SciDAC-3 Institute for Sustained Performance, Energy, and Resilience

http://super-scidac.org

Bob Lucas, PI
University of Southern California

Argonne PI: *Boyana Norris*
norris@mcs.anl.gov

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 Science, Advanced Scientific Computing Research

SciDAC Performance Efforts

- SciDAC has always had an effort focused on performance
- Performance Evaluation Research Center (PERC)
  - Benchmarking, modeling, and understanding
- Performance Engineering Research Institute (PERI)
  - Performance engineering, modeling, and engagement
  - Three SciDAC-e projects
- Institute for Sustained Performance, Energy, and Resilience (SUPER)
  - Performance engineering
  - Energy minimization
  - Resilient applications
  - Optimization of the above

SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Outline

SUPER Team

Research Directions

Broader Engagement

SUPER

SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

SUPER Team

ANL Paul Hovland Boyana Norris Stefan Wild Argonne National Laboratory	**LBNL** David Bailey Lenny Oliker Sam Williams Berkeley Lab	**LLNL** Bronis de Supinski Daniel Quinlan Lawrence Livermore National Laboratory	**Oregon** Allen Malony Sameer Shende University of Oregon	**ORNL** Gabriel Marin Philip Roth Patrick Worley Oak Ridge National Laboratory	**UCSD** Laura Carrington Ananta Tiwari UCSD
UMD Jeffrey Hollingsworth University of Maryland 1856	**UNC** Rob Fowler Allan Porterfield Renci Renaissance Computing Institute	**USC** Jacque Chame Pedro Diniz Bob Lucas (PI) USC University of Southern California	**UTEP** Shirley Moore UTEP	**UTK** Dan Terpstra ICL UT	**Utah** Mary Hall The University of Utah

SUPER

Broadly Based Effort

- **All PIs have independent research projects**
  - SUPER money alone isn't enough to support any of its investigators
  - SUPER leverages other work and funding

- **SUPER contribution is integration, results beyond any one group**
  - Follows successful PERI model (tiger teams and autotuning)
  - Collaboration extends to others having similar research goals
    - John Mellor-Crummey at Rice is an active collaborator
    - Other likely collaborators include LANL, PNNL, Portland State, and UT San Antonio
    - Perhaps Juelich and Barcelona too?

Outline

SUPER Team

Research Directions

Broader Engagement

SUPER Research Objectives

INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Performance Optimization of DOE Applications

- Automatic performance tuning
- Energy minimization
- Resilient computing
- Optimization of the above

Collaboration with SciDAC-3 Institutes

Engagement with SciDAC-3 Applications

SUPER

Performance Engineering Research

- **Measurement and monitoring**
  - Adopting University of Oregon's TAU system
  - Building on UTK's PAPI measurement library
  - Also collaborating with Rice and its HPCToolkit
- **Performance Database**
  - Extending TAUdb to enable online collection and analysis
- **Performance modeling**
  - PBound and Roofline models to bound performance expectations
  - MIAMI to model impact of architectural variation
  - PSINS to model communication

INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

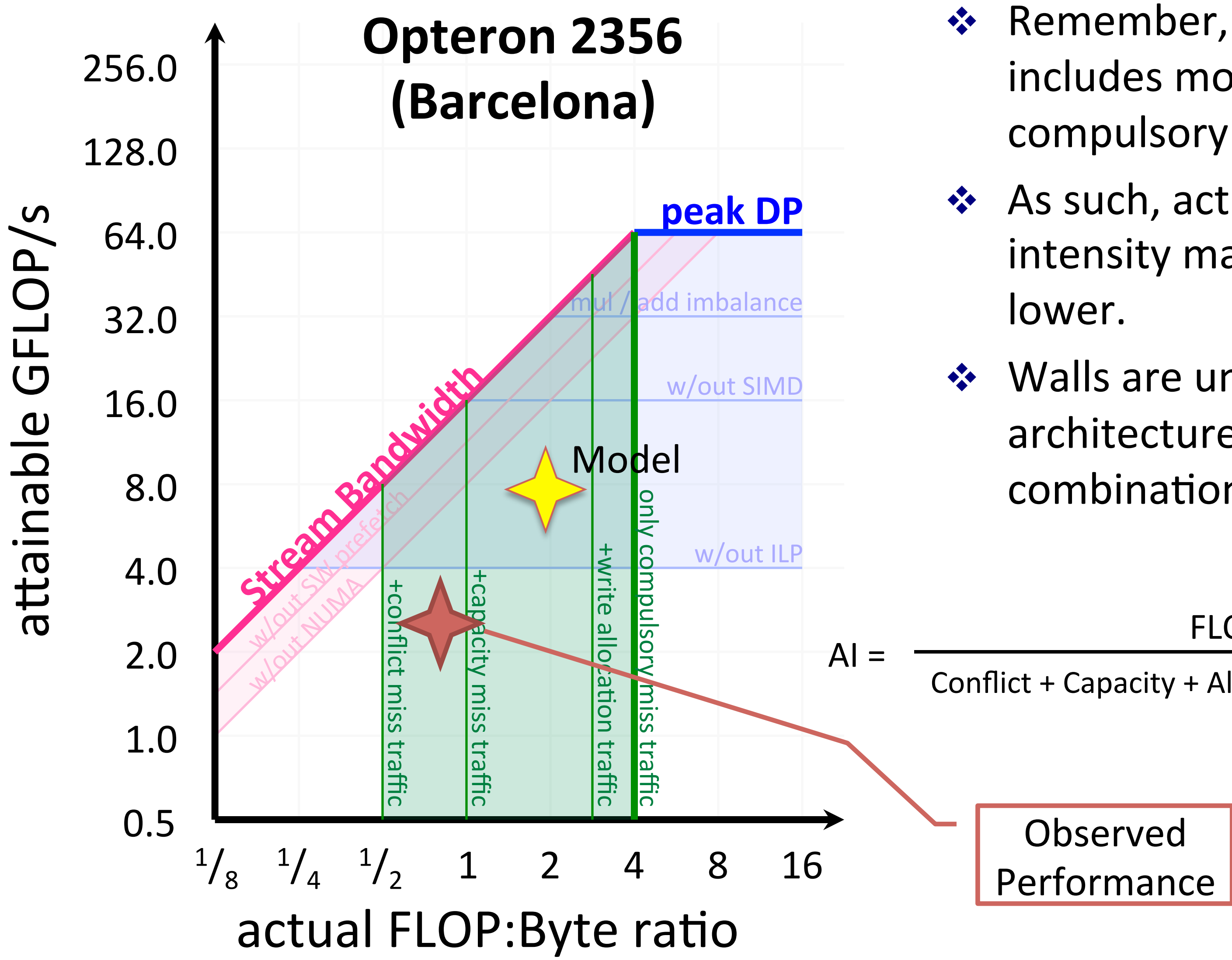
Generating Performance Models with PBound

- Use source code analysis to generate performance bounds (best/worst case scenarios)
- Can be used for
  - Understanding performance problems on current architectures
    - For example, when presented in the context of roofline models (introduced by Sam Williams, LBNL)
  - Projecting performance to hypothetical architectures

SUPER

Example Roofline Model

locality walls



- Remember, memory traffic includes more than just compulsory misses.
- As such, actual arithmetic intensity may be substantially lower.
- Walls are unique to the architecture-kernel combination

$$AI = \frac{\text{FLOPs}}{\text{Conflict + Capacity + Allocations + Compulsory}}$$

Sample PBound Output

```
void axpy4(int n, double *y, double a1, double *x1, double a2,
double *x2, double a3, double *x3, double a4, double *x4)
{
  register int i;
  for (i=0; i<=n-1; i++)
    y[i]=y[i]+a1*x1[i]+a2*x2[i]+a3*x3[i]+a4*x4[i];
}
```

```
#include "pbound_list.h"
void axpy4(int n, double *y, double a1, double *x1, double a2,
double *x2, double a3, double *x3, double a4, double *x4)
{
#ifdef pbound_log
  pboundLogInsert("axpy.c@6@5",2,0,5 * ((n - 1) + 1) + 4,     1 *
((n - 1) + 1),3 * ((n - 1) + 1) + 1,4 * ((n - 1) + 1));
#endif
}
```

SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Autotuning

- Led by Mary Hall, University of Utah
- Extend PERI autotuning system for future architectures
  - New TAU front-end for triage
  - CUDA-CHiLL, Orio to target GPUs
  - OpenMP-CHiLL for SMP multicores
  - Active Harmony provides search engine
    - Drive empirical autotuning experimentation
    - Balance threads and MPI ranks in hybrids of OpenMP and MPI
    - Extend to surface/area to volume, or halo size, experiments
  - Targeted Autotuning
    - Domain-specific languages
    - Users write simple code and leave the tuning to us
  - Whole program autotuning
    - Parameters, algorithm choice, libraries linked, etc.

SUPER

Autotuning PETSc Solvers on GPUs with Orio
B. Norris and A. Mametjanov, Argonne

SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

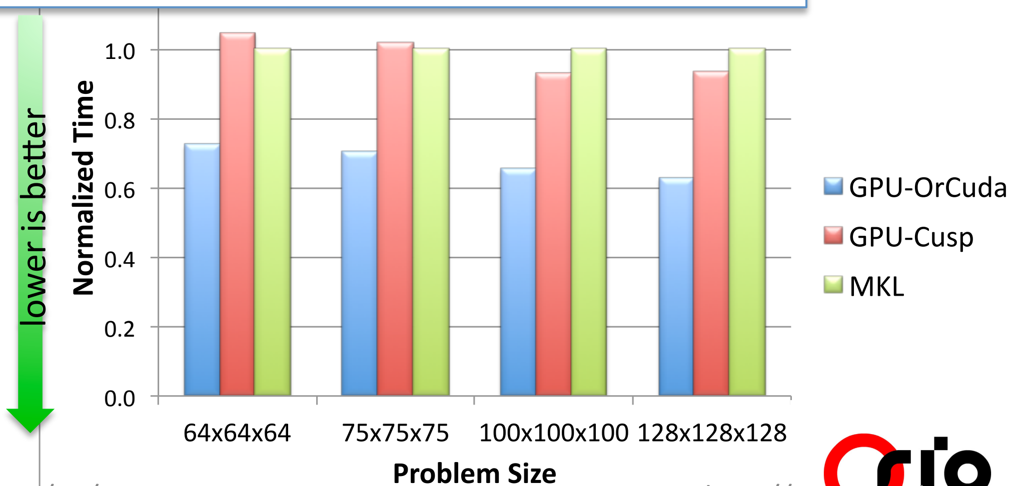
Objectives

- Generate high-performance implementations sparse matrix algebra kernels defined using a high-level syntax
    - Optimize performance automatically on different architectures, including cache-based multicore CPUs and heterogeneous CPU/GPU systems
    - Integrate autotuning into the Portable Extensible Toolkit for Scientific Computing (PETSc)
- Efficiently search the space of possible optimizations

Impact

- Optimizations for multicore CPUs and GPUs from the same high-level input.
- Improved code readability and maintainability.
- Performance of autotuned kernels typically exceeds that of tuned vendor numerical library implementations.

Example: Autotuning a structured-grid PDE application using PETSc for GPUs (solid fuel ignition).



Intel Xeon 2.8 GHz E5462 (8 cores), 2.8GHz; GPU: NVIDIA Fermi C2070

Accomplishments FY2012

- Developed new transformations for sparse matrix operations targeting GPUs.
- Autotuning resulted in improvements ranging from **1.5x** to **2x** over tuned vendor libraries on different GPUs for a solid fuel ignition PETSc application.
- "Autotuning stencil-based computations on GPUs." A. Mametjanov, D. Lowell, C.-C. Ma, and B. Norris. *Proceedings of IEEE Cluster 2012*.

Orio
http://tinyurl.com/OrioTool

SUPER

SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Autotuning Framework

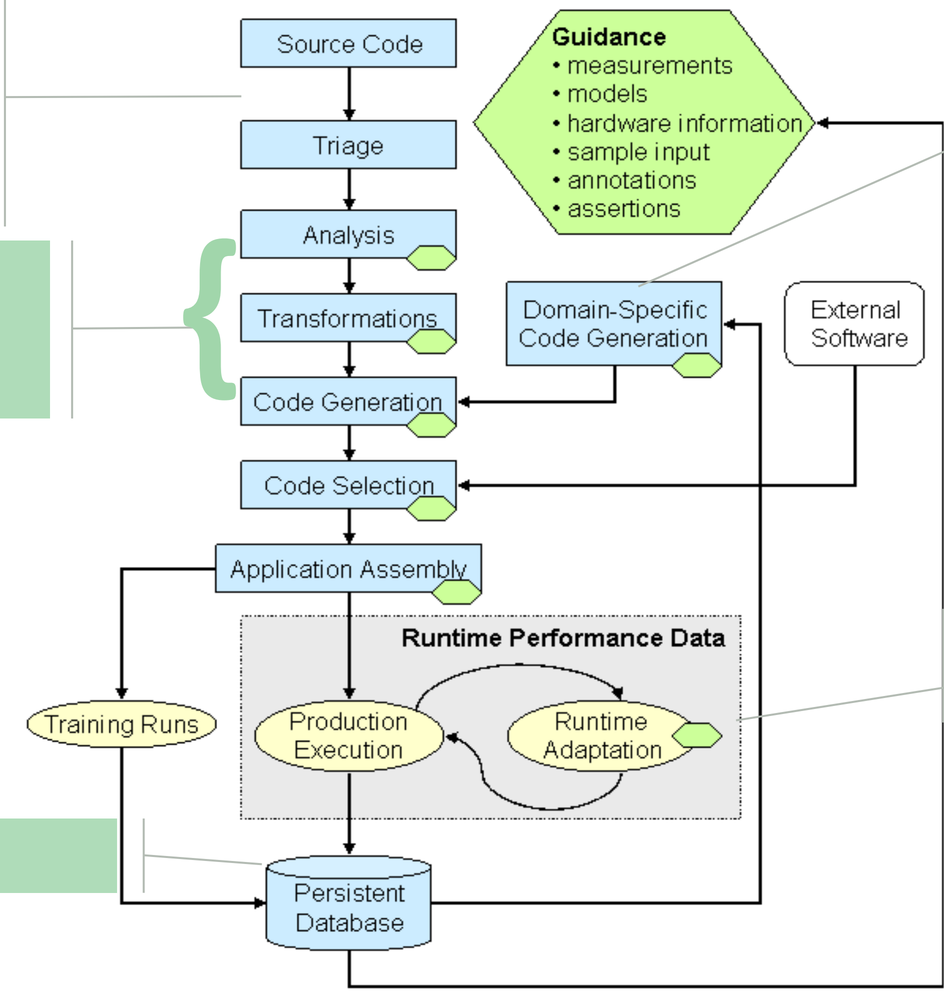
TAU (Oregon)
HPC Toolkit (Rice)
ROSE (LLNL)
MIAMI (ORNL)
PAPI (UTK)

CHiLL (USC/ISI and Utah)
ROSE (LLNL)
Orio (Argonne)

OSKI (LBNL)

Active Harmony (UMD)
GCO (UTK)

PerfTrack (LBNL, SDSC, RENCI)



SUPER

Energy Minimization

- Led by Laura Carrington, University of California at San Diego
- Develop new energy aware APIs for users
  - I know the processor on the critical path in my multifrontal code
- Obtain more precise data regarding energy consumption
  - Extend PAPI to sample hardware power monitors
  - Build new generation of PowerMon devices
  - Extend performance models
- Transform codes to minimize energy consumption
- Inform systems to allow them to exploit DVFS

Green Queue: Customized Large-scale Clock Frequency Scaling (Carrington and Tiwari)

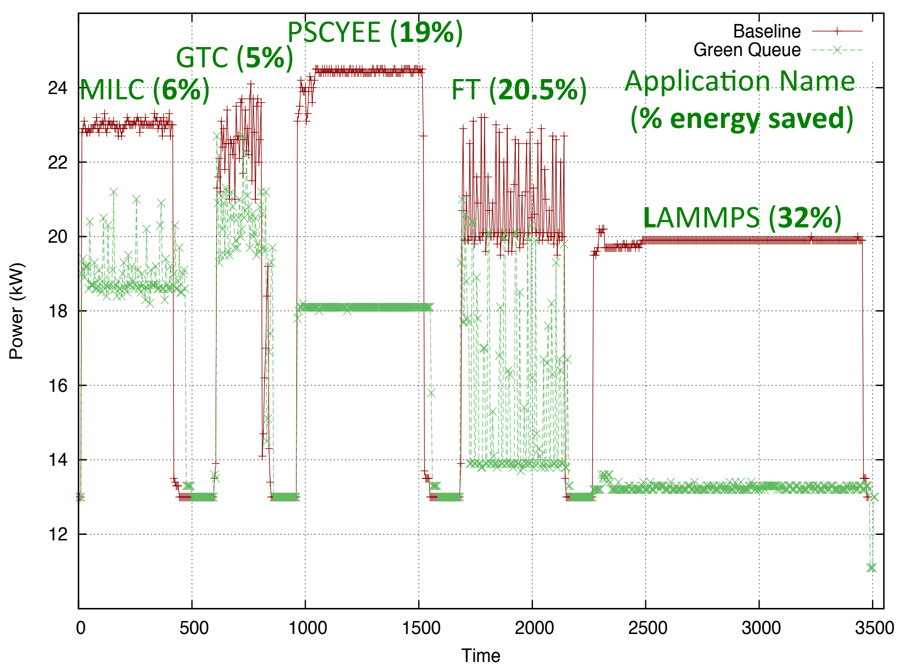
SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Overall Objectives

- **The *power wall*** is one of the major factors impeding the development and deployment of *exascale* systems
- Solutions are needed that reduce energy from both the hardware and software side
- **Solutions must be easy to use, fully automated, scalable and application-aware**

Research Plan

- Develop static and runtime application analysis tools to quantify application phases that affect both system-wide power draw and performance
- Use the properties of these application phases to develop models of the power draw
- Combine power and performance models to create phase-specific Dynamic Voltage-Frequency Scaling (DVFS) settings



Power monitoring during application runs.
Numbers in parenthesis show energy savings.

Progress & Accomplishments

- **Developed Green Queue, a framework that can make CPU clock frequency changes based on intra-node (e.g., memory latency bound) and inter-node (e.g., load imbalance) aspects of an application**
- **Deployed Green Queue on SDSC's Gordon supercomputer using one full rack (1024 cores) and a set scientific applications**
- **Green Queue shown to reduce the operational energy required for HPC applications by an average of 12% (maximum 32%)**
- **Deployed a test system for SUPER team members to facilitate collaboration on energy efficiency research and tool integration**

SUPER

SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Resilient Computing

- Led by Bronis de Supinski, Livermore National Laboratory
- Investigate directive-based API for users
  - Enable users to express their knowledge w/r resilience
  - Not all faults are fatal errors
  - Those that can't be tolerated can often be ameliorated
- Automating vulnerability assessment
  - Modeled on success of PERI autotuning
  - Conduct fault injection experiments
  - Determine which code regions or data structures fail catastrophically
  - Determine what transformations enable them to survive
- In either event, extend ROSE compiler to implement transformations

SUPER

INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Fault Resilience of an Algebraic Multi-Grid (AMG) Solver

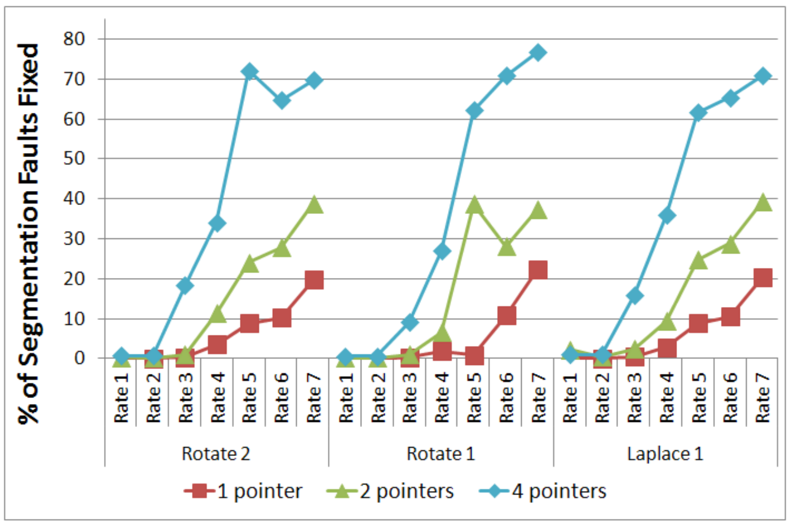
(Casas Guix and de Supinski)

General AMG Features

- AMG solves linear systems of equations derived from the discretization of partial differential equations.
- AMG is an iterative method that operates on nested grids of varying refinement.
- Two operators (restriction and interpolation) propagate the linear system through the different grids.

Goals of the SUPER resilience effort

- Identify vulnerable data and code regions.
- Design and implement simple and effective resilience strategies to improve vulnerability of sensitive pieces of code.
- Long term: develop a general methodology to automatically improve the reliability of generic HPC codes.



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

Progress & Accomplishments

- **Developed of a methodology to automatically inject faults to assess the vulnerability of codes to soft errors.**
- **Performed a vulnerability study of AMG.**
- **Determined that AMG is most vulnerable to soft errors in pointer arithmetic, which lead to fatal segmentation faults.**
- **Demonstrated that triple modular redundancy in pointer calculations reduces the vulnerability of AMG to soft errors**
- **Presented at ICS 2012**

SUPER

Optimization

Led by Paul Hovland, Argonne National Laboratory

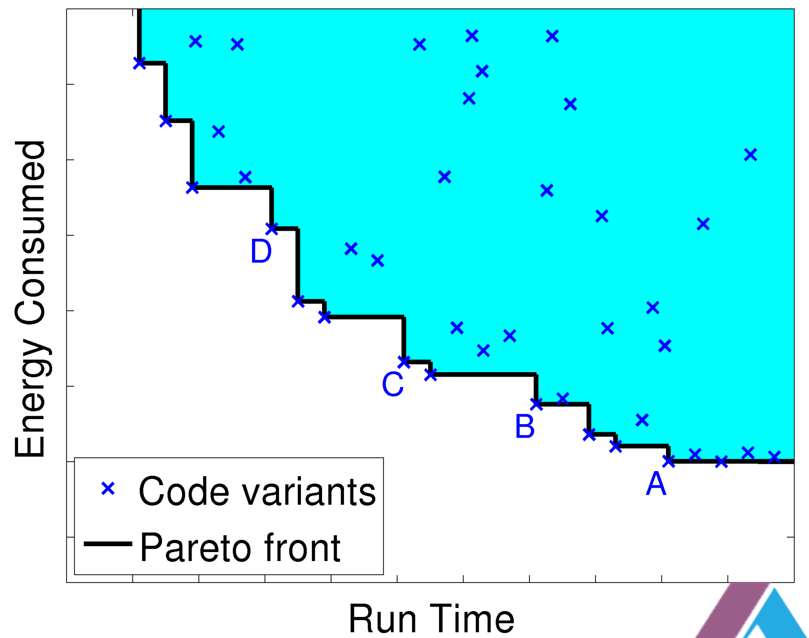
Performance, energy, and resilience are implicitly related and require *simultaneous* optimization

- E.g., Processor pairing covers soft errors, but halves throughput

Results in a stochastic, mixed integer, nonlinear, multi-objective, optimization problem

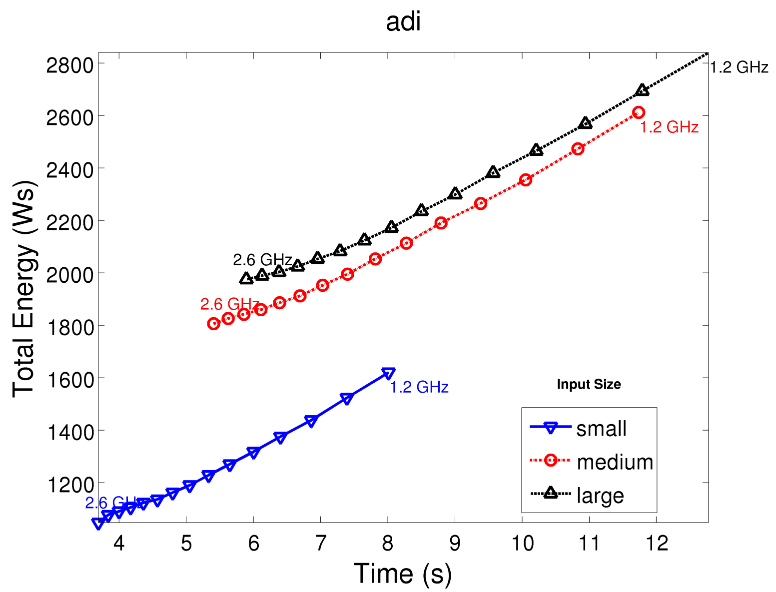
Only sample small portion of search space:

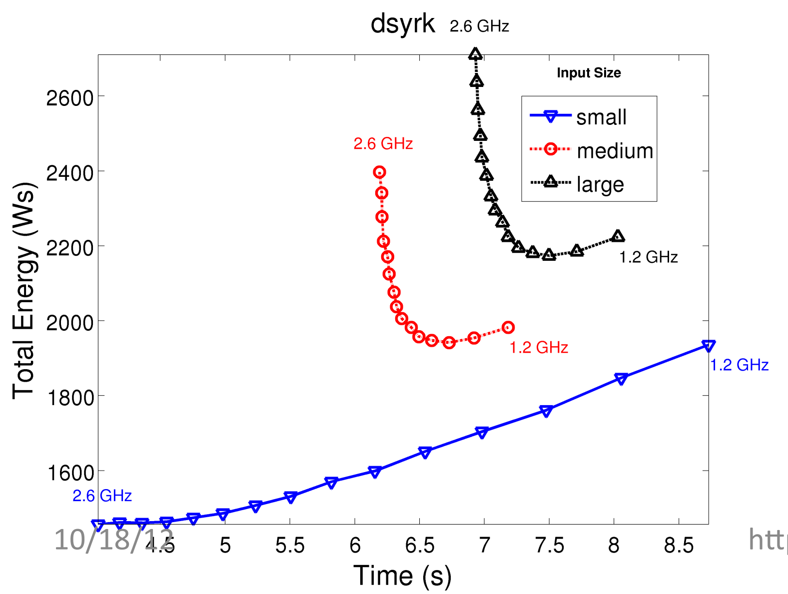
- Requires efficient derivative-free numerical optimization algorithms
- Need to adapt algorithms from continuous to discrete autotuning domain



SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Optimization: Time vs. Energy





- Conventional wisdom suggests that best strategy to minimize energy is to minimize execution time
- However, in practice, real tradeoffs between time and energy are observed
  - Depends on computation being performed
  - Depends on problem size

SUPER

Outline

SUPER Team

Research Directions

Broader Engagement

SUPER

SciDAC-3 Application Partnerships

- Collaboration with SciDAC Application Partnerships is expected
  - Yet SUPER funding is spread very thin
- SUPER investigators included in 12 Application Partnerships
  - Our time costs money, like everybody else
- Principles used to arrange teams
  - Technical needs of the proposal
  - Familiarity of the people
  - Proximity

SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Application Engagement

- Led by Pat Worley, Oak Ridge National Laboratory
- PERI strategy: proactively identify application collaborators
  - Based on comprehensive application survey at beginning of SciDAC-2
  - Exploited proximity and long-term relationships
- SUPER strategy: broaden our reach
  - Key is partnering with staff at ALCF, OLCF, and NERSC
  - Augment TAUdb to capture data from applications from centers
  - Initial interactions between Oregon and ORNL with and LCFs
  - Collaborate with other SciDAC-3 investigators
- Focused engagement as requested by DOE

SUPER

SUPER

INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

SUPER Summary

- **Research components**
  - Automatic performance tuning
    - New focus on portability
  - Addressing the "known unknowns"
    - Energy minimization
    - Resilient computing
  - Optimization of the above

- **Near-term impact on DOE computational science applications**
  - Application engagement coordinated with ALCF, NLCF, and NERSC
  - Tool integration, making research artifacts more approachable
  - Participation in SciDAC-3 Application Partnerships
  - Outreach and tutorials

SUPER