XSCI

eXtreme Scale Computing Initiative

Initiative Leadership

T.P. Straatsma
Initiative Lead
(509) 375-2802
tps@pnl.gov

Darren Kerbyson
Initiative Chief Scientist
(509) 375-2603
darren.kerbyson@pnl.gov

Upcoming Events

Date: May 16-20

Event: IEEE International Parallel & Distributed Processing Symposium 2011

Location: Anchorage, Alaska

Date: May 17

Event: Global Arrays Tutorial: Parallel Programming Using the Global Arrays Toolkit: Now and into the Future.

Location: Anchorage, Alaska

Date: May 27

Time: 12-1 pm

Event: Brown bag: Exploring Architectures Suitable for Scientific Applications at Exascale Levels

Location: BSF/Darwin room

Pacific Northwest
NATIONAL LABORATORY

Proudly Operated by Battelle *Since 1965*

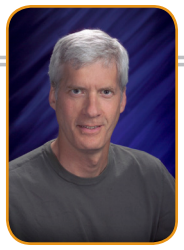
Leadership Notes:

In the second of our new series of brown bags on the Initiative, Sriram Krishnamoorthy held a general brown bag on the accomplishments on his project, which includes the scalable trace compression algorithms he has developed. In addition, he presented a well-attended special tutorial on how to get started in the use of the GA/ARMCI performance diagnostic capabilities as they are currently available through the Tuning and Analysis Utilities in TAU. Another highlight from the XSCI projects is that Karol Kowalski achieved a new scalability milestone of his EOM-CCSD(T) triples code, which now scales to more than 200,000 cores on Jaguar, the peta-scale machine at OLCF.

The upcoming IEEE International Parallel and Distributed Processing Symposium, which will be held May 16-20, 2011, in Anchorage, AK, will feature a number of XSCI related activities. A tutorial by Bruce Palmer, Manoj Krishnan and Abhinav Vishnu on "Parallel Programming Using the Global Arrays Toolkit: Now and Into the Future" will provide an overview of the Global Arrays (GA) programming toolkit and describe its capabilities, performance, and the use of GA in high-performance computing applications. Darren Kerbyson is the Co-Chair for the "Workshop on Large-Scale Parallel Processing," and Adolfy Hoisie, Darren Kerbyson and TP Straatsma will organize a meeting to discuss "Future Approaches to Data Centric Programming for Exascale."

In an effort to provide GA capabilities for applications in disciplines other than molecular and subsurface science, a new activity has started to identify capability requirements for Lattice Boltzmann codes as well as sparse linear algebra kernels for PDE solvers, and to implement these capabilities into the GA toolkit. This activity will allow extending the use of GA into new application areas.

New Projects



» Developing Functionality and Performance Enhancements to the Global Array Toolkit, led by Bruce Palmer

Meet the Team

T.P. Straatsma, Molecular Dynamics
Abhinav Vishnu, Scalable Fault Tolerance
Steven Yabusaki, Subsurface
Manoj Krishnan, Data Decomposition
Karol Kowalski, Molecular Modeling
Oreste Villa, Architecture Exploration
Daniel Chavarria, Performance Optimizations
Sriram Krishnamoorthy, Performance Diagnostics
Kevin Barker, Power and Performance Modeling
Huub Van Dam, Molecular Modeling
Bruce Palmer, Functionality & Performance Enhancements

Feature of the Quarter

Advanced Scalability for STOMP: Subsurface Simulation and Characterization at Extreme Resolution

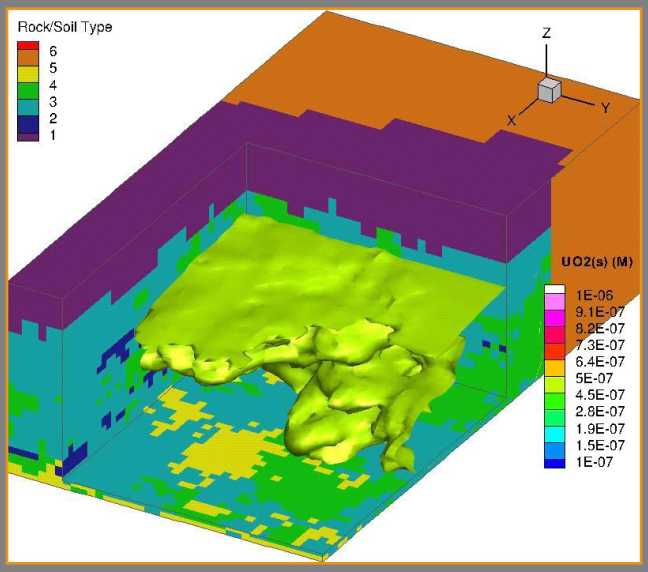
A highly scalable, next-generation version of PNNL's subsurface simulator, STOMP, is currently being developed by PNNL researchers to address detailed subsurface processes, properties, and conditions at extreme-scale resolutions. This capability is needed to solve diverse problems in DOE mission-critical areas, including the accurate assessment of risks associated with nuclear waste management, effective remediation design, geologic carbon sequestration, and enhanced hydrocarbon recovery. To date, this code, referred to as eSTOMP, has been successfully tested on as many as 131,040 processor cores. Re-engineering eSTOMP with the PNNL-developed GA Toolkit provided an efficient parallel implementation based on one-sided communication and a shared memory programming interface. The GA toolkit was used to implement a distributed, unstructured data model that allowed the reuse of a large fraction of the original source code. Those familiar with the original STOMP code and input will find eSTOMP easy to run and enhance code capabilities without having to learn new languages, data structures, or programming styles.

Demonstrated applications include uranium biogeochemistry and transport at two of the three U.S. Department of Energy (DOE) Office of Science Integrated Field Research Challenge (IFRC) sites (Hanford 300 Area and Rifle, CO) and supercritical CO_2 injection into a heterogeneous sandstone formation in the Illinois Basin for deep geologic carbon sequestration. An uncertainty quantification application performed 317 simulations of 3-D, transient, variably saturated flow in the 300 Area vadose zone-aquifer system driven by hourly Columbia River stage fluctuations in 23 hours, whereas a comparable single simulation using a desktop computer takes one week.

Recently, the eSTOMP code was one of 4 codes selected by the DOE Office of Advanced Scientific Computing Research to participate in the Joule Metric for Computational Effectiveness. For this activity, PNNL was awarded 25 million node-hours on Jaguar, the largest computer in the DOE Leadership Computing Facility at Oak Ridge National Laboratory. The targeted application is a field-scale simulation of engineered bioremediation in a groundwater plume at the Rifle IFRC site where acetate is injected to stimulate metal-reducing bacteria to immobilize uranium. The 3-D simulations will account for impacts on uranium behavior due to water table change and spatially variable physical and reactive material properties. While there are over 100 biological and geochemical species included in the reaction network, the reaction model will test a novel genome-scale metabolic ("*in silico*") model of *Geobacter sulfurreducens* that accounts for over 500 intracellular reactions by itself.

Major Accomplishments:

- Current capabilities address variably saturated flow, transport, and reactions and nonisothermal, multiphase flow with supercritical CO_2.
- The new code, re-engineered with PNNL's GA Toolkit and a new grid component, has demonstrated very good scalability. In particular, the eSTOMP routines, excluding third-party libraries, have exhibited linear speedup for large processor core counts.
- Componentization of eSTOMP has provided for significant reuse of the original STOMP source code, simpler code maintenance, and easier extension of eSTOMP capabilities.
- Saturated flow module demonstrated at 131,040 cores.



Zone of bioreduced uranium after 10 days of acetate injection. Spatially variable material properties and a declining water table result in preferential transport paths and reactions.

The XSCI is a quarterly publication covering research being conducted in PNNL's eXtreme Scale Computing Initiative. For more information, contact the initiative leads or visit our web page at *http://xsci.pnl.gov* .

Operational Team

Monica Moffett, Administrator

Rebecca Wattenburger, Finance

Jessica Garza, Communications

Website

http://xsci.pnl.gov