EMSL

Scientific Innovation Through Integration

NWChem - Computational Chemistry at Scale

Bert de Jong
bert.dejong@pnl.gov

Pacific Northwest NATIONAL LABORATORY

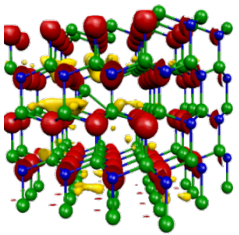
Proudly Operated by Battelle Since 1965

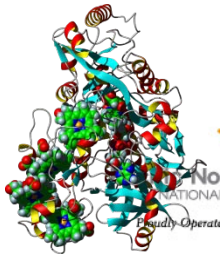
U.S. DEPARTMENT OF ENERGY

Delivering Computational Chemistry

EMSL

- DOE's Premier Computational Chemistry Software
- Provides one-of-a-kind computational chemistry modeling solutions for molecular science that are scalable with respect to scientific challenge and hardware
- Provides major integrated modeling and simulation capability
  - Broad range of molecules, including biomolecules, nanoparticles and heavy elements
  - Electronic structure of molecules
  - Extensive solid state capability (DFT plane-wave, CPMD)
  - Molecular dynamics, molecular mechanics





Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

U.S. DEPARTMENT OF ENERGY

Computational Chemistry at Scale

EMSL

- Computational chemistry solutions scalable with chemical system and computing capability
- Massively parallel computational chemistry software designed to take advantage of large computing resources
- Scalability demonstrated to tens of thousands of processors
- World-wide distribution (>2400 groups and computing centers)

MS^3
MOLECULAR SCIENCE SOFTWARE SUITE

NWCHEM
HIGH-PERFORMANCE COMPUTATIONAL CHEMISTRY SOFTWARE

GA TOOLS
PARALLEL COMPUTING LIBRARIES AND SOFTWARE TOOLS

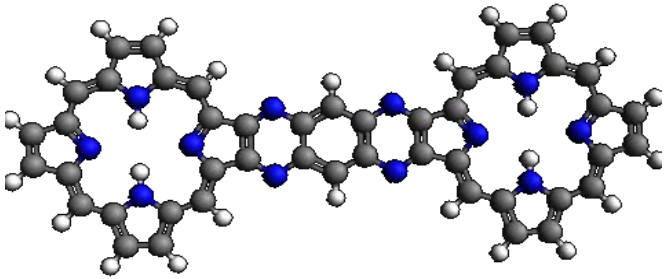
ECCE
EXTENSIBLE COMPUTATIONAL CHEMISTRY ENVIRONMENT

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

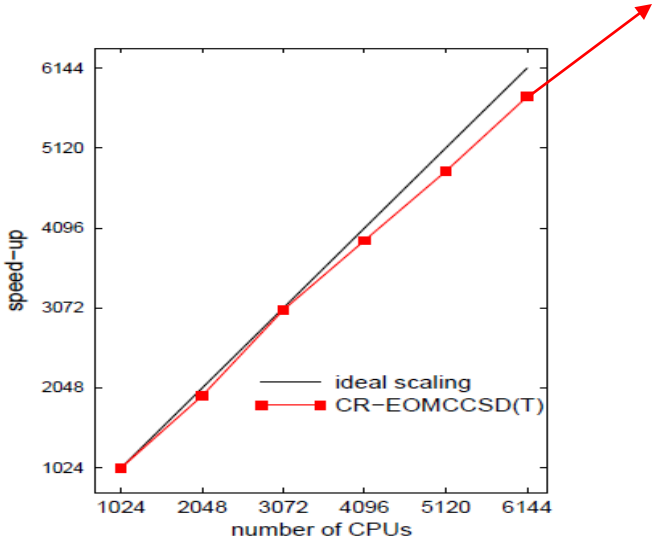
U.S. DEPARTMENT OF ENERGY

New Science with High Accuracy Methods

- Designing light harvesting systems and molecular switches requires a highly accurate and comprehensive understanding of the behavior of electrons in the system.



P_2TA molecule forms a basis for molecular switches and wires in nanoelectronics

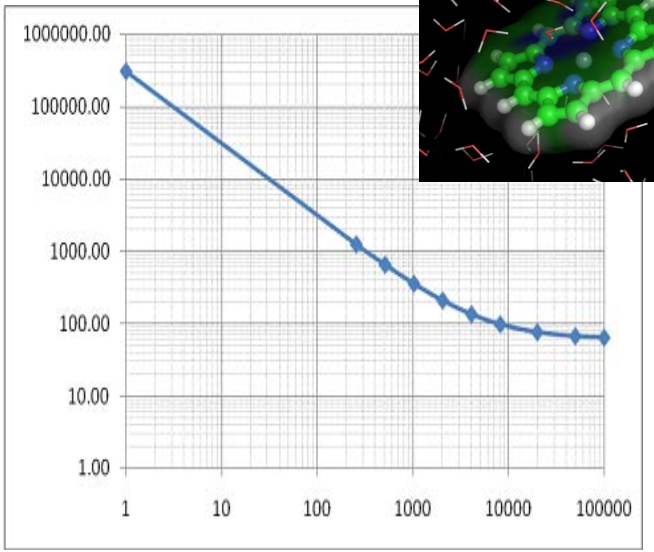


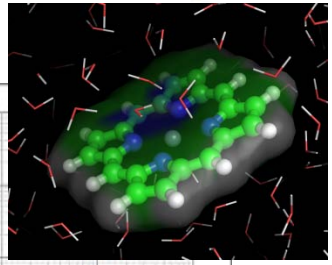
Ideal scaling of high accuracy CR-EOMCCSD(T) tripes for the P_2TA molecule

High Accuracy to the Petascale

EMSL

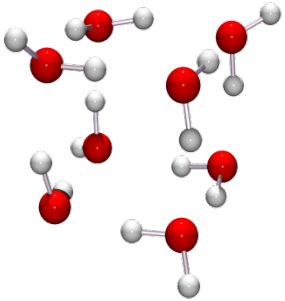
- Improving the parallel efficiency for coupled cluster algorithms
  - Balancing communication and operations
  - Better task scheduling based on knowledge of computational costs and overhead
  - Heterogeneous computing using processor groups
  - Improving data-localization reduces communication





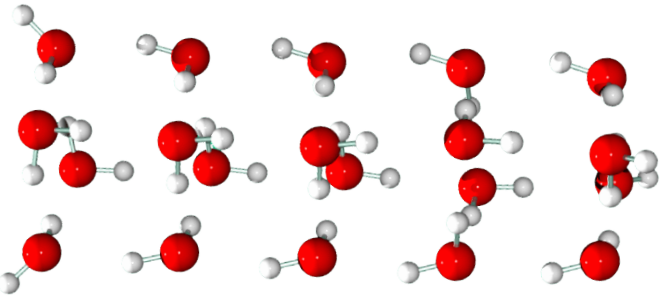
Performance model for the light harvesting Zn-porphyrine molecule

NWChem - Gordon Bell Finalist



2007

2009



8 water molecules on EMSL's MPP2 achieve 7 TFlops (63% of peak) on 1,840 processors

24 water molecules on ORNL's Jaguar achieve 1.31 PFlops (57% of peak) on 224,196 processors

ORNL – PNNL Collaboration

Improving our understanding of water's properties

Pacific Northwest NATIONAL LABORATORY

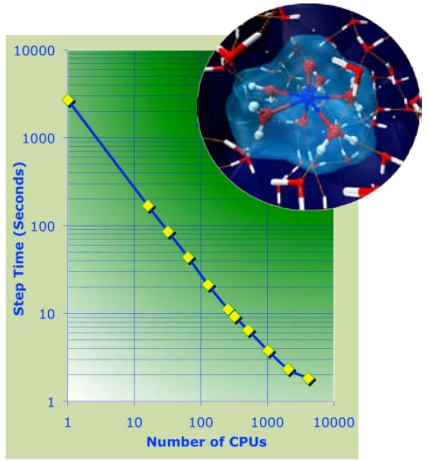
Proudly Operated by Battelle Since 1965

U.S. DEPARTMENT OF ENERGY

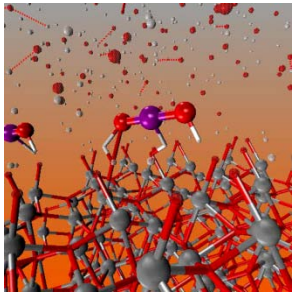
Chemistry at Interfaces Needs Petascale

EMSL

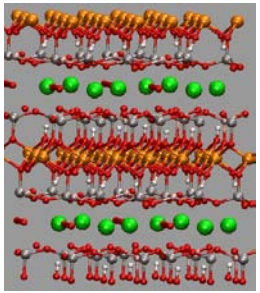
- Catalysis, hydrogen production, environmental remediation driven by chemistry at interfaces
- Chemistry at interfaces involves dynamical processes of complex chemical transformations



Car-Parrinello plane wave simulation of uranium oxide behavior in solution.



Molecules interacting with a reactive surface



Complex materials and interfaces

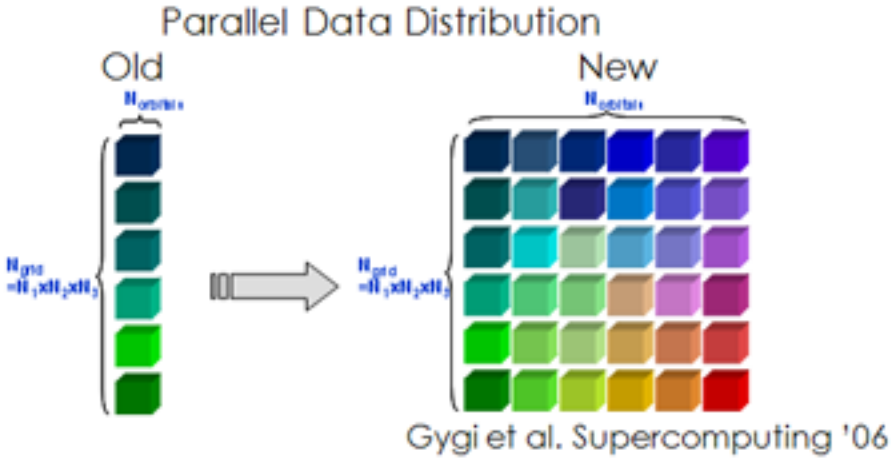
Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

U.S. DEPARTMENT OF ENERGY

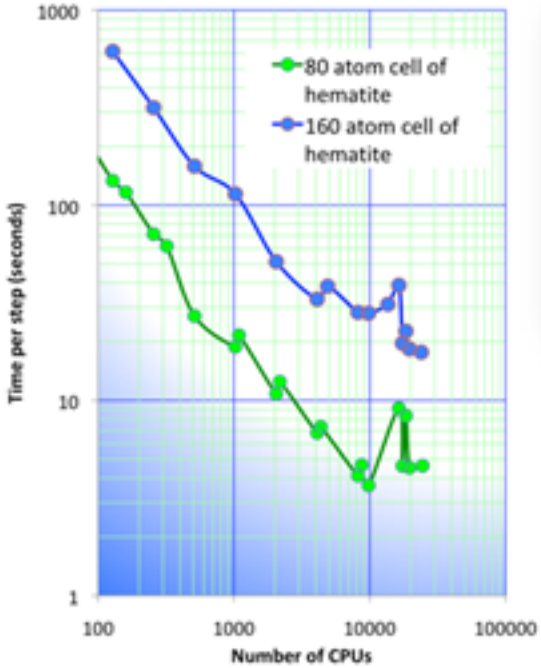
Hard Petaflop Scaling of Plane Waves

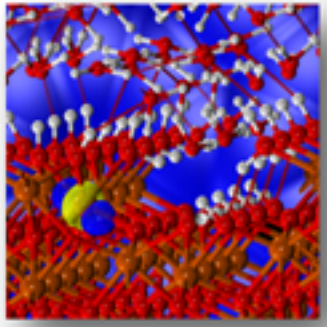
EMSL

- Smart data distribution and smart communication algorithms enable hybrid-density functional theory to scale to large numbers of processors



Gygi et al. Supercomputing '06





Hybrid density functional theory plane wave calculation on hematite scaling on NERSC's Franklin

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965

U.S. DEPARTMENT OF ENERGY

Getting Ready for Extreme Scale Chemistry

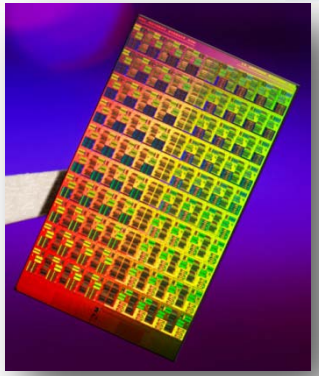
EMSL

- Efficient NWChem on extreme scale platform enables our users to
  - Tackle larger molecular complexes with greater accuracy
  - Integrate of methods and dynamics
  - Obtain their results with a faster time-to-solution on large computing platforms
- Goal is to enable NWChem to run effectively on extreme scale platforms
  - Mitigating scalability issues in computational chemistry
  - Software infrastructure and architecture (re)design
    - 10,000s to 100,000s processors
    - Computing paradigm is changing drastically
    - Actor based models for efficient load balancing
    - Applications need fault resilience and/or tolerance

NWChem - Chemistry on New Architectures

EMSL

- Large multicore compute nodes
  - Heterogeneous nodes and/or heterogeneous computing in a node
- GPGPU
  - Rapidly evolving field
- Heterogeneous networks
  - Align data movement with network topology/bandwidth



Intel's Teraflop Research Chip has 80 cores and 1 TFlops performance

Barracuda: NWChem development cluster funded by Recovery Act

NVIDIA Tesla S1070
4 TFlops in 1U

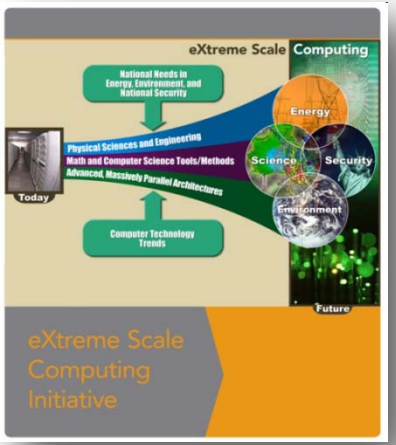
Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965
U.S. DEPARTMENT OF ENERGY

Achieving Scalability Takes a Team

EMSL

- Extreme scaling requires a collaborative multidisciplinary development team
  - Computational chemists
  - Mathematicians
  - Computer scientists

- PNNL's eXtreme Scale Computing Initiative leads the way
  - Scalable computational chemistry methods
  - Scalable software infrastructure
  - Fault tolerance at the application level



NWChem Is Going Open-Source

EMSL

- NWChem Consortium to deliver code and infrastructure for computational chemistry community to build upon
- Greatly expands user base through ease of access by students
- Expanding developers base by leveraging development and resources of universities and national labs
- Establish more collaborative development environment

NWChem - Delivering Computational Chemistry at Scale

Questions?

NWChem brochures available.