

Strategic Plan for Scientific Computing Facilities (2007-2009)

D. Petravick, et al.

10/1/06

Mission

Fermilab has a long history of technical leadership with respect to scientific facilities. Fermilab's Scientific Computing Facilities organization works proactively with experiments to plan, provision and operate, and assess computer systems and services for scientific computations.

Context and Assessment of Current State

As of early FY 2007, Fermilab is providing productive facilities. The facilities themselves are well aligned with experiments needs, and indeed have often helped define the mainstream facility model for computing for HEP and LQCD. There is an extant body of best practices, which include benchmarking, system installation, system patching, and batch system monitoring.

The elements of strategy that has lead to the current state include:

- Appropriate and close relationships with the scientific organizations.
- Awareness of the technical evolution of facility components.
- Adherence to price-performance as a prime metric in procurement decisions.
- Willingness to determine a good balance between self-integration and procured systems.
- Proper attention to aspects of the facility's need to sustain its operational effectiveness, for example, backups and security. These attributes are not always obvious to end-user stakeholders.
- Willingness to out-source work within and outside Fermilab in order to obtain efficiencies. (examples the are the outsourcing the Scientific Linux and of repairs to Decision One).

Fermilab facilities are organized into three departments, aligns with major scientific efforts: CMS; Running Experiments; and Lattice QCD. The operational concept of each department differs, because the computing models of each experimental group differ.

The Running Experiment facilities provide well-architected ensembles of networked computers possessing well-administered operating systems; connectivity to storage systems; basic infrastructure middleware infrastructure, for example batch systems; and support for experiment run-grid software. The general purpose farm is completely integrated to Fermigrid.

Lattice QCD facilities supply an HPC environment apropos to the Lattice community, The which includes a login service and support for the Lattice data grid.

CMS offers a login service, however the majority of its resources are offered behind the still-developing WLCG (including the OSG) grid service abstractions. In addition the CMS facility provides both batch and interactive capacity to the CMS Center. There is a significant understanding of and feedback to the CMS computing model. The model extends beyond Fermilab, and demands excellent connectivity to every CMS T2 center world wide, and expert support of designated Tier 2 facilities.

All facilities are on a track to provide opportunistic, reciprocal access of their facilities using grid methods.

Lastly, the facilities have played an important role in developing general methods of organizing large distributed collaborations. The methods implemented in the context of grid technologies, but the intellectual underpinnings are technology invariant. Examples include: Organization of users, Identification of the trust that facilities place in experiments and the trust that experiments place in facilities, and efficient, end to end problem resolution.

Vision

We expect a center of excellence in provisioning computational and data management facilities to the extended High Energy Physics community – second to none in the world.

- Run II computing to be in a steady and stable state.
- To have a fully provisioned CMS facility premiere grid facility and interactive facility for the LPC.
- To have an important role in the Lattice Facilities
- To be an important contributor to HEP's general computing needs.
- To be a Full partner in national and international computing efforts.
- To be Recognized strength for hosting the ILC.

Stakeholders

The particle physics community (including accelerator-based experiments, Lattice Computing, the Grid, and particle astrophysics) is FNAL's customer.

The Facilities are stakeholders to the Division's service architecture efforts, and to infrastructure providers outside of Fermilab, most notably the Open Science Grid and the WLCG.

Goals and Objectives

1. Maintain a world-wide reputation as a leading center for scientific data. Provide facilities that are responsive to the Experiments needs, interacting with their computing models and scientists.
2. Sustain excellent capacity and usability, including capacity planning, including the potential for provisioning capacity on contingency, with provisioning and operations at justifiable costs.
3. Further development and deployment of the grid service model, development and population of the Open Science Enclave, and continuing development of a secure, usable security model
4. Further development of the Computing Division's internal service model.
5. Integrate with experiment, national, international and HEP cyber-infrastructures, with awareness of HEP funding agencies and peers.

Strategies

1. Interact with experiment's computing models and scientists.
2. Apply effort to metrics on and feedback to software and service providers, and develop new service suppliers.
3. Work jointly with the FNAL CST, and grid security organizations to define and populate the FNAL Open Science Enclave.
4. Measure all costs, maintain comparables, and clearly state business methods.
5. Represent Fermilab facilities at appropriate technical and other fora.

Resource Needs

There is a shortage of computer room space. There is a shortage of system administrators and service managers. Skills to document and measure costs, provide metrics and state business models are short. Skill in helping set OSE policy and practice need to be developed.

Progress Indicators

1. *Increased understanding of the facilities utilization, and a linkage of facility use to scientific utility.*
2. *Increased commonality of approach in facility underpinnings.*
3. *Movement to a service-based interface to experiments.*
4. *Increased robustness of service offered, for a constant level of support effort.*





