

# OPEN

Compute Project







# Backpack: Facebook's 100G Modular Switch

Deep Dive

**Zhiping Yao**

NETWORK HARDWARE ENGINEER, FACEBOOK



# Backpack System Introduction



Backpack

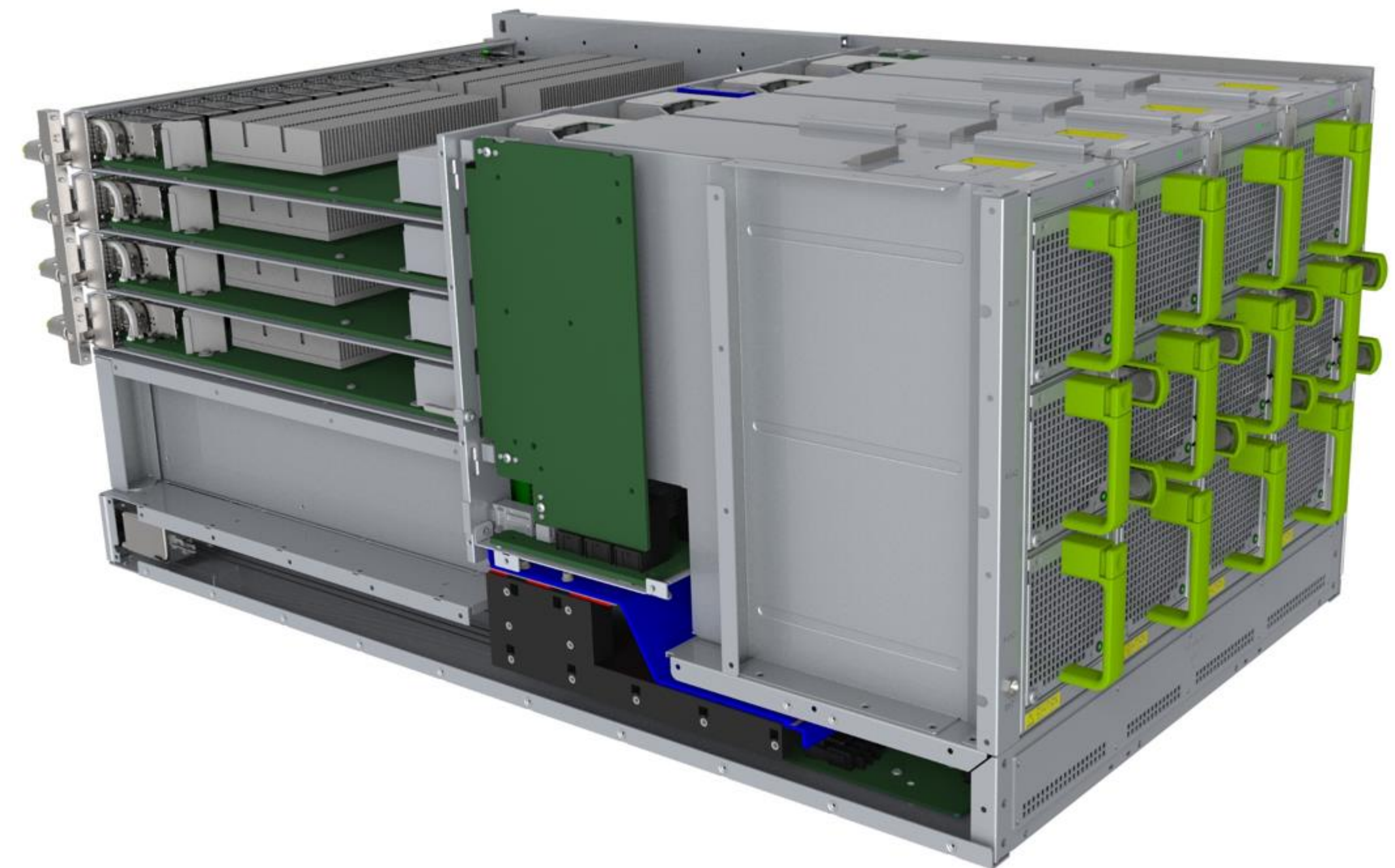
# Backpack : an open modular switch

- Hardware architecture: Ethernet only, fully open
- Network topology: dual stage spine-leaf
- Switch Software: FBOSS and OpenBMC
- Manageability: operated like server from BMC
- Density: 128 x QSFP28 100G support
- Building block: SWE (Switch Element)
  - Data plane: Tomahawk Switch ASIC
  - Control Plane: COM-e BayTrail CPU Module
  - Management Plane: BMC AST1250



# Backpack : FB Modular Switch Platform

- Orthogonal Direct Architecture
- Disaggregated data, control, and management plane design





# Innovative Design

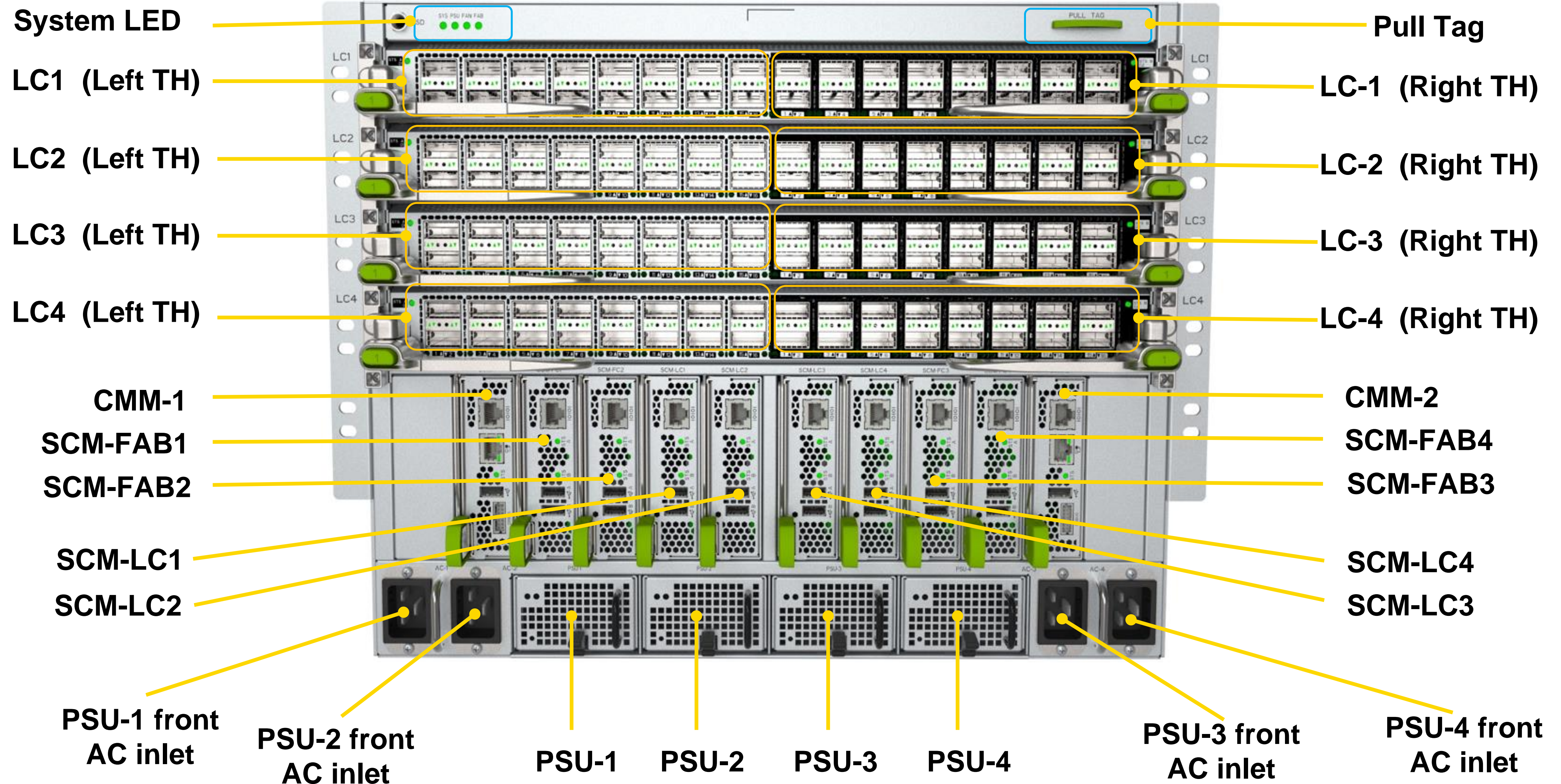
- Fully Disaggregated Architecture
  - 12 Switch Elements (SWE)
  - Separate SCM module for Control CPU
- An Orthogonal Direct Chassis Architecture
  - All major module cards are designed to be mated orthogonally
  - Open up more air channel for a better thermal performance
  - Reduce the PCB trace length for better signal integrity
  - Support future 8x16 OD connector to double port and speed
- A sophisticated thermal design to support low cost 55C CWDM4 optics

# Backpack System Components

- Line Cards (LC)
- Fabric Card (FAB)
- System Controller Module (SCM)
- Chassis Management Module (CMM)
- Horizontal Control Plane (HCP)
- Vertical Control Plane (VCP-L, VCP-R)
- Bus Bar Assembly (BBA)
  - Horizontal Bus Bar (HBAR)
  - Vertical Bus Bar (VBAR)
  - Horizontal Power Distribution Board (HPD)
- Fan Control Board (FCB)
- Power Distribution Board (PDB)

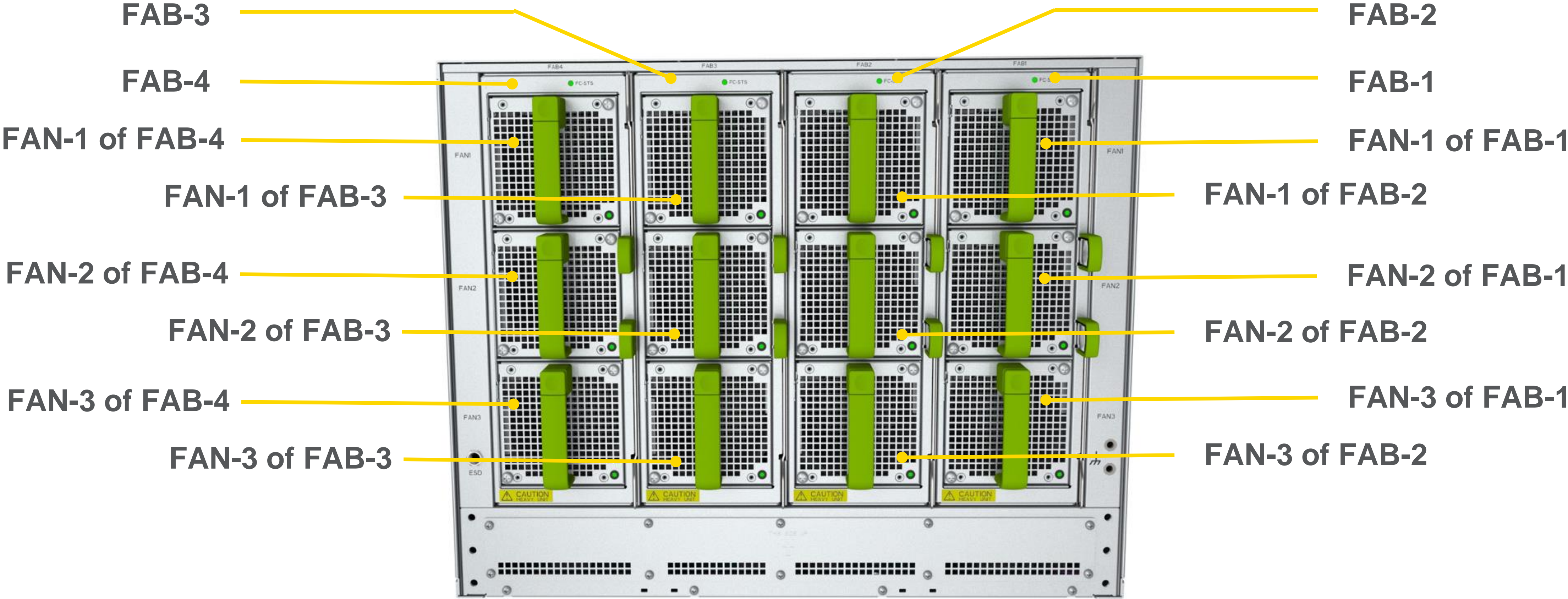


# Backpack Front View



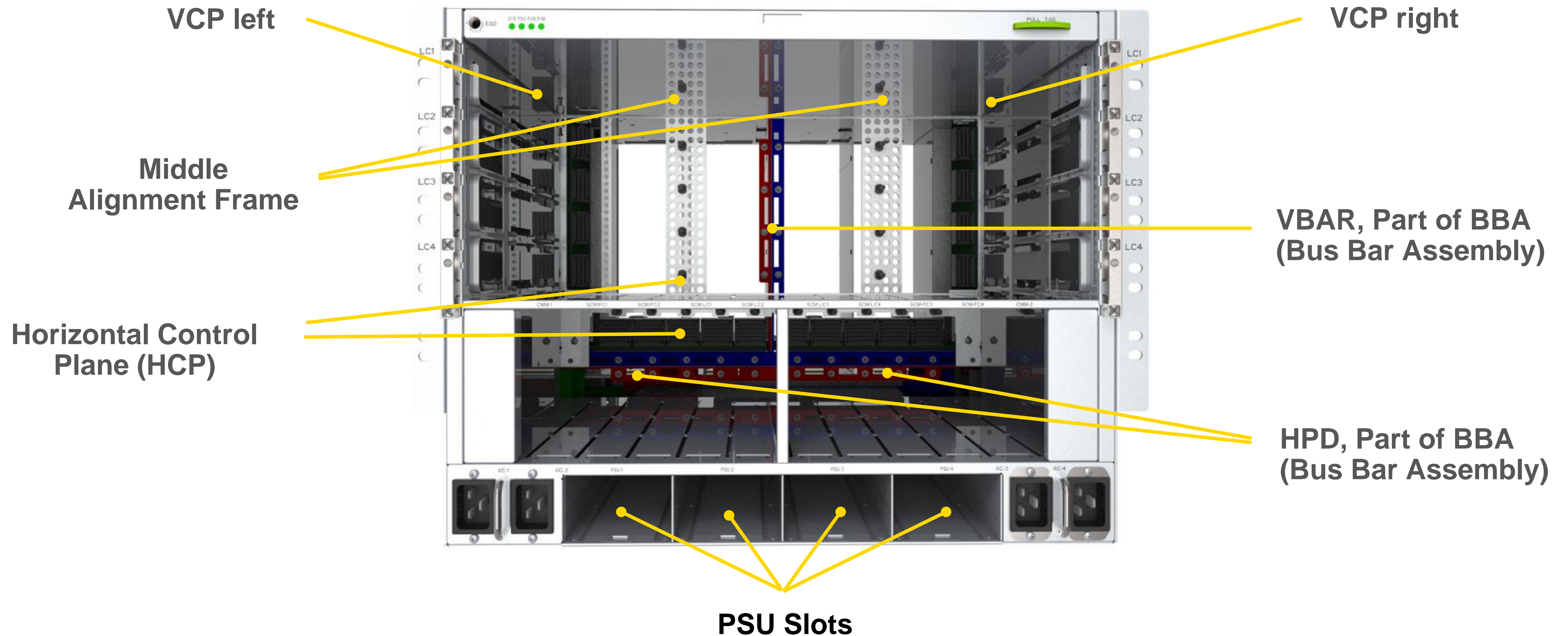


# Backpack Rear view



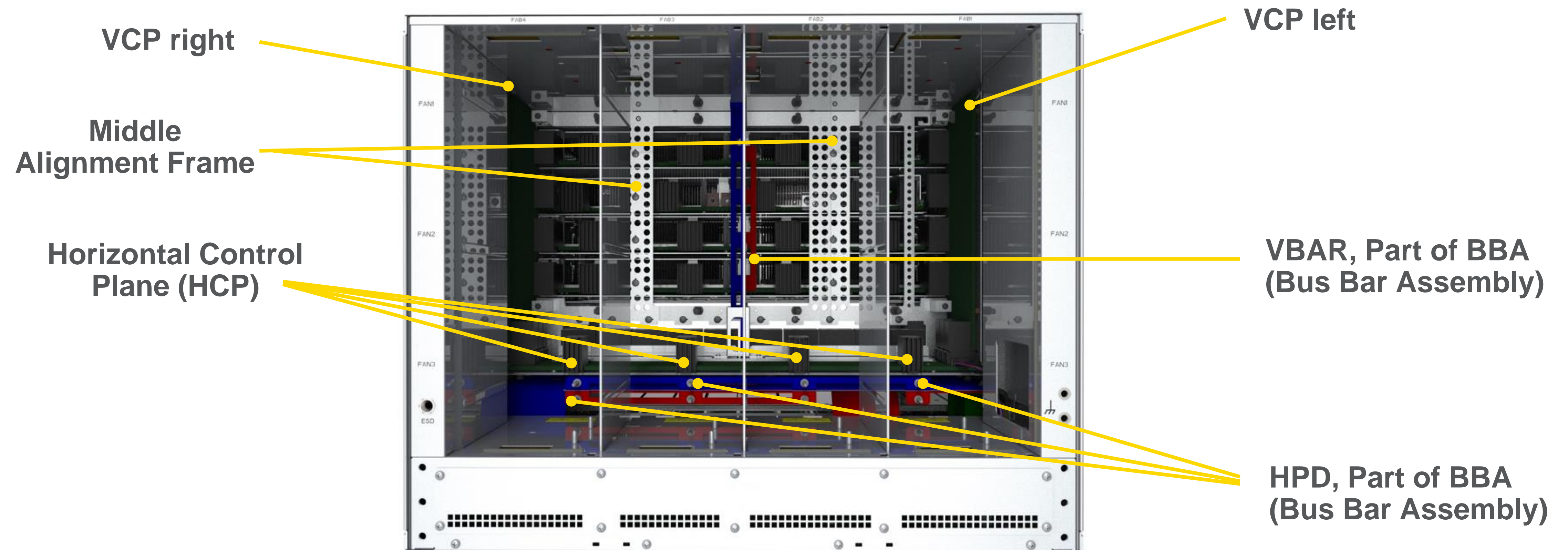


# Front view without LC and FAB



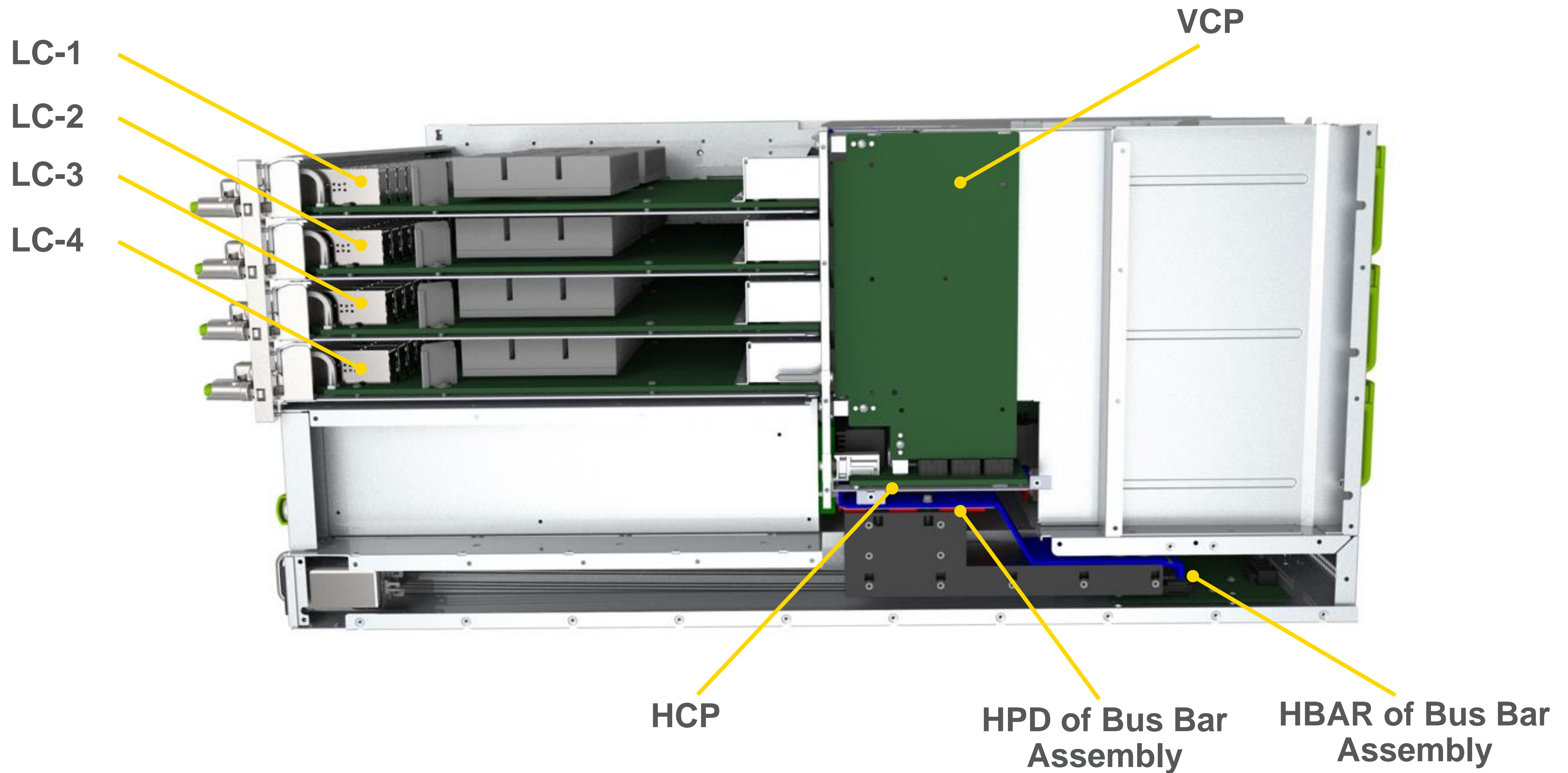


# Rear view without FAB





# Backpack Side View





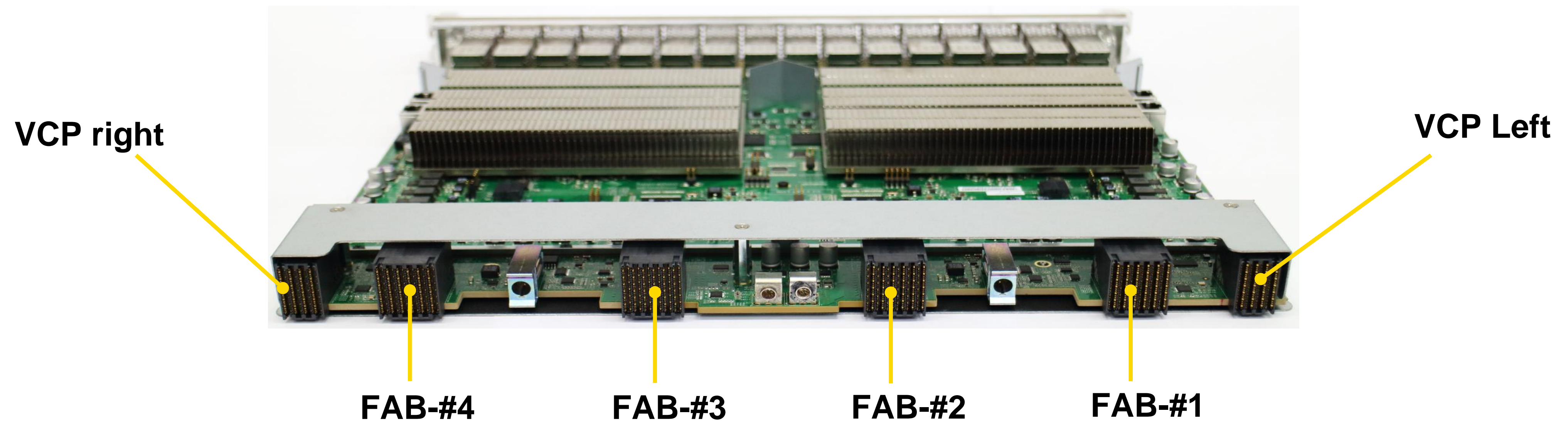
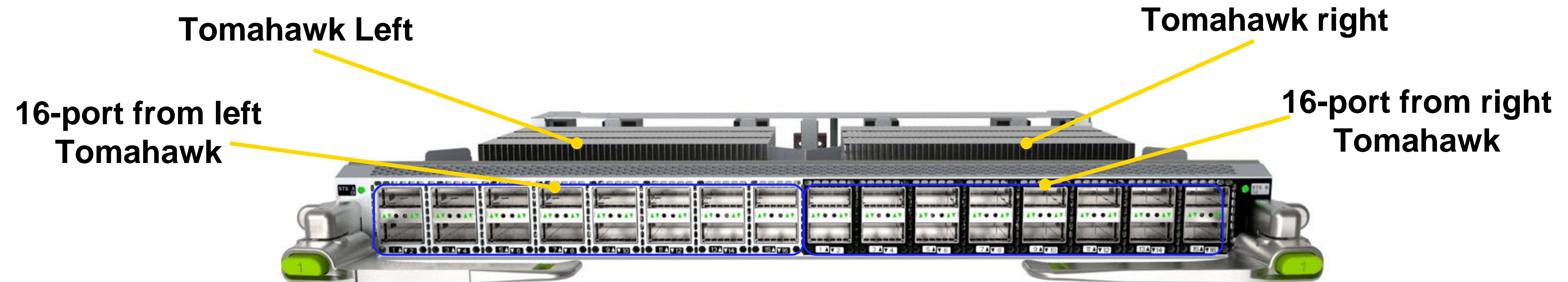
# Line Card (LC)

- Two Tomahawk 3.2T ASIC
- 32 QSFP28 100G Ports
- Four DMO Connector to FAB for data plane signals
- Two DMO Connector to VCP for control plane signals





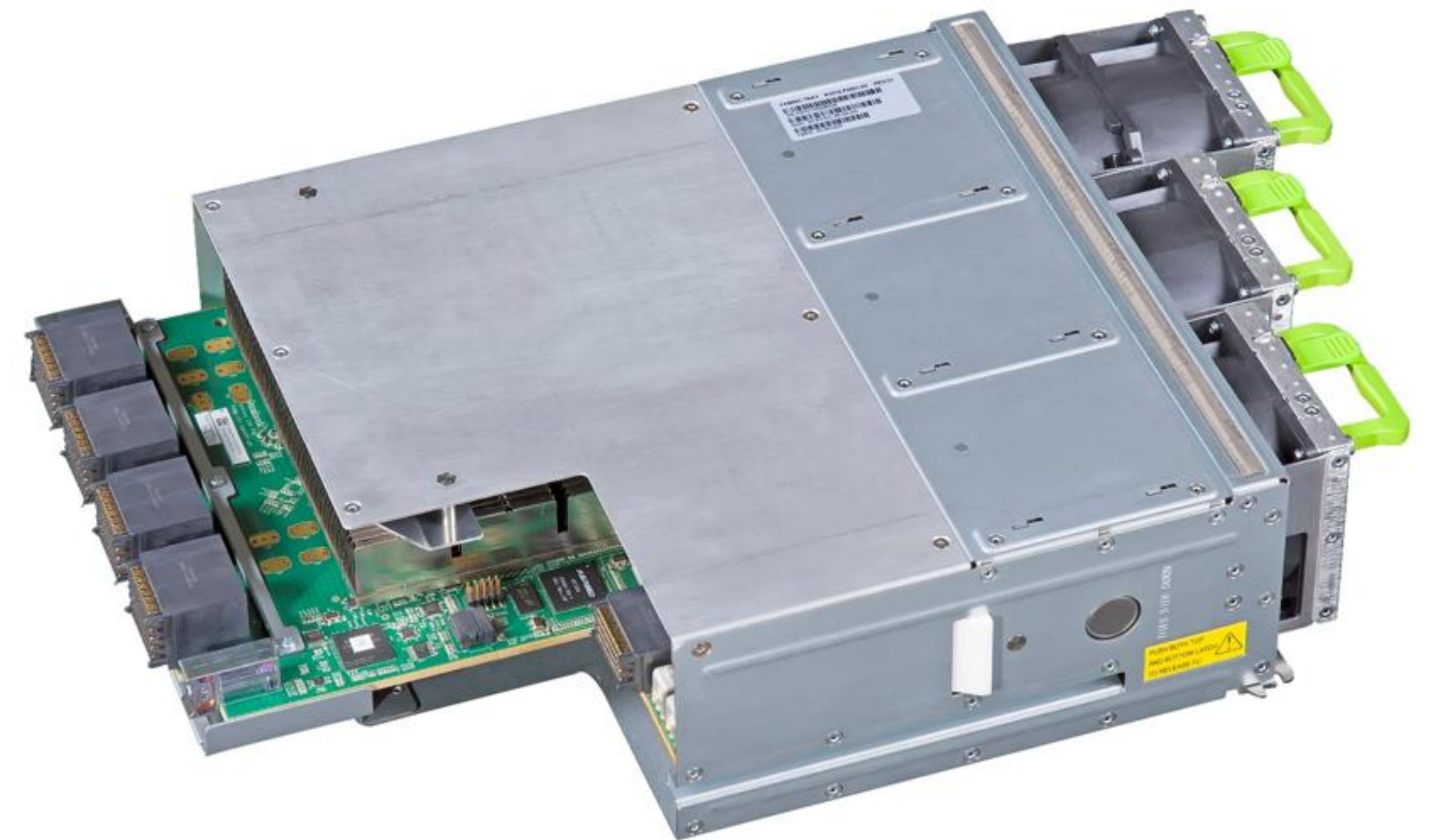
# Interface of Line Card (LC)





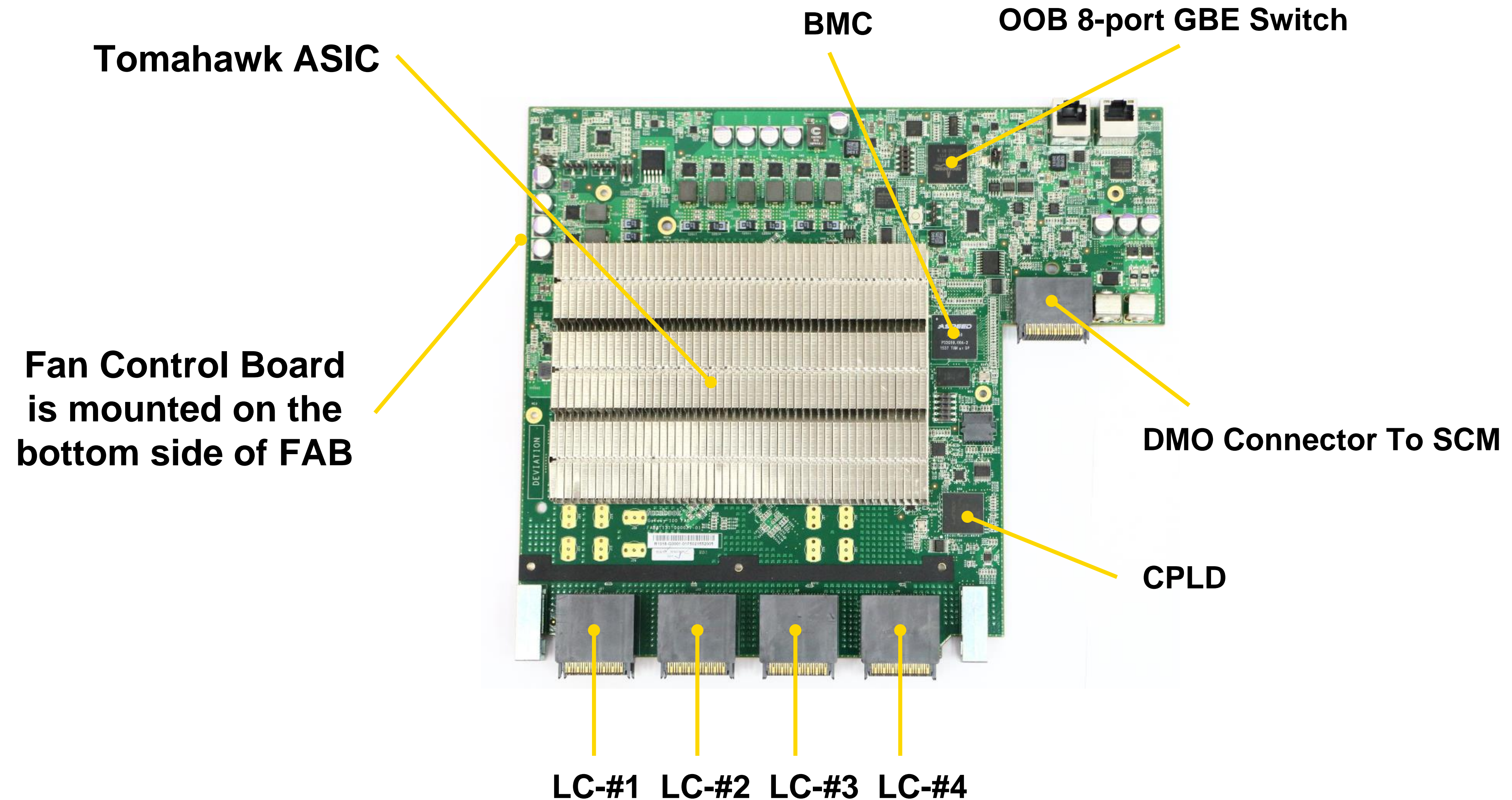
# Fabric Card (FAB)

- One Tomahawk 3.2T ASIC
- Fan Control Board (FCB) is mezzanine card of FAB
- Four 6 x 12 DMO Connector to Four LC for data plane signals
- One 6x8 DMO Connector to HCP for control plane signals





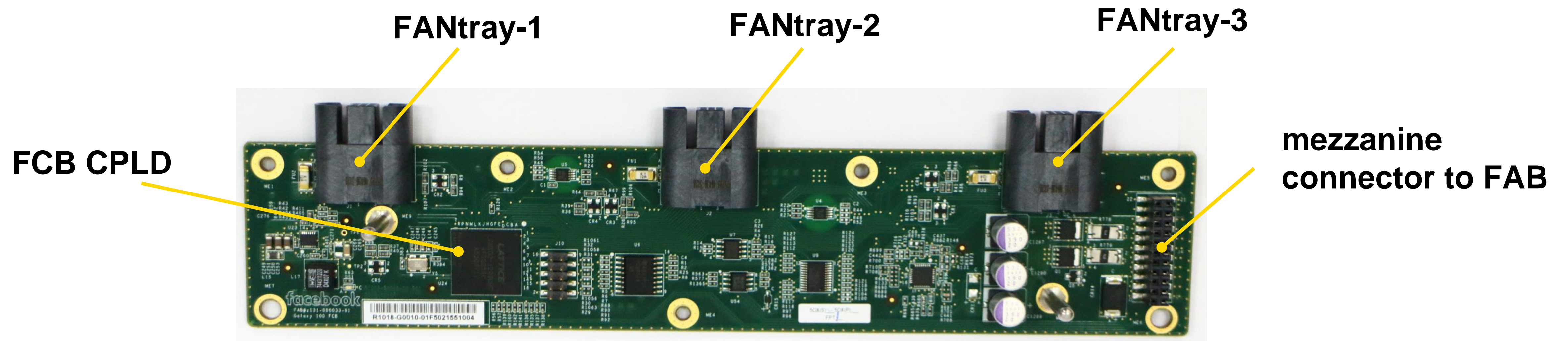
# Components of Fabric Card (FAB)





# Fan Control Board (FCB)

- Mezzanine Card of Fabric Card(FAB)
- Support 3 fan-tray
- FCB CPLD control fan speed, CMM can control FCB CPLD via system I2C bus





# Fabric Card (FAB) Assembly

- FAB Main Board + FCB Mezzanine Board
- Fan-tray plugged into FAB assembly and Mated to FCB





# Chassis Management Module (CMM)

- AS2540 BMC
- OOB 16 port Switch BCM5396
- Console UART MUX for all LC and FAB COM-e CPU and BMC CPU
- Chassis Management I2C bus





# System Control Module (SCM)

- Two COM-e CPU Module Cards
- COM-e CPU Module has one to one mapping to SWE
- 6x12 DMO connector to HCP
- SCM-LC has two COM-e mounted, SCM-FAB has only one COM-e mounted





# System Control Module (SCM)

→ **SCM-LC:**  
SCM for Line Card



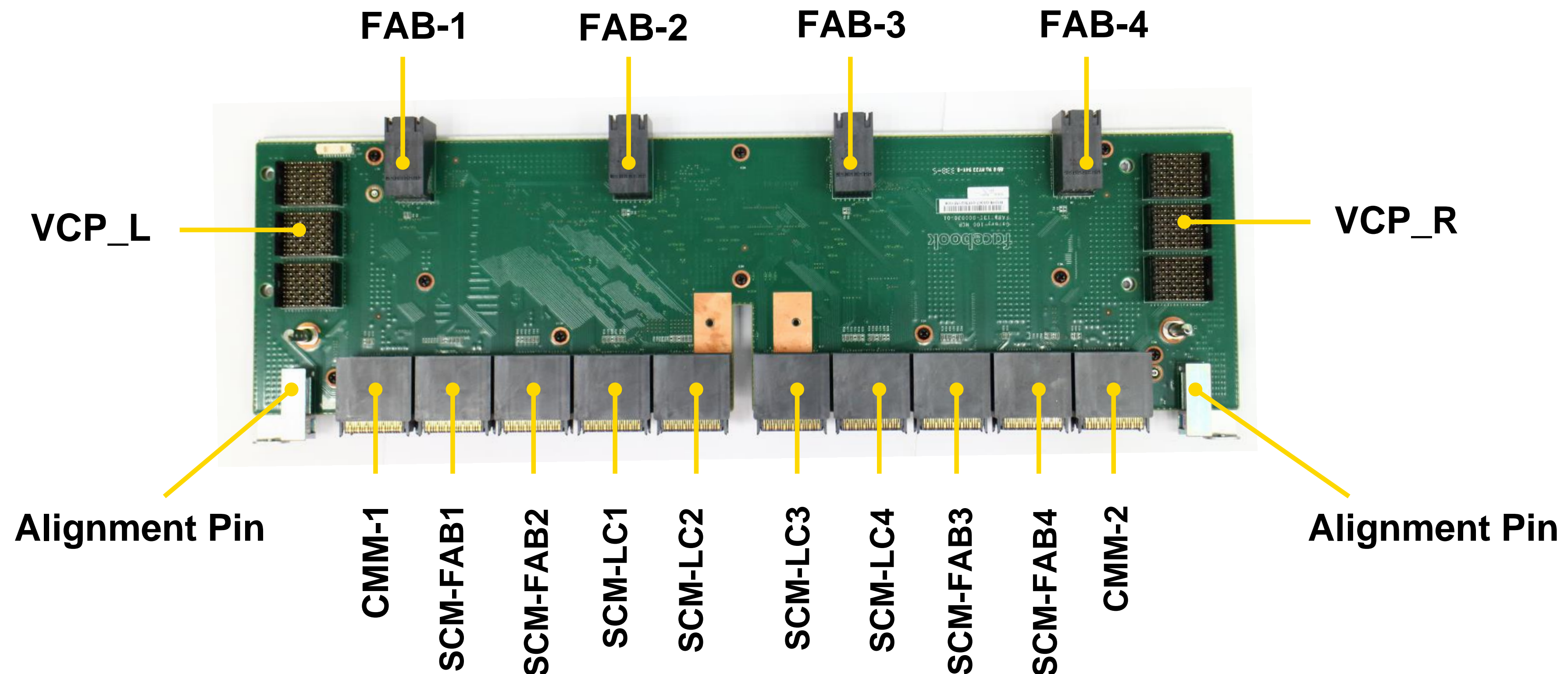
→ **SCM-FAB:**  
SCM for Fabric Card





# Horizontal Control Plane (HCP)

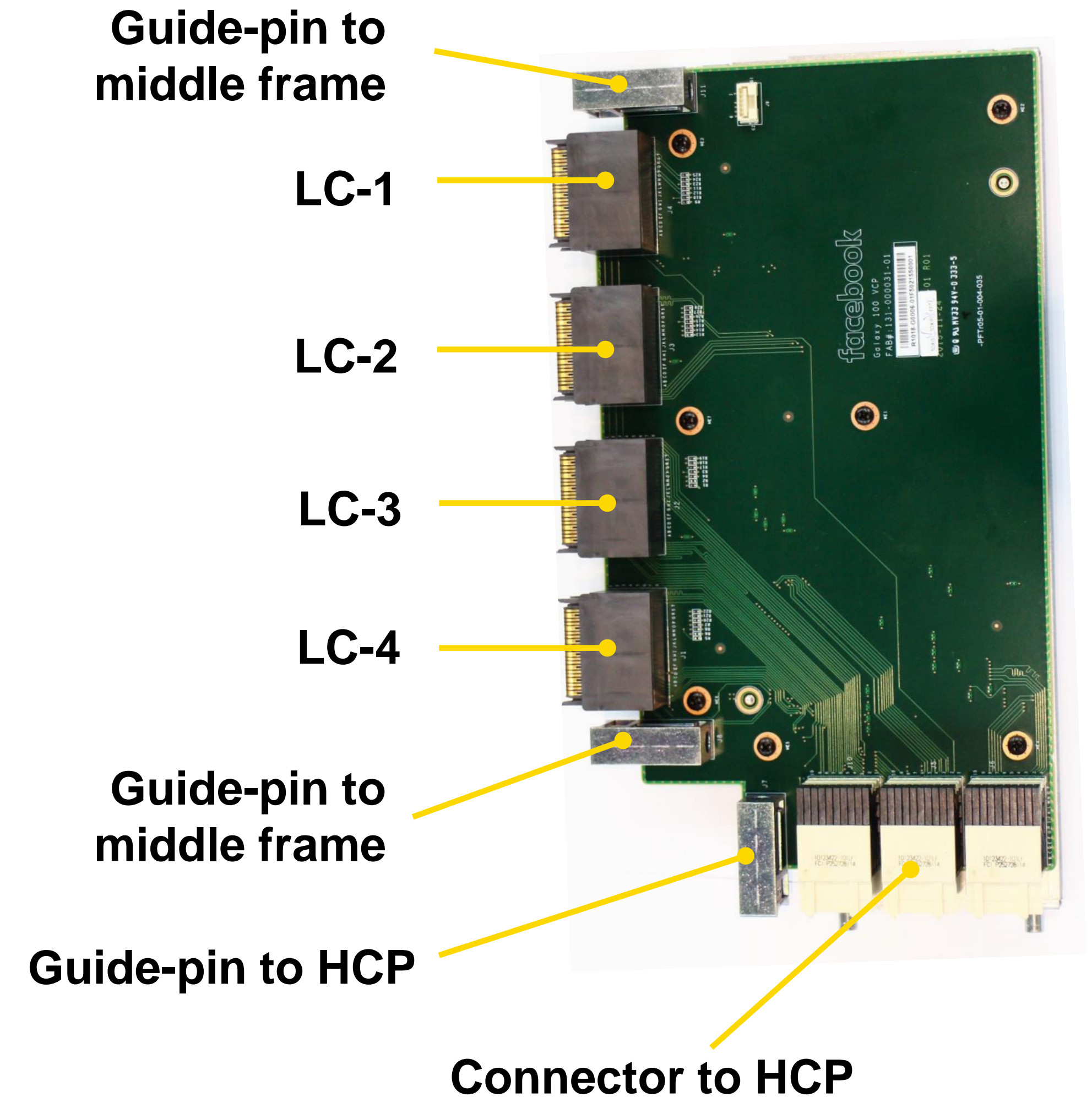
- Horizontal Control Plane for Galaxy Chassis
- Connect to two CMM, eight SCM and VCP-L, VCP-R





# Vertical Control Plane (VCP)

- Two VCP: VCP-L and VCP-R
- Connect to four LC
- Form the control plane of Backpack system together with HCP





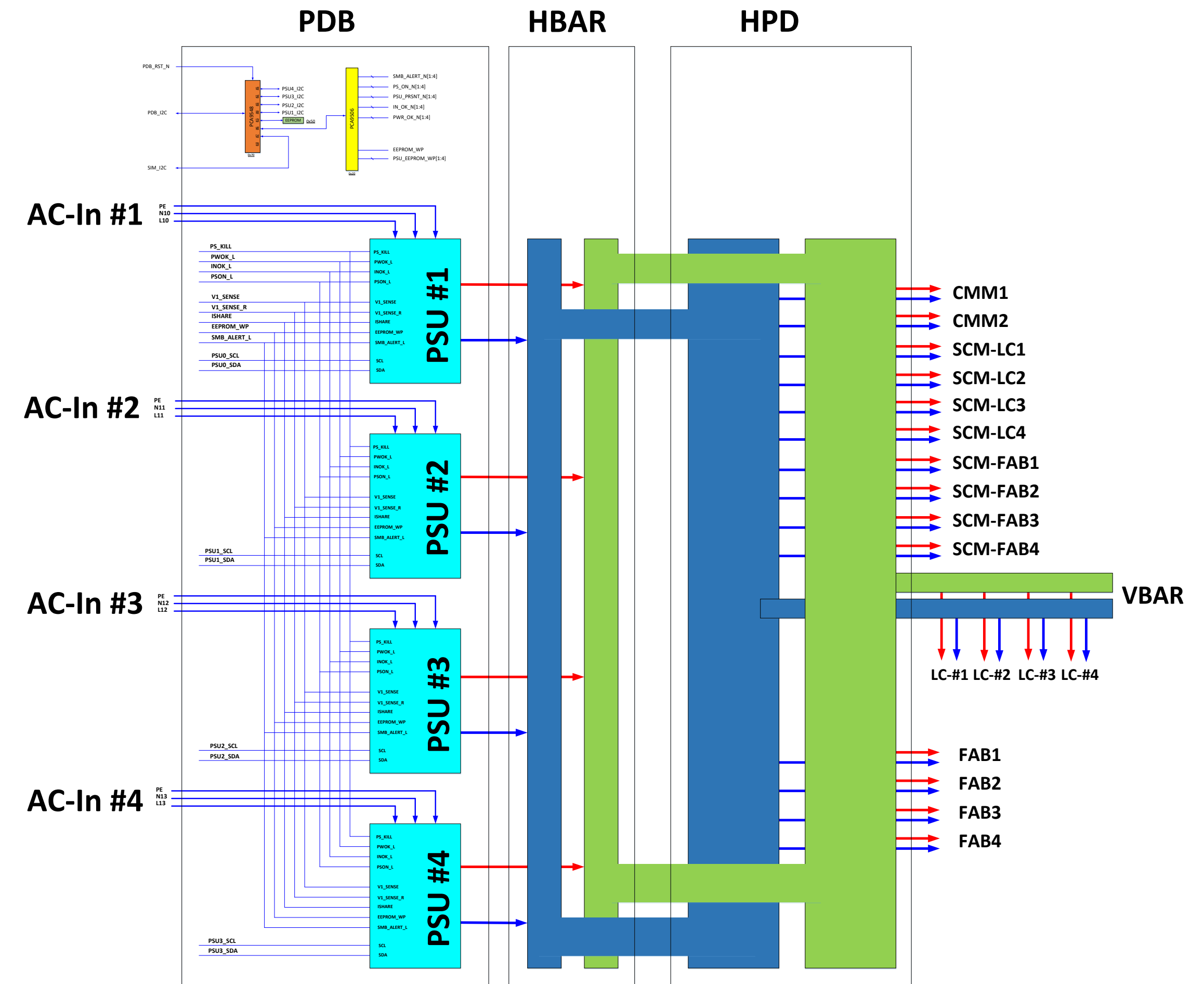
# Power distribution system

- ➔ PSU: PowerOne PFE3000-12
- ➔ BBA (Power Bus Bar Assembly)

- HBAR: Receive power output from PSU
- HPD: Provide power for CMM, SCM, FAB
- VBAR: Provide power for LC

- ➔ Power control and monitor cable

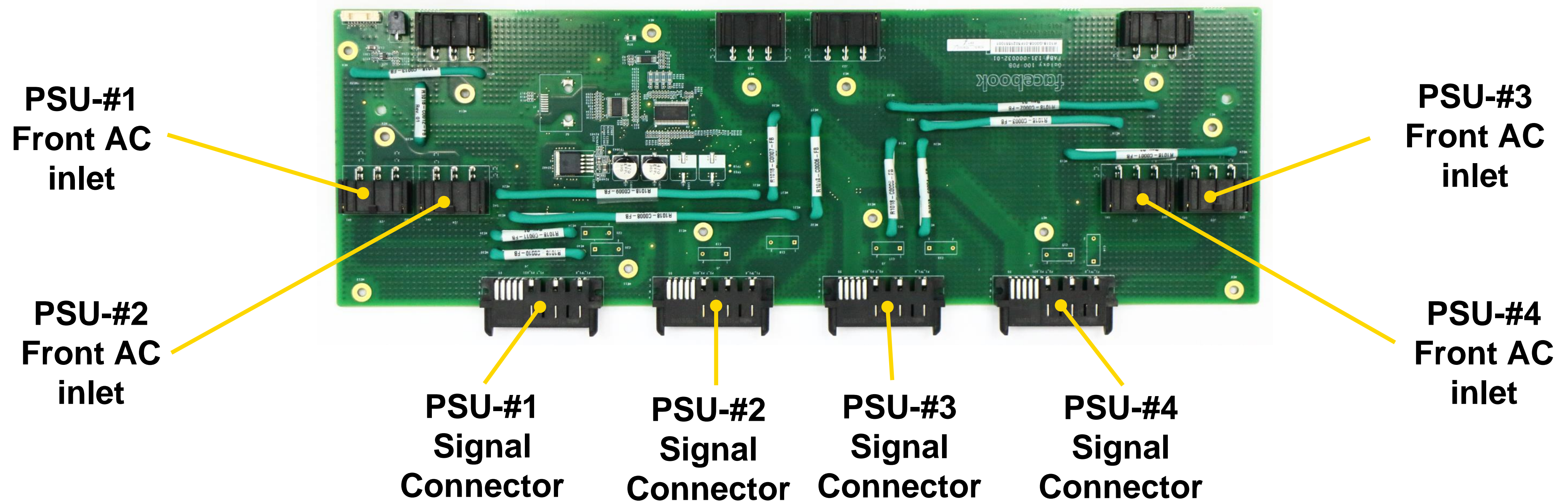
- CMM as controller
- CMM ⇔ HCP ⇔ PDB Jumper cable
- HCP ⇔ SIM for System LED





# Power Distribution Board (PDB)

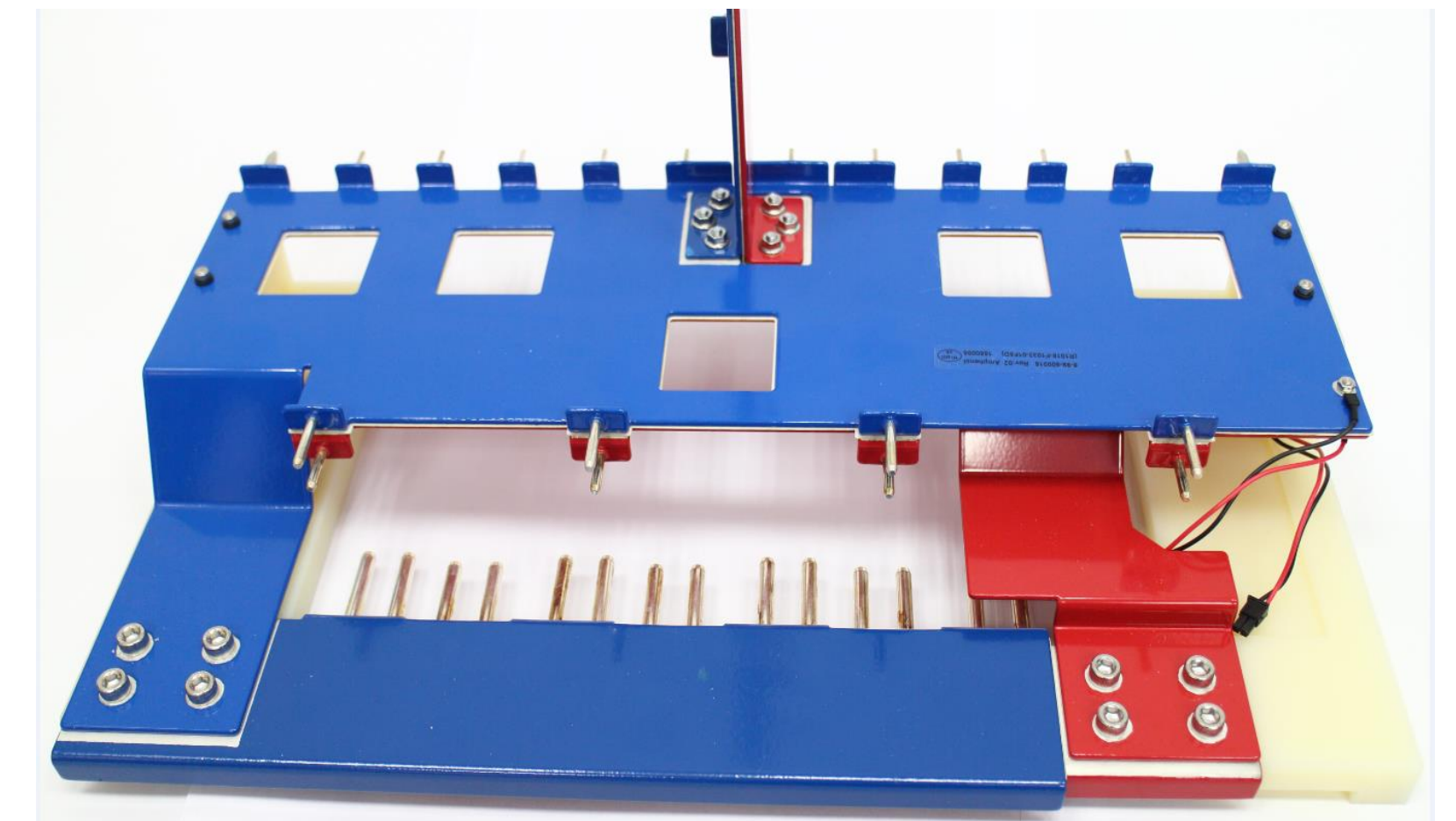
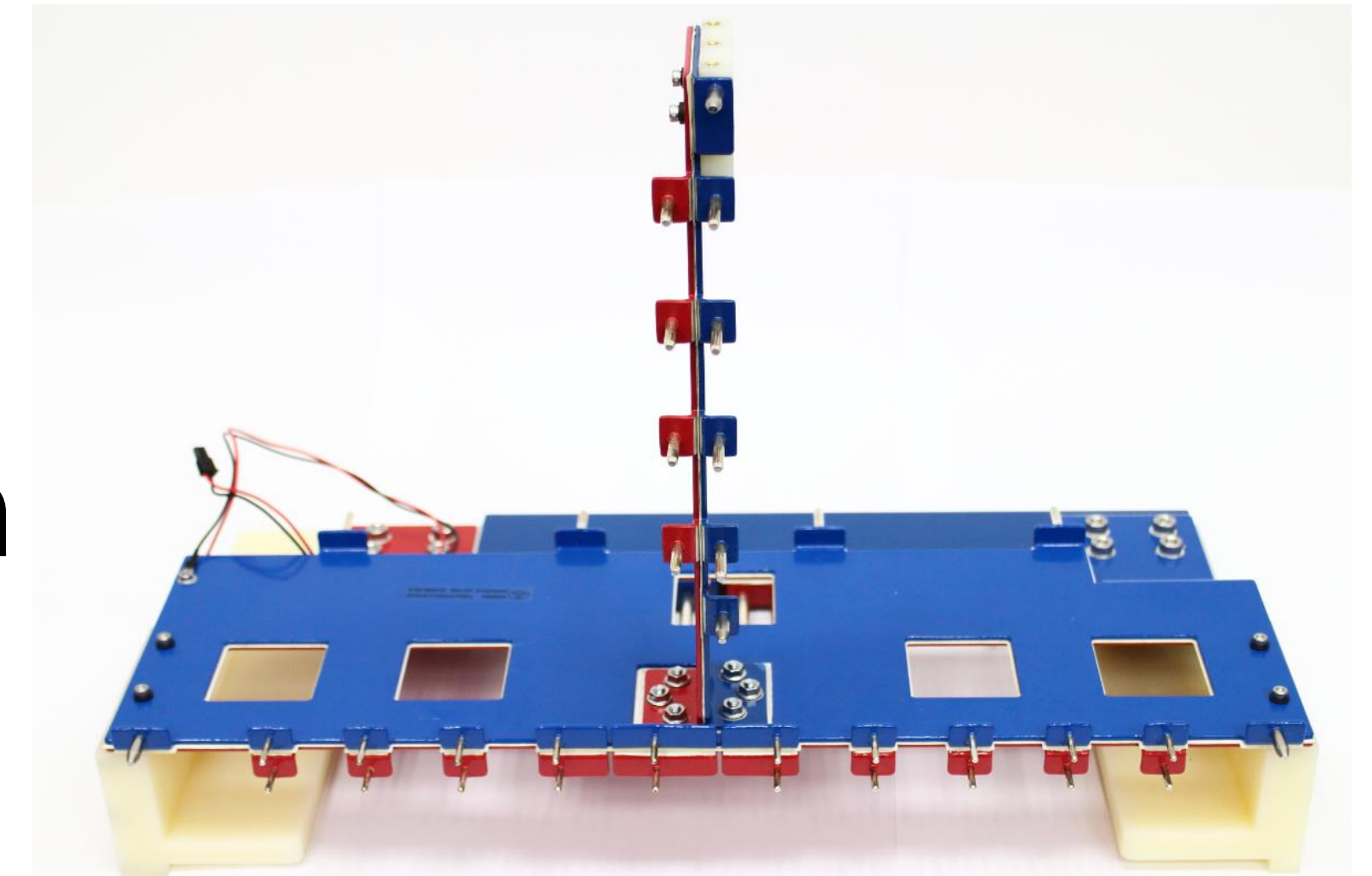
- Four AC inlet: front access
- CMM can access four PSU through PDB
- Output power is load sharing on HBAR





# Bus BAR Assembly

- BBA(Bus Bar Assembly) consists of
  - HBAR: Horizontal Bar
  - HPD: Horizontal Power Distribution
  - VBAR: Vertical Bar
- HBAR Connect to PSU  
HPD for CMM, SCM and FAB  
VBAR for LC
- Sense wire to PDB





# Typical Power Consumption

PWM	20%	30%	40%	50%	60%	70%	80%	90%	100%
Normal condition	2670	2697	2726	2800	2896	3124	3290	3550	3729
One fan failure condition	2550	2627	2703	2768	2822	3056	3223	3428	3645

➔ Unit is Watt

➔ Measured with 3.5W eload



# Thermal design

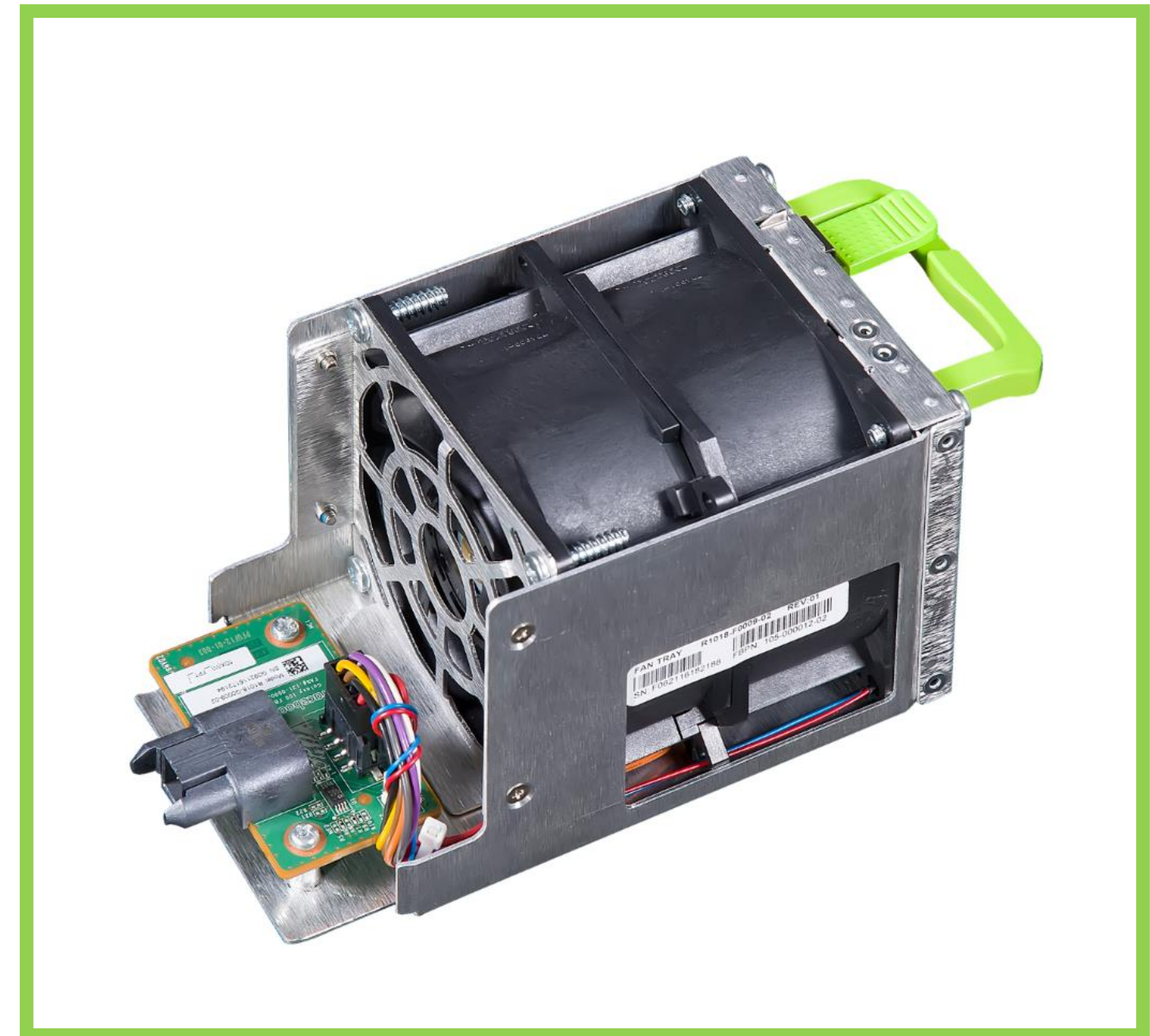
- Support 55C optic at ambient 35C environment
- 12 fan-tray on the rear panel
- Thermal Improved LC front panel design for better QSFP28 cooling
- Multiple on-board temperature sensor to monitor thermal healthy status of the system





# Fan-tray

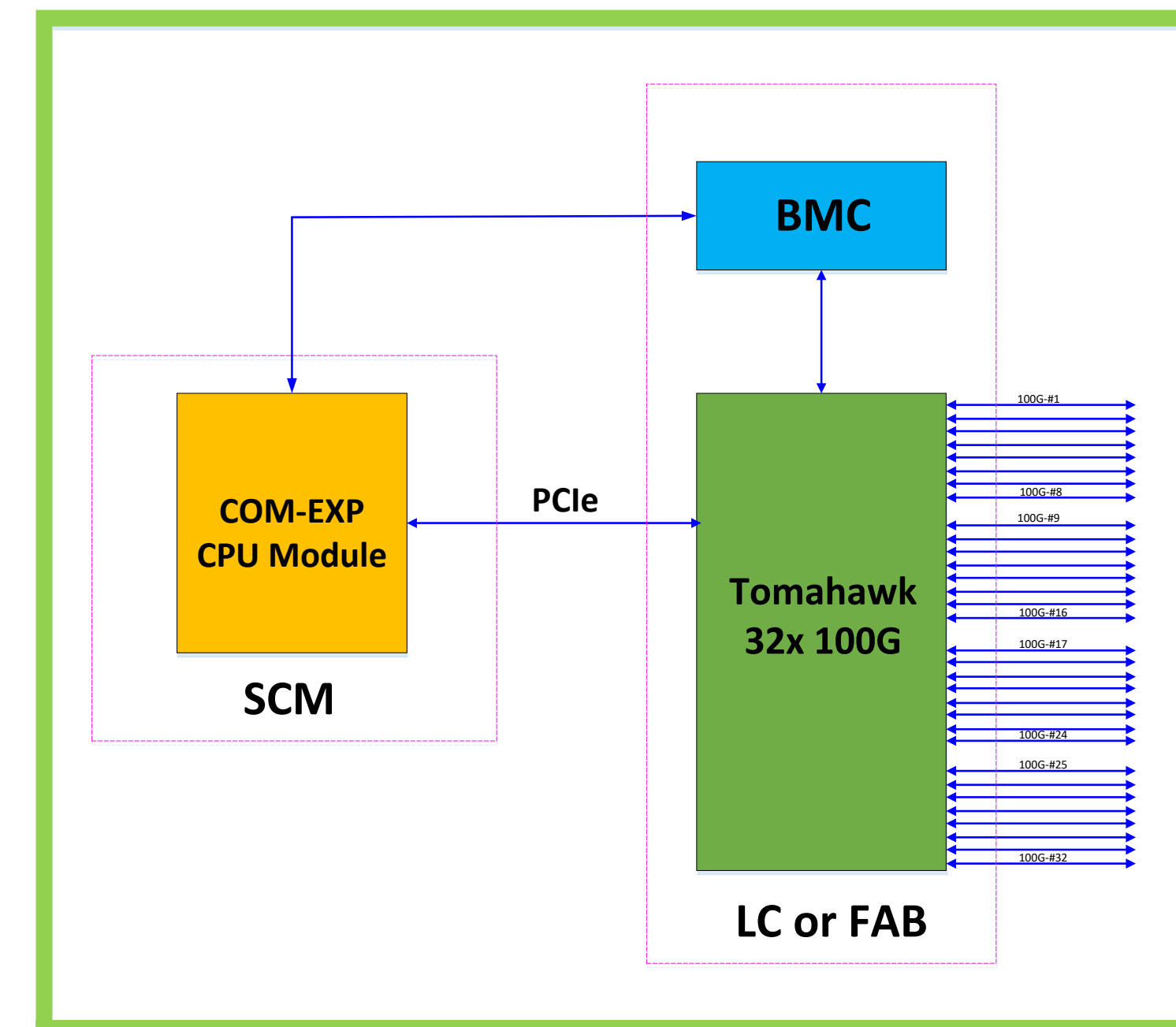
- Screw-less latch design for easy maintenance
- Powerful 80 x 80 mm CR fan
- Hot swappable
- LED on rear panel
- Each FAB carries 3 Fan-tray by FCB(Fan Control Board), total 12 fan-tray in galaxy chassis





# Switch Element (SWE)

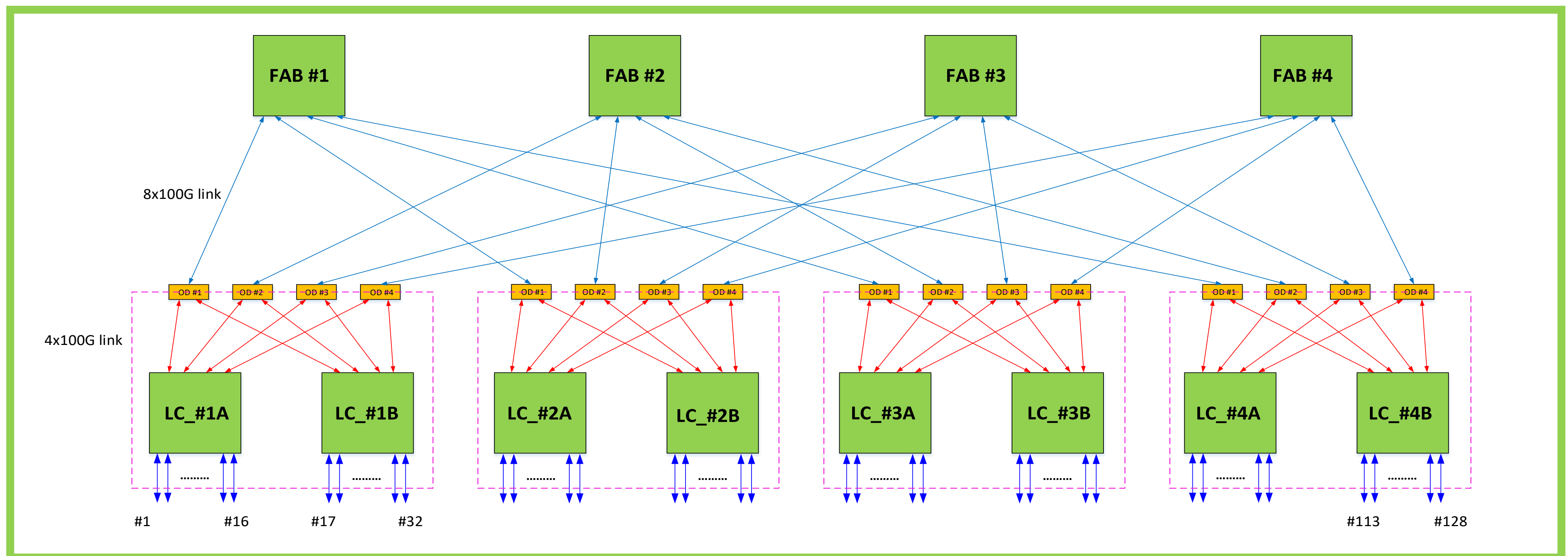
- Disaggregated architecture
- Switch element consists of three components
  - Tomahawk switch ASIC
  - COM-e CPU module
  - BMC
- 12 switch elements in Backpack
  - Each LC has 2 SWE
  - Each FAB has 1 SWE





# Fabric Topology: CLOS

- Two Stage Spine-leaf architecture
- Fully non-blocking
- 12 Switch Elements





# Optic transceiver

## → Backpack support QSFP28 100G optic

- CWDM4
- Can support other MSA, such as SR4, LR4, CLR4, etc.

## → Backpack support QSFP+ 40G optic

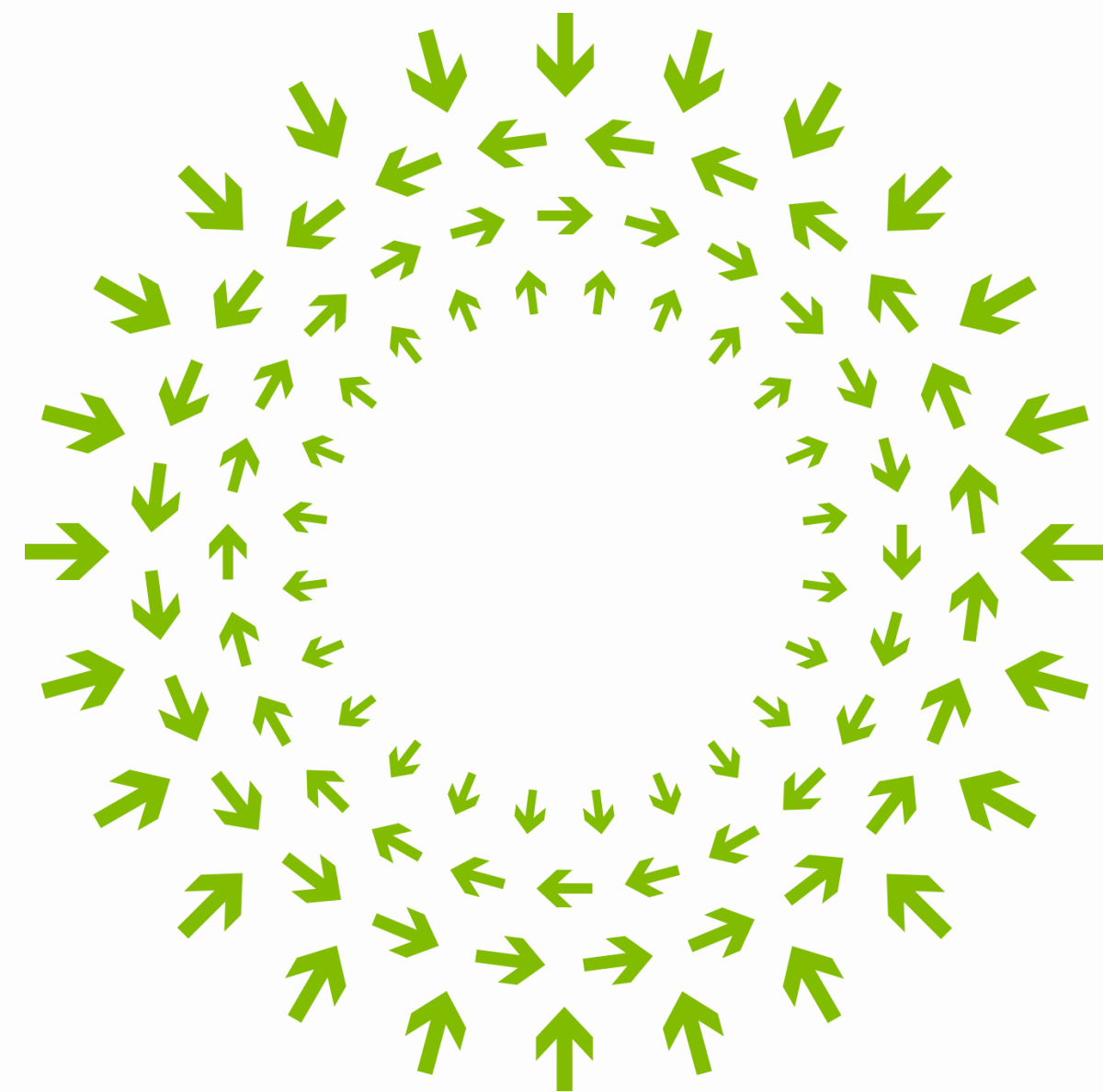
- QSFP+ 40G SR4 optic (multi-mode fiber OM3/OM4)
- QSFP+ 40G LR4 optic (single mode fiber)



# Q&A







# OPEN

Compute Project

