SciDAC Institute Poster Blitz

September 10, 2012

Poster Session

4:30-6:00pm in Roosevelt/Madison

Format

- 26 Posters
- 4 Institutes (FASTMath, SUPER, QUEST, SDAV)
- 1 slide each
- 60 seconds
  - ***Never send to know for whom the bell tolls, it tolls for thee*** -John Donne

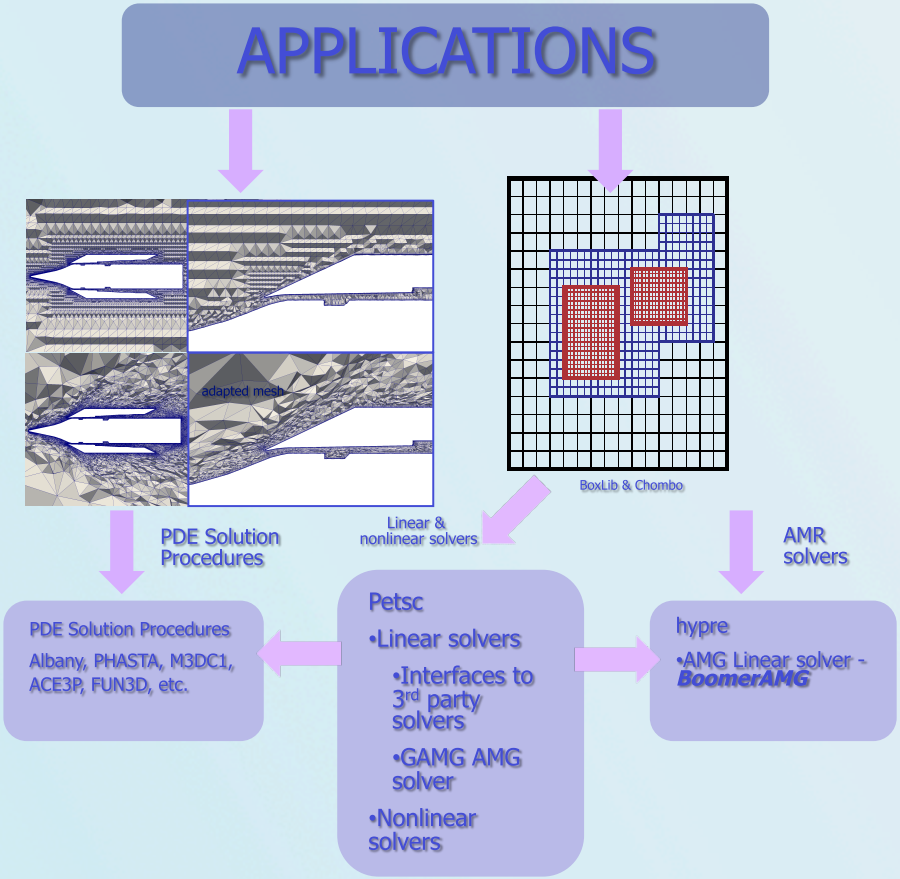
….Ready Mark?

Solver – Mesh interactions

FASTMath Team Members:

Structured and unstructured mesh interaction with equation solver libraries: BoxLib, Chombo & MeshAdapt with hypre, PETSc & Trilinos

- Mesh & discretization frameworks & tools
  - BoxLib – structured grid AMR framework
  - Chombo - structured grid AMR framework
  - AdaptMesh – unstructured grid AMR tool
- Equation solver libraries in FASTMath
  - HYPRE – algebraic and structure gird multigrid solvers
  - PETSc – numerical library w/ interfaces to 3rd party (AMG) solver (hypre, ML)
  - Trilinos – AMG solver ML
- Couple mesh/discretization to equation solvers
  - Linear sparse algebraic multigrid (AMG)
  - Linear structured grid solvers (GMG)
  - Nonlinear solvers
  - Time steppers



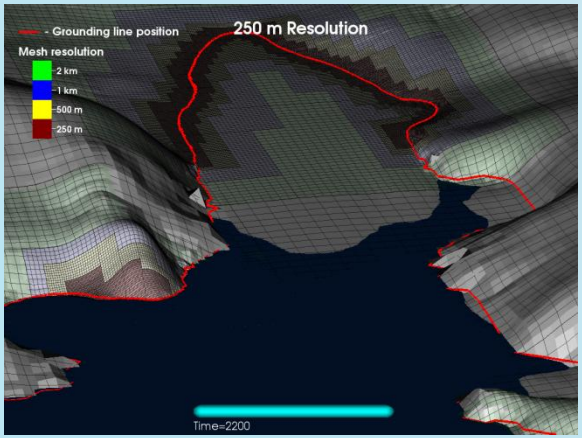
More Information: http://www.fastmath-scidac.org or contact Lori Diachin, LLNL, diachin2@llnl.gov, 925-422-7130

Block-structured AMR: Algorithms and Applications

FASTMath Team Members: Ann Almgren, John Bell, Mike Lijewski (BoxLib);
Phil Colella, Daniel Graves, Terry Ligocki, Mark Adams, Brian van Straalen (Chombo)

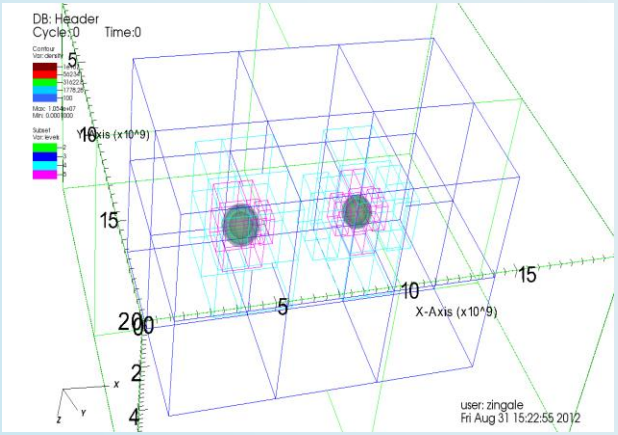
Block-structured adaptive mesh refinement (AMR) software frameworks available for the development of parallel, multiphysics applications

- Block-structured AMR software frameworks
  - BoxLib
  - Chombo
- Each is the basis for a variety of massively parallel application codes
- Available for development of new codes and algorithms (User Guides for both, as well as BoxLib/Tutorials)



BISICLES project – ice sheet modeling using Chombo

- Chombo: emphasis on embedded boundary representation of complex geometry

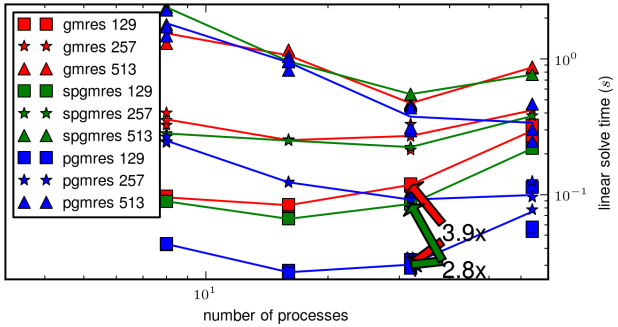


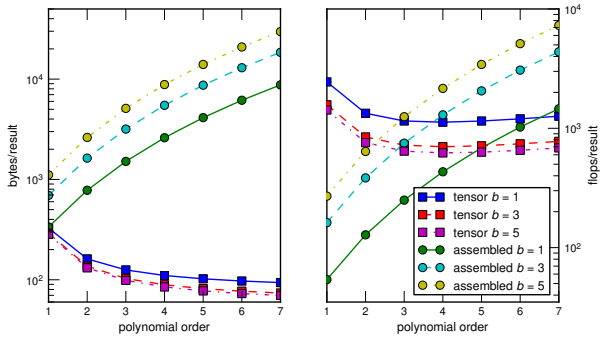
CASTRO code– modeling white dwarf mergers using BoxLib

- Support for finite difference / finite volume discretizations of hyperbolic, parabolic and elliptic PDE's
- Support for particle/mesh schemes (stand-alone or integrated with PDE solvers)
- Higher-order methods in space and time available
- BoxLib: allows rapid prototyping of new approaches to AMR and new multiphysics algorithms

More Information: http://www.fastmath-scidac.org or contact Lori Diachin, LLNL, diachin2@llnl.gov, 925-422-7130

Architecture-aware Themes





- Memory locality and the bandwidth wall
  - Allocation granularity
  - Cooperative prefetch
- Algorithmic adaptation
  - Pipelining to hide latency of collectives
  - Matrix-free and nonlinear methods
- Load balancing
  - Locality and granularity trade-offs
  - Dynamic versus static
  - Interleaved versus locality-enforced memory
- Multicore programming models
  - Library-friendly scoping
  - Sharing thread pools
  - Removing implicit synchronization (e.g. barriers)

Data Partitioning for Scientific Applications and Emerging Architectures

FASTMath Team Members: Karen Devine, Vitus Leung, and Siva Rajamanickam (SNL); Onkar Sahni, Mark Shephard, and Cameron Smith (RPI)

The FASTMath institute provides data-partitioning and load-balancing expertise and software for use in a wide range of scientific applications. Our research addresses data distribution and layout issues for heterogeneous multicore architectures and dynamic, unstructured applications.

Partitioning and Dynamic Load Balancing

- Balance process workloads; minimize communication

New data distribution challenges

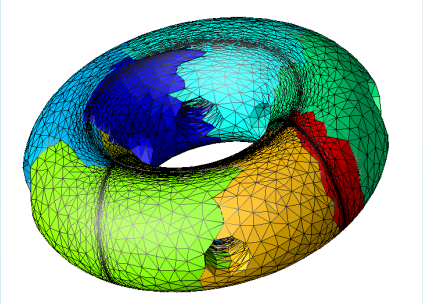
- Heterogeneous, hierarchical architectures
- More complicated, irregular data dependencies

Software:

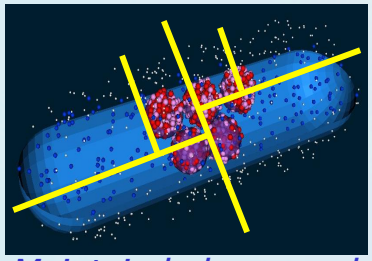
- **Zoltan Partitioning, Load Balancing & Ordering Toolkit**
  - Geometric and topologically-based methods
  - Support for many types of applications (particles, elements, matrices, circuits, etc.)
- **ParMA Partition Improvement using Mesh Adjacencies**
  - Improves partitioning metrics for mesh-based applications
  - Balance multiple constraints (e.g., # elements and # vertices per process)

Research:

- Partitioning for multicore: hierarchical methods, MPI+threads
- Predictive load balancing for use with adaptive mesh refinement



Dynamically redistribute work in adaptive finite element methods



Maintain balance and geometric locality in particle-based methods

More Information: http://www.fastmath-scidac.org or contact Karen Devine, kddevin@sandia.gov

FASTMath Nonlinear Solver and Time Integration Tools

Team Leads: Dan Reynolds [SMU], Andy Salinger [SNL], Barry Smith [ANL], Carol Woodward [LLNL]

Diverse needs of Science Applications for robustness and performance are met by FASTMath's composable software building blocks and expertise.

ODE Integrator

- Implicit (f_E=0)
  - Nonlinear Solve (f_I)
  - Estimate error
  - Adapt Δt

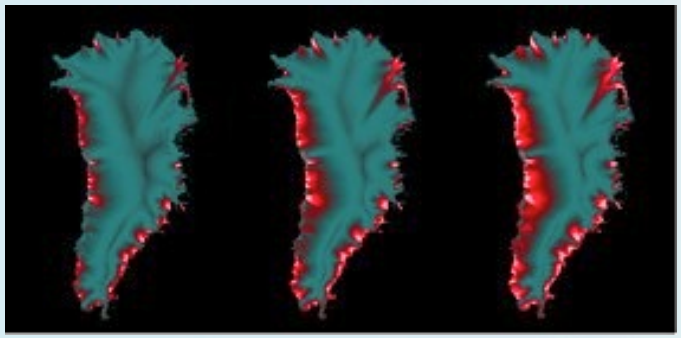
Nonlinear Solver

- Newton
  - Calc F
  - Calc A
  - Krylov
    - calc F'
    - Solve A
      - Precon Krylov

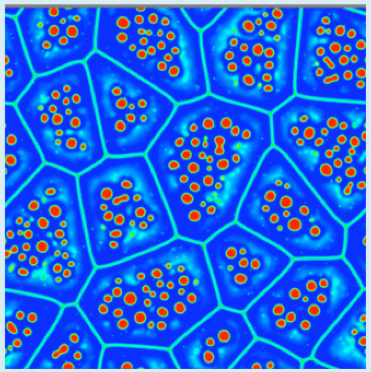
F

Application Impact

- Climate



- Material Science



- Subsurface Flows
- Fusion Energy

FASTMath Software: sundials, PETSc, Trilinos

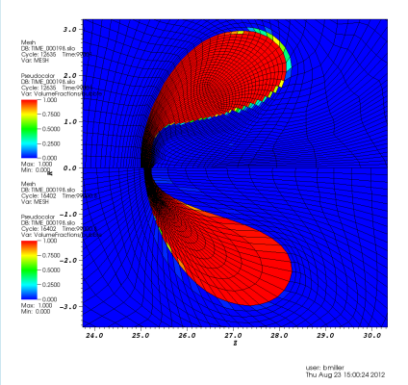
Unstructured Meshing Techniques

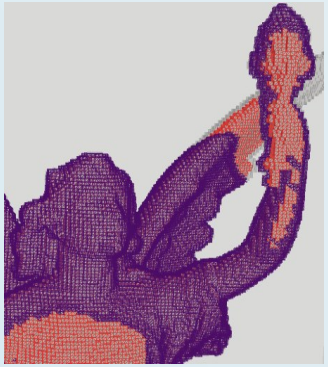
FASTMath Team Members:

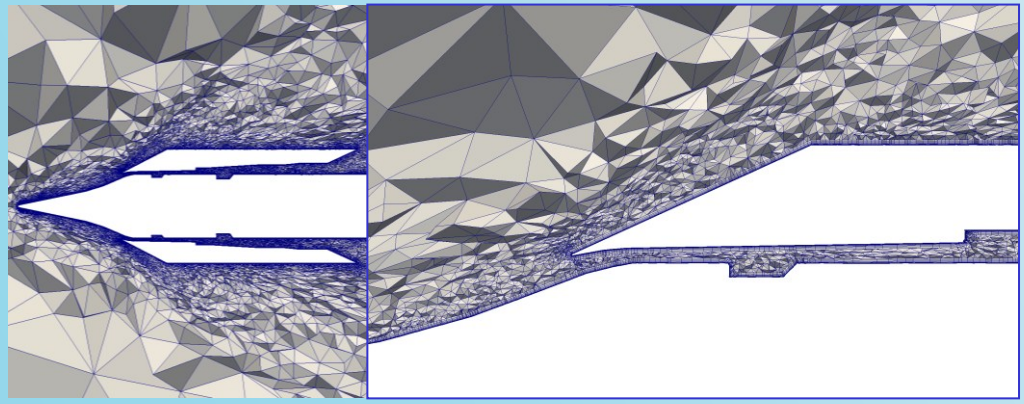
Providing the parallel unstructured mesh data structures and services needed by developers of PDE solution procedures targeted for exascale computers

Unstructured meshes can yield required levels of accuracy using fewer degrees of freedom at the cost of more complex parallel data structures and algorithms.

- FASTMath data structures and services provide PDE solution procedure developers needed tools
- Currently supporting 12 applications







Current FASTMath developments

- Parallel mesh infrastructure
- mesh-to-mesh solution transfer
- mesh optimization
- mesh adaptation
- massively parallel PDE solver
- Meshing operations library

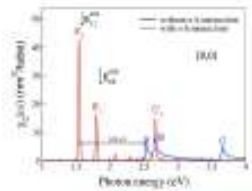
More Information: http://www.fastmath-scidac.org or contact Lori Diachin, LLNL, diachin2@llnl.gov, 925-422-7130

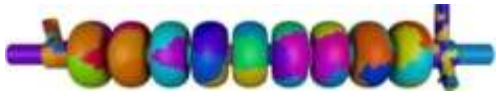
Numerical Algorithms for Large-scale Eigenvalue Calculation

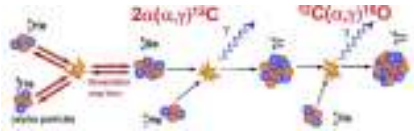
(C. Yang, M. H. Aktulga, E. Ng)

$$\boldsymbol{Ax} = \boldsymbol{\lambda B x};$$
$$\boldsymbol{T}(\boldsymbol{\lambda})\boldsymbol{x} = \boldsymbol{0};$$
$$\boldsymbol{H}(\boldsymbol{x})\boldsymbol{x} = \boldsymbol{\lambda x}$$

Goal: Develop efficient eigensolvers for large-scale linear and nonlinear eigenvalue problems arising in SciDAC applications







Technical Challenges:

- Improve convergence
- Exploit concurrency on leadership class machines
- Reduce communication overhead
- Reduce complexity
- Leverage highly efficient computational kernels (e.g. BLAS3)
- Maintain stability and reliability

Approaches:

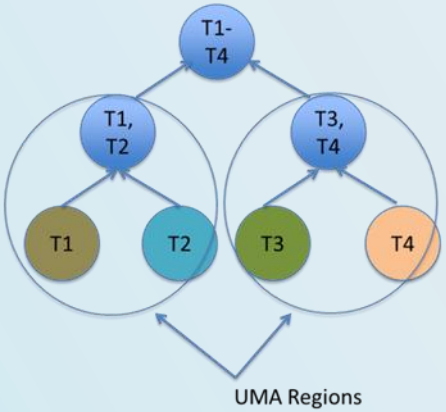
- Implicitly restarted Lanczos/Arnoldi iterations (PARPACK)
- Spectrum slicing using
  - Shift-invert
  - Contour integral projection

Linear System Solution in FASTMath Software

FASTMath Team Members: Jed Brown, Rob Falgout, Jonathan Hu, Sherri Li, Siva Rajamanickam, Jacob Schroder, Barry Smith, Ichitaro Yamazaki, Ulrike Yang

The efficient implementation of sparse linear solvers is crucial to enabling many large-scale simulations. The new generation of high performance computers with multi-/many core nodes and million-way parallelism presents new challenges for the linear solvers in the FASTMath software tools, PETSc, SuperLU, Trilinos and hypre.

- Multi-threading key matrix kernels
- Hiding communication in Krylov solvers



0	1	0	1	0	1	0	1	0	1	0	1
2	3	2	3	2	3	2	3	2	3	2	3
0	1	0	1	0	1	0	1	0	1	0	1
2	3	2	3	2	3	2	3	2	3	2	3
0	1	0	1	0	1	0	1	0	1	0	1
2	3	2	3	2	3	2	3	2	3	2	3

- Scheduling strategies and threading in sparse matrix factorizations
- Reducing communication in algebraic multigrid methods

More Information: http://www.fastmath-scidac.org or contact Ulrike Yang, LLNL, yang11@llnl.gov, 925-422-2850

SUPER's Green HPC: Improving the Energy-efficiency of DoE HPC

Laura Carrington (lead), Ananta Tiwari
UCSD/SDSC

Rob Fowler
RENCI

Bronis de Supinski, Barry Roundtree
LLNL

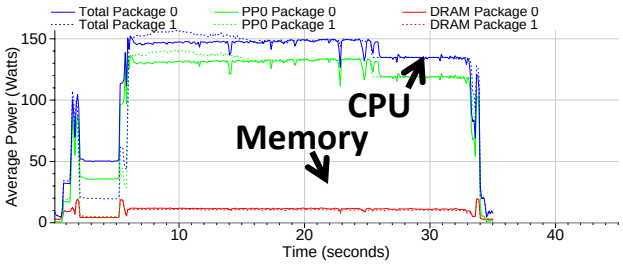
Dan Terpstra
UTK

SUPER's Energy Thrust: capture energy requirements of the different computation and communication patterns prevalent in HPC applications and use this for energy-efficient optimization strategies –

SUPER's Green HPC

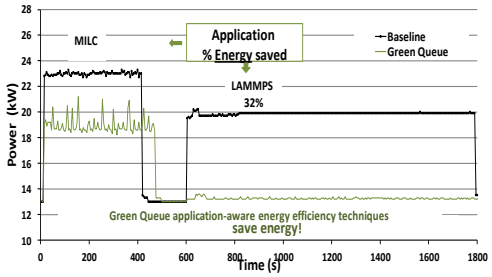
- Develop component level power measurement devices for HPC systems - PowerMon2 & more
- Develop an interface to access component level power measurements - PAPI
- Utilize the component level API and measurement device to research energy-efficient optimization strategies – Green Queue

PowerMon2 + PAPI



PLASMA Cholesky Factorization N=30,000 threads=16

Green Queue



Building a Resilience Autotuning Framework: SUPER's Integrated Resilience Strategy

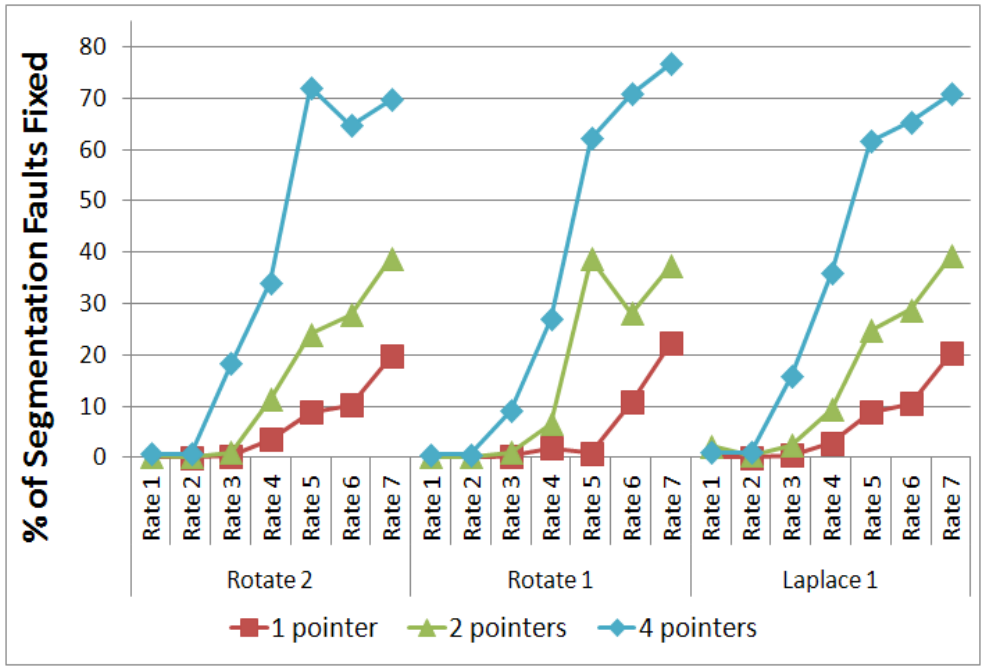
Bronis R. de Supinski, Marc Casas-Guix, Dan Quinlan, Chanua Liao
LLNL

Pedro Diniz, Bob Lucas
USC-ISI

Ganesh Gopalakrishnan
University of Utah

George Bosilca
University of Tennesse

- Mechanisms to assess application vulnerability
- Evaluate improvements from hand transformations
- ROSE translators to automate transformations
- User annotations to guide translators
- Autotune application of transformations
- Apply optimization techniques to balance performance and resilience



In three different experiments, pointer replication in AMG increasingly reduces the number of fatal segmentation faults.

SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

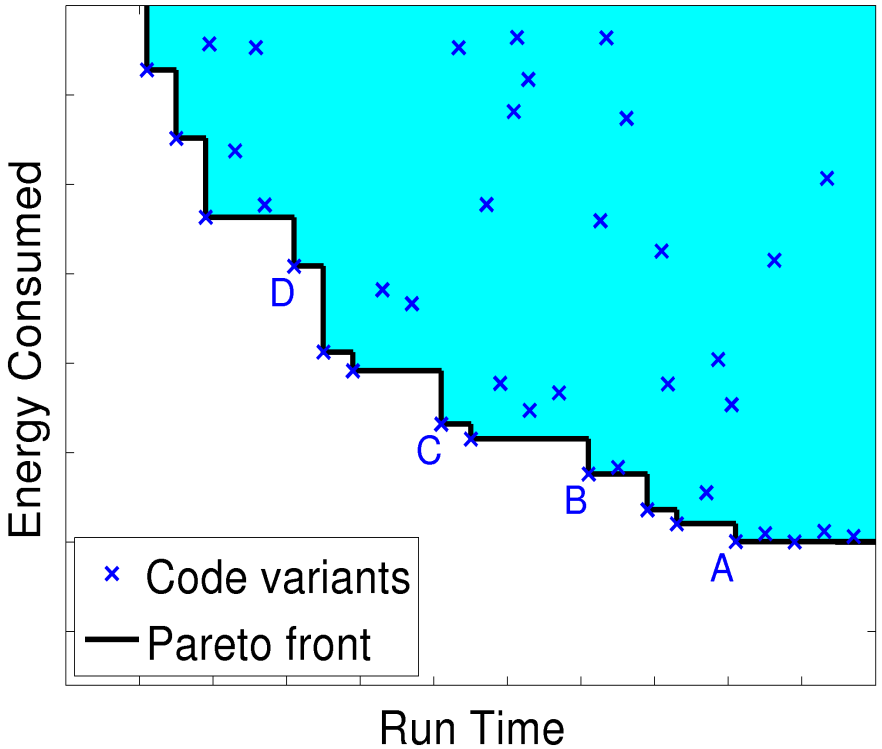
Tuning for Multiple Objectives: Power versus Performance

Prasanna Balaprakash, Paul Hovland (lead), Stefan Wild
Argonne National Laboratory

Jeff Hollingsworth
University of Maryland

Ananta Tiwari
San Diego Supercomputer Center

- Autotuning search problem = optimization problem
- Multiple objectives:
  - execution time
  - energy consumption
  - resilience to errors
  - power demands
  - memory footprint
- Multi-objective optimization problem
  - cannot pose as single objective problem
  - weights/constraints not known at search time
- Goal: *develop optimization framework for empirical search with multiple objectives*



Support for this work was provided through the Scientific Discovery through Advanced Computing (SciDAC) program funded by the U.S. Department of Energy Office of Advanced Scientific Computing Research.

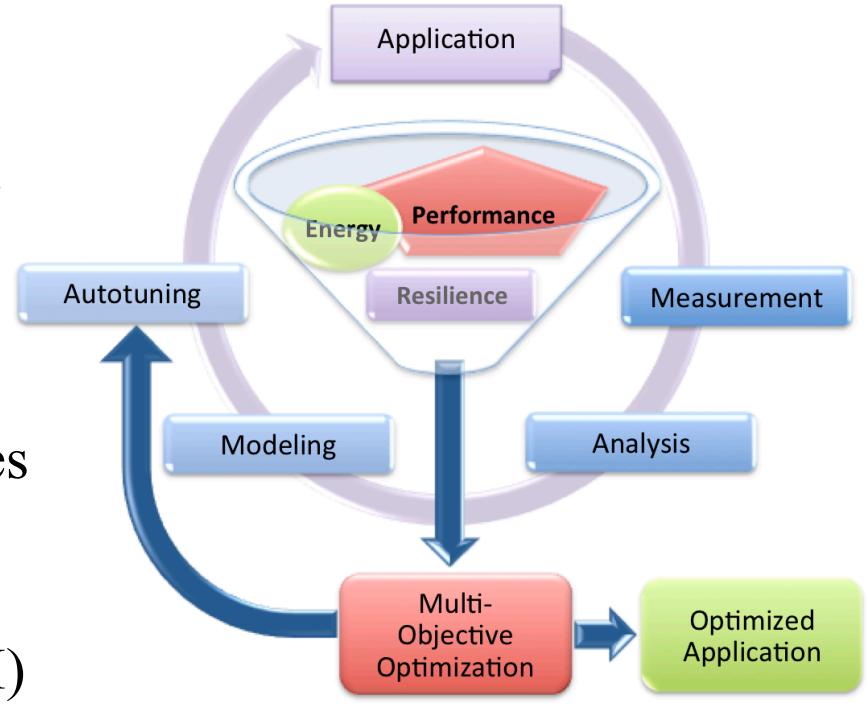
Tool Integration and Autotuning for SUPER Performance Optimization

SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Allen D. Malony, Nick Chaimov — University of Oregon
Mary Hall — University of Utah
Jeff Hollingsworth — University of Maryland
Boyana Norris — Argonne National Laboratory
Gabriel Marin — Oak Ridge National Laboratory

Methods for performance optimization of petascale applications must address the growing complexity of HPC environments.

Greater performance tool integration and tuning process automation are necessary.



SUPER is advancing autotuning capabilities through the coupling of:

- performance measurement, analysis, database, and triage tools (TAU, MIAMI)
- compiler and program translators (CHiLL, Orio)
- autotuning frameworks (Active Harmony, CHiLL, Orio)

The poster describes the ongoing integration of these tools in the SUPER ecosystem, and their use in performance optimization of SciDAC applications.

Support for this work was provided through the Scientific Discovery through Advanced Computing (SciDAC) program funded by the U.S. Department of Energy Office of Advanced Scientific Computing Research

Patrick Worley, Philip Roth
Oak Ridge National Laboratory
Kevin Huck, Scott Biersdorff, Chee Wai Lee, Allen Malony, Suzanne Millstein, Sameer Shende, Wyatt Spear
University of Oregon

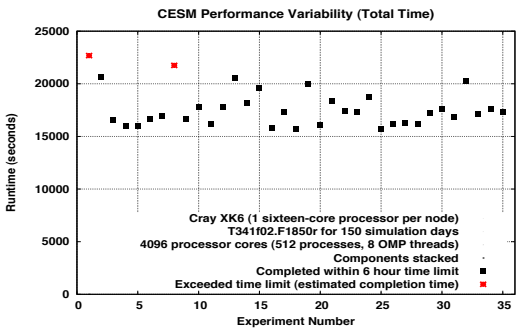
Tracking Performance

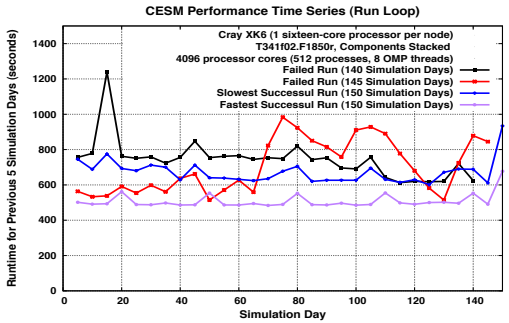
Anyone can do it

- simple low overhead techniques, suitable for production runs, available at all centers
- requires some (project-wide) discipline to save performance data and associated metadata

and everyone should

- for evaluating project throughput
- for identifying performance issues
- important mechanism for collaborating with SUPER





SUPER (and your SUPER liaison) is here to help, with

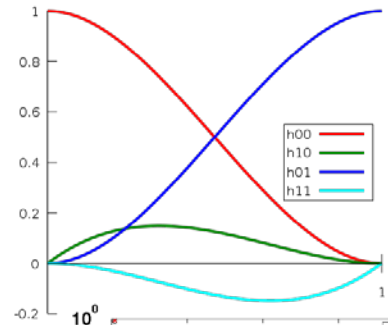
- advice on performance data capture methodology and tools
- sophisticated tools when have performance problems to address

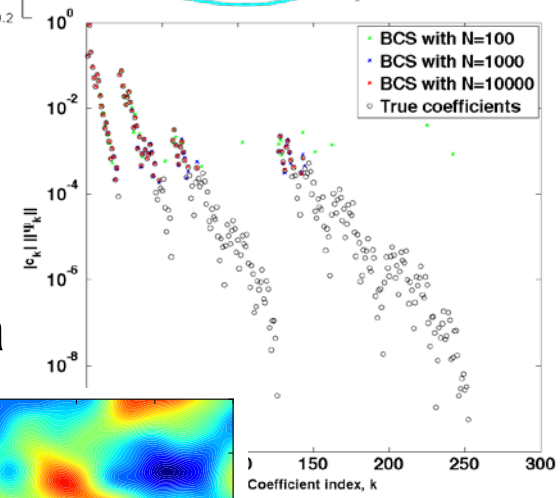
See poster for case study tracking performance for a suite of experiments, to determine nature of performance variability and to quantify impact.

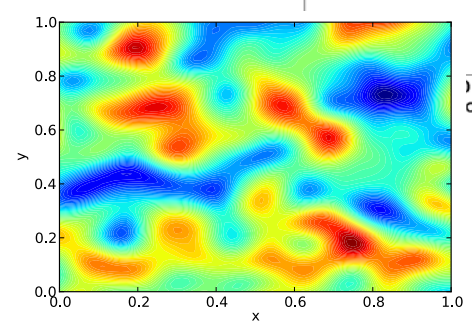
Algorithmic Developments for UQ in Extreme Scale Simulations

Bert Debusschere, Michael Eldred, John Jakeman, Khachik Sargsyan, Cosmin Safta, Jaideep Ray, Kenny Chowdhary, Habib Najm
Sandia National Laboratories
http://www.quest-scidac.org/
bjdebus@sandia.gov

- High Dimensionality:
  - Gradient-enhanced interpolants
  - Local error estimation with hierarchical value/gradient surpluses
  - Compressive sensing for sparsity detection
- Data characterization:
  - Data-Free Inference to handle missing data
  - Random field representations
- Predictive model calibration:
  - Subgrid model upscaling







Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

www.quest-scidac.org

QUEST Software Tools for Uncertainty Propagation and Inference

QUEST software tools support a range of:

- ***UQ studies:*** sensitivity analysis, uncertainty propagation, statistical inference
- ***Environments:*** rapid prototyping in interpreted languages ⟷ production computing in compiled languages on parallel platforms
- ***Intrusion:*** embedded ⟷ linked ⟷ black box

DAKOTA (SNL) is a production C++ application that provides non-intrusive algorithms for:

- design optimization
- model calibration
- uncertainty quantification
- global sensitivity analysis
- parameter studies
- solution verification

GPMSA (LANL) is a MATLAB toolbox that focuses on Bayesian inference using Gaussian process response surfaces to minimize the number of forward model calls required in the inference.

QUESO (UT) is an C++/MPI library that provides statistical algorithms for Bayesian inference, model calibration, model validation, and decision making under uncertainty. Its C++ abstractions enable the easy integration of new algorithms.

UQTk (SNL) is a library of C++ and Matlab functions for propagation of uncertainty through computational models

- Primarily relies on PCE for representing random variables and stochastic processes
- Targets rapid prototyping, algorithmic research, and outreach through tutorials and education

Reduced and Adapted Models for UQ

Evangelia Kalligiannaki, Ramakrishna Tipireddy, **Roger Ghanem** *
University of Southern California

Objective

- Constrain sochastic representations to specific domains in parameter space.
- Adapt stochastic models to low-dimensional quantities of interest.

Constrained Representations

Let $\mathcal{I}$ denotes a subset of $L_2(\Omega, \Sigma(H), P)$ defined by constraints.
$y(\omega) \in \mathcal{I}$ if $y(w)$ satisfies the constraints $\forall \omega$.

Given $y \in \mathcal{I}$ we construct

$$y(\xi) = \sum_{i=0}^{n} c_i \Psi_i(\xi)\,,$$

where
$c_i = \int_\Omega y(\xi(\omega))\Psi_i(\xi(\omega))dP(\omega) = \mathbb{E}_P[y\Psi_i]$,

Then the projection of y on $\mathcal{I}$ is given by

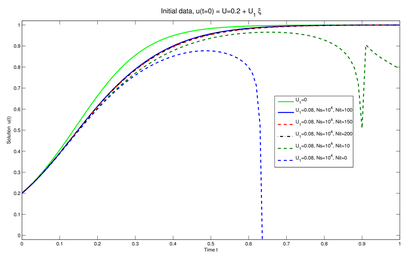
$$\tilde{y}(\xi) = \sum_{i=0}^{n} \tilde{c}_i \Psi_i(\xi)\,,$$

where
$\tilde{c}_i = \lim_{k\to\infty} \tilde{c}_i^{(k)}$
and
$\tilde{c}_i^{(k)} = \mathbb{E}[\chi_{A(\tilde{y})}\tilde{y}^{(k-1)}\Psi_i],\ k = 1, 2, \ldots$
$\tilde{y}^{(k)}(\xi) = \sum_{i=0}^{n} \tilde{c}_i^{(k)}\Psi_i(\xi)$
$\tilde{y}^{(0)} = y$
$A(\tilde{y}^{(k)}) := \{\omega \in \Omega : \tilde{y}^{(k)}(\omega) \in \mathcal{I}\}$
Ψ_i i^{th} Hermite polynomial.

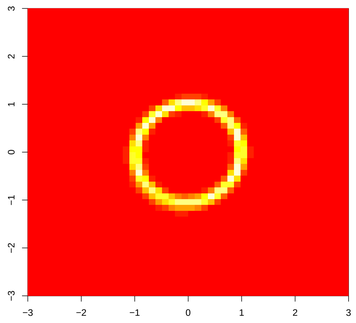
- $\tilde{y}$ is a fixed point of

$$G(y) = \sum_{i=1}^{n} \mathbb{E}_\xi[\chi_{A(y)}\Psi_i]\Psi_i\,.$$

- $\tilde{y}(\xi) \in \mathcal{I}$



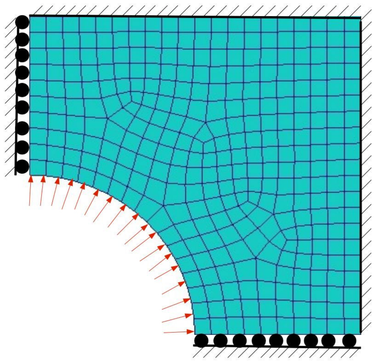
Divergence problems resolved (due to sampling unstable realizations) in solutions of stochastic ODEs.



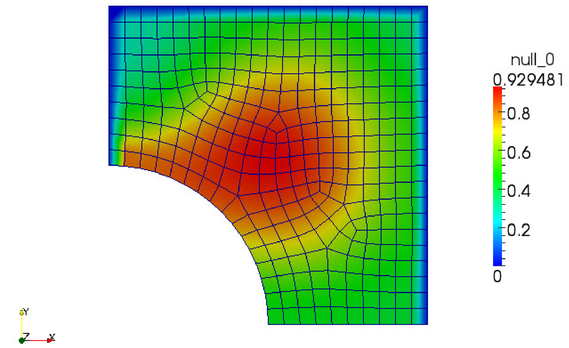
Restricting expansions to a ring in 2-dimensional parameter space.

Adaptation to QoI

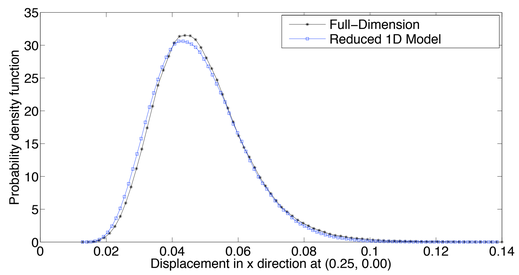
- Instead of expanding $u(\boldsymbol{\xi})$ along polynomials in $\boldsymbol{\xi}$, we expand it on polynomials in $\boldsymbol{\eta} = \boldsymbol{A}\boldsymbol{\xi}$ where $\boldsymbol{A}$ is an isometry.
- with proper choice of $\boldsymbol{A}$, the measure of the solution is concentrated along the dominant η_1 dimension.



Plane-Stress with Internal Pressure.



Spatial Variation of one coordinate of the Optimal 1D-Direction.



PDF at QOI with 10 parametric dimensions and 1-adapted dimension.

LANL UQ Interactions with other SciDAC Efforts

Collaboration with three SciDAC efforts:

- Nuclear Computational Low Energy Initiative;
- Plasma Surface Interactions: Bridging from the Surface to the Micron Frontier through Leadership Class Computing;
- Computation-Driven Discovery for the Dark Universe.

Common themes include:

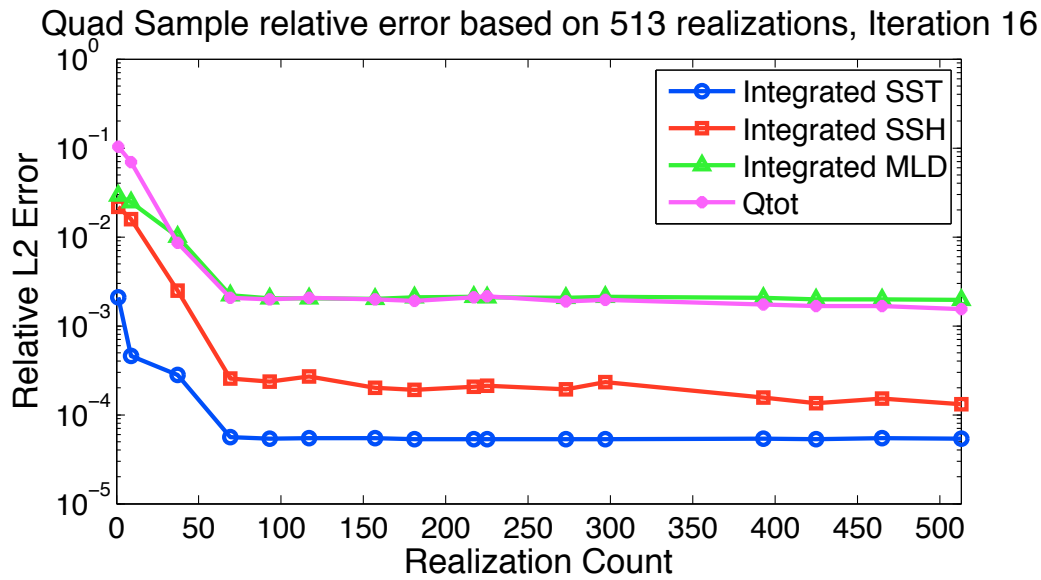
- Sensitivity analysis – understanding and quantifying the impact of model inputs on key model outputs;
- Model calibration – combining physical measurements with model runs to estimate/constrain important model parameters;
- Bridging scales – using a collection of models, or model fidelities, to bridge spatial (and/or temporal) scales for the physical processes of interest;
- Accounting for model discrepancy – modeling the difference between model predictions and the true physical process, producing more reliable uncertainties associated with model predictions.

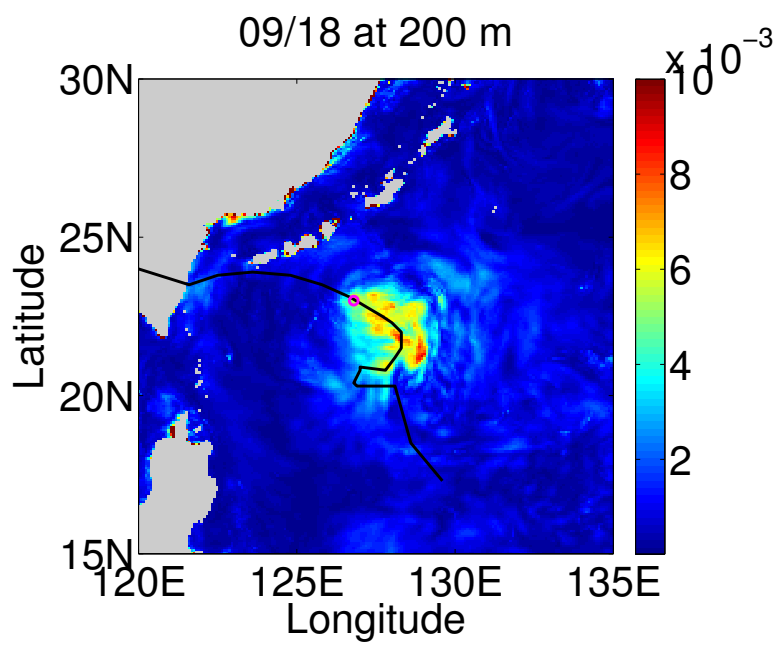
Poster 18 - Knio

Development and A Priori Testing of a Sparse Adaptive Pseudospectral PC Representation

J. Winokur, P. Conrad, I. Sraj, M. Iskandarani, Y. Marzouk, O. Knio

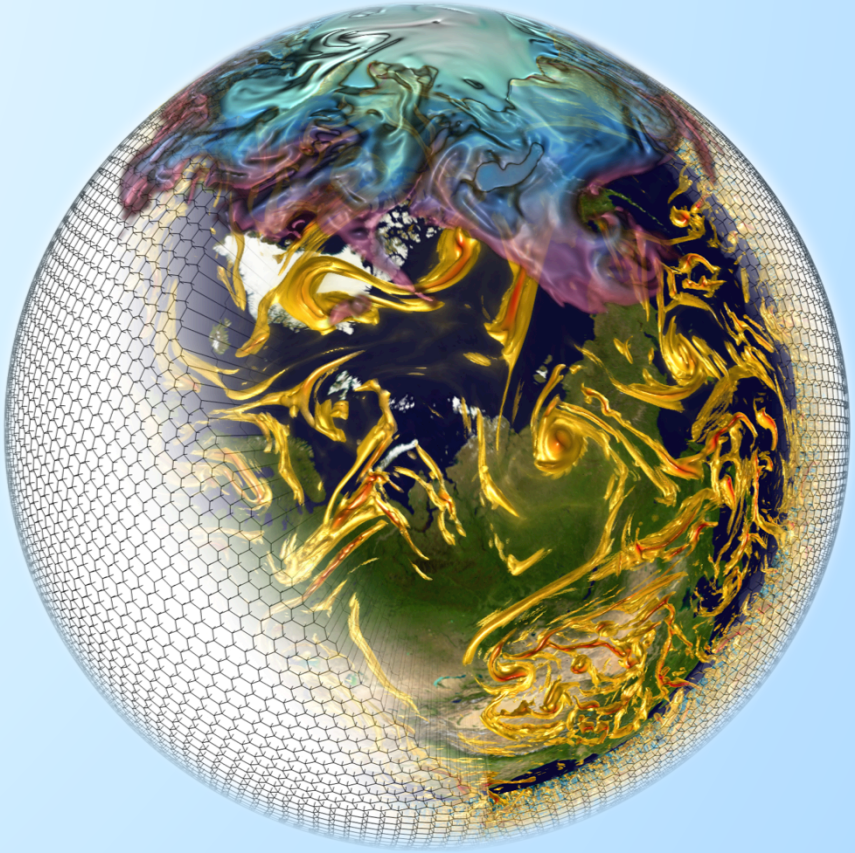
- We describe an adaptive sampling approach that
  - accommodates **arbitrary admissible sparse** grids
  - enables us to retain all polynomials that can be estimated **without internal aliasing**
- Assess the performance of the algorithm through a priori tests, based on existing OGCM databases
- Demonstrate dynamic implementation through extreme scale computations of the oceanic circulation in the Pacific Ocean





Technologies for Scientific Visualization

- **Delivering to the SciDAC community with production visualization tools**
  - VisIt and ParaView
- **Preparing for the near future with next-generation visualization technology**
  - In-situ visualization and analysis frameworks
  - New algorithms and languages
    - Efficient, Portable Performance on Next-Generation Architectures
  - Advanced/improved vector field visualization and rendering
  - Comparing scientific datasets and understanding uncertainty



Visualization of climate data on geodesic grid output by GCRM (Global Cloud Resolving Model) using a direct rendering solution. No pre-partitioning of the cells into tetrahedra is required, thus no storage overhead. The renderer is GPU accelerated.

Leads: James Ahrens, LANL, Wes Bethel, LBNL

SDAV
Scalable Data Management, Analysis, and Visualization

Recent Achievements in Visual Data Exploration and Analysis for Climate Science

Challenges:

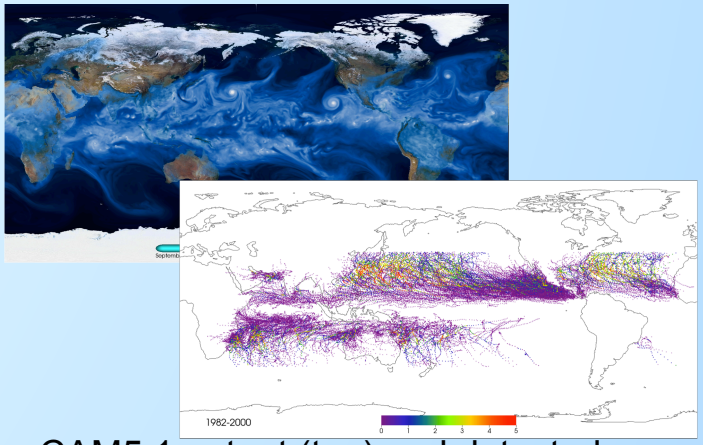
- Increasing model complexity, output data size confounds legacy software tools and existing approaches for analysis and visualization.

Approach:

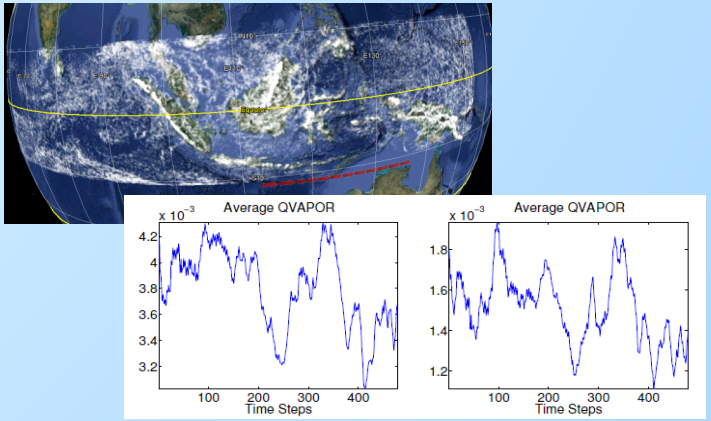
- Applied R&D, productization, leverage existing investments to maximize productivity.
- Close partnerships with climate science projects.

Impacts:

- New generation of HPC vis/analysis tools that accommodate larger, more complex data; more focused analysis; run on modern computational platforms.



CAM5.1 output (top) and detected cyclone tracks (bottom)



Cloud rendering on a virtual globe (top), average water vapor mixing rate (bottom)

Supporting Combustion Research with Visualization, Analysis and Data Movement

Combustion Scientists

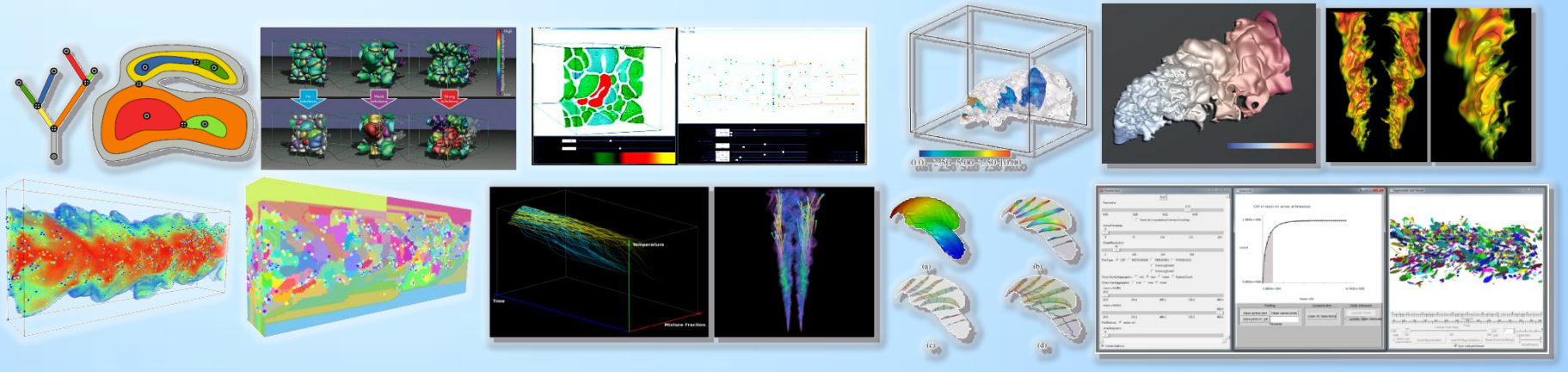
Dr. Jackie Chen at SNL, Dr. John Bell at LBNL

Goal: understanding turbulent combustion

- Combustion provides 85% of the US energy needs
- Low emission/temperature engines will operate at combustion regimes currently poorly understood
- Need new understanding of turbulence-chemistry to predict efficiency and pollutant emissions

Problem and Challenges

- Feature definitions require inspection of data
- Massive datasets are hard to manage
- Selection parameters are not predetermined
- Intermittent phenomenon



Data Movement

- Improving I/O performance with ADIOS
- Hybrid *in-situ/in-transit* framework

Analysis Techniques

- Topological approaches
- Time tracking
- Multi-field analysis
- Particle tracing
- Flame parameterization
- Shape characterization

Interactive Visualization

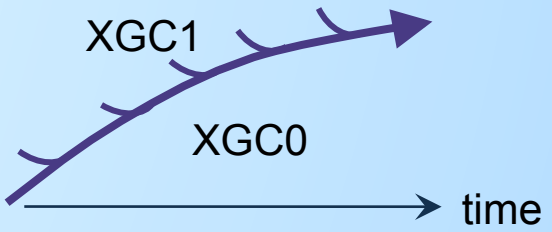
- Volume visualization
- Feature based visualization
- Parameter studies
- Feature-based statistics

SDAV
Scalable Data Management, Analysis, and Visualization

SDAV Technologies for the Next Generation Fusion Techniques

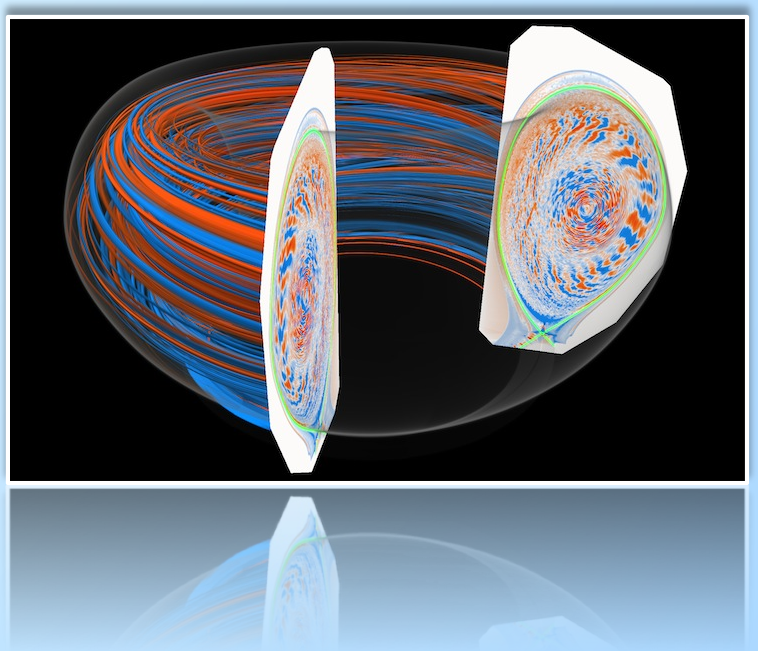
Problems:

- Need for a Collaborative Data Management, Analysis and Visualization Framework for petascale fusion simulations
- Scientific insights are hiding inside **too much data**
- Long term goal is a tight kinetic-kinetic coupling in XGC code, to extend the **first-principles** full-f simulation **to experimental time scale**



Solutions:

- ADIOS I/O to handle large datasets
- FastQuery, Range based queries to find/read data quickly
- EFFIS Framework for data management, simulation monitoring, analysis, code coupling
- New visualization techniques to emphasize the important details
- In Situ / In Transit technology to couple codes and perform in situ visualizations



SDAV
Scalable Data Management, Analysis, and Visualization

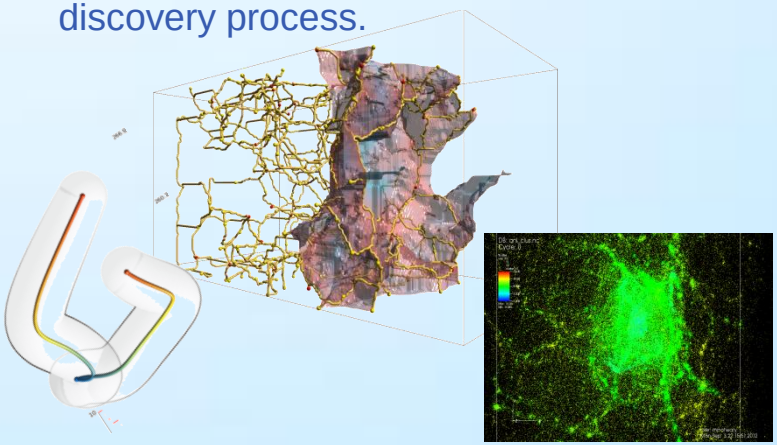
Advanced Data Analysis Techniques for Science Discovery

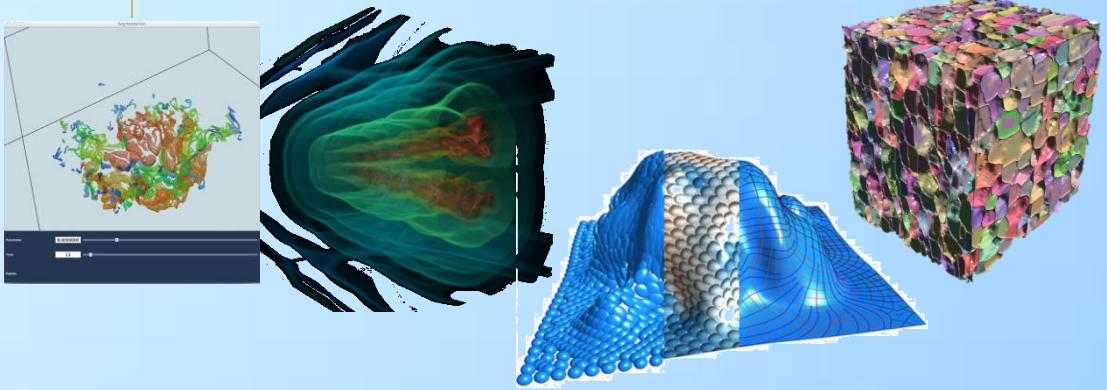
Problem:

- The evolution of modern science has been characterized by an exponential increase in use and production of data generated by simulations, experiments and sensing devices.
- Data analysis tools have become an essential component of the scientific investigation and discovery process.

Approach:

- SDAV serves a broad set of science needs with the deployment of new advanced techniques in diverse fields such as statistics, data mining, and topology.
- The tools are deployed in advanced software components that deal efficiently with massive data sets and provide effective solutions for the challenges of modern data intensive science.





- TALASS: Feature-based analysis and tracking
- MSEXPLORER: Structural topological analysis
- High dimensional data analysis and exploration

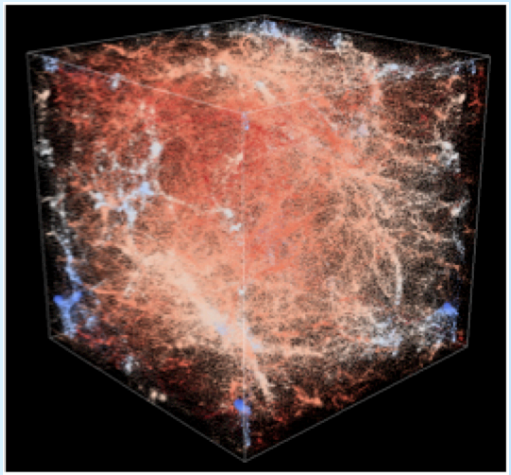
- STPMiner: A spatiotemporal pattern miner
- Parallel distance field computations
- Dual space analysis of particle data

- DBSCAN: Parallel density based data clustering algorithm
- Metric tensor flow surface analysis
- Streaming, feature-based statistics
- Parallel computation of Finite-Time Lyapunov Exponent (FTLE)

SDAV
Scalable Data Management, Analysis, and Visualization

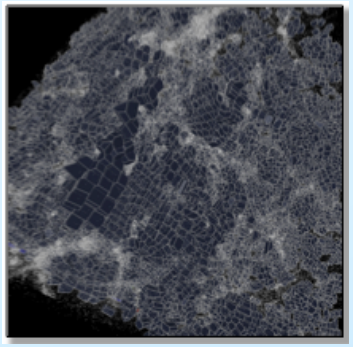
Visual and Numerical Data Analysis of Large-Scale Cosmology and Astrophysics Simulations

The identification of large-scale cosmological structures enables simulated matter tracer particles to be correlated with telescope surveys, inform how gravity drives the distribution of galaxies in the universe, and infer properties of dark energy and dark matter.



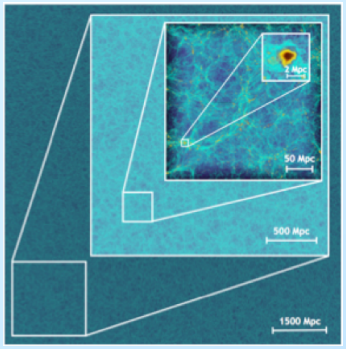
Halos, voids, filaments, and walls are visible in this simulation. Our goal is to analyze, visualize, and store such structures efficiently at scale. (Image courtesy Jay Takle)

In Situ Data Analysis



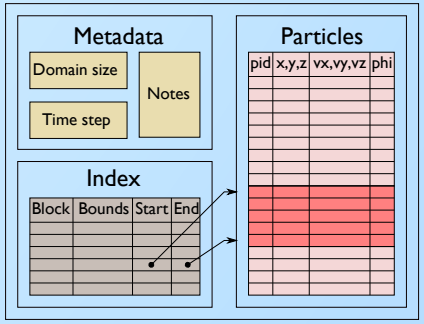
In situ Voronoi tessellation derived from tracer particle distribution.

Visualization



Visualization of HACC simulation using vl3.

Data Management



Schema for analysis output enables spatial querying in parallel netCDF.

A set of in situ, coprocessing, and postprocessing tools for scientists to locate, track, and evaluate the evolution of large-scale structures.

SDAV
Scalable Data Management, Analysis, and Visualization

Yes, there are more SDAV posters!

Managing data movement and orchestrating data analysis in large-scale systems is a complex and critical component of effective computational science.

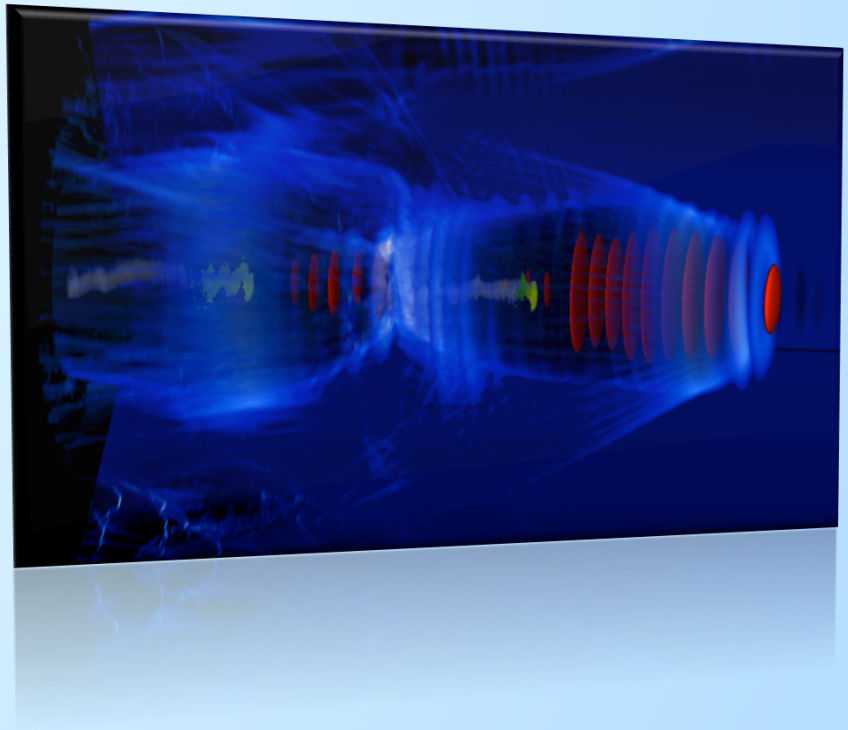
Technologies for Extreme-Scale Data Management covers:

- Frameworks for data movement and in situ and co-analysis
- Software for indexing, reorganizing, and compressing scientific datasets
- Tools for observing the I/O behavior of production simulation runs

Scalable Data Management, Analysis and Visualization of Particle Accelerator Simulation Data

Problem:

- Accurate simulation of particle acceleration processes requires the use of $>10^8$ up to 10^{12} particles, e.g., to control particle shot noise and to avoid overestimation of the microbunching instability.
- Simulation data is extremely large, of varying resolution, and heterogeneous.
- Particle features of interest are often small.
- The available I/O bandwidth and data storage capabilities are decreasing relative to computation.



Solution:

- Developed FastQuery, a distributed index/query system, making state-of-the art index/query technology based on FastBit accessible to scientific file formats such as HDF5.
- Integrate parallel index/query capabilities with the parallel visualization system VisIt.
- Develop methods for automatic detection of particle beams.
- Couple simulation with general purpose, state-of-the-art *in situ* technology using VisIt.

Thank You!

Poster Session

4:30-6:00pm in Roosevelt/Madison