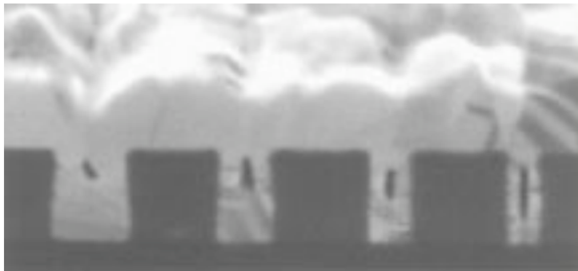
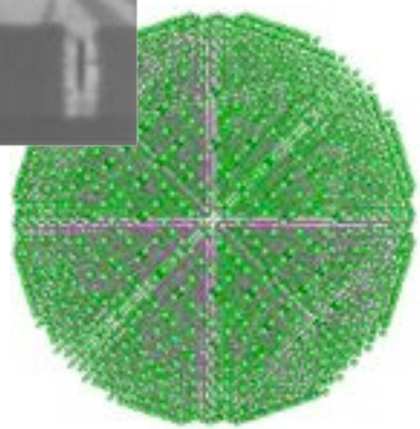
Computational Mathematics: Role, Impact, Challenges

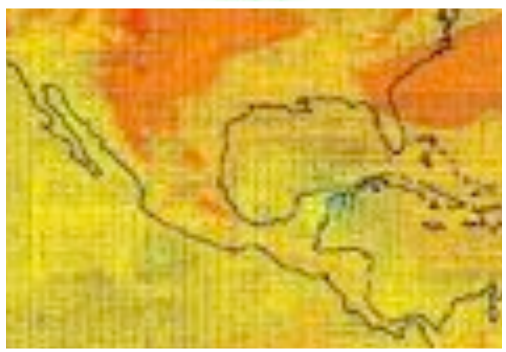
Juan Meza
Department Head
Lawrence Berkeley National Laboratory

NCN Excellence in Computer Simulation, Berkeley, CA   November 6, 2007

Advances across the board in computer simulation







- Level-set methods used to improve industrial and medical processes
- Improved computational methods for understanding material science properties of nanostructures
- Combustion simulations lead to a better understanding of turbulent flames
- Better earth systems modeling due to increased resolution models
- New data management techniques improve the efficiency of searching databases

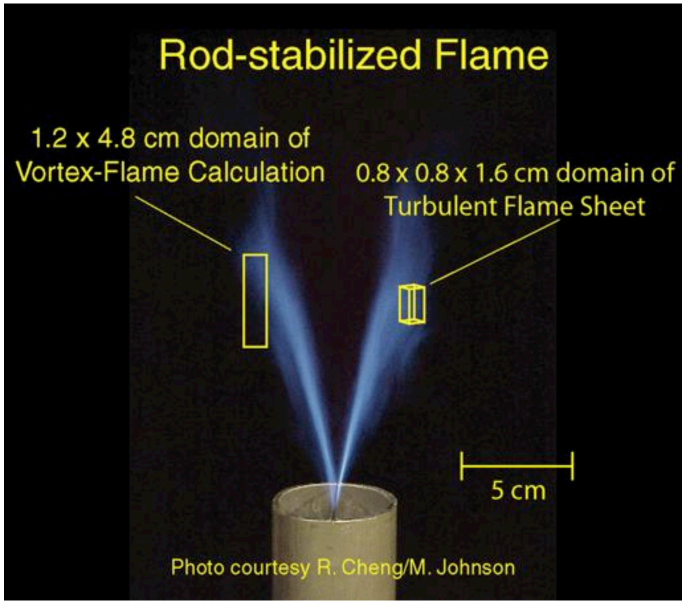
Enabling Mathematical Technologies

- Improved linear equation solvers directly leading to higher fidelity/resolution models
- Better understanding of discretization and higher-order methods
- Interior point revolution leading to more efficient optimization techniques and also allowing for wider use of optimization
- Implementation of all of these techniques in easy to use, robust, and portable math libraries
  - BLAS, Linpack, EISPACK,
  - LAPACK, ScaLAPACK, SuperLU
  - FFTW
  - AMR packages, e.g. Chombo, BoxLib

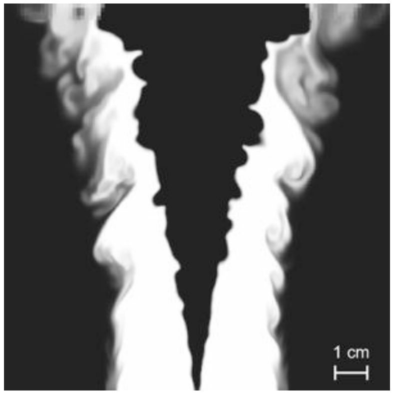
Simulation capability increased by a factor of 10,000

- Advances in applied mathematics have dramatically changed what we can accurately simulate
  - Mathematical formulation to exploit separation of scales
  - Specialized discretization methods
  - Adaptive mesh refinement
  - Parallel numerical algorithms

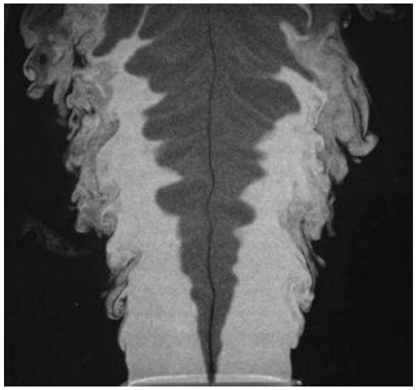
- These advances have improved computational efficiency by a factor of more than 10,000 allowing laboratory scale simulations



Simulation

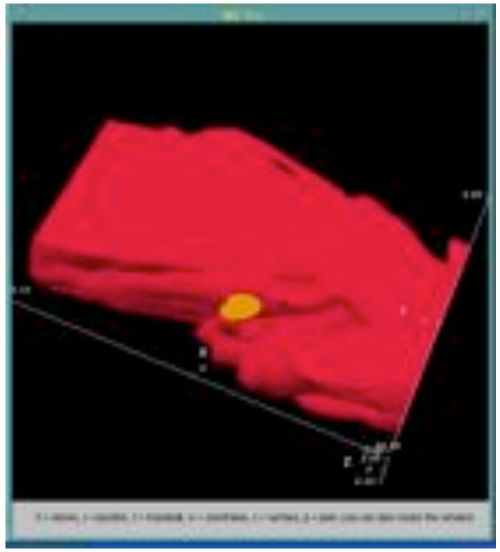


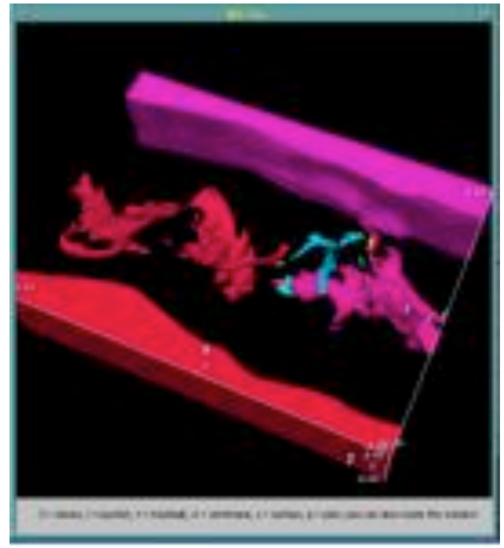
Experiment



Query-driven visualization of combustion data set

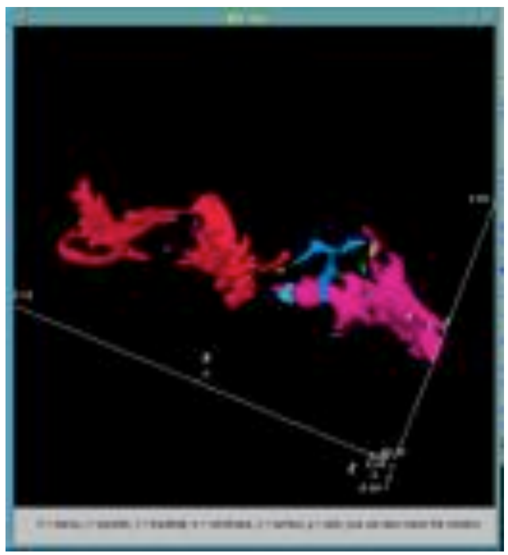
a) Query: CH4 > 0.3

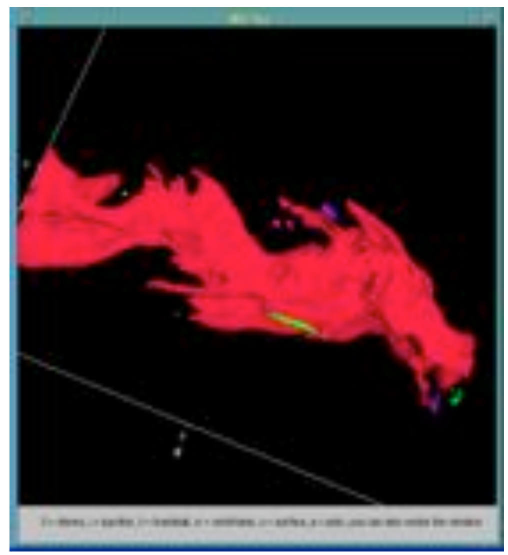




b) Q: temp < 3

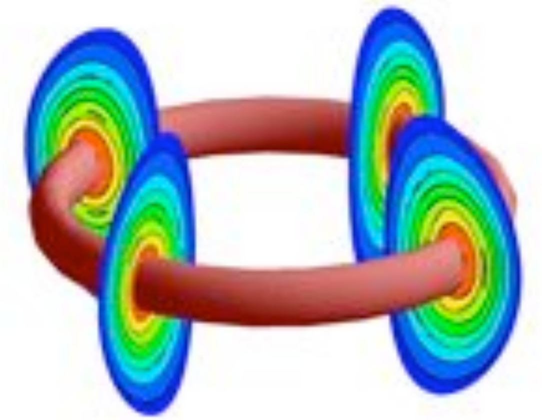
c) Q: CH4 > 0.3 AND temp < 3





d) Q: CH4 > 0.3 AND temp < 4

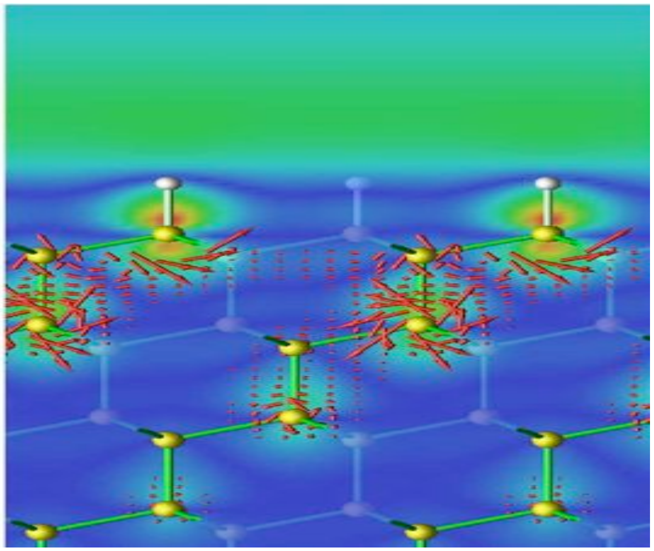
Parallel SuperLU Speeds Up Fusion Code



- NIMROD is a parallel simulation code for fluid-based modeling of nonlinear macroscopic electromagnetic dynamics in fusion plasmas.
- Joint work between CEMM and TOPS has led to an improvement in NIMROD by a factor of 5-10 on NERSC IBM SP, equivalent of 3-5 years progress in computing hardware.

- Parallel SuperLU, developed at LBNL, has been incorporated into NIMROD as an alternative linear solver.
  - SuperLU is >100x and 64x faster on 1 and 9 processors.
  - A much larger linear system has to be solved using the conjugate gradient method in the last time-advance.
  - SuperLU is used to factor a preconditioning matrix for the conjugate gradient iterations, resulting in a 10x speedup.

Material Science: PARATEC



- PARATEC performs first-principles quantum mechanical total energy calculation using pseudopotentials & plane wave basis set
- Density Functional Theory to calculate structure & electronic properties of new materials
- DFT calculations are one of the largest consumers of supercomputer cycles in the world

- PARATEC uses all-band CG approach to obtain wavefunction of electrons
- Part of calculation in real space, other in Fourier space using specialized 3D FFT to transform wavefunction
- Generally obtains high percentage of peak on different platforms
- Developed with Louie and Cohen's groups (UCB, LBNL), Raczkowski

PARATEC: Performance

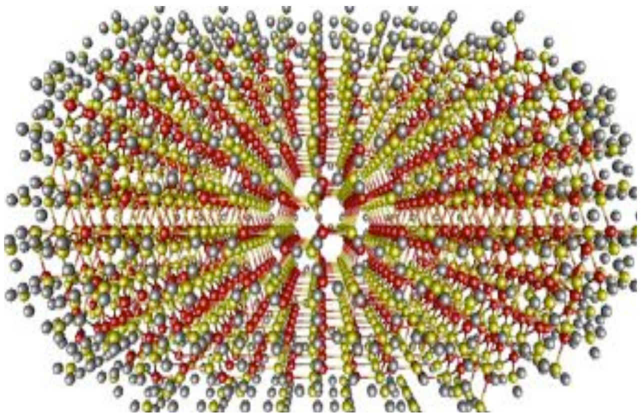
Data Size	P	Power 3		Power4		Altix		ES		X1	
		Gflops/P	% peak	Gflops/P	% peak	Gflops/P	% peak	Gflops/P	%peak	Gflops/P	% peak
432 Atom	32	0.95	63%	2.0	39%	3.7	62%	4.7	60%	3.0	24%
	64	0.85	57%	1.7	33%	3.2	54%	4.7	59%	2.6	20%
	128	0.74	49%	1.5	29%	---	---	4.7	59%	1.9	15%
	256	0.57	38%	1.1	21%	---	---	4.2	52%	---	---
	512	0.41	28%	---	---	---	---	3.4	42%	---	---
686 Atom	128							4.9	62%	3.0	24%
	256							4.6	57%	1.3	10%

- Most architectures achieve high performance
- Performance due to computational intensity of code (BLAS3, FFT)
- X1 shows lowest % of peak due to non-vectorizable code

Office of Science
U.S. DEPARTMENT OF ENERGY

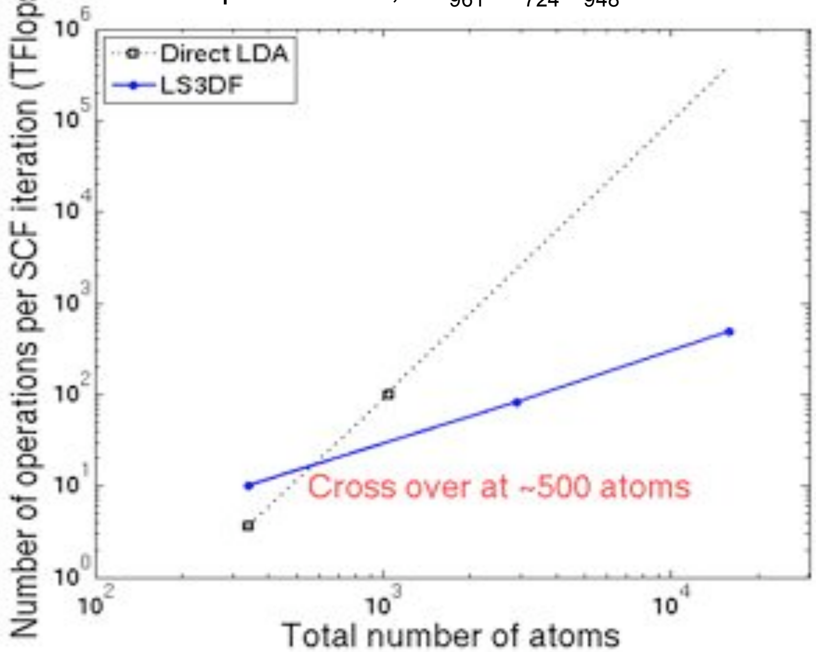
BERKELEY LAB

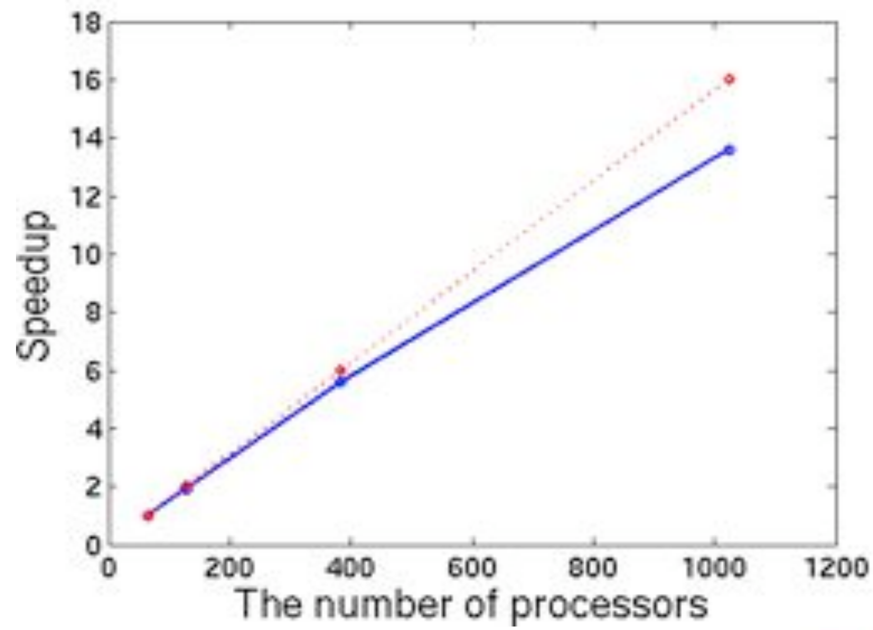
Dipole Moment calculation



The calculated dipole moment of a 2633 atom CdSe quantum rod, $Cd_{961}Se_{724}H_{948}$

- New linear scaling DFT algorithm allows simulation of larger systems
- Using 2560 processors at NERSC, calculation took about 30 hours
- Cross over occurs around 500 atoms





Some Lessons Learned

- Many success stories
- More interaction needed between computational mathematics and application scientists and experimentalists
- Better software infrastructure needed
  - Too much re-inventing the wheel
  - Debugging is a nightmare (e.g. C++ templates)
  - Better software engineering practices
- Should consider whether you want a more computationally efficient slower algorithm or a more scalable algorithm - time to solution best metric

New Frontiers

- Data, Data, Data
  - New field of Computational ***Data*** Sciences?
- Massively parallel; soon >1 million processors
  - Need architecture aware algorithms
- Probabilistic - the new “iterative”?
  - “*overwhelming probability*”
- Uncertainty quantification
  - Earth systems modeling
- Optimization
  - Better operating points
  - Comparison of competing designs
  - Prediction of experimental properties

Thank you

Office of Science
U.S. DEPARTMENT OF ENERGY

BERKELEY LAB