Big Computing and the Mitchell Institute for Fundamental Physics and Astronomy

Texas A&M

Big Data Workshop

David Toback

January 2015

Outline

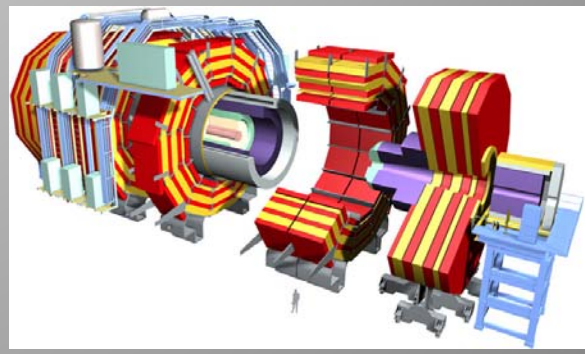
- Overview of Particle Physics and Astronomy and the Mitchell Institute
- Collider Physics Computing Needs
- Dark Matter Computing Need
- Phenomenology
- Requirements and Lessons Learned
- Status and Future

Overview of Particle Physics and Astronomy and the Mitchell Institute

- **Mitchell Institute covers both Particle Physics and Astronomy**
  - **Both theory and experiment in both, includes String Theory**
- **Many of these groups need significant computing, although in many different ways**
- **Our primary work horse has been the Brazos Cluster (brazos.tamu.edu)**
- **Big usage from 4 user/groups**
  - **CMS Experiment at the Large Hadron Collider (LHC)**
  - **Dark Matter Search using the CDMS Experiment**
  - **High Energy Phenomenology**
  - **Other (mostly CDF experiment at Fermilab, and Astronomy group)**

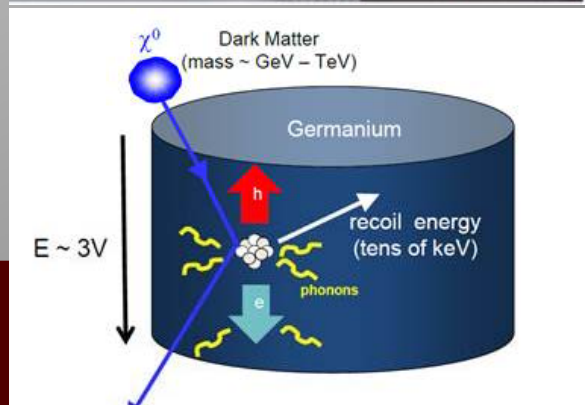
CMS Experiment at the LHC

- **Collider Physics at CERN/Fermilab have often been the big computing drivers in the world (brought us the WWW)**
- **LHC was a driving force in creating Grid computing**
- **LHC experiments have a 3-tiered, distributed computing model which requires 10's of petabytes and millions of CPU hours (T1, T2 and T3 for CMS of which A&M is a member institution)**
- **Not well suited to Supercomputing at A&M because of interesting Requirements**
  - **Ability to run on the GRID as an international collaboration**
    - **Jobs and data from around the world**
    - **Firewall issues for external users**
    - **Automated data distribution and local jobs regularly accessing remote databases**
  - **The computing needs here are high THROUGHPUT, not high PERFORMANCE**
    - **Just run LOTS of independent jobs on multi-Tb datasets**



Dark Matter Searches with the CDMS Experiment

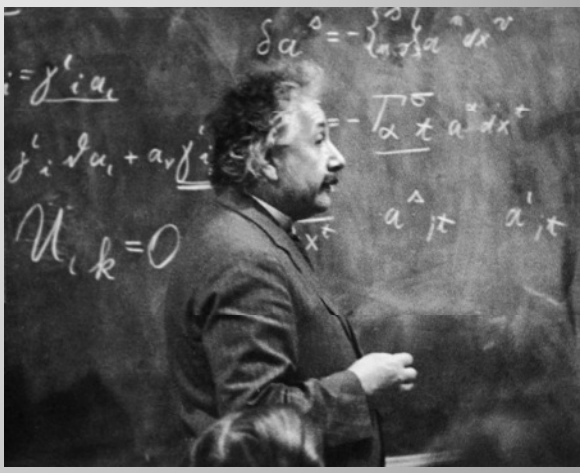
- **Much smaller experiment (~100 scientists)**
- **Smaller scale computing, but many of the same issues as we interface with other national labs**
  - **Stanford Linear Accelerator Center (SLAC)**
  - **Fermi National Accelerator Laboratory (FNAL or Fermilab)**
  - **Sudbury Neutrino Observatory (SNOLab)**

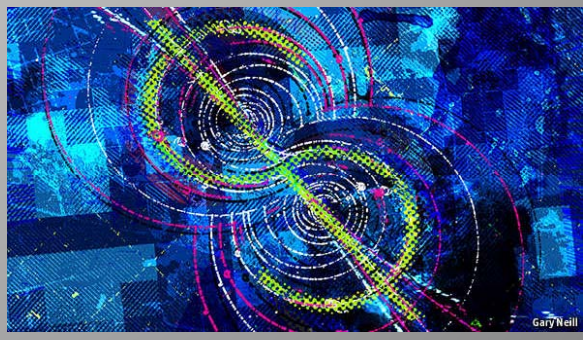


David Toback
Big Computing and the Mitchell Institute for

Particle Theory/ Phenomenology

- **Particle Phenomenologist's at MIST do calculations, but also do VERY large simulations of collider data to see what can be discovered at LHC**
- **Again, just need high throughput and lots of memory**
  - **Jobs don't need to talk to each other**





Overview of Brazos and Why we use it (and not somewhere else)

- **We bought into the *Stakeholder* model on Brazos:**
  - **309 Compute nodes/3096 cores for the cluster, Institute owns 700 cores, can run on other cores opportunistically(!)**
  - **~200 Tb of disk, Institute owns about 100Tb, can use extra space if available**
- **Can get ~1Tb/hour from Fermilab, 0.75Tb/hr from SLAC**
  - **Links to places around the world (CERN-Switzerland, DESY-Germany, CNAF-Spain, UK, FNAL-US, Pisa, CalTech, Korea, France Etc.)**
- **Big wins for us**
  - **Well suited for High Throughput (lots of parallel jobs looking at millions of separate collisions to see if they look like a new particle was produced)**
  - **Accepts jobs from Open Science Grid (OSG)**
  - **Excellent support from Admins – Almes (PI), Dockendorf and Johnson**
  - **Fun use case: Typically bring big data copies from CERN, run on it to make local processed data, then delete the local copy of the raw data**

More detail on how we run at: http://collider.physics.tamu.edu/mitchcomp

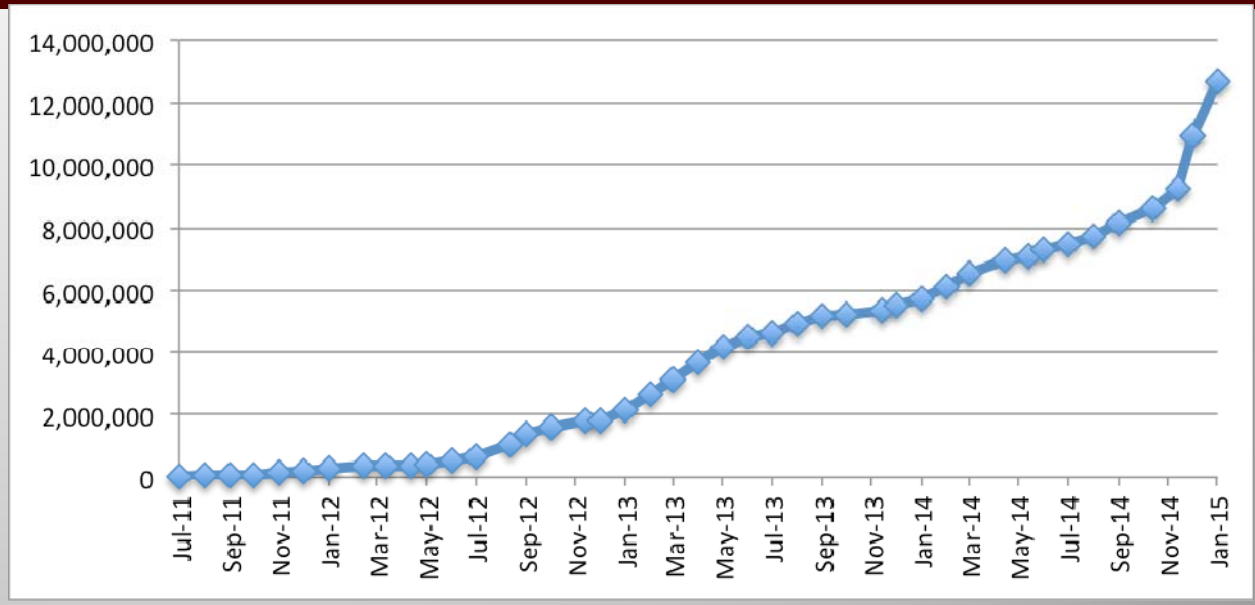
Online Monitoring

- Constantly interrogate the system
  - Disks up? Jobs running? Small data transfers working?
- Run short dummy-jobs for various test cases
  - Both run local jobs as well as accept automated jobs from outside
- Automated alarms for the “first line of defense” team (not Admins), as well as the Admin team
  - Send email as well as make the monitoring page Red

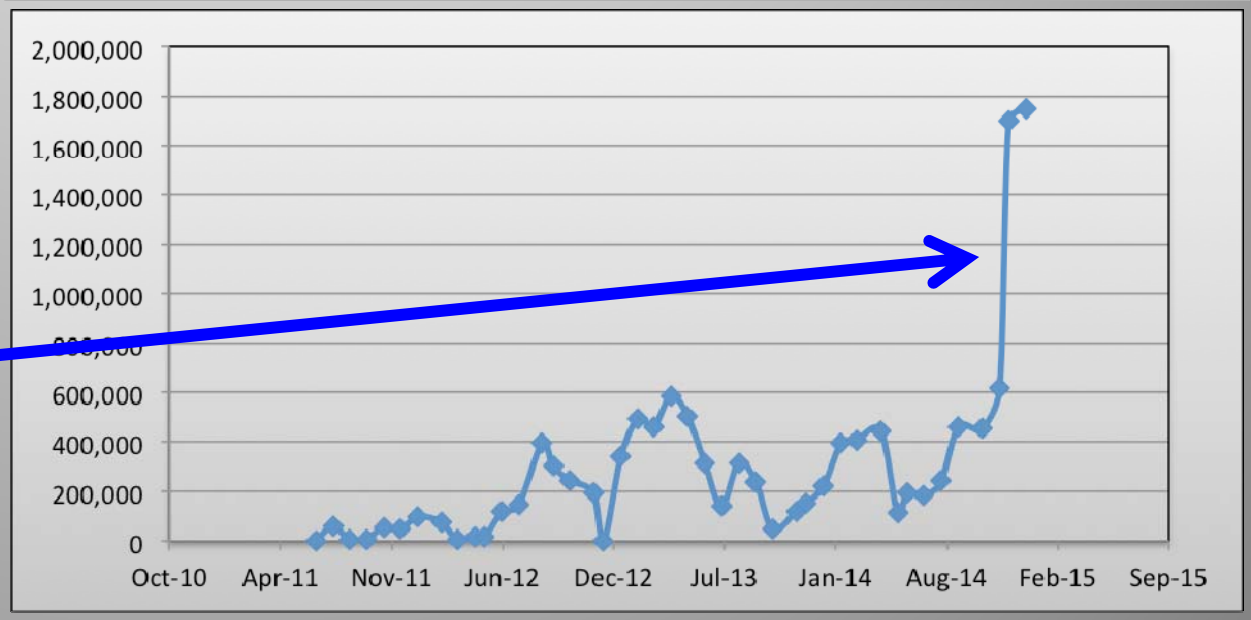
More detail about our monitoring at http://hepx.brazos.tamu.edu/all.html

Fun Plots about how well we've done

- **Cumulative CPU Hours**



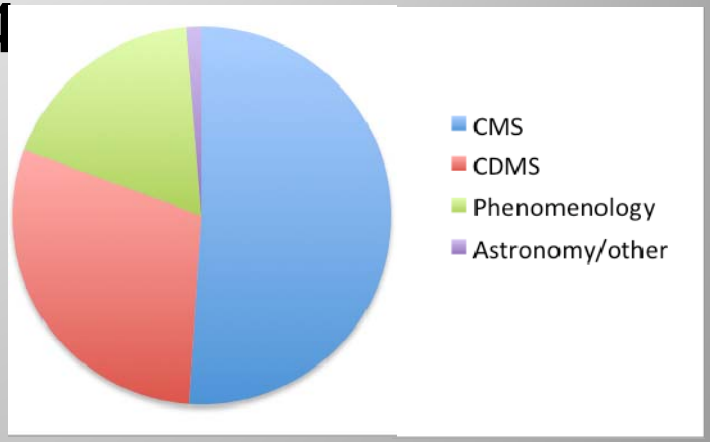
- **CPU-hrs per month**
  - **Picked up speed with new operating system and sharing rules**



More fun numbers

Top 5 Months

- 1. 1,750,719 core hours - January 2015
- 2. 1,702,351 core hours - December 2014
- 3. 670,049 core hours - November 2014
- 4. 584,213 core hours - April 2013
- 5. 503,757 core hours - May 2013



Top 5 Users of the Month

- 1. 1,487,169 core hours – K. Colletti (CDMS)- Jan 2015
- 2. 1,234,387 core hours – K. Colletti (CDMS)- Dec 2014
- 3. 476,798 core hours – K. Colletti (CDMS)- Nov 2014
- 4. 439,777 core hours – S. Wu (Pheno)- Dec 2014
- 5. 382,759 core hours – A. Perloff (CMS)- Apr 2013

Some Lessons Learned

- **Monitoring how quickly data gets transferred can tell you if there are bad spots in the network locally as well as around the world**
  - **Found multiple bad/flaky boxes in Dallas using PerfSonar**
- **Monitoring how many jobs each user has running tells you how well the batch system is doing fair-share and load balancing**
  - **Much harder than it looks, especially since our users are very "bursty": They don't know exactly when they need to run, and when they need to run they have big needs NOW (telling them to plan doesn't help)**
- **Experts that know both the software *and* the Admin is a huge win**
  - **Useful to have users interface with local software experts (my students) as the first line of defense before bugging Admins**
- **National labs are much better set up for "collaborative work" since they trust collaborations**
  - **Upside to working at the lab: Much more collective disk and CPU, important data stored locally**
  - **Downside: No one gets much of the disk or CPU (most of our users could use both, but choose to work locally if they can)**
  - **Different balancing security with ability to get work done is difficult**

Conclusions

- **Mitchell Institute physicists effectively use the Brazos cluster for our High Throughput needs**
  - **More disk would be good, but the ability to get data here quickly has ameliorated that (and been the envy of our colleagues)**
  - **More CPU/Supercomputing would be good, but the ability to accept jobs from the GRID is more important; running opportunistically has worked fine**
  - **The amount of red tape to get jobs in, and allow our non-A&M colleagues to run, has been significant (but not insurmountable)**
- **Bottom line: Been happy with the Brazos cluster (thanks Admins!) as they helped us discover the Higgs Boson**
- **Looking forward to when the LHC turns on in March!**