



MLNXSM (InfiniBand Subnet Manager) Utility Release Notes v5.16.0

Table of Contents

| | | |
|----------|---|-----------|
| 1 | Release Notes Update History | 3 |
| 2 | Overview | 4 |
| 2.1 | Software Download | 4 |
| 2.2 | Packages..... | 4 |
| 3 | Changes and New Features | 5 |
| 3.1 | Parameter Changes | 13 |
| 4 | Bug Fixes | 18 |
| 5 | Legal Notices and 3rd Party Licenses | 21 |

1 Release Notes Update History

| Revision | Date | Description |
|----------|-----------------|---|
| 5.16.0 | August 10, 2023 | Initial release of this Release Notes version. This version introduces Changes and New Features and Bug Fixes . |

2 Overview

MLNXSM is an InfiniBand Subnet Manager and Subnet Administrator based on OpenSM.

2.1 Software Download

Please visit [InfiniBand Management Tools](#) page.

2.2 Packages

| Package | Version |
|-----------|---------------|
| MLNX_OFED | 23.07-0.5.0.0 |
| UFM | 6.14.x |

3 Changes and New Features

| v5.16.0 | |
|--|---|
| Congestion Control | Added support for configuring CC buffer thresholds according to the switch capabilities. |
| Performance | Improved performance of routing calculation for routers. |
| Unicast Route Rebalancing | Added the option to avoid unicast route rebalancing when HBF is enabled on all SLs. |
| Port GUID | Added the option to write the destination port GUID when logging direct routed SMPs. |
| Systemd | Added support for Systemd. |
| v5.15.0 | |
| Programmable Congestion Control | Extended IBCC to support programmable congestion control. Feature control parameters: <ul style="list-style-type: none"> <code>ppcc_algo_dir</code> - Path to directory with PPCC algorithm profile files |
| Fast Recovery | Added Fast Recovery support to configure a policy on switches for reporting ports as unhealthy, and to support isolating the unhealthy ports. Feature control parameters: <ul style="list-style-type: none"> <code>fast_recovery_enabled</code> - Enable Fast Recovery feature 0 - SM will not send configuration, and will not isolate ports reported by switch (default) 1 - Disable the feature 2 - Enable the feature <code>fast_recovery_conf_file</code> - Path to fast recovery policy file |
| General | Added support for running SM with a topology specification file Feature control parameters: <ul style="list-style-type: none"> <code>topo_config_enabled</code> - Enable/disable the feature <code>topo_config_file</code> - Path to topology specification file |
| General | Added support for generating SM performance report. the report file is created in the SM logs directory with the name <code>"opensm-perflog.json"</code> . Feature control parameters: <ul style="list-style-type: none"> <code>enable_performance_logging</code> - Enable/disable the feature (enabled by default) <code>osm_perflog_dump_limit</code> - Limit performance log file (20MB by default) |
| General | Added port label for supporting NDR switches and CAs. |
| General | Added support for additional predefined port groups: <ul style="list-style-type: none"> <code>ALL_ANS</code> - All aggregation nodes <code>ALL_SWITCH_TO_SWITCH</code> - All switch ports connected to other switches <code>ALL_SWITCH_TO_CA</code> - All switch ports connected to CAs |
| General | Limited the number of simultaneous SL2VL and VLArb MADs sent per device. |
| General | Enabled dumping FLID ranges to the <code>opensm-router.dump</code> file |
| Routing | Added support for running an SM root detection algorithm when root GUIDs file is invalid. |
| Routing | Improved FLID routing calculation time. |
| v5.14.0 | |

| | |
|--|--|
| General | Updated the subnet configuration flow as follow: <ul style="list-style-type: none"> • Switch-to-switch links are updated prior to updating routing tables • ARLFT tables are set prior to setting the AR Group tables • Routing tables of switches with topology changes are updated prior to tables of switches without topology changes |
| General | Added support for multithreaded SL2VL calculation. |
| Routers | Added support for AR over routers (FLIDs). |
| DFP Routing Engine | Added support for persistent AR groups for DF+ 1 (dfp) routing engine. |
| UPDN and AR UPDN Routing Engine | Improved UPDN minhop tables calculation times. |
| v5.13.0 | |
| SL-to-VL Mapping Table | Added support for port masked optimized SLtoVLMappingTable programming. |
| DFP2 Routing Engine | Added support for PFRN on DFP2 routing engine. |
| PKey Validation Traps | Added support for suppressing multicast PKey validation traps. |
| General | Updated the default value of drop_event_subscriptions to TRUE. |
| General | Updated the default value of drop_subscr_on_report_fail to TRUE. |
| General | Enabled sending LFT/ARLFT entry for SMLID last. |
| MTU/Rate | Enabled MTU and Rate calculation for router PathRecords according to route. |
| v5.12.0 | |
| Self Healing Network with PFRN | Self healing network with PFRN is now at GA level for AR_UPDN routing engine. |
| Hash Based Forwarding (HBF) | Hash Based Forwarding (HBF) is now at GA level |
| Adaptive Routing Engine | Improved routing calculation time for AR_UPDN routing engine. |
| Adaptive Routing | A dedicated AR group ID per leaf is now assigned also when SHIELD is disabled. |
| General | Added the option to avoid initializing links marked for port resets. |
| General | Updated root_guid_file parameter description. |
| General | Removed from guid2lid vPorts with LID required set to 0. |
| DragonFly+ Topologies | Improved root detection algorithm for DragonFly+ topologies to support leaf switches without hosts. |
| v5.11.0 | |

| | |
|--|---|
| Self healing network with PFRN | <p>[Beta] PFRN is used for fast link fault recovery. If a link fails or disconnects, switches send messages to the peer switches to update the routing tables. This feature is supported only on <code>ar_updn</code> and <code>ar_ftree</code> routing engine and if all fabric switches support PFRN (NVIDIA Quantum and NVIDIA Quantum-2 switch systems only).</p> <p>Feature control parameters:</p> <ul style="list-style-type: none"> <code>pfrn_sl</code> - SL for PFRN messages (default SL 0). <code>pfrn_mask_clear_timeout</code> - Time-out since the last PFRN message received by the switch for an AR group, after which unused port masks will be cleared. The value is in multipliers of 30 seconds. (default 180) <code>pfrn_mask_force_clear_timeout</code> - Time-out since last mask clear operation, after which unused port masks are cleared by the switch. The value is in multipliers of 240 seconds. (default 720) <p>In order to disable PFRN, set <code>shield_mode</code> value to 2.</p> |
| Multiport high availability | <p>Allows SM to failover to another port in the case of SM link failure. It requires configuring more than one port GUID in the GUID parameter.</p> <p>Feature control parameter:</p> <ul style="list-style-type: none"> <code>enable_sm_port_failover</code> - Enable or disable failover (default FALSE). |
| Hash Based Forwarding (HBF) | <p>Allows selection of the switch outgoing port for statically routed packets based on the packet's parameters (ECMP like).</p> <p>With <code>dfp2</code> routing engine, non-minhop routes will be used for static routing as well as for Adaptive Routing.</p> <p>Feature control parameters:</p> <ul style="list-style-type: none"> <code>hbf_sl_mask</code> - SLs supporting HBF (default 0). <code>hbf_hash_type</code> - Hash function for HBF <ul style="list-style-type: none"> 0 - CRC (default). 1 - XOR. <code>hbf_seed_type</code> - Hash seed type: <ul style="list-style-type: none"> 0 - Seed (default). 1 - Random. <code>hbf_seed</code> - Hash seed, 32 bit number: <ul style="list-style-type: none"> 0xffffffff - Use switch GUID for seed (default). 0x0-0xffffffffe - Specific seed value. <code>hbf_hash_fields</code> - Fields of packet for hash calculation (default 0x40F00C0F). <code>hbf_weights</code> - Weights ratio between ports of different AR sub-groups: <ul style="list-style-type: none"> auto - SM/routing engine decision (default). <sg0>,<sg1>,<sg2> - User defined weights for subgroup 0 to subgroup 2. |
| SA response time | Improved SA response time for multicast join requests during routing calculation. |
| Persistent mapping | Added support for persistent mapping between AR group ID and the destination switch GUID. |
| Switch SMA response MADs | Switch SMA response MADs are now routed using PLFT0 to overcome a firmware limitation in <code>dfp2</code> . |
| SM ports table | Added SM ports table to SMDB. |
| Log message verbosity | Changed verbosity of log message when toggling ports to INFO. |
| Port state report | Added the option to report to the log when failing to update port state from ARM to ACTIVE. |
| Virtualization traps | Added details for virtualization traps to the log file. |
| Statistics dump file per SM | Enabled statistics dump file per SM port by default. |
| Asymmetric flow algorithm for trees | Enabled asymmetric flow algorithm for trees (<code>ar_updn</code> and <code>ar_ftree</code>) by default. |

| v5.10.0 | |
|--|--|
| Adaptive Timeout SL Mask | Added support for Adaptive Timeout SL mask. Feature control parameter: <ul style="list-style-type: none"> <code>adaptive_timeout_sl_mask</code> - an adaptive timeout enabled SLs mask. (Default <code>0xFFFF</code>) |
| IB Router QoS | Extended the QoS policy file to support subnet prefixes and port GIDs for inter subnet QoS. This improvement enables the definition of SL/rate/MTU/packet-life for cross-subnet paths. For further information, refer to the <code>doc/QoS_management_in_OpenSM.txt</code> document. |
| Routing Engine | Added a new root detection algorithm in UPDN and <code>ar_updn</code> routing engines. Feature control parameters: <ul style="list-style-type: none"> <code>find_roots_color_algorithm</code> - enables/disables the feature. (Default is <code>TRUE</code>) <code>max_cas_on_spine</code> - sets the maximum number of CAs on a switch to allow considering it as a spine instead of a leaf by the routing algorithm. |
| Routing Engine | Changed the default routing engine to be <code>ar_updn</code> instead of <code>minhop</code> . |
| Vendor Specific (VS) Key | Added support for Vendor Specific (VS) key. The following are the parameters related to the feature: <ul style="list-style-type: none"> <code>vs_key_enable</code> - enables VS key configuration: <ul style="list-style-type: none"> 0 - ignore 1 - disable 2 - enable <code>vs_key_lease_period</code> - the lease period used for VS keys in [sec]. <code>vs_key_ci_protect_bits</code> - the protection level for the key: <ul style="list-style-type: none"> 1 - protected 0 - unprotected (The response Key Info exposes the key). <code>vs_max_outstanding_mads</code> - the maximum number of outstanding VS MADs in the network at once. <code>key_mgr_seed</code> - used by the key manager for VS key configuration. |
| SA Response Time | Improved SA response time during routing calculation. Feature control parameter: <ul style="list-style-type: none"> <code>enable_queries_during_routing</code> - enables SA queries during routing calculation. (Default is <code>TRUE</code>) |
| Report Duplicated GUIDs | Added support for reporting duplicated GUIDs to UFM. |
| Switch Reboot | Added the option to report switch reboots to UFM. |
| Long Transaction Timeout | Enabled the option to use long transaction timeout for PI for port 0 MADs. |
| SMDB Dump File | Added subnet prefix to SMDB dump file. |
| SM Binding Port Information | Added SM binding port information to the MAD details in the timeout message, dumped to the SM log file. |
| OpenSM Start Time | Added SM start time to the SMDB dump file. |
| Dump MAD Statistics per SM Port | Added the option to enable dump MAD statistics per SM port. Feature control parameter: <ul style="list-style-type: none"> <code>osm_stats_dump_per_sm_port</code> - enables/disables the feature. (Default is <code>FALSE</code>) |
| Adaptive Routing (AR) Group IDs | Made the process of selecting AR (Adaptive Routing) group IDs deterministic in each run of the SM on the same fabric. |

| v5.9.1 | |
|---|---|
| Link Speed | Added support for NDR InfiniBand link speed in SM, |
| Configuration File Validation | Added a new command line option <code>--validate_conf_files</code> to enable SM to only validate configuration files and exit afterwards. Note: This version of the tool supports only the validation partition file part. |
| Persistent Multicast (MC) Trees | This capability enables reading MulticastForwardingTables tables upon SM startup/fail-over to ensure the new MASTER SM does not break multicast routing. To enable/disable it use the <code>get_mft_tables</code> parameter (default TRUE). |
| DragonFly+ Topologies | Added SHIELD support for dfp2 routing engine for DragonFly+ topologies. |
| SM Allowed GUIDs List | This new capability enables the user to specify the list of GUIDs allowed to run SM in the fabric. When the list is provided, the master SM will avoid handover to ports that are not specified in the list. To enable this feature use the <code>allowed_sm_guids</code> parameter. When set to "(null)", the feature disabled. |
| Limiting the Number of VLS for Long Distance Links | This new capability enables the user to set the maximum operational VL per port by a new file specified by the <code>device_configuration_file</code> parameter in the OpenSM configuration file. To provide per port configuration use the <code>device_configuration_file</code> . For more information, see <i>doc/device_configuration.md</i> . |
| Send ClientReregister after Subnet Configuration | This new capability enables the user to send ClientReregister after subnet configuration to prevent the hosts from sending SA requests to the SM before the SM is ready to respond to them. This feature can be controlled using the following parameters: <code>client_rereg_mode</code> - Control modes of sending ClientReregister. Supported values: <ul style="list-style-type: none"> • 0 - Do not send client re-registration. • 1 - Send client re-registration during LID assignment (previous default behavior). • 2 - [Default] Send client re-registration after routing and QoS configuration from link manager. The new parameter replaces the deprecated <code>no_clients_rereg</code> parameter. |
| kDOR Generalized Hypercube Engine | Added kDOR Generalized Hypercube engine. |
| General | <ul style="list-style-type: none"> • Added the option to print a summary of AR and DragonFly+ supported switches to the log • Improved performance of NR lookup by LID • Changed the verbosity of port group creation messages to be in INFO level • Added new statistics counters to opensm-statistics.dump • Added the option to consider affinity when calculating number of cores |
| v5.8.1 | |

| | |
|---|---|
| Asymmetric trees | <p>The feature is applicable to <code>ar_updn</code> and <code>ar_ftree</code> routing engines. It reduces congestion in asymmetric tree topologies with missing uplinks on leaf switches.</p> <p>To enable/disable the feature, use the <code>ar_tree_asymmetric_flow</code> parameter . The supported values are:</p> <ul style="list-style-type: none"> • 0 - Disable the feature (default). • 1 - Enable the feature using single AR subgroup. Note: Recommended for asymmetric tree topologies with 1000-2000 leaf switches. • 2 - Enable the feature using two AR subgroups. Note: Recommended for asymmetric tree topologies with less than 1000 leaf switches. |
| Selecting LID for Master SM | <p>This feature prevents SM LID changes upon fail-over.</p> <p>To set the LID for master SM, use the <code>master_sm_lid</code> parameter. The supported values are:</p> <ul style="list-style-type: none"> • 0 - Disable the feature (default). • 1-0xBFFF - LID to set to SM port when in MASTER state. |
| Root GUIDs file for Dragonfly+ Routing Engines | <p>This feature enables root GUIDs file for Dragonfly+ topology Routing Engines (<code>dfp</code> and <code>dfp2</code>).</p> <p>To set the file with GUIDs of root switches of the topology use the <code>root_guid_file</code> parameter.</p> |
| Dragonfly+ Routing Engine | <p>Added new routing engine (<code>dfp2</code>) for Dragonfly+ topologies. This engine supports Dragonfly+ topologies with any kind of tree topology islands. If the topology contains an island with more than 2 tree levels, the root GUIDs file, including the root switches of all Dragonfly+ islands should be provided.</p> <p>To add the <code>dfp2</code> new Routing Engine, use the <code>routing_engine</code> parameter.</p> |
| Maximum Operational VLs for Ca, Routers and Switches | <p>This feature enables the user to configure different <code>max_op_vls</code> for CAs, Routers and Switches.</p> <p>To set the maximum operational VLs per device type, use the following parameters:</p> <ul style="list-style-type: none"> • <code>max_op_vls_ca</code> - Maximum operational VLs for CAs. When 0, use value <code>max_op_vls</code>. (default 0) • <code>max_op_vls_rtr</code> - Maximum operational VLs for routers. When 0, use value <code>max_op_vls</code>. (default 0) • <code>max_op_vls_sw</code> - Maximum operational VLs for switches. When 0, use value <code>max_op_vls</code>. (default 0) |
| “VL packing” for Dragonfly+ and KDOR Routing Engines. | <p>Added support for “VL packing” for Dragonfly+ and KDOR routing engines. This feature reduces the maximum operational VLs for CAs to half of subnet <code>max_op_vls</code> when using <code>dfp/dfp2/kdor_hc</code> routing engines.</p> <p>To enable/disable the feature, use the <code>enable_vl_packing</code> parameter.</p> <p>The following is an example of “VL packing”:</p> <ul style="list-style-type: none"> • <code>enable_vl_packing</code> set to TRUE • <code>max_op_vls</code> set to 3 (Enable 4 VLs) • <code>max_op_vls_ca</code> set to 2 (Use 2 VLs for CAs) |
| Support SRP target on HCAs with Socket-Direct architecture/Virtual Machines. | <p>Enabled returning PortInfoRecord and NodeRecord for virtual ports and reporting virtual port capability changes To enable/disable the feature (default TRUE), use the <code>enable_virt_rec_ext</code> parameter.</p> |

| | |
|--|---|
| General | <ul style="list-style-type: none"> • Improved balancing of direct routes calculated for multi-port • Set HCA-grp in port groups parser to be an optional parameter • Added support for router alias GUIDs configuration for virtual ports • Added NDR speed port info capability bit to <code>ib_types.h</code> • Avoided sending client reregister to vport index 0 • Extended <code>rpg_byte_reset</code> to 19 bits in Congestion Control • Enabled <code>crashd</code> by default • Added report auxiliary port state changes • OpenSM now overrides the attributes of IPoIB multicast groups loaded from the SADB with broadcast group • Limited the number of PortInfo and MEPI MADs sent per device simultaneously • Configured switch AR SL mask according to the <code>ar_sl_mask</code> configuration parameter • Updated GeneralInfo device IDs list to include NVIDIA Quantum-2 and future ConnectX family devices • Updated the man page with AR routing engines • Avoided exiting OpenSM when failing to bind to auxiliary port |
| v5.7.2 | |
| General | <ul style="list-style-type: none"> • Added support for MCMR join/leave requests with default subnet prefix • Set Enabled AR SL mask on switches according to <code>ar_sl_mask</code>. |
| v5.7.1 | |
| Multiple ports | <p>Allows MLNXSM to use multiple ports for sending Subnet configuration MADs. Feature control parameters:</p> <ul style="list-style-type: none"> • <code>guid</code> - Comma separated list of MLNXSM port GUIDs. First port GUID specifies primary port which used for Subnet Management (discovery, traps) and Subnet Administration. Additional port GUIDs are used for sending subnet configuration (SMP Set MADs). Configuration file example: <code>guid 0x10001,0x10002</code> |
| Extend router selection algorithm | <p>Supports specifying hash function, seed and additional hash function arguments for router selection during path records calculation. Feature control parameters:</p> <ul style="list-style-type: none"> • <code>rtr_selection_function</code> - Hash function to be used by router selection algorithm. • Supported values - <code>crc32</code> (default). • <code>rtr_selection_seed</code> - Seed for router selection algorithm. (default 0) • <code>rtr_selection_algo_parameters</code> - Comma separated list of parameters for router selection algorithm. Supported values: <code>sgid</code>, <code>dgid</code>. (default <code>sgid</code>, <code>dgid</code>) |
| LMC for routers and number of LIDs allowed per router for inter-subnet path records | <p>Feature control parameters:</p> <ul style="list-style-type: none"> • <code>lids_per_rtr</code> - Defines number of Router LIDs to be used for inter-subnet path records. When set to 0, MLNXSM will use number of LIDs per router according to global LMC. (default 0) When set to non-zero, MLNXSM will set LMC for router ports according to the value of this parameter (minimal N such that $2^N \geq \text{lids_per_rtr}$). If global LMC is not zero, <code>lids_per_rtr</code> is ignored. When <code>lids_per_rtr</code> is set to non-zero value, <code>updn/ar_updn/chain</code> with <code>updn</code> routing engines should be used. |

| | |
|---|---|
| Congestion Control | <p>Feature control parameters:</p> <ul style="list-style-type: none"> • <code>mlnx_congestion_control</code> - Enabled/Disable Mellanox Congestion Control. Supported values: 0 - Do not configure congestion control (default). 1 - Disable congestion control on the subnet. 2 - Configure congestion control according to policy file. • <code>congestion_control_policy_file</code> - Path to congestion control policy file. <p>For additional information, please review congestion_control.md file provided with MLNXSM.</p> |
| LIDs range in Routing Chains | <p>Replaces path-bit qualifier in routing chain configuration by min-path-bit and max-path-bit qualifiers. (path-bit is still supported for backward compatibility).</p> <p>Example of usage:</p> <ul style="list-style-type: none"> • <code>min-path-bit: 1</code> • <code>max-path-bit: 3</code> |
| Controlling maximum number of MADs on wire per destination | <p>Feature control parameters:</p> <ul style="list-style-type: none"> • <code>max_wire_smpls_per_device</code> - Number of MADs on the wire per device. (default 2) |
| Configuring service keys to service name | <p>Service keys' configuring service names.</p> <p>Feature control parameters:</p> <p><code>service_name2key_map_file</code> - Path to service name to service key map file. File contains mapping from service name to service key which is specified in IPv6 format.</p> <p>For example, map service name <SERVICE NAME> service key 0::1 by adding the following line to the file: <SERVICE NAME> 0::1</p> |
| General | <ul style="list-style-type: none"> • Disabled the option to send PortInfo MADs to switch ports that did not change their state from the previous sweep. • Enabled Adaptive Routing for all SLs on switches. • Set limit to SMInfo dispatcher queue. • Improved performance of missing routes calculation for trees. • Improved performance of <code>ar_updn</code> and <code>ar_ftree</code> routing engines. • Improved performance of inter-subnet path record calculation. • Added log number of link resets by MLNXSM at the end of heavy sweep. • Disabled creating subnet LST file as default. <p>Feature control parameters:</p> <ul style="list-style-type: none"> • <code>enable_lst_file</code> - Controls dumping subnet LST file of the topology If set to <code>TRUE</code>, LST file is created after heavy sweep. (default <code>FALSE</code>) • Removed LMC support from verbosity bypass. • Enabled empty port groups file in routing chains. • Aligned index table columns in SMDB file. • UPDN LID tracking - Added the option to give precedence to exit ports leading to switch with lower total number of routes over exit ports leading to switch with less routes to the switch of the destination LID. <p>Feature control parameters:</p> <ul style="list-style-type: none"> • <code>updn_lid_tracking_prefer_total_routes</code> If set to <code>TRUE</code>, enable the feature. (default <code>FALSE</code>) • UPDN LID tracking - Improved routing algorithm to improve routing utilization and routes balancing. • UPDN LID tracking - Updated routing engine to support LMC. |

| | |
|--------------------------------------|--|
| Default Configuration Changes | <ul style="list-style-type: none"> • Changed the default values of: <ul style="list-style-type: none"> • <code>max_topologies_per_sw</code> from 1 to 4 • <code>scatter_ports</code> from 0 (disabled) to 8 • <code>log_flash</code> from FALSE to TRUE • Disabled dumping subnet LST file by default. |
|--------------------------------------|--|

3.1 Parameter Changes

| Parameter Name | Status | Type | Description |
|---|--------|---------|--|
| 5.15.0 | | | |
| <code>max_wire_smps</code> | Update | Numeric | Change default value to 32 |
| <code>max_wire_smps2</code> | Update | Numeric | Change default value to 32 |
| <code>hbf_sl_mask</code> | Update | Numeric | Change default value to 0xffff |
| <code>topo_config_enabled</code> | New | Boolean | Enable running SM with topology spec file (default is FALSE) |
| <code>topo_config_file</code> | New | Path | Path to topology specification file (default is "(null)") |
| <code>fast_recovery_enabled</code> | New | Numeric | Enable fast recovery feature (default is 0) |
| <code>fast_recovery_conf_file</code> | New | Path | Path to fast recovery file (default is "(null)") |
| <code>ppcc_algo_dir</code> | New | Path | Path to directory with PPCC algorithm profile files (default is <code>/etc/opensm/ppcc_algo_dir</code>) |
| <code>enable_performance_logging</code> | New | Boolean | Enable generating SM performance report (default is TRUE) |
| <code>osm_perflog_dump_limit</code> | New | Numeric | Limit performance log file size in megabytes (default is 20) |
| <code>max_seq_redis</code> | Update | Numeric | Change default value to 4 |
| <code>qos_policy_file</code> | Update | Path | Change default value to "(null)" |
| <code>routing_engine</code> | Update | String | Update description that dfp2 is not experimental |
| 5.14.0 | | | |
| <code>qos</code> | Update | Boolean | Changed default to TRUE |
| <code>use_optimized_slvl</code> | Update | Boolean | Changed default to TRUE |
| <code>use_optimized_port_mask_slvl</code> | Update | Boolean | Changed default to TRUE |
| <code>long_transaction_timeout</code> | Update | Numeric | Changed default from 500 milliseconds to 1000 milliseconds |
| 5.13.0 | | | |
| <code>root_guid_file</code> | Update | Path | Updated description to include all supported routing engines |

| | | | |
|-------------------------------|--------|---------|--|
| suppress_mc_pkey_traps | New | Boolean | Suppress Multicast PKey violation traps (default is TRUE) |
| drop_subscr_on_report_fail | Update | Boolean | Changed default to TRUE |
| drop_event_subscriptions | Update | Boolean | Changed default to TRUE |
| use_optimized_slvl | Update | Boolean | Updated description that parameter control wild carded optimization |
| use_optimized_port_mask_slvl | New | Boolean | Enable port masked optimized SLtoVLMMappingTable programming (default is FALSE) |
| rtr_pr_mtu | Update | Numeric | Change default to 255 (Calculate PathRecord MTU according to route) |
| rtr_pr_rate | Update | Numeric | Change default to 255 (Calculate PathRecord Rate according to route) |
| 5.12.0 | | | |
| root_guid_file | Update | Path | Updated description to include all supported routing engines. |
| 5.11.0 | | | |
| pfrn_sl | New | Number | SL for PFRN messages. Default 0 |
| pfrn_mask_clear_timeout | New | Number | Time since last PFRN for an AR group to clear unused port masks. Default 180 |
| pfrn_mask_force_clear_timeout | New | Number | Time since last mask clear, after which unused port mask is cleared by the switch. Default 720 |
| n2n_key_enable | New | Number | Enable Node-to-Node Key (management class 0xC) configuration. Default 0 (Ignore) |
| n2n_key_protect_bit | New | Number | Protection level for class 0xC. Default 1 |
| n2n_key_lease_period | New | Number | Lease period for class 0xC key. Default 60 |
| n2n_max_outstanding_mads | New | Number | Maximum number of N2N MADs in the network at once. Default 500 |
| enable_sm_port_failover | New | Boolean | Enable SM fail over to another port in case of link failure. Default is False |
| hbf_sl_mask | New | Boolean | SL mask for HBF. Default 0x0000 (Disabled) |
| hbf_hash_type | New | Number | HBF hash type. Default 0 (CRC) |
| hbf_seed_type | New | Number | HBF seed type. Default 0 (User defined seed) |
| hbf_seed | New | Number | Seed for HBF. Default 0xFFFFFFFF (Use switch GUID) |
| hbf_hash_fields | New | Number | Fields of packet for hash calculation. Default 0x40F00C0F |
| hbf_weights | New | String | Weights ratio between ports of different groups. Default 'auto' (Routing algorithm decision) |

| | | | |
|--------------------------------------|--------|---------|--|
| cache_ar_group_id | New | Boolean | Load GUID to AR group ID cache file on startup. Default TRUE |
| ar_tree_asymmetric_flow | Update | Number | Changed default to 1. |
| sm_stats_dump_per_sm_port | Update | Boolean | Changed default to TRUE, |
| 5.10.0 | | | |
| adaptive_timeout_sl_mask | New | Number | Define a adaptive timeout SL mask of the port. Default 0xFFFF |
| routing_engine | Update | String | Changed default value from (<i>null</i>) to <i>ar_updn</i> |
| find_roots_color_algorithm | New | Boolean | Find root using coloring algorithm for tree based topologies. Default is TRUE. |
| max_cas_on_spine | New | Boolean | The maximum number of CAs on a switch to allow considering it as a spine instead of a leaf by the routing algorithm. |
| hm_num_traps | Update | Number | Changed default value from 250 to 60. |
| hm_num_traps_period_secs | Update | Number | Changed default value from 60 to 90 seconds. |
| 5.9.1 | | | |
| allowed_sm_guids | New | String | Define list of allowed SM port GUIDs |
| device_configuration_file | New | String | Path to device configuration file |
| client_rereg_mode | New | Number | Control sending ClientReregister to devices |
| max_rate_enum | New | Number | Define maximal supported rate in SA records |
| gmp_traps_threads_num | New | Number | Number of threads for processing GMP traps |
| get_mft_tables | New | Boolean | Enable/Disable reading MFT tables on first master sweep |
| routing_engine | Update | String | Support <i>kdor-ghc</i> for Generalized Hypercube routing engine |
| mepi_cache_enabled | Update | Boolean | Changed default from <i>FALSE</i> to <i>TRUE</i> |
| no_clients_rereg | Update | Boolean | Deprecated by <i>client_rereg_mode</i> |
| use_original_extended_sa_rate_s_only | Update | Boolean | Deprecated by <i>max_rate_enum</i> |
| dfp_down_up_turns_mode | Update | Number | Changed default from 0 to 2 (disable down/up turns) |
| routing_threads_num | Update | Number | Changed default value from 1 to 0 |
| force_link_speed_ext | Update | Number | Support NDR speeds |
| 5.8.1 | | | |
| max_wire_smps | Update | Number | Change default from 4 to 16 |
| max_wire_smps2 | Update | Number | Change default from 4 to 16 |

| | | | |
|---------------------------------------|----------|---------|--|
| max_smpls_timeout | Update | Number | Change default from 600000 to 300000 milliseconds |
| max_msg_fifo_timeout | Update | Number | Change default from 10000 to 5000 milliseconds |
| transaction_timeout | Update | Number | Change default from 200 to 100 milliseconds |
| enable_crashd | Update | Boolean | Change default from FALSE to TRUE |
| routing_engine | Update | Text | Support <i>dfp2</i> routing engine |
| master_sm_lid | New | LID | LID for local SM when in MASTER state |
| enable_virt_rec_ext | New | Boolean | Enable PortInfoRecord/NodeRecord for virtual ports/nodes |
| ar_tree_asymmetric_flow | New | Number | AR Asymmetric trees max flow algorithm |
| max_op_vls_ca | New | Number | max_op_vl for CAs |
| max_op_vls_sw | New | Number | max_op_vl for switches |
| max_op_vls_rtr | New | Number | max_op_vl for routers |
| enable_vl_packing | New | Boolean | Enable VL packing |
| 5.7.2 | | | |
| ar_sl_mask | Existing | Number | Modified behavior: Parameter controls AR SL mask both in switches and HCAs |
| 5.7.1 | | | |
| enable_lst_file | New | Boolean | Control dumping subnet LST file of the topology |
| lids_per_rtr | New | Number | Control number of LIDs per router of inter-subnet path record |
| max_wire_smpls_per_device | New | Number | Control maximum number of MADs on wire per device |
| service_name2key_map_file | New | Path | Path to service name to service key map file |
| rtr_selection_function | New | String | Hash function to be used by router selection algorithm |
| rtr_selection_seed | New | Number | Seed for router selection algorithm |
| rtr_selection_algo_parameters | New | String | Comma separated list of parameters for router selection algorithm |
| updn_lid_tracking_prefer_total_routes | New | Boolean | Control UPDN LID tracking exit port selection criteria |
| mlnx_congestion_control | New | Number | Control Mellanox Congestion Control enablement |
| congestion_control_policy_file | New | Path | Path to Congestion Control policy file |
| guid | Update | List | Changed the type from GUID to list of commas separated GUIDs |
| scatter_ports | Update | Number | Changed default value from 0 (disabled) to 8 |

| | | | |
|-----------------------|--------|---------|--|
| log_flash | Update | Boolean | Changed default value from FALSE to TRUE |
| max_topologies_per_sw | Update | Number | Changed default value from 1 to 4 |

4 Bug Fixes

| Version | Description |
|---------|---|
| 5.16.0 | <ul style="list-style-type: none"> Fixed a race condition when parsing configuration file that can result in SM crash Fixed an issue with reconfiguring switch after removed from held-back list Fixed a redundant heavy sweep when suspecting a port is unhealthy Fixed the handling process of switches with all port marked as 'no_discover' Fixed an issue with writing 'unknown vendor' to screen |
| 5.15.0 | <ul style="list-style-type: none"> Fixed the reporting of PortInfo validation failure for rebooted HCA ports Fixed the writing of VS and CC MADs in debug verbosity Fixed an issue related to sending SL2VL MADs for unhealthy ports Fixed an issue related to the initiating heavy sweep process when reporting new unhealthy port |
| 5.14.0 | <ul style="list-style-type: none"> Fixed a crash that occurred when Incremental Multicast Routing was enabled Fixed an issue related to the enabled PLFT2 on DF+2 when the "dive-ins" are not permitted Marked ports that did not respond to the NI as unhealthy instead of entire node Fixed root detection algorithm in trees with missing links Fixed root detection algorithm in DF+ with roots without leaves |
| 5.13.0 | <ul style="list-style-type: none"> Fixed an issue related to removing ServiceRecords when the port is disconnected. Set threads affinity according to the scheduler affinity. Enabled dumping SMDB file when ucast cache feature is enabled. Fixed event reporting to untrusted subscribers. Fixed an issue related to creating service records with P_Key 0. Fixed the handling of routers marked as "unhealthy" when using fat-tree routing engine. Fixed an issue related to creating a dump files directory when it did not exist on startup. |
| 5.12.0 | <ul style="list-style-type: none"> Fixed a crash that occurred when <i>drop_subscr_on_report_fail</i> was enabled. Fixed a case that caused FRN to fail when there were isolated/heldback switches. Fixed a memory leak when changed the list of routing engines during runtime. Fixed an issue that prevented from ports to be directly activated in INIT state. Fixed an issue that prevented activating virtual ports on first time master sweep when running with <i>--once</i>. Fixed a memory leak when parsing QoS policy file with errors. Fixed an issue that prevented the incrementation of of outstanding AN2AN/VS/CC MADs when no response was expected. Added support for routers in FTREE routing engine. Fixed an issue related to AR LFT in trees that had entries with FREE state and empty group 0. |
| 5.11.0 | <ul style="list-style-type: none"> Fixed unconditional jump on uninitialized value when in <code>dfp2</code> when <code>ar_sl_mask</code> is set to 0. Fixed a case of duplicated LIDs when persistent SM LID feature is enabled. Fixed invalidating ucast cache when discovering faulty switch. Fixed a crash when detecting two ports of the same node with different port GUID but the same port number. Fixed traps 1310 and 1311 (duplicate GUIDs) type to 'security'. Fixed reporting trap 1312 to UFM. |

| | |
|----------------------|--|
| <p>5.10.0</p> | <ul style="list-style-type: none"> • Fixed a crash that occurred during a race between the LFT record get query and routing configuration. • Fixed a non-generic notices statistics counters in the dump file. • Fixed the postponing isolation and reporting process of the noisy ports. • Fixed an issue related to the selecting of the held back/isolated switches as roots for multicast trees. • Fixed an issue that caused the unresponsive links to remain in Active state. • Fixed an issue that affected the writing of invalid AN2AN links to SMDB dump file. • Fixed the IPoIB traffic loss after changing the subnet prefix and loading the MC groups from SADB upon SM restart. • Fixed the SM build on Debian with libibumad from rdma-core. • Fixed the way how port capability changes are handled during runtime. • Fixed an incorrect endianness issue in error log message 0F29. • Fixed an incorrect log message when enabling SHARP on the device. • Fixed the statistics counters race condition with SM multi port. • Fixed rewriting of the statistics file when the existing file had different header than the current. In case the previous header is different from the current, a backup of the old file is created as well as the updated statistics file. |
| <p>5.9.1</p> | <ul style="list-style-type: none"> • Fixed a crash incident when isolating the switch using: <ul style="list-style-type: none"> • the "held_back_sw_guid" file while running SM with updn/ar_updn • using GUIDs order file with a port group that includes HCAs that are connected to a held-back switch • Fixed an issue that resulted in breaking routing for virtual port LIDs upon failover/restart • Fixed an issue that caused ar_ftree to create non-credit loop free routing between IO nodes • Fixed an issue that resulted in continuation of the discovery stage during subnet configuration stage • Fixed an issue that missed getting MEPI after switch reset • Fixed multicast group leak when handling leave of SendOnlyFullMembers of multicast groups • Fixed a leak when spoofing notice 144 for virtual ports |
| <p>5.8.1</p> | <ul style="list-style-type: none"> • Enabled SA requests with default subnet prefix in GRH on subnet with non-default subnet prefix • Fixed a crash when processing virtual ports after aborted heavy sweep • Fixed a wrong direct route for GeneralInfo MADs after coming out-of-standby • Fixed a crash in UPDN LID tracking that happened when multithreading was enabled • Fixed file descriptor leakage when running with crashd • Fixed an issue that resulted in setting default pkey at index 0 on invalid partitions.conf • Fixed an issue that prevented setting ar_sl_mask on hosts when running with armgr plugin • Freed alias GUIDs resources when deleting virtual port object • Fixed checking 2x link width capability • Enabled handling MCMemberRecord request with default subnet prefix on subnet with non-default subnet prefix |

5.7.1

- Fixed memory overflow upon virtual ports removal from the Subnet when using Adaptive Routing.
- Fixed handling ‘;’ and ‘:’ in nodes names in port groups policy file parser.
- Fixed missing routes-to-routers after recovery the routing engine in Dragonfly+.
- Fixed port_search_order usage when LMC is enabled.
- Fixed SA LinkRecords and MultipathRecords LMC support.
- Fixed partition checking for LinkRecord and PortInfoRecord queries.
- Fixed dedicated groups calculation for switches with ANs when FRN enabled.
- Fixed router support in port groups.
- Fixed an issue that prevented SADB dumping when updating service records.

5 Legal Notices and 3rd Party Licenses

| Product | Version | Legal Notices and 3rd Party Licenses |
|---------|---------|---|
| OpenSM | 5.16.0 | <ul style="list-style-type: none">• 3rd Party Unify Notice• 3rd Party Notice |

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

© 2024 NVIDIA Corporation & affiliates. All Rights Reserved.

