemetry 324
System Health 324
UFM Health Tab 325
UFM Logs Tab 328
Event Logs 330
Subnet Manager (SM) Logs 330
UFM Logs 330
Fabric Health Tab 330
Daily Reports Tab 338
Activating and Deactivating the Daily Report 338
Saving Daily Reports 339
Other Daily Report Configurations 339
Report Content 341
Topology Compare Tab 361
Overview 361
Topology Comparison Flow 364

Fabric Validation Tab 366
IBDiagnet Tab 371
Jobs 379
Settings 381
Events Policy 382
SNMP Settings 385
Configuring Email-on-Events 385
Device Access 389
Network Management 391
UFM Routing Protocols 392
Configuring Routing Protocol 393
Subnet Manager Tab 396
SM Keys Configuration 396
SM Limits Configuration 398
SM Lossy Manager Configuration 401
SM SL2VL Mapping Configuration 402
SM Sweep Configuration 402
SM Handover Configuration 404
SM Threading Configuration 405
SM Logging Configuration 407
SM Miscellaneous Settings 408

SM QoS Configuration 410
SM Congestion Control Configuration 411
SM Adaptive Routing Configuration 412
Non-Optimal Links 412
Email 415
Remote Location 419
Data Streaming 421
Topology Compare 423
Token-based Authentication 424
Plugin Management 425
User Preferences 433
User Management 436
User Authentication 436
Creating User Accounts 436
User Account Management 437
UFM Communication Requirements 439
UFM Server Communication with Clients 439
UFM Server Communication with UFM Web UI Client 439
UFM Server Communication with SNMP Trap Managers 440
Summary of UFM Server Communication with Clients 440
UFM Server Communication with InfiniBand Switches 440

UFM Server InfiniBand Communication with Switch 441
UFM Server Communication with Switch Management Software (Optional) 441
UFM Server Communication with Externally Managed Switches (Optional) 442
Summary of UFM Server Communication with InfiniBand Switches 442
UFM Server Communication with InfiniBand Hosts 442
UFM Server InfiniBand Communication with HCAs 443
UFM Server Communication with Host Management (Optional) 443
Summary of UFM Server Communication with InfiniBand Hosts 444
UFM Server High Availability (HA) Active—Standby Communication 444
UFM Server HA Active—Standby Communication 445
UFM Server Health Monitoring 446
UFM Health Configuration 446
UFM Core Files Tracking 450
Example of Health Configuration 451
Event Burst Management 452
Recovery from Consecutive Failures 452
Cable Transceiver Temperatures 454
GUI Views 454
Alarms 454
Event Policy 454
Table Enhancements 455

Look and Feel Improvements 455
Displayed Columns 456
Export All Data as CSV 457
Time Zone Converter 459
Multi-Subnet UFM 461
Overview 461
Setting Up Multi-Subnet UFM 461
Functionality 462
UFM Plugins 490
rest-rdma Plugin 490
Deployment Server 491
Deploy Plugin on UFM Appliance 491
Deployment Client 491
How to Run 492
Server 492
Client 492
Authentication Configuration 493
Remote ibdiagnet Request 493
Authenticated Remote ibdiagnet Request 494
Rest Request with Username/Password Authentication 495
Rest Request with Client Certificate Authentication 495

Rest Request with Token Authentication 496
NDT Plugin 496
Overview 496
Deployment 497
Authentication 497
REST API 497
Topodiff 498
Subnet Merger 498
NDT Format – Topodiff 498
Switch to Switch NDT 499
Switch to Host NDT 500
Other 501
NDT Format – Subnet Merger 502
Functionality 502
Subnet Merger Flow 503
Subnet Merger UI 506
Extending the InfiniBand Setup via Subnet Merger 513
UFM Telemetry FluentD Streaming (TFS) Plugin 514
Overview 514
Deployment 514
Authentication 515

Rest API 515
UFM Events Fluent Streaming (EFS) Plugin 515
Overview 515
Deployment 516
Authentication 516
Rest API 516
UFM Bright Cluster Integration Plugin 517
Overview 517
Deployment 517
Authentication 517
GUI Screens 518
Rest API 521
UFM Cyber-AI Plugin 522
Overview 522
Deployment 522
Autonomous Link Maintenance (ALM) Plugin 525
Overview 525
Schematic Flow 526
Deployment 528
Data Collection 529
ALM Configuration 533

ALM Jobs 534
DTS Plugin 534
Overview 534
Deployment 535
DPU Requirements 535
Installation 535
GUI Screens 537
Info 537
Health 538
Telemetry 539
Data Sources 540
GRPC-Streamer Plugin 540
Authentication 540
Create a Session to UFM from GRPC 541
Create New Subscription 542
Edit Known Subscription 543
Get List of Known Subscribers 544
Delete a Known Subscriber 544
Run a Known Subscriber Once 545
Run Streamed Data of a Known Subscriber 546
Run a New Subscriber Once 547

Run New Subscriber Streamed Data 548
Run A Serialization on All the Running Streams 549
Stop a Running Stream 549
Run a subscribe stream 549
Get the variables from a known subscriber 550
Get Help / Version 551
Sysinfo Plugin 551
Overview 551
Deployment 551
REST API 552
Sysinfo Query Format 552
SNMP Plugin 553
Deployment 553
Authentication 553
REST API 554
Usage 554
Other 557
Packet Mirroring Collector (PMC) Plugin 557
Overview 557
Deployment 558
Installation 558

GUI Screens 560
Network Event Collector Display 560
Network Event Collector Configuration 561
PDR Deterministic Plugin 562
Deployment 562
Default Configuration 563
Isolation Decisions 563
Calculating BER Counters 564
Dry Run Mode 565
GNMI-Telemetry Plugin 565
Authentication 565
Secure Server 566
Capability Request 566
Get Request 566
Subscribe Stream Request 567
Subscribe On-Change Request 568
Messages Data Format 569
Inventory Requests 569
Events Requests 570
Appendixes 572
Appendix – SM Default Files 572

Appendix – UFM Subnet Manager Default Properties 573
Configuring UFM for SR-IOV 583
Isolating Switch From Routing 584
Appendix – Enhanced Quality of Service 585
Enhanced QoS Policy File 585
Notes 586
Policy File Example 586
Appendix – Partitioning 587
Appendix – SM Activity Report 588
SM Supported Activities 588
Appendix – SM Partitions.conf File Format 590
Appendix - Supported Port Counters and Events 594
InfiniBand Port Counters 595
Supported Traps and Events 597
Appendix – Diagnostic Utilities 612
InfiniBand Diagnostics Commands 612
Diagnostic Tools 614
Dependencies 614
Utilities Descriptions 616
Appendix – Device Management Feature Support 638
Appendix – Used Ports 641

Appendix – Routing Chains 642
Configuring Routing Chains 642
Defining Port Groups 642
Defining Port Group Policy File 643
Configuring Port Group Policy 643
Port Group Qualifiers 643
Rule Qualifiers 644
Predefined Port Groups 647
Port Groups Policy Examples 647
Defining Topologies Policy File 648
Configuring Topology Policy 648
Topology Qualifiers 648
Configuration File per Routing Engine 649
Defining Routing Chain Policy File 649
First Routing Engine in Chain 650
Configuring Routing Chains Policy 650
Routing Engine Qualifiers 650
Dump Files per Routing Engine 652
Appendix – Adaptive Routing 653
Appendix – Configuration Files Auditing 653
Appendix - Managed Switches Configuration Info Persistency 655

Appendix – Security Features 656
SA Enhanced Trust Model (SAETM) 656
Set of Untrusted SA Requests Allowed 657
Proxy SA Requests 658
Registration Limits 658
SAETM Logging 658
SGID Spoofing 659
M_Key Authentication 659
M_Key Authentication Enablement 659
M_Key Per Port 660
Subnet Manager Protection 661
Appendix – NVIDIA SHARP Integration 661
NVIDIA Scalable Hierarchical Aggregation and Reduction Protocol (SHARP)™ 661
NVIDIA SHARP Aggregation Manager 662
NVIDIA SHARP AM Prerequisites 662
NVIDIA SHARP AM Configuration 663
Running NVIDIA SHARP AM in UFM 663
Operating NVIDIA SHARP AM with UFM 664
Monitoring NVIDIA SHARP AM by UFMHealth 666
Managing NVIDIA SHARP AM by UFM High Availability (HA) 666
NVIDIA SHARP AM Logs 666

NVIDIA SHARP AM Version 666
Appendix – AHX Monitoring 666
Configuring AHX Monitoring 667
AHX Monitoring Events 668
Appendix – UFM Event Forwarder 669
Appendix – UFM Migration 671
Overview 671
Backup UFM configuration 671
UFM (Bare Metal) 672
UFM Docker Container 672
UFM Appliance 672
Restore UFM Configuration 673
UFM Bare Metal 673
UFM Docker Container 674
UFM Appliance 674
Appendix - Switch Grouping 675
UI Presentation 677
Appendix – Secondary Telemetry Fields 685
Document Revision History 692
EULA, Legal Notices and 3rd Party Licenses 695
Legal Notice 695

Third-Party Licenses 695
License Agreement 695
    Exhibit A 699

About This Document

NVIDIA® UFM® for Software Defined Networks (UFM®-SDN) appliance enables data center operators to efficiently provision, monitor and operate large-scale compute and storage data center interconnect infrastructures. UFM eliminates the complexity of fabric management, while also providing deep visibility into traffic and optimizing fabric performance.

Software Download

To download the UFM software, please visit NVIDIA's Licensing Portal.

If you do not have a valid license, please fill out the NVIDIA Enterprise Account Registration form to get a UFM evaluation license.

Document Revision History

For the list of changes made to this document, refer to Document Revision History.

Release Notes

NVIDIA® UFM®-SDN Appliance is a powerful platform for managing InfiniBand scale-out computing environments. UFM enables data center operators to efficiently monitor and operate the entire fabric, boost application performance and maximize fabric resource utilization.

Key Features

UFM-SDN Appliance provides a central management console, including the following main features:

- Fabric dashboard including congestion detection and analysis
- Advanced real-time health and performance monitoring
- Fabric health reports
- Threshold-based alerts
- Fabric segmentation/isolation
- Quality of Service (QoS)
- Routing optimizations
- Central device management
- Task automation
- Logging
- High availability
- Daily report: Statistical information of the fabric during the last 24 hours
- Event management
- Switch auto-provisioning
- UFM-SDN Appliance in-service software upgrade
- Fabric validation tests
- Client certificate authentication
- IPv6 on management ports

Prior to installation, please verify that all prerequisites are met. Please refer to System Requirements.

The Logical Server Model Management feature is going to be deprecated in UFM v6.12.0.

Changes and New Features

This section lists the new and changed features in this software version.

> ⚠ For an archive of changes and features from previous releases, please refer to Changes and New Features History.

> ⚠ The items listed in the table below apply to all UFM license types.

Feature	Description
Defining Node Description	Added the option to establish a blacklist containing possible node descriptions that should be avoided when grouping Multi-NIC HCAs during host startup. For more information, refer to Defining Node Description Black-List.
Network Reports	Added the ability to view topology change events related to devices and links. For more information, refer to Events History, Device Status Events and Link Status Events.
Client Certificate Authentication	Added support for pinning SAN with RegEx.
Secondary Telemetry	Added the ability to expose SHARP telemetry in UFM Telemetry. For more information, refer to Exposing Switch Aggregation Nodes Telemetry.
	Added the ability to stop SHARP telemetry endpoint using CLI commands. For more information, refer to Stopping Telemetry Endpoint Using CLI Command.

Feature	Description
REST APIs	Enhanced the logging REST API by adding the ability to get event logs in JSON file format. For more information, refer to Get Events as JSON.
	Added the ability to expose managed switch power consumption in Web UI. For more information, refer ro Get Managed Switches Power Consumption.
	Added ability to filter the event logs by source. For more information, refer to Create Log History.
	Added the support for TBD Usage Statistics REST API
	Added the ability to generate enterprise network reports. For more information, refer to Events History, Device Status Events and Link Status Events.
	Introduced REST APIs for various authentication types. For more information, refer to Examples of REST APIs Using Various Authentication Types.
	Added the ability to update UFM Configuration REST API.
	Added the option to expose cable information. For more information, refer to Get Ports with Cable Information.
	Improved dynamic telemetry by adding the ability to instantiate a new instance and delete a running instance. For more information, refer to UFM Dynamic Telemetry Instances REST API.
	Added the option to set “down” ports as unhealthy. For more information, refer to Unhealthy Ports REST API.
	Added forge InfiniBand anti-spoofing support. For more information, refer to Forge InfiniBand Anti-Spoofing REST API.
	The "`site_name`" field can be exposed in all the REST APIs requests. To expose it, perform the following: • In `gv.cfg` and in the [Server] section, set a value for "`site_name`" • In `gv.cfg` and in the [Server] section, set "`expose_site_name`" to "`true`" • Restart UFM for the changes to take effect
Plugins	Added support for the gNMI-Telemetry plugin that employs the gNMI protocol to stream data from UFM telemetry. In addition, added support for secure mode based on client authentication. For more information, refer to the GNMI-Telemetry Plugin.
	Added support for ALM configuration for controlling isolation/de-isolation. For more information, refer to ALM Configurations.
	REST over RDMA Plugin: Moved to Ubuntu 22-based docker container, OFED 5.8-3.0.7.0, ucx_py 0.35.0 and Python 3.10.
Supported Transceivers	Added support for FR4 transceivers

⚠ The items listed in the table below apply to all UFM license types.

⚠ For bare metal installation of UFM, it is required to install MLNX_OFED 5.X (or newer) before the UFM installation.
Please make sure to use the UFM installation package that is compatible with your setup, as detailed in Bare Metal Deployment Requirements.

Unsupported Functionalities/Features

The following are the unsupported functionality/features in UFM®-SDN Appliance:

- UFM Appliance Gen1
- NVIDIA Care (MCare) Integration
- UFM on VM (UFM with remote fabric collector)
- The UFM Multi-site portal feature is no longer supported. The Multi-Subnet feature can be used instead
- Logical Elements tab - Removed as of UFM Enterprise v4.11.0

Installation Notes

Supported Devices

Supported NVIDIA Externally Managed Switches

Type	Model	Latest Tested Firmware Version
NDR switches	• MQM9790	31.2010.6102
HDR switches	• MQM8790	27.2010.6102
EDR switches	• SB7790 • SB7890	15.2010.5108
FDR switches	• SX6025 • SX6015 • SX6005	11.2000.1142

Supported NVIDIA Internally Managed Switches

Type	Model	Latest Tested OS Version
NDR switches	• MQM9700	MLNX-OS 3.11.1014
HDR switches	• MQ8700 • MCS8500 • TQ8100-HS2F	MLNX-OS 3.11.1014

Type	Model	Latest Tested OS Version
	• TQ8200-HS2F	
EDR switches	• SB7700 • SB7780 • SB7800 • CS7500 • CS7510 • CS7520	MLNX-OS 3.10.5002
FDR switches	• SX6012 • SX6018 • SX6036 • SX6506 • SX6512 • SX6518 • SX6536 • SX1012 • SX6710 • SX6720 • SX1700 • SX1710	MLNX-OS 3.8.1054
Long-haul	• MTX6100 • MTX6240 • MTX6280	MLNX-OS 3.6.8008
InfiniBand-Ethernet Gateways	SX6036G (FDR)	MLNX-OS 3.6.8008

⚠ UFM-SDN Appliance no longer supports the following QDR switches management over SNMP, which means that the internal modules of these switches, such as fans, power supply, etc., will not be discovered by UFM. In addition, health events for such modules will not be triggered by UFM:

- Grid Director™ 4036
- Grid Director™ 4036E
- Grid Director™ 4200

- Grid Director™ 4700

⚠ For supported HCAs per MLNX_OFED version, please refer to MLNX_OFED Release Notes.

System Requirements

UFM GUI Client Requirements

The platform and GUI requirements are detailed in the following tables:

Platform	Details
Browser	Edge, Internet Explorer, Firefox, Chrome, Opera, Safari
Memory	• Minimum: 8 GB • Recommended: 16 GB

MFT Package Version

Platform	Details
MFT	Integrated with MFT version mft-4.26.0-93

UFM SM Version

Platform	Type and Version
SM	UFM package includes SM version 5.17.0

UFM NVIDIA SHARP Software Version

Platform	Type and Version
NVIDIA® Scalable Hierarchical Aggregation and Reduction Protocol (SHARP)™	UFM package includes NVIDIA SHARP software version 3.5.0

Bug Fixes in This Release

Ref #	Description
3644553	**Description**: When querying the ports, adding a `cable_info=true` as an argument will give cable information per port
	Keywords: Ports, Query, `cable_info=true`
	Discovered in release: v4.13.0
3604212	**Description:** Broken UFM REST API
	Keywords: REST API
	Discovered in release: v4.12.1
3604183	**Description:** UFM error interpretation
	Keywords: Error, Interpretation

Ref #	Description
	Discovered in release: v4.13.1
3587849	**Description:** OpenSM restart when backup UFM loses power
	Keywords: OpenSM, Restart
	Discovered in release: v4.8.0
3577427	**Description:** UFM REST API returns wrong model for NDR unmanaged switch
	Keywords: Unmanaged Switch, NDR, REST API
	Discovered in release: v4.12.1
3575882	**Description:** UFM event is not generated for a switch down
	Keywords: UFM Event, Switch Down
	Discovered in release: v4.12.1
3628421	**Description:** UFM timezone issue
	Keywords: Timezone
	Discovered in release: v4.13.1

Known Issues in This Release

N/A

Ref #	Issue

For a list of known issues from previous releases, please refer to Known Issues History.

Changes and New Features History

Feature	Description
v4.13.1 (UFM Enterprise v6.14.1)	
Cable and Transceivers Burning	UFM supports second-source cable transceivers burn.
Module REST API	Added HW revision field in GET module REST API response.
Telemetry	Added support for the MRCS register read in UFM Telemetry.
UFM Reports	UFM Daily report will be disabled by default after upgrade or clean installation.
v4.12.0 (UFM Enterprise v6.13.0)	
Network Fast Recovery	Added the ability to automatically isolate a malfunctioning switch port as detected by the switch. Refer to Enabling Network Fast Recovery
Multi-Subnet UFM	Added support for multiple UFM instances, wherein multiple instances are aggregated, managed and controlled by a centralized UFM instance. Refer to Multi-Subnet UFM.
Switch ASIC Failure Detection	Added support for a new indication (UFM event) that identifies a failure of a specific switch ASIC. Refer to Configuring Partial Switch ASIC Failure Events.
Automatic Switch Grouping	Added support for enabling automatic grouping of 1U switches by UFM, as per a pre-defined user-configured mapping. Refer to Appendix - Switch Grouping.
SHARP Trees APIs	Incorporated support for a new UFM REST API that presents the current active SHARP trees. Refer to NVIDIA SHARP Resource Allocation REST API.

SHARP Reservation APIs	Added support for SHARP Reservation API enhancements. Refer to NVIDIA SHARP Resource Allocation REST API.
Email Time-Zone Settings	Added the ability to configure time-zone settings for UFM email notifications, ensuring that sent events or daily reports align with the configured time zone. Refer to Email.
Switch Connectivity Failure Indication	Incorporated support for a new UFM event indication that identifies failed communication with a specified managed switch. Refer to Supported Traps and Events.
Dynamic Telemetry	Added APIs that enable the creation and management of UFM Telemetry instances, allowing users to select desired counters and ports as per their requirements. Refer to UFM Dynamic Telemetry Instances REST API.
TFS (Telemetry Fluent Streaming) Plugin	Added support for UFM telemetry data streaming from multiple endpoints to Fluent Bit. Refer to Telemetry to Fluent Streaming (TFS) Plugin REST API.
	Added support for enabling white/black counters lists within the TFS Plugin. Refer to Telemetry to Fluent Streaming (TFS) Plugin REST API.
DTS (DPU Telemetry) Plugin	Added support for displaying DPUs data within the UFM Web UI. Refer to DTS Plugin.
Cyber-AI Plugin	Added support for displaying Cyber-AI software within the UFM Web UI. Refer to UFM Cyber-AI Plugin.
Packet Mirroring Collector (PMC) Plugin	Added the Packet Mirroring Collector (PMC) plugin that allows users to catch and collect mirrored pFRN and congestion notifications from switches for enhanced real-time network visibility. Refer to Packet Mirroring Collector (PMC) Plugin.
SNMP Traps Listener Plugin	Added the capability to enable registration and monitoring of SNMP traps from managed switches, in addition to updating UFM with the relevant trap information. Refer to SNMP Plugin.
Bright Cluster Integration Plugin	Added support for integration of data from Bright Cluster Manager (BCM) into UFM, providing a more comprehensive network perspective. Refer to UFM Bright Cluster Integration Plugin.
Expanding Non-Blocking Fabric (NDT Plugin extension)	Added a feature that facilitates seamless expansion of the IB fabric, ensuring uninterrupted functionality and optimal performance throughout the fabric. Refer to NDT Format – Merger.

PDR (Packet Drop Rate) Plugin	Added a new functionality that enables automatic detection and isolation of port failures through monitoring of PDR (Packet Drop Rate), BER (Bit Error Rate), and high cable temperatures. Refer to PDR Deterministic Plugin.
v4.11.0 (UFM Enterprise v6.12.0)	
Managed Switches - Sysinfo Mechanism	Added the ability to save switches inventory data into JSON format files and present the latest fetched switches data upon UFM start-up. The saved switches data is available UFM upon system dump. Refer to Appendix - Managed Switches Configuration Info Persistency
REST over RDMA Plugin	Introduced security improvements (allowed read-only options in remote ibdiagnet) and added support for Telemetry API. Refer to rest-rdma Plugin
Events and Notifications	Added support for indicating potential switch ASIC failure by detecting a defined percentage of unhealthy switch ports. Refer to Additional Configuration (Optional)
SHARP AM Multi-Port	Added support for detecting IB fabric interface failure and automatic failover to an alternative active port in SHARP Aggregation Manager (AM). Refer to Multi-port SM
UFM System Dump	Added support for downloading the generated UFM system dump. Refer to UFM System Dumb Tab
UFM REST API	Added support for adding or removing hosts to Partition key (PKey) assignments (when adding/removing hosts, all the related host GUIDs are assigned to/removed from the PKey). Refer to Add Host REST API
	UFM System Dump Improvements including Creating New System Dump API
Managed Switch Operations	Added support for persistent enablement/disablement of managed switches ports. Refer to Ports Window
UFM SDK	Created a script to get TopX data by category. Refer to UFM Aggregation TopX README.md file
Proxy Authentication	Added option to delegate authentication to a proxy. Refer to Delegate Authentication to a Proxy
UFM Initial Settings	Removed the requirement to set the IPoIB address to the main IB interface used by UFM/SM (gv.cfg → fabric_interface)
MLNX_OFED Package	Upgraded the MLNX_OFED version to v5.8-1.0.1
MFT Package	Integrated with MFT version 4.23.0-104
v4.10.0 (UFM Enterprise v6.11.0)	

UFM Discovery and Device Management	• InBand autosicovery of switchs' IP addresses using `ibdiagnet` • Discovering the device's PSID and FW version using `ibdiagnet` by default instead of using an SM vendor plugin
CPU Affinity	Enabling the user to control CPU affinity of UFM's major processes
gRPC API	Added support for streaming UFM REST API data over gRPC as part of new UFM plugin. Refer to GRPC-Streamer Plugin.
Telemetry	• Added support for flexible counters infrastructure (ability to change counter sets that are sampled by the UFM) • Updated the set of available counters for Telemetry (removed General counters from default view: Row BER, Effective BER and Device Temperature. Now available through the secondary telemetry instance). Refer to Secondary Telemetry
EFS UFM Plugin	Added support for streaming UFM events data to FluentD destination as part of a new UFM plugin. Refer to UFM Telemetry FluentD Streaming (TFS) Plugin
General UI Enhancements	• Displayed columns of all tables are persistent per user, with the option to restore defaults. Refer to Displayed Columns • Improved look and feel in Network Map. Refer to Network Map • Added Reveal Uptime to the general tab in the devices information tabs. Refer to Device General Tab
REST APIs	Added support for PKey filtering for default session data. Refer to Get Default Monitoring Session Data by PKey Filtering.
	Added support for filtering session data by groups. Refer to Monitoring Sessions REST API.
	Added support for resting all unhealthy ports at once. Refer to Mark All Unhealthy Ports as Healthy at Once
	Added support for presenting system uptime in UFM REST API. Refer to Systems REST API.
Deployment Installation	UFM installation is now based on Conda-4.12 (or newer) for Python v3.9 environment and third party packages deployments
MLNX_OFED Package	Upgraded the MLNX_OFED version to v5.7-1.0.2.0
Diagnostics Utilities	Added CLI commands for new diagnostic utilities
NVIDIA SHARP Software	Updated NVIDIA SHARP software version to v3.1.1.
UFM Logical Elements	UFM Logical Elements (Environments, Logical Servers, Networks) views are deprecated and will no longer be available starting from UFM v4.11.0 (January 2023 release)

v4.9.0 (UFM Enterprise v6.10.0)	
UFM Package	Integrated with UFM Enterprise v6.10.0
System health enhancements	Add support for the periodic fabric health report, and reflected the ports' results in UFM's dashboard
UFM Plugins Management	Add support for plugin management via UFM web UI
UFM Extended Status	Added REST API for exposing UFM readiness
Failover to Other Ports	Add support for SM and UFM Telemetry failover to other ports on the local machine
UFM Appliance Upgrade	Added a set of REST APIs for supporting the UFM Appliance upgrade
Configuration Audit	Add support for tracking changes made in major UFM configuration files (UFM, SM, SHARP, Telemetry)
UFM Plugins	Add support for new SDK plugins
Telemetry	Add support for statistics processing based on UFM telemetry csv format
UFM High Availability Installation	UFM high availability installation has changed and it is now based on an independent high availability package which should be deployed in addition to the UFM Enterprise standalone package. for further details about the new UFM high availability installation, please refer to - Installing UFM Server Software for High Availability
Network Map	Add support for network map topology persistency on server side
UFM Web UI	Add option to copy and paste tables content (GUIDS and LIDS) via UFM Web UI
UFM Authentication	Add support for token based authentication
SM Configuration	Setting AR (Adaptive Routing) Up Down as the default routing configuration in UFM / SM (for new UFM installations)
UFM REST API	Add Support for CloudX API in UFM for OpenStack integration and allow auto provisioning of the InfiniBand fabric

NDR Support	Add support for discovering and monitoring Nvidia NDR switches
UFM Plugins	Added a support to deploy and run UFM plugins.
MLNX_OFED Package	Upgraded the MLNX_OFED version integrated with UFM Appliance to v5.5-1.0.3

Bug Fixes History

Ref. #	Description
3590777	**Description**: After upgrading UFM new telemetry data is not being collected and presented in UI Telemetry tab.
	Keywords: Telemetry, Coredump
	Discovered in release: 4.13.0
3461058	**Description:** When using the Dynamic Telemetry API to create a new telemetry instance, the log rotation mechanism will not be applied for the newly generated logs of the UFM Telemetry instance
	Keywords: Dynamic, Telemetry, Log-rotate
	Discovered in Release: 4.12.0
3383916	**Description**: Fixed Client CTRL+C server disruption
	Keywords: Client CTRL+C, Server functionality
	Discovered in Release: Rest Over RDMA Image 1.0.0-21
3375414	**Description:** Fixed improper functionality of UFM UI Dashboard
	Keywords: UI Dashboard
	Discovered in Release: 4.10.0

Ref. #	Description
3342713	**Description:** Fixed UFM Health configuration for periodic restarts of the telemetry
	Keywords: UFM Health, Telemetry, Periodic restarts
	Discovered in Release: 4.10.0
3457714	**Description**: The Subnet Merger function in the NDT plugin is not supported in UFM Appliance.
	Keywords: UFM Subnet Merger; Plugin
	Discovered in Release: 4.11.0
3361160	Description: Fixed UFM long upgrade time due to a large historical Telemetry database file
	Keywords: Long Upgrade Time, Historical Telemetry, Database File
	Discovered in Release: 4.10.0
3268270	Description: Show managed switches inventory data (Sysinfo) immediately after UFM initialization
	Keywords: Managed Switches, Inventory, Sysinfo
	Discovered in Release: 4.10.0
3338600	Description: Fixed UFM UI lockdown by adding protection to the failed path on backend side
	Keywords: UFM UI, lockdown
	Discovered in Release: 4.10.0
3276163	Description: Fixed remote syslog configuration in UFM Web UI to be persistent
	Keywords: Remote Syslog, Web UI
	Discovered in Release: 4.10.0
3268270	Description: Show managed switches inventory data (Sysinfo) immediately after UFM initialization
	Keywords: Managed Switches, Inventory, Sysinfo
	Discovered in Release: 4.10.0
3338600	Description: Fixed UFM UI lockdown by adding protection to the failed path on backend side

Ref. #	Description
	Keywords: UFM UI, lockdown
	Discovered in Release: 4.10.0
3276163	Description: Fixed remote syslog configuration in UFM Web UI to be persistent
	Keywords: Remote Syslog, Web UI
	Discovered in Release: 4.10.0
3266866	Description: Failure to collect system dump for unmanaged switch
	Keywords: System Dump; Unmanaged Switch
	Discovered in Release: 4.10.0
3234082	Description: UFM WebUI unresponsive after failover issue
	Keywords: UFM, WebUI, failover
	Discovered in Release: 4.9.0
3199572	Description: Incorrect Tier reporting in the UFM events
	Keywords: Tier, Incorrect Report
	Discovered in Release: 4.9.0
3107006	Description: Using GET All Modules REST API (GET /ufmRest/resources/modules), returns N/A in device_name.
	Keywords: Modules, N/A, device_name
	Discovered in Release: 4.8
3076817	Description: Upgrading to the latest UFM version (UFMAPL_4.8.0.6_UFM_6.9.0.7), the UFM WEB UI shows log and error messages with "invalid date."
	Keywords: WEB UI, "invalid date"
	Discovered in Release: 4.8
3060127	Description: UFM WEB UI - Ports REST API returns tier parameters as N/A in response

Ref. #	Description
	Keywords: WEB UI, tier, N/A
	Discovered in Release: 4.8
3052660	Description: UFM monitoring mode is not working
	Keywords: Monitoring, mode
	Discovered in Release: 4.8
3031121	Description: Network map showing a link between QM8790 and Manta Ray leaf having BW of >20,000 Gb/s
	Keywords: Network Map, BW, 20,000
	Discovered in release: 4.7.0
3003366	Description: UFM Starting and Stopping On Its Own Since Merge
	Keywords: Start, Stop
	Discovered in release: 4.6.0
2968236	Description: Fabric health Old Alerts and events do not clear
	Keywords: Fabric Health, Alerts, clear
	Discovered in release: 4.7.0
2957984	Description: BER Not Being Read or Reported
	Keywords: BER, Not, Reported
	Discovered in release: 4.7.0

Ref. #	Description
3032227	Description: UFM UFMAPL_4.7.0.3_UFM_6.8.0.6 lists one of my skyways as "host" instead of "gateway"
	Keywords: skyway, gateway, host
	Discovered in release: 4.7.0
2966472	Description: UFM Fabric health BER_CHECK warnings
	Keywords: Fabric Health, BER, check
	Discovered in release: 4.7.0
2968236	Description: All ibdiagnet runs fail if multi-port SM is enabled
	Keywords: Multi-port SM, Fabric Health report, ibdiagnet
	Fixed in release: 4.7.1
2972007	Description: UFM cannot start if peer machine is not reachable
	Keywords: High availability, UFM start
	Fixed in release: 4.7.1
2957984	Description: High BER ports are not being read or reported.
	Keywords: High BER ports
	Fixed in Release: 4.7.1
-	Description: UFM does not automatically isolate high Effective BER ports.
	Keywords: Unhealthy ports, auto-isolated high BER ports
	Fixed in release: 4.7.0

Ref. #	Description
2705974	Description: If new ports are added after UFM startup, the default session REST API (GET /ufmRest/monitoring/session/0/data) will not include port statistics for the newly added ports.
	Keywords: Default session; REST API; missing ports
	Fixed in release: 4.7.0
2804807	Description: UFM WEB GUI becomes unresponsive and event/REST API log stops printing.
	Keywords: Web UI, unresponsive
	Fixed in release: 4.7.0
2801258	Description: UFM failed to serve incoming REST API requests.
	Keywords: REST API, hang, unresponsive
	Fixed in release: 4.7.0
2744757	Description: CheckSMConnectivityOnStandby should consider multiple GUIDs on a port.
	Keywords: UFM Health, Multi-port SM
	Fixed in release: 4.6.0
2589617	Description: UFM stopped to discover topology on SuperPOD environment.
	Keywords: Stopped; discover
	Fixed in release: 4.6.0
2615514	Description: UFM software no longer supports license type "UFM APPLIANCE".
	Keywords: License; UFM Appliance
	Fixed in release: 4.6.0
2677501	Description: UFM triggers "Cooling Device Monitoring Error" event when AHX monitoring is configured.
	Keywords: AHX monitoring
	Fixed in release: 4.6.0

Ref. #	Description
2790127	Description: UFM Fabric Health report aborted with a timeout.
	Keywords: Fabric Health report, ibdiagnet
	Fixed in release: 4.6.0
2789146	The CLI command "ib topo-file generate" not working.
	Fabric Topology file
	Fixed in release: 4.6.0
2782069	Description: (None) values are reported in "High BER port" UFM events.
	Keywords: High BER Port
	Fixed in release: 4.6.0
2482401	Description: Telemetry is not reporting multi-HCA ports due to naming convention
	Keywords: Telemetry: multi-HCA
	Fixed in release 4.5.1
2638032	Description: Wrong label in "Effective BER" event.
	Keywords: Effective BER; events
	Fixed in release 4.5.1
2288773	Description: NVIDIA-SuperPOD UFM 6.5.0-4 incorrectly reports HCA names in GUI
	Keywords: WebUI
	Discovered in Release: 4.3.0
	Fixed in Release: 4.4.0
1967348	Description: Email sender address cannot contain more than one period (“.”) in the domain name.
	Keywords: Email, sender, period
	Discovered in Release: 4.1.5

Ref. #	Description
	Fixed in Release: 4.2
2069425	Description: SMTP server username cannot have more than 20 characters.
	Keywords: Email
	Discovered in Release: 4.1.5
	Fixed in Release: 4.2
2197199	Description: Auto-provisioning cannot communicate with Mellanox switch over HTTPS.
	Keywords: Auto-provisioning
	Discovered in Release: 4.1.5
	Fixed in Release: 4.2
2192488	Description: FW upgrade as part of UFM Appliance upgrade causes to QP reset fail.
	Keywords: UFM Appliance upgrade
	Discovered in Release: 4.1.5
	Fixed in Release: 4.2
1783048	Description: Opening UFM web UI in monitoring mode is now supported.
	Keywords: Web UI, monitoring mode

Ref. #	Description
	Discovered in Release: 4.1.5
	Fixed in Release: 4.2
1691882	Description: UFM Agent now is now part of the UFM web UI.
	Keywords: UFM Agent
	Discovered in Release: 4.1.5
	Fixed in Release: 4.2
1793244	Description: UFM/module temperature thresholds notifications.
	Keywords: Temperature module
	Discovered in Release: 4.1.5
	Fixed in Release: 4.2
1358248	Description: Fixed the issue where ibdiagnet’s unresponsiveness when using the get_physical_info flag caused UFM to hang.
	Keywords: ibdiagnet
	Discovered in Release: 3.7.0
	Fixed in Release: 4.1.5
1706226	Description: Fixed an issue where MLNX_OS credentials were missing at the device "access_credentials" menu (the issue was detected on old Java based GUI). At the new UFM Web UI – MLNX_OS credentials are represented by HTTP credentials.

Ref. #	Description
	Keywords: MLNX_OS, credentials
	Discovered in Release: 4.0.0
	Fixed in Release: 4.1.5
1914379	Description: MellanoxCare service can now communicate with UFM (valid only when http communication is configured between MCare and UFM).
	Keywords: MellanoxCare, http, https
	Discovered in Release: 4.0.0
	Fixed in Release: 4.1.5
1779694	Description: Fixed the issue where UFM REST API "get_all_pkeys" broke backward compatibility.
	Keywords: UFM REST API, PKey Management
	Discovered in Release: 4.0.0
	Fixed in Release: 4.0.1
1450154	Description: Fixed the issue where partition configuration was lost after UFM restart.
	Keywords: partitions.conf, PKey, configuration
	Discovered in Release: 3.5.0
	Fixed in Release: 3.7.2

Ref. #	Description
1498857	Description: Fixed the issue where the SM used to crash during tear down when SM port GUID is 0.
	Keywords: Subnet Manager
	Discovered in Release: 3.5.0
	Fixed in Release: 3.7.2
1422797	Description: Added missing width_enabled and width_supported attributes to the ports REST APIs.
	Keywords: REST
	Discovered in Release: 3.5.0
	Fixed in Release: 3.7.2
1294010	Description: Fixed the issue where partition configuration was lost after upgrading to UFM version 5.9.6 and restarting the server.
	Keywords: partitions.conf, PKey, configuration
	Discovered in Release: 3.6.0
	Fixed in Release: 3.7.0
1276539	Description: Updated report execution command in order to avoid the following false warning of wrong link speed during topology comparison: "Wrong link speed on:Sf4521403000f0c10/Nf4521403000f0c10/P1. Expected:FDR10 got:10"
	Keywords: Topology compare report
	Discovered in Release: 3.6.0

Ref. #	Description
	Fixed in Release: 3.7.0
1131286	Description: Fixed a memory leak of UFM's main process when running multiple reports periodically.
	Keywords: Memory leak, reports
	Discovered in Release: 3.5
	Fixed in Release: 3.6.1
1064349	Description: Fixed the issue where UFM health reported false unresponsiveness of OpenSM.
	Keywords: UFM health, OpenSM
	Discovered in Release: 3.0
	Fixed in Release: 3.5
920493	Description: Fixed the issue where UFM health used to crash when restarting history mysql daemon.
	Keywords: UFM health, history mysql
	Discovered in Release: 2.5
	Fixed in Release: 3.0
958428	Description: Fixed the issue where UFM fabric health report used to crash when "Cable type check & cable diagnostics" was checked in settings.
	Keywords: UFM fabric health reports

Ref. #	Description
	Discovered in Release: 2.5
	Fixed in Release: 3.0
798876	Description: Fixed the issue where there was no indication of removing an electric cable out of a power supply module.
	Keywords: UFM events, device management
	Discovered in Release: 2.5
	Fixed in Release: 3.0
-	Description: Fixed the issue of when one of the ports of ConnectX-4 card was removed and the other remained connected, a switch module removal event was created.
	Keywords: UFM events, Device Management
	Discovered in Release: 2.5
	Fixed in Release: 3.0
-	Description: Fixed the issue where UFM CPU usage test was greater than the threshold (100%).
	Keywords: UFM CPU
	Discovered in Release: 2.5
	Fixed in Release: 3.0
688875	Description: Fixed the invalid format (Grid [Grid [Grid]]) of the Event Logs.

Ref. #	Description
	Keywords: UFM events log
	Discovered in Release: 2.5
	Fixed in Release: 3.0
-	Description: Fixed the issue where UFM GUI launch used to fail due to Java security block.
	Keywords: UFM GUI
	Discovered in Release: 2.5
	Fixed in Release: 3.0
-	Description: Fixed the issue where Mellanox Care failed to complete scan (IB Simulator failure).
	Keywords: Mellanox Care
	Discovered in Release: 2.5
	Fixed in Release: 3.0
766828	Description: Fixed the issue of a false "SM LID Changed" event was created in monitoring mode.
	Keywords: UFM events, Monitoring mode
	Discovered in Release: 2.5
	Fixed in Release: 3.0

Ref. #	Description
690862	Description: Fixed the issue where ConnectX-4 grouping on UFM was corrupted when using the Unhealthy Ports feature.
	Keywords: Grouping, unhealthy, ports
	Discovered in Release: 2.5
	Fixed in Release: 3.0

Known Issues History

Ref. #	Issue
N/A	**Description:** Execution of UFM Fabric Health Report (via UFM Web UI / REST API) will trigger ibdiagnet to use SLRG register which might cause some of the Switch and HCA's firmware to stuck and cause the HCA's ports to stay at "Init" state.
	Keywords:
	Discovered in Release: 6.14.0
3538640	**Description:** Fixed ALM plugin log rotate function

Ref. #	Issue
	Keywords: ALM, Plugin, Log rotate
	Discovered in Release: 4.12.1
3532191	**Description**: Fixed UFM hanging (database is locked) after corrective restart of UFM health.
	Keywords: Hanging, Database, Locked
	Discovered in Release: 4.12.1
3555583	**Description:** Resolved REST API links inability to return hostname for computer nodes
	Keywords: REST API, Links, Hostname, Computer Nodes
	Discovered in Release: 4.11.1
3547517	**Description:** Fixed UFM logs REST API returning empty result when SM logs exist on the disk
	Keywords: Logs, SM logs, Empty
	Discovered in Release: 4.10.0
3546178	**Description:** Fixed SHARP jobs failure when SHARP reservation feature is enabled
	Keywords: SHARP, Jobs, Reservation
	Discovered in Release: 4.12.1
3541477	**Description:** Fixed UFM module temperature alerting on wrong thresholds

Ref. #	Issue
	Keywords: Module Temperature, Alert Threshold
	Discovered in Release: 4.12.1
3191419	**Description:** Fixed UFM default session API returning port counter values as NULL
	Keywords: Null, Port Counter, Value, API
	Discovered in Release: 4.8.0
3560659	**Description**: Fixed proper update in [MngNetwork] mtu_limit in gv.cfg when restarting UFM.
	Keywords: mtu_limit, gv.cfg, Update, UFM restart
	Discovered in Release: 4.12.1
3496853	**Description:** Fixed daily report not being sent properly.
	Keywords: Daily Report, Failure
	Discovered in Release: 4.12.1
3469639	**Description**: Fixed REST RDMA server failure every couple of days, causing inability to retrieve ibdiagnet data.
	Keywords: REST RDMA, ibdiagnet
	Discovered in Release: 4.12.1
3455767	**Description:** Fixed incorrect combination of multiple devices in monitoring.
	Keywords: Monitoring, Incorrect combination
	Discovered in Release: 4.12.1
3511410	**Description**: Collect system dump for DGX host does not work due to missing sshpass utility.
	Workaround: Install sshpass utility on the DGX .

Ref. #	Issue
	Keywords: System Dump, DGX, sshpass utility
3432385	**Description**: UFM does not support HDR switch configured with hybrid split mode, where some of the ports are split and some are not.
	Workaround: UFM can properly operate when **all** or **none** of the HDR switch ports are configured as split.
	Keywords: HDR Switch, Ports, Hybrid Split Mode
3461658	**Description**: After the upgrade from UFM-SDN Appliance v4.12.0 GA to UFM-SDN Appliance v4.12.1 FUR, the network fast recovery path in `opensm.conf` is not automatically updated and remains with a null value (`fast_recovery_conf_file (null)`)
	Workaround: If you wish to enable the network fast recovery feature in UFM, make sure to set the appropriate path for the current fast recovery configuration file (`/opt/ufm/files/conf/opensm/fast_recovery.conf`) in the opensm.conf file located at `conf/opensm`, before starting UFM.
	Keywords: Network fast recovery, Missing, Configuration
3361160	**Description**: Upgrading UFM-SDN Appliance from versions 4.7.0, 4.8.0 and 4.9.0 results in cleanup of UFM-SDN Appliance historical telemetry database (due to schema change). This means that the new telemetry data will be stored based on the new schema.
	Workaround: To preserve the historical telemetry database data while upgrading from UFM-SDN appliance version 4.7.0, 4.8.0 and 4.9.0, perform the upgrade in two phases. First, upgrade to UFM-SDN v4.10.0, and then upgrade to the latest UFM version (UFM-SDN appliance 4.11.0 or newer). It is important to note that the upgrade process may take longer depending on the size of the historical telemetry database.
	Keywords: UFM Historical Telemetry Database, Cleanup, Upgrade
3346321	**Description:** In some cases, when multiport SM is configured in UFM, a failover to the secondary node might be triggered instead of failover to the local available port
	Workaround: N/A
	Keywords: Multiport SM, Failover, Secondary port

Ref. #	Issue
N/A	**Description**: Enabling a port for a managed switch fails in case that port is not disabled in a persistent way (this may occur in ports that were disabled on a previous versions of UFM - prior to UFM v4.11.0)
	Workaround: Set "persistent_port_operation=false" in `gv.cfg` to use non-persistent (legacy) disabling or enabling of port. UFM restart is required.
	Keywords: Disable, Enable, Port, Persistent
N/A	**Description**: Enabling a port for a managed switch fails in case that port is not disabled in a persistent way (this may occur in ports that were disabled on a previous versions of UFM - prior to UFM v4.11.0)
	Workaround: Set "persistent_port_operation=false" in `gv.cfg` to use non-persistent (legacy) disabling or enabling of port. UFM restart is required.
	Keywords: Disable, Enable, Port, Persistent
N/A	**Description**: Running UFM software with external UFM-SM is no longer supported
	Workaround: N/A
	Keywords: External UFM-SM
N/A	**Description:** If using SM mkey per port, several UFM operations might fail (get cable info, get system dump, switch FW upgrade).
	Workaround: N/A
	Keywords: SM, mkey per port
N/A	**Description:** If using SM mkey per port, several UFM operations might fail (get cable info, get system dump, switch FW upgrade).
	Workaround: N/A
	Keywords: SM, mkey per port
	Discovered in release: 4.7.1

Ref. #	Issue
2796317	**Description:** SHARP jobs may hang when running in reservations mode (i.e. SHARP allocation is enabled), and reservation is created with limited PKEY
	Workaround: The PKEY used for creating the reservation should be "full" (the most significant bit should be on e.g. 0x805c)
	Keywords: SHARP AM, allocation, reservation, PKEY
	Discovered in release: 4.6.0
-	**Description:** Auto-isolated high Symbol BER ports reported by UFM as unhealthy are not automatically set back as healthy when high symbol BER is elapsed.
	Workaround: Manually set auto-isolated ports as healthy.
	Keywords: Unhealthy ports, auto-isolated high BER ports
	Discovered in release: 4.6.0
2694977	**Description:** Adding a large number of devices (~1000) to a group or a logical server, on large scale setup takes ~2 minutes.
	Workaround: N/A
	Keywords: Add device; group; logical server; large scale
	Discovered in release: 4.6.0
2710613	**Description:** Periodic topology compare will not report removed nodes if the last topology change included only removed nodes.
	Workaround: N/A
	Keywords: Topology comparison
	Discovered in release: 4.6.0
2698055	**Description:** UFM, configured to work with telemetry for collecting historical data, is limited to work only with the configured HCA port. If this port is part of a bond interface and a failure occurs on the port, collection of telemetry data via this port stops.
	Workaround: Reconfigure telemetry with the new active port and restart it within UFM.
	Keywords: Telemetry; history; bond; failure

Ref. #	Issue
	Discovered in release: 4.6.0
2304264	**Description:** The option to collect system dump is only supported for hosts containing the CURL utility which supports the scp and sftp protocols.
	Workaround: To check the protocols supported by CURL, run: `curl -V` If scp and sftp are not supported, install a CURL version that supports these protocols.
	Keywords: System dump, host, CURL
	Discovered in release: 4.5.0
2480430	**Description:** Mellanox SHARP AM does not run with `smx_sock_port` value less than 1024 or greater than 49151.
	Workaround: N/A
	Keywords: Mellanox SHARP; aggregation manager
	Discovered in release: 4.5.0
2288038	**Description:** When the user try to collect system dump for UFM Appliance host, the job will be completed with an error with the following summary: "Running as a none root user Please switch to root user (super user) and run again."
	Workaround: N/A
	Keywords: System dump, UFM Appliance host
	Discovered in release: 4.4.0
2384211	**Description:** MLNX-OS version 3.9.2002 does not support SHARP allocation.
	Walkaround: Downgrade the switch to MLNX-OS version 3.9.1906.
	Keywords: SHARP allocation, MLNX-OS
	Discovered in release: 4.4.0

Ref. #	Issue
2366031	**Description:** When upgrading a switch with MLNX-OS version 3.9.1932 and later, you must make sure to comply with the new password requirements for admin and monitor users. • Password must contain 8-64 characters • Password must be different than username • Password must be different than 5 previous passwords • Password must contain at least one of each of the following: Lowercase, uppercase and digits
	Workaround: N/A
	Keywords: User password, switch OS
	Discovered in release: 4.4.0
2100564	**Description:** For modular dual-management switch systems, switch information is not presented correctly if the primary management module fails and the secondary takes over.
	Workaround: To avoid corrupted switch information, it is recommended to manually set the virtual IP address (box IP address) for the switch as the managed switch IP address (manual IP address) within UFM.
	Keywords: Modular switch, dual-management, virtual IP, box IP
	Discovered in release: 4.3.0
2135272	**Description:** UFM does not support hosts equipped with multiple HCAs of different types (e.g. a host with ConnectX®-3 and ConnectX-4/5/6) if multi-NIC grouping is enabled (i.e. multinic_host_enabled = true).
	Workaround: All managed hosts must contain HCAs of the same type (either using ConnectX-3 HCAs or use ConnectX-4/5/6 HCAs).
	Keywords: Multiple HCAs

Ref. #	Issue
	Discovered in release: 4.3.0
2063266	**Description:** Firmware upgrade for managed hosts with multiple HCAs is not supported. That is, it is not possible to perform FW upgrade for a specific host HCA.
	Workaround: Running software (MLNX_OFED) upgrade on that host will automatically upgrade all the HCAs on this host with the firmware bundled as part of this software package.
	Keywords: FW upgrade, multiple HCAs
	Discovered in release: 4.3.0
-	**Description:** When upgrading from software version 4.1.x or older to 4.2.x or later, in an intermediary step, where the standby UFM appliance is upgraded and the master UFM appliance is not, some CLI commands from the master will not be operational. This happens because, in that moment, the standby appliance is in a higher SSH security level.
	Workaround: After upgrading the master appliance to the latest version, CLI commands resume operations normally.
	Keywords: Upgrade, high availability, SSH
	Discovered in release: 4.2.0
2130688	**Description:** Registering an external SM system with two different IP addresses is not supported.
	Workaround: Before registering an external SM system with a new IP address, it is required to unregister the old IP address for that system.
	Keywords: External SM

Ref. #	Issue
	Discovered in release: 4.2.0
1895385	**Description:** QoS parameters (mtu, sl and rate_limit) change does not take effect unless OpenSM is restarted.
	Workaround: N/A
	Keywords: QoS, PKey, OpenSM
	Discovered in release: 4.2.0
-	**Description:** Management PKey configuration (e.g. MTU, SL) can be performed only using PKey management interface (via GUI or REST API).
	Workaround: N/A
	Keywords: PKey, Management PKey, REST API
	Discovered in release: 4.2.0
-	**Description:** The hostname or/and IP address of UFM HA server is used in the HA configuration.
	Workaround: Do not change hostname or IP address of UFM HA server unless you wish to reconfigure the HA mechanism.
	Keywords: High availability
	Discovered in release: 4.2.0

Ref. #	Issue
-	**Description:** UFM-SDN Appliance supports limited number of login sessions. When the limit is reached, any available client application (GUI, Multisite Portal or SDK) will not be able to connect to UFM during the login session timeout (default timeout is 10 minutes).
	Workaround: When using SDK, do not exceed 5 logins per minute.
	Keywords: UFM Server
-	**Description:** Configuration from lossy to lossless requires device reset.
	Workaround: Reboot all relevant devices after changing behavior from lossy to lossless.
	Keywords: Lossy configuration

Overview

Scale-Out Your Fabric with Unified Fabric Manager

NVIDIA's Unified Fabric Manager (UFM®) is a powerful platform for managing scale-out computing environments. UFM enables data center operators to efficiently monitor and operate the entire fabric, boost application performance and maximize fabric resource utilization.

While other tools are device-oriented and involve manual processes, UFM's automated and application-centric approach bridges the gap between servers, applications and fabric elements, thus enabling administrators to manage and optimize from the smallest to the largest and most performance-demanding clusters.

UFM Benefits

Central Console for Fabric Management

UFM provides all fabric management functions in one central console.

The ability to monitor, troubleshoot, configure and optimize all fabric aspects is available via one interface. UFM's central dashboard provides a one-view fabric-wide status view.

In-Depth Fabric Visibility and Control

UFM includes an advanced granular monitoring engine that provides real-time access to switch and host data, enabling cluster-wide monitoring of fabric health and performance, real-time identification of fabric-related errors and failures, quick problem resolution via granular threshold-based alerts and a fabric utilization dashboard.

Advanced Traffic Analysis

Fabric congestion is difficult to detect when using traditional management tools, resulting in unnoticed congestion and fabric under-utilization. UFM's unique traffic map quickly identifies traffic trends, traffic bottlenecks, and congestion events spreading over the fabric, which enables the administrator to identify and resolve problems promptly and accurately.

Enables Multiple Isolated Application Environments on a Shared Fabric

Consolidating multiple clusters into a single environment with multi-tenant data centers and heterogeneous application landscapes requires specific policies for the different parts of the fabric. UFM enables segmentation of the fabric into isolated partitions, increasing traffic security and application performance.

Service-Oriented Automatic Resource Provisioning

UFM uses a logical fabric model to manage the fabric as a set of business-related entities, such as time critical applications or services. The logical fabric model enables fabric monitoring and performance optimization on the application level rather than just at the individual port or device level. Managing the fabric using the logical fabric model provides improved visibility into fabric performance and potential bottlenecks, improved performance due to application-centric optimizations, quicker troubleshooting and higher fabric utilization.

Quick Resolution of Fabric Problems

UFM provides comprehensive information from switches and hosts, showing errors and traffic issues such as congestion. The information is presented in a concise manner over a unified dashboard and configurable monitoring sessions. The monitored data can be correlated per job and customer, and threshold-based alarms can be set.

Seamless Failover Handling

Failovers are handled seamlessly and are transparent to both the user and the applications running on the fabric, significantly lowering downtime. The seamless failover makes UFM in conjunction with other Mellanox products, a robust, production-ready solution for the most demanding data center environments.

Open Architecture

UFM provides an advanced Web Service interface and CLI that integrate with external management tools. The combination enables data center administrators to consolidate management dashboards while flawlessly sharing information among the various management applications, synchronizing overall resource scheduling, and simplifying provisioning and administration.

Main Functionality Modules

Fabric Dashboard

UFM's central dashboard provides a one-view fabric-wide status view. The dashboard shows fabric utilization status, performance metrics, fabric-wide events, and fabric health alerts.

The dashboard enables you to efficiently monitor the fabric from a single screen and serves as a starting point for event or metric exploration.

Fabric Design and Configuration

In the Design view you can define the fabric logical model on which UFM bases configuration of the fabric. The correlation of all UFM functionality to the logical model enables simple and quick fabric configuration, easy change management and business oriented analysis as follows:

- Monitored data and fabric events are correlated to the logical model.
- Fabric and host configuration is performed according to the logical model.

- Performance optimization is derived from the logical model.

Fabric Discovery and Physical View

UFM discovers the devices on the fabric and populates the views with the discovered entities. In the physical view of the fabric, you can view the physical fabric topology, model the data center floor, and manage all the physical-oriented events.

Central Device Management

UFM provides the ability to centrally access switches and hosts, and perform maintenance tasks such as firmware and software upgrade, shutdown and restart.

Monitoring

UFM includes an advanced granular monitoring engine that provides real time access to switch and server data. Fabric and device health, traffic information and fabric utilization are collected, aggregated and turned into meaningful information.

Configuration

In-depth fabric configuration can be performed from the Settings view, such as routing algorithm selection and access credentials.

The Event Policy Table, one of the major components of the Configuration view, enables you to define threshold-based alerts on a variety of counters and fabric events. The fabric administrator or recipient of the alerts can quickly identify potential errors and failures, and actively act to solve them.

Logging

The Logging view enables you to view detailed logs and alarms that are filtered and sorted by category, providing visibility into traffic and device events as well as into UFM server activity history.

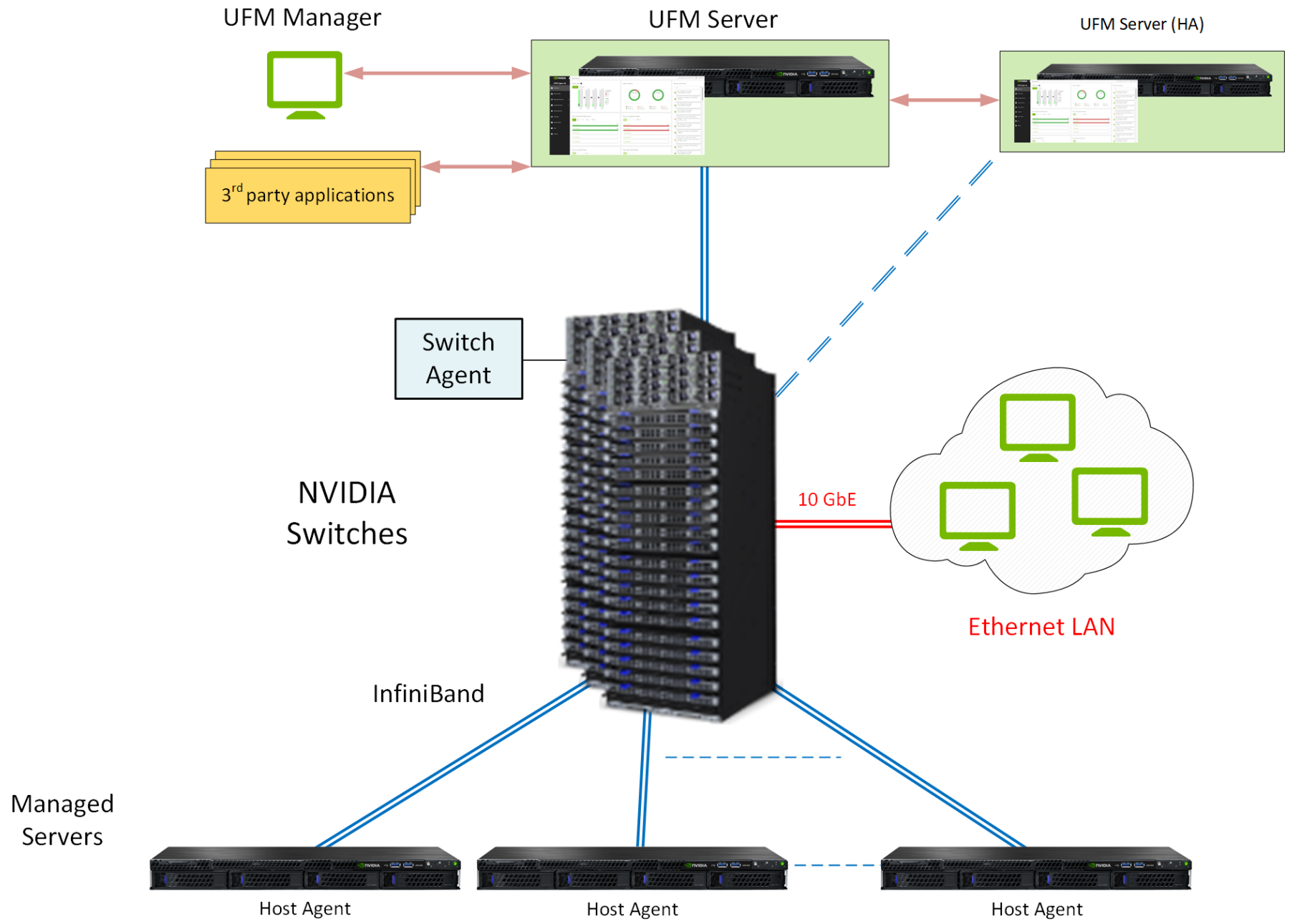
High Availability

In the event of a failover, when the primary (active) UFM®-SDN Appliance goes down or is disconnected from the fabric, UFM's High Availability (HA) capability allows for a secondary (standby) UFM®-SDN Appliance to immediately and seamlessly take over fabric management tasks. Failovers are handled seamlessly and are transparent to both the user and the applications running in the fabric. UFM’s High Availability capability, when combined with Mellanox's High Availability switching solutions allows for non-disruptive operation of complex and demanding data center environments.

InfiniBand Fabric Managed by UFM

NVIDIA® UFM®-SDN Appliance provides all management functionalities required for managing fabrics.

Fabric Topology with UFM-SDN Appliance


UFM Manager
UFM Server
UFM Server (HA)
3rd party applications
Switch Agent
NVIDIA Switches
10 GbE
Ethernet LAN
InfiniBand
Managed Servers
Host Agent
Host Agent
Host Agent

UFM-SDN Appliance comes with pre-installed UFM and has complete visibility over the fabric to manage routing on all devices.

Managed Switching Devices are fabric switches and gateways that UFM manages.

Managed Servers are the compute nodes in the fabric on which the various applications are running, and UFM manages all servers connected to the fabric.

UFM Host Agent is an optional component that can be installed on the Managed Servers. UFM Host Agent provides local host data and host device management functionality.

The UFM Host Agent provides the following functionality:

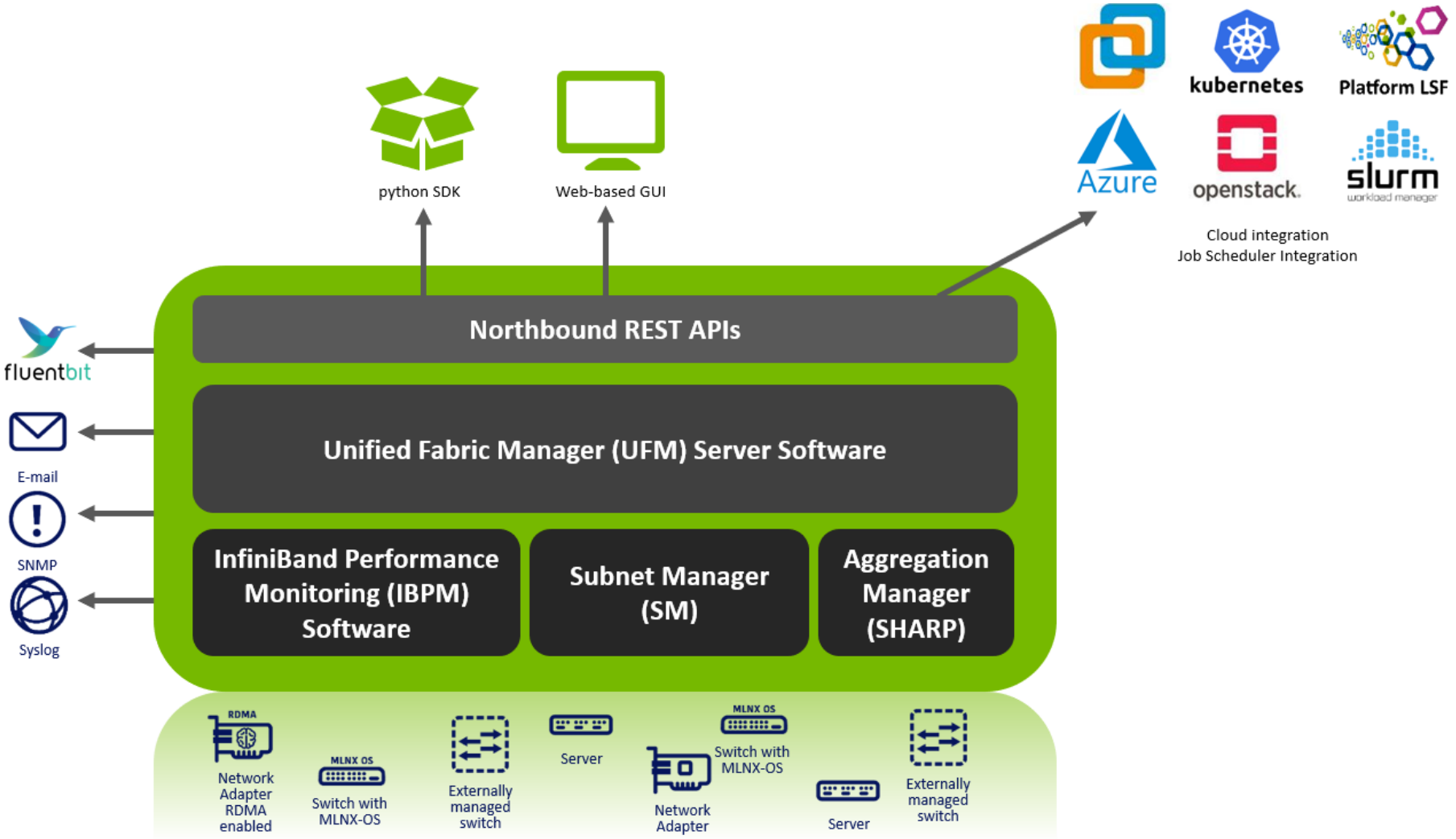
- Discovery of IP address, CPU, and memory parameters on host
- Collection of CPU/Memory/Disk performance statistics on host
- Upgrading HCA Firmware and OFED remotely
- Creating an IP interface on top of the InfiniBand partition

UFM Switch Agent is an embedded component in NVIDIA switches that allows IP address discovery on the switch and allows UFM to communicate with the switch.

UFM Software Architecture

The following figure shows the UFM high-level software architecture with the main software components and protocols. Only the main logical functional blocks are displayed and do not necessarily correspond to system processes and threads.

UFM High-Level Software Architecture


python SDK
Web-based GUI
kubernetes
Platform LSF
Azure
openstack
slurm
workload manager
Cloud integration
Job Scheduler Integration
Northbound REST APIs
fluentbit
E-mail
SNMP
Syslog
Unified Fabric Manager (UFM) Server Software
InfiniBand Performance Monitoring (IBPM) Software
Subnet Manager (SM)
Aggregation Manager (SHARP)
RDMA
Network Adapter RDMA enabled
MLNX OS
Switch with MLNX-OS
Externally managed switch
Server
Network Adapter
MLNX OS
Switch with MLNX-OS
Server
Externally managed switch

Graphical User Interface

UFM User Interface is a web application based on JavaScript and Angular JS, which is supported by any Web Browser. The Web application uses a standard REST API provided by the UFM server.

Client Tier API

Third-party clients are managed by the REST API.

Client Tier SDK Tools

Support for UFM's API and a set of tools that enhance UFM functionality and interoperability with third-party applications are provided as part of UFM.

UFM Server

UFM server is a central data repository and management server that manages all physical and logical data. UFM-SDN Appliance receives all data from the Device and Network tiers and invokes Device and Network tier components for management and configuration tasks. UFM-SDN Appliance uses a database for data persistency. The UFM-SDN Appliance is built on the Python twisted framework.

Subnet Manager

Subnet Manager (SM) is the InfiniBand "Routing Engine", a key component used for fabric bring-up and routing management.
UFM uses the Open Fabric community OpenSM Subnet Manager. UFM uses a plug-in API for runtime management and fabric data export.

NVIDIA Scalable Hierarchical Aggregation and Reduction Protocol (SHARP)™ Aggregation Manager

NVIDIA Scalable Hierarchical Aggregation and Reduction Protocol (SHARP) is a technology that improves the performance of mathematical and machine learning applications by offloading collective operations from the CPU to the switch network.

Aggregation Manager (AM) is a key component of NVIDIA SHARP software, used for NVIDIA SHARP resources management.

For further information about NVIDIA SHARP AM, refer to Appendix - NVIDIA SHARP Integration.

Performance Manager

The UFM Performance Manager component collects performance data from the managed fabric devices and sends the data to the UFM-SDN Appliance for fabric-wide analysis and display of the data.

Device Manager

The Device Manager implements the set of common device management tasks on various devices with varying management interfaces. The Device Manager uses SSH protocol and operates native device CLI (command-line interface) commands.

UFM Switch Agent

UFM Switch Agent is an integrated part of NVIDIA switch software. The agent supports system parameter discovery and device management functionality on switches.

Communication Protocols

UFM uses the following communication protocols:

- Web UI communicates with the UFM server utilizing **Web Services** carried on **REST API**.
- The UFM server communicates with the switch Agent located on managed switches by proprietary **TCP/UDP**-based discovery and monitoring protocol and **SSH**.
- Monitoring data is sent by the switch Agent to UFM server Listener by a proprietary **TCP**-based protocol.

UFM Installation and Initial Configuration

UFM® software includes Server and Agent components. UFM Server software should be installed on a central management node. For optimal performance, and to minimize interference with other applications, it is recommended to use a dedicated server for UFM. The UFM Agent is an optional component and should be installed on fabric nodes. The UFM Agent should not be installed on the Management server.

The following sections provide step-by-step instructions for installing and activating the license file, installing the UFM server software, and installing the UFM Agent.

- Additional Configuration (Optional)
- Installing License
- Historical Telemetry Collection in UFM

Additional Configuration (Optional)

General Settings in gv.cfg

Configure general settings in the `conf/gv.cfg` file.

> ⚠ When running UFM in HA mode, the gv.cfg file is replicated to the standby server.

Enabling SHARP Aggregation Manager

SHARP Aggregation Manager is disabled by default. To enable it, set:

```
[Sharp]
sharp_enabled = true
```

⚠ Upon startup of UFM or SHARP Aggregation Manager, UFM will resend all existing tenant allocations to SHARP AM.

Running UFM in Monitoring Mode

```
monitoring_mode = yes
```

⚠ For more information, see Running the UFM Software in Monitoring Mode.

Enabling Predefined Groups

```
enable_predefined_groups = true
```

⚠ By default, pre-defined groups are enabled. In very large-scale fabrics, pre-defined groups can be disabled in order to allow faster startup of UFM.

Enabling Multi-NIC Host Grouping

```
multinic_host_enabled = true
```

⚠ Upon first installation of UFM 6.4.1 and above, multi-NIC host grouping is enabled by default. However, if a user is upgrading from an older version, then this feature will be disabled for them.

⚠ It is recommended to set the value of this parameter before running UFM for the first time.

Defining Node Description Black-List

> ⚠ Node descriptions from the black-list should not be used for Multi-NIC grouping.

During the process of host reboot or initialization/bringup, the majority of HCAs receive a default label rather than an actual, real description. To prevent the formation of incorrect multi-NIC groups based on these default labels, this feature offers the option to establish a blacklist containing possible node descriptions that should be avoided when grouping Multi-NIC HCAs during host startup. Once a legitimate node description is assigned to the host, the HCAs are organized into multi-NIC hosts based on their respective descriptions. It is recommended to configure this parameter before initiating the UFM for the first time.

For instance, nodes initially identified with descriptions listed in the `exclude_multinic_desc` will not be initially included in Multi-NIC host groups until they obtain an updated, genuine node description.

Modify the `exclude_multinic_desc` parameter in the cv.fg file:

```
exclude_multinic_desc = localhost,generic_name_1,generic_name_2
```

Running UFM Over IPv6 Network Protocol

The default multicast address is configured to an IPv4 address. To run over IPv6, this must be changed to the following in section UFMAgent of gv.cfg.

```
[UFMAgent]
...
# if ufmagent works in ipv6 please set this multicast address to FF05:0:0:0:0:0:0:15F
mcast_addr = FF05:0:0:0:0:0:0:15F
```

Adding SM Plugin (e.g. lossymgr) to event_plugin_name Option

The following options allow users to set the SM plugin options via the UFM configuration. Once SM is started by UFM, it will start the SM plugin with the specified options.

```
# Event plugin name(s)
event_plugin_name osmufmpi lossymgr
```

Add the plug-in options file to the `event_plugin_options` option:

```
# Options string that would be passed to the plugin(s)
event_plugin_options --lossy_mgr -f <lossy-mgr-options-file-name>
```

These plug-in parameters are copied to the opensm.conf file in Management mode only.

Multi-port SM

SM can use up to eight-port interfaces for fabric configuration. These interfaces can be provided via /opt/ufm/conf/gv.cfg. The users can specify multiple IPoIB interfaces or bond interfaces in /opt/ufm/conf/gv.cfg, subsequently, the UFM translates them to GUIDs and adds them to the SM configuration file (/opt/ufm/conf/opensm/opensm.conf). If users specify more than eight interfaces, the extra interfaces are ignored.

```
[Server]

# disabled (default) | enabled (configure opensm with multiple GUIDs) | ha_enabled (configure multiport SM with
high availability)
 multi_port_sm = disabled
 # When enabling multi_port_sm, specify here the additional fabric interfaces for OpenSM conf
 # Example: ib1,ib2,ib5 (OpenSM will support the first 8 GUIDs where first GUID will
 # be extracted the fabric_interface, and remaining GUIDs from additional_fabric_interfaces
 additional_fabric_interfaces =
```

⚠ UFM treats bonds as a group of IPoIB interfaces. So, for example, if bond0 consists of the interfaces ib4 and ib8, then expect to see GUIDs for ib4 and ib8 in opensm.conf.

⚠ Duplicate interface names are ignored (e.g. ib1,ib1,ib1,ib2,ib1 = ib1,ib2).

Configuring UDP Buffer

This section is relevant only in cases where `telemetry_provider=ibpm`. (By default, telemetry_provider=telemetry).

To work with large-scale fabrics, users should set the set_udp_buffer flag under the [IBPM] section to "yes" for the UFM to set the buffer size (default is "no").

```
# By deafult, UFM does not set the UDP buffer size. For large scale fabrics
# it is recommended to increase the buffer size to 4MB (4194304 bits).
set_udp_buffer = yes
# UDP buffer size
udp_buffer_size = 4194304
```

Virtualization

This allows for supporting virtual ports in UFM.

```
[Virtualization]
# By enabling this flag, UFM will discover all the virtual ports assigned for all hypervisors in the fabric
enable = false
# Interval for checking whether any virtual ports were changed in the fabric
interval = 60
```

Static SM LID

Users may configure a specific value for the SM LID so that the UFM SM uses it upon UFM startup.

```
[SubnetManager]
# 1- Zero value (Default): Disable static SM LID functionality and allow the SM to run with any LID.
#    Example: sm_lid=0
# 2- Non-zero value: Enable static SM LID functionality so SM will use this LID upon UFM startup.
sm_lid=0
```

⚠ To configure an external SM (UFM server running in sm_only mode), users must manually configure the opensm.conf file (`/opt/ufm/conf/opensm/opensm.conf`) and align the value of `master_sm_lid` to the value used for `sm_lid` in gv.cfg on the main UFM server.

Configuring Log Rotation

This section enables setting up the log files rotate policy. By default, log rotation runs once a day by cron scheduler.

```
[logrotate]
#max_files specifies the number of times to rotate a file before it is deleted (this definition will be applied to
#SM and SHARP Aggregation Manager logs, running in the scope of UFM).
#A count of 0 (zero) means no copies are retained. A count of 15 means fifteen copies are retained (default is 15)
max_files = 15
#With max_size, the log file is rotated when the specified size is reached (this definition will be applied to
#SM and SHARP Aggregation Manager logs, running in the scope of UFM). Size may be specified in bytes (default),
#kilobytes (for example: 100k), or megabytes (for exapmle: 10M). if not specified logs will be rotated once a day.
max_size = 3
```

Configuration Examples in gv.cfg

The following show examples of configuration settings in the gv.cfg file:

- Polling interval for Fabric Dashboard information

  ```
  ui_polling_interval = 30
  ```

- [**Optional**] UFM Server local IP address resolution (by default, the UFM resolves the address by gethostip). UFM Web UI should have access to this address.

  ```
  ws_address = <specific IP address>
  ```

- HTTP/HTTPS Port Configuration

```
# WebServices Protocol (http/https) and Port
ws_port = 8088
ws_protocol = http
```

- Connection (port and protocol) between the UFM server and the APACHE server

```
ws_protocol = <http or https>
ws_port = <port number>
```

For more information, see Launching a UFM Web UI Session.

- SNMP `get-community` string for switches (fabric wide or per switch)

```
# default snmp access point for all devices
[SNMP]
port = 161
gcommunity = public
```

- Enhanced Event Management (Alarmed Devices Group)

```
[Server]
auto_remove_from_alerted = yes
```

- Log verbosity

```
[Logging]
# optional logging levels
#CRITICAL, ERROR, WARNING, INFO, DEBUG
level = INFO
```

For more information, see "UFM Logs".

- Settings for saving port counters to a CSV file

```
[CSV]
write_interval = 60
```

```
ext_ports_only = no
```

For more information, see "Saving the Port Counters to a CSV File".

- Max number of CSV files (UFM Advanced)

```
[CSV]
max_files = 1
```

For more information, see "Saving Periodic Snapshots of the Fabric (Advanced License Only)".

> ⚠ The access credentials that are defined in the following sections of the conf/gv.cfg file are used only for initialization:
>
> - SSH_Server
> - SSH_Switch
> - TELNET
> - IPMI
> - SNMP
> - MLNX_OS
>
> To modify these access credentials, use the UFM Web UI. For more information, see "Device Access".

- Configuring the UFM communication protocol with MLNX-OS switches. The available protocols are:
  - http
  - https (default protocol for secure communication)

For configuring the UFM communication protocol after fresh installation and prior to the first run, set the MLNX-OS protocol as shown below.

Example:

```
[MLNX_OS]
protocol = https
port = 443
```

Once UFM is started, all UFM communication with MLNX-OS switches will take place via the configured protocol.

For changing the UFM communication protocol while UFM is running, perform the following:

1. Set the desired protocol of MLNX-OS in the conf/gv.cfg file (as shown in the example above).
2. Restart UFM.
3. Update the MLNX-OS global access credentials configuration with the relevant protocol port. Refer to "Device Access" for help.
   For the http protocol - default port is 80.
   For the https protocol - default port is 443.
4. Update the MLNX-OS access credentials with the relevant port in all managed switches that have a valid IP address.

SM Trap Handler Configuration

The SMTrap handler is the SOAP server that handles traps coming from OpenSM.

There are two configuration values related to this service:

- `osm_traps_debounce_interval` - defines the period the service holds incoming traps
- `osm_traps_throttle_val` - once `osm_traps_debounce_interval` elapses, the service transfers `osm_traps_throttle_val` to the Model Main

> ⚠ By default, the SM Trap Handler handles up to 1000 SM traps every 10 seconds.

CPU Affinity on UFM

This feature allows setting the CPU affinity for the major processes of the UFM (such as ModelMain, SM, SHARP, Telemetry).

In order to increase the UFM's efficiency, the number of context-switches is reduced. When each major CPU is isolated, users can decrease the number of context-switches, and the performance is optimized.

The CPU affinity of these major processes is configured in the following two levels:

- Level 1- The major processes initiation.
- Level 2- Preceding initiation of the model's main subprocesses which automatically uses the configuration used in level 1 and designates a CPU for each of the sub-processes.

According to user configuration, each process is assigned with affinity.

By default, this feature is disabled. In order to activate the feature, configure `Is_cpu_affinity_enabled` with true, check how many CPUs you have on the machine, and set the desired affinity for each process.

For example:

```
[CPUAffinity]
Is_cpu_affinity_enabled=true
Model_main_cpu_affinity=1-4
Sm_cpu_affinity=5-13
SHARP_cpu_affinity=14-22
Telemetry_cpu_affinity=22-23
```

The format should be a comma-separated list of CPUs. For example: 0,3,7-11.

The ModelMain should have four cores, and up to five cores. The SM should have as many cores as you can assign. You should isolate between the ModelMain cores and the SM cores.

SHARP can be assigned with the same affinity as the SM. The telemetry should be assigned with three to four CPUs.

Quality of Service (QoS) Support

Infiniband Quality of Service (QoS) is disabled by default in the UFM SM configuration file.

To enable it and benefit from its capabilities, set the qos flag to TRUE in the /opt/ufm/files/conf/opensm/opensm.conf file.

Example:

```
# Enable QoS setup
qos FALSE
```

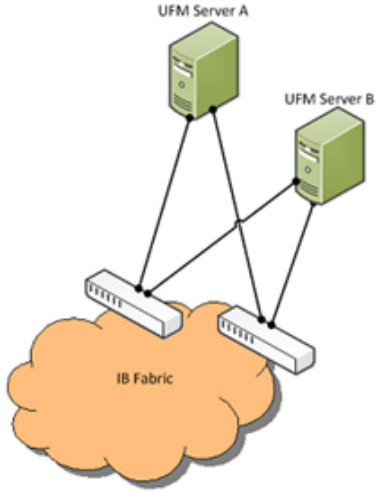
> ⚠ The QoS parameters settings should be carefully reviewed before enablement of the qos flag. Especially, sl2vl and VL arbitration mappings should be correctly defined.

For information on Enhanced QoS, see Appendix – SM Activity Report.

UFM Failover to Another Port

When the UFM Server is connected by two or more InfiniBand ports to the fabric, you can configure UFM Subnet Manager failover to one of the other ports. When failure is detected on an InfiniBand port or link, failover occurs without stopping the UFM Server or other related UFM services, such as mysql, http, DRDB, and so on. This failover process prevents failure in a standalone setup, and preempts failover in a High Availability setup, thereby saving downtime and recovery.

Network Configuration for Failover to IB Port



⚠ UFM SM failover is not relevant for Monitoring mode, because in this mode, UFM must be connected to the fabric over ib0 only.

To enable UFM failover to another port:

- Configure bonding between the InfiniBand interfaces to be used for SM failover. In an HA setup, the UFM active server and the UFM standby server can be connected differently; but the bond name must be the same on both servers.

- Set the value of fabric_interface to the bond name. using the /opt/ufm/scripts/*change_fabric_config.sh* command as described in Configuring General Settings in gv.cfg. If ufma_interface is configured for IPoIB, set it to the bond name as well. These changes will take effect only after a UFM restart. For example, if bond0 is configured on the ib0 and ib1 interfaces, in gv.cfg, set the parameter fabric_interface to bond0.
- If IPoIB is used for UFM Agent, add bond to the ufma_interfaces list as well.

When failure is detected on an InfiniBand port or link, UFM initiates the give-up operation that is defined in the Health configuration file for OpenSM failure. By default:

- UFM discovers the other ports in the specified bond and fails over to the first interface that is up (SM failover)
  If no interface is up:
  - In an HA setup, UFM initiates UFM failover
  - In a standalone setup, UFM does nothing

If the failed link becomes active again, UFM will select this link for the SM only after SM restart.

Delegating Authentication to a Proxy

To allow a custom user authentication, you can configure UFM to delegate the user authentication to a remote Proxy server. The remote Proxy server is written by the user, thus, allowing flexibility on deciding how the authentication is performed.

By default, the feature is disabled. To activate the feature, configure `auth_proxy_enabled` with `true`.

Proxy should use ufmRestV3 to send requests to UFM. The request header should contain a username and role. The available roles are System_Admin, Fabric_Admin, Fabric_Operator, and Monitoring_Only. If the request header is sent without a username or a role, it is rejected by the UFM.

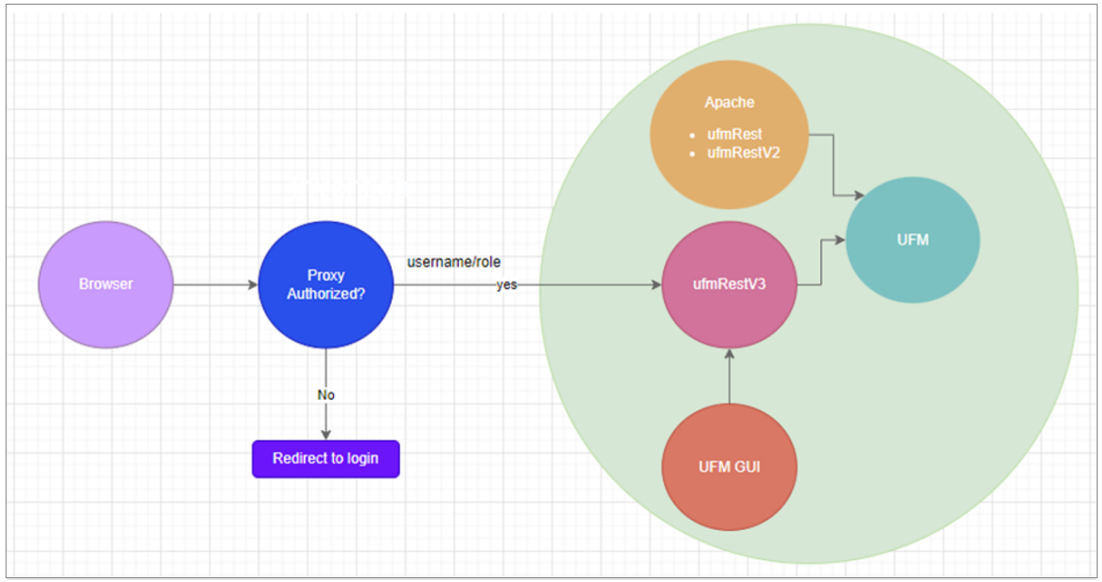
For example:

```
[AuthProxy]
# Defaults to false, but set to true to enable this feature
auth_proxy_enabled = true
# HTTP Header name that will contain the username
auth_proxy_header_name = X_WEBAUTH_USER
# HTTP Header name that will contain the user roles. The available roles are as follows: System_Admin,
Fabric_Admin, Fabric_Operator, and Monitoring_Only
auth_proxy_header_role = X_WEBAUTH_ROLE

# Set to `true` to enable auto sign up of users who do not exist in UFM DB. Defaults to `true`.
```

```
auth_proxy_auto_sign_up = true
# Limit where auth proxy requests come from by configuring a list of IP addresses.
# This can be used to prevent users spoofing the X_WEBAUTH_USER header.
# This option is required
# Example `whitelist = 192.168.1.1, 192.168.1.0/24, 2001::23, 2001::0/120`
auth_proxy_whitelist =
```

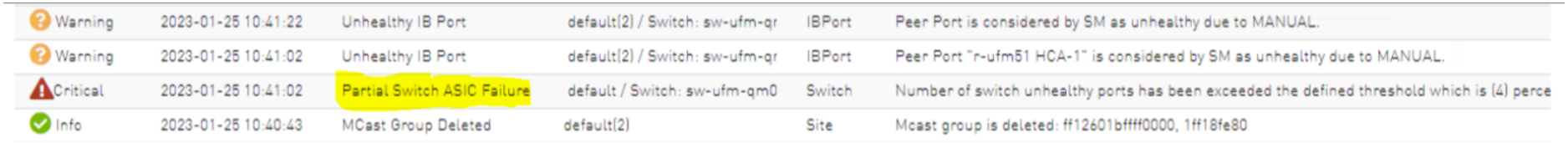
The following chart describes the flow:



Configuring Partial Switch ASIC Failure Events

UFM can identify switch ASIC failure by detecting pre-defined portion of the switch ports, reported as unhealthy. By default, this portion threshold is set to 20% of the total switch ports. Thus, the UFM will trigger the partial switch ASIC event in case the number of unhealthy switch ports exceeds 20% of the total switch ports.

You can configure UFM to control Partial Switch ASIC Failure events. To configure, you may use the CLI command `ufm unhealthy-ports switch-ports-threshold`.For an example, in case the switch has 32 ports, once 7 ports are detected as unhealthy ports, the UFM will trigger the partial switch ASIC event. Example:



Warning	2023-01-25 10:41:22	Unhealthy IB Port	default(2) / Switch: sw-ufm-qr	IBPort	Peer Port is considered by SM as unhealthy due to MANUAL.
Warning	2023-01-25 10:41:02	Unhealthy IB Port	default(2) / Switch: sw-ufm-qr	IBPort	Peer Port "r-ufm51 HCA-1" is considered by SM as unhealthy due to MANUAL.
Critical	2023-01-25 10:41:02	Partial Switch ASIC Failure	default / Switch: sw-ufm-qm0	Switch	Number of switch unhealthy ports has been exceeded the defined threshold which is (4) perce
Info	2023-01-25 10:40:43	MCast Group Deleted	default(2)	Site	Mcast group is deleted: ff12601bffff0000, 1ff18fe80

Enabling Network Fast Recovery

> ⚠ To enable the Network Fast Recovery feature, ensure that all switches in the fabric use the following MLNX-OS/firmware versions:
>
> - MLNX-OS version 3.10.6004 and up
> - Quantum firmware versions:
>   - Quantum FW v27.2010.6102 and up
>   - Quantum2 FW v31.2010.6102 and up

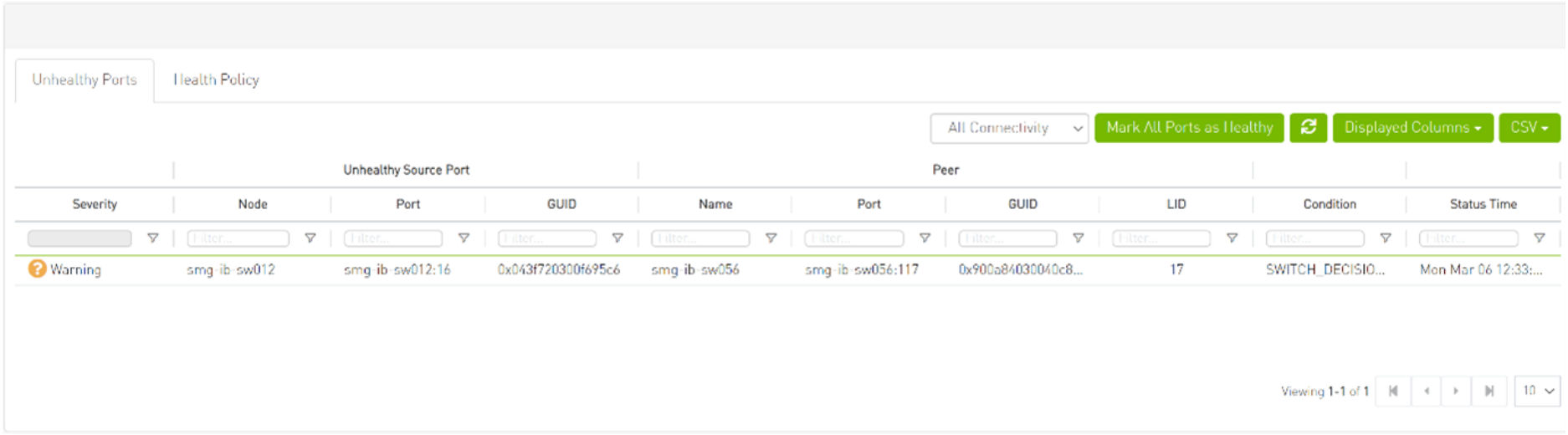
Fast recovery is a switch-firmware based facility for isolation and mitigation of link-related issues. This system operates in a distributed manner, where each switch is programmed with a simple set of rule-based triggers and corresponding action protocols. These rules permit the switch to promptly react to substrandard links within its locality, responding at a very short reaction time - as little as approximately 100 milliseconds. The policy is provided and managed via the UFM & SM channel. Moreover, every autonomous action taken by a switch in the network is reported to the UFM.

The immediate reactions taken by the switch enable SHIELD and pFRN. These mechanisms collaborate to rectify routing within the proximity of the problematic link before it can disrupt transactions at the transport layer. Importantly, this process occurs rapidly, effectively limiting the spreading of congestion to a smaller segment of the network.

To use the Network Fast Recovery feature, you need to enable the designated trigger in the `conf/gv.cfg` file. By doing this, you can specify which triggers the UFM will support. Alternatively, you may run the following command:

```
ufm network-fast-recovery enable
```

The "Unhealthy Ports" page provides visibility of these ports. If desired, the user can mark a port as healthy, triggering a restart of that specific port on the switch.
The trigger that initiated the isolation of ports can be viewed under the "Condition" column, as seen below.



Disabling Rest Roles Access Control

By default, the Rest Roles Access Control feature is enabled. It can be disabled by setting the `roles_access_control_enabled` flag to false:

```
[RolesAccessControl]
roles_access_control_enabled = true
```

Setting up Telemetry in UFM

Setting up telemetry deploys UFM Telemetry as bare metal on the same machine. Historical data is sent to SQLite database on the server and live data becomes available via UFM UI or REST API.

Enabling UFM Telemetry

The UFM Telemetry feature is enabled by default and the provider is the UFM Telemetry. The user may change the provider via flag in `conf/gv.cfg`

The user may also disable the History Telemetry feature in the same section.

```
[Telemetry]
history_enabled=True
```

Changing UFM Telemetry Default Configuration

There is an option to configure parameters on a telemetry configuration file which takes effect after restarting the UFM or failover in HA mode.

The `launch_ibdiagnet_config.ini` default file is located under `/opt/ufm/conf/telemetry_defaults` and is copied to the telemetry configuration location ((`/opt/ufm/conf/telemetry`) upon startup UFM.

All values taken from the default file take effect at the deployed configuration file except for the following:

Note that normally the user does not have to do anything and they get two pre-configured instances – one for low frequency and one for higher-frequency sampling of the network.

Value	Description
`hca`	-
`scope_file`	-
`plugin_env_PROMETHEUS_ENDPOINT`	The port on which HTTP endpoint is configured
`plugin_env_PROMETHEUS_INDEXES`	Configures how data is indexed and stored in memory
`config_watch_enabled=true`	Configures network watcher to inform ibdiagnet that network topology has changed (as ibdiagnet lacks the ability to re-discover network changes)

Value	Description
plugin_env_PROMETHEUS_CSET_DIR	Specifies where the counterset files, which define the data to be retrieved and the corresponding counter names.
num_iterations	The number of iterations to run before 'restarting', i.e. rediscovering fabric.
plugin_env_CLX_RESTART_FILE	A file that is 'touched' to indicate that an ibdiagnet restart is necessary

The following attributes are configurable via the gv.cfg:

- sample_rate (gv.cfg → dashboard_interval) - only if manual_config is set to false
- prometheus_port

Supporting Generic Counters Parsing and Display

As of UFM v6.11.0, UFM can support any numeric counters from the HTTP endpoint. The list of supported counters are fetched upon starting the UFM from all the endpoints that are configured.

Some of the implemented changes are as follows:

1. Counter naming - all counters naming convention is extracted from the HTTP endpoint. The default `cset` file is configured as follows: "`Infiniband_LinkIntegrityErrors=^LocalLinkIntegrityErrorsExtended$`" to get this name to the UFM.
   Counters received as floats should contain an "_f" suffix such as: `Infiniband_CBW_f=^infiniband_CBW$`
2. Attribute units - To see units of a specific counter on the UI graphs, configure the `cset` file to have the counter returned as "`counter_name_u_unit`" .
3. Telemetry History:

   The SQLite history table `(ufm_telemetry.db - telemetry_calculated)`, contains the new naming convention of the telemetry counters.In the case of an upgrade, all previous columns that were configured are renamed following the new naming convention, and then, the data is saved.if a new counter that is not in the table needs to be supported, the table is altered upon UFM start.
4. New counter/`cset` to fetch - if there is a new `cset`/counter that needs to be supported AFTER the UFM already started, preform system restart.

5. Created New API/UfmRestV2/telemetry/counters for the UI visualization. This API returns a dictionary containing the counters that the UFM supports, based on the fetched URLs and their units (if known).

Supporting Multiple Telemetry Instances Fetch

This functionality allows users to establish distinct Telemetry endpoints that are defined to their preferences.

Users have the flexibility to set the following aspects:

- Specify a list of counters they wish to pull. This can be achieved by selecting from an existing, predefined counters set (`cset file`) or by defining a new one.
- Set the interval at which the data should be pulled.

Upon initiating the Telemetry endpoint, users can access the designated URL to fetch the desired counter data.

To enable this feature, under the [Telemetry] section in `gv.cfg`,the flag named "`additional_cset_url`" holds the list of additional URLs to be fetched. the URLs should be separated by " " (with a space) and should follow the following format: http://<IP>:<PORT>/csv/<CSET_NAME>. For example http://10.10.10.10:9001/csv/minimal http://10.10.10.10:9002/csv/test.

> ⚠ Only csv extensions are supported.

Each UFM Telemetry instance run by UFM can support multiple cset (counters set) in parallel.If the user would like to have a second cset file fetched by UFM and exposed by the same UFM Telemetry instance, the new cset file should be placed under `conf/telemetry/prometheus_configs/cset/` and configured in gv.cfg to fetch its data as described above.

Secondary Telemetry

As a default configuration, a second UFM Telemetry instance runs, granting access to an extended set of counters that are not available in the default telemetry session. The default telemetry session is used for the UFM Web UI dashboard and user-defined telemetry views. These additional counters can be accessed via the following API endpoint: **http://<UFM_IP>:9002/csv/xcset/low_freq_debug**. It is important to note that these exposed counters are not accessible through UFM's REST APIs.

All the configurations for the second telemetry can be found under `conf/secondary_telemetry/`, where the defaults are located under `conf/secondary_telemetry_defaults/`. The second telemetry instance also allows telemetry data to be exposed on disabled ports, although this feature can be disabled if desired.

The relevant flags in the gv.cfg file are as follows:

- secondary_telemetry = true (To enable or disable the entire feature)
  Alternatively, secondary telemetry can be enabled or disabled via the following CLI command:

  ```
  [no] ufm low-frequency-telemetry enable
  ```

  The `no` form disables the feature.
  UFM restart is required for changes to take effect. Run:

  ```
  ufm process model restart
  ```

- secondary_endpoint_port = 9002 (The endpoint's exposed port)
- secondary_disabled_ports = true (If set to true, secondary telemetry will expose data on disabled ports)

For the list of secondary telemetry fields and available counters, please refer to Appendix – Secondary Telemetry Fields.

Stopping Telemetry Endpoint Using CLI Command

To stop secondary telemetry endpoint only using the CLI you may run the following command:

```
no ufm process telemetry start
```

Exposing Switch Aggregation Nodes Telemetry

To expose switches SHARP aggregation nodes telemetry, follow the below steps:

- Configure the secondary telemetry instance. Run: Edit the configuration file: `conf/secondary_telemetry_defaults/launch_ibdiagnet_config.ini`
- Set the following:

    - arg_16=--sharp --sharp_opt dsc
    - plugin_env_CLX_EXPORT_API_SKIP_SHARP_PM_COUNTERS=0
- Add the wanted attributes to the default `xcset` or to a new one:
    - New `xcset` -
        - Edit the configuration file: `conf/secondary_telemetry/prometheus_configs/cset/<name for your choise>.xcset`
        - After restarting, query curl `http://<UFM_IP>:9002/csv/xcset/<chosen_name>`
    - Existing `xcset` -
        - Edit the configuration file: `conf/secondary_telemetry/prometheus_configs/cset/low_freq_debug.xcset`
    - Add the following attributes:
        - packet_sent
        - ack_packet_sent
        - retry_packet_sent
        - rnr_event
        - timeout_event
        - oos_nack_rcv
        - rnr_nack_rcv
        - packet_discard_transport
        - packet_discard_sharp
        - aeth_syndrome_ack_packet
        - hba_sharp_lookup
        - hba_received_pkts
        - hba_received_bytes
        - hba_sent_ack_packets
        - rcds_sent_packets
        - hba_sent_ack_bytes
        - rcds_send_bytes
        - hba_multi_packet_message_dropped_pkts
        - hba_multi_packet_message_dropped_bytes
- Restart telemetry. Run:

```
ufm process telemetry restart
```

Installing License

1. Log into the CLI as admin and use the selected password.
2. Install the license.

```
ufm-appliance [ mgmt-sa ] (config) # ufm license install ftp://admin@192.168.11.1/home/admin/licenses/volt-ufm-advanced.lic
```

Historical Telemetry Collection in UFM

Storage Considerations

UFM periodically collects fabric port statistics and saves them in its SQLite database. Before starting up UFM Enterprise, please consider the following disk space utilization for various fabric sizes and duration.
The measurements in the table below were taken with sampling interval set to once per 30 seconds.

⚠ Be aware that the default sampling rate is once per 300 seconds. Disk utilization calculation should be adjusted accordingly.

Number of Nodes	Ports per Node	Storage per Hour	Storage per 15 Days	Storage per 30 Days
16	8	1.6 MB	576 MB (0.563 GB)	1152 MB (1.125 GB)
100	8	11 MB	3960 MB (3.867 GB)	7920 MB (7.734 GB)
500	8	50 MB	18000 MB (17.58 GB)	36000 MB (35.16 GB)
1000	8	100 MB	36000 MB (35.16 GB)	72000 MB (70.31 GB)

UFM-SDN Appliance Installation and Initial Configuration

- Configuring UFM-SDN Appliance for the First Time
- Obtaining the License
- Initial UFM-SDN Application Configuration
- Installing UFM-SDN Appliance License
- Setting Up Telemetry in UFM-SDN Appliance
- Installing and Running UFM-SDN Agent Software
- Upgrading UFM-SDN Appliance Software

Configuring UFM-SDN Appliance for the First Time

The appliance has multiple Ethernet management interfaces. The primary management interface is eth0. An additional interface exists, for connecting to a remote management controller (It usually connects to the same management network as eth0). For using out-of-the-box DHCP settings: Default hostname for the appliance (over eth0) is "ufm-appliance-[MAC ADDRESS]". The MAC address for eth0 is available on the pull-tab and can be configured in the DHCP server. To use the remote management controller with DHCP, the free-range IP allocation must be enabled on the DHCP server. A static IP address for remote management interface can be configured via the CLI ("`chassis remote-management ip`" command).

The following are two methods of connecting UFM SDN Appliance:

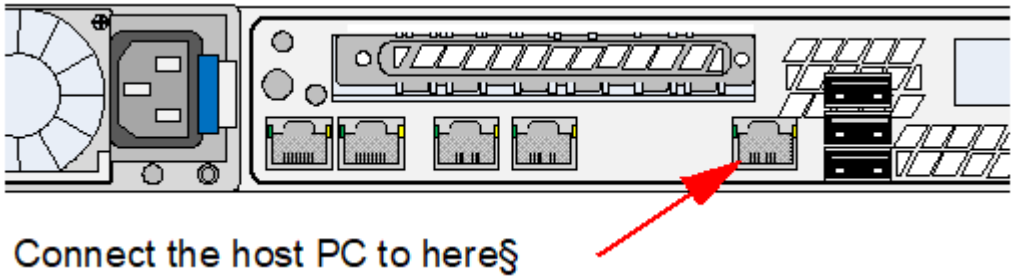
- "Connecting UFM®-SDN Appliance via a Serial Port"
- "Connecting UFM®-SDN Appliance via Remote Management"

Connecting UFM-SDN Appliance via Serial Port

The appliance supports a direct connection via a serial port. Please follow the steps below to connect to the appliance.

> ⚠ Configuration via a serial port is only required if you want to use a static IP address and not the out-of-the-box DHCP setting for eth0. Otherwise, an IP is assigned by the DHCP server and you can login to the CLI over LAN.

1. Connect the host PC to the (RJ-45) port of the appliance system using the supplied cable. The Console ports for the appliance is shown below.



> ⚠ Make sure to connect to the Serial RJ-45 port of the appliance and not to the (Ethernet) eth0 port.

2. Configure a serial terminal program (for example, HyperTerminal, minicom, or Tera Term) on your host PC with the settings described in the following table.

Serial Terminal Program Configuration

Parameter	Setting
Baud Rate	Gen1: 9600 Gen2: 115200
Data bits	8
Stop bits	1
Parity	None
Flow Control	None

Connecting UFM-SDN Appliance via Remote Management

UFM-SDN Appliance includes an Intel® motherboard and a built-in remote management controller. The Remote Management Controller module features remote KVM access and control through LAN or Internet. The Remote Management can be configured either via the UFM-SDN Appliance CLI or the Remote Management Web Console.

For further information, please refer to the UFM-SDN Command Reference Guide, section UFM-SDN Appliance Chassis Management, and UFM-SDN Appliance Remote Management User Manual.

Configuring UFM-SDN Appliance Using CLI Wizard

1. Login as admin and use "admin" as the password.
   This starts the Mellanox configuration wizard.
2. Go through the configuration wizard. The table below shows an example of a wizard session.

Mellanox UFM-SDN Appliance Configuration Wizard

Wizard Session Display (Example)	Comments
Do you want to use the wizard for initial configuration? yes	You must perform this configuration the first time you operate the appliance or after resetting the appliance to the factory defaults. Type ‘y’ and then press <Enter>.
Step 1: Hostname? [ufm-appliance-898b4e]	If you wish to accept the default hostname, then press <Enter>. Otherwise, type a different hostname and press <Enter>.
Step 2: Use DHCP on eth0 interface? [yes]	Perform this step to obtain an IP address for the appliance. (eth0 is the management port of the appliance.) If you wish the DHCP server to assign the IP address, type ‘yes’ and press <Enter>. If you type ‘no’ (no DHCP), then you will be asked whether you wish to use the ‘zeroconf’ configuration or not. If you enter ’yes’ (yes Zeroconf), the session will continue as shown in ***Configuration Wizard Session - IP Zeroconf Configuration*** table below. If you enter "no" (no Zeroconf), then you need to enter a static IP, and the session will continue as shown in the ***Static IP configuration*** table below.
Step 3: Enable IPv6 [yes]	Perform this step to enable IPv6 on management ports. If you wish to enable IPv6, type "yes" and press <Enter>. If you enter "no" (no IPv6), then you will automatically be referred to Step 5.

Wizard Session Display (Example)	Comments
Step 4: Enable IPv6 autoconfig (SLAAC) on eth0 interface	Perform this step to enable StateLess address autoconfig on external management port. If you wish to enable it, type "yes" and press <Enter>. If you wish to disable it, enter "no".
Step 5: Use DHCPv6 on eth0 interface? [yes]	Perform this step to enable DHCPv6 on the eth0 interface.
Step 6: Admin password (Must be typed)? **Step 6**: Confirm admin password?	To avoid illegal access to the machine, please type a password for the admin user and then press <Enter>. Confirm the password by re-entering it. ⚠ Note that password characters are not printed.
Step 7: Monitor password (Must be typed)? **Step 7**: Confirm monitor password?	To avoid illegal access to the machine, please type a password for the monitor user and then press <Enter>. Confirm the password by re-entering it. ⚠ Note that password characters are not printed.
Step 8: bond0 IP address and masklen? [0.0.0.0/0]	Allows defying IP address for IPoIB bond (bond is pre-configured with ib0 & ib1 ports).
You have entered the following information: 1. Hostname: ufm-appliance-898b4e 2. Use DHCP on eth0 interface: yes 3. Enable IPv6: yes 4. Enable IPv6 autoconfig (SLAAC) on eth0 interface: no 5. Enable DHCPv6 on eth0 interface: yes 6. Admin password (Must be typed): (CHANGED) 7. Monitor password (Must be typed): (CHANGED)	The wizard displays a summary of your choices and then asks you to confirm the choices or to re-edit them. Either press <Enter> to save changes and exit, or enter the configuration step number that you wish to return to. To run the command "configuration jump-start", you must be in Config mode.

Wizard Session Display (Example)	Comments
8. bond0 IPv4 address and masklen: 1.1.1.1/24 To change an answer, enter the step number to return to. Otherwise hit <enter> to save changes and exit. Choice: Configuration changes saved. To return to the wizard from the CLI, enter the "configuration jump-start" command from configure mode. Launching CLI...	

Configuration Wizard Session - IP Zeroconf Configuration

Wizard Session Display - IP Zeroconf Configuration (Example)
Step 1: Hostname? [r-ufm236] Step 2: Use DHCP on eth0 interface? [no] Step 3: Use zeroconf on eth0 interface? [no] Step 4: Primary IPv4 address and masklen? [0.0.0.0/0] Step 5: Default gateway? Step 6: Primary DNS server? Step 7: Domain name? Step 8: Enable IPv6? [yes] Step 9: Enable IPv6 autoconfig (SLAAC) on eth0 interface? [no] Step 10: Enable DHCPv6 on eth0 interface? [yes] Step 11: Admin password (Enter to leave unchanged)?

Wizard Session Display - IP Zeroconf Configuration (Example)

```
Step 12: Monitor password (Enter to leave unchanged)?
Step 13: bond0 IPv4 address and masklen? [1.1.1.1/24]

You have entered the following information:

 1. Hostname: r-ufm236
 2. Use DHCP on eth0 interface: no
 3. Use zeroconf on eth0 interface: no
 4. Primary IPv4 address and masklen: 0.0.0.0/0
 5. Default gateway:
 6. Primary DNS server:
 7. Domain name:
 8. Enable IPv6: yes
 9. Enable IPv6 autoconfig (SLAAC) on eth0 interface: no
 10. Enable DHCPv6 on eth0 interface: yes
 11. Admin password (Enter to leave unchanged): (unchanged)
 12. Monitor password (Enter to leave unchanged): (unchanged)
 13. bond0 IPv4 address and masklen: 1.1.1.1/24

To change an answer, enter the step number to return to.
Otherwise hit <enter> to save changes and exit.


Choice:

Configuration changes saved.

To return to the wizard from the CLI, enter the "configuration jump-start"
command from configure mode. Launching CLI...
<hostname> [standalone: master] >
```

Static IP Configuration

Wizard Session Display - Static IP Configuration (Example)

```
Do you want to use the wizard for initial configuration? y

Step 1: Hostname? [r-ufm236]
Step 2: Use DHCP on eth0 interface? [no]
Step 3: Use zeroconf on eth0 interface? [no]
Step 4: Primary IPv4 address and masklen? [0.0.0.0/0] 10.209.36.38/24
Step 5: Default gateway? 10.209.36.1
Step 6: Primary DNS server?
Step 7: Domain name?
Step 8: Enable IPv6? [yes]
Step 9: Enable IPv6 autoconfig (SLAAC) on eth0 interface? [no]
Step 10: Enable DHCPv6 on eth0 interface? [yes]
Step 11: Admin password (Enter to leave unchanged)?
Step 12: Monitor password (Enter to leave unchanged)?
Step 12: Confirm monitor password?
Step 13: bond0 IPv4 address and masklen? [1.1.1.1/24]

You have entered the following information:

 1. Hostname: r-ufm236
 2. Use DHCP on eth0 interface: no
 3. Use zeroconf on eth0 interface: no
 4. Primary IPv4 address and masklen: 10.209.36.38/24
 5. Default gateway: 10.209.36.1
 6. Primary DNS server:
 7. Domain name:
 8. Enable IPv6: yes
 9. Enable IPv6 autoconfig (SLAAC) on eth0 interface: no
 10. Enable DHCPv6 on eth0 interface: yes
 11. Admin password (Enter to leave unchanged): (unchanged)
 12. Monitor password (Enter to leave unchanged): (CHANGED)
 13. bond0 IPv4 address and masklen: 1.1.1.1/24

To change an answer, enter the step number to return to.
```

Wizard Session Display - Static IP Configuration (Example)
Otherwise hit <enter> to save changes and exit. Choice: Configuration changes saved. To return to the wizard from the CLI, enter the "configuration jump-start" command from configure mode. Launching CLI... <hostname>[standalone: master] >

Obtaining the License

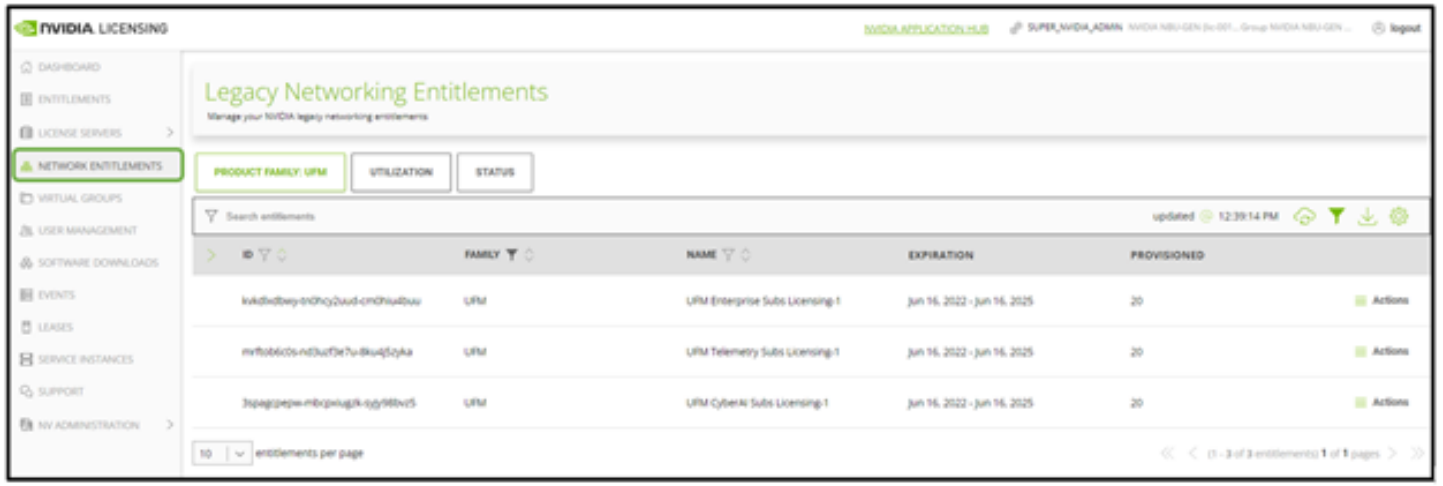
UFM-SDN Appliance is licensed per managed servers according to the UFM license agreement.
When you purchase UFM-SDN Appliance, you will receive an email with instructions on obtaining your product license. A valid license is a prerequisite for the installation and operation of UFM-SDN Appliance.

UFM licenses are per managed node and are aggregative. If you install an additional license, the system adds the previous node number and the new node number and manages the sum of the nodes. For example, if you install a license for 10 managed nodes and an additional license for 15 nodes, UFM will be licensed for up to 25 managed nodes.

To obtain the license:

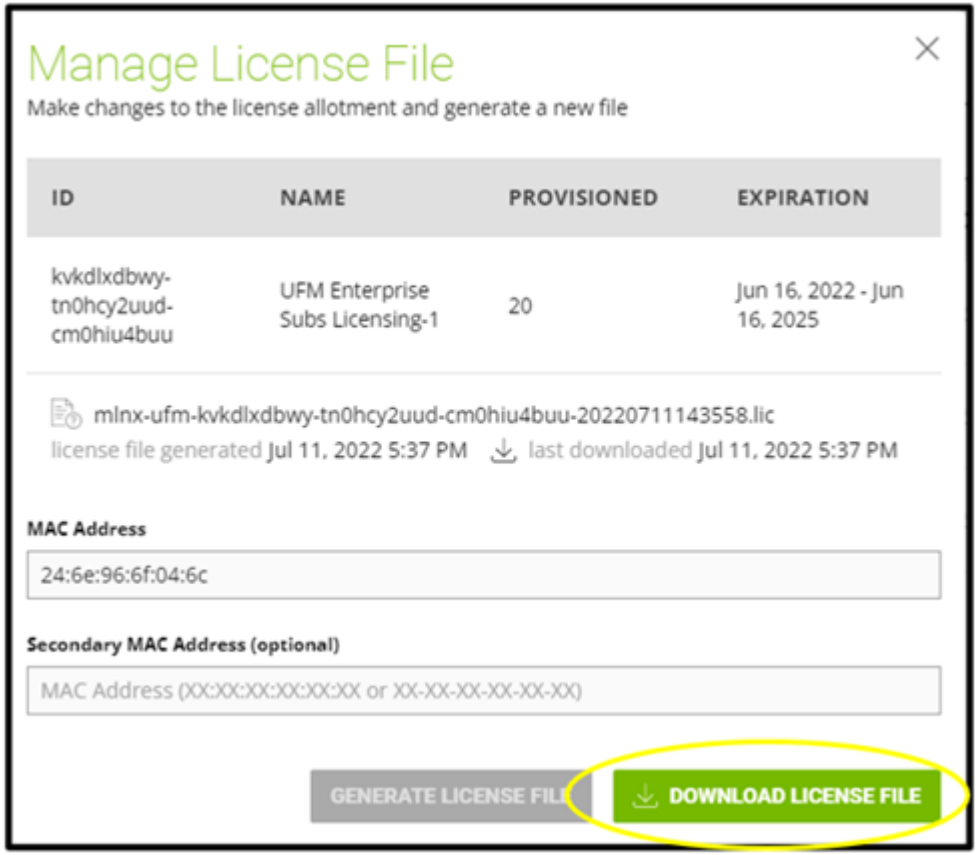
1. Go to NVIDIA's Licensing and Download Portal and log in as specified in the licensing email you received.
    - If you did not receive your NVIDIA Licensing and Download Portal login information, contact your product reseller.

2. If you purchased UFM directly from NVIDIA and you did not receive the login information, contact enterprisesupport@nvidia.com. Click on the Network Entitlements tab. You'll see a list with the serial licenses of all your software products and software product license information and status.



3. Select the license you want to activate and click on the “Actions” button.

4. In the MAC Address field, enter the MAC address of the delegated license-registered host. If applicable, in the HA MAC Address field, enter your High Availability (HA) server MAC address. If you have more than one NIC installed on a UFM Server, use any of the MAC addresses.



5. Click on Generate License File to create the license key file for the software.
6. Click on Download License File and save it on your local computer.

If you replace your NIC or UFM server, repeat the process of generating the license to set new MAC addresses. You can only regenerate a license two times. To regenerate the license after that, contact NVIDIA Sales Administration at enterprisesupport@nvidia.com.

Initial UFM-SDN Application Configuration

- Changing UFM Running Mode
- Modifying UFM Configuration Files
- Configuring General Settings in gv.cfg
- Configuring Two Appliances to Run in High Availability
- Configuring UFM-SDN Appliance with External UFM-SMs
- CPU Affinity on UFM-SDN Appliance
- Quality of Service (QoS) Support
- UFM Failover to Another Port
- Starting UFM
- User Authentication
- Licensing
- Launching a UFM Web UI Session

Changing UFM Running Mode

> ⚠ In Management or High Availability modes, ensure that all Subnet Managers in the fabric are disabled before running UFM. Any remaining active Subnet Managers will prevent UFM from running.

To change the UFM running mode:

1. Log into the `CLI` as admin and use the selected password.
2. Change the running mode.

```
ufm-appliance > enable
ufm-appliance # configure terminal
ufm-appliance [ mgmt-sa ] (config) # ufm mode <mode>
```

UFM running mode:

- mgmt. - management mode
- mon - monitoring mode
- mgmt.-allow-sm - management mode with "External SMs" in the fabric
- remote-mon - connects UFM-SDN Appliance to the Remote Fabric Collector
- sm-only - make appliance work as an "External SM" machine
  For further information, please refer to the UFM-SDN Appliance Command Reference Guide.

Running the UFM Server in Monitoring Mode

Run UFM in Monitoring mode while running concurrent instances of Subnet Manager on Mellanox switches. Monitoring and event management capabilities are enabled in this mode. UFM non-monitoring features such as provisioning and performance optimization are disabled in this mode.

The following table describes whether features are enabled or disabled in Monitoring mode.

Features Enabled/Disabled in Monitoring Mode

Feature	Enabled/Disabled in Monitoring Mode
Fabric Discovery	Enabled
Topology Map	Enabled
Fabric Dashboard	Enabled
Fabric Monitoring	Enabled
Alerts and Thresholds (inc. SNMP traps)	Enabled
Fabric Logical Model	Enabled
Device Management	Enabled
Subnet Manager and plugins	Disabled

Feature	Enabled/Disabled in Monitoring Mode
Subnet Manager Configuration	Disabled
Automatic Fabric Partitioning	Disabled
Central Device Management	Disabled
Quality of Service	Disabled
Failover (High Availability mode)	Disabled
Traffic Aware Routing Algorithm	Disabled
Fabric Collective Accelerator (FCA)	Disabled
Integration with Schedulers	Disabled
Unhealthy Ports	Disabled

In Monitoring mode, UFM periodically discovers the fabric and updates the topology maps and database.

For Monitoring mode, connect UFM to the fabric using port ib0 only. The fabric must have a subnet manager (SM) running on it (on another UFM, HBSM, or switch SM).

⚠ When UFM is running in Monitoring mode, the internal OpenSM is not sensitive to changes in OpenSM configuration (opensm.conf).

⚠ When running in Monitoring mode, the following parameters are automatically overwritten in the conf/opensm/opensm_mon.conf file on startup:

- event_plugin_name osmufmpi
- event_plugin_options --vendinfo -m 0

Any other configuration is not valid for Monitoring mode.

To run in Monitoring mode:

1. Stop the UFM service. Run:

```
ufm-appliance [ mgmt-sa ] (config) # no ufm start
```

2. Change the UFM mode. Run:

```
ufm-appliance [ mgmt-sa ] (config) # ufm mode mon
```

3. (Optional) In the *conf/gv.cfg* configuration file:
   - If desired, change **mon_mode_discovery_period** (the default is 60 seconds)
   - Set **reset_mode** to no_reset
     It is recommend to use this setting when running multiple instances of UFM so that each port counter is not reset by different UFM instances. For more information, see Resetting Physical Port Counters.
4. Start the UFM service. Run:

```
ufm-appliance [ mgmt-sa ] (config) # ufm start
```

The running mode is set to Monitoring, and the frequency of fabric discovery is updated according to the setting of mon_mode_discovery_period.

Note that a monitor icon will appear at the top of the navigation bar indicating that monitoring mode is enabled:

Modifying UFM Configuration Files

To modify UFM configuration files:

1. Export the configuration to a "zip" file.

```
ufm-appliance > enable
ufm-appliance # configure terminal
ufm-appliance [ mgmt-sa ] (config) # ufm configuration export
```

2. Upload the file to a USB storage device or to a remote server.

```
ufm-appliance [ mgmt-sa ] (config) # ufm configuration upload ufm-config-20121128-180857.zip scp://
mlnx:123456@172.30.3.201/tmp
```

3. Extract the "zip" file ("conf" folder).
4. Perform any required modification to the files placed in the "conf" folder.
5. Pack the "conf" folder back into the "zip" file.
6. Fetch the file to the appliance.

```
ufm-appliance [ mgmt-sa ] (config) # ufm configuration fetch scp://mlnx:123456@172.30.3.201/tmp/ufm-
config-20121128-180857-modified.zip
```

7. [Optional] Stop UFM server.
   Some parameters can be modified and have effect without restarting UFM®-SDN SM.

```
ufm-appliance [ mgmt-sa ] (config) # no ufm start
```

8. Import the configuration from the "zip" file.

```
ufm-appliance [ mgmt-sa ] (config) # ufm configuration import ufm-config-20121128-180857-modified.zip
```

   Due to the configuration import, a signal is sent to the SM.
9. [Optional, only if Step 5 is performed] Start UFM server.

```
ufm-appliance [ mgmt-sa ] (config) # ufm start
Starting UFM, please wait...
```

```
ufm-appliance [ mgmt-sa ] (config) # show ufm status
UFM mode: Management
Local - ufmapl_3467ed (172.30.30.8)
=======================================
UFM Running
SM Running
IBPM Running
Health Running
MySQL Running
ufm-appliance [ mgmt-sa ] (config) #
```

For further information, please refer to the UFM®-SDN Appliance Command Reference Guide.

Configuring General Settings in gv.cfg

Configure general settings in the *conf/gv.cfg* file. When running UFM in HA mode, the gv.cfg file is replicated to the standby server.

Configuring Fabric Interface

Fabric interface should be set to one of the InfiniBand IPoIB interfaces which connect the UFM/SM to the fabric:

```
fabric_interface = ib0
```

- By default, `fabric_interface` is set to `ib0`
- `fabric_interface` must be up and running before UFM startup, otherwise, UFM will not be able to run
- `fabric_interface` must be configured with a valid IPv4 address before UFM startup, otherwise, UFM will not be able to run

Running UFM in Monitoring Mode

```
monitoring_mode = yes
```

For more information, see Changing UFM Running Mode.

Enabling Predefined Groups

```
enable_predefined_groups = true
```

⚠ By default, pre-defined groups are enabled. In very large scale fabrics, pre-defined groups can be disabled in order to allow faster startup of UFM.

Enabling Multi-NIC Host Grouping

```
multinic_host_enabled = true
```

⚠ Upon first installation of UFM 4.6.1 and above, mutli-NIC host grouping is enabled by default. However, if a user is upgrading from an older version, then this feature will be disabled for them.

⚠ It is recommended to set the value of this parameter before running UFM for the first time.

Running UFM SM Only (UFM HA with Additional SMs)

```
management_mode = sm_only
```

For more information, see Configuring UFM-SDN Appliance with External UFM-SMs.

Running UFM Over IPv6 Network Protocol

The default multicast address is configured to an IPv4 address. To run over IPv6, this must be changed to the following in section UFMAgent of gv.cfg.

```
[UFMAgent]
...
# if ufmagent works in ipv6 please set this multicast address to FF05:0:0:0:0:0:0:15F
mcast_addr = FF05:0:0:0:0:0:0:15F
```

Adding SM Plug-in to event_plugin_name Option

Add `ccmgr` for example:

```
# Event plugin name(s)
event_plugin_name osmufmpi ccmgr
```

Add the plug-in options file to the `event_plugin_options` option:

```
# Options string that would be passed to the plugin(s)
event_plugin_options  ccmgr --conf_file /opt/ufm/files/conf/cc-mgr.conf
```

These plug-in parameters are copied to the opensm.conf file in Management mode only.

Enabling SHARP Aggregation Manager

⚠ Mellanox SHARP AM is disabled by default.

```
[Sharp]
sharp_enabled = true
```

⚠ Upon startup of UFM or SHARP Aggregation Manager, UFM will resend all existing persistent allocation to SHARP AM.

Multi-port SM

OpenSM supports up to 8 GUIDs being configured in opensm.conf file. This allows users to configure several fabric interfaces for OpenSM in UFM configuration file conf/gv.cfg. This means that users can specify multiple IPoIB interfaces or bonds in UFM, and UFM will translate them to GUIDs and add them to the OpenSM configuration file (conf/opensm/opensm.conf). If users specify more than 8 interfaces, the extra interfaces are ignored.

```
[Server]
# True/false flag to configure OpenSM with multiple GUIDs
enable_multi_port_sm = false
# When enabling multi_port_sm, specify here the additional fabric interfaces for opensm.conf
# Example: ib1,ib2,ib5 (OpenSM will support the first 8 GUIDs where first GUID are extracted
# from the fabric_interface field. The remaining GUIDs are taken from the
# additional_fabric_interfaces field.
additional_fabric_interfaces =
```

⚠ UFM treats bonds as a group of IPoIB interfaces. So, for example, if bond0 consists of the interfaces ib0 and ib1, then expect to see GUIDs for ib0 and ib1 in opensm.conf.

⚠ Duplicate interface names are ignored (e.g. ib1,ib1,ib1,ib2,ib1 = ib1,ib2).

Setting UDP Buffer

To work with large scale fabrics, users should set the set_udp_buffer flag under the [IBPM] section to yes in order for UFM to set the buffer size (default is no).

```
# By deafult, UFM does not set the UDP buffer size. For large scale fabrics
# it is recommended to increase the buffer size to 4MB (4194304 bits).
set_udp_buffer = yes
```

```
# UDP buffer size
udp_buffer_size = 4194304
```

Virtualization

Virtualization allows for supporting virtual ports in UFM.

```
[Virtualization]
# By enabling this flag, UFM will discover all the virtual ports assigned for all hypervisors in the fabric
enable = false
# Interval for checking whether any virtual ports were changed in the fabric
interval = 60
```

Static SM LID

Users may configure a static value for the SM LID in order so UFM SM runs on this LID upon UFM startup.

```
[SubnetManager]
# 1- Zero value (Default): Disable static SM LID functionality and allow the SM to run with any LID.
#    Example: sm_lid=0
# 2- Non-zero value: Enable static SM LID functionality so SM will use this LID upon UFM startup.
sm_lid=0
```

⚠ To configure an external SM (UFM server running in sm_only mode), users must manually configure the opensm.conf file (`conf/opensm/opensm.conf`) and align the value of `master_sm_lid` to the value used for `sm_lid` in gv.cfg on the main UFM server.

Maximum Live Telemetry Sessions

In the gv.cfg configuration file, it is possible to set a limit on the number of live telemetry sessions running in parallel using the field `max_live_sessions`.

```
[Telemetry]
# max parallel user live sessions
max_live_sessions=3
# UFM's provider of telemetry (counters). possible values:telemetry, ibpm
telemetry_provider=telemetry
```

Alternatively, you may run the command `ufm telemetry collector max-live-sessions`.

Configuring Log Rotation

This section enables setting up the log files rotate policy. By default, log rotation runs once a day by cron scheduler.

⚠ Alternatively, you can run the following commands:

- ufm logging files rotation criteria
- ufm logging files rotation max-num

```
[logrotate]
#max_files specifies the number of times to rotate a file before it is deleted (this definition will be applied to
#SM and SHARP Aggregation Manager logs, running in the scope of UFM).
#A count of 0 (zero) means no copies are retained. A count of 15 means fifteen copies are retained (default is 15)
max_files = 15
#With max_size, the log file is rotated when the specified size is reached (this definition will be applied to
#SM and SHARP Aggregation Manager logs, running in the scope of UFM). Size may be specified in bytes (default),
#kilobytes (for example: 100k), or megabytes (for exapmle: 10M). if not specified logs will be rotated once a day.
max_size = 3
```

Configuration Examples in gv.cfg

The following show examples of configuration settings in the gv.cfg file:

- Polling interval for Fabric Dashboard information

```
ui_polling_interval = 30
```

- [**Optional**] UFM Server local IP address resolution (by default, the UFM resolves the address by gethostip). UFM Web UI should have access to this address.

```
ws_address = <specific IP address>
```

- HTTP/HTTPS Port Configuration

```
# WebServices Protocol (http/https) and Port
ws_port = 8088
ws_protocol = http
```

- Connection (port and protocol) between the UFM server and the APACHE server

```
ws_protocol = <http or https>
ws_port = <port number>
```

For more information, see Launching a UFM Web UI Session.

- SNMP `get-community` string for switches (fabric wide or per switch)

```
# default snmp access point for all devices
[SNMP]
port = 161
gcommunity = public
```

- Configuring Monitoring History

```
[MonitoringHistory]
# history_report_timeout = 600
```

⚠ Other [Monitoringhistory] parameters are updated implicitly via the configuration script and should NOT be modified in the gv.cfg file directly.

- Enhanced Event Management (Alarmed Devices Group)

```
[Server]
auto_remove_from_alerted = yes
```

- Log verbosity:

```
[Logging]
# optional logging levels
#CRITICAL, ERROR, WARNING, INFO, DEBUG
level = INFO
```

For more information, see UFM Logs.

- Settings for saving port counters to a CSV file

```
[CSV]
write_interval = 60
ext_ports_only = no
```

For more information, see Saving the Port Counters to a CSV File.

- Max number of CSV files (UFM Advanced)

```
[CSV]
max_files = 1
```

For more information, see Saving Periodic Snapshots of the Fabric (Advanced License Only).

⚠ The access credentials that are defined in the following sections of the conf/gv.cfg file are used only for initialization:
- SSH_Server

- SSH_Switch
- TELNET
- IPMI
- SNMP
- MLNX_OS

To modify these access credentials, use the UFM Web UI. For more information, see Device Access.

- Configuring the UFM communication protocol with MLNX-OS switches. The available protocols are:
  - http (default protocol)
  - https (for secure communication)

For configuring the UFM communication protocol after fresh installation and prior to the first run, set the MLNX-OS protocol as shown below.

Example:

```
[MLNX_OS]
protocol = https
port = 443
user = admin
credentials = admin
```

Once UFM is started, all UFM communication with MLNX-OS switches will take place via the configured protocol.

For changing the UFM communication protocol while UFM is running, perform the following:

1. Set the desired protocol of MLNX-OS in the conf/gv.cfg file (as shown in the example above).
2. Restart UFM.
3. Update the MLNX-OS global access credentials configuration with the relevant protocol port. Refer to Device Access for help.
   For the http protocol - default port is 80.
   For the https protocol - default port is 443.
4. Update the MLNX-OS access credentials with the relevant port in all managed switches that have a valid IP address.

SM Trap Handler Configuration

The SMTrap handler is the SOAP server that handles traps coming from OpenSM.

There are two configuration values related to this service:

- `osm_traps_debounce_interval` - defines the period the service holds incoming traps
- `osm_traps_throttle_val` - once `osm_traps_debounce_interval` elapses, the service transfers `osm_traps_throttle_val` to the Model Main

> ⚠ By default SM Trap Handler handles up to 1000 SM traps every 10 seconds.

Configuring Two Appliances to Run in High Availability

Prerequisites for UFM High Availability (HA) Installation

- Reliable and high capacity out-of-band IP connectivity between the UFM Primary and Secondary UFM-SDN Appliance (1 Gb Ethernet is recommended). This connectivity is used for HA monitoring and replication of data and configuration.
- In fabrics consisting of multiple tiers of switches, it is recommended that the InfiniBand ports of the primary and secondary UFM-SDN Appliance be connected to different fabric switches on the same tier (the outermost edge in CLOS 5 designs).

To configure two appliances to run in High Availability (HA):

1. Stop the UFM server on both machines.

```
ufm-appliance > enable
ufm-appliance # configure terminal
ufm-appliance [ mgmt-sa ] (config) # no ufm start
```

2. Verify the UFM license is installed.

```
ufm-appliance [ mgmt-sa ] (config) # show ufm license
```

If the license is not installed, please do so. For further information, please refer to the section "Obtaining the License".

3. Configure both Appliances to run in HA configuration.

```
ufm-appliance [ mgmt-sa ] (config) # ufm ha configure <peer IP> <virtual IP> [interface]
```

⚠ You must run the "`ufm ha configure`" command from one of the machines, it is not required to run it on both appliances.

To return to a standalone configuration:

1. Stop the UFM server.

```
ufm-appliance > enable
ufm-appliance # configure terminal
ufm-appliance [ mgmt-ha-active ] (config) # no ufm start
```

2. Verify the DRBD disk state is up-to-date on both appliances.

```
ufm-appliance [ mgmt-ha-active ] (config) # show ufm status
```

⚠ If the status is not synchronized, wait for a couple of more minutes and then run the "`show ufm status`" command.

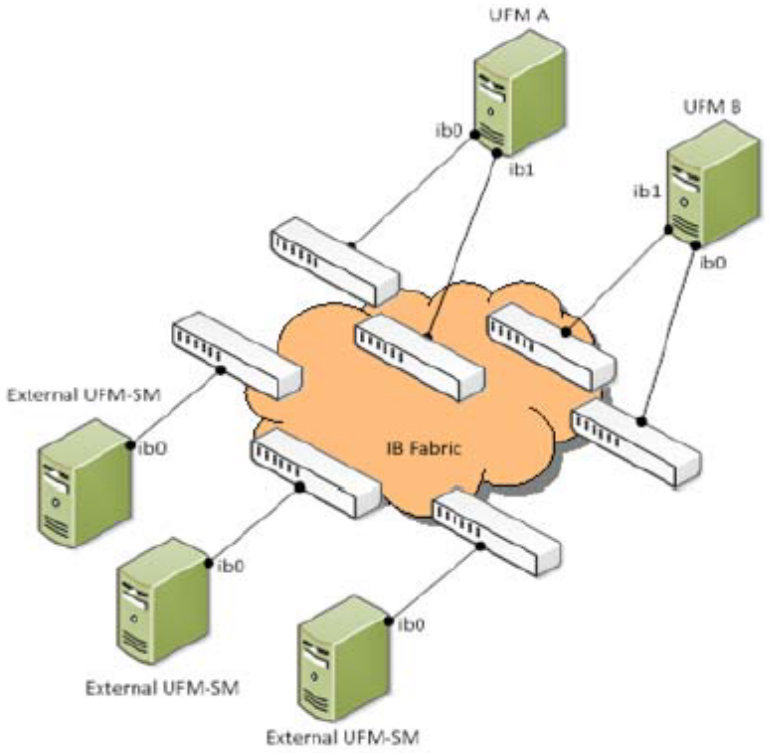
3. Disable High Availability configuration on both machines.

```
ufm-appliance [ mgmt-ha-active ] (config) # no ufm ha
```

Configuring UFM-SDN Appliance with External UFM-SMs

The main UFM®-SDN Appliance should be set to Standalone with external Sms or with High Availability external SMs (see “High Availability” on page 27).
External UFM SM requires additional UFM-SDN Appliance configured in sm-only mode.
Adding External SMs provides additional Subnet Managers for failover. Moreover, it enables UFM®-SDN Appliance upgrade without fabric downtime (In-service upgrade).
While the main UFM-SDN Appliance is running, it synchronizes the SMs configuration files on all the external SMs. If the main UFM-SDN Appliance fails (or stops for maintenance operations) an External SM takes mastership and manages the fabric until the main UFM-SDN Appliance resumes operation.

UFM HA with Additional External SMs



Configuring External UFM Subnet Manager

To configure the External UFM® Subnet Manager (SM):

1. Set the External SM machines mode to “sm-only”.

```
ufm-appliance [ mgmt-sa ] (config) # ufm mode sm-only
ufm-appliance [ sm-only ] (config) #
```

2. Set the Active machine mode to “mgmt-allow-sm” to enable External SM machines registration.

```
ufm-appliance [ mgmt-sa ] (config) # ufm mode mgmt-allow-sm
ufm-appliance [ mgmt-allow-sm-sa ] (config) #
```

3. Register the External SM machines.

```
ufm-appliance [ mgmt-allow-sm-sa ] (config) # ufm external-sm register <UFM-SM IP address>
ufm-appliance [ mgmt-allow-sm-sa ] (config) #
```

4. Verify the registration succeeded.

```
ufm-appliance [ mgmt-allow-sm-sa ] (config) # show ufm external-sm
10.30.30.101
ufm-appliance [ mgmt-allow-sm-sa ] (config) #
```

5. Create a trust relationship between the Standby and the External SM machines in the HA configuration.
   This step is performed on the Standby machine.

```
ufm-appliance [ ha-standby ] (config) # ufm external-sm trust <UFM-SM IP address>
ufm-appliance [ ha-standby ] (config) #
```

6. Start UFM on both UFM and UFM-SM machines.
7. Verify the SM configuration is synchronized in the External SM machines.

```
ufm-appliance [ mgmt-allow-sm-sa ] (config) # show ufm external-sm status
10.30.30.19 0 ok Synchronized on: Sun Mar 10 11:26:22 GMT-2 2013
10.30.30.20 0 ok Synchronized on: Sun Mar 10 11:26:23 GMT-2 2013
ufm-appliance [ mgmt-allow-sm-sa ] (config) #
```

   This step is performed on the Active machine and the synchronization might take up to a few minutes.

CPU Affinity on UFM-SDN Appliance

To execute major processes on the system (UFM, SM, SHARP, Telemetry), sub and multi-threaded processes are run on the system.

In order to increase UFM's efficiency, the number of context-switches is reduced. When each major CPU is isolated, users can decrease the number of context-switches, and the performance is optimized.

The CPU affinity of these major processes is configured in the following two levels:

- Level 1- The major processes initiation.
- Level 2- Preceding initiation of the model's main subprocesses.

According to user configuration, each process is assigned with affinity.

For using this feature,

1. Run:

```
[CPUAffinity]
Is_cpu_affinity_enabled=true
Model_main_cpu_affinity=1-4
Sm_cpu_affinity=5-13
SHARP_cpu_affinity=14-22
Telemetry_cpu_affinity=22-23
```

2. Check how many CPUs you have on your machine.
3. Enable the feature on gv.cfg.
4. Using the comments within the CPUAffinity section on gv.cfg define for each app the CPU set
5. Restart UFM.

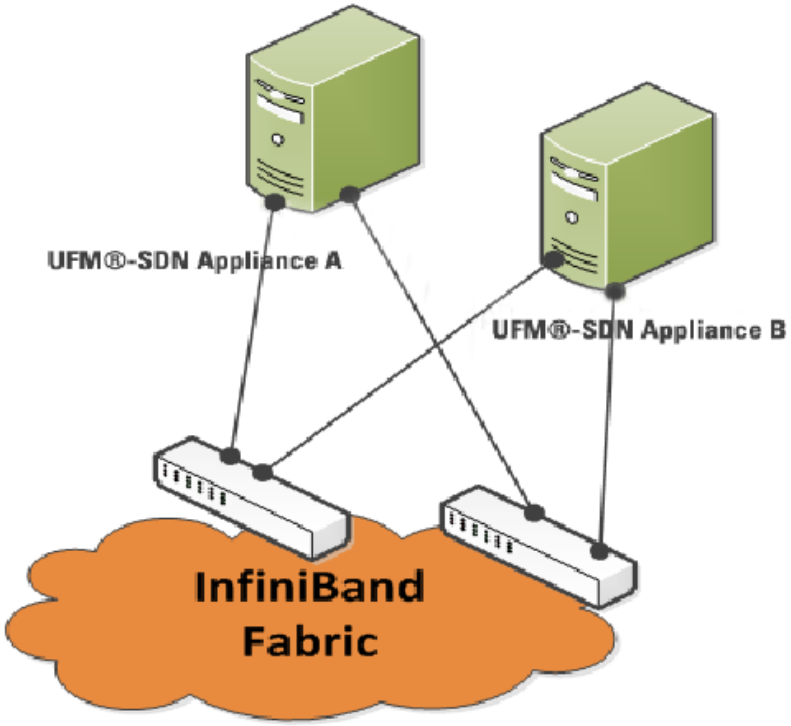
Quality of Service (QoS) Support

Quality of Service (QoS) is set as False by default.
To enable it and benefit from its capabilities, set it to True in the conf/opensm/opensm.conf file.

UFM Failover to Another Port

When failure is detected on an InfiniBand port or link, failover occurs without stopping the UFM®-SDN Appliance or other related UFM services, such as mysql, http, DRDB, and so on.
When the UFM®-SDN Appliance is connected by two or more InfiniBand ports to the fabric, you can configure UFM Subnet Manager failover to one of the other ports. When failure is detected on the InfiniBand port or link, failover occurs without stopping the UFM®-SDN Appliance or other related UFM services, such as mysql, http, DRDB, and so on. This failover process prevents failure in a standalone setup, and preempts failover in a High Availability setup, thereby saving downtime and recovery.

Network Configuration for Failover to IB Port



⚠ UFM SM failover is not relevant for Monitoring mode, because in this mode, UFM must be connected to the fabric over ib0 only.

When failure is detected on an InfiniBand port or link, UFM®-SDN Appliance initiates the give-up operation that is defined in the Health configuration file for OpenSM failure. By default:

- UFM-SDN Appliance discovers the other ports in the specified bond and fails over to the first interface that is up (SM failover)
- If no interface is up:
  - In an HA setup, UFM initiates UFM failover.

- In a standalone setup, UFM®-SDN Appliance does nothing.

If the failed link becomes active again, UFM®-SDN Appliance will select this link for the SM only after SM restart.

Starting UFM

1. Log into the CLI as admin and use the selected password.
2. Start UFM.

```
ufm-appliance > enable
ufm-appliance # configure terminal
ufm-appliance [ mgmt-sa ] (config) # ufm start
```

⚠ In order to view the status of UFM services, run the command “show ufm status”.

User Authentication

UFM User Authentication is based on standard Apache User Authentication. Each Web Service client application must authenticate against the UFM-SDN Appliance to gain access to the system.

The UFM software comes with one predefined user:

- Username: admin
- Password: admin

You can change the definition, add, or delete users.

Licensing

UFM license is subscription-based featuring the following subscription options:

- 1-year subscription
- 3-year subscription
- 5-year subscription
- Evaluation 30-day trial license

> ⚠ UFM will continue to support old license types, but they are no longer available to obtain.

2 months before the expiration of your subscription license, UFM will warn you that your license will expire soon. After the subscription expires, UFM will continue to work with the expired license for two months beyond its expiration.

During this extra two month period, UFM will generate a critical alarm indicating that UFM license has expired and that you need to renew your subscription. Failing to do so within that 2-month period activates UFM Limited Mode. Limited mode blocks all REST APIs as well as access to the UFM web UI.

UFM enables functionality based on the license that was purchased and installed. This license determines the functionality and the maximum allowed number of nodes in the fabric.

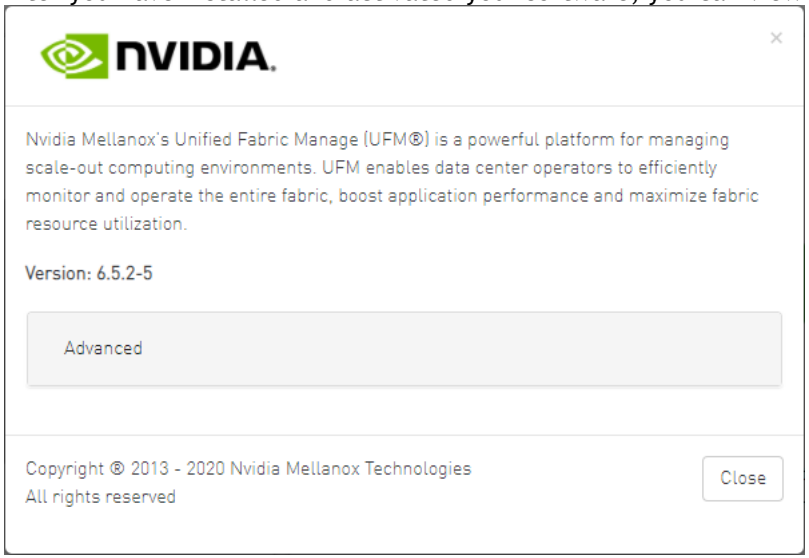
To renew your UFM subscription, purchase a new license and install the new license file by downloading the license file to a temp directory on the UFM master server and then copying the license file to /opt/ufm/files/licenses/ directory.

> ⚠ UFM may not detect new license files if downloaded directly to /opt/ufm/files/licenses. If UFM does not detect the new license file, a UFM restart may be required.

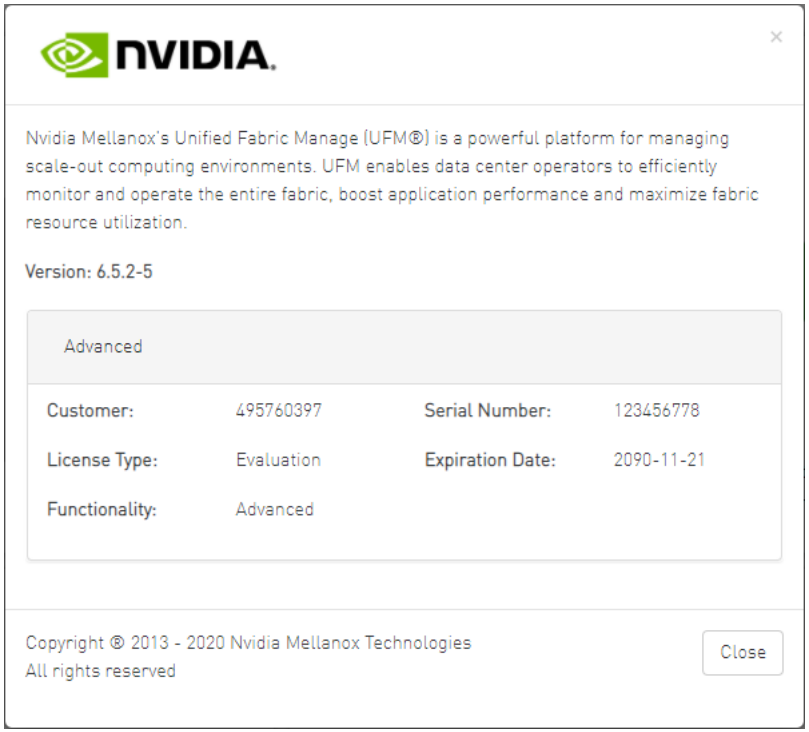
If several licenses are installed on the server (more than one license file exists under /opt/ufm/files/licenses/), UFM uses only the strongest license and takes into consideration the expiration date, and the managed device limits on it, regardless of any other licenses that may exist on the server.

To view license information from the Web UI:

1. After you have installed and activated your software, you can view your licenses in the Web UI by clicking the About icon ( ) in the main page.



2. The main About screen displays the current UFM version and build. To view the advanced license information, click the Advanced button. Advanced License Information includes:



- **Customer**: Customer number provided by Mellanox
- **Serial Number**: UFM serial number provided by Mellanox
- **License Type**: Permanent or evaluation
- **Expiration Date**: License expiration date (limited), unlimited License (nothing is written)
- **Functionality**: UFM Advanced

To view license information from the CLI:

Run CLI Command "**ufmlicense**" to display information about all installed licenses on the UFM server under /opt/ufm/files/licenses. This includes invalid and expired license information.

To remove a license:

Delete the license file from /opt/ufm/files/licenses.

Launching a UFM Web UI Session

Before accessing the UFM Web UI:

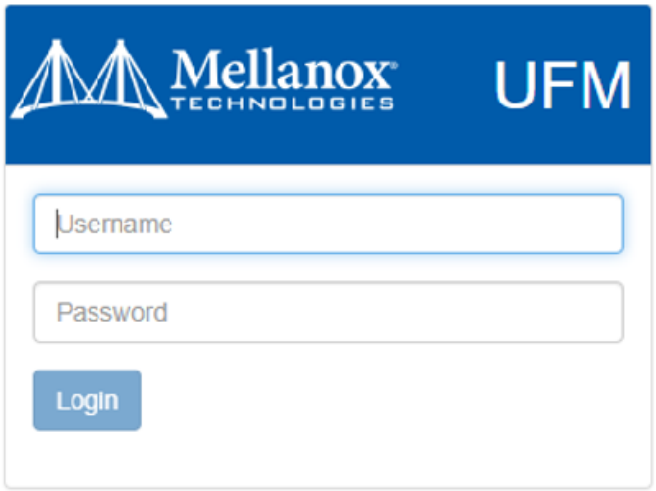
- If required, you can change the configuration of the connection (port and protocol) between the UFM and the APACHE server in the file Change to conf/gv.cfg.

```
ws_protocol = http or https
ws_port = port number
```

- To receive the configured IP in case you configured DHCP, run from the UFM®-SDN Appliance CLI the `show interfaces eth0` command.

To launch a UFM Web UI session

1. Launch the Web UI by entering the following URL in your browser:
   http://<UFM_appliance_IP> or https://<UFM_appliance_IP>



2. In the Login page, enter your User Name and your predefined user Password and click Login. Once you have entered your user name and password, the main page opens, showing the UFM Dashboard.

Installing UFM-SDN Appliance License

1. Log into the CLI as admin and use the selected password.
2. Install the license.

```
ufm-appliance [ mgmt-sa ] (config) # ufm license install ftp://admin@192.168.11.1/home/admin/licenses/volt-ufm-advanced.lic
```

Setting Up Telemetry in UFM-SDN Appliance

⚠ The steps described in this page are not mandatory.

Setting up telemetry deploys UFM Telemetry and UFM Telemetry Collector for collecting fabric port statistics and streaming them via the telemetry collector into the SQLite database in UFM. Once data is available on the database, users are able to generate reports and view statistical data using the UFM Web UI.

To enable UFM telemetry:

1. Pull the ufm-telemetry container from the UFM Telemetry docker hub repository. For more information, please refer to steps 1-4 under appendix UFM Telemetry Docker Container.
2. Pull the telemetry-collector container from docker hub. Run:

   ```
   ufmapl [ mgmt-ha-active ] (config) # docker pull mellanox/telemetry-collector
   ```

   ⚠ The previous two steps require an internet connection.

3. Enable the telemetry-collector. Run:

   ```
   ufmapl [ mgmt-ha-active ] (config) # ufm telemetry collector enable
   ```

4. Enable the ufm-telemetry run locally. Run:

   ```
   ufmapl [ mgmt-ha-active ] (config) # telemetry streaming enable
   ```

Installing and Running UFM-SDN Agent Software

The UFM agent is an optional component that you can install on the fabric nodes to collect local node information and perform local management tasks.

UFM Agent Installation Prerequisites

Before you install the UFM agent, ensure that the following packages are installed on your system:

- net-tools
- coreutils
- procps
- sysstat
- sed
- util-linux
- grep
- gawk
- OpenIPMI-tools
- Iproute
- ip-bonding

Most of these packages are installed by the default LINUX installation. The UFM agent installation script terminates the installation if packages are missing. However, package names might vary between Linux distributions.

> ⚠ TCP/UDP port 6306 is used for communication with the UFM server. This value might be overwritten (see installation option *ufma-ufm-mcast-port*).

Installing UFM Agent Software

The default UFM installation directory is */opt/ufm*.

To install the UFM Agent software, do the following:

1. Create a temporary directory (for example */tmp/ufm*).
2. Extract *ufma-3.0.0-XXX-<OS>-x86_64.tgz* to the temporary directory (where <OS> is *redhat* or *suse*).
3. Change directory to */tmp/ufm/ufma-3.0.0-XXX-<OS>-x86_64*.

4. From the temporary directory, run the following command as root: ./install.sh --ufma-mcast-ifaces *"ethxx,ibxx"*. You can use command options to override default installation parameters.
5. After installation you can remove the temporary directory.

UFMA Installation Parameters

General Installation Parameters

Parameter	Default	Notes
--autostart	y	Override this if you do not want to start the UFMA process automatically.
--start-ipmi	y	Override this option if you do not use IPMI in your fabric.
--scriptsonly	y	Use this option when you install UFMA on the UFM server.

UFMA Parameters

Parameter	Default	Notes
--ufma-cpu-affinity	None	CPU core number, starting from zero.
--ufma-ufm-mcast-addr	224.0.23.172	This address MUST be configured in UFM.
--ufma-ufm-mcast-port	6306	This port MUST be configured in UFM.
--ufma-mcast-ifaces	N/A, mandatory parameter	You can bind UFMA listener to specific interfaces; for example, eth0, ib3.
--ufma-dir	/opt/ufma	Use only if /opt/ is mounted as read-only.

Running UFM Agent Software

After installing the UFM Agent software, the UFM Agent runs automatically. The UFM Agent reports messages to */var/log/ufma.log*.

- To check the status, invoke:

```
/etc/init.d/ufmad status
```

- To set the verbosity of the log level, use the UFMA_LOG_LEVEL environment variable. Valid values are from 0 (low) to 6 (high).
  To activate the new setting, restart the agent using:

```
/etc/init.d/ufmad restart
```

  For troubleshooting, send the log files to Mellanox support.

InfiniBand Bonding on Managed Host

Bonding between two ports of the same HCA is fully supported when the UFM Agent is running on the host. The UFM Agent learns the bond interface configuration that was created manually, and can then configure partitioning, meaning a p-key-based interface, on top of the bond interface.

When the UFM Agent is not installed, UFM learns and presents bond slave interfaces as separate interfaces, i.e. bonding is not indicated.

UFM SNMP Agent

UFM supports system MIBs, and answers SNMP queries for the following parameters:

Parameter	Value
sysDescr	UFM Server 5.2

Parameter	Value
sysObjectID	1.3.6.1.4.1.5206.1.200
sysContact	www.mellanox.com

Default get community string is "public", default set community string is "private".

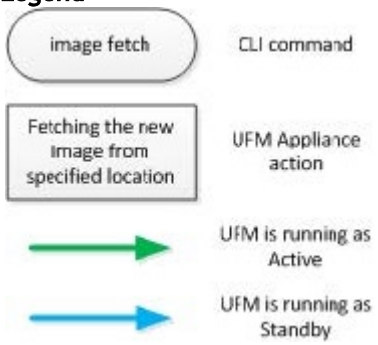
UFM Agent Upgrade

The UFM Agent is a stateless application. There is no dedicated upgrade procedure. The Agent installation process detects whether there is a previous version of the agent and automatically removes it.

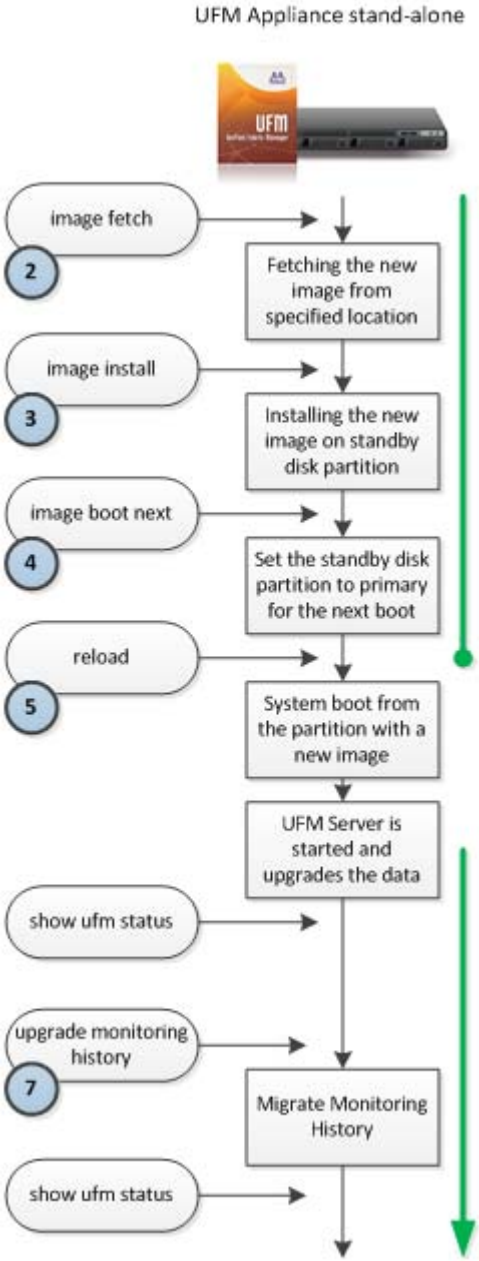
Upgrading UFM-SDN Appliance Software

The following figures describe various upgrade scenarios.

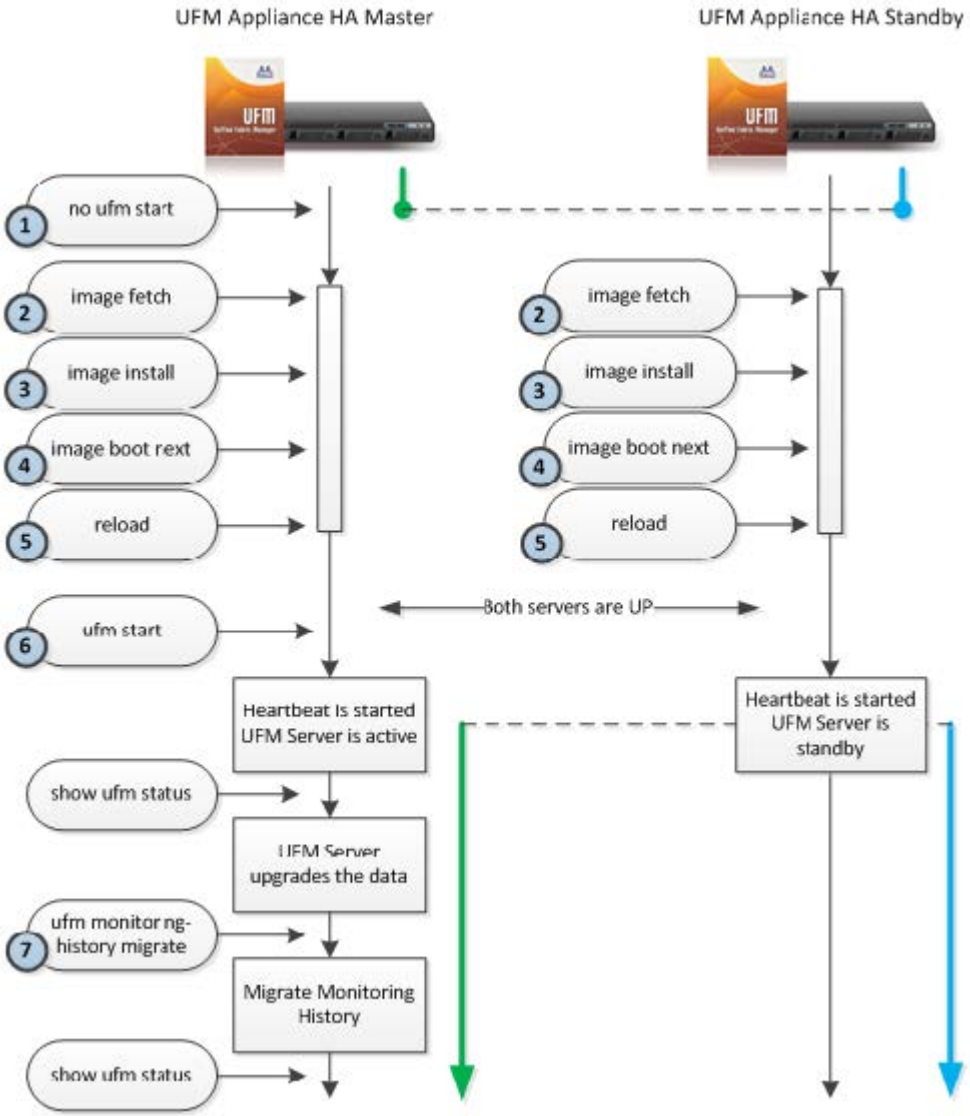
Legend



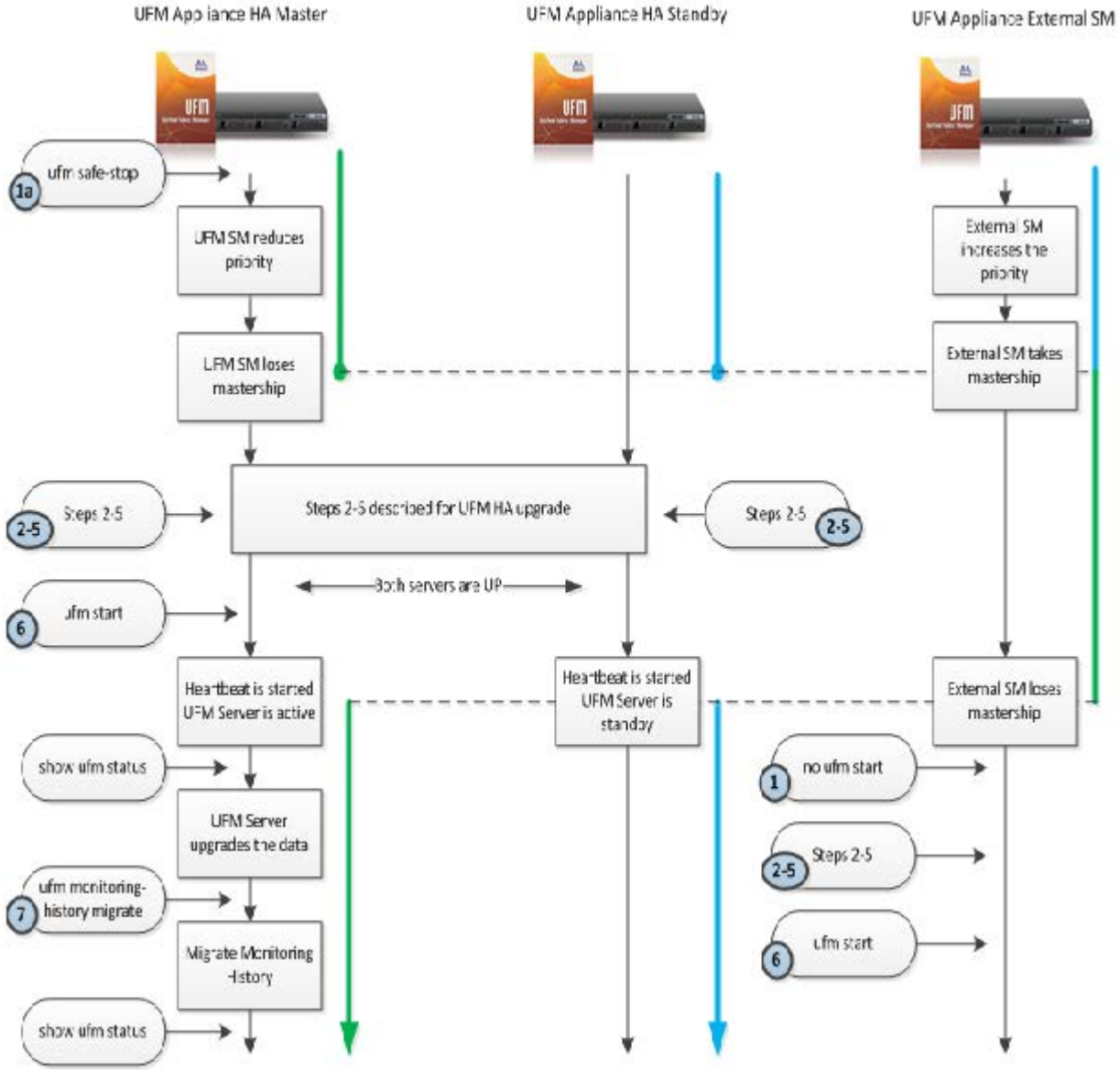
UFM-SDN Appliance Standalone Upgrade Flow


UFM Appliance stand-alone
image fetch
2
Fetching the new image from specified location
image install
3
Installing the new image on standby disk partition
image boot next
4
Set the standby disk partition to primary for the next boot
reload
5
System boot from the partition with a new image
UFM Server is started and upgrades the data
show ufm status
upgrade monitoring history
7
Migrate Monitoring History
show ufm status

UFM-SDN Appliance High Availability Out-Of-Service Upgrade Flow


UFM Appliance HA Master
UFM Appliance HA Standby
1
no ufm start
2
image fetch
3
image install
4
image boot next
5
reload
2
image fetch
3
image install
4
image boot next
5
reload
Both servers are UP
6
ufm start
Heartbeat Is started
UFM Server is active
Heartbeat is started
UFM Server is
standby
show ufm status
UFM Server
upgrades the data
7
ufm monitor ng-
history migrate
Migrate Monitoring
History
show ufm status

UFM-SDN Appliance High Availability External UFM-SM In-Service Upgrade Flow


UFM Appliance HA Master
UFM Appliance HA Standby
UFM Appliance External SM
ufm safe-stop
1a
UFM SM reduces priority
UFM SM loses mastership
External SM increases the priority
External SM takes mastership
Steps 2-5
2-5
Steps 2-5 described for UFM HA upgrade
Steps 2-5
2-5
Both servers are UP
ufm start
6
Heartbeat is started UFM Server is active
Heartbeat is started UFM Server is standby
External SM loses mastership
show ufm status
UFM Server upgrades the data
no ufm start
1
Steps 2-5
2-5
ufm monitoring-history migrate
7
Migrate Monitoring History
ufm start
6
show ufm status

To upgrade UFM-SDN Appliance software on your system:

Obtain the available image (.img file) and place it on either a USB storage device or an ftp/scp server and follow the steps below.

High Availability Out-Of-Service Upgrade

1. Stop UFM.

```
ufm-appliance [ mgmt-ha-active ] (config) # no ufm start
```

   Alternatively, if external UFM-SM is available, run:

```
ufm-appliance [ mgmt-ha-active ] (config) # ufm safe-stop
```

2. Fetch the new software image.

```
ufm-appliance [ mgmt-ha-active ] (config) # image fetch scp://<username>@192.168.10.125/var/www/html/
<image_name>
Password ******
100.0%[##################################################################]
```

3. Install the new image.

```
ufm-appliance [ mgmt-ha-active ] (config) # image install image-ufm_appliance-x86_64-
x86_64-20121205-152024.img
Step 1 of 4: Verify Image
100.0%
[##################################################################]
Step 2 of 4: Uncompress Image
100.0%
[##################################################################]
Step 3 of 4: Create Filesystems
100.0%
```

```
[##################################################################]
Step 4 of 4: Extract Image
100.0%
[##################################################################]
```

4. Set next boot location as the other slot.

```
ufm-appliance [ mgmt-ha-active ] (config) # image boot next
```

5. Save configuration. Run:

```
ufm-appliance [ mgmt-ha-active ] (config) # configuration write
```

6. Reload UFM-SDN Appliance.

```
ufm-appliance [ mgmt-ha-active ] (config) # reload
```

7. Start UFM.

```
ufm-appliance [ mgmt-ha-active ] (config) # ufm start
```

High Availability In-Service Upgrade

1. Run on both master and standby:
    a. Fetch the new software image. Run:

```
ufm-appliance [ mgmt-ha-active ] (config) #  image fetch <download URL>
```

    b. Install the new image. Run:

```
ufm-appliance [ mgmt-ha-active ] (config) #  image install <image name>
```

c. Set next boot location as the other slot. Run:

```
ufm-appliance [ mgmt-ha-active ] (config) #  image boot next
```

d. Save the configuration changes. Run:

```
ufm-appliance [ mgmt-ha-active ] (config) #  configuration write
```

2. Reload the standby UFM-SDN Appliance. Run:

```
ufm-appliance [ ha-standby ] (config) # reload
```

3. Once the standby is up and running, reload the master UFM-SDN Appliance. Run:

```
ufm-appliance [ mgmt-ha-active ] (config) # reload
```

As a result, a failover will occur from the master node to the standby node.

UFM-SDN Appliance Common Use Cases

Restoring UFM Application Data on the UFM®-SDN Appliance

Box configuration (e.g. interfaces' IP addresses, users, AAA, etc…) should be backed up and restored separately via CLI. For further information, please refer to the UFM®-SDN Command Reference Guide, section System Management.

The following are different UFM data restoration scenarios:

Software Version	Backed Up Data Version	Action Required
1.6	1.6	Restore the backed up data
1.5	1.6	1. Upgrade the software 2. Restore the backed up data
1.6	1.5	Contact NVIDIA Mellanox Support

Restoring Data on Standalone Configuration

1. Stop the UFM server.

```
ufm-appliance > enable
ufm-appliance # configure terminal
ufm-appliance [ mgmt-sa ] (config) # no ufm start
```

2. Restore the backed up UFM data.

```
ufm-appliance [ mgmt-sa ] (config) # ufm configuration import <zip-file>
ufm-appliance [ mgmt-sa ] (config) # ufm database import <zip-file>
```

3. Start the UFM server.

Restoring Data on High Availability Configuration

To restore the data on the High Availability configuration when only the Standby machine has been replaced:

1. Stop the UFM server.

```
ufm-appliance [ mgmt-ha-active ] > enable
ufm-appliance [ mgmt-ha-active ] # configure terminal
ufm-appliance [ mgmt-ha-active ] (config) # no ufm start
```

2. Re-configure both Appliances to run in HA configuration. This command is performed on the Active machine.

```
ufm-appliance [ mgmt-ha-active ] (config) # ufm ha configure <peer-IP> <virtual-IP> [interface]
```

   As a result, the UFM data is synchronized by the DRBD.
3. Start the UFM server.

To restore the data on the High Availability configuration when both the Standby and the Active machine have been replaced:

1. Restore the backed up UFM data on one of the machines.

```
ufm-appliance [ mgmt-sa ] (config) # ufm configuration import <zip-file>
ufm-appliance [ mgmt-sa ] (config) # ufm database import <zip-file>
```

2. Re-configure both Appliances to run in HA configuration. This command is performed on the Active machine.

```
ufm-appliance [ mgmt-sa ] (config) # ufm ha configure <peer-IP> <virtual-IP> [interface]
```

   As a result, the UFM data is synchronized on the Standby machine by the DRBD.
3. Start the UFM server.

Backing Up the UFM Database

> ⚠ The following process backs up the UFM database, and not the Monitoring History database.

To back up the UFM database:

1. Export the database to a "zip" file.

```
ufm-appliance > enable
ufm-appliance # configure terminal
ufm-appliance [ mgmt-sa ] (config) # ufm database export
```

2. Upload the file to either a USB storage device or an ftp/scp server.

```
ufm-appliance [ mgmt-sa ] (config) # ufm database upload ufm-database-180857.zip scp://
mlnx:123456@172.30.3.201/tmp
```

Backing Up the UFM Configuration

1. Export the configuration to a "zip" file.

```
ufm-appliance > enable
ufm-appliance # configure terminal
ufm-appliance [ mgmt-sa ] (config) # ufm configuration export
```

2. Upload the file to either a USB storage device or an ftp/scp server.

```
ufm-appliance [ mgmt-sa ] (config) # ufm configuration upload ufm-config-20121128-180857.zip scp://
mlnx:123456@172.30.3.201/tmp
```

Generating a Debug Dump File

1. Generate a debug dump file.

```
ufm-appliance > enable
ufm-appliance # configure terminal
ufm-appliance [ mgmt-sa ] (config) # debug generate dump
```

2. Upload the file to either a USB storage device or an ftp/scp server.

```
ufm-appliance [ mgmt-sa ] (config) # file debug-dump upload sysdump-ufm-appliance-8990b8-20120710-102624.tgz
ftp://admin@172.30.3.36/home/admin
```

Creating Users

UFM®-SDN Appliance supports local and centrally managed users (Using LDAP, RADIUS, TACACS).

UFM®-SDN Appliance has predefined users. For the complete list of users and their roles, please refer to the UFM®-SDN Command Reference Guide, section User Roles (Capabilities).

Creating Local User

1. Log into the CLI as admin and use the selected password.
2. Create the users with a specific capability (role).

```
ufm-appliance > enable
ufm-appliance # configure terminal
ufm-appliance [ mgmt-sa ] (config) # username <username> capability <capability>
```

3. Set the user's password.

```
ufm-appliance [ mgmt-sa ] (config) # username <username> password <password>
```

Linking Users to LDAP Server

1. Create a group for UFM®-SDN Appliance users on the ActiveDirectory server.
2. Assign users to the group.
3. [Optional] Add the attribute "localUserName" to the user schema for correlating it to a local username on the UFM®-SDN Appliance (If this attribute is not added, all users will be correlated with the default user - "admin").
4. Add LDAP as the authentication method.

```
ufm-appliance [ mgmt-sa ] (config) # aaa authentication login default local ldap
```

5. Configure the LDAP server.

```
ufm-appliance [ mgmt-sa ] (config) # ldap base-dn <string>
ufm-appliance [ mgmt-sa ] (config) # ldap bind-dn <string>
ufm-appliance [ mgmt-sa ] (config) # ldap bind-password <string>
ufm-appliance [ mgmt-sa ] (config) # ldap host <IP Address>
```

For further information please refer to section "User Management and AAA" of the *UFM®-SDN Command Reference Guide*.

Updating HCA Firmware

When logging into the CLI, and a similar upgrade notification as the below is displayed, you need to update your firmware version.

```
$ ssh admin@172.30.30.17
Password:
Last login: Sun Feb 24 17:27:59 2013 from 172.30.3.201
Mellanox UFM Appliance
UFM is configured as standalone.
UFM mode: Management.
```

```
HCA Firmware: 2.10.600 (Requires update)
```

To update the HCA firmware version:

1. Install the updated HCA firmware supplied in the software image.

   ```
   ufm-appliance [ mgmt-sa ] (config) # firmware install
   ```

2. Reboot the system.
3. Verify the firmware was updated correctly.

   ```
   ufm-appliance [ mgmt-sa ] (config) # show firmware
   Installed firmware version:              2.11.500
   Running firmware version:                2.11.500
   Image available for installation:        2.11.500
   ```

Recovering Primary and Secondary Servers Failure in Standby-Standby Mode

1. Generate debug dump from both UFM Appliances.

   ```
   ufm-appliance [ mgmt-sa ] (config) # debug generate dump
   ufm-appliance [ mgmt-sa ] (config) # file debug-dump upload sysdump-ufm-appliance-8990b8-20120710-102624.tgz
   ftp://admin@172.30.3.36/home/admin
   ```

2. Investigate the root cause of the issue.
   If the root cause related to the UFM License of UFM Configuration you need to perform the following steps:
   a. Look into the debug dump file to recognize the initial primary server (ufm-sysdump-<datetime>.tar.gz ./ufm_sys_conf/haresources).

      ```
      ufm-appliance-8a447a hb_track_start check_sm 172.30.46.100/16/eth0/172.30.255.255 drbddisk::ufmdb
      Filesystem::/dev/drbd0::/opt/ufm/files::ext3 ufmd::hamode fix_arp hb_track_end
      ```

   b. Disable the UFM High Availability configuration on both appliances.

```
ufm-appliance [ mgmt-sa ] (config) # no ufm ha
```

c. Update the UFM configuration or install the UFM License on the Primary Appliance.
d. Configure HA on the Primary Appliance.

```
ufm-appliance [ mgmt-sa ] (config) # ufm ha configure <peer IP> <virtual IP> [interface]
```

e. Wait until the mode is changed to "active" on the Primary Appliance and "standby" on the Secondary Appliance.
f. Verify the DRBD synchronization is completed.

```
ufm-appliance [ mgmt-sa ] (config) # show ufm status
```

g. Start UFM on the Primary Appliance.

```
ufm-appliance [ mgmt-sa ] (config) # ufm start
```

Getting Familiar with UFM's Data Model

Overview of Data Model

UFM enables the fabric administrator to manage the fabric based on discovery data collected from the fabric. This data is mapped into model elements (objects) available to the end user via UFM REST API and UFM Web UI.

UFM Model Basics

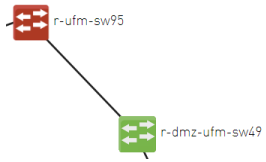
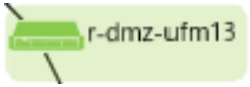
The fabric managed by UFM consists of a set of physical and logical objects, including their connections. The Object Model has a hierarchical object-oriented tree structure with objects as the tree elements. Each object defines an abstraction for physical or logical fabric elements.

Physical Model

The Physical Model represents the physical resources and connectivity topology of the Network. UFM enables discovery, monitoring and configuration of the managed physical objects.

Physical Objects

Icon	Name	Description
N/A	Port Object	Represents the external physical port on switch or on Host Channel Adapter (HCA). A port is identified by its number. UFM provides InfiniBand standard management and monitoring capabilities on the port level.
N/A	Module Object	Represents the Field Removable Unit, Line card, and Network card on switch or HCA on host. For NVIDIA Switches, Line and Network Cards are modeled as modules.

Icon	Name	Description
	Link Object	Represents the physical connection between two active ports.
N/A	Cable Object	Represents the physical cable or the transceiver connected to one of the link edges.
	Computer Object	Represents the computer (host) connected to the Fabric. The UFM Agent installed on the host provides extended monitoring and management capabilities. Hosts without agents are limited to InfiniBand standard management and monitoring capabilities.
	Switch Object	Represents the switch chassis in the Fabric. A Switch object is created for every NVIDIA Switch. Switches of other vendors are represented as InfiniBand Switches and limited by InfiniBand standard management and monitoring capabilities.
	Rack Object	Represents the arbitrary group of switches or computers. When linked devices are shown as a group, the link is shown between the group and the peer object.

UFM Web UI

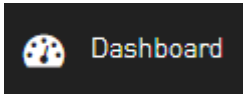
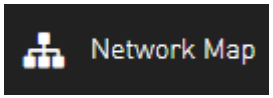
This section is constituted by the following sub-sections:

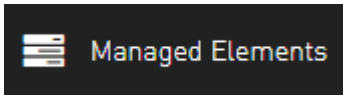
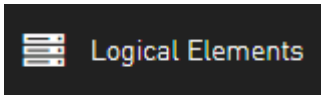
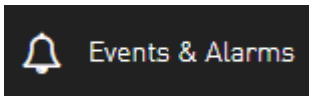
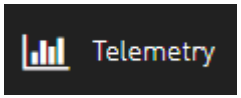
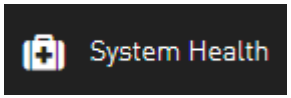
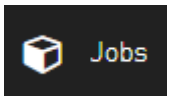
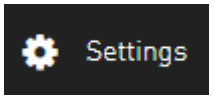
- UFM Web UI Main Navigation Buttons
- Fabric Dashboard
- Network Map
- Managed Elements
- Events & Alarms
- Telemetry
- System Health
- Jobs
- Settings

UFM Web UI Main Navigation Buttons

UFM software consists of several main web UI windows, accessible from a sidebar menu on the left side of the screen.

Navigator Tabs

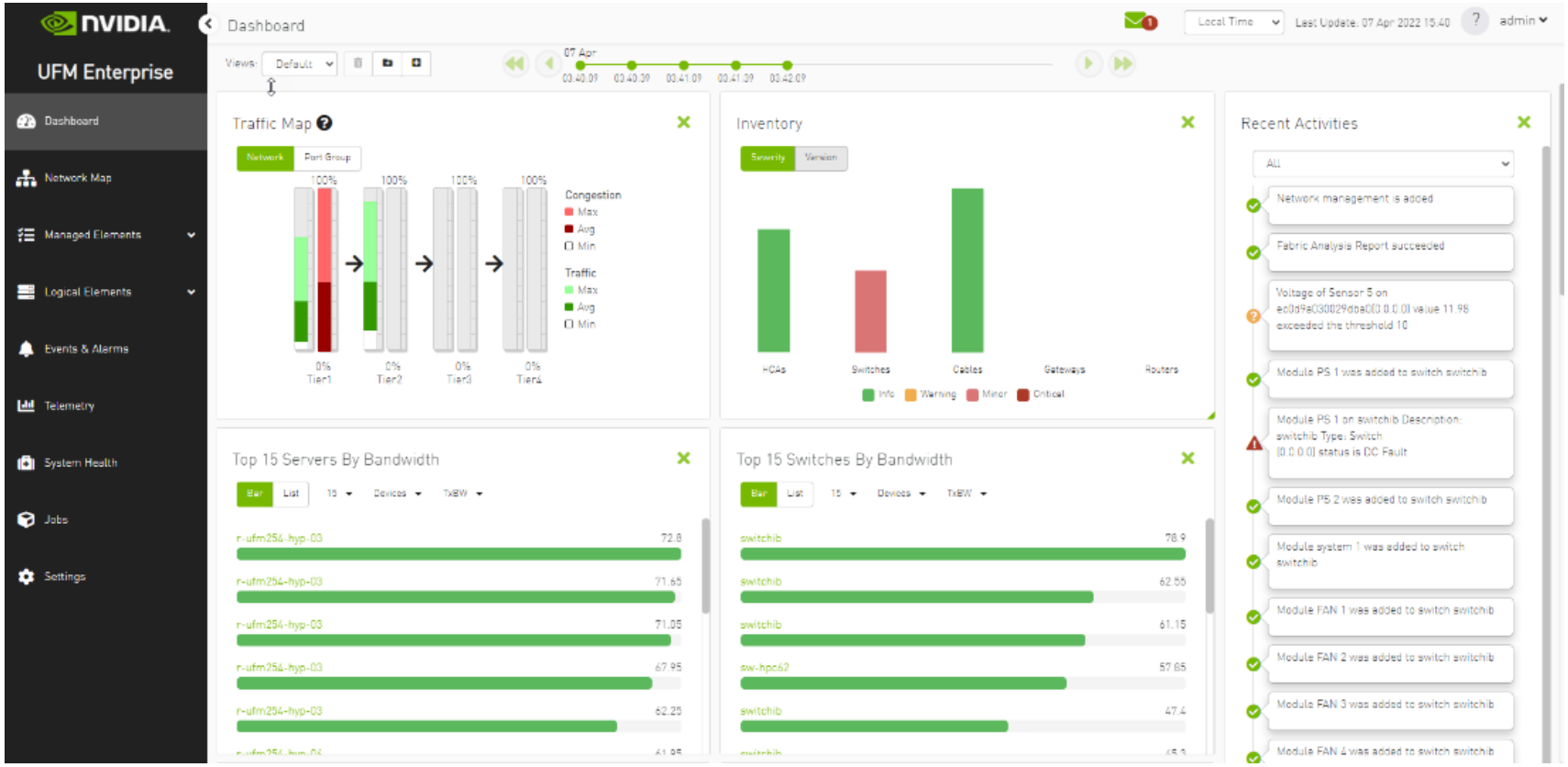
Tab Icon	Description
	Provides a summary view of the fabric status.
	Provides a hierarchical topology view of the fabric.

Tab Icon	Description
	Provides information on all fabric devices. This information is presented in a table format.
	Provides information on all logical servers. This information is presented in a table format.
	Provides information on the events & alarms generated by the system.
	Enables establishing monitoring sessions on devices or ports.
	Enables running and viewing fabric reports, UFM reports, and system logs. You can also back up UFM configuration files.
	Provides information on all jobs created, as a result of UFM actions.
	Enables configuring UFM server and UFM fabric settings, including events policy, device access, network management, subnet manager, and user management

Fabric Dashboard

The dashboard window summarizes the fabric's status, including events, alarms, errors, traffic and statistics.

Fabric Dashboard View

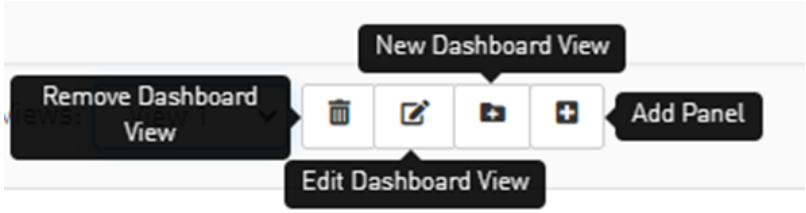


The Fabric Dashboard view consists of the following six dashboards, which provide real-time information about the fabric.

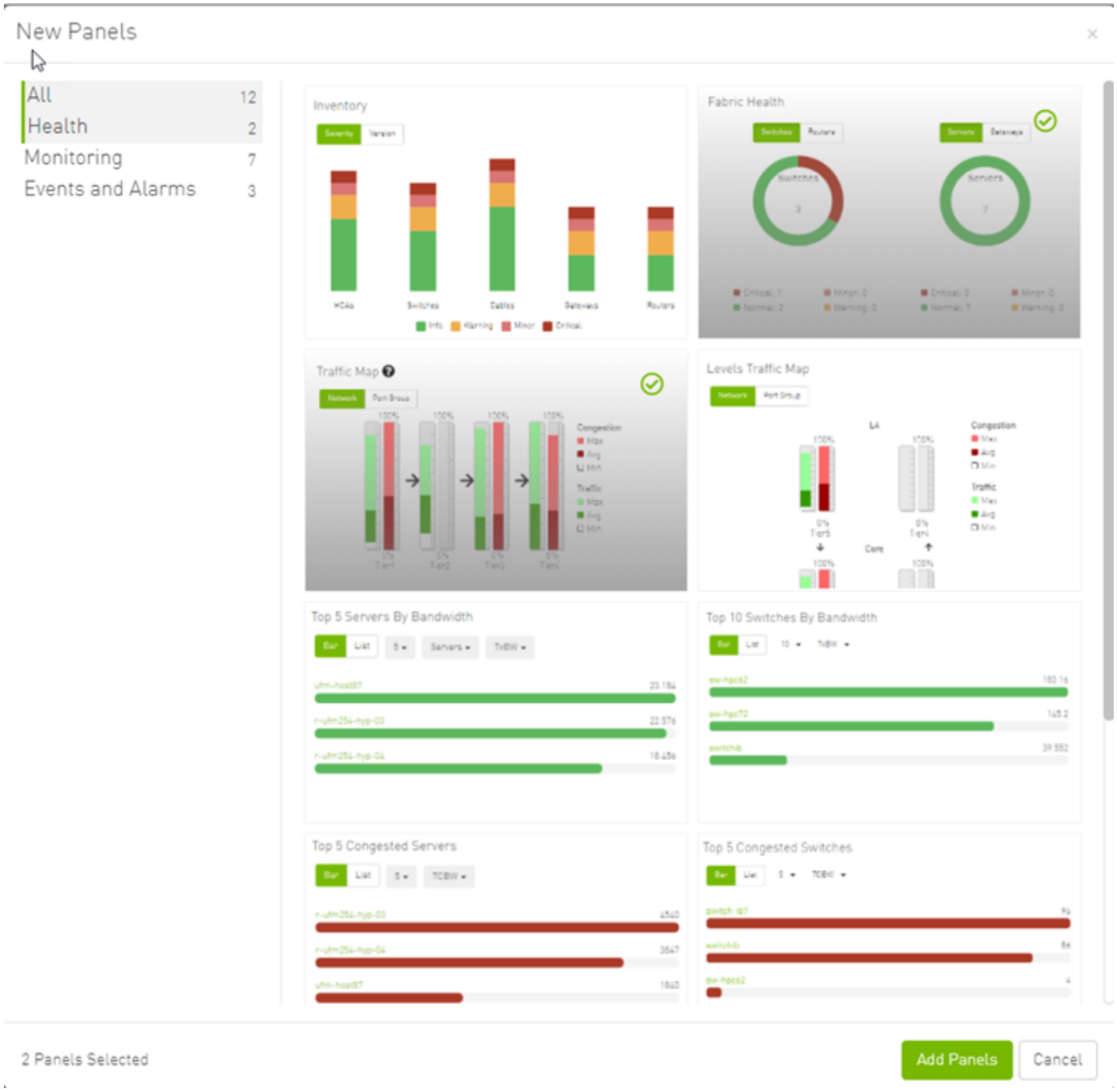
Dashboard Views and Panel Management

UFM is installed with a default view of the most important panels. These panels are resizable and draggable. Users can customize their default view or create new views altogether

The dashboard views and panels are managed by a set of action buttons appearing at the top of the main dashboard screen:



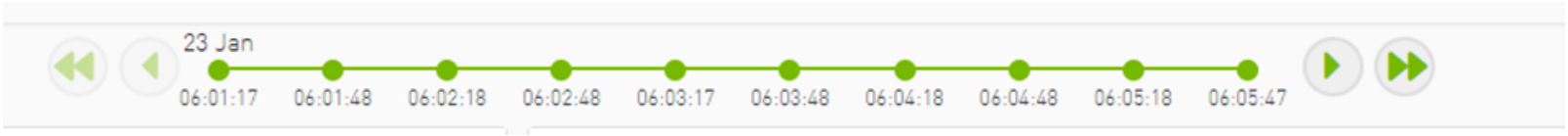
Clicking on the Add Panel button will show a model to select which panels you wish to add to the current dashboard view.


New Panels
All 12
Health 2
Monitoring 7
Events and Alarms 3
Inventory
Fabric Health
Traffic Map
Levels Traffic Map
Top 5 Servers By Bandwidth
Top 10 Switches By Bandwidth
Top 5 Congested Servers
Top 5 Congested Switches
2 Panels Selected
Add Panels
Cancel

Dashboard Timeline Snapshots

Once the user is logged into the UFM Enterprise, the UFM will start recording snapshots of the dashboard panel data every 30 seconds.

The user is able to navigate between these snapshots and load the dashboard data of a specific data snapshot.



Dashboard Panels

The Fabric Dashboard view consists of the following 12 panels, which are categorized into 3 main categories and provide real-time information about the fabric.

- Health:
  - Inventory
  - Fabric Health
- Monitoring:
  - Traffic Map
  - Levels Traffic Map
  - Top X Servers by bandwidth
  - Top X Switches by bandwidth
  - Top X congested servers
  - Top X congested switches
  - Top X utilized Pkeys
- Events and Alarms:
  - Recent Activities
  - Top X alarmed servers
  - Top X alarmed switches
  - Events History

Top N Servers/Switches by Rx or Tx Bandwidth

The Top N servers/switches by Rx or Tx Bandwidth component shows the top elements that are transmitting or receiving the most bandwidth per second. These elements are classified top-down according the defined Transmit (Tx) or Receive (Rx) bandwidth (MB/sec Rate).

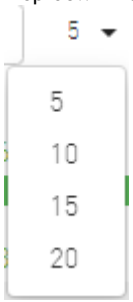
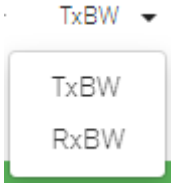
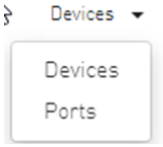
Bandwidth is measured as a rate in bytes/sec.

- Transmitted (Tx) bandwidth is measured by N server/switch ports in MB/sec
- Received (Rx) bandwidth is measured by N server/switch ports in MB/sec

> ⚠ N can be 5, 10, 15, or 20.

The following table lists the icons of this component:

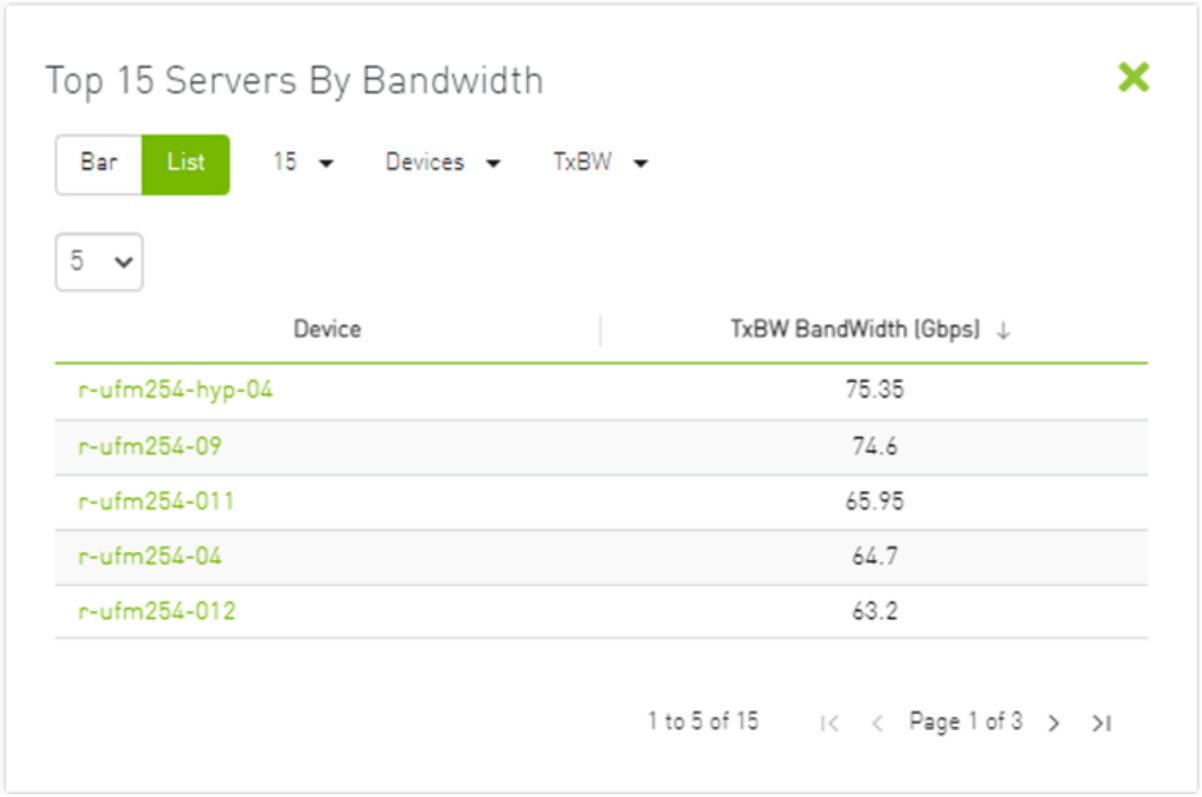
Options	Description
List view 	Shows the top N elements as a list Each element is shown in a row with the name of the element and the bandwidth rate
Bar view 	Shows the top N nodes as a bar graph • X axis shows the rate as a value • Y axis shows the Node (server) name

Options	Description
Drop-down menu 	Selects the number of items to display Default: 10 nodes
Monitoring attributes 	Selects the attribute for monitoring: • TxBW – Transmit Bandwidth • RxBW – Receive Bandwidth
View by port/element 	Switches view to top 5 elements by bandwidth or top 5 ports by bandwidth. Nodes view is presented by default. • Clicking a specific port in the ports view under the port column redirects to the ports table and highlights that particular port • Clicking a specific device in the devices view under the device column redirects to the Devices table and highlights that particular node
Filter toggle	Toggles the filter textbox

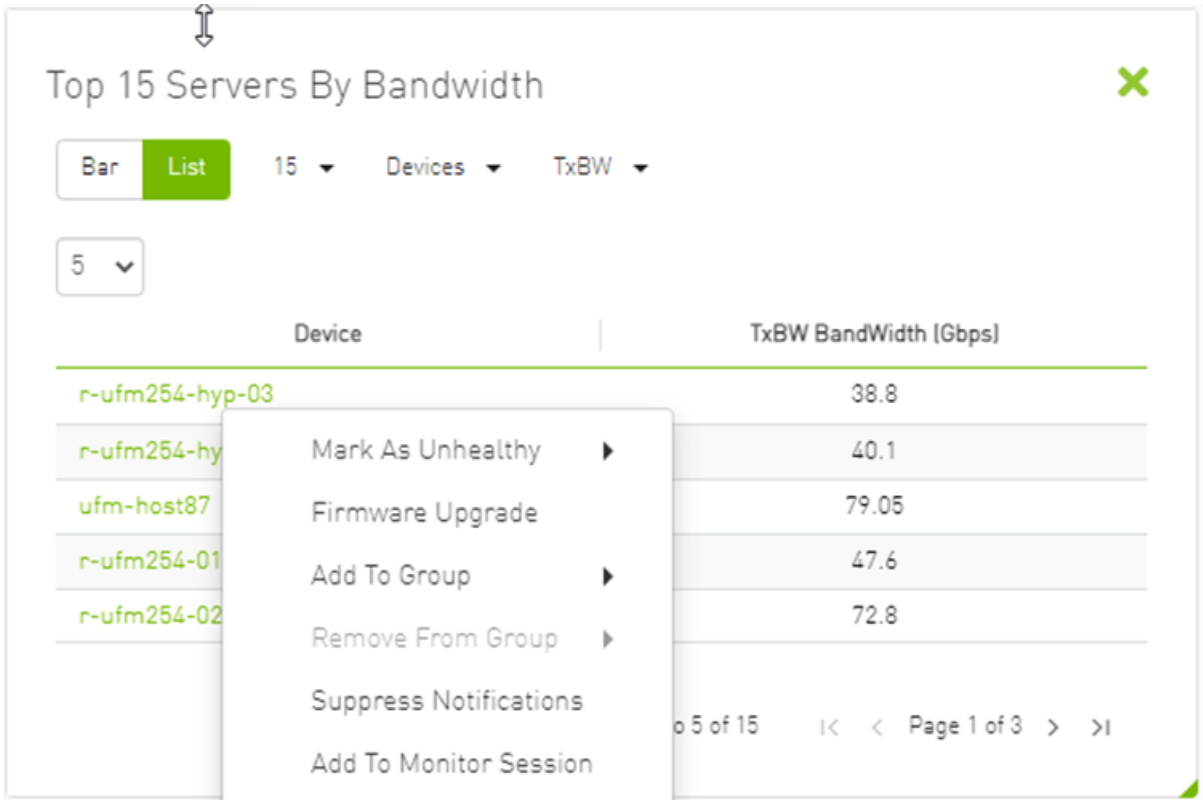
Top Servers/Switches by Bandwidth—Bar View

Top 15 Servers By Bandwidth
Bar
List
15
Devices
TxBW
r-ufm254-hyp-03
79.1
r-ufm254-hyp-03
72.65
r-ufm254-hyp-03
72.25
r-ufm254-hyp-03
70.35
r-ufm254-hyp-03
70.1

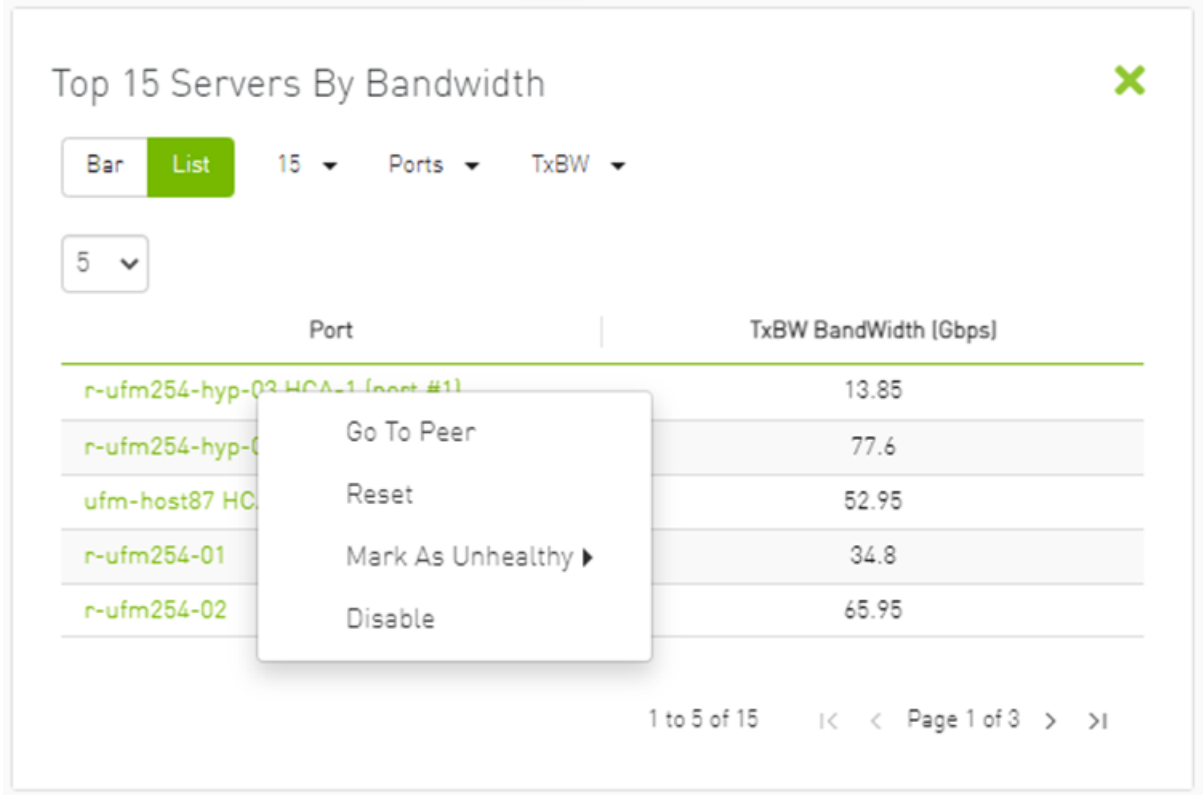
Top Servers/Switches by Bandwidth—List View



Right-clicking a device displays a list of the actions that can be performed. These actions (shown in the following screenshot) are the same actions available in the devices table (see Devices Actions table under Devices Window).



Right-clicking a port displays a list of the actions that can be performed. These actions (shown in the following screenshot) are the same actions available in the Ports table (see Ports Window for more information).



Top N Congested Servers/Switches by Rx/Tx Bandwidth

The Top N Congested devices by Rx or Tx Bandwidth component shows the top congested devices, classified top-down according to the defined Transmit (Tx) or Receive (Rx) bandwidth.

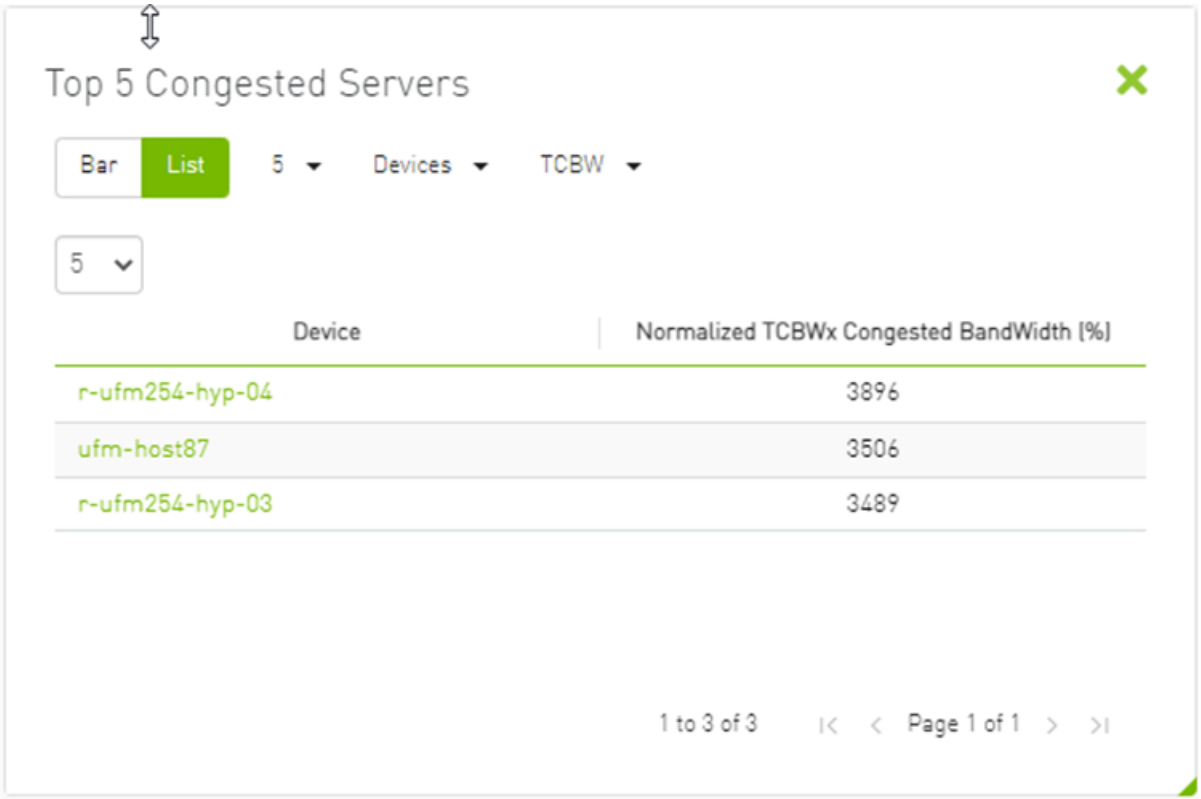
Bandwidth is measured as congestion bandwidth rate (CBW) by percentage.

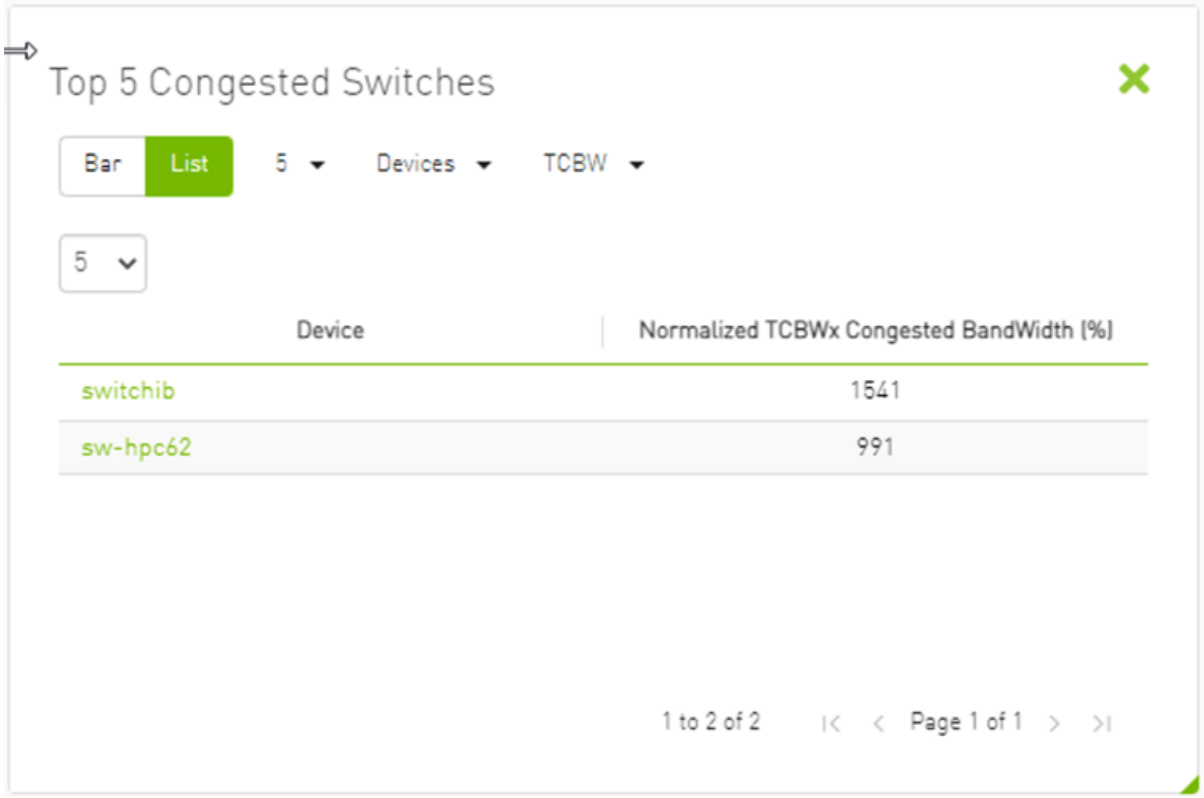
- For Tx, congestion is measured by N HCA ports.

- For Rx, congestion is measured by N switch ports connected to HCAs.

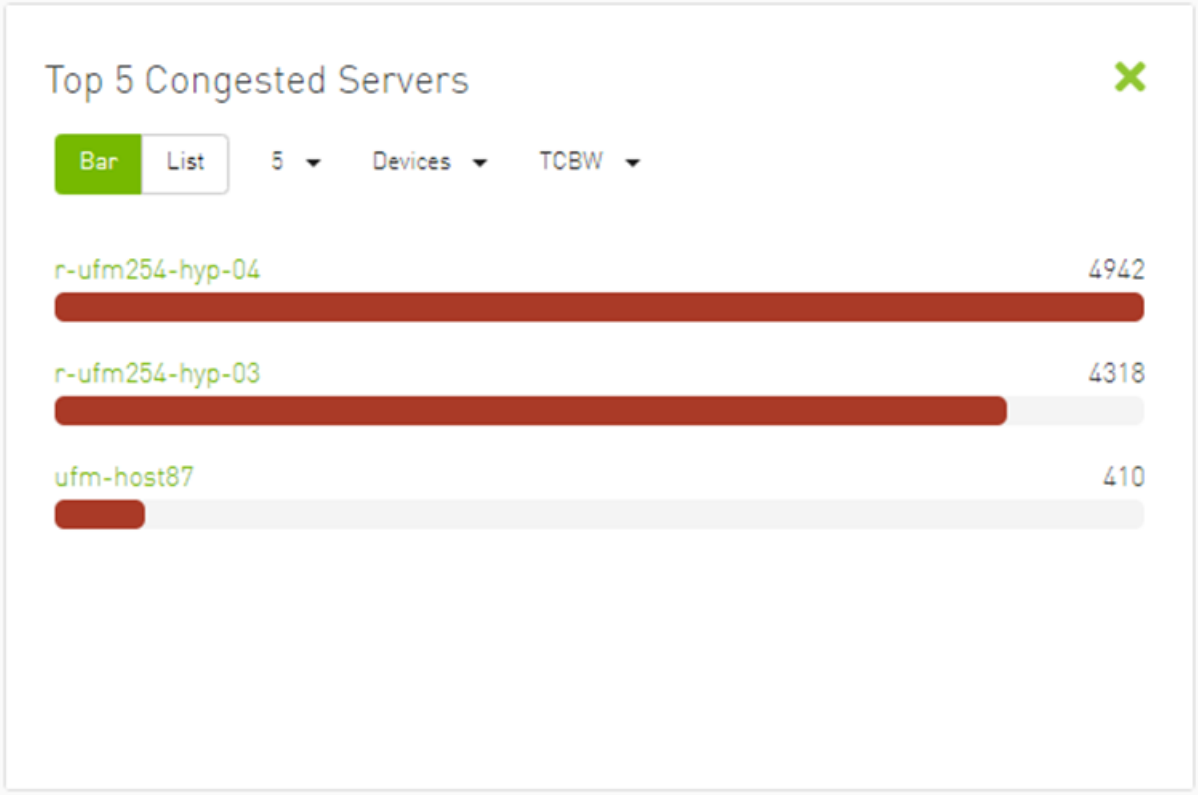
> ⚠ N can be 5, 10, 15, or 20.

Top N Congested Servers by Bandwidth—List View





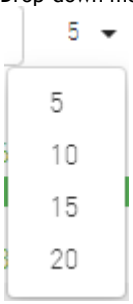
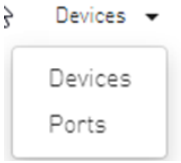
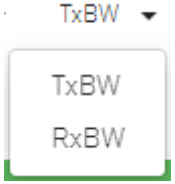
Top N Congested Servers/Switches by Bandwidth—Bar View



The following table describes the options available in this component.

Top N Congested Devices by Rx/Tx Bandwidth

Options	Description
Bar view Bar List	Shows the top N congested devices as a bar graph • X axis shows the rate as a percentage

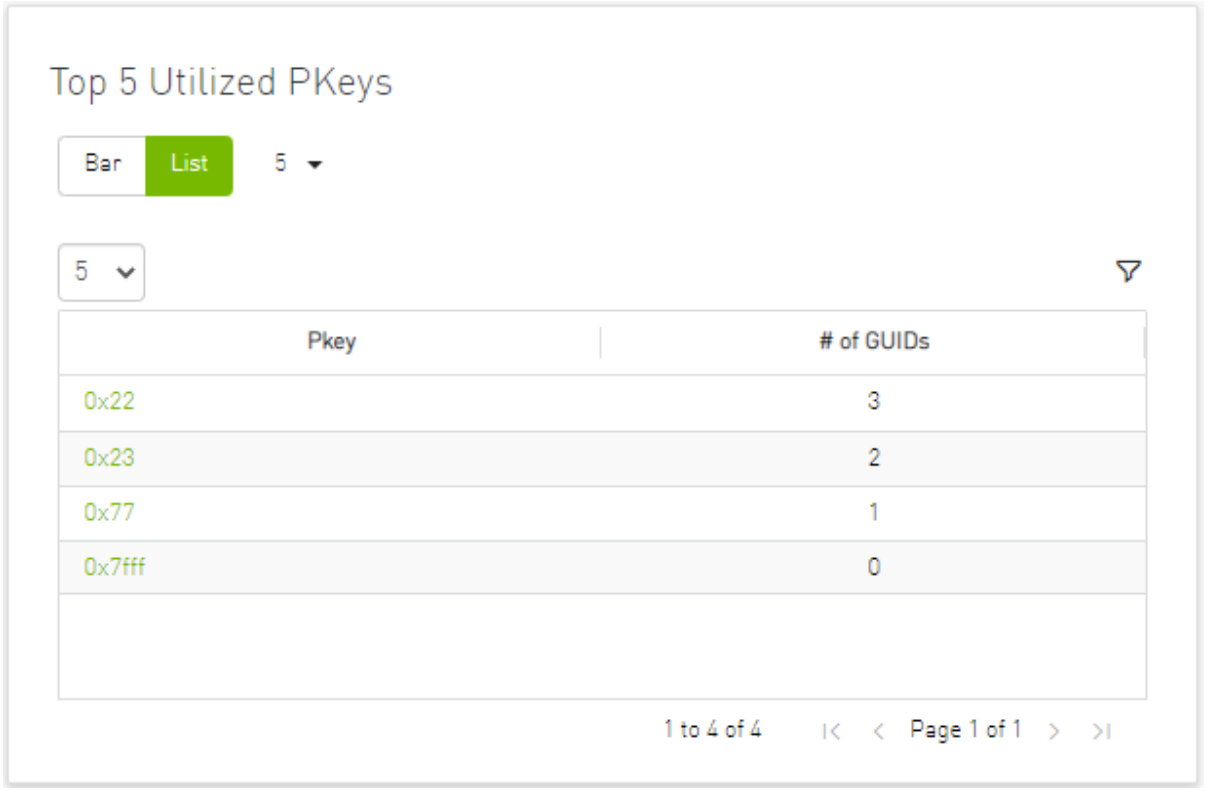
Options	Description
	• Y axis shows the congested Node (server) name
List view 	Shows the top N congested nodes as a list Each congested node is shown in a row with the name of the node and its picture. It also shows the bandwidth rate
Drop-down menu 	Enables selecting the number of top N congested nodes Default: 10 nodes
View by port/element 	Switches view to Top 5 elements By Bandwidth or Top 5 Ports By Bandwidth. Devices view is presented by default. • Clicking a specific port in the Ports view under the Port column redirects to the Ports table and highlights that particular port • Clicking a specific device in the Nodes view under the Device column redirects to the Devices table and highlights that particular node
Monitoring attributes 	• RCBW – Receive Congested Bandwidth (percentage) • TCBW – Transmit Congested Bandwidth (percentage)

Top N Utilized PKeys

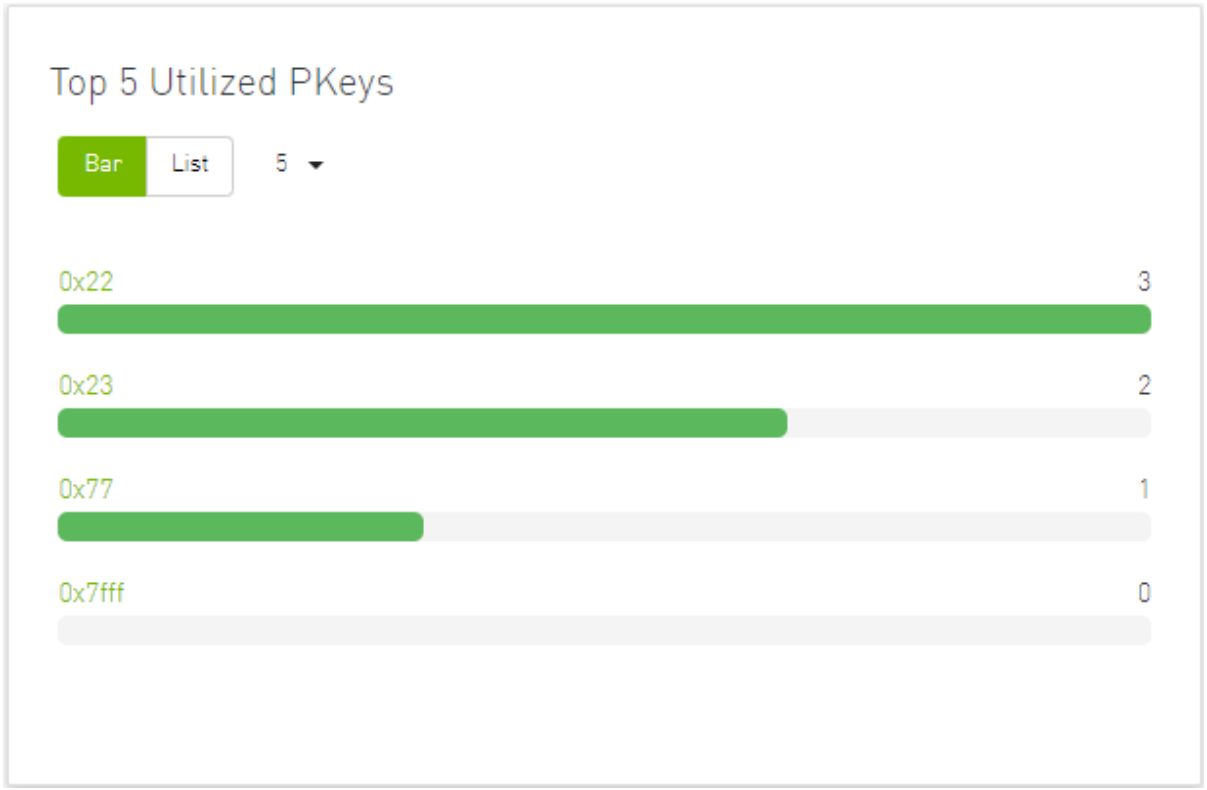
Top N Utilized PKeys displays the top utilized PKeys based on the number of the PKey members.

> ⚠ N can be 5, 10, 15, or 20.

Top N Utilized PKeys—List View

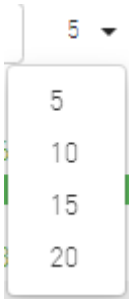


Top N Utilized PKeys—Bar View



The following table describes the options available in this component.

Top N Utilized PKeys

Options	Description
Bar view 	Shows the top N ***Utilized Pkeys*** as a bar graph • X axis shows the number of members • Y axis shows the names of the PKeys
List view 	Shows the top N ***Utilized PKeys*** as a list Each PKey is shown in a row with the name of the PKey and the number of its members
Drop-down menu 	Enables selecting the number of top N ***Utilized PKeys*** Default: 10 ***Utilized PKeys***

Top N Alarmed Servers/Switches

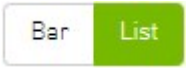
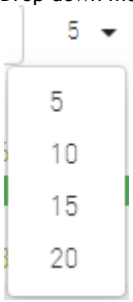
The Top N Alarmed Servers/Switches component shows the top nodes with alarms classified in a descending order. Alarmed nodes are measured according to the following:

- Severity – only the top nodes, in order of severity:
  - Critical
  - Minor
  - Warning

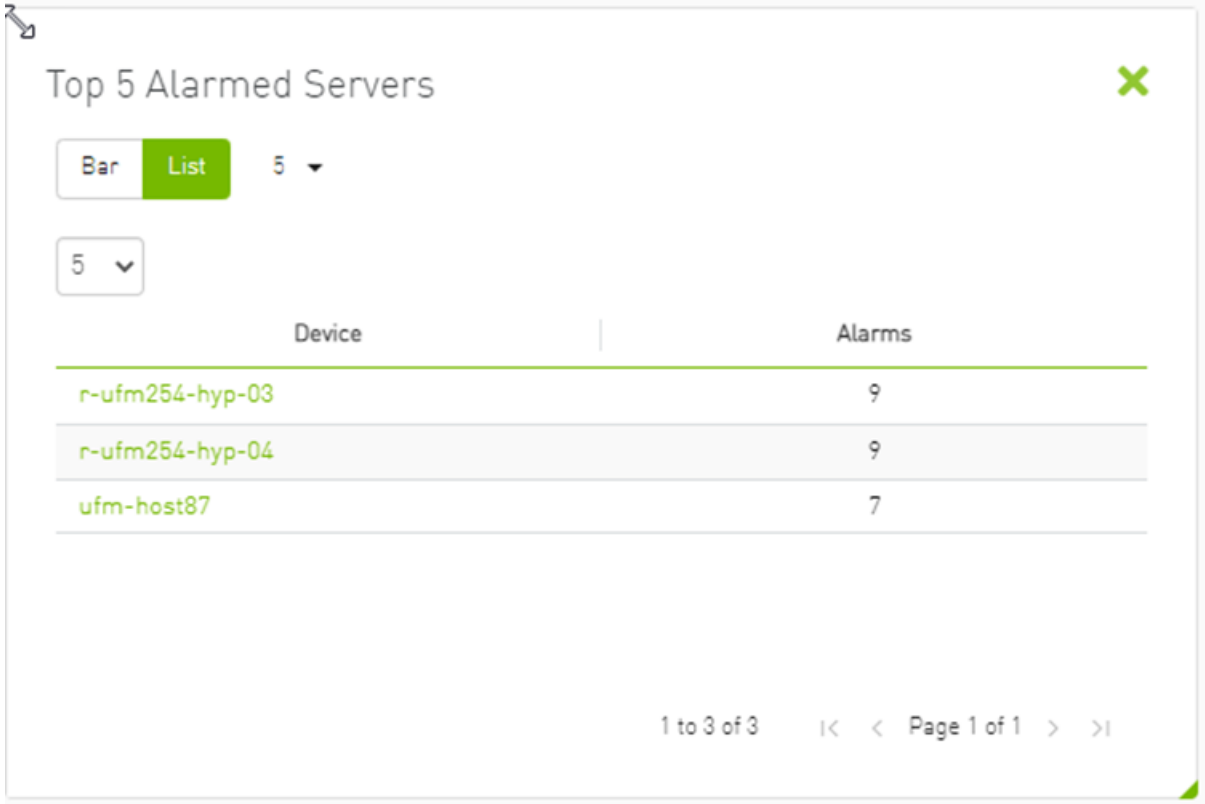
    - Normal
- Alarm – numbers (N can be 5, 10, 15, or 20)

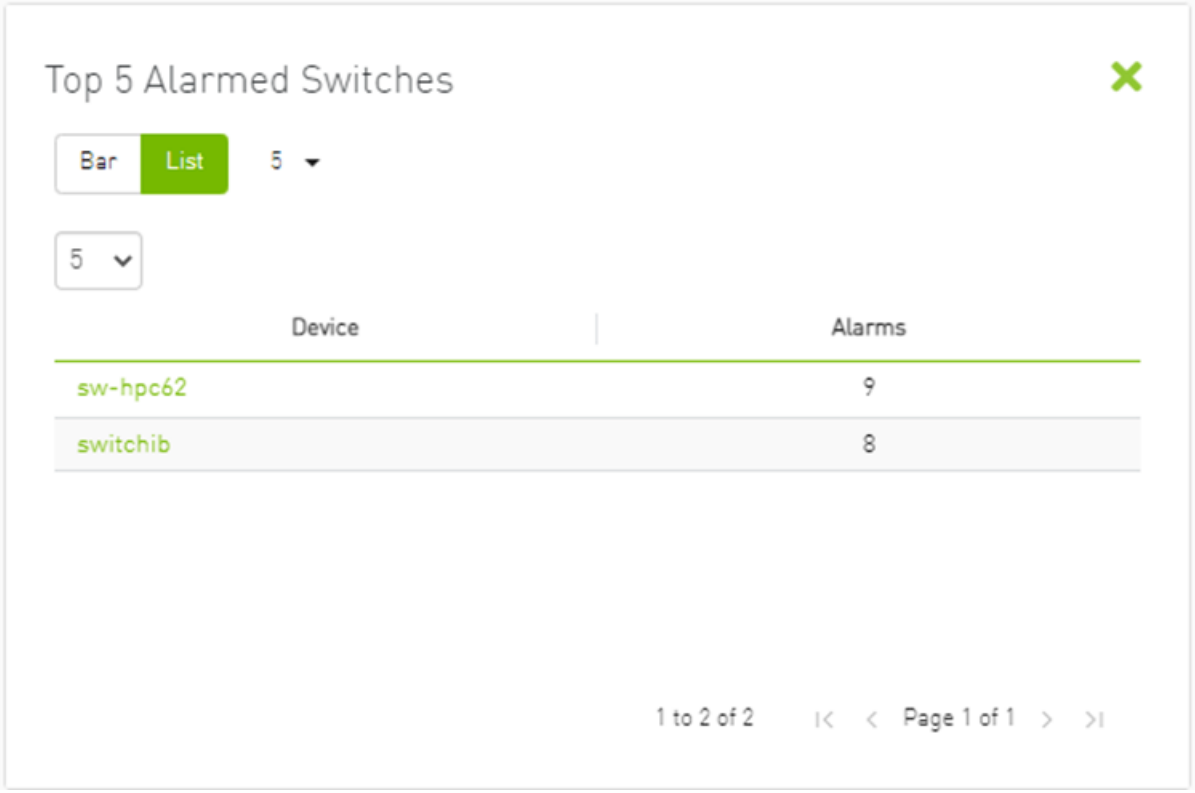
The following table lists the components.

Top N Alarmed Servers/Switches

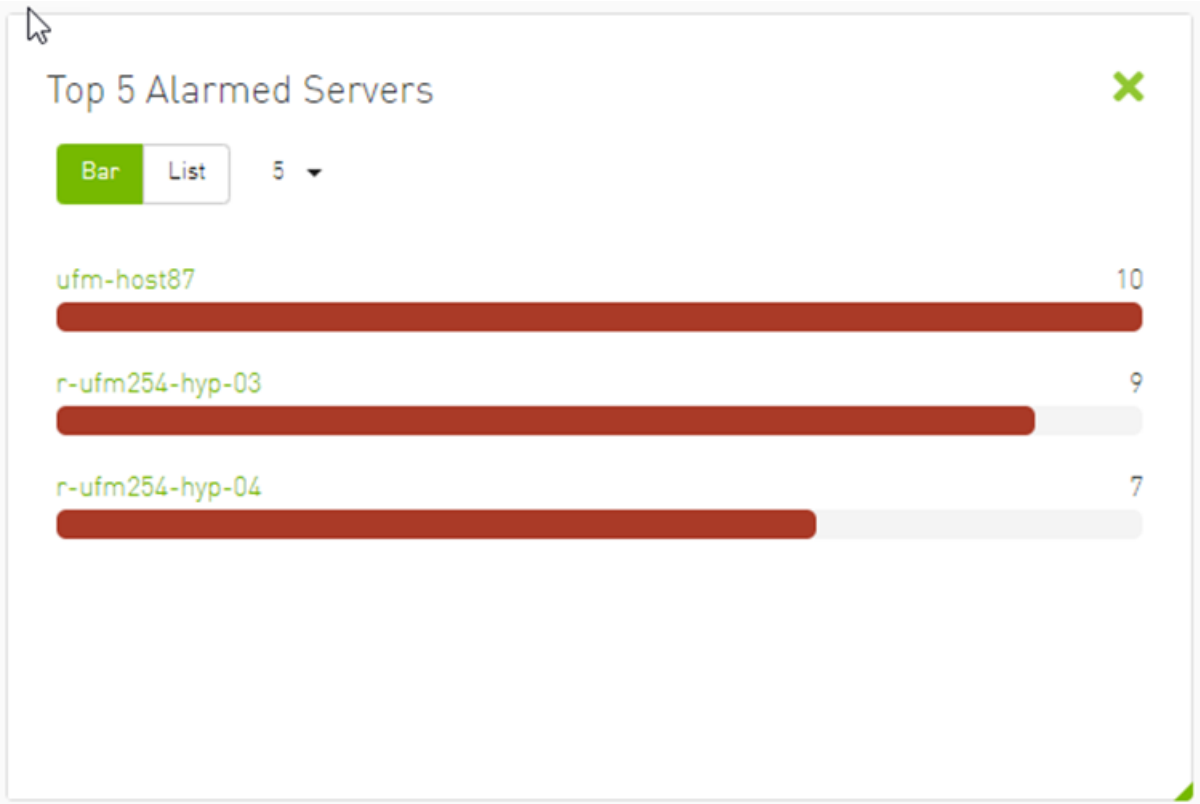
Options	Description
List view 	Shows the top N alarmed servers/switches as a list. Each alarmed device is shown in a row with the name of the node and the number of alarms.
Bar view 	Shows the top N alarmed devices as a bar graph. • X axis shows the number of alarms • Y axis shows the names of the alarmed nodes (servers)
Drop down menu 	Enables selecting the number of top N alarmed nodes. Selects the number of items to display. Default: 10 alarmed nodes
Filter toggle 	Toggles the Filter textbox

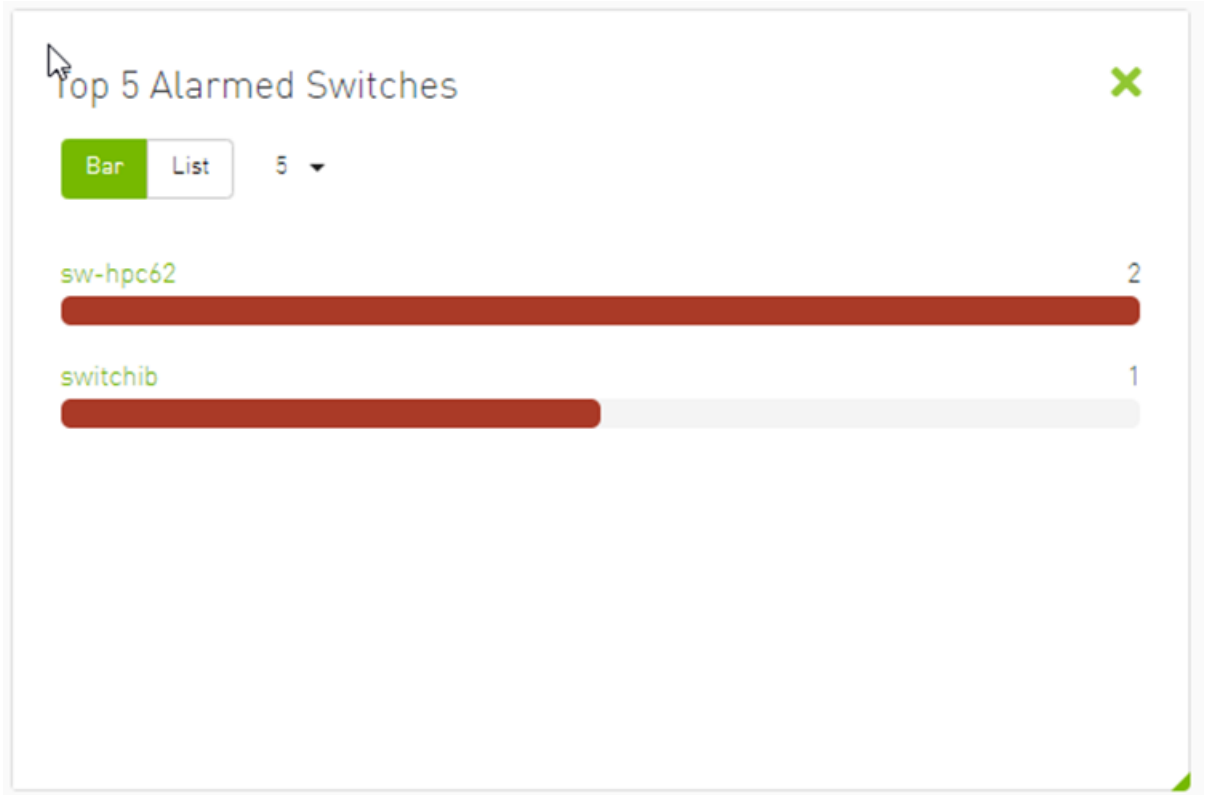
Top Alarmed Servers/Switches—List View


Top 5 Alarmed Servers
Bar
List
5
5
Device
Alarms
r-ufm254-hyp-03
9
r-ufm254-hyp-04
9
ufm-host87
7
1 to 3 of 3
Page 1 of 1



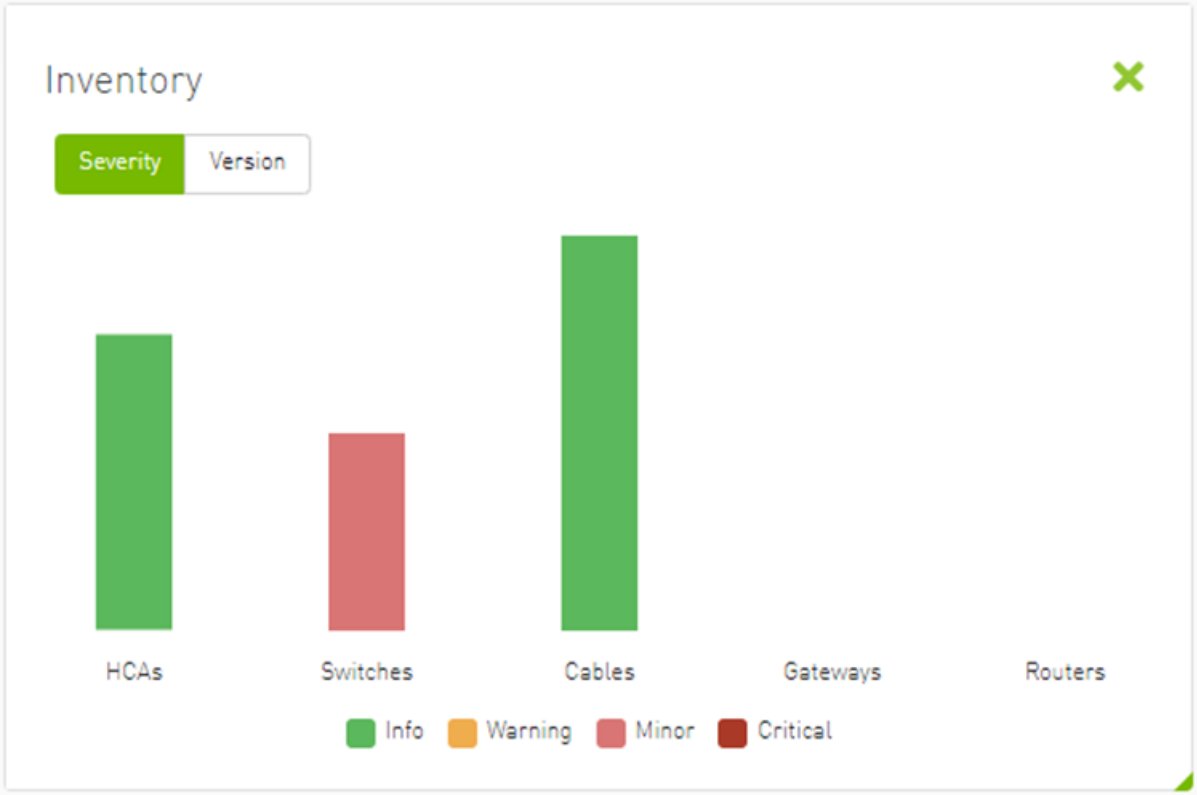
Top N Alarmed Servers/Switches—Bar View

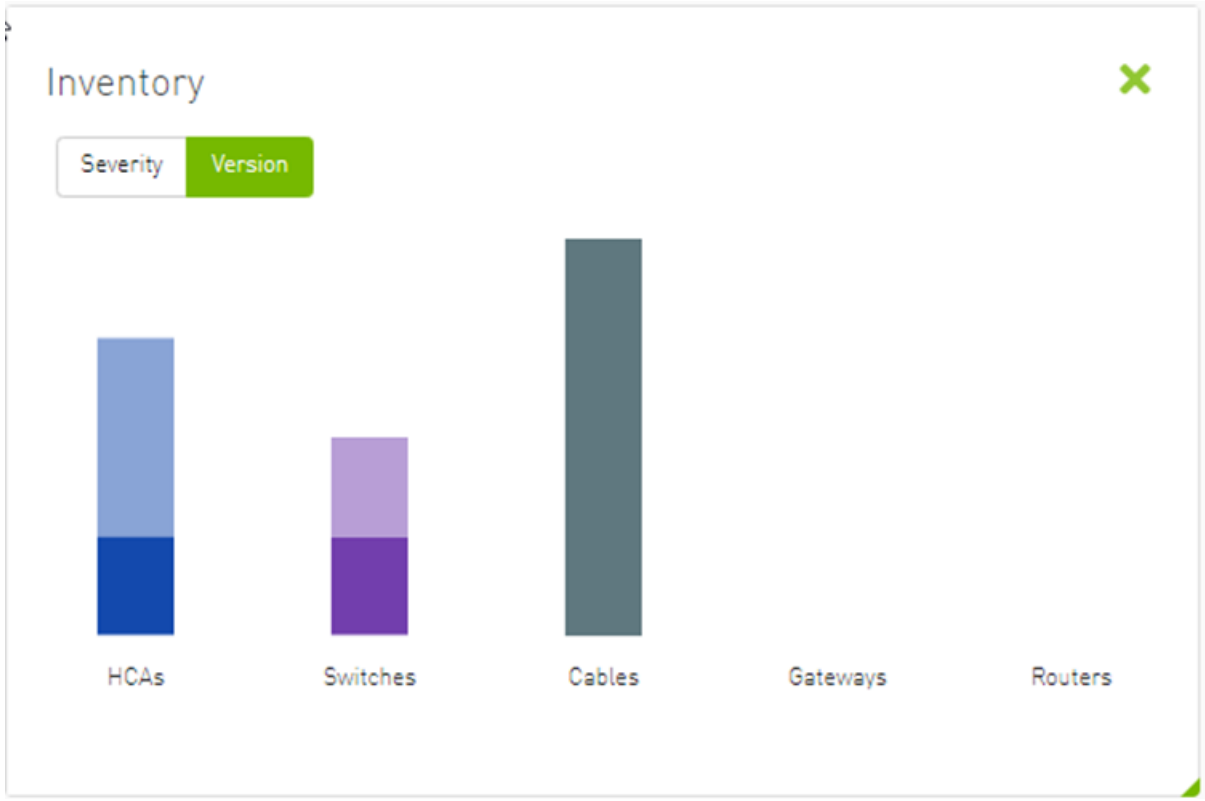

Top 5 Alarmed Servers
Bar
List
5
ufm-host87
10
r-ufm254-hyp-03
9
r-ufm254-hyp-04
7



Inventory Summary

The Fabric Inventory Summary component shows a summary of your fabric inventory (HCAs, Switches, Gateways, Routers and Cables) categorized by the element's severity or firmware version.


Inventory
Severity
Version
HCAs
Switches
Cables
Gateways
Routers
Info
Warning
Minor
Critical



Clicking on one bar element with specific severity/firmware version will redirect you to the clicked element's table.

Fabric Utilization

The Fabric Utilization component shows the number of alarmed objects, categorized by the alarm's severity. They are as follows:

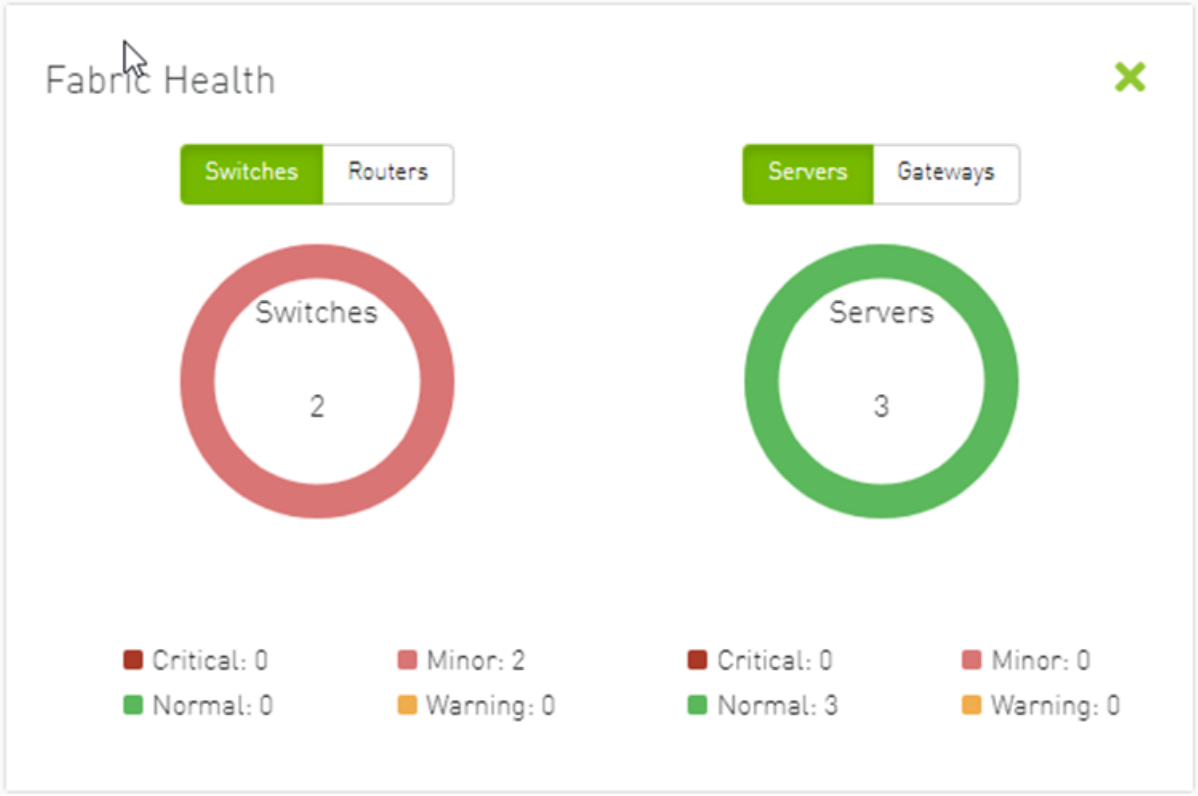
1. Warning
2. Minor

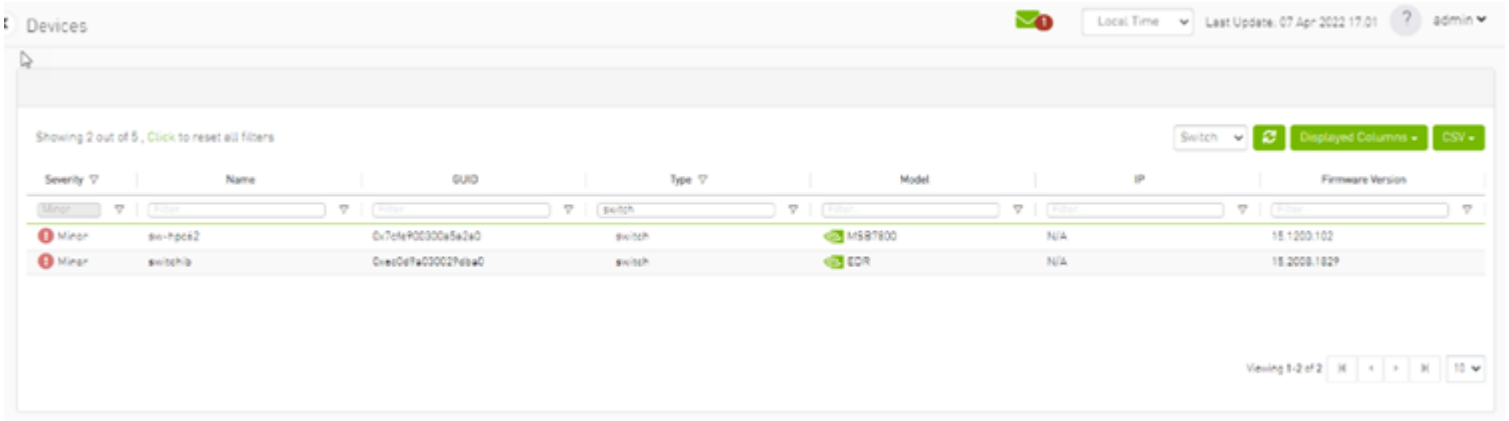
3. Normal
4. Critical

If Server X has 2 minor alarms, 1 warning alarm and 2 critical alarms, and Server Y has 0 minor alarms, 2 warning alarms and 1 critical alarm, the **Fabric Resource Utilization** pie chart will show 2 servers in the critical slice, 2 servers in the warning slice and 1 server in the minor slice.

You can filter for both switches and nodes of a specific severity level by clicking the specific pie slice indicating the severity.

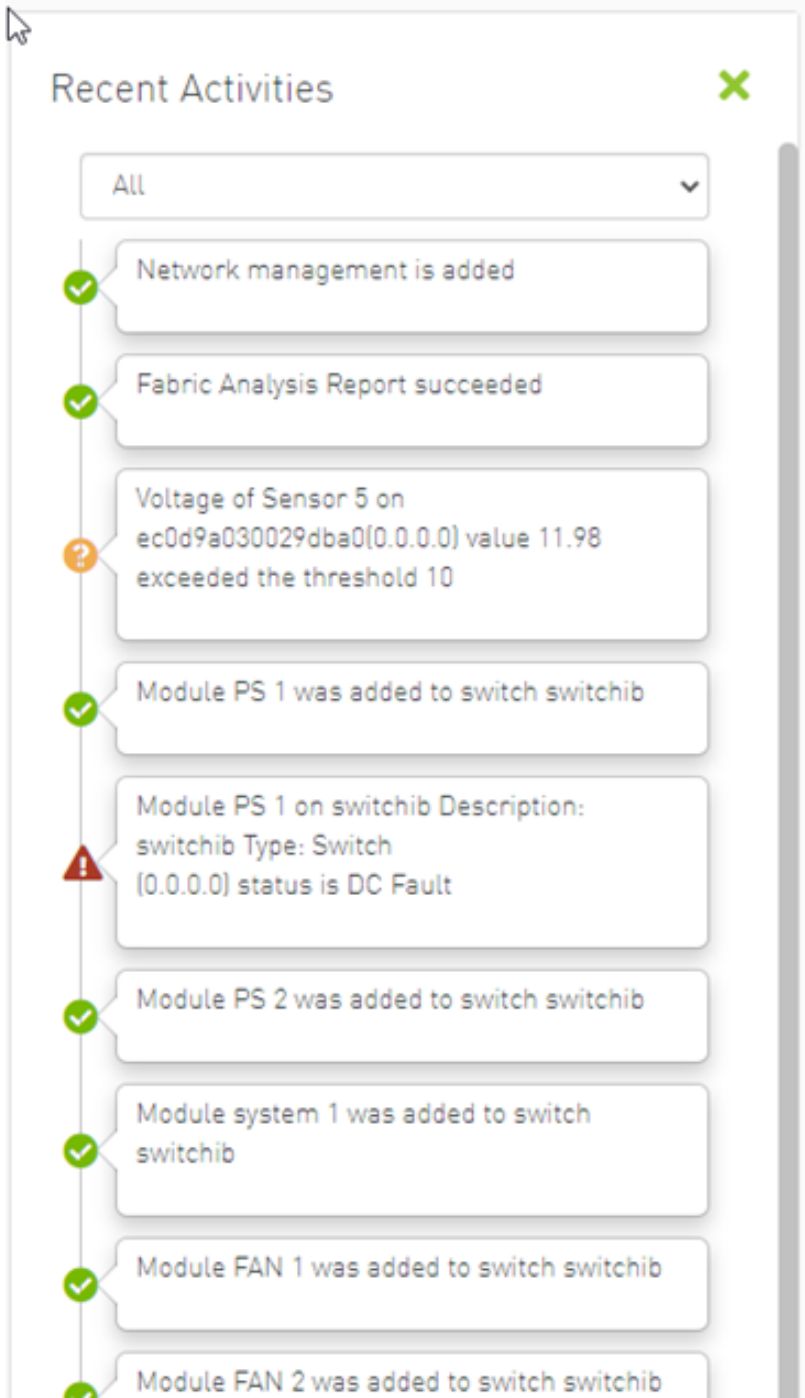
In the example below, the Devices table lists all the switches of severity level "Minor" after clicking the red (Minor) slice from the Switches pie chart.





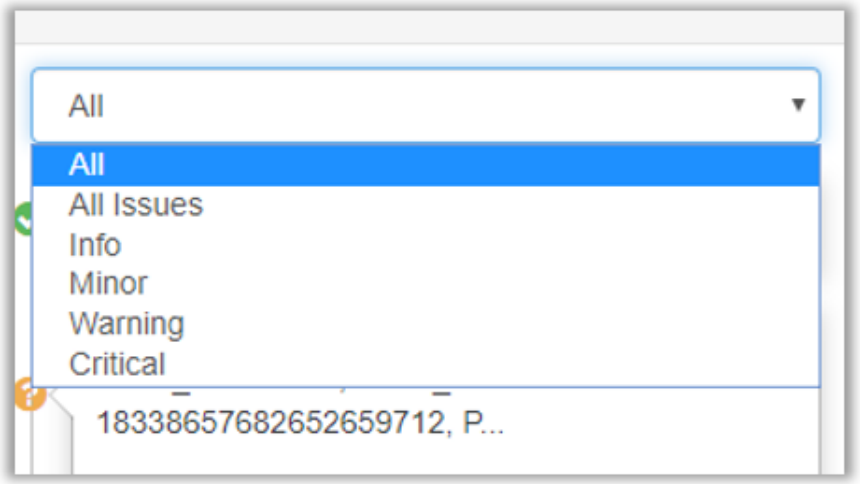
Recent Activities

The Recent Activities component lists the recent events detected by the UFM system.


Recent Activities
All
Network management is added
Fabric Analysis Report succeeded
Voltage of Sensor 5 on ec0d9a030029dba0(0.0.0.0) value 11.98 exceeded the threshold 10
Module PS 1 was added to switch switchib
Module PS 1 on switchib Description: switchib Type: Switch (0.0.0.0) status is DC Fault
Module PS 2 was added to switch switchib
Module system 1 was added to switch switchib
Module FAN 1 was added to switch switchib
Module FAN 2 was added to switch switchib

You can filter for the events you would like to see in one list using the drop-down menu that provides the following options:

- All – shows all recent activities
- All issues – shows all non-Info activities
- Info – shows all activities with Info severity or higher
- Minor – shows you all activities with Minor severity or higher
- Warning – shows you all activities with Warning severity or higher
- Critical – shows you all activities with Critical severity

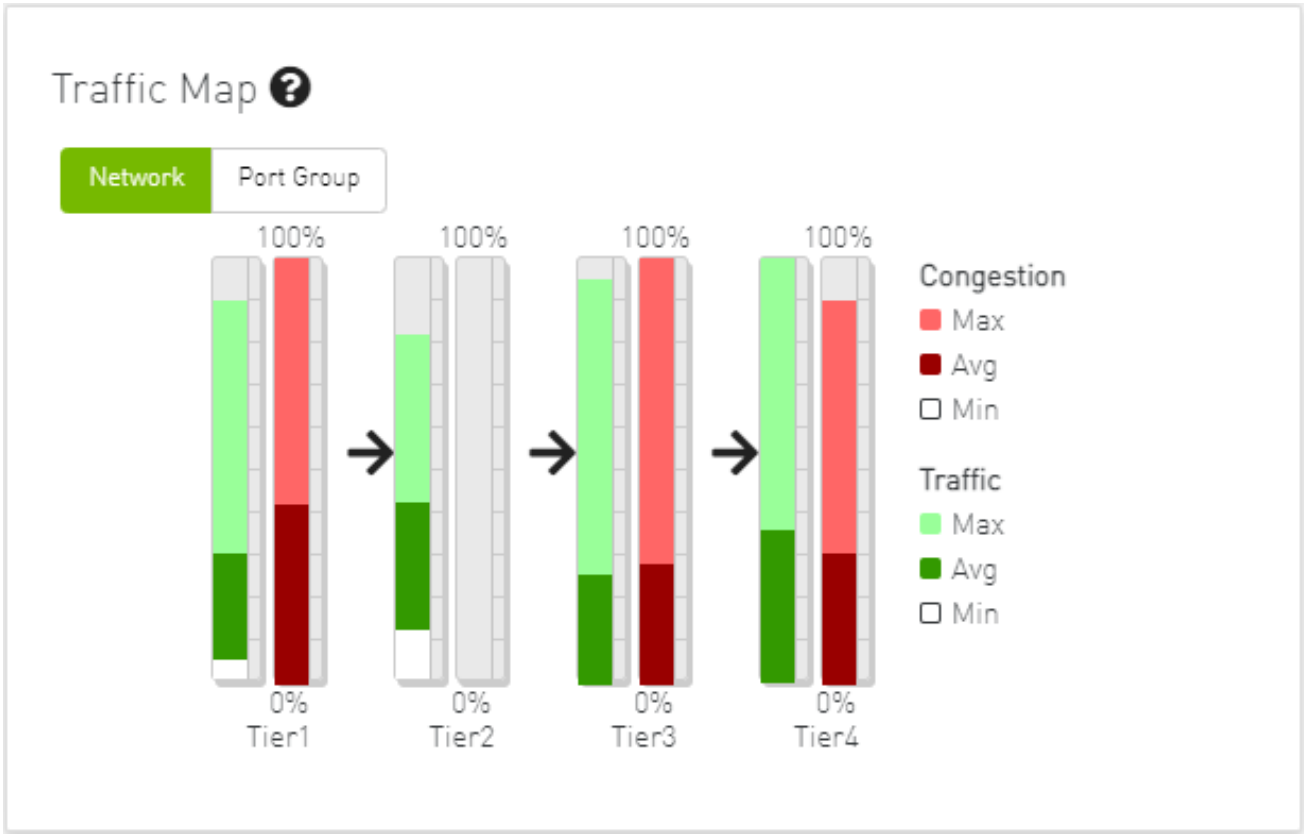


Traffic Map

The Traffic Map dashboard shows the normal traffic versus congested traffic distributed on switch tiers and on port groups. This view, together with the **Top N Congestion** dashboard, gives a full status of the traffic congestion of the fabric.

Network Traffic Map

Four double bars represent the transmitted bandwidth (normalized transmit data) and normalized congested bandwidth (CBW), both measured in bytes/sec with minimum, average, and maximum bandwidth values.



An explanatory window on traffic map opens once clicked on the icon.

Traffic Map Guide

Mellanox's unique Traffic Map provides a valuable real-time aggregate view of the fabric performance by showing the overall bandwidth utilization per switching tier coupled with congestion information.

Reading the Traffic Map Chart

The Traffic Map contains four tiers; each tier is represented by a green and a red bar, as shown in the following Traffic Map Chart :

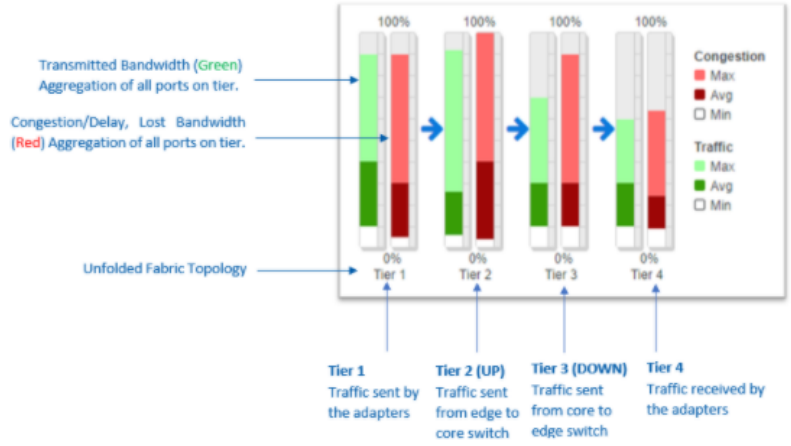


Figure 1: Traffic Map Chart

Color coding for each tier is as follows:

- The green the percentage of overall bandwidth generated by the specific tier. This bar is divided in light and dark green colors.
  - The light green indicates the peak port utilization.
  - The dark green indicates the average utilization.
- The red bar indicates the percentage of congestion (also referred to as lost bandwidth) in the specific tier. This bar is divided in red and dark red colors.
  - Red indicates the peak port congestion.
  - Dark red indicates the average congestion.

Close

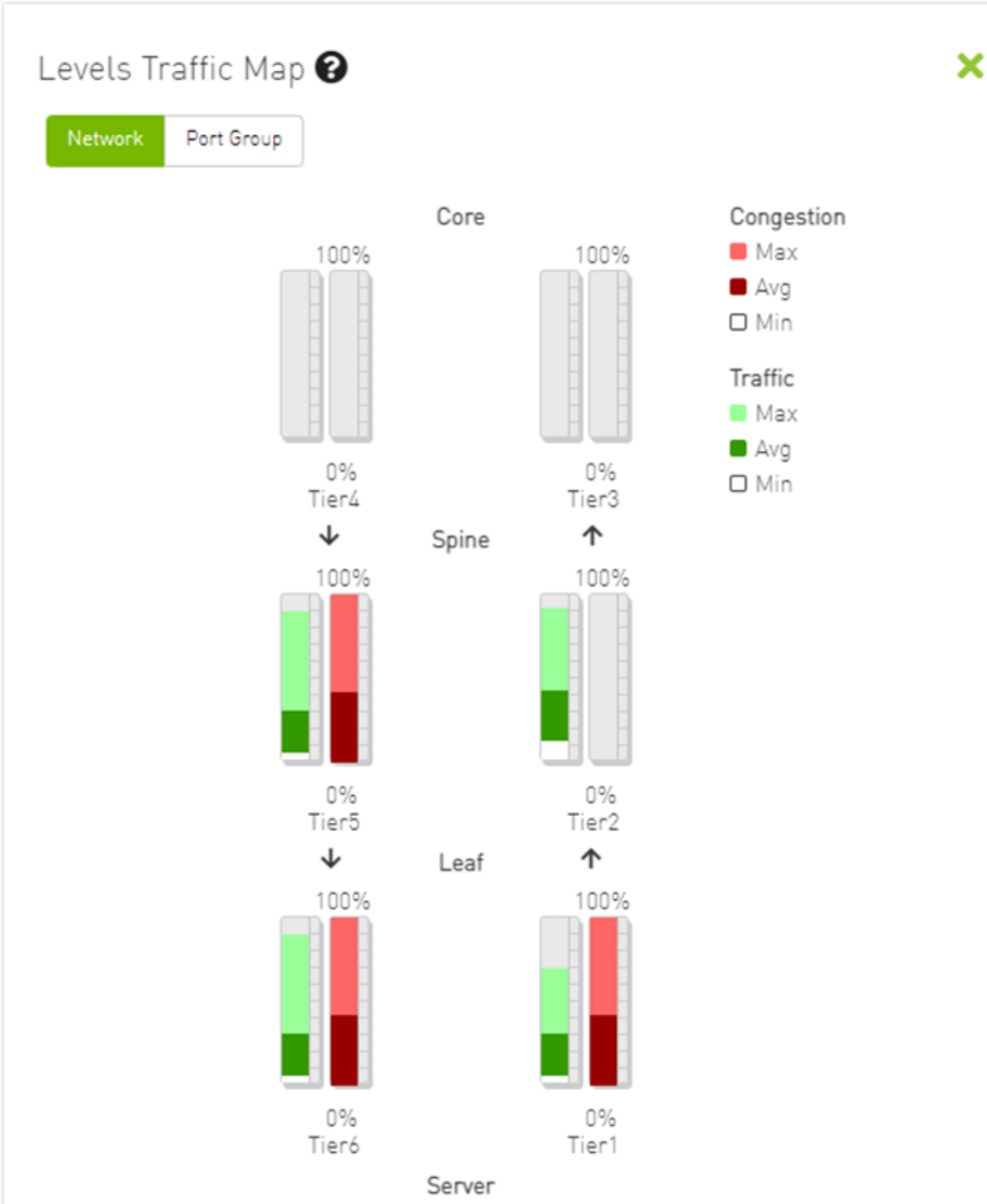
The percentage of total theoretical bandwidth (TBW) is calculated based on the underlying InfiniBand technology (SDR, DDR, QDR, FDR or EDR). The speed can be viewed when checking the ports.

- The vertical axis shows the following:
  - Bandwidth (BW) is represented by a green bar and is measured in percentages
  - Congested Bandwidth (CBW) is represented by a red bar and is measured in percentages
  - Minimum, average, and maximum bandwidth are represented in each bar by a subset color
- The horizontal axis represents the tiers.
  The bottom of the dashboard represents the tier-related transmitted traffic, which is divided into four segments by measurement ports:
  - Tier 1 – represents the traffic injected by all adapters
  - Tier 2 – represents the traffic sent from the edge switches to the core of the fabric (in case of a single Director switch, this tier indicates traffic utilization inside the Director between the line and fabric boards)
  - Tier 3 – represents the traffic sent from the core to the edge switches
  - Tier 4 – represents the traffic sent from the edge switch to the adapters

> ⚠ The illustrations at the bottom of the tiers show a four-tier topology:
> Server [tier 1] Switch [tier 2] Director Switch [tier 3] Switch [tier 4] Server.

Levels Network Traffic Map

Different representation of the fabric traffic map that based on the devices/ports levels.


Levels Traffic Map
Network
Port Group
Core
100%
100%
0%
Tier4
0%
Tier3
Spine
100%
100%
0%
Tier5
0%
Tier2
Leaf
100%
100%
0%
Tier6
0%
Tier1
Server
Congestion
Max
Avg
Min
Traffic
Max
Avg
Min

The level of the device/port is the distance between the device and the nearest server/gateway.

Levels Calculations:

- The levels calculations are configurable from the `gv.cfg` file under TopologyLevels section enable item and it is disabled by default.
- The levels names are configurable from the `gv.cfg` file under TopologyLevels section levels item and by default we are defining up to 4 levels levels equals server, leaf, spine, core
    - Server: hosts and gateways.
    - Leaf: switches and routers that are directly connected to the server
    - Spine: switches and routers that are directly connected to the leaf
    - Core: switches and routers that are directly connected to the spine

If the fabric has more than 4 levels, the level value will be L + distance e.g., L4, L5, L(N), and if levels was empty, the levels will start from L0, L1, L2, etc.

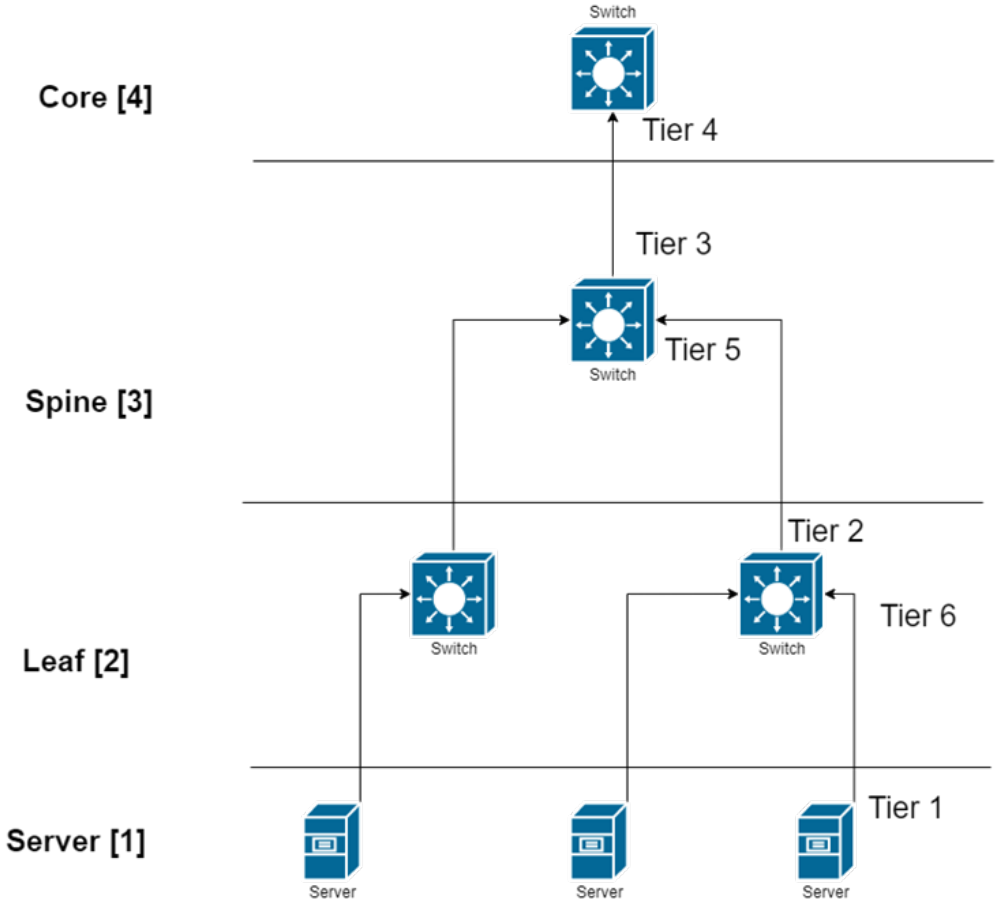
The levels calculations are done at either the discovery stage or once the topology changes.

Ports Tiers calculations based on the levels:

If the levels calculations is enabled, the port’s tier will be calculated as the following steps:

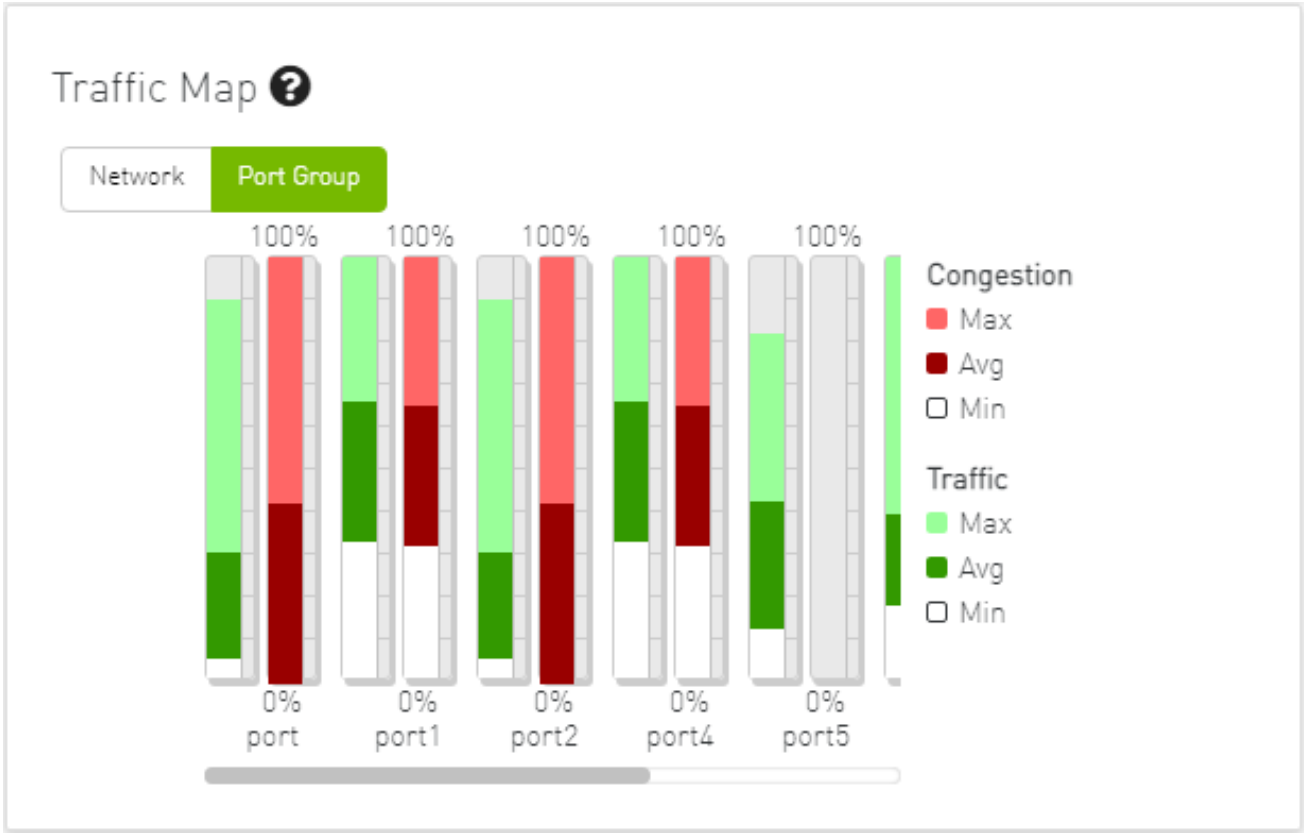
1. Get the level for both port's parent device and port's peer parent device
2. Decide whether the port’s data flow is the up or down direction, by checking the order of the parent and peer parent level:
    a. If the parent's level order is less than or equals the parent peer level, then the port's flow is up and tier is the parent level order
    b. If the port's flow is down and the tier is the distance between the host to the root device and the distance between the root to the parent device

Example:



If the level calculations are disabled, the tier calculations will be done as mentioned in this section.

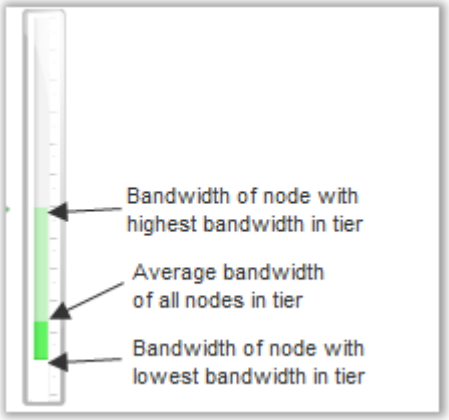
Port Group Traffic Map



Traffic Map Bar Chart

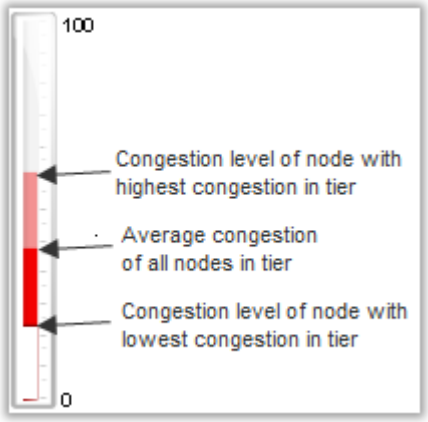
- **Bandwidth Bars**
  The bandwidth graph shows how traffic is traversing the fabric and how traffic is being transmitted between the servers. For example, the following considerations could be evaluated:
  - The size of the difference between max bandwidth and min bandwidth.
  - The traffic that is flowing in the middle tiers and whether it would be more efficient to move the traffic to the edges to save the uplinks.

  Bandwidth levels are measured in percentages, as shown below:



- **Congestion Bars**
  The Congestion graph shows where congestion starts. For example, the following considerations could be evaluated:
  - If congestion is in the first or second tier, there is probably a routing problem
  - If there is no red bar, it means that there is no congestion or no routing problems

Congestion levels are measured in percentages, as shown:

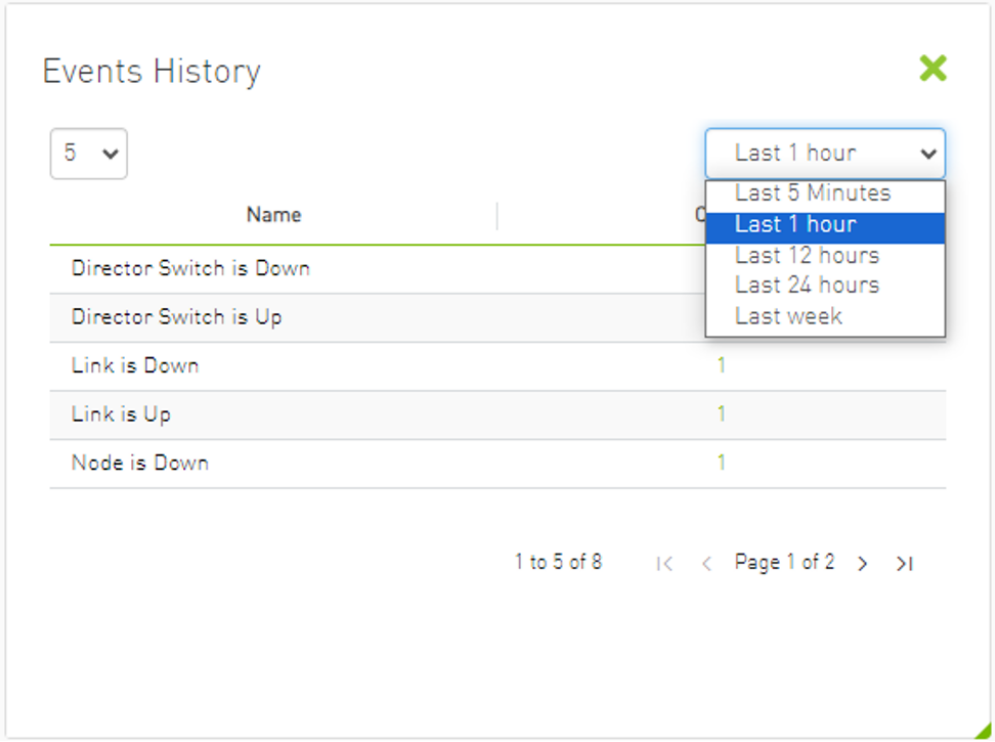


Events History

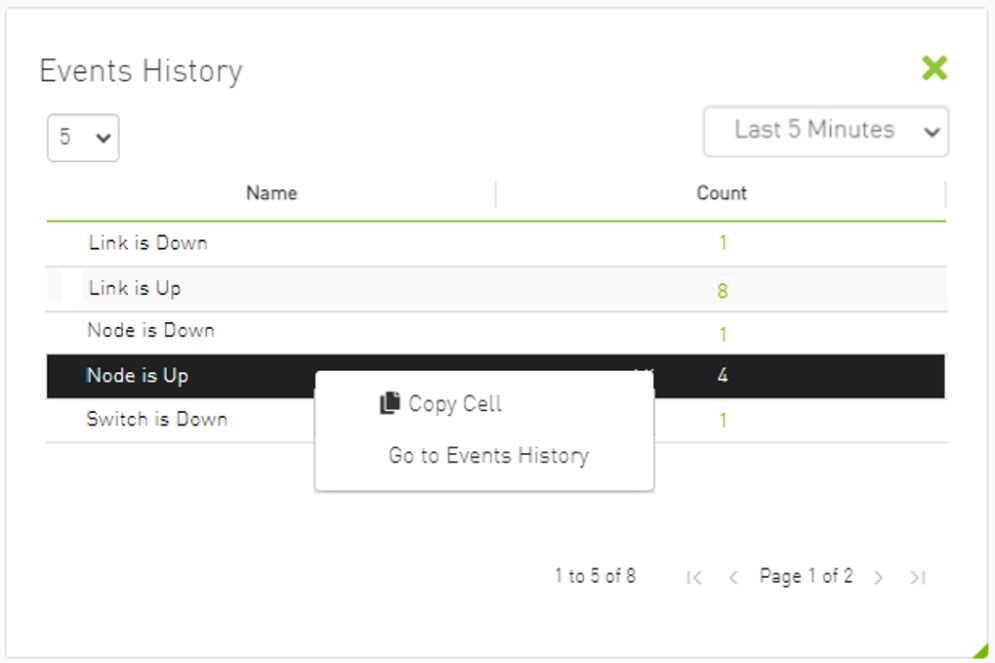
> ⚠ To view the Event History panel in the dashboard, the System Monitoring feature must be enabled. Otherwise, the panel will be hidden. Users can enable System Monitoring by setting the `system_monitoring_metrics` flag under the `SystemMonitoring` section in the `gv.cfg` file to true.

The Events History panel presents the topology change events in a table along with their respective counts.

The user can filter the event count by selecting the desired time interval.

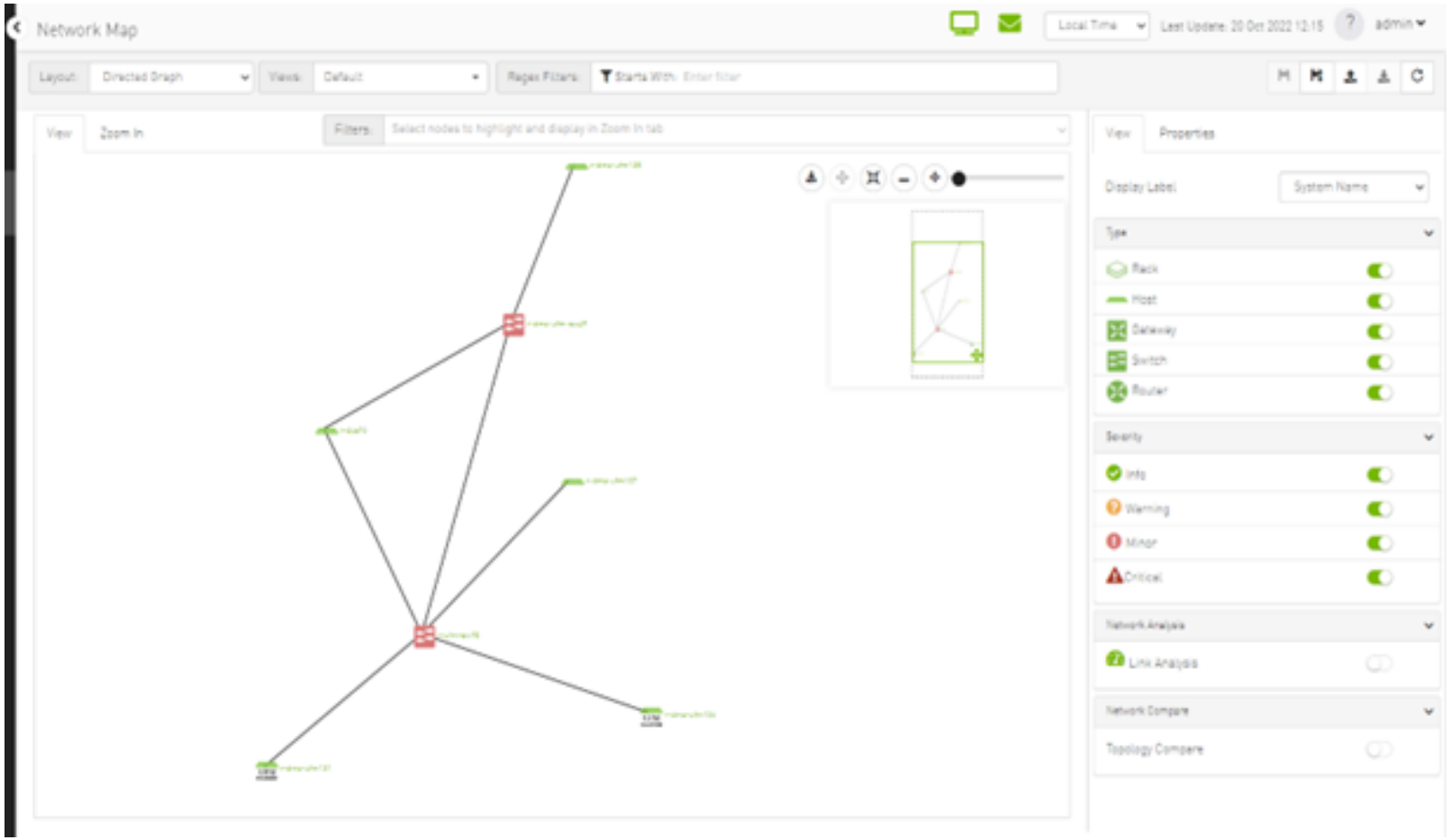


Users can navigate to the ‘Device/Link Status Events’ tabs by either clicking on the counter value or by right-clicking and selecting 'Go to Events History'.



Network Map

The Network Map window shows the fabric, its topology, elements and properties. UFM performs automatic fabric discovery and displays the fabric elements and their connectivity. In the Network Map window, you can see how the fabric and its elements are organized (e.g., switches and hosts).


Network Map
Local Time
admin
Layout:
Directed Graph
Views:
Default
Regex Filters:
Starts With
View
Zoom In
Filters:
Select nodes to highlight and display in Zoom In tab
View
Properties
Display Label:
System Name
Type
Rack
Host
Gateway
Switch
Router
Severity
Info
Warning
Minor
Critical
Network Analysis
Link Analysis
Network Compare
Topology Compare

Network Map Components

Component	Icon	Description
Switches		Represents third party switches discovered/managed by UFM
Hosts		Represents the computer (host) connected to the discovered/managed switches
Routers		Represents third party routers discovered/managed by UFM
Gateways		Represents third party gateways discovered/managed by UFM
Links		Represents the connections between devices on the fabric
Racks		Represents all nodes (hosts) physically connected to a switch

⚠ The level of severity of devices affects the color they are displayed in. For further information, refer to table "Device Severity Levels".

- To zoom in/out of the map, scroll the mouse wheel up and down or using the slider on the right top corner
- To move around in the map, press and hold down the left key while you move sideways and up/down
- To see the hosts inside a rack, right-click the Rack icon and click "Expand Hosts"

Selecting Map Elements

Users are able to select elements from the Network Map. Right-clicking an element opens a context menu which allows users to perform actions on it.

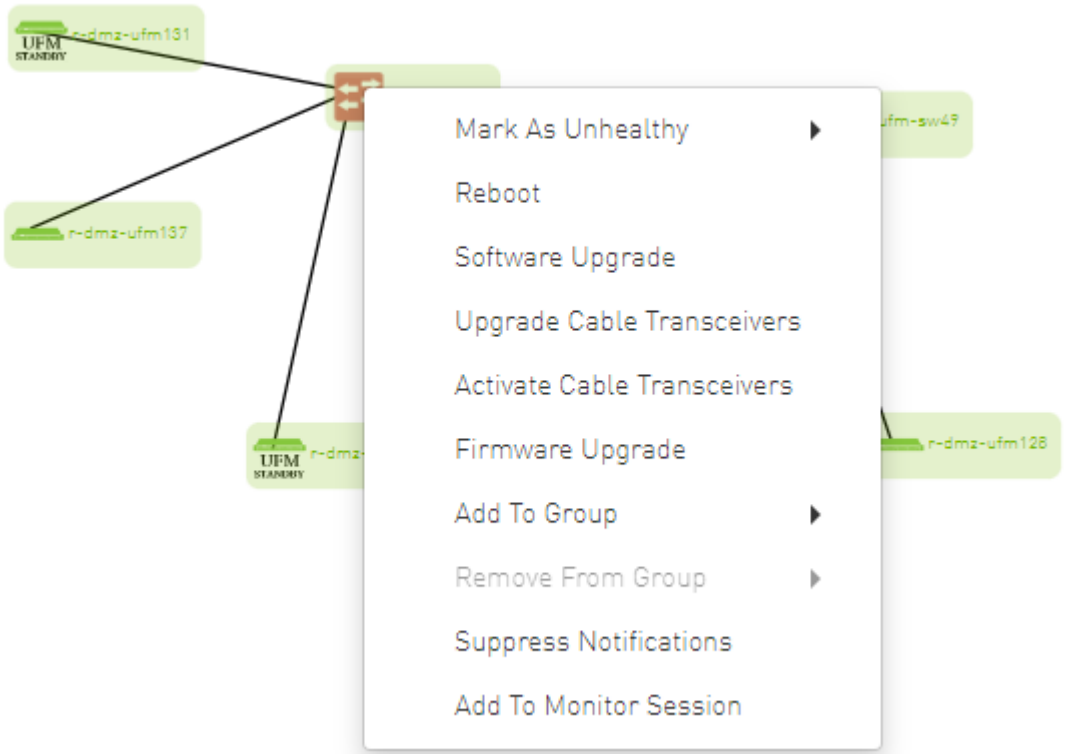
It is possible to select multiple elements at once using any of the following methods:

- By holding down Ctrl or Shift and dragging their mouse across the map.

  > ⚠ Please note that Ctrl starts new selection, while Shift adds to the current selection.

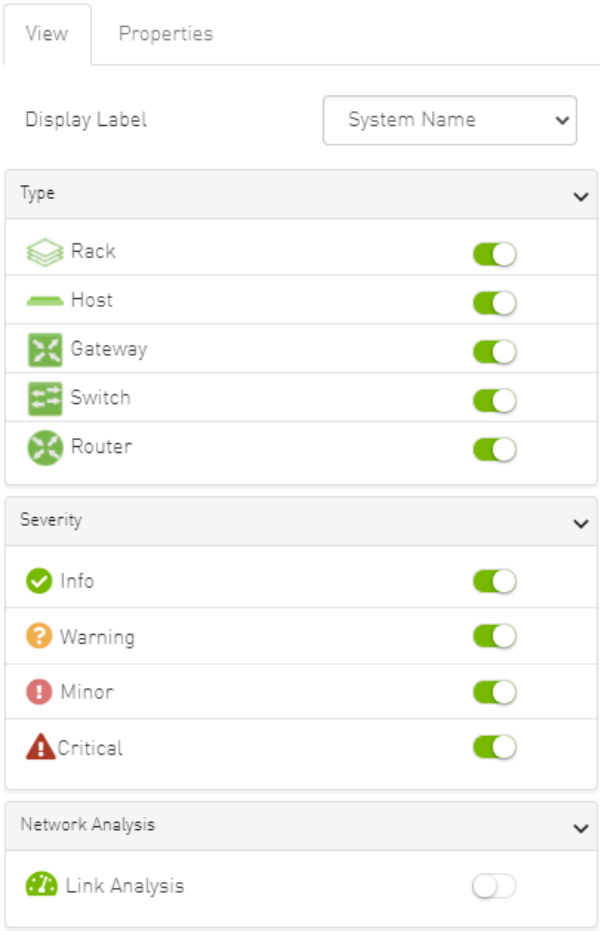
- By holding down Shift and clicking a new element on the map.

Multi-select makes it possible for users to perform actions on multiple devices with one right-click rather than repeating the same process per device.

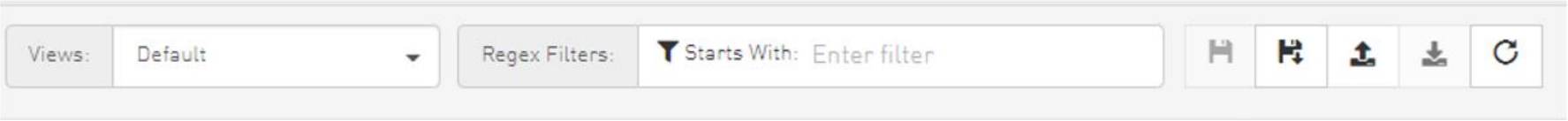


Map Information and Settings

The right pane of the Network Map view enables you to control the view settings, as well as obtain further information on selected elements from the map.



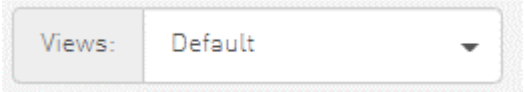
The customized views created using the type and severity filters, selected fabric nodes, zoom level, and Expand/Collapse All Racks options can be saved for later access. These customized views can be saved and accessed using the bar available on top of the Network Map:



- "Save As" icon ( ) saves newly created customized views
- "Save" icon ( ) saves edits performed on existing views
- “Import” icon ( ) import map from local device. The file format should be txt

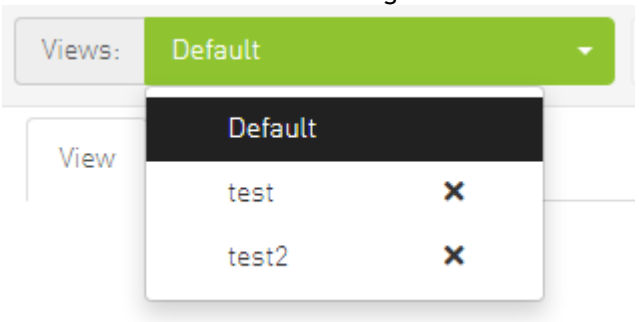
- “Export” icon ( ) export network as text file
- To reload/refresh the network map, use the refresh icon ( ).

- Drop down menu gives access to all previously saved views

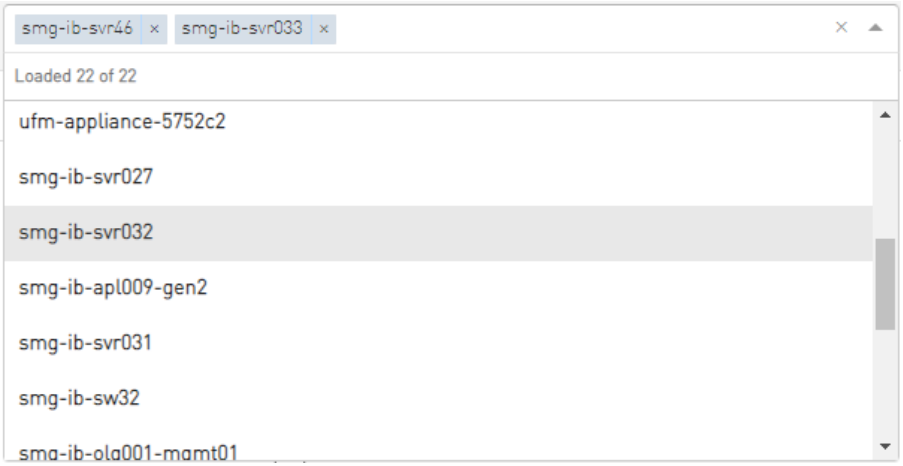


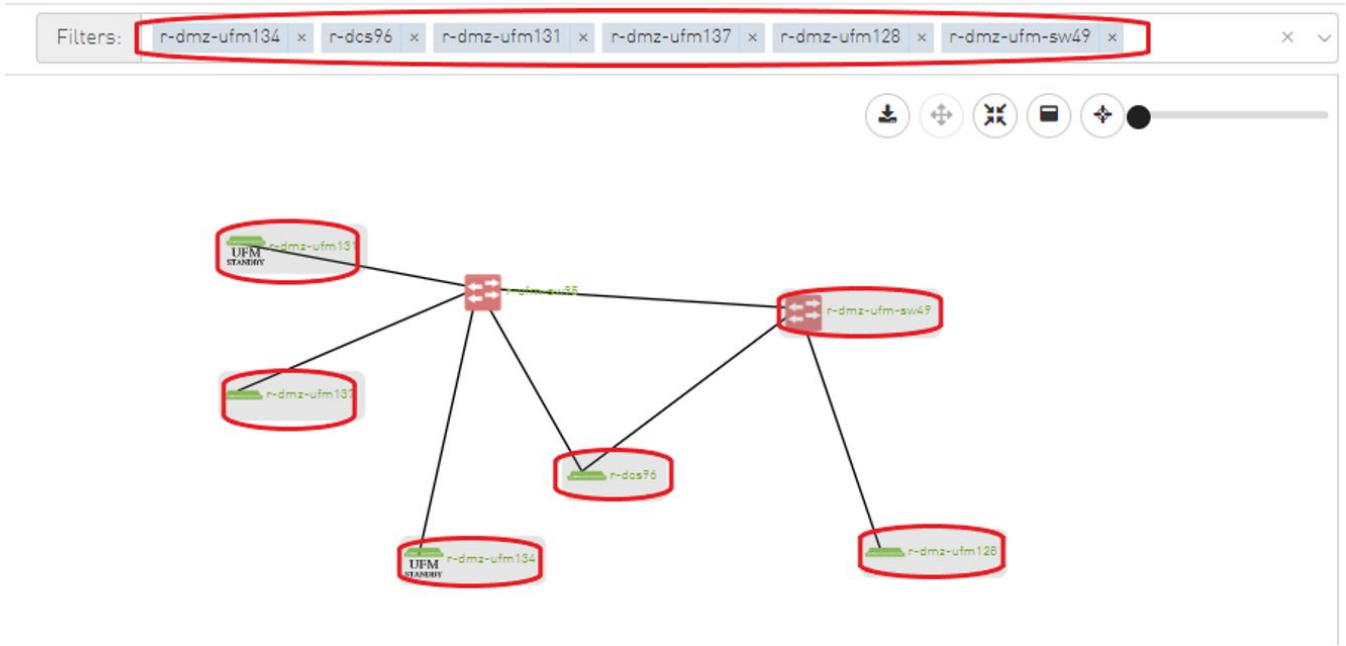
- "Default" view is a predefined view where nodes are positioned randomly, all filters are enabled, and all racks are collapsed. Changes made to this view cannot be saved unless under a new view name using the "Save As" icon.

- Saved views can be deleted using the "x" button.



You can select a node from the dropdown menu located above the Network Map view in order to highlight/display them in the "Zoom In" tab.


smg-ib-svr46
smg-ib-svr033
Loaded 22 of 22
ufm-appliance-5752c2
smg-ib-svr027
smg-ib-svr032
smg-ib-apl009-gen2
smg-ib-svr031
smg-ib-sw32
smg-ib-ola001-mgmt01



Map View Tab

The Network Map "View" tab displays the fabric containing all nodes (e.g. switches, racks including the hosts, etc).

If your fabric consists of more than 500 nodes, please note that:

- The "View" tab will show only the switches in your fabric. Therefore, "Expand all racks" and "Rack filter" functions will be disabled.
- Link analysis will be disabled.

To have a better experience in this instance, you can switch to the "Zoom In" tab.

Map Zoom In Tab

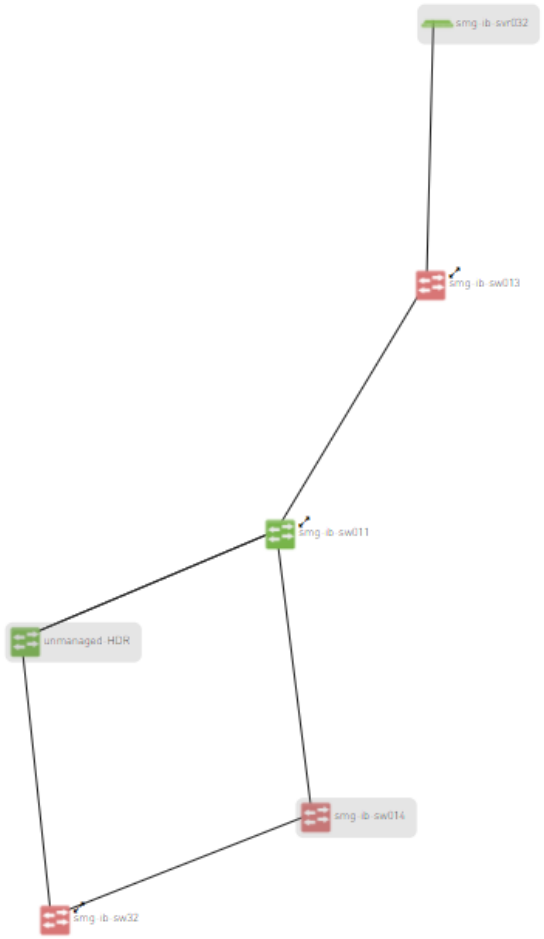
The Network Map "Zoom In" tab displays only the selected nodes from the dropdown menu above the map view and the nodes directly connected to the selected nodes.

Network Map

Views: Default

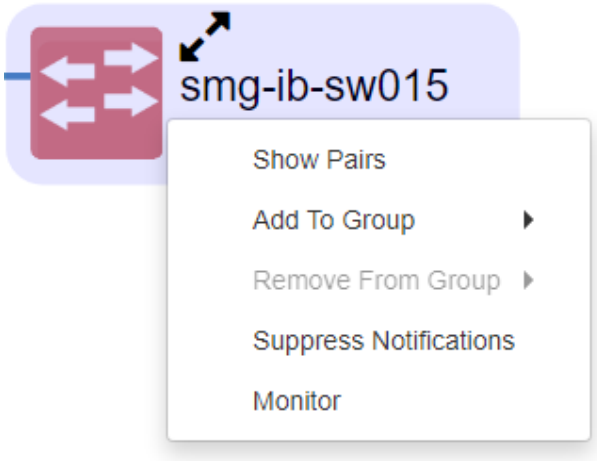
unmanaged-HDR × smg-ib-sw014 × smg-ib-svr032 ×

View | Zoom In



If some switches still have hidden connected nodes, you will see the following icon:

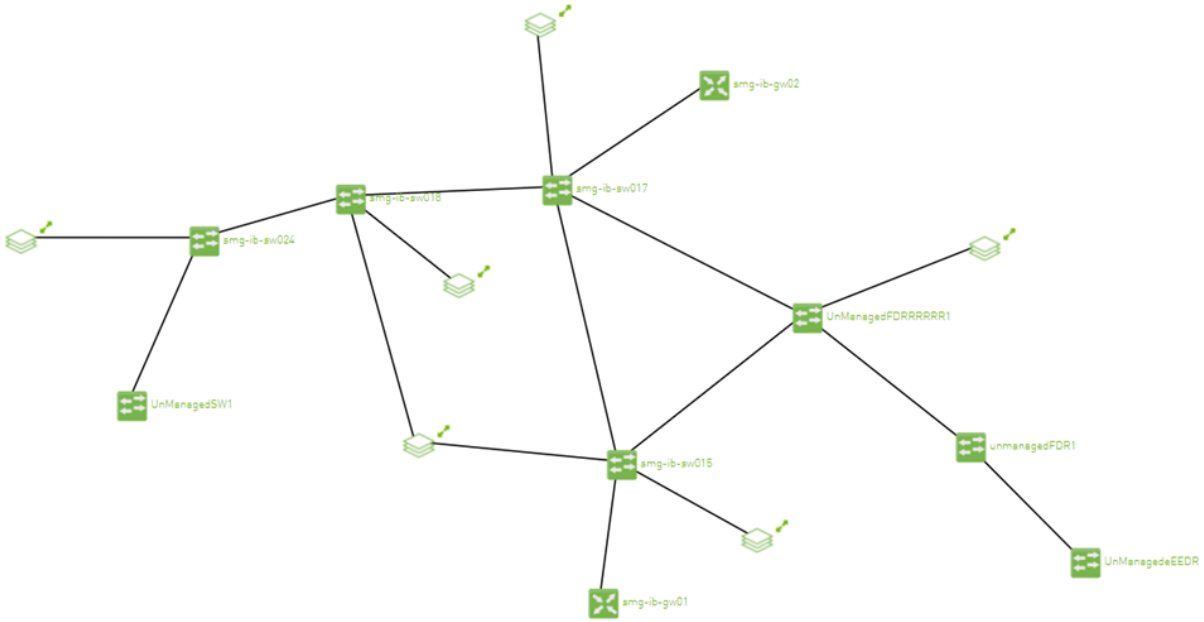
To reveal the hidden nodes connected to this switch, you can right-click it and select "Show Pairs" which adds this switch to the selected nodes list and shows the direct connected nodes to this switch.



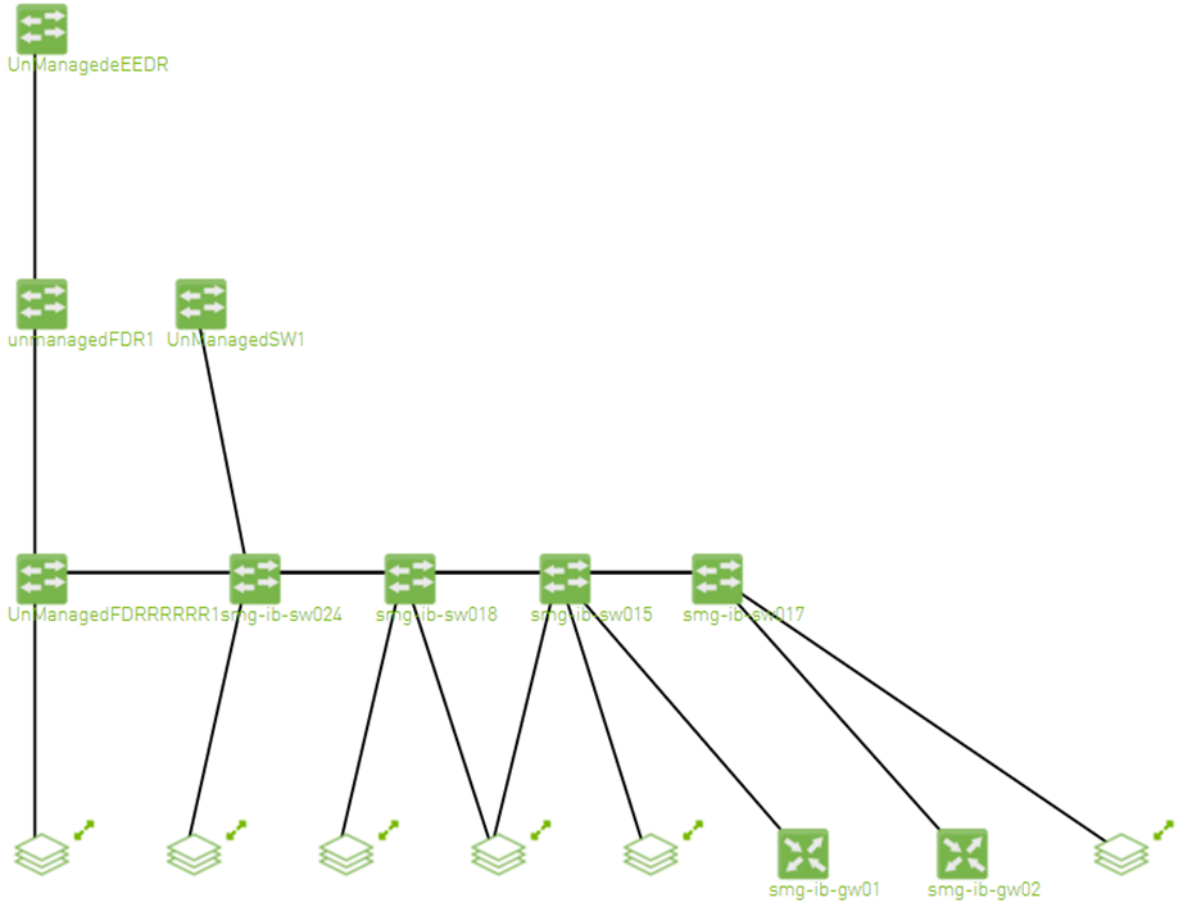
Map Layouts

Layout controls nodes positions in the map. UFM network map supports two types of layouts:

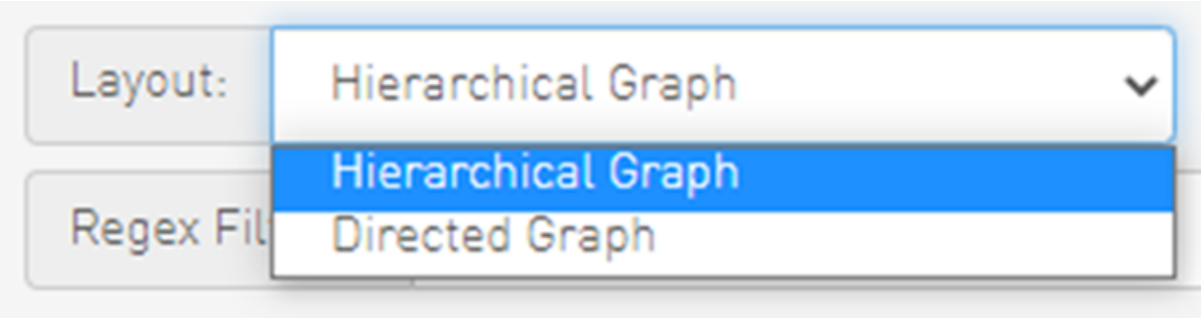
- Directed layout: the nodes are distributed depending on the connections between them so that the connected nodes will be near each other without conflict.



- Hierarchical layout: the nodes are distributed as layers; each layer will contain nodes that have the same level value.



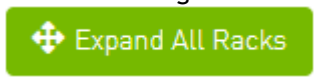
You can switch between layouts from the dropdown menu located above the Network Map view.



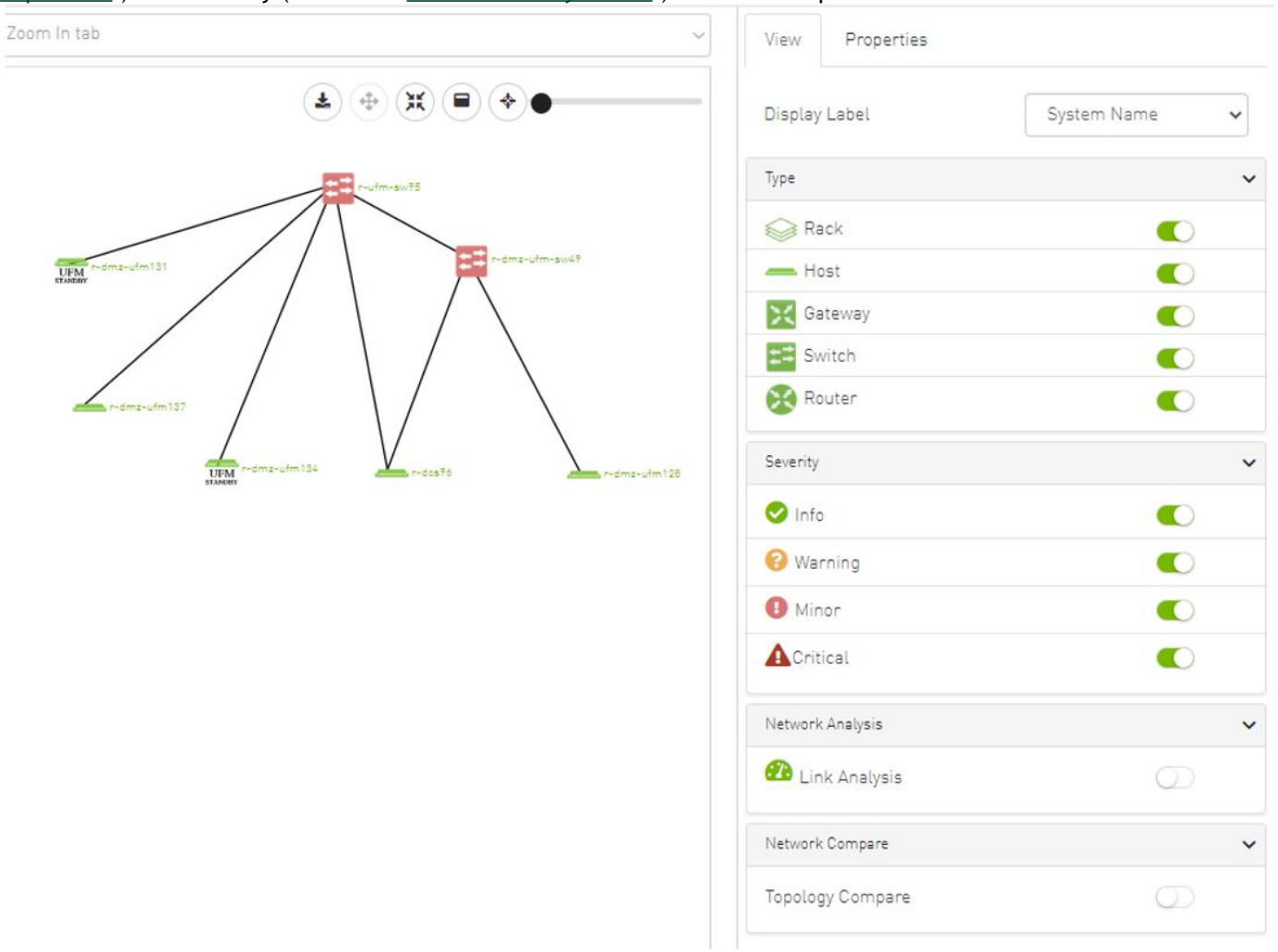
The default layout for small fabric (less than 30 nodes) is hierarchical and for large fabric is directed.

Information View Tab

- Enables searching for one or more elements in the map, by typing either their name or their GUID in the Search field. Note that the search mechanism is **not** case-sensitive.
- Enables displaying the elements either by their name, GUID, or IP.
- Enables viewing all hosts of all racks in the fabric using the "Expand All Racks" button.



- Enables customizing the view of the map by filtering for certain elements to appear in the map using the Type (see table "Network Map Components") and Severity (see table "Device Severity Levels") filters. Example:



Device Severity Levels

Component	Description
	Info
	Critical
	Minor
	Warning

Link Analysis

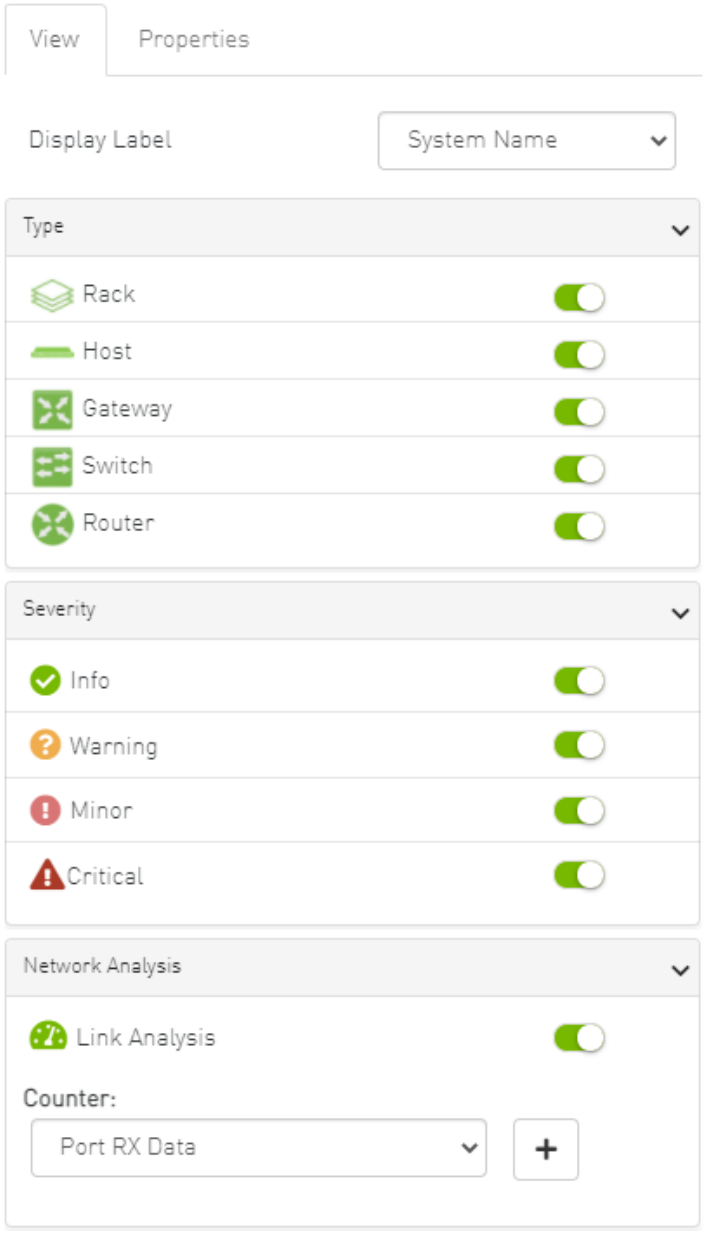
Link analysis allows the user to display the link analytics according to a selected static counter, and define the conditions on which the analysis is based. The links are colored according to the specified conditions. It is possible to define up to five conditions per counter.

The counter’s conditions are applied on four values:

- The source values of the selected counter
- The destination value of the selected counter
- The source value of the opposite of the selected counter
- The destination value of the opposite of the selected counter

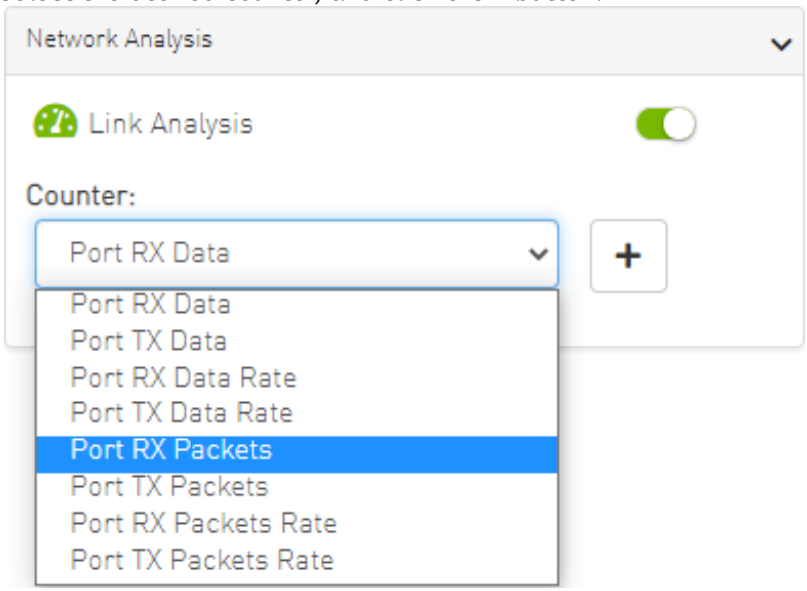
The worst matched value between these four is taken into consideration.

The "Network Analysis" section on the right side under the View tab contains a radio button to enable/disable the link analysis.

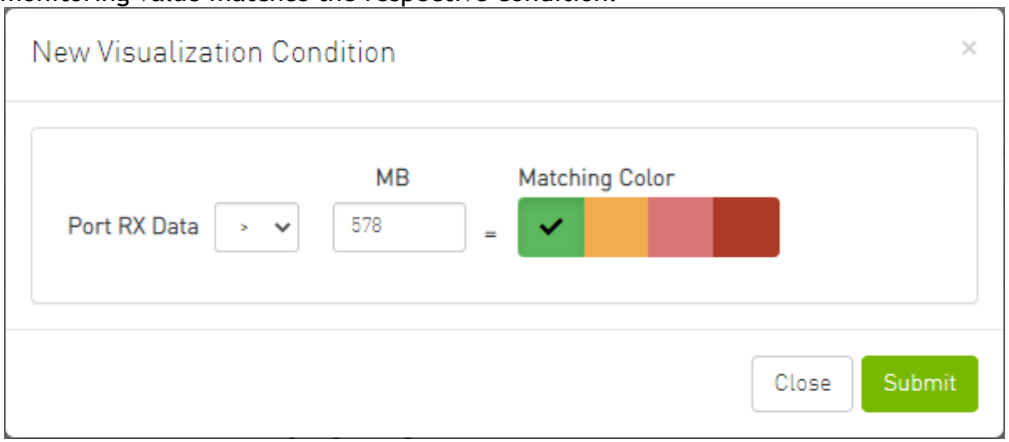

View
Properties
Display Label
System Name
Type
Rack
Host
Gateway
Switch
Router
Severity
Info
Warning
Minor
Critical
Network Analysis
Link Analysis
Counter:
Port RX Data

To define a condition:

1. Select the desired counter, and click the + button.



2. Select the appropriate operator, and define the desired threshold and color on the form that pops up. This color is applied on the link if the link monitoring value matches the respective condition.

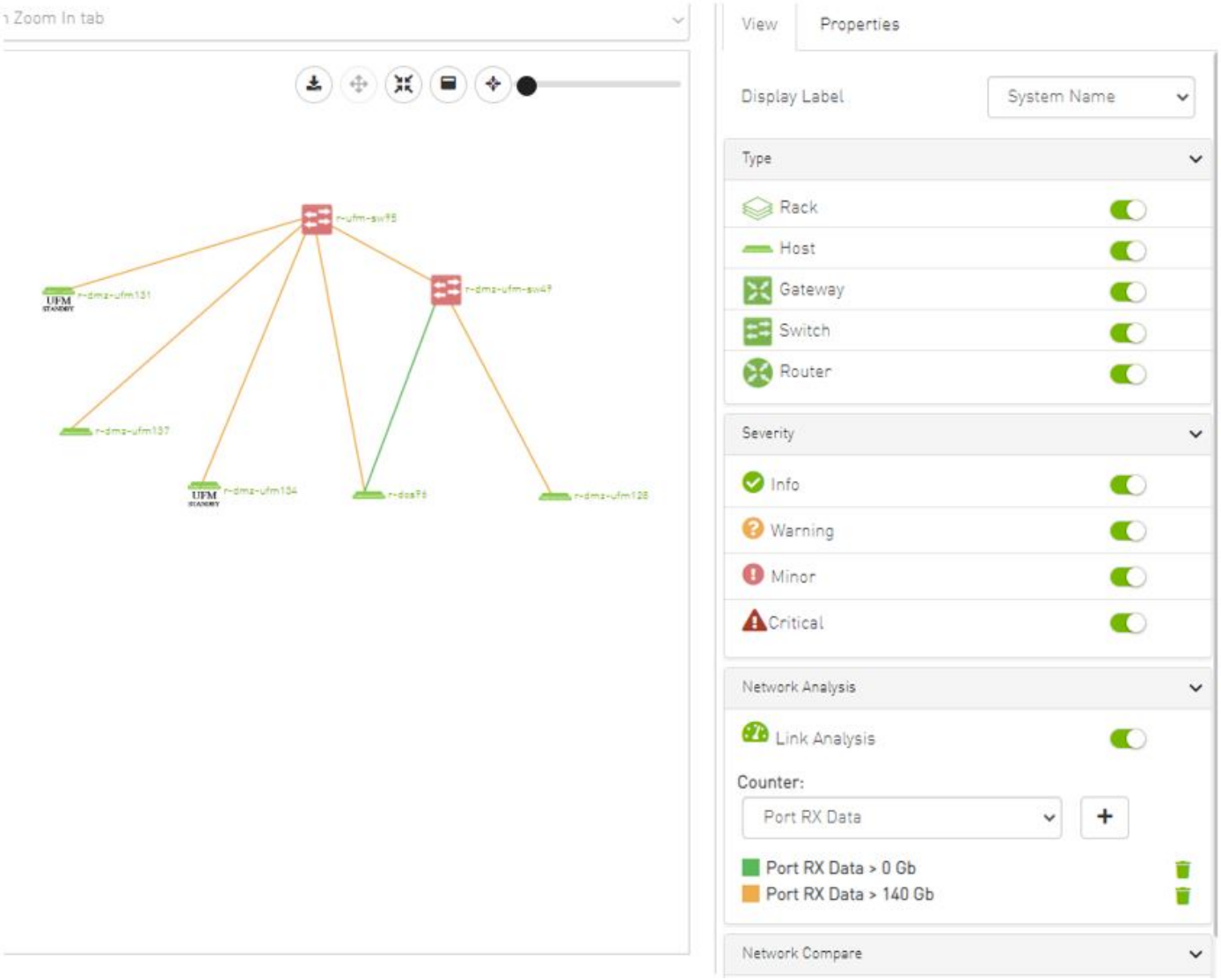


⚠ The colors are sorted from the lowest to the highest priority (i.e from left to right, green to red).

⚠ The counter's conditions are sorted based on the threshold values:
- Ascending if the operator is greater than (>)
- Descending if the operator is smaller than (<)

Last matched condition's color are taken into consideration in the link coloring.

3. Once the condition is set, the network map lights up the links that meet your condition.


Zoom In tab
r-ufm-sw95
r-dmz-ufm-sw47
r-dmz-ufm131
r-dmz-ufm137
r-dmz-ufm134
r-dgx96
r-dmz-ufm128
View
Properties
Display Label
System Name
Type
Rack
Host
Gateway
Switch
Router
Severity
Info
Warning
Minor
Critical
Network Analysis
Link Analysis
Counter:
Port RX Data
Port RX Data > 0 Gb
Port RX Data > 140 Gb
Network Compare

⚠ Note how the added conditions are listed in the Network Analysis section, if Link Analysis is enabled, and they are colored accordingly.

View | Properties

Link 1

Link/Port Properties

Property	Source	Destination
System GUID	0x0002c903007b78b0	0xb8599f0300fc6de4
Port	1	3
MTU	4096	4096
Width	4X	4X
Speed	FDR	FDR
Port RX Data	20379.85 Gb	5.9 Gb
Port TX Data	18.05 Gb	6134.55 Gb
Port RX Data Rate	0 Gb/s	0 Gb/s
Port TX Data Rate	0 Gb/s	0 Gb/s
Port RX Packets	1285841763 Packets	7796207 Packets
Port TX Packets	22720574 Packets	386937725 Packets
Port RX Packets Rate	2.9 Packets/s	2.9 Packets/s
Port TX Packets Rate	2.9 Packets/s	2.9 Packets/s

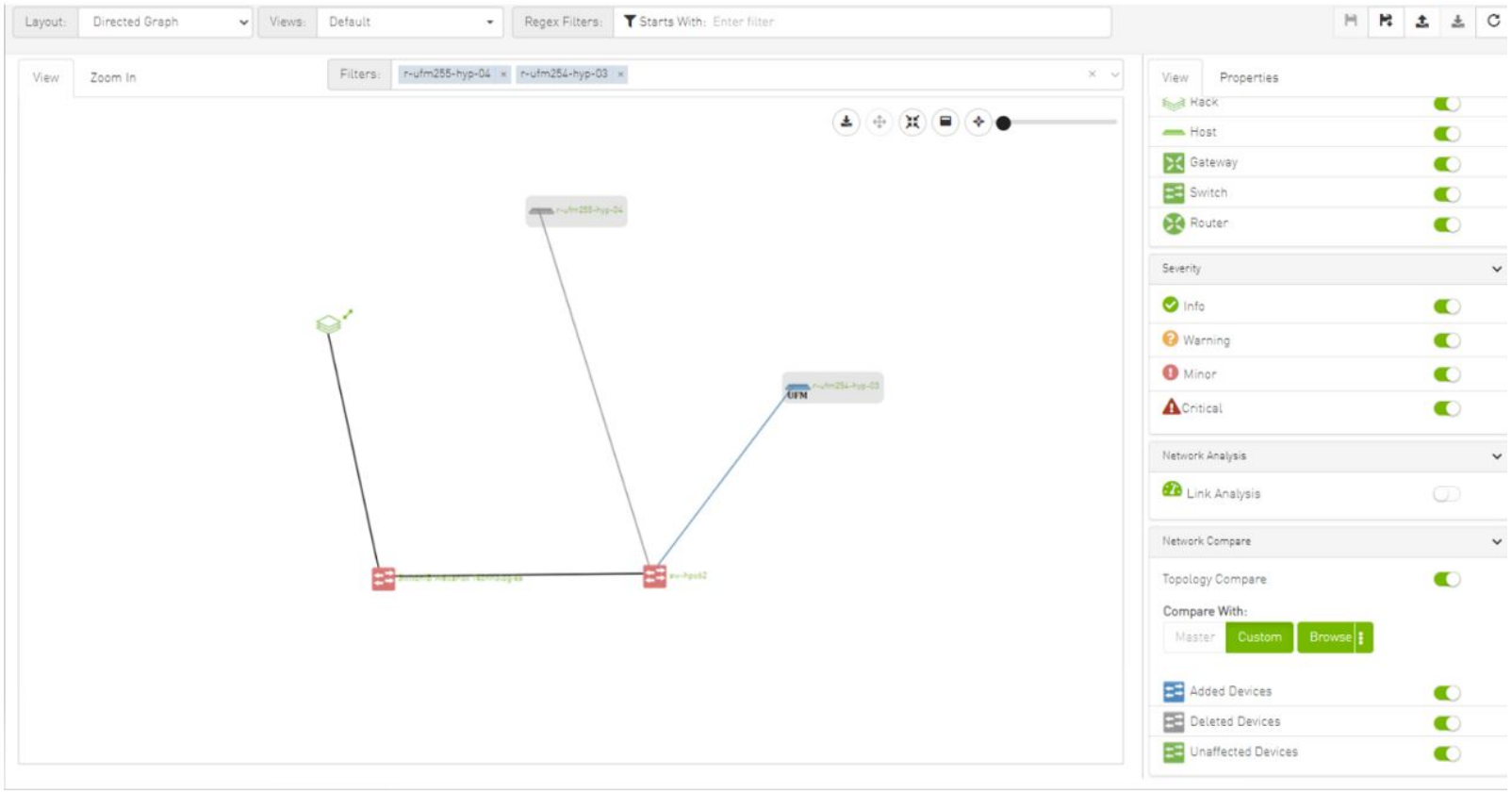
Cable Info

⚠ Notice how the monitored counter is presented in boldface, and the background color is presented with the worst matched condition.

Please note that if the current layout and view are saved, the defined conditions are saved inside the view being saved.

Topology Compare

It is possible to enable the Topology Compare feature from the View tab in the right-hand pane. When the radio button is enabled, it is possible to compare the current topology with the master topology or with a custom topology whose `.topo` file you may upload.



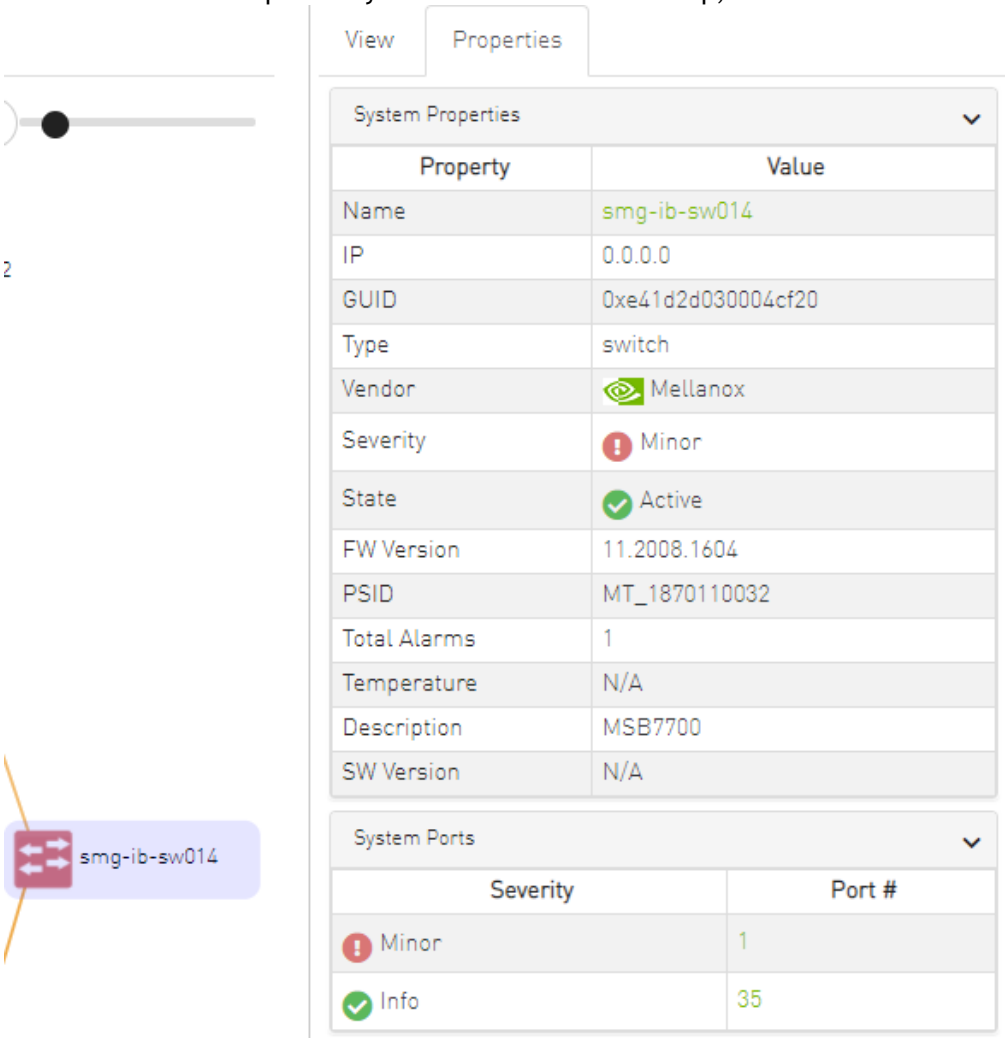
Topology compare key:

- A blue node signifies an added node
- A gray host signifies a deleted node
- A gray and black line signifies that some links were deleted and others were unchanged
- A gray and blue line signifies that some links were deleted, and others were added
- A gray, blue, and black line signifies that some links were deleted, some were added, and some were unchanged

- A blue and black line signifies that some links were added, and some were unchanged

Properties Tab

- Provides details on a specific system selected from the map, as shown in the following example:



- Provides link/port properties and cable info on a specific link selected from the map, including destination and source ports, as shown in the following example:

View | Properties

Link 1

Collect System Dump

Link/Port Properties

Property	Source	Destination
System GUID	0x0008f105002020fb	0x248a070300f88fe0
Port	18	1
MTU	4096	4096
Width	4X	4X
Speed	EDR	EDR
Port RX Data	614 MB	164 MB
Port TX Data	164 MB	614 MB
Port RX Data Rate	0 MB/s	0 MB/s
Port TX Data Rate	0 MB/s	0 MB/s
Port RX Packets	1662888 Packets	597647 Packets
Port TX Packets	597646 Packets	1662723 Packets
Port RX Packets Rate	0.45 Packets/s	0.25 Packets/s
Port TX Packets Rate	0.25 Packets/s	0.45 Packets/s

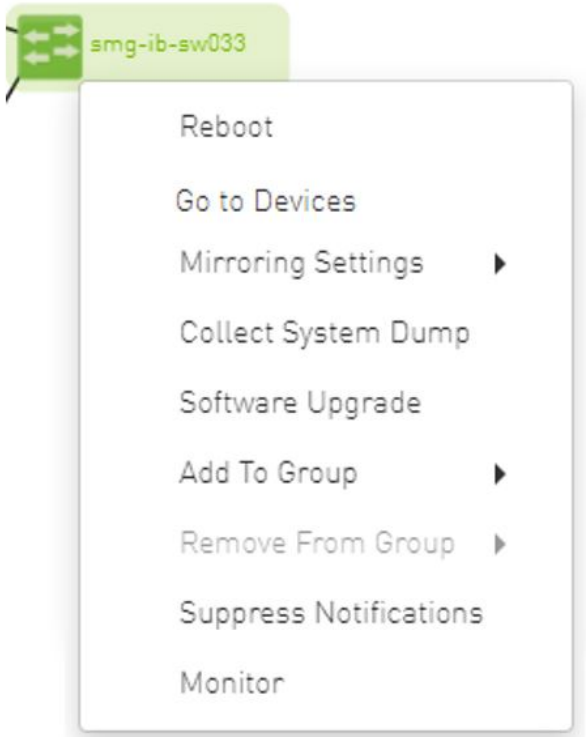
Cable Info

Property	Value
Part Number	MCP1600-E00A
Length	1 m
Serial Number	MT1714VS00778
Identifier	QSFP+
Technology	Copper cable- unequalized
Revision	A2

Network Map Elements Actions

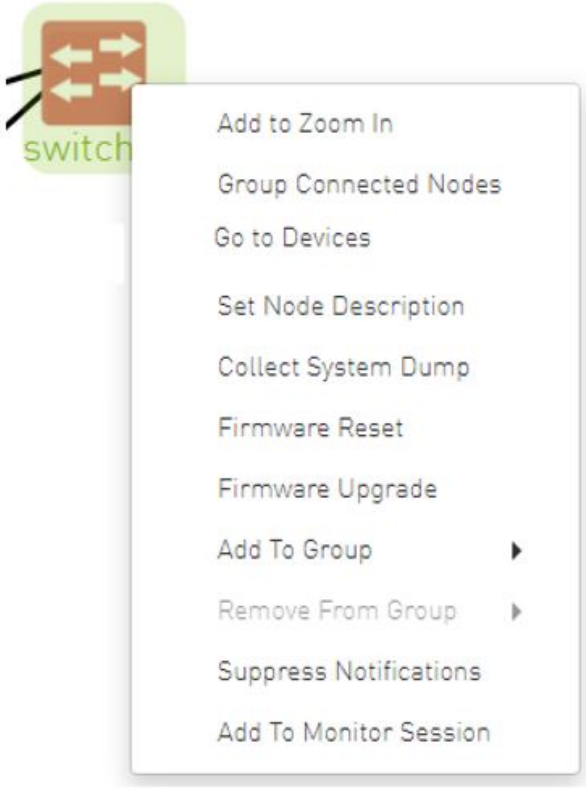
In the Network Map, a right-click on any of the elements enables performing a set of actions depending on the element type and its capabilities. See the list of available actions for each element type in the tables below.

Supported Actions for Internally Managed Switches



Element Type	Supported Actions	Description
Managed Switch	Reboot	Reboot the switch software
	Mirroring Settings	Set the mirroring configuration for the switch
	Collect System Dump	Collect system dump from the device
	Software Upgrade	Perform switch software upgrade
	Add to Group	Add switch to logical group
	Remove from Group	Remove switch from logical group
	Suppress Notification	Suppress all event notifications for the switch
	Monitor	Configure and activate switch monitoring
	Go to Devices	Go to devices page and select the device

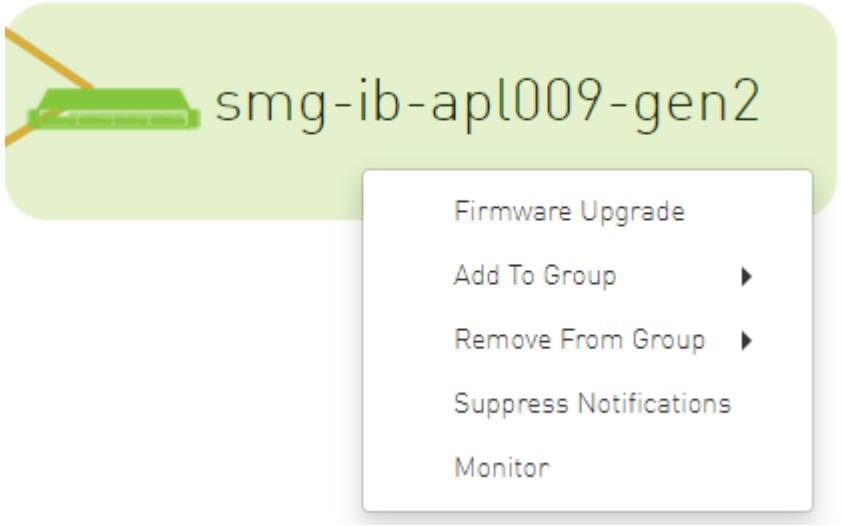
Supported Actions for Externally Managed Switches



Element Type	Supported Actions	Description
Externally Managed Switch	Set Node Description	Sets description for specific node
	Firmware Reset	Perform switch firmware reset

Element Type	Supported Actions	Description
	Firmware Upgrade	Perform switch firmware upgrade
	Add to Group	Add switch to logical group
	Remove from Group	Remove switch from logical group
	Suppress Notification	Suppress all event notifications for the switch
	Monitor	Configure and activate switch monitoring
	Go To Devices	Go to devices page and select the device

Supported Actions for Hosts



Element Type	Supported Actions	Description
Hosts	Firmware Upgrade	Perform switch firmware upgrade
	Add to Group	Add host to logical group
	Remove from Group	Remove host from logical group
	Suppress Notification	Suppress all event notifications for the host
	Monitor	Configure and activate host monitoring

Managed Elements

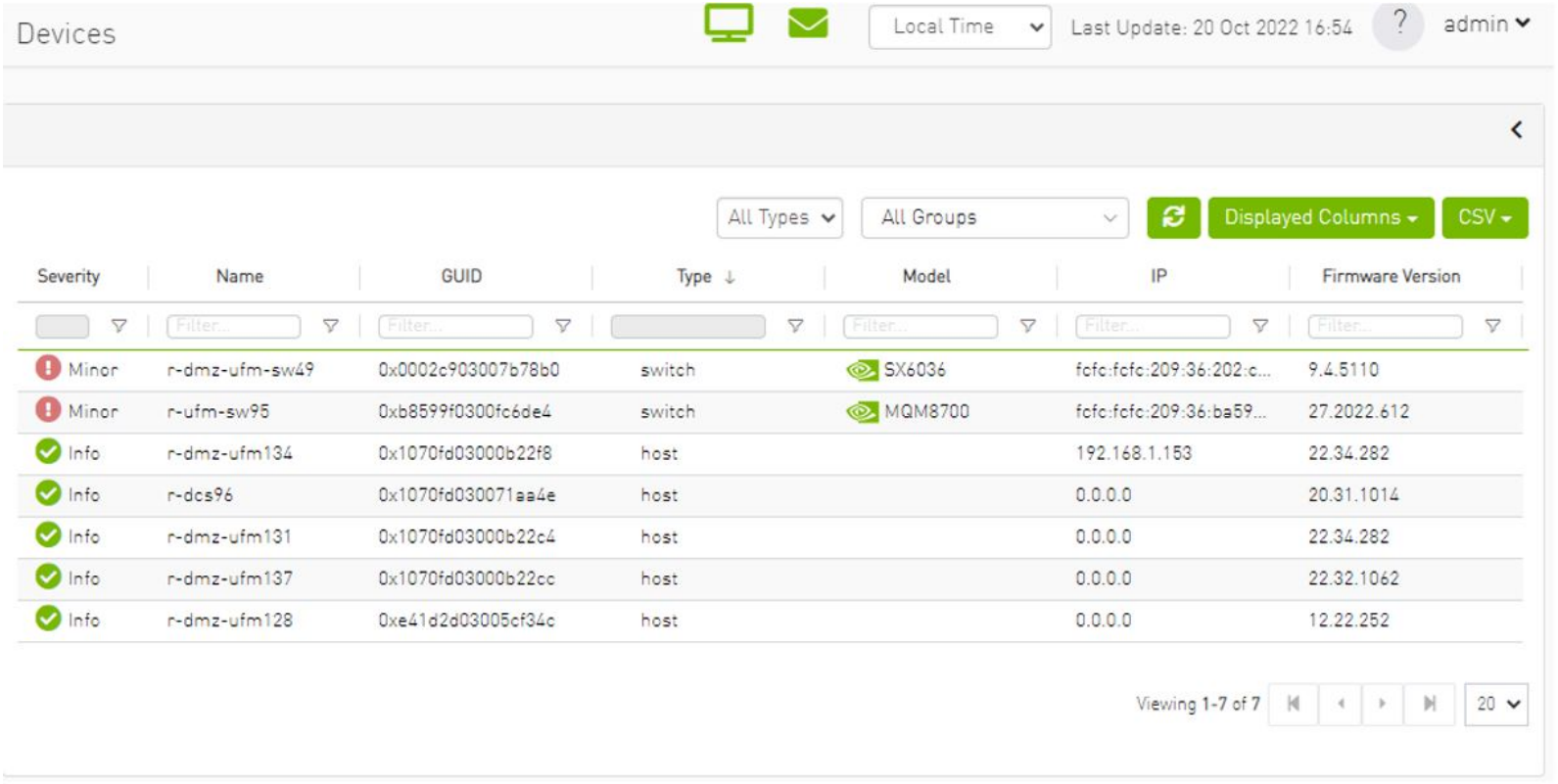
The UFM **Managed Elements** window allows you to obtain information on the fabric physical elements, such as devices, ports and cables.

> ⚠ All information provided in a tabular format in UFM web UI can be exported into a CSV file.

- Devices Window
- Ports Window
- Virtual Ports Window
- Unhealthy Ports Window
- Cables Window
- Groups Window
- Inventory Window
- PKeys Window
- HCAs Window

Devices Window

The Devices window shows data pertaining to the physical devices in a tabular format.



Devices Window Data

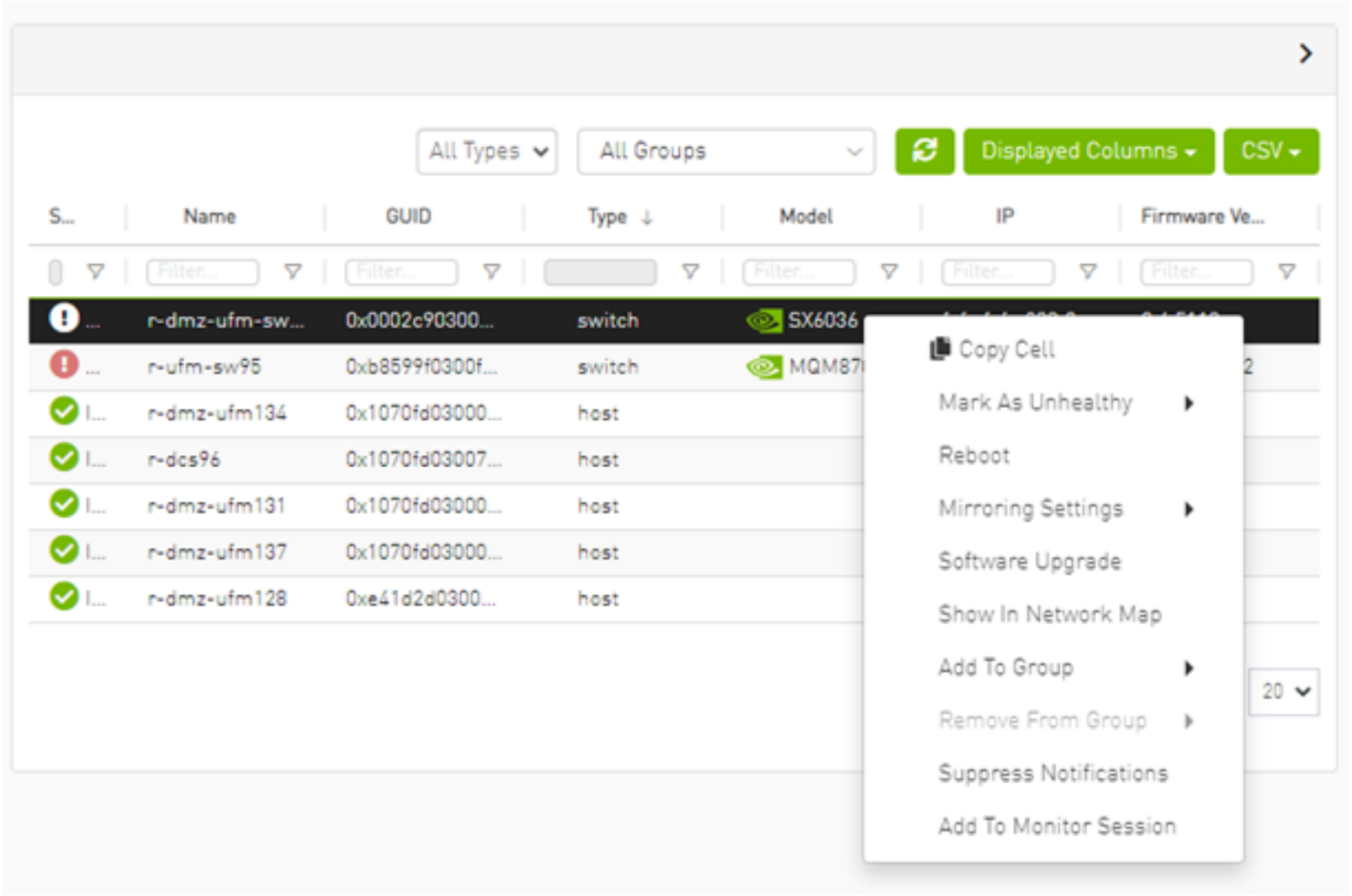
Data Type	Description
Health	Health of the device reflecting the highest alarm severity. Please refer to the Health States table.
Name	Name of the device If UFM Agent is running on a device, the following icon will appear next to the device name:
GUID	System GUID of the device
Type	Type of the device: switch, node, IB router, and getaway
IP	IP address of the device
Vendor	The vendor of the device
Firmware Version	The firmware version installed on the device

Health States

Icon	Name	Description
	Normal	Information/notification displayed during normal operating state or a normal system event.
	Critical	Critical means that the operation of the system or a system component fails.
	Minor	Minor reflects a problem in the fabric with no failure.

Icon	Name	Description
	Warning	Warning reflects a low priority problem in the fabric with no failure. A warning is asserted when an event exceeds a predefined threshold.

A right-click on the device name displays a list of actions that can be performed on it.



Devices Actions

Action	Description
Firmware Upgrade	Perform a firmware upgrade on the selected device
Firmware Reset	Reboot the device. This action is only applicable to unmanaged hosts (servers).
Set Node Description	Configure a description to this node
Collect System Dump	Collect the system dump log for a specific device
Add to Group	Add the selected device to a devices group
Remove from Group	Remove the selected device from a devices group
Suppress Notifications	Suppress all event notifications for the device
Add to Monitor Session	Configure and activate host monitoring
Show in Network Map	Move to Zoom In tab in network map and add the selected device to filter list

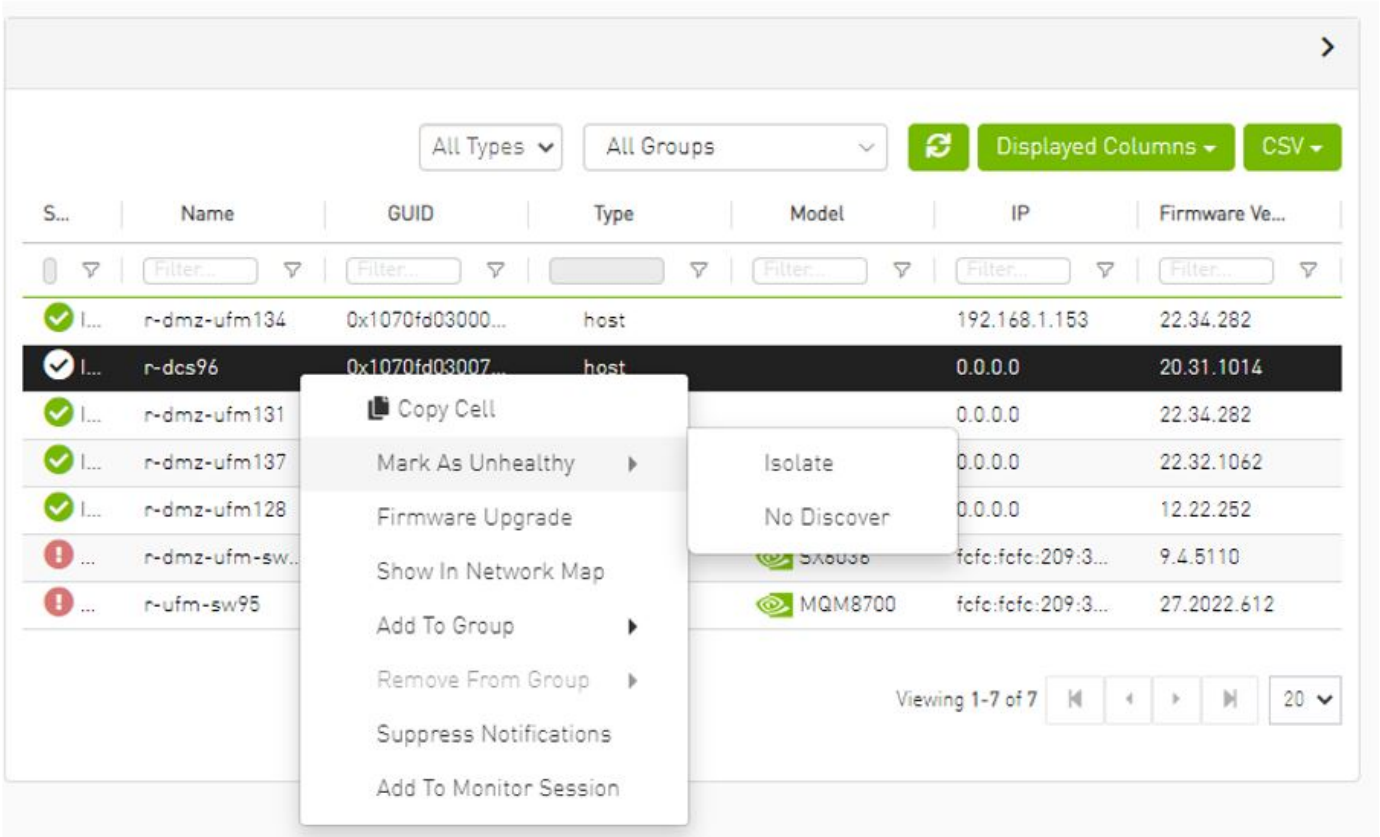
⚠ Collecting system dump for hosts, managed by UFM, is available only for hosts which are set with a valid IPv4 address and installed with MLNX_OFED.

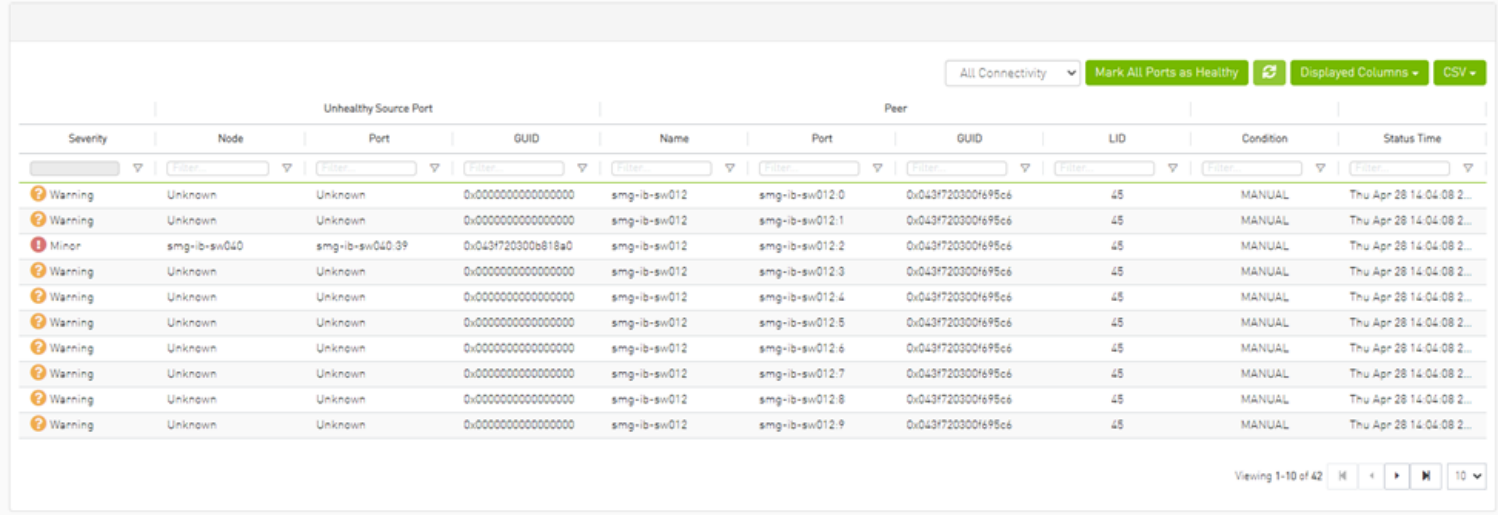
Mark Device as Unhealthy

From the Devices table, it is possible to mark devices as healthy or unhealthy using the context menu (right-click).

There are two options for marking a device as unhealthy:

- Isolate
- No Discover


All Types
All Groups
Displayed Columns
CSV
S...
Name
GUID
Type
Model
IP
Firmware Ve...
r-dmz-ufm134
0x1070fd03000...
host
192.168.1.153
22.34.282
r-dcs96
0x1070fd03007...
host
0.0.0.0
20.31.1014
r-dmz-ufm131
0.0.0.0
22.34.282
r-dmz-ufm137
0.0.0.0
22.32.1062
r-dmz-ufm128
0.0.0.0
12.22.252
r-dmz-ufm-sw...
fcfc:fcfc:209:3...
9.4.5110
r-ufm-sw95
MQM8700
fcfc:fcfc:209:3...
27.2022.612
Viewing 1-7 of 7
20
Copy Cell
Mark As Unhealthy
Isolate
No Discover
Firmware Upgrade
Show In Network Map
Add To Group
Remove From Group
Suppress Notifications
Add To Monitor Session



Server: `conf/opensm/opensm-health-policy.conf` content:

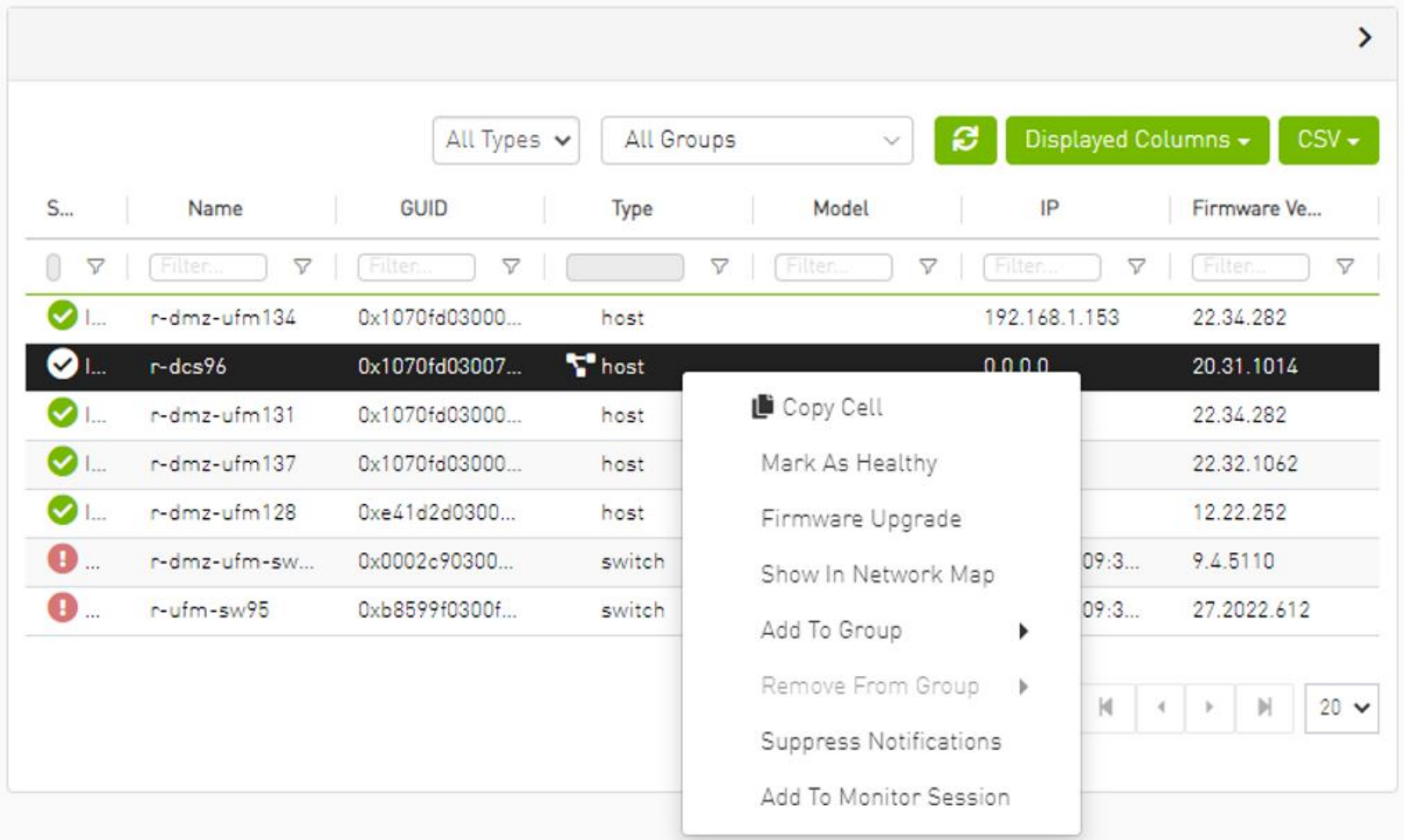
```
0xe41d2d030003e3b0 34 UNHEALTHY isolate
0xe41d2d030003e3b0 19 UNHEALTHY isolate
0xe41d2d030003e3b0 3 UNHEALTHY isolate
0xe41d2d030003e3b0 26 UNHEALTHY isolate
0xe41d2d030003e3b0 0 UNHEALTHY isolate
0xe41d2d030003e3b0 27 UNHEALTHY isolate
0xe41d2d030003e3b0 7 UNHEALTHY isolate
0xe41d2d030003e3b0 10 UNHEALTHY isolate
0xe41d2d030003e3b0 11 UNHEALTHY isolate
0xe41d2d030003e3b0 22 UNHEALTHY isolate
0xe41d2d030003e3b0 18 UNHEALTHY isolate
0xe41d2d030003e3b0 29 UNHEALTHY isolate
0xe41d2d030003e3b0 8 UNHEALTHY isolate
0xe41d2d030003e3b0 5 UNHEALTHY isolate
0xe41d2d030003e3b0 17 UNHEALTHY isolate
0xe41d2d030003e3b0 23 UNHEALTHY isolate
0xe41d2d030003e3b0 15 UNHEALTHY isolate
0xe41d2d030003e3b0 24 UNHEALTHY isolate
```

```
0xe41d2d030003e3b0 2 UNHEALTHY isolate
0xe41d2d030003e3b0 16 UNHEALTHY isolate
0xe41d2d030003e3b0 13 UNHEALTHY isolate
0xe41d2d030003e3b0 14 UNHEALTHY isolate
0xe41d2d030003e3b0 32 UNHEALTHY isolate
0xe41d2d030003e3b0 33 UNHEALTHY isolate
0xe41d2d030003e3b0 35 UNHEALTHY isolate
0xe41d2d030003e3b0 20 UNHEALTHY isolate
0xe41d2d030003e3b0 21 UNHEALTHY isolate
0xe41d2d030003e3b0 28 UNHEALTHY isolate
0xe41d2d030003e3b0 1 UNHEALTHY isolate
0xe41d2d030003e3b0 9 UNHEALTHY isolate
0xe41d2d030003e3b0 4 UNHEALTHY isolate
0xe41d2d030003e3b0 31 UNHEALTHY isolate
0xe41d2d030003e3b0 30 UNHEALTHY isolate
0xe41d2d030003e3b0 36 UNHEALTHY isolate
0xe41d2d030003e3b0 12 UNHEALTHY isolate
0xe41d2d030003e3b0 25 UNHEALTHY isolate
0xe41d2d030003e3b0 6 UNHEALTHY isolate
```

`/opt/ufm/files/log/opensm-unhealthy-ports.dump` content:

```
# NodeGUID, PortNum, NodeDesc, PeerNodeGUID, PeerPortNum,  PeerNodeDesc, {BadCond1, BadCond2, ...}, TimeStamp
0x248a070300ef19a0, 23, "MF0;r-hyp-sw-01:MSB7700/U1", 0x7cfe900300292356, 1, "ufm-host86 HCA-1", {FLAPPING}, Thu Nov 18 15:05:02 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 0, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 1, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 2, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 3, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 4, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 5, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 6, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 7, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 8, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 9, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 10, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 11, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 12, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 13, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 14, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 15, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 16, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 17, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 18, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 19, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 20, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 21, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 22, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 23, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 24, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 25, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 26, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 27, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 28, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 29, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 30, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 31, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 32, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 33, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 34, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 35, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
0x0000000000000000, 0, "(null)", 0xe41d2d030003e3b0, 36, "(null)", {MANUAL}, Thu Nov 18 16:27:34 2021
~
```

Mark Device as Healthy



Server `/opt/ufm/files/conf/opensm/opensm-health-policy.conf` content:

```
0xe41d2d030003e3b0 15 HEALTHY
0xe41d2d030003e3b0 25 HEALTHY
0xe41d2d030003e3b0 35 HEALTHY
0xe41d2d030003e3b0 0 HEALTHY
0xe41d2d030003e3b0 11 HEALTHY
0xe41d2d030003e3b0 21 HEALTHY
0xe41d2d030003e3b0 28 HEALTHY
0xe41d2d030003e3b0 7 HEALTHY
0xe41d2d030003e3b0 17 HEALTHY
0xe41d2d030003e3b0 14 HEALTHY
0xe41d2d030003e3b0 24 HEALTHY
0xe41d2d030003e3b0 34 HEALTHY
0xe41d2d030003e3b0 3 HEALTHY
0xe41d2d030003e3b0 10 HEALTHY
0xe41d2d030003e3b0 20 HEALTHY
0xe41d2d030003e3b0 31 HEALTHY
0xe41d2d030003e3b0 6 HEALTHY
0xe41d2d030003e3b0 16 HEALTHY
0xe41d2d030003e3b0 27 HEALTHY
0xe41d2d030003e3b0 2 HEALTHY
0xe41d2d030003e3b0 13 HEALTHY
0xe41d2d030003e3b0 23 HEALTHY
0xe41d2d030003e3b0 33 HEALTHY
0xe41d2d030003e3b0 30 HEALTHY
0xe41d2d030003e3b0 9 HEALTHY
0xe41d2d030003e3b0 19 HEALTHY
0xe41d2d030003e3b0 26 HEALTHY
0xe41d2d030003e3b0 36 HEALTHY
0xe41d2d030003e3b0 5 HEALTHY
0xe41d2d030003e3b0 12 HEALTHY
0xe41d2d030003e3b0 22 HEALTHY
0xe41d2d030003e3b0 32 HEALTHY
0xe41d2d030003e3b0 1 HEALTHY
0xe41d2d030003e3b0 8 HEALTHY
0xe41d2d030003e3b0 18 HEALTHY
0xe41d2d030003e3b0 29 HEALTHY
0xe41d2d030003e3b0 4 HEALTHY
```

`/opt/ufm/files/log/opensm-unhealthy-ports.dump` content:

```
# NodeGUID, PortNum, NodeDesc, PeerNodeGUID, PeerPortNum, PeerNodeDesc, {BadCond1, BadCond2, ...}, timestamp
```

Upgrading Software and Firmware for Hosts and Externally Managed Switches

Software/Firmware Upgrade via FTP

Software and firmware upgrade over FTP is enabled by the UFM Agent. UFM invokes the Software/Firmware Upgrade procedure locally on switches or on hosts. The procedure copies the new software/firmware file from the defined storage location and performs the operation on the device. UFM sends the set of attributes required for performing the software/firmware upgrade to the agent.

The attributes are:

- File Transfer Protocol – default FTP
  - The Software/Firmware upgrade on InfiniScale III ASIC-based switches supports FTP protocol for transmitting files to the local machine.
  - The Software/Firmware upgrade on InfiniScale IV-based switches and hosts supports TFTP and protocols for transmitting files to the local machine.
- IP address of file-storage server
- Path to the software/firmware image location
  The software/firmware image files should be placed according to the required structure under the defined image storage location. Please refer to section Devices Window.
- File-storage server access credentials (User/Password)

In-Band Firmware Upgrade

You can perform in-band firmware upgrades for externally managed switches and HCAs. This upgrade procedure does not require the UFM Agent or IP connectivity, but it does require current PSID recognition. Please refer to section PSID and Firmware Version In-Band Discovery. This feature requires that the Mellanox Firmware Toolkit (MFT), which is included in the UFM package, is installed on the UFM server. UFM uses flint from the MFT for in-band firmware burning.

Before upgrading, you must zip all subdirectories and then import the firmware repository to the UFM-SDN Appliance. A subdirectory should be created for each PSID and one firmware image should be placed under it. For example:

```
     MT_0D80110009
            fw-ConnectX2-rel-2_9_1000-MHQH29B-XTR_A1.bin
     MT_0F90110002
            fw-IS4-rel-7_4_2040-MIS5023Q_A1-A5.bin
```

1. Create a zip file in the above mentioned hierarchy with zip containing folders that are PSID named which contain the needed binary.
2. `# ufm firmware fetch` <download URL of the zip file e.g. scp://...>
3. `# ufm firmware import` <zip file>

Directory Structure for Software or Firmware Upgrade Over FTP

Before performing a software or firmware upgrade, you must create the following directory structure for the upgrade image. The path to the `<ftp user home>/<path>/` directory should be specified in the upgrade dialog box.

```
<ftp user home>/<path>/
     InfiniScale3 - For anafa based switches Software/Firmware upgrade images
            voltaire_fw_images.tar - firmware image file
            ibswmpr-<s/w version>.tar - software image file
    InfiniScale4 - For InfiniScale IV based switches Software/Firmware upgrade images
            firmware_2036_4036.tar - Firmware image file
            upgrade_2036_4036.tgz - Software image file
    OFED /* For host SW upgrade*/
            OFED-<OS label>.tar.bz2
    <PSID>* - For host FW upgrade
            fw_update.img
```

The <PSID> value is extracted from the `mstflint` command:

```
mstflint -d <device> q
```

The device is extracted from the `lspci` command. For example:

```
# lspci
06:00.0 InfiniBand: Mellanox Technologies MT25208 InfiniHost III Ex
# mstflint -d 06:00.0 q | grep PSID
```

```
PSID: VLT0040010001
```

PSID and Firmware Version In-Band Discovery

The device PSID and device firmware version are required for in-band firmware upgrade and for the correct functioning of Subnet Manager plugins, such as Congestion Control Manager and Lossy Configuration Management. For most devices, UFM discovers this information and displays it in the Device Properties pane. The PSID and the firmware version are discovered by the Vendor-specific MAD.

By default, the gv.cfg file value for event_plugin_option is set to (null). This means that the plugin is disabled and opensm does not send MADs to discover devices' PSID and FW version. Therefore, values for devices' PSID and FW version are taken from `ibdiagnet` output (section NODES_INFO).

The below is an example of the default value:

```
event_plugin_options = (null)
```

To enable the vendor-specific discovery by opemsm, in the gv.cfg configuration file, change the value of event_plugin_option to (--vendinfo -m 1), as shown below:

```
event_plugin_options = --vendinfo -m 1
```

If the value is set to -vendinfo -m 1, the data should be supplied by opensm, and in this case the `ibdiagnet` output is ignored.

⚠ In some firmware versions, the information above is currently not available.

Switch Management IP Address Discovery

From NVIDIA switch FM version 27.2010.3942 and up, NVIDIA switches support switch management IP address discovery using MADs. This information can be retrieved as part of `ibdiagnet` run (`ibdiagnet` output), and assigned to discover switches in UFM.

There is an option to choose the IP address of which IP protocol version that is assigned to the switch: IPv4 or IPv6.

The discovered_switch_ip_protocol key, located in the gv.cfg file in section [FabricAnalysys], is set to 4 by default. This means that the IP address of type IPv4 is assigned to the switch as its management IP address. In case this value is set to 6, the IP address of type IPv6 is assigned to the switch as its management IP address. Alternatively, you may run the command "`ufm discovered-switch-ip-version <ipv4/ipv6>`".

After changing the discover_switch_ip_protocol value in gv.cfg, the UFM Main Model needs to be restarted for the update to take effect. The discovered IP addresses for switches are not persistent in UFM – every UFM Main Model restarts the values of management IP address which is assigned from the `ibdiagnet` output.

Upgrading Server Software

The ability to update the server software is applicable only for hosts (servers) with the UFM Agent.

To upgrade the software:

1. Select a device.
2. From the right-click menu, select Software Update.
3. Enter the parameters listed in the following table.

Parameter	Description
Protocol	Update is performed via FTP protocol
IP	Enter the host IP
Path	Enter the parent directory of the FTP directory structure for the Upgrade image. The path should not be an absolute path and should not contain the first slash (/) or trailer slash.
User	Name of the host username
Password	Enter the host password

4. Click Submit to save your changes.

Upgrading Firmware

You can upgrade firmware over FTP for hosts and switches that are running the UFM Agent, or you can perform an in-band upgrade for externally managed switches and HCAs.

Before you begin the upgrade ensure that the new firmware version is in the correct location. For more information, please refer to section In-Band Firmware Upgrade.

To upgrade the firmware:

1. Select a host or server.
2. From the right-click menu, select Firmware Upgrade.
3. Select protocol In Band.
4. For upgrade over FTP, enter the parameters listed in the following table.

Parameter	Description
IP	Enter device IP
Path	Enter the parent directory of the FTP directory structure for the Upgrade image. The path should not be an absolute path and should not contain the first slash (/) or trailer slash.
Username	Name of the host username
Password	Enter the host password

5. Click submit to save your changes.

> ⚠ The firmware upgrade takes effect only after the host or externally managed switch is restarted.

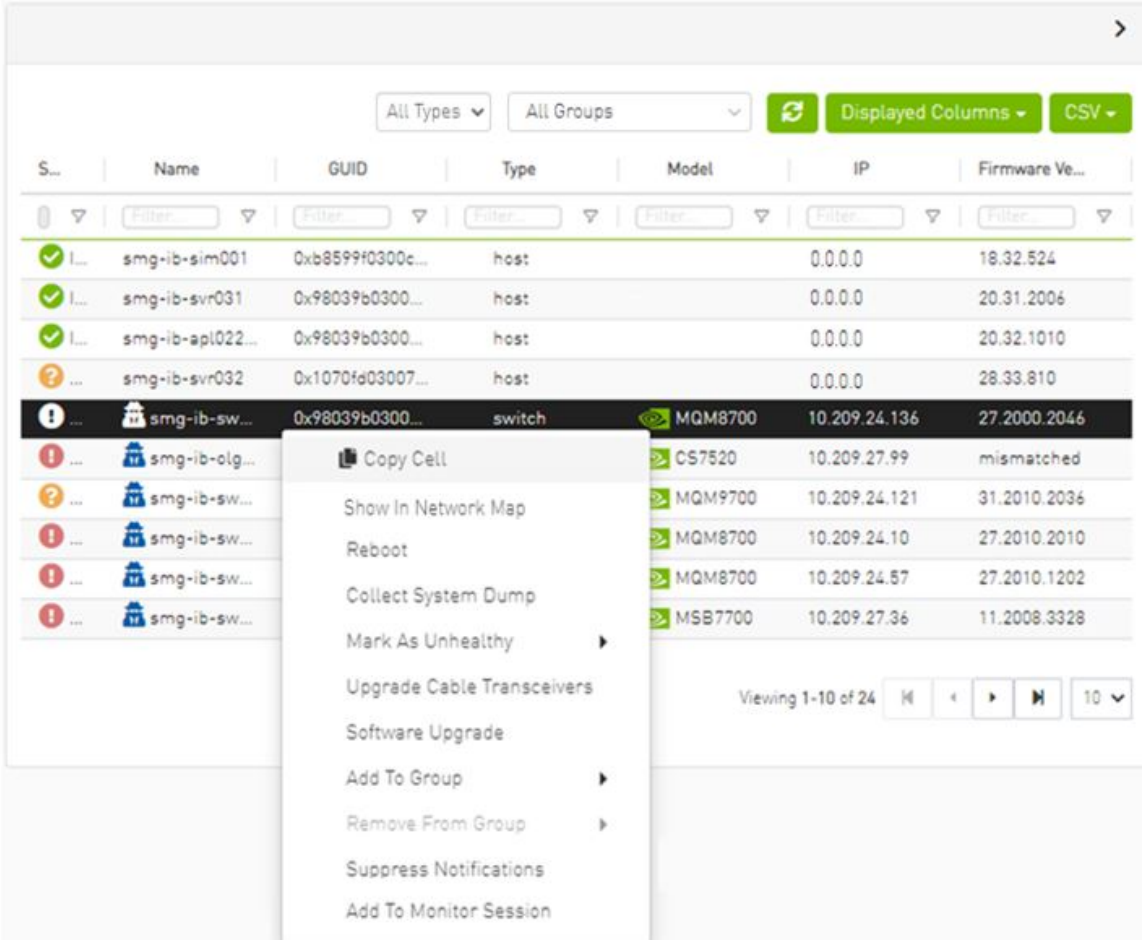
Upgrade Cables Transceivers Firmware Version

The main purpose of this feature is to add support for burning of multiple cables transceiver types on multiple devices using linkx tool which is part of flint. This needs to be done from both ends of the cable (switch and HCA/switch).

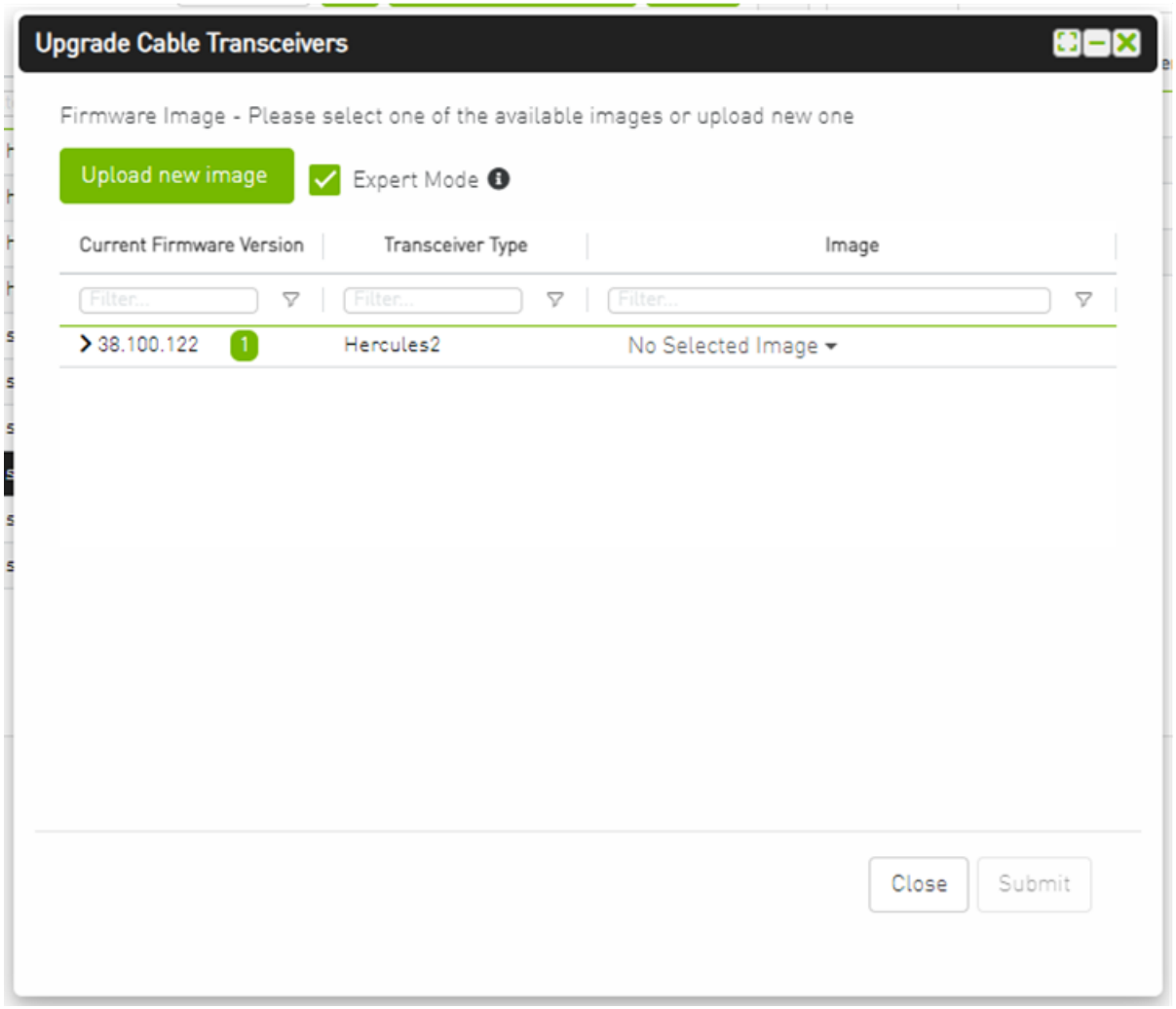
To upgrade cables transceivers FW version:

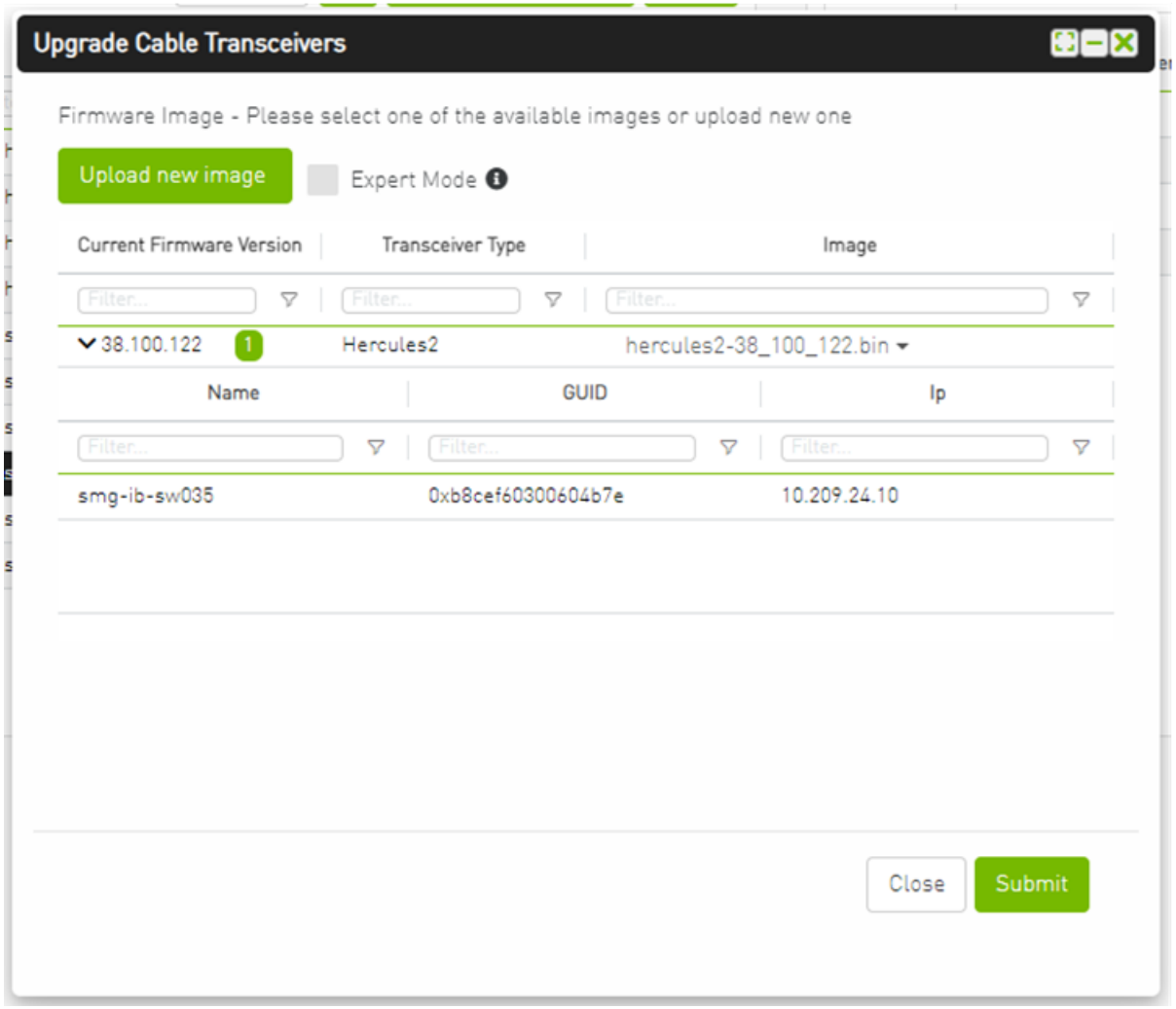
1. Navigate to managed elements page

2. select the target switches and click on Upgrade Cable Transceivers option

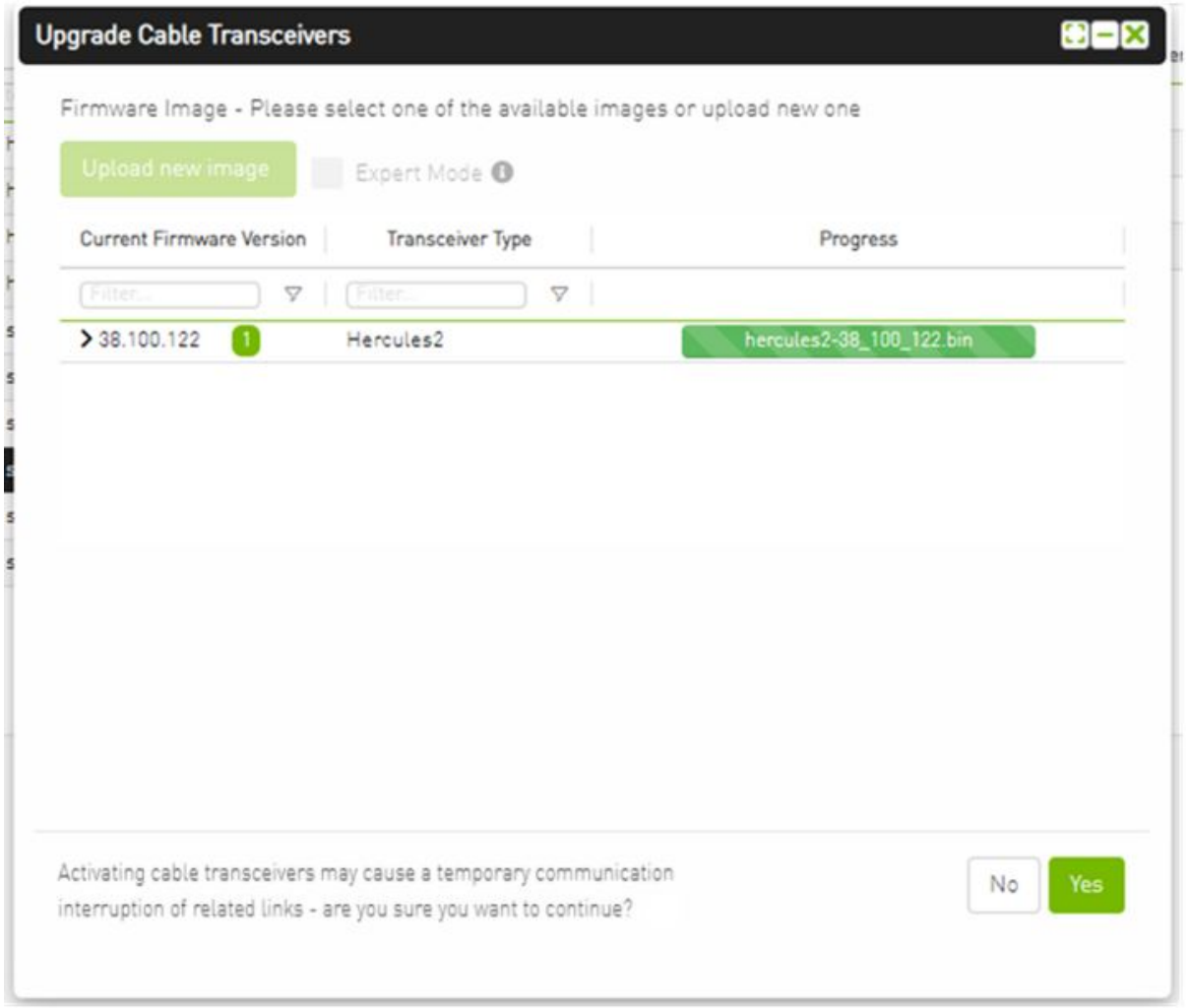


3. A model will be shown containing list of the active firmware versions for the cables of the selected switches, besides the version number, a badge will show the number of matched switches:

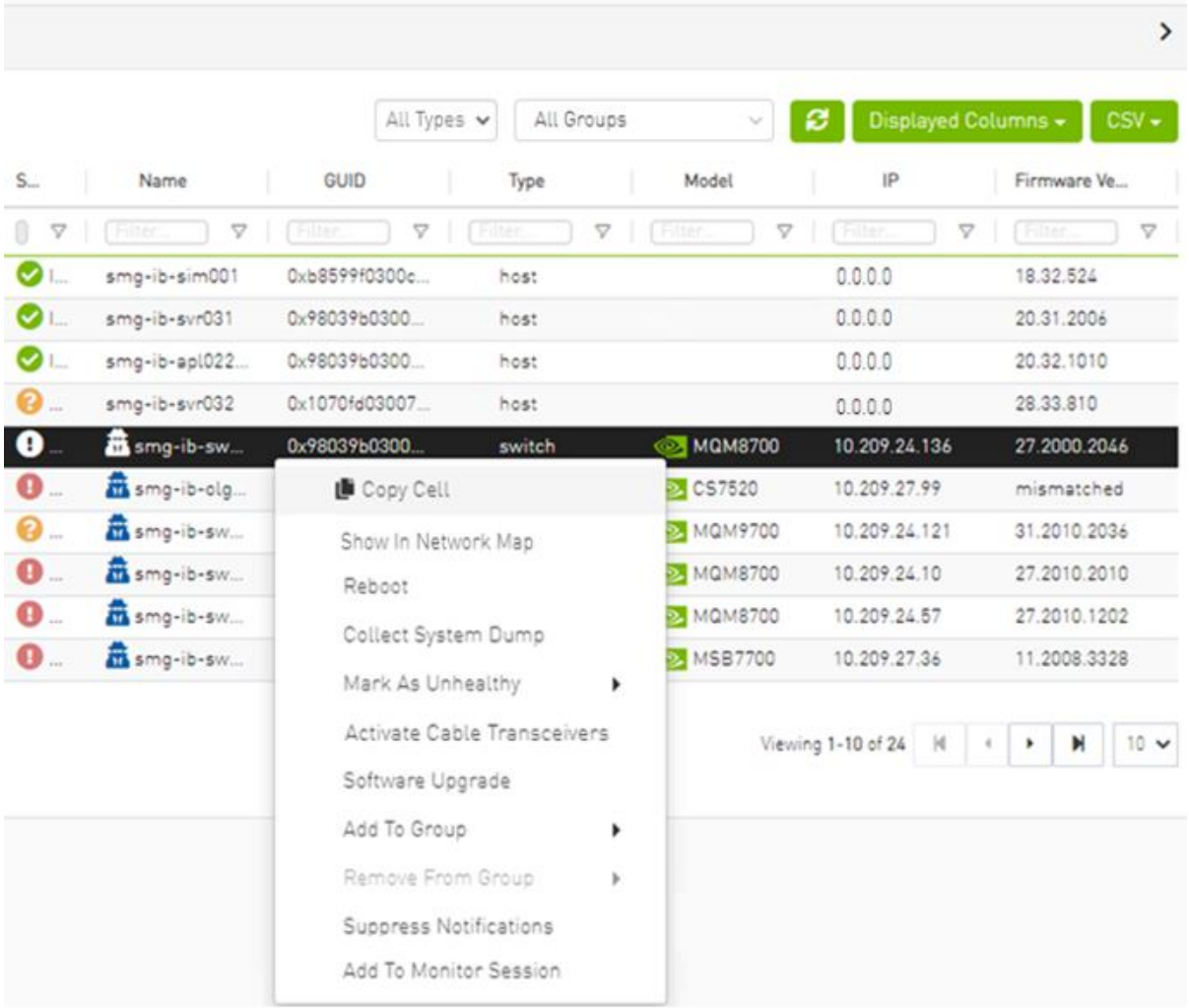

Upgrade Cable Transceivers
Firmware Image - Please select one of the available images or upload new one
Upload new image
Expert Mode
Current Firmware Version
Transceiver Type
Image
Filter...
Filter...
Filter...
38.100.122
1
Hercules2
No Selected Image
Close
Submit



4. After the user clicks Submit, the GUI will start sending the selected binaries with the relevant switches sequentially, and a model with a progress bar will be shown (this model can be minimized):

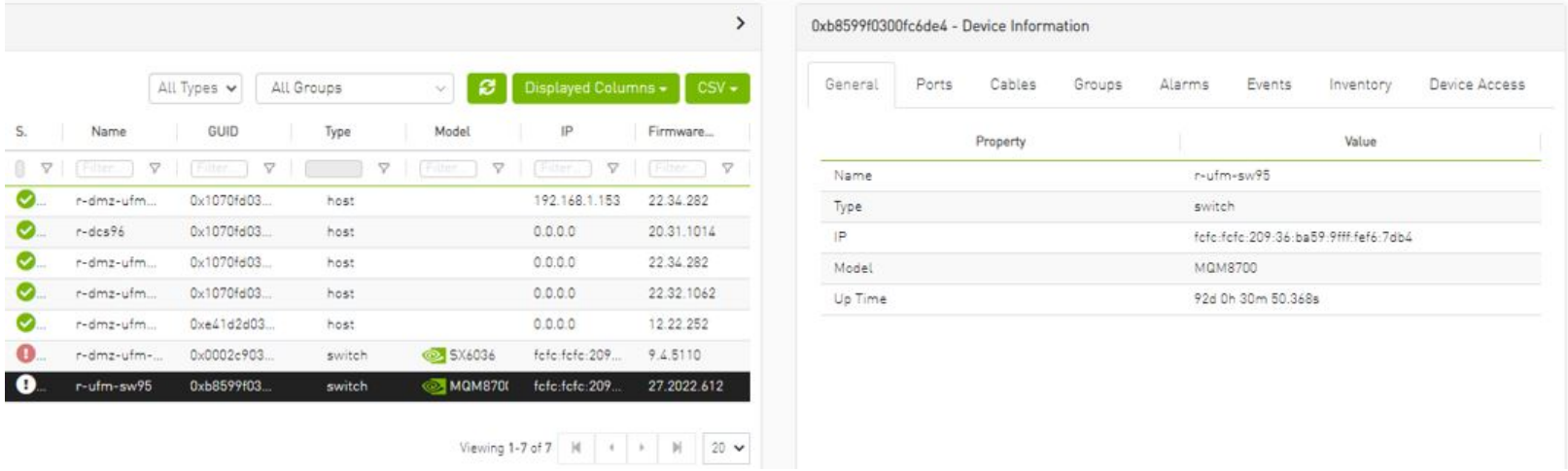

Upgrade Cable Transceivers
Firmware Image - Please select one of the available images or upload new one
Upload new image
Expert Mode
Current Firmware Version
Transceiver Type
Progress
Filter...
Filter...
38.100.122
1
Hercules2
hercules2-38_100_122.bin
Activating cable transceivers may cause a temporary communication interruption of related links - are you sure you want to continue?
No
Yes

5. After the whole action is completed successfully, you will be able to see the following message at the model bottom The upgrade cable transceivers completed successfully, do you want to activate it? by clicking the yes button it will run a new action on all the burned devices to activate the new uploaded binary image.
6. Another option to activate burned cables transceivers you can go to the Groups page and right click on the predefined Group named Devices Pending FW Transceivers Reset or you can right click on the upgraded device from managed element page and select Activate cable Transceivers action.


All Types
All Groups
Displayed Columns
CSV
S...
Name
GUID
Type
Model
IP
Firmware Ve...
smg-ib-sim001
0xb8599f0300c...
host
0.0.0.0
18.32.524
smg-ib-svr031
0x98039b0300...
host
0.0.0.0
20.31.2006
smg-ib-apl022...
0x98039b0300...
host
0.0.0.0
20.32.1010
smg-ib-svr032
0x1070fd03007...
host
0.0.0.0
28.33.810
smg-ib-sw...
0x98039b0300...
switch
MQM8700
10.209.24.136
27.2000.2046
smg-ib-olg...
CS7520
10.209.27.99
mismatched
smg-ib-sw...
MQM9700
10.209.24.121
31.2010.2036
smg-ib-sw...
MQM8700
10.209.24.10
27.2010.2010
smg-ib-sw...
MQM8700
10.209.24.57
27.2010.1202
smg-ib-sw...
MSB7700
10.209.27.36
11.2008.3328
Viewing 1-10 of 24
10
Copy Cell
Show In Network Map
Reboot
Collect System Dump
Mark As Unhealthy
Activate Cable Transceivers
Software Upgrade
Add To Group
Remove From Group
Suppress Notifications
Add To Monitor Session

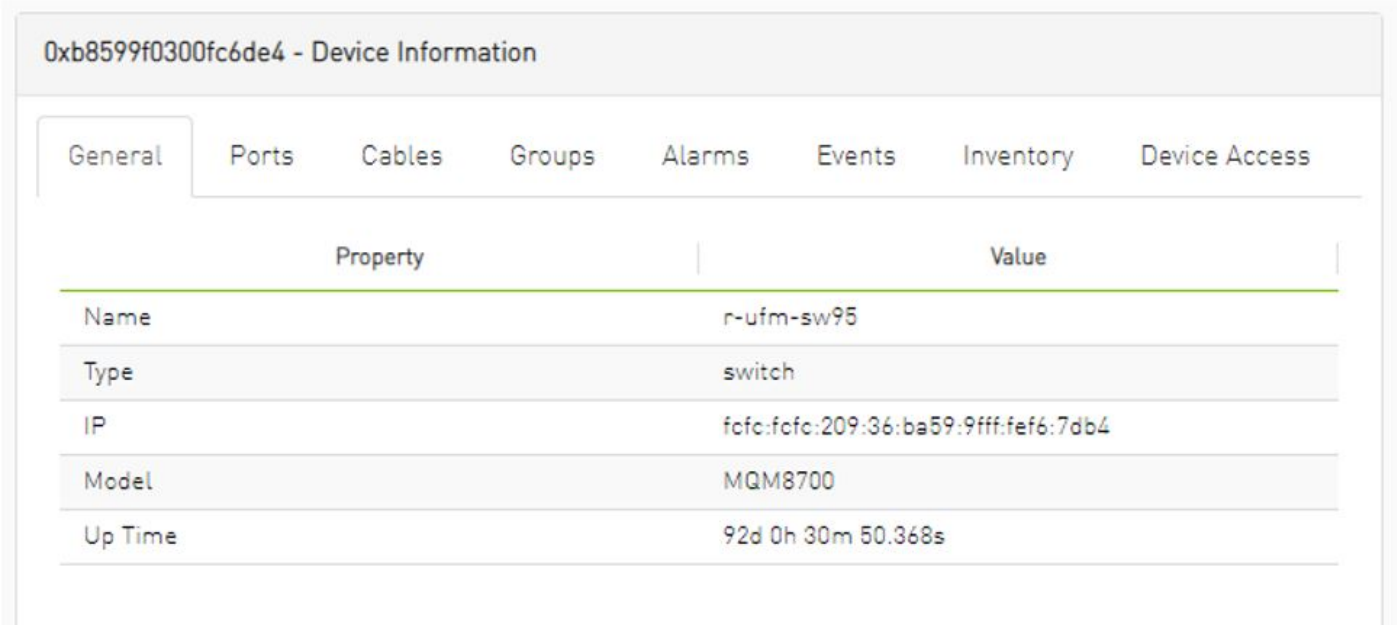
Device Information Tabs

Selecting a device from the Devices table reveals the **Device Information** table on the right side of the screen. This table provides information on the device's ports, cables, groups, events, alarms, inventory, and device access.



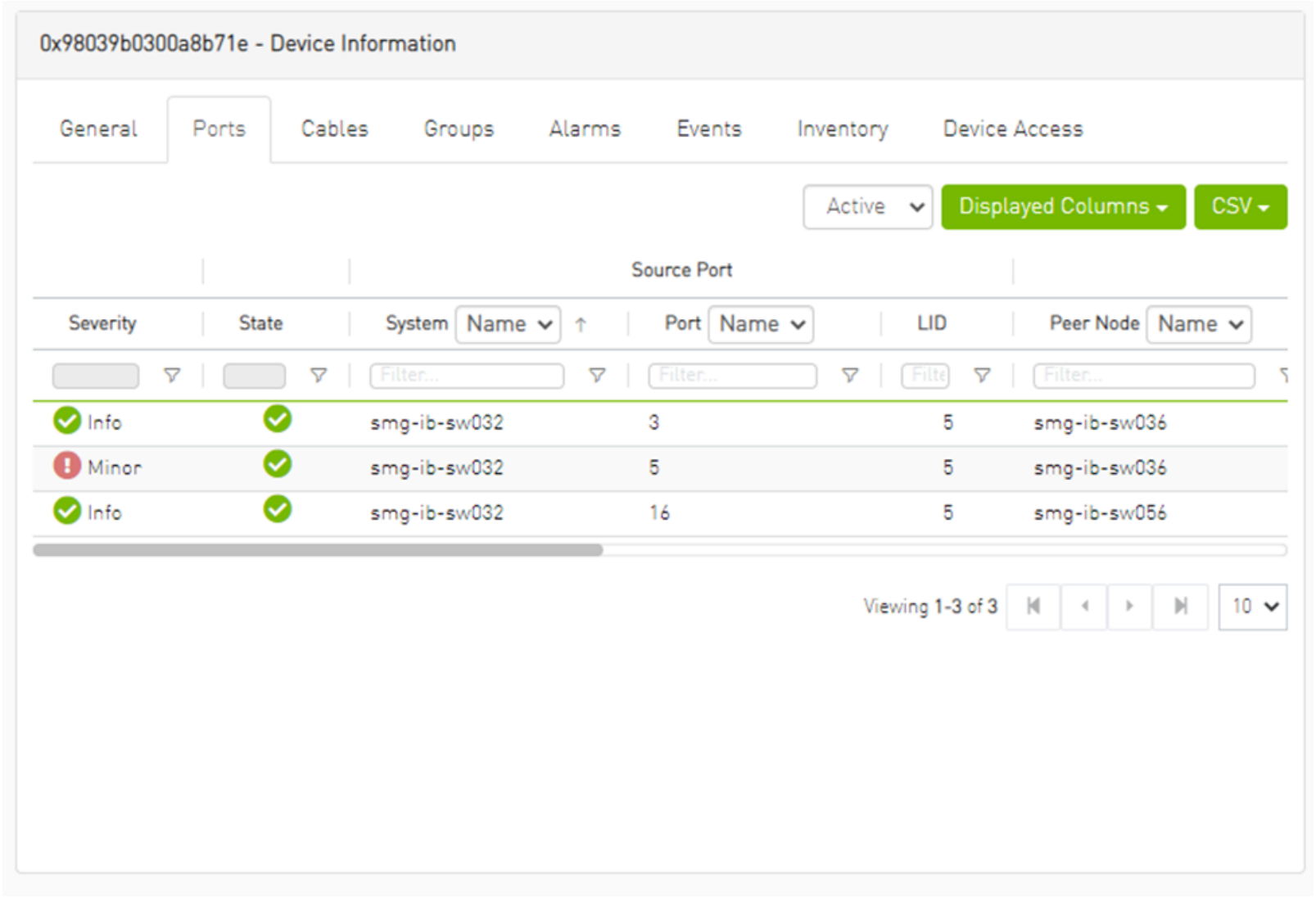
General Tab

Provides general information on the selected device.



Ports Tab

This tab provides a list of the ports connected to this device in a tabular format.

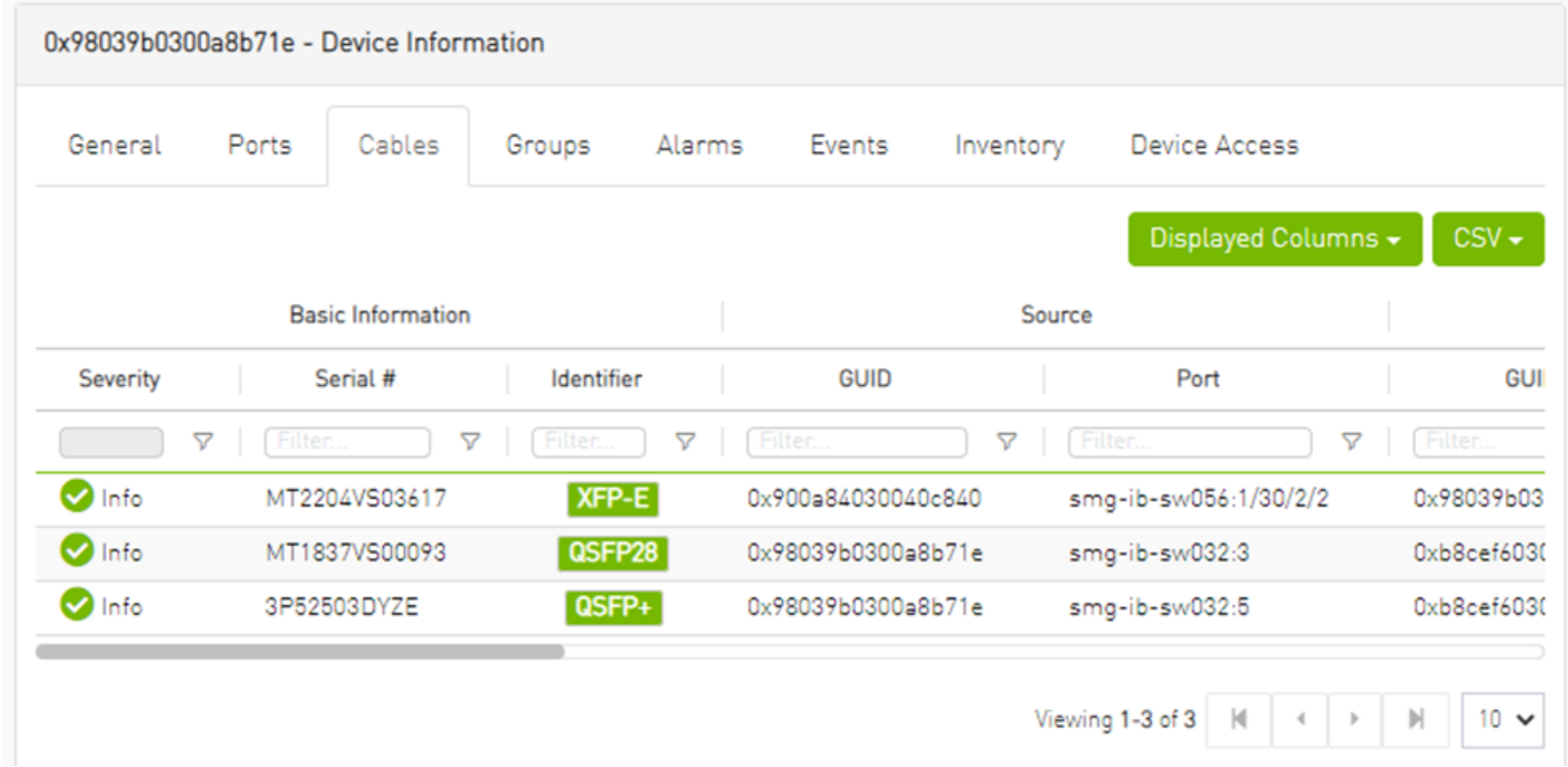

0x98039b0300a8b71e - Device Information
General
Ports
Cables
Groups
Alarms
Events
Inventory
Device Access
Active
Displayed Columns
CSV
Source Port
Severity
State
System Name
Port Name
LID
Peer Node Name
Filter...
Info
smg-ib-sw032
3
5
smg-ib-sw036
Minor
smg-ib-sw032
5
5
smg-ib-sw036
Info
smg-ib-sw032
16
5
smg-ib-sw056
Viewing 1-3 of 3
10

Ports Data

Data Type	Description
Port Number	The number of ports on device.
Node	The node name/GUID/IP that the port belongs to. Note that you can choose the node label (name/GUID/IP) using the drop-down menu available above the Ports data table.
Health	Health of the port reflecting the highest alarm severity. Please refer to the Health States table.
State	Indicates whether the port is connected (active or inactive).
LID	The local identifier (LID) of the port.
MTU	Maximum Transmission Unit of the port.
Speed 	Lists the highest value of active, enabled and supported speeds in icons indicating their status: • Dark green – active speed • Light green – enabled speed • Grey – supported yet disabled speed
Width 	Lists the highest value of active, enabled and supported widths in icons indicating their status: • Dark green – active width • Light green – enabled width • Grey – supported yet disabled width
Peer	The GUID of the device the port is connected to.
Peer Port	The name of the port that is connected to this port.

Cables Tab

This tab provides a list of the cables connected to this device in a tabular format.



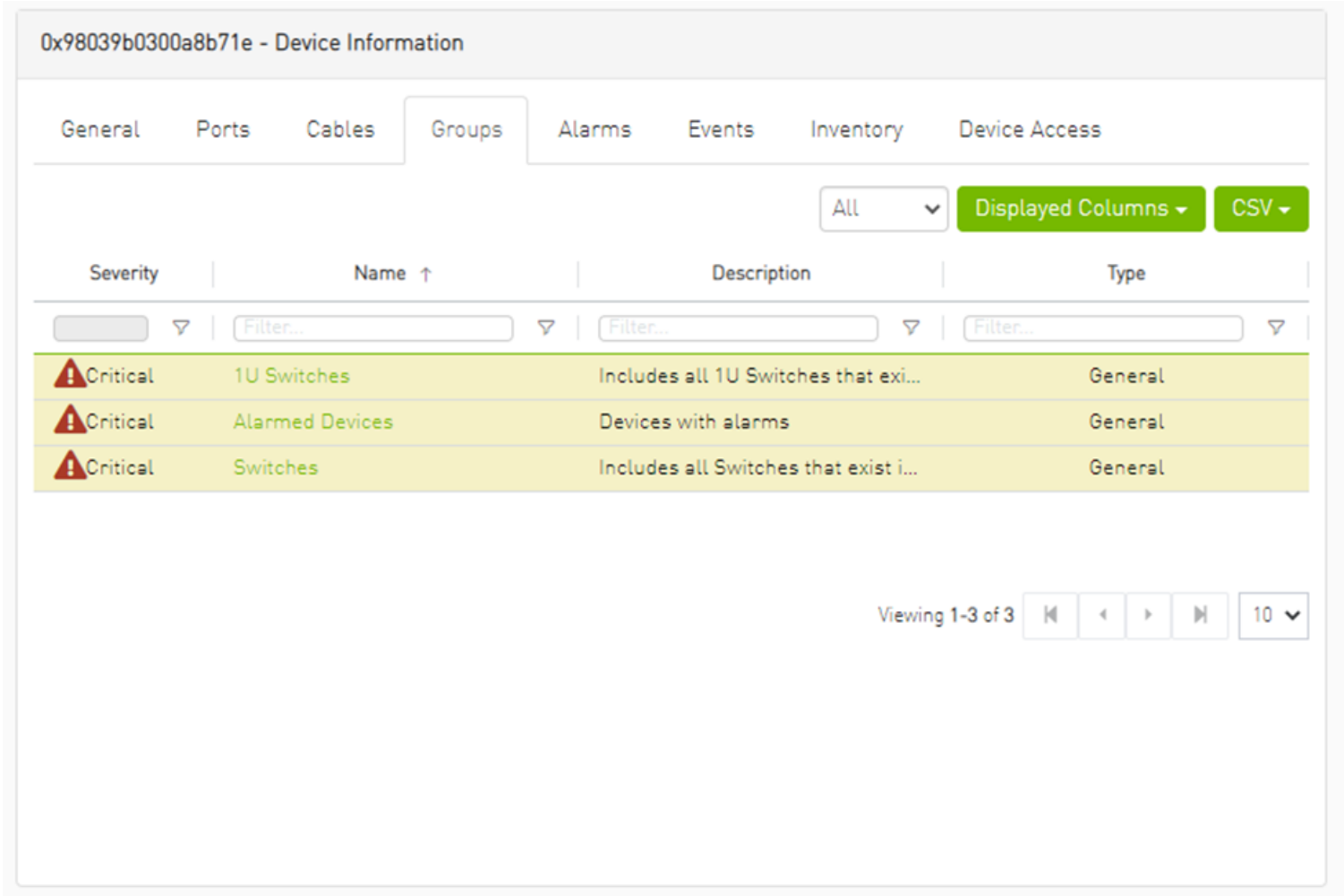
Cables Data

Data Type	Description

Basic Information	
Health	Health of the cable reflecting the highest alarm severity. Please refer to the Health States table.
Serial Number	Serial number of the cable.
Identifier	Identifier of the cable.
Source Port Information	
Source GUID	GUID of the source port the cable is connected to.
Source Port	The number of the source port the cable is connected to.
Destination Port Information	
Destination GUID	GUID of the destination port the cable is connected to.
Destination Port	The number of the destination port the cable is connected to.
Advanced Information	
Revision	Revision of the cable.
Link Width	The maximum link width of the cable.
Part Number	Part number of the cable.
Technology	The transmitting medium of the cable: copper/optical/etc.
Length	The cable length in meters.

Groups Tab

This tab provides a list of the groups to which the selected device belongs.

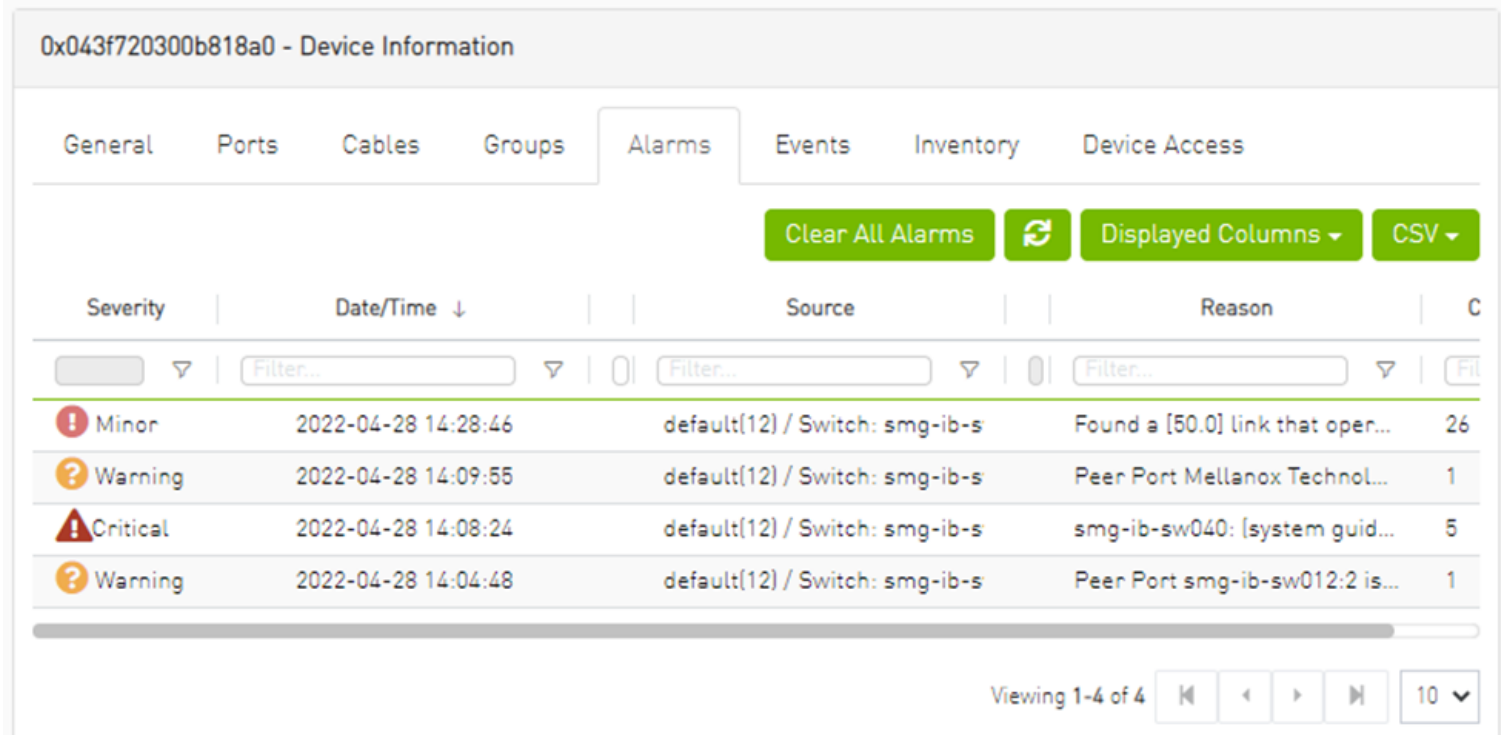

0x98039b0300a8b71e - Device Information
General
Ports
Cables
Groups
Alarms
Events
Inventory
Device Access
All
Displayed Columns
CSV
Severity
Name
Description
Type
Filter...
Filter...
Filter...
Critical
1U Switches
Includes all 1U Switches that exi...
General
Critical
Alarmed Devices
Devices with alarms
General
Critical
Switches
Includes all Switches that exist i...
General
Viewing 1-3 of 3
10

Groups Data

Data Type	Description
Severity	Aggregated severity level of the group (the highest severity level of all group members).
Name	Name of the group.
Description	Description of the group.
Type	Type of the group: General/Rack.

Alarms Tab

This tab provides a list of all UFM alarms related to the selected device.



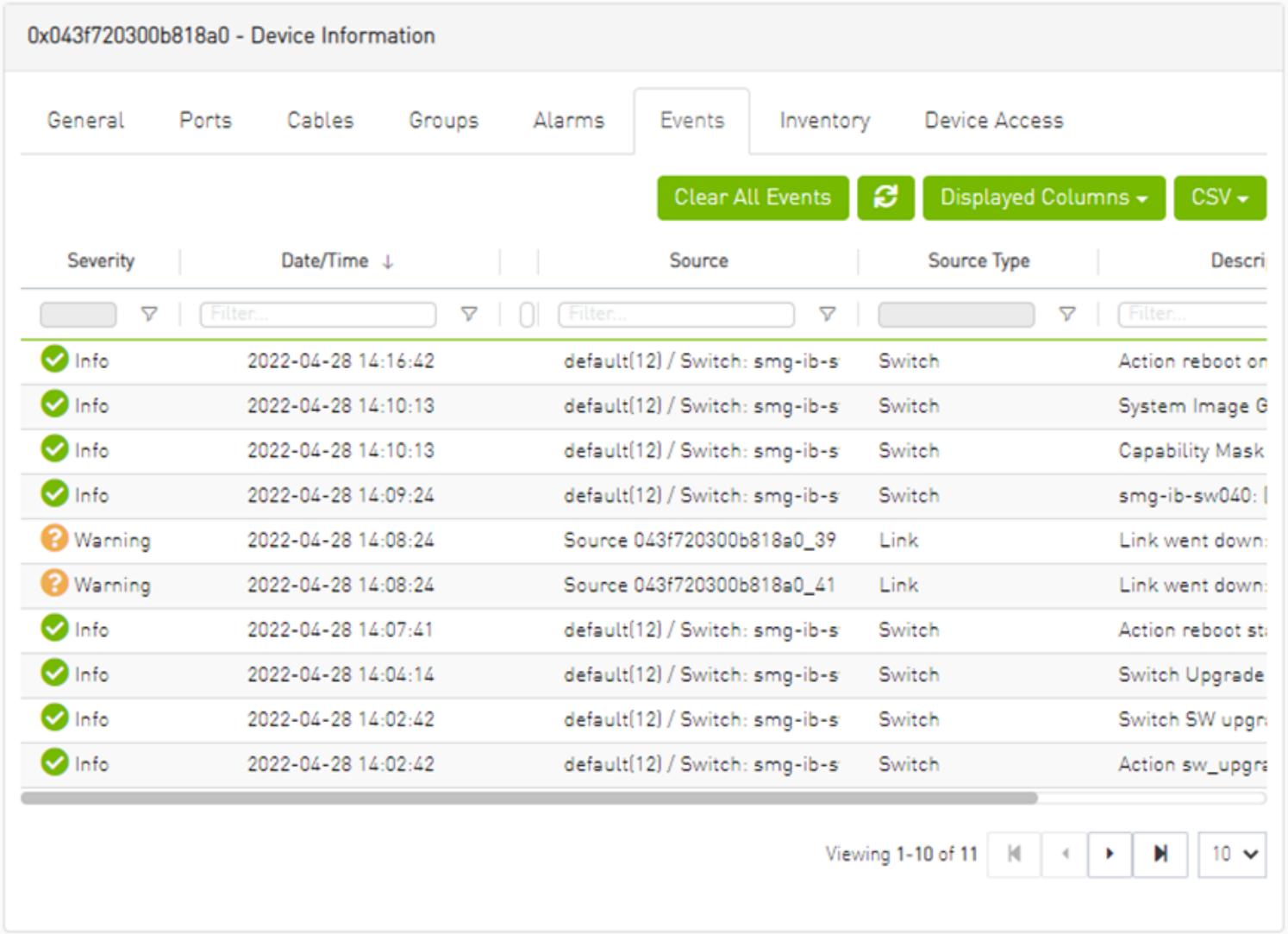
Alarms Data

Data Type	Description
Alarms ID	Alarm identifier.
Source	Source object (device/port) on which the alarm was triggered.
Severity	The severity of the alarm.

Data Type	Description
Description	Description of the alarm.
Date/Time	The time when the alarm was triggered.
Reason	Reason for the alarm.
Count	Number of instances that the alarm occurred on the related source object.

Events Tab

This tab provides a list of the UFM events that are related to the selected device.


0x043f720300b818a0 - Device Information
General
Ports
Cables
Groups
Alarms
Events
Inventory
Device Access
Clear All Events
Displayed Columns
CSV
Severity
Date/Time
Source
Source Type
Descri
Filter...
Info 2022-04-28 14:16:42 default(12) / Switch: smg-ib-s Switch Action reboot on
Info 2022-04-28 14:10:13 default(12) / Switch: smg-ib-s Switch System Image G
Info 2022-04-28 14:10:13 default(12) / Switch: smg-ib-s Switch Capability Mask
Info 2022-04-28 14:09:24 default(12) / Switch: smg-ib-s Switch smg-ib-sw040:
Warning 2022-04-28 14:08:24 Source 043f720300b818a0_39 Link Link went down:
Warning 2022-04-28 14:08:24 Source 043f720300b818a0_41 Link Link went down:
Info 2022-04-28 14:07:41 default(12) / Switch: smg-ib-s Switch Action reboot st
Info 2022-04-28 14:04:14 default(12) / Switch: smg-ib-s Switch Switch Upgrade
Info 2022-04-28 14:02:42 default(12) / Switch: smg-ib-s Switch Switch SW upgr
Info 2022-04-28 14:02:42 default(12) / Switch: smg-ib-s Switch Action sw_upgr
Viewing 1-10 of 11
10

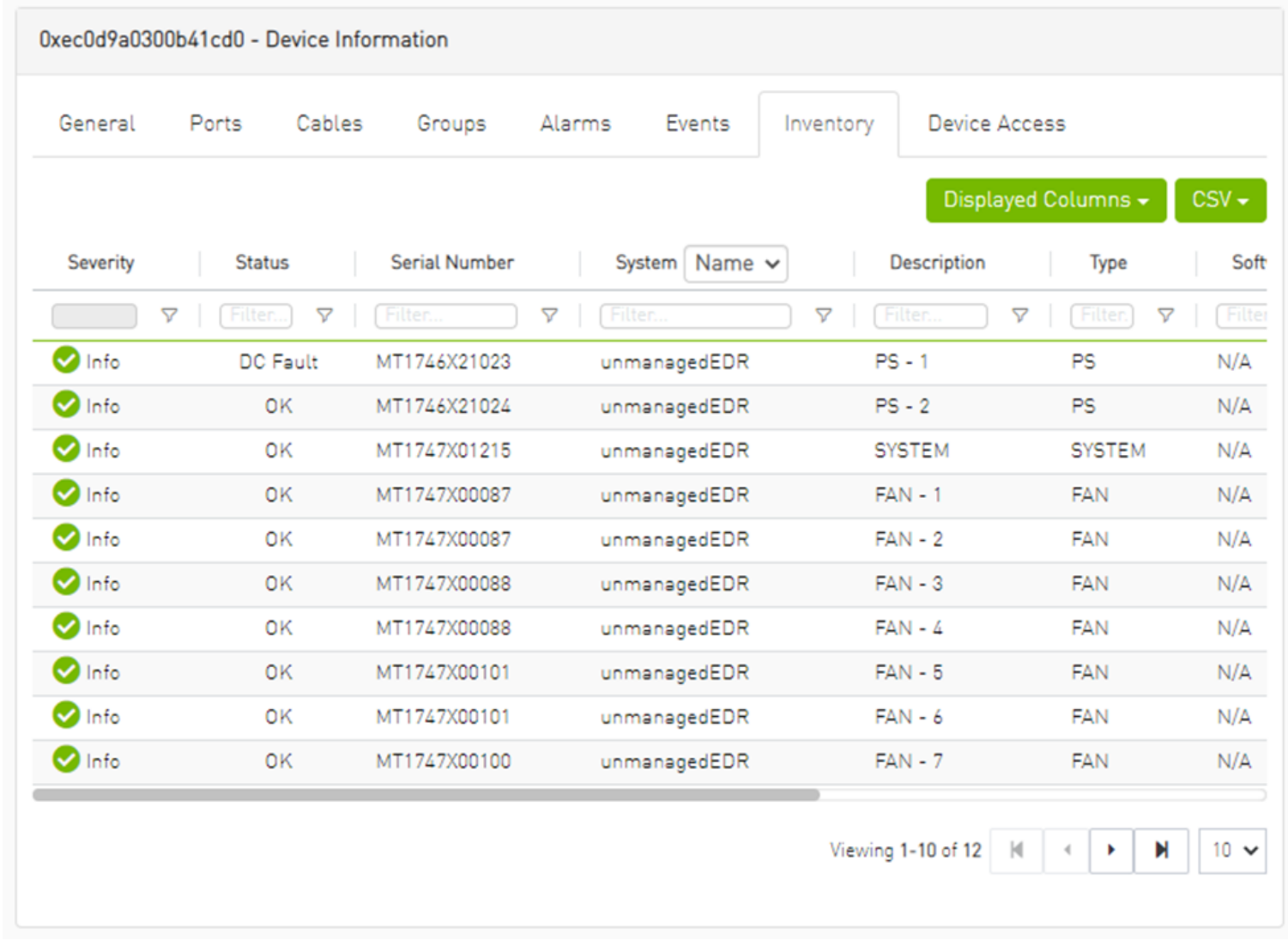
Events Data

Data Type	Description
Severity	Event severity - Info, Warning, Error, Critical or Minor.
Event Name	The name of the event.
Source	The source object (device/port) on which the event was triggered.
Date/Time	The time when the event was triggered.
Category	The category of the event indicated by icons. Hovering over the icon will display the category name.
Description	Description of the event. Full description can be displayed by hovering over the text.

Inventory Tab

This tab provides a list of the device's modules with information in a tabular format.

> ⚠ This tab is available for switches only.


0xec0d9a0300b41cd0 - Device Information
General
Ports
Cables
Groups
Alarms
Events
Inventory
Device Access
Displayed Columns
CSV
Severity
Status
Serial Number
System Name
Description
Type
Soft
Info DC Fault MT1746X21023 unmanagedEDR PS - 1 PS N/A
Info OK MT1746X21024 unmanagedEDR PS - 2 PS N/A
Info OK MT1747X01215 unmanagedEDR SYSTEM SYSTEM N/A
Info OK MT1747X00087 unmanagedEDR FAN - 1 FAN N/A
Info OK MT1747X00087 unmanagedEDR FAN - 2 FAN N/A
Info OK MT1747X00088 unmanagedEDR FAN - 3 FAN N/A
Info OK MT1747X00088 unmanagedEDR FAN - 4 FAN N/A
Info OK MT1747X00101 unmanagedEDR FAN - 5 FAN N/A
Info OK MT1747X00101 unmanagedEDR FAN - 6 FAN N/A
Info OK MT1747X00100 unmanagedEDR FAN - 7 FAN N/A
Viewing 1-10 of 12
10

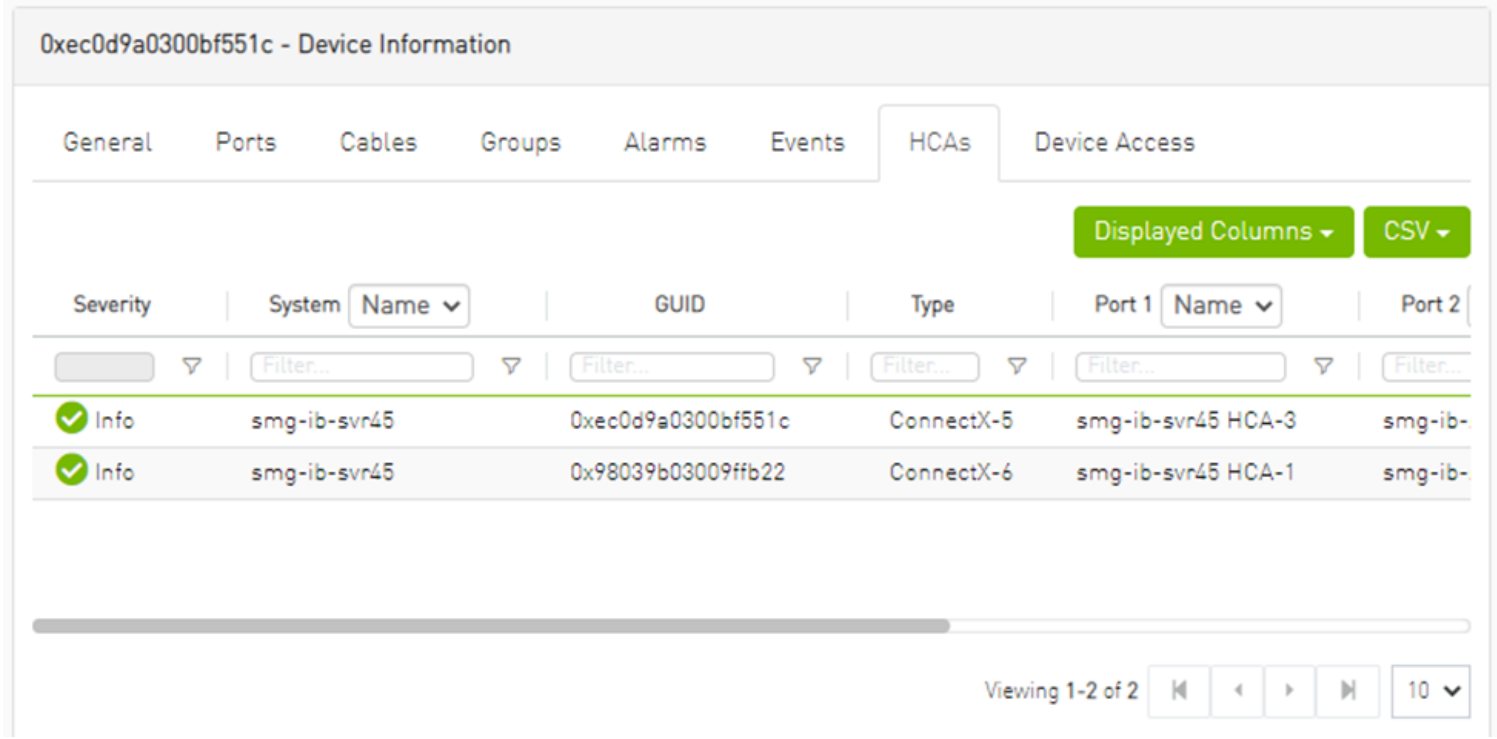
Inventory Data

Data Type	Description
Health	Health of the module reflecting the highest alarm severity. Please refer to the Health States table.
Status	The module status.
Serial Number	Serial number of the module.
Name	Name of the device.
Description	Description of the module.
Type	Type of the module: spine/line/etc.
Firmware Version	Firmware version installed on the module.
Hardware Version	Hardware version of the module.
Temperature	Temperature of the module.

HCAs Tab

This tab provides a list of the device's HCAs with information in a tabular format.

> ⚠ This tab is available for hosts only.

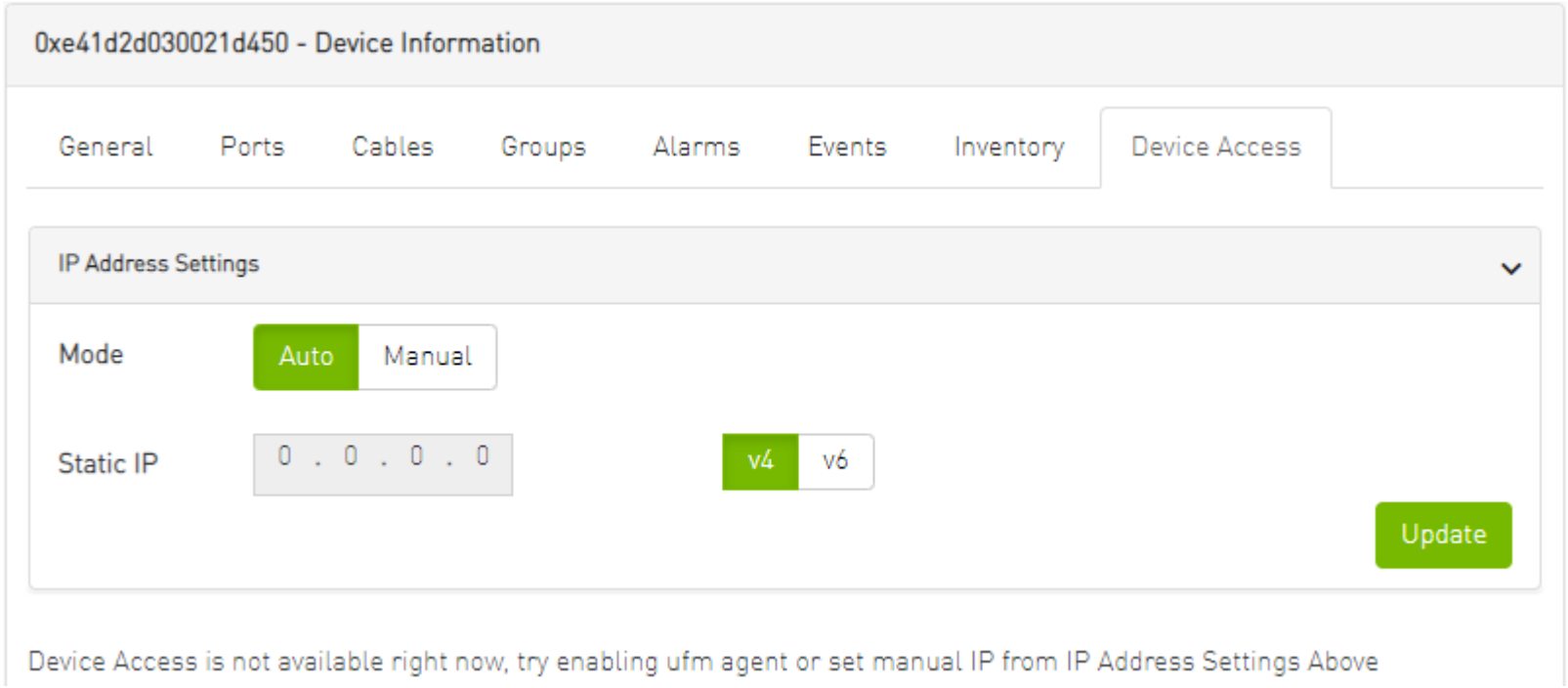


Data Type	Description
Health	Health of the HCA reflecting the highest alarm severity. Please refer to the Health States table.
Name	HCA Index
GUID	HCA GUID

Data Type	Description
Type	HCA Type
Port GUID	HCA ports GUIDs
PSID	HCA PSID
FW Version	HCA firmware version

Device Access Tab

This tab allows for managing the access credentials of the selected device for remote accessibility. To be able to set access credentials for the device, a device IP must be set either by installing UFM Agent on the device, or by manually setting the IP under **IP Address Settings** (IP is now supported with v4 and v6).

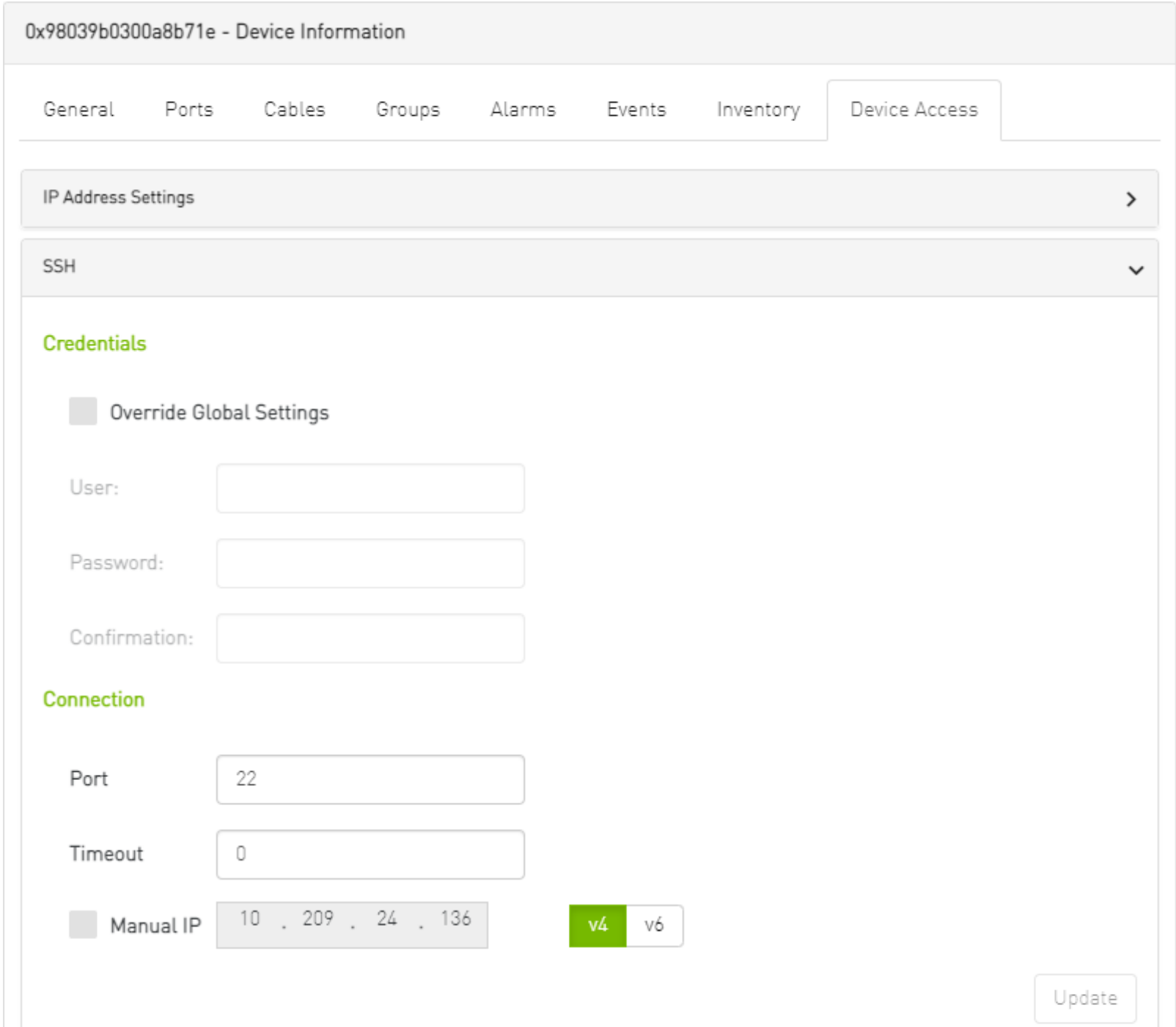


> ⚠ After manually setting the IP address of NVIDIA® Mellanox® InfiniScale IV® and SwitchX® based switches, UFM will first validate the new IP before setting it.

To edit your device access credentials

1. Select the preferred protocol tab:
   - **SSH** - allows you to define the SSH parameters to open an SSH session on your device (available for **nodes** and **switches**)
   - **IPMI** - allows you to set the IPMI parameters to open an IPMI session on your device for remote power control (available for **nodes only**)
   - **HTTP** - allows you to define the HTTP parameters to open an HTTP session on your device (available for **switches only**)

2. Click Update to save your changes.

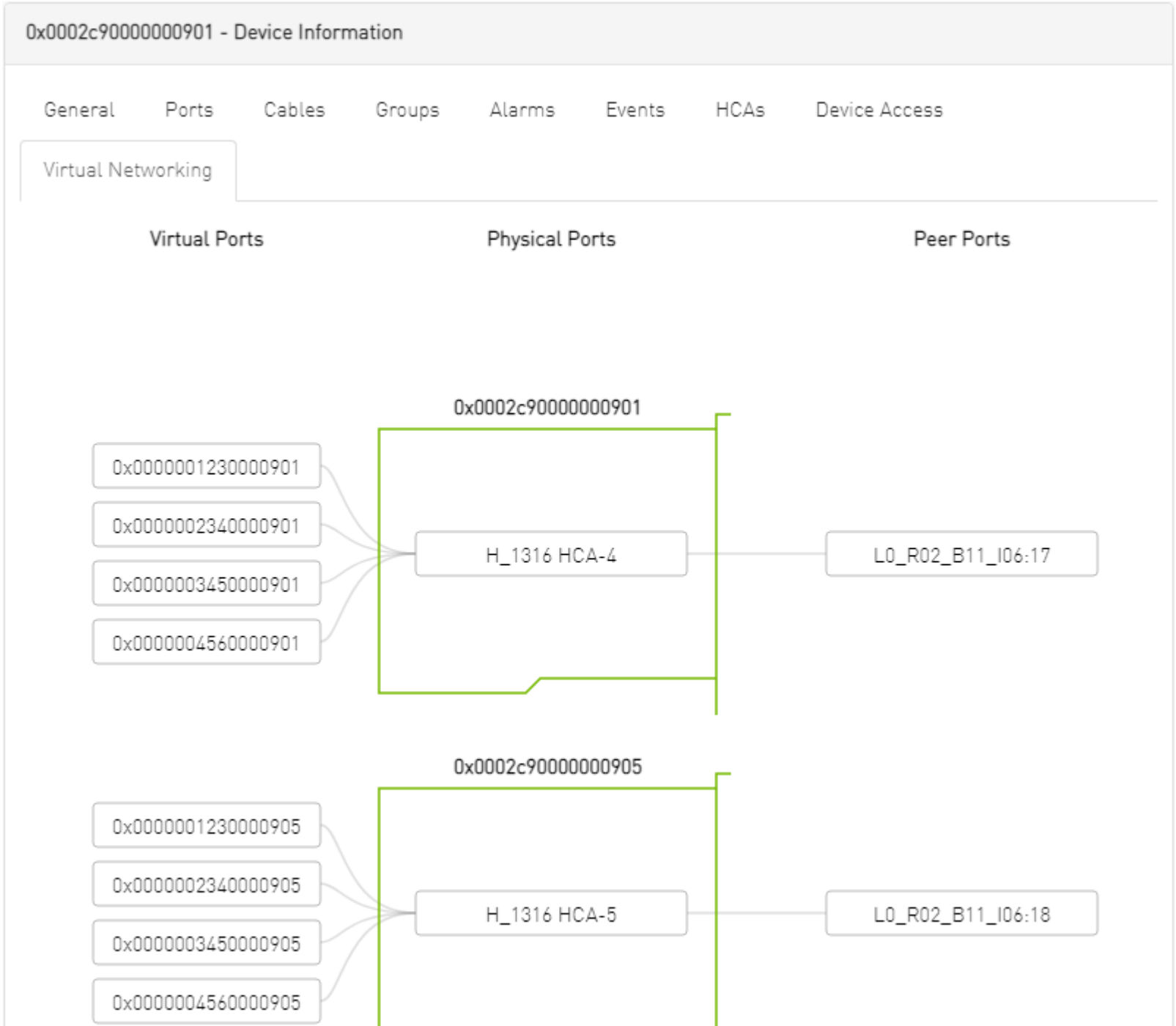

0x98039b0300a8b71e - Device Information
General
Ports
Cables
Groups
Alarms
Events
Inventory
Device Access
IP Address Settings
SSH
Credentials
Override Global Settings
User:
Password:
Confirmation:
Connection
Port
22
Timeout
0
Manual IP
10 . 209 . 24 . 136
v4
v6
Update

Device Access Credentials Parameters

Field	Description
User	Fill in or edit the computer user name.
Password	Enter the device password.
Confirmation	Enter the device password a second time to confirm.
Manual IP	Enter the device IP address (could be IPv4/IPv6).
Port	Enter the port number.
Timeout	Enter the connection timeout (in seconds) for the device specific protocol (SSH/HTTP/IPMI).

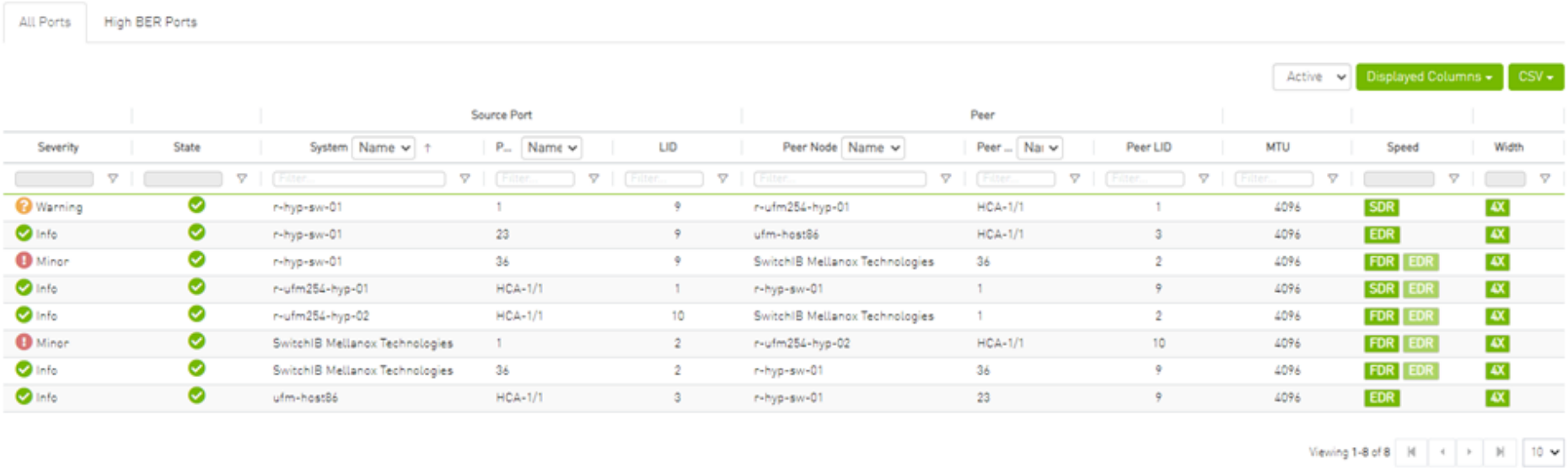
Virtual Networking Tab

This tab displays a map containing the HCAs for the selected device, and the ports and virtual ports it is connected to.


0x0002c90000000901 - Device Information
General
Ports
Cables
Groups
Alarms
Events
HCAs
Device Access
Virtual Networking
Virtual Ports
Physical Ports
Peer Ports
0x0002c90000000901
0x0000001230000901
0x0000002340000901
0x0000003450000901
0x0000004560000901
H_1316 HCA-4
L0_R02_B11_I06:17
0x0002c90000000905
0x0000001230000905
0x0000002340000905
0x0000003450000905
0x0000004560000905
H_1316 HCA-5
L0_R02_B11_I06:18

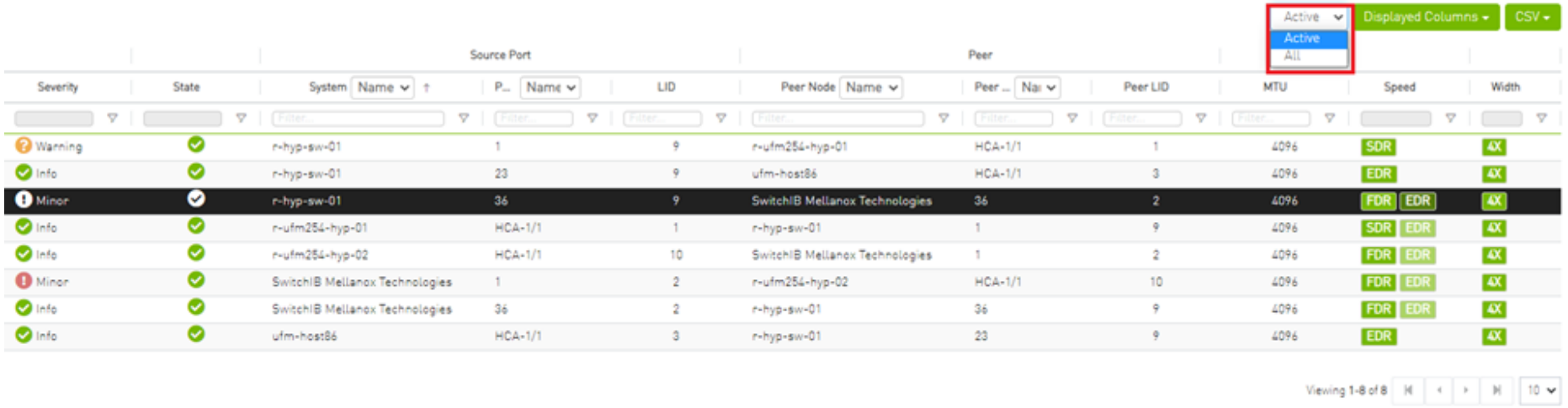
Ports Window

Provides a list of all ports in UFM.

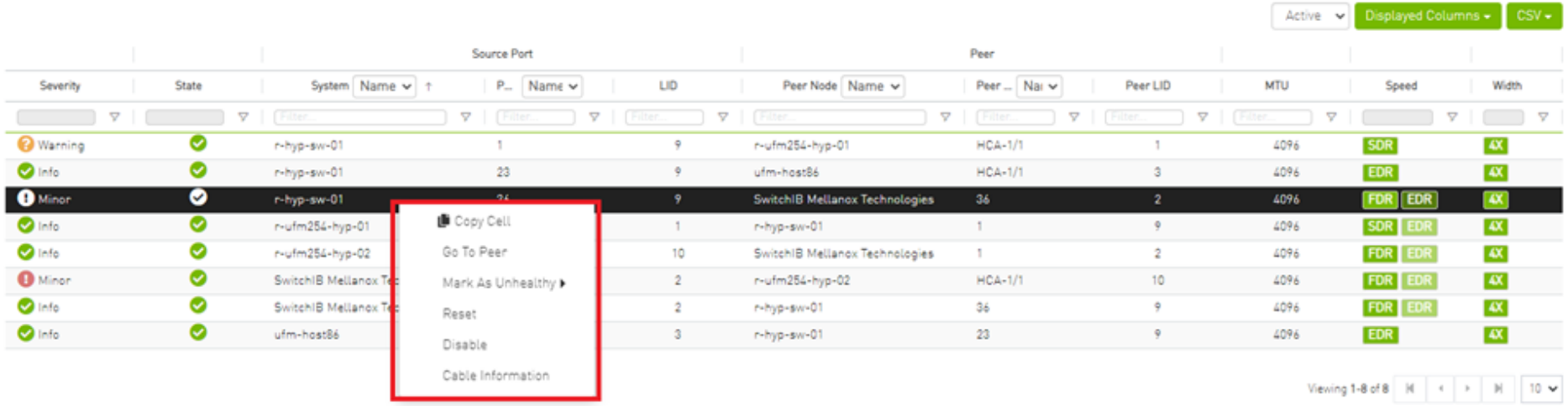


The table can be filtered by port state. The filter contains two options:

- Active – only active ports
- All – all ports

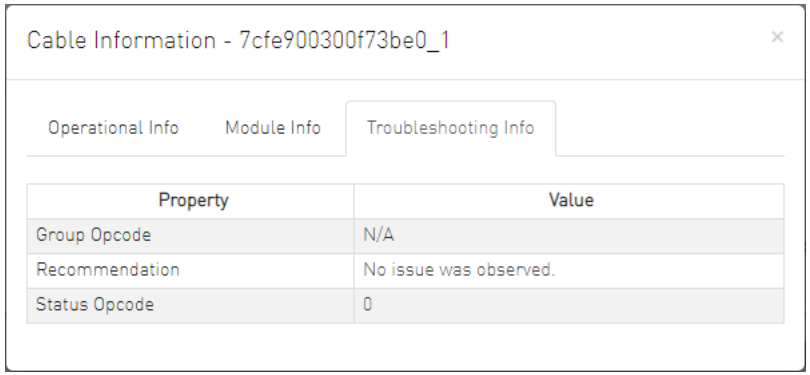


When right-clicking one of the available ports, the following actions appear:



All enable/disable actions on managed switches' ports are persistent. Thus, if a managed switch port is disabled, the port remains disabled even when rebooting the switch.

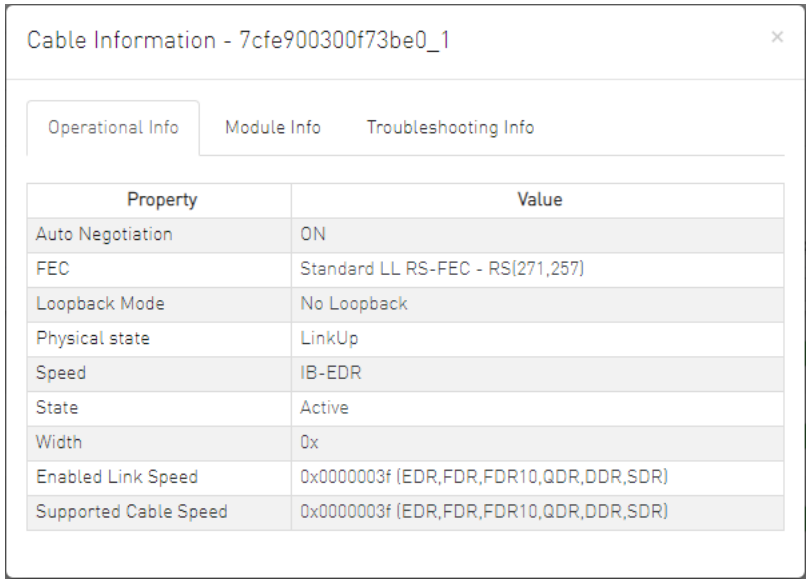
Clicking "Cable Information" opens up a window which provides data on operational, module, and troubleshooting information as shown in the following:



Cable Information - 7cfe900300f73be0_1

Operational Info | Module Info | Troubleshooting Info

Property	Value
Vendor Serial Number	MT1515VS07837
Vendor Part Number	MCP1600-E001
Vendor Name	Mellanox
Attenuation (5g,7g,12g) [dB]	4,5,9
Bias Current [mA]	N/A
Cable Technology	Copper cable unequalized
Cable Type	Passive copper cable
CDR RX	N/A
CDR TX	N/A
Compliance	N/A
Digital Diagnostic Monitoring	No
FW Version	N/A
Identifier	QSFP+
LOS Alarm	N/A
OUI	Mellanox
Power Class	1.5 W max
Rev	A2
Rx Power Current [dBm]	N/A
Temperature [C]	N/A
Transfer Distance [m]	1
Tx Power Current [dBm]	N/A
Voltage [mV]	N/A
Wavelength [nm]	N/A

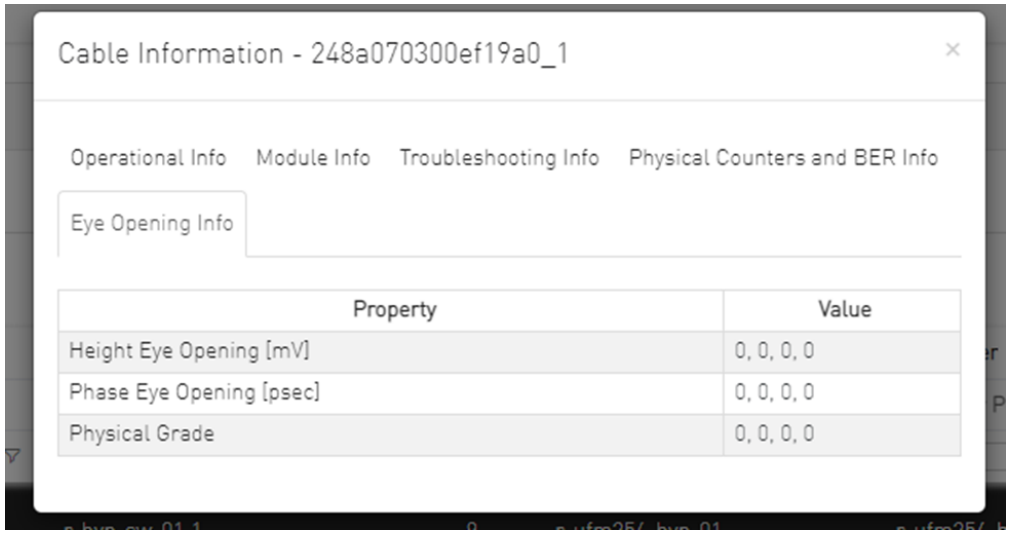


Physical Grade and Eye Opening Information

Eye opening information contains the following data:

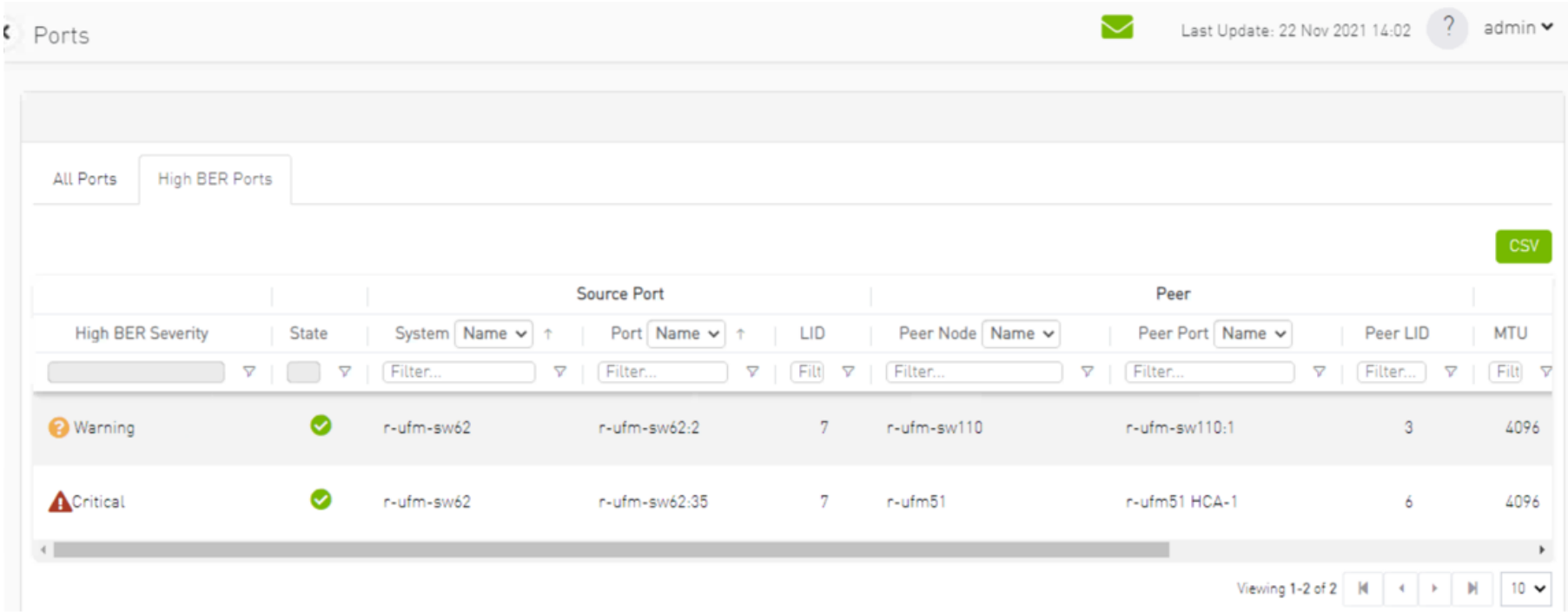
- Physical Grade: [Grade0, Grade1, Grade2, Grade3]
- Height Eye Opening [mV]: [Height0, Height1, Height2, Height3]
- Phase Eye Opening [psec]: [Phase0, Phase1, Phase2, Phase3]

A new tab called Eye Information was added under cable information modal in ports table.



Auto-isolation of High-BER Ports

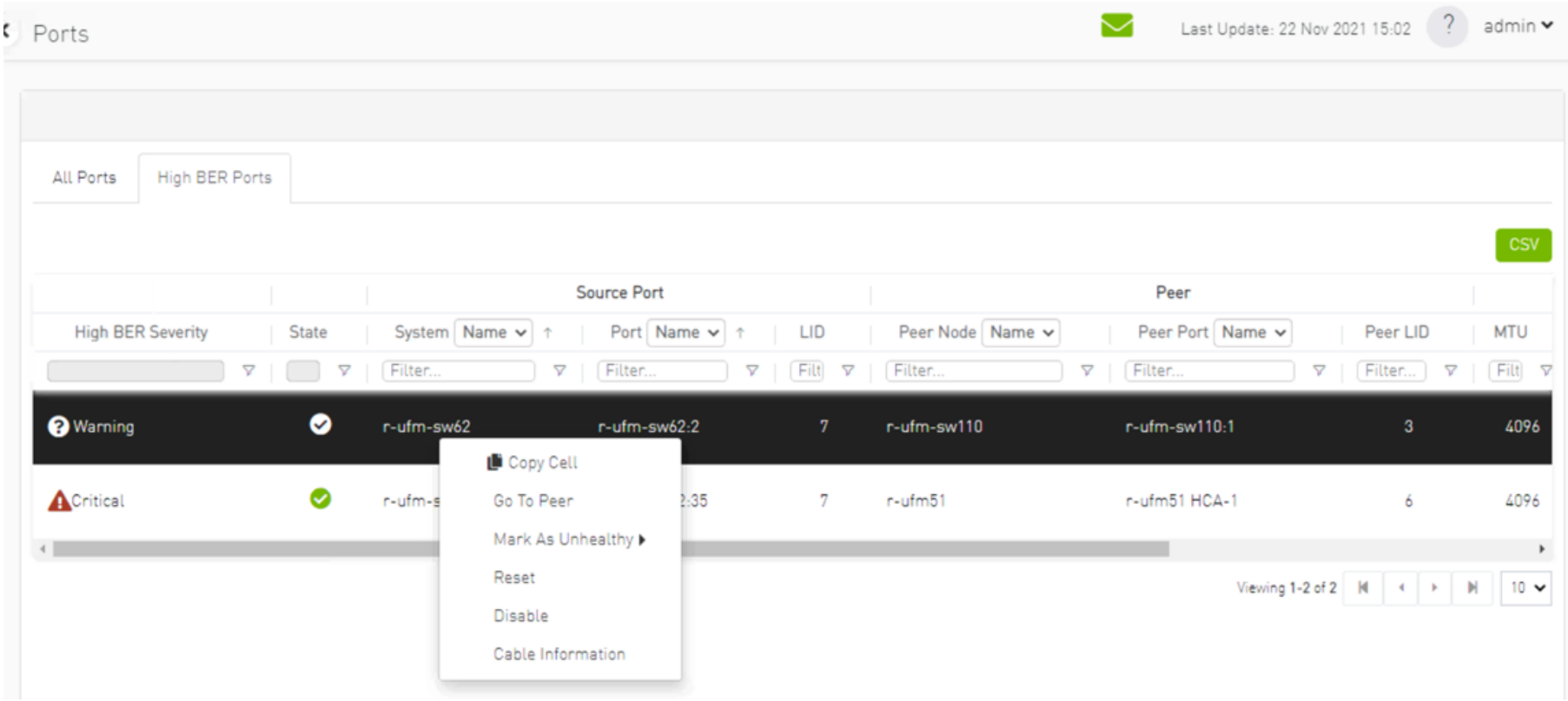
The High BER Ports tab lists all high-BER ports in the fabric.



The flags `high_ber_ports_auto_isolation` must be configured in the `gv.cfg` file to enable this feature.

For each port discovered as a high-BER port, a new event is triggered in the Events table.

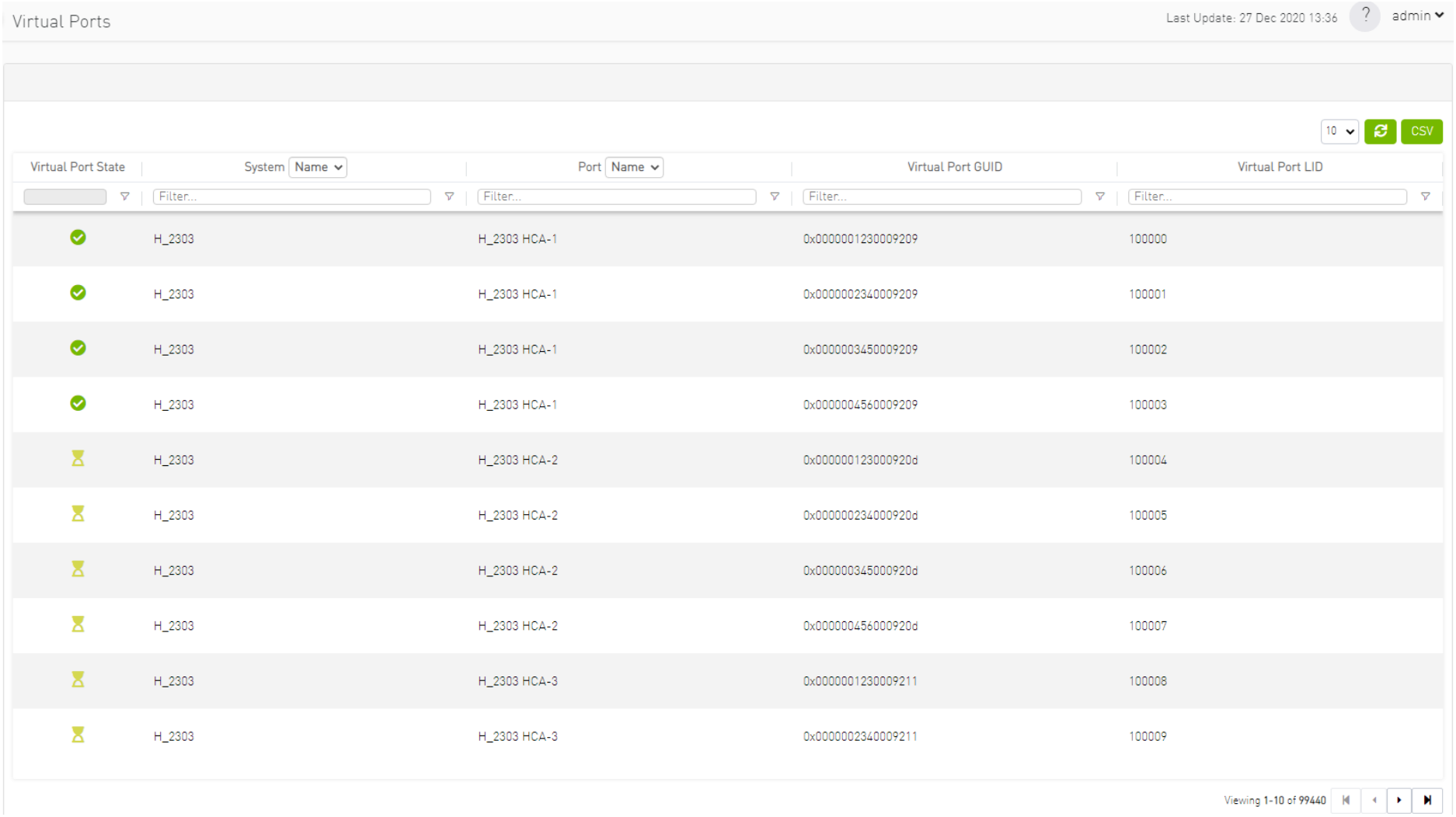
Marking the high-BER port as unhealthy suppresses all events and notifications related to the auto-isolated port.



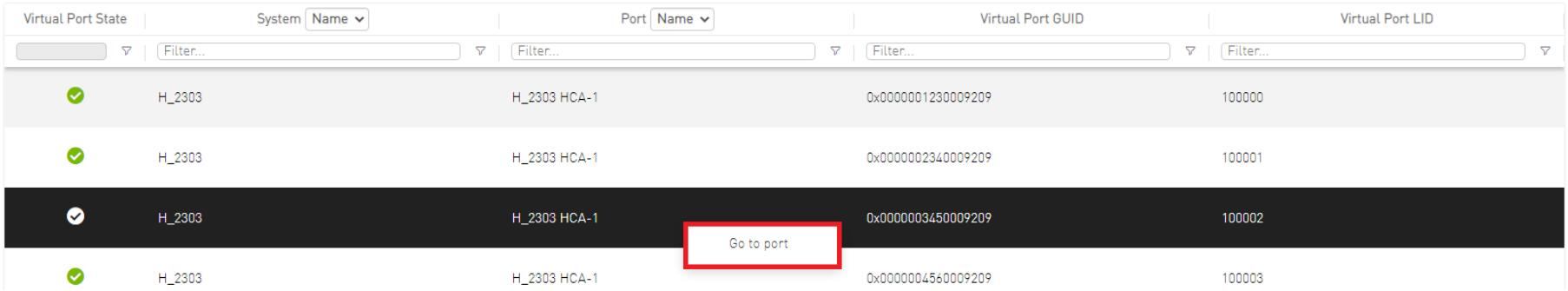
Virtual Ports Window

> ⚠ This page is only available if Virtualization is enabled in gv.cfg.

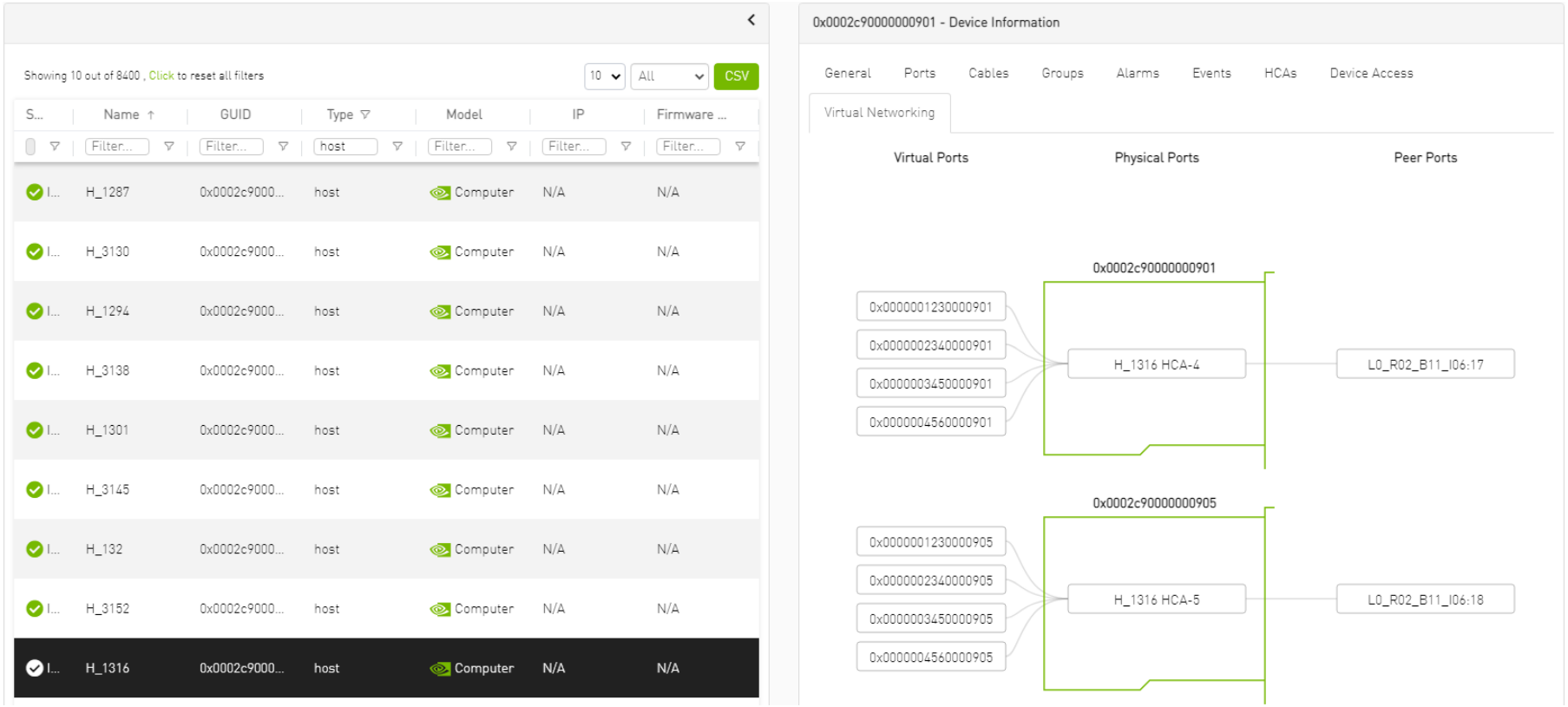
Provides a list of all virtual ports in UFM.



Right-clicking a virtual port allows navigation to the physical port mapped it is mapped to.



Clicking "Go to port" navigates to the Virtual Networking tab of the Device Information screen.



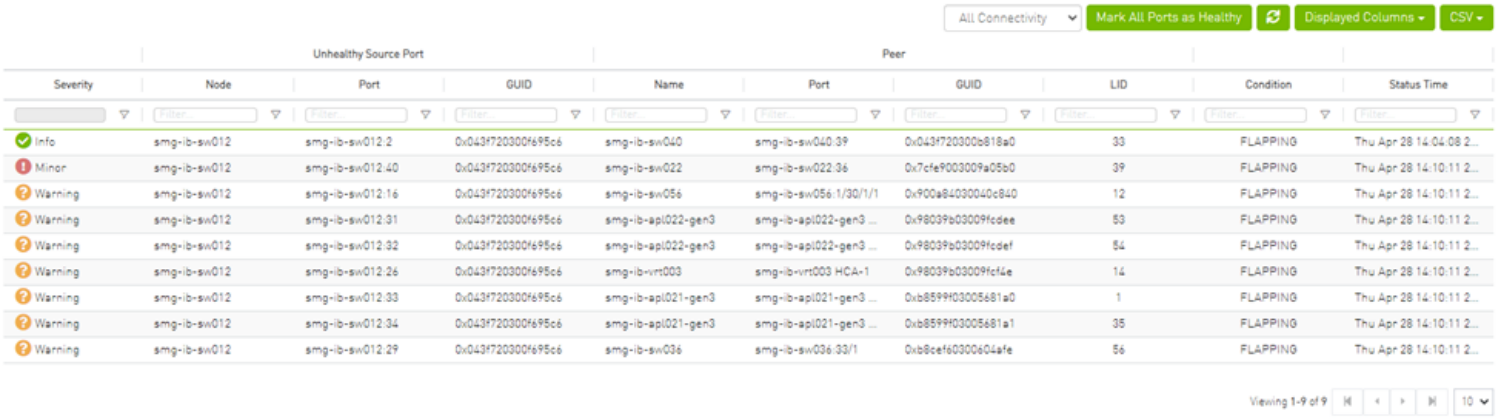
Unhealthy Ports Window

The Unhealthy Ports view shows all the unhealthy nodes in the fabric and the OpenSM health policy of the healthy/unhealthy nodes.

After the Subnet Manager examines the behavior of subnet nodes (switches and hosts) and discovers that a node is “unhealthy” according to the conditions specified below, the node is displayed in the Unhealthy Ports window. Once a node is declared as “unhealthy”, Subnet Manager can either ignore, report,

isolate or disable the node. The user is provided with the ability to control the actions performed and the phenomena that declares a node “unhealthy." Moreover, the user can “clear” nodes that were previously marked as “unhealthy."

The information is displayed in a tabular form and includes the unhealthy port’s state, source node, source port, source port GUID, peer node, peer port, peer GUID, peer LID, condition, and status time.



> **Warning** The feature requires OpenSM parameter hm_unhealthy_ports_checks to be set to TRUE (default).

> **Warning** This feature is not available in the "Monitoring Only Mode."

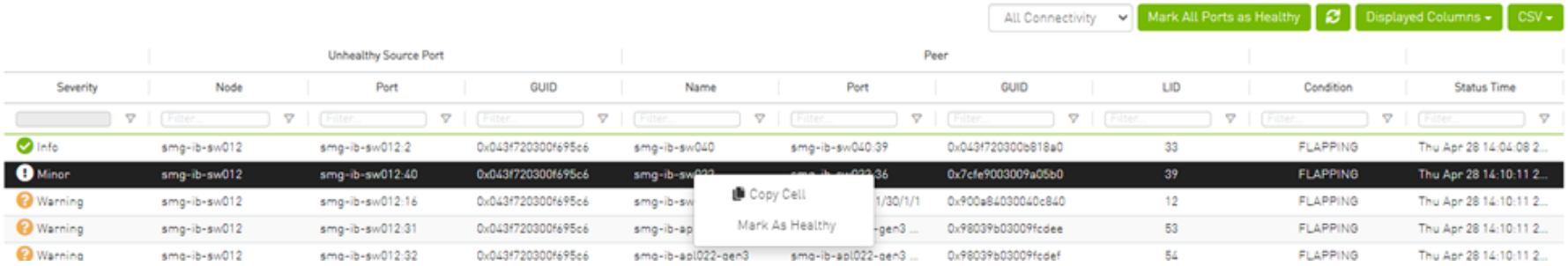
The following are the conditions that would declare a node as “unhealthy”:

- Reboot - If a node was rebooted more than 10 times during last 900 seconds
- Flapping - If several links of the node found in Initializing state in 5 out of 10 previous sweeps
- Unresponsive - A port that does not respond to one of the SMPs and the MAD status is TIMEOUT in 5 out of 7 previous SM sweeps
- Noisy Node - If a node sends traps 129, 130 or 131 more than 250 traps with interval of less than 60 seconds between each two traps
- Seterr - If a node respond with bad status upon SET SMPs (PortInfo, SwitchInfo, VLArb, SL2VL or Pkeys)
- Illegal - If illegal MAD fields are discovered after a check for MADs/fields during receive_process
- Manual - Upon user request mark the node as unhealthy/healthy
- Link Level Retransmission (LLR) – Activated when retransmission-per-second counter exceeds its threshold

All conditions except LLR generate Unhealthy port event, LLR generates a High Data retransmission event.

To clear a node from the Unhealthy Ports Tab, do the following:

1. Go to the Unhealthy Ports window under Managed Elements.
2. From the Unhealthy Ports table, right click the desired port it and mark it as healthy.



To mark a node as permanently healthy, do the following:

1. Create a file named opensm-health-policy.conf.user_ext on a remote host.
2. Enter the node and the port information and set it as "Healthy".

```
0x0002c903005dd832 6 Healthy
```

3. Import the file.

```
ufm-appliance [ mgmt-sa ] (config) # ib sm configuration import opensm-health-policy-user-ext scp://
root:123456@192.168.1.3/tmp/health-policy.conf.user_ext
```

4. Make the changes effective by running "ib sm opensm-health-policy-merge".

```
ufm-appliance [ mgmt-sa ] (config) # ib sm opensm-health-policy-merge
```

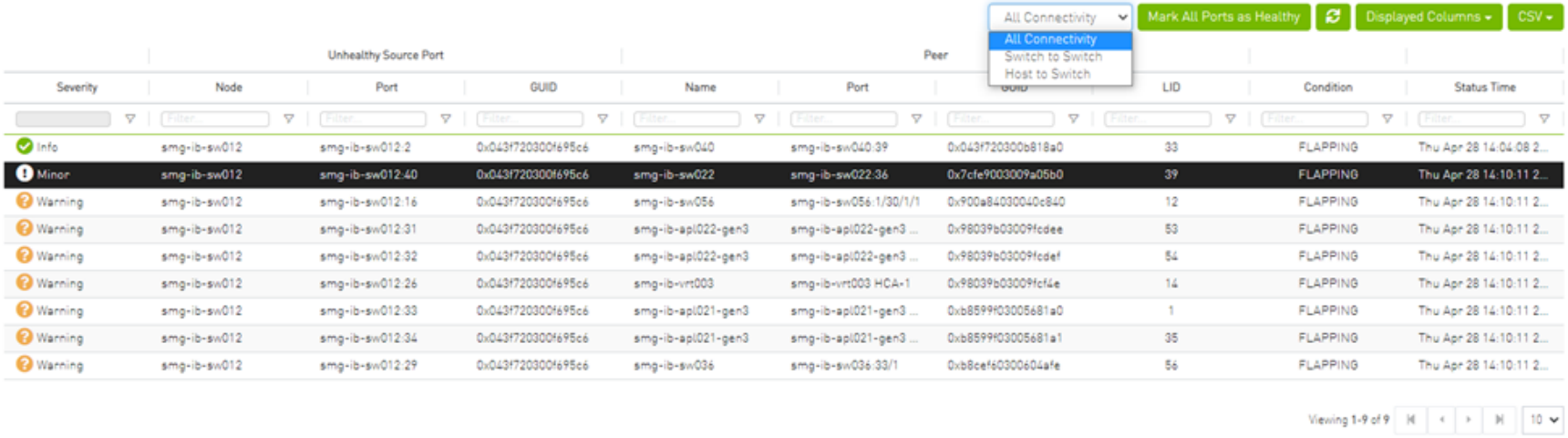
⚠ To control Partial Switch ASIC Failure event:
Trigger Partial Switch ASIC Failure whenever number of unhealthy ports exceed the defined percent of the total number of the switch ports.
The switch_asic_fault_threshold flag (under the UnhealthyPorts section in gv.cfg file) default value is 20.

Unhealthy Port Connectivity Filter

It is possible to to filter the Unhealthy Ports table by connectivity (all, host-to-switch, or switch-to-host).

Filtering the Unhealthy Ports table is possible from the dropdown options at the top of the table which includes

- All Connectivity
- Switch to Switch
- Host to Switch

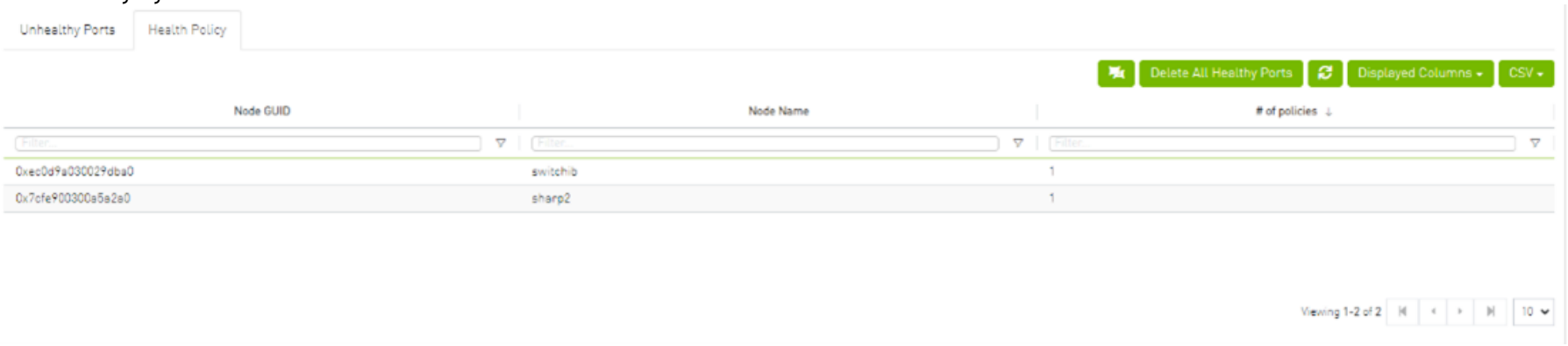


Health Policy Management

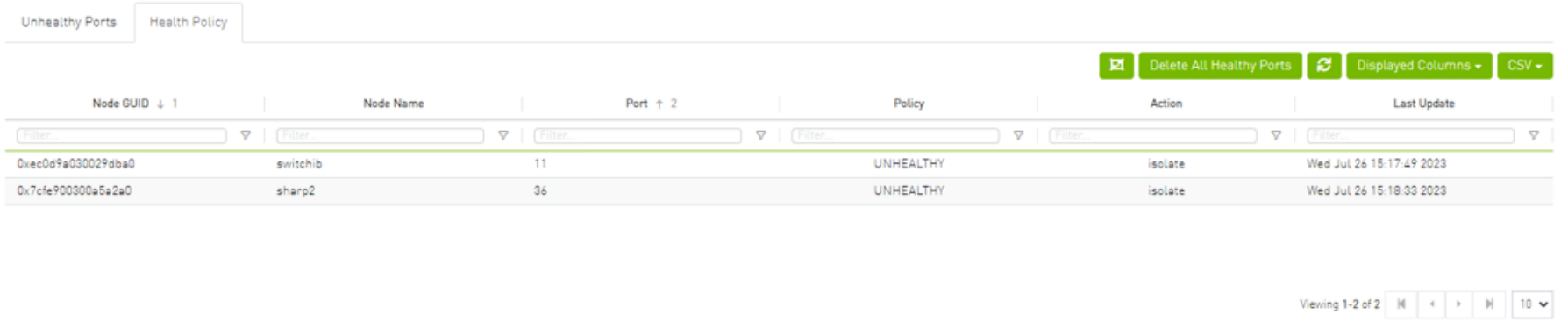
This view manages the OpenSM health policy for the healthy/unhealthy nodes and ports. The OpenSM health policy is stored in the /opt/ufm/files/conf/opensm/opensm-health-policy.conf file.

The information is displayed in a tabular form, with an option to group it either by devices or ports, and includes the health nodes/ports details (GUID, Name, policy [healthy/unhealthy])

1. Health Policy by devices:



2. Health Policy by ports:



To switch between the above views, simply click on the control button located at the top right corner of the table. By default, the devices view will be shown.

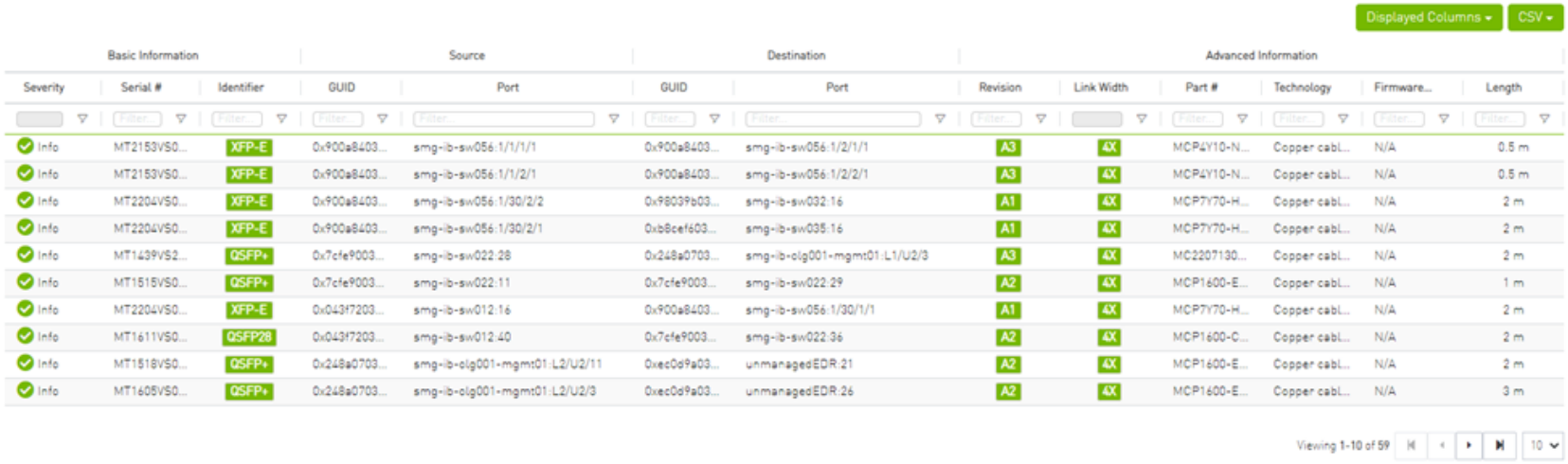
The health policy supports the following capabilities. When you select a policy and right-click, you can perform the following actions:

1. Delete the Policy
2. Mark the selected healthy policies as unhealthy (Isolate/No discover)
3. Mark the selected unhealthy policies as healthy

If you wish to delete all the healthy ports from the health policy, click on the 'Delete All Healthy Ports' option situated at the top right corner of the policy table.

Cables Window

Provides a list of all cables in UFM. For more information, see Device's Cables Tab.



Right-clicking a cable from the list allows users to Collect System Dump for the endpoints of the link.

Groups Window

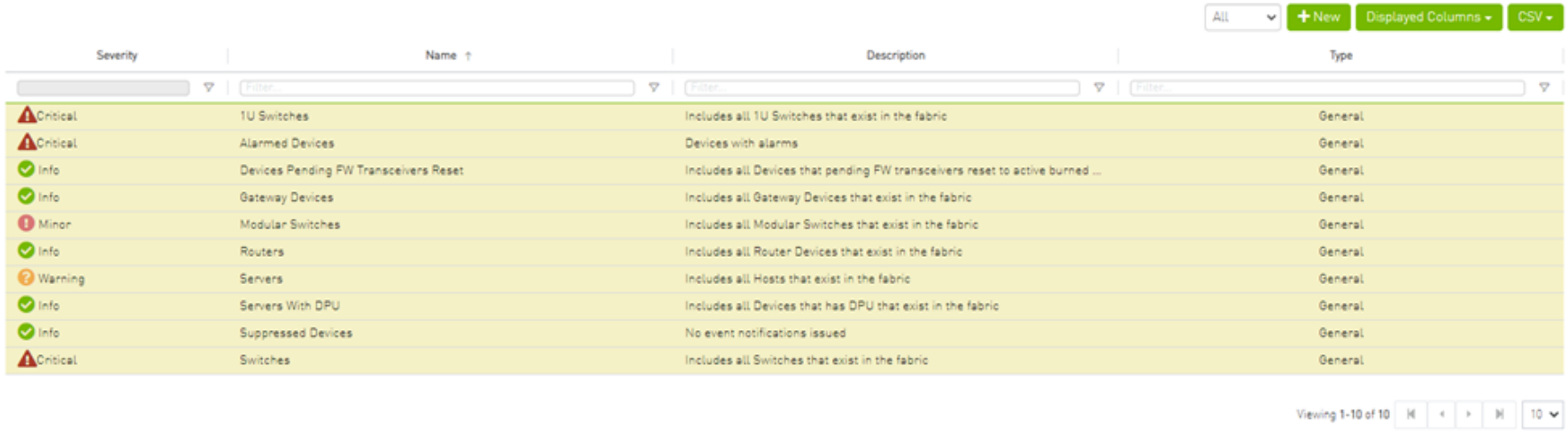
The Groups window allows users to create new groups of devices and provides information about existing groups.

> ⚠ All predefined groups have Read permissions only, except Suppressed_Devices to/from which the user is also able to add/remove members or devices.

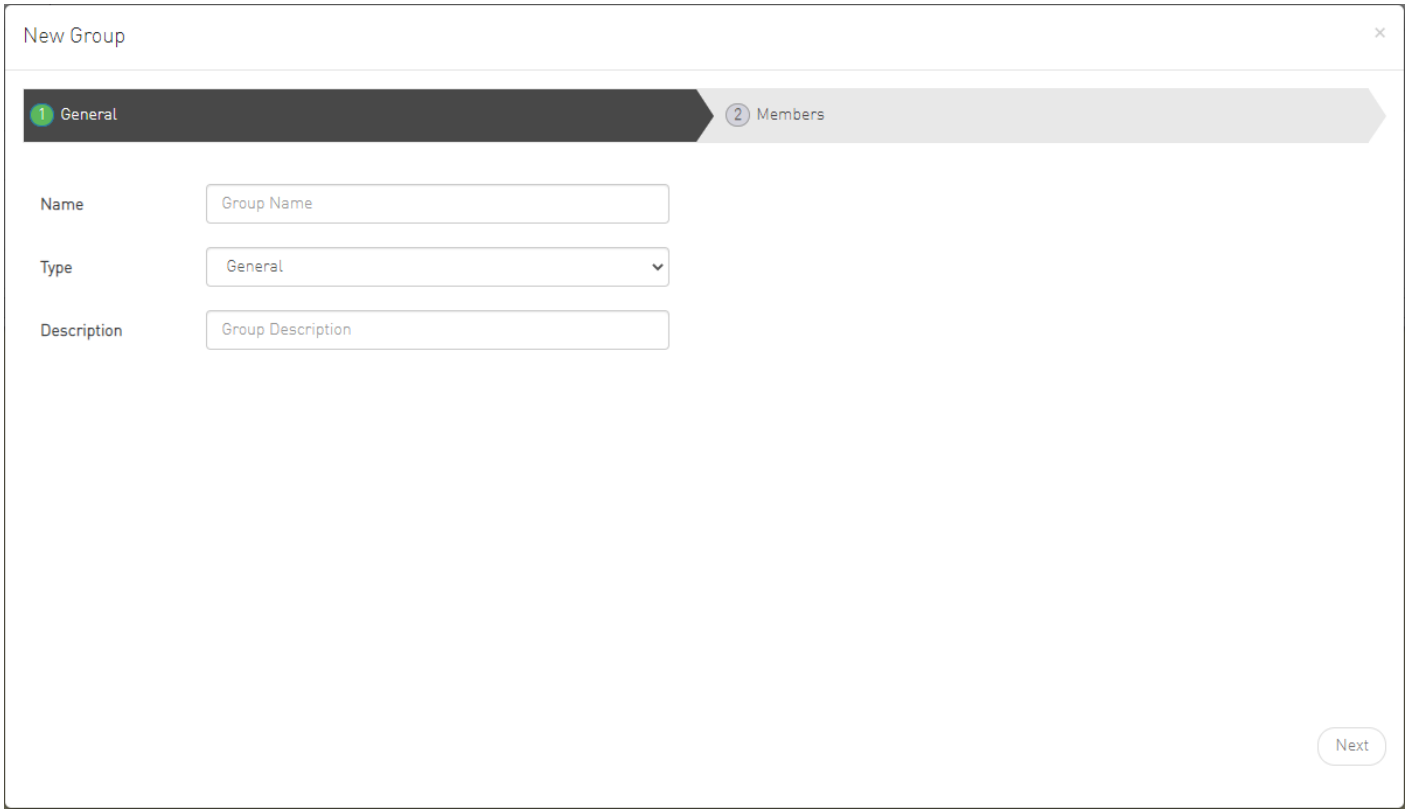
⚠ The following predefined groups auto-populate upon UFM startup: Switches, 1U_Switches, Modular_Switches, Gateway_Devices, and Hosts.

To create a group of devices, do the following:

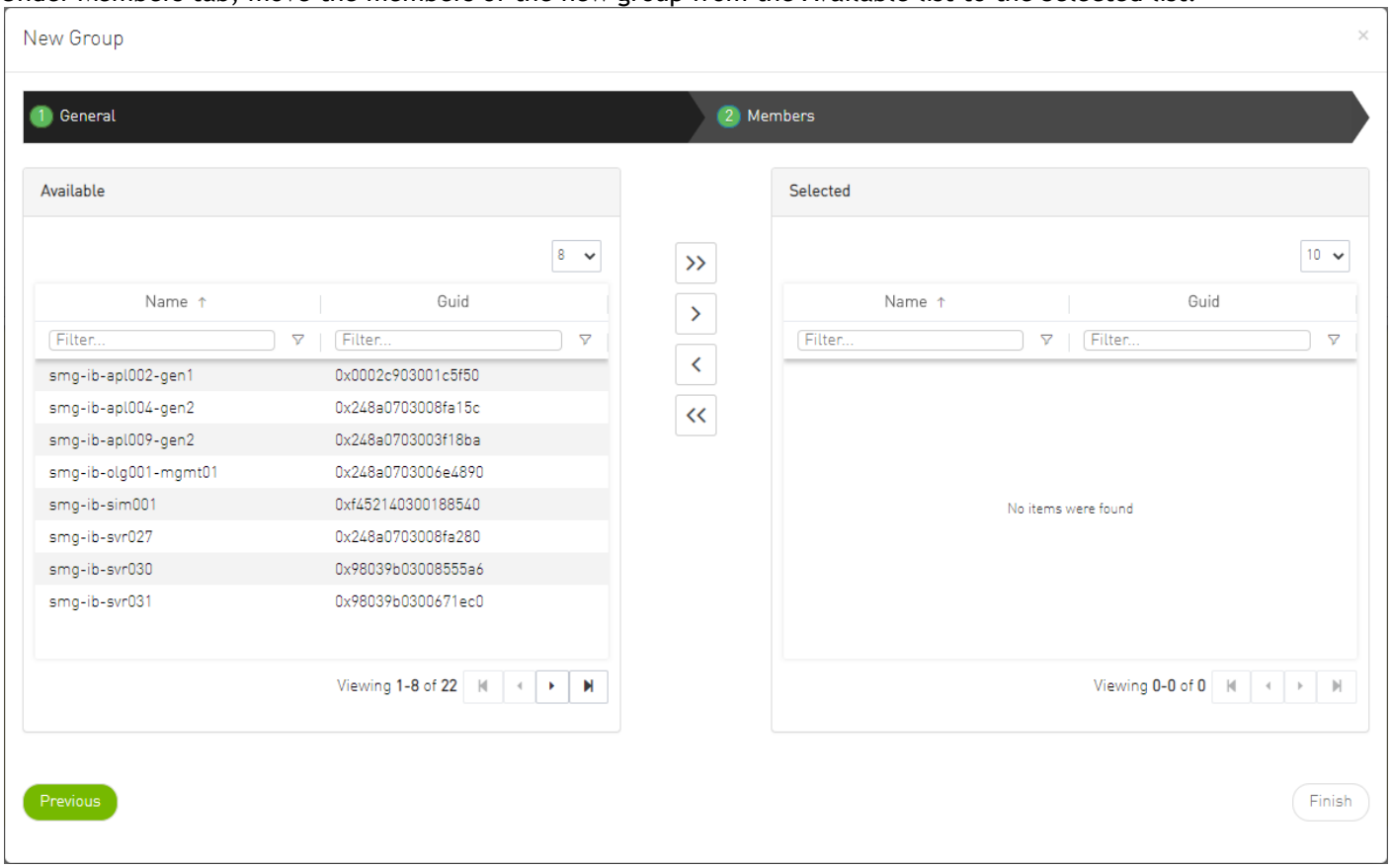
1. Click “New” under “Groups.”



2. In the New Group wizard, fill in the required information under the General tab: Name (must be between 4-20 characters), Type (General/Rack/Port), and Description (optional), and click Next.

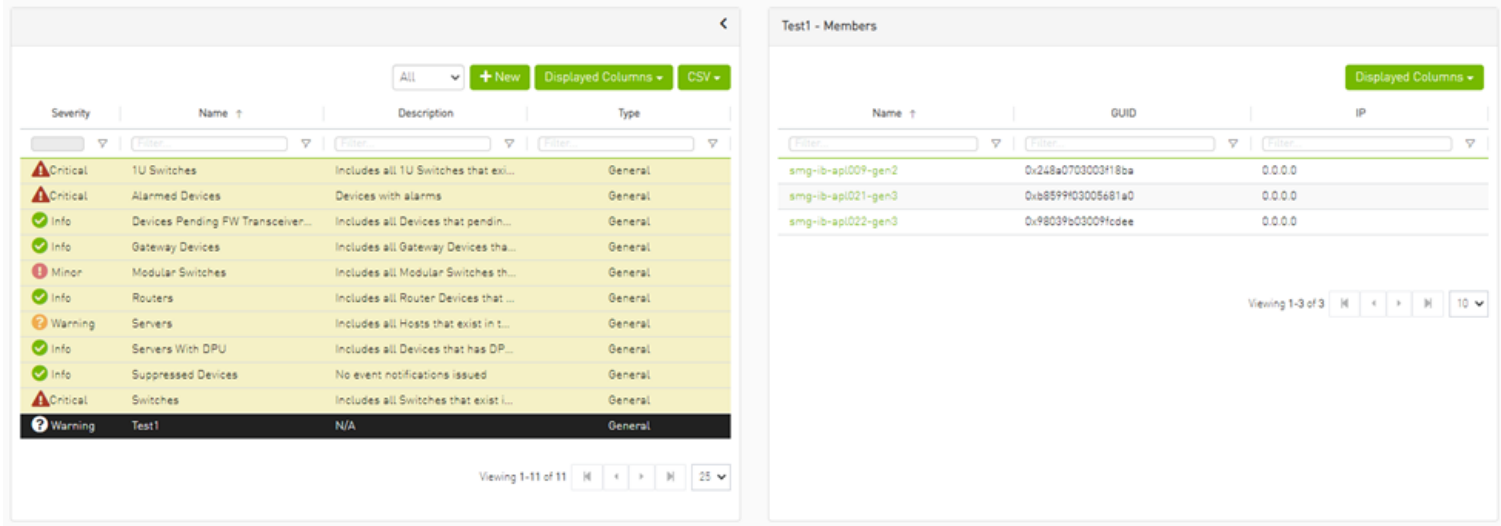

New Group
1 General
2 Members
Name
Group Name
Type
General
Description
Group Description
Next

3. Under Members tab, move the members of the new group from the Available list to the Selected list.



4. Click “Finish” and the new group will appear under the Groups window.

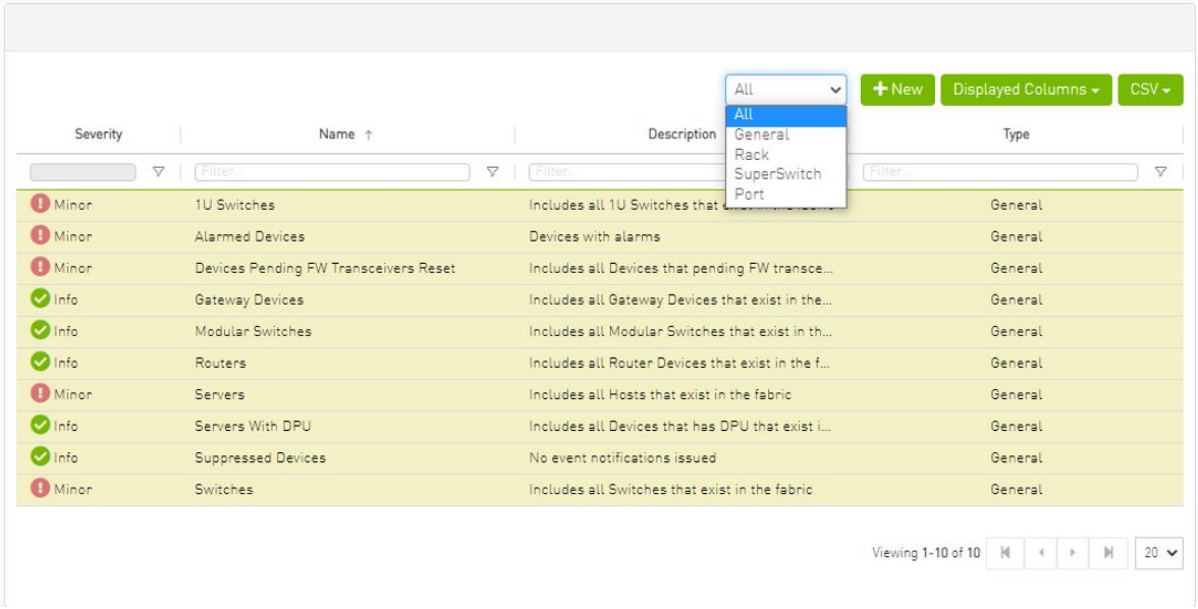
Group members details – port’s hostname, port’s GUID, and device’s IP address – can be viewed when selecting the group from the list of all groups available.



Group Actions

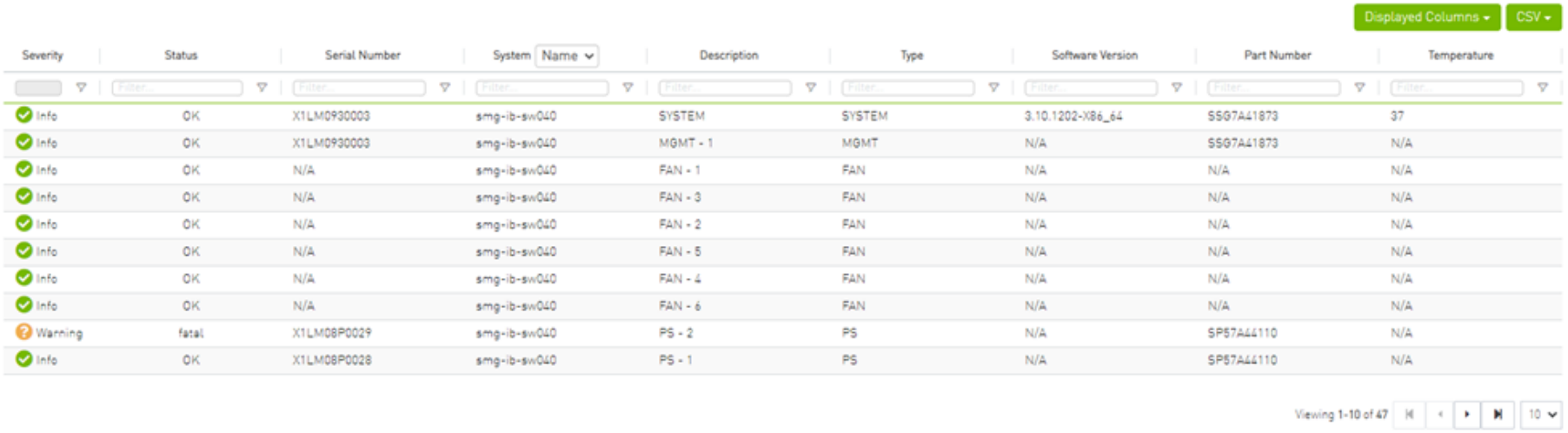
Right-clicking a group enables performing the following actions:

- **Edit** – groups can be modified either by editing the group description under **General** tab, or substituting group members under **Members** tab
- **Delete** – existing groups can be deleted from the list
- **Remove All Members** – all members of an existing group can be removed at once
- **Collect System Dump** – sysdump may be generated for all members of an existing group
  The user can filter group by type (General, Rack, Super Switch and Port)



Inventory Window

Provides a list of all modules in UFM. For more information, see Device's Inventory Tab.



PKeys Window

The PKeys window allows users to create new groups of ports and provides information about existing PKeys.

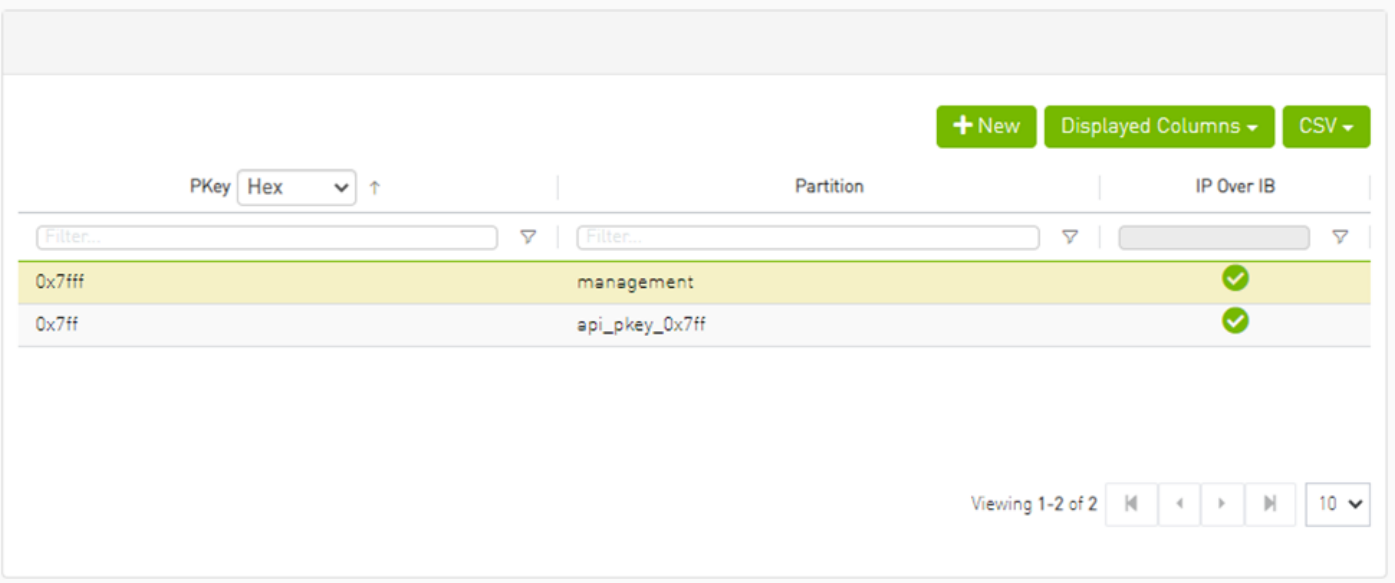
> ⚠ This window offers one predefined PKey (highlighted in the list of PKeys): Management key 0x7fff with Read permissions only.

For further information about InfiniBand partitioning (Pkeys management), please refer to the Partitioning Appendix.

Creating New PKey

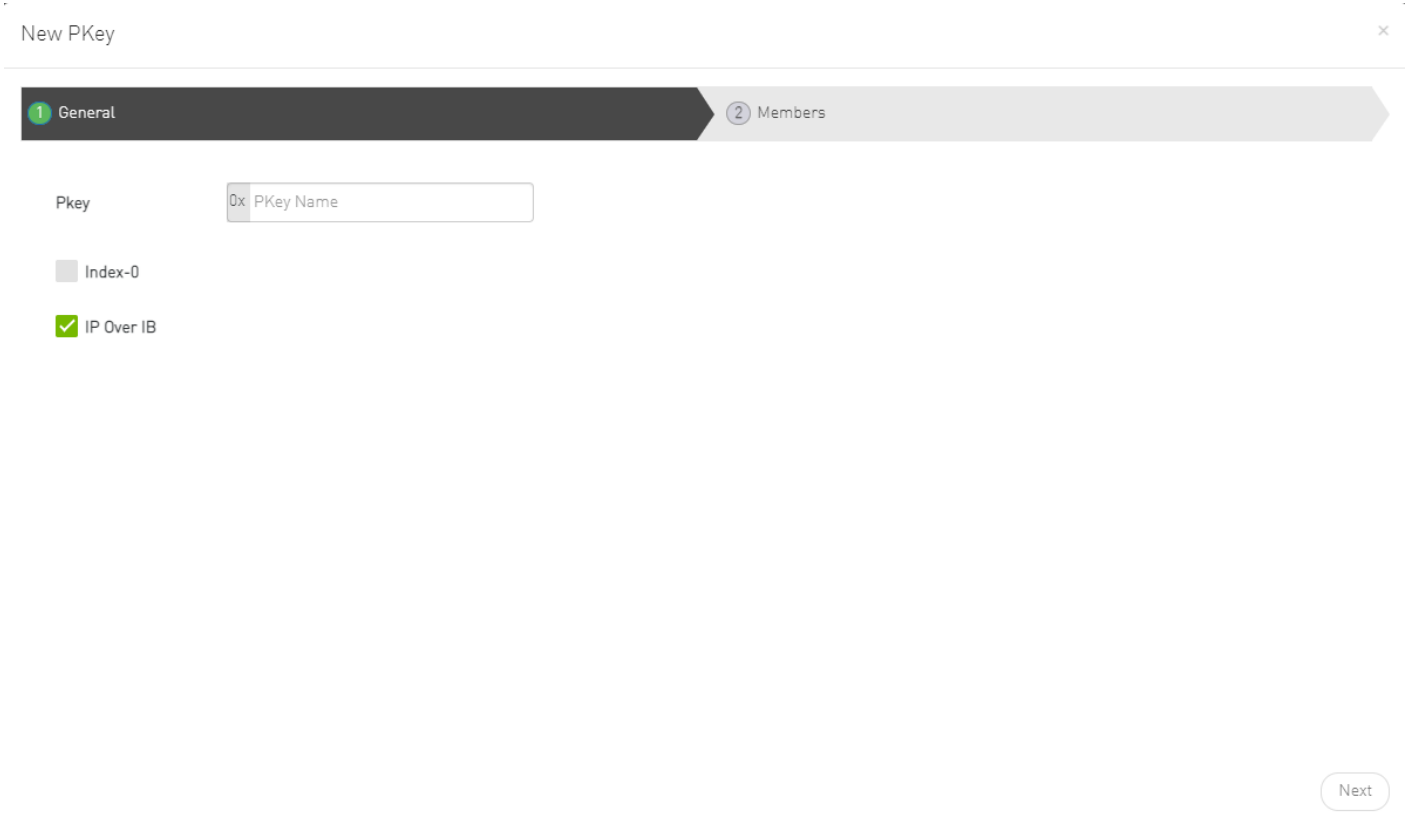
To create a PKey:

1. Click the “New” button under “PKeys".
   Please note that the yellow highlighted PKeys are predefined ones.

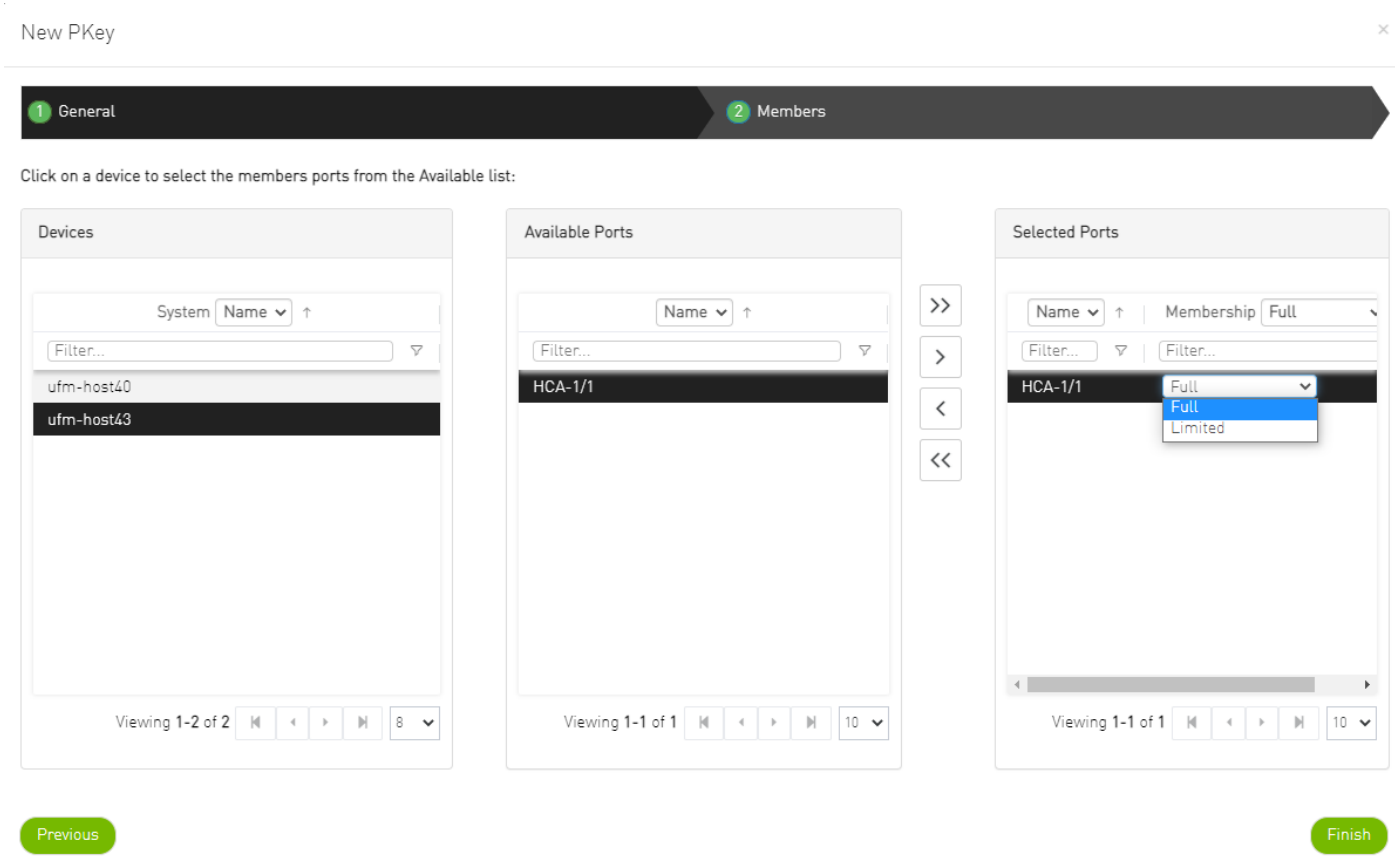


2. In the New PKey wizard, fill in the required information under the General tab:
    - Name—must be between 0x1 and 0x7fff, inclusive
    - Index-0 attribute—True/False

- IP Over IB attribute—True/False



3. Click "Next."
4. Under Members tab, select the device of which ports you would like to group in one PKey, and move the members (ports) from the Available list to the Selected list. For each member (port) you may specify a membership type (Full/limited).

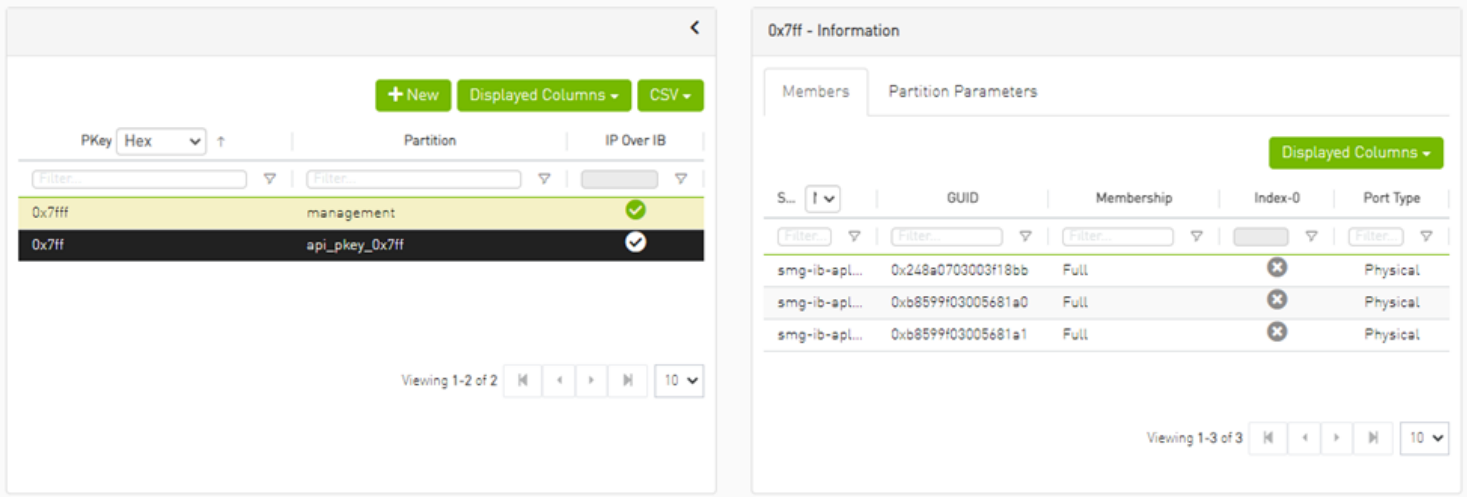


5. Click “Finish”. The new PKey will become available under the PKey window.

When selecting a PKey from the PKeys table, **PKey Information** table will appear on the right side of the screen. This table provides information on the PKey's members and QoS settings.

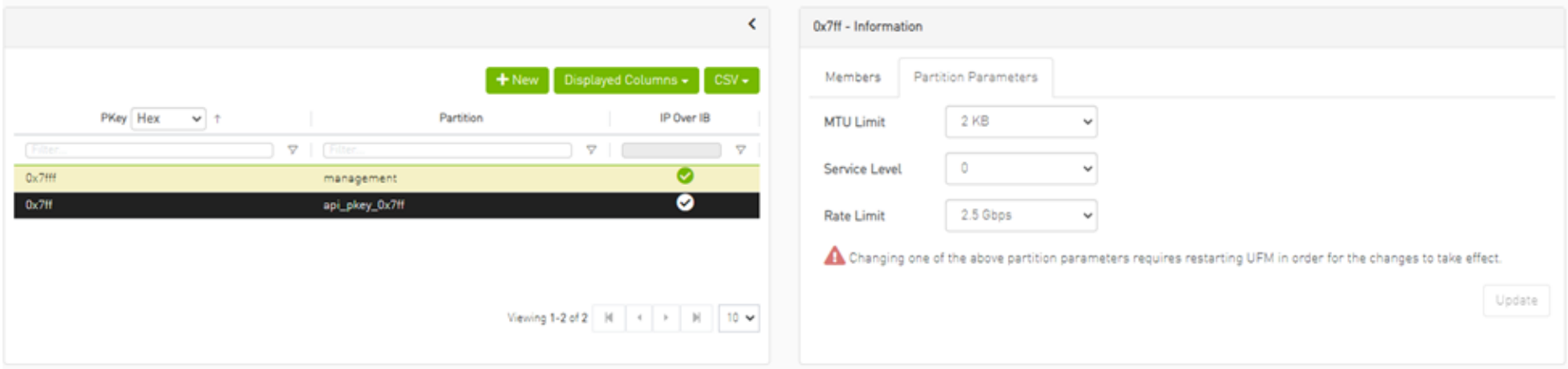
PKey Members Tab

Provides details on the PKey members: port’s hostname (node), device's IP address, port GUID, port number, membership and index-0 attributes values.



PKey QoS Tab

Displays the current partitioning parameter settings of the selected PKey: MTU Limit, Service Level and Rate limit. These settings can be modified by the user.

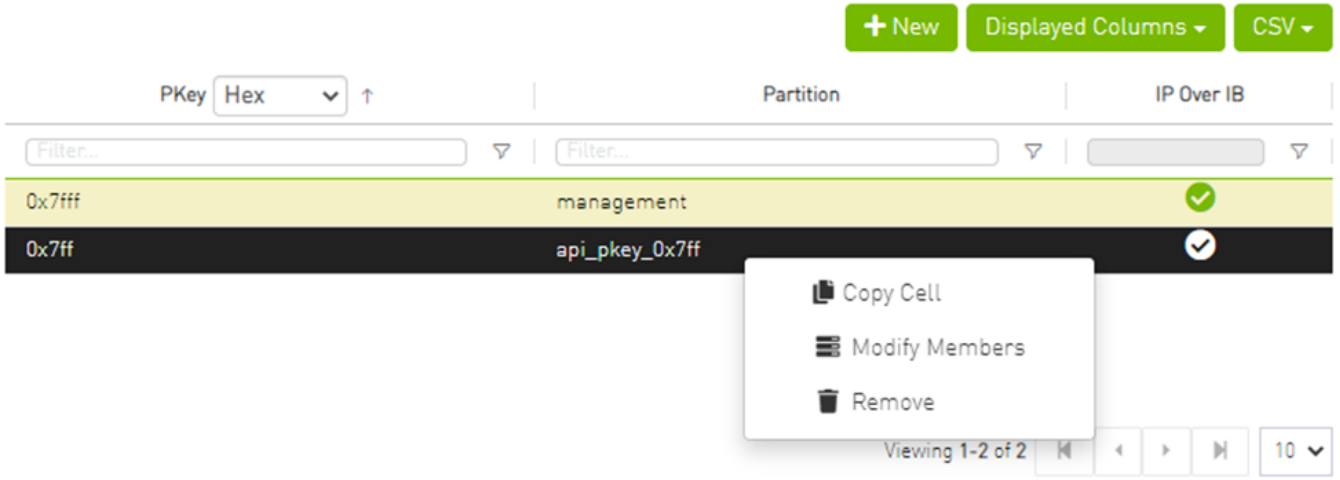


PKey Actions

Right-clicking one PKey from the list enables performing the following actions:

- **Modify Members** – PKeys can be modified either by editing the attributes under **General** tab, or updating the members under **Members** tab. Including updating ports memberships.

- **Remove** – existing PKeys can be deleted from the list.



> For information on partitioning, refer to Appendix – Partitioning.

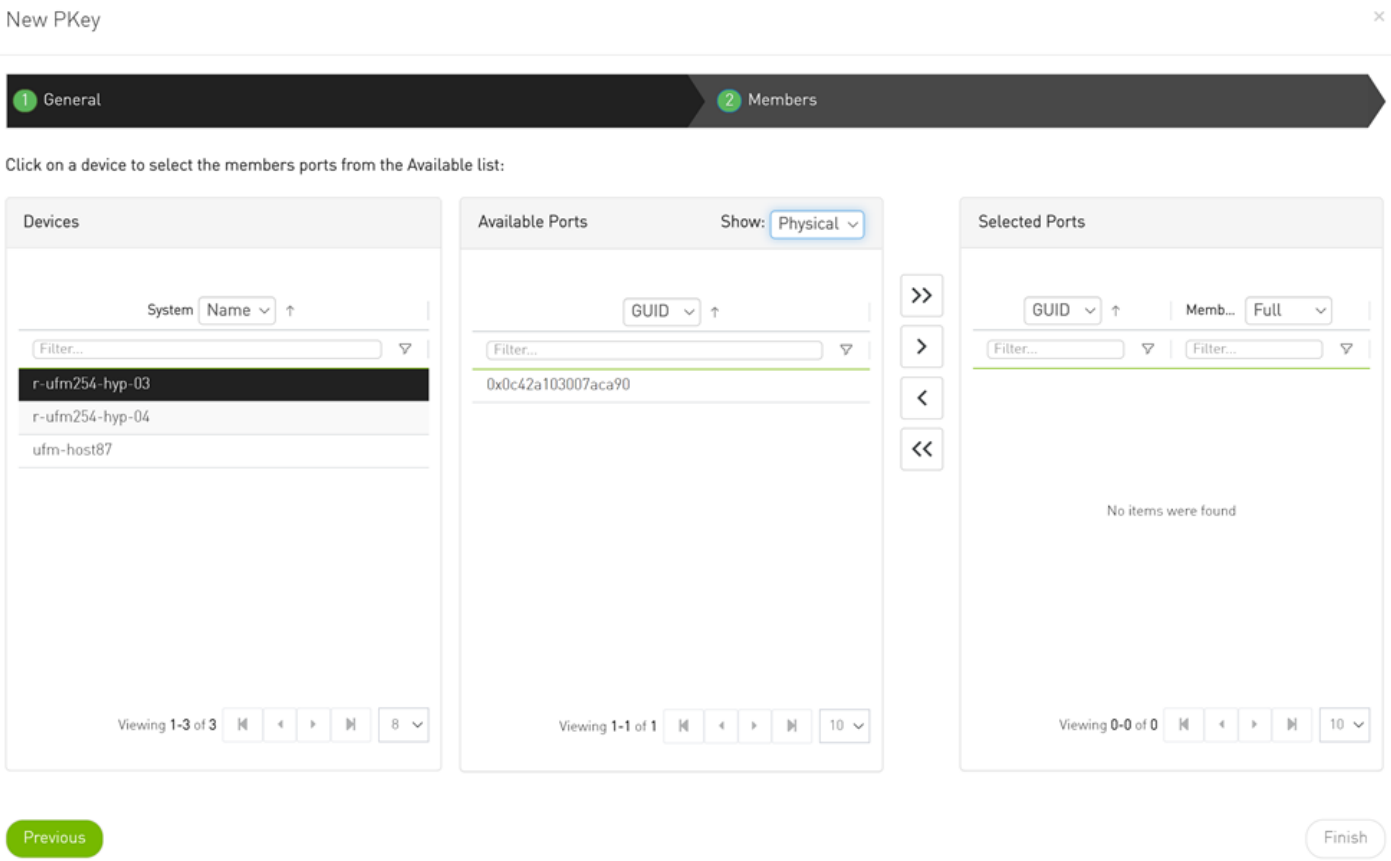
> Note that restarting OpenSM is required for the QoS parameters change to take effect.

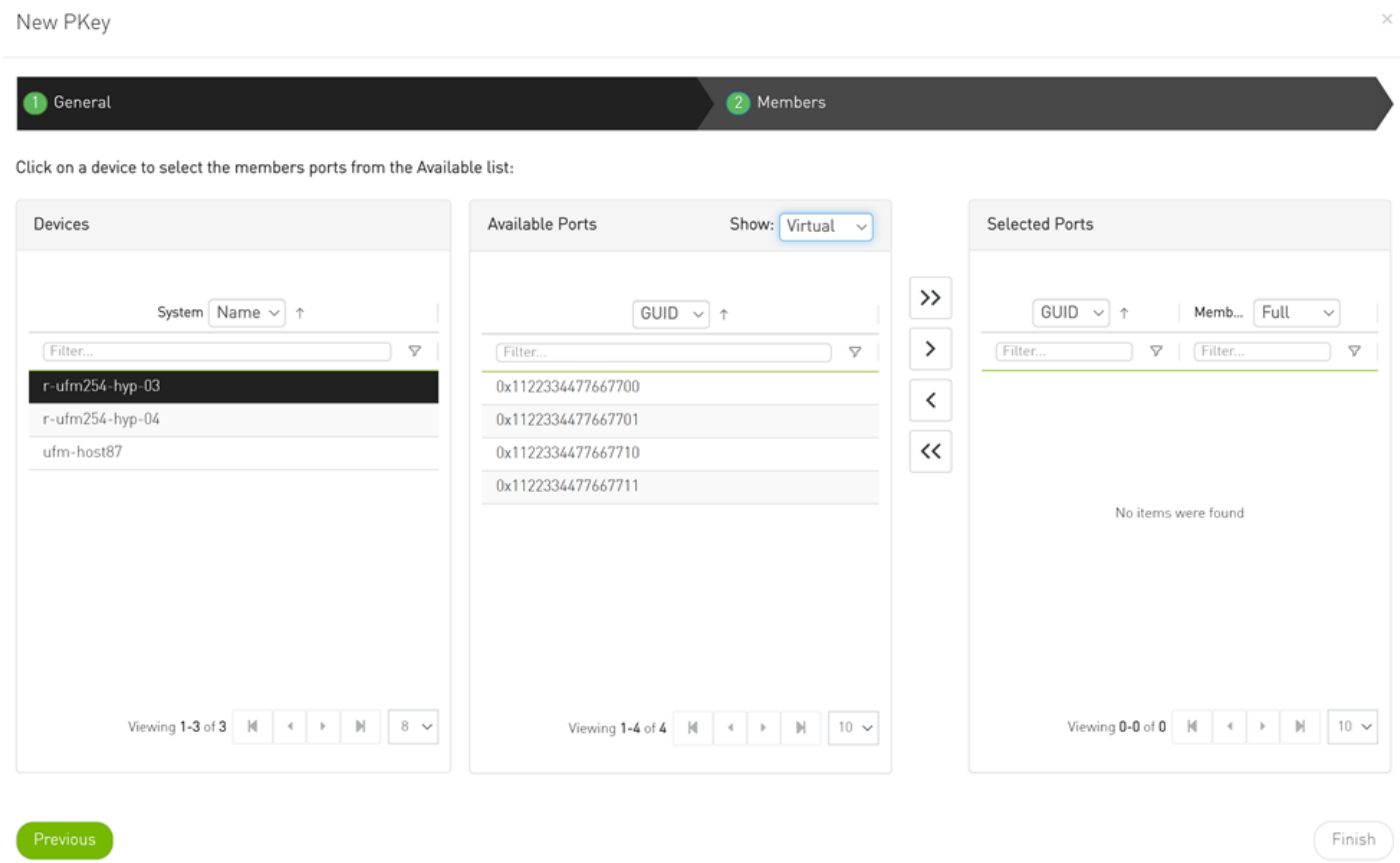
Support Pkey with Virtual Ports

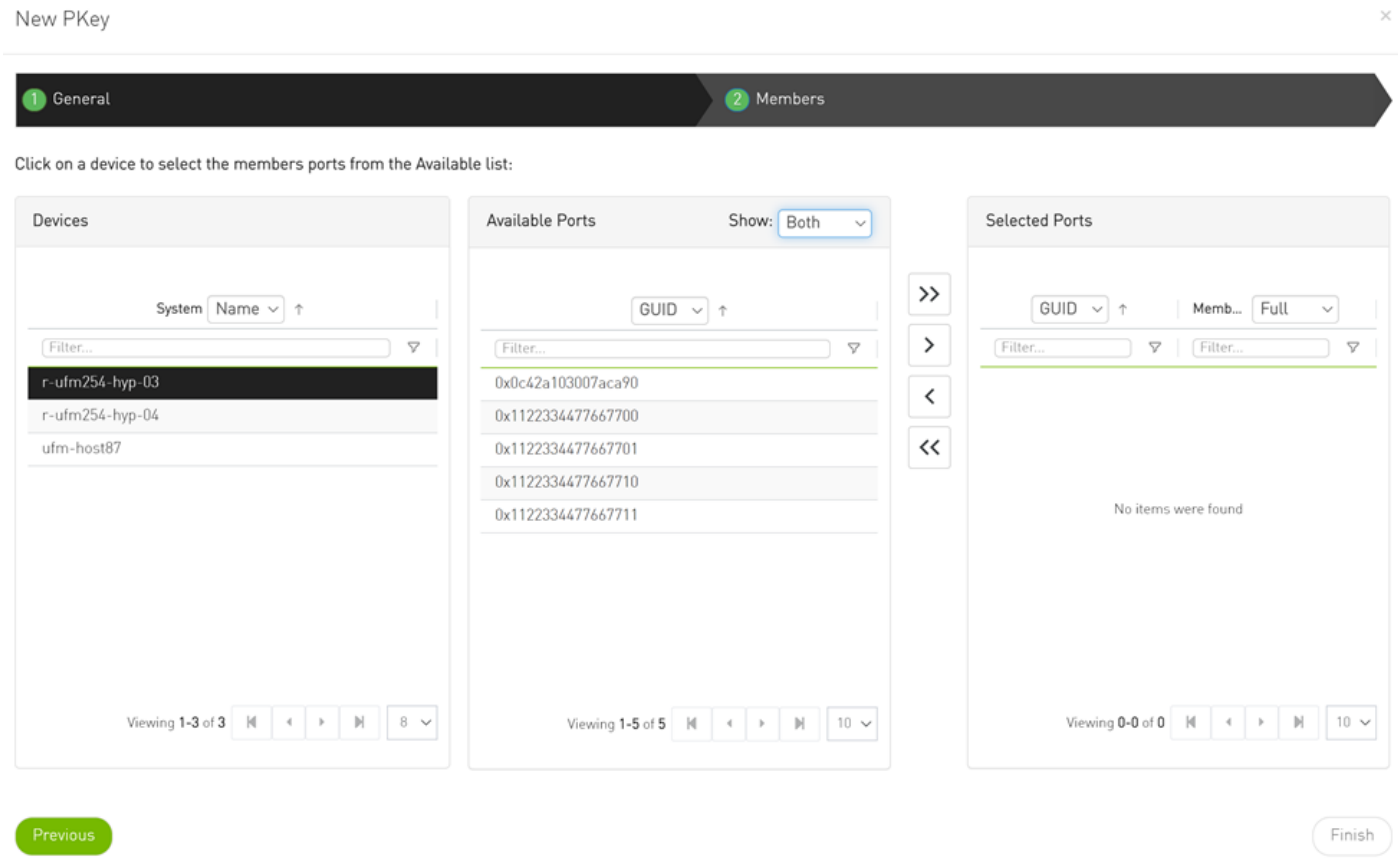
Creating a pkey with virtual ports is supported, so pkey can contain the following types of port:

- Physical
- Virtual
- Both physical and virtual

The create new pkey wizard dropdown includes port types.

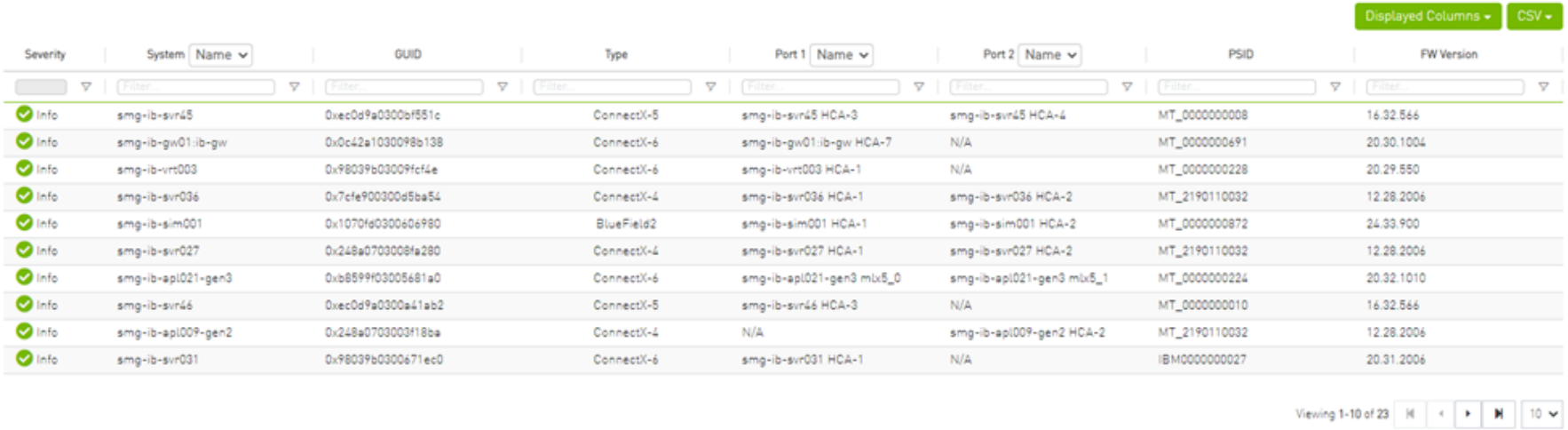

New PKey
1 General
2 Members
Click on a device to select the members ports from the Available list:
Devices
System Name
Filter...
r-ufm254-hyp-03
r-ufm254-hyp-04
ufm-host87
Viewing 1-3 of 3
8
Available Ports
Show: Physical
GUID
Filter...
0x0c42a103007aca90
Viewing 1-1 of 1
10
Selected Ports
GUID
Memb... Full
Filter...
Filter...
No items were found
Viewing 0-0 of 0
10
Previous
Finish


New PKey
1 General
2 Members
Click on a device to select the members ports from the Available list:
Devices
System Name
Filter...
r-ufm254-hyp-03
r-ufm254-hyp-04
ufm-host87
Viewing 1-3 of 3
8
Available Ports
Show: Virtual
GUID
Filter...
0x1122334477667700
0x1122334477667701
0x1122334477667710
0x1122334477667711
Viewing 1-4 of 4
10
Selected Ports
GUID
Memb... Full
Filter...
Filter...
No items were found
Viewing 0-0 of 0
10
Previous
Finish



HCAs Window

Provides a list of all the HCAs of the hosts in UFM. For more information, see section "HCAs Tab".

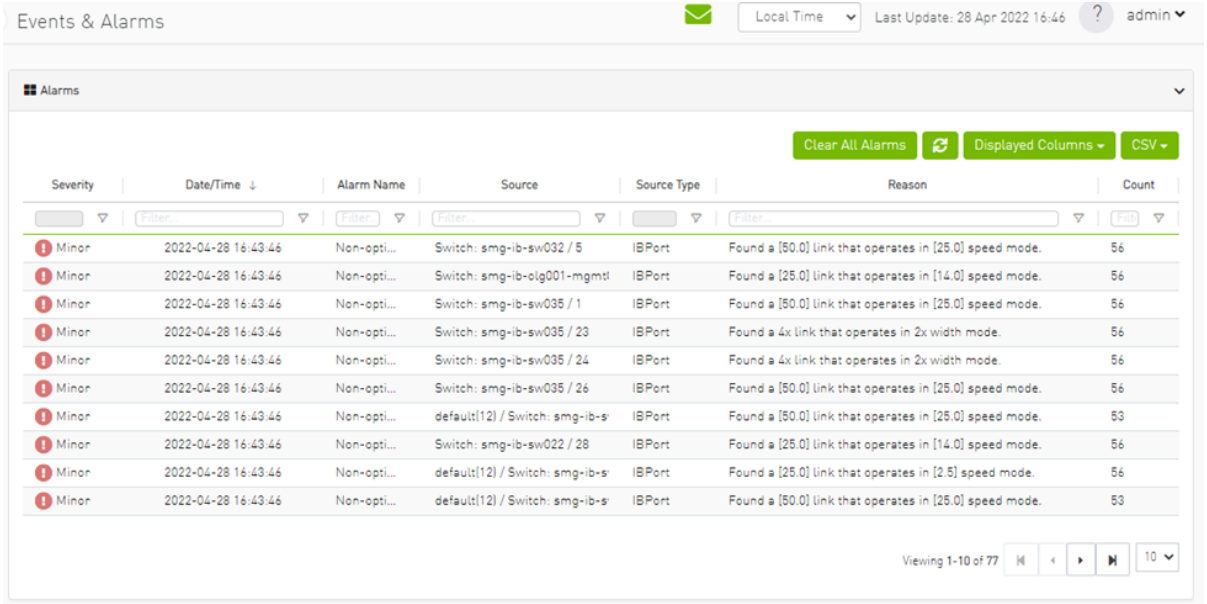


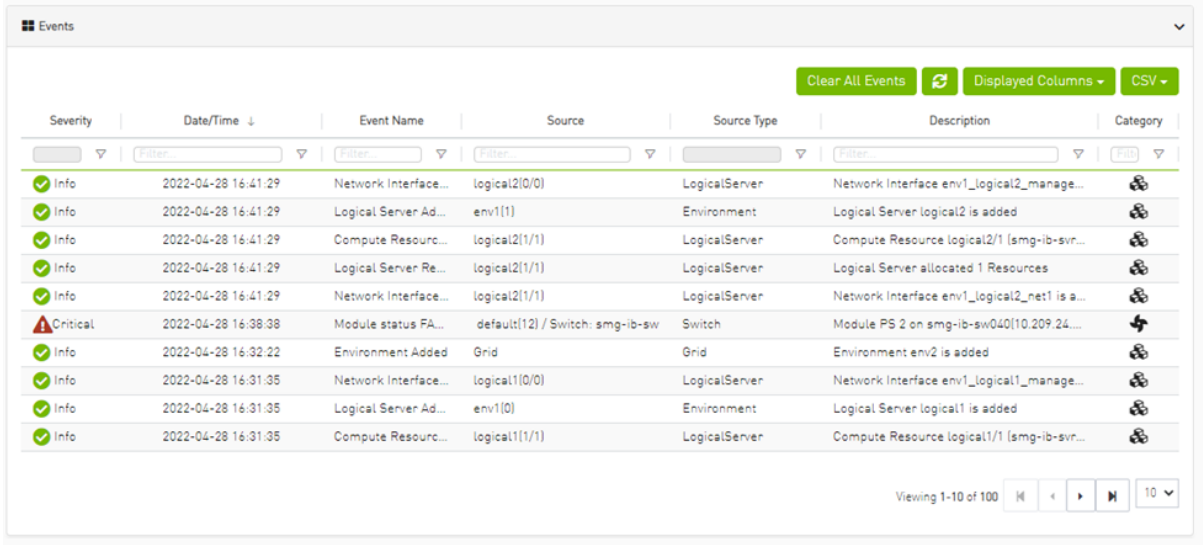
Events & Alarms

> ⚠ All information provided in a tabular format in UFM web UI can be exported into a CSV file.

UFM allows you to identify any problem including ports and device connectivity using events and alarms. Problems can be detected both prior to running applications and during standard operation.

Events trigger alarms (except for "normal" events. i.e., Info events) when they exceed a predefined threshold. Events and alarms can be configured under Events Policy tab under Settings window. For more information, refer to Events Policy Tab.


Events & Alarms
Local Time
Last Update: 28 Apr 2022 16:46
admin
Alarms
Clear All Alarms
Displayed Columns
CSV
Severity
Date/Time
Alarm Name
Source
Source Type
Reason
Count
Minor 2022-04-28 16:43:46 Non-opti... Switch: smg-ib-sw032 / 5 IBPort Found a [50.0] link that operates in [25.0] speed mode. 56
Minor 2022-04-28 16:43:46 Non-opti... Switch: smg-ib-olg001-mgmtl IBPort Found a [25.0] link that operates in [14.0] speed mode. 56
Minor 2022-04-28 16:43:46 Non-opti... Switch: smg-ib-sw035 / 1 IBPort Found a [50.0] link that operates in [25.0] speed mode. 56
Minor 2022-04-28 16:43:46 Non-opti... Switch: smg-ib-sw035 / 23 IBPort Found a 4x link that operates in 2x width mode. 56
Minor 2022-04-28 16:43:46 Non-opti... Switch: smg-ib-sw035 / 24 IBPort Found a 4x link that operates in 2x width mode. 56
Minor 2022-04-28 16:43:46 Non-opti... Switch: smg-ib-sw035 / 26 IBPort Found a [50.0] link that operates in [25.0] speed mode. 56
Minor 2022-04-28 16:43:46 Non-opti... default[12] / Switch: smg-ib-s IBPort Found a [50.0] link that operates in [25.0] speed mode. 53
Minor 2022-04-28 16:43:46 Non-opti... Switch: smg-ib-sw022 / 28 IBPort Found a [25.0] link that operates in [14.0] speed mode. 56
Minor 2022-04-28 16:43:46 Non-opti... default[12] / Switch: smg-ib-s IBPort Found a [25.0] link that operates in [2.5] speed mode. 56
Minor 2022-04-28 16:43:46 Non-opti... default[12] / Switch: smg-ib-s IBPort Found a [50.0] link that operates in [25.0] speed mode. 53
Viewing 1-10 of 77
10



Users can enable the events persistency mechanism from the gv.cfg. This allows the user to see the events in the case of restarting the UFM or in HA mode.

> ⚠ Alternatively you can run the following commands:
> - ufm events persistency enable
> - ufm events max-restored

The persistency is deactivated by default and can be enabled by the following controlled parameters in the config file:

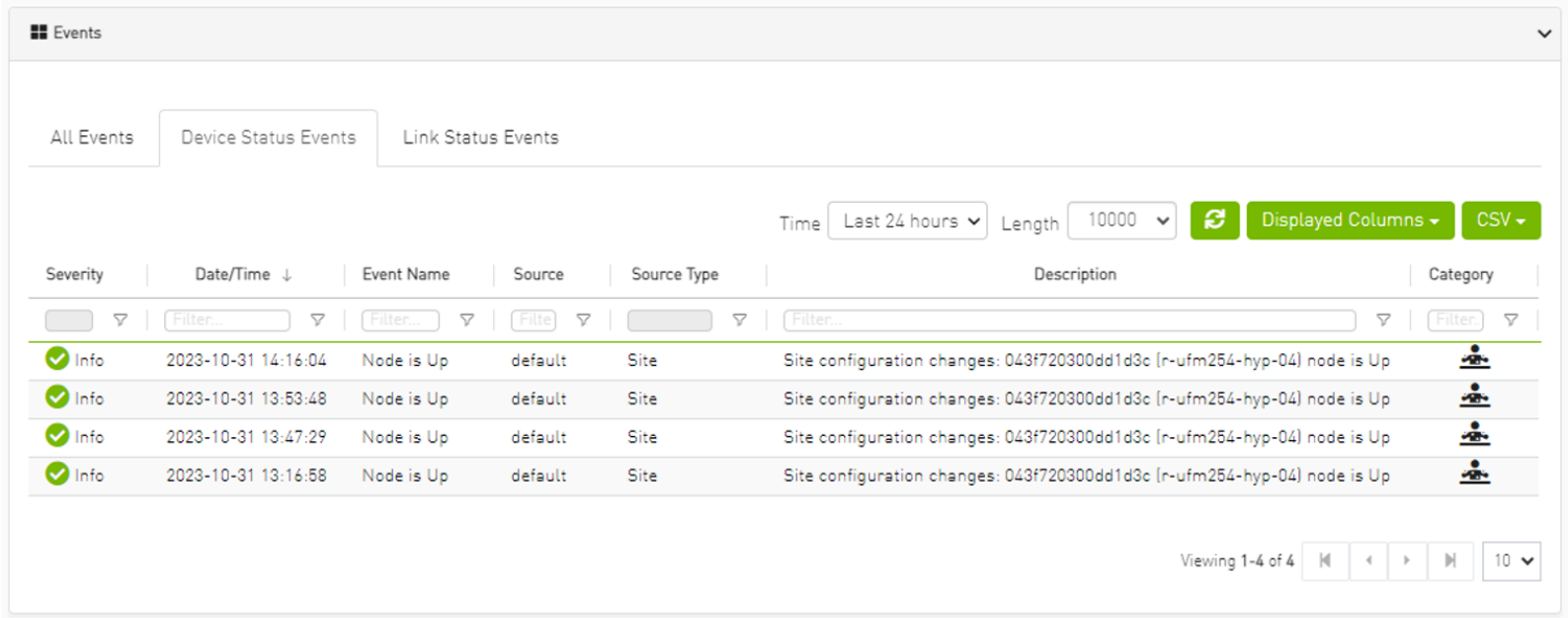
- `max_restored_events = 50 #` - will determine the number of events to restore
- `events_persistency_enabled = true #` - will set to true for the feature to work

Device Status Events

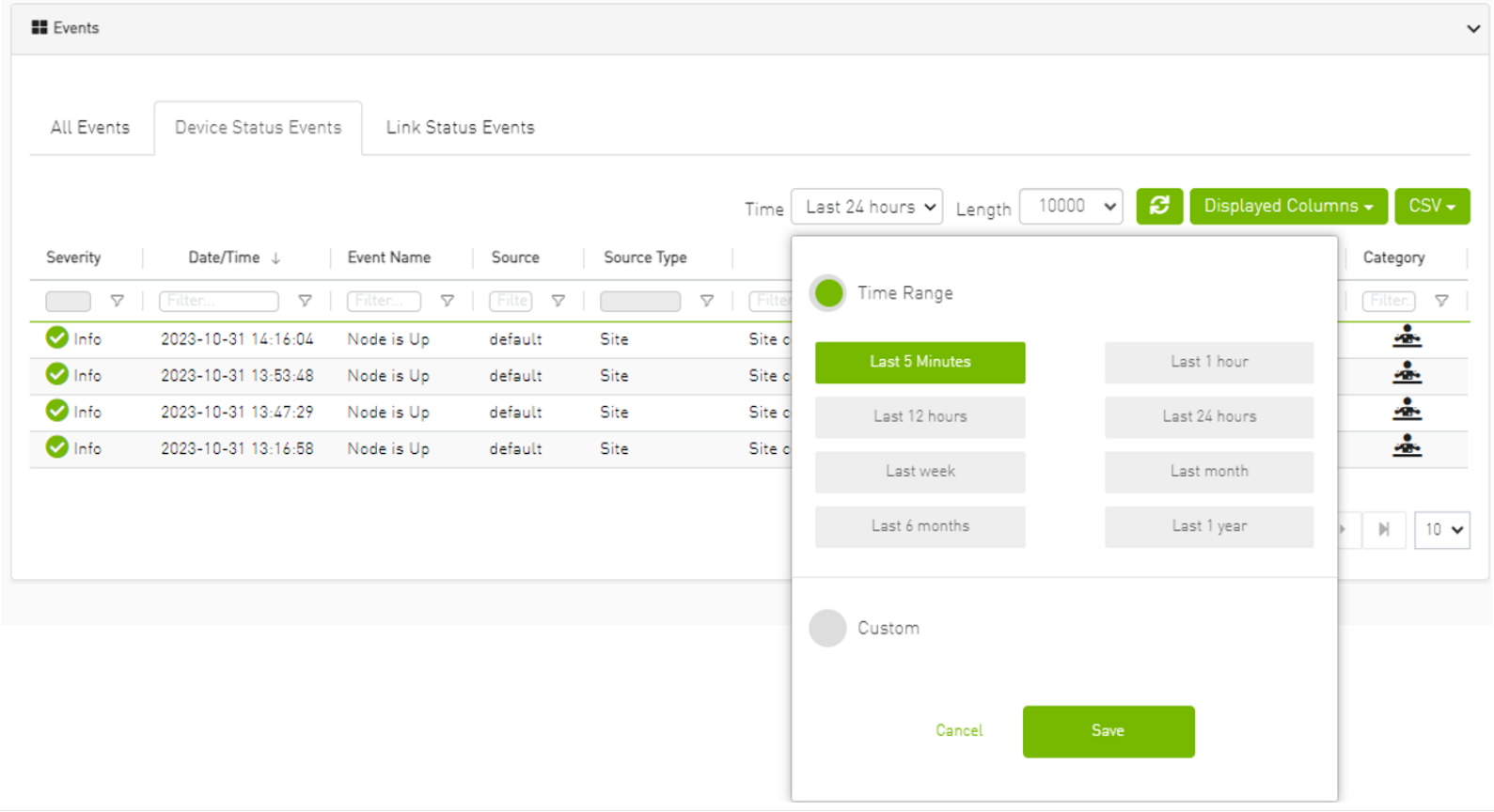
The Device Status Events tab displays topology change events related to devices in a table. it will support the following event types:

- None is Up/Down

- Switch is Up/Down
- Director Switch is Up/Down



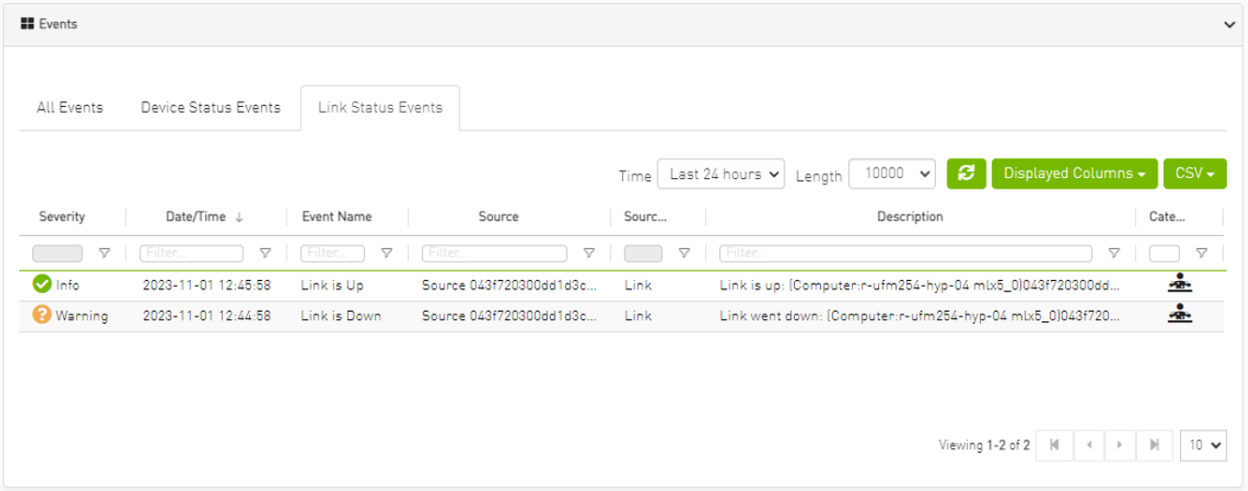
Filters are be provided to allow events filtering by the desired time interval with a length limit.



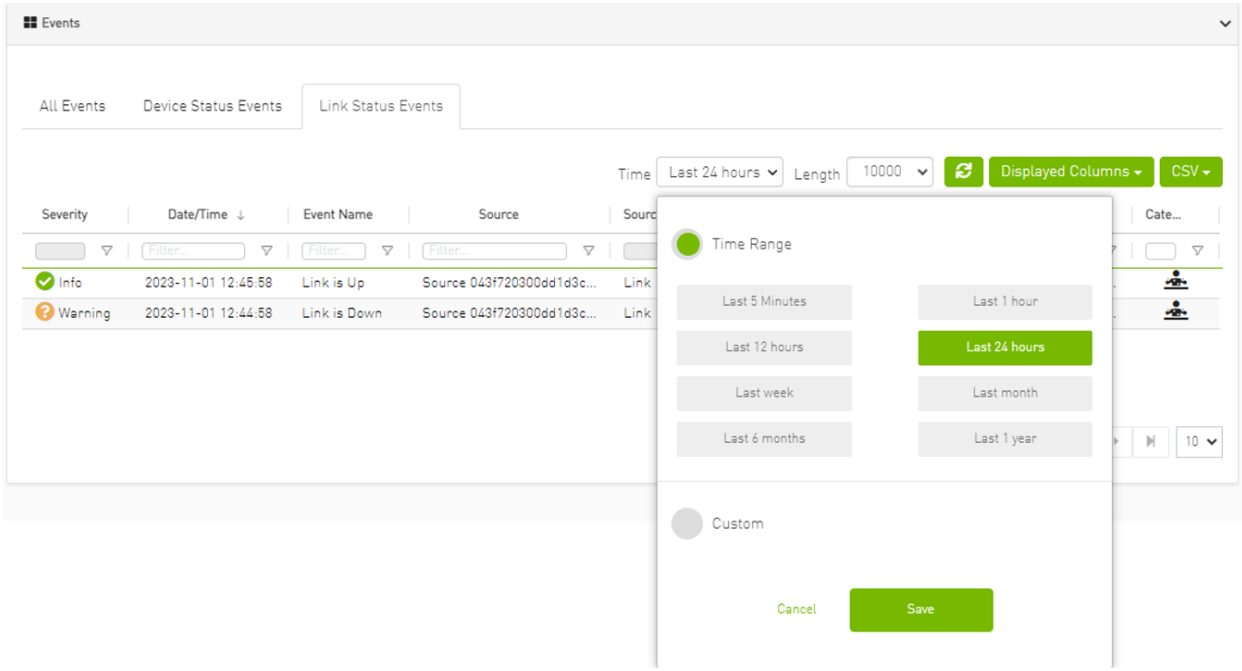
Link Status Events

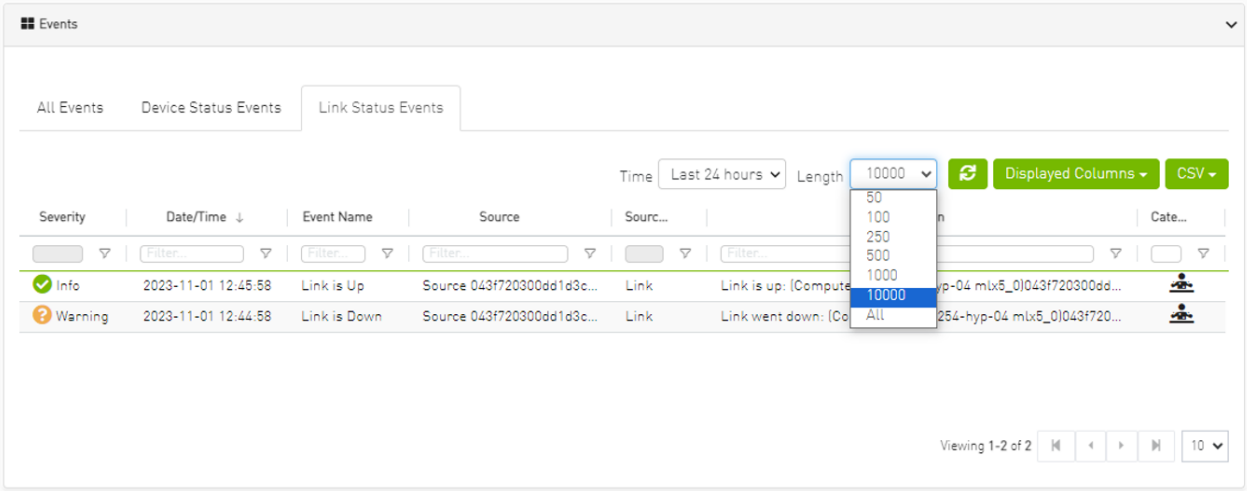
The Link Status Events tab displays topology change events related to links in a table. It supports the following event type:

- Link is Up/Down



Filters are provided to allow filtering by the desired time interval in a time range.


Events
All Events
Device Status Events
Link Status Events
Time
Last 24 hours
Length
10000
Displayed Columns
CSV
Severity
Date/Time
Event Name
Source
Cate...
Info
2023-11-01 12:45:58
Link is Up
Source 043f720300dd1d3c...
Warning
2023-11-01 12:44:58
Link is Down
Source 043f720300dd1d3c...
Time Range
Last 5 Minutes
Last 1 hour
Last 12 hours
Last 24 hours
Last week
Last month
Last 6 months
Last 1 year
Custom
Cancel
Save



> **Warning** Rge related switch context menu is displayed only if the event type is 'Switch is Up/Down'. Other event types show the default context menu, which is 'Copy Cell'.

Telemetry

Error: null

System Health

The System Health window enables running and viewing reports and logs for monitoring and analyzing UFM server and fabric health through the following tabs: UFM Health, UFM Logs, UFM Snapshot, Fabric Health, Daily Reports and Topology Compare.

- UFM Health Tab
- UFM Logs Tab
- Fabric Health Tab
- Daily Reports Tab

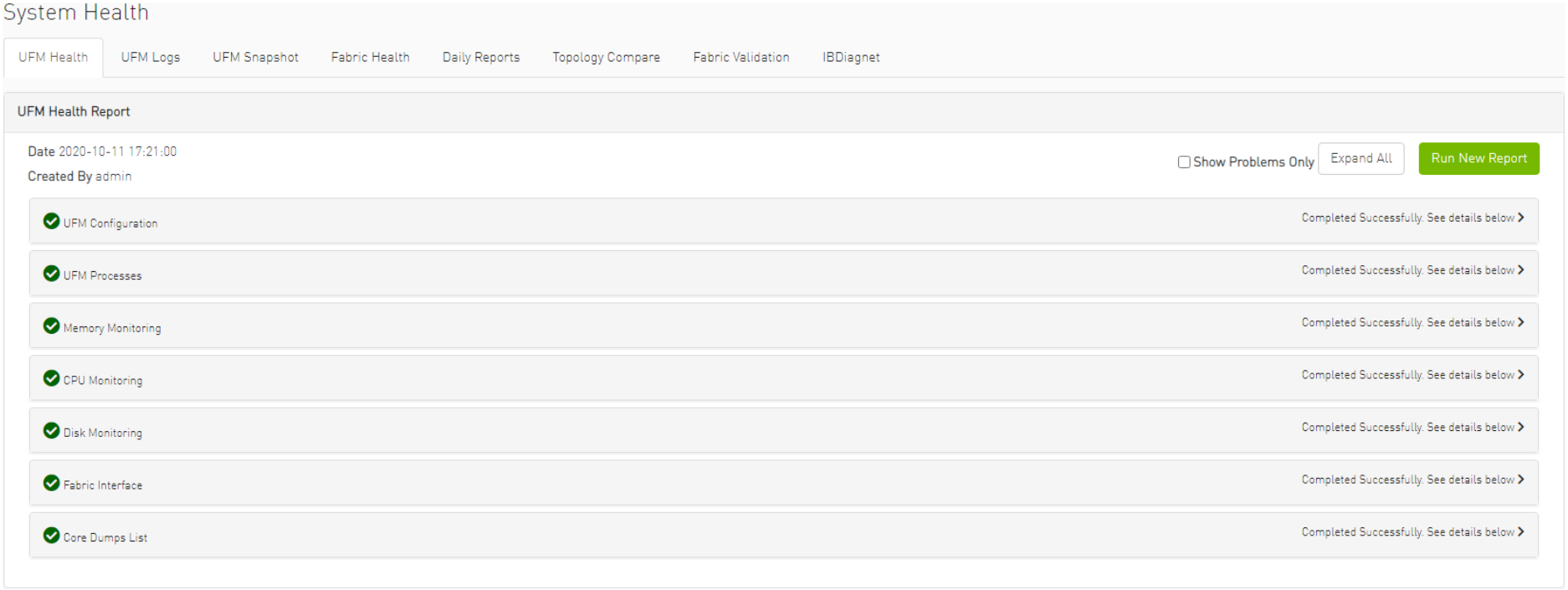
- Topology Compare Tab
- Fabric Validation Tab
- IBDiagnet Tab

UFM Health Tab

Through **UFM Health** tab, you can create reports that run a series of checks on the UFM server.

Each check that is run for a report triggers a corresponding event. Events are also triggered when a report starts and ends. For more information, see Events & Alarms.

To run a new report, click “Run New Report”. Results will be displayed inline automatically.



You can expand the results of each check or expand the results of all checks at once by clicking the "Expand All" button.

To view only the errors of the report results, click the "Show Problems Only" checkbox.

The following tables describe the checks included in the report.

UFM Health Report Checks

UFM Configuration	
Check	**Description**
Release Number	UFM software version and build.
License Type	Type of license, permanent or evaluation.
License Customer Number	The customer number provided by NVIDIA.
License UID	The UFM serial number provided by NVIDIA.
License Expiration Date	License expiration date for limited licenses.
License Functionality	Level of functionality enabled for the end-user, standard or advanced.
License Devices Limit	The maximum number of devices that UFM is licensed to manage. Note that it displays the current active and valid UFM licenses (not the sum of all valid licenses devices)
Running Mode	UFM running mode, Standalone or High Availability (HA). When UFM is in HA mode, additional information is displayed for the master and standby servers.

UFM Processing	
Check	**Description**
OpenSM	Status of the OpenSM service.
ibpm	Status of the ibpm (Performance Manager) service.
ModelMain	Status of the main UFM service.
httpd	Status of the httpd service.
MySql	Status of the MySql service.

Memory Monitoring	
Check	**Description**
Total memory usage	Percentage of total memory usage.
UFM memory usage	Percentage of UFM memory usage

CPU Monitoring	
Check	**Description**
Total CPU Capacity	Percentage of CPU capacity available
CPUs Number	Number of CPUs

CPU Monitoring	
Check	**Description**
Total CPU utilization	Percentage of total CPU utilization.
UFM CPU utilization	Percentage of UFM CPU utilization.

Disk Monitoring	
Check	**Description**
Disk <diskname>	Percentage of disk usage.

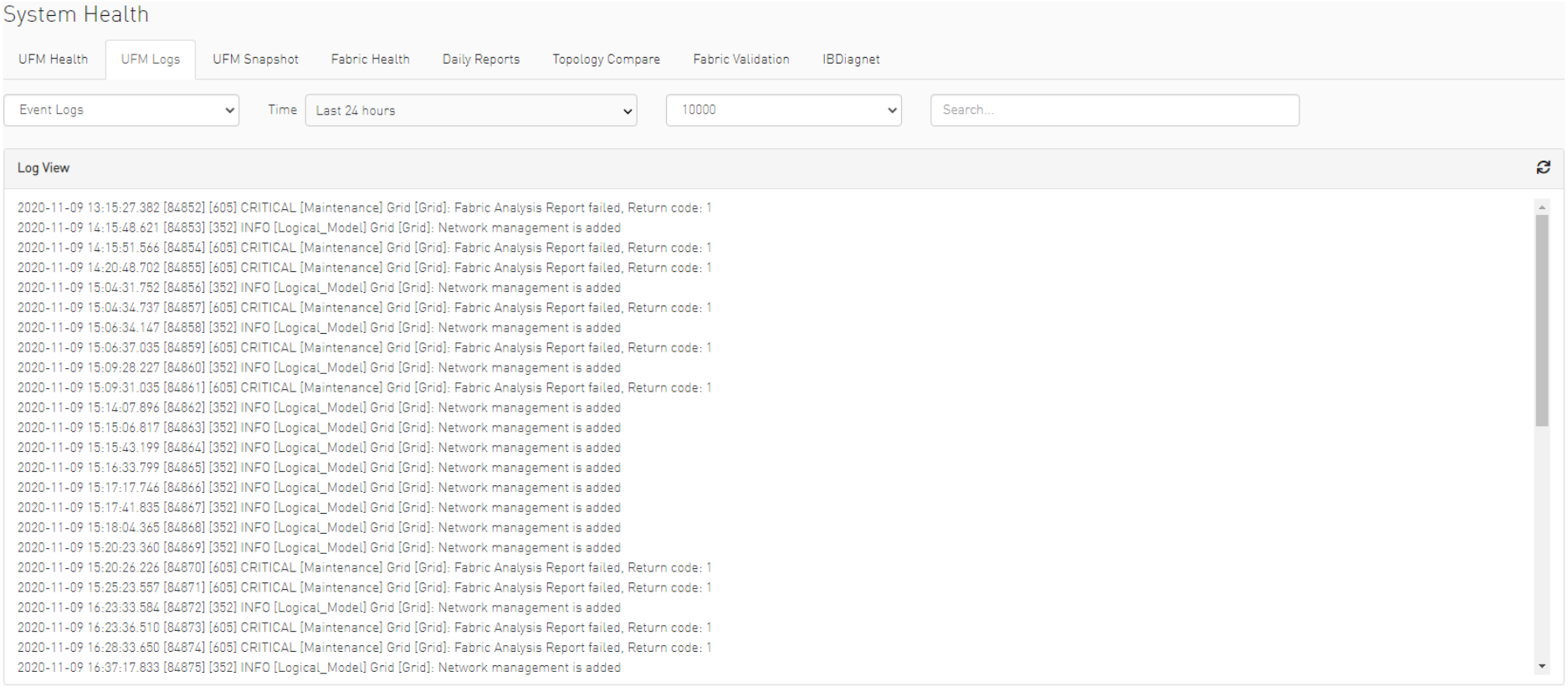
Fabric Interface	
Check	**Description**
Fabric Interface	Name and state of fabric interface.

UFM Logs Tab

UFM logging records events and actions that can serve to identify fabric and UFM server issues and assist in troubleshooting.

The logs are categorized into three files according to the activities they record: **Event** logs, **SM** logs, and **UFM** logs.

To view the log files, select the desired log file from the drop-down menu. Log data will be displayed:



In the Logs window, you can do the following:

- Refresh the data using the Refresh button on the right-hand side of the screen
- Search for a specific value using the Search bar
- Limit the display to a specific time period using the Time drop-down menu
- Limit the display to a specific number of lines using the drop-down menu (use "All" option to display all lines)
- Control the display of log occurrences by either showing all lines or hiding the duplicated ones.

Event Logs

Event Logs show the history of fabric events detected and initiated by the UFM server. The timestamp and severity of an event is indicated as well as the cause of the event and additional relevant information. *The Event log is kept on the UFM server in the events.log* file. Events can be configured whether to appear in the log files under the Events Policy tab in the Settings window. For more information, see Events Policy.

See "Appendix - Supported Port Counters and Events" for a comprehensive list of Events.

Subnet Manager (SM) Logs

SM Logs show messages of the Subnet Manager and communication plug-in.

The log verbosity is defined by selecting the Log Levels in the Subnet Manager tab under Settings window. For more information, see Subnet Manager Tab.

UFM Logs

UFM Logs is a general log of UFM Server. The log saves a history of user actions, events, polling results and other server activities and errors. Log verbosity is defined on start-up in the configuration file conf/gv.cfg:

```
[Logging]
# optional logging levels
#CRITICAL, ERROR, WARNING, INFO, DEBUG
level = WARNING
```

The default verbosity level is WARNING.

Fabric Health Tab

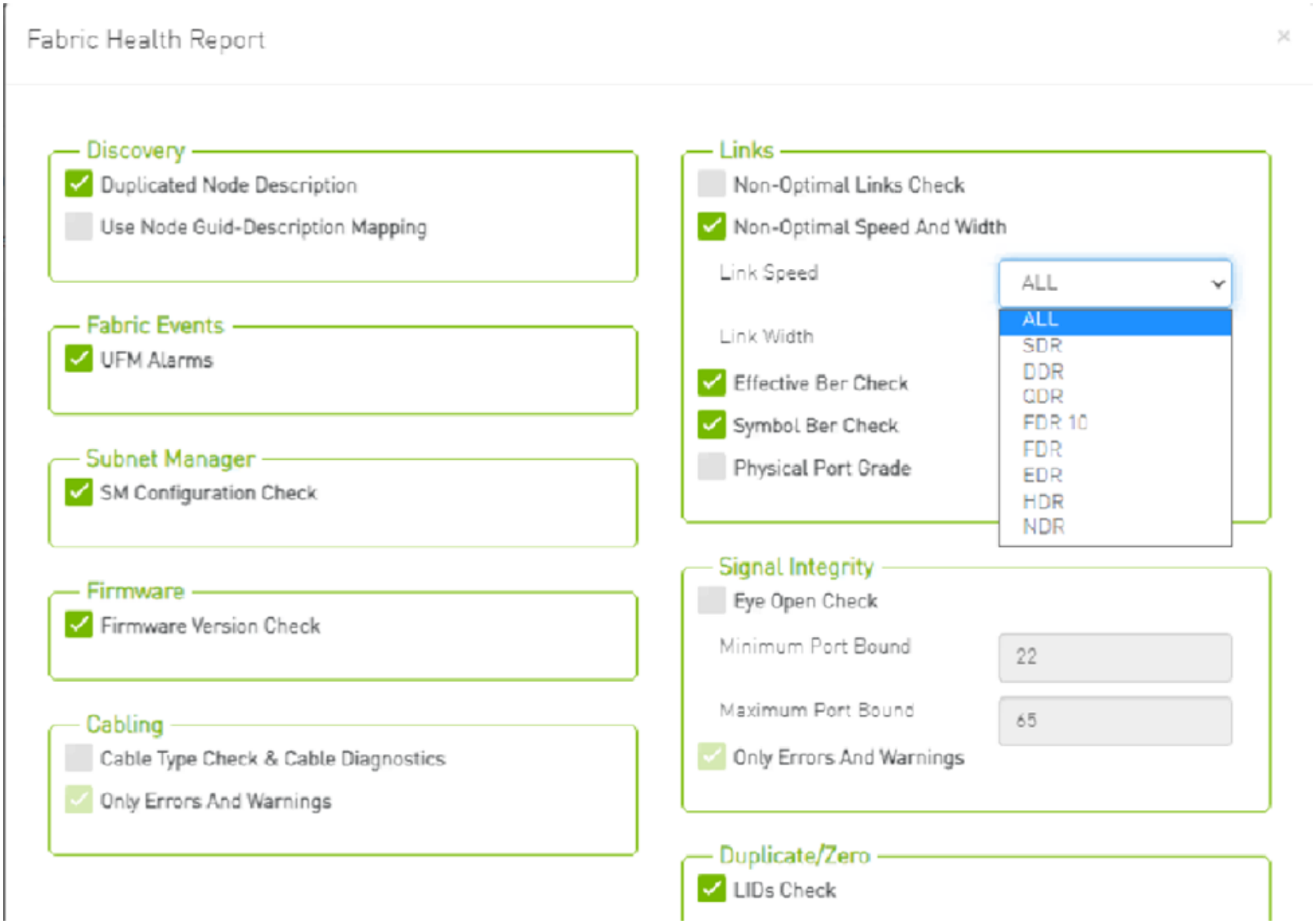
Through **Fabric Health** tab, you can create reports that run a series of checks on the fabric.

Each check that is run for a report triggers a corresponding event. Events are also triggered when a report starts and ends. For more information, see Events & Alarms.

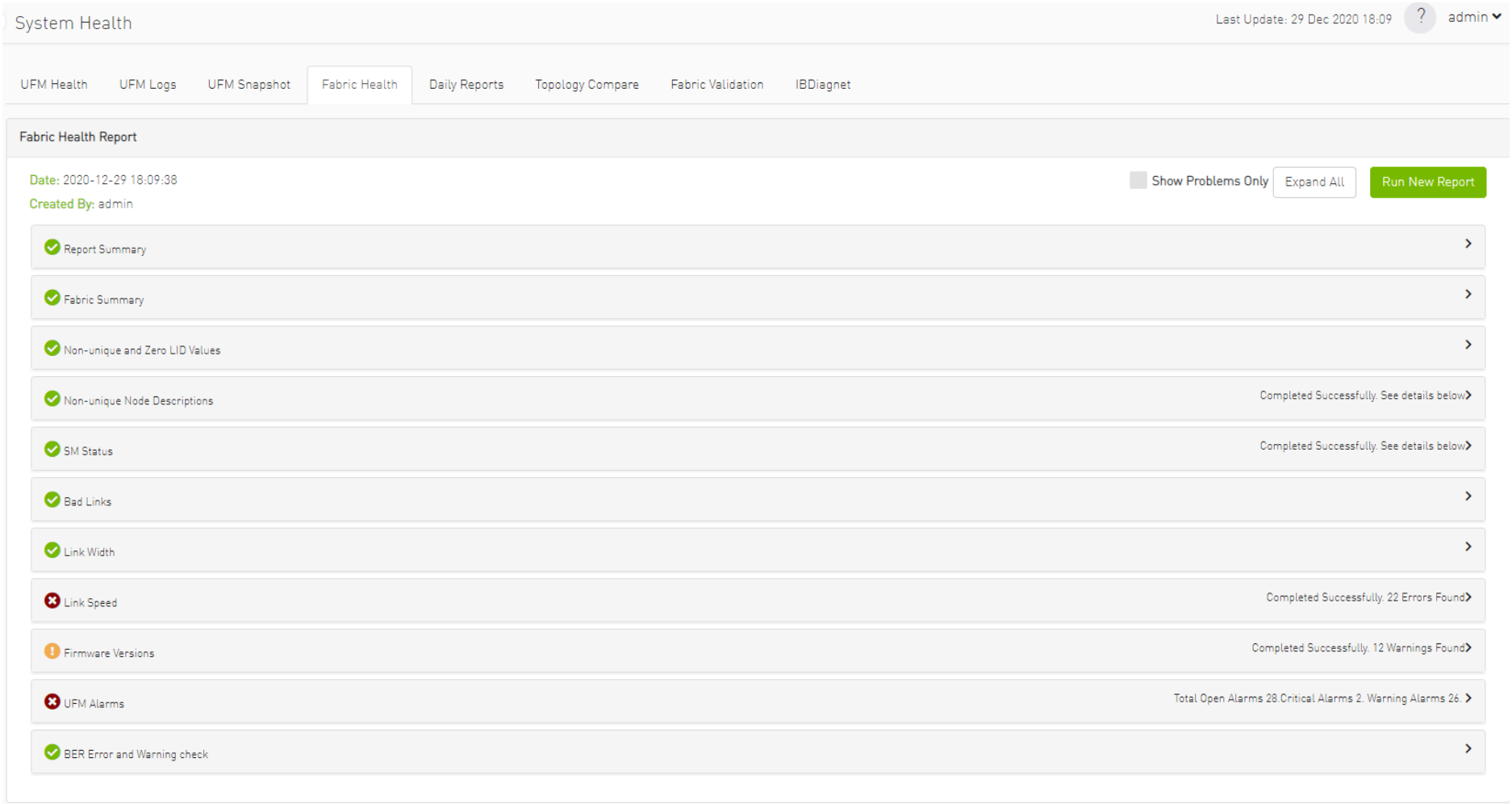
To run a new report, do the following:

1. Click “Run New Report."

2. Select the desired fabric health checks to run in the Fabric Health Report window and click “Run Report."


Fabric Health Report
Discovery
Duplicated Node Description
Use Node Guid-Description Mapping
Fabric Events
UFM Alarms
Subnet Manager
SM Configuration Check
Firmware
Firmware Version Check
Cabling
Cable Type Check & Cable Diagnostics
Only Errors And Warnings
Links
Non-Optimal Links Check
Non-Optimal Speed And Width
Link Speed
ALL
Link Width
ALL
SDR
DDR
QDR
FDR 10
FDR
EDR
HDR
NDR
Effective Ber Check
Symbol Ber Check
Physical Port Grade
Signal Integrity
Eye Open Check
Minimum Port Bound
22
Maximum Port Bound
65
Only Errors And Warnings
Duplicate/Zero
LIDs Check

Results will be displayed automatically:

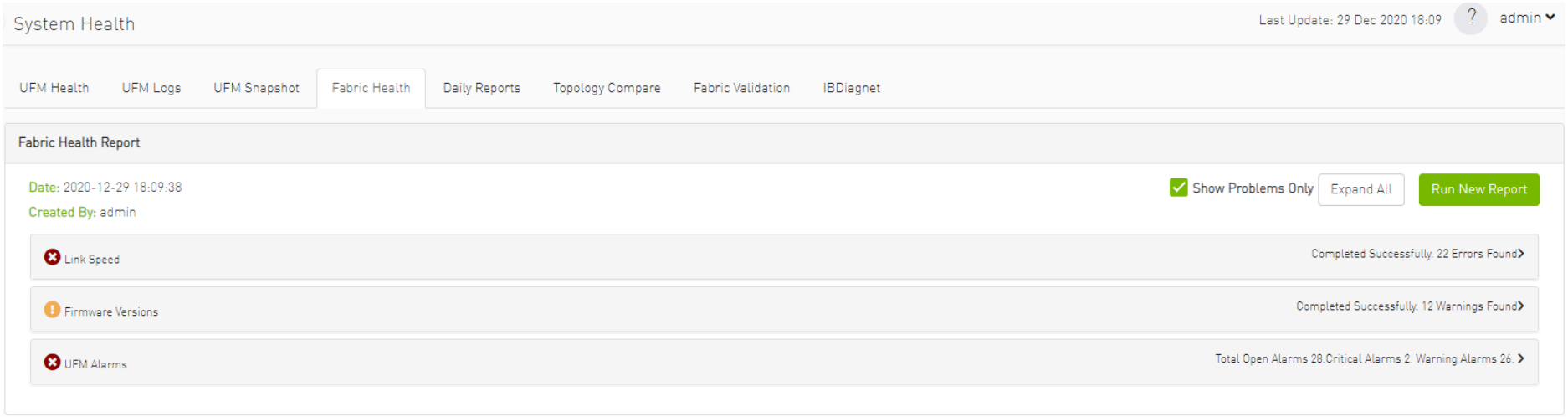


The report displays, the following:

- A report summary table of the errors and warnings generated by the report.
- A fabric summary of the devices and ports in the fabric.
- Details of the results of each check run by the report.

You can expand the view of each check or expand the view of all checks at once by clicking “Expand All."

To view only the errors of the report results, click the "Show Problems Only" checkbox.



The following table describes the checks included in the report.

Fabric Health Report Checks

Check	Description	To run, select:
Duplicate/Zero LID Check	Lists all ports with same LID or zero LID value.	LIDs Check Default: Selected

Check	Description	To run, select:
Duplicated Node Description	Lists all nodes with same node description. Does not include switches with the same description.	Duplicated Node Description Default: Selected
Use Node GUID-Description Mapping	Enables the usage of a mapping file (between node GUID and node description) when running duplicate node description analysis of the fabric. This file is located on the UFM server side at: *conf/sm_guid_desc_mapping.cfg*, and uses the following format (node_guid → description): *0x248a070300702710 "Desc1"* *0x248a0703007026f0 "Desc2"* *0x0002c90300494100 "Desc3"*	**Use Node GUID-Description Mapping** Default: Unchecked **Note**: In order for this checkbox to be available, the Duplicated Node Description checkbox should also be selected. Otherwise, this checkbox will be greyed-out.
SM Check	Checks that: • There is one and only one active (master) Subnet Manager in the fabric. • The master is selected according to highest priority and lowest port GUID. The report lists all SMs in the fabric with their attributes.	SM Configuration Check Default: Selected
Bad Links Check	Performs a full-fabric discovery and reports “non-responsive” ports with their path.	Non-Optimal Links Check Default: Selected
Link Width	Checks if link width is optimally used. • When a width is selected, the report lists the active links that do not meet the optimum for the selection. • When no width is selected (All), the test checks whether the enabled width on both sides of the link equals the configured maximum (confirms that auto-negotiation was successful).	None-Optimal Speed and Width Default: Selected Link Width: The default is ALL.

Check	Description	To run, select:
Link Speed	Checks if link speed is optimally used. • When a speed is selected, the report lists the active links that do not meet the optimum for the selection. • When no speed is selected (All), the test checks whether the enabled speed on both sides of the link equals the configured maximum (confirms that auto-negotiation was successful).	None-Optimal Speed and Width Default: Selected Link Speed: The default is ALL.
Effective Ber Check	Provides a BER test for each port, calculates BER for each port and check no BER value has exceeded the BER thresholds. In the results, this section will display all ports that has exceeded the BER thresholds. Note that there are two levels of threshold: Warning threshold (default=1e-13) and Error threshold (default=1e-8).	Effective Ber Check Default: Selected
Effective Port Grade	Provides a grade per port lane in the fabric, which indicates the current port lane quality.	Physical Port Grade Default: Not Selected
Firmware Check	Checks for firmware inconsistencies. For each device model in the fabric, the test finds the latest installed version of the firmware and reports devices with older versions.	Firmware Version Check Default: Selected
Eye Open Check	(For QDR only) Lists Eye-Opener information for each link. When minimum and maximum port bounds are specified, the report lists the links with eye size outside of the specified bounds.	Eye Open Check Default: Selected Minimum and Maximum port bound: By default no bounds are defined.
Cable Information	Reports cable information as stored in EPROM on each port: cable vendor, type, length and serial number.	Cable Type Check & Cable Diagnostics Default: NOT selected because this test might take a long time to complete (40 msec per port)
UFM Alarms	Lists all open alarms in UFM.	UFM Alarms Default: Selected

Daily Reports Tab

The Daily Report feature collects, analyzes, and reports the most significant issues of the fabric in the last 24 hours (from 00:00 to 24:00). The reports present statistical information such as Summary of Traffic, Congestions and UFM events that occurred during the last 24 hours. These statistics are sent to a pre-defined recipients list on a daily basis. It is also possible to specify a non-24-hour range, by updating the UFM configuration file—see section Other Daily Report Configurations for details.

The following are the formats of the Daily Report:

- Interactive—opened via the browser. The charts are displayed in SVG format. This format can be accessed from the UFM Web UI and is also sent by email as an attachment (see Daily Report View in the Web UI section below).
- Static—opened via mail client (Outlook, Gmail, Hotmail, etc). The charts are displayed in PNG format.

Activating and Deactivating the Daily Report

Daily Report can be activated/deactivated via the `conf/gv.cfg` file.

> ⚠ Daily Reports mechanism is activated by default.

To deactivate the Daily Report, do the following:

1. Open the `conf/gv.cfg` file.
2. Find the DailyReport section.
3. Set the daily_report_enabled option to false.

```
daily_report_enabled = false
```

To re-activate the Daily Report:, do the following:

1. Open the `conf/gv.cfg` file.

2. Find the DailyReport section.

3. Set the daily_report_enabled option to true.

```
daily_report_enabled = true
```

Saving Daily Reports

By default, the maximum number of reports that will be saved is 365 (one per day).

To configure the maximum number of reports to save, do the following:

1. Open the `conf/gv.cfg` file.

2. Find the DailyReport section.

3. Set the max_reports option to the desired value. A count of 0 (zero) means no copies are retained. (default and max is 365).

4. Restart UFM.

Other Daily Report Configurations

All the Daily Report configuration parameters can be found in the "DailyReport" section in gv.cfg configuration file.

The following are additional Daily Report configurations options:

- top_x option specifies the number of results in the "Top X" charts. Max number can be 20. (Default value is 10). top_x value will be applied to all charts existing in the Daily Report.

- mail_send_interval option specifies the epoch in minutes after midnight that the report can be emailed. By default, if UFM was down during midnight, and was restarted after 1:00, the report of the previous day will be generated and saved, but will not be emailed. This can be changed by editing the mail_send_interval. (default value is 60 minutes, meaning that the report will be send only between 00:00 to 1:00).
- log_level option specifies the Daily Report log verbosity. Default value is INFO (optional values: INFO, WARNING and ERROR).
- attach_fabric_health_report option indicates whether or not to add the fabric health report as attachment to the mail. Default value is true (optional values: true or false).
- fabric_health_report_timeout specifies the max time in seconds, to wait for fabric health report generation. Default value is 900 seconds (15 minutes).
  In case of large fabrics, fabric health report might take longer than the default 15 minutes. User can enlarge the timeout for fabric health report to complete.
- **max_attached_file_size** specifies the maximum file size in Bytes for each email attachment that can be sent. Default value is 2 Megabytes.
  If the size of a certain file has exceeded this value, the file will not be sent as an attachment in the Daily Report mail.

```
[DailyReport]
# top_x specifies the number of results per each top x chart.
# max number can be 20.(default is 10)
top_x=10
# max_reports specifies the number of reports to save.
# A count of 0 (zero) means no copies are retained.(default and max is 365)
max_reports = 365
#time interval in minutes after midnight
#when passed mail will not be sent
mail_send_interval=60
log_level = INFO
daily_report_enabled = true
attach_fabric_health_report = true
fabric_health_report_timeout = 900
# max attached file size in bytes, default is 2M (2097152 Bytes)
max_attached_file_size = 2097152
```

- **max_attached_file_size** specifies the maximum file size in Bytes for each email attachment that can be sent. Default value is 2 Megabytes.
- The **start_hour** and **end_hour** options enable selecting a sub-range of the day, during which, the relevant report data will be collected. Since by default this option is configured to collect data from the last 24 hours, the default start_hour is set to 0 (or 00), and the default end_hour is set to 24.

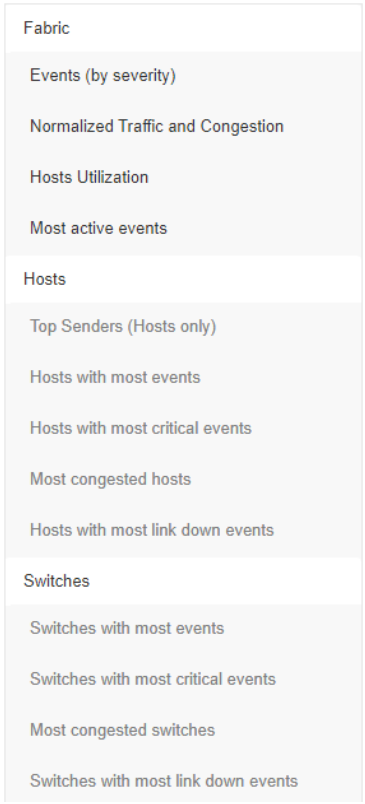
If these options are configured to different values, the generated report will include data from the specified interval only. The start_hour values range is 00 to 23, and the end_hour values range is 00 to 24. The specified end_hour must be greater than the specified start_hour. If, for example, the start_hour is configured to 08, and the end_hour is configured to 10, the generated report will include data collected between 08:00-10:00 (excluding 10:00).

Report Content

Sidebar

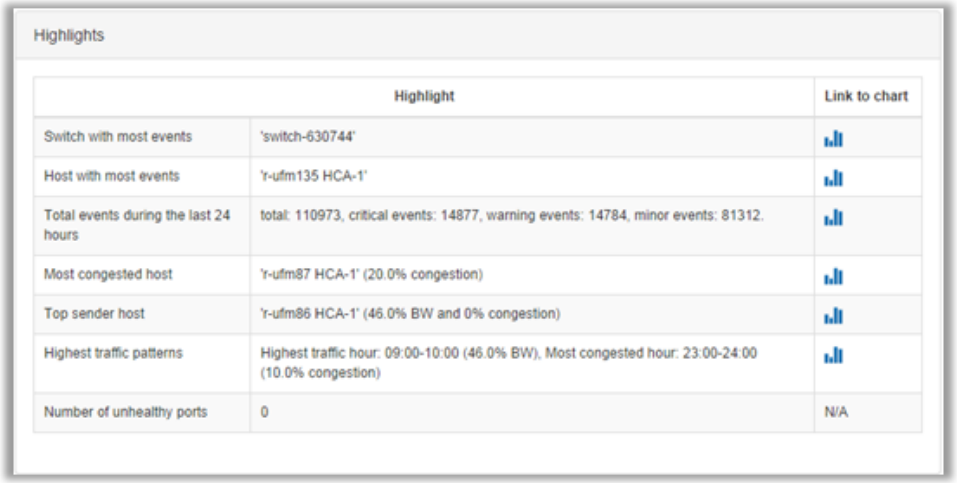
The Sidebar includes general information regarding the fabric, such as: the site name, number of switches and hosts in the fabric, and the dates on which the report was generated.

Navigation between the charts can be done via the menu charts on the sidebar.



Daily Report Highlights

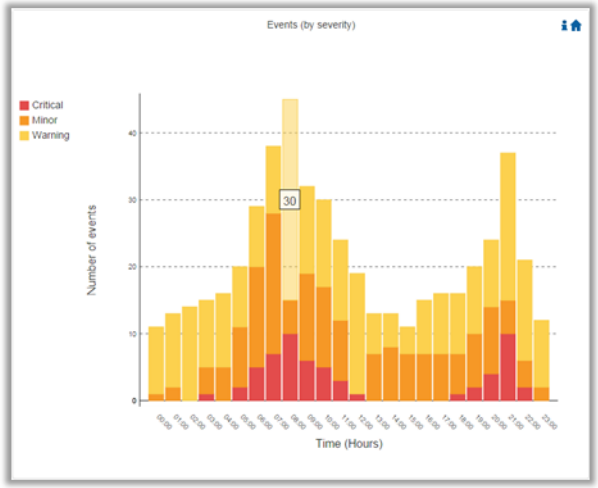
The top of the report shows highlight activities of the network, such as: the host with the most events, the most congested host and switch, and top sender host. To see the related chart of each highlight, click the corresponding icon in the "Link to chart column.



Available Charts

Events by Severity

Events by Severity displays in a graphical view the distribution of all the UFM events that occurred during each hour. Events are separated into the following severity levels: Critical, Minor, and Warning.



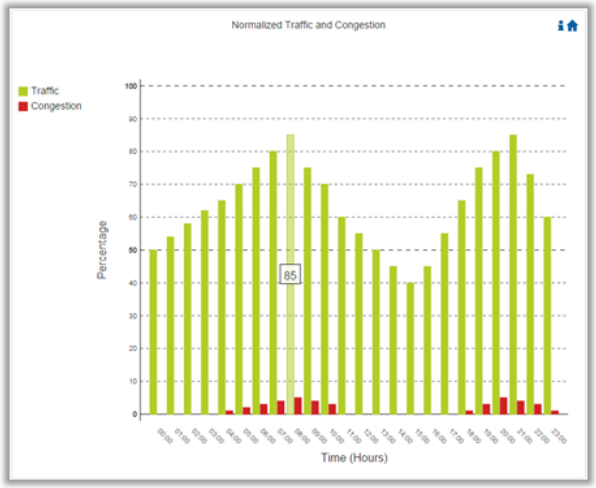
Hovering over the bars in the interactive report displays the amount of events per hour.

Normalized Traffic and Congestion

Normalized Traffic and Congestion displays in a graphical view the normalized traffic and congestions of the fabric. This graph displays the accumulated data for the Senders in the fabric (not including switches).

Congestion normalization is based on the number of delayed packets (packets that wait in the queue) and bandwidth loss.

The graph displays the percentage of the traffic utilization in green and the percentage of the congestion in red.



⚠ Hovering over the bars in the interactive report displays the percentage of the traffic/congestion per hour.

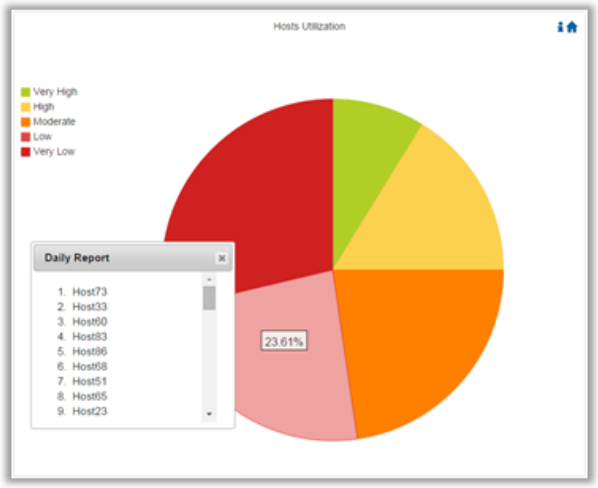
Hosts Utilization Distribution

Hosts Utilization Distribution displays in a graphical view the groups of hosts, where each host belongs to a specific group according to its utilization status.

To see the hosts in each group, click on the pie chart (at the interactive report).

The utilization groups are:

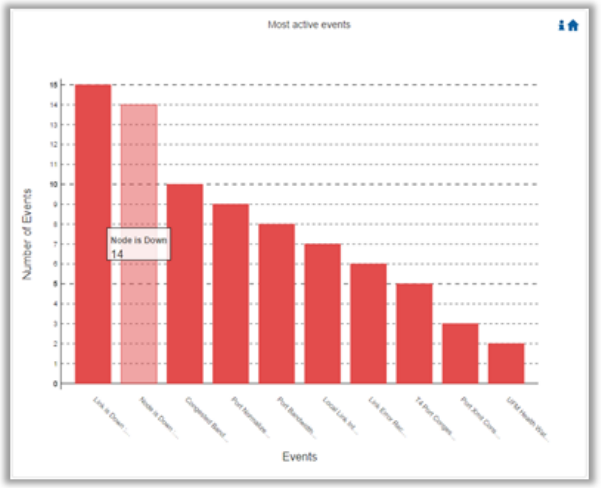
- Very low—up to 20% utilized
- Low—20-40% utilized
- Moderate—40-60% utilized
- High—60-80% utilized
- Very high—80-100% utilized



⚠ Hovering over the slices in the interactive report displays the percentage of hosts in this group.

Most Active Events

Most Active Events displays in a graphical view the most active events, ordered by the number of occurrences during the last 24 hours.

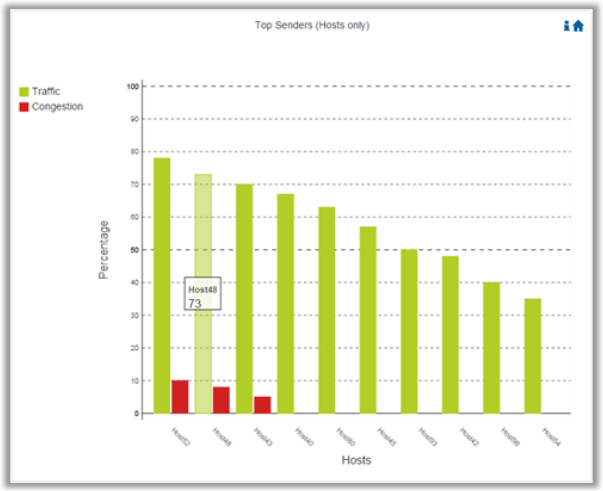


Hovering over the bars in the interactive report displays the number of occurrences for each active event, and hovering on each event's name displays a tooltip with the event's description.

Top Senders

Top Senders displays in a graphical view the normalized traffic and congestions of the top sender hosts. Congestion normalization is based on the number of the delayed packets (packets that wait in queue) and bandwidth loss.

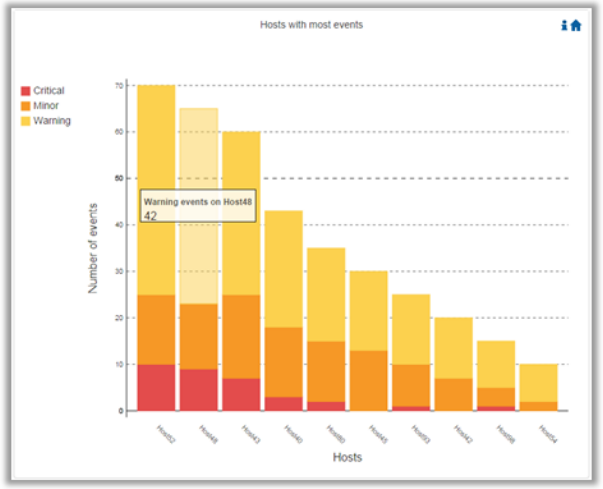
The graph displays the percentage of the traffic utilization in green and the percentage of the congestion in red.



⚠ Hovering over the bars in the interactive report displays the percentage of the traffic/congestion for a selected host.

Hosts with Most Events

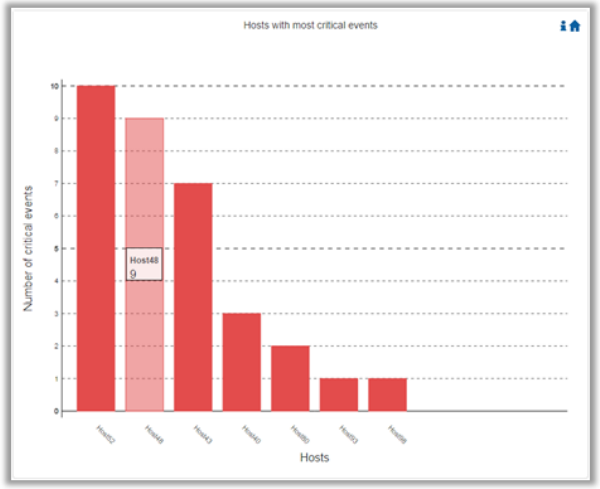
Hosts with Most Events displays in a graphical view the hosts with the most events. Events are separated into the following severity levels: Critical, Minor, and Warning.



⚠ Hovering over the bars in the interactive report displays the amount of events per severity for a selected host.

Hosts with Most Critical Events

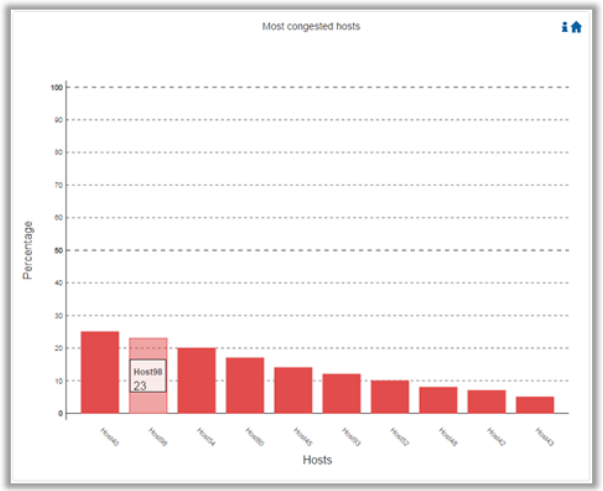
Hosts with Most Critical Events displays in a graphical view the hosts with the most critical events.



⚠ Hovering over the bars in the interactive report displays the amount of critical events for a selected host.

Most Congested Hosts

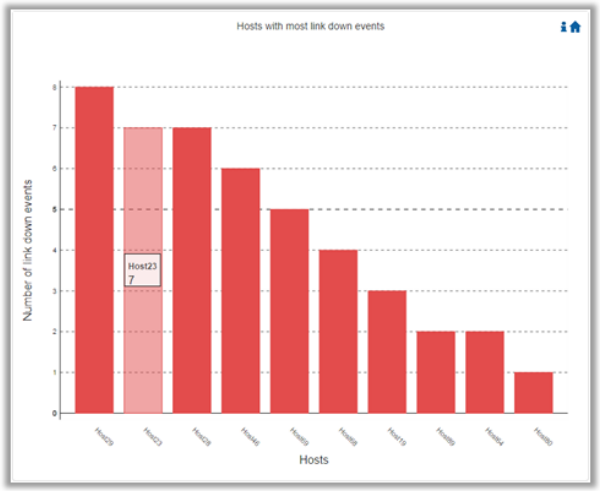
Most Congested Hosts displays in a graphical view the normalized congestions of the most congested hosts. Congestion normalization is based on the number of the delayed packets (packets that wait in queue) and bandwidth loss.



Hovering over the bars in the interactive report displays the percentage of the congestion for a selected host.

Hosts with Most Link Down Events

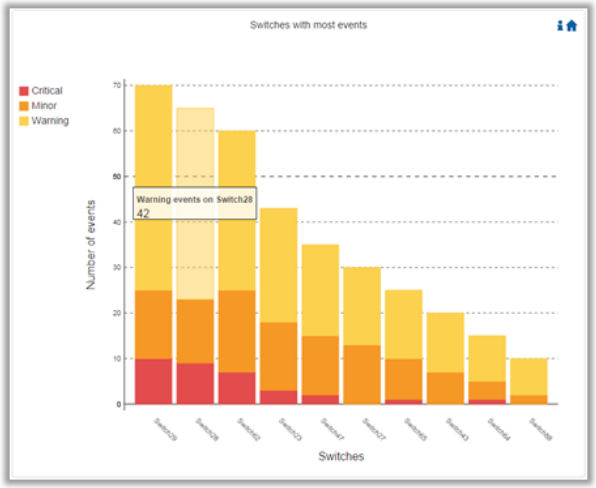
Hosts with Most Link Down Events displays in a graphical view the list of the hosts with the most link down events during the last 24 hours.



Hovering over the bars in the interactive report displays the amount of link-down events for a selected host.

Switches with Most Events

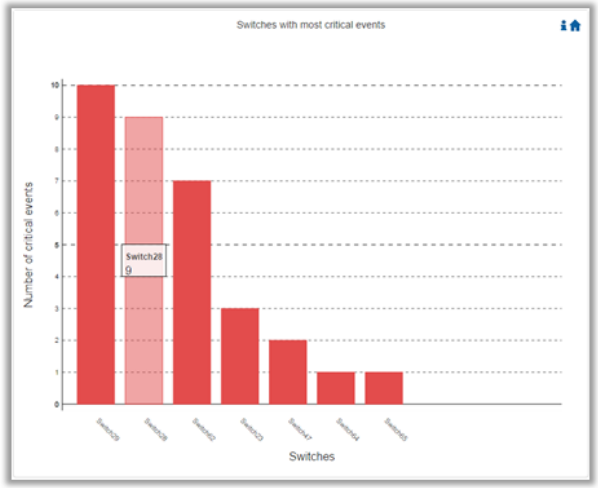
Switches with Most Events displays in a graphical view the switches with the most events. Events are separated into the following severity levels: Critical, Minor, and Warning.



⚠ Hovering over the bars in the interactive report displays the amount of events per severity for a selected switch.

Switches with Most Critical Events

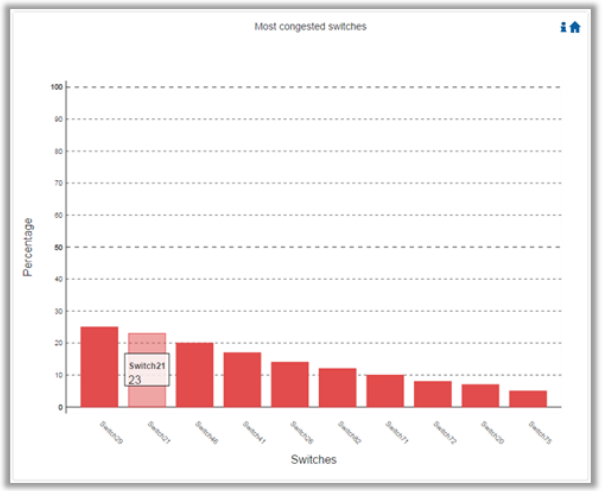
Switches with Most Critical Events displays in a graphical view the switches with the most critical events.



⚠ Hovering over the bars in the interactive report displays the amount of critical events for a selected switch.

Most Congested Switches

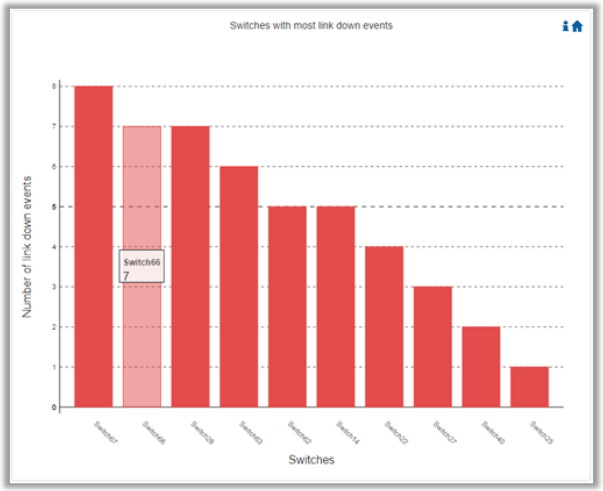
Most Congested Switches displays in a graphical view the normalized congestions of the most congested switches. Congestion normalization is based on the number of delayed packets (packets that wait in queue) and bandwidth loss.



Hovering over the bars in the interactive report displays the percentage of the congestion for a selected switch.

Switches with Most Link Down Events

Switches with Most Link Down Events displays in a graphical view the list of the switches with the most link down events during the last 24 hours.



⚠ Hovering over the bars in the interactive report displays the amount of link-down events for a selected switch.

⚠ Clicking on the “help” icon in the upper right corner of each chart, in the interactive report , will display a short description of the chart.
Clicking on the “home” icon in the upper right corner of each chart, in the interactive report , will move the display to the beginning of the report.

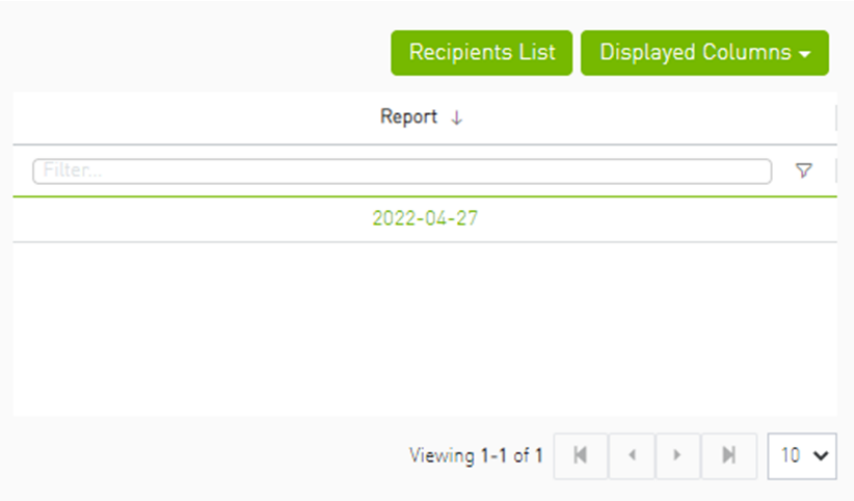
⚠ On charts: “Events by Severity”, “Hosts with Most Events”, and “Switches with Most Events”, if the maximum value in the Y-axis is less than 5, an “m” unit will appear and stand for “milli".

⚠ For all charts, if the value is higher than 1000 in the Y-axis, a “k” unit will appear and stand for “killo”.

Daily Report View in the Web UI

In this tab, you can select the UFM daily reports that you wish to view and you can specify the recipients to which these daily reports will be sent.

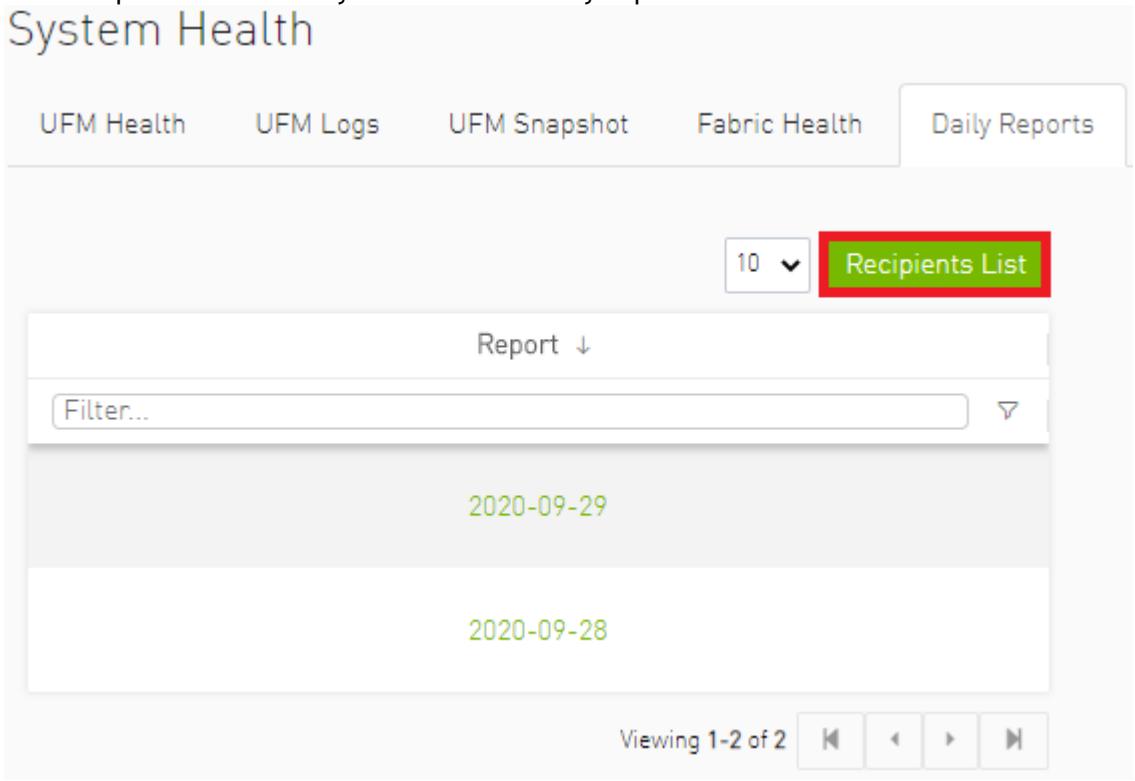
➢ ***To view a specific daily report, click the relevant report date from the list of available daily reports.***



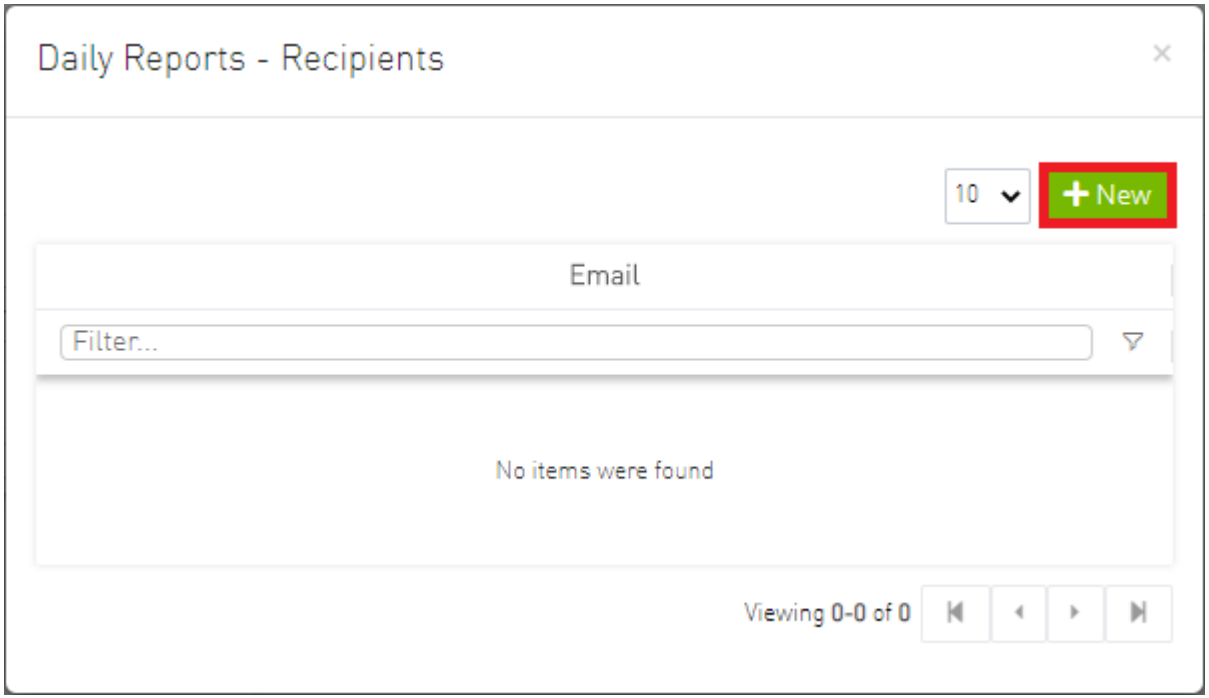
The specified report content will be displayed when clicking the report (see Activating and Deactivating the Daily Report).

➢ ***To configure the Recipients list for the daily reports, do the following:***

1. Click Recipients List under System Health → Daily Reports tab.

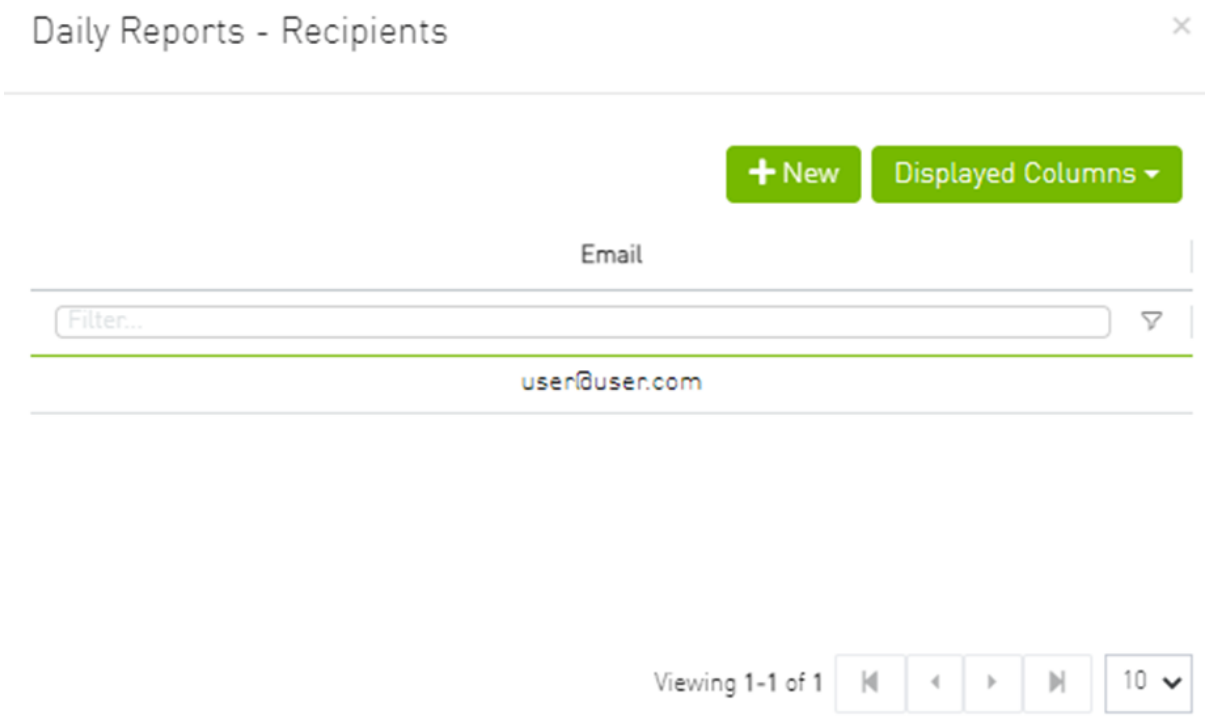


2. Click New.


Daily Reports - Recipients
10
New
Email
Filter...
No items were found
Viewing 0-0 of 0

3. In the Recipients List window, enter valid recipient email addresses, comma-separated, and click Submit.

The new recipient/recipients will be added to the Daily Reports Recipients list.



These recipients will automatically start receiving the UFM daily reports.

Topology Compare Tab

Overview

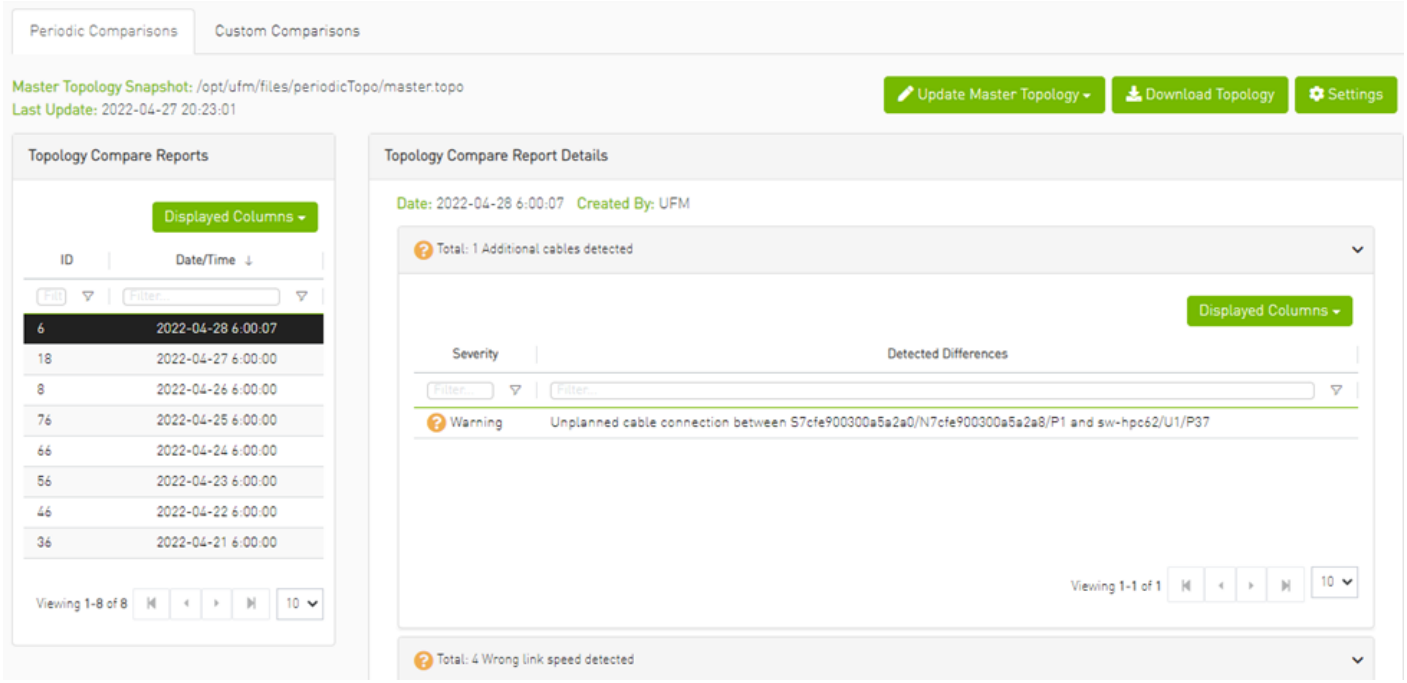
The Topology Compare tab allows two methods of topology comparison:

- Periodic Comparison

- Custom Comparison

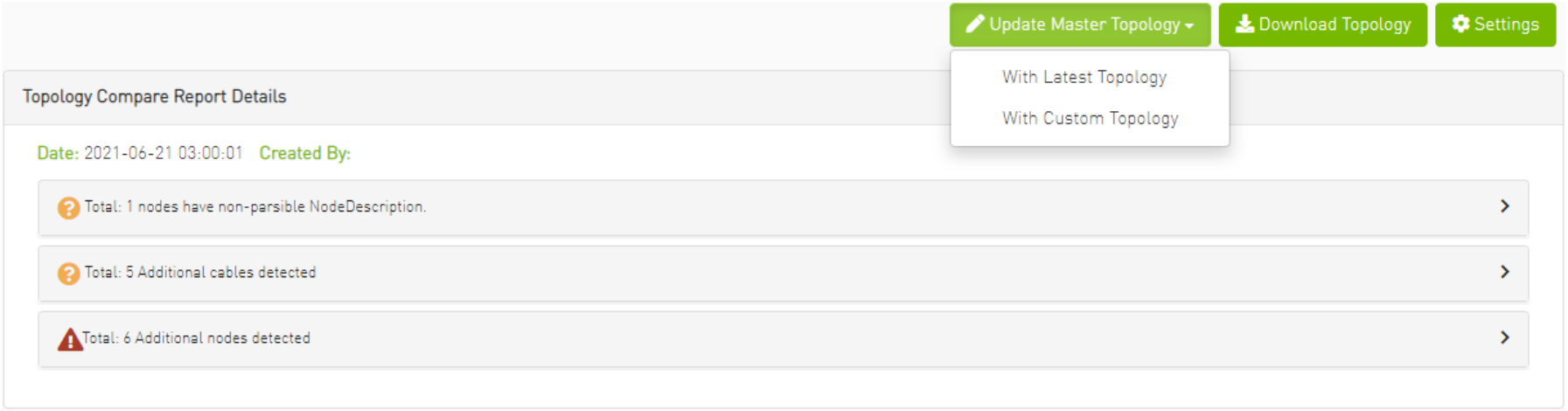
Periodic Comparison

Periodic comparison allows users to compare the current fabric topology with a preset master topology. The master topology may be set either by selecting the current topology or uploading a predefined custom topology.



When a report is selected from the "Topology Compare Reports" table, its result are displayed on the right side under "Topology Compare Report Details".

- To update the master topology with the latest (current) topology or a custom topology saved in external file, click the "Updated Master Topology" dropdown button.



- To download the current topology as a `.topo` file, click the "Download Topology" button.
- The Settings button navigates to the Topology Compare tab of the Settings view which allows users to configure periodic comparison settings.

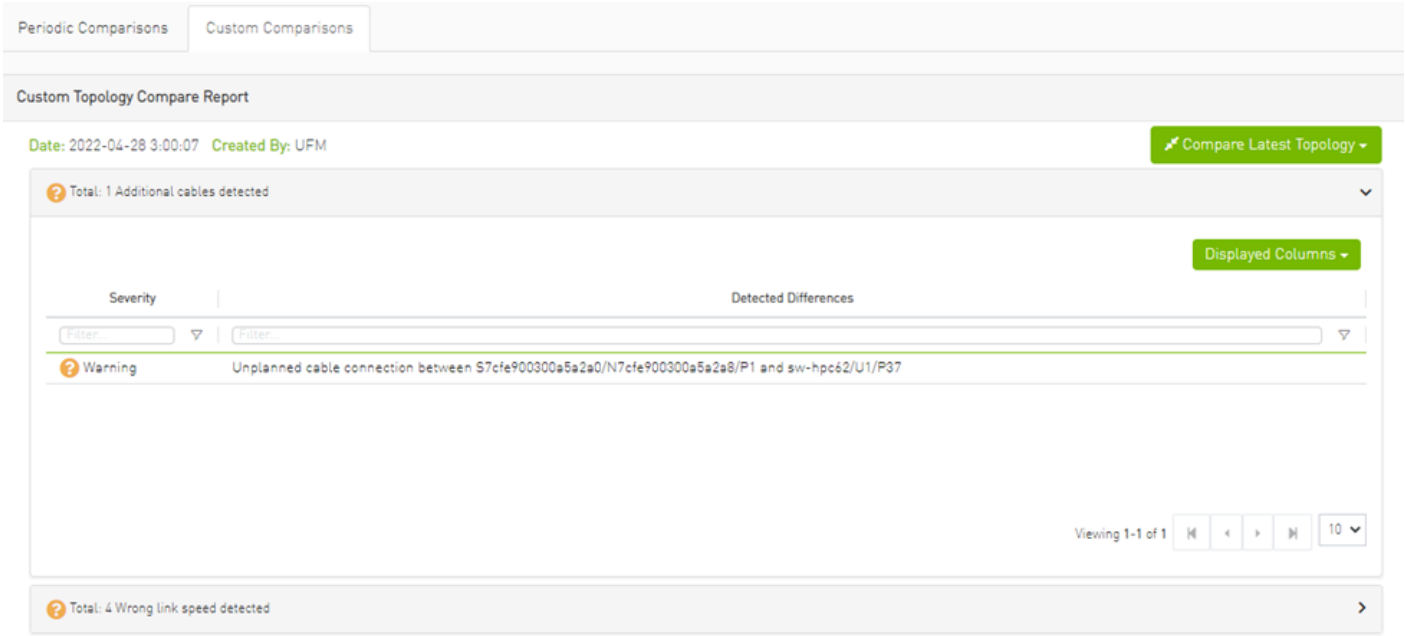
Custom Comparison

Custom comparison compares user-defined topology with the current fabric topology. UFM compares the current fabric topology to a topology snapshot (of the same setup) and reports any differences between them.

To be able to use the UFM topology comparison mechanism, first you need to create a TOPO file that defines the current topology of the fabric.

> Ideally, the topology snapshot (`.topo` file) should be taken after the setup bring-up phase has been completed so that no more topology changes are expected to take place.

Once the TOPO file is created, you can use the topology comparison mechanism to compare the current fabric topology to the one in the TOPO file and view their differences (if found).



To compare the current topology with the master topology or a custom topology (external file), make a selection from the "Compare Latest Topology" dropdown button and upload the `.topo` file to compare against.

Topology Comparison Flow

To create the topology file for later comparison with the current topology, do the following:

1. Run the following CLI command on UFM Appliance server machine:

```
ufmapl [ mgmt-ha-active ] (config) # ib topo-file generate
```

Once command execution is completed, the new topology file will be created. Several .ibnl files that were (optionally) created will also be available in the defined output directory. These `.ibnl` files will be used when comparing any topology file to the current fabric topology.

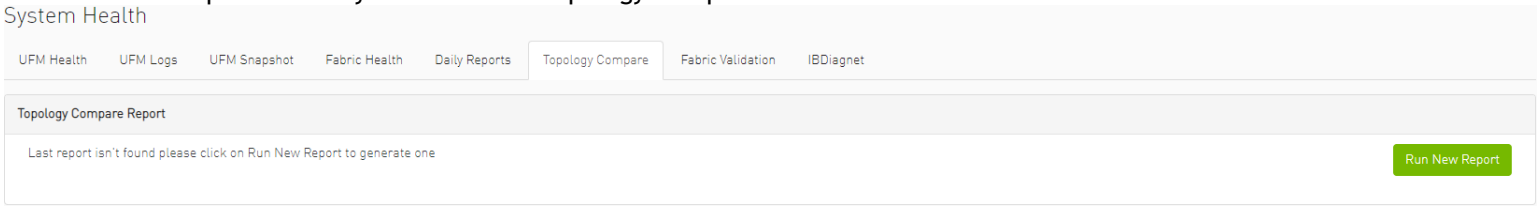
2. Run the following CLI command in order to upload the created topology file (`fabric.topo`) for later comparison with the current fabric topology:

```
ufmapl [ mgmt-ha-active ] (config) # ib topo-file upload scp://root@192.168.1.200/tmp
Password (if required): ******
```

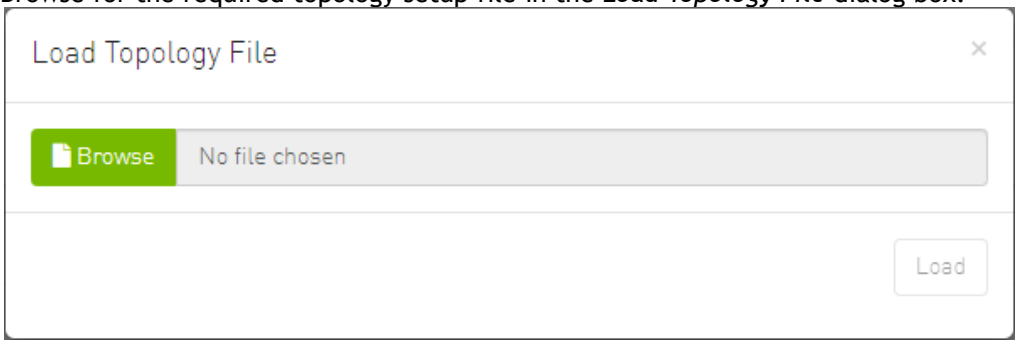
At any time during your UFM session, you can view the last generated report through the UFM Web UI or in HTML format in a browser window.

To perform topology comparison, do the following:

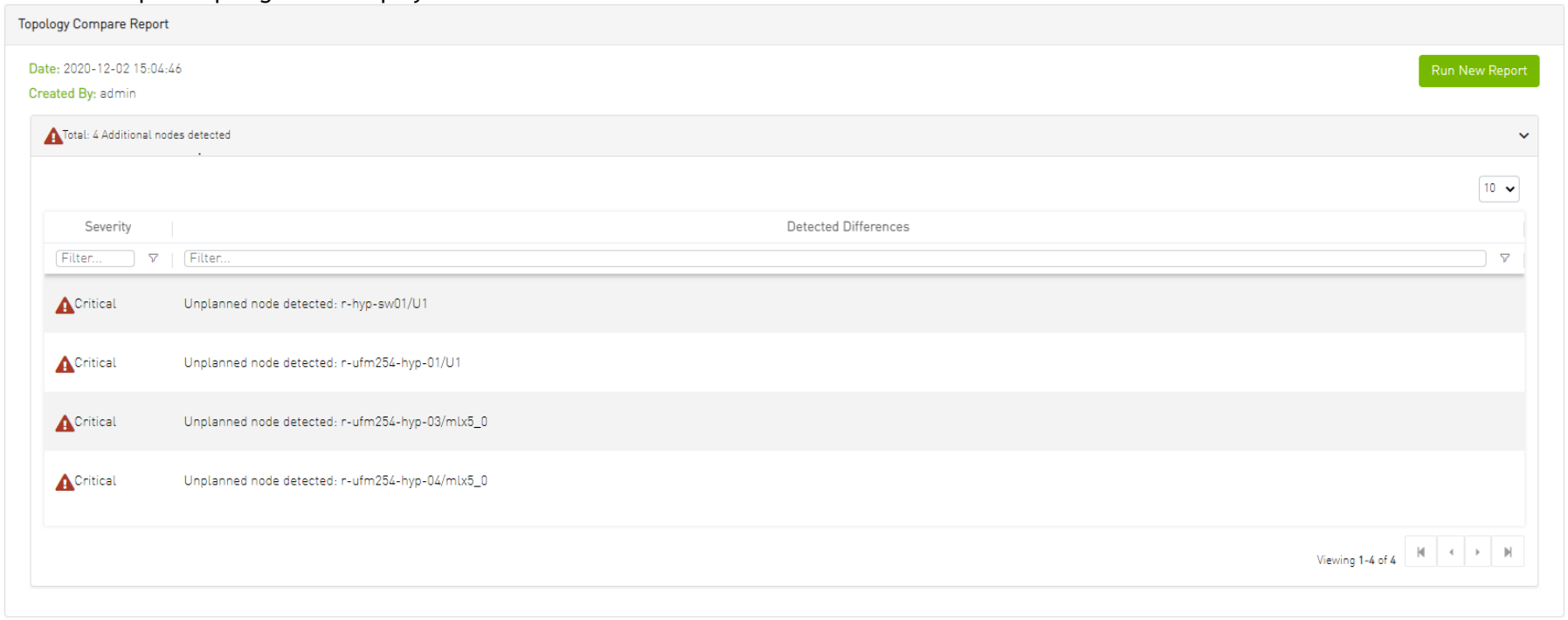
1. Click Run Now Report under System Health à Topology Compare.



2. Browse for the required topology setup file in the *Load Topology File* dialog box.

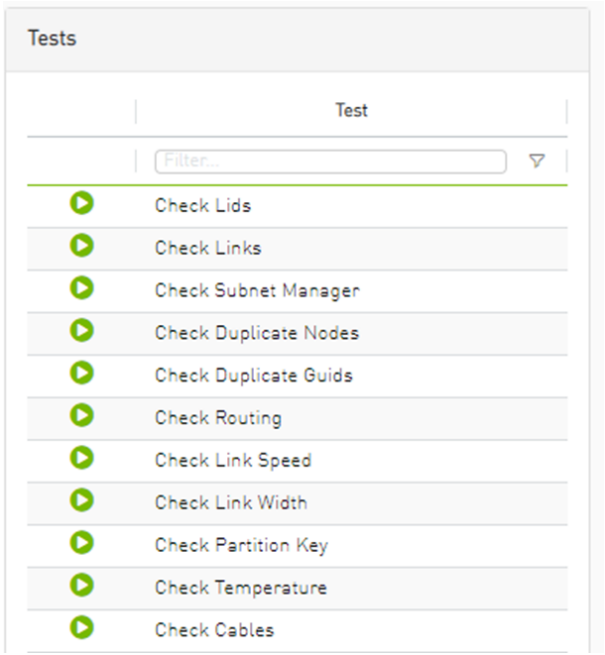


3. Click Load.
   UFM will compare topologies and display the results.



Fabric Validation Tab

The Fabric Validation tab displays the fabric validation tests and gives the ability to run the test and receive/view the summary as a job output. Summary of the job contains all errors and warnings that were found during the test execution.

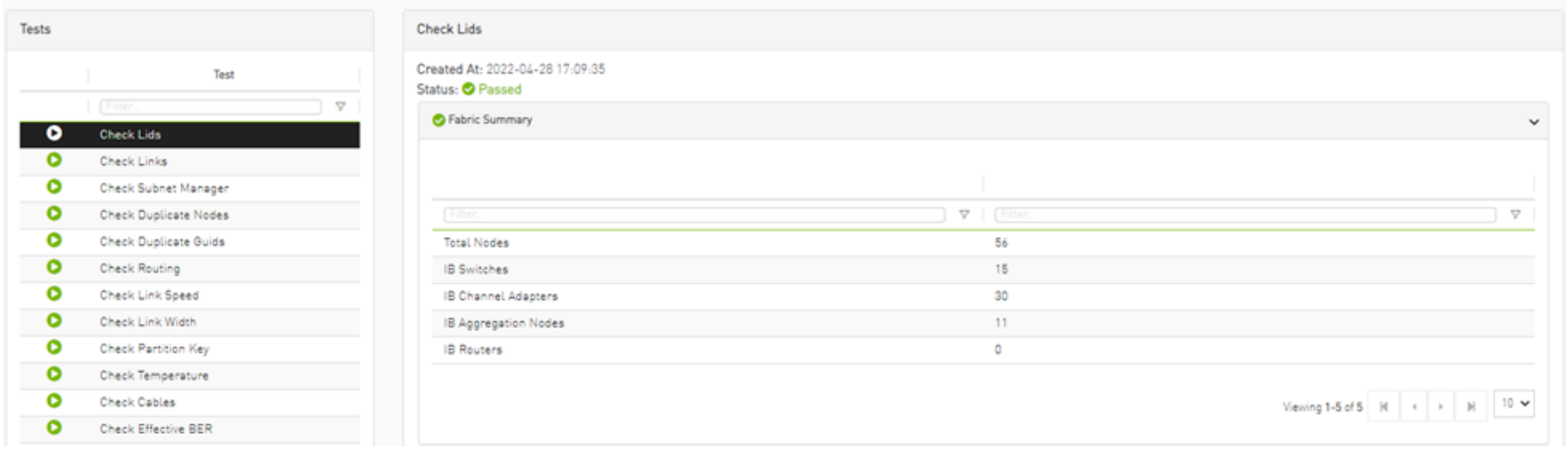


Test	Description
Check Lids	Checks for bad lids. Possible lid errors are: • zero lid • lid duplication
Check Links	Checks for connectivity issues where all ports connected are not in the same state (active)
Check Subnet Manager	Checks for errors related to subnet manager. Possible SM errors are: • Failed to get SMInfo Mad • SM Not Found

Test	Description
	• SM Not Correct (master SM with wrong priority) • Many master SMs exists
Check Duplicate Nodes	Checks for duplications in nodes description
Check Duplicate Guids	Checks for GUIDs duplications
Check Routing	Checks for failures in getting routing MADs
Check Link Speed	Checks for errors related to link speed. Possible link speed errors are: • Different speed between ports • Wrong configuration – 'enable' not part of the 'supported' • Unexpected speed
Check Link Width	Checks for errors related to link width. Possible link width errors are: • Different width between ports • Wrong configuration – 'enable' not part of the 'supported' • Unexpected width
Check Partition Key	Checks for errors related to PKey. Possible PKey errors are: • Failed to get Pkey Tables • Mismatching pkeys between ports
Check Temperature	Checks for failure in getting temperature sensing.
Check Cables	Checks for errors related to cables. Possible cable errors are: • This device does not support cable info capability • Failed to get cable information (provides a reason)
Check Effective BER	Checks that the Effective BER does not exceed the threshold
Dragonfly Topology Validation	Validate if the topology is Dragonfly
SHARP Fabric Validation	Checks for SHARP Configurations in the fabric
Tree Topology Validation	Checks if the fabric is a tree topology

Test	Description
Socket Direct Mode Reporting	Presents the inventory of fabric HCAs that are using socket direct

To run a specific test, click the play button. The job will be displayed once completed.

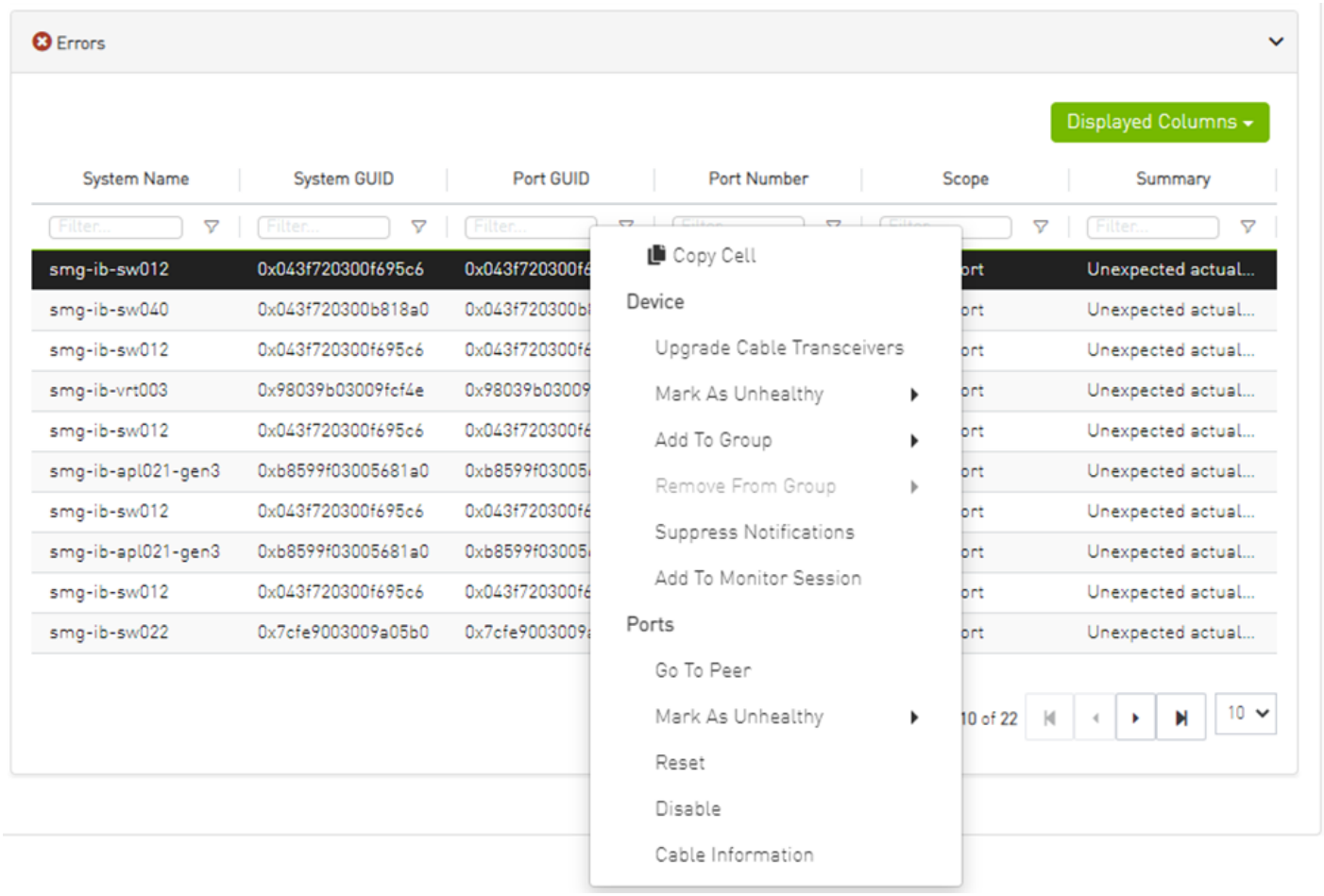


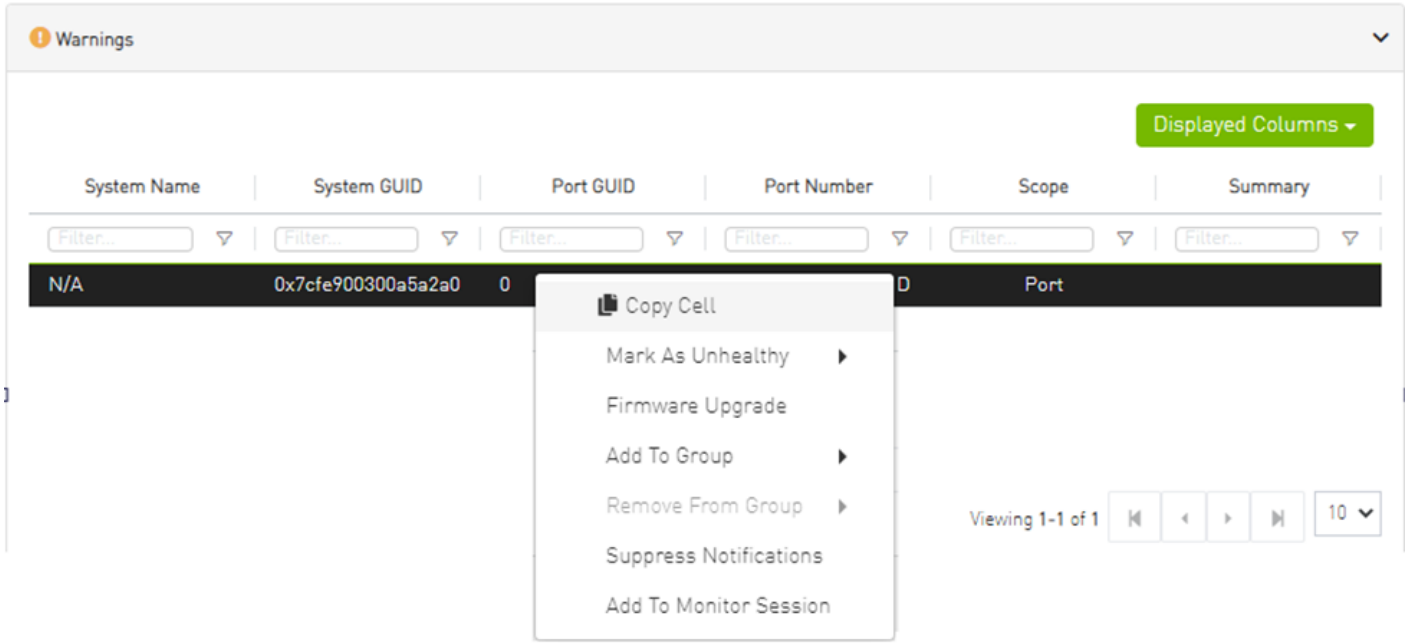
The job will also be displayed in the Jobs window.

Some validation tests contain data related to devices or ports like device GUID and port GUID.

Depending on that information a context menu for each related device/port can be shown.

If the data is related to a port the context menu will contain both port and device options.





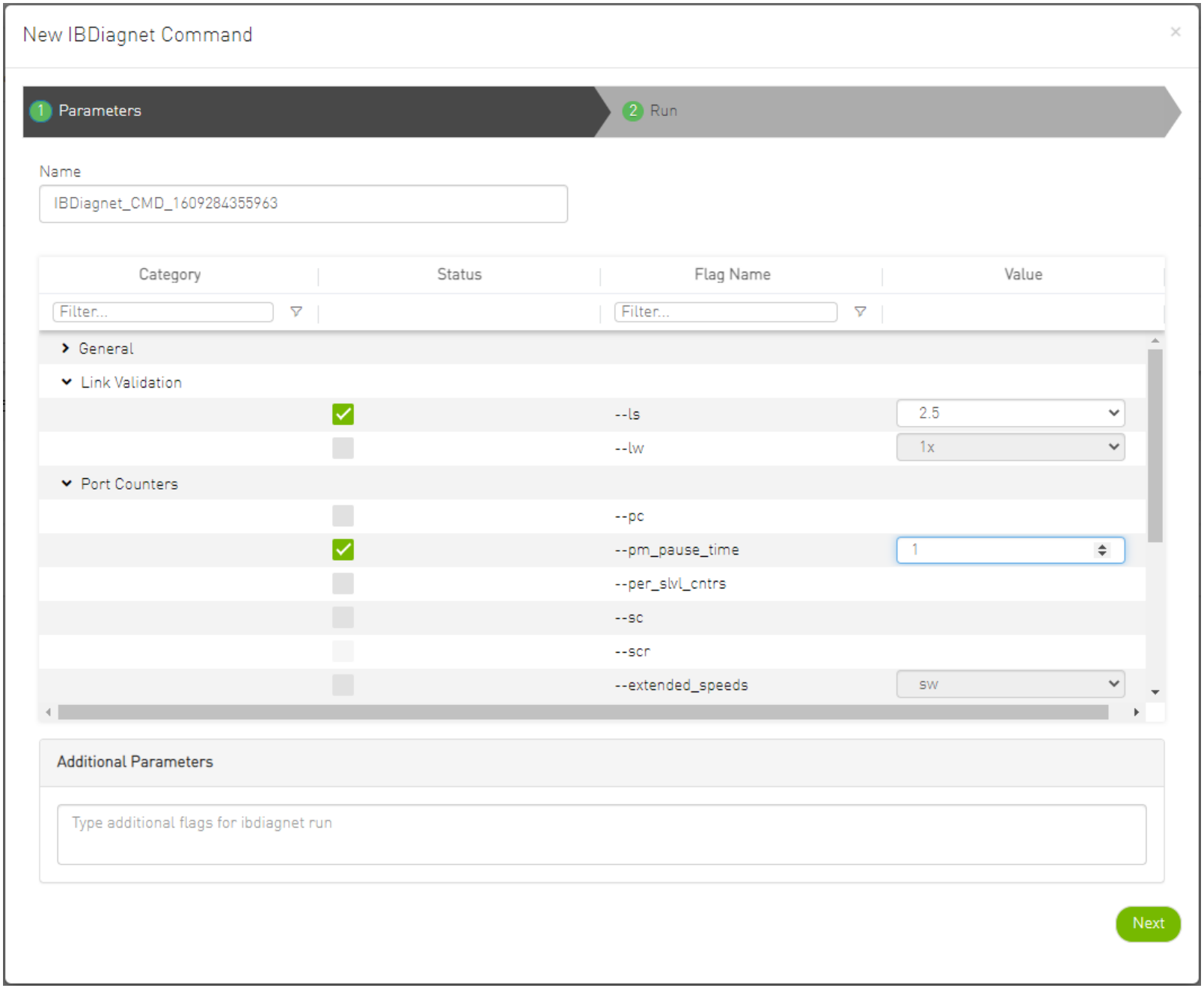
IBDiagnet Tab

The periodic IBDiagnet tab allows users to create scheduled ibdiagnet tasks on their devices using any of the defined parameters.

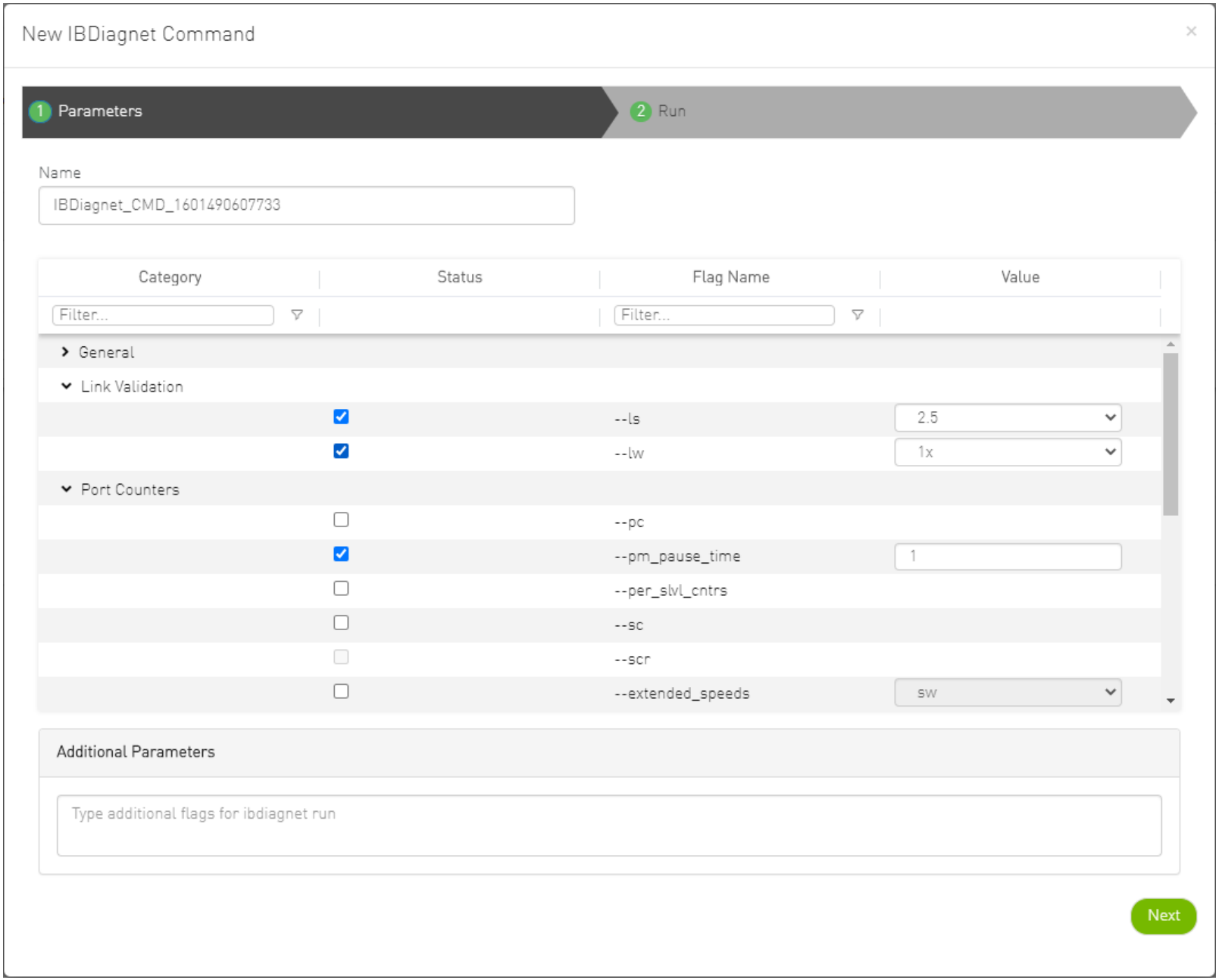
Users can also configure a remote location to save the ibdiagnet output to.

To create a new ibdiagnet command:

1. Click the New button on the top right of the IBDiagnet tab to open the “New IBDiagnet Command” wizard.


New IBDiagnet Command
1 Parameters
2 Run
Name
IBDiagnet_CMD_1609284355963
Category
Status
Flag Name
Value
Filter...
Filter...
General
Link Validation
--ls
2.5
--lw
1x
Port Counters
--pc
--pm_pause_time
1
--per_slvl_cntrs
--sc
--scr
--extended_speeds
sw
Additional Parameters
Type additional flags for ibdiagnet run
Next

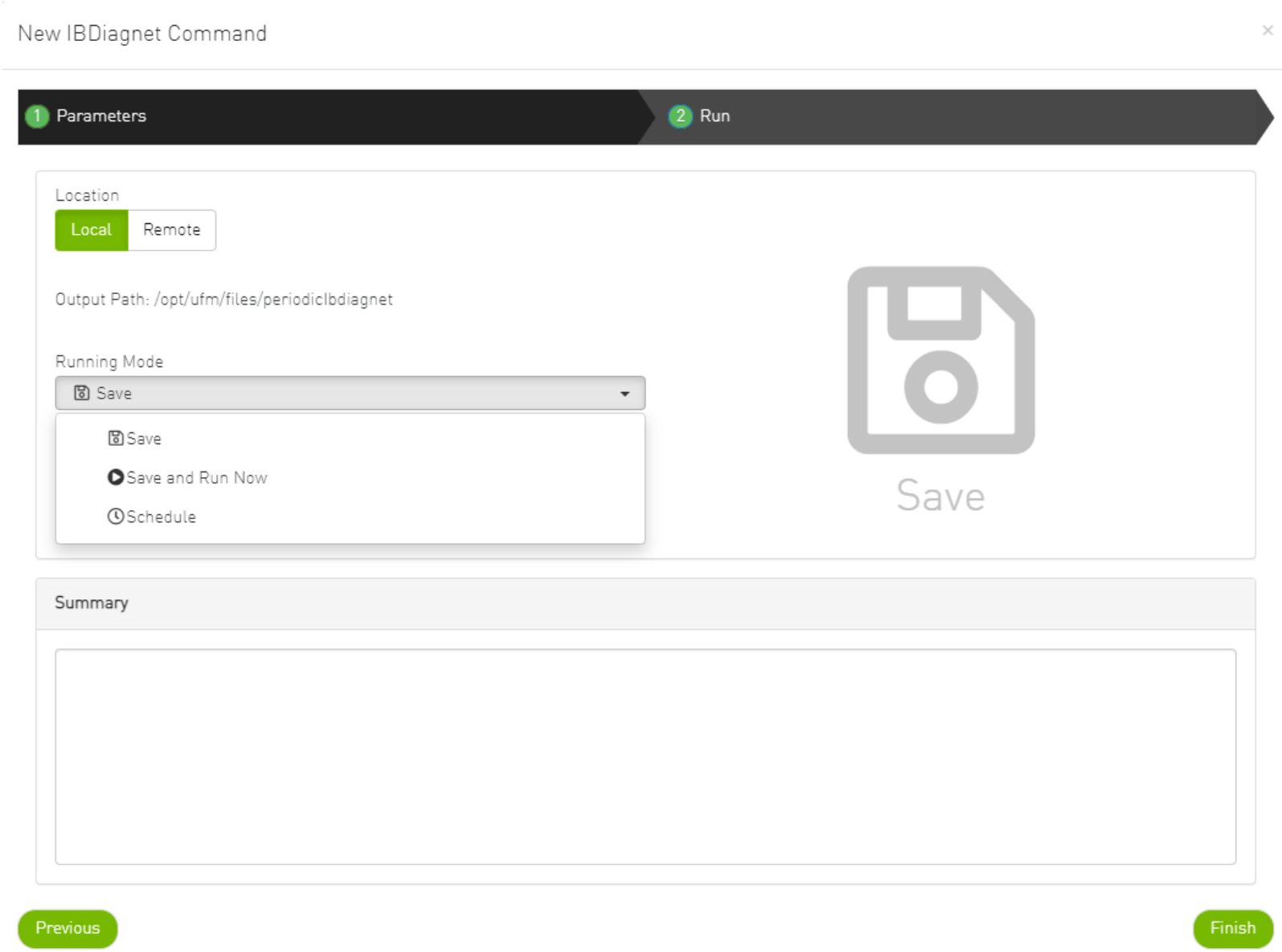
2. Select the desired ibdiagnet flags for your command by selecting the listed flags (categories are expandable), or by manually adding the desired flags into the Additional Parameters box below, and then click Next.


New IBDiagnet Command
1 Parameters
2 Run
Name
IBDiagnet_CMD_1601490607733
Category
Status
Flag Name
Value
Filter...
Filter...
General
Link Validation
--ls
2.5
--lw
1x
Port Counters
--pc
--pm_pause_time
1
--per_slvl_cntrs
--sc
--scr
--extended_speeds
sw
Additional Parameters
Type additional flags for ibdiagnet run
Next

> ⚠ It is possible to use the filters at the top of the Category and Flag Name columns in order to search for flags.

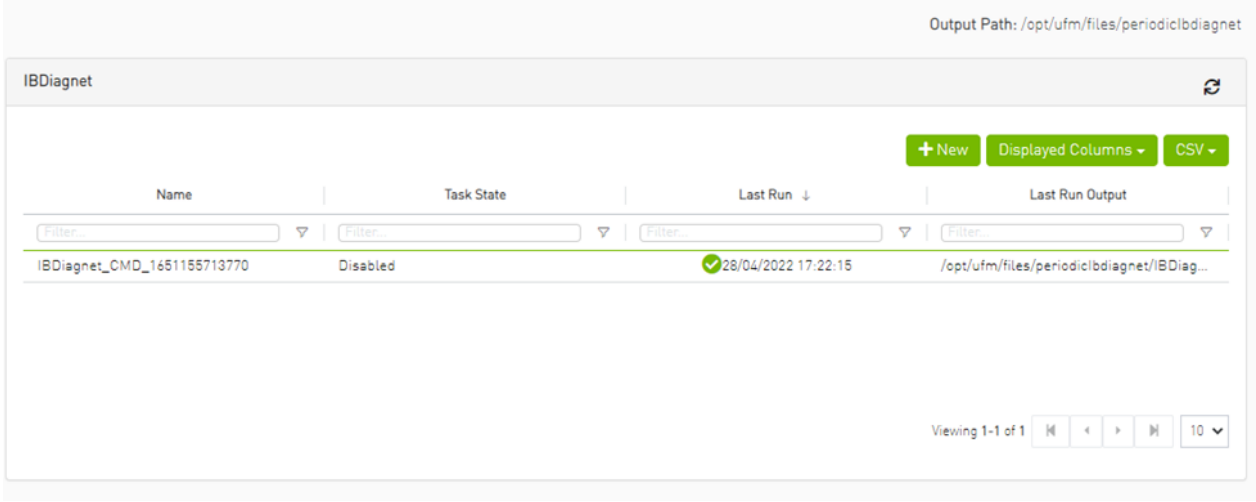
3. In the Run screen:

a. Select whether you would like to save this run for later (Save), run it immediately (Save and Run Now), or schedule it for a later time (Schedule) and then click Finish.


New IBDiagnet Command
1 Parameters
2 Run
Location
Local
Remote
Output Path: /opt/ufm/files/periodicIbdiagnet
Running Mode
Save
Save
Save and Run Now
Schedule
Save
Summary
Previous
Finish

> ⚠ Note that you can see the summary of your chosen flags for this run in the Summary panel.

You will then be able to see run results on the tab which will display where the output is saved on the server.



It is also optional to edit/activate/deactivate/delete a running task using right-click.

Under gv.cfg, it is possible to configure other parameters.

```
[PeriodicIbdiagnet]
# Directory location where outputs are written
periodic_ibdiagnet_dir_location=/opt/ufm/files/periodicIbdiagnet
# Minimum time between two tasks (in minutes)
minimum_task_interval=60
# Maximum number of tasks running simultaneously
max_optional_tasks=5
# Maximum number of outputs to save per task (oldest gets deleted)
max_saved_outputs=5
# Percentage threshold for disk usage from which UFM deletes old task results
disk_usage_threshold=80
```

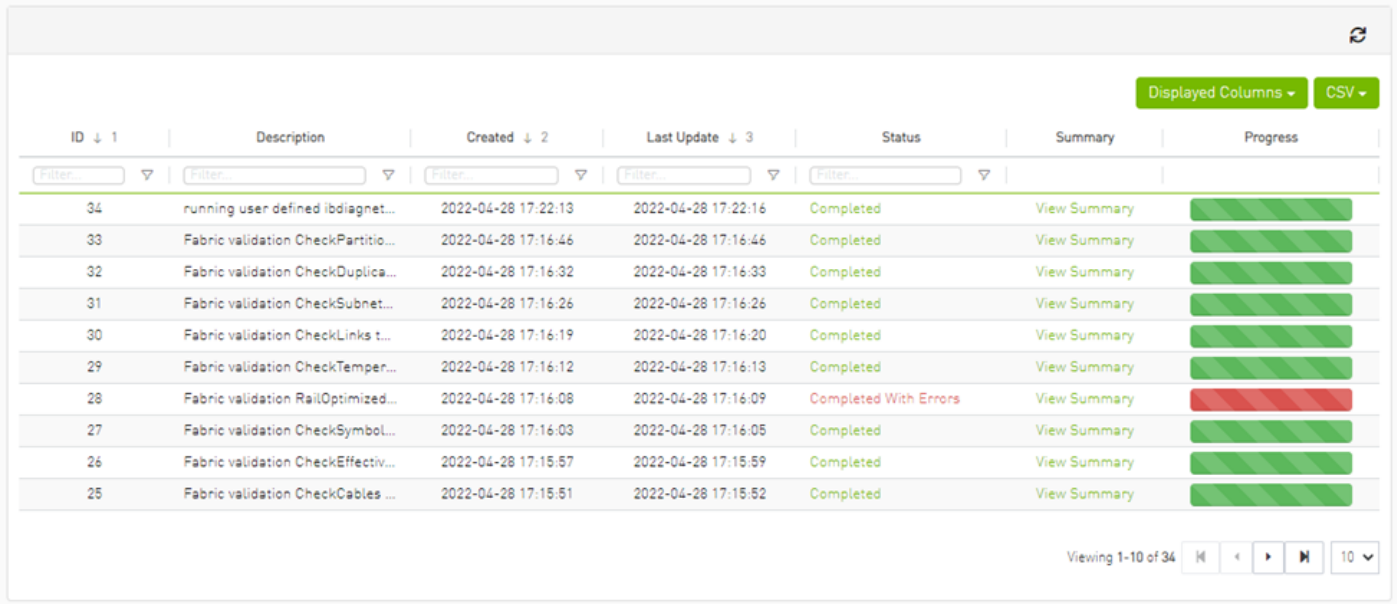
⚠ UFM restart is required for these changes to take effect.

Jobs

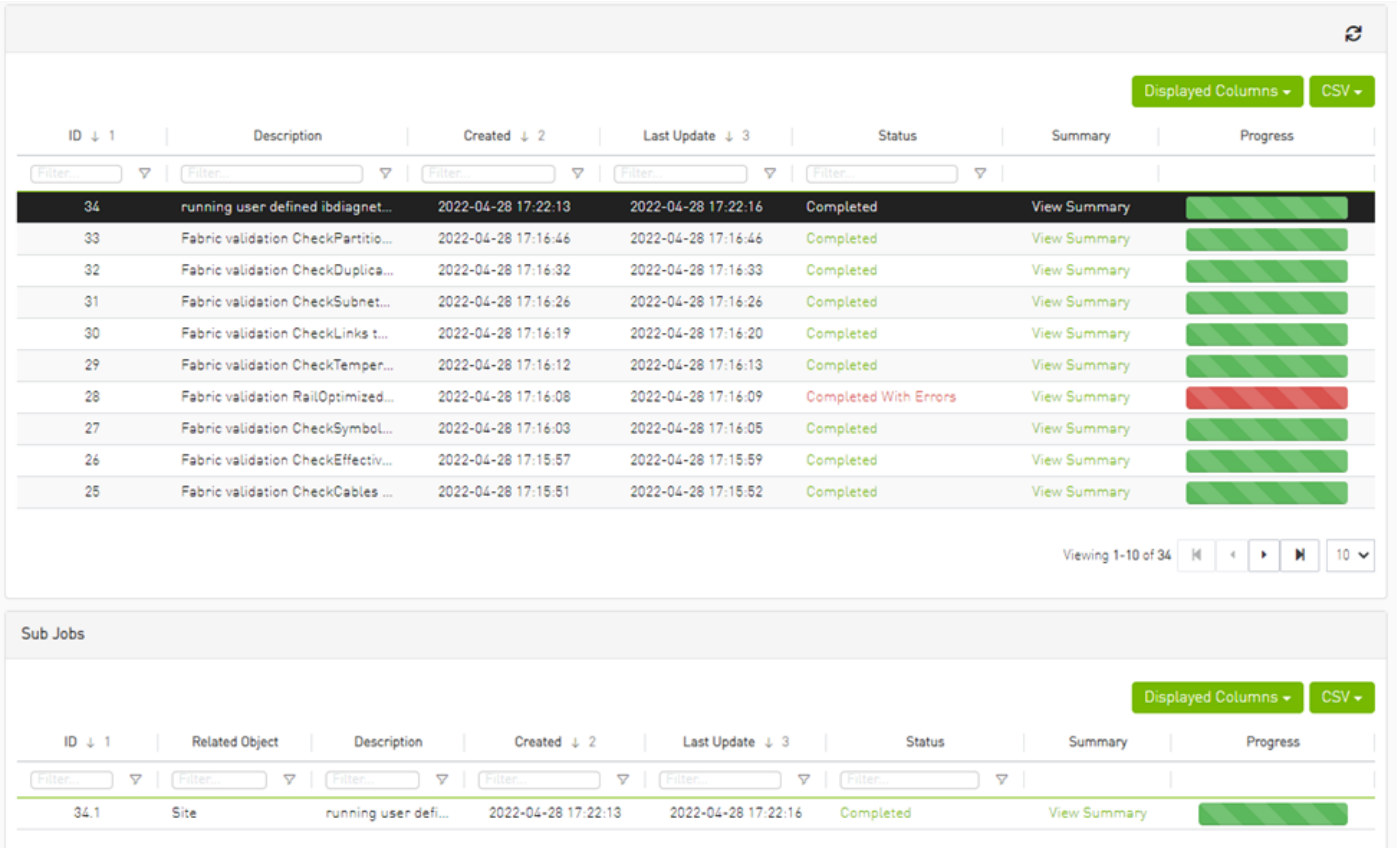
⚠ All information provided in a tabular format in UFM web UI can be exported into a CSV file.

The Jobs window displays all of UFM running Jobs. A Job is a running task defined by the user and applied on one or more of the devices (provisioning, software upgrade, firmware upgrade, reboot, etc.).

UFM users can monitor the progress of a running job, as well as the time it was created, its last update description and its status. The status value can be “Running” (during operation) “Completed with Errors”, in case an error has occurred, and “Completed."



When selecting a job from the main Jobs table, its related sub jobs will be displayed in the Sub Jobs table below.



Settings

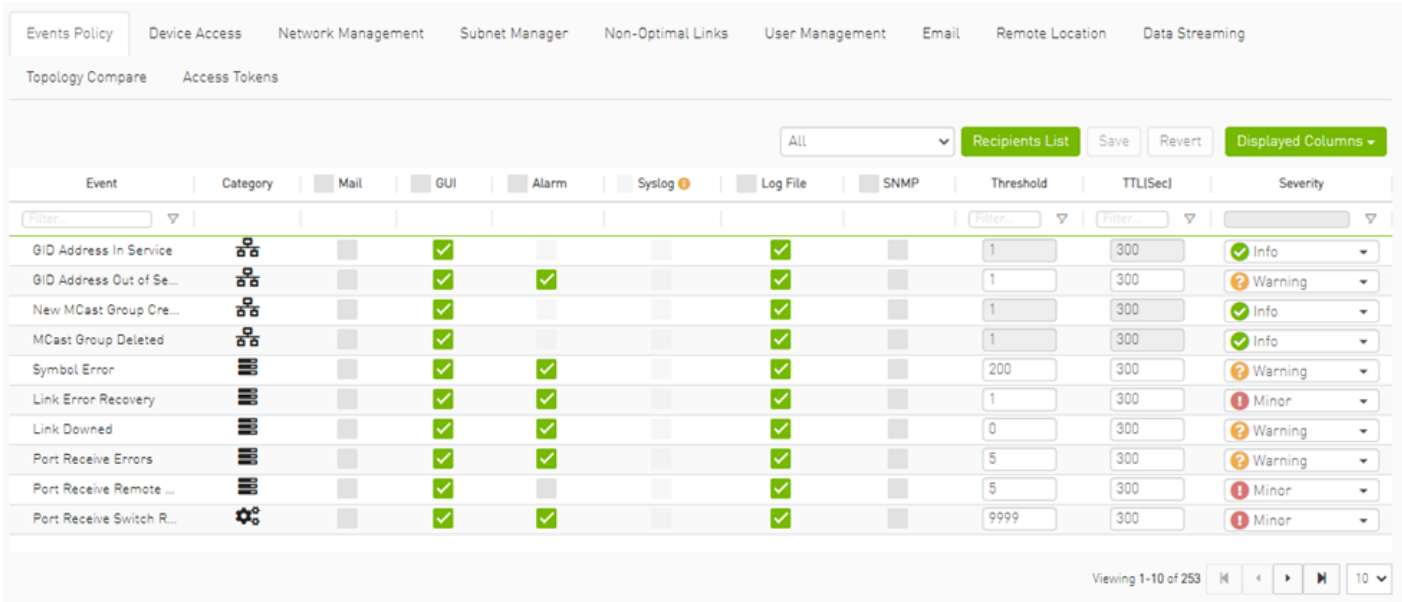
> ⚠ All information provided in a tabular format in UFM web UI can be exported into a CSV file.

This window enables configuring the following UFM server and fabric-related settings:

- Events Policy
- Device Access
- Network Management
- Subnet Manager Tab
- Non-Optimal Links
- Email
- Remote Location
- Data Streaming
- Topology Compare
- Token-based Authentication
- Plugin Management
- User Preferences

Events Policy

The Events Policy tab allows you to define how and when events are triggered for effective troubleshooting and fabric maintenance.



Events are reported by setting the following parameters:

Option	Description/Instructions
Event	Event description.
Category	Event category, such as Communication Error and Hardware represented by icons.
Mail	When selected, the corresponding events will be sent a list of recipients according to Configuring Email-on-Events.
Web UI	When selected, the corresponding events are displayed in the Events & Alarms window in the Web UI.
Alarm	Select the Alarm option to trigger an alarm for a specific event. When selected, the alarms will appear in the Events & Alarms window in the Web UI.

Option	Description/Instructions
Syslog	When checked along with the Log file option, the selected events will be written to Syslog.
Log File	Select the Log File option if you would like the selected event to be reported in a log file.
SNMP	The UFM Server will send events to third-party clients by means of SNMP traps. Select the event SNMP check box option to enable the system to send an SNMP trap for the specific event. The SNMP trap will be sent to the port defined in Configuration file located under: /opt/ufm/conf/gv.cfg. For further information, refer to SNMP Settings.
Threshold	An event will be triggered when the traffic/error rate exceeds the defined threshold. For example: when PortXmit Discards is set to 5 and the counter value grows by 5 units or more between two sequential reads, an event is generated.
TTL (Sec)	TTL (Alarm Time to Live) sets the time during which the alarm on the event is visible on UFM Web UI. TTL is defined in seconds. CAUTION: Setting the TTL to 0 makes the alarm permanent, meaning that the alarm does not disappear from the Web UI until cleared manually.
Action	The action that will be executed in case the event which has triggered the action can be none or isolated (make the port unhealthy or isolated). This attribute can be set only for ports event policy.
Severity	Select the severity level of the event and its alarm from the drop-down list: Info, Warning, Minor, and Critical.

- Category column in the Events Policy table indicates to which category the event belongs. These categories are defined in the event configuration file and cannot be modified. Categories are: Hardware, Fabric Configuration, Communication Error, Fabric Notification, Maintenance, Logical Model, Fabric Topology, Gateway, Module Status, and UFM Server.
- Event logs can still be checked even if the events.log file checkbox was not checked during Syslog configuration.
- For a certain event to be sent to Syslog, both the Syslog and the Log File checkboxes must be checked. Otherwise, the selected events will not be sent to Syslog.

See Appendix - Supported Port Counters and Events for detailed information on port counters and events.

SNMP Settings

When UFM is running, the Web UI Policy Table shows the SNMP traps. You can then modify and save an SNMP Trap flag for each event. SNMP settings are enabled only after the installation of the UFM license.

UFM sends SNMP Trap using version SNMPV2 to the default port 162.

To set the SNMP properties:

1. Open the `conf/gv.cfg` configuration file.
2. Under the [Notifications] line (see the following example):
    a. Set the (snmp_listeners) IP addresses and ports
    b. Port is optional - the default port number is 162
    c. Use a comma to separate multiple listeners

    Format:

    ```
    snmp_listeners = <IP Address 1>[:<port 1>][,<IP Address 2>[:<port 2>]…]
    ```

    Example:

    ```
    [Notifications]
    snmp_listeners = host1, host2:166
    ```

Configuring Email-on-Events

UFM enables you to configure each event to be sent by email to a list of pre-defined recipients. Every 5 minutes (configurable) UFM will collect all “Mail” selected events and send them to the list of pre-defined recipients. By default, the maximum number of events which can be sent in a single email is 100 (configurable, should be in the range of 1-1000)

The order of events in the email body can be set as desired. The available options are: order by severity or order by time (by default: order by severity)

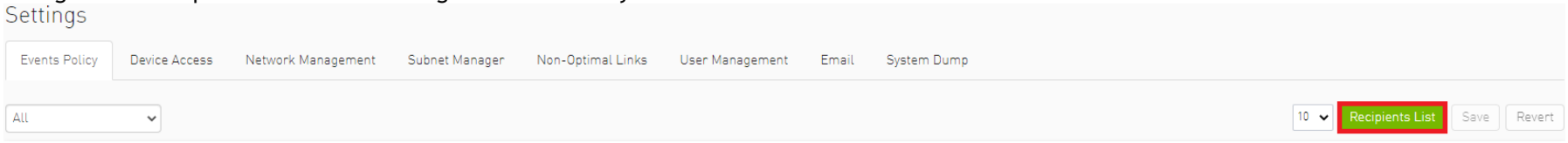
To change email-on-events setting, do the following:

1. Edit the `conf/gv.cfg` file.
2. Go to section "[Events]" and set the relevant parameters:
    - sending_interval (default=5)—Time interval for keeping events (minimum 10 seconds, maximum 24 hours)
    - sending_interval_unit (default = minute)—Optional units: minute, second, hour
    - cyclic_buffer (default=false)—If the cyclic buffer is set to true, older events will be dropped, otherwise newer events will be dropped (if reaches max count)
    - max_events (default=100)—Maximum number of events to be sent in one mail (buffer size), should be in the range of 1-1000
    - group_by_severity (default=true)—Group events in mail by severity or by time

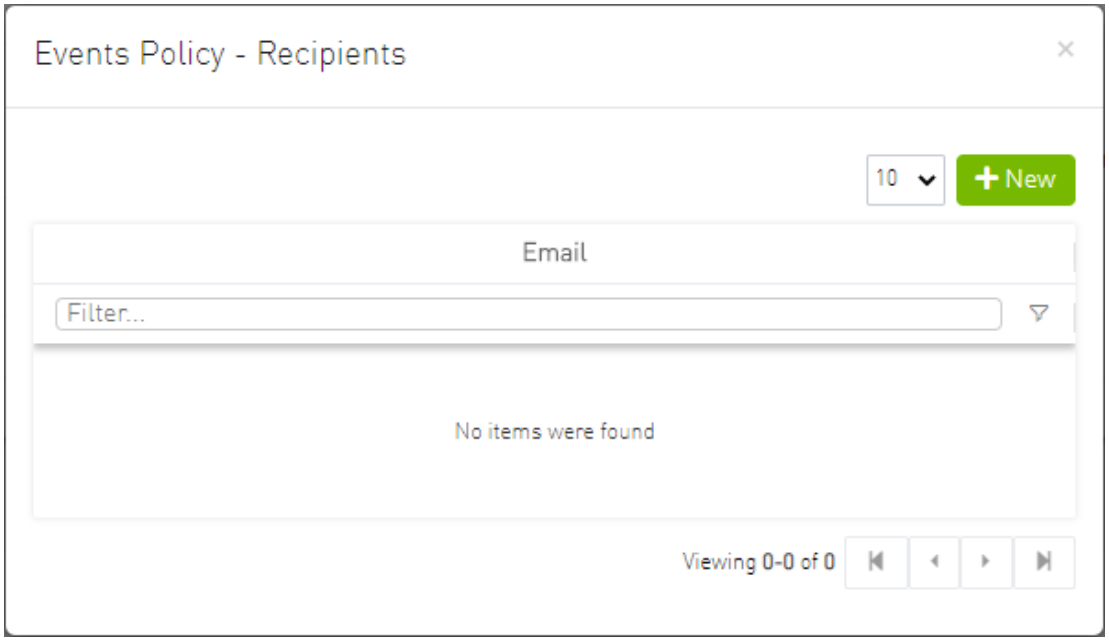
To receive the email-on-events, do the following:

⚠ Configure SMTP settings under Settings window → Email tab - see Email Tab.

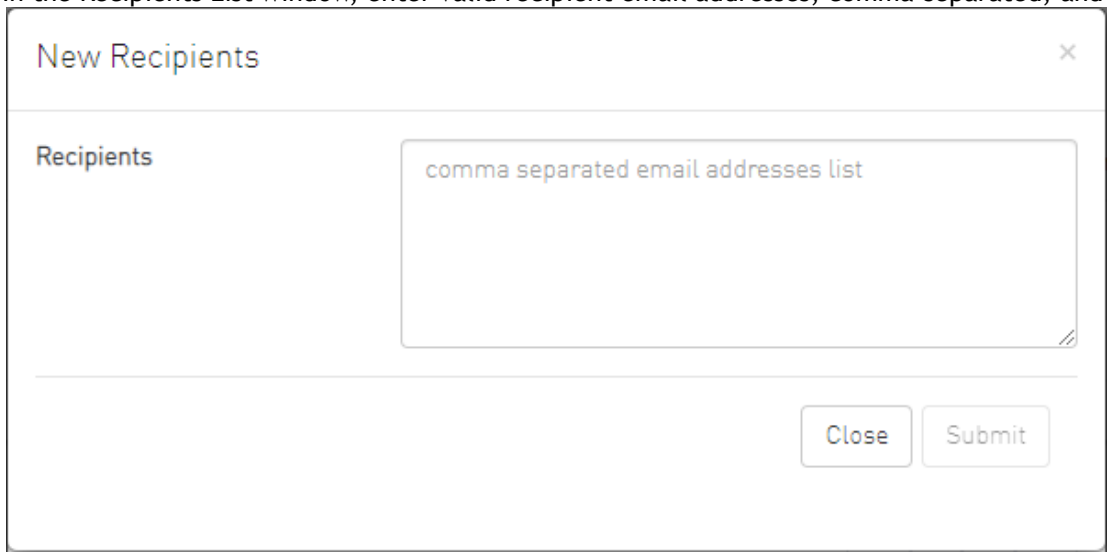
1. Configure the Recipients List under Settings → Events Policy.



2. Click New.

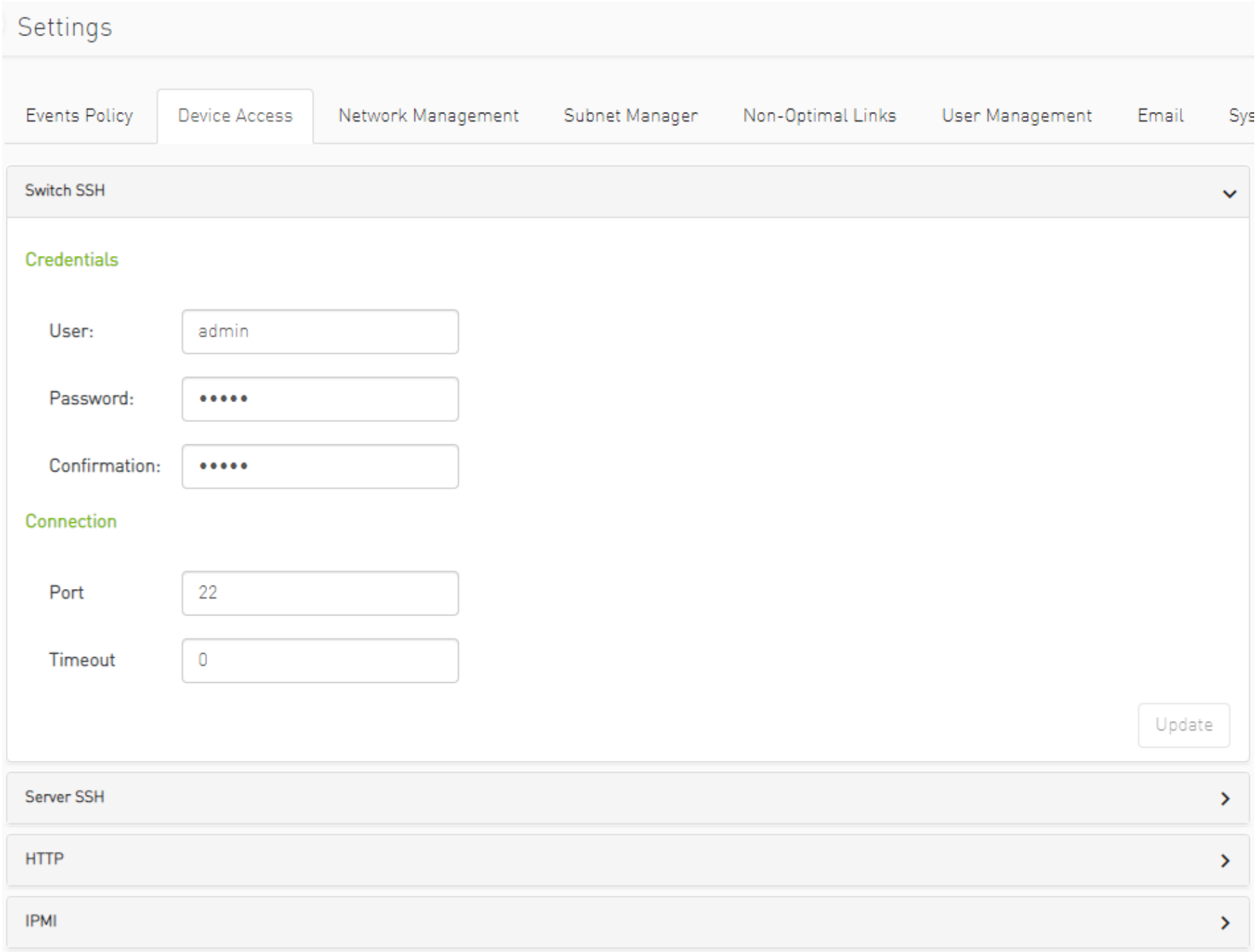

Events Policy - Recipients
10
New
Email
Filter...
No items were found
Viewing 0-0 of 0

3. In the Recipients List window, enter valid recipient email addresses, comma-separated, and click Submit.



The new recipients are then added to the Events Policy Recipients list.
These recipients automatically start receiving emails on the events for which the Mail checkbox is checked in the table under Events Policy.

Device Access


Settings
Events Policy
Device Access
Network Management
Subnet Manager
Non-Optimal Links
User Management
Email
Sys
Switch SSH
Credentials
User:
admin
Password:
•••••
Confirmation:
•••••
Connection
Port
22
Timeout
0
Update
Server SSH
HTTP
IPMI

You can configure default access parameters for remote administration via the following protocols:

- **Switch/Server SSH** - allows you to define the SSH parameters to open an SSH session on your device
- **IPMI** - allows you to set the IPMI parameters to open an IPMI session on your device for remote power control
- **HTTP** - allows you to define the HTTP parameters to open an HTTP session on your device
  Default credentials are applicable to all switches and servers in the fabric.

> ⚠ The default SSH (CLI) switch credentials match the Grid Director series switch. To change the credentials for IS5030/IS5035 edit the [SSH_Switch] section in the gv.cfg file.

Define access parameters for the remote user as described in the following table.

Site Access Credential Parameters

Parameter	Description
User	The name of the user allowed remote access.
Password	Enter the user password.
Confirmation	Re-enter the password.
Port	Each communication protocol has a default port for connection. You can modify the port number, if required.
Timeout	Each communication protocol has a default timeout, i.e. the maximum time, in seconds, to wait for a response from the peer. You can modify the timeout, if required.

Network Management

UFM achieves maximum performance with latency-critical tasks by implementing traffic isolation, which minimizes cross-application interference by prioritizing traffic to ensure critical applications get the optimal service levels.

UFM Routing Protocols

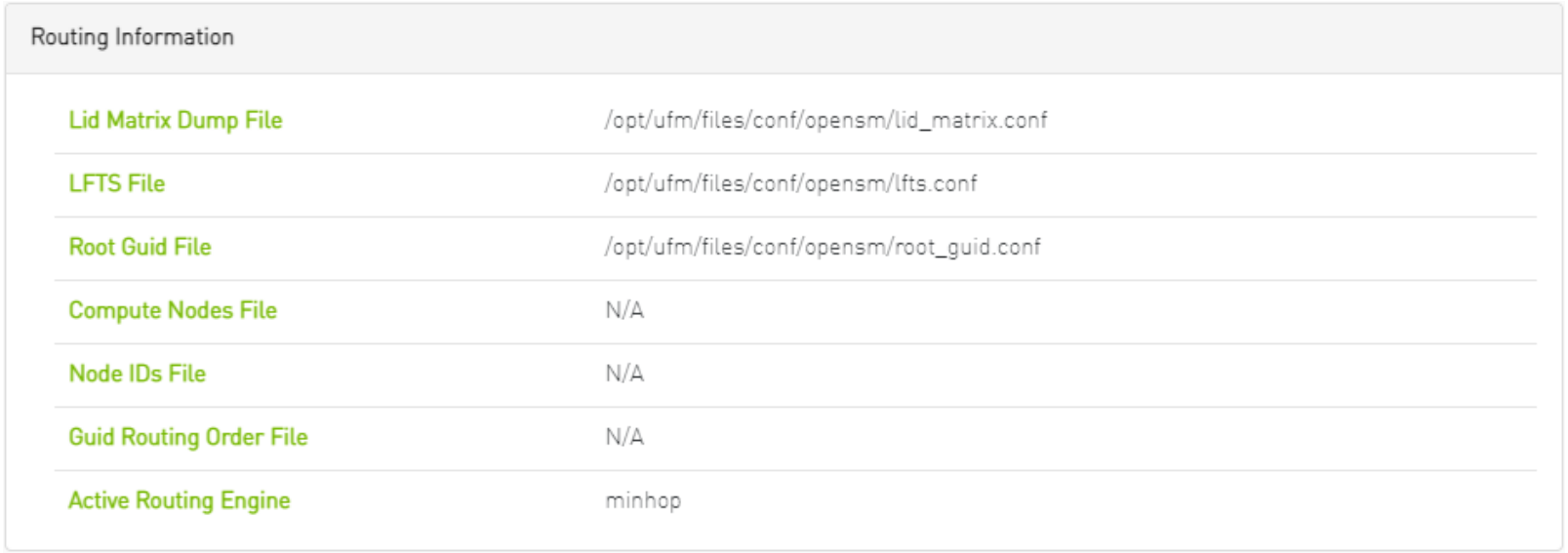
UFM web UI supports the following routing engines:

- MINHOP – based on the minimum hops to each node where the path length is optimized (i.e., shortest path available).
- UPDN – also based on the minimum hops to each node but it is constrained to ranking rules. Select this algorithm if the subnet is not a pure Fat Tree topology and deadlock may occur due to a credit loops in the subnet.
- File-Based (FILE) – The FILE routing engine loads the LFTs from the specified file, with no reaction to real topology changes.
- Fat Tree – an algorithm that optimizes routing for congestion-free "shift" communication pattern.
  Select Fat Tree algorithm if a subnet is a symmetrical or almost symmetrical fat-tree. The Fat Tree also optimizes K-ary-N-Trees by handling non-constant K in cases where leafs (CAs) are not fully staffed, and the algorithm also handles any Constant Bisectional Bandwidth (CBB) ratio. As with the UPDN routing algorithm, Fat Tree routing is constrained to ranking rules.
  - Quasi Fat Tree – PQFT routing engine is a closed formula algorithm for two flavors of fat trees
  - Quasi Fat Tree (QFT)
  - Parallel Ports Generalized Fat Tree (PGFT)
    PGFT topology may use parallel links between switches at adjacent levels, while QFT uses parallel links between adjacent switches in different sub-trees. The main motivation for that is the need for a topology that is not just optimized for a single large job but also for smaller concurrent jobs.
- Dimension Order Routing (DOR) – based on the Min Hop algorithm, but avoids port equalization, except for redundant links between the same two switches. The DOR algorithm provides deadlock-free routes for hypercubes, when the fabric is cabled as a hypercube and for meshes when cabled as a mesh.
- Torus-2QoS – designed for large-scale 2D/3D torus fabrics. In addition, you can configure Torus-2QoS routing to be *traffic aware*, and thus optimized for neighbor-based traffic.
- Routing Engine Chain (Chain) – an algorithm that allows configuring different routing engines on different parts of the IB fabric.
- Adaptive Routing (AR) – enables the switch to select the output port based on the port's load. This option is not available via UFM Web UI.
  - AR_UPDN
  - AR_FTREE
  - AR_TORUS
  - AR_DOR
- Dragonfly+ (DFP, DPF2)

Configuring Routing Protocol

Network Management tab enables setting the preferred routing protocol supported by the UFM software, as well as routing priority.

To set the desired routing protocol, move one routing protocol or more from the **Available** list to the **Selected** list, and click "Save" in the upper right corner.

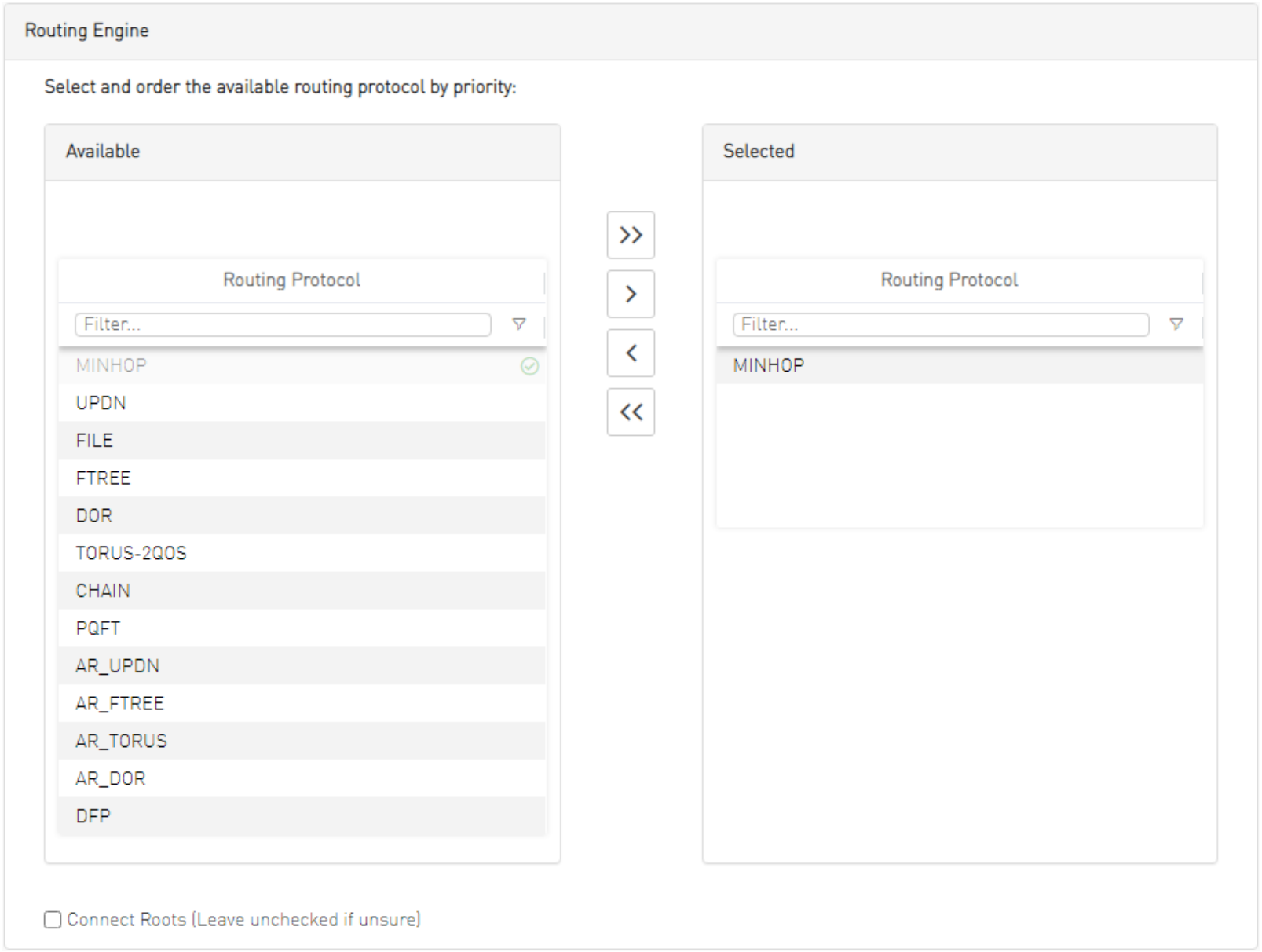


The protocol at the top of the list has the highest priority and will be chosen as the **Active Routing Engine**. If the settings for this protocol are not successful, UFM takes the next available protocol.

Routing Information is listed on the top of the screen:

Field/Box	Description
LID Matrix Dump File	File holding the LID matrix dump configuration

Field/Box	Description
LFTS File	File holding the LFT routing configuration
Root GUID File	File holding the root node GUIDS (for fat-tree or Up/Down)
Compute Nodes File	File holding GUIDs of compute nodes for fat-tree routing algorithm
GUID Routing Order File	File holding the routing order GUIDs (for MinHop and Up/Down)
Node IDs File	File holding the node IDs
Active Routing Engine	The current active routing algorithm used by the managing OpenSM


Routing Engine
Select and order the available routing protocol by priority:
Available
Routing Protocol
Filter...
MINHOP
UPDN
FILE
FTREE
DOR
TORUS-2QOS
CHAIN
PQFT
AR_UPDN
AR_FTREE
AR_TORUS
AR_DOR
DFP
Selected
Routing Protocol
Filter...
MINHOP
Connect Roots (Leave unchecked if unsure)

Subnet Manager Tab

UFM is a management platform using a user-space application for InfiniBand fabric management. This application is developed within the context of an open-source environment. This application serves as an InfiniBand Subnet Manager and a Subnet Administration tool.

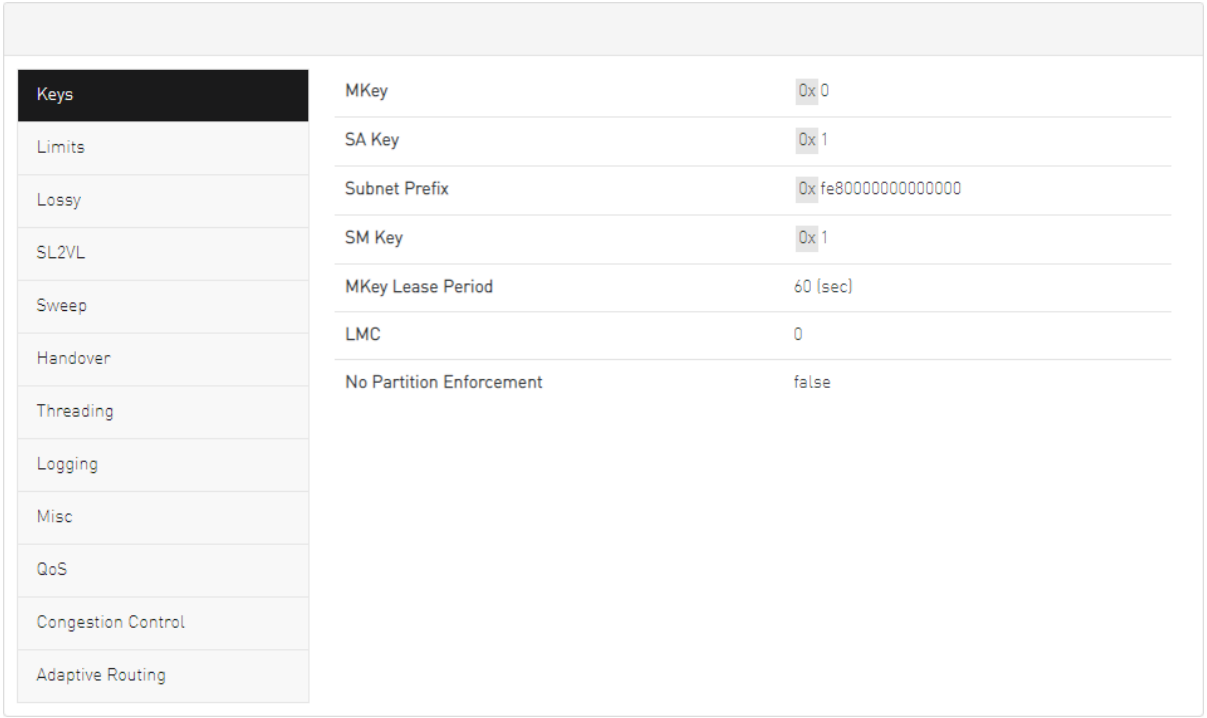
The UFM Subnet Manager (SM) is a centralized entity running on the server that discovers and configures all the InfiniBand fabric devices to enable traffic flow throughout the fabric.

To view and configure SM parameters in the ***Subnet Manager*** tab, select the relevant tab according to the required configuration.

For more information, please refer to Appendix – Enhanced Quality of Service.

SM Keys Configuration

The SM Keys tab enables you to view the Subnet Manager Keys. You cannot change the configuration in this tab.

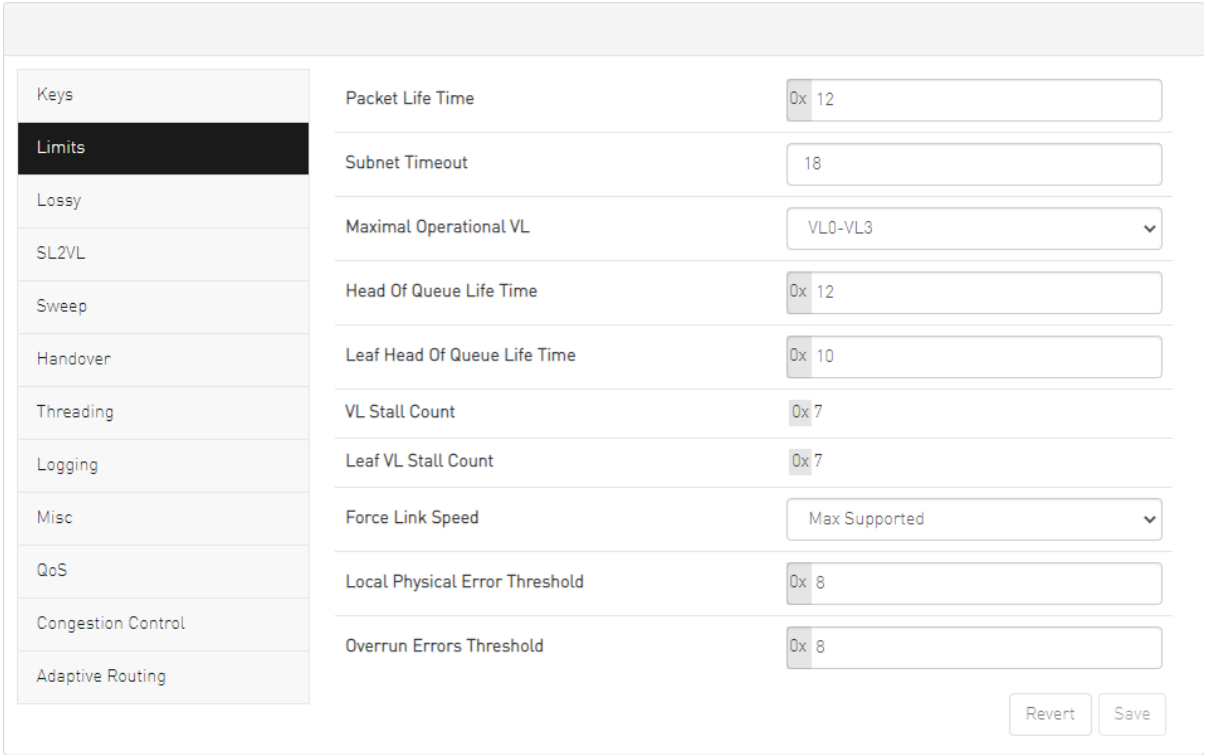


Field	Description	Default
MKey	A field that allows you to view or edit the M_Key value sent to all ports to qualify all the set (PortInfo). Authentication is performed by the management entity at the destination port and is achieved by comparing the key contained in the SMP with the key (the M_Key Management key) residing at the destination port.	0x0000000000000000
SA Key	Shows the SM_Key value to qualify the receive SA queries as 'trusted'.	0x0000000000000001
Subnet Prefix	An identifier of the subnet. The subnet prefix is used as the most significant 64 bit of the GID of each InfiniBand node in the subnet.	0xfe80000000000000

Field	Description	Default
SM Key	Read-only field that displays the Key of the Subnet Manager (SM).	0x0000000000000001
MKey Lease Period	A field that allows you to view or edit the lease period used for the M_Key on this subnet in [sec].	0
LMC	Defines the LID Mask Control value for the SM. Possible values are 0 to 7. LID Mask Control (LMC) allows you to assign more than one LID per port. NOTE: Changes to the LMC parameter require a UFM restart.	0
No Partition Enforcement	Disables partition enforcement by switches.	Disabled

SM Limits Configuration

The SM Limits tab enables you to view and set the Subnet Manager Limits.



To configure SM Limits, set the fields as described in the table below, and click “Save."

Field	Description	Default
Packet Life Time	A field that allows you to view and/or edit the code of maximum lifetime a packet in a switch. The actual time is 4.096 usec * 2^<packet_life_time>. The value 0x14 disables this mechanism	0x12

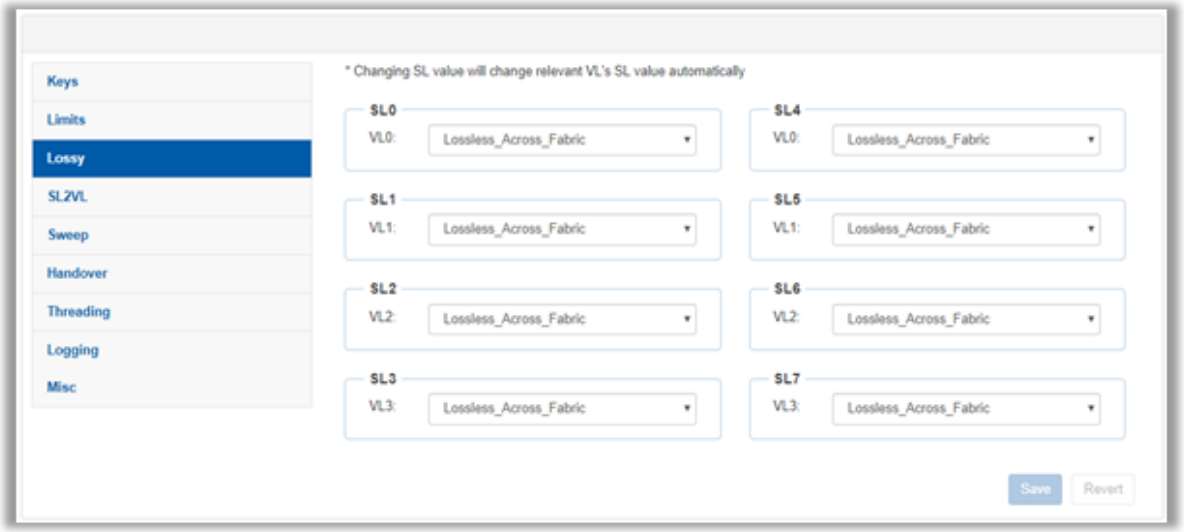
Field	Description	Default
Subnet Timeout	A field that allows you to view and/or edit the subnet_timeout code that will be set for all the ports. The actual timeout is 4.096usec * 2^<subnet_timeout>	18
Maximal Operational VL	A field that allows you to view and/or edit the limit of the maximal operational VLs: • 0: NO_CHANGE • 1: VL0 1 • 2: VL0_VL1 • 3: VL0_VL3 • 4: VL0_VL7 • 5: VL0_VL14	3
Head of Queue Life Time	A field that allows you to view and/or edit the code of maximal time a packet can wait at the head of transmission queue. The actual time is 4.096usec * 2^<head of queue lifetime> The value 0x14 disables this mechanism.	0x12
Leaf Head of Queue Life Time	A field that allows you to view and/or edit the maximum time a packet can wait at the head of queue on a switch port connected to a CA or gateway port.	0x10
VL Stall Count	A field that allows you to view the number of sequential packets dropped that cause the port to enter the VLStalled state. The result of setting this value to zero is undefined.	0x07
Leaf VL Stall Count	This field allows you to view the number of sequential packets dropped that cause the port to enter the VLStalled state. This value is for switch ports driving a CA or gateway port. The result of setting the parameter to zero is undefined.	0x07
Force Link Speed	A parameter that allows you to modify the PortInfo:LinkSpeedEnabled field on switch ports. If 0, do not modify. • Values are: • 1: 2.5 Gbps • 3: 2.5 or 5.0 Gbps • 5: 2.5 or 10.0 Gbps • 7: 2.5 or 5.0 or 10.0 Gbps • 2,4,6,8-14 Reserved	15 By default, UFM sets the enabled link speed equal to the supported link speed.

Field	Description	Default
	• 15: set to PortInfo:LinkSpeedSupported	
Local Physical Error Threshold	A field that allows you to view and/or edit the threshold of local phy errors for sending Trap 129.	0x08
Overrun Errors Threshold	A field that allows you to view and/or edit the threshold of credit overrun errors for sending Trap 130.	0x08

SM Lossy Manager Configuration

> ⚠ This tab is available to users with an advanced license only.

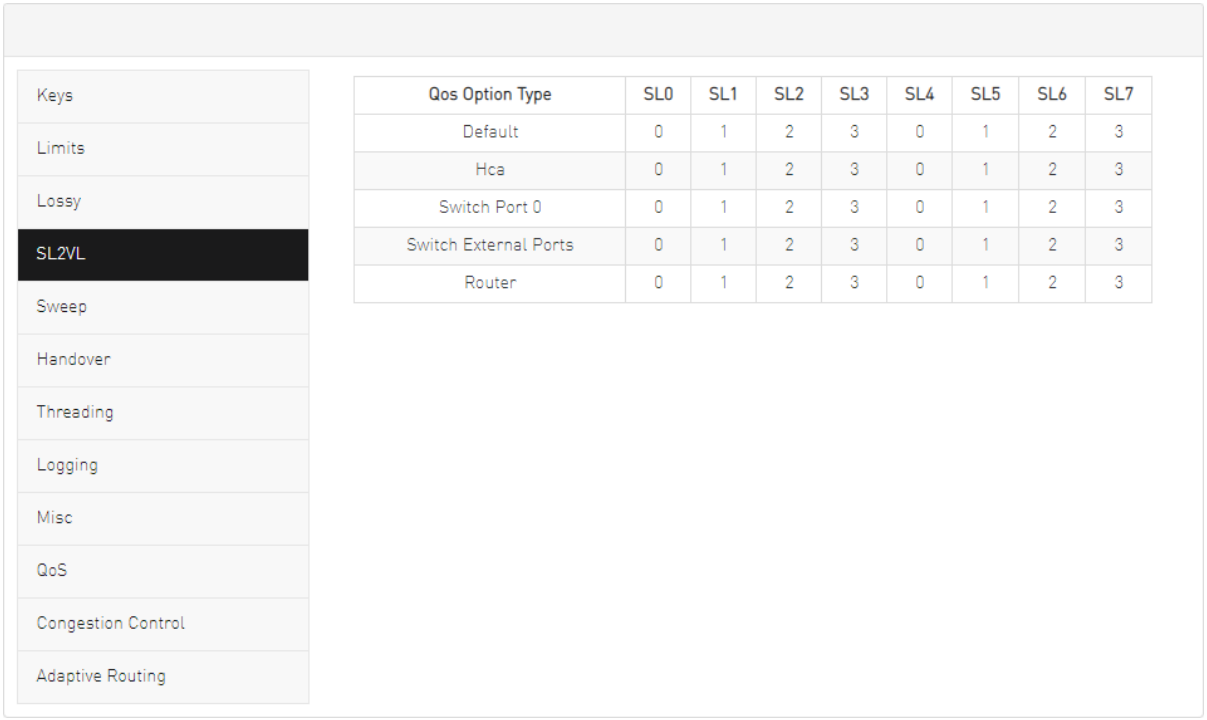
The SM Lossy tab enables you to view and set the Lossy Configuration Manager options after Lossy Configuration has been enabled.



SM SL2VL Mapping Configuration

The SM SL2VL tab enables you to view the SL (service level) to VL (virtual lane) mappings and the configured Lossy Management. You cannot change the configuration in this tab.

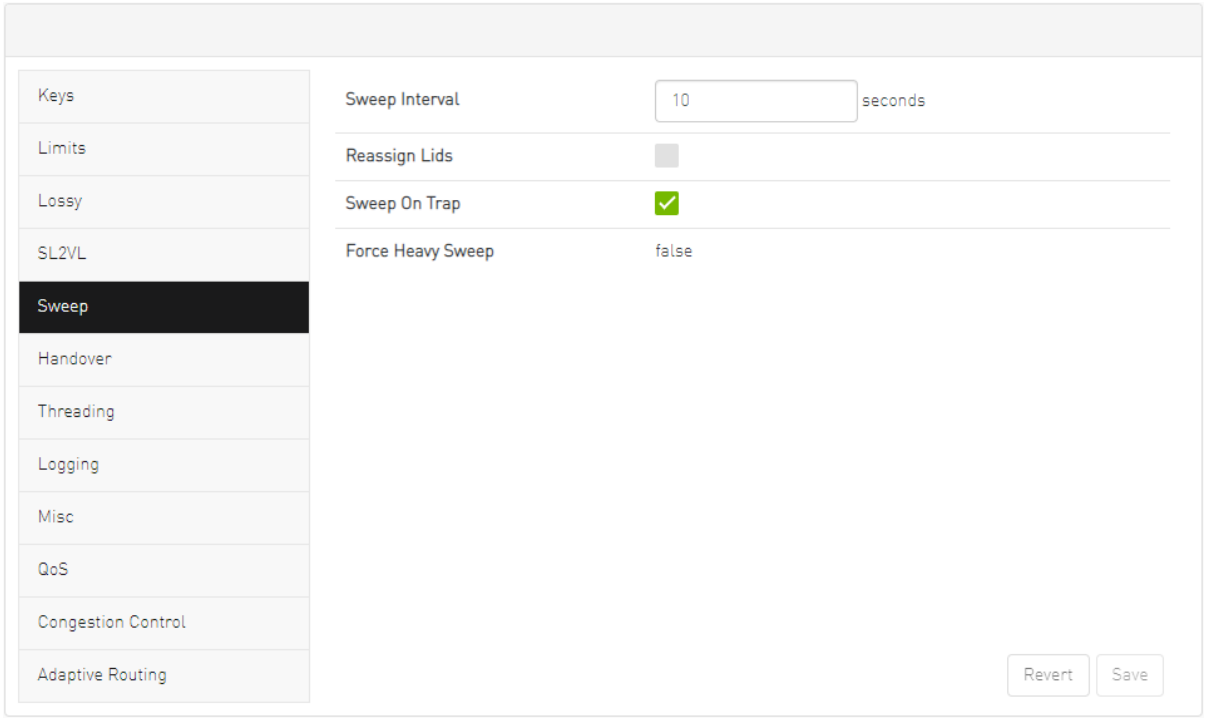
However, you can change it in the previous SM Lossy Manager Configuration (Advanced License only) tab.



Qos Option Type	SL0	SL1	SL2	SL3	SL4	SL5	SL6	SL7
Default	0	1	2	3	0	1	2	3
Hca	0	1	2	3	0	1	2	3
Switch Port 0	0	1	2	3	0	1	2	3
Switch External Ports	0	1	2	3	0	1	2	3
Router	0	1	2	3	0	1	2	3

SM Sweep Configuration

The Sweep tab enables you to view and/or set the Subnet Manager Sweep Configuration parameters.

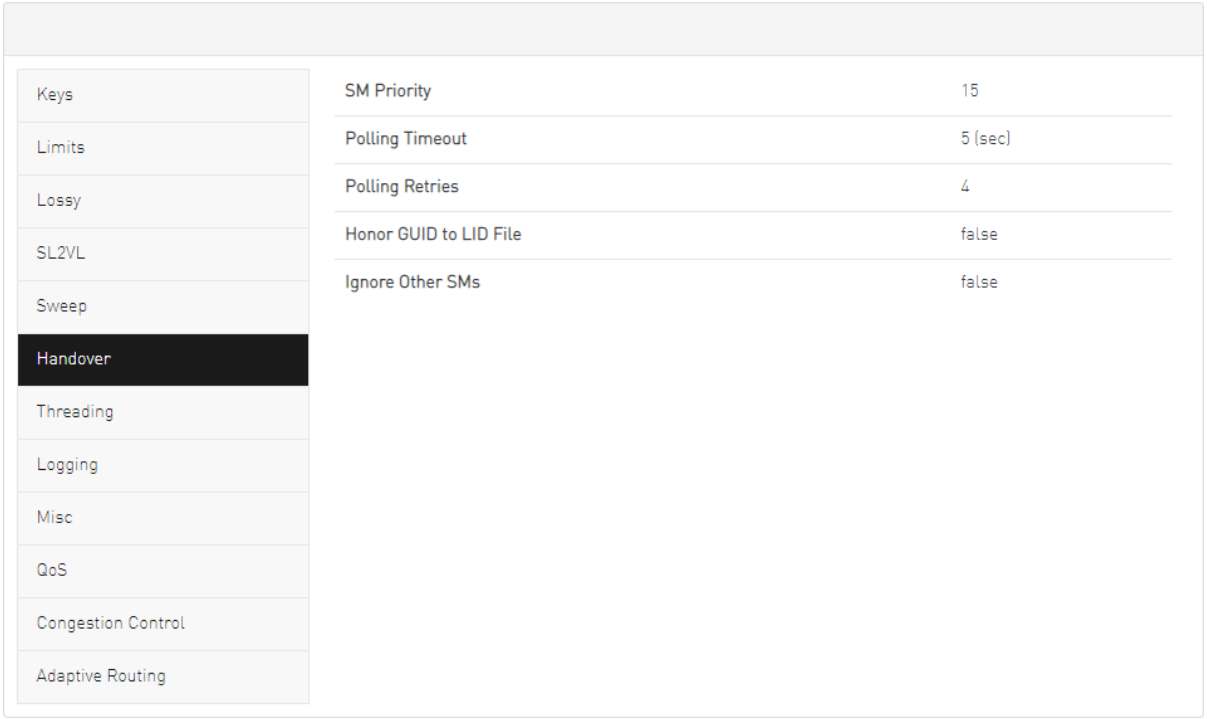


To configure SM Sweep, set the fields as described in the table below and click “Save."

Field/Box	Description	Default
Sweep Interval	A field that allows you to view and/or edit the number of seconds between light sweeps (0 disables it).	10
Reassign LIDs	If enabled, causes all LIDs to be reassigned.	Disabled
Sweep on Trap	If enabled, traps 128 and 144 will cause a heavy sweep.	Enabled
Force Heavy Sweep	If enabled, forces every sweep to be a heavy sweep.	Disabled

SM Handover Configuration

The SM Handover tab enables you to view the Subnet Manager Handover Configuration parameters. You cannot change the configuration in this tab.

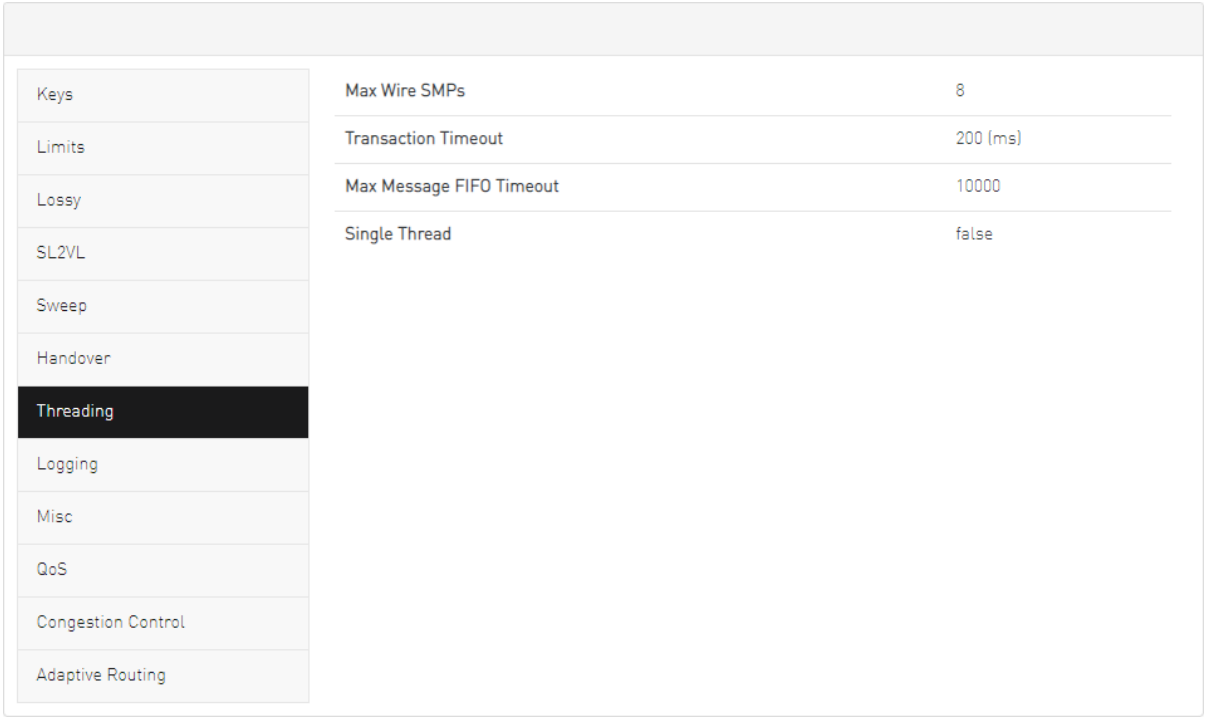


Field/Box	Description	Default
SM Priority	A field that shows the SM priority used for determining the master. Range is 0 (lowest priority) to 15 (highest). Note: Currently, these settings may not be changed.	15
Polling Timeout	A field that shows the timeout in [sec] between two polls of active master SM.	Range=10000

Field/Box	Description	Default
Polling Retries	Number of failing polls of remote SM that declares it "not operational."	4
Honor GUID to LID File	If enabled, honor the guid2lid file when coming out of standby state, if the file exists and is valid.	Disabled
Ignore other SMs	If enabled, other SMs on the subnet are ignored.	Disabled

SM Threading Configuration

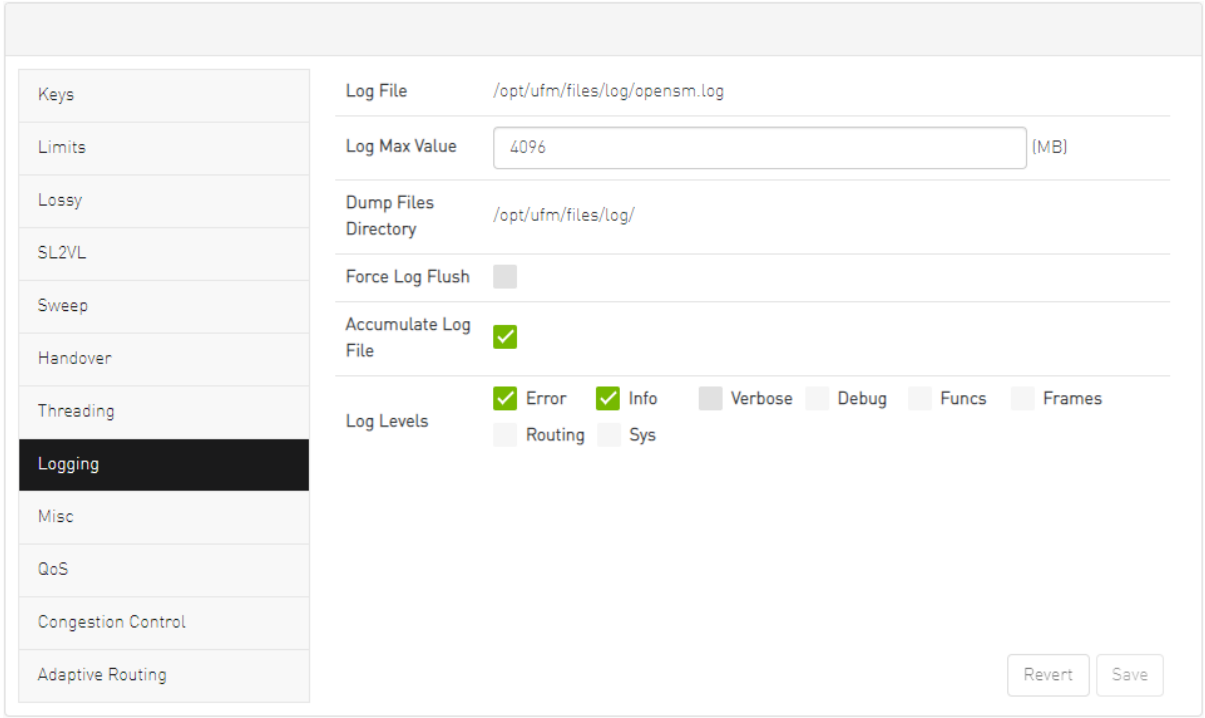
The SM Threading tab enables you to view the Subnet Manager Timing and Threading Configuration parameters. You cannot change the configuration in this tab.



Field/Box	Description	Default
Max Wire SMPs	A field that shows the maximum number of SMPs sent in parallel.	4
Transaction Timeout	A field that shows the maximum time in [msec] allowed for a transaction to complete.	200
Max Message FIFO Timeout	A field that shows the maximum time in [msec] a message can stay in the incoming message queue.	10000
Single Thread	When enabled, a single thread is used for handling SA queries.	Disabled

SM Logging Configuration

The SM Logging tab enables you to view and/or set the **Subnet Manager Logging Configuration** parameters.



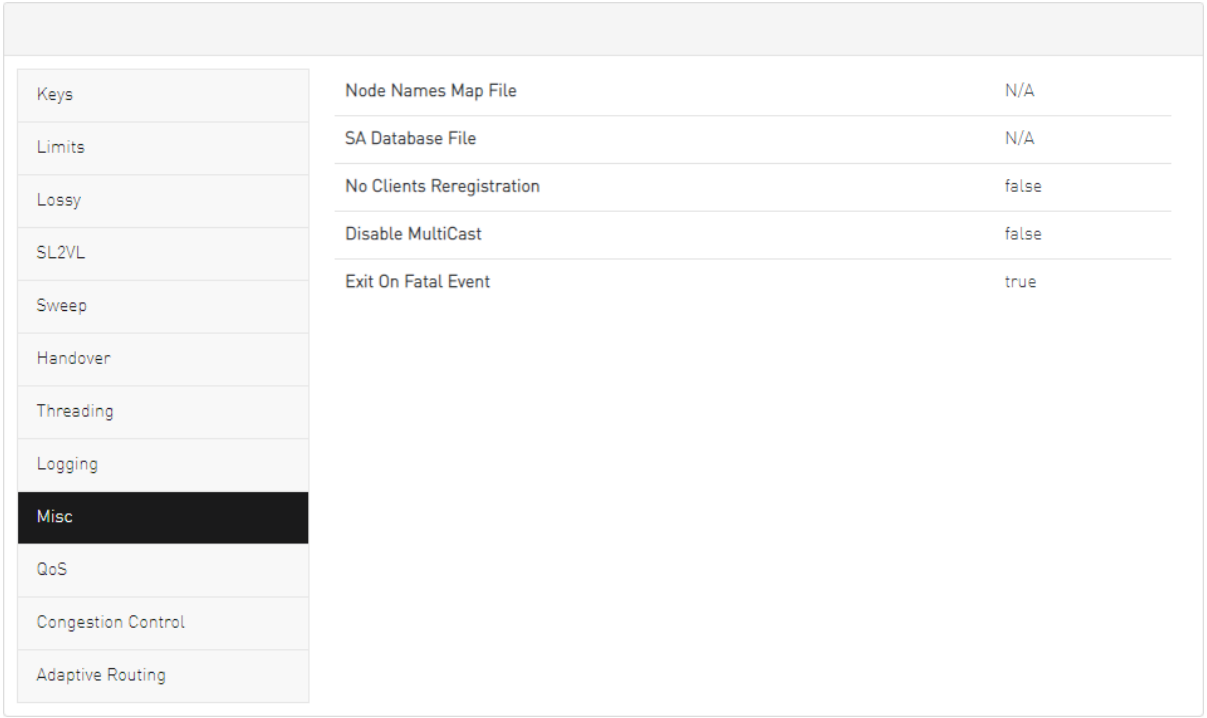
To configure SM Logging, set the fields as described in the table below and click “Save."

Field/Box	Description	Default
Log File	Path of the Log file to be used.	`opensm.log`

Field/Box	Description	Default
Log Max Size	A field that allows you to view and/or edit the size limit of the log file in MB. If overrun, the log is restarted.	4096
Force Log Flush	Force flush to the log file for each log message.	Disabled
Accumulate Log File	If enabled, the log accumulates over multiple SM sessions.	Enabled
Log Levels	Available log levels: Error, Info, Verbose, Debug, Funcs, Frames, Routing, and Sys.	Error and Info

SM Miscellaneous Settings

The Misc tab enables you to view additional **Subnet Manager Configuration** parameters. You cannot change the configuration in this tab.

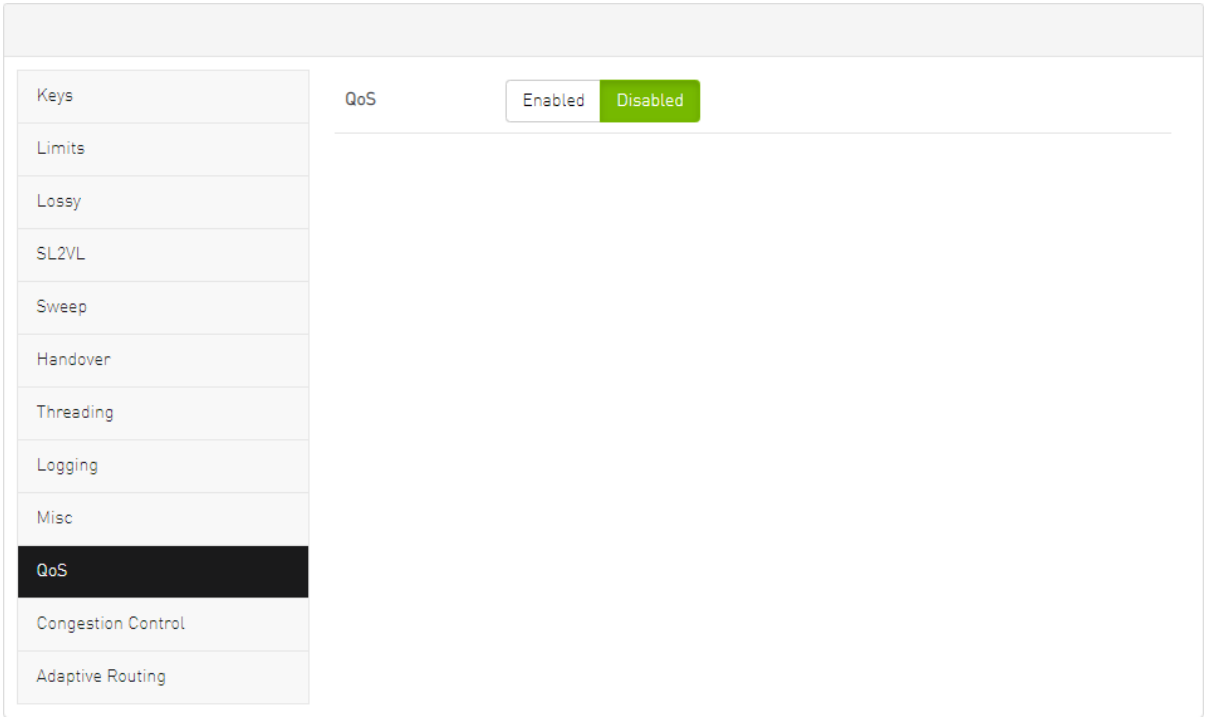


Field/Box	Description	Default
Node Names Map File	A field that allows you to view and/or set the node name map for mapping nodes to more descriptive node descriptions.	None
SA Database File	SA database file name	None
No Clients Reregistration	If enabled, disables client re-registration.	Disabled
Disable Multicast	If enabled, the SM disables multicast support and no multicast routing is performed.	Disabled

Field/Box	Description	Default
Exit on Fatal Event	If enabled, the SM exits on fatal initialization issues.	Enabled

SM QoS Configuration

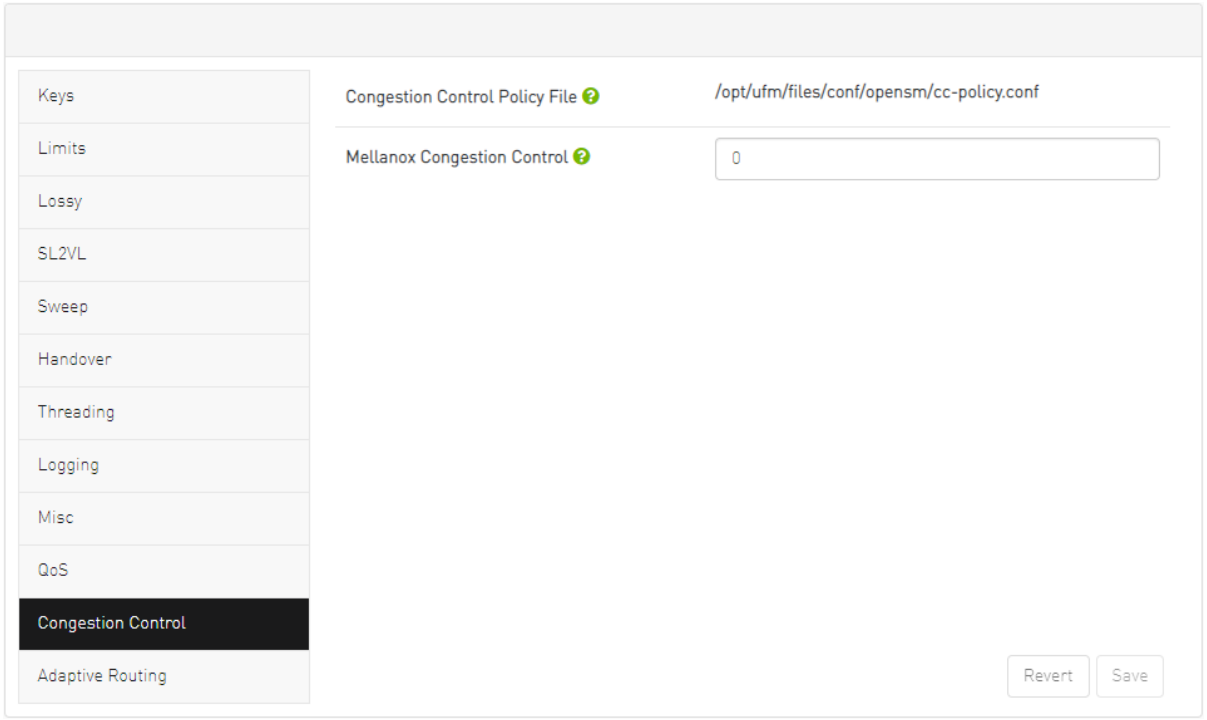
The QoS tab allows you to enable or disable QoS functionality. QoS is disabled by default.



SM Congestion Control Configuration

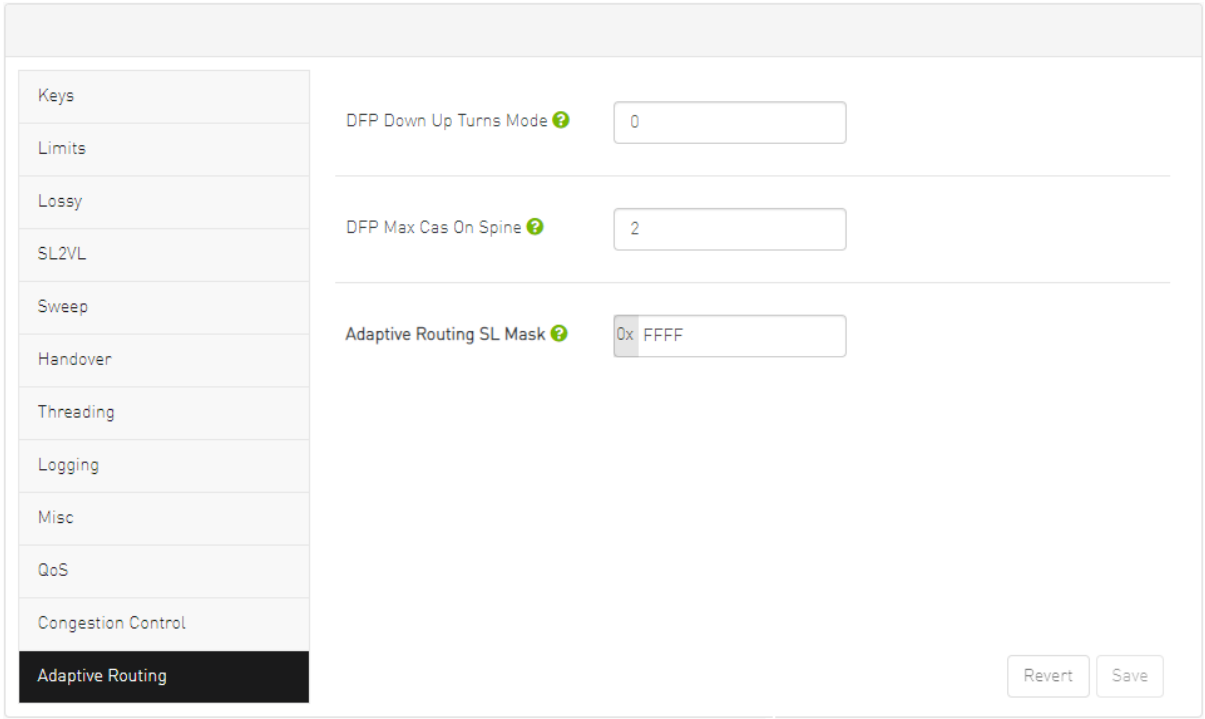
The Congestion Control tab allows you to enable, disable, or ignore congestion control.

- 0 - Ignore (default)
- 1 - Enable
- 2 - Disable



SM Adaptive Routing Configuration

The Adaptive Routing tab allows you to configure adaptive routing parameters.



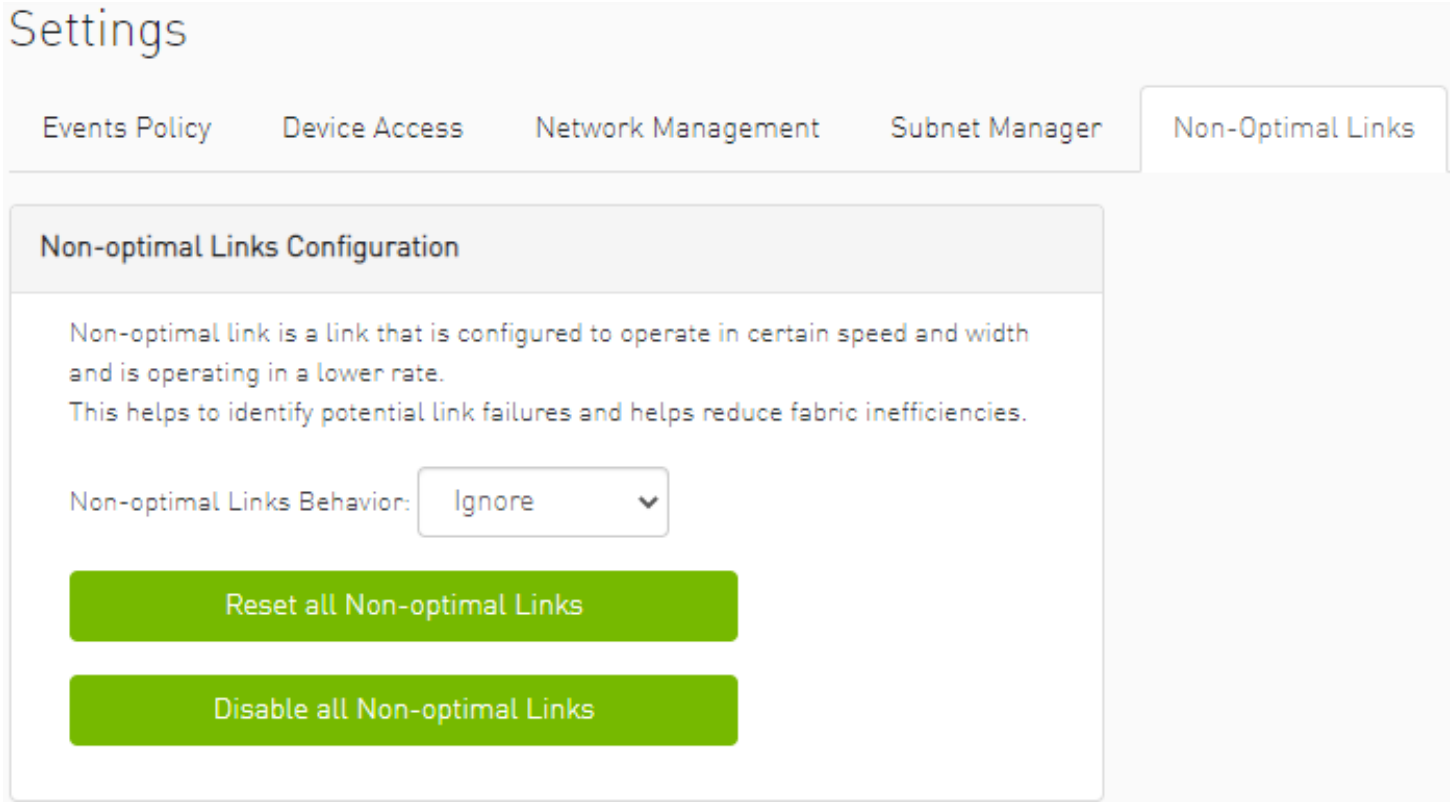
Non-Optimal Links

A non-optimal link is a link between two ports that is configured to operate at a certain speed and width and is operating at a lower rate. The Non-optimal links feature helps you identify potential link failures and reduce fabric inefficiencies.

Non-optimal links can be any of the following:

- NDR links that operate in HDR, EDR, FDR, QDR, DDR or SDR mode
- HDR links that operate in EDR, FDR, QDR, DDR or SDR mode
- EDR links that operate in FDR, QDR, DDR or SDR mode
- FDR links that operate in QDR, DDR or SDR mode
- QDR links that operate in DDR or SDR mode
- 4X links that operate in 1X mode

The Non-Optimal Links window allows you to set the preferred action for non-optimal links.



To set the non-optimal links policy:

From the drop-down menu, select the action for Non-optimal Links behavior.

The drop-down menu defines the default behavior. Options are: **Ignore** (default), **Disable**, and **Reset**.

Option	Description
Ignore	Ignore the non-optimal links

Option	Description
Reset	Reset all non-optimal links ports
Disable	Disable all non-optimal links ports

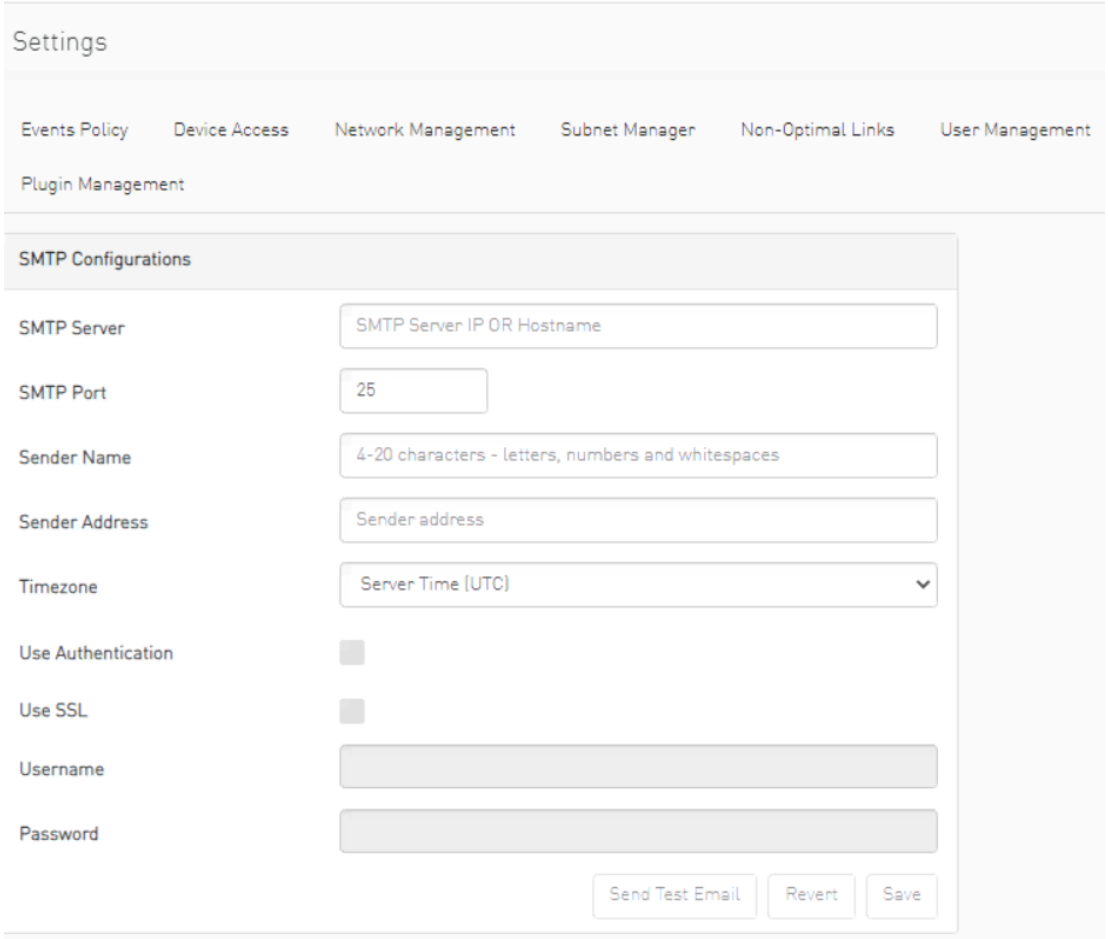
Reset all Non-Optimal Links allows users to reset all current non-optimal links ports on-demand.

Disable all Non-Optimal Links allows users to disable all current non-optimal links ports on-demand.

Email

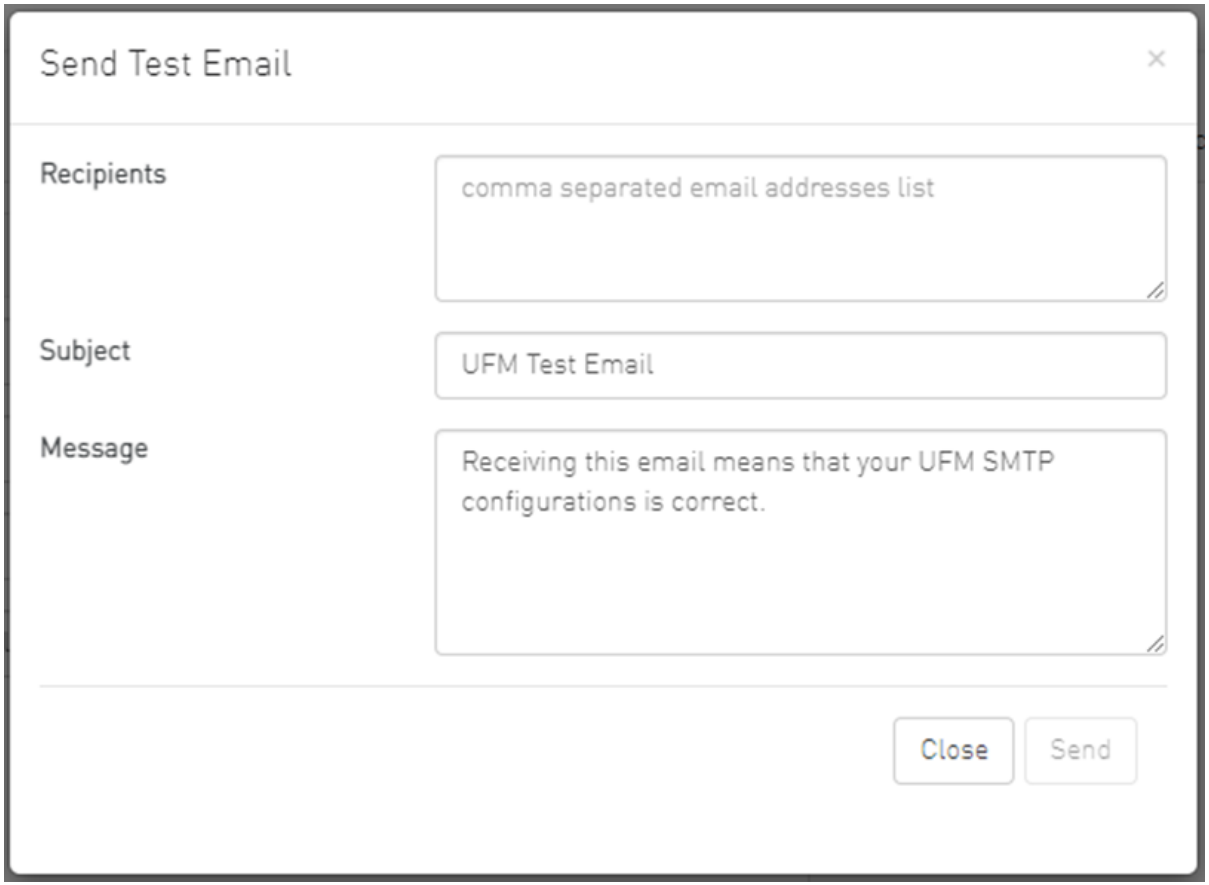
SMTP configuration is required to set both the Daily Reports Tab and the Email-on-Events features.

1. In the SMTP Configuration dialogue window, enter the following information:


Settings
Events Policy
Device Access
Network Management
Subnet Manager
Non-Optimal Links
User Management
Plugin Management
SMTP Configurations
SMTP Server
SMTP Server IP OR Hostname
SMTP Port
25
Sender Name
4-20 characters - letters, numbers and whitespaces
Sender Address
Sender address
Timezone
Server Time (UTC)
Use Authentication
Use SSL
Username
Password
Send Test Email
Revert
Save

Attribute	Description
SMTP Server	The IP or host name of the SMTP server. **Examples**: • If mail service is installed, localhost is a valid value for this field, but usually it cannot send mails outside the local domain. • smtp.gmail.com
SMTP Port	Default value – 25
Sender Name	The name that will be displayed in the email header
Sender Address	A valid email address that will be displayed in the email header
Time Zone	The default time zone for receiving sent emails is the server time zone. Users have the option to specify a different preferable time zone
Use Authentication	By default, this field is unchecked. If checked, you must supply a username and password in the respective fields
Use SSL	Default value is false – not using SSL
Username	SMTP account username
Password	SMTP account password

2. Click “Save." All configuration of the SMTP server will be saved in the UFM Database.
   Click “Send Test Email” to test the configuration and the following model will appear:



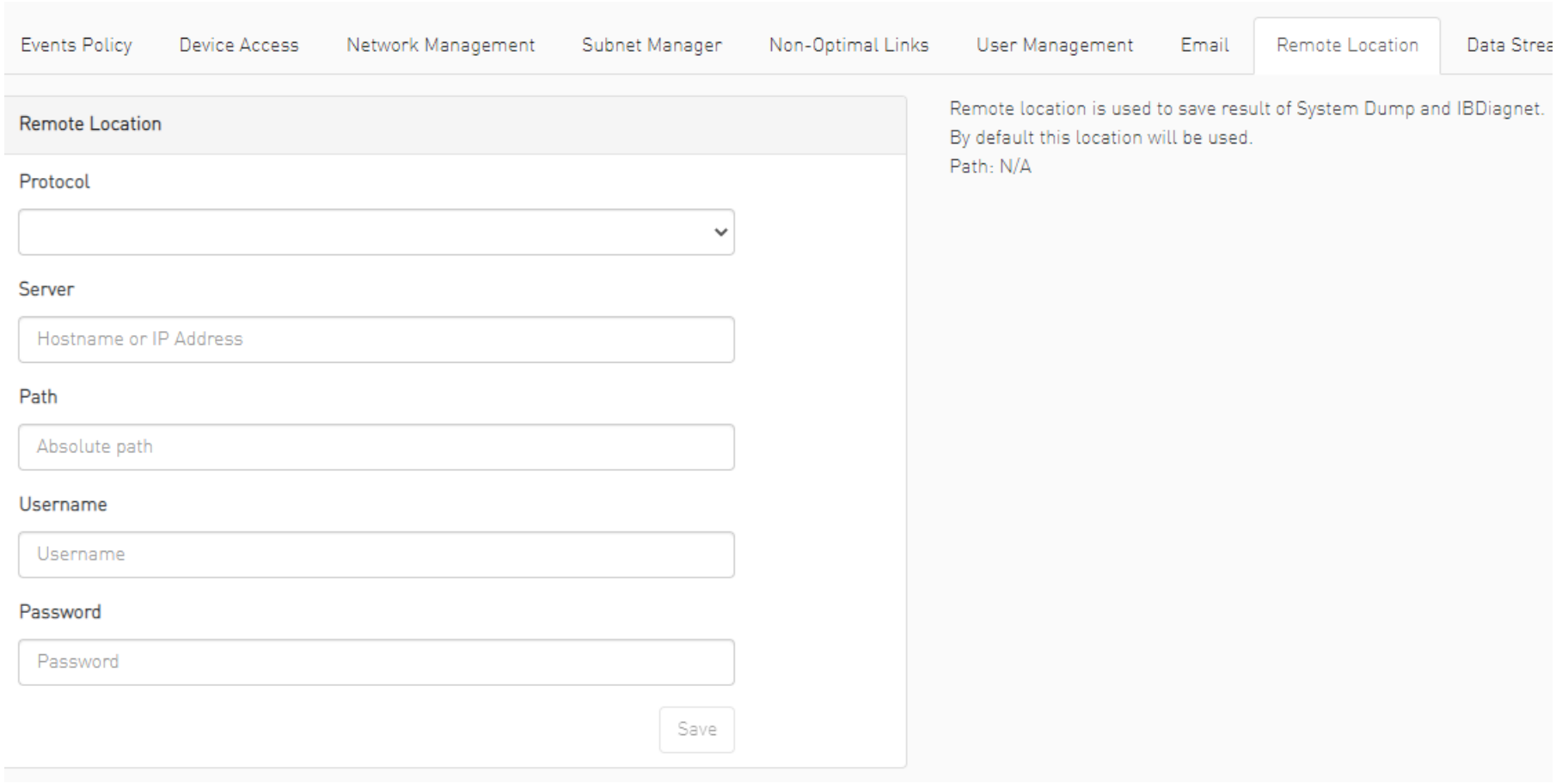
Attribute	Description
Recipients	User can choose email from event policy and daily report recipients or enter any email
Subject	Email subject

Attribute	Description
Message	Email message

The System Health window enables running and viewing reports and logs for monitoring and analyzing UFM server and fabric health through the following tabs: UFM Health, UFM Logs, UFM Snapshot, Fabric Health, Daily Reports and Topology Compare.

Remote Location

Remote location tab is used to set a predefined remote location for the results of System Dump action on switches and hosts and for IBDiagnet executions.



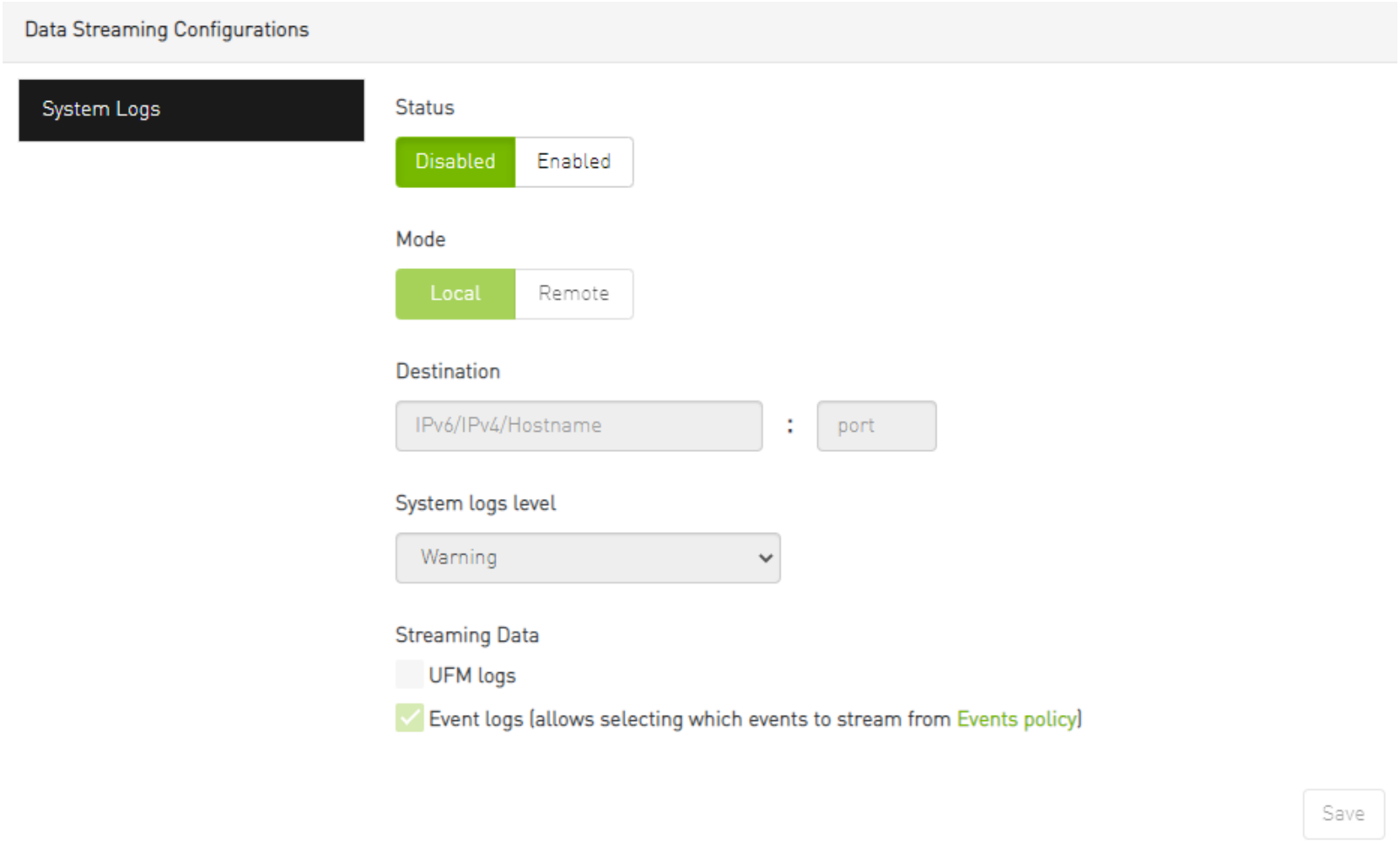
Field	Description
Protocol	The protocol to use to move the dump file to the external storage (scp/sftp)
Server	Hostname or IP address of the server

Field	Description
Path	The path where dump files are saved
Username	Username for the server
Password	Respective password

After configuring these parameters, it would be possible for users to collect sysdumps for specific devices, groups, or links (through Network Map/Cables Window) by right-clicking the item and selecting System Dump.

Data Streaming

This section allows users to configure System Logs settings via web UI.

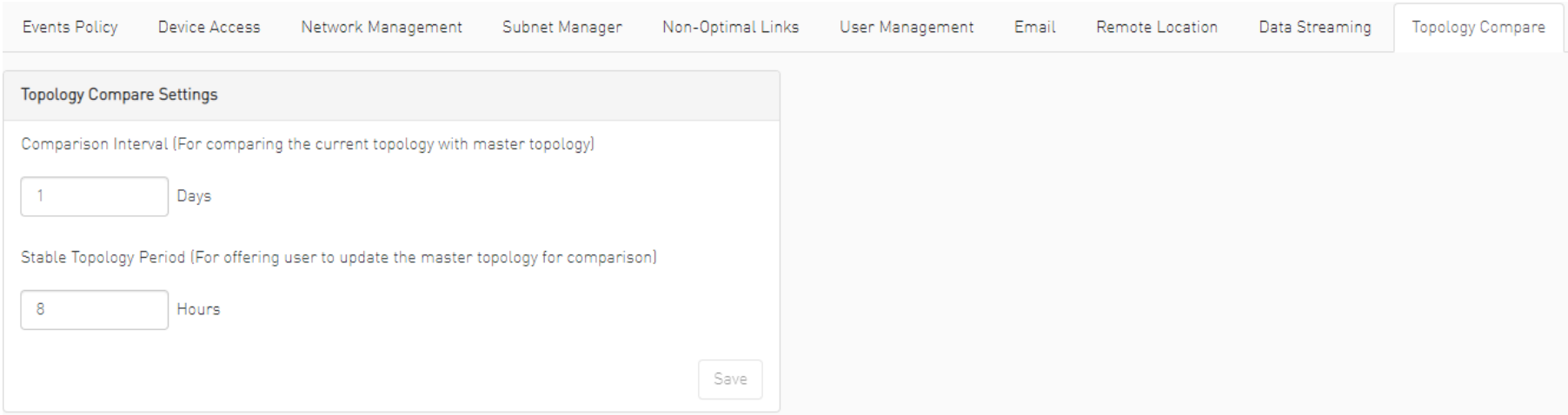


Field	Description
Status	Enable/disable exporting UFM logs to system logs

Field	Description
Mode	Export logs to local or remote system logs
Destination	Remote server IP/hostname and port
System Logs Level	Log level to export
Streaming Data	Logs to export to system logs. ⚠ Events logs are selected one by one from Events Policy settings when the system logs feature is enabled.

Topology Compare

This tab controls the settings for the Periodic Topology Comparison feature.



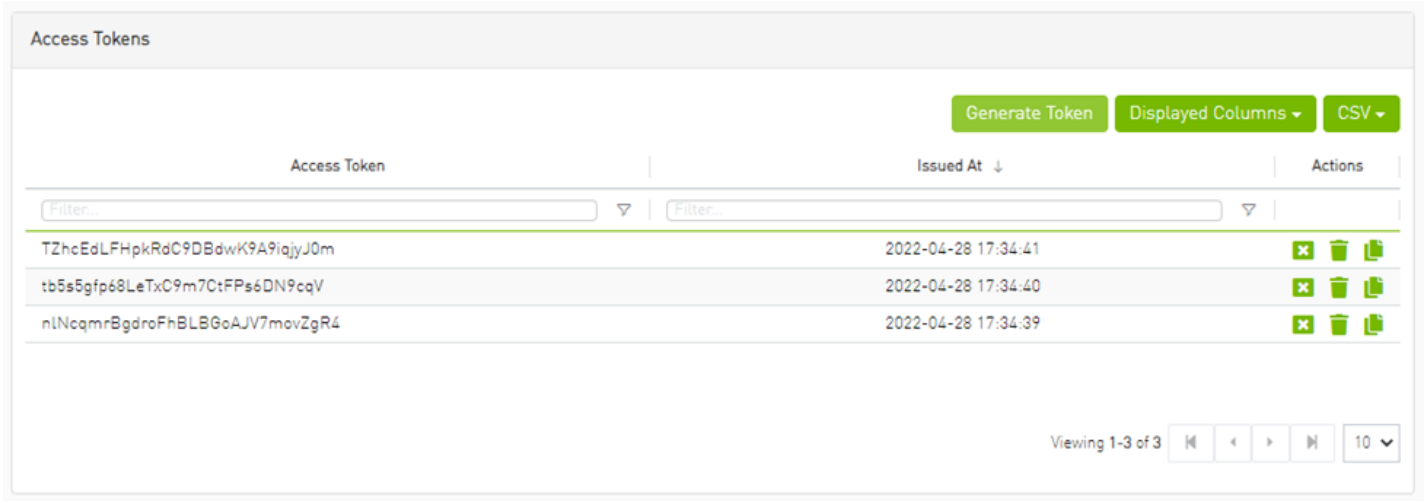
- Comparison Interval - determines how often the current topology is compared against the master topology
- Stable Topology Period - determines how long a topology must be stable before it is designated the new master topology

Token-based Authentication

Token-based authentication is a protocol which allows users to verify their identity, and in return receive a unique access token. During the life of the token, users then access the UFM APIs that the token has been issued for, rather than having to re-enter credentials each time they need to use any UFM API.

> ⚠ Under the Settings section there is a tab titled called “Access Tokens”.

The functionality of the added tab is to give the user the ability to create new tokens & manage the existing ones (list, copy, revoke, delete):



Actions:

Name	Icon	Description
Revoke		Revoke a specific token. ⚠ The revoked token will no longer be valid.
Delete		Delete a specific token.
Copy		Copy specific token into the clipboard.

⚠ Each user is able to list and manage only the tokens that have been created by themselves. Only the users with system_admin role will be able to create tokens.

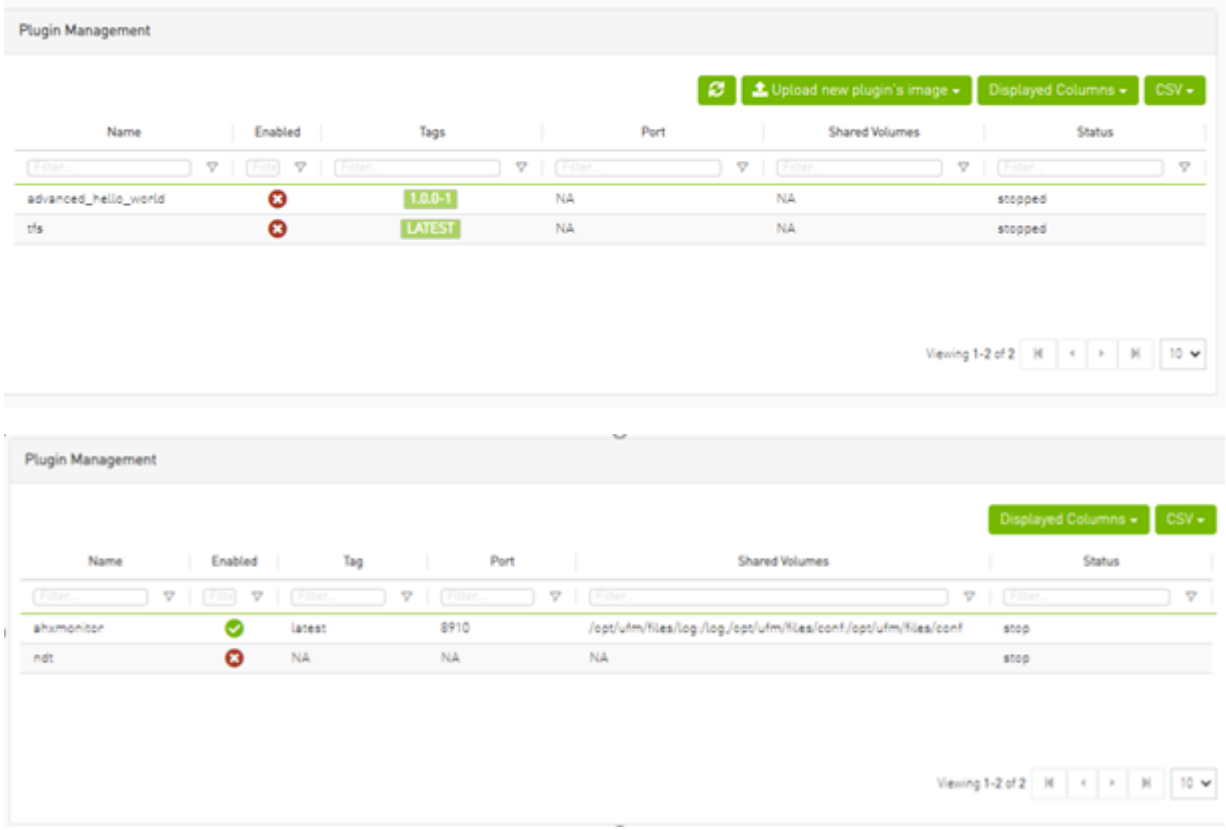
Plugin Management

Plugin management allows users to manage UFM plugins without using CLI commands. Under "Settings", there is a tab titled “Plugin Management”.

The functionality of the "Plugin Management" tab is to give the user the ability to add, remove, disable and enable plugins.

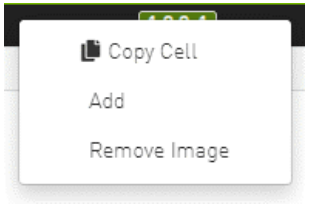
Furthermore, the plugin management feature allows loading a plugin's image in two ways: either by remotely pulling it from a Docker Hub repository or by directly uploading the image file from the user's local machine.

By default, the option to load plugin images is hidden. However, users can enable this functionality by modifying the 'gv.cfg' file. Specifically, they need to set '`upload_plugins_images_via_gui`' under the Plugins section to '`true`'.

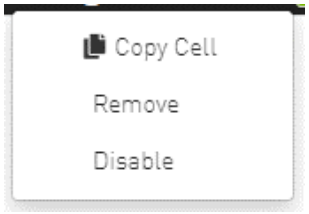

Plugin Management
Upload new plugin's image
Displayed Columns
CSV
Name
Enabled
Tags
Port
Shared Volumes
Status
advanced_hello_world
1.0.0-1
NA
NA
stopped
tfs
LATEST
NA
NA
stopped
Viewing 1-2 of 2
10
Plugin Management
Displayed Columns
CSV
Name
Enabled
Tag
Port
Shared Volumes
Status
ahxmonitor
latest
8910
/opt/ufm/files/log:/log,/opt/ufm/files/conf:/opt/ufm/files/conf
stop
ndt
NA
NA
NA
stop
Viewing 1-2 of 2
10

Actions:

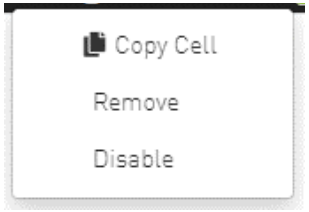
- Add – Used to add a selected plugin, opens a model to select the needed tag.



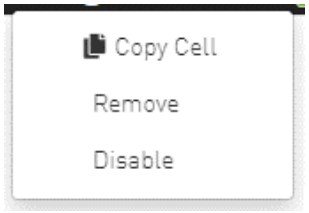
- Remove – Used to remove a selected plugin.



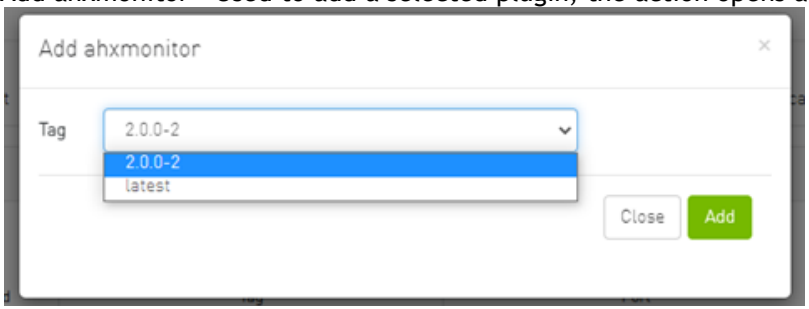
- Disable – Used to disable a selected plugin, so the plugin is disabled once the UFM is disabled.



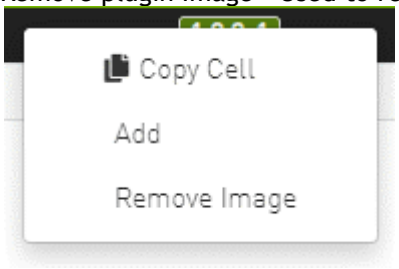
- Enable – Used to enable a selected plugin, so the plugin is enabled once the UFM is enabled.



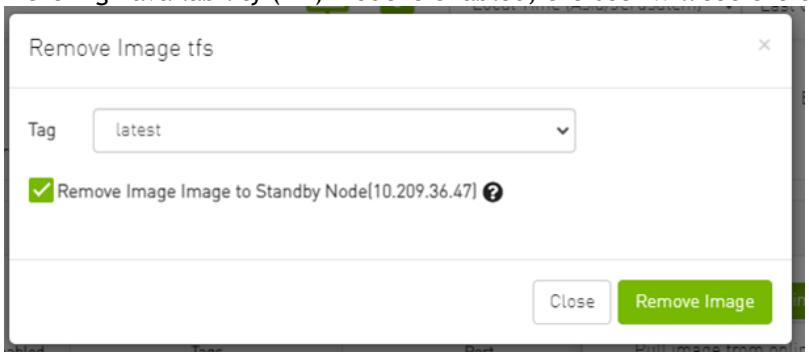
- Add ahxmonitor – Used to add a selected plugin; the action opens a modal to select the requested tag.



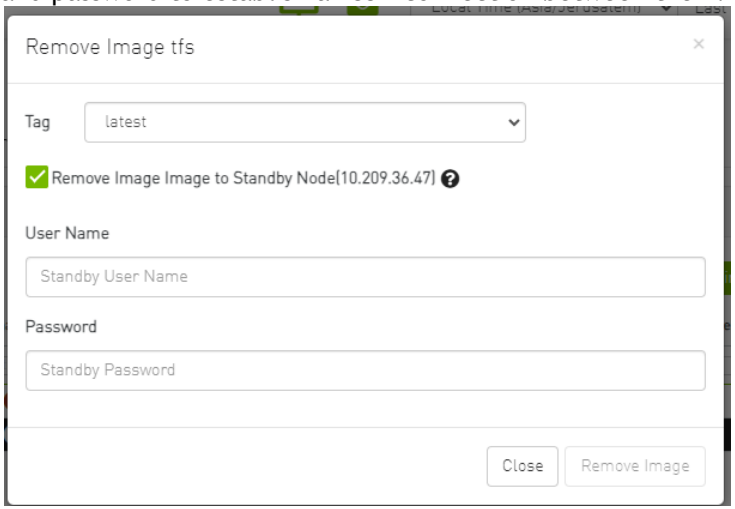
- Remove plugin Image – Used to remove plugin image



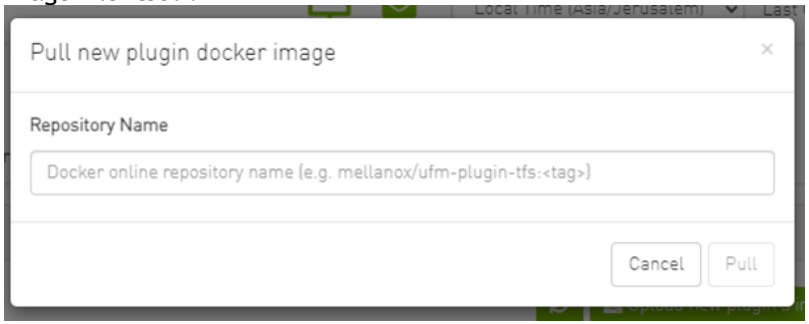
If the high availability (HA) mode is enabled, the user will see the option to remove the image from the standby node as well.



In cases where there is no established trust communication between the master and standby nodes, the user will be required to provide a username and password to establish an SSH connection between them.

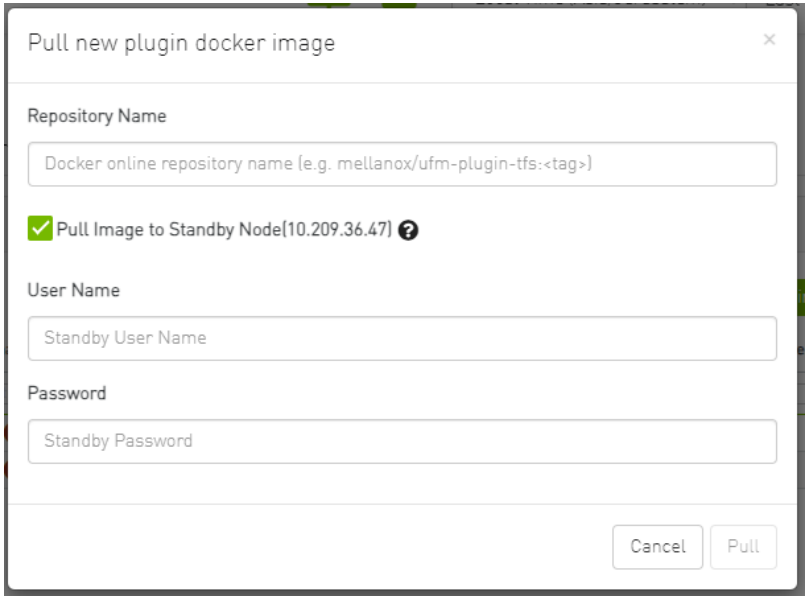


- Pull plugin Image – Used to pull plugin image remotely (e.g. from a Docker Hub repository) or by loading it from user local machine by uploading the image file itself.

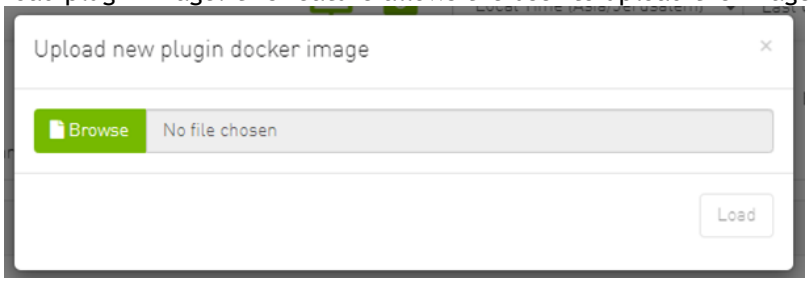


If the high availability (HA) mode is active, the user will be presented with the choice to pull the image to the standby node as well.

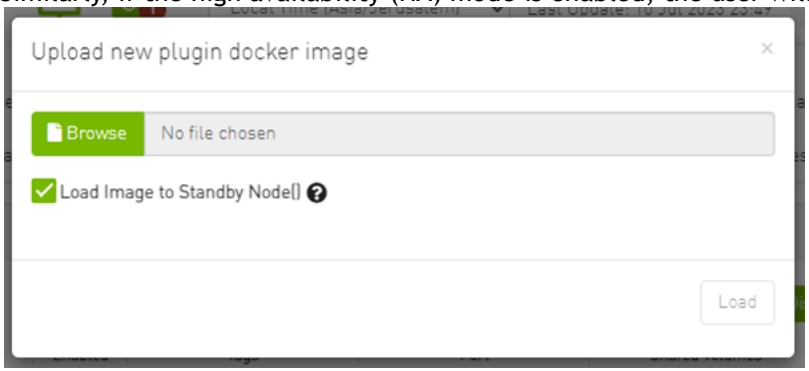
Once again, in the absence of trusted communication between the master and standby nodes, the user will need to input a username and password to create an SSH connection between the nodes.



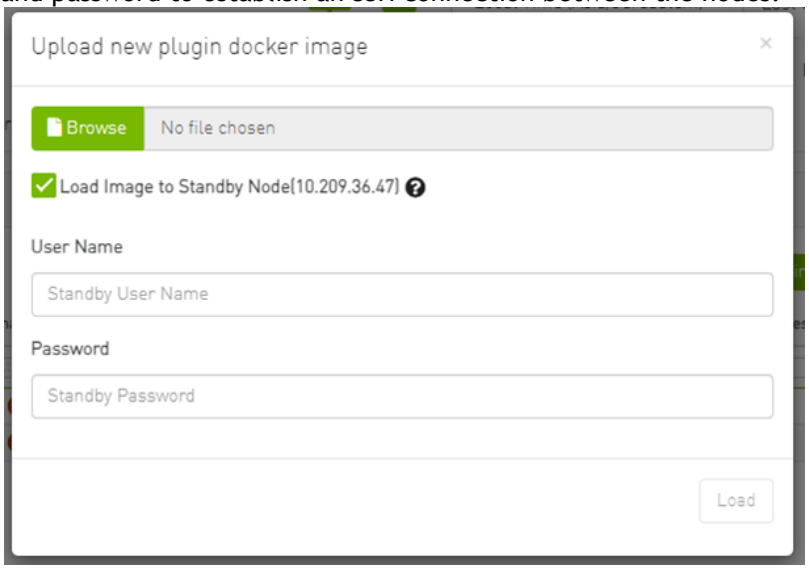
- Load plugin Image: this feature allows the user to upload the image file from their local machine directly.



Similarly, if the high availability (HA) mode is enabled, the user will have the option to load the image to the standby node too.

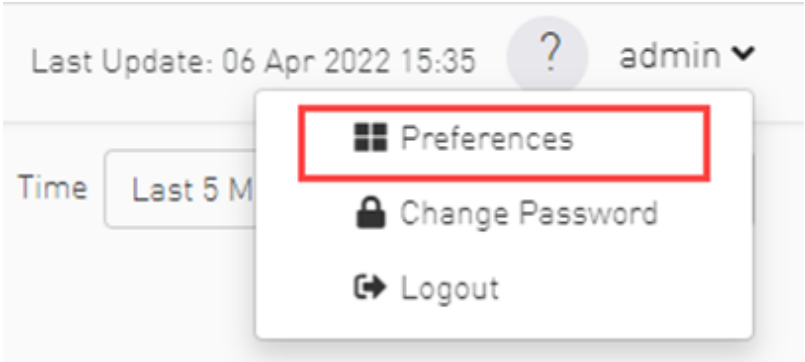


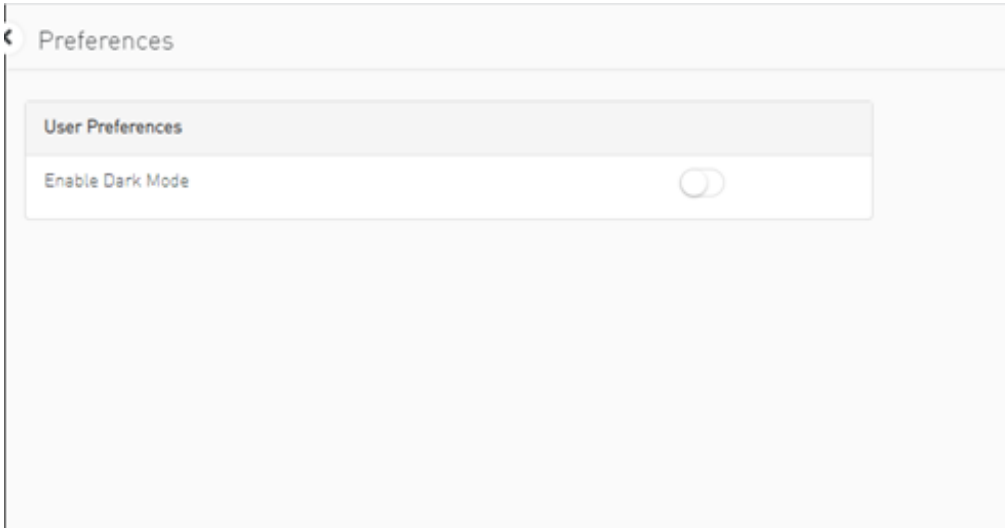
And, as mentioned earlier, if there is no trusted communication between the master and standby node, the user will need to provide a username and password to establish an SSH connection between the nodes.



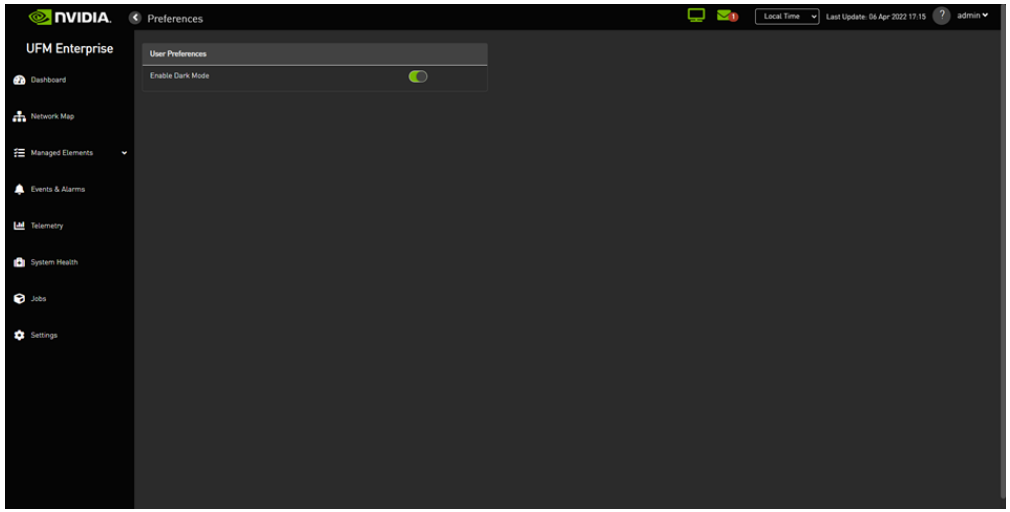
User Preferences

This page allows user to change UI preferences in general.





When user enables dark mode, the UFM is presented in dark theme.


NVIDIA
Preferences
Local Time
Last Update: 06 Apr 2022 17:15
admin
UFM Enterprise
Dashboard
Network Map
Managed Elements
Events & Alarms
Telemetry
System Health
Jobs
Settings
User Preferences
Enable Dark Mode

User Management

User Authentication

UFM®-SDN Appliance provides an authentication, authorization and accounting (AAA) mechanism which enables management of users with different roles (capabilities).

It supports local users created on the appliance, and remote (centrally managed) users created on central authentication servers such as LDAP.

Creating User Accounts

UFM®-SDN Appliance supports local and centrally managed users (Using LDAP, RADIUS, TACACS).

UFM®-SDN Appliance has predefined users. For the complete list of users and their roles, please refer to the UFM®-SDN Command Reference Guide, section *Users and Roles (Capabilities)*.

To create a local user:

1. Log into the CLI as admin and use the selected password.
2. Enter the “config” configuration mode.

```
ufm-appliance > enable
ufm-appliance # configure terminal
```

3. Creates the users with a specific capability (role).

```
ufm-appliance [ mgmt-sa ] (config) # username <username> capability <capability>
```

4. Set the user’s password.

```
ufm-appliance [ mgmt-sa ] (config) # username <username> password <password>
```

To link users to an LDAP server (such as ActiveDirectory):

1. Create a group for UFM®-SDN Appliance users on the ActiveDirectory server
2. Assign users to the group.
3. [Optional] Add the attribute "localUserName" to the user schema for correlating it to a local username on the UFM®-SDN Appliance (If this attribute is not added, all users will be correlated with the default user - "admin")
4. Add LDAP as the authentication method.

```
ufm-appliance [ mgmt-sa ] (config) # aaa authentication login default local ldap
```

5. Configure the LDAP server.

```
ufm-appliance [ mgmt-sa ] (config) # ldap base-dn <string>
ufm-appliance [ mgmt-sa ] (config) # ldap bind-dn <string>
ufm-appliance [ mgmt-sa ] (config) # ldap bind-password <string>
ufm-appliance [ mgmt-sa ] (config) # ldap host <IP Address>
```

For further information, please refer to the UFM®-SDN Command Reference Guide, section *User Management and AAA*.

User Account Management

The default user (admin) has System Administration rights. A user with system Administration rights can manage other users' accounts, including creation, deletion, and modification of accounts.
A UFM user can belong to one of the following groups:

- Monitoring Only – Users can see the fabric configuration, open monitoring sessions, define monitoring templates, and export monitoring data to CSV files. Monitoring only mode enables device hardware alerts and other device management capabilities with pre-defined device credentials.
- Fabric Operator – Users can perform all operations allowed to Monitoring group users, and can also configure fabric, modify the fabric design, define logical objects, and allocate resources. Management group users cannot create, delete or modify environments or global networks.

- Fabric Administrator – Users can perform all operations allowed to Management group users, and can also create, delete, and modify environments and global networks. Users in this group cannot manage other users' accounts.
- System Administrator – Users can perform all operations allowed to Fabric Administration group, and can also manage other users' accounts.

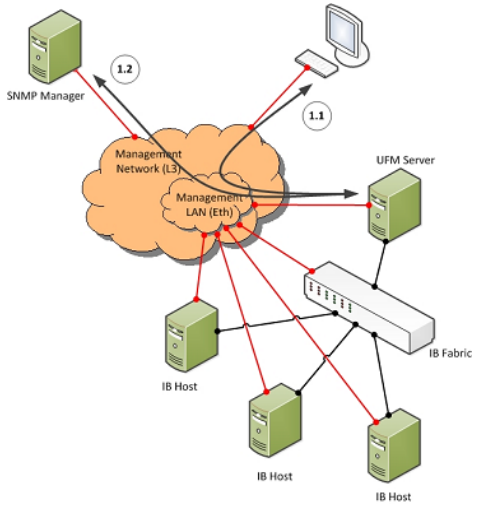
UFM Communication Requirements

This chapter describes how the UFM server communicates with InfiniBand fabric components.

UFM Server Communication with Clients

The UFM Server communicates with clients over IP. The UFM Server can belong to a separate IP network, which can also be behind the firewall.

UFM Server Communication with Clients



UFM Server Communication with UFM Web UI Client

Communication between the UFM Server and the UFM web UI client is HTTP(s) based. The only requirement is that TCP port 80 (443) must not be blocked.

UFM Server Communication with SNMP Trap Managers

The UFM Server can send SNMP traps to configured SNMP Trap Manager(s). By default, the traps are sent to the standard UDP port 162. However, the user can configure the destination port. If the specified port is blocked, UFM Server traps will not reach their destination.

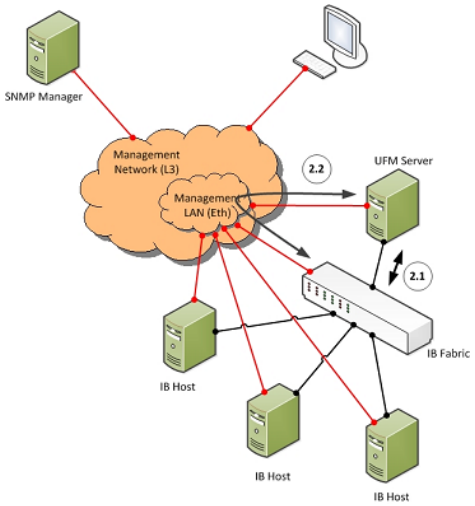
Summary of UFM Server Communication with Clients

Affected Service	Network	Address / Service / Port	Direction
Web UI Client	Out-of-band management*	HTTP / 80 HTTPS / 443	Bi-directional
SNMP Trap Notification	Out-of-band management*	UDP / 162 (configurable)	UFM Server to SNMP Manager

*If the client machine is connected to the IB fabric, IPoIB can also be used.

UFM Server Communication with InfiniBand Switches

UFM Server Communication with InfiniBand Switches



UFM Server InfiniBand Communication with Switch

The UFM Server must be connected directly to the InfiniBand fabric (via an InfiniBand switch). The UFM Server sends the standard InfiniBand Management Datagrams (MAD) to the switch and receives InfiniBand traps in response.

UFM Server Communication with Switch Management Software (Optional)

The UFM Server auto-negotiates with the switch management software on Mellanox Grid Director switches. The communication is bound to the switch Ethernet management port.

The UFM Server sends a multicast notification to MCast address 224.0.23.172, port 6306 (configurable). The switch management replies to UFM (via port 6306) with a unicast message that contains the switch GUID and IP address. After auto-negotiation, the UFM server uses Switch JSON API (HTTPS based) to retrieve inventory data and to apply switch actions (software upgrade and reboot) on the managed switch.

The following Device Management tasks are dependent on successful communication as described above:

- Switch IP discovery

- FRU Discovery (PSU, FAN, status, temperature)
- Software and firmware upgrades

The UFM Server manages IB Switch Devices over **HTTPS** (default port **443** – configurable) and / or SSH (default port 22 – configurable).

UFM Server Communication with Externally Managed Switches (Optional)

UFM server uses Ibdiagnet tool to discover chassis information (PSU, FAN, status, temperature) of the externally managed switches.

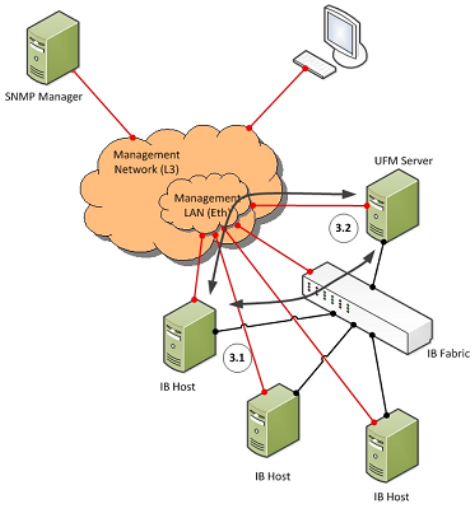
By monitoring chassis information data, UFM can trigger selected events when module failure occurs or a specific sensor value is above threshold.

Summary of UFM Server Communication with InfiniBand Switches

Affected Service	Network	Address / Service / Port	Direction
InfiniBand Management / Monitoring	InfiniBand	Management Datagrams	Bi-directional
Switch IP Address Discovery (auto-negotiation with switch management software)	Out-of-band management	Multicast 224.0.23.172, TCP / 6306 (configurable)	Multicast: UFM Server to switch TCP: Bi-directional
Switch Chassis Management / Monitoring	Out-of-band management	TCP / UDP / 6306 (configurable) SNMP / 161 (configurable) SSH / 22 (configurable)	Bi-directional

UFM Server Communication with InfiniBand Hosts

UFM Server Communication with InfiniBand Hosts



UFM Server InfiniBand Communication with HCAs

The UFM Server must be connected directly to the InfiniBand fabric. The UFM Server sends the standard InfiniBand Management Datagrams (MADs) to the Host Card Adapters (HCAs) and receives InfiniBand traps.

UFM Server Communication with Host Management (Optional)

The UFM Server auto-negotiates with the UFM Agent on a Host. The UFM Host Agent can be bound to the management Ethernet port or to an IPoIB interface (configurable). The UFM Server sends a multicast notification to MCast address 224.0.23.172, port 6306 (configurable). The UFM Agent replies to UFM (port 6306) with a unicast message that contains the host GUID and IP address. After auto-negotiation, the UFM Server and UFM Agent use XML-based messaging.

The following Device Management tasks are dependent on successful communication as described above:

- Host IP discovery
- Host resource discovery and monitoring: CPU, memory, disk
- Software and firmware upgrades

⚠ UFM 3.6 supports in-band HCA FW upgrade. This requires enabling FW version and PSID discovery over vendor-specific MADs. for more information, see the UFM User Manual.

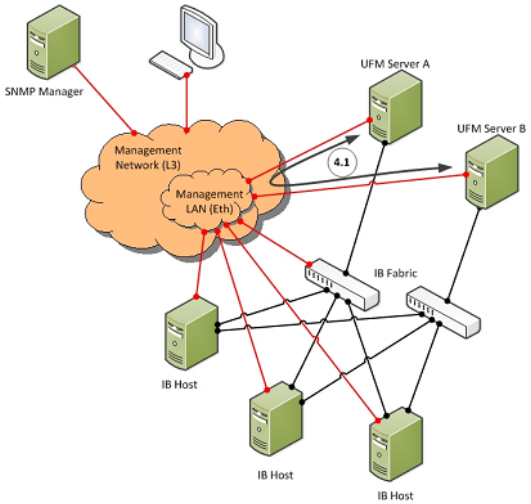
The UFM Server connects to the hosts over SSH (default port 22 - configurable) with root credentials, which are located in the UFM Server database.

Summary of UFM Server Communication with InfiniBand Hosts

Affected Service	Network	Address / Service / Port	Direction
InfiniBand Management / Monitoring	InfiniBand	Management Datagrams	Bi-directional
Host IP Address Discovery (auto-negotiation with UFM Host Agent)	Out-of-band management or IPoIB	Multicast 224.0.23.172, TCP / 6306 (configurable)	Multicast: UFM Server to UFM Agent TCP: Bi-directional
Host OS Management / Monitoring	Out-of-band management or IPoIB	TCP / UDP / 6306 (configurable) SSH / 22 (configurable)	Bi-directional

UFM Server High Availability (HA) Active—Standby Communication

UFM Server HA Active—Standby Communication



UFM Server HA Active—Standby Communication

UFM Active—Standby communication enables two services: heartbeat and DRBD.

- *heartbeat* is used for auto-negotiation and keep-alive messaging between active and standby servers. *heartbeat* uses port 694 (udp).
- DRBD is used for low-level data (disk) synchronization between active and standby servers. DRBD uses port 8888 (tcp).

Affected Service	Network	Address / Service / Port	Direction
UFM HA heartbeat	Out-of-band management*	UDP / 694	Bi-directional
UFM HA DRBD	Out-of-band management*	TCP / 8888	Bi-directional

*An IPoIB network can be used for HA, but this is not recommended, since any InfiniBand failure might cause split brain and lack of synchronization between the active and standby servers.

UFM Server Health Monitoring

The UFM Server Health Monitoring module is a standalone module that monitors UFM resources and processes according to the settings in the */opt/ufm/files/conf/UFMHealthConfiguration.xml* file.

For example:

- Each monitored resource or process has its own failure condition (number of retries and/or timeout), which you can configure.
- If a test fails, UFM will perform a *corrective operation*, if defined for the process, for example, to restart the process. You can change the configured corrective operation. If the corrective operation is set to "None", after the defined number of failures, the *give-up* operation is performed.
- If a test reaches the configured threshold for the number of retries, the health monitoring initiates the *give-up* operation defined for the process, for example, UFM failover or stop.
- By default, events and alarms are sent when a process fails, and they are also recorded in the internal log file.

Each process runs according to its own defined schedule, which you can change in the configuration file.

Changes to the configuration file take effect only after a UFM Server restart. (It is possible to kill and run in background the process nohup python /opt/ufm/ufmhealth/UfmHealthRunner.pyo &.)

You can also use the configuration file to improve disk space management by configuring:

- How often to purge MySQL binary log files.
- When to delete compressed UFM log files (according to free disk space).

The settings in the */opt/ufm/files/conf/UFMHealthConfiguration.xml* file are also used to generate the UFM Health Report.

The following section describes the configuration file options for UFM server monitoring.

UFM Health Configuration

The UFM health configuration file contains three sections:

- Supported Operations—This section describes all the operations that can be used in tests, and their parameters.
- Supported Tests—This section describes all the tests. Each test includes:
    - The main test operation.

- A corrective operation, if the main operation fails.
- A give-up operation, if the main operation continues to fail after the corrective operation and defined number of retries.

The number of retries and timeout is also configured for each test operation.

- Test Schedule - This section lists the tests in the order in which they are performed and their configured frequency.

The following table describes the default settings in the */opt/ufm/files/conf/UFMHealthConfiguration.xml* file for each test. The tests are listed in the order in which they are performed in the default configuration file.

You might need to modify the default values depending on the size of your fabric.

For example, in a large fabric, the SM might not be responsive for *sminfo* for a long time; therefore, it is recommended to increase the values for timeout and number of retries for **SMResponseTest**.

Recommended configurations for *SMResponseTest* are:

- For a fabric with 5000 nodes:
  - Number of retries = 12
  - Frequency = 10
- For a fabric with 10000 nodes:
  - Number of retries = 12
  - Frequency = 20

Test Name / Description	Test Operation	Corrective Operation (if Test Operation fails)	No. Retries / Give-up Operation	Test Frequency
CpuUsageTest Checks total CPU utilization.	CPUTest Tests that overall CPU usage does not exceed 80% (this percentage is configurable).	None If UFM Event Burst Management is enabled, it is automatically initiated when the test operation fails	1 Retry None	1 minute
AvailableDiskSpaceTest Checks available disk space.	FreeDiskTest Tests that disk space usage for */opt/ufm* does not exceed 90% (this percentage is configurable).	CleanDisk Delete compressed UFM log files under */opt/ufm*	3 Retries None	1 hour

Test Name / Description	Test Operation	Corrective Operation (if Test Operation fails)	No. Retries / Give-up Operation	Test Frequency
CheckIBFabricInterface Checks state of active fabric interface.	IBInterfaceTest Tests that active fabric interface is up.	BringUpIBFabricInterface Bring up the fabric interface	3 Retries SMOrUFMFailoverOrDoNothing	35 seconds
CheckIBFabricInterfaceStandby (HA only) Checks state of fabric interface on standby.	IBInterfaceTestOnStandby Tests that fabric interface on standby is up.	None	1 Retry None	1 minute
MemoryTest Checks total memory usage.	MemoryUsageTest Tests that memory usage does not exceed 90% (this percentage is configurable).	None	1 Retry None	1 minute
SMProcessTest Checks status of the OpenSM service.	SMRunningTest Tests that the SM process is running.	RestartProcess Restart the SM process	1 Retry UFMFailoverOrDoNothing	10 seconds
SMResponseTest Checks responsiveness of SM (when SM process is running).	SMTest Tests SM responsiveness by sending the sminfo query to SM.	None	9 Retries UFMFailoverOrDoNothing	10 seconds
IbpmTest Checks status of the IBPM (Performance Manager) service.	ProcessIsRunningTest Tests that the IBPM service is running.	RestartProcess Restart the IBPM service	3 Retries None	1 minute
ModelMainTest Checks status of the main UFM service	ProcessIsRunningTest Tests that the UFM service is running.	RestartProcess Restart the UFM service	3 Retries UFMFailoverOrDoNothing	20 seconds
HttpdTest Checks status of the httpd service.	ProcessIsRunningTest Tests that the httpd service is running.	RestartProcess Restart the httpd service	3 Retries None	20 seconds

Test Name / Description	Test Operation	Corrective Operation (if Test Operation fails)	No. Retries / Give-up Operation	Test Frequency
MySqlTest Checks status of the MySql service.	ConnectToMySql Tests that the MySql service is running.	None	1 Retry UFMFailoverOrDoNothing	20 seconds
CleanMySql Purges MySql Logs	AlwaysFailTest Fails the test in order to perform the corrective action.	PurgeMySqlLogs Purge all MySql Logs on each test	1 Retry None	24 hours
UFMServerVersionTest Checks UFM software version and build.	UfmVersionTest Returns UFM software version information.	None	1 Retry None	24 hours
UFMServerLicenseTest Checks UFM License information.	UfmLicenseTest Returns UFM License information.	None	1 Retry None	24 hours
UFMServerHAConfigurationTest (HA only) Checks the configuration on master and standby.	UfmHAConfigurationTest Returns information about the master and standby UFM servers.	None	1 Retry None	24 hours
UFMMemoryTest Checks available UFM memory.	UfmMemoryUsageTest Tests that UFM memory usage does not exceed 80% (this percentage is configurable).	None	1 Retry None	1 minute
UFMCpuUsageTest Checks UFM CPU utilization.	CPUTest Tests that UFM CPU usage does not exceed 60% (this percentage is configurable).	None	1 Retry None	1 minute

Test Name / Description	Test Operation	Corrective Operation (if Test Operation fails)	No. Retries / Give-up Operation	Test Frequency
CheckDrbdTcpConnectionPerformanceTest (HA only) Checks the tcp connection between master and standby	TcpConnectionPerformanceTest Tests that bandwidth is greater than 100 Mb/sec and latency is less than 70 usec (configurable).	None	2 Retry None	10 minute

⚠ The Supported Operations section of the configuration file includes additional optional operations that can be used as corrective operations or give-up operations.

UFM Core Files Tracking

To receive a notification every time OpenSM or ibpm creates a core dump, please refer to the list of all current core dumps of OpenSM and ibpm in the UFM health report.

To receive core dump notifications, do the following:

1. Set the core_dumps_directory field in the gv.cfg file to point to the location where all core dumps are created (by default, this location is set to /tmp).

   ```
   core_dumps_directory = /tmp
   ```

2. Set the naming convention for the core dump file. The name must include the directory configured in the step above.
   The convention we recommend is:

   ```
   echo "/tmp/%t.core.%e.%p.%h" > /proc/sys/kernel/core_pattern
   ```

3. Make sure core dumps directory setting is persistent between reboots. Add the kernel.core_pattern parameter with the desired file name format to the /etc/systctl.conf file. Example:

```
kernel.core_pattern=/tmp/%t.core.%e.%p.%h
```

4. Configure the core file size to be unlimited.

```
ulimit -c unlimited
```

5. (Only on UFM HA master) Update the UFM configuration file gv.cfg to enable core dump tracking.

```
track_core_dumps = yes
```

Example of Health Configuration

The default configuration for the overall memory test in the *opt/ufm/files/conf/UFMHealthConfiguration.xml* file is:

```
<Test Name="MemoryTest" NumOfRetriesBeforeGiveup="3" RetryTimeoutInSeconds="10">
    <TestOperation Name="MemoryUsageTest">
        <Parameters>
            <Parameter Name="ThresholdInPercents" Value="90"/>
        </Parameters>
    </TestOperation>
    <CorrectiveOperation Name="None"/>
    <GiveupOperation Name="None"/>
</Test>
```

This configuration tests the available memory. If memory usage exceeds 90%, the test is repeated up to 3 times at 10 second intervals, or until memory usage drops to below 90%. No corrective action is taken and no action is taken after 3 retries.

To test with a usage threshold of 80%, and to initiate UFM failover or stop UFM after three retries, change the configuration to:

```
<Test Name="MemoryTest" NumOfRetriesBeforeGiveup="3" RetryTimeoutInSeconds="10">
    <TestOperation Name="MemoryUsageTest">
        <Parameters>
            <Parameter Name="ThresholdInPercents" Value="80"/>
```

```
        </Parameters>
    </TestOperation>
    <CorrectiveOperation Name="None"/>
    <GiveupOperation Name="UFMFailoverOrStop"/>
</Test>
```

Event Burst Management

UFM event burst management can lower the overall CPU usage following an event burst by suppressing events. Event burst management is configured in the *gv.cfg* configuration file.

When the overall CPU usage exceeds the threshold configured by the CpuUsageTest in the */opt/ufm/files/conf/UFMHealthConfiguration.xml* file, a High CPU Utilization event occurs.

This event initiates the UFM event burst management, which:

- Suppresses events. The default level of suppression enables critical events only.
- If, after a specified period of time (30 seconds, by default), no further High CPU Utilization event occurs, the UFM server enables all events.

To modify Event burst management configuration, change the following parameters in the gv. cfg file:

```
# The events' level in case events are suppressed (the possible levels are disable_all_events,
enable_critical_events, and enable_all_events)
# The entire feature can be turned off using the level "enable_all_events"
suppress_events_level = enable_critical_events
# The amount of time in seconds which events are suppressed
suppress_events_timeout = 30
```

Recovery from Consecutive Failures

UFM Server Health Monitor might restart or trigger a failover in order to recover from specific failures. In case a re-start or failover fails, UFM Server Health Monitor tries the operation again. Upon a number of consecutive failure attempts to restart or failover, UFM Server Health Monitor stops trying to restart Model Main and allows OpenSM to run without intervention. The behavior maximum number of consecutive restart attempts is defined in the configuration file /opt/ufm/files/conf/UFMHealthConfiguration.xml:

```
<Parameter Name="RestartAttempts" Value="8"/>
<Failover MaxAllowedAttempts="6"/>
```

Cable Transceiver Temperatures

The UFM has alarms that notify the user in cases where an active cable overheats/overcools.

The UFM uses ibdiagnet to get cable temperature analysis and report exceptions via the Alarms view.

Related events:

- 919 for high cable temperature
- 920 for low cable temperature

GUI Views

Alarms

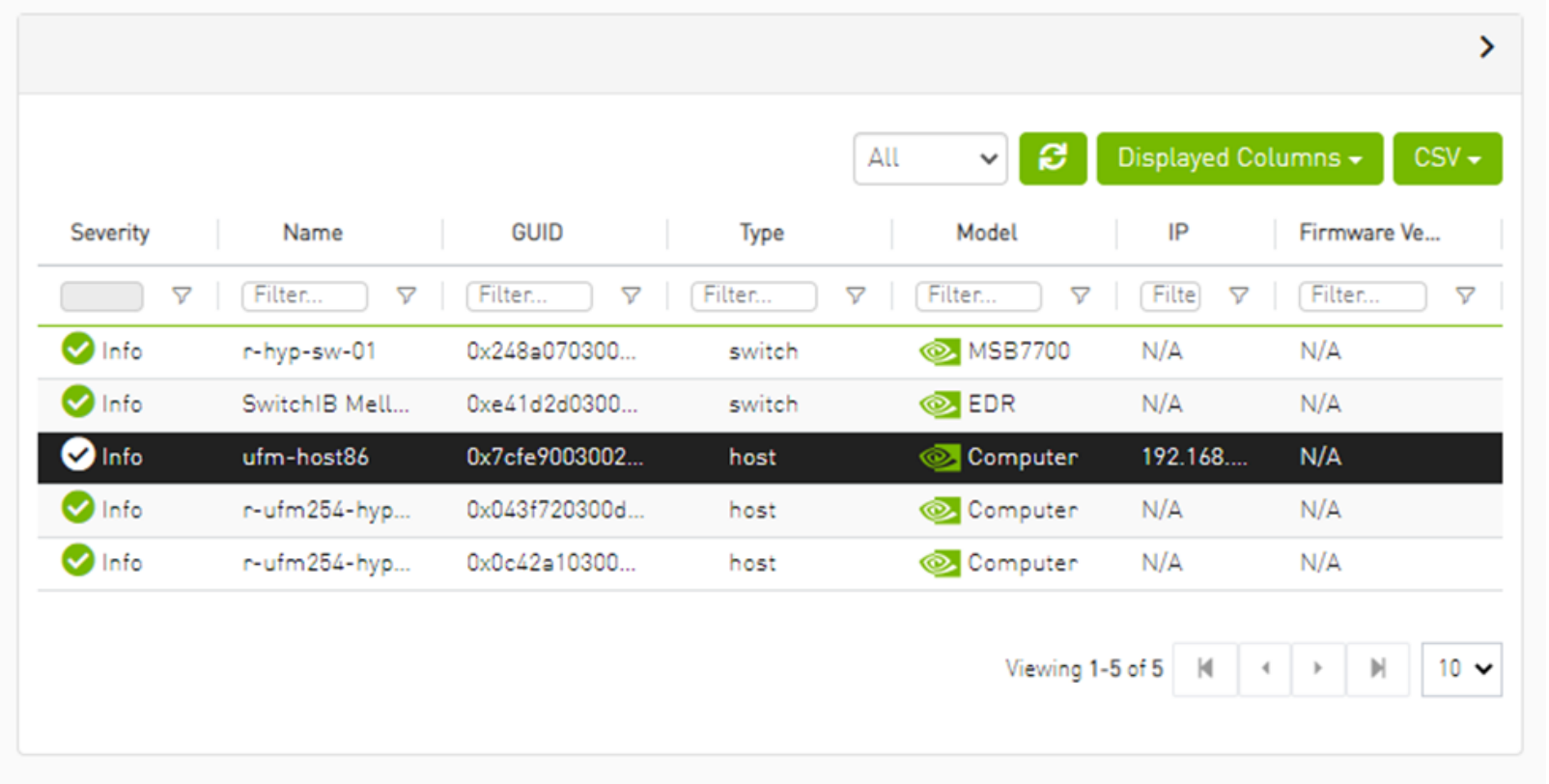
Severity	Date/Time	Alarm Name	Source	Sourc...	Reason	Count
	Filter...	Filter...	Filter...		cable	Filter
Critical	2022-03-12 23:25:09	Cable Temperature High	default[3] / Switch: r-hyp-sw-l	IBPort	Cable High Temperature Alarm reported- current temperature: 116C- threshold: 70C	1
Critical	2022-03-12 23:25:09	Cable Temperature Low	default[3] / Computer: r-ufm2	IBPort	Cable Low Temperature Alarm reported- current temperature: 50C- threshold: 90C	1

Event Policy

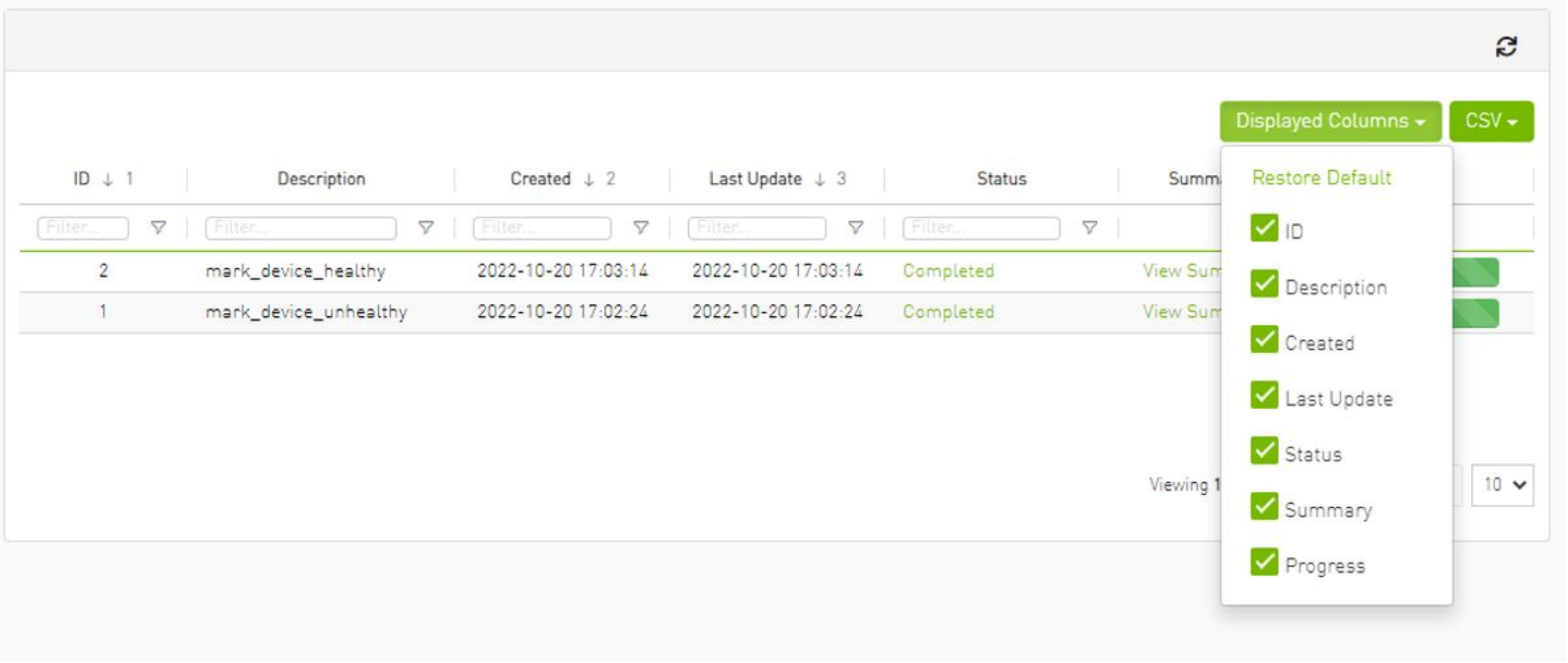
Event	Category	Mail	GUI	Alarm	Syslog	Log File	SNMP	Threshold	TTL[Sec]	Severity
cable temp								Filter...	Filter...	
Cable Temperature High									0	Critical
Cable Temperature Low									0	Critical

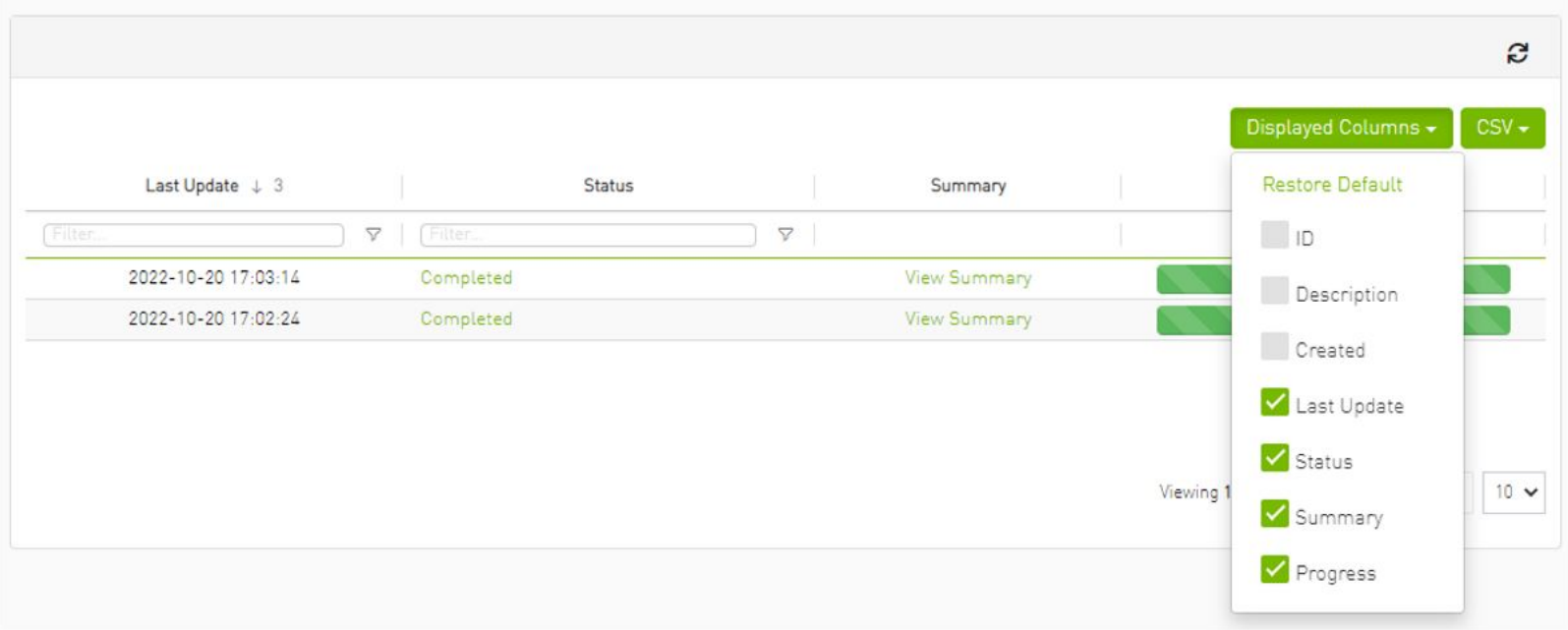
Table Enhancements

Look and Feel Improvements



Displayed Columns



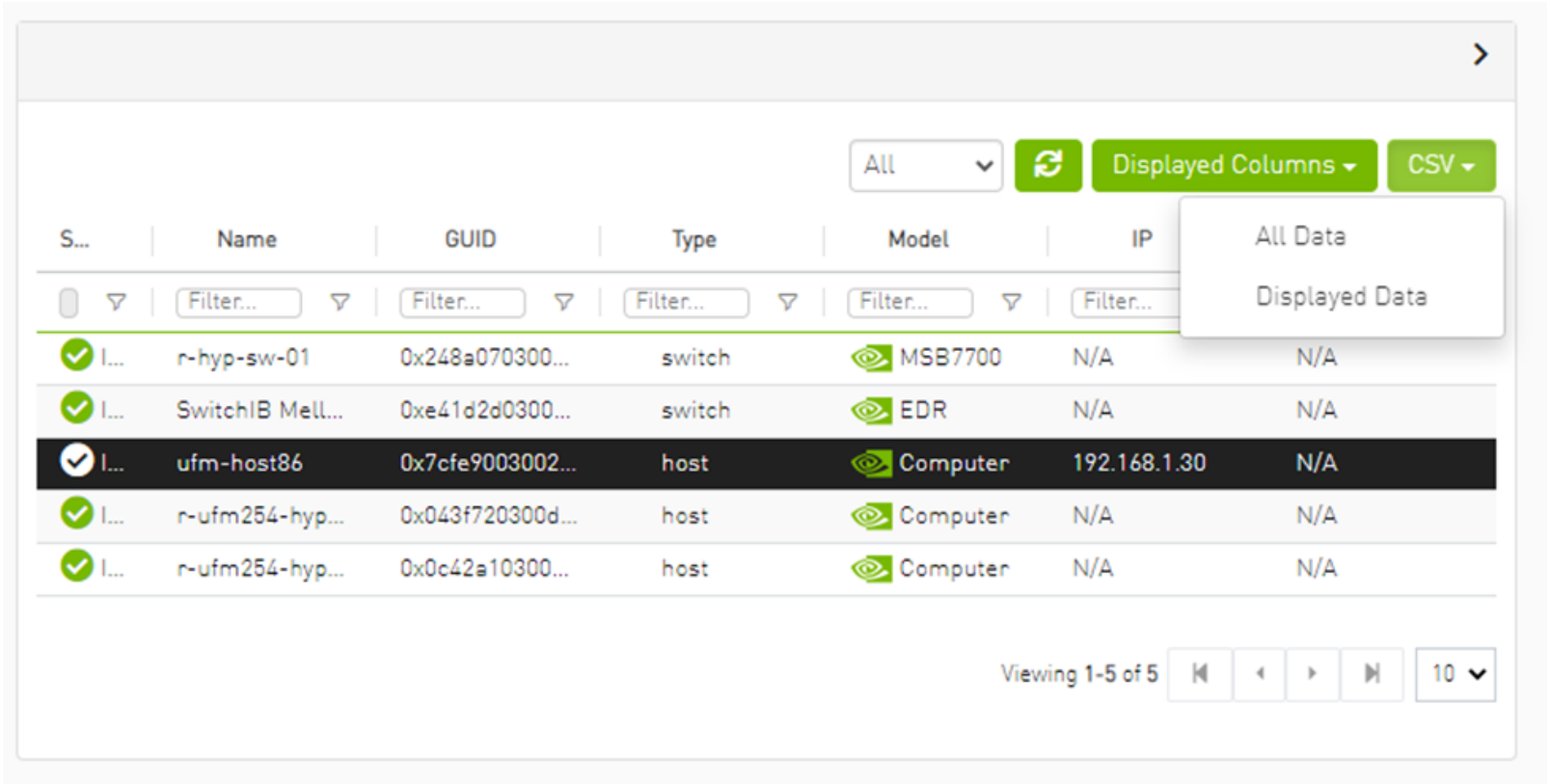


⚠ Displayed columns of all tables are persistent per user, with the option to restore defaults.

Export All Data as CSV

There are two options for exporting as CSV

- All Data: all data returned from server.
- Displayed Data: only displayed rows.

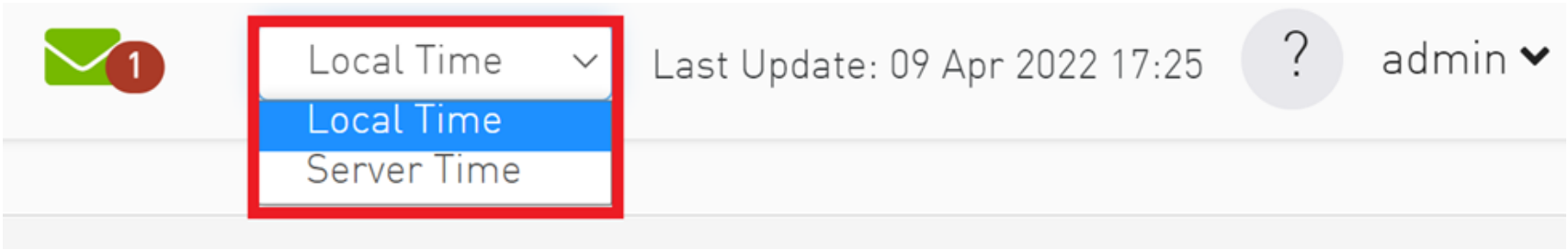

All
Displayed Columns
CSV
All Data
Displayed Data
S...
Name
GUID
Type
Model
IP
Filter...
I...
r-hyp-sw-01
0x248a070300...
switch
MSB7700
N/A
N/A
I...
SwitchIB Mell...
0xe41d2d0300...
switch
EDR
N/A
N/A
I...
ufm-host86
0x7cfe9003002...
host
Computer
192.168.1.30
N/A
I...
r-ufm254-hyp...
0x043f720300d...
host
Computer
N/A
N/A
I...
r-ufm254-hyp...
0x0c42a10300...
host
Computer
N/A
N/A
Viewing 1-5 of 5
10

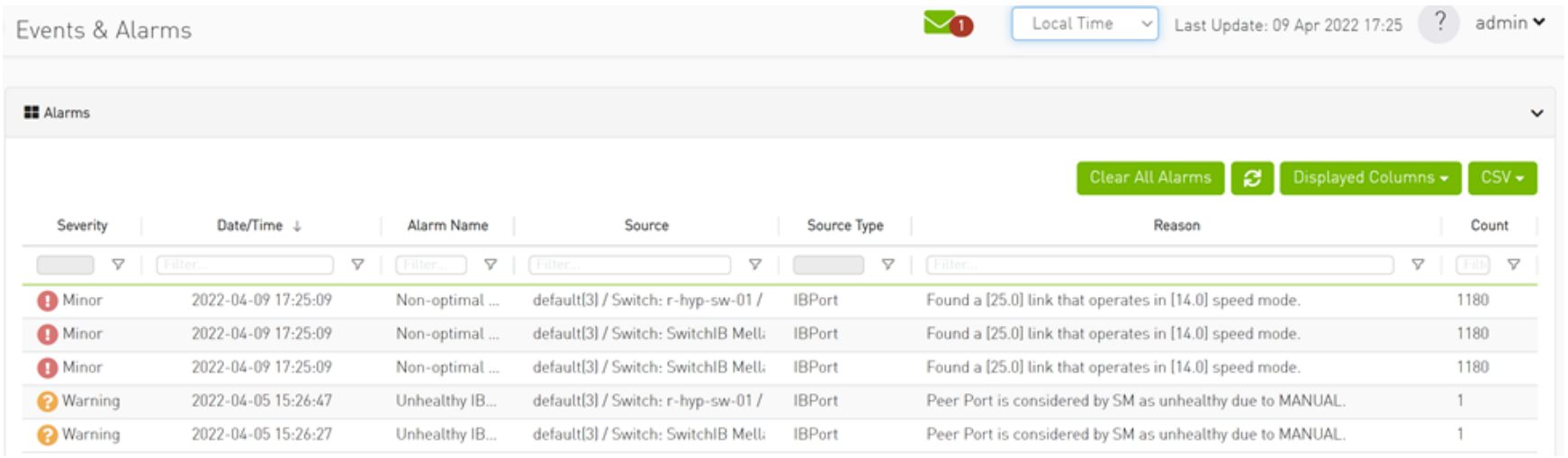
Time Zone Converter

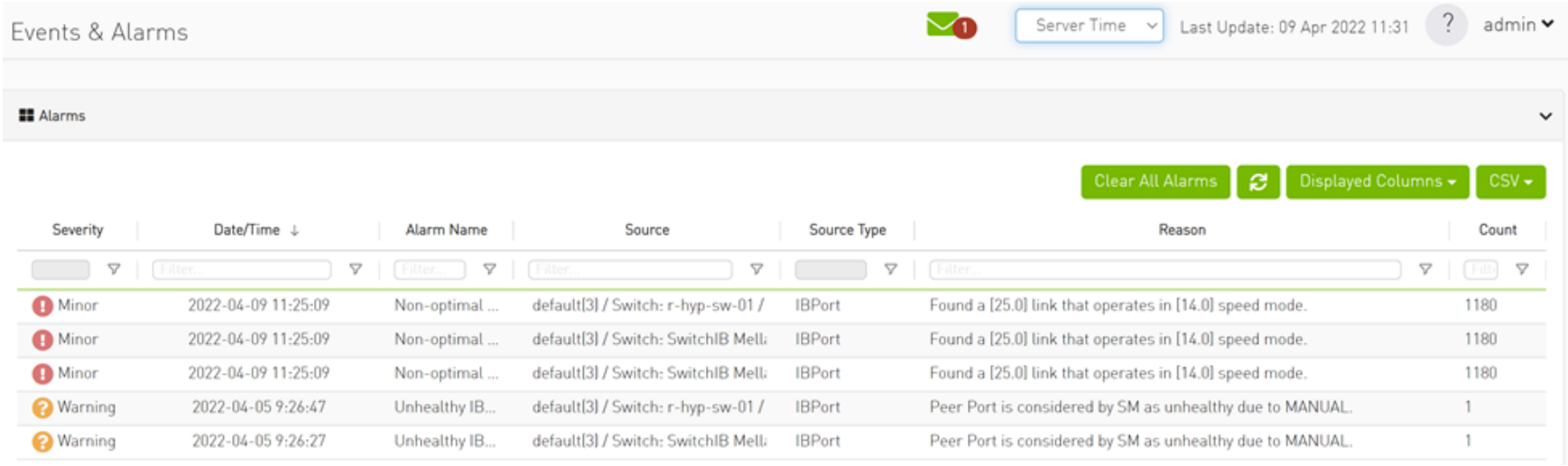
Time zone converter provides the ability to unify all times in UFM like events and alarms, ibdiagnet, telemetry and logs.

The user can switch between local and machine time.

There is a drop-down menu in the status bar to switch between local and server/machine time.







⚠ In the screenshots, the difference between Server Time and Local Time is 6 hours.

Multi-Subnet UFM

Overview

The Multi-Subnet UFM feature allows for the management of large fabrics, consisting of multiple sites, within a single product, namely Multi-Subnet UFM.

This feature is comprised of two layers: UFM Multi-Subnet Provider and UFM Multi-Subnet Consumer.

The UFM Provider functions as a Multi-Subnet Provider, exposing all local InfiniBand fabric information to the UFM consumer. On the other hand, the UFM Consumer acts as a Multi-Subnet Consumer, collecting and aggregating data from currently configured UFM Providers, enabling users to manage multiple sites in one place. While UFM Consumer offers similar functionality to regular UFM, there are several behavioral differences related to aggregation.

Setting Up Multi-Subnet UFM

In `/opt/ufm/files/conf/gv.cfg`, fill in the section named [`Multisubnet`] for UFM Multi-Subnet Provider and Consumer.

To set up UFM as a Multi-Subnet Provider, perform the following:

- Set `multisubnet_enabled` to `true`
- Set `multisubnet_role` to `provider`
- Set `multisubnet_site_name` (optional, if not set, it will be randomly generated); e.g., *`provider_1`*
- Start UFM

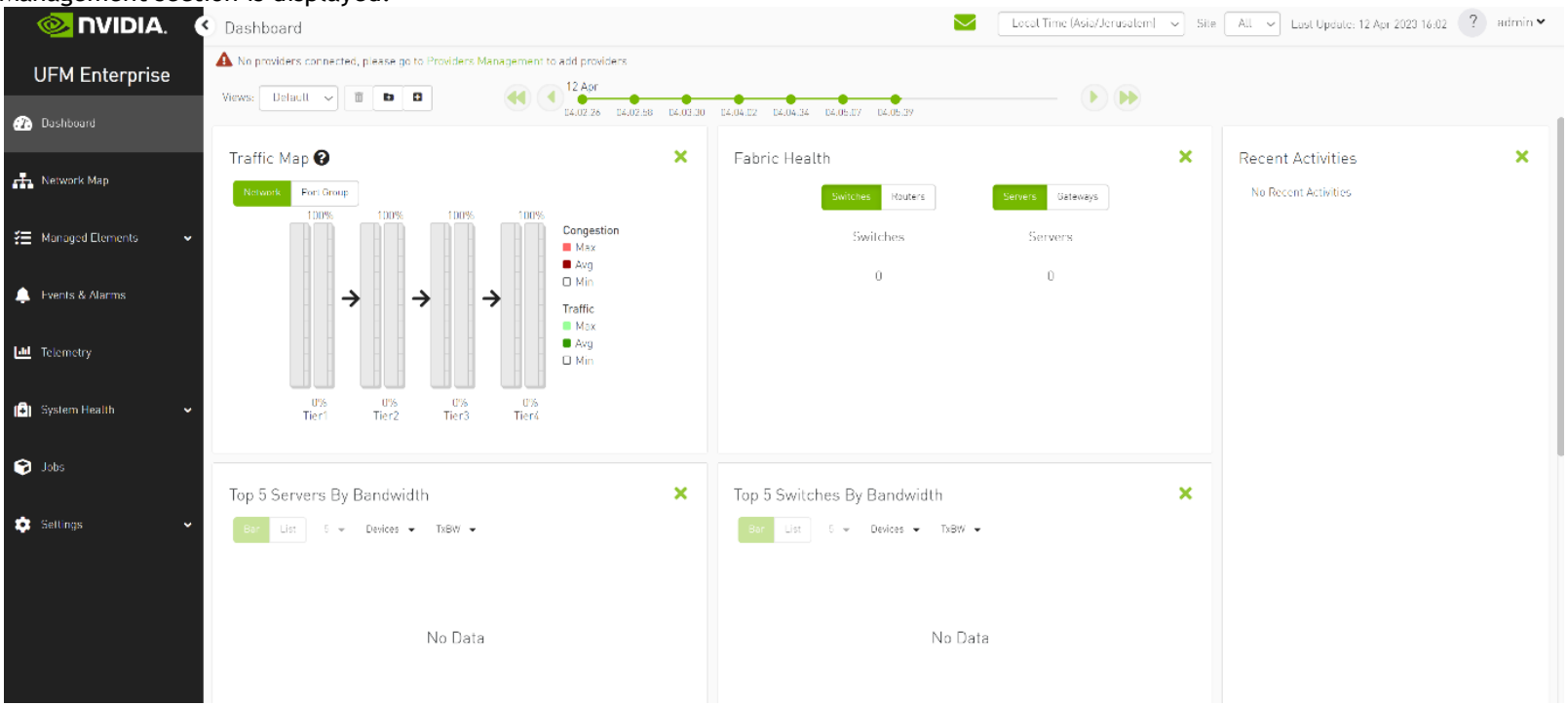
To set up UFM as a Multi-Subnet Consumer, perform the following:

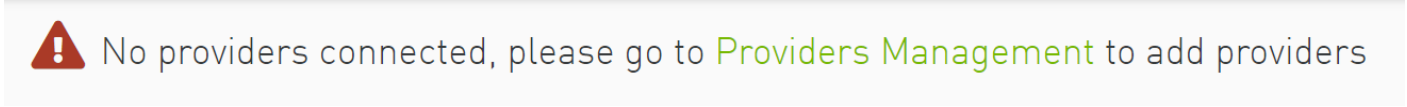
- Set `multisubnet_enabled` to `True`
- Set `multisubnet_role` to `consumer`
- Start UFM

It is important to note that UFM Multi-Subnet Consumer can be configured on a machine or VM without an established InfiniBand connectivity. Additionally, users may customize UFM Provider and Consumer using optional configuration parameters found in the [`Multisubnet`] section of `/opt/ufm/files/conf/gv.cfg`.

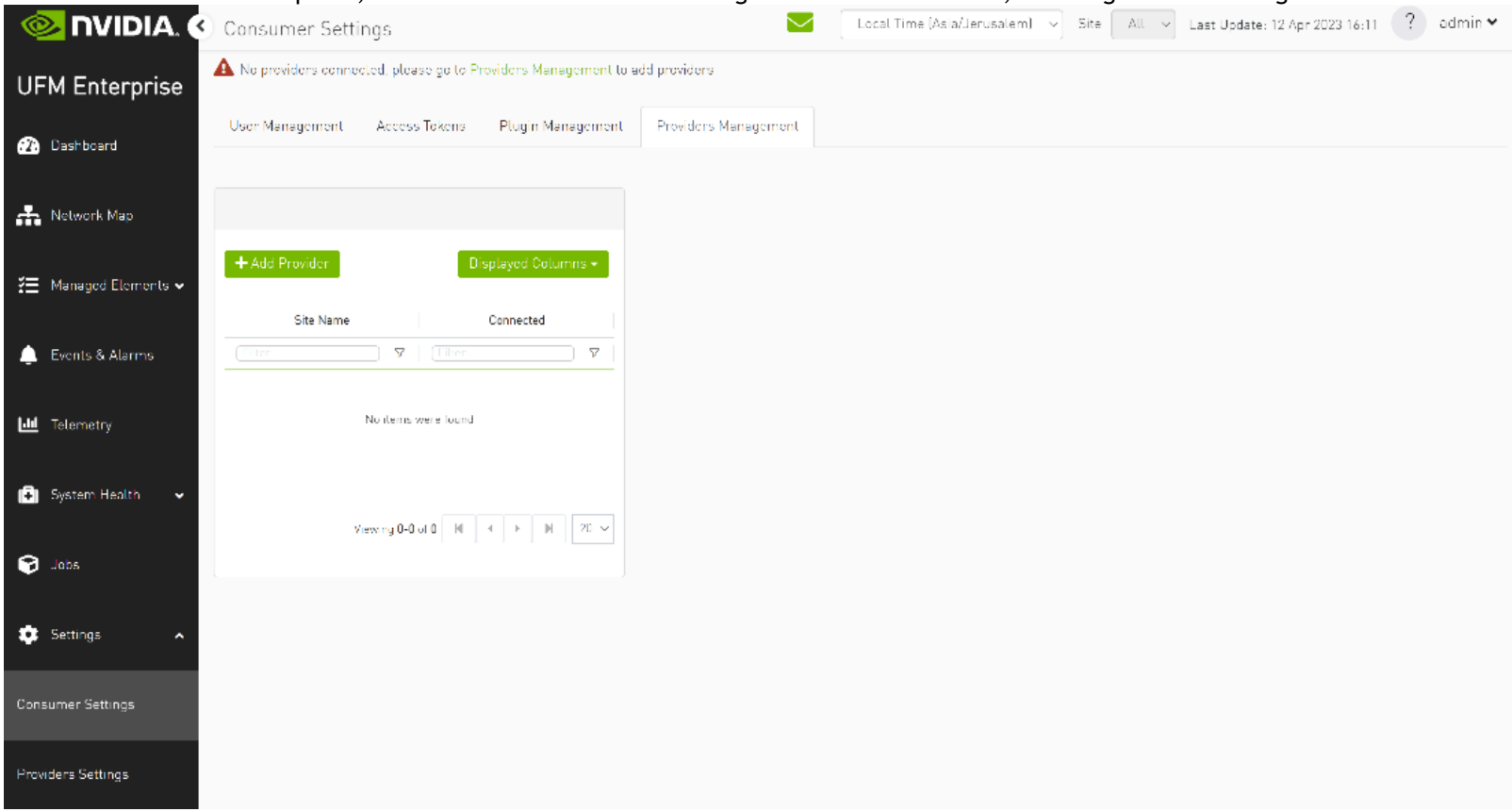
Functionality

1. Following the initial launch of the Consumer, the Dashboard view is devoid of data, and a message containing a hyperlink leading to the Provider Management section is displayed.

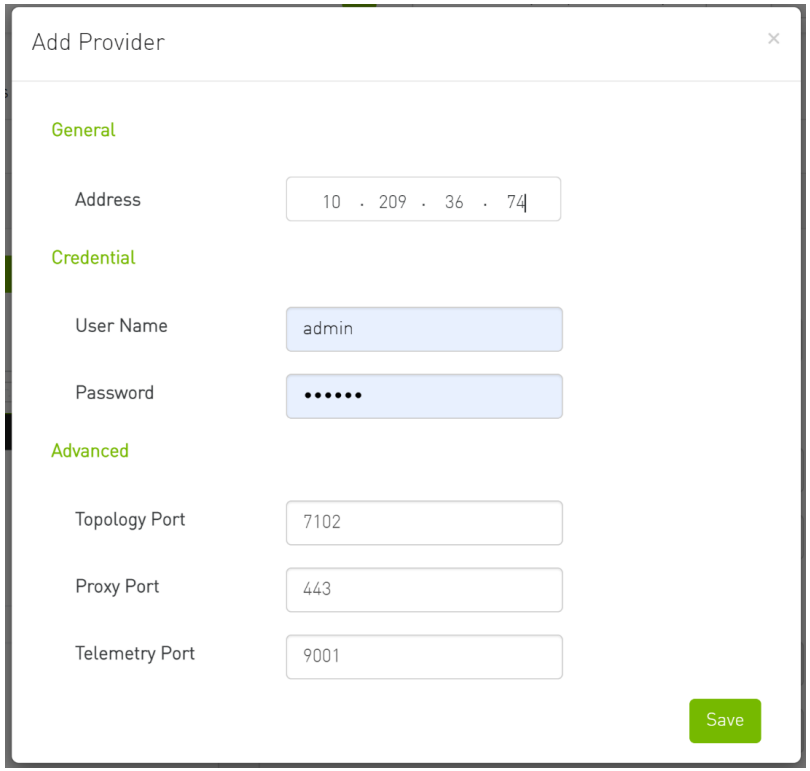




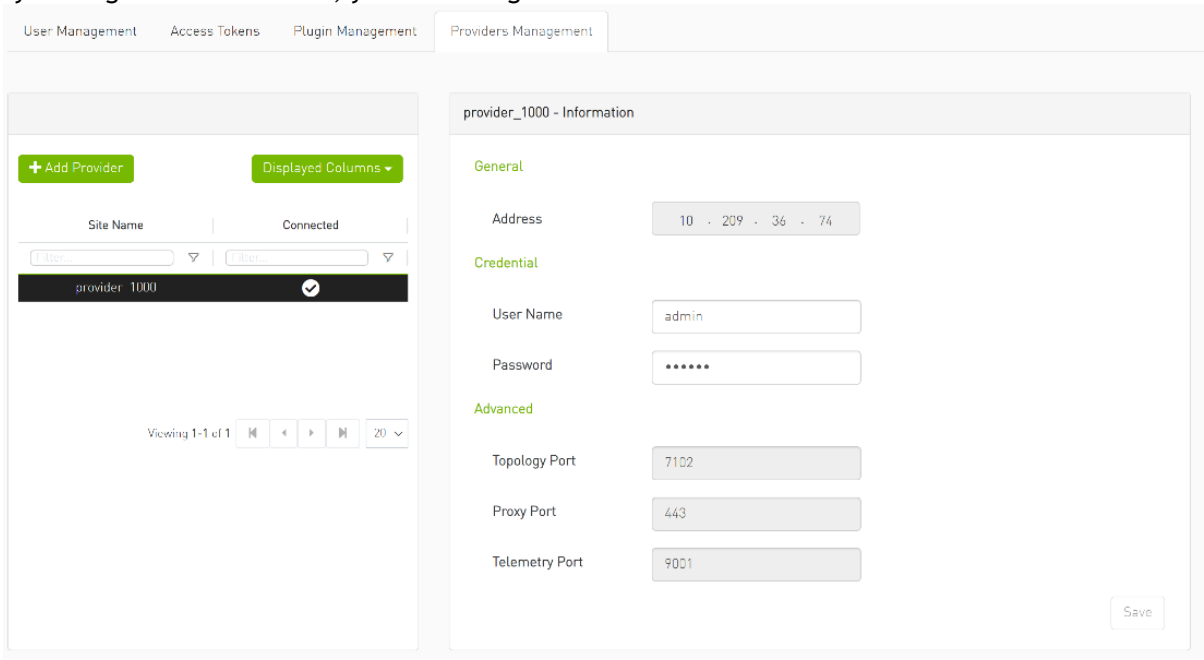
2. As shown in the below snapshot, a new section for Provider Management has been added, enabling users to configure UFM Providers.



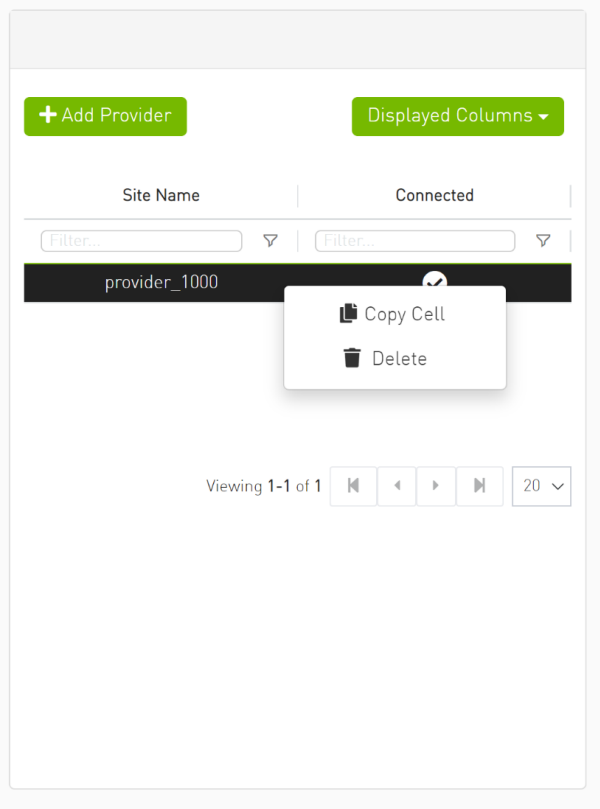
   a. To add a provider, the user is required to enter its IP address and credentials. Unless there are multiple instances of UFM providers on a single machine, the advanced section parameters should be set with default values. However, if there are multiple instances, the advanced parameters may be set per Provider and then be configured in the Providers Management view.


Add Provider
General
Address
10 . 209 . 36 . 74
Credential
User Name
admin
Password
Advanced
Topology Port
7102
Proxy Port
443
Telemetry Port
9001
Save

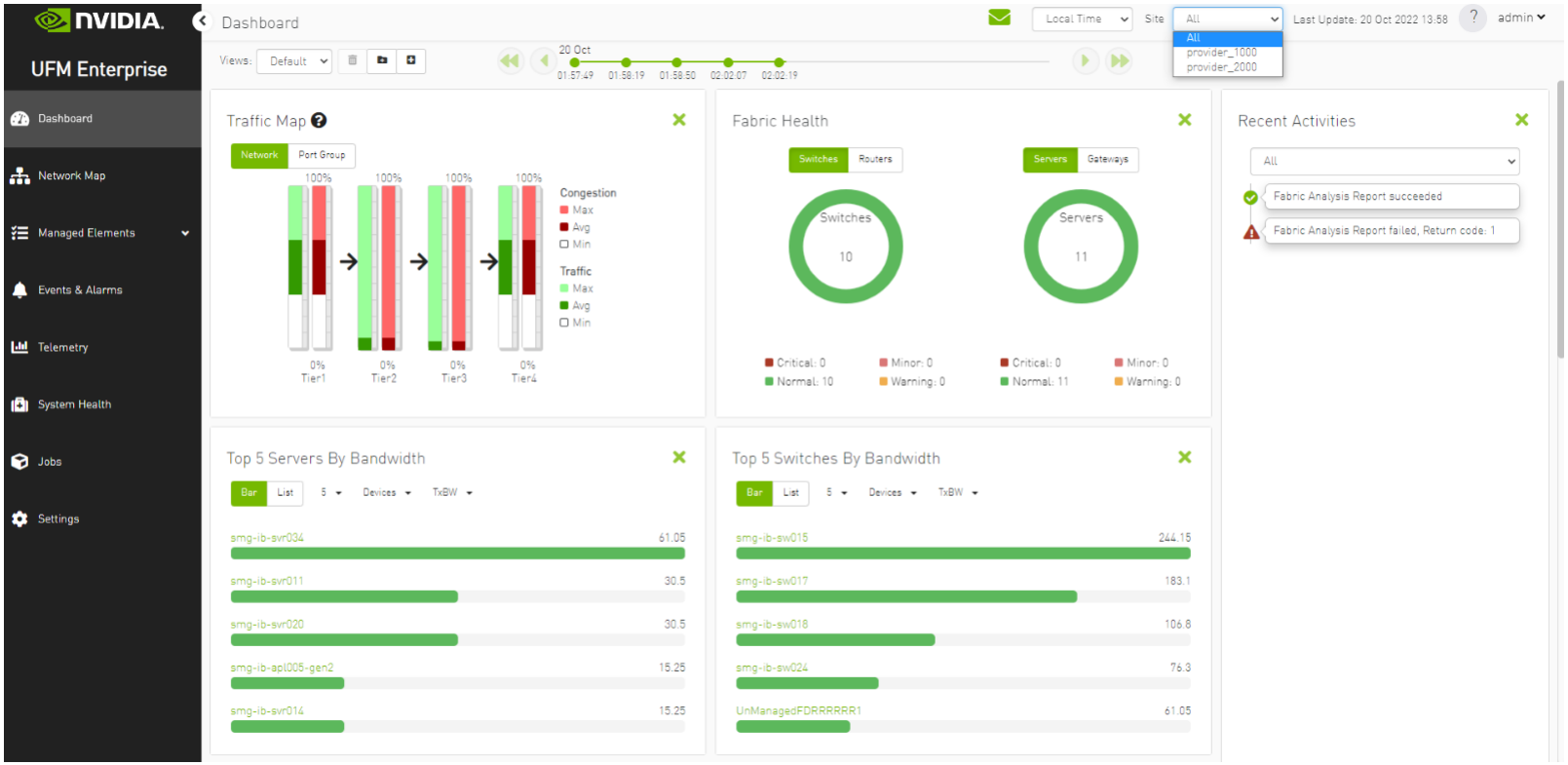
b. By editing the Provider view, you can change Provider's credentials.

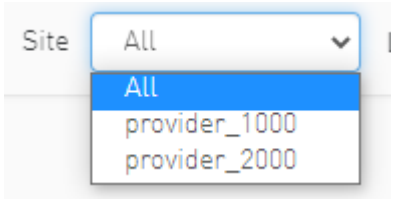


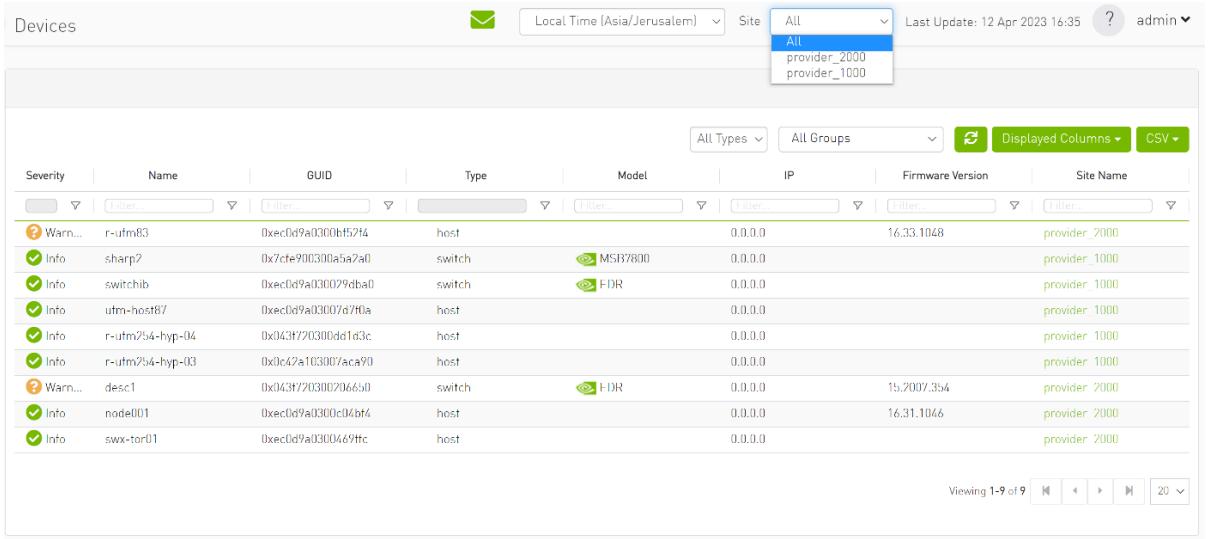
c. The "Delete Provider" function removes the selected Provider from the Consumer. Please note that this action may take some time to complete, and changes may only be reflected in the view after approximately 30 seconds.

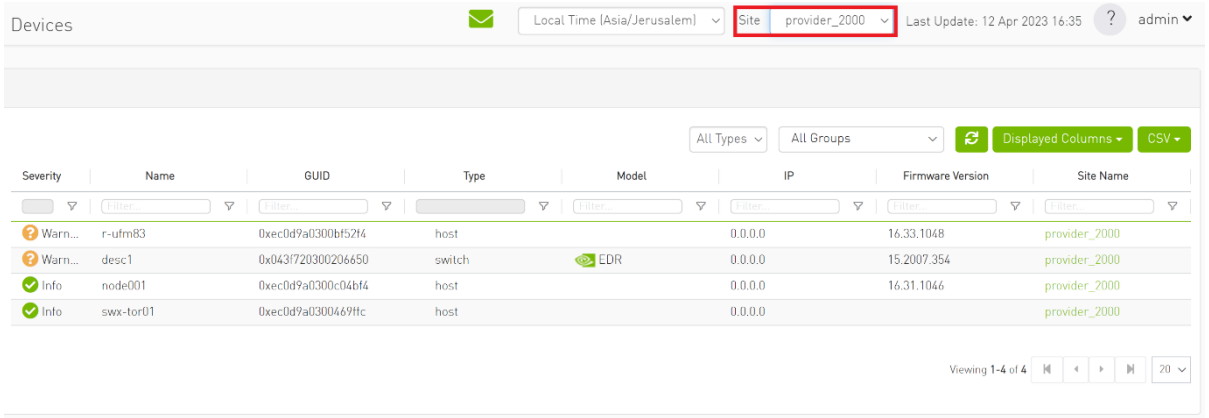


3. A general filter has been added to the top right corner of the page, enabling users to filter displayed data by site.

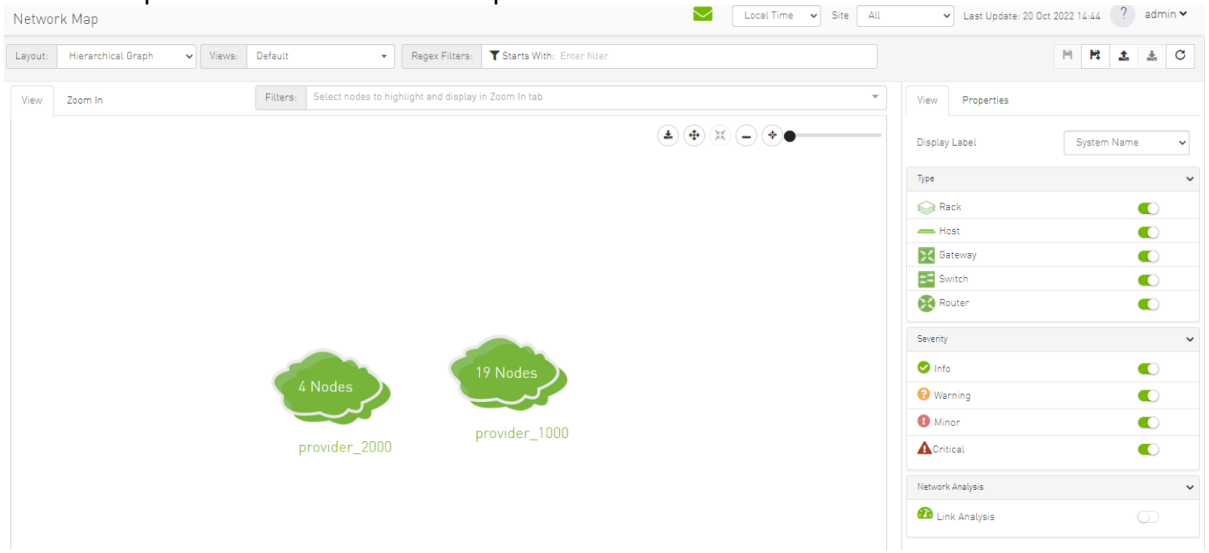

NVIDIA
UFM Enterprise
Dashboard
Network Map
Managed Elements
Events & Alarms
Telemetry
System Health
Jobs
Settings
Views: Default
20 Oct
01:57:49
01:58:19
01:58:50
02:02:07
02:02:19
Local Time
Site
All
provider_1000
provider_2000
Last Update: 20 Oct 2022 13:58
admin
Traffic Map
Network
Port Group
100%
0%
Tier1
Tier2
Tier3
Tier4
Congestion
Max
Avg
Min
Traffic
Fabric Health
Switches
Routers
Servers
Gateways
10
11
Critical: 0
Minor: 0
Normal: 10
Warning: 0
Critical: 0
Minor: 0
Normal: 11
Warning: 0
Recent Activities
All
Fabric Analysis Report succeeded
Fabric Analysis Report failed, Return code: 1
Top 5 Servers By Bandwidth
Bar
List
5
Devices
TxBW
smg-ib-svr034
61.05
smg-ib-svr011
30.5
smg-ib-svr020
30.5
smg-ib-apl005-gen2
15.25
smg-ib-svr014
15.25
Top 5 Switches By Bandwidth
smg-ib-sw015
244.15
smg-ib-sw017
183.1
smg-ib-sw018
106.8
smg-ib-sw024
76.3
UnManagedFDRRRRRR1
61.05

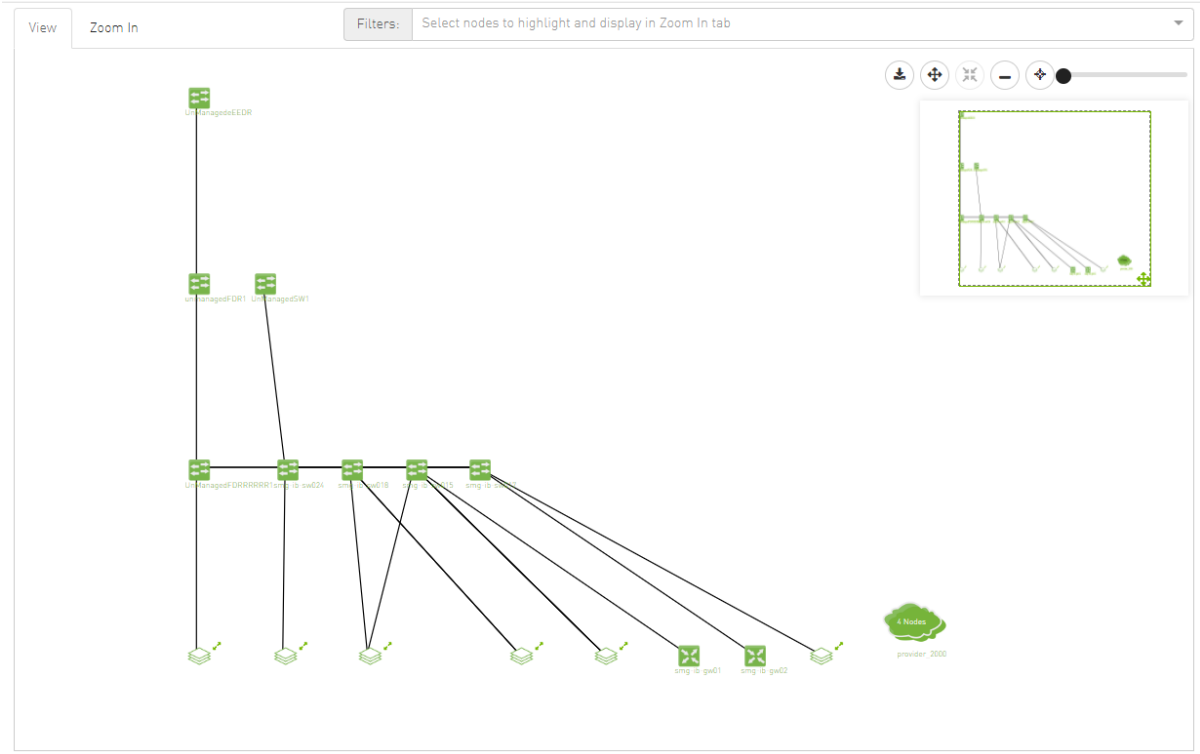

Site
All
All
provider_1000
provider_2000


Devices
Local Time (Asia/Jerusalem)
Site
All
Last Update: 12 Apr 2023 16:35
admin
All
provider_2000
provider_1000
All Types
All Groups
Displayed Columns
CSV
Severity
Name
GUID
Type
Model
IP
Firmware Version
Site Name
Warn... r-ufm83 0xec0d9a0300bf52f4 host 0.0.0.0 16.33.1048 provider_2000
Info sharp2 0x7cfe900300a5a2a0 switch MSB7800 0.0.0.0 provider_1000
Info switchib 0xec0d9a03002 9dba0 switch EDR 0.0.0.0 provider_1000
Info ufm-host87 0xec0d9a03007d7f0a host 0.0.0.0 provider_1000
Info r-ufm254-hyp-04 0x043f720300dd1d3c host 0.0.0.0 provider_1000
Info r-ufm254-hyp-03 0x0c42a103007aca90 host 0.0.0.0 provider_1000
Warn... desc1 0x043f720300206650 switch EDR 0.0.0.0 15.2007.354 provider_2000
Info node001 0xec0d9a0300c04bf4 host 0.0.0.0 16.31.1046 provider_2000
Info swx-tor01 0xec0d9a0300469ffc host 0.0.0.0 provider_2000
Viewing 1-9 of 9
20

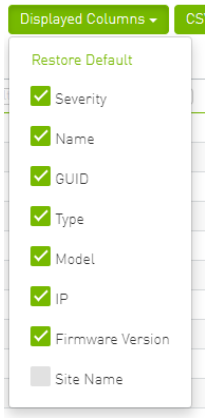

Devices
Local Time (Asia/Jerusalem)
Site
provider_2000
Last Update: 12 Apr 2023 16:35
admin
All Types
All Groups
Displayed Columns
CSV
Severity
Name
GUID
Type
Model
IP
Firmware Version
Site Name
Warn... r-ufm83 0xec0d9a0300bf52f4 host 0.0.0.0 16.33.1048 provider_2000
Warn... desc1 0x043f720300206650 switch EDR 0.0.0.0 15.2007.354 provider_2000
Info node001 0xec0d9a0300c04bf4 host 0.0.0.0 16.31.1046 provider_2000
Info swx-tor01 0xec0d9a0300469ffc host 0.0.0.0 provider_2000
Viewing 1-4 of 4
20

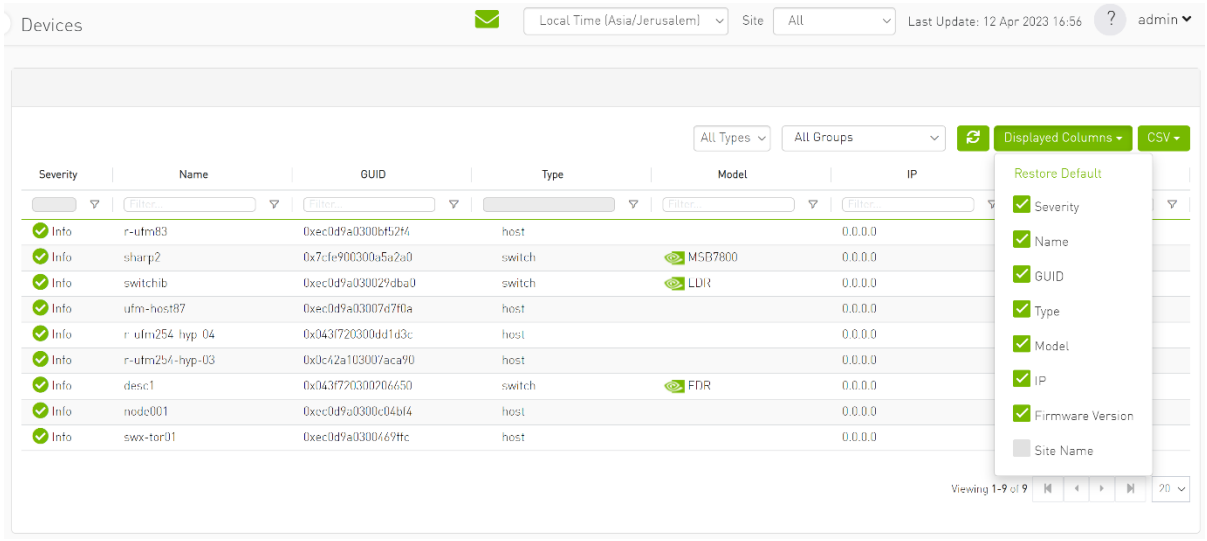
4. Network map contains “clouds” for each provider.

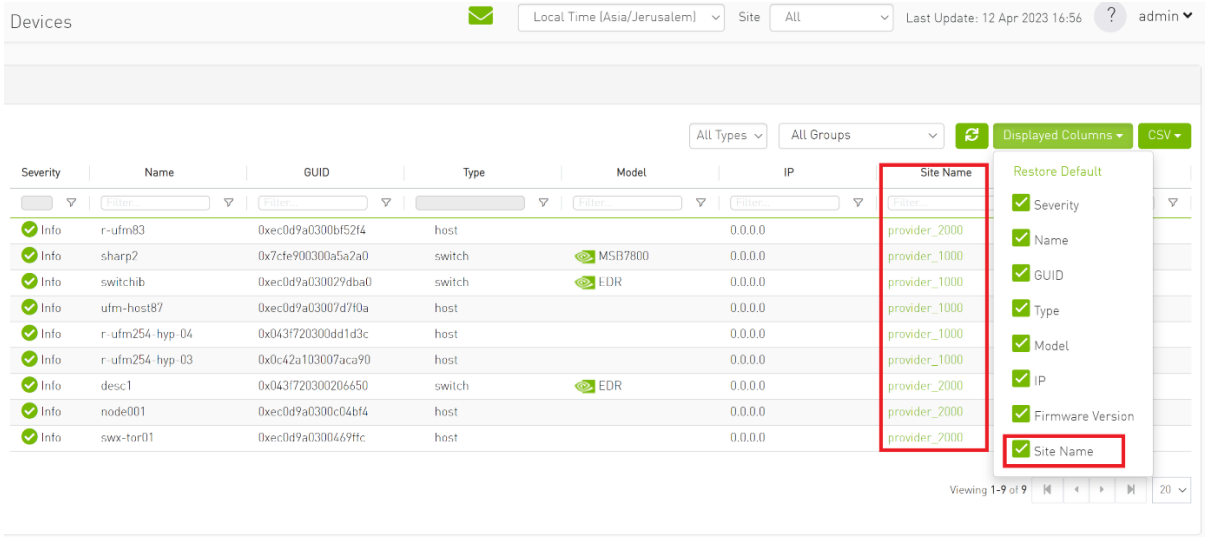



View
Zoom In
Filters:
Select nodes to highlight and display in Zoom In tab
UnManagedeEEDR
UnManagedFDR1
UnManagedSW1
UnManagedFDRRRRRR1
smg-ib-sw024
smg-ib-gw01
smg-ib-gw02
4 Nodes
provider 2000

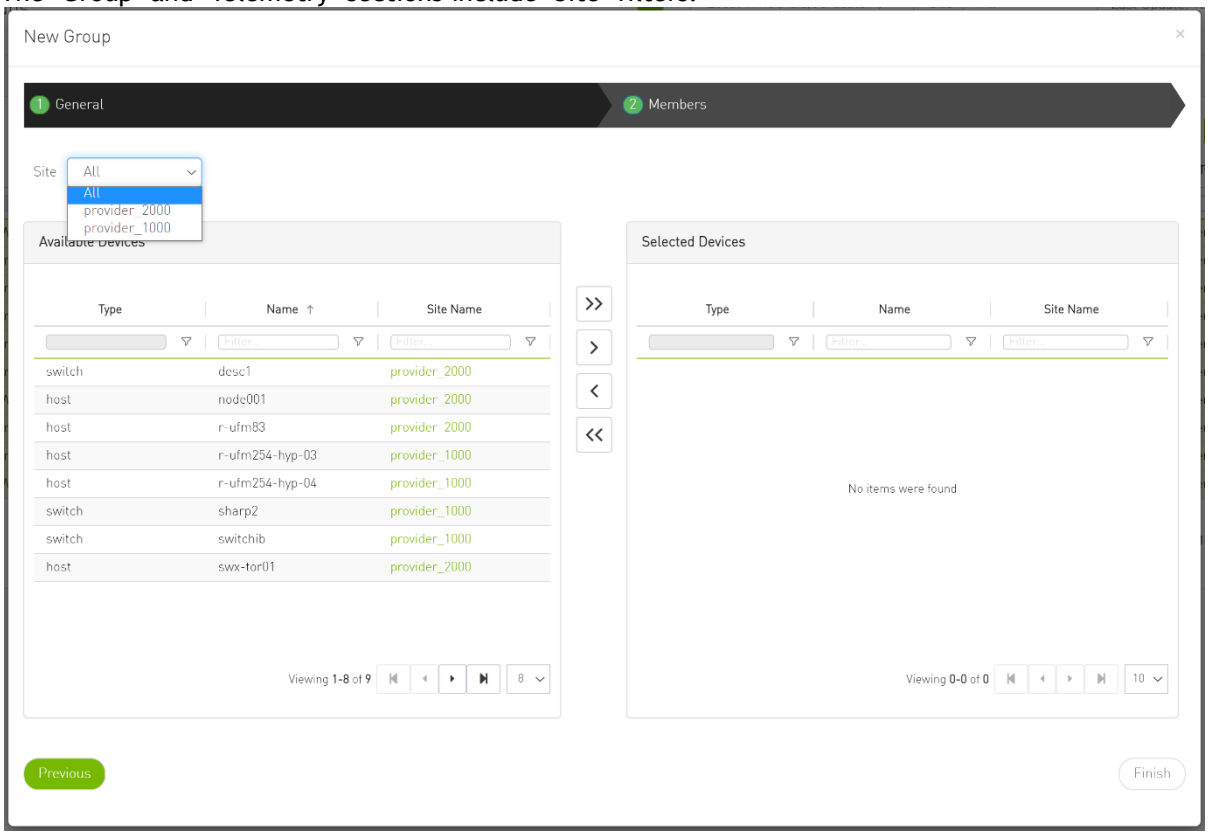
5. A "Site Name" column is present in all Managed Elements sections. The column is disabled (hidden) by default.



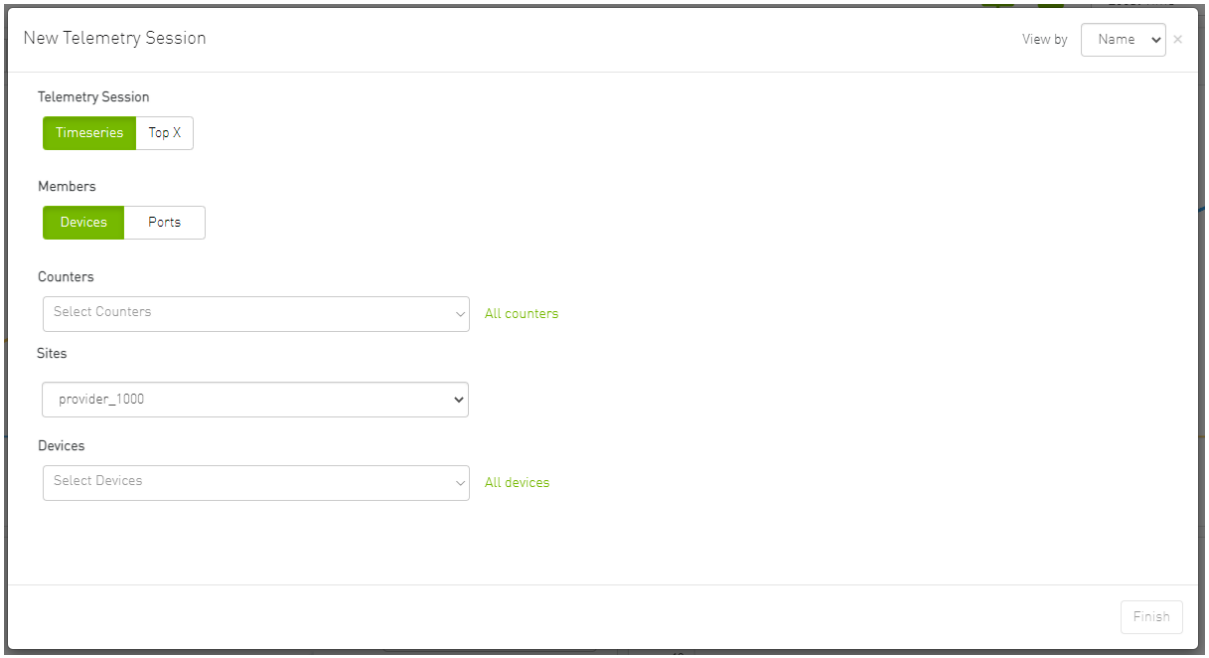


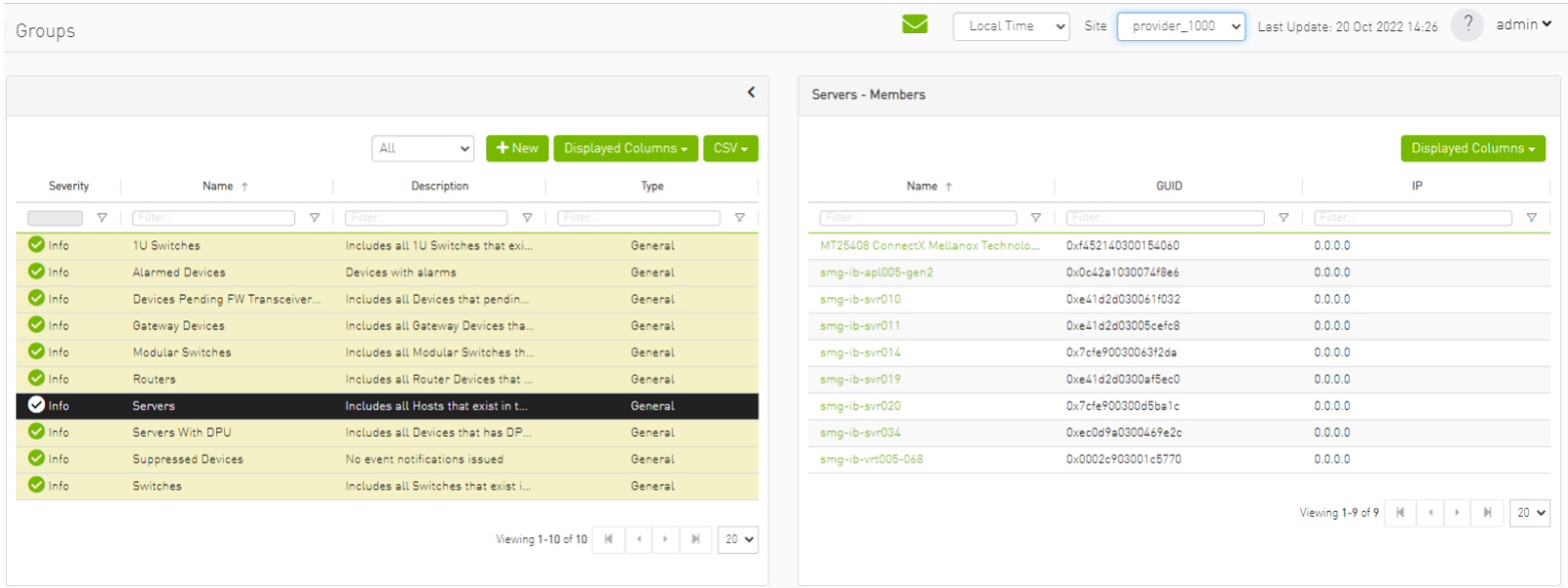

Devices
Local Time (Asia/Jerusalem)
Site
All
Last Update: 12 Apr 2023 16:56
admin
All Types
All Groups
Displayed Columns
CSV
Restore Default
Severity
Name
GUID
Type
Model
IP
Site Name
Firmware Version
Info r-ufm83 0xec0d9a0300bf52f4 host 0.0.0.0 provider_2000
Info sharp2 0x7cfe900300a5a2a0 switch MSB7800 0.0.0.0 provider_1000
Info switchib 0xec0d9a030029dba0 switch EDR 0.0.0.0 provider_1000
Info ufm-host87 0xec0d9a03007d7f0a host 0.0.0.0 provider_1000
Info r-ufm254-hyp-04 0x043f720300dd1d3c host 0.0.0.0 provider_1000
Info r-ufm254-hyp-03 0x0c42a103007aca90 host 0.0.0.0 provider_1000
Info desc1 0x043f720300206650 switch EDR 0.0.0.0 provider_2000
Info node001 0xec0d9a0300c04bf4 host 0.0.0.0 provider_2000
Info swx-tor01 0xec0d9a0300469ffc host 0.0.0.0 provider_2000
Viewing 1-9 of 9
20

6. The "Group" and "Telemetry" sections include "Site" filters.



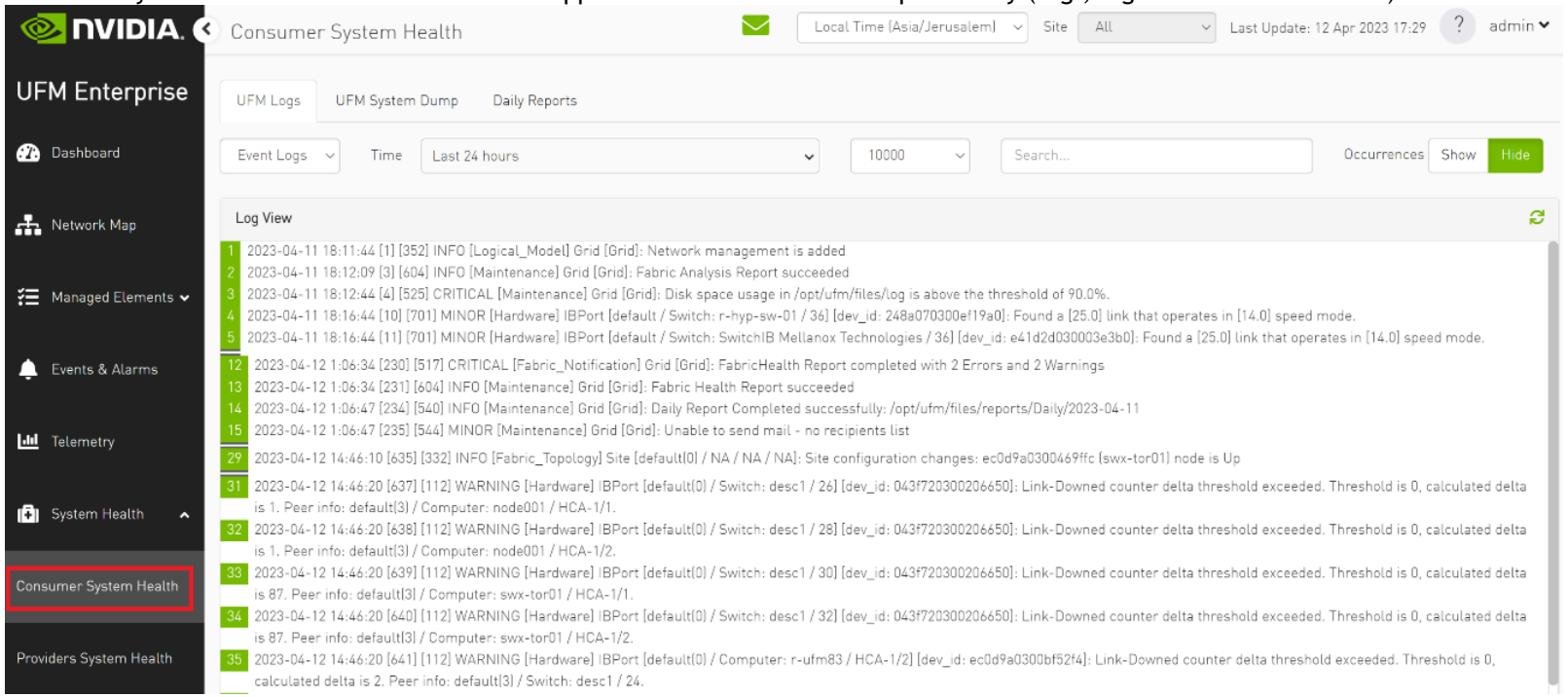
7. The filter in "Groups" impacts the Members table only.


New Telemetry Session
View by
Name
Telemetry Session
Timeseries
Top X
Members
Devices
Ports
Counters
Select Counters
All counters
Sites
provider_1000
Devices
Select Devices
All devices
Finish

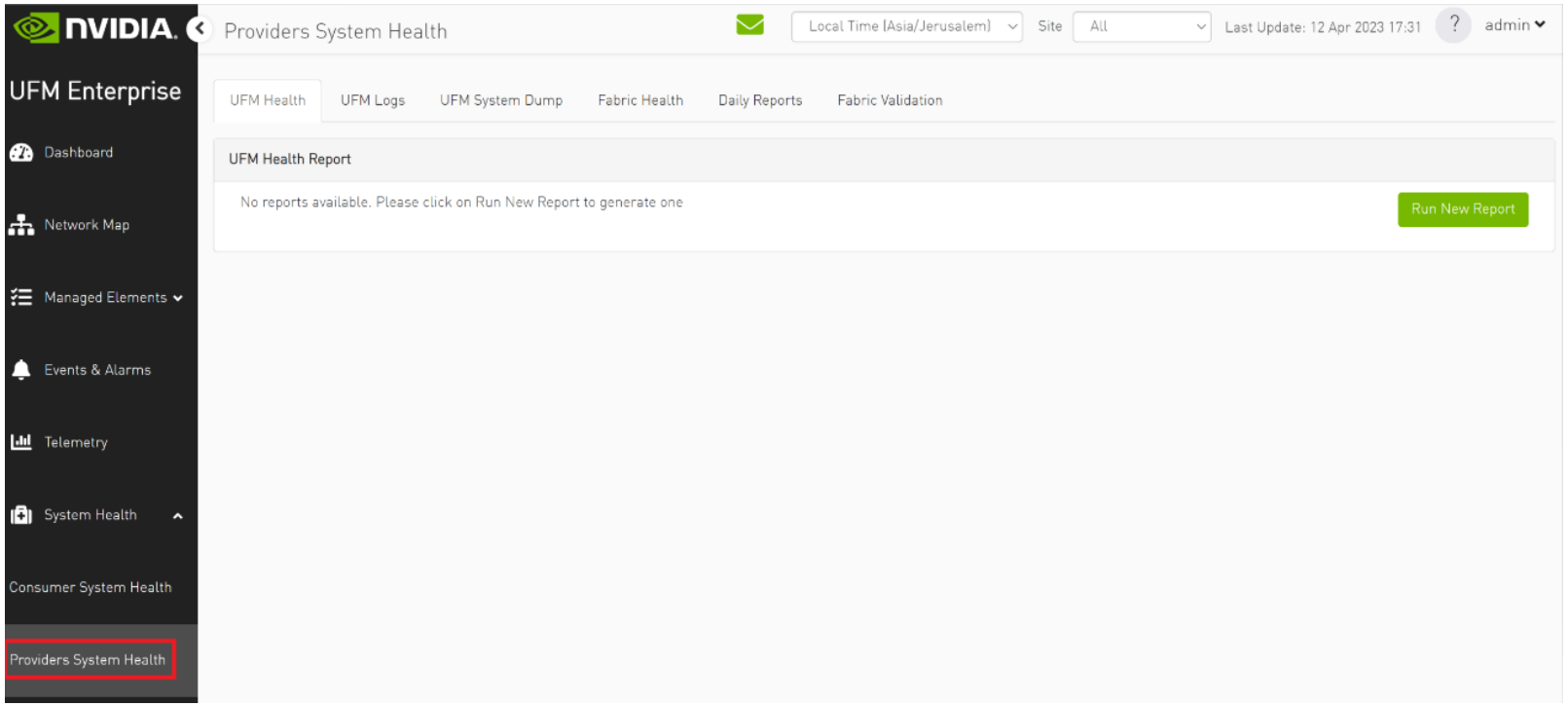


8. In the System Health tab, subsections for Consumer and Provider are available.

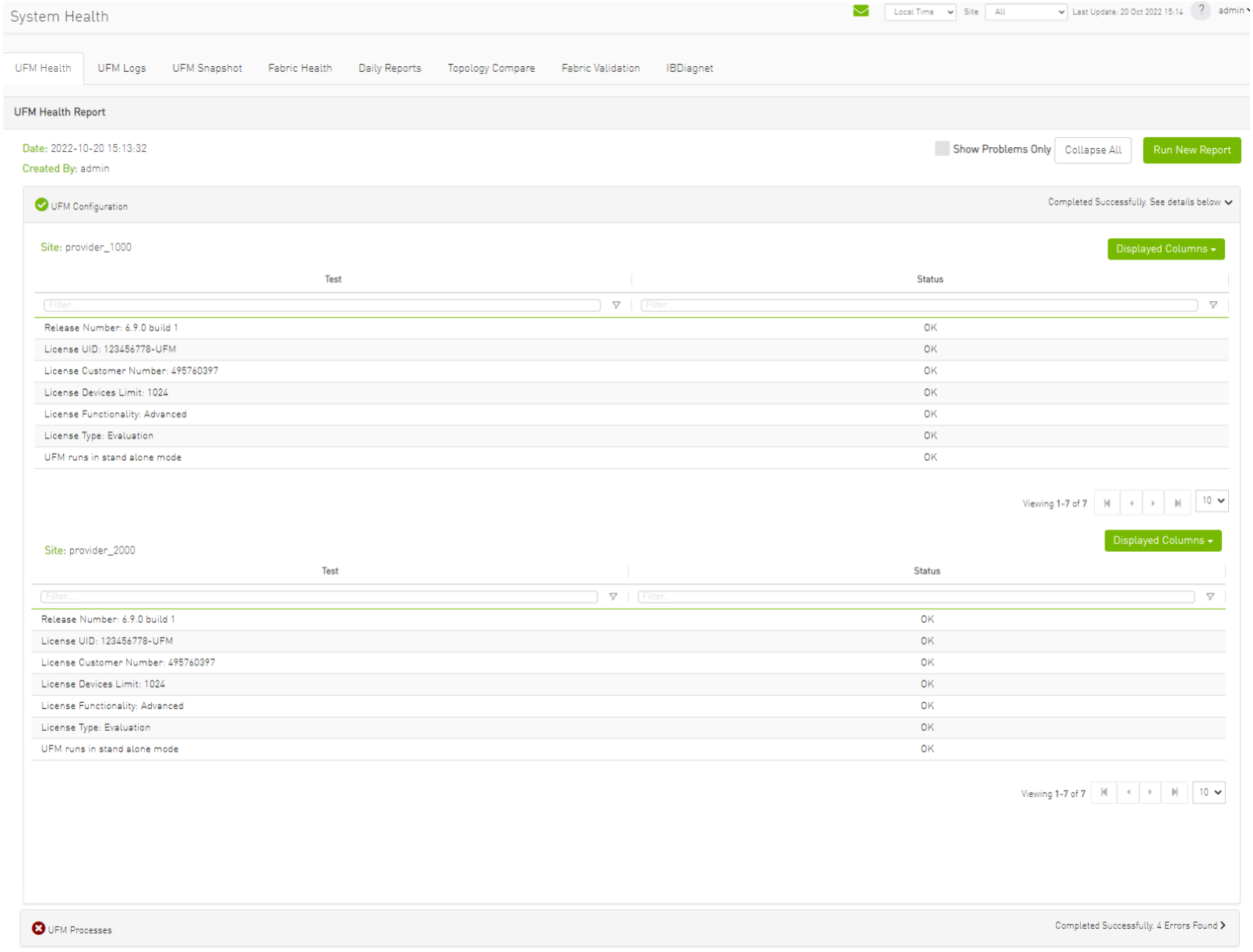
a. Consumer System Health tab contains sections applicable to Consumer UFM specifically (e.g., logs from Consumer UFM).



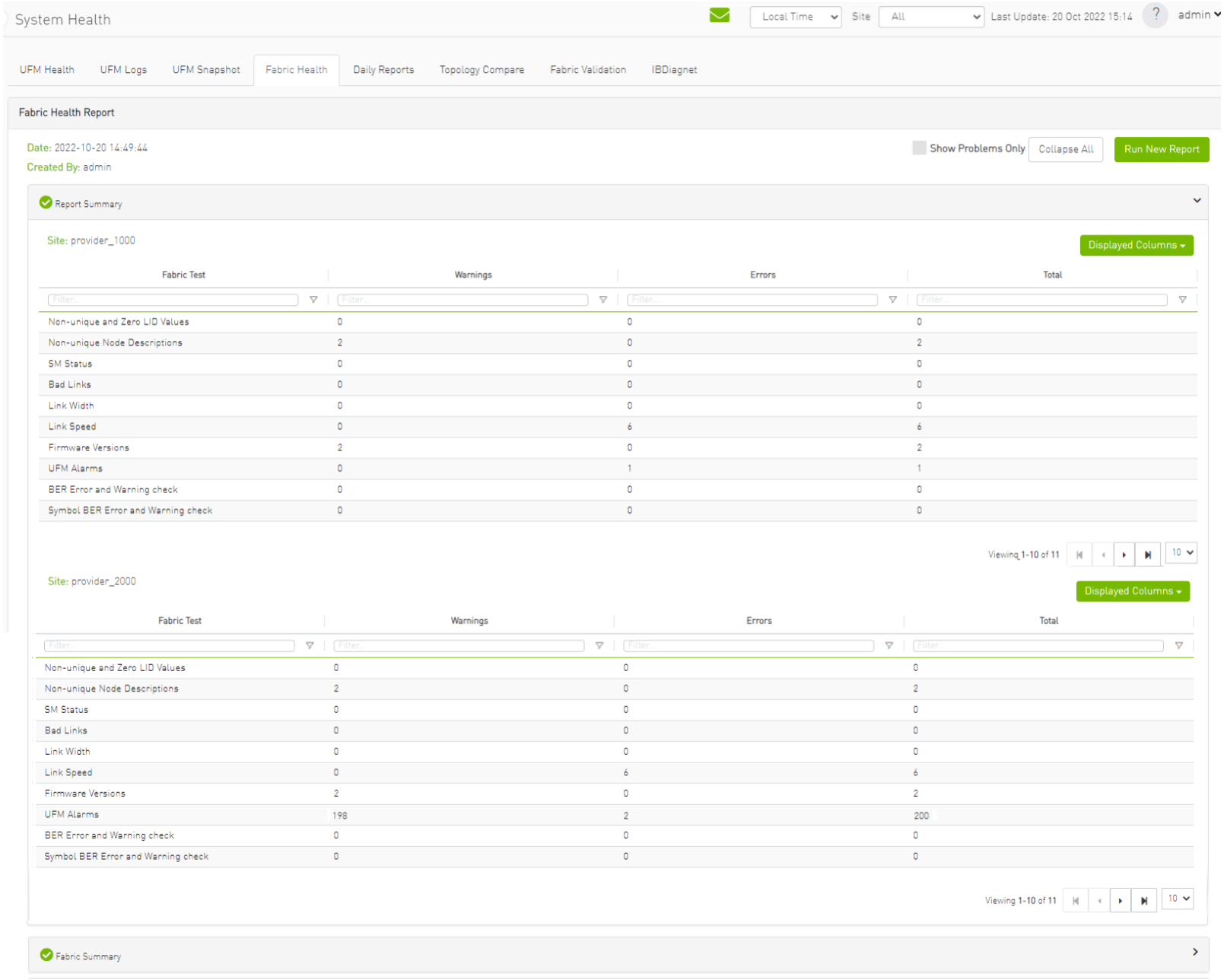
b. Provider System Health contains sections applicable to one or multiple providers (e.g., Fabric Health Report can be triggered on multiple Providers from the Consumer).


NVIDIA
Providers System Health
Local Time (Asia/Jerusalem)
Site
All
Last Update: 12 Apr 2023 17:31
admin
UFM Enterprise
Dashboard
Network Map
Managed Elements
Events & Alarms
Telemetry
System Health
Consumer System Health
Providers System Health
UFM Health
UFM Logs
UFM System Dump
Fabric Health
Daily Reports
Fabric Validation
UFM Health Report
No reports available. Please click on Run New Report to generate one
Run New Report

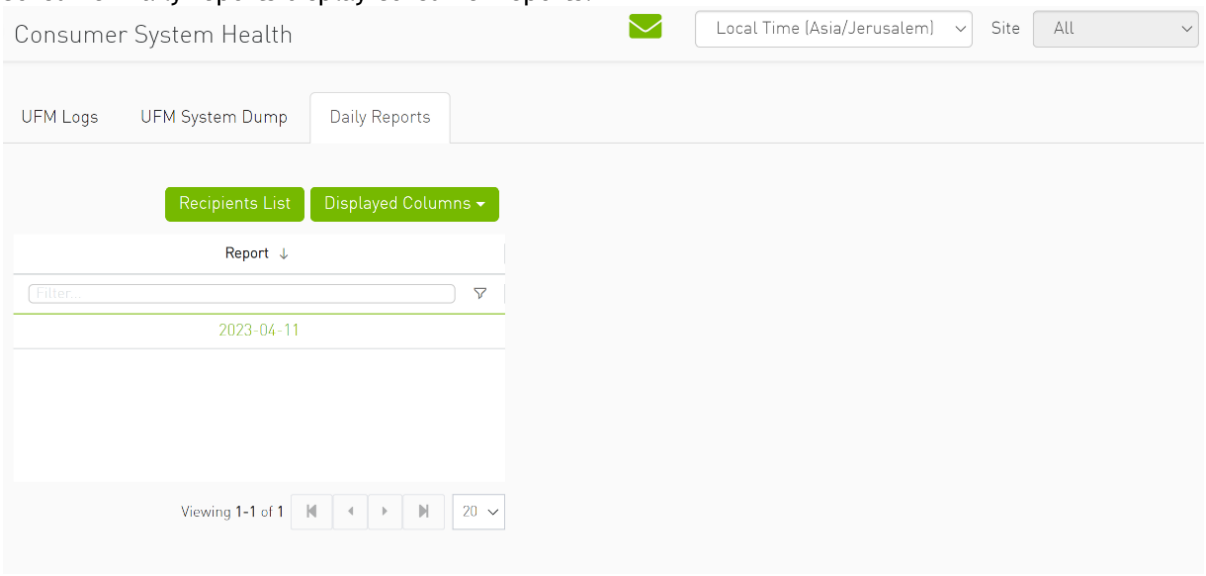
9. UFM Health tab contains sub report tables for each provider.


System Health
Local Time
Site
All
Last Update: 20 Oct 2022 15:14
admin
UFM Health
UFM Logs
UFM Snapshot
Fabric Health
Daily Reports
Topology Compare
Fabric Validation
IBDiagnet
UFM Health Report
Date: 2022-10-20 15:13:32
Created By: admin
Show Problems Only
Collapse All
Run New Report
UFM Configuration
Completed Successfully. See details below
Site: provider_1000
Displayed Columns
Test
Status
Release Number: 6.9.0 build 1 OK
License UID: 123456778-UFM OK
License Customer Number: 495760397 OK
License Devices Limit: 1024 OK
License Functionality: Advanced OK
License Type: Evaluation OK
UFM runs in stand alone mode OK
Viewing 1-7 of 7
10
Site: provider_2000
Displayed Columns
Test
Status
Release Number: 6.9.0 build 1 OK
License UID: 123456778-UFM OK
License Customer Number: 495760397 OK
License Devices Limit: 1024 OK
License Functionality: Advanced OK
License Type: Evaluation OK
UFM runs in stand alone mode OK
Viewing 1-7 of 7
10
UFM Processes
Completed Successfully. 4 Errors Found

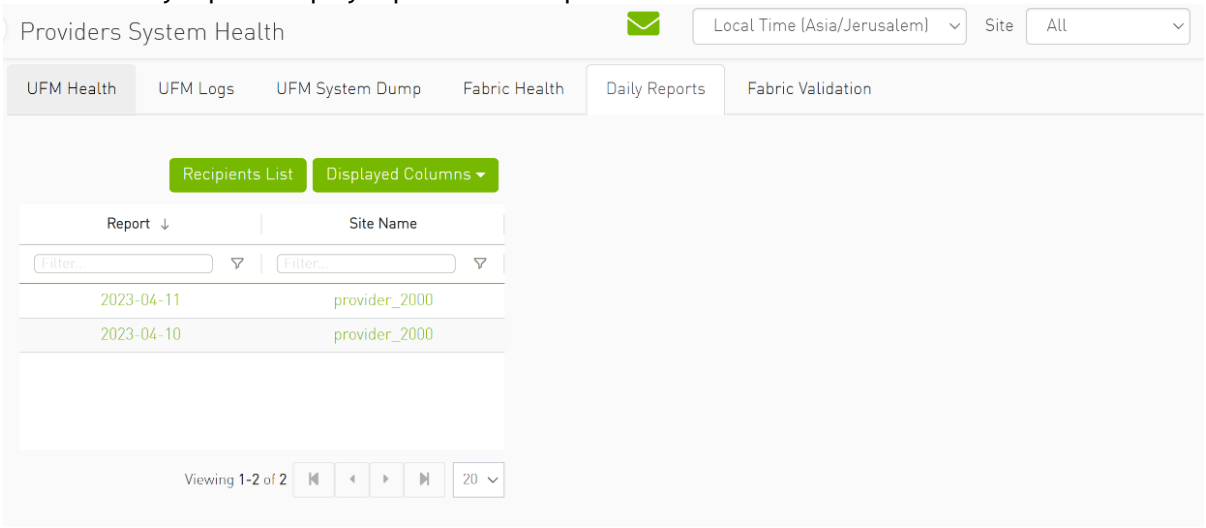
10. Fabric Health contains sub report tables for each provider.


System Health
Local Time
Site
All
Last Update: 20 Oct 2022 15:14
admin
UFM Health
UFM Logs
UFM Snapshot
Fabric Health
Daily Reports
Topology Compare
Fabric Validation
IBDiagnet
Fabric Health Report
Date: 2022-10-20 14:49:44
Created By: admin
Show Problems Only
Collapse All
Run New Report
Report Summary
Site: provider_1000
Displayed Columns
Fabric Test
Warnings
Errors
Total
Non-unique and Zero LID Values 0 0 0
Non-unique Node Descriptions 2 0 2
SM Status 0 0 0
Bad Links 0 0 0
Link Width 0 0 0
Link Speed 0 6 6
Firmware Versions 2 0 2
UFM Alarms 0 1 1
BER Error and Warning check 0 0 0
Symbol BER Error and Warning check 0 0 0
Viewing 1-10 of 11
10
Site: provider_2000
Displayed Columns
Fabric Test
Warnings
Errors
Total
Non-unique and Zero LID Values 0 0 0
Non-unique Node Descriptions 2 0 2
SM Status 0 0 0
Bad Links 0 0 0
Link Width 0 0 0
Link Speed 0 6 6
Firmware Versions 2 0 2
UFM Alarms 198 2 200
BER Error and Warning check 0 0 0
Symbol BER Error and Warning check 0 0 0
Viewing 1-10 of 11
10
Fabric Summary

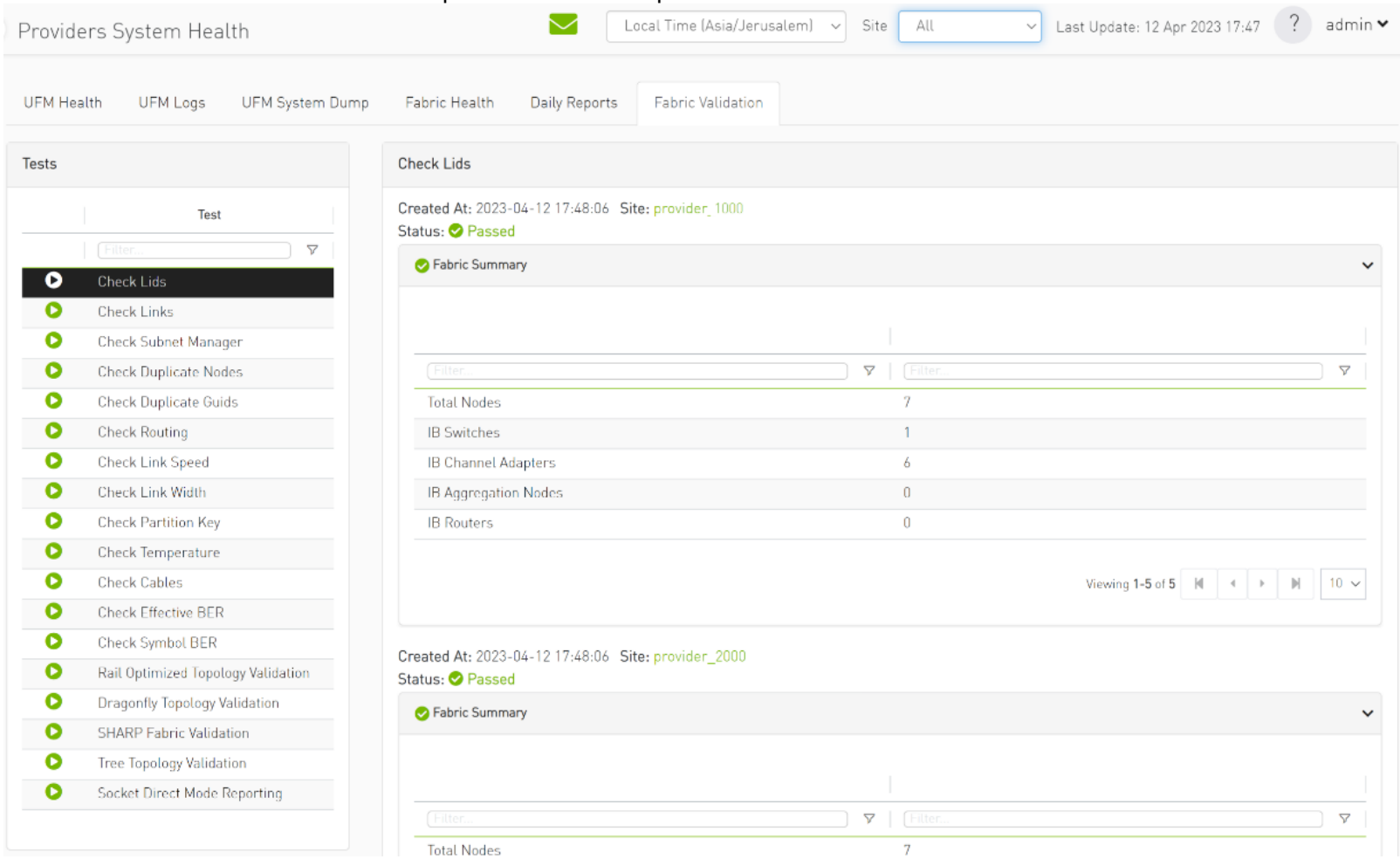
11. Daily Reports:
    a. Consumer Daily reports display consumer reports.



b. Providers Daily reports display reports from all providers.

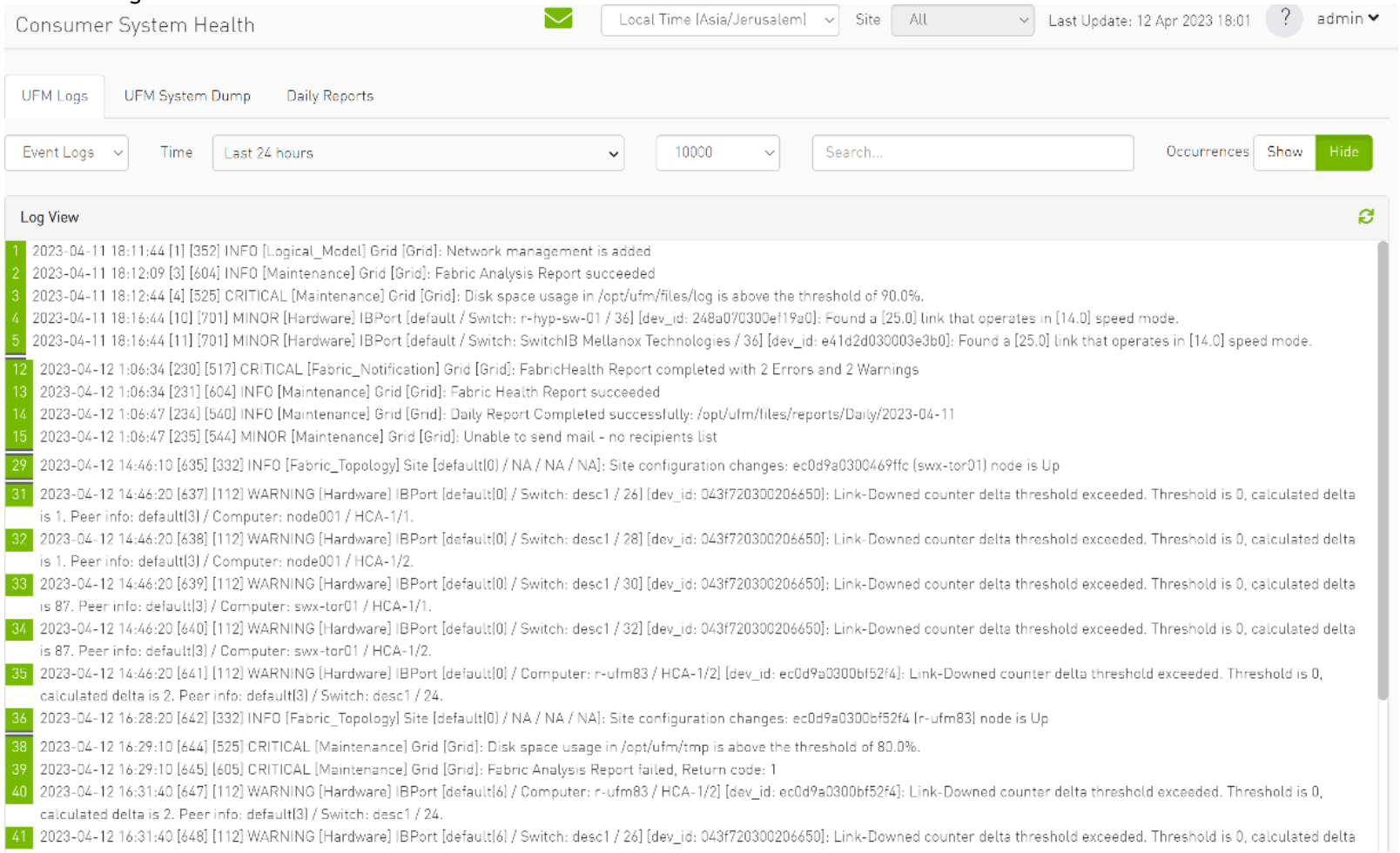


12. The "Fabric Validation" tab contains sub report tables for each provider.

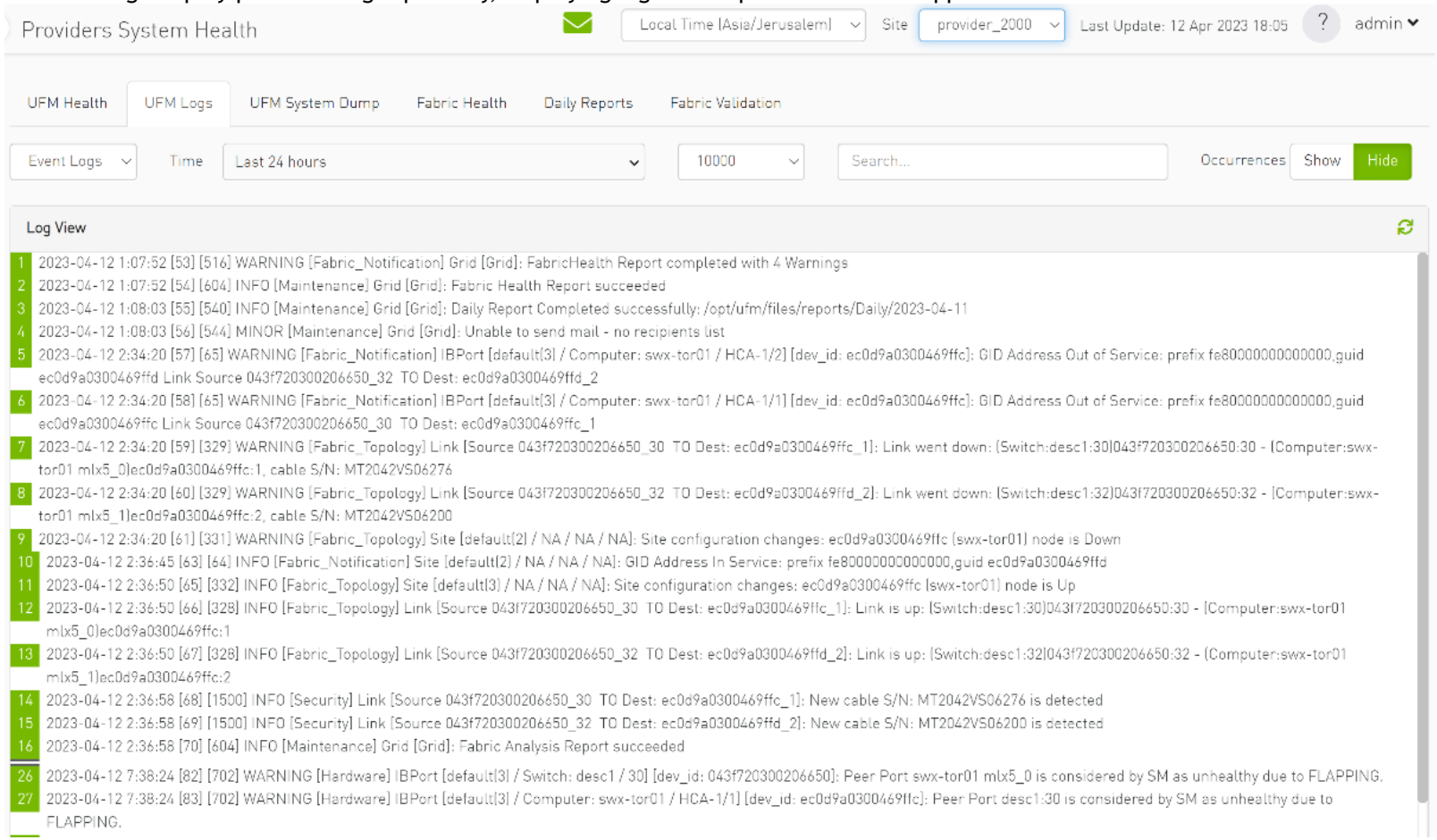


13. In "UFM Logs" Tab:

a. Consumer logs:

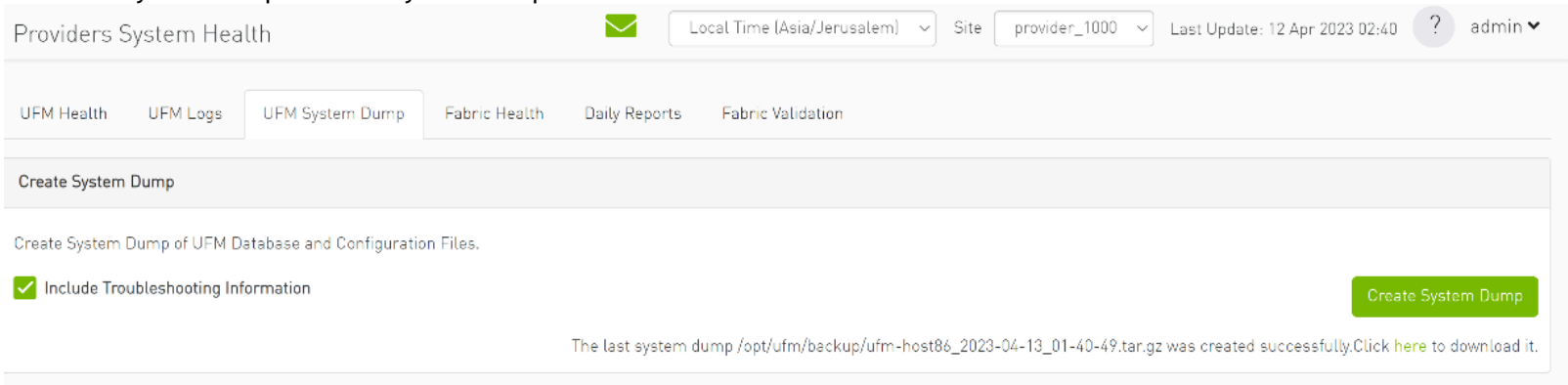


b. Providers logs display providers log separately, displaying logs for all providers is not supported.

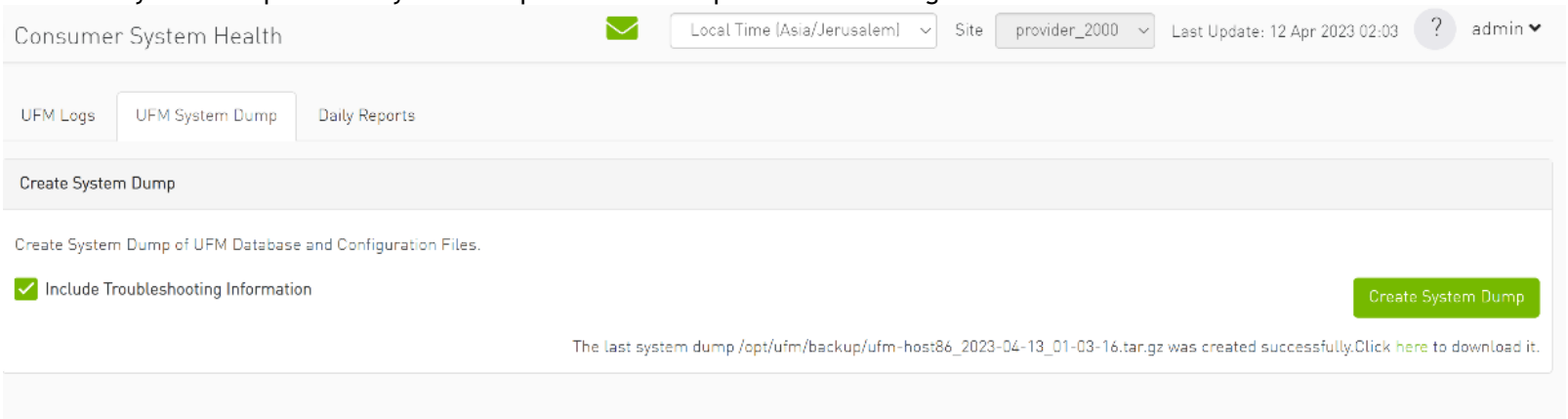


14. In the "System Dump" tab:

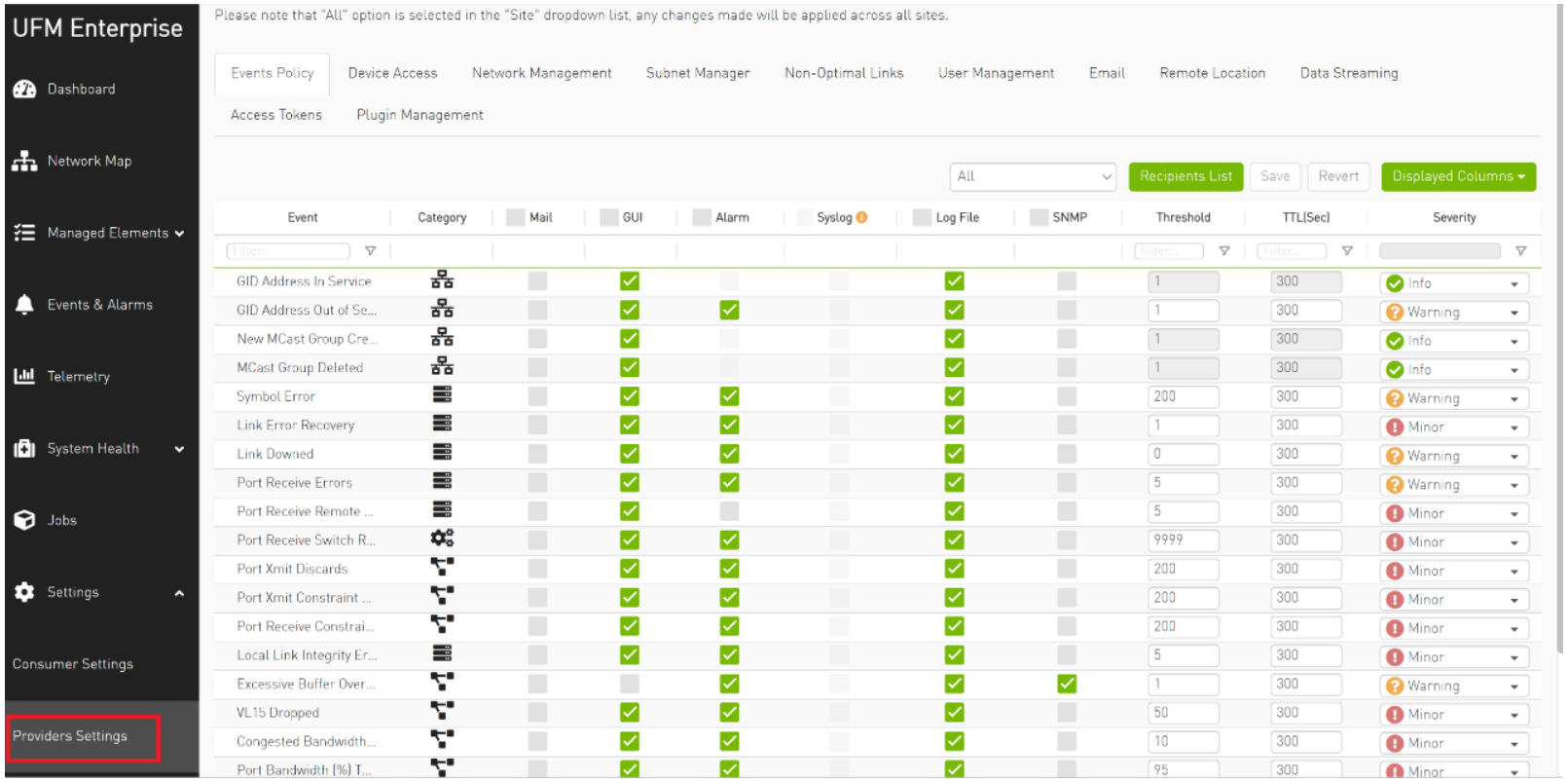
a. "Consumer System Dump" collects system dump for consumer



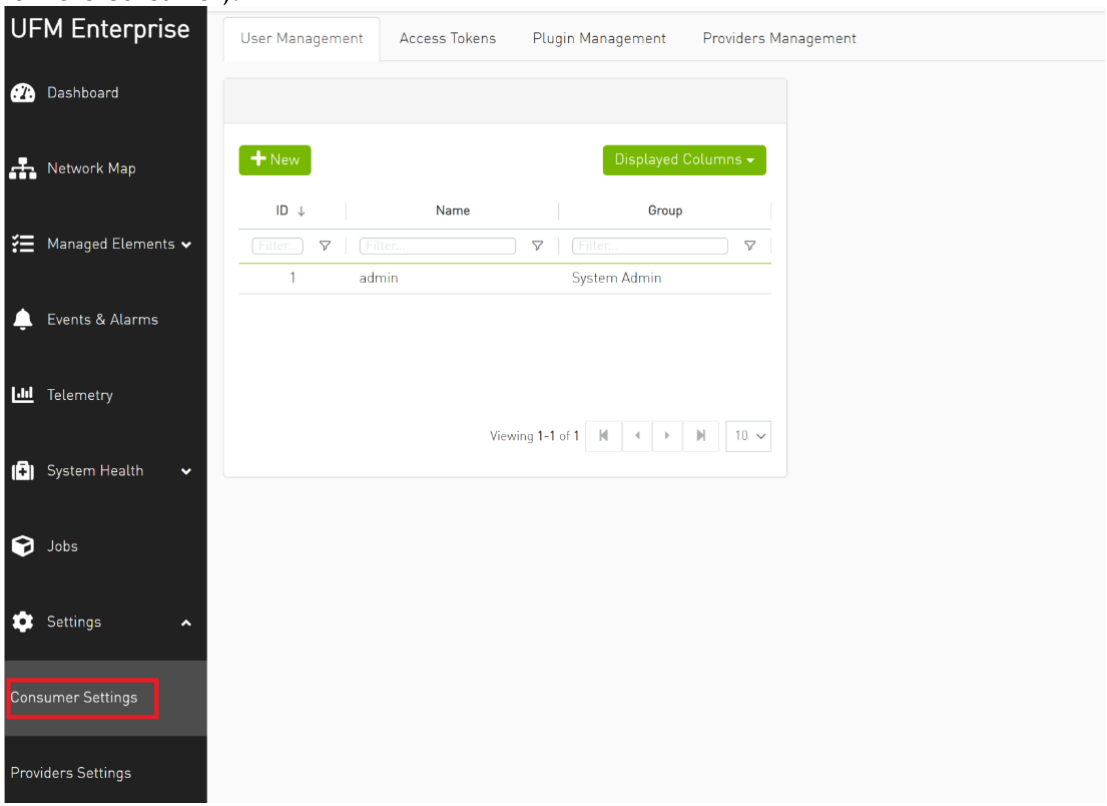
b. "Providers System Dump" collect system dumps for one or all providers and mergeS them into one folder



15. Under "Settings", subsections for Consumer and Provider are available.
    a. "Consumer Settings" contain sections applicable to Consumer UFM specifically (e.g., creation of access tokens for UFM consumer authentication);


UFM Enterprise
Dashboard
Network Map
Managed Elements
Events & Alarms
Telemetry
System Health
Jobs
Settings
Consumer Settings
Providers Settings
Please note that "All" option is selected in the "Site" dropdown list, any changes made will be applied across all sites.
Events Policy
Device Access
Network Management
Subnet Manager
Non-Optimal Links
User Management
Email
Remote Location
Data Streaming
Access Tokens
Plugin Management
All
Recipients List
Save
Revert
Displayed Columns
Event
Category
Mail
GUI
Alarm
Syslog
Log File
SNMP
Threshold
TTL(Sec)
Severity
GID Address In Service
GID Address Out of Se...
New MCast Group Cre...
MCast Group Deleted
Symbol Error
Link Error Recovery
Link Downed
Port Receive Errors
Port Receive Remote ...
Port Receive Switch R...
Port Xmit Discards
Port Xmit Constraint ...
Port Receive Constrai...
Local Link Integrity Er...
Excessive Buffer Over...
VL15 Dropped
Congested Bandwidth...
Port Bandwidth (%) T...

b. "Provider Settings" contain sections applicable to one or multiple providers (e.g., Event Policies can be changed for multiple Providers at once from the Consumer).



UFM Plugins

- rest-rdma Plugin
- NDT Plugin
- UFM Telemetry FluentD Streaming (TFS) Plugin
- UFM Events Fluent Streaming (EFS) Plugin
- UFM Bright Cluster Integration Plugin
- UFM Cyber-AI Plugin
- Autonomous Link Maintenance (ALM) Plugin
- DTS Plugin
- GRPC-Streamer Plugin
- Sysinfo Plugin
- SNMP Plugin
- Packet Mirroring Collector (PMC) Plugin
- PDR Deterministic Plugin
- GNMI-Telemetry Plugin

rest-rdma Plugin

rest-rdma is a tool designed for sending requests over InfiniBand to the UFM server. These REST requests can fall into three categories:

1. UFM REST API requests
2. ibdiagnet requests
3. Telemetry requests

The rest-rdma utility is distributed as a Docker container, capable of functioning both as a server and a client.

Deployment Server

Deploy Plugin on UFM Appliance

1. Log into your UFM as admin.
2. Enter config mode. Run:

```
enable
config terminal
```

> ⚠ Make sure that UFM is running with `show ufm status`. If UFM is down, then run with `ufm start`.

3. Ensure that rest-rdma plugin is disabled with the `show ufm plugin` command.
4. Pull the plugin container with `docker pull mellanox/ufm-plugin-rest-rdma:[version]`.
5. Run `ufm plugin rest-rdma add tag [version]` to enable the plugin.
6. Check that plugin is up and running with `show ufm plugin`

Deployment Client

Run the following command to pull the image from the docker hub:

```
docker pull mellanox/ufm-plugin-rest-rdma:[version]
```

Verify that the `/tmp/ibdiagnet` directory exists on the client's computer. If not - create it.

To start container as client (on any host in the same fabric as UFM server) run:

```
docker run -d --network=host --privileged --name=ufm-plugin-rest-rdma --rm -v /tmp/ibdiagnet:/tmp/ibdiagnet
mellanox/ufm-plugin-rest-rdma:[version] client
```

To check that plugin is up and running, run:

```
docker ps
```

How to Run

Server

In server mode `ufm_rdma.py` is started automatically and is restarted if exited. If the `ufm_rdma.py` server is not running - enter to the docker and run the following commands to start the server:

```
 cd /opt/ufm/src/ufm-plugin-ufm-rest
./ufm_rdma.py -r server
```

Client

There are three options to run client. Running the client from inside the Docker container, using a custom script from the hosting server to execute the client or using the "docker exec" command from the hosting server.

1. Option 1: Run the client from inside the Docker container
    a. Enter the docker container using `docker exec -it ufm-plugin-rest-rdma bash`
    b. Then, run `cd /opt/ufm/src/ufm-plugin-rest-rdma`
    c. Use the `-h` help option to see the available parameters
       `./ufm_rdma.py -h`
2. Option 2: From the host server, the scripts can be located at `/opt/ufm/ufm-plugin-ufm-rest/ directory` inside the docker container. They can copied using the following command:

> ⚠ cp <containerId>:/opt/ufm/ufm-plugin-ufm-rest/[script name] /host/path/target

Example:

> ⚠ cp <containerId>:/opt/ufm/ufm-plugin-ufm-rest/ufm-rest-rdma_client.sh /host/path/target

a. To see the available options, run:
   `./ufm-rest-rdma_client.sh -h`

3. Option 3: From hosting server, use the `docker exec` command.

> ⚠ To run from inside docker, run:
> `docker exec ufm-plugin-rest-rdma` prior to the command.
> For example: `docker exec ufm-plugin-rest-rdma /opt/ufm/ufm-plugin-ufm-rest/src/ufm_rdma.py -r client -u admin -p password -t simple -a GET -w ufmRest/app/ufm_version`

Authentication Configuration

Telemetry and ibdiagnet request authentication options could be enabled or disabled (enabled by default - set to True) in `ufm_rdma.ini file` in [Server] section on the server. The `rest_rdma` server performs simple requests to UFM server, using supplied credentials to verify that the user is allowed to run telemetry or ibdiagnet requests.

```
[Server]
use_ufm_authentication=True
```

Remote ibdiagnet Request

The following two user scripts can run on the hosting server.

- remote_ibdiagnet_auth.sh
- remote_ibdiagnet.sh

These scripts should be copied from the container to the hosting server using the following command:

```
cp <containerId>:/opt/ufm/ufm-plugin-ufm-rest/[script name] /host/path/target
```

Example :

```
cp <containerId>:/opt/ufm/ufm-plugin-ufm-rest/remote_ibdiagnet_auth.sh /host/path/target
```

The `remote_ibdiagnet.sh` script does not require authentication as the server side can run on a machine which does not run UFM (which is responsible for the authentication). This means it can run from the hosting server.

```
/remote_ibdiagnet.sh [options]
```

Authenticated Remote ibdiagnet Request

The `remote_ibdiagnet_auth.sh` script can receive parameters as credentials for authentication with UFM server.

```
/remote_ibdiagnet_auth.sh [options]
```

To get all the options, run the following command:

```
/remote_ibdiagnet_auth.sh -h
```

⚠ Important Note:
When using remote_ibdiagnet.sh, authentication is not required and the the ibdiagnet parameters should be sent in ibdiagnet format.
Example: `./remote_ibdiagnet.sh --get_phy_info`

When using the remote_ibdiagnet_**auth**.sh, the ibdiagnet parameters should be sent using the `-l` key.
Example without credentials: `./remote_ibdiagnet_auth.sh -l '--get_phy_info'`

Example with credentials: `./remote_ibdiagnet_auth.sh -u username -p password -l '-get_phy_info'`

Please use the -h option to see the examples of credential usage.

Rest Request with Username/Password Authentication

To get the UFM version from inside the docker:

```
./ufm_rdma.py -r client -u admin -p admin_pwd -t simple -a GET -w ufmRest/app/ufm_version
```

To get the UFM version from hosting server using script:

```
./ufm_rest_rdma_client.sh -u admin -p admin_pwd -t simple -a GET -w ufmRest/app/ufm_version
```

For telemetry:

```
./ufm_rdma.py -r client -u admin -p admin_pwd -t telemetry -a GET -g 9001 -w /csv/enterprise
```

To get ibdiagnet run result using UFM REST API from inside the docker:

```
./ufm_rdma.py -r client -u admin -p admin_pwd -t ibdiagnet -a POST -w ufmRest/reports/ibdiagnetPeriodic -l
'{"general": {"name": "IBDiagnet_CMD_1234567890_199_88", "location": "local", "running_mode": "once"},
"command_flags": {"--pc": ""}}'
```

Rest Request with Client Certificate Authentication

```
need to pass path to client certificate file and name of UFM server machine:
6.  ./ufm_rdma.py -r client -t simple -a GET -w ufmRest/resources/modules -d /path/to/certificate/file/ufm-
client.pfx -s ufm.azurehpc.core.azure-test.net
for telemetry if need authentication from inside the docker
```

```
./ufm_rdma.py -r client  -t telemetry -a GET -g 9001 -w csv/enterprise -d /path/to/certificate/file/ufm-client.pfx
-s ufm.azurehpc.core.azure-test.net
```

⚠ Client certificate file should be located INSIDE the docker container.

Rest Request with Token Authentication

```
need to pass token for authentication
./ufm_rdma.py -r client -k OGUY7TwLvTmFkXyTkcsEWD9KKNvq6f -t simple -a GET -w ufmRestV3/app/ufm_version
for telemetry if need to perform authentication
./ufm_rdma.py -r client -k 4rQRf7i7wEeliuJEurGbeecc210V6G  -t telemetry -a GET -g 9001 -w /csv/enterprise
```

⚠ Token could be generated using UFM UI.

⚠ If a token is used for client authentication, `ufmRestV3` must be used.

NDT Plugin

Overview

NDT plugin is a self-contained Docker container with REST API support managed by UFM. The NDT plugin introduces the following capabilities:

a. NDT topology comparison: Allows the user to compare InfiniBand fabric managed by the UFM and NDT files which are used for the description of InfiniBand clusters network topology.
   - Verifies the IB fabric connectivity during cluster bring-up.
   - Verifies the specific parts of IB fabric after component replacements.
   - Automatically detects any changes in topology.

b. Subnet Merger - Expansion of the fabric based on NDT topology files
   Allows users to gradually extend the InfiniBand fabric without causing any disruption to the running fabric. The system administrator should prepare the NDT topology files, which describe the InfiniBand fabric extensions. Then, an intuitive and user-friendly UI wizard facilitates the topology extension process with a step-by-step guidance for performing necessary actions.
   - The Subnet Merger tool verifies the fabric topology within a predefined NDT file, and reports issues encountered for immediate resolution.
   - Once the verification results are acceptable by the network administrator, the tool creates a topoconfig file to serve as input for OpenSM. This allows setting the physical port states of the designated boundary ports as desired (physical ports can be set as disabled or no-discover).
   - Once the topoconfig file is deployed, the IB network can be extended and verified for the next IB extension.

Deployment

The following are the possible ways NDT plugin can be deployed:

1. On UFM Appliance
2. On UFM Software

For detailed instructions on how to deploy the NDT plugin refer to this page.

Authentication

Following authentication types are supported:

- basic (/ufmRest)
- client (/ufmRestV2)
- token (/ufmRestV3)

REST API

The following REST APIs are supported:

Topodiff

- GET /help
- GET /version
- POST /upload_metadata
- GET /list
- POST /compare
- POST /cancel
- GET /reports
- GET /reports/<report_id>
- POST /delete

Subnet Merger

- GET /merger_ndts_list
- GET /merger_ndts_list/<ndt_file_name>
- POST /merger_upload_ndt
- POST /merger_verify_ndt
- GET /merger_verify_ndt_reports
- GET /merger_verify_ndt_reports/<report_id>
- POST /merger_update_topoconfig
- POST /merger_deploy_ndt_config
- POST /merger_update_deploy_ndt_config
- POST /merger_delete_ndt
- GET /merger_deployed_ndt
- POST /merger_create_topoconfig

For detailed information on how to interact with NDT plugin, refer to the NVIDIA UFM Enterprise > Rest API > NDT Plugin REST API.

NDT Format – Topodiff

NDT is a CSV file containing data relevant to the IB fabric connectivity. The NDT plugin extracts the IB connectivity data based on the following fields:

1. Start device
2. Start port
3. End device
4. End port
5. Link type

Switch to Switch NDT

By default, IB links are filtered by:

- Link Type is Data
- Start Device and End Device end with IBn, where n is a numeric value.

For TOR switches, Start port/End port field should be in the format **Port N**, where **N** is a numeric value.

For Director switches, Start port/End port should be in the format **Blade N_Port i/j**, where **N** is a leaf number, **i** is an internal ASIC number and **j** is a port number.

Examples:

Start Device	Start Port	End Device	End Port	Link Type
DSM07-0101-0702-01IB0	Port 21	DSM07-0101-0702-01IB1	Blade 2_Port 1/1	Data
DSM07-0101-0702-01IB0	Port 22	DSM07-0101-0702-01IB1	Blade 2_Port 1/1	Data
DSM07-0101-0702-01IB0	Port 23	DSM07-0101-0702-02IB1	Blade 3_Port 1/1	Data
DSM09-0101-0617-001IB2	Port 33	DSM09-0101-0721-001IB4	Port 1	Data
DSM09-0101-0617-001IB2	Port 34	DSM09-0101-0721-001IB4	Port 2	Data
DSM09-0101-0617-001IB2	Port 35	DSM09-0101-0721-001IB4	Port 3	Data

Switch to Host NDT

NDT is a CSV file containing data not only relevant to the IB connectivity.

Extracting the IB connectivity data is based on the following five fields:

1. Start device
2. Start port
3. End device
4. End port
5. Link type

IB links should be filtered by the following:

- Link type is "Data".
- "Start Device" or "End Device" end with **IBN**, where **N** is a numeric value.
    - The other Port should be based on persistent naming convention: **ibpXsYfZ**, where **X**, **Y** and **Z** are numeric values.

For TOR switches, Start port/End port field will be in the format Port n, where n is a numeric value.

For Director switches, Start port/End port will be in the format **Blade N_Port i/j**, where **N** is a leaf number, **i** is an internal ASIC number and **j** is a port number.

Examples:

Start Device	Start Port	End Device	End Port	Link Type
DSM071081704019	DSM071081704019 ibp11s0f0	DSM07-0101-0514-01IB0	Port 1	Data
DSM071081704019	DSM071081704019 ibp21s0f0	DSM07-0101-0514-01IB0	Port 2	Data
DSM071081704019	DSM071081704019 ibp75s0f0	DSM07-0101-0514-01IB0	Port 3	Data

Other

Comparison results are forwarded to syslog as events. Example of `/var/log/messages` content:

1. Dec 9 12:32:31 <server_ip> ad158f423225[4585]: NDT: missing in UFM "SAT111090310019/SAT111090310019 ibp203s0f0 - SAT11-0101-0903-19IB0/15"
2. Dec 9 12:32:31 <server_ip> ad158f423225[4585]: NDT: missing in UFM "SAT11-0101-0903-09IB0/27 - SAT11-0101-0905-01IB1-A/Blade 12_Port 1/9"
3. Dec 9 12:32:31 <server_ip> ad158f423225[4585]: NDT: missing in UFM "SAT11-0101-0901-13IB0/23 - SAT11-0101-0903-01IB1-A/Blade 08_Port 2/13"

For detailed information about how to check syslog, please refer to the NVIDIA UFM-SDN Appliance Command Reference Guide > UFM Commands > UFM Logs.

Minimal interval value for periodic comparison in five minutes.

In case of an error the clarification will be provided.

For example, the request "`POST /compare`" without NDTs uploaded will return the following:

- URL: https://<server_ip>/ufmRest/plugin/ndt/compare
- response code: 400
- Response:

```
{
  "error": [
    "No NDTs were uploaded for comparison"
  ]
}
```

Configurations could be found in "`ufm/conf/ndt.conf`"

- Log level (default: INFO)
- Log size (default: 10240000)
- Log file backup count (default: 5)
- Reports number to save (default: 10)
- NDT format check (default: enabled)
- Switch to switch and host to switch patterns (default: see NDT format section)

For detailed information on how to export or import the configuration, refer to the NVIDIA UFM-SDN Appliance Command Reference Guide > UFM Commands > UFM Configuration Management.

Logs could be found in “`ufm/logs/ndt.log`”.

For detailed information on how to generate a debug dump, refer to the NVIDIA UFM-SDN Appliance Command Reference Guide > System Management > Configuration Management > File System.

NDT Format – Subnet Merger

The Subnet Merger tool facilitates the seamless expansion of the InfiniBand fabric based on Non-Disruptive Topology (NDT) files. This section outlines the process of extending the fabric while ensuring uninterrupted operation. The tool operates through an intuitive UI wizard, guiding users step-by-step in extending the fabric topology.

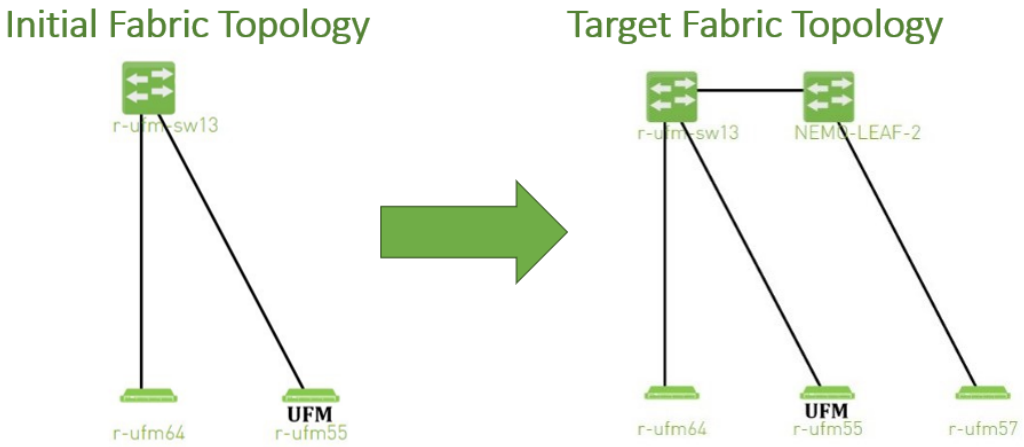
The Subnet Merger tool enables the gradual expansion of the InfiniBand fabric without causing disruptions to the existing network. To achieve this, system administrators need to prepare NDT topology files that describe the planned fabric extensions. The tool offers an intuitive UI wizard that simplifies the extension process.

Functionality

1. NDT Topology File Verification: The Subnet Merger tool verifies the InfiniBand fabric topology specified in a predefined NDT file. During this verification, any issues encountered are reported to the user for immediate resolution. This step ensures the integrity of the planned fabric extension.
1. Topology Extension Preparation: Upon successful verification of the NDT topology file, the tool generates a comprehensive verification report. The network administrator reviews this report and ensures its acceptability.
1. Topoconfig File Generation: After obtaining acceptable verification results, the tool generates a topoconfig file. This file serves as input for OpenSM, the Subnet Manager for InfiniBand fabrics. The topoconfig file allows the network administrator to define the desired physical port states for designated boundary ports. These states include "disabled" or "no-discover."
1. Fabric Extension and Verification: With the topoconfig file prepared, the Subnet Merger tool initiates the deployment of the extended fabric configuration. The tool ensures that the defined physical port states are implemented. Once the extension is in place, the IB network can be extended further as needed. The fabric extension is executed while maintaining the operational stability of the existing network.

1. Conclusion: The Subnet Merger tool offers a reliable and user-friendly solution for expanding InfiniBand fabrics using NDT topology files. By following the steps provided in the intuitive UI wizard, system administrators can seamlessly extend the fabric while adhering to predefined physical port states. This tool ensures the smooth operation of the fabric throughout the expansion process, eliminating disruptions and enhancing network scalability.

Subnet Merger Flow



1. Create NDT, file that describes initial topology with definition of boundary ports. Boundary ports – switch ports that will be used for fabric extension. In our case it will be r-ufm-sw13 switch ports number 1 and 3. In NDT file those ports should be defined as boundary and disabled:

```
rack #,U height,#Fields:StartDevice,StartPort,StartDeviceLocation,EndDevice,EndPort,EndDeviceLocation,U
height_1,LinkType,Speed,_2,Cable Length,_3,_4,_5,_6,_7,State,Domain
,,MF0;r-ufm-sw13:MQM8700/U1,Port 1,,,,,,,,,,,,,,,Disabled,Boundary
,,MF0;r-ufm-sw13:MQM8700/U1,Port 30,,r-ufm55 mlx5_1,Port 1,,,,,,,,,,,,Active,In-Scope
,,MF0;r-ufm-sw13:MQM8700/U1,Port 29,,r-ufm55 mlx5_0,Port 1,,,,,,,,,,,,Active,In-Scope
,,MF0;r-ufm-sw13:MQM8700/U1,Port 26,,r-ufm64 mlx5_0,Port 1,,,,,,,,,,,,Active,In-Scope
,,MF0;r-ufm-sw13:MQM8700/U1,Port 3,,,,,,,,,,,,,,,Disabled,Boundary
```

2. Upload a new NDT topology file which describes the desired topology. Before deploying to UFM, the new NDT topology file should be verified against the existing topology – to find out mismatches and problems.
   After the verification, the plugin generates reports including information about:
   - Duplicated GUIDs
   - Misswired links
   - Non-existent links in the pre-defined NDT files
   - Links that exist in the fabric and not in the NDT file
2. Following the issues detected in the plugin reports, the network administrator changes the NDT file or the fabric. The verification process can be repeated as many times as necessary until the network administrator is satisfied with the results.
3. If the NDT verification results are satisfactory, a topoconfig file is generated and can be deployed to the UFM server to be used as configuration input for OpenSM. Topoconfig file should be located at /opt/ufm/files/conf/opensm/topoconfig.cfg on UFM server. By sending SIGHUP signal to opensm it forced to read configuration and to deploy it. In topoconfig file at this stage boundary ports will be defined as Disabled.
   Example of topoconfig.cfg:

```
0xb83fd2030080302e,1,-,-,Any, Disabled
0xb83fd2030080302e,30,0xf452140300280081,1,Any,Active
0xb83fd2030080302e,29,0xf452140300280080,1,Any,Active
0xb83fd2030080302e,26,0xf452140300280040,1,Any,Active
0xb83fd2030080302e,3,-,-,Any, Disabled
```

4. Next stage is to extend the fabric. Prepare separately new subnet that will be added to the existing fabric and, once it is ready, connect to the boundary ports, that are defined as Disabled in configuration file, so newly added subnet will not be discovered by opensm and will not affect in any way current setup functionality.
5. Once new subnet connected to the fabric - prepare next NDT file, that contains setup, that describes current fabric with extended, when previously defined as boundary ports defined as Active and if planned to continue with extension new ports defined as boundary.
   For example port number 9 of switch r-ufm-sw13:

```
rack #,U height,#Fields:StartDevice,StartPort,StartDeviceLocation,EndDevice,EndPort,EndDeviceLocation,U
height_1,LinkType,Speed,_2,Cable Length,_3,_4,_5,_6,_7,State,Domain
,,MF0;r-ufm-sw13:MQM8700/U1,Port 1,,NEMO-LEAF-2,Port 1,,,,,,,,,,,Active,In-Scope
,,MF0;r-ufm-sw13:MQM8700/U1,Port 30,,r-ufm55 mlx5_1,Port 1,,,,,,,,,,,Active,In-Scope
,,MF0;r-ufm-sw13:MQM8700/U1,Port 29,,r-ufm55 mlx5_0,Port 1,,,,,,,,,,,Active,In-Scope
,,NEMO-LEAF-2,Port 11,,r-ufm57 mlx5_0,Port 1,,,,,,,,,,,Active,In-Scope
,,MF0;r-ufm-sw13:MQM8700/U1,Port 26,,r-ufm64 mlx5_0,Port 1,,,,,,,,,,,Active,In-Scope
```

```
,,NEMO-LEAF-2,Port 1,,MF0;r-ufm-sw13,Port 1,,,,,,,,,,,Active,In-Scope
,,MF0;r-ufm-sw13:MQM8700/U1,Port 3,,NEMO-LEAF-2,Port 3,,,,,,,,,,,Active,In-Scope
,,NEMO-LEAF-2,Port 3,,MF0;r-ufm-sw13,Port 3,,,,,,,,,,,Active,In-Scope
,,MF0;r-ufm-sw13:MQM8700/U1,Port 9,,,,,,,,,,,,,Disabled,Boundary
```

6. After new subnet connected physically to the fabric, in opensm configuration file (topoconfig.cfg) boundary ports previously defined as Disabled should be set as No-discover. Example:

```
0xb83fd2030080302e,1,-,-,Any,No-discover
0xb83fd2030080302e,30,0xf452140300280081,1,Any,Active
0xb83fd2030080302e,29,0xf452140300280080,1,Any,Active
0xb83fd2030080302e,26,0xf452140300280040,1,Any,Active
0xb83fd2030080302e,3,-,-,Any,No-discover
```

7. Updated file should be deployed to UFM. In case boundary ports will be defined as No-discover – fabric, connected beyond those ports will not be discovered by opensm, but all the ibutils (ibdiagnet…) could send mads beyond those ports to newly added subnet - so NDT file verification for extended setup could be performed.
8. Upload new NDT file and run verification for this file. Fix problems detected by verification. Once satisfied with results – deploy configuration to UFM.
   Example of topoconfig file for extended setup:

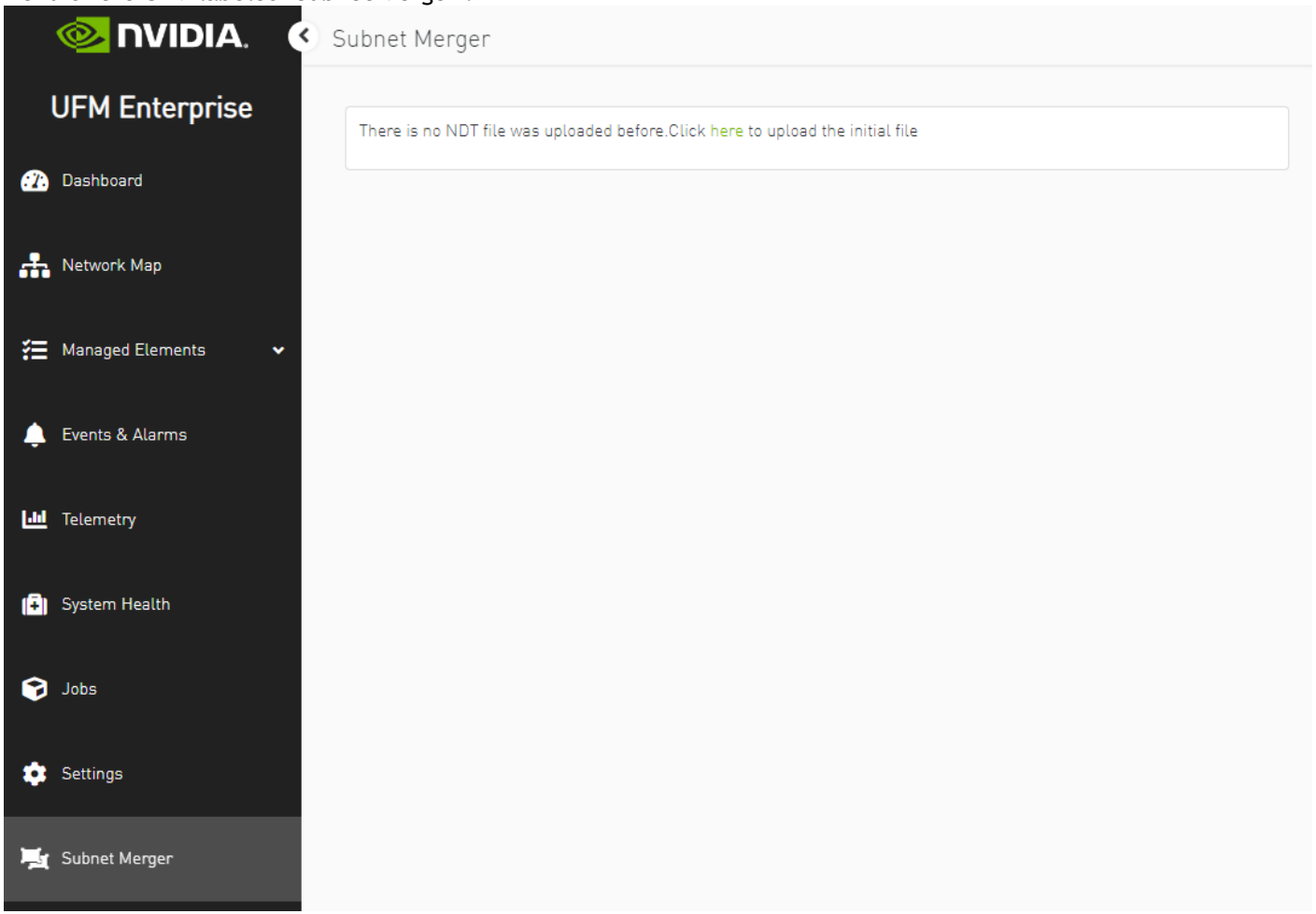
```
0xb83fd2030080302e,1,0x98039b0300867bba,1,Any,Active
0xb83fd2030080302e,30,0xf452140300280081,1,Any,Active
0xb83fd2030080302e,29,0xf452140300280080,1,Any,Active
0x98039b0300867bba,11,0x248a0703009c0066,1,Any,Active
0xb83fd2030080302e,26,0xf452140300280040,1,Any,Active
0x98039b0300867bba,1,0xb83fd2030080302e,1,Any,Active
0xb83fd2030080302e,3,0x98039b0300867bba,3,Any,Active
0x98039b0300867bba,3,0xb83fd2030080302e,3,Any,Active
0xb83fd2030080302e,9,-,-,Any,Disabled
```

9. Repeat previous steps if need to perform additional setup extension.

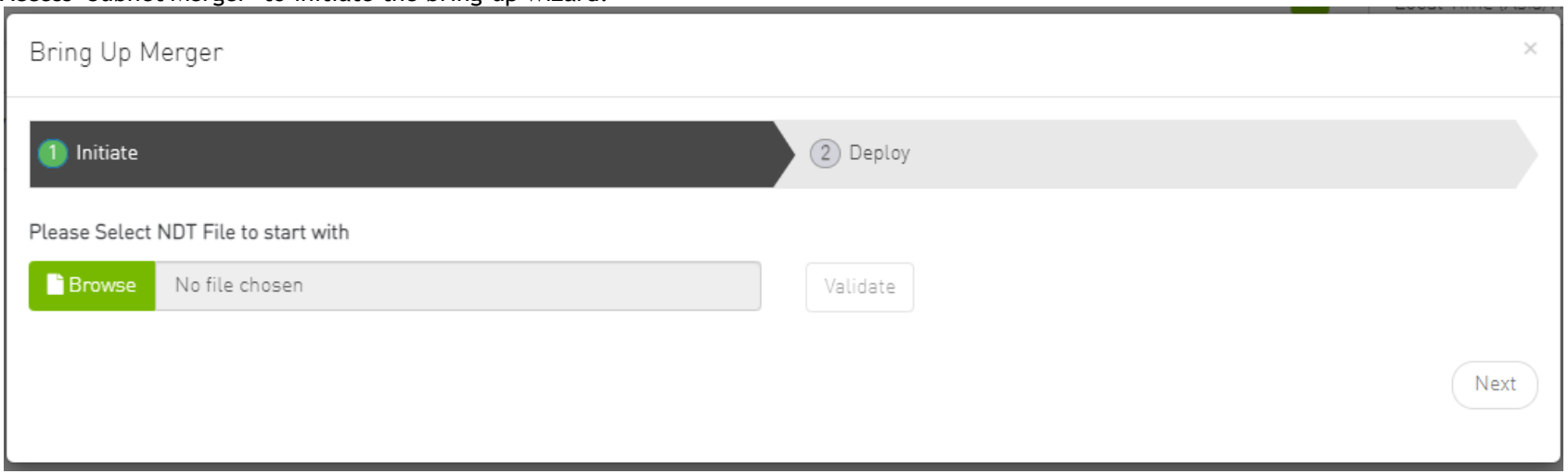
Subnet Merger UI

Bring-Up Merger Wizard

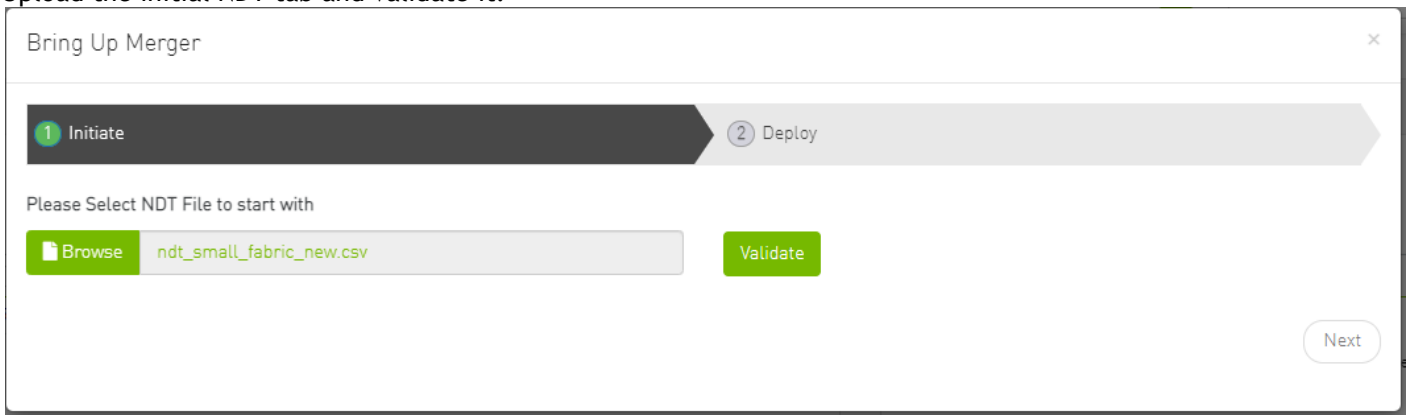
1. Add the NDT plugin to UFM by loading the plugin's image through Settings->Plugins Management. A new item will appear in the main left navigator menu of the UFM labeled "Subnet Merger".

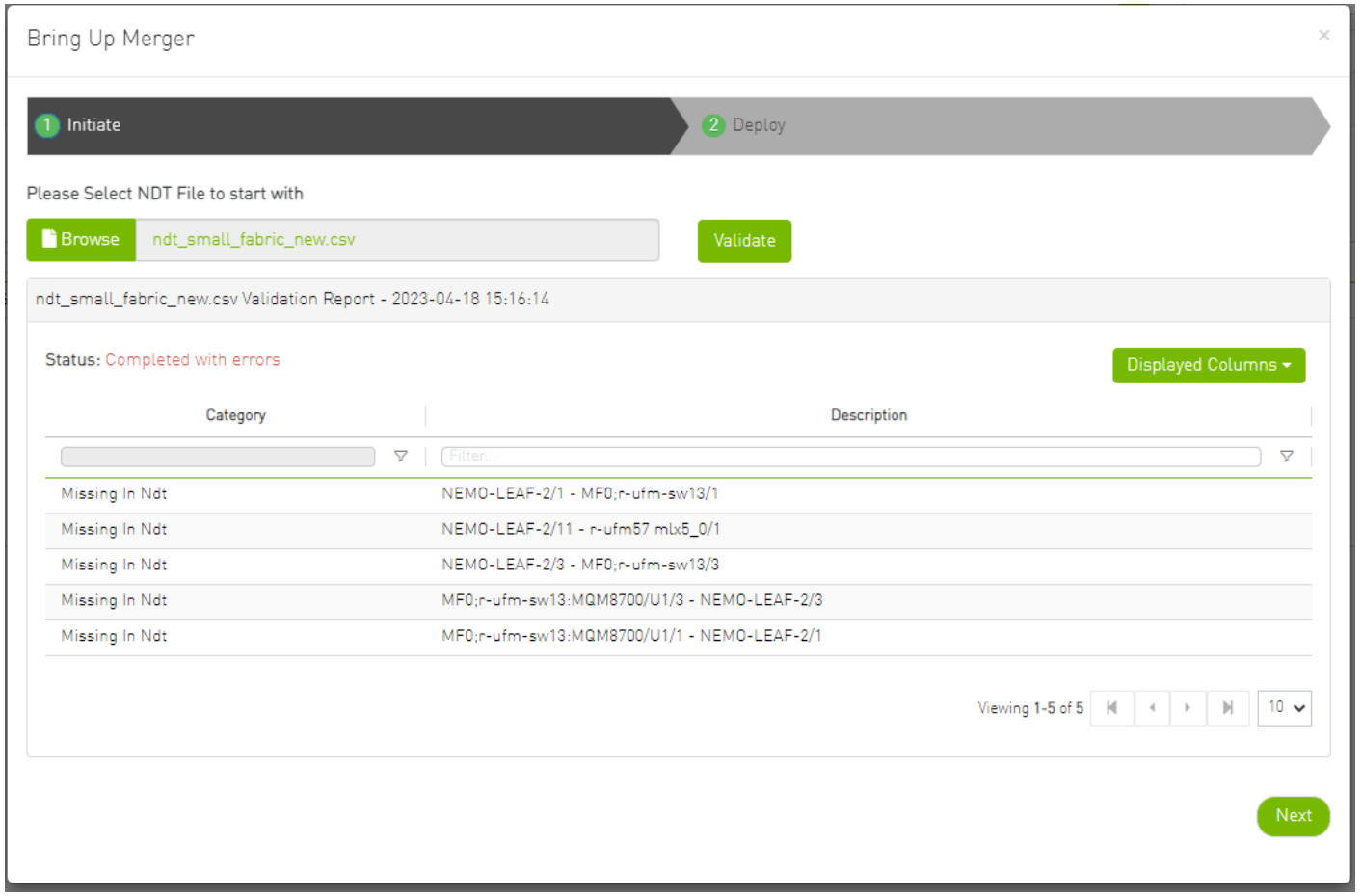


2. Access "Subnet Merger" to initiate the bring-up wizard.

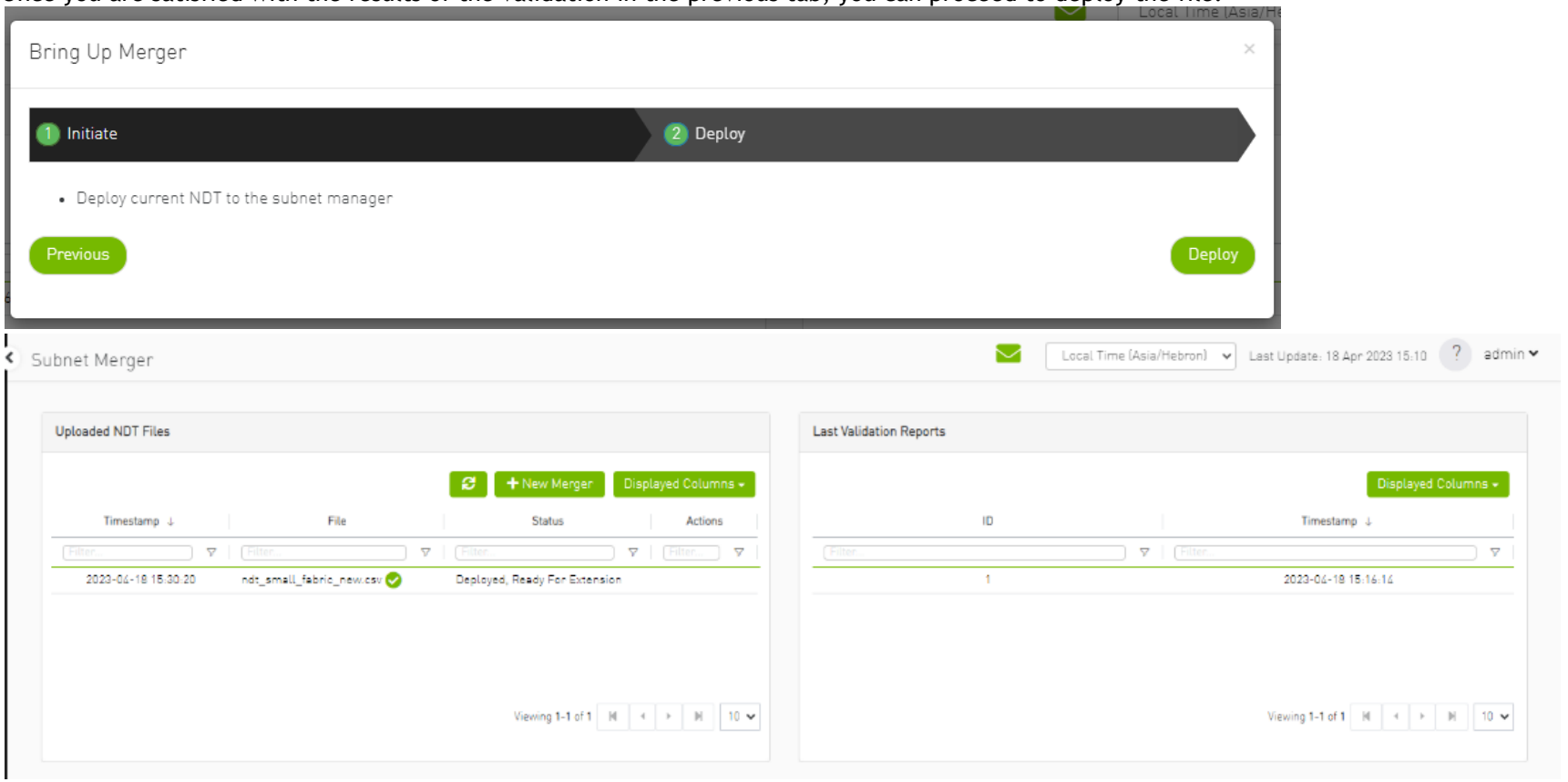


3. The wizard will guide you through the process, containing the following steps:
    a. Upload the initial NDT tab and validate it.



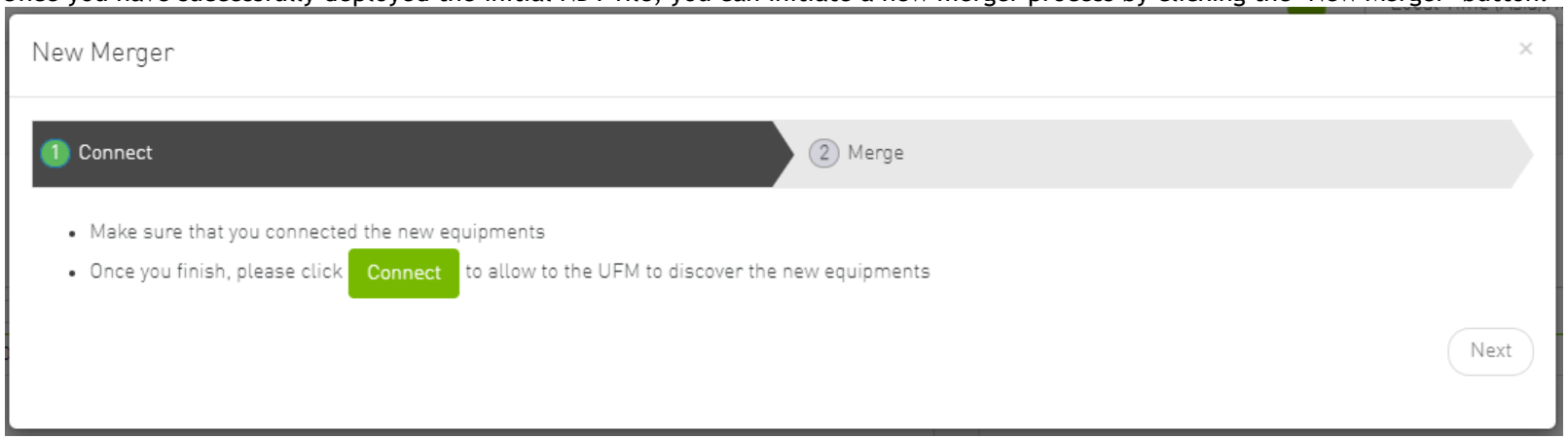

Bring Up Merger
1 Initiate
2 Deploy
Please Select NDT File to start with
Browse
ndt_small_fabric_new.csv
Validate
ndt_small_fabric_new.csv Validation Report - 2023-04-18 15:16:14
Status: Completed with errors
Displayed Columns
Category
Description
Filter...
Missing In Ndt
NEMO-LEAF-2/1 - MF0;r-ufm-sw13/1
Missing In Ndt
NEMO-LEAF-2/11 - r-ufm57 mlx5_0/1
Missing In Ndt
NEMO-LEAF-2/3 - MF0;r-ufm-sw13/3
Missing In Ndt
MF0;r-ufm-sw13:MQM8700/U1/3 - NEMO-LEAF-2/3
Missing In Ndt
MF0;r-ufm-sw13:MQM8700/U1/1 - NEMO-LEAF-2/1
Viewing 1-5 of 5
10
Next

b. Once you are satisfied with the results of the validation in the previous tab, you can proceed to deploy the file.

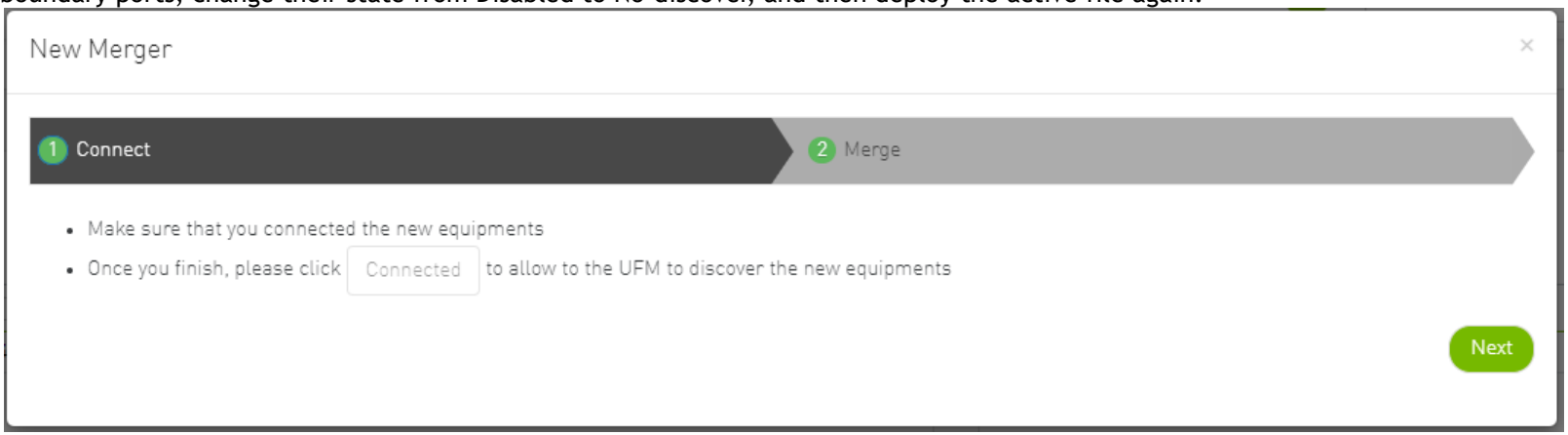


New Subnet Merger

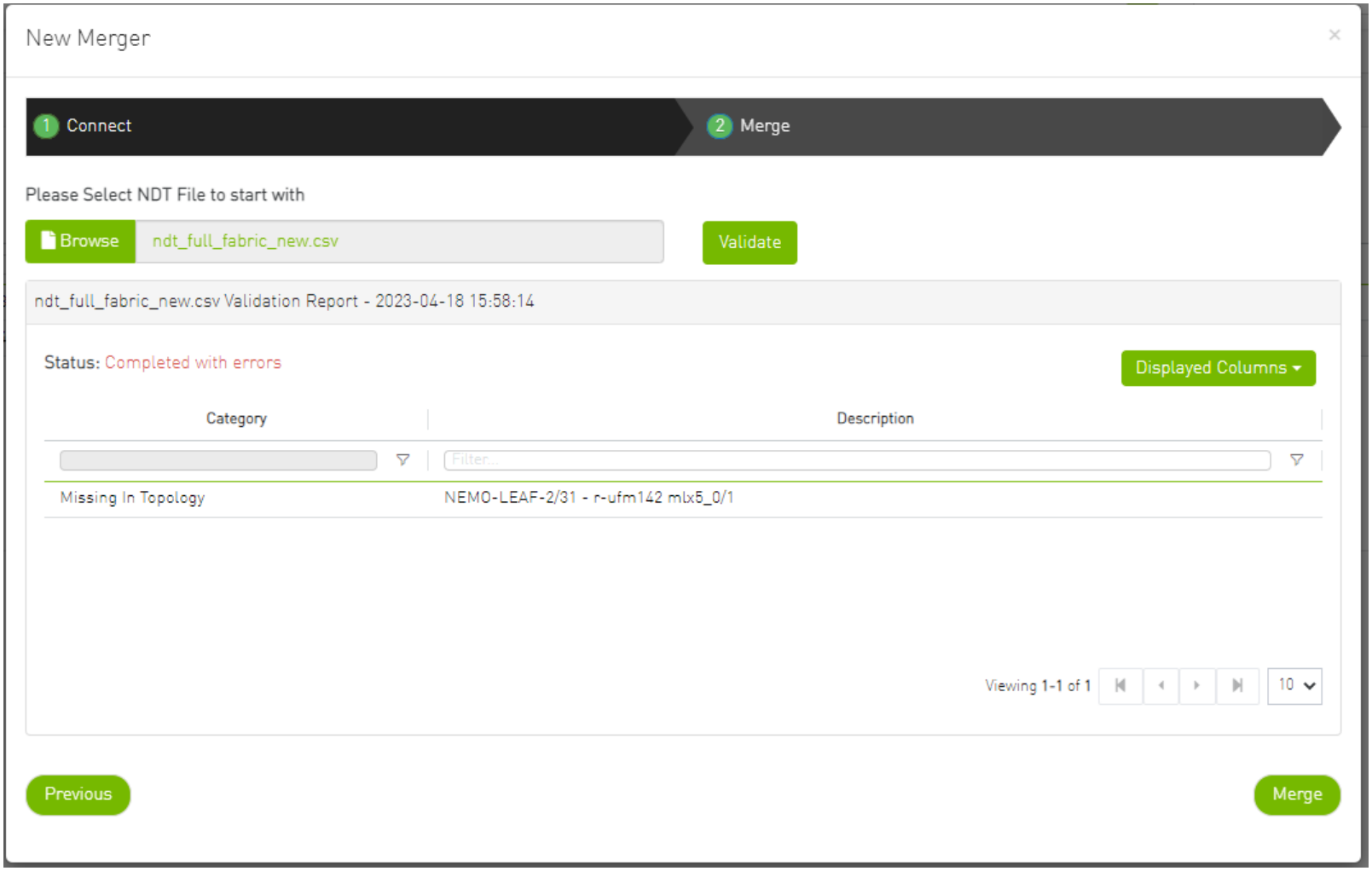
Once you have successfully deployed the initial NDT file, you can initiate a new merger process by clicking the "New Merger" button.



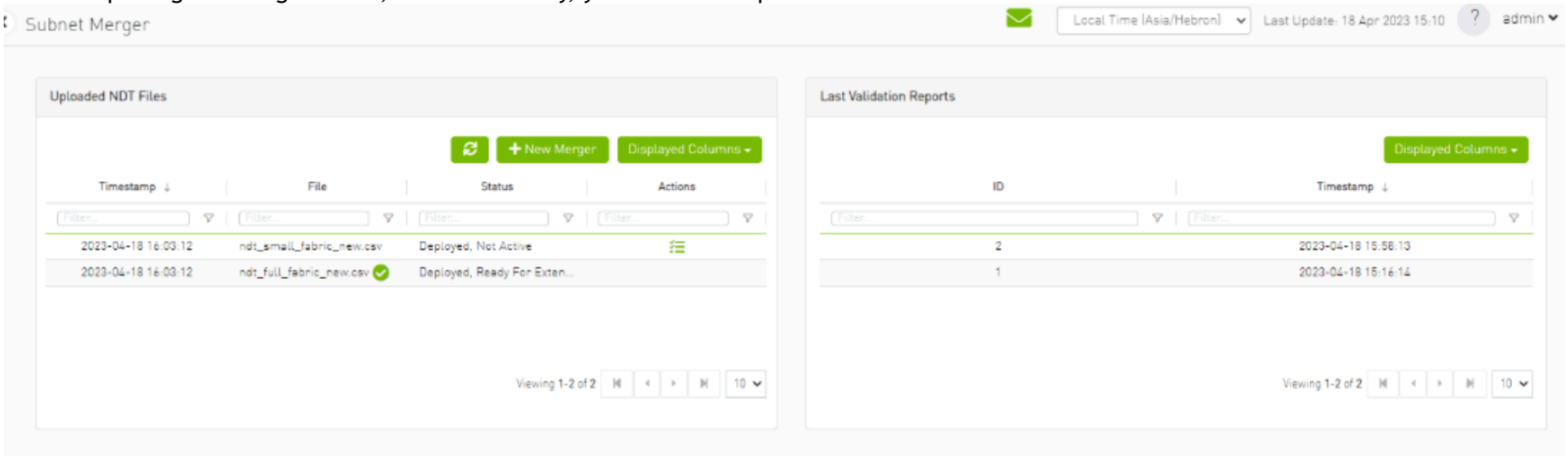
1. "Connect" Tab, it is important to physically connect the new equipment and confirm the connection. Then, click on a button which will open the boundary ports, change their state from Disabled to No-discover, and then deploy the active file again.



2. "Merge" Tab: Once the new equipment is connected and the boundary ports are updated, upload a new NDT file that includes both the current and newly added equipment, along with their boundary ports for future merges. Please note that you cannot merge the file if there are duplicate GUIDs in the report's results.


New Merger
1 Connect
2 Merge
Please Select NDT File to start with
Browse
ndt_full_fabric_new.csv
Validate
ndt_full_fabric_new.csv Validation Report - 2023-04-18 15:58:14
Status: Completed with errors
Displayed Columns
Category
Description
Filter...
Missing In Topology
NEMO-LEAF-2/31 - r-ufm142 mlx5_0/1
Viewing 1-1 of 1
10
Previous
Merge

3. After completing the merge wizard, and if necessary, you can further proceed to extend the IB fabric.



Extending the InfiniBand Setup via Subnet Merger

The following instructions outline the necessary steps for expanding the InfiniBand setup or fabric using subnet merging.

1. Step 1: NDT File Upload (Repeatable)
   Upload the NDT file, performing this action as many times as required, especially when addressing file-related issues.
2. Step 2: NDT File Validation and Verification (Repeatable)
   Validate the NDT file, a process that can be repeated multiple times, particularly after fixing fabric topology or NDT file errors. After initiating this call, you will obtain a validation report ID. The progress of this process is asynchronous, with the report's status initially indicated as "running." Once the report is completed, the status will change to either "Successfully completed" or "Completed with errors."
3. Step 3: Retrieving and Monitoring the Validation Report
   Retrieve the validation report by its corresponding ID, running this step through continuous polling until the report reaches completion.
4. Step 4: Review and Potential Fixes
   Inspect the report and address any necessary fixes to either the NDT file or the topology. Should changes be made to the file, upload the corrected NDT file anew. Alternatively, in case of topology has changed, repeat the verification process.

5. Step 5: Topology Deployment to UFM
   Deploy the verified topology to UFM once you are satisfied with the verification outcomes.
6. Step 6: Adjusting Boundary Ports and Deployment
   Following the physical connection of the setup extension, change the boundary ports' state from "Disabled" to "No-discover."
7. Step 7: Uploading Updated Topoconfig File
   Deploy the updated topoconfig file to the UFM server.
8. Step 8: Next NDT File Upload (Combined Fabric and Extension)
   Upload the next NDT file, which consolidates the current fabric and extension components.
9. Step 9: NDT File Verification
   Conduct the NDT file verification process.
10. Step 10: Reviewing Verification Report
    Review the verification report.
11. Step 11: Addressing Setup or NDT File Issues
    If necessary, make necessary adjustments to the setup or NDT file.
12. Step 12: Final Configuration Deployment
    Once content with the modifications, proceed to deploy the configuration to UFM.
13. Step 13: Iterative Workflow
    Repeat this flow as many times as needed to further the expansion process.

UFM Telemetry FluentD Streaming (TFS) Plugin

Overview

TFS plugin is a self-contained Docker container with REST API support managed by UFM. TFS plugin provides Telemetry counters streaming to FluentD capability. As a fabric manager, the UFM Telemetry holds real-time network telemetry information of the network topology. This information changes over time and is reflected to the telemetry console. In order to do so, we present a stream of the UFM Telemetry data to the FluentD plugin.

Deployment

The following are the possible ways the TFS plugin can be deployed:

1. On UFM Appliance
2. On UFM Software

For complete instructions on deploying the TFS plugin, refer to UFM Telemetry endpoint stream To Fluentd endpoint (TFS).

Authentication

The following authentication types are supported:

- basic (/ufmRest)
- client (/ufmRestV2)
- token (/ufmRestV3)

Rest API

The following REST APIs are supported:

- POST /plugin/tfs/conf
- GET /plugin/tfs/conf
- POST /plugin/tfs/conf/attributes
- GET /plugin/tfs/conf/attributes

For detailed information on interacting with TFS plugin, refer to the NVIDIA UFM Enterprise > Rest API > TFS Plugin REST API.

UFM Events Fluent Streaming (EFS) Plugin

Overview

EFS plugin is a self-contained Docker container with REST API support managed by UFM. EFS plugin extracts the UFM events from UFM Syslog and streams them to a remote FluentD destination. It also has the option to duplicate current UFM Syslog messages and forward them to a remote Syslog destination. As a fabric manager, it will be useful to collect the UFM Enterprise events/logs, stream them to the destination endpoint and monitor them.

Deployment

The following are the ways EFS plugin can be deployed:

1. On UFM Appliance
2. On UFM Software

For detailed instructions on how to deploy EFS plugin, refer to UFM Event Stream to FluentBit endpoint (EFS).

Authentication

The following authentication types are supported:

- basic (/ufmRest)
- client (/ufmRestV2)
- token (/ufmRestV3)

Rest API

The following REST APIs are supported:

- PUT /plugin/efs/conf
- GET /plugin/efs/conf

For detailed information on how to interact with EFS plugin, refer to the NVIDIA UFM Enterprise > Rest API > EFS Plugin REST API.

UFM Bright Cluster Integration Plugin

Overview

The Bright Cluster Integration plugin is a self-contained docker container managed by UFM and is managed by the REST APIs. It enables integrating data from Bright Cluster Manager (BCM) into UFM, providing a more comprehensive network perspective. This integration improves network-centered Root Cause Analysis (RCA) tasks and enables better scoping of workload failure domains.

Deployment

The Bright Cluster Integration plugin can be deployed either on the UFM Appliance or on UFM Software.

For detailed instructions on Bright Cluster Integration plugin deployment, refer to UFM Bright Cluster Integration Plugin.

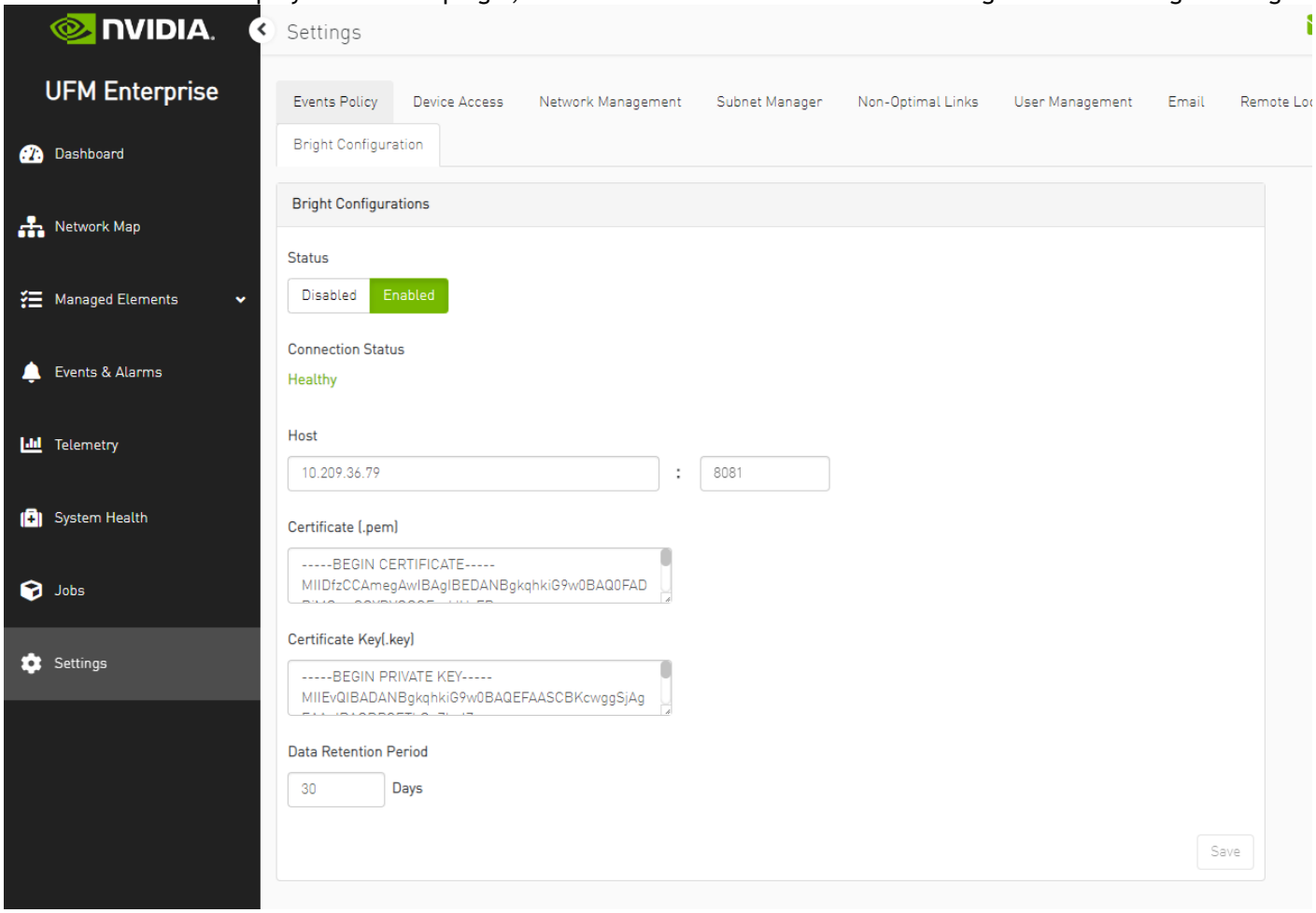
Authentication

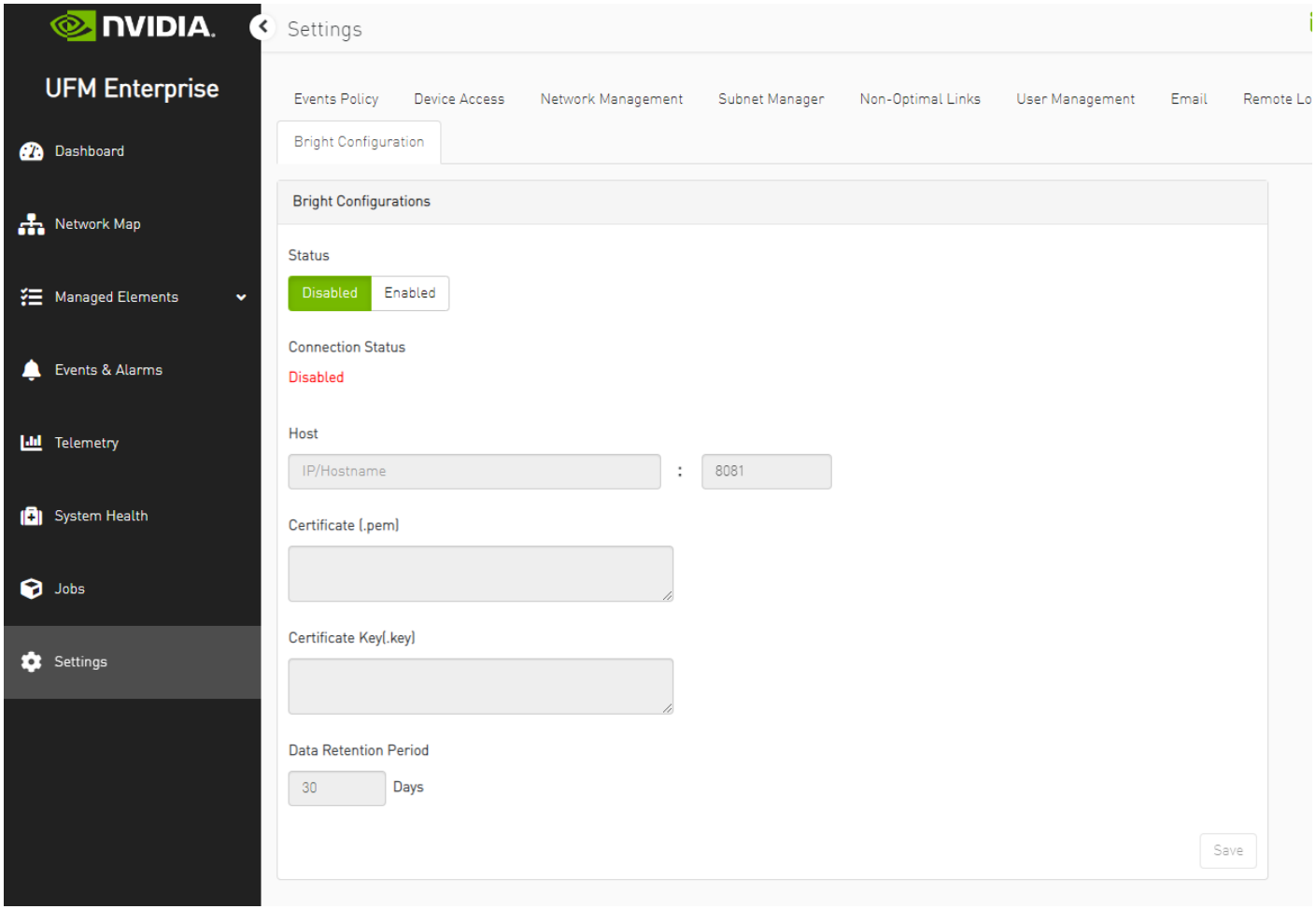
The following authentication types are supported:

- basic (/ufmRest)
- client (/ufmRestV2)
- token (/ufmRestV3)

GUI Screens

1. After the successful deployment of the plugin, a new tab is shown under the UFM settings section for bright configurations management:

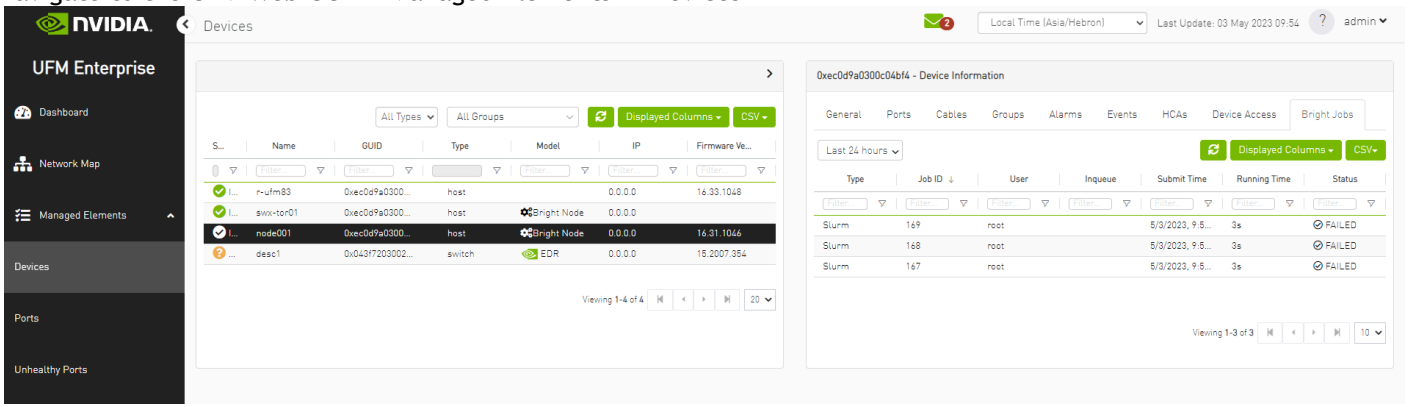




Fill the below required configurations:

Parameter	Description
Host	Hostname or IP of the BCM server
Port	Port of the BCM server, is typically 8081
Certificate	BMC client certificate content that could be located in the BMC server machine under `.cm/XXX.pem`
Certificate key	BMC client certificate key that could be located in the BMC server machine under `.cm/XXX.key`
Data retention period	UFM erases the data gathered in the database after the configured retention period. By default, after 30 days.

2. After you ensure you have successfully completed the plugin configuration, and that you have established a healthy connection with the BMC, navigate to the UFM Web GUI -> Managed Elements -> Devices



Rest API

The following REST APIs are supported:

- PUT plugin/bright/conf
- GET plugin/bright/conf

- GET plugin/bright/data/nodes
- GET plugin/bright/data/jobs

For detailed information on how to interact with bright plugin APIs, refer to NVIDIA UFM Enterprise > Rest API > UFM Bright Cluster Integration Plugin REST API.

UFM Cyber-AI Plugin

Overview

The primary objective of this plugin is to integrate the UFM CyberAI product into the UFM Enterprise WEB GUI. This integration would result in both products being available within a single application.

Deployment

The following are the ways UFM CyberAI plugin can be deployed:

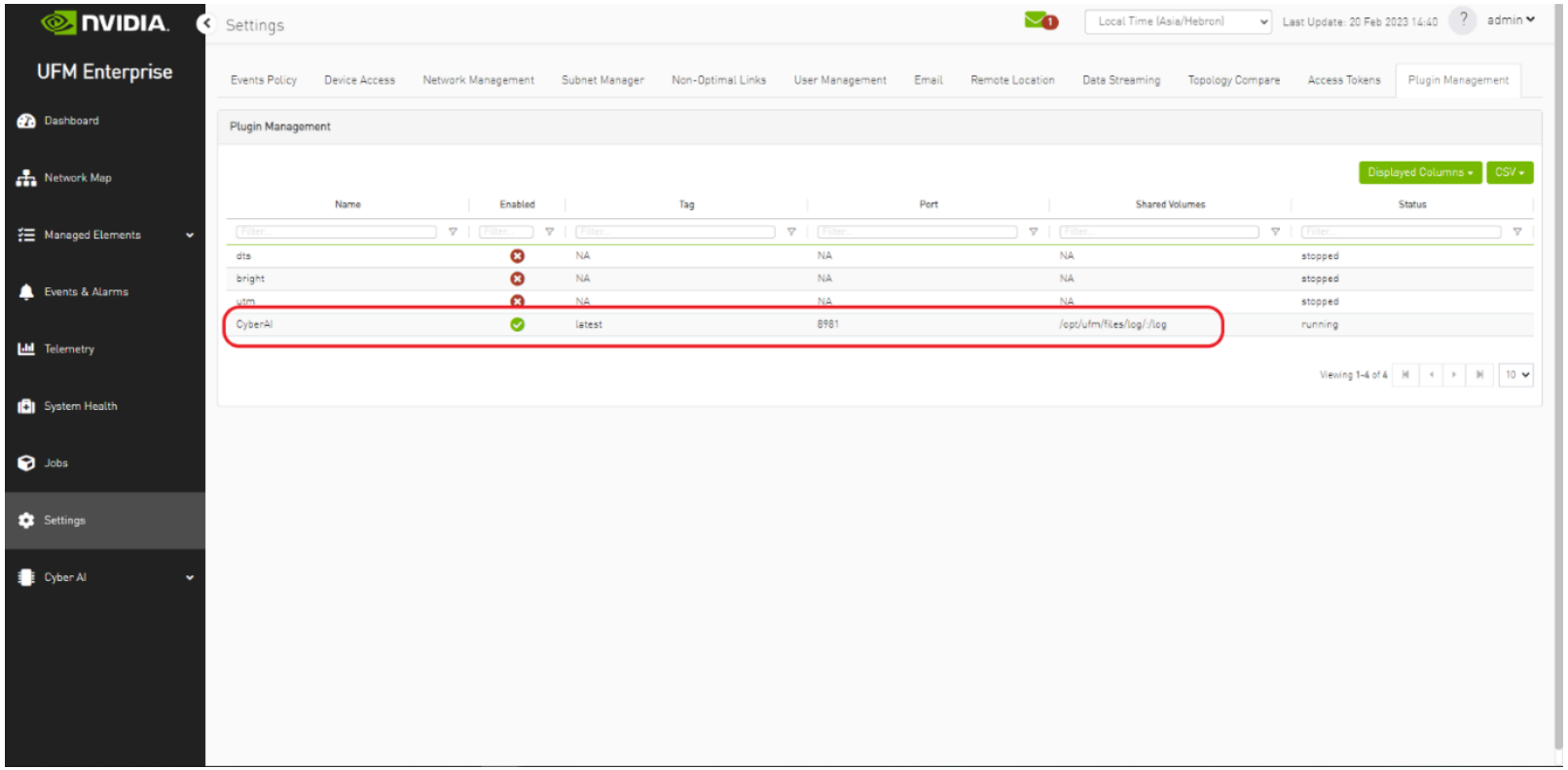
1. On UFM Appliance
2. On UFM Software

First, download the `ufm-plugin-cyberai-image` from the NVIDIA License Portal (NLP), then load the image on the UFM server, using the UFM GUI -> Settings -> Plugins Management tab or by loading the image via the following command:

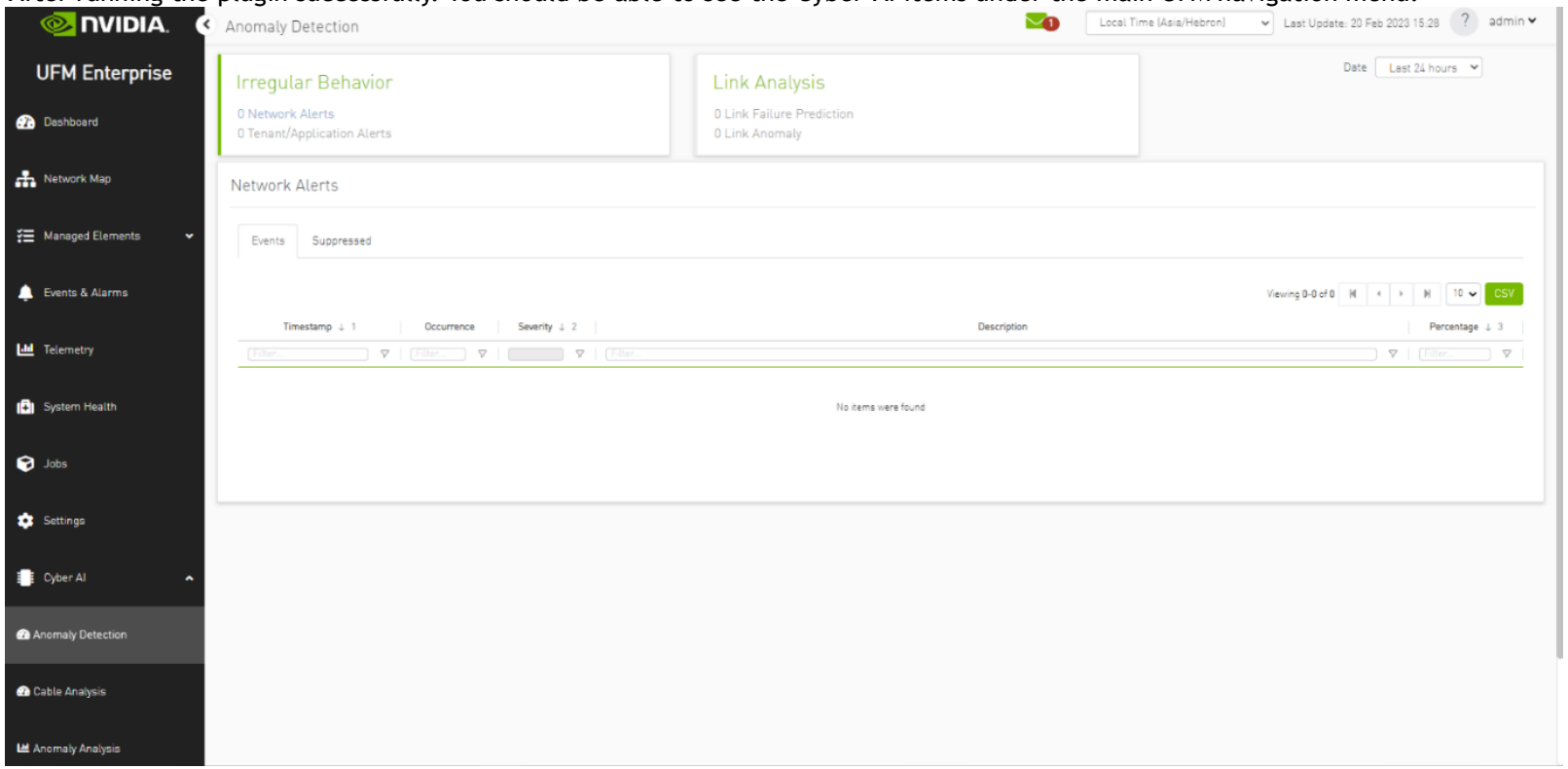
1. Login to the UFM server terminal.
2. Run:

```
docker load -I <path_to_image>
```

Once the plugin's image has been successfully loaded, you can locate the plugin in the Plugins management table within the UFM GUI. You can then run the plugin by right-clicking on the row associated with the plugin.


NVIDIA
UFM Enterprise
Settings
Local Time (Asia/Hebron)
Last Update: 20 Feb 2023 14:40
admin
Events Policy
Device Access
Network Management
Subnet Manager
Non-Optimal Links
User Management
Email
Remote Location
Data Streaming
Topology Compare
Access Tokens
Plugin Management
Dashboard
Network Map
Managed Elements
Events & Alarms
Telemetry
System Health
Jobs
Settings
Cyber AI
Plugin Management
Displayed Columns
CSV
Name
Enabled
Tag
Port
Shared Volumes
Status
dts
NA
NA
NA
stopped
bright
NA
NA
NA
stopped
ufm
NA
NA
NA
stopped
CyberAI
latest
8981
/opt/ufm/files/log/:/log
running
Viewing 1-4 of 4
10

After running the plugin successfully. You should be able to see the Cyber-AI items under the main UFM navigation menu:



For more details, please refer to the UFM Cyber-AI User Manual

Autonomous Link Maintenance (ALM) Plugin

Overview

The primary objective of the Autonomous Link Maintenance (ALM) plugin is to enhance cluster availability and improve the rate of job completion. This objective is accomplished by utilizing machine learning (ML) models to predict potential link failures. The plugin then isolates the expected failing links, implements maintenance procedures on them, and subsequently restores the fixed links to their original state by removing the isolation.

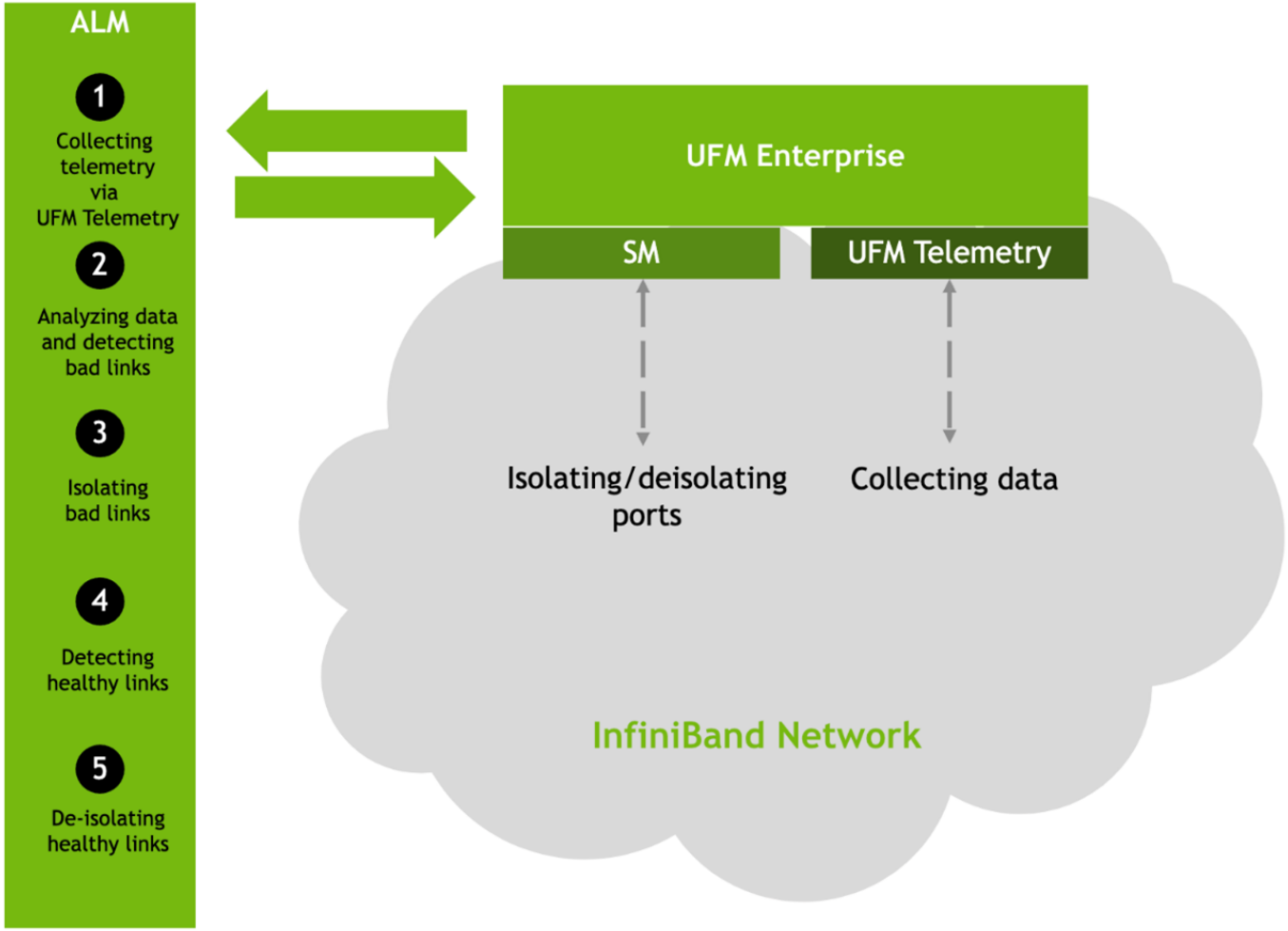
The ALM plugin performs the following tasks:

1. Collects telemetry data from UFM and employs ML jobs to predict which ports need to be isolated/de-isolated
2. Identifies potential link failures and isolates them to avert any interruption to traffic flow
3. Maintains a record of maintenance procedures that can be executed to restore an isolated link
4. After performing the required maintenance, the system verifies if the links can be de-isolated and restored to operational status (brought back online)

The ALM plugin operates in the following two distinct modes:

1. Shadow mode
    - Collects telemetry data, runs ML prediction jobs, and saves the predictions to files.
2. Active mode
    - Collects telemetry data, runs ML prediction jobs, and saves the predictions to files.
    - Automatically isolates and de-isolates based on predictions.
    - It is essential to note that a subset of the links must be specified in the allow list to enable this functionality.

Schematic Flow


ALM
1
Collecting telemetry via UFM Telemetry
2
Analyzing data and detecting bad links
3
Isolating bad links
4
Detecting healthy links
5
De-isolating healthy links
UFM Enterprise
SM
UFM Telemetry
Isolating/deisolating ports
Collecting data
InfiniBand Network

Deployment

The Autonomous Link Maintenance (ALM) plugin can be deployed using the following methods:

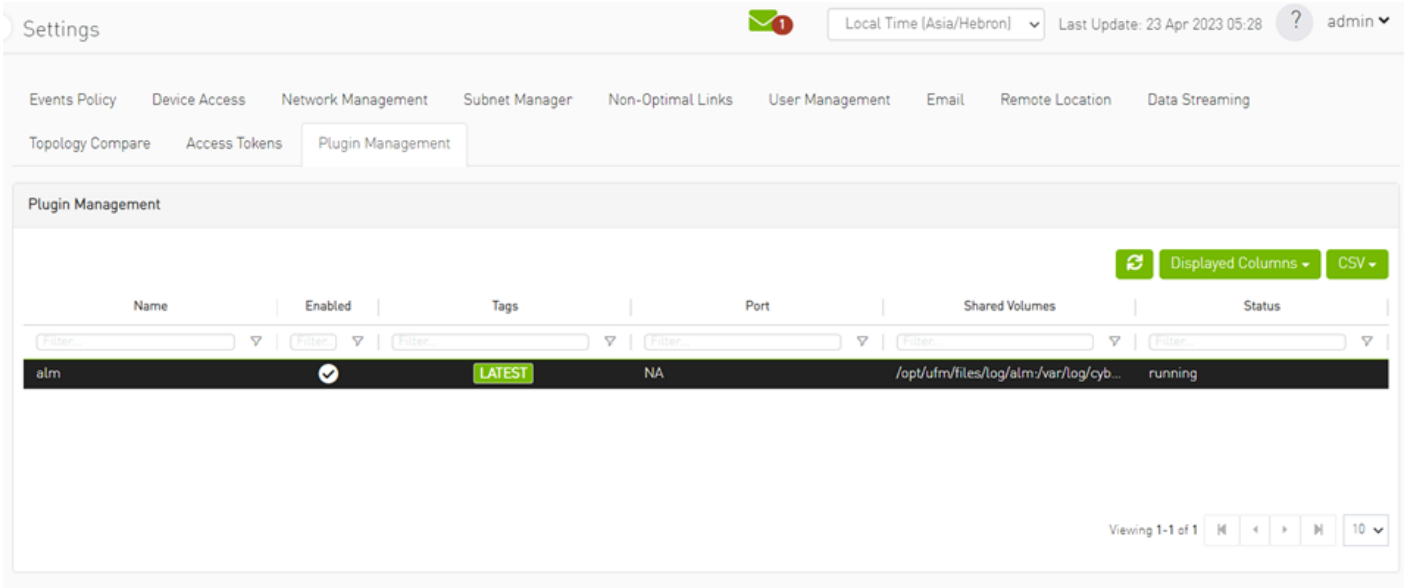
1. On the UFM Appliance
2. On the UFM Software

To deploy the plugin, follow these steps:

1. Download the ufm-plugin-alm-image from the NVIDIA License Portal (NLP).
2. Load the downloaded image onto the UFM server. This can be done either by using the UFM GUI by navigating to the Settings -> Plugins Management tab or by loading the image via the following instructions:
3. Log in to the UFM server terminal.
4. Run:

```
docker load -I <path_to_image>
```

5. After successfully loading the plugin image, the plugin should become visible within the plugins management table within the UFM GUI. To initiate the plugin's execution, simply right-click on the respective in the table.



Data Collection

The ALM plugin collects data from the UFM Enterprise appliance in the following two methods:

1. Low-frequency collection: This process occurs every 0 minutes and gathers data for the following counter: hist0, hist1, hist2, hist3, hist4, phy_effective_errors, phy_symbol_errors
2. High-frequency collection : This process occurs every 10 seconds and gathers data for the following counters: phy_state,logical_state,link_speed_active,link_width_active,fec_mode_active, raw_ber,eff_ber,symbol_ber,phy_raw_errors_lane0,phy_raw_errors_lane1,phy_raw_errors_lane2, phy_raw_errors_lane3,phy_effective_errors,phy_symbol_errors,time_since_last_clear, hist0,hist1,hist2,hist3,hist4,switch_temperature,CableInfo.temperature,link_down_events, plr_rcv_codes,plr_rcv_code_err,plr_rcv_uncorrectable_code,plr_xmit_codes,plr_xmit_retry_codes, plr_xmit_retry_events,plr_sync_events,hi_retransmission_rate,fast_link_up_status, time_to_link_up,status_opcode,status_message,down_blame,local_reason_opcode,

remote_reason_opcode,e2e_reason_opcode,num_of_ber_alarams,PortRcvRemotePhysicalErrorsExtended,
PortRcvErrorsExtended,PortXmitDiscardsExtended,PortRcvSwitchRelayErrorsExtended,PortRcvConstraintErrorsExtended,
VL15DroppedExtended,PortXmitWaitExtended,PortXmitDataExtended,PortRcvDataExtended,PortXmitPktsExtended,
PortRcvPktsExtended,PortUniCastXmitPktsExtended,PortUniCastRcvPktsExtended,PortMultiCastXmitPktsExtended,PortMultiCastRcvPktsExtended

3. The collected counters can be configurable and customized to suit your requirements. The counters can be found at /opt/ufm/conf/plugins/alm/counters.cfg

```
root@r-ufm116:~# cat  /opt/ufm/conf/plugins/alm/counters.cfg
[HighFreq]
phy_state = last_update_value
logical_state = last_update_value
link_speed_active = last_update_value
link_width_active = last_update_value
fec_mode_active = last_update_value
raw_ber = last_update_value
eff_ber = last_update_value
symbol_ber = last_update_value
phy_raw_errors_lane0 = delta
phy_raw_errors_lane1 = delta
phy_raw_errors_lane2 = delta
phy_raw_errors_lane3 = delta
phy_effective_errors = delta
phy_symbol_errors = delta
time_since_last_clear = last_update_value
hist0 = delta
hist1 = delta
hist2 = delta
hist3 = delta
hist4 = delta
switch_temperature = last_update_value
CableInfo.Temperature = last_update_value
link_down_events = delta
plr_rcv_codes = delta
plr_rcv_code_err = delta
plr_rcv_uncorrectable_code = delta
plr_xmit_codes = delta
plr_xmit_retry_codes = delta
plr_xmit_retry_events = delta
plr_sync_events = delta
hi_retransmission_rate = delta
fast_link_up_status = last_update_value
time_to_link_up = last_update_value
status_opcode = last_update_value
status_message = last_update_value
down_blame = last_update_value
local_reason_opcode = last_update_value
remote_reason_opcode = last_update_value
e2e_reason_opcode = last_update_value
num_of_ber_alarams = delta
PortRcvRemotePhysicalErrorsExtended = delta
PortRcvErrorsExtended = delta
PortXmitDiscardsExtended = delta
PortRcvSwitchRelayErrorsExtended = delta
```

ALM Configuration

The ALM configuration is used for controlling isolation/de-isolation. The configuration can be found under `/opt/ufm/cyber-ai/conf/cyberai.cfg`.

Name	Section name	Description
`mode`	CyberAi	The mode can be active or shadow The active mode means the alm will apply isolation/deisolation rule omn all ports exceptin the port in the expect list And the shadow mode mean the alm will apply isolation/deisolation rules on the ports on the except list The mode can be either "active" or "shadow." In active mode, the ALM will enforce isolation/deisolation rules on all ports except those listed in the "expect" list. In shadow mode, the ALM will enforce isolation/deisolation rules on the ports listed in the "except" list.
`except_list`	CyberAi	Includes the ports that receive the opposite treatment compared to the mode. Format: portguid_number, portguid_portnumber2
`max_per_hour`	Isolation	The maximum number of ports that can be isolated in a hour
`max_per_week`	Isolation	Maximum number of ports that can be isolated in a week
`max_per_month`	Isolation	Maximum number of the ports that can be isolated in a month
`Deisolation_time`	DeIsolation	The waiting time before deisolate the isolated port
`max_per_hour`	DeIsolation	The maximum number of deisolated port per hour

Name	Section name	Description
absolute_threshold_of_isolated_ports	Isolation	The maximum number of ports than can be isolated in one sample

ALM Jobs

The table presented below displays the names and descriptions of ALM jobs. These jobs are designed to predict the ports that require isolation/de-isolation. Upon enabling the ALM plugin, these ALM jobs run periodically.

ALM Job Name	Description	Frequency
Port_hist	By using the low frequency bit error histogram counters, the ALM job identifies the ports that will be monitored at high frequency in the next time interval. The job generates an output file that is later read by the high frequency telemetry monitoring job. It prioritizes links that are more susceptible to failure.	600 seconds
Low_freq_predict	Predicts the likelihood of a port failure by analyzing input data from low frequency telemetry, while only utilizing physical layer counters. The prediction works for isolated ports as well. The resulting output from this task serves as a critical input for determining whether to isolate or de-isolate ports.	10 seconds

DTS Plugin

Overview

The DTS Monitor can be run either as a standalone tool or as a plugin within UFM. It collects all the endpoint information for DPUs and consolidates it into a single interface.

Deployment

DPU Requirements

- OS: ubuntu 20/22
- BlueField: BlueField-2 or BlueField-3
- DTS: version > 1.12
- DPE service up and running
- yaml configured with "DTS_CONFIG_DIR=ufm"
    - Add to the following line in `file doca_telemetry_standalone.yaml`
        - Command:

```
/bin/bash", "-c", "/usr/bin/telemetry-init.sh && /usr/bin/enable-fluent-forward.sh
```

        - Command:

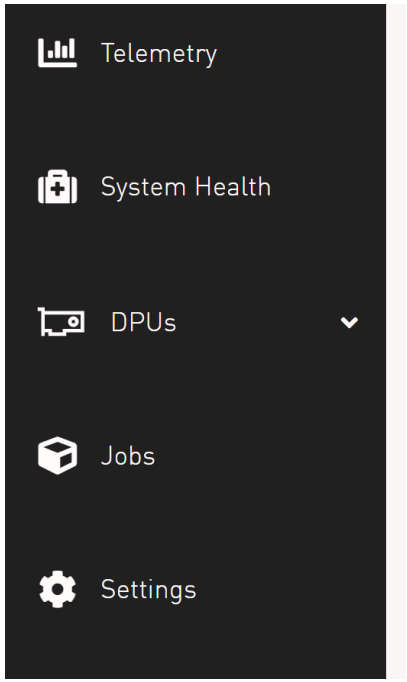
```
/bin/bash", "-c", " DTS_CONFIG_DIR=ufm /usr/bin/telemetry-init.sh && /usr/bin/enable-fluent-forward.sh
```

Installation

you need to load the image on the UFM server; either using the UFM GUI -> Settings -> Plugins Management tab or by loading the image via the following command:

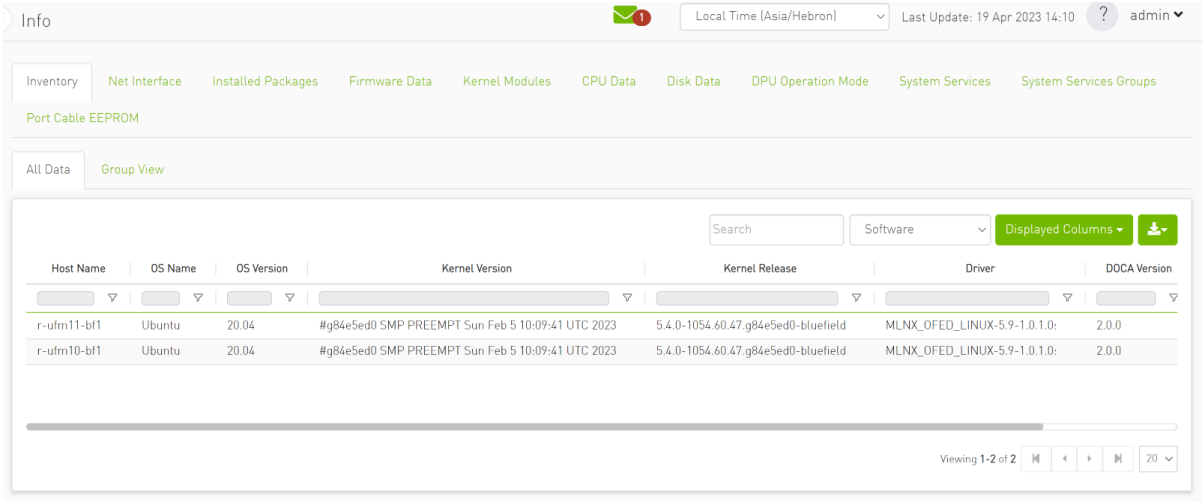
1. Login to the UFM server terminal.
2. Run: *docker load -I <path_to_image>*

After completing the plugin addition and refreshing the UFM GUI, a new menu item, titled DPUs, will be added to the left navigation bar.

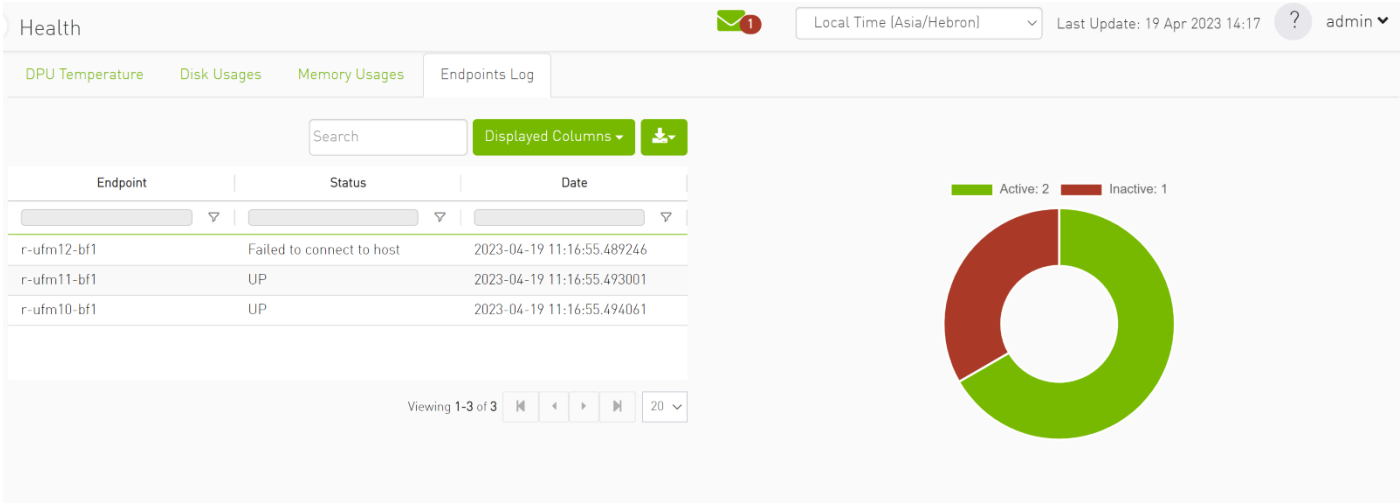

Telemetry
System Health
DPUs
Jobs
Settings

GUI Screens

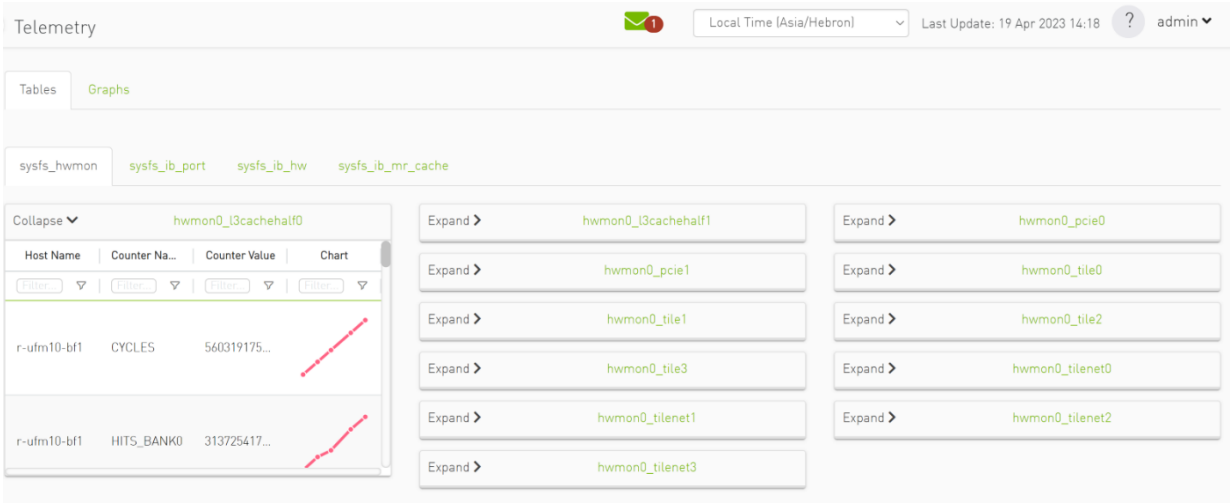
Info



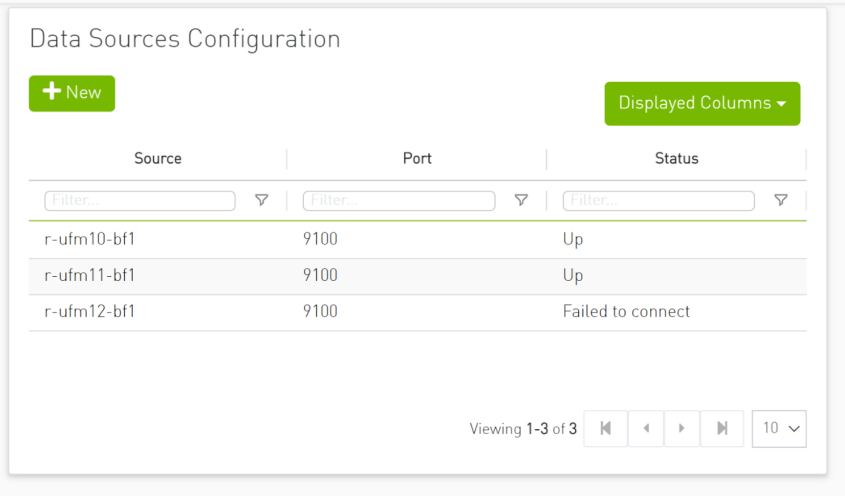
Health


Health
Local Time (Asia/Hebron)
Last Update: 19 Apr 2023 14:17
admin
DPU Temperature
Disk Usages
Memory Usages
Endpoints Log
Search
Displayed Columns
Endpoint
Status
Date
r-ufm12-bf1
Failed to connect to host
2023-04-19 11:16:55.489246
r-ufm11-bf1
UP
2023-04-19 11:16:55.493001
r-ufm10-bf1
UP
2023-04-19 11:16:55.494061
Viewing 1-3 of 3
20
Active: 2
Inactive: 1

Telemetry


Telemetry
Local Time (Asia/Hebron)
Last Update: 19 Apr 2023 14:18
admin
Tables
Graphs
sysfs_hwmon
sysfs_ib_port
sysfs_ib_hw
sysfs_ib_mr_cache
Collapse
hwmon0_l3cachehalf0
Host Name
Counter Na...
Counter Value
Chart
r-ufm10-bf1
CYCLES
560319175...
r-ufm10-bf1
HITS_BANK0
313725417...
Expand
hwmon0_l3cachehalf1
hwmon0_pcie1
hwmon0_tile1
hwmon0_tile3
hwmon0_tilenet1
hwmon0_tilenet3
hwmon0_pcie0
hwmon0_tile0
hwmon0_tile2
hwmon0_tilenet0
hwmon0_tilenet2

Data Sources



GRPC-Streamer Plugin

Authentication

The following authentication types are supported:

- Basic (/ufmRest)
- Token (/ufmRestV3)

Create a Session to UFM from GRPC

Description: Creates a session to receive REST API results from the UFM's GRPC server. After a stream or one call, the session is deleted so the server would not save the authorizations.

- Call: CreateSession in the grpc
- Request Content Type - message SessionAuth
- Request Data:

```
message SessionAuth{
  string job_id=1;
  string username = 2;
  string password = 3;
  optional string token = 4;
}
```

- Job_id - The unique identifier for the client you want to have
- Username - The authentication username
- Password - The authentication password
- Token - The authentication token
- Response:

```
message SessionRespond{
  string respond=1;
}
```

- Respond types:
  - Success - Ok.
  - ConnectionError - UFM connection error (bad parameters or UFM is down).
  - Other exceptions - details sent in the respond.
- Console command:

```
client session --server_ip=server_ip --id=client_id --auth=username,password --token=token
```

Create New Subscription

- Description: Only after the server has established a session for this grpc client,add all the requested REST APIs with intervals and delta requests.
- Call: AddSubscriber
- Request Content Type - Message SubscriberParams
- Request Data:

```
message SubscriberParams{
  message APIParams {
    string ufm_api_name = 1;
    int32 interval = 2;
    optional bool only_delta = 3;
  }
  string job_id = 1;
  repeated APIParams apiParams = 2;
}
```

- Job_id - A unique subscriber identifier
- apiParams - The list of apiParams from the above message above:
  - ufm_api_name - The name from the known to server request api list
  - interval - The interval between messages conducted in a stream run. Presented in seconds.
  - only_delta - Receives the difference between the previous messages in a stream run.
- Response content type:

```
message SessionRespond{
  string respond=1;
}
```

- Respond Types:
  - Created a user with session and added new IP- Ok.
  - Cannot add subscriber that do no have an established session - need to create a session before creating subscriber.
  - The server already have the ID - need to create new session and new subscriber with a new unique ID.
- Console command:

```
client create --server_ip=localhost --id=client_id --apis=events;40;True,links,alarms;10
```

The API's list is separated by commas, and each modifier for the REST API is separated by a semi comma.

If the server is not given a modifier, default ones are used (where only_delta is False and interval is based on the API).

Edit Known Subscription

- Description: Changes a known IP. Whether the server has the IP or not.
- Call: AddSubscriber
- Request Content Type - Message SubscriberParams
- Request Data:

```
message SubscriberParams{
  message APIParams {
    string ufm_api_name = 1;
    int32 interval = 2;
    optional bool only_delta = 3;
  }
  string job_id = 1; //unique identifier for this job
  repeated APIParams apiParams = 2;
}
```

- Job_id - The subscriber unique identifier
- apiParams - A list of apiParams from the above message.
    - ufm_api_name - name from the known to server request api list
    - interval - The interval between messages conducted in a stream run. Presented in seconds.
    - only_delta - Receives the difference between the previous messages in a stream run.
- Response content type:

```
message SessionRespond{
  string respond=1;
}
```

- Respond Types:
    - Created user with new IP- Ok.
    - Cannot add subscriber without an established session – need to create a session before creating subscriber.
    - Cannot add subscriber illegal apis – cannot create subscriber with empty API list, call again with correct API list.

Get List of Known Subscribers

- Description: Gets the list of subscribers, including the requested list of APIs.
- Call: ListSubscribers
- Request Content Type: google.protobuf.Empty
- Response:

```
message ListSubscriberParams{
  repeated SubscriberParams subscribers = 1;
}
```

- Console command: server subscribes --server_ip=server_ip

Delete a Known Subscriber

- Description: Deletes an existing subscriber and removes the session.
- Call: DeleteSubscriber
- Request Content Type: Message gRPCStreamerID
- Request Data:

```
message gRPCStreamerID{
string job_id = 1;
}
```

- Response:protobuf.Empty

Run a Known Subscriber Once

- Description: Runs the Rest API list for a known subscriber once and returns the result in message runOnceRespond, and then delete the subscriber's session.
- Call: RunOnceJob
- Request Content Type: Message gRPCStreamerID
- Request Data:

```
message gRPCStreamerID{
string job_id = 1;
}
```

- Response content type:

```
message runOnceRespond{
  string job_id=1;
  repeated gRPCStreamerParams results = 2;
}
```

- Job_id- The first message unique identifier.
- Results – list of gRPCStreamerParams contains results from each REST API
- Responses:
  - Job id - Cannot run a client without an established session. Empty results – an existing session for this client is not found, and the client is not known to the server.
  - Job id - Cannot run the client without creating a subscriber. Empty results – a session was created for the client but the subscription is not created.
  - Job_id - Cannot connect to the UFM. empty result – the GRPC server cannot connect to the UFM machine and receive empty results, because it cannot create a subscriber with an empty API list. This means that the UFM machine is experiencing a problem.
  - Job_id - The first unique message identifier of the messages. Not empty results – Ok
- Console command:

```
client once_id --server_ip=server_ip --id=client_id
```

Run Streamed Data of a Known Subscriber

- Description: Run a stream of results from the Rest API list for a known Subscriber and return the result as interator, where each item is message gRPCStreamerParams. at the end, delete the session.
- Call: RunStreamJob
- Request Content Type: Message gRPCStreamerID
- Request Data:

```
message gRPCStreamerID{
string job_id = 1;
}
```

- Response content type: iterator of messages gRPCStreamerParams

```
message gRPCStreamerParams{
  string message_id = 1; // unique identifier for messages
  string ufm_api_name = 2; // what rest api receive the data from
  google.protobuf.Timestamp timestamp = 3; //what time we created the message, can be converted to Datetime
  string data = 4; // data of rest api call
}
```

- Response:
  - One message only containing "Cannot run a client without a session" - A session has not been established
  - No message - A session and/or a subscriber with this ID does not exist.
  - Messages with interval between with the modifiers - Ok
- Console command:

```
client stream_id --server_ip=server_ip --id=client_id
```

Run a New Subscriber Once

- Description: After ensuring that a session for this specific job ID is established, the server runs the whole REST API list for the new subscriber once and returns the following result in message `runOnceRespond`. This action does not save the subscribe ID or the established session in the server.
- Call: RunOnce
- Request Content Type: Message SubscriberParams
- Request Data:

```
message SubscriberParams{
  message APIParams {
    string ufm_api_name = 1;
    int32 interval = 2;
    optional bool only_delta = 3;
  }
  string job_id = 1; //unique identifier for this job
  repeated APIParams apiParams = 2;
}
```

- Response content type:

```
message runOnceRespond{
  string job_id=1;
  repeated gRPCStreamerParams results = 2;
}
```

- Responses:
  - Job id = Cannot run a client without an established session. Empty results - no session for this client.
  - Job_id = 0 - The GRPC server cannot connect to the UFM machine and receive empty results, or it cannot create a subscriber with an empty API list.
  - Job_id = The messages' first unique identifier, and not an empty result - Ok.
- Console command:

```
client once --server_ip=server_ip --id=client_id --auth=username,password --token=token --apis=events;40;True
,links;20;False,alarms;10
```

- The console command creates a session for this specific client.
- A token or the basic authorization is needed, not both.

Run New Subscriber Streamed Data

- Description: After the server checks it has a session for this job ID, Run a stream of results from the Rest API list for a new Subscriber and return the result as interator, where each item is message gRPCStreamerParams. at the end, delete the session.
- Call: RunPeriodically
- Request Content Type: Message SubscriberParams
- Request Data:

```
message SubscriberParams{
  message APIParams {
    string ufm_api_name = 1;
    int32 interval = 2;
    optional bool only_delta = 3;
  }
  string job_id = 1; //unique identifier for this job
  repeated APIParams apiParams = 2;
}
```

- Response content type: iterator of messages gRPCStreamerParams
- Response:
  - Only one message with data equals to Cant run client without session – no session
  - Messages with intervals between with the modifiers – Ok
- Console command:

```
client stream --server_ip=server_ip --id=client_id --auth=username,password --token=token --apis=events;40;True,li
nks;20;False,alarms;10
```

- console command also create session for that client.

- no need for both token and basic authorization, just one of them.

Run A Serialization on All the Running Streams

- Description: Run a serialization for each running stream. The serialization will return to each of the machines the results from the rest api list.
- Call: Serialization
- Request Content Type: google.protobuf.Empty
- Response: google.protobuf.Empty

Stop a Running Stream

- Description: Cancels running stream using the client id of the stream and stop it from outside, If found stop the stream.
- Call: StopStream
- Request Content Type: Message gRPCStreamerID
- Request Data:

```
message gRPCStreamerID{
string job_id = 1;
}
```

- Response: google.protobuf.Empty

Run a subscribe stream

- Description: Create a subscription to a client identifier, all new messages that go to that client, will be copied and also sent to this stream.
- Call: Serialization
- Request Content Type: message gRPCStreamerID
- Response: iterator of messages gRPCStreamerParams

```
message gRPCStreamerParams{
  string message_id = 1; // unique identifier for messages
  string ufm_api_name = 2; // what rest api receive the data from
  google.protobuf.Timestamp timestamp = 3; //what time we created the message, can be converted to Datetime
  string data = 4; // data of rest api call
}
```

- the identifier may or may not be in the grpc server.
- Cannot be stop streamed using StopStream.
- Console command:

```
client subscribe --server_ip=server_ip --id=client_id
```

Get the variables from a known subscriber

- Description: Get the variables of known subscriber if found, else return empty variables.
- Call: GetJobParams
- Request Content Type: message gRPCStreamerID
- Response:

```
message SubscriberParams{
  message APIParams {
    string ufm_api_name = 1; //currently the list of api from ufm that are supported are [Jobs, Events,
Links, Alarms]
    int32 interval = 2;
    optional bool only_delta = 3;
  }
  string job_id = 1; //unique identifier for this job
  repeated APIParams apiParams = 2;
}
```

Get Help / Version

- Description: Get help and the version of the plugin, how to interact with the server. What stages need to be done to extract the rest apis (Session>run once/stream or Session>AddSubscriber>once_id/stream_id)
- Call: Help or Version
- Request Content Type: google.protobuf.Empty
- Response:

```
message SessionRespond{
  string respond=1;
}
```

Sysinfo Plugin

Overview

The Sysinfo plugin is a Docker container that is managed by UFM and comes with REST API support. Its purpose is to allow users to run commands and extract information from managed switches. This feature enables users to schedule runs at regular intervals and execute commands on switches directly from UFM.

The plugin takes care of managing sessions to the switches and can extend them if necessary. It also enables users to send both synchronous and asynchronous commands to all the managed switches. Additionally, it can intersect the given switches with the running UFM to ensure that only those switches that are on the UFM are activated.

Deployment

The following are the possible ways plugin plugin can be deployed:

1. On UFM Appliance
2. On UFM Software.

3. Authentication

Following authentication types are supported:

- basic (/ufmRest)
- client (/ufmRestV2)
- token (/ufmRestV3)

REST API

The following REST APIs are supported:

- GET /help
- GET /version
- POST /query
- POST /update
- POST /cancel
- POST /delete

Sysinfo Query Format

The Sysinfo plugin is responsible for extracting basic data needed to create a query. This is done using the following five fields:

1. Switches - An array of switch IP addresses. If this field is left empty, the plugin will gather all switches from the running UFM.
2. Callback - The URL location to which the answers should be sent.
3. Commands - An array of commands that need to be executed.
4. Schedule_run - An optional field used to set intervals for running the commands. The interval can be specified in seconds and can be set to run until a certain duration or end time. The start time can also be controlled.

There are additional flags for a configurable query:

- `ignore_ufm=True`: Does not check the UFM for switches or intersect it with given switches
- `username`: Overrides the switches' default username
- `password`: Overrides the switches' default password

- `is_async`: Rather than attempting to execute all commands simultaneously at the switch, the commands are executed one after the other in sequence.
- `one_by_one=False`:Instead of sending results from each switch as soon as information is obtained, all data is sent at once to the callback. This change eliminates multiple small sends and replaces them with a single large send.

For detailed information on how to interact with Sysinfo plugin, refer to the NVIDIA UFM Enterprise > Rest API > Sysinfo Plugin REST API.

SNMP Plugin

The SNMP plugin is a self-contained Docker container that includes REST API support and is managed by UFM. Its primary function is to receive SNMP traps from switches and forward them to UFM as external events. This feature enhances the user experience by providing additional information about switches in the InfiniBand fabric via UFM events and alarms.

Deployment

There are two potential deployment options for the SNMP plugin:

- On UFM Appliance
- On UFM Software

For detailed instructions on how to deploy the SNMP plugin, refer to this page.

Authentication

The following authentication types are supported:

- basic (/ufmRest)
- client (/ufmRestV2)
- token (/ufmRestV3)

REST API

The following REST API are supported:

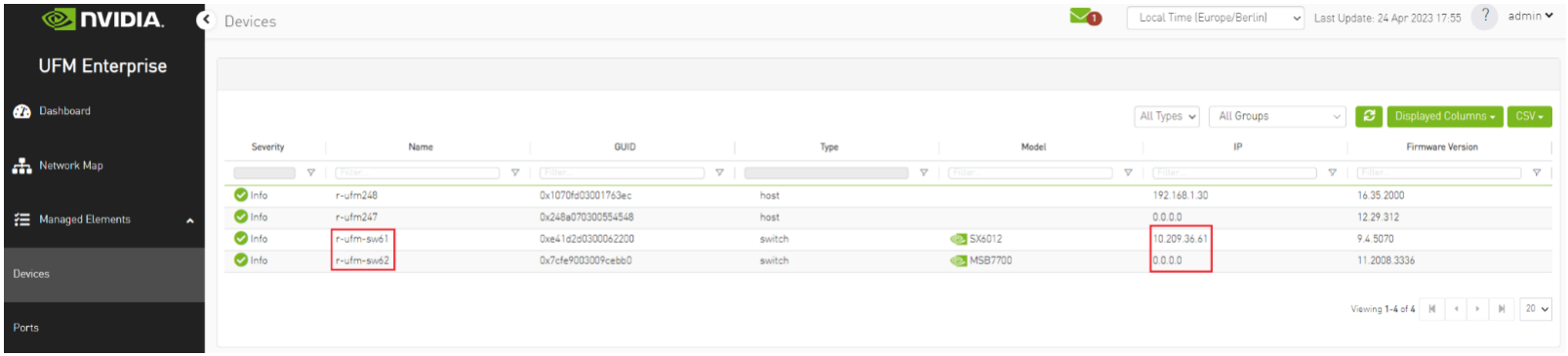
- GET /switch_list
- GET /trap_list
- POST /register
- POST /unregister
- POST /enable_trap
- POST /disable_trap
- GET /version

For more information, please refer to UFM Enterprise Documentation → UFM REST API → SNMP Plugin REST API.

Usage

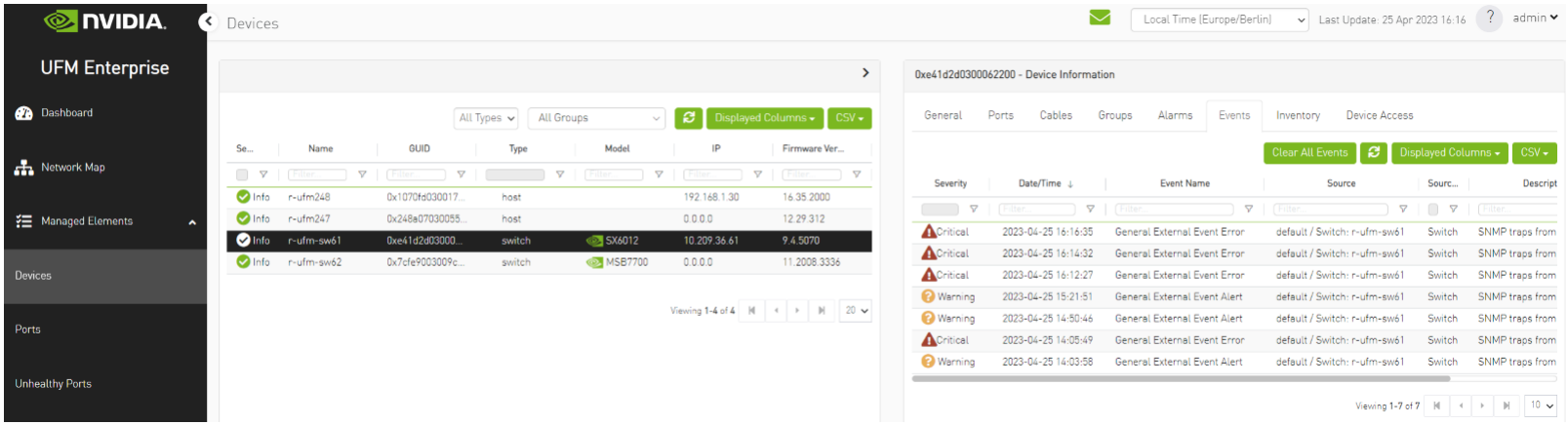
By default, upon initialization, the SNMP plugin captures traps from all switches within the fabric. However, this behavior can be modified through configuration settings utilizing the "snmp_mode" option, with available values of "auto" or "manual".

It is important to ensure that the switch is visible to UFM and has a valid IP address. As illustrated in the following example, switch traps will only be received from "r-ufm-sw61".



The following is an instance of a trap received by the SNMP plugin and displayed as a UFM event:

Additionally, there is an option to verify events/alarms for a particular switch:



The SNMP plugin performs a periodic check of the fabric every 180 seconds, allowing for prompt receipt of traps from new switches or updated IP addresses of existing switches in under 180 seconds. This interval may be adjusted via the "ufm_switches_update_interval" option. To manually register or unregister a switch, please refer to the UFM Enterprise Documentation → UFM REST API → SNMP Plugin REST API.

The SNMP plugin employs the most up-to-date SNMP v3 protocol, which incorporates advanced security measures such as authentication and encryption. The "snmp_version" option enables the selection of SNMP versions “1” or “3”. It is essential to note that only switch-exposed traps will be transmitted to UFM as events.

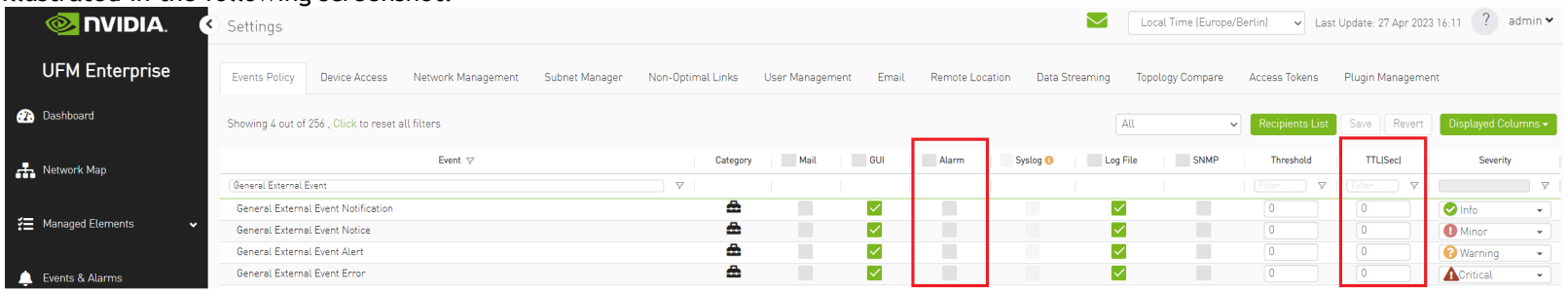
OID	Name	Description	Status	Severity
MELLANOX-EFM-MIB::testTrap	send-test	A test trap ordered by the system administrator	Enabled	Warning
MELLANOX-EFM-MIB::asicChipDown	asic-chip-down	ASIC (Chip) Down	Enabled	Critical
MELLANOX-EFM-MIB::cpuUtilHigh	cpu-util-high	CPU utilization has risen too high	Enabled	Warning
MELLANOX-EFM-MIB::diskSpaceLow	disk-space-low	Filesystem free space has fallen too low	Enabled	Warning
MELLANOX-EFM-MIB::expectedShutdown	expected-shutdown	Expected system shutdown	Enabled	Info
MELLANOX-EFM-MIB::systemHealthStatus	health-module-status	Health module Status	Enabled	Critical
MELLANOX-EFM-MIB::insufficientFans	insufficient-fans	Insufficient amount of fans in system	Enabled	Warning
MELLANOX-EFM-MIB::insufficientFansRecover	insufficient-fans-recover	Insufficient amount of fans in system recovered	Enabled	Info
MELLANOX-EFM-MIB::insufficientPower	insufficient-power	Insufficient power supply	Enabled	Warning
RFC1213::linkdown	interface-down	An interface's link state has changed to down	Enabled	Minor
RFC1213::linkup	interface-up	An interface's link state has changed to up	Enabled	Info
MELLANOX-EFM-MIB::unexpectedShutdown	unexpected-shutdown	Unexpected system shutdown	Enabled	Minor
SNMPv2-MIB::coldStart	cold-start	SNMP entity reinitialized	Enabled	Info

To learn more about how to enable or disable a specific trap, please refer to the UFM Enterprise Documentation → UFM REST API → SNMP Plugin REST API.

If some traps are not included in the default list, they may be added using the "snmp_additional_traps" option. The SNMP plugin will consider these traps as "enabled" and transmit them to UFM as events with an "Info" severity level.

To ensure the uninterrupted reception of traps from switches within a large fabric, changes must be made to the UFM configuration in the [/opt/ufm/conf/gv.cfg] file's [Events] section. Specifically, the "max_events" option should be raised from 100 to 1000, while "medium_rate_threshold" and "high_rate_threshold" should both be set to 500. To implement configuration adjustments, disable and then enable the plugin.

In case of an event storm, it is necessary to adjust the Event Policy settings such that General Events are non-alarmable and the TTL is set to zero, as illustrated in the following screenshot:



Other

Additional configurations are located in "/opt/ufm/conf/plugins/snmp/snmp.conf". To implement configuration adjustments, disable and then enable the plugin. For instructions on modifying the appliance, please refer to the UFM-SDN App CLI Guide.

Logs for the SNMP plugin are stored in "/opt/ufm/logs/snmptrap.log". For guidance on accessing logs on the appliance, please refer to the UFM-SDN App CLI Guide.

Packet Mirroring Collector (PMC) Plugin

Overview

The Packet Mirroring Collector/Controller plugin facilitates the configuration of pFRN and Congestion mirroring on switches and subsequently captures mirrored packets, enabling users to conduct real-time monitoring of network events.

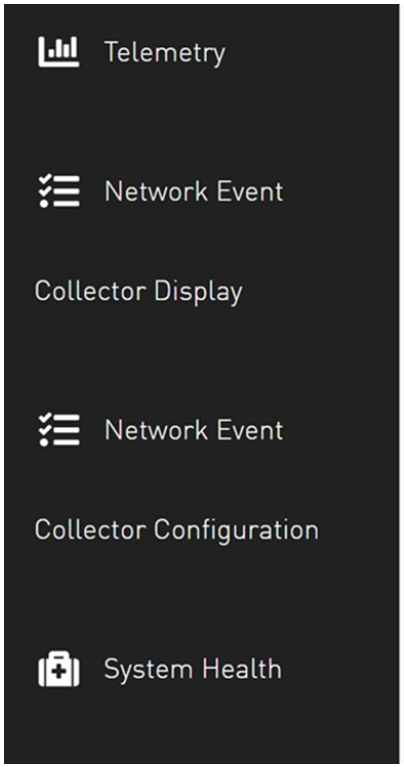
Deployment

Installation

Load the image on the UFM server; either using the UFM GUI -> Settings -> Plugins Management tab, or by loading the image via the following command:

1. Login to the UFM server terminal.
2. Run

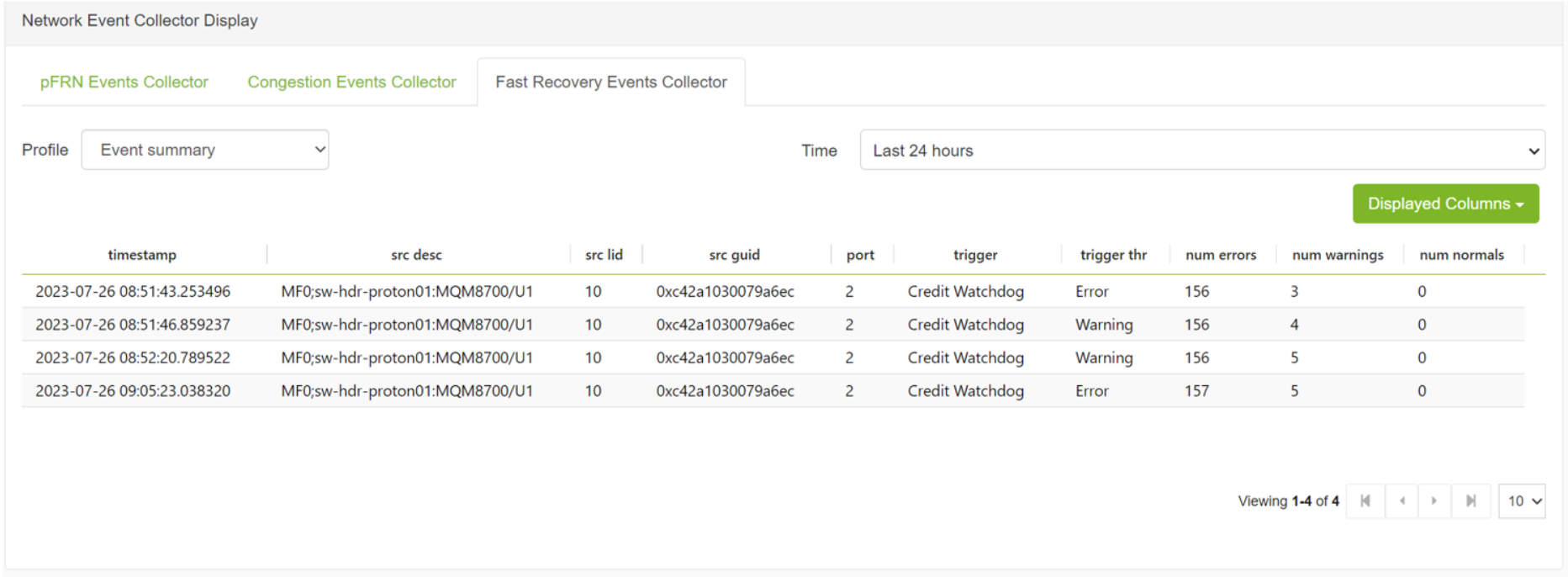
```
docker load -I <path_to_image>
```



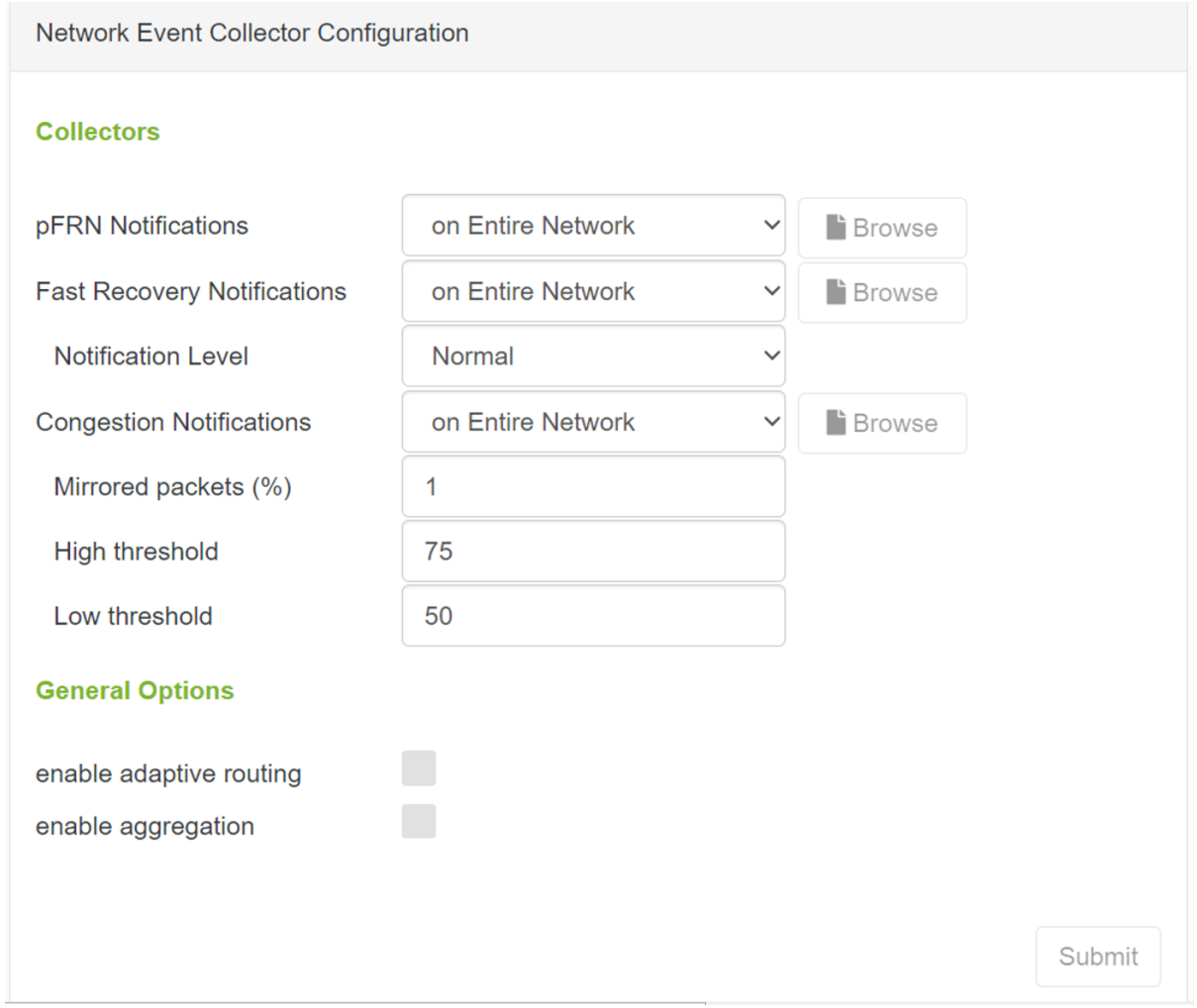
Upon completion of the plugin addition and subsequent refresh of the UFM GUI, the left navigation bar will display two new menu items. These two tabs can be observed in the following GUI screenshots

GUI Screens

Network Event Collector Display



Network Event Collector Configuration



PDR Deterministic Plugin

Overview

The PDR Deterministic plugin is a Docker container that is managed by the UFM and is designed to manage port isolation instead of the UFM automatic isolation. In order to perform port isolation, the PDR plugin utilizes an algorithm that depends on telemetry data provided by UFM Telemetry and monitors packet drop rate (PDR), BER counter values, and cable temperature. Additionally, the plugin can operate in a "dry run" mode, which enables writing to the log without initiating port isolation.

Deployment

1. Install UFM with the latest software version.
2. Run:

```
ufmapl [ mgmt-sa ] (config) # ufm start
```

3. To get PDR plugin image, please contact the NVIDIA Support team. After that, load the plugin using this command:
   When working with UFM in HA mode, load the plugin on the standby node.

```
docker load -i ufm-plugin-pdr-determinitic.tar
```

4. Run the following command. Add `-p pdr-determinitic` to enable the plugin:

```
ufmapl [ mgmt-sa ] (config) # ufm plugin pdr-determinitic add
```

5. Ensure that the plugin is up and running. Run:

```
ufmapl [ mgmt-sa ] (config) # show ufm plugin
```

Default Configuration

The following table lists the default configuration when running the plugin. These configurations can be changed via the pdr_deterministic.conf file.

Value	Default Value	Description
T_ISOLATE	300	Interval for requesting telemetry counters in seconds
MAX_NUM_ISOLATE	10	Maximum number of ports to be isolated. Max(10,0.5% * fabric_size)
TMAX	70	The maximal nominal operating temperature for fabric devices and cables (minimum of the two) Value is in Celsius.
D_TMAX	10	The maximum allowed temperature change within T_ISOLATE interval. Value is in Celsius.
MAX_PDR	1e-12	The maximum allowed Packet Drop Rate.
CONFIGURED_BER_CHECK	True	Indicates whether to check BER counters thresholds
DRY_RUN	False	Isolation decisions are only logged and will not take affect
DEISOLATE_CONSIDER_TIME	5	Consideration time for port de-isolation (in minutes)
AUTOMATIC_DEISOLATE	True	automatically performs de-isolation, even if a port is not set as "treated"
DO_DEISOLATION	True	If set to false, the plugin does not perform de-isolation

⚠ BER thresholds will be taken from the `Field_BER_Thresholds.csv` file.

Isolation Decisions

The plugin's purpose is to isolate malfunctioning ports using the isolation API from the UFM. A port is set as isolated if the values of its counter pass the thresholds of its cable temperature, effective BER, symbol BER, raw BER, or packet drop rate. A port can be de-isolated if its values are back to normal for 5 minutes (configurable).

The primary objective of the plugin is to utilize the isolation API provided by the UFM to isolate malfunctioning ports. A port is set as "isolated" when the values of its counter surpass the predetermined thresholds for parameters such as temperature, effective BER, symbol BER, raw BER, or packet drop rate.

Calculating BER Counters

For calculating BER counters, the plugin extracts the maximum window it needs to wait for calculating the BER value, using the following formula:

$$seconds = \frac{max_BER_target^{-1}}{min_port_rate}$$

Example:

	Rate	BER Target	Minimum Bits	Minimum Time in Seconds	In min
HDR	2.00E+11	1.00E-12	1.00E+12	5	0.083333
HDR	2.00E+11	1.00E-13	1.00E+13	50	0.833333
HDR	2.00E+11	1.00E-14	1.00E+14	500	8.333333
HDR	2.00E+11	1.00E-16	1.00E+16	50000	833.3333

BER counters are calculated with the following formula:

$$BER = \frac{error\ bits_i - error\ bits_{i-1}}{total\ bits_i - total\ bits_{i-1}} = \frac{error\ bits_i - error\ bits_{i-1}}{Link\ data\ rate * (time_i - time_{i-1})}$$

The following telemetry counters are used:

- Symbol: phy_symbol_errors_high/low
- Effective: phy_effective_errors_high/low
- raw: sum(phy_raw_errors_lane<i>_high/low)

Data is kept in memory and is saved for the largest window period.

Dry Run Mode

The plugin can simulates port isolation without actually executing it for the purpose of analyzing the algorithm's performance and decision-making process in order to make future adjustments. This behavior is achieved through the implementation of a "`dry_run`" flag that changes the plugin's behavior to solely record its port "isolation" decisions in the log, rather than invoking the port isolation API. All decisions will be recorded in the plugin's log.

GNMI-Telemetry Plugin

The GNMI Telemetry Plugin functions as a server that employs the gNMI protocol to stream data from UFM telemetry. Users can select what data to stream, specify the intervals, and choose whether to include only deltas (on-change mode).

The GNMI server is designed to support four functions: capability, get, subscribe, and set. However, it should be noted that the server does not currently support the "set" function, only "capability," "get," and "subscribe."

The streamed data is delivered in CSV format. Headers are initially provided in the first message, and subsequently, they are included in every other message. The data is presented in hex format to conserve space for data that remains unchanged. The values are presented as an array of strings, each representing a unique identifier (GUID) and port.

Depending on the selected mode, the values may have missing rows if there have been no changes in the GUID and port.

Furthermore, the plugin has the capability to stream UFM's metadata by providing an inventory of it. While the provided examples will use the gNMIc client for convenience, this functionality can work with any gNMI client.

Authentication

The server's authentication is determined by the gNMI protocol, and whether it is secured or unsecured is specified in the configuration. Two configurable items require authentication: the UFM Telemetry URL and the UFM inventory IP. Both of these items must be configured in the configuration file.

- Authentication is not necessary for the UFM telemetry URL. Therefore, only the telemetry URL is required.
- By default, the inventory is sourced from the UFM of the local host. However, it is possible to change the UFM inventory location to a different machine in the config file. To do so, token access to that machine is necessary.

Secure Server

The server can be secured by using certificates. To secure the server, modify the "`secure_mode_enabled`" flag to "`true`" in the configuration.

Upon initialization, the gNMI server retrieves the UFM certificates from the `/var/opt/ufm/webclient/` folder, utilizing both the server certificates and CA certificates. It is possible to change the certificate folder by changing the shared volume.

The server will requires certificates for client calls and grants access only if the client certificates match its own. The gNMI server periodically examines its certificates for updates and ensures that they remain up to date.

Capability Request

Description: The capability request provides information about the Yang files that the server supports, including their versions. This request can be fulfilled without the need for a connection to the telemetry or inventory.

Example:

```
gnmic -a localhost:9339 capability
```

Get Request

The Get request retrieves data at a specified path. If the telemetry is devoid of information, the server will respond with an empty response. Otherwise, it will respond with counters it can locate.

The path construction follows these steps:

1. Begin with "`nvidia/ib`"
2. Specify the `node_guid` that the user wants to select, with an asterisk (*) representing a selection of all nodes.
3. Choose the desired ports for the selected nodes.
4. Select "`amber`" and the desired counters group, and then specify the counter.

Example:

```
gnmic -a localhost:9339 --insecure get --path nvidia/ib/guid[guid=0x5255456]/port[port_number=2]/amber/
port_counters/hist0
```

The request from the above example is run from node_guid `0x5255456`, in port number 2, and the queried counter is hist0.

Example 2:

```
gnmic -a localhost:9339 --insecure get --path nvidia/ib/guid[guid=*]/port[port_number=*]/amber/port_counters/hist0
```

The request from the above example is run from all the node_guids, in all ports, and the queried counter is hist0.

Example3:

```
gnmic -a localhost:9339 --insecure get --path nvidia/ib/guid[guid=0x5255456]/port[port_number=2]/amber/*
```

The request from the above example is run from node_guid `0x5255456`, port 2, and all its counters.

Subscribe Stream Request

The subscribe request, similar to the get request, provides data from the specified path. When the telemetry is empty, the server responds with an empty result. However, if there is data available, the server responds with the counters it can locate. The stream delivers information at intervals corresponding to the requested interval. If a user fails to specify an interval, the server will transmit the information as soon as it becomes available. The path construction follows the same pattern as the get request.

Example:

```
gnmic -a localhost:9339 --insecure sub --path nvidia/ib/guid[guid=0x5255456]/port[port_number=2]/amber/
port_counters/hist0 -i 30s
```

TBD: This request from node_guid 0x5255456 port 2 the counter hist0 and set the interval to 30 seconds.

If the user wants to test the stream, the stream mode can be set to once, and after that one respond, the stream will be stopped.

Example:

```
gnmic -a localhost:9339 --insecure sub --path nvidia/ib/guid[guid=0x5255456]/port[port_number=2]/amber/
port_counters/hist0 -i 30s --mode once
```

TBD:This request is run from node_guid 0x5255456, port 2 the counter hist0 once, and then shut the stream off, much like a get request.

Subscribe On-Change Request

The subscribe on-change request, much like the standard subscribe request, provides data from the specified path. In the event that the telemetry lacks data, the server responds with an empty result. However, when data is available, the server responds with the counters it can locate. The stream delivers information according to the interval specified in the request, but only if there is new information to transmit. Otherwise, it will wait for the next interval to check the telemetry for updates. The path construction follows the same pattern as the get request.

Importantly, only the data that has been updated will be included in the response; all other parts will be empty but retain the specified format. Similarly, only the nodes that have been updated will be included in the response.

Example:

```
gnmic -a localhost:9339 --insecure  sub --path nvidia/ib/guid[guid=0x5255456]/port[port_number=2]/amber/
port_counters/hist0  --stream-mode on-change --heartbeat-interval 1m
```

TBD: This request from node_guid 0x5255456 port 2 the counter hist0, every minute it will check for changes, if there are it will send the new value.

Example:

```
gnmic -a localhost:9339 --insecure  sub --path nvidia/ib/guid[guid=*]/port[port_number=*]/amber/port_counters/* --
stream-mode on-change --heartbeat-interval 1m
```

This request involves all nodes and ports, aiming to retrieve all counters from the telemetry. It periodically checks for changes every minute, and when changes are detected, it promptly sends the updated values.

Messages Data Format

Telemetry messages consist of two key components: Headers and Values, both representing the telemetry data in CSV format. When utilizing a subscribe request, the headers transition to a string hash format after the second message, primarily to conserve message size. In the case of on-change subscribe messages, there is an additional adjustment where only nodes that have undergone changes are included, along with their corresponding modified values. All other counters for that node will remain empty.

Each value within the "Values" section starts with a timestamp, followed by the node_guid and port number, and then the value of the counter, maintaining the same order as the headers. If a specific counter is not present for the node, it will remain empty in the message.

Example:

```
gnmic -a localhost:9339 --insecure sub --path nvidia/ib/guid[guid=*]/port[port_number=*]/amber/port_counters/hist0
--path nvidia/ib/guid[guid=*]/port[port_number=*]/amber/port_counters/hist1 -i 30s
[{ "source": "localhost:9339",
  "subscription-name": "default-1690282472",
  "timestamp": 1690282475124352063,
  "time": "2023-07-25T13:54:35.124352063+03:00",
  "updates": [{ "Path": "hist0", "values": { "hist0": {
          "Headers": "timestamp,guid,port,hist0,hist1",
          "Values": ["240771222771818,0x8168793592c6a790,1,,2",
            "240771222771818,0x47a67159c915493f,1,1,2",
…
            "240771222771818,0x667203ac69f3f2bf,1,2,",
            "240771222771818,0x113cd807bfed3853,1,0,"
          ]}}}]}
```

TBD: The second message and on the headers will be set to hash values.

Inventory Requests

Inventory messages are conveyed in separate updates, presenting the inventory details of the UFM associated with the provided IP. These messages display comprehensive information, including the total count of various components within the UFM, such as switches, routers, servers, and more, along with details about active ports and the total number of ports, including disabled ones. In cases where the plugin is unable to establish contact with the UFM, it

will revert to using default values defined in the configuration file. It is worth noting that the path for inventory requests differs from the conventional path structure, as they do not rely on specific nodes or ports. Consequently, inventory requests are initiated after "`nvidia/ib`."

Example:

```
gnmic -a localhost:9339 --insecure get -path nvidia/ib/inventory/*
```

Response:

```
[{
    "source": "localhost:9339",
    "timestamp": 1698842237536878067,
    "time": "2023-11-01T09:37:17.536878067+02:00",
    "updates": [{
        "Path": "nvidia/ib/inventory",
        "values": {
          "nvidia/ib/inventory": {
            "ActivePorts": 4, "Cables": 2, "Gateways": 0, "HCAs": 2, "Routers": 0,”Servers": 2,
"Switches": 1, "TotalPorts": 38, "timestamp": 1698842211535069000}
     }}]}]
```

Events Requests

Events messages are provided in separate updates, offering insights into the events occurring within the UFM associated with the specified IP. Given that the event metadata remains consistent, even when numerous events are part of a request, the message format adopts a CSV-like structure. The Headers section contains essential metadata regarding UFM events, while the Values section contains the raw event data. Users can subscribe to these events with the on-change feature enabled, receiving only the events triggered within the subscription interval. Notably, the path structure for event requests differs from the typical node or port-based structure and is requested after "`nvidia/ib`."

Example:

```
gnmic -a localhost:9339 --insecure get -path nvidia/ib/events/*
```

Response:

```
[ {
    "source": "localhost:9339",
    "timestamp": 1698824809647515575,
    "time": "2023-11-01T09:46:49.647515575+02:00",
    "updates": [  {
        "Path": "nvidia/ib/events",
        "values": {
          "nvidia/ib/events": {
            "Headers": [ "id","object_name","write_to_syslog","description","type","event_type","severity","timest
amp","counter","category","object_path","name"],
            "Values": [
              "7718,Grid,false,Disk space usage in /opt/ufm/files/log is above the threshold of
90.0%.,Grid,525,Critical,2023-11-01 07:25:54,N/A,Maintenance,Grid,Disk utilization threshold reached",
              "7717,Grid,false,Disk space usage in /opt/ufm/files/log is above the threshold of
90.0%.,Grid,525,Critical,2023-11-01 07:24:54,N/A,Maintenance,Grid,Disk utilization threshold reached",
              "7716,Grid,false,Disk space usage in /opt/ufm/files/log is above the threshold of
90.0%.,Grid,525,Critical,2023-11-01 07:23:54,N/A,Maintenance,Grid,Disk utilization threshold reached",
….
              "7491,ec0d9a0300d42e54,false,Mcast group is deleted: ff12601bffff0000,
00000002,Computer,67,Info,2023-10-31 06:39:21,N/A,Fabric Notification,default / Computer: r-ufm59,MCast Group
Deleted"]       }
        }  }   ]  }]
```

Appendixes

- Appendix – SM Default Files
- Appendix – UFM Subnet Manager Default Properties
- Appendix – Enhanced Quality of Service
- Appendix – Partitioning
- Appendix – SM Activity Report
- Appendix – SM Partitions.conf File Format
- Appendix - Supported Port Counters and Events
- Appendix – Diagnostic Utilities
- Appendix – Device Management Feature Support
- Appendix – Used Ports
- Appendix – Routing Chains
- Appendix – Adaptive Routing
- Appendix – Configuration Files Auditing
- Appendix - Managed Switches Configuration Info Persistency
- Appendix – Security Features
- Appendix – NVIDIA SHARP Integration
- Appendix – AHX Monitoring
- Appendix – UFM Event Forwarder
- Appendix – UFM Migration
- Appendix - Switch Grouping
- Appendix – Secondary Telemetry Fields

Appendix – SM Default Files

The SM default files are located under the following paths:

- Default SM configuration file – conf/opensm/opensm.conf
- Default node name map file – conf/opensm/ib-node-name-map
- Default partition configuration file – conf/opensm/partitions.conf
- Default QOS policy configuration file – conf/opensm/qos-policy.conf
- Default prefix routes file – conf/opensm/prefix-routes.conf

Appendix – UFM Subnet Manager Default Properties

The following table provides a comprehensive list of UFM SM default properties.

Category	Property	Config File Attribute	Default	Mode/ Field	Description
Generic	Subnet Prefix	subnet_prefix	0xfe80000000000000	RW	Subnet prefix used on the subnet 0xfe80000000000000
	LMC	lmc	0	RW	The LMC value used on the subnet: 0-7 Changes to the LMC parameter require a UFM restart.
	SM LID	master_sm_lid	0		Force LID for local SM when in MASTER state Selected LID must match configured LMC 0 disables the feature
Keys	M_Key	m_key	0x0000000000000000	RW	M_Key value sent to all ports -used to qualify the set(PortInfo)
	M_Key Lease Period	m_key_lease_period	0	RW	The lease period used for the M_Key on the subnet in [sec]
	SM_Key	sm_key	0x0000000000000001	RO	SM_Key value of the SM used for SM authentication
	SA_Key	sa_key	0x0000000000000001	RO	SM_Key value to qualify rcv SA queries as 'trusted'
	Partition enforcement	part_enforce	• Out • In • Both (default- outbound and inbound enforcement enabled)	RO	Partition enforcement type (for switches)

Category	Property	Config File Attribute	Default	Mode/ Field	Description
	MKEY lookup	m_key_lookup	FALSE	RW	If FALSE, SM will not try to determine the m_key of unknown ports.
	M_Key Per Port	m_key_per_port	FALSE	RW	When m_key_per_port is enabled, OpenSM will generate an M_Key for each port
Limits	Packet Life Time	packet_life_time	0x12	RW	The maximum lifetime of a packet in a switch. The actual time is 4.096usec * 2^<packet_life_time> The value 0x14 disables the mechanism
	VL Stall Count	vl_stall_count	0x07	RO	The number of sequential packets dropped that cause the port to enter the VL Stalled state. The result of setting the count to zero is undefined.
	Leaf VL Stall Count	leaf_vl_stall_count	0x07	RO	The number of sequential packets dropped that causes the port to enter theleaf VL Stalled state. The count is for switch ports driving a CA or gateway port. The result of setting the count to zero is undefined.
	Head Of Queue Life time	head_of_queue_lifetime	0x12	RW	The maximum time a packet can wait at the head of the transmission queue. The actual time is 4.096usec * 2^<head_of_queue_lifetime> The value 0x14 disables the mechanism
	Leaf Head Of Queue Life time	leaf_head_of_queue_lifetime	0x10	RW	The maximum time a packet can wait at the head of queue on a switch port connected to a CA or gateway port.
	Maximal Operational VL	max_op_vls	3	RW	Limit of the maximum operational VLs

Category	Property	Config File Attribute	Default	Mode/ Field	Description
	Force Link Speed	force_link_speed	15 (Do NOT change)	RO	Force PortInfo: LinkSpeedEnabled on switch ports. If 0, do not modify. Values are: 1: 2.5 Gbps 3: 2.5 or 5.0 Gbps 5: 2.5 or 10.0 Gbps 7: 2.5 or 5.0 or 10.0 Gbps 2,4,6,8-14 Reserved 15: set to PortInfo: LinkSpeedSupported
Limits	Subnet Timeout	subnet_timeout	18 (1second)	RW	The subnet_timeout code that will be set for all the ports. The actual timeout is 4.096usec * 2^<subnet_timeout>
	Local PHY Error Threshold	local_phy_errors_threshold	0x08	RW	Threshold of local phy errors for sending Trap 129
	Overrun Errors Threshold	overrun_errors_threshold	0x08	RW	Threshold of credit overrun errors for sending Trap 130
Sweep	Sweep Interval	sweep_interval	10	RW	The time in seconds between subnet sweeps (Disabled if 0)
	Reassign Lids	reassign_lids	FALSE (disabled)	RW	If TRUE (enabled), all LIDs are reassigned
	Force Heavy Sweep	force_heavy_sweep_window	-1	RW	Forces heavy sweep after number of light sweeps (-1 disables this option and 0 will cause every sweep to be heavy)

Category	Property	Config File Attribute	Default	Mode/ Field	Description
	Sweep On trap	sweep_on_trap	TRUE (enabled)	RW	If TRUE every trap 128 and 144 will cause a heavy sweep
	Alternative Route Calculation	max_alt_dr_path_retries	4	RW	Maximum number of attempts to find an alternative direct route towards unresponsive ports
	Fabric Rediscovery	max_seq_redisc	2	RW	Max Failed Sequential Discovery Loops
	Offsweep Rebalancing Enable	offsweep_balancing_enabled	FALSE	RW	Enable/Disable idle time routing rebalancing
	Offsweep Rebalancing Window	offsweep_balancing_window	180	RW	Set the time window in seconds after sweep to start rebalancing
Handover	SM Priority	sm_priority	15	RO	SM (enabled). The priority used for deciding which is the master. Range is 0 (lowest priority) to 15 (highest)
	Ignore Other SMs	ignore_other_sm	FALSE (disabled)	RO	If TRUE other SMs on the subnet should be ignored
	Polling Timeout	sminfo_polling_timeout	10	RO	Timeout in seconds between two active master SM polls
	Polling Retries	polling_retry_number	4	RO	Number of failing remote SM polls that declares it non-operational
	Honor GUID-to-LID File	honor_guid2lid_file	FALSE (disabled)	RO	If TRUE, honor the guid2lid file when coming out of standby state, if the guid2lid file exists and is valid

Category	Property	Config File Attribute	Default	Mode/ Field	Description
	Allowed SM GUID list	allowed_sm_guids	(null) (disabled)		List of Host GUIDs where SM is allowed to run when specified. OpenSM ignores SM running on port that is not in this list. If 0, does not allow any other SM. If null, the feature is disabled.
Threading	Max Wire SMPs	max_wire_smps	8	RW	Maximum number of SMPs sent in parallel
	Transaction Timeout	transaction_timeout	200	RO	The maximum time in [msec] allowed for a transaction to complete
	Max Message FIFO Timeout	max_msg_fifo_timeout	10000	RO	Maximum time in [msec] a message can stay in the incoming message queue
	Routing Threads	routing_threads_num	0	RW	Number of threads to be used for parallel minhop/updn calculations. If 0, number of threads will be equal to number of processors.
	Routing Threads Per Core	max_threads_per_core	0	RW	Max number of threads that are allowed to run on the same processor during parallel computing. If 0, threads assignment per processor is up to operating system initial assignment.
Logging	Log File	log_file	/opt/ufm/files/log/opensm.log	RO	Path of Log file to be used
	Log Flags	log_flags	Error and Info 0x03	RW	The log flags, or debug level being used.

Category	Property	Config File Attribute	Default	Mode/ Field	Description
	Force Log Flush	force_log_flush	FALSE (disabled)	RO	Force flush of the log file after each log message
	Log Max Size	log_max_size	4096	RW	Limit the size of the log file in MB. If overrun, log is restarted
	Accumulate Log File	accum_log_file	TRUE (enabled)	RO	If TRUE, will accumulate the log over multiple OpenSM sessions
	Dump Files Directory	dump_files_dir	/opt/ufm/files/log	RO	The directory to hold the file SM dumps (for multicast forwarding tables for example). The file is used collects information.
	Syslog log	syslog_log	0x0	RW	Sets a verbosity of messages to be printed in syslog
Misc	Node Names Map File	node_name_map_name	Null	RW	Node name map for mapping node's to more descriptive node descriptions
	SA database File	sa_db_file	Null	RO	SA database file name
	No Clients Reregistration	no_clients_rereg	FALSE (disabled)	RO	If TRUE, disables client reregistration
	Exit On Fatal Event	exit_on_fatal	TRUE (enabled)	RO	If TRUE (enabled), the SM exits for fatal initialization issues
	Switch Isolation From Routing	held_back_sw_file	Null	RW	File that contains GUIDs of switches isolated from routing

Category	Property	Config File Attribute	Default	Mode/ Field	Description
	Enable NVIDIA SHARP support	sharp_enabled	Enabled	RW	Defines whether to enable/disable NVIDIA SHARP on supporting ports.
Multicast	Disable Multicast	disable_multicast	FALSE (disabled)	RO	If TRUE, OpenSM should disable multicast support and no multicast routing is performed
	Multicast Group Parameters	default_mcg_mtu	0	RW	Default MC group MTU for dynamic group creation. 0 disables this feature, otherwise, the value is a valid IB encoded MTU
Multicast	Multicast Group Parameters	default_mcg_rate	0	RW	Default MC group rate for dynamic group creation. 0 disables this feature, otherwise, the value is a valid IB encoded rate
Multicast	Enable incremental multicast routing	enable_inc_mc_routing	FALSE	RW	Enable incremental multcast routing
Multicast	MC root file	mc_roots_file	null	RW	Specify predefined MC groups root guids
QoS	Settings	qos	FALSE (disabled) *From UFM v3.7 and on	RW	If FALSE (disabled), SM will not apply QoS settings
Unhealthy Ports	Enabling Unhealthy Ports	hm_unhealthy_ports_checks	TRUE	RW	Enables Unhealthy Ports configuration
	Configuration file	hm_ports_health_policy_file	null	RW	Specifies configuration file for health policy

Category	Property	Config File Attribute	Default	Mode/ Field	Description
	Unhealthy actions	hm_sw_manual_action	no_discover	RW	Specifies what to do with switch ports which were manually added to health policy file
	MADs validation	validate_smp	TRUE	RW	If set to TRUE, opensm will ignore nodes sending non-spec compliant MADs. When set to FALSE, opensm will log the warning in the opensm log file about non-compliant node
Routing	Unicatst Routingengine	routing_engine	(null)	RW	By default, ar_updn routing engine is used by the SM. Supported routing engines are minhop, updn, dnup, ftree, dor, torus-2QoS, kdor-hc, kdor-ghc , dfp, dfp2, ar_updn, ar_ftree and ar_dor.
	Randomization	scatter_ports	8	RW	Assigns ports in a random order instead of round-robin. If 0, the feature is disabled, otherwise use the value as a random seed. Applicable to the MINHOP/UPDN routing algorithms
	Randomization	guid_routing_order_no_scatter	TRUE	RO	Do not use scatter for ports defined in guid_routing_order file
	Unicast Routing Caching	use_ucast_cache	TRUE	RW	Use unicast routing cache for routing computation time improvement
	GUID Ordering During Routing	guid_routing_order_file	NULL	RW	The file holding guid routing order of particular guids (for MinHop, Up/Down)
	Torus Routing	torus_config	/opt/ufm/files/conf/opensm/ torus-2QoS.con	RW	Torus-2QoS configuration file name
	Routing Chains	pgrp_policy_file	NULL	RW	The file holding the port groups policy
		topo_policy_file	NULL	RW	The file holding the topology policy
		rch_policy_file	NULL	RW	The file holding the routing chains policy

Category	Property	Config File Attribute	Default	Mode/ Field	Description
		max_topologies_per_sw	1	RO	Defines maximal number of topologies to which a single switch may be assigned during routing engine chain configuration.
	Incremental Multicast Routing (IMR)	enable_inc_mc_routing	TRUE	RW	If TRUE, MC nodes will be added to the MC tree incrementally. When set to FALSE, the tree will be recalculated per eachg change.
	MC Global root	mc_primary_root_guid/ mc_secondary_root_guid	0x0000000000000000 (for both)	RW	Primary and Secondary global mc root guid
	Scatter ports	use_scatter_for_switch_lid	FALSE	RW	Use scatter when routing to the switch's LIDs
	updn lid tracking mode	updn_lid_tracking_mode	FALSE	RW	Controls whether SM will use LID tracking or not when updn or ar_updn routing engine is used
Events	Event Subscription Handling	drop_subscr_on_report_fail	FALSE	RW	Drop subscription on report failure (o13-17.2.1)
	Event Subscription Handling	drop_event_subscriptions	TRUE	RW	Drop event subscriptions (InformInfo and ServiceRecords) on port removal and SM coming out of STANDBY
Virtualization	Virtualization enabled	virt_enabled	Enabled	RW	Enables/disables virtualization support
	Maximum ports in virtualization process	virt_max_ports_in_process	64	RW	Sets a number of ports to be handled on each virtualization process cycle
Router	Router aguid enable	rtr_aguid_enable	0 (Disabled)	RW	Defines whether the SM should create alias GUIDs required for router support for each HCA port

Category	Property	Config File Attribute	Default	Mode/ Field	Description
	Router path record flow label	rtr_pr_flow_label	0	RW	Defines flow label value to use in multi-subnet path query responses
	Router path record tclass	rtr_pr_tclass	0	RW	Defines tclass value to use in multi-subnet path query responses.
	Router path record sl	rtr_pr_sl	0	RW	Defines sl value to use in multi-subnet path query responses
	Router path record MTU	rtr_pr_mtu	4 (IB_MTU_LEN_2048)	RW	Define MTU value to use in multi-subnet path query responses
	Router path record rate	rtr_pr_rate	16 (IB_PATH_RECORD_RATE_100_GBS)	RW	Defines rate value to use in multi-subnet path query responses
SA Security	SA Tnhanced Trust Model (SAETM)	sa_enhanced_trust_model	FALSE	RW	Controls whether SAETM is enabled.
	Untrusted GuidInfo records	sa_etm_allow_untrusted_guidinfo_rec	FALSE	RW	Controls whether to allow Untrusted Guidinfo record requests in SAETM.
	Guidinfo record requests by VF	sa_etm_allow_guidinfo_rec_by_vf	FALSE	RW	Controls whether to allow Guidinfo record requests by vf in SAETM.
	Untrusted proxy requests	sa_etm_allow_untrusted_proxy_requests	FALSE	RW	Controls whether to allow Untrusted proxy requests in SAETM.
	Max number of multicast groups	sa_etm_max_num_mcgs	128	RW	Max number of multicast groups per port/vport that can be registered.

Category	Property	Config File Attribute	Default	Mode/ Field	Description
	Max number of service records	sa_etm_max_num_srvcs	32	RW	Max number of service records per port/vport that can be registered.
	Max number of event subscriptions	sa_etm_max_num_event_subs	32	RW	Max number of event subscriptions (InformInfo) per port/ vport that can be registered.
	SGID spoofing	sa_check_sgid_spoofing	TRUE	RW	If enabled, the SA checks for SGID spoofing in every request with GRH included, unless the SLID is from a router port at that request.

Configuring UFM for SR-IOV

Single-root I/O virtualization (SR-IOV) enables a PCI Express (PCIe) device to appear to be multiple separate physical PCIe devices.

UFM is ready to work with SR-IOV devices by default. You can fine-tune the configuration using the SM configuration.

The following arguments are available for ConnectX-5 and later devices:

Argument	Value	Description
virt_enabled	• 0 - no virtualization support • 1 - disable virtualization on all virtualization supporting ports • 2 - enable virtualization on all virtualization supporting ports (default)	Virtualization support
virt_max_ports_in_process	Possible values: 0-65535; where 0 processes all pending ports Default: 64	Maximum number of ports to be processed simultaneously by the virtualization manager

Argument	Value	Description
virt_default_hop_limit	Possible values: 0-255 Default: 2	Default value for hop limit to be returned in path records where either the source or destination are virtual ports

Isolating Switch From Routing

UFM can isolate particular switches from routing in order to perform maintenance of the switches with minimal interruption to the existing traffic in the fabric.

Isolating a switch from routing is done via UFM Subnet Manager as follows:

1. Create a file that includes either the node GUIDs or the system GUID of the switches under maintenance. For example:

```
0x1234566
0x1234567
```

2. Import the file into the UFM Appliance. Run:

```
ib sm configuration import held-back-sw
```

Once SM completes rerouting, the traffic does not go through the ports of isolated switches.

To attach the switch to the routing:

1. Remove the GUID of the switch from the list of isolated switches defined in Step 1 of the isolation process.
2. Import the file into the UFM Appliance. Run:

```
ib sm configuration import held-back-sw
```

Once SM completes rerouting, traffic will go through the switch.

Appendix – Enhanced Quality of Service

Enhanced QoS provides a higher resolution of QoS at the service level (SL). Users can configure rate limit values per SL for physical ports, virtual ports, and port groups, using enhanced_qos_policy_file configuration parameter.

Valid values of this parameter:

- Full path to the policy file through which Enhanced QoS Manager is configured
- "null" – to disable the Enhanced QoS Manager (default value)

⚠ To enable Enhanced QoS Manager, QoS must be enabled in SM configuration file.

Enhanced QoS Policy File

The policy file is comprised of two sections:

- **BW_NAMES**: Used to define bandwidth setting and name (currently, rate limit is the only setting). Bandwidth names are defined using the syntax:

  ```
  <name> = <rate limit in 1Mbps units>
  ```

  Example:

  ```
  My_bandwidth = 50
  ```

- **BW_RULES**: Used to define the rules that map the bandwidth setting to a specific SL of a specific GUID. Bandwidth rules are defined using the syntax:

  ```
  <guid>|<port group name> = <sl id>:<bandwidth name>, <sl id>:<bandwidth name>…
  ```

  Examples:

```
0x2c90000000025 = 5:My_bandwidth, 7:My_bandwidth
Port_grp1 = 3:My_bandwidth, 9:My_bandwidth
```

Notes

- **Rate limit = 0** represents unlimited rate limit.
- Any unspecified SL in a rule will be set to 0 (unlimited) rate limit automatically.
- “default” is a well-known name which can be used to define a default rule used for any GUID with no defined rule (If no default rule is defined, any GUID without a specific rule will be configured with unlimited rate limit for all SLs).
- Failure to complete policy file parsing leads to an undefined behavior. User must confirm no relevant error messages in SM log in order to ensure Enhanced QoS Manager is configured properly.
- An empty file with only ‘BW_NAMES’ and ‘BW_RULES’ keywords configures the network with an unlimited rate limit.
- The VPORT_BW_RULES section is optional and includes virtual port GUIDs only (including the vport0 GUID). Physical port GUIDs added to this section are treated as vport0 GUIDs.

Policy File Example

The below is an example of configuring all ports in the fabric with rate limit of 50Mbps on SL1, except for GUID 0x2c90000000025, which is configured with rate limit of 100Mbps on SL1. In this example, all SLs (other than SL1) are unlimited.

```
-----------------------------------------------------------------------
BW_NAMES
bw1 = 50
bw2 = 100
BW_RULES
default: 1:bw1
0x2c90000000025: 1:bw2
-----------------------------------------------------------------------
VPORT_BW_RULES
default = all:DEF_BW_2
```

Appendix – Partitioning

Partitioning enforces isolation of the fabric. The default partition is created on all managed devices. Devices that are running an SM, all switches, and the 10GbE gateway are added to the default partition with full membership. By default, all the HCA ports are also added to the default partition with FULL membership.

Partitioning is provisioned to the Subnet Manager via the partitions.conf configuration file, which cannot be removed or modified.

> ⚠ For those who use NVIDIA gateway systems, for proper system functionality, disable the automatic partitioning by changing the attribute gateway_port_partitioning = none in the conf/gv.cfg configuration. Restart UFM for the change to take effect.

If required, you can add an extension to the *partitions.conf* file that is generated by UFM. You can import the file *conf/partitions.conf.user_ext* by using the CLI command "ib sm configuration import partition-config-user-ext" and the content of this extension file will be added to the *partitions.conf* file. Files synchronization is done by UFM on every logical model change. However, it can also be triggered manually by running the CLI command "ib sm partition-config-merge". The command merges the *conf/partitions.conf.user_ext* file into the *conf/opensm/partitions.conf* file and starts the heavy sweep on the Subnet Manager.

> ⚠ The maximum length of the line in the partitions.conf file is 4096 characters. However, to enable long PKeys, it is possible to split the pkey membership to multiple lines:
> IOPartition=0x4, ipoib, sl=0, defmember=full : <port-guid1> , <port-guid2> ;
> IOPartition=0x4, ipoib, sl=0, defmember=full : <port-guid3> , <port-guid4> ;

The *partitions.conf.user_ext* uses the same format as the *partitions.conf* file. See SM Partitions.conf File Format for the format of the *partitions.conf* file.

For example, to add server ports to PKey 4:

```
IOPartition=0x4, ipoib, sl=0, defmember=full : 0x8f10001072a41;
```

Appendix – SM Activity Report

SM can produce an activity report in a form of a dump file that details the different activities done in the SM. Activities are divided into subjects. The table below specifies the different activities currently supported in the SM activity report.

Reporting of each subject can be enabled individually using the configuration parameter activity_report_subjects:

- Valid values:
  Comma-separated list of subjects to dump. The current supported subjects are:
  - "mc" – activity IDs 1, 2 and 8
  - "prtn" – activity IDs 3, 4, and 5
  - "virt" – activity IDs 6 and 7
  - "routing" – activity IDs 8-12

  Two predefined values can be configured as well:
  - "all" – dump all subjects
  - "none" – disable the feature by dumping none of the subjects
  - Default value: "none"

SM Supported Activities

Activity ID	Activity Name	Additional Fields	Comments	Description
1	mcm_member	- MLid - MGid - Port Guid - Join State	Join state: 1 - Join -1 - Leave	Member joined/left MC group

Activity ID	Activity Name	Additional Fields	Comments	Description
2	mcg_change	- MLid - MGid - Change	Change: 0 - Create 1 - Delete	MC group created/deleted
3	prtn_guid_add	- Port Guid - PKey - Block index - Pkey Index		Guid added to partition
4	prtn_create	-PKey - Prtn Name		Partition created
5	prtn_delete	- PKey - Delete Reason	Delete Reason: 0 - empty prtn 1 - duplicate prtn 2 - sm shutdown	Partition deleted
6	port_virt_discover	- Port Guid - Top Index		Port virtualization discovered
7	vport_state_change	- Port Guid - VPort Guid - VPort Index - VNode Guid - VPort State	VPort State: 1 - Down 2 - Init 3 - ARMED 4 - Active	Vport state changed

Activity ID	Activity Name	Additional Fields	Comments	Description
8	mcg_tree_calc	- mlid		MCast group tree calculated
9	routing_succeed	routing engine name		Routing done successfully
10	routing_failed	routing engine name		Routing failed
11	ucast_cache_invalidated			ucast cache invalidated
12	ucast_cache_routing_done			ucast cache routing done

Appendix – SM Partitions.conf File Format

This appendix presents the content and format of the SM *partitions.conf* file.

```
OpenSM Partition configuration
==============================

The default partition will be created by OpenSM unconditionally even
when partition configuration file does not exist or cannot be accessed.

The default partition has P_Key value 0x7fff. OpenSM's port will always
have full membership in default partition. All other end ports will have
full membership if the partition configuration file is not found or cannot
be accessed, or limited membership if the file exists and can be accessed
but there is no rule for the Default partition.

Effectively, this amounts to the same as if one of the following rules
below appear in the partition configuration file:
In the case of no rule for the Default partition:
Default=0x7fff : ALL=limited, SELF=full ;
In the case of no partition configuration file or file cannot be accessed:
Default=0x7fff : ALL=full ;
```

```
File Format
===========

Comments:

Line content followed after \'#\' character is comment and ignored by
parser.

General file format:

<Partition Definition>:[<newline>]<Partition Properties>;

     Partition Definition:
       [PartitionName][=PKey][,ipoib_bc_flags][,defmember=full|limited]

        PartitionName  - string, will be used with logging. When omitted
                         empty string will be used.
        PKey           - P_Key value for this partition. Only low 15 bits will
                         be used. When omitted will be autogenerated.
        ipoib_bc_flags - used to indicate/specify IPoIB capability of this partition.

        defmember=full|limited - specifies default membership for port guid
                          list. Default is limited.

     ipoib_bc_flags:
        ipoib_flag|[mgroup_flag]*

        ipoib_flag - indicates that this partition may be used for IPoIB, as
             a result the IPoIB broadcast group will be created with
             the flags given, if any.

     Partition Properties:
       [<Port list>|<MCast Group>]* | <Port list>

     Port list:
        <Port Specifier>[,<Port Specifier>]

     Port Specifier:
        <PortGUID>[=[full|limited]]
```

```
   PortGUID         - GUID of partition member EndPort. Hexadecimal
                      numbers should start from 0x, decimal numbers
                      are accepted too.
   full or limited  - indicates full or limited membership for this
                      port.  When omitted (or unrecognized) limited
                      membership is assumed.

MCast Group:
   mgid=gid[,mgroup_flag]*<newline>

           - gid specified is verified to be a Multicast address
         IP groups are verified to match the rate and mtu of the
         broadcast group.  The P_Key bits of the mgid for IP
         groups are verified to either match the P_Key specified
         in by "Partition Definition" or if they are 0x0000 the
         P_Key will be copied into those bits.

mgroup_flag:
   rate=<val>  - specifies rate for this MC group
                 (default is 3 (10GBps))
   mtu=<val>   - specifies MTU for this MC group
                 (default is 4 (2048))
   sl=<val>    - specifies SL for this MC group
                 (default is 0)
   scope=<val> - specifies scope for this MC group
                 (default is 2 (link local)).  Multiple scope settings
                 are permitted for a partition.
         NOTE: This overwrites the scope nibble of the specified
           mgid.  Furthermore specifying multiple scope
           settings will result in multiple MC groups
           being created.
   qkey=<val>      - specifies the Q_Key for this MC group
                     (default: 0x0b1b for IP groups, 0 for other groups)
                     WARNING: changing this for the broadcast group may
                  break IPoIB on client nodes!!!
   tclass=<val>    - specifies tclass for this MC group
                     (default is 0)
   FlowLabel=<val> - specifies FlowLabel for this MC group
                     (default is 0)

newline: '\n'
```

```
Note that values for rate, mtu, and scope, for both partitions and multicast
groups, should be specified as defined in the IBTA specification (for example,
mtu=4 for 2048).

There are several useful keywords for PortGUID definition:

 - 'ALL' means all end ports in this subnet.
 - 'ALL_CAS' means all Channel Adapter end ports in this subnet.
 - 'ALL_SWITCHES' means all Switch end ports in this subnet.
 - 'ALL_ROUTERS' means all Router end ports in this subnet.
 - 'SELF' means subnet manager's port.

Empty list means no ports in this partition.



Notes:
-----

White space is permitted between delimiters ('=', ',',':',';').

PartitionName does not need to be unique, PKey does need to be unique.
If PKey is repeated then those partition configurations will be merged
and first PartitionName will be used (see also next note).

It is possible to split partition configuration in more than one
definition, but then PKey should be explicitly specified (otherwise
different PKey values will be generated for those definitions).



Examples:
--------

 Default=0x7fff : ALL, SELF=full ;
 Default=0x7fff : ALL, ALL_SWITCHES=full, SELF=full ;

 NewPartition , ipoib : 0x123456=full, 0x3456789034=limited, 0x2134af2306 ;

 YetAnotherOne = 0x300 : SELF=full ;
 YetAnotherOne = 0x300 : ALL=limited ;
```

```
 ShareIO = 0x80 , defmember=full : 0x123451, 0x123452;
 # 0x123453, 0x123454 will be limited
 ShareIO = 0x80 : 0x123453, 0x123454, 0x123455=full;
 # 0x123456, 0x123457 will be limited
 ShareIO = 0x80 : defmember=limited : 0x123456, 0x123457, 0x123458=full;
 ShareIO = 0x80 , defmember=full : 0x123459, 0x12345a;
 ShareIO = 0x80 , defmember=full : 0x12345b, 0x12345c=limited, 0x12345d;

 # multicast groups added to default
 Default=0x7fff,ipoib:
        mgid=ff12:401b::0707,sl=1 # random IPv4 group
        mgid=ff12:601b::16     # MLDv2-capable routers
        mgid=ff12:401b::16     # IGMP
        mgid=ff12:601b::2      # All routers
        mgid=ff12::1,sl=1,Q_Key=0xDEADBEEF,rate=3,mtu=2 # random group
        ALL=full;


Note:
----

The following rule is equivalent to how OpenSM used to run prior to the
partition manager:

Default=0x7fff,ipoib:ALL=full;
```

Appendix - Supported Port Counters and Events

Port counters and events are available in the following views:

- Events and Port Counters area, at the bottom of the UFM window
- Error window (Error tab) in the Manage Devices tab
- In the New Monitoring Session window, in the Monitor tab, when clicking Create New Session
- Event Log in the Log tab (click Show Event Log)

InfiniBand Port Counters

The following tables list and describe the port counters and events currently supported:

- InfiniBand Port Counters
- Calculated Port Counters

InfiniBand Port Counters	
Counter	**Description**
Xmit Data (in bytes)	Total number of data octets, divided by 4, transmitted on all VLs from the port, including all octets between (and not including) the start of packet delimiter and the VCRC, and may include packets containing errors. All link packets are excluded. Results are reported as a multiple of four octets.
Rcv Data (in bytes)	Total number of data octets, divided by 4, received on all VLs at the port. All octets between (and not including) the start of packet delimiter and the VCRC are excluded and may include packets containing errors. All link packets are excluded. When the received packet length exceeds the maximum allowed packet length specified in C7-45: the counter may include all data octets exceeding this limit. Results are reported as a multiple of four octets.
Xmit Packets	Total number of packets transmitted on all VLs from the port, including packets with errors and excluding link packets.
Rcv Packets	Total number of packets, including packets containing errors and excluding link packets, received from all VLs on the port.
Rcv Errors	Total number of packets containing errors that were received on the port including: • Local physical errors (ICRC, VCRC, LPCRC, and all physical errors that cause entry into the BAD PACKET or BAD PACKET DISCARD states of the packet receiver state machine) • Malformed data packet errors (LVer, length, VL) • Malformed link packet errors (operand, length, VL) • ackets discarded due to buffer overrun (overflow)
Xmit Discards	Total number of outbound packets discarded by the port when the port is down or congested for the following reasons: • Output port is not in the active state

InfiniBand Port Counters	
Counter	**Description**
	• Packet length has exceeded NeighborMTU • Switch Lifetime Limit exceeded • Switch HOQ Lifetime Limit exceeded, including packets discarded while in VLStalled State.
Symbol Errors	Total number of minor link errors detected on one or more physical lanes.
Link Error Recovery	Total number of times the Port Training state machine has successfully completed the link error recovery process.
Link Error Downed	Total number of times the Port Training state machine has failed the link error recovery process and downed the link.
Local Integrity Error	The number of times that the count of local physical errors exceeded the threshold specified by LocalPhyErrors
Rcv Remote Physical Error	Total number of packets marked with the EBP delimiter received on the port.
Xmit Constraint Error	Total number of packets not transmitted from the switch physical port for the following reasons: • FilterRawOutbound is true and packet is raw • PartitionEnforcementOutbound is true and packet fails partition key check or IP version check
Rcv Constraint Error	Total number of packets received on the switch physical port that are discarded for the following reasons: • FilterRawInbound is true and packet is raw • PartitionEnforcementInbound is true and packet fails partition key check or IP version check
Excess Buffer Overrun Error	The number of times that OverrunErrors consecutive flow control update periods occurred, each having at least one overrun error
Rcv Switch Relay Error	Total number of packets received on the port that were discarded when they could not be forwarded by the switch relay for the following reasons: • DLID mapping • VL mapping • Looping (output port = input port)
VL15 Dropped	Number of incoming VL15 packets dropped because of resource limitations (e.g., lack of buffers) in the port
XmitWait	The number of ticks during which the port selected by PortSelect had data to transmit but no data was sent during the entire tick because of insufficient credits or of lack of arbitration.

InfiniBand Calculated Port Counters	
Counter	**Description**
Normalized XmitData	Effective port bandwidth utilization in % XmitData incremental/ Link Capacity
Normalized Congested Bandwidth	Amount of bandwidth that was suppressed due to congestion (XmitWait incremental/ Time) * Link Capacity Separate counters are used for Tier 4 ports and for the rest of the ports.

Supported Traps and Events

Device events are listed as VDM or CDM in the Source column of the Events table in the UFM GUI. For information about defining event policy, see Configuring Event Management.

Alarm ID	Alarm Name	To Log	Alarm	Default Severity	Default Threshold	Default TTL	Related Object	Category	Source
64	GID Address In Service	1	0	Info	1	300	Port	Fabric Notification	SM
65	GID Address Out of Service	1	0	Warning	1	300	Port	Fabric Notification	SM
66	New MCast Group Created	1	0	Info	1	300	Port	Fabric Notification	SM
67	MCast Group Deleted	1	0	Info	1	300	Port	Fabric Notification	SM
110	Symbol Error	1	1	Warning	200	300	Port	Hardware	Telemetry
111	Link Error Recovery	1	1	Minor	1	300	Port	Hardware	Telemetry
112	Link Downed	1	1	Critical	1	300	Port	Hardware	Telemetry

Alarm ID	Alarm Name	To Log	Alarm	Default Severity	Default Threshold	Default TTL	Related Object	Category	Source
113	Port Receive Errors	1	1	Minor	5	300	Port	Hardware	Telemetry
114	Port Receive Remote Physical Errors	0	0	Minor	5	300	Port	Hardware	Telemetry
115	Port Receive Switch Relay Errors	1	1	Minor	999	300	Port	Fabric Configuration	Telemetry
116	Port Xmit Discards	1	1	Minor	200	300	Port	Communication Error	Telemetry
117	Port Xmit Constraint Errors	1	1	Minor	200	300	Port	Communication Error	Telemetry
118	Port Receive Constraint Errors	1	1	Minor	200	300	Port	Communication Error	Telemetry
119	Local Link Integrity Errors	1	1	Minor	5	300	Port	Hardware	Telemetry
120	Excessive Buffer Overrun Errors	1	1	Minor	100	300	Port	Communication Error	Telemetry
121	VL15 Dropped	1	1	Minor	50	300	Port	Communication Error	Telemetry
122	Congested Bandwidth (%) Threshold Reached	1	1	Minor	10	300	Port	Hardware	Telemetry
123	Port Bandwidth (%) Threshold Reached	1	1	Minor	95	300	Port	Communication Error	Telemetry
130	Non-optimal link width	1	1	Minor	1	0	Port	Hardware	SM

Alarm ID	Alarm Name	To Log	Alarm	Default Severity	Default Threshold	Default TTL	Related Object	Category	Source
134	T4 Port Congested Bandwidth	1	1	Warning	10	300	Port	Communication Error	Telemetry
141	Flow Control Update Watchdog Timer Expired	1	0	Warning	1	300	Port	Hardware	SM
144	Capability Mask Modified	1	0	Info	1	300	Port	Fabric Notification	SM
145	System Image GUID changed	1	0	Info	1	300	Port	Communication Error	SM
156	Link Speed Enforcement Disabled	1	0	Critical	0	300	Site	Fabric Notification	SM
250	Running in Limited Mode	1	1	Critical	1	0	Grid	Maintenance	Licensing
251	Switching to Limited Mode	1	1	Critical	1	0	Grid	Maintenance	Licensing
252	License Expired	1	1	Warning	1	0	Grid	Maintenance	Licensing
253	Duplicated licenses	1	0	Critical	1	0	Grid	Maintenance	Licensing
254	License Limit Exceeded	1	0	Critical	1	0	Grid	Maintenance	Licensing
255	License is About to Expire	1	0	Warning	1	0	Grid	Maintenance	Licensing
256	Bad M_Key	1	0	Minor	1	300	Port	Security	SM
257	Bad P_Key	1	0	Minor	1	300	Port	Security	SM
258	Bad Q_Key	1	0	Minor	1	300	Port	Security	SM
259	Bad P_Key Switch External Port	1	0	Critical	1	300	Port	Security	SM

Alarm ID	Alarm Name	To Log	Alarm	Default Severity	Default Threshold	Default TTL	Related Object	Category	Source
328	Link is Up	1	0	Info	1	0	Link	Fabric Topology	SM
329	Link is Down	1	0	Warning	1	0	Site	Fabric Topology	SM
331	Node is Down	1	0	Warning	1	0	Site	Fabric Topology	SM
332	Node is Up	1	0	Info	1	300	Site	Fabric Topology	SM
336	Port Action Succeeded	1	0	Info	1	0	Port	Maintenance	UFM
337	Port Action Failed	1	0	Minor	1	0	Port	Maintenance	UFM
338	Device Action Succeeded	1	0	Info	1	0	Port	Maintenance	UFM
339	Device Action Failed	1	0	Minor	1	0	Port	Maintenance	UFM
344	Partial Switch ASIC Failure	1	1	Critical	1	0	Switch	Maintenance	UFM
370	Gateway Ethernet Link State Changed	1	0	Warning	1	0	Gateway	Gateway	SM
371	Gateway Reregister Event Received	1	0	Warning	1	0	Gateway	Gateway	SM
372	Number of Gateways Changed	1	0	Warning	1	0	Gateway	Gateway	SM
373	Gateway will be Rebooted	1	0	Warning	1	0	Gateway	Gateway	SM
374	Gateway Reloading Finished	1	0	Info	1	0	Gateway	Gateway	SM
380	Switch Upgrade Error	1	1	Critical	1	0	Switch	Maintenance	UFM

Alarm ID	Alarm Name	To Log	Alarm	Default Severity	Default Threshold	Default TTL	Related Object	Category	Source
381	Switch Upgrade Failed	1	0	Info	1	0	Switch	Maintenance	UFM
328	Module status NOT PRESENT	1	1	Warning	1	420	Switch	Module Status	UFM
383	Host Upgrade Failed	1	0	Info	1	0	Computer	Maintenance	UFM
384	Switch Module Powered Off	1	1	Info	1	420	Switch	Module Status	UFM
385	Switch FW Upgrade Started	1	0	Info	1	0	Switch	Maintenance	UFM
386	Switch SW Upgrade Started	1	0	Info	1	0	Switch	Maintenance	UFM
387	Switch Upgrade Finished	1	0	Info	1	0	Switch	Maintenance	UFM
388	Host FW Upgrade Started	1	0	Info	1	0	Computer	Maintenance	UFM
389	Host SW Upgrade Started	1	0	Info	1	0	Computer	Maintenance	UFM
391	Switch Module Removed	1	0	Info	1	0	Switch	Fabric Notification	Switch
392	Module Temperature Threshold Reached	1	0	Info	40	0	Module	Hardware	Switch
393	Switch Module Added	1	0	Info	1	0	Switch	Fabric Notification	Switch
394	Module Status FAULT	1	1	Critical	1	420	Switch	Module Status	Switch
395	Device Action Started	1	0	Info	1	0	Port	Maintenance	UFM
396	Site Action Started	1	0	Info	1	0	Port	Maintenance	UFM

Alarm ID	Alarm Name	To Log	Alarm	Default Severity	Default Threshold	Default TTL	Related Object	Category	Source
397	Site Action Failed	1	0	Minor	1	0	Port	Maintenance	UFM
398	Switch Chip Added	1	0	Info	1	0	Switch	Fabric Notification	Switch
399	Switch Chip Removed	1	0	Critical	1	0	Switch	Fabric Notification	Switch
403	Device Pending Reboot	1	1	Warning	0	300	Device	Maintenance	UFM
404	System Information is missing	1	1	Warning	1	300	Switch	Communication Error	UFM
405	Switch Identity Validation Failed	1	1	Warning	1	300	Switch	Communication Error	UFM
406	Switch System Information is missing	1	1	Waring	1	300	Switch	Communication Error	UFM
407	COMEX Ambient Temperature Threshold Reached	1	1	Minor	60	300	Switch	Hardware	Switch
408	Switch is Unresponsive	1	1	Critical	1	300	Switch	Communication Error	UFM
502	Device Upgrade Finished	1	0	Info	1	300	Device	Maintenance	UFM
506	Device Upgrade Finished	1	0	Info	1	300	Device	Maintenance	UFM
508	Core Dump Created	1	1	Info	1	300	Grid	Maintenance	UFM
510	SM Failover	0	1	Critical	1	300	Grid	Fabric Notification	SM
511	SM State Change	0	1	Info	1	300	Grid	Fabric Notification	SM

Alarm ID	Alarm Name	To Log	Alarm	Default Severity	Default Threshold	Default TTL	Related Object	Category	Source
512	SM UP	0	1	Info	1	300	Grid	Fabric Notification	SM
513	SM System Log Message	0	1	Minor	1	300	Grid	Fabric Notification	SM
514	SM LID Change	0	1	Warning	1	300	Grid	Fabric Notification	SM
515	Fabric Health Report Info	1	1	Info	1	300	Grid	Fabric Notification	UFM
516	Fabric Health Report Warning	1	1	Warning	1	300	Grid	Fabric Notification	UFM
517	Fabric Health Report Error	1	1	Critical	1	300	Grid	Fabric Notification	UFM
518	UFM-related process is down	1	1	Critical	1	300	Grid	Maintenance	UFM
519	Logs purge failure	1	1	Minor	1	300	Grid	Maintenance	UFM
520	Restart of UFM-related process succeeded	1	1	Info	1	300	Grid	Maintenance	UFM
521	UFM is being stopped	1	1	Critical	1	300	Grid	Maintenance	UFM
522	UFM is being restarted	1	1	Critical	1	300	Grid	Maintenance	UFM
523	UFM failover is being attempted	1	1	Info	1	300	Grid	Maintenance	UFM
524	UFM cannot connect to DB	1	1	Critical	1	300	Grid	Maintenance	UFM
525	Disk utilization threshold reached	1	1	Critical	1	300	Grid	Maintenance	UFM
526	Memory utilization threshold reached	1	1	Critical	1	300	Grid	Maintenance	UFM

Alarm ID	Alarm Name	To Log	Alarm	Default Severity	Default Threshold	Default TTL	Related Object	Category	Source
527	CPU utilization threshold reached	1	1	Critical	1	300	Grid	Maintenance	UFM
528	Fabric interface is down	1	1	Critical	1	300	Grid	Maintenance	UFM
529	UFM standby server problem	1	1	Critical	1	300	Grid	Maintenance	UFM
530	SM is down	1	1	Critical	1	300	Grid	Maintenance	UFM
531	DRBD Bad Condition	1	1	Critical	1	300	Grid	Maintenance	UFM
532	Remote UFM-SM Sync	1	1	Info	1	0	Grid	Maintenance	UFM
533	Remote UFM-SM problem	1	1	Critical	1	0	Site	Maintenance	UFM
535	MH Purge Failed	1	1	Warning	1	300	Grid	Maintenance	UFM
536	UFM Health Watchdog Info	1	1	Info	1	300	Grid	Maintenance	UFM
537	UFM Health Watchdog Critical	1	1	Critical	1	300	Grid	Maintenance	UFM
538	Time Diff Between HA Servers	1	1	Warning	1	300	Grid	Maintenance	UFM
539	DRBD TCP Connection Performance	1	1	Warning	1	900	Grid	Maintenance	UFM
540	Daily Report Completed successfully	1	0	Info	1	300	Grid	Maintenance	UFM
541	Daily Report Completed with Error	1	0	Minor	1	300	Grid	Maintenance	UFM
542	Daily Report Failed	1	0	Critical	1	300	Grid	Maintenance	UFM

Alarm ID	Alarm Name	To Log	Alarm	Default Severity	Default Threshold	Default TTL	Related Object	Category	Source
543	Daily Report Mail Sent successfully	1	0	Info	1	300	Grid	Maintenance	UFM
544	Daily Report Mail Sent Failed	1	0	Minor	1	300	Grid	Maintenance	UFM
545	SM is not responding	1	1	Critical	1	300	Grid	Maintenance	UFM
560	User Connected							Security	UFM
561	User Disconnected							Security	UFM
602	UFM Server Failover	1	1	Critical	1	0	Site	Fabric Notification	UFM
603	Events Suppression	1	0	Critical	0	300	Site	Maintenance	UFM
604	Report Succeeded	1	1	Info	1	300	Grid	Maintenance	UFM
605	Report Failed	1	1	Critical	1	300	Grid	Maintenance	UFM
606	Correction Attempts Paused	1	0	Warning	1	0	Site	Fabric Notification	UFM
701	Non-optimal Link Speed	1	1	Minor	1	0	Port	Hardware	UFM
702	Unhealthy IB Port	1	1	Warning	1	0	Port	Hardware	SM
703	Fabric Collector Connected	1	0	Info	1	0	Grid	Maintenance	UFM
704	Fabric Collector Disconnected	1	1	Critical	1	0	Grid	Maintenance	UFM
750	High data retransmission count on port	1	1	Warning	500	1	Port	Hardware	SM

Alarm ID	Alarm Name	To Log	Alarm	Default Severity	Default Threshold	Default TTL	Related Object	Category	Source
901	Fabric Configuration Started	0	1	Info	1	0	Grid	Fabric Notification	UFM
902	Fabric Configuration Completed	0	1	Info	1	0	Grid	Fabric Notification	UFM
903	Fabric Configuration Failed	0	1	Critical	1	0	Grid	Fabric Notification	UFM
904	Device Configuration Failure	0	1	Critical	1	0	Device	Fabric Notification	UFM
905	Device Configuration Timeout	0	1	Critical	1	0	Device	Fabric Notification	UFM
906	Provisioning Validation Failure	0	1	Critical	1	0	Grid	Fabric Notification	UFM
907	Switch is Down	1	1	Critical	1	0	Site	Fabric Topology	UFM
908	Switch is Up	1	1	Info	1	300	Site	Fabric Topology	UFM
909	Director Switch is Down	1	1	Critical	1	300	Site	Fabric Topology	UFM
910	Director Switch is Up	1	1	Info	1	0	Site	Fabric Topology	UFM
911	Module Temperature Low Threshold Reached	1	1	Warning	60	300	Module	Hardware	Telemetry
912	Module Temperature High Threshold Reached	1	1	Critical	60	300	Module	Hardware	Telemetry
913	Module High Voltage	1	1	Warning	10	420	Switch	Module Status	Telemetry
914	Module High Current	1	1	Warning	10	420	Switch	Module Status	Telemetry
915	BER_ERROR	1	1	Critical	1e-8	420	Port	Hardware	Telemetry

Alarm ID	Alarm Name	To Log	Alarm	Default Severity	Default Threshold	Default TTL	Related Object	Category	Source
916	BER_WARNING	1	1	Warning	1e-13	420	Port	Hardware	Telemetry
917	SYMBOL_BER_ERROR	1	1	Critical	10	420	Port	Hardware	Telemetry
918	High Symbol BER reported	1	1	Warning	10	420	Port	Hardware	Telemetry
919	Cable Temperature High	1	1	Critical	0	0	Port	Hardware	Telemetry
920	Cable Temperature Low	1	1	Critical	0	0	Port	Hardware	Telemetry
1300	SM_SAKEY_VIOLATION	1	1	Warning		5300	Port	Security	SM
1301	SM_SGID_SPOOFED	1	1	Warning		5300	Port	Security	SM
1302	SM_RATE_LIMIT_EXCEEDED	1	1	Warning		5300	Port	Security	SM
1303	SM_MULTICAST_GROUPS_LIMIT_EXCEEDED	1	1	Warning		5300	Port	Security	SM
1304	SM_SERVICES_LIMIT_EXCEEDED	1	1	Warning		5300	Port	Security	SM
1305	SM_EVENT_SUBSCRIPTION_LIMIT_EXCEEDED	1	1	Warning		5300	Port	Security	SM
1306	Unallowed SM was detected in the fabric	1	1	Warning	0	300	Port	Fabric Notification	SM
1307	SMInfo SET request was received from unallowed SM	1	1	Warning	0	300	Port	Fabric Notification	SM
1309	SM was detected with non-matching SMKey	1	1	Warning	0	300	Port	Fabric Notification	SM
1310	Duplicated node GUID was detected	1	1	Critical	1	0	Device	Fabric Notification	SM

Alarm ID	Alarm Name	To Log	Alarm	Default Severity	Default Threshold	Default TTL	Related Object	Category	Source
1311	Duplicated port GUID was detected	1	1	Critical	1	0	Port	Fabric Notification	SM
1312	Switch was Rebooted	1	1	Info	1	0	Device	Fabric Notification	UFM
1315	Topo Config File Error	1	1	Critical	1	0	Grid	Fabric Notification	UFM
1316	Topo Config Subnet Mismatch	1	1	Critical	1	0	Grid	Fabric Notification	Topodiff
1400	High Ambient Temperature	1	1	Warning	0	86400	Switch	Hardware	Switch
1401	High Fluid Temperature	1	1	Warning	0	86400	Switch	Hardware	Switch
1402	Low Fluid Level	1	1	Warning	0	86400	Switch	Hardware	Switch
1403	Low Supply Pressure	1	1	Warning	0	86400	Switch	Hardware	Switch
1404	High Supply Pressure	1	1	Warning	0	86400	Switch	Hardware	Switch
1405	Low Return Pressure	1	1	Warning	0	86400	Switch	Hardware	Switch
1406	High Return Pressure	1	1	Warning	0	86400	Switch	Hardware	Switch
1407	High Differential Pressure	1	1	Warning	0	86400	Switch	Hardware	Switch
1408	Low Differential Pressure	1	1	Warning	0	86400	Switch	Hardware	Switch
1409	System Fail Safe	1	1	Warning	0	86400	Switch	Hardware	Switch
1410	Fault Critical	1	1	Critical	0	86400	Switch	Hardware	Switch

Alarm ID	Alarm Name	To Log	Alarm	Default Severity	Default Threshold	Default TTL	Related Object	Category	Source
1411	Fault Pump1	1	1	Critical	0	86400	Switch	Hardware	Switch
1412	Fault Pump2	1	1	Critical	0	86400	Switch	Hardware	Switch
1413	Fault Fluid Level Critical	1	1	Critical	0	86400	Switch	Hardware	Switch
1414	Fault Fluid Over Temperature	1	1	Critical	0	86400	Switch	Hardware	Switch
1415	Fault Primary DC	1	1	Critical	0	86400	Switch	Hardware	Switch
1416	Fault Redundant DC	1	1	Critical	0	86400	Switch	Hardware	Switch
1417	Fault Fluid Leak	1	1	Critical	0	86400	Switch	Hardware	Switch
1418	Fault Sensor Failure	1	1	Critical	0	86400	Switch	Hardware	Switch
1419	Cooling Device Monitoring Error	1	0	Critical	0	1	Grid	Hardware	Switch
1420	Cooling Device Communication Error	1	1	Critical	0	86400	Switch	Hardware	Switch
1500	New cable detected	1	0	Info	1	0	Link	Security	UFM
1502	Cable detected in a new location	1	0	Warning	1	0	Link	Security	UFM
1503	Duplicate Cable Detected	1	0	Critical	1	0	Link	Security	UFM
1315	Topo Config File Error	1	1	Critical	1	0	Grid	Fabric Notification	UFM
1504	SHARP Allocation Succeeded	1	1	Info	1	0	Grid	SHARP	SHARP

Alarm ID	Alarm Name	To Log	Alarm	Default Severity	Default Threshold	Default TTL	Related Object	Category	Source
1505	SHARP Allocation Failed	1	0	Warning	1	0	Grid	SHARP	SHARP
1506	SHARP Deallocation Succeeded	1	0	Info	1	0	Grid	SHARP	SHARP
1507	SHARP Deallocation Failed	1	0	Warning	1	0	Grid	SHARP	SHARP
1508	Device Collect System Dump Started	1	0	Info	1	300	Device	Maintenance	UFM
1509	Device Collect System Dump Finished	1	0	Info	1	300	Device	Maintenance	UFM
1510	Device Collect System Dump Error	1	0	Critical	1	300	Device	Maintenance	UFM
1511	Virtual Port Added	1	0	Info	1	0	Port	Fabric Notification	SM
1512	Virtual Port Removed	1	0	Warning	1	0	Port	Fabric Notification	SM
1513	Burn Cables Transceivers Started	1	0	Info	1	0	Device	Maintenance	UFM
1514	Burn Cables Transceivers Finished	1	0	Info	1	0	Device	Maintenance	UFM
1515	Burn Cables Transceivers Failed	1	0	Warning	1	0	Device	Maintenance	UFM
1516	Activate Cables Transceivers FW Finished	1	0	Info	1	0	Device	Maintenance	UFM
1517	Activate Cables Transceivers FW Failed	1	0	Warning	1	0	Device	Maintenance	UFM
1520	Aggregation Node Discovery Failed	1	0	Critical	1	0	SHARP AM	SHARP	SHARP
1521	Job Started	1	0	Info	1	0	SHARP AM	SHARP	SHARP

Alarm ID	Alarm Name	To Log	Alarm	Default Severity	Default Threshold	Default TTL	Related Object	Category	Source
1522	Job Ended	1	0	Info	1	0	SHARP AM	SHARP	SHARP
1523	Job Start Failed	1	0	Critical	1	0	SHARP AM	SHARP	SHARP
1524	Job Error	1	0	Critical	1	0	SHARP AM	SHARP	SHARP
1525	Trap QP Error	1	0	Critical	1	0	SHARP AM	SHARP	SHARP
1526	Trap Invalid Request	1	0	Critical	1	0	SHARP AM	SHARP	SHARP
1527	Trap Sharp Error	1	0	Critical	1	0	SHARP AM	SHARP	SHARP
1528	Trap QP Alloc timeout	1	0	Critical	1	0	SHARP AM	SHARP	SHARP
1529	Trap AMKey Violation	1	0	Critical	1	0	SHARP AM	SHARP	SHARP
1530	Unsupported Trap	1	0	Critical	1	0	SHARP AM	SHARP	SHARP
1531	Reservation Updated	1	0	Info	1	0	SHARP AM	SHARP	SHARP
1532	Sharp is not Responding	1	0	Critical	1	0	SHARP AM	SHARP	SHARP
1533	Agg Node Active	1	0	Info	1	0	SHARP AM	SHARP	SHARP
1534	Agg Node Inactive	1	0	Warning	1	0	SHARP AM	SHARP	SHARP
1535	Trap AMKey Violation Triggered by AM	1	0	Warning	1	0	SHARP AM	SHARP	SHARP
1550	Guids Were Added to Pkey	1	0	Info	1	0	Port	Fabric Notification	UFM

Alarm ID	Alarm Name	To Log	Alarm	Default Severity	Default Threshold	Default TTL	Related Object	Category	Source
1551	Guids Were Removed from Pkey	1	0	Info	1	0	Port	Fabric Notification	UFM
1600	VS/CC Classes Key Violation							Security	SM
1602	PCI Speed Degradation Warning	1	1	Warning	1	0	Port	Fabric Notification	UFM
1603	PCI Width Degradation Warning	1	1	Warning	1	0	Port	Fabric Notification	UFM

For a list of AHX related events, please refer to "AHX Monitoring Events".

Appendix – Diagnostic Utilities

> ⚠ For UFM-SDN Appliance, all the below diagnostics commands have ib prefix.
> For example, for UFM-SDN Appliance, the command `ibstat` is `ib ibstat`.

InfiniBand Diagnostics Commands

Command	Description
ibstat	Shows the host adapters status.
ibstatus	Similar to ibstat but implemented as a script.
ibnetdiscover	Scans the topology.
ibaddr	Shows the LID range and default GID of the target (default is the local port).

Command	Description
ibroute	Displays unicast and multicast forwarding tables of the switches.
ibtracert	Displays unicast or multicast route from source to destination.
ibping	Uses vendor MADs to validate connectivity between InfiniBand nodes. On exit, (IP) ping-like output is shown.
ibsysstat	Obtains basic information for the specific node which may be remote. This information includes: hostname, CPUs, memory utilization.
sminfo	Queries the SMInfo attribute on a node.
smpdump	A general purpose SMP utility which gets SM attributes from a specified SMA. The result is dumped in hex by default.
smpquery	Enables a basic subset of standard SMP queries including the following: node info, node description, switch info, port info. Fields are displayed in human readable format.
perfquery	Dumps (and optionally clears) the performance counters of the destination port (including error counters).
ibswitches	Scans the net or uses existing net topology file and lists all switches.
ibhosts	Scans the net or uses existing net topology file and lists all hosts.
ibnodes	Scans the net or uses existing net topology file and lists all nodes.
ibportstate	Gets the logical and physical port states of an InfiniBand port or disables or enables the port (only on a switch). Note: This tool can change port settings. Should be used with caution.
saquery	Issues SA queries.
ibdiagnet	ibdiagnet scans the fabric using directed route packets and extracts all the available information regarding its connectivity and devices.

Command	Description
ibnetsplit	Automatically groups hosts and creates scripts that can be run to split the network into sub-networks each containing one group of hosts.
Ibqueryerrors	Queries IB spec-defined errors from all fabric ports. Note: This tool can change reset port counters Should be used with caution.
smparquery	Queries adaptive-routing related settings from a particular switch. Note: This tool can change reset port counters Should be used with caution.

Diagnostic Tools

Model of operation: All utilities use direct MAD access to operate. Operations that require QP 0 mads only, may use direct routed mads, and therefore may work even in subnets that are not configured. Almost all utilities can operate without accessing the SM, unless GUID to lid translation is required.

Dependencies

Multiple port/Multiple CA support:

When no InfiniBand device or port is specified (as shown in the following example for "Local umad parameters"), the tools select the interface port to use by the following criteria:

1. The first InfiniBand ACTIVE port.
2. If not found, the first InfiniBand port that is UP (physical link up).

If a port and/or CA name is specified, the **tool** attempts to fulfill the user's request and will fail if it is not possible.

For example:

```
ibaddr      # use the 'best port'
ibaddr -C mthca1    # pick the best port from mthca1 only.
ibaddr -P 2     # use the second (active/up) port from the first available IB device.
ibaddr -C mthca0 -P 2   # use the specified port only.
```

Common Options & Flags

Most diagnostics take the following flags. The exact list of supported flags per utility can be found in the usage message and can be shown using util_name -h syntax.

```
# Debugging flags
-d   raise the IB debugging level. May be used several times (-ddd or -d -d -d).
-e   show umad send receive errors (timeouts and others)
-h   show the usage message
-v   increase the application verbosity level.
     May be used several times (-vv or -v -v -v)
-V   show the internal version info.
```

```
# Addressing flags
-D          use directed path address arguments.
            The path is a comma separated list of out ports.
            Examples:
            "0"   # self port
            "0,1,2,1,4"   # out via port 1, then 2, ...
-G          use GUID address arguments.
            In most cases, it is the Port GUID.
            Examples:
            "0x08f1040023"
-s <smlid>  use 'smlid' as the target lid for SA queries.
```

```
# Local umad parameters:
-C <ca_name>     use the specified ca_name.
-P <ca_port>     use the specified ca_port.
-t <timeout_ms>  override the default timeout for the
                 solicited mads.
```

CLI notation: all utilities use the POSIX style notation, meaning that all options (flags) must precede all arguments (parameters).

Utilities Descriptions

ibstatus

A script that displays basic information obtained from the local InfiniBand driver. Output includes LID, SMLID, port state, link width active, and port physical state.

Syntax

```
ibstatus [-h] [devname[:port]]
```

Examples:

```
ibstatus                      # display status of all IB ports
ibstatus mthca1               # status of mthca1 ports
ibstatus mthca1:1 mthca0:2    # show status of specified ports
```

See also: ibstat

ibstat

Similar to the ibstatus utility but implemented as a binary and not as a script. Includes options to list CAs and/or ports.

Syntax

```
ibstat [-d(ebug) -l(ist_of_cas) -p(ort_list) -s(hort)] <ca_name> [portnum]
```

Examples:

```
ibstat            # display status of all IB ports
ibstat mthca1     # status of mthca1 ports
ibstat mthca1 2   # show status of specified ports
ibstat -p mthca0  # list the port guids of mthca0
```

```
ibstat -l         # list all CA names
```

See also: ibstatus

ibroute

Uses SMPs to display the forwarding tables (unicast (LinearForwardingTable or LFT) or multicast (MulticastForwardingTable or MFT)) for the specified switch LID and the optional lid (mlid) range. The default range is all valid entries in the range 1...FDBTop.

Syntax

```
ibroute [options] <switch_addr> [<startlid> [<endlid>]]
```

Nonstandard flags:

```
-a              show all lids in range, even invalid entries.
-n              do not try to resolve destinations.
-M              show multicast forwarding tables. In this case the range
                parameters are specifying mlid range.
node-name-map   node name map file
```

Examples:

```
ibroute 2            # dump all valid entries of switch lid 2
ibroute 2 15         # dump entries in the range 15...FDBTop.
ibroute -a 2 10 20   # dump all entries in the range 10..20
ibroute -n 2         # simple format
ibroute -M 2         # show multicast tables
```

See also: ibtracert

ibtracert

Uses SMPs to trace the path from a source GID/LID to a destination GID/LID. Each hop along the path is displayed until the destination is reached or a hop does not respond. By using the -m option, multicast path tracing can be performed between source and destination nodes.

Syntax

```
ibtracert [options] <src-addr> <dest-addr>
```

Nonstandard flags:

```
-n              simple format; don't show additional information.
-m <mlid>       show the multicast trace of the specified mlid.
-f <force>      force
node-name-map   node name map file
```

Examples:

```
ibtracert 2 23          # show trace between lid 2 and 23
ibtracert -m 0xc000 3 5 # show multicast trace between lid 3
and 5 for mcast lid 0xc000.
```

smpquery

Enables a basic subset of standard SMP queries including the following node info, node description, switch info, port info. Fields are displayed in human readable format.

Syntax

```
smpquery [options] <op> <dest_addr> [op_params]
```

Currently supported operations and their parameters:

```
nodeinfo <addr>
nodedesc <addr>
portinfo <addr> [<portnum>]   # default port is zero
switchinfo <addr>
pkeys <addr> [<portnum>]
s12vl <addr> [<portnum>]
vlarb <addr> [<portnum>]
GUIDInfo (GI) <addr>
MlnxExtPortInfo (MEPI) <addr> [<portnum>]
Combined (-c) : use Combined route address argument
node-name-map : node name map file
```

```
extended (-x) :  use extended speeds
```

Examples:

```
smpquery nodeinfo 2       # show nodeinfo for lid 2
smpquery portinfo 2 5     # show portinfo for lid 2 port 5
```

smpdump

A general purpose SMP utility that gets SM attributes from a specified SMA. The result is dumped in hex by default.

Syntax

```
smpdump [options] <dest_addr> <attr> [mod]
```

Nonstandard flags:

```
-s       show output as string
```

Examples:

```
smpdump -D 0,1,2 0x15 2       # port info, port 2
smpdump 3 0x15 2              # port info, lid 3 port 2
```

ibaddr

Can be used to show the LID and GID addresses of the specified port or the local port by default. This utility can be used as simple address resolver.

Syntax

```
ibaddr [options] [<dest_addr>]
```

Nonstandard flags:

```
gid_show  (-g) : show gid address only
lid_show   (-l) : show lid range only
```

```
Lid_show (-L) : show lid range (in decimal) only
```

Examples:

```
ibaddr                # show local address
ibaddr 2               # show address of the specified port lid
ibaddr -G 0x8f1040023  # show address of the specified port guid
```

sminfo

Issues and dumps the output of an sminfo query in human readable format. The target SM is the one listed in the local port info or the SM specified by the optional SM LID or by the SM direct routed path.

> CAUTION: Using sminfo for any purpose other than a simple query might result in a malfunction of the target SM.

Syntax

```
sminfo [options] <sm_lid|sm_dr_path> [sminfo_modifier]
```

Nonstandard flags:

```
-s <state>          # use the specified state in sminfo mad
-p <priority>       # use the specified priority in sminfo mad
-a <activity>       # use the specified activity in sminfo mad
```

Examples:

```
sminfo         # show sminfo of SM listed in local portinfo
sminfo 2       # query SM on port lid 2
```

perfquery

Uses PerfMgt GMPs to obtain the PortCounters (basic performance and error counters) from the Performance Management Agent (PMA) at the node specified. Optionally show aggregated counters for all ports of node. Also, optionally, reset after read, or only reset counters.

```
perfquery [options]  [<lid|guid> [[port] [reset_mask]]]
```

Nonstandard flags:

```
-a                      Shows aggregated counters for all ports of the destination lid.
-r                      Resets counters after read.
-R                      Resets only counters.
Extended (-x)           Shows extended port counters
Xmtsl (-X)              Shows Xmt SL port counters
Rcvsl  ,( -S)           Shows Rcv SL port counters
Xmtdisc  (-D)           Shows Xmt Discard Details
rcverr, (-E)    Shows Rcv Error Details
extended_speeds (-T)    Shows port extended speeds counters
oprcvcounters   Shows Rcv Counters per Op code
flowctlcounters Shows flow control counters
vloppackets Shows packets received per Op code per VL
vlopdata    Shows data received per Op code per VL
vlxmitflowctlerrors Shows flow control update errors per VL
vlxmitcounters  Shows ticks waiting to transmit counters per VL
swportvlcong    Shows sw port VL congestion
rcvcc   Shows Rcv congestion control counters
slrcvfecn   Shows SL Rcv FECN counters
slrcvbecn   Shows SL Rcv BECN counters
xmitcc  Shows Xmit congestion control counters
vlxmittimecc    Shows VL Xmit Time congestion control counters
smplctl (-c)    Shows samples control
loop_ports (-l)     Iterates through each port
```

Examples:

```
perfquery                # read local port's performance counters
perfquery 32 1          # read performance counters from lid 32, port 1
perfquery -a 32         # read from lid 32 aggregated performance counters
perfquery -r 32 1      # read performance counters from lid 32 port 1 and reset
perfquery -R 32 1      # reset performance counters of lid 32 port 1 only
perfquery -R -a 32     # reset performance counters of all lid 32 ports
perfquery -R 32 2 0xf000  # reset only non-error counters of lid 32 port 2
```

ibping

Uses vendor mads to validate connectivity between InfiniBand nodes. On exit, (IP) ping like output is show. ibping is run as client/server. The default is to run as client. Note also that a default ping server is implemented within the kernel.

Syntax

```
ibping [options] <dest lid|guid>
```

Nonstandard flags:

```
-c <count>      stop after count packets
-f              flood destination: send packets back to back w/o delay
-o <oui>        use specified OUI number to multiplex vendor MADs
-S              start in server mode (do not return)
```

ibnetdiscover

Performs InfiniBand subnet discovery and outputs a human readable topology file. GUIDs, node types, and port numbers are displayed as well as port LIDs and node descriptions. All nodes (and links) are displayed (full topology). This utility can also be used to list the current connected nodes. The output is printed to the standard output unless a topology file is specified.

Syntax

```
ibnetdiscover [options] [<topology-filename>]
```

Nonstandard flags:

```
l   Lists connected nodes
H   Lists connected HCAs
S   Lists connected switches
g   Groups
full (-f)   Shows full information (ports' speed and width, vlcap)
show (-s)   Shows more information
Router_list  (-R)   Lists connected routers
node-name-map   Nodes name map file
cache   filename to cache ibnetdiscover data to
load-cache  filename of ibnetdiscover cache to load
diff    filename of ibnetdiscover cache to diff
diffcheck   Specifies checks to execute for --diff
ports : (-p)    Obtains a ports report
max_hops (-m)   Reports max hops discovered by the library
outstanding_smps (-o)   Specifies the number of outstanding SMP's which should be issued during the scan
```

ibhosts

Traces the InfiniBand subnet topology or uses an already saved topology file to extract the CA nodes.

Syntax

```
ibhosts [-h] [<topology-file>]
```

Dependencies: ibnetdiscover, ibnetdiscover format

ibswitches

Traces the InfiniBand subnet topology or uses an already saved topology file to extract the InfiniBand switches.

Syntax

```
ibswitches [-h] [<topology-file>]
```

Dependencies: ibnetdiscover, ibnetdiscover format

ibportstate

Enables the port state and port physical state of an InfiniBand port to be queried or a switch port to be disabled or enabled.

Syntax

```
ibportstate [-d(ebug) -e(rr_show) -v(erbose) -D(irect) -G(uid) -s smlid -V(ersion) -C ca_name -P ca_port -t
timeout_ms] <dest dr_path|lid|guid> <portnum> [<op>]
```

Supported ops: enable, disable, query, on, off, reset, speed, espeed, fdr10, width, down, arm, active, vls, mtu, lid, smlid, lmc, mkey, mkeylease, mkeyprot

Examples:

```
ibportstate 3 1 disable                 # by lid
ibportstate -G 0x2C9000100D051 1 enable # by guid
ibportstate -D 0 1                      # by direct route
```

ibnodes

Uses the current InfiniBand subnet topology or an already saved topology file and extracts the InfiniBand nodes (CAs and switches).

Syntax

```
ibnodes [<topology-file>]
```

Dependencies: ibnetdiscover, ibnetdiscover format

ibqueryerrors

Queries or clears the PMA error counters in PortCounters by walking the InfiniBand subnet topology.

```
ibqueryerrors [options]
```

Syntax

```
Options:
```

```
--suppress, -s <err1,err2,...>  suppress errors listed
--suppress-common, -c   suppress some of the common counters
--node-name-map <file>  node name map file
--port-guid, -G <port_guid>  report the node containing the port
                             specified by <port_guid>
--, -S <port_guid>      Same as "-G" for backward compatibility
--Direct, -D <dr_path>  report the node containing the port specified
                        by <dr_path>
--skip-sl               don't obtain SL to all destinations
--report-port, -r       report port link information
--threshold-file <val>  specify an alternate threshold file, default: /etc/infiniband-diags/error_thresholds
--GNDN, -R              (This option is obsolete and does nothing)
--data                  include data counters for ports with errors
--switch                print data for switches only
--ca                    print data for CA's only
--router                print data for routers only
--details               include transmit discard details
--counters              print data counters only
--clear-errors, -k      Clear error counters after read
--clear-counts, -K      Clear data counters after read
--load-cache <file>     filename of ibnetdiscover cache to load
--outstanding_smps, -o <val>  specify the number of outstanding SMP's
                              which should be issued during the scan
--config, -z <config>   use config file, default: /etc/infiniband-diags/ibdiag.conf
--Ca, -C <ca>           Ca name to use
--Port, -P <port>       Ca port number to use
--timeout, -t <ms>      timeout in ms
--m_key, -y <key>       M_Key to use in request
--errors, -e            show send and receive errors
--verbose, -v           increase verbosity level
--debug, -d             raise debug level
--help, -h              help message
--version, -V           show version
```

smparquery

Issues Adaptive routing-related queries to the fabric switch.

Syntax

```
```

```
Supported ops (and aliases, case insensitive):
    ARInfo (ARI) <addr>
    ARGroupTable (ARGT) <addr> [<plft>] [<group_table>] [<blocknum>]
    ARLFTTable (ARLT) <addr> [<plft>] [<blocknum>]
    PLFTInfo (PLFTI) <addr>
    PLFTDef (PLFTD) <addr> [<blocknum>]
    PLFTMap (PLFTM) <addr> [<plft>] [<control_map>]
    PortSLToPLFTMap (PLFTP) <addr> [<blocknum>]
    RNSubGroupDirectionTable (DIRT) <addr> [<blocknum>]
    RNGenStringTable (GSTR) <addr> [<plft>] [<blocknum>]
    RNGenBySubGroupPriority (GSGP) <addr>
    RNRcvString (RSTR) <addr> [<blocknum>]
    RNXmitPortMask (RNXM) <addr> [<blocknum>]
    PortRNCounters (RNPC) <addr>


Options:
    Main
        -C|--Ca <ca>                : Ca name to use
        -P|--Port <port>            : Ca port number to use
        -D|--Direct                 : use Direct address argument
        -L|--Lid                    : use LID address argument
        -h|--help                   : help message
        -V|--version                : show version
        -d|--debug                  : Print debug logs
```

saquery

Issues SA queries.

Syntax

```
saquery [-h -d -P -N -L -G -s -g][<name>]
```

Queries node records by default.

d	Enables debugging
P	Gets PathRecord info
N	Gets NodeRecord info
L (-L)	Returns just the Lid of the name specified
G (-G)	Returns just the Guid of the name specified
S (-S)	Returns the PortInfoRecords with isSM capability mask bit on
G (-g)	Gets multicast group info
L (-l)	Returns the unique Lid of the name specified
O (-O)	Returns name for the Lid specified
m(-m)	Gets multicast member info (if multicast group specified, list
x (-x)	member GIDs only for group specified for example 'saquery -m
c (-c)	0xC000')
S (-S)	Gets LinkRecord info"
I (-I)	Gets the SA's class port info
list (-D)	Gets ServiceRecord info
src-to-dst (<src:dst>)	Gets InformInfoRecord (subscription) info
sgid-to-dgid (<sgid-dgid>)	the node desc of the CA's
node-name-map	Gets a PathRecord for <src:dst> where src and dst are either node
smkey <val>	names or LIDs
slid <lid>	Gets a PathRecord for <sgid-dgid> where sgid and dgid are addresses
dlid <lid>	in IPv6 format
mild <lid>	Specifies a node name map file
sgid <gid>	SA SM_Key value for the query. If non-numeric value (like 'x') is
dgid <gid>	specified then saquery will prompt for a value. Default (when not
gid <gid>	specified here or in ibdiag.conf) is to use SM_Key == 0 (or
mgid <gid>	\"untrusted\")
Reversible", 'r', 1, NULL"	Source LID (PathRecord)
numb_path ", 'n', 1, NULL"	Destination LID (PathRecord)
pkey: P_Key (PathRecord, MCMemberRecord).	Multicast LID (MCMemberRecord)
qos_class (-Q)	Source GID (IPv6 format) (PathRecord)
sl	Destination GID (IPv6 format) (PathRecord)
mtu : (-M)	Port GID (MCMemberRecord)
rate (-R)	Multicast GID (MCMemberRecord)
pkt_lifetime	

qkey (-q) (PathRecord, MCMemberRecord). tclass (-T) flow_label : (-F) hop_limit : (-H) scope join_state (-J) proxy_join (-X) service_id	Reversible path (PathRecord) Number of paths (PathRecord) QoS Class (PathRecord) Service level (PathRecord, MCMemberRecord) MTU and selector (PathRecord, MCMemberRecord) Rate and selector (PathRecord, MCMemberRecord) Packet lifetime and selector (PathRecord, MCMemberRecord) If non-numeric value (like 'x') is specified then saquery will prompt for a value. Traffic Class (PathRecord, MCMemberRecord) Flow Label (PathRecord, MCMemberRecord) Hop limit (PathRecord, MCMemberRecord) Scope (MCMemberRecord) Join state (MCMemberRecord) Proxy join (MCMemberRecord) ServiceID (PathRecord)

Dependencies: OpenSM libvendor, OpenSM libopensm, libibumad

ibsysstat

```
ibsysstat [options] <dest lid|guid> [<op>]
```

Nonstandard flags:

```
Current supported operations:
  ping - verify connectivity to server (default)
  host - obtain host information from server
  cpu  - obtain cpu information from server
  -o <oui>    use specified OUI number to multiplex vendor mads
  -S          start in server mode (do not return)
```

ibnetsplit

Automatically groups hosts and creates scripts that can be run in order to split the network into sub-networks containing one group of hosts.

Syntax

- Group:

```
ibnetsplit [-v][-h][-g grp-file] -s <.lst|.net|.topo> <-r head-ports|-d max-dist>
```

- Split:

```
ibnetsplit [-v][-h][-g grp-file] -s <.lst|.net|.topo>
-o out-dir
```

- Combined:

```
ibnetsplit [-v][-h][-g grp-file] -s <.lst|.net|.topo> <-r head-ports|-d max-dist> -o out-dir
```

Usage

- Grouping:
  The grouping is performed if the -r or -d options are provided.
  - If the -r is provided with a file containing group head ports, the algorithm examines the hosts distance from the set of node ports provided in the head-ports file (these are expected to be the ports running standby SM's).
  - If the -d is provided with a maximum distance of the hosts in each group, the algorithm partition the hosts by that distance.

    ⚠ This method of analyzation may not be suitable for some topologies.

  The results of the identified groups are printed into the file defined by the -g option (default ibnetsplit.groups) and can be manually edited. For groups where the head port is a switch, the group file uses the FIRST host port as the port to run the isolation script from.

- Splitting:
  - If the -o flag is included, this algorithm analyzes the MinHop table of the topology and identifies the set of links and switches that may potentially be used for routing each group ports. The cross-switch links between switches of the group to other switches are declared as split-

links and the commands to turn them off using Directed Routes from the original Group Head ports are written into the out-dir provided by the -o flag.

Both stages require a subnet definition file to be provided by the -s flag. The supported formats for subnet definition are:

- *.net - for ibnetdiscover
- *.lst - for opensm-subnet.lst or ibiagnet.lst
- *.topo - for a topology file

HEAD PORTS FILE

This file is provided by the user and defines the ports by which grouping of the other host ports is defined.

Format:

Each line should contain either the name or the GUID of a single port. For switches the port number shall be 0.

```
<node-name>/P<port-num>|<PGUID>
```

GROUPS FILE

This file is generated by the program if the head-ports file is provided to it. Alternatively it can be provided (or edited) by the user if different grouping is desired. The generated script for isolating or connecting the group should be run from the first node in each group.

Format:

Each line may be either:

```
GROUP: <group name>
<node-name>/P<port-num>|<PGUID>
```

ibdiagnet

ibdiagnet scans the fabric using directed route packets and extracts all the available information regarding its connectivity and devices.

It then produces the following files in the output directory (see below):

- "ibdiagnet2.log" - A log file with detailed information.
- "ibdiagnet2.db_csv" - A dump of the internal tool database.
- "ibdiagnet2.lst" - A list of all the nodes, ports and links in the fabric.
- "ibdiagnet2.pm" - A dump of all the nodes PM counters.
- "ibdiagnet2.mlnx_cntrs" - A dump of all the nodes Mellanox diagnostic counters.
- "ibdiagnet2.net_dump" - A dump of all the links and their features.

- "ibdiagnet2.pkey" - A list of all pkeys found in the fabric.
- "ibdiagnet2.aguid" - A list of all alias GUIDs found in the fabric.
- "ibdiagnet2.sm" - A dump of all the SM (state and priority) in the fabric.
- "ibdiagnet2.fdbs" - A dump of unicast forwarding tables of the fabric switches.
- "ibdiagnet2.mcfdbs" - A dump of multicast forwarding tables of the fabric switches.
- "ibdiagnet2.slvl" - A dump of SLVL tables of the fabric switches.
- "ibdiagnet2.nodes_info" - A dump of all the nodes vendor specific general information for nodes who supports it.
- "ibdiagnet2.plft" - A dump of Private LFT Mapping of the fabric switches.
- "ibdiagnet2.ar" - A dump of Adaptive Routing configuration of the fabric switches.
- "ibdiagnet2.vl2vl" - A dump of VL to VL configuration of the fabric switches.

Load plugins from:

/tmp/ibutils2/share/ibdiagnet2.1.1/plugins/

You can specify additional paths to be looked in with "IBDIAGNET_PLUGINS_PATH" env variable.

```
Plugin Name                                   Result     Comment
libibdiagnet_cable_diag_plugin-2.1.1          Succeeded  Plugin loaded
libibdiagnet_phy_diag_plugin-2.1.1            Succeeded  Plugin loaded
```

Syntax

```
[-i|--device <dev-name>] [-p|--port <port-num>]
[-g|--guid <GUID in hex>] [--skip <stage>]
[--skip_plugin <library name>] [--sc]
[--scr] [--pc] [-P|--counter <<PM>=<value>>]
[--pm_pause_time <seconds>] [--ber_test]
[--ber_thresh <value>] [--llr_active_cell <64|128>]
[--extended_speeds <dev-type>] [--pm_per_lane]
[--ls <2.5|5|10|14|25|FDR10|EDR20>]
[--lw <1x|4x|8x|12x>] [--screen_num_errs <num>]
[--smp_window <num>] [--gmp_window <num>]
[--max_hops <max-hops>] [--read_capability <file name>]
[--write_capability <file name>]
[--back_compat_db <version.sub_version>]
[-V|--version] [-h|--help] [-H|--deep_help]
```

```
[--virtual] [--mads_timeout <mads-timeout>]
[--mads_retries <mads-retries>] [-m|--map <map-file>]
[--vlr <file>] [-r|--routing] [--r_opt <[vs,][mcast,]>]
[--sa_dump <file>] [-u|--fat_tree]
[--scope <file.guid>] [--exclude_scope <file.guid>]
[-w|--write_topo_file <file name>]
[-t|--topo_file <file>] [--out_ibnl_dir <directory>]
[-o|--output_path <directory>]
Cable Diagnostic (Plugin)
[--get_cable_info] [--cable_info_disconnected]
Phy Diagnostic (Plugin)
[--get_phy_info] [--reset_phy_info]
```

Options

```
-i|--device <dev-name>        : Specifies the name of the device of the port
                                used to connect to the IB fabric (in case
                                of multiple devices on he local system).
-p|--port <port-num>          : Specifies the local device's port number
                                used to connect to the IB fabric.
-g|--guid <GUID in hex>       : Specifies the local port GUID value of the
                                port used to connect to the IB fabric. If
                                GUID given is 0 than ibdiagnet displays
                                a list of possible port GUIDs and waits
                                for user input.
--skip <stage>                : Skip the executions of the given stage.
                                Applicable skip stages (vs_cap_smp
                                vs_cap_gmp | links | pm |
                                speed_width_check | all).
--skip_plugin <library name>  : Skip the load of the given library name.
                                Applicable skip plugins:
                                (libibdiagnet_cable_diag_plugin-2.1.1 |
                                 libibdiagnet_phy_diag_plugin-2.1.1).
--sc                           : Provides a report of Mellanox counters
--scr                          : Reset all the Mellanox counters (if -sc
                                 option selected).
--pc                           : Reset all the fabric PM counters.
-P|--counter <<PM>=<value>>    : If any of the provided PM is greater then
                                 its provided value than print it.
--pm_pause_time <seconds>      : Specifies the seconds to wait between
                                 first counters sample and second counters
```

```
                                  sample. If seconds given is 0 than no
                                  second counters sample will be done.
                                  (default=1).
--ber_test                       :Provides a BER test for each port.
                                  Calculate BER for each port and check no
                                  BER value has exceeds the BER threshold.
                                  (default threshold="10^-12").
--ber_thresh <value>              :Specifies the threshold value for the
                                  BER test. The reciprocal number of the
                                  BER should be provided. Example: for
                                  10^-12 than value need to be
                                  1000000000000 or 0xe8d4a51000
                                  (10^12).If threshold given is 0 than all
                                  BER values for all ports will be
                                  reported.
--llr_active_cell <64|128>        : Specifies the LLR active cell size
                                    for BER test, when LLR is active in the
                                    fabric.
--extended_speeds <dev-type>        : Collect and test port extended speeds
                                    counters. dev-type: (sw | all).
--pm_per_lane                       : List all counters per lane (when
                                    available).
--ls <0|2.5|5|10|14|25|50|100|FDR10>    : Specifies the expected link speed.
--lw <1x|4x|8x|12x>                 : Specifies the expected link width.
--screen_num_errs <num>             : Specifies the threshold for printing
                                      errors to screen. (default=5).
--smp_window <num>                  : Max smp MADs on wire. (default=8).
--gmp_window <num>                  : Max gmp MADs on wire. (default=128).
--max_hops <max-hops>               : Specifies the maximum hops for the
                                      discovery process. (default=64).
--read_capability <file name>       : Specifies capability masks
                                      configuration file, giving capability
                                      mask configuration for the fabric.
                                      ibdiagnet will use this mapping for
                                      Vendor Specific MADs sending.
--write_capability <file name>      : Write out an example file for
                                      capability masks configuration,
                                      and also the default capability
                                      masks for some devices.
--back_compat_db <version.sub_version> : Show ports section in
                                        "ibdiagnet2.db_csv" according to
                                        given version. Default version 2.0.
```

```
-V|--version                          : Prints the version of the tool.
-h|--help                             : Prints help information (without
                                        plugins help if exists).
-H|--deep_help                        : Prints deep help information
                                        (including plugins help).
--virtual                             : Discover VPorts during discovery
                                        stage.
--mads_timeout <mads-timeout>         : Specifies the timeout (in
                                         milliseconds) for sent and received
                                         mads. (default=500).
--mads_retries <mads-retries>         : Specifies the number of retreis for
                                        every timeout mad. (default=2).
-m|--map <map-file>                   : Specifies mapping file, that maps
                                        node guid to name
                                        (format: 0x[0-9a-fA-F]+ "name").
                                        Maping file can also be specified by
                                        Environment variable
                                        "IBUTILS_NODE_NAME_MAP_FILE_PATH".
--src_lid <src-lid>                   : source lid
--dest_lid <dest-lid>                 : destination lid
--dr_path <dr-path>                   : direct route path
-o|--output_path <directory>          : Specifies the directory where the
                                        Output files will be placed.
                                        (default="/var/tmp/ibdiagpath/").

Cable Diagnostic (Plugin)
--get_cable_info                      : Indicates to query all QSFP cables
                                        for cable information. Cable
                                        information will be stored
                                        in "ibdiagnet2.cables".
--cable_info_disconnected             : Get cable info on disconnected
                                        ports.

Phy Diagnostic (Plugin)
--get_phy_info                        : Indicates to query all ports for phy
                                        information.
--reset_phy_info                      : Indicates to clear all ports phy
                                        information.
```

ibdiagpath

ibdiagpath scans the fabric using directed route packets and extracts all the available information regarding its connectivity and devices. It then produces the following files in the output directory (see below):

- "ibdiagnet2.log" - A log file with detailed information.
- "ibdiagnet2.db_csv" - A dump of the internal tool database.
- "ibdiagnet2.lst" - A list of all the nodes, ports and links in the fabric.
- "ibdiagnet2.pm" - A dump of all the nodes PM counters.
- "ibdiagnet2.mlnx_cntrs" - A dump of all the nodes Mellanox diagnostic counters.
- "ibdiagnet2.net_dump" - A dump of all the links and their features.

Cable Diagnostic (Plugin):

This plugin performs cable diagnostic. It can collect cable info (vendor, PN, OUI etc..) on each valid QSFP cable, if specified.

It produces the following files in the output directory (see below):

- "ibdiagnet2.cables" - In case specified to collect cable info, this file will contain all collected cable info.

Phy Diagnostic (Plugin)

This plugin performs phy diagnostic.

Load Plugins from:

```
/tmp/ibutils2/share/ibdiagnet2.1.1/plugins/
```

You can specify additional paths to be looked in with "IBDIAGNET_PLUGINS_PATH" env variableLoad plugins from:

```
Plugin Name                                   Result     Comment
libibdiagnet_cable_diag_plugin-2.1.1          Succeeded  Plugin loaded
libibdiagnet_phy_diag_plugin-2.1.1            Succeeded  Plugin loaded
```

Syntax

```
[-i|--device <dev-name>] [-p|--port <port-num>]
[-g|--guid <GUID in hex>] [--skip <stage>]
[--skip_plugin <library name>] [--sc]
[--scr] [--pc] [-P|--counter <<PM>=<value>>]
[--pm_pause_time <seconds>] [--ber_test]
[--ber_thresh <value>] [--llr_active_cell <64|128>]
```

```
[--extended_speeds <dev-type>] [--pm_per_lane]
[--ls <2.5|5|10|14|25|FDR10|EDR20>]
[--lw <1x|4x|8x|12x>] [--screen_num_errs <num>]
[--smp_window <num>] [--gmp_window <num>]
[--max_hops <max-hops>] [--read_capability <file name>]
[--write_capability <file name>]
[--back_compat_db <version.sub_version>]
[-V|--version] [-h|--help] [-H|--deep_help]
[--virtual] [--mads_timeout <mads-timeout>]
[--mads_retries <mads-retries>] [-m|--map <map-file>]
[--src_lid <src-lid>] [--dest_lid <dest-lid>]
[--dr_path <dr-path>] [-o|--output_path <directory>]

Cable Diagnostic (Plugin)
[--get_cable_info] [--cable_info_disconnected]
Phy Diagnostic (Plugin)
[--get_phy_info] [--reset_phy_info]
```

Options

-i|--device <dev-name>	:Specifies the name of the device of the port used to connect to the IB fabric (in case of multiple devices on the local system).
-p|--port <port-num>	:Specifies the local device's port number used to connect to the IB fabric.
-g|--guid <GUID in hex>	:Specifies the local port GUID value of the port used to connect to the IB fabric. If GUID given is 0 than ibdiagnet displays a list of possible port GUIDs and waits for user input.
--skip <stage>	:Skip the executions of the given stage. Applicable skip stages: (vs_cap_smp | vs_cap_gmp | links | pm | speed_width_check | all).
--skip_plugin <library name>	:Skip the load of the given library name. Applicable skip plugins: (libibdiagnet_cable_diag_plugin-2.1.1 | libibdiagnet_phy_diag_plugin-2.1.1).
--sc	:Provides a report of Mellanox counters
--scr	:Reset all the Mellanox counters (if -sc option selected).
--pc	:Reset all the fabric PM counters.
-P|--counter <<PM>=<value>>	:If any of the provided PM is greater then its provided value than print it.
--pm_pause_time <seconds>	:Specifies the seconds to wait between first counters sample and
--ber_test	
--ber_thresh <value>	
--llr_active_cell <64|128>	
--extended_speeds <dev-type>	
--pm_per_lane	
:List all counters per lane (when available).	
--ls <2.5|5|10|14|25|FDR10|EDR20>	
--lw <1x|4x|8x|12x>	

```
--screen_num_errs <num>
--smp_window <num>
--gmp_window <num>
--max_hops <max-hops>
--read_capability <file name>
--write_capability <file name>
--back_compat_db <version.sub_version>
-V|--version
-h|--help
-H|--deep_help
--virtual
--mads_timeout <mads-timeout>
--mads_retries <mads-retries>
-m|--map <map-file>
--src_lid <src-lid>
--dest_lid <dest-lid>
--dr_path <dr-path>
-o|--output_path <directory>
Cable Diagnostic (Plugin)
--get_cable_info
--cable_info_disconnected
Phy Diagnostic (Plugin)
--get_phy_info
--reset_phy_info
```

```
second counters sample. If seconds given is 0 than no second
counters sample will be done. (default=1).
:Provides a BER test for each port. Calculate BER for each port and
check no BER value has exceeds the BER threshold.(default
threshold="10^-12").
:Specifies the threshold value for the BER test. The reciprocal
number of the BER should be provided. Example: for 10^-12 than value
need to be 1000000000000 or 0xe8d4a51000(10^12).If threshold given
is 0 than all BER values for all ports will be reported.
:Specifies the LLR active cell size for BER test, when LLR is active
in the fabric.
:Collect and test port extended speeds counters. dev-type: (sw |
all).


:Specifies the expected link speed.
:Specifies the expected link width.
:Specifies the threshold for printing errors to screen. (default=5).
:Max smp MADs on wire. (default=8).
:Max gmp MADs on wire. (default=128).
:Specifies the maximum hops for the discovery process.(default=64).
:Specifies capability masks configuration file, giving capability
mask configuration for the fabric. ibdiagnet will use this mapping
for Vendor Specific MADs sending.
:Write out an example file for capability masks configuration, and
also the default capability masks for some devices.
:Show ports section in "ibdiagnet2.db_csv" according to given
version. Default version 2.0.
:Prints the version of the tool.
:Prints help information (without plugins help if exists).
:Prints deep help information (including plugins help).
:Discover VPorts during discovery stage.
:Specifies the timeout (in milliseconds) for sent and received mads.
(default=500).
:Specifies the number of retries for every timeout mad.(default=2).
```

	:Specifies mapping file, that maps node guid to name (format: 0x[0-9a-fA-F]+ "name"). Mapping file can also be specified by environment variable "IBUTILS_NODE_NAME_MAP_FILE_PATH". :source lid destination lid :direct route path :Specifies the directory where the output files will be placed. (default="/var/tmp/ibdiagpath/"). :Indicates to query all QSFP cables for cable information. Cable information will be stored in "ibdiagnet2.cables". :Get cable info on disconnected ports. :Indicates to query all ports for phy information. :Indicates to clear all ports phy information.

Appendix – Device Management Feature Support

The following table describes the management features available on supported devices.

Feature	10 Gb Ethernet Gateway Module	Grid Director 4700/ 4200/ 4036/ 4036E v3.5	Managed IS5000 Switchesv	Managed SX6000 Switches	Externally Managed IS5000 / SX6000 Switches	Gateway BX5020	HP C-Class	Linux Hosts	Windows Hosts
Discovery									
IB L2 Discovery	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Advanced Discovery (IP, hostname, Hosts: CPU, memory, FW version)	Yes	Yes	No	Yes	No	No	No	Yes with UFM Host Agent	No

Ethernet access Management interface	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes
Provisioning/ Configuration									
IB Partitioning (pkey)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
QoS: SL (SM configuration)	N/A	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
QoS: Rate Limit (SM configuration)	N/A	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Interface/VIF Configuration (IP, hostname, mtu, Bonding)	N/A	N/A	N/A	N/A	N/A	No	N/A	Yes with UFM Host Agent	No
Device Monitoring									
Device Resources: CPU, Memory, Disk	No	Yes	No	No	No	No	No	Yes with UFM Host Agent	No
Get device alerts (Temperature, PS, Fan) **Note:** This feature is not supported on Switch-X switches.	Yes	Yes	No	Yes	Yes	No	No	No	No
L1 (Physical Port) - Monitoring	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
L2-3 (Interface/VIF) - Monitoring	No	No	No	No	No	No	No	Yes with UFM Host Agent	No

Congestion Monitoring per port (enables congestion map)	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Congestion Monitoring per flow (Advanced Package)	No	Yes	No	No	No	No	No	No	No
Device Management									
Add/remove to/from Rack	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Add/remove to/from Logical Server	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Yes	Yes
View/clear Alarms	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
SSH terminal to device	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes
Power On	No	No	No	No	No	No	No	Yes with IPMI	No
Reboot	No	No	No	Yes (SX3606 only)	No	No	No	Yes with IPMI	No
Shutdown	No	No	No	No	No	No	No	Yes with IPMI	No
Port Enable/Disable	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firmware Upgrade (HCA & switch)	No	Yes	No	Yes (Upon SW upgrade - SX6036 only)	No	No	No	Yes	No
Inband Firmware Upgrade (over InfiniBand connection)	No	No	No	No	Yes	No	No	Yes	Yes

Software Upgrade (OFED & switch)	No	Yes	No	Yes (SX3606 only)	No	No	No	Yes with UFM Host Agent	No
Protocols									
Communication UFM Server – Device	IB/SNMP	IB/UDP /SSH	IB	IB/HTTP/ SSH	IB	IB	IB	IB, SSH, IPMI, UDP	IB

1. For a full list of supported IS5000 switches, see Supported IS5000 Switches.
2. QoS Rate Limit (SM configuration): On ConnectX HCAs-only, for hosts.
3. XmitWait counter monitoring requires ConnectX HCAs with firmware version 2.6 and above.
4. This feature requires that the IP address is configured.

Appendix – Used Ports

The following is the list of ports used by the UFM Server for internal and external communication:

Port	Purpose
80(tcp), 443(tcp)	Used by WS clients (Apache Web Server)
694(udp)	Used by Heartbeat - communication between UFM Primary and Standby server
8000(udp)	Used for UFM server listening for REST API requests (redirected by Apache web server)
6306(udp)	Used for Multicast requests – communication with latest UFM Agents
8005(udp)	Used as UFM monitoring listening port

Port	Purpose
8089(tcp)	Used for internal communication between UFM server and MonitoirngHistoryEngine
8888(tcp)	Used by DRBD – communication between UFM Primary and Standby server
15800(tcp)	Used for communication with legacy UFM Agents on Mellanox Grid Director DDR switches
8081(tcp), 8082(tcp)	Used for internal communication with Subnet Manager

Appendix – Routing Chains

The routing chains feature is offering a solution that enables one to configure different parts of the fabric and define a different routing engine to route each of them. The routings are done in a sequence (hence the name "chains") and any node in the fabric that is configured in more than one part is left with the last routing engine updated for it.

Configuring Routing Chains

The configuration for the routing chains feature consists of the following steps:

1. Define the port groups.
2. Define topologies based on previously defined port groups.
3. Define configuration files for each routing engine.
4. Define routing engine chains over defined topologies.

Defining Port Groups

The basic idea behind the port groups is the ability to divide the fabric into sub-groups and give each group an identifier that can be used to relate to all nodes in this group. The port groups are used to define the participants in each of the routing algorithms.

Defining Port Group Policy File

In order to define a port group policy file, set the parameter 'pgrp_policy_file' in the opensm configuration file, as follows:

```
/opt/ufm/files/conf/opensm/port_groups_policy_file.conf
```

Configuring Port Group Policy

The port groups policy file details the port groups in the fabric. The policy file should be composed of one or more paragraphs that define a group. Each paragraph should begin with the line 'port-group' and end with the line 'end-port-group'.

For example:

```
port-group
…port group qualifiers…
end-port-group
```

Port Group Qualifiers

⚠ Unlike the port group's begining and ending which do not require a colon, all qualifiers must end with a colon (':'). Also - a colon is a predefined mark that must not be used inside qualifier values. An inclusion of a colon in the name or the use of a port group, will result in the policy's failure.

Table 62: Port Group Qualifiers

Parameter	Description	Example
name	Each group must have a name. Without a name qualifier, the policy fails.	name: grp1

Parameter	Description	Example
use	‘use’ is an optional qualifier that one can define in order to describe the usage of this port group (if undefined, an empty string is used as a default).	use: first port group

Rule Qualifiers

There are several qualifiers used to describe a rule that determines which ports will be added to the group. Each port group may contain one or more rules of the rule qualifiers in Table 63 (at least one rule shall be defined for each port group).

Table 63: Rule Qualifiers

Parameter	Description	Example
guid list	Comma separated list of guids to include in the group. If no specific physical ports were configured, all physical ports of the guid are chosen. However, for each guid, one can detail specific physical ports to be included in the group. This can be done using the following syntax: • Specify a specific port in a guid to be chosen port-guid: 0x283@3 • Specify a specific list of ports in a guid to be chosen port-guid: 0x286@1/5/7 • Specify a specific range of ports in a guid to be chosen port-guid: 0x289@2-5 • Specify a list of specific ports and ports ranges in a guid to be chosen port-guid: 0x289@2-5/7/9-13/18 • Complex rule port-guid: 0x283@5-8/12/14, 0x286, 0x289/6/8/12	port-guid: 0x283, 0x286, 0x289
port guid range	It is possible to configure a range of guids to be chosen to the group. However, while using the range qualifier, it is impossible to detail specific physical ports. Note: A list of ranges cannot be specified. The below example is invalid and will cause the policy to fail: port-guid-range: 0x283-0x289, 0x290-0x295	port-guid-range: 0x283-0x289
port name	One can configure a list of hostnames as a rule. Hosts with a node description that is built out of these hostnames will be chosen. Since the node description contains the network card index as well, one might also specify a network card index and a physical port to be	port-name: hostname=kuku; port=2; hca_idx=1

Parameter	Description	Example
	chosen. For example, the given configuration will cause only physical port 2 of a host with the node description ‘kuku HCA-1’ to be chosen. port and hca_idx parameters are optional. If the port is unspecified, all physical ports are chosen. If hca_idx is unspecified, all card numbers are chosen. Specifying a hostname is mandatory. One can configure a list of hostname/port/hca_idx sets in the same qualifier as follows: port-name: hostname=kuku; port=2; hca_idx=1 , hostname=host1; port=3, hostname=host2 Note: port-name qualifier is not relevant for switches, but for HCA’s only.	
port regexp	One can define a regular expression so that only nodes with a matching node description will be chosen to the group	port-regexp: SW.*
	It is possible to specify one physical port to be chosen for matching nodes (there is no option to define a list or a range of ports). The given example will cause only nodes that match physical port 3 to be added to the group.	port-regexp: SW.*:3
union rule	It is possible to define a rule that unites two different port groups. This means that all ports from both groups will be included in the united group.	union-rule: grp1, grp2
subtract rule	One can define a rule that subtracts one port group from another. The given rule, for example, will cause all the ports which are a part of grp1, but not included in grp2, to be chosen. In subtraction (unlike union), the order does matter, since the purpose is to subtract the second group from the first one. There is no option to define more than two groups for union/subtraction. However, one can unite/subtract groups which are a union or a subtraction themselves, as shown in the port groups policy file example.	subtract-rule: grp1, grp2

Predefined Port Groups

There are 3 predefined port groups that are available for use, yet cannot be defined in the policy file (if a group in the policy is configured with the name of one of these predefined groups, the policy fails) -

- ALL - a group that includes all nodes in the fabric
- ALL_SWITCHES - a group that includes all switches in the fabric.
- ALL_CAS - a group that includes all HCA's in the fabric.

Port Groups Policy Examples

```
port-group
name: grp3
use: Subtract of groups grp1 and grp2
subtract-rule:  grp1, grp2
end-port-group

port-group
name: grp1
port-guid: 0x281, 0x282, 0x283
end-port-group

port-group
name: grp2
port-guid-range: 0x282-0x286
port-name: hostname=server1 port=1
end-port-group

port-group
name: grp4
port-name: hostname=kika port=1 hca_idx=1
end-port-group

port-group
name: grp3
```

```
union-rule: grp3, grp4
end-port-group
```

Defining Topologies Policy File

In order to define a port group policy file, set the parameter 'topo_policy_file' in the opensm configuration file.

```
/opt/ufm/files/conf/opensm/topo_policy_file.conf
```

Configuring Topology Policy

The topologies policy file details a list of topologies. The policy file should be composed of one or more paragraphs which define a topology. Each paragraph should begin with the line 'topology' and end with the line 'end-topology'.

For example:

```
topology
…topology qualifiers…
end-topology
```

Topology Qualifiers

> ⚠ Unlike topology and end-topology which do not require a colon, all qualifiers must end with a colon (':'). Also – a colon is a predefined mark that must not be used inside qualifier values. An inclusion of a column in the qualifier values will result in the policy's failure.

All topology qualifiers are mandatory. Absence of any of the below qualifiers will cause the policy parsing to fail.

Parameter	Description	Example
id	Topology ID. Legal Values – any positive value. Must be unique.	id: 1
sw-grp	Name of the port group that includes all switches and switch ports to be used in this topology.	sw-grp: some_switches
hca-grp	Name of the port group that includes all HCA's to be used in this topology.	hca-grp: some_hosts

Configuration File per Routing Engine

Each engine in the routing chain can be provided by its own configuration file. Routing engine configuration file is the fraction of parameters defined in the main opensm configuration file.

Some rules should be applied when defining a particular configuration file for a routing engine:

- Parameters that are not specified in specific routing engine configuration file are inherited from the main opensm configuration file.
- The following configuration parameters are taking effect only in the main opensm configuration file:
- qos and qos_* settings like (vl_arb, sl2vl, etc.)
- lmc
- routing_engine

Defining Routing Chain Policy File

In order to define a port group policy file, set the parameter 'rch_policy_file' in the opensm configuration file, as follows:

```
/opt/ufm/files/conf/opensm/routing_chains_policy.conf
```

First Routing Engine in Chain

The first unicast engine in a routing chain must include all switches and HCA's in the fabric (topology id must be 0). The path-bit parameter value is path-bit 0 and it cannot be changed.

Configuring Routing Chains Policy

The routing chains policy file details the routing engines (and their fallback engines) used for the fabric's routing. The policy file should be composed of one or more paragraphs which defines an engine (or a fallback engine). Each paragraph should begin with the line 'unicast-step' and end with the line 'end-unicast-step'.

For example:

```
unicast-step
…routing engine qualifiers…
end-unicast-step
```

Routing Engine Qualifiers

⚠ Unlike unicast-step and end-unicast-step which do not require a colon, all qualifiers must end with a colon (':'). Also – a colon is a predefined mark that must not be used inside qualifier values. An inclusion of a colon in the qualifier values will result in the policy's failure.

Parameter	Description	Example
id	'id' is mandatory. Without an id qualifier for each engine, the policy fails. • Legal values – size_t value (0 is illegal).	is: 1

Parameter	Description	Example
	• The engines in the policy chain are set according to an ascending id order, so it is highly crucial to verify that the id that is given to the engines match the order in which you would like the engines to be set.	
engine	This is a mandatory qualifier that describes the routing algorithm used within this unicast step. Currently, on the first phase of routing chains, legal values are minhop/ftree/updn.	engine: minhop
use	This is an optional qualifier that enables one to describe the usage of this unicast step. If undefined, an empty string is used as a default.	use: ftree routing for cluster 1
config	This is an optional qualifier that enables one to define a separate opensm config file for a specific unicast step. If undefined, all parameters are taken from main opensm configuration file.	config: /etc/config/opensm2.cfg
topology	Define the topology that this engine uses. • Legal value – id of an existing topology that is defined in topologies policy (or zero that represents the entire fabric and not a specific topology). • Default value – If unspecified, a routing engine will relate to the entire fabric (as if topology zero was defined). • Notice: The first routing engine (the engine with the lowest id) MUST be configured with topology: 0 (entire fabric) or else, the routing chain algorithm will fail.	topology: 1
fallback-to	This is an optional qualifier that enables one to define the current unicast step as a fallback to another unicast step. This can be done by defining the id of the unicast step that this step is a fallback to. • If undefined, the current unicast step is not a fallback. • If the value of this qualifier is a non-existent engine id, this step will be ignored. • A fallback step is meaningless if the step it is a fallback to did not fail. • It is impossible to define a fallback to a fallback step (such definition will be ignored)	-

Parameter	Description	Example
path-bit	This is an optional qualifier that enables one to define a specific lid offset to be used by the current unicast step. Setting lmc > 0 in main opensm configuration file is a prerequisite for assigning specific path-bit for the routing engine. Default value is 0 (if path-bit is not specified)	Path-bit: 1

Dump Files per Routing Engine

Each routing engine on the chain will dump its own data files if the appropriate log_flags is set (for instance 0x43).

- The files that are dumped by each engine are:
  - opensm-lid-matrix.dump
  - opensm-lfts.dump
  - opensm.fdbs
  - opensm-subnet.lst

These files should contain the relevant data for each engine topology.

⚠ sl2vl and mcfdbs files are dumped only once for the entire fabric and NOT by every routing engine.

- Each engine concatenates its ID and routing algorithm name in its dump files names, as follows:
  - opensm-lid-matrix.2.minhop.dump
  - opensm.fdbs.3.ftree
  - opensm-subnet.4.updn.lst
- If a fallback routing engine is used, both the routing engine that failed and the fallback engine that replaces it, dump their data.

If, for example, engine 2 runs ftree and it has a fallback engine with 3 as its id that runs minhop, one should expect to find 2 sets of dump files, one for each engine:

- opensm-lid-matrix.2.ftree.dump
- opensm-lid-matrix.3.minhop.dump
- opensm.fdbs.2.ftree

- opensm.fdbs.3.munhop

Appendix – Adaptive Routing

> ⚠ As of UFM v6.4, Adaptive Routing plugin is no longer required for Adaptive Routing and SHIELD configuration. AR is now part of the core Subnet Manager implementation. However, upgrading UFM to v6.4 from an earlier version using the AR plugin will remain possible.

For information on how to set up AR and SHIELD, please refer to How-To Configure Adaptive Routing and Self Healing Networking.

Appendix – Configuration Files Auditing

The main purpose of this feature is to allow users to track changes made to selected configuration files. When activating the feature, all the changes are reflected in specific log files which contain information about the changes and when they took place.

To activate this feature:

In *TrackConfig* section in gv.cfg, file value of *track_config* key should be set to **true** and value of *track_conf_files* key should contain a comma-separated list of defined conf files to be tracked.
By default - ALL conf-files are tracked. To activate the feature, after *track_config* key is set to true, the UFM server should be restarted.

Example:

```
[TrackConfig]
# track config files changes
track_config = true
# Could be selected options (comaseparated) UFM, SM, SHARP, Telemetry. Or ALL for all the files.
track_conf_files = ALL
```

Alternatively, you may run the command "ufm track-conf-changes enable."

The below lists the configuration files that can be tracked:

Conf File Alias	Configuration Files
UFM	conf/gv.cfg
SM	conf/opensm/opensm.conf
SHARP	conf/sharp2/sharp_am.cfg
Telemetry	conf/telemetry/launch_ibdiagnet_config.ini
ALL	All the above configuration files.

Once the feature is activated and the UFM server is restarted, the UFM generates file which list the changes made in each of the tracked conf files. These files are located in the auditing folder which is part of the UFM sys-dump. The file naming convention is as follows: original conf file name with audit.log suffix.

Example: For gv.cfg, the name of the changes-tracking file is gv.cfg.audit.log. Changes are stored in auditing files in “linux diff”-like format.

Example:

```
cat /opt/ufm/files/auditing/gv.cfg.audit.log
=== Change occurred at 2022-07-24 07:31:48.679247 ===
---
+++
@@ -45,7 +45,7 @@
mon_mode_discovery_period = 60
check_interface_retry = 5
# The number of times to try if the InfiniBand fabric interface is down. The duration of each retry is 1 second.
-ibport_check_retries = 90
+ibport_check_retries = 92
ws_address = UNDEFINED
ws_port = 8088
ws_protocol = https
```

Appendix - Managed Switches Configuration Info Persistency

UFM uses a periodic system information-pulling mechanism to query managed switches inventory data. The inventory information is saved in local JSON files for persistency and tracking of managed switches' status.

Upon UFM start up, UFM loads the saved JSON files to present them to the end user via REST API or UFM WEB UI.

After UFM startup is completed, UFM pulls all managed switches data and updates the JSON file and the UFM model periodically (the interval is configurable). In addition, the JSON files are part of UFM system dump.

The following parameters allow configuration of the feature via gv.cfg fie:

```
[SrvMgmt]
# how often UFM should send json requests for sysinfo to switches (in seconds)
systems_poll = 180
# To create UFM model in large setups might take a lot of time.
# This is an initial delay (in minutes) before starting to pull sysinfo from switches.
systems_poll_init_timeout = 5
# to avoid sysinfo dump overloading and multiple writing to host
# switches sysinfo will be dumped to disc in json format every set in this variable
# sysinfo request. If set to 0 - will not be dumped, if set to 1 - will be dumped every sysinfo request
# this case (as example defined below) dump will be created every fifth sysinfo request, so if system_poll is 180
 sec (3 minutes) sysinfo dump to the file will e performed every 15 minutes.
sysinfo_dump_interval = 5
# location of the sysinfo dump file (it is in /opt/ufm/files/logs (it will be part of UFM dump)
sysinfo_dump_file_path = /opt/ufm/files/log/sysinfo.dump
```

Appendix – Security Features

SA Enhanced Trust Model (SAETM)

Standard SA has a concept of trust-based requests on the SA_Key that is part of each SA MAD. A **trusted request** is when the SA_Key value is not equal to zero but equals the SA configured value, while an **untrusted request** is when the SA_Key value equals zero in the request. If a request has a non-zero SA_Key value that is different from the configured SA key, it will be dropped and reported.

When SAETM is enabled, the SA limits the set of untrusted requests allowed. Untrusted requests that are not allowed according to SAETM will be silently dropped (for the set of untrusted requests allowed, see the following section below).

SAETM feature is disabled by default. To enable it, set the sa_enhanced_trust_model parameter to TRUE.

Additional SAETM Configuration Parameters

Parameter	Description
sa_etm_allow_untrusted_guidinfo_rec	Defines whether to allow GUIDInfoRecord as part of the SAETM set of untrusted requests allowed (see section below)
sa_etm_allow_guidinfo_rec_by_vf	Defines whether to drop GUIDInfoRecord from non-physical ports (see section below)
sa_etm_allow_untrusted_proxy_requests	Defines the behavior for proxy requests (see section below)
sa_etm_max_num_mcgs/ sa_etm_max_num_srvcs/ sa_etm_max_num_event_subs	Defines the registration limits in SAETM (see section below)

Set of Untrusted SA Requests Allowed

The following table lists the untrusted requests allowed when SAETM is enabled:

Request	Request Type
MCMemberRecord	Get/Set/Delete
PathRecord	Get
PathRecord	GetTable (only if both destination and source are specified (e,g. only point to point))
ServiceRecord	Get/Set/Delete
ClassPortInfo	Get
InformInfo	Set (for non-SM security traps)
GUIDInfoRecord	Set/Delete - this request can only be part of this set depending on the values of sa_etm_allow_untrusted_guidinfo_rec and sa_etm_allow_guidinfo_rec_by_vf - see elaboration below.

When sa_etm_allow_untrusted_guidinfo_rec is set to FALSE (and SAETM is enabled), the SA will drop GUIDInfoRecord Set/Delete untrusted requests.

When sa_etm_allow_guidinfo_rec_by_vf is set to FALSE (and SAETM is enabled), the SA will drop GUIDInfoRecord Set/Delete requests from non-physical ports.

If sa_etm_allow_untrusted_guidinfo_rec=FALSE, GUIDInfoRecord Set/Delete requests will become part of the SAETM set of untrusted requests allowed. Note that if sa_etm_allow_guidinfo_rec_by_vf=FALSE, the requests will only be allowed from physical ports.

Proxy SA Requests

SA modification request (SET/DELETE) is identified as a proxy operation when the port corresponding with the requester source address (SLID from LRH/ SGID from GRH) is diffident than the port for which the request applies:

- For MCMemberRecord, when the MCMemberRecord.PortGID field does not match the requester address
- For ServiceRecord, when the ServiceRecord.ServiceGID field does not match requester address
- For the GUIDInfoRecord, when the LID field in the RID of the record does not match the requester address

When sa_etm_allow_untrusted_proxy_requests is set to FALSE and SAETM is enabled, untrusted proxy requests will be dropped.

Registration Limits

When any of sa_etm_max_num_mcgs, sa_etm_max_num_srvcs or sa_etm_max_num_event_subs parameters is set to 0, the number of this parameter's registrations can be unlimited. When the parameter's value is different than 0, attempting to exceed the maximum number of registrations will result in the request being silently dropped. Consequently, the requester and request info will be logged, and an event will be generated for the Activity Manager.

The following parameters control the maximum number of registrations:

Parameter	Description
sa_etm_max_num_mcgs	Maximum number of multicast groups per port/vport that can be registered.
sa_etm_max_num_srvcs	Maximum number of service records per port/vport that can be registered.
sa_etm_max_num_event_subs	Maximum number of event subscriptions (InformInfo) per port/vport that can be registered.

SAETM Logging

When requesting an operation that is not part of the SAETM set of untrusted requests, it will be silently dropped and eventually written to the SM log.

The logging of the dropped MADs is repressed to not overload the OpenSM log. If the request that needs to be dropped was received from the same requester many times consecutively, OpenSM logs it only if the request number is part of the following sequence:

0, 1, 2, 5, 10, 20, 50, 100, 200... (similar to the trap log repression).

SGID Spoofing

SA can validate requester addresses by comparing the SLID and SGID of the incoming request. SA determines the requester port by the SLID and SGID field of the request. SGID spoofing is when the SGID and SLID do not match.

When sa_check_sgid_spoofing parameter is enabled, SA checks for SGID spoofing in every request that includes GRH, unless the SLID belongs to a router port in that same request. In case the request SGID does not match its SLID, the request will be dropped. The default value of this parameter is TRUE.

M_Key Authentication

M_Key Authentication Enablement

In order to enable M_Key authentication in the InfiniBand fabric, the following parameters must be set in opensm.conf:

Argument	Value	Description
m_key	64-bit integer Default: 0	The value must be set to some random number.
m_key_protection	0-2 Default: 0	• 0 – weakest level of protection SubnGet(*) shall succeeds for any key in the MADHeader:M_Key and SubnGetResp(PortInfo) shall return the contents of the PortInfo:M_Key component. SubnSet(*) and SubnTrapRepress(*) shall fail if MADHeader:M_Key does not match the PortInfo:M_Key component in the port. • 1 SubnGet(*) shall succeed for any key in the MADHeader:M_Key and SubnGetResp(PortInfo) shall return the contents of the PortInfo:M_Key component set to zero if MADHeader:M_Key does not match the PortInfo:M_Key component in the port.

Argument	Value	Description
		SubnSet(*) and SubnTrapRepress(*) shall fail if MADHeader: M_Key does not match the PortInfo:M_Key component in the port. • 2 SubnGet(*), SubnSet(*), and SubnTrapRepress(*) shall fail if MADHeader:M_Key does not match the PortInfo:M_Key component in the port.
m_key_lease_period	0-65535 Default: 0	The lease period used for the M_Key on this subnet in seconds. Recommended value is 60 seconds.
m_key_lookup	TRUE/FALSE Default: FALSE	Must be enabled when M_key is non-zero

M_Key Per Port

This feature increases protection on the fabric as a unique M_Key is generated and set for each HCA, router, or switch port.

OpenSM calculates an M_Key per port by performing a hash function on the port GUID of the device and the M_Key configured in opensm.conf.

To enable M_Key per port, set the parameter below in addition to the parameters listed in the previous section:

```
m_key_per_port TRUE
```

Once enabled, OpenSM forces the values of the following parameters:

```
m_key 0x1 (unless configured to non-zero value)
m_key_protection to 2 (unless configured to 3)
m_key_lookup to TRUE
```

Subnet Manager Protection

To protect UFM subnet manager from a hostile SM that may be enabled in the fabric, the SM_Key parameter must be set to some random value in addition to the M_Key protection described before:

```
sm_key <random_64b_integer>
```

Once a hostile SM is detected and queried by UFM SM, UFM SM compares the SM_Key provided by the hostile SM to the SM_Key configured in UFM opensm.conf.

As UFM SM_Key is a random 64-bit number, there is a high probability that the SM_Key provided by hostile SM will not match the UFM SM_Key.

As a result UFM SM, ignores hostile SMs and reports them in opensm.log and the syslog.

Example from opensm.log:

```
ERR 2F18: Got SM <direct_path_to_the_hostile_SM_node> with sm_key <hostile_SM_KEY> that doesn't match our local
sm_key. Ignoring SMInfo.
```

Example from syslog:

```
Found remote SM <direct_path_to_the_remote_SM> with non-matching sm_key
```

Appendix – NVIDIA SHARP Integration

NVIDIA Scalable Hierarchical Aggregation and Reduction Protocol (SHARP)™

NVIDIA SHARP is a technology that improves the performance of MPI operation by offloading collective operations from the CPU and dispatching to the switch network, and eliminating the need to send data multiple times between endpoints. This approach decreases the amount of data traversing the network as aggregation nodes are reached, and dramatically reduces the MPI operation time.

NVIDIA SHARP software is based on:

- Hardware capabilities in Switch-IB™ 2
- Hierarchical communication algorithms (HCOL) library into which NVIDIA SHARP capabilities are integrated
- NVIDIA SHARP daemons, running on the compute nodes
- NVIDIA SHARP Aggregation Manager, running on UFM

1. These components should be installed from HPCX or MLNX_OFED packages on compute nodes. Installation details can be found in SHARP Deployment Guide.

NVIDIA SHARP Aggregation Manager

Aggregation Manager (AM) is a system management component used for system level configuration and management of the switch-based reduction capabilities. It is used to set up the NVIDIA SHARP trees, and to manage the use of these entities.

AM is responsible for:

- NVIDIA SHARP resource discovery
- Creating topology aware NVIDIA SHARP trees
- Configuring NVIDIA SHARP switch capabilities
- Managing NVIDIA SHARP resources
- Assigning NVIDIA SHARP resource upon request
- Freeing NVIDIA SHARP resources upon job termination

AM is configured by a topology file created by Subnet Manager (SM): subnet.lst. The file includes information about switches and HCAs.

NVIDIA SHARP AM Prerequisites

In order for UFM to run NVIDIA SHARP AM, the following conditions should be met:

- Managed InfiniBand fabric must include at least one of the following Switch-IB 2 switches with minimal firmware version of 15.1300.0126:
  - CS7500
  - CS7510
  - CS7520
  - MSB7790

    - MSB7800
- NVIDIA SHARP software capability should be enabled for all Switch-IB 2 switches in the fabric (a dedicated logical port #37, for NVIDIA SHARP packets transmission, should be enabled and should be visible via UFM).
- UFM OpenSM should be running to discover the fabric topology.

NVIDIA SHARP AM is tightly dependent on OpenSM as it uses the topology discovered by OpenSM.

- NVIDIA SHARP AM should be enabled in UFM configuration by running:

```
[Sharp]
sharp_enabled = true
```

  Alternatively, you may run the command "ib sharp enable".

NVIDIA SHARP AM Configuration

By default, when running NVIDIA SHARP AM by UFM, there is no need to run further configuration. To modify the configuration of NVIDIA SHAPR AM, you can edit the following NVIDIA SHARP AM configuration file:`conf/sharp/sharp_am.cfg`.

Running NVIDIA SHARP AM in UFM

To run NVIDIA SHARP AM within UFM, do the following:

1. Make sure that the root GUID configuration file (root_guid.conf) exists in conf/opensm. This file is required for activating NVIDIA SHARP AM.
2. Enable NVIDIA SHARP in conf/opensm/opensm.conf OpenSM configuration file by running "ib sm sharp enable" or by setting the sharp_enabled parameter to 2:

```
# SHArP support
# 0: Ignore SHArP - No SHArP support
# 1: Disable SHArP - Disable SHArP on all supporting switches
# 2: Enable SHArP - Enable SHArP on all supporting switches
sharp_enabled 2
```

3. Make sure that port #6126 (on which NVIDIA SHARP AM is communicating with NVIDIA SHARP daemons) is not being used by any other application. If the port is being used, you can change it by modifying smx_sock_port parameter in the NVIDIA SHARP AM configuration file: conf/sharp2/sharp_am.cfg or via the command "ib sharp port".
4. Enable NVIDIA SHARP AM in conf/gv.cfg UFM configuration file by running the command "ib sharp enable" or by setting the sharp_enabled parameter to true (it is false by default):

```
[Sharp]
sharp_enabled = true
```

5. (Optional) Enable NVIDIA SHARP allocation in conf/gv.cfg UFM configuration file by setting the sharp_allocation_enabled parameter to true (it is false by default):

```
[Sharp]
sharp_allocation_enabled = true
```

> ⚠ If the field sharp_enabled, and sharp_allocation_enabled are both set as true in gv.cfg, UFM sends an allocation (reservation) request to NVIDIA SHARP Aggregation Manager (AM) to allocate a list of GUIDs to the specified PKey when a new “Set GUIDs for PKey” REST API is called. If an empty list of GUIDs is sent, a PKEY deallocation request is sent to the SHARP AM.
> NVIDIA SHARP allocations (reservations) allow SHARP users to run jobs on top of these resource (port GUID) allocations for the specified PKey. For more information, please refer to the *UFM REST API Guide* under Actions REST API → PKey GUIDs → Set/Update PKey GUIDs.

Operating NVIDIA SHARP AM with UFM

If NVIDIA SHARP AM is enabled, running UFM will run NVIDIA SHARP AM, and stopping UFM will stop NVIDIA SHARP AM.

To start UFM with NVIDIA SHARP AM (enabled):

```

```

```
/etc/init.d/ufmd start
```

The same command applies to HA, using /etc/init.d/ufmha.

Upon startup of UFM or SHARP Aggregation Manager, UFM will resend all existing persistent allocation to SHARP AM.

- **To stop UFM with NVIDIA SHARP AM (enabled):**

```
/etc/init.d/ufmd stop
```

- **To stop only NVIDIA SHARP AM while leaving UFM running:**

```
/etc/init.d/ufmd sharp_stop
```

- **To start only NVIDIA SHARP AM while UFM is already running:**

```
/etc/init.d/ufmd sharp_start
```

Upon startup of UFM or SHARP Aggregation Manager, UFM will resend all existing persistent allocation to SHARP AM.

- **To restart only NVIDIA SHARP AM while UFM is running:**

```
/etc/init.d/ufmd sharp_restart
```

Upon startup of UFM or SHARP Aggregation Manager, UFM will resend all existing persistent allocation to SHARP AM.

- **To display NVIDIA SHARP AM status while UFM is running:**

```
/etc/init.d/ufmd sharp_status
```

Monitoring NVIDIA SHARP AM by UFMHealth

UFMHealth monitors SHARP AM and verifies that NVIDIA SHARP AM is always running. When UFMHealth detects that NVIDIA SHARP AM is down, it will try to re-start it, and will trigger an event to the UFM to notify it that NVIDIA SHARP AM is down.

Managing NVIDIA SHARP AM by UFM High Availability (HA)

In case of a UFM HA failover or takeover, NVIDIA SHARP AM will be started on the new master node using the same configuration that was used prior to the failover/takeover.

NVIDIA SHARP AM Logs

NVIDIA SHARP AM log file (sharp_am.log) at /opt/ufm/files/log.

NVIDIA SHARP AM log files are rotated by UFM logrotate mechanism.

NVIDIA SHARP AM Version

NVIDIA SHARP AM version can be found at /opt/ufm/sharp/share/doc/SHARP_VERSION.

Appendix – AHX Monitoring

AHX monitoring enables monitoring HDR director switch cooling devices (i.e. AHX) and sending events to the UFM. The events are triggered on the switch associated with the cooling device if the monitoring utility encounters an issue.

The monitoring utility runs periodically and communicates with the AHX devices over the Modbus protocol (TCP port 502).

Configuring AHX Monitoring

1. Enable AHX monitoring. Run:

```
ufmapl [ mgmt-ha-active ] (config) # ib managed-switch ahx-monitor enable
```

2. Set the AHX monitoring interval. Run:

```
ufmapl [ mgmt-ha-active ] (config) # ib managed-switch ahx-monitor interval
```

3. Add AHX devices for monitoring. Run:

```
ufmapl [ mgmt-ha-active ] (config) # ib managed-switch ahx-monitor device
```

4. [Optional] Review the settings. Run:

```
ufmapl [ mgmt-ha-active ] (config) # show ib managed-switch ahx-monitor
  AHX Monitoring:
    Enabled : Yes
    Interval: 1m


  AHX Devices:
    Switch name          : switch-01
    Primary IP address   : 10.10.1.11
    Secondary IP address: 10.11.1.11
```

AHX Monitoring Events

Alarm ID	Alarm Name	To Log	Alarm	Severity	Threshold	TTL	Related Object	Category	Message
1400	COOLING_DEV_HIGH_AMBIENT_TEMP	1	1	Warning		86400	Switch	Hardware	High Ambient Temperature
1401	COOLING_DEV_HIGH_FLUID_TEMP	1	1	Warning		86400	Switch	Hardware	High Fluid Temperature
1402	COOLING_DEV_LOW_FLUID_LEVEL	1	1	Warning		86400	Switch	Hardware	Low Fluid Level
1403	COOLING_DEV_LOW_SUPPLY_PRESS	1	1	Warning		86400	Switch	Hardware	Low Supply Pressure
1404	COOLING_DEV_HIGH_SUPPLY_PRESS	1	1	Warning		86400	Switch	Hardware	High Supply Pressure
1405	COOLING_DEV_LOW_RETURN_PRESS	1	1	Warning		86400	Switch	Hardware	Low Return Pressure
1406	COOLING_DEV_HIGH_RETURN_PRESS	1	1	Warning		86400	Switch	Hardware	High Return Pressure
1407	COOLING_DEV_HIGH_DIFF_PRESS	1	1	Warning		86400	Switch	Hardware	High Differential Pressure
1408	COOLING_DEV_LOW_DIFF_PRESS	1	1	Warning		86400	Switch	Hardware	Low Differential Pressure
1409	COOLING_DEV_SYSTEM_FAIL_SAFE	1	1	Warning		86400	Switch	Hardware	System Fail Safe
1410	COOLING_DEV_FAULT_CRITICAL	1	1	Critical		86400	Switch	Hardware	Fault Critical
1411	COOLING_DEV_FAULT_PUMP1	1	1	Critical		86400	Switch	Hardware	Fault Pump1
1412	COOLING_DEV_FAULT_PUMP2	1	1	Critical		86400	Switch	Hardware	Fault Pump2

Alarm ID	Alarm Name	To Log	Alarm	Severity	Threshold	TTL	Related Object	Category	Message
1413	COOLING_DEV_FLUID_LEVEL_CRIT	1	1	Critical		86400	Switch	Hardware	Fault Fluid Level Critical
1414	COOLING_DEV_FLUID_OVERTEMP	1	1	Critical		86400	Switch	Hardware	Fault Fluid Over Temperature
1415	COOLING_DEV_FAULT_PRIMARY_DC	1	1	Critical		86400	Switch	Hardware	Fault Primary DC
1416	COOLING_DEV_FAULT_REDUND_DC	1	1	Critical		86400	Switch	Hardware	Fault Redundant DC
1417	COOLING_DEV_FAULT_FLUID_LEAK	1	1	Critical		86400	Switch	Hardware	Fault Fluid Leak
1418	COOLING_DEV_SENSOR_FAILURE	1	1	Critical		86400	Switch	Hardware	Fault Sensor Failure
1419	COOLING_DEV_MONITOR_ERROR	1	0	Critical		1	Grid	Hardware	Cooling Device Monitoring Error
1420	COOLING_DEV_COMM_ERROR	1	1	Critical		86400	Switch	Hardware	Cooling Device Communication Error

Appendix – UFM Event Forwarder

UFM event forwarding enables forwarding the UFM events to Fluentd.

```
2020-10-13 19:40:10.183000000 +0000 fluent_bit:
{"pri":"10","logger":"eventlog","pid":"119194","severity":"CRITICAL","time":"2020-10-13
19:40:10.183","index":"17","id":"394","event_severity":"CRITICAL","site":"ufm-nvidia-
test","category":"Module_Status","object":"Switch","object_desc":"[default(4) / Switch: r-ufm-sw121 / NA / NA]
[dev_id: 0c42a1030060acd0]","message":"Module PS 1 on r-ufm-sw121(10.209.36.90) status is fatal\u0000"}
```

To configure UFM events forwarder:

1. Enable sending UFM events to UFM events forwarder (internal) container. Run:

```
ufmapl [mgmt-ha-active] (config) # ufm logging syslog enable
ufmapl [mgmt-ha-active] (config) # ufm logging syslog ufm-events enable
```

2. Set the UFM events forwarder (internal) container IPv4 address and listening port. Run:

```
ufmapl [mgmt-ha-active] (config) ufm logging syslog 127.0.0.1:5140
```

3. Enable docker service. Run:

```
ufmapl [mgmt-ha-active] (config) # no docker shutdown
```

4. Deploy the container. Run:

```
ufmapl [mgmt-ha-active] (config) # docker pull mellanox/ufm-events-forwarder
```

5. Enable forwarding the UFM events. Run:

```
ufmapl [mgmt-ha-active] (config) # ufm events-forwarder enable
```

6. Set the Fluentd IPv4 address and listening port. Run:

```
ufmapl [mgmt-ha-active] (config) # ufm events-forwarder server ip 10.215.1.7 port 24224
```

7. Set UFM site name. Run:

```
ufmapl [mgmt-ha-active] (config) # ufm site-name ufm.nvidia-test.net
```

8. Save the configuration changes. Run:

```
ufmapl [mgmt-ha-active] (config) # configuration write
```

9. Once all configurations are set, start the UFM service. Run:

```
ufmapl [mgmt-ha-active] (config) # ufm start
```

For more information regarding the commands used, please refer to the UFM Appliance Command Reference Guide.

Appendix – UFM Migration

Overview

UFM migration enables backup and restores UFM configuration files.

Backup UFM configuration

By default, the following folders (placed in `/opt/ufm/files`) are being backed up:

- conf
- dashboardViews
- licenses
- networkViews
- scripts
- sqlite
- templates/user-defined
- ufmhealth/scripts
- userdata
- users_preferences

⚠ The user may also backup the UFM historical telemetry data ("-t" argument).

UFM (Bare Metal)

```
/opt/ufm/scripts/ufm_backup.sh --help
usage: ufm_backup.pyc [-h] [-f BACKUP_FILE] [-t]
```

Optional Arguments

-h	--help	show this help message and exit
-f	--backup-file BACKUP_FILE	full path of zip file to be generated
-t	--telemetry	backup UFM historical telemetry

UFM Docker Container

1. Backup UFM configuration. Run:

   ```
   docker exec ufm /opt/ufm/scripts/ufm_backup.sh
   ```

2. Copy the backup file from UFM docker container to the host. Run:

   ```
   docker cp ufm:/root/<backup file> <path on host>
   ```

UFM Appliance

1. Backup UFM configuration. Run:

```
ufm data backup [with-telemetry]
```

2. Upload the backup file to a remote host. Run:

```
ufm data upload <backup file> <upload URL>
```

⚠ More details can be found in the log file `/tmp/ufm_backup.log`.

Restore UFM Configuration

⚠ All folders which are a part of the UFM backup are restored (filter is done during the backup stage).

UFM Bare Metal

```
/opt/ufm/scripts/ufm_restore.sh --help
usage: ufm_restore.pyc [-h] -f BACKUP_FILE [-u] [-v]
```

Optional Arguments

-h	--help	show this help message and exit
-f BACKUP_FILE	--backup-file BACKUP_FILE	full path of zip file generated by backup script
-u	--upgrade	upgrades the restored UFM files
-v	--verbose	makes the operation more talkative

UFM Docker Container

1. Stop UFM. Run:

```
docker exec ufm /etc/init.d/ufmd stop
```

2. Copy the backup file from the host into UFM docker container. Run:

```
docker cp <backup file> ufm:/tmp/<backup file>
```

3. Restore UFM configuration. Run:

```
docker exec ufm /opt/ufm/scripts/ufm_restore.sh -f /tmp/<backup file> [--upgrade]
```

4. Start UFM. Run:

```
docker exec ufm /etc/init.d/ufmd start
```

UFM Appliance

1. Stop UFM. Run:

```
no ufm start
```

2. Copy the backup file from a remote host into UFM appliance. Run:

```
ufm data fetch <download URL>
```

3. Restore UFM configuration. Run:

```
ufm data restore <backup file>
```

4. Start UFM. Run:

```
ufm start
```

> ⚠ When restoring the UFM configuration from host to a container, the following parameters in `/opt/ufm/files/conf/gv.cfg` may be reset the following:
> - fabric_interface
> - ufma_interfaces
> - mgmt_interface

> ⚠ UFM configuration upgrade during restore is not supported in UFM Appliance GEN2/GEN2.5

More details can be found in the log files `/tmp/ufm_restore.log` and `/tmp/ufm_restore_upgrade.log`

Appendix - Switch Grouping

To facilitate the logical grouping of 1U switches into a "director-like switch" group, the UFM implements a special dedicated group of interconnected 1U switches based on a YAML configuration file. This group, which is of type "superswitch", only includes 1U switches connected to each other, with some functioning as lines and others as spines.

To access the configuration file for superswitches, users can define the path in the [SubnetManager] section of the `gv.cfg` file, using the variable name "`super_switch_config_file_path`". For instance, the path can be specified as follows: `super_switch_config_file_path=/opt/ufm/files/conf/super_switches_configuration.cfg`.

It is important to note that the file must be located in the `/opt/ufm/files file` tree, as it should be replicated between master and slave UFM servers in a high-availability configuration.

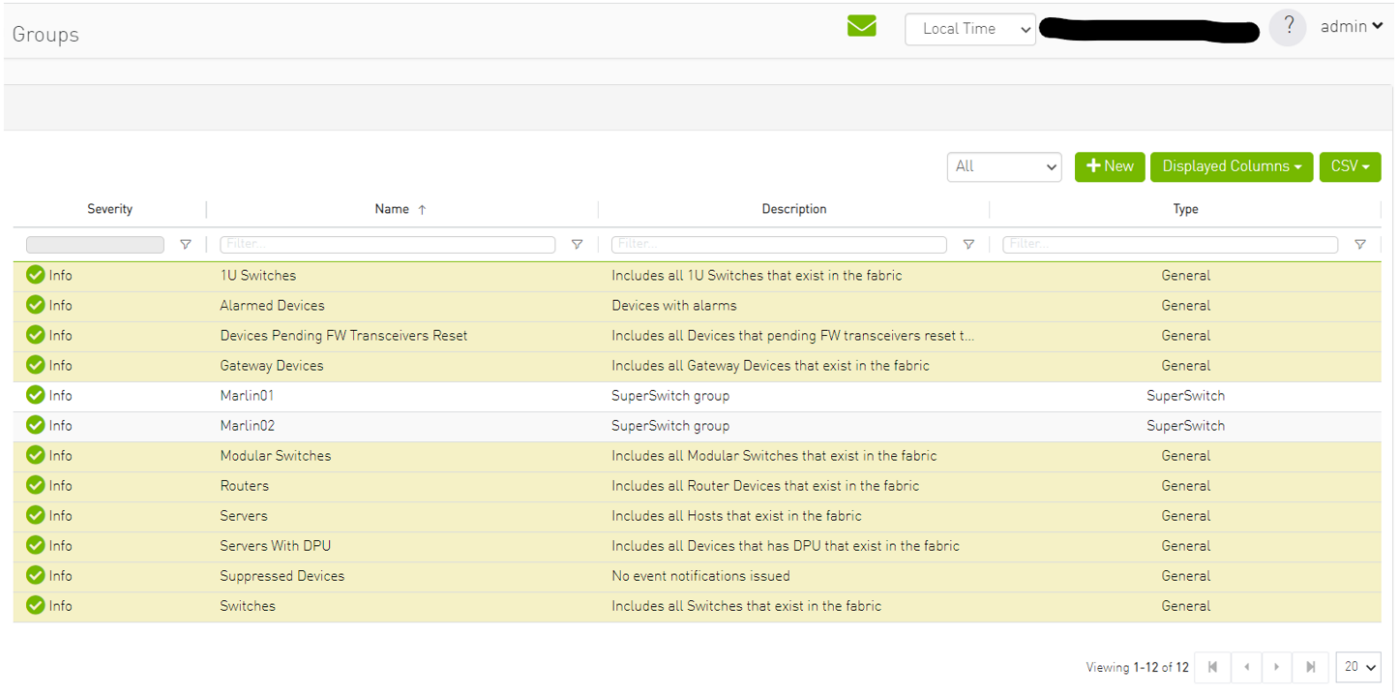
The structure of the superswitch definition should be as follows, based on the following example:

```
superswitch:
  - name: "Marlin01" # Director switch name
    description: "primary dc switch" # Free text with the customer facing description
    location: "US, NC, DC01" # Director switch location (global location, includes all racks/switches)
    racks: # Director switch Racks definitions
      #Rack definition
      - name: "rack A" # Director switch rack name
        location:
          dc-grid-row: "A" # formalized rack location in DC
          dc-grid-column: "1" # formalized
          comments: "left-most rack in the line" #Cutomer facing commnent on the rack
        leafs: # List of Director switch leafs (for the rack specified)
          - guid: "0x043f720300922a00" #required filed. Switch GUID.
            location-u: 1 # required field. Device location in rack: "U#"
            description: "MF0;gorilla-01:MQM9700/U1" # optional field.
          - guid: "0x043f720300899cc0" #required filed. Switch GUID.
            location-u: XX # required field. Device location in rack: "U#"
            description: "MF0;gorilla-01:MQM9700/U2" # optional field.
        spines: # List of Director switch spines (for the rack specified)
          - guid: "0x043f720900922a00" #required filed. Switch GUID.
            location-u: 10 # required field. Device location in rack: "U#"
            description: "MF0;gorilla-02:MQM9700/U1" # optional field.
          - guid: "0x043f720900899cc0" #required filed. Switch GUID.
            location-u: XX # required field. Device location in rack: "U#"
            description: "MF0;gorilla-02:MQM9700/U2" # optional field.
  - name: "Marlin02" # Director switch name
    description: "primary dc switch" # Free text with the customer facing description
    location: "US, NC, DC01" # Director switch location (global location, includes all racks/switches)
    racks: # Director switch Racks definitions
      #Rack definition
      - name: "rack B" # Director switch rack name
        location:
          dc-grid-row: "B" # formalized rack location in DC
          dc-grid-column: "1" # formalized
          comments: "left-most rack in the line" #Cutomer facing commnent on the rack
        leafs: # List of Director switch leafs (for the rack specified)
          - guid: "0x093f720300922a00" #required filed. Switch GUID.
            location-u: 1 # required field. Device location in rack: "U#"
            description: "MF0;gorilla-03:MQM9700/U1" # optional field.
          - guid: "0x093f720300899cc0" #required filed. Switch GUID.
```

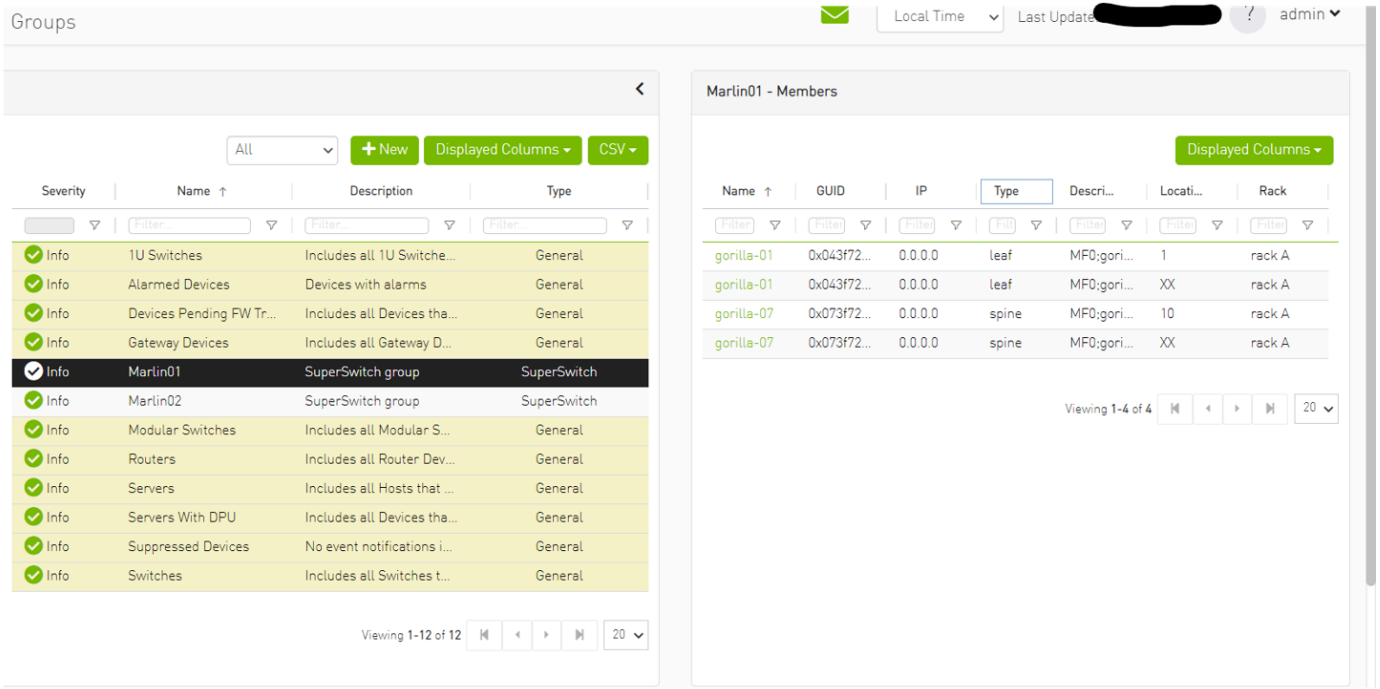
```
    location-u: XX # required field. Device location in rack: "U#"
    description: "MF0;gorilla-03:MQM9700/U2" # optional field.
spines: # List of Director switch spines (for the rack specified)
  - guid: "0x093f720900922a00" #required filed. Switch GUID.
    location-u: 10 # required field. Device location in rack: "U#"
    description: "MF0;gorilla-04:MQM9700/U1" # optional field.
  - guid: "0x093f720900899cc0" #required filed. Switch GUID.
    location-u: XX # required field. Device location in rack: "U#"
    description: "MF0;gorilla-04:MQM9700/U2" # optional field
```

UI Presentation

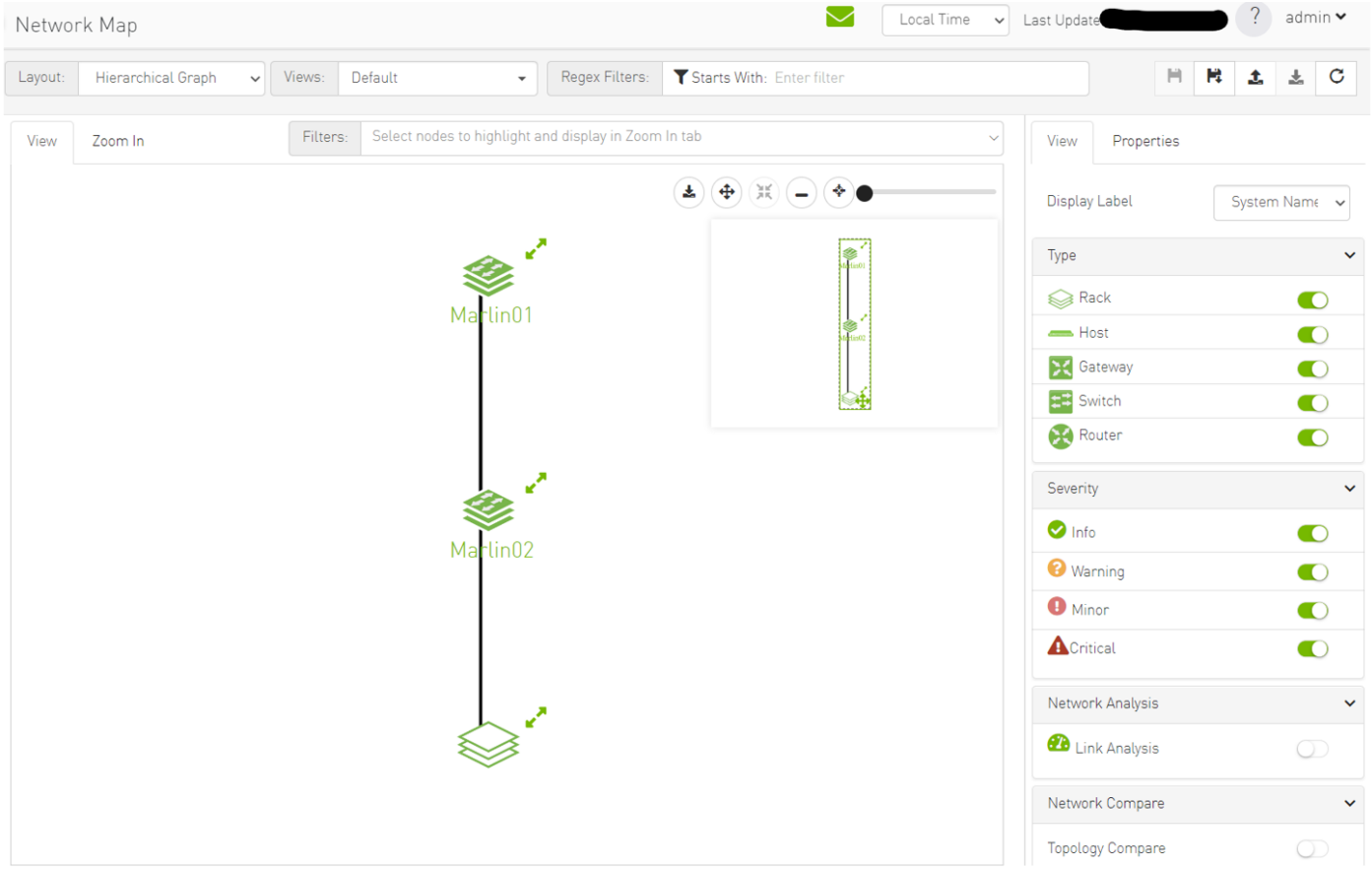
The logical grouping can be accessed under the "Groups" view, specifically listed as "SuperSwitch group" type.



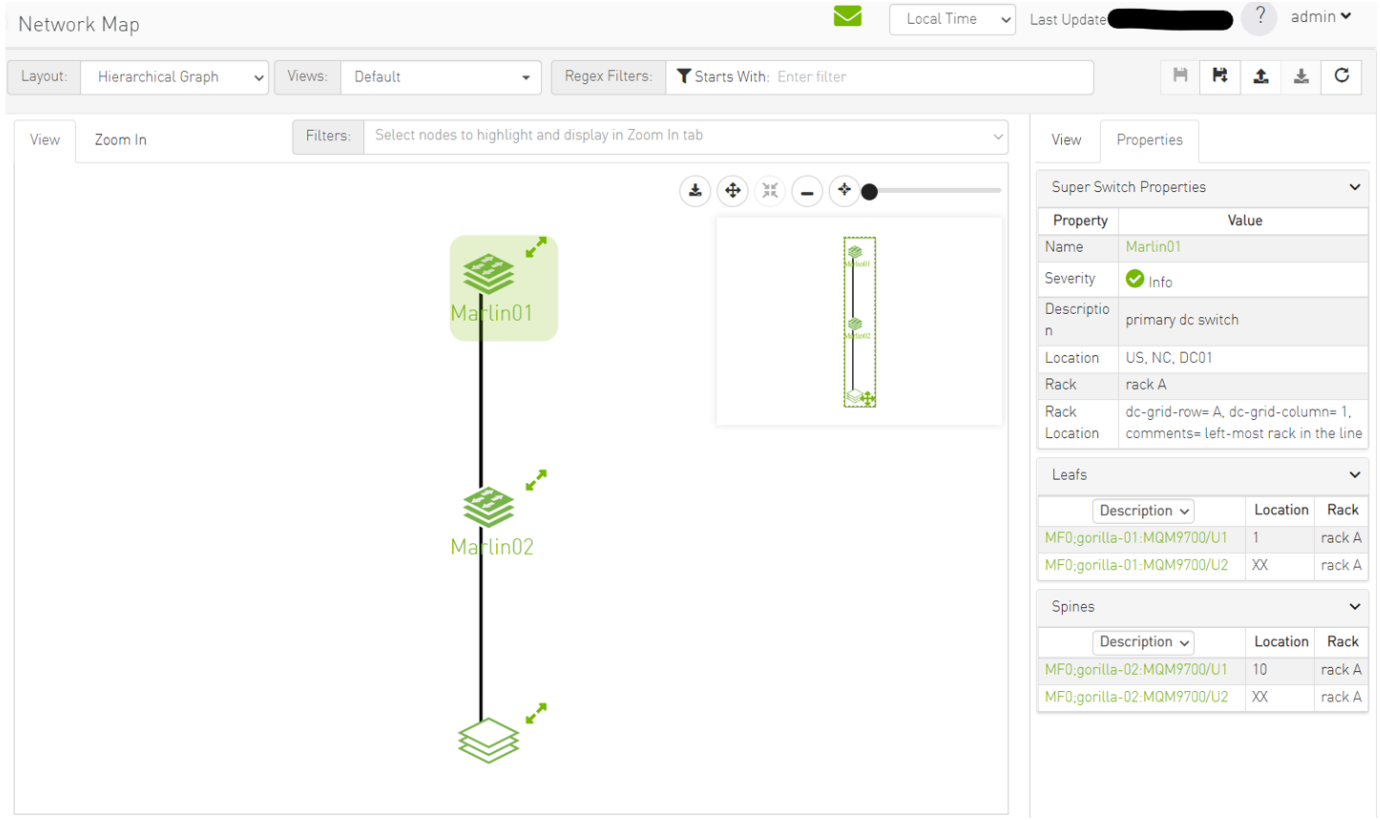
Upon selecting the group type SuperSwitch, additional columns containing information related to the SuperSwitch are added to the details view.



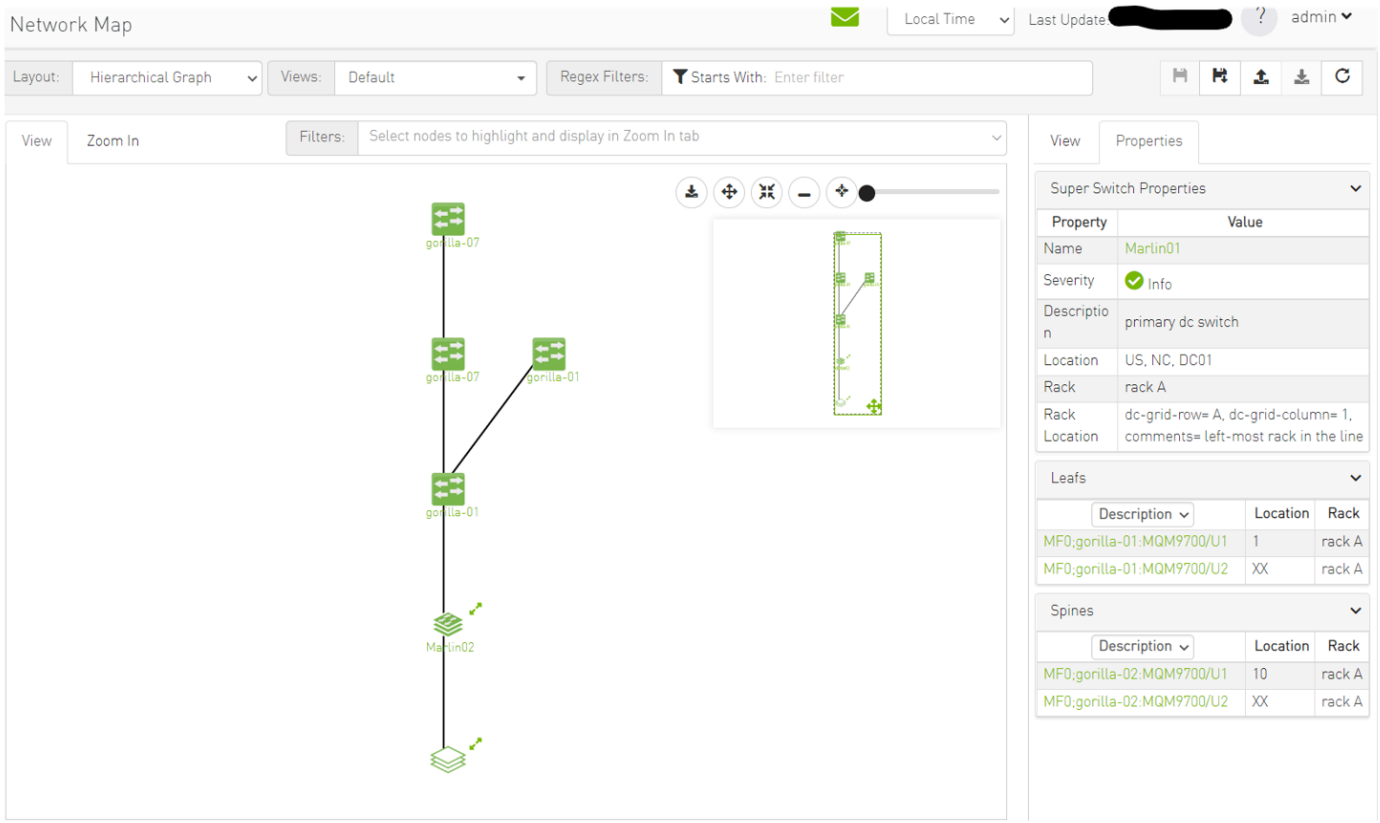
An icon for the SuperSwitch group in its collapsed view exists on the network map.

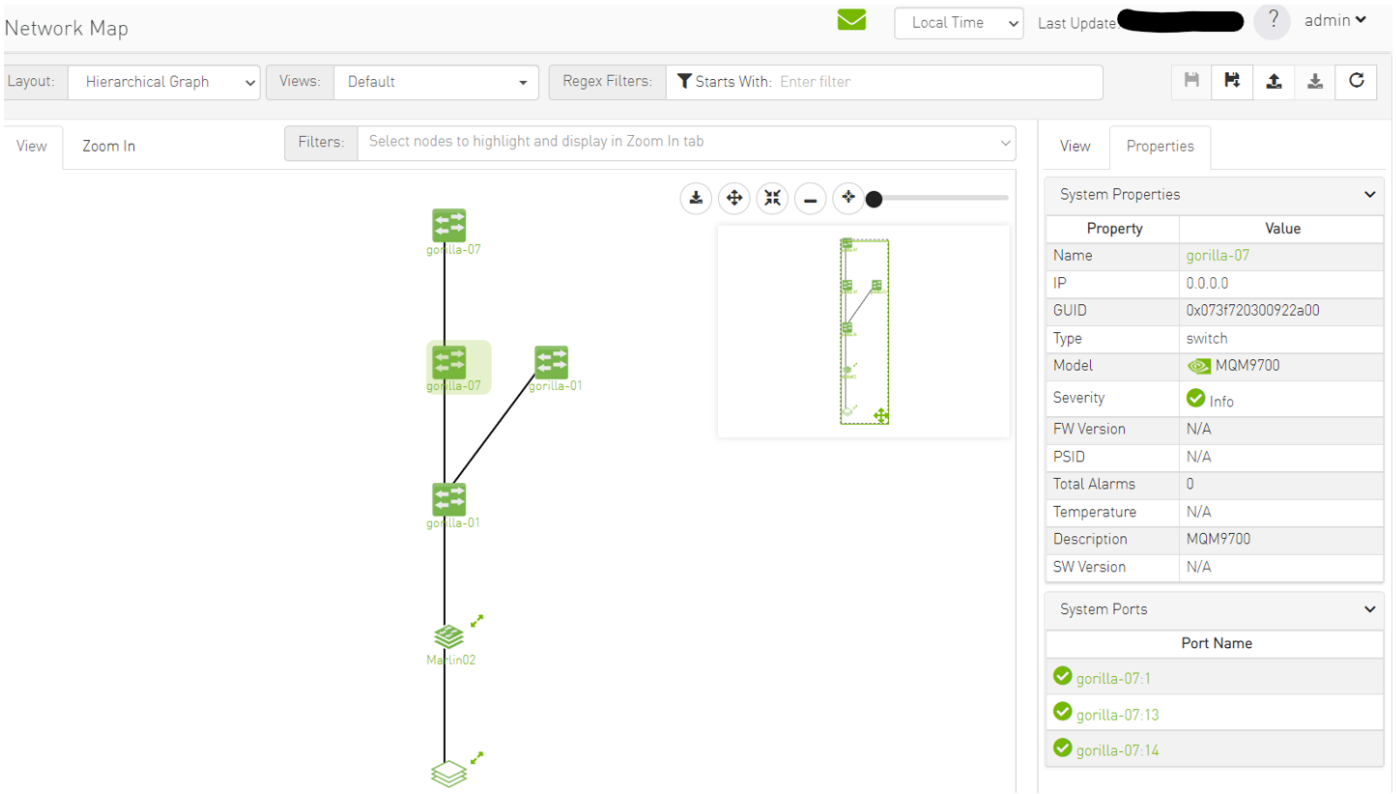


Upon selecting the SuperSwitch group, all of its properties can be viewed in the details view.

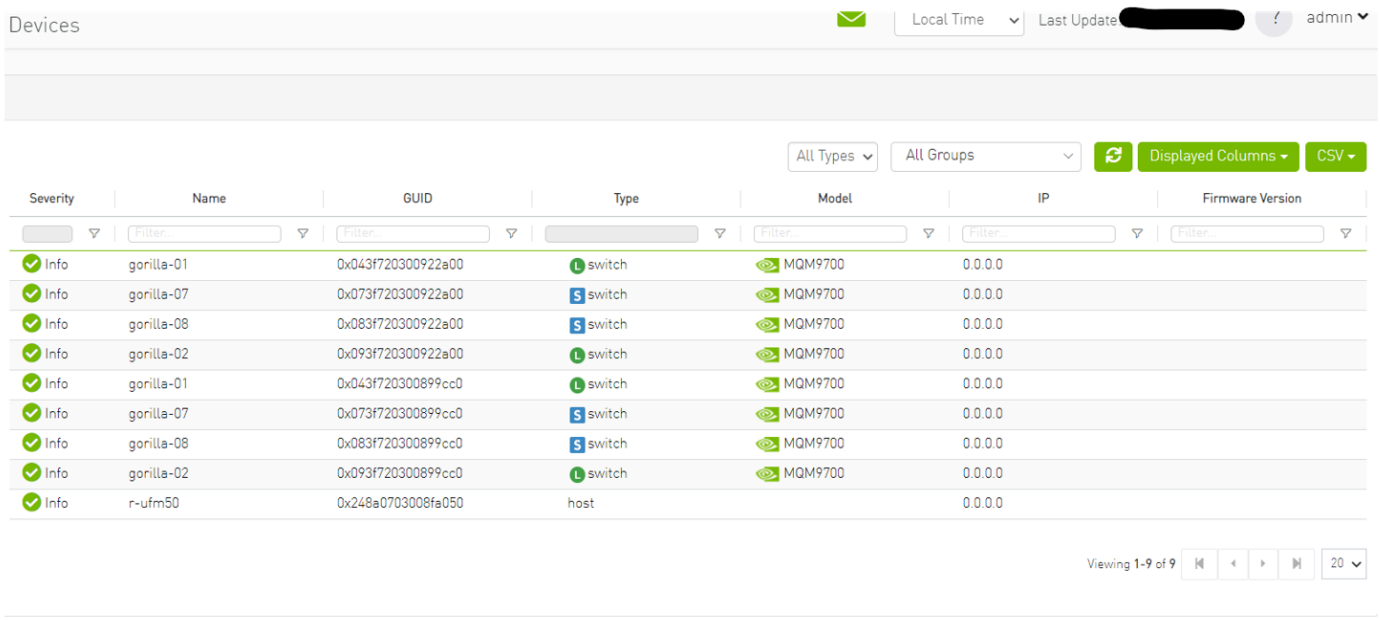


Expanding the SuperSwitch group icon displays all the switches included in the group as separate 1U switches, along with their respective properties.

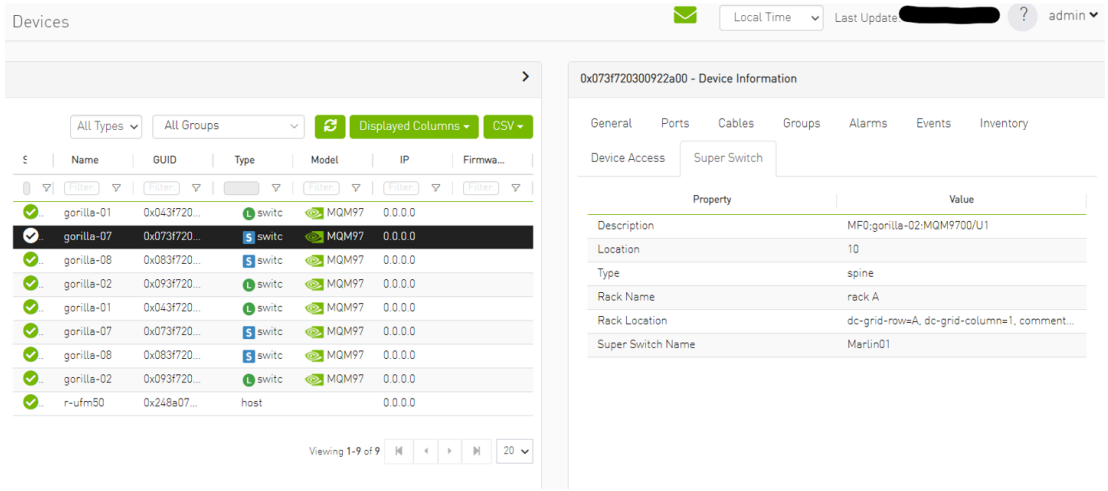

Network Map
Local Time
Last Update
admin
Layout:
Hierarchical Graph
Views:
Default
Regex Filters:
Starts With:
Enter filter
View
Zoom In
Filters:
Select nodes to highlight and display in Zoom In tab
gorilla-07
gorilla-07
gorilla-01
gorilla-01
Marlin02
View
Properties
Super Switch Properties
Property
Value
Name
Marlin01
Severity
Info
Description
primary dc switch
Location
US, NC, DC01
Rack
rack A
Rack Location
dc-grid-row= A, dc-grid-column= 1, comments= left-most rack in the line
Leafs
Description
Location
Rack
MF0;gorilla-01:MQM9700/U1
1
rack A
MF0;gorilla-01:MQM9700/U2
XX
rack A
Spines
Description
Location
Rack
MF0;gorilla-02:MQM9700/U1
10
rack A
MF0;gorilla-02:MQM9700/U2
XX
rack A



On the devices view, switches that are part of the SuperSwitch group are marked with an additional icon that indicates their role in the group. The "S" icon denotes spines, while the "L" icon denotes lines.



Selecting a switch that belongs to the SuperSwitch group in the properties view allows you to view all the switch properties related to the SuperSwitch group.



⚠ Each SuperSwitch definition can include one or more racks where each embedded rack can include multiple leafs and spines switches.

Appendix – Secondary Telemetry Fields

The following is a list of available counters which includes a variety of metrics related to timestamps, port and node information, error statistics, firmware versions, temperatures, cable details, power levels, and various other telemetry-related data.

Field Name	Description
Node_GUID	node GUID
Device_ID	PCI device ID
node_description	node description

Field Name	Description
lid	lid
Port_Number	port number
port_label	port label
Phy_Manager_State	FW Phy Manager FSM state
phy_state	physical state
logical_state	Port Logical link state
Link_speed_active	ib link active speed
Link_width_active	ib link active width
Active_FEC	Active FEC
Total_Raw_BER	Pre-FEC monitor parameters
Effective_BER	Post FEC monitor parameters
Symbol_BER	BER after all phy correction mechanism: post FEC + PLR monitor parameters
Raw_Errors_Lane_[0-3]	This counter provides information on error bits that were identified on lane X. When FEC is enabled this induction corresponds to corrected errors. In PRBS test mode, indicates the number of PRBS errors on lane X.
Effective_Errors	This counter provides information on error bits that were not corrected by FEC correction algorithm or that FEC is not active.
Symbol_Errors	This counter provides information on error bits that were not corrected by phy correction mechanisms.

Field Name	Description
Time_since_last_clear_[Min]	The time passed since the last counters clear event in msec. (physical layer statistical counters)
hist[0-15]	Hist[i] give the number of FEC blocks that had RS-FEC symbols errors of value i or range of errors
FW_Version	Node FW version
Chip_Temp	switch temperature
Link_Down	Perf.PortCounters(LinkDownedCounter)
Link_Down_IB	Total number of times the Port Training state machine has failed the link error recovery process and downed the link.
LinkErrorRecoveryCounter	Total number of times the Port Training state machine has successfully completed the link error recovery process.
PlrRcvCodes	Number of received PLR codewords
PlrRcvCodeErr	The total number of rejected codewords received
PlrRcvUncorrectableCode	The number of uncorrectable codewords received
PlrXmitCodes	Number of transmitted PLR codewords
PlrXmitRetryCodes	The total number of codewords retransmitted
PlrXmitRetryEvents	The total number of retransmitted event
PlrSyncEvents	The number of sync events
HiRetransmissionRate	Recieved bandwidth loss due to codes retransmission
PlrXmitRetryCodesWithinTSecMax	The maximum number of retransmitted events in t sec window

Field Name	Description
link_partner_description	node description of the link partner
link_partner_node_guid	node_guid of the link partner
link_partner_lid	lid of the link partner
link_partner_port_num	port number of the link partner
Cable_PN	Vendor Part Number
Cable_SN	Vendor Serial Number
cable_technology	
cable_type	Cable/module type
cable_vendor	
cable_length	
cable_identifier	
vendor_rev	Vendor revision
cable_fw_version	
rx_power_lane_[0-3]	RX measured power
tx_power_lane_[0-3]	TX measured power
Module_Voltage	Internally measured supply voltage

Field Name	Description
Module_Temperature	Module temperature
fast_link_up_status	Indicates if fast link-up was performed in the link
time_to_link_up_ext_msec	Time in msec to link up from disable until phy up state. While the phy manager did not reach phy up state the timer will return 0.
Advanced_Status_Opcode	Status opcode: PHY FW indication
Status_Message	ASCII code message
down_blame	Which receiver caused last link down
local_reason_opcode	Opcde of link down reason - local
remote_reason_opcode	Opcde of link down reason - remote
e2e_reason_opcode	see local_reason_opcode for local reason opcode for remote reason opcode: local_reason_opcode+100
PortRcvRemotePhysicalErrors	Total number of packets marked with the EBP delimiter received on the port.
PortRcvErrors	Total number of packets containing an error that were received on the port
PortXmitDiscards	Total number of outbound packets discarded by the port because the port is down or congested.
PortRcvSwitchRelayErrors	Total number of packets received on the port that were discarded because they could not be forwarded by the switch relay.
ExcessiveBufferOverrunErrors	The number of times that OverrunErrors consecutive flow control update periods occurred, each having at least one overrun error

Field Name	Description
LocalLinkIntegrityErrors	The number of times that the count of local physical errors exceeded the threshold specified by LocalPhyErrors
PortRcvConstraintErrors	Total number of packets received on the switch physical port that are discarded.
PortXmitConstraintErrors	Total number of packets not transmitted from the switch physical port.
VL15Dropped	Number of incoming VL15 packets dropped due to resource limitations (e.g., lack of buffers) in the port
PortXmitWait	The time an egress port had data to send but could not send it due to lack of credits or arbitration - in time ticks within the sample-time window
PortXmitDataExtended	Transmitted data rate per egress port in bytes passing through the port during the sample period
PortRcvDataExtended	The received data on the ingress port in bytes during the sample period
PortXmitPktsExtended	Total number of packets transmitted on the port.
PortRcvPktsExtended	Total number of packets received on the port
PortUniCastXmitPkts	Total number of unicast packets transmitted on all VLs from the port. This may include unicast packets with errors, and excludes link packets
PortUniCastRcvPkts	Total number of unicast packets, including unicast packets containing errors, and excluding link packets, received from all VLs on the port.
PortMultiCastXmitPkts	Total number of multicast packets transmitted on all VLs from the port. This may include multicast packets with errors.
PortMultiCastRcvPkts	Total number of multicast packets, including multicast packets containing errors received from all VLs on the port.
SyncHeaderErrorCounter	Count of errored block sync header on one or more lanes

Field Name	Description
PortSwLifetimeLimitDiscards	Total number of outbound packets discarded by the port because the Switch Lifetime Limit was exceeded. Applies to switches only.
PortSwHOQLifetimeLimitDiscards	Total number of outbound packets discarded by the port because the switch HOQ Lifetime Limit was exceeded. Applies to switches only.
rq_num_wrfe	Responder - number of WR flushed errors
rq_num_lle	Responder - number of local length errors
sq_num_wrfe	Requester - number of WR flushed errors
Temp_flags	Latched temperature flags of module
Vcc_flags	Latched VCC flags of module
device_hw_rev	Node HW Revision
sw_revision	switch revision
sw_serial_number	switch serial number

Document Revision History

Release	Date	Description
4.14.0	Nov 5, 2023	Updated: • UFM Logs Tab - Added log occurrences display Added • Events History • Device Status Events • Link Status Events • GNMI-Telemetry Plugin • In Secondary Telemetry, added instructions on Exposing Switch Aggregation Nodes Telemetry and Stopping Telemetry Endpoint Using CLI Command
4.13.1	Aug 31, 2023	Updated: • Changes and New Features • Bug Fixes in This Release
4.13.0	Aug 10, 2023	Updated: • Changes and New Features • Bug Fixes in This Release • Known Issues in This Release • Plugin Management • Secondary Telemetry • PDR Deterministic Plugin - Updated step 3 in "Deployment". • rest-rdma Plugin • NDT Plugin • Autonomous Link Maintenance (ALM) Plugin • Appendix - Supported Port Counters and Events - Added alarm ID # 1602 and 1603 and status column for all alarm IDs. Added: • Health Policy Management

Release	Date	Description

4.12.1	May 18, 2023	Updated
		• Bug Fixes in This Release
		• Known Issues in This Release
		Added:
		• NDT Format Merger
		• Subnet Merger UI
		• Added the following Plugins:
		  • Autonomous Link Maintenance (ALM) Plugin
		  • PDR Deterministic Plugin

4.12.0	Feb 19, 2023	Updated
		• Changes and New Features
		• Bug Fixes in This Release
		• Known Issues in This Release
		• Email - Added time zone preference
		• NDT Plugin
		• UFM Telemetry FluentD Streaming (TFS) Plugin - Updated REST APIs
		• Appendix - Supported Port Counters and Events
		Added:
		• Multi-Subnet UFM
		• Enable Network Fast Recovery
		• Added the following Plugins:
		  • UFM Bright Cluster Integration Plugin
		  • UFM Cyber-AI Plugin
		  • DTS Plugin
		  • Sysinfo Plugin
		  • SNMP Plugin
		  • Packet Mirroring Collector (PMC) Plugin

4.11.1	Feb 19, 2023	Updated • Changes and New Features • Bug Fixes in This Release • Known Issues in This Release
	Mar 16, 2023	Updated Changes and New Features - Added MFT package integration details

Release	Date	Description
4.11.0	Feb 2, 2023	Updated: • Changes and New Features • Bug Fixes in This Release • Known Issues in This Release • Configuring Partial Switch ASIC Failure Events • Updated example in Multi-port SM • Appendix – Used Ports Added: • Added a note under Ports Window • Added a note under Unhealthy Ports Window • Delegate Authentication to a Proxy Removed: • UFM Logical Elements tab from the Web UI
4.10.1	Dec 1, 2022	Updated the following sections: • Changes and New Features to include the upgrade of NVIDIA SHARP SW version • Installation Notes • Bug Fixes in This Release
	Dec 6, 2022	Updated In-Band Firmware Upgrade
	Dec 19, 2022	Updated Changes and New Features

EULA, Legal Notices and 3rd Party Licenses

Legal Notice

Third-Party Licenses

License Agreement

This license is a legal agreement (“Agreement”) between you and Mellanox Technologies, Ltd. ("NVIDIA") and governs the use of the NVIDIA UFM software and materials provided hereunder (“SOFTWARE”). If you are entering into this Agreement on behalf of a company or other legal entity, you represent that you have the legal authority to bind the entity to this Agreement, in which case “you” will mean the entity you represent.

You agree to use the SOFTWARE only for purposes that are permitted by (a) this license, and (b) any applicable law, regulation, or generally accepted practices or guidelines in the relevant jurisdictions.

1. License. Subject to the terms and conditions of this Agreement and payment of applicable subscription fee, NVIDIA MELLANOX grants you a personal, non-exclusive, non-sublicensable (except as provided in this Agreement), non-transferable, non-commercial license to install and use the Software for your internal business purposes for configuring, operating, and managing your InfiniBand network and not for further distribution.
2. Authorized Users. You may allow access and use of the Software to: (i) employees and contractors of your entity provided that the access and use of the Software is made from your secure network to perform work on your behalf and (ii) If you are an academic institution you may allow users enrolled or employed by the academic institution to access and use the Software from your secure network (“Authorized Users”). You hereby undertake to be responsible and liable for any non-compliance with the terms of this Agreement by your Authorized Users. You further agree to immediately resolve any non-compliance by your Authorized Users of which you become aware and endeavor take necessary steps to prevent any new occurrences.
3. Limitations Your license to use the SOFTWARE is restricted as follows:
   3.1 The SOFTWARE is licensed for your use in systems with the registered NVIDIA Host Channel Adapter (HCA) Products or related adapter products.
   3.2 Each copy of the SOFTWARE shall be limited to the number of HCAs indicated in the applicable purchase order.
   3.3 You may use software back-up utilities to make one back-up copy of the Software Product. You may use the back-up copy solely for archival purposes

3.4 You may not use the SOFTWARE in conjunction with a number of managed nodes or managed devices which is beyond the allowable limit or copy the SOFTWARE on additional hardware. You shall not use any features which are not included in the scope of this Agreement as described in the accompanying documentation.
3.5 You may not reverse engineer, decompile or disassemble, or remove copyright or other proprietary notices from any portion of the SOFTWARE or copies of the SOFTWARE.
3.6 You may not disclose the results of benchmarking, competitive analysis, regression, or performance data relating to the SOFTWARE without the prior written permission from NVIDIA Mellanox.
3.7 Except as expressly provided in this license, you may not copy, sell, rent, sublicense, transfer, distribute, modify, or create derivative works of any portion of the SOFTWARE. For clarity, unless, you have an agreement with NVIDIA Mellanox for this purpose you may not distribute or sublicense the SOFTWARE as a stand-alone product.
3.8 You may not bypass, disable, or circumvent any technical limitation, encryption, security, digital rights management, or authentication mechanism in the SOFTWARE.
3.9 You may not use the Software in any manner that would cause it to become subject to an open source software license. As examples, licenses that require as a condition of use, modification, and/or distribution that the Software be: (i) disclosed or distributed in source code form; (ii) licensed for the purpose of making derivative works; or (iii) redistributable at no charge.
3.10 Unless you have an agreement with NVIDIA Mellanox for this purpose, you may not use the Software with any system or application where the use or failure of the system or application can reasonably be expected to threaten or result in personal injury, death, or catastrophic loss. Examples include use in avionics, navigation, military, medical, life support or other life critical applications. NVIDIA Mellanox does not design, test, or manufacture the Software for these critical uses and NVIDIA Mellanox shall not be liable to you or any third party, in whole or in part, for any claims or damages arising from such uses.
3.11 You agree to defend, indemnify and hold harmless NVIDIA Mellanox and its affiliates, and their respective employees, contractors, agents, officers and directors, from and against any and all claims, damages, obligations, losses, liabilities, costs or debt, fines, restitutions and expenses (including but not limited to attorney's fees and costs incident to establishing the right of indemnification) arising out of or related to your use of the Software outside of the scope of this license, or not in compliance with its terms.

4. Updates. NVIDIA Mellanox may, at its option, make available patches, workarounds, or other updates to this Software. Unless the updates are provided with their separate governing terms, they are deemed part of the Software licensed to you as provided in this license. You agree that the form and content of the Software that NVIDIA Mellanox provides may change without prior notice to you. While NVIDIA Mellanox generally maintains compatibility between versions, NVIDIA Mellanox may in some cases make changes that introduce incompatibilities in future versions of the SOFTWARE.
5. Pre-Release Versions. Software versions identified as alpha, beta, preview, early access or otherwise as pre-release may not be fully functional, may contain errors or design flaws, and may have reduced or different security, privacy, availability, and reliability standards relative to commercial versions of NVIDIA Mellanox software and materials. You may use a pre-release Software version at your own risk, understanding that these versions are not intended for use in production or business-critical systems. NVIDIA Mellanox may choose not to make available a commercial version of any

pre-release Software. NVIDIA Mellanox may also choose to abandon development and terminate the availability of a pre-release Software at any time without liability.

6. Third-Party Components. The Software may include third-party components with separate legal notices or terms as may be described in proprietary notices accompanying the Software or as provided in an Exhibit to this Agreement. If and to the extent there is a conflict between the terms in this license and the third-party license terms, the third-party terms control only to the extent necessary to resolve the conflict. For details regarding the third party components, please review Exhibit A.
7. OWNERSHIP
   7.1 NVIDIA Mellanox or its licensors reserves all rights, title, and interest in and to the Software not expressly granted to you under this license NVIDIA Mellanox and its suppliers hold all rights, title, and interest in and to the Software, including their respective intellectual property rights. The Software is copyrighted and protected by the laws of the United States and other countries, and international treaty provisions.
   7.2 Subject to the rights of NVIDIA Mellanox and its suppliers in the Software, you hold all rights, title, and interest in and to your applications and your derivative works of the sample source code delivered in the Software including their respective intellectual property rights.
8. You may, but are not obligated to, provide to NVIDIA Mellanox Feedback. “Feedback” means suggestions, fixes, modifications, feature requests or other feedback regarding the Software. Feedback, even if designated as confidential by you, shall not create any confidentiality obligation for NVIDIA Mellanox. NVIDIA Mellanox and its designees have a perpetual, non-exclusive, worldwide, irrevocable license to use, reproduce, publicly display, modify, create derivative works of, license, sublicense, and otherwise distribute and exploit Feedback as NVIDIA Mellanox sees fit without payment and without obligation or restriction of any kind on account of intellectual property rights or otherwise.
9. No Warranties. THE SOFTWARE IS PROVIDED AS-IS. TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW NVIDIA MELLANOX AND ITS AFFILIATES EXPRESSLY DISCLAIM ALL WARRANTIES OF ANY KIND OR NATURE, WHETHER EXPRESS, IMPLIED OR STATUTORY, INCLUDING, BUT NOT LIMITED TO, WARRANTIES OF MERCHANTABILITY, NON-INFRINGEMENT, OR FITNESS FOR A PARTICULAR PURPOSE. NVIDIA MELLANOX DOES NOT WARRANT THAT THE SOFTWARE WILL MEET YOUR REQUIREMENTS OR THAT THE OPERATION THEREOF WILL BE UNINTERRUPTED OR ERROR-FREE, OR THAT ALL ERRORS WILL BE CORRECTED.
10. Limitations of Liability. TO THE MAXIMUM EXTENT PERMITTED BY APPLICABLE LAW NVIDIA MELLANOX AND ITS AFFILIATES SHALL NOT BE LIABLE FOR ANY SPECIAL, INCIDENTAL, PUNITIVE OR CONSEQUENTIAL DAMAGES, OR FOR ANY LOST PROFITS, PROJECT DELAYS, LOSS OF USE, LOSS OF DATA OR LOSS OF GOODWILL, OR THE COSTS OF PROCURING SUBSTITUTE PRODUCTS, ARISING OUT OF OR IN CONNECTION WITH THIS LICENSE OR THE USE OR PERFORMANCE OF THE SOFTWARE, WHETHER SUCH LIABILITY ARISES FROM ANY CLAIM BASED UPON BREACH OF CONTRACT, BREACH OF WARRANTY, TORT (INCLUDING NEGLIGENCE), PRODUCT LIABILITY OR ANY OTHER CAUSE OF ACTION OR THEORY OF LIABILITY, EVEN IF NVIDIA MELLANOX HAS PREVIOUSLY BEEN ADVISED OF, OR COULD REASONABLY HAVE FORESEEN, THE POSSIBILITY OF SUCH DAMAGES. IN NO EVENT WILL NVIDIA MELLANOX AND ITS AFFILIATES TOTAL CUMULATIVE LIABILITY UNDER OR ARISING OUT OF THIS LICENSE EXCEED US$10.00. THE NATURE OF THE LIABILITY OR THE NUMBER OF CLAIMS OR SUITS SHALL NOT ENLARGE OR EXTEND THIS LIMIT.
11. T Your rights under this license will terminate automatically without notice from NVIDIA Mellanox (a) upon expiration of your subscription, (b) if you fail to comply with any term and condition of this license including non-payment of applicable fees, or (c) if you commence or participate in any legal proceeding against NVIDIA Mellanox with respect to the Software. NVIDIA Mellanox may terminate this license with advance written notice to you, if NVIDIA Mellanox decides to no longer provide the Software in a country or, in NVIDIA Mellanox’s sole discretion, the continued use of it is no

longer commercially viable. Upon any termination of this license, you agree to promptly discontinue use of the Software and destroy all copies in your possession or control. All provisions of this license will survive termination, except for the license granted to you.

12. Product Support. Product support for the Software Product is provided by NVIDIA Mellanox or its authorized agents under the applicable subscription license, in accordance with NVIDIA Mellanox's standard support and maintenance terms and conditions. For product support, please refer to NVIDIA Mellanox support number provided in the documentation.
13. Applicable Law. This license will be governed in all respects by the laws of the United States and of the State of Delaware, without regard to the conflicts of laws principles. The United Nations Convention on Contracts for the International Sale of Goods is specifically disclaimed. You agree to all terms of this license in the English language. The state or federal courts residing in Santa Clara County, California shall have exclusive jurisdiction over any dispute or claim arising out of this license. Notwithstanding this, you agree that NVIDIA Mellanox shall still be allowed to apply for injunctive remedies or urgent legal relief in any jurisdiction.
14. No Assignment. This license and your rights and obligations thereunder may not be assigned by you by any means or operation of law without NVIDIA Mellanox's permission. Any attempted assignment not approved by NVIDIA MELLANOX in writing shall be void and of no effect. NVIDIA Mellanox may assign, delegate, or transfer this license and its rights and obligations, and if to a non-affiliate you will be notified.
15. E The Software is subject to United States export laws and regulations. You agree to comply with all applicable U.S. and international export laws, including the Export Administration Regulations (EAR) administered by the U.S. Department of Commerce and economic sanctions administered by the U.S. Department of Treasury's Office of Foreign Assets Control (OFAC). These laws include restrictions on destinations, end-users and end-use. By accepting this license, you confirm that you are not currently residing in a country or region currently embargoed by the U.S. and that you are not otherwise prohibited from receiving the Software.
16. Government Use. The Software is, and shall be treated as being, "Commercial Items" as that term is defined at 48 CFR § 2.101, consisting of "commercial computer software" and "commercial computer software documentation", respectively, as such terms are used in, respectively, 48 CFR § 12.212 and 48 CFR §§ 227.7202 & 252.227-7014(a)(1). Use, duplication or disclosure by the U.S. Government or a U.S. Government subcontractor is subject to the restrictions in this license pursuant to 48 CFR § 12.212 or 48 CFR § 227.7202. In no event shall the US Government user acquire rights in the Software beyond those specified in 48 C.F.R. 52.227-19(b)(1)-(2).
17. Please direct your legal notices or other correspondence to NVIDIA Corporation, 2788 San Tomas Expressway, Santa Clara, CA, 95051 United States of America, Attention: Legal Department and to: NBU-Legal_Notices@exchange.nvidia.com
18. Entire Agreement. This license is the final, complete, and exclusive agreement between the parties relating to the subject matter of this license and supersedes all prior or contemporaneous understandings and agreements relating to this subject matter, whether oral or written. If any court of competent jurisdiction determines that any provision of this license is illegal, invalid, or unenforceable, the remaining provisions will remain in full force and effect. Any amendment or waiver under this license shall be in writing and signed by representatives of both parties.

(v APR. 28, 2022)

Exhibit A

SOFTWARE includes the following open source/ freeware that are subject to specific license conditions listed in the table below, which may be updated from time to time by NVIDIA Mellanox or the Open Source provider. The below table is current as of December 2021. To obtain source code for software provided under licenses that require redistribution of source code, including the GNU General Public License or for update queries contact: http://www.mellanox.com/page/gnu_code_request .This offer is valid for a period of three (3) years from the date of the distribution of this product by NVIDIA Mellanox.

Component name	Version	Home Page	License
@candlefw/wick	0.8.12	https://github.com/galactrax/cfw-wick#readme	MIT License
ABSender	master-20121122	https://github.com/100Continue/ABSender	Apache License 2.0
APBS	apbs-0.3.1	https://sourceforge.net/projects/apbs	GNU General Public License v2.0 or later
Amazon Kindle Source Code	6.2	http://www.amazon.com/gp/help/customer/display.html?nodeId=200203720	Apache License 2.0
Amiga Research OS	20120217	https://aros.sourceforge.io/license.html	Aros Public License V 1.1
Apache ActiveMQ	2.2.2	http://activemq.apache.org/	Apache License 2.0
Apache HTTP Server	1.3.7, 1.3.8	http://httpd.apache.org/	Apache License 1.0
Apache HTTP Server	2, 2.0.11, 2.0.23, 2.0.25, 2.0.26, 2.0.30, 2.0.33, 2.0.35, 2.0.36,2.0.38, 2.0.39, 2.0.40, 2.0.41, 2.0.43, 2.1.0	http://httpd.apache.org/	Apache License 1.1

Component name	Version	Home Page	License
Apache HTTP Server	2.0.59, 2.1.1, 2.1.10, 2.1.2, 2.1.3, 2.1.4, 2.1.5, 2.1.6, 2.1.7, 2.1.8, 2.1.9, 2.2.1, 2.2.2 2.2.12, 2.2.13, 2.2.14, 2.2.15, 2.2.16, 2.2.17, 2.2.22, 2.2.26, 2.2.3, 2.2.4, 2.2.5, 2.2.6, 2.2.7, 2.2.9, 2.3.0, 2.3.1, 2.3.4	http://httpd.apache.org/	Apache License 2.0
Apache HTTP Server	STRIKER_2_1_0_RC1	http://httpd.apache.org/	Apache License 2.0
Apache Portable Runtime	0.9.13, 0.9.15, 1.2.0, 1.2.10, 1.2.11, 1.2.12, 1.2.7, 1.2.8, 1.2.9, 1.3.0, 1.3.1, 1.3.10, 1.3.12, 1.3.2, 1.3.3, 1.3.4, 1.3.5, 1.3.7, 1.3.8, 1.3.9, 1.4.7, 1.5.1, 1.5.2; APR_1_0_RC2; JCW_0_9_5_PRE1	http://apr.apache.org/	Apache License 2.0
Apache Portable Runtime	0.9.4 APACHE_2_0_37 APACHE_2_0_40 APACHE_2_0_44 APACHE_2_0_48	http://apr.apache.org/	Apache License 1.1
Apache Portable Runtime	APU_1_0_RC1	http://apr.apache.org/	(MIT License AND RSA Message-Digest License AND Apache License 2.0 AND Beerware License AND RSA

Component name	Version	Home Page	License
			MD4 or MD5 Message-Digest Algorithm License AND Christian Michelsen Research License AND Apache License 1.1)
Apache Tomcat	1.1.0, 6.0.24	http://tomcat.apache.org/	Apache License 2.0
BIND9 (Berkeley Internet Name Domain)	9.9.11	https://www.isc.org/wordpress/software/bind/	Mozilla Public License 2.0
Berkeley DB	4.5.20	http://www.oracle.com/technology/products/berkeley-db/db/index.html	BSD 3-clause "New" or "Revised" License
Chromium (Google Chrome)	32.0.1700.102	http://code.google.com/chromium/	BSD 3-clause "New" or "Revised" License
Cinder	v0.8.0	http://libcinder.org	BSD 3-clause "New" or "Revised" License
Clonezilla	1.2.10	http://clonezilla.org/	GNU General Public License v3.0 or later
Cron	3.0pl1	https://alioth.debian.org/projects/pkg-cron/	Cron License
CyanogenMod - android_external_busybox	cm-10.1-M1, cm-10.1-M2	https://github.com/CyanogenMod/android_external_busybox/blob/cm-12.0/LICENSE	GNU General Public License v2.0 or later
D-Bus	1.2.6	http://www.freedesktop.org/wiki/Software/dbus	Academic Free License v2.1
DHCP (ISC)	4.3.6	http://www.isc.org/downloads/dhcp/	ISC License
Darik's Boot and Nuke	dban-2.0.0	http://sourceforge.net/projects/dban	(GNU Lesser General Public License v3.0 or later AND GNU

Component name	Version	Home Page	License
			General Public License v3.0 or later)
Debian Games	11.04.1+repack	http://wiki.debian.org/Games	BSD 3-clause "New" or "Revised" License
FLAC - Free Lossless Audio Codec	flac-1.1.1-beta1-src	http://flac.sourceforge.net	BSD 3-clause "New" or "Revised" License
FarGroup/FarManager	builds/3.0.2890	https://github.com/FarGroup/FarManager/blob/master/LICENSE	BSD 3-clause "New" or "Revised" License
FreeBSD	5.5, 6, 9.0-BETA1, release/11.2.0,12.2, 2.2.0, 2.2.6, 5.0.0cvs	https://github.com/trueos/trueos	BSD 2-clause "Simplified" License
FreeBSD	bsd_44_lite	https://github.com/trueos/trueos	BSD 4-clause "Original" or "Old" License
FreeBSD Ports	RELEASE_4_5_0 RELEASE_4_6_0	https://www.freebsd.org/ports/	BSD 2-clause FreeBSD License
FreeNAS	0.7	https://www.freenas.org/	BSD 3-clause "New" or "Revised" License
GD	2.0.1beta, 2.0.32, 2.0.33, 2.0.34RC1, 2.0.35, 2.0.35RC5	http://www.libgd.org	GD License
GD	2.0.36_rc1	http://www.libgd.org	(X11 License OR MIT License)
GLib	1.2.3, 2.14.6, 2.19.5	http://library.gnome.org/devel/glib/	Apache License 2.0
GNU Compiler Collection	4.7.0	http://gcc.gnu.org/	(GD License OR Unknown License)

Component name	Version	Home Page	License
GNU Libtool	1.4.1	http://www.gnu.org/software/libtool/	BSD 3-clause "New" or "Revised" License
GNU Parted	1.8.1, 2.4	http://www.gnu.org/software/parted	GNU General Public License v2.0 or later
GNU Parted	2.4	http://www.gnu.org/software/parted	GNU General Public License v3.0 or later
Gentoo Linux	release_1_3_17	https://www.gentoo.org/	GNU General Public License v2.0 or later
Heimdal Kerberos	heimdal-0.0n	http://www.h5l.org/	BSD 3-clause "New" or "Revised" License
HipHop Virtual Machine for PHP	HHVM-3.1.0	https://github.com/facebook/hhvm	(PHP License v3.01 AND Zend License v2.0)
Kablink	1.1 Alpha1	https://www.kablink.org/	Apache License 2.0
Less	374	http://www.greenwoodsoftware.com/less/	BSD 2-clause "Simplified" License
Less	429	http://www.greenwoodsoftware.com/less/	GNU General Public License v2.0 or later OR Less License
LineageOS	cm-10.1.0-RC1	https://lineageos.org/	(FSF Unlimited License AND BSD 3-clause "New" or "Revised" License)
Linux Test Project	2004	https://github.com/linux-test-project/ltp	GNU General Public License v2.0 or later

Component name	Version	Home Page	License
Linux-Pam	0.59, 0.72, 0.74, 0.76, 0.99.1.0, 0.99.2.0, 0.99.4.0, 0.99.5.0, 0.99.6.1, 0.99.6.2, 1.0.0	http://www.linux-pam.org	BSD 3-clause "New" or "Revised" License
Linux-Pam	1.0.1	http://www.linux-pam.org	(X11 License AND FSF Unlimited License)
MapServer	rel-1-0-0	http://mapserver.org	(X11 License AND MIT License)
Merruk-Technology	2.0-20121113	http://www.merruk.ma	GNU General Public License v2.0 only
MinGW - Minimalist GNU for Windows	binutils-2.20	http://mingw.sourceforge.net/	Public Domain
MythTV	v0.13	http://www.mythtv.org	GNU General Public License v2.0 or later
NFS	1.0.6	http://linux-nfs.org/	GNU General Public License v2.0 or later
Net-SNMP	5.0.9, 5.4.2.1, 5.5.2.pre1, 5.7.3, END-UCD-SNMP. Ext-5-3-cvs20050331, JBPN-CBL-1, 5.0.11.1, 5.2.2	http://www.net-snmp.org	(CMU License AND BSD 3-clause "New" or "Revised" License)
Net-SNMP	5.1.2, Ext-5-0, Ext-5-0-2, Ext-5-0-4, Ext-5-4-1-1, V4-2-patches-merge2	http://www.net-snmp.org	(Diffstat License OR BSD 3-clause "New" or "Revised" License)
Net-SNMP	Ext-5-0, Ext-5-0-4	http://www.net-snmp.org	(Diffstat License AND BSD 3-clause "New" or "Revised" License AND Christian Michelsen Research License)

Component name	Version	Home Page	License
Net-SNMP	Ext-5-4-1-1	http://www.net-snmp.org	(Diffstat License AND BSD 3-clause "New" or "Revised" License AND Christian Michelsen Research License AND Bzip2 License)
Net-SNMP	V4-2-patches-merge2	http://www.net-snmp.org	Diffstat License AND Christian Michelsen Research License)
Net-SNMP	5.2.4 source code, 5.2.5 pre-releases, 5.3.1, 5.3.2 pre-releases, 5.4.2 pre-releases, 5.5, Ext-4-0-pre5, Ext-4-1-pre1, Ext-5-0-2-pre1,Ext-5-0-7-pre1, Ext-5-0-8-pre1, Ext-5-2-2rc6, Ext-5-2-pre2, Ext-5-2-pre3, Ext-5-3-pre1, Ext-5-3-pre3, Ext-5-3-pre4, Ext-5-4-1-pre1, Ext-5-4-1-pre3, Ext-5-4-pre1, Ext-5-4-pre1, Ext-5-4-pre4, Ext-5-5-pre1, Ext-5-5-pre2, Ext-5-5-pre3, Ext-5-5-rc1, Ext-5-5-rc3, 5.3.0.1. 5.8.1.pre1, 5.8.1.pre2	http://www.net-snmp.org	BSD 3-clause "New" or "Revised" License
NetBSD	1.1, 1.5, 2	http://www.netbsd.org	BSD 3-clause "New" or "Revised" License
OpenFabrics Enterprise Distribution - OFED	1.2, 1.5, 3.3.2018	https://www.openfabrics.org/downloads/rdmacm/	BSD 2-clause "Simplified" License
OpenFabrics Enterprise Distribution - OFED	3.1.8	https://www.openfabrics.org/downloads/rdmacm/	BSD 3-clause "New" or "Revised" License
OpenLDAP	2.4.44	http://www.openldap.org/	Open LDAP Public License v2.8

Component name	Version	Home Page	License
OpenSSH	5.3p1, 7.4p1,7.7, 7.7p1, 7.8, 7.8p1, 7.9, 7.9p1, 8.0p1, pre-reorder	http://www.openssh.com/	BSD 3-clause "New" or "Revised" License
OpenSSH	7.2p2, 7.6p1	http://www.openssh.com/	X11 License
OpenWrt	12.09, 14.07	http://openwrt.org/	GNU General Public License v2.0 or later
PCRE	7.1, 7.4, 7.6	http://www.pcre.org/	PCRE License
PCRE	4, 7.6, 7.7, 7.8	http://www.pcre.org/	BSD 3-clause "New" or "Revised" License
PHP	MERGE_FROM_NEW_LOOK_2001_TAG_1	http://svn.php.net	BSD 2-clause "Simplified" License
PortableApps.com	WinMerge 2.10.0 , 2.6.12Source	http://portableapps.com/	Apache License 2.0
Python programming language	v2.4a2	https://www.python.org	Python Software Foundation License 2.0
Qualcomm Kernel Tree for MSM/ QSD family and Android 4.4	ath-201808291719	https://www.codeaurora.org/projects/all-active-projects/linux-msm	ISC License
TACACS+ client library and PAM module	1.2.10, 1.2.9	https://sourceforge.net/projects/tacplus	BSD 3-clause "New" or "Revised" License
Stephane-D/SGDK	V1.62	https://github.com/Stephane-D/SGDK/blob/master/license.txt	MIT License
TACACS+ client library and PAM module	1.3.2	https://sourceforge.net/projects/tacplus	GNU General Public License v2.0 or later
Tarifa	Tarifa019.tar	http://sourceforge.net/projects/tarifa	GNU General Public License v2.0 or later

Component name	Version	Home Page	License
Tcl/Tk	8.1.1	http://www.tcl.tk/	TCL/TK License
Tecla Library	1.2.3, 1.4.0, 1.4.1, 1.5.0, 1.6.0, 1.6.2	http://www.astro.caltech.edu/~mcs/tecla/index.html	MIT License
The GWARE Project	2.10.2	http://sourceforge.net/projects/gware	GNU Lesser General Public License v2.1 or later
TizenRT	1.1_Public_Release	https://github.com/Samsung/TizenRT	Apache License 2.0
UC-7402.7408.7410.7420-LX Plus Source	20100210	http://www.moxa.com/product/UC-7408.htm	GNU General Public License v2.0 only
WinMerge	2.11.1.7	https://winmerge.org/	Apache License 2.0
XAMPP	1.4.5, 1.6.4	https://www.apachefriends.org/index.html	BSD 3-clause "New" or "Revised" License
XAMPP	1.6.4	https://www.apachefriends.org/index.html	GNU General Public License v2.0 or later
XQilla	1.1.0	http://xqilla.sourceforge.net	BSD 3-clause "New" or "Revised" License
YaST	broken/svn/openSUSE-9_3	http://opensuse.org/YaST	MIT License
Zile (Zile is Lossy Emacs)	1.4, 1.5, 1.5.2, 1.5.3, 1.6, 1.6.1, 1.6.2	http://zile.sourceforge.net	GNU General Public License v2.0 or later
afwall	V2.6.0.1, v2.8.0, v2.9.0, v2.9.1, v2.9.4	https://github.com/ukanth/afwall	MIT License
alcatel	20	http://www.alcatel-mobilephones.com/	Apache License 2.0

Component name	Version	Home Page	License
alcatel	4/18/2012, 20120601, 918	http://www.alcatel-mobilephones.com/	GNU General Public License v2.0 or later
appweb	3.0B.0-0	http://code.google.com/p/appweb	Apache License 2.0
asuswrt-merlin	376.48, 376.48, 380.62	https://github.com/RMerl/asuswrt-merlin	Artistic License 1.0
asuswrt-merlin	378.51, 380.62	https://github.com/RMerl/asuswrt-merlin	GNU General Public License v2.0 or later
avahi	v0.6	http://avahi.org	GNU Lesser General Public License v2.1 or later
awokengazebo-lfi	lfi-20080723	http://www.awokengazebo.com/software/lfi/	BSD 4-clause "Original" or "Old" License
beefproject	beef-0.4.3.1	http://beefproject.com	Apache License 2.0
bitswitcher	0.2.0, 0.3.0, 0.3.3	http://sourceforge.net/projects/bitswitcher	GNU General Public License v2.0 or later
buildroot-kindle	master-20130206	https://github.com/twobob/buildroot-kindle	GNU General Public License v2.0 or later
busybox	1.10.0, 1.12.0, 1.2.0, 1.4.0, 1.5.0, 1.8.0, 1_11_0, 1_13_0, 1 14_1, 1_16_0, 1_17_1 17 1, 1_17_2, 1_18_0, 1_18_2, 1_19_0, 1_19_1, 1_19_4, 1_20_2, 1_21_0, 1_24_0, 1_29_0, 1_3_0, 1_7_0	https://github.com/mirror/busybox	GNU General Public License v2.0 only
busybox	1_14_0, 1_15_0, 1_17_0, 1_19_2, 1_19_3, 1_20_0, 1_20_1, 1_28_0,	https://github.com/mirror/busybox	GNU General Public License v2.0 or later

Component name	Version	Home Page	License
catboost/catboost	v0.2	https://catboost.ai	Apache License 2.0
curl	7.16.0	https://curl.se/	curl License
decorator-ko	26, 28	http://jinself.tistory.com/372	Public Domain
file	5.22	http://www.darwinsys.com/file/	Fine Free File Command License
fluxcapacitor	0	https://github.com/majek/fluxcapacitor	MIT License
fvpatwds : fvpat Webdev Server	fvpatwds v0.1.4	http://sourceforge.net/projects/fvpatwds	Apache License 2.0
generator-minxing	1.0.2	https://github.com/yeoman/generator-minxing#readme	Apache License 2.0
geonkick	2.3.6	https://github.com/iurie-sw/geonkick	GNU General Public License v3.0 or later
hostap-ct	lf-5.1.7, lf-5.3.3, lf-5.3.3b, lf-5.3.4, lf-5.3.5	https://github.com/greearb/hostap-ct	BSD 3-clause "New" or "Revised" License
hostapd	hostap_0_5_2, hostap_0_5_3, hostap_0_5_6,	http://w1.fi/hostapd/	GNU General Public License v2.0 or later
howl	0.9.4, 0.9.6, 0.9.7, 0.9.9, 1.0.0,0.9.3, 0.9.1	https://howl.io	BSD 3-clause "New" or "Revised" License
illumos-joyent	20121101	http://www.illumos.org/projects/illumos-gate	Common Development and Distribution License 1.0
krb5/krb5	1.0-alpha3, 1.0-beta2, 1.0-beta5	https://github.com/krb5/krb5	Krb5-MIT License

Component name	Version	Home Page	License
libevent - an event notification library	0.1, 1.0d, 1.0e,1.4.1-beta	http://libevent.org/	BSD 3-clause "New" or "Revised" License
libexpat	1.95.0, 1.95.1, 1.95.2, 2.0.0, v19991013	http://www.libexpat.org/	Expat License
libexpat	V19991013	http://www.libexpat.org/	Mozilla Public License 1.1
linux-yocto-dev	v2.6.12	http://git.yoctoproject.org/cgit/cgit.cgi/linux-yocto-dev/	GNU General Public License v2.0 with Linux Syscall Note
littlekernel-m900-eclair	master-20110326	http://github.com/LouZiffer/littlekernel-m900-eclair	GNU General Public License v2.0 only
lmdb	0.9.18	http://symas.com/mdb/	Open LDAP Public License
math-linux	0.0.1	http://sourceforge.net/projects/math-linux	GNU General Public License v3.0 or later
mod_dup	2.5.0	http://github.com/Orange-OpenSource/mod_dup/	Apache License 2.0
ngx_pagespeed	1.9.32.4-dbg-ssl-crash	https://github.com/pagespeed/ngx_pagespeed	Apache License 2.0
nss_ldap	253	https://github.com/PADL/nss_ldap	GNU Library General Public License v2 or later
opensm	3.3.17	http://www.openfabrics.org/	BSD 2-clause "Simplified" License
pGina	Plugin Bundle 05-11-2006	http://pgina.org/	MIT License
pam_radius	release_2_0_0	http://freeradius.org/pam_radius_auth/	GNU General Public License v2.0 only

Component name	Version	Home Page	License
protovis	3.3.1	http://mbostock.github.io/protovis/	BSD 3-clause "New" or "Revised" License
root-project	5-13-04e	https://root.cern	(GNU Lesser General Public License v2.1 or later AND MIT License AND GNU General Public License v2.0 or later)
rsyslog	sysklogd-141-import	https://www.rsyslog.com/	GNU General Public License v2.0 or later
rtems-libbsd	5.1	http://git.rtems.org/rtems-libbsd.git/	Apache License 2.0
rtl8186 - toolchain	0.5.5_src	http://rtl8186.sourceforge.net	GNU General Public License v2.0 or later
snake-os	0.9	http://code.google.com/p/snake-os/	GNU General Public License v2.0 or later
ssmtp	2.61	http://packages.qa.debian.org/s/ssmtp.html	GNU General Public License v2.0 or later
svn://svn.tug.org/texlive/trunk	texlive-2009.0	http://www.tug.org/texlive/	LaTeX Project Public License - Version Unspecified
util-linux	2.11q, 2.11w, 2.12a, 2.13-pre1	https://en.wikipedia.org/wiki/Util-linux	GNU General Public License v2.0 or later
videolan/vlc	0.5.0	https://github.com/videolan/vlc	(GNU Lesser General Public License v2.1 or later AND GNU General Public License v2.0 or later)
wakame-vdc	v13.06.0	http://wakame.axsh.jp/	Unknown License

Component name	Version	Home Page	License
wpa_supplicant - IEEE 802.1X, WPA, WPA2, RSN, IEEE 802.11i	0.5.0, 0.5.3, 0.5.5, 0.5. 6, 0.5.8, 0.6.0, 0.6.10, 0.6.2, 0.6.3, 0.6.4, 0.6.8, 0.7.0, 0.7.1, 0.7.2, 0.7.3, 1, 2, 2.1, 2.2, 2.3, 2.4, 2.5, 2.6, 2.7, 2.7+git20190108+11ce7a1, , 2.7~git20180504+60a5737, 2.7~git20180606+b915f2c, 2.7~git20180706+420b5dd	http://w1.fi/wpa_supplicant/	BSD 3-clause "New" or "Revised" License
xorp.ct	1.5, xorp-1-7	http://www.candelatech.com/xorp.ct	MIT License
zeroconf	0.9	https://files.pythonhosted.org/packages/20/d7/418ff6c684ace0f5855ec56c66cfa99ec50443c41693b91e9abcccfa096c/zeroconf-0.20.0.tar.gz	GNU General Public License v2.0 or later

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. Neither NVIDIA Corporation nor any of its direct or indirect subsidiaries and affiliates (collectively: “NVIDIA”) make any 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, and Mellanox are trademarks and/or registered trademarks of NVIDIA Corporation and/or Mellanox Technologies Ltd. in the U.S. and in other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright
© 2023 NVIDIA Corporation & affiliates. All Rights Reserved.