Outline

• System architecture
  • Gemini network
  • Compute resources
  • Storage
• Programming environment
• Support
Blue Waters System

Total Compute Nodes (XE + XK): 25,712
Aggregate Memory: 1.5 PB

SCUBA Subsystem - Storage Configuration for User Best Access
National Petascale Computing Facility

• Modern Data Center
  • 90,000+ ft² total
  • 30,000 ft² 6 foot raised floor
  • 20,000 ft² machine room gallery with no obstructions or structural support elements

• Energy Efficiency
  • LEED certified Gold
  • Power Utilization Efficiency, PUE = 1.1–1.2
  • 24 MW current capacity – expandable
  • Highly instrumented
Cray XE6/XK7 - 276 Cabinets

Cray XE6 Compute Nodes
- 5,688 Blades – 22,640 Nodes – 362,240 Cores

Cray XK7 GPU Nodes
- 768 Blades – 3,072 Nodes
- 24,576 Cores – 3,072 GPUs

Note: HSN not needed for access to login nodes and storage
High Speed Network: Gemini Interconnect

Topology: 3D Torus
Current size: 23 x 24 x 24
Will soon be: 24 x 24 x 24
Nodes per gemini: 2
Peak inj. bandwidth: 9.6 GB/s
Gemini Torus

VMD image
gray – xe nodes
red – xk nodes
blue – service nodes
**XE6 Node**

Blue Waters contains 22,640 XE6 compute nodes

<table>
<thead>
<tr>
<th>Node Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Core Modules*</td>
</tr>
<tr>
<td>Peak Performance</td>
</tr>
<tr>
<td>Memory Size</td>
</tr>
<tr>
<td>Memory Bandwidth (Peak)</td>
</tr>
<tr>
<td>Interconnect Injection Bandwidth (Peak)</td>
</tr>
</tbody>
</table>

*Each core module includes 1 256-bit wide FP unit and 2 integer units. This is often advertised as 2 cores, leading to a 32 core node.
### XK7 Node

Blue Waters currently contains 3,072 NVIDIA Kepler (GK110) GPUs. Another 1,152 will be added soon.

### XK7 Compute Node Characteristics

<table>
<thead>
<tr>
<th>Feature</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Host Processor</td>
<td>AMD Series 6200 (Interlagos)</td>
</tr>
<tr>
<td>Host Processor Performance</td>
<td>156.8 Gflops/sec</td>
</tr>
<tr>
<td>Kepler Peak (DP floating point)</td>
<td>&gt; 1.3 Tflops/sec</td>
</tr>
<tr>
<td>Host Memory</td>
<td>32GB 51 GB/sec</td>
</tr>
<tr>
<td>Kepler Memory</td>
<td>6GB GDDR5 capacity &gt; 180 GB/sec</td>
</tr>
</tbody>
</table>
Storage – Lustre Disk

- Cray Sonexion with Lustre for all filesystems
- All filesystems visible from compute nodes
  - Can run application from any filesystem, but scratch highly preferred, especially for heavy I/O
- 30-day scratch purge policy
- Users may submit quota increase requests
- /home and /projects backed up daily – saved for 30 days

<table>
<thead>
<tr>
<th>Filesystem</th>
<th>Total Usable Space</th>
<th>Quota</th>
<th>OSTs (Object Storage Target)</th>
</tr>
</thead>
<tbody>
<tr>
<td>/home</td>
<td>2.2 PB</td>
<td>1 TB user</td>
<td>144</td>
</tr>
<tr>
<td>/projects</td>
<td>2.2 PB</td>
<td>5 TB group</td>
<td>144</td>
</tr>
<tr>
<td>/scratch</td>
<td>22 PB</td>
<td>500 TB group</td>
<td>1440</td>
</tr>
</tbody>
</table>
Storage – Nearline HPSS

- Spectra Logic T-Finity
  - Dual-arm robotic tape libraries
  - High availability and reliability, with built-in redundancy
- Blue Waters Archive
  - Capacity: 380 PBs \((raw)\), 300 PBs \((usable)\)
  - Disk cache: 1.2 PB
  - Bandwidth: 100 GB/sec \((sustained)\)
  - RAIT – Redundant Arrays of Independent Tapes for increased reliability
- Same /home and /projects directory structure as Lustre
- 5 TB user quota, 50 TB group quota
**GridFTP / Globus Online (GO)**

- GridFTP client on IE and HPSS nodes
  - Must be used to access HPSS
- GO interface
  - Blue Waters Portal ([https://go-bluewaters.ncsa.illinois.edu](https://go-bluewaters.ncsa.illinois.edu))
  - `globus-url-copy` command-line
  - Create your own endpoint with Globus Connect
- GO also recommended for transferring large files between Lustre filesystems
Programming Environment

Languages
- Fortran
- C
- C++
- Python
- UPC

Compilers
- Cray Compiling Environment (CCE)
- GNU

Programming Models
- Distributed Memory (Cray MPT)
  - MPI
  - SHMEM
- Shared Memory
  - OpenMP 3.0
- PGAS & Global View
  - UPC (CCE)
  - CAF (CCE)

IO Libraries
- NetCDF
- HDF5
- ADIOS

Tools
- Environment setup
- Modules
- Debugging Support Tools
  - Fast Track Debugger (CCE w/ DDT)
  - Abnormal Termination Processing

Resource Manager

Optimized Scientific Libraries
- LAPACK
- ScaLAPACK
- BLAS (libgoto)
- Iterative Refinement Toolkit
- Cray Adaptive FFTs (CRAFFT)
- FFTW
- Cray PETSc (with CASK)
- Cray Trilinos (with CASK)

Debuggers
- Allinea DDT
- Igdb
- Prog. Env.
- PAPI
- PerfSuite
- Eclipse
- Traditional

Performance Analysis
- Cray Performance Monitoring and Analysis Tool
- NCSA supported
- Cray added value to 3rd party

Cray Comparative Debugger

Visualisation
- VisIt
- Paraview
- YT

HPSS

3rd party packaging
- NCSA supported
- Licensed ISV SW

Cray Linux Environment (CLE)/SUSE Linux
Blue Waters Support

• Documentation
  • BW Portal (https://bluewaters.ncsa.illinois.edu/)
    • Documentation => User Guide

• System status
  • Portal
  • MOTD (Message Of The Day)
  • Broadcast e-mails from admins

• Help – SEAS team
  • Phone, chat, e-mail
    • JIRA
Blue Waters Support (continued)

- SEAS team (Science and Engineering Applications Support)
  - Phone*: (217) 244-6689
  - Chat (portal)*: Your Blue Waters => Live Chat
  - JIRA ticket system
    - Portal: Your Blue Waters => Your Tickets
    - E-mail: help+bw@ncsa.illinois.edu
- Multiple support levels
  - Basic (logging in) to advanced (software debugging and optimization)

* Manned M-F 9am – 5pm Central Time