Edison Overview

Richard Gerber
NERSC User Services Deputy Group Lead

February 14, 2013
Edison Phase I is Here!

From a User Perspective Edison is very similar to Hopper!
# Edison Addresses the NERSC Workload Needs

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Description</th>
<th>Comment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Processor</td>
<td>Intel Ivy Bridge (Phase 1: Sandy Bridge)</td>
<td>Fast, cutting-edge, commodity processor</td>
</tr>
<tr>
<td>Node</td>
<td>Dual-socket, 64 GB 1866 MHz memory</td>
<td>Large memory per node Excellent memory bandwidth</td>
</tr>
<tr>
<td>Interconnect</td>
<td>Cray Aries, dragonfly topology</td>
<td>Excellent latency &amp; bandwidth Excellent scaling Adaptive routing eases congestion</td>
</tr>
<tr>
<td>Storage</td>
<td>6.48 PB 140 GB/sec I/O bandwidth, 3 file systems</td>
<td>Large, dedicated scratch storage High bandwidth; better metadata</td>
</tr>
</tbody>
</table>
Edison - Cray XC30 Phase 1 [Phase 2]

- 10K [104K] compute cores
- Cray ”Aires” interconnect
- Two 8-core Intel ’Sandy Bridge' 2.6 GHz processors per node [Phase 2: Ivy Bridge]
- 16 [TBA] processor cores per node
- 64 GB of memory per node @1866 MHz
- 42 [333] TB of aggregate memory
- Edison Phase 1 access Feb. 2013 [Phase 2 in Summer or Fall 2013]

- 4 [TBA] GB memory per core for applications
- 1.6 / 6.4 PB of scratch disk
- CCM compatibility mode available
- Intel, Cray, GNU compilers
Installation Timeline

- **Phase 0**
  - 6 login nodes ✓
  - 1 file system (35GB/s, 1.6PB) ✓

- **Phase 1 (Now)**
  - 4 cabinets with ~10K Sandy Bridge processors ✓
  - Start Phase 1 acceptance test in February ✓
  - Start DARPA Mission Partner Access ✓
  - Add 2 file systems (105GB/s, 4.8PB)

- **Phase 2 (early July)**
  - Add 24 cabinets with Ivy Bridge processors
  - Upgrade first 4 cabinets with Ivy Bridge
  - About 100K cores total
  - Will require downtime of a few weeks
**Cascade Compute Blade**

- Faceplate assembly
- Card Edge Stiffener (2)
- QPDC 0 PCIe Gen3 Connections
- QPDC 1 PCIe Gen3 Connections
- Aries
- FPGA
- Tolopia
- 52V-12V Bus Converter (2)
- Power and Backplane Connector
- Inter-group Cable Connectors

**SPECIFICATIONS**
- Module power: 2014 Watts
- PDC max. power: 900 Watt
- Air flow req.: 275 cfm
- Size: 2.125 in x 12.95 in x 33.5 in
- Weight: <40 lbm
Aries Network

- Dragonfly Architecture
- Configured in 3 ranks:
  - Rank-1 is chassis level
  - Rank-2 is cabinet level
  - Rank-3 is system level
- Global bandwidth tuned by number of optical cables (rank-3)
- Within a 2 cabinet group
  - Minimal routing – 2 hops
  - Non-minimal routing – 4 hops
Edison Network Configuration

• **Phase 1**
  – 4-cabinet system consists of 2 fully-populated two-cabinet groups
  – Rank-3: 42-wide bundles of optical cables (42 Optical Cables)

• **Phase 2**
  – 28-cabinet system consists of 14 fully-populated two-cabinet groups
  – Rank-3: 6-wide bundles of optical cables (546 Optical Cables)
Edison Scratch File System (March)

3 Scratch File Systems

2 - 1.6 PB @ 35 GB/sec
1 - 3.2 PB @ 70 GB/sec

• Multiple file systems to reduce contention, improve metadata performance
• One 2X Hopper’s size & performance for I/O intensive applications

Edison is also connected to NERSC Global Homes, Global Scratch and Project file systems

Cray Sonexon 1600 Lustre Appliance
External Login Nodes

- Phase 1: 6 nodes, Phase 2: 12 nodes
- Quad processor Sandy Bridge @ 2 GHz
- 512 GB DDR3 memory
- 2 dual-port 10GB ethernet
- 2 dual-port FDR InfiniBand (storage)
Application and Development Environment

• Edison programming environment is remarkably similar to that of Hopper

• Supports production software applications, libraries, and tools needed by the entire NERSC workload

• Enables effective application performance at scale, single node (high-throughput computing), and everything in between
Quick Status

• Edison Phase I is up and running
• Added first early users on Feb. 4
  – We’ll add you now if you’re not on yet. Just ask.
• System has been extremely stable
  – Batch system software issues persist (not Cray software)
• Early tests: about 2X Hopper performance per core
• Compute
  – Meeting or exceeding expectations
• I/O
  – Not completely there yet (mostly reads and metadata)
Open Issues

• **Intel programming environment is still being configured.**
  – LibSci is not available and users need to use MKL to access LAPCK, SCALAPACK routines.

• **Batch queue issue (Moab)**
  – Some completed jobs stay in the queue long time

Please let us know if you find anything is missing in the Intel programming environment