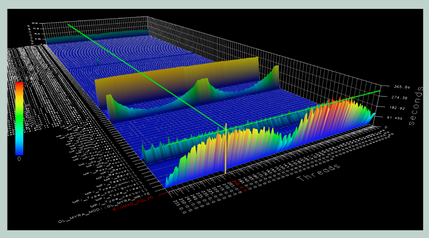
Performance Measurement and Analysis using the TAU Performance System

Allen D. Malony, Sameer Shende, Kevin Huck, Jeffrey S. Vetter, Joel Denny, Jungwon Kim, Seyong Lee, Kei Davis, Hal Finkel

Funded through ECP 2.3.2.09 (PROTEAS) and ECP 2.2.6.08 (CODAR)

What is TAU?



- Tuning and Analysis Utilities (24+ year project)
- http://tau.uoregon.edu
- Comprehensive performance profiling and tracing
  - ✓ Integrated, scalable, flexible, portable
  - ✓ Targets all parallel programming/execution paradigms
- Integrated performance toolkit
  - ✓ Instrumentation, measurement, analysis, visualization
  - ✓ Timers, samples, counters, integrated tool callback support, hardware counter support
  - ✓ Widely-ported performance profiling and tracing system
  - ✓ Performance data management and data mining
  - ✓ Open source (BSD-style license)

PROTEAS Project Goals

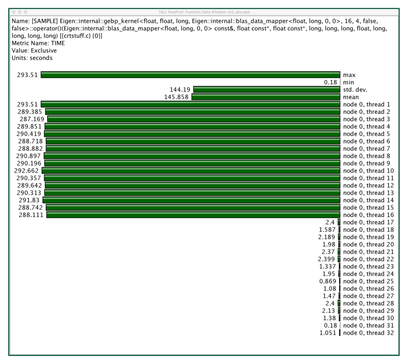
(PROgramming Toolchain for Emerging Architectures and Systems)

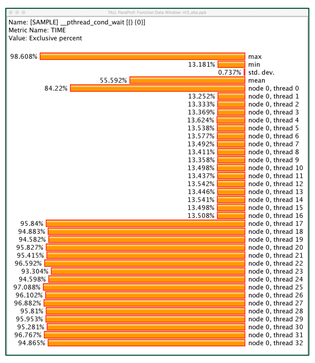
Programmer productivity and performance portability are two of the most important challenges facing applications targeting future exascale computing platforms. The PROTEAS project is a strategic response to the continuous changes in architectures and hardware (e.g., heterogeneous computing, deep memory hierarchies, nonvolatile memory) that are defining the landscape for emerging ECP systems. PROTEAS is a flexible programming framework and integrated toolchain that will provide ECP applications the opportunity to work with programming abstractions and to evaluate solutions that address the exascale programming challenges they face.

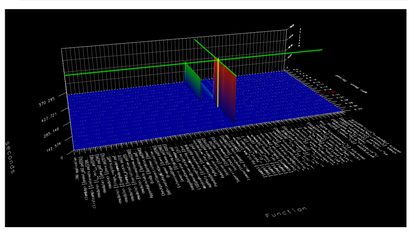
Key Capabilities: LLVM; **OpenACC**, **CUDA**, **OpenCL**; Performance tools with **TAU**; Expertise and software systems for **heterogeneous computing** (GPUs, FPGAs, Manycore) and **deep memory hierarchies** including **nonvolatile memory**; **Performance portability** metrics, tools, and strategies.

New Capabilities in TAU from PROTEAS

- **LLVM Integration** : TAU LLVM plugin was built to provide selective instrumentation at the routine level for CLANG compilers
- **OpenACC Updates** : Tracking data transfers between host and GPU with variable names
- **OpenMP Updates** : OpenMP Tools Interface TR6 support based on library replacement integrated in TAU, with finer resolution and configurable measurements
- **Kokkos** : TAU was updated to support Kokkos profiling interface with source code information about kernels
- **Sample Resolution** : Configurable sample resolution to simplify profile measurement at the line, routine, and file level
- **ROCm** : Initial TAU support for OpenCL and ROCm released in TAU v2.28
- **CUDA 10** : Added support in TAU for CUDA 10 events, multithreaded kernel launch – see figure

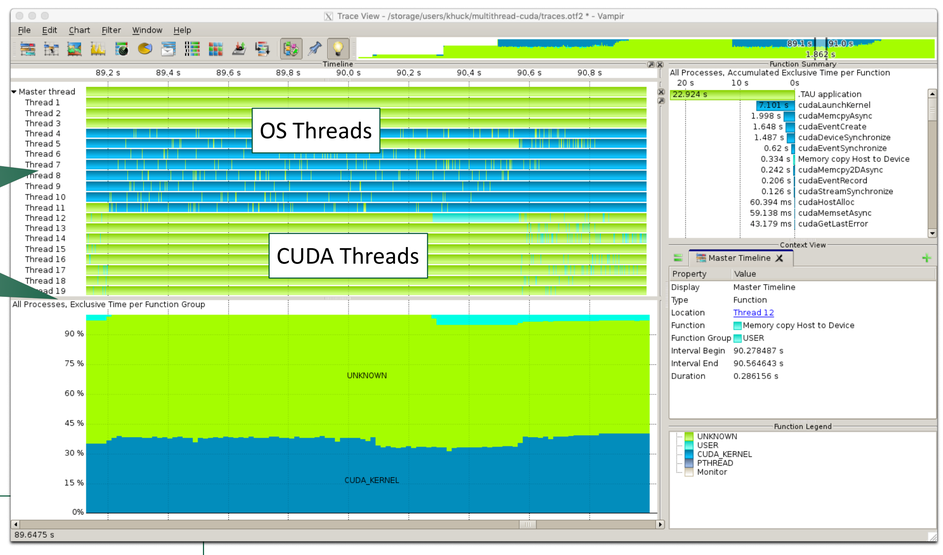






Figures: CANDLE application measurement, showing bimodal performance distribution

TAU Tutorial:
Thursday, Jan 17
1:30 PM



Above figure: Multithreaded CUDA application traced with TAU, visualized in Vampir – both CPU and GPU activity in one timeline

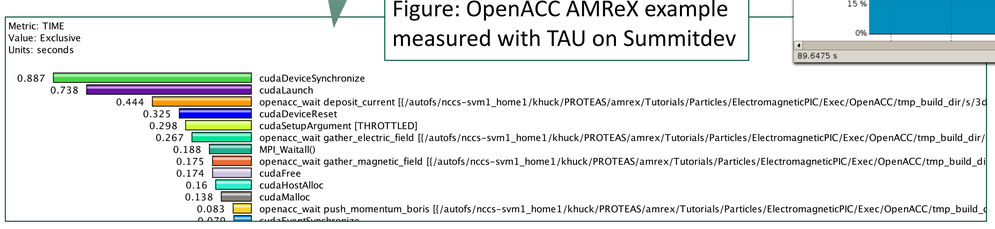


Figure: OpenACC AMReX example measured with TAU on Summitdev

ECP Collaborators

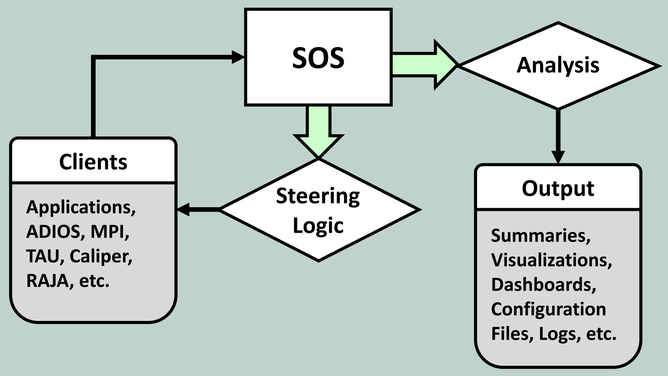
- Co-Design Center for Online Data Analysis and Reduction at the Exascale (**CODAR**)
- ADIOS Framework for Scientific Data on Exascale Systems (**ADIOS**)
- Exascale Deep Learning and Simulation Enabled Precision Medicine for Cancer (**CANDLE**) – see figures
- Tackling Chemical, Materials and Biomolecular Challenges in the Exascale Era (**NWChemEx**)
- Block-Structured AMR Co-Design Center (**AMReX**)
- ECP Applications Effective use of Kokkos to Achieve Performance Portability (**Kokkos**)
- Others...

TAU / PROTEAS: Next Steps

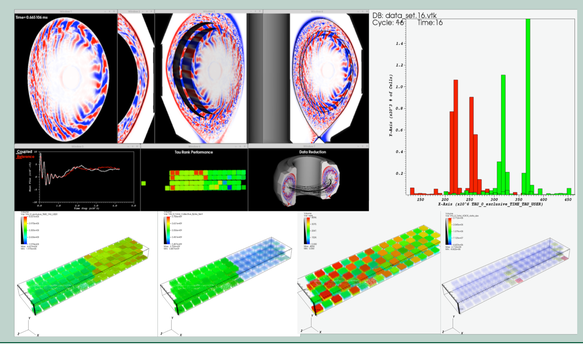
- **ROCm** : Integration with rocProfiler and rocTracer for access to PAPI counters for AMD GPUs
- **NVMe support in TAU** : LIKWID and PAPI support for memory hierarchy counters
- **Improved LLVM integration**
  - Support for F18/Flang compilers for compiler-based instrumentation
  - Improved selective instrumentation using TAU LLVM plugin

Related Project: Scalable Observation System

Runtime data aggregation for monitoring and control



Figures: Coupled fusion application measured by TAU and monitored by SOS with VTK visualization data



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