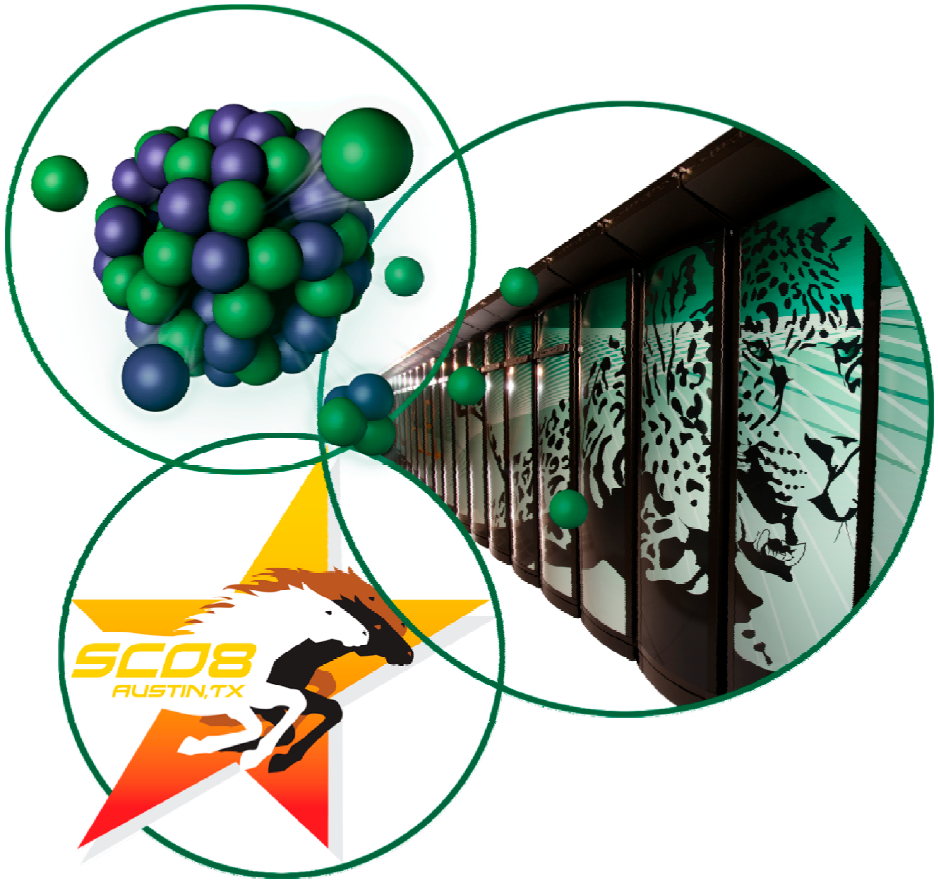
Performance Evaluation and Analysis Consortium (PEAC) End Station



Presented by

Patrick H. Worley

Computational Earth Sciences 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
Performance tools	• Port performance-related tools and system infrastructure to LCF systems and make them available to National Center for Computational Sciences (NCCS) 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 performance characterization and prediction methodologies • Modify methodologies as necessary to improve their utility for predicting resource requirements for production runs on LCF systems

Overview (continued)

Consortium goals (continued)	
Application analysis and optimization	• Analyze performance • Help optimize current and candidate LCF application codes
Performance and application community support	• 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 NCCS 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 performance community:

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

Status as of 9/1/08

11 active ALCF users, 25 active NCCS users

Working on

- System evaluation (15)
- Application performance analysis and optimization (10)
- Tool development (9)
- Application modeling (7)
- Infrastructure development (3)

(many working in multiple areas)

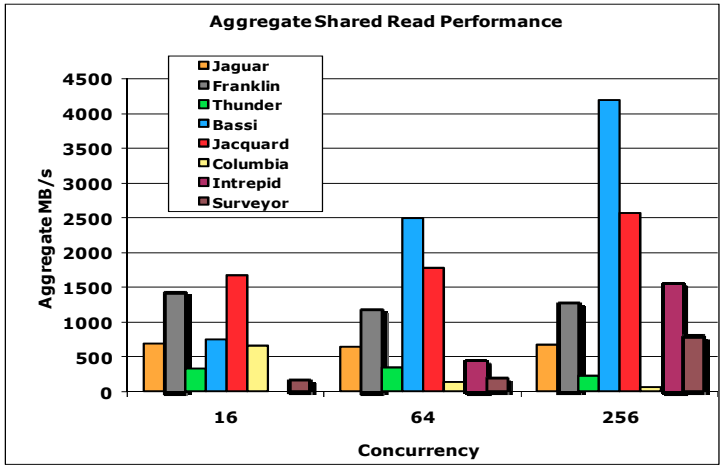
Consuming

- XT4: 400,000 processor hours
- BG/P: 3,200,000 processor hours

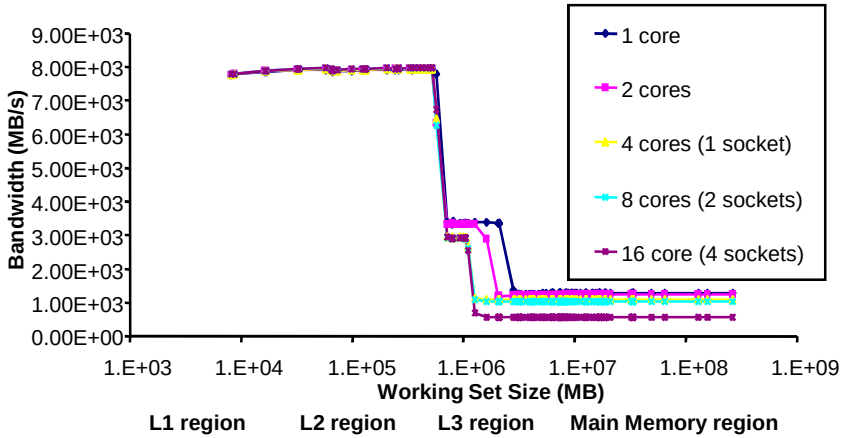
Contributing to

- Three journal papers
- One invited proceedings paper
- Eleven contributed proceedings papers (most refereed)
- Three poster presentations
- Numerous oral presentations
- Five code releases

Subsystem evaluations



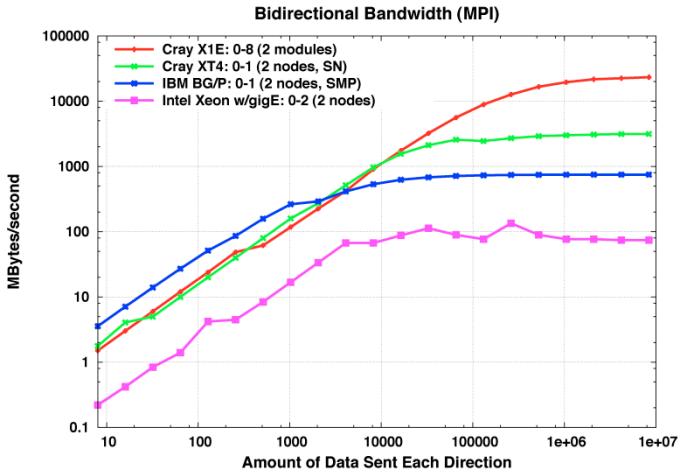
I/O performance characterization (LBL)



Effect of dual- to quad- core upgrade on memory hierarchy bandwidths (SDSC)

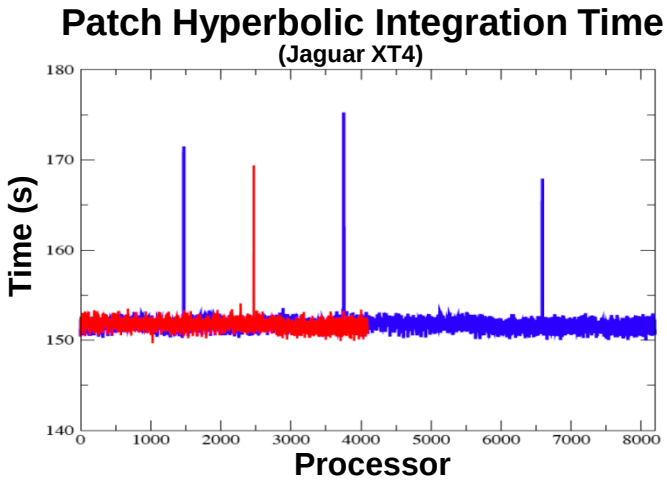
System	4 Neighbors		8 Neighbors	
		Periodic		Periodic
BG/L	2.24		2.01	
BG/L, VN	1.46		1.81	
XT3	7.5	8.1	9.08	9.41
XT4	10.7	10.7	13.0	13.7
XT4 SN	5.47	5.56	6.73	7.06

Identifying performance anomalies (ANL)

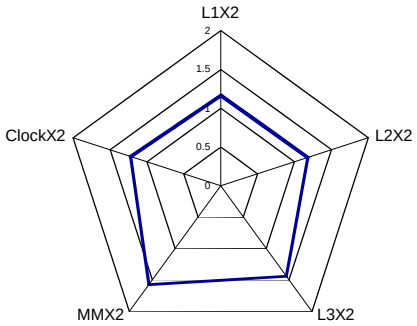


MPI performance evaluation (ORNL)

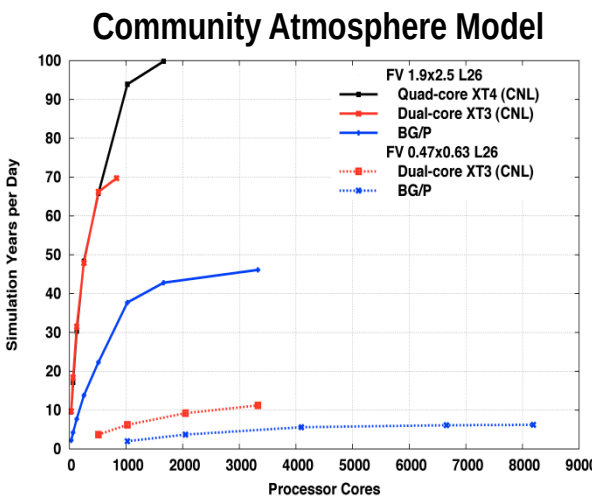
Application analyses and benchmarks

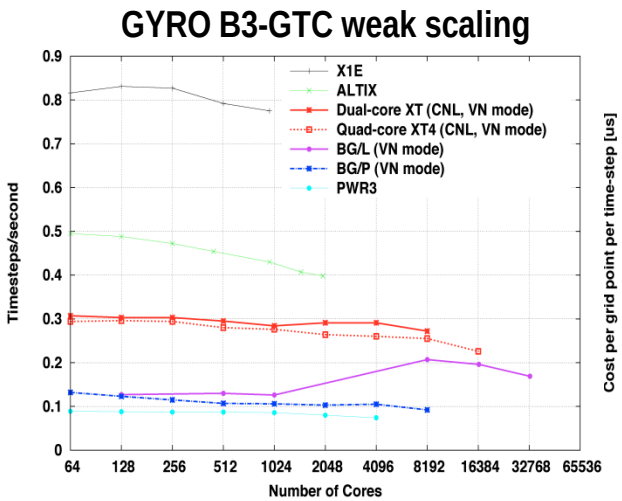


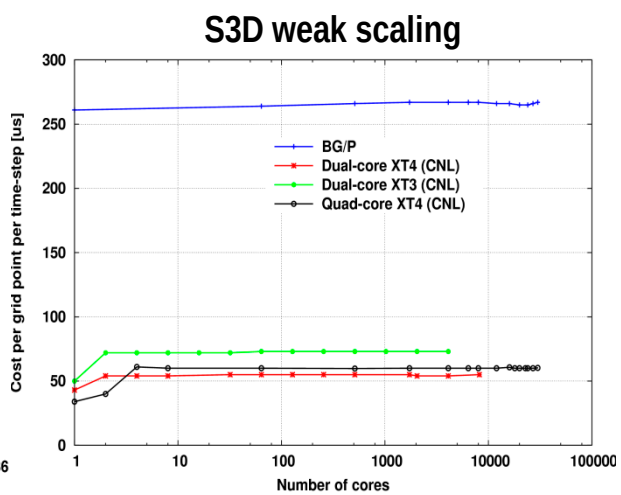
Identifying performance anomalies (LBL)



Performance sensitivities (SDSC)

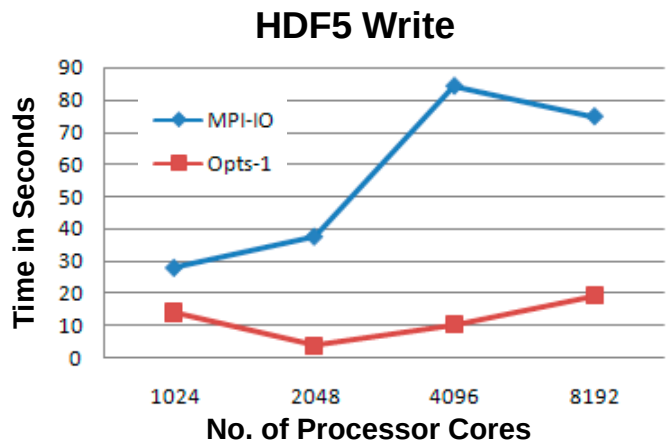




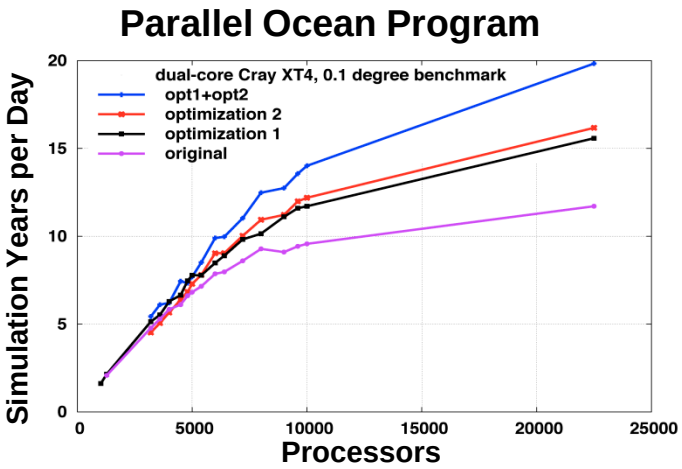


Application benchmarks (ORNL)

Application porting and optimization

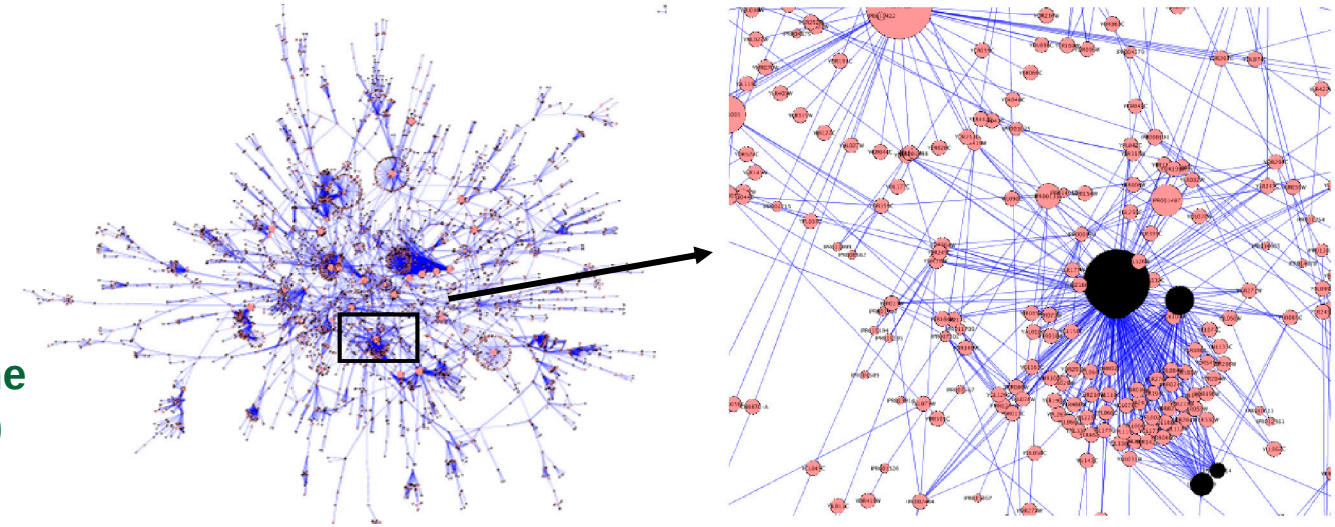


I/O optimization in PFLOTRAN (NCSU)



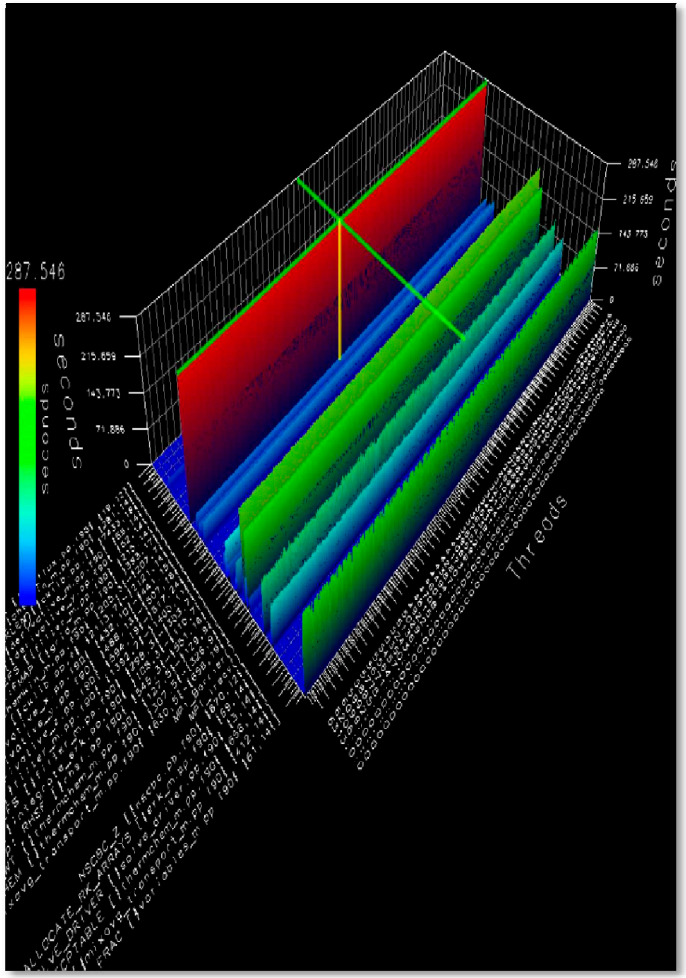
Algorithm optimization (ORNL)

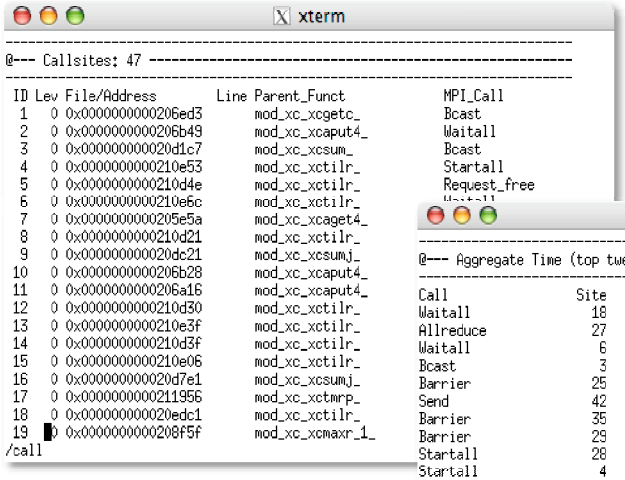
Processing of genomes into domain maps: need improved load balancing that takes into account scale-free nature of the graphs (RENCI/NCSA)



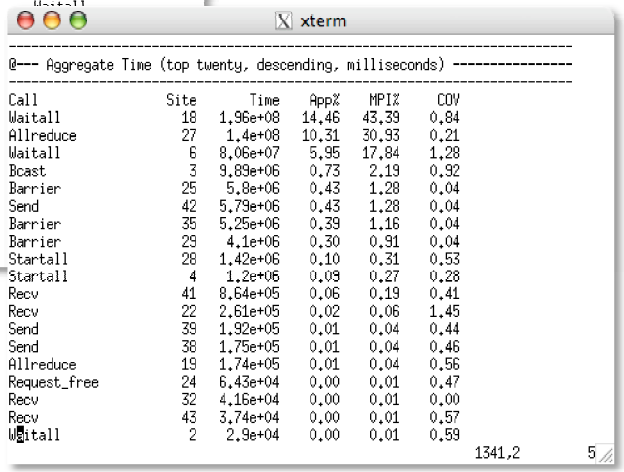
Tool development

TAU graphical display of profile data by thread (UOregon)



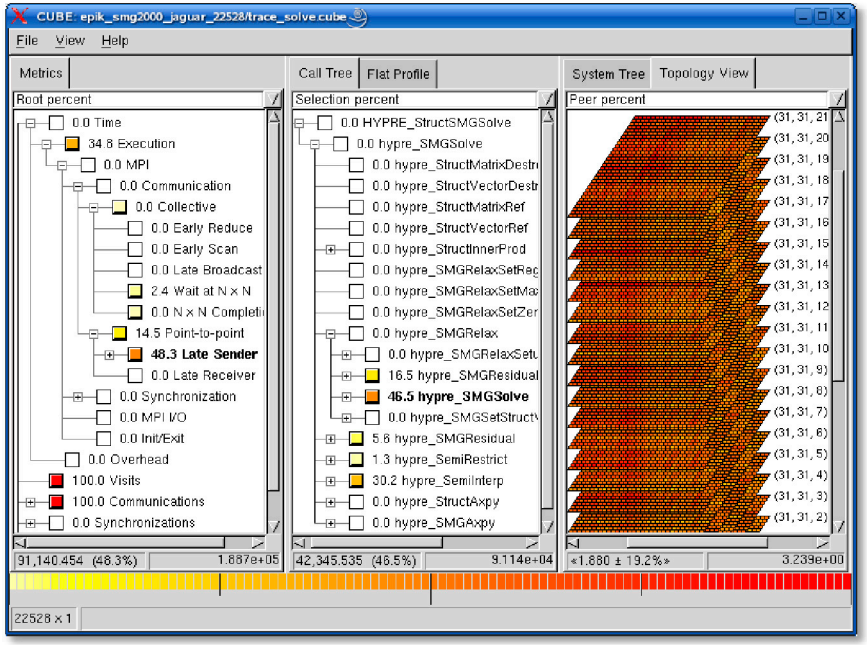


```
xterm
-----------------------------------------------------------------------
@--- Callsites: 47 ---------------------------------------------------
-----------------------------------------------------------------------
 ID Lev File/Address        Line Parent_Funct             MPI_Call
  1   0 0x0000000000206ed3       mod_xc_xcgetc_           Bcast
  2   0 0x0000000000206b49       mod_xc_xcaput4_          Waitall
  3   0 0x0000000000020d1c7      mod_xc_xcsum_            Bcast
  4   0 0x0000000000210e53       mod_xc_xctilr_           Startall
  5   0 0x0000000000210d4e       mod_xc_xctilr_           Request_free
  6   0 0x0000000000210e6c       mod_xc_xctilr_
  7   0 0x0000000000205e5a       mod_xc_xcaget4_
  8   0 0x0000000000210d21       mod_xc_xctilr_
  9   0 0x0000000000020dc21      mod_xc_xcsumj_
 10   0 0x0000000000206b28       mod_xc_xcaput4_
 11   0 0x0000000000206a16       mod_xc_xcaput4_
 12   0 0x0000000000210d30       mod_xc_xctilr_
 13   0 0x0000000000210e3f       mod_xc_xctilr_
 14   0 0x0000000000210d3f       mod_xc_xctilr_
 15   0 0x0000000000210e06       mod_xc_xctilr_
 16   0 0x0000000000020d7e1      mod_xc_xcsumj_
 17   0 0x0000000000211956       mod_xc_xctmrp_
 18   0 0x0000000000020edc1      mod_xc_xctilr_
 19   0 0x0000000000208f5f       mod_xc_xcmaxr_1_
/call
```



```
xterm
-----------------------------------------------------------------------
@--- Aggregate Time (top twenty, descending, milliseconds) ------------
-----------------------------------------------------------------------
Call                Site       Time    App%    MPI%     COV
Waitall               18   1.96e+08   14.46   43.39    0.84
Allreduce             27    1.4e+08   10.31   30.93    0.21
Waitall                6   8.06e+07    5.95   17.84    1.28
Bcast                  3   9.89e+06    0.73    2.19    0.92
Barrier               25    5.8e+06    0.43    1.28    0.04
Send                  42   5.79e+06    0.43    1.28    0.04
Barrier               35   5.25e+06    0.39    1.16    0.04
Barrier               29    4.1e+06    0.30    0.91    0.04
Startall              28   1.42e+06    0.10    0.31    0.53
Startall               4    1.2e+06    0.09    0.27    0.28
Recv                  41   8.64e+05    0.06    0.19    0.41
Recv                  22   2.61e+05    0.02    0.06    1.45
Send                  39   1.92e+05    0.01    0.04    0.44
Send                  38   1.75e+05    0.01    0.04    0.46
Allreduce             19   1.74e+05    0.01    0.04    0.56
Request_free          24   6.43e+04    0.00    0.01    0.47
Recv                  32   4.16e+04    0.00    0.01    0.00
Recv                  43   3.74e+04    0.00    0.01    0.57
Waitall                2    2.9e+04    0.00    0.01    0.59
                                                        1341,2        5
```

mpiP call site profiling (LLNL/ORNL)



SCALASCA trace-based performance analysis (FZ-Jülich, UTenn)

OAK RIDGE National Laboratory

Co-principal investigators

University of Illinois at Urbana-Champaign
William Gropp

Lawrence Berkeley National Laboratory
David Bailey
Leonid Oliker

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

Computational Earth Sciences Group
Computer Science and Mathematics Division
(865) 574-3128
worleyph@ornl.gov

Barbara Helland

Office of Advanced Scientific Computing Research
DOE Office of Science