McGill Cluster Facilities

Andreas Warburton
McGill

CDF IPP-Canada Workshop
University of Toronto
15 April 2003

Outline

- Brief Description of McGill Beowulf clusters
  - Astrophysics Pulsar Beowulf “Borg”
  - Condensed Matter Physics CFI Beowulf
- Status of McGill’s new Computation Facility
  - AC
  - Power
  - Cluster installations
- Status and Timetable
- Network Situation

Astrophysics Pulsar Beowulf “Borg”

- Owners/Operators: Vicky Kaspi’s pulsar astrophysics group
- ~2 years old
- 104 Athlon processors in 52 nodes
- ~1.6 TByte on two file servers
  - Plan is to get further ~2 TByte of TornadoRAID F4 disk
- Until recently, had a 10 Mbit/s connection to McGill backbone
- Pulsar data sets: ~100 @ ~100 GByte each
- Data get shipped from Arecibo to McGill on DLT
- Duty cycle for astro applications is high
- CDF MC prospects: fair
  - CPU/disk resources are currently somewhat limited
  - Short term: possible use for development and testing
  - Longer term: possible CPU available (if CDF can help with disk resources)

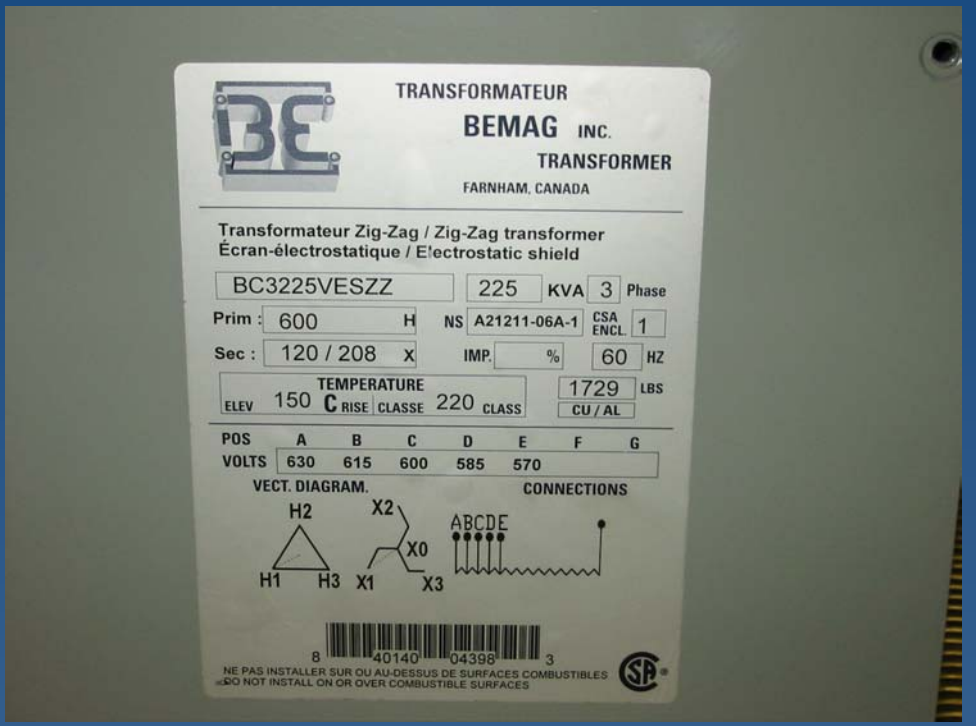
Condensed Matter Physics (Nanotools) CFI Beowulf

- Principals: Martin Grant and Hong Guo (CMP theorists)
- Currently undergoing construction
- 700 Athlon processors (each 2 GHz) in 350 nodes
- (For now) largest such cluster in Canada?
- 40 Gbyte disk per node
- 2 Gbyte RAM per node
- CMP cluster design not optimized for HEP Monte Carlo applications:
  - CMP theoretical applications are mostly CPU intensive
  - Large file servers not required for staging
  - Network bandwidth not especially critical

Air Handlers

- Expect 50 ton heat load with everything running
- 3 AC units x 20 tons/unit = 60 tons of Air Conditioning
- Heat Exchangers located on roof of Rutherford Physics Building
- Air-flow volume: 40,000 CFM (room atmosphere replaced 3x per minute)

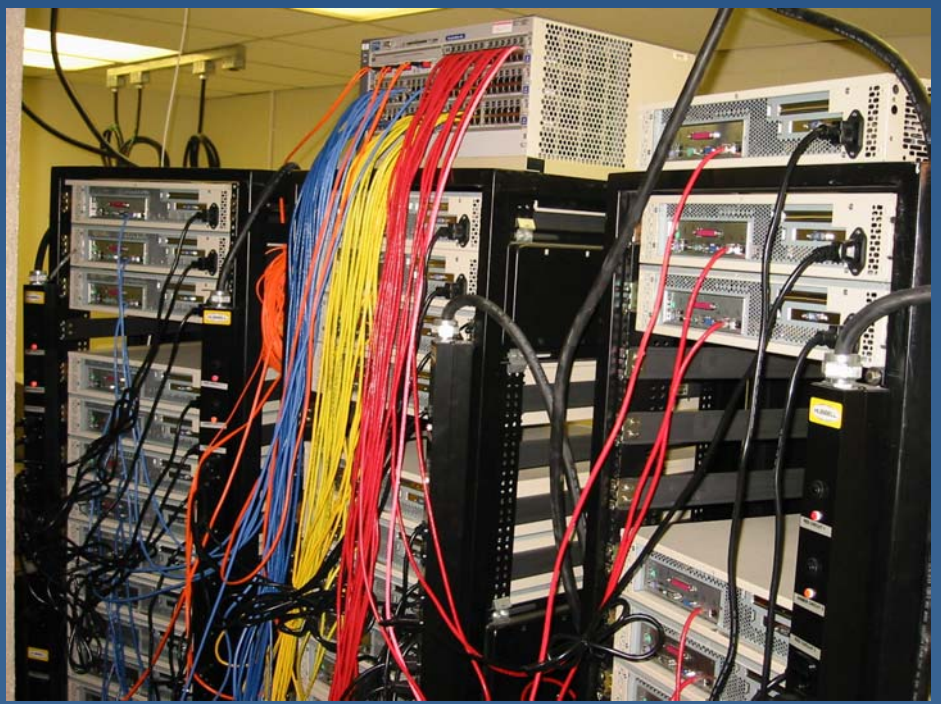
Power Transformation and Delivery



- Room supply: 600 V, 200 A, 3-phase
- Transform down to 208/120 Vac
- Maximum power draw expected: 180 kVA
- Transformer rating: 225 kVA @ >95% efficiency
- Transformer dissipation: ~3 kW
- Specially configured to handle switched-mode PSs in computers
- Separate grounding from regular mains

Pulsar Beowulf Status

- Installed in new room
- Up and running



Nanotools Beowulf: Partially Constructed

- working nodes have Debian (Woody) installed
- switches need configuration/synchronization
- remaining nodes need installation
- burn-in testing yet to be done
- benchmarking yet to be done

Status of Nanotools Beowulf

- Still undergoing construction (<1 week remains)
- 14 out of 350 found to be dead (bad batch)
- 10 Black Diamond switches need synchronization (technician scheduled for 04/22)
- Original benchmarking deadline of 04/15 missed
- New benchmarking target/deadline: end of April

Some Notes on Networking

- New: the Pulsar and Nanotools Beowulf clusters now will be connected to the McGill backbone by 100 Mbit/s connections.
- McGill Backbone: a series of 2-10 Gbit/s links (new, expensive, heavily loaded, and “badly designed”).
- Link to outside world: 1 Gbit/s link to RISQ/CA*net/Gigapop (Quebec research network), followed by 1 Gbit/s link to CA*net “East-West” system, followed by “Internet2”.
- Possibility exists to rent 1 Gbit/s switch to McGill backbone.
- Real question: Is backbone bandwidth adequate? Needs assessment.

MCGILL-CDF1 --> fcdfsgi2.fnal.gov (15 hops)

```
chavin.physics.mcgill.ca
james-rsm.gw.mcgill.ca
james-core-msfc.GW.McGill.CA
burnside-core-msfc.GW.McGill.CA
132.216.248.33
c4-gw.risq.net
c4-mon01.canet4.net
c4-tor01.canet4.net
chi-gev140-canet-toranto.es.net
chicr1-ge1-chirt1.es.net
fnal-pos-chi.es.net
ge6-0-0.r-s-frw.fnal.gov
vlan300.r-s-hub-fcc.fnal.gov
vlan304.r-cdf-cas-fcc2e.fnal.gov
fcdfsgi2.fnal.gov
```

CDF Opportunities and Plans

- In principle, 20% of Nanotools cluster is available to "outside users", including CDF MC production
- How is percentage of a cluster measured? Management procedure not yet worked out:
  - Nodes or CPUs?
  - Percentage of CPU load?
  - Real time?
- Nanotools disks have been partitioned 4 ways
  - Plan: cloned FNAL/Redhat will be installed in one of the partitions
  - Will be done on ~20% of the nodes

Next Steps

- Take a quantitative assessment of McGill/FNAL network/bandwidth situation
- Explore and implement networking options
- Once Nanotools is commissioned, up, and running, set up CDF-II code
- Explore setting up CAFtools/fbsng/GRID