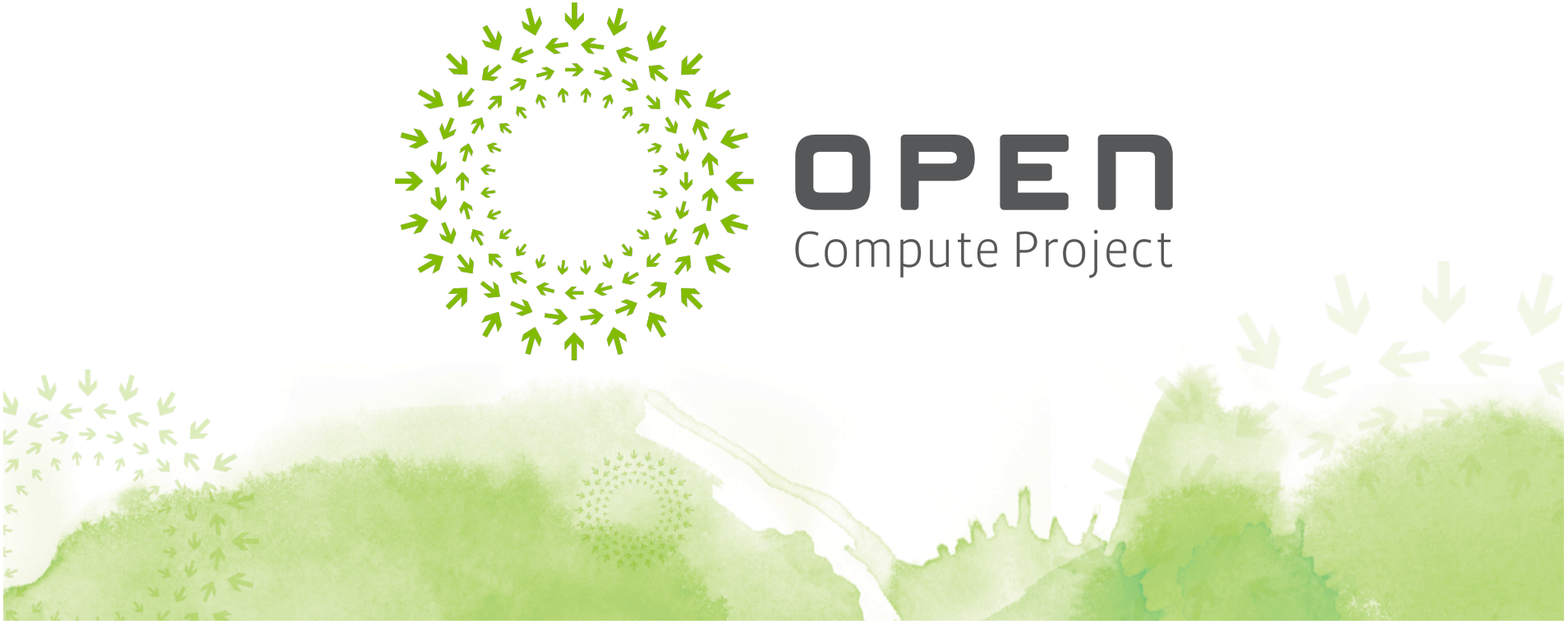
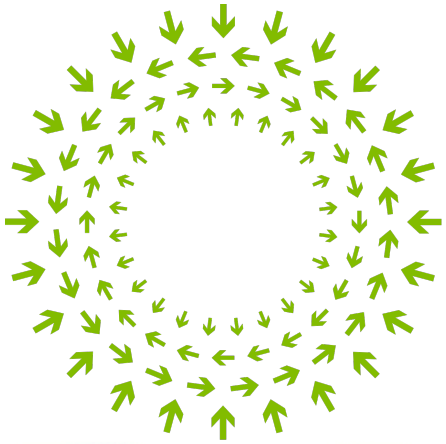

OPEN
Compute Project



OPEN

Compute Project

Platinum Sponsor

OCP U.S. SUMMIT 2016

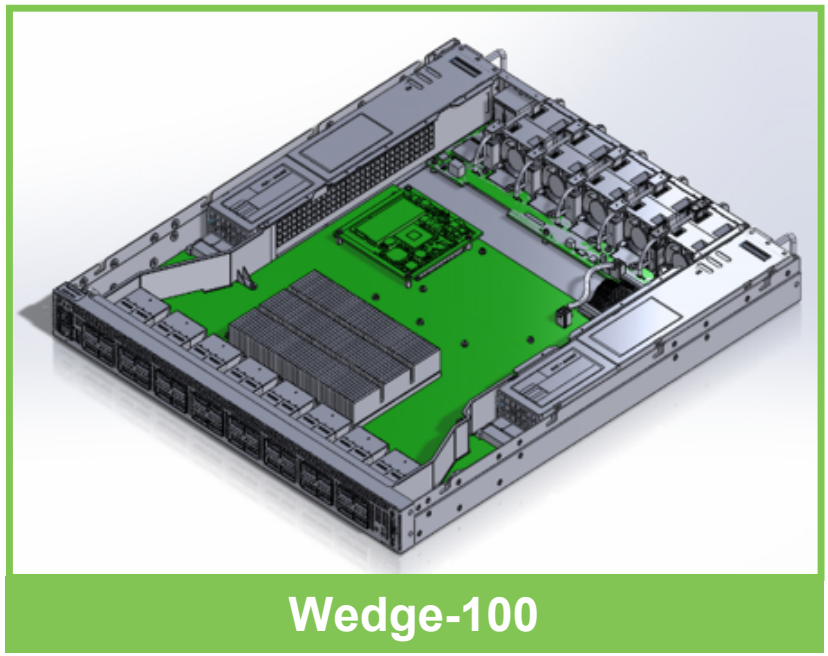
Deep dive of FB network hardware

6-pack:128x40G Modular Switch
Wedge-100: 32x100G TOR Switch

Zhiping Yao
NETWORK HARDWARE ENGINEER OF FACEBOOK

Facebook network hardware

6-pack



Wedge-100

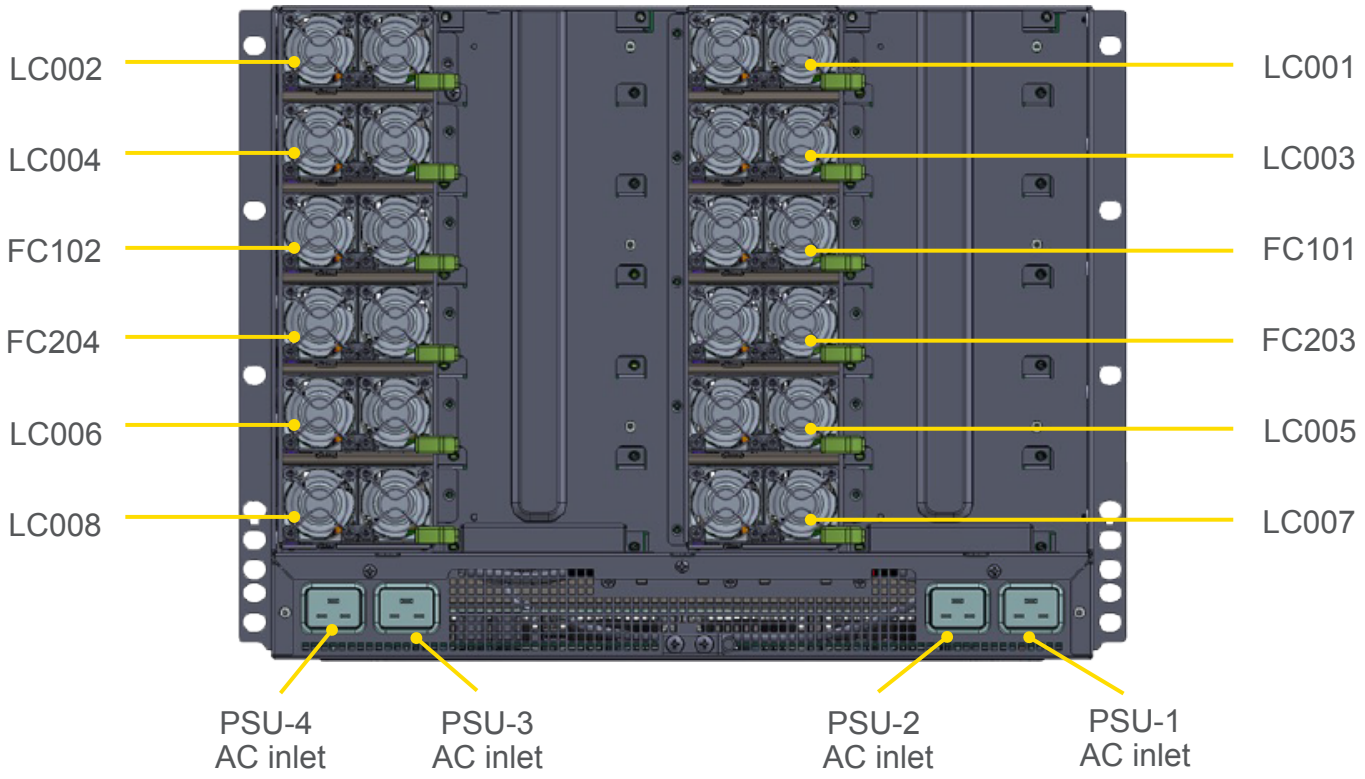
6-pack system introduction

6-pack: an open hardware modular switch

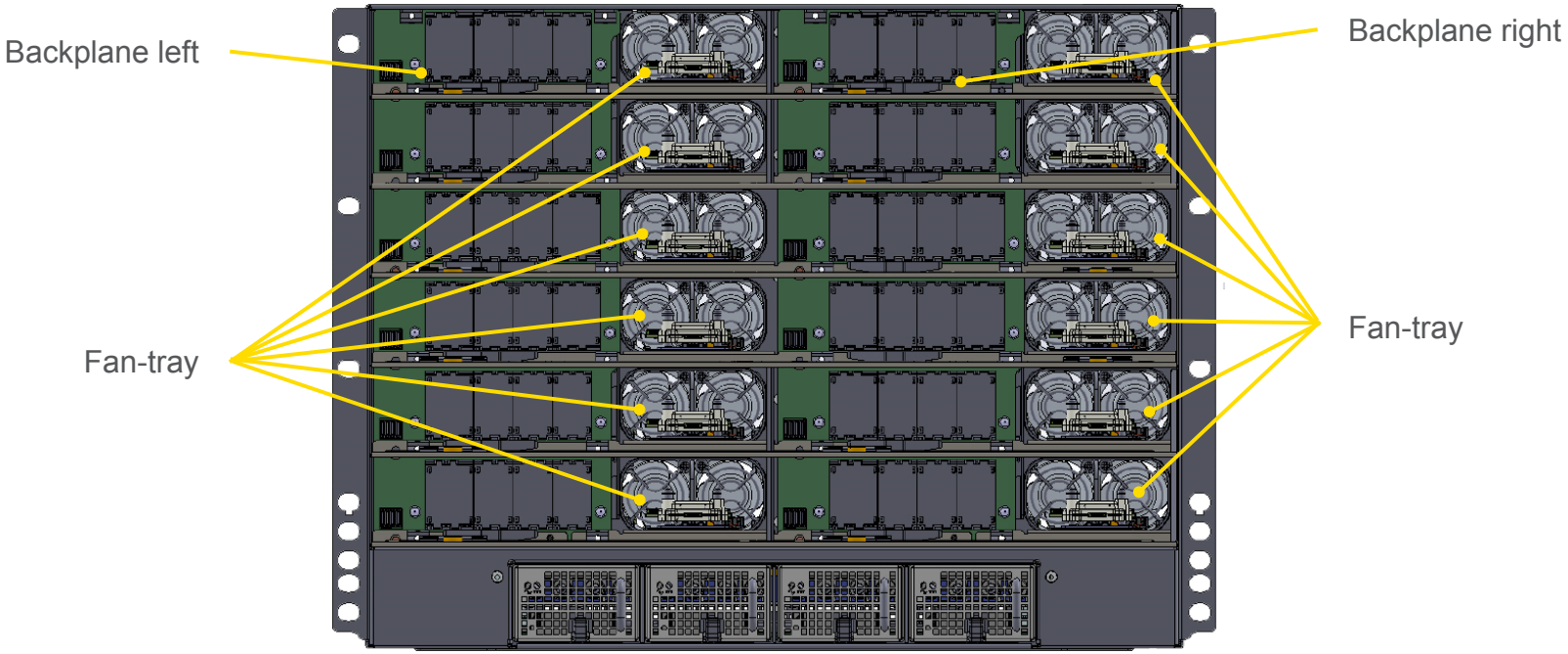
- Hardware architecture: ethernet only, fully open
- Network topology: dual stage spine-leaf
- Switch Software: FBOSS
- Manageability: operated like server from BMC
- Density: 128xQSFP+ 40G support
- Building block: SWE (switch element)
  - Trident-2 switch ASIC
  - Panther+ MicroServer
  - BMC

Front view of 6-pack

Rear view of 6-pack

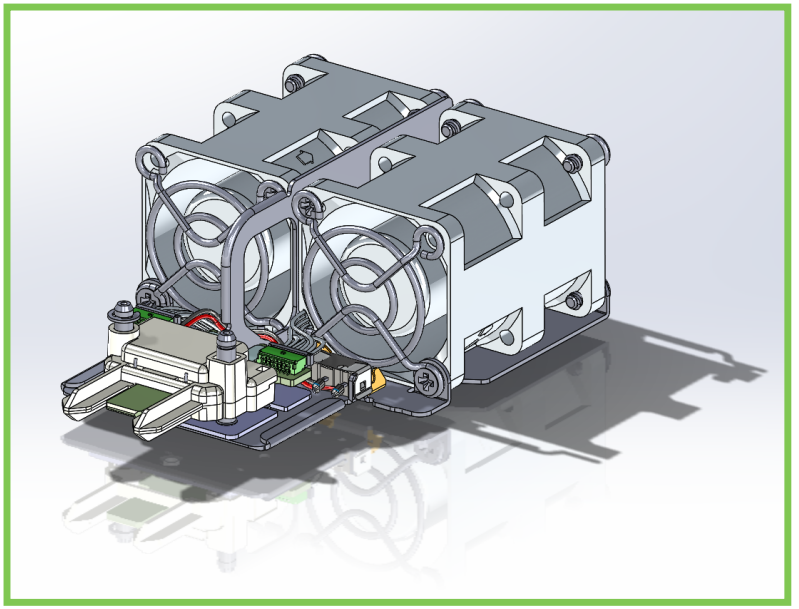


Front view without LC and FAB



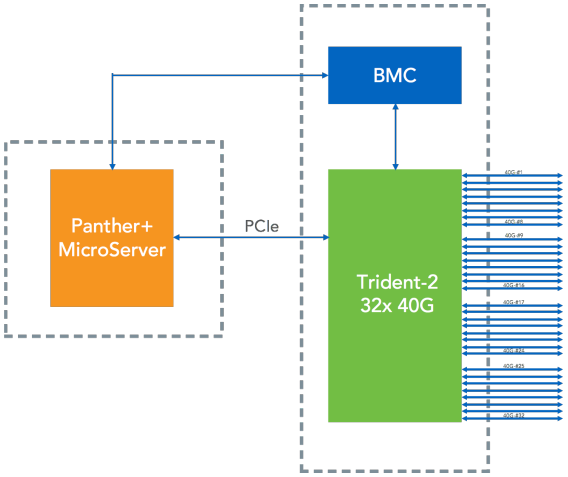
6-pack fan-tray

- 6-pack fan-tray
  - Two 40x40x56mm CR fan
  - Hot-swap
- One fan-tray per switch element
  - One fan-tray per line card
  - Two fan-tray per fabric card

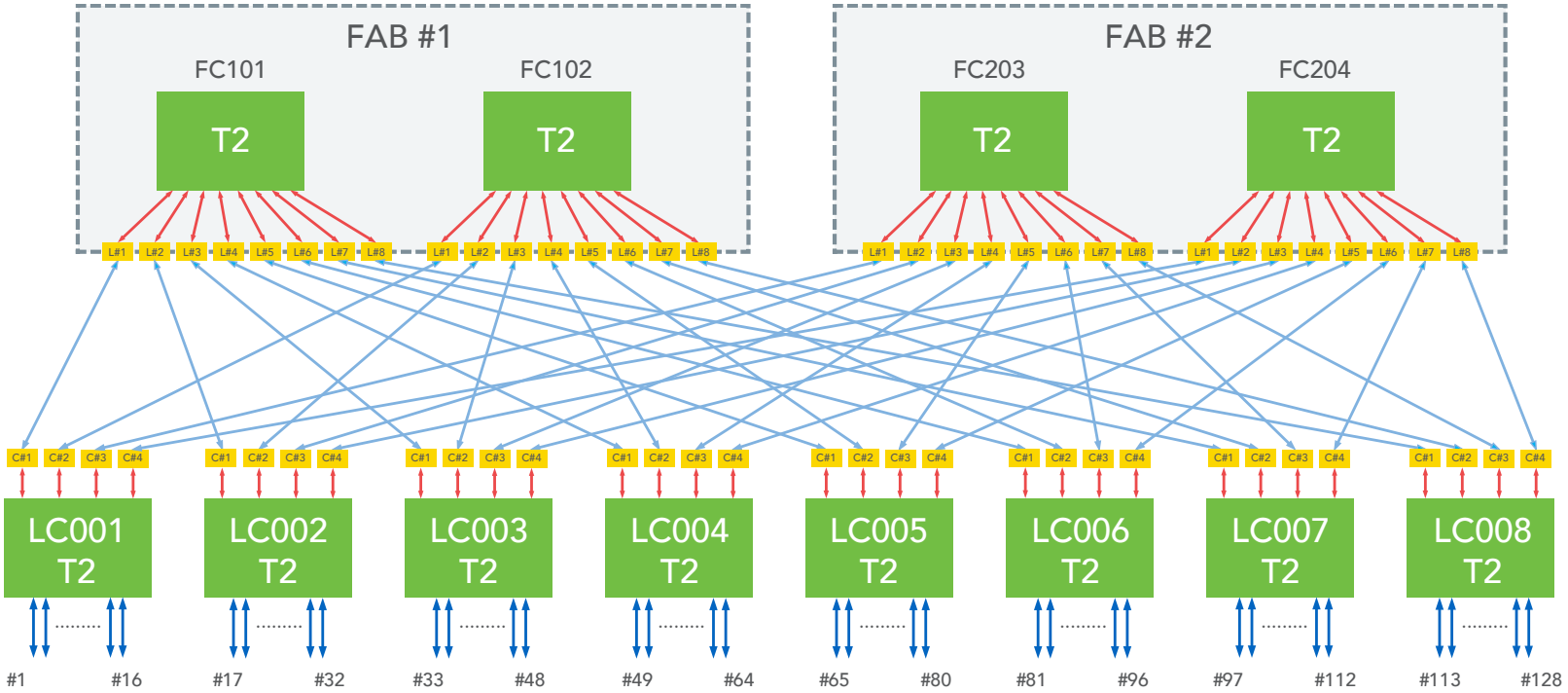


Switch element (SWE)

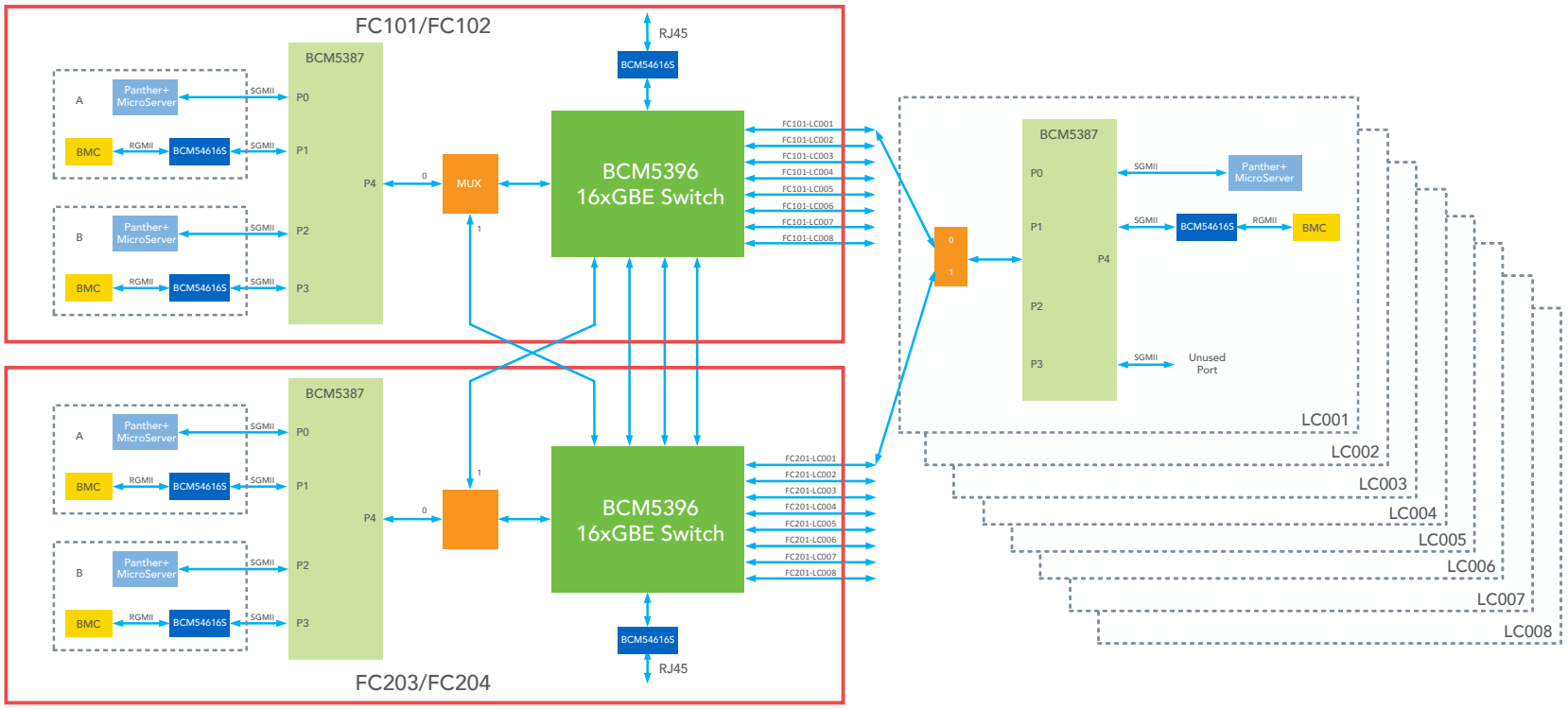
- Disaggregated architecture
- Switch element consists of three components
  - Trident-2 switch ASIC
  - Panther+ CPU module
  - BMC
- 12 switch elements in 6-pack
  - Each LC has 1 SWE
  - Each FAB has 2 SWE



6-pack fabric architecture

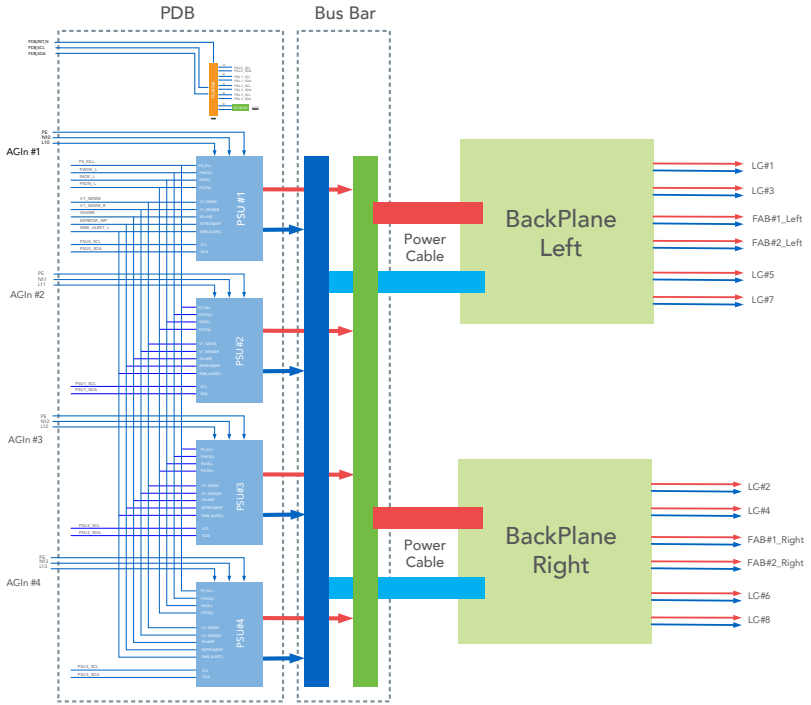


6-pack out-of-band ethernet architecture



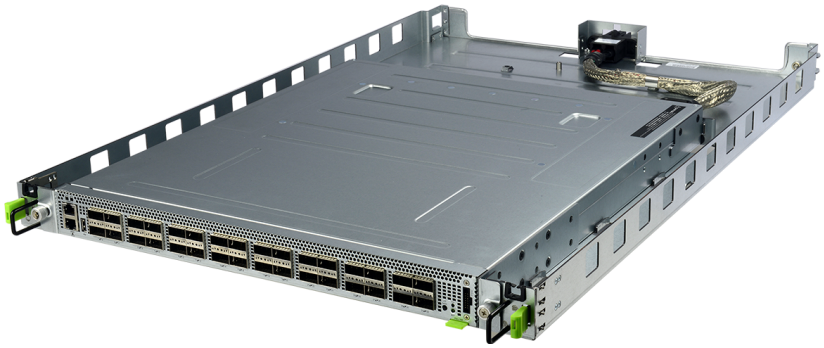
6-pack power distribution system

- PSU: PowerOne PFE3000-12
- Power bus bar assembly
  - Bus bar to combine output from four PSU
  - Power cable to backplane
- Power control and monitor cable
  - FAB ⇔ BP ⇔ PDB



Wedge-100 TOR Switch introduction

19-in SKU



21-in SKU

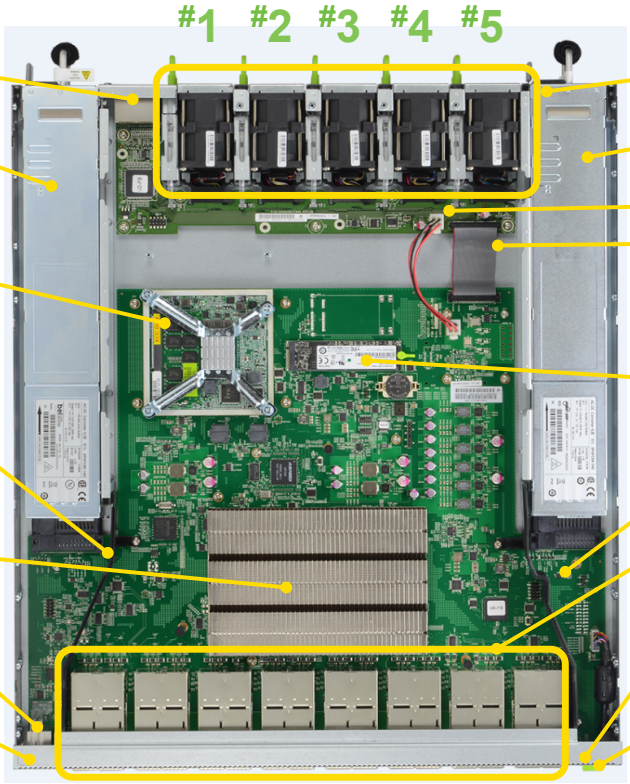
Wedge-100: Open 32x100G TOR Switch

- Facebook's second generation Open TOR Switch based on TOMAHAWK Switch ASIC
- Support Open BMC and FBOSS
- Standard COM-E CPU module as control plane CPU
- Support OpenRack V2 bus bar and rackmon function
- Support 100G QSFP28 DAC cable and 55C CWDM4 optic transceiver

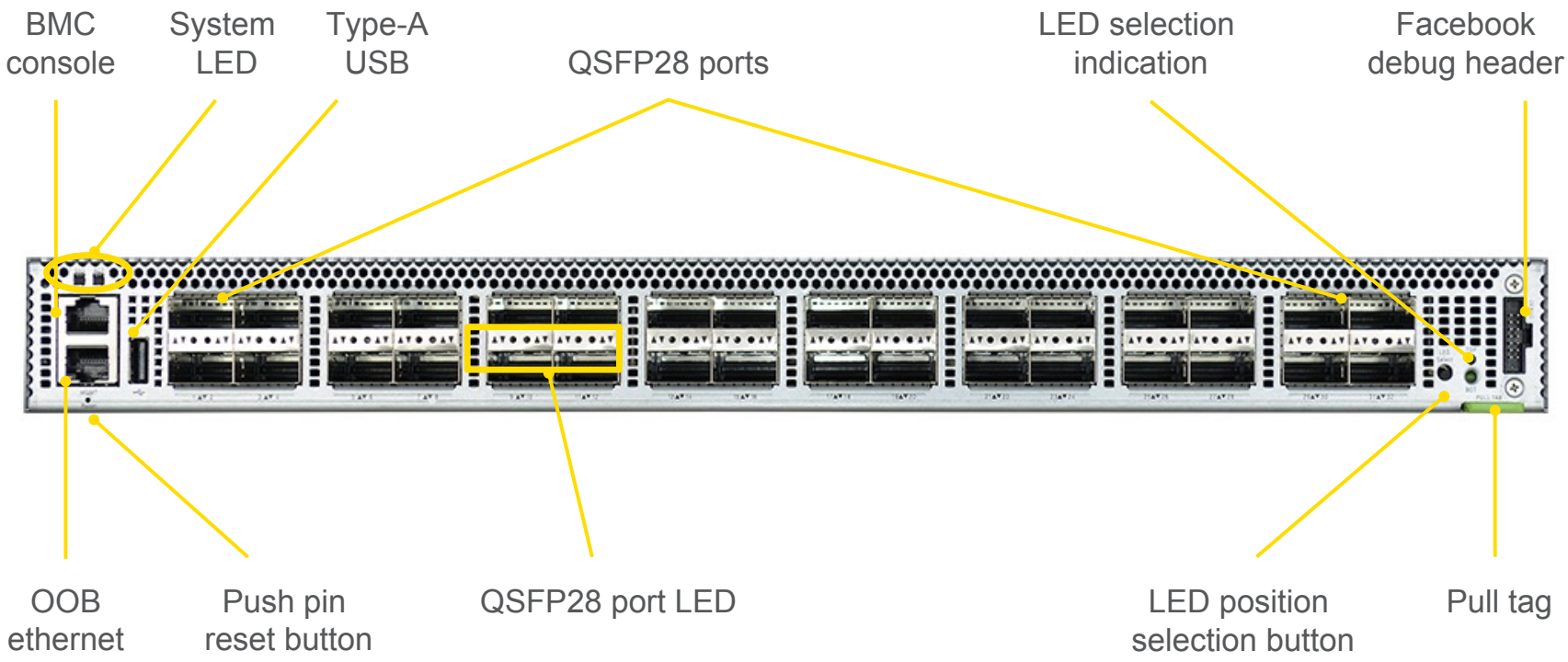
Benefits and applications of wedge-100

- ➔ Enable next generation servers
  - Increased compute density
  - Lower overhead costs
  - Increased network speed with 50G and 25G NIC
- ➔ Support OpenRack V2
- ➔ Improved thermal design for 100G optics
  - 55C CWDM4 optic transceiver

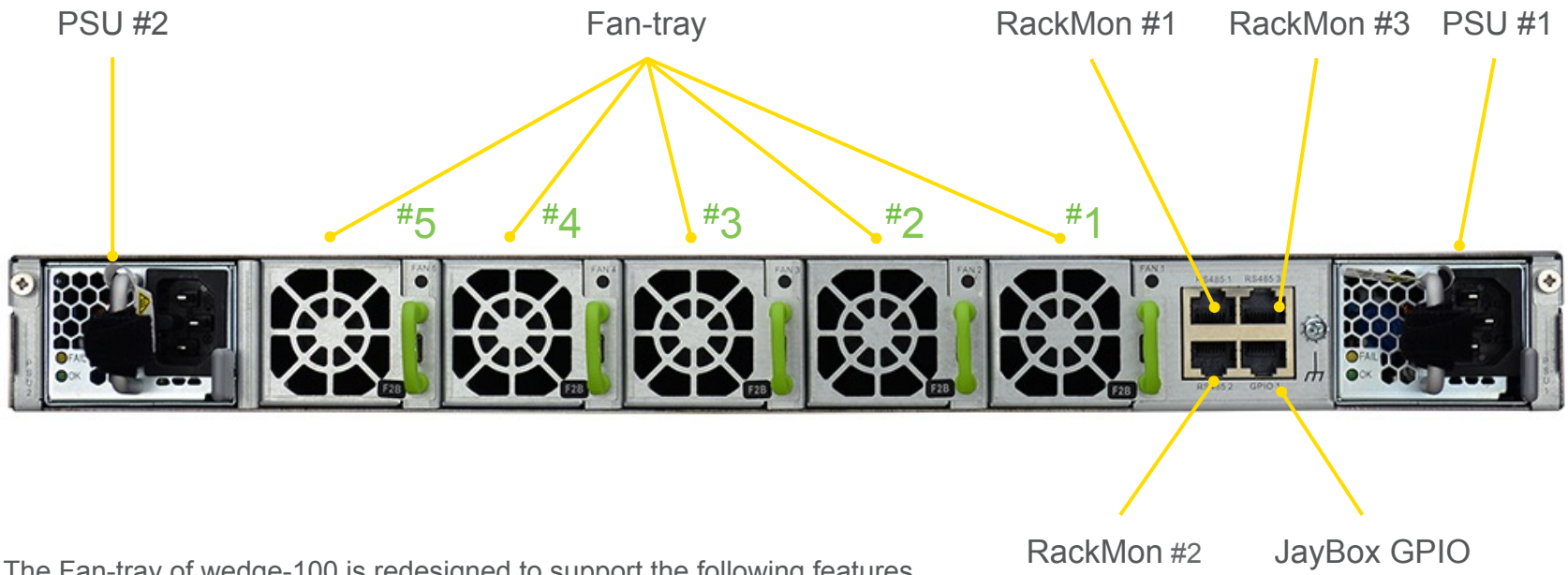
Top view (19-in SKU)



Front panel


BMC console
System LED
Type-A USB
QSFP28 ports
LED selection indication
Facebook debug header
OOB ethernet
Push pin reset button
QSFP28 port LED
LED position selection button
Pull tag

Rear panel



The Fan-tray of wedge-100 is redesigned to support the following features
1: Change to latch design for tooless purpose
2: Fan LED now can be displayed on rear panel to indicate fan-tray failure

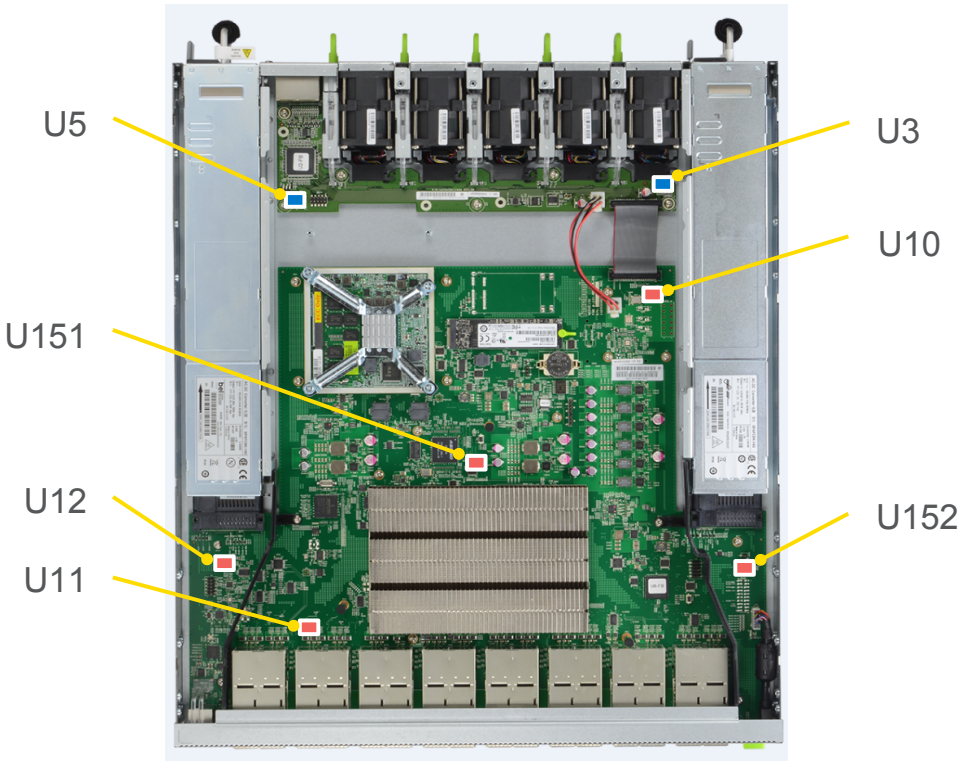
Thermal design

- Support 55C optic at ambient 35C environment
- Five fan-tray on the rear panel
- PSU has separate air channel to avoid recycling
- Multiple on-board temperature sensor to monitor thermal healthy status of the system

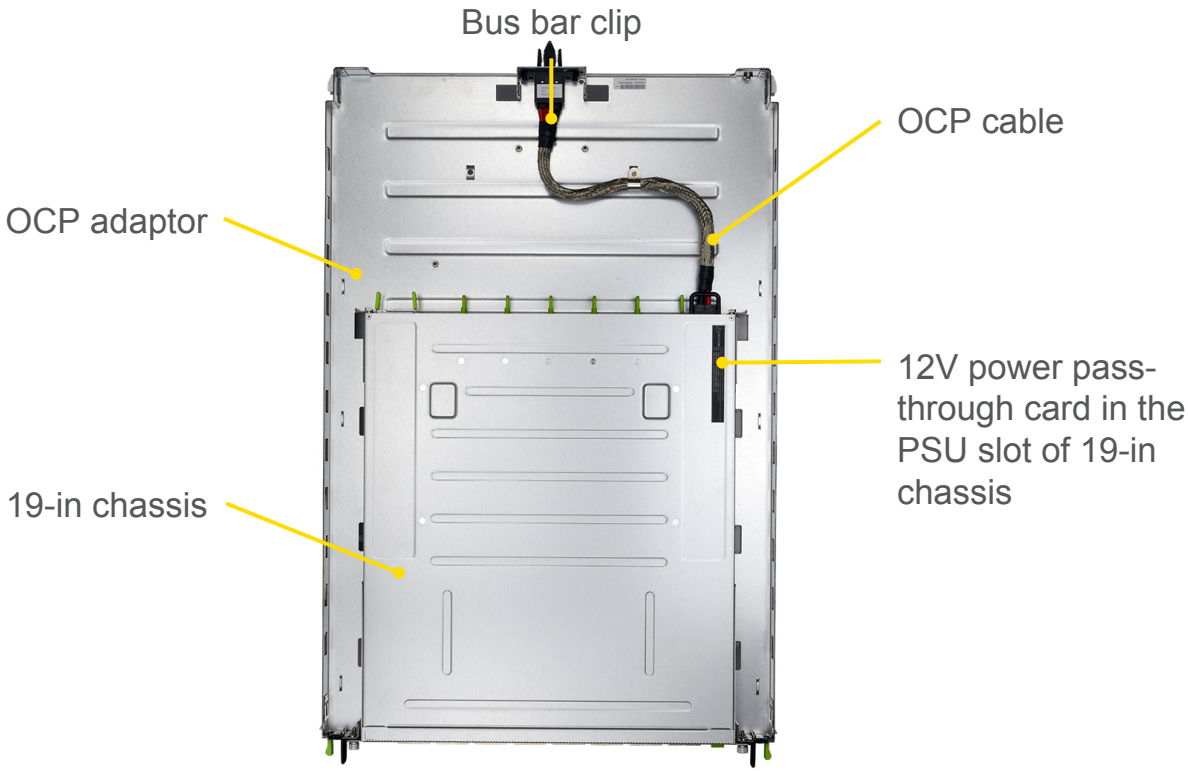
Fan-tray

- Screw-less latch design for easy maintenance
- Powerful CR fan
- LED on rear panel

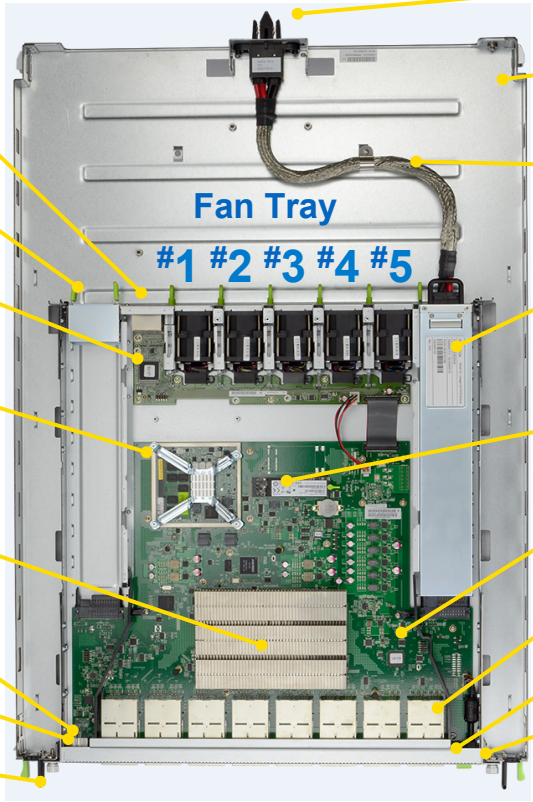
Placement of temperature sensors



OpenRack V2 solution



Top view (21-in SKU)



RackMon 2x2 RJ45 interface

PSU is not used in 21-in SKU,
a dummy filler panel is needed here

Fan_ctrl and RackMon card

COM-E CPU module

Tomahawk ASIC

Type-A USB

1: BMC console
2: OOB ethernet

21-In tray latch

Bus bar clip kit

21-in adaptor box for 19-in chassis

12V DC power cable to bus bar clip

Power bar pass-through card

M.2 128MB SSD

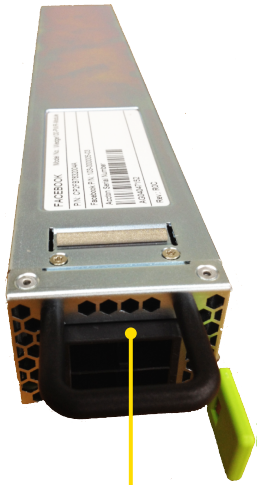
Wedge-100 main board

QSFP28 ports

LED selection button and status LED

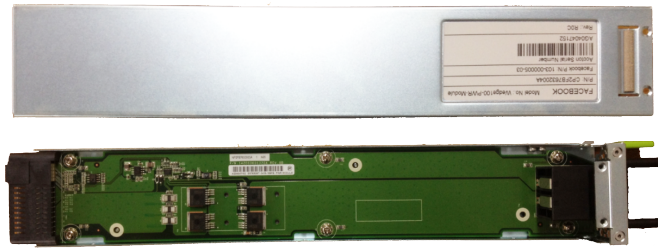
Facebook debug connector

OCP power passthrough card

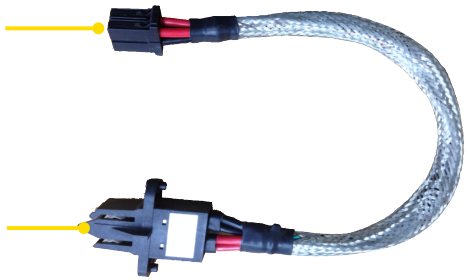


12V bus bar power input connector

Power passthrough card is the same form factor as PSU



OCP power cable



12V power to wedge-100 main board

Port mode

The QSFP28 port of wedge-100 can support the following mode

- **1x100G**: use CWDM4 QSFP28 optic connect to 100G aggregation switch
- **2x50G**: use QSFP28-2xQSFP28 DAC cable connect to 50G NIC
- **4x25G**: use QSFP28-4xSFP28 fan-out cable connect to 25G NIC
- **1x40G**: support QSFP+ 40G SR4 optic and QSFP+ 40G LR4 optic
- **4x10G**: use QSFP-4xSFP fan-out cable connect to 10G SFP+ NIC

100G optic transceiver

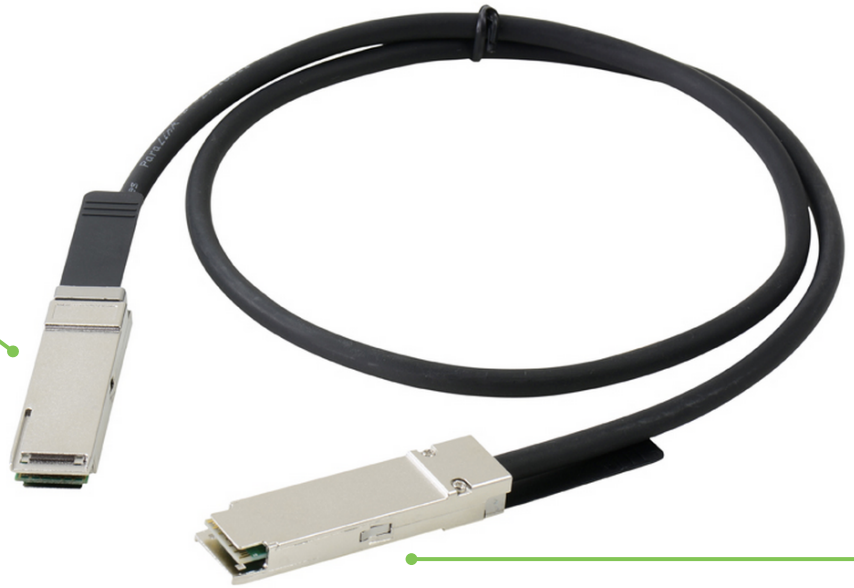
- Wedge-100 support QSFP28 100G optic
  - CWDM4
  - Can support other MSA, such as SR4, LR4, CLR4, etc.
- Wedge-100 support QSFP+ 40G optic
  - QSFP+ 40G SR4 optic (multi-mode fiber OM3/OM4)
  - QSFP+ 40G LR4 optic (single mode fiber)

Wedge-100 DAC cable

- Wedge-100 support the following 3 type of cable
  - QSFP28 100G to QSFP28 100G cable, 1M, 2M, and 3M
  - QSFP28 100G to two QSFP28 50G cable, 1M, 2M, and 3M
  - QSFP28 100G to four SFP28 25G cable, 1M, 2M, and 3M
- Wedge-100 support old wedge-40 DAC if ports are configured at 1x40G or 4x10G mode
  - QSFP+ to 4xSFP+ 40G fanout cable

QSFP28 100G to QSFP28 100G cable

- Four lanes, QSFP28 at both ends
- 25Gbps speed each lane

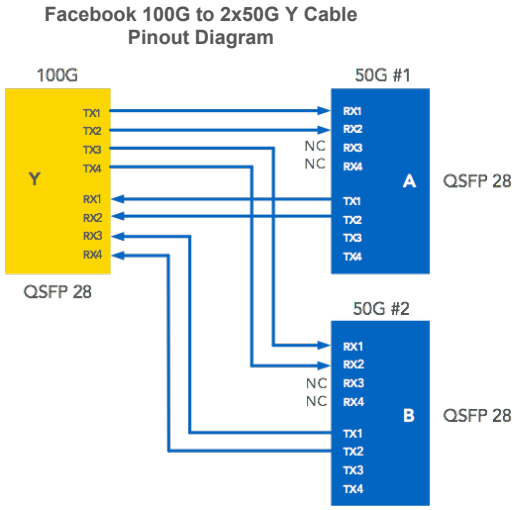


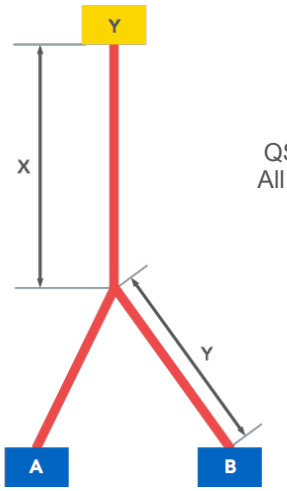
QSFP28 100G
All 4 lanes used

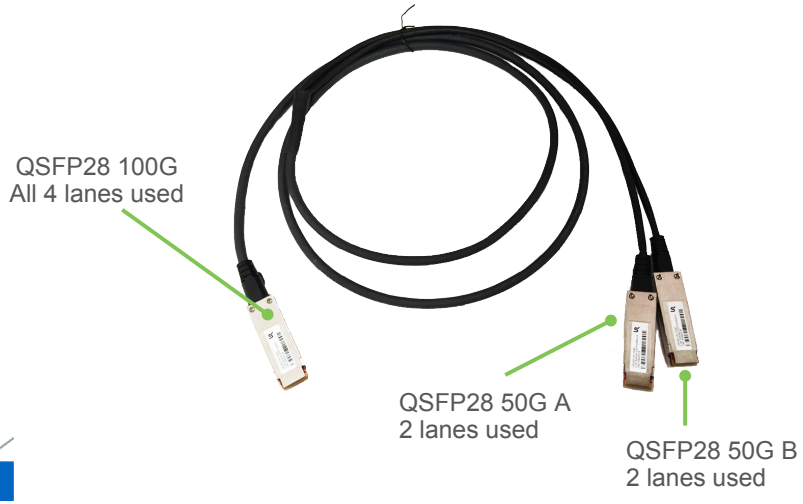
QSFP28 100G
All 4 lanes used

QSFP28 100G to 2x QSFP28 50G cable

- a.k.a Y-Cable, this is a cable for wedge-100 to inter-op with Yosemite 50G NIC CX4
- Yosemite 50G NIC CX4 use QSFP28 form factor, but only 2 lanes used
- Header Y use all 4 lanes, header A, and B use 2 lanes
- All lanes are 25Gbps

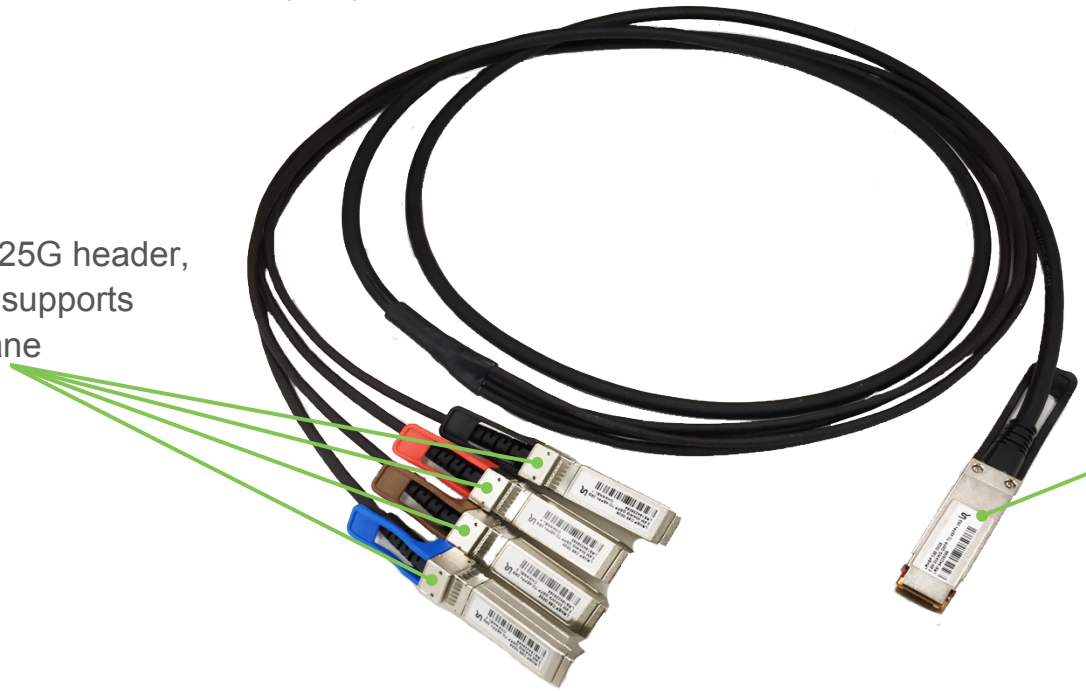






QSFP28 100G to 4x SFP28 25 cable

- Four lanes in QSFP28 side, 25Gbps speed each lane
- One lane at SFP28 side, 25Gbps speed
- 1M, 2M, 3M



Four SFP28 25G header,
each SFP28 supports
25G single lane

QSFP28 100G
All 4 lanes used

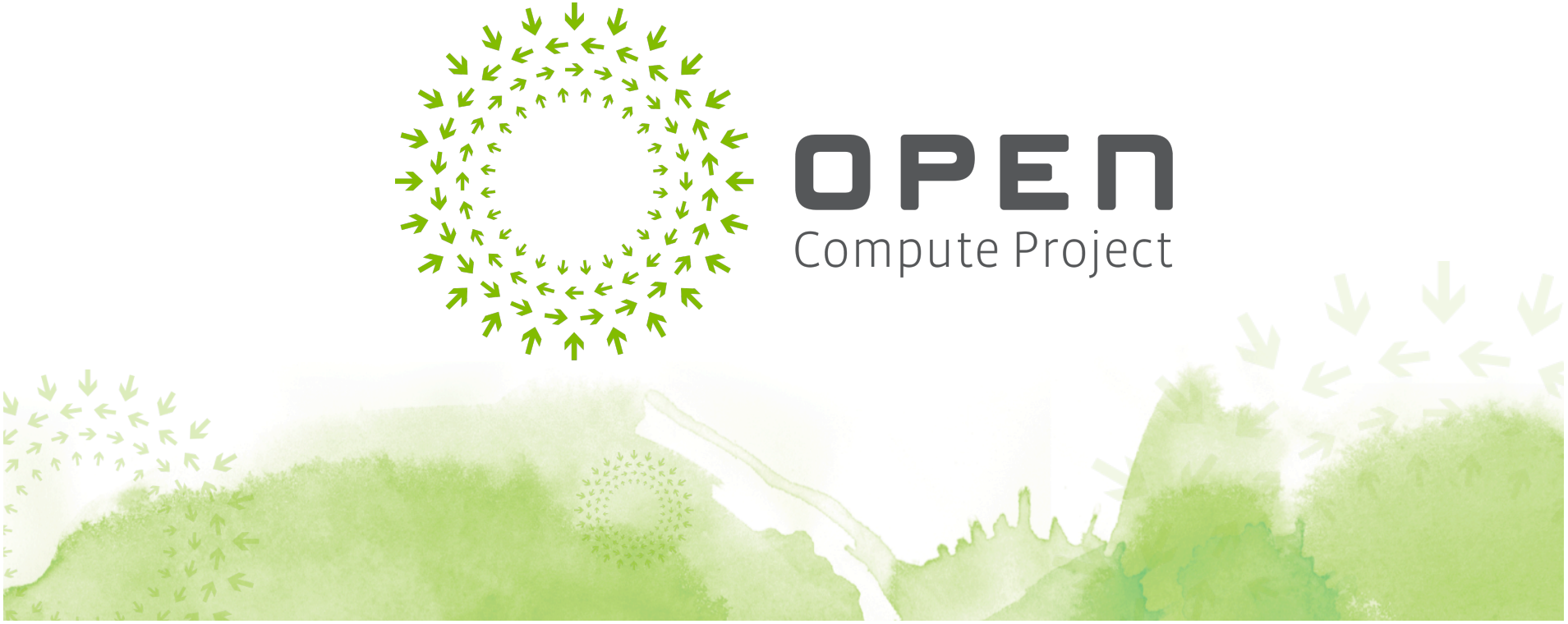
OCP Contribution

- 6-pack and wedge-100 OCP contribution
  - 6-pack specification
  - 6-pack design package
  - Wedge-100 specification
  - Wedge-100 design package

- Our ODM partner: Accton

http://www.opencompute.org/wiki/Networking/SpecsAndDesigns

Q&A


OPEN
Compute Project