Project Olympus
Overview

Mark Shaw
Hardware Engineering
Director, Azure Cloud
Hardware Infrastructure

Badriddine Khessib
Firmware Engineering
Director, Azure Cloud
Hardware Infrastructure
<table>
<thead>
<tr>
<th>Next-gen Cloud Hardware</th>
<th>Development Model</th>
<th>Industry Ecosystem</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open sourcing leading edge Hyperscale cloud hardware currently under development at Microsoft</td>
<td>New collaboration model with OCP community – co-develop open hardware at cloud speed</td>
<td>Bootstrap a vibrant ecosystem in OCP for the next generation of datacenter hardware</td>
</tr>
</tbody>
</table>

https://github.com/opencomputeproject/Project_Olympus
1U SERVER

PROJECT OLYMPUS SERVER SETS THE STANDARD

Next generation Intel® Xeon® Processors, codename Skylake

GbE Redfish Management

GbE Redfish Management

NCSI, KVM, OCP mezz enabled

Up to 16 M.2 CloudSSD

3 PCI-E x16 FHHL slots

DDR4 memory
POWER SUPPLY

OPTIMIZED FOR EFFICIENCY, RELIABILITY, HYPERSCALE DATACENTERS

• Three x 340W PSUs Fully Integrated
  - Three-phase balanced AC power
  - 680W N+1 (1020W total)
  - Dual-feed auto-selection (IVS)

• Fault Mode Resiliency
  - AC feed failure, automatic fail over
  - PSU failure caps power if necessary
  - N+1 HA => no repair on failure
  - Double fault will be extremely rare
ECOSYSTEM MOTHERBOARDS

DROP IN OR ADAPT TO PROJECT OLYMPUS SERVER CHASSIS
UNIVERSAL POWER

A/C POWER ADAPTED TO YOUR DATA CENTER

• Supports high availability data centers
  − Dual-feed, three-phase A/C
  − Rack power monitoring and capping
• Universal Power Distribution
  − Distributes power and management
  − Integrated Rack Manager
  − Supports blind-mate servers
• A/C power cord adaptation
  − Racks never change, only A/C cord
    • 208V-30A, 208V-50A, 415V-30A, 400V-32A
19” EIA RACK

NO-FRILLS HIGH CAPACITY RACKS

• EIA-310 Standard
  - 600mm wide, 1200mm deep
  - 42U & 48U options
  - 3,000 lb capacity

• Features
  - Integrates with Project Olympus modules
  - Front & rear locking door, sidewall options
  - Three EIA rails for standard equipment
  - Baffles and air blocking panel options
PROJECT OLYMPUS
RACK MANAGEMENT
RACK MANAGEMENT

DRIVING UP DATA CENTER UTILIZATION

• Rack Management
  - Restful API I/F or Redfish via Ethernet
  - Rack Manager (RM) ARM CPU
  - Integrated into PDU

• Blade Management flexible to your needs
  - GbE I/F to each blade’s BMC
  - NCSI enabled, cable to OCP Mezz Carrier
  - KVM enabled on motherboard

• Standalone, 1U rack mount version
  - For hardware that does not use the PDU
RACK MANAGEMENT: GUIDING PRINCIPLES

- Simplicity
- Scalability
- Cost
- Security
- Openness
Rack Manager Overview

- **Rack Manager Functions**
  - Power Management
  - Out-of-band Server Management
- **RM Instances in Cloud**
  - Discrete (Consumes 1U)
  - Integrated into PMDU
- **Communication**
  - Network
  - TTY Console
- **Hardware Signaling**
  - Server Presence
  - Server On/Off
  - Server Throttle
- **Power Metering & Control**
- **Remote Debug**
- **Remote Media**
- **Out of band FW update and recovery**
  - UEFI, CPLD, FPGA, PSU
Rack Manager Block Diagram

- LEDs - Attention, Power, Debug, Status
- PCIe x16 edge finger & PCIe x8 edge finder to Interface with PMDU or backplane
- Temp Sensor
- Humidity Sensor
- FRU
- GPIO Buffers
  - x48 for Blade presence
  - x48 for Blade enable
  - GPIOs for Boot strap, Throttle bypass, power control etc.
- QSPI
- eMMC
- 1GB DDR3L
- To Fabric
- To Mgmt. Switch
- To DIGI

Debug
1GB DDR3L
PCIe x16 edge finger & PCIe x8 edge finder to Interface with PMDU or backplane
GPIO Buffers
x48 for Blade presence
x48 for Blade enable
GPIOs for Boot strap, Throttle bypass, power control etc.
RACK MANAGEMENT: Lessons Learned

- ARM vs. X86
- Linux vs. Windows
- GbE vs. RS232
- Rack vs. 12U Chassis
- Redfish vs. custom API
RACK MANAGEMENT: WHAT’S NEW

• Rack/Row/DC level Dynamic Power Capping
• Presence and Slotld
• Remote JTAG debug
• OOB FW update
• Odata client for auto-discovery
Auth at Interface:
- Ethernet / REST
- CLI / TTY

Privilege:
- Permissions token
- Log / Audit

Execution:
- Task execution
- Logical to Physical

Action:
- Hardware Control
STANDALONE RACK MANAGER

- 1U Rack Manager Assembly
- Reuse of Rack Manager
- Supports Non-WCS Rack Management
- Support Row Management
- Redundant AC Input
- Single SW image
PROJECT OLYMPUS
ADD-ON MODULES
• 4U Flexible and Configurable Chassis
  - Eight NVIDIA SXM2 high-performance GPU
  - Eight PCI-E based GPU cards
  - Four additional x16 I/O slots

• Scalability
  - Up to 32 GPUs with four chassis communicating across PCI-Express fabric

• Flexible Topologies
  - Four PCI-E switches, eight x16 cable ports
  - Config Peer-to-Peer Bandwidth via cables
HDD STORAGE

HIGH DENSITY, HIGH RELIABILITY

• 4U JBOD – 88 HDDs / chassis
  - Cold-aisle hot repair service
• Robust Feature Set
  - Minimized Rotational Vibration, Acoustic Noise
  - BMC/Fans ensure <51C Temp
  - Individual HDD on/off to minimize NTF
• Partitionable with Project Olympus Servers
  - 88 HDDs, 1.2PB, on one server
  - 44 HDDs, 600TB, on two servers
  - 22 HDDs, 300TB, on four servers
FLASH STORAGE

HIGH PERFORMANCE

• 1U JBOF – 64 nvme M.2s / chassis
  - 64TB – 256TB capacity
  - 16 hot-plug M.2 carriers, 4 M.2’s each
  - Cold-aisle hot repair service
• Robust Feature Set
  - OpenBMC manages chassis, fans, power
  - Individual power domain per carrier

Partner soon to be announced
HOW TO DOWNLOAD

COMMUNITY PARTICIPATION

https://github.com/Project-Olympus/rackmanager

https://github.com/opencomputeproject/Project_Olympus
<table>
<thead>
<tr>
<th>Day</th>
<th>Time</th>
<th>Title</th>
<th>Presenter</th>
<th>Session Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weds, March 8th</td>
<td>12:30pm</td>
<td>Designing Open Source Hardware to Inspire an Ecosystem</td>
<td>Mark Shaw</td>
<td>Expo Hall Theatre</td>
</tr>
<tr>
<td></td>
<td>2:35pm</td>
<td>Enabling Cloud Workloads Through Innovations in Silicon</td>
<td>Leendert van Doorn</td>
<td>Executive Track</td>
</tr>
<tr>
<td>Thurs, March 9th</td>
<td>9:00am</td>
<td>Microsoft Project Olympus Overview</td>
<td>Mark Shaw, Siamak Tavallaei</td>
<td>Workshop</td>
</tr>
<tr>
<td></td>
<td>10:00am</td>
<td>Microsoft Project Olympus Servers</td>
<td>Mark A. Shaw</td>
<td>Workshop</td>
</tr>
<tr>
<td></td>
<td>11:00am</td>
<td>Microsoft Project Olympus High Density Flash</td>
<td>Mark A. Shaw</td>
<td>Workshop</td>
</tr>
<tr>
<td></td>
<td>11:30am</td>
<td>Microsoft Project Olympus Storage JBOD</td>
<td>Bruce Hoch</td>
<td>Workshop</td>
</tr>
<tr>
<td></td>
<td>1:00pm</td>
<td>Microsoft Project Olympus GPU Expansion</td>
<td>Siamak Tavallaei</td>
<td>Workshop</td>
</tr>
</tbody>
</table>