Portable Performance Characterization of the CCSM with TAU

SC'05

November 12-18, 2005

Seattle, WA

ornl

UNIVERSITY OF OREGON

NCAR

Co-Authors

- National Center for Atmospheric Research (NCAR)
  - George R. Carr Jr, gcarr@ucar.edu
- Oak Ridge National Laboratory (ORNL)
  - John B. Drake, drakejb@ornl.gov
  - Michael W. Ham, hammw@ornl.gov
  - Forrest M. Hoffman, forrest@climate.ornl.gov
  - Patrick H. Worley, worleyph@ornl.gov
- University of Oregon
  - Sameer Shende, sameer@cs.uoregon.edu

... and too many others to list

ornl

UNIVERSITY OF OREGON

NCAR

Acknowledgements

- The National Center for Atmospheric Research is funded by the National Science Foundation.
- ORNL research sponsored by the Atmospheric and Climate Research Division and the Office of Mathematical, Information, and Computational Sciences, Office of Science, U.S. Department of Energy under Contract No. DE-AC05-00OR22725 with UT-Battelle, LLC.
- These slides have been coauthored by a contractor of the U.S. Government under contract No. DE-AC05-00OR22725. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes
- Oak Ridge National Laboratory is managed by UT-Battelle, LLC for the United States Department of Energy under Contract No. DE-AC05-00OR22725.
- The research at the University of Oregon is supported by the Office of Science, U.S. Department of Energy under Contract No. DE-FG02-05ER23680 and DE-FG03-01ER25501.

ornl

UNIVERSITY OF OREGON

NCAR

Overview

- CCSM3 and Standalone CAM Introduction
- Tau Introduction
- Machine Targets
- Some Examples of Analysis

ornl

UNIVERSITY OF OREGON

NCAR

CAM Introduction

- Community Atmospheric Model
- SPMD (single binary)
- Incorporates some land, sea ice, and ocean aspects
- Can be configured as Atmospheric Model for use with CCSM

ornl

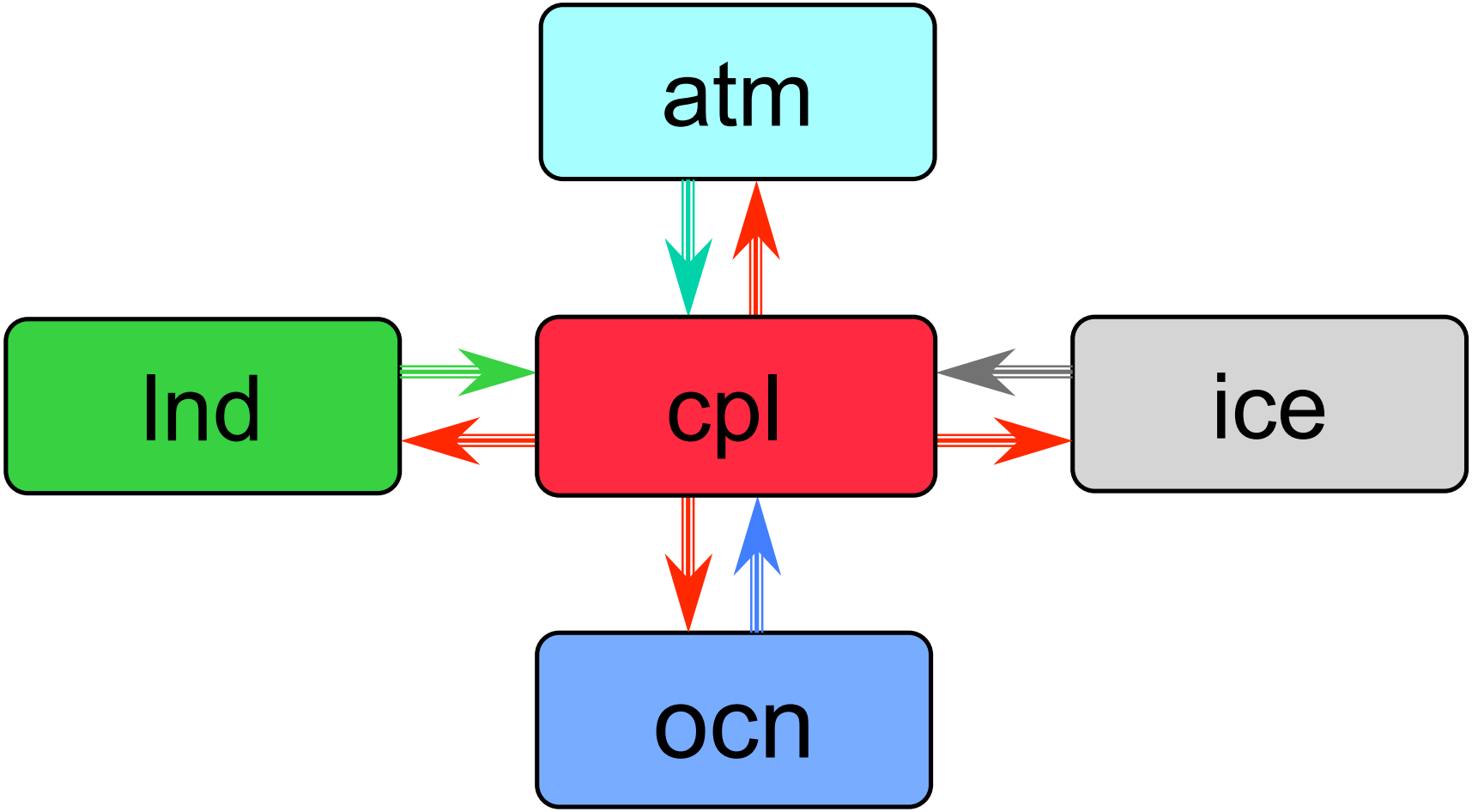
UNIVERSITY OF OREGON

NCAR

CCSM Introduction

- CCSM, the Community Climate System Model, is a coupled model for simulating the earth's climate system.
- Components in CCSM3 include
  - Atmospheric Model – CAM 3.0
  - Ocean Model – modified version of POP 1.4.3
  - Sea Ice Model – CSIM5
  - Land Model – CLM3
  - Coupler - CPL6
- All components use MPI
- Some can use OpenMP on platforms where available
- MPMD (multi-program, multi-data) - 5 different binaries all running on separate processor sets under one MPI environment
- Each component has different scaling and performance aspects

CCSM Hub and Spoke



Supported Machines

- IBM Power3, Power4 - fully validated
- Xeon Linux Clusters (GigE and Myrinet) - validated T31x3
- Cray X1 - recently validated T31x3, just starting T85x1
- SGI Altix - baseline validation complete for T31x3
- Earth Simulator - fully validated on pre-release, update planned
- Opteron Linux Clusters (Myrinet) - work underway
- Xeon Linux Clusters (InfiniBand) - work continuing
- Cray XT3 and XD1 - work begun

NCAR's Bluesky

IBM clustered SMP system

- 1.3 GHz Power4 CPUs
- Colony Switch
- 76 8-way SMP nodes
- 25 32-way SMP nodes
- 2GB memory per CPU

Total of 1600 CPUs,
3.2TBytes of memory,
and 8.32 peak TFLOPs

ornl

UNIVERSITY OF OREGON

NCAR

Phoenix - Cray X1E

Cray X1E with 256 SMP nodes

- 4 Multi-Streaming Processors (MSP) per node
- 4 Single Streaming Processors (SSP) per MSP
- Two 32-stage 64-bit wide vector units running at 1.13 GHz and one 2-way superscalar unit running at 565 MHz per SSP
- 2 MB Ecache per MSP
- 8 GB of memory per node

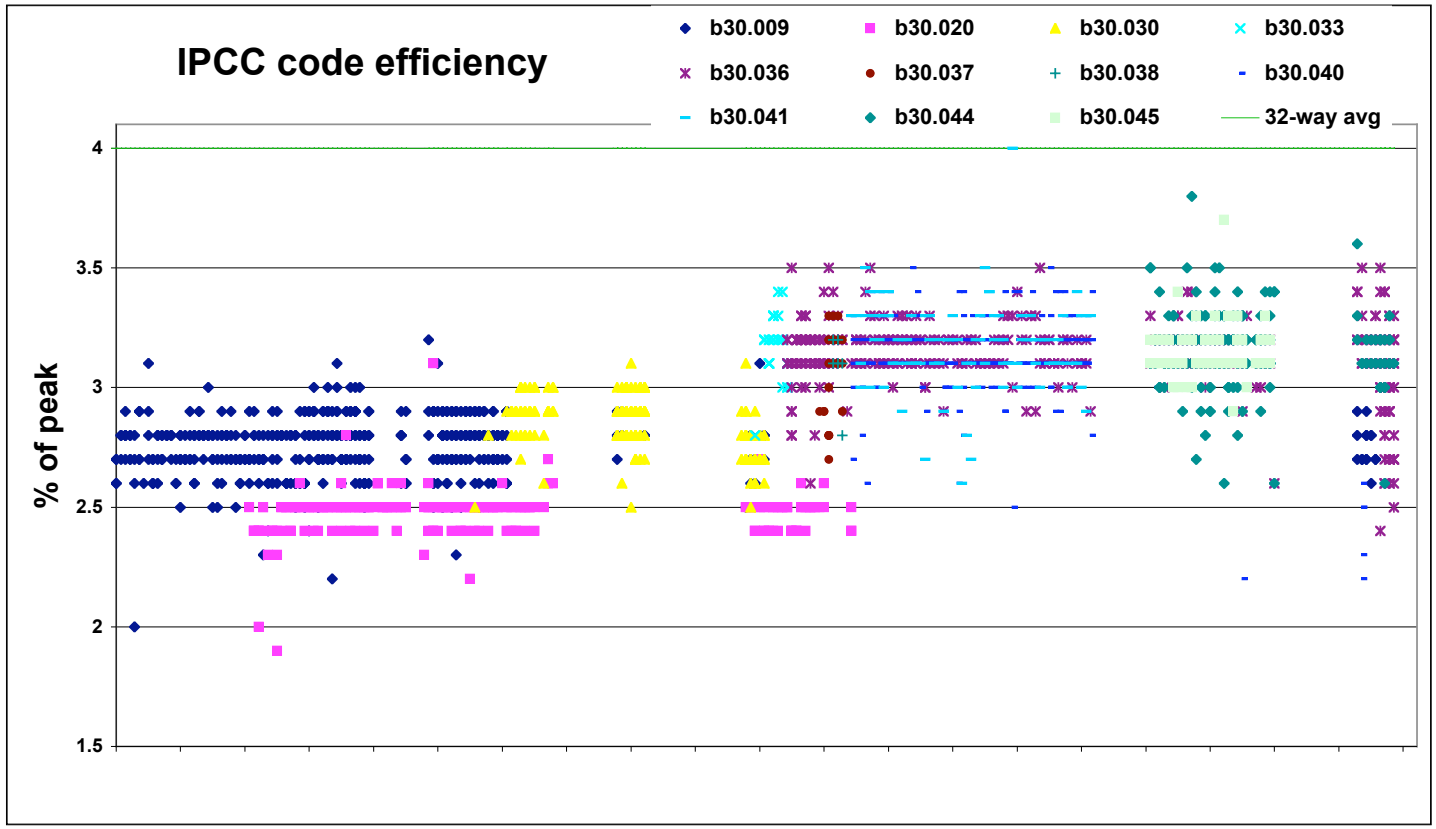
Total of 1024 processors (MSPs), 2 TB of memory, and 18 TF/s peak performance.

ornl

UNIVERSITY OF OREGON

NCAR

CCSM Bluesky Percent of Peak



ornl

UNIVERSITY OF OREGON

NCAR

The Perfomance Characterization Problem

- CCSM is supported on many platforms
  - Need machine independent way to evaluate performance
  - Need easy way to automate "standard" performance characterization as well as complex in-depth analysis
  - CCSM is a complex MPMD application
  - CCSM is currently undergoing massive science changes (finite volume, biogeochemistry)

ornl

UNIVERSITY OF OREGON

NCAR

A Solution: TAU

- Supports both auto instrumentation and user defined events
- Support for MPI and OpenMP
- Available on multiple platforms
- Low impact instrumentation using hardware counter where available
- Easily works into application build process

ornl

UNIVERSITY OF OREGON

NCAR

Project Status

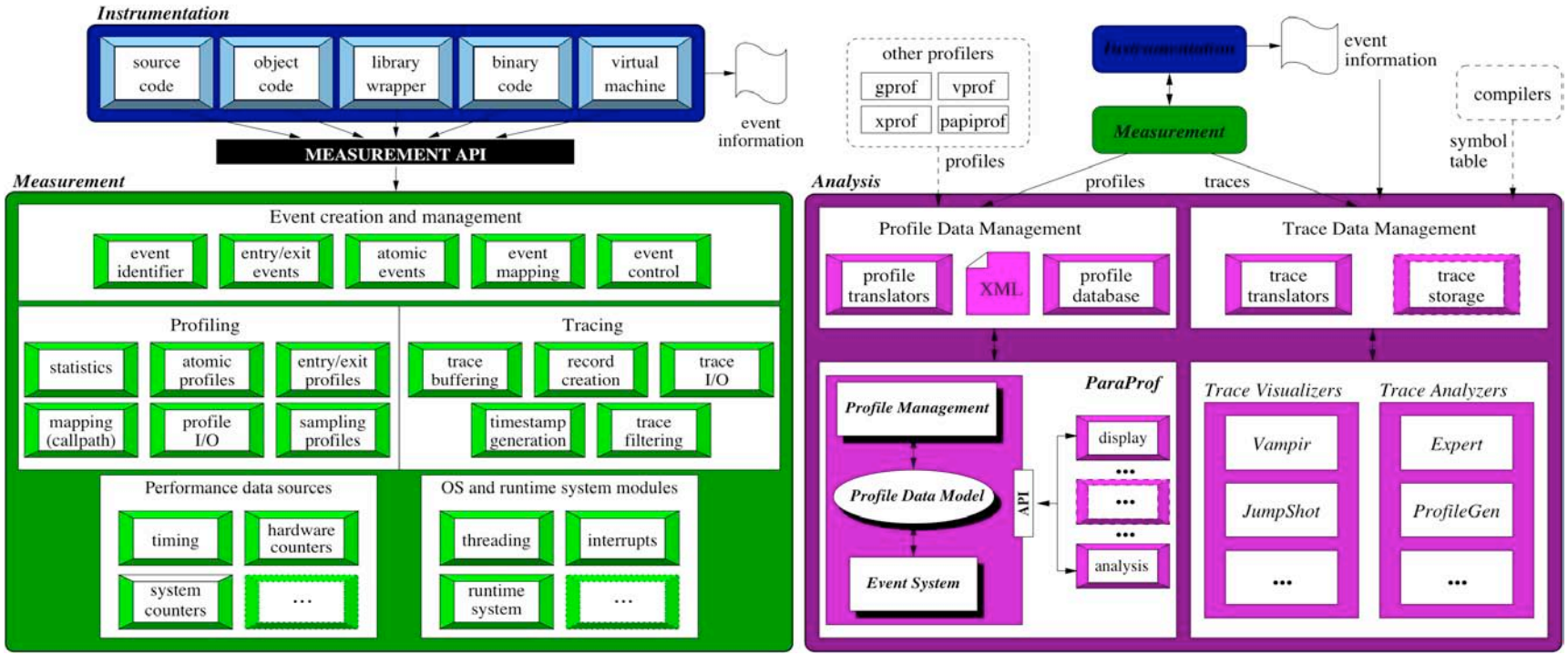
- Tau installed on
  - NCAR's Bluesky (IBM Power4)
  - ORNL's Cheetah (IBM Power4)
  - ORNL's Phoenix (Cray X1E)
  - University of Oregon's NIC (IBM Power4) (???? Regatta)
- In 4 hours on Bluesky
  - Installed Tau
  - Performed Basic Functionality Test
  - Performed First Instrumentation of CAM
- Tau tested
  - CAM MPI only
  - CAM MPI/OMP begun
  - CCSM MPI only

ornl

UNIVERSITY OF OREGON

NCAR

TAU Performance System



- Portable Profiling and Tracing Tools
- Automatic Instrumentation of Code

ornl

UNIVERSITY OF OREGON

NCAR

TAU Features

- Call path profiling
- Compensation of profiling overhead
- Selective instrumentation (exclude/include)
- Throttling of lightweight routines at runtime
- Support for PAPI hardware counters
- Visualizers for TAU traces: Vampir, Jumpshot
- ParaProf scalable profile browser
- Modular, configurable measurement library
- Easy integration in build systems

ornl

UNIVERSITY OF OREGON

NCAR

TAU Modifications in CAM

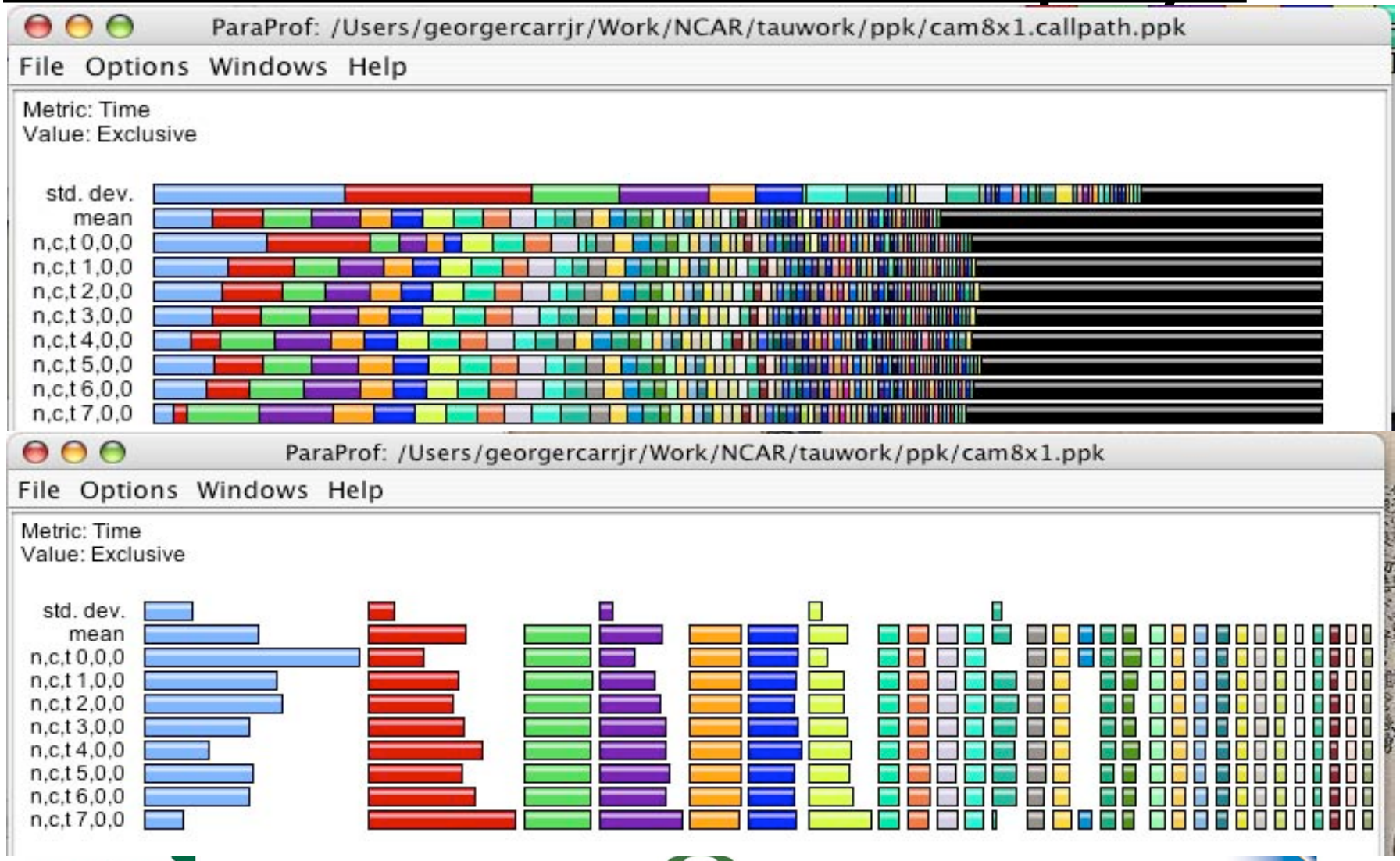
- Choose TAU measurement library/stub makefile:
  - `export TAU_MAKEFILE=/home/bluesky/sshende/tau2/ibm64/lib/Makefile.tau-nocomm-mpi-pdt`
  - `export TAU_OPTIONS='-optVerbose -optPdtF95Opts="-p -DCAM -DNO_SHR_VMATH -DHIDE_SHR_MSG -DAIX" -optTauSelectFile=/home/bluesky/sshende/camscripts/select.tau'`
- Throttle instrumentation in lightweight routines
  - `export TAU_THROTTLE=1`
  - `export TAU_THROTTLE_NUMCALLS=100000`
  - `export TAU_THROTTLE_PERCALL=10`
  - Disables instrumentation in any routine that executes over 100000 times and has an inclusive time per call of less than 10 microseconds!
- Choose TAU's compiler script
  - `Original: $cfgdir/configure -dyn eul -res 64x128 -spmd`
  - `With TAU: $cfgdir/configure -dyn eul -res 64x128 -spmd -fc tau_f90.sh`
  - No Changes to source code or configuration system!

ornl

UNIVERSITY OF OREGON

NCAR

Normal and Stacked Displays

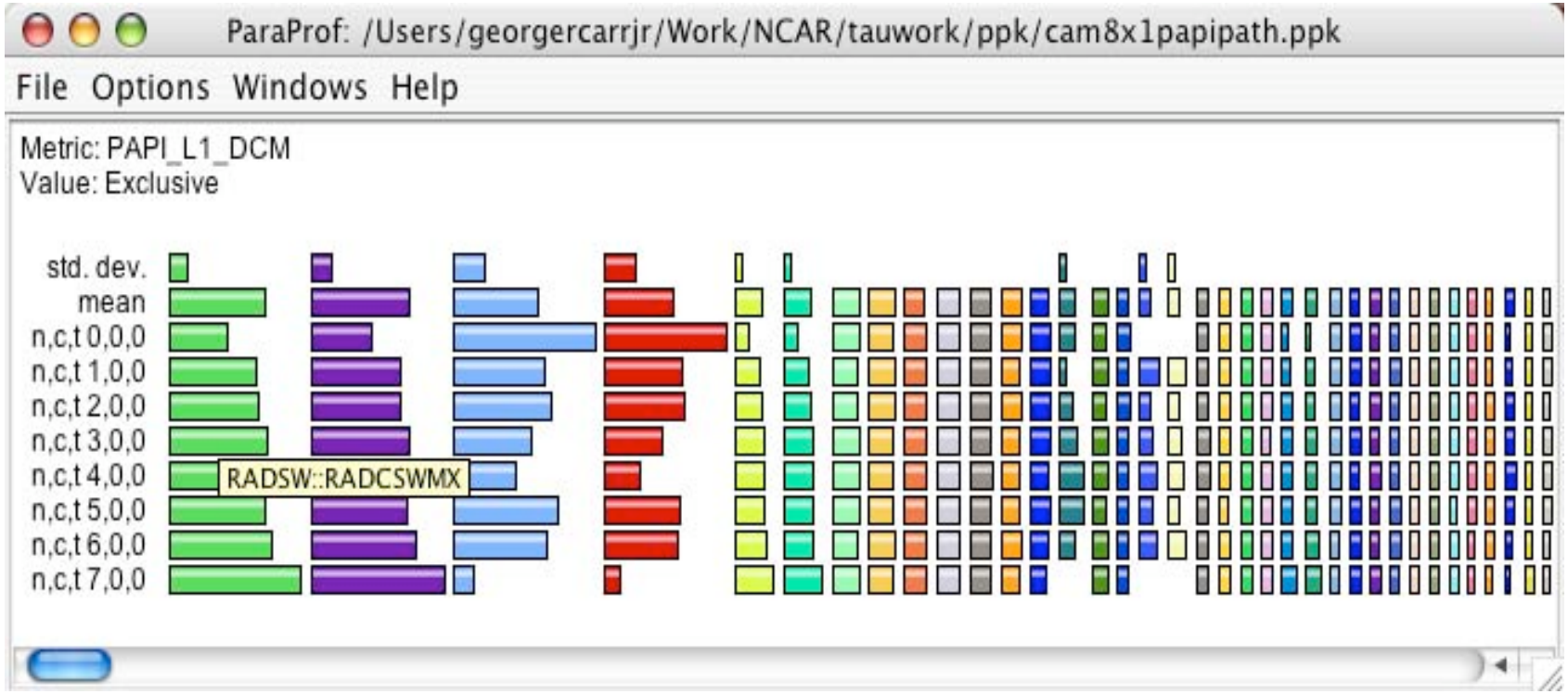


ornl

UNIVERSITY OF OREGON

NCAR

Using PAPI to Get Hardware Counters

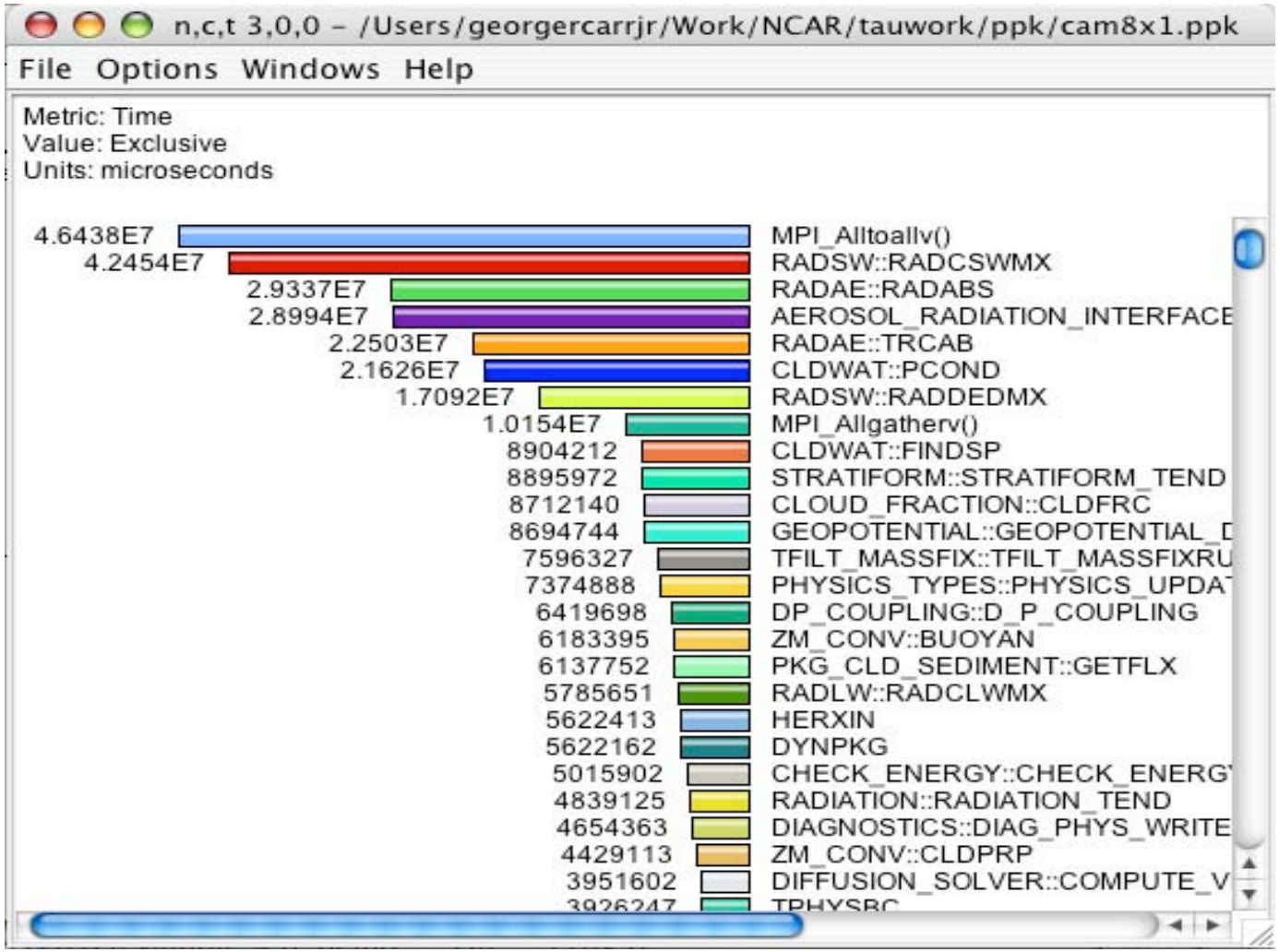


ornl

UNIVERSITY OF OREGON

NCAR

Call Thread Bar Graph



ornl

UNIVERSITY OF OREGON

NCAR

CallPath Display

Mean Call Path Data – cam8x1.callpath.comp.ppk/

File Options Windows Help

Metric Name: Time
Sorted By: Exclusive
Units: seconds

Parents

Routine

Children

```
      Exclusive        Inclusive      Calls/Tot.Calls     Name[id]
      ------------------------------------------------------------------------------

         46.233           46.233      361/1445            CLM_CAMMOD::LP_ALL2ALL_CHUNK_TO_CLUMP[1431]
          0.636            0.636      361/1445            CLM_CAMMOD::LP_ALL2ALL_CLUMP_TO_CHUNK[1432]
        7.8E-4           7.8E-4       1/1445              CLM_CAMMOD::LP_ALL2ALL_CLUMP_TO_CHUNK_INIT[1433]
          6.89             6.89       722/1445            MPIALLTOALLV[1693]
-->      53.759           53.759       1445                MPI_Alltoallv()[1710]

          43.15            94.325      7680/7680           RADIATION::RADIATION_TEND[1824]
-->       43.15            94.325      7680                RADSW::RADCSWMX[1829]
           0.36             0.36       7342.5/7342.5       AEROSOL_RADIATION_INTERFACE::AER_LAYER_MASS_GET[25]
         28.439           28.439       139508/139508       AEROSOL_RADIATION_INTERFACE::AER_LAYER_RAD_PROPS_GET[26]
          0.087            0.087       7680/7680           AEROSOL_RADIATION_INTERFACE::SW_DIAGNOSTICS_INIT[30]
          0.071            0.14        7342.5/7342.5       CMPARRAY_MOD::CMPDAYNITE_1D_I[1442]
          0.314            0.314       44055/44055         CMPARRAY_MOD::CMPDAYNITE_1D_R_COPY[1443]
          0.071            0.361       7342.5/7342.5       CMPARRAY_MOD::CMPDAYNITE_2D_R[1444]
          0.673            0.673       73425/73425         CMPARRAY_MOD::CMPDAYNITE_2D_R_COPY[1445]
          0.076            0.154       7342.5/7342.5       CMPARRAY_MOD::EXPDAYNITE_1D_I[1448]
          1.186            1.186       29370/60090         CMPARRAY_MOD::EXPDAYNITE_2D_R[1449]
          0.501            1.154       58375.4/430597      QUICK_SORT::QUICK_SORT_1[1800]
         18.308           18.308       139508/139508       RADSW::RADDEDMX[1830]

         29.495           58.421       640/640             RADLW::RADCLWMX[1827]
-->      29.495           58.421       640                 RADAE::RADABS[1802]
         25.211           25.211       449280/449280       RADAE::TRCAB[1807]
          3.715            3.715       66560/66560         RADAE::TRCABN[1808]

         28.439           28.439       139508/139508       RADSW::RADCSWMX[1829]
-->      28.439           28.439       139508              AEROSOL_RADIATION_INTERFACE::AER_LAYER_RAD_PROPS_GET[26]
```

ornl

UNIVERSITY OF OREGON

NCAR

Thread Statistics Text

n,c,t, 3,0,0 - /Users/georgercarrjr/Work/NCAR/tauwork/ppk/cam8x1.ppk

File Options Windows Help

Metric: Time
Sorted By: Exclusive
Units: microseconds

%Total Time	Exclusive	Inclusive	#Calls	#Child Calls	Inclusive/Call	Name
9.8	4.6438E7	4.6438E7	1445	0	32136.857	MPI_Alltoallv()
19.1	4.2454E7	9.0665E7	7680	498394	11805.28	RADSW::RADCSWMX
11.6	2.9337E7	5.5137E7	640	515840	86150.85	RADAE::RADABS
6.1	2.8994E7	2.8994E7	145901	0	198.721	AEROSOL_RADIATION_IN
4.7	2.2503E7	2.2503E7	449280	0	50.086	RADAE::TRCAB
6.6	2.1626E7	3.1227E7	23104	123105	1351.569	CLDWAT::PCOND
3.6	1.7092E7	1.7092E7	145901	0	117.145	RADSW::RADDEDMX
2.1	1.0154E7	1.0154E7	722	0	14063.115	MPI_Allgatherv()
1.9	8904212	8904212	23104	0	385.397	CLDWAT::FINDSP
13.8	8895972	6.5369E7	23104	323456	2829.337	STRATIFORM::STRATIFO
2.0	8712140	9652591	46208	46208	208.894	CLOUD_FRACTION::CLDF
1.8	8694744	8694744	300352	0	28.949	GEOPOTENTIAL::GEOPOT
1.8	7596327	8408582	2888	8664	2911.559	TFILT_MASSFIX::TFILT
3.4	7374888	1.607E7	462080	300352	34.777	PHYSICS_TYPES::PHYSI
1.7	6419698	7942025	361	92416	22000.069	DP_COUPLING::D_P_COU
1.3	6183395	6183395	23104	0	267.633	ZM_CONV::BUOYAN
2.0	6137752	9356218	46208	46208	202.48	PKG_CLD_SEDIMENT::GE
13.8	5785651	6.5378E7	7680	40686	8512.7	RADLW::RADCLWMX
1.2	5622413	5622413	8760	0	641.828	HERXIN
16.5	5622162	7.8591E7	361	40889	217702.795	DYNPKG
1.1	5015902	5015902	138624	0	36.184	CHECK_ENERGY::CHECK
35.7	4839125	1.6973E8	23104	269185	7346.204	RADIATION::RADIATION

ornl

UNIVERSITY OF OREGON

NCAR

User Event Statistics

n,c,t, 3,0,0 - /Users/georgercarrjr/Work/NCAR/tauwork/ppk/cam8x1.ppk

File Options Windows Help

Sorted By: Number of Samples

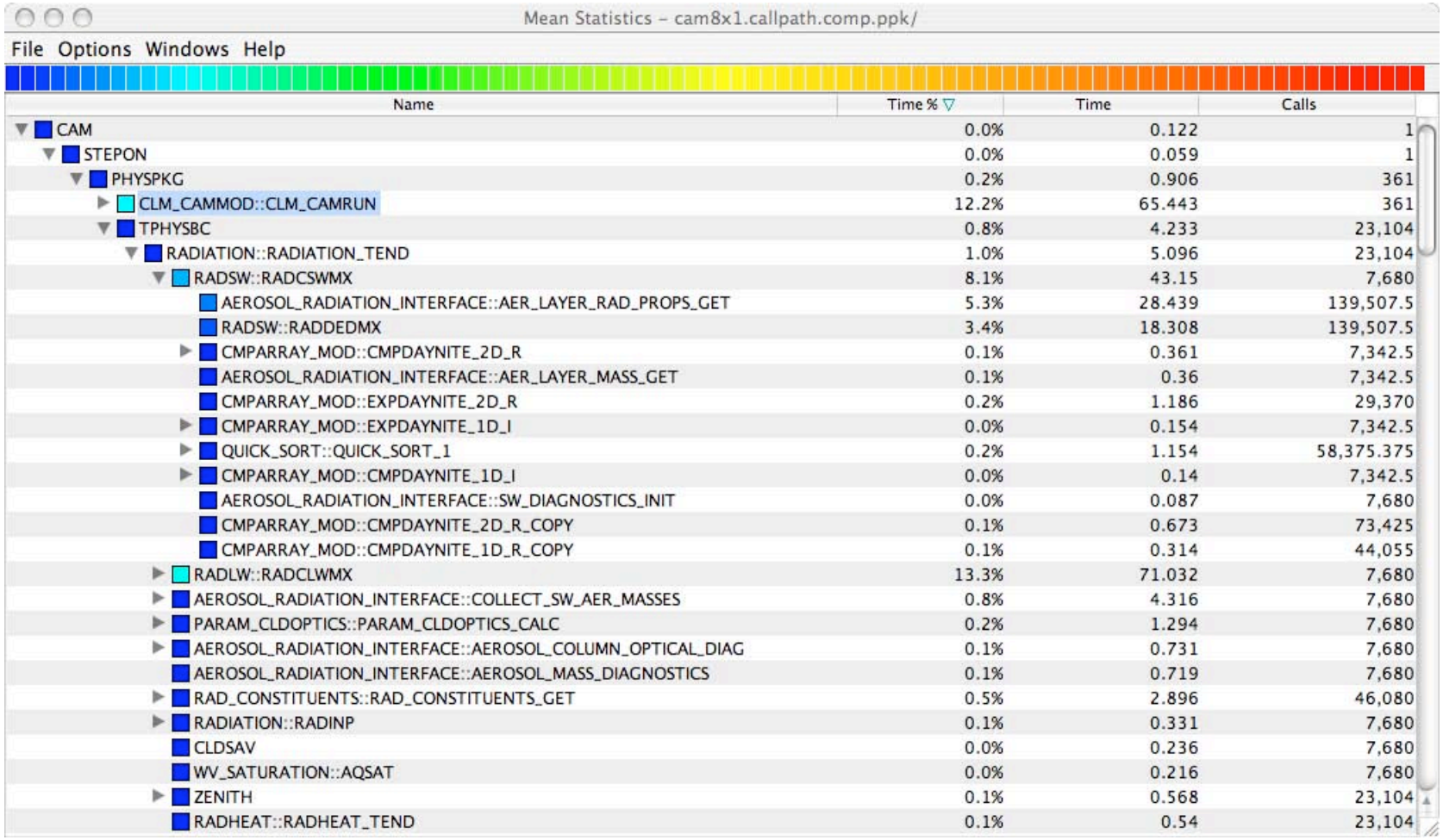
NumSamples	Max	Min	Mean	Std. Dev	Name
1449	135680	4	71412	63118	Message size for all-to-all
796	2940000	4	38891	297852	Message size for broadcast
723	5312	4	4090.3	1224.6	Message size for all-gather
556	61680	4	18417	10370	Message size for gather
57	483328	8192	148749	124331	Message size for scatter

ornl

UNIVERSITY OF OREGON

NCAR

Thread Statistics Table



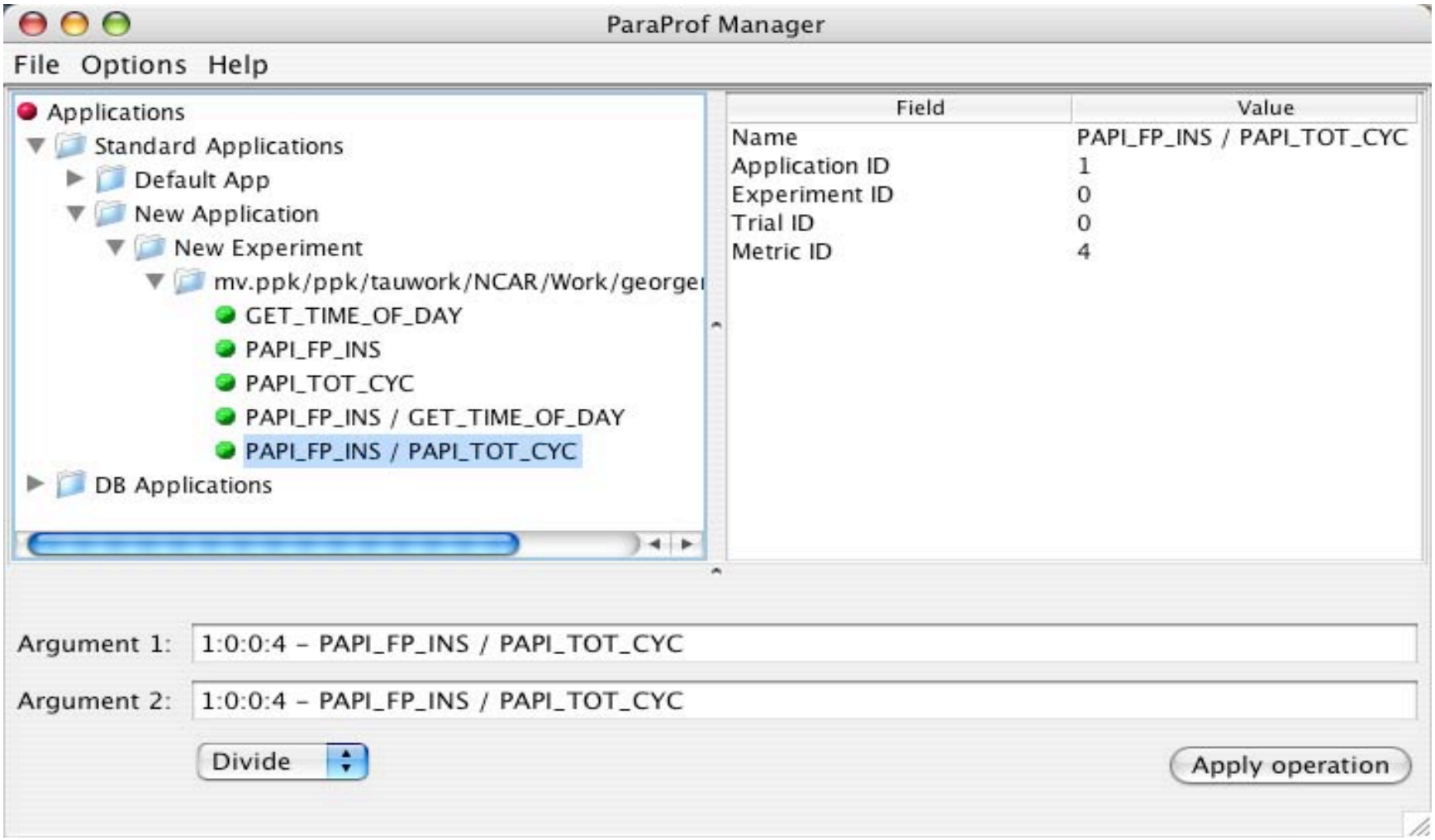
Name	Time % ▽	Time	Calls
CAM	0.0%	0.122	1
STEPON	0.0%	0.059	1
PHYSPKG	0.2%	0.906	361
CLM_CAMMOD::CLM_CAMRUN	12.2%	65.443	361
TPHYSBC	0.8%	4.233	23,104
RADIATION::RADIATION_TEND	1.0%	5.096	23,104
RADSW::RADCSWMX	8.1%	43.15	7,680
AEROSOL_RADIATION_INTERFACE::AER_LAYER_RAD_PROPS_GET	5.3%	28.439	139,507.5
RADSW::RADDEDMX	3.4%	18.308	139,507.5
CMPARRAY_MOD::CMPDAYNITE_2D_R	0.1%	0.361	7,342.5
AEROSOL_RADIATION_INTERFACE::AER_LAYER_MASS_GET	0.1%	0.36	7,342.5
CMPARRAY_MOD::EXPDAYNITE_2D_R	0.2%	1.186	29,370
CMPARRAY_MOD::EXPDAYNITE_1D_I	0.0%	0.154	7,342.5
QUICK_SORT::QUICK_SORT_1	0.2%	1.154	58,375.375
CMPARRAY_MOD::CMPDAYNITE_1D_I	0.0%	0.14	7,342.5
AEROSOL_RADIATION_INTERFACE::SW_DIAGNOSTICS_INIT	0.0%	0.087	7,680
CMPARRAY_MOD::CMPDAYNITE_2D_R_COPY	0.1%	0.673	73,425
CMPARRAY_MOD::CMPDAYNITE_1D_R_COPY	0.1%	0.314	44,055
RADLW::RADCLWMX	13.3%	71.032	7,680
AEROSOL_RADIATION_INTERFACE::COLLECT_SW_AER_MASSES	0.8%	4.316	7,680
PARAM_CLDOPTICS::PARAM_CLDOPTICS_CALC	0.2%	1.294	7,680
AEROSOL_RADIATION_INTERFACE::AEROSOL_COLUMN_OPTICAL_DIAG	0.1%	0.731	7,680
AEROSOL_RADIATION_INTERFACE::AEROSOL_MASS_DIAGNOSTICS	0.1%	0.719	7,680
RAD_CONSTITUENTS::RAD_CONSTITUENTS_GET	0.5%	2.896	46,080
RADIATION::RADINP	0.1%	0.331	7,680
CLDSAV	0.0%	0.236	7,680
WV_SATURATION::AQSAT	0.0%	0.216	7,680
ZENITH	0.1%	0.568	23,104
RADHEAT::RADHEAT_TEND	0.1%	0.54	23,104

ornl

UNIVERSITY OF OREGON

NCAR

Creating Derived Metrics

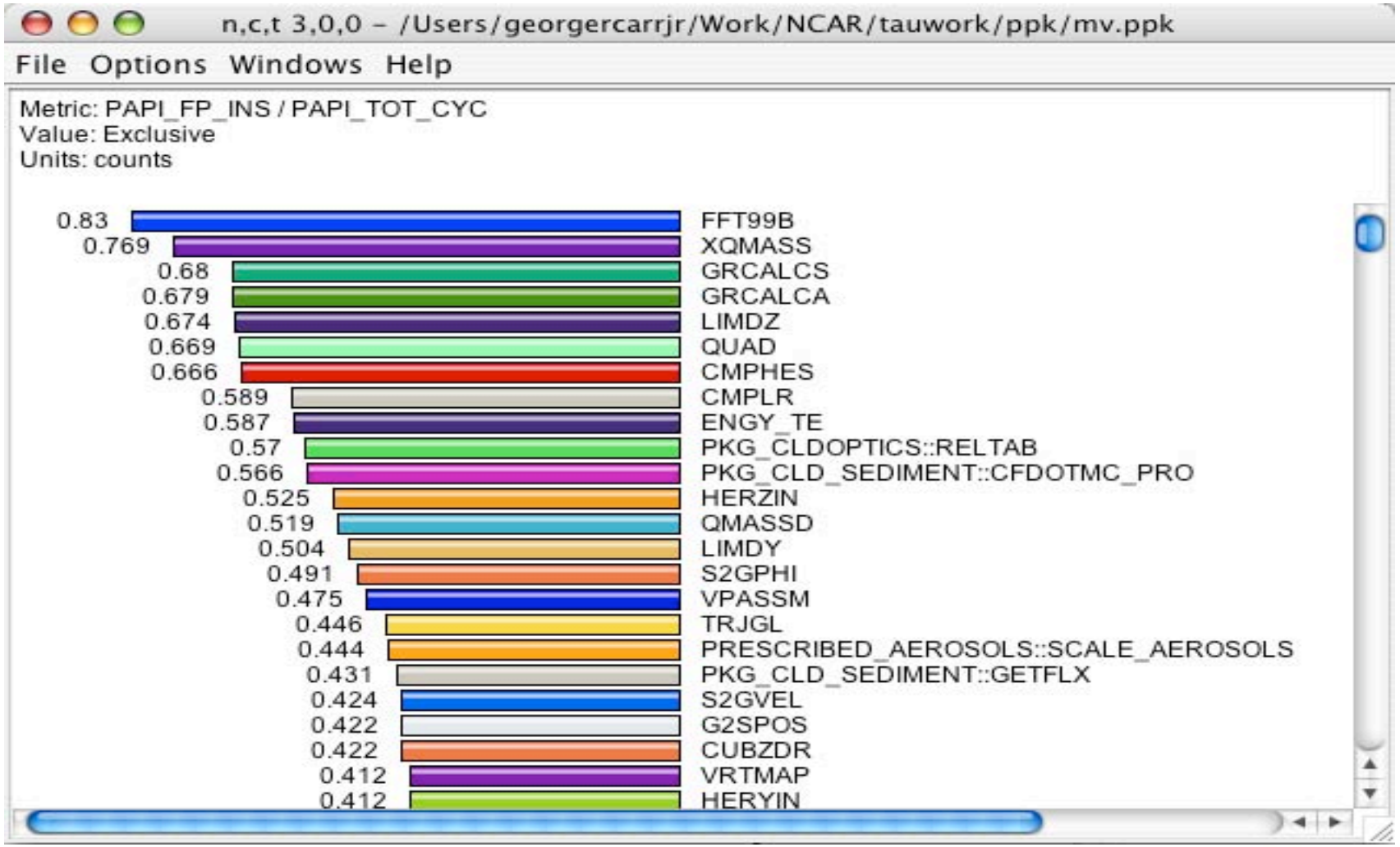


ornl

UNIVERSITY OF OREGON

NCAR

Display of Derived Metric

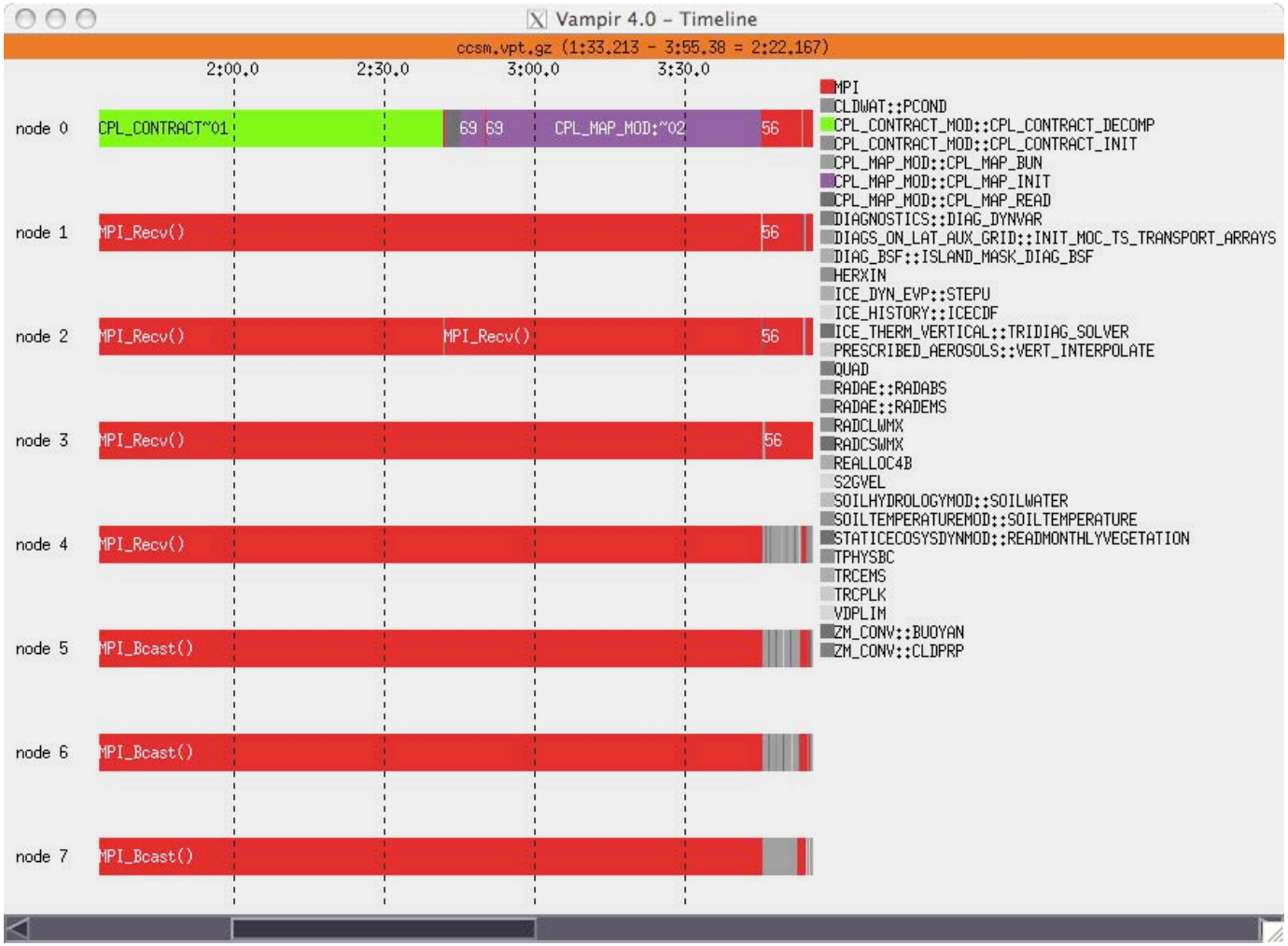


ornl

UNIVERSITY OF OREGON

NCAR

Load Imbalance: Timeline

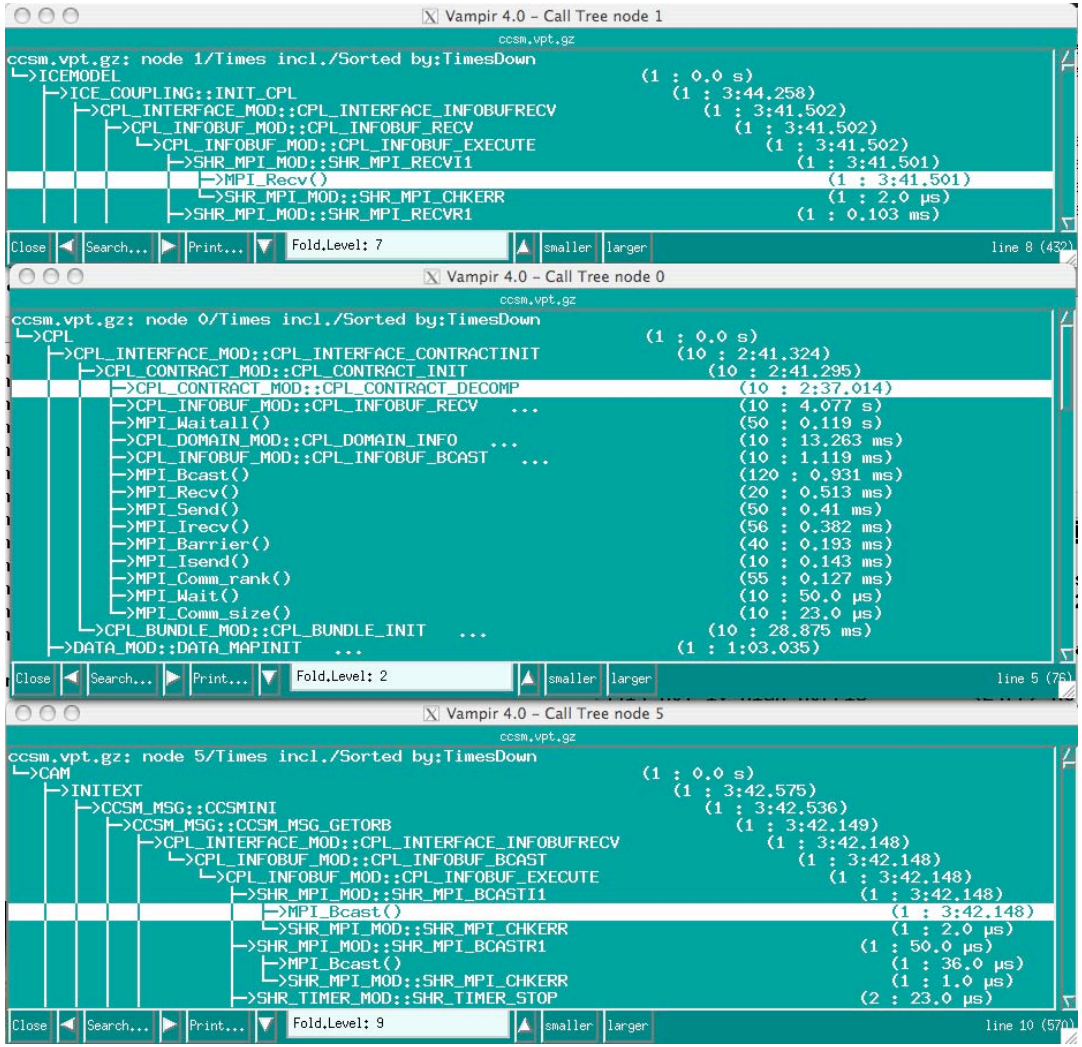


ornl

UNIVERSITY OF OREGON

NCAR

CCSM CallTree: TAU +Vampir

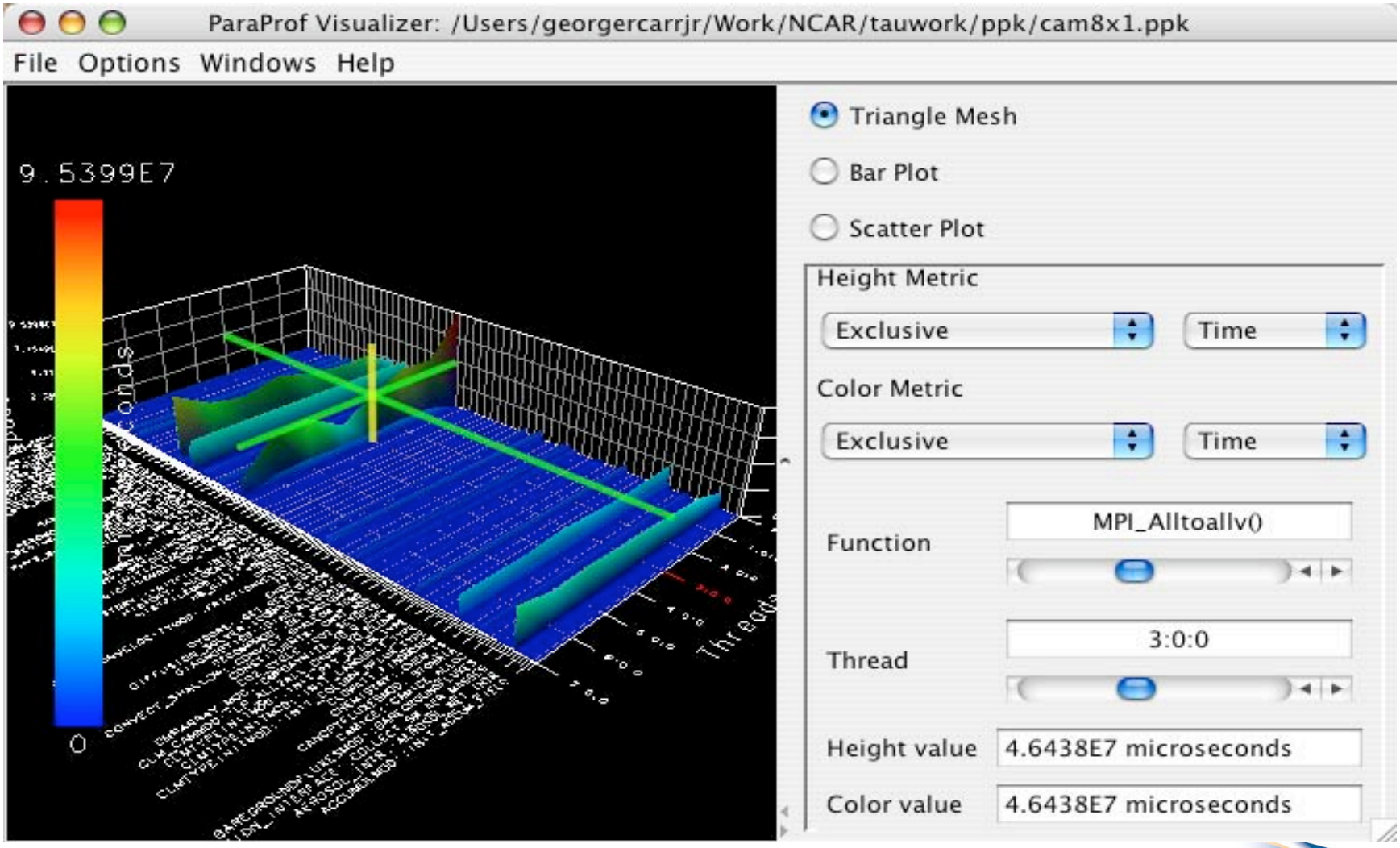


ornl

UNIVERSITY OF OREGON

NCAR

3D Visualization

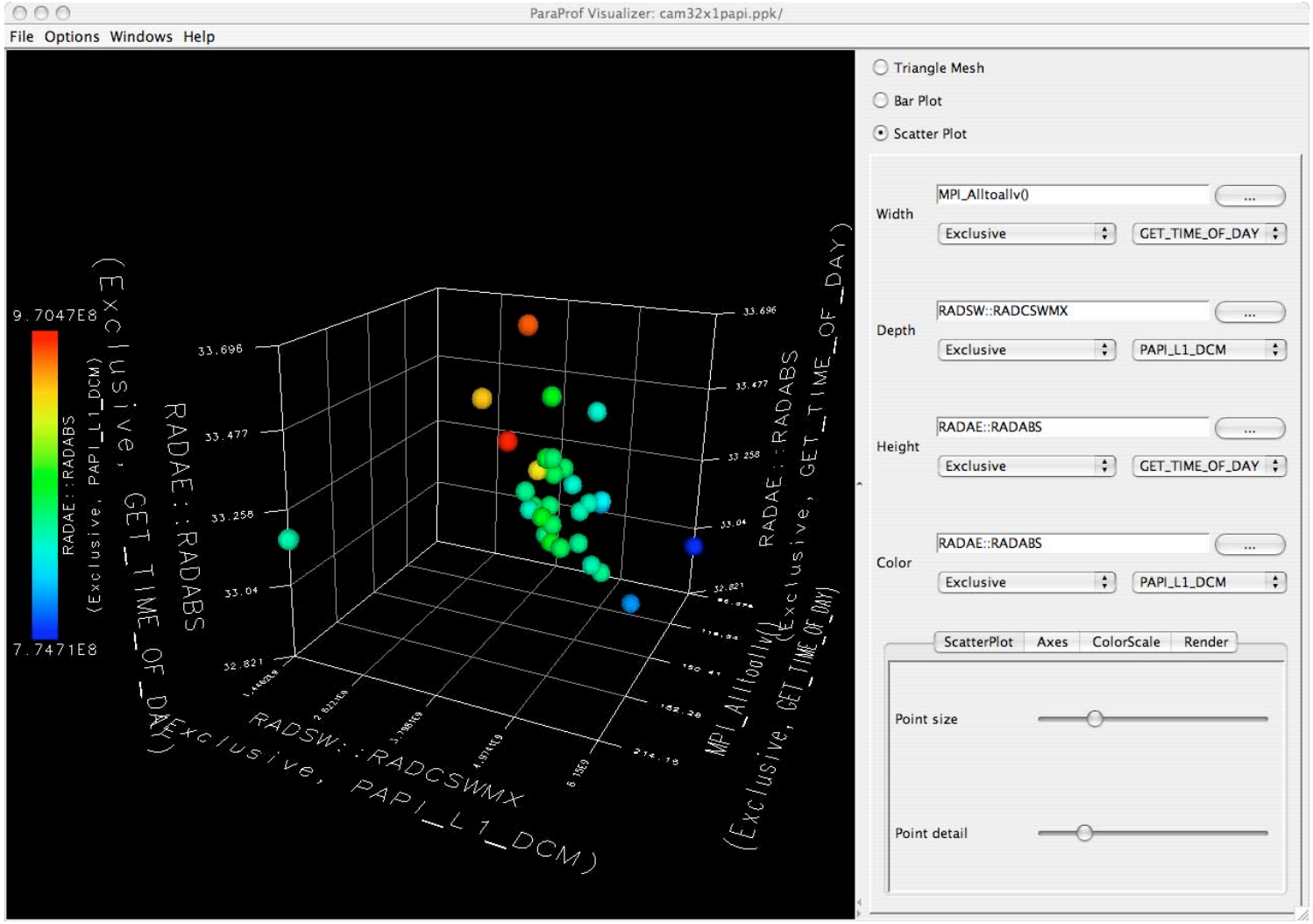


ornl

UNIVERSITY OF OREGON

NCAR

3D ScatterPlot



ornl

UNIVERSITY OF OREGON

NCAR

For More Information

See:

- NCAR's Bluesky
  - http://www.cisl.ucar.edu/computers/bluesky/
- The Community Climate System Model (CCSM)
  - http://www.ccsm.ucar.edu/models/ccsm3.0/
- ORNL's Phoenix
  - http://www.ccs.ornl.gov/Phoenix/Phoenix.html
- Tuning and Analysis Utilities (TAU)
  - http://www.cs.uoregon.edu/research/tau/home.php
- Program Database Toolkit (PDT)
  - http://www.cs.uoregon.edu/research/pdt/home.php
- Perfromance Data Standard and API
  - http://icl.cs.utk.edu/papi/
- Kit for Objective Judgement and Knowledge-based Detection of Performance Bottlenecks (KOJAK)
  - http://icl.cs.utk.edu/kojak/
- Vampir – commercial trace visualization tool
  - http://www.vampir-ng.de/
- OpenMP Pragma and Region Instrumentor (OPARI)
  - http://www.fz-juelich.de/zam/kojak/opari

ornl

UNIVERSITY OF OREGON

NCAR