OPEN.
FOR BUSINESS.
Facebook Latest server/ML hardware Overview

Whitney Zhao/Hardware Eng./Facebook Inc.
Agenda

- Overview
- 1S Server Yosemite V2
- 2S Server Tioga Pass
- AI/ML Server Big Basin V2
<table>
<thead>
<tr>
<th>Year</th>
<th>Facebook</th>
<th>Messenger</th>
<th>Instagram</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010</td>
<td>600M</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2012</td>
<td>1B</td>
<td>Intro</td>
<td>Acquisition</td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>1.3B</td>
<td>200M</td>
<td>200M</td>
<td></td>
</tr>
<tr>
<td>2016</td>
<td>1.65B</td>
<td>900M</td>
<td>500M</td>
<td>1B</td>
</tr>
<tr>
<td>2018</td>
<td>2.2B</td>
<td>1.3B</td>
<td>1B</td>
<td>1.5B</td>
</tr>
</tbody>
</table>
A history of building for scale

2011
- Data Center
- Triplet Rack
- Battery Cabinet
- Freedom Servers
- Spitfire Server (AMD)
- Power Supply

2012
- Windmill (Intel)
- Winterfell
- Mezzanine Card V1

2013
- Knox
- Open Rack V2

2014
- BluRay
- Mezzanine Card V2
- Open Rack V1

2015
- Leopard
- Cold Storage
- Micro Server (Panther)

2016
- Wedge 100
- Big Sur
- Honey Badger

2017
- Backpack
- Lightning
- Yosemite V2

2018
- Big Basin V2
- Twin Lake
- OCP NIC 3.0
- Tioga Pass
Modular Design 1S Server Yosemite V2
Yosemite V2: Modular platform
System Overview

- Flexible Multi-Node 1S platform
- Twin Lakes 1S server
- Open Rack V2 compatible with new 40U vCubby design
- Support 4x 1S server cards or 2x 1S server + 2x Device cards
- Single smart baseboard to handle different configurations
- 100G/50G Multi-Host Network options
Twin Lakes is architected for high compute performance

http://files.opencompute.org/oc/public.php?service=files&t=9d240277a11b83f9764420e99e18ac38

<table>
<thead>
<tr>
<th></th>
<th>Intel® Xeon® D-2100 Processor</th>
<th>Intel® Xeon® D-1500 Processor</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CPU Cores</strong></td>
<td>Up to 18 with Intel® HT</td>
<td>Up to 16 with Intel® HT</td>
</tr>
<tr>
<td><strong>Cache</strong></td>
<td>LLC: 1.375 MB/Core</td>
<td>LLC: 1.5KB/Core</td>
</tr>
<tr>
<td></td>
<td>MLC: 1MB/Core</td>
<td>MLC: 256K/Core</td>
</tr>
<tr>
<td><strong>Memory</strong></td>
<td>4 Channels,</td>
<td>2 Channels</td>
</tr>
<tr>
<td></td>
<td>DDR4 1866/2133/2400/2666*</td>
<td>DDR4/DDR3L 1600/1866/2133</td>
</tr>
<tr>
<td><strong>PCIe</strong></td>
<td>CPU: x32 PCIe Gen 3</td>
<td>CPU: x24 PCIe Gen 3 lanes</td>
</tr>
<tr>
<td></td>
<td>• Twin Lakes uses all 32 lanes</td>
<td>FlexIO: x8 PCIe Gen 2 lanes</td>
</tr>
<tr>
<td></td>
<td>FlexIO: x20 PCIe Gen 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Twin Lakes uses 9 FlexIO PCIe lanes</td>
<td></td>
</tr>
<tr>
<td><strong>Acceleration Engines</strong></td>
<td>AVX512 Intel® QuickAssist Technology</td>
<td>AVX256 No built-in Intel® QAT</td>
</tr>
<tr>
<td></td>
<td>• Up to 100 Gbps Crypto/Compression</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• 100 KOpS PKE 2K</td>
<td></td>
</tr>
</tbody>
</table>

*Twinkle validated at 2400 MT/s
General Purpose 2S Server Tioga Pass
Overview

- Facebook general purpose OCP 2S server
- Higher performance over previous gen
- Modularity
- Easy to service
- Open

https://www.opencompute.org/products/108/wynn-tioga-pass-standard-sx7220g3-1-2u-ocp-server-up-to-768gb-8gb-16gb-32gb-ddr4-up-to-2666mts-12-dimm-slots
http://files.opencompute.org/ou/public.php?service=files&file=wiynn-tioga-pass-standard-sx7220g3-1-2u-ocp-server-up-to-768gb-8gb-16gb-32gb-ddr4-up-to-2666mts-12-dimm-slots
System Overview

- Intel Skylake-SP Processor up to 165w
- Single/Double Sided DIMM w/ 12 memory channels
- PCIe X32 to front I/O
- 3.5”HDD/Nvme M.2 ssd as boot drive
- Support up to 100G OCP NIC2.0
- Half width/High density
- Open Rack V2
Configuration

- Compute servers
- Compute server w/ PCIe AIC
- Headnode for PCIe expansion box
- Boot drive
  - 3.5” HDD
  - Nvme M.2 SSD
- NIC Card
  - OCP NIC2.0 25G/50G/100G
  - Intel KR Mezz Card
Optimize Power & Performance for FB workload

In search of best efficiency for scaling out

F(X)

F(X) = System performance / \( w \)

\( X = \{ x_1, x_2, \ldots, x_n \} \)

Parameters to explore across different SKUs

<table>
<thead>
<tr>
<th>( X_1 )</th>
<th>TDP</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X_2 )</td>
<td>Core Count</td>
</tr>
<tr>
<td>( X_3 )</td>
<td>Frequency</td>
</tr>
<tr>
<td>( X_4 )</td>
<td>Tcase</td>
</tr>
<tr>
<td>( X_5 )</td>
<td>Others</td>
</tr>
</tbody>
</table>
Big Basin Machine learning Training Server
Impact

Facebook’s commitment to developing AI & advancing ML
Goal

- Open, full contribution to OCP
- Disaggregation/Modularity
- Serviceability
Big Basin V2 Overview

- 3 OU chassis
- Open Rack v2 compatible
- 8x Nvidia Tesla V100 GPUs; NVLink capable
- 300W TDP for each Tesla V100 GPU
- Facebook 2S Server Tioga Pass as Head node
A deeper look into Big Basin

- Baseboard on sliding tray
- IO board
- Baseboard
- Midplane board
### Performance

- **Comparisons of GPU Hardware**

<table>
<thead>
<tr>
<th>Metrics</th>
<th>NVIDIA V100</th>
<th>NVIDIA P100</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Performance</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FP-32</td>
<td>15 TFLOPS</td>
<td>10.6 TFLOPS</td>
<td>1.42x</td>
</tr>
<tr>
<td>FP-16</td>
<td>30 TFLOPS</td>
<td>21.2 TFLOPS</td>
<td></td>
</tr>
<tr>
<td>TensorCore</td>
<td>125 TFLOPS</td>
<td>NA</td>
<td>Up to 5x</td>
</tr>
<tr>
<td>Mem Bandwidth</td>
<td>900 GB/s</td>
<td>720 GB/s</td>
<td>1.25x</td>
</tr>
<tr>
<td>NVLink</td>
<td>300 GB/s</td>
<td>160 GB/s</td>
<td>1.88x</td>
</tr>
<tr>
<td><strong>Power</strong></td>
<td>300 W</td>
<td>300 W</td>
<td></td>
</tr>
</tbody>
</table>
Performance

• Comparisons of GPU Hardware

• Head-node upgrade: Tioga Pass
  • New CPU architecture: Broadwell-SP to Skylake-SP
  • Double PCIe bandwidth
  • Upgraded 100G NIC

• CUDA 9 + cudnn 7: faster libraries, etc.
OPEN. FOR BUSINESS.