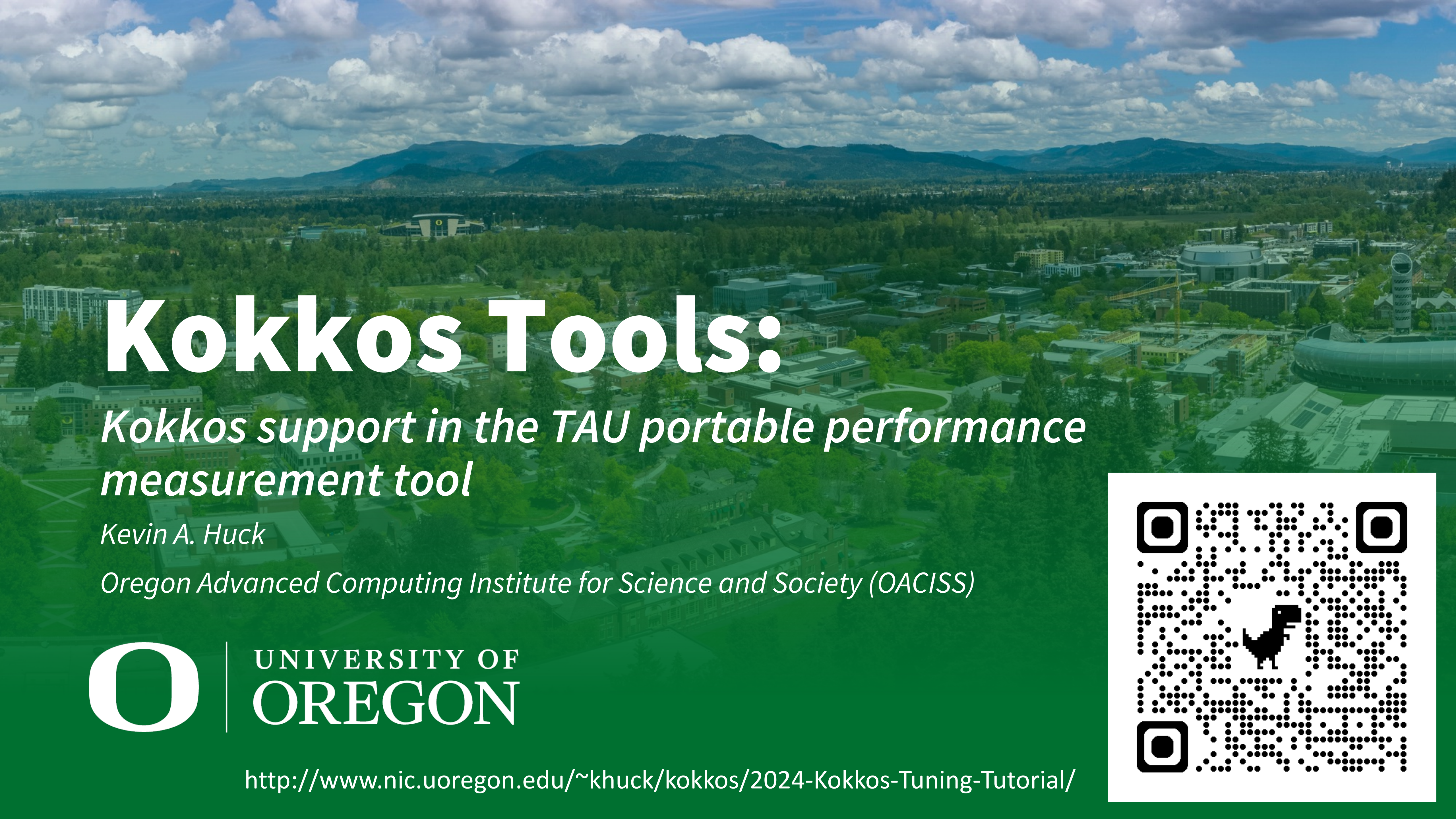


Kokkos Tools:

Kokkos support in the TAU portable performance measurement tool

Kevin A. Huck

Oregon Advanced Computing Institute for Science and Society (OACISS)

UNIVERSITY OF OREGON

http://www.nic.uoregon.edu/~khuck/kokkos/2024-Kokkos-Tuning-Tutorial/

TAU Performance System

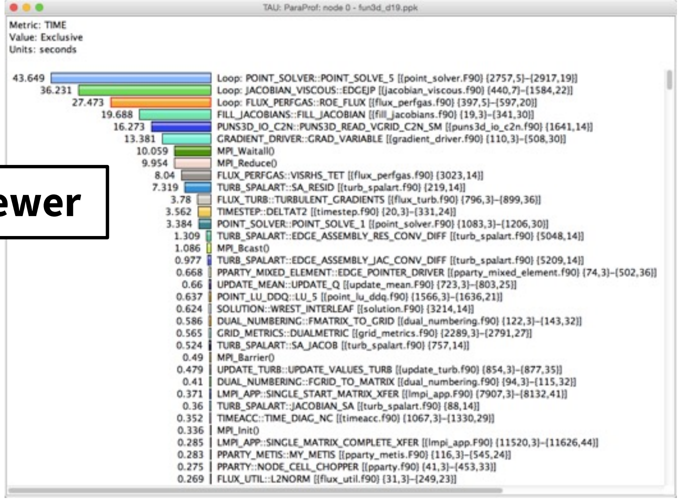
- **T**uning and **A**nalysis **U**tilities (29+ year project)
- Integrated performance toolkit:
  - Multi-level performance instrumentation
  - Highly configurable
  - Widely ported performance profiling / tracing system
  - Portable (java, python) visualization / exploration / analysis tools
- Supports all major HPC programming models
- MPI/SHMEM, OpenMP, OpenACC, CUDA, HIP, SYCL/OneAPI, **Kokkos**...
- Support for ML/AI frameworks: TensorFlow, pyTorch, Horovod
- Integrated with PAPI, LIKWID for hardware counter support
- https://tau.uoregon.edu or https://github.com/UO-OACISS/tau2 (public mirror)

Performance Measurement

- **Timers**
  - Requires instrumentation of some kind
    - Manual, automated
    - Source, compiler provided, binary
    - **Library callbacks**, API wrappers, weak symbol replacement
  - Simple to implement
- **Sampling**
  - Requires specialized system libraries / support
    - Periodic signals, signal handler
    - Call stack unwinding
  - No modification to executable/library needed
  - Potential to interfere with system support (signal handlers)
  - Can mix with timers to generate a hybrid profile

Profiling and Tracing

- **Profiling**: how much time was spent in each measured function on each thread in each process?
  - Collapses the time axis
  - No ordering or causal event information
  - Small summary per thread/process, regardless of execution time – only grows with number of timers & threads/processes
- **Tracing**: record all function entry & exit events on a timeline
  - Detailed view of what happened
  - The longer the program runs, the bigger the trace

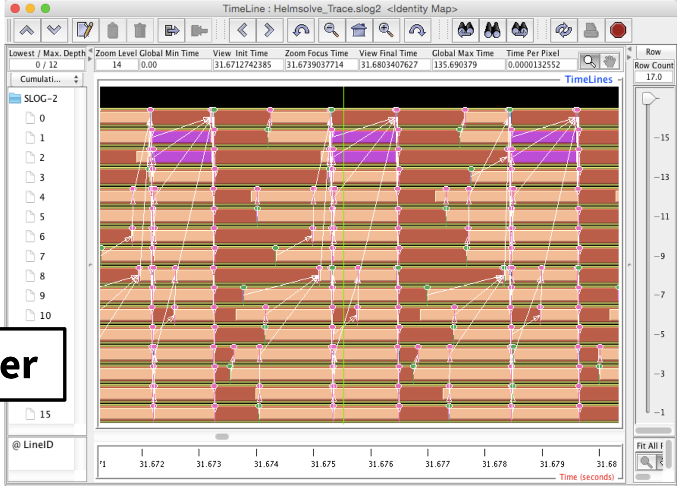


ParaProf profile viewer

TAU: ParaProf: node 0 - fun3d_d19.ppk

```
Metric: TIME
Value: Exclusive
Units: seconds

43.649  Loop: POINT_SOLVER::POINT_SOLVE_5 [{point_solver.F90} {2757,5}-{2917,19}]
36.231  Loop: JACOBIAN_VISCOUS::EDGEJP [{jacobian_viscous.f90} {440,7}-{1584,22}]
27.473  Loop: FLUX_PERFGAS::ROE_FLUX [{flux_perfgas.f90} {397,5}-{597,20}]
19.688  FILL_JACOBIANS::FILL_JACOBIAN [{fill_jacobians.f90} {19,3}-{341,30}]
16.273  PUNS3D_IO_C2N::PUNS3D_READ_VGRID_C2N_SM [{puns3d_io_c2n.f90} {1641,14}]
13.381  GRADIENT_DRIVER::GRAD_VARIABLE [{gradient_driver.f90} {110,3}-{508,30}]
10.059  MPI_Waitall()
 9.954  MPI_Reduce()
  8.04  FLUX_PERFGAS::VISRHS_TET [{flux_perfgas.f90} {3023,14}]
 7.319  TURB_SPALART::SA_RESID [{turb_spalart.f90} {219,14}]
  3.78  FLUX_TURB::TURBULENT_GRADIENTS [{flux_turb.f90} {796,3}-{899,36}]
 3.562  TIMESTEP::DELTAT2 [{timestep.f90} {20,3}-{331,24}]
 3.384  POINT_SOLVER::POINT_SOLVE_1 [{point_solver.F90} {1083,3}-{1206,30}]
 1.309  TURB_SPALART::EDGE_ASSEMBLY_RES_CONV_DIFF [{turb_spalart.f90} {5048,14}]
 1.086  MPI_Bcast()
 0.977  TURB_SPALART::EDGE_ASSEMBLY_JAC_CONV_DIFF [{turb_spalart.f90} {5209,14}]
 0.668  PPARTY_MIXED_ELEMENT::EDGE_POINTER_DRIVER [{pparty_mixed_element.f90} {74,3}-{502,36}]
  0.66  UPDATE_MEAN::UPDATE_Q [{update_mean.F90} {723,3}-{803,25}]
 0.637  POINT_LU_DDQ::LU_5 [{point_lu_ddq.f90} {1566,3}-{1636,21}]
 0.624  SOLUTION::WREST_INTERLEAF [{solution.F90} {3214,14}]
 0.586  DUAL_NUMBERING::FMATRIX_TO_GRID [{dual_numbering.f90} {122,3}-{143,32}]
 0.565  GRID_METRICS::DUALMETRIC [{grid_metrics.f90} {2289,3}-{2791,27}]
 0.524  TURB_SPALART::SA_JACOB [{turb_spalart.f90} {757,14}]
  0.49  MPI_Barrier()
 0.479  UPDATE_TURB::UPDATE_VALUES_TURB [{update_turb.f90} {854,3}-{877,35}]
  0.41  DUAL_NUMBERING::FGRID_TO_MATRIX [{dual_numbering.f90} {94,3}-{115,32}]
 0.371  LMPI_APP::SINGLE_START_MATRIX_XFER [{lmpi_app.F90} {7907,3}-{8132,41}]
  0.36  TURB_SPALART::JACOBIAN_SA [{turb_spalart.f90} {88,14}]
 0.352  TIMEACC::TIME_DIAG_NC [{timeacc.f90} {1067,3}-{1330,29}]
 0.336  MPI_Init()
 0.285  LMPI_APP::SINGLE_MATRIX_COMPLETE_XFER [{lmpi_app.F90} {11520,3}-{11626,44}]
 0.283  PPARTY_METIS::MY_METIS [{pparty_metis.F90} {116,3}-{545,24}]
 0.275  PPARTY::NODE_CELL_CHOPPER [{pparty.f90} {41,3}-{453,33}]
 0.269  FLUX_UTIL::L2NORM [{flux_util.f90} {31,3}-{249,23}]
```

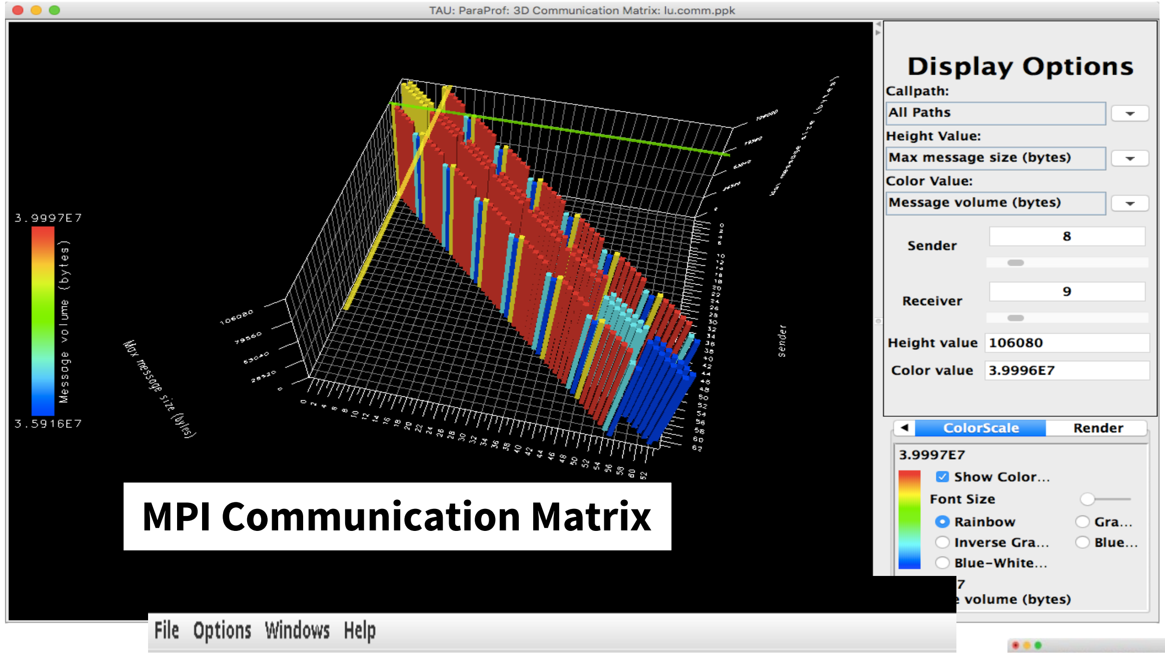


JumpShot trace viewer

TimeLine : Helmsolve_Trace.slog2 <Identity Map>

Lowest / Max. Depth	Zoom Level	Global Min Time	View Init Time	Zoom Focus Time	View Final Time	Global Max Time	Time Per Pixel	Row Count
0 / 12	14	0.00	31.6712742385	31.6739037714	31.6803407627	135.690379	0.0000132552	17.0

TAU Analysis Tools: ParaProf, Vampir



Display Options

Callpath: All Paths

Height Value: Max message size (bytes)

Color Value: Message volume (bytes)

Sender: 8

Receiver: 9

Height value: 106080

Color value: 3.9996E7

ColorScale | Render

3.9997E7

☑ Show Color...

Font Size

● Rainbow ○ Gra...

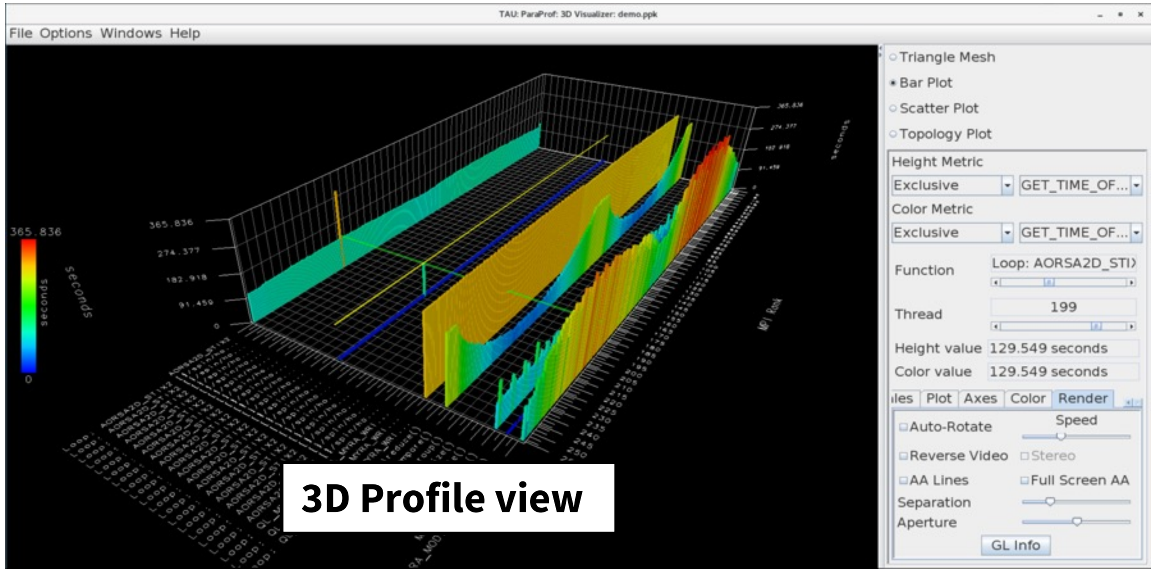
○ Inverse Gra... ○ Blue...

○ Blue-White...

...7

...volume (bytes)

MPI Communication Matrix


TAU: ParaProf: 3D Visualizer: demo.ppk
File Options Windows Help
365.836
274.377
182.918
91.459
0
seconds
MPI Rank

○ Triangle Mesh

● Bar Plot

○ Scatter Plot

○ Topology Plot

Height Metric: Exclusive | GET_TIME_OF...

Color Metric: Exclusive | GET_TIME_OF...

Function: Loop: AORSA2D_STI>

Thread: 199

Height value: 129.549 seconds

Color value: 129.549 seconds

...les | Plot | Axes | Color | Render

□ Auto-Rotate | Speed

□ Reverse Video □ Stereo

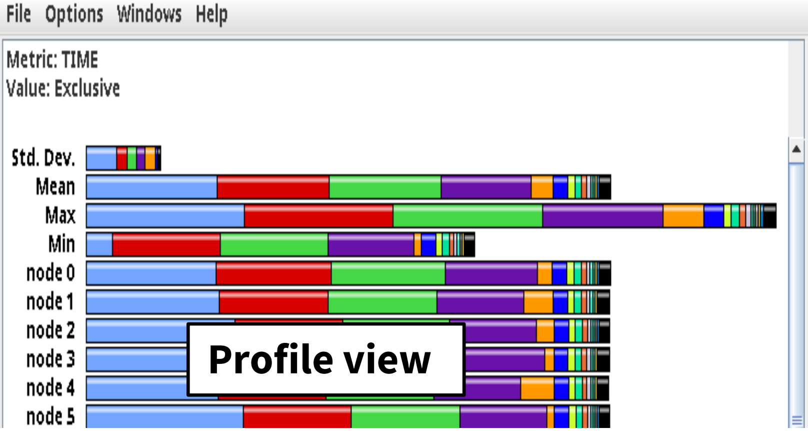
□ AA Lines □ Full Screen AA

Separation

Aperture

GL Info

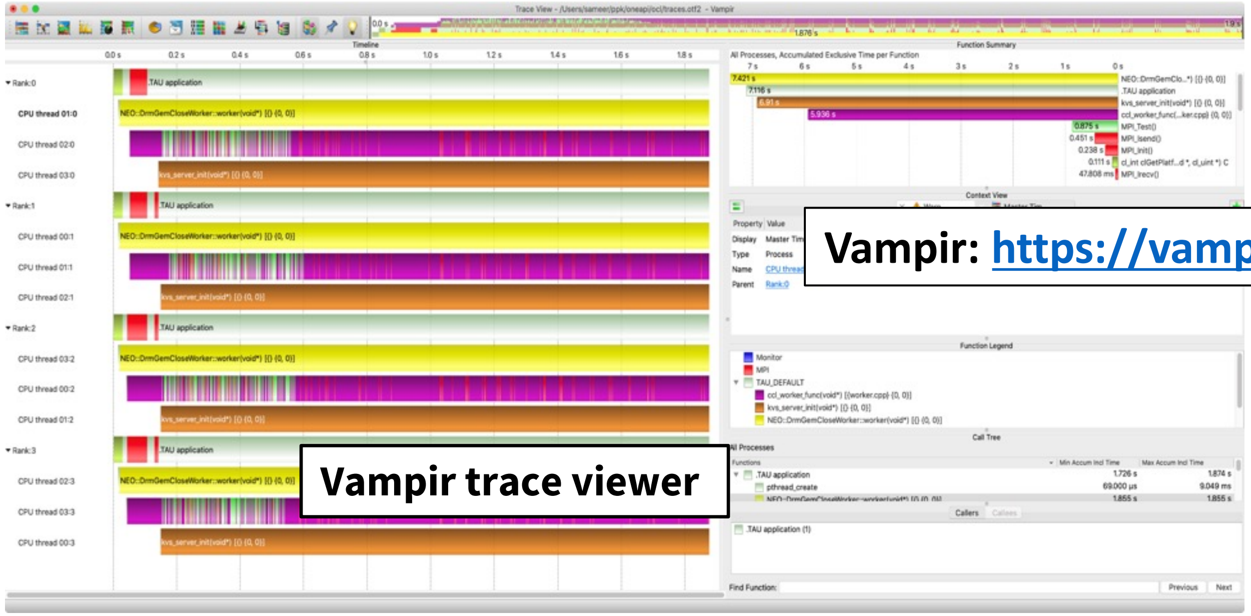
3D Profile view



Metric: TIME

Value: Exclusive

Profile view



Vampir: https://vampir.eu

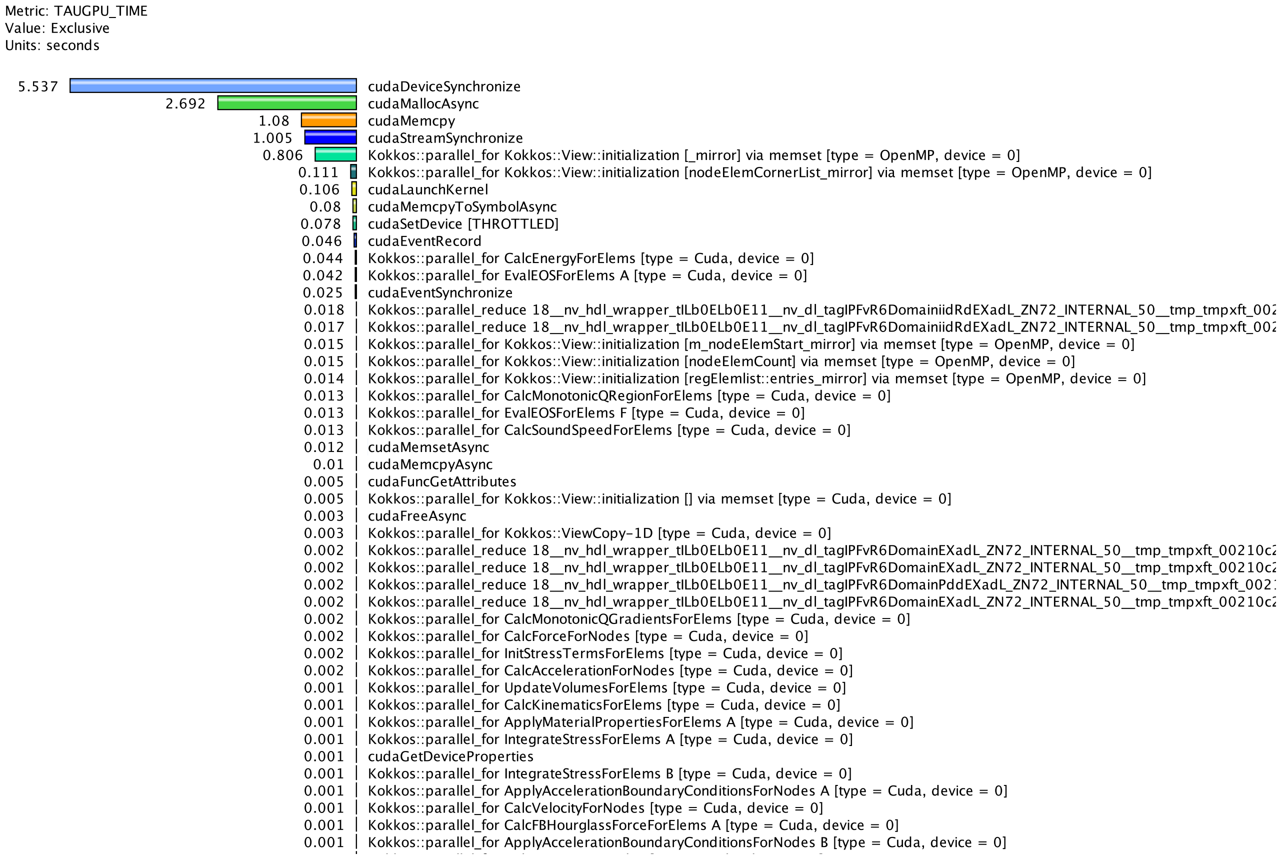
Vampir trace viewer

UNIVERSITY OF OREGON

Kokkos support in TAU – since February, 2017

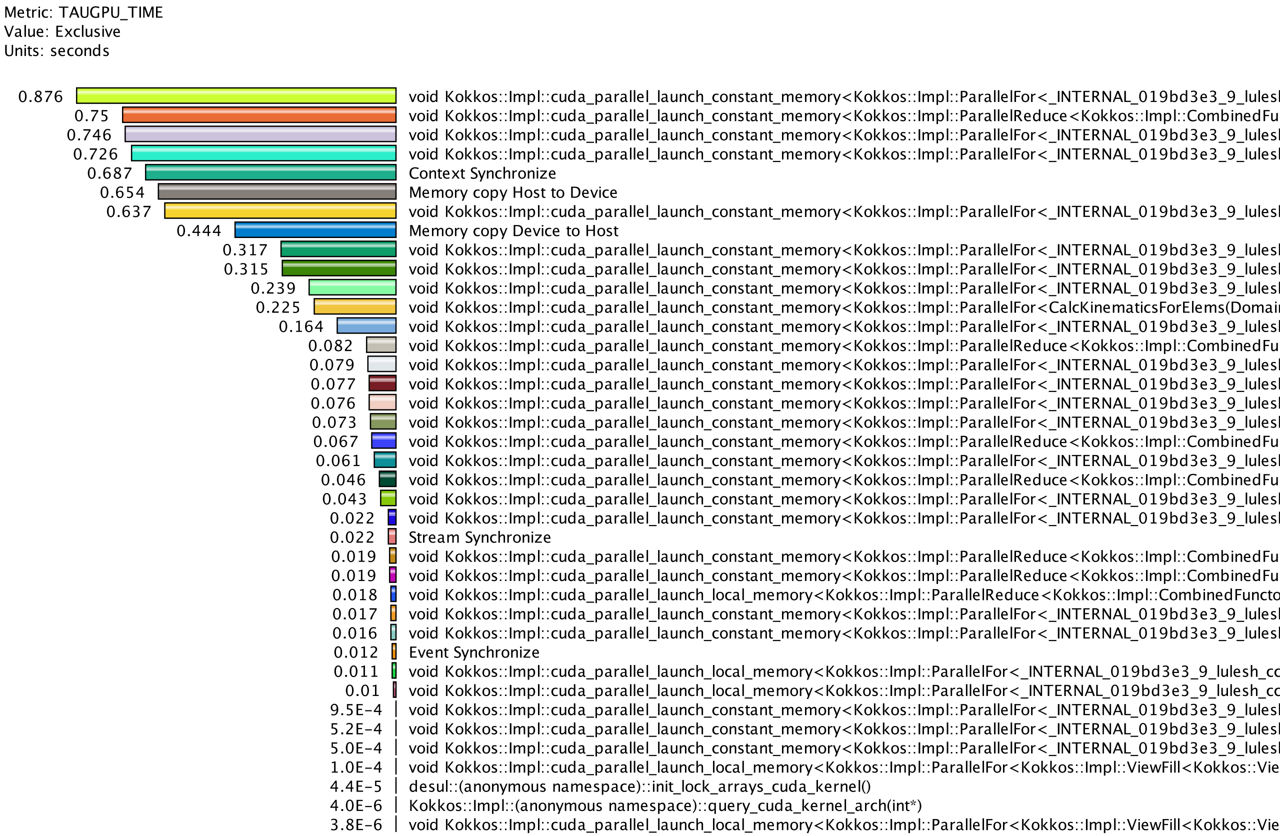
- TAU implements the Kokkos Profiling API (`Kokkos_Profiling_C_Interface.h`)
- TAU sets an environment variable KOKKOS_TOOLS_LIBS to tell Kokkos that it should enable profiling and enable function callbacks to the TAU implementations
- TAU implements
  - `kokkosp_[init|finalize]_library`
  - `kokkosp_[begin|end]_parallel_[for|scan|reduce]`
  - `kokkosp_[push|pop]_profile_region`
- Names for regions are passed to the tools to provide intelligent labels
- In addition, TAU also implements support for native Pthreads, OpenMP, OpenACC, CUDA, HIP, SYCL back-end measurement – no code changes necessary
- Fun fact: if you have a Raja application, and Raja is configured with `-DRAJA_ENABLE_RUNTIME_PLUGINS`, Raja implements the same callback API!

TAU Example – Kokkos Lulesh (from kokkos-miniapps)



```
Metric: TAUGPU_TIME
Value: Exclusive
Units: seconds

5.537   cudaDeviceSynchronize
2.692   cudaMallocAsync
1.08    cudaMemcpy
1.005   cudaStreamSynchronize
0.806   Kokkos::parallel_for Kokkos::View::initialization [_mirror] via memset [type = OpenMP, device = 0]
0.111   Kokkos::parallel_for Kokkos::View::initialization [nodeElemCornerList_mirror] via memset [type = OpenMP, device = 0]
0.106   cudaLaunchKernel
0.08    cudaMemcpyToSymbolAsync
0.078   cudaSetDevice [THROTTLED]
0.046   cudaEventRecord
0.044   Kokkos::parallel_for CalcEnergyForElems [type = Cuda, device = 0]
0.042   Kokkos::parallel_for EvalEOSForElems A [type = Cuda, device = 0]
0.025   cudaEventSynchronize
0.018   Kokkos::parallel_reduce 18__nv_hdl_wrapper_tILb0ELb0E11__nv_dl_tagIPFvR6DomainiidRdEXadL_ZN72_INTERNAL_50__tmp_tmpxft_002
0.017   Kokkos::parallel_reduce 18__nv_hdl_wrapper_tILb0ELb0E11__nv_dl_tagIPFvR6DomainiidRdEXadL_ZN72_INTERNAL_50__tmp_tmpxft_002
0.015   Kokkos::parallel_for Kokkos::View::initialization [m_nodeElemStart_mirror] via memset [type = OpenMP, device = 0]
0.015   Kokkos::parallel_for Kokkos::View::initialization [nodeElemCount] via memset [type = OpenMP, device = 0]
0.014   Kokkos::parallel_for Kokkos::View::initialization [regElemlist::entries_mirror] via memset [type = OpenMP, device = 0]
0.013   Kokkos::parallel_for CalcMonotonicQRegionForElems [type = Cuda, device = 0]
0.013   Kokkos::parallel_for EvalEOSForElems F [type = Cuda, device = 0]
0.013   Kokkos::parallel_for CalcSoundSpeedForElems [type = Cuda, device = 0]
0.012   cudaMemsetAsync
0.01    cudaMemcpyAsync
0.005   cudaFuncGetAttributes
0.005   Kokkos::parallel_for Kokkos::View::initialization [] via memset [type = Cuda, device = 0]
0.003   cudaFreeAsync
0.003   Kokkos::parallel_for Kokkos::ViewCopy-1D [type = Cuda, device = 0]
0.002   Kokkos::parallel_reduce 18__nv_hdl_wrapper_tILb0ELb0E11__nv_dl_tagIPFvR6DomainEXadL_ZN72_INTERNAL_50__tmp_tmpxft_00210c2
0.002   Kokkos::parallel_reduce 18__nv_hdl_wrapper_tILb0ELb0E11__nv_dl_tagIPFvR6DomainEXadL_ZN72_INTERNAL_50__tmp_tmpxft_00210c2
0.002   Kokkos::parallel_reduce 18__nv_hdl_wrapper_tILb0ELb0E11__nv_dl_tagIPFvR6DomainPddEXadL_ZN72_INTERNAL_50__tmp_tmpxft_0021
0.002   Kokkos::parallel_reduce 18__nv_hdl_wrapper_tILb0ELb0E11__nv_dl_tagIPFvR6DomainEXadL_ZN72_INTERNAL_50__tmp_tmpxft_00210c2
0.002   Kokkos::parallel_for CalcMonotonicQGradientsForElems [type = Cuda, device = 0]
0.002   Kokkos::parallel_for CalcForceForNodes [type = Cuda, device = 0]
0.002   Kokkos::parallel_for InitStressTermsForElems [type = Cuda, device = 0]
0.002   Kokkos::parallel_for CalcAccelerationForNodes [type = Cuda, device = 0]
0.001   Kokkos::parallel_for UpdateVolumesForElems [type = Cuda, device = 0]
0.001   Kokkos::parallel_for CalcKinematicsForElems [type = Cuda, device = 0]
0.001   Kokkos::parallel_for ApplyMaterialPropertiesForElems A [type = Cuda, device = 0]
0.001   Kokkos::parallel_for IntegrateStressForElems A [type = Cuda, device = 0]
0.001   cudaGetDeviceProperties
0.001   Kokkos::parallel_for IntegrateStressForElems B [type = Cuda, device = 0]
0.001   Kokkos::parallel_for ApplyAccelerationBoundaryConditionsForNodes A [type = Cuda, device = 0]
0.001   Kokkos::parallel_for CalcVelocityForNodes [type = Cuda, device = 0]
0.001   Kokkos::parallel_for CalcFBHourglassForceForElems A [type = Cuda, device = 0]
0.001   Kokkos::parallel_for ApplyAccelerationBoundaryConditionsForNodes B [type = Cuda, device = 0]
```



```
Metric: TAUGPU_TIME
Value: Exclusive
Units: seconds

0.876   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.75    void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelReduce<Kokkos::Impl::CombinedFu
0.746   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.726   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.687   Context Synchronize
0.654   Memory copy Host to Device
0.637   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.444   Memory copy Device to Host
0.317   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.315   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.239   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.225   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<CalcKinematicsForElems(Domai
0.164   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.082   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelReduce<Kokkos::Impl::CombinedFu
0.079   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.077   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.076   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.073   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.067   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelReduce<Kokkos::Impl::CombinedFu
0.061   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.046   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelReduce<Kokkos::Impl::CombinedFu
0.043   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.022   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.022   Stream Synchronize
0.019   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelReduce<Kokkos::Impl::CombinedFu
0.019   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelReduce<Kokkos::Impl::CombinedFu
0.018   void Kokkos::Impl::cuda_parallel_launch_local_memory<Kokkos::Impl::ParallelReduce<Kokkos::Impl::CombinedFuncto
0.017   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.016   void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
0.012   Event Synchronize
0.011   void Kokkos::Impl::cuda_parallel_launch_local_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesh_cc
0.01    void Kokkos::Impl::cuda_parallel_launch_local_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesh_cc
9.5E-4  void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
5.2E-4  void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
5.0E-4  void Kokkos::Impl::cuda_parallel_launch_constant_memory<Kokkos::Impl::ParallelFor<_INTERNAL_019bd3e3_9_lulesl
1.0E-4  void Kokkos::Impl::cuda_parallel_launch_local_memory<Kokkos::Impl::ParallelFor<Kokkos::Impl::ViewFill<Kokkos::Vie
4.4E-5  desul::(anonymous namespace)::init_lock_arrays_cuda_kernel()
4.0E-6  Kokkos::Impl::(anonymous namespace)::query_cuda_kernel_arch(int*)
3.8E-6  void Kokkos::Impl::cuda_parallel_launch_local_memory<Kokkos::Impl::ParallelFor<Kokkos::Impl::ViewFill<Kokkos::Vie
```

Main thread launching kernels

Virtual thread with CUDA activity

UNIVERSITY OF OREGON

Intel Xeon system with NVIDIA A100, size 256, 100 iterations, one rank

PerfStubs side note…

- PerfStubs is a “frictionless” instrumentation library
  - https://github.com/UO-OACISS/perfstubs
  - One source file, three headers
  - Provides a plugin interface for performance tools
  - Can be compiled away if desired
- Integrated into several libraries (so far) as a git submodule
  - CAMTIMERS
  - PETSc
  - Ginkgo
  - ADIOS2
  - Others?
- Provides runtime integration with TAU & APEX

Boehme, Huck, Madsen, Weidendorfer,
“The Case for a Common Instrumentation Interface for HPC Codes”
https://doi.org/10.1109/ProTools49597.2019.00010, 2019

Acknowledgements

Parts of 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.

This research used resources of the Oak Ridge Leadership Computing Facility at the Oak Ridge National Laboratory, which is supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

U.S. DEPARTMENT OF ENERGY | Office of Science

RAPIDS

ECP EXASCALE COMPUTING PROJECT

Thanks! Questions?