



Telemetry

Table of contents

High-Frequency (Primary) Telemetry Fields	5
Low-Frequency (Secondary) Telemetry Fields	9

UFM Telemetry allows the collection and monitoring of InfiniBand fabric port statistics, such as network bandwidth, congestion, errors, latency, and more.

UFM provides a range of telemetry capabilities:

- Real-time monitoring views
- Monitoring of multiple attributes
- Intelligent Counters for error and congestion counters
- InfiniBand port-based error counters
- InfiniBand congestion XmitWait counter-based congestion measurement
- InfiniBand port-based bandwidth data

The telemetry session panels support the following actions:

- Rearrangement via a straightforward drag-and-drop function
- Resizing by hovering over the panel's border

UFM Telemetry data is collected via UFM telemetry instances invoked during UFM startup.

Telemetry Instance	Description	REST API
High-Frequency (Primary) Telemetry Instance	<p>A default telemetry session that collects a predefined set of ~30 counters covering bandwidth, congestion, and error metrics, which UFM analyzes and reports. These counters are used for:</p> <ul style="list-style-type: none">• Default Telemetry Session - An ongoing session used by the UFM to display UFM WebUI dashboard charts information and for monitoring and analyzing ports threshold events (the session interval is 30 secs by default)• Real-Time Telemetry - allows users to define live telemetry sessions for monitoring small subsets of	<p>For Default and Real-time Telemetry: Monitoring REST API</p> <p>For Historical Telemetry: History Telemetry Sessions REST API</p>

Telemetry Instance	Description	REST API
	<p>devices or ports and a selected set of counters. For more information, refer to Telemetry - User-Defined Sessions</p> <ul style="list-style-type: none"> • Historical Telemetry - based on the primary telemetry and collects statistical data from all fabric ports and stores them in an internal UFM SQLite database (the session interval is 5 mins by default) 	
Low-Frequency (Secondary) Telemetry Instance	Operates automatically upon UFM startup, offering an extended scope of 120 counters. For a list of the Secondary Telemetry Fields, refer to Low-Frequency (Secondary) Telemetry Fields .	N/A

For direct telemetry endpoint access, which exposes the list of supported counters:

For the **High-Frequency (Primary) Telemetry Instance**, run the following command:

```
curl http://r-ufm114:9001/csv/cset/converted_enterprise
```

For the **Low-Frequency (Secondary) Telemetry Instance**, run the following command:

```
curl http://r-ufm114:9002/csv/xset/low_freq_debug
```

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.

i Note

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)

High-Frequency (Primary) 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
timestamp	
source_id	
tag	
node_guid	node GUID
port_guid	Port GUID
port_num	Port Number
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
SymbolErrorCounterExtended	This counter provides information on error bits that were not corrected by phy correction mechanisms.
LinkErrorRecoveryCounterExtended	Total number of times the Port Training state machine has successfully completed the link error recovery process.
LinkDownedCounterExtended	Perf.PortCounters

Field Name	Description
PortRcvErrorsExtended	Total number of packets containing an error that were received on the port
PortRcvRemotePhysicalErrorsExtended	Total number of packets marked with the EBP delimiter received on the port.
PortRcvSwitchRelayErrorsExtended	Total number of packets received on the port that were discarded because they could not be forwarded by the switch relay.
PortXmitDiscardsExtended	Total number of outbound packets discarded by the port because the port is down or congested.
PortXmitConstraintErrorsExtended	Total number of packets not transmitted from the switch physical port.
PortRcvConstraintErrorsExtended	Total number of packets received on the switch physical port that are discarded.
LocalLinkIntegrityErrorsExtended	The number of times that the count of local physical errors exceeded the threshold specified by LocalPhyErrors
ExcessiveBufferOverflowErrorsExtended	The number of times that OverrunErrors consecutive flow control update periods occurred, each having at least one overrun error
VL15DroppedExtended	Number of incoming VL15 packets dropped due to resource limitations (e.g., lack of buffers) in the port
PortXmitWaitExtended	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
hist[0-4]	Hist[i] give the number of FEC blocks that had RS-FEC symbols errors of value i or range of errors
infiniband_CBW	
Normalized_CBW	
NormalizedXW	
Normalized_XmitData	

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
timestamp	
source_id	
tag	
node_guid	node GUID
port_guid	Port GUID
port_num	Port Number
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
SymbolErrorCounterExtended	
LinkErrorRecoveryCounterExtended	
LinkDownedCounterExtended	
PortRcvErrorsExtended	
PortRcvRemotePhysicalErrorsExtended	
PortRcvSwitchRelayErrorsExtended	
PortXmitDiscardsExtended	
PortXmitConstraintErrorsExtended	

Field Name	Description
PortRcvConstraintErrorsExtended	
LocalLinkIntegrityErrorsExtended	
ExcessiveBufferOverflowErrorsExtended	
VL15DroppedExtended	
PortXmitWaitExtended	
hist[0-4]	Hist[i] give the number of FEC blocks that had RS-FEC symbols errors of value i or range of errors
infiniband_CBW	
Normalized_CBW	
NormalizedXW	
Normalized_XmitData	

Low-Frequency (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
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 widthsource_id
Active_FEC	Active FEC
Total_Raw_BER	Pre-FEC monitor parameters

Field Name	Description
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.
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_I_B	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
PlrRcvCodeError	The total number of rejected codewords received
PlrRcvUncorrectableCode	The number of uncorrectable codewords received

Field Name	Description
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
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	

Field Name	Description
vendor_rev	Vendor revision
cable_fw_version	
rx_power_lane_[0-7]	RX measured power
tx_power_lane_[0-7]	TX measured power
Module_Voltage	Internally measured supply voltage
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.

Field Name	Description
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.
ExcessiveBufferOverflowErrors	The number of times that OverrunErrors consecutive flow control update periods occurred, each having at least one overrun error
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.

Field Name	Description
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
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_wrf_e	Responder - number of WR flushed errors
rq_num_lle	Responder - number of local length errors
sq_num_wrf_e	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
measured_freq_[0-1]	Clock frequency measurement in last 100msec
min_freq_[0-1]	Minutes of clock frequency measured. Units of 0.1 KHz
max_freq_[0-1]	Max of clock frequency measured. Units of 0.1 KHz

Field Name	Description
max_delta_freq_[0-1]	Observed max delta frequency in window of 100msec. Units of 0.1 KHz
snr_media_lane_[0-7]	SNR value on the media lane <i>. In unit scale of 1/256 dB. The SNR value represents the electrical signal-to-noise ratio on an optical lane, and is defined as the minimum of the three individual eye SNR values.
snr_host_lane_[0-7]	SNR value on the host lane <i>. In unit scale of 1/256 dB. The SNR value represents the electrical signal-to-noise ratio on an optical lane, and is defined as the minimum of the three individual eye SNR values.
tx_cdr_lol	Bitmask for latched Tx cdr loss of lock flag per lane.
rx_cdr_lol	Bitmask for latched Rx cdr loss of lock flag per lane.
tx_los	Bitmask for latched Tx loss of signal flag per lane.
rx_los	Bitmask for latched Rx loss of signal flag per lane.
phy_received_bits	This counter provides information on the total amount of traffic (bits) received
rq_general_error	The total number of packets that were dropped since it contained errors. Reasons for this include: Dropped due to MPR mismatch.

© Copyright 2024, NVIDIA. PDF Generated on 08/14/2024