BES Computing Requirements

Robert J. Harrison (UT/ORNL)

with contributions from

Steven G. Louie (LBNL),
Marvin L. Cohen (UCB),
M. Mavrikakis (UW-Madison),
Gregory Newman (ESD)
P.K. Yeung (GA Tech.)

Materials Theory at NERSC: Recent results and future directions

PI's: **Steven G. Louie**
Marvin L. Cohen
Materials Sciences Division, LBNL
Department of Physics, UC-Berkeley

First-principles methods

INPUT

Atomic positions

Parameter-free

First-principles METHOD

- **Density functional theory (LDA, GGA, etc)**
- **Many-body perturbation theory**
- **Nonequilibrium transport**

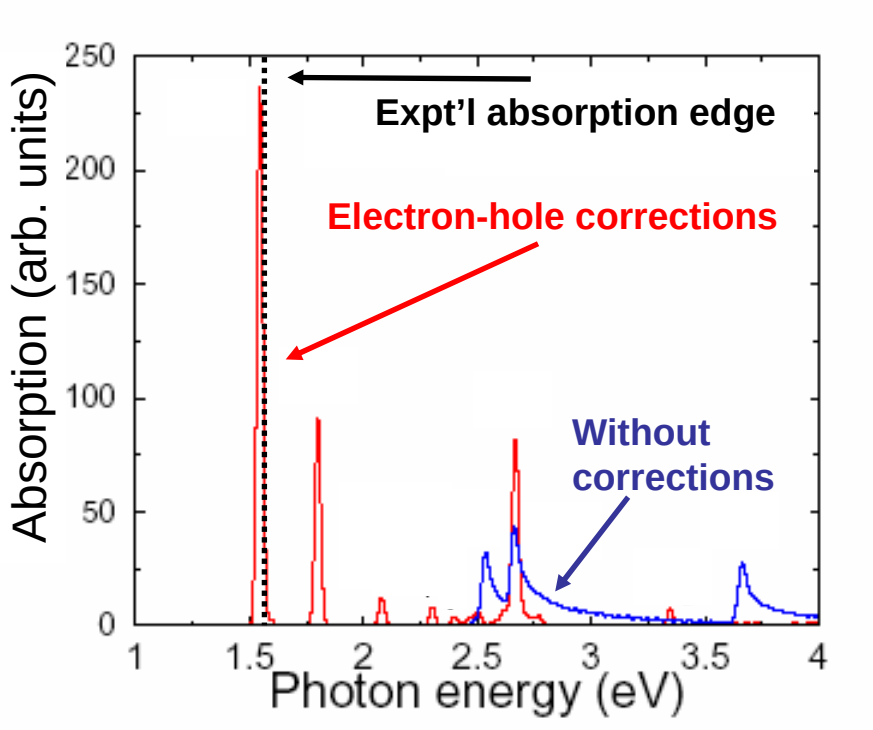
OUTPUT

Experimental observables:

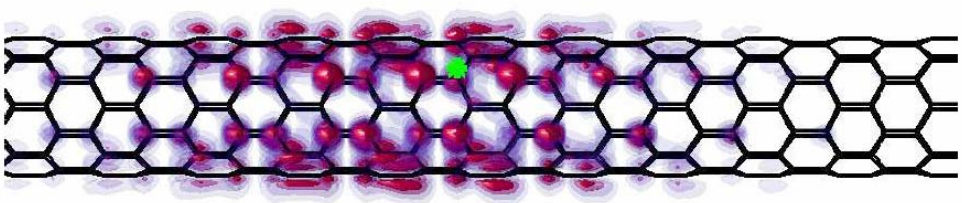
Atomic structure
Chemical bonding
Optical absorption
Strength
Resistance
Etc.

First-principles calculations are parameter-free and require only an initial guess at the atomic positions as input. Depending on the level of approximation, a variety of materials properties can be computed and compared directly with experiment. Currently, our methods are parallel up to 8 nodes, involve many linear algebra and FFT operations, and require large memory and ultrafast communication between processors.

Beyond conventional approximations: many-body perturbation theory



Calculated photon absorption spectrum of a (8,0) carbon nanotube



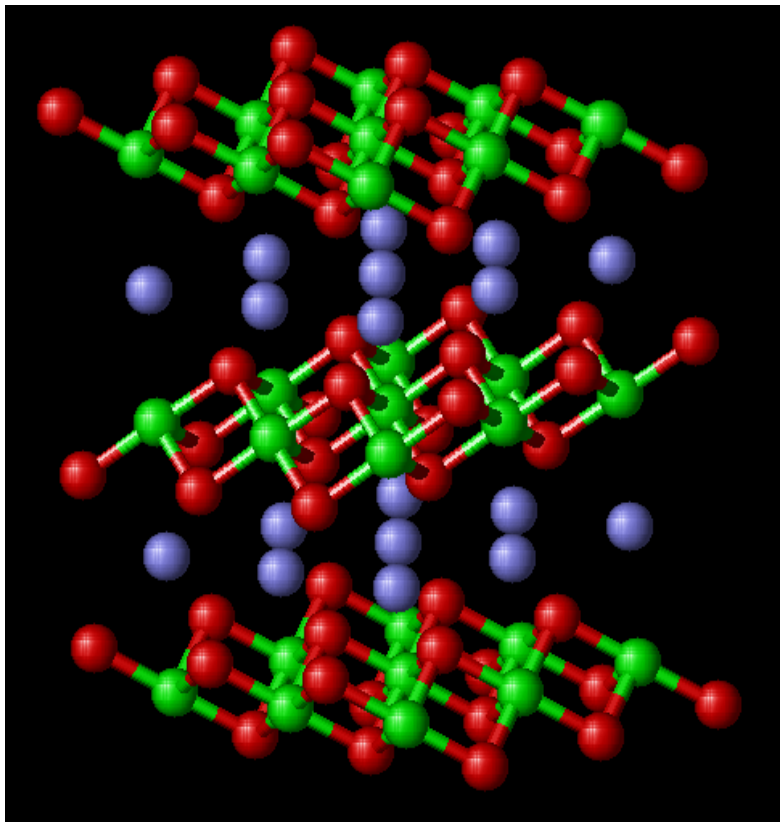
Density plot of exciton wavefunction at the absorption edge

Including many-electron effects when calculating optical properties for nanostructures is essential. At left, we found that without the electron-hole interaction, the absorption spectrum of a carbon nanotube disagrees with experiment by a factor of two. Obtaining this spectrum required more than 200,000 hours at NERSC using between 4 and 8 nodes.

Spataru, Ismail-Beigi & Louie, Phys. Rev. Lett. (2004)

Predicting and understanding novel properties and materials

O Co Na



Calculated crystal structure of $NaCoO_2$

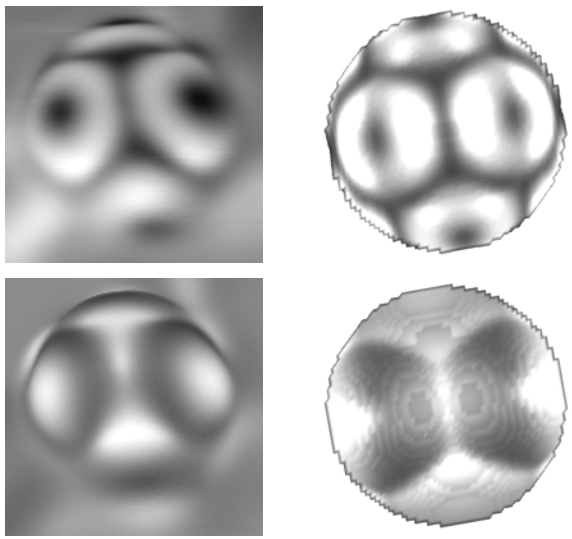
At left is the calculated crystal structure of Na_xCoO_2 an exotic thermoelectric oxide that also exhibits superconductivity. The strong electron correlation present in this material requires an intermediate level of theory, “LSDA + U”. Calculation of its structural, electronic, vibrational properties required 50,000 CPU hours incrementally, using between 2 and 4 nodes.

Zhang, Luo, Crespi, Cohen & Louie, Phys. Rev. B (2004)

Helping to interpret experiments

C_{60} on Ag(100)

Experiment **Theory**



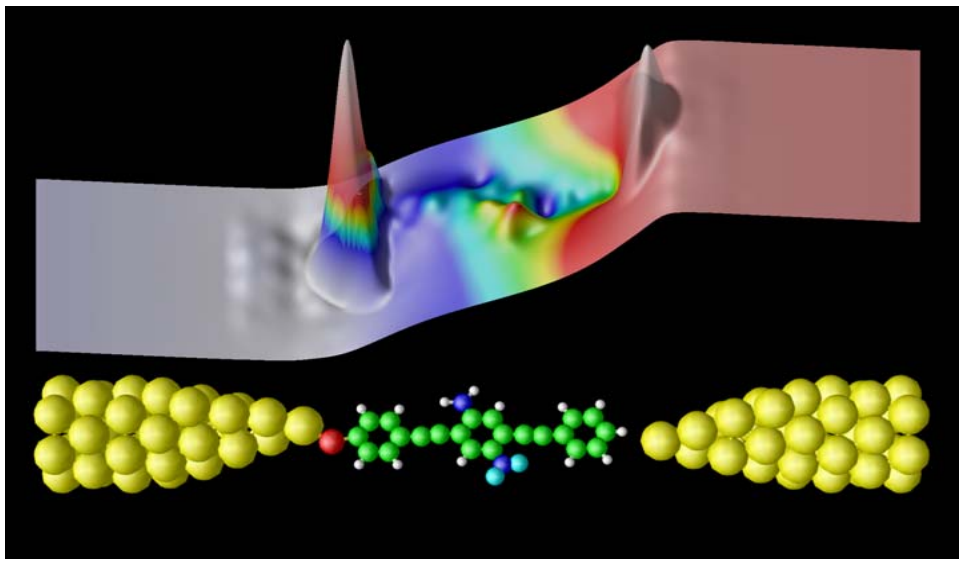
Comparison of experimental STM images of a single C_{60} molecule on Ag(100) with theory.

Lu, Grobis, Khoo, Louie & Crommie, Phys. Rev. Lett. (2003).

Using density functional theory we have simulated STM images of C_{60} molecules on a silver surface and compared with high-resolution experiments. The important roles of the Ag(100) substrate and the tip trajectory in producing the observed images are isolated and explained by our theory. The calculations involved about 250 atoms, and required about 5,000 CPU hours at NERSC with 1 or 2 nodes.

Guiding the development of future technology

Computational design of new systems for molecular electronics



Structure and potential energy profile at finite bias of a ‘Tour’ wire, an organic molecule connected by gold leads

We have recently developed a new computational approach for predicting how the resistance of single molecules is affected by the presence of an applied voltage. Treating both the molecular wire and metallic leads quantum mechanically requires many atoms and electrons (100-500). These state-of-the-art calculations, still in progress, have required 100,000 hours on NERSC thus far.

Choi, Neaton, Son, Khoo, Cohen & Louie, in progress

Future goals of our materials theory program

Scientific directions

• Computational modeling of complex materials, such as nanostructures, in different environments, under a variety of thermodynamic conditions, and with an applied voltage bias

• More atoms, new methodologies, and the development of novel, efficient algorithms

Computational requirements

• “First-principles” calculations demand high-speed internode communication and lots of memory, and therefore run most efficiently on fewer nodes. A more flexible queuing system with enhanced priority for smaller (1-8 node) batch jobs is thus highly desirable.

• Internal collaborations toward further optimization of our existing algorithms for NERSC architecture.

DOE-BES
Chemical Sciences
(M. Mavrikakis – UW-Madison)

- Atomic-scale design of heterogeneous catalysts directly from *first-principles*
- Example applications:
  - Fuel Cells for mobile and stationary applications
  - Hydrogen Production from renewable feedstocks
  - Production of high value intermediates for pharmaceuticals

What can NERSC do?

(M. Mavrikakis – UW-Madison)

- Keep upgrading parallel computing environment every 2-3 years. SP's are very reliable platforms and should be continued.

- Over the last few years, there is a general trend to give higher priority to super-large jobs. However, state-of-the-art static planewave codes do NOT parallelize well beyond O(100) CPUs. Devise and maintain a queuing policy that would make our lives easier!

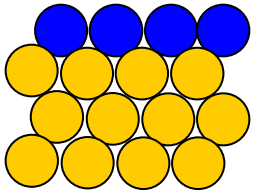
- Either establish separate hardware for medium range parallel calculations (up to O(100) CPUs/job) or have a special queue on a large parallel computer for typical computational chemistry jobs.

Design and Screening of Promising Heterogeneous Catalysts for FUEL CELLS

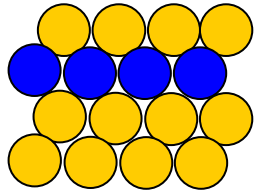
Using modern Computational Chemistry methods

(M. Mavrikakis – UW-Madison)

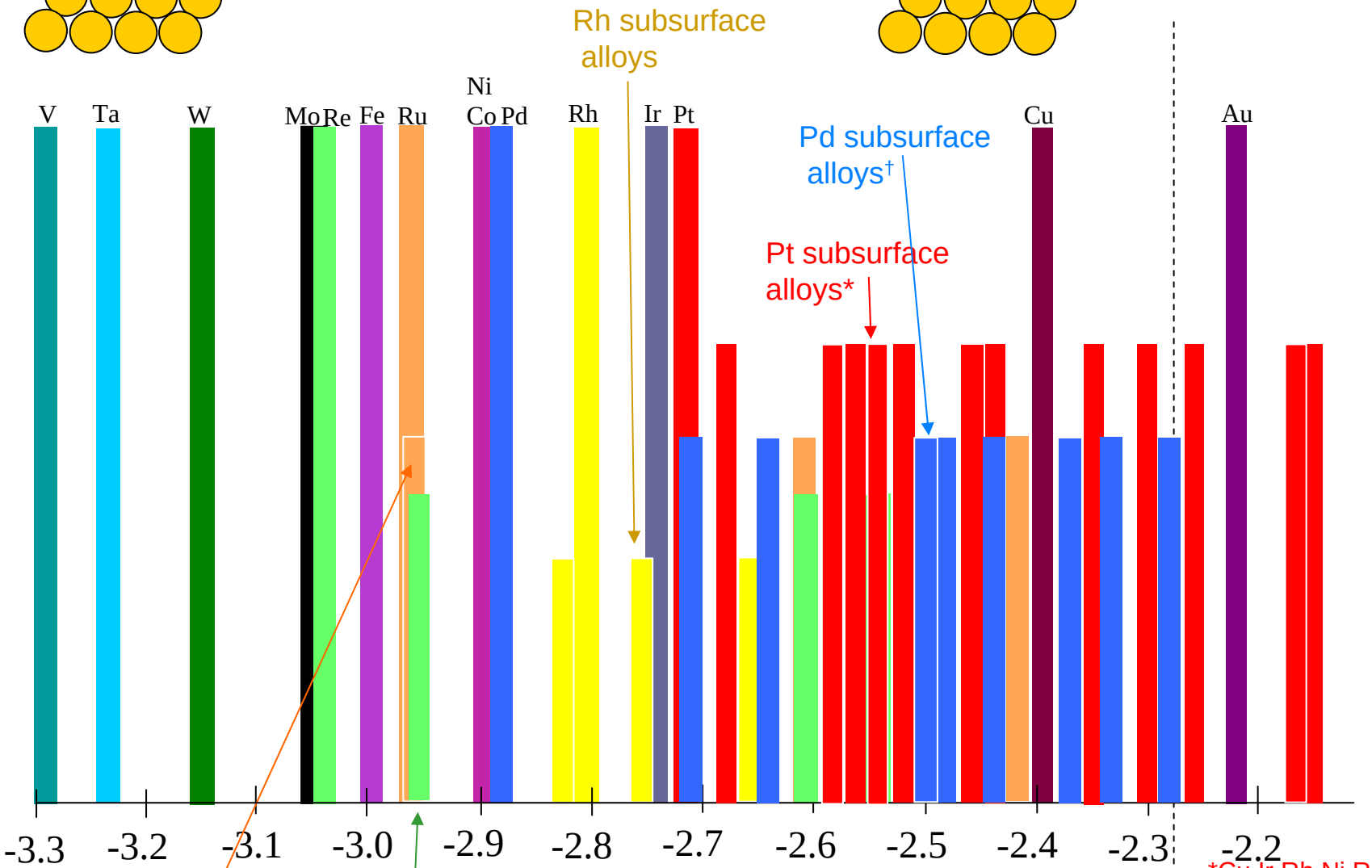
Hydrogen Binding Energy (M. Mavrikakis – UW-Madison)



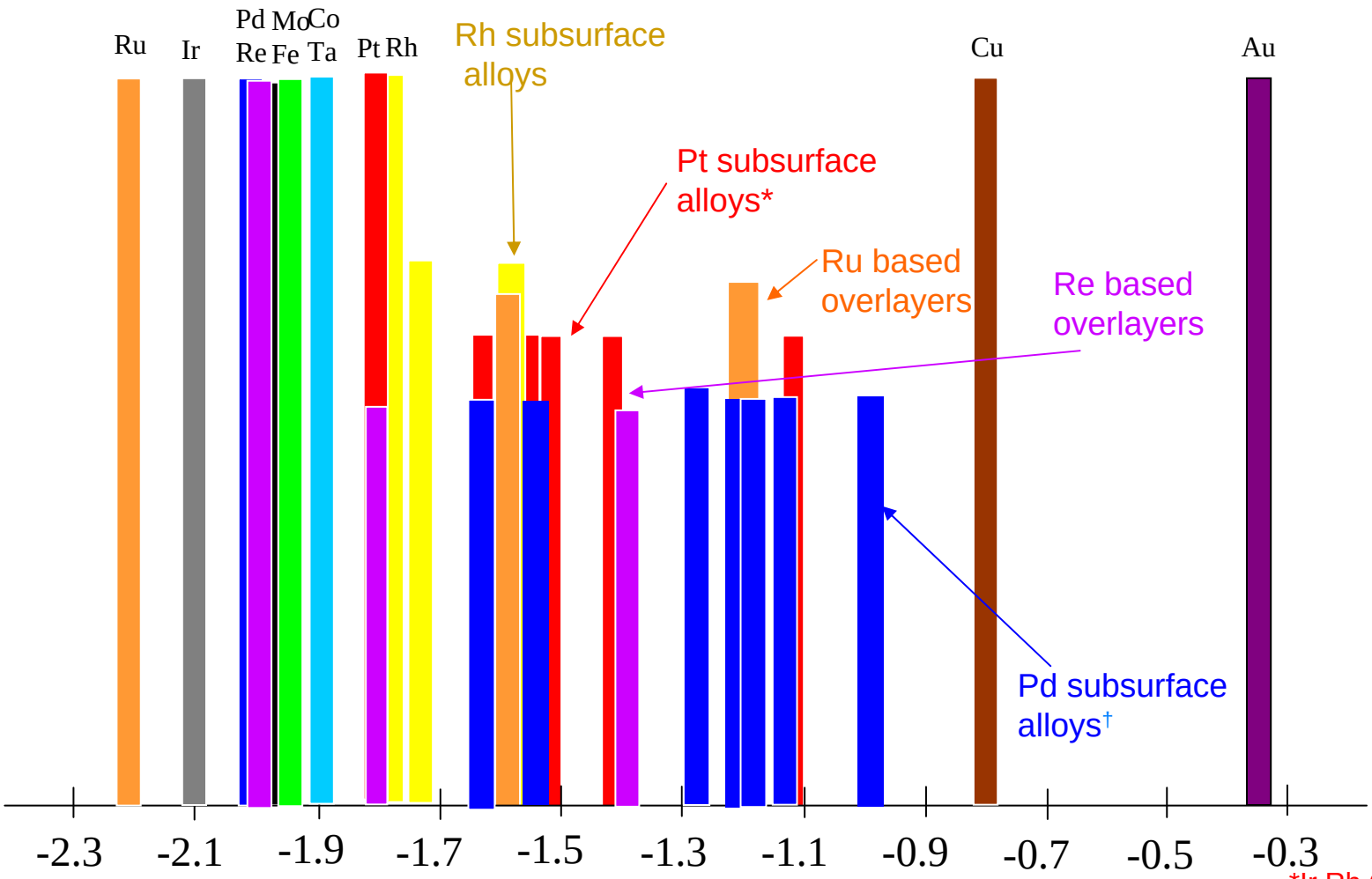
Overlayers



Subsurface Alloys



CO Binding Energy:
Application to Hydrogen Fuel Cells

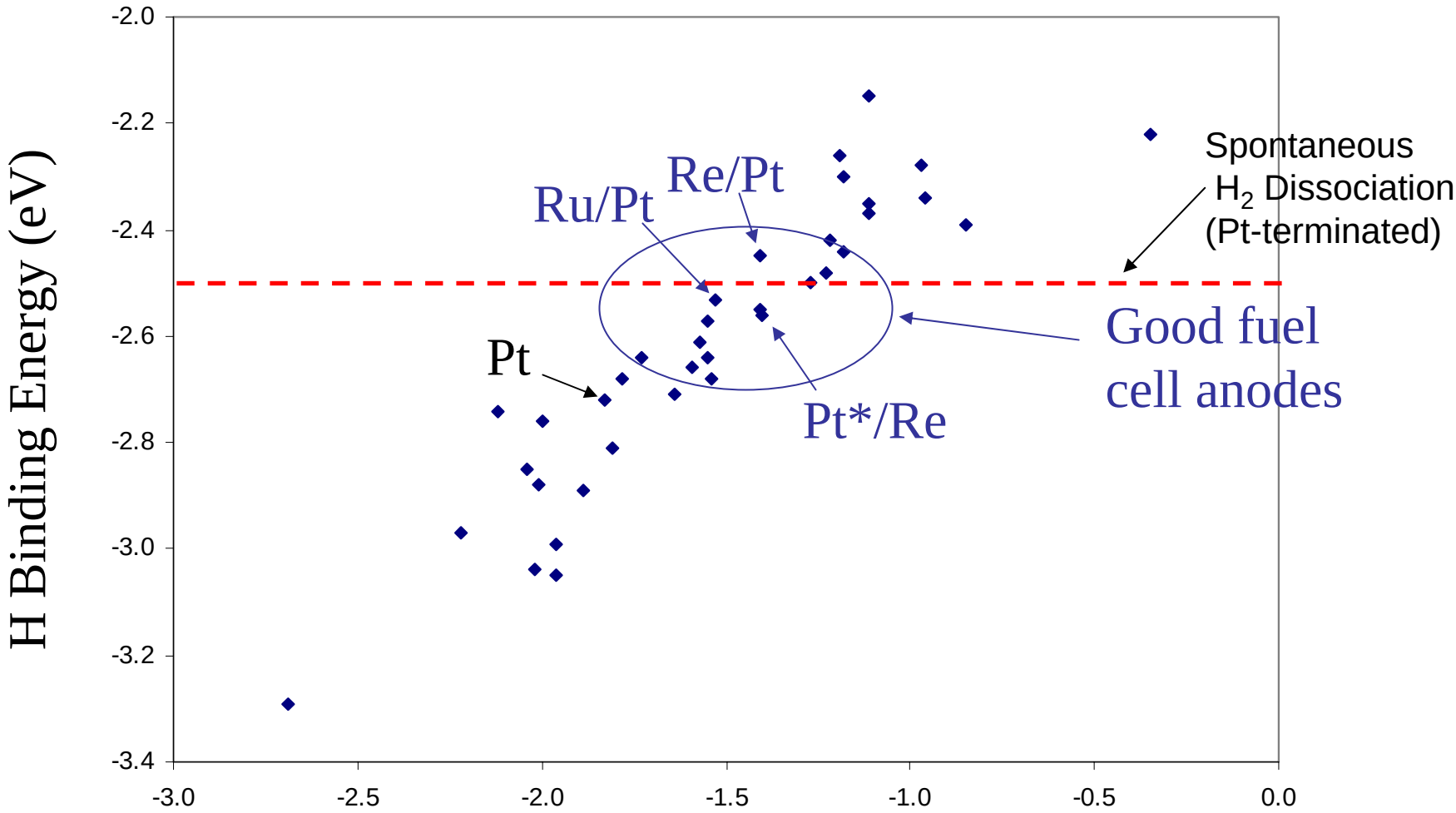


(M. Mavrikakis – UW-Madison)

*Ir,Rh,Cu,Ru,Re,Ni Mo,Fe

†Ir,Ru,Re,Fe,Mo,W, V,Ta

Atomic-scale Design of Fuel Cell Anode Catalysts



(M. Mavrikakis – UW-Madison)

High Resolution Geophysical Imaging on Distributed Computing Architectures

Gregory Newman

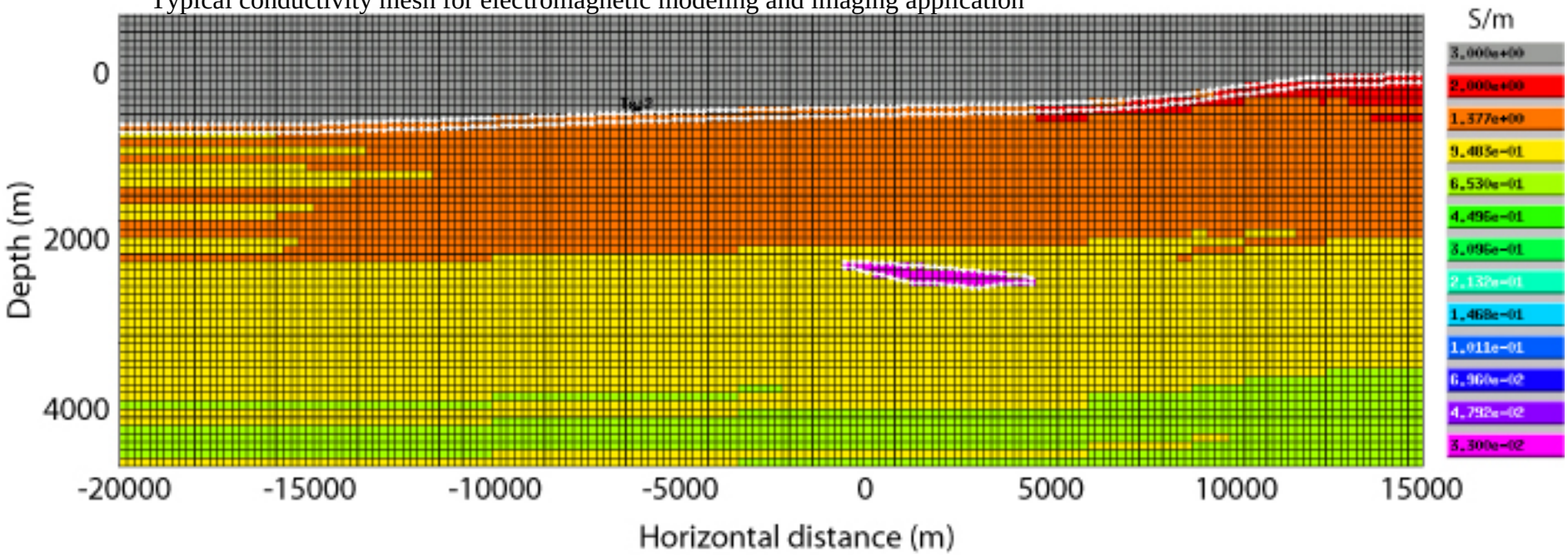
ESD

June 21, 2004

High Resolution 3D Imaging

- Provides unprecedented level of detail on subsurface geological processes
- Applications already exist – oil & gas exploration, environmental, geothermal resource evaluation and geological hazards
- Problem is computationally intensive

Typical conductivity mesh for electromagnetic modeling and imaging application



Large Scale 3D Geophysical Imaging
(electromagnetics)

Forward Problem – one transmitter - one frequency

Meshes ~ 27 million elements

- sparse linear systems ~ 80 million unknowns
- solved using Krylov-methods 10^{13} Flops

  cost 2000 matrix-vector multiples

Inverse Problem – one transmitter - one frequency

- must solve forward problem multiple times
- 4 fwd solves per inversion iteration
- 50 inversion iterations: 2 x10^{15} Flops

Large-Scale 3D EM Inversion

Image Fidelity Requires:

- multiple transmitters ~100 (lateral resolution)
- multiple harmonics ~10 (depth resolution)

Inversion Cost 2×10^{17} flops

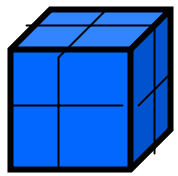
Large Scale 3D EM Inversion

Multi-Source Inversion Highly Parallel

- domain decomposition

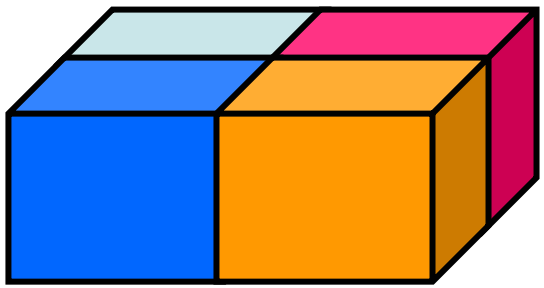
distribute fwd problem on a group of nodes for fixed source & harmonic

- data decomposition



8 nodes per group

distribute copies of fwd problem on different groups of nodes



4 groups with 32 nodes

Large Scale 3D EM Inversion

Multi-Source Inversion Considerations

– global communication

needed amongst the various data processor banks

several dot products per inversion iteration

– Main computational burden – Fwd Solves

each solve is independent of the others

– Estimated Time to Solution on 10 Tflop Platform

T= $2 \times 10^{17} / 10^{12}$= 200,000 sec or 55.55 hours

Large Scale Geophysical Imaging

– Not possible to solve problem on NERSC

current allocation for all geosciences ~500,000 hours

– Additional NERSC resources are required !!!

Engineering Physics at NERSC

P.K. Yeung, Georgia Tech

K.R. Sreenivasan, U. of Maryland & ICTP, Italy

- The computational science
- Numerical and computational issues
- Hardware requirements
- Software and system environment
- Consulting and allocations

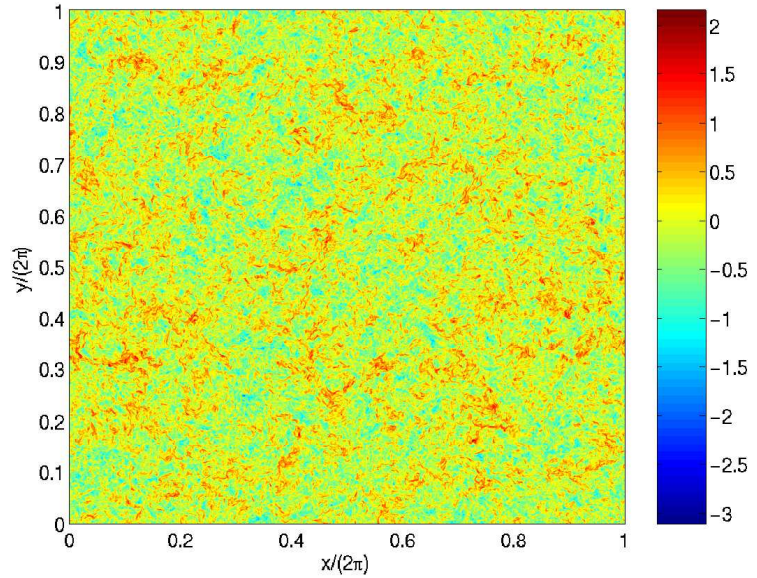
The Computational Science

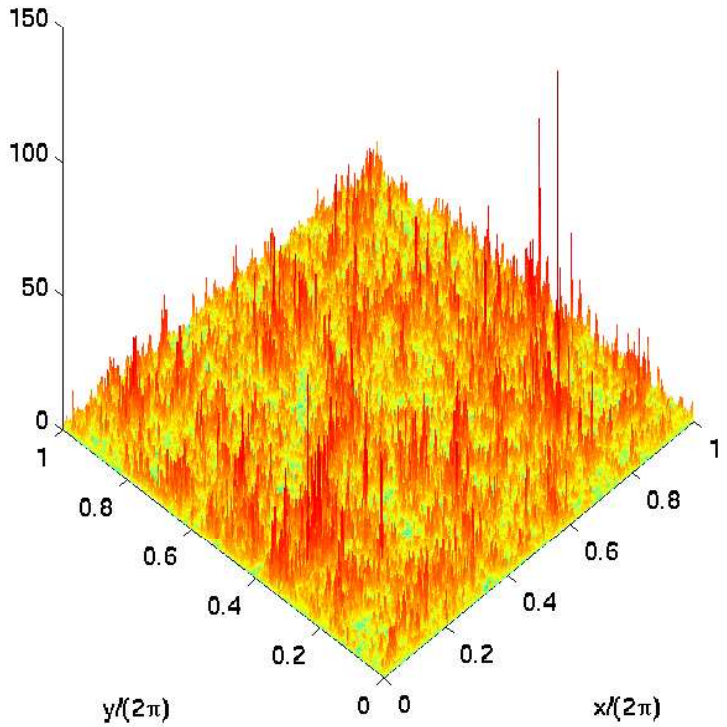
- numerical simulations based on exact equations
  provide immense detail under controlled conditions
- to obtain fundamental physical understanding
- to test/improve models for practical applications
- as exemplified by study of fluid turbulence:
  - unsteady, three-dimensional, wide range of scales
  - striving towards the previously "impossible"
    (4096^3 on *Earth Simulator* in Japan; 2048^3 in US)

Turbulence in 2048^3 simulation:

Left: 2D contours of logarithm of dissipation rate

Right: isosurfaces showing localized, intermittent peaks





Numerical and Computational Issues

- partial differential equations in time and (3D) space
- numerical accuracy, stability and efficiency, e.g:
  - pseudo-spectral methods for derivatives in space
  - step-size restrictions for explicit time integration
  - fast or machine-optimized software libraries for the most time-consuming operations
- use of massively parallel (Terascale) computers:
  - bigger problem sizes via domain decomposition
  - scalability is limited by communication calls

Hardware Requirements

- Desire a factor of 16 advance over next 5 years:
  - faster processors with more memory
  - faster communication (especially large messages among large number of processors)
- Address architecture-dependent opportunities for overlap between computation and communication
- Moderate number of powerful processors is better than large number of less powerful ones
- Large scratch disks and reliable mass storage

Software and System Requirements

- provide highly optimized software libraries
  (e.g. Fast Fourier Transform, eigenvalue solvers)
- optimization tools and performance monitors
- scheduler that manages job queues efficiently:
  - should favor the largest jobs
  - use backfilling to let small jobs run
  - provide users with maximum information

Consulting and Allocations

- access to dedicated consultants is important:
  - fast response, detailed follow-through
  - willingness to look into user codes
  - assign "strategic" consultants to high-end users
- when system problems occur:
  - encourage users to report problems (24x7)
  - keep users informed of status
- the annual resource allocation process:
  - include reviewers from PI's discipline (?)

Chemical Discovery through Advanced Computing

- A consortium of SciDAC chemistry projects computing at NERSC
- Harrison (ORNL), Gordon (Ames), Schaefer (UGA), Ermler (Memphis), Wagner (ANL), Head-Gordon (Berkeley), Thompson (OK)
- Allocation is a paltry 160K hours
  - 18 (Power3!) processors for 1 year

ChemDAC - II

- **Objectives**

1) To develop highly correlated methods for the prediction of accurate energetics, and subsequently dynamics for complex chemical processes.
2) To address the soot formation problem by focusing ab initio research on issues important to aromatization reactions
3) High fidelity simulations of combustion devices and other chemical processes using existing or emerging algorithms for the Cumulative Reaction Probability (CRP).
4) To develop computational methods that will permit the treatment of heavy-element-containing systems with accuracy comparable to that of systems containing light elements (i.e., first and second row atoms).
5) To perform all-electron calculations free of basis-set error upon the ground and excited state energies and properties of a range of combustion related and other molecular systems.
6) Electron correlations on the Si(100) surface using density functional theory (DFT) and multireference (MR) plus perturbation theory approaches.

ChemDAC codes

http://www.msg.ameslab.gov/GAMESS/GAMESS.html
http://www.q-chem.com/
http://www.emsl.pnl.gov:2080/docs/nwchem/nwchem.html
http://www.itc.univie.ac.at/~hans/Columbus/columbus_main.html
http://zopyros.ccqc.uga.edu/psi/psi.html
http://www.gaussian.com/
http://www.tc.bham.ac.uk/molpro/

These codes are predominantly written in Fortran, C and C++.
The MRA code is written in Python/C/C++.
Time-independent CRP: parallelization via PETSc and MPI in Fortran
Time-dependent CRP: sequential code in Fortran parallelized via MPI
Trajectory: sequential code in Fortran
Time-independent CRP: parallelization via PETSc, non-PETSc in Fortran
Time-dependent CRP: OpenMP code in Fortran for SMP nodes, now being adapted to MPI parallelization between nodes.

Chemistry future Science

- Chemical catalysis at the nanoscale
  - Fuel cells
  - Biological and bio-mimetic systems
  - Chemical synthesis
- Molecular electronics
  - Self assembly, electron transport
- Heavy-element chemistry
  - Rigorous relativistic methods
- Higher precision and larger systems
- Automatically generated codes (e.g., TCE)

Advanced Methods for Electronic Structure

George I. Fann[1], Zhengting Gan[1], Robert J. Harrison[1,2], So Hirata[3], Takeshi Yanai[1]

[1]Oak Ridge National Laboratory

[2]Pacific Northwest National Laboratory

[3]University of Tennessee, Knoxville

MICS SAP

BES SciDAC

In collaboration with

Gregory Beylkin[4], P. Sadayappan[5], Marcel Nooijen[6], David Bernholdt[1], Russ Pitzer[5], and others

[4]University of Colorado

[5]Ohio State University

[6]University of Waterloo

harrisonrj@ornl.gov

SciDAC
Scientific Discovery through Advanced Computing

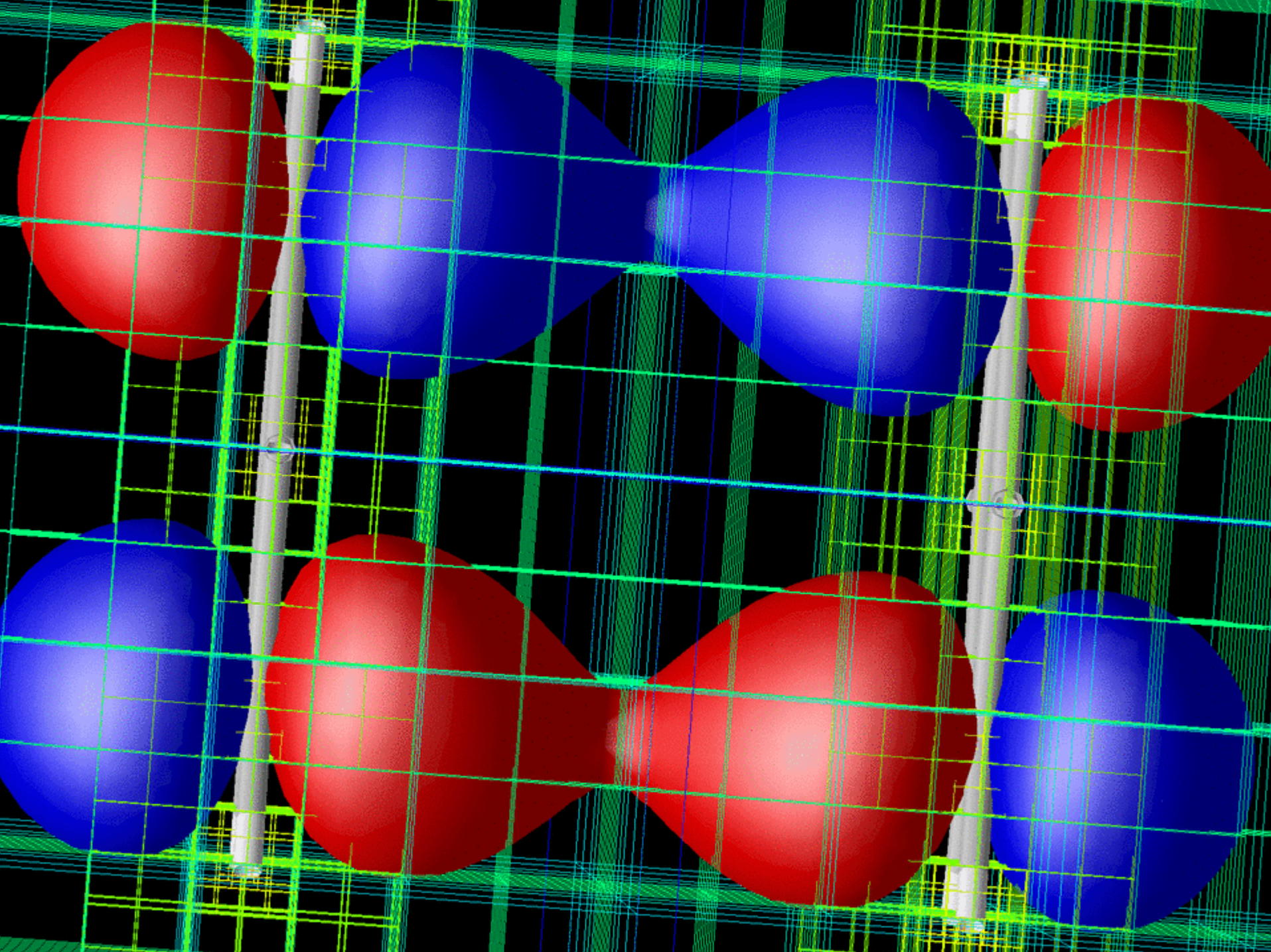
UT-BATTELLE
Oak Ridge National Laboratory

MADNESS

Multiresolution
Adaptive
Numerical
Scientific
Simulation

Multiresolution chemistry objectives

- Complete elimination of the basis error
  - One-electron models (e.g., HF, DFT)
  - Pair models (e.g., MP2, CCSD, …)
- Correct scaling of cost with system size
- General approach
  - Readily accessible by students and researchers
  - Higher level of composition
  - No two-electron integrals – replaced by fast application of integral operators
- New computational approaches
- *Fast algorithms with guaranteed precision*



Synthesis of High Performance Algorithms for Electronic Structure Calculations

http://www.cis.ohio-state.edu/~gb/TCE

- Collaboration between DOE/SciDAC, NSF/ITR and ORNL/LDRD
- Objective: *develop a high level programming tool that translates many-body quantum theory into efficient massively parallel codes. This is anticipated to revolutionize the rate of progress in this field by eliminating man-years of programming effort.*
- NSF Project:

  Sadayappan (PI), Baumgartner, Cociorva, Pitzer (OSU)
  Bernholdt, Harrison (unfunded) (ORNL)
  Ramanujam (LSU)
  Nooijen (Waterloo)
- DOE SciDAC: Harrison (PI), Hirata (PNNL)
- DOE ORNL/LDRD: Bernholdt (PI, 2002-3)
- Other SciDAC projects adopting this tool: Piecuch, Gordon

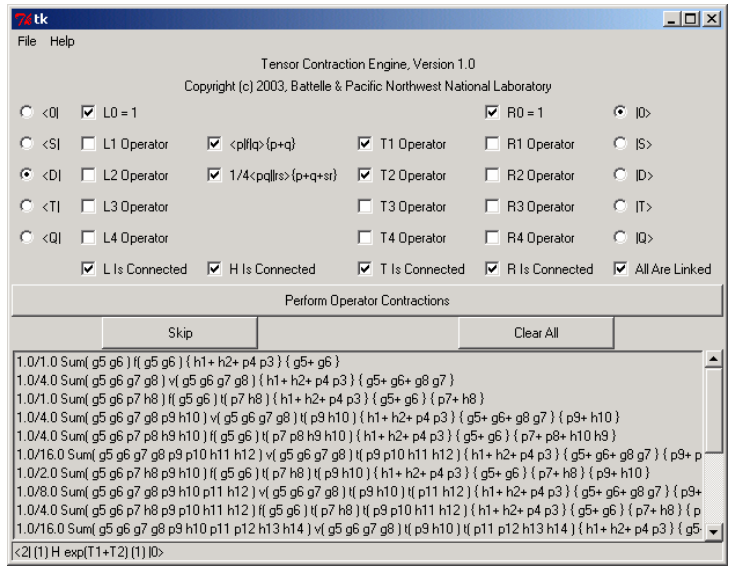
Tensor Contraction Engine

Definition of a many-electron theory

$$E = \left\langle \Phi_0 \left| \left[e^{-T_1-T_2} H e^{T_1+T_2} \right]_C \right| \Phi_0 \right\rangle$$

$$0 = \left\langle \Phi_i^a \left| \left[e^{-T_1-T_2} H e^{T_1+T_2} \right]_C \right| \Phi_0 \right\rangle$$

$$0 = \left\langle \Phi_{ij}^{ab} \left| \left[e^{-T_1-T_2} H e^{T_1+T_2} \right]_C \right| \Phi_0 \right\rangle$$



Mathematical expressions

$$(\Xi_1)_{h_1p_9}^{h_7h_{10}} = +t_{h_1}^{p_5} v_{p_5p_9}^{h_7h_{10}}$$

$$(\Xi_2)_{h_1h_2}^{h_{10}h_{11}} = -\tfrac{1}{2} t_{h_1h_2}^{p_7p_8} v_{p_7p_8}^{h_{10}h_{11}}$$

$$(\Xi_3)_{h_1p_5}^{h_{10}p_3} = -\tfrac{1}{2} t_{h_1}^{p_6} v_{p_5p_6}^{h_{10}p_3}$$

$$(\Xi_4)_{p_5}^{h_{10}} = -t_{h_7}^{p_6} v_{p_5p_6}^{h_7h_{10}}$$

$$(\xi_{222})_{h_1p_5}^{h_{10}h_{11}} = +v_{h_1p_5}^{h_{10}h_{11}} + \tfrac{1}{2} (\Xi_1)_{h_1p_5}^{h_{10}h_{11}}$$

$$(\xi_{22})_{h_1h_2}^{h_{10}h_{11}} = -v_{h_1h_2}^{h_{10}h_{11}} + P_2 t_{h_1}^{p_5} (\xi_{222})_{h_2p_5}^{h_{10}h_{11}} + (\Xi_2)_{h_1h_2}^{h_{10}h_{11}}$$

$$(\xi_{23})_{h_1p_5}^{h_{10}p_3} = +v_{h_1p_5}^{h_{10}p_3} + (\Xi_3)_{h_1p_5}^{h_{10}p_3}$$

$$(\xi_{24})_{p_5}^{h_{10}} = +f_{p_5}^{h_{10}} + (\Xi_4)_{p_5}^{h_{10}}$$

$$(\xi_{25})_{h_1p_9}^{h_7h_{10}} = +v_{h_1p_9}^{h_7h_{10}} + (\Xi_1)_{h_1p_9}^{h_7h_{10}}$$

$$(\xi_2)_{h_1h_2}^{h_{10}p_3} = +v_{h_1h_2}^{h_{10}p_3} + \tfrac{1}{2} t_{h_{11}}^{p_3} (\xi_{22})_{h_1h_2}^{h_{10}h_{11}} - P_2 t_{h_1}^{p_5} (\xi_{23})_{h_2p_5}^{h_{10}p_3}$$
$$- t_{h_1h_2}^{p_3p_5} (\xi_{24})_{p_5}^{h_{10}} + P_2 t_{h_1h_7}^{p_3p_9} (\xi_{25})_{h_2p_9}^{h_7h_{10}} + \tfrac{1}{2} t_{h_1h_2}^{p_5p_6} v_{p_5p_6}^{h_{10}p_3}$$

$$(\xi_3)_{h_1p_5}^{p_3p_4} = +v_{h_1p_5}^{p_3p_4} - \tfrac{1}{2} t_{h_1}^{p_6} v_{p_5p_6}^{p_3p_4}$$

$$(\xi_{42})_{p_8}^{h_9} = +f_{p_8}^{h_9} + (\Xi_4)_{p_8}^{h_9}$$

$$(\xi_4)_{h_1}^{h_9} = +f_{h_1}^{h_9} + t_{h_1}^{p_8} (\xi_{42})_{p_8}^{h_9} - t_{h_7}^{p_6} v_{h_1p_6}^{h_7h_9} - \tfrac{1}{2} t_{h_1h_8}^{p_6p_7} v_{p_6p_7}^{h_8h_9}$$

$$(\xi_5)_{p_5}^{p_3} = +f_{p_5}^{p_3} - t_{h_7}^{p_6} v_{p_5p_6}^{h_7p_3} - \tfrac{1}{2} t_{h_7h_8}^{p_3p_6} v_{p_5p_6}^{h_7h_8}$$

$$(\xi_{62})_{h_1p_8}^{h_9h_{11}} = +v_{h_1p_8}^{h_9h_{11}} + \tfrac{1}{2} (\Xi_1)_{h_1p_8}^{h_9h_{11}}$$

$$(\xi_6)_{h_1h_2}^{h_9h_{11}} = -v_{h_1h_2}^{h_9h_{11}} + P_2 t_{h_1}^{p_8} (\xi_{62})_{h_2p_8}^{h_9h_{11}} + (\Xi_2)_{h_1h_2}^{h_9h_{11}}$$

$$(\xi_7)_{h_1p_5}^{h_6p_3} = +v_{h_1p_5}^{h_6p_3} + 2 (\Xi_3)_{h_1p_5}^{h_6p_3} - \tfrac{1}{2} t_{h_1h_8}^{p_3p_7} v_{p_5p_7}^{h_6h_8}$$

Hirata, PNNL

A parallel computer program

```
 nprocs = GA_NNODES()
 count = 0
 next = NXTVAL(nprocs)
 DO p3b = noab+1,noab+nvab
 DO h6b = 1,noab
 DO h1b = 1,noab
 DO p5b = noab+1,noab+nvab
 IF (next.eq.count) THEN
 IF ((.not.restricted).or.(int_mb(k_spin+p3b-1)+int_mb(k_spin+h6b-1
&)+int_mb(k_spin+h1b-1)+int_mb(k_spin+p5b-1).ne.8)) THEN
 IF (int_mb(k_spin+p3b-1)+int_mb(k_spin+h6b-1) .eq. int_mb(k_spin+h
&1b-1)+int_mb(k_spin+p5b-1)) THEN
 IF (ieor(int_mb(k_sym+p3b-1),ieor(int_mb(k_sym+h6b-1),ieor(int_mb(
&k_sym+h1b-1),int_mb(k_sym+p5b-1)))) .eq. ieor(irrep_v,irrep_t)) TH
&EN
 DO p7b = noab+1,noab+nvab
 DO h8b = 1,noab
 IF (int_mb(k_spin+p3b-1)+int_mb(k_spin+p7b-1) .eq. int_mb(k_spin+h
&1b-1)+int_mb(k_spin+h8b-1)) THEN
 IF (ieor(int_mb(k_sym+p3b-1),ieor(int_mb(k_sym+p7b-1),ieor(int_mb(
&k_sym+h1b-1),int_mb(k_sym+h8b-1)))) .eq. irrep_t) THEN
 IF ((restricted).and.(int_mb(k_spin+p3b-1)+int_mb(k_spin+p7b-1)+in
&t_mb(k_spin+h1b-1)+int_mb(k_spin+h8b-1).eq.8)) THEN
```

Chemistry and Materials future computing requirements - I

- More ... lots more ... at least 100x more by CY2006
  - Current allocations are much too small to support current science let alone explore new frontiers
- Capacity rather than capability is the big limitation
  - Unlike some other disciplines, we need to run only a few very large benchmark calculations for calibration, and then run a much larger number of fewer-process jobs
  - Each job might run a long time, but not scale to more than ~100 processors
- Capacity computing might cost 2-3x less per peak FLOP than capability computing
  - Materials and chemistry actually run efficiently (good fraction of peak cpu speed) on many commodity systems

Chemistry and Materials future computing requirements - II

- Ratio of capability to capacity requirements
  - 1:1 to 1:10 depending upon who you ask
  - Sensitive to actual target applications
- Why should NERSC provide capacity computing?
  - Cost savings to DOE from scale, eliminating redundancy across sites, better scheduling of load, continuous technology refresh, sharing of expertise, ...
  - Co-location with capability resources for data sharing
  - “Sexy” peak TFLOP numbers ... 100-1000TF in
  - Usable as capability resource by some apps.
  - End-to-end execution of entire scientific project ... ease of use for users, “ownership” for NERSC

Specific machine characteristics

- Low-latency & high-bandwidth inter-process communication
  - Objectives: calculations scale to more processors; good performance for irregular computations; increased programmer productivity
    - E.g., long-time classical and ab initio dynamics
    - Future calculations will be increasingly irregular and use sparse data structures
    - Poorly balanced machines require much more effort from programmers to obtain good performance

Specific machine characteristics - II

- Faster processors that are near state-of-the-art when purchased
  - Objectives: decrease wall time of simulations; reduced need to scale to 1000+ processors; longer lifetime of the system
  - E.g., except for DGEMM, about 6 Power3 processors equals 1 Opteron (Spec CPU 2000FP)
  - 1000 Power3 → 166 Opterons

Specific machine characteristics - III

- Some nodes with larger memories and greater I/O capacity
  - Objective: Many-body electronic structure calculations in particular need much more memory and I/O for intermediates (some integrals, iterative subspace information) ... some data cannot be efficiently distributed
  - Probably not cost effective to configure the whole machine in this fashion
- High-performance I/O on scratch data does not have to be to a shared file system
  - Pick the most cost effective solution

Specific machine characteristics - IV

- Balance costs $$ ... let the science drive the choice between FLOPs/s and bytes ...
  - To date FLOPs has always won out against scientific requirements
  - NERSC cannot just look at current and historic usage to estimate the future requirements ...
- Why increasingly few chemists compute at NERSC
  - Chemists no longer need supercomputers? NO!
  - Chemists have enough computer resources elsewhere? NO!
  - Chemists don't have very scalable codes? True for some types of simulation, but there are codes such as NWChem and GAMESS that do scale well and are efficient for many common methods.
  - It is impossible to get a large allocation of time? YES!
  - The current resources at NERSC are not well balanced for chemistry ... slow I/O & communications, limited memory? YES!

Challenges for DOE

- Improved complex wide management and allocation of resources
  - Coordinate allocation of capability and capacity resources
- Improved understanding of actual computational requirements, not just the “sexy” hero-calculations
  - Capacity computing (in appropriate proportion to capability) not only can, but is *required* to push scientific frontiers
- Enable computer time allocations at scale
  - NERSC should be encouraging and supporting *only* projects that users cannot do at home, even if this means being more selective
  - E.g., all supported projects should be 1+M node hours or have other special requirements (exceptions are inevitable)
  - Insight should be allocating projects in the 10 M node hour range (for a modern computational node)