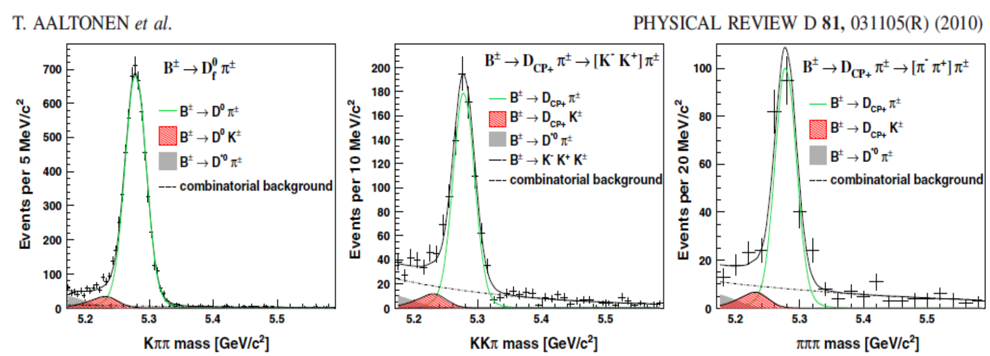
Large Scale Computing Efforts and Data Preservation

Stephen Wolbers

CD/SCF (Scientific Computing Facilities)

DOE Site Visit

August 24, 2010



Large Scale Computing Efforts

- Many large-scale computing efforts are common for all of the scientific computing activities – Run 2, Intensity Frontier, CMS, Astroparticle Physics, Accelerator Physics, Lattice QCD, Theory, etc.
  - Storage
  - GRID/Clouds
  - Sharing
  - Efficient Operations
  - Facilities

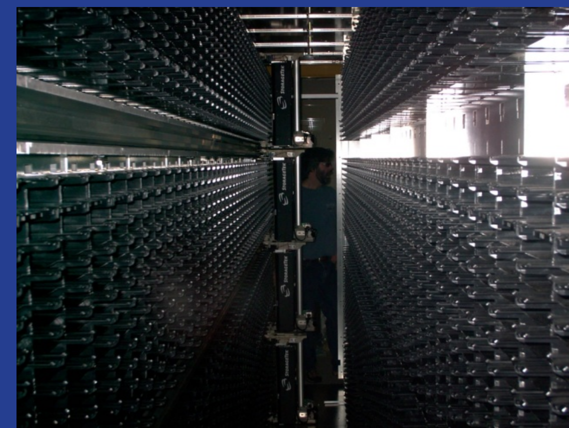
Storage

- Storage (tape and disk) is critical for all scientific activities
- Tape-based robotics
  - 6 STK (ORACLE/SUN) SL8500 tape libraries
    - 10000 tapes each
    - LTO3, LTO4, LTO5 (soon) tape technology
    - Enstore hierarchical mass storage software
    - Currently total of 24 PB for all FNAL
  - All old robots now decommissioned (as of end of July, 2010) (Big accomplishment!)
    - Large migration activity
    - No bytes left behind (see "Data Preservation" – mentioned elsewhere)

Tape Libraries – refresh of technology

2009 decommissioning of ADIC AML/2

2010 decommissioning of STK 9310



Current Libraries:

Oracle SL8500

Current capacity:

6 libraries x 8 Petabyes

Storage

- Future of mass storage.
- We must constantly move forward with mass storage systems.
- Will investigate other robotic systems and vendors (would like to benefit from competition).
- We are working on enhancements to Enstore to handle small files, scaling to larger datasets and data-rates (especially CMS), etc.
- We have to think about the future (5-10 years) of mass storage software and architecture.
- These all take time and most especially effort.

Storage (II)

- **Disk Storage**
  - Large resources, many technologies and software packages, including BlueArc, dCache, Hadoop, LUSTRE.
  - Investigations at Fermilab and elsewhere help determine how to effectively and efficiently use multi-PB (or smaller!) disk storage space for production, analysis, and other uses.
  - Will continue to grow in importance as disk becomes a larger component (100%?) of storage resources.
  - CD engages in many of the investigations and storage development that allow HEP to take advantage of newer technologies and to optimize our use of all of the many technologies available.

GRID/Clouds

- FermiCloud
  - A System has been acquired and is in initial testing and use.
  - Initial capabilities and testing are targeted for Phase 1.
  - Phase 2 targets low-cpu-load servers.
  - This all provisions FNAL to make use of DOE "Cloud" Computing and to help maximize our use of on-site or commercial computing.
  - This is a fairly complicated activity and requires a reasonable amount of effort to carry out the necessary investigations.
- GRID covered by Ruth

Sharing

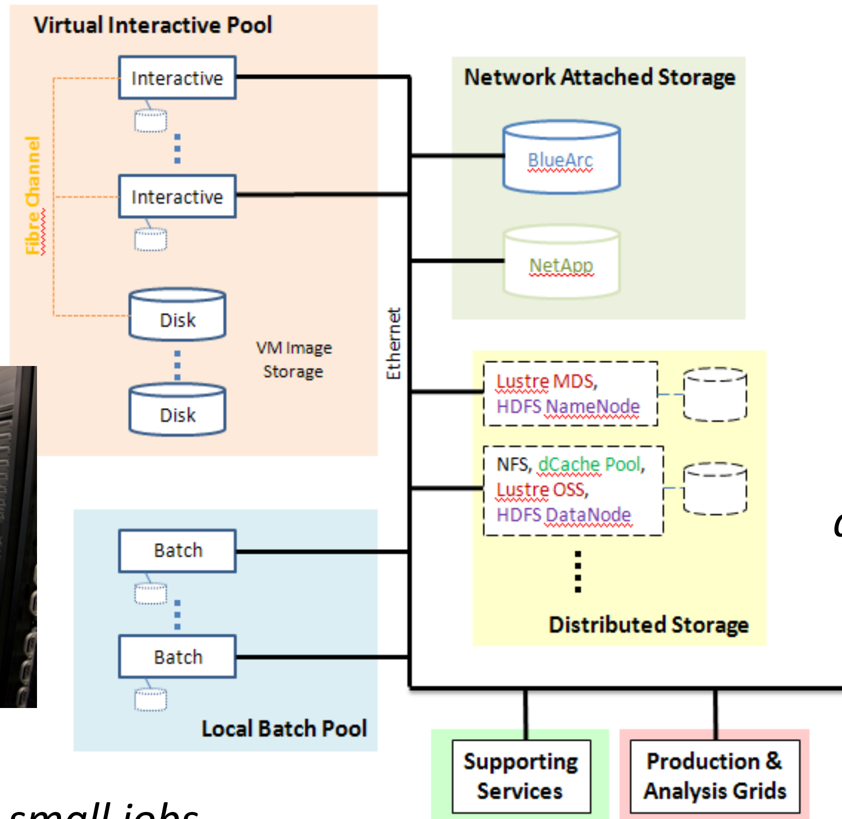
- There is a strong history of sharing large resources and computational techniques at Fermilab. These include:
  - Farms. Sharing has been a feature of the large computing farms since the early 1990's. This makes the most effective use of the systems and also satisfies the priorities of the laboratory's physics program.
  - New HPC cluster to share resources for TD, AD and CD accelerator research is acquired and will go into production soon.
  - Mass Storage.
    - Shared software (Enstore/dCache)
    - Shared libraries (Currently 6 SL8500)
  - Management/Monitoring
    - Trying to share or use similar or common software.

Experiments share common resources

Adjustable and flexible allocations

Interactive resources in Virtual Machines

General Physics Computing Facility (GPCF)



Network attached common storage for user and project data

Highly distributed storage for high I/O analysis requirements (future)

Local batch pool for small jobs and as "training ground" for grid

Production and large analysis jobs on grid (FermiGrid)

Supporting services reside in virtualized servers (FermiCloud)

New HPC cluster

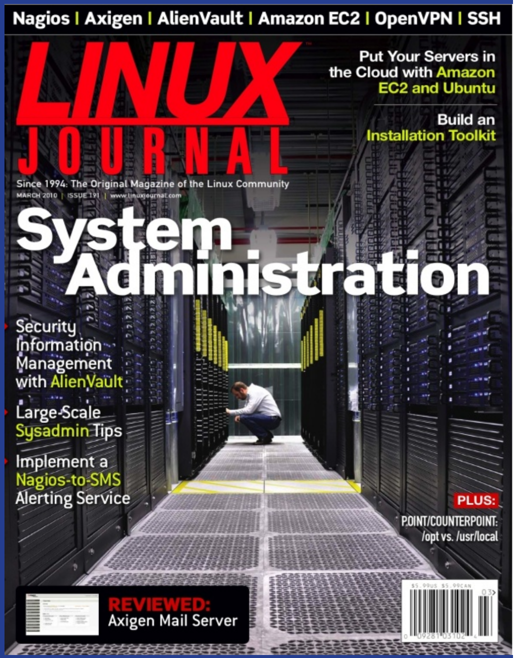
- A new HPC (parallel processing) cluster has been acquired using TD, AD and CD funds.
- Initially targeted at accelerator simulations, SCRF cavity design, and related calculations.
- There may be a need for more general purpose tightly-coupled computing of this type and we will have the basis for expanding the system accordingly
- We also have the possibility of sharing this resource in various ways (GRID, Cloud, etc.)

Efficient Operations

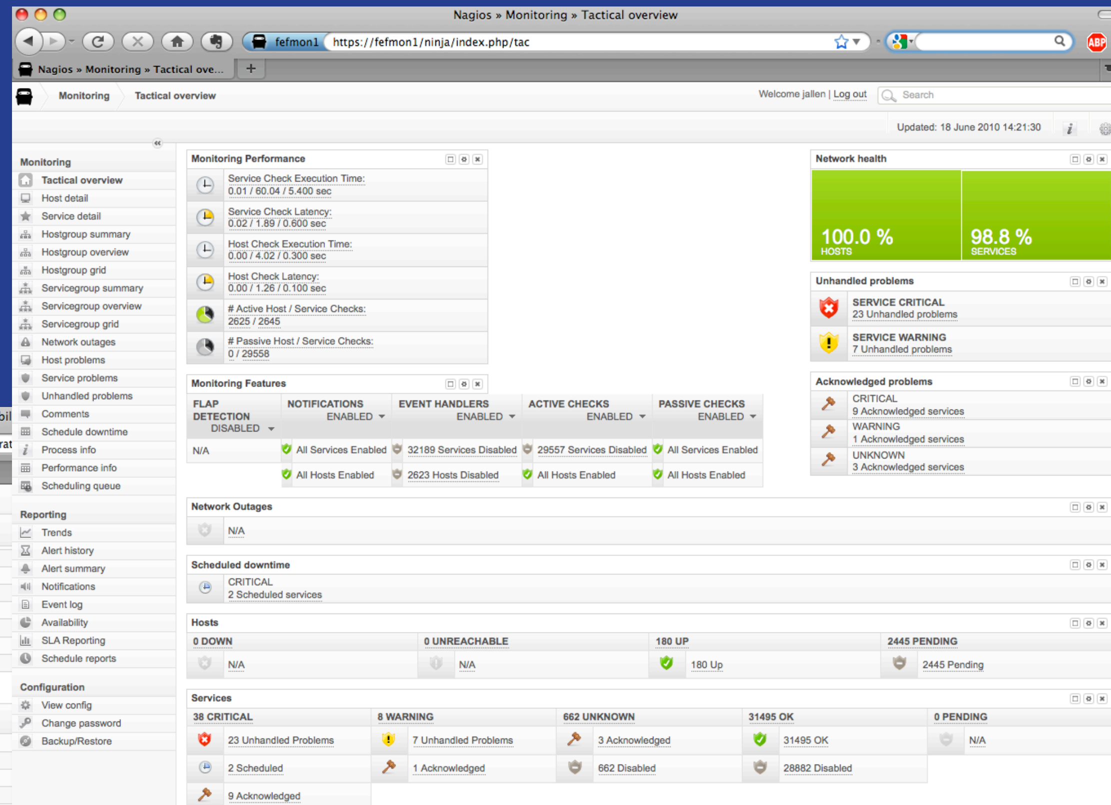
- One common theme in SCF (Scientific Computing Facilities) is efficient operations. This is accomplished via:
  - Common purchasing wherever possible, saving a great amount of effort in specifying, testing, installing, and debugging hardware.
  - Common monitoring and management tools.
  - Common administration staff.
  - Common investigations of new technologies and techniques (e.g., disk and virtualization).
  - Sharing knowledge

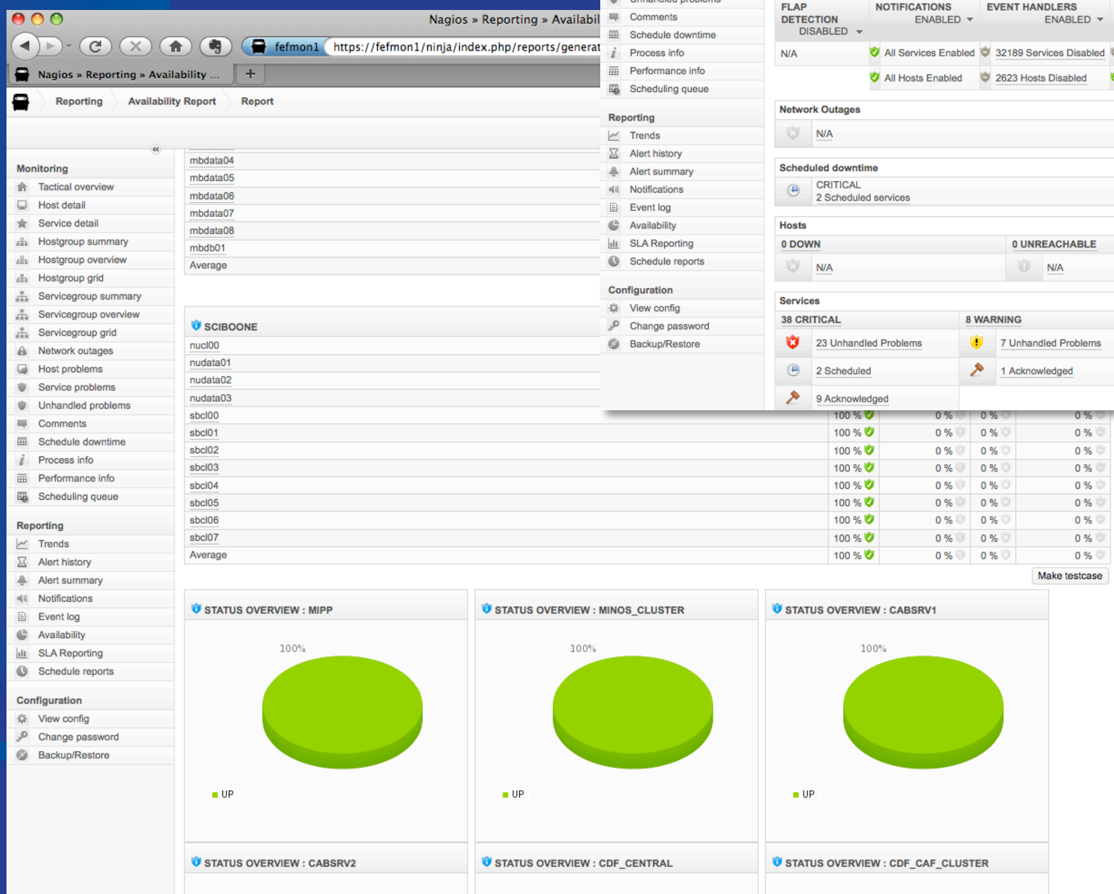
Sustaining Run II + growing Intensity Frontier computing support

- Operations are almost as efficient as they can be
  - Leverage Open Source tools to overcome the challenges of large scale system management.
  - Recruit and retain top IT talent for our large complex environment



Using Nagios to monitor almost 4000 systems and 35000 services (excl.CMS)





Integrated with asset management database and Service Desk for automatic paging and ticket generation

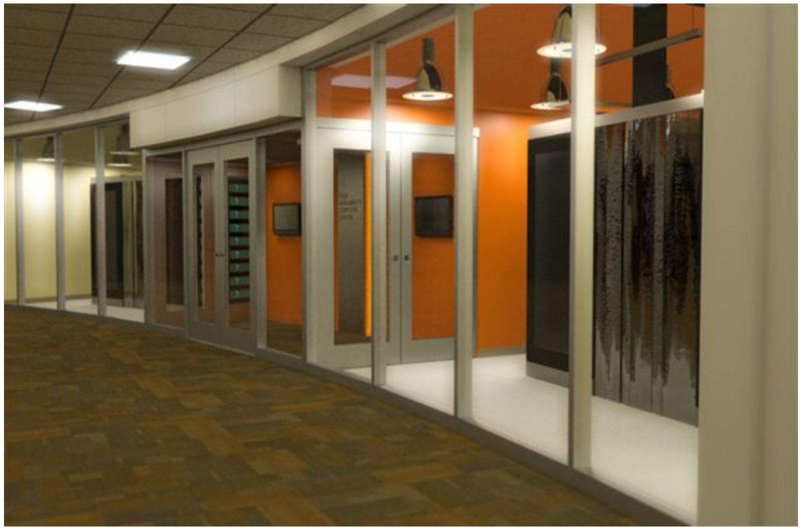
Facilities

- All of the computing has to "live" someplace.
- The computer rooms have to have certain characteristics:
  - Space
  - Power and power density
  - Cooling and cooling density
  - Networking
- 3 (soon 4) major facilities:
  - FCC (mainly FCC2)
  - GCC
  - LCC
  - FCC3 (in October 2010)

Facilities

- **Facilities require:**
  - Excellent and specialized staff
    - Just recently created a new family of "Data Center Facilities Operations" job titles.
  - Sufficient space, power, cooling
  - Uptime requirements
    - Lead to other requirements, such as good monitoring, good maintenance programs, etc.
    - UPS, Generators, dual-power of certain services, etc.
  - Division of types of services into types of facilities
  - Excellent networking that allows for flexibility in the use of the facilities.

Computing Centers



Data Preservation

Data Preservation

- Data Preservation is an activity which can be loosely defined as:
  - Retaining the capability to analyze experiments after the data-taking period ends.
  - Exploring the possibility of extending that period beyond the lifetime of the collaboration.
  - Exploring the possibility of other access to the data, including other experiments (for combinations or updating results with newer models/theories), other scientific users and the general public.
  - Building a framework for accomplishing the above in a uniform way across HEP.

Data Preservation

- **Status of the International Effort**
  - Data Preservation workshops have been held at DESY, SLAC, CERN (2009) and KEK (2010).
    - Fermilab has been represented at all the workshops, with CD, CDF and D0 sending reps in most cases.
    - DOE has also been observing.
  - The workshops are aiming to provide a document "the DPHEP blueprint" to ICFA, funding agencies, laboratories, etc.
  - Latest iteration was presented to ICFA on July 24.
  - Next workshop tbd – an informal request to FNAL to host in spring 2011 has been received.

Data Preservation

- **FNAL activities:**
  - Participation in workshops.
  - Discussions in CD, CDF, D0.
  - Plan for 5 year analysis capability beyond the end of Run 2.
  - Migration of all data (in the robots) to new media is automatic and will continue indefinitely.
- **Future**
  - Study the efforts of BaBar and H1/ZEUS as they move forward.
  - Maintain participation in Data Preservation workshops.
  - Understand the complex requirements for maintaining analysis capability for code, documentation, calibrations, associated information.
  - Utilize computing techniques such as virtualization to enable some of the capabilities.

Conclusions

- There is ample opportunity to partner with others where appropriate to explore new technologies, new operational techniques, solve common problems, and take advantage of the knowledge and resources available in many places.
- In this discussion I mentioned a few but there are undoubtedly others. Some could become real, some might never work out for various reasons.