

# **Enabling Data Intensive Collaborative Science**

**Dr. Tefvik Kosar**

**LSU Computer Science Department and CCT**

Date: Feb 8th, 2008  
Time: 3:00 PM  
Venue: 256 Coates Hall

## **Abstract**

The trend in scientific as well as commercial applications from a diverse range of fields has been towards being more and more data-intensive over time. This increase in the demand for large-scale data processing has necessitated collaboration and sharing among the world's leading education, research, and industrial institutions and use of distributed resources owned by collaborating parties. In a widely distributed environment, data is often not locally accessible and has thus to be remotely retrieved and stored. While traditional distributed systems work well for computation that requires limited data handling, they fail in unexpected ways when the computation accesses, creates, and moves large amounts of data especially over wide-area networks. Scientists, researchers, and application developers are often forced to spend a great deal of time and energy on solving basic data-handling issues, such as the physical location of data, how to access it, and/or how to move it to visualization and/or compute resources for further analysis. In this talk, I will present our most recent research in solving these data-handling issues in a distributed environment and our vision of a new paradigm called data-aware distributed computing.