Hewlett Packard Enterprise

HPE Slingshot Monitoring Software: Actionable Insights for HPC and AI Systems

Sahil Patel, *Technical Product Manager, HPE*

CUG 2025

Scan the QR code to access the frequent Q&A’.

Additionally includes HPE Slingshot AIOps resources and contact details for the Telemetry-focused members of Slingshot Product Management and the Development team.

Agenda

- Introduction
- Overview of Slingshot Telemetry & More
- HPE Slingshot Monitoring Software Overview
- Key Features & Capabilities
- Summary & Q&A

You’ve got Slingshot. But are you leveraging its full power?

Awareness is power. Most systems aren’t using what’s already available.

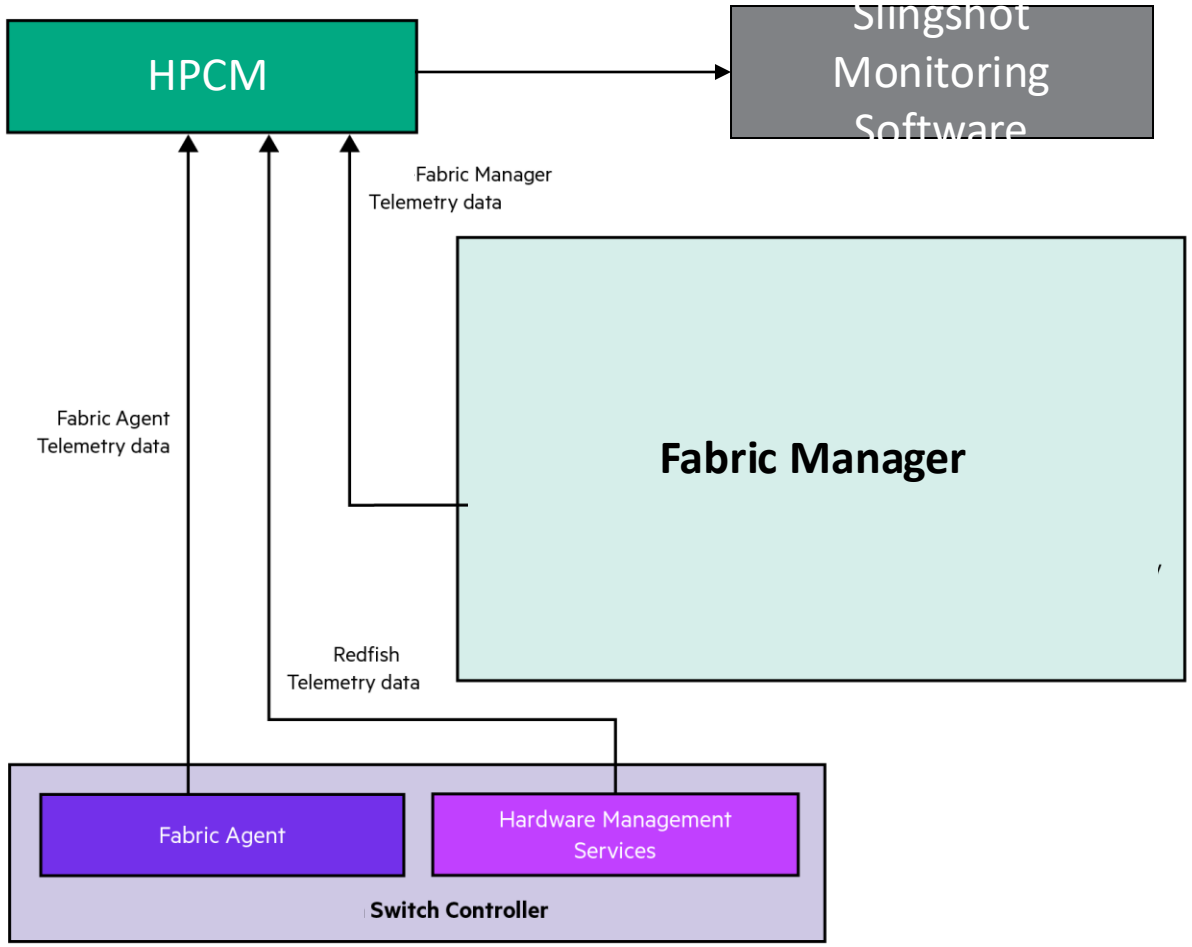
Deeper Switch Insights:
Port Conditions
ASIC Errors

Evolved HPCM Monitoring:
Unified features across releases
System-aware Data

New Unified Interface:
Easier
Interactive
Action-focused

Your systems are capable of much more, lets unlock that value.

Slingshot Monitoring Software Telemetry



- **Sources of Telemetry Data:**
  - *Fabric Manager (FM)*
    - –Generates health events when it detects an info, warning, or error condition
  - *Switch Controller*
    - –Fabric Agent
      - –Collects ASIC specific metrics (i.e., performance counters, routing errors, port errors, etc)
      - –Generates health events when it detects info, warning, or error condition
    - –Hardware Management Services
      - – Periodically collects switch environmental metrics
      - – Generates switch environmental or components redfish events

Telemetry – Slingshot
Metrics

Metrics

- Fabric performance telemetry
- Fabric Critical Telemetry
- Switch hardware telemetry

90+ metrics/counters

Fabric Critical Telemetry

Message Id	Metrics/Counters	Collection Rate	Definition
PortErrors	`pcs_corrected_cw`	On config timer	The number of corrected FEC errors.
PortErrors	`pcs_uncorrected_cw`	On config timer	The number of uncorrected FEC errors.
RoutingErrors	`frf_empty_route_uf_cntr`	On config timer	An unexpected frame was received. The frame was dropped.
RoutingErrors	`frf_empty_route_cntr`	On config timer	A received frame was not able to be routed. The frame was dropped.
RoutingErrors	`frf_empty_route_edge_cnt`	On config timer	A received frame, directed to an edge port, was dropped because none of the candidate ports

Switch Hardware Telemetry

Message Id	Metrics	Collection Rate	Definition
Current	`PowerSupply`	1 Hz	The current, in Amps, of all power supplies.
Current	`VoltageRegulator`	1 Hz	The current, in Amps, of a voltage regulator.
Power	`PowerSupply`	1 Hz	The input or output power, in Watts, from a power supply.

Telemetry – Slingshot

CrayFabricHealth events

Events

- CrayFabricHealth events
- ASIC error flag events
- Port Error Events
- Redfish events
- High speed network link events

300+ events

CrayFabricHealth events

Fabric Health events are data generated by the Fabric Manager to report the health of the system. Fabric Health events are posted on health check occurrence - including when the error clears and a healthy state is reached. Generally, the Fabric Health events cover following areas:

- **Traffic**: Health events impacting passing traffic or caused by detecting issues in the data-plane of the Fabric
- **Configuration**: Health events impacting configuration of the Fabric policies or caused by detecting misconfigurations of the Fabric
- **Runtime**: Health events impacting the health and management state of the Fabric Manager itself, including resource or management network issues.
- **Security**: Health events indicating configuration issues related to security aspects of Fabric Manager.

With Health Monitors support, Health events can have custom categories and hence are not limited to the above mentioned categories.

Currently system has few default Health Monitors and they report events under following categories:

1. All ASIC errors are reported under “AsicErrorFlagHealth” category.
2. All Port error events are reported under “PortErrorEvent” category.

Health event includes an associated Severity:

- **Critical**: The impacted object or area is unhealthy and immediate action should be taken to repair or mitigate.
- **Warning**: The impacted object may be unhealthy, but the issue may not require immediate attention and may already be mitigated.
- **Info**: The impacted object has updated information, and no action is required.
- **Healthy**: A specific health check was performed on the object resulting in no detected error for that type.

Health events may be temporary due to an intermittent or operational change on the system. The Fabric Health engine maintains an active operational window:

- configurable by the administrator.
- which reports current active alerts and health issues affecting the fabric. The operational window is a moving window with a range of time, and refreshed on a frequent basis.

Events generated by the Fabric Manager and Fabric Agent on the Rosetta

- **Event Subcategory**: Defines the subcategory of the event
- **Runtime**: Includes fabric policy, telemetry, Fabric Manager host health, and Fabric Manager node file system health
- **Configuration**: Includes switch, port and fabric policies, and configuration
- **Traffic**: Includes port status, link status, and routing related issues
- **Security**: Includes authentication
- **Firmware**: Includes switch and ASIC firmware related issues
- **Event Name**: Defines the name of the event
- **Cause**: Possible cause for the event
- **Recommended Action**: Remediation action in case of failures

Telemetry – Slingshot

ASIC error flag events

Events

ASIC error flag events

Health monitors are enabled by default for all ASIC error categories and also run for the `AsicErrorFlags.frf.empty_route` ASIC error.

The following events of type Critical or Warning are generated for ASIC errors and accessible under `/metrics` on the Rosetta switch.

Events are categorized as Critical or Warning health events based on the values set for the warningHealthAlertLevel and criticalHealthAlertLevel which in turn decides the type of monitor. The warningHealthAlertLevel and criticalHealthAlertLevel attributes for an event can be modified in `/fabric/asic-error-flag-policies/standard` to raise or lower the severity of a health alert. However, it is recommended to use the default values for these attributes.

NOTE: All of the error flag events in the following table meet the following criteria:

- **Subcategory:** CrayFabricHealth
- **Severity:** Critical or Warning
- **Service Name** RosettaAsicErrorFlagAgentsService
- **Name** AsicErrorFlagHealth.critical.<CATEGORY> **OR** AsicErrorFlagHealth.warning.<CATEGORY>
  - NOTE: Only the <CATEGORY> portion of the event name is shown in the table

Prefix location with:

Frequency/Trigger	Name	Location
Critical or Warning type EC_CRIT monitor for ASIC error flags belonging to EC_CRIT class as per the conditions specified in the healthmonitor rules	EC_CRIT	/fabric/health-monitors/asic-errors-general-critical-ec_crit /fabric/health-monitors/asic-errors-general-warning-ec_crit
Critical or Warning type EC_UNCOR_S monitor for ASIC error flags belonging to EC_UNCOR_S class as per the conditions specified in the healthmonitor rules	EC_UNCOR_S	/fabric/health-monitors/asic-errors-general-critical-ec_uncor_s /fabric/health-monitors/asic-errors-general-warning-ec_uncor_s

Telemetry – Slingshot

Redfish events

Events

- CrayFabricHealth events
- ASIC error flag events
- Port Error Events
- Redfish events
- High speed network link events

Redfish events

This section includes information for interpreting Redfish events generated by the Redfish APIs.

The switch controller implements a Redfish endpoint, which conforms to the standard Redfish API. Refer to the following external documentation for more information.

In addition to standard events, HPE implements the OEM extensions outlined in the following subsections.

Cray Uefi

Cray Uefi Message Registry	Description
FatalUefiError	Indicates that a critical error was encountered.
NonFatalUefiError	Indicates that a non-fatal error was encountered.
CorrectableUefiError	Indicates that a correctable error was encountered.
PreviousBootUefiError	Indicates that a previous boot error was encountered.
SMMBPingTimeout	The SMMB ping timed out
GeneralUefiError	Indicates that a UEFI error was encountered.

Cray Alert

Cray Alert Message Registry	Description
ResourcePowerStateChanged	The power state of resource %1 has changed to type %2. The state types shall be used from Resource.PowerState.
HardwareFailoverDetected	The primary hardware associated with a resource has failed. A backup device has been enabled to allow normal operation.

Cray Alert Message Registry	Description
HardwareDegradationDetected	A hardware resource is operational but is functioning at a reduced capacity.
HardwareDegradationCleared	A hardware resource has regained full operational capacity.
HardwareFailureDetected	A hardware resource has failed. The system is still operational but immediate action should be taken to avoid downtime.
HardwareDisconnectDetected	A hardware resource has been disconnected. The system is still operational but immediate action should be taken to avoid downtime.
HardwareDisconnectCleared	A hardware resource has been reconnected.
HardwareEmergencyPowerOffDetected	A hardware resource has powered down to protect itself from damage.

PowerFailureDetected	Indicates that a power failure has occurred.
HsnTransceiverRemoved	Indicates that a high-speed-network transceiver was removed or powered off.
HsnTransceiverInstalled	Indicates that a high-speed-network transceiver is installed.
HsnTransceiverMgmtError	Indicates the management interface of a high-speed-network transceiver is in error.
HsnTransceiverEPO	Indicates a high-speed-network transceiver emergency power off (EPO) event.
HsnLinkUpDetected	Indicates that a high-speed-network link changed state to up.
HsnLinkDownDetected	Indicates that a high-speed-network link changed state to down.
HsnLinkErrorDetected	Indicates that a high-speed-network link changed state to error.

Telemetry – Slingshot

High speed network link & port error events

Events

- CrayFabricHealth events
- ASIC error flag events
- Port Error Events
- Redfish events
- High speed network link events

High speed network link events

Asynchronous link-events are posted by the agent when there is a change in state of a link. The following link-events are posted.

Supported events

The following event types are supported:

- Link-Status Up: Link has met quality standard and is in use, included in the routing tables.
- Link-Status Down: Link is enabled but has not come up, or went down from operational, includes a description of why the link failed.
- Link-Status Error: Link has encountered an error while trying to come up or while in operation.
- Link Flap: A link went down and back up within a short period of time. Includes a description of why the link failed.

Port state information

Port state information including MAC, MTU, Type, SubType, and Speed is posted in the Link-Status Up event:

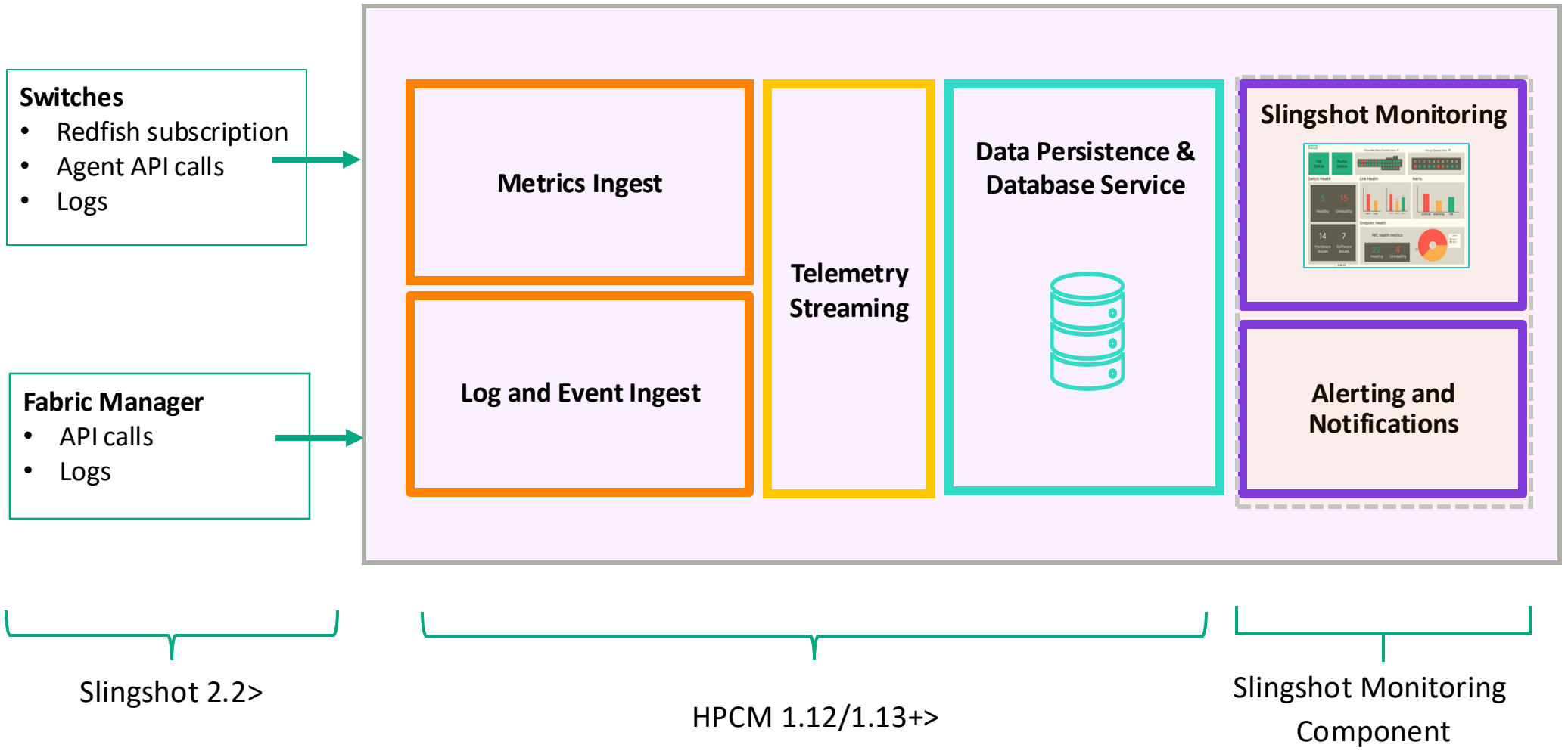
- MTU: Maximum Transmission Unit (MTU) is the maximum frame size supported by the interface.
- MAC: MAC address of the interface.
- Type: Type information for the interface, such as Ethernet or Fabric.
- SubType: SubType information of the interface, such as Edge, local, or global.
- Speed: Speed of the interface in Mbps.

Port error events

Slingshot port error events are generated by the switch agents for critical port errors reported by Rosetta hardware. These errors are posted on occurrence.

Message Id	Metrics/Counters	Collection Rate	Definition
PortErrorEvent	pcs_corrected_cw	On occurrence	Posted when the number of corrected FEC errors exceeds a threshold.
PortErrorEvent	pcs_uncorrected_cw	On occurrence	Posted when the number of uncorrected FEC errors exceeds a threshold.
PortErrorEvent	pcs_llr_replay	On occurrence	Posted when the number of llr_tx_replay_event and llr_rx_replay_event errors exceeds a threshold.
PortErrorEvent	ibuf_full	On occurrence	Posted when the number of ibuf_full errors exceeds a threshold.
PortErrorEvent	ageq_stall_obuf	On occurence	Indicates excessive blocking on the egress port, This may be caused by a very high error rate on the link preventing transmission through the LLR. Could be caused by bad configuration.
PortErrorEvent	ageq_stall_ibuf_vc_0[0-3]	On occurence	Indicates the link partner's buffer space is not being released. Likely indicates a critical problem on another Rosetta. Could be caused by very high levels of network congestion caused by a wide uncontrolled incast.

Slingshot Monitoring Architecture and Deployment Model



Customer Struggles, Needs, and Wants

Pain Points

- Immense amount of manual processing of data to draw insights
- Several engineering hours invested in root cause analysis of issues
- Uncertainty about what HPE Slingshot components to monitor leads to blind spots and inefficiencies
- Building custom diagnostic tools (via CLI) is slow and time-consuming process, delaying issue resolution and increasing operational overhead

User Considerations

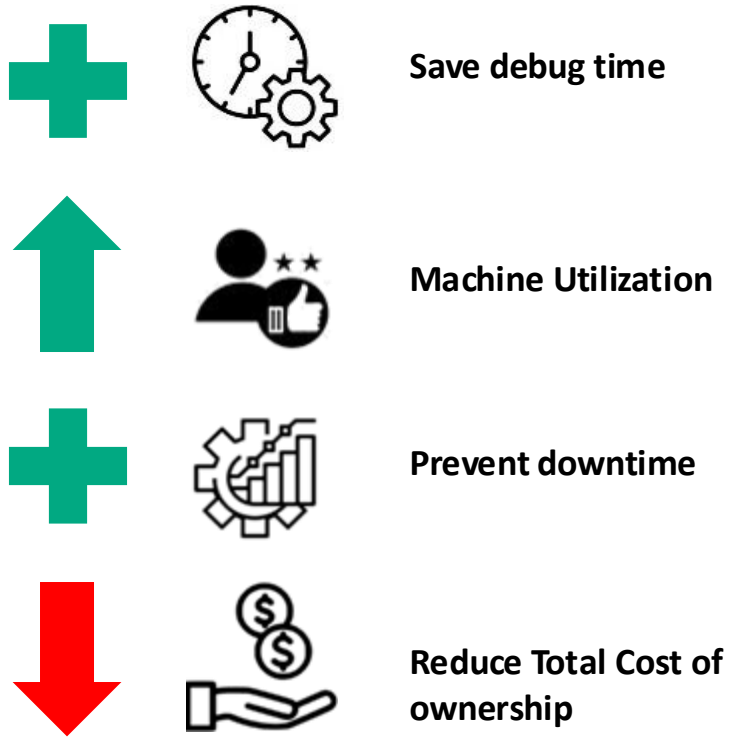
- Is my fabric functional?
- How many switches are healthy and unhealthy?
- How many fabric and end ports are online and offline?
- Are there issues with the links? Any ASIC errors?

Outcomes

- Quickly summarize and synthesize insights in real-time
- Reduce the manual time taken for the root cause analysis
- Improve the overall utilization of their system through data-driven decision-making

Target User

- HPC / AI customers using Slingshot
- System and network administrators



HPE Slingshot Monitoring Software: Unifying the Monitoring Experience

HPE Slingshot Monitoring Software is designed to **generate real-time insights,** offering a quick and clear synthesis of the fabric's overall health.

Offering **real-time notifications** to identify key issues and to alert customers, **easy-to-read dashboards reduce root cause identification time and manual troubleshooting.**

End users can greatly **improve / maintain fabric utilization**, ensuring all components operate optimally while swiftly detecting and resolving any degradation or failures.

Welcome to Grafana

Email or username

email or username

Password

password

Log in

Forgot your password?

FM Status - fmn1

fmn1 HEALTHY

FM Status - fmn2

fmn2 HEALTHY

Total Unhealthy Links

1

Unhealthy Links by Type

Total Unhealthy Switches

9

Switch Group Health

Active Alerts

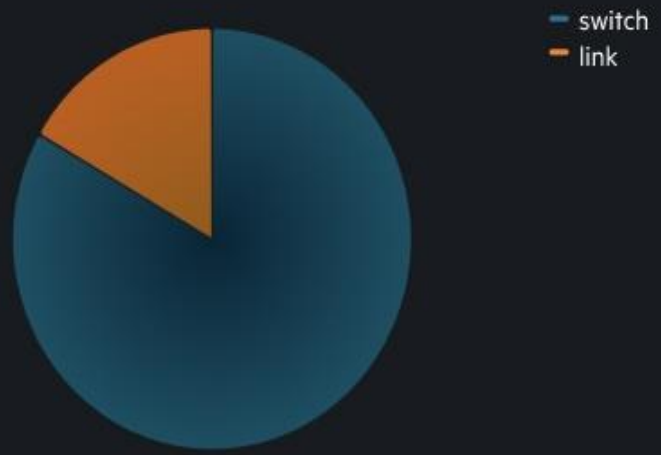
7

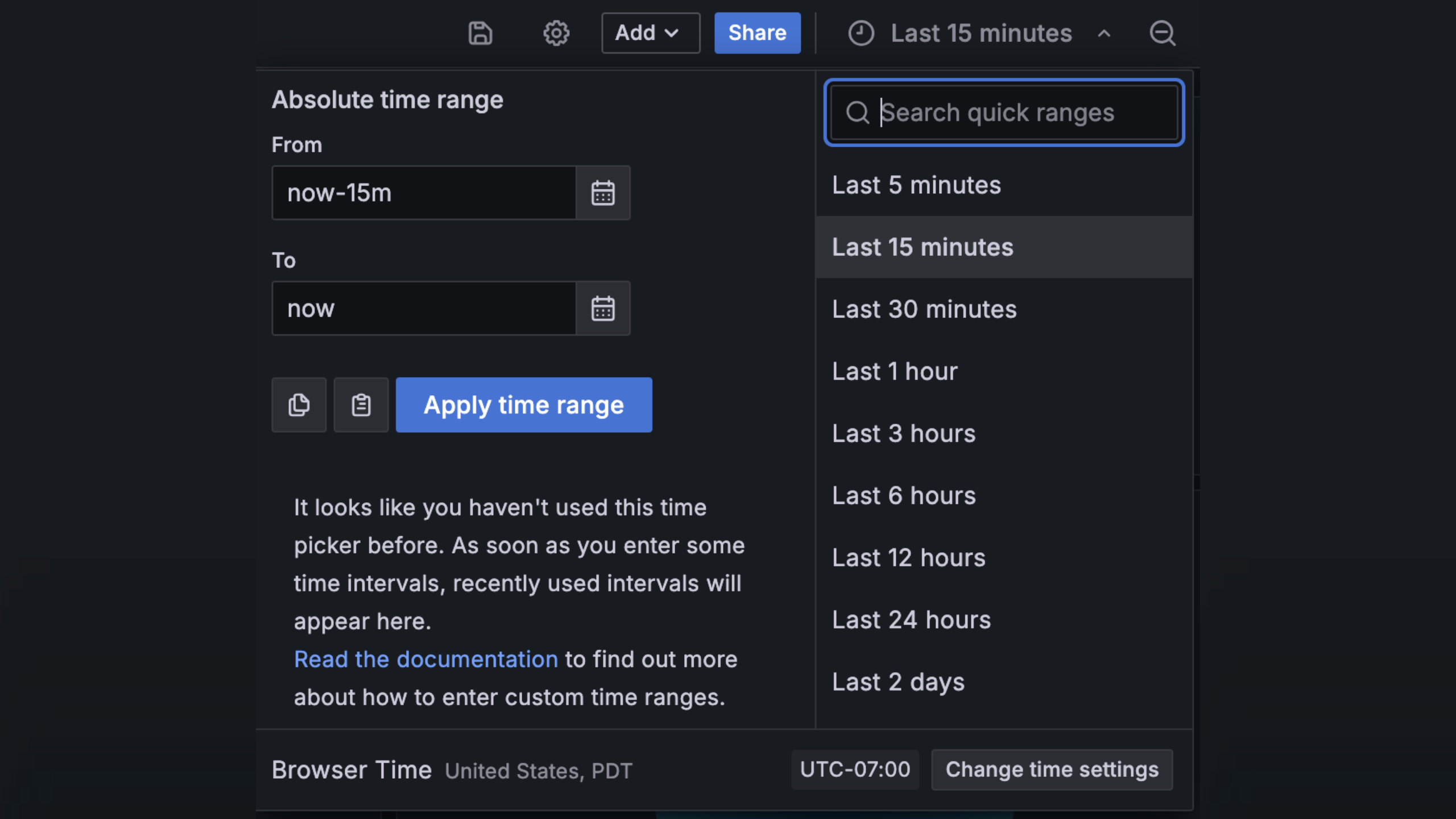
Silenced Alerts

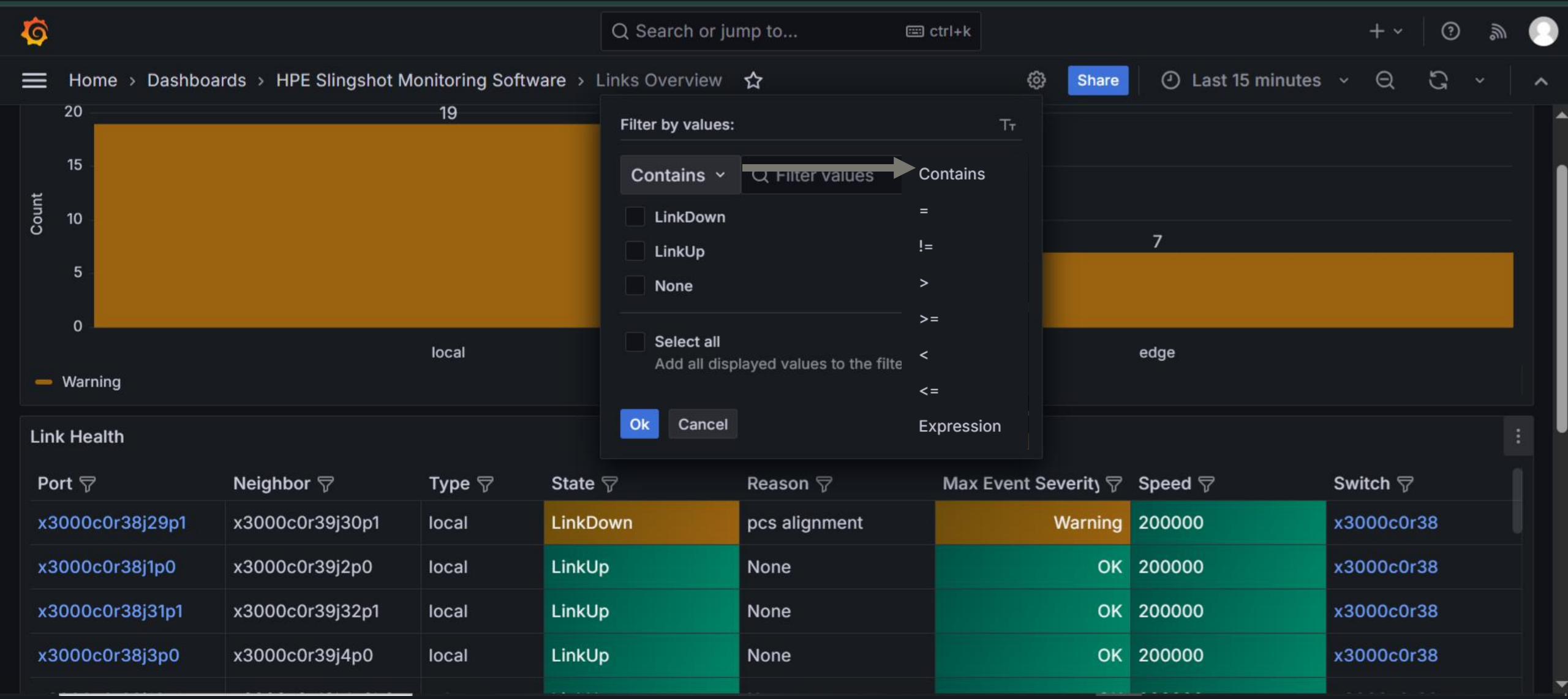
0

Active Alerts by Severity

Alerts by Component Type

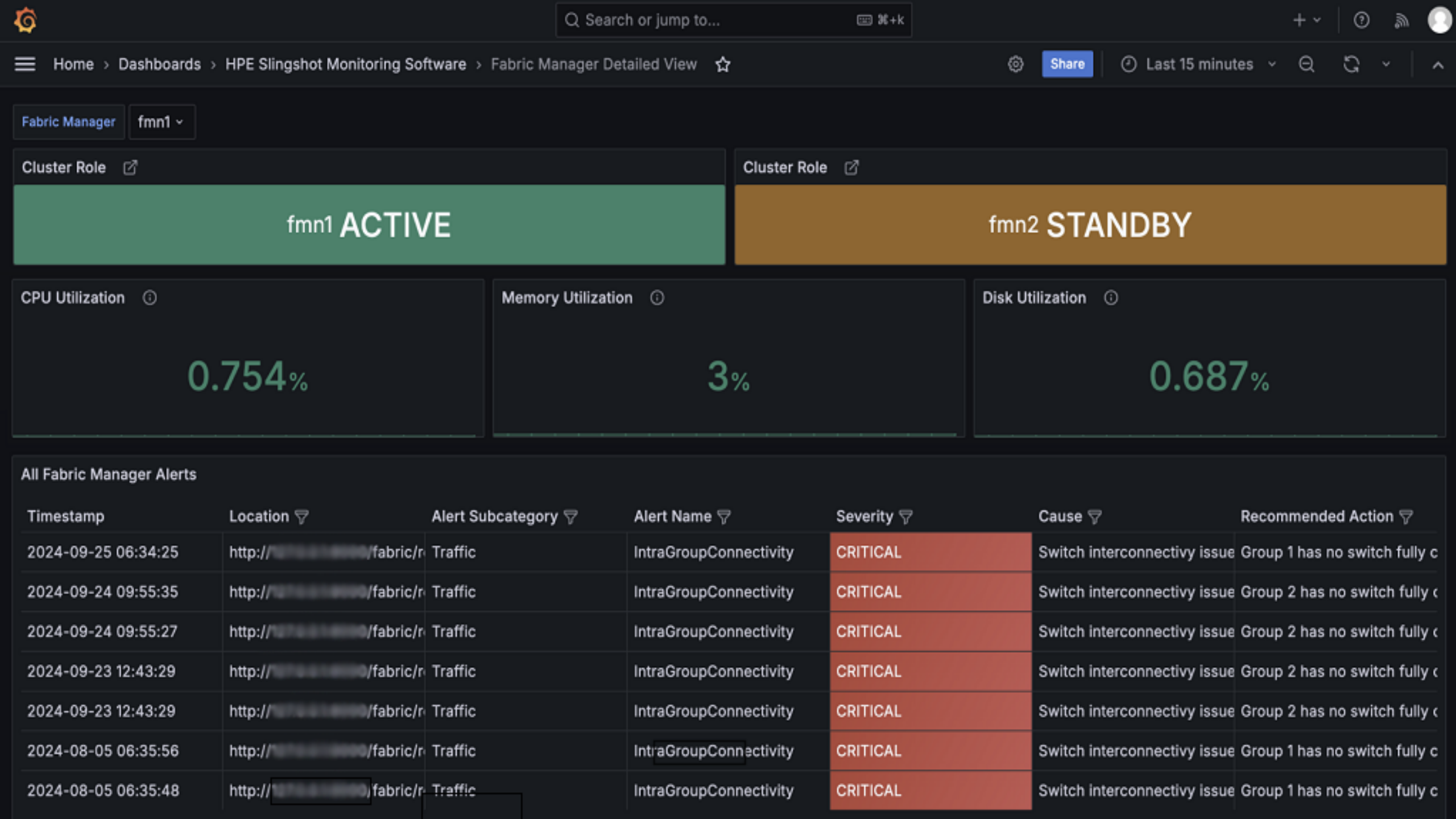



Add
Share
Last 15 minutes
Absolute time range
From
now-15m
To
now
Apply time range
It looks like you haven't used this time picker before. As soon as you enter some time intervals, recently used intervals will appear here.
Read the documentation to find out more about how to enter custom time ranges.
Search quick ranges
Last 5 minutes
Last 15 minutes
Last 30 minutes
Last 1 hour
Last 3 hours
Last 6 hours
Last 12 hours
Last 24 hours
Last 2 days
Browser Time United States, PDT
UTC-07:00
Change time settings



Link Health

Port	Neighbor	Type	State	Reason	Max Event Severity	Speed	Switch
x3000c0r38j29p1	x3000c0r39j30p1	local	LinkDown	pcs alignment	Warning	200000	x3000c0r38
x3000c0r38j1p0	x3000c0r39j2p0	local	LinkUp	None	OK	200000	x3000c0r38
x3000c0r38j31p1	x3000c0r39j32p1	local	LinkUp	None	OK	200000	x3000c0r38
x3000c0r38j3p0	x3000c0r39j4p0	local	LinkUp	None	OK	200000	x3000c0r38


Search or jump to... ⌘+k
Home › Dashboards › HPE Slingshot Monitoring Software › Fabric Manager Detailed View
Share
Last 15 minutes
Fabric Manager
fmn1
Cluster Role
fmn1 ACTIVE
Cluster Role
fmn2 STANDBY
CPU Utilization
0.754%
Memory Utilization
3%
Disk Utilization
0.687%
All Fabric Manager Alerts
Timestamp
Location
Alert Subcategory
Alert Name
Severity
Cause
Recommended Action
2024-09-25 06:34:25
http://[redacted]/fabric/r
Traffic
IntraGroupConnectivity
CRITICAL
Switch interconnectivy issue
Group 1 has no switch fully c
2024-09-24 09:55:35
http://[redacted]/fabric/r
Traffic
IntraGroupConnectivity
CRITICAL
Switch interconnectivy issue
Group 2 has no switch fully c
2024-09-24 09:55:27
http://[redacted]/fabric/r
Traffic
IntraGroupConnectivity
CRITICAL
Switch interconnectivy issue
Group 2 has no switch fully c
2024-09-23 12:43:29
http://[redacted]/fabric/r
Traffic
IntraGroupConnectivity
CRITICAL
Switch interconnectivy issue
Group 2 has no switch fully c
2024-09-23 12:43:29
http://[redacted]/fabric/r
Traffic
IntraGroupConnectivity
CRITICAL
Switch interconnectivy issue
Group 2 has no switch fully c
2024-08-05 06:35:56
http://[redacted]/fabric/r
Traffic
IntraGroupConnectivity
CRITICAL
Switch interconnectivy issue
Group 1 has no switch fully c
2024-08-05 06:35:48
http://[redacted]/fabric/r
Traffic
IntraGroupConnectivity
CRITICAL
Switch interconnectivy issue
Group 1 has no switch fully c

Severity: All | Component Type: All

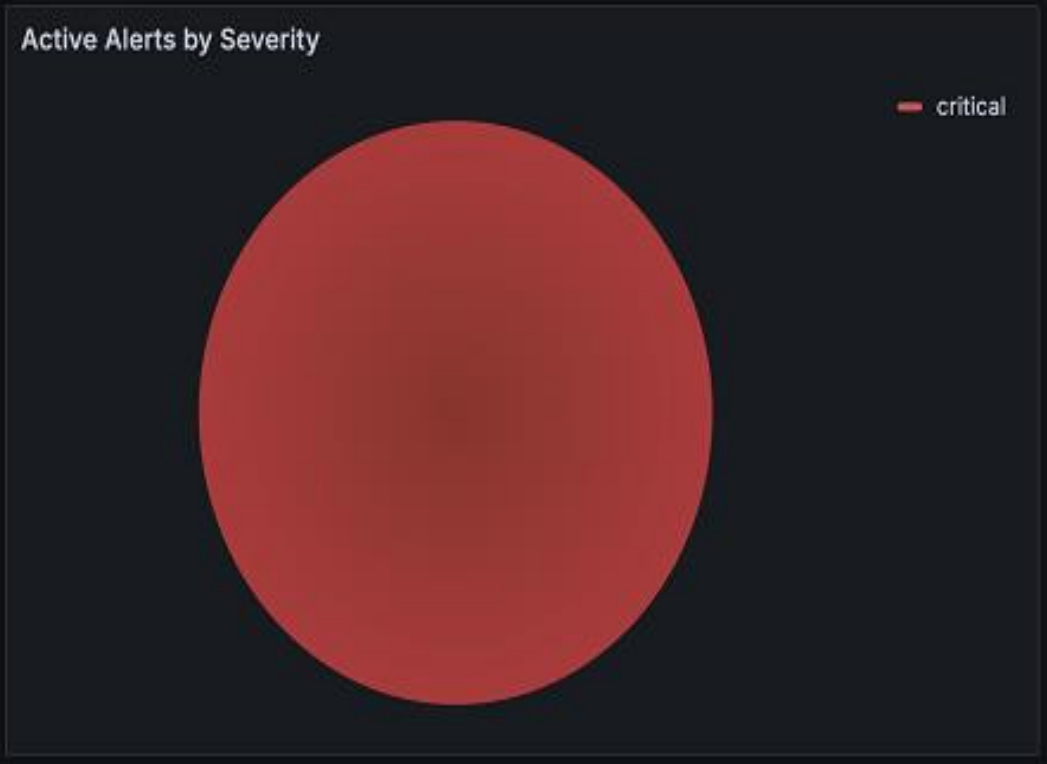
Active Alerts

4

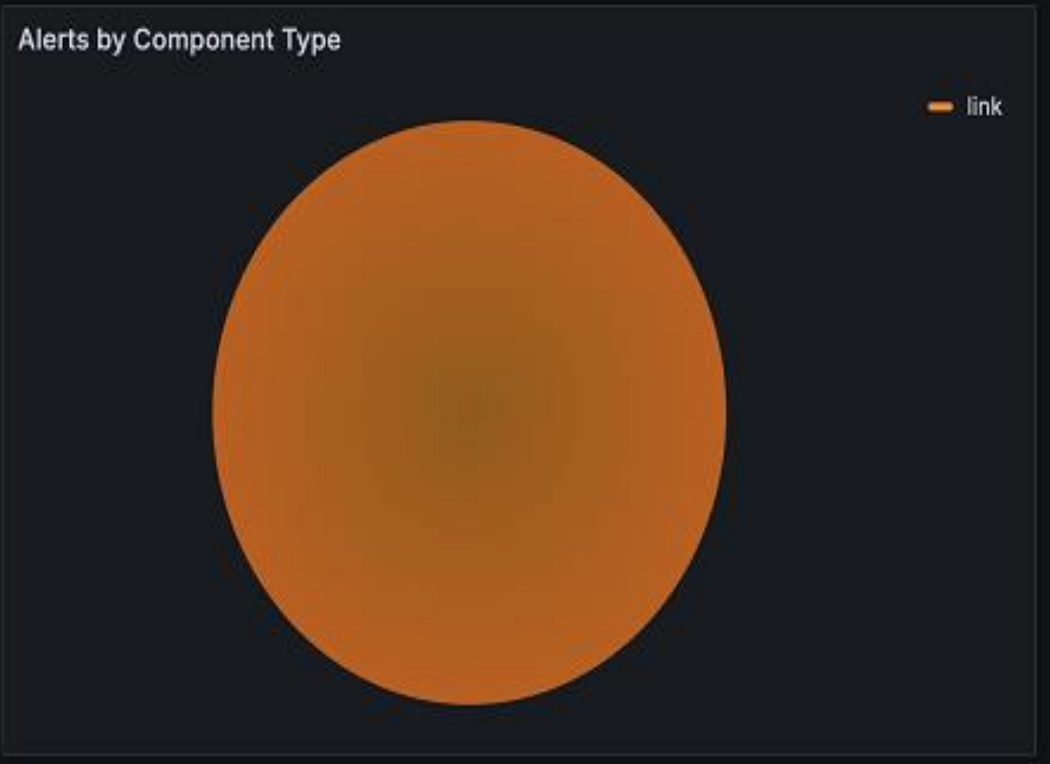
Silenced Alerts

0

Active Alerts by Severity



Alerts by Component Type



Timestamp	Name	Type	Severity	Component Type	Location	Message
2024-11-04 09:50:49.001	Slingshot Link Health Critica	LinkErrors.MultiBitError	critical	link	x3000c0r38j2p0	value=Recommend: Multi-bit
2024-11-04 09:50:48.996	Slingshot Link Health Critica	LinkErrors.MultiBitError	critical	link	x3000c0r38j14p1	value=Recommend: Multi-bit
2024-11-04 09:50:48.991	Slingshot Link Health Critica	LinkErrors.MultiBitError	critical	link	x3000c0r38j14p0	value=Recommend: Multi-bit
2024-11-04 09:50:49.005	Slingshot Link Health Critica	LinkErrors.MultiBitError	critical	link	x3000c0r38j2p1	value=Recommend: Multi-bit

Alertmanager Alerts Silences Status Settings Help

New Silence

Start 2025-04-27T22:48:06.446Z

Duration 2h

End 2025-04-28T00:48:06.446Z

Matchers Alerts affected by this silence

+

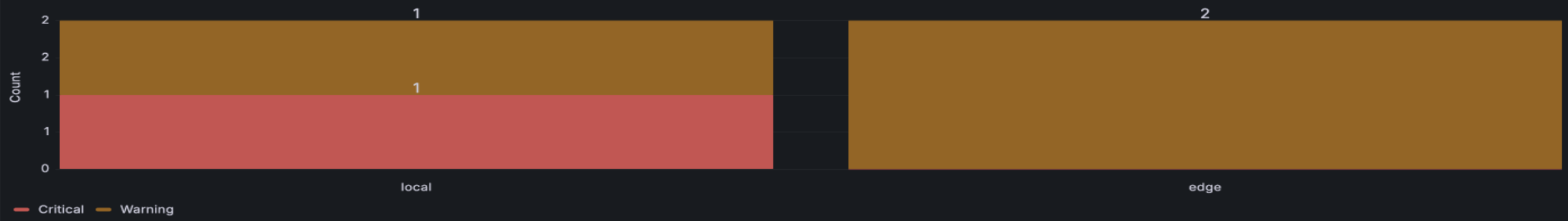
Custom matcher, e.g. env="production"

Creator

Comment

Preview Alerts Create Reset

Unhealthy Links by Type



Link Health

Port	Neighbor	Type	State	Reason	Max Event Severity	Speed	Switch
x3000c0r38j3p1	x3000c0r39j4p1	local	LinkError	None	Critical	200000	x3000c0r38
x3000c0r38j4p1	x3000c0r39j3p1	local	LinkUp	None	OK	200000	x3000c0r38
x3000c0r38j2p1	x3000c0r39j1p1	local	LinkUp	None	OK	200000	x3000c0r38
x3000c0r38j2p0	x3000c0r39j1p0	local	LinkUp	None	OK	200000	x3000c0r38
x3000c0r38j14p1	gaudi2-hsn3	edge	LinkDown	uncorrectable MBE detect	OK	200000	x3000c0r38
x3000c0r38j14p0	gaudi2-hsn2	edge	LinkDown	uncorrectable MBE detect	OK	200000	x3000c0r38
x3000c0r38j15p0	None	None	LinkDown	uncorrectable MBE detect	Warning	200000	x3000c0r38
x3000c0r38j30p0	x3000c0r39j29p0	local	LinkUp	None	OK	200000	x3000c0r38
x3000c0r39j31p1	x3000c0r38j32p1	local	LinkUp	None	OK	200000	x3000c0r39
x3000c0r39j31p0	x3000c0r38j32p0	local	LinkUp	None	OK	200000	x3000c0r39
x3000c0r39j29p0	x3000c0r38j30p0	local	LinkUp	None	OK	200000	x3000c0r39
x3000c0r39j16p0	x3000c0s18b4n0h1	edge	LinkUp	None	OK	200000	x3000c0r39
x3000c0r39j18p1	x3000c0s18b1n0h1	edge	LinkUp	None	OK	200000	x3000c0r39
x3000c0r39j18p0	x3000c0s18b2n0h1	edge	LinkUp	None	OK	200000	x3000c0r39

link x3000c0r38j4p1

Link Type: local

Status: LinkDown	BER: 0	Status: LinkDown	BER: 0
Link Speed: 200 Gbps	FEC: 0	Link Speed: 200 Gbps	FEC: 0

Link Status Reason / Recommended Action

Reason
pcs link down

Recommended Action
The signal quality between the two endpoints is poor and insufficient for PCS alignment. See: link hardware debug sequence.

Link Status Reason / Recommended Action

Reason
headshell absent

Recommended Action
Examine the jack and check for loose or missing cable headshell. If problem persists, see HPE Slingshot Troubleshooting guide.

LinkDown Events by Cause

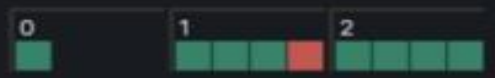
HSN Link Events

Timestamp	Event Type	Severity	Message	Cause
2024-11-04 10:07:58.169	HsnLinkDownDetected	Warning	The link x3000c0r38b0:j4p1 has change	pcs link down
2024-11-04 10:07:50.493	HsnLinkUpDetected	OK	The link x3000c0r38b0:j4p1 has change	None
2024-11-04 10:07:04.687	HsnLinkDownDetected	Warning	The link x3000c0r38b0:j4p1 has change	pcs link down

Link-Related Fabric Health Events

Timestamp	Alert Category	Alert Name	Severity	Cause	Recommended Action
2024-08-20 10:50:35	LinkErrors	cassinipause	CRITICAL	Port has entered error state cassi	Inspect the switch port. If issue p
2024-08-20 09:50:31	LinkErrors	cassinipause	CRITICAL	Port has entered error state cassi	Inspect the switch port. If issue p
2024-08-20 07:50:23	LinkErrors	cassinipause	CRITICAL	Port has entered error state cassi	Inspect the switch port. If issue p
2024-08-20 06:50:19	LinkErrors	cassinipause	CRITICAL	Port has entered error state cassi	Inspect the switch port. If issue p
2024-08-20 05:50:15	LinkErrors	cassinipause	CRITICAL	Port has entered error state cassi	Inspect the switch port. If issue p
2024-08-20 04:50:11	LinkErrors	cassinipause	CRITICAL	Port has entered error state cassi	Inspect the switch port. If issue p

Switch Health



Switch	IP	Status	ASIC Temperature	Max Event Severit	Firmware	Uptime	Group ID	Switch ID
x3001c0r40	[redacted]	up	38	OK	sc.2.2.0-72-slingshot	38 days, 12:27	0	0
x9000c1r1	[redacted]	up	39	OK	sc.2.2.0-72-slingshot	40 days, 5:16	1	0
x9000c1r3	[redacted]	up	35	OK	sc.2.2.0-72-slingshot	21 days, 12:39	1	1
x9000c1r5	[redacted]	up	34	OK	sc.2.2.0-72-slingshot	21 days, 12:39	1	2
x9000c3r1	[redacted]	up	37	OK	sc.2.2.0-72-slingshot	42 days, 3:03	2	0
x9000c3r3	[redacted]	up	38	OK	sc.2.2.0-72-slingshot	42 days, 3:03	2	1
x9000c3r5	[redacted]	up	35	OK	sc.2.2.0-72-slingshot	41 days, 1:55	2	2
x9000c3r7	[redacted]	up	35	OK	sc.2.2.0-72-slingshot	42 days, 3:03	2	3
x9000c1r7	[redacted]	down		OK	None	None	1	3

switch: x3000c0r38

Local Port Health

Edge Port Health

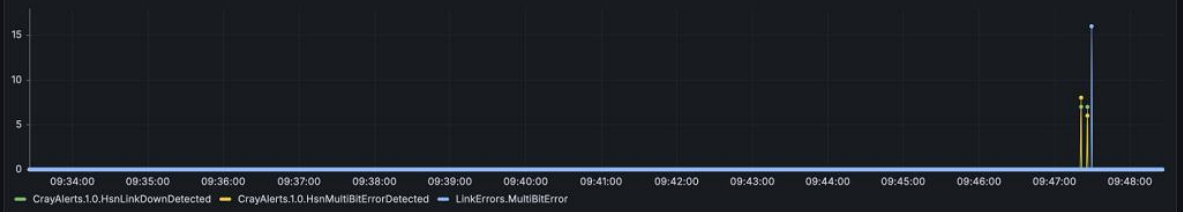
Global Port Health

Firmware version: sc.2.2.0-72-slingshot-release

IP address: 172.23.0.4

Uptime: 31 days 14:51

All Alerts



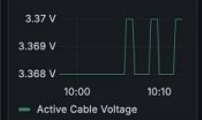
ASIC Temperature

ASIC Core Voltage

ASIC SerDes Voltage

ASIC SerDes I/O Voltage

Active Cable Voltage



Health Alerts

Timestamp	Location	Alert Subcategory	Alert Name	Severity	Cause	Recommended Action
2024-11-04 09:47:29.747	x3000c0r38j2p0	LinkErrors	MultiBitError	CRITICAL	Port has encountered multi-l	Multi-bit errors are detected
2024-11-04 09:47:29.742	x3000c0r38j2p1	LinkErrors	MultiBitError	CRITICAL	Port has encountered multi-l	Multi-bit errors are detected
2024-11-04 09:47:29.733	x3000c0r38j14p1	LinkErrors	MultiBitError	CRITICAL	Port has encountered multi-l	Multi-bit errors are detected
2024-11-04 09:47:29.728	x3000c0r38j14p0	LinkErrors	MultiBitError	CRITICAL	Port has encountered multi-l	Multi-bit errors are detected

HSN Link Events

Timestamp	Jack	Event Type	Severity	Message	Cause
2024-11-04 09:47:21	j8p0	CrayAlerts.1.0.HsnLinkDownDetec	Warning	The link x3000c0r38b0:j8p0 has	uncorrectable MBE detected
2024-11-04 09:47:21	j8p1	CrayAlerts.1.0.HsnLinkDownDetec	Warning	The link x3000c0r38b0:j8p1 has (	uncorrectable MBE detected
2024-11-04 09:47:21	j10p1	CrayAlerts.1.0.HsnLinkDownDetec	Warning	The link x3000c0r38b0:j10p1 has	uncorrectable MBE detected
2024-11-04 09:47:21	j10p0	CrayAlerts.1.0.HsnLinkDownDetec	Warning	The link x3000c0r38b0:j10p0 has	uncorrectable MBE detected
2024-11-04 09:47:21	j11p0	CrayAlerts.1.0.HsnLinkDownDetec	Warning	The link x3000c0r38b0:j11p0 has	uncorrectable MBE detected
2024-11-04 09:47:21	j5p1	CrayAlerts.1.0.HsnLinkDownDetec	Warning	The link x3000c0r38b0:j5p1 has (	uncorrectable MBE detected

Redfish Alerts

Timestamp	Location	Alert Name	Severity	Message
2024-11-04 09:47:21.439	x3000c0r38	HsnMultiBitErrorDetected	Critical	The link x3000c0r38b0:j8p1 has encoun
2024-11-04 09:47:21.439	x3000c0r38	HsnMultiBitErrorDetected	Critical	The link x3000c0r38b0:j10p1 has encou
2024-11-04 09:47:21.439	x3000c0r38	HsnMultiBitErrorDetected	Critical	The link x3000c0r38b0:j10p0 has encou
2024-11-04 09:47:21.439	x3000c0r38	HsnMultiBitErrorDetected	Critical	The link x3000c0r38b0:j11p0 has encou
2024-11-04 09:47:21.439	x3000c0r38	HsnMultiBitErrorDetected	Critical	The link x3000c0r38b0:j11p1 has encour
2024-11-04 09:47:21.438	x3000c0r38	HsnMultiBitErrorDetected	Critical	The link x3000c0r38b0:j5p1 has encoun

← Back to Link Detailed View for x3000c0r38j4p1

Common Link Troubleshooting Procedures

Link Hardware Debug Sequences:

Steps to debug Mountain Cabinet downed links between NIC and switch:

1. Validate both switch and NIC are properly configured and attempting to bring up the HSN link. For more information, see sections 3.2.2, 3.4.2.5, and 3.4.2.6 of the HPE Slingshot Troubleshooting Guide.
2. Perform "group hug": apply pressure to both the compute and switch blades simultaneously to seat the ExaMax connectors more firmly.
3. Reseat the switch blade (this will act as an ASIC reset and reboot as well).
4. Reseat the compute blade.
5. Swap NIC mezzanine cards between nodes. Observe if the failure follows the NIC or stays with the cable. If failure follows the NIC, replace the NIC. If failure stays with the L0 cable, replace the L0 cable.
6. Replace node card with known node card that has good link on the reporting errored HSN link.
7. Replace switch.

Steps to debug River Cabinet downed links between NIC and switch:

1. Validate both switch and NIC are properly configured and attempting to bring up the HSN link. For more information, see sections 3.2.2, 3.4.2.5, and 3.4.2.6 of the HPE Slingshot Troubleshooting Guide.
2. Reseat cable between NIC card and switch.
3. Reboot compute node.
4. Reset ASIC by using command `fmn-reset-switch -r -i <xname>` and reboot switch.
5. Reseat cable between NIC card and switch a second time.
6. Replace cable between NIC card and switch.
7. Replace NIC card.
8. Replace switch.

Steps to debug downed links between switches:

1. Validate both switch and NIC are properly configured and attempting to bring up the HSN link. For more information, see sections 3.2.2, 3.4.2.5, and 3.4.2.6 of the HPE Slingshot Troubleshooting Guide.
2. Reseat cable between switches.
3. Reset asic by using command `fmn-reset-switch -r -i <xname>`
4. Slot power cycle.
5. Reseat cable between switches a second time.
6. Replace cable between switches.
7. Replace local switch.
8. Replace remote switch.

Customer Support Contact:

Open a Hewlett Packard Enterprise Support Center issue and provide the .tar file(s) generated by running

```
hsn_triage_capture
```

from each switch.

Bounce Port Protocol:

Run the following command from the active Fabric Manager:

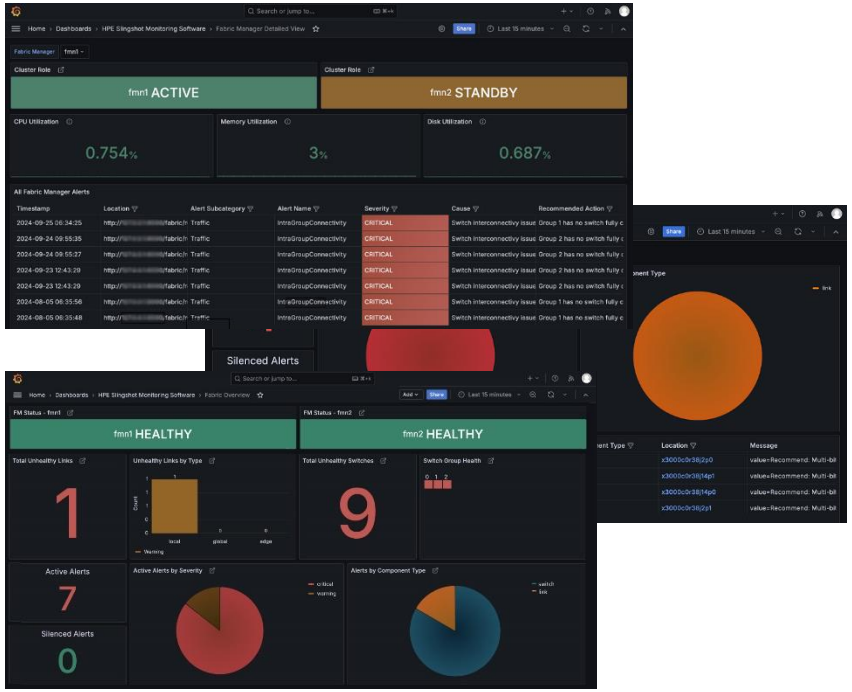
```
fmn-reset-port -n <port xname>
```

HPE Slingshot Monitoring Software: Summarizing and Synthesizing

- HPE Slingshot Monitoring Software is designed to generate real-time insights, offering a **quick and clear synthesis of the fabric's overall health**.
  - **Top-down view of fabric health** presents most critical info on homepage.
  - Drill-downs available for each section to **guide directed debugging**.
    – Links, ports, switches, and the fabric manager.

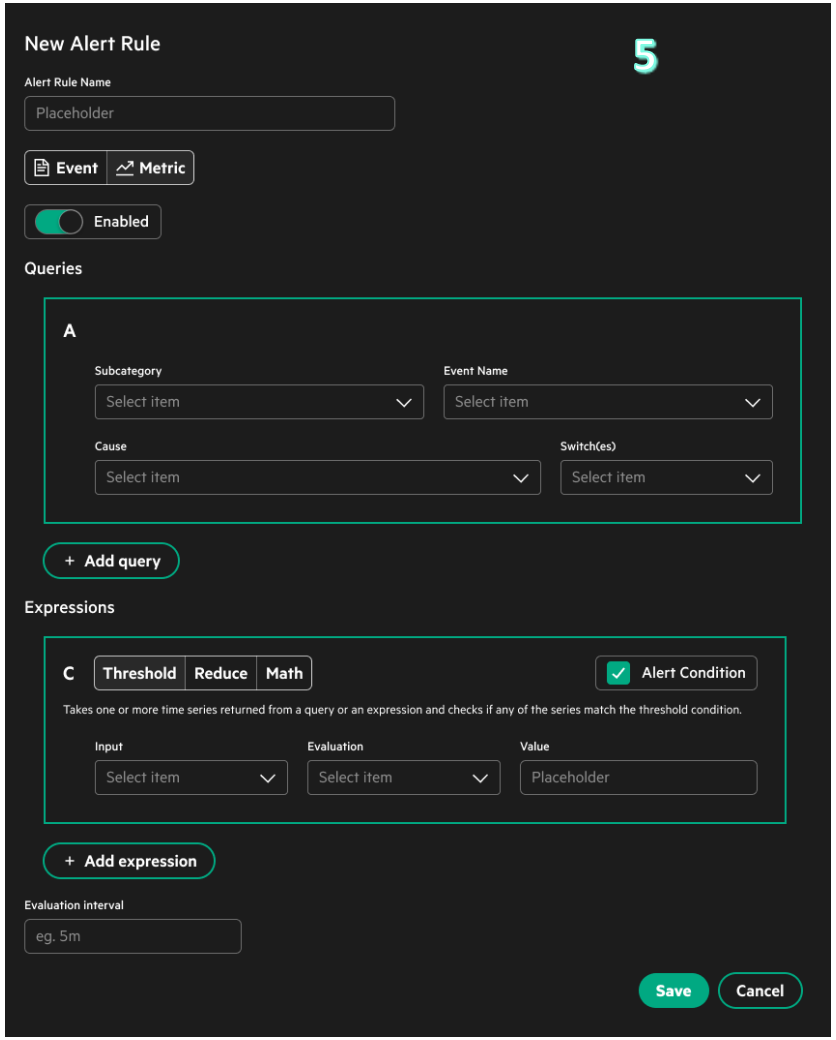
- The tool enables users to **reduce manual time taken for the root cause analysis** of the overall system and **improves overall utilization of the system** through data-driven decision-making
  - Machine down time is limited

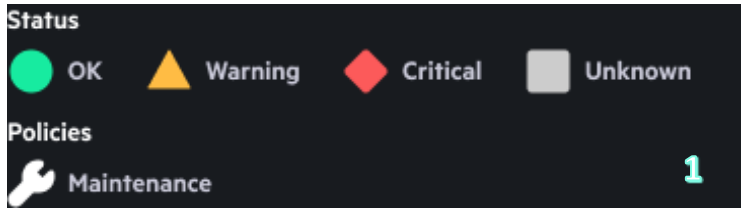
- Users are provided the capabilities to
  - Quickly **summarize and synthesize insights** in real-time
  - **Reduce manual time** taken for the root cause analysis
  - **Improve overall utilization** of their system through data-driven decision-making

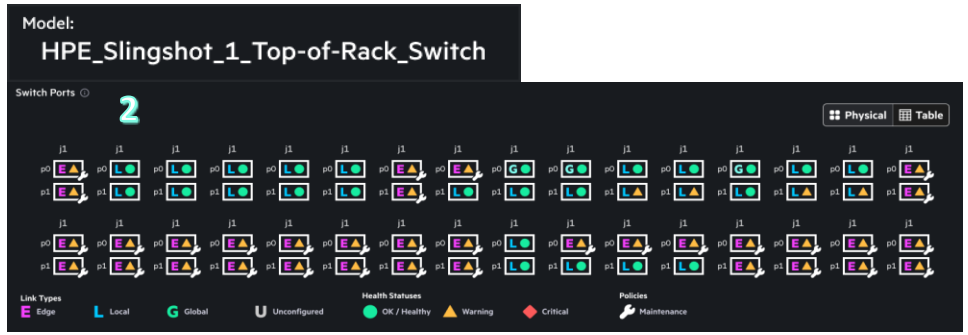


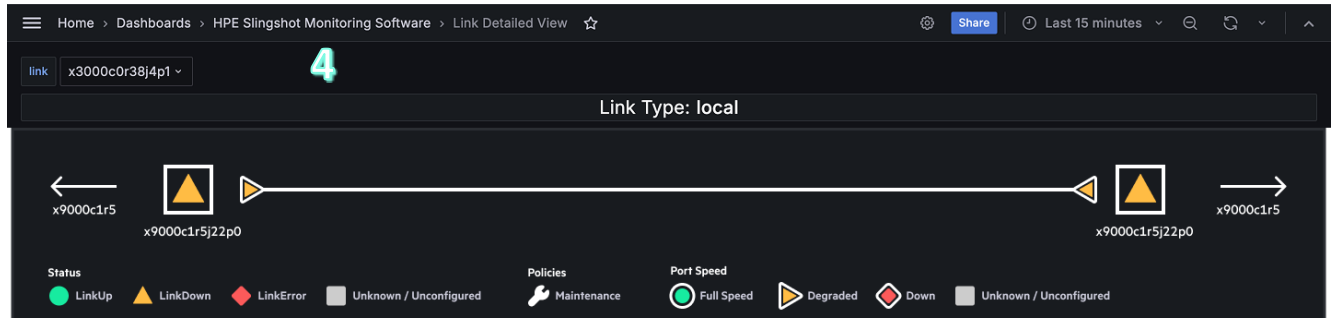
Thrive in the exascale era and beyond with HPE Slingshot Monitoring Software- HPE Delivers Today!

Slingshot Monitoring Software – Upcoming Features

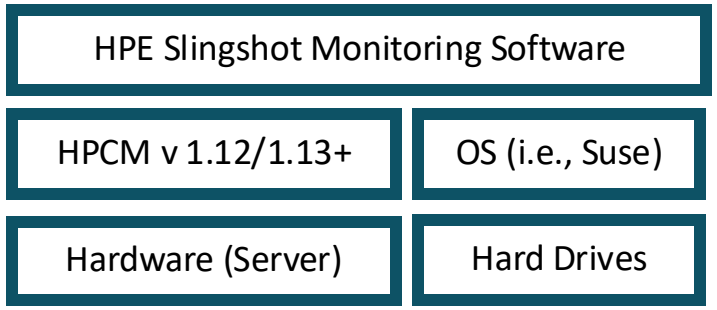








Ready to Get Started with SMS? Here's What You Need to Know.



- **HPE Slingshot Monitoring Software:**
  - SMS has no SKU, any customer with existing HPCM 1.12/1.13 and above, can enable and start using Slingshot Monitoring.

- **Please Note:**
  - SMS is available at no cost for Slingshot Customers.
  - However, to collect Slingshot telemetry - the HPE Performance Cluster Manager (HPCM) software version 1.12 or higher needs to be available.
  - For customers that don't want to use HPCM for cluster management (ex. have different Cluster Manager), or customers that do not wish to upgrade their existing HPCM, a standalone/monitoring version of HPCM can be enabled with purchase of only one HPCM license.

	HPCM Customer	Non-HPCM Customer (ex CSM or other CM)
Use Case	Customer wants to enable SS monitoring and advanced telemetry on their new/existing HPCM.	Customer does not have HPCM and requires stand-alone SS monitoring and telemetry.
What to offer	• SMS • HPCM	• SMS • Instance of HPCM 1.12/1.13+ only to collect telemetry
Dependencies	• HPCM 1.12/1.13+ (VictoriaMetrics) • Slingshot 2.2 >	• HPCM 1.12/1.13+ (VictoriaMetrics) • Slingshot 2.2 >

Licensing requirement	• 1x HPCM license
	  • SMS has no licensing
	• Linux license (Suse)
		• Only 1 HPCM license to enable monitoring
		  • SMS has no licensing
		• Linux license (Suse etc)

**HPCM 1.12/1.13+ enables VictoriaMetrics Database required by SMS*

Thank you!

Sahil Patel, *Technical Product Manager, HPE*
Contact: spatel@hpe.com