Computing and Data Grids:
Infrastructure for Large-Scale Collaborative Science

William E. Johnston, Tom Hinke, Tony Lissota, Piyush Mehrotra, and Arsi Vaziri
Information and Computing Sciences, Lawrence Berkeley National Lab

As the practice of science moves beyond the single investigator, due to the complexity of the problems that now dominate science, large collaborative and multi-institutional teams are needed to address these problems.

In order to support this shift in science, the computing and data handling infrastructure that is essential to most of modern science must also change in order to support this increased complexity. This is the goal of computing and data Grids: Software infrastructure - middleware - that facilitates solving large-scale science problems by providing the mechanisms to access, aggregate, and manage the computer network based infrastructure of science. This infrastructure includes computing systems, data archive systems, scientific instruments, and computer mediated human collaborations.

In this talk we will examine several large-scale science problems, and their requirements for computing and data Grid infrastructure. We will also discuss the basic Grid services, and their evolution toward Web services, that are designed to provide the necessary computing and data functionality for the routine operation of wide spread, collaborative science.