Performance Evaluation of GPU-accelerated HPC and Al applications using HPCToolkit, TAU, and ParaTools Pro for $E4S^{TM}$

.....

Sameer Shende

Research Professor, University of Oregon

President and Director

ParaTools, Inc.

Presentation at EnergyHPC conference, Rice U., Houston

February 27, 2025

https://tinyurl.com/e4stut



TAU Performance System®

- Versatile profiling and tracing toolkit that supports:
 - MPI, ROCm, CUDA, DPC++/SYCL (Level Zero), OpenCL, and OpenMP (OpenMP Tools Interface for Target Offload), OpenACC
- Scalable, portable, performance evaluation toolkit for HPC and AI/ML workloads that supports:
 - C++/C/DPC++, Fortran, Python
- Interfaces with Program Database Toolkit (PDT) and SALT-FM for instrumentation of source code.
- Supports PAPI, Likwid for hardware performance counter information
- Instrumentation includes support for PETSc (Perfstubs), XGC (CAMTIMERS), Kokkos, MPI, pthread, event-based sampling, GPU runtimes
- A single tool (tau_exec) is used to launch un-instrumented, un-modified binaries
- TAU's paraprof, pprof, perfexplorer for profile analysis; Vampir, Jumpshot, Perfetto.dev for traces
- <u>http://tau.uoregon.edu</u>



Using TAU's Runtime Preloading Tool: tau_exec

- Preload a wrapper that intercepts the runtime system call and substitutes with another
 - **MPI**
 - o OpenMP
 - o POSIX I/O
 - Memory allocation/deallocation routines
 - Wrapper library for an external package
- No modification to the binary executable!
- Enable other TAU options (communication matrix, OTF2, event-based sampling)

TAU: Quickstart Guide

Profiling:

MPI: % mpirun -np 16 tau exec -ebs ./a.out

- Pthread: % mpirun -np 16 tau exec -T mpi,pthread -ebs ./a.out
- CUDA: % mpirun -np 16 tau_exec -T cupti -cupti ./a.out
- ROCM: % mpirun -np 16 tau exec -rocm ./a.out

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

Tracing:

- Vampir: MPI: % export TAU_TRACE=1; export TAU_TRACE_FORMAT=otf2
 % mpirun -np 16 tau exec ./a.out; vampir traces.otf2 &
- Chrome/Jumpshot: % export TAU_TRACE=1; mpirun -np 64 tau_exec ./a.out
 % tau_treemerge.pl;

Chrome: % tau_trace2json tau.trc tau.edf -chrome -ignoreatomic -o app.json Chrome browser: chrome://tracing (Load -> app.json) or Perfetto.dev

• Jumpshot: tau2slog2 tau.trc tau.edf -o app.slog2; jumpshot app.slog2

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 –l0 ./a.out % 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,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 export TAU_METRICS=TIME,PAPI_L1_DCM... -ebs_resolution=<file | function | line>

Advanced Instrumentation Features of TAU

TAU Performance System[®]

Parallel performance framework and toolkit

- Supports all HPC platforms, compilers, runtime system
- Provides portable instrumentation, measurement, analysis





Application Performance Engineering using TAU

- How much time is spent in each application routine and outer *loops*? Within loops, what is the contribution of each *statement*? What is the time spent in OpenMP loops? In kernels on GPUs. How long did it take to transfer data between host and device (GPU)?
- How many instructions are executed in these code regions? Floating point, Level 1 and 2 data cache misses, hits, branches taken? What is the extent of vectorization for loops?
- How much time did my application spend waiting at a barrier in MPI collective operations?
- What is the memory usage of the code? When and where is memory allocated/de-allocated? Are there any memory leaks? What is the memory footprint of the application? What is the memory high water mark?
- How much energy does the application use in Joules? What is the peak power usage?
- What are the I/O characteristics of the code? What is the peak read and write *bandwidth* of individual calls, total volume?
- How does the application *scale*? What is the efficiency, runtime breakdown of performance across different core counts?

Instrumentation

- Direct and indirect performance observation
- Instrumentation invokes performance measurement
- Direct measurement with probes
- Indirect measurement with periodic sampling or hardware performance counter overflow interrupts
- Events measure performance data, metadata, context, etc.
- User-defined events

Interval (start/stop) events to measure exclusive & inclusive duration
 Atomic events take measurements at a single point
 Measures total, samples, min/max/mean/std. deviation statistics
 Context events are atomic events with executing context
 Measures above statistics for a given calling path

Instrumentation

Add hooks in the code to perform measurements

Source instrumentation using a preprocessor

- Add timer start/stop calls in a copy of the source code.
- Use SALT-FM or 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.
- TAU LLVM plugin for selective instrumentation at the LLVM IR level before codegen.
- Requires recompiling using TAU compiler scripts (tau_cc.sh, tau_f90.sh...)
- Runtime preloading of TAU's Dynamic Shared Object (DSO)
 - No need to recompile code! Use **mpirun tau_exec ./app** with options.

Using Instrumentation of Source Code in TAU using PDT

TAU supports several measurement and thread options

Phase profiling, profiling with hardware counters, MPI library, CUDA...

Each measurement configuration of TAU corresponds to a unique stub makefile and library that is generated when you configure it

To instrument source code automatically using PDT

Choose an appropriate TAU stub makefile in <arch>/lib:

(or module load tau...)

% export TAU_MAKEFILE=\$TAU/Makefile.tau-papi-mpi-pdt

% export TAU_OPTIONS= '-optVerbose ...' (see tau_compiler.sh)

% export PATH=\$TAUDIR/x86_64/bin:\$PATH

Use tau_f90.sh, tau_cxx.sh, tau_upc.sh, or tau_cc.sh as F90, C++, UPC, or C compilers respectively:

% mpif90 foo.f90 changes to

% tau_f90.sh foo.f90

Set runtime environment variables, execute application and analyze performance data:

% pprof (for text based profile display)

% paraprof (for GUI)

TAU Configurations available

% module load tau % Is STAU/Makefile* /packages/tau-2.34/x86 64/lib/Makefile.tau-intel-papi-mpi-pthread-pdt /packages/tau-2.34/x86 64/lib/Makefile.tau-intel-papi-ompt-mpi-pdt-openmp /packages/tau-2.34/x86 64/lib/Makefile.tau-papi-mpi-pdt /packages/tau-2.34/x86 64/lib/Makefile.tau-papi-pdt /packages/tau-2.34/x86 64/lib/Makefile.tau-papi-pthread-pdt /packages/tau-2.34/x86 64/lib/Makefile.tau-papi-tbb-pdt For an uninstrumented binary: % mpirun – np 16 tau exec – T mpi, papi, pdt ./a.out Picks the configuration represented by /packages/tau-2.34/x86 64/lib/Makefile.tau-intel-papi-mpi-pdt To use OpenMP instrumentation: % export TAU OMPT SUPPORT LEVEL=full % export OMP_NUM_THREADS=<N> % mpirun –np 16 tau exec –T ompt, mpi –ompt –ebs ./a.out To use TAU's source code instrumentation: % export TAU MAKEFILE=/packages/tau-2.34/x86 64/lib/Makefile.tau-intel-papi-mpi-pdt % make CC=tau cc.sh F90=tau f90.sh CXX=tau cxx.sh ; mpirun -np 16./a.out % pprof –a | more % paraprof % paraprof --pack foo.ppk

TAU's Legacy Static Analysis System: Program Database Toolkit (PDT)



Automatic Source Instrumentation using PDT



Problems with PDT

- PDT has usability issues
 - Not compatible with any C++ features beyond C++11
 - Closed source so have to ship large binary blobs to end users
 - PDB file requirement to serialize all types to textual description presents problems with modern C++ with deeply-templated types (large number of types with long description length)
- PDT has maintainability issues
 - Closed source developers must sign NDA, limits number of developers involved
 - PDT instrumentor mixes EDG parser and PDT client code in same codebase difficult to upgrade to new versions
 - Difficult to support for new C++ features
 - Major EDG release makes numerous breaking changes to API, which would require a from-scratch rewrite to adopt
- Fortran support
 - Based on older GFortran parser that does not support modern Fortran features

Flang-based Instrumentation



SALT plugin for Flang: NASA SBIR Phase I

- Phase I proof-of-concept implements instrumentation for performance monitoring.
- SALT Visitor visits each node of parse tree
 - Identifies programs/subprograms for instrumentation
 - Identifies source locations for inserting variable declarations and timer start/stop code for TAU timers
 - Not all parse tree node types in Flang have associated source position data.
 - Scans into children of a given node to find source position data in subnodes.
- Phase I focused specifically on instrumentation.

Profiling and Tracing

Description Construction 23.697 Construction 36.697 Construction <

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

Tracing shows you when the events take place on a timeline



Inclusive vs. Exclusive values

- Inclusive
 - Information of all sub-elements aggregated into single value
- Exclusive
 - Information cannot be subdivided further



How much data do you want?



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



Instrumentation

 $t_{3} t_{4}$ $t_{6} t_{7} t_{8}$ t₉ ι₅ τ₁₀ ι₂ 11¹12¹13 ۲₁₄ Time foo(0) foo(1) foo(2) int main() main Measurement int i; TAU START ("main"); Measurement code is inserted such that every event of for (i=0; i < 3; i++) interest is captured directly foo(i); TAU_STOP("main"); Can be done in various ways return 0; Advantage: Much more detailed information void foo(int i) Disadvantage: TAU_START ("foo"); Processing of source-code / executable if (i > 0)necessary foo(i - 1);Large relative overheads for small functions TAU_STOP("foo");

ParaProf Profile Browser

File Options Windows Help Metric: TIME Value: Exclusive Std. Dev. Mean Max Min node 0 node 1 node 2 node 3 node 4 node 5 node 6 node 7 node 8 node 9 node 10 node 11 node 12 node 13 node 14 node 15 node 16 node 17 node 18 node 19 node 20 node 21 node 22 node 23 node 24 node 25 node 26 node 27 node 28 node 29 node 30 node 31 -

% paraprof

ParaProf Profile Browser



Inclusive Measurements

	TAU: ParaProf: node 0 - fun3d_d19.ppk	
Metric: TIME		
Units: seconds		
221,305	.TAU application	
221.304	NODET [{main.f90} {4.1}-{35.17}]	
197.989	FLOW::ITERATE [{flow.F90} {1692,14}]	
195.577	FLOW::STEP_SOLVER [{flow.F90} {1845,14}]	
195.569	RELAX_STEADY::RELAX [{relax_steady.f90} {30,3}-{307,22}]	
61.927	UPDATE_MEAN::UPDATE_LINEAR_SYSTEM_MEAN [{update_mean.F90} {195,3}-{275,42}]	
61.28	UPDATE_MEAN::UPDATE_JACOBIAN_DRIVER_MEAN [{update_mean.F90} {460,3}-{505,44}]	
61.275	UPDATE_MEAN::UPDATE_JACOBIAN [{update_mean.F90} {513,3}-{588,32}]	
61.258	FILL_JACOBIANS::FILL_JACOBIAN [{fill_jacobians.f90} {19,3}-{341,30}]	
59.068	GCR_SOLVE::GCR_SOLVER_QSET [{gcr_solve.f90} {47,3}-{415,32}]	
57.635	GCR_SOLVE_UTIL::GCR_PRECONDITIONER_QSET [{gcr_solve_util.f90} {40,3}-{131,40}]	
57.152	POINT_SOLVER::POINT_SOLVE [{point_solver.F90} {31,3}-{214,28}]	
56.882	UPDATE_MEAN::UPDATE_RHS_MEAN [{update_mean.F90} {102,3}-{185,32}]	
54.402	KELAX_MEAN::KELAX [{relax_mean.f90} {22,3}-{84,22}]	
53.103	LINEARSOLVE_NODIVCHECK:NODIVCHECK_RELAX_Q [{IInearSolve_nodivcheck.r90} {56,14}]	
52.807	$= \text{CPDATE_MEANRESIDUAL_S [(update_mean.r90) (42,5)=(94,27)]}$ $= \text{ELN2D_RES_ELOW(upds_ELOW [(fup2d_ros_flow f00) (27,2)=(270,25)]}$	
52.800	ELIV: DESIDIAL COMPRESSIBLE [[flux fon) (25,3)-(27,3)-(27,3,25)]	
52.730	POINT SOLVER: POINT SOLVE 5 [(noint solver EQ0) (2700 3)-(2921 30)]	
52.744	Loop: POINT_SOLVER::POINT_SOLVE_5 [{point_solver.F90} {2757.5}-{2917.19}]	
36.232	ACOBIAN VISCOUS::VISCOUS ACOBIAN [{iacobian viscous.f90} {20.14}]	
36.231	ACOBIAN VISCOUS::EDGEIP [{iacobian viscous.f90} {324.14}]	
36.231	Loop: JACOBIAN_VISCOUS::EDGEJP [{jacobian_viscous.f90} {440,7}-{1584,22}]	
27.474	FLUX_PERFGAS::INVISCID_FLUX_DRIVER [{flux_perfgas.f90} {37,14}]	
27.474	FLUX_PERFGAS::ROE_FLUX [{flux_perfgas.f90} {236,14}]	
27.473	Loop: FLUX_PERFGAS::ROE_FLUX [{flux_perfgas.f90} {397,5}-{597,20}]	
22.707	FLOW::INITIALIZE_DATA [{flow.F90} {465,14}]	
22.694	FLOW::INITIALIZE_DATA2 [{flow.F90} {663,14}]	
20.916	PPARTY_PREPROCESSOR::PPARTY_PREPROCESS [{pparty_preprocessor.f90} {28,14}]	
16.726	PPARTY_PREPROCESSOR::PPARTY_READ_GRID [{pparty_preprocessor.f90} {735,14}]	
16.726	PUNS3D_IO_C2N::PUNS3D_READ_VGRID_C2N [{puns3d_io_c2n.f90} {1543,14}]	
16.657	PUNS3D_IO_C2N::PUNS3D_READ_VGRID_C2N_SM [{puns3d_io_c2n.f90} {1641,14}]	
14.159	GRADIENT_DRIVER::GRAD_VARIABLE [{gradient_driver.190} {110,3}-{508,30}]	
13.852	UPDATE_TUKB::UPDATE_KHS_TUKB [{update_turb.190} {742,3}-{845,32}]	

Exclusive Time



Callpath Profiling

TAU: ParaProf: Statistics for: node 0 - fun3d_d19.ppk				
Name	Exclusive… ♥ In	clusive	Calls (Child
TAU application	0.001	221.305	1	1
NODET [{main.f90} {4,1}-{35,17}]	0	221.304	1	105
FLOW::INITIALIZE_PROJECT [{flow.F90} {366,14}]	0	0.517	1	9
FLOW::ITERATE [{flow.F90} {1692,14}]	0	197.989	100	500
FLOW::STEP_POST [{flow.F90} {2098,14}]	0.001	2.394	100	1,202
FLOW::STEP_SOLVER [{flow.F90} {1845,14}]	0.001	195.577	100	702
RELAX_STEADY::RELAX [{relax_steady.f90} {30,3}-{307,22}]	0.049	195.569	100	800
UPDATE_TURB::UPDATE_VALUES_TURB [{update_turb.f90} {854,3}-{877,35}]	0.479	0.737	100	300
RELAX_TURB::RELAX [{relax_turb.f90} {22,3}-{68,22}]	0.024	4.77	100	300
RELAX_MEAN::RELAX [{relax_mean.f90} {22,3}-{84,22}]	0.002	54.402	100	300
WU_DEFS::TIMES [{wu_defs.f90} {59,3}-{174,22}]	0.003	0.065	200	200
GCR_SOLVE::GCR_SOLVER_QSET [{gcr_solve.f90} {47,3}-{415,32}]	0.002	54.334	100	801
GCR_UTIL::RES_RMS_QSET [{gcr_util.f90} {375,3}-{395,29}]	0.001	0.15	100	100
GCR_UTIL::MATRIX_TO_GRID_RES [{gcr_util.f90} {313,3}-{336,35}]	0.001	0.536	100	100
GCR_UTIL::MATRIX_TO_GRID_DQ [{gcr_util.f90} {282,3}-{305,34}]	0.001	0.195	100	100
GCR_UTIL::GRID_TO_MATRIX_RES [{gcr_util.f90} {344,3}-{367,35}]	0	0.341	100	100
GCR_SOLVE_UTIL::GCR_PRECONDITIONER_QSET [{gcr_solve_util.f90} {40,3}-{131,40}]	0	53.104	100	100
LINEARSOLVE_NODIVCHECK::NODIVCHECK_RELAX_Q [{linearsolve_nodivcheck.F90} {56,14}]	0.008	53.103	100	4,900
WU_DEFS::TIMES [{wu_defs.f90} {59,3}-{174,22}]	0.02	0.34	3,200	3,200
POINT_SOLVER::POINT_SOLVE [{point_solver.F90} {31,3}-{214,28}]	0.004	52.751	1,500	1,500
POINT_SOLVER::POINT_SOLVE_5 [{point_solver.F90} {2700,3}-{2921,30}]	0.003	52.747	1,500	1,500
Loop: POINT_SOLVER::POINT_SOLVE_5 [{point_solver.F90} {2757,5}-{2917,19}]	43.649	52.744	1,500	36,000
LMPI_APP::SINGLE_START_MATRIX_XFER [{Impi_app.F90} {7907,3}-{8132,41}]	0.271	0.512	18,000	85,500
LMPI_APP::SINGLE_MATRIX_COMPLETE_XFER [{Impi_app.F90} {11520,3}-{11626,44}]	0.228	8.583	18,000	30,000
LMPI::LMPI_WAITALL [{lmpi.F90} {20175,3}-{20200,29}]	0.139	8.355	30,000	30,000
MPI_Waitall()	8.217	8.217	30,000	0
LMPI::INTEGR_SCALAR_REDUCE [{Impi.F90} {4584,3}-{4611,37}]	0	0.002	100	100
IINEAR_SPECTRAL::SET_FIELD_POINTS [{linear_spectral.f90} {173,3}-{184,33}]	0	0.002	100	200

% export TAU_CALLPATH=1 % export TAU_CALLPATH_DEPTH=100

ParaProf Function Window



Callsite Profiling and Tracing

TAU: ParaProf: n,c,t 0,0,0 - lu callsite.ppk Metric: TIME Value: Exclusive Units: seconds 14.695 MPI Recv() 12.627 [CALLSITE] MPI Recv() [@] [exchange 1] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/exchange 1.f} {68}] 6.987 .TAU application 4.097 MPI Send() 2.068 [CALLSITE] MPI_Recv() [@] [exchange_1_] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/exchange_1.f] {86}] 2.053 [CALLSITE] MPI_Send() [@] [exchange_1_] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/exchange_1.f} {113}] 1.926 [CALLSITE] MPI_Send() [@] [exchange_1_] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/exchange_1.f} {130}] 0.365 MPI Finalize() 0.365 [CALLSITE] MPI_Finalize() [@] [MAIN_] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/lu.f} {161}] 0.218 MPI Init() 0.218 [CALLSITE] MPI_Init() [@] [init_comm_] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/init_comm.f} {31}] 0.188 🛽 MPI_Wait() 0.104 [[CALLSITE] MPI_Wait() [@] [exchange_3_] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/exchange_3.f] {152}] 0.08 [CALLSITE] MPI_Wait() [@] [exchange_3_] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/exchange_3.f} {288}] 0.063 [CALLSITE] MPI_Send() [@] [exchange_3_] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/exchange_3.f] {209]] 0.057 MPI Irecv() 0.055 [[CALLSITE] MPI_Send() [@] [exchange_3_] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/exchange_3.f] {73}] 0.03 [CALLSITE] MPI_Irecv() [@] [exchange 3] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/exchange 3.f} {108]] 0.026 [CALLSITE] MPI_Irecv() [@] [exchange_3_] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/exchange_3.f} {244}] 0.012 | MPI Bcast() 0.009 [CALLSITE] MPI Bcast() [@] [bcast inputs] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/bcast inputs.f} {28}] 0.008 MPI Allreduce() 0.005 | [CALLSITE] MPI Allreduce() [@] [l2norm] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/l2norm.f} {55}] [CALLSITE] MPI_Bcast() [@] [bcast_inputs_] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/bcast_inputs.f} {33}] 0.002 0.001 [CALLSITE] MPI_Allreduce() [@] [ssor_] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/ssor.f} {205}] 0.001 [CALLSITE] MPI_Wait() [@] [exchange_4_] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/exchange_4.f} {59}] 0.001 [CALLSITE] MPI Wait() [@] [exchange 6] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/exchange 6.f} {54}] [CALLSITE] MPI Wait() [@] [exchange 5] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/exchange 5.f} {54}] 0.001 7.6E-4 [CALLSITE] MPI Wait() [@] [exchange 4] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/exchange 4.f} {104}] 4.9E-4 [[CALLSITE] MPI_Irecv() [@] [exchange_4.] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/exchange_4.f} {50}] 4.9E-4 MPI Barrier() 4.9E-4 [CALLSITE] MPI Barrier() [@] [ssor] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/ssor.f} {70}] 4.2E-4 [CALLSITE] MPI Allreduce() [@] [error] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/error.f} {53}] 4.0E-4 MPI Comm size() 2.6E-4 [[CALLSITE] MPI Comm size() [@] [read input] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/read input.f} {78}] 2.4E-4 MPI Comm rank() 2.4E-4 | [CALLSITE] MPI_Allreduce() [@] [pintgr_] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/pintgr.f] {272}] 2.4E-4 [[CALLSITE] MPI_Comm_rank() [@] [init_comm_] [{/lus/theta-fs0/projects/Tools/tau/workshop/NPB3.1/LU/init_comm.f} {36}]

ParaProf Callpath Thread Relations Window

•••	TAU: ParaProf: Call Path Data n,c,t, 5,0,0 - fun3d_d19.ppk				
Metric I Sorted I Units: s	Name: TIME By: Exclusive econds				
	Exclusive	Inclusive	Calls/Tot.Calls	Name[id]	
>	46.323 46.323 0.16 0.166	53.014 53.014 6.379 0.311	1500/1500 1500 18000/30100 18000/30100	<pre>POINT_SOLVER::POINT_SOLVE_5 [{point_solver.F90} {2700,3}-{2921,30}] Loop: POINT_SOLVER::POINT_SOLVE_5 [{point_solver.F90} {2757,5}-{2917,19}] LMPI_APP::SINGLE_MATRIX_COMPLETE_XFER [{lmpi_app.F90} {11520,3}-{11626,44}] LMPI_APP::SINGLE_START_MATRIX_XFER [{lmpi_app.F90} {7907,3}-{8132,41}]</pre>	
>	36.386 36.386	36.386 36.386	78/78 78	JACOBIAN_VISCOUS::EDGEJP [{jacobian_viscous.f90} {324,14}] Loop: JACOBIAN_VISCOUS::EDGEJP [{jacobian_viscous.f90} {440,7}-{1584,22}]	
>	4.7E-4 0.242 0.002 16.689 0.013 0.587 0.003 1.521 1.2E-5 0.012 0.055 10.328 0.005 29.457	4.7E-4 0.242 0.002 16.689 0.013 0.587 0.003 1.521 1.2E-5 0.012 0.055 10.328 0.005 29.457	45/27956 777/27956 2/27956 31/27956 48/27956 197/27956 197/27956 2/27956 1040/27956 3/27956 18941/27956 10/27956 27956	<pre>PPARTY_METIS::MY_METIS [{pparty_metis.F90} {116,3}-{545,24}] LMPI::LOGICL_SCALAR_BCAST [{lmpi.F90} {3691,3}-{3727,36}] LMPI::DOUBLE_TENSOR_BCAST [{lmpi.F90} {4185,3}-{4214,36}] LMPI::INTEGG_MATRIX_BCAST [{lmpi.F90} {3240,3}-{3276,36}] LMPI::INTEGG_VECTOR_BCAST [{lmpi.F90} {3376,3}-{3412,36}] LMPI::INTEGR_VECTOR_BCAST [{lmpi.F90} {3196,3}-{3232,36}] LMPI::CHARACTER_BCAST [{lmpi.F90} {4096,3}-{4132,36}] LMPI::DOUBLE_VECTOR_BCAST [{lmpi.F90} {4096,3}-{4132,36}] LMPI::DOUBLE_VECTOR_BCAST [{lmpi.F90} {4051,3}-{4087,36}] LMPI::INTEGS_SCALAR_BCAST [{lmpi.F90} {3331,3}-{3367,36}] LMPI::INTEGS_SCALAR_BCAST [{lmpi.F90} {3151,3}-{3187,36}] LMPI::DOUBLE_MATRIX_BCAST [{lmpi.F90} {4140,3}-{4176,36}] MPI_Bcast()</pre>	
>	27.158 27.158	27.158 27.158	100/100 100	<pre>FLUX_PERFGAS::ROE_FLUX [{flux_perfgas.f90} {236,14}] Loop: FLUX_PERFGAS::ROE_FLUX [{flux_perfgas.f90} {397,5}-{597,20}]</pre>	

Shows the contribution of parents and children for each routine (marked by an arrow)

ParaProf Callpath Thread Relations Window

•••				TAU: ParaProf: Call Path Data n.c.t, 13,0,0 - fun3d_d19.ppk	
Metric Sorted Units: s	Name: TIME By: Exclusive econds				
	Exclusive	Inclusive	Calls/Tot.Calls	Name[id]	
>	45.642 45.642 0.299 0.6	52.774 52.774 6.259 0.873	1500/1500 1500 18000/30100 18000/30100	<pre>POINT_SOLVER::POINT_SOLVE_5 [{point_solver.F90} {2700,3}-{2921,30}] Loop: POINT_SOLVER::POINT_SOLVE_5 [{point_solver.F90} {2757,5}-{2917,19}] LMPI_APP::SINGLE_MATRIX_COMPLETE_XFER [{lmpi_app.F90} {11520,3}-{11626,44}] LMPI_APP::SINGLE_START_MATRIX_XFER [{lmpi_app.F90} {7907,3}-{8132,41}]</pre>	
>	37.689 37.689	37.689 37.689	78/78 78	JACOBIAN_VISCOUS::EDGEJP [{jacobian_viscous.f90} {324,14}] Loop: JACOBIAN_VISCOUS::EDGEJP [{jacobian_viscous.f90} {440,7}-{1584,22}]	
>	28.431 28.431	28.431 28.431	100/100 100	<pre>FLUX_PERFGAS::ROE_FLUX [{flux_perfgas.f90} {236,14}] Loop: FLUX_PERFGAS::ROE_FLUX [{flux_perfgas.f90} {397,5}-{597,20}]</pre>	
>	0.003 0.542 0.033 0.005 16.724 0.032 1.48 1.5E-5 0.002 0.013 6.1E-4 5.481 0.243 24.557	$\begin{array}{c} 0.003\\ 0.542\\ 0.033\\ 0.005\\ 16.724\\ 0.032\\ 1.48\\ 1.5E-5\\ 0.002\\ 0.013\\ 6.1E-4\\ 5.481\\ 0.243\\ 24.557\end{array}$	1197/27956 489/27956 3/27956 10/27956 31/27956 5371/27956 2/27956 2/27956 48/27956 48/27956 45/27956 18941/27956 777/27956 27956	LMPI::CHARACTER_BCAST [{lmpi.F90} {3100,3}-{3136,32}] LMPI::INTEGR_VECTOR_BCAST [{lmpi.F90} {3196,3}-{3232,36}] LMPI::INTEG8_SCALAR_BCAST [{lmpi.F90} {3331,3}-{3367,36}] LMPI::DOUBLE_MATRIX_BCAST [{lmpi.F90} {4140,3}-{4176,36}] LMPI::DOUBLE_MATRIX_BCAST [{lmpi.F90} {4051,3}-{4087,36}] LMPI::DOUBLE_VECTOR_BCAST [{lmpi.F90} {4056,3}-{4132,36}] LMPI::DOUBLE_VECTOR_BCAST [{lmpi.F90} {3736,3}-{3772,36}] LMPI::DOUBLE_TENSOR_BCAST [{lmpi.F90} {3376,3}-{3412,36}] LMPI::INTEG8_VECTOR_BCAST [{lmpi.F90} {3151,3}-{3412,36}] LMPI::INTEG8_SCALAR_BCAST [{lmpi.F90} {3151,3}-{3167,36}] LMPI::INTEG8_SCALAR_BCAST [{lmpi.F90} {3151,3}-{3727,36}] MPI_BCASC [{lmpi.F90} {3151,3}-{3727,36}]	
>	20.045 20.045 1.4E-4 0.006 0.003 0.008 3.2E-4 0.443	$\begin{array}{c} 61.19\\ 61.19\\ 1.4E-4\\ 2.491\\ 0.003\\ 0.008\\ 37.689\\ 0.445\end{array}$	78/78 78 78/78 3822/16665 3822/8622 7644/17444 78/78 78/123	UPDATE_MEAN::UPDATE_JACOBIAN [{update_mean.F90} {513,3}-{588,32}] FILL_JACOBIANS::FILL_JACOBIAN [{fill_jacobians.f90} {19,3}-{341,30}] SOURCE::SOURCE_JACOBIAN [{source.f90} {93,3}-{168,32}] LMPI::LMPI_CONDITIONAL_STOP [{lmpi.F90} {611,3}-{672,38}] BC_NAMES::BC_HAS_PRESSURE_CLOSURE [{bc_names.f90} {1618,3}-{1693,38}] BC_NAMES::ELEMENT_BASED_BC [{bc_names.f90} {1390,3}-{1439,31}] JACOBIAN_VISCOUS::VISCOUS_JACOBIAN [{jacobian_viscous.f90} {20,14}] TIMEACC::TIME_DIAG_MC [{timeacc.f90} {1067,3}-{1330,29}]	

Identifying Collective Wait States: Thread Callpath Relations Window

Metric Name: TIME Sorted By: Exclusive Units: seconds Inclusive Calls/Tot.Calls Exclusive Name[id] 1/1 1099.614 1191.772 i:SETUP 1099.614 1191.772 1 i:LOAD --> MPI_Allreduce() 92.158 3/9543 0.006 9.8E-4 9.8E-4 11/15177 MPI_Gatherv() 1.448 1.448 43/15177 MPI Gather() 15.353 15.353 46/15177 MPI Alltoall() 89.821 89.821 4311/15177 MPI_Bcast() 6.777 6.777 MPI Allgather() 195/15177 68.678 68.678 991/15177 MPI Reduce() 9.179 9.179 12/15177 MPI_Comm_dup() MPI_Allgatherv() 0.125 0.125 25/15177 382.861 382.861 9543/15177 MPI Allreduce() 574.243 574.243 15177 MPI Collective Sync --> DISTRIBUTE_FØG 2.507 2.508 10/186 2.433 2.434 10/186 F UPD FØ SP 5.156 5.158 20/186 FØ CHARGE SEARCH INDEX 5.507 5.505 22/186 PULLBACK WEIGHT 24.86 24.872 UPDATE_PTL_WEIGHT 102/186 0.473 0.473 2/186 MAIN LOOP 4.975 4.977 DIAG f0 PORT1 PTL 20/186 45.91 45.93 186 copy_ptl_to_device --> Kokkos::parallel_for set_buffer_particles_d [type = Cuda, device = 0] 0.02 0.02 186/272

TAU: ParaProf: Call Path Data n,c,t, 118,0,0 - 128_d3d.ppk

MPI Collective Sync is the time spent in a barrier operation inside a collective

ParaProf Thread Comparison Window



Comparing Rank 118 with 22. Right click on "node 118" -> Add node to comparison window

ParaProf Function Window



Tracing: Jumpshot (ships with TAU)



% export TAU_TRACE=1; mpirun -np 256 tau_exec ./a.out % tau_treemerge.pl; tau2slog2 tau.trc tau.edf -o app.slog2

% jumpshot app.slog2
Tracing: Chrome Browser or Perfetto.dev

Chrome://tracing × +						
$\leftarrow \rightarrow \mathbb{C}$ (Ghrome chrome://tracing						९ 🛧 🥥 😝 :
Record Save Load matmult.json					Processes View Options	$\leftarrow \rightarrow $ » ?
	380 m	·	385 ms		· · ·	390 ms
						×
~ 0			.TAU app	lication		ie
			MAIN [{matmult.f90]	} {39,7}-{132,22}]		Size
	MPI_Recv() M MPI MPI M	PI_Recv() MPI_Recv() M	MPI MPI_Re MPI_Recv()	MPI_Recv() M	M MPI_Recv() MPI_Recv()	MPI_Recv() M M MPI_Recv() MPI
						X S
~ 0			.TAU app	lication		
			MAIN [{matmult.f90]	} {39,7}-{132,22}]		
	MULTIPLY_MATRICES [{matmult.f90} {25,7	-{37,38}] MULTIPLY_MATR	ICES [{matmult.f90} {25,7}-{37,38}]	MULTIPLY_MATRIC	CES [{matmult.f90} {25,7}-{37,38}]	MULTIPLY_MATRICES [{matmult.f90} {25,7}-{37,38}]
						X
~ 0			.TAU app	lication		
			MAIN [{matmult.f90]	} {39,7}-{132,22}]		Fra
	MULTIPLY_MATRICES [{matmult.f90}	MULTIPLY_MATRICES [{matmult.	.f90} {25,7}-{37,38}] MULTIPLY	_MATRICES [{matmult.f9	90} {25,7}-{37,38}] MULTIPLY	2_MATRICES [{matmult.f90} {25,7}-{37,38}] MUL ∃
✓ Process 3						X Dat
▼ 0			.TAU app	lication		
			MAIN [{matmult.f90]	} {39,7}-{132,22}]		~
	MULTIPLY_MATRICES [{matmult.f90} {25,7]	-{37,38}] MULTIPLY_MATRI	CES [{matmult.f90} {25,7}-{37,38}]	MULTIPLY_MATRIC	ES [{matmult.f90} {25,7}-{37,38}]	MULTIPLY_MATRICES [{matmult.fs + 7}-{37,38}]
✓ Process 4						
▼ 0			.TAU app	lication		
				{39,7}-{132,22}]		
	MOLTIPLT_MAT MOLTIPLT_MATH	ICES [{matmult.190} {25,7}-{37,36}]	MOLTIPLY_MATRICES [{matrices	un:190} {25,7}-{37,36}]	MOLTIPLE MATRICES (maine	MULTIP 1 11 HICES
79 items selected. Slices (79)						
Name 🗸	Wall Duration 🔝	Self time 🤝 Average Wall	Duration v Occurrences v	Event(s)	Link	
MULTIPLY_MATRICES [{matmult.f90} {25,7	<u>}-{37,38}]</u> 🔍 📕 114.459 ms	114.459 ms	4.239 ms 27	Incoming flow	<u>MPI</u>	
MPI_Send()_	0.157 ms	0.157 ms	0.007 ms 23	Incoming flow	MPI	
MPI_Recv()_	0.994 ms	0.994 ms	0.043 ms 23	Incoming flow	<u>MPI</u>	
.TAU application	2,274.680 ms	0.174 ms	758.227 ms 3	Incoming flow	MPI	
MAIN [{matmult.f90} {39,7}-{132,22}]	2,274.506 ms	13.845 ms	758.169 ms 3	Incoming flow	MPI	
Totals	4,664.796 ms	129.629 ms	59.048 ms 79	Incoming flow	MPI	
				Incoming flow	MPI	
Selection start			0.196 ms	Incoming flow	MPI	
Selection extent			759.696 ms	Incoming flow	MPI	
				Incoming flow	MPI	
				Incoming flow	MPI	
Snapz Pro	x			Incoming flow	MPI	

% 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)

Perfetto.dev (open the UI)

Perfetto.dev Trace Browser: Kokkos

🔍 🔍 🔍 🕼 trace_events.0.json.ş	gz (1 MB) · × +							~
← → C 🔒 ui.perfetto.dev/#	#!/viewer?local_cache_key=00000000	-0000-0000	0-8c28-e38	b6fa725bb				ė ☆ 🛊 🏵 E
Perfetto ≡					Q Search			1
		0.5		52.4 ms	104.9 ms 157.3 ms	209.7 ms 262.2 ms	314.6 ms 367.1 ms	419.5 ms 471.9 ms
Navigation								
🖿 Open trace file								
Open with legacy UI	1647381313.4 s + 485.9 ms	+108.5 us	+308.5 us +508	5 us +708.3 us +	08.4 us +1.1 ms +1.3 ms +1.5 ms +1.7 ms +1.9 ms +2.1 ms +2.3 ms +2.5 ms	-2.7 ms +2.9 ms +3.1 ms +3.3 ms +3.5 ms +3.7 ms +	3.9 ms +4.1 ms +4.3 ms +4.5 ms +4.7 ms +4.9 ms +5.1 ms +5.3 ms	+5.5 ms +5.7 ms +5.9 ms +6.1 ms +6.3 ms +6.5 ms +6.7 ms +6.9 ms
Record new trace	×							
	 Process 0 0 							
Current Trace	CPU Thread 0.0					APEX MAIN Main Thread: APEX main thread		
trace_events.0.json.gz (1 MB)		hipMemcpyTo	oSymbol h	hi	241		pthread_join	
Show timeline	GPU: Bytes Allocated value	10 M						
🞍 Download	GPU: Bytes Freed value							
👶 Query (SQL)	GPU: CopyDeviceToHost Bytes value	0						
() Metrics	GPU: FillBuffer Bytes value	0						
Info and stats	GPU: Host Bytes Allocated value	2.5 M						
Ť	GPU: Total Bytes Occupied on Device value	100 M						
Convert trace	value	2.5 M				OS Thra	sad' APEY nthread wranner	
	CDI Thread 2.2					ad Kokkos	Idra<0x7efe007e2dd0>	Kokkosmarallel reduce Type:HIP Dev
Switch to legacy UI	GFO THEAD 5 5					hipFuncGetAttributes	him hiph	ostMalloc h h h h h hipStreamCreate
Convert to Json						OS TI	hread: APEX pthread wrapper	
Everyole Transa	CPU Thread 4 4				hipMalloc h. hipMal. h K. h. h		addr≈≺0x7efe007e2dd0> Kokkos∷parallel_for [Type:HIP, Device: 0] initialize	Kokkos::parallel_reduce [Type:HIP, D.,
Example fraces							h_ h	II
Open Android example						05	Thread: APEX pthread wrapper	
Open Chrome example							addr= <ux e2ddu="" eleuu=""></ux>	Malla susceptibility (see 1995)
	CPU Thread 5 5				прмат. п к. пр. п прм. п к.	Kol	kkos:::parallel_for [Type:HIP, Device: 0] initialize	Ko Kokkos::parallel_reduce [Type:HiP, D
	CPU Thread 5 5				hpmai. n k. hp. n hpm. n k.	Ко	kkos_parallel_for (Type:HIP, Device: 0) initialize	Kow Kokkos:paranet_reduce type:hii/ D
Support	CPU Thread 5 5 HIP [0:00000] 65536					Kol	kkos:parallel_for(Type:HIP, Device: 0) initialize	
Support	CPU Thread 5 5 HIP (0:00000) 65536 HIP (0:00001) 131072					Kol	kon:parallel. for (Type:HID Device: 0) initialize	
Support (?) Keyboard shortcuts (?) Documentation	CPU Thread 5 5 HIP (0:00000) 65536 HIP (0:00001) 131072 Current Selection Flow Events					Kol	Assessment of the state of initialize	
Support Keyboard shortcuts Documentation Firstags	CPU Thread 5 5 HIP (0:00000) 65536 HIP (0:00001) 131072 Current Selection Flow Events Slice Details						Adoc_parallel.for (Type sill) Device: 0 initialize	
Support Keyboard shortcuts Documentation Flags Report a bug	CPU Thread 5 5 HIP (20000) 65536 HIP (20000) 131072 Current Selection Flow Events Slice Details Name Category	OS	Thread: APEX	 pthread wrappe			koc-paralel.for (Type SII) Device: 0) initialize	
Support Keyboard shortcuts Documentation Filags Report a bug	CPU Thread 5 5 HIP (0:0000) 65536 HIP (0:0000) 131072 Current Selection Flow Events Slice Details Name Category Start time	OS CP 48	Thread: APEX PU 7ms 53us 394	pthread wrappe			koc-paralel.for (Typestil) Device: 0) initialize	
Support Keyboard shortcuts Documentation Flags Report a bug Sample queries	CPU Thread 5 5 HIP (0:0000) 65536 HIP (0:0000) 131072 Current Selection Flow Events Silce Details Name Category Start time Duration Thread duration	OS CP 48: 377 0s	S Thread: APEX PU 77ms 53us 394 ms 130us	pthread wrappe			koc-paralel.for (Type SII) Device: 0) initialize	
Support Keyboard shortcuts Documentation Flags Report a bug Sample queries Compute summary statistics	CPU Thread 5 5 HIP (0:0000) 65586 HIP (0:0000) 131072 Current Selection Elew Events Silce Details Name Category Start time Duration Thread duration Silce ID ame CIID	0S CP 48: 37/1 0S 57 7	3 Thread: APEX 70 77ms 53us 3944 ms 130us	pthread wrappo			koc-paralel.for (Type Sil) Device: 0) initialize	
Support	CPU Thread 5 5 HIP (0.0000) 65586 HIP (0.0000) 131072 Current Selection Flow Events Silce Details Name Category Start time Duration Thread duration Silce ID args. R2int GUID args. R2int GUID	0S 0P 481 37/1 0S 57 7-464 40	S Thread: APEX PU 7ms 53us 394 ims 130us 511686018427	pthread wrappe 15 387903			koc-parallel.for (Type SII) Device: 0) initialize	

- % export TAU_TRACE=1; mpirun -np 64 tau_exec -rocm ./a.out;
- % tau_treemerge.pl;
- % tau_trace2json tau.trc tau.edf -chrome -ignoreatomic -o app.json

Perfetto.dev

🗕 🔍 🜒 External trace (RPC) - Per	fetto × +												~
\leftrightarrow \rightarrow C \cong ui.perfetto.dev/#!/	/viewer?local_cache_key										Ô ☆	🍖 🖈 🛛	@ :
\bigcirc Perfetto =				Q Search									
Navigation ^		0 s 5.4	\$	10.7 s	16.1 s	21.4 s	26.8 s	32.1	s 3:	7.5 s	42.8 s	48.2 s	
 Open with legacy UI Record new trace 	0 s + 18.9 s * Process 9	+16.4 ms	+66.4 ms	+116.4 ms	+166.4 ms	+216.4 ms	+266.4 ms	+316.4 ms	+366.4 ms	+416.4 ms	+466.4 ms	+516.4 ms	+566.4 r
Current Trace	Thread 9						.TAU application	I					
Show timeline							.TAU application taupreload_main main Simulation						
← ↔ Query (SQL) ⊘ Metrics	Thread 0	C	_			Spat sem translate	Propagate ilal-space ilag-trans proton				c	Update system V P Star com	m bounda rt comm o M M M
i Info and stats		transfer-stencil-data M MPI_Waitall()	compute-mapping-	M M.	MP M	if-stencil-data-x MPI_Waitall()	Compute-mapping-	M update_remote- M	M MPI_Waitall()	MP	apping-y update		
Convert trace	Thread 1						.TAU application						
▲ Convert to json	Current Selection						.TAU application						<u></u> ↑ ~
Example Traces	Slice Details Name Category Start time Duration	MPL_Waitall() TAU 19s 149ms 472us 9ms 469us				Argumer	nts						

Perfetto.dev Trace Browser



Vampir [TU Dresden] Timeline: Kokkos

the Edit Chard	Elber Window Hale			N Trac	e View - /home/sss	hend/workshop/ko	kkos/kokkos-miniap	ps/lulesh-2.0/kokk	s-minimal-cpu/traces	s.otf2 * - Vampir		
te gat ghan	t Eiter Window Help	No. 14, 155	ten Mail de la	5					N TRATEGICAL CONTRACTOR OF A CONTRACT OF	NTEL O MARKATANA ANA ANA ANA ANA ANA ANA ANA ANA AN	69 156791 s 69 156791 s 69 156886 s d	
1988 EEE 🛤		m m 19		0				THE CHARGE A				ALL AND MALE AND A
	69.15680 s	69.15681 s	69.15682 s	69.15683 s	69.15684 s	69.15685 s	69.15686 s	69.15687 s	69.15.688 s	All Processes, Acc	Function Summary umulated Inclusive Time per Function	
Master thread	_			-						×		
Thread 1	TAU application						TAU application		_	Kokkos::pa[dev	rice=0] [87 µs (2.13%)]	
Thread 2	TAU application						TAU appli	cation		Open0}] [465	μs (11.36%)]	
Thread 3	TAU application		_			-		TAU applicati	in			
Thread 4	TAU application							TAU application	-			
Thread 5	TAU application		_				TAU	application				
Thread 6	TAU application		-	_				TAU application				No.
Thread 7	TAU application				_		TAU app	lication				
Thread 8	TAU application							TAU	application	ř.		
Thread 9	TAU application						TAU app	lication				(2.269 ms (55.4
Thread 10	TAU application		-				TAU	application		11 1 4 4 4 1 4 1 4 1		
Thread 11	TAU application							TAU application		[1.109 ms (27	.1%)]	
Thread 12	TAU application	-		-				TAU a	plication	t l		
Thread 13	TAU application			-			TAU applicati	on		6 - E		
Thread 14	TAU application		-				TAU a	pplication		ř.		
Thread 15	TAU application							TAU applica	tion			
Thread 16	TAU application		_						-			
Thread 17	TAU application						TAU applica	tion	_		Context View	
Thread 18	TAU application	-					TAU application		-	😑 🥌 Function	Summary 🔲	
Thread 19	TAU application	-		-	the second s	-	TAU applicati	on	_	Property	Value	
Thread 20	TAU application				-	_		TAU	application	Display	Function Summary	
	N	all Tree - /home/s	sshand/workshon/k	okkos/kokkos-mini	anos/lulesh-2.0/kok	kos-minimal-cou/	traces off2 * - Vamo	ir		Filter	All Processes	F Idealan Al
		an tree - fridingta	and the second	and a long a line in	abibairaraari kraficar	con the state	or accounter a range		×	Function	Kokkos::parallel_for CalcEnergyForElems	E [device=0]
Processes						All the back share T	teres in the second			TAU_KOKKO	Function Legend	
 TAU applic 	cation					Min Inclusive II	4.527 μs	we time	94.527 µs	Kokkos:	parallel_for CalcKinematicsForElems [device=0] parallel_for CalcMonotonicOGradientsForElems [device	e=01
Kokkos	parallel for CalcEnergyF	orElems E (device	=0)	a lulesh cc caa5di	ic5CalcEnermEor	8	7.000 µs		87.000 µs	Kokkos:	parallel for CalcMonotonicQRegionForElems [device=0	01
	ThenMP IMPLICIT TACK KA	kkosulmeluParalli	ALENT INTERNAL O	Tulach er caasde	S-CalcEnarmEntE	7	2 000 UE		72 000 115	Kokkos:	parallel for CalcPressureForElems A [device=0]	
Callers Calle	es									Kokkos:: Kokkos::	parallel_for CalcPressureForElems B [device=0] parallel_for CalcSoundSpeedForElems [device=0]	
OpenMP_PAR	ALLEL_REGION: Kokkos::Im	pl::ParallelFor<_It	VTERNAL 9 Julesh o	c_caa5d6c5::CalcE	nergyForElems(do	uble*, double*, do	uble*, double*, dou	ble*, double*, dou	ble*, double*, double	Kokkos:	parallel for CalcVelocityForNodes [device=0]	
										Kokkos:	parallel for EvalEOSForElems B [device=0]	
4									. F	Kokkos:	parallel for EvalEOSForElems C [device=0] parallel for EvalEOSForElems D [device=0]	
											The second s	

% export TAU_TRACE=1; export TAU_TRACE_FORMAT=otf2 % tau_exec -T ompt _ompt ./a.out

% vampir traces.otf2 &

Event Based Sampling (EBS)



% mpirun -n 16 tau exec -ebs a.out

ParaProf

Click on Columns: to sort by incl time

000

Open binvcrhs Click on Sample

File Options Windows Help				
Name	Exclusive TIME	Inclusive TIME 🗸	Calls	Child Calls
Ŷ-	9.167	9.36	1	2,432
🖕 🔤 [CONTEXT] . TAU application	0	9.019	901	0
[SUMMARY] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ,	2.89	2.89	288	0
SUMMARY] matmul_sub_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT	1.27	1.27	127	0
SUMMARY] x_solve_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/x	1.16	1.16	116	0
SUMMARY] z_solve_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/z]	1.08	1.08	108	0
SUMMARY] y_solve_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/y	1.08	1.08	108	0
SUMMARY] compute_rhs_ ({/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/B	0.83	0.83	83	0
SUMMARY] matvec_sub_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-	0.49	0.49	49	0
SUMMARY] Ihsinit_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/in	0.08	0.08	8	0
– [SAMPLE] add_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/add.f}	0.05	0.05	5	0
SUMMARY] binvrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/s	0.04	0.04	4	0
SUMMARY] exact_solution_ ({/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/	0.02	0.02	2	0
— [SAMPLE] copy_x_face [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ	0.01	0.01	1	0
SUMMARY] exact_rhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-M;	0.01	0.01	1	0
SAMPLE] initialize_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/in	0.009	0.009	1	0
MPI_Init_thread()	0.155	0.155	1	0
- MPI_Finalize()	0.022	0.022	1	0
► MPI_Waitall()	0.018	0.018	804	0
► MPI_Irecv()	0.004	0.004	804	0
MPI_Isend()	0.001	0.001	804	0
MPI_Comm_split()	0	0	1	0
MPI_Bcast()	0	0	9	0
- MPI_Reduce()	0	0	3	0
MPI_Barrier()	0	0	2	0
- MPI_Comm_size()	0	0	1	0
MPI_Comm_rank()	0	0	2	0

X TAU: ParaProf: Statistics for: node 0 - /rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/bin

% export TAU_SAMPLING=1



.

X TAU: ParaProf: Statistics for: node 0 - /rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/bin

File Options Windows Help

Name	Exclusive TIME	Inclusive TIME 🗸	Calls	Child Calls
- 🗖 .TAU application	9.167	9.368	1	2,432 📤
🖕 🔤 [CONTEXT] .TAU application	0	9.019	901	0
[SUMMARY] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f}]	2.89	2.89	288	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f}_{228}]	0.14	0.14	14	0
— 📕 [SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} Show Sou	urce Code 0.09	0.09	9	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} Show In S	Statistics Table 0.09	0.09	9	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} Show Fur	nction Histogram 0.06	0.06	6	0
[SAMPLE] binvcrhs_ ({/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} Show Fur	nction Bar Chart 0.06	0.06	6	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} Assign Fu	unction Color 0.06	0.06	6	0
SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} Reset to	Default Color 0.06	0.06	6	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {z444}]	0.05	0.05	5	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {332}]	0.05	0.05	5	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {275}]	0.05	0.05	5	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {331}]	0.04	0.04	4	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {445}]	0.04	0.04	4	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {254}]	0.04	0.04	4	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {314}]	0.04	0.04	4	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {343}]	0.04	0.04	4	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {403}]	0.04	0.04	4	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {389}]	0.03	0.03	3	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {415}]	0.03	0.03	3	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {247}]	0.03	0.03	3	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {300}]	0.03	0.03	3	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {309}]	0.03	0.03	3	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {444}]	0.03	0.03	3	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {468}]	0.03	0.03	3	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {242}]	0.03	0.03	3	0
[SAMPLE] binvcrhs_ [{/rwthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f} {407}]	0.03	0.03	3	0
- ISAMPLE1 binvcrhs_[{/nvthfs/rz/cluster/work/hpclab17/NPB3.3-MZ-MPI/BT-MZ/solve_subs.f}_{412}]	0.03	0.03	3	0 /

Callstack Sampling in TAU

TAU: ParaProf: Statistics for: n,c,t 2,0,0 - gamess_unw_call_ebs.ppk		
Name	Inclusive TIME V	Calls
TAU application	79.592	1
▼ MPI_Recv()	75.607	6,870
CONTEXT] MPI_Recv()	74.848	1,497
[UNWIND] /gpfs/mira-home/sameer/gamess-theta-tau/object/unport.f.410 [@] MAIN [{/gpfs/mira-home/sameer/gamess-the	ta-t 26.196	524
UNWIND] /gpfs/mira-home/yuri/dist/Github/gamess-theta-tau/ddi/src/ddi_fortran.c.67 [@] beging_ [{/gpfs/mira-home/samee	r/g: 21.7	434
[UNWIND] /gpfs/mira-home/sameer/gamess-theta-tau/object/gamess.f.538 [@] main [{/gpfs/mira-home/sameer/gamess-theta	-ta 11.85	237
[UNWIND] /gpfs/mira-home/yuri/dist/Github/gamess-theta-tau/ddi/src/ddi_init.c.113 [@] ddi_init_ [{/gpfs/mira-home/yuri/dis	t/Gi 8.701	174
[UNWIND] /gpfs/mira-home/yuri/dist/Github/gamess-theta-tau/ddi/src/ddi_server.c.99 [@] DDI_Init [{/gpfs/mira-home/yuri/di	st/C 5.75	115
[UNWIND] /lib64/libc-2.22.so.0 [@] _start [{/home/abuild/rpmbuild/BUILD/glibc-2.22/csu//sysdeps/x86_64/start.S} {118}]	0.2	4
[SAMPLE] GNII_DlaProgress [{/opt/cray/ugni/6.0.14-6.0.4.0_14.1_ge7db4a2.ari/lib64/libugni.so.0.6.0} {0}]	0.2	4
[UNWIND] [/opt/cray/ugni/6.0.14-6.0.4.0_14.1_ge7db4a2.ari/lib64/libugni.so.0.6.0.0] [@] UNRESOLVED UNKNOWN	0.15	3
[SAMPLE] GNI_CqGetEvent [{/opt/cray/ugni/6.0.14-6.0.4.0_14.1_ge7db4a2.ari/lib64/libugni.so.0.6.0} {0}]	0.051	1
[UNWIND] /opt/cray/pe/mpt/7.6.3/gni/mpich-intel/16.0/lib/libmpich_intel.so.3.0.1.0 [@] MPIDI_CH3I_Progress [{/opt/cray/pe/mpich_intel.so.3.0.1.0 [@] MPIDI_CH3I_Progress [{/opt/cray/pe/mpich_intel.so.3.0.0 [@] MPIDI_CH3I_Progress [{/opt/cray/pe/m	ot/7 0.05	1
MPI_Finalize()	3.601	1
MPI_Send()	0.122	6,866
MPI_Init_thread()	0.112	1
CONTEXT] .TAU application	0.05	1
MPI_Bcast()	0.014	6
MPI_Allgather()	0.004	3
MPI_Barrier()	0.003	7
MPI_Comm_create()	0.002	4
MPI_Gather()	0.002	1
MPI_Comm_split()	0.002	1
MPI_Group_intersection()	0.001	1
MPI_Comm_group()	0.001	1
MPI_Group_incl()	0	3
MPI_Comm_rank()	0	6
MPI_Comm_size()	0	2

% export TAU_SAMPLING=1; export TAU_EBS_UNWIND=1

TAU – Callstack Sampling

Name	Inclusive	Calls ⊽
TAU application	34.979	1
▶ [CONTEXT] .TAU application	31.647	632
void shmem_barrier_all_()	1.219	46,029
[CONTEXT] void shmem_barrier_all_()	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_Open	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	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/v	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 shmem_real8_max_to_all_(void *, void *, int *, int *, int *, int *, void *, long *)	0.325	2,000
[CONTEXT] void shmem_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_to	0.45	9
🔻 🗖 [UNWIND] /notbackedup/tmp/ulib/mpt/nightly/7.2/062215-RC/sma_dmapp/src/shmem_opt_reduction.h.107 [@] _smai_opt_double_ma	0.45	9
[SAMPLE] _smai_smp_reduce_double_max [{/notbackedup/tmp/ulib/mpt/nightly/7.2/062215-RC/sma_dmapp/src/shmem_opt_reduc	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 Context Event Window

	TAU: ParaProf: Context Eve	ents for: node 0, thread 0 -	exafel1_230cores.ppk			
Name 🗠	Total	NumSamples	MaxValue	MinValue	MeanValue	Std. Dev.
<pre><module> [{step5_batch.py}{1}]</module></pre>						
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
cuMemcpyDtoH_v2						
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				
	For all said and		Calle	
Name	Exclusive V	Inclusive	Calls	Child
►	19.845	20.166	303	10,914
Interposed in the second se	16.672	133.715	1	1,066
▼□ <mark>MPI_Bcast()</mark>	12.263	12.263	2	0
[CONTEXT] MPI_Bcast()	0	12.21	407	0
[SAMPLE] PAMI_Context_lock [{/autofs/nccs-svm1_sw/summit/.swci/1-compute/opt/space	3.27	3.27	109	0
[SAMPLE] pthread_spin_lock [{/usr/lib64/libpthread-2.17.so} {0}]	2.34	2.34	78	0
[SAMPLE] start_libcoll_blocking_collective [{/autofs/nccs-svm1_sw/summit/.swci/1-computers/section/	1.89	1.89	63	0
[SAMPLE] PAMI::Device::IBV::Device::advance() [{/autofs/nccs-svm1_sw/summit/.swci/1-cc	1.56	1.56	52	0
[SAMPLE] PAMI_Context_advancev [{/autofs/nccs-svm1_sw/summit/.swci/1-compute/opt	0.69	0.69	23	0
[SAMPLE] UNRESOLVED /usr/lib64/libmlx5.so.1.0.0	0.51	0.51	17	0
SUMMARY] LIBCOLL_Advance_pami [{/SMPI_build_dir/ibmsrc/r	0.42	0.42	14	0
[SAMPLE] LIBCOLL_Advance_pami [{/SMPI_build_dir/ibmsrc/n	0.42	0.42	14	0
[SAMPLE] PAMI_Context_unlock [{/autofs/nccs-svm1_sw/summit/.swci/1-compute/opt/s	0.39	0.39	13	0
[SAMPLE] pthread_spin_unlock [{/usr/lib64/libpthread-2.17.so} {0}]	0.36	0.36	12	0
[SAMPLE]memcpy_power7 [{} {0}]	0.33	0.33	11	0
[SAMPLE] 000003d.plt_call.PAMI_Context_lock [{} {0}]	0.15	0.15	5	0
[SAMPLE] verbs_get_exp_ctx [{pami.cc} {0}]	0.09	0.09	3	0
[SAMPLE] PAMI_Context_trylock_advancev [{/autofs/nccs-svm1_sw/summit/.swci/1-comp	0.06	0.06	2	0
[SAMPLE] 000003d.plt_call.PAMI_Context_unlock [{} {0}]	0.06	0.06	2	0
[SAMPLE] opal_progress [{/autofs/nccs-svm1_sw/summit/.swci/1-compute/opt/spack/20	0.03	0.03	1	0
[SAMPLE] 00000052.plt_call.PAMI_Context_advancev [{} {0}]	0.03	0.03	1	0
SUMMARY] CCMI::Executor::ShmemBroadcastT <false, ccmi::executor::shmematomicbarrie<="" p=""></false,>	0.03	0.03	1	0
[SAMPLE] CCMI::Executor::ShmemBroadcastT < false, CCMI::Executor::ShmemAtomicBarrie	0.03	0.03	1	0
Init [{initpy}{150}]	11.518	15.698	101	1,010
Channel_pixels [{step5_pad.py}{79}]	10.949	106.61	100	13,358
uMemcpyDtoH_v2	9.433	9.433	500	0

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

Deep Learning: Tensorflow

TAU: ParaProf: Statistics for: node 0, thread 8 - nt3_baseline_keras2.ppk		
Name	Inclusiv	Calls ⊽
TAU application	519.211	1
CONTEXT] .TAU application	509.222	50,915
□[SAMPLE] Eigen::internal::gebp_kernel <float, 0="" 0,="" eigen::internal::blas_data_mapper<float,="" float,="" long,="">,</float,>	240.632	24,089
[SAMPLE]pthread_cond_wait [{} {0}]	86.384	8,634
[SAMPLE] Eigen::internal::gemm_pack_rhs <float, eigen::internal::tensorcontractionsubmapper<float,="" long,="" lor<="" p=""></float,>	51.345	5,135
[SAMPLE] Eigen::internal::gemm_pack_rhs <float, eigen::internal::tensorcontractionsubmapper<float,="" long,="" lor<="" p=""></float,>	24.375	2,416
[SAMPLE] void tensorflow::SpatialMaxPoolWithArgMaxHelper <eigen::threadpooldevice, float="">(tensorflow::OpK</eigen::threadpooldevice,>	16.301	1,630
[SAMPLE]memset_sse2 [{} {0}]	13.446	1,336
[SAMPLE] Eigen::TensorEvaluator <eigen::tensorcontractionop<eigen::array<eigen::indexpair<long>, 1ul> co</eigen::tensorcontractionop<eigen::array<eigen::indexpair<long>	5.99	599
[SAMPLE] long Eigen::internal::operator/ <long, false="">(long const&, Eigen::internal::TensorIntDivisor<long, false)<="" p=""></long,></long,>	5.843	585
[SAMPLE] std::_Function_handler <void (long,="" eigen::internal::tensorexecutor<eigen::tensorassignop<i<="" long),="" p=""></void>	5.377	538
[SAMPLE] floatvector Eigen::TensorEvaluator <eigen::tensorbroadcastingop<eigen::indexlist<int, eigen::typ<="" p=""></eigen::tensorbroadcastingop<eigen::indexlist<int,>	4.862	487
[SAMPLE] Eigen::TensorEvaluator <eigen::tensorcontractionop<eigen::array<eigen::indexpair<long>, 1ul> co</eigen::tensorcontractionop<eigen::array<eigen::indexpair<long>	4.775	478
[SAMPLE] Eigen::TensorEvaluator <eigen::tensorassignop<eigen::tensormap<eigen::tensor<float, 1,="" long=""></eigen::tensorassignop<eigen::tensormap<eigen::tensor<float,>	4.037	404
[SAMPLE] Eigen::internal::gemm_pack_lhs <float, eigen::internal::tensorcontractionsubmapper<float,="" lon<="" long,="" p=""></float,>	3.679	367
[SAMPLE] Eigen::internal::EvalRange <eigen::tensorevaluator<eigen::tensorassignop<eigen::tensormap<eigen< p=""></eigen::tensorevaluator<eigen::tensorassignop<eigen::tensormap<eigen<>	2.981	298
[SAMPLE] tensorflow::MaxPoolingOp <eigen::threadpooldevice, float="">::SpatialMaxPool(tensorflow::OpKernelCo</eigen::threadpooldevice,>	2.915	295
[SAMPLE] std::_Function_handler <void (long,="" eigen::internal::tensorexecutor<eigen::tensorassignop<i<="" long),="" p=""></void>	2.91	291
[SAMPLE] std::_Function_handler <void (long,="" eigen::internal::tensorexecutor<eigen::tensorassignop<i<="" long),="" p=""></void>	2.772	277
[SAMPLE] Eigen::internal::gemm_pack_lhs <float, eigen::internal::tensorcontractionsubmapper<float,="" lon<="" long,="" p=""></float,>	2.481	248
[SAMPLE] std::_Function_handler <void (long,="" eigen::internal::tensorexecutor<eigen::tensorassignop<i<="" long),="" p=""></void>	2.148	215
[SAMPLE] void Eigen::internal::call_dense_assignment_loop <eigen::map<eigen::matrix<float, -1,="" -1<="" 0,="" p=""></eigen::map<eigen::matrix<float,>	2.008	197
[SAMPLE] Eigen::NonBlockingThreadPoolTempl <tensorflow::thread::eigenenvironment>::WorkerLoop(int) [{/ho</tensorflow::thread::eigenenvironment>	1.999	200
[SAMPLE] Eigen::internal::ptranspose(Eigen::internal::PacketBlock < floatvector, 4>&) [{crtstuff.c} {0}]	1.919	192
[SAMPLE] Eigen::internal::gemm_pack_rhs <float, eigen::internal::tensorcontractionsubmapper<float,="" long,="" lor<="" p=""></float,>	1.607	160
[SAMPLE] Eigen::TensorEvaluator <eigen::tensorcontractionop<eigen::array<eigen::indexpair<long>, 1ul> co</eigen::tensorcontractionop<eigen::array<eigen::indexpair<long>	1.518	152

% tau_python -ebs nt3_baseline_keras2.py (CANDLE)

TAU – Context Events

00	TAU: ParaProf: Context Events for th	nread: n,c,t, 1,0,0 -	samarc_obe_4p	_iomem_cp.ppl	k			
	Name 🗸		Total	MeanValue	NumSamples Min	Value	MaxValue	Std. Dev.
▼ .TAU application								0
▶ read()								
▶ fopen64()	Mrita hand	width n	ar fila					
▶ fclose()		WIGUNDE						
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()								L
malloc size			3,877	323.083	12	32	981	252.72
free size			1,536	219.429	7	32	464	148.122
▶ fopen64()								
▶ fclose()		- Rvto	e writt	on to	oach filo			
<pre><module> [{obe.py}{8}]</module></pre>		Dyici			cault me	;		
writeRestartData [{samarcInterfact	e.py}{145}]							
samarcWriteRestartData								
▼ write()								
WRITE Bandwidth (MB/s)	<file="samarc <="" nodes.00004="" restore.00002="" td=""><td>/proc.00001"></td><td></td><td>74.565</td><td>117</td><td>0</td><td>2,156.889</td><td>246.386</td></file="samarc>	/proc.00001">		74.565	117	0	2,156.889	246.386
WRITE Bandwidth (MB/s)	<file="samarc <="" nodes.00004="" restore.00001="" td=""><td>/proc.00001"></td><td></td><td>77.594</td><td>117</td><td>0</td><td>1,941.2</td><td>228.366</td></file="samarc>	/proc.00001">		77.594	117	0	1,941.2	228.366
WRITE Bandwidth (MB/s)			\vee	76.08	234	0	2,156.889	237.551
Bytes Written <file="same< td=""><td>arc/restore.00002/nodes.00004/proc.00002</td><td>1"></td><td>2,097,552</td><td>17,927.795</td><td>117</td><td>1</td><td>1,048,576</td><td>133,362.946</td></file="same<>	arc/restore.00002/nodes.00004/proc.00002	1">	2,097,552	17,927.795	117	1	1,048,576	133,362.946
Bytes Written <file="same< td=""><td>arc/restore.00001/nodes.00004/proc.00002</td><td>1"></td><td>2,097,552</td><td>17,927.795</td><td>117</td><td>1</td><td>1,048,576</td><td>133,362.946</td></file="same<>	arc/restore.00001/nodes.00004/proc.00002	1">	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()								

% mpirun –np 16 tau_python -io ./foo.py

ParaProf 3D Profile Browser



TAU – ParaProf 3D Visualization



% paraprof app.ppk Windows -> 3D Visualization -> Bar Plot (right pane)

TAU – 3D Communication Window



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

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

Intel[®] OneAPI

Level Zero

Track time spent in kernels executing on GPU

Track time spent in OneAPI runtime calls

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 + HIP

PMPI + Roctracer interfaces

ROCprofiler+ MPI

ROCm Rocprofiler+ PMPI MPI interface

Kokkos + OpenMP

Kokkos profiling API + OMPT to transparently track events *Kokkos* + *pthread* + *MPI*

Kokkos + pthread wrapper interposition library + PMPI layer *Python* + *ROCTracer* + *MPI*

Python + ROCm Roctracer + MPI profiling interface

MPI + OpenCL

PMPI + OpenCL profiling interfaces

Binary instrumentation of libraries: Work in progress

- % tau_run a.out -o a.inst
 - instruments a binary. Other flags –T <tags>, -f <selective instrumentation file>
- % tau_run -l /path/to/libhdf5.so.310 -o libhdf5.so.310 instruments a DSO
- % tau_exec ./a.out

executes the uninstrumented application with the instrumented shared object.

Works on x86_64. Issues with aarch64:

https://github.com/dyninst/dyninst/issues/1708 and https://github.com/dyninst/dyninst/pull/1712

To use with DyninstAPI 13 on x86_64:

- 1. Load spack: spack/share/spack/setup-env.sh
- 2. Install dyninst: spack install dyninst@13 %gcc@11
- 3. Configure tau with dyninst:
 - 3.1 spack find -p dyninst boost tbb elfutils
 - 3.2 Copy the paths for each package into the configure line
- 3.3./configure -bfd=download -dyninst=<dir> -tbb=<dir> -boost=<dir> -elf=<dir>; <set paths>; make install

Binary instrumentation of libraries: HDF5



\$ pprof

Reading Profile files in profile.*

NODE 0;CONTEXT 0;THREAD 0:

%Time	Exclusive msec	Inclusive total msec	#Call	#Subrs	Inclusive usec/call	Name
100.0	0.272	 68	1	1	68245	.TAU application
99.6	1	67	1	26	67973	taupreload_main
65.8	0.008	44	6	1	7484	H5open
65.8	6	44	2	14	22448	H5_init_library
36.0	4	24	1	12	24563	H5VL_init_phase2
27.8	1	18	1	319	18943	H5T_init
19.8	0.193	13	179	179	76	H5Tregister_int
19.5	0.302	13	179	310	74	H5Tregister
19.0	4	12	155	2555	84	H5Tpath_find_real
13.0	2	8	1	79	8857	H5P_init_phase1
12.7	0.663	8	2	51	4349	H5F_open
11.2	0.348	7	1	6	7610	H5Fcreate
10.5	0.386	7	1	6	7138	H5Fcreate_api_commor
9.8	0.406	6	1	2	6707	H5VL_file_create
9.2	0.005	6	1	1	6299	H5VLnative_file_crea
7.1	1	4	488	976	10	Н5Т_сору
6.5	1	4	1	363	4452	H5E_init
5.6	0.013	3	4	12	956	H5I_dec_app_ref
5.6	0.013	3	2	10	1896	H5Fclose
5.5	0.009	3	2	4	1878	H5Fclose_cb
5.5	0.01	3	2	6	1868	H5VL_file_close
5.4	0.013	3	2	4	1852	H5VLnative_file_clos
5.4	0.019	3	4	8	924	H5F_try_close.localali

TAU – Event Based Sampling (EBS)



% export TAU_SAMPLING=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 [{Impi.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 [{Impi.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 [{Impi.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></file=stdout>	911	62	21	1	14.694	7.441
Bytes Written <file=pipe></file=pipe>	22	22	1	1	1	0
Bytes Written <file=process_output velrsdl.dat=""></file=process_output>	7,826	100	302	76	78.26	22.487
Bytes Written <file=process_output momrsdl.dat=""></file=process_output>	7,826	100	302	76	78.26	22.487
Bytes Written <file=process_output massrsdl.dat=""></file=process_output>	11,325	100	435	110	113.25	32.337
Bytes Written <file=grid_output bodybndry.dat=""></file=grid_output>	9,724	5	8,192	4	1,944.8	3,174.201
Bytes Written <file= <="" apps="" case_catalog="" home="" n="" rotcfd_regression="" sameer="" sukra="" td="" uns2d=""><td>45</td><td>1</td><td>45</td><td>45</td><td>45</td><td>0</td></file=>	45	1	45	45	45	0
Bytes Written <file=. naca0012_largegrid_00010.rst="" restart_history="" restarts=""></file=.>	44,619,720	5,484	8,192	4	8,136.346	640.325
Bytes Written <file=. naca0012_largegrid_00005.rst="" restart_history="" restarts=""></file=.>	44,619,720	5,484	8,192	4	8,136.346	640.325
Bytes Written <file=. naca0012_largegrid.rst="" restarts=""></file=.>	44,619,720	5,484	8,192	4	8,136.346	640.325
Bytes Written <file=. process_output="" turbrsdl.dat=""></file=.>	4,271	. 72	224	57	59.319	19.544
Bytes Written <file=. process_output="" solver.out=""></file=.>	2,039	13	797	43	156.846	191.359
Bytes Written <file=. field_solutions="" naca0012_largegrid_00010.sln="" solution_history=""></file=.>	4,356,976	534	8,192	4	8,159.131	501.319
Bytes Written <file=. field_solutions="" naca0012_largegrid_00005.sln="" solution_history=""></file=.>	4,356,976	534	8,192	4	8,159.131	501.319
Bytes Written <file=. field_solutions="" naca0012_largegrid.sln=""></file=.>	4,356,976	534	8,192	4	8,159.131	501.319
Bytes Written <file=. body_pressure="" naca0012_largegrid_00010_body.prs=""></file=.>	65,986	i 9	8,190	1,300	7,331.778	2,133.204
Bytes Written <file=. body_pressure="" naca0012_largegrid_00005_body.prs=""></file=.>	65,986	9	8,190	1,300	7,331.778	2,133.204
Bytes Written <file=. body_pressure="" frcmnt.out=""></file=.>	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

\varTheta 🔿 🔿 TAU: ParaProf	f: Context Events for	thread: n,c,t, 1,0,0 - s	samarc_obe_4p	_iomem_cp.ppl	k			
Name 🗸			Total	MeanValue	NumSamples	MinValue	MaxValue	Std. Dev.
▼ .TAU application								1
read()								
▶ fopen64()								
▶ fclose()								
OurMain()	Write b	andwidth	per					
malloc size			P 0 .	,097.174	23	11	12,032	2,851.143
free size	file			,746.692	13	11	12,032	3,660.642
OurMain [{wrapper.py}{3}]								
▶ read()								U.
malloc size			3,877	323.083	12	32	981	252.72
free size			1,536	219.429	7	32	464	148.122
▶ fopen64()								
► fclose()								
<pre><module> [{obe.py}{8}]</module></pre>		Dutoo	writtor	a ta aa	ab fila			
writeRestartData [{samarcInterface.py}{145}]		Dytes	writter	110 69				
samarcWriteRestartData								
<pre>vrite()</pre>	\checkmark	,						
WRITE Bandwidth (MB/s) <file="samarc restore.0"<="" td=""><td>0002/nodes.0000</td><td>4/proc.00001"></td><td></td><td>74.565</td><td>117</td><td>0</td><td>2,156.889</td><td>246.386</td></file="samarc>	0002/nodes.0000	4/proc.00001">		74.565	117	0	2,156.889	246.386
WRITE Bandwidth (MB/s) <file="samarc restore.0<="" td=""><td>0001/nodes.0000</td><td>4/proc.00001"></td><td></td><td>77.594</td><td>117</td><td>0</td><td>1,941.2</td><td>228.366</td></file="samarc>	0001/nodes.0000	4/proc.00001">		77.594	117	0	1,941.2	228.366
WRITE Bandwidth (MB/s)			\checkmark	76.08	234	0	2,156.889	237.551
Bytes Written <file="samarc node<="" restore.00002="" td=""><td>s.00004/proc.000</td><td>01"></td><td>2,097,552</td><td>17,927.795</td><td>117</td><td>1</td><td>1,048,576</td><td>133,362.946</td></file="samarc>	s.00004/proc.000	01">	2,097,552	17,927.795	117	1	1,048,576	133,362.946
Bytes Written <file="samarc node<="" restore.00001="" td=""><td>s.00004/proc.000</td><td>01"></td><td>2,097,552</td><td>17,927.795</td><td>117</td><td>1</td><td>1,048,576</td><td>133,362.946</td></file="samarc>	s.00004/proc.000	01">	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()								5

ParaProf Comparison Window



Comparing Rank 0 with 5. Right click on "node 5" -> Add node to comparison window

ParaProf Comparison Window



ParaProf's Topology Display Window



ParaProf Topology Display

Comparison of the second	all/qball_512cores.ppk
	 Triangle Mesh Bar Plot Scatter Plot Topology Plot
	Exclusive TIME event2 Timer O [SAMPLE] fftwi_no_twiddle_15 [{/opt/cray/pe/fftw/2.1.5.9/lib
	Exclusive TIME
58. 239 (30) (30	event3 (Color) Timer O MPI_Allreduce()
10.757	ScatterPlot Axes ColorScale Render
	Auto-Rotate Speed
	Reverse Video Stereo
	AA Lines Full Screen AA
	Separation
	Aperture
	GL Info

ParaProf's Scalable 3D Visualization



786,432 ranks

TAU Hands-On

TAU Exercise #1: Event Based Sampling (EBS)

Using ParaTools Pro for E4S image on AWS with Adaptive Computing's On-Demand Data Center (ODDC)

STEP 1: Go to <u>https://tinyurl.com/e4stut</u>

STEP 2: Reserve an instance and login to: <u>https://paratools.adaptivecomputing.com</u> with the credentials. Firefox recommended.

CoMD: TAU with event-based sampling (EBS)



% cd examples/CoMD/src-mpi % make; cd ../bin

CoMD: TAU with event-based sampling (EBS)


CoMD: TAU's paraprof visualizer



% paraprof &

73

CoMD: TAU's paraprof visualizer

C → C Activities C edu	HPC Cloud On-Demand Data Co X A https://vnc-paratoo	urboVNC: e4s-24-05-aws:1 (to X	+ /vnc.html?resi tutorial10@	ze=remote&path=novnc/websockify?token=6626dd00a6069602bd7559 🏠 Jun 13 02:35 De4s-24-05-aws: ~/examples/CoMD/bin		¥	٢	≻ £ ∢>
Ctivities	A https://vnc-paratoc	Is.adaptivecomputing.com/vnc,	/vnc.html?resi tutorial10@	ze=remote&path=novnc/websockify?token=6626dd00a6069602bd7559 🏠 Jun 13 02:35 De4s-24-05-aws: ~/examples/CoMD/bin		⊻	٢	ඩ ∢ >
ctivities T edu File Optic Applicat C C C	Huoregon-tau-paraprof-ParaProf	TALI: ParaProf Manager	tutorial10@	Jun 13 02:35 De4s-24-05-aws: ~/examples/CoMD/bin	0 =			(پ
File Optic Applicat Contemportation	ions Help	TAU: ParaProf Manager	tutorial10@	De4s-24-05-aws: ~/examples/CoMD/bin				
File Optic	ions Help	TAU: ParaProf Manager						
File Option	ions Help	in an		×				
● Applicat † ☐ Star † ☐ D	iona nep			TAU: ParaProf: /home/tutorial10/examples/CoMD/bin		- (<
	tions ndard Applications Default App ☐ Default Exp ∲ ∯ bin/CoMD/examples/tutorial10/hon ↓ ∯ TIME	TrialField Name Application ID Experiment ID Trial ID CPU Cores CPU MHz CPU Vendor CPUs Allowed CPUs Allowed CPUs Allowed CCPUs Allowed Ending Timestamp Executable File Type Index File T	bin/CoM 0 0 16 3100.2; Intel(R) Genuin 000000 0-1 /home/t 171824 /home/t 1 TAU pro ac-5901 2024-01 ac-5901 000000 0 0 130390 ac-5901	File Options Windows Help Metric: TIME Value: Exclusive Std. Dev. Mean Max Min node 0, thread.0 node 0, thr Show Thread Bar Chart node 1, thr Show Thread Statistics Text Window node 1, thr Show Thread Statistics Table node 2, thr Show Thread Call Graph node 2, thr Show Thread Call Graph node 3, thr Show User Event Bar Chart node 3, thr Show User Event Bar Chart node 3, thr Show User Event Window Show Metadata for Thread Add Thread to Comparison Window				ĩ
		OS Name OS Release OS Version Starting Timestamp	Linux 5.19.0- #30~2 171824	1029-aws 2.04.1-Ubuntu SMP Thu Jul 6022396106				

Right click on Node 0, Thread 0 and choose Show Thread Statistics Table (third option)

TAU's ParaProf Profile Browser: Thread Statistics Table

\rightarrow (A https://vnc-paratools.adaptivecomputing.com/vnc/vnc.html?resize=remote&path=novnc/webs	ockify?token=6626	dd00a6069602bd7559 රූ	\bigtriangledown	⊻ : ⊻ Ξ
tivities	€ edu-uoregon-tau-paraprof-ParaProf Jun 13 02:35				م ە (ئ
	TAU: ParaProf: Statistics for: pode 0, thread 0 - /home/tut	orial 10/examples	/CoMD/bin		X
File	Ontions Windows Help	onder toy examples			
	Name Exclusive	TIME	Inclusive TIME V	Calls	Child Calls
	.TAU application	0	26.582	1	1
9	taupreload_main	26.174	26.582	1	1,255
	e CONTEXT] taupreload_main	0	25.65	855	0
	• [SUMMARY] ljForce [{/home/tutorial10/examples/CoMD/src-mpi/ljForce.c}]	24.27	24.27	809	0
	[SAMPLE] IJForce [{/home/tutorial10/examples/CoMD/src-mpi/IJForce.c} {198}]	4.8	4.8	160	0
	[SAMPLE] IjForce [{/home/tutorial10/examples/CoMD/src-mpi/IjForce.c} {200 show to statistics Table		3.78	126	0
	[SAMPLE] IjForce [{/home/tutorial10/examples/CoMD/src-mpi/IjForce.c} { 199 Show Function Historia	3.66	3.66	122	0
	[SAMPLE] IJForce [{/home/tutorial10/examples/CoMD/src-mpi/IJForce.c} { 189 Show Function Bar Char	t 2.76	2.70	92	0
	[SAMPLE] IjForce [{/home/tutorial10/examples/CoMD/src-mpi/IjForce.c} {202 Assign Function Color	1.98	1.98		0
	[SAMPLE] IjForce [{/home/tutorial10/examples/CoMD/src-mpi/IjForce.c} {193 Reset to Default Color	1.56	1.56	52	0
	[SAMPLE] IjForce [{/home/tutorial10/examples/CoMD/src-mpi/IjForce.c} {208}]	1.47	1.47	49	0
	[SAMPLE] ijForce [{/home/tutorial10/examples/CoMD/src-mpi/ijForce.c} {207}]	1.38	1.38	46	0
	[SAMPLE] IjForce [{/home/tutorial10/examples/CoMD/src-mpi/IjForce.c} {224}]	0.66	0.66	22	0
	[SAMPLE] ij-orce [{/home/tutorial10/examples/CoMD/src-mpi/ij-orce.c} {200}]	0.63	0.63	21	0
	[SAMPLE] IJ-orce [{/home/tutorial10/examples/CoMD/src-mpi/IJ-orce.} {223}]	0.39	0.39	13	0
	[SAMPLE] ij-orce [{/nome/tutorial10/examples/CoMD/src-mp//ji-orce.c] {185}]	0.27	0.27	9	0
2	[SAMPLE] ij-orce [{/nome/tutorial10/examples/CoMD/src-mp//ij-orce.} {210}]	0.24	0.24	8	0
A	[SAMPLE] ij-orce [{/nome/tutorial10/examples/CoMD/src-mp//ji-orce.c} {220}]	0.18	0.18	6	0
	SAMPLEJ IJ-Force [{/nome/tutorial10/examples/CoMD/src-mp//IJ-orce.c} {214}]	0.15	0.15	5	0
	SAMPLEJ IJ-Force [{/nome/tutoriai10/examples/CoMD/src-mp//JForce.c} {181}]	0.12	0.12	4	0
	SAMPLE I Force [{/nome/tutorial10/examples/CoMD/src-mp//jForce.c} {1/5}]	0.09	0.09	3	0
	SAMPLE I Force [{/nome/tutorial10/examples/CoMD/src-mpi/iForce.c} {159}]	0.06	0.06	2	0
	SAMPLE i Force [{/norme/cutorial10/examples/CoMD/stc-mp/ijForce.c} {18/}]	0.08	0.08	2	0
	SAMPLEJ (Force [{/nome/tutoriai10/examples/comb/src-mp/i/Force.c} { 150}]	0.03	0.03	1	0
		0.30	0.30	12	0
		0.21	0.21	7	0
		0.21	0.21	/	0

Click on columns to sort (e.g., Inclusive)

Expand nodes and right click on a sample and

Select "Show Source Code"

TAU's ParaProf Profile Browser: Source Code Browser

•••	49 HPC Cloud On-Demand Data Cr × TurboVNC: e4s-24-05-aws:1 (tr × +					~					
< → C	C A https://vnc-paratools.adaptivecomputing.com/vnc/vnc.html?resize=remot	te&path=novnc/webs	ockify?token=6626	dd00a6069602bd7559 😭	${igsidential}$	⊻ : ≦ =					
Activities	t edu-uoregon-tau-paraprof-ParaProf Ju	in 13 02:36				∢ » ()					
TAL	TAU: ParaProf: Source Browser: /home/tutorial10/examples/CoMD/src-mpi/ljForce.c – 🗆 × ad 0 - /home/tutorial10/examples/CoMD/bin – 🗆 ×										
File	Help										
166		·									
167	int nbrBoxes[2/];	Exclusive	e TIME	Inclusive TIME 🗸	Calls	Child Calls					
168	for (int iBox=0: iBoxes->hoxes->hoxes->hoxes: iBox++)		0	26.582	1	1 -					
169	{		26.174	26.582	1	1,255					
170	<pre>int nIBox = s->boxes->nAtoms[iBox];</pre>		0	25.65	855	0					
- 172	if (nIBox == 0) continue;		24.27	24.27	809	0					
172	<pre>int nNbrBoxes = getNeighborBoxes(s->boxes, iBox, nbrBoxes);</pre>		4.8	4.8	160	0					
174	// loop over neighbors of iBox		3.78	3.78	126	0					
175	for (int jimp=0; jimp <nnbrboxes; jimp++)<="" td=""><td></td><td>3.66</td><td>3.66</td><td>122</td><td>0</td></nnbrboxes;>		3.66	3.66	122	0					
176	int iBox = nbrBoxes[iImp]:		2.76	2.76	92	0_					
177			1.98	1.98	66	0					
178	assert(jBox>=0);		1.56	1.56	52	0					
180			1.47	1.47	49	0					
181	<pre>int nJBox = s->boxes->nAtoms[jBox];</pre>		1.38	1.38	46	0					
182	if (nJBox == 0) continue;		0.66	0.66	22	0					
7 183	//loop over atoms in iBox		0.63	0.63	21	0					
184	for (int iOff=iBox*MAXATOMS.ii=0: ii <nibox: ii++.ioff++)<="" td=""><td></td><td>0.39</td><td>0.39</td><td>13</td><td>0</td></nibox:>		0.39	0.39	13	0					
185	{		0.27	0.27	9	0					
186	<pre>int iId = s->atoms->gid[i0ff];</pre>		0.24	0.24	8	0					
188	// loop over atoms in jBox		0.18	0.18	6	0					
189	<pre>for (int jOff=MAXATOMS*jBox,ij=0; ij<njbox; ij++,joff++)<="" pre=""></njbox;></pre>		0.15	0.15	5	0					
> 190	t dr[3].		0.12	0.12	4	0					
191	int iId = s->atoms->qid[i0ff]:	-	0.09	0.09	3	0					
192	if (jBox < s->boxes->nLocalBoxes && jId <= iId)		0.06	0.06	2	0					
	continue; // don't double count local-local pairs.		0.06	0.06	2	0					
194	real_t r2 = 0.0;		0.03	0.03	1	0					
196	for (int m=0; m<3; m++)	c}	0.36	0.36	12	Ő					
197	drimi - c.satome.srij0fflimi.c.satome.srij0fflimi.	_	0.21	0.21	7	0					
198	r2+=dr[m]*dr[m]:	qe	0.21	0.21	7	0					
199	}	11	0.12	0.12	1						
200	,										
201	if (r2 > rCut2) continue;										
202											

The application spent

4.8 seconds at line 198 in IjForce.c in MPI rank 0. TAU collected 160 samples at this line of code.

It is within five levels of for loops!

There was no change to source code, build system, or the application binary!

TAU Exercise #2: Instrumenting PETSc application using TAU's Perfstubs interface

Launching the binary using tau_exec -ebs



cd ~/examples/petsc-cpu ./compile.sh

qsub tau.qsub qstat After it completes: paraprof &

TAU's ParaProf Profile Browser: Source Code Browser

	HPC Cloud On-Demand Data Ce X	TurboVNC: e4s-24-05-aws:1 (t. × +			~	
$- \rightarrow C$	https://vnc-parato	ols.adaptivecomputing.com/vnc/vnc.h	tml?resize=remote&path=novnc/websockify?token=662	edd00a6069602bd7559 🖒	> ⊻ @ 1 =	
Activities 🗈 Term	inal		Jun 13 02:03		()) ()	
	R		tutorial10@e4s-24-05-aws: ~/examples/petsc-	ри		
	11540000.e4s-2 00:02:00 R tutorial10@e4s	4-05-aws tutoria 00:00:26 -24-05-aws:~/exam	ll0 e4s-24-0 ex50 ples/petsc-cpu\$ qstat -u ;	3795 2 \$USER	4	qsub tau.qsub qstat -u \$USER
2	e4s-24-05-aws: Req'd Job ID Time S	Elap Usernam Time	e Queue Jobname	SessID NDS	Req'd TSK Memory	# After it completes
	11540000.e4s-2 00:02:00 C	4-05-aws tutoria 	l10 e4s-24-0 ex50	3795 2	4	paraprof &
0	tutorial10@e4s clean.sh compile.sh ex50 ex50.c ex50.e11540000 tutorial10@e4s	-24-05-aws:~/exam ex50.oll540000 ex50.gsub ex50.sbatch makefile profile.0.0.0 -24-05-aws:~/exam	<pre>ples/petsc-cpu\$ ls profile.0.0.1 profile.2 profile.0.0.2 profile.2 profile.1.0.0 profile.2 profile.1.0.1 profile.3 profile.1.0.2 profile.3 ples/petsc-cpu\$ paraprof</pre>	.0.0 profile.3.0.2 .0.1 run-single-nod .0.2 run-tau-oddc.s .0.0 .0.1	le.sh ;h	
					Home	

....

TAU's paraprof browser with PETSc performance profile



Using pprof: TAU's text based profile browser

	🗇 🧏 нр	C Cloud On-Demand	Data Co X 🚾 TurboVNC: e4s	s-24-05-aws:1 (tu × +				~	
$\leftarrow \rightarrow$	C	O A ht	tps://vnc-paratools.adaptive	computing.com/vnc/vnc.html?re	size=remote&path=novnc/	/websockifv?token=66	26dd00a6069602bd75	59 ☆ ♡ ½ © ♪ =	
Activities	🕒 Terminal				Jun 13 02:04			Image: Image	
		F		ti	utorial10@e4s-24-05-aws	s: ~/examples/petsc-	-сри	Q = _ = ×	
		tutoria [1] 152 tutoria Reading	110@e4s-24-0 2885 110@e4s-24-0 9 Profile file	pprof –a more					
		NODE 0; %Time	CONTEXT 0;TH Exclusive msec	READ 0: Inclusive total msec	#Call	#Subrs	Inclusive usec/call	Name	Here we see PETSc timers translated into TAU timers using the Perfstubs library.
?		100.0 100.0 65.2 18.1	0.266 597 13,080 1,615	20,080 20,080 13,095 3,629	1 1 1 1 1	1 159 2 284	20080268 20080002 13095901 3629490	.TAU application taupreload_main MPI_Finalize() Main Stage	No modification to the source,
•		13.4 8.7 8.7 8.0	2,700 1 1,512 0	2,700 1,756 1,754 1,607	1 1 1 44	0 27 10258 0	2700776 1756523 1754558 36527	<pre>MPI_Init_thread() PCSetUp MatLUFactor MatLUFactor => [CONTEXT_</pre>	build bystern, or the bindry.
] MatLl 8.0 7.8 Main S	JFactor 0 0 Stage	1,607 1,567	44 12	0 0	36527 130617	[CONTEXT] MatLUFactor Main Stage => [CONTEXT]	
								Home	

Generating Traces

- % cp /tmp/trace.qsub .
- % ./compile.sh
- % qsub trace.qsub
- % qstat

% firefox <u>https://perfetto.dev</u> & Click -> Trace Viewer -> Open the UI -> Open Trace File -> navigate to app.json Open the four trace rows (one for each rank) Use wasd keys to widen/shrink/left/right scroll

What it was doing

cd ~/examples/petsc-cuda; vi ex50.qsub # Comment out previous CALLPATH options export TAU_TRACE=1

- % qsub ex50.qsub
- % tau_treemerge.pl
- % tau_trace2json tau.trc tau.edf -chrome \
 -ignoreatomic -o app.json

Open Firefox, load Perfetto.dev

trace visualizer and open app.json
wasd keys to widen/shrink/left/right

Visualizing Traces with https://Perfetto.dev

● ● ●	~	ᅇ Priv	ate browsing
C A https://vnc-paratools.adaptivecomputing.com/vnc/vnc.html?resize=remote&path=novnc/websockify?token=37eda972003a64808d4c5b0f5fe2b90cd309a5bee5f34988ee11eae3		⊠ ©	മ ≡
Activities 👹 Firefox Feb 27 15:52			•)) (J)
General app.json (16 MB) - Perf × Firefox Privacy Notice - × +	~	×	
← → C O A https://ui.perfetto.dev/#!/viewer?local_cache_key=0000000-0000-017c-93b8d32f69c3 ☆	♥ 🌒	රු ≡	×
AP Perfetto ■ Search or type '>' for commands or ':' for SQL mode		Ŀ	-
	' ₀	00:00:00	
Open trace file 00:00:00 000 0000 000 + 00:00:00 330 000 000 00:00:00 335 000 000 00:00:00 340 000 000 00:00:00 345 000 000 00:00:00 355 000 000 00:00:00 00:00:	0	00:00 365 0	0
O) Record new trace X ≡ Y Default Workspace Y ≡			
Current Trace 027713 C .TAU application taupreload_main			
PD:json (16 MB) Mait Stage MatView MatView MatView	MPI_	Finalize() pthread	
Show timeline MatLUFactor PCApply PCApply MPLAllreduce()			
Lownload MPICo			
Cuery (SQL)			
Viz 12/804 □		.TAU applicat	
Metrics ∧ 0.27713 □			
Info and stats 0 27713 (main thread) taupreload_main			
Convert trace A Convert trace A MatView M MatUFactor PCApply PCApply A A A A A A A A A A A A A A A A A A A		pthrea	
Switch to legacy UI			
🛃 Convert to .json			
1 27804		.TAU applicat	or
Example Traces ^ 127804			01
Open Android example		.TAU applicat	
Den Chrome example		↑ ^	

wasd W = widen S = Shrink A = LeftD = Right

Generating Traces with Vampir

- % cp /tmp/vampir.qsub .
- % ./compile.sh (only needs to be done once)
- % qsub vampir.qsub
- % export PATH=/tmp/vampir/bin:\$PATH
- % vampir traces.otf2 &
 (No need to merge or convert traces)

Vampir [TU Dresden]



```
Generating callpath profiles
 Edit tau.qsub
 # add
 export TAU CALLPATH=1
 export TAU CALLPATH DEPTH=100
 export TAU PROFILE FORMAT=merged
 mpirun ...
 % qsub tau.qsub
 %
  qstat
```

Generating Traces

```
# Edit tau.qsub
# Comment out previous CALLPATH options
export TAU_TRACE=1
export TAU_TRACE_FORMAT=otf2
```

```
% qsub tau.qsub
```

```
% export PATH=/tmp/vampir/bin:$PATH
```

```
% vampir traces.otf2 &
(No need to merge or convert traces)
```

TAU Exercise #5: CUDA instrumentation using CUPTI (CUDA Profiling Tools Interface)

- % cd ~/examples/petsc-cuda
- % cp /tmp/cupti.qsub .
- % qsub cupti.qsub
- % qstat
- After it completes:
- % paraprof tauprofile.xml &

PETSc CUDA: What it was doing

```
% cd ~/examples/petsc-cuda
% ./compile.sh
Edit the qsub file:
spack load tau+cuda
export TAU PROFILE FORMAT=merged
mpiexec tau exec -T cupti, mpi -cupti -ebs ./ksp...
% qsub *.qsub
```

% paraprof tauprofile.xml &

Generating Traces

- % cp /tmp/trace-cuda.qsub .
- % qsub trace-cuda.qsub
- % qstat
- % firefox https://perfetto.dev & Click -> Trace Viewer -> Open the UI -> Open Trace File -> navigate to app.json Open the four trace rows (one for each rank) Use wasd keys to widen/shrink/left/right scroll

Visualizing Traces with https://Perfetto.dev

🔴 🔴 🌒 💰 HPC Cloud On-I	Demand Data C∈ X	🚾 e4s-24-11-aws-2:1 (†	training0) - × 🕂							~	<mark>00</mark> f	Private bro	wsing
$\leftarrow \rightarrow C$	O A https://	vnc-paratools.adapti	vecomputing.com/vnc/	vnc.html?resize=rer	note&path=novnc	:/websockify?to	oken=37eda972003	a64808d4c5b0f5fe	2b90cd309a5bee5f3	84§ ☆		گ	≡
Activities 🛛 😢 Firefox					Feb 27 16	:29						(1)	С
🖻 🚱 bench.json (34 MB)	- P(× +										`	/	×
← → C	O A https:/	//ui.perfetto.dev/a	#!/viewer?local_ca	che_key=00000	000-0000-0000)-a3a3-38ae	9804d3fb			\$		9 Î	=
A Perfetto ≡				Q Search o	or type '>' for cor	mmands or ':'	for SQL mode						
Navigation ^		00:00:00	00:00:01	00:00:01	1 1 1	00:00:02	00:00:02	00:00:	33 1 1 0	1 1 1 1 10:00:03	00:00:04		01
🗁 Open trace file	00:00:00 + 000 000 000		00:00:03 607 000 000	00:00:03 608 000 000	00:00:03 609 000 000	06	0:00:03 10 000 000	00:00:03 611 000 000	00:00:03 612 000 000	00:00:03 613 000 000	00:00:03 614 000 0	00	
O) Record new trace	$x = \nabla$ Default	Workspace 🗸 🗧											
	1 27768												
Current Trace	1 27/68	(main thread)				(PT	TAU appli. HREAD] addr=<0x7ff558eb	cation 1ca0> [{(unknown)} {0, 0}]					
bench.json (34 MB)	0 27713						.TAU appli	cation					
🚃 Show timeline							taupreioad Main St	_main age					
لع Download			MatMult	Ma	tMult		Step2 - Ma MatMult		MatMul	t	Mati	Ault	
曼 Query (SQL)			VecScatterBe Ve	VecScatter	Begin Async	Ve	VecScatterBegin cudaMemonyAsync	Ve	VecScatterBegin	N Ve.	VecScat cudaMem	erBegin covAsvoc	
🟠 Viz				ouddineniopy			ouddineniopjicojno	1	e du la			opyroyne	
 Metrics 	Thread 27716					5	.TAU appli	cation					
 Info and stats 	3 27771						.TAU appli	cation					
			void	void cus	parse::c		void cusparse::cs		void cusparse	e::cs		void cuspar	selic
Convert trace	4 31753												
_	^ 2 27716												
Switch to legacy UI	2 27716	(main thread)					.TAU appli	cation					
🛃 Convert to .json	0 27713						.TAU appli	cation					
							taupreload Main St	_main age					
Example Traces ^			MatMult	Ма	tMult		Step2 - Ma MatMult		MatMu	lt	Mat	Mult	
Doen Android example			VecScatterEnd	VecScatt	VecScatterEnd	VecScatte.	VecScatte	erEnd Vec	Scatte Vec	ScatterEnd	VecScatte	VecScatter	End
Open Chrome example						cudamen.	mri_waita		ameni / MP1		uuamenic	wir'i_walla	ii)
	1 27768						TAII annli	cation					
Support						[PT	HREAD] addr=<0x7f4785cb	1ca0> [{(unknown)} {0, 0}]					
Support	3 27771		void cusparsettes		void cuspa	irse::cs	.TAU appli	cation void cusparse::cs		void cusparserics			
⑦ Keyboard shortcuts	Current Select	ion	tota adaptiracinadia		Tota caspa			tera sasparacinosi				-	↑ ^

wasd W = widen S = Shrink A = LeftD = Right

TAU Exercise #6: paraprof 3D display

TAU paraprof



TAU paraprof 3D visualization



TAU paraprof: 3D Scatter Plot



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 zxf tau.tgz; cd tau-2.<ver>
- wget http://tau.uoregon.edu/ext.tgz ; tar xf ext.tgz
- ./configure -bfd=download -pdt=<dir> -papi=<dir> -pthread -c++=mpicxx -cc=mpicc -fortran=mpif90
 -dwarf=download -unwind=download -otf=download
 -iowrapper -papi=<dir>
- make install

•Using TAU:

- export TAU_MAKEFILE=<taudir>/x86_64/lib/Makefile.tau-<TAGS>
- make CC=tau_cc.sh CXX=tau_cxx.sh F90=tau_f90.sh

Compile-Time Options

•Optional parameters for the TAU_OPTIONS environment variable: % tau_compiler.sh

-optVerbose		Turn on verbose debugging messages
-optCompInst		Use compiler based instrumentation
-optNoCompInst		Do not revert to compiler instrumentation if source instrumentation fails.
-optTrackIO	Wrap POS (Requires	SIX I/O call and calculates vol/bw of I/O operations TAU to be configured with <i>—iowrapper</i>)
-optTrackGOMP		Enable tracking GNU OpenMP runtime layer (used without -opari)
-optMemDbg	Enable rur	ntime bounds checking (see TAU_MEMDBG_* env vars)
-optKeepFiles		Does not remove intermediate .pdb and .inst.* files
-optPreProcess		Preprocess sources (OpenMP, Fortran) before instrumentation
-optTauSelectFile	=" <file>"</file>	Specify selective instrumentation file for tau_instrumentor
-optTauWrapFile=	=" <file>"</file>	Specify path to link_options.tau generated by tau_gen_wrapper
-optHeaderInst		Enable Instrumentation of headers
-optTrackUPCR		Track UPC runtime layer routines (used with tau_upc.sh)
-optLinking=""		Options passed to the linker. Typically \$(TAU_MPI_FLIBS) \$(TAU_LIBS) \$(TAU_CXXLIBS)
-optCompile=""		Options passed to the compiler. Typically \$(TAU_MPI_INCLUDE) \$(TAU_INCLUDE) \$(TAU_DEFS)
-optPdtF95Opts=		Add options for Fortran parser in PDT (f95parse/gfparse)

Compile-Time Options (contd.)

•Optional parameters for the TAU_OPTIONS environment variable: % tau_compiler.sh

-optShared	Use TAU's shared library (libTAU.so) instead of static library (default)					
-optPdtCxxOpts=""	Options for C++ parser in PDT (cxxparse).					
-optPdtF90Parser=""	Specify a different Fortran parser					
-optPdtCleanscapeParser	Specify the Cleanscape Fortran parser instead of GNU gfparser					
-optTau=""	Specify options to the tau_instrumentor					
-optTrackDMAPP Enable instrumentation of low-level DMAPP API calls on Cray						
-optTrackPthread	Enable instrumentation of pthread calls					

See tau_compiler.sh for a full list of TAU_OPTIONS.

. . .

TAU's Runtime Environment Variables

Environment Variable	Default	Description
TAU_TRACE	0	Setting to 1 turns on tracing
TAU_CALLPATH	0	Setting to 1 turns on callpath profiling
TAU_TRACK_MEMORY_FOO TPRINT	0	Setting to 1 turns on tracking memory usage by sampling periodically the resident set size and high water mark of memory usage
TAU_TRACK_POWER	0	Tracks power usage by sampling periodically.
TAU_CALLPATH_DEPTH	2	Specifies depth of callpath. Setting to 0 generates no callpath or routine information, setting to 1 generates flat profile and context events have just parent information (e.g., Heap Entry: foo)
TAU_SAMPLING	1	Setting to 1 enables event-based sampling.
TAU_TRACK_SIGNALS	0	Setting to 1 generate debugging callstack info when a program crashes
TAU_COMM_MATRIX	0	Setting to 1 generates communication matrix display using context events
TAU_THROTTLE	1	Setting to 0 turns off throttling. Throttles instrumentation in lightweight routines that are called frequently
TAU_THROTTLE_NUMCALLS	100000	Specifies the number of calls before testing for throttling
TAU_THROTTLE_PERCALL	10	Specifies value in microseconds. Throttle a routine if it is called over 100000 times and takes less than 10 usec of inclusive time per call
TAU_CALLSITE	0	Setting to 1 enables callsite profiling that shows where an instrumented function was called. Also compatible with tracing.
TAU_PROFILE_FORMAT	Profile	Setting to "merged" generates a single file. "snapshot" generates xml format
TAU_METRICS	TIME	Setting to a comma separated list generates other metrics. (e.g., ENERGY,TIME,P_VIRTUAL_TIME,PAPI_FP_INS,PAPI_NATIVE_ <event>:<subevent>)</subevent></event>

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 –optMemDbg or tau_exec –memory)
TAU_EBS_SOURCE	TIME	Allows using PAPI hardware counters for periodic interrupts for EBS (e.g., TAU_EBS_SOURCE=PAPI_TOT_INS when TAU_SAMPLING=1)
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 TAU_MEMDBG_PROTECT_*)
TAU_MEMDBG_OVERHEAD	0	Specifies the number of bytes for TAU's memory overhead for memory debugging.
TAU_MEMDBG_PROTECT_BELOW/AB OVE	0	Setting to 1 enables tracking runtime bounds checking below or above the array bounds (requires –optMemDbg while building or tau_exec –memory)
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 –optMemDbg or tau_exec – memory)
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

Support Acknowledgements

- 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
- NASA SBIR
- Department of Defense (DoD)
 - PETTT, HPCMP
- National Science Foundation (NSF)
 - SI2-SSI, Glassbox
- CEA, France
- Industry: AWS, AMD, IBM, ARM, Intel, NVIDIA
- Partners:
 - -University of Oregon
 - -The Ohio State University
 - -ParaTools, Inc.
 - -University of Tennessee, Knoxville
 - -T.U. Dresden, GWT
 - -Jülich Supercomputing Center





Acknowledgment

 This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Next-Generation Scientific Software Technologies program, under contract number DE-AC02-06CH11357. DOE SBIR DE-SC0022502.



- https://science.osti.gov/ascr
- <u>https://pesoproject.org</u>
- https://ascr-step.org
- <u>https://hpsf.io</u>
- <u>https://www.energy.gov/technologytransitions/sbirsttr</u>



Thank you

https://www.exascaleproject.org

This research was supported by the Exascale Computing Project (17-SC-20-SC), a joint project of the U.S. Department of Energy's Office of Science and National Nuclear Security Administration, responsible for delivering a capable exascale ecosystem, including software, applications, and hardware technology, to support the nation's exascale computing imperative.



Thank you to all collaborators in the ECP and broader computational science communities. The work discussed in this presentation represents creative contributions of many people who are passionately working toward next-generation computational science.

