January 2010

Center for Adaptive Supercomputing Software

www.cass-mt/pnl.gov

CASS-MT Research Areas

Applications

- Social Media Analysis
- Contingency Analysis for the Electric Power Grid
- High-Throughput Video Analysis
- Understanding Text Documents

Architectures

- Architectural Studies on Multithreaded Architectures

Languages

- Chapel for hybrid systems

Methods

- Dynamic Network Analysis
- Social Network Analysis
- Irregular 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*

CONNECTING THE DOTS

John Feo, Director CASS-MT, and Peter Ungaro, Cray CEO, moderated the second annual Cray XMT Birds-of-a-Feather (BOF) session at the annual SC09 conference in Portland, Oregon. The session attracted more than double the number of participants that attended the first BOF at SC08, reflecting growing interest in the use of Cray's multithreaded architecture to tackle data intensive problems such as "connecting the dots" in the analysis of social networks.

The session included four presentations followed by a general discussion session.

All of the presentations can be found on the CASS-MT website: *http://cass-mt.pnl.gov/news.aspx#46*

System Architecture and Future Plans

Shoaib Mufti and Jim Harald from Cray Inc reviewed the current system architecture and software changes since the last BOF. There are now five Cray XMT sites with eight machines, and Cray has a 512 processor system running in house. Shoaib shared that Cray is designing Cray XMT-2 and plans to ship in 2010. This system will provide a memory subsystem that is several times larger and faster than the current system.

Multithreaded Graph Library (MTGL)

Greg Mackey from Sandia National Laboratory discussed the Multithreaded Graph Library (MTGL) that Sandia is developing for the Cray XMT. This library is a collection of commonly used graph routines written in C++. It is also available for download on the CASS-MT website. Greg explained that MTGL supports user defined graphs each with their own traits, API, and property maps. MTGL is simpler, but less generic, than the Boost Graph Library. He also demonstrated several real program examples and how they ran as fast as hand coded versions.

GraphCT – Analyzing large static graphs

David Bader from Georgia Tech took the floor next. There was general amusement among the audience when he began by showing his Facebook page. He quickly made the point by showing Facebook metrics that more than 300 million users were currently active, making Facebook an interesting and relevant social network upon which to base research. He described GraphCT, a tool to summarize and analyze large static graphs. The tool can compute a variety of graph statistics and characteristics, such as diameter, degree distribution, betweenness centrality, and st-connectivity. The tool's front end runs on any workstation, but the statistic and character kernels run on the XMT. They are highly optimized for maximum performance and scalability. Dr. Bader also described a generic graph data structure, STINGER, which he is developing to support dynamic changes to real-time graphs.

H1N1 – Theoretical modeling

Shahid Bokhari from Ohio State University gave an entertaining talk on the serious subject of the H1N1 virus. He presented an algorithm to compute a theoretical model of the virus' evolution over time. The algorithm builds a multipartite graph whose nodes are virus or reassortment events and whose edges are mutation or reassortment choices. Dr. Bokari pointed out that an early version of the code performed poorly due to a memory hotspot, but with help from Cray, he eliminated the hotspot and achieved linear scalability. Dr. Bokari's colleagues are using the code to predict the evolution of the H1N1 virus.

Following the general discussion, John Feo thanked Peter Ungaro, Cray's CEO, for hosting the BOF as part of Cray's whisper suites. Peter expressed his gratification at the turnout and was encouraged to see from the talks that real applications were starting to emerge for the XMT. The next BOF is tentatively planned for the IEEE International Parallel and Distributed Processing Symposium in Atlanta in April

If you have an interest in attending the next BOF or receiving further CASS-MT newsletters and information, please contact John Feo at *john.feo@pnl.gov*.

CASS-MT is dedicated to research on systems software, programming environments, and applications in a High-Performance Computing (HPC) multithreaded architecture environment.

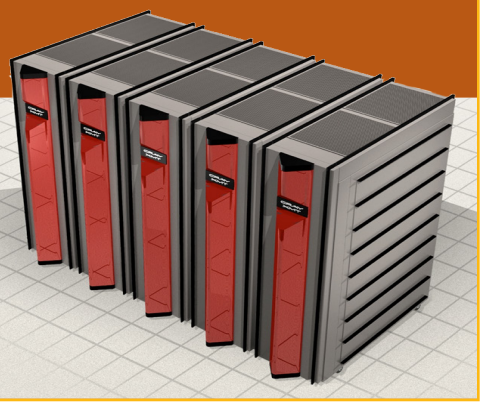
We offer the only Open Science Cray XMT system, a one-of-a-kind supercomputer consisting of 128 multithreaded processors, 1 TB RAM, and a 7.7 TB Lustre parallel filesystem.

The Cray XMT supercomputer has the potential to substantially accelerate data analysis and predictive analytics beyond the limitations of traditional computing. Multithreaded processors allow multiple, simultaneous processing, helping researchers find solutions to the world's most complex challenges faster. The XMT can process irregular, data-intensive applications that have random memory access patterns. Unlike many applications where data delivery is dependent on memory speed, the Cray XMT's multi-threaded architecture tolerates memory access latencies by switching context between multiple threads that work continuously, overlapping the memory latency and preventing the processor from being held up while it waits for data to arrive.

The multithreaded technology powering our Cray XMT is ideally suited to perform pattern matching, scenario development, behavioral prediction, anomaly identification, and graph analysis.

Try it for yourself. We seek to create collaborations and provide expertise for porting and optimizing applications. The opportunity to use our Cray XMT system is available to internal and external research partners.

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



Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle *Since 1965*