SAND2011-3571C

Exascale computing: An application perspective

Intel HPC Roundtable
Exascale Panel
May 25, 2011

Curtis Janssen
Sandia National Laboratories
Livermore, CA
cljanss@sandia.gov

Chemistry Contributions: Robert Harrison et al.
Combustion Contributions: Jackie Chen et al.

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company,
for the United States Department of Energy's National Nuclear Security Administration
under contract DE-AC04-94AL85000.

Sandia National Laboratories

Co-design is a strategy to deliver exascale hardware/software

- Exascale will not be business as usual
  - Applications will have to change:
    - Billions of threads, faults, energy to solution, and time to solution are all critical issues
  - Lower hardware design margins
    - Provide what the application needs but no more
- An integrated *exascale co-design* effort is needed
  - Allow evaluation of cost-benefit trade-off in a coupled hardware and software development process
- DOE/ASCR is establishing domain-specific areas for exascale development
  - Chemistry, magnetic fusion, high energy density physics, materials, climate, nuclear energy, combustion
- Similar activities in NNSA/ASC.

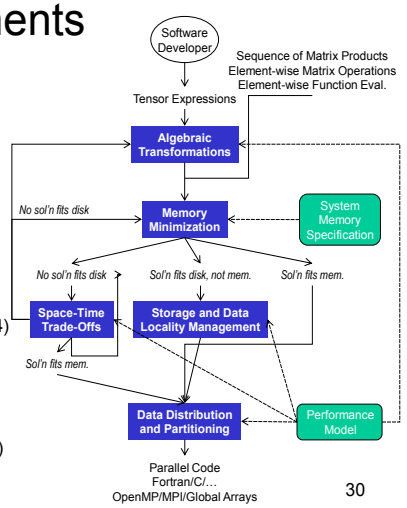
Chemistry co-design drivers (1)

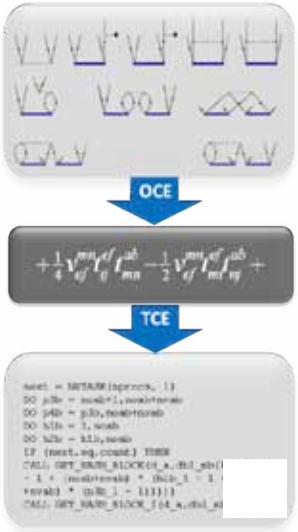
- Chemistry has unique co-design issues: CECC is the only team in this space
  - We (mostly) don't use grids – atom-centered functions (3D), $f(\mathbf{r}_1\text{-}\mathbf{r}_2)$ (6D), and other (nD)
  - We (mostly) don't use MPI – Global Arrays and similar models dominate scalable codes
  - We extensively employ DSLs and code transformation – essential for science, boosts productivity, rapidly targets new architectures, injecting instrumentation – IBM, Intel, etc., *very* enthusiastic
- Algorithms to solve many-body Schrödinger equation – $O(10^7)$ LOC
  - *Background:* Lots of data but many more FLOPs, block-sparse DGEMM-like kernels
  - *Science objective:* Simulating energy applications accurately with validation+UQ
  - *Issues:* Resilience, async. comm. (RMA+RMI), memory, reduce algorithm scaling, power

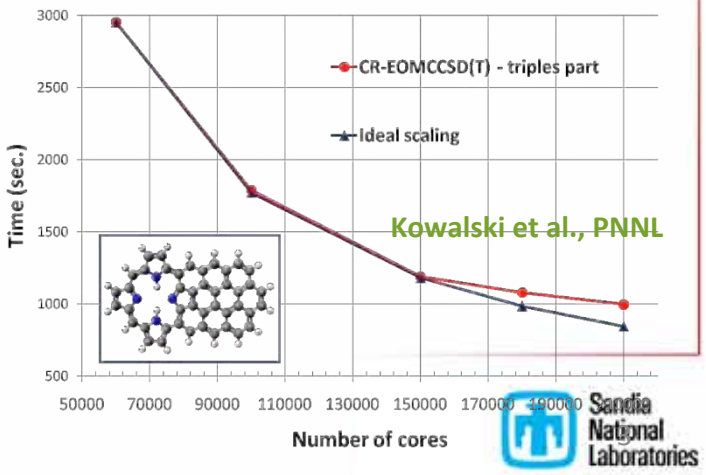
Our Tensor Contraction Engine (DSL) now defines the state of the art in many-body simulation.

TCE Components

- Algebraic Transformations
  – Minimize operation count
- Memory Minimization
  – Reduce intermediate storage via loop fusion (LCPC'03)
- Space-Time Transformation
  –Trade-offs between storage and recomputation (PLDI'02)
- Data Locality Optimization
  – Optimize use of storage hierarchy via tiling (ICS'01, HiPC'03, IPDPS'04)
- Data Dist./Comm. Optimization
  – Optimize parallel data layout (IPDPS'03)
- Integrated System (SuperComputing'02, Proc. IEEE 05)

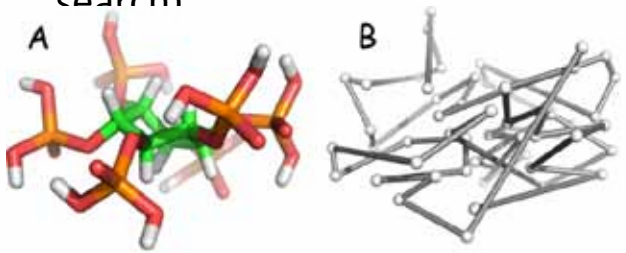




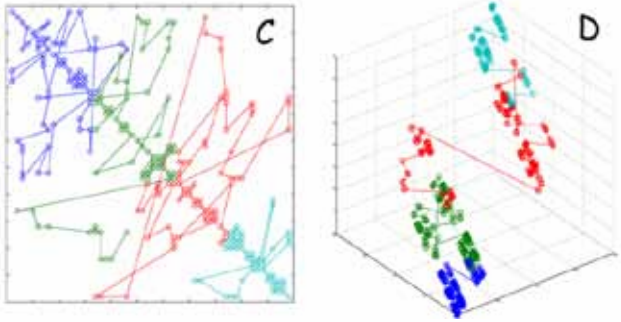


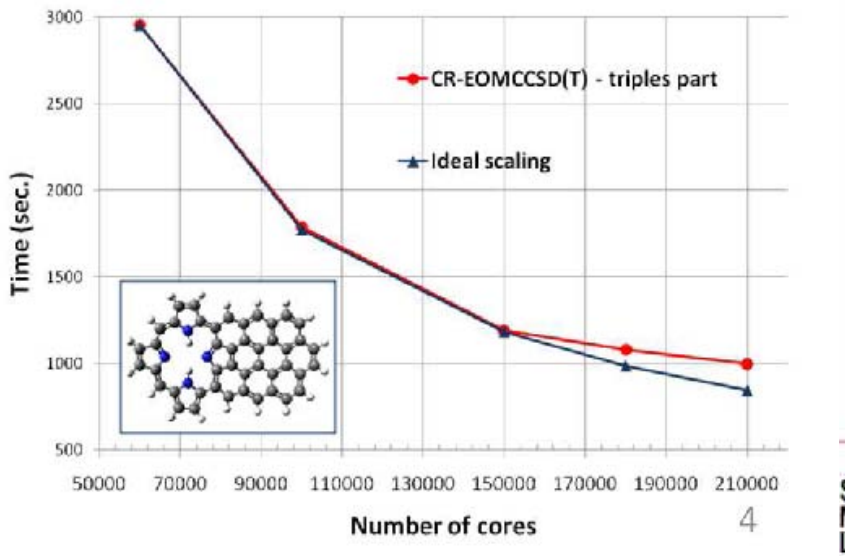
Chemistry co-design drivers (2)

- Robust and power efficient algorithms for one-body Schrodinger – O(10^5) LOC
  - *Background:* Density functional theory in atomic orbitals, block-sparse trees with fast summation
  - *Science objective:* Run at scaling limit for thermodynamic integration of energy-related materials
  - *Issues:* Interconnect, power, resilience, scaling, numerical robustness, at scaling limit data motion dominates, irregular and small non-square matrices
- Efficient and resilient algorithms to evaluate two-electron integrals – O(10^5) LOC
  - *Background:* Multiple algorithms – recursion, special functions, quadrature; near min.op. algorithms obtain ~40% peak on x86-64, but no satisfactory solution yet on current accelerators
  - *Science objective:* Increased accuracy and speed, more types of bases and integral
  - *Issues:* CPU/memory architecture, resilience, power, optimal algorithm hard to find (graph search)



Quantum locality can be exploited for data- and load-balancing via space-filling curves, from atoms (A-B) through matrices (C) to the product space (D).





Sandia National Laboratories

Combustion co-design drivers

Three central co-design efforts:

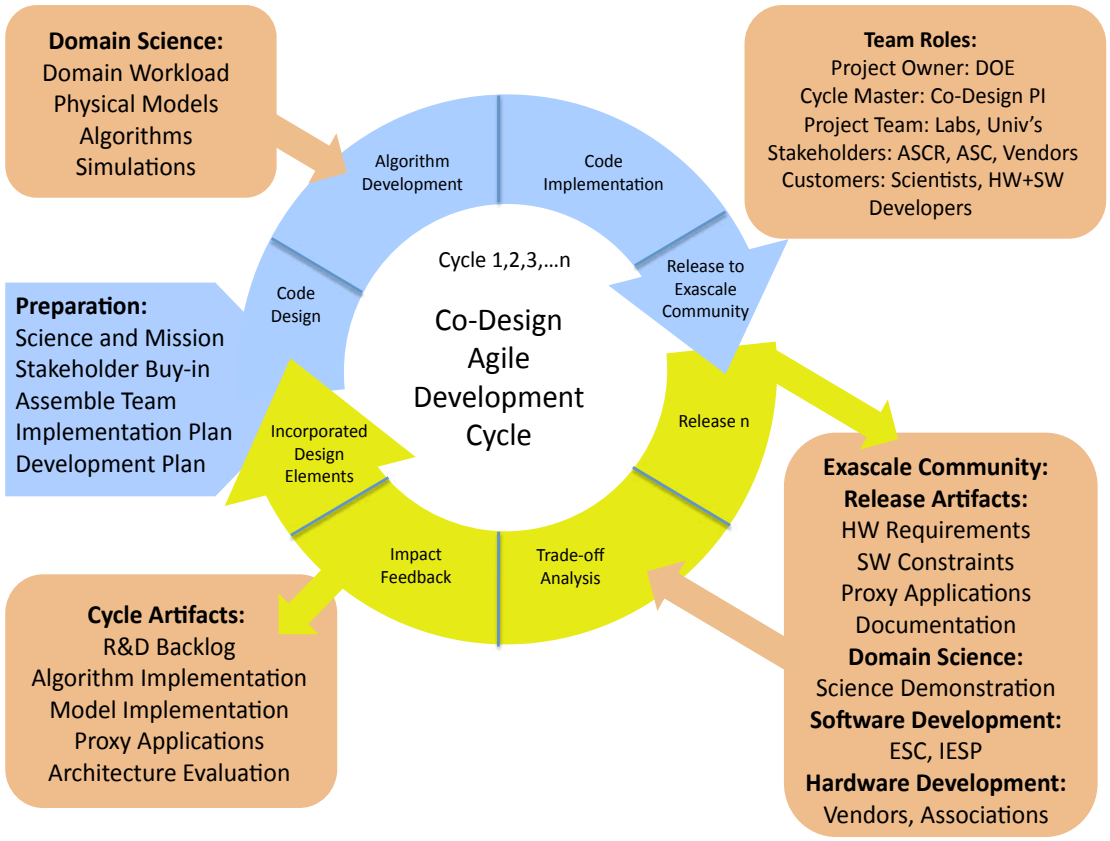
- **PDE methodology for direct numerical simulation**
  - Require AMR to meet spatial resolution requirements
  - Must support both low Mach number and fully compressible formulations
- ***In situ* data analytics**
  - Data rates too high for deferred analysis – emphasis on data reduction and steering
  - Support for data layout, volume and particle visualization, topological feature tracking, pathlines, local flame coordinates, feature-base statistics
- **Embedded uncertainty quantification (UQ)**
  - Impact of uncertainty of chemical parameters on predictive capability
  - Quantitative comparisons with experimental data

Research areas

- How can programming models be used to: exploit fine-grained parallelism in PDE algorithms, express data movement vs. floating point operations in designing numerical algorithms, expose issues of fault tolerance and energy use?
- What is the most effective strategy for in situ data analysis? Shared work on nodes or staging? Use scratchpad memory for analysis? How do we balance simulation with analysis? What is the optimal data structure?
- How can we best formulate UQ problems for complex multiphysics problems with "chaotic" dynamics such as turbulent combustion? Hardware support?

Sandia National Laboratories

Co-design implementation: Spiral model with evolving software and hardware surrogates



Exascale Co-Design Consortium (ECDC) Whitepaper on Application-Driven Co-Design December 7, 2010

Software Surrogates

- **Compact Apps**
  - small program capturing some aspect of full app. Generates a result.
- **Mini Apps**
  - small program capturing simplified aspect of full app. Perhaps no meaningful result.
- **Skeleton Apps**
  - captures control flow and communication pattern of app. Runs in simulator.
- **Kernels**
  - Capture node-level aspects of an algorithm.

Hardware Surrogates

- Cycle-accurate emulators
- Coarse-grained simulators
- Testbeds/proto-exascale machines
- Both open-source and proprietary

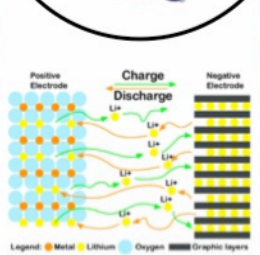
Sandia National Laboratories

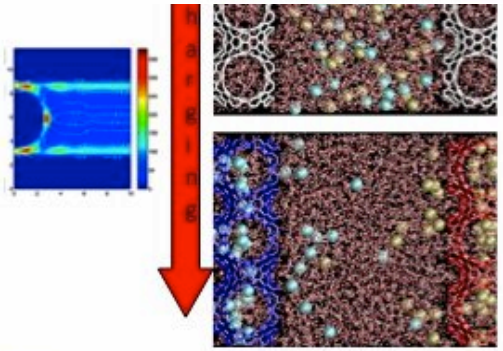
Supplemental Slides

Sandia National Laboratories

CECC Science Theme

The central science theme of the CECC is understanding, controlling, and ultimately designing chemically and electrically active interfaces that are relevant to diverse battery technologies, ultra-capacitors, fuel cells, environmental chemistries, and catalytic processes for sustainable energy conversion including biomass conversion.





Center for Molecular Electrocatalysis (PNNL)

Center for Electrical Energy Storage (ANL)

Argonne-Northwestern Solar Energy Research Center (Northwestern)

HeteroFoaM Center (USC)

Center for Defect Physics In Structural Materials (ORNL)

CINT (LANL)

CNM (ANL)

CNMS (ORNL)

BATT Program (EERE)

CECC directly coordinates with BES and EERE centers to advance their science

Center of Chemical Hydrogen Storage Center of Excellence (PNNL)

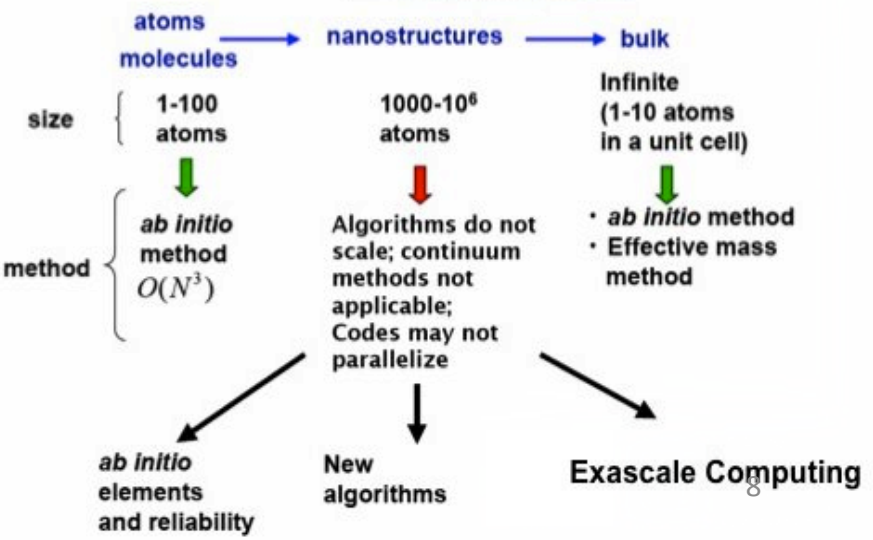
Center for Atomic-Level Catalyst Design (LSU)

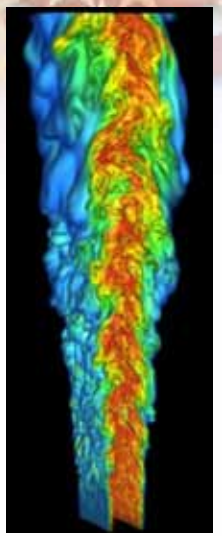
Center on Nanostructuring for Efficient Energy Conversion (Stanford)

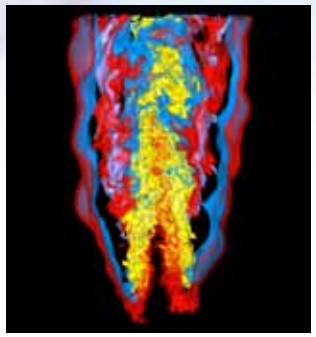
Fluid Interface Reactions, Structures and Transport Energy (ORNL)

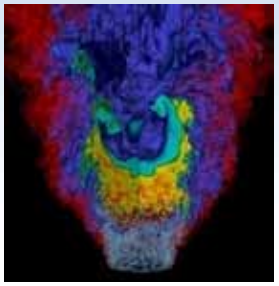
Polymer-Based Materials For Harvesting Solar Energy (Umass)

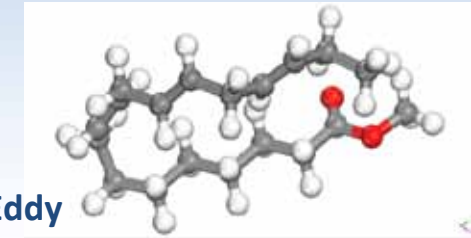
Institute for Atom-Efficient Chemical Transformations (ANL)











Exaflops Large-Eddy Simulation (LES)

Petaflops Sub-Model Development

Bio-Diesels

Methyl Decanoate

Teraflops Direct Numerical Simulation (DNS)

Methyl Butanoate

Alcohols

Unsteady Larminar Flame Simulations

Butanol

Gigaflops

Turbulent Flame Experiments

Di-methyl Ether

Quantum chemical Kinetics Simulations

Foundational fuels (C1-C4)

Mechanism Validation Experiments

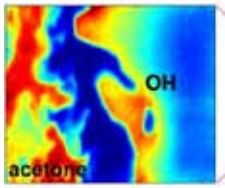
Ethanol

Ethylene

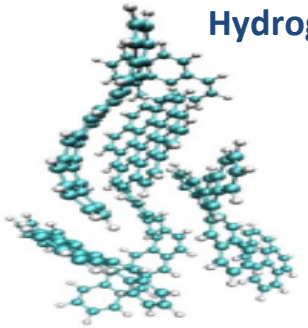
Methane

Laminar Flame/Ignition Experiments

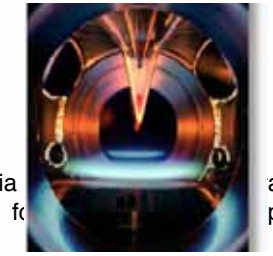
Syngas

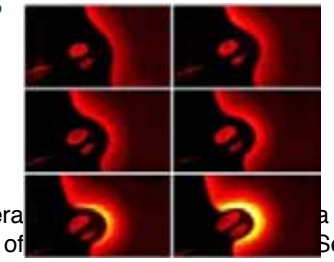


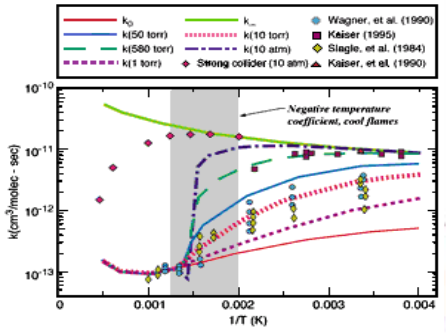
Hydrogen



Chemical Kinetic Experiments







Sandia
atory opera
a Lo
partment of
Sec
under contract DE-AC04-94AL85000.

Sandia National Laboratories