The Magic Cluster

Tuesday, September 08, 2009

Structure of the Cluster -- Hardware

- Head Node – Magic.cse.buffalo.edu
  - Hardware Profile
    - Model – Dell PowerEdge 1950
    - CPU - two Dual Core Xeon Processors (5148LV) operating at 2.33GHz
    - Memory - 16 GB Memory
    - Special Notes:
      - This is the gateway in to the magic cluster
      - This machine does not have a Tesla Co-Processor attached to it. This machine will run CUDA code in emulation mode only.
  - OS
    - Red Hat Enterprise Linux 5.3

Structure of the Cluster -- Hardware

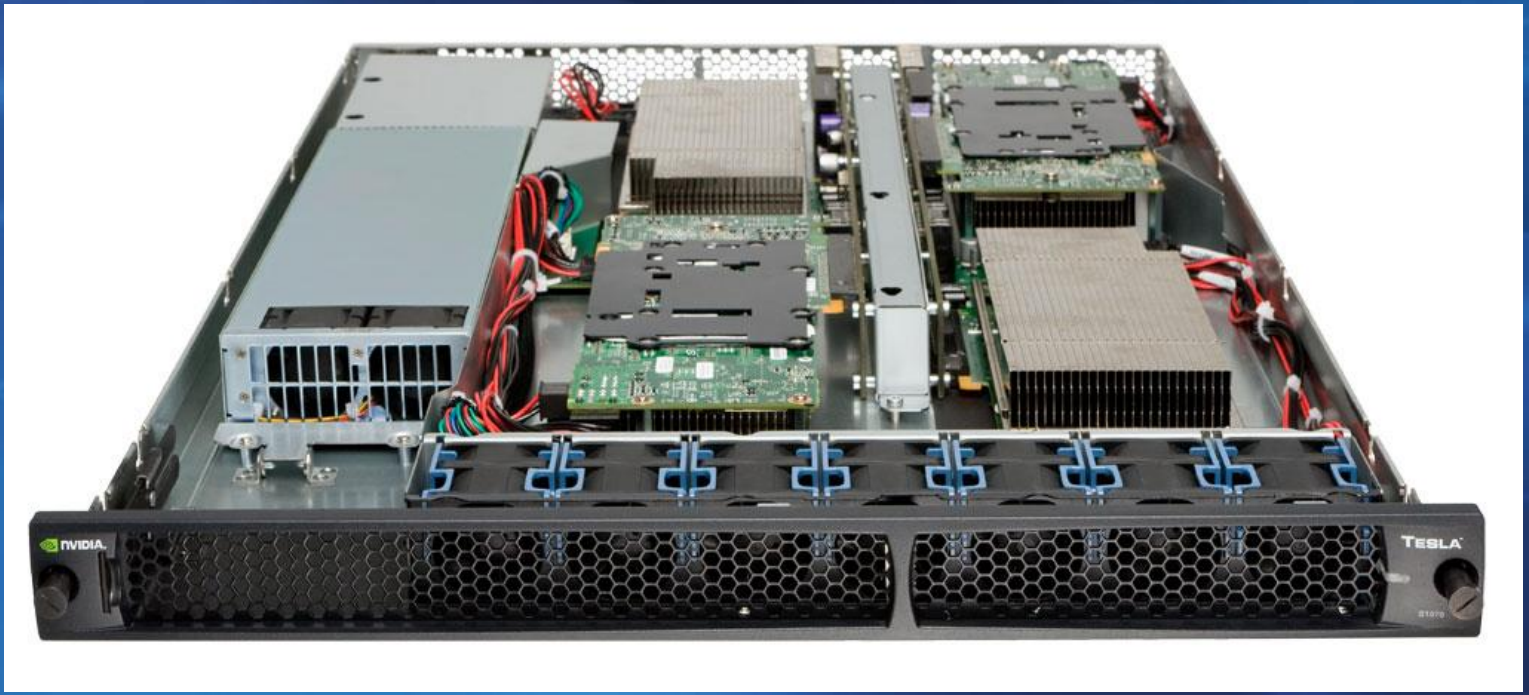
- Worker Nodes (ci-xeon-1 – ci-xeon-8)
  - Hardware Profile
    - Model - Dell PowerEdge 1950
    - CPU - Dual Core Xeon (X5260) Processor operating at 3.33GHz
    - Memory - 4 GB Memory
  - OS
    - Red Hat Enterprise Linux 5.3
  - Co-Processor
    - Nvidia Tesla 1070S

Structure of the Cluster -- Hardware

- Worker Nodes (ci-xeon-9 – ci-xeon-13)
  - Hardware Profile
    - Model - Dell PowerEdge 1950
    - CPU - two Quad Core Xeon (E5430) Processors operating at 2.66GHz
    - Memory - 8 GB Memory
  - OS
    - Red Hat En terprise Linux 5.3
  - Co-Processor
    - Nvidia Tesla 1070S

Structure of the Cluster – Tesla GPUs

- Tesla 1070S Coprocessor

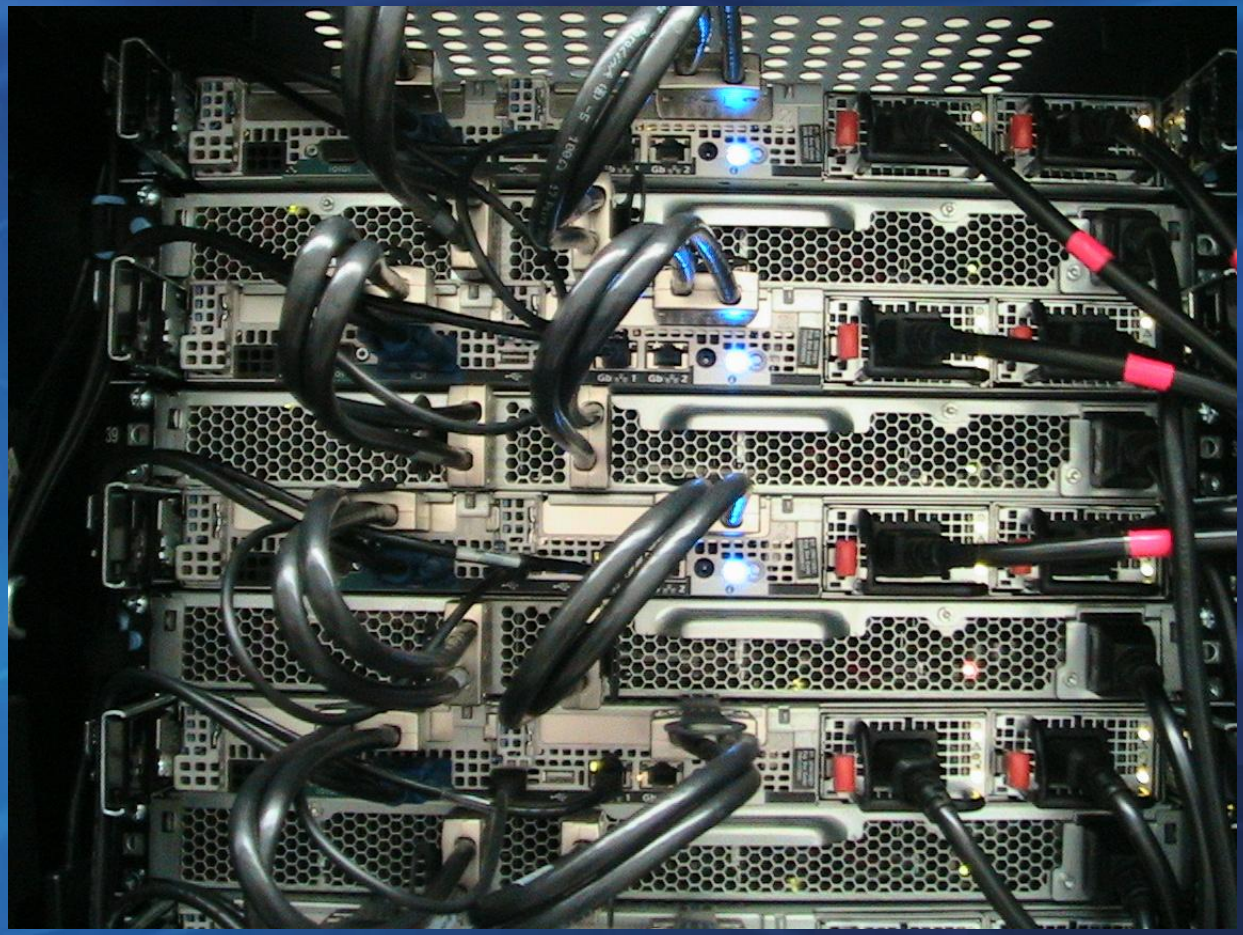


Structure of the Cluster – Tesla GPUs

# of Tesla GPUs	4
# of Streaming Processor Cores	960 (240 per processor)
Frequency of processor cores	1.296 to 1.44 GHz
Single Precision floating point performance (peak)	3.73 to 4.14 TFlops
Double Precision floating point performance (peak)	311 to 345 GFlops
Floating Point Precision	IEEE 754 single & double
Total Dedicated Memory	16 GB
Memory Interface	512-bit
Memory Bandwidth	408 GB/sec
Max Power Consumption	800 W
System Interface	PCIe x16 or x8

Structure of the Cluster – Tesla GPUs

- Each Tesla Coprocessor is connected to a compute node via 2 PCI-X cards.



Structure of the Cluster – Tesla GPUs

- Tesla 1070S Coprocessor
  - References:
    - http://www.nvidia.com/object/product_Tesla_s1070_us.html
    - http://www.transtec.co.uk/GB/E/products/personal_computer/supercomputing_pc.html#

Structure of the Cluster – Network

- NAT
  - Magic is the head node with a publicly accessible IP address.
    - All worker nodes are hidden from the world behind the firewall (magic).
    - Magic “masquerades” as a worker node when a worker node needs to talk to the internet.
    - Internet access from the worker nodes should not occur that often.

Structure of the Cluster – Network

Internet

CSE Firewall

Building Switch

Lab Switch

Magic (NAT)

Worker Nodes

Cluster Switch

Colosseum (Storage)

Structure of the Cluster – Network

- Network
  - All network connections are Gigabit Ethernet
  - There are two network switches, one is public facing, the other is private facing.
  - The only way to hop from the public to the private switch is through the NAT firewall on magic. Magic is duo-homed with a NIC on the public switch and a NIC on the private switch.
  - Colosseum also is duo-homed with a NIC on the public switch and a NIC on the private switch but does not have any NAT functionality.
  - There is a host-based firewall between the cluster and the outside world. This firewall will drop inbound connections to any machine (internal) from off campus.
    - If you wish to access the machine from off campus you can use the UBVPN. This will then allow you to access the machine.

Structure of the Cluster – Network

- Network
  - All Edge switches connect to the UB Backbone through fiber channel 1Gb/s connections. This is in the process of being upgraded 10 Gb/s connections.
  - The firewall between the Building Edge switch is currently 100Mb/s. This is due to be upgraded during the winter break.
  - UB's backbone connects to the internet through the NYSERNet ISP Across 2 router links.
  - Speed to the internet has been clock at 6.50 Mb/S

Structure of the Cluster – Storage

- The Storage Server is colosseum.cse.buffalo.edu.
- You should only log in to this machine to transfer large data sets via SCP, SFTP, or Windows File Sharing (available on campus only).
  - Programs such as filezilla or winSCP should be used to move large data sets.
- All storage space on this machine is exported to every machine in the cluster via NFS.
- File Systems
  - /shared-space is an area for software installation shared across all worker nodes.
  - /scratch is a read/write area for temporary files shared across all worker nodes. Files will be purged after 1 week.
  - /home is where user home directories reside. Quota space is 15 Gb per user.
  - /projects is an area for all Cyber Infrastructure specific storage.

Structure of the Cluster – Storage

Filesystem	Size	Used	Avail	Use%	Mounted on
colosseum:/projects	7.9T	175M	7.5T	1%	/projects
colosseum:/scratch	2.9T	405M	2.8T	1%	/scratch
magic:/home	262G	16G	233G	7%	/home
magic:/shared-space	262G	16G	233G	7%	/shared-space

Structure of the Cluster – Power

- All 13 nodes, head node, storage node, and disk shelves are on uninterruptable power supplies. These 3 power supplies will tolerate around 10 minutes of power failure before nodes shut down.
- The Tesla units are on wall power. If the power fails in Furnas hall the Tesla units will shut down!

Tools -- CUDA

- NVIDIA's CUDA development tools consist of three key components to help you get started:
  - The latest CUDA driver – installed on all machines
  - A complete CUDA toolkit
    - NVCC C compiler
      - Found at /usr/local/cuda/bin on each machine.
    - CUDA FFT and BLAS libraries for the GPU
    - Profiler
    - gdb debugger for the GPU
    - CUDA runtime driver (also available in the standard NVIDIA GPU driver)
    - CUDA programming manual
    - http://www.nvidia.com/object/Tesla_software.html
    - http://www.nvidia.com/object/cuda_learn.html

Tools -- CUDA

- NVIDIA's CUDA development tools consist of three key components to help you get started:
  - Developer SDK
    - examples with source code
    - Built-in to /shared-space/CUDA-SDK on all machines

Tools – Languages and Compiles

- Fortran
- C++
- Java

Accessing the Cluster

- SSH to magic.cse.buffalo.edu
  - Windows -- use the Putty program available on the UBIT website.
    - http://ubit.buffalo.edu/software/win/putty.php
  - Mac/Linux – use your built-in SSH client
    - ssh Username@magic.cse.buffalo.edu
  - If you are off campus be sure to be using the UBVPN client
    - http://ubit.buffalo.edu/software/

Accessing the Cluster

- If you need a graphical connection
  - Windows -- use the xwin32 program available on the UBIT website.
    - http://ubit.buffalo.edu/software/win/XWin32/
  - Mac/Linux – use your built-in SSH client
    - ssh –X Username@magic.cse.buffalo.edu
  - Note: Be careful what you do graphically! Encrypting a graphics tunnel is expensive to the network bandwidth. Consider if you really need to run something graphically.

Accessing the Cluster

- The Queuing System
  - Nodes should be access using the batch system, torque.
    - The advantage of the queuing system is that you will be placed on a node with the greatest amount of available resources. At any given time users from the grid, or other researchers could be running jobs.
    - An interactive session can be started by entering “qsub –I”.
    - Use the following reference for instructions on interacting with the queuing system.
      - http://www.clusterresources.com/products/torque/docs/ in particular Section 2.0, Submitting and Managing Jobs.

Accessing the Cluster

- You MUST properly configure your ssh keys before using Torque. To do so execute:
  - ssh-kegen -t rsa (select all defaults, including blank password)
  - cp ~/.ssh/id_rsa.pub ~/.ssh/authorized_keys
  - Copy the /local/torque/dot-rhosts to ~/.rhosts

Accessing the Cluster

- Accounts
  - Accounts will be created based on your UBIT name and your CSE password.
  - Accounts on this cluster are managed separately from UBIT and CSE machines. Onetime password synchronization will occur between your CSE account and your account on magic. From then on, password changes from UBIT or CSE will propagate to magic.
  - You username and password are the same on every node of the cluster.

Contact Information

- kpcleary@cse.buffalo.edu
  - For quick questions

OR

- Cse-consult@cse.buffalo.edu