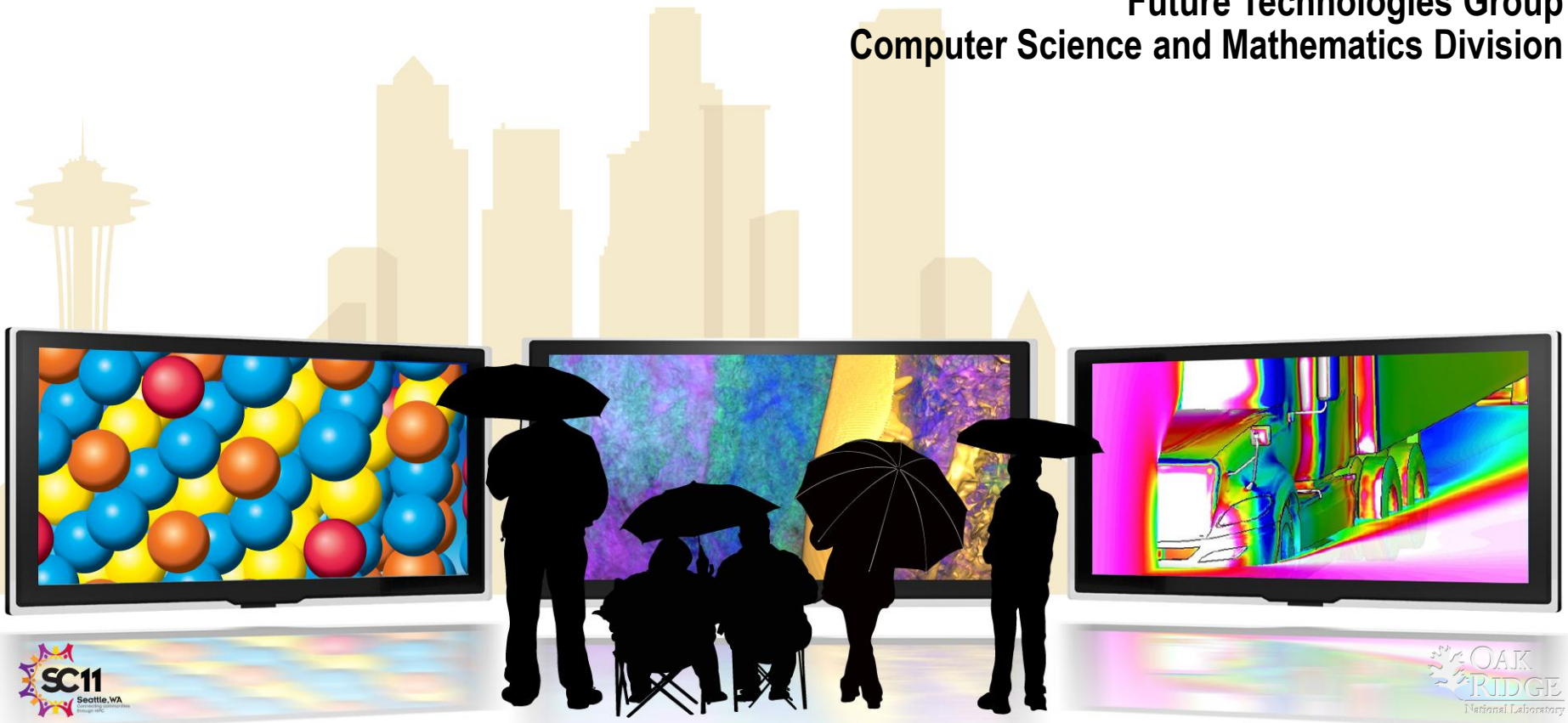


Performance Evaluation and Analysis Consortium (PEAC) End Station

Presented by

Patrick H. Worley

Future Technologies Group
Computer Science and Mathematics Division



Overview

The PEAC End Station provides the performance evaluation and performance tool developer communities access to the Leadership Computing Facility (LCF) systems

Consortium goals

System evaluation

- Evaluate the performance of LCF systems using standard and custom micro-, kernel, and application benchmarks
- Determine efficient usage techniques and identify and characterize system performance issues

Performance tools

- Port performance-related tools and system infrastructure to LCF systems and make them available to Oak Ridge Leadership Computing Facility (OLCF) and Argonne Leadership Class Facility (ALCF) users
- Further develop the tools to take into account the scale and unique features of LCF systems

Performance modeling

- Validate the effectiveness of and further develop performance characterization and prediction methodologies for LCF systems

Overview (continued)

Consortium goals (continued)	
Application analysis and optimization	<ul style="list-style-type: none">• Analyze and help optimize performance of current and candidate LCF application codes
Performance and application community support	<ul style="list-style-type: none">• Provide access to other performance researchers who are interested in contributing to the performance evaluation of the LCF systems or in porting complementary performance tools of use to the OLCF user community• Provide access to application developers who wish to evaluate the performance of their codes on LCF systems

All of this must be accomplished while adhering to the “Golden Rules” of the PEAC community:

- **Low visibility (no production runs!)**
- **Open and fair evaluations**
- **Timely reporting of results**

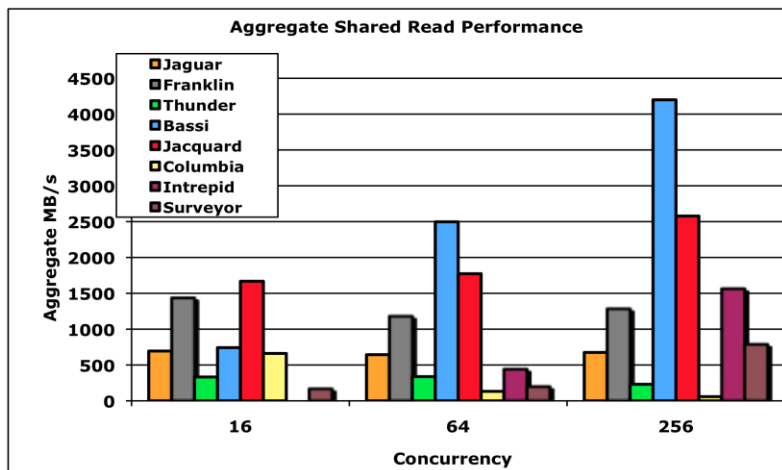
Status as of September 1, 2011

23 active ALCF users, 40 active OLCF users

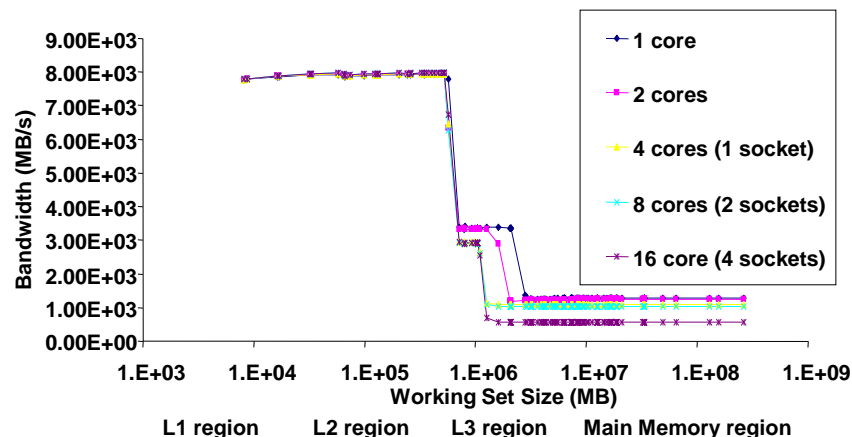
Working on	Consuming	Contributing to (at least)
<ul style="list-style-type: none">• Application performance analysis and optimization (18)• Application performance modeling (10)• Algorithm development (7)• Performance tool development (37)• Performance-related system infrastructure development (16)• System performance evaluation (10)	<ul style="list-style-type: none">• XT5: 7,000,000 processor hours out of an allocation of 20,000,000 hours• BG/P: 2,000,000 processor hours out of an allocation of 10,000,000 hours <p><i>since January 1, 2011</i></p>	<ul style="list-style-type: none">• Two book chapters• Four journal articles• Eighteen proceedings papers, including: 2 CUG, 1 Euro-Par, 1 ICPP, 3 ICS, 3 IPDPS, 5 SC11• Two oral presentations• Two code releases <p><i>appearing since January 1, 2011</i></p>

many working in multiple areas

System evaluation

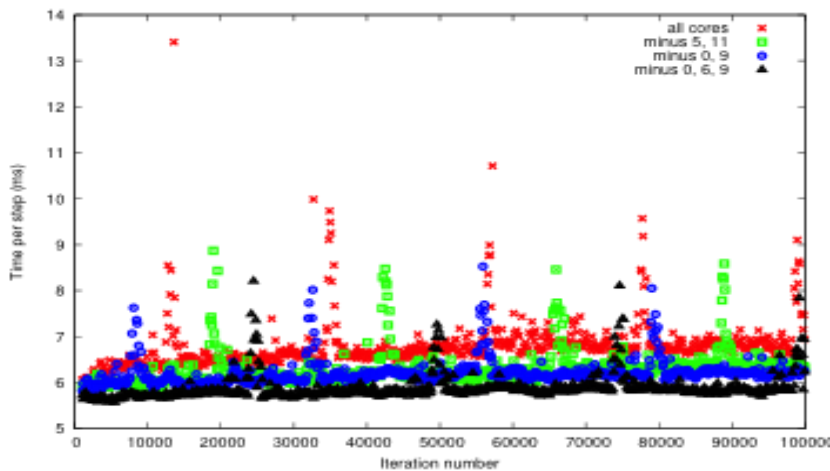


I/O performance characterization [LBL]



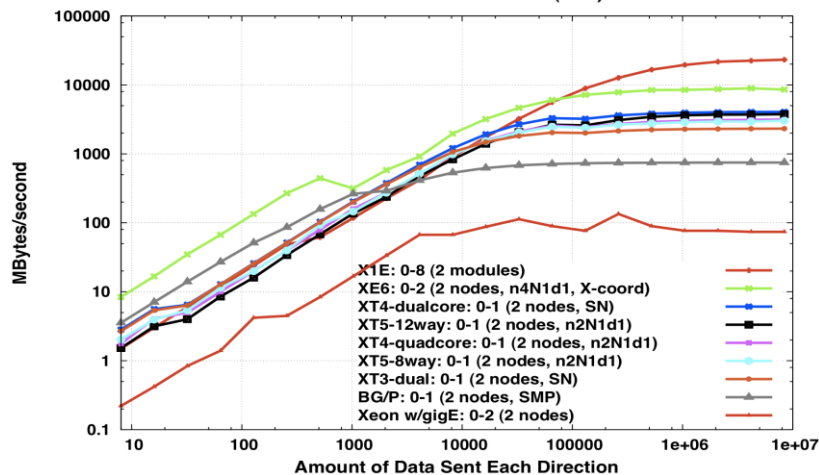
Effect of dual- to quad-core upgrade on memory hierarchy bandwidths [SDSC]

NAMD on Cray XT5 (IAPP)



Measuring impact of system noise [UIUC]

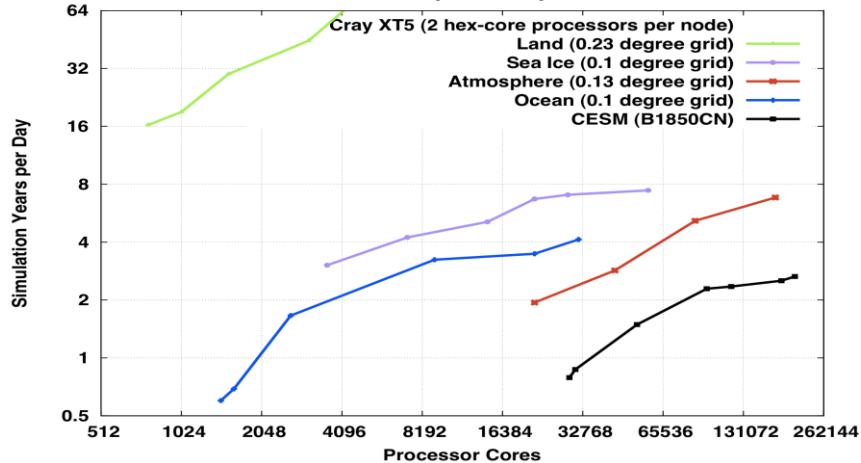
Bidirectional Bandwidth (MPI)



MPI performance evaluation [ORNL]

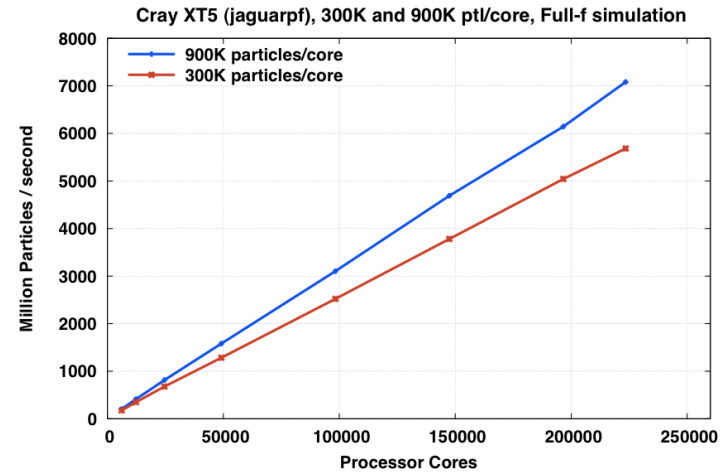
Application analysis and evaluation

CESM (Community Earth System Model)



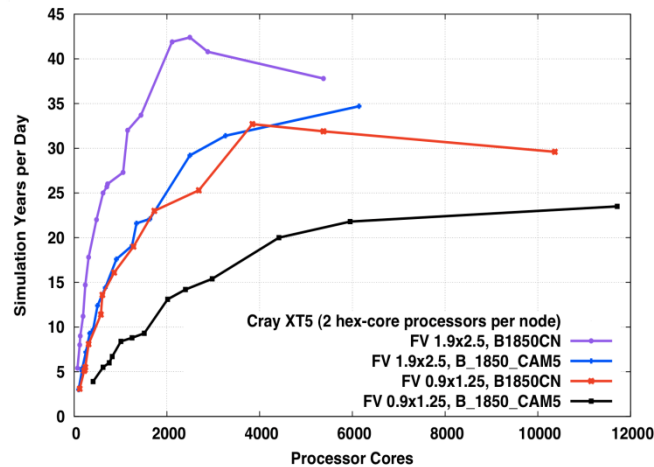
Strong scaling (high resolution) [ORNL/SNL]

XGC1 performance on 3mm ITER grid



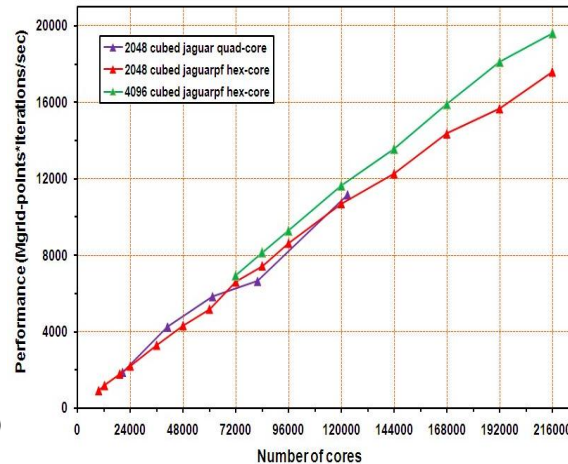
Full system XGC1 weak scaling [ORNL]

CESM (Community Earth System Model)



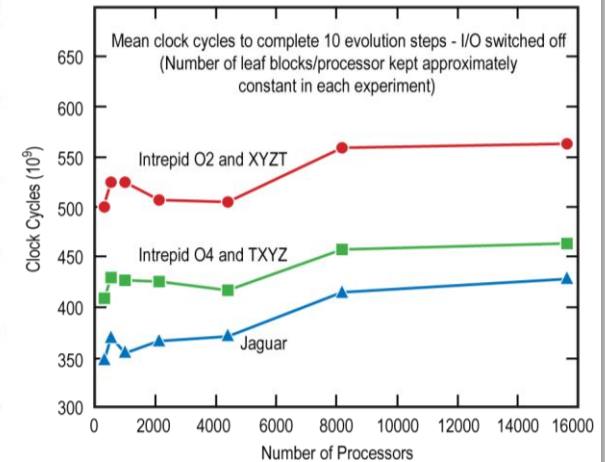
Strong scaling (low resolution) [ORNL]

SBLI (Shock/Boundary Layer Interaction code)



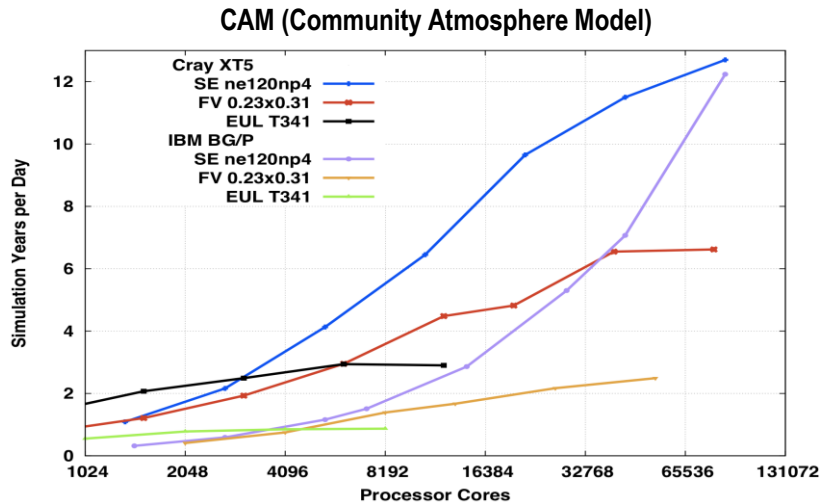
SBLI strong scaling [Daresbury]

FLASH normalized weak scaling

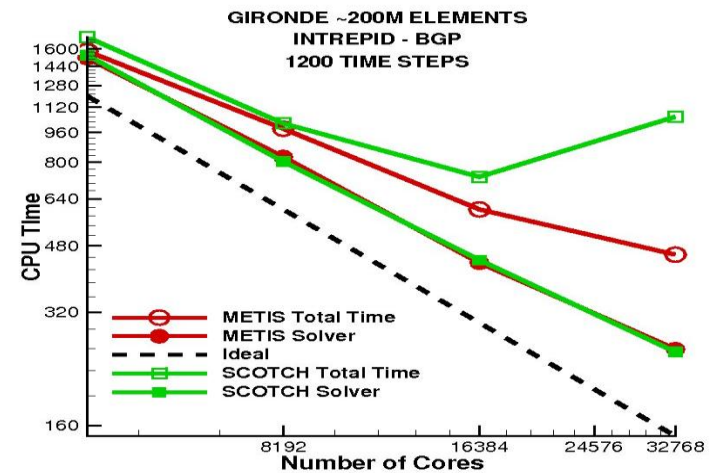


[Univ. of Chicago]

Algorithm analysis and comparison

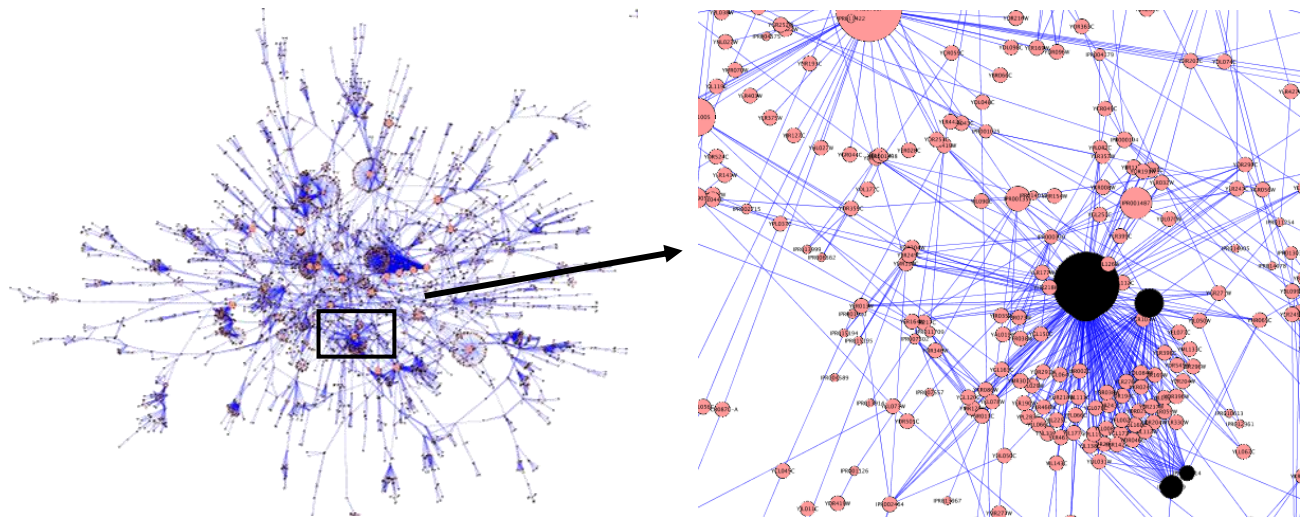


Algorithm and platform comparison [LLNL/ORNL/SNL]



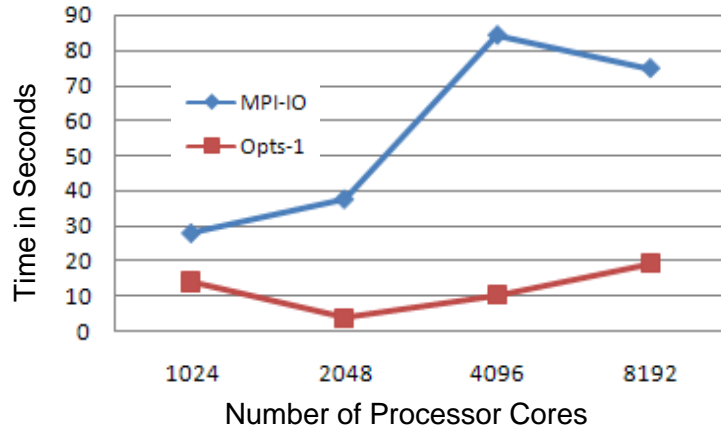
Parallel algorithm comparison [Daresbury]

Processing of genomes into domain maps: need improved load balancing that takes into account scale-free nature of the graphs [RENCI, NCSA]



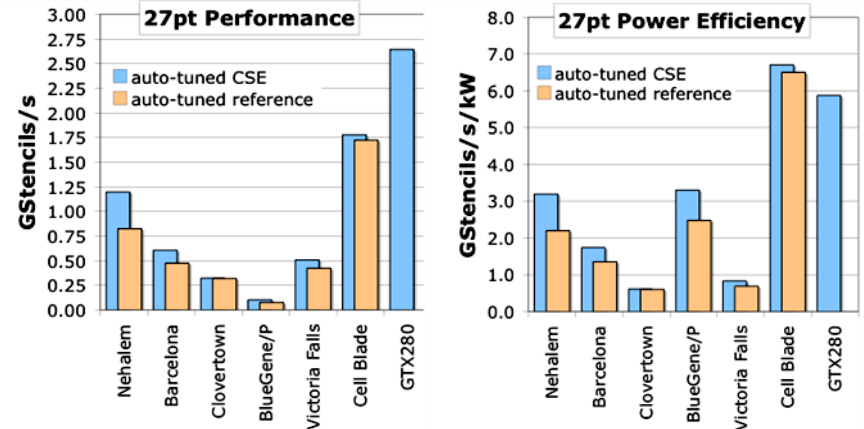
Optimization and optimization technology

HDF5 Write



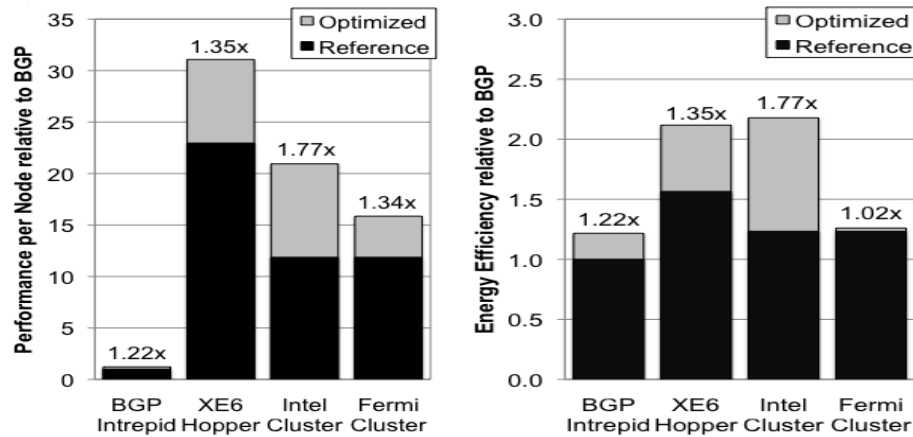
I/O optimization in PFLOTRAN [NCSU]

Finite difference Laplacian stencil benchmark



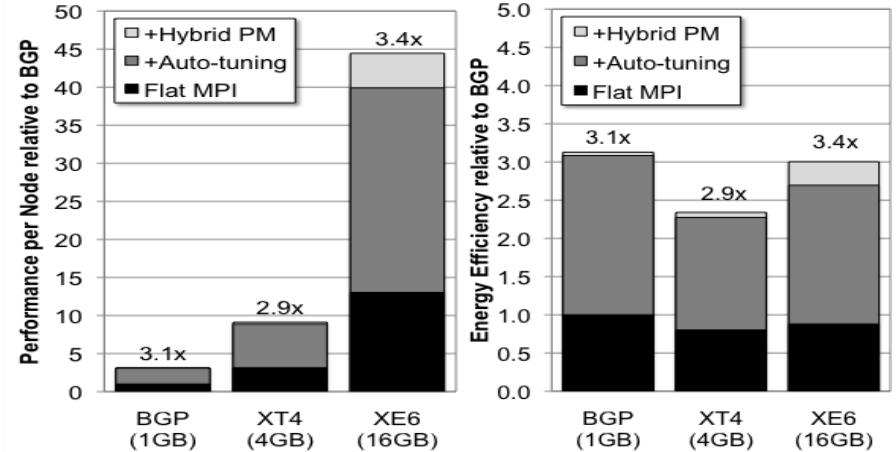
Performance impact of auto-tuning [LBL]

GTC gyrokinetic turbulence simulation code



Optimizing for CPUs and GPUs [LBL]

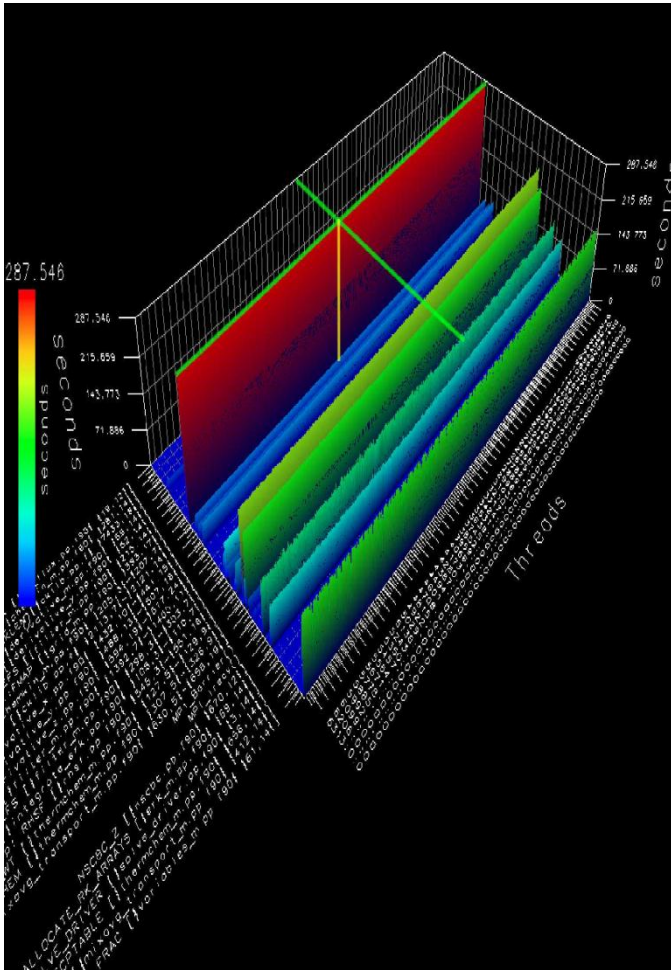
LBMHD Lattice Boltzmann magneto-hydrodynamics code



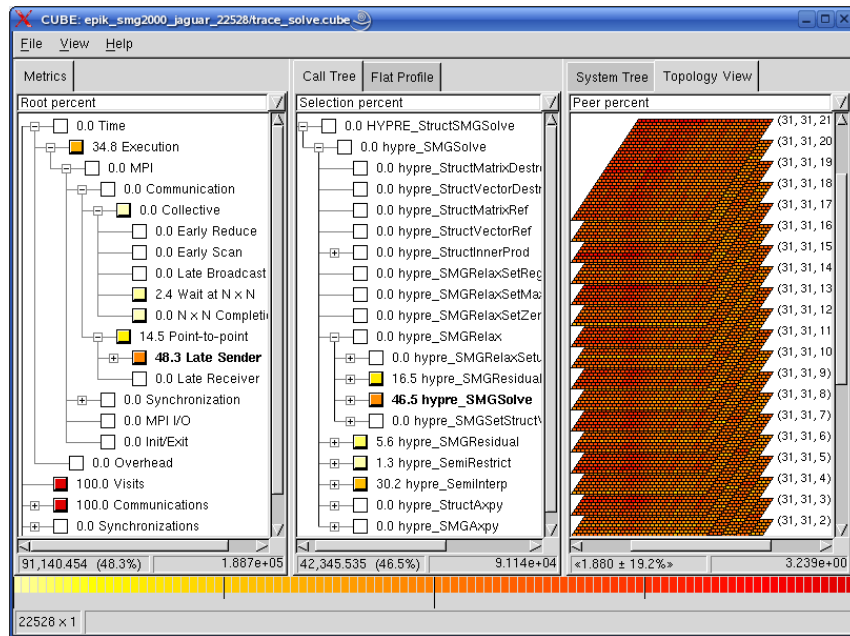
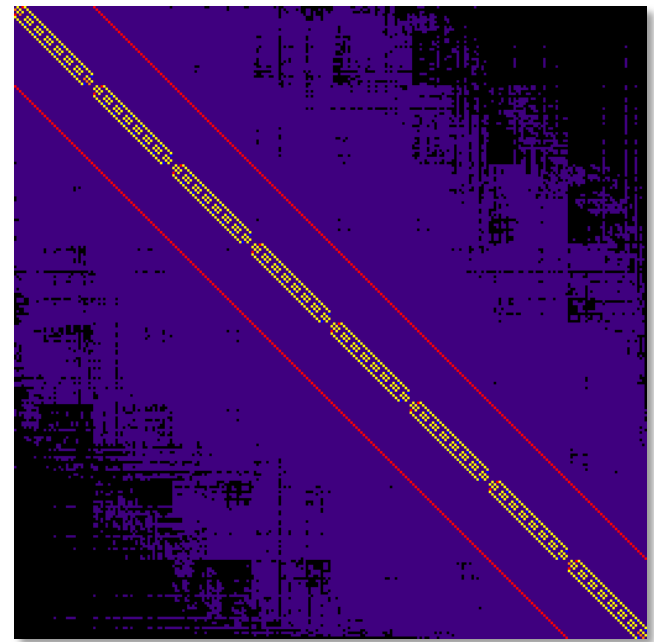
Application of auto-tuning [LBL]

Tool development

TAU graphical display of profile data by thread
(Univ. Oregon)



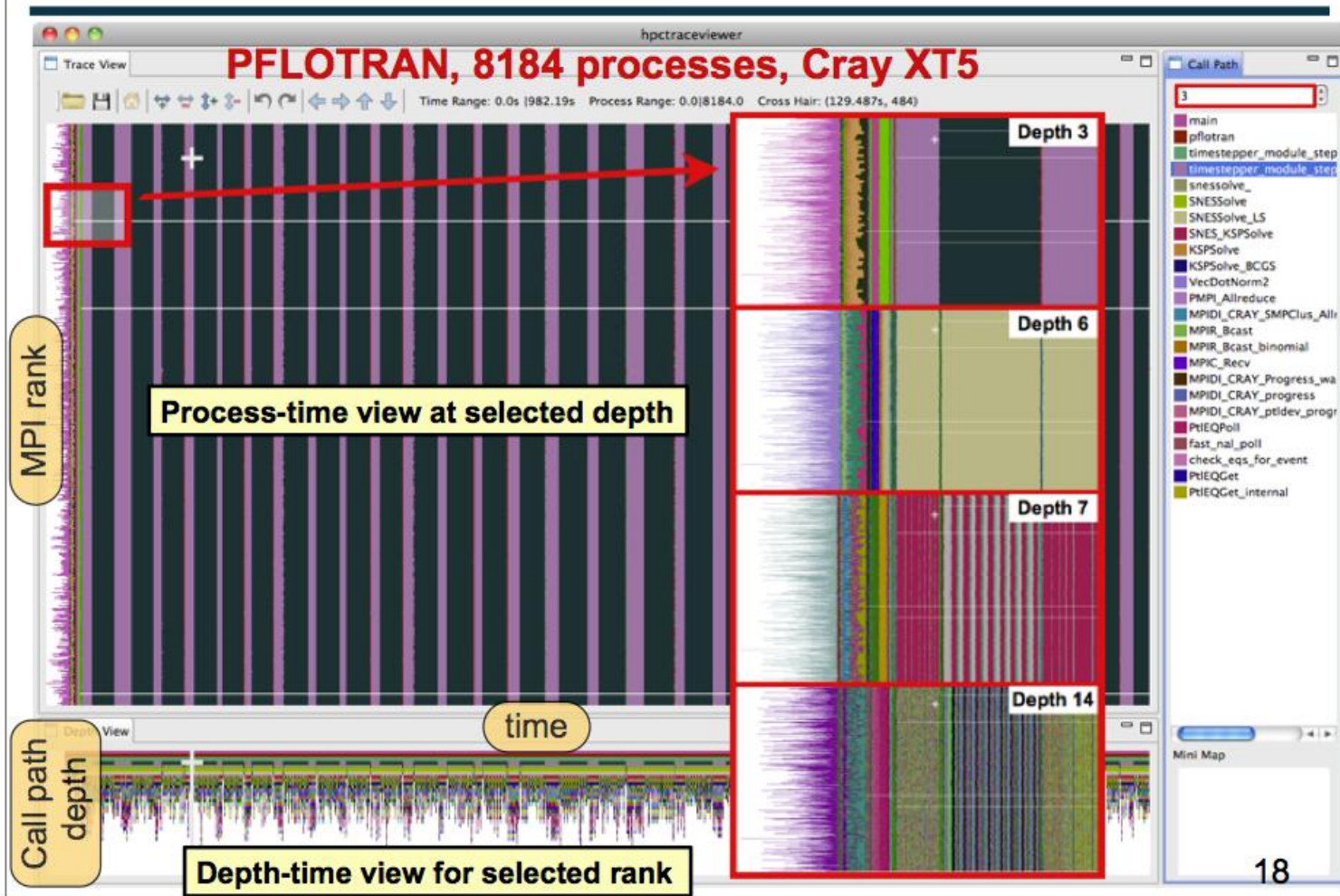
Visualization of point-to-point
communication topology
collected with mpiP
(LLNL, ORNL)



SCALASCA
trace-based
performance
analysis
(FZ-Jülich,
Univ. Tenn.)

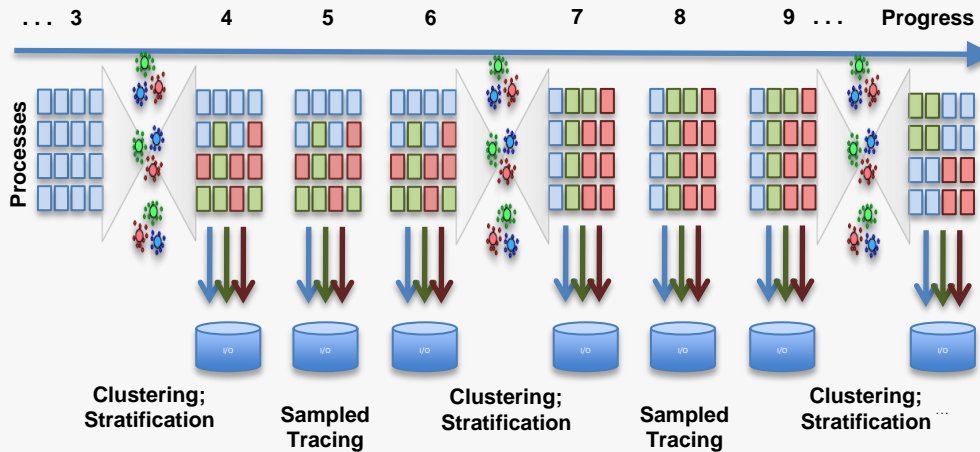
Tool development

Exposing Temporal Call Path Patterns



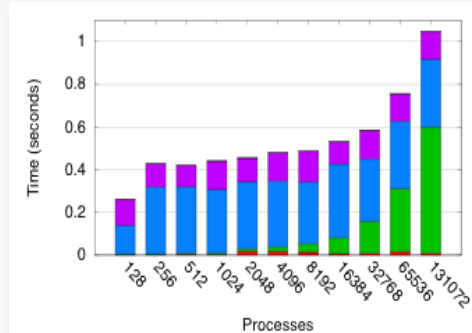
Sample-based tracing and visualization using HPCToolkit [Rice]

Tool development



Using Clustering Algorithm with Parallel Extensions to K-Medoids (CAPEK) to adaptively sample performance traces

(improving performance tracing at scale)
[LLNL, RENC1, Microsoft]



Using performance modeling to predict where performance optimization might be most effective [SDSC, ORNL]

MaxGainExtraIP – improvement potential from increased ILP

reaction.F90

```

2719
2720 ! add contribution to derivatives of total with respect to free
2721 ! bear in mind that the water density portion is scaled below
2722 do j = 1, ncomp
2723   jcomp = reaction%eqcplxspecid(j,icplx)
2724   tmpreal = reaction%eqcplxstoich(j,icplx)*exp(lnQK-ln_conc(jcomp)) / &
2725           rt_auxvar%sec_act_coef(icplx)
2726   do i = 1, ncomp
2727     icomp = reaction%eqcplxspecid(i,icplx)
2728     rt_auxvar%dtotal(icomp,jcomp,iphase) = rt_auxvar%dtotal(icomp,jcomp,iphase) + &
2729             reaction%eqcplxstoich(i,icplx)*tmpreal
2730   enddo
2731 enddo
2732

```

routine total accounts for 36% of improvement potential;
- loop computing dtotal accounts for 22% of improv. potential

false recurrence on dtotal;
- icomp/jcomp indices take distinct values but are loaded from another array

Scopes	MaxGainExtraIP	Exec time	CPU Time
Experiment Aggregate Metrics	5.15e10 100.0	2.69e11 100.0	9.42e10 100.0
reaction_module.rotal_	1.86e10 36.1%	2.94e10 11.1%	2.39e10 25.4%
loop at reaction.F90: 2695-2732	1.80e10 35.0%	2.40e10 9.1%	2.29e10 24.4%
loop at reaction.F90: 2722-2731	1.14e10 22.1%	1.38e10 5.2%	1.38e10 14.6%
loop at reaction.F90: 2715-2718	1.82e09 3.5%	2.07e09 0.8%	2.07e09 2.2%
loop at reaction.F90: 2705-2708	5.55e08 1.1%	9.88e08 0.4%	8.51e08 0.9%
Path 1 (x): 733500.0	7.10e01 0.0%	1.14e02 0.0%	1.00e02 0.0%
Path 3: 1.6137E7	6.90e01 0.0%	1.17e02 0.0%	1.02e02 0.0%
Path 2: 4.76775E7	6.50e01 0.0%	1.09e02 0.0%	9.50e01 0.0%

Co-principal investigators

University of
Illinois at
Urbana-
Champaign

William Gropp



Lawrence
Berkeley
National
Laboratory

David Bailey
Leonid Oliner



Lawrence
Livermore
National
Laboratory

Bronis
de Supinski



Oak Ridge
National
Laboratory

Jeffrey Vetter
Patrick Worley (PI)



Rice
University

John
Mellor-Crummey



University of
California–
Berkeley and
NERSC

Kathy Yelick



University of
California–
San Diego

Allan Snavely



University
of Maryland

Jeffrey
Hollingsworth



University
of Oregon

Allen Malony



University of
Tennessee

Jack Dongarra



University of
Wisconsin

Barton Miller



Contacts

Patrick H. Worley

Future Technologies Group
Computer Science and Mathematics Division
(865) 574-3128
worleyph@ornl.gov

Barbara Helland (OLCF Program Manager)

Betsy Riley (ALCF Program Manager)

Office of Advanced Scientific Computing Research
DOE Office of Science