

eXtreme Scale Computing



Pacific Northwest
NATIONAL LABORATORY

*Proudly Operated by **Battelle** Since 1965*

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. PNNL is located in Richland, WA.



Technical Challenges of Exascale Computing. The core mission of the eXtreme Scale Computing Initiative (XSCI) is to develop methods, algorithms, and software that can be efficiently used on the next generation of supercomputers with 1000-fold increase in computational power compared to what is available today. The primary driver for PNNL to invest in high-performance computing is to solve problems our society is facing in areas such as energy, the environment, and national security. Major advancements are needed for PNNL to maintain its existing leadership and establish new capabilities in specific computational science applications areas, to enable the use of the next-generation massively parallel computer hardware, and to take advantage of upcoming new funding opportunities in scientific and engineering modeling and simulation.

The initiative is developing and demonstrating solutions that address the following technical challenges for any modeling and simulation capability at extreme scale:

- » **Scalability**– The next-generation computers will be truly massively parallel, with millions of processing threads simultaneously working on a single computational problem. Algorithms need to be redesigned to optimally lay out the data, avoid collective operations, and provide effective load balancing. We are developing new capabilities with our Global Arrays programming model for domain science application programmers to effectively design and implement their computational-intensive kernels to scale to the largest possible number of processing cores.
- » **Architectures**– The supercomputers of the future are expected to rely on a hierarchy of heterogeneous architectures, which will include multiple many-core sockets on a node augmented with accelerators. The effective programming of such hardware is a complex task, and the Initiative is providing the support to applications to appropriately map the algorithmic concurrencies onto the hardware components best suited for the task to be performed.
- » **Fault Tolerance**– The sheer number of electronic components on large-scale machines will lead to a strong reduction in the mean time between failures. The result is that applications are almost guaranteed to encounter hardware failures during a calculation. The checkpoint-restart approach of the past will no longer be reasonable for many applications. The Initiative is designing fault-tolerance capabilities within the Global Arrays framework to deal with faults during a calculation without the need to abort and restart the job, through data replication and new data management processes.
- » **Productivity**– Effective program design, implementation, and debugging is becoming a prohibitively complex process. Existing methods to evaluate parallel performance and identify bottlenecks to scalability will not be suitable for the large processor count machines of the future. The Initiative is developing scalable trace compression algorithms to manage the amount of trace data and will provide analysis and visualization tools for the application programmer to understand the performance behavior of his code and to help identify opportunities for performance improvement.
- » **Power**– The exascale computer of the future will have extreme power requirements, making the operation of such machines costly. While in the past achieving floating point performance was the metric of a well-written code, in the future achieving power efficiency may well become the metric. The Initiative is developing new capabilities for integrated performance and power modeling.

The development of high-performance computers, the networks to connect them, and the scalable mathematics and software to run them are crucial for science-based solutions in energy, the environment, and national security. XSCI at PNNL helps make this happen through innovative research to discover the mathematical algorithms and methods and the computer science and computational tools that enable researchers in many scientific disciplines to analyze, model, simulate, and predict complex phenomena.

The XSCI is building the capabilities needed to enable scientific advancements and breakthroughs in selected domain sciences through computational modeling and simulation on next-generation, extreme-scale computers. The Initiative consists of an integrated research program with an interdisciplinary approach that brings together high-performance computer science, applied mathematics, and computational domain science to develop next-generation, extreme-scale modeling, and simulation applications.

The research plan for this Initiative is focused on scalability, methodology accuracy, and fault resiliency through design, implementation, and evaluation of new algorithms for computer systems with capabilities beyond what is currently available. An important component is the evaluation of computational kernels on the very high-end computer resources currently available, and the development of performance prediction models for next-generation, extreme-scale systems.

The objective of the XSCI is to integrate our computational sciences, computer science, applied mathematics, and scientific data management expertise to enable extreme-scale applications to address the major scientific challenges in energy, the environment, and national security. The Initiative is emphasizing environmental subsurface and molecular modeling as scientific domain areas in which PNNL has a recognized leadership position.

High-performance computing. Our approach merges science and technology by employing hardware that maximizes processor speed, memory and interconnect bandwidth, efficient use of secondary storage, and reliability; developing algorithms that are scalable, resource-efficient, load-balanced, and manage computational complexity and exploits space-time locality; and creating programming models, numerical libraries, communication libraries, compilers, and debuggers that support data decomposition, low communication overhead, and portability. One of our award-winning programming models, Global Arrays, enables researchers to more efficiently access global data, run bigger models, and simulate larger systems, resulting in a better understanding of the data and process being evaluated.

Computer science. We focus on two main areas, Programming Models and Performance Analysis and Modeling. Programming Models

will allow us to develop software infrastructure for parallel applications to interface efficient, massively parallel hardware while taking into consideration the architectural characteristics and application requirements. The four elements and planned activities are:

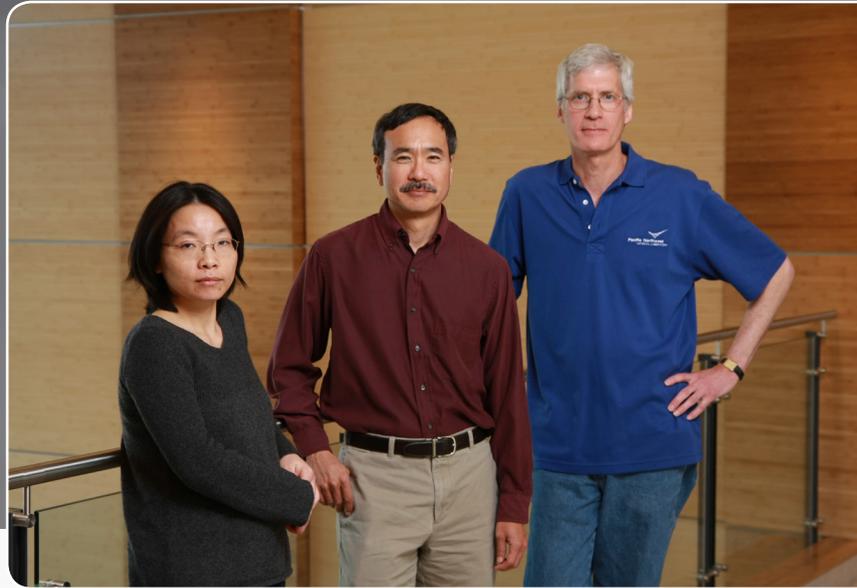
- » **Communication** - Tools need to be developed to diagnose specific bottlenecks in inter-processor communication, and optimized algorithms are being designed for time-stepping algorithms with heavy beyond-nearest-neighbor communication.
- » **Distributed Data Structures** - Certain applications will benefit from scalable grid generation packages (developed in-house or customized for our problem areas). Data structures optimized for locality and efficient computation on multi-core processors, and direct data access integration with solvers can boost performance.
- » **Domain-Specific Tools** - Parallel programming frameworks are being designed that hide architectural complexities and increase programmer productivity, and include libraries for communication, and data transformations optimized for scalability and performance.
- » **Input/Output and Storage** - High-performance I/O libraries are needed to address application needs, Active Storage methods are being investigated for reducing data movement across networks, and tools are being designed for scalable metadata management.



XSCI is conducting research using Chinook, a 163-teraflop supercomputer at the Environmental Molecular Sciences Laboratory, a U.S. Department of Energy (DOE) national scientific user facility located at PNNL.

STOMP

Subsurface Transport Over Multiple Phases, or STOMP, is a software tool that provides multi-dimensional modeling of subsurface flow and reactive transport phenomena. STOMP can accurately track and predict the subsurface migration of specific materials over time, and over multiple phases, for many applications.



Performance modeling and analysis. Both will be applied to the developed domain applications and are critical to the success of the initiative in the absence of next-generation, massively parallel systems. The overall approach is to develop analytical performance models that will be validated based on experimental data obtained on the existing hardware, followed by extrapolation to include characteristics of the upcoming massively parallel architectures. The three elements and their activities are:

- » **Architecture Characteristics** - Extrapolation approaches are being developed based on architectural trends for network performance (latency, bandwidth) relative to the processor performance, analysis of network topologies, and expected performance of future many-core processors (byte/flop/op ratios).
- » **Integrated Models** - Application models are being mapped onto architectural characteristics, with validation of performance and scalability models on the available architectures, and identification of performance and scalability bottlenecks.

- » **Modeling of Application Performance** - This activity is necessary to develop performance assessment methodology for existing codes, and to build performance and scalability models for applications and computational kernels.

Computational Science. We focus on two computational domain application areas for initial scalability development, Molecular and Subsurface Modeling.

- » **Molecular Modeling** - A wide range of molecular processes require highly diversified methodologies to describe systems characterized by different time and spatial scales. This problem in a vast majority of situations entails a need for having multi-scale methods that are built upon more basic fundamental solvers, whose scalability and accuracy define the final performance of combined approaches. Given the unprecedented increases in computer technology over the last decade, we expect that exascale computing, through the design



Extreme-scale computing is essential to enable breakthroughs in areas such as accurate weather and climate prediction, biotechnology, and new energy sources.

of innovative computational techniques for incorporating phenomena at the fundamental time, space and complexity scales are needed to make transformational breakthroughs and scientific discoveries. Specific goals for this area include providing the ability for:

- Modeling a wide range of molecular processes including excited states at high accuracy
 - Parallel-in-time methodology for molecular modeling
 - Simulating biomolecular systems at biologically relevant length and time scales
 - Incorporating phenomena at fundamental time and spatial scales into models that are relevant up to macro scales.
- » **Subsurface Modeling** - A reliable means of quantitatively predicting the movement of fluids and the fate and transport of contaminants in the subsurface environment is of significant value to the DOE and to our nation as a whole. DOE's Hanford Site is one of the most contaminated sites in the world and requires particular attention in addressing predictability of the transport of contaminants for supporting scientifically defensible

environmental management strategies. This capability also is needed for diverse problems including enhanced recovery of hydrocarbons, geologic sequestration of carbon, effective design of subsurface remediation, and accurate assessment of risks associated with subsurface disposal of nuclear wastes. While major advances have already been made in our ability to numerically simulate complex hydrologic, geochemical, and microbiological processes of interest, accurate and detailed characterization of spatial property distributions in the subsurface remains the

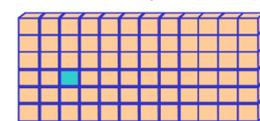
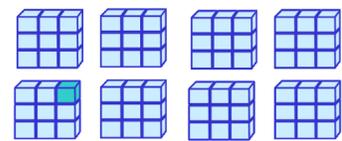
primary challenge facing reactive transport modeling efforts. Specific goals for this are:

- Inverse modeling for systems with billions of unknowns
- Uncertainty quantification and propagation
- Highly scalable forward modeling code on more than 100,000 cores
- Parallel-in-time codes for reservoir and geochemical modeling
- Coupled multi-scale methods for surface chemistry
- Multi-phase, pore-scale simulators coupled to physical/chemical processes.

GLOBAL ARRAYS TOOLKIT

Global Arrays has been designed to dramatically simplify writing code for supercomputers. The toolkit helps scientists to translate their ideas into highly efficient software that allows mathematical computations to run independently using subsets of processors of the supercomputer. The programmer is free to use both the shared-memory and message-passing paradigms in the same program, and to take advantage of existing message-passing software libraries. The Global Arrays toolkit has been in the public domain since 1994. It has been actively supported and employed in several large codes since then.

Physically distributed data



Global Address Space



The High-Performance Computing System-3, Chinook, is a balanced supercomputer that has been tailored to meet the current and future operational needs of users. Chinook is located in EMSL on the PNNL campus.

Power and performance modeling.

Supercomputer systems have primarily been designed with performance as the primary criterion that must be satisfied. However, as we move into the exascale era, power consumption will increasingly provide the pressure that will drive system design. Simply scaling up current technology will not provide an acceptable solution; extrapolating from today's largest machines indicates that on the order of 10GW of power will be required for an exascale system of similar design. New approaches and technology will be required to achieve such performance within an acceptable power budget.

XSCI is an integrated performance and power-modeling methodology. Such a tool will allow researchers to quantitatively analyze the trade-off between workload performance and power consumption on current and future architectures extending into the exascale regime. Within the scope of this work, researchers are developing a power-modeling methodology analogous to existing and proven performance modeling techniques. This power-modeling methodology will then be integrated with performance modeling to allow rapid study of architecture and workload designs.

EXCITED-STATE, COUPLED-CLUSTER CALCULATIONS

Researchers are developing software for the description and prediction of the structure and dynamics of molecules in excited electronic states. The accurate description of such systems allows for the understanding and design of molecular systems that harvest light and convert captured photons into chemical or electrical energy. Due to the importance of these processes, the ability is required to routinely carry out excited-state calculations of systems with hundreds of atoms. One of the goals of XSCI is to develop electronic structure software using coupled-cluster methodologies for excited states, which will scale to exascale computational resources. This is accomplished by redesigning and optimizing local memory management, employing a novel global addressing strategy within Global Arrays to deal with the large intermediate data, and restructuring of the coupled-cluster equations to allow better load balancing in a task-based execution model.

The new tools being developed for performing excited-state exascale calculations are based on the development of active-space equation-of-motion coupled-cluster corrections. The principal idea behind this development is the merging of the iterative approaches, which represents the Schrödinger equation in a small space providing a reliable description of the excited state of interest, with more expensive methods capable of restoring missing correlation effects in a non-iterative way. In this approach the reduced iterative-core approach is represented by the EOMCCSd method, i.e., EOMCC with all singles and a selected set of double excitations defined by the amplitudes with active spinorbital indices only. In addition, a nested variant of the EOMCCSD(T) approach is being implemented, which provides a very balanced description of excited and ground-state correlation effects. This results in a very accurate description of excitation energies.

Classical Molecular Dynamics Simulations

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 essential features. This applies to the challenge of understanding protein folding, protein docking, complex enzymatic reactions, and the association and function of large protein and protein-DNA complexes. The common challenge in all of these cases is the need for sufficiently large ensembles of conformations to capture the relevant events from non-deterministic Monte Carlo simulations with broad sampling of the available configuration space, or from extended deterministic molecular dynamics simulations, which also provides the kinetics of events.

A new high-performance molecular dynamics software implementation for biomolecular modeling and simulation is being developed, which is aware and takes advantage of the specific characteristics of the underlying architecture to achieve scalable performance on the very large processor counts that can be expected to exist in the next-generation extreme-scale supercomputers. The new implementation strategies will focus on latency hiding, reducing communication, data structure optimizations, and single-node performance optimization.

The Initiative is developing core computer science capabilities as well as new domain science methods, algorithms, and software that can effectively be used on the computing resources that will soon be three orders of magnitude more capable than those available today. By successfully addressing the key issues affecting extreme-scale computing, this will establish PNNL among the leaders in large-scale computational applications required to

address and help solve our society's problems in areas such as energy, the environment, and national security. The fundamental capabilities developed by the Initiative will be able to support improving scalability and fault tolerance in applications beyond the molecular and subsurface science applications that are specifically targeted in the Initiative, and for which PNNL already holds a leadership position.

For more information about the **eXtreme
Scale Computing Initiative**, contact:

TP Straatsma
Pacific Northwest National Laboratory
(509) 372-4720
tps@pnl.gov

www.pnl.gov
http://xsci.pnl.gov/



*Proudly Operated by **Battelle** Since 1965*



U.S. DEPARTMENT OF
ENERGY