INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

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

Bob Lucas
University of Southern California
July 31, 2014

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

SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Outline

Institute Awareness

Architecture Awareness

Application Awareness

Management Strategy

Science Pipeline

SUPER

Intra-Institute Awareness

- SciDAC has always had a performance focus
- Performance Evaluation Research Center (PERC)
  - Performance measurement, modeling and benchmarking
  - Making a science of performance understanding
- Performance Engineering Research Institute (PERI)
  - Automatic performance tuning
- Institute for Sustained Performance, Energy, and Resilience
  - Heterogeneous systems
  - Energy
  - Resilience
  - Multi-objective Optimization

SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

SUPER Team

Paul Hovland K. Narayanan Stefan Wild Argonne National Laboratory	Lenny Oliker Sam Williams Berkeley Lab	B. de Supinski Todd Gamblin Chun Leo Liao Kathryn Mohror Daniel Quinlan Lawrence Livermore National Laboratory	Kevin Huck Allen Malony Boyana Norris Sameer Shende University of Oregon	Eduardo D'Azevedo Philip Roth Patrick Worley Oak Ridge National Laboratory	Laura Carrington Ananta Tiwari UCSD
Rey Chen J. Hollingsworth Sukhyun Song University of Maryland 1856	Rob Fowler Anirban Mandal Allan Porterfield Raul Ruth Renci Renaissance Computing Institute	Jacque Chame Pedro Diniz Bob Lucas (PI) USC University of Southern California	Shirley Moore UTEP	George Bosilca Anthony Danalis Jack Dongarra Heike McCraw ICL UT	G. Gopalakrishnan Mary Hall The University of Utah

SUPER

SUPER Expertise

Autotuning

Active Harmony, CHiLL, GCO, Orio, ROSE

Measurement of performance and energy

Dyninst, mpiP, PAPI, PEBIL, pcubed, RCRToolkit, TAU

Modeling of performance and energy

Green Queue, PMAC, Roofline

Parallel Programming APIs

OpenMP, OpenMPI

Resilience

ROSE

Multi-Objective Optimization

TAO

INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Broadly Based Effort

All PIs have independent research projects

- SUPER alone isn't enough to support any of its investigators
- SUPER leverages other work and funding, our science pipeline

SUPER contribution is integration, results beyond any one group

- Follows SciDAC-2 PERI model (tiger teams and autotuning)
- Collaboration extends to others having similar research goals
  - John Mellor-Crummey of Rice University hosted a SUPER all-hands meeting

SUPER

INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Inter-Institute Awareness

- There are three other SciDAC-3 institutes
  - FASTMath
  - QUEST
  - SDAV
- Management coordination
  - Face to face meetings and phone calls
  - Opportunistic scheduling:
    - SciDAC conferences
    - When I visit Livermore
- Participation in all-hands meetings
  - SUPER personnel have traveled to QUEST and SDAV meetings
  - SUPER personnel have presented to a FASTMath conference call
  - FASTMath, QUEST, and SDAV people have briefed at SUPER meetings

FASTMath

- **Inter-ASCR interactions discouraged in SciDAC-1 and SciDAC-2**
  - **Nevertheless, we regularly encountered mathematical artifacts**
- **Explicit collaboration is encouraged in SciDAC-3**
  - **NUCLEI**
    - **Performance optimization of LOBPCG block eigensolver**
  - **ParaDIS**
    - **Performance analysis of the KINSOL solver**
  - **PISCEES**
    - **Performance optimization of Trilinos solvers in ice sheet models**
  - **EPSi**
    - **Additional OpenMP parallelism led to need for thread-safe PETSc**
  - **Multiscale Climate**
    - **Zoltan and hypergraph partitioning being evaluated for halo load imbalance**
  - **PDSLin**
    - **Performance analysis of hybrid sparse solver**
  - **ACES4BGC**
    - **Performance characterization of the MOAB library**

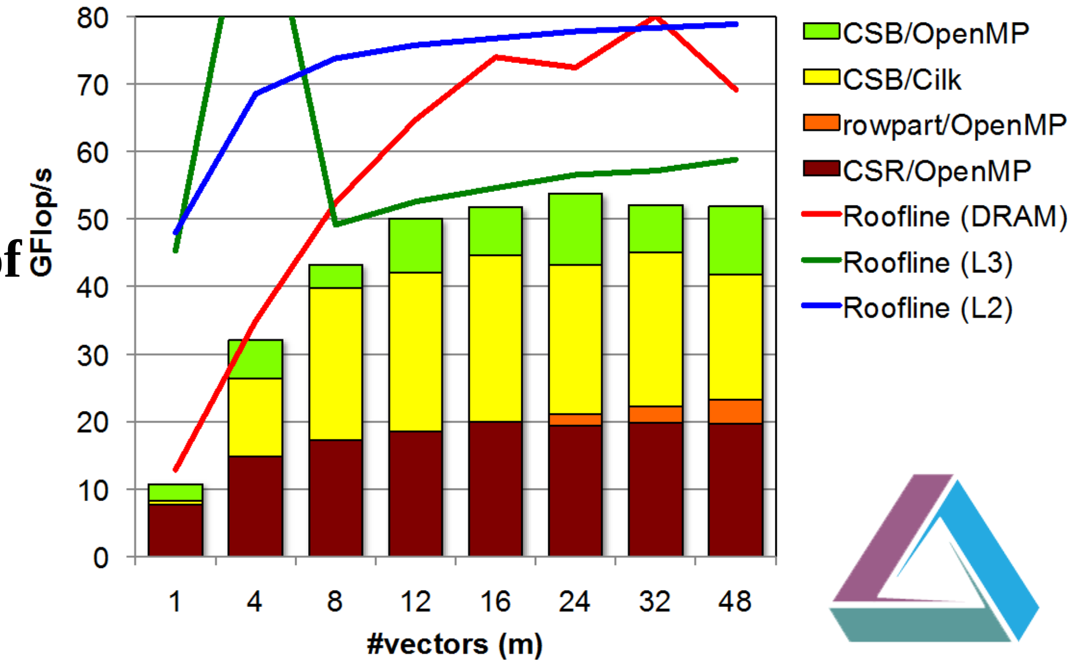
Activities and Impacts

Example Impact of Performance-Modeling Guided Development: NUCLEI / SPMM

Investigating performance advantage of replacing Lanczos with a block eigensolver, requiring an efficient implementation of Sparse Matrix-Dense Matrix Multiplication (SpMM) and a transposed variant.

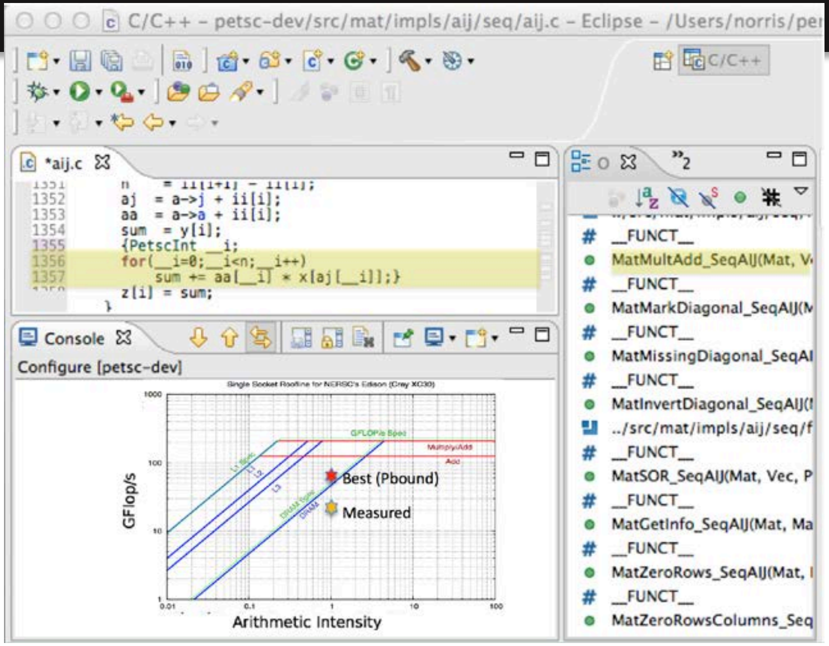
Initial work has improved performance of combined SpMM and its transpose by 3X. Further work being guided by Roofline analysis of performance bounds and transition zones.

(a collaboration with FASTMath)



SUPER

Roofline Toolkit



Enable construction, use, and interpretation of performance models by a broad range of DOE SciDAC scientists and beyond (figure shows tool sketch)

Modifications to SUPER and FASTMath at ANL and LBNL

- **Oregon also participating**

Roofline Toolkit is targeting downloadable collection of programs installable across a variety of systems, consisting of:

- **Hardware characterization, via portable, instrumented microbenchmarks and tools for many and multicore systems**
- **Software characterization via static analysis/modeling of source code, and performance counter execution instrumentation**
- **Data manipulation and visualization interface via TAU TAUdb database query system augmented for roofline analysis**

SUPER

QUEST

SUPER performance measurement and analysis of QUEST software.

In progress with UNC and Sandia (DAKOTA)

Collaboration to use SUPER tools and methods to achieve end-to-end performance and resilience in QUEST workflows at extreme scale.

Improvement of application codes in inner loops of UQ workflows.

Application of UQ methods to improve SUPER autotuning and optimization.

Using UQ methods to derive multi-scale surrogate models to complement SUPER optimization.

SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

SDAV

Interest from both SDAV and SUPER personnel in visualizing performance and other application analysis data

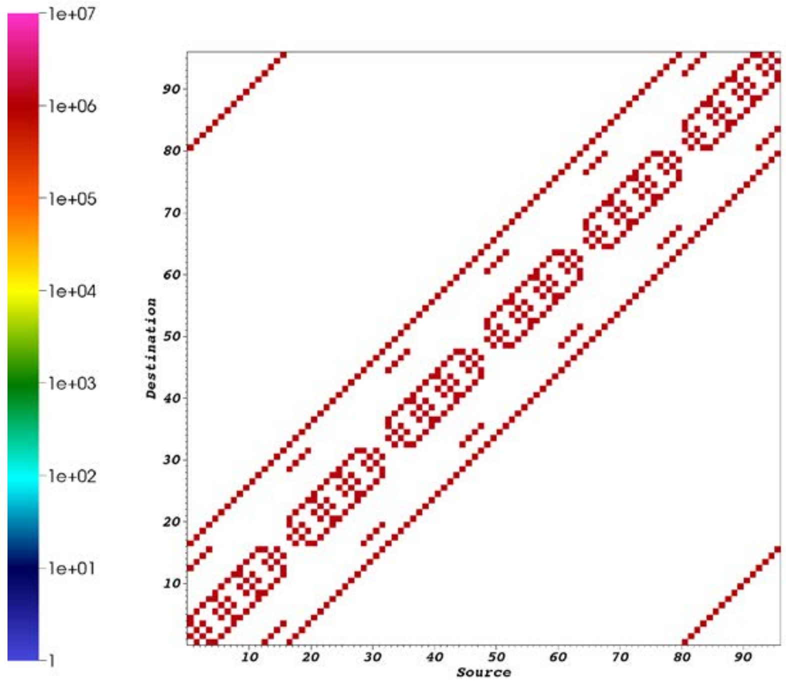
- **Visualization of communication topology data (figure on right)**
- **Visualization techniques for Value Influence Analysis results**
- **Discussions at the SDAV institute all-hands meeting in Spring 2014 regarding visualization techniques for SDAV analysis approaches**

Resilience

- **SDAV project team members led by Wes Bethel have and interest in resilience techniques**
- **SDAV participated in the last SUPER resilience team telecon and a ROSE fault injection briefing planed for an SDAV call.**

Collaboration on performance engineering

- **Empirical modeling of VPIC and VORPAL codes**



SUPER

Outline

Institute Awareness

Architecture Awareness

Application Awareness

Management Strategy

Science Pipeline

Architecture Aware Research

Measurement and Modeling

Performance Engineering

Performance Portability and Autotuning

Energy

Resilience

Multi-Objective Optimization

End-to-End Performance

Measurement and Modeling

Performance Measurement

Dyninst, mpiP, PAPI, PEBIL, pcubed, TAU

Performance Modeling

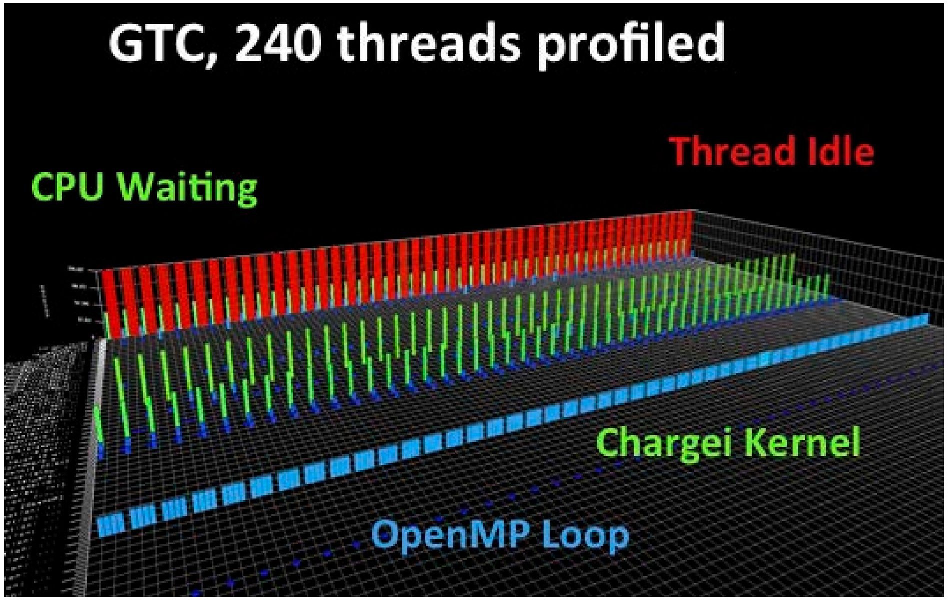
Pbound, PMAC, and Roofline

Energy Measurement

PAPI (RAPL) and RCRToolkit

Energy Modeling

Green Queue



TAU profile of GCT chargei kernel

Hierarchical Auto-Tuning for Lattice Boltzmann Simulation

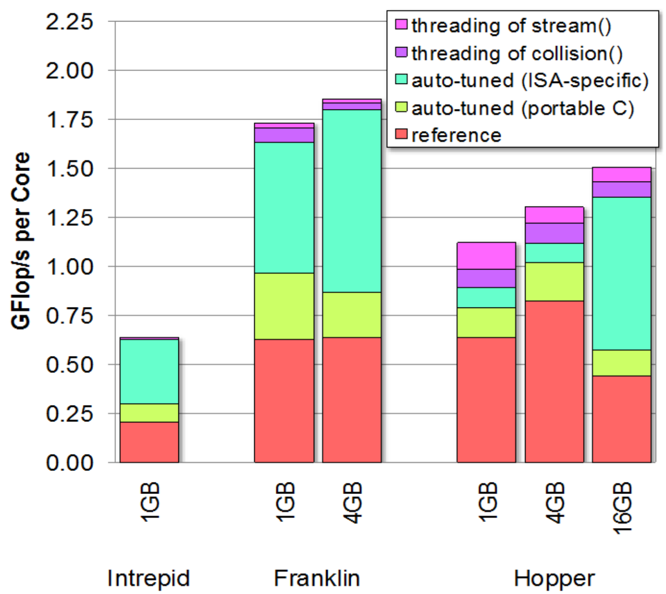
Lenny Oliker and Sam Williams

Objectives

- **Explore optimization techniques to effectively utilize complex and rapidly evolving hybrid HPC systems.**
- **Investigate Lattice Boltzmann computation that simulates homogeneous isotropic turbulence in MHD.**
- **Develop distributed auto-tuning approach, with intelligent search scheme to minimize intra-node performance variation and communication.**

Goals

- **Demonstrate that hierarchical and distributed auto-tuning effectively extracts performance.**
- **Achieve performance portability on emerging supercomputers by leveraging the system's hybrid communication capabilities.**
- **Validate approach for large-scale simulations on variety of leading supercomputing platforms.**

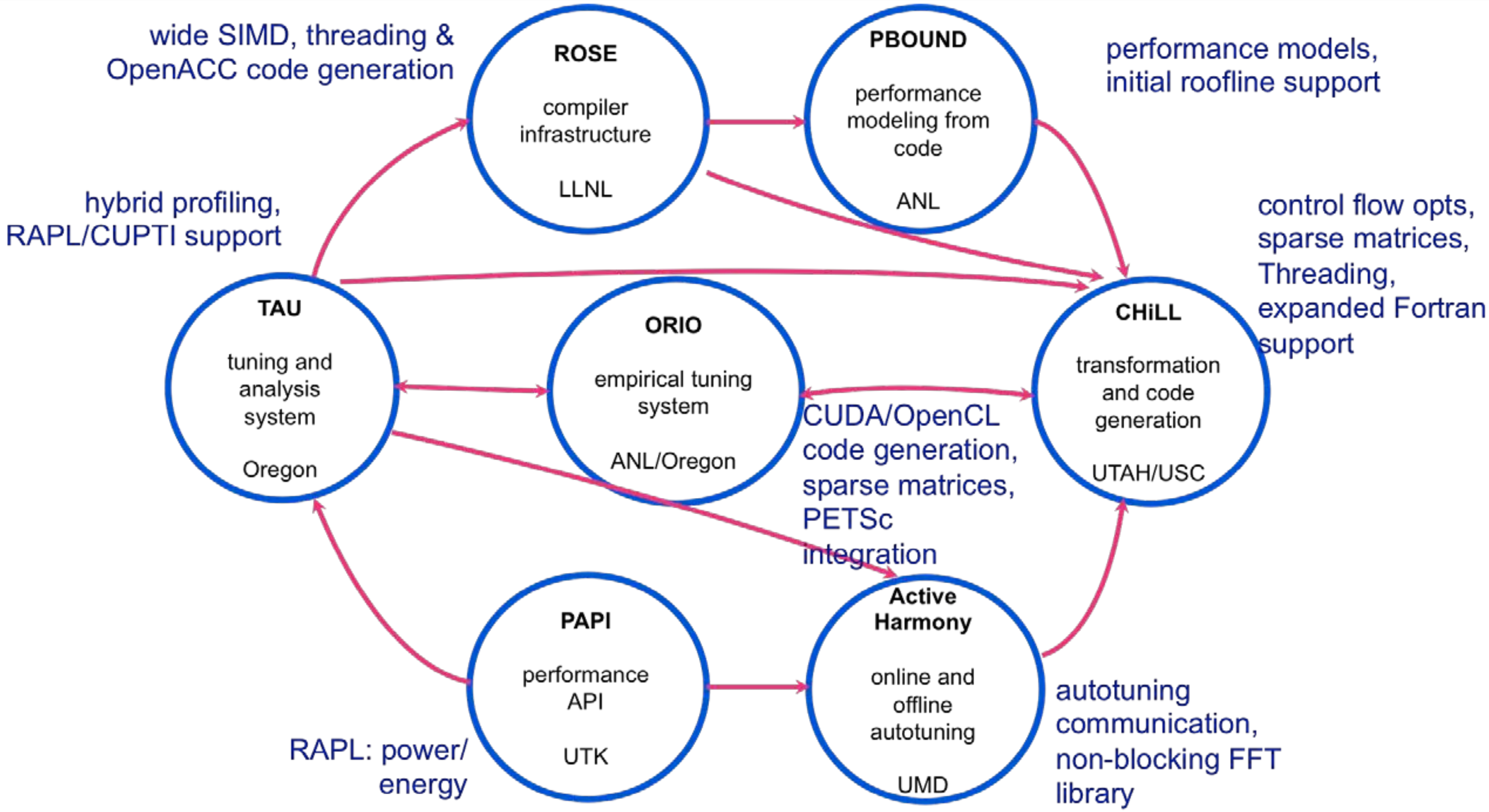


Results on up to 49,152 cores – showing auto-tuned speedup of 3.4x on Hopper versus optimized reference version

Progress & Accomplishments

- **New version compared with previously optimized SC05 Gordon Bell Finalist reference version.**
- **Methodology resulted in substantial reduction in runtime overhead, while providing performance portability on Intrepid, Franklin, and Hopper.**
- **Results show that our unique tuning approach improves performance and energy by up to 3.4x on large-scale simulations versus reference code.**
- **Approach applicable to variety of important numerical methods such as CFD and QCD on forthcoming HPC systems.**

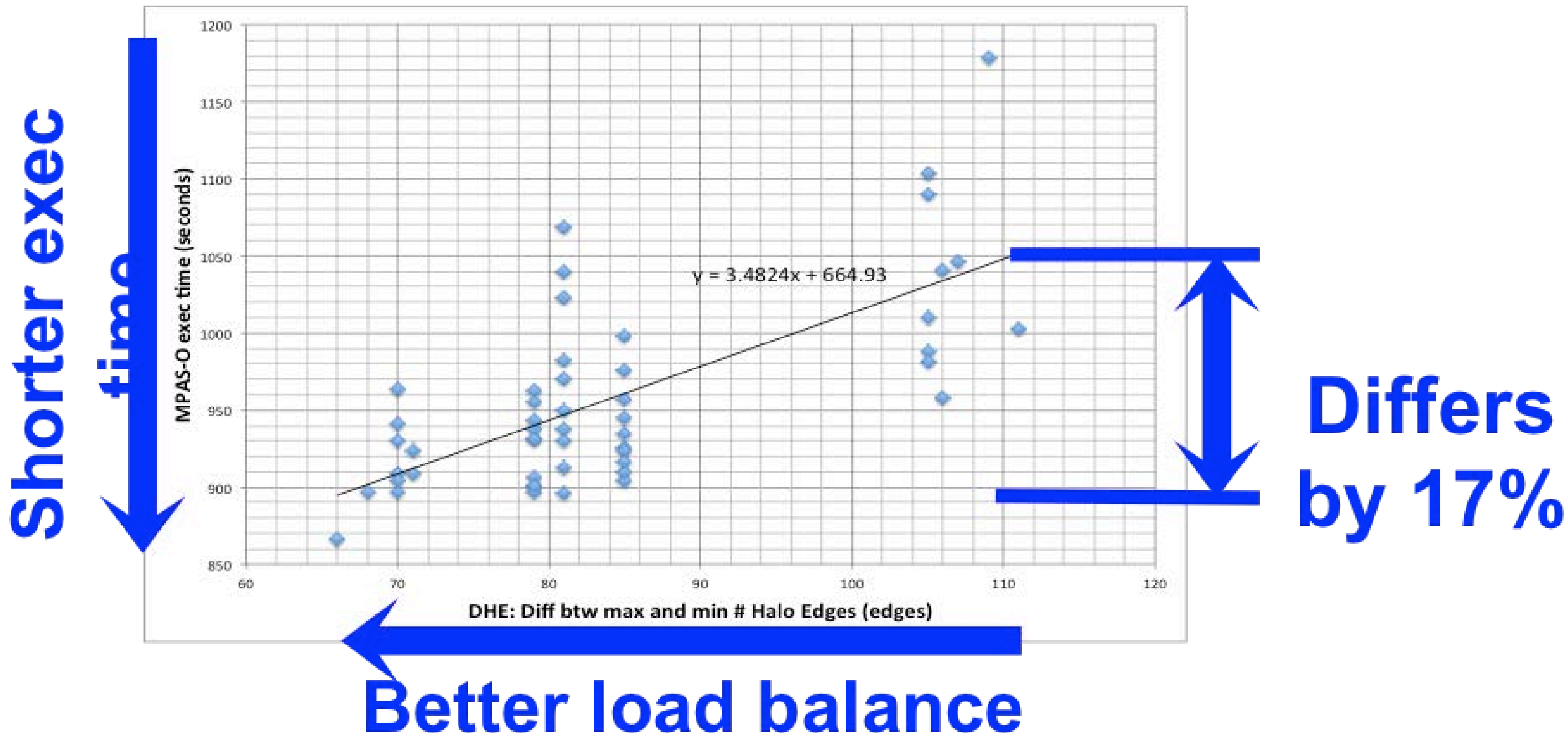
Autotuning



Tools must support architectural changes and application requirements.

Tools are extended in response to application tuning.

Partitioning in MPAS-Ocean



Add TAU_METADATA calls to capture sizes of structures
Explore different partitions with METIS

Energy Optimization

Understand how scientific computations affect the overall power demands of HPC systems

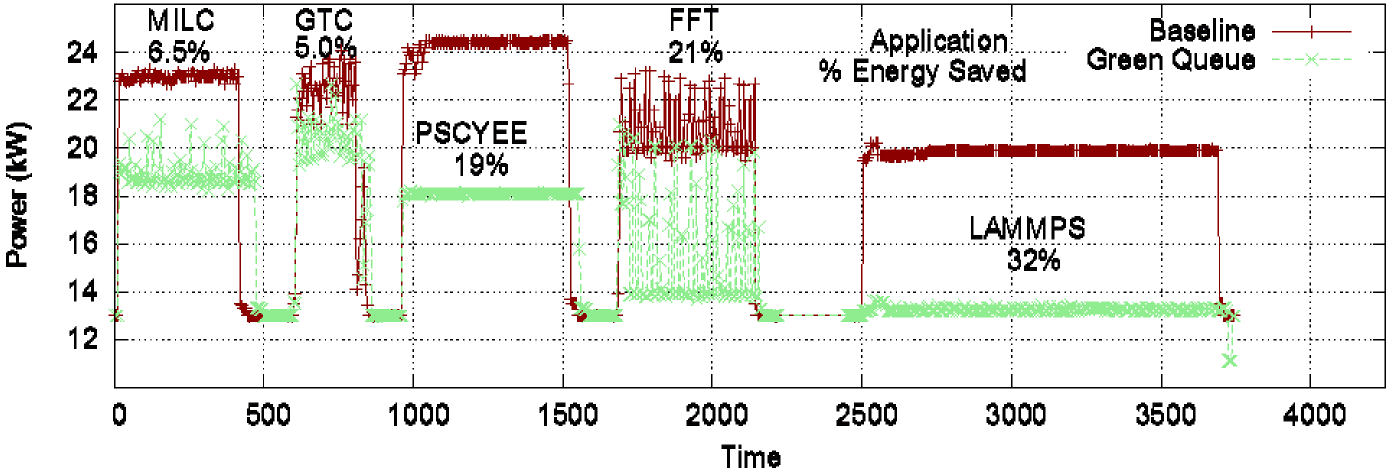
- Design application- and architecture-aware optimization techniques that enable efficient computing within a power budget or reduced energy usage

Green Queue Framework

- Automate the development and deployment of customized architecture- and application-aware power savings recipes for large-scale HPC applications

RCR Tool Suite

- Automate energy optimizations that simultaneously target multiple hardware domains to enable computing within a specified (restricted) power budget



Resilient Computing

- **Checkpointing**
  - **Optimizing checkpoint intervals**
  - **Infrastructure for runtime tracking**
    - **Checkpoint prior to to exceeding time limit**
- **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 fault 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**

Multi-Objective Optimization

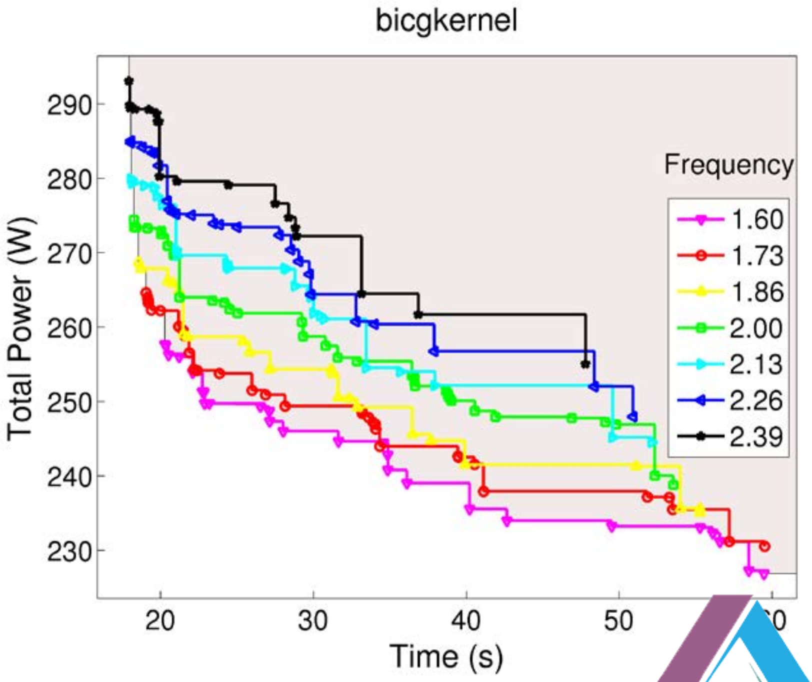
Performance, energy, and resilience are implicitly related and require *simultaneous* optimization

- **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 Tools Diagnose and Quantify Impact of Performance Variability

P. Worley (ORNL) and K. Huck (Univ. of Oregon)

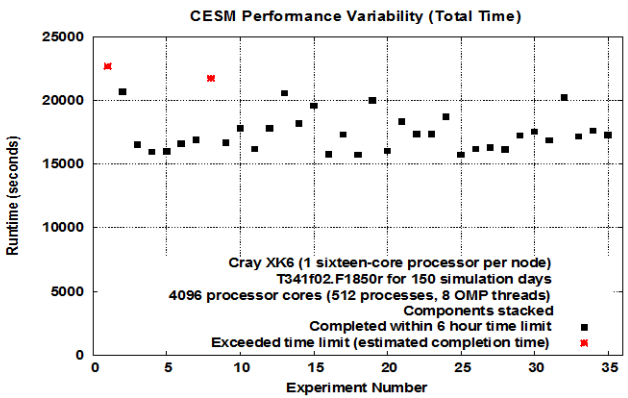
Objective

- Automate capture and analysis of performance data for production workflows, project-wide, for large-scale scientific collaborations.
  - Enable accurate assessment of project computer usage and allocation requirements
  - Identify opportunities for performance optimization, in individual simulations, in job scheduling, and in pre- and post-processing
  - Identify and quantify ways in which external factors degrade performance, enabling system-level optimizations by LCF staff

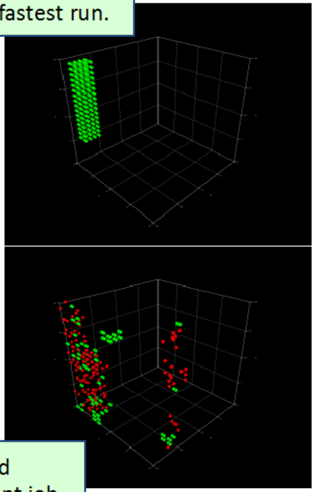
Issues

- Performance variability causes job failure when requested time is exceeded.
- Performance variability implies decreased performance and is wasteful of a project's time allocation.
- Performance variability makes both resource allocation estimation and performance optimization more difficult.

Example: Variability in performance within a suite of climate model experiments on the Cray XT5 at ORNL.



XT5 nodes allocated for the fastest run.



XT5 nodes allocated for the slowest run. Red nodes share a network switch with a different job.

Approach and Initial Results

- Nonintrusive and easy to adopt tools
  - Modified job submission scripts used to collect both system and job information and to archive data automatically
  - Application performance instrumentation data archived also when available
- Observations
  - Contention from other jobs is a major source of performance variability at NERSC and ORNL for some applications.
  - Node allocation policy can be used to mitigate this sensitivity.

U.S. DEPARTMENT OF ENERGY | Office of Science

INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Outline

Institute Awareness

Architecture Awareness

Application Awareness

Management Strategy

Science Pipeline

SUPER

SciDAC-3 Application Partnerships

- Collaboration with SciDAC Application Partnerships is expected
  - Yet SUPER funding is spread very thin
- SUPER investigators included in 14 Application Partnerships
  - Our time costs money, like everybody else
- Common features of our successful collaborations with SciDAC applications
  - A motivated member of the application team
  - A motivated member of the SUPER team
  - A specific need to improve performance of some aspect of execution, or a port to a new architecture
  - A computational kernel that represents the performance issue, including a representative input and output, validation and instructions for building and running it

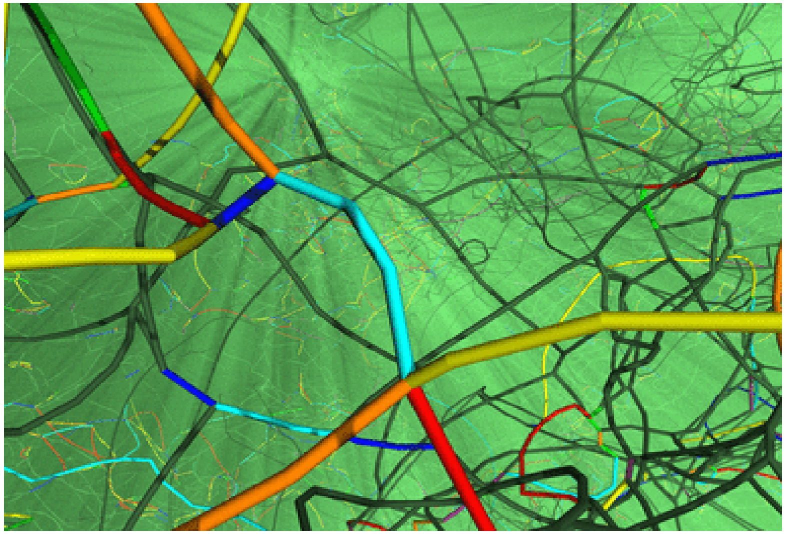
SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Participation in 12 SciDAC-3 Application Partnerships

BER	**Applying Computationally Efficient Schemes for BioGeochemical Cycles (ORNL)**
BER	**MultiScale Methods for Accurate, Efficient, and Scale-Aware Models of the Earth Sys. (LBNL)**
BER	**Predicting Ice Sheet and Climate Evolution at Extreme Scales (LBNL & ORNL)**
BES	**Developing Advanced Methods for Excited State Chemistry in the NWChem S/W Suite (LBNL)**
BES	**Optimizing Superconductor Transport Properties through Large-scale Simulation (ANL)**
BES	**Simulating the Generation, Evolution and Fate of Electronic Excitations in Molecular and Nanoscale Materials with First Principles Methods (LBNL)**
FES	**Partnership for Edge Plasma Simulation (ORNL)**
FES	**Plasma Surface Interactions (ORNL)**
HEP	**Community Petascale Project for Accelerator Science and Simulation (ANL)**
NP	**A MultiScale Approach to Nuclear Structure and Reactions (LBNL)**
NP	**Computing Properties of Hadrons, Nuclei and Nuclear Matter from QCD (UNC)**
NP	**Nuclear Computational Low Energy Initiative (ANL)**

SUPER

Collaboration with ASC



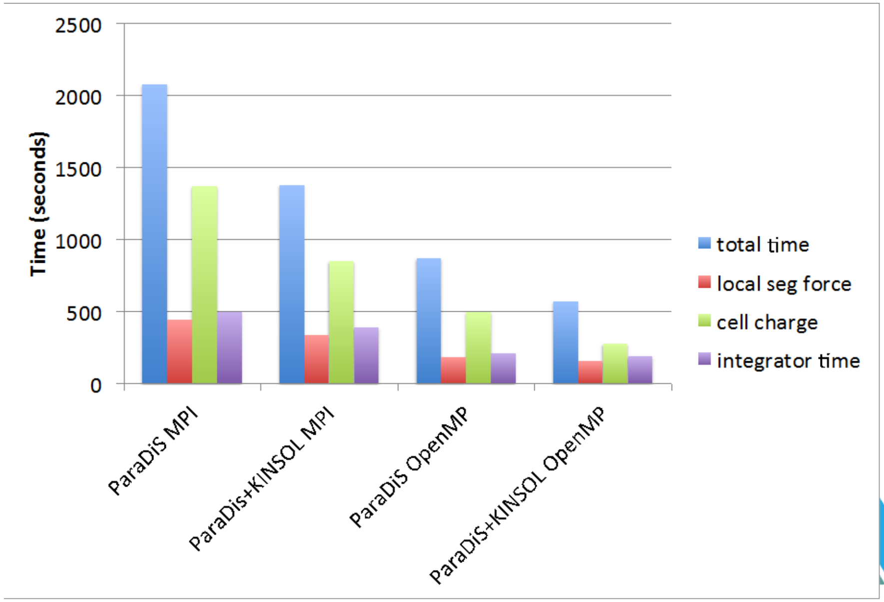
ParaDiS models dislocation lines in a physical volume

- The volume is divided into several computational domains

The dislocations interact and move in response to physical forces

- External stress and inter-dislocation interactions determine the forces

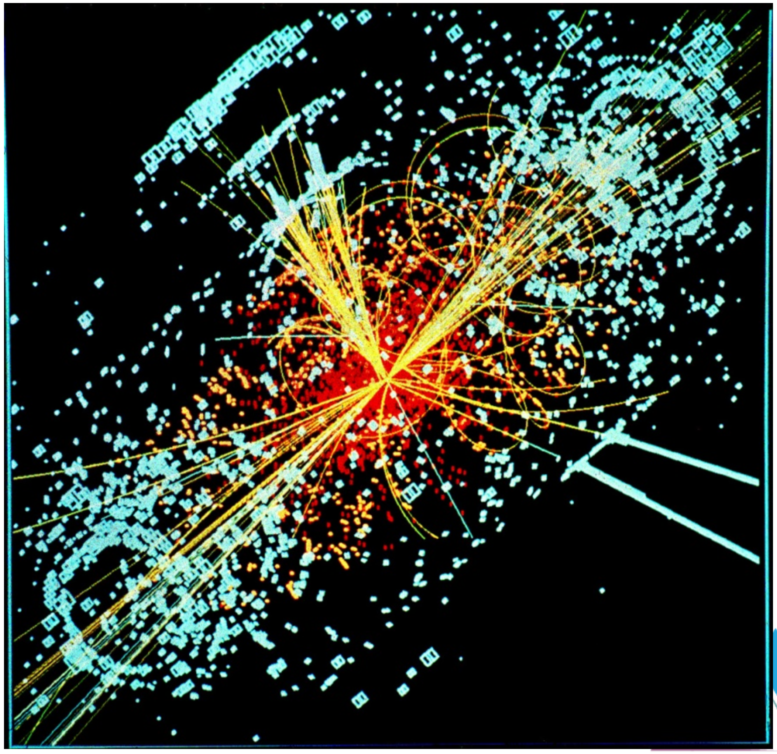
Porting ParaDIS with KINSOL (FASTMath) to BG/Q and characterizing performance at different scales, with and without OpenMP.



Geant4

Exploring the Transformation of Geant4 for the Future

- Geant4 is extremely important to the design and execution of HEP experiments
- How can it adapt to future multicore devices?



Partnership with HEP

- Lali Chatterjee is the HEP PM
- Ceren Suset is the ASCR PM

Modification to SUPER at ANL

- Oregon, UNC and USC also contributing
- FNAL and SLAC contributing for HEP

Outline

Institute Awareness

Architecture Awareness

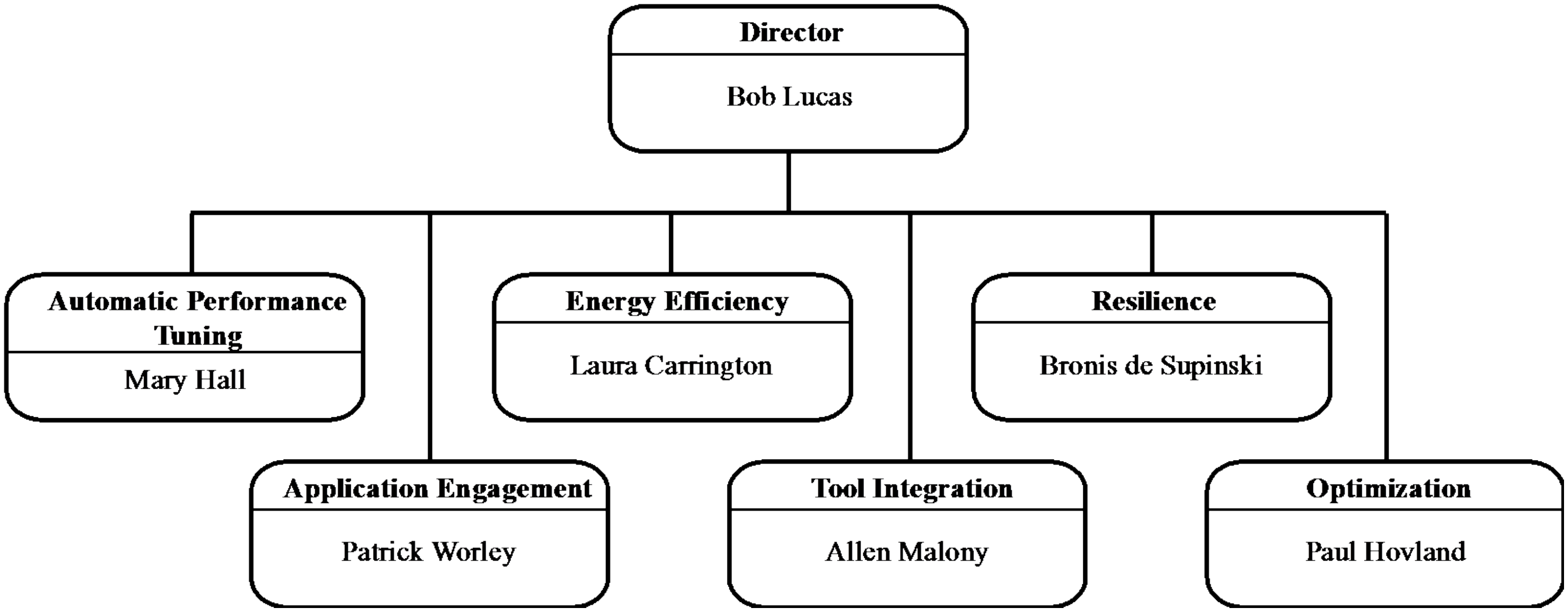
Application Awareness

Management Strategy

Science Pipeline

Posters

SUPER Organization



Relatively flat organization

- David Bailey initially helped manage overall project
- Lenny Oliker stepped in when David retired

Functions well after a dozen year history of collaboration

INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Weekly Calls

Wednesday, 9:00 AM PDT

All hands (Bob)

S/W integration (Al)

Resilience (Bronis)

Optimization (Paul)

Tuesday, 11:30 AM PDT

Autotuning (Mary)

Energy (Laura)

End-to-end performance (Pat)

Measurement and modeling (Al)

The schedule adapts as needed

- **Support focus on MPAS-Ocean code**
- **Preparation for midterm review**

SUPER

Meeting Schedule

University of Oregon	Sept. 21-22, 2011
University of North Carolina RENCI	March 29-30, 2012
Argonne National Laboratory	Sept. 26-27, 2012
Rice University	April 4-5, 2013
Lawrence Berkeley National Laboratory	Sept. 18-19, 2013
University of Utah	March 27-28, 2014
University of Maryland	TBD Sept. 2014
University of California, San Diego SDSC	TBD March 2015
Oak Ridge and University of Tennessee	TBD Sept. 2015
University of Southern California ISI	TBD March 2016

SUPER Infrastructure

- **Web site**
  - **Shirley Moore (UTEP) manages**
- **Wiki**
  - **Shirley Moore (UTEP) manages**
- **SVN (code, data, reports, talks)**
  - **Rob Fowler (UNC) hosts**
- **Performance database**
  - **Al Malony (Oregon) hosts**
- **NERSC project account**
  - **LBNL arranged**
- **PEAC INCITE allocation**
  - **Initially led by Pat Worley (ORNL)**
  - **Now led by Lenny Oliker (LBNL)**
  - **Broader collaboration than just SUPER**

Outline

Institute Awareness

Architecture Awareness

Application Awareness

Management Strategy

Science Pipeline

SUPER

INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Science Pipeline

SUPER funnels the rich intellectual products borne from a history of research and development in performance areas into an effective performance engineering center of mass for the SciDAC program

SUPER pulls from prior investments by ASCR and others the technology and expertise that past efforts produced, especially with respect to methodologies, tools, and integration across performance engineering areas

- measurement, analysis, modeling
- program analysis, optimization and tuning
- resilience

SUPER focuses on integration of expertise for addressing performance engineering problems across the SciDAC landscape, leveraging the robust performance tools available

SUPER

Pipeline to Tools/Technology Integration and Application

SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Tools / Technologies

TAU
PAPI
mpiP
GPTL
RCRToolkit
PBound
Roofline
PEBIL
PSiNtracer
ROSE
CHiLL
Active Harmony
Orio

ASCR funding

- Performance
- Modeling
- Reliability
- Code analysis
- Autotuning

Integration
End-to-end

Center of mass for performance engineerng

- Autotuning
- Optimization
- ***SciDAC applications***
- Resilience
- Energy

NSF, DARPA, and other funding

SUPER

SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

TAU Development Pipeline

Funding pipeline: 2001 – 2011

DOE

Evolution — Flexible performance measurements, Performance mapping in software layers

CCA — Automated source instrumentation, Modeling and computational QoS

ZeptoOS — Kernel-level measurement, Runtime scalable monitoring

Productive — Source code analysis (PDT), Performance data management (PerfDMF)

Knowledge — Performance knowledge, Performance data mining (PerfExplorer)

PRIMA — Measurement infrastructure refactor, TAU + Scalasca → Score-P

MOGO — Parallel performance visualization, Automatic library wrapping

**Vancouve
r** — Heterogeneous performance, Accelerator analysis

NSF

POINT — Open source interoperation, Performance engineering

Glassbox — Cross-layer Integration

Posters

SUPER Performance Auto-Tuning and MPAS-Ocean Optimization

Multi-objective Optimization of Power, Energy, Performance, and Resilience

SUPER Power Awareness for HPC

SUPER Engagement: Activities and Impacts

Resilience Assessment and Enhancement

Transforming Geant4 for the Future

SUPER
INSTITUTE FOR SUSTAINED PERFORMANCE, ENERGY, AND RESILIENCE

Summary

SciDAC-3 Research

- Performance Engineering of Scientific Software
- Automatic performance tuning
  - New focus on portability
- Addressing the “known unknowns”
  - Energy minimization
  - Resilient computing
- Optimization of the above

Impact on DOE computational science applications

- Tool integration, making research artifacts more approachable
- Participation in SciDAC-3 Application Partnerships
- Engagement with broader DOE HPC community

SUPER