April 2011

Center for Adaptive Supercomputing Software

http://cass-mt.pnl.gov

CASS Research Areas

Architecture Studies

- Next generation hardware designs
- Next generation hybrid systems
- Emulating the XMT on commodity systems

System Software

- Compilers
- Runtime systems
- Programming tools
- Communication libraries

Algorithms

- Semantic databases
- Structural analysis
- Social networks
- Bayesian networks
- N-X contingency analysis
- Text analysis
- Clustering
- Real-time methods
- Dynamic data structures

N-x Contingency Analysis in Electric Power Grids:

What is the impact on a power grid when Generator X goes down? Or, Transmission-line Y goes down? These are the questions that concern power grid operators, and a systematic study of component failures is known as the Contingency Analysis. It is the focus of "Task 6" within CASS. Considering a single component-failure at a time is known as the "N-1" analysis, and studying multiple, say x, concurrent failures is known as the "N-x" analysis.

For decades, ideas and methods have moved from physics, biology and mathematics to social sciences. However, only recently, ideas and tools from social science are moving back to natural science – a flow that Linton Freeman calls "the Wrong Way on a One-Way street." The concept of betweenness-centrality is one such "Wrong Way" tool that has found applications in several areas outside of social science. CASS researchers embraced betweenness centrality to address the challenging problem of Contingency Analysis. The need for unconventional approaches to this problem is driven by the fact that conventional approaches tend to be computationally infeasible even for moderate sizes of power grids. Moreover, a power grid is a dynamic and complex system with a need to conduct contingency analysis after every few seconds of operation.

Networks are ubiquitous – power grids, transportation, information, human interaction, and chemical interaction of molecules, just to name a few. However, these networks could differ in their fundamental behavior. For example, in human interactions gossip can spread in a purely random fashion, but the laws of physics govern the flow of current in an electrical network. CASS researchers understood these limitations and adapted their tools accordingly. However, these changes were not sufficient. Working within a generic framework, they are now exploring the diverse fields of electrical engineering, graph theory, social network analysis, and spectral graph theory to build a comprehensive set of tools to tame the challenging "N-x" contingency problem. Within this context, they have developed a comprehensive framework not only to validate the results obtained from various tools, but also to select a particular tool based on input parameters such as computational resources available and desired accuracy. In addition, they are also developing advanced visualization techniques for the contingency problem.

Scalable Combinatorial Algorithms on the XMT:

Combinatorial (graph) algorithms play a critical enabling role in scientific and high performance computing. For example, graph coloring is used in compilers for register allocation and for efficient computation of sparse Jacobian and Hessian matrices in the context of automatic differentiation; graph matching is used in permutation and scaling in the solution of a system of linear equations and during coarsening phase of a multi-level algorithms used in many contexts.

Working with collaborators from Purdue University (Ariful Azad, Assefaw Gebremedhin and Alex Pothen), The Ohio State University (Ümit V. Çatalyürek), and Conviva Inc. (Florin Dobrian), researchers from CASS (John Feo, Mahantesh Halappanavar) have developed scalable implementations for distance-1 coloring, ½-approx weighted matching and maximum matching on the Cray XMT. Using a comprehensive set of synthetic graphs with up to a billion edges and novel algorithms based on dataflow principles, they showcased the power and efficacy of the XMT for executing graph algorithms. Recently, they were joined by two other CASS members, Oreste Villa and Antonino Tumeo, in implementing some of these key algorithms on emerging manycore architectures such as Nvidia's general purpose graphics processing units. They have demonstrated significant speedups on state-of-the-art multicore systems, GPUs, and the XMT. Their work has been submitted to leading journals for publication.

Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle *Since 1965*

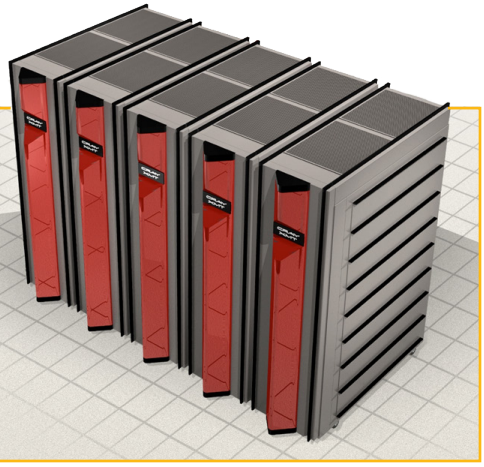
RECENT EVENTS

- PNNL's Mahantesh Halappanavar and John Feo, along with their collaborators Assefaw Gebremedhin and Alex Pothen of Purdue University and Umit Catalyurek of The Ohio State University, gave the invited talk "Multithreaded Algorithms for Graph Coloring" at the Society for Industrial and Applied Mathematics (SIAM) Conference on Computational Science and Engineering in Reno, Nevada in March. The talk discussed two kinds of multithreaded coloring algorithms, one for use on the Cray-XMT and one for general shared-memory architectures.
- John Feo and Jace Mogill, also of PNNL, gave an invited talk at the SIAM conference titled "System Requirements for Enabling Graph Algorithms on Massively Parallel Computers." The talk discussed the requirements for graph-based applications motivated by a clustering algorithm for generating Delaunay meshes.
- Mahantesh Halappanavar presented a joint work titled "Multithreaded Algorithms for Approximation and Exact Matching in Graphs" at the Workshop on Manycore and Accelerator-based High-performance Scientific Computing organized by the International Center for Computational Science in Berkeley, Calif. in January.
- PNNL's George Chin presented the paper "Fault Detection in Distributed Climate Sensor Networks using Dynamic Bayesian Networks" at the 2010 IEEE E-Science Conference in Brisbane, Australia in December. The paper described research conducted with a Cray XMT-based Bayesian network algorithm and its application to distributed climate sensor networks.
- David Bader and David Ediger of Georgia Tech presented the tutorial "Parallel Programming for Graph Analysis" at the ACM SIGPLAN Annual Symposium on Principles and Practice of Parallel Programming in San Antonio in February. The tutorial discussed the challenges in analyzing the "data deluge" of today's world and presented algorithms and data structures capable of analyzing spatio-temporal data at massive scale on parallel systems. Jason Riedy, also of Georgia Tech, also contributed to the development of the material.

UPCOMING EVENTS

- The workshop "Semantic Graph Database Search Patterns" will be held in Seattle April 25-26. The purpose of the workshop is to bring together technical leaders to develop a set of proposed abstract graph query patterns which when instantiated against a set of large triplebases produce compelling standard queries. The overall goal is to support future efforts in developing a robust capability to allow for the testing and comparison of both domain-dependent queries against real-world data as a function of relative size, and domain-independent queries against arbitrary SGDBs. The workshop will be held at Cray's Headquarters Building located at 901 5th Avenue in Seattle.
- The workshop "Scalable Graph Libraries" will be held in Atlanta June 29-30. The purpose of the workshop is to characterize the requirements of graph algorithm software packages by answering multiple questions and developing a white paper describing the requirements of graph libraries and programming frameworks, as well as the issues and challenges facing developers. The workshop will be help at Georgia Tech University in Atlanta.

John Feo,
Director of CASS
(509) 375-3768
John.feo@pnl.gov
cass-mt.pnl.gov/



Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle *Since 1965*

PNNL-SA-78980