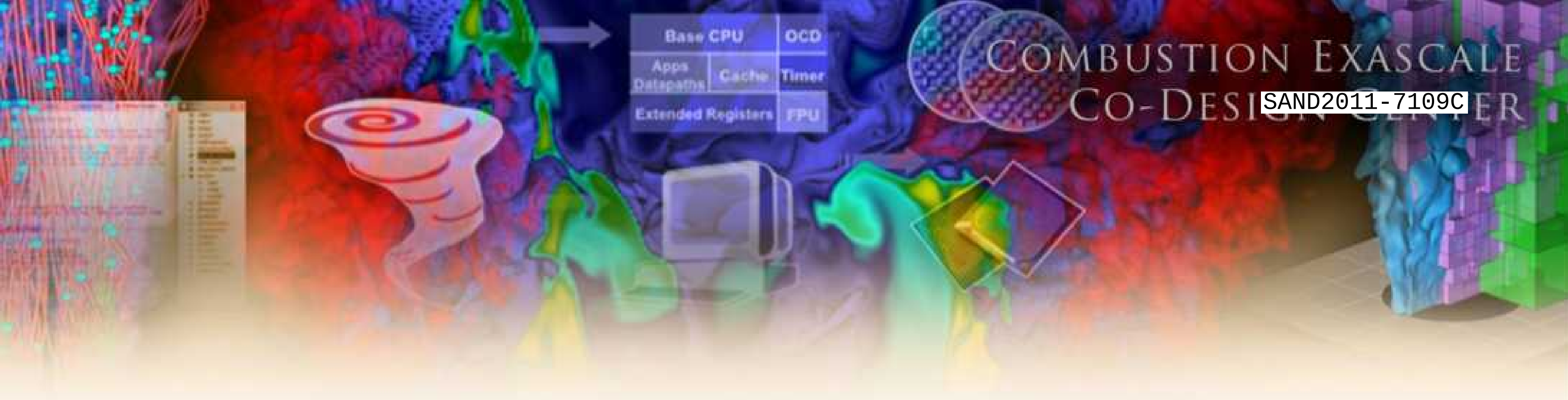


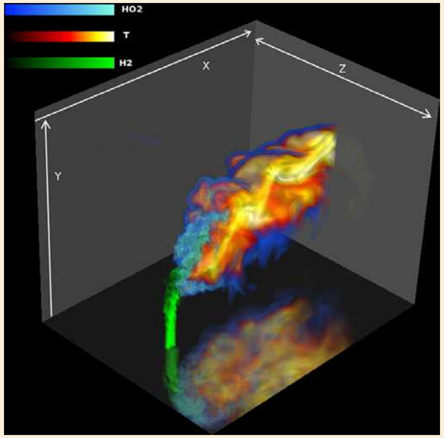
SAND2011-7109C

CECDC Overview and Combustion Science Driver

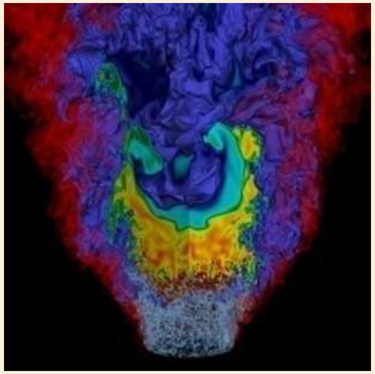
Jacqueline H. Chen
CECDC Kick-off Meeting
Sandia National Laboratories
Livermore, CA 94550
September 20-22, 2011

Why Combustion

- 83% of U.S. energy comes from combustion of fossil fuels
- National goals
  - Reduce greenhouse gas emissions by 80% by 2050
  - Reduce petroleum usage by 25% by 2020
- Meeting these goal requires a new generation of high-efficiency, low emission combustion systems
  - New designs for IC engines, turbines and burners
  - New fuels
- Examples
  - Engine designs to burn biodiesel for transportation
  - Fuel flexible turbines for power generation
- Why exascale
  - Current design methodologies are largely phenomenological
  - Lack the science base needed to develop new devices / fuels
  - Exascale computing will enable direct simulations of complex fuels at relevant turbulence/pressure conditions with quantified uncertainty
  - Key element needed to develop and calibrate models needed for engineering design
- Use cases based on increasing pressure and fuel complexity



Simulation of a nitrogen-diluted hydrogen jet in crossflow



Simulation of NOx emissions from a low swirl injector fuels by H2

Combustion Co-Design Project

- **Goals**
  - Redesign all aspects of the combustion simulation process from core simulation and analysis methodology to programming models and languages to enable high-fidelity combustion simulations on exascale architectures
  - Identify and influence key hardware features that have a critical impact on combustion simulation performance
- **Objectives**
  - Redesign PDE solution algorithms in terms of a hybrid parallelization model to expose fine-grained parallelism and increase concurrency
  - Analyze alternative PDE discretization approaches in terms of data movement and power usage instead of flops
  - Develop programming models that enable developers to specify data locality
  - Develop / deploy communication avoiding linear solver concepts to reduce synchronization costs
  - Develop *in situ* analysis methodology (hardware and software) to perform initial stages of the analysis "on the fly" to reduce I/O requirements
  - Develop UQ methodologies (hardware and software) appropriate to turbulent combustion to quantify predictive capabilities of simulations
  - Demonstrate combustion computations on unreliable processors and memory
  - Demonstrate the productivity of DSL in the constrained domain of combustion algorithms

Combustion Co-Design Project

- **What questions must be answered to build the code capability to perform this type of simulation**
  - **What type of PDE discretization methodology will be effective on many-core exascale architectures**
  - **What hardware features are needed for these algorithms to run efficiently**
  - **How do we best integrate UQ into the simulations**
  - **Are there specialized hardware features that facilitate UQ algorithms**
  - **What algorithmic approaches are needed to integrate *in situ* analysis and viz into the simulation**
  - **What types of hardware / system features are needed to support *in situ* analytics**
  - **What type of programming models and languages do we need to effectively express these types of algorithms**
  - **How do we effectively compute in a high failure rate environment**

Project Organization

- Project organized around two co-design themes
  - PDE/UQ
  - *In situ* analysis and visualization/UQ
  - Overall design coordination task
  - Interdisciplinary teams for each co-design theme

- Cross-cutting CS activities
  - Abstractions, algorithms, programming models, languages and optimization
  - Architecture simulation
  - In situ scientific data management and analytics
  - Hardware resiliency and software fault tolerance
  - Vendor interface

CECDC Organizational Structure

Internal Steering Com.

Chen
Bell
Moser
McCormick
Klasky
Pascucci
Quinlan
Janssen
Shalf

Management Team

Director: Chen
Deputy Director: Bell
Project Manager: Moreno

External Advisory Board

Kathy Yelick (Chair)
Marsha Berger
Marc Snir
Kwan-Liu Ma
TBD

Combustion Science Driver and Requirements
Lead: Chen

PDE/UQ Solver Co-Design Team (Co-Leads: Bell/Moser)
In situ Scientific Data Management and Analytics Co-Design Team (Co-Leads: Klasky/Pascucci)

Math
Lead: Bell

- PDE Solver Methods
- Linear Solvers
- Data Structure/AMR Metadata
- Single point Thermo-Chemistry Kernel

UQ
Lead: Moser

- In Situ UQ
- Smart Sampling
- Polynomial Chaos
- Transient Adjoints
- Chaotic Systems

SDMA Co-Leads: Klasky/Pascucci

- Data Staging Infrastructure
- Workflow Scheduling And Resiliency
- In situ Topological Analysis/Viz.
- In situ Statistical Analysis and Reduction
- Data Layout

Programming Models/Autotuning
Lead: McCormick

- Combustion Abstractions
- Domain Specific Programming Language
- Run-time Systems
- Autotuning
- Fault Tolerance

Simulators
Lead: Shalf

- Simulation of Skeletal, Node and Compact Apps
- Architectural Simulation
- Compiler Support for Hardware/Software Simulation
- SST Macro
- SST Micro

Teams

Math*

John Bell -- Lead
Vincent Beckner
Marc Day
Rob Falgout
Ray Grout
Mike Lijewski
Ulrike Yang

UQ*

Bob Moser -- Lead
Omar Ghattas

Simulators*

John Shalf -- Lead
David Donofrio
Dan Quinlan
Peter Pirkelbauer
Arun Rodrigues
Curtis Janssen

Programming models*

Pat McCormick – Lead
Vince Schuster
Pat Hanrahan
Alex Aiken
Dan Quinlin
Peter Pirkelbauer
Sam William
Erich Strohmaier
Paul Hargrove

***Requirements/validation oversight by combustion scientists**

Jackie Chen -Lead
John Bell
Marc Day
Ray Grout
Hemanth Kolla

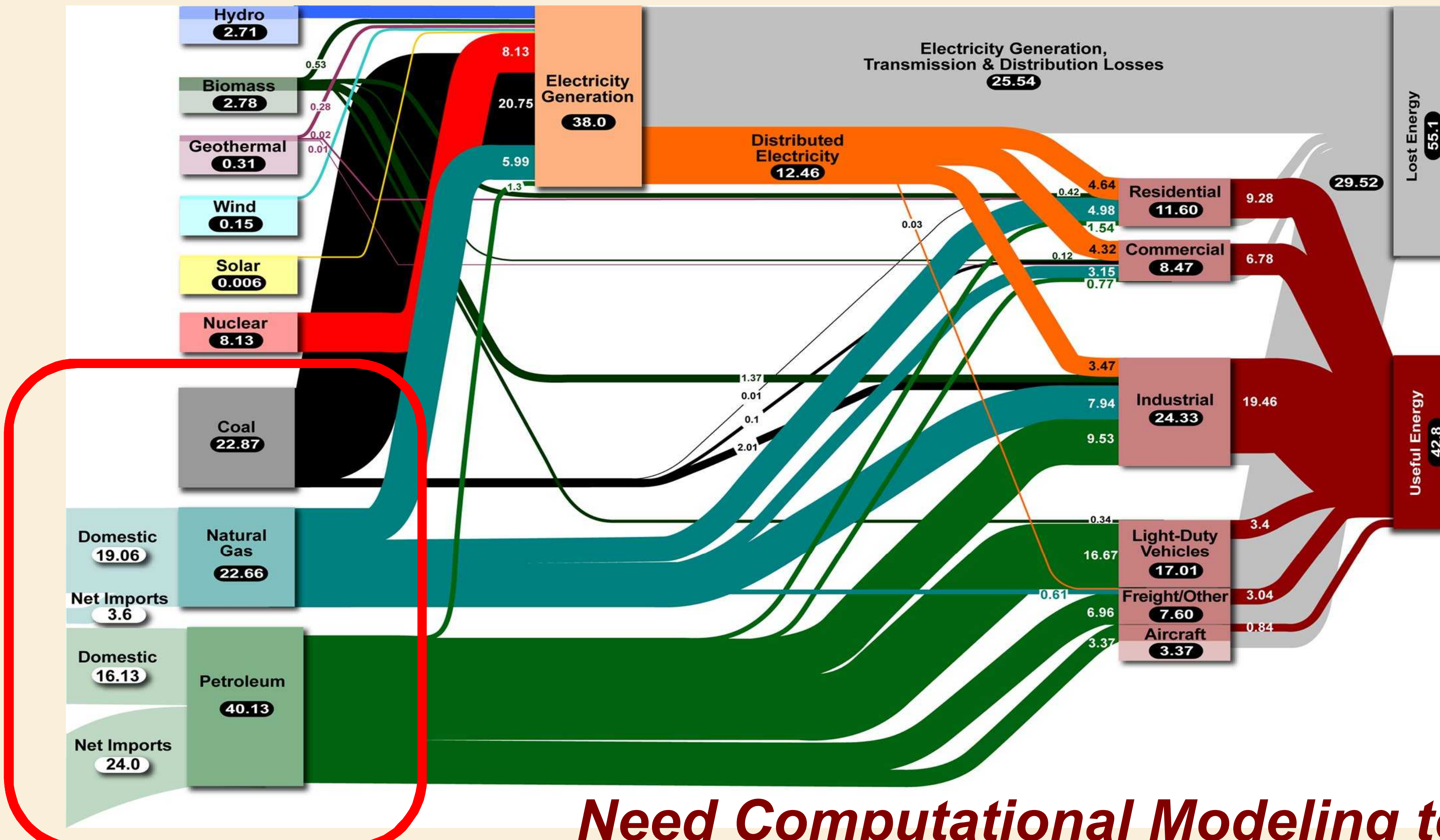
Data Management and analytics team*

Jackie Chen
Scott Klasky -- Lead
Hasan Abbasi
Valerio Pascucci--Lead
Manish Parashar
Karsten Schwan
Janine Bennett
Hongfeng Yu
Hemanth Kolla
Attila Gyulassy
Peer-Timo Bremer

Combustion Driver

- Motivation
- DNS approach
- Why exascale?
- Community Data
- Petascale DNS
- Mountains of Data – Analytics/Viz
- Combustion Use Cases for co-design
  - HCCI combustion
  - Lifted Biofuel jet flame

Combustion accounts for 85% of the energy used in the United States



Need Computational Modeling to enable efficient combustion systems

Soucrce: LLNL

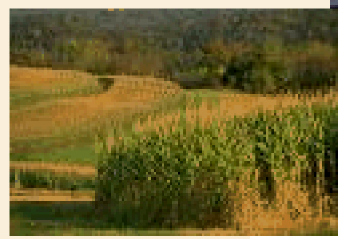
Motivation: Changing World of Fuels and Engines

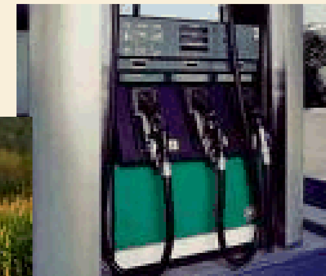
- Fuel streams are rapidly evolving
  - Heavy hydrocarbons
    - Oil sands
    - Oil shale
    - Coal
  - New renewable fuel sources
    - Ethanol
    - Biodiesel

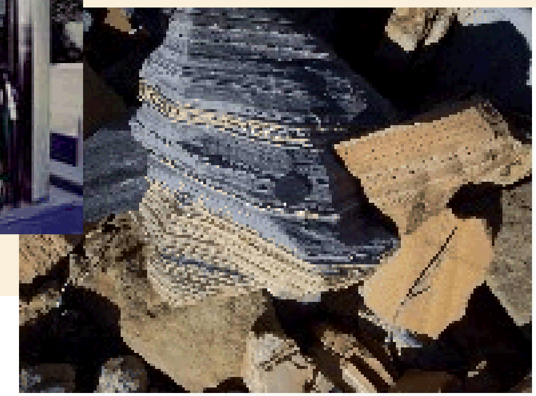
- New engine technologies
    - Direct Injection (DI)
    - Homogeneous Charge Compression Ignition (HCCI)
    - Low-temperature combustion

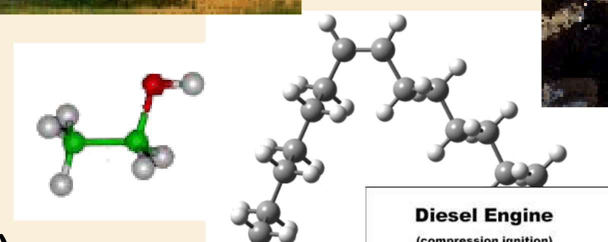
- Mixed modes of combustion (dilute, high-pressure, low-temp.)

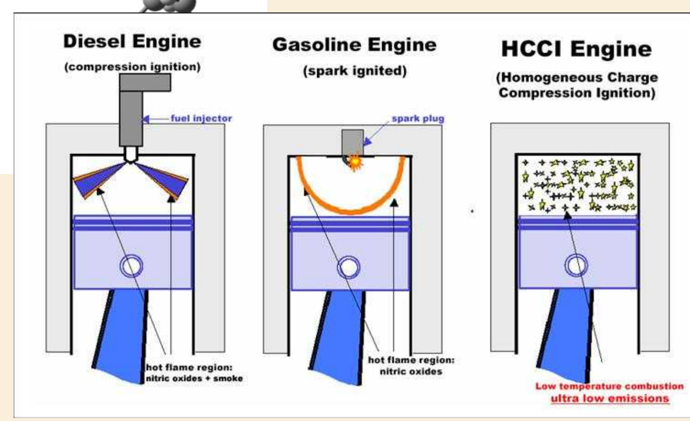
- Sound scientific understanding is necessary to develop predictive, validated multi-scale models!









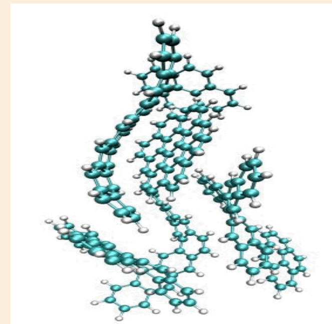


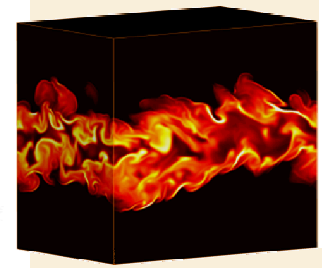
Multi-scale Modeling of IC engine processes

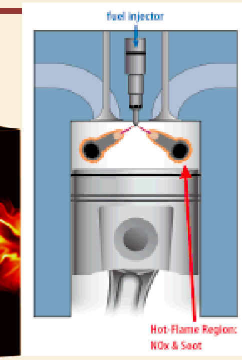
•Multi-scale modeling describes IC engine processes, from quantum scales up to device-level, continuum scales

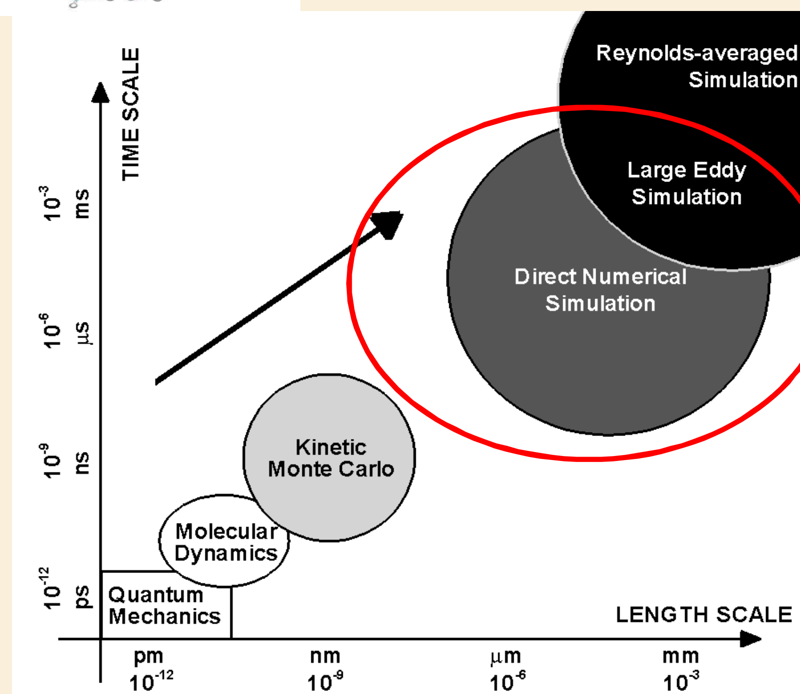
• Needs:

- •Develop a general theoretical framework for transfer of information from one scale to the next
- •Use HPC to bridge the current gap between coarse-grained atomistic approaches and fine-grained continuum approaches





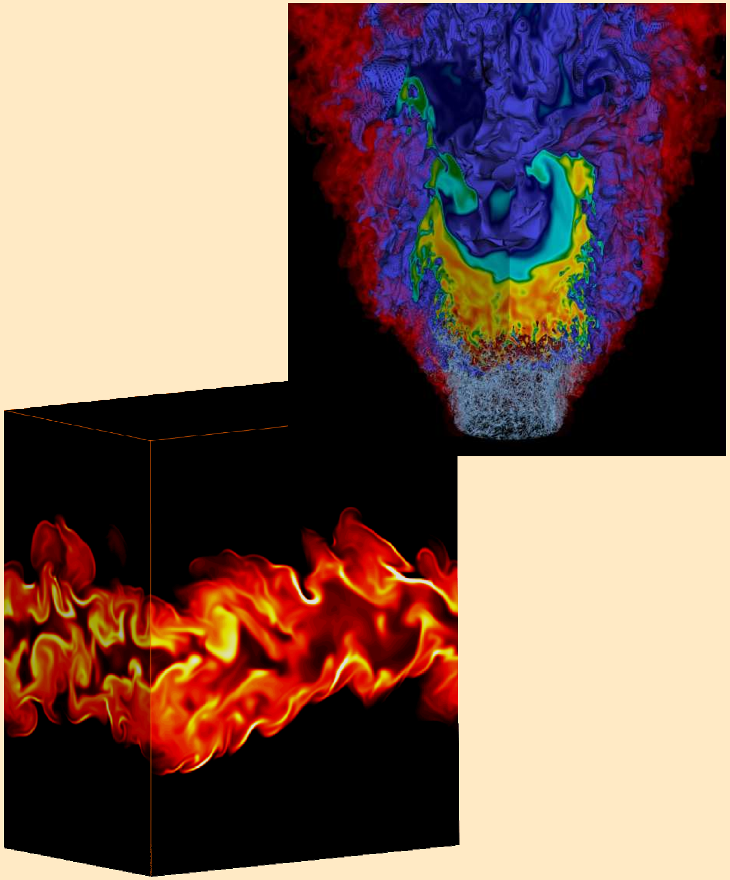




Direct Numerical Simulation

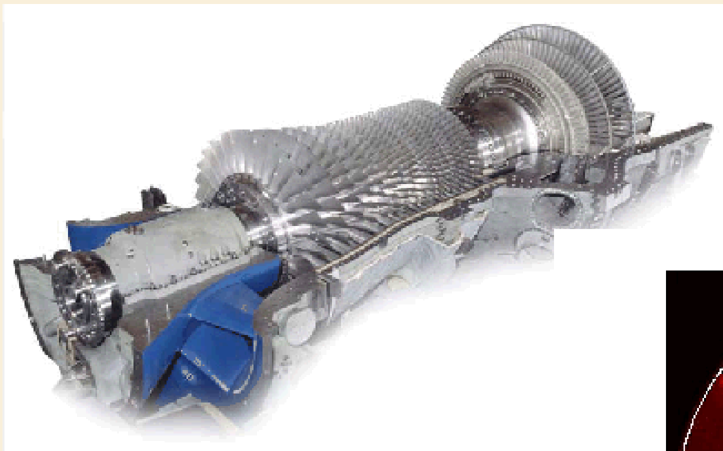
- Used to perform first-principles-based DNS of reacting flows
- Solves compressible or low-Mach reacting Navier-Stokes equations
- High-fidelity numerical methods (high-order finite difference or AMR)
- Detailed reaction kinetics and molecular transport models
- Multi-physics (sprays, radiation and soot)
- Runs on all major platforms, scales well on petascale machines

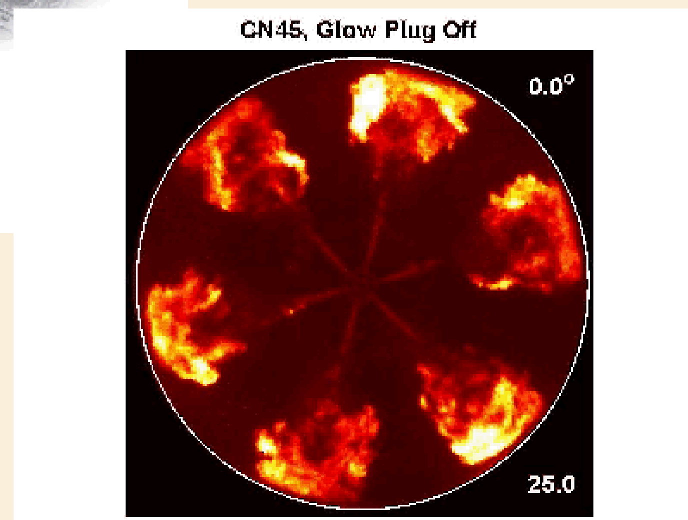
DNS ⟷ *Physical models* ⟷ *Engineering CFD codes (RANS, LES)*



DNS provides unique fundamental insight into gas-phase chemistry-turbulence interactions

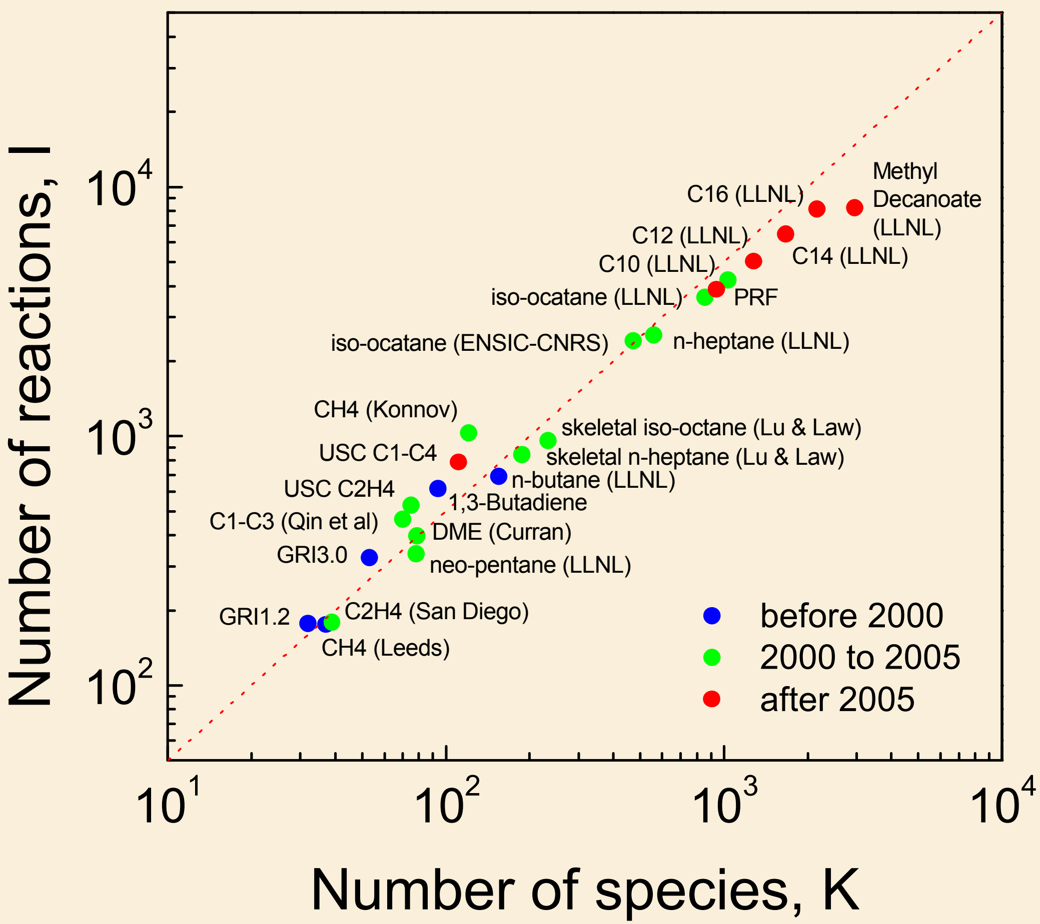
Why exascale? Parameter space of practical combustors





Increase Reynolds number, pressure, and chemical complexity
$N_{xyz} \alpha\ Re_t^{\ 9/4}$ $Re_t \sim O(1000)$ in spark ignited engine
need O(billion grids)

Chemical Complexity of Hydrocarbons



For explicit solver chemistry cost scales linearly with number of species (Species:9 hydrogen, 22 ethylene, 30 di-methyl ether, 73 bio-fuel)

Disparity in Flame &Turbulence Scales with Pressure

$Ka = \delta_L^2/\eta^2 \propto P^{0.7}$ $Re_t \propto P$

$\eta \propto P^{-3/4}$ $S_L \propto P^{-1/2}$

$\delta_L \propto D/S_L \propto P^{-1} P^{-1/2} \propto P^{-1/2}$

- Eddies get smaller as pressure increases
- Ratio of turbulent integral scale to flame thickness increases
- Ratio of turbulence intensity to laminar flame speed increases
- Increases propensity for burning in thin reaction zones regime

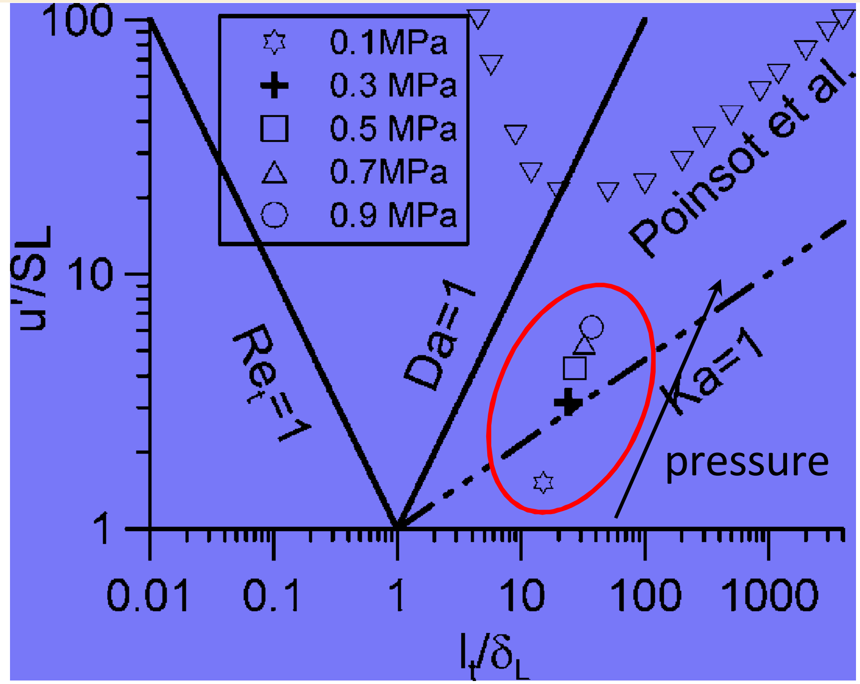
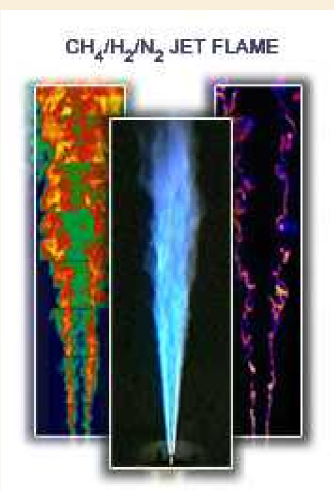


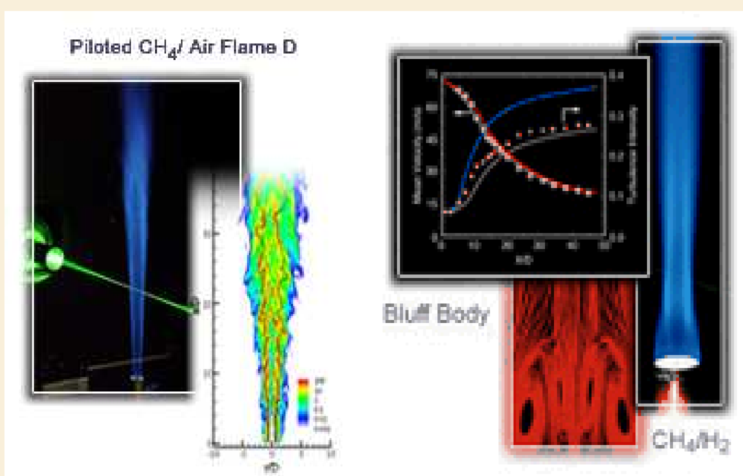
Fig. 3. Turbulent combustion regimes of the investigated flames.

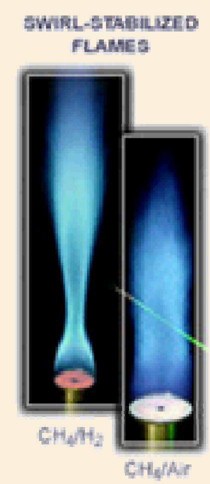
Increased disparity in turbulence and flame scales at high pressure requires adaptive mesh refinement

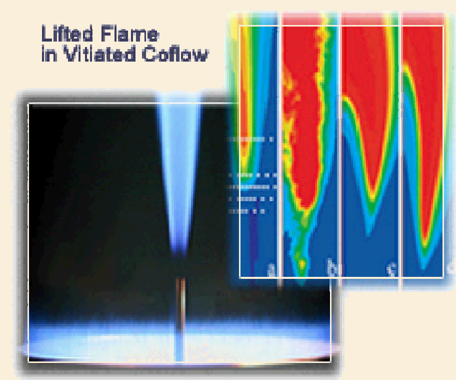
Community Benchmark Data Sets

- Turbulent Nonpremixed Flame Workshop Series and International Workshop on Premixed Turbulent Flames
- International collaboration of experimental and computational researchers
- Addition of high-fidelity numerical benchmarks for model development and validation
- Need for data mining, data sharing and visualization tools for petabytes of data

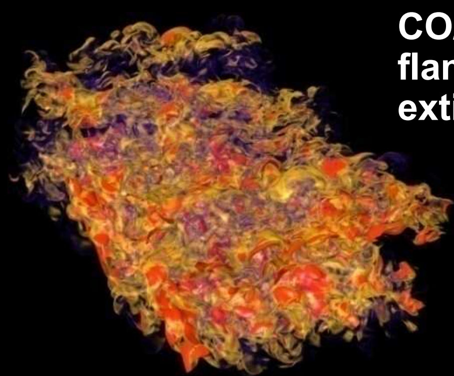






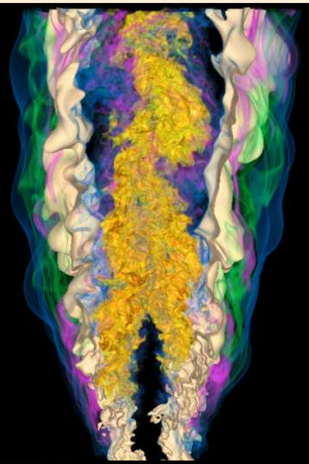


DNS Benchmarks for Model Development

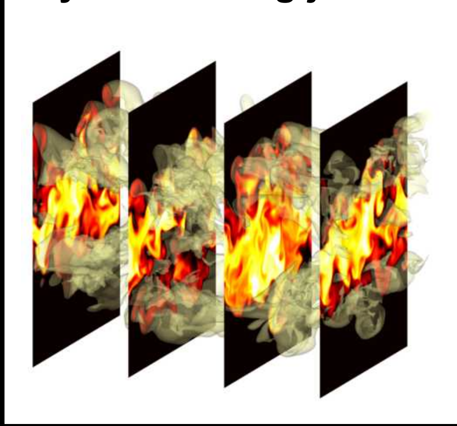


CO/H_2 and C2H4 jet flames extinction/reignition

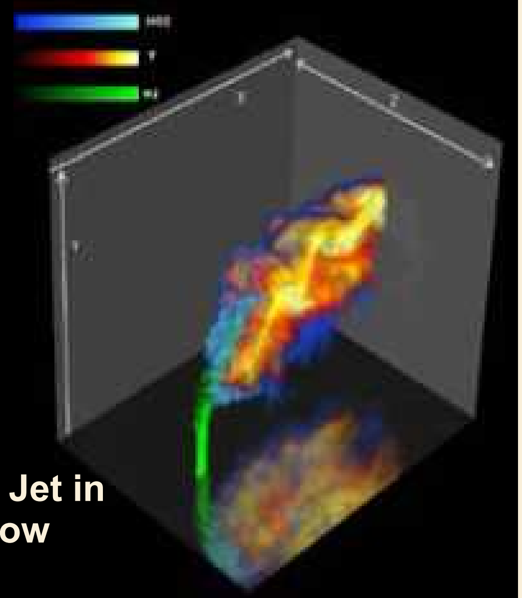
Lifted H2 and C2H4 jet flames in hot coflow



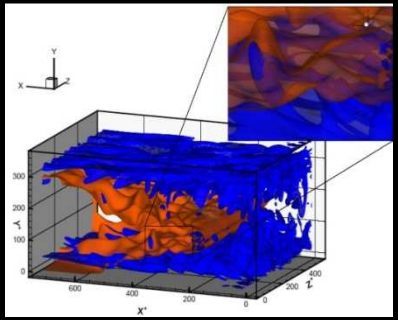
Ethylene sooting jet flame



Lean premixed and stratified Bunsen flames



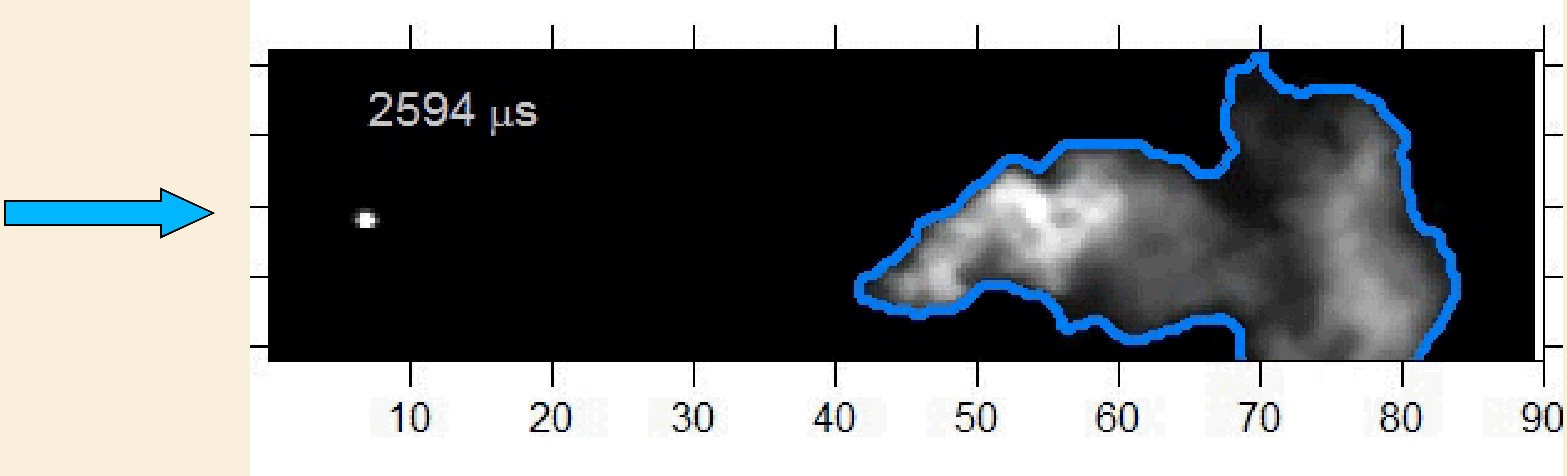
H_2 Fuel Jet in Crossflow



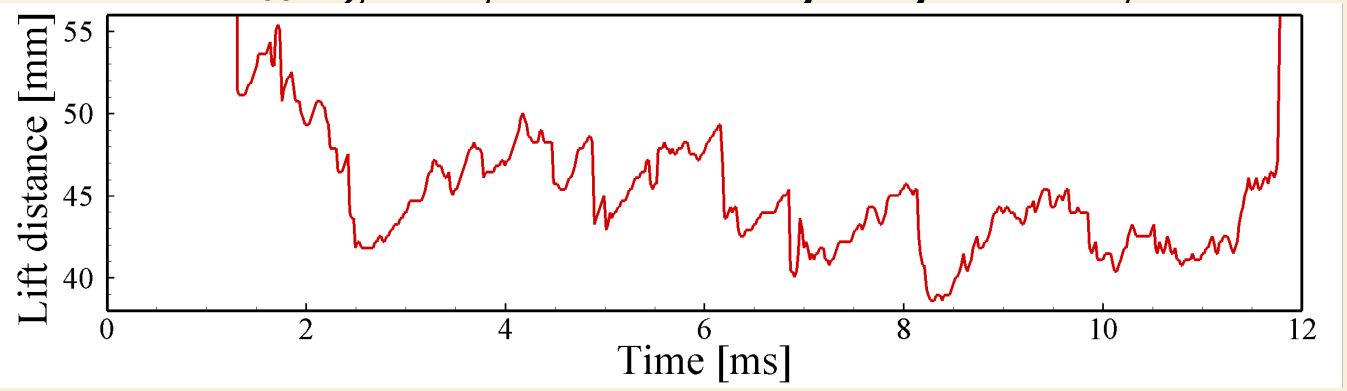
H2 Flame-wall interaction

Motivation: understanding stabilization of lifted flames in heated coflow

What is the role of ignition in lifted flame stabilization?



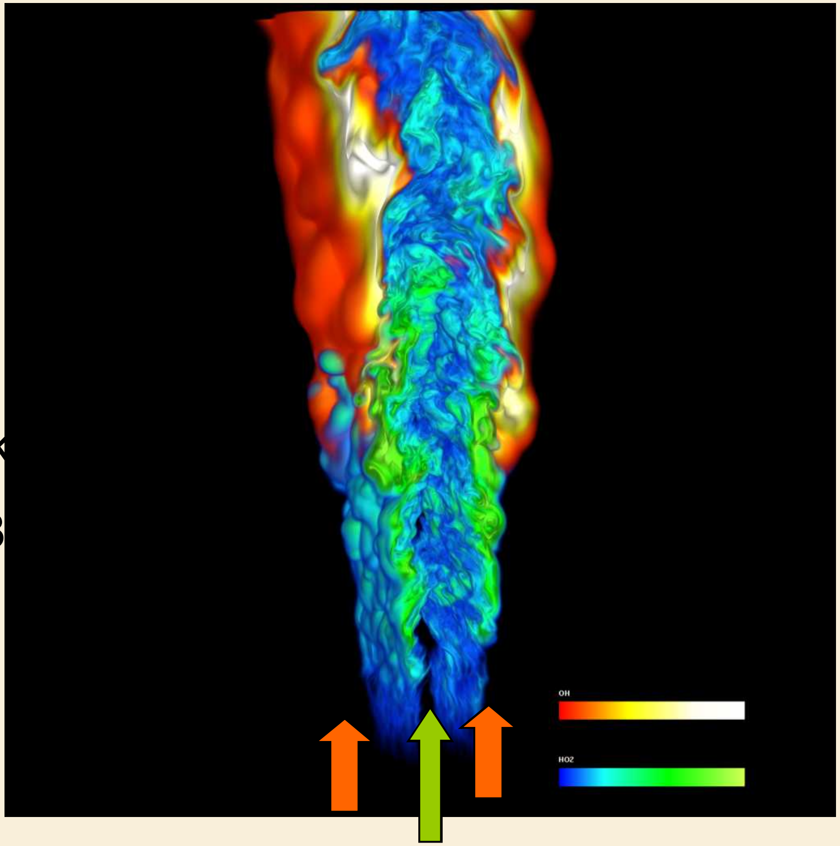
Chemiluminescence from diesel lift-off stabilization for #2 diesel, ambient 21% O_2, 850K, 35 bar courtesy of Lyle Pickett, SNL



Petascale DNS of a Lifted Autoignitive Ethylene-Air Jet Flame

- 3D slot burner configuration:
  - $L_x \times L_y \times L_z = 30 \times 40 \times 6\ \text{mm}^3$ with
  - 1.28 billion grid points
  - High fuel jet velocity (204m/s); coflow velocity (20m/s)
  - Nozzle size for fuel jet, H = 2.0mm
  - Re_{jet} = 10,000; τ_j = 0.15ms; 6 flow through times
  - Cold fuel jet (18% C_2H_4 + 82% N_2) at 550K
  - Detailed C_2H_4/air chemistry, 22 species 18 global reactions, 201 steps
  - Hot coflow air at 1,550K
- Performed on CrayXT4 at ORNL on 30,000 cores and 14 million cpu-hrs
  - 240 TB field data, 50TB particle data

Ethylene-air lifted jet flame at Re=10000



Mountains of Data - Reduction

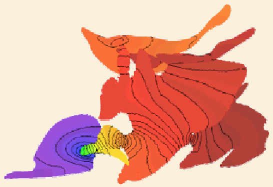
- At the exascale I/O speeds will lag behind compute speeds – fewer checkpoints and analysis to persistent storage

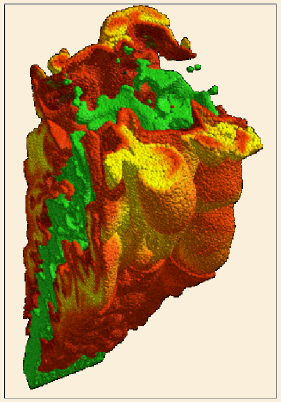
- In situ analysis/viz/UQ – minimize data movement
  - Topological data analysis for concise multi-scale representation of combustion features
  - Feature tracking and visualization capture highly intermittent transient phenomena (e.g. ignition, extinction)
  - Uncertainty propagation in unsteady chaotic reactive flows to provide sensitivities of heat release and emissions to key chemical parameters

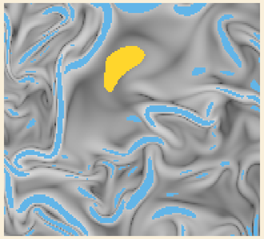
- Current state-of-the-art is analytics as ‘postprocessing’.

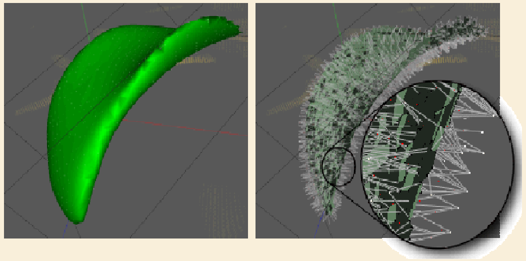
In Situ Analytics for Turbulent Combustion

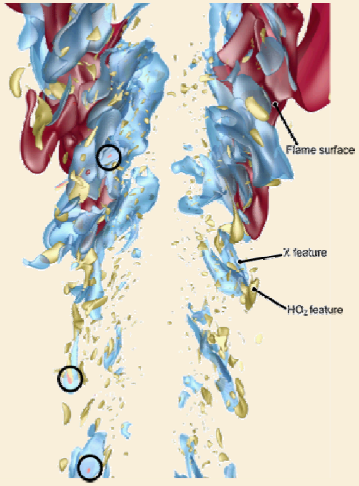
- *In situ* volume and particle rendering
- Lagrangian particle querying and analysis
- Topological segmentation
  - Contour trees
  - Morse-Smale complex
  - Jacobi sets
- Scalar field comparison
- Shape analysis
- Feature tracking
- Statistical analysis
- Statistical dimensionality reduction

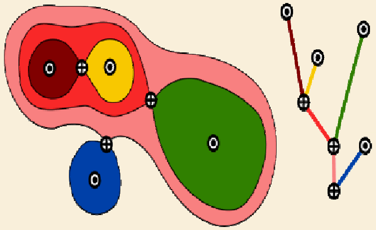






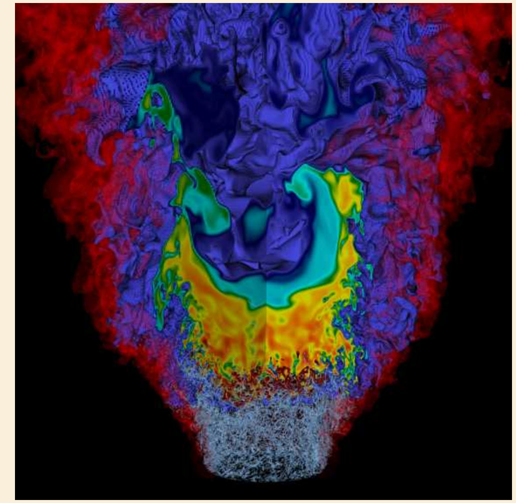


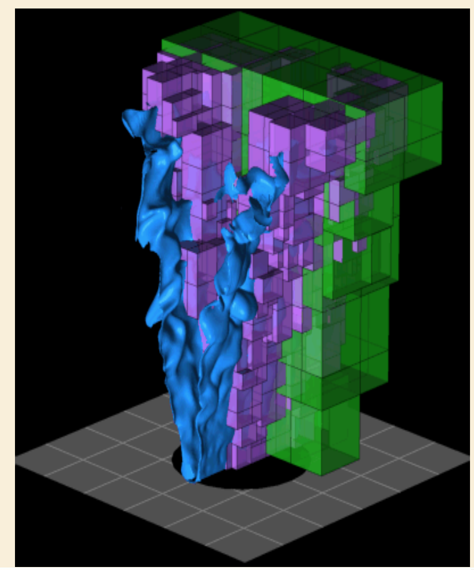




High-dimensional Multivariate Analysis for Block Structured AMR Data

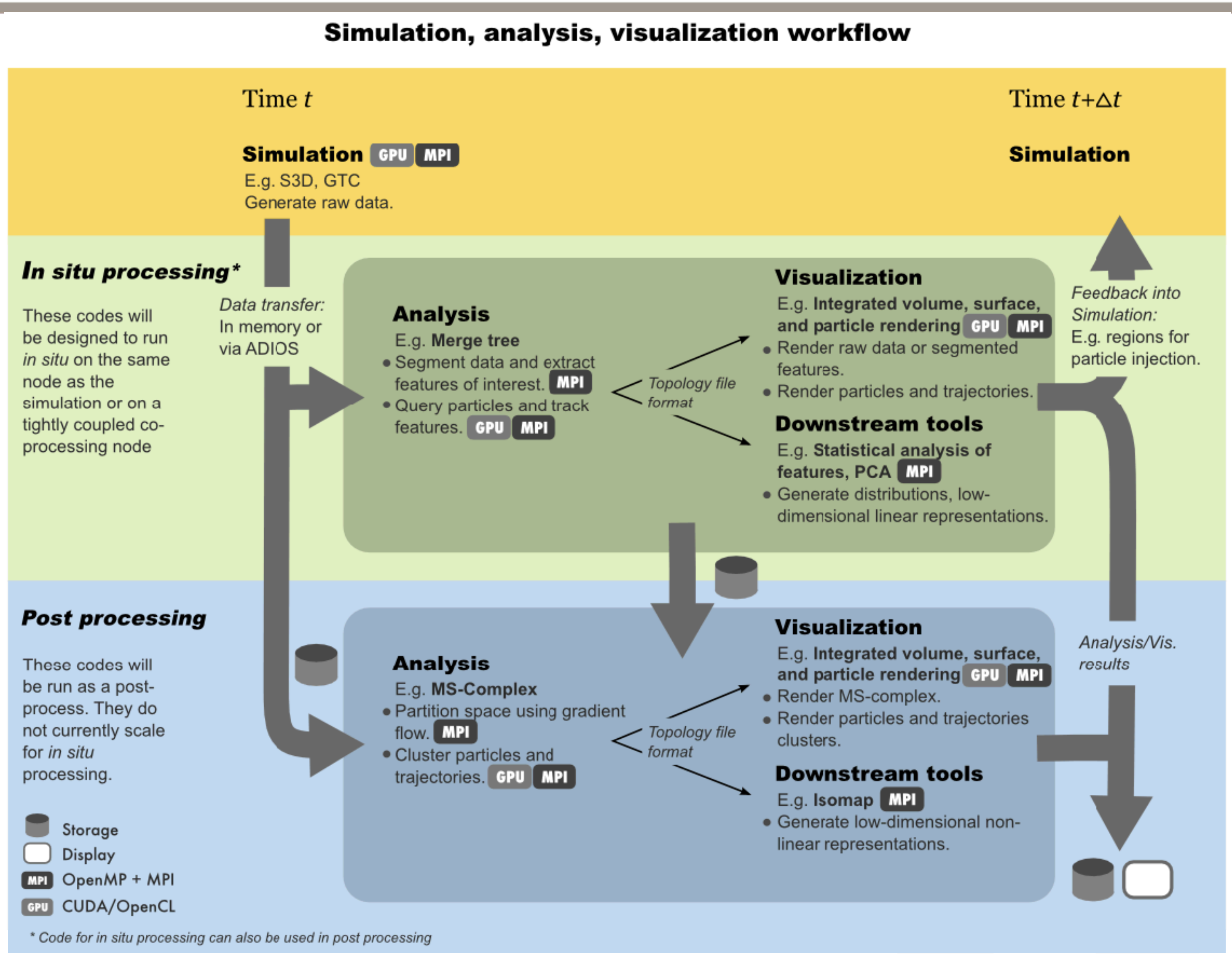
- Diagnostics of flame-fluid system
  - conditional data sampling, statistics processing
  - Construction of "flame-centric" local coordinates
  - Quantified chemical pathway analysis
  - Lagrangian sampling (passive particle traces)
  - Topological analysis (flame/emissions structures)
  - Construction of these in an AMR hierarchy
- Challenges at the exascale
  - Infeasible to dump sufficient data to reconstitute discretization-specific information
    - Temporal and spatial derivatives
    - Intermediate values (chemical rates, transport)
  - Reduced I/O via ***in-situ*** diagnostics
  - Hardware will influence implementation
  - Auxiliary threads w/data access, on-the-fly
    - Morse-Smale complex or Reeb graph dynamics
    - Local flamelet analysis, chemical flux balances
    - Accumulation of joint probability distributions





AMR grids (each with 32^3 - 64^3 cells) dynamically resolve "interesting" solution features

Simulation, analysis and visualization workflow : towards in situ at exascale

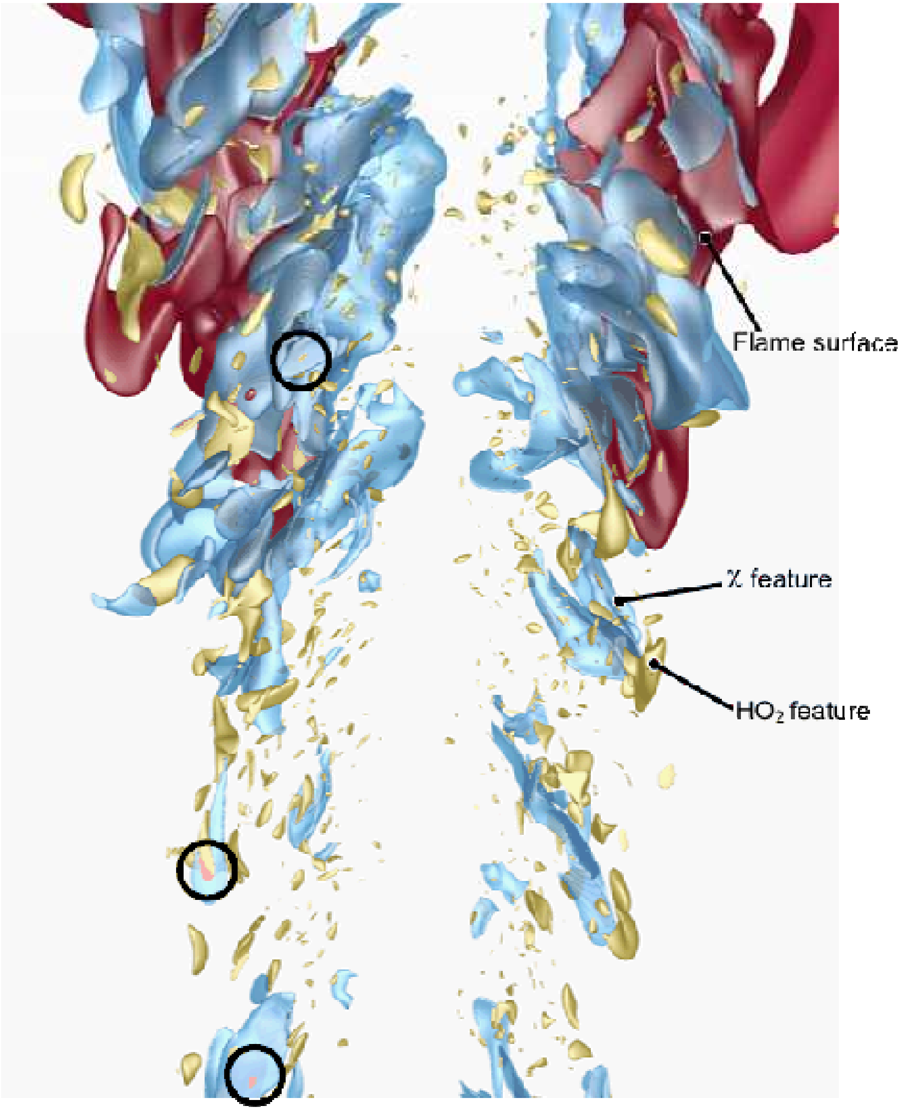


3D Topological Segmentation

Mascarenhas, Grout, Pascucci, Peer-Timo Bremer and Chen 2009

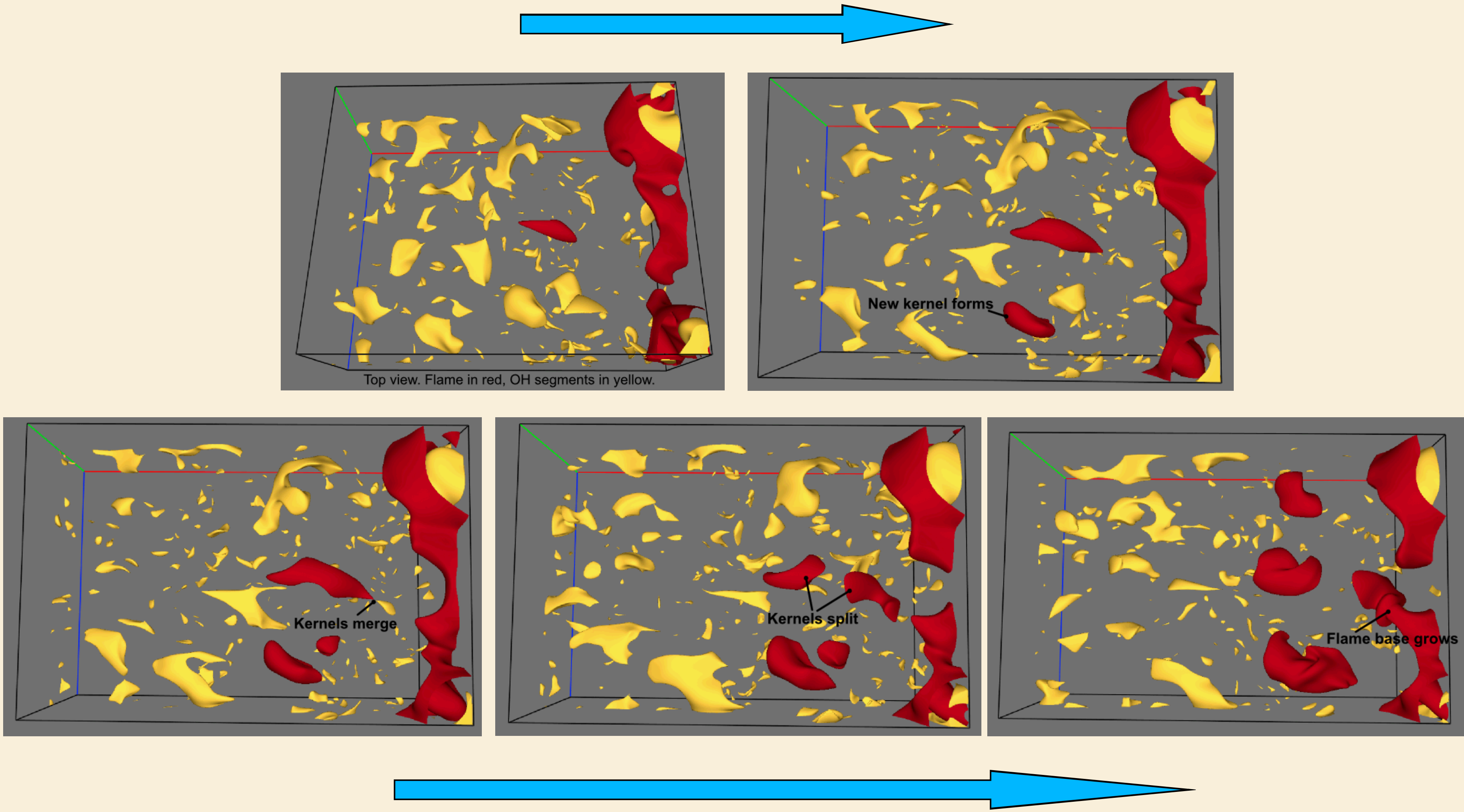
- A topological segmentation of the χ field consistent with its underlying Morse complex is possible using a combinatorial method (Bremer et al.)
- Briefly:
  - Grow region starting from local maxima.
  - Remove noise by cancelling maximum-saddle pairs with small difference in χ value.
  - Grow an isosurface around each remaining local maxima until the isovalue reaches a percentage (e.g. 50%) of the local maxima.

Relationship Between Mixing Rate and Ignition in a Lifted Turbulent Jet Flame



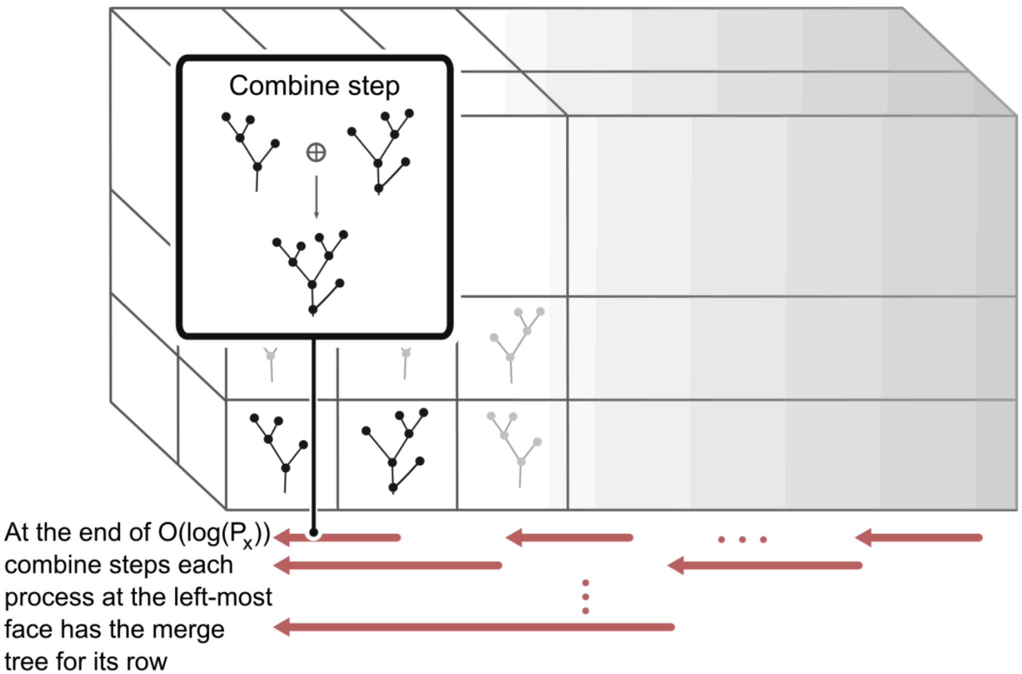
- Maxima of HO_2 species can indicate auto-ignition kernels
- These kernels are expected to live outside high-χ regions.

Tracking Kernel Evolution



Distributed Parallel Computation of Merge Tree

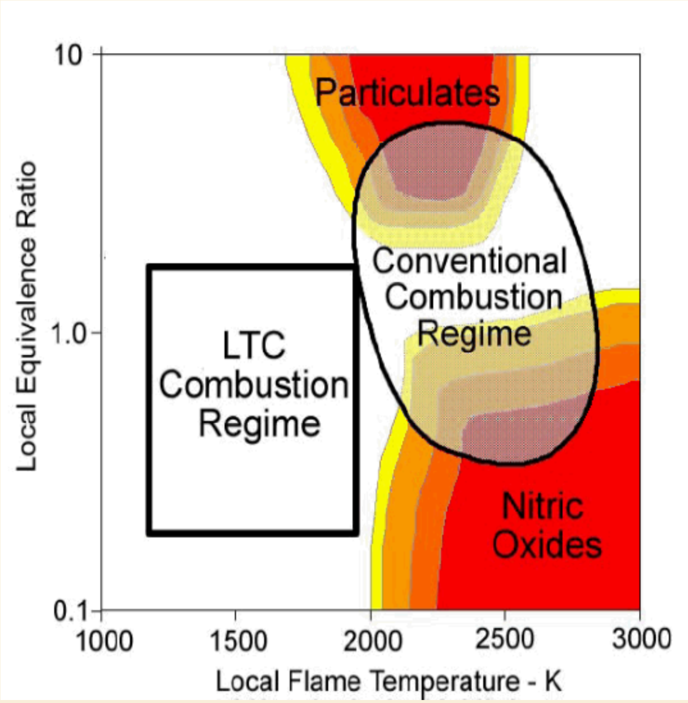
- We are working on a distributed parallel algorithm to compute the merge tree of massive combustion data.
  - Tests on DNS combustion data achieve run times < 10 seconds.

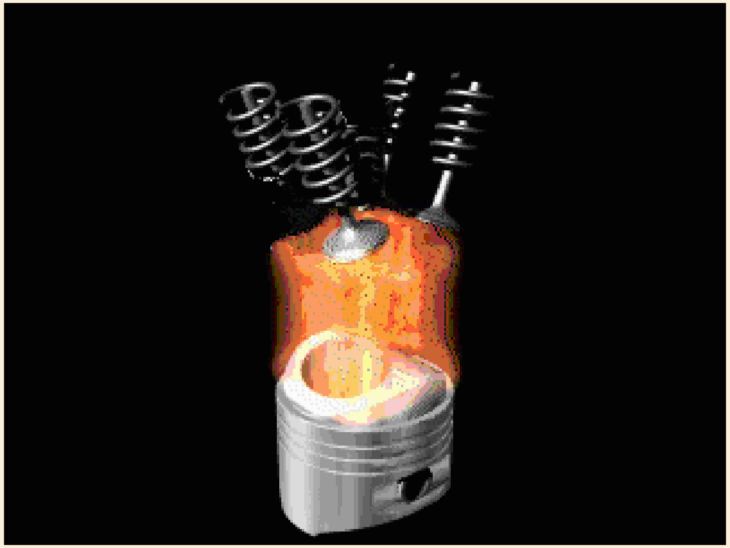


Alg. Phase	2.5K	10K	30K	129.6K
SINE				
Loc. MT.	21.858	4.739	1.408	0.363
Bound. crit. pts.	0.944	0.370	0.167	0.070
Combine	0.127	0.323	0.612	1.787
Total	22.929	5.432	2.187	2.220
CHI				
Alg. Phase	2.5K	10K	30K	129.6K
Loc. MT.	0.356	0.221	0.106	0.053
Bound. crit. pts.	0.861	0.340	0.154	0.071
Combine	4.970	4.880	4.966	4.658
Total	**6.187**	**5.441**	**5.226**	**4.782**

Homogeneous Charge Compression Ignition (HCCI) Engines

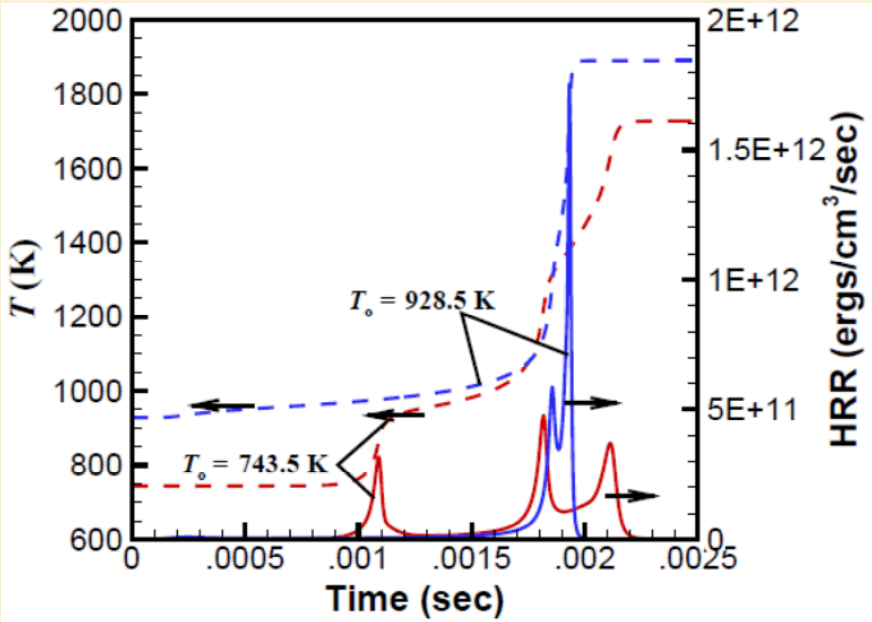
- Operating fuel-lean and at low temperatures reduces soot and NO_x and has potential for high diesel efficiency





Combustion initiation via autoignition: challenge to control ignition timing and pressure rise via mixture preparation

Turbulent Autoignition DNS Use Case



- 1.7 billion grids
- Fuel: di-methyl ether, 30 species
- 2.5-3 ms total run time
- 2-5 ns timesteps
- 3 TB/timestep (6-stage RK)
- Ignition events occupy < 10% vol.
- Track for ~20% of total time
- ~100,000 to 200,000 time steps
- UQ: Sensitivity of heat release to O(10) reaction rate input parameters

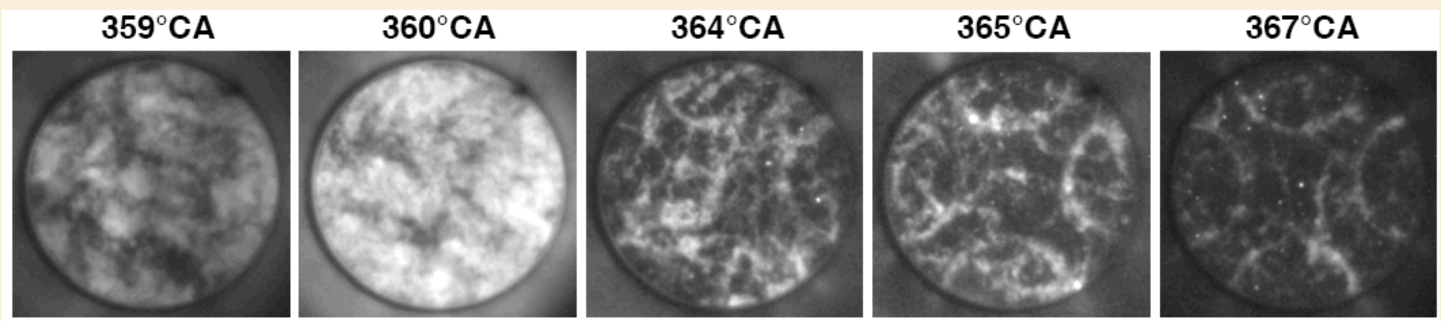


Figure S1. Chemiluminescence images of stratified sequential autoignition in an HCCI engine, courtesy John Dec, Sandia National Laboratories.

Combustion HCCI Use Case: SDMA/UQ Co-Design Issues

- SDMA middleware - explore staging requirements and functionality as they interact with requirements from combustion scientists and designers of the exascale hardware (e.g. limitations of cross-section communication bandwidth or resiliency).

- Sharing (memory) resources – work with programming models to develop lightweight integration with abstraction layers used by DNS solver to interface with the data (different data formats) for topology, viz and UQ. Investigate data partitioning and distribution schemes to minimize data replication taking into account simulation and analysis constraints.

- Extreme parallelism – Resolution of global features with topology methods will be done in an asynchronous staging area. Explore co-design tradeoffs with moving elements of topological computation to staging area.

- Data buffering requirements – large amounts of data required for backwards Integration for UQ

Turbulent Autoignitive Biofuel Lifted Jet Flame – an Exascale Use Case

- Simulation parameters: 4725*3200*540 ~ 8 billion grid points, 78 variables (73 species), $4.2*10^5$ timesteps, 850 million cpu-hrs.
- Sub-domain for UQ: 1350*1000*540 ~ 730 million grid points.
- Data per time step to be moved for UQ ~ 730*106*6*78*8 ~ 2.73 TB
- Typical duration of event of interest ~ 0.15 ms ~ 30,000 time steps.
- Total data to be moved to staging area ~ 82 PB.

Acknowledgments

- Combustion:
  - Chun Sang Yoo (UNIST)
  - Hemanth Kolla (SNL)
  - Gaurav Bansal (Intel)
- Computer Science:
  - Ajith Mascarenhas (Google)
  - Hongfeng Yu (SNL)
  - Janine Bennett (SNL)
  - Valerio Pascucci (Utah)

Questions

jhchen@sandia.gov