Grid Computing in High Energy Physics

I am a High Energy Physicist, doing my research with two major international laboratories, namely, the Fermi National Accelerator Laboratory (Fermilab) located in Batavia, Illinois and the European Center for Nuclear Research (CERN), situated in Geneva, Switzerland. The two projects that I am involved with at these two laboratories include the DZERO collaboration at Fermilab and CMS (Compact Muon Solenoid) at CERN.

At the foothills of the Jura Mountains, on the border between France and Switzerland is a 17 mile circular tunnel, currently under construction, called the Large Hadron Collider (LHC). Once complete, in 2007, it will smash particles, protons to protons with nearly the speed of light to create conditions that occurred a few seconds after the Big Bang. These collisions will happen at an unprecedented energy of 14 trillion electron volts. This is an enormous effort to attempt to answer fundamental questions in particle-physics that has boggled the minds of scientists for a long time. The theory that once was believed to explain sub-atomic interactions completely, known as the Standard Model, is now merely a stepping stone to a grand theory that would describe all physical phenomena in nature. The beam collisions would reveal physics beyond the Standard Model. The physics program includes the study of electroweak symmetry breaking, investigating the properties of the top quark, a search for new heavy gauge bosons, probing quark and lepton substructure, looking for supersymmetry and exploring other new phenomena.

Once operational we shall collect data of the order of 15 Petabytes (15 million Gigabytes) every year. This data shall be used by thousands of scientists from all around the world. One of the challenges that this scenario shall pose would be to build and maintain data storage and provide analysis infrastructure for the entire high energy physics community. The current model adopted at the LHC includes a four-tier Grid Structure which shall distribute data worldwide.

CERN will serve as the "Tier-0" center with a primary backup recorded on tape. This data, after initial processing, will be distributed to a series of Tier-1 centers, which are large computer centers with sufficient storage capacity for a large fraction of the data, and with round-the-clock support for the Grid. The Tier-1 centers will make data available to Tier-2 centers, each consisting of one or several collaborating computing facilities, which can store sufficient data and provide adequate computing power for specific analysis tasks. Individual scientists will access these facilities through Tier-3 computing resources, which can consist of local clusters in a University Department or even individual PCs, and which may be allocated to the Grid on a regular basis.

I will talk about how the technique of grid computing is already very successful in high energy physics with focus on the Northwest Indiana Computational Grid Consortium, which consists of 4 institutions collaborating on this major effort for high performance computing.

Dr. Neeti Parashar
Assistant Prof. of Physics
Purdue University Calumet
Indiana, USA