Examples of Advanced PDE Applications

David E. Keyes

Department of Applied Physics & Applied Mathematics
Columbia University

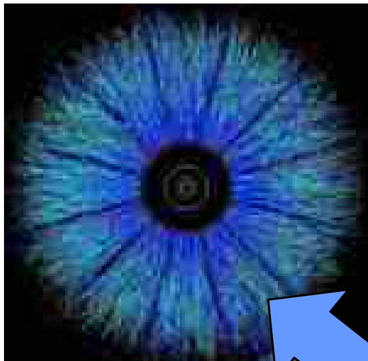
Institute for Scientific Computing Research
Lawrence Livermore National Laboratory

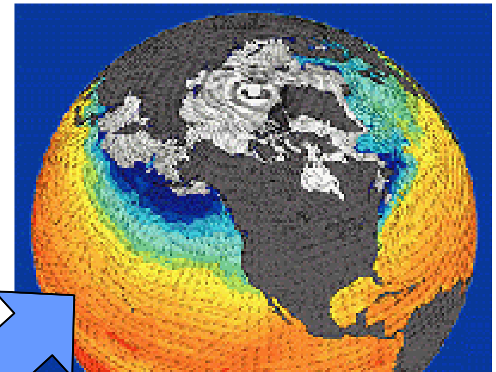
Motivation

- **Parallel solver performance a major concern for PDE simulations, e.g., of the U.S. DOE Scientific Discovery through Advanced Computing (SciDAC) program**
- **For target applications, implicit solvers may require 50% to 95% of execution time *(at least before expert overhaul for algorithmic optimality and implementation performance)***
- **Even after a "best manual practice" overhaul, solver may still require 20% to 50% of execution time**
- **The solver may hit up against both the processor scalability limit and the memory bandwidth limitation of the PDE-based application, before any other part of the code**

SciDAC apps and infrastructure

4 projects in high energy and nuclear physics

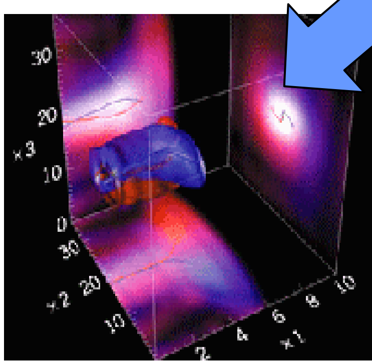


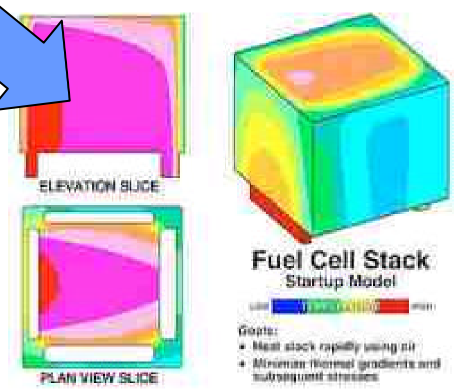


14 projects in biological and environmental research

17 projects in scientific software and network infrastructure

5 projects in fusion energy science





10 projects will in basic energy sciences

Toolchain for PDE Solvers in TOPS* project

- **Design and implementation of "solvers"**
  - **Time integrators**

    (w/ sens. anal.)

    $$f(\dot{x}, x, t, p) = 0$$

  - **Nonlinear solvers**

    (w/ sens. anal.)

    $$F(x, p) = 0$$

  - **Constrained optimizers**

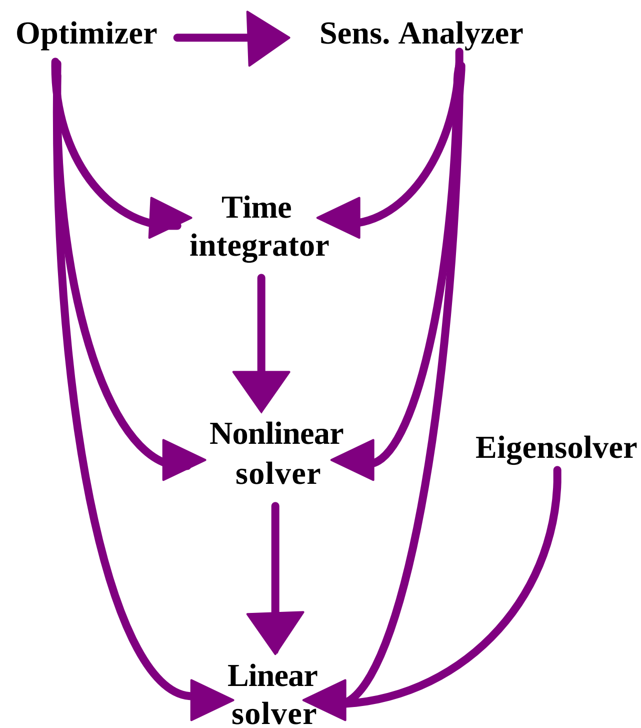
    $$\min_{u} f(x,u)\ s.t.\ F(x,u) = 0, u \geq 0$$

  - **Linear solvers**

    $$Ax = b$$

  - **Eigensolvers**

    $$Ax = lBx$$



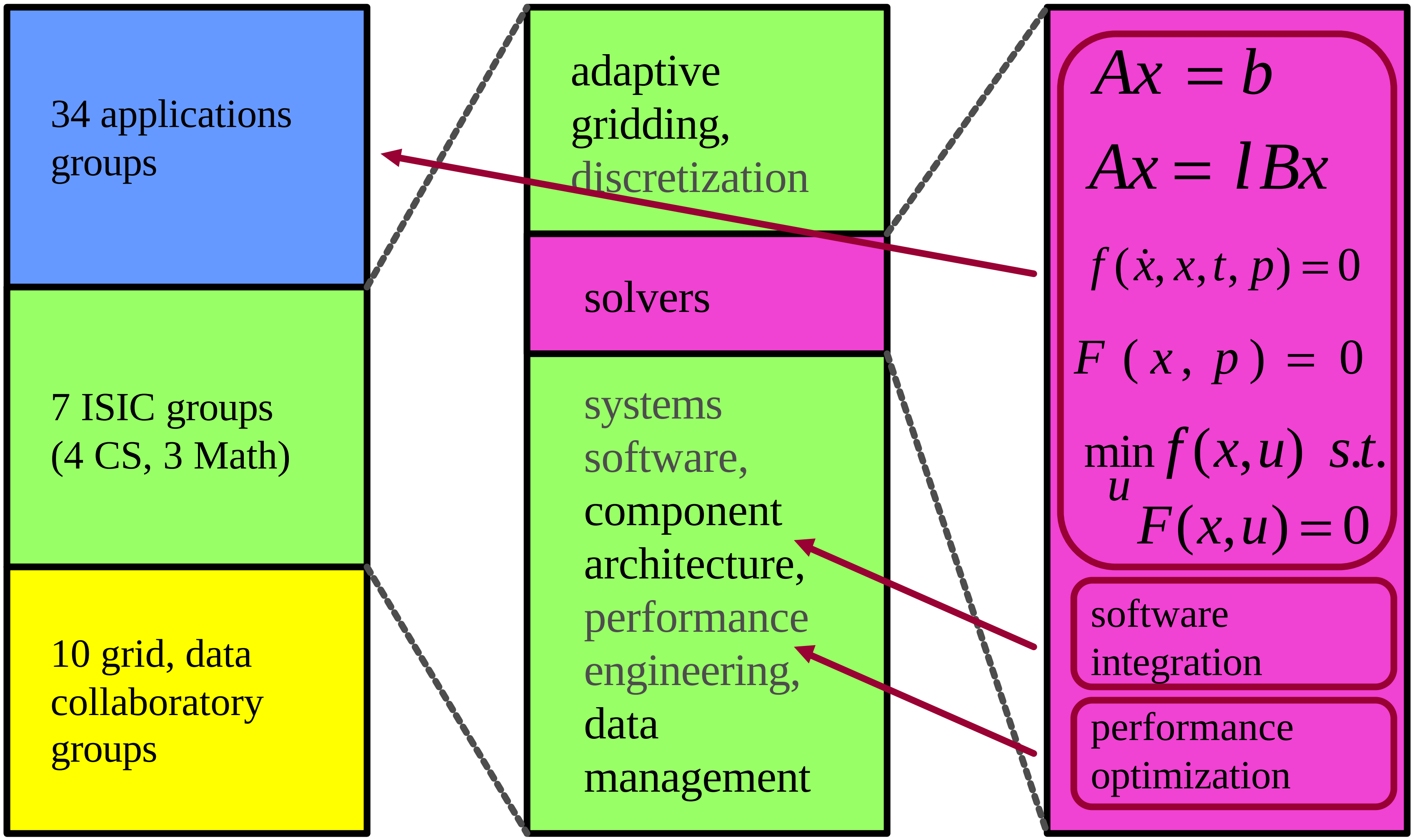
- **Software integration**
- **Performance optimization**

Indicates dependence

*Terascale Optimal PDE Simulations

SciDAC

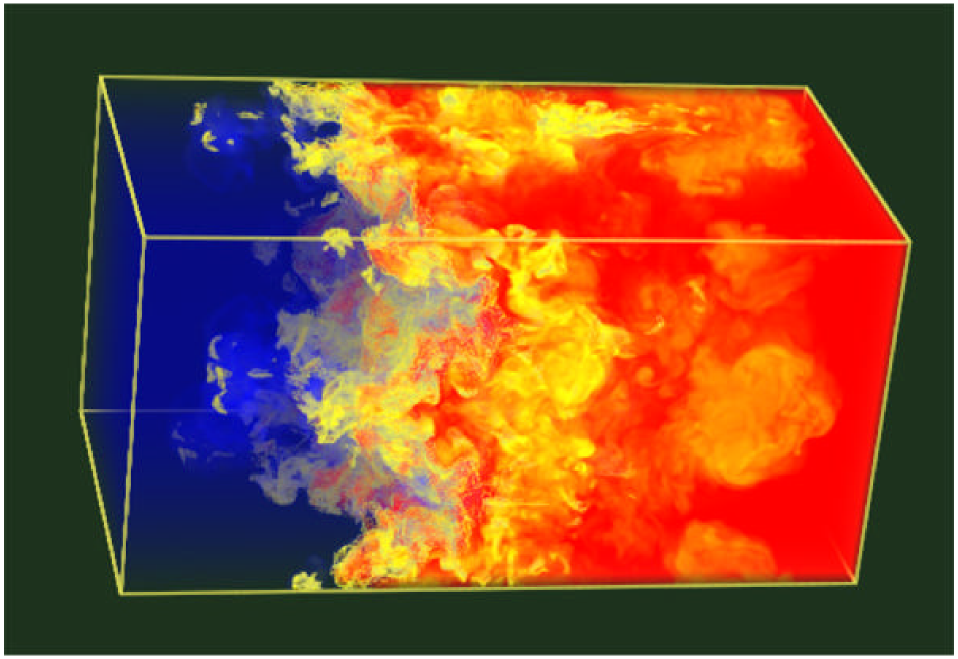
Scientific Discovery through Advanced Computing



(See posters at DD-15 Thursday evening!)

Imperative: multiple-scale applications

- Multiple spatial scales
  - interfaces, fronts, layers
  - thin relative to domain size, $d << L$
- Multiple temporal scales
  - fast waves
  - small transit times relative to convection or diffusion, $t << T$

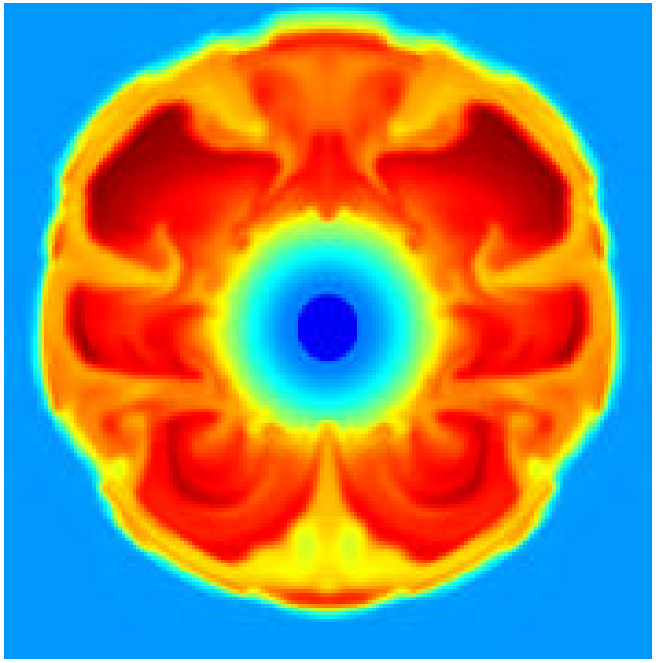


Richtmeyer-Meshkov instability, c/o A. Mirin, LLNL

- Analyst must *isolate dynamics of interest* and *model the rest* in a system that can be discretized over more modest range of scales (beyond scope of this lecture – application specific)
- May lead to local discontinuity or infinitely "stiff" subsystem requiring special treatment by the solution method (our scope)

Examples: multiple-scale applications

- **Biopolymers, nanotechnology**
  - 10^{12} range in time, from 10^{-15} sec (quantum fluctuation) to 10^{-3} sec (molecular folding time)
  - typical computational model ignores smallest scales, works on classical dynamics only, but scientists increasingly want both
- **Galaxy formation**
  - 10^{20} range in space from binary star interactions to diameter of universe
  - heroic computational model handles all scales with localized adaptive meshing



Supernova simulation, c/o A. Mezzacappa, ORNL

- **Supernovae simulation**
  - massive ranges in time and space scales for radiation, turbulent convection, diffusion, chemical reaction, nuclear reaction

Problem characteristics

- Multiple time scales
- Multiple spatial scales
- Linear ill conditioning
- Complex geometry and severe anisotropy
- Coupled physics, with essential nonlinearities
- Ambition for predictability and design from simulation

Our “pat” answers

- **Multiple time scales**
- **Multiple spatial scales**
- **Linear ill conditioning**
- **Complex geometry and severe anisotropy**
- **Coupled physics, with essential nonlinearities**
- **Ambition for predictability and design**

Stiff integrators

Adaptivity

Optimal solvers

Advanced meshing

Physics-based preconditioning

Sensitivity tools

Multiscale stress on computer architecture

- Spatial resolution stresses memory size
  - number of floating point words
  - precision of floating point words
- Temporal resolution stresses processor clock rate and/or memory bandwidth
- Both stress interprocessor latency, and *together* they *severely* stress memory bandwidth
- Less severely stressed, in principle, are memory latency and interprocessor bandwidth *(another talk)*
- But "brute force" not an option; need "good" algorithms:
  - Multiscale representations, adaptive meshes, optimal solvers, scalable everything, …

Multiscale stress on algorithms

- Spatial resolution stresses condition number
  - Ill-conditioning: small error in input may lead to large error in output
  - For self-adjoint linear systems, cond no. $k = \|A\| \cdot \|A^{-1}\|$ — related to ratio of max to min eigenvalue
  - For discrete Laplacian, $k = O(h^{-2})$
  - With improved resolution we approach the continuum limit of an unbounded inverse
- Standard iterative methods fail due to growth in iterations like $O(k)$ or $O(\sqrt{k})$; we seek $O(1)$
- Direct methods fail due to memory growth and bounded concurrency
- Solution is *domain decomposed multilevel methods*

Multiscale stress on algorithms, cont.

- **Temporal resolution stresses stiffness**
  - **Stiffness: failure to track fastest mode may lead to exponentially growing error in other modes, related to ratio of max to min eigenvalue of A in $y_t = Ay$**
  - **By definition, multiple timescale problems contain phenomena of very different relaxation rates**
  - **Certain idealized systems (e.g., incomp NS) are infinitely stiff**
- **Number of steps to finite simulated time grows, to preserve stability, regardless of accuracy requirements**
- **A solution is to *step over* fast modes by assuming quasi-equilibrium**
- **Throws temporally stiff problems into spatially ill-conditioned regime**

Newton-Krylov-Schwarz: a PDE applications "workhorse"

$$F(u) \approx F(u_c) + F'(u_c)du = 0$$
$$u = u_c + l\ du$$

$$J du = -F$$
$$du = \underset{x \in V \equiv \{F, JF, J^2F, \cdots\}}{\operatorname{argmin}} \{Jx + F\}$$

$$M^{-1} J du = -M^{-1} F$$
$$M^{-1} = \sum_i R_i^T (R_i J R_i^T)^{-1} R_i$$

Newton

nonlinear solver

asymptotically quadratic

Krylov

accelerator

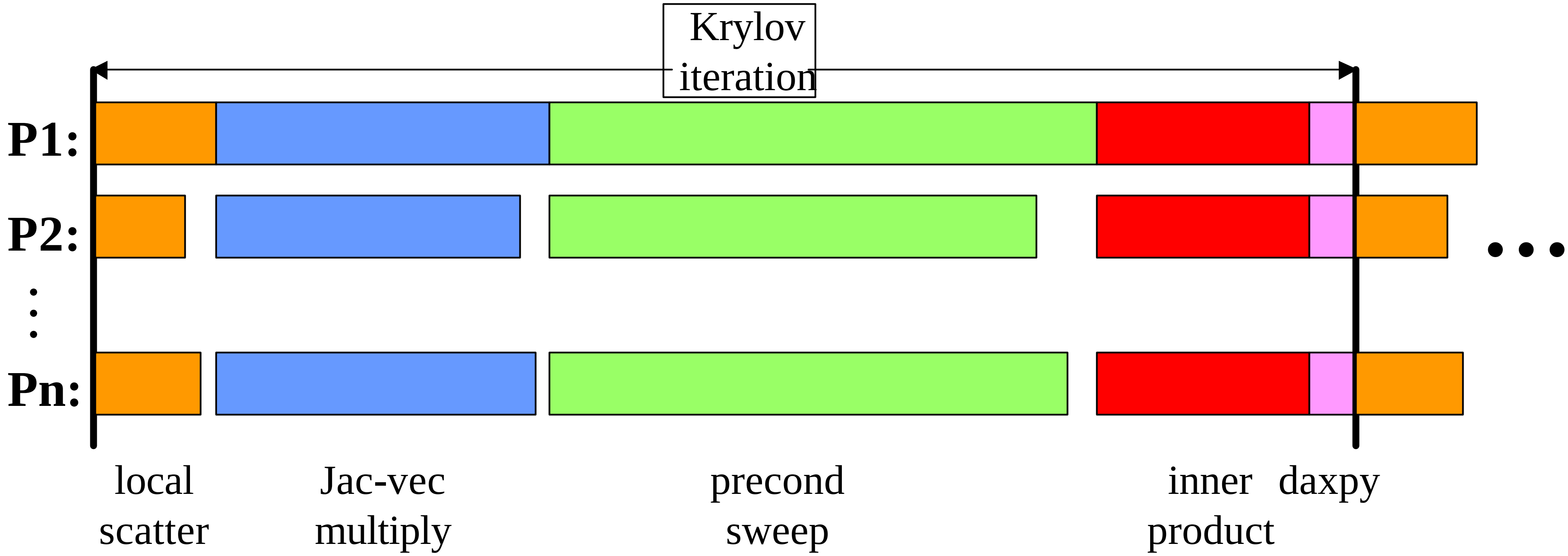
spectrally adaptive

Schwarz

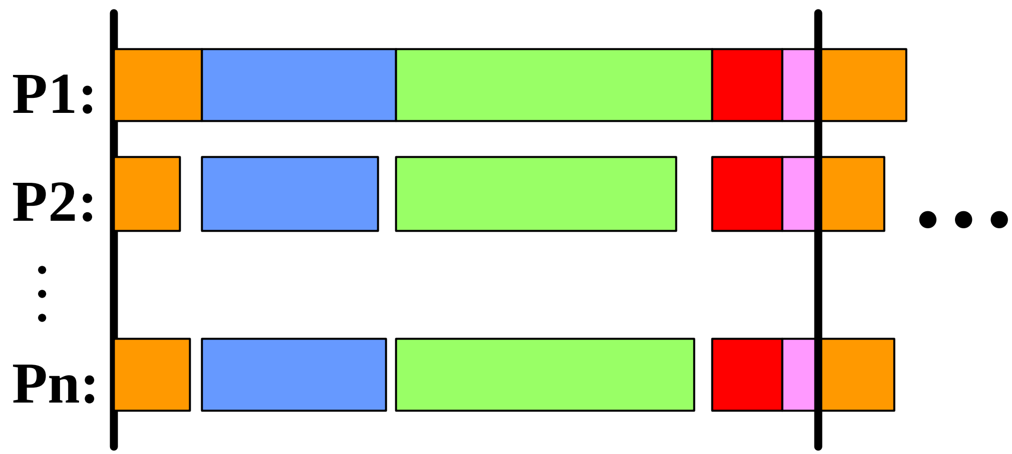
preconditioner

parallelizable

(N)KS kernel in parallel



What happens if, for instance, in this (schematicized) iteration, arithmetic speed is ***doubled***, scalar all-gather is ***quartered***, and local scatter is ***cut by one-third***? Each phase is considered separately. Answer is to the right.



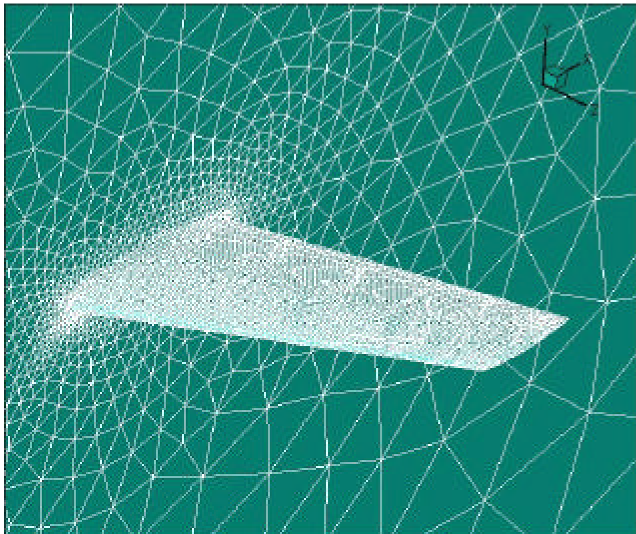
Background of FUN3D Application

- Tetrahedral vertex-centered unstructured grid code developed by W. K. Anderson (LaRC) for steady compressible and incompressible Euler and Navier-Stokes equations (with one-equation turbulence modeling)
- Used in airplane, automobile, and submarine applications for analysis and design
- Standard discretization is 2nd-order Roe scheme for convection and Galerkin for diffusion
- Newton-Krylov-Schwarz solver with global point-block-ILU preconditioning, with false timestepping for nonlinear continuation towards steady state; competitive with FAS multigrid in practice
- Legacy implementation/ordering is vector-oriented

Features of this 1999 Submission

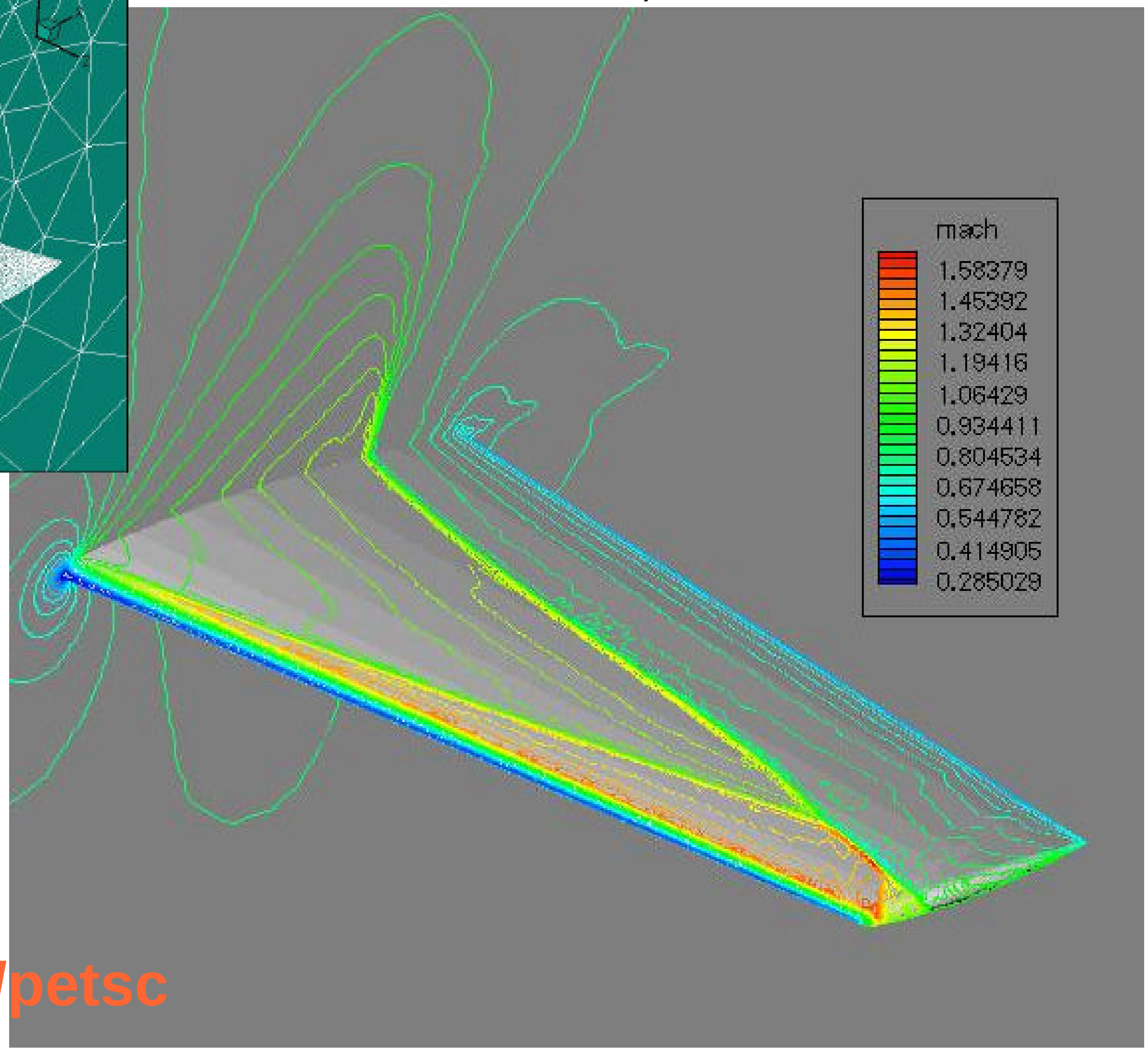
- Based on “legacy” (but contemporary) CFD application with significant F77 code reuse
- Portable, message-passing library-based parallelization, run on NT boxes through Tflop/s ASCI platforms
- Simple multithreaded extension (for ASCI Red)
- Sparse, unstructured data, implying memory indirection with only modest reuse - nothing in this category had ever advanced to Bell finalist round
- Wide applicability to other implicitly discretized multiple-scale PDE workloads - of interdisciplinary interest
- Extensive profiling has led to follow-on algorithmic research

Computational Aerodynamics



mesh c/o D. Mavriplis, ICASE

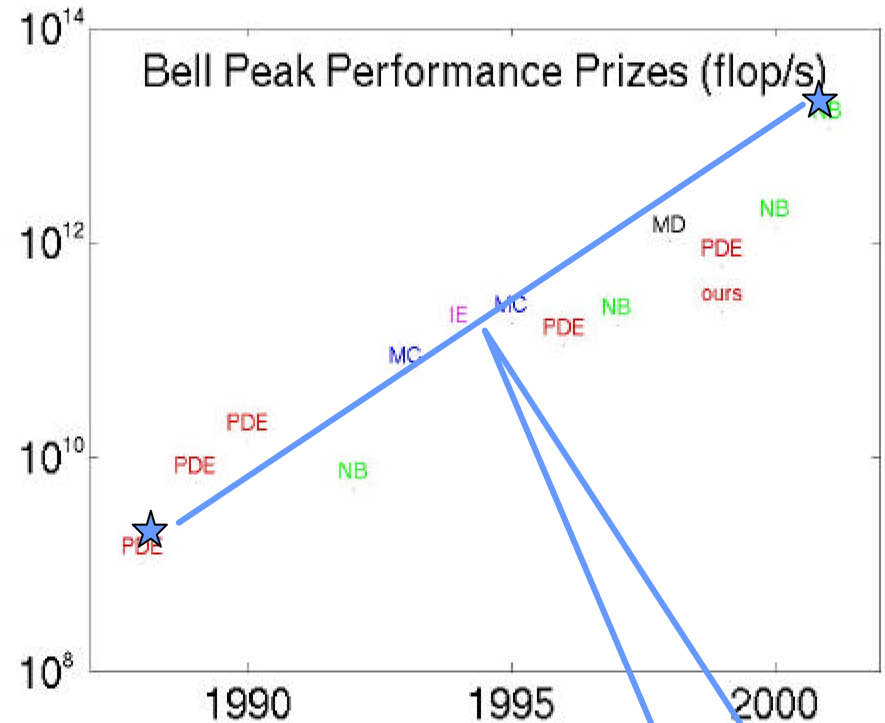
Transonic “Lambda” Shock, Mach contours on surfaces



Implemented in PETSc

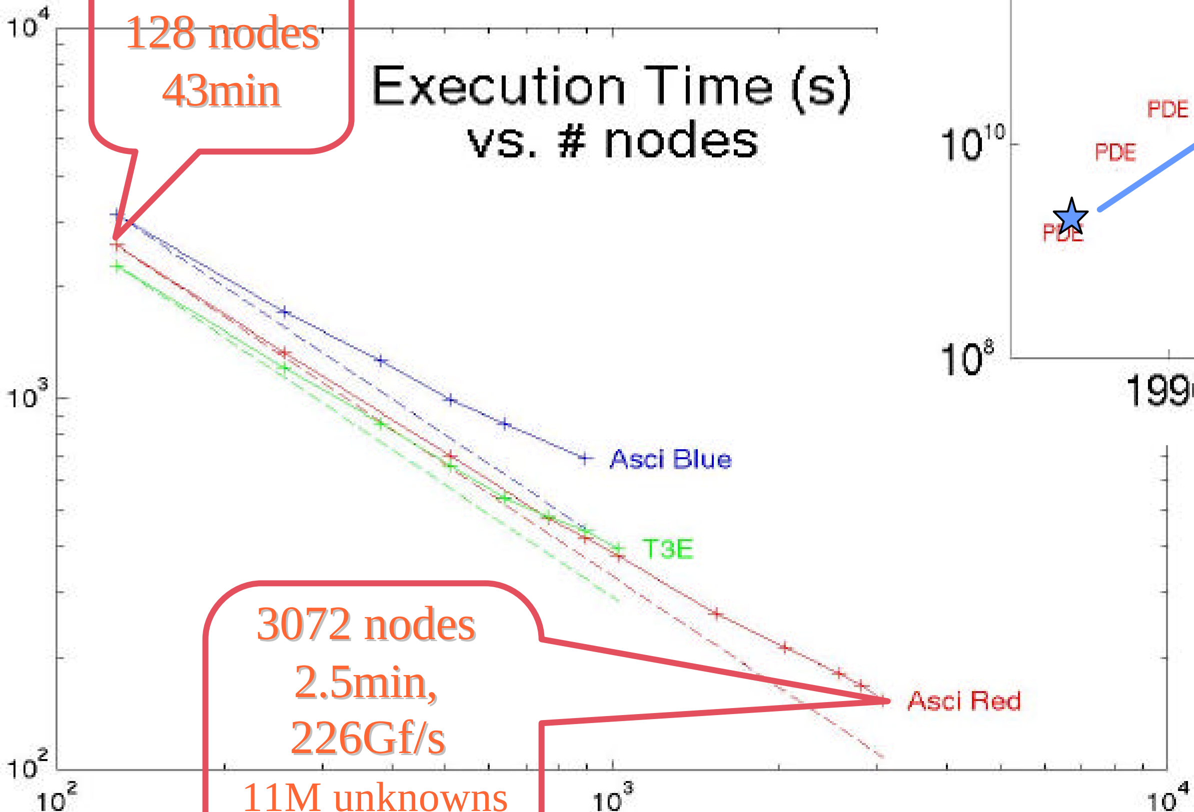
www.mcs.anl.gov/petsc

Fixed-size Parallel Scaling Results



Four orders of magnitude in 13 years

128 nodes
43min



3072 nodes
2.5min,
226Gf/s
11M unknowns
15µs/unknown
70% efficient

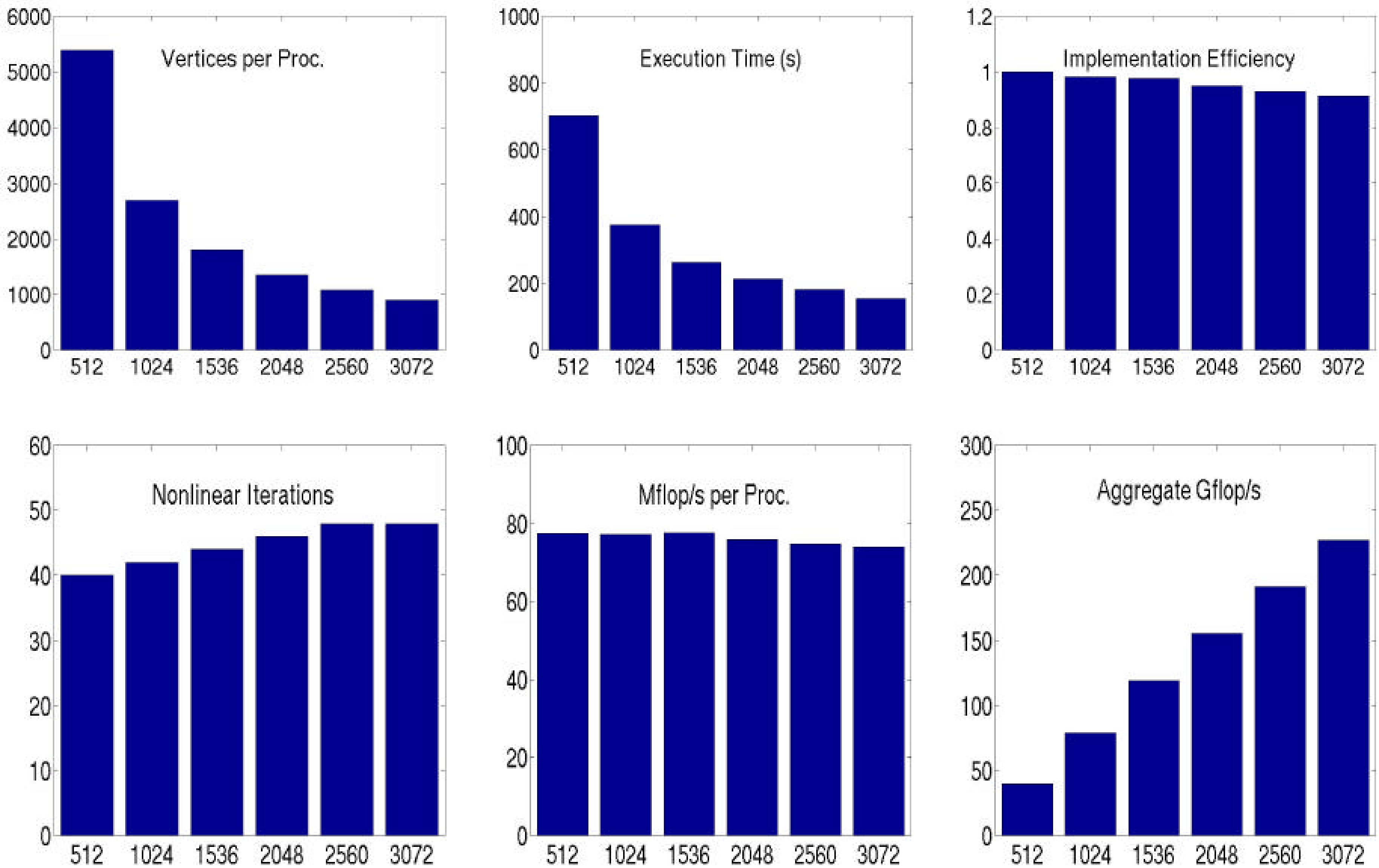
c/o K. Anderson, W. Gropp, D. Kaushik, D. Keyes and B. Smith

Bell Prize Performance History

Year	*Type*	*Application*	*Gflop/s*	*System*	*No. Procs*
1988	PDE	Structures	1.0	Cray Y-MP	8
1989	PDE	Seismic	5.6	CM-2	2,048
1990	PDE	Seismic	14	CM-2	2,048
1992	NB	Gravitation	5.4	Delta	512
1993	MC	Boltzmann	60	CM-5	1,024
1994	IE	Structures	143	Paragon	1,904
1995	MC	QCD	179	NWT	128
1996	PDE	CFD	111	NWT	160
1997	NB	Gravitation	170	ASCI Red	4,096
1998	MD	Magnetism	1,020	T3E-1200	1,536
1999	PDE	CFD	627	ASCI BluePac	5,832
2000	NB	Gravitation	1,349	GRAPE-6	96
2001	NB	Gravitation	11,550	GRAPE-6	1,024
2002	PDE	Climate	26,500	Earth Sim	5,120

Fixed-size parallel scaling results on ASCI Red

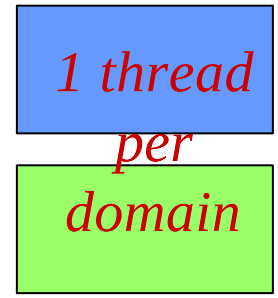
ONERA M6 Wing Test Case, Tetrahedral grid of 2.8 million vertices on up to 3072 ASCI Red Nodes (Pentium Pro 333 MHz processors)

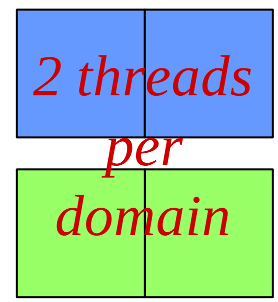


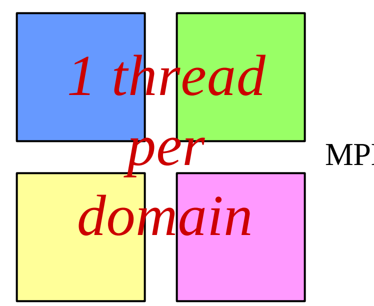
Choices in subdomain slackness

- **Table shows execution times of residual flux evaluation phase for NKS Euler simulation on ASCI Red (2 processors per node)**
- **In each paradigm, the second processor per node contributes another load/store unit and processing while sharing fixed memory bandwidth**
- **Note that 1 thread is worse than 1 MPI process, but that 2-thread performance eventually surpass 2-process performance as subdomains become small**

Nodes	MPI/OpenMP		MPI	
	1 Thr	2 Thr	1 Proc	2 Proc
256	483s	261s	456s	258s
2560	76s	39s	72s	45s
3072	66s	33s	62s	40s







Effect of overlap and fill on convergence

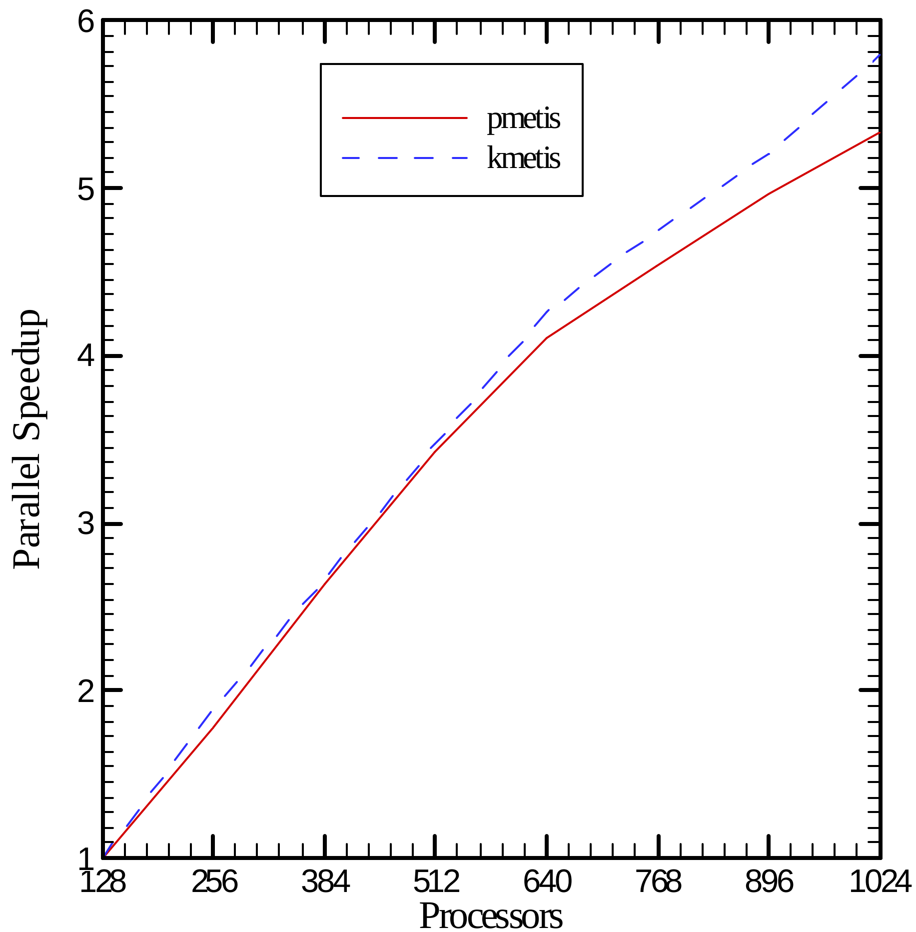
- **Convergence rate of the additive Schwarz Method (ASM) generally improves with overlap**
- **Convergence rate of ILU generally improves with higher fill level**
- **However, we need to use a proper combination of these two parameters that results in the *smallest execution time***
- **For ILU(1) on each subdomain, ONERA M6 aerodynamics problem:**

Number of Processors	Overlap					
	0		1		2	
	Time	Linear Iterations	Time	Linear Iterations	Time	Linear Iterations
32	598	674	564	549	617	532
64	334	746	335	617	359	551
128	177	807	178	630	200	555

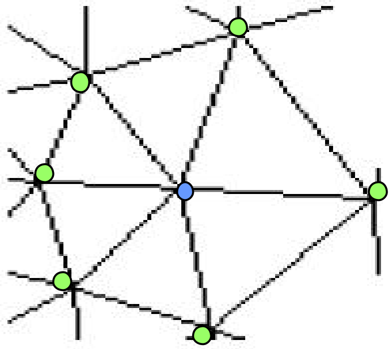
Effect of Data Partitioning Strategies

Grid of 2.8 million vertices on Cray T3E-1200

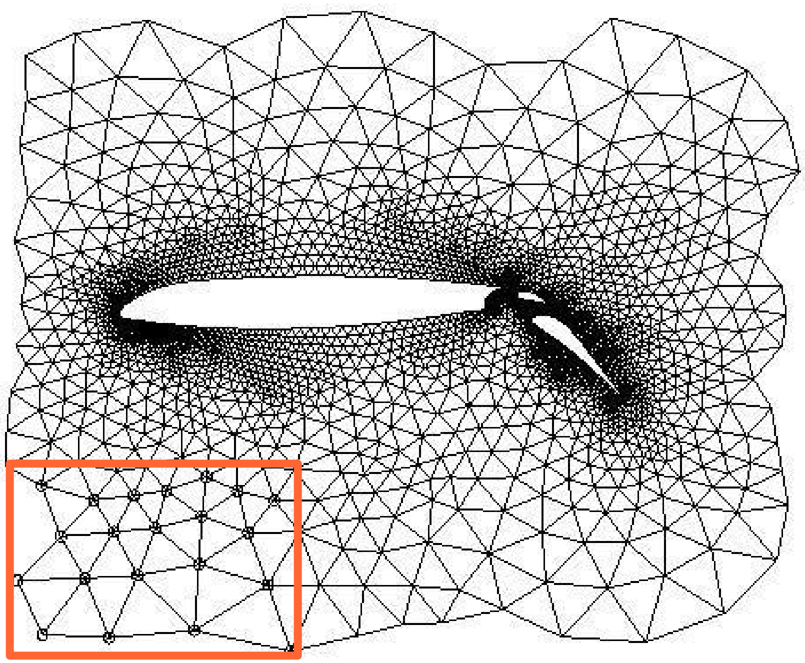
- *pmetis* attempts to balance the number of nodes and edges on each partition
- *kmetis* tries to reduce the number of non-contiguous subdomains and connectivity of the subdomains
- *kmetis* gives slightly better scalability at high end

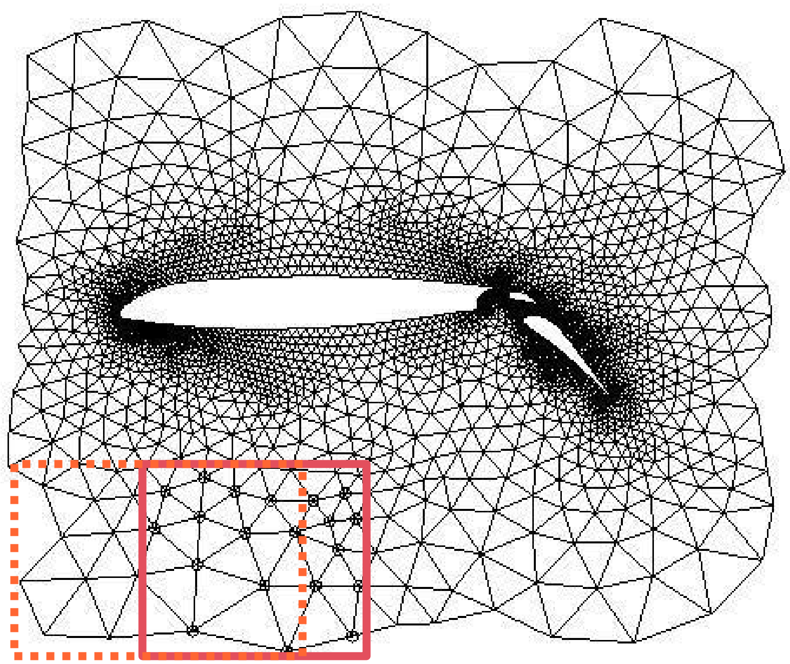


PDE workingsets

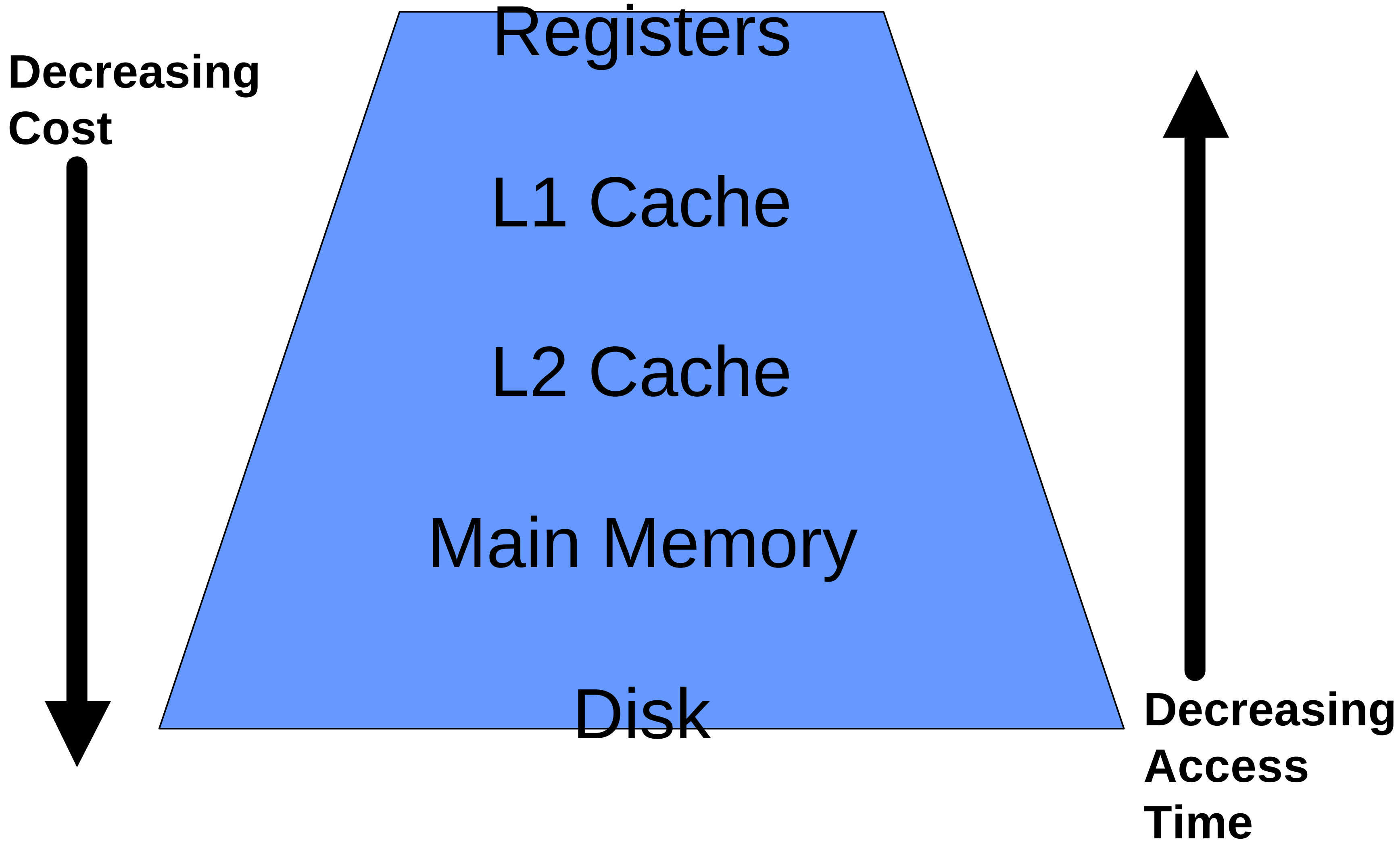


- Smallest: data for single stencil
- Largest: data for entire subdomain
- Intermediate: data for a neighborhood collection of stencils, reused as possible



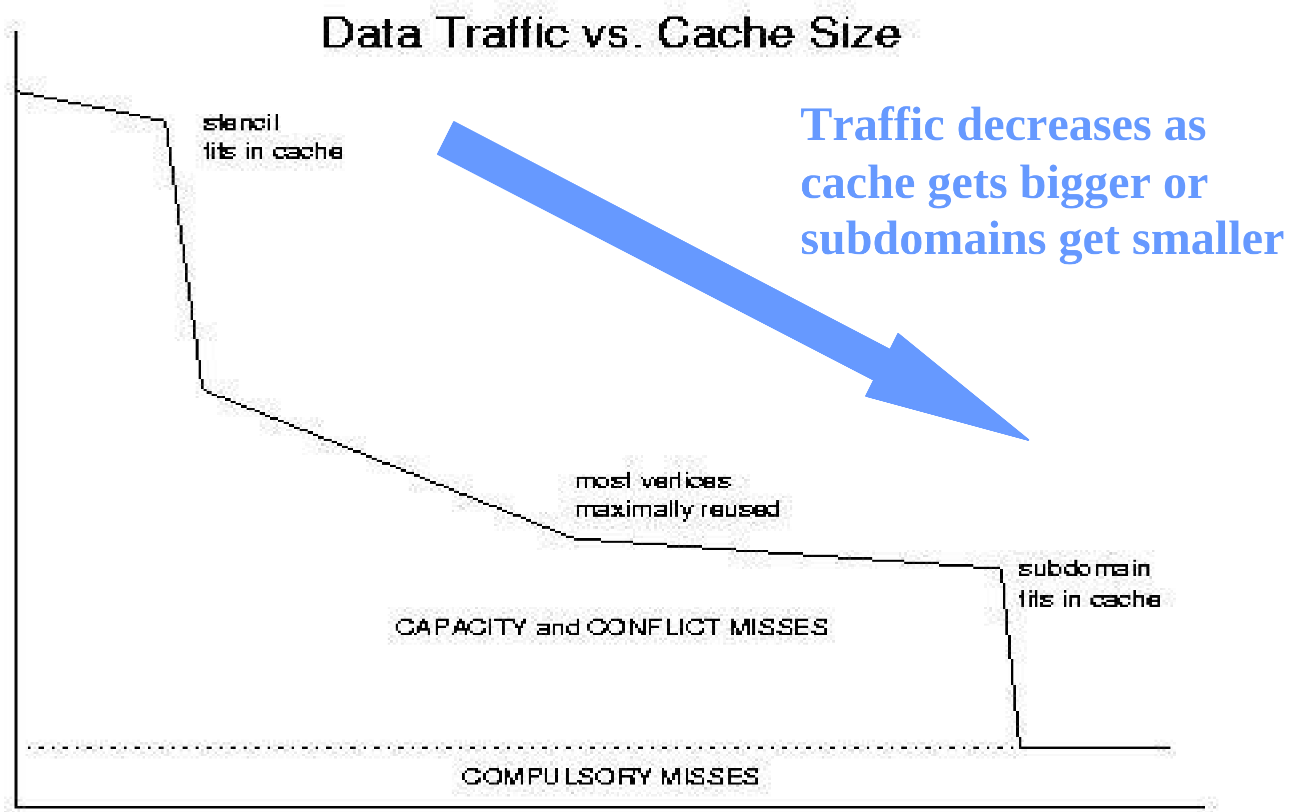


Memory Hierarchy Trapezoid



Cache traffic for PDEs

- **As successive workingsets "drop" into a level of memory, capacity (and with effort conflict) misses disappear, leaving only compulsory, reducing demand on main memory bandwidth**



Improvements from locality-based reordering

Factor of Five!

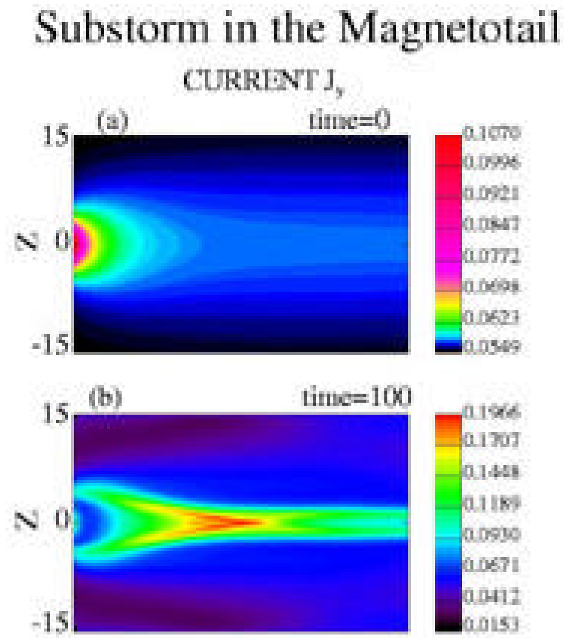
Processor	Clock MHz	Peak Mflop/s	Opt. % of Peak	Opt. Mflop/s	Reord. Only Mflop/s	Interl. only Mflop/s	Orig. Mflop/s	Orig. % of Peak
R10000	250	500	25.4	127	74	59	26	5.2
P3	200	800	20.3	163	87	68	32	4.0
P2SC (2 card)	120	480	21.4	101	51	35	13	2.7
P2SC (4 card)	120	480	24.3	117	59	40	15	3.1
604e	332	664	9.9	66	43	31	15	2.3
Alpha 21164	450	900	8.3	75	39	32	14	1.6
Alpha 21164	600	1200	7.6	91	47	37	16	1.3
Ultra II	300	600	12.5	75	42	35	18	3.0
Ultra II	360	720	13.0	94	54	47	25	3.5
Ultra II/HPC	400	800	8.9	71	47	36	20	2.5
Pent. II/LIN	400	400	20.8	83	52	47	33	8.3
Pent. II/NT	400	400	19.5	78	49	49	31	7.8
Pent. Pro	200	200	21.0	42	27	26	16	8.0
Pent. Pro	333	333	18.8	60	40	36	21	6.3

Two fusion energy applications

- **Center for Extended MHD Modeling (CEMM, based at PPPL)**
  - **Realistic toroidal geom., unstructured mesh, hybrid FE/FD discretization**
  - **Fields expanded in scalar potentials, and streamfunctions**
  - **Operator-split, linearized, w/11 potential solves in each poloidal cross-plane/step (90% exe. time)**
  - **Parallelized w/PETSc (Tang *et al.*, SIAM PP01, Chen *et al.*, SIAM AN02, Jardin *et al.*, SIAM CSE03)**
- **Want from us:**
  - **Now: scalable linear implicit solver for much higher resolution (and for AMR)**
  - **Later: fully nonlinearly implicit solvers and coupling to other codes**
- **Center for Magnetic Reconnection Studies (CMRS, based at Iowa)**
  - **Idealized 2D Cartesian geom., periodic BCs, simple FD discretization**
  - **Mix of primitive variables and streamfunctions**
  - **Sequential nonlinearly coupled explicit evolution, w/2 Poisson inversions on each timestep**
- **Want from us:**
  - **Now: scalable Jacobian-free Newton-Krylov nonlinearly implicit solver for higher resolution in 3D (and for AMR)**
  - **Later: physics-based preconditioning for nonlinearly implicit solver**

Magnetic Reconnection problem

Magnetic Reconnection: Applications to Sawtooth Oscillations, Error Field Induced Islands and the Dynamo Effect



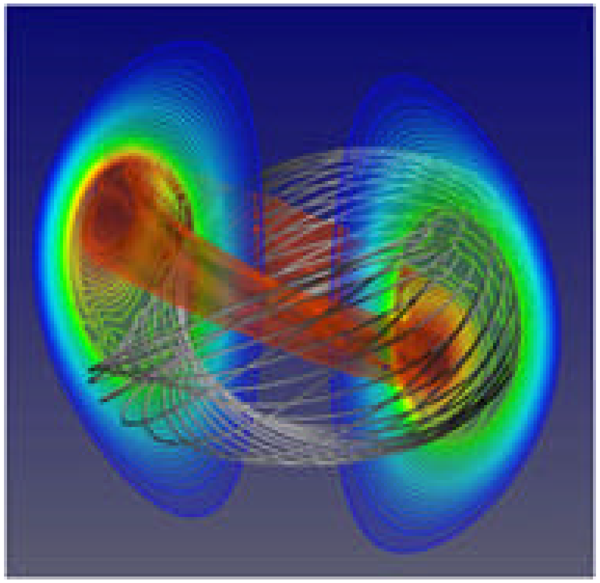
The research goals of this project include producing a unique high performance code and using this code to study magnetic reconnection in astrophysical plasmas, in smaller scale laboratory experiments, and in fusion devices. The modular code that will be developed will be a fully three-dimensional, compressible Hall MHD code with options to run in slab, cylindrical and toroidal geometry and flexible enough to allow change in algorithms as needed. The code will use adaptive grid refinement, will run on massively parallel computers, and will be portable and scalable. The research goals include studies that will provide increased understanding of sawtooth oscillations in tokamaks, magnetotail substorms, error-fields in tokamaks, reverse field pinch dynamos, astrophysical dynamos, and laboratory reconnection experiments.

Amitava Bhattacharjee
University of Iowa

from SciDAC webpages

Sawtooth Instability problem

Center for Extended Magnetohydrodynamic Modeling



This research project will develop computer codes that will enable a realistic assessment of the mechanisms leading to disruptive and other stability limits in the present and next generation of fusion devices. With an improvement in the efficiency of codes and with the extension of the leading 3D nonlinear magneto-fluid models of hot, magnetized fusion plasmas, this research will pioneer new plasma simulations of unprecedented realism and resolution. These simulations will provide new insights into low frequency, long-wavelength non-linear dynamics in hot magnetized plasmas, some of the most critical and complex phenomena in plasma and fusion science. The underlying models will be validated through cross-code and experimental comparisons.

Steve Jardin
PPPL

from SciDAC webpages

2D Hall MHD sawtooth instability (PETSc examples ex29.c and ex31.c)

(Porcelli *et al.*, 1993, 1999)

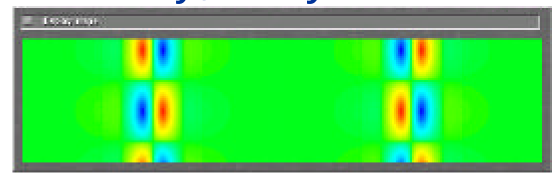
Model equations:

$$
\begin{aligned}
\frac{\partial F}{\partial t} + [\phi, F] &= \rho_s^2 [U, \psi] \\
\frac{\partial U}{\partial t} + [\phi, U] &= [J, \psi] \\
F &= \psi + d_e^2 J \\
J &= -\nabla^2 \psi \\
U &= \nabla^2 \phi
\end{aligned}
$$

$$\text{with} [A, B] = \hat{z} \cdot \nabla A \times \nabla B$$

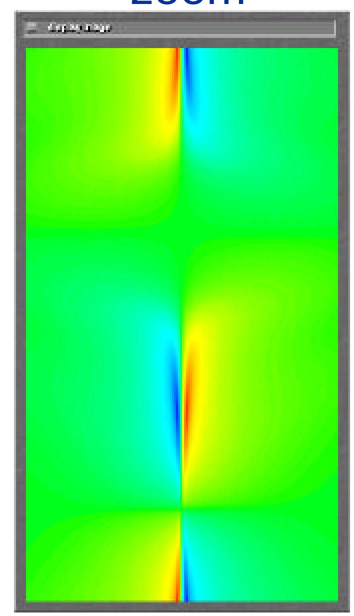
$$
\begin{aligned}
\vec{B} &= B_0 \hat{z} + \nabla \psi \times \hat{z} \\
\vec{v} &= \hat{z} \times \nabla \phi
\end{aligned}
$$

Vorticity, early time



Vorticity, later time

zoom



Equilibrium:

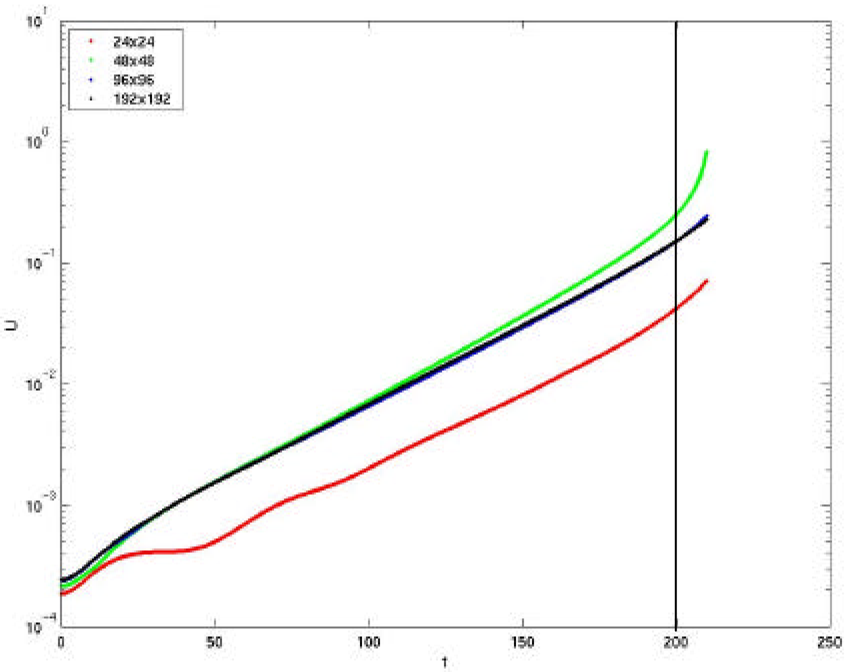
$$\phi_{eq} = U_{eq} = 0$$

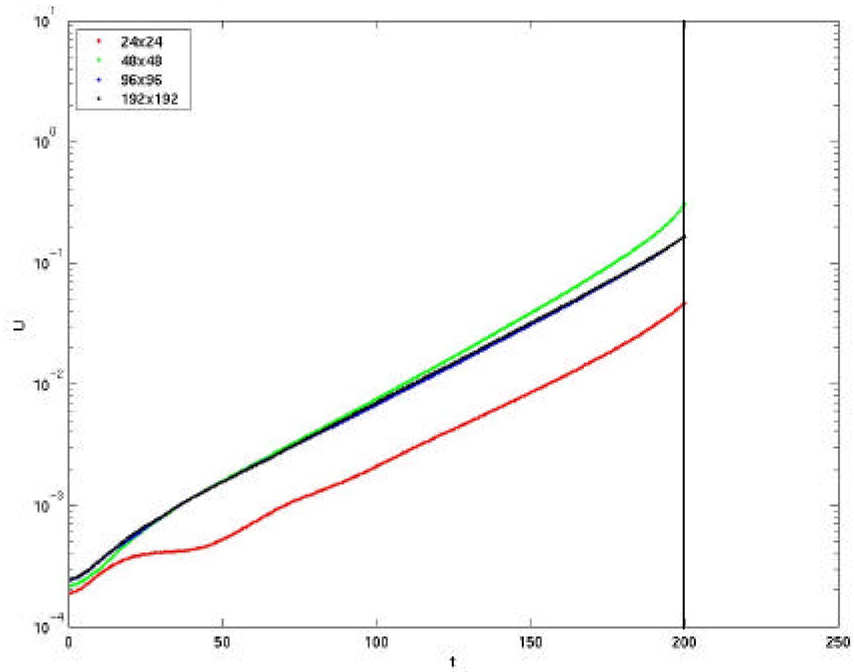
$$\psi_{eq} = J_{eq} = \cos x \quad , \quad F_{eq} = (1 + d_e^2) \cos x$$

figures c/o A. Bhattacharjee, CMRS

PETSc's DMMG in Hall MR application

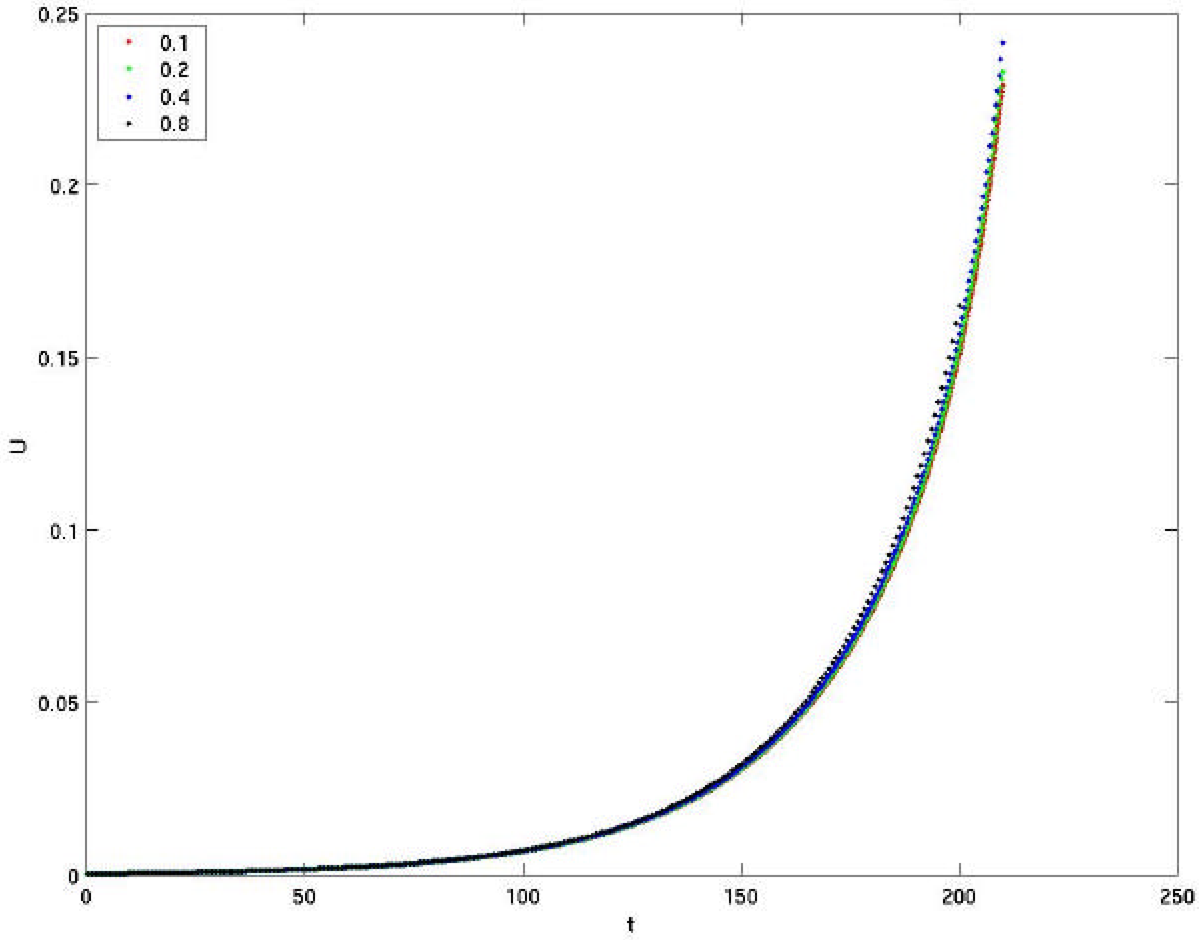
- Mesh and time refinement studies of CMRS Hall magnetic reconnection model problem (4 mesh sizes, dt=0.1 (nondimensional, near CFL limit for fastest wave) on left, dt=0.8 on right)
- Measure of functional inverse to thickness of current sheet versus time, for $0<t<200$ (nondimensional), where singularity occurs around $t=215$





PETSc's DMMG in Hall MR app., cont.

- **Implicit timestep increase studies of CMRS Hall magnetic reconnection model problem, on finest (192´192) mesh of previous slide, in absolute magnitude, rather than semi-log**



Full set of MHD equations

Physical models for these *macroscopic* dynamics are based on fluid-like magnetohydrodynamic (MHD) descriptions

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E}$$

$$\mathbf{E} = -\mathbf{V} \times \mathbf{B} + h\mathbf{J}$$

$$m_0 \mathbf{J} = \nabla \times \mathbf{B}$$

$$\frac{\partial n}{\partial t} + \nabla \cdot (n\mathbf{V}) = 0$$

$$r\left(\frac{\partial \mathbf{V}}{\partial t} + \mathbf{V} \cdot \nabla \mathbf{V}\right) = \mathbf{J} \times \mathbf{B} - \nabla p + \nabla \cdot n r \nabla \mathbf{V}$$

$$\frac{n}{g-1}\left(\frac{\partial T}{\partial t} + \mathbf{V} \cdot \nabla T\right) = -p\nabla \cdot \mathbf{V} + \nabla \cdot n\left[\left(c_{\parallel} - c_{\perp}\right)\hat{\mathbf{b}}\hat{\mathbf{b}} + c_{\perp}\mathbf{I}\right] \cdot \nabla T + Q$$

Challenges in magnetic fusion

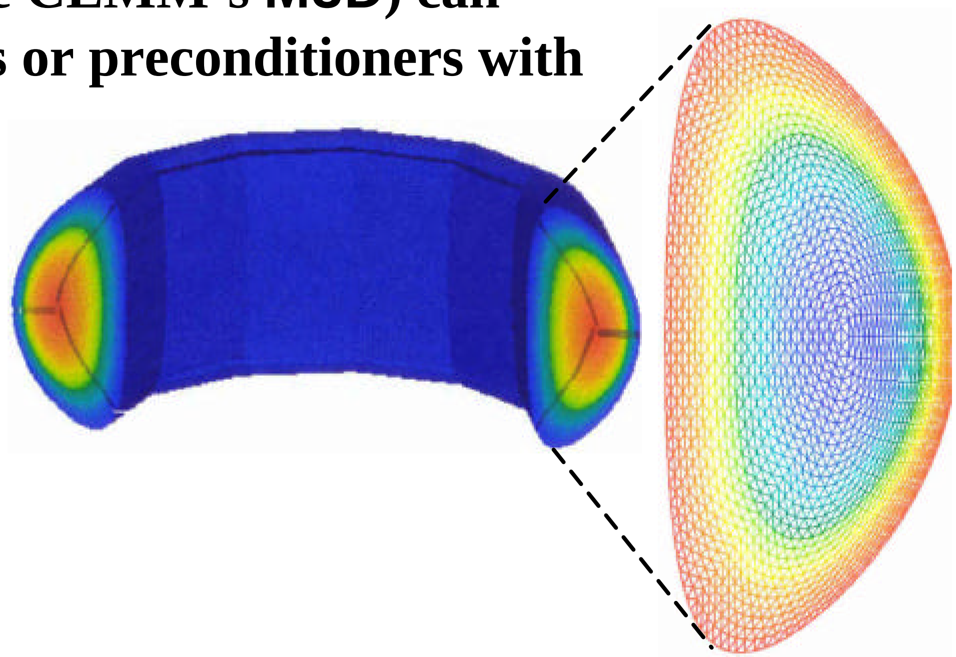
- **Conditions of interest possess two properties that pose great challenges to numerical approaches—anisotropy and stiffness.**

    - **Anisotropy produces subtle balances of large forces, and vastly different parallel and perpendicular transport properties.**

    - **Stiffness reflects the vast range of time-scales in the system: targeted physics is slow (~transport scale) compared to waves**

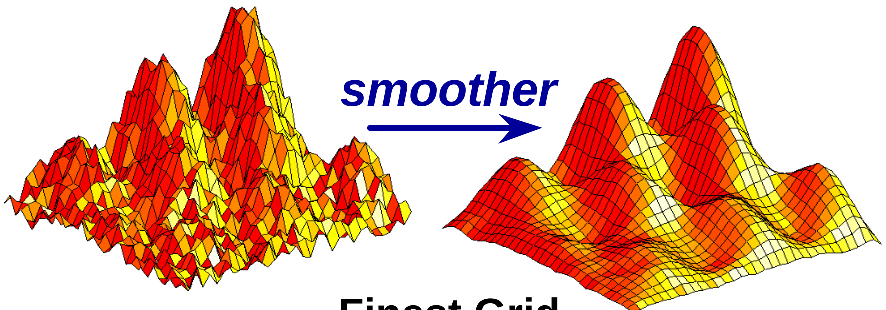
Solver interoperability accomplishments

- **Hypre in PETSc**
  - **codes with PETSc interface (like CEMM's M3D) can invoke Hypre routines as solvers or preconditioners with command-line switch**



- **SuperLU_DIST in PETSc**
  - **as above, with SuperLU_DIST**
- **Hypre in AMR Chombo code**
  - **so far, Hypre is level-solver only; its AMG will ultimately be useful as a bottom-solver, since it can be coarsened indefinitely without attention to loss of nested geometric structure; also FAC is being developed for AMR uses, like Chombo**

Multilevel preconditioning

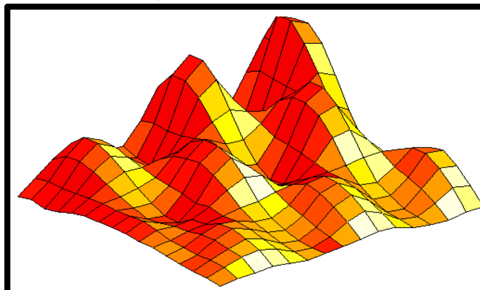


Finest Grid

A Multigrid V-cycle

Restriction
transfer from fine to coarse grid

coarser grid has fewer cells (less work & storage)



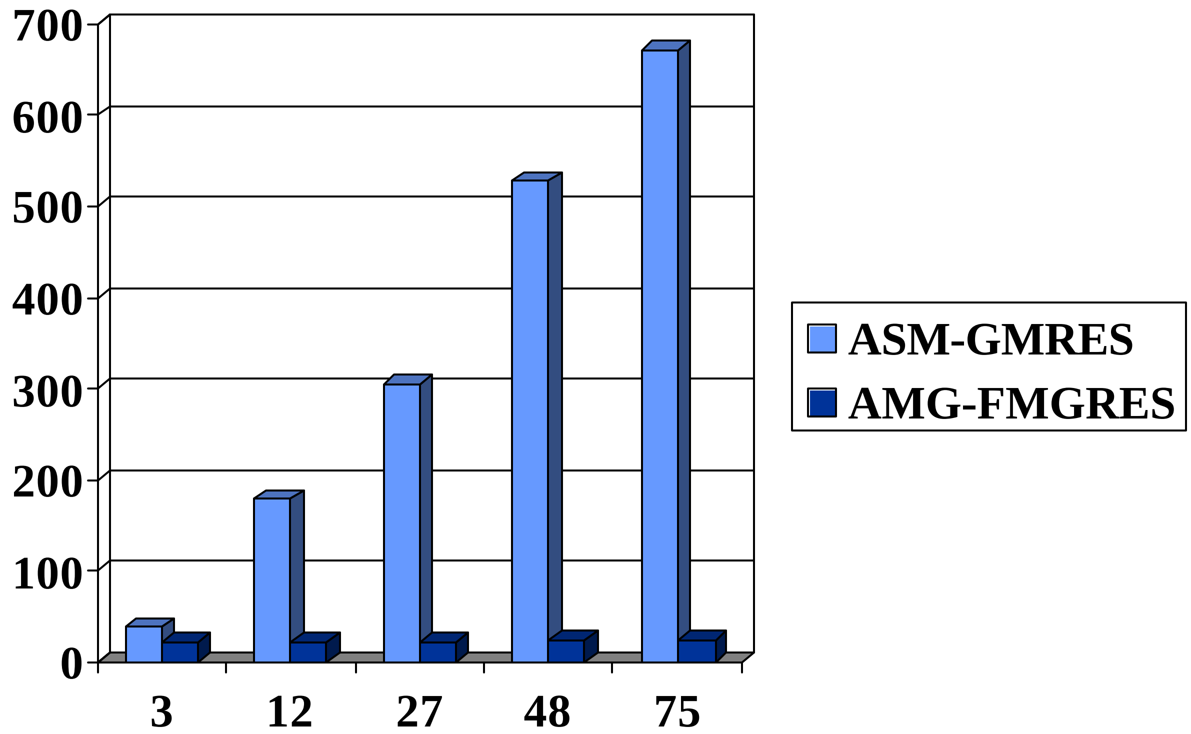
First Coarse Grid

Prolongation
transfer from coarse to fine grid

Recursively apply this idea until we have an easy problem to solve

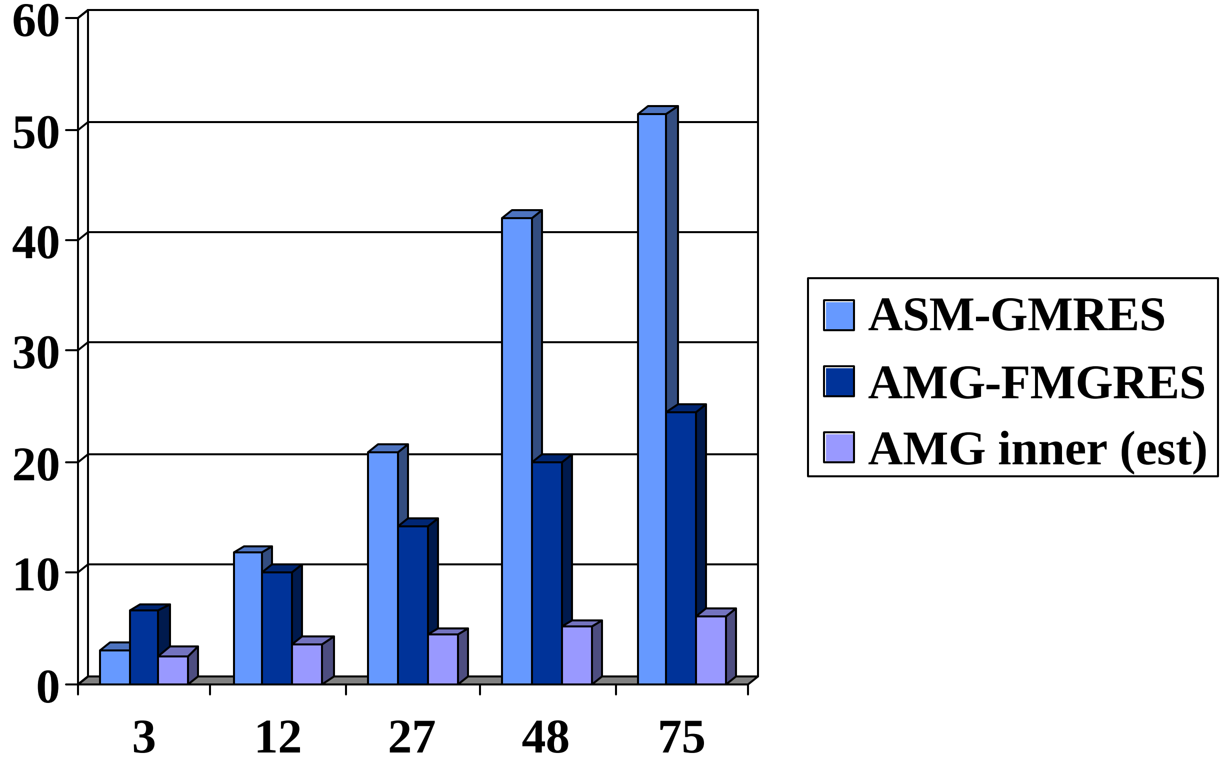
Hypre's AMG in SciDAC app

- PETSc-based PPPL code M3D has been retrofit with Hypre's algebraic MG solver of Ruge-Steuben type
- Iteration count results below are averaged over 19 different PETSc SLESSolve calls in initialization and one timestep loop for this operator split unsteady code, abcissa is number of procs in scaled problem; problem size ranges from 12K to 303K unknowns (approx 4K per processor)



Hypre's AMG in SciDAC app, cont.

- Scaled speedup timing results below are summed over 19 different PETSc SLESSolve calls in initialization and one timestep loop for this operator split unsteady code
- Majority of AMG cost is coarse-grid formation (preprocessing) which does not scale as well as the inner loop V-cycle phase; in production, these coarse hierarchies will be saved for reuse (same linear systems are called in each timestep loop), making AMG much less expensive and more scalable



Background of Hypre Library

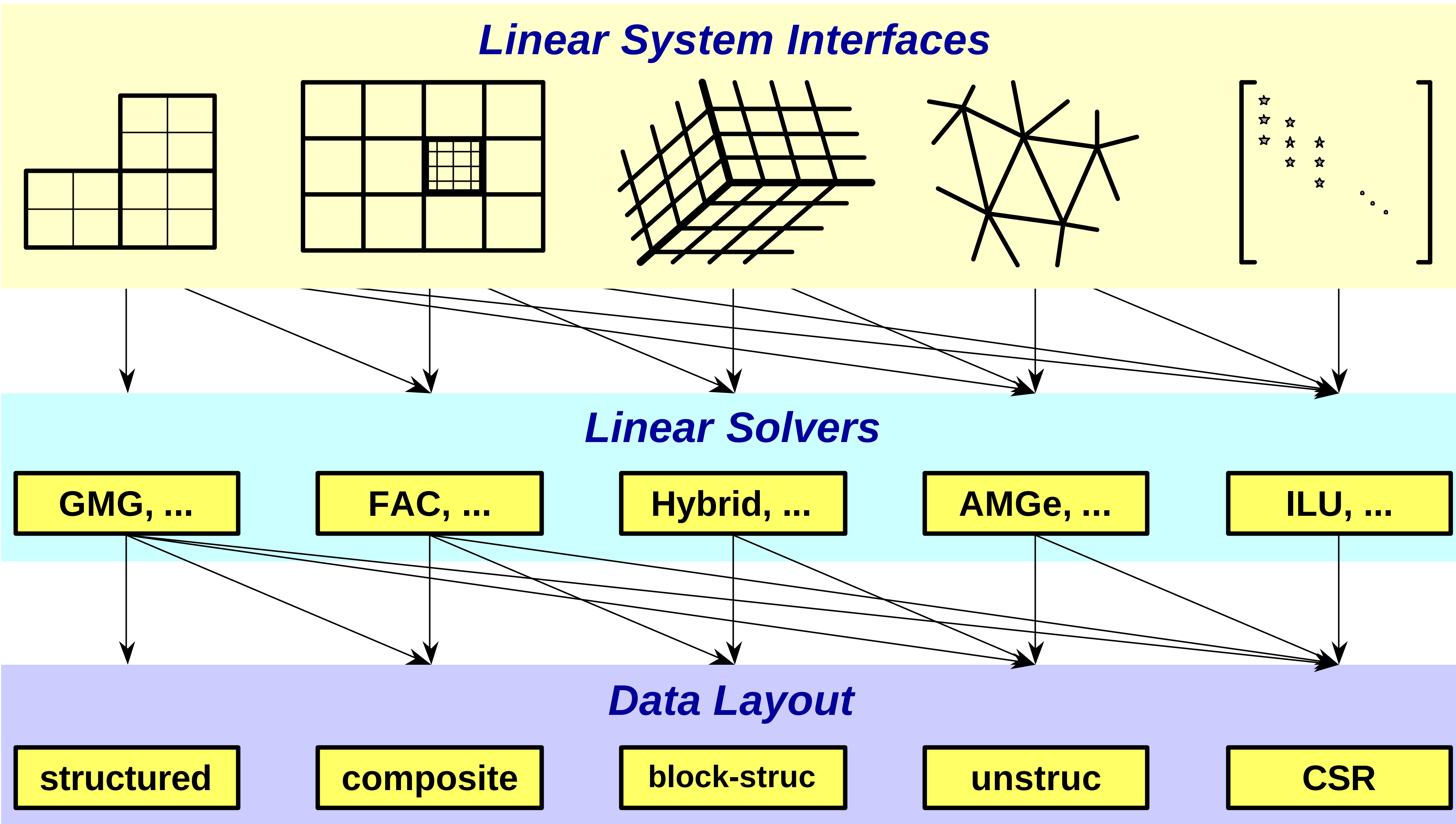
(combined with PETSc under SciDAC)

- **Developed by Chow, Cleary & Falgout (LLNL) to support research, prototyping, and production parallel solutions of operator equations in message-passing environments; now joined by seven additional staff (Henson, Jones, Lambert, Painter, Tong, Treadway, Yang) under ASCI and SciDAC**
- **Object-oriented design similar to PETSc**
- **Concentrates on linear problems only**
- **Richer in preconditioners than PETSc, with focus on algebraic multigrid**
- **Includes other preconditioners, including sparse approximate inverse (Parasails) and parallel ILU (Euclid)**

hypre
high performance preconditioners

See **http://www.llnl.gov/CASC/hypre/**

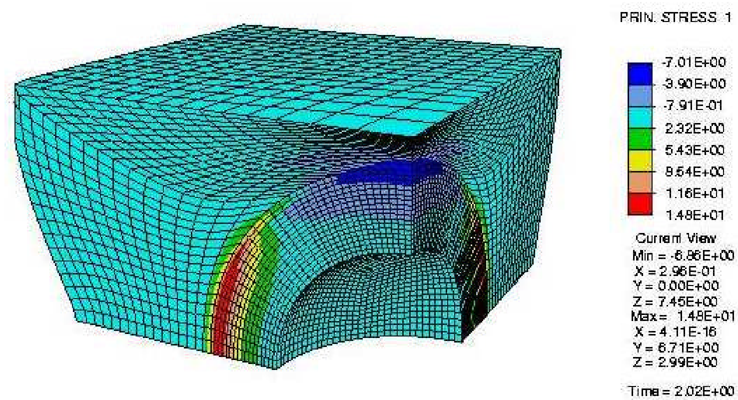
Hypre's "Conceptual Interfaces"



Nonlinear (plastic) deformation

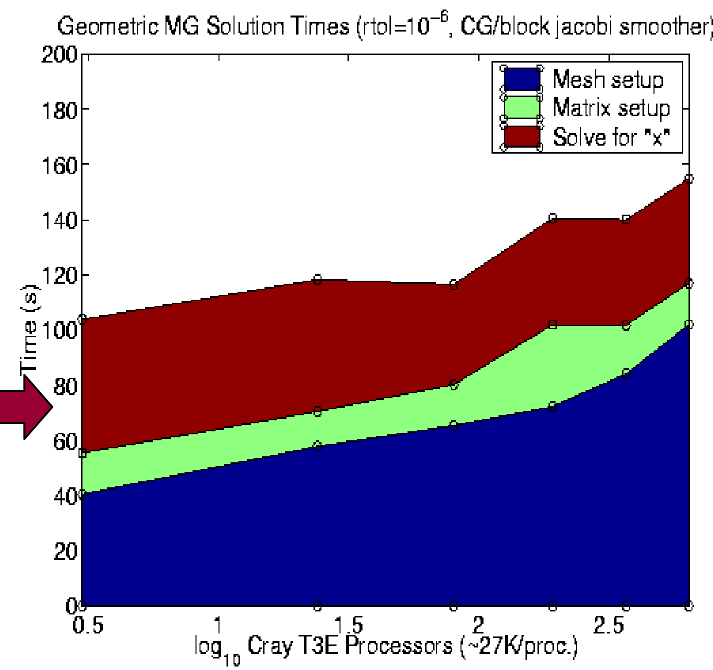
- **PETSc was used in "Prometheus", a multilevel nonlinear mechanics code to scalably compute plastic deformation of a stiff steel-rubber system**
- **Solver scalability actually improves with better resolution**

c/o M. Adams, Berkeley-Sandia



Prometheus steel/rubber composite

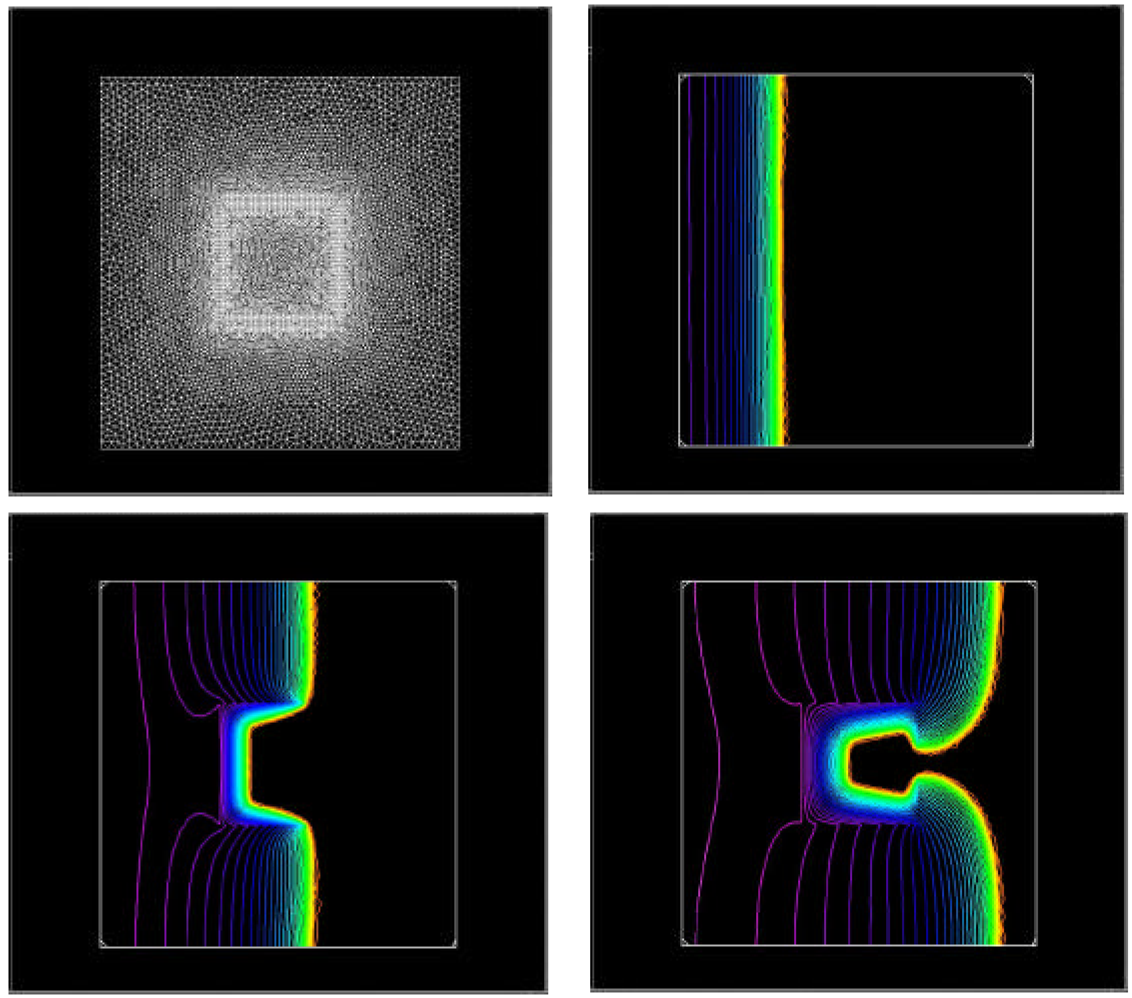
The solver is a key part, but not the only part, of simulation that needs to be scalable



Radiation transport

PETSc was used to simulate "flux-limited diffusion" transport of radiative energy in inhomogeneous materials, as shown in 2D cross-section, with a high-Z material surrounded by low-Z material

Scaling was similar to aerodynamic application in 3D

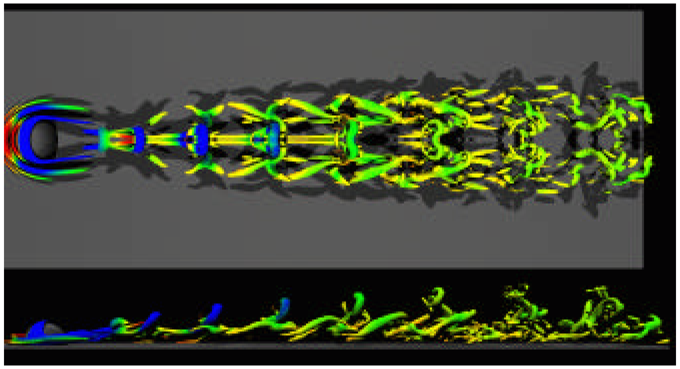


Joint work between ODU, ICASE, and LLNL

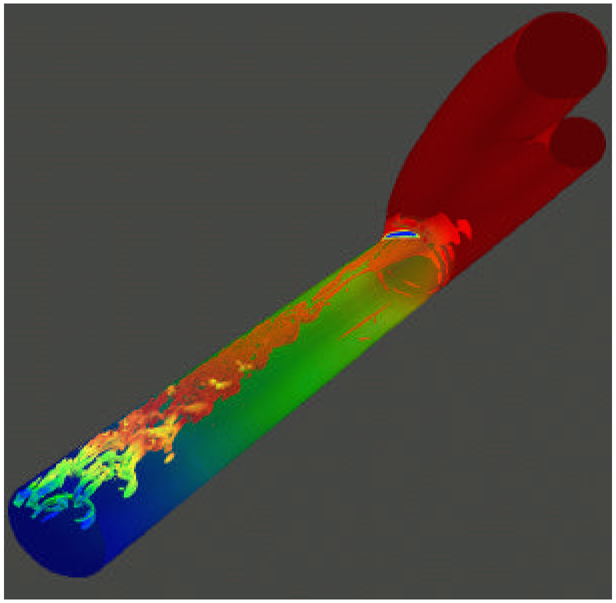
Incompressible flow

Nek5000: unstructured spectral element code (not PETSc, but …)

- **unsteady Navier-Stokes solver**
  - **high-order tensor-product polynomials in space (N~5-15)**
  - **high-order operator splitting in time**
- **two-level overlapping additive Schwarz for pressure**
  - **spectral elements taken as subdomains**
  - **fast local solves (tensor-product diagonalization)**
  - **fast coarse-grid solver**
    - **sparse exact factorization yields $\underline{x}_0 = A_0^{-1}\underline{b} = XX^T\underline{b}$ as parallel matrix-vector product**
    - **low communication; scales to 10,000 processors**



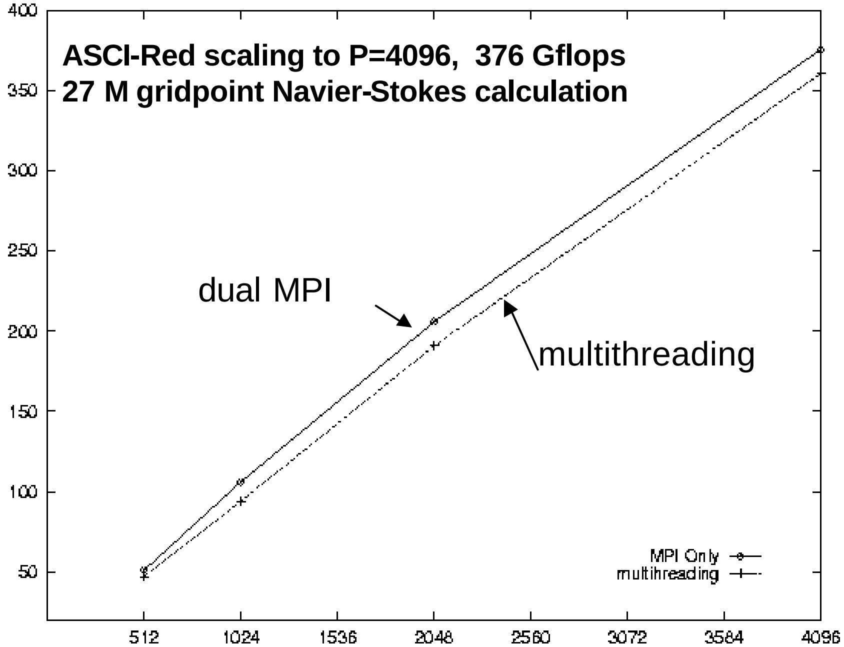
Transition near roughness element



ansition in arterio-venous graft, Re_v=2700

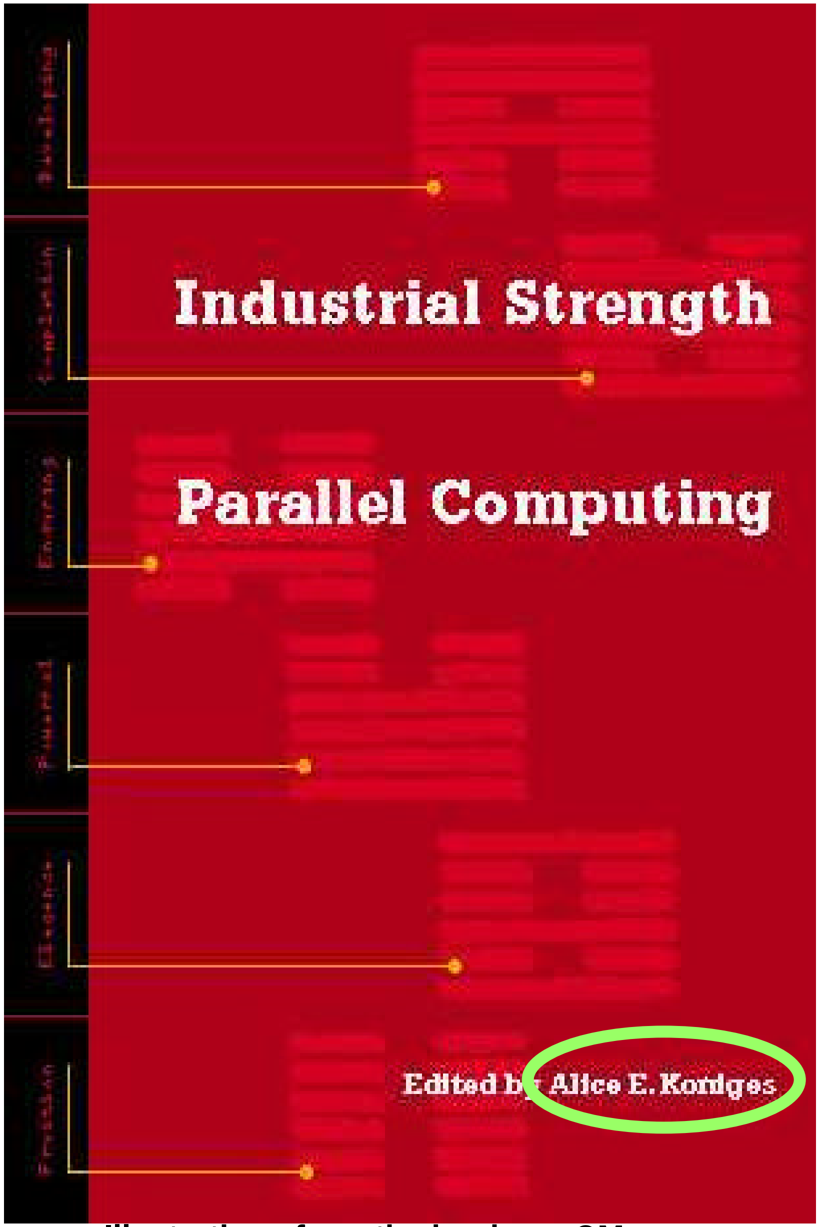
c/o P. Fischer and H. Tufo

www.mcs.anl.gov/appliedmath/Flow/cfd.html



Book of parallel computing case studies

- **Introductory material overviews the typical MPP environment for real applications and the special challenges encountered, such as scheduling and performance optimization issues**
- **Applications contributions (17)**
  - **Industrial/scientific considerations**
  - **Computational considerations**
  - **Computational results**
- **Future of industrial parallel supercomputing**

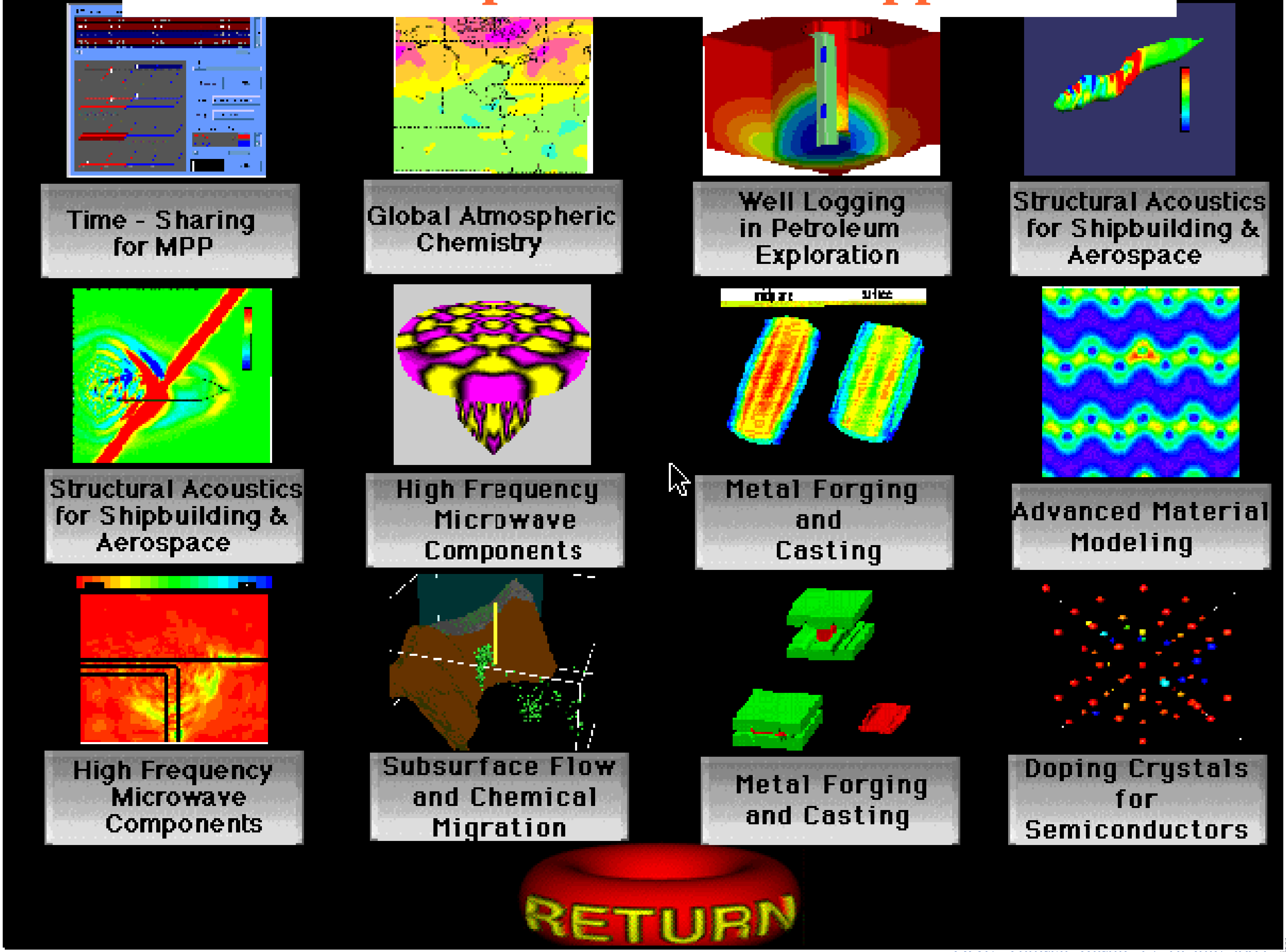


Morgan Kaufmann Publishers

Distinguished professional and educational publications in computer science, engineering, and information technology.

ISBN1-55860-540-1

Pictorial Sample of Parallel Applications



EOF