

Cray XK6 Overview

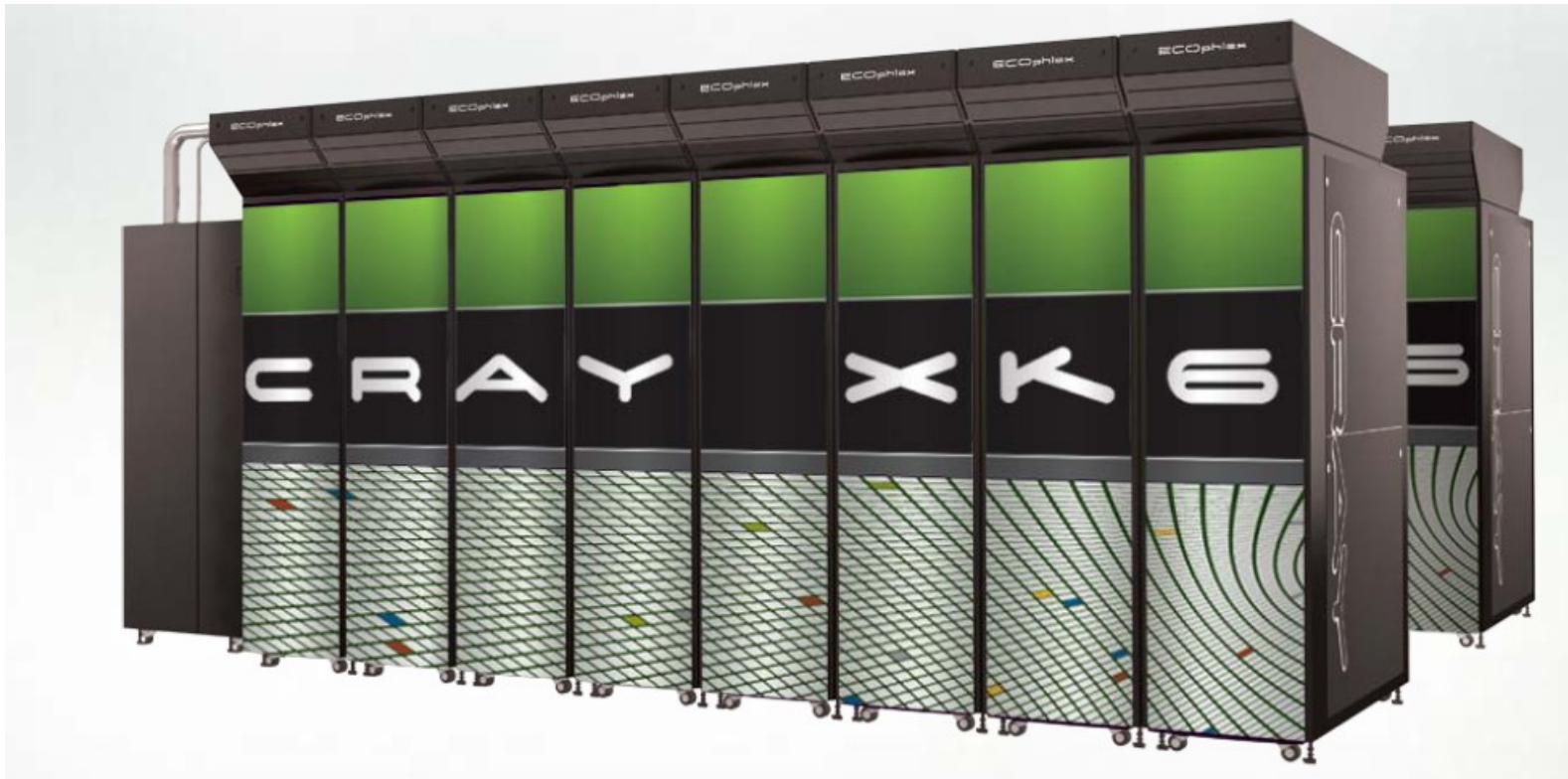
Basel, 27 October 2011

National Supercomputing Service @CSCS

System Engineer

Nicola Bianchi

The beast ... and ...



... back to the reality



TODI: the CSCS XK6

- 2 cabinets (racks)
- 6 cages
- 192 nodes
 - 16 service nodes
 - 176 compute nodes
- Air Cooled
- Lustre scratch FS



The machine's numbers

- 2'816 Cores (176 CPU AMD Interlagos)
- 5'632 GB DDR3 RAM (32GB per node)
- 176 GPUs (NVIDIA x2090)
- 1'056 GB GDDR5 GPU Memory (6GB per node)
- 130 TF peak performance

XIO Blades

- Service Nodes
 - Login nodes
 - DVS
 - Lustre routers
 - SLURM
- 1 CPU AMD 6 cores
- 16GB DDR2 RAM
- 1 PCIe BUS

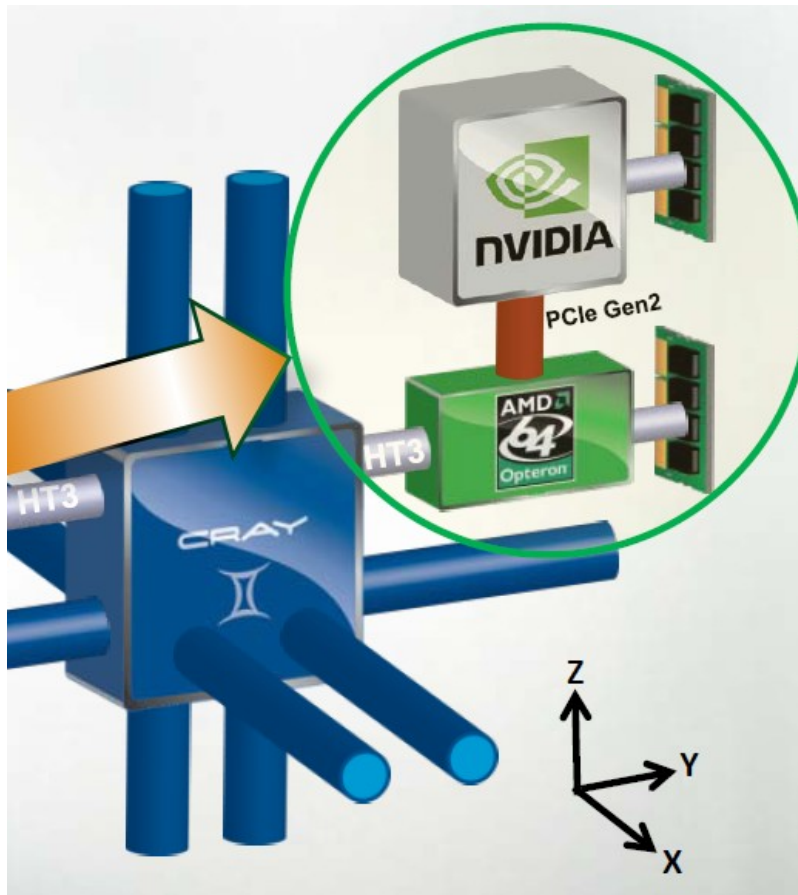
XK6 Compute blade



XK6 Compute nodes

- 1 CPU AMD Opteron 6200 (Interlagos, socket G34) 16 cores
- 1 GPU NVIDIA Tesla X2090 (Fermi architecture)
- 32GB DDR3 RAM (1600MHz) (latency <100 nanoseconds)
- 6GB GDDR5 GPU memory

GEMINI Interconnect



- 3D TORUS topology

Gemini Interconnect

- The AMD CPU is connected to the GEMINI through HT3 (Hyper Transport 3), no PCIe bottleneck
- 15'000'000 of MPI messages per second
- internode latency about 1.3 microseconds
- 10.4 GB/s bandwidth
- Adaptive routing
- Scales to over 100,000 endpoints

Hands on

- Xtcli status s0 -t accell
- Xtnodestat
- Xtprocadmin
- Module avail



xtnodestat

```
File Edit View Search Terminal Help
[nbianchi @ tod11]-[09:54:02]-[~]:-) xtnodestat
Current Allocation Status at Thu Oct 27 09:54:07 2011
```

```

      C0-0      C1-0
n3  ;;;;;;;;;; ;;;;;;;;;;S;
n2  ;;;;;;;;;; ;;;;;;;;;;S;
n1  ;;;;;;;;;; ;;;;;;;;;;S;
c2n0 ;;;;;;;;;; ;;;;;;;;;;S;
n3  a; ;;;S;; ;;;;;;;;;;
n2  a; ;;;S;; ;;;;;;;;;;
n1  a; ;;;S;; ;;;;;;;;;;
c1n0 a; ;;;S;; ;;;;;;;;;;
n3  S; ;;;S;; ;S; ;;;S;
n2  S; ;;;S;; ;S; ;;;S;
n1  S; ;;;S;; ;S; ;;;S;
c0n0 S; ;;;S;; ;S; ;;;S;
      s01234567 01234567
```

Legend:

```

nonexistent node          S  service node
; free interactive compute node - free batch compute node
A allocated interactive or ccm node ? suspect compute node
W waiting or non-running job   X  down compute node
Y down or admin down service node Z  admin down compute node
```

```
Available compute nodes:      172 interactive,      0 batch
```

Job ID	User	Size	Age	State	command line
a 2596	robinson	4	0h01m	run	namd2

```
[nbianchi @ tod11]-[09:54:07]-[~]:-)
```

xtcli status s0 -t accel

```
File Edit View Search Terminal Help
crayadm@todismw:~> xtcli status s0 -t accel
Network topology: class 0
Network type: Gemini
Nodeid: Service Core Arch| Comp state [Flags]
-----
c0-0c0s1n0a0: - OP| ready [warn]
c0-0c0s1n1a0: - OP| ready [warn]
c0-0c0s1n2a0: - OP| ready [warn]
c0-0c0s1n3a0: - OP| ready [warn]
c0-0c0s2n0a0: - OP| ready [warn]
c0-0c0s2n1a0: - OP| ready [warn]
c0-0c0s2n2a0: - OP| ready [warn]
c0-0c0s2n3a0: - OP| ready [warn]
c0-0c0s3n0a0: - OP| ready [warn]
c0-0c0s3n1a0: - OP| ready [warn]
c0-0c0s3n2a0: - OP| ready [warn]
c0-0c0s3n3a0: - OP| ready [warn]
c0-0c0s4n0a0: - OP| ready [warn]
c0-0c0s4n1a0: - OP| ready [warn]
c0-0c0s4n2a0: - OP| ready [warn]
c0-0c0s4n3a0: - OP| ready [warn]
c0-0c0s5n0a0: - OP| ready [warn]
c0-0c0s5n1a0: - OP| ready [warn]
c0-0c0s5n2a0: - OP| ready [warn]
c0-0c0s5n3a0: - OP| ready [warn]
c0-0c0s6n0a0: - OP| ready [warn]
c0-0c0s6n1a0: - OP| ready [warn]
c0-0c0s6n2a0: - OP| ready [warn]
c0-0c0s6n3a0: - OP| ready [warn]
c0-0c0s7n0a0: - OP| ready [warn]
c0-0c0s7n1a0: - OP| ready [warn]
```



Thank You.



Questions

&

Answers
