
NVIDIA
IB NETWORKING PRACTICE – RUN/DEBUG/MAINTAIN

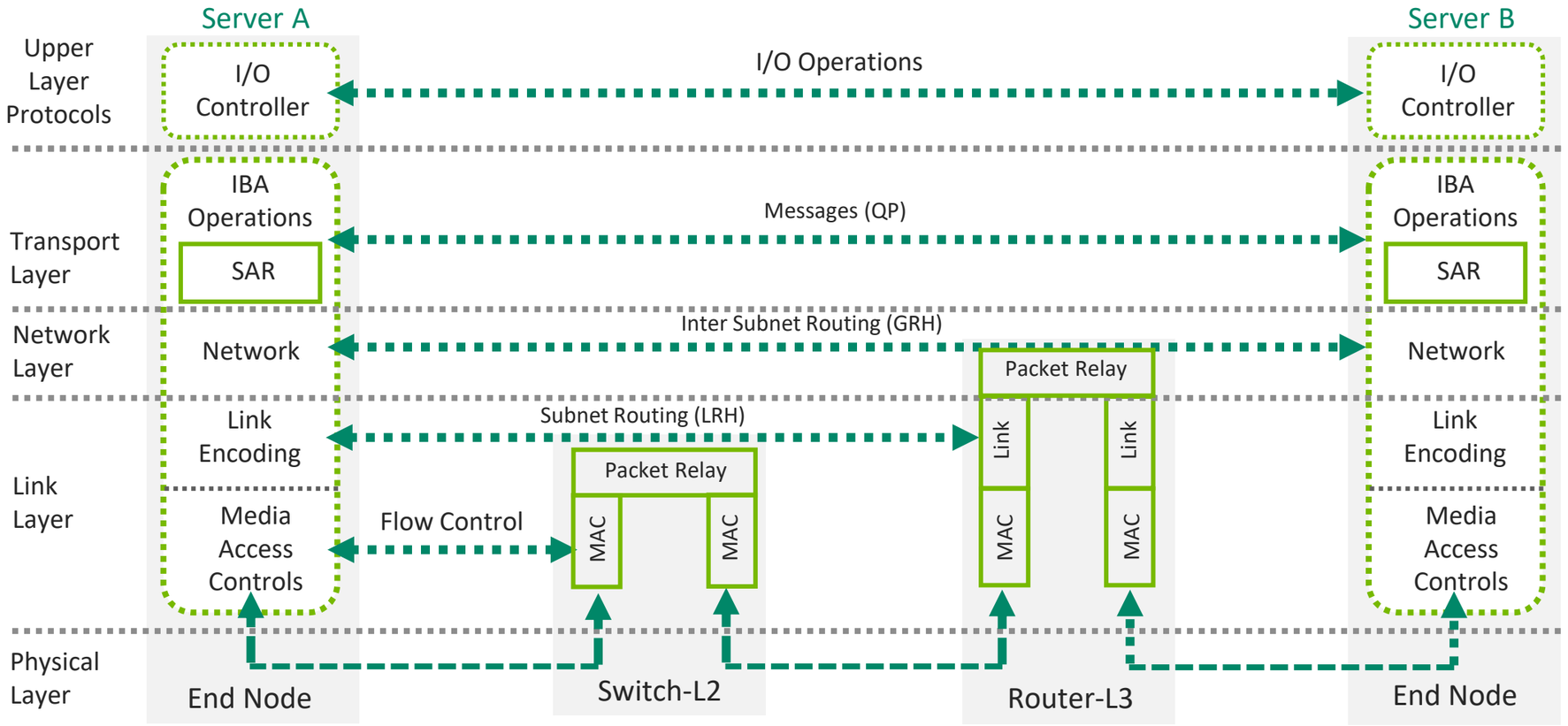
AGENDA

- IB Basic Brief
- OpenSM
- Ibdiagnet
- IPoIB
- UFM
- Q&A
- Reference

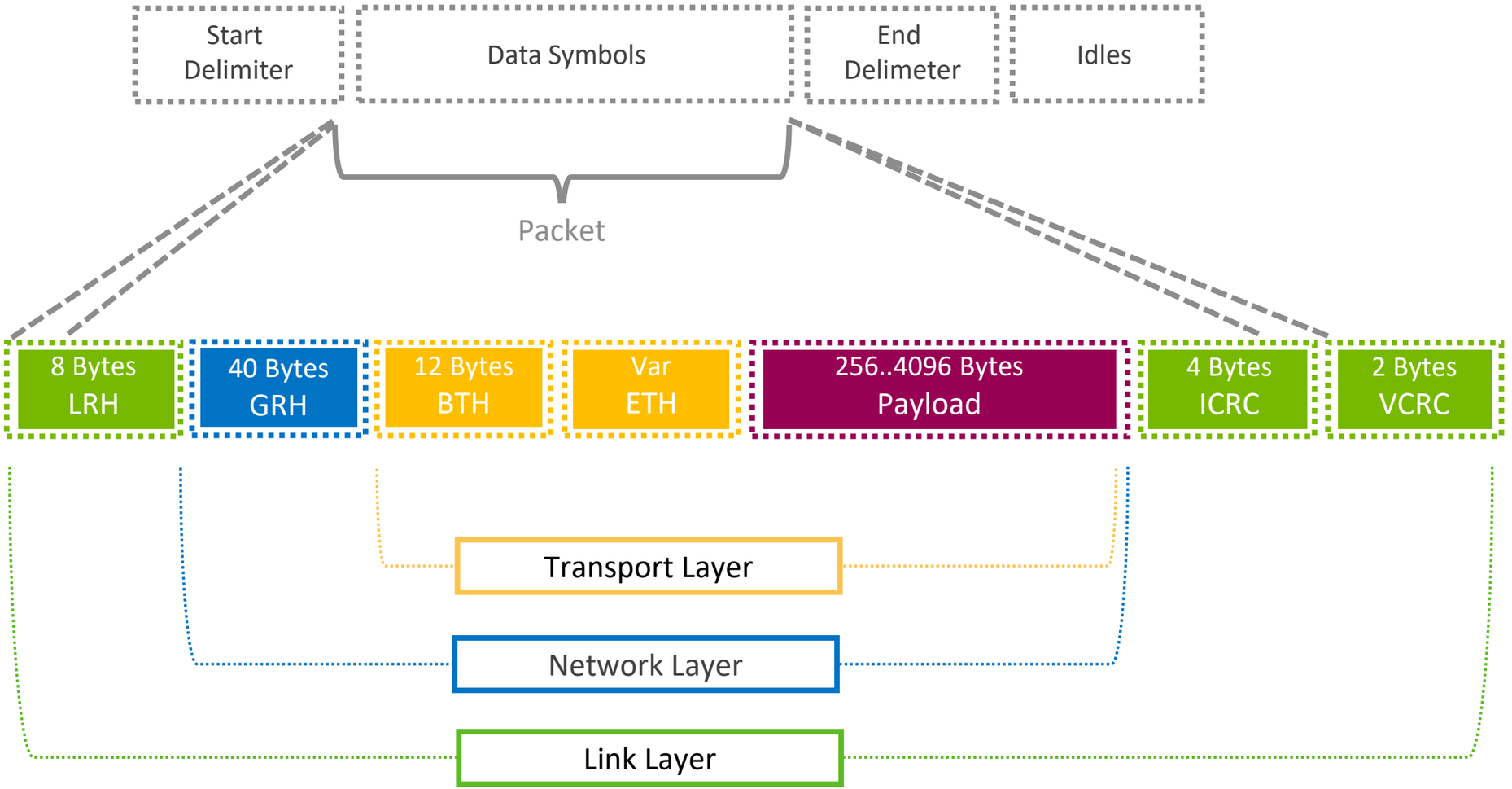
IB BASIC

INFINIBAND NETWORK STACK

- InfiniBand uses a multi-layer processing stack to transfer data between nodes
- Provides CPU offloads functions
- Offers greater adaptability through a variety of services and protocols



DATA PACKET STRUCTURE



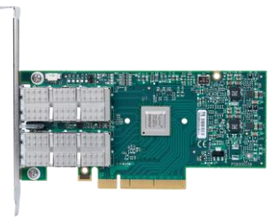
LAYER 2 ADDRESSING - LIDS

- LID (Local Identifier) is a 16-bit Layer 2 address
- LIDs are assigned by the Subnet Manager when a node becomes active
- HCAs are assigned with a LID per port
- Switches
  - 1 IC switches are assigned with a single LID
  - Director switches are assigned with a LID per switch module in the chassis

All ports share the Node LID

Director switch

Switch 1
Node LID



LID Port 1

LID Port 2

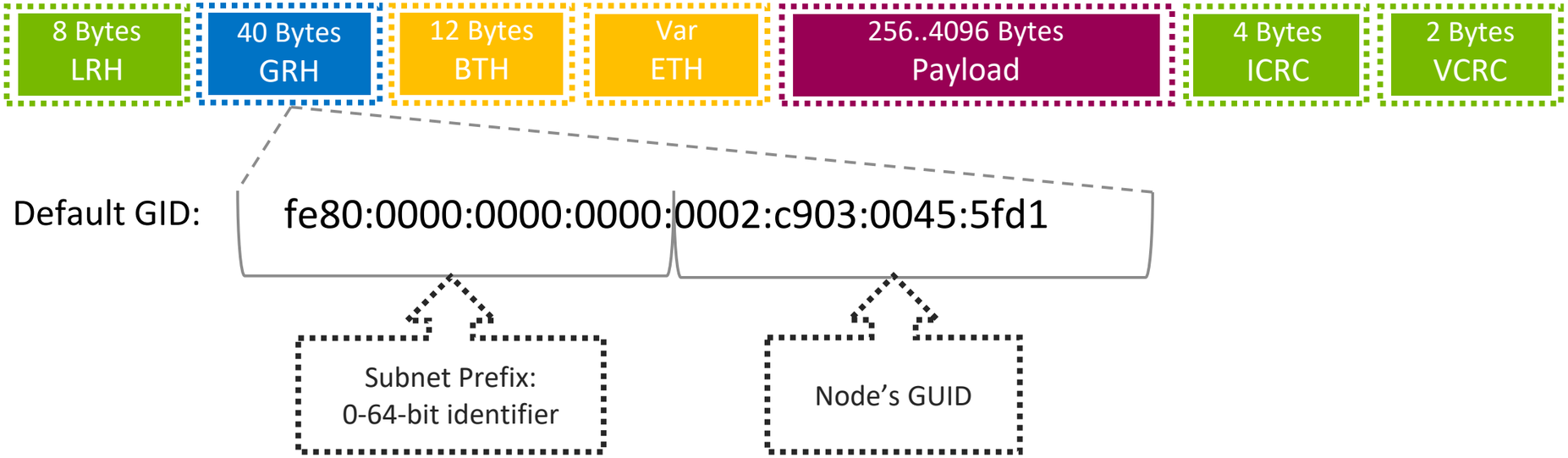
An HCA has a System image GUID a Node GUID and a Port GUID

The LID that was assigned by the SM to the HCA port

```
[root@mtlacad03 ~]# ibstat
CA 'mlx5_0'
        CA type: MT4115
        Number of ports: 1
        Firmware version: 12.21.2010
        Hardware version: 0
        Node GUID: 0x7cfe9003005d7b2e
        System image GUID: 0x7cfe9003005d7b2e
        Port 1:
                State: Active
                Physical state: LinkUp
                Rate: 56
                Base lid: 5
                LMC: 0
                SM lid: 25
                Capability mask: 0x2651e84a
                Port GUID: 0x7cfe9003005d7b2e
                Link layer: InfiniBand
```

LAYER 3 ADDRESSING - GID

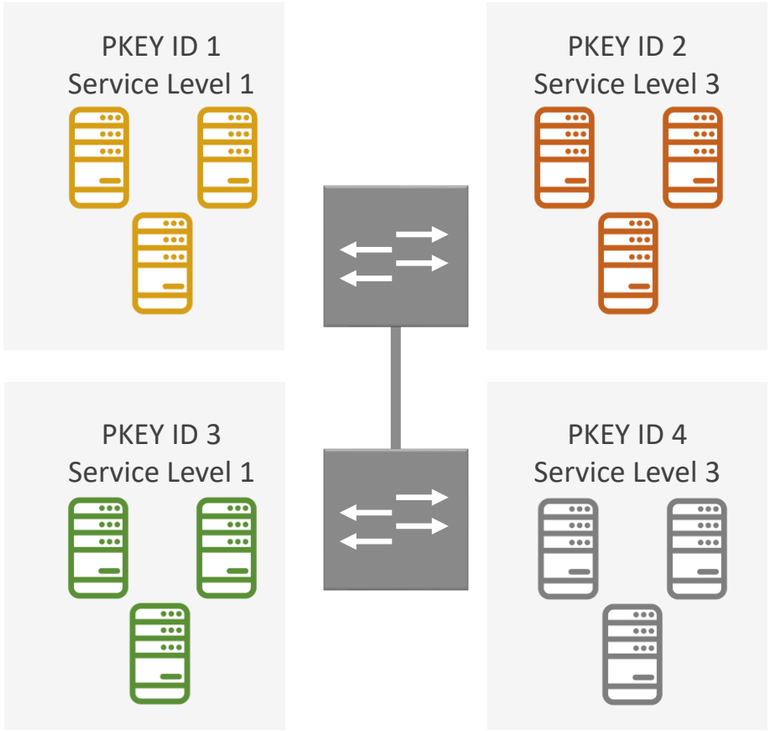
- **GID (Global Identifier)** is a 128-bit field in the Global Routing Header (GRH) used to identify a single end port or a multicast group
- GIDs are globally unique (across multiple subnets)
- GID's structure:
  - Based on a subnet prefix (a 64-bit identifier) combined with the Port GUID
  - IPv6 type address
- Each HCA port is automatically assigned with a default GID that can be used only in the local subnet



FABRIC SEGMENTATION - PARTITIONS

- A partition describes a set of end nodes within the fabric that may communicate
- When is port is assigned to a partition its **membership type** can be set to:
  - Full membership – can communicate with any other node in the partition
  - Limited membership – can communicate only with nodes in full membership in the partition
- Nodes may be members of multiple partitions at once
- PKEY – a field in BTH header used for membership in a partition

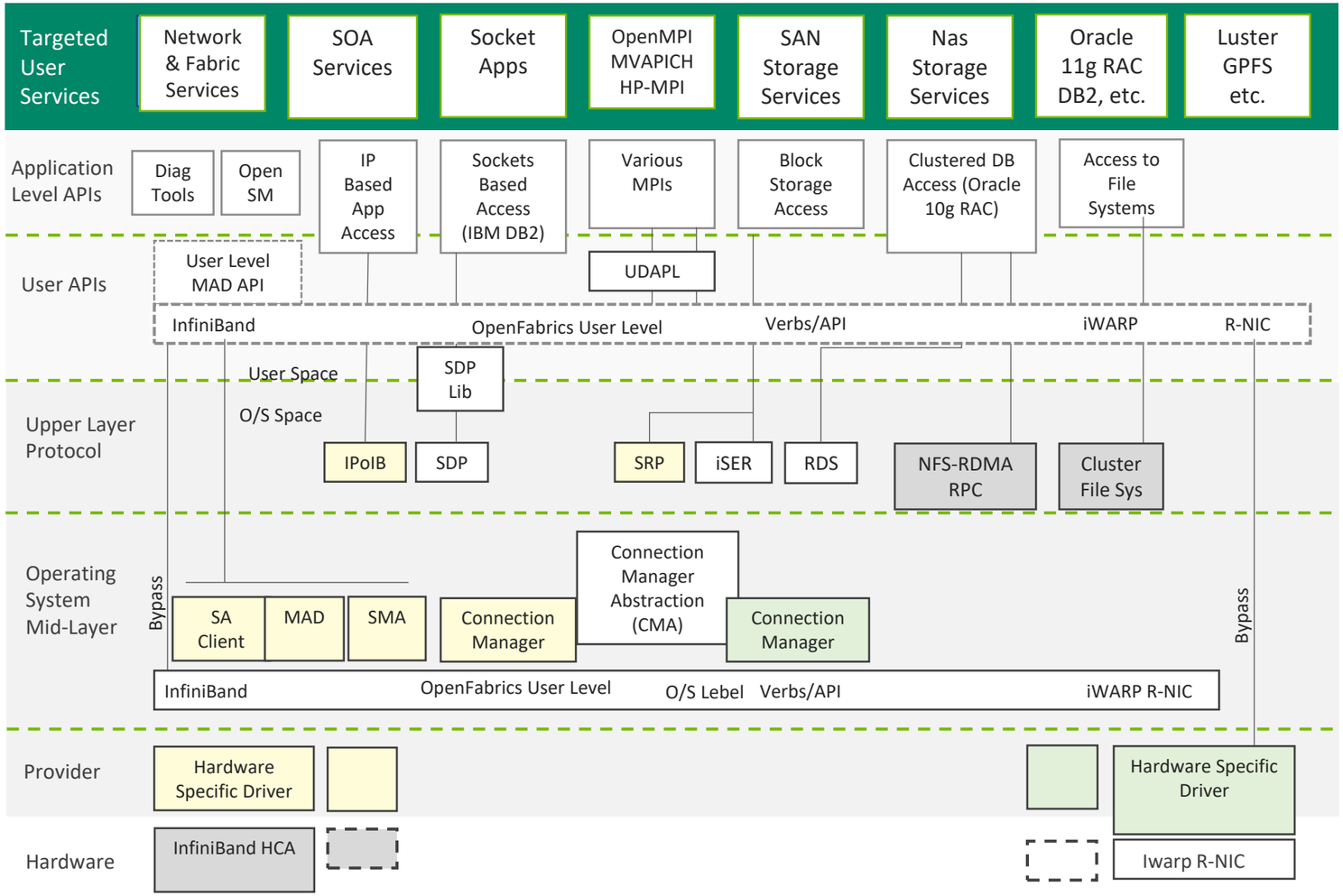
```
[root@mtbc-r740-06 ~]# smpquery pkeys 8
   0: 0xffff 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000
   8: 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000
  16: 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000
  24: 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000
  32: 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000
  40: 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000
  48: 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000
  56: 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000
  64: 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000
  72: 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000
  80: 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000
  88: 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000
  96: 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000
 104: 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000
 112: 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000
 120: 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000 0x0000
128 pkeys capacity for this port
```



? For which purposes should you use partitioning in the fabric?

OPENFABRICS ENTERPRISE DISTRIBUTION (OFED™)

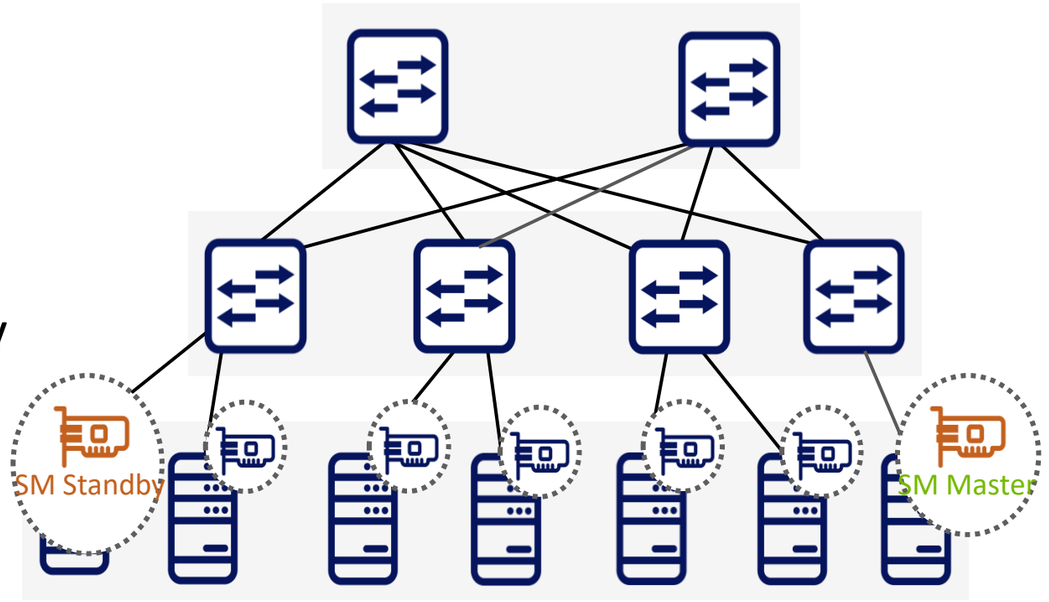
- OFED is the **open-source** softwarestack for **RDMA and kernel bypass** applications
- OFED provides high performance computing sites and enterprise data centers with flexibility and investment protection.
- The OFED architecture defines the means of interaction and creates a common language between different protocols, drivers and kernels in order to establish RDMA connectivity.



OPENSM

SUBNET MANAGER - RULES & ROLES

- **The Subnet Manager (SM) is responsible for:**
  - Discovering the topology
  - Assigning LIDs to devices
  - Calculating and programming switch forwarding tables
  - Managing all the elements in the fabric
  - Monitoring changes in subnet
- **SM can be implemented on any node** in the fabric: server, switch, or specialized device
- **Only one master SM is allowed**. The master SM that is configured as the Master SM, other SMs are in Standby mode
- **SM runs from a managed switch : < 2000 nodes**
- **SM runs from a server or from UFM: > 2000 nodes**



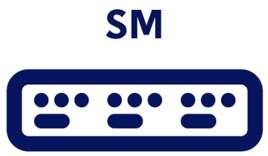
OPENSM CONFIGURATION FILE

- If file does not exist, create it with the command: `opensm -c /etc/opensm/opensm.conf`
  - This flag will export the OpenSM configurations to a file
  - When working with Mellanox OFED the location is: **/etc/opensm/opensm.conf**
  - When working with UFM, the location is: **/opt/ufm/files/conf/opensm/opensm.conf**

```
[root@mtlacad07 ~]# more /etc/opensm/opensm.conf
#
# DEVICE ATTRIBUTES OPTIONS
#
# The port GUID on which the OpenSM is running
guid 0xe41d2d030067b161

# M_Key value sent to all ports qualifying all Set(PortInfo)
m_key 0x0000000000000000

# The lease period used for the M_Key on this subnet in [sec]
m_key_lease_period 0
```



SM FAILOVER

- **SM failover -** When the master SM fails, one of the standbys is elected as the new master
- **SM handover -** If the failed master is up, it is re-elected as the master, also known as **double-failover**
  - Whenever a new master is elected it rediscovers the fabric from scratch
- Should avoid SM handovers for stability, decrease overhead and improve overall performance.
- To avoid double failover, '**master_sm_priority**' parameter need to be configured to value 15 on both SMs.
  - The parameter 'master_sm_priority' is used by OpenSM , only when it becomes the master

```
sm_priority 14

master_sm_priority 15
```

```
sm_priority 13

master_sm_priority 15
```

- UFM running in management mode, SM priority is **automatically configured as the** highest
  - Other SMs (servers or switches) should be manually configured to lower priorities

OPENSM MESSAGE LOOK UP

- Useful Log Messages

Message Type	Description
ERR	Error messages
tables	Routing engine
Log_trap_info	Switch/Host trap messages
-I removed	Removed ports
-I Timeout	MAD timeout

```
grep <Message Type> /var/log/opensm.log
grep <Message Type> /opt/ufm/log/opensm.log
```

OPENSM MESSAGE EXAMPLE

- Verify that OpenSM has successfully activated the subnet:

  If OpenSM was able to setup the subnet correctly, its log file should include the message "SUBNET UP".

```
[root@mtlacad02 ~]# cat /var/log/opensm.log | grep 'SUBNET UP'

Jul 20 14:10:27 564635 [C1598700] 0x02 -> SUBNET UP
```

Use `locate` **`opensm.log`** to find it location (or **`find`** command)

- Verify routing engine convergence:

```
[root@mtlacad02 ~]# grep table /var/log/opensm.log
Jul 20 14:10:27 556075 [C1598700] 0x02 -> osm_ucast_mgr_process: minhop tables configured
on all switches
[root@mtlacad02 ~]# cat /var/log/opensm.log | grep updn
Jul 20 14:10:27 555993 [C1598700] 0x02 -> updn_lid_matrices: disabling UPDN algorithm, no
root nodes were found
```

```
[root@mtlacad02 ~]# cat /var/log/opensm.log | grep table
Jul 20 14:38:07 714802 [ECD47700] 0x02 -> osm_ucast_mgr_process: updn tables configured on
all switches
```

OPENSM LOGS CONFIGURATION

- The SM log file size can be changed
- Choose how often a new SM log file will be created: daily, weekly (default), monthly.
- The SM log file will reach its maximum log size, or it will obey the rotational periodically order.

1) Modify the OpenSM log maximum file size:

```
vi /etc/opensm/opensm.conf
vi /opt/ufm/files/conf/opensm/opensm.conf
```

Look for log_max_size and change the size

2) Modify the OpenSM log frequency rotation:

```
vi /etc/logrotate.d/opensm
vi /opt/ufm/files/conf/logrotate.conf
```

Look for daily, weekly, monthly and change accordingly

SM INFORMATION

- Find the SM:

```
sminfo
sm lid 573 sm guid 0x2c90300fe2ed1, activity count
26181972 priority 15 state 3 SMINFO_MASTER
```

- Query node description:

```
smpquery nd 573
Node Description:.....................sm2 HCA-1
```

- Make sure the routing algorithm converged as expected:

```
grep table opensm.log
Feb 19 10:42:25 488716 [321B1700] 0x02 -> osm_ucast_mgr_process: updn tables configured on all
switches
```

SA QUERY

- You can query nodes using the **saquery** tool.

```
[root@mtlacad02 ~]# saquery 8
NodeRecord dump:
                lid.....................8
                reserved................0x0
                base_version............0x1
                class_version...........0x1
                node_type...............Channel Adapter
                num_ports...............2
                sys_guid................0xf45214030033d083
                node_guid...............0xf45214030033d080
                port_guid...............0xf45214030033d081
                partition_cap...........0x80
                device_id...............0x1003
                revision................0x1
                port_num................1
                vendor_id...............0x2C9
                NodeDescription.........mtlacad01 HCA-1
```

```
[root@mtbc-r740-06 log]# saquery 5
NodeRecord dump:
                lid.....................5
                reserved................0x0
                base_version............0x1
                class_version...........0x1
                node_type...............Switch
                num_ports...............81
                sys_guid................0xb8599f0300d52326
                node_guid...............0xb8599f0300d52326
                port_guid...............0xb8599f0300d52326
                partition_cap...........0x8
                device_id...............0xD2F0
                revision................0xA0
                port_num................0
                vendor_id...............0x2C9
                NodeDescription.........BD-BDDWDC4-SPG-
383-QM8790-SL20N3
```

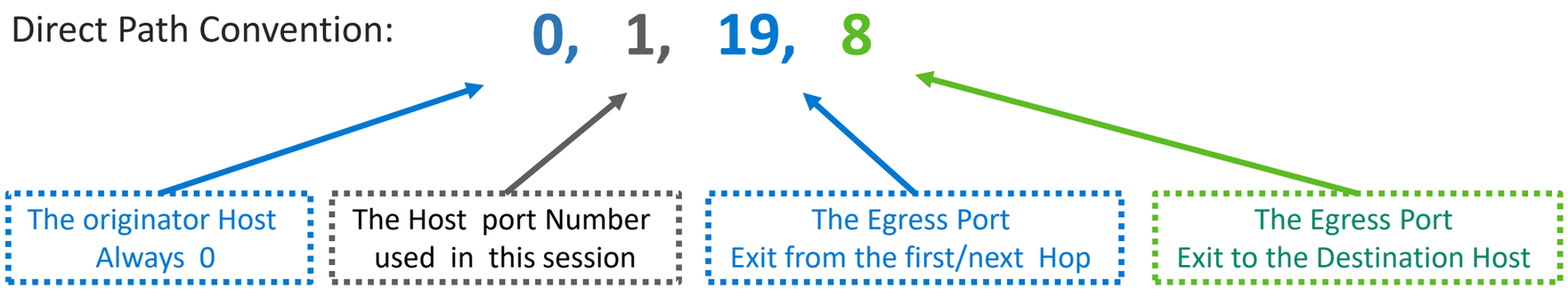
ERR 3113 – MAD COMPLETED IN ERROR

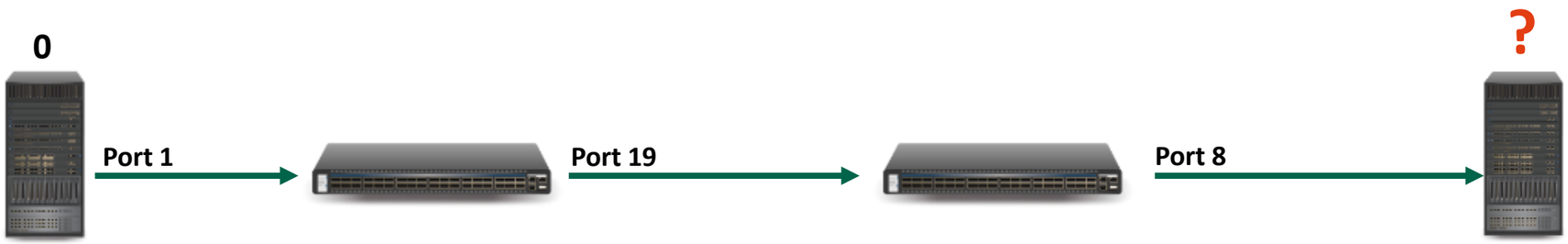
```
Feb 17 10:56:41 275109 [309AE700] 0x01 -> Received SMP on a 4 hop path: Initial path =
0,1,20,17,9, Return path  = 0,0,0,0,0
Feb 17 10:56:41 275116 [309AE700] 0x01 -> sm_mad_ctrl_send_err_cb: ERR 3113: MAD completed
in error (IB_TIMEOUT): SubnGet(PortInfo), attr_mod 0x0, TID 0x11ebace2
Feb 17 10:56:41 275133 [309AE700] 0x01 -> log_send_error: ERR 5411: DR SMP Send completed
with error (IB_TIMEOUT) – dropping
```

- A MAD sent to the destination was not completed
- Use smpquery to demand link state and node description:
  - `smpquery pi -D 0,1,20,17,9`
  - `smpquery nd -D 0,1,20,17,9`

ERR 3315 - UNKNOWN REMOTE SIDE FOR NODE

```
Feb 17 10:56:40 433133 [321B1700] 0x01 -> state_mgr_light_sweep_start: ERR 3315: Unknown remote side for
node 0x0002c90300619ab0 (MF0;io01-ib1:SXX536/L18/U1) port 8.
Adding to light sweep sampling list
Feb 17 10:56:40 433140 [321B1700] 0x01 -> Directed Path Dump of 3 hop path: Path = 0,1,19,18
```





IDENTIFYING THE IMPACTED “LOST“ REMOTE PORT

1. Trace the last component responses the MADS:

```
[root@v-sup11 log]# smpquery -D ND 0,1
Node Description:.Mellanox 4036E # v-sup-sw02-4036E
[root@v-sup11 log]# smpquery -D ND 0,1,18
Node Description:.Mellanox sLB-4018    Line 1  Chip 1 4700 #4700-B9B8
[root@v-sup11 log]# smpquery -D ND 0,1,18,21
ibwarn: [22153] mad_rpc: _do_madrpc failed; dport (DR path slid 65535
smpquery: iberror: failed: operation ND: node info query failed
```

2. We have detected the last working component in that route, we can identify the component GUID, as we already know its node description
   - For that purpose you can sort `ibhosts` or `ibswitches` or `ibnetdiscover` results

```
[root@v-sup11 log]# ibswitches
Switch  : 0x0008f105006000de ports 36 "Mellanox sLB-4018    Line 18
Switch  : 0x0008f10500650c4a ports 36 "Mellanox sLB-4018    Line 4
Switch  : 0x0008f10500650390 ports 36 "Mellanox sLB-4018    Line 1
```

3. Now you will use `iblinkinfo -S <switch GUID>:` in order to check the **status of the port** disconnecting the “chain“ in the heart of this log

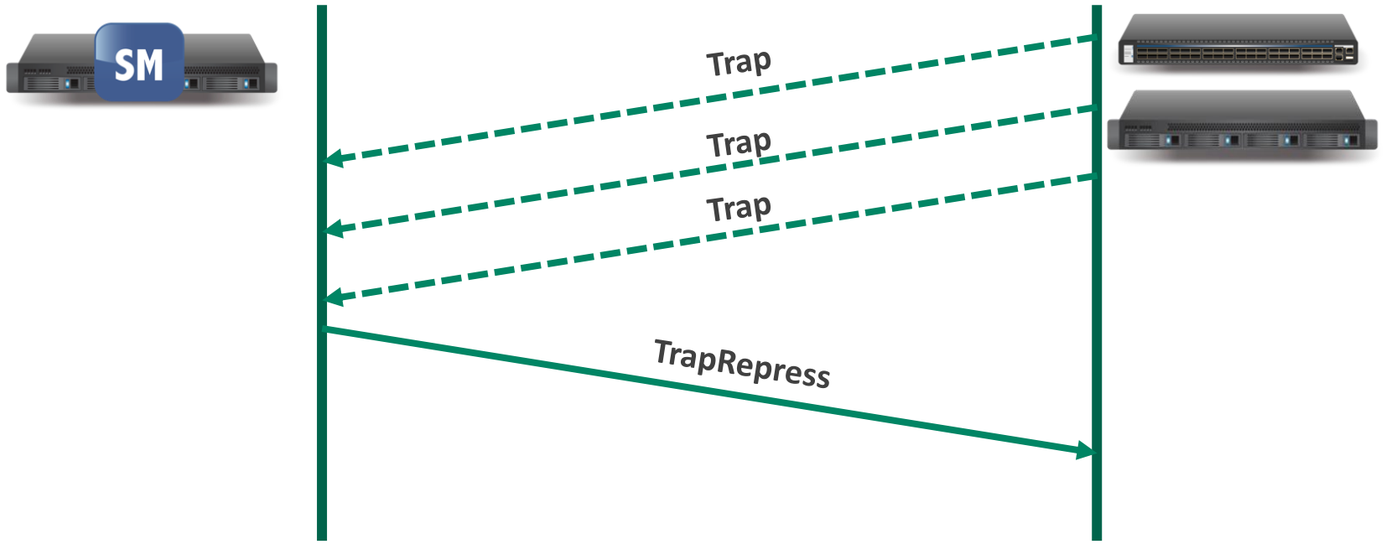
```
[root@v-sup11 log]# iblinkinfo -S 0x0008f10500650390
Switch 0x0008f10500650390 Mellanox sLB-4018    Line 1  Chip 1 4700 #4700-B9B8:

 12    11[  ] ==( 4X          10.0 Gbps Active/  LinkUp)==>      10    1[
 )
 12    12[  ] ==( 4X          10.0 Gbps Active/  LinkUp)==>      10    2[

12    20[ 2] ==( 4X          10.0 Gbps Active/  LinkUp)==>      26    18[  ]
12    21[ 3] ==(                Down/ Polling)==>             [  ] "" ( )
```

MANAGEMENT TRAPS

- Message sent by a management agent to SM when certain asynchronous management events occur (such as protocol violations)
  - The trap message has its data in the form of a notice attribute
- IB devices send traps to the SM when certain events occur
  - Event for example: a switch detects a state change on one of its ports (i.e., a topology change and/or device joining or leaving)
  - The node stops sending the trap when it receives a **TrapRepress**



TRAPS 64-67 - INFORMATIONAL TRAPS FROM THE SM

In and Out of Service Traps

- Trap 64: <GIDADDR> is now in service
- Trap 65: <GIDADDR> is out of service
  - The events by traps 64-65 indicate that an endport's reachability from a subscribing endport has changed
  - Endport A is reachable from subscriber endport B if there is a PathRecord with source B and destination A Endport
  - reachability may change as a result of:
    - changes to restrictions on access through partitioning
    - endports commencing or ceasing participation on the subnet

```
Dec 30 22:29:00 438533 [BE933700] 0x02 -> log_notice: Reporting Generic Notice type:3 num:64 (GID in service) from LID:8 GID:fe80::b859:9f03:f9:90be
Dec 30 22:31:12 085637 [BE933700] 0x02 -> log_notice: Reporting Generic Notice type:3 num:65 (GID out of service) from LID:8 GID:fe80::b859:9f03:f9:90be
```

Multicast Group Create/Delete Traps

- Trap 66: New multicast group with multicast address <GIDADDR> is now created
- Trap 67: Multicast group with multicast address <GIDADDR> is now deleted

```
Jan 04 00:08:20 353525 [E1B79700] 0x02 -> log_notice: Reporting Generic Notice type:3 num:66 (New mcast group created) from LID:8 GID:ff12:601b:ffff::2
Jan 04 00:09:43 343709 [E1B79700] 0x02 -> log_notice: Reporting Generic Notice type:3 num:67 (Mcast group deleted) from LID:8 GID:ff12:601b:ffff::2
```

TRAP 128

Num 128: Link state change

- Always generated by a switch due to a port status change
- Used mainly as a trigger for the SM to start a sweep

```
log_trap_info: Received Generic Notice type:1 num:128 (Link state change) Producer:2 (Switch) from
LID:58 TID:0x000000000000010e
```

- Finding the link that failed:

```
[root@mtbc-r740-06 log]# for i in {1..36};do echo Port:$i;perfquery 3 $i | grep
LinkDownedCounter;done
Port:8
LinkDownedCounter:..............53
```

* Remember, link down isn't necessarily an error, it will also appear when a server is rebooted

TRAP 129

```
Feb 21 08:53:04 342099 [369BA700] 0x01 -> log_trap_info: Received Generic Notice type:1 num:129
(Local Link integrity threshold reached) Producer:2 (Switch) from LID:376 Port 3
TID:0x0000000000000095
```

perfquery

```
perfquery 376 3
    # Port counters: Lid 376 port 3 (CapMask: 0x1300)
    PortSelect:......................3
    CounterSelect:..................0x0000
    SymbolErrorCounter:.............0
    LinkErrorRecoveryCounter:........174
    LinkDownedCounter:..............1
    PortRcvErrors:..................7311
    PortRcvConstraintErrors:.........0
    CounterSelect2:.................0x00
    LocalLinkIntegrityErrors:........12
    ExcessiveBufferOverrunErrors:....0
```

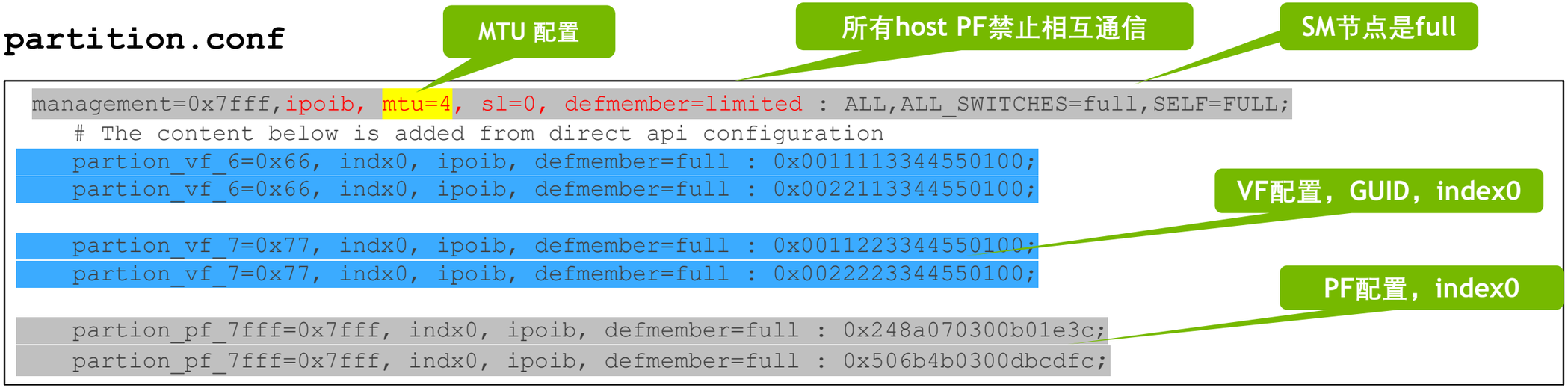
OPENSM CONFIG

opensm.conf (sriov)

```
routing_engine updn,minhop
part_enforce off
virt_enabled 2
use_ucast_cache TRUE
sweep_every_hup_signal TRUE
```

```
opensm -B -F /etc/opensm/opensm.conf -P /etc/opensm/partition.conf -v 1
```

partition.conf



IBDIAGNET

IB COMMON COMMANDS

Command	Description	Examples
ethtool -i ib0	驱动和FW版本	
ibstat/ibstatus	ib端口状态	
ibdev2netdev	RDMA和ib端口的对应关系	
Ibhost / ibnodes / ibswitches	显示网络中所有的网卡交换机	
sminfo	显示网络中激活SM信息	
ibnetdiscover	查询网络连接信息	`ibnetdiscover -p`
iblinkinfo	查询网络连接信息	`iblinkinfo -S 0xb8599f0300d52326`
saquery	网络相关信息查询	`saquery 8 -K` `saquery -m`

smpquery	节点相关信息查询	`smpquery pi -D 0,1,79`
		`smpquery nd -D 0,1,79`
		`smpquery pkeys 8`

perfquery	查询端口统计信息	`perfquery -x <lid> <port>`
ibtracert <lid-from> <lid-to>	查询两个节点间的路径信息	`ibtracert 14 8`
ibdiagnet	Ib 网卡诊断	`ibdiagnet -lw 4x -P all=1 --pm_pause_time 1200 --get_cable_info`

ibportstate	查询和配置端口状态	`ibportstate 5 79`
		`ibportstate 5 79 reset`
		`ibportstate 3 1 disable/enable`

ibdump	IB报文抓包	`Ibdump -i mlx5_0`
ib_write/read/send_bw/lat	IB点对点性能测试	`Ib_write_lat -d mlx5_0 / Ib_write_lat -d mlx5_0 <ip>`

IBDIAGNET

- An integrated IB fabric diagnostics command line tool
- Scans the fabric using directed/lid routed packets
- Extracts the available information regarding
  - Connectivity
  - Device status
- Checks for errors in the following scopes:
  - Ports – counters thresholds, port state
  - Nodes – firmware versions, LID assignments
  - Links – links speed and width
  - Fabric - topology matching, Subnet Manager, routing
- Errors are reported to screen and saved in a log file

TESTS AND FEATURES

- Fabric discovery
- Duplicated GUIDs detection
- Duplicated node names
- Links operational state
- LIDs check:
  - No duplicated LIDs
  - No zero LIDs assigned
- SM check:
  - There is one master Subnet Manager
  - The master Subnet Manager is the correct one
- Port counters check:
  - No overflowed error counters are found
  - Modified threshold can be given for each counter
- BER test:
  - Bit Error Rate calculation per time
- Firmware check:
  - Verifies FW generations and validity for cluster components
- Speed/Width checks:
  - Verifying that actual active speed and width matches its maximum supported capability

USING IBDIAGNET

Options for Advanced Cluster Analysis & optional flags

- ibdiagnet -h

```
[root@ib-cert-sv01 ~]# ibdiagnet -h
NAME
    ibdiagnet
SYNOPSYS
    Main
        [-i|--device <dev-name>] [-p|--port <port-num>]
        [-g|--guid <GUID in hex>] [--vlr <file>]
        [-r|--routing] [-u|--fat_tree] [-o|--output_path <directory>]
        [--skip <stage>] [--skip_plugin <library name>]
        [--pc] [-P|--counter <<PM>=<value>>]
        [--pm_pause_time <seconds>] [--ber_test]
        [--ber_use_data] [--ber_thresh <value>]
        [--extended_speeds <dev-type>] [--pm_per_lane]
        [--ls <2.5|5|10|14|25|FDR10>] [--lw <1x|4x|8x|12x>]
        [-w|--write_topo_file <file name>]
        [-t|--topo_file <file>] [--out_ibnl_dir <directory>]
        [--screen_num_errs <num>] [--smp_window <num>]
        [--gmp_window <num>] [--max_hops <max-hops>]
        [-V|--version] [-h|--help] [-H|--deep_help]
```

UNDERSTANDING IBDIAGNET REPORT

Report Summary and Messages Severity Level

- I = Informative
- W = Warning
- E = Error

```
Summary
-I- Stage                          Warnings   Errors     Comment
-I- Discovery                      0          0
-I- Lids Check                     0          0
-I- Links Check                    0          0
-I- Subnet Manager                 0          0
-I- Port Counters                  4          0
-I- Nodes Information              1          5
-I- Speed / Width checks           0          1
-I- Partition Keys                 0          0
-I- Alias GUIDs                    0          1

-I- You can find detailed errors/warnings in: /var/tmp/ibdiagnet2/ibdiagnet2.log


-I- ibdiagnet database file   : /var/tmp/ibdiagnet2/ibdiagnet2.db_csv
-I- LST file                  : /var/tmp/ibdiagnet2/ibdiagnet2.lst
-I- Subnet Manager file       : /var/tmp/ibdiagnet2/ibdiagnet2.sm
-I- Ports Counters file       : /var/tmp/ibdiagnet2/ibdiagnet2.pm
-I- Nodes Information file    : /var/tmp/ibdiagnet2/ibdiagnet2.nodes_info
-I- Partition keys file       : /var/tmp/ibdiagnet2/ibdiagnet2.pkey
-I- Alias guids file          : /var/tmp/ibdiagnet2/ibdiagnet2.aguid
```

DISCOVERED COMPONENTS

Discovery

- -I- Discovering ... 13 nodes (5 Switches & 8 CA-s) discovered.
- -E- FW Check finished with errors

```
-E- FW Check finished with errors
-W- mtlacad08/U2 - Node with Devid:4115(0x1013),PSID:MT_2190110032 has FW version 12.24.1000 while
the latest FW version for the same Devid/PSID on this fabric is 12.26.1040
-W- mtlacad04/U2 - Node with Devid:4115(0x1013),PSID:MT_2190110032 has FW version 12.24.1000 while
the latest FW version for the same Devid/PSID on this fabric is 12.26.1040
```

NODES INFORMATION & SYNCHRONIZED FW

Nodes Information

```
Nodes Information
-I- Retrieving ... 11/11 nodes (4/4 Switches & 7/7 CA-s) retrieved.
-W- Nodes Info retrieving finished with errors
-W- 6036B19GW/GW - The firmware of this device does not support general info capability
-E- FW Check finished with errors
-E- 6036A17/U1 - Node has wrong FW version 9.2.4340. Maximum available FW version for this device in the fabric is 9.2.7300
-E- S0002c903004b6883/N0002c903004b6880 - The firmware of this device returned invalid general info data
-E- ib-cert-sv02/U1 - The firmware of this device returned invalid general info data
-E- S0002c903004b6d53/U1 - The firmware of this device returned invalid general info data
-E- ib-cert-sv03/U1 - The firmware of this device returned invalid general info data
```

```
-E- FW Check finished with errors
-W- mtlacad08/U2 - Node with Devid:4115(0x1013),PSID:MT_2190110032 has FW version 12.24.1000 while the latest FW
version for the same Devid/PSID on this fabric is 12.26.1040
-W- mtlacad04/U2 - Node with Devid:4115(0x1013),PSID:MT_2190110032 has FW version 12.24.1000 while the latest FW version for
the same Devid/PSID on this fabric is 12.26.1040
```

PORT COUNTERS

```
-I- Going to sleep for 1 seconds until next counters sample


-I- Ports counters retrieving (second time) finished successfully


-E- Ports counters value Check finished with errors
-E- lid=0x00e0 dev=51000 ibsw01-1/S01/U1/P31
        Performance Monitor counter         : Value
        max_retransmission_rate             : 65535      (overflow)
-E- lid=0x0121 dev=51000 ibsw01-1/L17/U1/P8
        Performance Monitor counter         : Value
        max_retransmission_rate             : 2964       (threshold=500)


-E- Ports counters Difference Check (during run) finished with errors
```

SPEED / WIDTH CHECKS

- The local IB port that is used to connect to the fabric is specified by one the following:
  - `-lw <1x|4x|8x|12x> - specifies the link width`
  - `-ls <2.5|5|10|14|25|50> - specifies the link speed`

```
Speed / Width checks
-I- Link Speed Check (Compare to supported link speed)
-E- Links Speed Check finished with errors
-E- Link: 6036B19GW/U1/P6<-->S0002c903004b6883/N0002c903004b6880/P1 - Unexpected actual link speed 2.5
(enable_speed1="2.5 or 5 or 10", enable_speed2="2.5 or 5 or 10" therefore final speed should be 10)

-I- Link Width Check (Compare to supported link width)
-I- Links Width Check finished successfully
```

DEEPER SEARCH FOR WITHIN LOG FILE

- **cat /var/tmp/ibdiagnet2/ibdiagnet2.log**

```
-I- Link Speed Check (Compare to supported link speed)
-E- Links Speed Check finished with errors
-E- Link: IBLF10/U1/P5<-->mtlacad07/U2/P1 - Unexpected actual link speed 14
(enable_speed1="2.5 or 5 or 10 or 14 or 25 or FDR10", enable_speed2="2.5 or 5 or 10 or 14 or
25 or 50" therefore final speed should be 25)
```

- [root@mtlacad01 ~]# **ibswitches**

```
Switch  : 0x7cfe9003009a0550 ports 36 "MF0;IBLF10:MSB7700/U1" enhanced port 0 lid 11

ibportstate -C mlx5_1 -L 11 5
LinkState:.....................Active
PhysLinkState:.................LinkUp
LinkSpeedExtEnabled:...........14.0625 Gbps or 25.78125 Gbps or undefined (7)
LinkSpeedExtActive:............14.0625 Gbps
ibwarn: [21824] validate_extended_speed: Peer ports operating at active extended speed 1
rather than 2 (25.78125 Gbps)
```

ROUTING COMPONENTS COLLECTION

```
Routing

-I- EXT switch info retrieving finished successfully


-I- Adaptive Routing is enabled on 6 switches.
-I- AR data retrieving finished successfully

-I- Retrieving ... 23/23 nodes (6/6 Switches & 17/17 CA-s)
retrieved.
-I- Unicast FDBS Info retrieving finished successfully
```

FABRIC DEBUG

1. **`ibdiagnet -pc`** first clear all counters
2. Wait for a certain time interval (between 30-60~ mins). Running MPI traffic on all nodes will also help
3. Check for errors that exceed the allowed threshold during this X time: **`ibdiagnet -ls 10 -lw 4x -P all=1`**
4. Fix problematic links (reseat or swap cables, replace switch ports or HCAs, etc.)
5. Go back to step 1 until fabric displays no further errors

- **`ibdiagnet -ls 25 -lw 4x -P all=1 --pm_pause_time 30`**
  - Check information provide from all counters and display each one of them crossing threshold of 1

```
Port Counters
-I- Retrieving PMClassPortInfo ... 11/11 nodes (4/4 Switches & 7/7 CA-s) retrieved.
-I- Retrieving ... 11/11 nodes (4/4 Switches & 7/7 CA-s) retrieved.
-W- Ports counters retrieving finished with errors
-W- 4036E-v-sup-sw01/IPR - This device does not support xmit wait counter capability
-W- 4036E-v-sup-sw01/IPR - This device does not support extended port counters capability
-W- 6036B19GW/GW - This device does not support xmit wait counter capability
-W- 6036B19GW/GW - This device does not support extended port counters capability

-I- Going to sleep for 600 seconds until next counters sample
-I- Time left to sleep ... 574 seconds.
```

BER TESTING*

- **--ber_test** – perform BER test
- **--ber_use_data** – use actual data to calculate BER for each port
- **--ber_thresh <value>** – specify the BER threshold value
  - The reciprocal number of the BER should be provided. For example:
    For a BER of 10^-12, use **--ber_thresh 1000000000000**
  - If threshold given value marked as zero , all fabric ports BER values will appear in the log file
  - 10^-12 would be the default value if no value is specified
- The time between the two samples is set by the **--pm_pause_time** option

```
ibdiagnet --ber_test --ber_thresh 100000000000000 (10^-12)
```

*** Supported in QDR and FDR generations**

```
Port Counters
-I- Retrieving PMClassPortInfo ... 11/11 nodes (4/4 Switches & 7/7 CA-s)
retrieved.
-I- Retrieving ... 11/11 nodes (4/4 Switches & 7/7 CA-s) retrieved.
-W- Ports counters retrieving finished with errors
-W- 4036E-v-sup-sw01/IPR - This device does not support xmit wait counter
capability
-W- 4036E-v-sup-sw01/IPR - This device does not support extended port counters
capability
-W- 6036B19GW/GW - This device does not support xmit wait counter capability
-W- 6036B19GW/GW - This device does not support extended port counters capability
-I- Going to sleep for 1 seconds until next counters sample
-I- Time left to sleep ... 1 seconds.
-I- Retrieving ... 11/11 nodes (4/4 Switches & 7/7 CA-s) retrieved.
-I- Ports counters retrieving (second time) finished successfully
-I- Ports counters value Check finished successfully
-I- Ports counters Difference Check (during run) finished successfully
-I- BER Check finished successfully
```

EFFECTIVE BER CHECK

```
ibdiagnet --ber_test --ber_thresh 100000000000000 --ber_use_data --pm_pause_time 30
```

```
-I- Effective BER Check finished successfully
-E- Effective BER Check 2 finished with errors
-W- S1c34da030047c144/N1c34da030047c144/P10/1 - BER exceeds threshold - BER value = 2.000000e-13 / threshold = 1.000000e-13
-W- S1c34da030047c244/N1c34da030047c244/P8/2 - BER exceeds threshold - BER value = 5.000000e-13 / threshold = 1.000000e-13
```

CREATE A TOPOLOGY FILE

- **ibdiagnet –w <top_file>**

```
[root@ib-cert-sv01 ~]# ibdiagnet -w oded.top
----------
Load Plugins from:
/usr/share/ibdiagnet2.1.1/plugins/
(You can specify more paths to be looked in with
"IBDIAGNET_PLUGINS_PATH" env variable)
```

```
-I- ibdiagnet database file   : /var/tmp/ibdiagnet2/ibdiagnet2.db_csv
-I- LST file                  : /var/tmp/ibdiagnet2/ibdiagnet2.lst
-I- Topology file             : oded.top
-I- Subnet Manager file       : /var/tmp/ibdiagnet2/ibdiagnet2.sm
-I- Ports Counters file       : /var/tmp/ibdiagnet2/ibdiagnet2.pm
-I- Nodes Information file    :
/var/tmp/ibdiagnet2/ibdiagnet2.nodes_info
-I- Partition keys file       : /var/tmp/ibdiagnet2/ibdiagnet2.pkey
-I- Alias guids file          : /var/tmp/ibdiagnet2/ibdiagnet2.aguid
```

```
MSB7700 IBLF09
   P1 -4x-25G-> HCA_2 mtlacad01 U2/P1
   P11 -4x-25G-> MSB7700 IBSP07 P11
   P12 -4x-25G-> MSB7700 IBSP08 P12
   P2 -4x-25G-> HCA_2 mtlacad02 U2/P1
   P5 -4x-25G-> HCA_2 mtlacad03 U2/P1
   P6 -4x-25G-> HCA_2 mtlacad04 U2/P1
   P7 -4x-25G-> HCA_2 mtlacad09 U2/P1
   P8 -4x-25G-> HCA_2 mtlacad10 U2/P1

MSB7700 IBLF10
   P1 -4x-25G-> HCA_2 mtlacad05 U2/P1
   P11 -4x-25G-> MSB7700 IBSP08 P11
   P12 -4x-25G-> MSB7700 IBSP07 P12
   P2 -4x-25G-> HCA_2 mtlacad06 U2/P1
   P5 -4x-25G-> HCA_2 mtlacad07 U2/P1
   P6 -4x-25G-> HCA_2 mtlacad08 U2/P1
   P7 -4x-25G-> HCA_2 mtlacad11 U2/P1
   P8 -4x-25G-> HCA_2 mtlacad12 U2/P1
```

ANALYZING FABRIC CONGESTION

Using PM Counters

Create a congestion map using PM (Performance Monitor) counters.

- Use ibsiganet2 ibdiagnet2.db_csv output file:
  - Step 1: Copy the “PM_INFO” data from the file to an Excel sheet
  - Step 2: Calculate the congestion index = **XmitWait** / **XmitPkt**
  - Step 3: Complete data and analyze results
- Congestion index Definition:
  Normalized XmitWait = ΔXmitWait / ΔXmitPackets
  Ports with congestion index >= 10 , should be treated as congested
- See relevant example in the next slide

```
[root@mtlacad01 ~]# perfquery

LocalLinkIntegrityErrors:........0
ExcessiveBufferOverrunErrors:....0
QP1Dropped:......................0
VL15Dropped:.....................0
PortXmitData:....................7064064
PortRcvData:.....................6380064
PortXmitPkts:....................98112866
PortRcvPkts:.....................88612
PortXmitWait:....................25888
```

PM COUNTERS AND CONGESTION

example

Port 1 Example:

- port_xmit_wait=0x07888c91 -> 126,389,393
- port_xmit_pkts=0x0485e2ae -> 75,883,182
- port_xmit_wait/port_xmit_pkts=126,389,393/75,833,182=1.66
- Congestion_index=1.66

Port 2 Example:

- port_xmit_wait=0xffffffff -> 4,294,967,295
- port_xmit_pkts=0x043e779e -> 71,202,718
- port_xmit_wait/port_xmit_pkts= 4,294,967,295/ 71,202,718=60.32
- Congestion_index=60.32

```
[root@mtlacad01 ~]# perfquery
# Port counters: Lid 20 port 1 (CapMask: 0x5A00)
PortSelect:......................1
CounterSelect:...................0x0000
SymbolErrorCounter:..............0
LocalLinkIntegrityErrors:........0
ExcessiveBufferOverrunErrors:....0
QP1Dropped:......................0
VL15Dropped:.....................0
PortXmitData:....................7064064
PortRcvData:.....................6380064
PortXmitPkts:....................98112866
PortRcvPkts:.....................88612
PortXmitWait:....................25888
```

Ports with congestion index >= 10 , should be treated as congested

http://www.binaryhexconverter.com/hex-to-decimal-converter

HAVING EXTRA UNKNOWN CABLE EDGE?

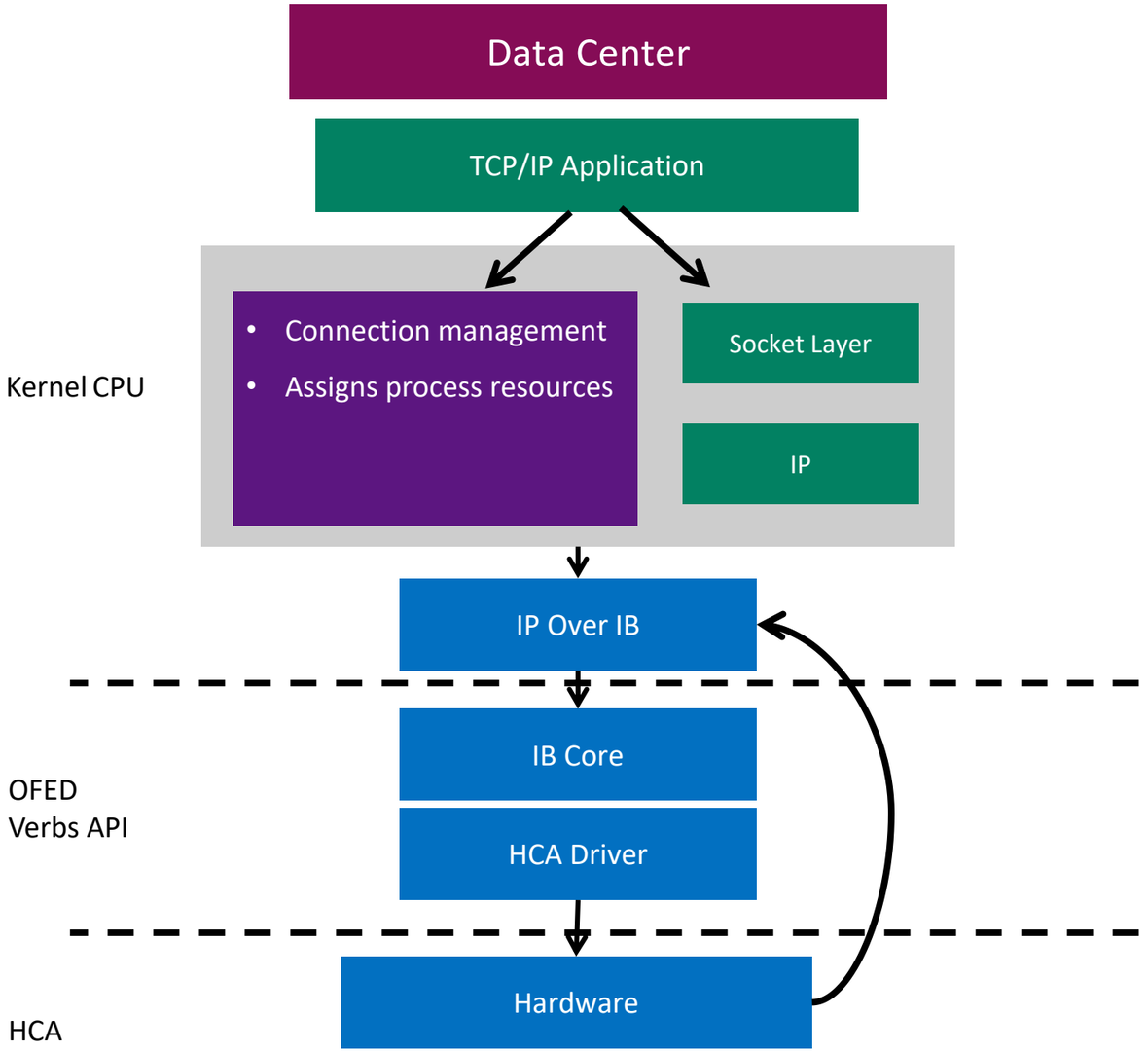
```
ibdiagnet --get_cable_info
```

```
-I- Cables Information                    : /var/tmp/ibdiagnet2/ibdiagnet2.cables

----------------------------------------------------------
Port=1 Lid=0x0003 GUID=0x1c34da030060cd61 Port Name=mtlacad10/U2/P1
----------------------------------------------------------
Vendor: Mellanox
OUI: 0x2c9
PN: MCP1600-E002
SN: MT1504VS00108
Rev: A2
Length: 2m
Type: Copper cable- unequalized
SupportedSpeed: SDR/DDR/QDR/FDR/EDR
Temperature: N/A
```

IPoIB

IPoIB PACKET FLOW

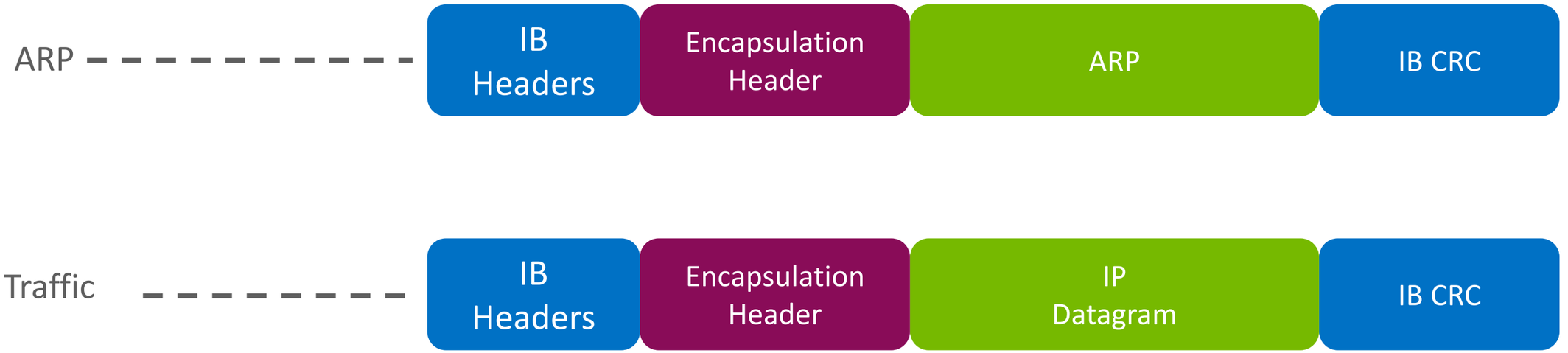


IP OVER IB (IPOIB)

- IPoIB allows to run TCP/IP over your InfiniBand network, enables to run non-InfiniBand aware applications
- IPoIB uses IB as “layer 2” for IP.
- IPoIB supports:
  - Unreliable Datagram (UD) service for UDP applications
  - Reliable Connections (RC) service for TCP applications
  - IPv4, IPv6, ARP and DHCP
  - Multicast

IPOIB PACKET FORMAT

- Before sending messages using IPoIB, an **A**ddress **R**esolution **P**rotocol (ARP) message is sent to discover the "physical address" of the host
- ARP is used to resolve an IP address into a "physical address"
- An IPoIB ARP message is sent over a specific Multicast Group created for that purpose

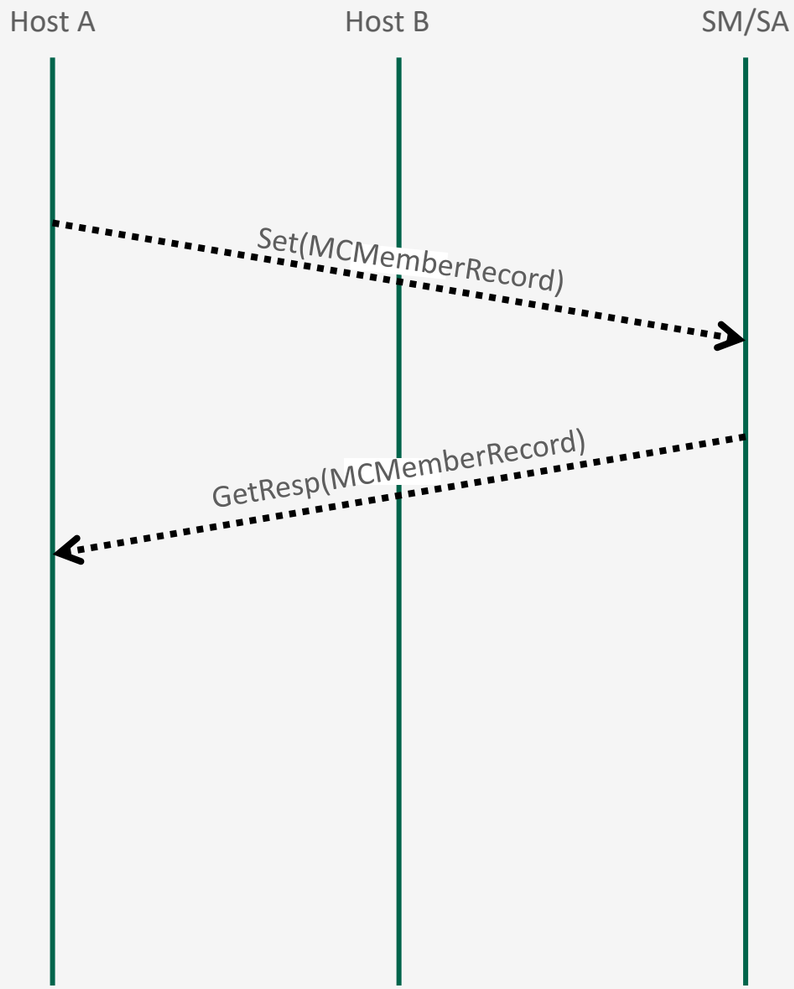


ENHANCED IPOIB (DATAGRAM MODE ONLY)

- Enables offloading ULP basic capabilities in order to optimize IPoIB data path.
- Supports multiple stateless offloads, such as RSS/TSS,
- Enabling IPoIB datagram to reach peak performance in both bandwidth and latency.
- Multi queues
- Interrupt moderation
- Multi partitions optimizations
- Sharing send/receive Work Queues
- Supported on ConnectX-4 adapter cards family and above only.

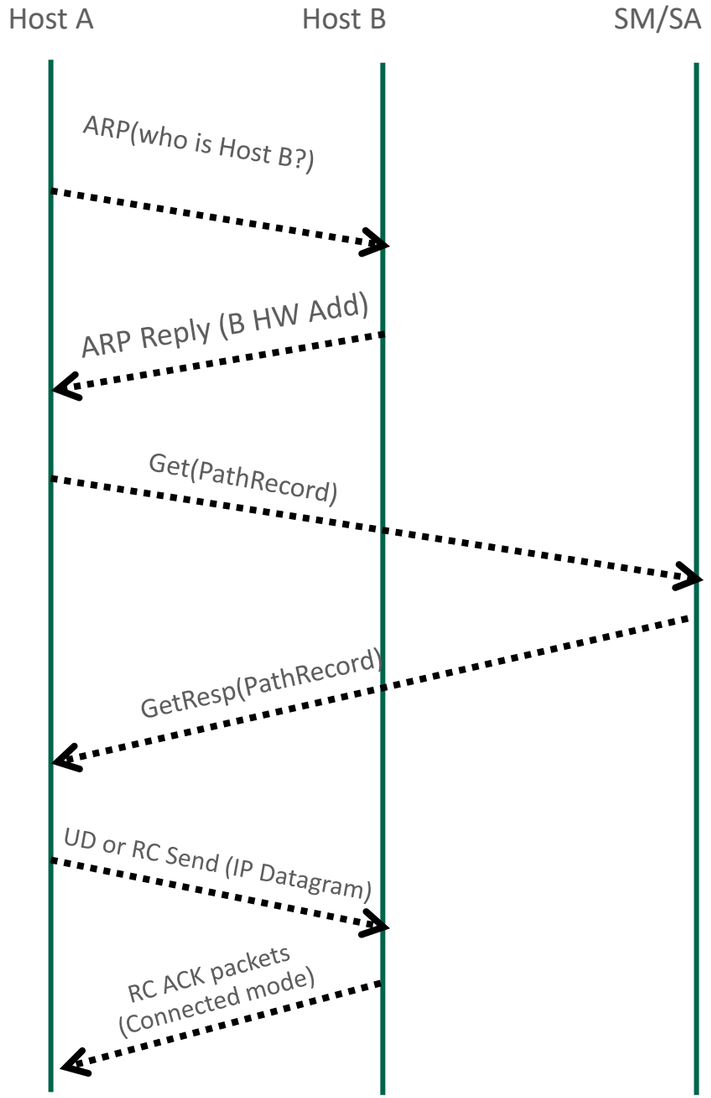
IPOIB PACKET FLOW 1

- Creating a Multicast Group of the hosts that have an IPoIB interface in the fabric:
- Host A sends the SA (Subnet Administrator) a Set(MCMemberRecord) command with the group's MGID
- The SA sends an acknowledgement message back to host A with the group's MLID



IPOIB PACKET FLOW 2

1. Creating a Multicast Group of the hosts in the fabric (that have IPoIB interface)
2. Host A sends an ARP message to the group in order to get the requested host's details.
3. Getting Host B's ARP response
4. Host A Requesting Host B's LID address from the SA
   - Host A sends the SM Host B's port GID and QPn
   - SM responses with Host B's LID extracted from the SA GID to LID data base
5. Host A:
   - Adds an IPoIB encapsulation header
   - Sends Work Request to the relevant QP
   - Starts sending traffic messages towards DLID of Host B
6. Host B:
   - Sends Host A RC ACK packets (Connected Mode)



IPOIB WORKING MODES

- Connected Mode (CM)
  - Uses reliable connection QP
  - IP packet MTU up to 65KB – messages are segmented
- Unreliable Datagram (UD)
  - Uses unreliable connection QP
  - IP packet MTU is limited by InfiniBand MTU (4KB) – segmentation is not supported
- To check the mode, use the following command:

```
[root@mtlacad02 ~]# cat /sys/class/net/ib0/mode
datagram
```

- To change the mode edit the file ‘ **/etc/infiniband/openib.conf** ’.
  - Set ‘**SET_IPOIB_CM=no**’ for datagram mode
  - Set ‘**SET_IPOIB_CM=yes**’ for connected mode
  - The **SET_IPOIB_CM** parameter is set to “**auto**” by default to :
    - Enhanced IPoIB (that is based on datagram but with HW offloads) for ConnectX-4/5/6

3 OPTIONS FOR IPOIB WORKING MODES

There are 3 ways to modify IPoIB working mode :

1. **Per interface mode**
   - The connection mode is defined on the interface file :

     /etc/sysconfig/network-scripts/ifcfg-ibX

     CONNECTED_MODE=no (yes)

2. **The global mode** (In case it is not defined Per interface mode)
   - Changing the mode in the :

     /etc/infiniband/openib.conf

     CONNECTED_MODE=no

3. **During RUNTIME per interface**
   - Stop ibX interface: ifdown ibx
   - Change interface IPoIB mode :
     - echo datagram > /sys/class/net/ib0/mode <- sets the mode of ib0 to UD
     - echo connected > /sys/class/net/ib0/mode < - sets the mode ib0 to Connected

IPOIB CONNECTED MODE MTU REACHES 65K

- Segmentations in supported
- IP MTU > IB MTU

```
[root@mtlacad02 ~]# cat /sys/class/net/ib0/mode
connected
```

```
[root@mtlacad02 ~]# ifconfig ib0
Ifconfig uses the ioctl access method to get the full address information, which limits hardware addresses to 8 bytes.
Because Infiniband address has 20 bytes, only the first 8 bytes are displayed correctly.
Ifconfig is obsolete! For replacement check ip.
ib0       Link encap:InfiniBand  HWaddr A0:00:00:27:FE:80:00:00:00:00:00:00:00:00:00:00:00:00:00:00
          inet addr:192.168.51.2  Bcast:192.168.51.255  Mask:255.255.255.0
          inet6 addr: fe80::e61d:2d03:fd:c6e8/64 Scope:Link
          UP BROADCAST RUNNING MULTICAST  MTU:65520  Metric:1
          RX packets:0 errors:0 dropped:0 overruns:0 frame:0
          TX packets:5 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:1024
          RX bytes:0 (0.0 b)  TX bytes:380 (380.0 b)
```

IPOIB 常见问题

大规模组网环境下的DHCP request风暴

协议栈里IP的租赁时间很短，在大规模组网时，会导致每个IPOIB同时发起DHCP请求而产生风暴，因为SM会集中处理，又会导致SM的负载过高，解决方法设置比较长的IP租赁时间，将SM的并发处理能力提高（通过opensm配置）

IPoIB的DHCP失败

在实际中DPCP申请IP地址失败，需要检查pkey是否匹配/multicast table是否成功，tracert查看整条路径是否正常

IPoIB的性能

在connected模式TCP的性能一般比diagram下好，原因是connected下的MTU可以达到64K，对大包或者比较以后，diagram模型因为某些就网卡不支持LRO等，性能不理想。采用enhanced diagram模式性能会比较好

IPoIB tx queue timeout

在connected的模式下，对应的QP是RC模式，某些情况下会有CQ error造成TX queue timeout，RC是有状态的，被datagram处理起来复杂很多，建议是使用enhanced diagram模式，这样避免RC的弊端

UFM

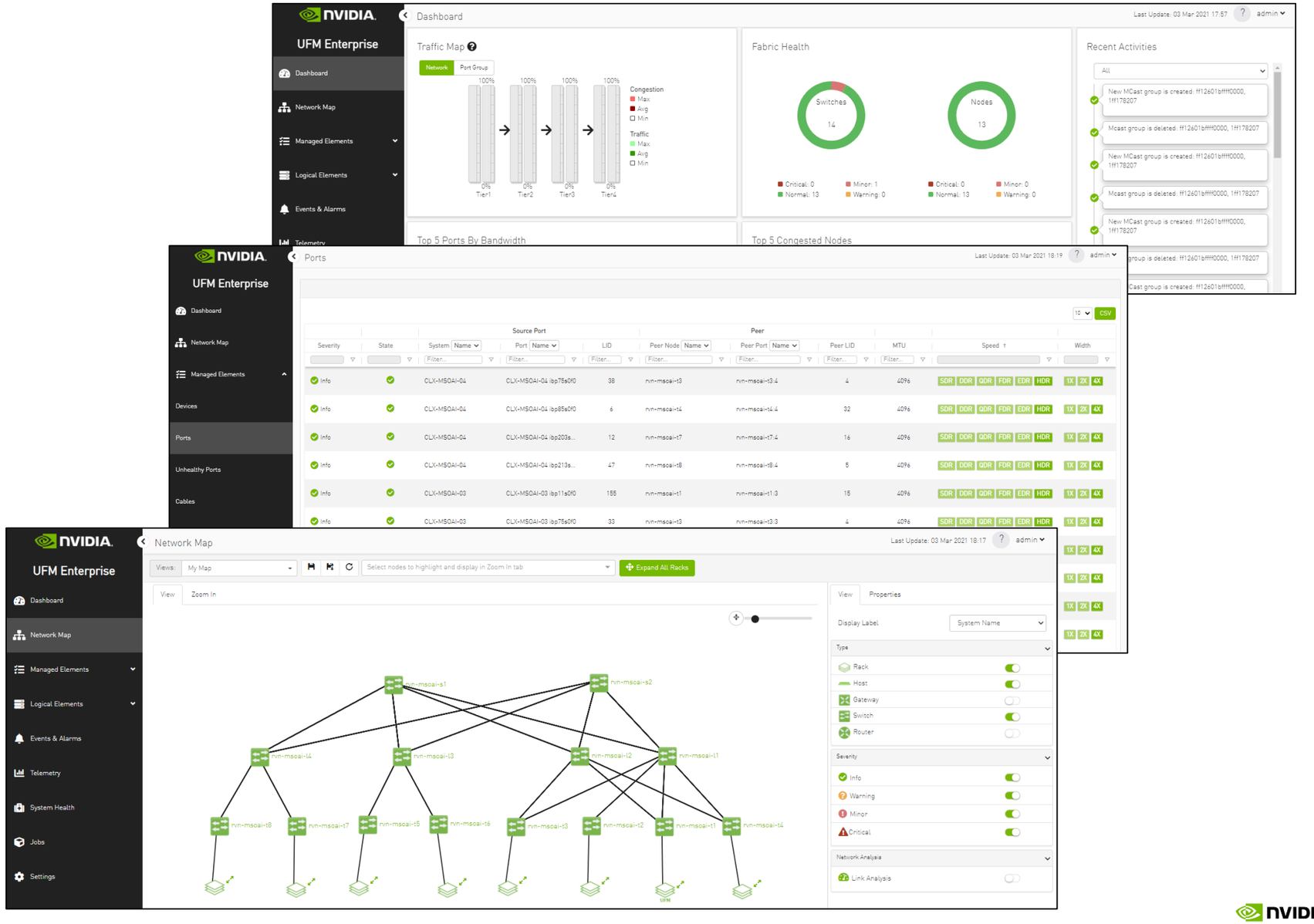
UFM IN THE FABRIC

- Manages Subnet Manager and Sharp Services
- Software or appliance form factor
- High availability - 2 or more
- Switch and adapter management
- Full management or monitoring only
- Layer 2 level monitoring
- REST API for configuration/monitoring
- Single Interface for all network

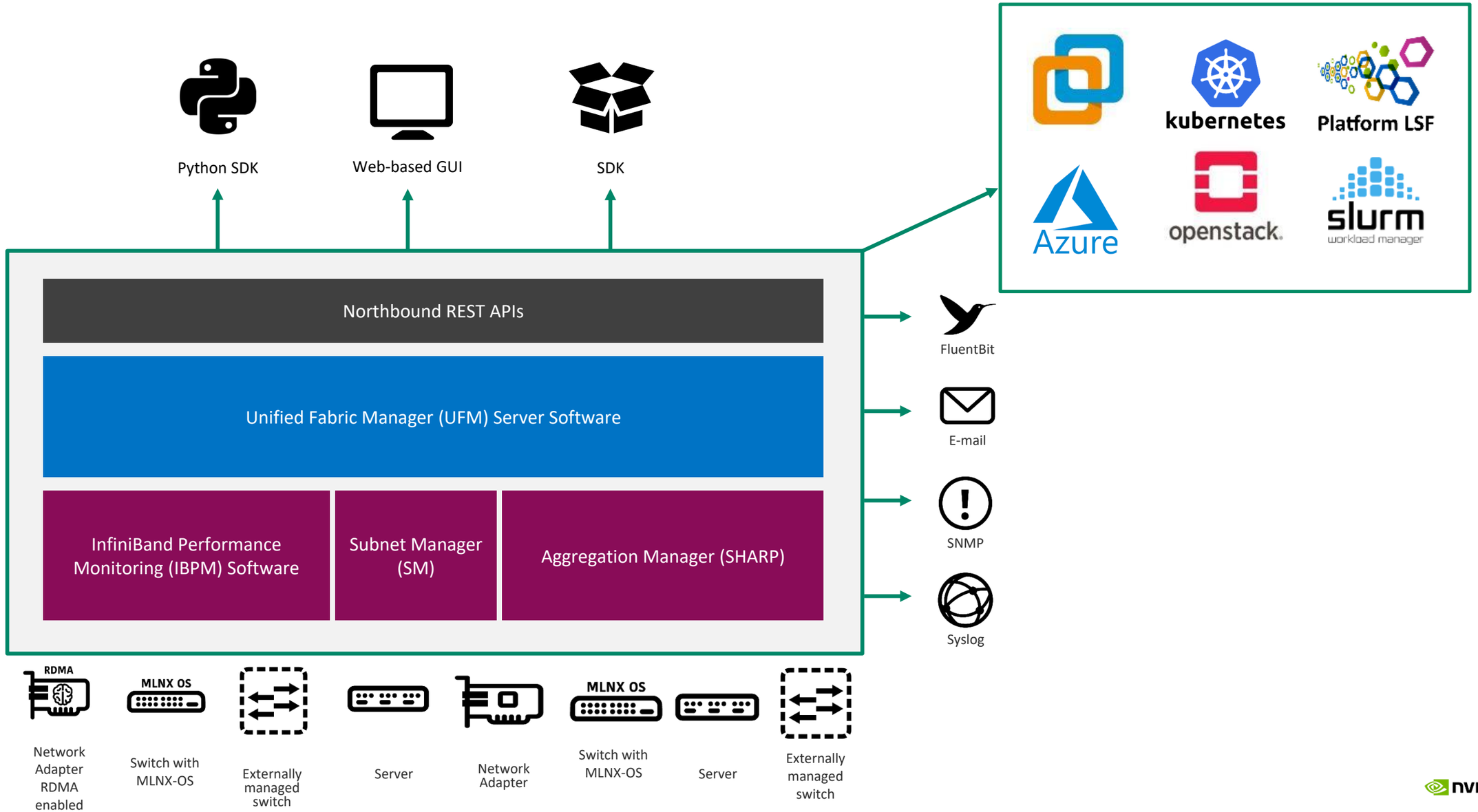
NVIDIA Fabric

KEY FEATURES

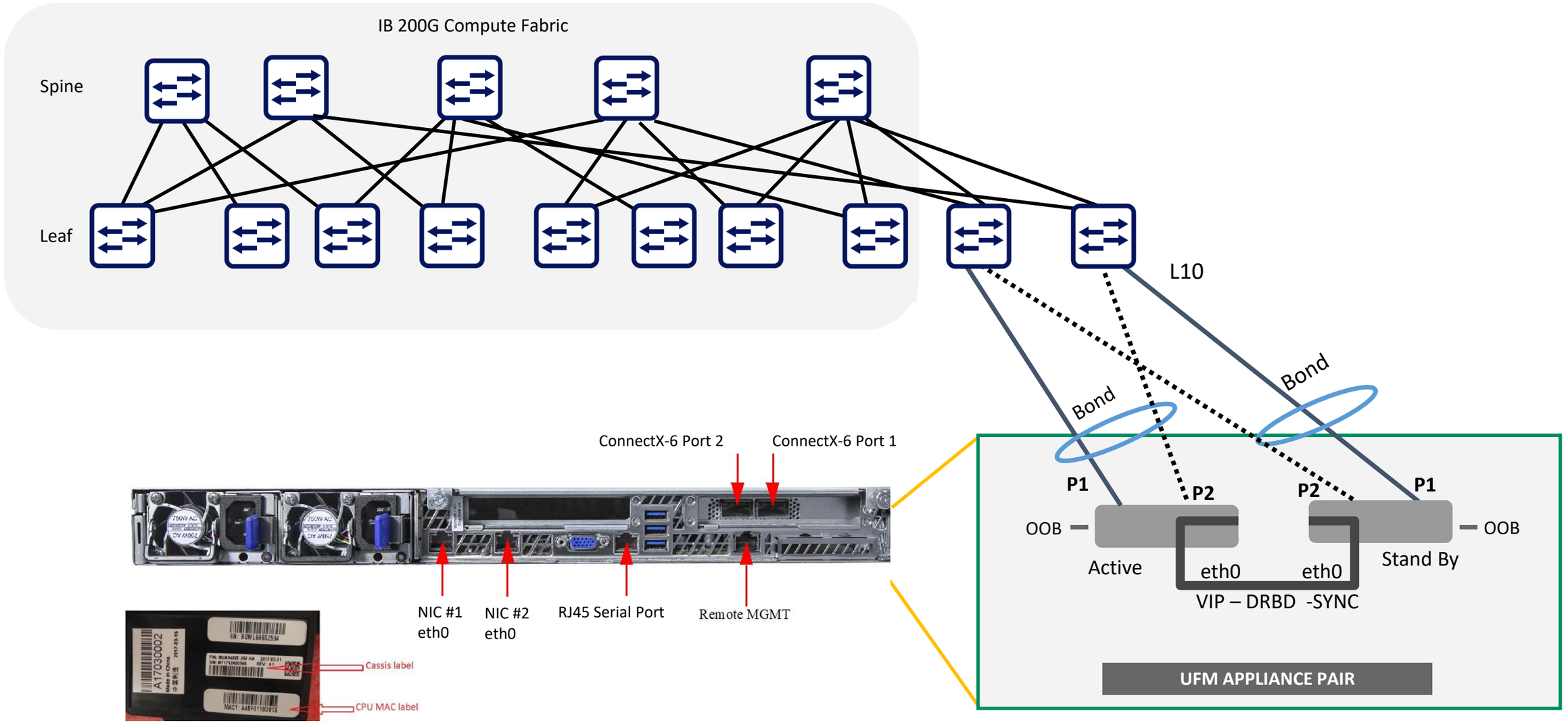
- Automated Network Discovery
- Centralized Device Management
- Automated Network Provisioning
- Software/FW Upgrades
- Fabric and Cluster Validation
- Network Telemetry
- Traffic Monitoring
- Performance Monitoring
- Health/Fault Monitoring
- Advanced Reporting
- Comprehensive REST APIs
- Rich Web-based UI



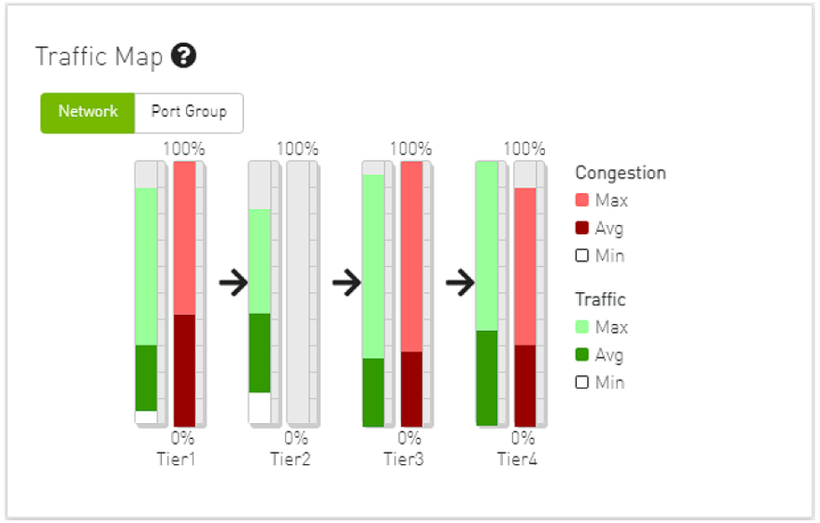
UFM SOFTWARE ARCHITECHTURE

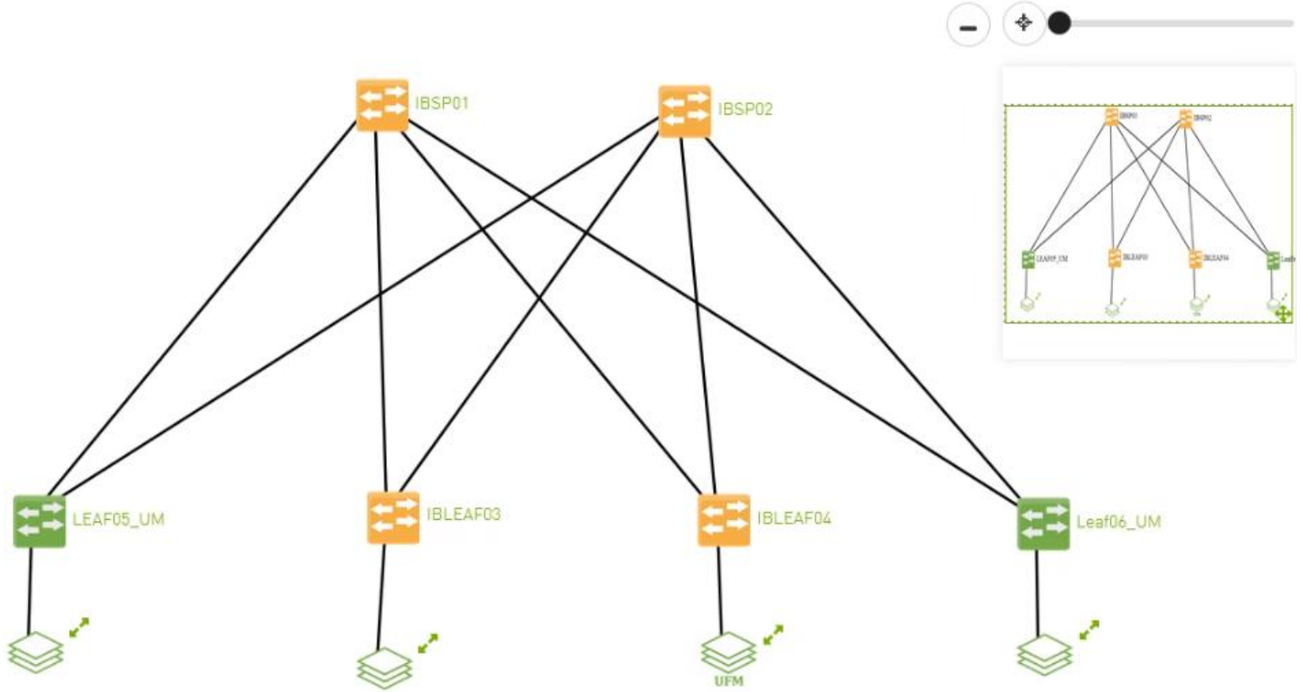


UFM INFINIBAND MANAGEMENT PAIR



UFM DASHBOARD





TRAFFIC MAP TIERS

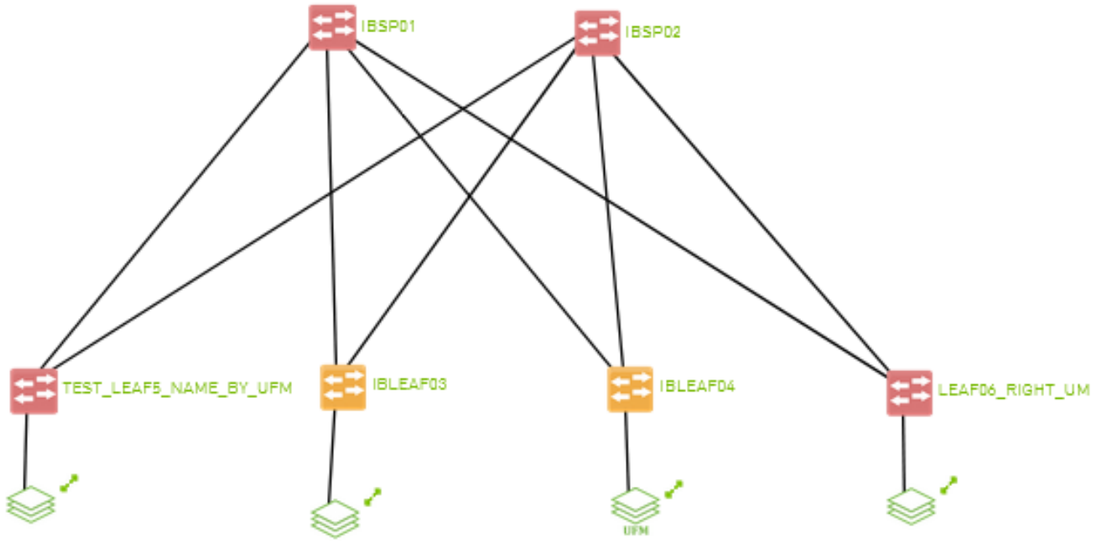
Traffic map Tiers significance :

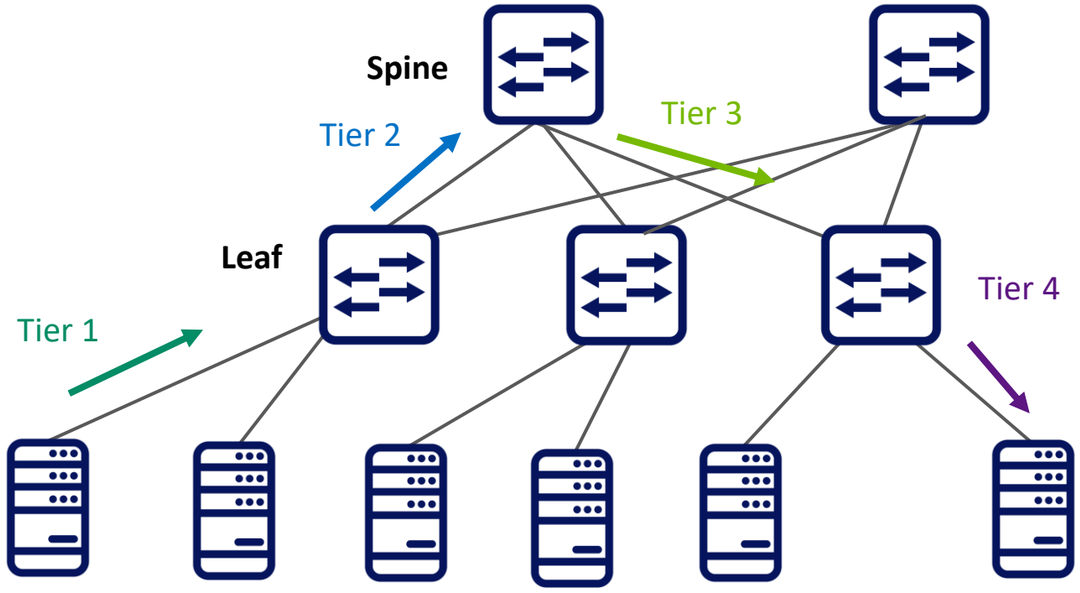
Tier 1 – represents the traffic injected by all adapters

Tier 2 – represents the traffic sent from the edge switches to the core of the fabric
(in case of a 3 layers topology this represents the traffic between the edge leaves and the next topology layers – the spines)

Tier 3 – represents the traffic sent from the core/spine to the edge switches

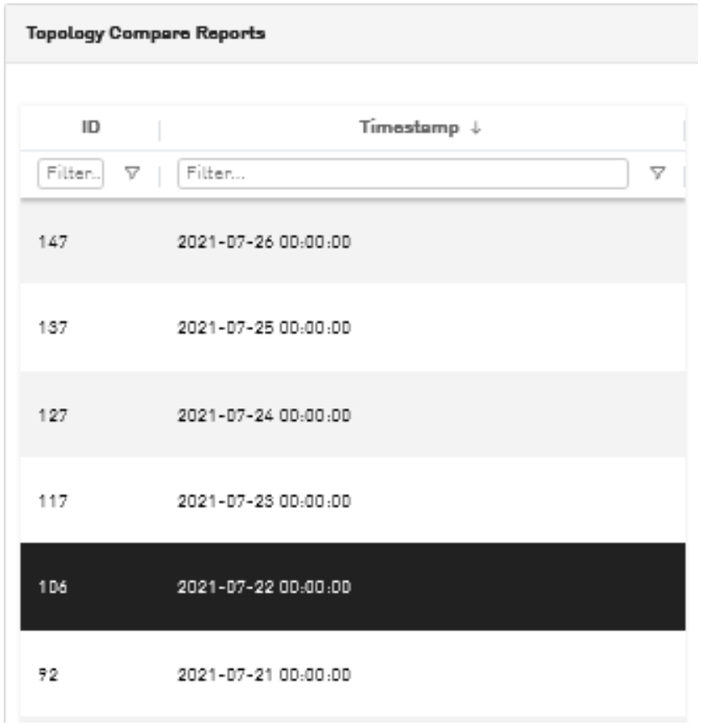
Tier 4 – represents the traffic sent from the edge switch to the adapters





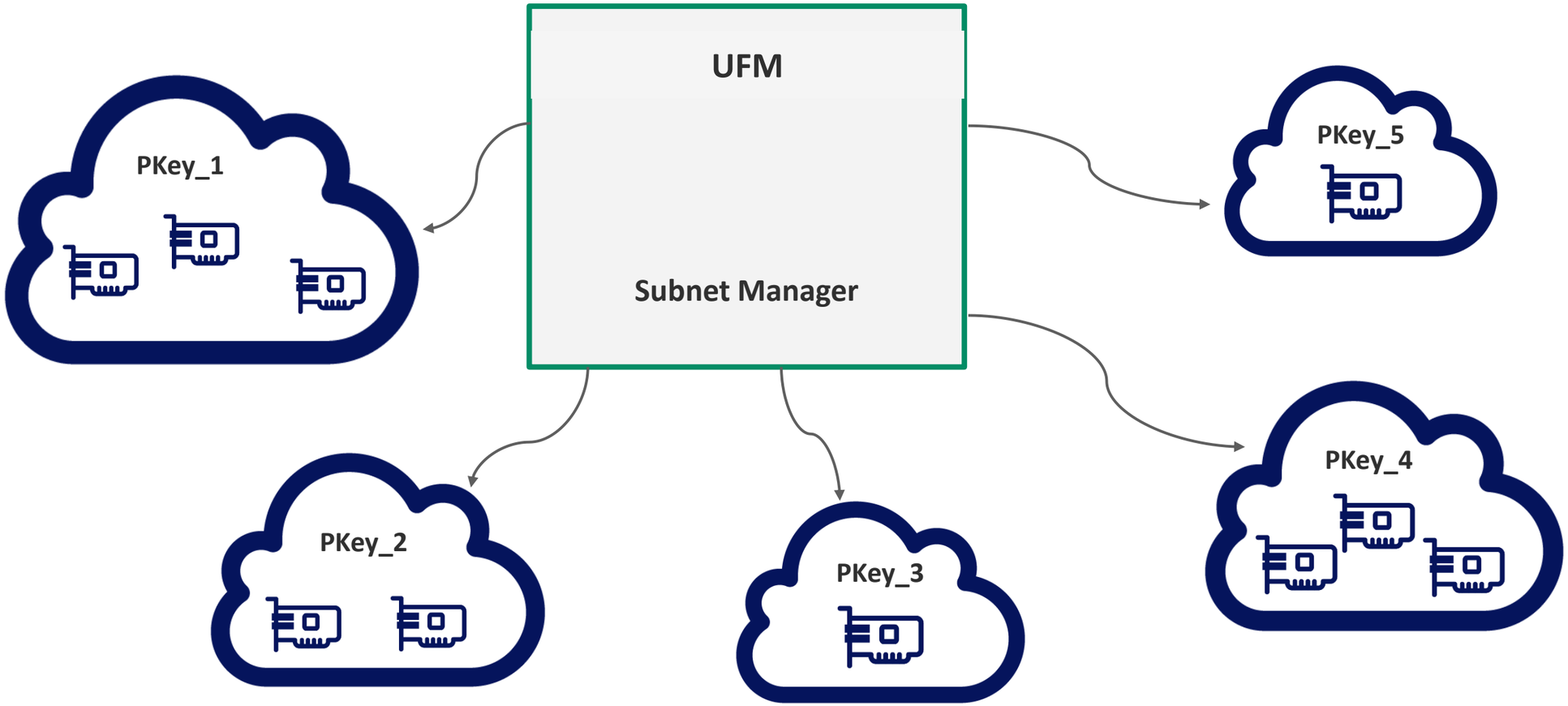
TOPOLOGY COMPARE REPORTS

Master Topology Snapshot:
/opt/ufm/files/periodicTopo/master.topo
Last Update: 2021-07-12 15:50:06

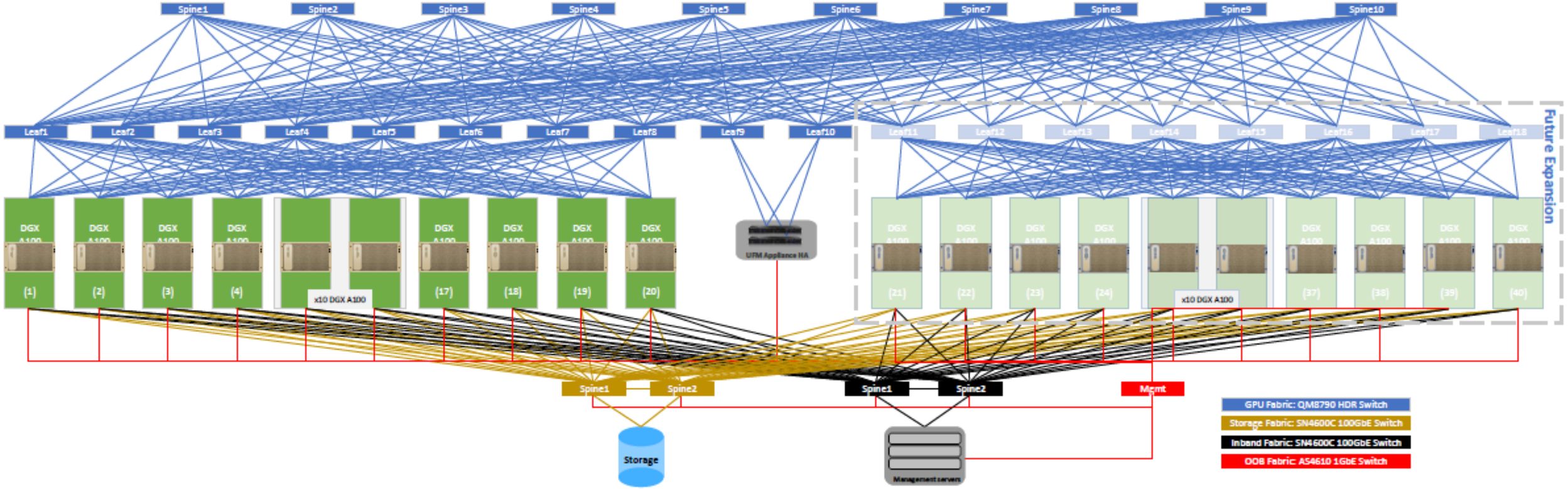


CLOUD NETWORKING API

- Allows operators to create/remove/update tenants
- Manages PKey GUIDs by getting, adding, and removing GUIDs from PKeys
- Isolate tenant networks



COMPUTE FABRIC TOPOLOGY




Q&A

Q&A

- **How to check RDMA traffic performance in IB cluster**
  check real time performance on UFM, for realtime of some port, need using perfquery -x
- **How to check RDMA counters of nic and ib switch**
  show_counter on host for nic, perfquery –x, ethtool -S ib0 | grep rdma
- **Where run opensm? Host or switch**
  Host, managed switch or applicance, refer to page 11
- **How to change MTU of IB and IPoIB**
  refer to page 26/59/60
- **What is the performance of IPoIB?**
  connected mode can get line rate, enhanced datagram also can get line rate
- **How to configure bond for IB port ?**
  IB RDMA device doesn't support bond, IPoIB ports support active-backup mode bonding. Applications should take care of the requirement by support multiple RDMA port or by multi-rail.
- **ibstat shows active and link-up, what's the difference**
  active state is logic state for up layer, usually is related by OpenSM. link-up is for physical layer state, means link is up, not managed by OpenSM.

REFERENCE

REFERENCE

IBTA offical webpage
https://www.infinibandta.org/

Inifiband official webpage
https://www.nvidia.com/en-us/networking/products/infiniband/

INFINIBAND/VPI SOFTWARE - MLNX_OFED
https://developer.nvidia.com/networking/infiniband-software

INFINIBAND Academy courses
https://academy.nvidia.com/en/training-by-topic/?training_by_topic=9&subt=54

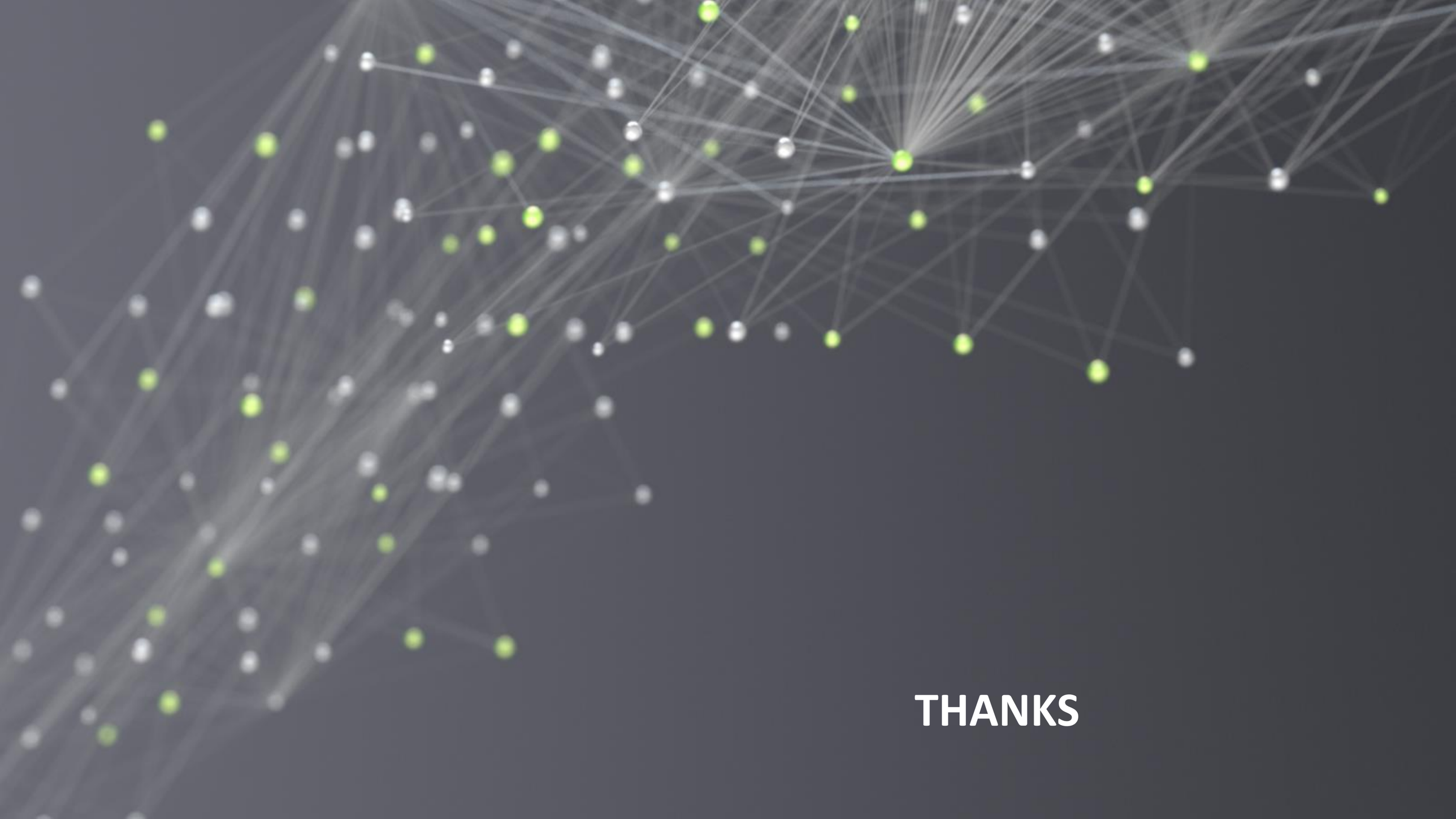
Community for questions and articles
https://community.mellanox.com/s/

Documents for solutions of networking products
https://docs.nvidia.com/networking/

QSG: Kubernetes Cluster Deployment on InfiniBand Fabric with RDMA Shared Device Plugin.
https://docs.nvidia.com/networking/pages/releaseview.action?pageId=18481842

How-to: Deploy RDMA accelerated Docker container over InfiniBand fabric.
https://docs.nvidia.com/networking/pages/releaseview.action?pageId=15049785

RDG: Virtualizing GPU-Accelerated HPC & AI Workloads on OpenStack Cloud over InfiniBand Fabric.
https://docs.nvidia.com/networking/pages/viewpage.action?pageId=30608172


THANKS