

Big Computing in High Energy Physics

College of Science
Big Data Workshop

David Toback

Department of Physics and Astronomy
Mitchell Institute for Fundamental Physics and Astronomy

January 2016



Outline

- **Particle Physics and Big Computing**
- **Big Collaborations and their needs:**
 - **What makes it hard? What makes it easy?**
 - **Individual Experiments**
 - Collider Physics/LHC/CMS
 - Dark Matter/CDMS Experiment
 - Phenomenology
- **How we get our needs met: Brazos Cluster**
- **Accomplishments and Status**
- **Lessons learned and Requirements for the future**
- **Conclusions**

Particle Physics and Big Computing

High Energy Physics = Particle Physics

- All activities within the *Mitchell Institute*
 - Includes both theory and experiment, as well as other things like String Theory and Astronomy
- *Big picture of science: Making discoveries at the interface of Astronomy, Cosmology and Particle Physics*
- Big Computing/Big Data needs 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)

What makes it Easy?

What makes it Hard?

Advantages:

- **Physics goals well defined**
- **Algorithms well defined, much of the code is common**
 - **We already have the benefits of lots of high quality interactions with scientists around the world**
- **Lots of world-wide infrastructure and brain power for data storage, data management, shipping and processing**
- **Massively parallel data (high throughput, not high performance)**

Disadvantages

- **Need more disk space than we have (VERY big data)**
- **Need more computing than we have (VERY big computing)**
- **Need to move the data around the world for analysis, for 1000's of scientists at hundreds of institutions**
 - **Political and security issues**
- **Need to support lots of code we didn't write**

Being able to bring, store and process lots of events locally is a major competitive science advantage

- **Better students, more and higher profile results, more funding etc.**
- **National labs provide huge resources, but they aren't enough**

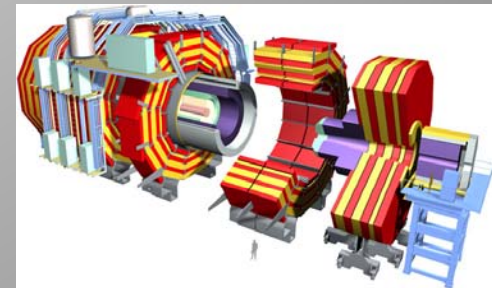
CMS Experiment at the LHC

(Discovery of the Higgs)

- Collider Physics at CERN/Fermilab have often been the big computing drivers in the world (brought us the WWW and still drive Grid computing worldwide)
- Experiments have a 3-tiered, distributed computing model on the Open Science Grid to handle the 10's of petabytes and millions of CPU hours each month
 - Fermilab alone runs more than 50M CPU-hrs/month
- Taking data now

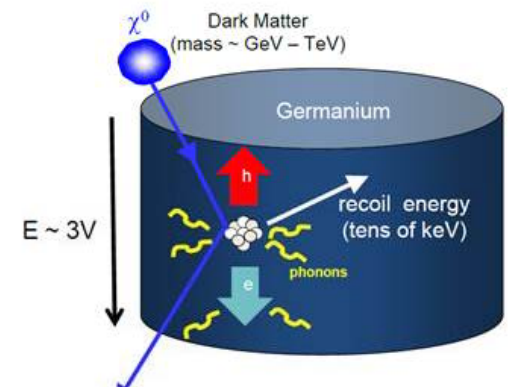
At A&M:

- Run jobs as one of many GRID sites as part of an international collaboration (we are a Tier 3)
- 4 Faculty, with ~20 postdocs + students



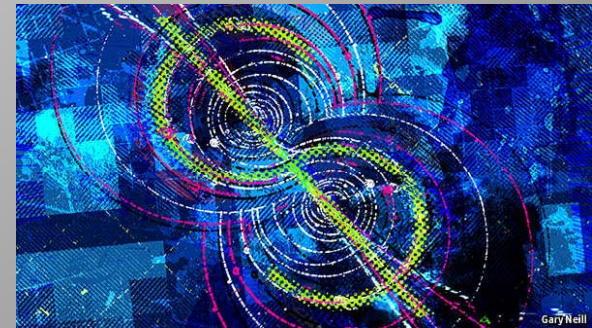
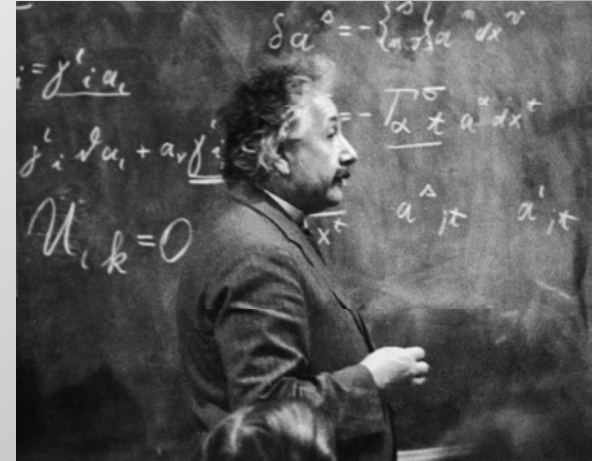
Dark Matter Searches with the CDMS Experiment

- Much smaller experiment (~100 scientists), but same computing issues
 - Only 100's of Tb, and just millions of CPU-hrs /Month
 - 4 faculty, ~20 postdoc + students
- Today: Most computing done at A&M, most data storage at Stanford Linear Accelerator Center (SLAC)
 - Big roles in Computing Project Management
- Future: Will start taking data again in 2020
 - Experiment will produce 10's Tb/month
 - Will share event processing with national labs (SLAC, Fermilab, SNOlab and PNNL)
 - We will be a Tier 2 center



Particle Theory/ Phenomenology

- Particle Phenomenologist's at MIST do extensive theoretical calculations and VERY large simulations of experiments:
 - Collider data to see what can be discovered at LHC
 - Dark Matter detectors for coherent neutrino scattering experiments
 - Interface between Astronomy, Cosmology and Particle Physics
- Just need high throughput, a few Gb of memory and a few 10's of Tb
 - Jobs don't talk to each other

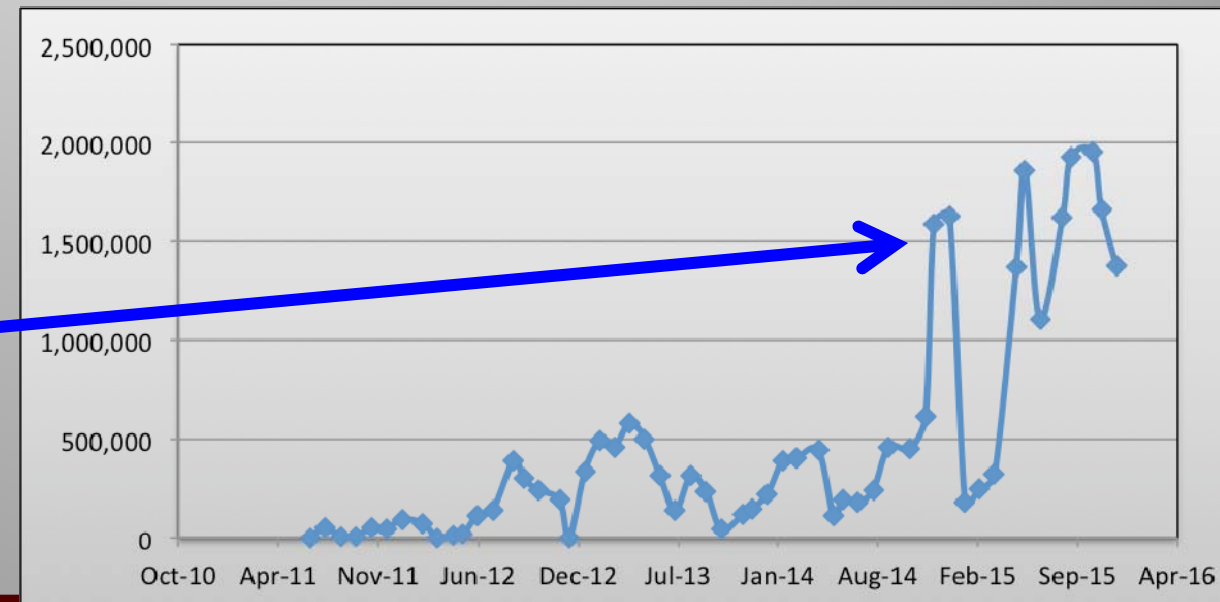
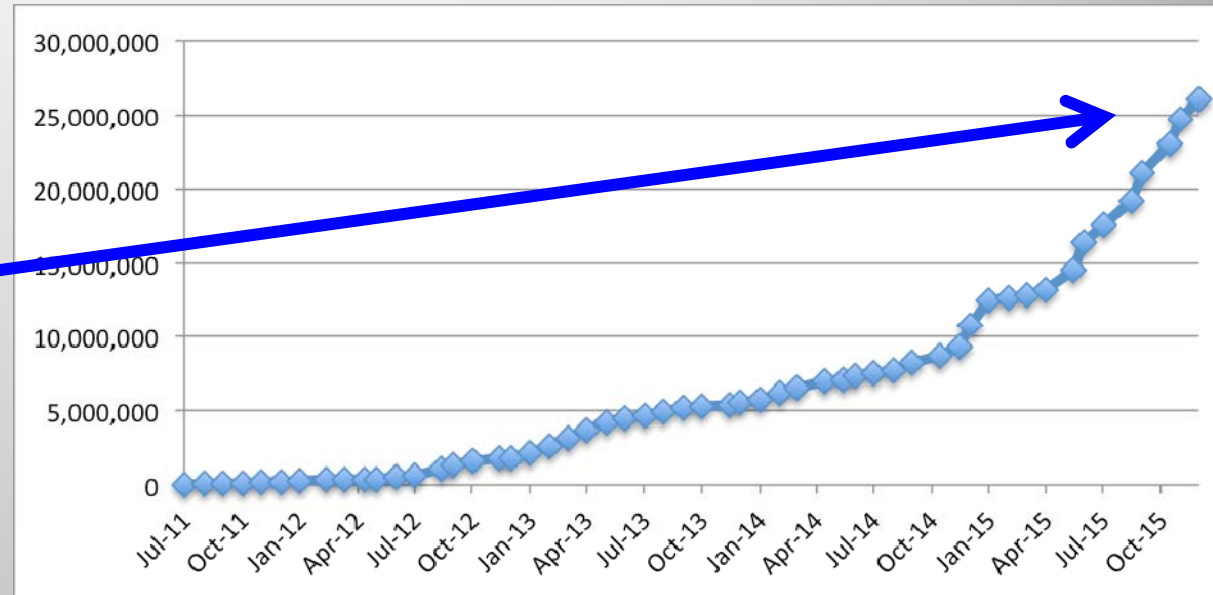


Overview of Brazos* and Why we use it (and not somewhere else)

- **HEP not well suited to existing Supercomputing at A&M because of experimental requirements**
 - **Jobs and data from around the world**
 - **Firewall issues for external users**
 - **Automated data distribution and local jobs regularly accessing remote databases**
- **Well matched to the Brazos cluster: high THROUGHPUT, not high PERFORMANCE**
 - **Just run LOTS of independent jobs on multi-Tb datasets**
 - **Have gotten a big bang for our by buck by cobbling together money to become a stakeholder**
 - **Purchased ~\$200k of computing/disk over the last few years from the Department of Energy, ARRA, NHARP, Mitchell Institute and College of Science**
 - **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.)**
 - **Accepts jobs from Open Science Grid (OSG)**
 - **Excellent support from Admins – Almes (PI), Dockendorf and Johnson**
- **More detail on how we run at: <http://collider.physics.tamu.edu/mitchcomp>**
- ***Special Thanks to Guy Almes for his Leadership, Management and wise council**

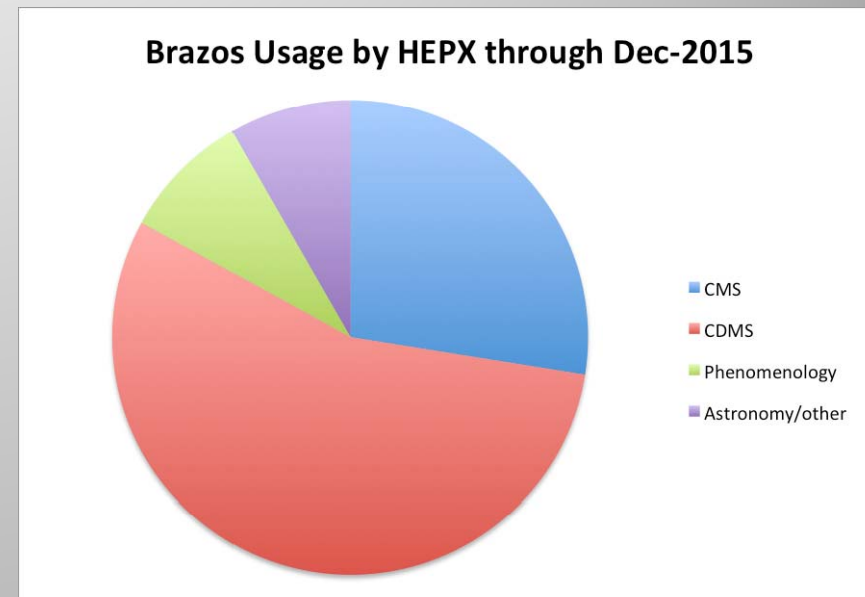
Fun Plots about how well we've done

- **Cumulative CPU Hours**
 - More than 25M core-hours used!
- **CPU-hrs per month**
 - Picked up speed with new operating system and sharing rules
 - 7 Months over 1.5M core-hours/month



More fun numbers: Sharing the Wealth

- **October 2015**
 - **1.95M Total core-hours used**
 - **1,628,972 core hours**
 - Katrina Colletti (CDMS)
- **Lots of individual users:**
 - **3 over 1M integrated core-hrs**
 - **23 over 100k**
 - **39 over 10k**
 - **49 over 1k**
- **Well spread over Dark Matter, LHC, Pheno and other**



Students

- **Already had 4 PhD's using the system, with more graduations coming in the next year or so**
- **Two of my students who helped bring up the system have gone on to careers in Scientific Computing:**
 - **Mike Mason: Master's 2010**
 - Now Computing professional at Los Alamos
 - **Vaikunth Thukral: Master's 2011**
 - Now Computing professional at Stanford Linear Accelerator Center (SLAC) on LSST (Astronomy)

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 some 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**
- **Hard to compete with national labs as they are better set up for “collaborative work” since they trust collaborations, but we can’t rely on them alone**
 - **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 of security with ability to get work done is difficult**

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>

Conclusions

- **High Energy Particle Physicists are (and have been) Leaders in Big Data/Big Computing for decades**
- **Local users continue this tradition at A&M by effectively using the *Brazos cluster* for our High Throughput needs**
- **We are happy to help others**
- **Want to help us?**
 - **The ability to get data here quickly has ameliorated short term problems for now, but we need much more disk and CPU**
 - **Have used Brazos and Ada for Simulations, but Ada has limited our ability to store data and run jobs. Getting Ada on the OSG would be helpful**
 - **The amount of red tape to get jobs in, and allow our non-A&M colleagues to run, has been significant (but not insurmountable). Slowly getting better**
 - **Provide software support in addition to Admins**
- **Bottom line: Been happy with the Brazos cluster (thanks Admins!) as they helped us discover the Higgs Boson, but our needs have been growing faster than the CPU and Disk availability, and we've struggled to support the software**