



NWChem, An ANL Legacy

Transition to Parallel Computing in Chemistry

Thom H. Dunning, Jr.

National Center for Supercomputing Applications
Department of Chemistry, School of Chemical Sciences
University of Illinois at Urbana-Champaign

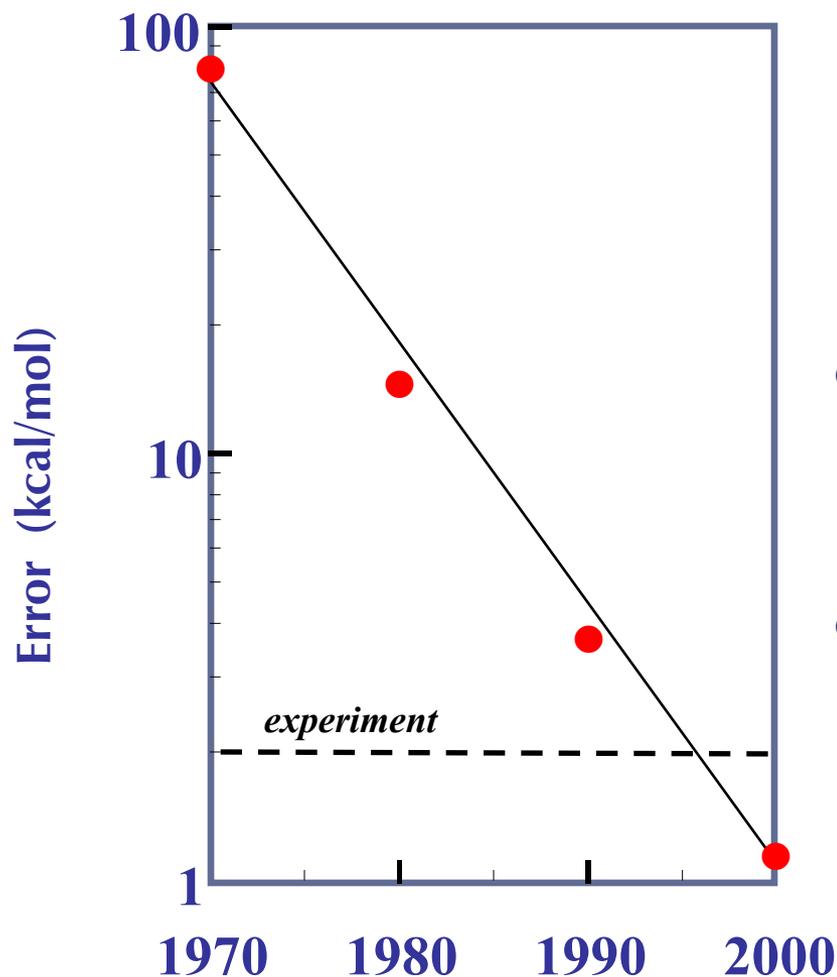


National Center for Supercomputing Applications
University of Illinois at Urbana-Champaign

Overview of Presentation

- **Sixty Years of Computational Chemistry**
 - Increasing accuracy of bond energies
 - Increasing complexity of molecular systems
- **Computational Chemistry @ Argonne in 1980s**
 - Focus on combustion chemistry
 - Exploration of new computing technologies
- **NWChem, A Legacy of Argonne**
 - PNNL MSRC/EMSL Project
 - PNNL recognized importance of molecular science
 - PNNL worked with BES to create new research center
 - Theory, Modeling & Simulation
 - NWChem & ECCE

Increasing Accuracy of Simulations



- **Bond energies**

- Critical for describing many chemical phenomena
- Difficult to determine experimentally

- **Accuracy of predictions**

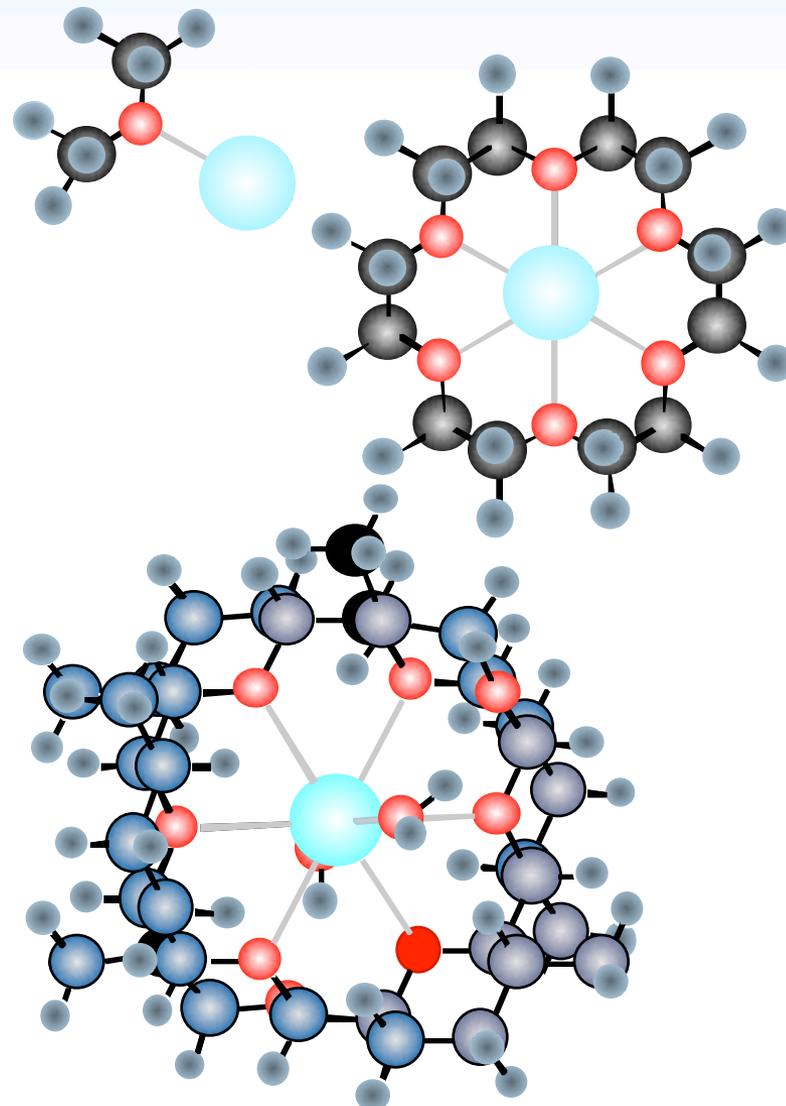
- Increased dramatically from 1970-2000

- **Combination of advances in**

- Theoretical approaches
- Computational techniques
- Computing power

Increasing Reach of Simulations

- **In 1990**
 - Model systems, e.g., ether–alkali ion complexes
- **In 2000**
 - Model separations agents, e.g., 18-crown-6–alkali ion complexes
- **In 2010**
 - Real-world separations agents, e.g., Still’s crown ether–ion complexes



Every field of computational science has a similar story to tell!

The purpose of computing is insight, not numbers.

Richard W. Hamming, 1962

*The purpose of computing is numbers as well as
insight.*

with apologies to Dr. Hamming

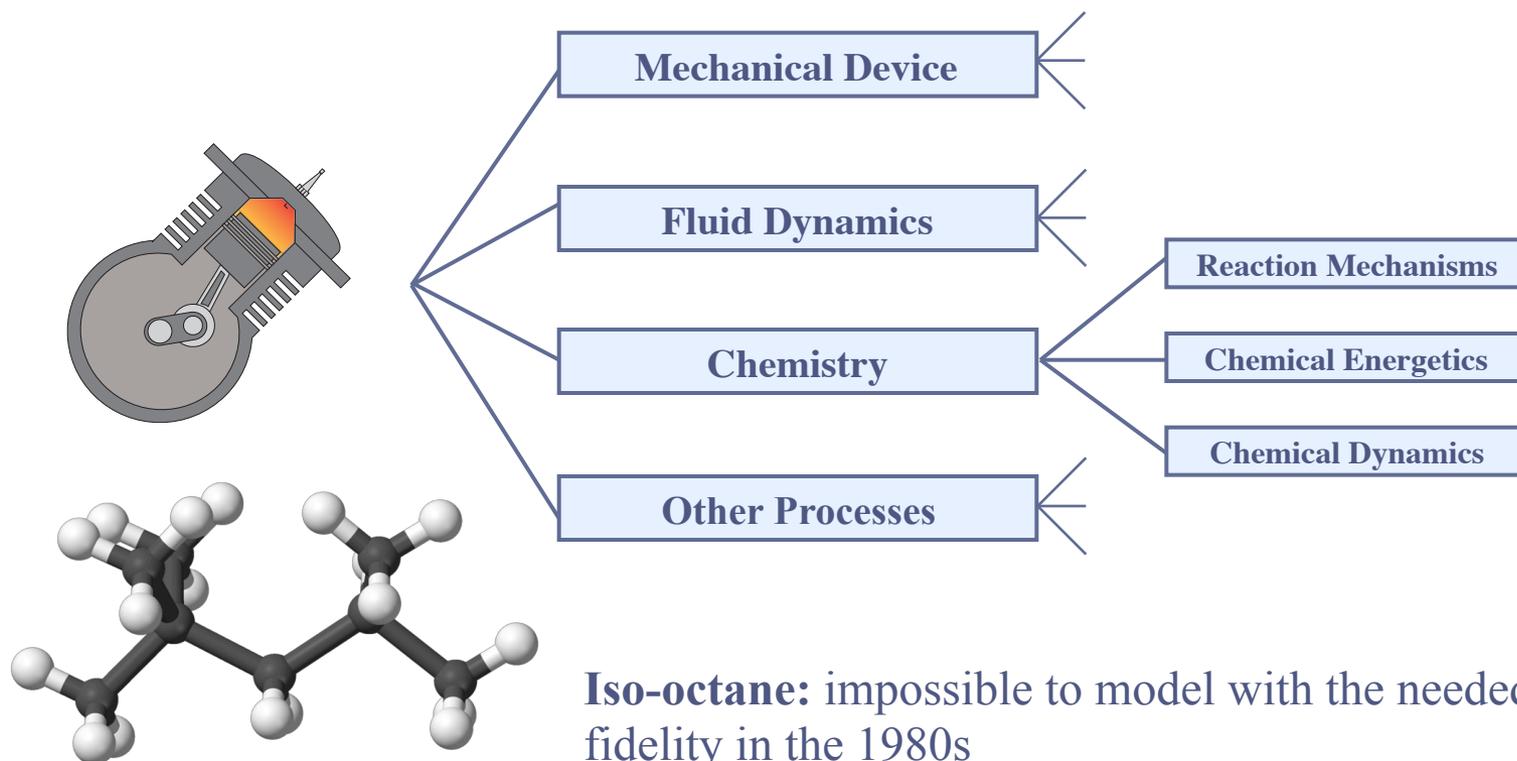
Computational Chemistry @ ANL

From the 1960s on, the Chemistry Division @ ANL had a strong program in theoretical and computational chemistry

Computational Chemistry @ ANL in 1980s

- **Focus**

- Chemistry problems critical to mission of DOE, namely, combustion of hydrocarbon fuels



Computational Chemistry @ ANL in 1980s

- **Focus**

- Chemistry problems critical to mission of DOE, e.g., combustion of hydrocarbon fuels
- Computational challenges led to an effort to explore computing technologies for computational chemistry

- **Argonne Computing Resources**

- Lab: IBM 370/195 and successors
- Division: VAX 11/780
- T&CC Group:
 - Constant interactions with MCS
 - FPS-164 (vector mini-supercomputer)
 - Alliant FX/8 (parallel computer)



Development of NWChem

A new computational capability in chemistry funded by PNNL's MSRC/EMSL Project, with supplemental funding from OASCR

MSRC/EMSL Project: A Golden Opportunity

- **Vision**

- Bill Wiley, PNNL director, who had seen opportunities in his own field of biology created by molecular science research
- Bob Marianelli, director of BES Chemical Sciences, who understood the importance of molecular sciences to DOE mission

- **PNNL recruited THDJr to lead Theory, Modeling & Simulation Directorate**

- Funded computer hardware: prototypes and HPCS-1
- Funded chemistry software development: NWChem & ECCE

- **Turbulent History**

- Started as a BES project, transferred to DOE-EM, returned to BER



NWChem Project Goals

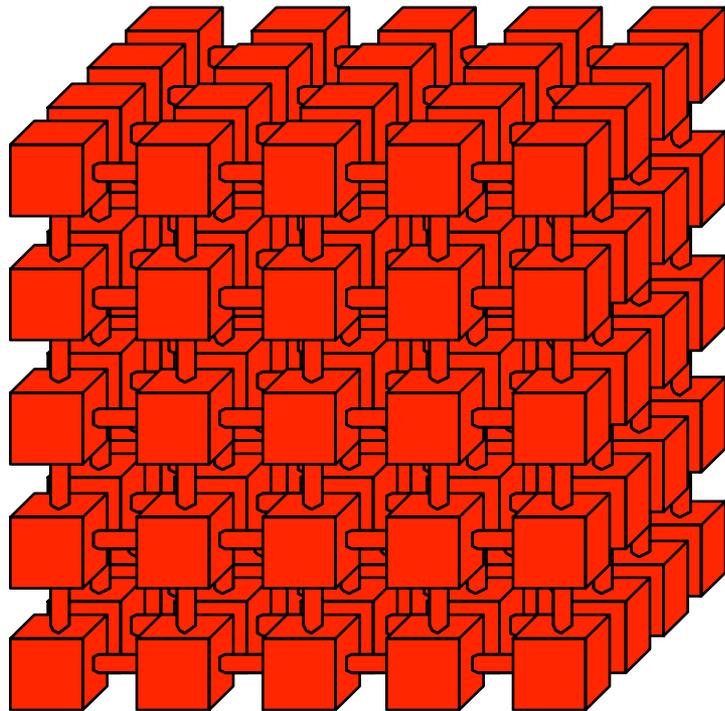
- **Capabilities**

- Designed from the ground up for parallel computers
- Full range of basic computational chemistry capabilities
- Modular and extensible
- Scalable to “large” number of processors

- **Implementation**

- Affordable to develop, maintain and extend
- As independent of parallel computer architecture as possible
 - But, broadly tunable to critical parameters of computer system

Issues in Parallel Computing



Hardware Issues

Processor speed
Memory latency and bandwidth
Communications latency and bandwidth
...

Software Issues

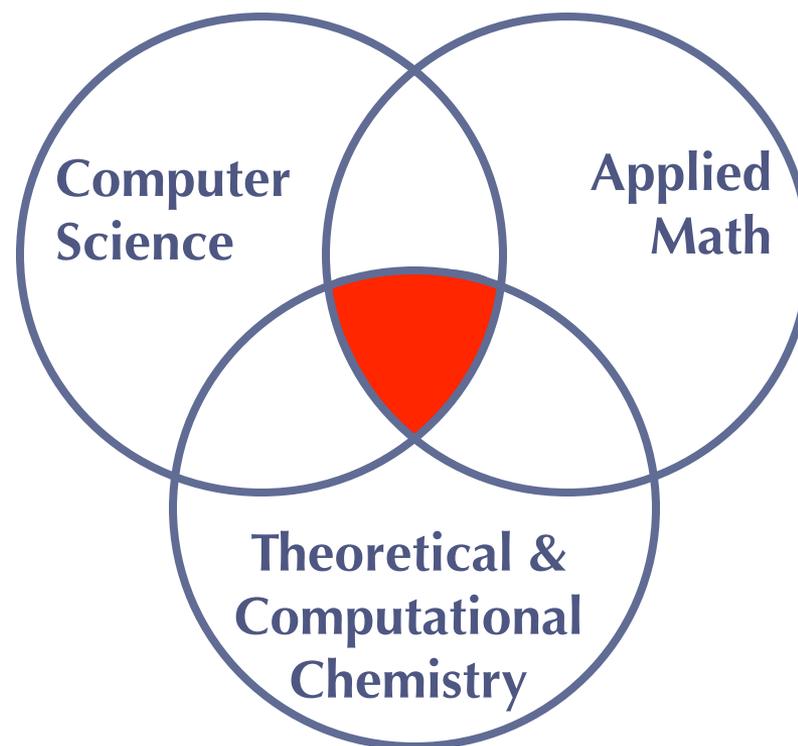
Single processor performance
Load balancing
Scalability
Portability
...

A Multidisciplinary NWChem Team

Brought theoretical and computational chemists, computer scientists, and applied mathematicians, together in a **long term collaboration**.

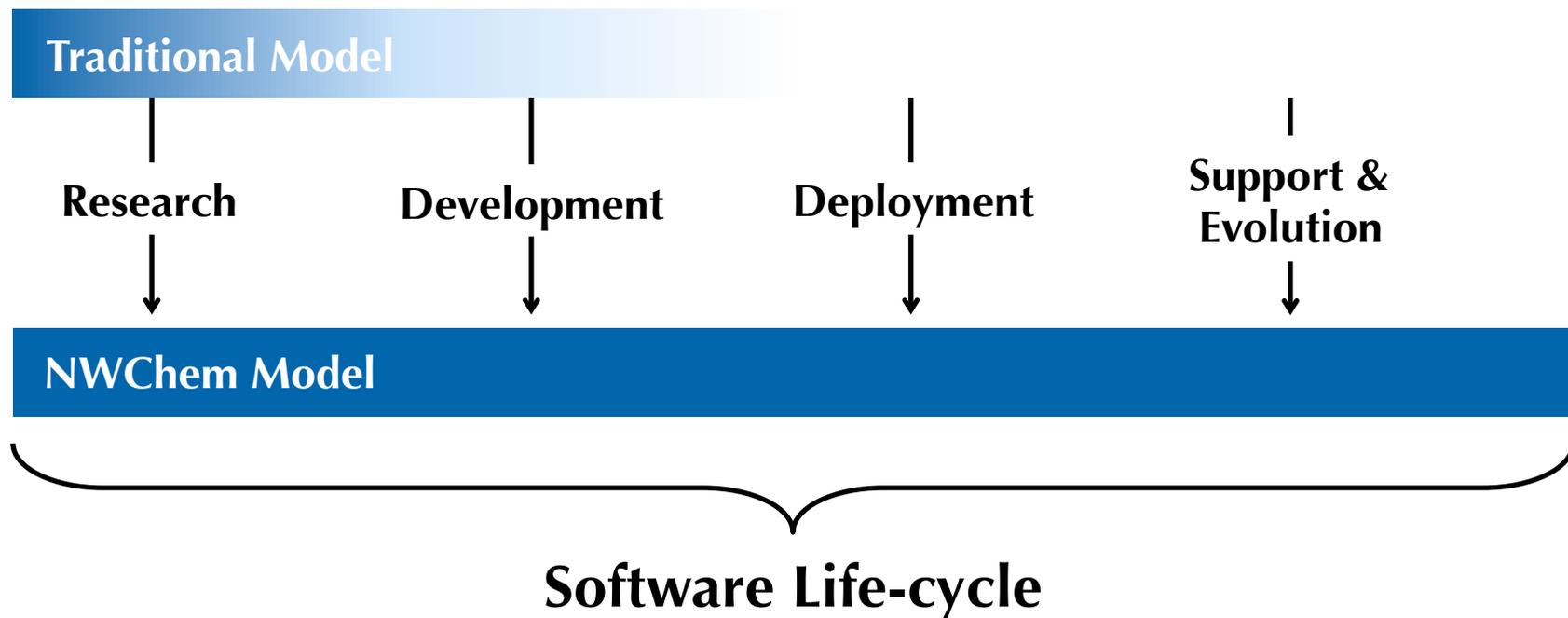
The collaboration is still thriving, although some of the players have changed.

NWChem just went *open source!*

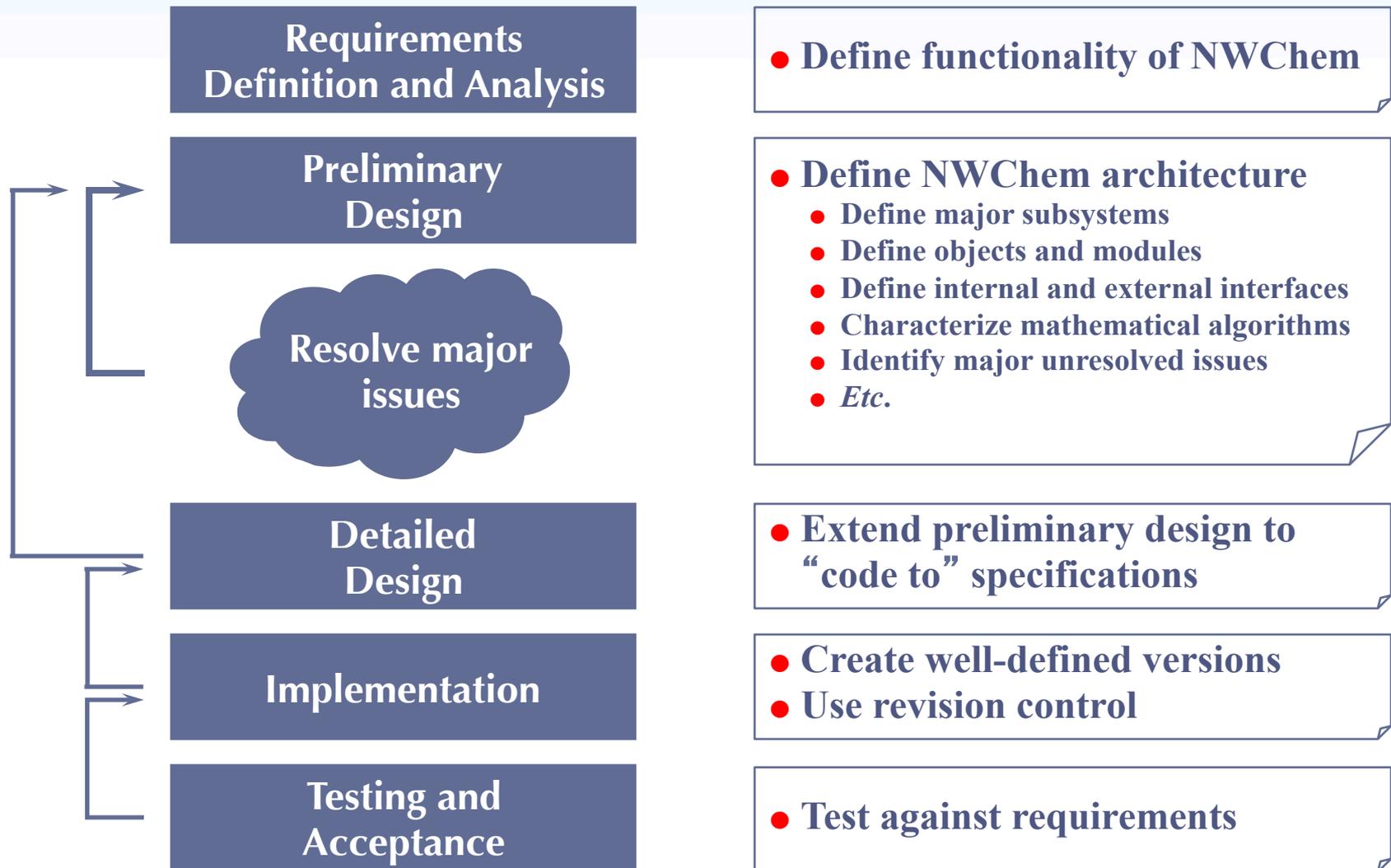


Focus on Software Life Cycle

Funded and integrated activities ranging from **research and development to deployment, support and evolution** to ensure that NWChem had and continued to have innovative, yet robust simulation capabilities



Modern Software Engineering Practices



NWChem Team

- **PNNL Core Team**

- Ray Bair, Chemist (Founded Effort)
- Jeff Nichols, Chemist, Team Leader
- Robert Harrison, Chief Chemistry Architect
- Ricky Kendall, Chemist
- TJ Straatsma, Chemist
- Jarek Nieplocha, Chief Computer Scientist
- Rik Littlefield, Computer Scientist, Applied Mathematician
- George Fann, Applied Mathematician

- **Postdoctoral Fellows**

- J. Anchell, E. Apra, D. Bernholdt, P. Borowski, T. Clark, H. Dachsel, M. Deegan, H. Fruchtl, R. Kutteh, X. Long, B. Meng, M. Stave, H. Taylor, A. Wong

NWChem Team (continued)

- **Visiting Scientists**

- K. Dyllal, M. Gutowski, K. Wolinski (Poland)

- **PNNL Collaborators**

- M. Dupuis, D. Elwood (CS), M. Feyereisen, A. Hess, J. Jaffe, J. Nicholas, M. Rosing (CS), G. Thomas (CS), M. Thompson

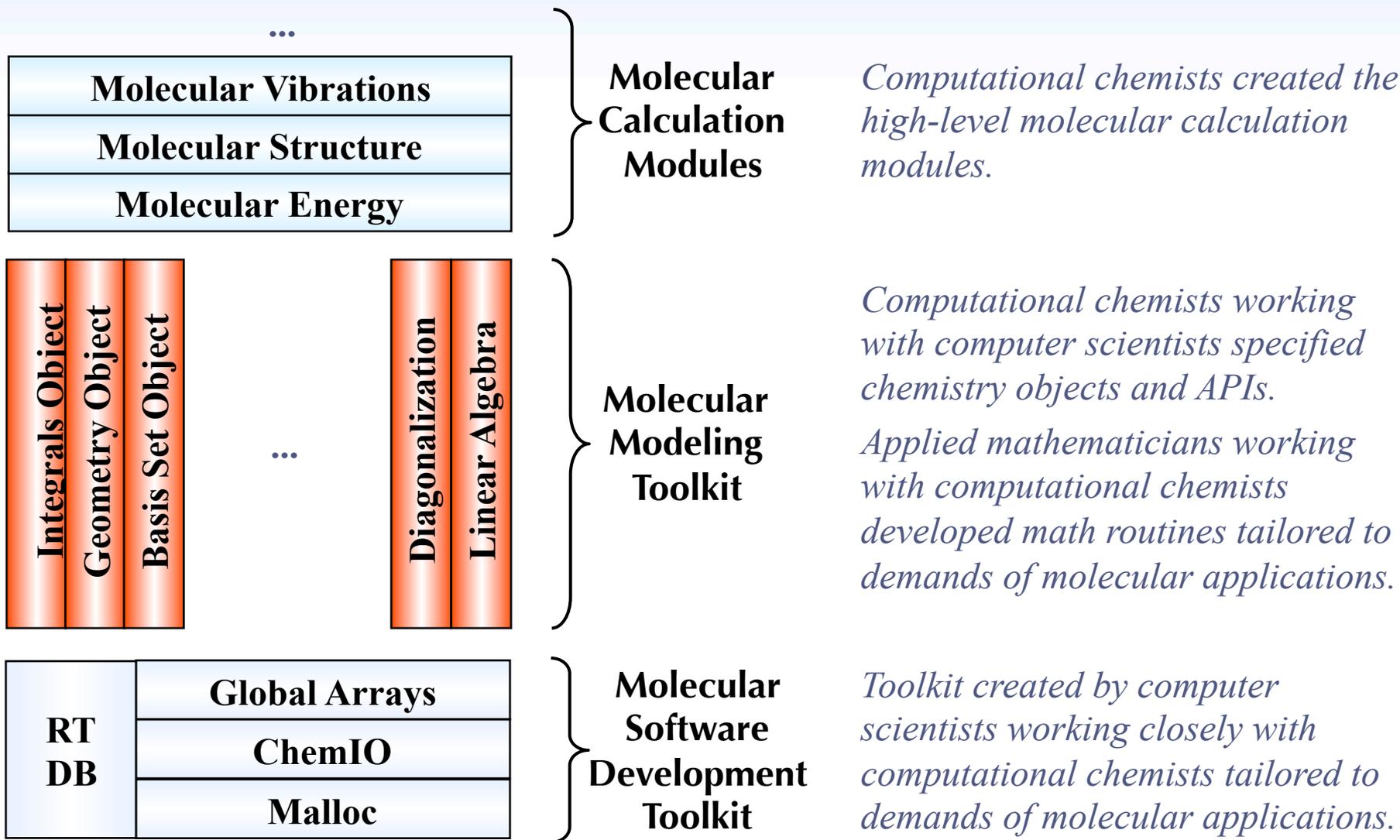
- **U.S. Collaborators**

- K. Jordan (U. Pittsburgh), A. McCammon (UC, San Diego), E. McCullough (Utah State U.), S. Plimpton (SNL), P. Pulay (U. Arkansas), R. Shepard (ANL), J. Tilson (ANL)

- **Foreign Collaborators**

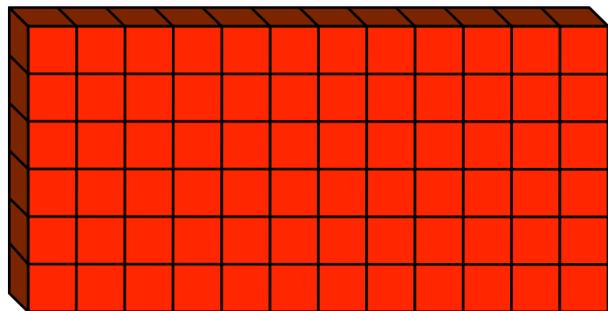
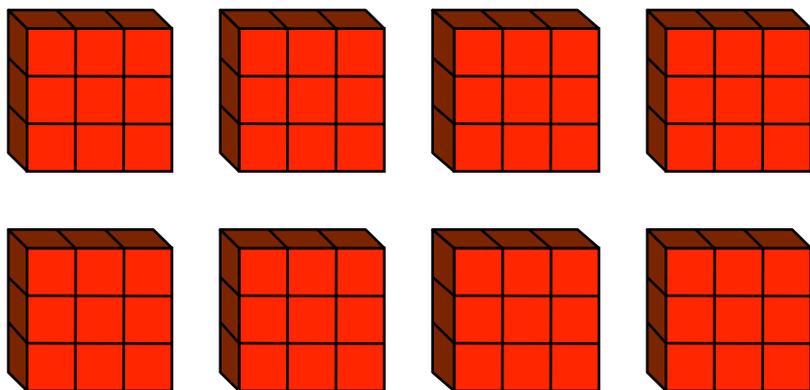
- M. Guest (UK), P. Knowles (UK), H. Lischka (Austria), A. Rendell (Australia), P. Sherwood (UK), W. Smith (UK), R. Wade (Germany)

NWChem Architecture



NWChem Innovation: Global Arrays

physically distributed data sub-blocks



single logical data structure

“Shared memory” access to physically-distributed arrays

- Single logical array, physically-distributed array sub-blocks
- Fast local access; one-sided asynchronous remote access
- Accounts for non-uniform memory access (NUMA)
- Portable—implemented on most existing shared memory and distributed memory computers
- Through ARMCI has been extended to SMP clusters

NWChem Innovations

- **Performance Features**

- **Plug & Play**—new computational advances can readily be incorporated into the program
- **Adaptive**—the optimum algorithm is used for the given number of processors, memory, local & remote disk storage

- **Portability Features**

- To port NWChem to another computer, **the only major changes that have to be made** are in the bottom layer
- **However**, the higher layers of NWChem are usually **tuned** to maximize performance

NWChem: Evolution and open-source!



NWChem 3.X
(1995-2001)

NWChem 4.X
(2001-2006)

NWChem 5.X
(2006-2008)

NWChem 6.X
2008 - 2012

Complete basic suite of tools

- ◆ High accuracy
- ◆ Density Functional Theory
- ◆ *Ab-initio* MD
- ◆ Classical MD
- ◆ Basic NMR

Expanding capability suite

- ◆ Environmental models
- ◆ Spectroscopic properties
- ◆ TDDFT excited states
- ◆ Relativistic effects
- ◆ POLYRATE direct dynamics



Fast time-to-solution

- ◆ High accuracy reaches 1.3 Pflop/s

Partners contribute

- ◆ Q-HOP in MD  UNIVERSITÄT DES SAARLANDES
- ◆ VENUS interface  TEXAS TECH UNIVERSITY
- ◆ Constrained DFT



New functionality

- ◆ DFT with long range dispersion

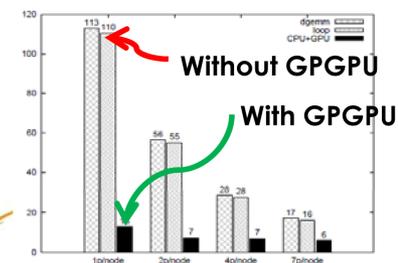
Open-source release

- ◆ Education Community License 2.0

Unique tools

- ◆ Real-time TDDFT response
- ◆ Optical rotation

Scalability on new hardware



Pacific Northwest
NATIONAL LABORATORY



Thank You!

