Duke Atlas Tier 3 Site

Doug Benjamin
(Duke University)

Duke Atlas Group

Personnel

- 4 Tenured Faculty
- 1 Research Professor, 2 Research Scientists
- 2 PostDocs (located at CERN)
- 2 PostDocs (who will eventually transition from CDF at the Tevatron to Atlas)
- Several students

Duke Tier 3 site (current HW configuration)

- 2 TB Fileserver (Centos 4.6 64 bit OS)
  - Data and user home areas exported to other nodes in cluster
- 1 condor head node (for future expansion)
- 3 interactive nodes
  - 2 dual core AMD , SL4.4 (32bit) Ram 4 GB
  - 500 GB local storage
- one node interactive node doubles as Atlas code server
- 2 nodes at CERN (SLC 4.5 (32 bit), 1 dual CPU, 2 GB RAM , 320-500 GB local storage each

Duke Tier 3 Software/Support Personnel

- Software installed
  - Several recent Atlas releases from kits (12.0.6,13.0.30 and 13.0.40)
  - And the occasionally nightly release from kits as needed for TRT code development
  - Condor (not used at moment), OSG client
- Support Personnel
  - Duke Physics department system managers maintain OS and interactive user accounts
  - Duke HEP Research scientist (me) Installs/maintains Atlas, OSG and Condor software
    - ~ 0.25 FTE available for these activities
    - Implies the code must be simple to install, debug and operate

Upgrade plans

Near term (within 6 months)

- Additional computing
  - 2-3 8 core machines, 2 GB/core
- Additional Disk -
  - up to 10 TB standalone fileservers

Longer term (minimal survival mode)

(next year or so depending on delivered Luminosity)

- a 8 core machine per group member (~10 machines)
- ~ 25 TB disk storage

Our current analysis model

Based on a small Tier 3

- Heavily depends on outside computing
- Fraction of data copied to local storage
- Test programs with local machines
- Send to grid - larger scale jobs
- Resulting histogram files returned to local machines

- Disadvantage -
  - Very reliant on availability on outside CPU and storage at the Tier1 and Tier2 sites

Upgrade Plan (ideal model)

- ~25 8 core machines
- ~ 100 TB disk space
- Analysis model
  - Copy most of DPD data locally
  - Analyze locally except for the very large jobs
  - Would allow us quickly analyze the data as it is available

Comparison w/ Duke computing @ CDF

Duke HEP has the equivalent of an Atlas Tier 3 at FNAL

- 26 cores on 10 machines (40% are desktops)
  - 14 cores are recently purchased (< 1 year) on 4 machines
- 15.4 TB Raid 5/6 on 6 file servers + local storage on compute nodes
- Make extensive use of collaboration wide CDF resources

- > 50% all CDF computing is made available to general users for their own analysis uses

  (vital since an institution Duke's size can not afford to have a large farm - our computing budget is too small)

- For Integrated Luminosity of ~ 2.4 fb^{-1}

  High Pt e's - Data: 122 M evts volume - 19 TB 158 kB/evt

  analysis ntuple: 122 M evts volume - 6 TB 52 kB/evt

  MC W(enu)/[Z(ee)] 36 [23] M evts - 6 [3.9] TB 163[167] kB/evt

  associated ntuples: 28 [20] M evts - 2.1[1.6] TB 82[52] kB/evt

Challenges

- Interaction with Duke Computing infrastructure
  - Campus Firewalls
    - open ports normally closed - slow process
  - Campus networking
    - “Last mile problem” - Duke U is on Internet 2 and Lambda Light Rail - Physics department network is not.
- Funding for Computing
  - In past 2 years - received no funds from DOE for computing
  - Any computing came from small grant the University
  - Need to find additional resources if we are to grow to the size we envision we need to be effective
  - US ATLAS should make it very clear to funding agencies the working model for tier 3 sites and available user support (CPU - storage) at Tier 2 sites

Conclusions

- Duke Tier 3 could be an example of a small tier site being designed to provide enough resources to allow group members to analyze the Atlas data effectively
- Without sufficient funding to Tier 3 sites, it implicitly assumes that sufficient resources will be made available to the general users at the Tier1 and Tier 2 sites.