

November 2010

Center for Adaptive Supercomputing Software

<http://cass-mt.pnl.gov>

PNNL Wins Second Place in 2010 "Billion Triple Challenge"

An entry from Pacific Northwest National Laboratory, Sandia National Laboratories, and Cray, Inc., received second place in the Billion Triple Challenge (BTC) at the 2010 International Semantic Web Conference (ISWC) held in Shanghai in November. The winning team was from the Hasso-Plattner Institute in Potsdam, Germany. The feedback from the jury panel was that the decision was difficult; both of the top two entries were very strong.

ISWC is the premier international forum for the presentation of research in "Semantic Web" technologies, which seek to bring intelligence and human accessibility of meaningful information to web-based databases and information systems. The BTC specifically is the leading test database against which international groups demonstrate capability and innovation for dealing with very large semantic databases. Participants are required to make use of this data set of 3.2 billion semantic database assertions (triples).

The PNNL/SNL/Cray entry "High Performance Semantic Factoring of Giga-Scale Semantic Graph Databases" took advantage of the Cray XMT architecture, which allowed them to hold all 624 GB of the input data in RAM. They were then able to scalably perform a variety of novel tasks for descriptive analysis of the inherent semantics in the BTC dataset, including the identification of its extant ontological structure, the sensitivity of its connectivity with respect to the semantic relations present, and the interaction among different organizations' contributions to the dataset. All of these analytical explorations were made possible by the unique strengths of the Cray XMT.

CASS Research Areas

Applications

- » Social Media Analysis
- » Contingency Analysis for the Electric Power Grid
- » Mesh Generation
- » Understanding Text Documents

Architectures

- » Architectural Studies on Multithreaded Architectures

Languages

- » Chapel for hybrid systems

Methods

- » Dynamic Network Analysis
- » Social Network Analysis
- » Semantic Database Systems

Runtime Systems

- » Compiler and Runtime System
- » Performance Analysis and Tools
- » Communication software for hybrid systems



Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by **Battelle** Since 1965



RECENT EVENTS

- John Feo participated in the Data Intensive Computing panel at the HPC User Forum “Using HPC to Drive Economic and Scientific Competitiveness” in Seattle in September. John presented the topic “Graph algorithms for irregular, unstructured data.”
- John Feo also provided an overview of the “Center for Adaptive Supercomputing Software” at the Gordon Grand Challenges in Data-Intensive Discovery Workshop at the University of California – San Diego, at the University of Maryland, and at George Mason University.
- Mahantesh Halappanavar gave the presentation “Massively Multithreaded Architectures and Algorithms for Graph Problems” at Old Dominion University in October.

- Alex Pothen from Purdue University presented CASS work on coloring algorithms at the 35th Woudschoten Conference in the Netherlands in October.

UPCOMING EVENTS

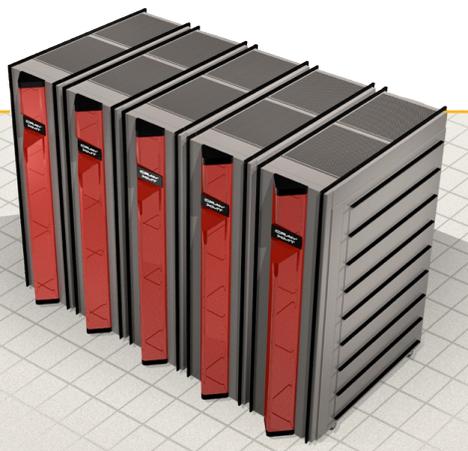
Supercomputing 2010 November 13-17 in New Orleans

- David Haglin, David Mizell, and Eric Goodman will lead a Birds of a Feather session at SC’10 titled “Semantic Graph/ Database Processing.” As data on the web becomes larger and more often available in Resource Description Framework (RDF) triples, it is becoming increasingly apparent that manipulation of these large knowledge bases requires high performance computing (HPC). This BOF aims to bring together researchers and practitioners from the two worlds of Semantic Databases (SDB) and HPC to discuss the common issues and how to use synergy moving forward.

- John Feo will lead a Birds of a Feather session on the “Cray XMT User Community.” This Birds-of-a-Feature will bring together the massive-multithreading community to discuss results, current status, and future machine designs driven by application requirements. The meeting will feature talks on application performance, demonstrations of programming tools, and position papers on expectations of next-generation multithreaded systems.



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