Overview

Timeline
- Ongoing project with yearly direction from DOE

Budget
- FY09 funding: $1M
- FY10 funding: $1M

Split among 3 projects:
  - Combustion Numerics
  - HECC Simulation
  - Exhaust Speciation

Barriers
- Inadequate understanding of the fundamentals of HECC
- Inadequate understanding of the fundamentals of mixed mode operation
- Computational expense of HECC simulations

Partners
- Sandia, Oak Ridge, Los Alamos
- Ford
- UC Berkeley, University of Wisconsin, University of Michigan, Lund Institute of Technology, Chalmers University, Tianjin University
- FACE working group, AEC MOU, SAE
Relevance to DOE objectives

- By 2015, improve the fuel economy of light-duty gasoline vehicles by 25 percent and of light-duty diesel vehicles by 40 percent, compared to the baseline 2009 gasoline vehicle.
  - Light-duty research focuses on reducing fuel consumption through investigating HCCI and PCCI part load, and transition to SI or CIDI for full load operation

- By 2015, improve heavy truck efficiency to 50 percent with demonstration in commercial vehicle platforms. This represents about a 20 percent improvement over current engine efficiency.
  - Heavy-engine research directed towards high efficiency strategies, such as Partially Premixed Combustion and Low-temperature Diesel Combustion

- By 2018, further increase the thermal efficiency of a heavy truck engine to 55 percent which represents about a 30 percent improvement over current engines.
  - We continue to provide the engine research community with insight and simulation tools for advanced combustion concepts
Objective: Enhance understanding of clean and efficient engine operation through detailed numerical modeling

Chemical kinetics

Fluid mechanics
Milestones: We have developed and experimentally validated detailed engine modeling tools

- **11x speedup demonstrated for GPU zero-dimensional ignition delay calculations (Feb. 2011)**

- **Developed and Deployed Parallel Multi-zone Combustion Model in Kiva4-mpi (Dec. 2010)**

- **Completed License agreement for Multi-zone Combustion Model with Convergent Science Inc. (Oct. 2010)**

- **Up to 300x speedup demonstrated for GPU based thermodynamic property calculations (Aug. 2010)**
Approach: Collaborate with industry, academia and national labs in the development of analysis tools leading to clean, efficient engines

- Gain fundamental and practical insight into HECC regimes through numerical simulations and experiments

- Develop and apply numerical tools to simulate HECC by combining multidimensional fluid mechanics with chemical kinetics

- Reduce computational expense for HECC simulations

- Democratize simulation: bring computational tools to the Desktop PC
Technical Accomplishments: We have made significant progress in improving and applying our advanced simulation tools to HECC.

- Improved Numerics
  - KIVA4-mpi multi-zone development

- New Computing Architectures
  - Parallel CFD and MZ-chemistry
Chemical kinetics involves solving systems of non-linear differential equations.

The Newton-Raphson method efficiently solves nonlinear equations. For 1 equation, 1 unknown, the formula is:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$
When solving a system of differential equations, the **Jacobian matrix** plays the role of the derivative.

\[ J = \frac{\partial f_i}{\partial y_j} \]

The Jacobian is the key player in solving chemical kinetics ODEs.
>95% of the chemistry solution CPU cost is spent constructing and solving the Jacobian system

During ignition:

\[ \Delta t \text{ (explicit)} = 10^{-12} \text{ to } 10^{-15} \text{ s} \]
\[ \Delta t \text{ (implicit)} = 10^{-6} \text{ to } 10^{-8} \text{ s} \]

Explicit Update
(lower cpu/step)

\[
\frac{\partial x_1}{\partial t} = f_1(t, x_1, \ldots, x_N) \\
\frac{\partial x_2}{\partial t} = f_2(t, x_1, \ldots, x_N) \\
\vdots \\
\frac{\partial x_N}{\partial t} = f_N(t, x_1, \ldots, x_N).
\]

Implicit Update
(more trajectory data)

\[
J = \begin{pmatrix}
\frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_N} \\
\frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_N} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial f_N}{\partial x_1} & \frac{\partial f_N}{\partial x_2} & \cdots & \frac{\partial f_N}{\partial x_N}
\end{pmatrix}
\]
Opportunities for 1000x speedup in computational chemistry cost through applied mathematics

- Perturbation methods
- Eigenstructure Analysis
- Pre-conditioners
- Sparse Solvers
- Jacobian Reuse
- Adaptive Sampling
- Mode Splitting
- Multi-zone ODE system
Constant volume ignition delay is the basis for numerical chemistry development.

The Constant Volume Reactor is the basic unit for chemistry in multidimensional CFD codes.

\[ T_0 = 1000 \, \text{K} \]
\[ P_0 = 20 \, \text{bar} \]
Applied mathematics techniques identify opportunities for improved solver convergence

Eigenvalues of the Jacobian determine how quickly an ODE solution will converge

- **Fast convergence**
  - ODEs with tightly clustered eigenvalues far from the origin converge faster.

- **Slow convergence**
Rapid convergence achieved with less than 10% of the Jacobian information.

Heuristic 1: only include the off-diagonal Jacobian terms corresponding to the fastest reaction rates.

Rapid convergence achieved with less than 10% of the Jacobian information.
Direct handling of certain unimolecular reactions gives much faster convergence.

Unimolecular reactions can produce eigenvalues near zero (after scaling) that slow convergence.

Direct inversion of the five 2x2 sub-matrices with the smallest eigenvalues reduces the number of iterations by a factor of 3 with less than 1% of the Jacobian information.

For this mechanism, direct handling of up to 20 unimolecular species gives significant speedup.

Ignition Delay with isooctane for different equivalence ratios:

- $\phi = 1$
- $\phi = 0.5$
- $\phi = 0.25$
Graphical Processing Units (GPUs) can bring supercomputing to the desktop workstation

- ½ Teraflop for $500
- 480 parallel processors
- Codes must be redesigned to take advantage of architecture
- Massively parallel computing on the desktop
- Fortran/C++ Compilers designed for GPUs now available

**LLNL is a center for research on using GPU architectures for large-scale scientific simulations**

<table>
<thead>
<tr>
<th><strong>LLNL Edge Cluster</strong></th>
<th><strong>Off-the-shelf Desktop</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>GPUs:</strong></td>
<td><strong>GPUs:</strong></td>
</tr>
<tr>
<td>412</td>
<td>1</td>
</tr>
<tr>
<td><strong>Type:</strong></td>
<td><strong>Type:</strong></td>
</tr>
<tr>
<td>Tesla M2050</td>
<td>Tesla C2050</td>
</tr>
<tr>
<td><strong>Cores:</strong></td>
<td><strong>Cores:</strong></td>
</tr>
<tr>
<td>185,000 GPU</td>
<td>448 GPU</td>
</tr>
<tr>
<td>2,500 CPU</td>
<td>8 CPU</td>
</tr>
<tr>
<td><strong>Tflop/s:</strong></td>
<td><strong>Tflop/s:</strong></td>
</tr>
<tr>
<td>212 (64-bit)</td>
<td>0.5 (64-bit)</td>
</tr>
<tr>
<td><strong>Price:</strong></td>
<td><strong>Price:</strong></td>
</tr>
<tr>
<td>4 - 5 M$</td>
<td>5.7 K$</td>
</tr>
</tbody>
</table>
Getting the most out of the GPU involves designing suitable algorithms

- Thermodynamic property evaluation illustrates algorithmic design
- Specific heat
  - Two polynomials: Low and high temperature

\[
C_p = \begin{cases} 
  a_1 + a_2 T + a_3 T^2 + a_4 T^3 + a_5 T^4 , & T < 10000K \\
  b_1 + b_2 T + b_3 T^2 + b_4 T^3 + b_5 T^4 , & T > 10000K 
\end{cases}
\]
Thermodynamic property evaluation illustrates GPU algorithm design

\[ C_p = \begin{cases} 
  a_1 + a_2 T + a_3 T^2 + a_4 T^3 + a_5 T^4, & T < 1000 K \\
  b_1 + b_2 T + b_3 T^2 + b_4 T^3 + b_5 T^4, & T > 1000 K 
\end{cases} \]

- Scheme v2.5 (shared memory, both high and low T calculated)
- Schemes v2.0 – v2.3 (shared memory)
- Scheme v2.4 (switched dimensions)
- Schemes v1.0, v1.1 (no shared memory, brute force)
Enthalpy and entropy calculations have greater speedup with the GPU; GPUs perform best with more arithmetic operations per memory access.
We are doing ignition delay calculations with large mechanisms to determine best practices for GPUs.
Direct matrix inversion is the most effective solver strategy on the GPU, but performs poorly on the CPU.

GPUs can do many high repetitive calculations with little computational effort.

CPUs get bogged down by these kinds of calculations.
We have achieved up to 11x speedup with GPU for ignition delay with large mechanisms.
We have done extensive KIVA-4mpi submodel development and testing to enable large-scale parallel CFD for engine simulation.
KIVA-4mpi-Multizone is tested with a well-characterized 3D benchmark case, square-bowl HCCI

Grid 1: ~25,000 cells
Grid 2: ~70,000 cells
Grid 3: ~200,000 cells
Kiva-3V Grid: ~400,000 cells
Kiva-4mpi-MZ allows us to investigate more complex geometries with large detailed mechanisms.
We consider it vital to conduct simulations with close coupling to experimental data.

Sandia NVO PCCI

Sandia DI PCCI

Lund HCCI

NREL IQT (DI Diesel)
We completed a 5 year licensing agreement for the LLNL Multi-zone Model with Convergent Science Inc. (CSI)

- CONVERGE from CSI is high performance parallel CFD solver widely used in US industry
- LLNL has CONVERGE-MZ licenses for complex 3D problems on our large-scale parallel computers
- LLNL is working with CSI to implement and test the CONVERGE multi-zone model
We are benchmarking KIVA-4mpi-multi-zone and CONVERGE-multi-zone

Kiva4-mpi-multi-zone

CONVERGE-multi-zone
Collaboration: We have ongoing interactions Industry, National Labs, and Universities

- **Convergent Science Inc** - Multi-zone license agreement
- **Advanced Engine Combustion (AEC) working group** (Industry, National labs, Univ. of Wisc.) - semiannual presentations
- **Fuels for Advanced Combustion Engines (FACE)** - working group
- **Sandia National Laboratory** - HCCI and PCCI, gaseous injection
- **Oak Ridge National Laboratory** - SI-HCCI transition and $^{14}$C exhaust analysis for HCCI and Diesel engines
- **Los Alamos National Laboratory** - Kiva4 development
- **Other Universities** - UC Berkeley, University of Wisconsin, University of Michigan, Lund Institute, Chalmers University, Tianjin University
- **Ford** - gaseous direct injection
- **Delphi** - direct injection
Future Work: We will explore strategies for improving efficiency of CFD and chemistry simulations

- **Improved computational chemistry solvers**
  - Sparse solvers (CPU & GPU)
  - More efficient data structures
  - Hybrid solver solutions
  - Solver parallelization compatibility
  - Reaction sort with submatrix direct inversion.
  - New integration error control logic
  - Increase GPU shared memory reuse
  - GPU particle motion/collision algorithms

- **Improved parallel CFD with chemistry**
  - Multi-criteria multi-zone
  - Spray parcel models
  - Spray initialization
Future work: extend applicability and computational efficiency of analysis tools

Enable 3-D fluid mechanics and detailed kinetics in today’s desktop PCs

Continue to validate and develop parallel CFD, multi-zone, and ANN for highest fidelity in fluids and chemistry
Summary: we are enhancing our analysis capabilities and improving computational performance.