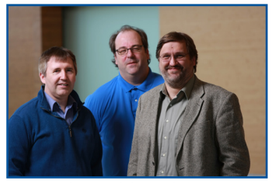
Development of Exascale Algorithms for Molecular Modeling

Lead investigator: Dr. Karol Kowalski


Dr. Karol Kowalski (right)

Researchers are developing a new generation of algorithms for molecular modeling, which will make effective use of emerging petascale and extreme-scale architectures. Key advances in molecular simulationas are being pursued in several areas—coupled-cluster algorithms describing excited-state correlation effects, algorithms for parallel-in-time dynamical calculations, and multiscale approaches for strongly interacting systems.

Major Accomplishments

» Demonstrated that the CCSD algorithms can scale across 34,000 CPUs.

Single Node Performance Optimizations for Extreme Scale Systems

Lead investigator: Dr. Daniel Chavarria

Dr. Daniel Chavarria (right)

Exascale systems will provide an unprecedented opportunity for researchers to attack larger and more complex problems, and reach solutions dramatically faster. Researchers are designing and implementing techniques to enhance GA capabilities to more efficiently use intra-node resources, including multi-core processors, NUMA shared memory, and other emerging node architectural features.

Goals:

» Development of new optimization methods and techniques for computation on emerging extreme-scale nodes.

» Demonstration at scale using a molecular dynamics driver application.

Visit our Web site at: **http://xsci.pnl.gov.**

Mapping Molecular Dynamics Algorithmic Parallelism to Heterogeneous Architectures

Lead investigator: Dr. TP Straatsma

A principal and general challenge for biomolecular modeling and simulation in support of environmental molecular science involves the generation of extended molecular trajectories using advanced simulation methodologies to capture the features central to the research questions being addressed. Using performance modeling and analysis, researchers are working to design new computer science and software engineering approaches to develop a high-performance molecular dynamics modeling and simulation implementation for biomolecular modeling research.

Goals:

» Deliver atomistic molecular dynamics simulation algorithm implementations providing the capability to simulate 250,000 to 1,000,000 atom chemical and biomolecular systems at 50+ ns/day.

» Effective scalability of large-scale molecular dynamics down to 10^4 pair interactions per core, which is 10 atoms per core.

About PNNL

Pacific Northwest National Laboratory is a Department of Energy Office of Science national laboratory where interdisciplinary teams advance science and technology and deliver solutions to America's most intractable problems in energy, the environment and national security. PNNL employs 4,900 staff, has an annual budget of nearly $1.1 billion, and has been managed by Ohio-based Battelle since the lab's inception in 1965.

FOR MORE INFORMATION, CONTACT:

TP Straatsma, Initiative Lead
509-375-2802
tps@pnl.gov
www.pnl.gov

PNNL-SA-73923

eXtreme Scale Computing Initiative

About the Initiative

Pacific Northwest National Laboratory (PNNL) is at the forefront of the development of next-generation modeling and simulation capabilities to address major scientific challenges in energy, climate, and the environment. The eXtreme Scale Computing Initiative (XSCI) is a co-design effort in computational domain science, high-performance computer science, applied mathematics, and performance modeling to build the capabilities that allow these modeling and simulation codes to perform at the extreme scale.

Approach

Immense computational resources are required to solve a range of problems our society is facing. We must enhance these codes by addressing bottlenecks in scalability, algorithm accuracy, and fault tolerance.

Initiative Goals

» Emphasize environmental subsurface and molecular modeling as scientific domain areas in which PNNL has a recognized leadership position.

» Develop methods, algorithms, and software for use in computers that are three orders of magnitude more capable than any systems available today.

» Help emerging application codes by providing software libraries and algorithmic capabilities that are critical for use on next-generation hardware.

» Advance scalable solvers, scalable I/O, dynamic load balancing, and adaptive grid-generation technologies.

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

Initiative Projects

A Scalable Fault-Tolerance Infrastructure and Algorithms with Programming Models and Scientific Applications

Lead investigator: Dr. Abhinav Vishnu

Dr. Abhinav Vishnu (left)

Researchers are designing and implementing a fault-tolerance framework for scientific applications and programming models. The project is targeting computationally intensive PNNL applications, with a focus on fault tolerance with low overhead using a Computational Chemistry Application, Global Address Space Programming Model, such as Global Arrays (GA), and a communication runtime system, such as Aggregate Remote Memory Copy Interface (ARMCI).

Major Accomplishments

» Successfully demonstrated the fault-tolerant capabilities of the overall framework using simple tests at different levels of the system software stack.

» Designed and developed a fault-tolerant version of the computational chemistry algorithm using a coupled cluster method as an example.

Scalable Performance Diagnostics for Massively Parallel Computers

Lead investigator: Dr. Sriram Krishnamoorthy

Dr. Sriram Krishnamoorthy (right)

Tools and techniques are being developed to identify performance inhibitors in applications that will aid in identifying performance bottlenecks within massively parallel systems. A scalable communication trace compression scheme, and a customized profiling framework are being developed to simplify the performance bottleneck discovery process. Researchers are pursuing multiple research directions to tackle challenges at scale, integrate support for ARMCI and GA in widely used tools, and evaluate applications.

Major Accomplishments

» Development of a scalable communication trace compression scheme.

Data Decomposition/Optimizations and Dynamic Load-Balancing Mechanisms for Extreme-Scale Computing in the Global Arrays Toolkit

Lead investigator: Dr. Manoj Krishnan

Dr. Manoj Krishnan (left)

Researchers are developing capabilities in the GA Toolkit to support flexible data decomposition, communication optimization, and dynamic load-balancing approaches for communication-intensive molecular simulations and subsurface modeling applications.

Goals:

» Implement flexible and scalable data decomposition methods.

» Provide dynamic load-balancing capabilities.

» Support fine-grained parallelism on multi-core architectures.

» Remove non-scaling data overheads in the GA libraries.

» Optimize nearest-neighbor communication.

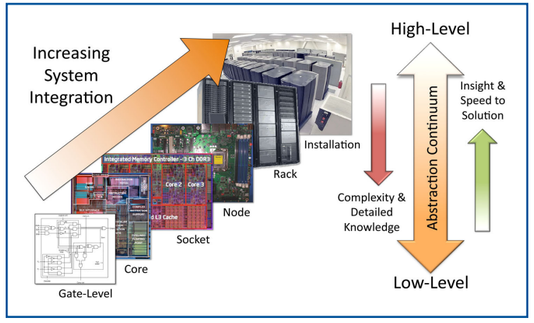
Integrating Power and Performance Modeling for Exascale Systems

Lead investigator: Dr. Kevin Barker

Future extreme-scale systems must be designed with not only performance criteria but also power consumption in mind. This project will develop quasi-analytical power modeling techniques to allow for rapid architectural space exploration.

Open Issues

1. Determining appropriate level of abstraction.
2. Integrating Performance Modeling methods methods to enable quantitative architectural explorations.
3. Obtaining Validation and Model Input data from existing systems.



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

Lead investigator: Steve Yabusaki

Steve Yabusaki (center)

Researchers are developing a highly scalable, next-generation version of PNNL's subsurface simulator, STOMP. The new implementation is able to deliver performance and still allow those familiar with the serial STOMP code to easily modify and enhance code capabilities without learning new languages, data structures, or programming styles. The project is focused on the re-engineering of STOMP operational modes, which address DOE missions in environmental restoration of nuclear production legacy waste sites, geologic sequestration of CO_2, and the development of alternative energy resources.

Major Accomplishments

» The STOMP routines, excluding third-party libraries, have exhibited linear speedup for large processor core counts.

» The new code, re-engineered with PNNL's GA Toolkit and a new grid component, has demonstrated very good scalability.

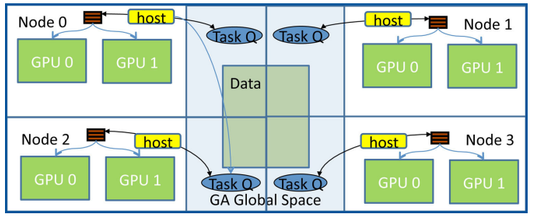
» Scalability of the saturated flow module demonstrated at 16,000 cores.

Exploring Architectures Suitable for Scientific Applications at the Extreme Scale

Lead investigator: Dr. Oreste Villa

Next-generation computer architectures have the potential to be part of future exascale-capable supercomputers. Among them, Graphic Processing Units (GPUs) are gaining attention due to the low cost and high-performance-per-watt ratio. The gap between the programmability of conventional architectures (CPUs) and GPUs is still large, making the design of complex scientific applications a hard problem that requires heavy modification to the existing implementations. This project is targeting PNNL applications, focusing on performance, programmability, and productivity, primarily on GPU-based clusters.

Major Accomplishments

» Development of a task-based execution framework that enables load balancing among multi-GPU systems.

» Acceleration by a factor of 10x for the quantum chemistry application CCSD(T) using a GPU-based cluster.