Open. Together.
Minipack, Low Power Fabric Switch

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

**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
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
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
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

Open. Together.
Minipack Front View with PIM-16Q

- System LED
- PSU left air channel
- SCM
- PSU right air channel
- Pull Tag

PIM-2  PIM-3  PIM-4  PIM-5  PIM-6  PIM-7  PIM-8  PIM-9
Minipack Front View with PIM-4DD

- System LED
- PSU left air channel
- SCM
- PSU right air channel
- Pull Tag

PIM-2 PIM-3 PIM-4 PIM-5 PIM-6 PIM-7 PIM-8 PIM-9
Minipack Rear View

- FAN-5
- FAN-7
- PSU-3
- PSU-4
- FAN-8
- FAN-6
- FAN-4
- FAN-2
- FAN-3
- FAN-1
- PSU-1
- PSU-2
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

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

- 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

100G CWDM4-Lite
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 6x 12 DMO connector to SCM
SMB Components

- Tomahawk-3 ASIC
- SYS CPLD
- 14 phase 400A VR
- BMC
- 16-port OOB GbE Switch
- IOB FPGA
- Power Connectors to FCM-T and FCM-B
- Power Connectors to FCM-T and FCM-B
- Power Connectors to PDB-L
- Power Connectors to PDB-R
- DMO connectors to SCM and PIM

Components:
- SCM
- PIM-2
- PIM-3
- PIM-4
- PIM-5
- PIM-6
- PIM-7
- PIM-8
- PIM-9
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

- 16x QSFP28 ports
- DOM FPGA
- Gearbox BCM81724
- CoolPower Power Connector
- PIM alignment Pin
- EXAMAX DMO Connector to SMB
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 regeneration
- 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

- EXAMAX DMO Connector to SMB
- SCM Ejector
- PIM Alignment Pin
- CoolPower Power Connector
- COM-e
- BDW-DE CPU
- SODIMM on COM-e Module
- SATA M.2 SSD
- NVMe M.2 SSD
- OOB PHY
- OOB SFP
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

MiniLake

BDW-DE SoC

SCM

PIM #1

I2C Master

MIIM Master

DOM Buffer

QSFP #1

... #16

Gearbox #1

... #4

PIM #8

I2C Master

MIIM Master

DOM Buffer

QSFP #1

... #16

Gearbox #1

... #4

DOM FPGA

PCIe Bridge

sLPC Bridge

IOB FPGA

SMB

BMC

I2C

I2C Master

MIIM Master

DOM Buffer

QSFP #1

... #16

Gearbox #1

... #4

I2C Master

MIIM Master

DOM Buffer

QSFP #1

... #16

Gearbox #1

... #4

DOM FPGA

PCIe Bridge

sLPC Bridge

IOB FPGA

SMB

BMC

I2C

I2C Master

MIIM Master

DOM Buffer

QSFP #1

... #16

Gearbox #1

... #4

I2C Master

MIIM Master

DOM Buffer

QSFP #1

... #16

Gearbox #1

... #4

DOM FPGA

PCIe Bridge

sLPC Bridge

IOB FPGA

SMB

BMC

I2C
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
## MiniLake Module Design
- **MiniLake Feature List**

<table>
<thead>
<tr>
<th>Spec</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Board size (W x L )</strong></td>
<td>95mm x 125mm (COM Express® Basic Module form factor)</td>
</tr>
<tr>
<td><strong>CPU</strong></td>
<td>Xeon® Processor D-1527 : 4C, 35W, 2.2 GHz (Broadwell-DE SoC, Grangeville Platform)</td>
</tr>
<tr>
<td><strong>Memory</strong></td>
<td>Two DDR4 SODIMM slots</td>
</tr>
<tr>
<td></td>
<td>• Support DDR4, speeds up to 2133 MT/s with ECC</td>
</tr>
<tr>
<td></td>
<td>• Max Memory Capacity: 32 GB (16GB / DIMM)</td>
</tr>
<tr>
<td><strong>Storage</strong></td>
<td>One SATA 3.0 Port (6Gb/s)</td>
</tr>
<tr>
<td><strong>Network</strong></td>
<td>One 1000BAST-T (Intel® Ethernet Controller I210)</td>
</tr>
<tr>
<td></td>
<td>Two 10GBASE-KR (Intel® Broadwell-DE integrated LAN controller)</td>
</tr>
<tr>
<td><strong>PCIe</strong></td>
<td>PCIe Gen3 x24 Lanes (up to Six PCIe Gen3 x4 ports)</td>
</tr>
<tr>
<td></td>
<td>PCIe Gen2 x4 Lanes (up to Four PCIe Gen2 x1 ports)</td>
</tr>
<tr>
<td></td>
<td>One PCIe CLK</td>
</tr>
<tr>
<td><strong>Management Port</strong></td>
<td>I2C is the primary sideband interface for server management functionality connection to BMC of carrier board</td>
</tr>
<tr>
<td><strong>Bridge IC</strong></td>
<td>TI Snowflake</td>
</tr>
<tr>
<td><strong>I/O</strong></td>
<td>Two USB2.0 I/F and Two USB3.0 (w/ USB 2.0)</td>
</tr>
<tr>
<td></td>
<td>One UART port for Console</td>
</tr>
<tr>
<td></td>
<td>One SPI I/F for 2nd boot flash</td>
</tr>
<tr>
<td></td>
<td>One LPC I/F</td>
</tr>
</tbody>
</table>

**Networking Specifications**

**Design Files**
MiniLake Block Diagram

- ECC SO-DIMM DDR4-2133
- Intel i210
- USB2.0 x4 / USB3.0 x2
- 10GBase-KR x2
- UART x1
- PCIe Gen2 x4
- PCIe Gen3 x24
- PCIe-RefCLK x1
- SATA 6Gb/s x1
- CPU XDP

Network Connectivity

Type 7 COMe

Specifications

Design Files
### MiniLake Module Design - Key Components

#### TOP Side

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

#### BOT Side

<table>
<thead>
<tr>
<th>Item</th>
<th>Component</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>DDR4 SODIMM CONN-1</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>XDP CONN</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>Temperature Sensor-2 (TMP75)</td>
<td>I2C Address: 0x9C</td>
</tr>
<tr>
<td>10</td>
<td>1G NIC (INTEL i210)</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>TPM2.0 (Infineon SLB9670)</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>COMe B2B CONN</td>
<td></td>
</tr>
</tbody>
</table>
## Design Power Budget Estimation

<table>
<thead>
<tr>
<th>Component</th>
<th>Sub power (W)</th>
<th>Quantity (pcs)</th>
<th>Utilization (%)</th>
<th>Total power (W)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU (BDX-DE)</td>
<td>35</td>
<td>1</td>
<td>85%</td>
<td>29.75</td>
</tr>
<tr>
<td>DDR4 SODIMM</td>
<td>2.4</td>
<td>2</td>
<td>85%</td>
<td>4.08</td>
</tr>
<tr>
<td>Embedded Controller (Snow flake)</td>
<td>1</td>
<td>1</td>
<td>85%</td>
<td>0.85</td>
</tr>
<tr>
<td>I210-AT</td>
<td>1</td>
<td>1</td>
<td>85%</td>
<td>0.85</td>
</tr>
<tr>
<td>Discrete IO</td>
<td>3</td>
<td>1</td>
<td>85%</td>
<td>2.55</td>
</tr>
<tr>
<td>VR Loss</td>
<td>10</td>
<td>1</td>
<td>85%</td>
<td>8.5</td>
</tr>
<tr>
<td>System total power Consumption Estimation</td>
<td></td>
<td></td>
<td></td>
<td>46.58</td>
</tr>
</tbody>
</table>
MiniLake Module Design
- Mechanical

- PCB Size: 95mm(W) x 125mm(L)

- Clearance area on the Carrier Board.
MiniLake in Minipack System

Minipack System
12.8T Switch

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
Open. Together.

OCP Global Summit | March 14–15, 2019