

An abstract graphic on the left side of the slide, composed of numerous thin, light green lines that curve and overlap to form a complex, organic shape resembling a stylized flower or a tunnel. The lines are set against a solid dark blue background.

Open. Together.

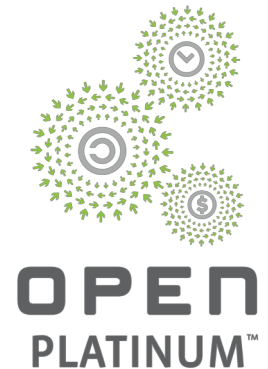


OCP
SUMMIT

Networking

Minipack, Low Power Fabric Switch

Sami Khan, Technical Program Manager
Xu Wang, Hardware Engineer
Hao Shen, Hardware Engineer
Facebook, Inc.



Open. Together.

Agenda

Minipack HW Overview

Minipack HW Deep Dive

MiniLake HW Deep Dive

Minipack – Next-Generation 128x 100G Switch

- **Adopt cutting edge switch ASIC**
- **Support mature optics – 100G CWDM4**
- **Lower Power / Smaller Size**

- **1/2 height, ~1/2 power** compared to existing Tomahawk based design – Backpack



Minipack

Switch ASIC: Broadcom Tomahawk-3

Size: 4RU (vs. 8RU Backpack)

Power: ~1.4KW budgetary (At full line rate, fully populated with 128x QSFP28 CWDM4 optics)

Radix: 128x



NETWORKING



Specifications



Design Files



Open. Together.

Minipack System Components

- ➔ Switch Main Board (SMB)
- ➔ System Control Module (SCM) - FRU
- ➔ Port Interface Module (PIM) - FRU
 - PIM-16Q: Port Interface Module with 16 x QSFP28 100G
 - PIM-4DD: Port Interface Module with 4 x QSFP56-DD 400G
- ➔ Fan Control Module (FCM): Top and Bottom
- ➔ Power Distribution System
 - Horizontal Bus Bar (HBAR)
 - Power Distribution Board (PDB): Left and Right
- ➔ DOM (Digital Optics Monitoring) FPGA



NETWORKING



Specifications

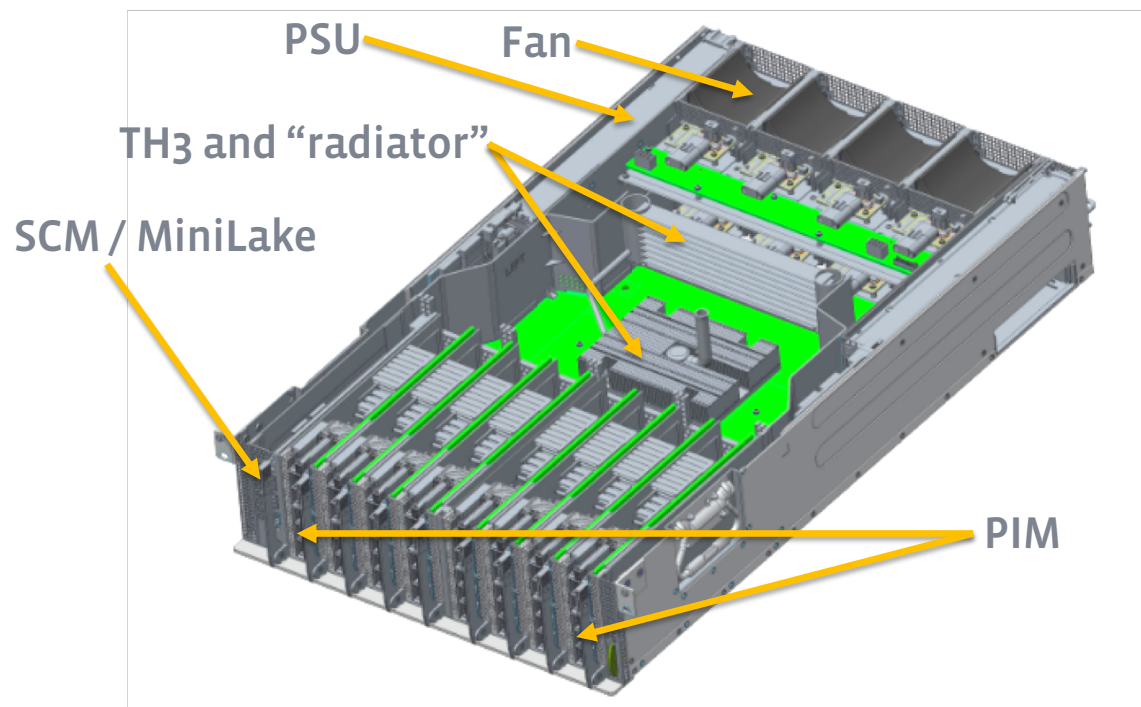


Design Files



Open. Together.

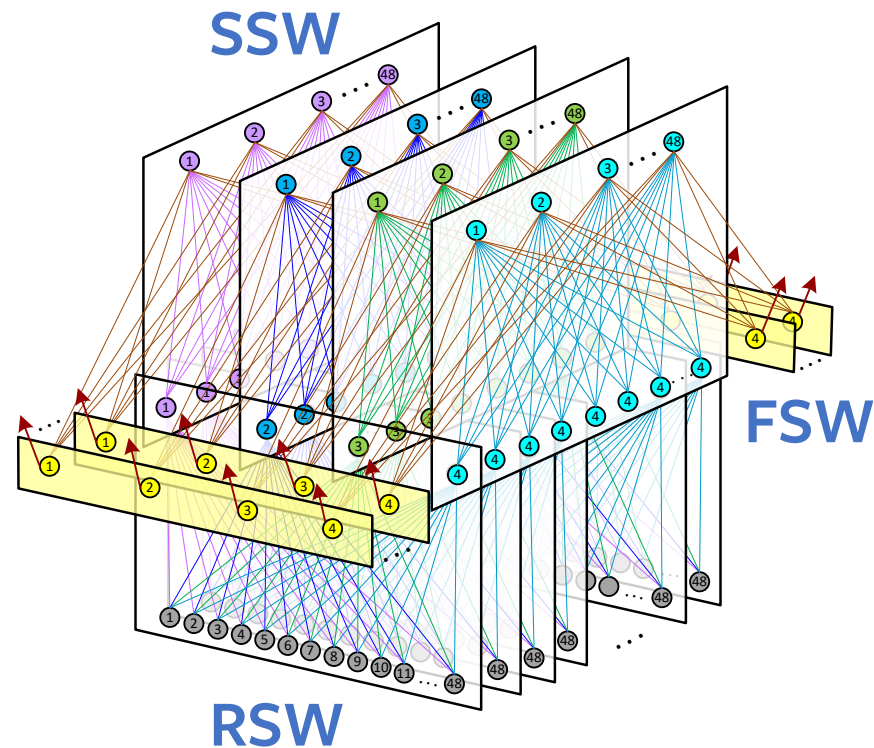
Minipack Chassis Architecture



- Orthogonal-direct architecture
 - Opens up airflow path for better thermal efficiency
 - Supports 100G CWDM4-Lite optics with 55 °C case temperature limit
 - Shortens PCB traces for lower loss
- FRU-able, modular PIM (line card)
 - Vertically oriented
 - PIM-16Q and PIM-4DD
 - Easy to explore other PIM options
- FRU-able SCM (micro-server carrier)

Where to use Minipack

- Minipack
 - FSW – Fabric Switch
 - SSW – Spine Switch
 - FA – Fabric Aggregator
- Supports multiple generations of rack switches



NETWORKING



Specifications



Design Files

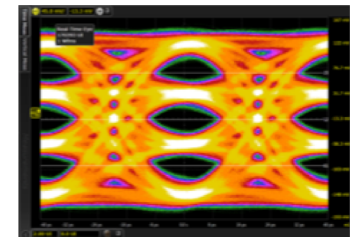
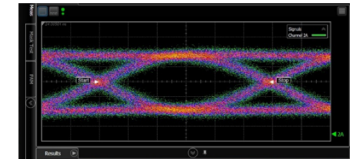
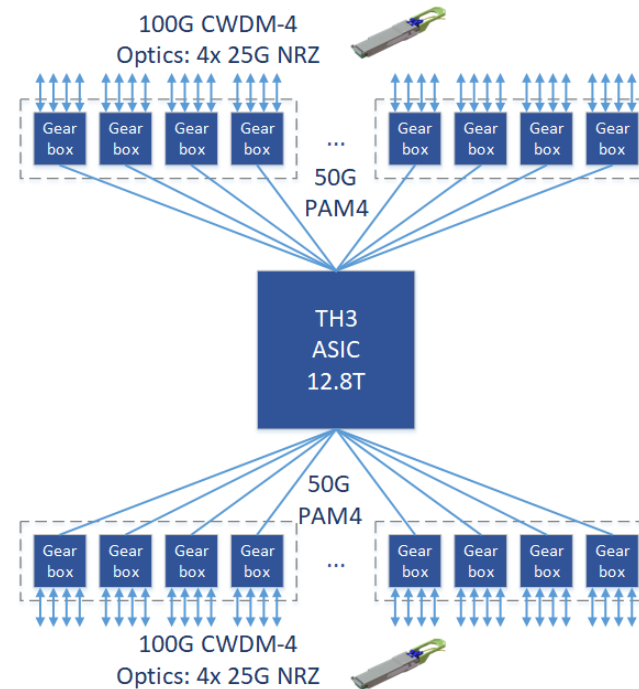
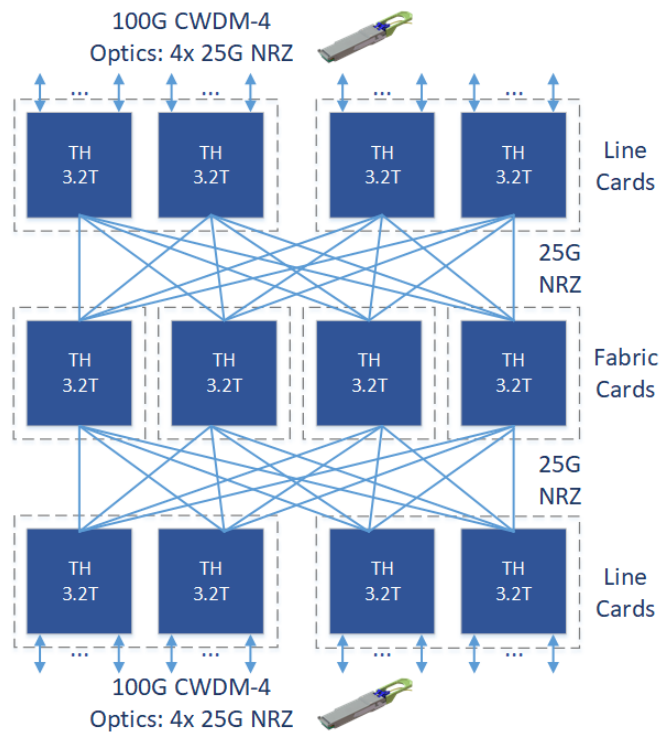
Agenda

Minipack HW Overview

Minipack HW Deep Dive

MiniLake HW Deep Dive

From Backpack to Minipack



Minipack Functional Block Diagram

→ SCM

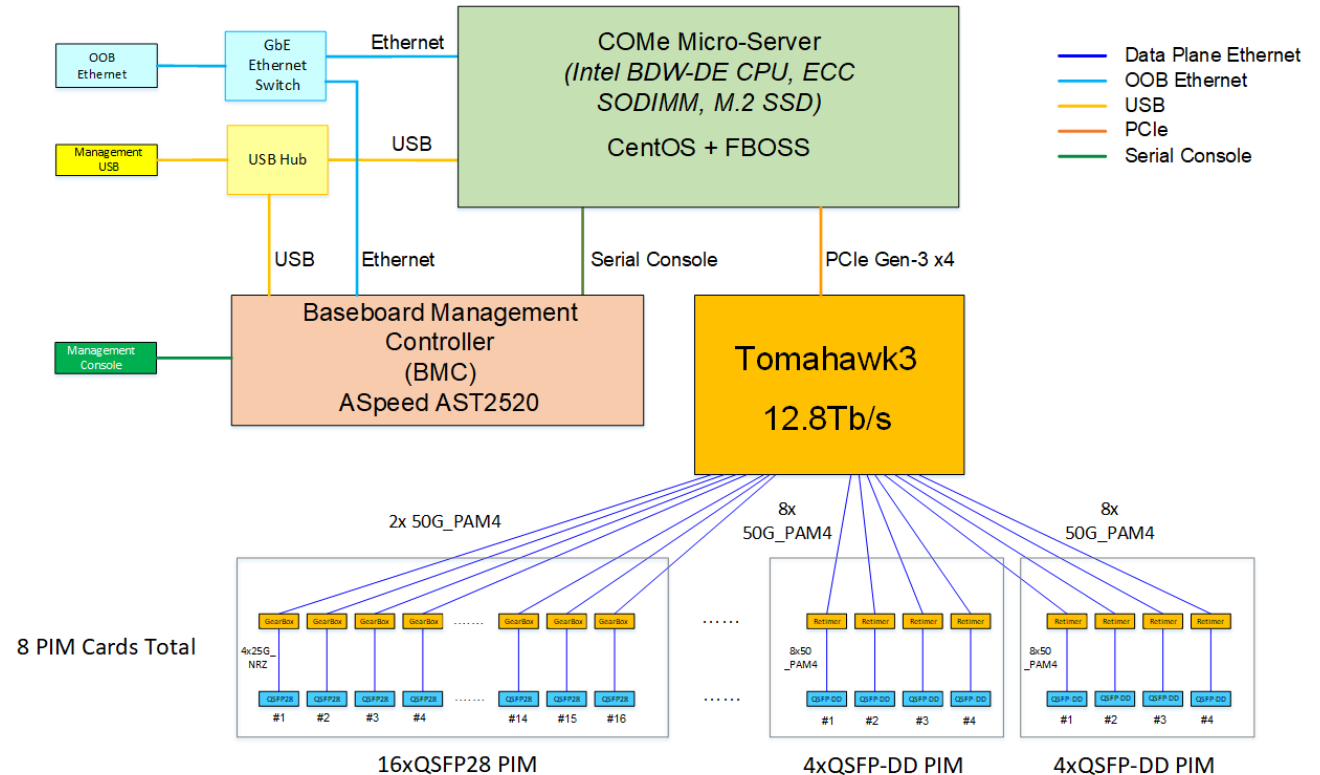
- Mini-Lake

→ SMB

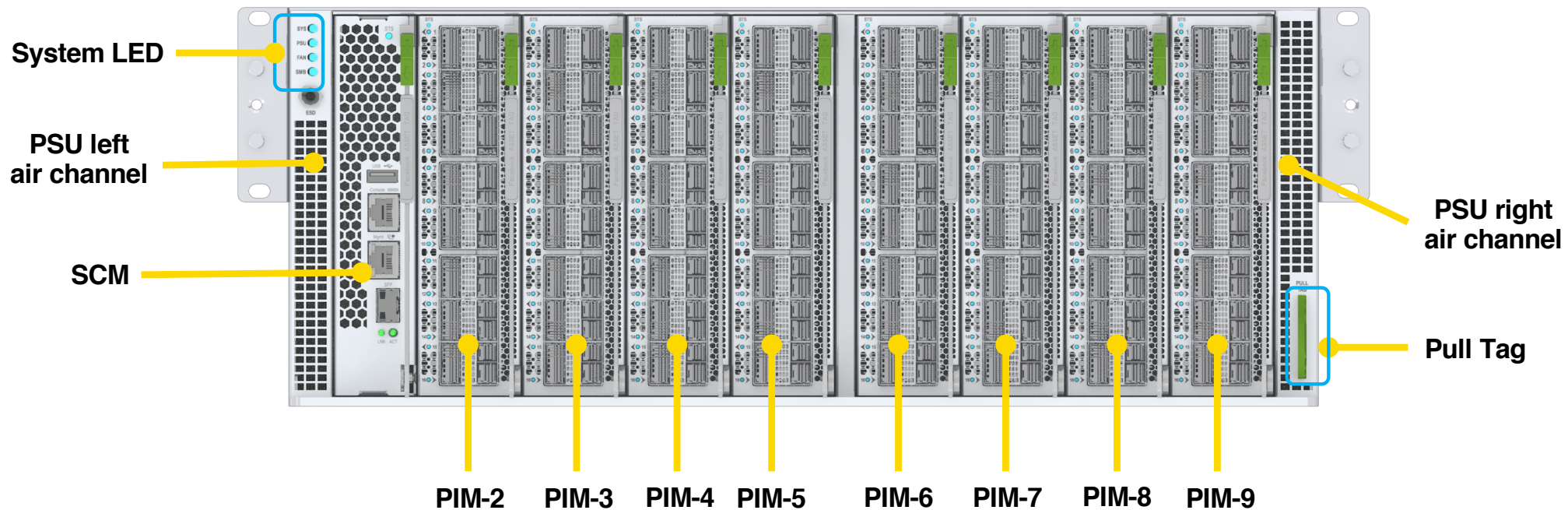
- TH3
- BMC

→ PIM

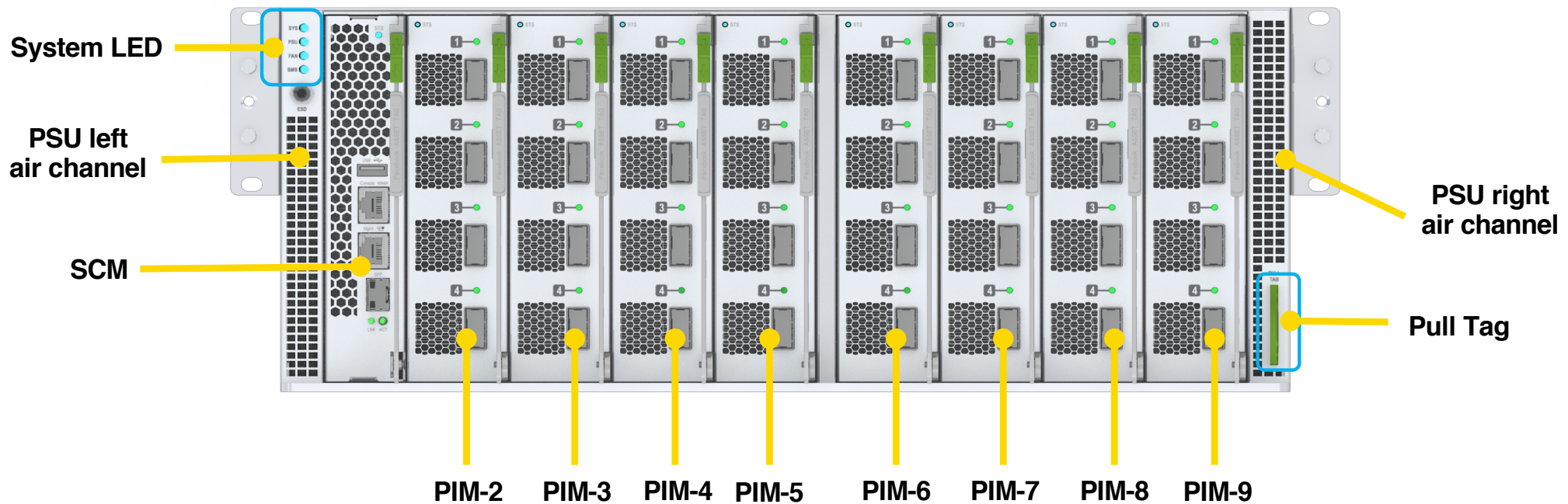
- PIM-16Q
- PIM-4DD



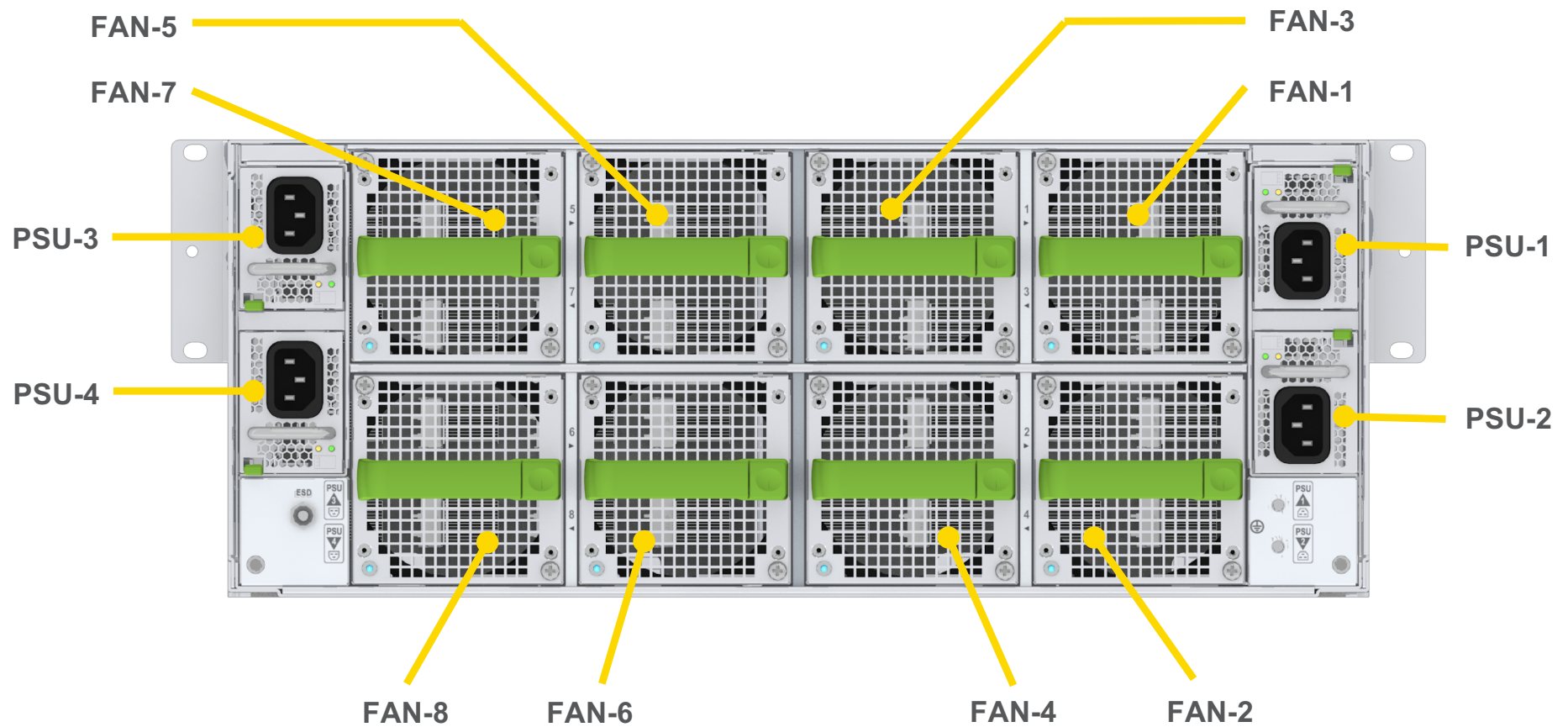
Minipack Front View with PIM-16Q



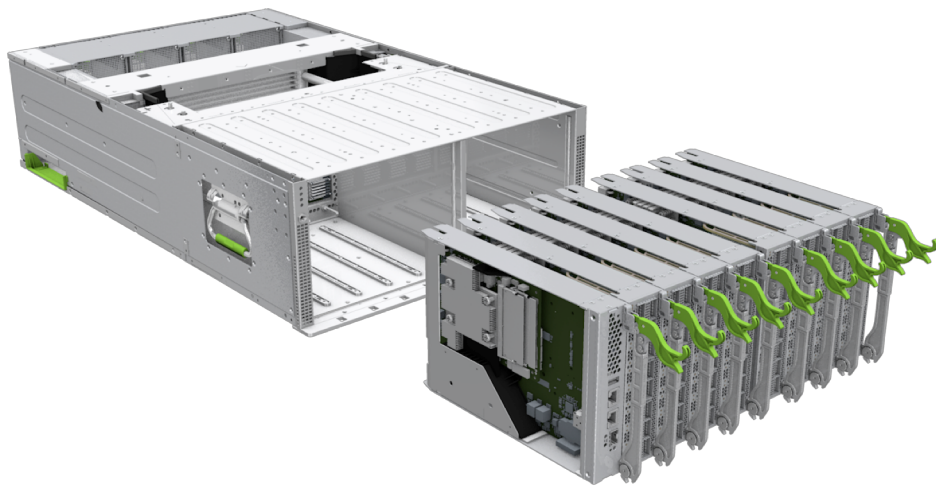
Minipack Front View with PIM-4DD



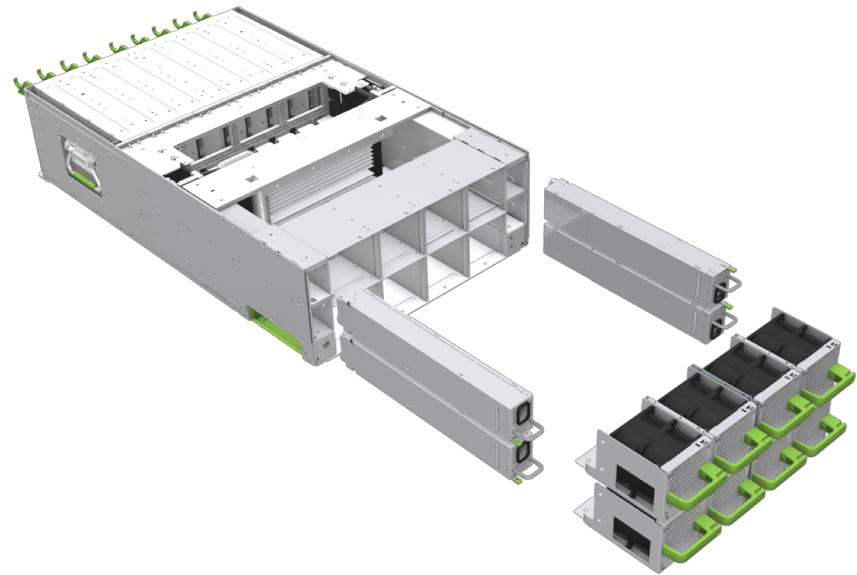
Minipack Rear View



Minipack FRUs

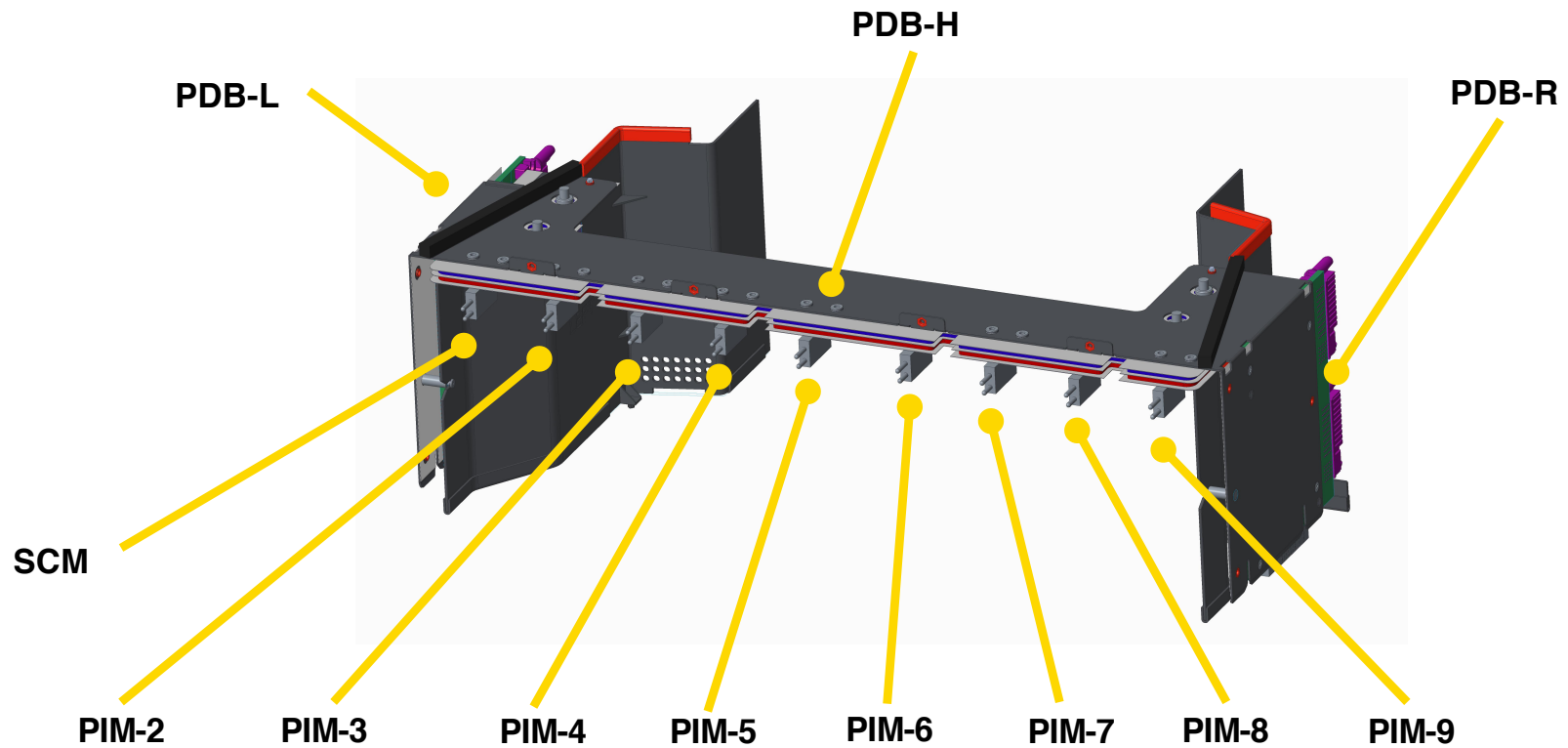


Front FRU



Rear FRU

Minipack 12V Bus-bar



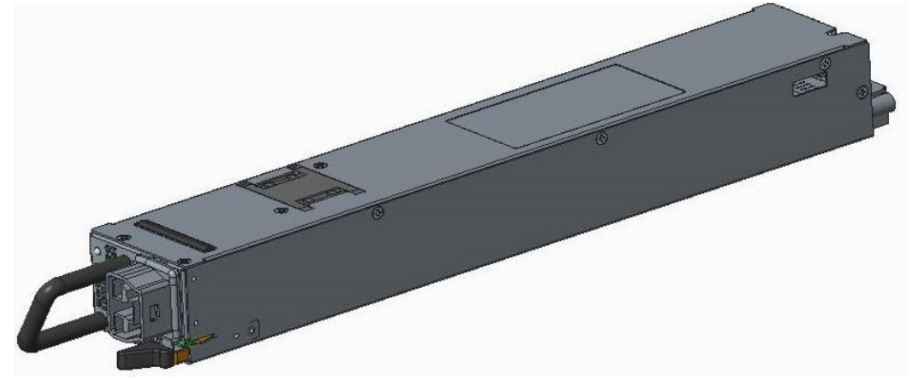
Minipack PSU

BelPower PFE1500



- Production now
- IEC-C16 AC inlet
- Single fuse: L-N
- Efficiency: 94%
- Input range: 208VAC – 230VAC

Delta HVAC



- NPI now
- HVAC inlet (Anderson)
- Dual fuse: L-L and L-N
- Efficiency: 96%
- Input range: 208VAC – 277VAC

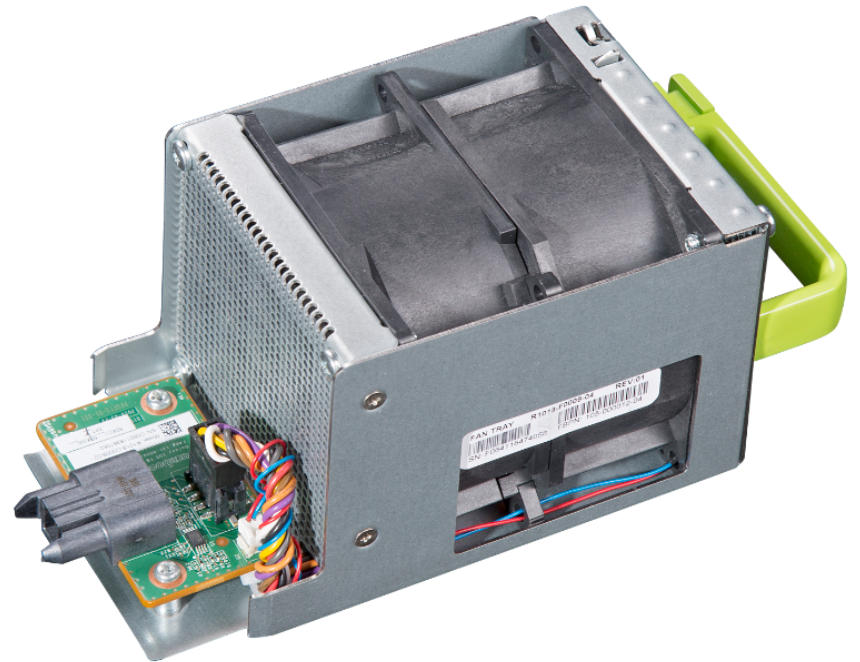
Minipack Power Consumption

<i>Inlet Temp °C</i>	20	23	26	29	31	33	35	37	39	N/A
<i>Fan PWM</i>	25%	29%	32%	35%	38%	41%	45%	48%	51%	100%
<i>Power (W)</i>	1271	1277	1283	1292	1295	1308	1325	1341	1356	1970

- ➔ Full line rate on all 128x 100G ports
- ➔ PSU AC input power. The unit is Watt
- ➔ All 8 PIM slots are populated with PIM-16Q and 3.5W electrical loopback modules

Minipack Fan Tray

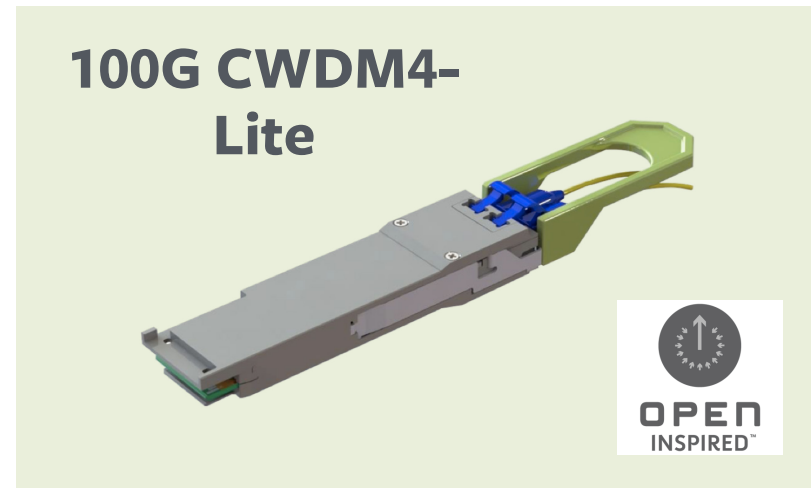
- Same as Backpack
- Screw-less latch for easy removal
- Powerful 80 x 80 x 80 mm counter-rotating fan
- Hot swappable
- LED built-in
- Each FCM carries 4 fan-trays. Total 8 fan-trays in Minipack chassis
- Use more fans to achieve better system thermal efficiency



Supported Optical Transceivers

- ➔ PIM-16Q
 - 100G QSFP28 CWDM4 optics
 - 40G QSFP+ optics
 - 200G QSFP56 FR4 optics at left column of ports on each PIM-16Q

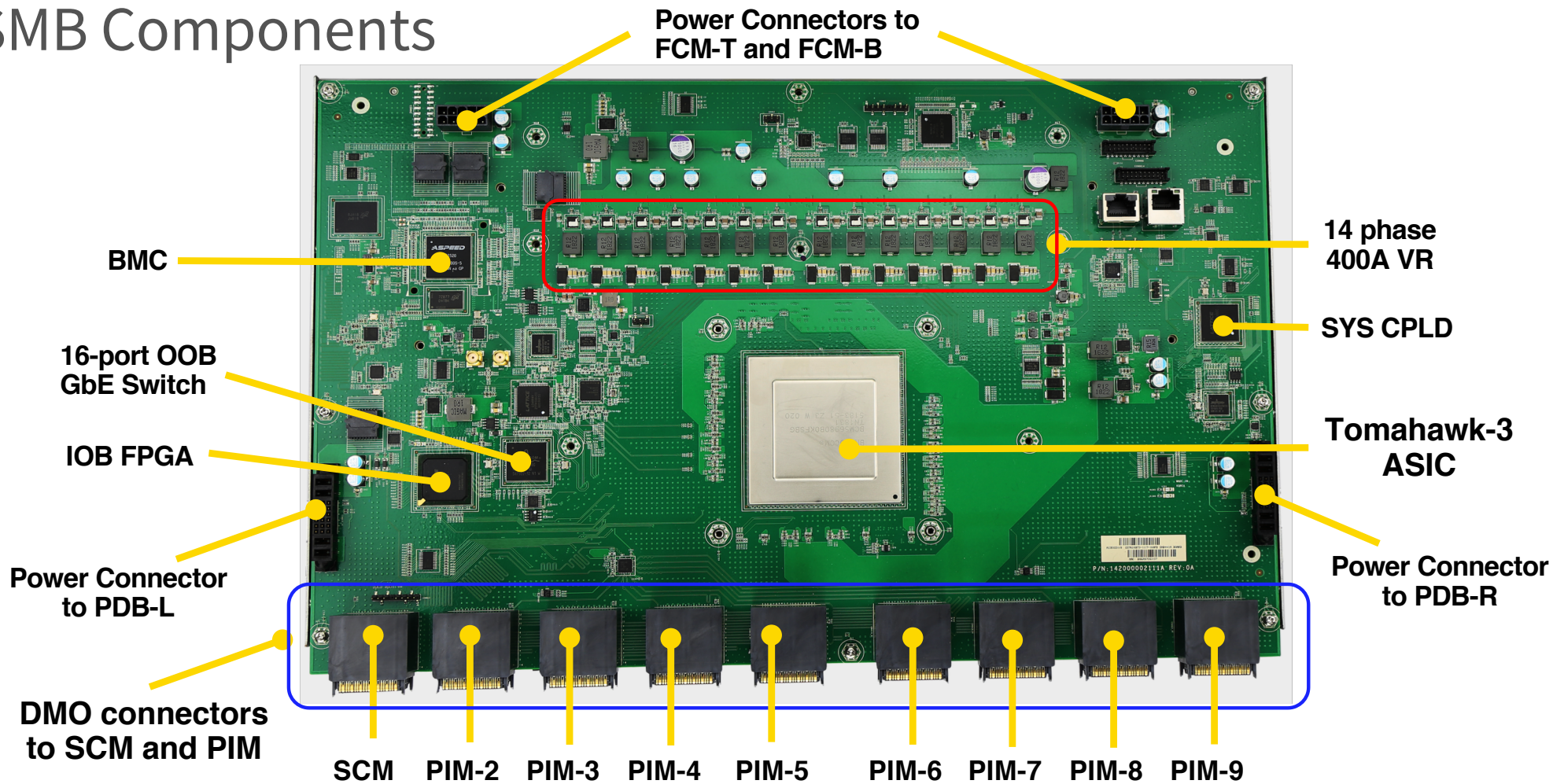
- ➔ PIM-4DD
 - 400G QSFP56-DD FR4 optics
 - 200G QSFP56 FR4 optics
 - 100G QSFP28 CWDM4 optics



Switch Main Board – SMB

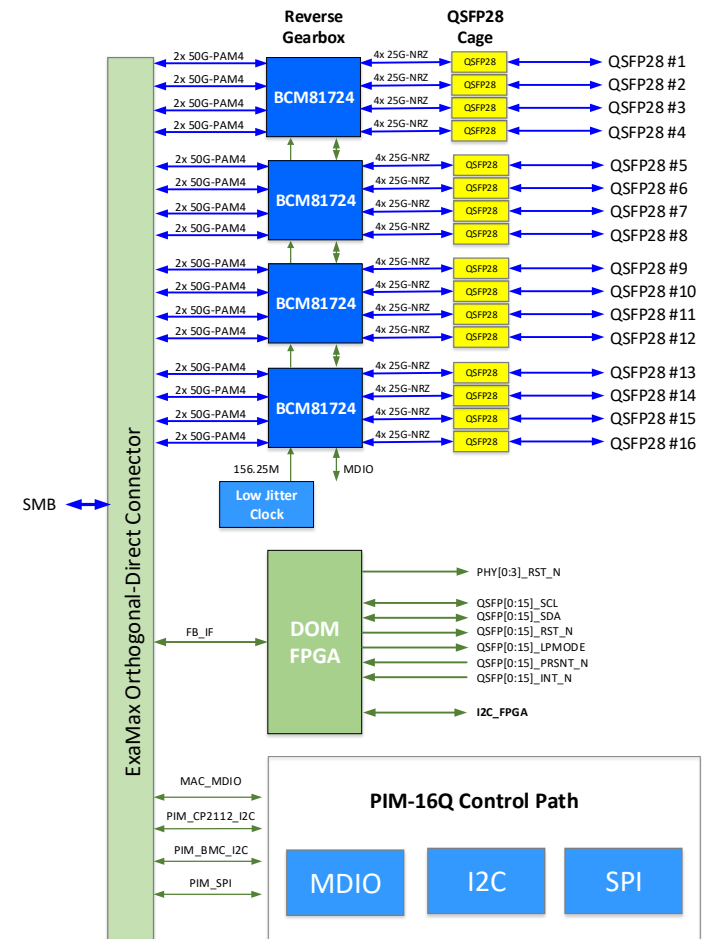
- TH3 12.8T switch ASIC
- BMC for SMB and chassis HW management
- 16-port switch for OOB / management GbE
- IOB FPGA as PCIe bridge to DOM (Digital Optics Monitoring) FPGAs
- Eight 6 x 12 DMO (Direct-Mate-Orthogonal) connectors to eight PIMs
- One 6 x 12 DMO connector to SCM

SMB Components

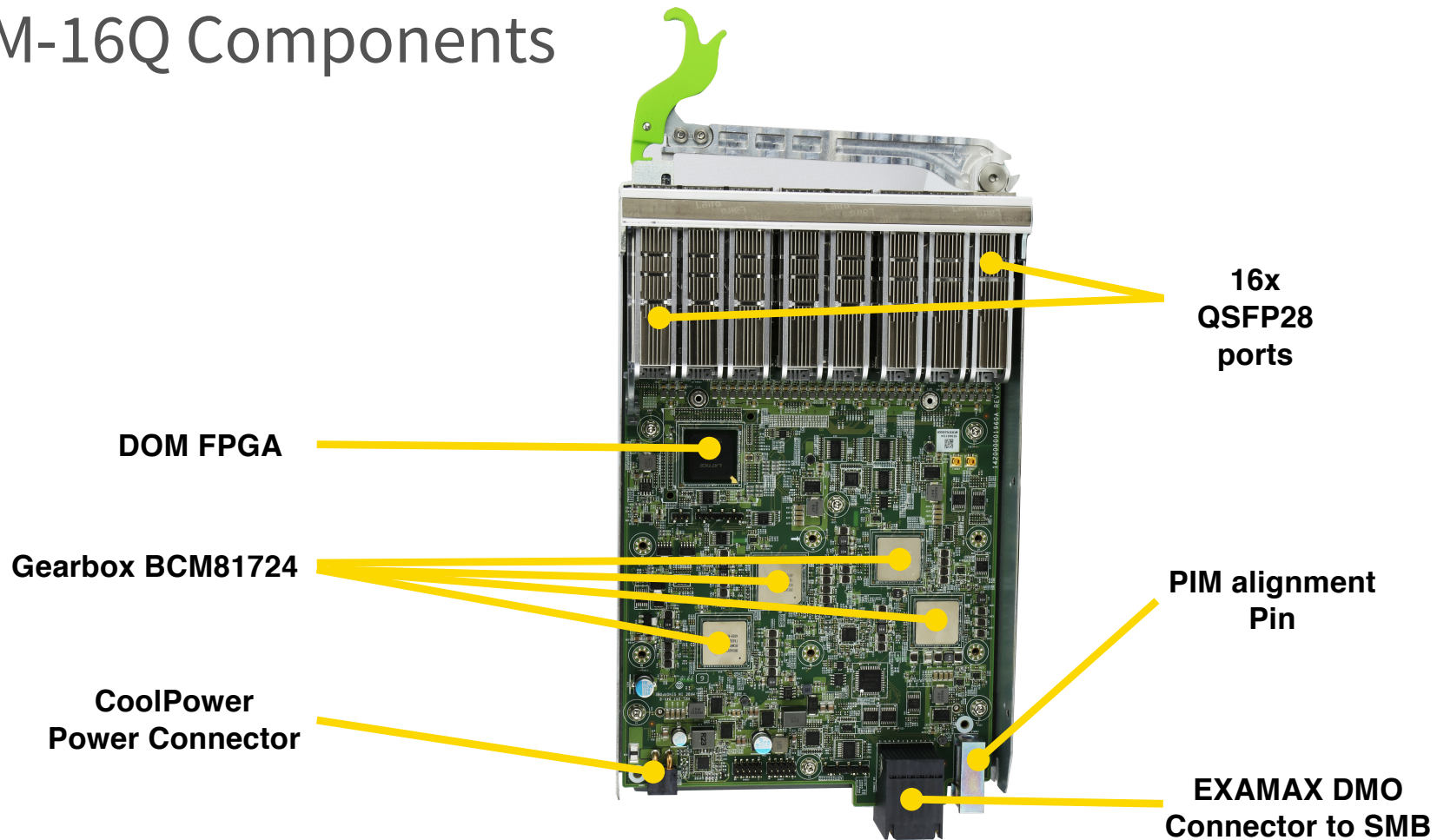


PIM-16Q Block Diagram

- ➔ Broadcom Reverse Gearbox BCM81724
- ➔ Translate 2x 50G PAM4 to 4x 25G NRZ
- ➔ RS-544 to RS-528 FEC translation (termination and re-generation)
- ➔ 16x QSFP28 100G ports
- ➔ DOM FPGA



PIM-16Q Components



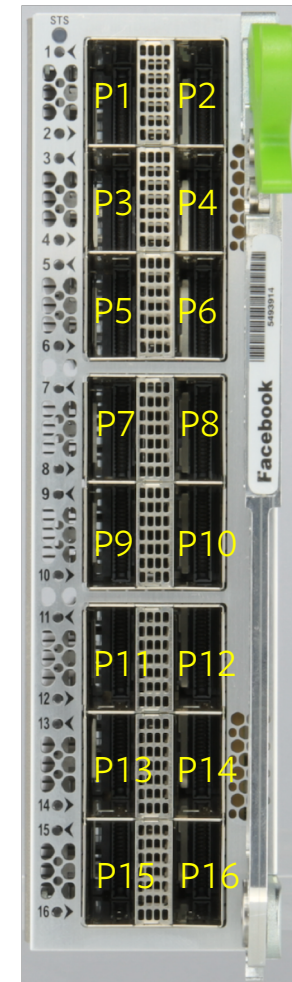
PIM-16Q – Supported Port Speeds

➔ Left column P1, P3, P5, P7, P9, P11, P13, P15

- 100G Mode
- 40G Mode
- **200G Mode**

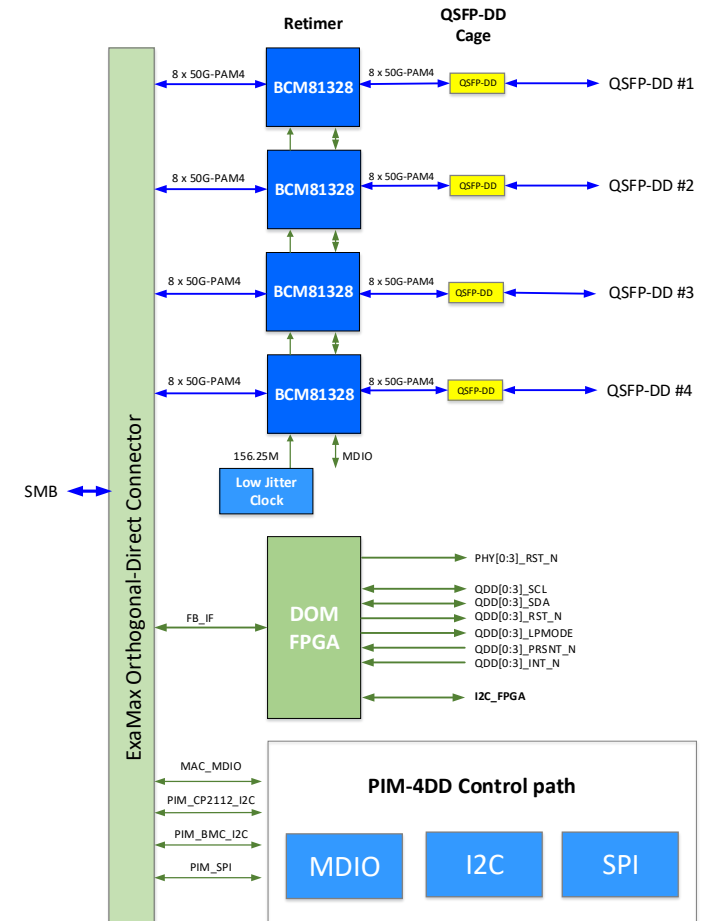
➔ Right column P2, P4, P6, P8, P10, P12, P14, P16

- 100G Mode
- 40G Mode



PIM-4DD Block Diagram

- ➔ Broadcom Retimer BCM81328
- ➔ Retiming 8x 50G PAM4 to 8x 50G PAM4
- ➔ No FEC termination and re-generation
- ➔ 4x QSFP56-DD 400G ports
- ➔ DOM FPGA



PIM-4DD Components

**EXAMAX DMO
Connector to SMB**

PIM Alignment Pin

Retimer Heatsink

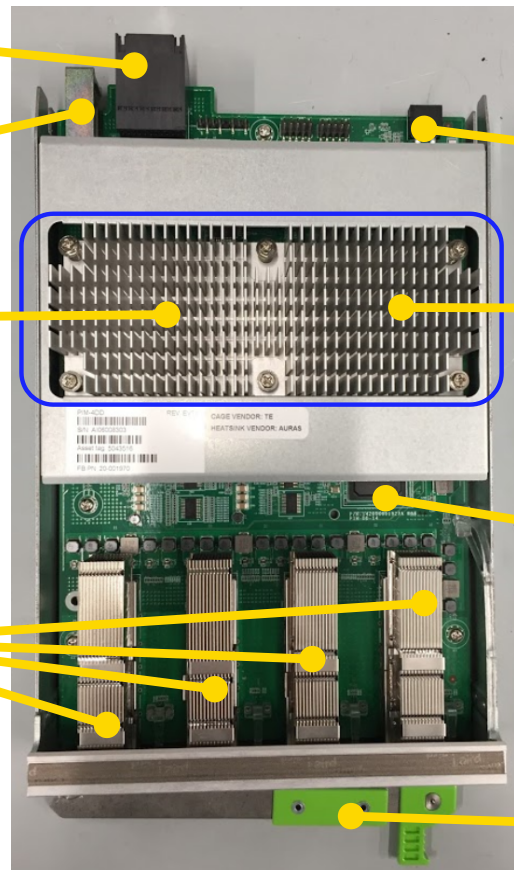
**4x QSFP-DD
Ports**

**CoolPower Power
Connector**

**Four BCM81328
Retimers**

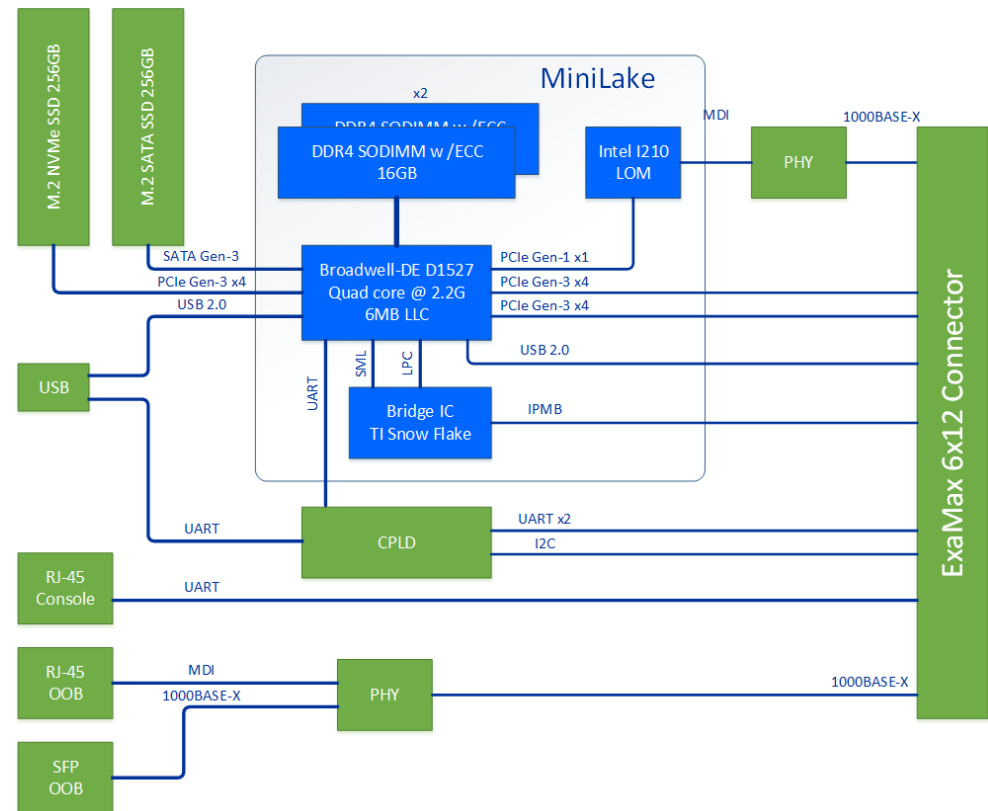
DOM FPGA

PIM Ejector

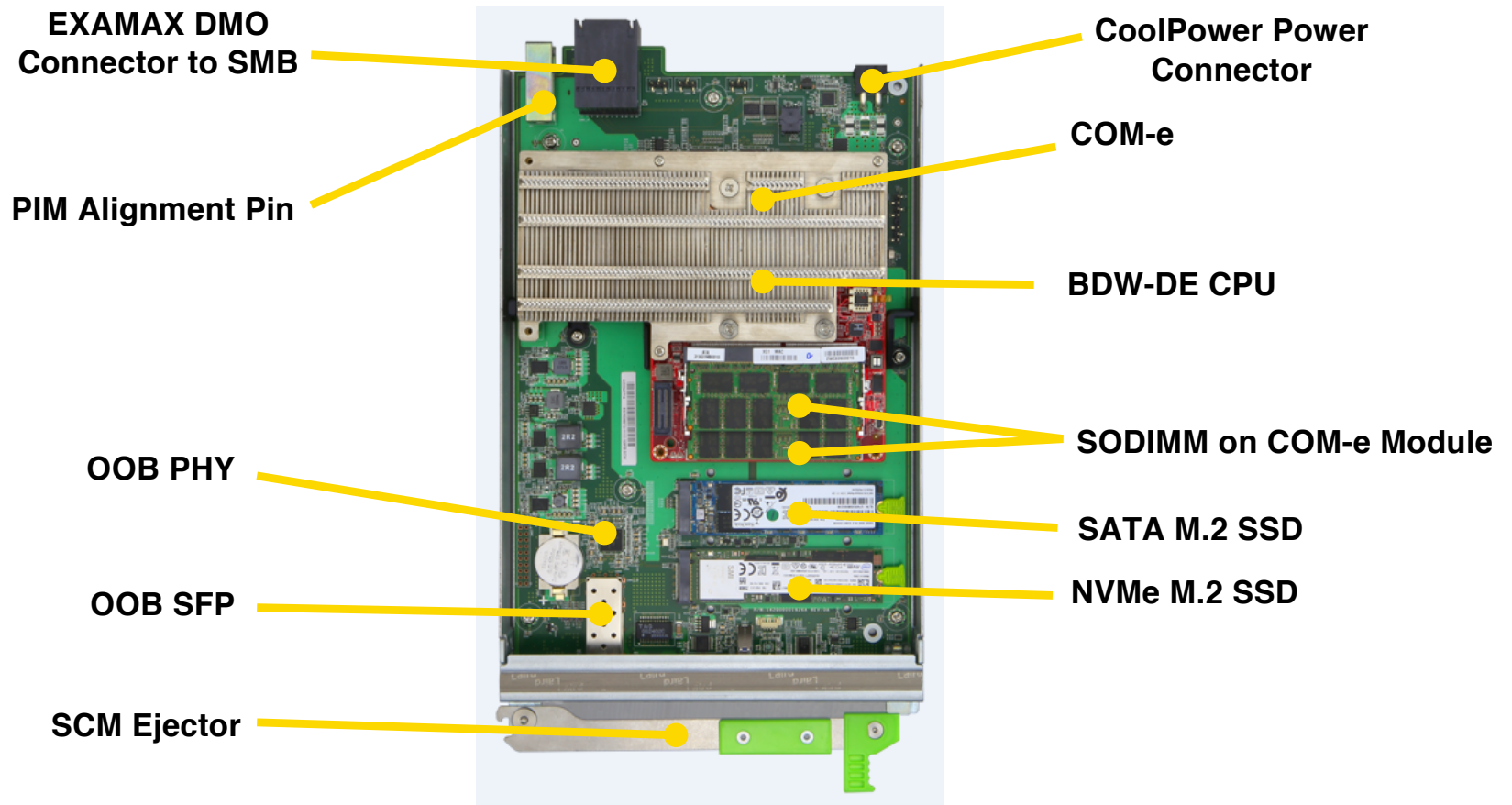


System Control Module – SCM

- ➔ Carrier board for MiniLake COM-e CPU module
- ➔ SATA and NVMe M.2 SSD
- ➔ External console, OOB (RJ-45 or SFP), and USB interfaces
- ➔ Supports Facebook USB debug dongle

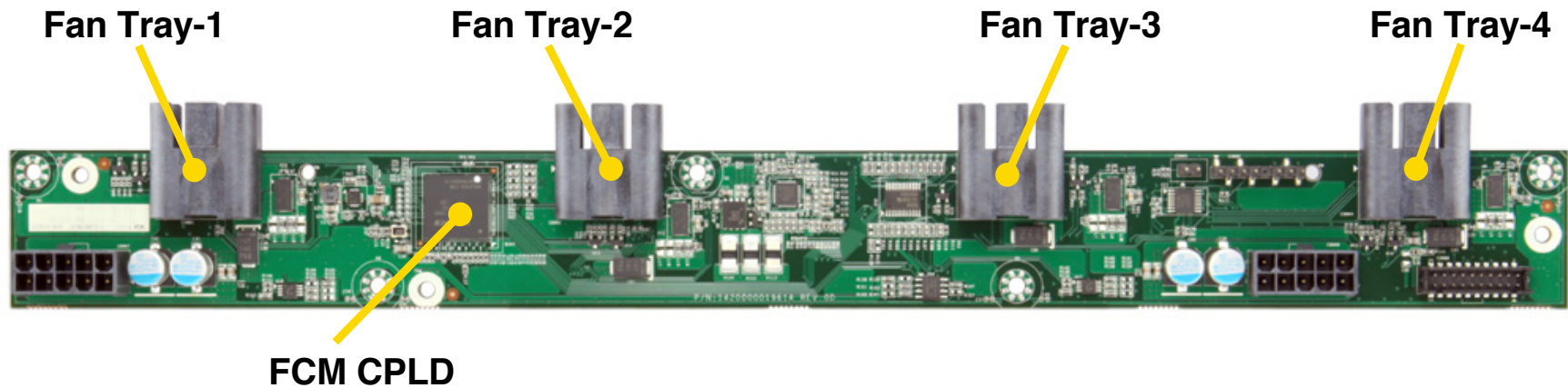


SCM Components



Fan Control Module – FCM

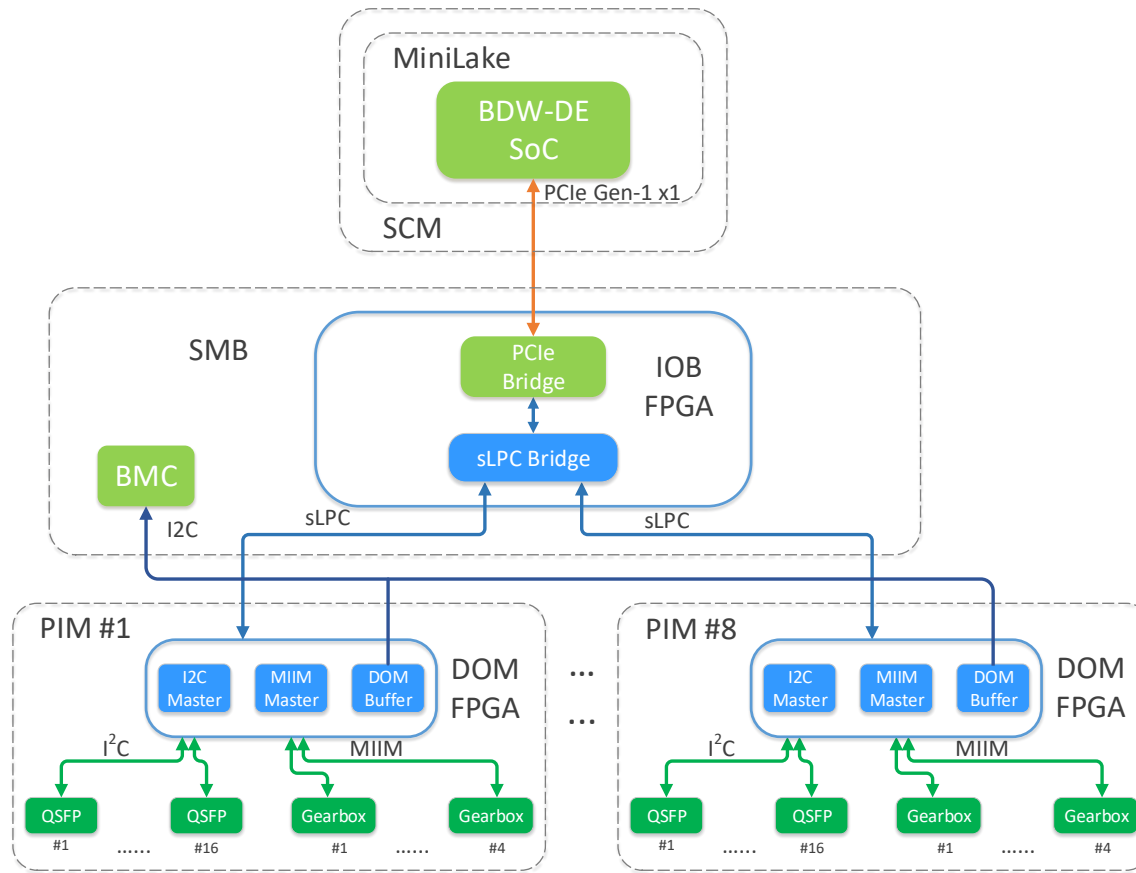
- ➔ Two FCMs in Minipack: FCM-T and FCM-B
- ➔ Each FCM supports 4 fan-trays
- ➔ BMC controls FCM CPLD via an I2C bus, which generates PWM signals to control fan speed



IOB and DOM FPGA

- Offloads CPU from direct I2C bus accesses to QSFP modules; Instead, CPU can read DOM information from FPGA buffers
- BMC also reads the optics temperature information through an I2C bus to FPGA buffers for the purpose of system fan control
- Provide MIIM (MDIO/MDC) master controllers for CPU to configure / monitor the gearbox chips on PIM-16Q (or retimers on PIM-4DD)

FPGA Block Diagram



Agenda

Minipack HW Overview

Minipack HW Deep Dive

MiniLake HW Deep Dive

MiniLake Introduction

- MiniLake is a COM-Express CPU Module
- MiniLake is currently implemented in Minipack Switch System
- Our Motivation to build MiniLake:
 1. Provide a compact solution with specified configuration
 2. Provide powerful IOs across different applications
 3. Better support on BIOS and OpenBMC features



NETWORKING



Specifications



Design Files



Open. Together.

MiniLake Module Design

- MiniLake Feature List

Spec	Description
Board size (W x L)	95mm x 125mm (COM Express® Basic Module form factor)
CPU	Xeon® Processor D-1527 : 4C, 35W, 2.2 GHz (Broadwell-DE SoC, Grangeville Platform)
Memory	Two DDR4 SODIMM slots
	•Support DDR4 ,speeds up to 2133 MT/s with ECC
	•Max Memory Capacity: 32 GB (16GB / DIMM)
Storage	One SATA 3.0 Port (6Gb/s)
Network	One 1000BASE-T (Intel® Ethernet Controller I210)
	Two 10GBASE-KR (Intel® Broadwell-DE integrated LAN controller)
PCIe	PCIe Gen3 x24 Lanes (up to Six PCIe Gen3 x4 ports)
	PCIe Gen2 x4 Lanes (up to Four PCIe Gen2 x1 ports)
	One PCIe CLK
Management Port	I2C is the primary sideband interface for server management functionality connection to BMC of carrier board
Bridge IC	TI Snowflake
I/O	Two USB2.0 I/F and Two USB3.0 (w/ USB2.0)
	One UART port for Console
	One SPI I/F for 2nd boot flash
	One LPC I/F



NETWORKING



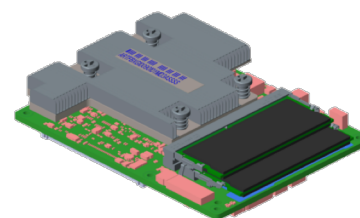
Specifications



Design Files



Open. Together.



NETWORKING



Specifications



Design Files

MiniLake Module Design

- Key Components



NETWORKING



Specifications

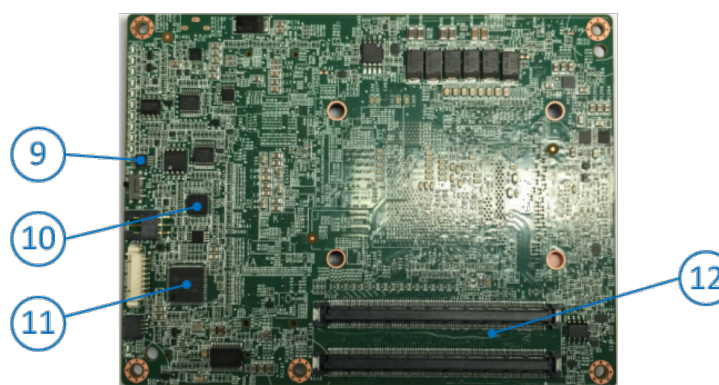


Design Files

TOP Side



BOT Side



Item	Component	Note
1	SOC (INTEL Processor D-1527)	Broadwell-DE (4C/35W/2.2GHz)
2	BIC (TI Snowflake)	Embedded Controller
3	Temperature Sensor-1 (TMP75)	I2C Address: 0x9E
4	CPLD (Altera EPM570M100C5N)	
5	BIOS Flash (with Socket)	
6	DDR4 SODIMM CONN-0	

Item	Component	Note
7	DDR4 SODIMM CONN-1	
8	XDP CONN	
9	Temperature Sensor-2 (TMP75)	I2C Address: 0x9C
10	1G NIC (INTEL i210)	
11	TPM2.0 (Infineon SLB9670)	
12	COMe B2B CONN	

MiniLake Module Design

- Power Budget



NETWORKING



Specifications



Design Files

Design Power Budget Estimation

Component	Sub power (W)	Quantity (pcs)	Utilization (%)	Total power (W)
CPU (BDX-DE)	35	1	85%	29.75
DDR4 SODIMM	2.4	2	85%	4.08
Embedded Controller (Snow flake)	1	1	85%	0.85
I210-AT	1	1	85%	0.85
Discrete IO	3	1	85%	2.55
VR Loss	10	1	85%	8.5
System total power Consumption Estimation				46.58

MiniLake Module Design

- Mechanical



NETWORKING

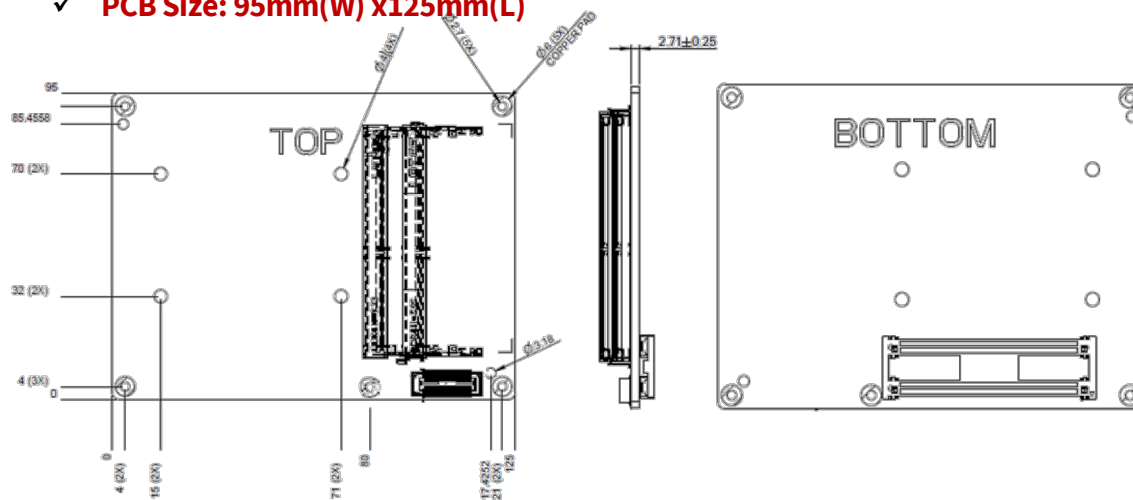


Specifications

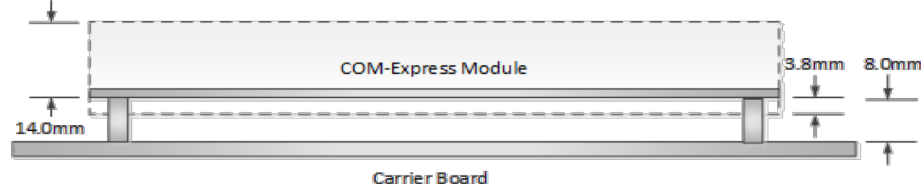


Design Files

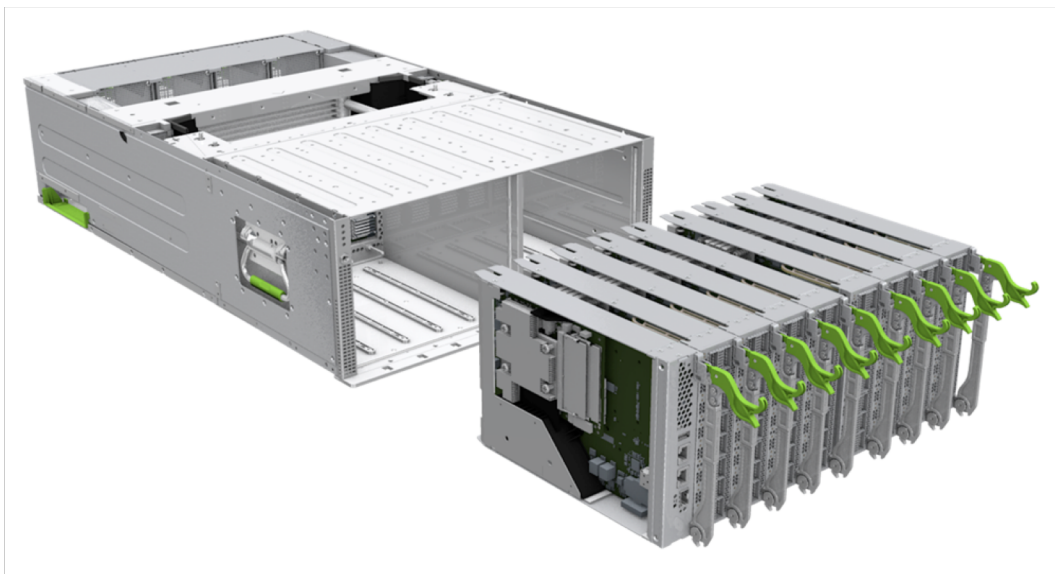
✓ **PCB Size: 95mm(W) x125mm(L)**



✓ **Clearance area on the Carrier Board.**



MiniLake in Minipack System



**Minipack System
12.8T Switch**



MiniLake

SCM: MiniLake Carrier



NETWORKING



Specifications



Design Files

Call to Action

- Contribution packages
 - EE design files
 - ME design files
 - OCP specifications
- Facebook's ODM partners
 - Edgecore: Minipack
 - Quanta: MiniLake



NETWORKING



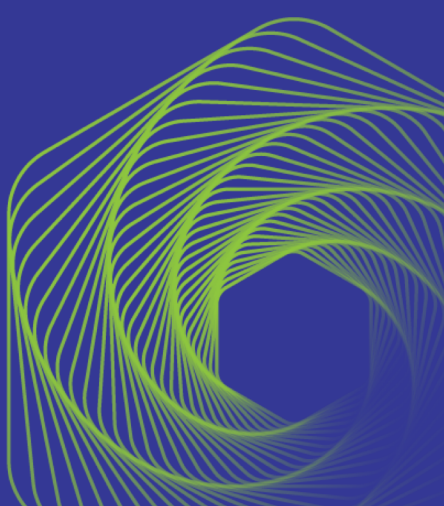
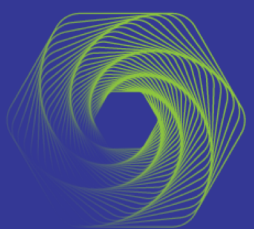
Specifications



Design Files



Open. Together.



Open. Together.

OCP Global Summit | March 14–15, 2019

