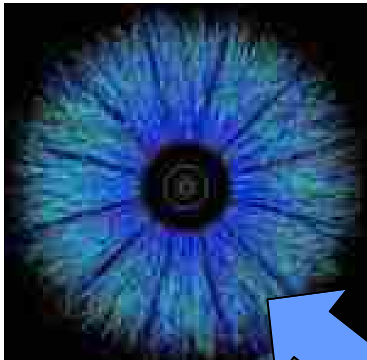
TOPS

Terascale Optimal PDE Simulations

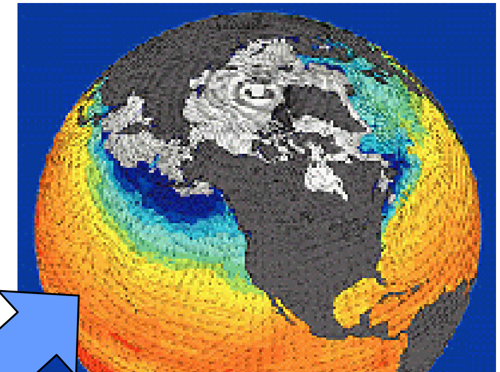
http://www.tops-scidac.org

"The partial differential equation entered theoretical physics as a handmaid, but has gradually become mistress." – **A. Einstein**



4 projects in high energy and nuclear physics

PDEs

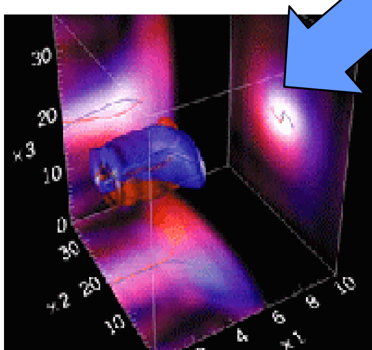


14 projects i
biological a
environmen
research

are dense

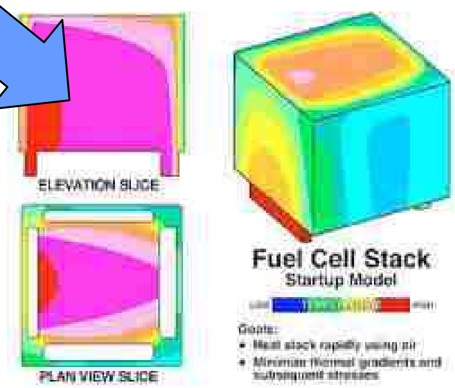
TOPS
Terascale Optimal PDE Simulations
& ISICs

in the SciDAC



5 projects in fusion energy science

portfolio



10 projects in basic energy sciences

Who we are in TOPS…

ARGONNE NATIONAL LABORATORY

Argonne National Laboratory is operated by The University of Chicago for the U.S. Department of Energy

… the **PETSc** and **TAO** people

Lawrence Livermore National Laboratory

… the **Hypre** and **SUNDIALS** people

BERKELEY LAB

Berkeley Lab

… the **SuperLU** and **PARPACK** people

Plus some university collaborators …

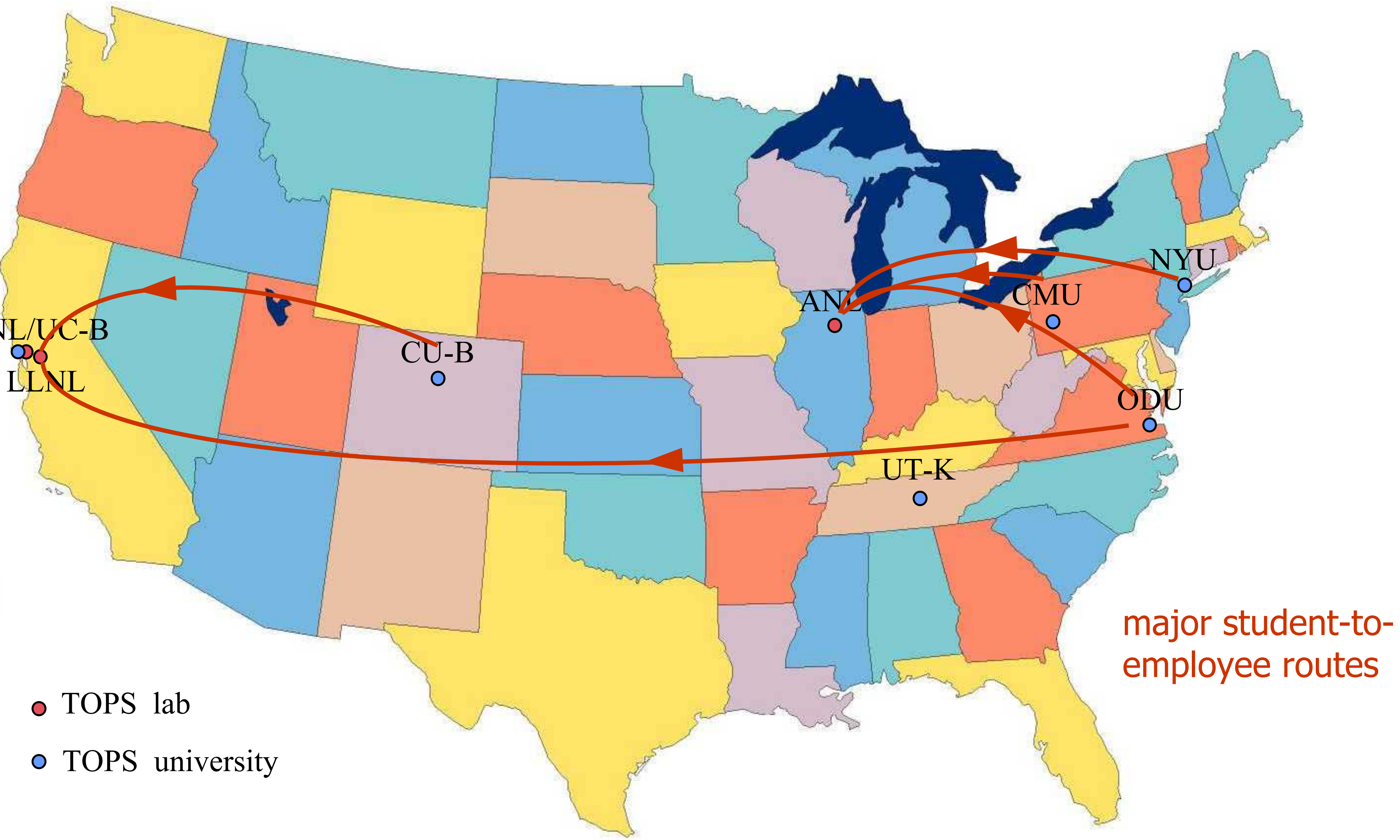
Berkeley
University of California

Carnegie Mellon

ICL UT

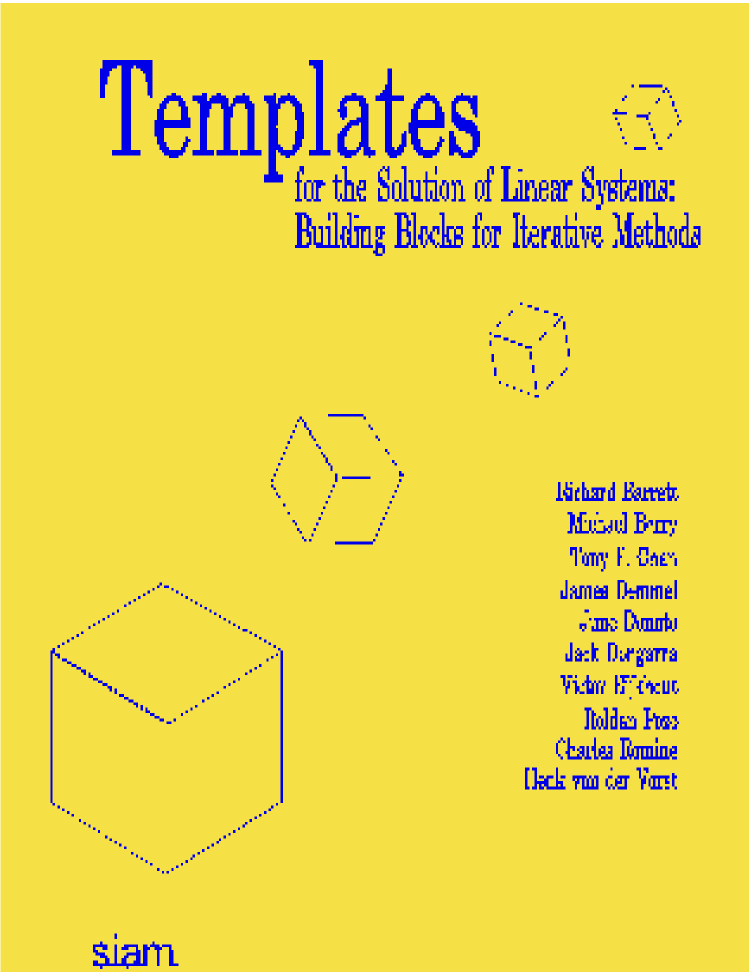
… with a history of lab collaborations in high performance computing

TOPS participants

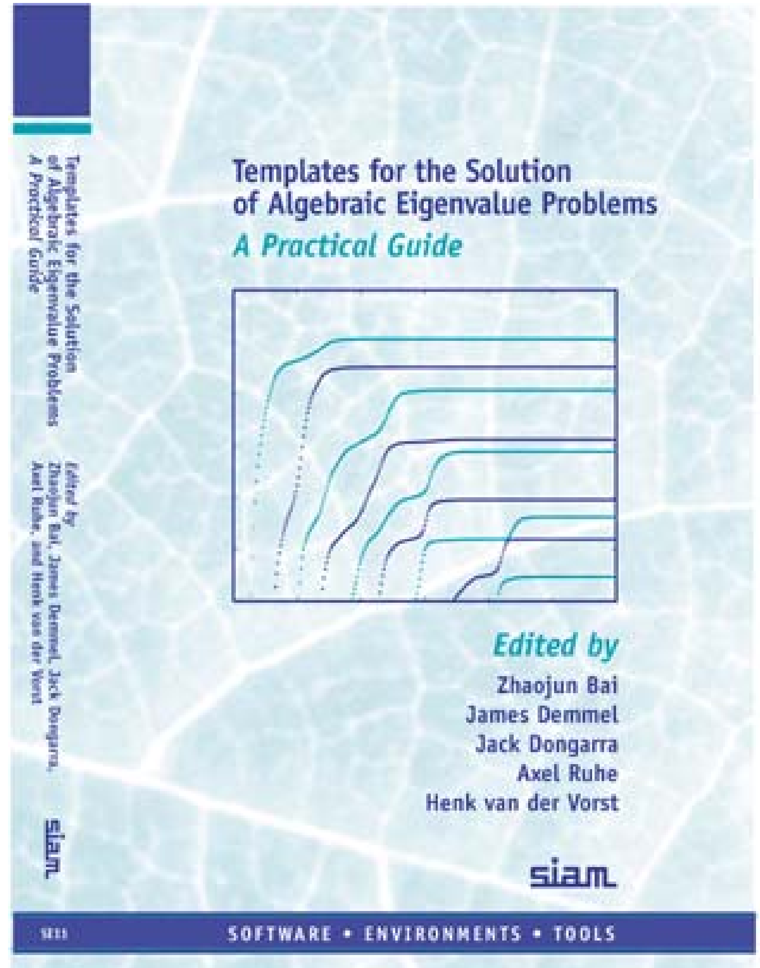


You may know the on-line “Templates” guides …

www.netlib.org/templates | **www.netlib.org/etemplates**



124 pp.



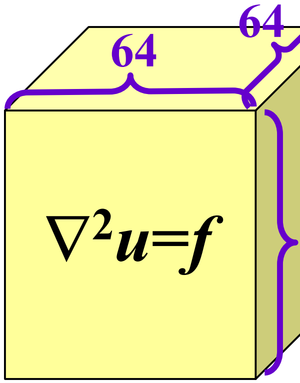
410 pp.

… these are good starts, but not adequate for SciDAC scales!

The power of optimal algorithms

- **Advances in algorithmic efficiency can rival advances in hardware architecture**
- **Consider Poisson's equation on a cube of size $N=n^3$**

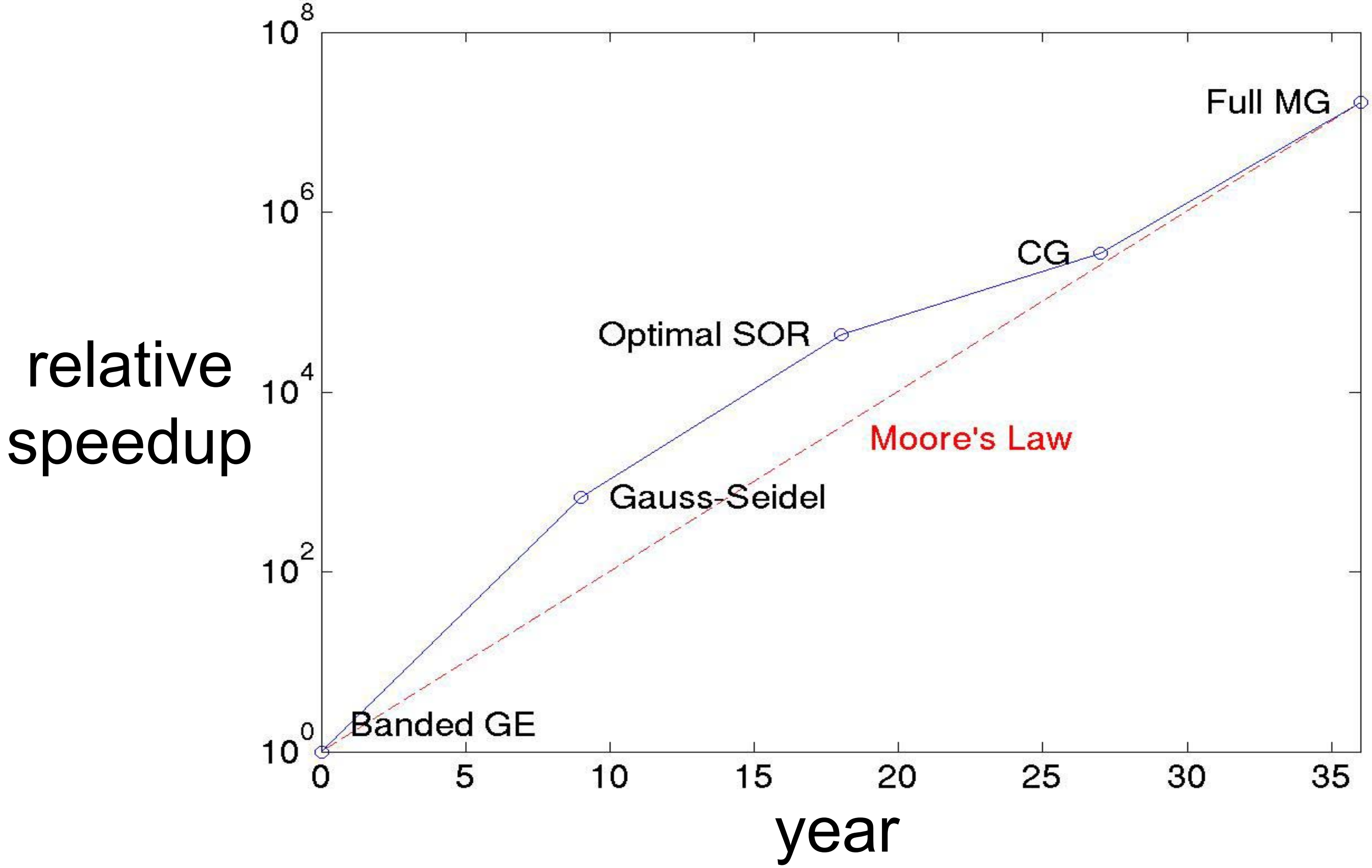
Year	*Method*	*Reference*	*Storage*	*Flops*
1947	GE (banded)	Von Neumann & Goldstine	n^5	n^7
1950	Optimal SOR	Young	n^3	$n^4 \log n$
1971	CG	Reid	n^3	$n^{3.5} \log n$
1984	Full MG	Brandt	n^3	n^3



- **If n=64, this implies an overall reduction in flops of ~16 million** *

Algorithms and Moore's Law

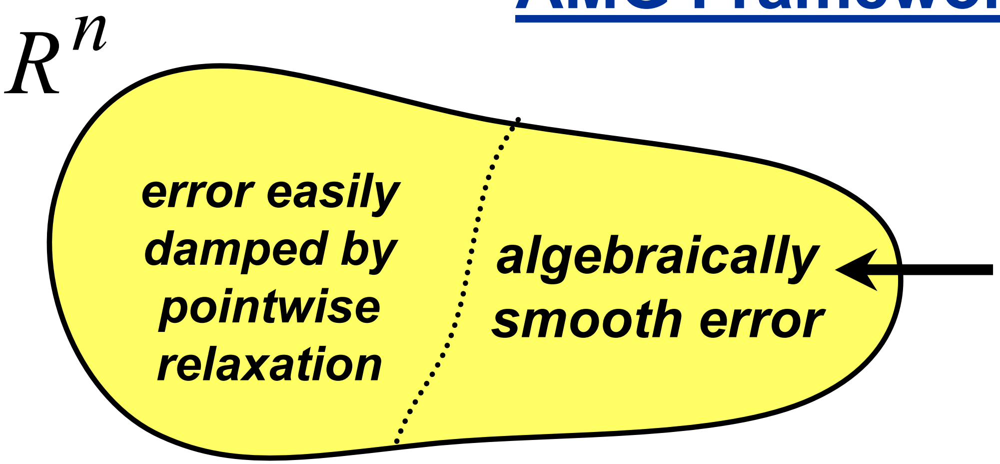
- **This advance took place over a span of about 36 years, or 24 doubling times for Moore's Law**
- **$2^{24} \approx 16$ million $\Rightarrow$ the same as the factor from algorithms alone!**



But where to go past *O(N)* ?

- Since *O(N)* is already optimal, there is nowhere further "upward" to go in efficiency, but one must extend optimality "outward," to more general problems
- Hence, for instance, algebraic multigrid (AMG) to seek to obtain *O(N)* in *indefinite, anisotropic,* or *inhomogeneous* problems on *irregular* grids

AMG Framework



Choose coarse grids, transfer operators, and smoothers to eliminate these "bad" components within a smaller dimensional space, and recur

TOPS is dedicated to the proposition that …

- **Not all problems are created equal**

  so a large variety of solvers should be callable from one interface

- **Solver software can rarely be thrown over the wall**

  so we are committed to collaborations with applied users

- **Discretization and solution rarely separate cleanly**

  so we are committed to collaborations with ISIC colleagues

- **Desire for resolution will grow without bound**

  so we concentrate on solvers that scale well (in the "weak" sense)

- **Solving the PDE well is only a beginning, not the end, in doing computational science**

  so we are providing a software "tool chain" of several links, which are implemented over common data structures and kernel functionality

TOPS has a dream that users will…

- **Understand range of algorithmic options w/tradeoffs**

  ***e.g.*, memory *vs.* time, comp. *vs.* comm., inner iteration work *vs.* outer**

- **Try all reasonable options "easily"**

  without recoding or extensive recompilation

- **Know how their solvers are performing**

  with access to detailed profiling information

- **Intelligently drive solver research**

  ***e.g.*, publish joint papers with algorithm researchers**

- **Simulate *truly new physics* free from solver limits**

  ***e.g.*, finer meshes, complex coupling, full nonlinearity**

A project like TOPS is needed because …

- **SciDAC applications are presently solver-bound**
  e.g., 90-95% of execution time in solver, limited to 1 or 2 dimensions
- **SciDAC ambitions are too low, focused too near**
  concentrated on getting a few big runs, without enough validation and verification, since iteration over the "forward" problem is costly
- **SciDAC community codes are hard to keep current**
  slow process to implement new algorithms, to port to new machines
- **SciDAC CS ISICS need good stepping stone to apps**
  solvers are good target for research in components and performance
- **SciDAC Math ISICs need good solvers, too**
  from scalable Poisson solves to mesh optimization, other ISICs have subproblems for TOPS, for which they are not otherwise funded

TOPS set out with certifiably good ingredients

- **Constituent software powers commercial toolkits**

  ***e.g.*, SUNDIALS (*Mathematica*), SuperLU (*Matlab*), PETSc (numerous)**

- **Constituent software powers major research codes**

  ***e.g.*, Hypre (ASCI), PETSc (NASA HPC, Harvard Medical)**

- **Constituent software has powered prizes**

  ***e.g.*, PETSc (Bell Prize), Veltisto ("Best Paper" at SC)**

- **… and science on covers of *Science* and *Nature***

  ***e.g.*, SuperLU, ScaLAPACK**

- **TOPS ingredients are continually being improved, in conjunction with thousands of computational scientists and engineers *around the world***

What value is being added by TOPS today?

- **Interoperability for new performance**

  e.g., Hypre preconditioners in PETSc

- **Interoperability for new functionality**

  e.g., PETSc in TAO and Veltisto for large-scale optimization

- **Interoperability for new CS research**

  e.g., componentization of PETSc and Hypre

- **Development and maintenance of core codes**
- **Expansion of user consulting capability**
- **Education and training of next generation of solver developers**
- **Outreach to applications community**

Outline for presentation

- Introduction & motivation (just completed ☺)
- TOPS scientific overview (broad and shallow) (D. Keyes)
  - Algorithmic research and development (5 areas)
  - Infrastructural research and development (2 areas)
  - Applications collaborations (3 major groups; 7 others)
  - Outreach (software, publications, presentations, service as "POC")
- Featured efforts (drilling down in a few spots)
  - Supporting SciDAC's Accelerator Science & Technology (E. Ng)
  - Componentizing optimization (J. Moré)
  - Bringing multilevel methods to the masses (R. Falgout)
  - Unifying solver frameworks (B. Smith)
- Wrap-up: supporting details, TOPS philosophy, and summary

Scope for TOPS

- **Design and implementation of "solvers"**
  - **Linear solvers**

    $$Ax = b$$

  - **Eigensolvers**

    $$Ax = \lambda Bx$$

  - **Nonlinear solvers** (w/ sens. anal.)

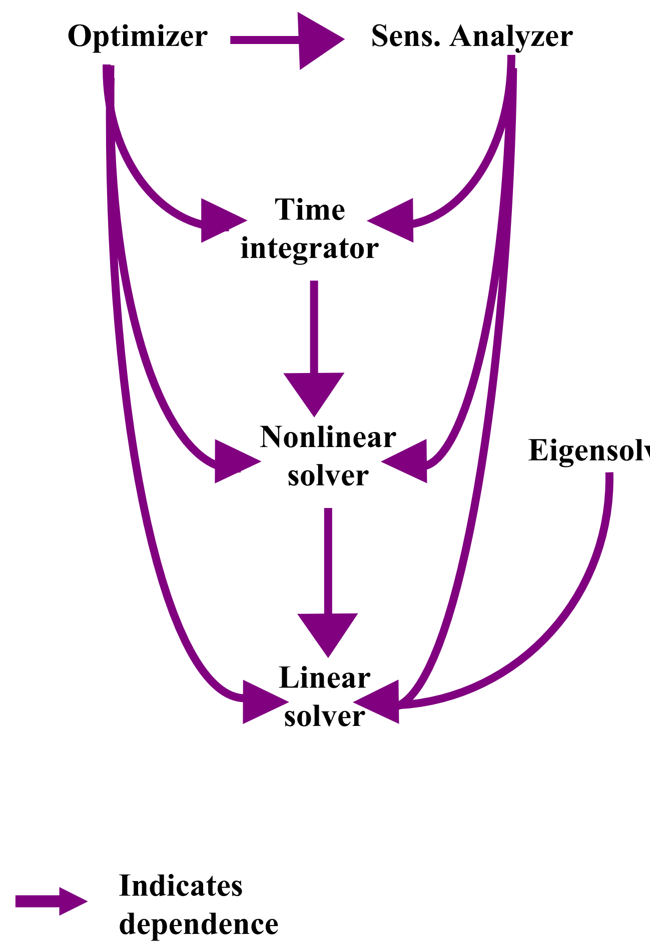
    $$F(x, p) = 0$$

  - **Time integrators** (w/ sens. anal.)

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

  - **Optimizers**

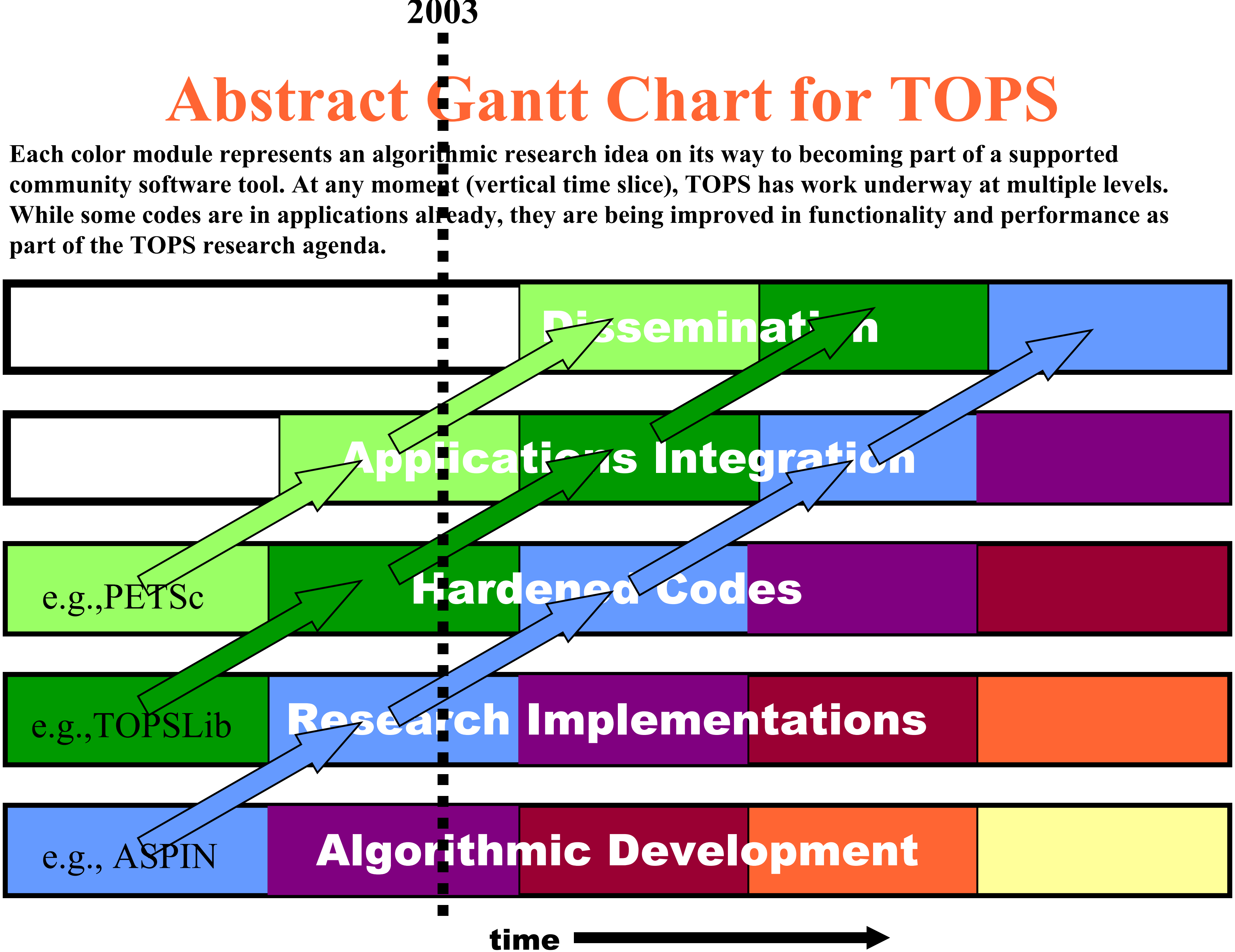
    $$\min_u \phi(x,u) \; s.t. \; F(x,u) = 0, u \geq 0$$



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

Abstract Gantt Chart for TOPS

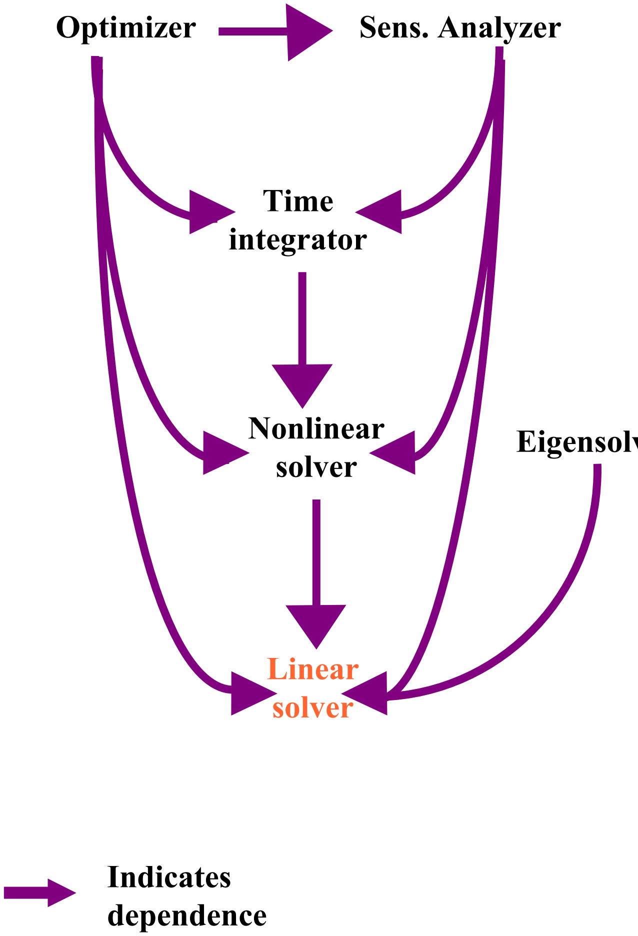
Each color module represents an algorithmic research idea on its way to becoming part of a supported community software tool. At any moment (vertical time slice), TOPS has work underway at multiple levels. While some codes are in applications already, they are being improved in functionality and performance as part of the TOPS research agenda.



Linear solvers

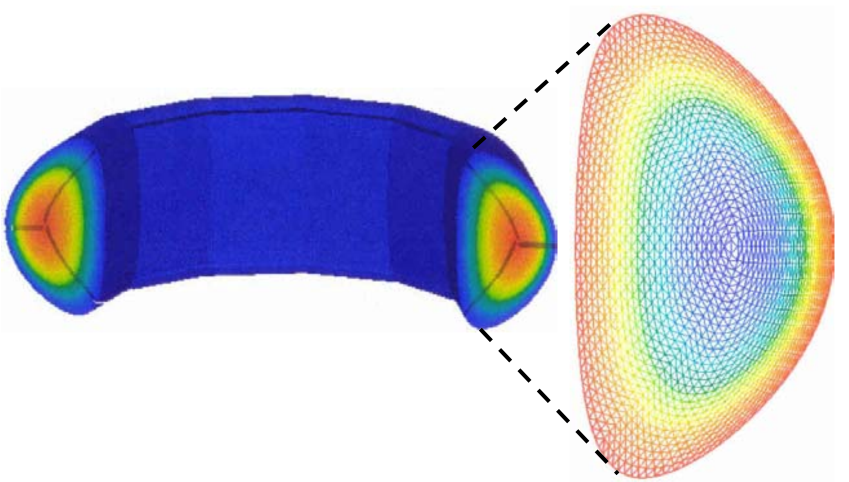
- One of the greatest investments in the history of DOE numerical computing and the largest and core part of TOPS
- TOPS features the workhorse combo of Krylov preconditioned with algebraic multigrid, geometric multigrid, various incomplete factorizations, as well as direct methods – all sparse oriented
- Also research on innovative methods like adaptive AMG, FOSLS, hierarchical ILU, and adaptive multi-method solvers
- Extensive research on scalability features and memory-adaptive versions of direct methods

$$Ax = b$$



Linear solvers progress

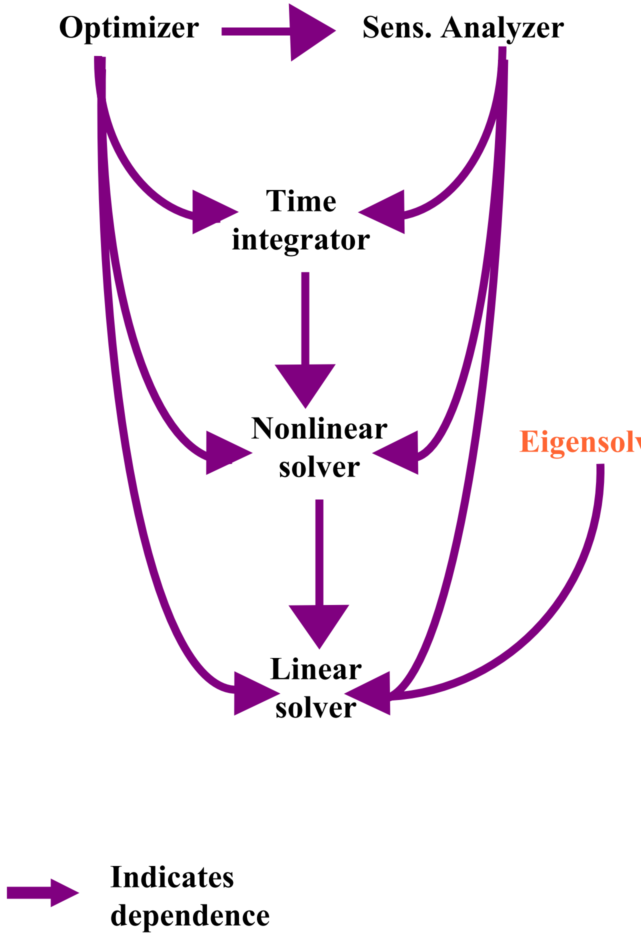
- Algebraic multigrid is dependent upon heuristics to make up for geometric information and to extend optimal convergence from the elliptic regime (where geometric and algebraic smoothness are the same) to more general problems
- When applying AMG anew, must occasionally extend the set of heuristics, sometimes using information beyond matrix alone; self-adaptive AMG a new holy grail (Falgout to discuss)
- On software side, also extending the set of interfaces to get closer to user data structures
- MG needs coarse solves, which is one reason for on-going direct methods research
- Sometimes no alternative to direct methods, including in shift-invert



Eigensolvers

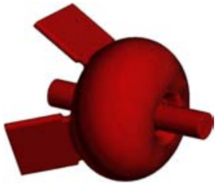
- Preferred eigenanalysis algorithm depends upon: system structure, computational resources, and portion of spectrum and invariant subspaces desired
- Based on customer, TOPS currently concentrates on sparse, symmetric, and small subrange of high-dimensional spectrum
- Exact-shift-invert Lanczos and Jacobi-Davidson both important, preference depending in part on memory available
- Innovative research also in multilevel eigensolvers and in sparse QR

$$Ax = \lambda Bx$$

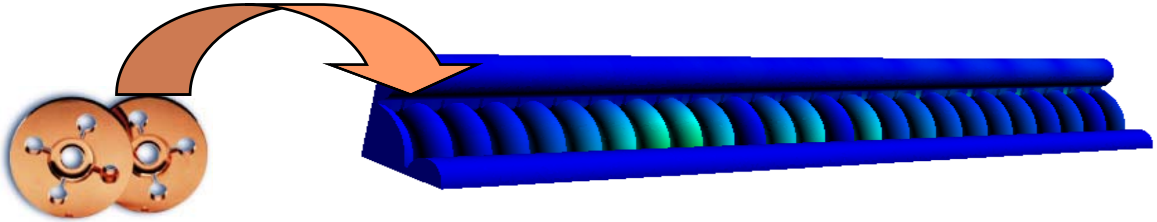


Eigensolvers progress

- AST's **Omega3P** is using TOPS software to find EM modes of accelerator cavities, currently lossless (lossy to come)



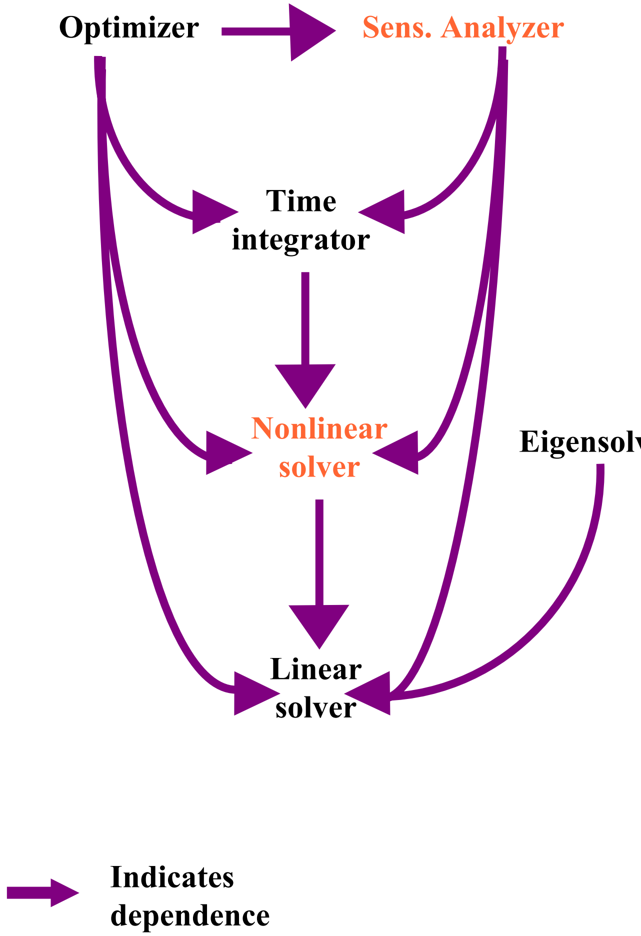
- Methods: Exact Shift-and-Invert Lanczos (ESIL), combining **PARPACK** with **SuperLU** when there is sufficient memory, and Jacobi-Davidson otherwise
- Current high-water marks (Ng to discuss):
  - 47-cell chamber, finite element discr. of Maxwell's eqs.
  - System dimension 1.3 million
  - 20 million nonzeros in system, 350 million in LU factors
  - *halved analysis time* on 48 processors, scalable to many hundreds



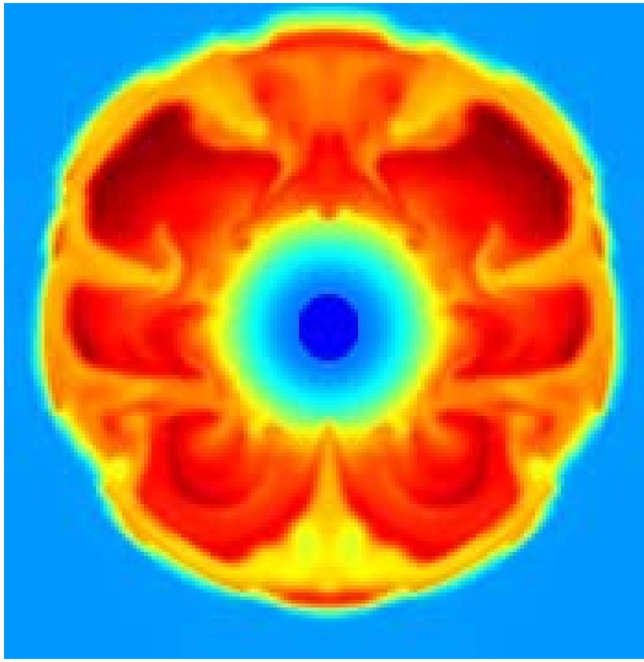
Nonlinear solvers

- TOPS features two "workhorse" methods, Newton-Krylov-Schwarz and Newton-Krylov-multigrid, plus two methods in research stages, nonlinear Schwarz (ASPIN) and nonlinear multigrid (FAS)
- Newton implies the ability to solve linear systems with the Jacobian, which leads instantly to sensitivity and optimization capabilities rarely present in legacy codes
- "Jacobian-free" versions of NKS and NK-MG do not require users to supply Jacobian evaluation routines
- Also researching nonlinear versions of substructuring DD methods

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



Nonlinear solver progress



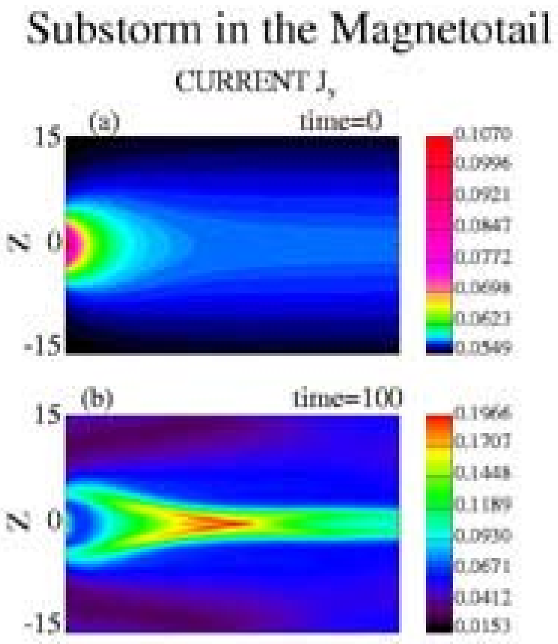
- **Mature algorithmic technology in mature software design, with hooks for user-supplied "physics-based" preconditioning**
- **Newton robustification required**
- **Pseudo-transient continuation, mesh sequencing, and mainstream algebraic techniques (linesearch and trustregion) available in PETSc and SUNDIALS**
- **Difficult "sell" to get users to embrace Newton after lifetime of splitting and linearization**
- **Built demo of a Hall MHD computation directly into PETSc release**

Next: 5-slide interlude on NK-MG

Ex.: nonlinear solvers in Hall MR

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.



PI: Amitava Bhattacharjee
University of Iowa

Status of CMRS collaboration

- CMRS team has provided TOPS with discretization of model 2D multicomponent MHD evolution code in PETSc's DMMG format using automatic differentiation for Jacobian objects
- TOPS has implemented fully nonlinearly implicit GMRES-MG-ILU parallel solver with deflation of nullspace in CMRS's doubly periodic formulation
- CMRS and TOPS reproduce the same dynamics on the same grids with the same time-stepping, up to a finite-time singularity due to collapse of current sheet (that falls below presently uniform mesh resolution)
- TOPS code, being implicit, can choose timesteps an order of magnitude larger, with potential for higher ratio in more physically realistic parameter regimes, but is presently slower in wall-clock time
- Plan: tune PETSc solver by profiling, blocking, reuse, etc.
- Plan: go higher-order in time
- Plan: identify the numerical complexity benefits from implicitness (in suppressing fast timescales) and quantify (explicit versus implicit)
- Plan (with APDEC team): incorporate AMR

2D Hall MHD sawtooth instability

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

Model equations:

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

ex29.c in PETSc 2.5.1

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

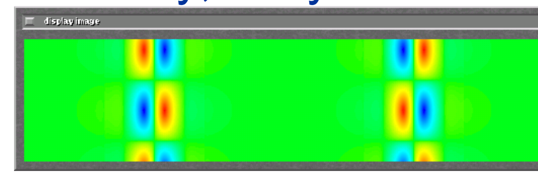
$$
\begin{aligned}
\vec{B} &= B_0 \hat{z} + \nabla \psi \times \hat{z} \\
\vec{v} &= \hat{z} \times \nabla \phi
\end{aligned}
$$

Equilibrium:

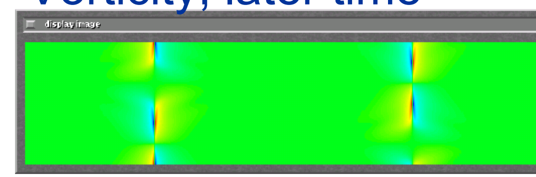
$$\phi_{eq} = U_{eq} = 0$$

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

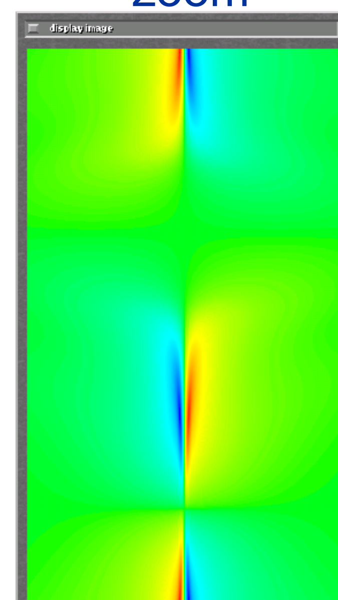
Vorticity, early time



Vorticity, later time

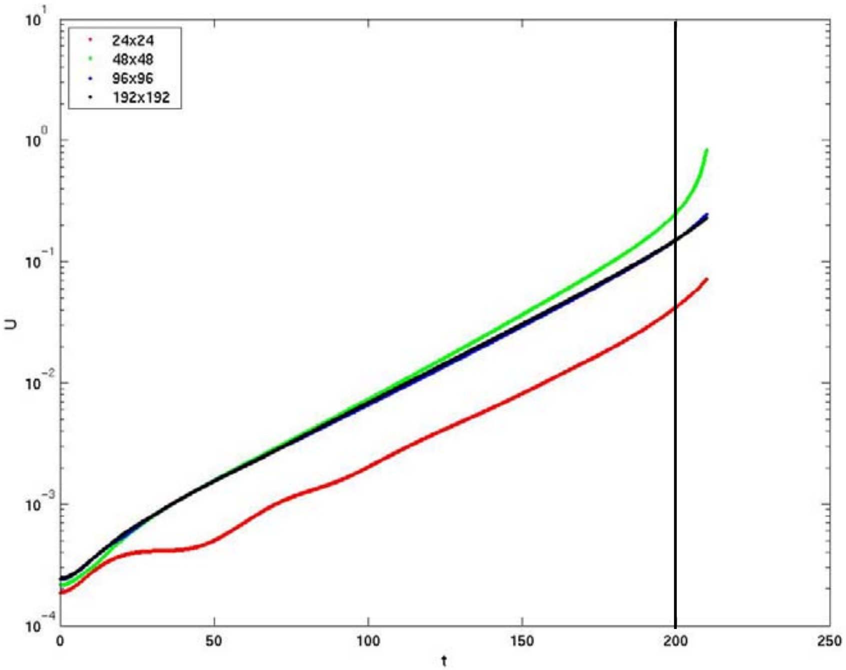


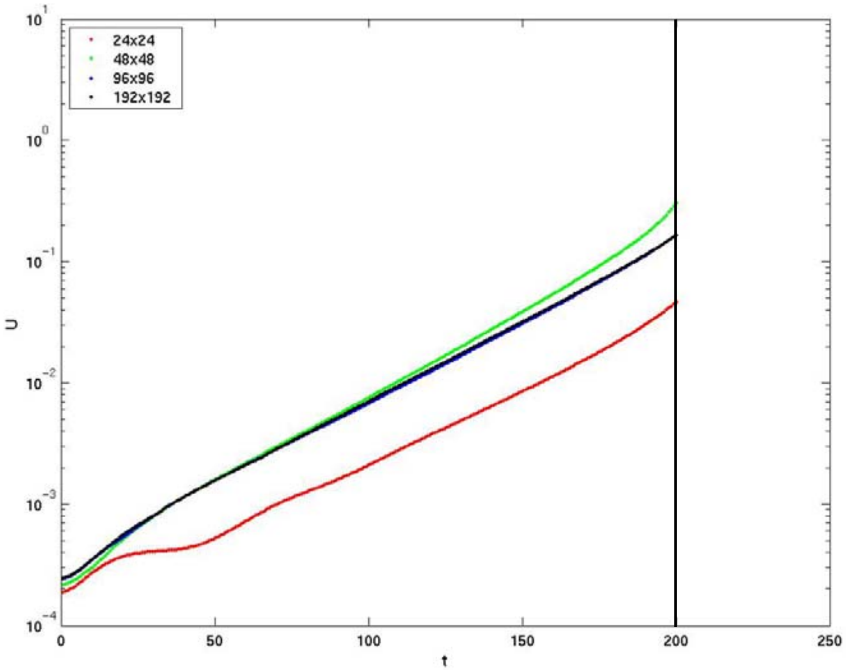
zoom



PETSc's DMMG in Hall MR application

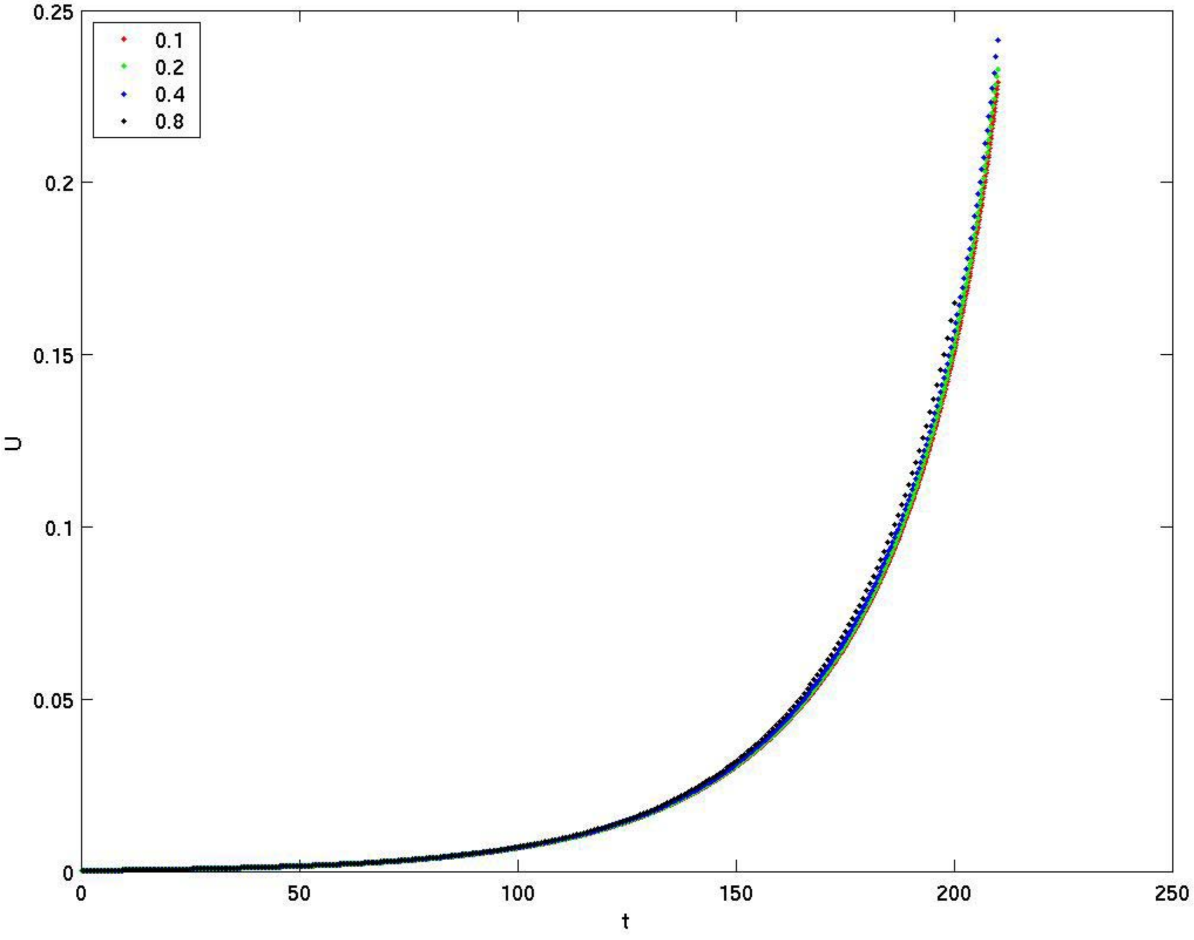
- Mesh and time refinement studies of CMRS Hall magnetic reconnection model problem (4 mesh sizes, dt=0.1 (nondimensional, near CFL limit for fastest wave) on left, dt=0.8 on right)
- Measure of functional inverse to thickness of current sheet versus time, for 0<*t*<200 (nondimensional), where singularity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





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

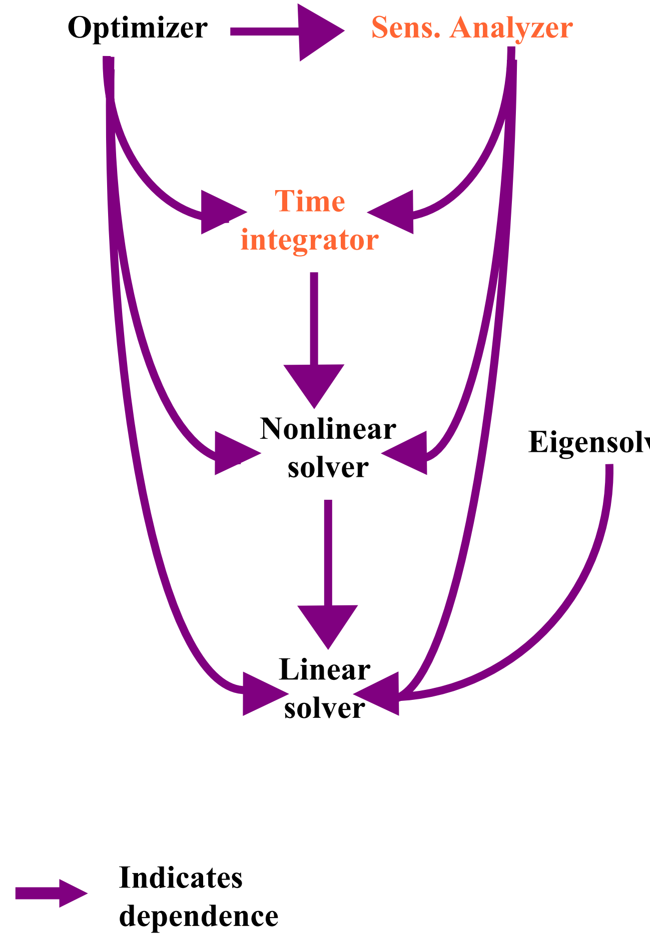
- Implicit timestep increase studies of CMRS Hall magnetic reconnection model problem, on finest (192×192) mesh of previous slide, in absolute magnitude, rather than semi-log



Time integrators w/ sensitivity analysis

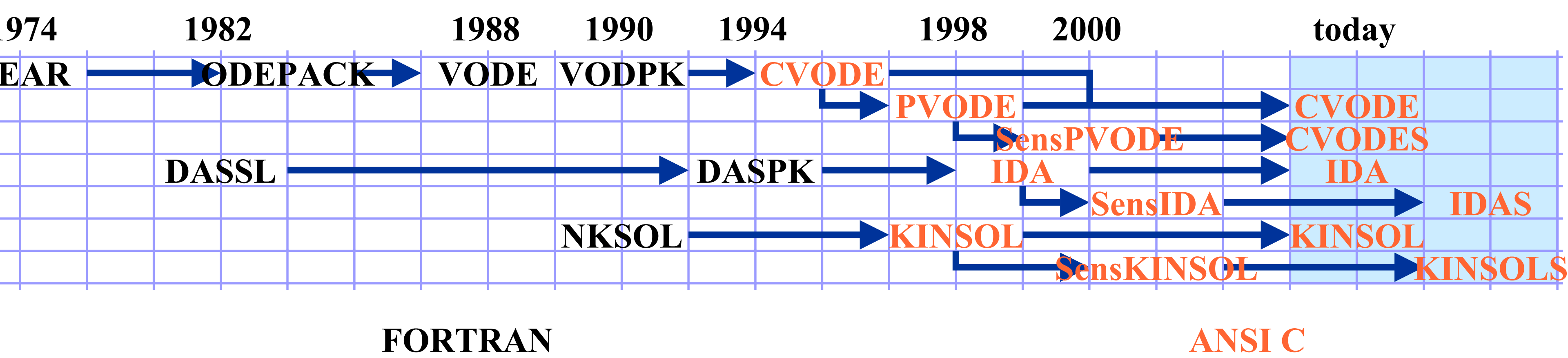
$$f(\dot{x}, x, t, p) = 0$$

- **Transient multirate problems require stiff integrators, a known art, assuming a powerful nonlinear solver capability**
- **SUNDIALS and PETSc both implement the PVODE backward differentiation schemes for temporal discretization**
- **PETSc supplies a variety of distributed data structures**
- **Users who want to use their own data structures, or to utilize built-in sensitivity estimation may prefer SUNDIALS**
- **Especially recommended for parameterized applications, requiring uncertainty quantification**



Integrators progress

- **PVODE, IDA, and KINSOL (an NK solver) now wrapped together in SUNDIALS and augmented with forward and adjoint sensitivity analysis capabilities**
- **Embodies decades of work in variable-order, variable-timestep method-of-lines and Newton-Krylov solvers at LLNL**



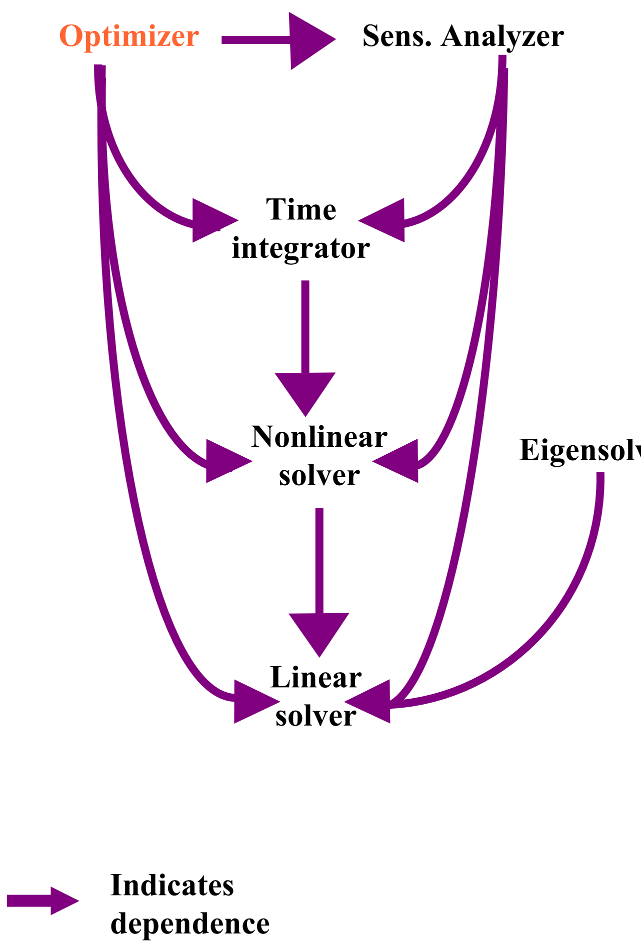
SUNDIALS

SUite of Nonlinear and DIfferential/ALgebraic equation Solvers

Optimizers

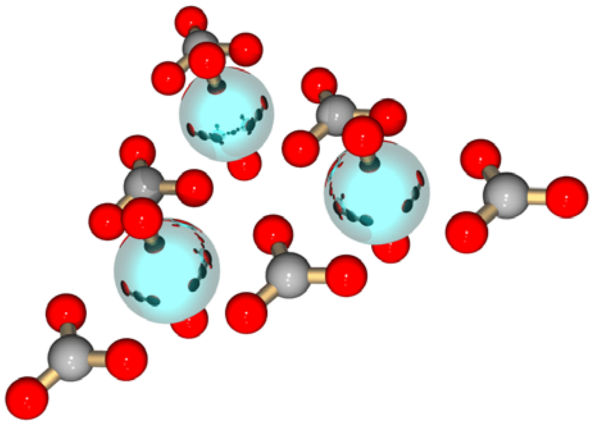
$$\min_{u} \phi(x,u) \; s.t. \; F(x,u) = 0, u \geq$$

- Many SciDAC simulations are properly posed as optimization problems, but this may not always be recognized
- Unconstrained or bound-contrained applications use **TAO**
- PDE-constrained problems use **Veltisto**
- Both are built on **PETSc** solvers (and **Hypre** preconditioners)
- **TAO** makes heavy use of AD, freeing user from much coding
- **Veltisto**, based on RSQP, switches as soon as possible to an "all-at-once" method and minimizes the number of PDE solution "work units"

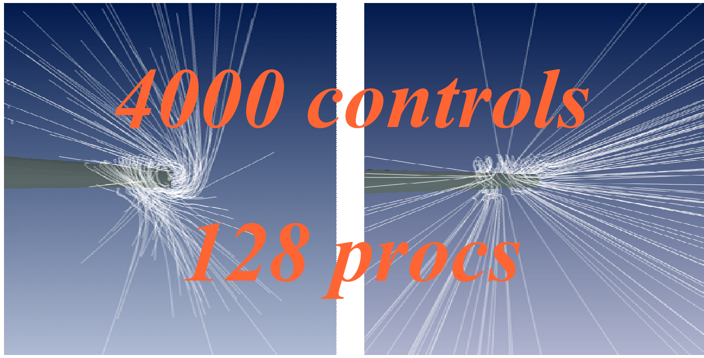


Optimizers progress

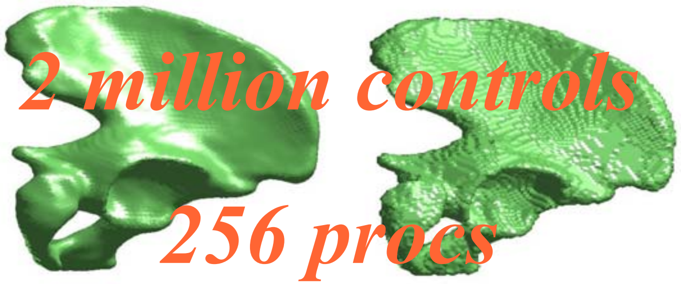
- **Unconstrained or bound-constrained optimization**
  - **TAO (interfaced in CCTTSS component framework) used in quantum chemistry energy minimization (Moré to discuss)**



- **PDE-constrained optimization**
  - **Veltisto used in flow control application, to straighten out wingtip vortex by wing surface blowing and sunction; performs full optimization in the time of just five N-S solves**

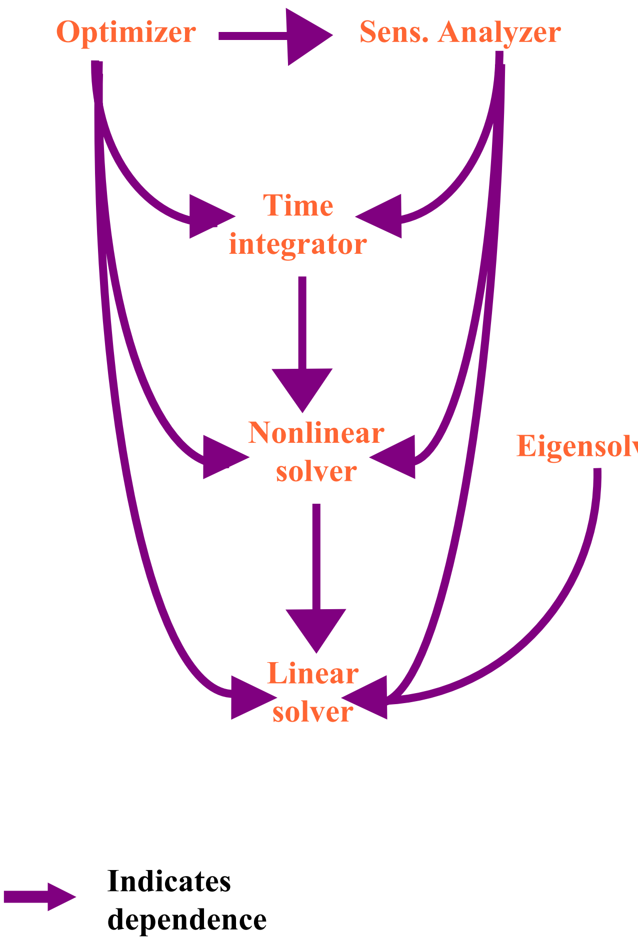


- **"Best technical paper" at SC2002 went to TOPS team**
  - **Inverse wave propagation employed to infer hidden geometry**



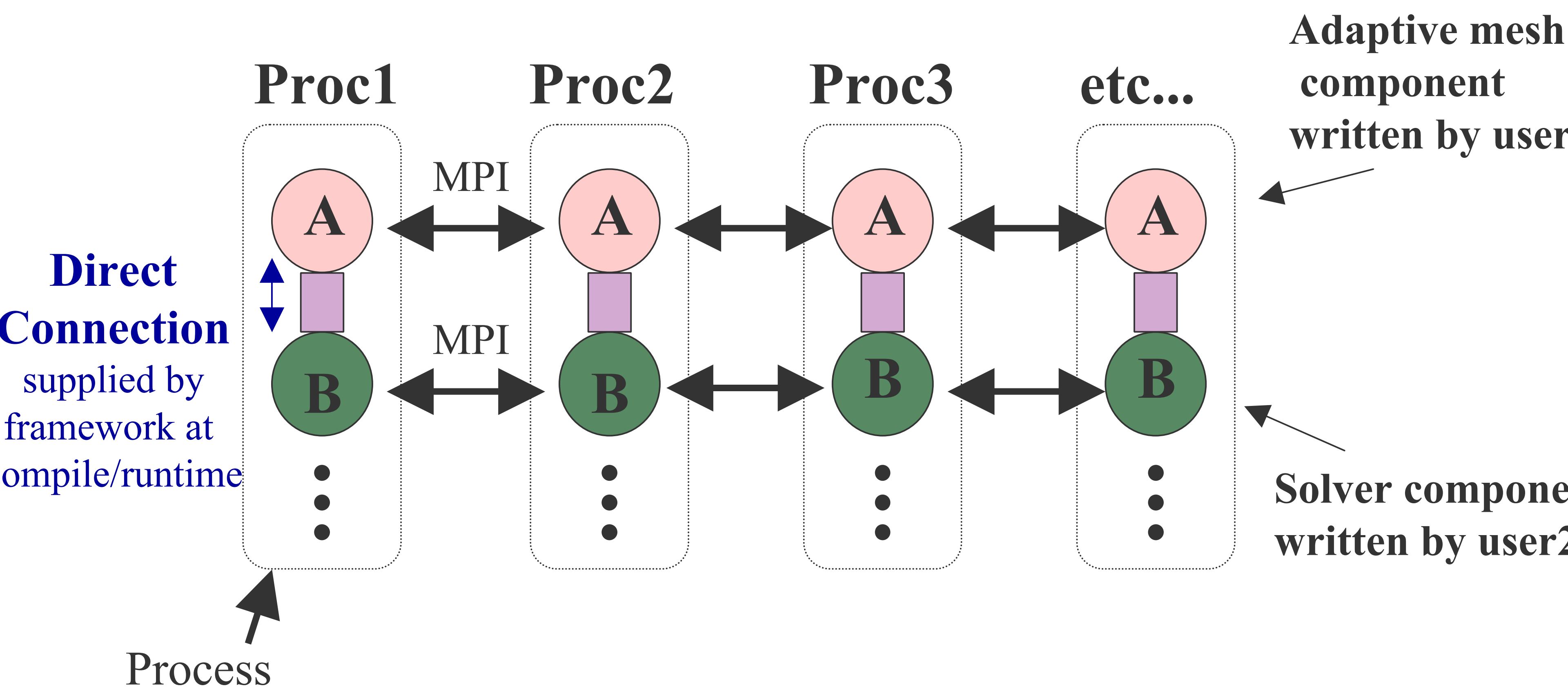
Software integration

- TOPS software achieves integration by supporting multiple interfaces
- Initially, this *N*-to-*N* compatibility is an $O(N^2)$ problem, dealt with case-by-case
- Once software is componentized and respects a standard interface, *N*-to-*N* compatibility reduces to an $O(N)$ problem
- Overhead cost depends upon how deep into inner loops component interfaces occur; experience shows that significant interoperability costs only 1-5% overhead
- Reduces risk to applications developer, since all solvers are available
- Parallel generalization is "SCMD" (single-component, multiple data)



Schematic of SCMD components

MPI application using CCA for interaction between components A and B within the same address space



Software integration progress

- **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 and Parallel_IC in PETSc**
  - invokable as above
- **Hypre in Chombo**
  - so far, **Hypre** is level-solver only; also FAC is being developed for AMR uses, like **Chombo**
- **Hypre and PETSc both being "SIDL'ized"**
  - one of TOPS' three foci of interaction with CCTTSS
- **TAO and PETSc componentized in early demonstration of CCA**
  - DOE "Top 10" award in 2002 recognized this effort, as part of larger componentization context

Performance optimization

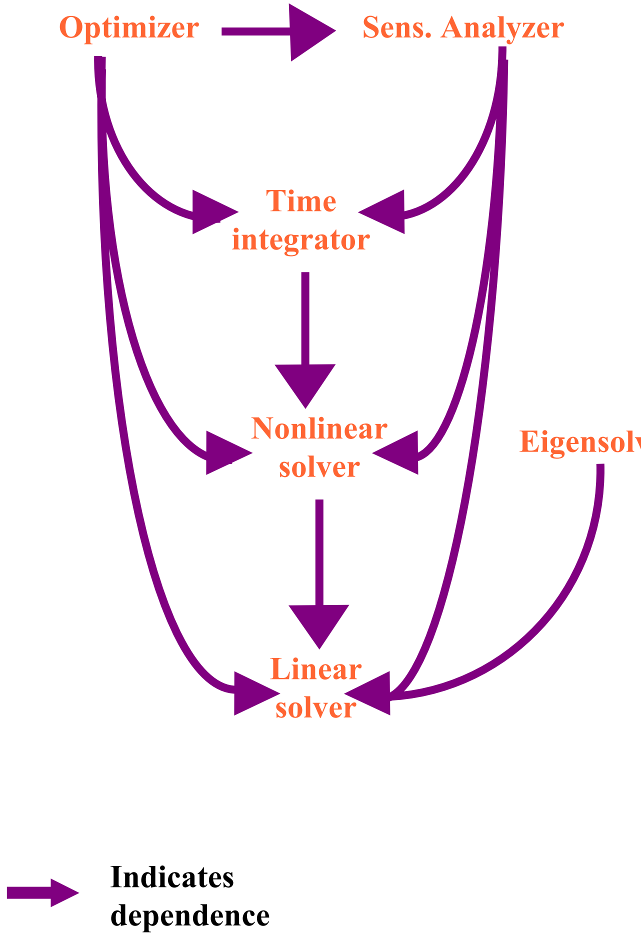
Optimal algorithms for large sparse matrices are prone to poor per-processor percentage of peak, since memory latency is ~100X processor clock period

Critical to block sparse computations for registers and for cache

TOPS leverages expertise that tuned dense kernels previously (ATLAS, PhiPAC)

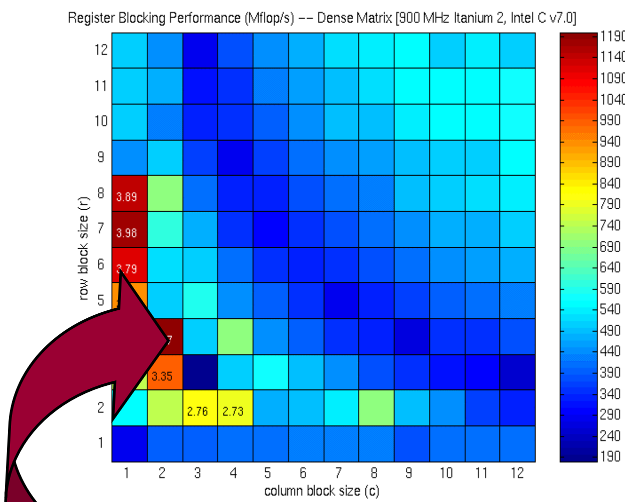
In 1999 TOPS researchers demonstrated gains of 2.5 to 7X over range of commercial microprocessors for NASA unstructured Euler code, from blocking and reordering (part of Bell Prize that year)

Current efforts include atomic composite operations, common in solvers, e.g., $A^T A x$



Performance optimization progress

- **TOPS has tuned sparse kernels**
  - (Jacobian) matrix-vector multiplication
  - sparse factorization
  - multigrid relaxation
- **Running on dozens of apps/platform combinations**
  - Power3 (NERSC) and Power4 (ORNL)
  - factors of 2 on structured (CMRS) and unstructured (CEMM) fusion apps
- **"Best student paper" at ICS2002 went to TOPS team**
  - theoretical model and experiments on effects of register blocking for sparse mat-vec



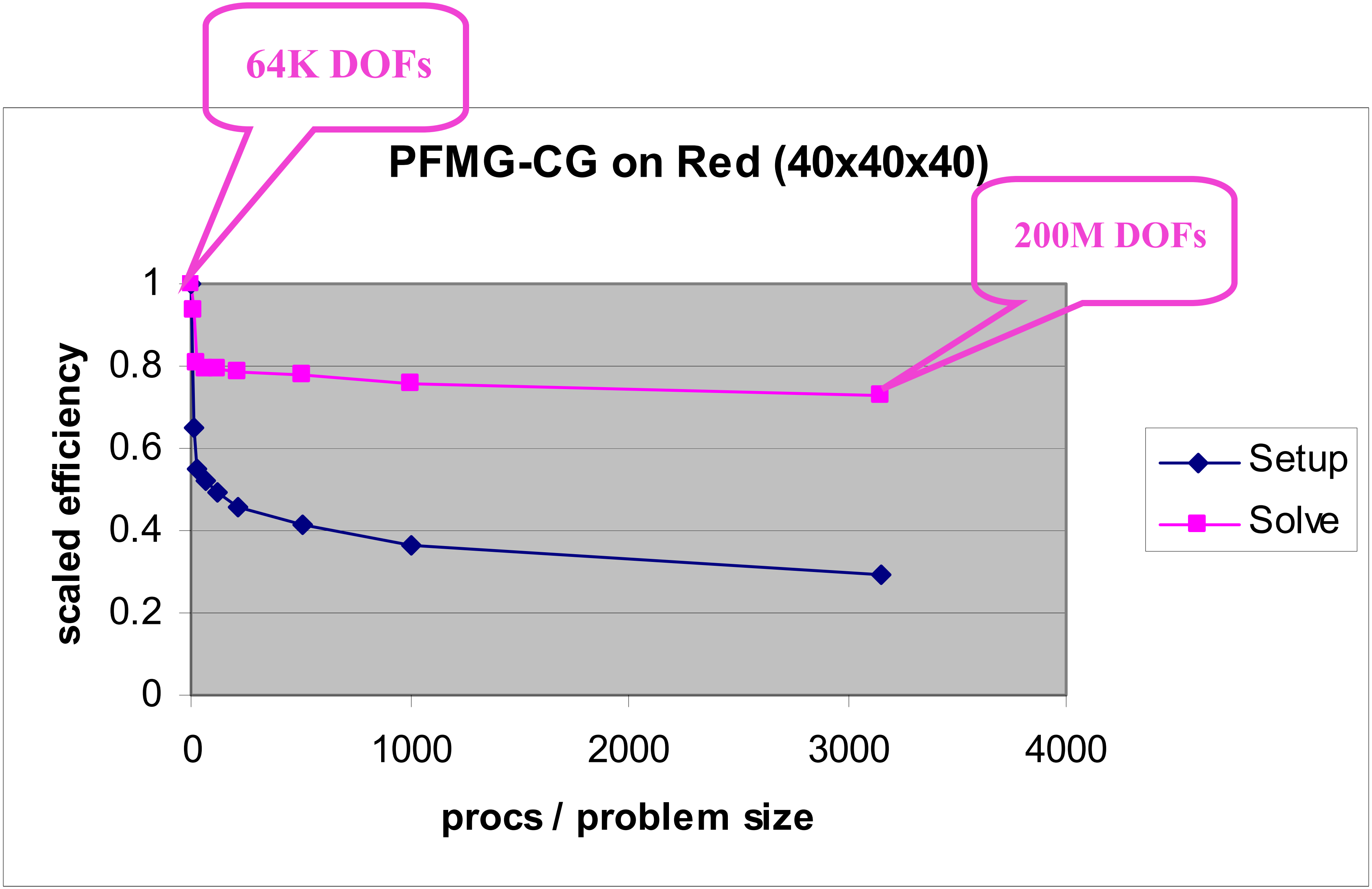
Blocking of 4 rows by 2 columns is 4.07 times faster on Itanium2 than default 1×1 blocks

We run on the actual SciDAC platforms …

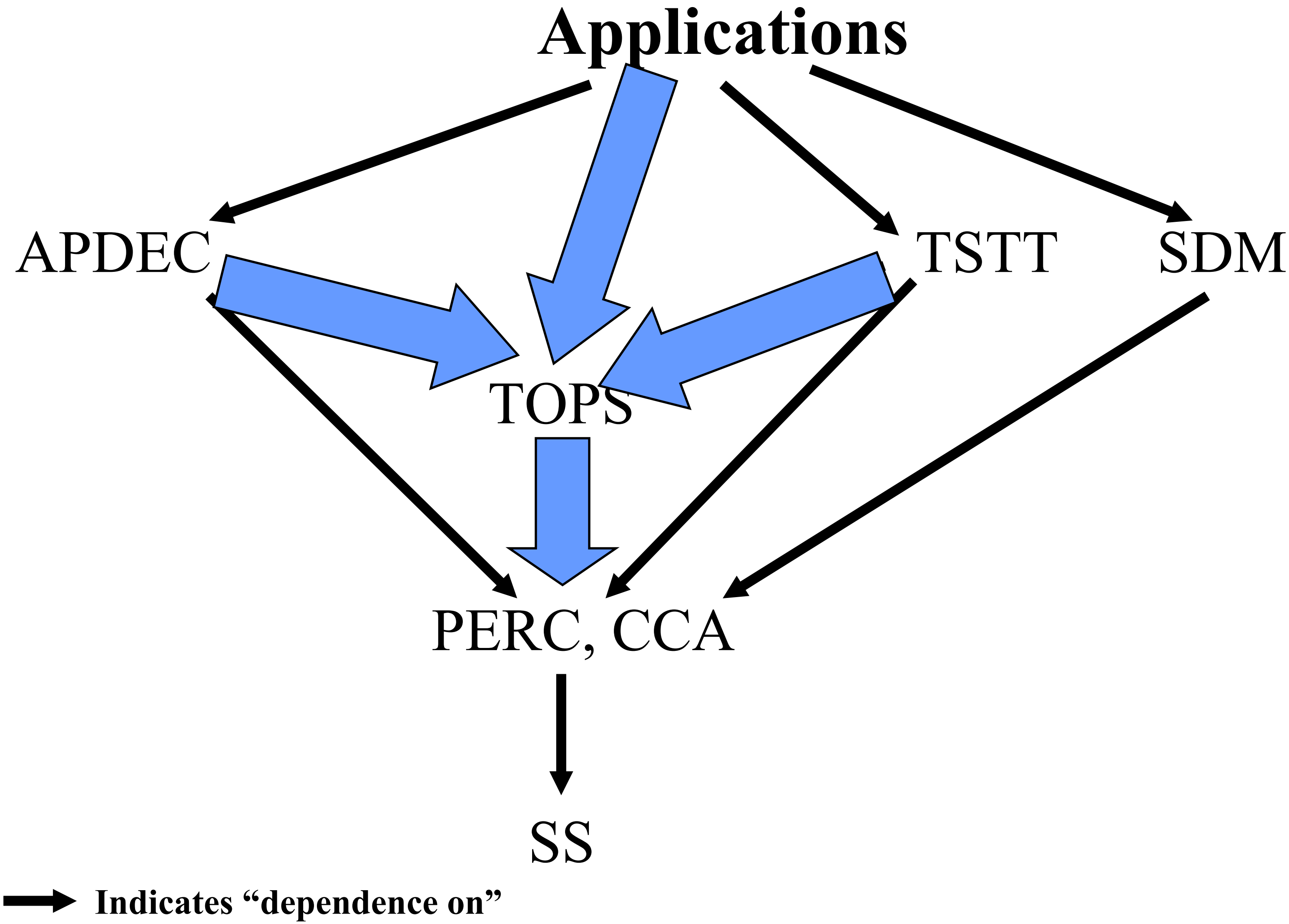
- IBM Power3+ SMP
- 16 procs per node
- 208 nodes
- 24 Gflop/s per node
- 5 Tflop/s (upgraded to 10, Feb 2003)

- IBM Power4 Regatta
- 32 procs per node
- 24 nodes
- 166 Gflop/s per node
- 4Tflop/s (10 in 2003)

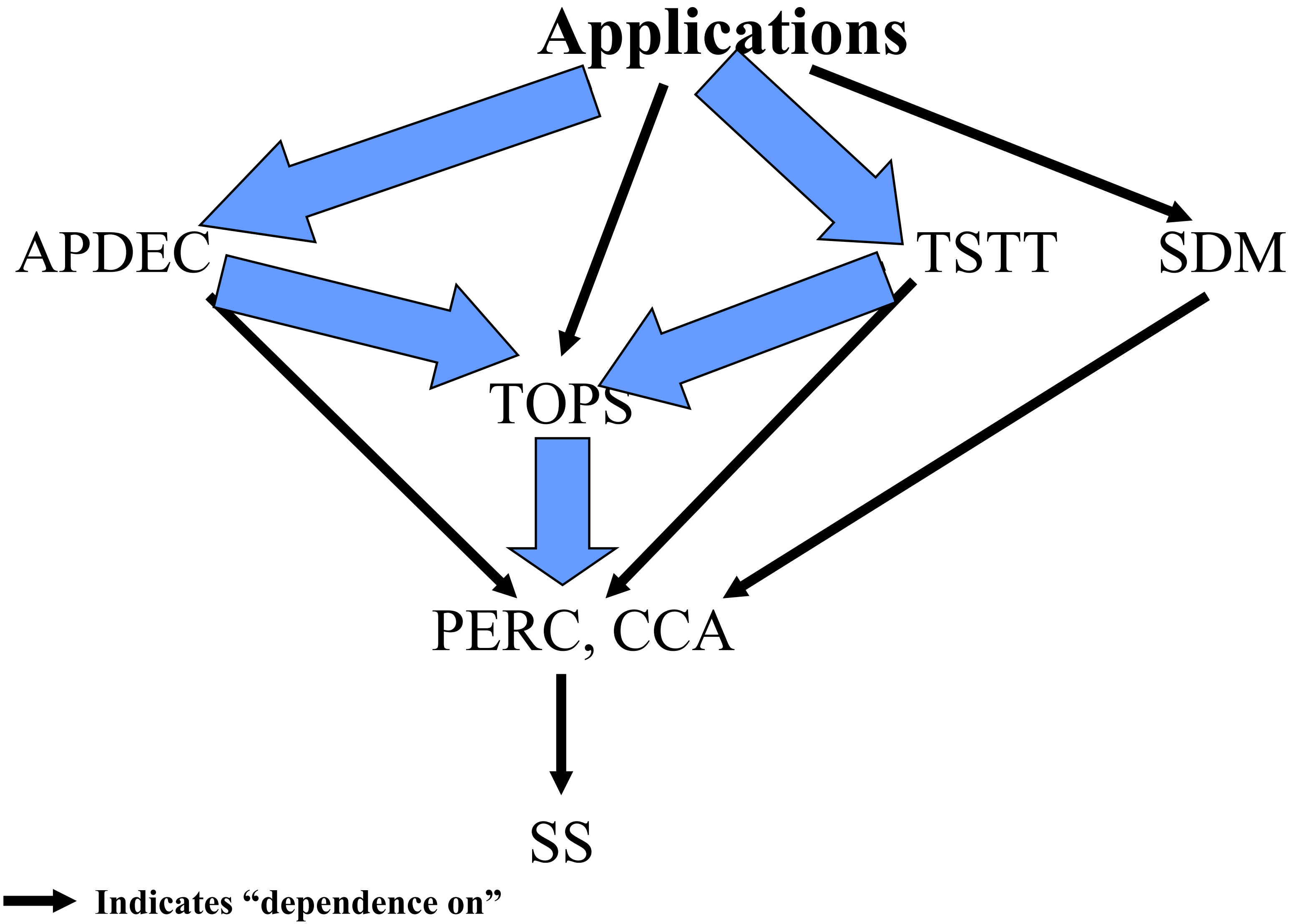
Parallel efficiency less a concern than serial!



Primary interaction pathways, 2003*



Primary interaction pathways, 2005*



Some high-impact TOPS "sure shots"

- **Robust and performant linear elliptic solves for composite-grid scalar systems in AMR via multigrid**
  - applications in combustion, fusion
- **Convenient sensitivity analysis for time-dependent and steady-state parameterized systems**
  - applications in astrophysics, climate, combustion
- **Expansion of spatial dimensionality (from 1 or 2 now to 3) in multigroup particle-rad-hydro via Newton-Krylov-MG**
  - application in astrophysics
- **Extension of discrete dimensionality in sparse generalized eigensystem analysis**
  - application in accelerator design

Some high-impact TOPS "medium shots"

- **Robust and performant solves for composite-grid coupled nonlinear systems in AMR via Newton-Krylov-MG**
  - applications in combustion, fusion
- **Robust and performant solves for high-order discretizations via multigrid**
  - applications in astrophysics, combustion, fusion
- **Robust and performant solves for *div-curl* discretizations via multigrid**
  - applications in climate
- **Robust and performant solves for *curl-curl* discretizations with high anisotropy via multigrid**
  - applications in fusion
- **Expansion of dimensionality up to 5 or 6 for phase-space Boltzmann methods through more optimal solvers**
  - applications in fusion

Some high-impact TOPS "long shots"

- **Breakthrough economizations in QCD lattice solver via multigrid**
  - **application to lattice gauge theory and QCDOC, BlueGene/L computers**
- **Breakthrough levels of accelerator efficiency and operational stability via shape optimization**
  - **application to accelerator design (*e.g.*, NLC)**
- **Breakthrough levels of tokamak and stellerator efficiency and operational stability via shape optimization**
  - **application to fusion energy device design (*e.g.*, ITER)**

Interactions: APDEC

- Main goal: provide full-depth multigrid solvers from Hypre for (often anisotropic) scalar linear problems on Chombo's composite AMR grids
- Chombo's native multigrid does not coarsen beyond geometric limits and does not always converge
- Hypre in Chombo now as a bottom solver; however its performance is poor relative to native
- Plan: improve performance of current bottom solver for AMR applications
- Plan: release specification and implementation for parallel semistructured AMG interface, add parallel FAC code

Interactions: TSTT

- **Main goal: provide multigrid and Newton-Krylov solvers for composite grid discretizations, through interfaces higher than default unstructured sparse graphs and matrices**
- **Besides supporting applications that call TOPS solvers *through* TSTT, TOPS also supports mesh optimization functions *within* TSTT**
- **Plan: collaborate on interface definitions**
- **Plan: develop multilevel solvers for high-order discretizations supported by TSTT**
- **Plan: reuse more of PETSc's distributed data structures in dynamic adaptive context, as performance may dictate**

Interactions: CCTTSS

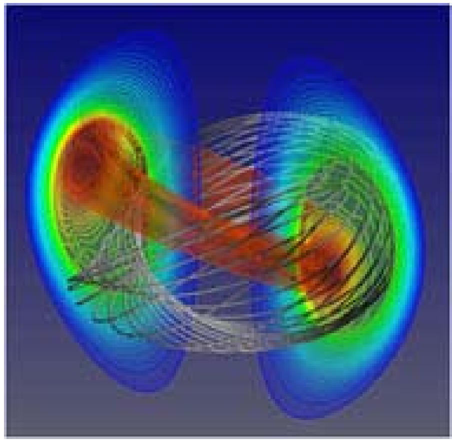
- Main goal: achieve language interoperability and improve software design via SIDL, and package TOPS solvers as components for CCA framework
- Richest interaction so far with any team – fundamental to TOPS, which is in turn driving SIDL development
- PETSc, Hypre both being SIDL'ized
- TOPS is an early demonstrator of the power of the CCA framework approach, within its own software domain (part of OASCR's only "DOE SC Top 10" achievement in 2002)
- Plan: develop abstract component interfaces for linear algebra (including eigenanalysis), nonlinear algebra, and unconstrained and constrained optimization

Interactions: PERC

- Main goal: use PERC's tools to understand, predict, tune, and improve performance efficiency of solvers
- Second richest interaction so far with any other team
- TOPS implicit solver examples provide simple free-standing code targets for PERC
- TOPS application partnerships provide relevant test data to PERC
- Plan: assist users to benefit from PERC's tools (as applied to solvers)
- Plan: create insertion path for TOPS' own successes in performance improvements for sparse kernels
- Plan: assist PERC to evaluate potential hardware acquisitions for DOE SC

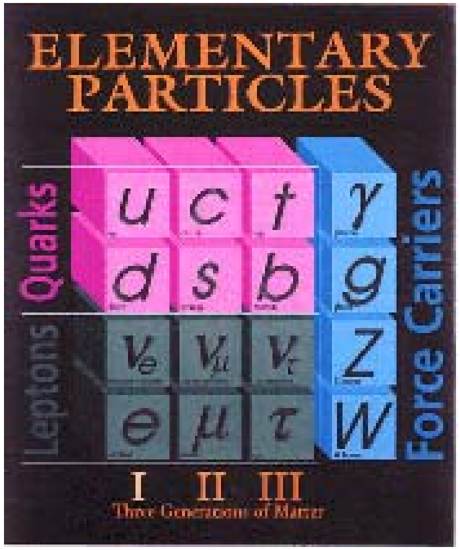
Interactions: FES projects

- *Projects: CEMM/M3D, CEMM/Nimrod, CMRS*
- **Main goal: lead fusion simulation program, with its multirate, multiphysics problems into nonlinearly implicit solvers**
- **In tight development cycle with M3D and CMRS projects already; Nimrod at a more exploratory stage**
- **Plan: work through existing PETSc interface in M3D to deliver successively more scalable implicit methods, beyond current "partially implicit"**
- **Plan: demonstrate cost-effectiveness of NK-MG directly in CMRS Hall MHD magnetic reconnection**
- **Plan: use Nimrod as driving app for MG for high-order discretization**



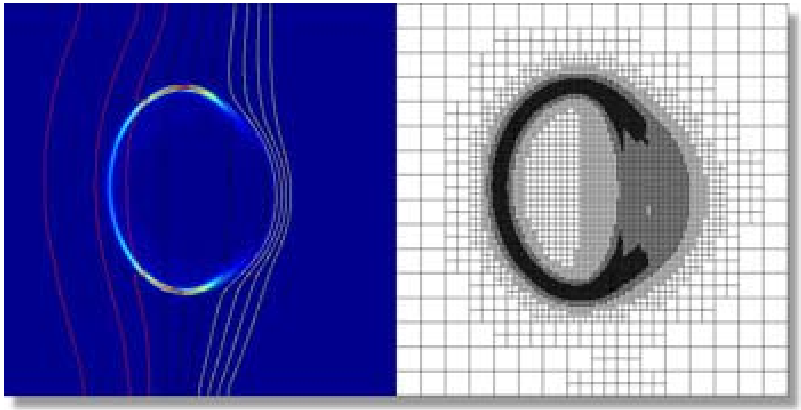
Interactions: HENP projects

- *Projects: AST, NCILGT, SSC, TSI*
- Main goal: support existing tight collaborations with AST (eigensolvers) and TSI (linear solvers)
- TSI and SSC have similar need to expand multiphysics simulations to higher dimensions
- Plan: provide many algorithmic options for AST
- Plan: lead TSI into higher dimensions with MG, and then into NK-MG
- Plan: leverage TSI for SSC
- Plan: seek breakthrough in NCILGT with AMG, replacing or preconditioning CG on the Wilson-Fermion operator



Interactions: BES projects

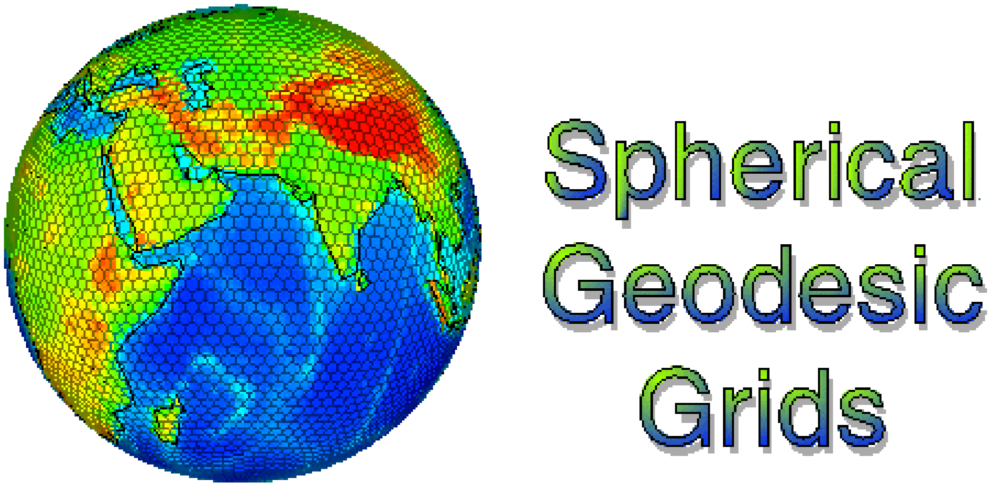
- *Projects: CFRFS, ASCTKD, NCMGO (CCTTSS)*
- **Main goal: support TOPS software (integrators, eigensolvers, optimizers) already in quotidian use in these three groups**



- **Plan: lead CFRFS into sensitivity analysis**
- **Plan: improve preconditioning for Jacobi-Davidson in ASCTKD through multigrid in SPAM**
- **Plan: help NCMGO to expand scaling beyond 10^4 atoms**

Interactions: BER projects

- ***Project: Geodesic Climate Model***
- **Main goal: assist with development of scalable *div-curl* solver for nonsimply connected domains embedded on the surface of a sphere**
- **Plan: provide on-going algorithmic consultation**



Lessons to date

- **Working with the same code on the same machine vastly speeds collaboration, as opposed to ftp'ing matrices around the country, etc.**
- **Exchanging codes better than exchanging papers**
- **Version control systems essential to having any last impact or "insertion path" for solver improvements**
- **"Doing physics" more fun than doing driven cavities**

TOPS outreach

- **Downloadable software**
- **Technical publications**
- **Technical and overview presentations**
- **Service as “point of contact” for scalable solvers**
  - **for individual PIs**
  - **in interdisciplinary meetings**
  - **in creation of whitepapers**

TOPS software outreach

- **Hypre** – scalable preconditioners
  www.llnl.gov/CASC/hypre
- **PARPACK** – scalable eigensolvers
  hpcf.nersc.gov/software/libs/math/parpack
- **PETSc** – scalable nonlinear and linear solvers
  www.mcs.anl.gov/petsc
- **SUNDIALS** – scalable ODE and nonlinear solvers
  www.llnl.gov/CASC/sundials
- **SuperLU** – parallel direct sparse LU methods
  www.nersc.gov/~xiaoye/SuperLU
- **TAO** – scalable general-purpose optimizers
  www.mcs.anl.gov/tao
- **Veltisto** – scalable PDE-constrained optimizers
  www.cs.nyu.edu/~biros/veltisto

TOPS publication outreach

- **Through April 2003, 29 TOPS-affiliated individuals had co-authored 77 works of scholarship under TOPS**
  - **42 journal papers**
    - **26 published or to appear**
    - **16 submitted and pending review**
  - **25 chapters in proceedings or solicited collections**
  - **4 technical reports**
  - **2 workshop reports**
  - **2 doctoral dissertations**
  - **2 edited proceedings**
- **At least one forthcoming major book (Widlund) will acknowledge TOPS**

TOPS presentation outreach

- **Tutorials**

  ACTS, SIAM PP, Supercomputing

- **Minisymposia**

  SIAM Annual, SIAM CS&E, USNCCM

- **Conf/workshop lectures**

  Advanced architecture, Domain decomposition, High energy physics, Magnetohydrodynamics, Multigrid, Multiscale, Nanotechnology, Parallel scaling, Salishan, Supercomputing, Supernovae

- **Lab seminars***

  ARL, BNL, General Atomics, LANL, NASA-Langley, ORNL, PPPL

- **University seminars***

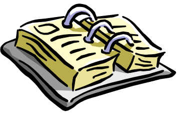
  Chicago, Columbia, Harvard, Houston, PennState, Stanford, UArkansas, UCalifornia, UIllinois, UKentucky

- **International**

  China, France, Germany, Mexico

***not including our own**

Ex.: Seven questions for users



- **Has your solver been unchanged for the past five or ten years?**
- **Is your solver running at 1-10% of machine peak?**

- **Do you spend more time in your solver than in your physics?**
- **Is your discretization or model fidelity limited by the solver?**

- **Is your time stepping limited by stability?**
- **Are you running loops *around* your analysis code?**

- **Do you care how sensitive to parameters your results are?**

If the answer to any of these questions is "yes", you may be a customer!

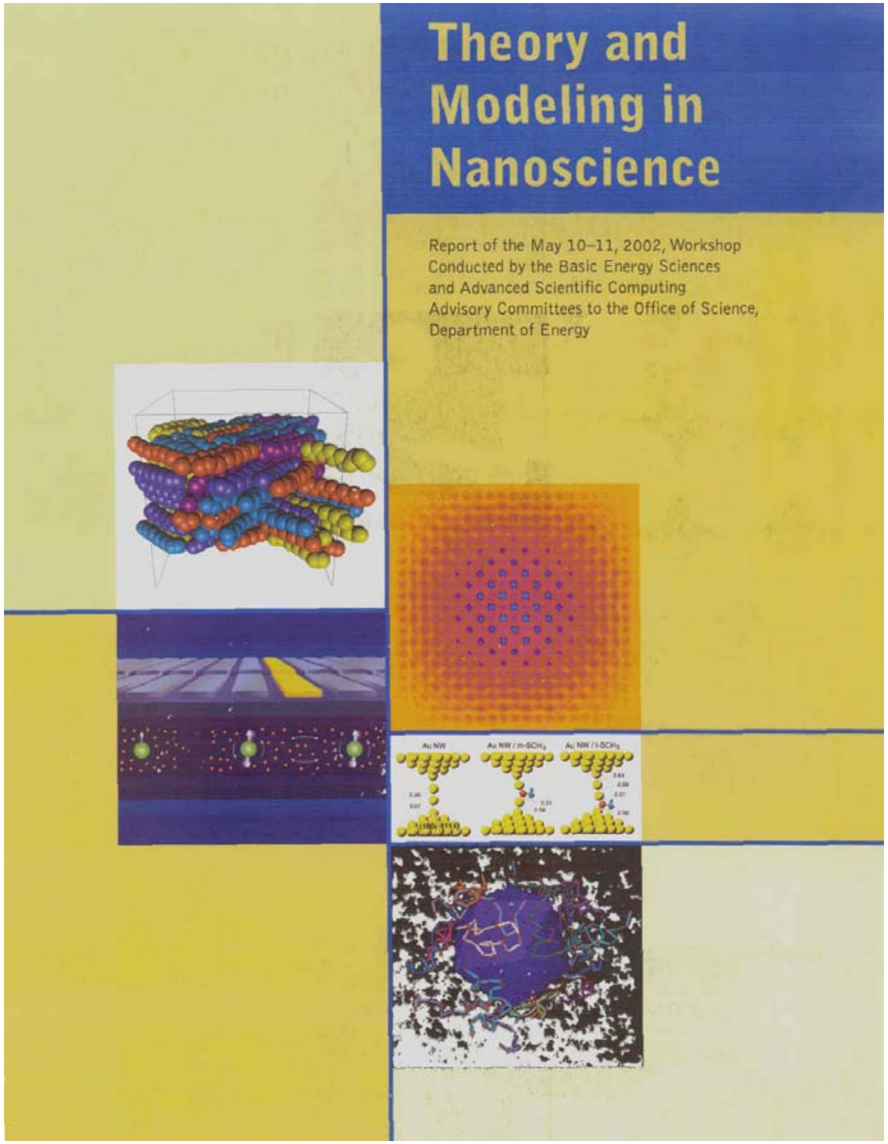
Ex.: Expectations of users

- **Be willing to experiment with novel algorithmic choices – optimality is *rarely* achieved beyond model problems without interplay between physics and algorithmics!**
- **Adopt flexible, extensible programming styles in which algorithmic and data structures are not hardwired**
- **Be willing to let us play with the real code you care about, but be willing, as well to abstract out relevant compact tests**
- **Be willing to make concrete requests, to understand that requests must be prioritized, and to work with us in addressing the high priority requests**
- **If possible, *profile* before seeking help**

TOPS service outreach

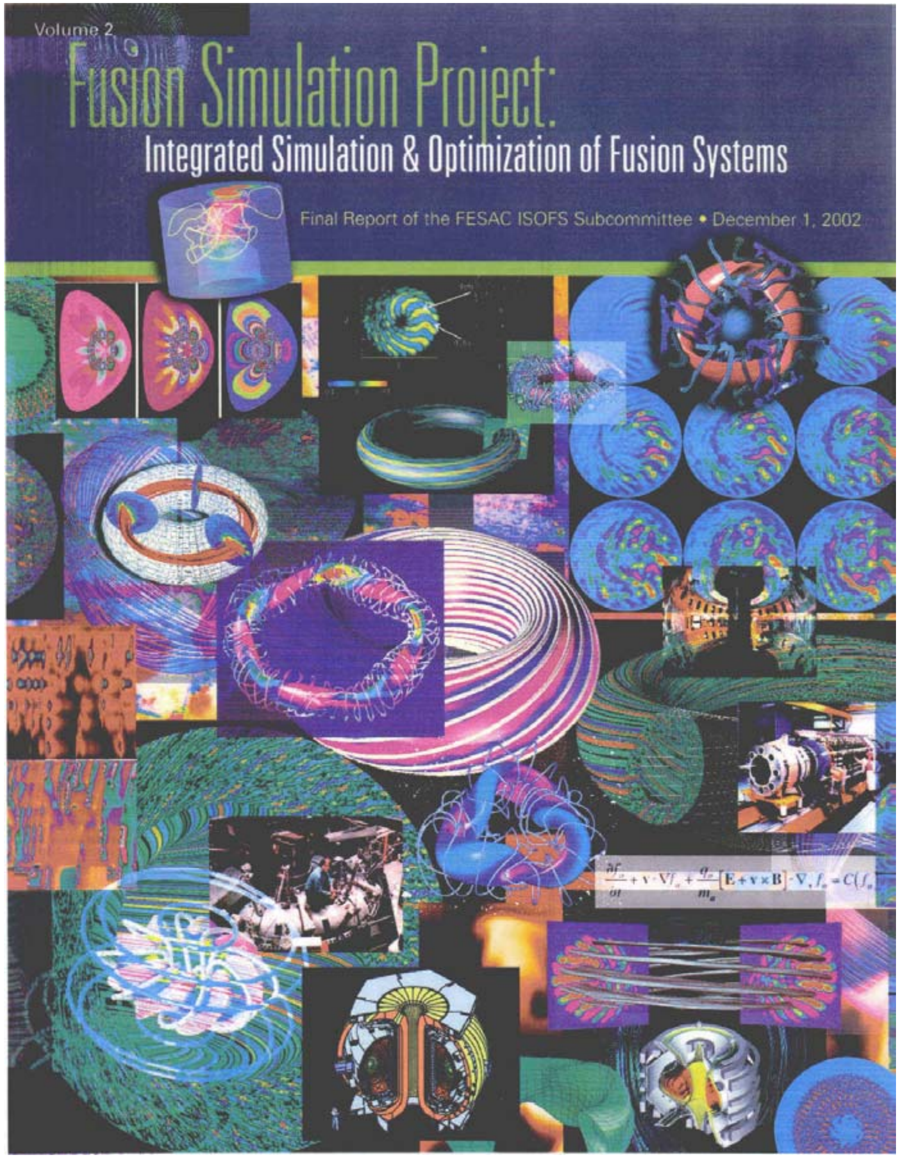
- Extensive e-mail and phone consulting by lab PIs on major codes “in the field” (Hypre, PETSc, SUNDIALS, SuperLU, TAO)
- Representation of the TOPS ISIC at interdisciplinary meetings (including the SciDAC PI meetings, themselves)
- Organization of and contribution to various research whitepapers which attach importance to scalable solvers

Outreach: nanoscience modeling



- Jul 2002 report to DOE
- Proposes $5M/year theory and modeling initiative to accompany the existing $50M/year experimental initiative in nano science
- Report lays out research in numerical algorithms and optimization methods on the critical path to progress in nanotechnology

Outreach: integrated fusion modeling



- **Dec 2002 report to DOE**
- **Currently DOE supports 52 codes in Fusion Energy Sciences**
- **US contribution to ITER will "major" in simulation**
- **Initiative proposes to use advanced computer science techniques and numerical algorithms to improve the US code base in magnetic fusion energy and allow codes to interoperate**

Outreach: Genomes to Life program

GENOMES to LIFE
BIOLOGICAL SOLUTIONS FOR ENERGY CHALLENGES
U.S. DEPARTMENT OF ENERGY
INNOVATIVE APPROACHES ALONG UNCONVENTIONAL PATHS

http://DOEGenomesToLife.org/compbio/

**Report on the Mathematics Workshop
for the
Genomes to Life Program**

U.S. Department of Energy
Gaithersburg, Maryland
March 18–19, 2002

Workshop Organizers
David L. Brown, Lawrence Livermore National Laboratory
John Guckenheimer, Cornell University
Esmond G. Ng, Lawrence Berkeley National Laboratory

Prepared by the Office of Advanced Scientific Computing Research
and
Office of Biological and Environmental Research
of the
U.S. Department of Energy
Office of Science

December 2002

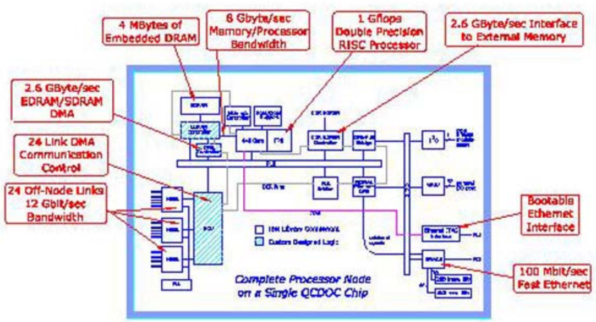
Office of Science
U.S. DEPARTMENT OF ENERGY

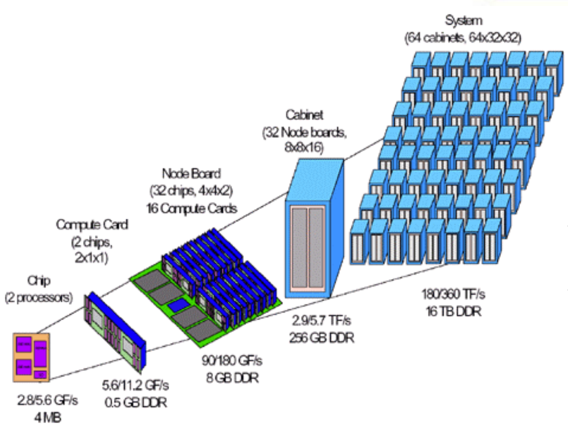
- Dec 2002 report to DOE
- DOE anticipates leadership in the microbial aspects of federal genomic research
- Applications to: climate change, bioremediation, and energy production, as well as health risks to humans
- Initiative will identify and characterize molecular machines of life, gene regulatory networks, and complex microbial communities
- Mathematics role includes PDEs

Outreach: new machines

First cabinet of Cray X-1, delivered to ORNL, March 2003

QCDOC, to be delivered to BNL, August 2003

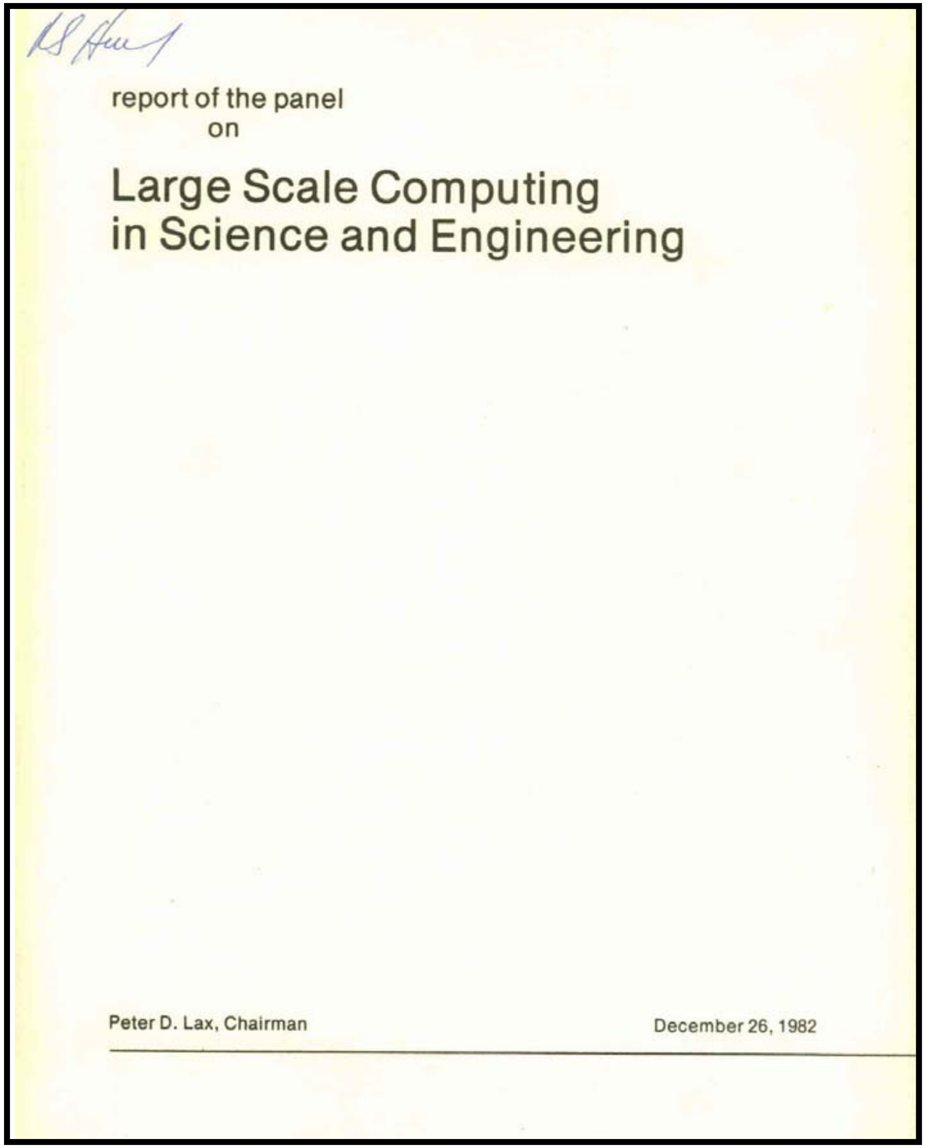




IBM BlueGene/L, to be delivered to LLNL, December 2004

- **Workshops throughout 2002**
- **DOE anticipates acquiring facilities in the 100 teraflop/s range**
- **PDE-based simulations will be among the prime jobs run on these new platforms**
- **TOPS-supported solver software has been identified as among the most important, after systems software, to get up and running**

Outreach: new report



- "Science Case for Large-scale Simulation"
- "Update of Lax" (shown)
- Extension of SciDAC
- Commissioned April 2003
- To be delivered July 2003
- Large degree of responsibility for this report has been put on the SciDAC Math ISIC PIs (and some of their reviewers ☺)

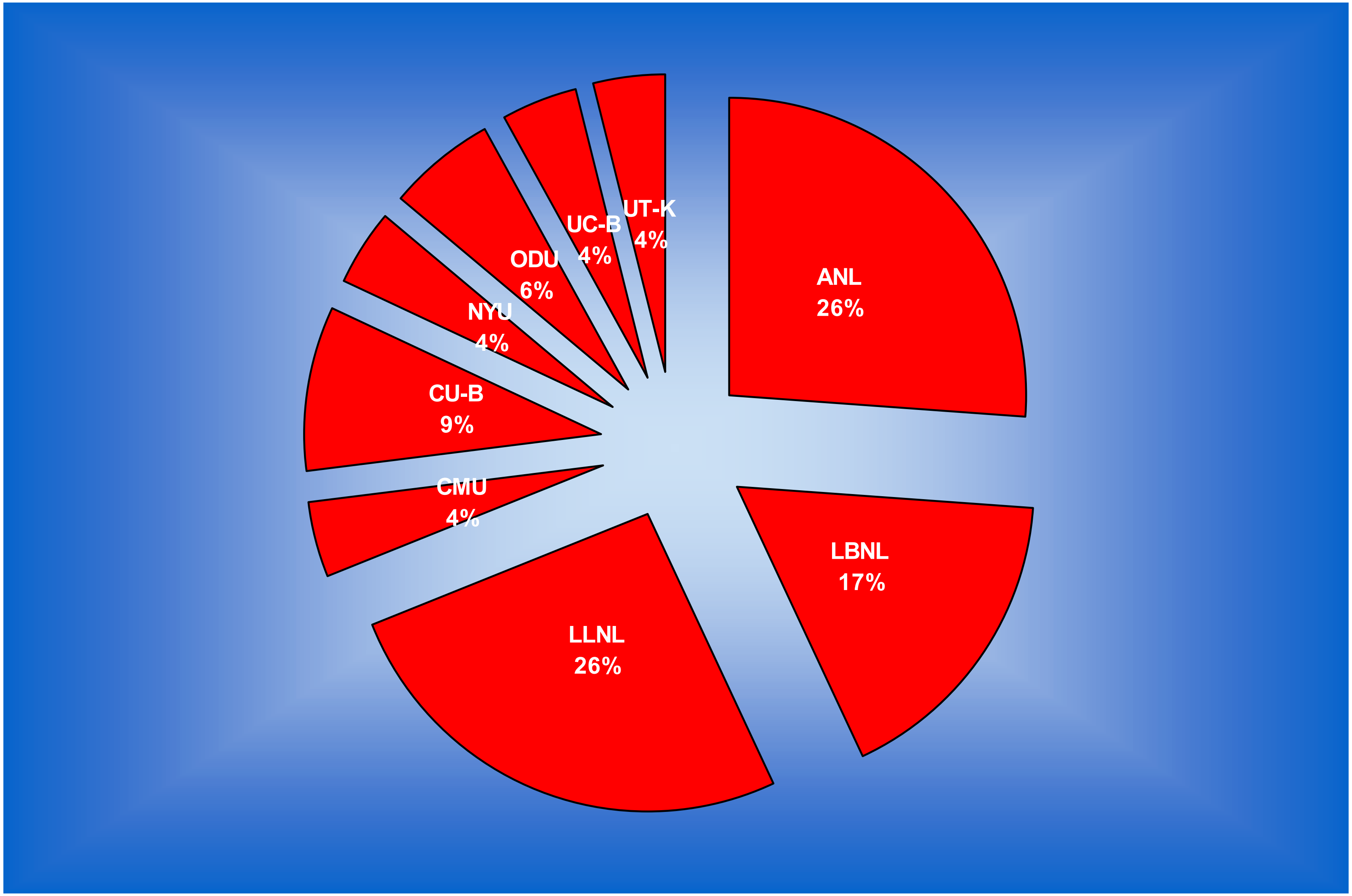
Time to drill down

- **Supporting Accelerator Science & Technology (E. Ng)**
  - featuring LBNL-led eigensolver and direct methods development
  - example of applications-driven TOPS work
- **Componentizing optimization (J. Moré)**
  - featuring ANL-led nonlinear optimization theory and development
  - example of componentization and CS ISIC-driven TOPS work
- **Bringing multilevel methods to the masses (R. Falgout)**
  - featuring LLNL-led multigrid theory and development
  - example of interoperability and math ISIC-driven TOPS work
- **Unifying solver frameworks (B. Smith)**
  - communicating the basis for the TOPS solver interface vision

Briefing book

- Algorithmic and software progress
- Research interactions
- Coordination of activities
- Appendices
  - Original proposal
  - 2-page summaries (5)
  - Posters (5)
  - Software
  - Publications

TOPS institutional allocations



Primary TOPS personnel*

- **ANL (7):** S. Benson, M. Knepley, M. Minkoff, J. Moré, T. Munson, B. Smith, H. Zhang
- **LBNL (6):** P. Husbands, X. Li, O. Marques, E. Ng, A. Pinar, C. Yang
- **LLNL (5):** E. Chow, R. Falgout, R. Serban, P. Vassilevski, C. Woodward
- **CMU (2):** V. Akcelik, O. Ghattas
- **CU-B (4):** X.-C. Cai, T. Manteuffel, S. McCormick, J. Ruge
- **NYU (3):** G. Biros, B. Hientzsch, O. Widlund
- **ODU (3):** F. Dobrian, D. Keyes, A. Pothen
- **UC-B (1):** J. Demmel
- **UT-K (2):** J. Dongarra, V. Eijkhout

*Does not include students, who circulate between universities and lab

Other participants in TOPS* work

- **Hosted by ANL:** Padma Raghavan, José Roman
- **Hosted by LBNL:** Tim Davis, Padma Raghavan
- **Hosted by LLNL:** Marian Brezina, Tim Chartier, Leszek Marcinkowski

*Supported either by TOPS or with other leveraging from host

What we believe

- **Many of us in TOPS came to work on solvers through interests in applications**
- **What we believe about …**
  - **applications**
  - **users**
  - **solvers**
  - **legacy codes**
  - **software**

… will impact how comfortable applications groups are collaborating with us

What we believe about *apps*

- **Solution of a system of PDEs is rarely a goal in itself**
  - Actual goal is characterization of a response surface or a design or control strategy
  - Solving the PDE is just one forward map in this process
  - Together with analysis, sensitivities and stability are often desired

⇒ **Software tools for PDE solution should also support related follow-on desires**

- **No general purpose PDE solver can anticipate all needs**
  - Why we have *national laboratories*, not *numerical libraries* for PDEs today
  - A PDE solver improves with user interaction
  - Pace of algorithmic development is very rapid

⇒ **Extensibility is important**

What we believe about *users*

- **Solvers are used by people of varying numerical backgrounds**
  - **Some expect MATLAB-like defaults**
  - **Others want to control everything, e.g., even varying the type of smoother and number of smoothings on different levels of a multigrid algorithm**

$\Rightarrow$ **Multilayered software design is important**

- **Users' demand for resolution is virtually insatiable**
  - **Relieving resolution requirements with modeling (e.g., turbulence closures, homogenization) only defers the demand for resolution to the next level**
  - **Validating such models requires high resolution**

$\Rightarrow$ **Processor scalability and algorithmic scalability (optimality) are critical**

What we believe about *legacy code*

- Porting to a scalable framework does not mean starting from scratch
  - High-value meshing and physics routines in original languages can be substantially preserved
  - Partitioning, reordering and mapping onto distributed data structures (that we may provide) adds code but little runtime

⇒ Distributions should include code samples exemplifying "separation of concerns"

- Legacy solvers may be limiting resolution, accuracy, and generality of modeling overall
  - Replacing the solver may "solve" several other issues
  - However, pieces of the legacy solver may have value as part of a preconditioner

⇒ Solver toolkits should include "shells" for callbacks to high value legacy routines

What we believe about *solvers*

- **Solvers are employed as part of a larger code**
  - Solver library is not only library to be linked
  - Solvers may be called in multiple, nested places
  - Solvers typically make callbacks
  - Solvers should be swappable

⇒ **Solver threads must not interfere with other component threads, including other active instances of themselves**

- **Solvers are employed in many ways over the life cycle of an applications code**
  - During development and upgrading, robustness (of the solver) and verbose diagnostics are important
  - During production, solvers are streamlined for performance

⇒ **Tunability is important**

What we believe about *software*

- A continuous operator may appear in a discrete code in many different instances
  - Optimal algorithms tend to be hierarchical and nested iterative
  - Processor-scalable algorithms tend to be domain-decomposed and concurrent iterative
  - Majority of progress towards desired highly resolved, high fidelity result occurs through cost-effective low resolution, low fidelity parallel efficient stages

⇒ Operator abstractions and recurrence are important

- Hardware changes many times over the life cycle of a software package
  - Processors, memory, and networks evolve annually
  - Machines are replaced every 3-5 years at major DOE centers
  - Codes persist for decades

⇒ Portability is critical

Summary

- Original vision of the TOPS team has been vindicated, as to its importance and timeliness, and expanded, as to its scope
- TOPS has already delivered on new theory, new algorithms, new code, new applications support, and extensive outreach for the larger SciDAC mission, and is actively collaborating on many fronts
- TOPS is integral to plans of two Math ISICs and two CS ISICS
- TOPS must now work hard to follow through on timely and performant delivery of solvers to its most important clients
- TOPS cannot presently meet all of the opportunities for bringing new solution techniques into DOE Office of Science computational practice and will have to narrow its focus if maintained at its current size; it awaits objective evaluation and guidance

"Knowing what is big and what is small is more important than being able to solve partial differential equations." **– S. Ulam**

TOPS

Terascale Optimal PDE Simulations

http://www.tops-scidac.org