Performance Optimization and Autotuning in the SUPER Project

Presenter: Shirley Moore
University of Texas at El Paso
October 7, 2014

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 18 56

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

www.super-scidac.org

SUPER

Research Themes

Architectures and applications are evolving rapidly

- Exponential growth in cores, which are becoming heterogeneous
- Scientific questions spanning multiples scales and phenomena
- High performance is increasingly difficult to achieve and maintain

Energy consumption has emerged as a major concern

- Cost of power and cooling is constraining science

Resilience will be a challenge in the very near future

- Shrinking VLSI geometries and voltages will reduce device reliability

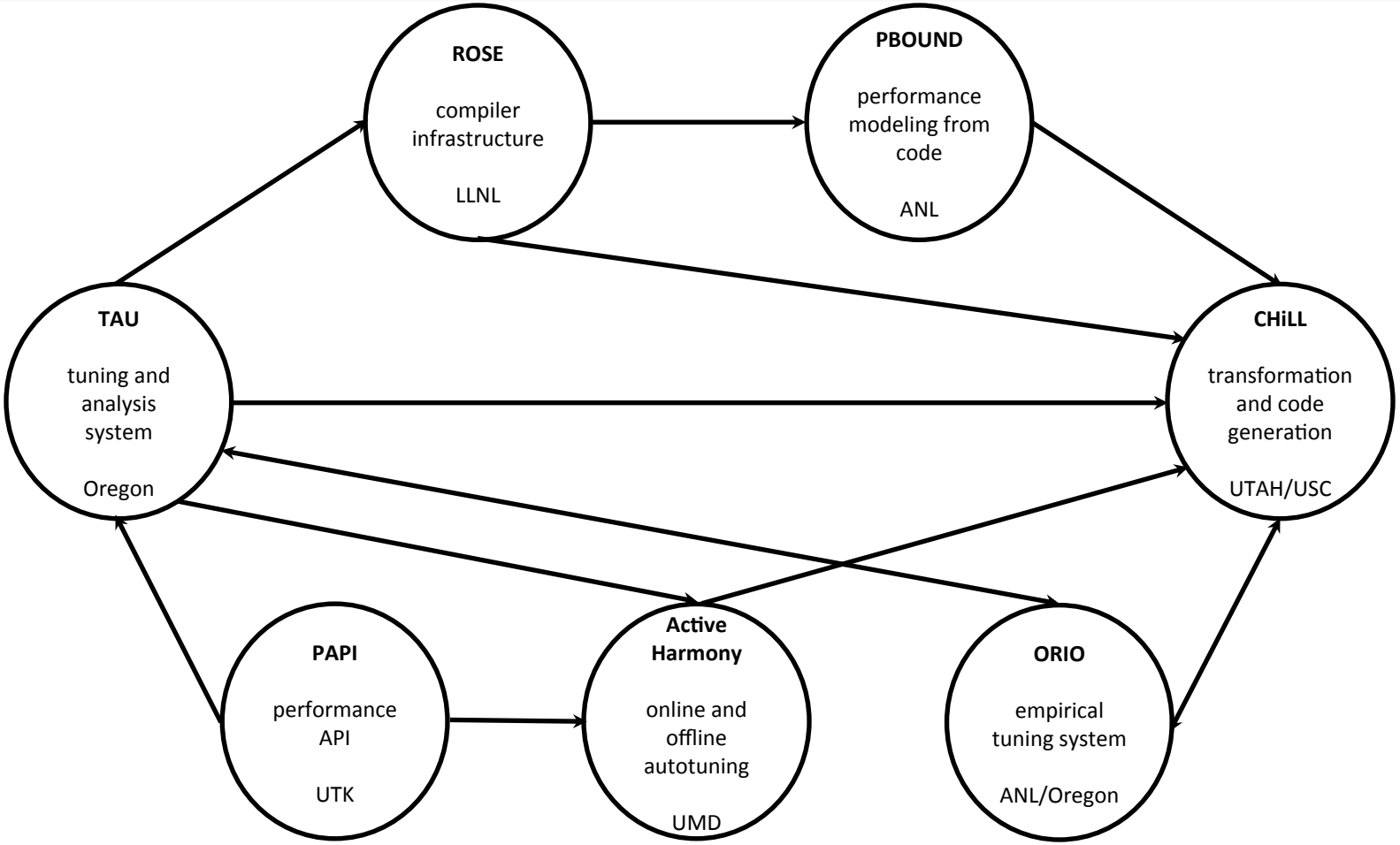
These can be traded off against each other

- Maximize scientific throughput per dollar or Watt

Presentation Focus

- Tool Integration and Development
- Optimization of MPAS-Ocean
- Summary of other Measurement/Analysis/ Optimization Efforts

SUPER Tool Integration

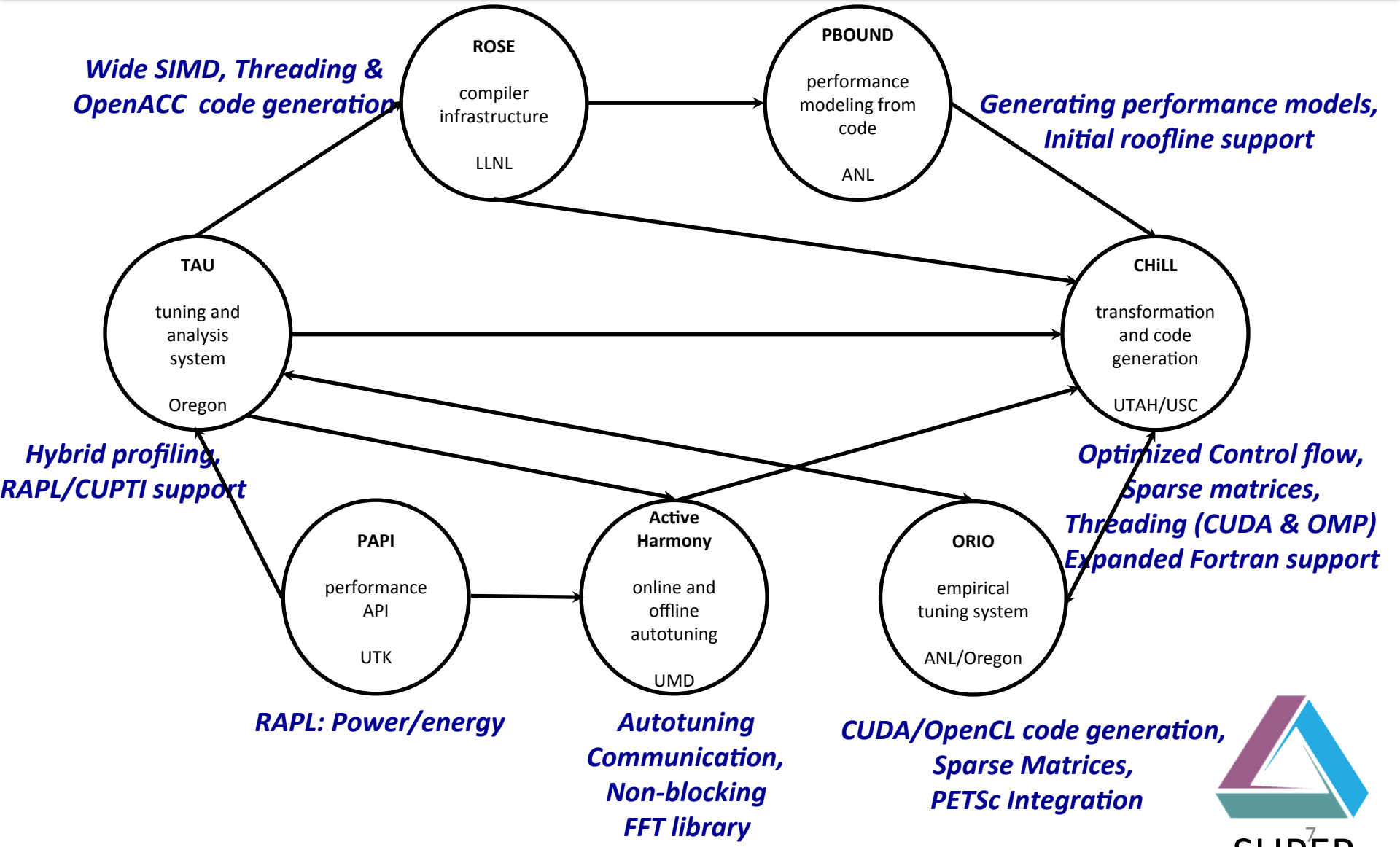


Focus of first year on tool integration.
Edges/arrows show SUPER tools that have been integrated.

SUPER Performance Focus

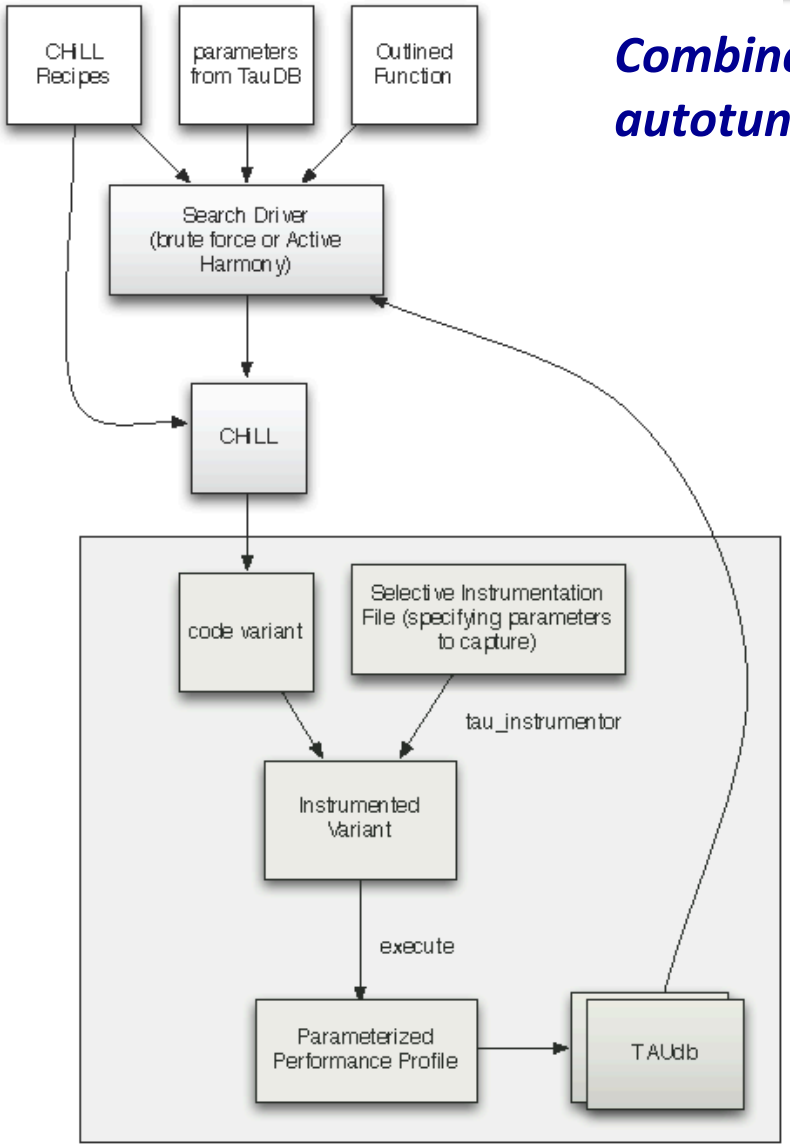
- Tools must support architectural changes…
  - Heterogeneous processing units
  - Hierarchical parallelism, mixed threads and processes
  - Wide SIMD
- … and application requirements
  - Scalability to many cores, processes
  - Dynamic parallelism
  - Batched linear algebra
  - Sparse matrices

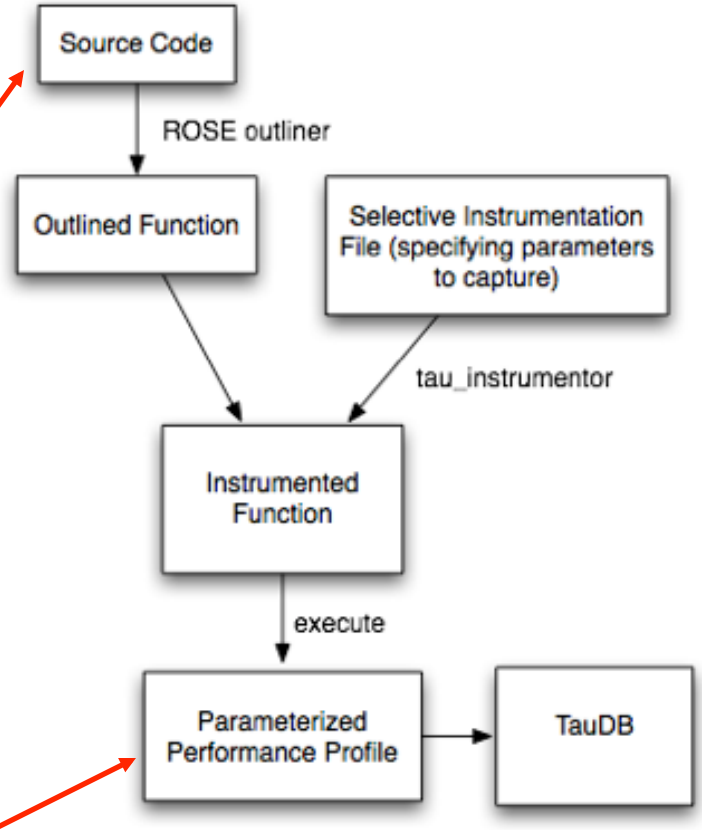
Progress on Tool Development



Example: TAU + ROSE + CHiLL + Active Harmony

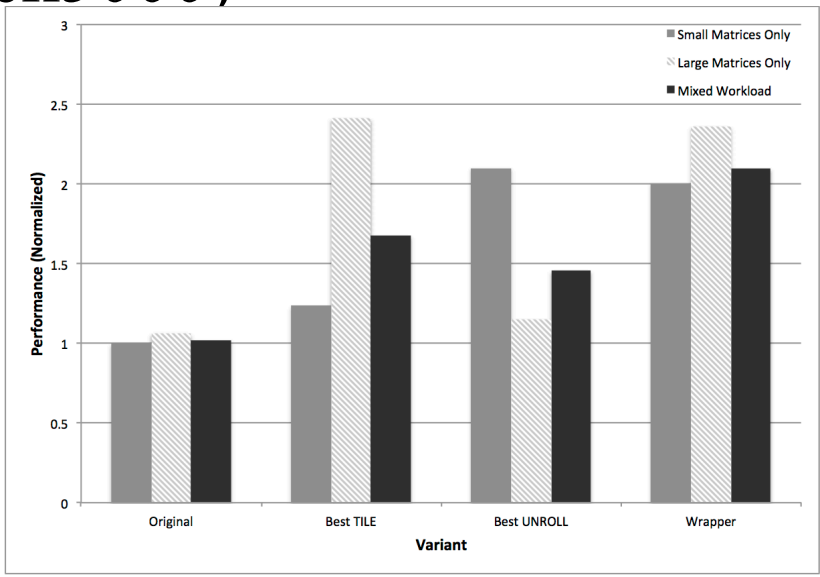
Combine tools to create, store and query results of autotuning experiments in TAUdb

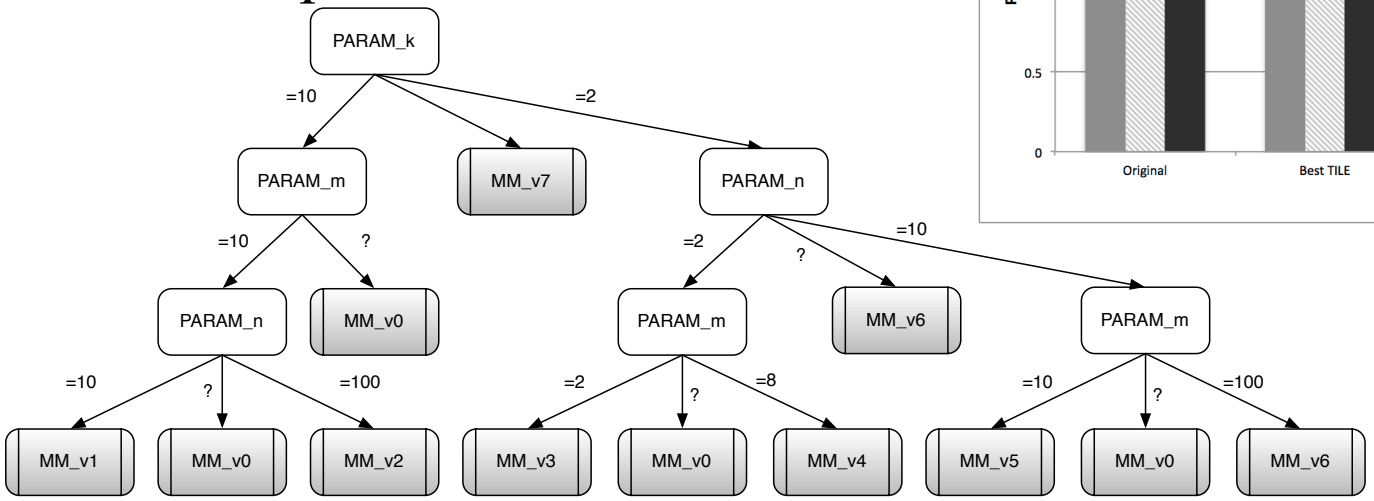




Example: TAU + ROSE + CHiLL + Active Harmony

- Apply machine learning to data stored in TAUdb
  - Generate decision trees based upon code features
- Consider matrix multiply (e.g., Nek5000)
  - Matrices of different sizes with different performance
  - Automatically generate function to select specialized variants





Applying Tools to SciDAC Codes

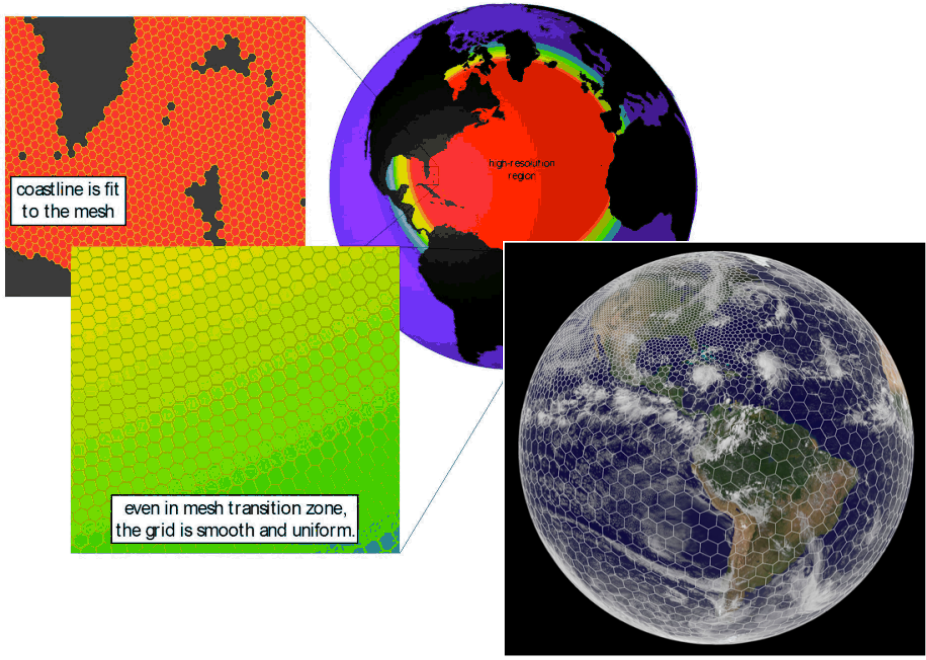
- SUPER interaction with SciDAC application teams
- Approach:
  - Collection of SUPER researchers examine application excerpt or full application to identify performance bottlenecks
  - Opportunities to improve code in different ways are explored.

MPAS-Ocean

- A new code ➔ an opportunity to have early impact
- Complex representation: variable resolution irregular mesh of hexagonal grid cells

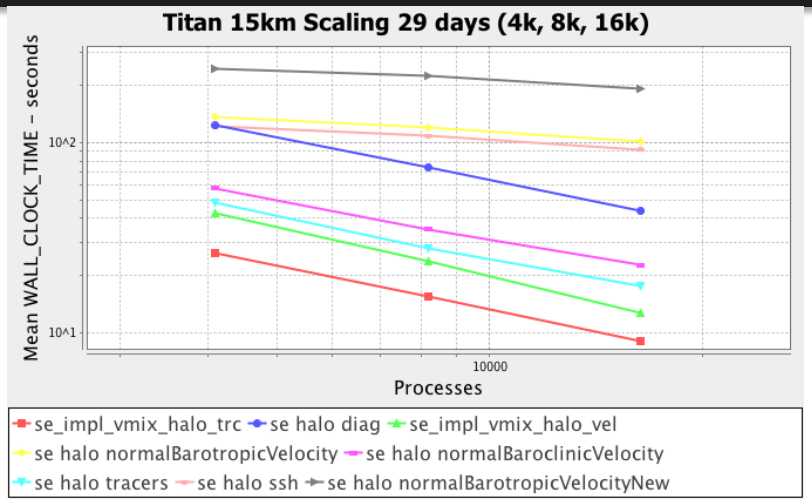
MPAS
Model for Prediction Across Scales

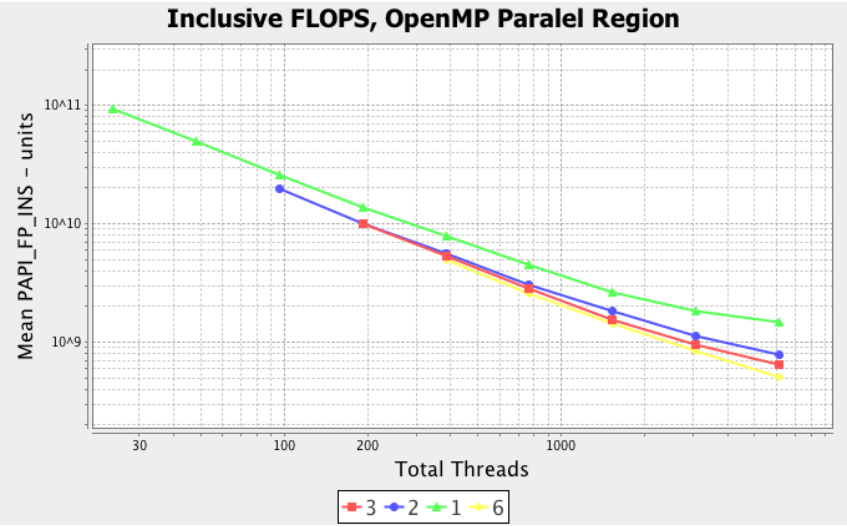
http://mpas-dev.github.io

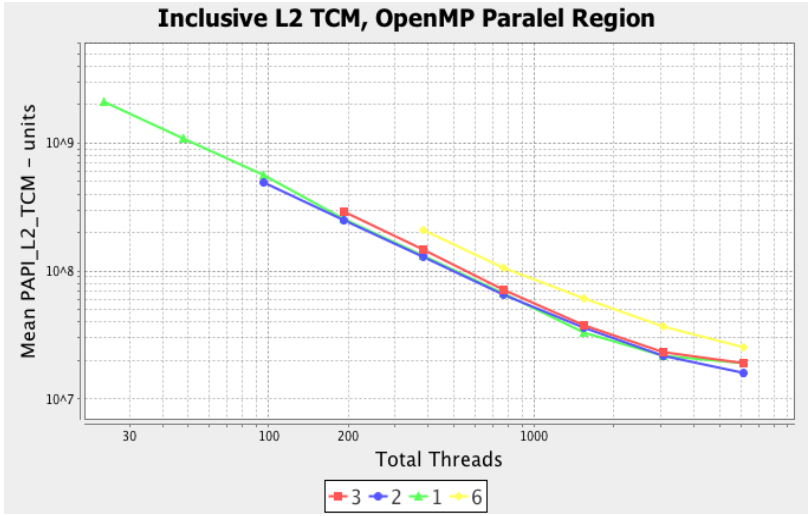


Issue 1: Limited Scalability

- MPI-only scaling leads to additional halo computation, communication
- OpenMP scaling promising, but reduction in FLOPs/core offset by increase in cache misses

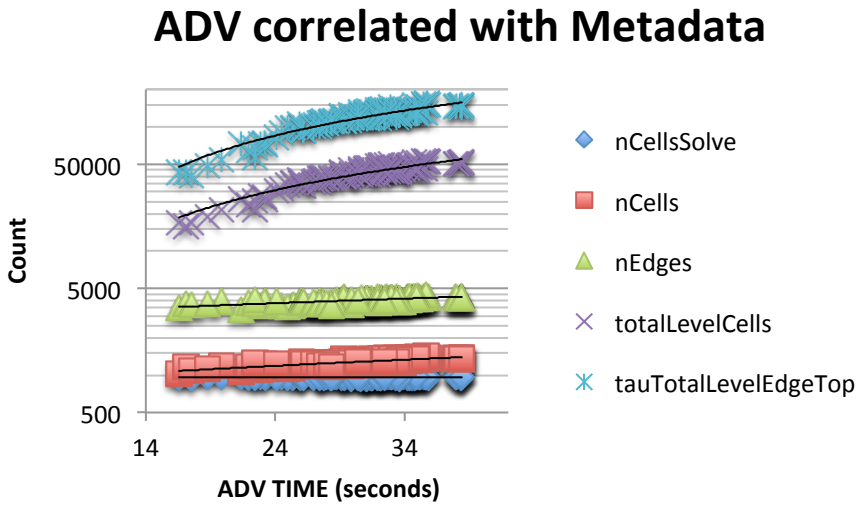


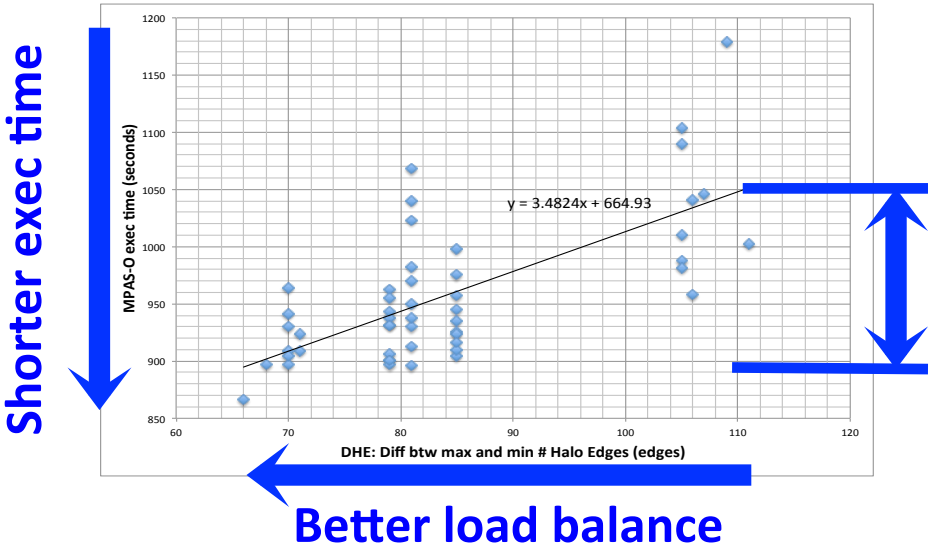




Issue 2: Load Imbalance Due to Partitioning

- Add TAU_METADATA calls to capture sizes of structures
  - Correlate computation balance with metadata fields
- Explore different partitions with METIS



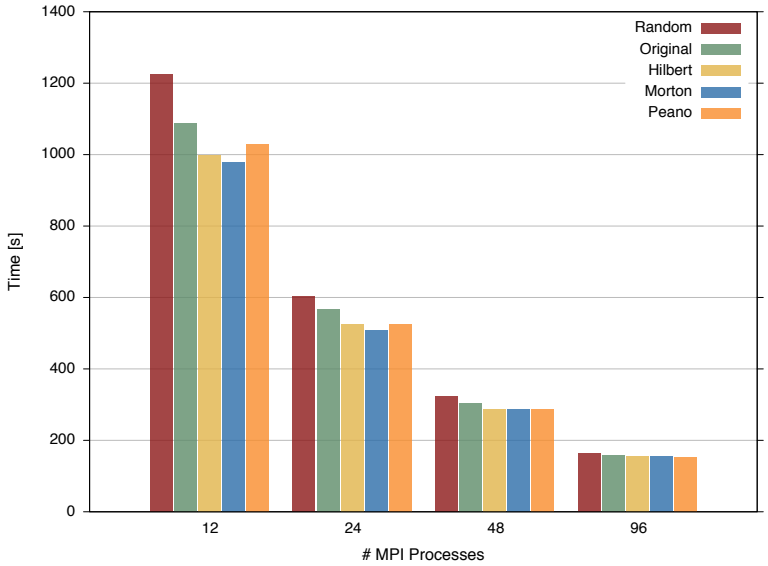


- Conclusion:
  - Partition on nCellsSolve, but value unknown before partitioning

SUPER

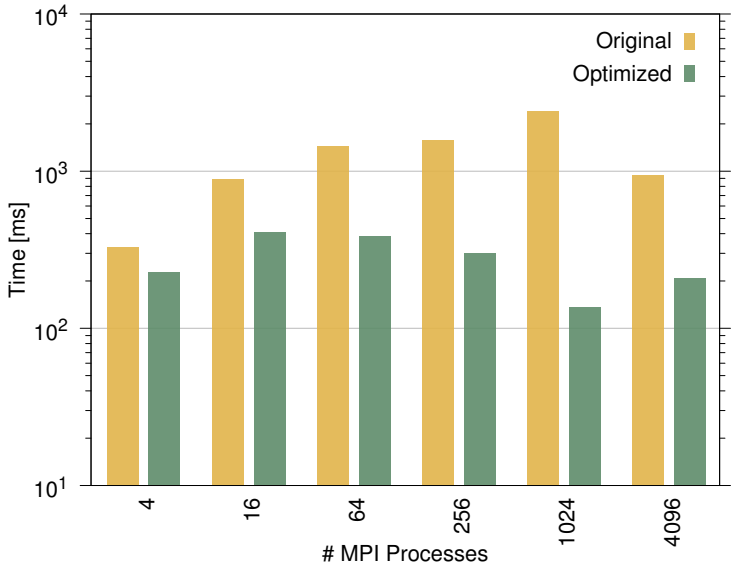
Issue 3: Communication Costs

Intra-Node:
Ordering of Cells Impacts Cache Behavior



Solution:
Use space-filling curves to explore different cell orderings

Inter-Node:
Communication Frequency Adds Overhead



Solution:
Aggregate communication, represented by benchmark

Issue 4: Indirection Increases Instruction Count

Performance Issue: Significant structure indirection inside loop nests

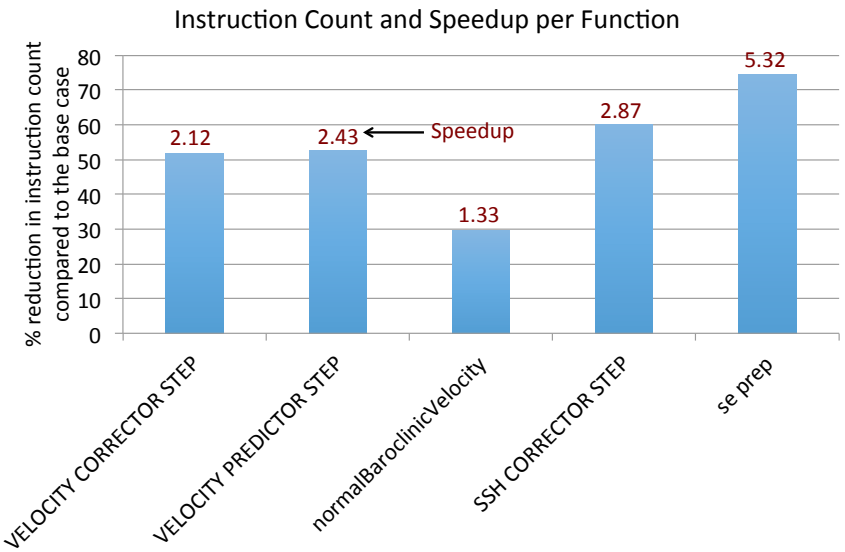
block%mesh%cellsOnEdge%array(1,iEdge) block%mesh%cellsOnEdge%array(2,iEdge)

Solution: Replace with pointer buffers that point directly to array

```
block => domain%blocklist
do while (associated(block))
  !...
    do iEdge=1,block%mesh%nEdges
      cell1 = block%mesh%cellsOnEdge%array(1,iEdge)
      cell2 = block%mesh%cellsOnEdge%array(2,iEdge)
      !...
  end do
  block => block % next
end do  ! block
```

Pointer buffer Auto-generator

```
! define Pointer buffers
real(kind=RKIND),dimension(:,:),pointer :: cellsOnEdge
block => domain%blocklist
do while (associated(block))
  ! associate pointer buffer to target variable
  cellsOnEdge => block%mesh%cellsOnEdge%array
  !...
    do iEdge=1,block%mesh%nEdges
      !Using pointer buffers to do calculation
      cell1 = cellsOnEdge(1,iEdge)
      cell2 = cellsOnEdge(2,iEdge)
      !...
  end do
  ! nullify the association when computation finished
  nullify(cellsOnEdge)
  ! deallocate buffers
  deallocate(cellsOnEdge)
  block => block % next
end do  ! block
```



Using pointer buffers in all functions: overall **1.16x** application speedup

SUPER

Applying Tools to SciDAC Codes, cont.

- Other code examples employing SUPER tools
  - XGC1
    - Performance measurement using TAU/CUPTI (also, HPC Toolkit and PerfExpert)
    - Conclusion: low IPC and high data accesses in PUSHE
    - Currently working on semi-automated generation of OpenACC directives
  - USQCD
    - Sub-optimal performance for quantum linear algebra
    - Rewrote to higher level and used CHiLL to regenerate low-level complex matrix-vector multiplies
    - Need to generalize to higher-dimensional lattices

NWChem TCE module

- Tensor Contraction Engine
  - implements approximations that converge toward exact solution of the **Schrödinger equation**
- libtensor
  - ***Key Challenge:*** Master-worker task parallelism of small BLAS calls, agnostic of NUMA and cache locality
  - Implementing NUMA measurement capability in PAPI to enable optimization of task scheduling

"Analysis and Tuning of Libtensor Library"
Khaled Ibrahim and Samuel Williams (Lawrence Berkeley National Lab)

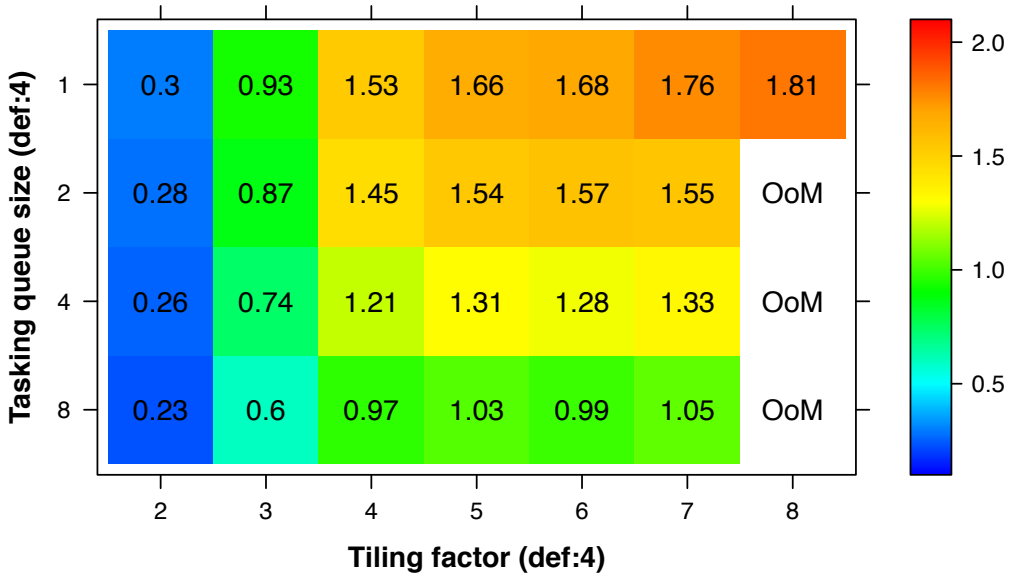
Objectives:

- Multicore architectures are imposing multiple challenges for scaling scientific applications.
- This work aims at identifying performance bottlenecks in multiple quantum chemistry chemistry frameworks.
- Expose challenges in quantum chemistry codes to computer scientist.

Impact:

- Tuning scientific codes leads to performance improvement and efficient execution.
- Make domain scientists focus efforts on the chemistry research rather than computational aspects.
- Allow computer scientists to explore efficient techniques to increase the productivity of developing codes for quantum chemistry computations.

Performance improvement on 24 core Intel Ivy-bridge (Edison)



Progress and Accomplishments:

- **More than 1.8X improvement in performance on 24-core Ivy-bridge system. More than 2X improvement on AMD Magny-Cours.**
- **Speedup improvement from 9.5x to 15.3x on 24 core Intel Ivy-bridge.**
- **Results and techniques to be published.**

Conclusion

- SUPER creates the opportunity for state-of-the-art performance tuning to be applied to SciDAC applications (e.g., MPAS-Ocean).
- SUPER fosters a community of performance tuning researchers, informed by the needs of application programmers.
- SUPER performance tools have been integrated, and are continuously extended to meet these needs.