Delivering High-Performance HPC Networking with Attractive Cost Models

Overview
Intel® Omni-Path Director Class Switches are the cornerstone of an end-to-end product family for HPC fabrics that delivers high performance with highly attractive cost models. Intel® Omni-Path Architecture (Intel® OPA) builds on proven technologies including the previous generation Intel® True Scale Architecture and the Cray® Aries interconnect and open source software to provide an evolutionary on-ramp to revolutionary new fabric capabilities.

Higher Performance at Lower Cost
Intel Omni-Path Director Class Switches deliver 100 Gbps port bandwidth with latency that stays low even at extreme scale. Based on Intel's 48-port switch silicon, these switches can lower fabric acquisition costs versus lower radix switches while simultaneously reducing space and power requirements. With these savings, you can achieve higher total cluster performance within the same hardware budget to expand and accelerate your research.

Flexible Fabrics at Every Scale
Intel Omni-Path Director Class Switches support mid-level clusters to supercomputers with 10s of thousands of servers. These switches can be used in combination with Intel Omni-Path Edge Switches to build low-latency, multi-tier fabrics with an exceptional set of features for high-speed networking.

Quality of Service (QoS)
All Intel OPA products feature excellent Quality of Service features including high availability that includes power, cooling and management module redundancy, component-level diagnostics and alarming, and out-of-band management. Beyond hardware availability, innovative Intel OPA features provide additional QoS options to the fabric. These include:

- Packet Integrity Protection (PIP) brings the ability to monitor for link errors with no added performance load, then resolves them.
- Traffic Fabric Optimization (TFO) can add traffic prioritization of data on an inter-switch link in 64-bit increments.
- Dynamic Lane Scaling (DLS) enables continuing operation of a link in a degraded mode, allowing the running job to complete, in the case of a lane failure.

Key Features
- Up to 28.8 Terabytes of aggregate BW
- Fully Redundant Subsystems
  - Spine/Management/Power/ Cooling

Next Generation Fabric Innovations
- Link Protection
  - Packet Integrity Protection (PIP)
  - "No-Load" Error Detection and Correction
  - Dynamic Lane Scaling (DLS)
  - Uptime Protection
- Quality of Service (QoS)
  - Traffic Fabric Optimization (TFO)
  - Larger MTUs (8/10K): Traffic Efficiency eg. Storage traffic

The Right Fabric for HPC and AI
Benefits
- Highly integrated design reduces space, power and cost
  - Up to 1,152 ports in 20U
  - Up to 288 ports in 7U
  - Expandable in 32/48 port
- Optimized for high message rates and low end-to-end MPI latency at scale
- Simple generational upgrades with:
  - Binary compatible applications
  - OFA Compliant, Same as InfiniBand*
  - FastFabric Installation/Tests Toolset
  - All open-source software
Physical Specifications

CHASSIS CONFIGURATIONS: 100SWD24 AND 100SWD06

<table>
<thead>
<tr>
<th>FEATURE</th>
<th>100SWD24</th>
<th>100SWD06</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>100 Gb port equivalents (min)</strong></td>
<td>32 (add in 32 or 48 port increments)</td>
<td>32 (add in 32 or 48 port increments)</td>
</tr>
<tr>
<td><strong>100 Gb port equivalents (max)</strong></td>
<td>1,152</td>
<td>288</td>
</tr>
<tr>
<td><strong>Total System Bandwidth (bi-di)</strong></td>
<td>28.8 TB/s</td>
<td>7.2 TB/s</td>
</tr>
<tr>
<td><strong>Chassis Height (+ Support Tray)</strong></td>
<td>20U (+0U)</td>
<td>7U</td>
</tr>
<tr>
<td><strong>Dimensions (w x h x d) (in/cm)</strong></td>
<td>17.6 x 35.0 x 29.5/44.7 x 88.9 x 74.9</td>
<td>17.6 x 12.2 x 29.5/44.7 x 31.0 x 74.9</td>
</tr>
<tr>
<td><strong>Weight (Empty FRU Chassis)</strong></td>
<td>256 lb/116.2 kg</td>
<td>103.5 lb/47 kg</td>
</tr>
<tr>
<td><strong>Weight (Base 1N: All Fans, 1 Mgmt Module, min PSUs)</strong></td>
<td>311.5 lb/141.3 kg</td>
<td>123.1 lb/55.9 kg</td>
</tr>
<tr>
<td><strong>Packaged Weight (fully configured)</strong></td>
<td>716 lb/324.8 kg</td>
<td>278.0 lb/126.1 kg</td>
</tr>
<tr>
<td><strong>Leaf Modules (max)</strong></td>
<td>24</td>
<td>6</td>
</tr>
<tr>
<td><strong>Cables</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32-Port Leaf</td>
<td>QSFP28</td>
<td>QSFP28</td>
</tr>
<tr>
<td>48-Port Leaf</td>
<td>QSFP-DD</td>
<td>Dual QSFP28</td>
</tr>
<tr>
<td><strong>Spine Modules (max)</strong></td>
<td>12 (48-port Leaf)</td>
<td>3 (48-port leaf)</td>
</tr>
<tr>
<td>8 (32-port Leaf)</td>
<td>2 (32-port Leaf)</td>
<td></td>
</tr>
<tr>
<td>12 (48/32 port mixed)</td>
<td>3 (48/32 port mixed)</td>
<td></td>
</tr>
<tr>
<td><strong>Fan Modules</strong></td>
<td>9</td>
<td>3</td>
</tr>
<tr>
<td><strong>Management Modules (standard/redundant)</strong></td>
<td>1 / 2</td>
<td>1 / 2</td>
</tr>
<tr>
<td><strong>Power Supplies (min/DC/AC redundancy)</strong></td>
<td>6 / 7 / 12</td>
<td>2 / 3 / 4</td>
</tr>
<tr>
<td><strong>Power (KW AC, typical-fully loaded, copper cables)</strong>†</td>
<td>8.5</td>
<td>1.98</td>
</tr>
<tr>
<td><strong>Power (KW AC, max-fully loaded, copper cables)††</strong></td>
<td>11.2</td>
<td>2.6</td>
</tr>
<tr>
<td><strong>Optical Transceiver—per port</strong></td>
<td>Class 2/3 (3.5W Max)</td>
<td>Class 2/3 (3.5W Max)</td>
</tr>
<tr>
<td>QSFP28</td>
<td>3.5W Max</td>
<td>3.5W Max</td>
</tr>
<tr>
<td>QSFP-DD</td>
<td>229 W DC</td>
<td>293.5 W AC</td>
</tr>
<tr>
<td><strong>Cooling (max required)</strong></td>
<td>1100 CFM at 40°C</td>
<td>390 CFM at 40°C</td>
</tr>
<tr>
<td><strong>Status LEDs††† (Ethernet/DC_On)</strong></td>
<td>2 / 1</td>
<td>2 / 1</td>
</tr>
</tbody>
</table>

† Typical: Fully populated chassis (max leafs, max spines, 2MM, AC Redundant PSUs), fully populated copper cables, running loop test (estimated).
†† Max: Fully populated chassis (max leafs, max spines, 2MM, AC Redundant PSUs), fully populated copper cables, worst process/temp/voltage (estimated).
††† Status LEDs - Ethernet Activity (Green), Ethernet Speed (Green/Orange) / DC_On (Green on push button).

**MODULES**

<table>
<thead>
<tr>
<th>MODULE</th>
<th>DIMENSIONS (W x H x D) (in/cm)</th>
<th>WEIGHT</th>
<th>PACKAGE DIM (W x H x D) (in/cm)</th>
<th>PACKAGED WEIGHT</th>
<th>POWER (W, typical)†</th>
<th>POWER (W, max)††</th>
</tr>
</thead>
<tbody>
<tr>
<td>QSFP 32p Leaf Module</td>
<td>6.9 x 2.2 x 12.3/17.5 x 5.6 x 31.2</td>
<td>8.5 lb/3.86 kg</td>
<td>11.5 x 6.5 x 17.0/29.2 x 16.5 x 43.2</td>
<td>10.0 lb/4.54 kg</td>
<td>175 W DC 189 W AC</td>
<td>229 W DC 247.5 W AC</td>
</tr>
<tr>
<td>QSFP-DD 48p Leaf Module</td>
<td>6.9 x 2.2 x 12.4/17.5 x 5.6 x 31.4</td>
<td>8.5 lb/3.86 kg</td>
<td>11.5 x 6.5 x 17.0/29.2 x 16.5 x 43.2</td>
<td>10.0 lb/4.54 kg</td>
<td>207 W DC 223.5 W AC</td>
<td>271 W DC 293.5 W AC</td>
</tr>
<tr>
<td>Spine Module</td>
<td>17.4 x 2.2 x 7.5/44.2 x 5.6 x 19.1</td>
<td>8.6 lb/3.9 kg</td>
<td>22.5 x 6.5 x 11.5/57.2 x 16.5 x 29.2</td>
<td>10.5 lb/4.76 kg</td>
<td>203 W DC 219 W AC</td>
<td>262 W DC 283 W AC</td>
</tr>
<tr>
<td>Fan Module</td>
<td>3.2 x 3.3 x 8.7/8.1 x 8.4 x 22.1</td>
<td>1.6 lb/0.73 kg</td>
<td>7.5 x 7.5 x 11.0/19.1 x 19.1 x 27.9</td>
<td>2.40 lb/1.09 kg</td>
<td>45 W DC 48.5 W AC</td>
<td>64 W DC 69 W AC</td>
</tr>
<tr>
<td>Power Supply Module†††</td>
<td>4.0 x 1.6 x 1.5/10.2 x 4.1 x 40.1</td>
<td>6.55 lb/2.98 kg</td>
<td>8.0 x 5.5 x 19.5/20.3 x 14.0 x 49.5</td>
<td>7.55 lb/3.42 kg</td>
<td>94% eff. at 230 V/50% load</td>
<td>91% eff. at 230 V/max load</td>
</tr>
<tr>
<td>Management Module</td>
<td>8.6 x 1.7 x 7.2/21.8 x 4.3 x 18.3</td>
<td>1.75 lb/0.79 kg</td>
<td>11.0 x 4.9 x 9.5/28.0 x 12.5 x 24.1</td>
<td>2.65 lb/1.2 kg</td>
<td>17W DC 18.5 W AC</td>
<td>25 W DC 27 W AC</td>
</tr>
</tbody>
</table>

† Typical: copper cables for Leaf (all estimated).
†† Max: copper cables for Leaf, worst process/temp/voltage (all estimated).
††† Weight includes AC Cord; When a form/fit option exists, use of a DC/DC supply may be possible.
Environmental Specifications
(planned for 48p Leaf Module#)

Temperature
Operating: 5° to 40° C, supporting AOC up to 2 W (Temperature de-rating 5°C/W up to 3 W AOC) supporting QSFP-DD AOC up to 3.5 W
Storage: -40° to 70° C

Humidity
Operating: 5% to 85% non-condensing
Storage: 5% to 95% non-condensing

Altitude
Operating: 0 – 10,000 feet (Temperature de-rating 1°C / 175 M above 900 M)
Storage: 0 – 40,000 feet

Shock
Unpackaged: Half-sine, 2g 11ms 300 pulses total
Packaged: 9” vertical and rotational drop

Vibration
Unpackaged: 5-500 Hz, 2.2 g RMS random
Packaged: 5-500 Hz, 1.09 g RMS random

Airflow
Front-to-back
1100 CFM maximum for 24 slot
390 CFM maximum for 6 slot

RoHS/REACH
Complies with RoHS II Directive 2011/65/EU of the European Parliament
Complies with REACH Regulation (EC) No 1907/2006

Acoustics
Less than 7.0 Bels

Agency Approvals - EMI and EMC
US/Canada
FCC Part 15, Subpart B, Class A
CAN ICES-3 (A)

Europe/International
CISPR22
CISPR32/EN55032
EN55024
EN61000-3-2
EN61000-3-3

Japan
VCCI, Class A

New Zealand/Australia
AS/NZS CISPR 22, Class A

Korea
RRA/KC (KN32, KN35), Class A

Taiwan
BSMI (CNS 13438), Class A

Customs Union: Russian Federation, Belarus and Kazakhstan
TR CU 020/2011 “Electromagnetic compatibility of technical equipment”

Agency Approvals – Safety (planned for 48p Leaf Module#)

US/Canada
TUV NRTL: UL 60950-1, CSA 22.1.No. 60950-1

Europe
TUV: EN60950-1

International
CB Scheme: IEC 60950-1

Customs Union: Russian Federation, Belarus and Kazakhstan
TR CU 004/2011 “On Safety of Low-Voltage Equipment”

Country of Origin
USA

Miscellaneous
Non-volatile memories
A link will be provided to the statement of volatility document

---

1 A sample configuration assumes a 750-node cluster, and number of switch chips required is based on a full bisectional bandwidth (FBB) Fat-Tree configuration. Intel® OPA uses one fully-populated 768-port director switch, and Mellanox EDR solution uses a combination of 648-port director switches and 36-port edge switches. Mellanox component pricing from www.kernelsoftware.com, with prices as of December 30, 2017.