BSC-CNS Networks

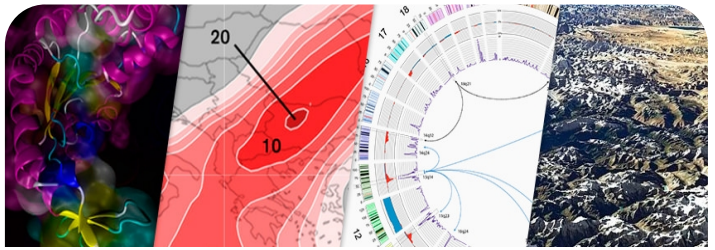
Barcelona
maria.manco@bsc.es
Network & Security
SIG-NOC 2025

OCT 15th 2025

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Objectives

Supercomputing services to Spanish and EU researchers



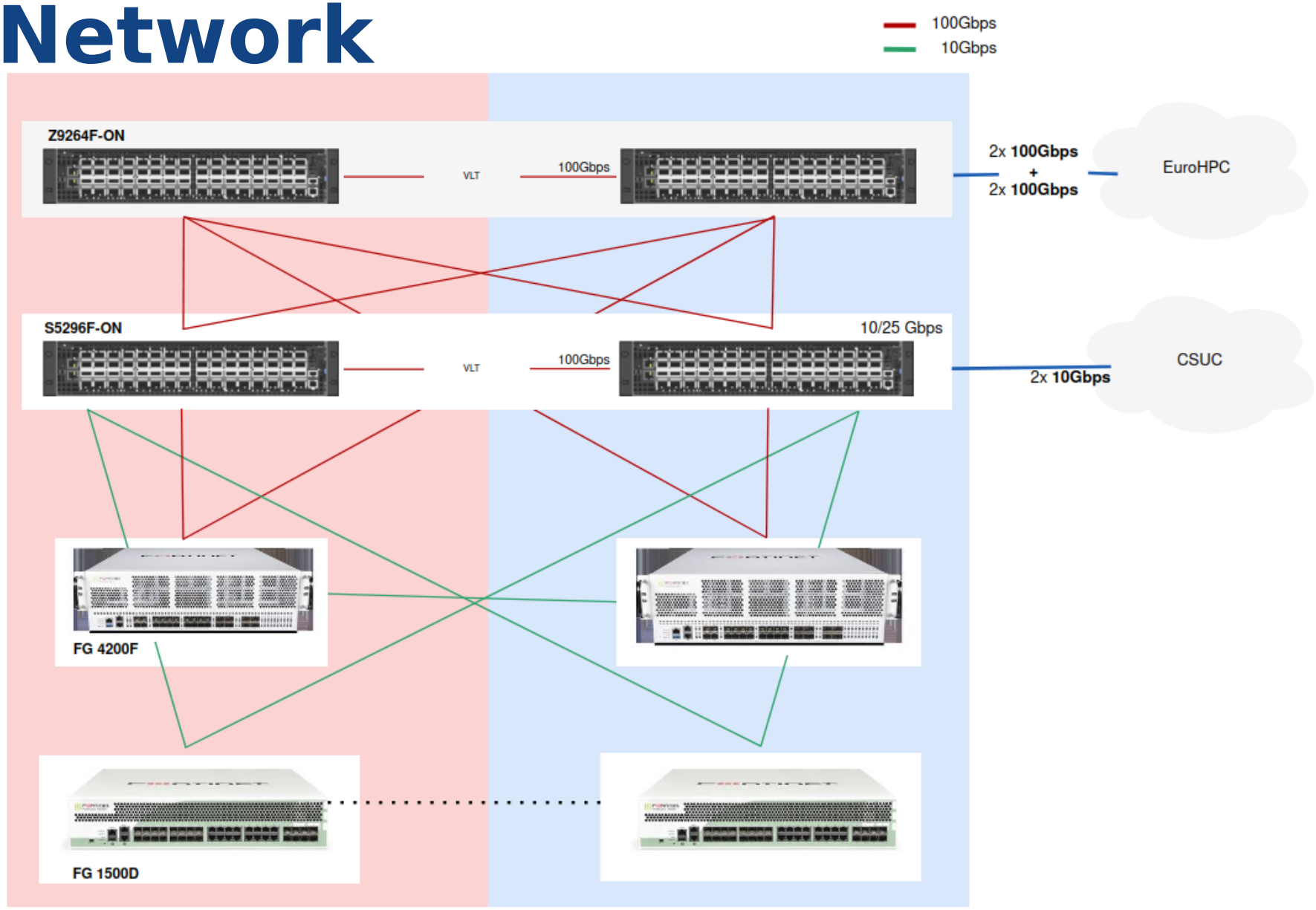
R&D in Computer, Life, Earth and Engineering Sciences

PhD programme, technology transfer, public engagement

Spanish Government	**60%**
Catalan Government	**30%**
UPC	**10%**

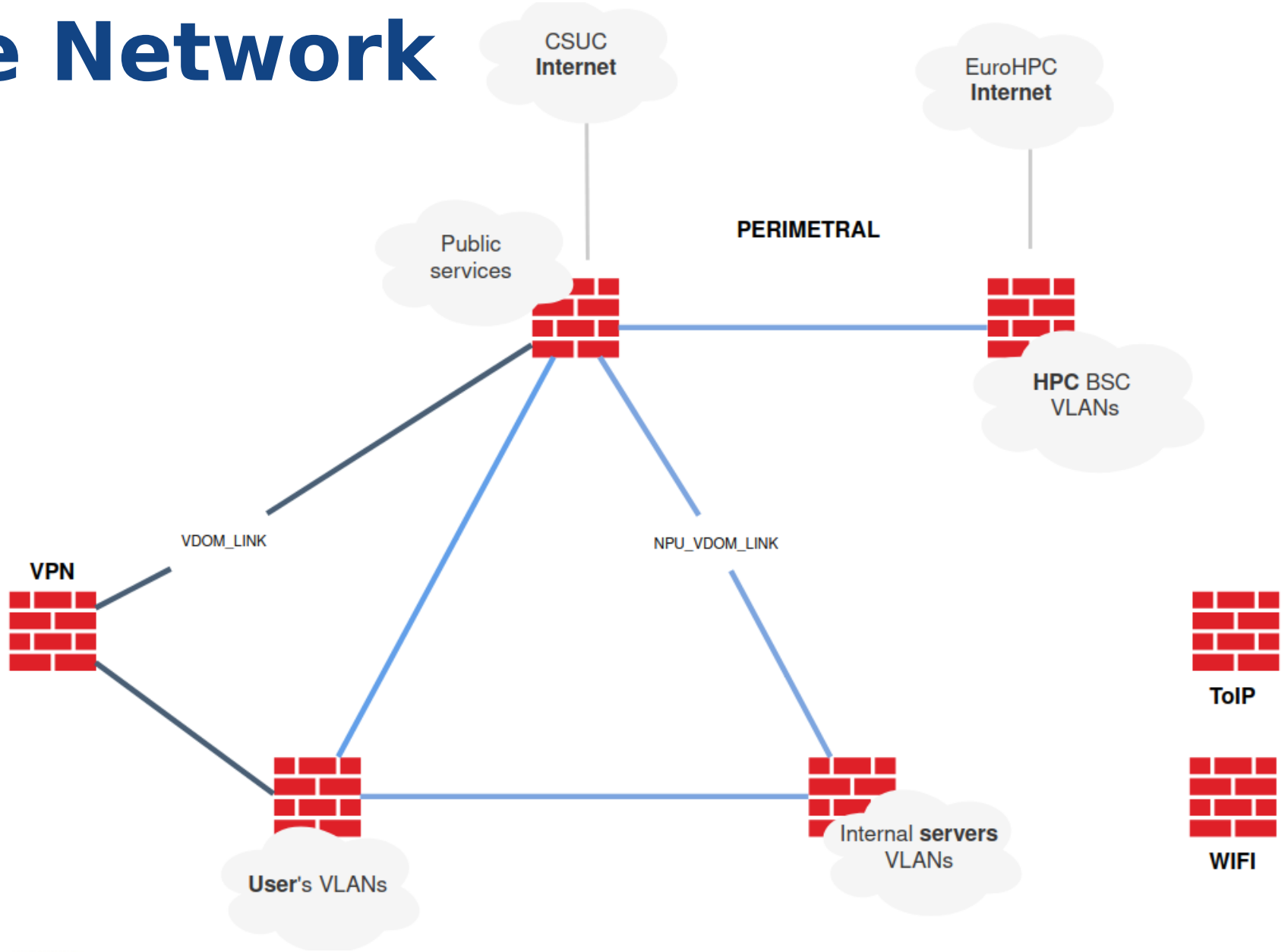
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Core Network



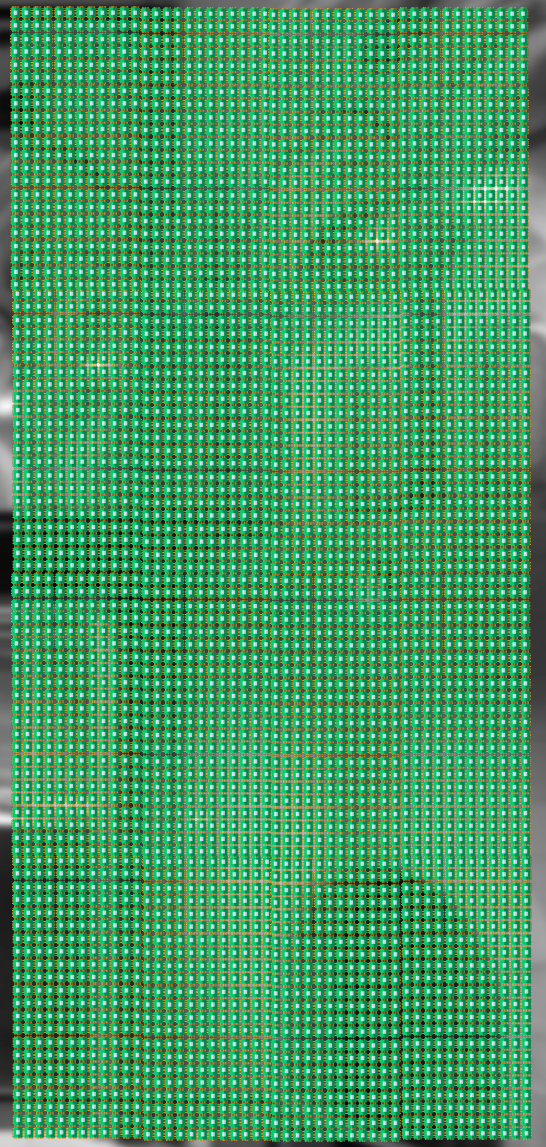
Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Core Network



MareNostrum 5

Total peak performance: **314** Pflops	
GPP	**45.4 Pflops**
ACC	**260 Pflops**
NGT GPP	**2.82 Pflops**
NGT ACC	**6 Pflops**



MareNostrum 1
2004 – 42.3 Tflops
1st Europe / 4th World
New technologies

MareNostrum 2
2006 – 94.2 Tflops
1st Europe / 5th World
New technologies

MareNostrum 3
2012 – 1.1 Pflops
12th Europe / 36th World

MareNostrum 4
2017 – 11.1 Pflops
2nd Europe / 13th World
New technologies

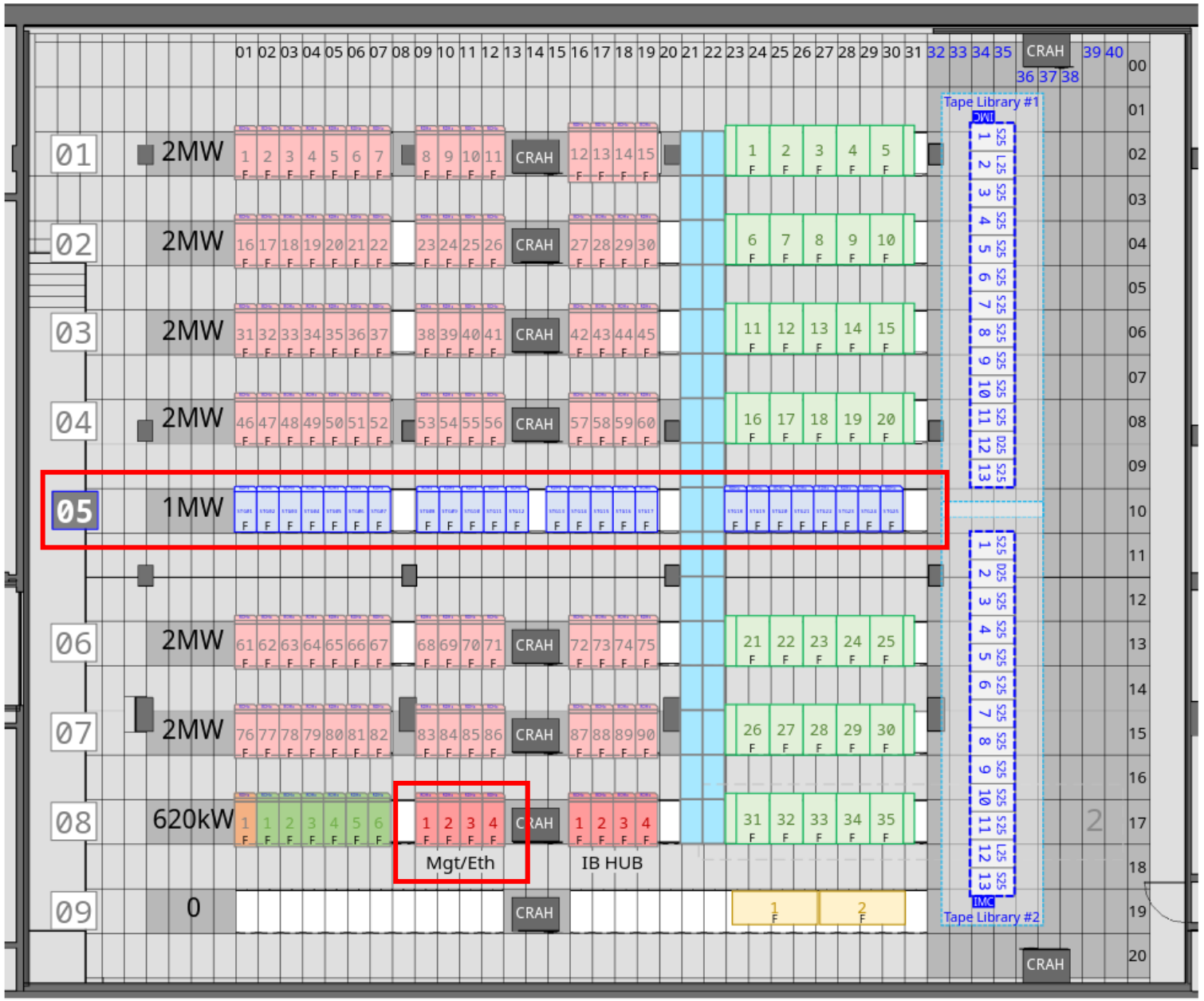
MareNostrum 5
2023 – 204.6 Pflops
3nd Europe / 8th World

MN5 Networks: Ethernet and Infiniband

- Location
- Hardware
- Purpose
- Topology
- Rack Connectivity
- Monitoring

Ethernet

Location



*Backed up by a UPS

Hardware

Nvidia SN**4600V**

64 100GbE QSFP56 ports

128 10/50/25/10/1GbE (breakout)

425ns latency

8.4B pps

Line-Rate switching

L2/L3

60W Typical

Cumulus

Nvidia SN**3700C**

32 100GbE QSFP28 ports

128 25/10/1GbE ports (breakout)

425ns latency

4.76B pps

Line-Rate switching

L2/L3

242W Typical

Cumulus

Nvidia SN**3420**

48 25/10/1GbE SFP+ ports

12 100GbE QSFP28 ports (uplinks)

425ns latency

3.58B pps

Line-Rate switching

L2/L3

242W Typical

Cumulus

Nvidia **AS4610-54t**

48 1GbE RJ45 ports

4 10GbE SFP+ ports (uplinks)

4us latency

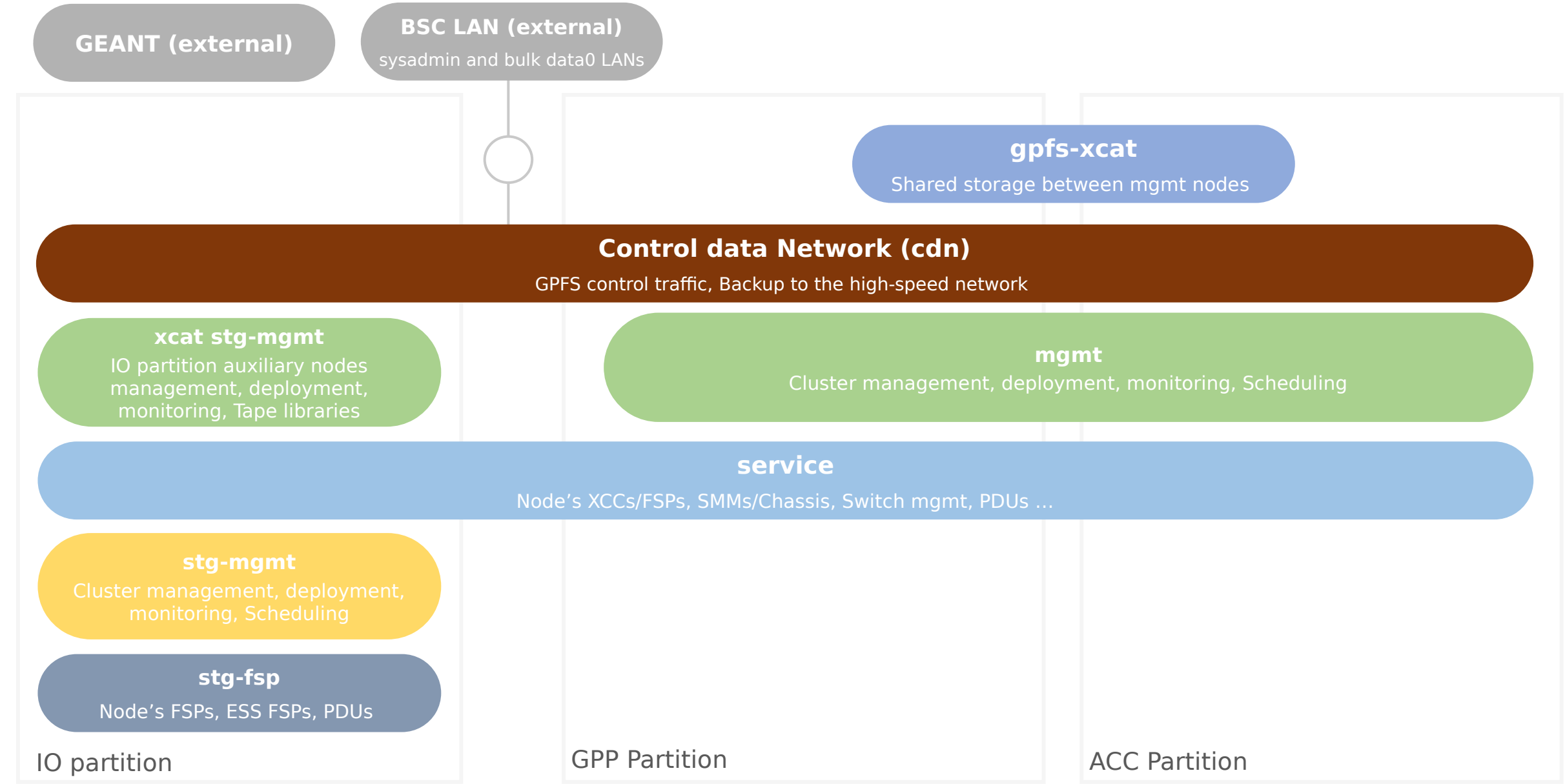
-

Line-Rate switching

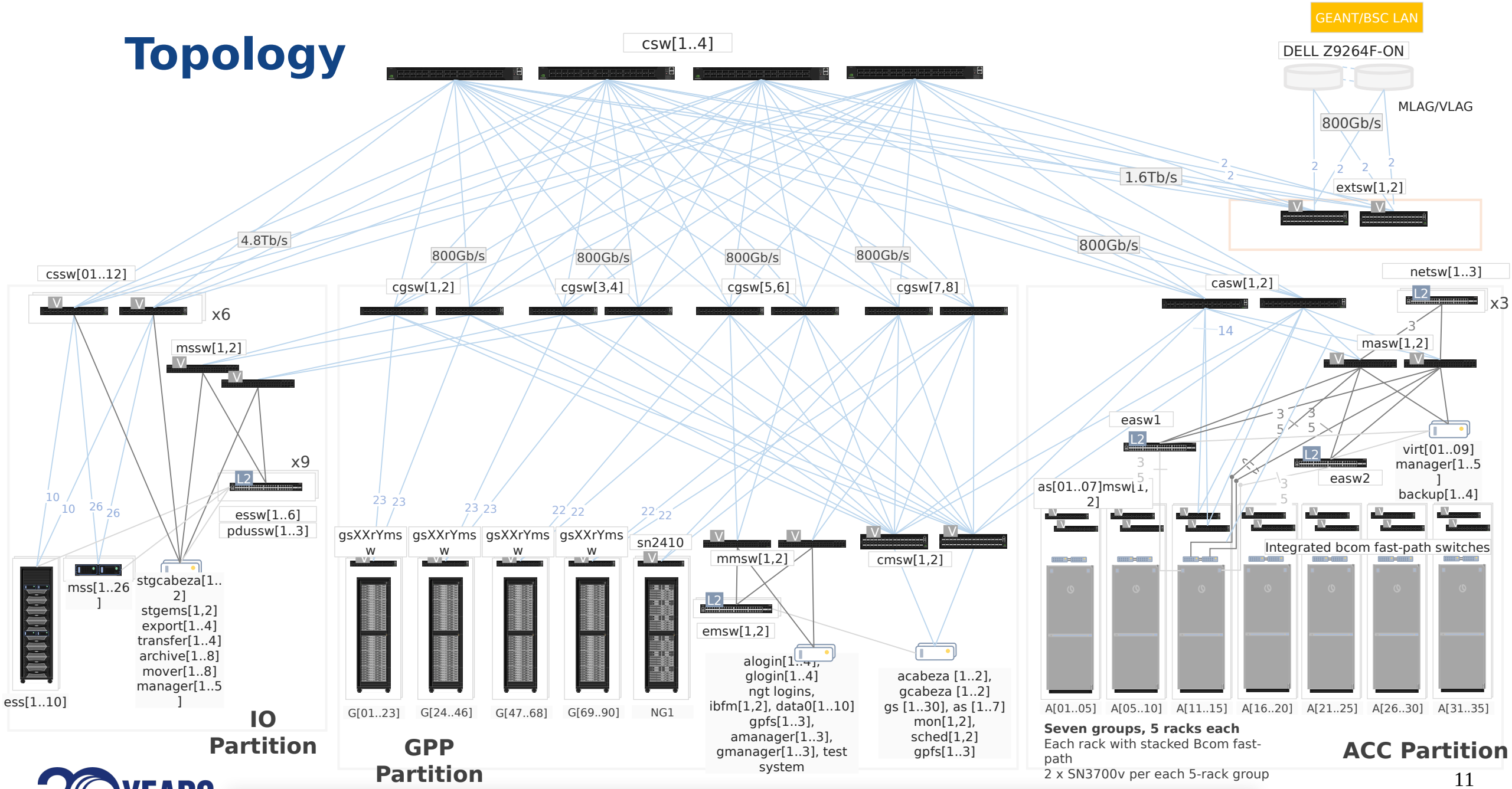
L2/L3

90W Typical

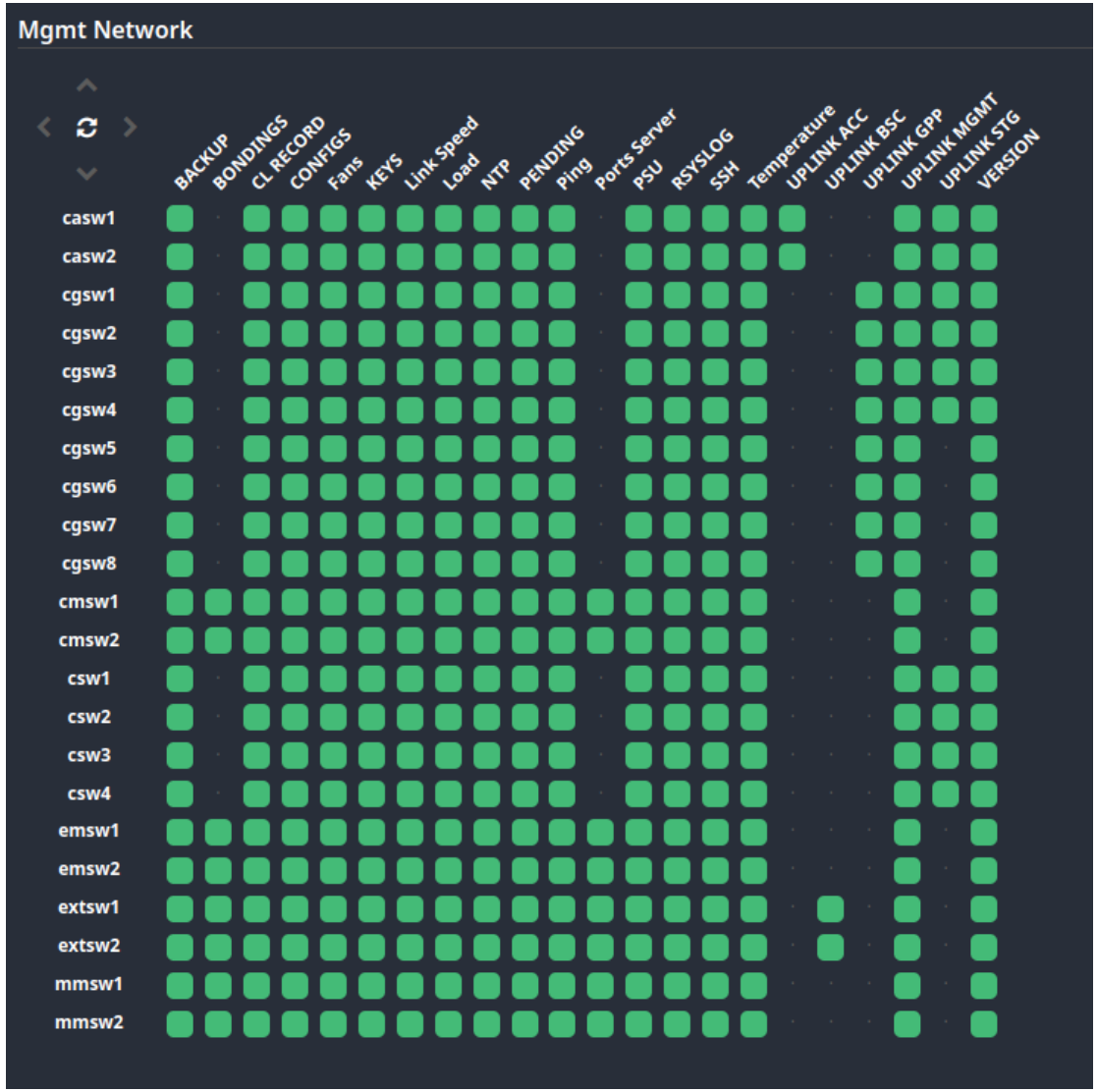
Cumulus

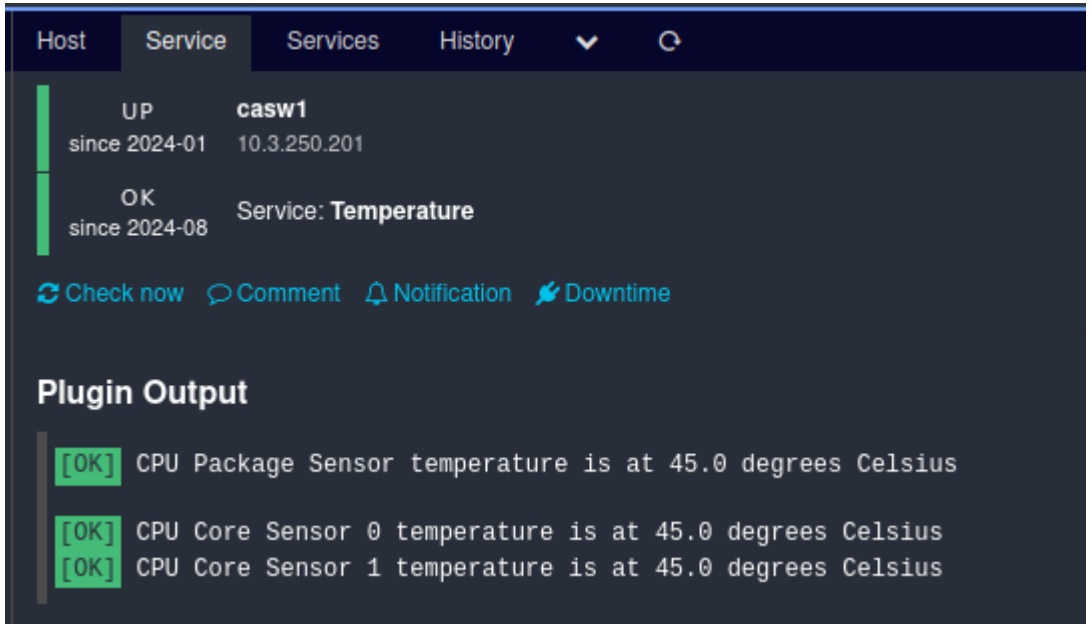

GEANT (external)
BSC LAN (external)
sysadmin and bulk data0 LANs
gpfs-xcat
Shared storage between mgmt nodes
Control data Network (cdn)
GPFS control traffic, Backup to the high-speed network
xcat stg-mgmt
IO partition auxiliary nodes management, deployment, monitoring, Tape libraries
mgmt
Cluster management, deployment, monitoring, Scheduling
service
Node's XCCs/FSPs, SMMs/Chassis, Switch mgmt, PDUs ...
stg-mgmt
Cluster management, deployment, monitoring, Scheduling
stg-fsp
Node's FSPs, ESS FSPs, PDUs
IO partition
GPP Partition
ACC Partition

Topology


GEANT/BSC LAN
DELL Z9264F-ON
MLAG/VLAG
800Gb/s
extsw[1,2]
1.6Tb/s
csw[1..4]
4.8Tb/s
800Gb/s
800Gb/s
800Gb/s
800Gb/s
800Gb/s
cssw[01..12]
x6
cgsw[1,2]
cgsw[3,4]
cgsw[5,6]
cgsw[7,8]
casw[1,2]
netsw[1..3]
x3
mssw[1,2]
masw[1,2]
x9
essw[1..6]
pdussw[1..3]
mss[1..26]
stgcabeza[1..2]
stgems[1,2]
export[1..4]
transfer[1..4]
archive[1..8]
mover[1..8]
manager[1..5]
ess[1..10]
gsXXrYmsw
gsXXrYmsw
gsXXrYmsw
gsXXrYmsw
sn2410
G[01..23]
G[24..46]
G[47..68]
G[69..90]
NG1
mmsw[1,2]
cmsw[1,2]
emsw[1,2]
alogin[1..4],
glogin[1..4]
ngt logins,
ibfm[1,2], data0[1..10]
gpfs[1..3],
amanager[1..3],
gmanager[1..3], test system
acabeza [1..2],
gcabeza [1..2]
gs [1..30], as [1..7]
mon[1,2],
sched[1,2]
gpfs[1..3]
easw1
easw2
as[01..07]msw[1,2]
virt[01..09]
manager[1..5]
backup[1..4]
Integrated bcom fast-path switches
A[01..05]
A[05..10]
A[11..15]
A[16..20]
A[21..25]
A[26..30]
A[31..35]
Seven groups, 5 racks each
Each rack with stacked Bcom fast-path
2 x SN3700v per each 5-rack group
IO Partition
GPP Partition
ACC Partition
Service
mgmt GPP
cdn
gpfs-xcat
xcat stg-mgmt
External
stg-fsp
stg-mgmt
V VTEP
X Free ports
L2 L2 Only
S Splitter cable
1GbE
25GbE
10GbE

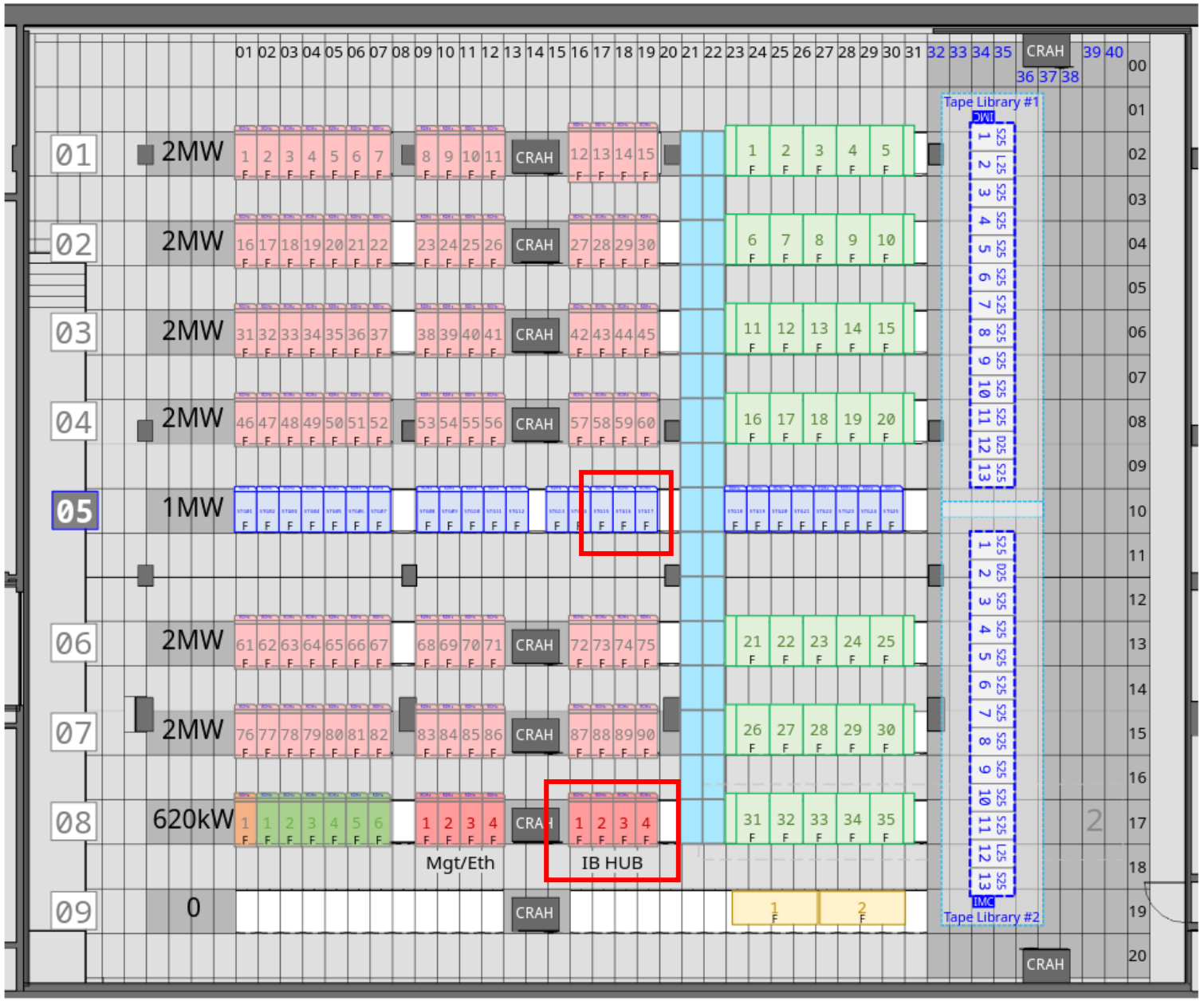
Monitoring


Mgmt Network
BACKUP
BONDINGS
CL RECORD
CONFIGS
Fans
KEYS
Link Speed
Load
NTP
PENDING
Ping
Ports Server
PSU
RSYSLOG
SSH
Temperature
UPLINK ACC
UPLINK BSC
UPLINK GPP
UPLINK MGMT
UPLINK STG
VERSION
casw1
casw2
cgsw1
cgsw2
cgsw3
cgsw4
cgsw5
cgsw6
cgsw7
cgsw8
cmsw1
cmsw2
csw1
csw2
csw3
csw4
emsw1
emsw2
extsw1
extsw2
mmsw1
mmsw2


Host
Service
Services
History
UP
since 2024-01
casw1
10.3.250.201
OK
since 2024-08
Service: Temperature
Check now
Comment
Notification
Downtime
Plugin Output
[OK] CPU Package Sensor temperature is at 45.0 degrees Celsius
[OK] CPU Core Sensor 0 temperature is at 45.0 degrees Celsius
[OK] CPU Core Sensor 1 temperature is at 45.0 degrees Celsius

High Speed Interconnect: **Infiniband**

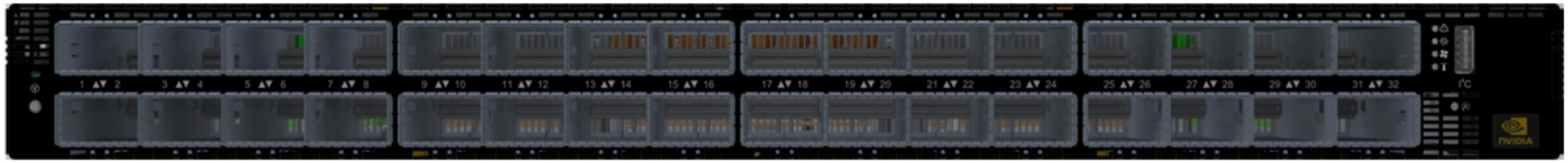
Location



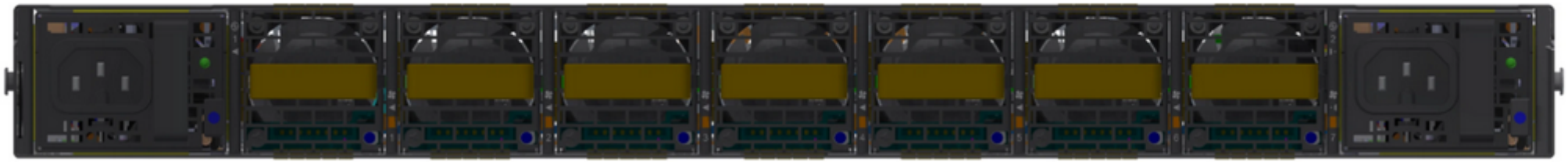
*Backed up by a UPS

Hardware

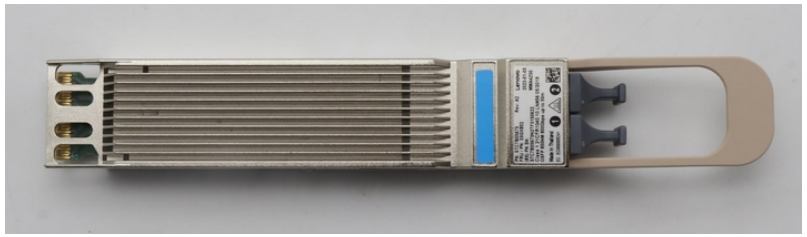
QM9790 Front View



QM9700 and QM9790 Rear View

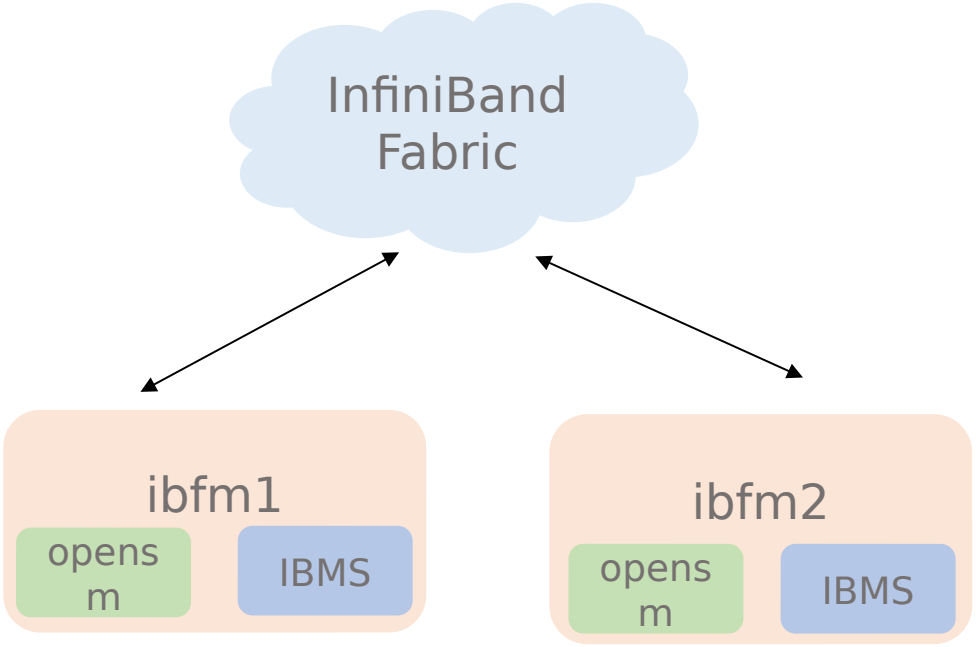


- 64 switch ports of **NDR** 400 Gb/s
- **2 x** redundant hot swap power supplies
- **7 x** redundant hot swap fans



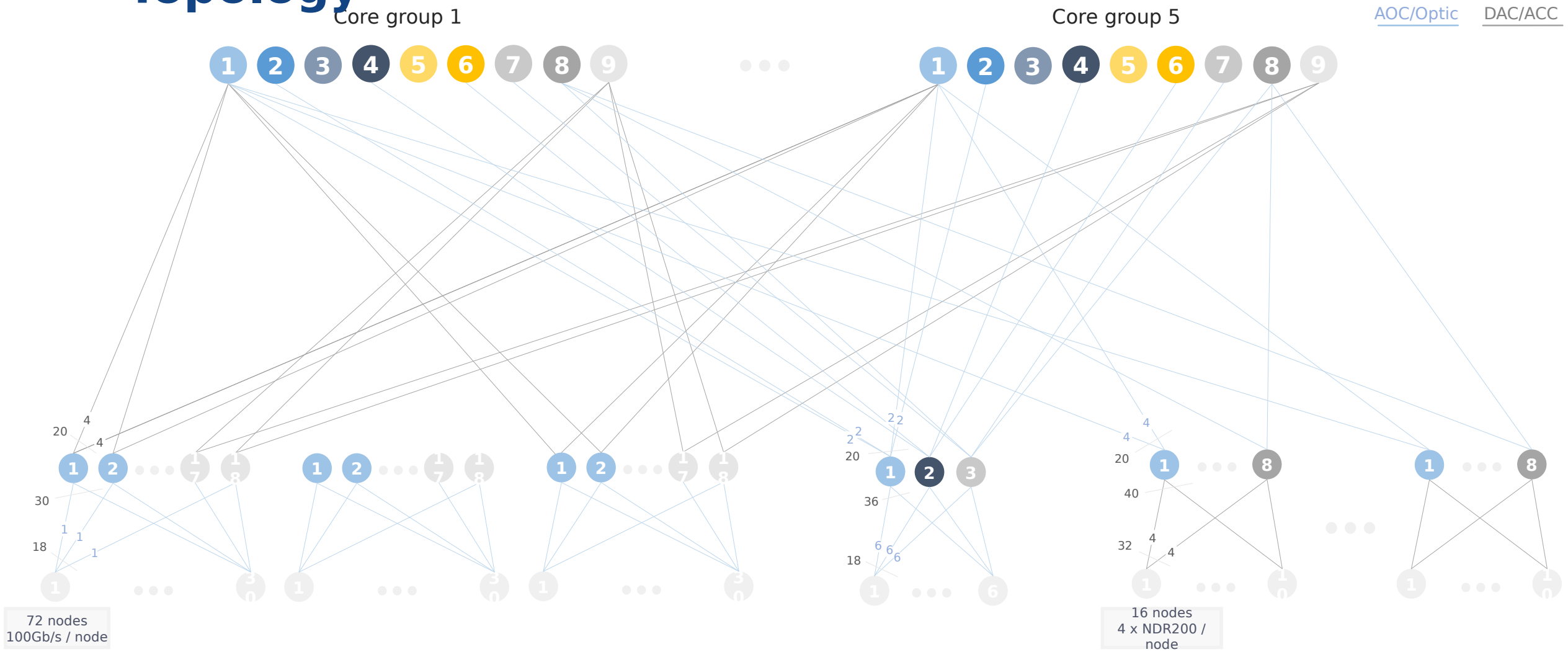
Purpose and Design

- **Low Latency**
- **High Bandwidth**
- **High Scalability**
- **RDMA Support**
- **Efficient Network Topology**



- Managed using **opensm** subnet manager service
- Eviden **IBMS** fabric manager software

Topology



GPP Island 1
30 compute racks
2,160 compute nodes

GPP Island 2
30 compute racks
2,160 compute nodes

GPP Island 3
30 compute racks
2,160 compute nodes

IO partition and
management island

ACC Island 1
5 compute racks
160 compute nodes

ACC Island 7
5 compute racks
160 compute nodes

Rack Connectivity - IB

Cores

702W

Avg typical power consumption per switch

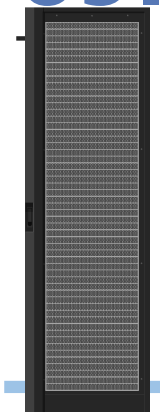
With 15 x twin NDR optical connections

17.4kW

Avg typical power consumption per rack

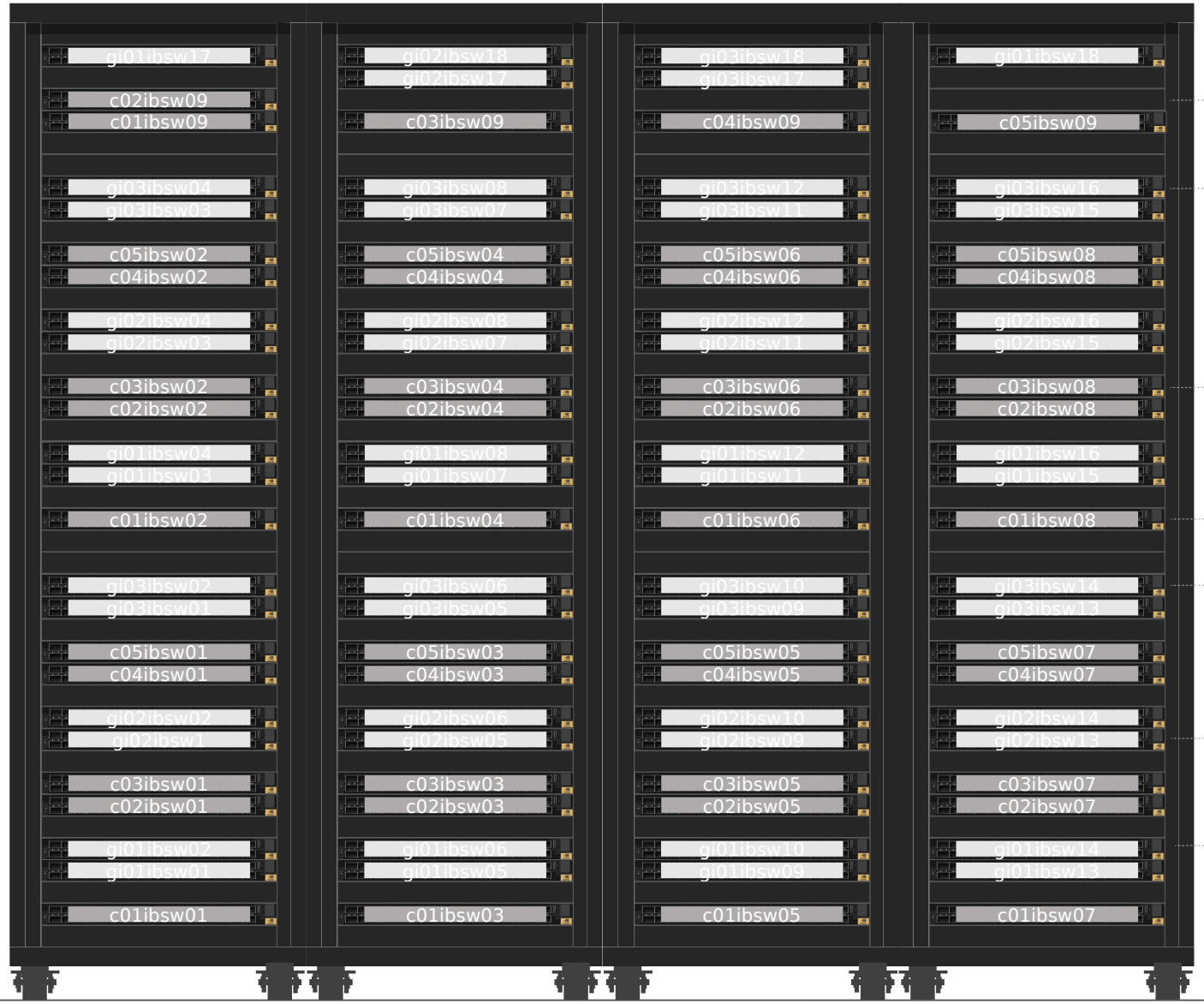
69.6kW

power consumption

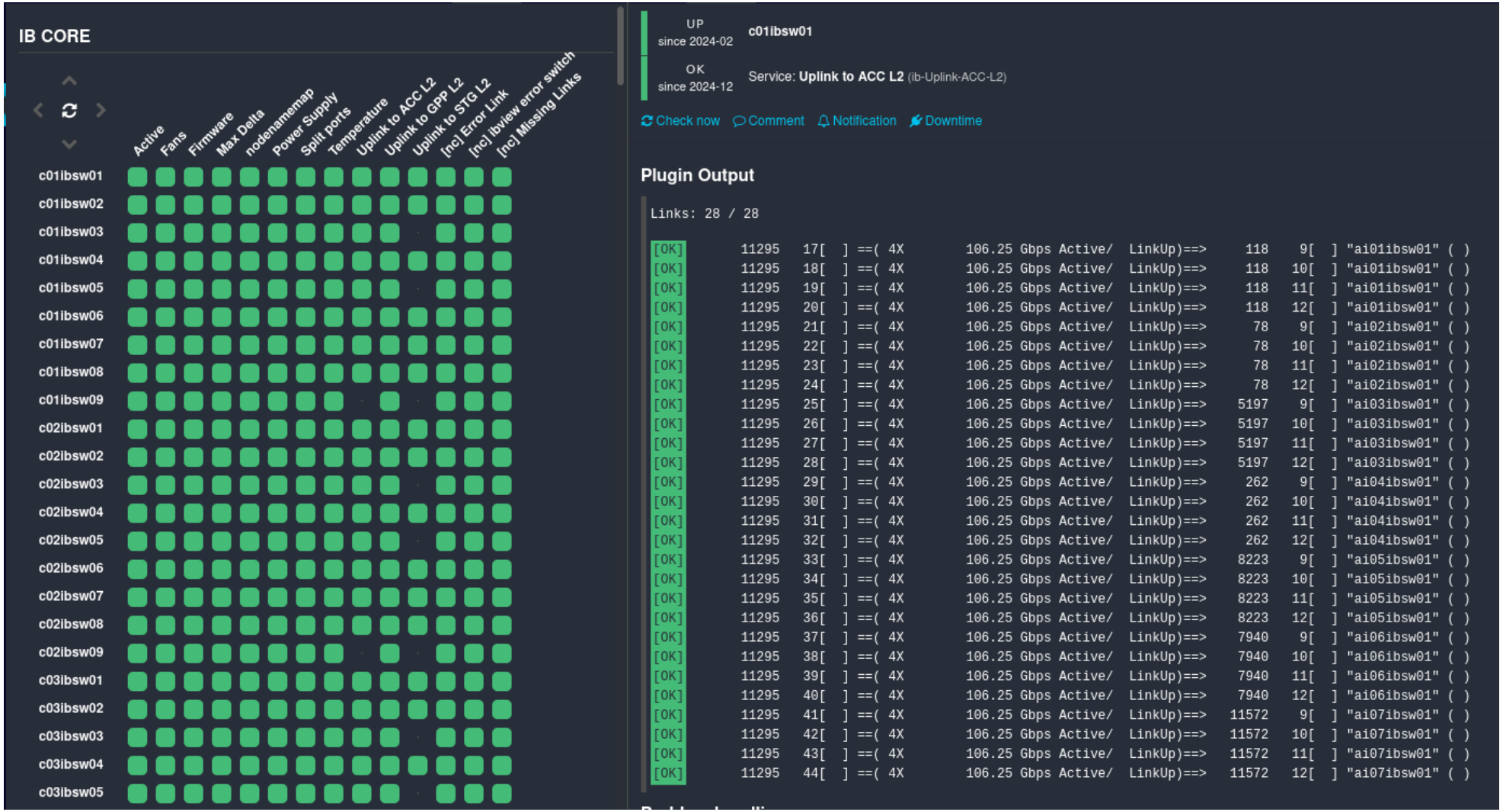


Heat to Cold Water Via Rear Door Heat Exchanger

IB01-IB04



Monitoring


IB CORE
Active
Fans
Firmware
Max Delta
nodenamemap
Power Supply
Split ports
Temperature
Uplink to ACC L2
Uplink to GPP L2
Uplink to STG L2
[nc] Error Link
[nc] ibview error switch
[nc] Missing Links
c01ibsw01
c01ibsw02
c01ibsw03
c01ibsw04
c01ibsw05
c01ibsw06
c01ibsw07
c01ibsw08
c01ibsw09
c02ibsw01
c02ibsw02
c02ibsw03
c02ibsw04
c02ibsw05
c02ibsw06
c02ibsw07
c02ibsw08
c02ibsw09
c03ibsw01
c03ibsw02
c03ibsw03
c03ibsw04
c03ibsw05
UP
since 2024-02
c01ibsw01
OK
since 2024-12
Service: Uplink to ACC L2 (ib-Uplink-ACC-L2)
Check now
Comment
Notification
Downtime
Plugin Output
Links: 28 / 28
[OK] 11295 17[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 118 9[ ] "ai01ibsw01" ( )
[OK] 11295 18[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 118 10[ ] "ai01ibsw01" ( )
[OK] 11295 19[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 118 11[ ] "ai01ibsw01" ( )
[OK] 11295 20[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 118 12[ ] "ai01ibsw01" ( )
[OK] 11295 21[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 78 9[ ] "ai02ibsw01" ( )
[OK] 11295 22[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 78 10[ ] "ai02ibsw01" ( )
[OK] 11295 23[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 78 11[ ] "ai02ibsw01" ( )
[OK] 11295 24[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 78 12[ ] "ai02ibsw01" ( )
[OK] 11295 25[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 5197 9[ ] "ai03ibsw01" ( )
[OK] 11295 26[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 5197 10[ ] "ai03ibsw01" ( )
[OK] 11295 27[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 5197 11[ ] "ai03ibsw01" ( )
[OK] 11295 28[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 5197 12[ ] "ai03ibsw01" ( )
[OK] 11295 29[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 262 9[ ] "ai04ibsw01" ( )
[OK] 11295 30[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 262 10[ ] "ai04ibsw01" ( )
[OK] 11295 31[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 262 11[ ] "ai04ibsw01" ( )
[OK] 11295 32[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 262 12[ ] "ai04ibsw01" ( )
[OK] 11295 33[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 8223 9[ ] "ai05ibsw01" ( )
[OK] 11295 34[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 8223 10[ ] "ai05ibsw01" ( )
[OK] 11295 35[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 8223 11[ ] "ai05ibsw01" ( )
[OK] 11295 36[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 8223 12[ ] "ai05ibsw01" ( )
[OK] 11295 37[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 7940 9[ ] "ai06ibsw01" ( )
[OK] 11295 38[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 7940 10[ ] "ai06ibsw01" ( )
[OK] 11295 39[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 7940 11[ ] "ai06ibsw01" ( )
[OK] 11295 40[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 7940 12[ ] "ai06ibsw01" ( )
[OK] 11295 41[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 11572 9[ ] "ai07ibsw01" ( )
[OK] 11295 42[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 11572 10[ ] "ai07ibsw01" ( )
[OK] 11295 43[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 11572 11[ ] "ai07ibsw01" ( )
[OK] 11295 44[ ] ==(4X 106.25 Gbps Active/ LinkUp)==> 11572 12[ ] "ai07ibsw01" ( )

BSC-CNS Networks
Thank you!

Barcelona
maria.manco@bsc.es
Network & Security
SIG-NOC 2025

OCT 15th 2025