Open. Together.
How Open Technology Helped DOE Labs Place Sixteen Supercomputers on the Top500

Sid Mair, SVP Federal Systems, Penguin Computing
About Penguin Computing

- 20 years of artificial intelligence (AI), engineering, and computer science for startups, Fortune 500, government, and academic organizations
- Specialized high-performance computing (HPC), bare metal HPC in the cloud, AI, and storage technologies
- Coupled with leading-edge design, implementation, hosting, and managed services including sys-admin and storage-as-a-service, and highly rated customer support
- More than 2,500 customers in 50 countries across nine major vertical markets
- Over 300 OCP racks delivered to date based on Tundra™ Extreme Scale Design

Since 2012
OCP Clusters in Top500

Sixteen supercomputers designed and built by Penguin Computing placed on the Top500 List since 2016

- All OCP based & deployed in U.S. national labs as part of the U.S. Department of Energy as an alternative to explosive test-based confidence
- Part of the DOE’s Advanced Simulation and Computing (ASC) program
- Provide simulation-based confidence in the nuclear stockpile, an alternative to explosive test-based confidence
CTS-1: A Perfect Opportunity for OCP Design

- Commodity Technology System 1 (CTS-1) supports National Nuclear Security Administration (NNSA) to ensure nuclear stockpile stewardship in compliance with the Comprehensive Nuclear-Test-Ban Treaty (CTBT) between the U.S. and the former Soviet Union.
- 30,000 Broadwell / Skylake dual processor nodes to date.
- Commodity clusters brought down the cost of HPC systems from approximately $100 million per teraFLOP in 1995 to less than $5,000 per teraFLOP today (factor of 20,000) with greater computing power and energy efficiency with each generation.
- Exemplifies how OCP-based technology can give organizations both value and performance.
- Provides flexibility in CPU Architectures, accelerators, and interconnects.
The Key -- Uniquely Flexible, Dense Tundra Design

Baseline Tundra™ Extreme Scale

• First was 1OU, 3 node, CPU compute, housed in v1 rack
• Supports up to 102 nodes per rack with switching
• May include GPGPU accelerated servers
• High-speed, low latency interconnects
  • Workloads synchronized within microseconds between nodes
• Air or liquid cooling (Direct-to-Chip or Rear Doors)
• Latest generation open bridge rack
• Accessible via cloud through Penguin Computing On-Demand® (POD)
• Multiple storage options
• (Details: https://www.opencompute.org/products/309/penguin-tundra-extreme-open-bridge-rack)
Vertiv HPC Power Shelf (3 x 12V DC Bus Bar)

- Shelf Dimensions: 3U (H) x 19in (W) x 25.8in (D)
- Many voltage options available (176-305Vac)
  - Nine (9) slots for 3300W Rectifiers and BBUs
  - 208V Single Power System Output (Max) = 16.5kW (N+1)
  - 277/480V Single Power System (Max) = 26.4kW (N+1)
  - 277/480V Dual Zone Configuration (Max) = 52.8kW (N+1)
- Redundant power options available
- Provides 3 Pair of DC Output Bus Bar Connections
- Temperature Env -10C to +45C (+14F to +113F)
- 2 AC Convenience Outlets for switching (Gen II)
Compute Node: Relion 1930e

CPU Processing
- Three Nodes in 1 OpenU Form Factor
- Dual Socket Intel® Xeon® E5-2600v4 per node
- Up to 1TB DDR4-2400MHz (8x DIMMs) per node
- Intel® C612 Chipset
- 1x Dedicated BMC
- 1x PCIe 3.0 x16
- 1x 2.5" Fixed SATA SSD
- Dual 1GbE/RJ45 LOM
- Optional 1x FDR LOM
- Supports Asetek Direct-to-Chip Cooling
- OCP-inspired product coming to the marketplace in Q2 2019
Compute Node: Relion XO1114GT

GPU Computing
• 1 OU Form Factor
• Dual Intel Xeon Cascade Lake-SP / Skylake-SP with Intel Omni-Path
• Up to 2TB DDR4-2933MHz (16x DIMMs)
• Intel® Lewisburg Chipset
• 1x Dedicated BMC
• Supports 4x GPGPU
• Nvidia Tesla Volta-Pcie
• 2x PCIe 3.0 Low profile (Speed depends on topology)
• 4x 2.5" SATA SSD
• Dual 1GbE/RJ45 LOM
• Support Asetek Direct-to-Chip Cooling (CPUs and GPUs)
• Flexible PCIe Topology

OCP-inspired product coming to the marketplace in Q2 2019
<table>
<thead>
<tr>
<th>LLNL</th>
<th>SNL</th>
<th>LANL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Opal(1)</td>
<td>Solo(2)</td>
<td>Kit (0.25)</td>
</tr>
<tr>
<td>Jade(14)</td>
<td>Cayenne/LC(6)</td>
<td>Fire(6)</td>
</tr>
<tr>
<td>Quartz(14) +2SU/LC</td>
<td>Serrano/LC(6)</td>
<td>Ice(6)</td>
</tr>
<tr>
<td>Shepards Crest(1)</td>
<td>DarkGhost(4)</td>
<td>Grizzly/LC(8)</td>
</tr>
<tr>
<td>Mica(2)</td>
<td>Eclipse/LC (8)</td>
<td>Snow/LC(2)</td>
</tr>
<tr>
<td>Nel(6)</td>
<td>Doom/LC(.5)</td>
<td>Hail(0.75)</td>
</tr>
<tr>
<td>Topaz/LC(4)</td>
<td>“H”/LC (8)</td>
<td>Lysander(0.25)</td>
</tr>
<tr>
<td>Pascal/LC(1.5)</td>
<td>Badger/LC(2.5)</td>
<td></td>
</tr>
<tr>
<td>Corona (1)</td>
<td>Kaleo/LC (2.5)</td>
<td></td>
</tr>
<tr>
<td>Pinot2 (1)</td>
<td>Sage (1.2)</td>
<td></td>
</tr>
<tr>
<td>GS#1 (1)</td>
<td>Cyclone (6)</td>
<td></td>
</tr>
<tr>
<td>GS#2 (1)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GS#3 (.5)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key - Number of “Scalable Units” = (X)
1 “Scalable Unit” = 3 OCP Racks plus management at approximately 62 nodes per rack
Flexible Architecture Enables Scalable Configurations

Tundra Extreme Scale, Xeon E5-2695v4 18C 2.1GHz, Intel Omni-Path

• LLNL “Quartz” - originally 14SU, expanding 2SU in 2019 -- TTL will be 16SU.
• LLNL “Jade” - 14SU
• LANL “Grizzly” w/ “Snow” is 10SU
• SNL “Serrano” and “Cayenne” - separate 6SU
• SNL “Eclipse” - base 6SU purchased in 2017, 2SU expanded in 2018
• LANL “Badger” - purchased in 2017, doubled in 2018
• LANL “Kodiak” (GPU cluster) - purchased in 2017, expanded in 2018
• LANL “Fire”, “Ice” and “Cyclone” - all separate 6SU
NEW Tundra-based “Corona” AI Cluster leverages flexibility, density, efficiency of OCP

- AMD EPYC™ processors, AMD Radeon™ Instinct™ GPU accelerators
- 383 teraFLOPS (floating point operations per second)
- 170 two-socket nodes incorporating 24-core AMD EPYC™ 7401 processors and a PCIe 1.6 Terabyte (TB) nonvolatile (solid-state) memory device
- Half of compute nodes utilize 4 AMD Radeon Instinct™ MI25 GPUs per node, delivering 4.2 petaFLOPS of FP32 peak performance
- Connected via a Mellanox HDR 200 Gigabit InfiniBand network
- Remaining compute nodes may be upgraded with future GPUs
- Uses OCP-inspired products coming to the marketplace in Q2 2019
Penguin Computing Tundra Extreme Scale
Learn more at:

www.penguincomputing.com/tundra

https://www.opencompute.org/products/309/penguin-tundra-extreme-open-bridge-rack