

# TAU and E4S

**Sameer Shende**

**Director, Performance Research Laboratory, University of Oregon**

**President, ParaTools, Inc.**

[sameer@cs.uoregon.edu](mailto:sameer@cs.uoregon.edu)

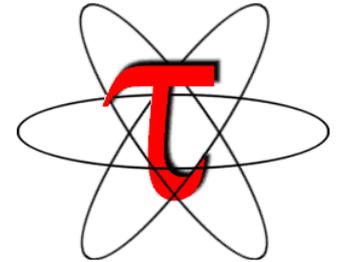
<http://tau.uoregon.edu>

**CERN, Monday, Jan 20, 2020**

**Slides: [http://tau.uoregon.edu/TAU\\_CERN20.pdf](http://tau.uoregon.edu/TAU_CERN20.pdf)**

# TAU Performance System<sup>®</sup>

<http://tau.uoregon.edu>

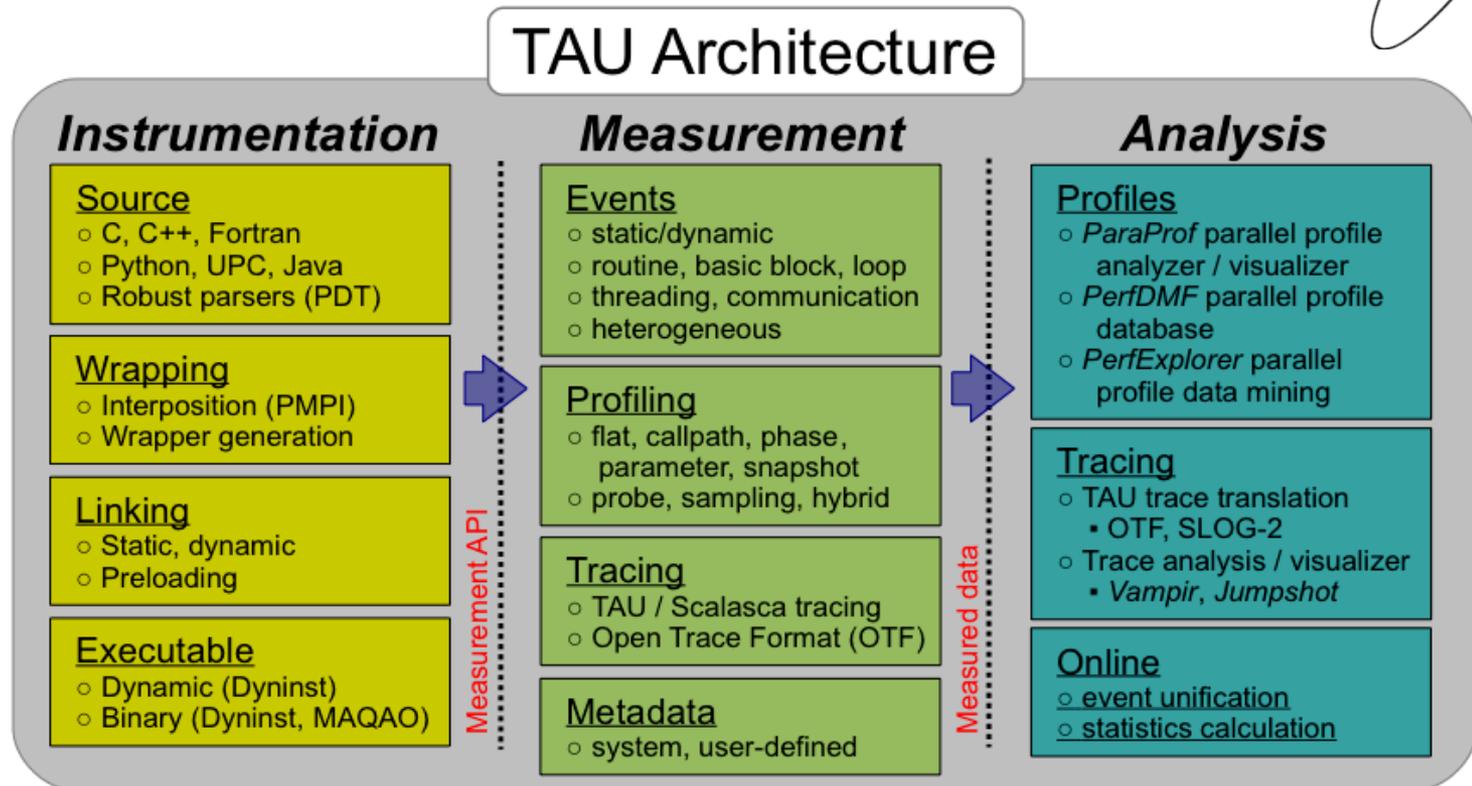
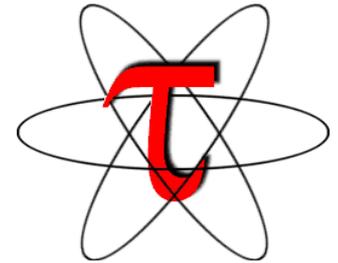


- **Tuning and Analysis Utilities (20+ year project)**
- **Comprehensive performance profiling and tracing**
  - Integrated, scalable, flexible, portable
  - Targets all parallel programming/execution paradigms
- **Integrated performance toolkit**
  - Instrumentation, measurement, analysis, visualization
  - Widely-ported performance profiling / tracing system
  - Performance data management and data mining
  - Open source (BSD-style license)
- **Integrates with application frameworks**

# Understanding Application Performance using TAU

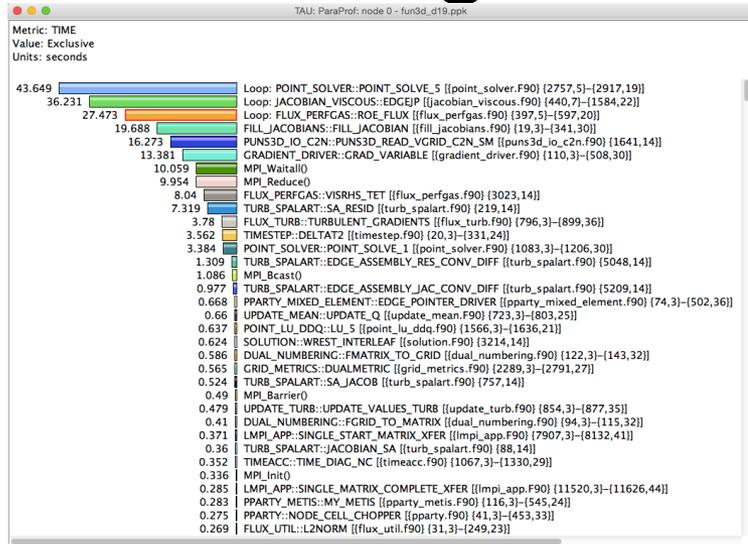
- **How much time** is spent in each application routine and outer *loops*? Within loops, what is the contribution of each *statement*?
- **How many instructions** are executed in these code regions? Floating point, Level 1 and 2 *data cache misses*, hits, branches taken?
- **What is the memory usage** of the code? When and where is memory allocated/de-allocated? Are there any memory leaks?
- **What are the I/O characteristics** of the code? What is the peak read and write *bandwidth* of individual calls, total volume?
- **What is the contribution of each phase** of the program? What is the time wasted/spent waiting for collectives, and I/O operations in Initialization, Computation, I/O phases?
- **How does the application scale?** What is the efficiency, runtime breakdown of performance across different core counts?

# TAU Architecture and Workflow

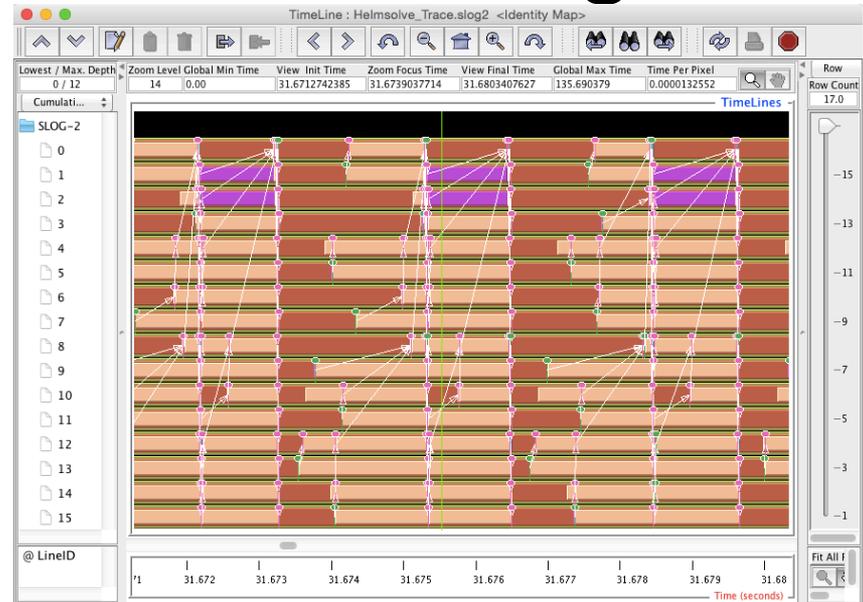


# Measurement: Profiling and Tracing

## Profiling



## Tracing



- Profiling and tracing

**Profiling** shows you **how much** (total) time was spent in each routine

**Tracing** shows you **when** the events take place on a timeline

# Instrumentation

## Source instrumentation using a preprocessor

- Add timer start/stop calls in a copy of the source code.
- Use Program Database Toolkit (PDT) for parsing source code.
- Requires recompiling the code using TAU shell scripts (tau\_cc.sh, tau\_f90.sh)
- Selective instrumentation (filter file) can reduce runtime overhead and narrow instrumentation focus.

## Compiler-based instrumentation

- Use system compiler to add a special flag to insert hooks at routine entry/exit.
- Requires recompiling using TAU scripts (tau\_cc.sh, tau\_f90.sh...)

## Python Interpreter Instrumentation

### Runtime preloading of TAU's Dynamic Shared Object (DSO)

- No need to recompile code! Use **mpirun tau\_exec ./app** with options.

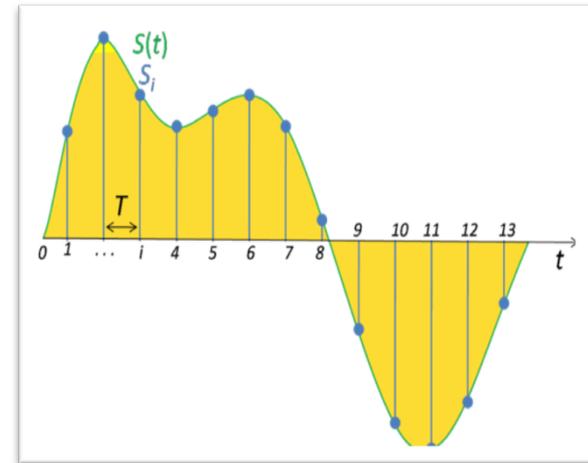
# Performance Data Measurement

## Direct via Probes

```
Call  
START ('potential')  
// code  
Call  
STOP ('potential')
```

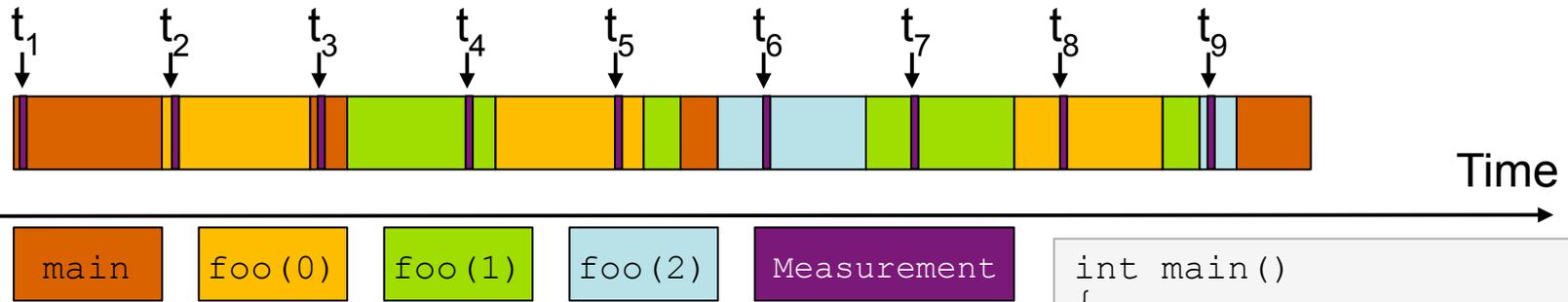
- Exact measurement
- Fine-grain control
- Calls inserted into code

## Indirect via Sampling



- No code modification
- Minimal effort
- Relies on debug symbols (**-g**)

# Sampling



## Running program is periodically interrupted to take measurement

- Timer interrupt, OS signal, or HWC overflow
- Service routine examines return-address stack
- Addresses are mapped to routines using symbol table information

## Statistical inference of program behavior

- Not very detailed information on highly volatile metrics
- Requires long-running applications

## Works with unmodified executables

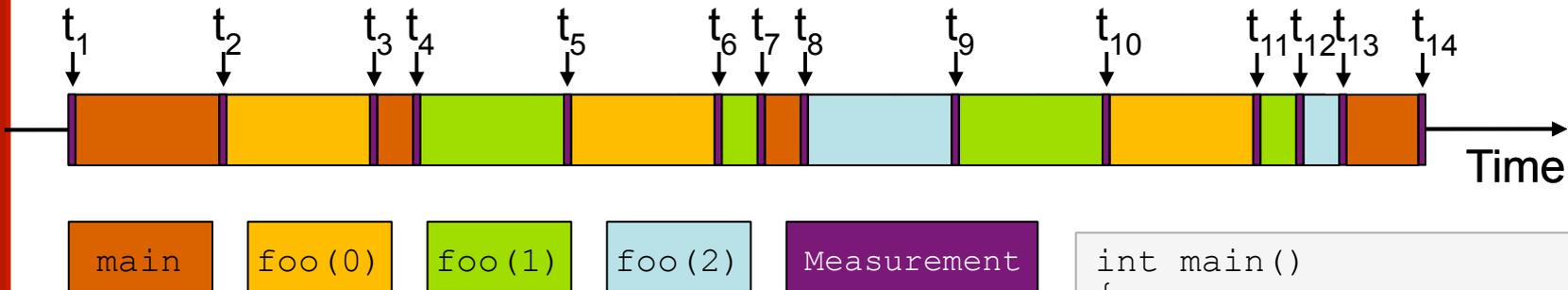
```
int main()
{
    int i;

    for (i=0; i < 3; i++)
        foo(i);

    return 0;
}

void foo(int i)
{
    if (i > 0)
        foo(i - 1);
}
```

# Instrumentation



**Measurement code is inserted such that every event of interest is captured directly**

- Can be done in various ways

**Advantage:**

- Much more detailed information

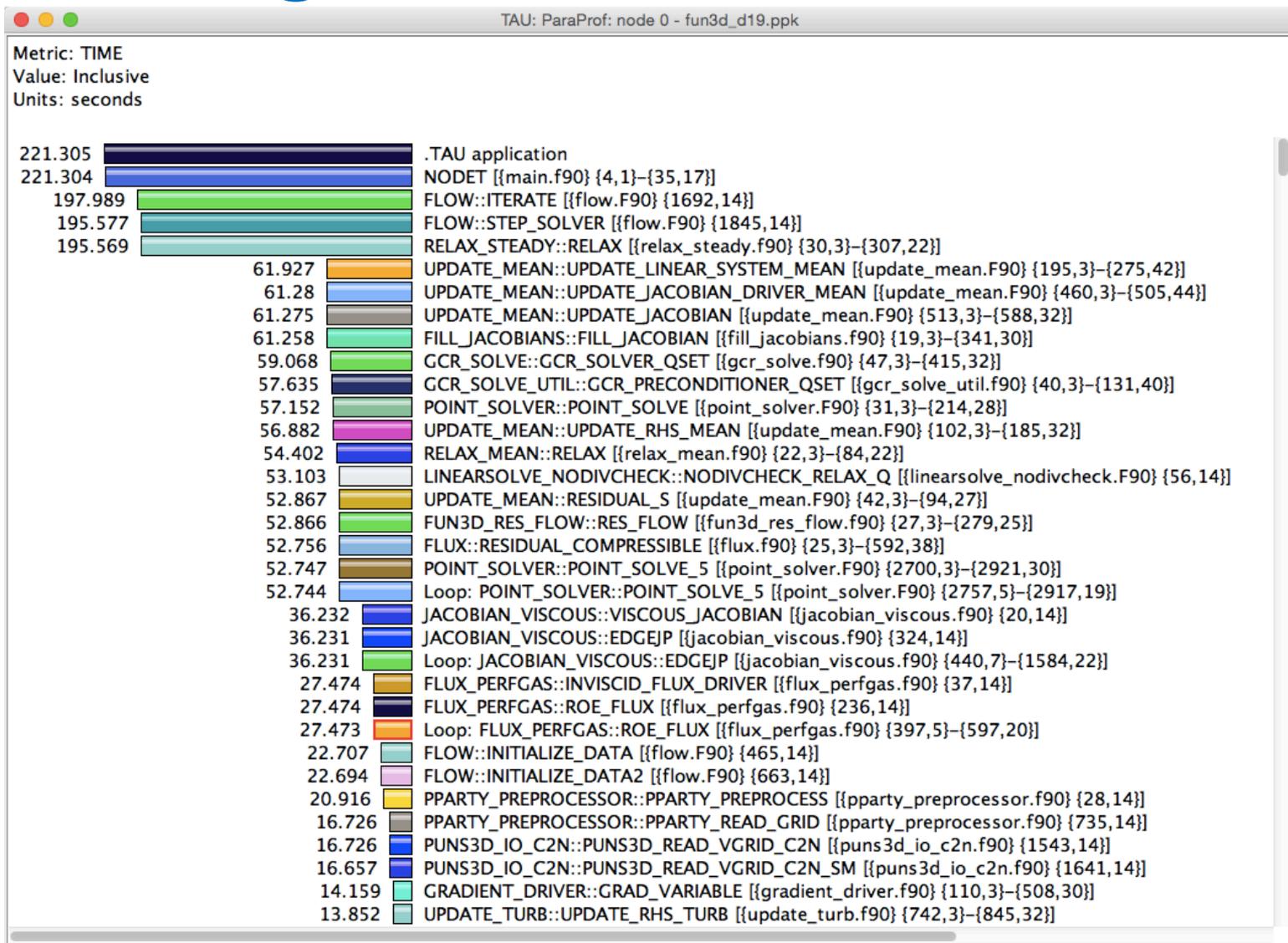
**Disadvantage:**

- Processing of source-code / executable necessary
- Large relative overheads for small functions

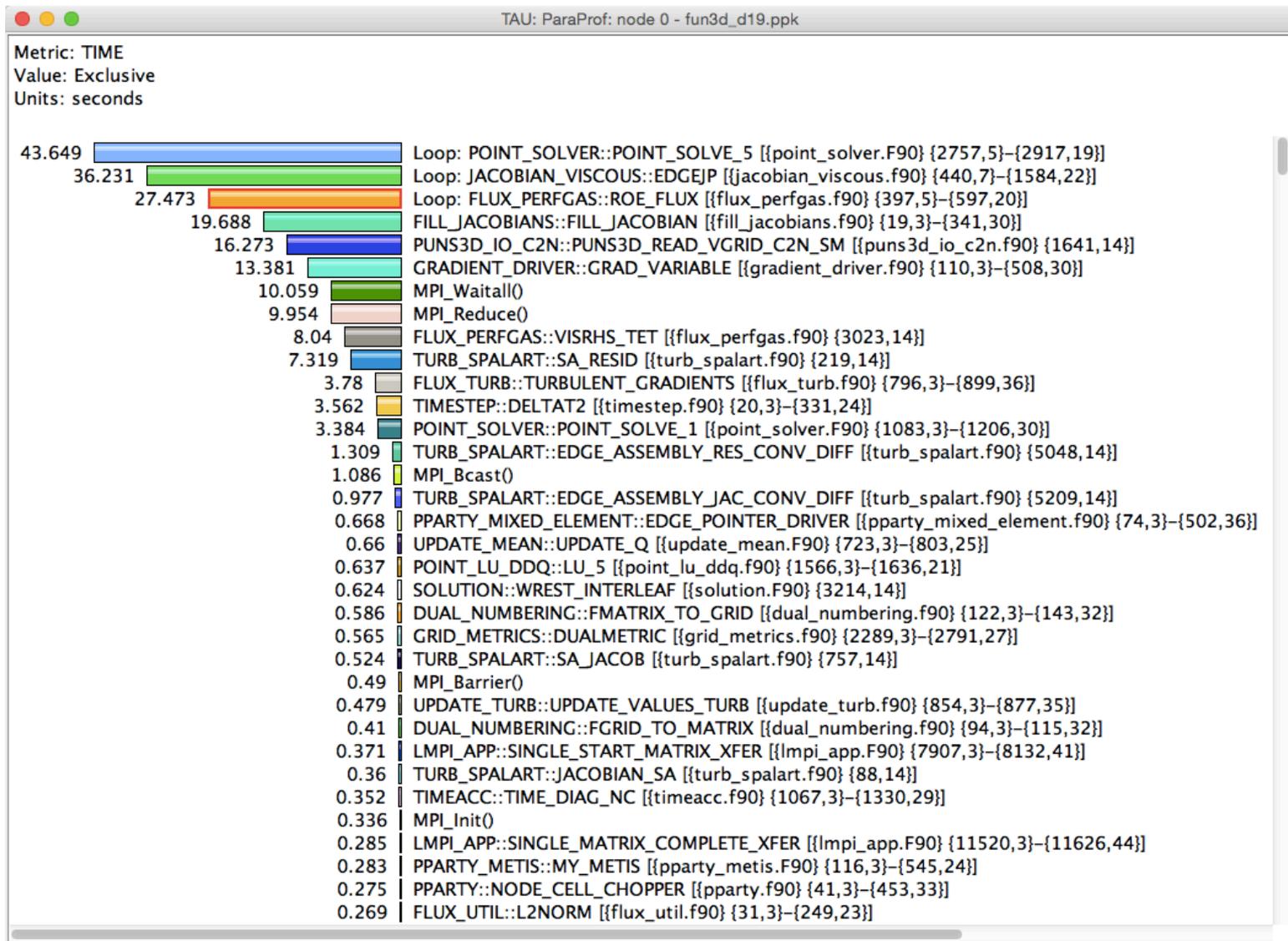
```
int main()
{
    int i;
    Start("main");
    for (i=0; i < 3; i++)
        foo(i);
    Stop("main");
    return 0;
}

void foo(int i)
{
    Start("foo");
    if (i > 0)
        foo(i - 1);
    Stop("foo");
}
```

# Profiling: Inclusive Measurements



# Profiling: Exclusive Time



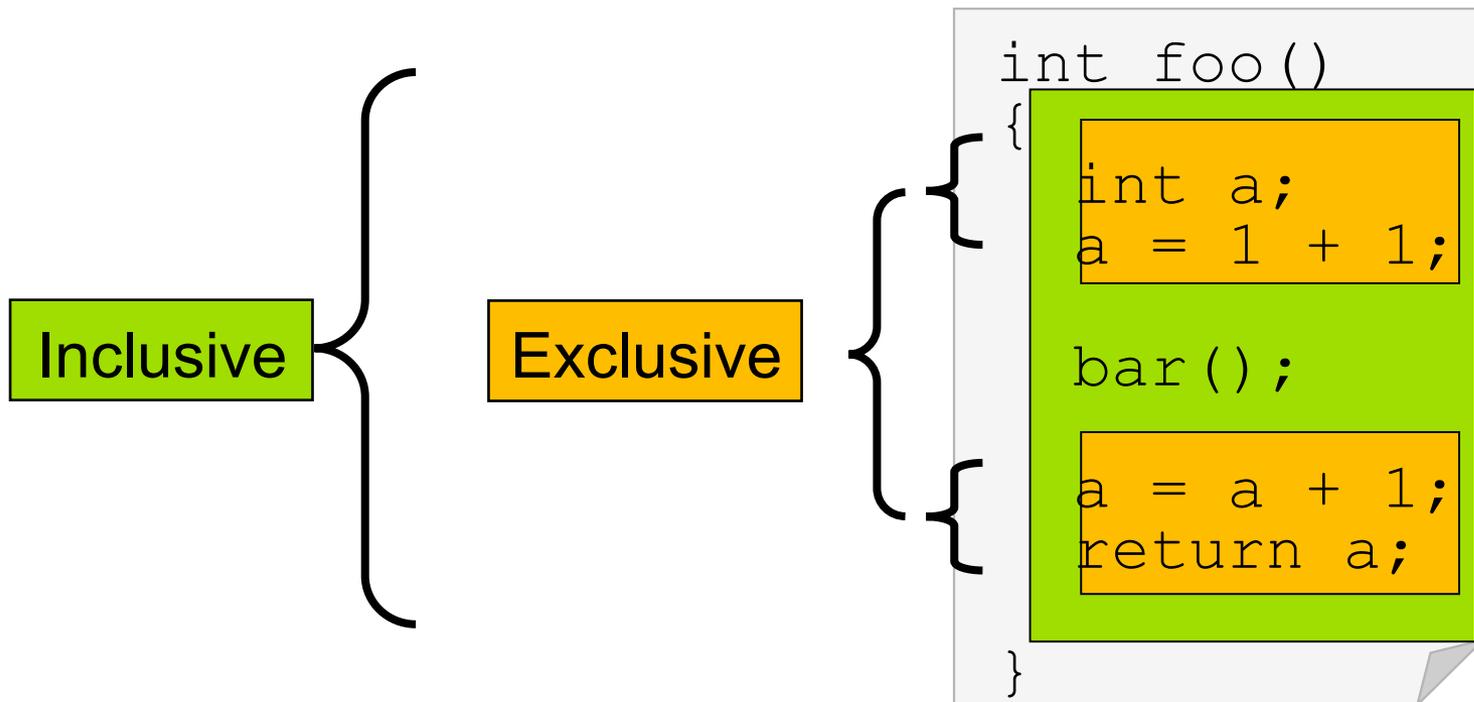
# Inclusive vs. Exclusive values

- **Inclusive**

- Information of all sub-elements aggregated into single value

- **Exclusive**

- Information cannot be subdivided further



# TAU: Quickstart Guide

## Setup:

- `% module load tau`

## Profiling with an un-instrumented application:

MPI: `% mpirun -np 64 tau_exec -ebs ./a.out`

- MPI+OpenMP: `% export TAU_OMPT_SUPPORT_LEVEL=full;`  
`% mpirun -np 64 tau_exec -T ompt,v5 -ompt ./a.out`
- Pthread: `% mpirun -np 64 tau_exec -T mpi,pthread -ebs ./a.out`
- Python+MPI+Sampling: `% mpirun -np 64 tau_python -ebs ./a.py`
- Python+MPI+CUDA+Sampling: `% mpirun -np 64 tau_python -cupti -ebs ./a.py`

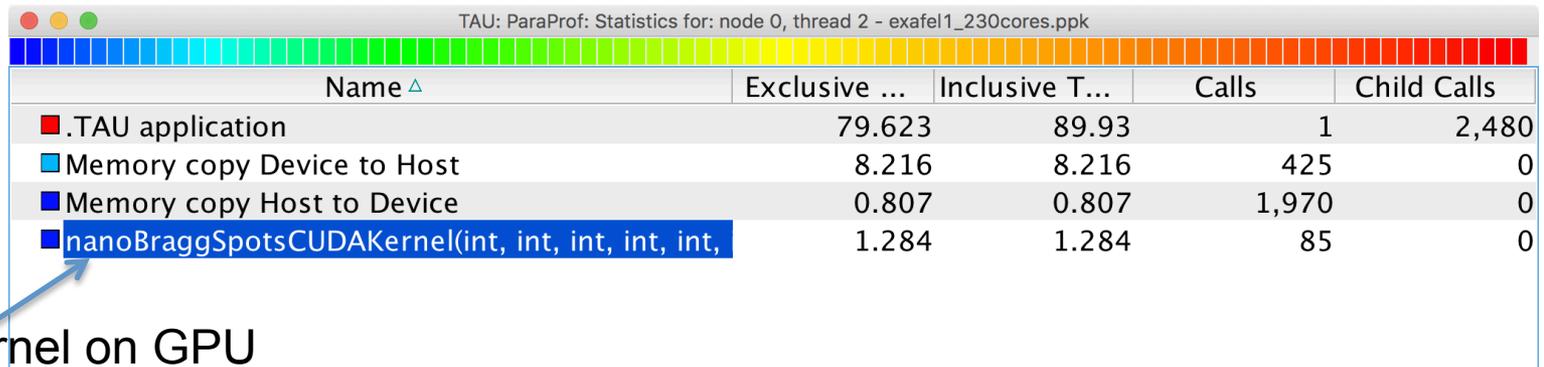
Analysis: `% pprof -a -m | more;` `% paraprof (GUI)`

## Tracing:

- Vampir: MPI: `% export TAU_TRACE=1; export TAU_TRACE_FORMAT=otf2`  
`% mpirun -np 64 tau_exec ./a.out; vampir traces.otf2 &`
  - Chrome: `% export TAU_TRACE=1; mpirun -np 64 tau_exec ./a.out`  
`% tau_treemerge.pl;`
- `% tau_trace2json tau.trc tau.edf -chrome -ignoreatomic -o app.json`  
Chrome browser: `chrome://tracing` (Load -> app.json)

# TAU supports Python, MPI, and CUDA

Without any modification to the source code or DSOs or interpreter, it instruments and samples the application using Python, MPI, and CUDA instrumentation.



TAU: ParaProf: Statistics for: node 0, thread 2 - exafel1\_230cores.ppk

Name $\Delta$	Exclusive ...	Inclusive T...	Calls	Child Calls
■ .TAU application	79.623	89.93	1	2,480
■ Memory copy Device to Host	8.216	8.216	425	0
■ Memory copy Host to Device	0.807	0.807	1,970	0
■ nanoBraggSpotsCUDAKernel(int, int, int, int, int, int)	1.284	1.284	85	0

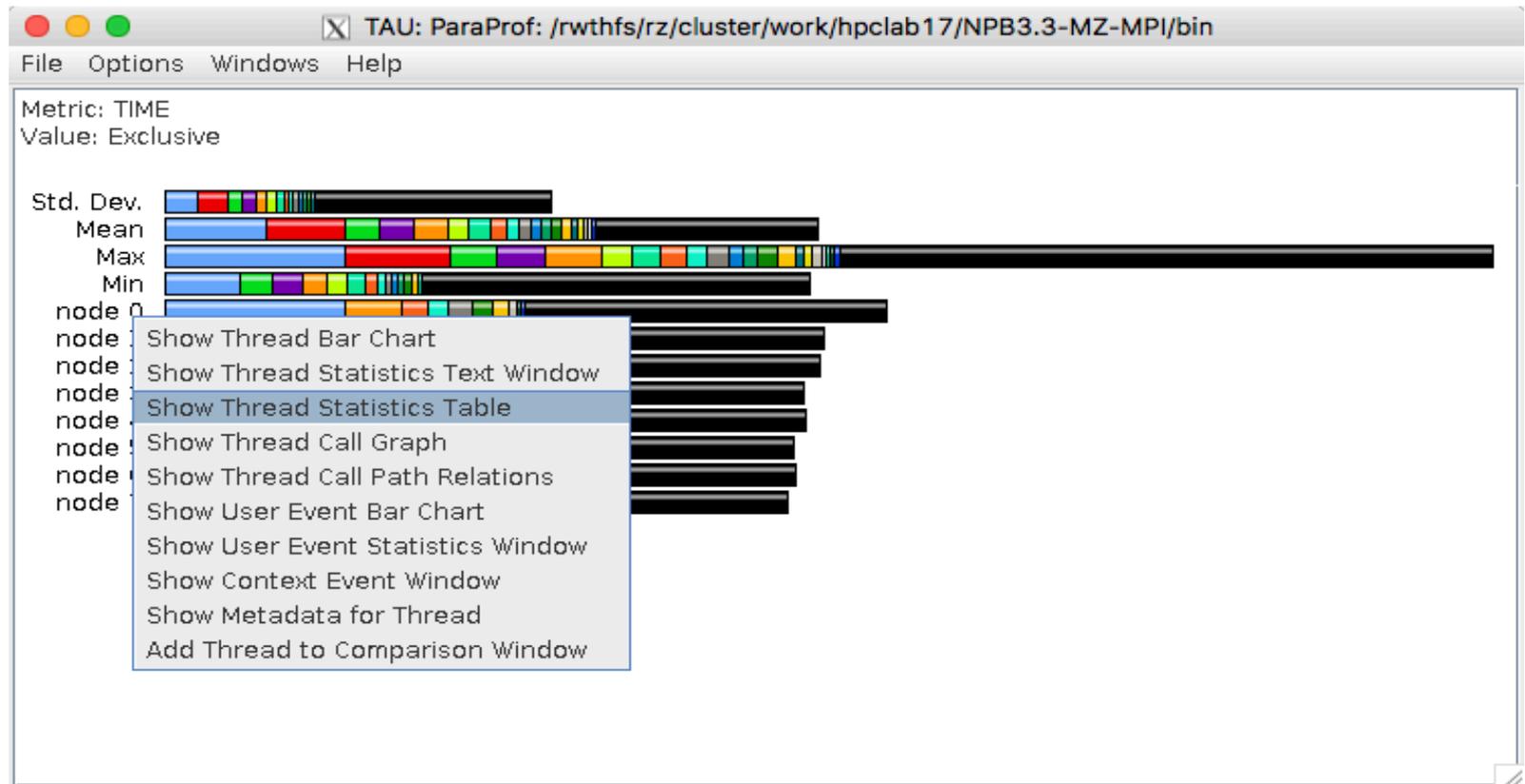
Kernel on GPU

```
% mpirun -np 230 tau_python -T cupti,mpi,pdt -ebs -cupti ./exafel.py
```

Instead of:

```
% mpirun -np 230 python ./exafel.py
```

# ParaProf Thread Statistics Table



Right click over “node X” and choose Show Thread Statistics Table

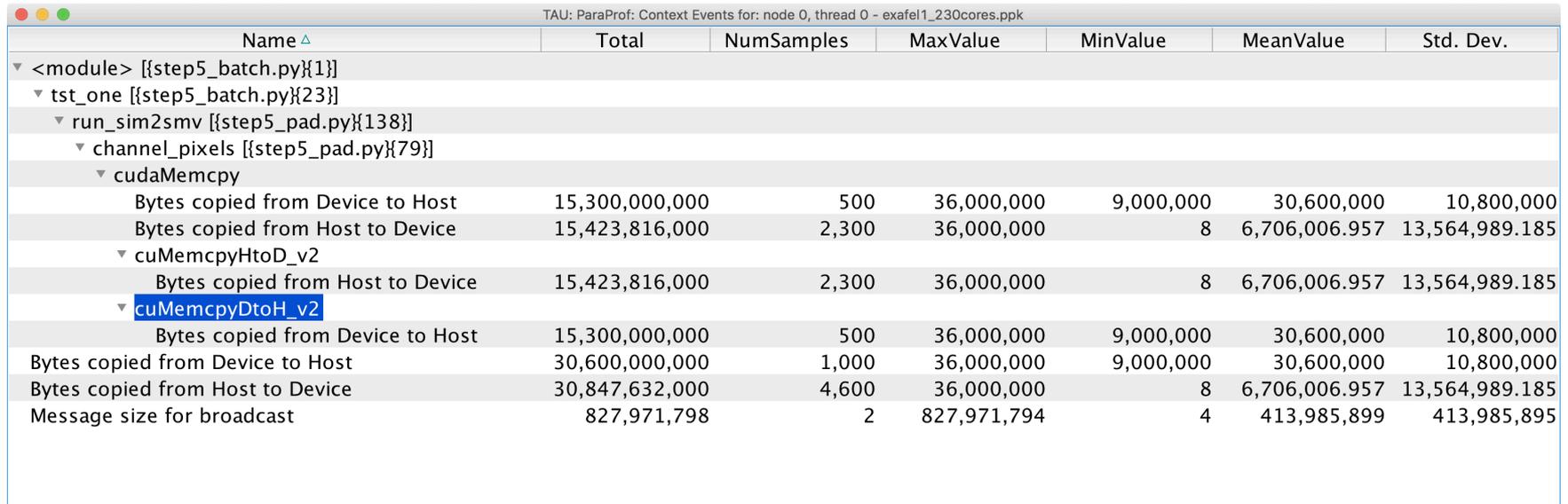
# TAU Thread Statistics Table

TAU: ParaProf: Statistics for: node 0, thread 0 - exafel1\_230cores.ppk

Name	Exclusive...	Inclusive ...	Calls	Child Calls
▸ <code>__init__</code> [from_scatterers_fft.py]{13}}	20.036	20.362	303	10,914
▸ <code>run_sim2smv</code> [step5_pad.py]{138}}	16.78	134.9	1	1,066
▸ <code>__init__</code> [__init__.py]{150}}	11.669	15.909	101	1,010
▾ <code>channel_pixels</code> [step5_pad.py]{79}}	11.029	107.657	100	13,358
▾ [CONTEXT] <code>channel_pixels</code> [step5_pad.py]{79}}	0	9.345	312	0
▾ [SAMPLE] <code>nanoBraggSpotsCUDA</code> [autofs/nccs-svm1_home1/iris/adse13_161/psana-legion/simtbx/sun	4.755	4.755	159	0
▾ [SAMPLE] <code>simtbx::nanoBragg::nanoBragg::add_nanoBragg_spots_cuda()</code> [autofs/nccs-svm1_home1/iris/	4.08	4.08	136	0
▾ [SAMPLE] <code>__memset_power8</code> [{}]{0}}	0.3	0.3	10	0
▾ [SAMPLE] <code>UNRESOLVED /usr/lib64/libc-2.17.so</code>	0.181	0.181	6	0
▸ [SUMMARY] <code>Tau_handle_driver_api_memcpy(void*, CUpti_CallbackDomain, unsigned int, CUpti_CallbackDz</code>	0.03	0.03	1	0
▸ <code>cuMemcpyDtoH_v2</code>	9.483	9.483	500	0
▸ <code>expand_to_p1_iselection</code> [__init__.py]{1376}}	7.349	7.35	101	606
▸ <code>load</code>	7.004	7.009	2	2,251
▸ <code>reset_wavelength</code> [util_fmodel.py]{121}}	6.197	6.553	100	47,550
▸ <code>is_unique_set_under_symmetry</code> [__init__.py]{790}}	5.913	5.915	202	808
▸ <code>__import__</code>	5.782	15.766	382	78
▸ <code>fp_fdp_at_wavelength</code> [fdp_plot.py]{44}}	5.616	5.723	800	1,600
▾ <code>MPI_Init_thread()</code>	4.987	4.987	1	0
▸ <code>cuDevicePrimaryCtxRetain</code>	4.735	4.735	2	0
▸ <code>&lt;module&gt;</code> [__init__.py]{1}}	4.255	23.888	85	756
▾ <code>MPI_Finalize()</code>	3.829	3.829	1	1
▸ <code>match_bijvoet_mates</code> [__init__.py]{1032}}	3.146	3.684	101	707
▸ <code>bcast</code>	3.073	3.448	1	9
▸ <code>__init__</code> [__init__.py]{20}}	3.011	3.399	101	149,196
▸ <code>compute_f_mask</code> [__init__.py]{299}}	2.897	18.853	101	707

Python, MPI, CUDA, and samples from DSOs are all integrated in a single view

# TAU Context Event Window



TAU: ParaProf: Context Events for: node 0, thread 0 - exafel1\_230cores.ppk

Name ^	Total	NumSamples	MaxValue	MinValue	MeanValue	Std. Dev.
<module> [{{step5_batch.py}}{1}]						
▼ tst_one [{{step5_batch.py}}{23}]						
▼ run_sim2smv [{{step5_pad.py}}{138}]						
▼ channel_pixels [{{step5_pad.py}}{79}]						
▼ cudaMemcpy						
Bytes copied from Device to Host	15,300,000,000	500	36,000,000	9,000,000	30,600,000	10,800,000
Bytes copied from Host to Device	15,423,816,000	2,300	36,000,000	8	6,706,006.957	13,564,989.185
▼ cuMemcpyHtoD_v2						
Bytes copied from Host to Device	15,423,816,000	2,300	36,000,000	8	6,706,006.957	13,564,989.185
▼ <b>cuMemcpyDtoH_v2</b>						
Bytes copied from Device to Host	15,300,000,000	500	36,000,000	9,000,000	30,600,000	10,800,000
Bytes copied from Device to Host	30,600,000,000	1,000	36,000,000	9,000,000	30,600,000	10,800,000
Bytes copied from Host to Device	30,847,632,000	4,600	36,000,000	8	6,706,006.957	13,564,989.185
Message size for broadcast	827,971,798	2	827,971,794	4	413,985,899	413,985,895

TAU tracks the data transfers between the host and the GPU.

# TAU's tracking of Python and MPI

TAU: ParaProf: Statistics for: node 1, thread 0 - exafel1\_230cores.ppk

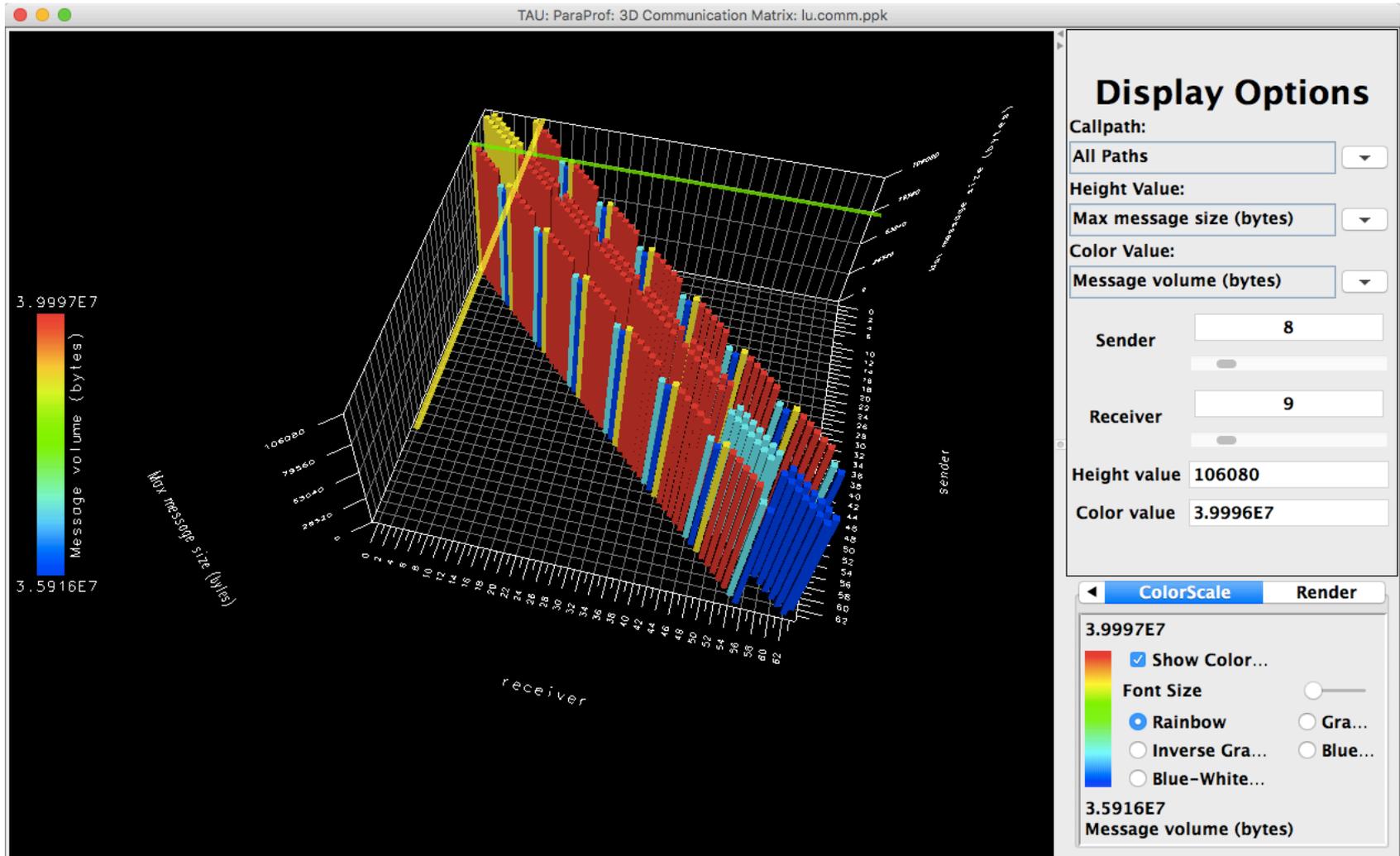
Name	Exclusive...	Inclusive ...	Calls	Child ...
▶ <code>__init__</code> [from_scatterers_fft.py]{13}	19.845	20.166	303	10,914
▶ <code>run_sim2smv</code> [step5_pad.py]{138}	16.672	133.715	1	1,066
▼ <code>MPI_Bcast()</code>	12.263	12.263	2	0
▼ [CONTEXT] <code>MPI_Bcast()</code>	0	12.21	407	0
■ [SAMPLE] <code>PAMI_Context_lock</code> [autofs/nccs-svm1_sw/summit/.swci/1-compute/opt/spac	3.27	3.27	109	0
■ [SAMPLE] <code>pthread_spin_lock</code> [usr/lib64/libpthread-2.17.so]{0}	2.34	2.34	78	0
■ [SAMPLE] <code>start_libcoll_blocking_collective</code> [autofs/nccs-svm1_sw/summit/.swci/1-compi	1.89	1.89	63	0
■ [SAMPLE] <code>PAMI::Device::IBV::Device::advance()</code> [autofs/nccs-svm1_sw/summit/.swci/1-cc	1.56	1.56	52	0
■ [SAMPLE] <code>PAMI_Context_advancev</code> [autofs/nccs-svm1_sw/summit/.swci/1-compute/opt	0.69	0.69	23	0
■ [SAMPLE] <code>UNRESOLVED /usr/lib64/libmlx5.so.1.0.0</code>	0.51	0.51	17	0
▼ [SUMMARY] <code>LIBCOLL_Advance_pami</code> [/_SMPI_build_dir_____/ibmsrc/r	0.42	0.42	14	0
■ [SAMPLE] <code>LIBCOLL_Advance_pami</code> [/_SMPI_build_dir_____/ibmsrc/n	0.42	0.42	14	0
■ [SAMPLE] <code>PAMI_Context_unlock</code> [autofs/nccs-svm1_sw/summit/.swci/1-compute/opt/sj	0.39	0.39	13	0
■ [SAMPLE] <code>pthread_spin_unlock</code> [usr/lib64/libpthread-2.17.so]{0}	0.36	0.36	12	0
■ [SAMPLE] <code>__memcpy_power7</code> [{}]{0}	0.33	0.33	11	0
■ [SAMPLE] <code>0000003d.plt_call.PAMI_Context_lock</code> [{}]{0}	0.15	0.15	5	0
■ [SAMPLE] <code>verbs_get_exp_ctx</code> [pami.cc]{0}	0.09	0.09	3	0
■ [SAMPLE] <code>PAMI_Context_trylock_advancev</code> [autofs/nccs-svm1_sw/summit/.swci/1-comp	0.06	0.06	2	0
■ [SAMPLE] <code>0000003d.plt_call.PAMI_Context_unlock</code> [{}]{0}	0.06	0.06	2	0
■ [SAMPLE] <code>opal_progress</code> [autofs/nccs-svm1_sw/summit/.swci/1-compute/opt/spack/2C	0.03	0.03	1	0
■ [SAMPLE] <code>00000052.plt_call.PAMI_Context_advancev</code> [{}]{0}	0.03	0.03	1	0
▼ [SUMMARY] <code>CCMI::Executor::ShmemBroadcastT&lt;false, CCMI::Executor::ShmemAtomicBarrie</code>	0.03	0.03	1	0
■ [SAMPLE] <code>CCMI::Executor::ShmemBroadcastT&lt;false, CCMI::Executor::ShmemAtomicBarrie</code>	0.03	0.03	1	0
▶ <code>__init__</code> [__init__.py]{150}	11.518	15.698	101	1,010
▶ <code>channel_pixels</code> [step5_pad.py]{79}	10.949	106.61	100	13,358
▶ <code>cuMemcpyDtoH_v2</code>	9.433	9.433	500	0

TAU can observe events in closed-source vendor libraries (e.g., `MPI_Bcast`)!



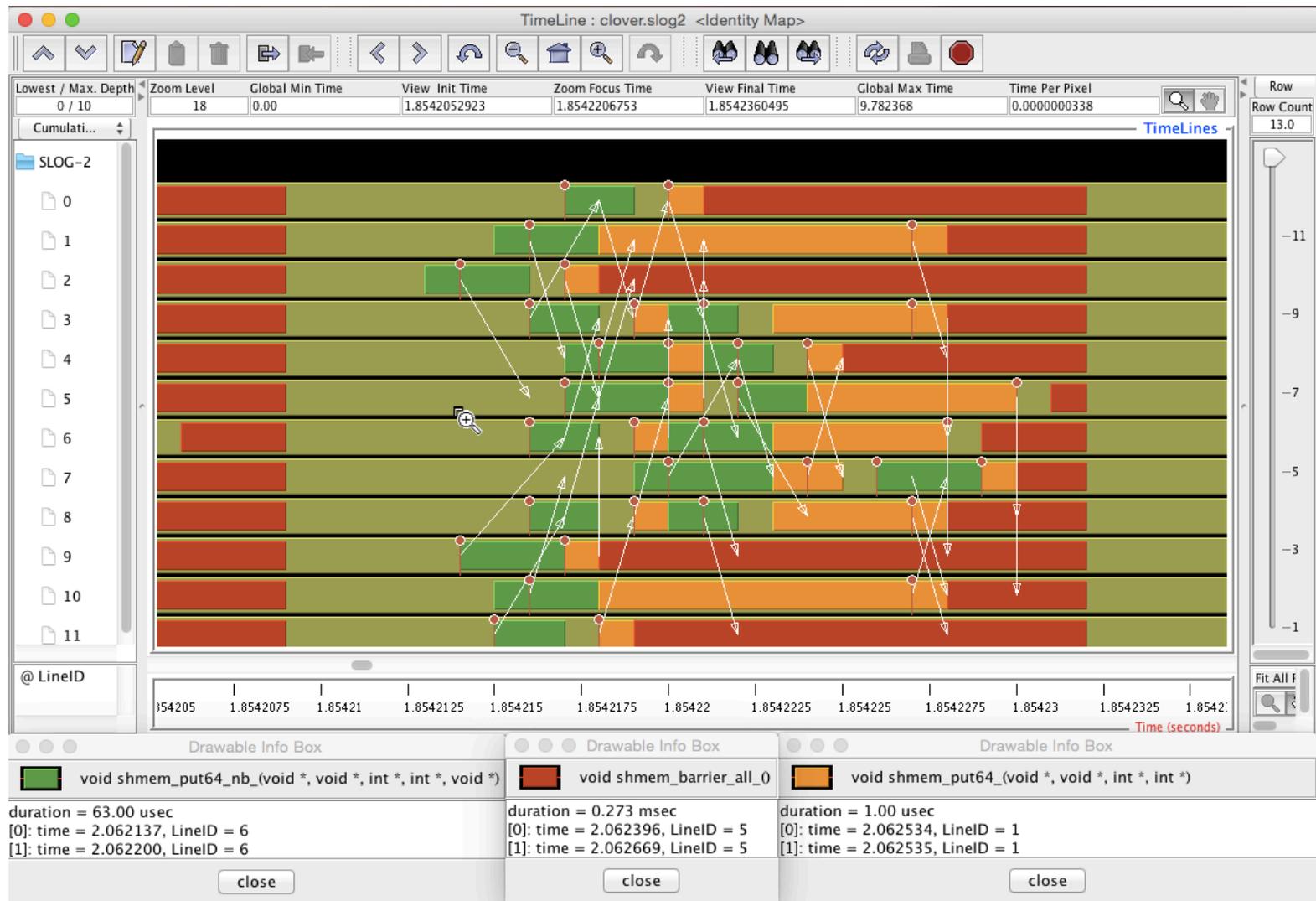


# TAU – 3D Communication Window

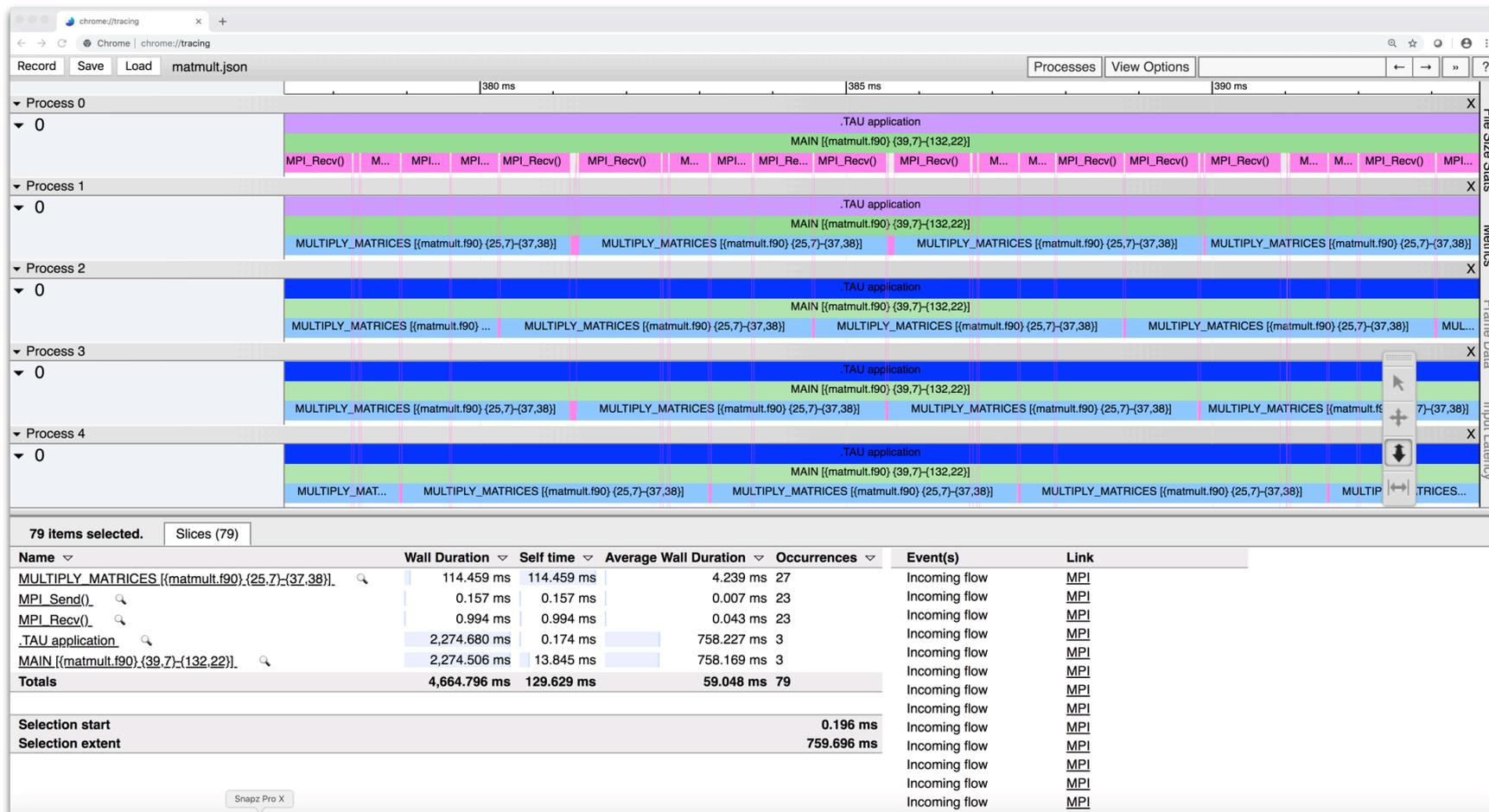


```
% export TAU_COMM_MATRIX=1; mpirun ... tau_exec ./a.out
% paraprof app.ppk; Windows -> 3D Communication Matrix
```

# Tracing: Jumpshot (ships with TAU)



# Tracing: Chrome Browser



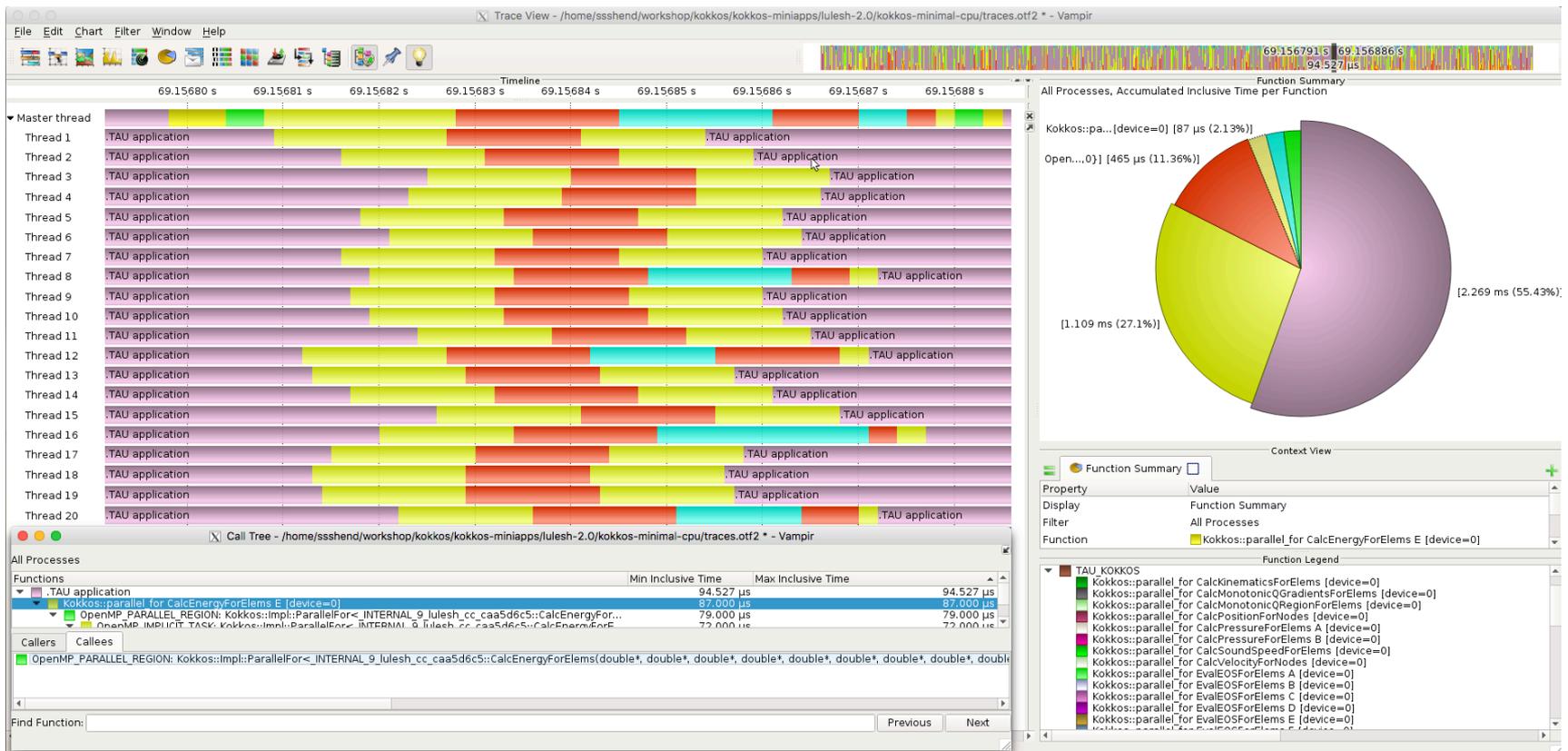
% export TAU\_TRACE=1

% mpirun -np 256 tau\_exec ./a.out

% tau\_treemerge.pl; tau\_trace2json tau.trc tau.edf -chrome -ignoreatomic -o app.json

Chrome browser: chrome://tracing (Load -> app.json)

# Vampir [TU Dresden] Timeline: Kokkos



```
% export TAU_TRACE=1; export TAU_TRACE_FORMAT=otf2
% tau_exec -ompt ./a.out
% vampir traces.otf2 &
```

# TAU's Support for Runtime Systems

## *MPI*

- PMPI profiling interface
- MPI\_T tools interface using performance and control variables

## *Pthread*

- Captures time spent in routines per thread of execution

## *OpenMP*

- OMPT tools interface to track salient OpenMP runtime events
- Opari source rewriter
- Preloading wrapper OpenMP runtime library when OMPT is not supported

## *OpenACC*

- OpenACC instrumentation API
- Track data transfers between host and device (per-variable)
- Track time spent in kernels

# TAU's Support for Runtime Systems (contd.)

## *OpenCL*

- OpenCL profiling interface
- Track timings of kernels

## *CUDA*

- Cuda Profiling Tools Interface (CUPTI)
- Track data transfers between host and GPU
- Track access to uniform shared memory between host and GPU

## *ROCm*

- Rocprofiler and Roctracer instrumentation interfaces
- Track data transfers and kernel execution between host and GPU

## *Kokkos*

- Kokkos profiling API
- Push/pop interface for region, kernel execution interface

## *Python*

- Python interpreter instrumentation API
- Tracks Python routine transitions as well as Python to C transitions

# Examples of Multi-Level Instrumentation

## ***MPI + OpenMP***

- MPI\_T + PMPI + OMPT may be used to track MPI and OpenMP

## ***MPI + CUDA***

- PMPI + CUPTI interfaces

## ***OpenCL + ROCm***

- Rocprofiler + OpenCL instrumentation interfaces

## ***Kokkos + OpenMP***

- Kokkos profiling API + OMPT to transparently track events

## ***Kokkos + pthread + MPI***

- Kokkos + pthread wrapper interposition library + PMPI layer

## ***Python + CUDA + MPI***

- Python + CUPTI + pthread profiling interfaces (e.g., Tensorflow, PyTorch) + MPI

## ***MPI + OpenCL***

- PMPI + OpenCL profiling interfaces

# TAU Execution Command (tau\_exec)

## Uninstrumented execution

- `% mpirun -np 256 ./a.out`

## Track GPU operations

- `% mpirun -np 256 tau_exec -rocm ./a.out`
- `% mpirun -np 256 tau_exec -cupti ./a.out`
- `% mpirun -np 256 tau_exec -cupti -um ./a.out` (for Unified Memory)
- `% mpirun -np 256 tau_exec -opencl ./a.out`
- `% mpirun -np 256 tau_exec -openacc ./a.out`

## Track MPI performance

- `% mpirun -np 256 tau_exec ./a.out`

## Track I/O, and MPI performance (MPI enabled by default)

- `% mpirun -np 256 tau_exec -io ./a.out`

## Track OpenMP and MPI execution (using OMPT for Intel v19+ or Clang 8+)

- `% export TAU_OMPT_SUPPORT_LEVEL=full;`
- `% mpirun -np 256 tau_exec -T ompt,v5,mpi -ompt ./a.out`

## Track memory operations

- `% export TAU_TRACK_MEMORY_LEAKS=1`
- `% mpirun -np 256 tau_exec -memory_debug ./a.out` (bounds check)

## Use event based sampling (compile with -g)

- `% mpirun -np 256 tau_exec -ebs ./a.out`
- Also `-ebs_source=<PAPI_COUNTER>` `-ebs_period=<overflow_count>`  
`-ebs_resolution=<file | function | line>`

# Configuration tags for tau\_exec

```
% ./configure -pdt=<dir> -mpi -papi=<dir>; make install
```

Creates in \$TAU:

```
Makefile.tau-papi-mpi-pdt (Configuration parameters in stub makefile)  
shared-papi-mpi-pdt/libTAU.so
```

```
% ./configure -pdt=<dir> -mpi; make install creates
```

```
Makefile.tau-mpi-pdt  
shared-mpi-pdt/libTAU.so
```

To explicitly choose preloading of shared-<options>/libTAU.so change:

```
% mpirun -np 256 ./a.out to
```

```
% mpirun -np 256 tau_exec -T <comma_separated_options> ./a.out
```

```
% mpirun -np 256 tau_exec -T papi,mpi,pdt ./a.out
```

Preloads \$TAU/shared-papi-mpi-pdt/libTAU.so

```
% mpirun -np 256 tau_exec -T papi ./a.out
```

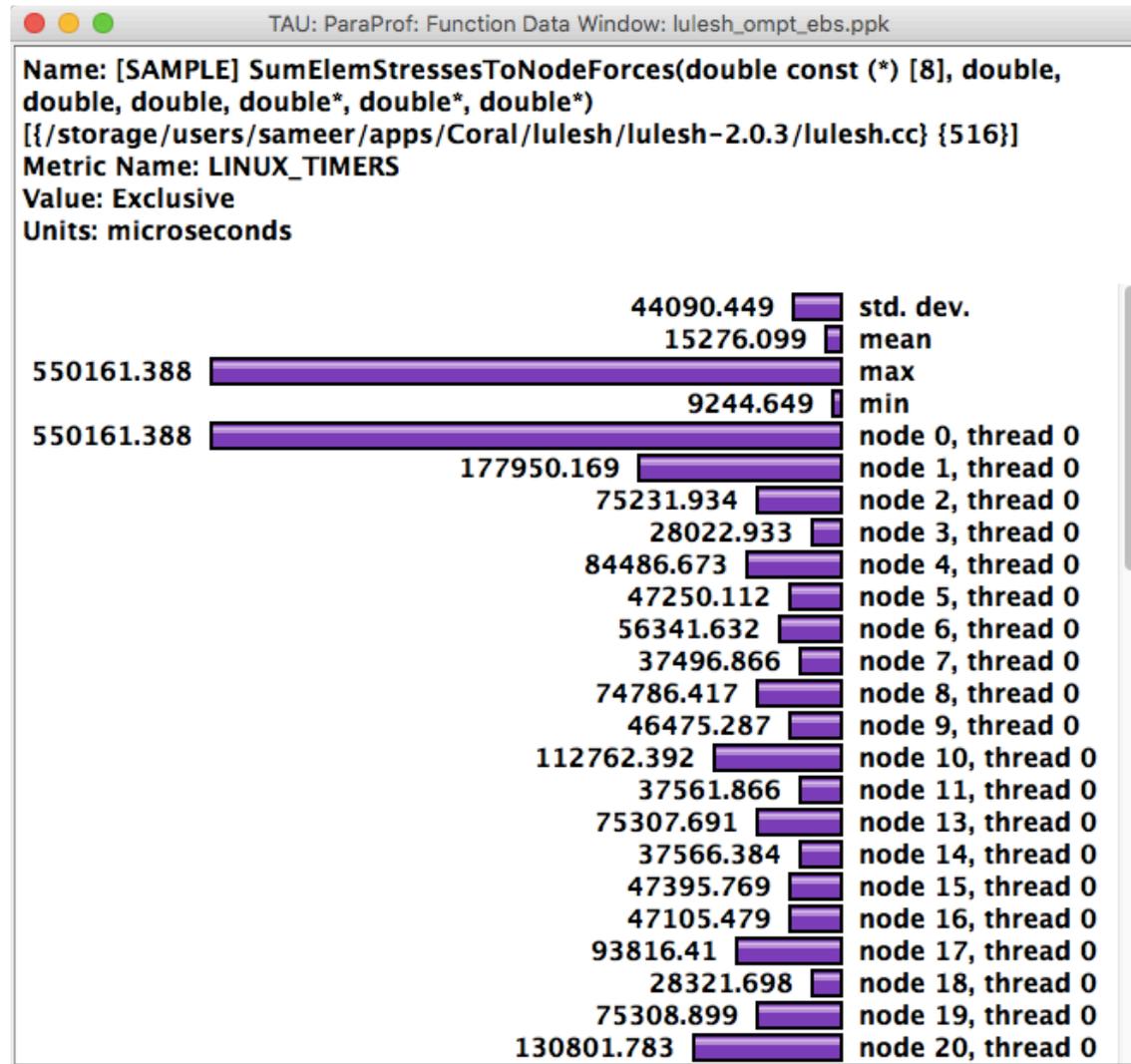
Preloads \$TAU/shared-papi-mpi-pdt/libTAU.so by matching.

```
% mpirun -np 256 tau_exec -T papi,mpi,pdt -s ./a.out
```

Does not execute the program. Just displays the library that it will preload if executed without the **-s** option.

**NOTE:** -mpi configuration is selected by default. Use **-T serial** for Sequential programs.

# TAU – Event Based Sampling (EBS)



% export TAU\_SAMPLING=1 or tau\_exec -ebs

# Python Instrumentation

TAU: ParaProf: Statistics for: node 0, thread 0 - amd\_resnet50\_fp\_1.ppk

Name	Exclusive TAUGPU_TIME	Inclusive TAUGPU_TIME	Calls	Child Calls
▾ .TAU application	0.575	182.783	1	6
▾ ▾ <module> [micro_benchmarking_pytorch.py]{1}	0.002	182.151	1	13
▾ ▾ ▾ main [micro_benchmarking_pytorch.py]{81}	0.002	168.702	1	1
▾ ▾ ▾ ▾ run_benchmarking [micro_benchmarking_pytorch.py]{40}	0.006	168.7	1	40
▾ ▾ ▾ ▾ ▾ forwardbackward [micro_benchmarking_pytorch.py]{33}	0.002	155.924	22	110
▾ ▾ ▾ ▾ ▾ ▾ backward [tensor.py]{79}	0.001	106.141	22	22
▾ ▾ ▾ ▾ ▾ ▾ ▾ backward [__init__.py]{38}	0.001	106.14	22	88
▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ run_backward	106.135	106.135	22	3
▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ pthread_create	0	0	3	0
▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ _make_grads [__init__.py]{20}	0.001	0.004	22	110
▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ isinstance	0	0	22	0
▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ len	0	0	22	0
▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ __call__ [module.py]{485}	0	49.77	22	110
▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ forward [container.py]{95}	0	49.768	22	66
▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ __call__ [module.py]{485}	0.001	49.767	44	220
▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ forward [resnet.py]{151}	0.003	49.765	22	484
▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ __call__ [module.py]{485}	0.006	49.759	220	1,100
▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ ▾ forward [container.py]{95}	0.002	45.622	88	440
▾ __call__ [module.py]{485}	0.007	45.616	352	1,760
▾ forward [resnet.py]{78}	0.071	45.598	352	6,600
▾ __call__ [module.py]{485}	0.07	45.495	3,256	16,280
▾ forward [conv.py]{319}	0.017	29.675	1,056	3,168
▾ conv2d	29.648	29.648	1,056	0
▾ __getattr__ [module.py]{523}	0.01	0.01	2,112	0
▾ forward [container.py]{95}	0.002	9.401	88	264
▾ forward [batchnorm.py]{59}	0.262	6.097	1,056	9,504

% tau\_python ./foo.py

# Identifying Wait States Using EBS

TAU: ParaProf: Statistics for: node 0, thread 0 - nt3\_baseline\_keras2.ppk

Name	Inclusive ...	Calls ▾
■ _do_call [{}session.py{}1348]	512.135	82
■ _run_fn [{}session.py{}1317]	512.134	82
▼ ■ TF_Run	512.093	82
▼ ■ [CONTEXT] TF_Run	512.173	51,211
■ [SAMPLE] __pthread_cond_wait [{} {} 0]	511.273	51,123
■ [SAMPLE] tensorflow::TensorBuffer* tensorflow::(anonymous namespace)::FromProtoField	0.42	42
■ [SAMPLE] __memcpy_ssse3_back [{} {} 0]	0.28	28
■ [SAMPLE] _int_free [{}malloc.c {} 0]	0.03	3
■ [SAMPLE] __GI___libc_malloc [{} {} 0]	0.02	2
■ [SAMPLE] std::basic_string<char, std::char_traits<char>, std::allocator<char> >::assign(s	0.02	1
■ [SAMPLE] google::protobuf::internal::MapField<tensorflow::NodeDef::NodeDef_AttrEntry,	0.02	1
■ [SAMPLE] __exchange_and_add [{} /home/msarahan/miniconda2/conda-bld/compiler_lin	0.01	1
■ [SAMPLE] void google::protobuf::internal::RepeatedPtrFieldBase::MergeFromInnerLoop<g	0.01	1
■ [SAMPLE] google::protobuf::internal::ArenaStringPtr::Destroy(std::basic_string<char, std::	0.01	1
■ [SAMPLE] std::_Hashtable<std::basic_string<char, std::char_traits<char>, std::allocator<	0.01	1
■ [SAMPLE] std::_Hashtable<tensorflow::Node*, std::pair<tensorflow::Node* const, tensorf	0.01	1
■ [SAMPLE] std::basic_string<char, std::char_traits<char>, std::allocator<char> >::_Rep::_I	0.01	1
■ [SAMPLE] std::_Hash_bytes(void const*, unsigned long, unsigned long) [{} /home/msaraha	0.01	1
■ [SAMPLE] std::basic_string<char, std::char_traits<char>, std::allocator<char> >::_M_cop	0.01	1
■ [SAMPLE] std::basic_string<char, std::char_traits<char>, std::allocator<char> >::size() cc	0.01	1
■ [SAMPLE] PyObject_Malloc [{} /home/nwani/m2u/conda-bld/python_1500576437846/woi	0.01	1
■ [SAMPLE] jemalloc_free [{} /home/nchaimov/candle/anaconda3/lib/python3.6/site-packag	0.01	1

```
% tau_python -ebs ./foo.py
```

# Kokkos and OpenMP Instrumentation

TAU: ParaProf: Statistics for: node 0, thread 0 - kokkos\_tau.ppk

Name	Exclusive...	Inclusive...	Calls	Child C...
▾ .TAU application	1.796	16.719	1	27
▾ Kokkos::parallel_for Kokkos::Example::BoxElemFixture<Kokkos::OpenMP, (Kokkos::Example::BoxElemPart::ElemOrder)0, ...	0.001	2.824	2	2
▾ OpenMP_PARALLEL_REGION: std::enable_if<std::is_same<Kokkos::Static, Kokkos::Static>::value, void>::type Kokkos::In	0	2.823	2	2
▾ OpenMP_IMPLICIT_TASK: std::enable_if<std::is_same<Kokkos::Static, Kokkos::Static>::value, void>::type Kokkos::Im	2.823	2.823	2	0
▾ Kokkos::parallel_reduce Kokkos::Example::FiniteElementIntegration<Kokkos::Example::BoxElemFixture<Kokkos::OpenMP	0	4.027	1	1
▾ OpenMP_PARALLEL_REGION: std::enable_if<std::is_same<Kokkos::Static, Kokkos::Static>::value, void>::type Kokkos::In	0	4.027	1	1
▾ OpenMP_IMPLICIT_TASK: std::enable_if<std::is_same<Kokkos::Static, Kokkos::Static>::value, void>::type Kokkos::Im	4.027	4.027	1	0
▾ Kokkos::parallel_reduce Kokkos::Example::FiniteElementIntegration<Kokkos::Example::BoxElemFixture<Kokkos::OpenMP	0.001	5.993	1	1
▾ OpenMP_PARALLEL_REGION: std::enable_if<std::is_same<Kokkos::Static, Kokkos::Static>::value, void>::type Kokkos::In	0	5.993	1	1
▾ OpenMP_IMPLICIT_TASK: std::enable_if<std::is_same<Kokkos::Static, Kokkos::Static>::value, void>::type Kokkos::Im	5.993	5.993	1	0
▾ Kokkos::parallel_reduce Kokkos::Example::LumpElemToNode<Kokkos::View<double* [8], Kokkos::OpenMP>, Kokkos::View	0.001	1.266	1	1
▾ OpenMP_PARALLEL_REGION: std::enable_if<std::is_same<Kokkos::Static, Kokkos::Static>::value, void>::type Kokkos::In	0	1.265	1	1
▾ OpenMP_IMPLICIT_TASK: std::enable_if<std::is_same<Kokkos::Static, Kokkos::Static>::value, void>::type Kokkos::Im	1.265	1.265	1	0
▾ Kokkos::parallel_reduce Kokkos::Example::LumpElemToNode<Kokkos::View<double* [8], Kokkos::OpenMP>, Kokkos::View	0.001	0.125	1	1
▾ OpenMP_PARALLEL_REGION: std::enable_if<std::is_same<Kokkos::Static, Kokkos::Static>::value, void>::type Kokkos::In	0	0.124	1	1
▾ OpenMP_IMPLICIT_TASK: std::enable_if<std::is_same<Kokkos::Static, Kokkos::Static>::value, void>::type Kokkos::Im	0.124	0.124	1	0
▸ OpenMP_PARALLEL_REGION: Kokkos::Impl::OpenMPExec::clear_scratch() [clone ._omp_fn.0] [{/home/users/sameer/apps	0	0	2	2
▸ OpenMP_PARALLEL_REGION: Kokkos::Impl::OpenMPExec::resize_scratch(unsigned long, unsigned long) [clone ._omp_fn.:	0	0	1	1
▸ OpenMP_PARALLEL_REGION: Kokkos::OpenMP::initialize(unsigned int, unsigned int, unsigned int) [clone ._omp_fn.2] [{/	0.017	0.025	1	1
▸ OpenMP_PARALLEL_REGION: Kokkos::OpenMP::initialize(unsigned int, unsigned int, unsigned int) [clone ._omp_fn.3] [{/	0	0	1	1
▸ OpenMP_PARALLEL_REGION: std::enable_if<std::is_same<Kokkos::Static, Kokkos::Static>::value, void>::type Kokkos::Impl	0	0.514	6	6
▸ OpenMP_PARALLEL_REGION: std::enable_if<std::is_same<Kokkos::Static, Kokkos::Static>::value, void>::type Kokkos::Impl	0	0.036	5	5
▸ OpenMP_PARALLEL_REGION: std::enable_if<std::is_same<Kokkos::Static, Kokkos::Static>::value, void>::type Kokkos::Impl	0	0.113	5	5

% tau\_exec -ompt ./a.out

# Event-based Sampling (EBS): CabanaMD

TAU: ParaProf: Statistics for: node 0, thread 0 - cabana.ppk

Name	Exclusive...	Inclusive...	Calls	Child Calls
└─ .TAU application	0.655	5.132	1	2,424
└─ Comm::update_halo	0.129	1.634	95	21,755
└─ [CONTEXT] Comm::update_halo	0	0.12	3	0
└─ [SAMPLE] __strlen_power8 [{} {0}]	0.09	0.09	2	0
└─ [SAMPLE] Kokkos::Impl::SharedAllocationRecord<void, void>::increment(Kokkos::Impl::SharedAllocationRecord<void, void>*) [{} /g/g20/reeve5/bin/CabanaMD]	0.03	0.03	1	0
└─ cudaDeviceSynchronize	0.991	0.991	3,043	0
└─ [CONTEXT] .TAU application	0	0.54	18	0
└─ [SUMMARY] LAMMPS_RandomVelocityGeom::reset(int, double*) [{} /g/g20/reeve5/pr/CabanaMD/src/input.h]	0.27	0.27	9	0
└─ [SAMPLE] LAMMPS_RandomVelocityGeom::reset(int, double*) [{} /g/g20/reeve5/pr/CabanaMD/src/input.h {128}]	0.09	0.09	3	0
└─ [SAMPLE] LAMMPS_RandomVelocityGeom::reset(int, double*) [{} /g/g20/reeve5/pr/CabanaMD/src/input.h {129}]	0.09	0.09	3	0
└─ [SAMPLE] LAMMPS_RandomVelocityGeom::reset(int, double*) [{} /g/g20/reeve5/pr/CabanaMD/src/input.h {130}]	0.06	0.06	2	0
└─ [SAMPLE] LAMMPS_RandomVelocityGeom::reset(int, double*) [{} /g/g20/reeve5/pr/CabanaMD/src/input.h {140}]	0.03	0.03	1	0
└─ [SUMMARY] Input::create_lattice(Comm*) [{} /g/g20/reeve5/pr/CabanaMD/src/input.cpp]	0.15	0.15	5	0
└─ [SAMPLE] Input::create_lattice(Comm*) [{} /g/g20/reeve5/pr/CabanaMD/src/input.cpp {745}]	0.03	0.03	1	0
└─ [SAMPLE] Input::create_lattice(Comm*) [{} /g/g20/reeve5/pr/CabanaMD/src/input.cpp {665}]	0.03	0.03	1	0
└─ [SAMPLE] Input::create_lattice(Comm*) [{} /g/g20/reeve5/pr/CabanaMD/src/input.cpp {721}]	0.03	0.03	1	0
└─ [SAMPLE] Input::create_lattice(Comm*) [{} /g/g20/reeve5/pr/CabanaMD/src/input.cpp {713}]	0.03	0.03	1	0
└─ [SAMPLE] Input::create_lattice(Comm*) [{} /g/g20/reeve5/pr/CabanaMD/src/input.cpp {714}]	0.03	0.03	1	0
└─ [SAMPLE] reference-cunsigned int, unsigned int, unsigned int> [{} /g/g20/reeve5/build_v100/install/kokkos/include/impl/Kokkos_ViewMapping.hpp {2740}]	0.06	0.06	2	0
└─ [SAMPLE] unsigned long Kokkos::Impl::ViewOffset<Kokkos::Impl::ViewDimension<0ul, 16ul, 3ul>, Kokkos::LayoutCabanaSlice<176, 16, 3, 0, 0, 0, 0>, void>::	0.03	0.03	1	0
└─ [SUMMARY] LAMMPS_RandomVelocityGeom::uniform0 [{} /g/g20/reeve5/pr/CabanaMD/src/input.h]	0.03	0.03	1	0
└─ [SAMPLE] LAMMPS_RandomVelocityGeom::uniform0 [{} /g/g20/reeve5/pr/CabanaMD/src/input.h {93}]	0.03	0.03	1	0
└─ Comm::exchange	0.024	0.392	6	3,371
└─ MPI_Finalize()	0.367	0.369	1	68
└─ Comm::exchange_halo	0.026	0.351	6	4,772
└─ MPI_Init()	0.323	0.323	1	0
└─ Cabana::Verlet	0.004	0.256	6	438
└─ Kokkos::parallel_for ForceLJCabanaNeigh::compute [device=0]	0.002	0.164	101	606
└─ MPI_Allreduce()	0.082	0.082	39	0
└─ [CONTEXT] MPI_Allreduce()	0	0.09	3	0
└─ [SAMPLE] __GI___sched_yield [{} {0}]	0.03	0.03	1	0
└─ [SAMPLE] pthread_spin_unlock [{} /usr/lib64/libpthread-2.17.so {0}]	0.03	0.03	1	0
└─ [SAMPLE] pthread_spin_lock [{} /usr/lib64/libpthread-2.17.so {0}]	0.03	0.03	1	0
└─ Kokkos::parallel_for Kokkos::View::initialization [device=0]	0.001	0.072	35	170
└─ Kokkos::parallel_for Kokkos::ViewFill-3D [device=0]	0.001	0.047	101	303
└─ Kokkos::parallel_reduce ForceLJCabanaNeigh::compute_energy [device=0]	0	0.042	11	77
└─ cudaLaunchKernel	0.015	0.028	527	1,581

Kokkos sample within Comm::update\_halo

Kokkos sample within top-level application code

Instrumented Kokkos::parallel\_for

Instrumented Kokkos::parallel\_reduce

EBS with Kokkos API



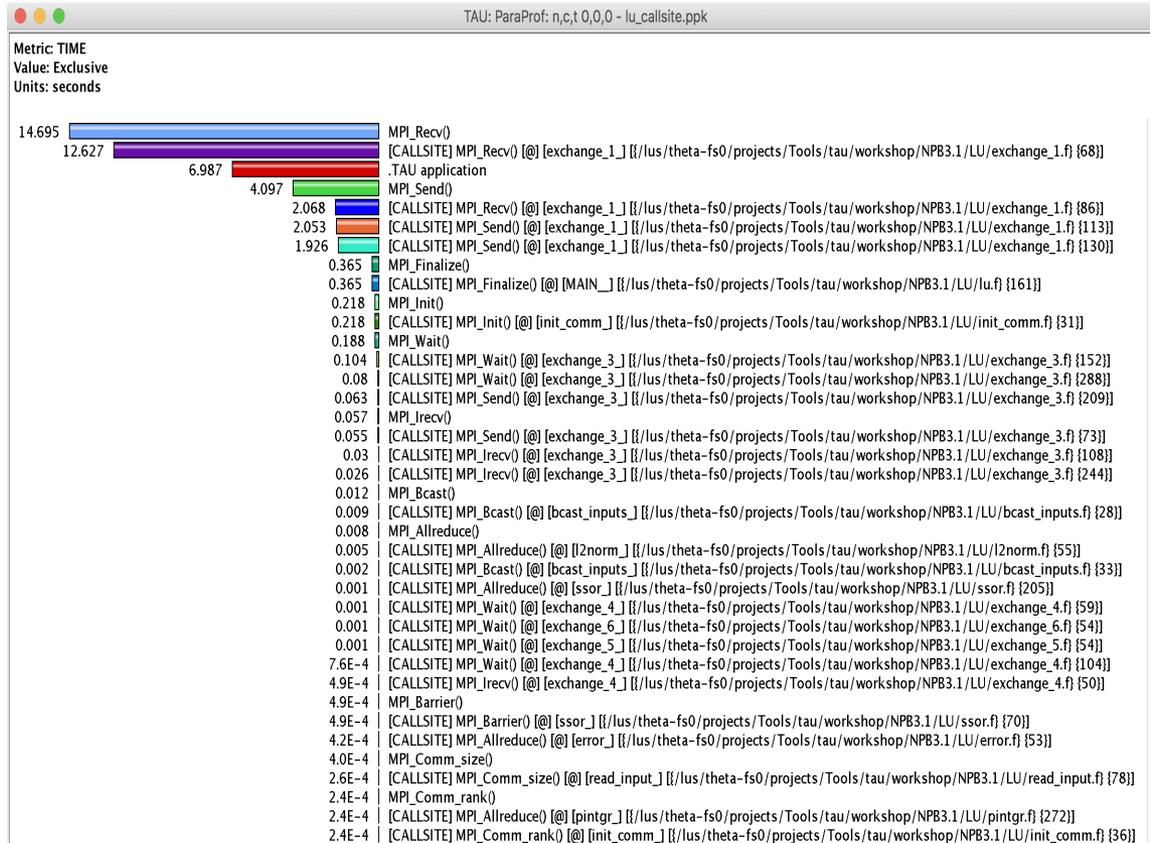
# TAU – Callsite Profiling

TAU: ParaProf: Statistics for: node 0 - clover\_callsite.ppk

Name	Excl...	Inclu...	Calls	Chil...
.TAU application	6.152	8.249	1	28,383
[CALLSITE] void start_pes_(int *) [@@] [/nfsprojects/volta-projects/tau/tau-2.24.1/craycnl/lib/libTAUsh-gnu-papi-shmem-pdt.so] UNRESOLVED ADDR	0.747	0.747	1	0
void start_pes_(int *)	0.747	0.747	1	0
void shmem_barrier_all_0	0.624	0.624	9,229	0
[CALLSITE] void shmem_barrier_all_0 [@@] [__clover_module_MOD_clover_exchange_message] [{/home/ssshend/CloverLeaf_OpenSHMEM/clover.f90} {572}]	0.401	0.401	4,610	0
[CALLSITE] void shmem_finalize_0 [@@] [/nfsprojects/volta-projects/tau/tau-2.24.1/craycnl/lib/libTAUsh-gnu-papi-shmem-pdt.so] UNRESOLVED ADDR	0.314	0.314	1	0
void shmem_finalize_0	0.314	0.314	1	0
[CALLSITE] void shmem_barrier_all_0 [@@] [__clover_module_MOD_clover_exchange_message] [{/home/ssshend/CloverLeaf_OpenSHMEM/clover.f90} {643}]	0.223	0.223	4,610	0
void shmem_put64_nb_(void *, void *, int *, int *, void *)	0.159	0.159	9,220	0
void shmem_put64_(void *, void *, int *, int *)	0.126	0.126	9,220	0
void shmem_real8_max_to_all_(void *, void *, int *, int *, int *, int *, void *, long *)	0.081	0.081	400	0
[CALLSITE] void shmem_put64_nb_(void *, void *, int *, int *, void *) [@@] [__clover_module_MOD_clover_exchange_message] [{/home/ssshend/CloverLeaf_	0.07	0.07	4,610	0
[CALLSITE] void shmem_put64_(void *, void *, int *, int *) [@@] [__clover_module_MOD_clover_exchange_message] [{/home/ssshend/CloverLeaf_OpenSHMEM	0.063	0.063	4,610	0
[CALLSITE] void shmem_real8_max_to_all_(void *, void *, int *, int *, int *, int *, void *, long *) [@@] [hydro_] [{/home/ssshend/CloverLeaf_OpenSHMEM/hydr	0.046	0.046	200	0
[CALLSITE] void shmem_real8_min_to_all_(void *, void *, int *, int *, int *, int *, void *, long *) [@@] [/nfsprojects/volta-projects/tau/tau-2.24.1/craycnl/lib/	0.04	0.04	200	0
void shmem_real8_min_to_all_(void *, void *, int *, int *, int *, int *, void *, long *)	0.04	0.04	200	0
[CALLSITE] void shmem_real8_max_to_all_(void *, void *, int *, int *, int *, int *, void *, long *) [@@] [hydro_] [{/home/ssshend/CloverLeaf_OpenSHMEM/hydr	0.036	0.036	200	0
[CALLSITE] void shmem_put64_nb_(void *, void *, int *, int *, void *) [@@] [__clover_module_MOD_clover_exchange] [{/home/ssshend/CloverLeaf_OpenSHME	0.028	0.028	601	0

% export TAU\_CALLSITE=1

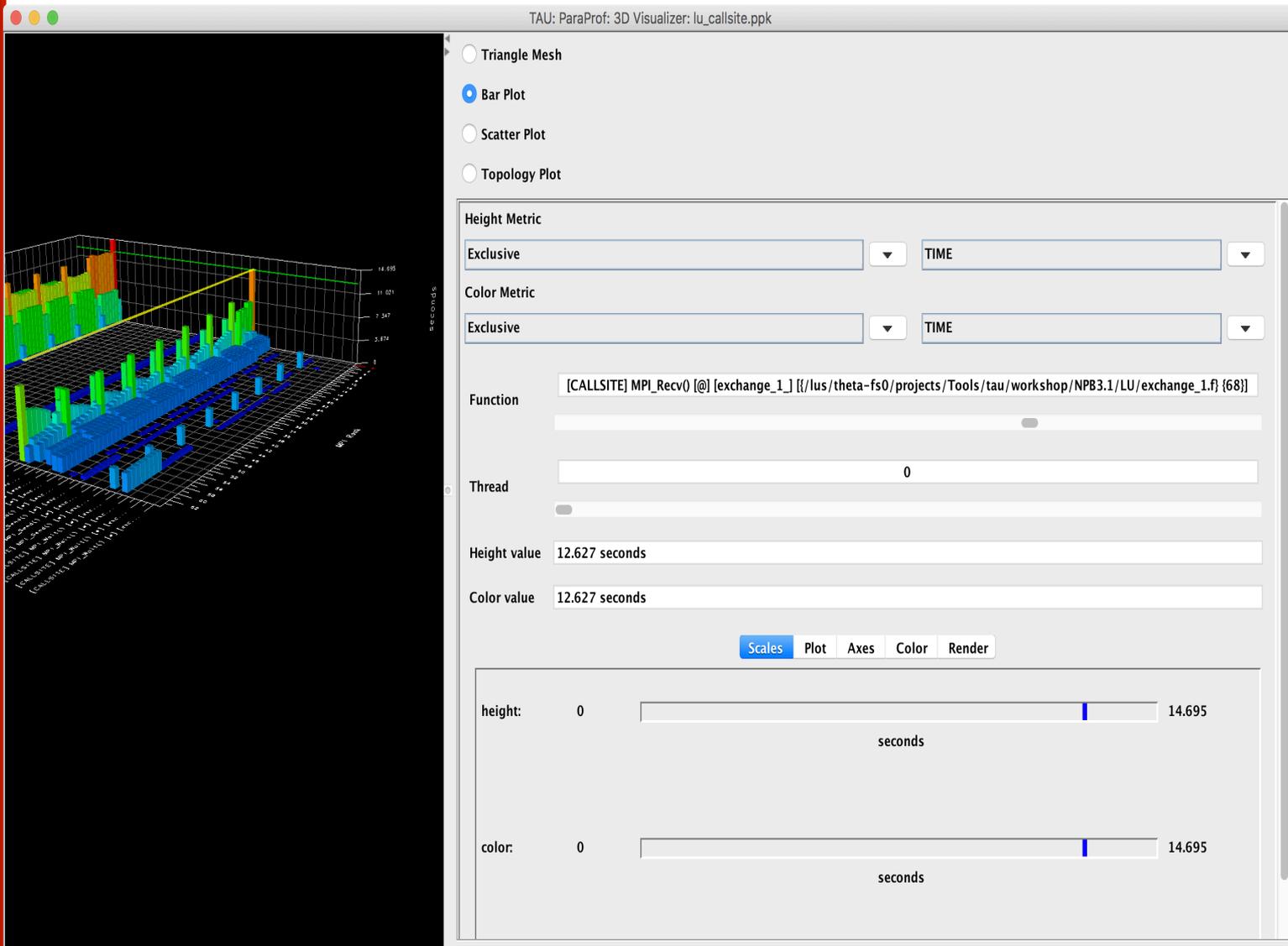
# Callsite Profiling and Tracing



% export TAU\_CALLSITE=1



# Callsite Profiling and Tracing



# TAU – Callstack Sampling

TAU: ParaProf: Statistics for: n,c,t 0,0,0 - clover\_gnu\_ebs\_unw\_call.ppk

Name	Inclusive...	Calls ▾
▾ .TAU application	34.979	1
▸ [CONTEXT] .TAU application	31.647	632
▾ void shmем_barrier_all_0	1.219	46,029
▾ [CONTEXT] void shmем_barrier_all_0	1.599	32
▾ [UNWIND] [/home/ssshend/CloverLeaf_OpenSHMEM/clover_leaf.f90.41] [ @ ] UNRESOLVED /lib64/libc-2.11.3.so	1.599	32
▾ [UNWIND] /home/ssshend/CloverLeaf_OpenSHMEM/hydro.f90.62 [ @ ] main [ { /home/ssshend/CloverLeaf_OpenSHMEM/clover_leaf.f90 } { 41 } ]	0.85	17
▾ [UNWIND] /home/ssshend/CloverLeaf_OpenSHMEM/advection.f90.102 [ @ ] hydro_ [ { /home/ssshend/CloverLeaf_OpenSHMEM/hydro.f90 } { 62 } ]	0.55	11
▾ [UNWIND] /home/ssshend/CloverLeaf_OpenSHMEM/update_halo.f90.36 [ @ ] __advection_module_MOD_advection [ { /home/ssshend/CloverLeaf_Ope	0.55	11
▾ [UNWIND] /home/ssshend/CloverLeaf_OpenSHMEM/clover.f90.292 [ @ ] __update_halo_module_MOD_update_halo [ { /home/ssshend/CloverLeaf_O	0.5	10
▾ [UNWIND] /home/ssshend/CloverLeaf_OpenSHMEM/clover.f90.572 [ @ ] __clover_module_MOD_clover_exchange [ { /home/ssshend/CloverLeaf_	0.5	10
▾ [UNWIND] UNRESOLVED [ @ ] __clover_module_MOD_clover_exchange_message [ { /home/ssshend/CloverLeaf_OpenSHMEM/clover.f90 } { 572 } ]	0.5	10
▾ [UNWIND] [/notbackedup/tmp/ulib/mpt/nightly/7.2/062215-RC/sma_dmapp/src/shmem_opt_barrier.c.118] [ @ ] UNRESOLVED /nfsproje	0.45	9
▾ [SAMPLE] _smai_smp_barrier_in [ { /notbackedup/tmp/ulib/mpt/nightly/7.2/062215-RC/sma_dmapp/src/shmem_opt_barrier.c } { 118 } ]	0.45	9
▸ [UNWIND] [/notbackedup/tmp/ulib/mpt/nightly/7.2/062215-RC/sma_dmapp/src/shmem_internal.h.88] [ @ ] UNRESOLVED /nfsprojects/\	0.05	1
▸ [UNWIND] /home/ssshend/CloverLeaf_OpenSHMEM/clover.f90.461 [ @ ] __update_halo_module_MOD_update_halo [ { /home/ssshend/CloverLeaf_O	0.05	1
▸ [UNWIND] /home/ssshend/CloverLeaf_OpenSHMEM/advection.f90.72 [ @ ] hydro_ [ { /home/ssshend/CloverLeaf_OpenSHMEM/hydro.f90 } { 62 } ]	0.15	3
▸ [UNWIND] /home/ssshend/CloverLeaf_OpenSHMEM/advection.f90.55 [ @ ] hydro_ [ { /home/ssshend/CloverLeaf_OpenSHMEM/hydro.f90 } { 62 } ]	0.15	3
▸ [UNWIND] /home/ssshend/CloverLeaf_OpenSHMEM/hydro.f90.52 [ @ ] main [ { /home/ssshend/CloverLeaf_OpenSHMEM/clover_leaf.f90 } { 41 } ]	0.5	10
▸ [UNWIND] /home/ssshend/CloverLeaf_OpenSHMEM/hydro.f90.54 [ @ ] main [ { /home/ssshend/CloverLeaf_OpenSHMEM/clover_leaf.f90 } { 41 } ]	0.25	5
▸ void start_pes_(int *)	0.508	1
▾ void shmем_real8_max_to_all_(void *, void *, int *, int *, int *, int *, void *, long *)	0.325	2,000
▾ [CONTEXT] void shmем_real8_max_to_all_(void *, void *, int *, int *, int *, int *, void *, long *)	0.5	10
▾ [UNWIND] [/home/ssshend/CloverLeaf_OpenSHMEM/clover_leaf.f90.41] [ @ ] UNRESOLVED /lib64/libc-2.11.3.so	0.5	10
▾ [UNWIND] /home/ssshend/CloverLeaf_OpenSHMEM/hydro.f90.58 [ @ ] main [ { /home/ssshend/CloverLeaf_OpenSHMEM/clover_leaf.f90 } { 41 } ]	0.45	9
▾ [UNWIND] /home/ssshend/CloverLeaf_OpenSHMEM/PdV.f90.107 [ @ ] hydro_ [ { /home/ssshend/CloverLeaf_OpenSHMEM/hydro.f90 } { 58 } ]	0.45	9
▾ [UNWIND] /home/ssshend/CloverLeaf_OpenSHMEM/clover.f90.740 [ @ ] __pdv_module_MOD_pdv [ { /home/ssshend/CloverLeaf_OpenSHMEM/PdV.f90	0.45	9
▾ [UNWIND] UNRESOLVED [ @ ] __clover_module_MOD_clover_check_error [ { /home/ssshend/CloverLeaf_OpenSHMEM/clover.f90 } { 740 } ]	0.45	9
▾ [UNWIND] [/notbackedup/tmp/ulib/mpt/nightly/7.2/062215-RC/sma_dmapp/src/shmem_reduction.h.207] [ @ ] UNRESOLVED /nfsprojects/vol	0.45	9
▾ [UNWIND] /notbackedup/tmp/ulib/mpt/nightly/7.2/062215-RC/sma_dmapp/src/shmem_opt_reduction.h.788 [ @ ] pshmem_double_max_tc	0.45	9
▾ [UNWIND] /notbackedup/tmp/ulib/mpt/nightly/7.2/062215-RC/sma_dmapp/src/shmem_opt_reduction.h.107 [ @ ] _smai_opt_double_m	0.45	9
▾ [SAMPLE] _smai_smp_reduce_double_max [ { /notbackedup/tmp/ulib/mpt/nightly/7.2/062215-RC/sma_dmapp/src/shmem_opt_reduci	0.45	9
▸ [UNWIND] /home/ssshend/CloverLeaf_OpenSHMEM/hydro.f90.54 [ @ ] main [ { /home/ssshend/CloverLeaf_OpenSHMEM/clover_leaf.f90 } { 41 } ]	0.05	1

% export TAU\_SAMPLING=1; export TAU\_EBS\_UNWIND=1

# TAU – Callpath Profiling

TAU: ParaProf: Statistics for: node 5 - fun3d\_d19.ppk

Name	Exclusive...	Inclusive...	Calls	Child...
▼ .TAU application	0	221.298	1	1
▼ NODET [{main.f90} {4,1}–{35,17}]	0	221.298	1	105
▶ FLOW::ITERATE [{flow.F90} {1692,14}]	0	197.989	100	500
▼ FLOW::INITIALIZE_DATA [{flow.F90} {465,14}]	0	22.707	1	2
▼ FLOW::INITIALIZE_DATA2 [{flow.F90} {663,14}]	0.002	22.705	1	197
▼ PPARTY_PREPROCESSOR::PPARTY_PREPROCESS [{pparty_preprocessor.f90} {28,14}]	0	20.897	1	23
▼ PPARTY_PREPROCESSOR::PPARTY_READ_GRID [{pparty_preprocessor.f90} {735,14}]	0	16.726	1	2
▼ PUNS3D_IO_C2N::PUNS3D_READ_VGRID_C2N [{puns3d_io_c2n.f90} {1543,14}]	0.011	16.725	1	11
▼ PUNS3D_IO_C2N::PUNS3D_READ_VGRID_C2N_SM [{puns3d_io_c2n.f90} {1641,14}]	0	16.656	1	5
▼ PUNS3D_IO_C2N::DISTRIBUTE_TET [{puns3d_io_c2n.f90} {1819,14}]	0.117	16.572	1	5
▼ LMPI::INTEGR_MATRIX_BCAST [{lmpi.F90} {3240,3}–{3276,36}]	0	16.448	4	4
■ MPI_Bcast()	16.448	16.448	4	0
▶ LMPI::LMPI_CONDITIONAL_STOP [{lmpi.F90} {611,3}–{672,38}]	0	0.007	1	2
▶ PUNS3D_IO_C2N::DISTRIBUTE_XYZ [{puns3d_io_c2n.f90} {2448,14}]	0.001	0.083	1	3
▶ LMPI::INTEGR_SCALAR_BCAST [{lmpi.F90} {3151,3}–{3187,36}]	0	0	3	3
▶ LMPI::LMPI_CONDITIONAL_STOP [{lmpi.F90} {611,3}–{672,38}]	0	0.058	1	2
▶ LMPI::INTEGR_SCALAR_BCAST [{lmpi.F90} {3151,3}–{3187,36}]	0	0	2	2
■ ALLOCATIONS::INTEGER_4_MY_ALLOC_PTR2 [{allocations.f90} {1010,3}–{1026,40}]	0	0	6	0
■ PUNS3D_IO_C2N::DISTRIBUTE_FAST_C2N [{puns3d_io_c2n.f90} {4226,14}]	0	0	1	0
▶ LMPI::LMPI_CONDITIONAL_STOP [{lmpi.F90} {611,3}–{672,38}]	0	0.001	1	2
▶ PPARTY_MIXED_ELEMENT::EDGE_POINTER_DRIVER [{pparty_mixed_element.f90} {74,3}–{50}]	0.65	0.873	1	174
▶ PPARTY::NODE_CELL_CHOPPER [{pparty.f90} {41,3}–{453,33}]	0.288	0.86	1	175
▶ PPARTY_PUNS3D::RAW_GRID_CHECKER [{pparty_puns3d.f90} {623,14}]	0.233	0.523	1	11
▶ PPARTY_METIS::MY_METIS [{pparty_metis.F90} {116,3}–{545,24}]	0.313	0.436	1	13,132
▶ PARTY_LMPI::PARTY_LMPI_SETUP_MPI_SM [{party_lmpi.f90} {613,3}–{686,40}]	0.006	0.337	1	10

% export TAU\_CALLPATH=1; export TAU\_CALLPATH\_DEPTH=100

# TAU Atomic Events

TAU: ParaProf: Context Events for: node 0 - /Users/sameer/tmp

Name ▾	Total	NumSamples	MaxValue	MinValue	MeanValue	Std. Dev.
Bytes Written <file=stdout>	911	62	21	1	14.694	7.441
Bytes Written <file=pipe>	22	22	1	1	1	0
Bytes Written <file=Process_Output/VelRsdL.dat>	7,826	100	302	76	78.26	22.487
Bytes Written <file=Process_Output/MomRsdL.dat>	7,826	100	302	76	78.26	22.487
Bytes Written <file=Process_Output/MassRsdL.dat>	11,325	100	435	110	113.25	32.337
Bytes Written <file=Grid_Output/bodyBndry.dat>	9,724	5	8,192	4	1,944.8	3,174.201
Bytes Written <file=/home/sameer/apps/sukra/RotCFD_Regression/case_catalog/UNS2D/N/	45	1	45	45	45	0
Bytes Written <file=./Restarts/Restart_History//NACA0012_LargeGrid_00010.Rst>	44,619,720	5,484	8,192	4	8,136.346	640.325
Bytes Written <file=./Restarts/Restart_History//NACA0012_LargeGrid_00005.Rst>	44,619,720	5,484	8,192	4	8,136.346	640.325
Bytes Written <file=./Restarts//NACA0012_LargeGrid.Rst>	44,619,720	5,484	8,192	4	8,136.346	640.325
Bytes Written <file=./Process_Output/TurbRsdL.dat>	4,271	72	224	57	59.319	19.544
Bytes Written <file=./Process_Output/Solver.out>	2,039	13	797	43	156.846	191.359
Bytes Written <file=./Field_Solutions/Solution_History/NACA0012_LargeGrid_00010.Sln>	4,356,976	534	8,192	4	8,159.131	501.319
Bytes Written <file=./Field_Solutions/Solution_History/NACA0012_LargeGrid_00005.Sln>	4,356,976	534	8,192	4	8,159.131	501.319
Bytes Written <file=./Field_Solutions/NACA0012_LargeGrid.Sln>	4,356,976	534	8,192	4	8,159.131	501.319
Bytes Written <file=./Body_Pressure/NACA0012_LargeGrid_00010_body.Prs>	65,986	9	8,190	1,300	7,331.778	2,133.204
Bytes Written <file=./Body_Pressure/NACA0012_LargeGrid_00005_body.Prs>	65,986	9	8,190	1,300	7,331.778	2,133.204
Bytes Written <file=./Body_Pressure/FrcMnt.out>	1,497	3	1,185	108	499	486.656
Bytes Written	147,107,546	18,550	8,192	1	7,930.326	1,420.552

# TAU – Context Events

TAU: ParaProf: Context Events for thread: n,c,t, 1,0,0 – samarc\_obc\_4p\_iomem\_cp.ppk

Name ▾	Total	MeanValue	NumSamples	MinValue	MaxValue	Std. Dev.
▾ .TAU application						
▶ read()						
▶ fopen64()						
▶ fclose()						
▾ OurMain()						
malloc size	25,235	1,097.174	23	11	12,032	2,851.143
free size	22,707	1,746.692	13	11	12,032	3,660.642
▾ OurMain [{{wrapper.py}}{3}]						
▶ read()						
malloc size	3,877	323.083	12	32	981	252.72
free size	1,536	219.429	7	32	464	148.122
▶ fopen64()						
▶ fclose()						
▾ <module> [{{obe.py}}{8}]						
▾ writeRestartData [{{samarcInterface.py}}{145}]						
▾ samarcWriteRestartData						
▾ write()						
WRITE Bandwidth (MB/s) <file="samarc/restore.00002/nodes.00004/proc.00001">		74.565	117	0	2,156.889	246.386
WRITE Bandwidth (MB/s) <file="samarc/restore.00001/nodes.00004/proc.00001">		77.594	117	0	1,941.2	228.366
WRITE Bandwidth (MB/s)		76.08	234	0	2,156.889	237.551
Bytes Written <file="samarc/restore.00002/nodes.00004/proc.00001">	2,097,552	17,927.795	117	1	1,048,576	133,362.946
Bytes Written <file="samarc/restore.00001/nodes.00004/proc.00001">	2,097,552	17,927.795	117	1	1,048,576	133,362.946
Bytes Written	4,195,104	17,927.795	234	1	1,048,576	133,362.946
▶ open64()						

Write bandwidth per file

Bytes written to each file

# E4S: Extreme-scale Scientific Software Stack

<https://e4s.io>



- **Spack** [<http://spack.io>] is the primary means for software delivery
- **SDKs**: collection of related ECP ST products where coordination across package teams will improve usability and practices, and foster community growth among teams that develop similar and complimentary capabilities. An SDK involves several products.
- **Containers of pre-built binaries of ECP ST products.**
- **Container runtimes supported**
  - Docker
  - Charliecloud
  - Shifter
  - Singularity
  - PCOCC at CEA
  - Inception at NCAR
- **VirtualBox Open Virtualization Appliance (OVA) image that contains these runtimes**
- **MPI replacement strategies to use native network interconnect**

# Integration and Interoperability: E4S

**E4S is released twice a year. SC19 release of E4S v1.0:**

- Containers and turn-key, from-source builds of 80+ popular HPC software packages
- 50 full release ECP ST products including:
  - MPI: MPICH and OpenMPI
  - Development tools: TAU, HPCToolkit, and PAPI
  - Math libraries: PETSc and Trilinos
  - Data and Viztools: Adios, HDF5
- Limited access to 10 additional ECP ST products
- Docker
- Singularity
- Shifter
- Charliecloud
- Inception
- Open Virtualization Appliance (OVA) for VirtualBox features Spack, E4S containers, and support for container environments

# Reproducible, Customizable Container Builds & Spack Mirrors

E4S provides base images and recipes for building Docker containers based on SDKs

- Git: <https://github.com/UO-OACISS/e4s>
- E4S provides **build caches for Spack for native bare-metal as well as container builds based installation** of ST products
- Build caches: <https://oaciss.uoregon.edu/e4s>
  - **The build cache model can be extended to target platforms**, and can be managed by facilities staff when appropriate.

# E4S Build Cache Binaries

<https://oaciss.uoregon.edu/e4s/inventory.html>

## E4S Build Cache for Spack 0.13.3

To use this build cache, you must set these environment variables and add the cache to Spack

```
AWS_ACCESS_KEY_ID=spack
AWS_SECRET_ACCESS_KEY=spack-public-user
S3_ENDPOINT_URL=https://instinct.nic.uoregon.edu:8083
spack mirror add e4s s3://uoregon/spack
```

Click on one of the packages below to see a list of all available variants.

Last updated: 01-11-2020 21:28 PST  
2352 Spack binaries in the build cache

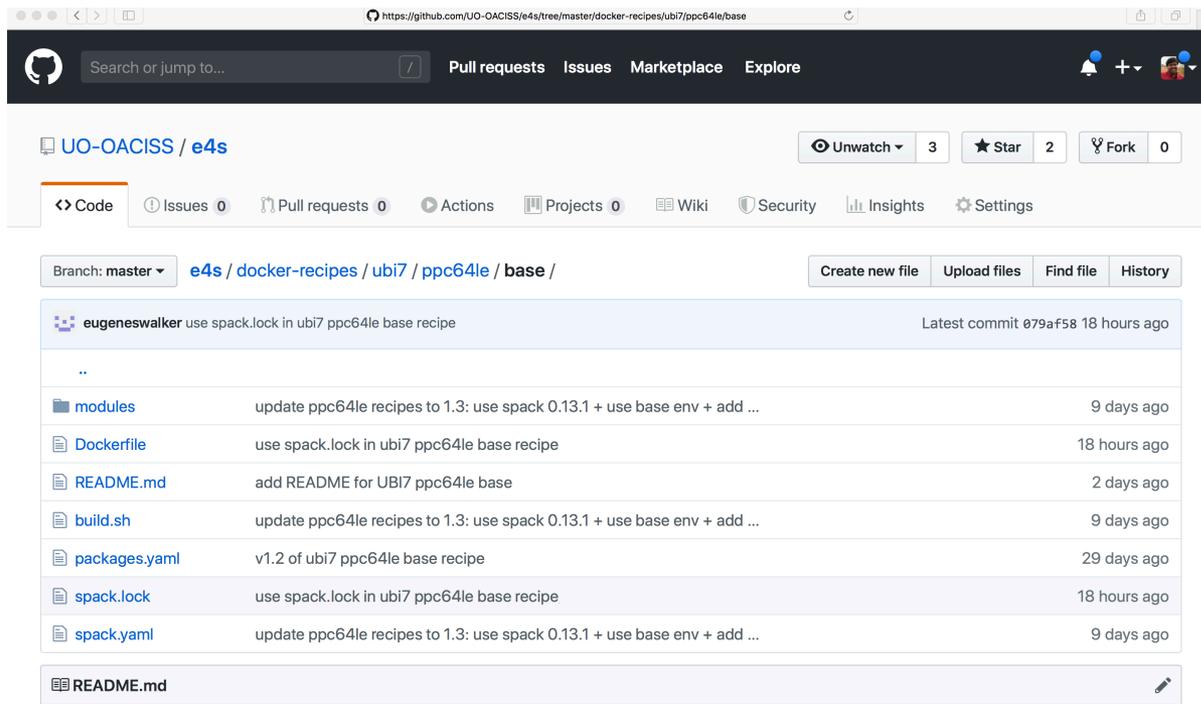
adios2@2.5.0 adios@1.13.1 adlbox@0.9.2 aml@0.1.0 ant@1.10.0 argobots@1.0rc1 autoconf@2.69 automake@1.16.1 axl@0.1.1 bdftopcf@1.0.5 binutils@2.31.1  
binutils@2.32 bison@3.4.2 bmi@develop bolt@1.0rc2 boost@1.70.0 bzip2@1.0.8 c-blosc@1.17.0 caliper@2.0.1 catalyst@5.6.0 cinch@develop cmake@3.15.4  
cuda@10.1.243 curl@7.63.0 darshan-runtime@3.1.7 darshan-util@3.1.7 diffutils@3.7 doxygen@1.8.15 dtcmp@1.1.0  
dyninst@10.1.0

Click on the full spec link to find out more.

Link	Arch	OS	Compiler	Created	Full Hash
<a href="#">Full Spec</a>	ppc64le	centos7	gcc@7.3.0	12-28-2019 11:28 PST	lqf23rj2bm25uisu7bhygaulsvfen3mp
<a href="#">Full Spec</a>	ppc64le	rhel7	gcc@7.3.0	12-28-2019 11:29 PST	gqv6ooyscns76p2ecw7orxtlcfdrukhlh
<a href="#">Full Spec</a>	ppc64le	ubuntu 18.04	gcc@7.3.0	12-28-2019 11:30 PST	g5spch6wljgioonp6oca6clfyz3ggesyh
<a href="#">Full Spec</a>	x86_64	centos7	gcc@7.3.0	12-28-2019 11:28 PST	ymtprwfbjfg2f653sjudziqqrbdwkezo
<a href="#">Full Spec</a>	x86_64	centos8	gcc@7.3.0	12-28-2019 11:29 PST	qaric7cxnuiho2eb2e2bjxkyjstpbihu
<a href="#">Full Spec</a>	x86_64	rhel7	gcc@7.3.0	12-28-2019 11:29 PST	cj5rmqj5xistrle6wyvbkcc33ih4c7zka
<a href="#">Full Spec</a>	x86_64	rhel8	gcc@7.3.0	12-28-2019 11:30 PST	dlkbegrw7gpy3gpgptymv3kyqieckleo
<a href="#">Full Spec</a>	x86_64	ubuntu 18.04	gcc@7.3.0	12-28-2019 11:30 PST	torjxbex4k4kuj66dge37dehjawi3fo

elfutils@0.177 environment-modules@4.3.1 er@0.0.3 exmcutils@0.5.7 expat@2.2.9 faodel@1.1906.1 findutils@4.6.0 flatcc@0.5.3 flecci@develop flex@2.6.4 font-util@1.3.2

# Reproducible Container Builds using E4S Base Images



<https://github.com/UO-OACISS/e4s>

- E4S base image (UBI 7.6) has Spack build cache mirror and GPG key installed.
- Base image has GCC and MPICH configured for MPICH ABI level replacement (with system MPI).
- **Customized container build using binaries from E4S Spack build cache for fast deployment.**
- **No need to rebuild packages from the source code.**
- Same recipe for container and native bare-metal builds with Spack!

# Docker Recipes on GitHub

The screenshot shows a GitHub repository page for the file `spack.yaml` in the `ubi7/x86_64/e4s` directory of the `docker-recipes` repository. The page includes a commit message by Wyatt Spear, 2 contributors, and the content of the `spack.yaml` file. The file content is as follows:

```
1  spack:
2    packages:
3      all:
4        compiler: [gcc@7.3.0]
5        variants: +mpi
6        providers:
7          mpi: [mpich]
8        buildable: true
9        version: []
10       paths: {}
11       modules: {}
```

<https://github.com/UO-OACISS/e4s>

- Create your own optimized container
- Base images
- SDKs

# Reproducible ppc64le Base Images on Dockerhub

The screenshot shows the Docker Hub interface. At the top, there's a navigation bar with the Docker Hub logo, a search bar, and links for Explore, Repositories, Organizations, and Get Help. The user profile 'exascaleproject' is visible. Below the navigation bar, the search results for 'ppc64le' are displayed. The results show three repositories: 'ubuntu1804\_ppc64le\_base', 'ubi7\_ppc64le\_base', and 'centos7\_ppc64le\_base', all updated 2 days ago. A tip at the bottom suggests switching namespaces. On the right, there's a sidebar with 'Organizations' (ecpcontainers, ecpe4s, ecpsdk) and a 'Download Docker Desktop' button.

Repository	Stars	Downloads	Visibility
ecpe4s / ubuntu1804_ppc64le_base	0	7	PUBLIC
ecpe4s / ubi7_ppc64le_base	0	7	PUBLIC
ecpe4s / centos7_ppc64le_base	0	10	PUBLIC

- Hub.docker.com
- ecpe4s
- Ubuntu 18.04
- RHEL/UBI 7.6
- Centos 7.6

# Integration and Interoperability: E4S on AWS

- E4S AWS public image ami-063e830287b86155c (US-West-2 Oregon) has following container runtimes:
  - Docker
  - Shifter
  - Singularity
  - Charliecloud
- Spack with base PMR components
- E4S full featured Singularity image
  - (exascaleproject/sdk:AHM19)
- Used in ISC-HPC 2019 tutorials
- **Used as base image for NASA GEOS-Chem E4S public image**
- Resources provided by AWS AI/ML team



The screenshot shows the AWS Management Console interface. The main content area displays a table of AMIs owned by the user. The table has columns for Name, AMI Name, AMI ID, Source, Owner, Visibility, and Status. The first row is selected, showing details for the AMI ami-016555a769a29afeb.

Name	AMI Name	AMI ID	Source	Owner	Visibility	Status
GEOS Chem E4S container with Spack, Docker, ...	E4S-GEOS-Chem	ami-016555a769a29afeb	792568971918/E4S-GEOS-Chem	792568971918	Public	available
SuperLU Tutorial E4S Singularity	E4S_SC_SuperLU_Tutorial	ami-07d0fb5dab32444f	792568971918/E4S_SC_SuperLU_Tutorial	792568971918	Public	available
ECP E4S image with AI and HPC software stacks...	import-ami-07af056b52139562	ami-063e830287b86155c	792568971918/import-ami-07af056b52139562	792568971918	Public	available
E4S container with Spack, Docker, Singularity, SH...	import-ami-0d68aa1dc4496567c	ami-0997e7325abb44f	792568971918/import-ami-0d68aa1dc4496567c	792568971918	Public	available

Image: ami-016555a769a29afeb			
Details	Permissions	Tags	
AMI ID	ami-016555a769a29afeb	AMI Name	E4S-GEOS-Chem
Owner	792568971918	Source	792568971918/E4S-GEOS-Chem
Status	available	State Reason	-
Creation date	July 30, 2019 at 2:04:49 PM UTC-7	Platform	Other Linux
Architecture	x86_64	Image Type	machine
Virtualization type	hvm	Description	E4S GEOS-Chem AWS AMI
Root Device Name	/dev/sda1	Root Device Type	eb6
RAM disk ID	-	Kernel ID	-
Product Codes	-	Block Devices	/dev/sda1-snap-06f38fc656d7760cb:80:false:gp2

# Support Acknowledgments

## US Department of Energy (DOE)

- ANL
- Office of Science contracts, ECP
- SciDAC, LBL contracts
- LLNL-LANL-SNL ASC/NNSA contract
- Battelle, PNNL and ORNL contract



## US Department of Defense (DoD)

- HPCMP



## National Science Foundation (NSF)

- SI2-SSI, CSSI



## NASA

## CEA, France



## Partners:

- University of Oregon
- The Ohio State University
- ParaTools, Inc.
- University of Tennessee, Knoxville



UNIVERSITY OF OREGON



THE OHIO STATE UNIVERSITY

ParaTools

THE UNIVERSITY of TENNESSEE

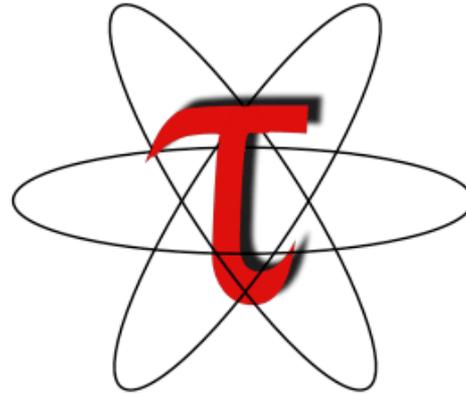
# Acknowledgment



**“This research was supported by the Exascale Computing Project (17-SC-20-SC), a collaborative effort of two U.S. Department of Energy organizations (Office of Science and the National Nuclear Security Administration) responsible for the planning and preparation of a capable exascale ecosystem, including software, applications, hardware, advanced system engineering, and early testbed platforms, in support of the nation’s exascale computing imperative.”**

<http://exascaleproject.org>

# Download TAU



<http://tau.uoregon.edu>

<http://taucommander.com>

<http://www.hpclinux.com> [OVA for VirtualBox]

<https://e4s.io>

[ Extreme-Scale Scientific Software Stack, Containers for HPC ]

Free download, open source, BSD license

# Reference

# Installing and Configuring TAU

## •Installing PDT:

- `wget tau.uoregon.edu/pdt_lite.tgz`
- `./configure --prefix=<dir>; make ; make install`

## •Installing TAU:

- `wget tau.uoregon.edu/tau.tgz; tar xzf tau.tgz; cd tau-2.<ver>`
- `wget http://tau.uoregon.edu/ext.tgz ; tar xf ext.tgz`
- `./configure -bfd=download -pdt=<dir>`  
`-iowrapper -mpi -dwarf=download -unwind=download -`  
`otf=download -papi=<dir>`  
`make install`

## •Using TAU:

- `export TAU_MAKEFILE=<taudir>/<arch>/lib/Makefile.tau-<TAGS>`
- `make CC=tau_cc.sh CXX=tau_cxx.sh F90=tau_f90.sh`

# Runtime Environment Variables

Environment Variable	Default	Description
TAU_TRACE	0	Setting to 1 turns on tracing
TAU_TRACE_FORMAT	Default	Setting to "otf2" turns on TAU's native OTF2 trace generation (configure with --otf=download)
TAU_EBS_UNWIND	0	Setting to 1 turns on unwinding the callstack during sampling (use with tau_exec -ebs or TAU_SAMPLING=1)
TAU_EBS_RESOLUTION	line	Setting to "function" or "file" changes the sampling resolution to function or file level respectively.
TAU_TRACK_LOAD	0	Setting to 1 tracks system load on the node
TAU_SELECT_FILE	Default	Setting to a file name, enables selective instrumentation based on exclude/include lists specified in the file.
TAU_OMPT_SUPPORT_LEVEL	basic	Setting to "full" improves resolution of OMPT TR6 regions on threads 1.. N-1. Also, "lowoverhead" option is available.
TAU_OMPT_RESOLVE_ADDRESS_EAGERLY	1	Setting to 1 is necessary for event based sampling to resolve addresses with OMPT. Setting to 0 allows the user to do offline address translation.

# Runtime Environment Variables

Environment Variable	Default	Description
TAU_TRACK_MEMORY_LEAKS	0	Tracks allocates that were not de-allocated (needs <code>-optMemDbg</code> or <code>tau_exec -memory</code> )
TAU_EBS_SOURCE	TIME	Allows using PAPI hardware counters for periodic interrupts for EBS (e.g., <code>TAU_EBS_SOURCE=PAPI_TOT_INS</code> when <code>TAU_SAMPLING=1</code> )
TAU_EBS_PERIOD	100000	Specifies the overflow count for interrupts
TAU_MEMDBG_ALLOC_MIN/MAX	0	Byte size minimum and maximum subject to bounds checking (used with <code>TAU_MEMDBG_PROTECT_*</code> )
TAU_MEMDBG_OVERHEAD	0	Specifies the number of bytes for TAU's memory overhead for memory debugging.
TAU_MEMDBG_PROTECT_BELOW/ ABOVE	0	Setting to 1 enables tracking runtime bounds checking below or above the array bounds (requires <code>-optMemDbg</code> while building or <code>tau_exec -memory</code> )
TAU_MEMDBG_ZERO_MALLOC	0	Setting to 1 enables tracking zero byte allocations as invalid memory allocations.
TAU_MEMDBG_PROTECT_FREE	0	Setting to 1 detects invalid accesses to deallocated memory that should not be referenced until it is reallocated (requires <code>-optMemDbg</code> or <code>tau_exec -memory</code> )
TAU_MEMDBG_ATTEMPT_CONTINUE	0	Setting to 1 allows TAU to record and continue execution when a memory error occurs at runtime.
TAU_MEMDBG_FILL_GAP	Undefined	Initial value for gap bytes
TAU_MEMDBG_ALINGMENT	Sizeof(int)	Byte alignment for memory allocations
TAU_EVENT_THRESHOLD	0.5	Define a threshold value (e.g., .25 is 25%) to trigger marker events for min/max