

# Integrated Performance Monitoring (IPM)

Karl Fuerlinger, Nicholas J. Wright, David Skinner, Katherine Yelick (UC Berkeley & Lawrence Berkeley National Lab)  
Allan Snaveley (San Diego Supercomputing Center)



<http://ipm-hpc.sf.net>

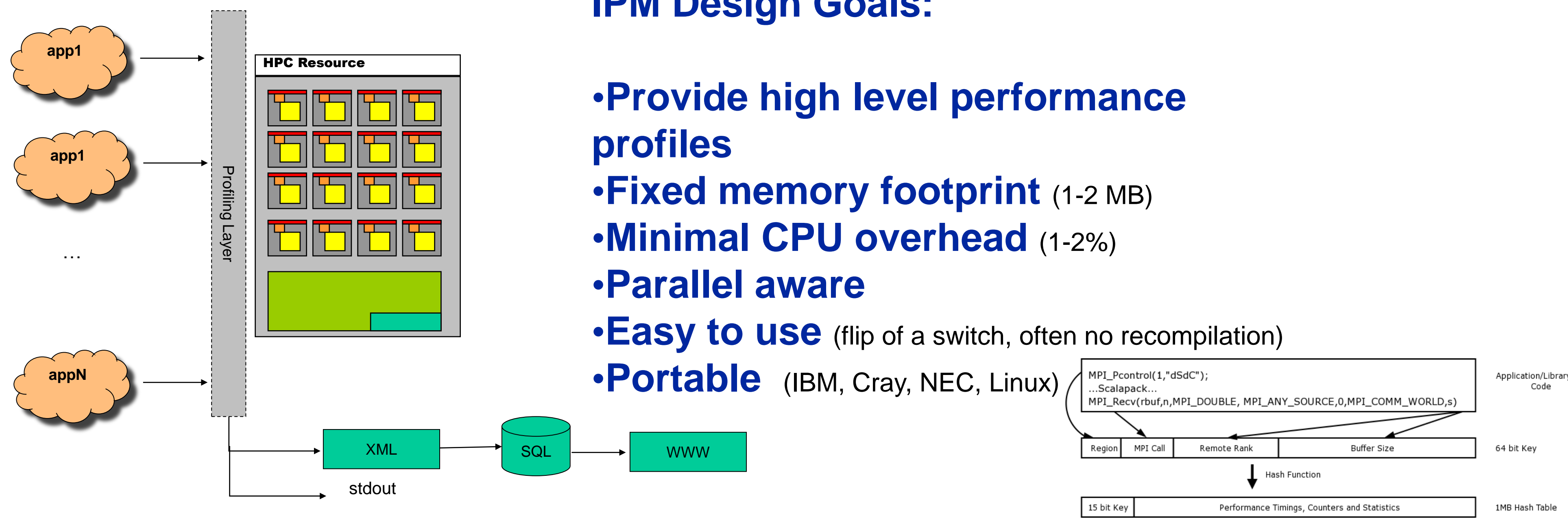


## What is IPM?

It is an approach to performance analysis that is focused on ease of use, scalable lightweight profiling and portability. It is less of a performance "tool" and more of a profiling infrastructure. It serves the needs of users and managers of HPC resources and is available under an open source license. **GPL**

### IPM Design Goals:

- Provide high level performance profiles
- Fixed memory footprint (1-2 MB)
- Minimal CPU overhead (1-2%)
- Parallel aware
- Easy to use (flip of a switch, often no recompilation)
- Portable (IBM, Cray, NEC, Linux)



## What do we mean by profile?

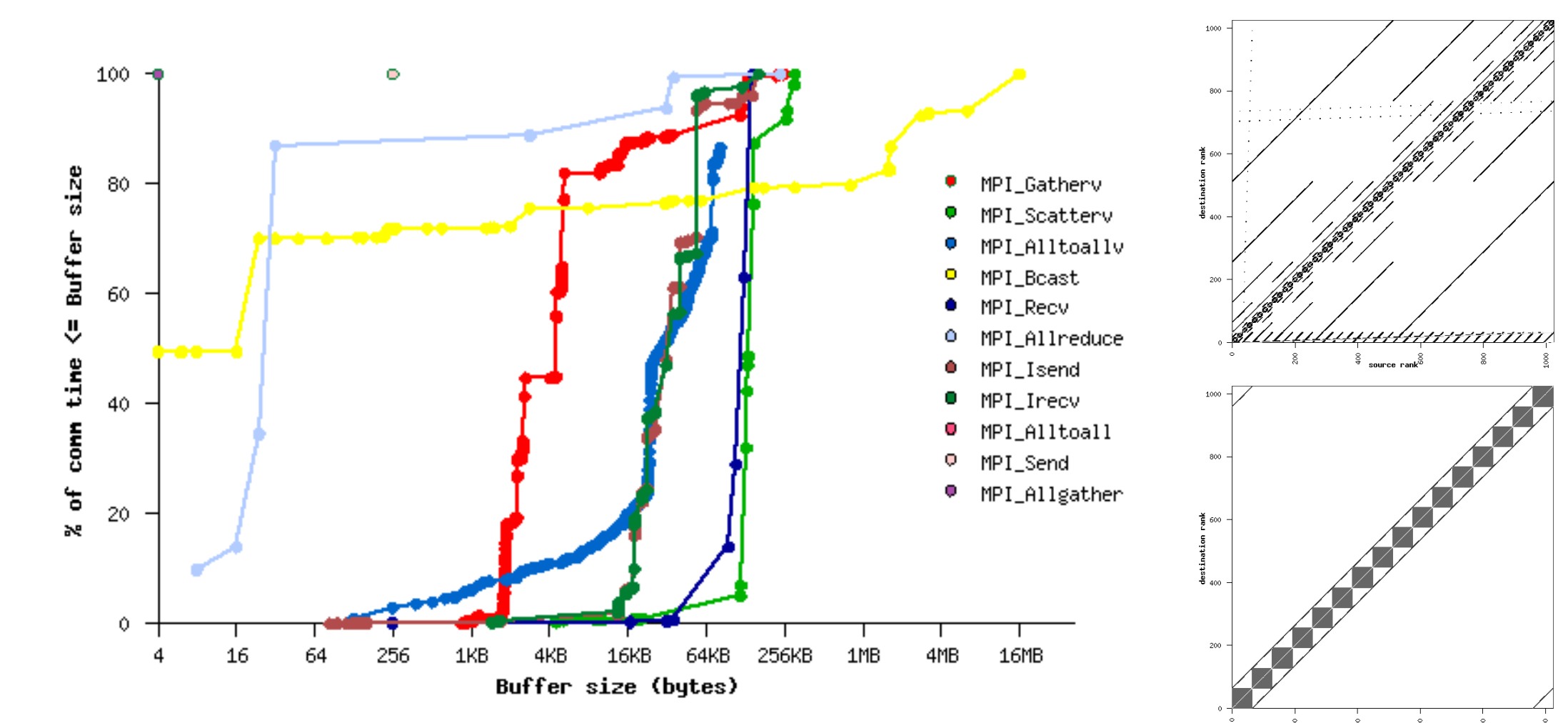
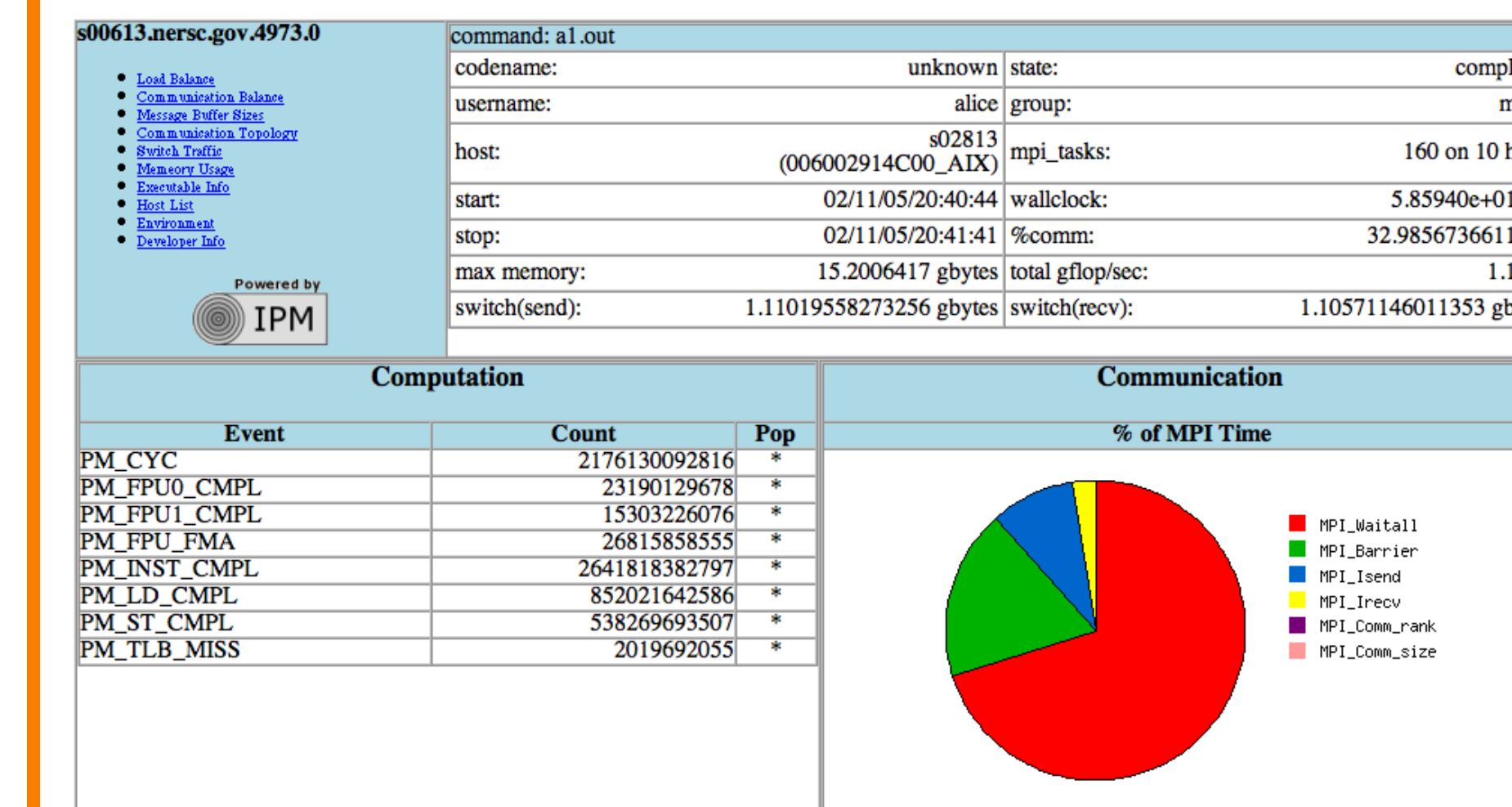


A profile should efficiently illuminate application and architectural aspects of performance for the developer/user in the right level of detail. The default profile is very simple and compact.

...more detail is available...

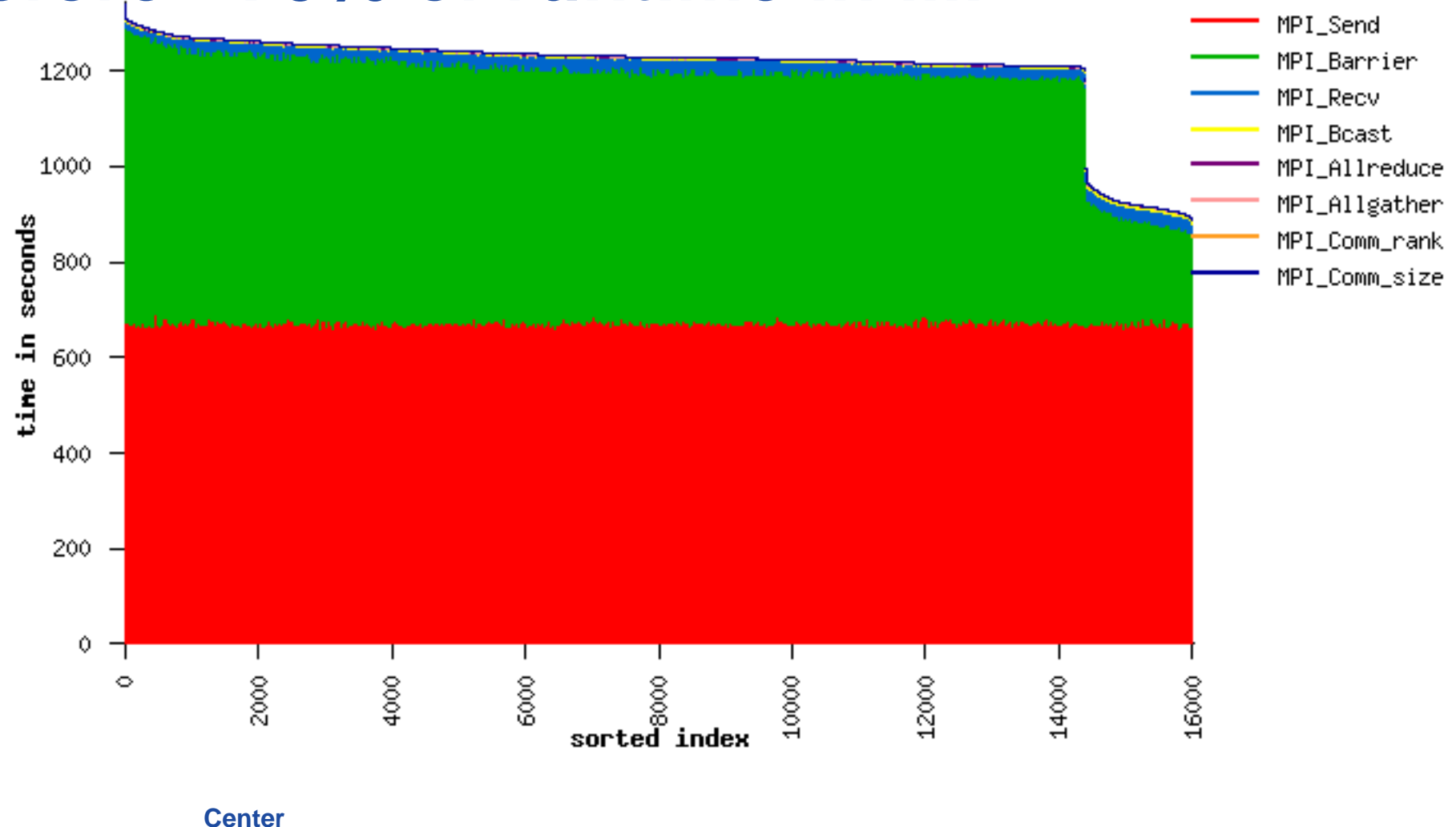
```
##IPMv0.85#####
# command : ./exe/pmcmd -O -c inpcrd -o res (completed)
# host : s05405          mpi_tasks : 64 on 4 nodes
# start : 02/22/05/10:03:55    wallclock : 24.278400 sec
# stop  : 02/22/05/10:04:17    %comm    : 32.43
# gbytes : 2.57604e+00 total    gflop/sec : 2.04615e+00 total
#
# [total]      <avg>      min      max
# wallclock    1373.67    21.1087   24.2784
# user         936.95     14.6398   20.3
# system      227.7      3.55781   5
# mpi         503.853    7.8727   4.2293   9.13725
# %comm       32.4268    17.42     41.407
# gflop/sec   2.04614    0.0319709 0.02724  0.04041
# gbytes      2.57604    0.0402507 0.0399284 0.0408173
# gbyte_tx    0.665125    0.0103926 1.09673e-05 0.0368981
# gbyte_rx    0.659763    0.0103088 9.83477e-07 0.0417372
#
# PM_CYC      3.00519e+11    4.69561e+09 4.50223e+09 5.83342e+09
# PM_FPU_FMA  1.03083e+10    1.61067e+08 1.36815e+08 1.96841e+08
# PM_INST_CMPL 3.33597e+11    5.21245e+09 4.33725e+09 6.44214e+09
# PM_TLB_MISS 1.67892e+08    2.62332e+06 1.16104e+06 2.36664e+07
#
# [time]      [calls]      <impi>      <swall>
# MPI_Bcast   352.365      2816        69.93      22.68
# MPI_Waitany 81.0002      185729     16.08      5.21
# MPI_Allreduce 38.6718     5184        7.68       2.49
# MPI_Allgather 14.7468     448         2.93       0.95
# MPI_Isend   12.9071     185729     2.56       0.83
#####
```

```
##IPMv0.85#####
# command : ./exe/pmcmd -O -c inpcrd -o res (completed)
# host : s05405          mpi_tasks : 64 on 4 nodes
# start : 02/22/05/10:03:55    wallclock : 24.278400 sec
# stop  : 02/22/05/10:04:17    %comm    : 32.43
# gbytes : 2.57604e+00 total    gflop/sec : 2.04615e+00 total
#####
```

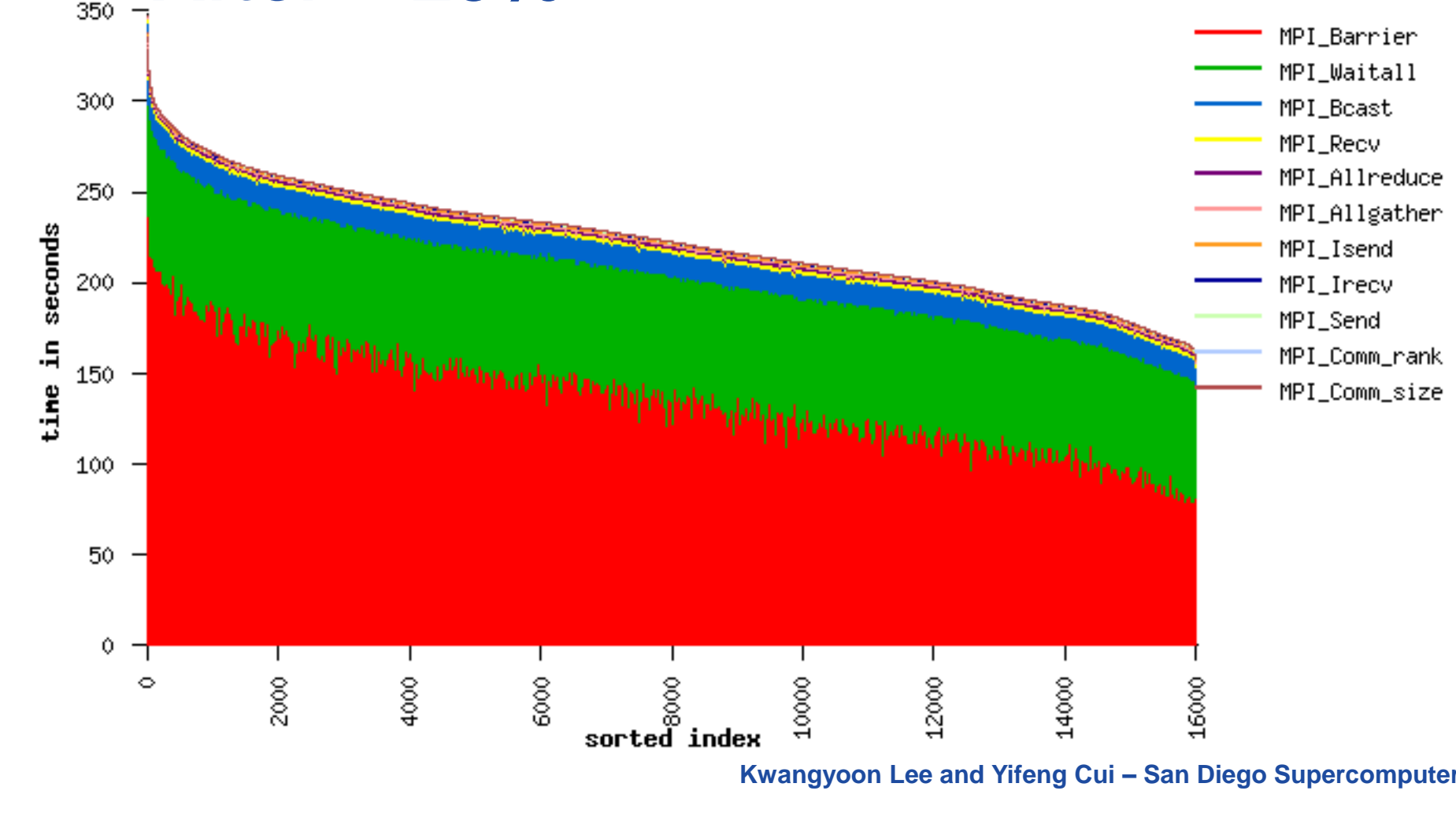


## Improving the Performance of the SCEC code ->2x faster!

Before - 70% of runtime in MPI



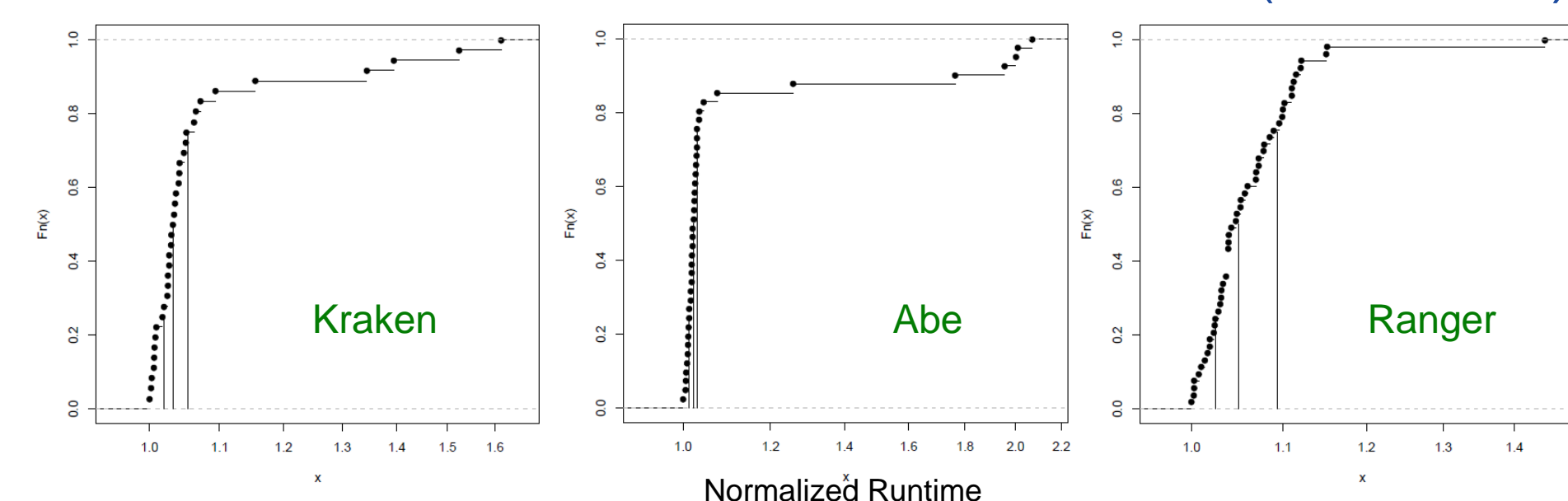
After - 25%



## Using IPM to Interpret Performance Variation

Using IPM and INCA ran WRF benchmark twice a day over the course of a month. Plots show cumulative probably function for the runtime. Analysis of IPM data showed that performance variation was due to network contention and differed significantly between machines and applications.

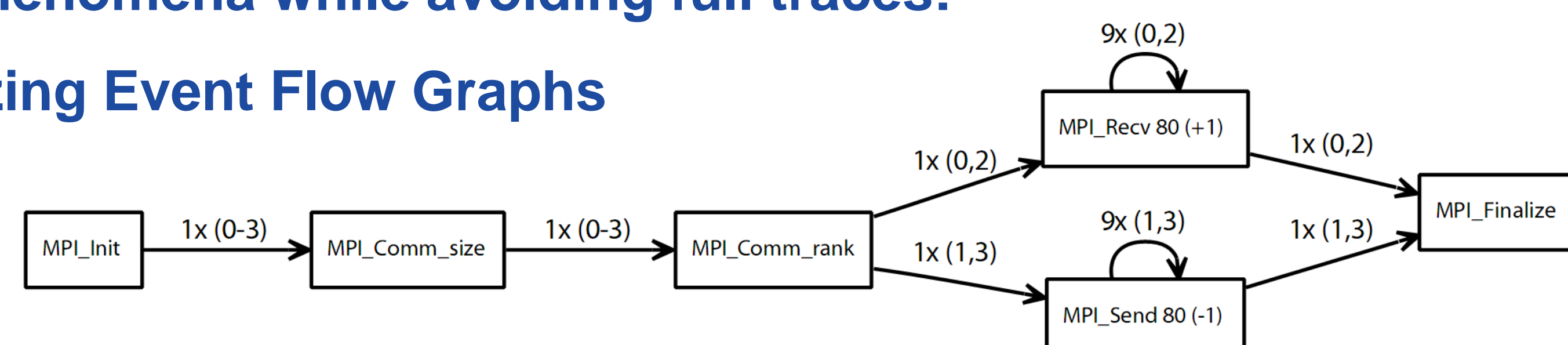
WRF 3D Numerical Weather Prediction (256 Tasks)



Nicholas J. Wright, Shava Smalen, Catherine Olschansky, Jim Hayes, and Allan Snaveley. Dept of Defense High Performance Computing Users Meeting, San Diego June 2009.

## Exploring temporal phenomena while avoiding full traces:

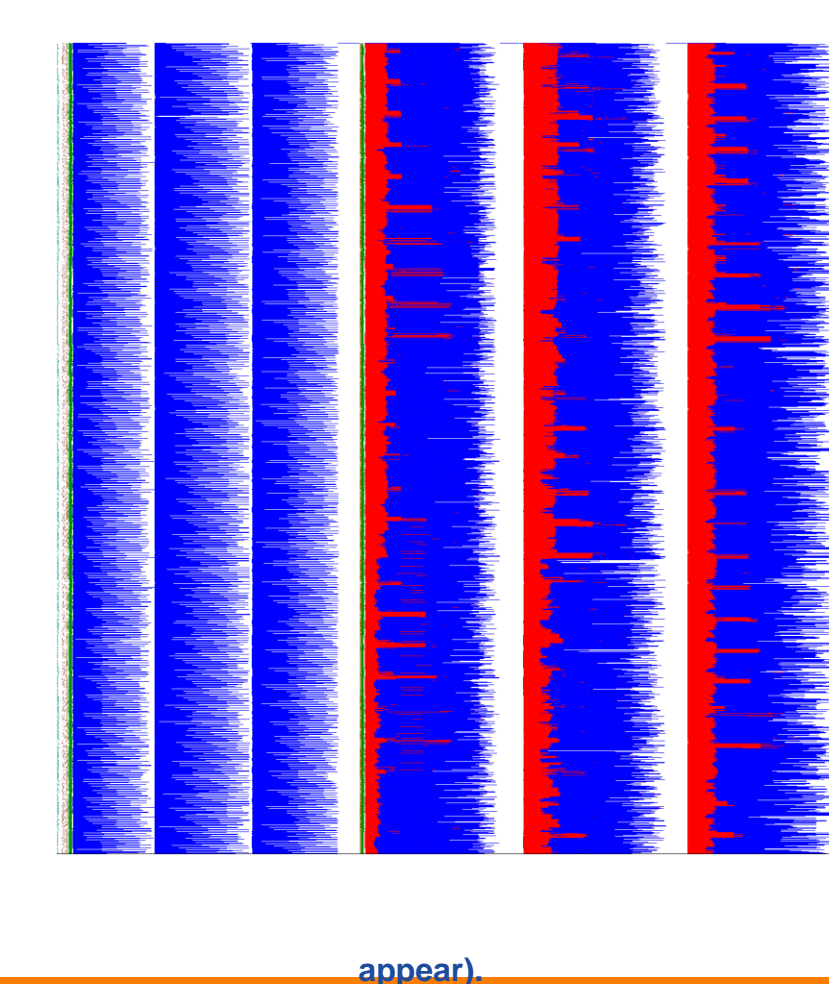
### Recording and Analyzing Event Flow Graphs



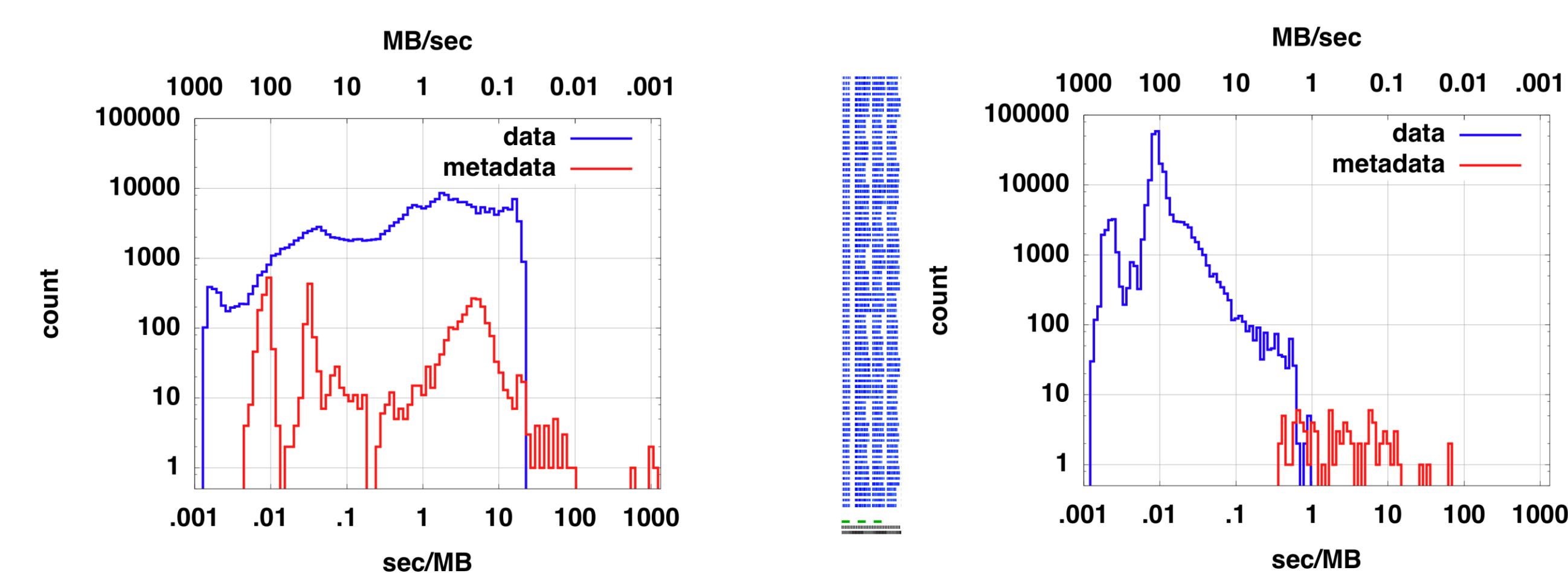
Karl Fuerlinger and David Skinner. Capturing and Visualizing Event Flow Graphs of MPI Applications. Workshop on Productivity and Performance (PROPER 2009) held in conjunction with Euro-Par 2009, Delft, The Netherlands, August 2009.

## Parallel I/O Performance from Events to Ensembles using IPM

Before:



After:



4X Speed-up in the I/O Performance of GCRM!

Andrew Uselton, Mark Howison, Nicholas J. Wright, David Skinner, Noel Keen, John Shalf, Karen L. Karavanic, Leonid Oliker. IPDPS 2010 (To

## SDCI Bringing IPM to the TeraGrid

Installed and tested on all TeraGrid resources.

Online Tutorial Produced in order to reach as many users as possible.

Combined DOE and NSF HPC knowledge

IPM database has 248K application performance profiles!