ISCR

Institute for Scientific Computing Research

Laboratory Directed Research and Development Project Abstracts

CASC

Center for Applied Scientific Computing

Novel Parallel Numerical Methods for Radiation and Neutron Transport

Peter N. Brown Center for Applied Scientific Computing

Abstract

We propose to design advanced numerical methods for the parallel solution of three-dimensional radiation and neutron transport problems on massively parallel computers. Our emphasis will be on the development and implementation of methods for the parallel solution of the Boltzmann transport equation, as well as related equations.

Recent developments in the areas of First-Order System Least-Squares (FOSLS) methods show great potential for providing more accurate and robust solution procedures than current approaches. The FOSLS-based approach also provides a natural scalable multilevel (i.e., multigrid) solution procedure for the resulting discretized problems, and is independent of the discretization used in phase space. We also propose to develop more accurate phase space discretization techniques, in space and direction (angle) as well as better time-stepping algorithms.

Large-Scale Scientific Data Analysis and Visualization

Mark Duchaineau Center for Applied Scientific Computing

Abstract

We propose a collaboration among CASC, University of California, Davis, University of Texas at Austin and LLNL Defense and Nuclear Technologies Department (DNT) scientists to develop and apply efficient multi-resolution and compression methods for high-performance visual exploration and precise quantitative analysis of extremely large scientific data (terabytes to petabytes).

In particular, we will investigate sparse (i.e., output-sensitive) wavelet transforms applied to variables defined over time-dependent, unstructured, adaptive 3D grids of mixed materials with general geometries, and fully output-sensitive, view-dependent optimization applied to the families of surfaces derived from such variables and grids. The algorithms will be fully scalable in the algorithmic sense, and efficiently implemented on distributed- or shared-memory parallel systems. Initially, the algorithms will be applied to ASCI and DNT physics simulations; a broad range of further applications is well known and will be aggressively pursued.

This project leverages considerable ASCI and Numerical Environment for Weapons Simulations (NEWS) funding, and is synergistic with other current and proposed CASC projects.

Applying Data Mining Technologies to Large-Scale Scientific Data Sets

Chandrika Kamath
Center for Applied Scientific Computing

Abstract

There is a rapidly widening gap between data collection capabilities and the abilities of scientists to explore, analyze, and understand this data. As a result, scientists at the Laboratory are expressing concern at the potential loss of useful information in this data. To address these problems, a new generation of computational techniques is needed to help automate the exploration and analysis of large-scale scientific data. Our research focuses on applying and extending ideas from the area of data mining, in particular pattern recognition, to improve the way in which scientists interact with large, multi-dimensional, time-varying data. As pattern recognition algorithms are common across problem domains, it will be possible to apply our research to many different applications across the Laboratory.

Data mining, though a field in its infancy, has shown promise of significant payoffs when applied to small, low-dimensional data. Several concerns, however, remain to be addressed in extending these algorithms to large-scale data, especially in the scientific domain. Our approach to scaling pattern recognition techniques to terabyte and petabyte data sets addresses the following issues.

- Effective preprocessing of the data using techniques such as dimension reduction and sampling to make the search for patterns in large, multi-dimensional data sets tractable.
- Development of new techniques to improve the effectiveness of pattern recognition algorithms.
- Error modeling to allow user control of accuracy versus time tradeoffs.
- Efficient parallel implementations to enable interactive exploration of data.

This research will combine advances in high-performance computing with techniques from machine learning, soft computing, and statistics. We will first demonstrate large-scale pattern recognition techniques in two test-bed applications, MACHO and FIRST. Then, in the follow-on years of this effort, we will apply our research to other relevant applications at the Laboratory, including ASCI.

Proposed work and anticipated results

During FY99 we propose to investigate several research ideas in image processing and pattern recognition algorithms. Using the image and catalog data from the FIRST project, we will

- Implement an algorithm to group radio sources.
- Create training and test data from existing known images of bent-doubles and sources other than bent doubles.
- Implement classification algorithms such as neural nets and decision trees to automatically identify bent doubles.
- Update the image catalog with additional data derived from the images.

We will also preprocess a subset of the MACHO images using image-processing techniques such as wavelets for de-noising and feature detection to generate an initial set of features that identify an asteroid.

For our future work, we will refine our image processing techniques as well as the clustering and classification algorithms using these initial applications, implement parallel versions in an object oriented framework, and couple our algorithms to the research on error modeling techniques.