Fermi National Accelerator Laboratory
P.O. Box 500 • Batavia, IL 60510
Office: 630/840-3921
December 2008

Application by Fermilab to Join the Department of Energy Computational Science Graduate Fellowship Program as a Practicum Site

Affiliation with the DOE.

Fermi National Accelerator Laboratory (Fermilab) advances the understanding of the fundamental nature of matter and energy by providing leadership and resources for qualified researchers to conduct basic research at the frontiers of high energy physics (HEP) and related disciplines. Fermilab is managed by the Fermi Research Alliance (FRA), a joint limited liability company of the Universities Research Association (URA) and the University of Chicago.

Effective in 2007, the U.S. Department of Energy (DOE) awarded to FRA the management and operating contract for Fermilab. FRA supports Fermilab's Laboratory Director Pier Oddone's long-term vision and total commitment to discovery, to the DOE scientific mission, and to ensuring our nation's continued preeminence in particle physics.

DOE provides the funding for Fermilab to: run the Tevatron, the highest energy accelerator in the United States, provide the data processing and distribution systems for the experiments using the Tevatron and related accelerators, beam lines for neutrino physics, and design new accelerators for the future; host the "Tier-1" and US LHC Physics Center for storage, distribution and analysis of data for the Large Hadron Collider (LHC) Compact Muon Solenoid (CMS) experiment (one of the largest science experiments in the world); house one of the "US top" computing farms in support of theoretical physics calculations; execute a program of neutrino and rare processes physics; and be a leading center for theoretical astrophysics in support of the local experimental astrophysics groups involved in the Dark Energy Survey, the Sloan Digital Sky Survey, and potentially the Joint Dark Energy Mission.

Computing Facilities, Computational Science and HPC Research

Fermilab's Scientific Computing Facilities provide the broad Fermilab user community (including accelerator and accelerator R&D scientists) with state-of-the-art large scale computing, storage, networking and database resources to meet their needs. The computing facilities include:

- A Lattice QCD facility which supplies high-performance computing services to the Lattice Quantum Chromodynamics US and international community.
- The LHC CMS experiment's Tier-1 in the US which archives data for later distribution to the several hundred CMS physicists in the US and several thousand members worldwide, as well as provides significant data reprocessing and analysis capabilities.

- High performance computing systems that serve astrophysics and accelerator modeling and simulation.
- Grid accessible high throughput computing farms which serve the data processing needs of all the users, and interface to the multi-disciplinary national grid infrastructure, the Open Science Grid.
- User accessible analysis facilities providing hundreds of users the ability to perform interactive analysis of the final data sets.
- Tape storage facilities for use by all Fermilab users and large disk farms as tape front-ends and online data caching systems.

The current total capacity of the facilities is 30,000 processing cores, 13 Petabytes of archival tape storage, easily expandable to 40 Petabytes, and more than 3.5 Petabytes of disk storage. The processors are used at 85% of capacity; Typical tape writing is at 100 Terabytes a day and reading is at >100 Terabytes a day. Data transfers are on average >2Gbit/sec to and from the Wide Area Network (with frequent bursts of >10Gb/sec). The monthly total of data transferred between Fermilab and the WAN is thus >0.65 Petabyes per month inbound and 2.2 Petabyes per month outbound.

Fermilab's computational activities include: modeling and simulation of the accelerators and the experiment detectors and physics processes; accelerator and detector readout and control systems; theoretical and experimental astrophysics; simulation, processing analysis, management, and distribution of the experimental and theoretical data; and massively parallel numerical simulations of Quantum Chromodynamics in support of theoretical high energy and nuclear physics programs. Fermilab does development and research of state-of-the art systems and software in support of these activities both in the methods and technologies for the scientific calculations as well as in the areas of networking, large scale distributed computing, scientific workflow and analysis frameworks.

The Fermilab Computing Division has a staff over more than two hundred and seventy. Members of the Computing Division include more than forty scientists and one hundred and eighty computing professionals, computer scientists and engineers. Scientists and engineers in the other divisions and sections in the Laboratory collaborate on many computational activities, especially the CMS Physics, Astrophysics and Accelerator Centers. The scale and complexity of the data in the physics and astrophysics programs provide real-time data acquisition, computational and data challenges that lend themselves to research and development from computer science and related areas of computational research. Programs at Fermilab that already include such activities are:

- *Accelerator modeling and simulation* as part of the SciDAC Compass project, which Fermilab leads. Specific challenges include Particle In Cell (PIC) algorithms and implementation, Parallel differential equation solvers and well as general infrastructure needs as described below.
- *Advanced scientific workflow* for the Lattice QCD scientists as part of the LQCD SciDAC program. Fault detection and mitigation systems for large-scale scientific computing facilities. Multicore multithread optimization techniques for massively

parallel computations. Investigations of Graphics Processing Units (GPUs) for high-performance scientific computing.

- *Theoretical astrophysics* for a wide range of projects in computational astrophysics, from modeling the large-scale distribution of dark matter and galaxies as probes of dark energy evolution, to simulations of individual galaxies and supermassive black holes. Computational physics challenges include adaptive mesh refinement techniques to model the evolution of dark matter, cosmic gas, star formation, and radiative transfer on massively parallel supercomputers. Numerical challenges include large dynamic range (1 million in space and 1 billion in time), rich multi-scale physics (from non-equilibrium ionization and thermodynamics to long-range gravitational interactions).
- *Advanced distributed systems research and development.* These include*: Advanced network technologies* and end-to-end deployments as part of the US CMS world-wide system (programs like the DOE ASCR funded project LambdaStation); *Computer security and trust analysis and modeling* for the distributed, collaborative systems deployed for experiment data processing and analysis (including investigations of the utility of social networks based security intrusion and incident pattern access recognition); *Data management and distribution* technologies (including collaborating on the SciDAC Center for Enabling Distributed Petascale Science).

Some specific computational science research challenges are given in the following table:

	Accelerator Modeling	**Theoretical Astrophysics**	**Lattice QCD**	**Accelerator Experiment methods**	**Grid Computing**
Particle In Cell (PIC) algorithms & implementation.	Y				
Parallel differential equation solvers.	Y		Y		
Workflow tools.	Y	Y	Y	Y	Y
Performance tools (single & parallel processing).	Y	Y	Y	Y	Y
Concurrent programming techniques (multi-threaded, multi-core).	Y	Y	Y	Y	Y
Large scale data management & storage.	Y	Y	Y	Y	Y
Advanced network management & integration.		Y	Y	Y	Y
Advanced security techniques for open science.	Y	Y	Y	Y	Y

Benefits to a Fellow doing a Practicum at Fermilab

Fermilab provides a unique environment for the deployment and use of new techniques and methods in high throughput and high performance computing in production environments. The practical needs and constraints of running and maintaining production quality systems and services drive and inform the research and development activities. The ongoing collaborations of developers and users provide iterative feedback and influence between the design and evolution of the research and experiences in the field. Fermilab's environment facilitates solutions that are both usable and effective. Fermilab provides a "hands-on" community where engaged groups of scientists and engineers work together to solve hard practical problems.

Fermilab has many existing collaborations between laboratory staff and computer and computational science faculty at universities. These will facilitate mentorship for and participation in the Computational Science Graduate Fellowship Program.

4. Practicum Coordinator

The Practicum Coordinator is Ruth Pordes, Associate Head of the Fermilab Computing Division and Executive Director of the Open Science Grid. Ruth has more than 20 years service in the Computing Division and has led multiple projects that are a collaboration between scientists, engineers – including computer scientists and computing professionals. Her resume is attached below.

5. Logistical Contact

Ruth Pordes will act as the logistical contact and will receive support from the Computing Division's Administrative Support Group led by Griselda Lopez as needed.