

BLUE WATERS

SUSTAINED PETASCALE COMPUTING

May 22, 2013

Introduction to Blue Waters

Ryan Mokos



GREAT LAKES CONSORTIUM
FOR PETASCALE COMPUTATION

CRAY®

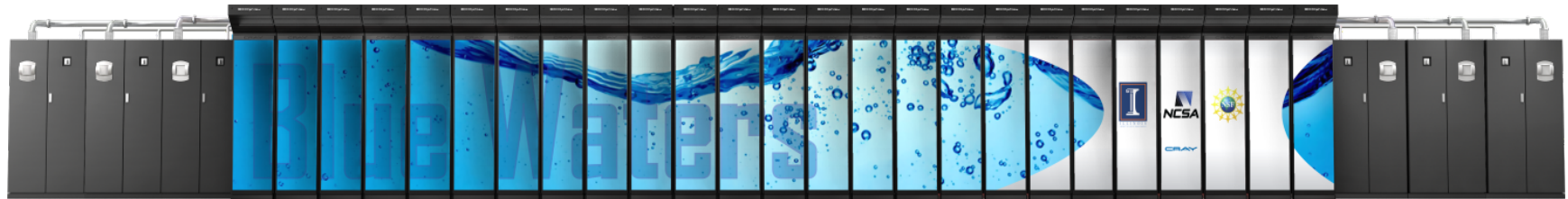
Outline

- 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

120+ Gb/sec

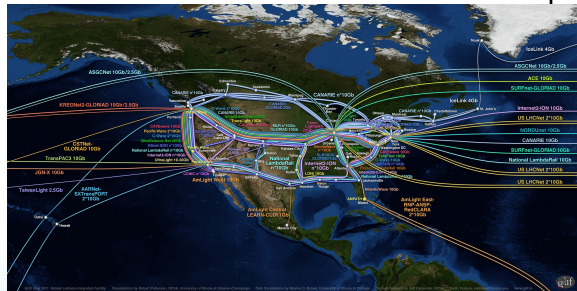
10/40/100 Gb
Ethernet Switch

External Servers

IB Switch

>1 TB/sec

100 GB/sec



100-300 Gbps WAN



Spectra Logic: 300 usable PB



Sonexion: 26 usable PB

National Petascale Computing Facility



- Modern Data Center

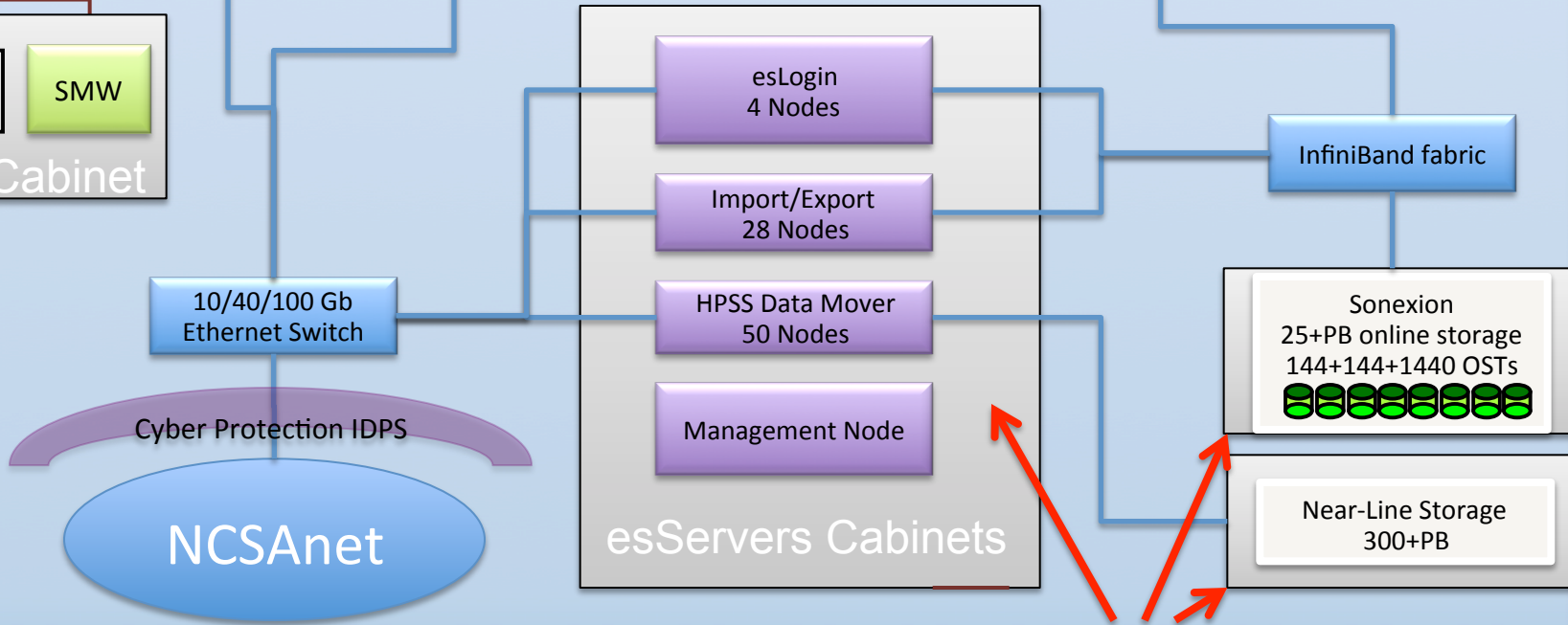
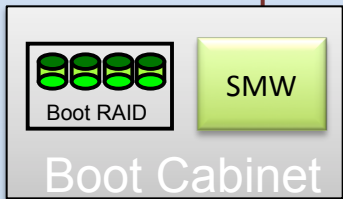
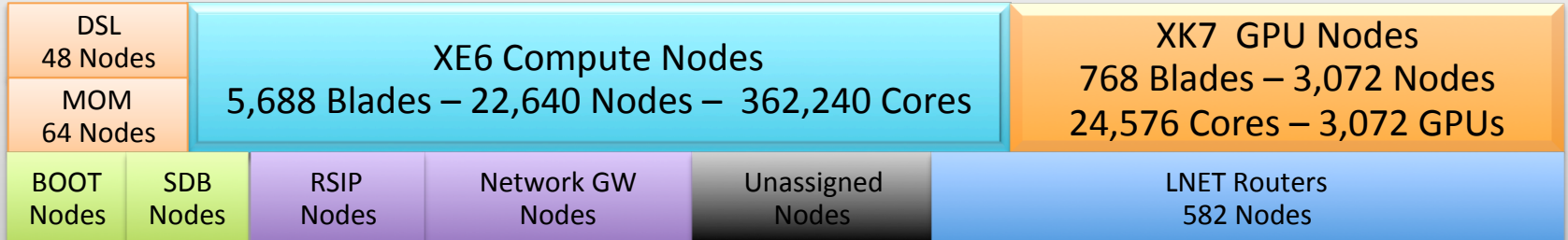
- 90,000+ ft² total
- 30,000 ft² 6 foot raised floor
- 20,000 ft² machine room gallery with no obstructions or structural support elements

- Energy Efficiency

- LEED certified Gold
- Power Utilization Efficiency, PUE = 1.1–1.2
- 24 MW current capacity – expandable
- Highly instrumented

Gemini Fabric (HSN)

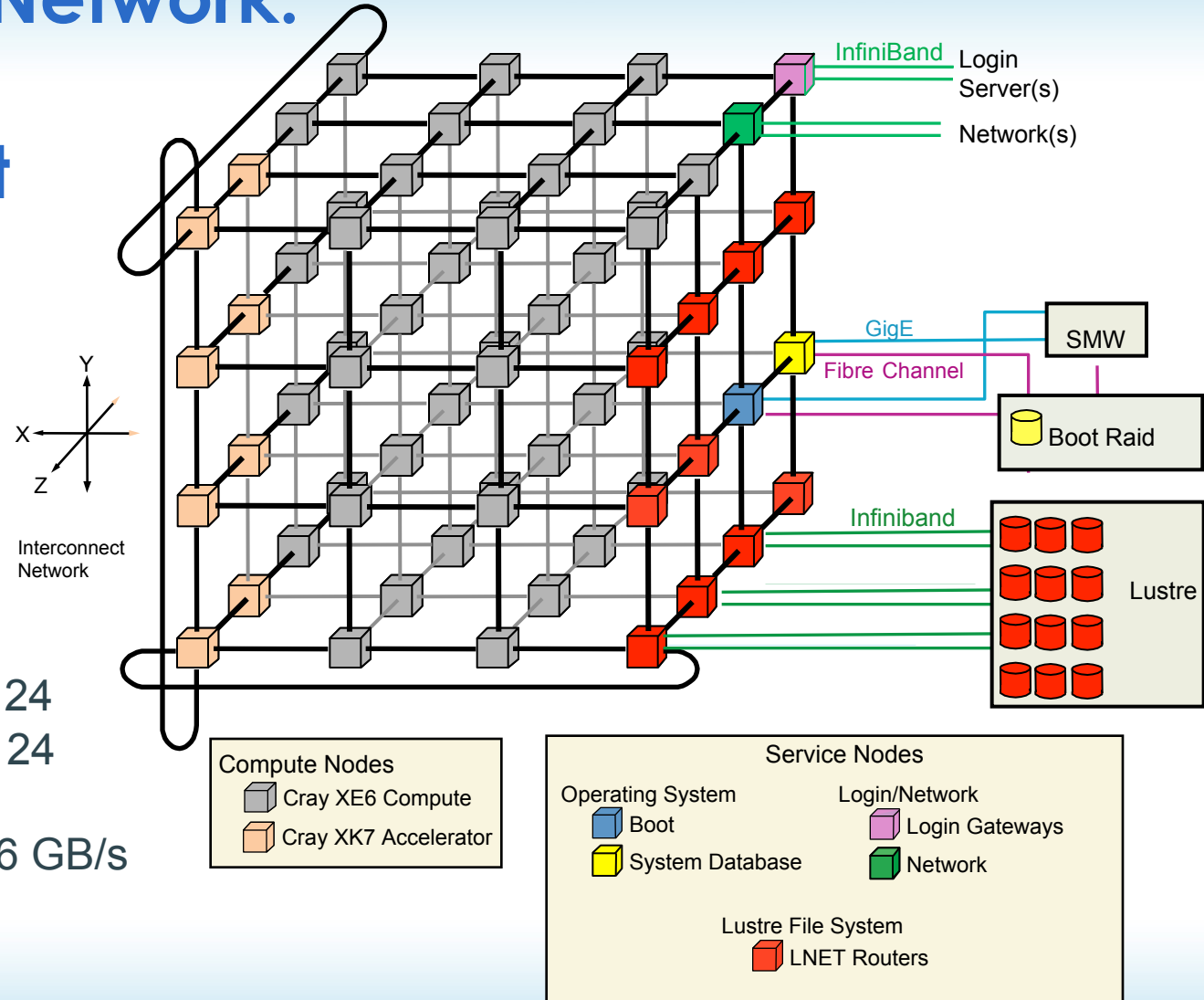
Cray XE6/XK7 - 276 Cabinets



NPCF

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

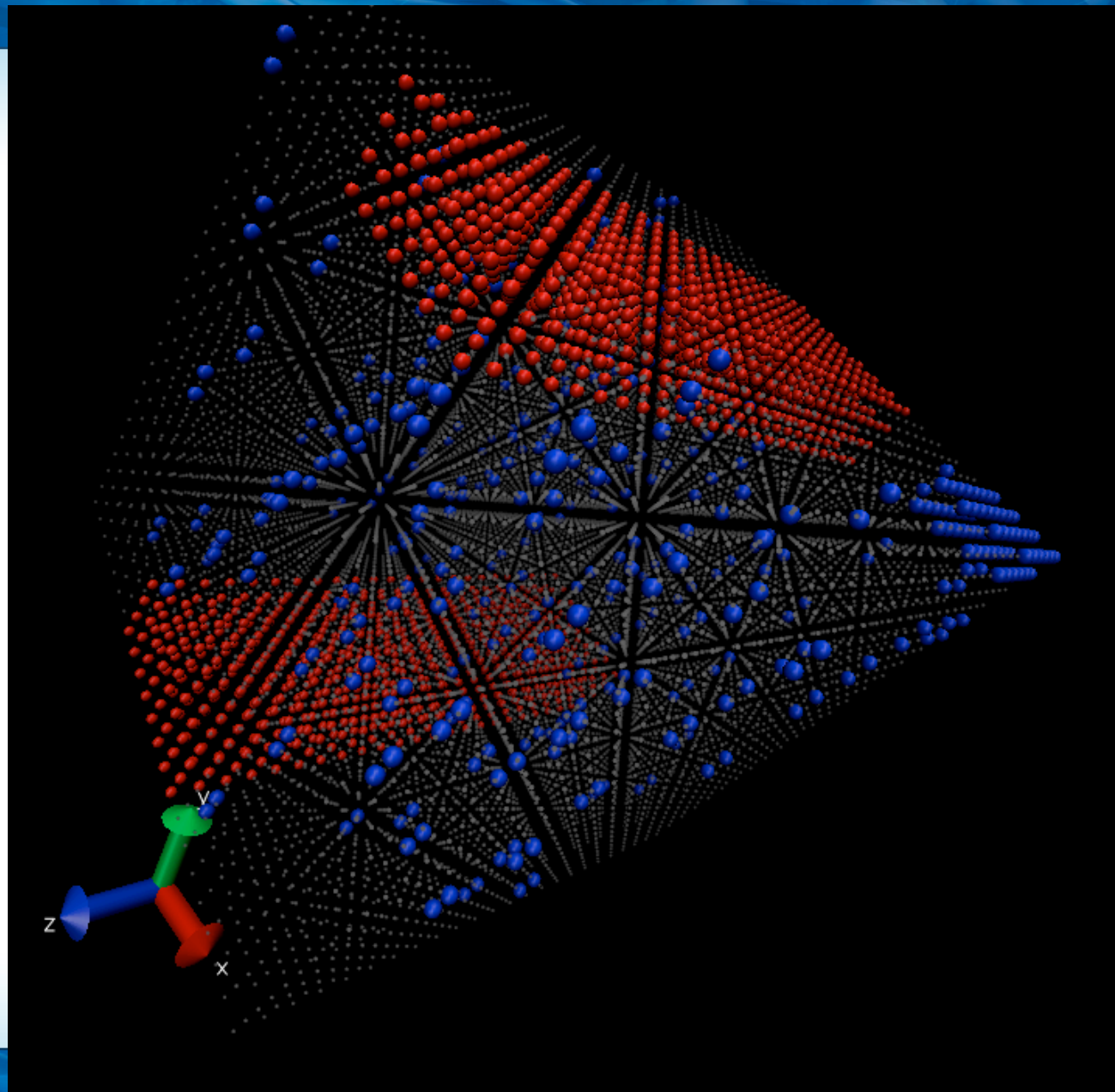
Gemini Torus

VMD image

gray – xe nodes

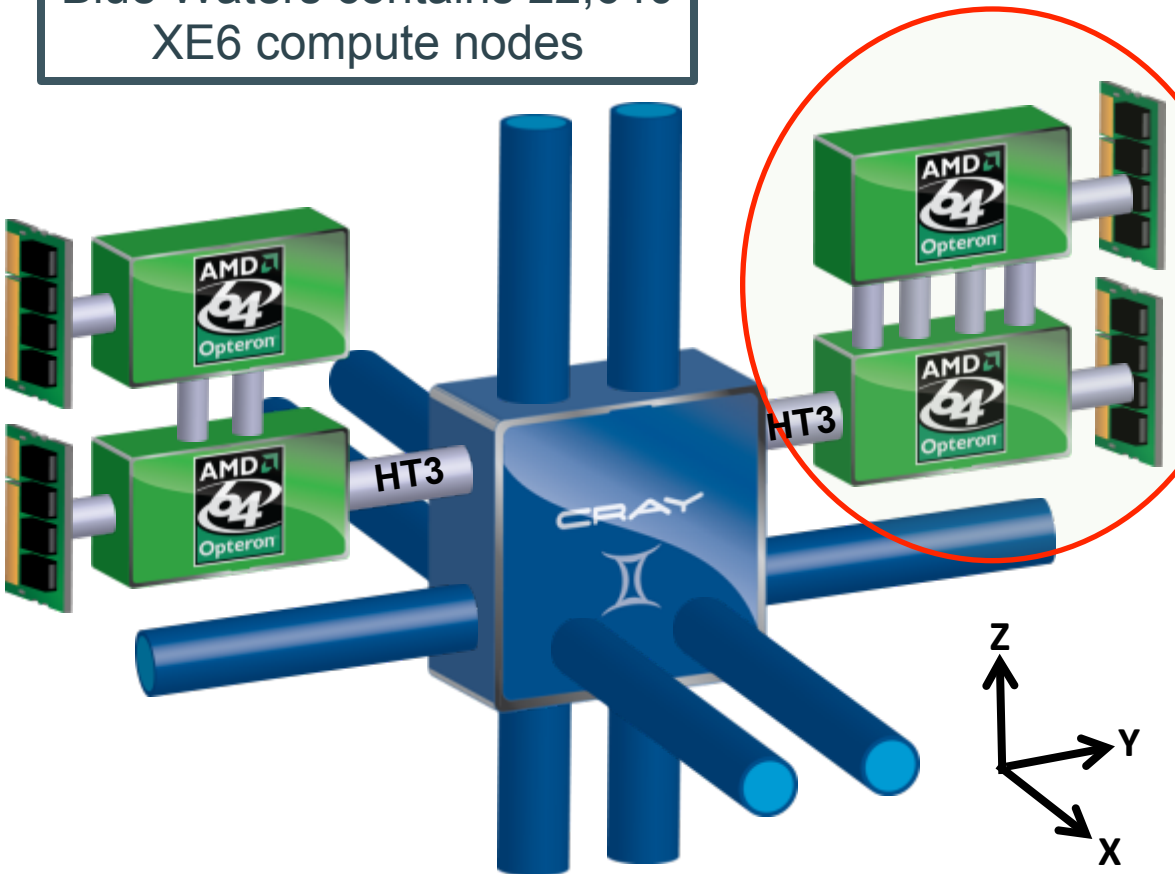
red – xk nodes

blue – service nodes



XE6 Node

Blue Waters contains 22,640 XE6 compute nodes



Node Characteristics

Number of Core Modules*	16
Peak Performance	313 Gflops/sec
Memory Size	64 GB per node
Memory Bandwidth (Peak)	102 GB/sec
Interconnect Injection Bandwidth (Peak)	9.6 GB/sec per direction

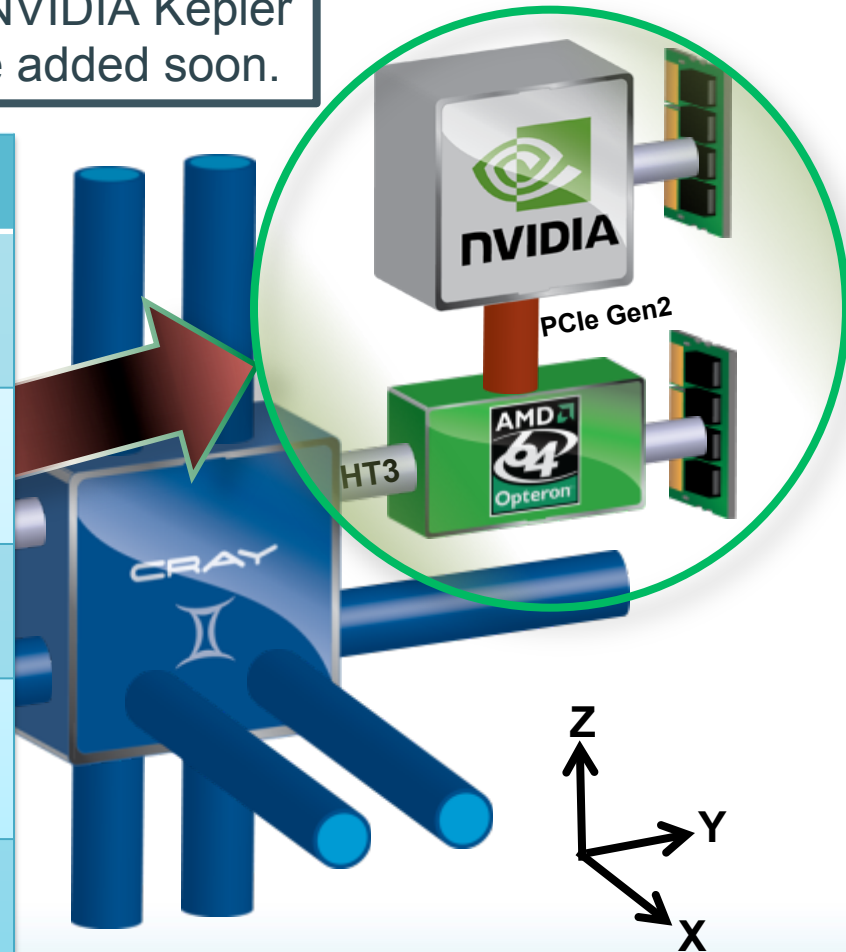
**Each core module includes 1 256-bit wide FP unit and 2 integer units. This is often advertised as 2 cores, leading to a 32 core node.*

XK7 Node

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

XK7 Compute Node Characteristics

Host Processor	AMD Series 6200 (Interlagos)
Host Processor Performance	156.8 Gflops/sec
Kepler Peak (DP floating point)	> 1.3 Tflops/sec
Host Memory	32GB 51 GB/sec
Kepler Memory	6GB GDDR5 capacity > 180 GB/sec



Storage – Lustre Disk

Filesystem	Total Usable Space	Quota	OSTs (Object Storage Target)
/home	2.2 PB	1 TB user	144
/projects	2.2 PB	5 TB group	144
/scratch	22 PB	500 TB group	1440

- 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

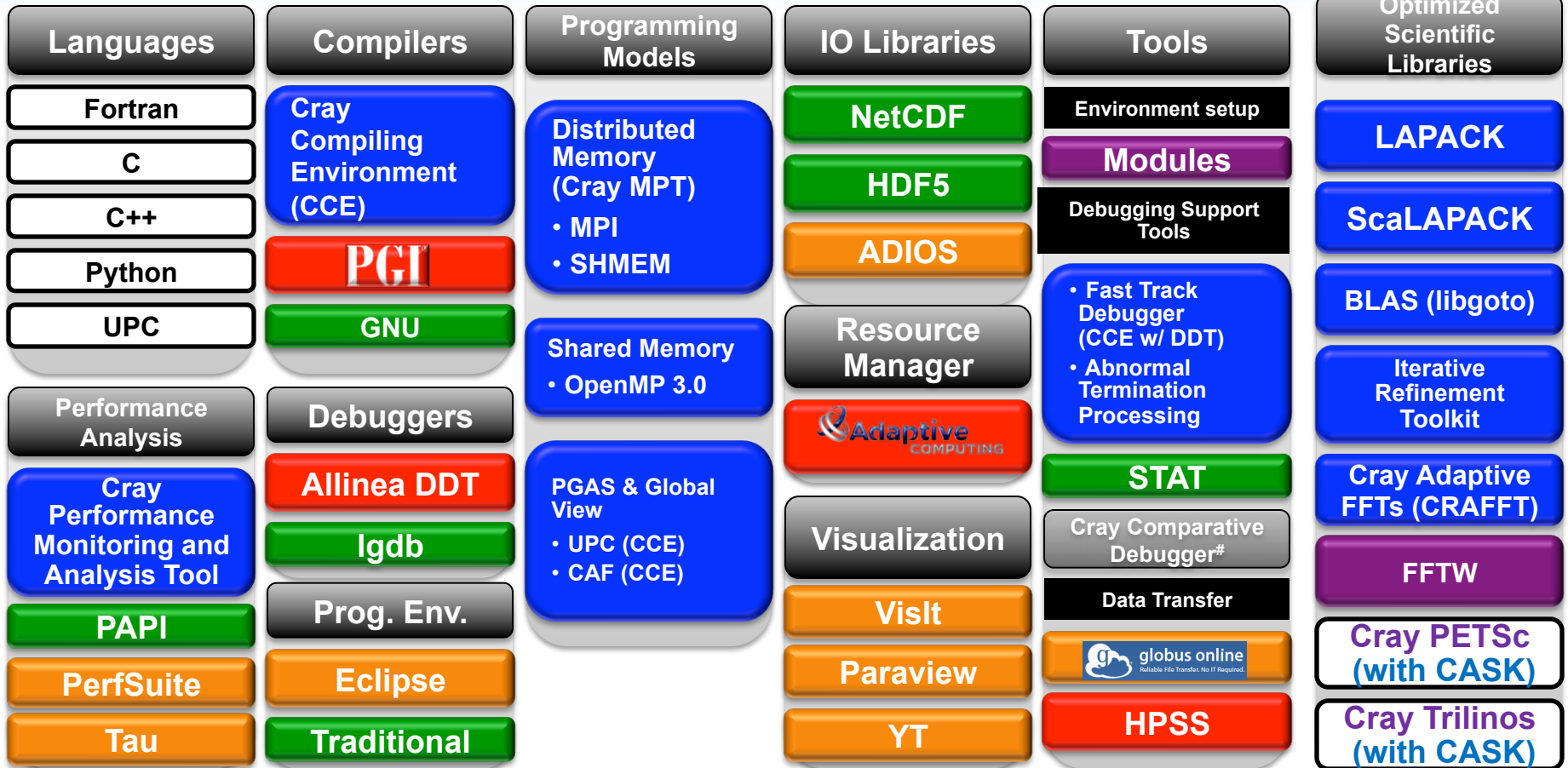
Storage – Nearline HPSS

- 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>)
 - globus-url-copy command-line
 - Create your own endpoint with Globus Connect
- GO also recommended for transferring large files between Lustre filesystems

Programming Environment



Cray Linux Environment (CLE)/SUSE Linux

Cray developed
Under development
Licensed ISV SW

3rd party packaging
NCSA supported
Cray added value to 3rd party

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