Special Tutorial : Tools and Services for Data Intensive Research
by
Roger S. Barga, PhD , Microsoft Research

Data Deluge in Scientific Research. Most scientific data analysis comprise analyzing voluminous data collected from various models and instruments. Efficient parallel/concurrent algorithms and frameworks are key to meeting scalability and performance requirements entailed in such scientific data analysis. Microsoft Research is working with researchers to apply and evaluate the same technology we routinely use in search to analyze petabytes of data on clusters of thousands of computers, to tackle data intensive research challenges. We are also providing this software, with programming guides and tutorials, for academic research and education. In this tutorial we will present a hands-on introduction to Microsoft Dryad, a high-performance general-purpose distributed computing engine that is designed to simplify the task of implementing distributed applications on clusters of Windows-based computers. DryadLINQ allows developers to implement Dryad applications in managed code by using an extended version of the LINQ programming model and API. Attendees will given a USB drive containing the Dryad software, along with programming guides and development documentation to take back to their lab for continued investigations and application.