AMR Algorithms for Multiphysics Applications: Current Practice and Future Directions

John Bell

Lawrence Berkeley National Laboratory

ICIS – Multiphysics Simulations: Challenges and Opportunities
Park City, UT
August 4, 2011

PDE-based multiphysics applications

Many PDE-based applications require high-fidelity simulation of multi-scale / multi-physics phenomena

- Combustion
- Subsurface flow
- Fusion
- Astrophysics
- Climate
- Fission

Characteristic of these problem areas is that they couple a number of different physical processes across a range of length and time scales

How can we exploit the structure of these problems in developing simulation methodology

- What are the characteristics of these type of problems
- What might the algorithms look like
- How can we implement them
- How can effectively utilize modern parallel architectures

CCSE Berkeley Lab

Lean Premixed Turbulent Combustion – An Example

Rod-stabilized V-flame

4-jet Low-swirl burner (LSB)

Slot burner

- Potential for efficient, low-emission power systems
- Design issues because of flame instabilities
- Want to simulate complex fuels at realistic pressure and turbulence conditions

CCSE Berkeley Lab

Scales

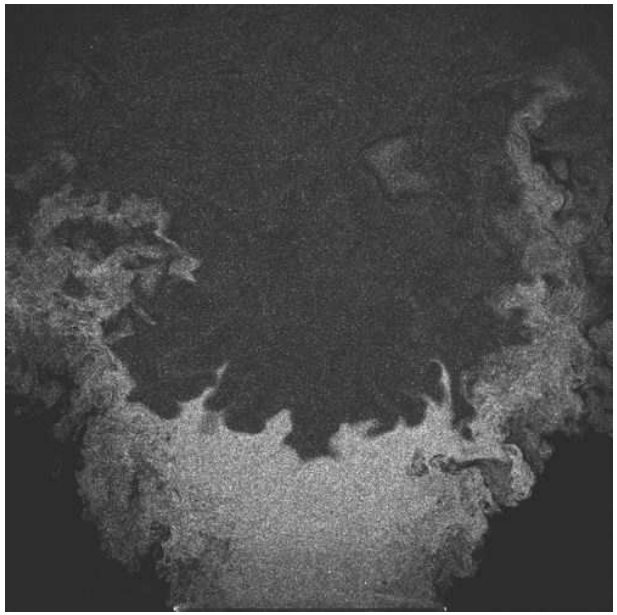
Spatial Scales

- Domain: $\approx$ 10 cm
- Flame thickness: $\delta_T \approx 1$ mm
- Integral scale: $\ell_t \approx 2 - 6$ mm

Temporal Scales

- Flame speed $O(10^2)$ cm/s
- Mean Flow: $O(10^3)$ cm/s
- Acoustic Speed: $O(10^5)$ cm/s

Fast chemical time scales



Mie Scattering Image

Strategies

- AMR to exploit varying spatial resolution requirements
- Issues in developing AMR depends on mathematical formulation of the problem
  - Approximation of the processes and how they coupling
  - Time step strategy

CCSE Berkeley Lab

Time step strategies

Scaling is paramount: Low communication, explicit discretizations, balanced work load – let the machine do the work

- Generic mathematical model
- Define spatial discretization – structured, unstructured, adaptive
- Identify time step based on stability requirements
- Integrate with explicit ODE algorithm
- Range of time scales determines performance

Coupling is paramount: Fully implicit, method of lines, iterative algorithms – preconditioners do the work

- Generic mathematical model
- Define spatial discretization – structured, unstructured, adaptive
- Identify time step based on accuracy requirements
- Integrate with implicit ODE algorithm (1 explicit scale)
- Efficiency of solver/preconditioner determines performance

Custom approach based on mathematical structure of the system

- Exploit the relationship between scales and processes in the problem

CCSE Berkeley Lab

Compressible Navier Stokes

Gas phase combustion – mixture model for diffusion

$$\begin{aligned}
\textbf{Mass} \quad & \rho_t + \nabla \cdot \rho U = 0 \\
\textbf{Momentum} \quad & (\rho U)_t + \nabla \cdot (\rho UU + p) = \rho \vec{g} + \nabla \cdot \tau \\
\textbf{Energy} \quad & (\rho E)_t + \nabla \cdot (\rho UE + pU) = \nabla \cdot \kappa \nabla T + \nabla \cdot \tau U \\
& \qquad + \sum_m \nabla \cdot (\rho h_m D_m \nabla Y_m) \\
\textbf{Species} \quad & (\rho Y_m)_t + \nabla \cdot (\rho U Y_m) = \nabla \cdot (\rho D_m \nabla Y_m) + \dot{\omega}_m
\end{aligned}$$

Augmented with

- Thermodynamics
- Reaction kinetics
- Transport coefficients

Need to preserve chemical and transport fidelity

CCSE Berkeley Lab

Mathematical formulation

Exploit natural separation of scales between fluid motion and acoustic wave propagation

Low Mach number model, $M = U/c \ll 1$ (Rehm & Baum 1978, Majda & Sethian 1985)

Start with the compressible Navier-Stokes equations for multicomponent reacting flow, and expand in the Mach number, $M = U/c$.

Asymptotic analysis shows that:

$$p(\vec{x}, t) = p_0(t) + \pi(\vec{x}, t) \quad \text{where} \quad \pi/p_0 \sim \mathcal{O}(M^2)$$

- p_0 does not affect local dynamics, π does not affect thermodynamics
- For open containers p_0 is constant
- Pressure field is instanteously equilibrated – removed acoustic wave propagation

Low Mach number formulation

Exploit separation of scales between fluid motion and acoustic wave propagation

Momentum $$\frac{\partial \rho U}{\partial t} + \nabla \cdot (\rho U U) = -\nabla \pi + \nabla \cdot \tau$$

Species $$\frac{\partial (\rho Y_m)}{\partial t} + \nabla \cdot (\rho U Y_m) = \nabla \cdot (\rho D_m \nabla Y_m) + \dot{\omega}_m$$

Mass $$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho U) = 0$$

Energy $$\frac{\partial \rho h}{\partial t} + \nabla \cdot \left(\rho h \vec{U} \right) = \nabla \cdot (\lambda \nabla T) + \sum_m \nabla \cdot (\rho h_m D_m \nabla Y_m)$$

Equation of state $p_0 = \rho \mathcal{R} T \sum_m \frac{Y_m}{W_m}$ constrains the evolution

Differentiation of EOS expresses constraint in the form

$$\nabla \cdot U = S$$

where S is a function of the thermodynamic state of the system

CCSE Berkeley Lab

Generalized projection formulation

Fractional step scheme

- Advance velocity and thermodynamic variables
  - Specialized advection algorithms
  - Diffusion
  - Stiff reactions
- Project solution back onto constraint – variable coefficient elliptic PDE, multigrid

Stiff kinetics relative to fluid dynamical time scales

$$\frac{\partial(\rho Y_m)}{\partial t} + \nabla \cdot (\rho U Y_m) = \nabla \cdot (\rho D_m \nabla Y_m) + \dot{\omega}_m$$

$$\frac{\partial(\rho h)}{\partial t} + \nabla \cdot (\rho U h) = \nabla \cdot (\lambda \nabla T) + \sum_m \nabla \cdot (\rho h_m D_m \nabla Y_m)$$

Operator split approach

- Chemistry $\Rightarrow \Delta t/2$
- Advection – Diffusion $\Rightarrow \Delta t$
- Chemistry $\Rightarrow \Delta t/2$

Decomposes problem into combination of processes, each discretized using algorithms designed to exploit mathematical structure of the process

CCSE Berkeley Lab

Block-structured AMR

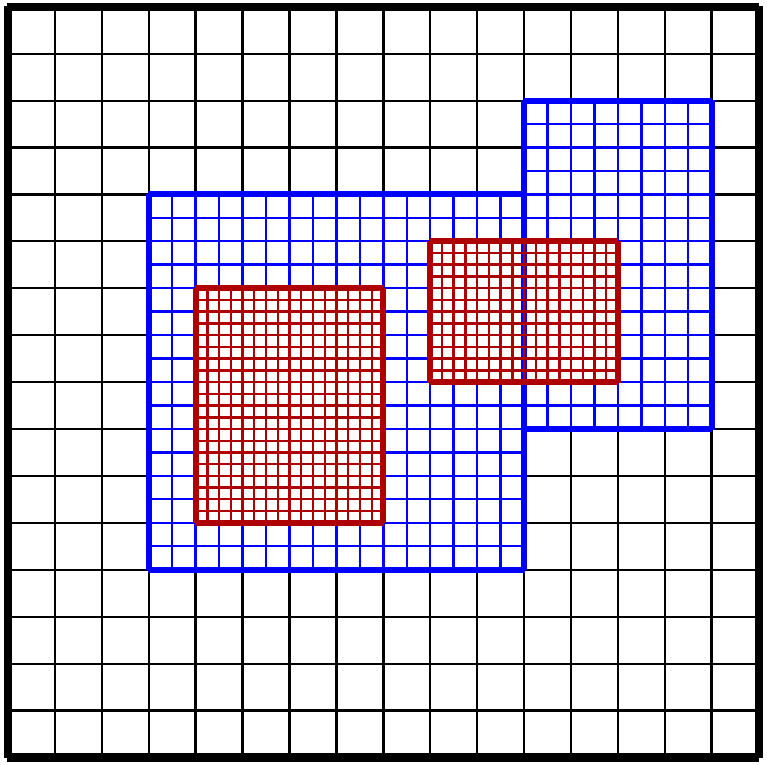
AMR – exploit varying resolution requirements in space and time

Block-structured hierarchical grids

- Amortize irregular work

Each grid patch (2D or 3D)

- Logically rectangular, structured
- Refined in space and (possibly) time by evenly dividing coarse grid cells
- Dynamically created/destroyed



2D adaptive grid hierarchy

How do we integrate different types of PDE's on this type of grid structure

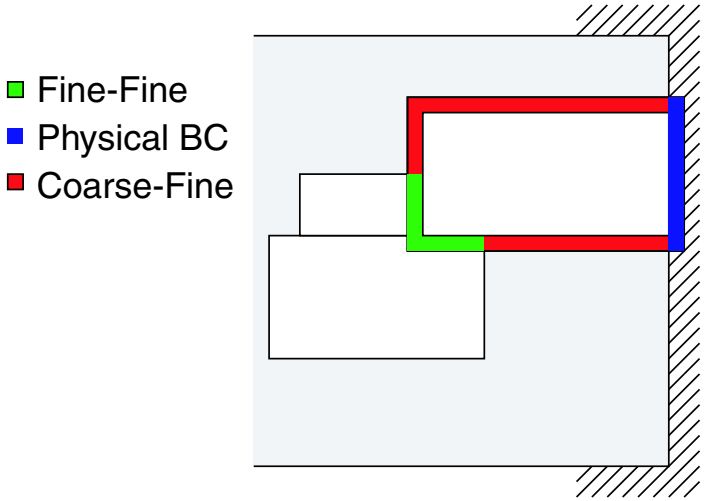
Consider a simple case – Hyperbolic Conservation Laws

CCSE Berkeley Lab

AMR for conservation laws

Conservative explicit finite volume scheme

$$U_{i,j}^{n+1} = U_{i,j}^{n} - \frac{\Delta t}{\Delta x}(F_{i+1/2,j} - F_{i-1/2,j}) - \frac{\Delta t}{\Delta y}(G_{i,j+1/2} - G_{i,j-1/2})$$



Recursive integration with subcycling in time

- Integrate each grid patch separately
- Fill ghost cells for next finer level, interpolating in space and time from coarser grid where needed
- Integrate fine grid for r time steps

- Berger and Colella, JCP 1989
- Bell, Berger, Saltzman, Welcome, JCP 1994

Coarse and fine grids are at the same time but the overall process isn't conservative.

At c-f edges flux used on the coarse grid and average of fine grid fluxes don't agree

Reflux to make overall integration conservative – update coarse grid with difference in coarse and fine fluxes

$$\Delta x_c \Delta y_c U^c = \Delta x_c \Delta y_c U^c - \Delta t^c A^c F^c + \sum \Delta t^f A^f F^f$$

CCSE Berkeley Lab

AMR Discretization Design

AMR discretization – solve on different levels separately

- Integrate on coarse grid
- Use coarse grid to supply Dirichlet data for fine grid at coarse / fine boundary
- Synchronize to correct errors that arise from advancing grids at different levels separately
  - Errors take the form of flux mismatches at the coarse/fine interface

Synchronization:

- Define what is meant by the solution on the grid hierarchy
- Identify the errors that result from solving the equations on each level of the hierarchy “independently” (motivated by subcycling in time)
- Solve correction equation(s) to “fix” the solution

CCSE Berkeley Lab

Elliptic AMR

Look at 1d (degnerate) example

$$-\phi_{xx} = \rho$$

where ρ is a discrete approximation to the derivative of a δ function at the center of the domain

$$\rho^f_J = -\alpha \qquad \rho^f_{J+1} = \alpha$$

but $\rho^c \equiv 0$

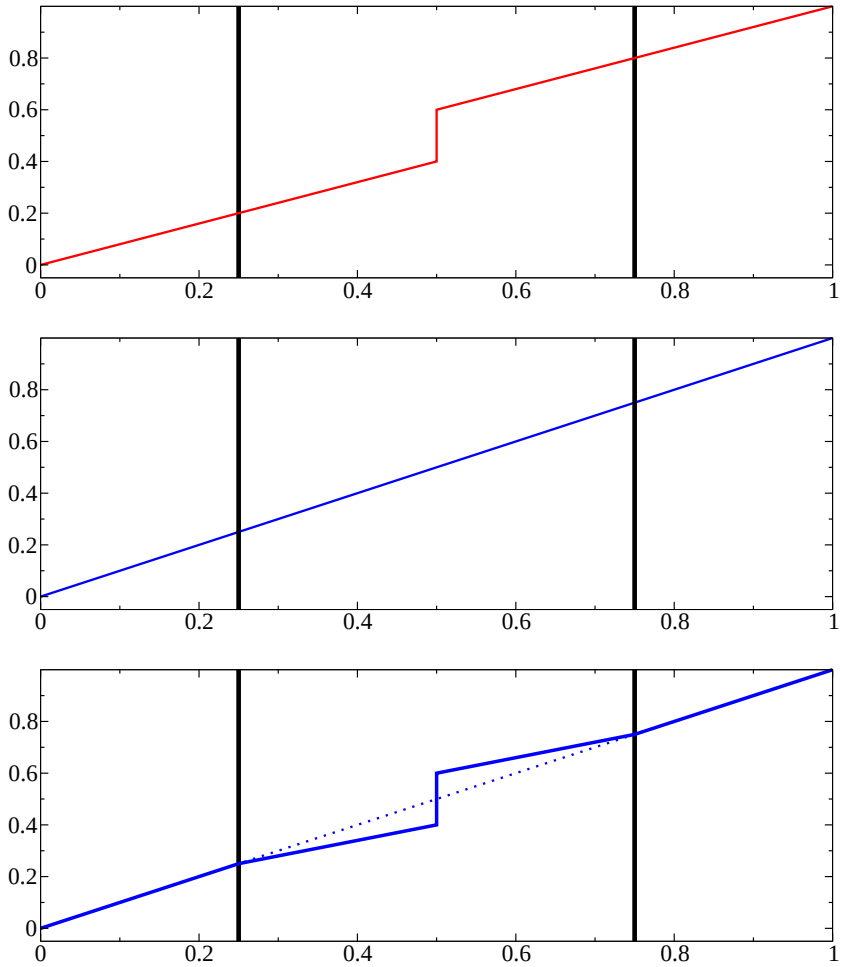
Define a composite discretization

$$L^{c-f}\phi^{c-f} = \rho^{c-f}$$

and solve

Apply design principles above

- Solve $L^c\bar{\phi}^c = \rho^c$
- Solve $L^f\bar{\phi}^f = \rho^f$ using Dirichlet boundary conditions at $c - f$ interface



Elliptic AMR – cont'd

How do we correct the solution

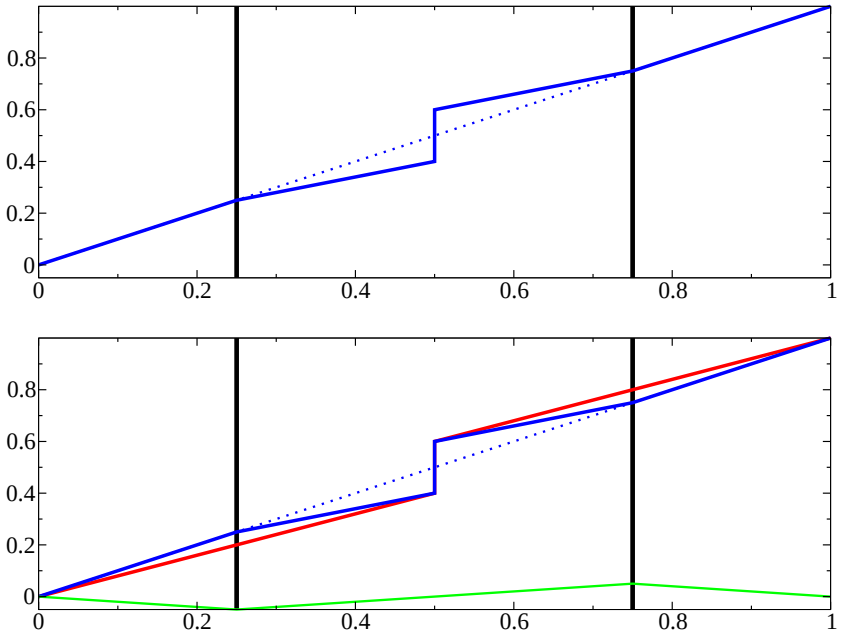
If we define $e = \phi - \bar{\phi}$ then

$$L^{c-f} e = R$$

where $R = 0$ except at $c - f$ boundary where the it is proportional to the jump in ϕ_x.

Solve for e and form $\phi = \bar{\phi} + e$

- e exactly corrects the mismatch
- Residual is localized to the $c - f$ boundary but correction is global
- The error equation is a discrete layer potential problem
- e is a discrete harmonic function on the fine grid $\rightarrow$ solve only on coarse grid and interpolate



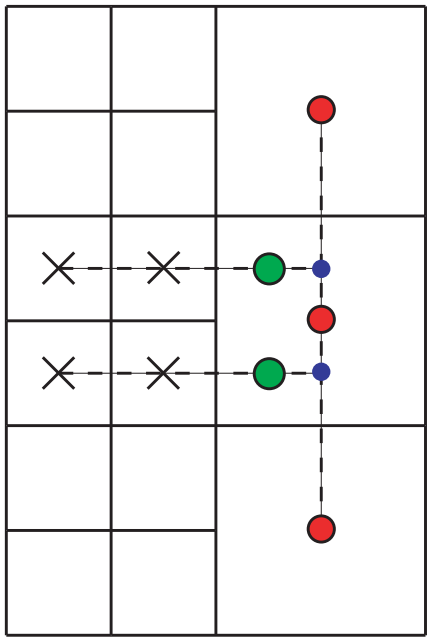
Spatial accuracy – cell-centered

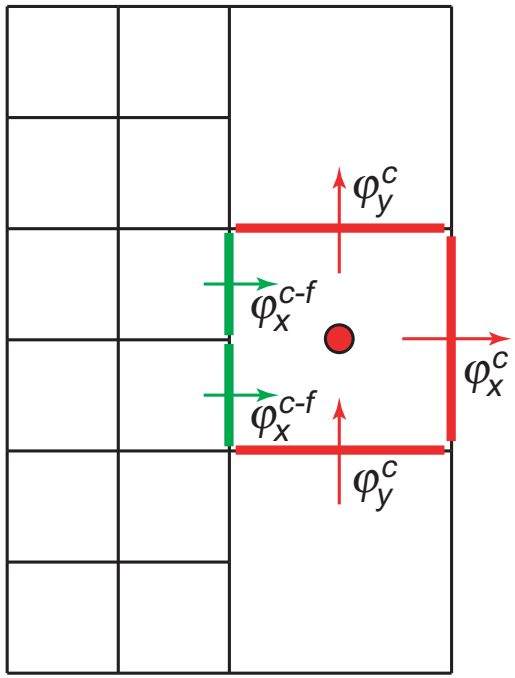
Modified equation gives

$$\psi^{comp} = \psi^{exact} + \Delta^{-1}\tau^{comp}$$

where τ is a *local* function of the solution derivatives.

Simple interpolation formulae are not sufficiently accurate for second-order operators





Hybrid algorithms

How far can we push this paradigm?

Develop a hybrid algorithm for fluid mechanics that couples a particle description to a continuum description

- Use expensive molecular model where needed
- Use a cheaper continuum model in the bulk of the domain

- AMR provides a framework for such a coupling
  - AMR for fluids except change to a particle description at the finest level of the heirarchy

- Use basic AMR design paradigm for development of a hybrid method
  - How to integrate a level
  - How to synchronize levels

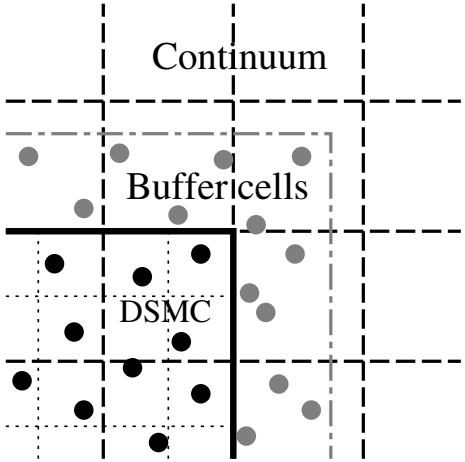
CCSE Berkeley Lab

Adaptive mesh and algorithm refinement

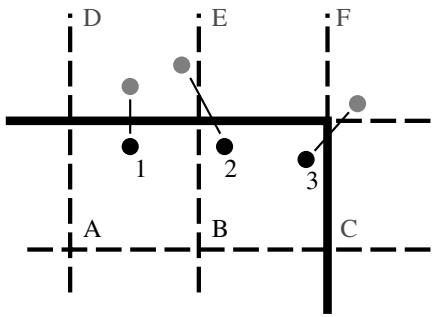
AMR approach to constructing hybrids – Garcia et al., JCP 1999

Hybrid algorithm – 2 level

- Advance continuum CNS solver
  - Accumulate flux F_C at DSMC boundary
- Advance DMSC region
  - Interpolation – Sampling from Chapman-Enskog distribution
  - Fluxes are given by particles crossing boundary of DSMC region
- Synchronize
  - Average down – moments
  - Reflux $\delta F = -\Delta t A F_C + \sum_p F_p$



DSMC boundary conditions

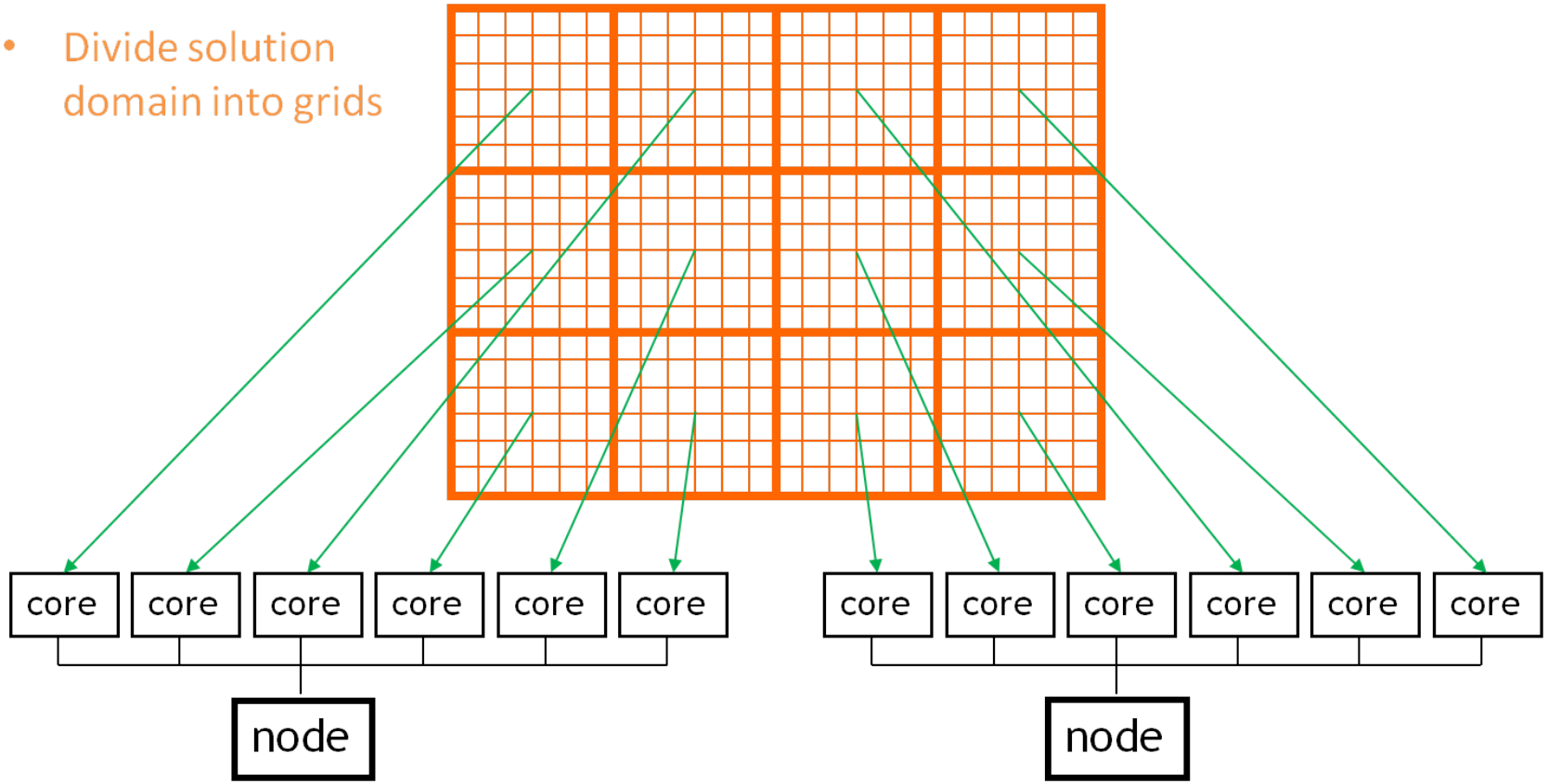


DSMC flux

Parallelization

Distribute grid patches in AMR to cores using MPI

- Divide solution domain into grids

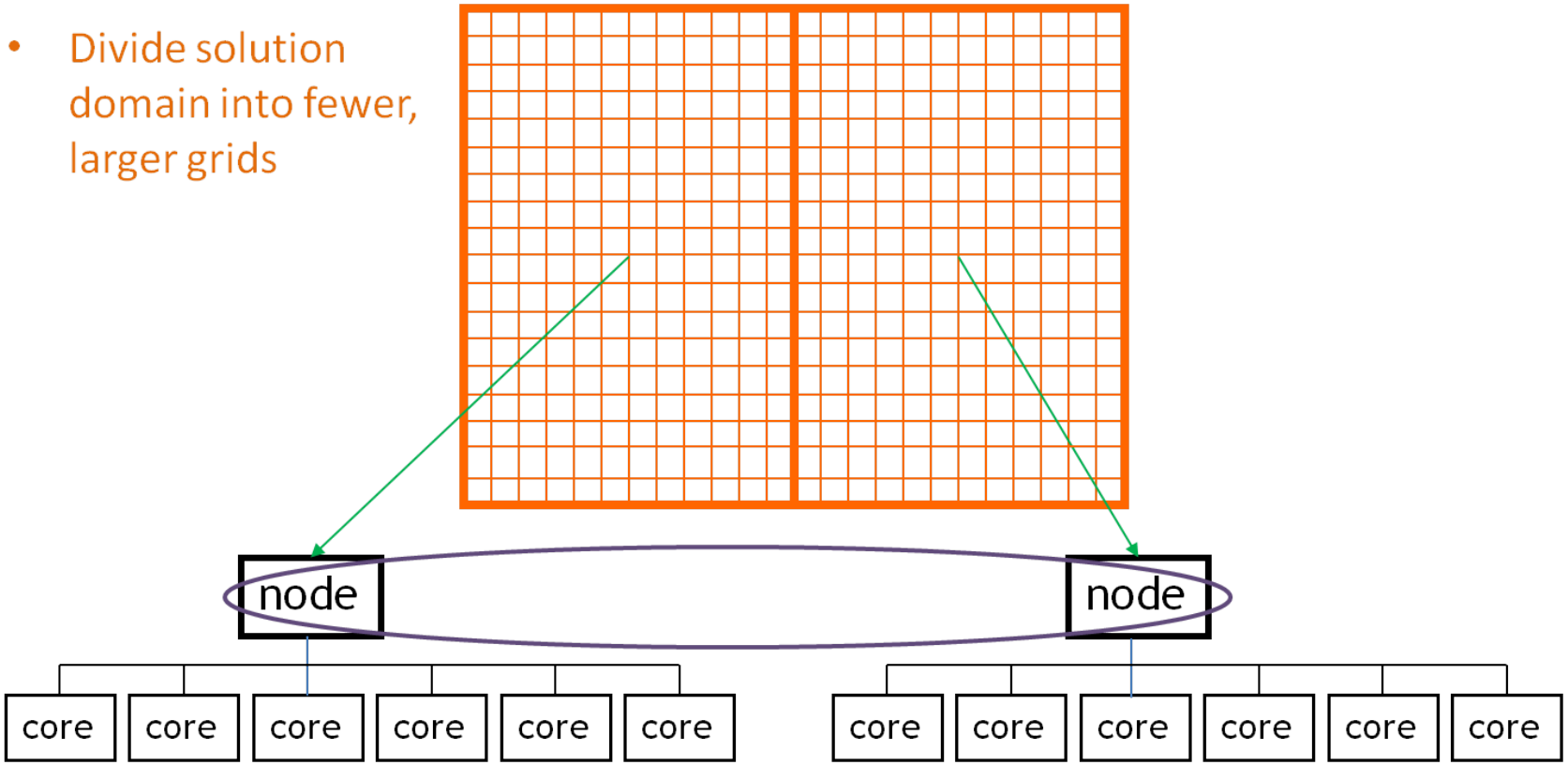


Each grid is assigned to a core

Cores communicate using MPI

Hybrid model

- Divide solution domain into fewer, larger grids



Block-structured AMR provides a natural framework for hybrid parallelization

Nodes communicate using MPI

OpenMP used to spawn threads so that cores within a node work on the grids simultaneously

CCSE Berkeley Lab

MPI versus Hybrid model

Advantages of hybrid model

- Fewer MPI processes lead to reduced communication time
- Less memory for storing ghost cell information
- Reduced work from larger grids – surface to volume effect

Disadvantages of hybrid model

- Spawning threads is expensive – makes performance worse for small core counts
- Can't hide parallelization from physics modules

With hybrid model, we have been able to scale multiphysics applications to 100K processors

CCSE Berkeley Lab

What can you build with these pieces?

LMC – Low Mach number combustion

PMAMR – AMR for subsurface (Darcy) flow (AMANZI-S – ASCEM)

- Elliptic / Parabolic pressure equation
- Component conservation laws
- Geochemistry

CASTRO – Self-gravity, radiation hydrodynamics

- Conservation laws – compressible flow with general EOS
- Elliptic self-gravity
- Flux-limited diffusion – nonlinear parabolic equations

NYX – Cosmology

- CASTRO + collisionless particles (dark matter)

MAESTRO – Low Mach number stratified flow

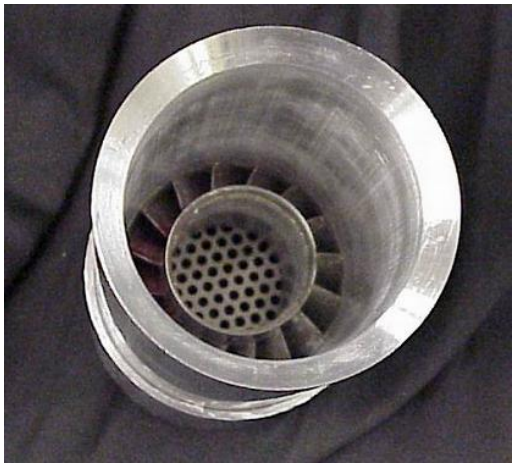
- Custom low Mach number formulation
- Compressible effects from thermal processes and stratification but no acoustic waves

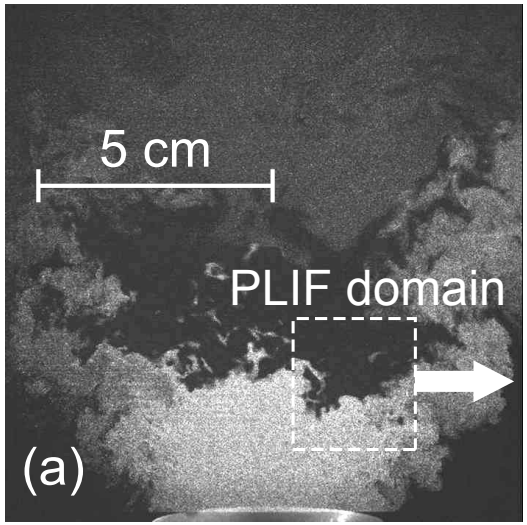
LLNS / DSMC hydrid

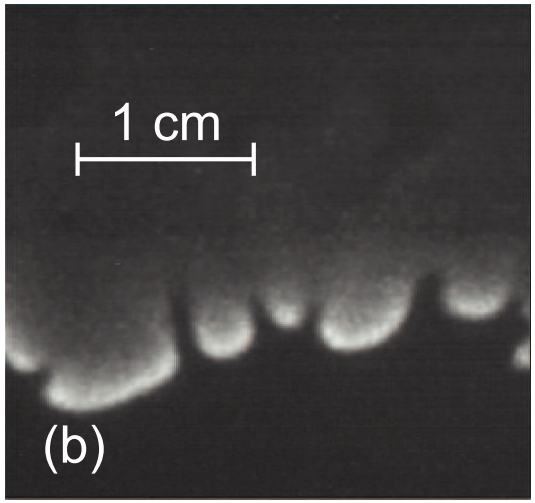
CCSE BERKELEY LAB

Hydrogen combustion

DOE's Office of Fossil Energy is interested in developing fuel-flexible turbines that can operate with hydrogen-rich fuels







Detailed simulation is needed to understand the structure of these flames

- OH PLIF shows gaps in the flame
- Standard flame models are not applicable
- Standard experimental diagnostics hard to interpret

CCSE Berkeley Lab

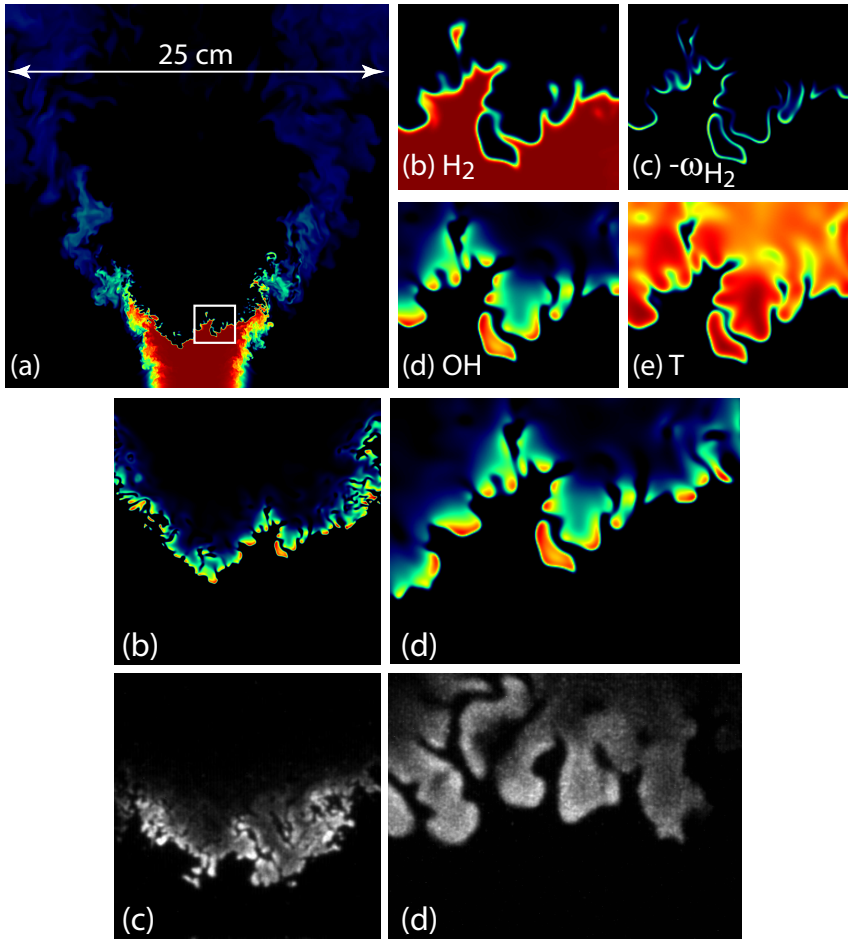
Low swirl burner simulation

Simulation of lean-premixed hydrogen flame stabilized on a low-swirl burner

- Detailed chemistry and transport
- No explicit models for turbulence or turbulence / chemistry interaction
- 25 cm x 25 cm x 25 cm domain
- Methodology enables simulation at effective resolution of 2048^3

Simulation captures cellular structure of thermodiffusively unstable lean hydrogen flames

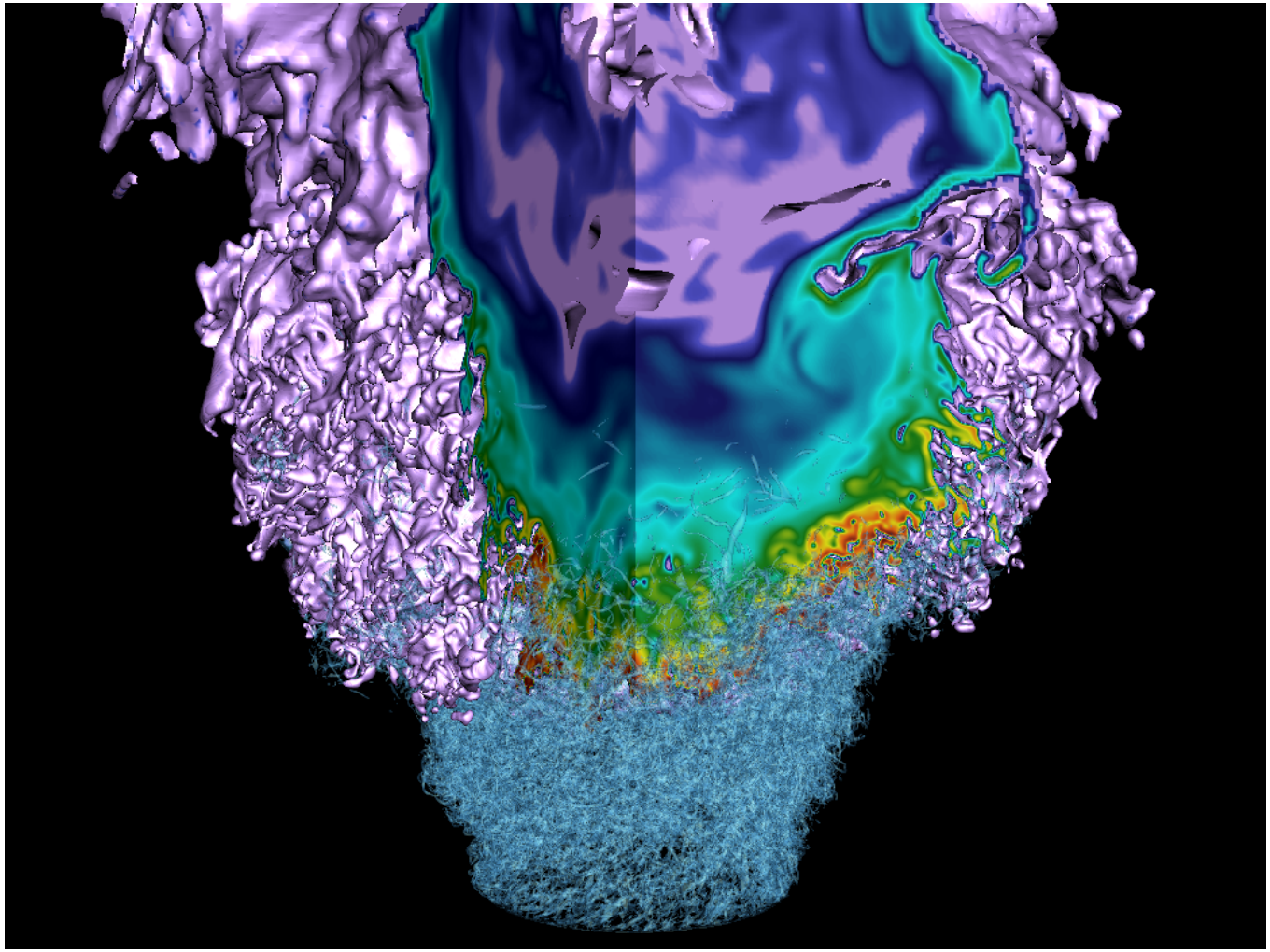
- Quantify enhanced burning from local enrichment of the fuel resulting for high H_2 diffusion
- Provide insight into the analysis of experimental diagnostics



Experiment vs. simulation

CCSE Berkeley Lab

Hydrogen low swirl burner



Animation of OH (flame marker) and vorticity

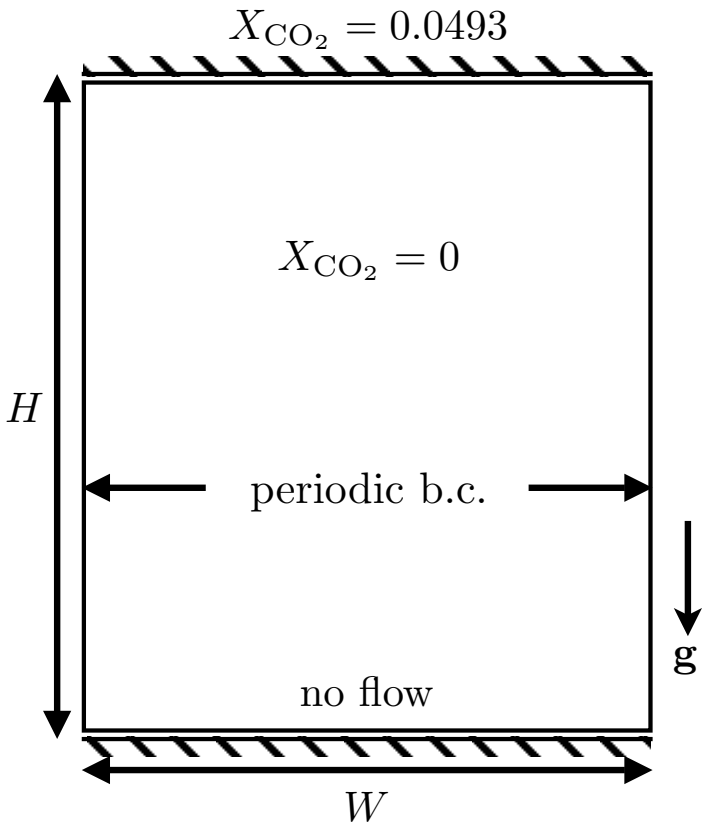
CCSE Berkeley Lab

Convection / diffusion in CO_2 sequestration

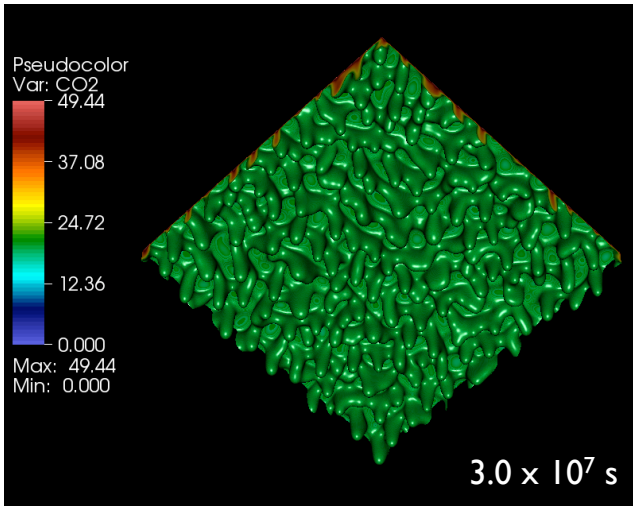
CO_2 is injected into the subsurface and forms a liquid layer on top of resident brine.

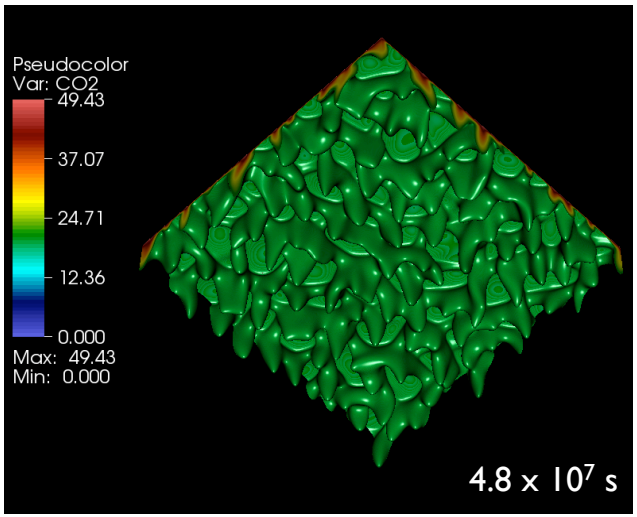
However, when CO_2 dissolves into the brine it increases the density, inducing gravity driven convection

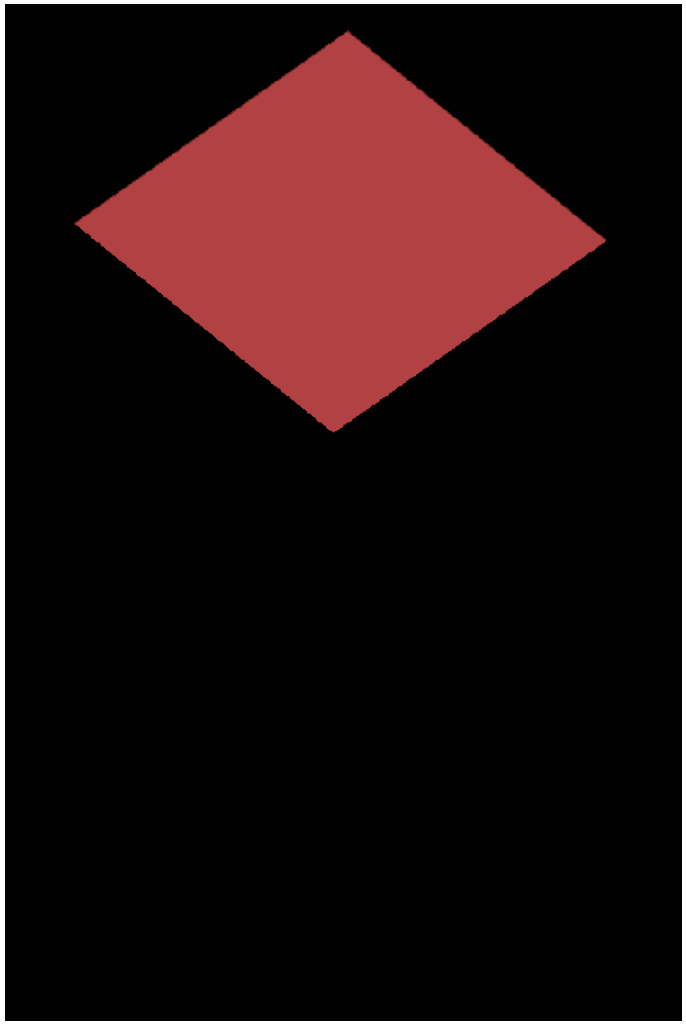
This effect can potentially increase the storage capacity of the formation



3D simulation



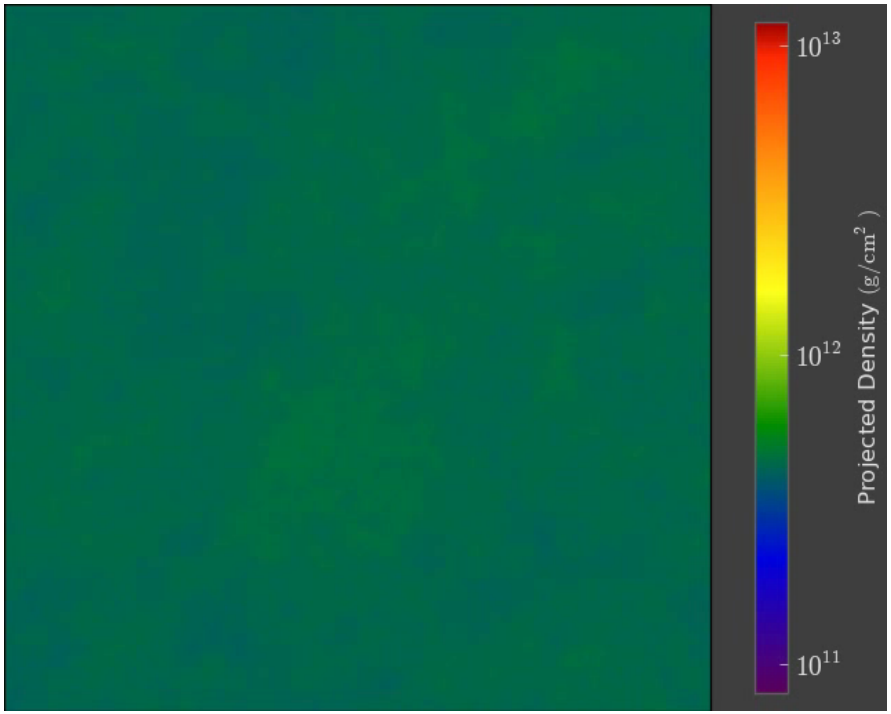




CCSE
Berkeley Lab

Cosmology

- Dark matter
  - Collisionless particles
  - Couple to baryonic matter through gravity
- Hydrodynamics equation in comoving coordinates
- Self gravity for baryonic and dark matter



Simulation of Santa Barbara cluster test problem

Animation courtesy of Casey Stark

CCSE Berkeley Lab

Exascale architectures

Power and cost constraints $\rightarrow$ a significant shift in architectural design for next generation systems

- Higher concurrency in low-power many-core, possibly heterogeneous nodes
- Much lower memory per core
- Performance based on memory access patterns and data movement, not FLOPS
- High synchronization costs

Rethink discretization methods for multiphysics applications

- More concurrency
- More locality with reduced synchronization
- Less memory / FLOP
- Analysis of algorithms currently based on a performance $\equiv$ FLOPS paradigm – can we analyze algorithms in terms of a more realistic performance model

Current programming models are inadequate for the task

- MPI reasonable for coarse-grained parallelism but at fine-grained level we write basically serial and add bandaids (OpenMP) to express parallelism
- We express codes in terms of FLOPS and let the compiler figure out the data movement
- Non-uniform memory access is already an issue but programmers can't easily control data layout

CCSE Berkeley Lab

What's next

Where to we need to go with algorithms?

Characteristics of current generation algorithms

- Second-order in space and time
- Strang split coupling between processes
- Lots of synchronization
- AMR metadata bottlenecks
  - Communication-rich multigrid
  - AMR synchronization points and bottlenecks

For exascale we would like things that are

- Higher-order in space and time
  - Implicitly requires more sophisticated coupling
  - Better way to deal with constrained systems
- Distributed AMR metadata
- Communication avoiding iterative methods
- Refactor AMR to reduce synchronization and bottlenecks

CCSE Berkeley Lab

Spectral deferred corrections – a framework for process coupling

Basic idea (Dutt, Greengard, Rokhlin): write solution of ODE, $u_t = f(t, u)$ on $[a, b]$ as

$$u(t) = u_a + \int_a^t f(\tau, u)\, d\tau \quad .$$

If we have an approximate solution $\hat{u}(t)$, we can define the residual

$$E(t, \hat{u}) = u_a + \int_a^t f(\tau, \hat{u})\, d\tau - \hat{u}(t) \quad .$$

Then, the error $\delta(t) \equiv u(t) - \hat{u}$ satisfies

$$\delta(t) = u(t) - \hat{u(t)} = (u_a + \int_a^t f(\tau, u)\, d\tau) - (u_a + \int_a^t f(\tau, \hat{u})\, d\tau - E(t, \hat{u}))$$

MISDC for advection/diffusion/reaction (Minion *et al.*):

- Treat each term separately using a simple approach
- Explicit advection, implicit diffusion, implicit reactions
- Use different time steps for each process
- Iterate SDC correction equation
  - Interpolating polynomial couples the processes

CCSE Berkeley Lab

SDC for ADR

Generalized SDC framework

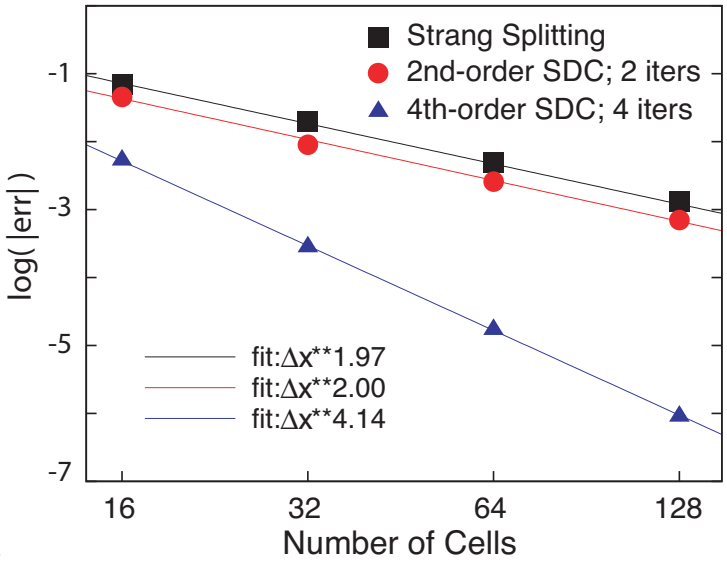
- Use different representations for each physical process
- Reuse existing components of the methodology
- Integrate reactions using VODE – Think of VODE as ”exact”

Consider a simple model problem

$$(u_i)_t + a(u_i)_x = D(u_i)_{xx} + R_i \quad i = 1, 4$$

where

$$\begin{aligned} R_1 &= -k_1 u_1 u_4 - k_2 u_1, \quad R_2 = k_1 u_1 u_4 \\ R_3 &= k_2 u_1, \quad R_4 = -k_1 u_1 u_4. \end{aligned}$$



Spectral Deferred Corrections

Develop a general framework for coupling processes in multiphysics applications

- Treat individual processes uses representation appropriate for that process
- Solver simpler subproblems but iterate to couple processes
- Higher-order in time is a change in quadrature rule
- Potential to evolve processes simultaneously

Need detailed understanding of SDC properties

- Accuracy and robustness of the overall discretization
- Convergence properties of the SDC iterations

and how these properties are related to

- Properties of processes
- Choice of quadrature rules
- Initialization and correction algorithms
- Potential acceleration strategies

Some work in this area by Minion and collaborators for ODE / DAE

This lays the ground work for higher-order temporal discretization

CCSE Berkeley Lab

SDC, AMR and Parallel

SDC and Parallel

- Simultaneous evaluation of different processes with best available approximation to other processes
- Initial iterations at lower resolution / lower fidelity
- SDC - Parallel in time (Minion, CAMCOS 2011)

Standard block-structured AMR integration advances levels sequentially from coarsest to finest

Use SDC ideas to restructure core AMR time-step strategy

- No need to complete iteration at a given level before starting the next level
- Use initial iterations on coarse grid to compute initial fine grid solutions
- This enables integration of different levels in the AMR hierarchy simultaneously
- Requires substantive changes to the underlying infrastructure to support efficient implementation

All of these ideas will reduce serial performance but they expose more concurrency and have potential for improving parallel performance

CCSE BERKELEY LAB

Summary and conclusions

Multiphysics applications characterized by a wide range of length and time scales

Approach to developing simulation methodology

- Analysis of relationship of temporal scales
- Mathematical formulation that exploits those relationships
- Numerics for each process that reflects character of the process
- Block-structured AMR for spatially varying resolution requirements
  - How to integrate PDEs in AMR grids
  - Framework for implementation of AMR algorithms
  - Structured grid AMR provides natural model for hierarchical parallelism

Shift in architectures as we move to the exascale

- FLOPS don't matter (much)
- Memory and data movement are the key
- Focus is on changes to the node-level architecture – issues are likely broader than just exascale

CCSE Berkeley Lab