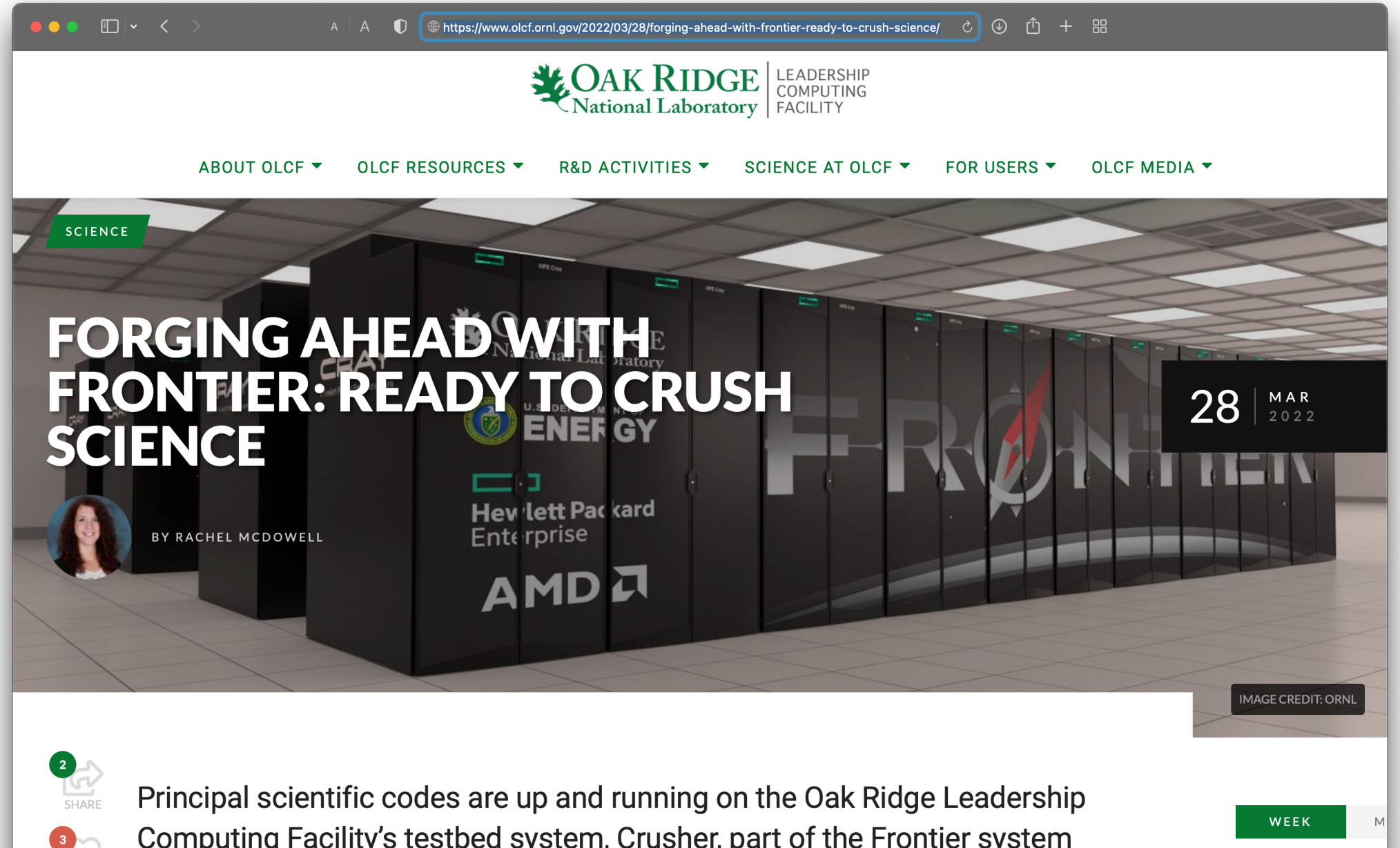
Hewlett Packard
Enterprise

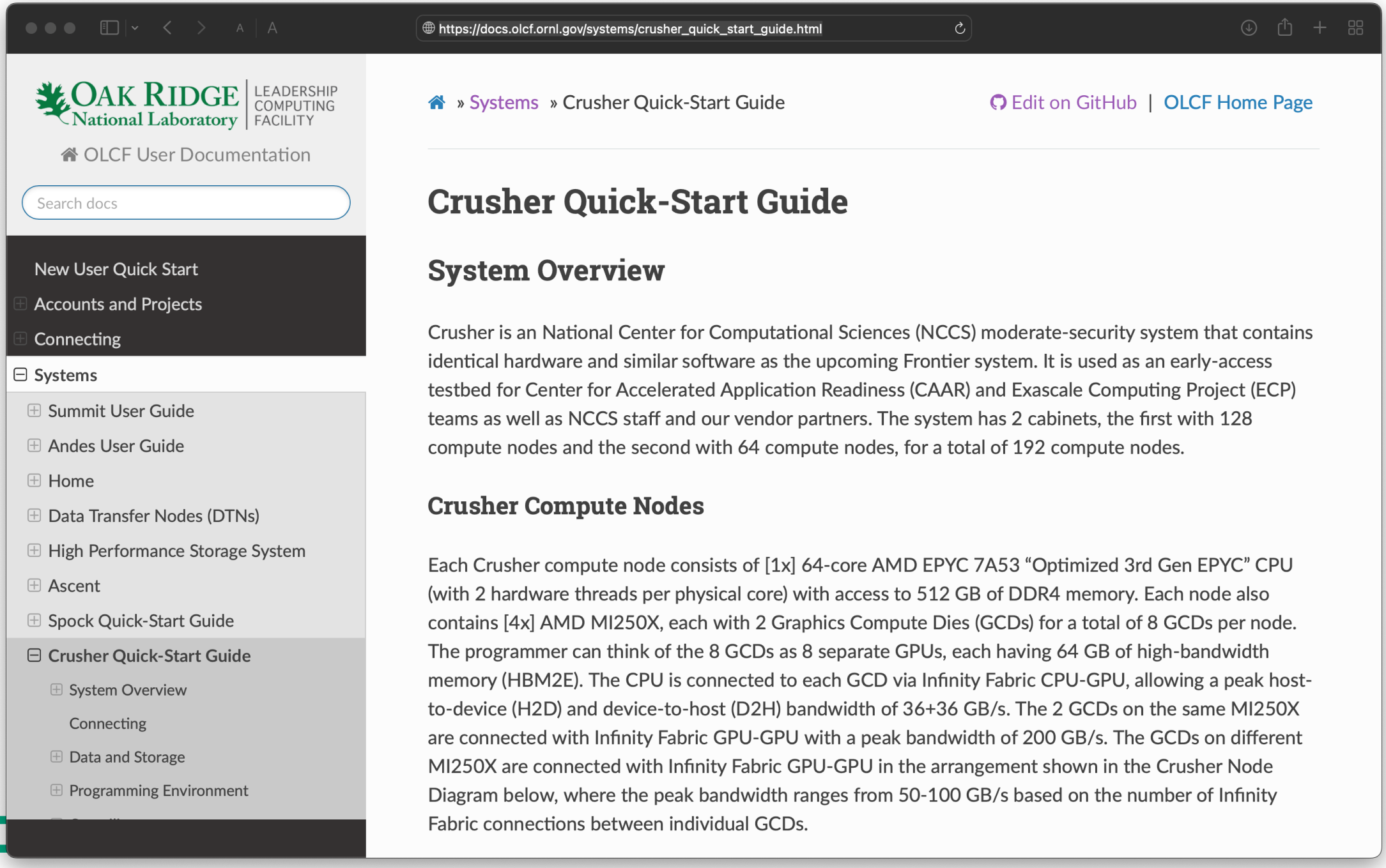
PART 2: PERFORMANCE PROFILING ON HPE CRAY SUPERCOMPUTERS WITH AMD GPUS

Trey White

May 2, 2022

THANKS TO OLCF FOR USE OF CRUSHER!




https://docs.olcf.ornl.gov/systems/crusher_quick_start_guide.html
OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY
OLCF User Documentation
Search docs
New User Quick Start
Accounts and Projects
Connecting
Systems
Summit User Guide
Andes User Guide
Home
Data Transfer Nodes (DTNs)
High Performance Storage System
Ascent
Spock Quick-Start Guide
Crusher Quick-Start Guide
System Overview
Connecting
Data and Storage
Programming Environment
» Systems » Crusher Quick-Start Guide
Edit on GitHub | OLCF Home Page
Crusher Quick-Start Guide
System Overview
Crusher is an National Center for Computational Sciences (NCCS) moderate-security system that contains identical hardware and similar software as the upcoming Frontier system. It is used as an early-access testbed for Center for Accelerated Application Readiness (CAAR) and Exascale Computing Project (ECP) teams as well as NCCS staff and our vendor partners. The system has 2 cabinets, the first with 128 compute nodes and the second with 64 compute nodes, for a total of 192 compute nodes.
Crusher Compute Nodes
Each Crusher compute node consists of [1x] 64-core AMD EPYC 7A53 “Optimized 3rd Gen EPYC” CPU (with 2 hardware threads per physical core) with access to 512 GB of DDR4 memory. Each node also contains [4x] AMD MI250X, each with 2 Graphics Compute Dies (GCDs) for a total of 8 GCDs per node. The programmer can think of the 8 GCDs as 8 separate GPUs, each having 64 GB of high-bandwidth memory (HBM2E). The CPU is connected to each GCD via Infinity Fabric CPU-GPU, allowing a peak host-to-device (H2D) and device-to-host (D2H) bandwidth of 36+36 GB/s. The 2 GCDs on the same MI250X are connected with Infinity Fabric GPU-GPU with a peak bandwidth of 200 GB/s. The GCDs on different MI250X are connected with Infinity Fabric GPU-GPU in the arrangement shown in the Crusher Node Diagram below, where the peak bandwidth ranges from 50-100 GB/s based on the number of Infinity Fabric connections between individual GCDs.

PERFTOOLS-LITE-GPU

or *An Epic Wrap Battle*

PRGENV-AMD BUILD

```
module load PrgEnv-amd
module load craype-accel-amd-gfx90a
module load rocm
export CXX='CC -x hip'
export CXXFLAGS='-ggdb -O3 -std=c++17 -Wall'
export LD='CC'
export LDFLAGS="${CXXFLAGS} -L${ROCM_PATH}/lib"
export LIBS='-lamdhip64'
make clean
make
```

PRGENV-AMD BUILD

use CCE wrappers with AMD Clang compilers
get HPE Cray MPI automatically
get latest Hip optimizations from AMD

```
module load PrgEnv-amd
module load craype-accel-amd-gfx90a
module load rocm
export CXX='CC -x hip'
export CXXFLAGS='-ggdb -O3 -std=c++17 -Wall'
export LD='CC'
export LDFLAGS="${CXXFLAGS} -L${ROCM_PATH}/lib"
export LIBS='-lamdhip64'
make clean
make
```

PRGENV-AMD BUILD

automatically target AMD MI250X accelerators (gfx90a)

```
module load PrgEnv-amd
module load craype-accel-amd-gfx90a
module load rocm
export CXX='CC -x hip'
export CXXFLAGS='-ggdb -O3 -std=c++17 -Wall'
export LD='CC'
export LDFLAGS="${CXXFLAGS} -L${ROCM_PATH}/lib"
export LIBS='-lamdhip64'
make clean
make
```

PRGENV-AMD BUILD

```
module load PrgEnv-amd
module load craype-accel-amd-gfx90a
module load rocm                                    use Hip
export CXX='CC -x hip'
export CXXFLAGS='-ggdb -O3 -std=c++17 -Wall'
export LD='CC'
export LDFLAGS="${CXXFLAGS} -L${ROCM_PATH}/lib"
export LIBS='-lamdhip64'
make clean
make
```

PRGENV-AMD RUN

```
module load PrgEnv-amd
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
```

PRGENV-AMD RUN

```
module load PrgEnv-amd
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
```

match build environment

PRGENV-AMD RUN

```
module load PrgEnv-amd
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
```

use accelerator memory for MPI buffers

PRGENV-AMD RUN

```
module load PrgEnv-amd
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
```

spread 64 MPI tasks
across 4 host numa domains
and across 8 accelerators
and across 4 network interfaces

PRGENV-AMD BUILD WITH PERFTOOLS-LITE-GPU

```
module load PrgEnv-amd
module load craype-accel-amd-gfx90a
module load rocm
module load perftools-lite-gpu
export PATH="${PATH}:${ROCM_PATH}/llvm/bin"
export CXX='CC -x hip'
export CXXFLAGS='-ggdb -O3 -std=c++17 -Wall'
export LD='CC'
export LDFLAGS="${CXXFLAGS} -L${ROCM_PATH}/lib"
export LIBS='-lamdhip64'
make clean
make
```

(path change is temporary workaround)

PRGENV-AMD BUILD WITH PERFTOOLS-LITE-GPU

```
module load PrgEnv-amd
module load craype-accel-amd-gfx90a
module load rocm
module load perftools-lite-gpu
export PATH="${PATH}:${ROCM_PATH}/llvm/bin"
export CXX='CC -x hip'
export CXXFLAGS='-ggdb -O3 -std=c++17 -Wall'
export LD='CC'
export LDFLAGS="${CXXFLAGS} -L${ROCM_PATH}/lib"
export LIBS='-lamdhip64'
make clean
make
```

new line in build log

```
INFO: creating the PerfTools-instrumented executable 'faces' (lite-gpu) ...OK
```

PRGENV-AMD RUN WITH PERFTOOLS-LITE-GPU

```
module load PrgEnv-amd
module load craype-accel-amd-gfx90a
module load rocm
module load perftools-lite-gpu
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
```

match build environment

PRGENV-AMD RUN WITH PERFTOOLS-LITE-GPU

New output!

```
...
0: time 2.818 avg 2.80146 min 2.82917 max
0:
0: ########################################################################
0: #                                                                      #
0: #            CrayPat-lite Performance Statistics                       #
0: #                                                                      #
0: ########################################################################
0:
0: CrayPat/X:  Version 21.12.0 Revision 543286d4e  11/23/21 01:35:38
0: Experiment:                        lite  lite-gpu
0: Number of PEs (MPI ranks):           64
0: Numbers of PEs per Node:              8  PEs on each of  8  Nodes
0: Numbers of Threads per PE:            1
0: Number of Cores per Socket:          64
0: Execution start time:  Thu Apr 14 00:04:38 2022
0: System name and speed:  crusher001  2.734 GHz (nominal)
0: AMD   Trento                    CPU  Family: 25  Model: 48  Stepping:  1
0: Core Performance Boost:  All 64 PEs have CPB capability
0:
...
```

HIPCC BUILD

```
module load craype-accel-amd-gfx90a
module load rocm
export CXX='hipcc'
export CXXFLAGS="-ggdb -O3 -std=c++17 -Wall \
  --offload-arch=gfx90a -I${CRAY_MPICH_DIR}/include"
export LD='hipcc'
export LDFLAGS="${CXXFLAGS} -L${CRAY_MPICH_DIR}/lib \
  ${PE_MPICH_GTL_DIR_amd_gfx90a}"
export LIBS="-lmpi ${PE_MPICH_GTL_LIBS_amd_gfx90a}"
make clean
make
```

HIPCC BUILD

target AMD MI250X (gfx90a)

```
module load craype-accel-amd-gfx90a
module load rocm
export CXX='hipcc'
export CXXFLAGS="-ggdb -O3 -std=c++17 -Wall \
  --offload-arch=gfx90a -I${CRAY_MPICH_DIR}/include"
export LD='hipcc'
export LDFLAGS="${CXXFLAGS} -L${CRAY_MPICH_DIR}/lib \
  ${PE_MPICH_GTL_DIR_amd_gfx90a}"
export LIBS="-lmpi ${PE_MPICH_GTL_LIBS_amd_gfx90a}"
make clean
make
```

HIPCC BUILD

```
module load craype-accel-amd-gfx90a
module load rocm                                    use HPE Cray MPI
export CXX='hipcc'
export CXXFLAGS="-ggdb -O3 -std=c++17 -Wall \
  --offload-arch=gfx90a -I${CRAY_MPICH_DIR}/include"
export LD='hipcc'
export LDFLAGS="${CXXFLAGS} -L${CRAY_MPICH_DIR}/lib \
  ${PE_MPICH_GTL_DIR_amd_gfx90a}"
export LIBS="-lmpi ${PE_MPICH_GTL_LIBS_amd_gfx90a}"
make clean
make
```

HIPCC RUN

```
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
```

no need for PrgEnv-amd
otherwise same

HIPCC BUILD WITH PERFTOOLS-LITE-GPU

as expected

```
module load perftools-lite-gpu
module load craype-accel-amd-gfx90a
module load rocm
export CXX='hipcc'
export CXXFLAGS="$(pat_opts include hipcc gpu) \
  $(pat_opts pre_compile hipcc gpu) -g -O3 -std=c++17 -Wall  \
  --offload-arch=gfx90a -I${CRAY_MPICH_DIR}/include \
  $(pat_opts post_compile hipcc gpu)"
export LD='hipcc'
export LDFLAGS="$(pat_opts pre_link hipcc gpu) ${CXXFLAGS} \
  -L${CRAY_MPICH_DIR}/lib ${PE_MPICH_GTL_DIR_amd_gfx908}"
export LIBS="-lmpi ${PE_MPICH_GTL_LIBS_amd_gfx908} \
  $(pat_opts post_link hipcc gpu)"
make clean
make
```

HIPCC BUILD WITH PERFTOOLS-LITE-GPU

```
module load perftools-lite-gpu
module load craype-accel-amd-gfx90a
module load rocm
export CXX='hipcc'
export CXXFLAGS="$(pat_opts include hipcc gpu) \
  $(pat_opts pre_compile hipcc gpu) -g -O3 -std=c++17 -Wall  \
  --offload-arch=gfx90a -I${CRAY_MPICH_DIR}/include \
  $(pat_opts post_compile hipcc gpu)"
export LD='hipcc'
export LDFLAGS="$(pat_opts pre_link hipcc gpu) ${CXXFLAGS} \
  -L${CRAY_MPICH_DIR}/lib ${PE_MPICH_GTL_DIR_amd_gfx908}"
export LIBS="-lmpi ${PE_MPICH_GTL_LIBS_amd_gfx908} \
  $(pat_opts post_link hipcc gpu)"
make clean
make
```

use pat_opts *to add compile and link arguments*

build phase *compiler* *"lite" experiment*

HIPCC BUILD WITH PERFTOOLS-LITE-GPU

```
module load perftools-lite-gpu
module load craype-accel-amd-gfx90a
module load rocm
export CXX='hipcc'
export CXXFLAGS="$(pat_opts include hipcc gpu) \
  $(pat_opts pre_compile hipcc gpu) -g -O3 -std=c++17 -Wall  \
  --offload-arch=gfx90a -I${CRAY_MPICH_DIR}/include \
  $(pat_opts post_compile hipcc gpu)"
export LD='hipcc'
export LDFLAGS="$(pat_opts pre_link hipcc gpu) ${CXXFLAGS} \
  -L${CRAY_MPICH_DIR}/lib ${PE_MPICH_GTL_DIR_amd_gfx908}"
export LIBS="-lmpi ${PE_MPICH_GTL_LIBS_amd_gfx908} \
  $(pat_opts post_link hipcc gpu)"
make clean
make
```

new line in build log

```
INFO: creating the PerfTools-instrumented executable 'faces' (lite-gpu) ...OK
```

HIPCC RUN WITH PERFTOOLS-LITE-GPU

```
module load perftools-lite-gpu
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
```

HIPCC RUN WITH PERFTOOLS-LITE-GPU

New output!

```
...
0: time 2.92938 avg 2.91265 min 2.93668 max
0:
0: ########################################################################
0: #                                                                      #
0: #            CrayPat-lite Performance Statistics                       #
0: #                                                                      #
0: ########################################################################
0:
0: CrayPat/X:  Version 21.12.0 Revision 543286d4e  11/23/21 01:35:38
0: Experiment:                    lite  lite-gpu
0: Number of PEs (MPI ranks):       64
0: Numbers of PEs per Node:          8  PEs on each of  8  Nodes
0: Numbers of Threads per PE:        1
0: Number of Cores per Socket:      64
0: Execution start time:  Fri Apr 15 14:39:27 2022
0: System name and speed:  crusher112  2.690 GHz (nominal)
0: AMD   Trento                  CPU  Family: 25  Model: 48  Stepping:  1
0: Core Performance Boost:  All 64 PEs have CPB capability
0:
...
```

PERFTOOLS-LITE-GPU OUTPUT

or *Everything But the Kitchen Sink*

HEADER

```
0: ########################################################################
0: #                                                                      #
0: #            CrayPat-lite Performance Statistics                       #
0: #                                                                      #
0: ########################################################################
0:
0: CrayPat/X:  Version 21.12.0 Revision 543286d4e  11/23/21 01:35:38
0: Experiment:                         lite  lite-gpu
0: Number of PEs (MPI ranks):            64
0: Numbers of PEs per Node:               8  PEs on each of  8  Nodes
0: Numbers of Threads per PE:             1
0: Number of Cores per Socket:           64
0: Execution start time:  Fri Apr 15 19:18:07 2022
0: System name and speed:  crusher001  2.685 GHz (nominal)
0: AMD   Trento                    CPU  Family: 25  Model: 48  Stepping:  1
0: Core Performance Boost:  All 64 PEs have CPB capability
```

I/O AND MEMORY STATS

```
0: Avg Process Time:          4.27 secs
0: High Memory:           20,086.0 MiBytes     313.8 MiBytes per PE
0: I/O Read Rate:     629.641824 MiBytes/sec
0: I/O Write Rate:    394.818
0: 571 MiBytes/sec
```

```
0: Table 1:  Profile by Function Group and Function
0:
0:   Time% |     Time |     Imb. |  Imb. |     Calls | Group
0:         |          |     Time | Time% |           |  Function=[MAX10]
0:         |          |          |       |           |   PE=HIDE
0:
0:  100.0% | 4.142693 |       -- |    -- | 848,867.0 | Total
0: |-----------------------------------------------------------------------
0: |  53.5% | 2.215179 |       -- |    -- | 291,668.0 | HIP
0: ||----------------------------------------------------------------------
0: ||  22.7% | 0.941439 | 0.473985 | 34.0% |  20,000.0 | hipStreamSynchronize
0: ||  19.1% | 0.792874 | 0.130271 | 14.3% |     141.0 | hipMemset
0: ||   4.1% | 0.169448 | 0.007623 |  4.4% |      20.0 | hipStreamCreate
0: ||   1.4% | 0.056639 | 0.008608 | 13.4% |  60,200.0 | __hipPushCallConfiguration
0: ||======================================================================
0: |  33.5% | 1.386777 |       -- |    -- | 545,708.0 | MPI
0: ||----------------------------------------------------------------------
0: ||  24.5% | 1.015371 | 0.242284 | 19.6% |  20,200.0 | MPI_Waitall
0: ||   7.0% | 0.289145 | 0.199618 | 41.5% | 262,600.0 | MPI_Isend
0: ||   1.9% | 0.080567 | 0.075127 | 49.0% | 262,600.0 | MPI_Irecv
0: ||======================================================================
0: |   9.1% | 0.377948 |       -- |    -- |  10,112.0 | USER
0: ||----------------------------------------------------------------------
0: ||   6.3% | 0.260671 | 0.015257 |  5.6% |     100.0 | Mugs::share
0: ||   2.3% | 0.094182 | 0.013986 | 13.1% |  10,000.0 | Faces::share
0: ||======================================================================
0: |   3.4% | 0.142381 |       -- |    -- |      81.0 | MPI_SYNC
0: ||----------------------------------------------------------------------
0: ||   3.4% | 0.139416 | 0.137014 | 98.3% |      66.0 | MPI_Barrier(sync)
0: |======================================================================
```

```
0: Table 1:  Profile by Function Group and Function
0:
0:   Time% |      Time |      Imb. |  Imb. |      Calls | Group
0:         |           |      Time | Time% |            |  Function=[MAX10]
0:         |           |           |       |            |   PE=HIDE
0:
0:
0:
0:
0:
0: ||  22.7% | 0.941439 | 0.473985 | 34.0% |  20,000.0 | hipStreamSynchronize
0: ||  19.1% | 0.792874 | 0.130271 | 14.3% |     141.0 | hipMemset
0: ||   4.1% | 0.169448 | 0.007623 |  4.4% |      20.0 | hipStreamCreate
0: ||   1.4% | 0.056639 | 0.008608 | 13.4% |  60,200.0 | __hipPushCallConfiguration
0: ||=================================================================
0: |  33.5% | 1.386777 |       -- |    -- | 545,708.0 | MPI
0: ||-----------------------------------------------------------------
0: ||  24.5% | 1.015371 | 0.242284 | 19.6% |  20,200.0 | MPI_Waitall
0: ||   7.0% | 0.289145 | 0.199618 | 41.5% | 262,600.0 | MPI_Isend
0: ||   1.9% | 0.080567 | 0.075127 | 49.0% | 262,600.0 | MPI_Irecv
0: ||=================================================================
0: |   9.1% | 0.377948 |       -- |    -- |  10,112.0 | USER
0: ||-----------------------------------------------------------------
0: ||   6.3% | 0.260671 | 0.015257 |  5.6% |     100.0 | Mugs::share
0: ||   2.3% | 0.094182 | 0.013986 | 13.1% |  10,000.0 | Faces::share
0: ||=================================================================
0: |   3.4% | 0.142381 |       -- |    -- |      81.0 | MPI_SYNC
0: ||-----------------------------------------------------------------
0: ||   3.4% | 0.139416 | 0.137014 | 98.3% |      66.0 | MPI_Barrier(sync)
0: |=================================================================
```

hides functions taking less than 1% of host runtime

```
0: Table 1:  Profile by Function Group and Function
0:
0:   Time% |      Time |      Imb. |   Imb. |      Calls | Group
0:         |           |      Time | Time% |            |  Function=[MAX10]
0:         |           |           |       |            |   PE=HIDE
0:
```

Imbalance Time = Maximum Time for any Single Task - Average Time across Tasks

```
0: ||  22.7% | 0.941439 | 0.473985 | 34.0% |  20,000.0 | hipStreamSynchronize
0: ||  19.1% | 0.792874 | 0.130271 | 14.3% |     141.0 | hipMemset
0: ||   4.1% | 0.169448 | 0.007623 |  4.4% |      20.0 | hipStreamCreate
```

*Imbalance Time % = (Imbalance Time / Max Time) * Tasks / (Tasks – 1)*

100% Imbalance Time means only one task spent time in that function

```
0: ||   2.3% | 0.094182 | 0.013986 | 13.1% |  10,000.0 | Faces::share
0: ||=============================================================================
0: |   3.4% | 0.142381 |       -- |    -- |      81.0 | MPI_SYNC
0: ||----------------------------------------------------------------------------
0: ||   3.4% | 0.139416 | 0.137014 | 98.3% |      66.0 | MPI_Barrier(sync)
0: |=============================================================================
```

```
0: Table 1:  Profile by Function Group and Function
0:
0:   Time% |     Time |     Imb. |  Imb. |     Calls | Group
0:         |          |     Time | Time% |           |  Function=[MAX10]
0:         |          |          |       |           |   PE=HIDE
0:
0:  100.0% | 4.142693 |       -- |    -- | 848,867.0 | Total
0: |-----------------------------------------------------------------------
0: |  53.5% | 2.215179 |       -- |    -- | 291,668.0 | HIP
0: ||----------------------------------------------------------------------
0: ||  22.7% | 0.941439 | 0.473985 | 34.0% |  20,000.0 | hipStreamSynchronize
0: ||  19.1% | 0.792874 | 0.130271 | 14.3% |     141.0 | hipMemset
0: ||   4.1% | 0.169448 | 0.007623 |  4.4% |      20.0 | hipStreamCreate
0: ||   1.4% | 0.056639 | 0.008608 | 13.4% |  60,200.0 | __hipPushCallConfiguration
0: ||======================================================================
0: |  33.5% | 1.386777 |       -- |    -- | 545,708.0 | MPI
0: ||----------------------------------------------------------------------
                                                  200.0 | MPI_Waitall
                                                  600.0 | MPI_Isend
                                                  600.0 | MPI_Irecv
                                                  ==============================
                                                   12.0 | USER
                                                  ------------------------------
                                                  100.0 | Mugs::share
0: ||   2.3% | 0.094182 | 0.013986 | 13.1% |  10,000.0 | Faces::share
0: ||======================================================================
0: |   3.4% | 0.142381 |       -- |    -- |      81.0 | MPI_SYNC
0: ||----------------------------------------------------------------------
0: ||   3.4% | 0.139416 | 0.137014 | 98.3% |      66.0 | MPI_Barrier(sync)
0: |======================================================================
```

host time spent on Hip API calls,
not kernel execution times on accelerator

MPI time on point to point
(and "active" work of collectives)

```
0: Table 1:  Profile by Function Group and Function
0:
0:   Time% |     Time |      Imb. |  Imb. |     Calls | Group
0:         |          |     Time  | Time% |           |  Function=[MAX10]
0:         |          |           |       |           |   PE=HIDE
                                                 48,867.0 | Total
                                                 291,668.0 | HIP
                                                  20,000.0 | hipStreamSynchronize
0: ||  19.1% | 0.792874 | 0.130271 | 14.3% |     141.0 | hipMemset
0: ||   4.1% | 0.169448 | 0.007623 |  4.4% |      20.0 | hipStreamCreate
0: ||   1.4% | 0.056639 | 0.008608 | 13.4% |  60,200.0 | __hipPushCallConfiguration
0: ||=========================================================================
0: |  33.5% | 1.386777 |       -- |    -- | 545,708.0 | MPI
0: ||-------------------------------------------------------------------------
0: ||  24.5% | 1.015371 | 0.242284 | 19.6% |  20,200.0 | MPI_Waitall
0: ||   7.0% | 0.289145 | 0.199618 | 41.5% | 262,600.0 | MPI_Isend
0: ||   1.9% | 0.080567 | 0.075127 | 49.0% | 262,600.0 | MPI_Irecv
0: ||=========================================================================
0: |   9.1% | 0.377948 |       -- |    -- |  10,112.0 | USER
0: ||-------------------------------------------------------------------------
0: ||   6.3% | 0.260671 | 0.015257 |  5.6% |     100.0 | Mugs::share
0: ||   2.3% | 0.094182 | 0.013986 | 13.1% |  10,000.0 | Faces::share
0: ||=========================================================================
0: |   3.4% | 0.142381 |       -- |    -- |      81.0 | MPI_SYNC
0: ||-------------------------------------------------------------------------
0: ||   3.4% | 0.139416 | 0.137014 | 98.3% |      66.0 | MPI_Barrier(sync)
0: |=========================================================================
```

MPI synchronization time
(Perftools adds a barrier before each collective)

```
0: Table 1:  Profile by Function Group and Function
0:
0:   Time% |      Time |      Imb. |   Imb. |      Calls | Group
0:         |           |     Time | Time% |            |  Function=[MAX10]
0:         |           |          |       |            |   PE=HIDE
0:
0:                                                        Total
0: |-------------------------------------------------------------------------
0: | HIP
0: |-------------------------------------------------------------------------
0: ||  19.1% | 0.792874 | 0.130271 | 14.3% |      141.0 | hipStreamSynchronize
0: ||   4.1% | 0.169448 | 0.007623 |  4.4% |       20.0 | hipMemset
0: ||   1.4% | 0.056639 | 0.008608 | 13.4% |   60,200.0 | hipStreamCreate
0: ||   1.4% | 0.056639 | 0.008608 | 13.4% |   60,200.0 | __hipPushCallConfiguration
0: ||=========================================================================
0: |  33.5% | 1.386777 |       -- |    -- | 545,708.0 | MPI
0: ||-------------------------------------------------------------------------
0: ||  24.5% | 1.015371 | 0.242284 | 19.6% |   20,200.0 | MPI_Waitall
0: ||   7.0% | 0.289145 | 0.199618 | 41.5% |  262,600.0 | MPI_Isend
0: ||   1.9% | 0.080567 | 0.075127 | 49.0% |  262,600.0 | MPI_Irecv
0: ||=========================================================================
0: |   9.1% | 0.377948 |       -- |    -- |   10,112.0 | USER
0: ||-------------------------------------------------------------------------
0: ||   6.3% | 0.260671 | 0.015257 |  5.6% |      100.0 | Mugs::share
0: ||   2.3% | 0.094182 | 0.013986 | 13.1% |   10,000.0 | Faces::share
0: ||=========================================================================
0: |   3.4% | 0.142381 |       -- |    -- |       81.0 | MPI_SYNC
0: ||-------------------------------------------------------------------------
0: ||   3.4% | 0.139416 | 0.137014 | 98.3% |       66.0 | MPI_Barrier(sync)
0: |=========================================================================
```

```
0: Observation:  MPI Grid Detection
0:
0:     There appears to be point-to-point MPI communication in a 4 X 4 X 4
0:     grid pattern. The 36.8% of the total execution time spent in MPI
0:     functions might be reduced with a rank order that maximizes
0:     communication between ranks on the same node. The effect of several
0:     rank orders is estimated below.
0:
0:     A file named MPICH_RANK_ORDER.Grid was generated along with this
0:     report and contains usage instructions and the Hilbert rank order
0:     from the following table.
0:
0:     Rank Order    On-Node     On-Node  MPICH_RANK_REORDER_METHOD
0:                  Bytes/PE  Bytes/PE%
0:                             of Total
0:                             Bytes/PE
0:
0:        Hilbert  3.776e+11      57.08%  3
0:            SMP  2.978e+11      45.02%  1
0:           Fold  2.576e+11      38.94%  2
0:     RoundRobin  2.346e+11      35.46%  0
```

CUTAWAY: MPICH_RANK_ORDER

or *Faster Route Now Available*

MPICH_RANK_ORDER.GRID

```
# The 'Grid' rank order in this file targets nodes with multi-core
# processors, based on Sent Msg Total Bytes collected for:
#
# Program:      .../faces+orig
# Ap2 File:     .../faces+19929-6079009t
# Number PEs:   64
# Max PEs/Node: 8
#
# To use this file, make a copy named MPICH_RANK_ORDER, and set the
# environment variable MPICH_RANK_REORDER_METHOD to 3 prior to
# executing the program.
#
# The following rank order was generated with the command:
#
#   grid_order -R -H -m 64 -n 8 -g 4x4x4 -c 1x1x1
#
0,1,17,16,20,21,5,4
8,24,28,12,13,29,25,9
10,26,30,14,15,31,27,11
7,3,2,6,22,18,19,23
39,35,34,38,54,50,51,55
59,43,47,63,62,46,42,58
57,41,45,61,60,44,40,56
52,53,37,36,32,33,49,48
```

HIPCC RUN WITH MPICH_RANK_REORDER

use same nodes for all runs

```
salloc -N 8 -t 30:00
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
cp faces+19929-6079009t/MPICH_RANK_ORDER.Grid ./MPICH_RANK_ORDER
export MPICH_RANK_REORDER_METHOD=3
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./face
```

HIPCC RUN WITH MPICH_RANK_REORDER

```
salloc -N 8 -t 30:00
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
cp faces+19929-6079009t/MPICH_RANK_ORDER.Grid ./MPICH_RANK_ORDER
export MPICH_RANK_REORDER_METHOD=3
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./face
```

two sets of three runs

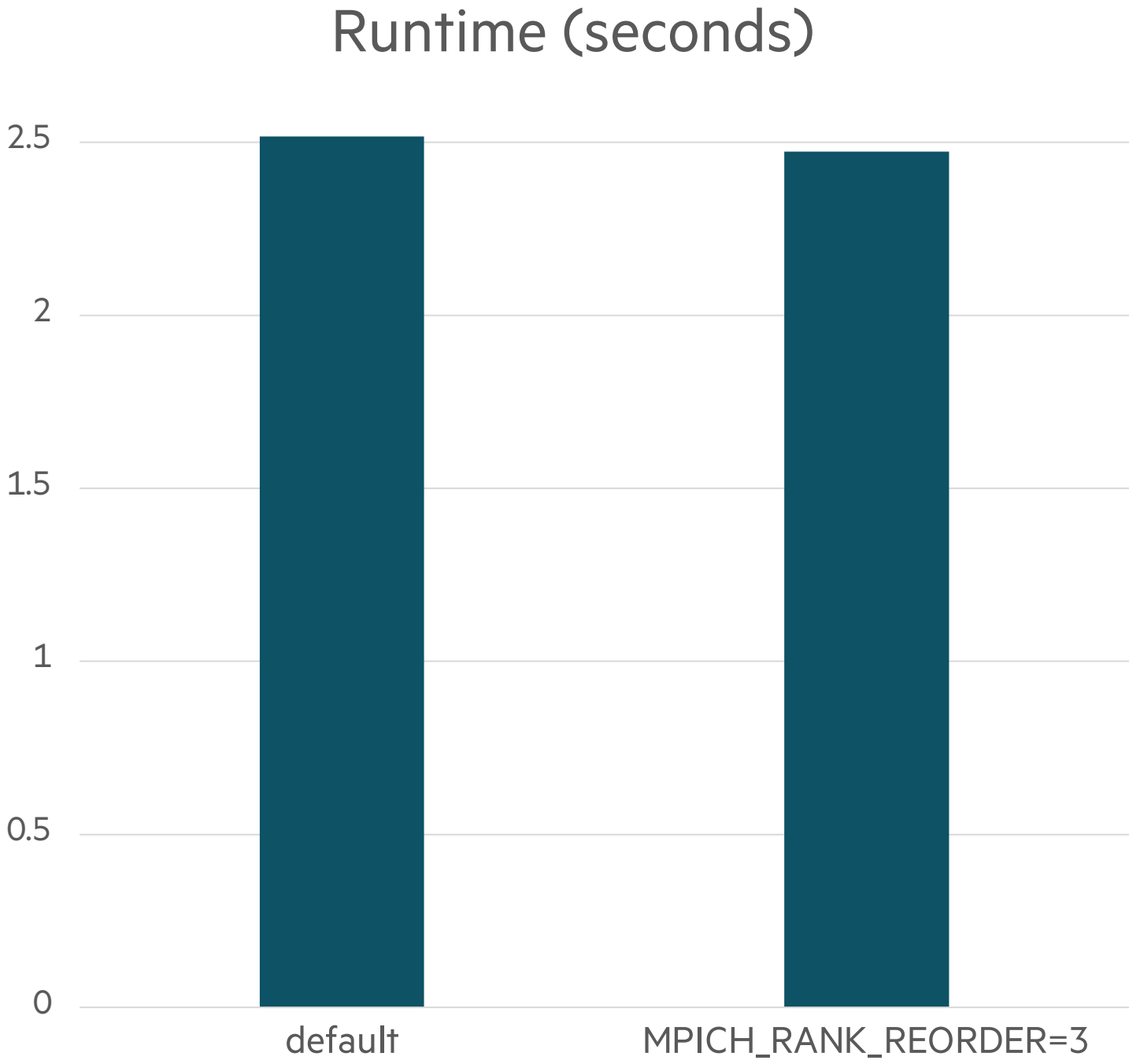
HIPCC RUN WITH MPICH_RANK_REORDER

```
salloc -N 8 -t 30:00
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
cp faces+19929-6079009t/MPICH_RANK_ORDER.Grid ./MPICH_RANK_ORDER
export MPICH_RANK_REORDER_METHOD=3
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./face
```

reorder MPI ranks of second set of runs

MPICH_RANK_ORDER

- 4x4x4 tasks (8 nodes)
- Average of 3 runs
- Same nodes
- 2% improvement
- But this is a small run



MORE PERFTOOLS-LITE-GPU OUTPUT

or *Back to Your Regularly Scheduled Program*

```
0: Table 2:  Time and Bytes Transferred for Accelerator Regions
0:
0:    Host | Host |    Acc |  Acc |  Acc Copy |  Events | Function=[max10]
0:   Time% | Time |  Time% | Time |       Out |         |  PE=HIDE
0:         |      |        |      | (MiBytes) |         |
0:
0:  100.0% | 1.93 | 100.0% | 2.68 |     36.00 | 111,066 | Total
0: |-------------------------------------------------------------------
0: |  49.7% | 0.96 |     -- |   -- |        -- |  20,000 | hipStreamSynchronize
0: |  34.5% | 0.67 |   0.0% | 0.00 |        -- |     141 | hipMemset
0: |   8.7% | 0.17 |     -- |   -- |        -- |      20 | hipStreamCreate
0: |   2.9% | 0.06 |  29.8% | 0.80 |        -- |  20,000 | hipKernel.gpuRun3x1<>
0: |   1.5% | 0.03 |  69.7% | 1.87 |        -- |  10,000 | hipKernel.gpuRun4x1<>
0: |===================================================================
```

```
0: Table 2:  Time and Bytes Transferred for Accelerator Regions
0:
0:    Host | Host |    Acc |  Acc |  Acc Copy |  Events | Function=[max10]
0:   Time% | Time |  Time% | Time |       Out |         |  PE=HIDE
0:         |      |        |      | (MiBytes) |         |
0:
0:  100.0% | 1.93 | 100.0% | 2.68 |     36.00 | 111,066 | Total
0: |-------------------------------------------------------------------
0: |  49.7% | 0.96 |     -- |   -- |        -- |  20,000 | hipStreamSynchronize
0: |  34.5% | 0.67 |   0.0% | 0.00 |        -- |     141 | hipMemset
0: |   8.7% | 0.17 |     -- |   -- |        -- |      20 | hipStreamCreate
0: |   2.9% | 0.06 |  29.8% | 0.80 |        -- |  20,000 | hipKernel.gpuRun3x1<>
0: |   1.5% | 0.03 |  69.7% | 1.87 |        -- |  10,000 | hipKernel.gpuRun4x1<>
0: |===================================================================
```

ordered by host time

host mostly waits for accelerator

```
0: Table 2:  Time and Bytes Transferred for Accelerator Regions
0:
0:    Host | Host |    Acc |  Acc |  Acc Copy |  Events | Function=[max10]
0:   Time% | Time |  Time% | Time |       Out |         |  PE=HIDE
0:         |      |        |      | (MiBytes) |         |
0:
0:  100.0% | 1.93 | 100.0% | 2.68 |     36.00 | 111,066 | Total
0: |-----------------------------------------------------------------------
0: |  49.7% | 0.96 |     -- |   -- |        -- |  20,000 | hipStreamSynchronize
0: |  34.5% | 0.67 |   0.0% | 0.00 |        -- |     141 | hipMemset
0: |   8.7% | 0.17 |     -- |   -- |        -- |      20 | hipStreamCreate
0: |   2.9% | 0.06 |  29.8% | 0.80 |        -- |  20,000 | hipKernel.gpuRun3x1<>
0: |   1.5% | 0.03 |  69.7% | 1.87 |        -- |  10,000 | hipKernel.gpuRun4x1<>
0: |=======================================================================
```

longest kernel ironically listed last

```
0: Table 3:  Program energy and power usage (from Cray PM)
0:
0:    Node | Node Power |  Process | Node Id=[mmm]
0:  Energy |         (W) |     Time |  PE=HIDE
0:     (J) |             |          |
0:
0:  48,007 | 11,105.819 | 4.322689 | Total
0: |-----------------------------------------
0: |  6,219 |  1,438.689 | 4.322687 | nid.5
0: |  5,979 |  1,383.223 | 4.322515 | nid.1
0: |  5,806 |  1,343.101 | 4.322832 | nid.6
0: |=========================================
```

```
0: Table 4:  File Input Stats by Filename
0:
0:  Avg Read | Avg Read |   Read Rate | Number |    Avg |   Bytes/ | File Name=!x/^/(proc|sys)/
0:  Time per |  MiBytes | MiBytes/sec |     of |  Reads |     Call |  PE=HIDE
0:    Reader |      per |             | Reader |    per |          |
0:      Rank |   Reader |             |  Ranks | Reader |          |
0:           |     Rank |             |        |   Rank |          |
0: |-----------------------------------------------------------------------
0: | 0.000081 | 0.000675 |    8.322485 |     64 |    3.0 |   236.00 | /opt/rocm-4.5.0/bin/.hipVersion
0: | 0.000016 | 0.000097 |    5.956450 |      1 |  102.0 |     1.00 | stdin
0: | 0.000009 | 0.004141 |  471.301375 |      1 |    1.0 | 4,342.00 | /tmp/comgr-f0045c/input/CompileSource
0: | 0.000008 | 0.004141 |  522.505222 |      1 |    1.0 | 4,342.00 | /tmp/comgr-dd1315/input/CompileSource
0: | 0.000008 | 0.004141 |  529.859742 |      1 |    1.0 | 4,342.00 | /tmp/comgr-ea4b92/input/CompileSource
0: | 0.000008 | 0.004141 |  533.340273 |      1 |    1.0 | 4,342.00 | /tmp/comgr-b59983/input/CompileSource
0: | 0.000008 | 0.004141 |  538.192602 |      1 |    1.0 | 4,342.00 | /tmp/comgr-5d3777/input/CompileSource
0: | 0.000007 | 0.004141 |  574.879062 |      1 |    1.0 | 4,342.00 | /tmp/comgr-68a2b0/input/CompileSource
0: | 0.000007 | 0.004141 |  587.938930 |      1 |    1.0 | 4,342.00 | /tmp/comgr-1a9bd1/input/CompileSource
0: | 0.000006 | 0.004141 |  680.949495 |      1 |    1.0 | 4,342.00 | /tmp/comgr-b33182/input/CompileSource
0: | 0.000005 | 0.004141 |  823.395085 |      1 |    1.0 | 4,342.00 | /tmp/comgr-9ba3ed/input/CompileSource
0: | 0.000005 | 0.004141 |  825.035641 |      1 |    1.0 | 4,342.00 | /tmp/comgr-0e2d22/input/CompileSource
0: | 0.000005 | 0.004141 |  836.705169 |      1 |    1.0 | 4,342.00 | /tmp/comgr-694329/input/CompileSource
0: | 0.000005 | 0.004141 |  902.540079 |      1 |    1.0 | 4,342.00 | /tmp/comgr-b67a4f/input/CompileSource
0: | 0.000004 | 0.004141 |  947.997684 |      1 |    1.0 | 4,342.00 | /tmp/comgr-e536c9/input/CompileSource
0: |=======================================================================
```

```
0: Table 4:  File Input Stats by Filename
0:
0:  Avg Read | Avg Read |   Read Rate | Number |    Avg |   Bytes/ | File Name=!x/^/(proc|sys)/
0:  Time per |  MiBytes | MiBytes/sec |     of |  Reads |     Call |  PE=HIDE
0:    Reader |      per |             | Reader |    per |          |
0:      Rank |   Reader |             |  Ranks | Reader |          |
0:           |     Rank |             |        |   Rank |          |
0: |-----------------------------------------------------------------------
0: | 0.000081 | 0.000675 |    8.322485 |     64 |    3.0 |   236.00 | /opt/rocm-4.5.0/bin/.hipVersion
0: | 0.000016 | 0.000097 |    5.956450 |      1 |  102.0 |     1.00 | stdin
0: | 0.000009 | 0.004141 |  471.301375 |      1 |    1.0 | 4,342.00 | /tmp/comgr-f0045c/input/CompileSource
0: | 0.000008 | 0.004141 |  522.505222 |      1 |    1.0 | 4,342.00 | /tmp/comgr-dd1315/input/CompileSource
0: | 0.000008 | 0.004141                                                -ea4b92/input/CompileSource
0: | 0.000008 | 0.004141                                                -b59983/input/CompileSource
0: | 0.000008 | 0.004141                                                -5d3777/input/CompileSource
0: | 0.000007 | 0.004141                                                -68a2b0/input/CompileSource
0: | 0.000007 | 0.004141 |  587.938930 |      1 |    1.0 | 4,342.00 | /tmp/comgr-1a9bd1/input/CompileSource
0: | 0.000006 | 0.004141 |  680.949495 |      1 |    1.0 | 4,342.00 | /tmp/comgr-b33182/input/CompileSource
0: | 0.000005 | 0.004141 |  823.395085 |      1 |    1.0 | 4,342.00 | /tmp/comgr-9ba3ed/input/CompileSource
0: | 0.000005 | 0.004141 |  825.035641 |      1 |    1.0 | 4,342.00 | /tmp/comgr-0e2d22/input/CompileSource
0: | 0.000005 | 0.004141 |  836.705169 |      1 |    1.0 | 4,342.00 | /tmp/comgr-694329/input/CompileSource
0: | 0.000005 | 0.004141 |  902.540079 |      1 |    1.0 | 4,342.00 | /tmp/comgr-b67a4f/input/CompileSource
0: | 0.000004 | 0.004141 |  947.997684 |      1 |    1.0 | 4,342.00 | /tmp/comgr-e536c9/input/CompileSource
0: |=======================================================================
```

each task checking in with Hip

```
0: Table 4:  File Input Stats by Filename
0:
0:  Avg Read | Avg Read |   Read Rate | Number |    Avg |   Bytes/ | File Name=!x/^/(proc|sys)/
0:  Time per |  MiBytes | MiBytes/sec |     of |  Reads |     Call |  PE=HIDE
0:    Reader |      per |             | Reader |    per |          |
0:      Rank |   Reader |             |  Ranks | Reader |          |
0:           |     Rank |             |        |   Rank |          |
0: |-----------------------------------------------------------------------
0: | 0.000081 | 0.000675 |    8.322485 |     64 |    3.0 |   236.00 | /opt/rocm-4.5.0/bin/.hipVersion
0: | 0.000016 | 0.000097 |    5.956450 |      1 |  102.0 |     1.00 | stdin
0: | 0.000009 | 0.004141 |  471.301375 |      1 |    1.0 | 4,342.00 | /tmp/comgr-f0045c/input/CompileSource
0: | 0.000008 | 0.004141 |  522.505222 |      1 |    1.0 | 4,342.00 | /tmp/comgr-dd1315/input/CompileSource
0: | 0.000008 | 0.004141 |                                                -ea4b92/input/CompileSource
0: | 0.000008 | 0.004141 |                                                -b59983/input/CompileSource
0: | 0.000008 | 0.004141 |                                                -5d3777/input/CompileSource
0: | 0.000007 | 0.004141 |                                                -68a2b0/input/CompileSource
0: | 0.000007 | 0.004141 |  587.938930 |      1 |    1.0 | 4,342.00 | /tmp/comgr-1a9bd1/input/CompileSource
0: | 0.000006 | 0.004141 |  680.949495 |      1 |    1.0 | 4,342.00 | /tmp/comgr-b33182/input/CompileSource
0: | 0.000005 | 0.004141 |  823.395085 |      1 |    1.0 | 4,342.00 | /tmp/comgr-9ba3ed/input/CompileSource
0: | 0.000005 | 0.004141 |  825.035641 |      1 |    1.0 | 4,342.00 | /tmp/comgr-0e2d22/input/CompileSource
0: | 0.000005 | 0.004141 |  836.705169 |      1 |    1.0 | 4,342.00 | /tmp/comgr-694329/input/CompileSource
0: | 0.000005 | 0.004141 |  902.540079 |      1 |    1.0 | 4,342.00 | /tmp/comgr-b67a4f/input/CompileSource
0: | 0.000004 | 0.004141 |  947.997684 |      1 |    1.0 | 4,342.00 | /tmp/comgr-e536c9/input/CompileSource
0: |=======================================================================
```

standard input from first task

AMD Rocm Code Object Manager

```
0: Table 4:  File Input S                                                    k/^/(proc|sys)/
0:
0:  Avg Read | Avg Read |                                                    |
0:  Time per |  MiBytes |                                                    |
0:    Reader |      per |                                                    |
0:      Rank |   Reader |                                                    |
0:           |     Rank |               |         |   Rank |           |
0: |-------------------------------------------------------------------------
0: | 0.000081 | 0.000675 |    8.322485 |     64 |    3.0 |   236.00 | /opt/rocm-4.5.0/bin/.hipVersion
0: | 0.000016 | 0.000097 |    5.956450 |      1 |  102.0 |     1.00 | stdin
0: | 0.000009 | 0.004141 |  471.301375 |      1 |    1.0 | 4,342.00 | /tmp/comgr-f0045c/input/CompileSource
0: | 0.000008 | 0.004141 |  522.505222 |      1 |    1.0 | 4,342.00 | /tmp/comgr-dd1315/input/CompileSource
0: | 0.000008 | 0.004141 |  529.859742 |      1 |    1.0 | 4,342.00 | /tmp/comgr-ea4b92/input/CompileSource
0: | 0.000008 | 0.004141 |  533.340273 |      1 |    1.0 | 4,342.00 | /tmp/comgr-b59983/input/CompileSource
0: | 0.000008 | 0.004141 |  538.192602 |      1 |    1.0 | 4,342.00 | /tmp/comgr-5d3777/input/CompileSource
0: | 0.000007 | 0.004141 |  574.879062 |      1 |    1.0 | 4,342.00 | /tmp/comgr-68a2b0/input/CompileSource
0: | 0.000007 | 0.004141 |  587.938930 |      1 |    1.0 | 4,342.00 | /tmp/comgr-1a9bd1/input/CompileSource
0: | 0.000006 | 0.004141 |  680.949495 |      1 |    1.0 | 4,342.00 | /tmp/comgr-b33182/input/CompileSource
0: | 0.000005 | 0.004141 |  823.395085 |      1 |    1.0 | 4,342.00 | /tmp/comgr-9ba3ed/input/CompileSource
0: | 0.000005 | 0.004141 |  825.035641 |      1 |    1.0 | 4,342.00 | /tmp/comgr-0e2d22/input/CompileSource
0: | 0.000005 | 0.004141 |  836.705169 |      1 |    1.0 | 4,342.00 | /tmp/comgr-694329/input/CompileSource
0: | 0.000005 | 0.004141 |  902.540079 |      1 |    1.0 | 4,342.00 | /tmp/comgr-b67a4f/input/CompileSource
0: | 0.000004 | 0.004141 |  947.997684 |      1 |    1.0 | 4,342.00 | /tmp/comgr-e536c9/input/CompileSource
0: |=========================================================================
```

```
0: Table 5:  File Output Stats by Filename
0:
0:       Avg |      Avg |   Write Rate | Number |    Avg |  Bytes/ Call | File Name=!x/^/(proc|sys)/
0:     Write |    Write |  MiBytes/sec |     of | Writes |              |  PE=HIDE
0:  Time per |  MiBytes |              | Writer |    per |              |
0:    Writer |      per |              |  Ranks | Writer |              |
0:      Rank |   Writer |              |        |   Rank |              |
0:           |     Rank |              |        |        |              |
0: |-------------------------------------------------------------------------------
0: | 0.006375 | 0.034980 |     5.486905 |     64 |  740.2 |        49.55 | _UnknownFile_
0: | 0.001101 | 2.747200 | 2,494.465733 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-ea8587/include/opencl1.2-c.pch
0: | 0.001087 | 2.747200 | 2,526.225974 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-1af37d/include/opencl1.2-c.pch
0: | 0.001071 | 2.747200 | 2,564.069071 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-5cd9b6/include/opencl1.2-c.pch
0: | 0.001062 | 2.747200 | 2,587.833464 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-e536c9/include/opencl1.2-c.pch
0: | 0.001050 | 2.747200 | 2,616.064545 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-ca74f7/include/opencl1.2-c.pch
0: | 0.001046 | 2.747200 | 2,625.319552 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-7b2ee9/include/opencl1.2-c.pch
0: | 0.001045 | 2.747200 | 2,627.838339 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-f0a91c/include/opencl1.2-c.pch
0: | 0.001039 | 2.747200 | 2,644.001971 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-0e5a38/include/opencl1.2-c.pch
0: | 0.001036 | 2.747200 | 2,650.885393 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-9ba3ed/include/opencl1.2-c.pch
0: | 0.001034 | 2.747200 | 2,655.795499 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-2c4c69/include/opencl1.2-c.pch
0: | 0.001034 | 2.747200 | 2,657.354845 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-e30237/include/opencl1.2-c.pch
0: | 0.001029 | 2.747200 | 2,670.357796 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-b59983/include/opencl1.2-c.pch
0: | 0.001012 | 2.747200 | 2,714.412622 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-d2c9be/include/opencl1.2-c.pch
0: | 0.001009 | 2.747200 | 2,723.524414 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-6d6ffb/include/opencl1.2-c.pch
0: |===============================================================================
```

```
0: Table 5:  File Output Stats by Filename
0:
0:       Avg |      Avg |   Write Rate | Number |    Avg |  Bytes/ Call | File Name=!x/^/(proc|sys)/
0:     Write |    Write |  MiBytes/sec |     of | Writes |              |  PE=HIDE
0:  Time per |  MiBytes |              | Writer |    per |              |
0:    Writer |      per |              |  Ranks | Writer |              |
0:      Rank |   Writer |              |        |   Rank |              |
0:           |     Rank |              |        |        |              |
0: |-------------------------------------------------------------------------
0: | 0.006375 | 0.034980 |     5.486905 |     64 |  740.2 |        49.55 | _UnknownFile_
0: | 0.001101 | 2.747200 | 2,494.465733 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-ea8587/include/opencl1.2-c.pch
0: | 0.001087 | 2.747200 | 2,526.225974 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-1af37d/include/opencl1.2-c.pch
0: | 0.001071 | 2.747200 | 2,564.069071 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-5cd9b6/include/opencl1.2-c.pch
0: | 0.001062 | 2.747200                                                    36c9/include/opencl1.2-c.pch
0: | 0.001050 | 2.747200                                                    74f7/include/opencl1.2-c.pch
0: | 0.001046 | 2.747200                                                    2ee9/include/opencl1.2-c.pch
0: | 0.001045 | 2.747200                                                    a91c/include/opencl1.2-c.pch
0: | 0.001039 | 2.747200                                                    5a38/include/opencl1.2-c.pch
0: | 0.001036 | 2.747200                                                    a3ed/include/opencl1.2-c.pch
0: | 0.001034 | 2.747200                                                    4c69/include/opencl1.2-c.pch
0: | 0.001034 | 2.747200                                                    0237/include/opencl1.2-c.pch
0: | 0.001029 | 2.747200 | 2,670.357796 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-b59983/include/opencl1.2-c.pch
0: | 0.001012 | 2.747200 | 2,714.412622 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-d2c9be/include/opencl1.2-c.pch
0: | 0.001009 | 2.747200 | 2,723.524414 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-6d6ffb/include/opencl1.2-c.pch
0: |=========================================================================
```

standard output from all tasks collected by Slurm

AMD Rocm Code Object Manager

```
0: Table 5:  File Output St                                                    /^/(proc|sys)/
0:
0:       Avg |       Avg |
0:     Write |     Write |
0:  Time per |  MiBytes |
0:    Writer |       per |
0:      Rank |   Writer |
0:           |     Rank |                  |         |         |                |
0: |-------------------------------------------------------------------------------
0: | 0.006375 | 0.034980 |     5.486905 |     64 |  740.2 |         49.55 | _UnknownFile_
0: | 0.001101 | 2.747200 | 2,494.465733 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-ea8587/include/opencl1.2-c.pch
0: | 0.001087 | 2.747200 | 2,526.225974 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-1af37d/include/opencl1.2-c.pch
0: | 0.001071 | 2.747200 | 2,564.069071 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-5cd9b6/include/opencl1.2-c.pch
0: | 0.001062 | 2.747200 | 2,587.833464 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-e536c9/include/opencl1.2-c.pch
0: | 0.001050 | 2.747200 | 2,616.064545 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-ca74f7/include/opencl1.2-c.pch
0: | 0.001046 | 2.747200 | 2,625.319552 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-7b2ee9/include/opencl1.2-c.pch
0: | 0.001045 | 2.747200 | 2,627.838339 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-f0a91c/include/opencl1.2-c.pch
0: | 0.001039 | 2.747200 | 2,644.001971 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-0e5a38/include/opencl1.2-c.pch
0: | 0.001036 | 2.747200 | 2,650.885393 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-9ba3ed/include/opencl1.2-c.pch
0: | 0.001034 | 2.747200 | 2,655.795499 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-2c4c69/include/opencl1.2-c.pch
0: | 0.001034 | 2.747200 | 2,657.354845 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-e30237/include/opencl1.2-c.pch
0: | 0.001029 | 2.747200 | 2,670.357796 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-b59983/include/opencl1.2-c.pch
0: | 0.001012 | 2.747200 | 2,714.412622 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-d2c9be/include/opencl1.2-c.pch
0: | 0.001009 | 2.747200 | 2,723.524414 |      1 |    2.0 | 1,440,324.00 | /tmp/comgr-6d6ffb/include/opencl1.2-c.pch
0: |===============================================================================
```


Finis

PLEASE, SIR, I WANT SOME MORE

- Run with *perftools-lite-gpu* to generate a default report at runtime
  - And save data to a directory: *./<exe>+<unique-ID>t* (*t* means it is a tracing experiment)
  - And generate *.ap2* files for *Apprentice2*
- Generate other reports with *pat_report*
- Visualize performance with *Apprentice2*

PAT_REPORT

or *The Kitchen Sink*

```
$ pat_report -O acc_time faces+19929-6079009t
...
Table 1:  Time and Bytes Transferred for Accelerator Regions

    Acc |  Acc | Host |  Acc Copy |  Events | Calltree
  Time% | Time | Time |       Out |         |  PE=HIDE
        |      |      | (MiBytes) |         |

 100.0% | 2.68 | 1.93 |     36.00 | 111,066 | Total
|-----------------------------------------------------------
| 100.0% | 2.68 | 1.93 |     36.00 | 111,066 | main
|  99.5% | 2.66 | 1.06 |        -- | 110,000 |  Faces::share
3  99.5% | 2.66 | 0.09 |        -- |  30,000 |   hipLaunchKernel
||||-----------------------------------------------------------
4|||  69.7% | 1.87 | 0.03 |        -- |  10,000 | hipKernel.gpuRun4x1<>
4|||  29.8% | 0.80 | 0.06 |        -- |  20,000 | hipKernel.gpuRun3x1<>
|===========================================================
...
```

```
$ pat_report -O acc_time faces+19929-6079009t
...
Table 1:  Time and Bytes Transferred for Accelerator Regions

    Acc |  Acc | Host |  Acc Copy |  Events | Calltree
  Time% | Time | Time |       Out |         |  PE=HIDE
        |      |      | (MiBytes) |         |

 100.0% | 2.68 | 1.93 |     36.00 | 111,066 | Total
|-----------------------------------------------------------
| 100.0% | 2.68 | 1.93 |     36.00 | 111,066 | main
|  99.5% | 2.66 | 1.06 |        -- | 110,000 |  Faces::share
3  99.5% | 2.66 | 0.09 |        -- |  30,000 |   hipLaunchKernel
||||-----------------------------------------------------------
4|||  69.7% | 1.87 | 0.03 |        -- |  10,000 | hipKernel.gpuRun4x1<>
4|||  29.8% | 0.80 | 0.06 |        -- |  20,000 | hipKernel.gpuRun3x1<>
|=============================================================
...
```

see significant data copies
(not much for this example)

```
$ pat_report -O acc_time faces+19929-6079009t
...
Table 1:  Time and Bytes Transferred for Accelerator Regions

    Acc |  Acc | Host |  Acc Copy |  Events | Calltree
  Time% | Time | Time |       Out |         |  PE=HIDE
        |      |      | (MiBytes) |         |

 100.0% | 2.68 | 1.93 |     36.00 | 111,066 | Total
|-----------------------------------------------------------
| 100.0% | 2.68 | 1.93 |     36.00 | 111,066 | main
|  99.5% | 2.66 | 1.06 |        -- | 110,000 |  Faces::share
3  99.5% | 2.66 | 0.09 |        -- |  30,000 |   hipLaunchKernel
||||-----------------------------------------------------------
4|||  69.7% | 1.87 | 0.03 |        -- |  10,000 | hipKernel.gpuRun4x1<>
4|||  29.8% | 0.80 | 0.06 |        -- |  20,000 | hipKernel.gpuRun3x1<>
|===========================================================
...
```

see kernel launches in order of accelerator runtime

```
$ pat_report -O acc_time -s show_ca=fu,so,li faces+19929-6079009t
...
Table 1:  Time and Bytes Transferred for Accelerator Regions

    Acc |  Acc | Host |  Acc Copy |  Events | Calltree
  Time% | Time | Time |       Out |         |  PE=HIDE
        |      |      | (MiBytes) |         |

 100.0% | 2.68 | 1.93 |     36.00 | 111,066 | Total
|-----------------------------------------------------------------------
|  99.5% | 2.66 | 1.06 |        -- | 110,000 | main:faces/hip/cug/main.cpp:line.165
||----------------------------------------------------------------------
||  69.7% | 1.87 | 0.03 |        -- |  30,000 | Faces::share:hip/cug/./gpu.hpp:line.189
3|  69.7% | 1.87 | 0.03 |        -- |  10,000 |  hipLaunchKernel:==NA==
4|        |      |      |           |         |   hipKernel.gpuRun4x1<>
||  29.8% | 0.80 | 0.07 |        -- |  60,000 | Faces::share:hip/cug/./gpu.hpp:line.153
3|  29.8% | 0.80 | 0.06 |        -- |  20,000 |  hipLaunchKernel:==NA==
4|        |      |      |           |         |   hipKernel.gpuRun3x1<>
|=======================================================================
...
```

See where the kernels were called?

TEMPLATES, TEMPLATES, EVERY WHERE

- C++ codes often launch kernels from template functions that take lambda arguments
- "Real" kernel code is in the user-provided lambdas
- Think "portability layers", Kokkos, Raja, Alpaka, Yakl, etc.
- Templates get inlined into user code
- Profiles show line numbers somewhere inside portability layers
  instead of line numbers in user code where lambdas appear

```
$ pat_report -O acc_time -s show_ca=fu,so,li faces+19929-6079009t
...
Table 1:  Time and Bytes Transferred for Accelerator Regions

    Acc |  Acc | Host |  Acc Copy |  Events | Calltree
  Time% | Time | Time |       Out |         |  PE=HIDE
        |      |      | (MiBytes) |         |

 100.0% | 2.68 | 1.93 |     36.00 | 111,066 | Total
|-----------------------------------------------------------------------
|  99.5% | 2.66 | 1.06 |        -- | 110,000 | main:faces/hip/cug/main.cpp:line.165
||----------------------------------------------------------------------
||  69.7% | 1.87 | 0.03 |        -- |  30,000 | Faces::share:hip/cug/./gpu.hpp:line.189
3|  69.7% | 1.87 | 0.03 |        -- |  10,000 |  hipLaunchKernel:==NA==
4|        |      |      |           |         |   hipKernel.gpuRun4x1<>
||  29.8% | 0.80 | 0.07 |        -- |  60,000 | Faces::share:hip/cug/./gpu.hpp:line.153
3|  29.8% | 0.80 | 0.06 |        -- |  20,000 |  hipLaunchKernel:==NA==
4|        |      |      |           |         |   hipKernel.gpuRun3x1<>
|=======================================================================
...
```

the call to Faces::share()

```
$ pat_report -O acc_time -s show_ca=fu,so,li faces+19929-6079009t
...
Table 1:  Time and Bytes Transferred for Accelerator Regions

    Acc |  Acc | Host |  Acc Copy |  Events | Calltree
  Time% | Time | Time |       Out |         |  PE=HIDE
        |      |      | (MiBytes) |         |

 100.0% | 2.68 | 1.93 |     36.00 | 111,066 | Total
|-----------------------------------------------------------------------
|  99.5% | 2.66 | 1.06 |        -- | 110,000 | main:faces/hip/cug/main.cpp:line.165
||----------------------------------------------------------------------
||  69.7% | 1.87 | 0.03 |        -- |  30,000 | Faces::share:hip/cug/./gpu.hpp:line.189
3|  69.7% | 1.87 | 0.03 |        -- |  10,000 |  hipLaunchKernel:==NA==
4|        |      |      |           |         |   hipKernel.gpuRun4x1<>
||  29.8% | 0.80 | 0.07 |        -- |  60,000 | Faces::share:hip/cug/./gpu.hpp:line.153
3|  29.8% | 0.80 | 0.06 |        -- |  20,000 |  hipLaunchKernel:==NA==
4|        |      |      |           |         |   hipKernel.gpuRun3x1<>
|=======================================================================
...
```

lines in the "portability layer"

WORKAROUND

- Turn off compiler inlining of top layer of portability-layer templates

```
template <typename F>
#ifdef CRAYPAT
__attribute__((noinline))
#endif
void gpuFor(const std::initializer_list<int> il0, F f,
  const hipStream_t stream = 0)
{
  ...
```

- Minimal impact on runtime
  - Kernel launches are much more expensive than host function calls
- But changes are in portability layer, not user code

```
$ pat_report -O acc_time -s show_ca=fu,so,li faces+8604-6077956t
...
Table 1:  Ti
```

changed code, recompiled, reran, and regenerated report

```
    Acc |  A
  Time% | Ti
        |       |       | (MiBytes) |          |

 100.0% | 2.64 | 1.81 |     36.00 | 111,066 | Total
|-------------------------------------------------------------------------
|  99.5% | 2.63 | 1.02 |        -- | 110,000 | main:faces/hip/cug/main.cpp:line.165
||------------------------------------------------------------------------
||  70.7% | 1.87 | 0.03 |        -- |  30,000 | Faces::share:faces/hip/cug/Faces.cpp:line.241
3|        |      |      |           |         |  gpuFor<>:hip/cug/./gpu.hpp:line.205
4|  70.7% | 1.87 | 0.03 |        -- |  10,000 |   hipLaunchKernel:==NA==
5|        |      |      |           |         |    hipKernel.gpuRun4x1<>
||  19.2% | 0.51 | 0.04 |        -- |  30,000 | Faces::share:faces/hip/cug/Faces.cpp:line.322
3|        |      |      |           |         |  gpuFor<>:hip/cug/./gpu.hpp:line.166
4|  19.2% | 0.51 | 0.03 |        -- |  10,000 |   hipLaunchKernel:==NA==
5|        |      |      |           |         |    hipKernel.gpuRun3x1<>
||   9.6% | 0.25 | 0.03 |        -- |  30,000 | Faces::share:faces/hip/cug/Faces.cpp:line.175
3|        |      |      |           |         |  gpuFor<>:hip/cug/./gpu.hpp:line.166
4|   9.6% | 0.25 | 0.03 |        -- |  10,000 |   hipLaunchKernel:==NA==
5|        |      |      |           |         |    hipKernel.gpuRun3x1<>
|=========================================================================
...
```

```
$ pat_report -O acc_time -s show_ca=fu,so,li faces+8604-6077956t
...
Table 1:  Time and Bytes Transferred for Accelerator Regions

    Acc |  Acc | Host |  Acc Copy |  Events | Calltree
  Time% | Time | Time |       Out |         |  PE=HIDE
        |      |      | (MiBytes) |         |

 100.0% | 2.64 | 1.81 |     36.00 | 111,066 | Total
|-----------------------------------------------------------------------
|  99.5% | 2.63 | 1.02 |        -- | 110,000 | main:faces/hip/cug/main.cpp:line.165
||----------------------------------------------------------------------
||  70.7% | 1.87 | 0.03 |        -- |  30,000 | Faces::share:faces/hip/cug/Faces.cpp:line.241
3|        |      |      |           |         |  gpuFor<>:hip/cug/./gpu.hpp:line.205
4|  70.7% | 1.87 | 0.03 |        -- |  10,000 |   hipLaunchKernel:==NA==
5|        |      |      |           |         |    hipKernel.gpuRun4x1<>
||  19.2% | 0.51 | 0.04 |        -- |  30,000 | Faces::share:faces/hip/cug/Faces.cpp:line.322
3|        |      |      |           |         |  gpuFor<>:hip/cug/./gpu.hpp:line.166
4|  19.2% | 0.51 | 0.03 |        -- |  10,000 |   hipLaunchKernel:==NA==
5|        |      |      |           |         |    hipKernel.gpuRun3x1<>
||   9.6% | 0.25 | 0.03 |        -- |  30,000 | Faces::share:faces/hip/cug/Faces.cpp:line.175
3|        |      |      |           |         |  gpuFor<>:hip/cug/./gpu.hpp:line.166
4|   9.6% | 0.25 | 0.03 |        -- |  10,000 |   hipLaunchKernel:==NA==
5|        |      |      |           |         |    hipKernel.gpuRun3x1<>
|=======================================================================
...
```

Where the lambdas are!

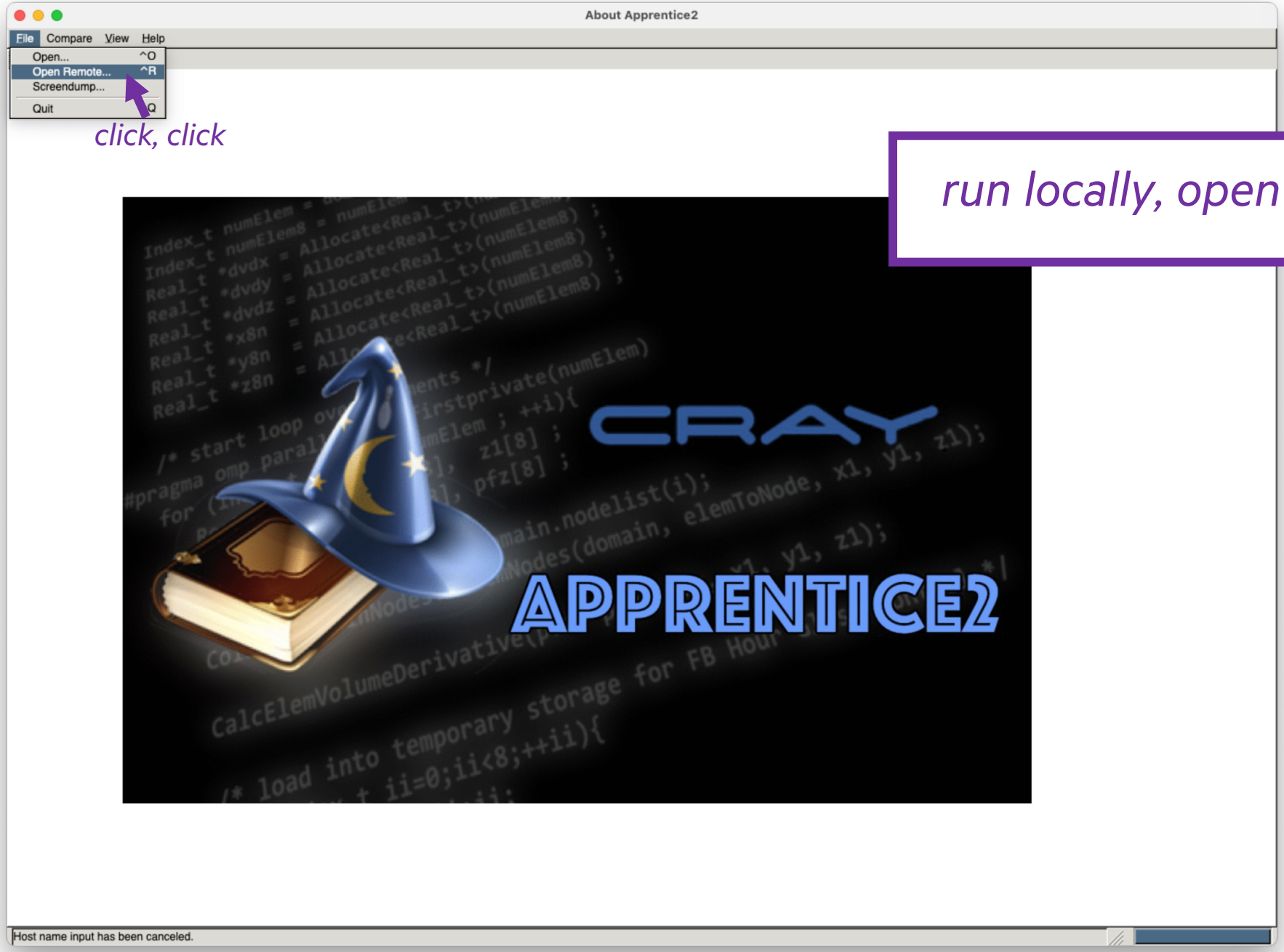
APPRENTICE2

or *Around the World in 80 Clicks*

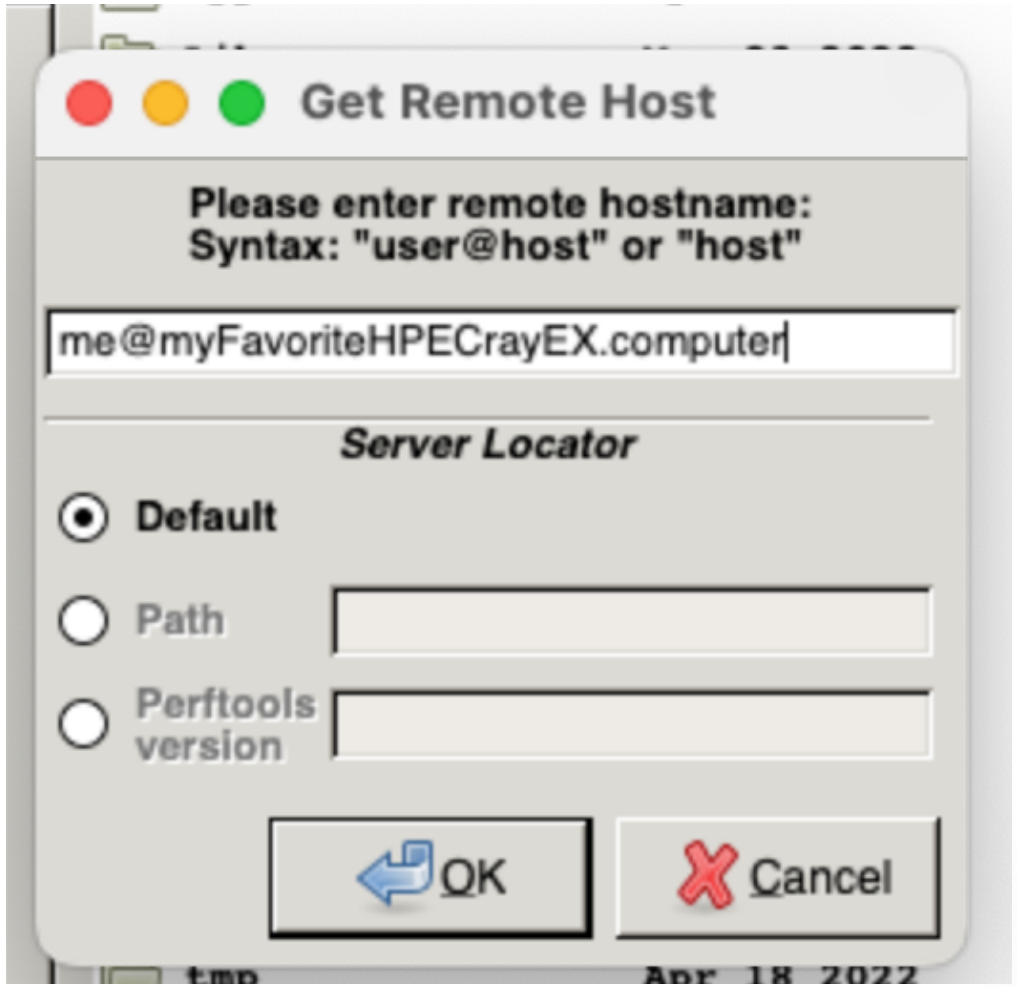
DOWNLOAD APPRENTICE2 TO YOUR LOCAL COMPUTER

```
$ ls -1 $CRAYPAT_ROOT/share/desktop_installers/App*
/opt/cray/pe/perftools/21.12.0/share/desktop_installers/Apprentice2Installer-21.12.0-4.dmg
/opt/cray/pe/perftools/21.12.0/share/desktop_installers/Apprentice2Installer-21.12.0-4.exe
```

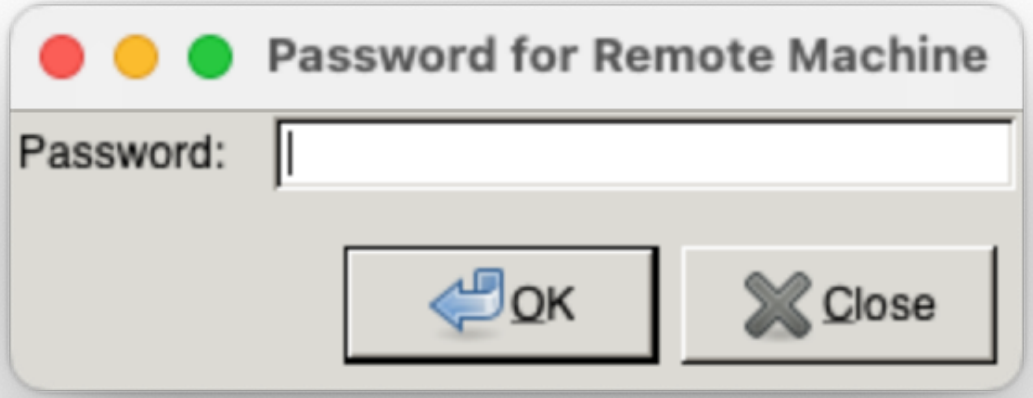

About Apprentice2
File Compare View Help
CRAY
APPRENTICE2

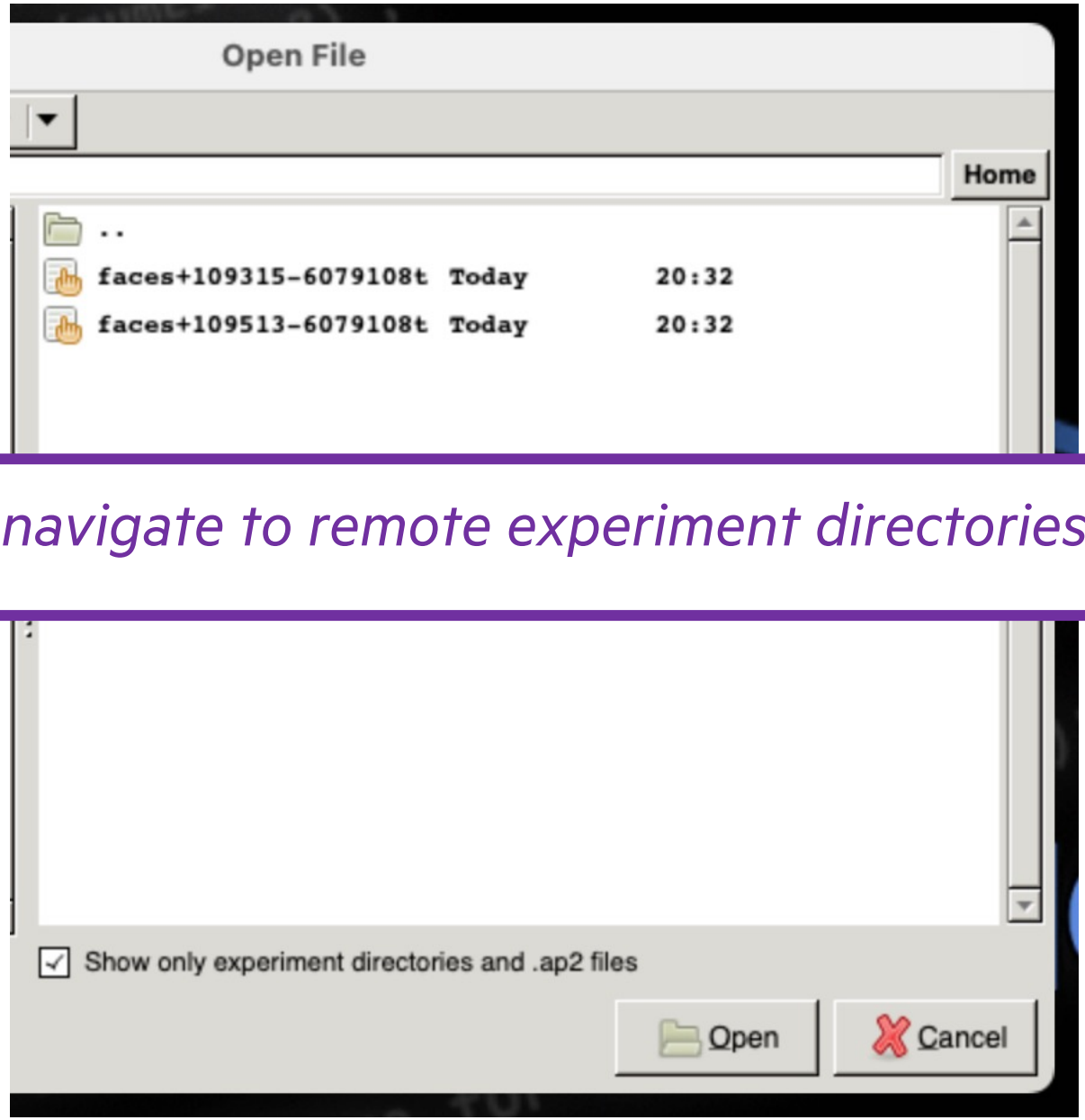


run locally, open remotely

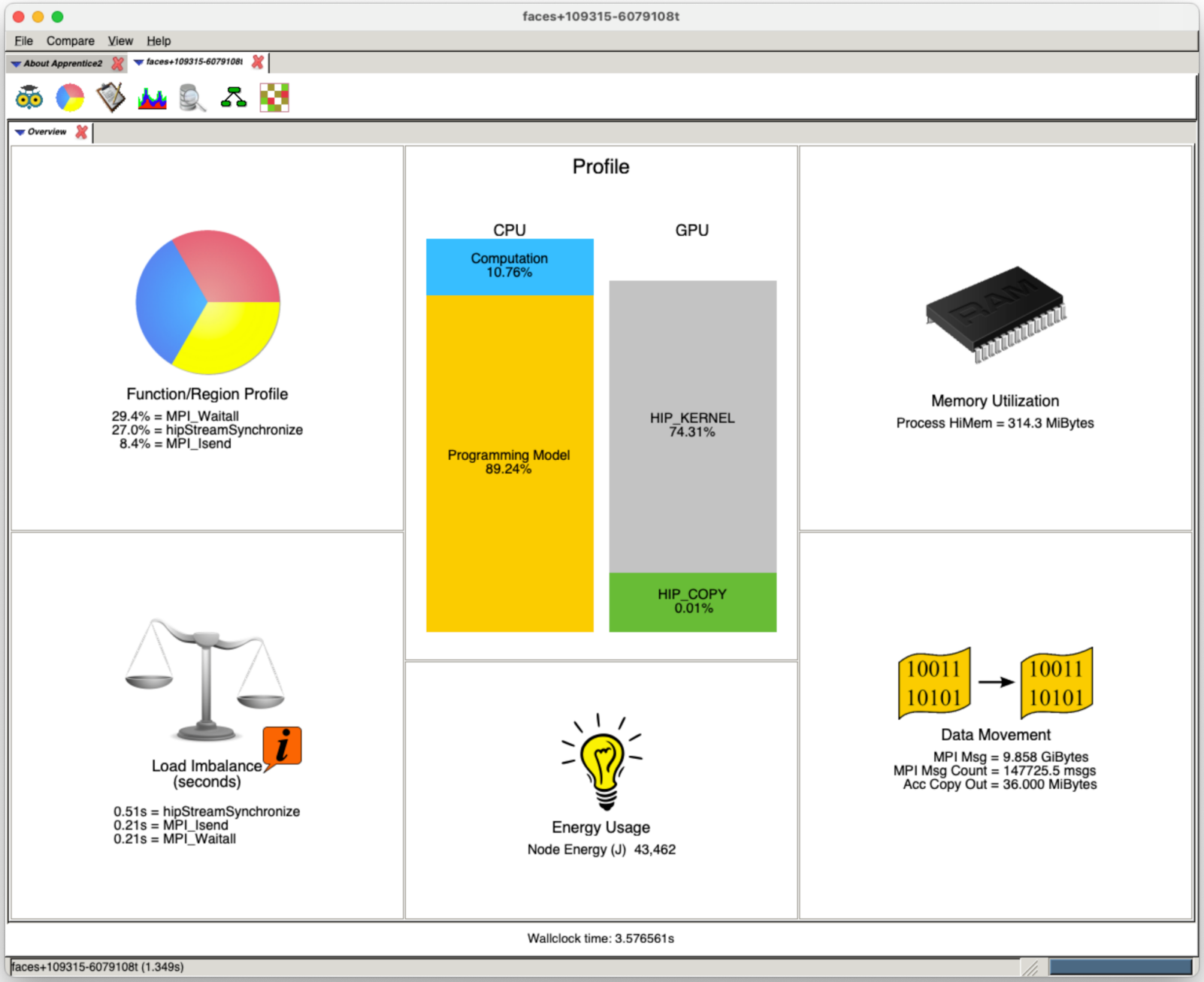


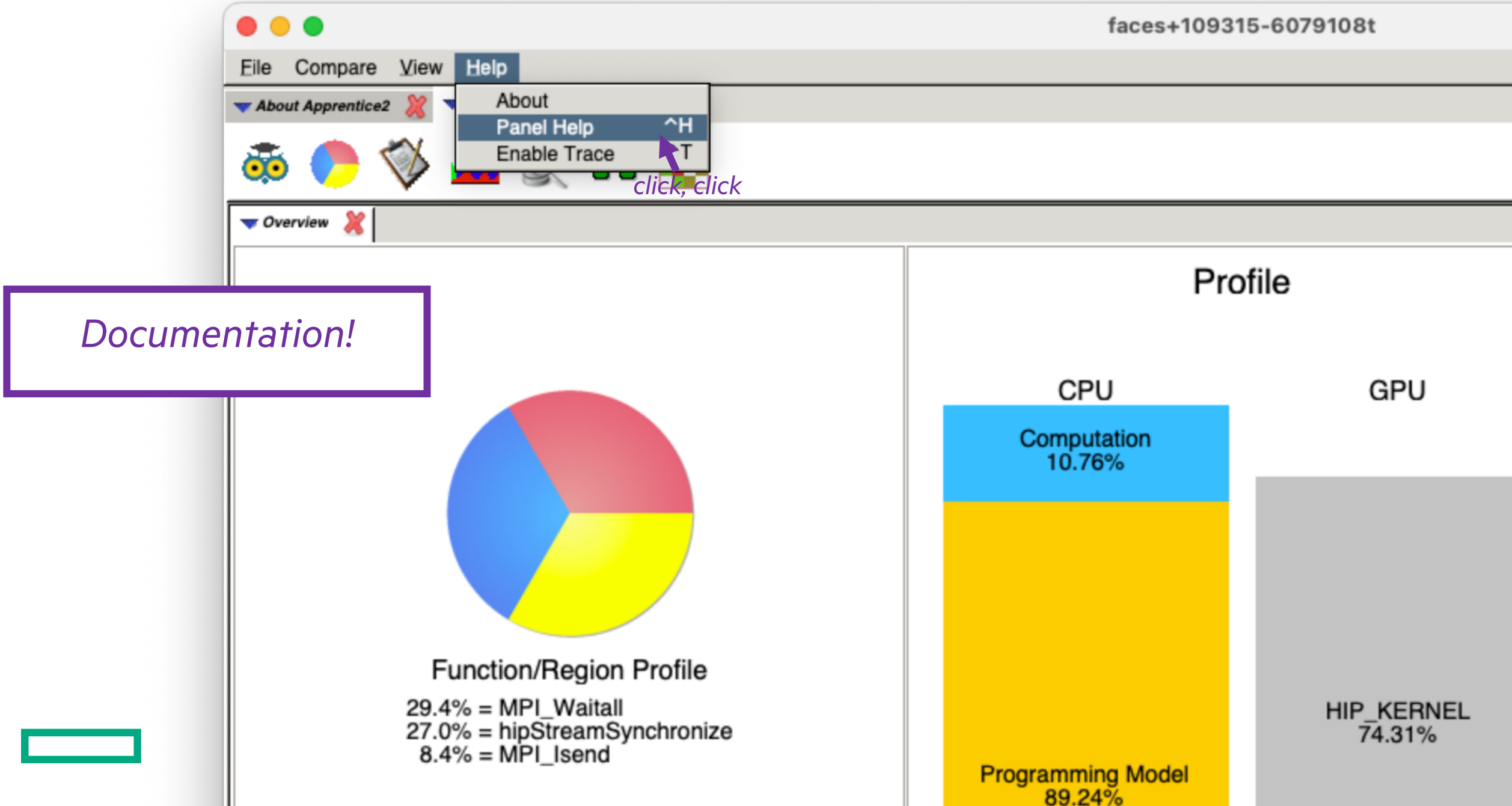
login to remote computer





navigate to remote experiment directories


faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
CPU
GPU
Computation
10.76%
Programming Model
89.24%
HIP_KERNEL
74.31%
HIP_COPY
0.01%
Function/Region Profile
29.4% = MPI_Waitall
27.0% = hipStreamSynchronize
8.4% = MPI_Isend
Memory Utilization
Process HiMem = 314.3 MiBytes
Load Imbalance
(seconds)
0.51s = hipStreamSynchronize
0.21s = MPI_Isend
0.21s = MPI_Waitall
Energy Usage
Node Energy (J) 43,462
10011
10101
10011
10101
Data Movement
MPI Msg = 9.858 GiBytes
MPI Msg Count = 147725.5 msgs
Acc Copy Out = 36.000 MiBytes
Wallclock time: 3.576561s
faces+109315-6079108t (1.349s)


faces+109315-6079108t
File
Compare
View
Help
About
Panel Help ^H
Enable Trace T
click, click
About Apprentice2
Overview
Documentation!
Function/Region Profile
29.4% = MPI_Waitall
27.0% = hipStreamSynchronize
8.4% = MPI_Isend
Profile
CPU
GPU
Computation
10.76%
Programming Model
89.24%
HIP_KERNEL
74.31%

Apprentice2Help.pdf
Page 1 of 35

Basic Help

Getting Started

Cray Apprentice2 is an interactive tool for visualizing and manipulating performance analysis data captured during program execution.

The number and appearance of the reports generated using Cray Apprentice2 is determined solely by the kind and quantity of data that has been captured. For example, if you use Cray Apprentice2 to analyze data that was captured using CrayPat on a Cray system, setting the environment variable **PAT_RT_SUMMARY** to **0** (zero) before executing the instrumented program will nearly double the number of reports available when analyzing the resulting data in Cray Apprentice2 -- but at the cost of greatly increased experiment data size due to the additional information being collected.

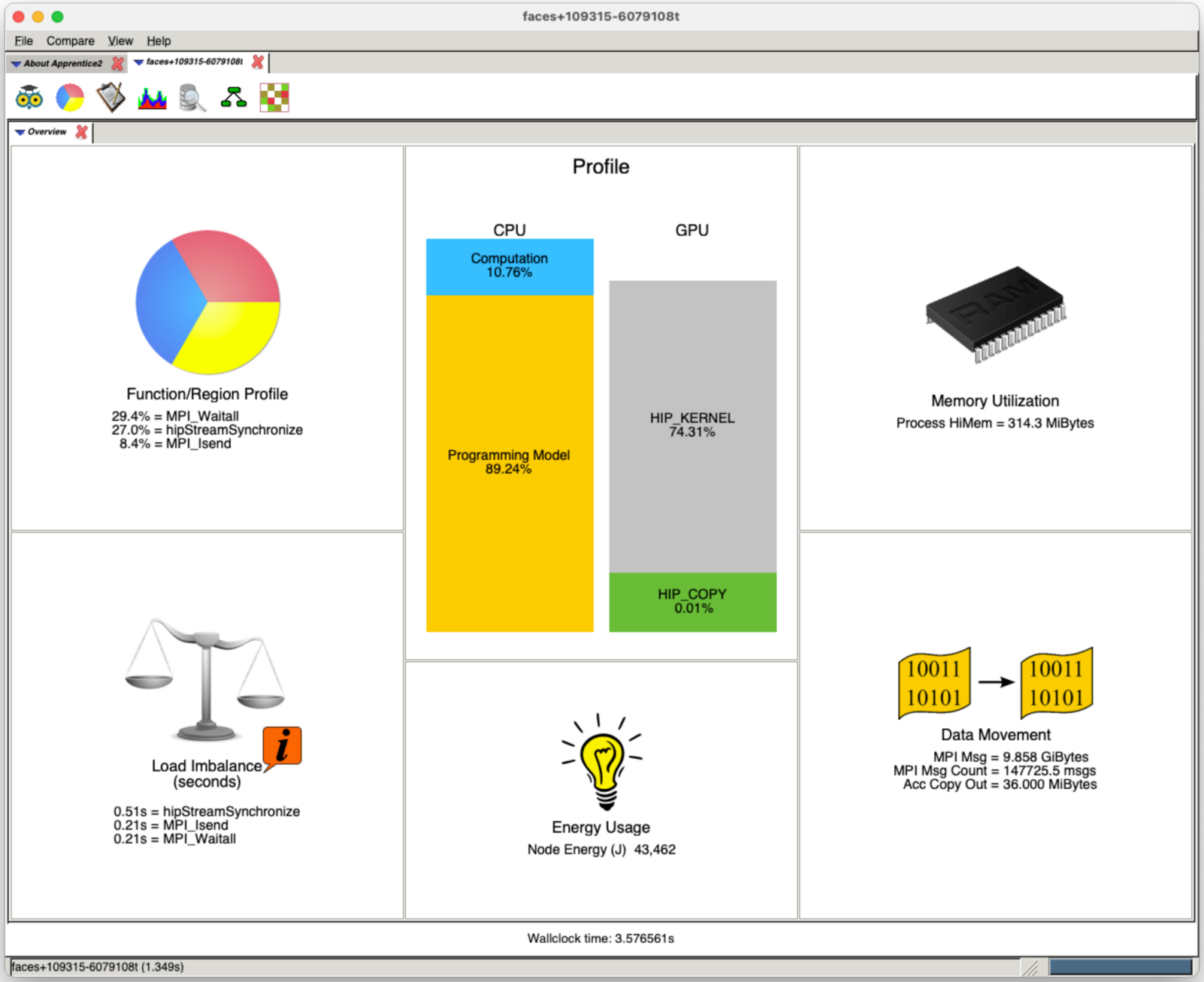
To begin using Cray Apprentice2, select an experiment data directory to open. After you select a experiment data, the data is read in and the Overview report is displayed.

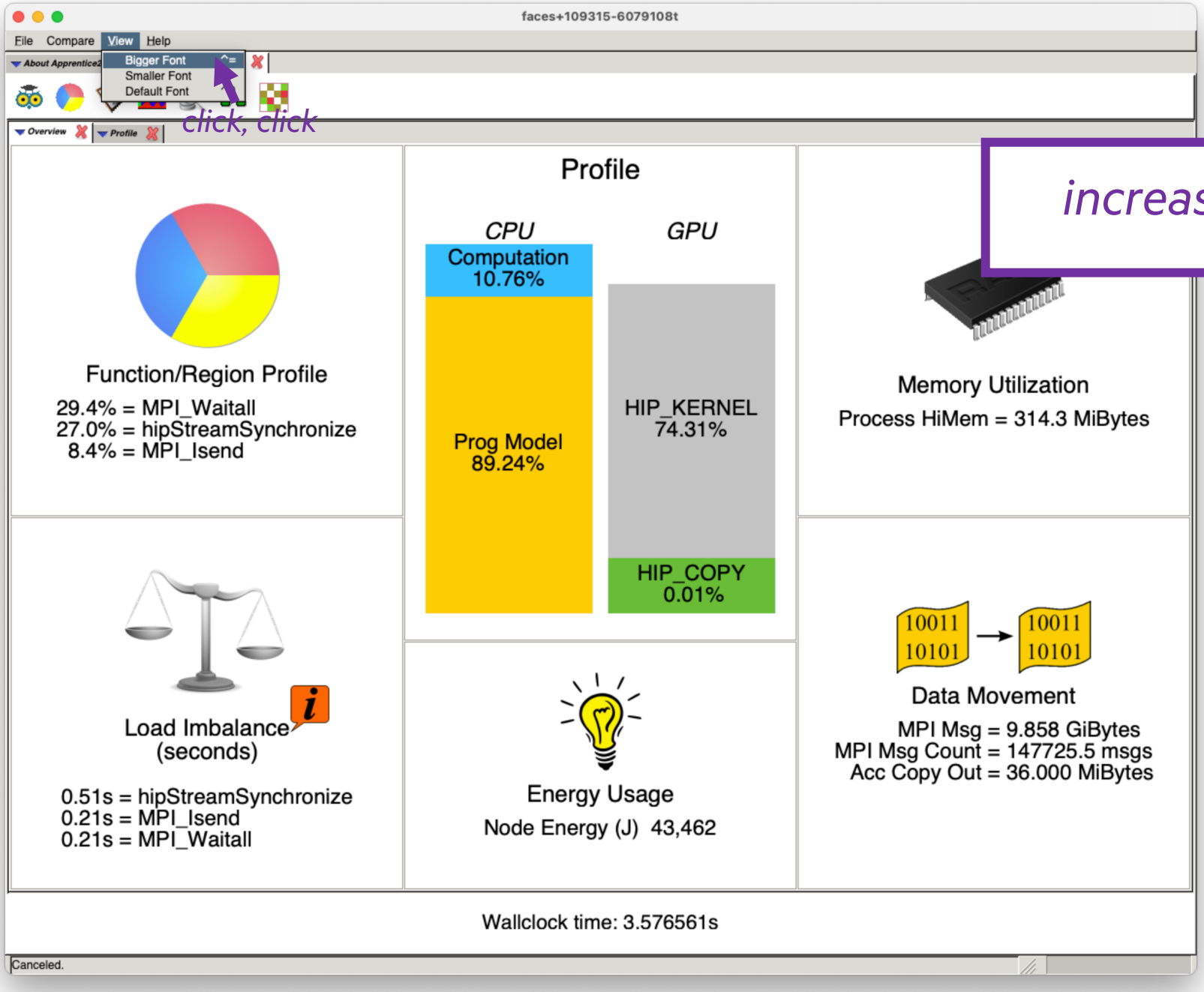
All reports share common navigation and user-interface functions and right-click menus.

Common Functions

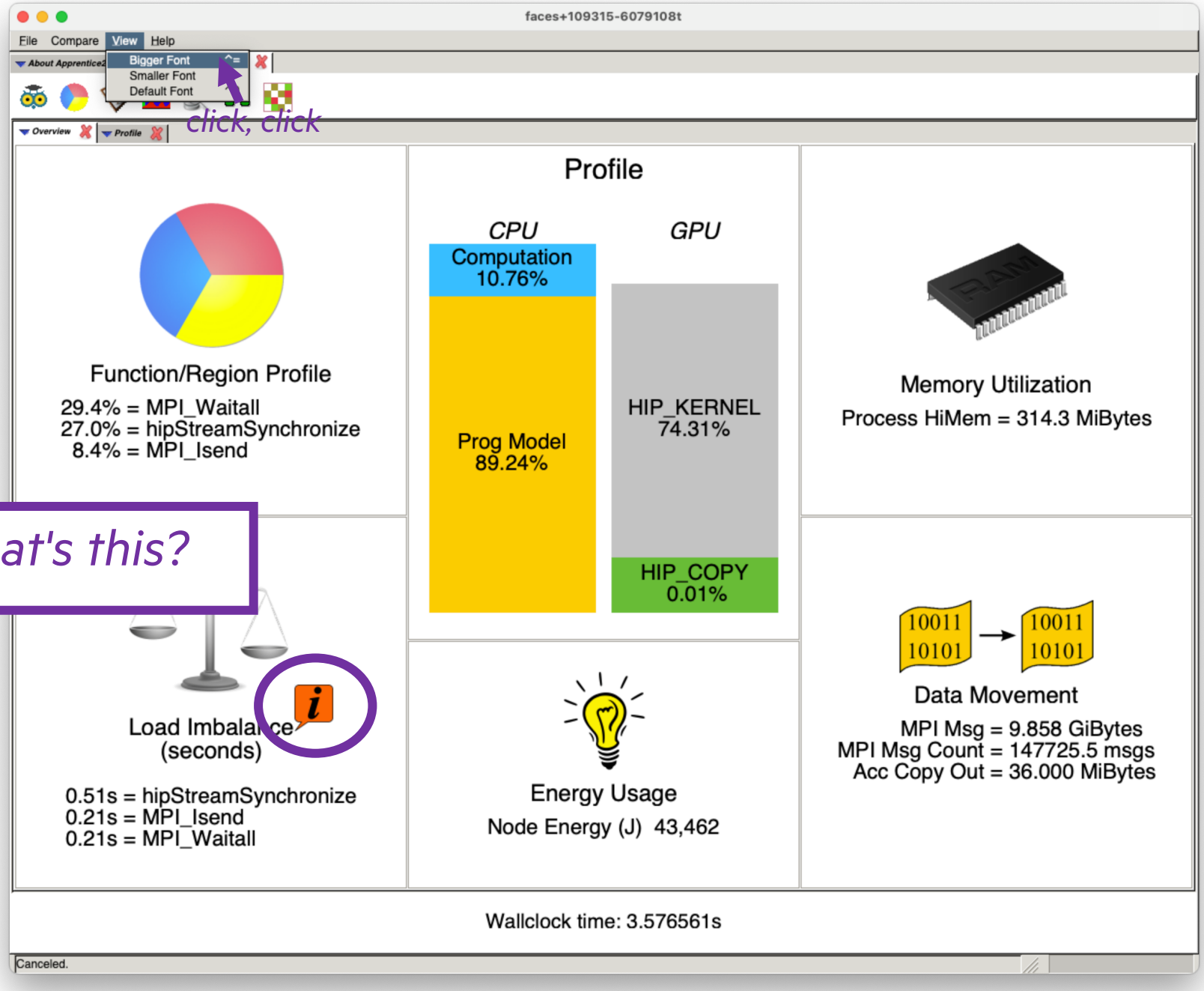
All reports share the following general navigation and user-interface features.

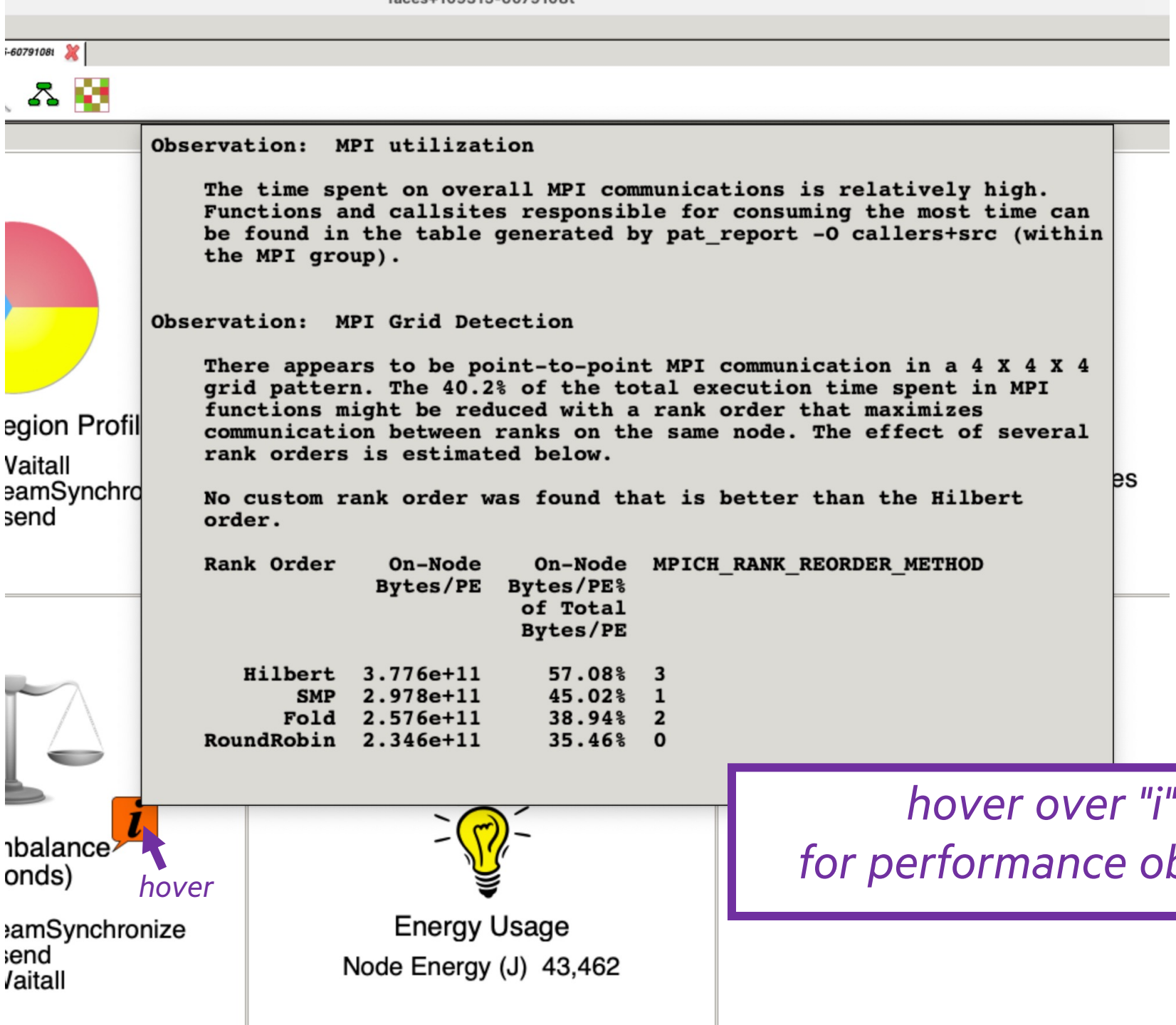
- The **File Menu** enables you to experiment data, capture the current screen display to a .png file, or exit Cray Apprentice2.
- The **View Menu** contains menu items for changing the size of the text for the various reports. You can choose **Bigger Font** or **Smaller Font** to make the report font sizes bigger or smaller, respectively, or choose **Default Font** to change the size back to the original size.
- The **Data File Tabs**, at the top of the display, show the names of the experiment data currently displayed. Multiple experiment data tabs can be open simultaneously. Click a

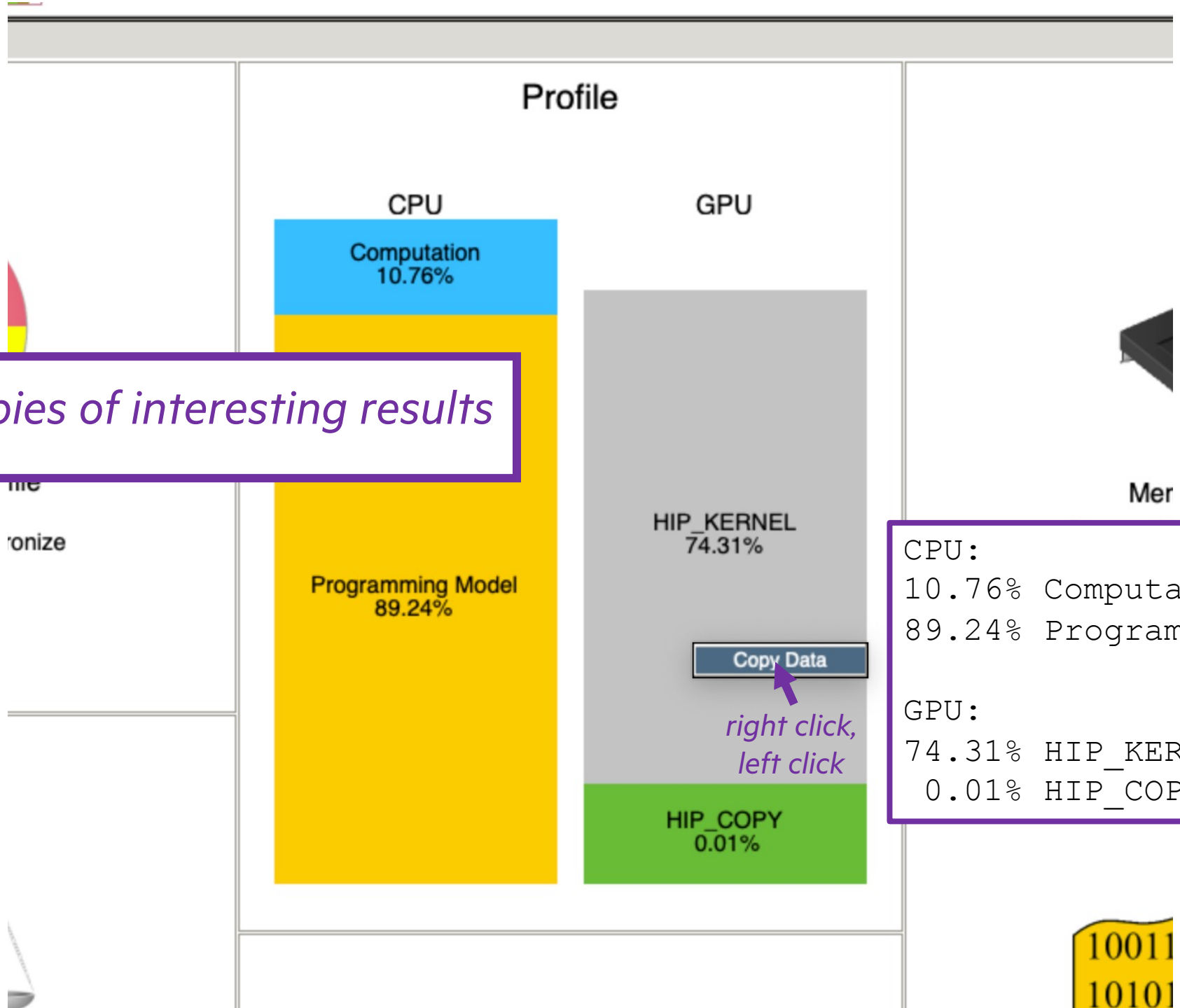

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Function/Region Profile
29.4% = MPI_Waitall
27.0% = hipStreamSynchronize
8.4% = MPI_Isend
Profile
CPU
GPU
Computation
10.76%
Programming Model
89.24%
HIP_KERNEL
74.31%
HIP_COPY
0.01%
Memory Utilization
Process HiMem = 314.3 MiBytes
Load Imbalance
(seconds)
0.51s = hipStreamSynchronize
0.21s = MPI_Isend
0.21s = MPI_Waitall
Energy Usage
Node Energy (J) 43,462
10011
10101
10011
10101
Data Movement
MPI Msg = 9.858 GiBytes
MPI Msg Count = 147725.5 msgs
Acc Copy Out = 36.000 MiBytes
Wallclock time: 3.576561s
faces+109315-6079108t (1.349s)



increase font size


faces+109315-6079108t
File Compare View Help
Bigger Font
Smaller Font
Default Font
click, click
Overview
Profile
Profile
CPU
GPU
Computation 10.76%
Prog Model 89.24%
HIP_KERNEL 74.31%
HIP_COPY 0.01%
Function/Region Profile
29.4% = MPI_Waitall
27.0% = hipStreamSynchronize
8.4% = MPI_Isend
Memory Utilization
Process HiMem = 314.3 MiBytes
What's this?
Load Imbalance (seconds)
0.51s = hipStreamSynchronize
0.21s = MPI_Isend
0.21s = MPI_Waitall
Energy Usage
Node Energy (J) 43,462
10011 10101
10011 10101
Data Movement
MPI Msg = 9.858 GiBytes
MPI Msg Count = 147725.5 msgs
Acc Copy Out = 36.000 MiBytes
Wallclock time: 3.576561s
Canceled.

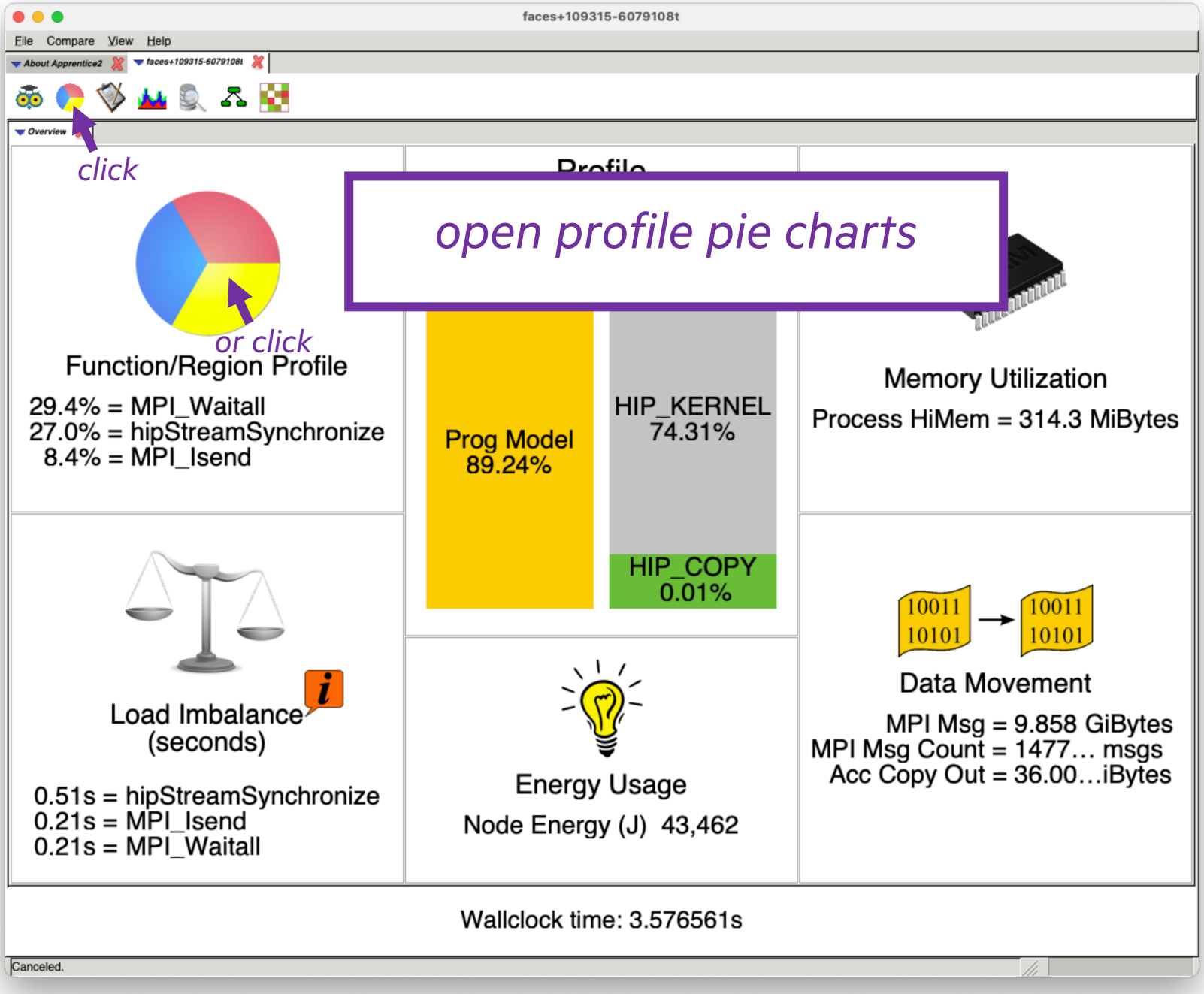

Observation: MPI utilization
The time spent on overall MPI communications is relatively high. Functions and callsites responsible for consuming the most time can be found in the table generated by pat_report -O callers+src (within the MPI group).
Observation: MPI Grid Detection
There appears to be point-to-point MPI communication in a 4 X 4 X 4 grid pattern. The 40.2% of the total execution time spent in MPI functions might be reduced with a rank order that maximizes communication between ranks on the same node. The effect of several rank orders is estimated below.
No custom rank order was found that is better than the Hilbert order.
Rank Order On-Node Bytes/PE On-Node Bytes/PE% of Total Bytes/PE MPICH_RANK_REORDER_METHOD
Hilbert 3.776e+11 57.08% 3
SMP 2.978e+11 45.02% 1
Fold 2.576e+11 38.94% 2
RoundRobin 2.346e+11 35.46% 0
hover
Energy Usage
Node Energy (J) 43,462
hover over "i" icons
for performance observations

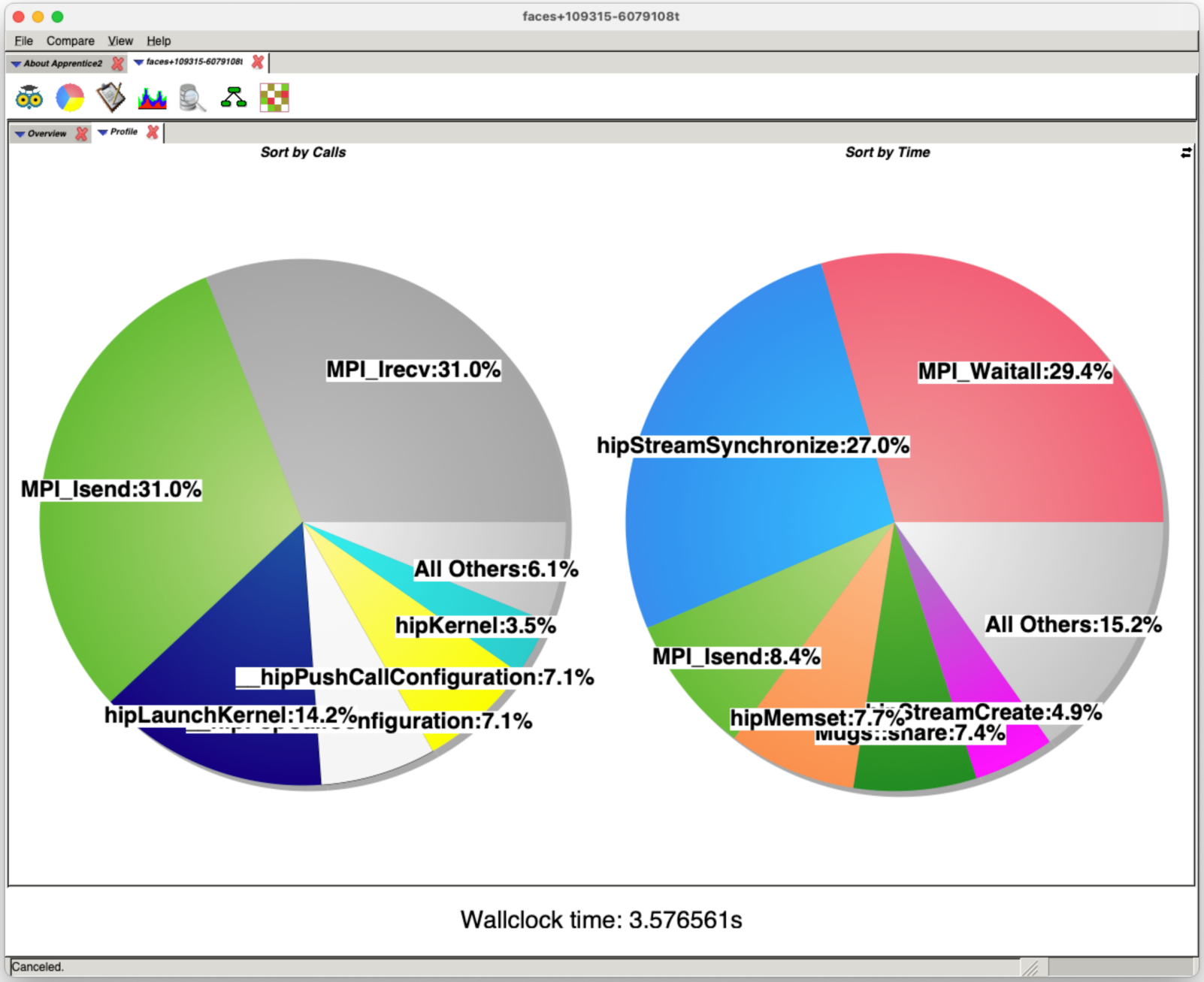


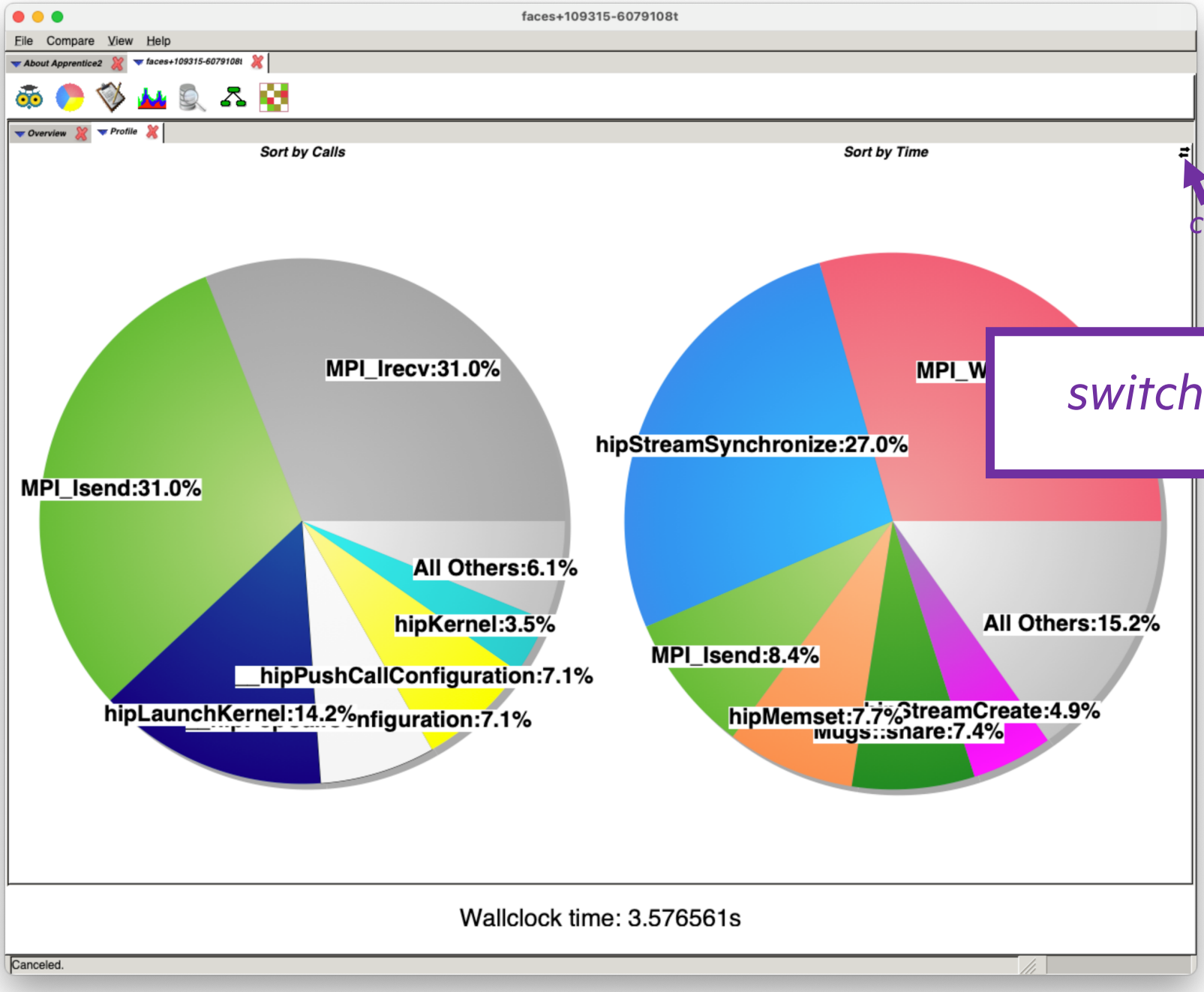
get text copies of interesting results

```
CPU:
10.76% Computation
89.24% Programming Model

GPU:
74.31% HIP_KERNEL
 0.01% HIP_COPY
```

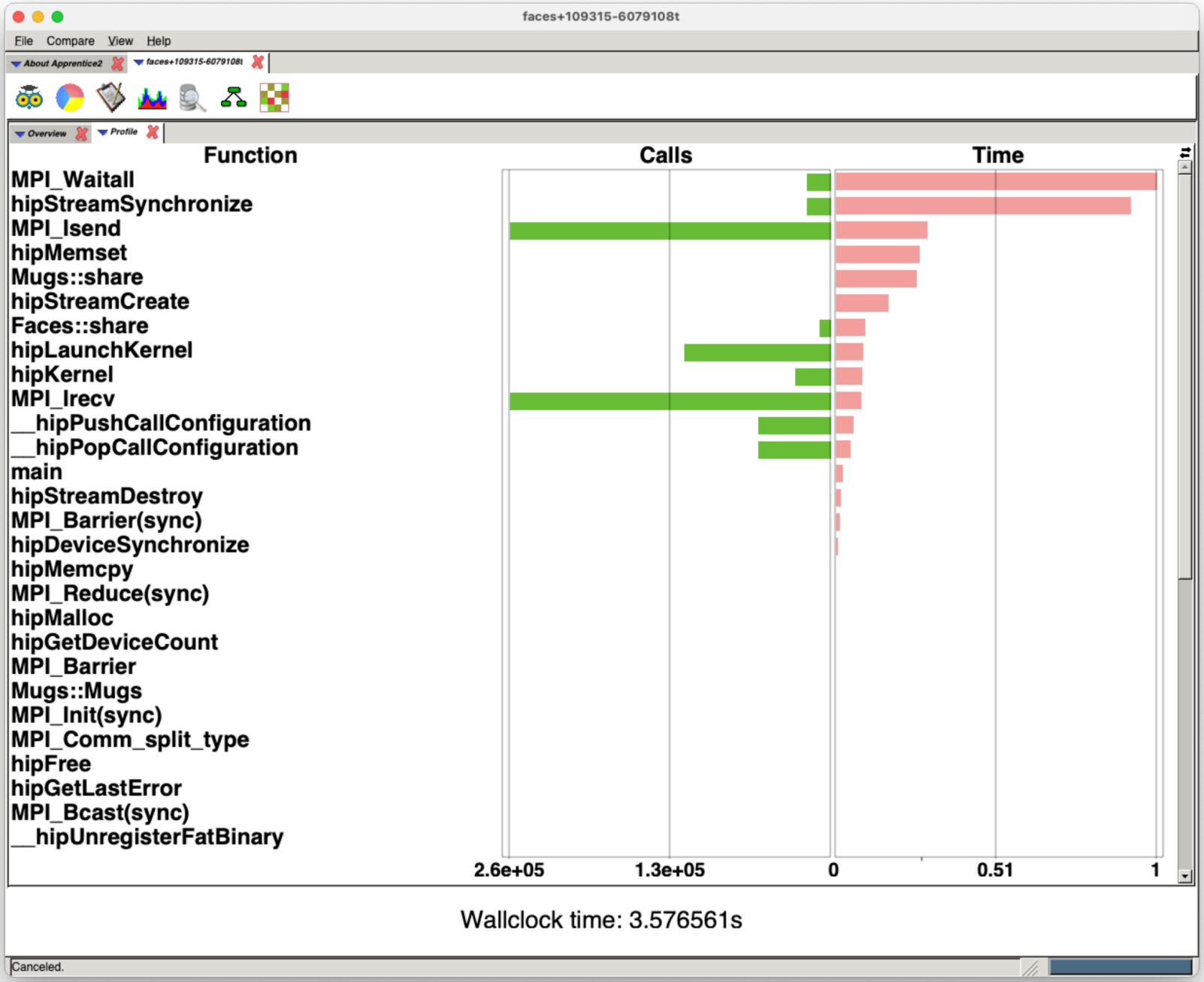

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
click
or click
open profile pie charts
Function/Region Profile
29.4% = MPI_Waitall
27.0% = hipStreamSynchronize
8.4% = MPI_Isend
Prog Model
89.24%
HIP_KERNEL
74.31%
HIP_COPY
0.01%
Memory Utilization
Process HiMem = 314.3 MiBytes
Load Imbalance
(seconds)
0.51s = hipStreamSynchronize
0.21s = MPI_Isend
0.21s = MPI_Waitall
Energy Usage
Node Energy (J) 43,462
10011
10101
10011
10101
Data Movement
MPI Msg = 9.858 GiBytes
MPI Msg Count = 1477… msgs
Acc Copy Out = 36.00…iBytes
Wallclock time: 3.576561s
Canceled.

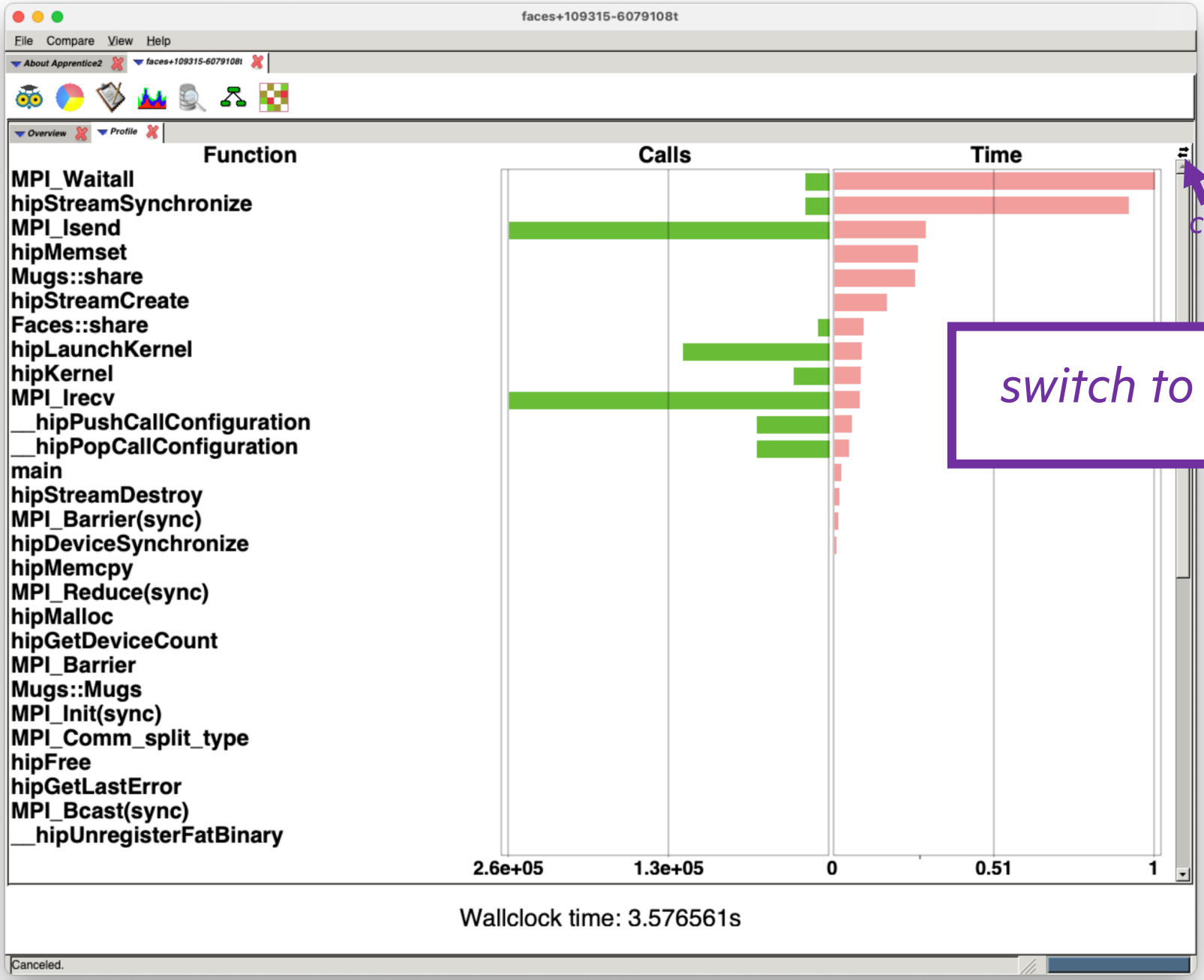

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
Sort by Calls
Sort by Time
MPI_Irecv:31.0%
MPI_Isend:31.0%
All Others:6.1%
hipKernel:3.5%
__hipPushCallConfiguration:7.1%
hipLaunchKernel:14.2%
nfiguration:7.1%
MPI_Waitall:29.4%
hipStreamSynchronize:27.0%
All Others:15.2%
MPI_Isend:8.4%
hipMemset:7.7%
StreamCreate:4.9%
Mugs::share:7.4%
Wallclock time: 3.576561s
Canceled.



click

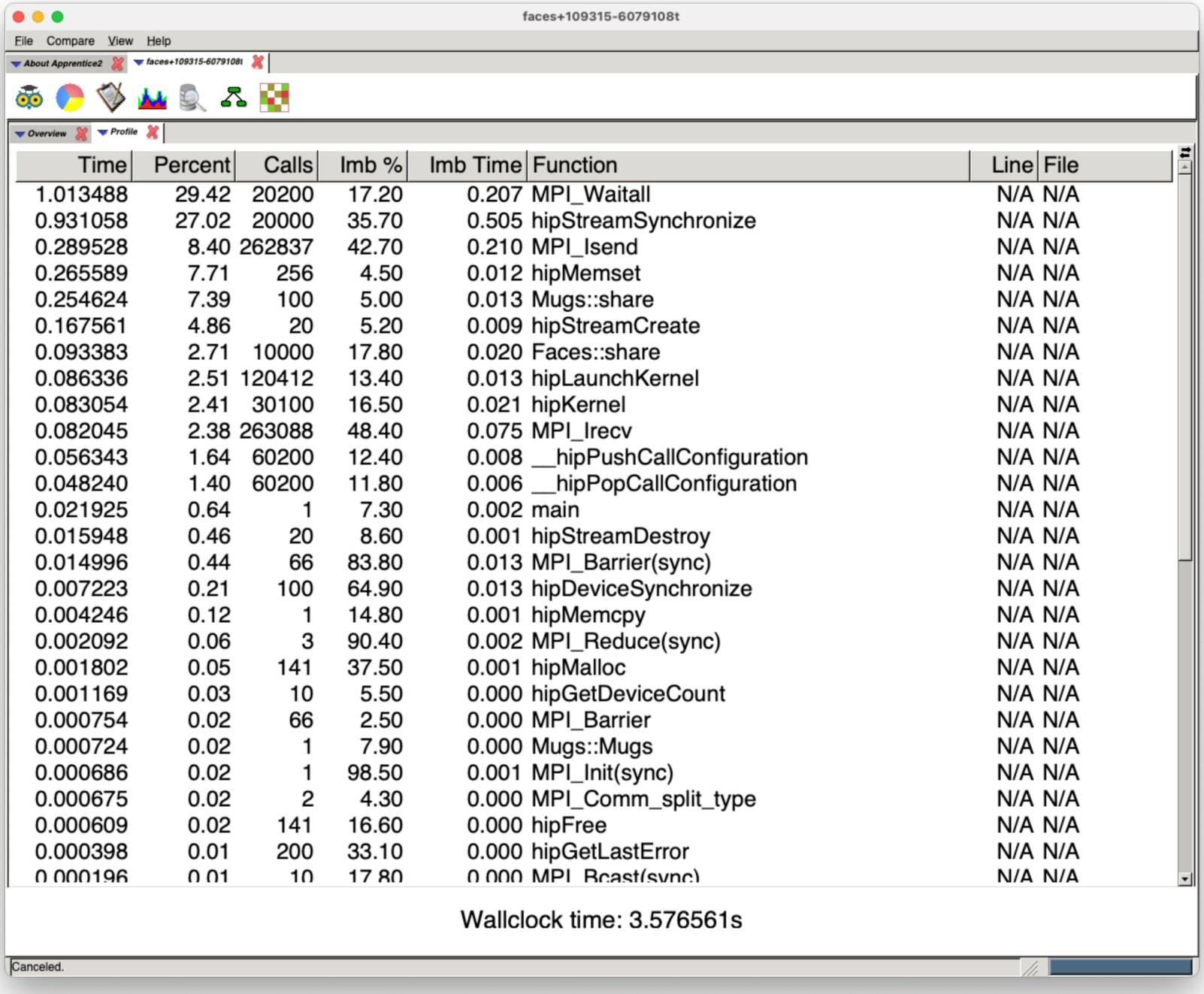
switch to bar chart

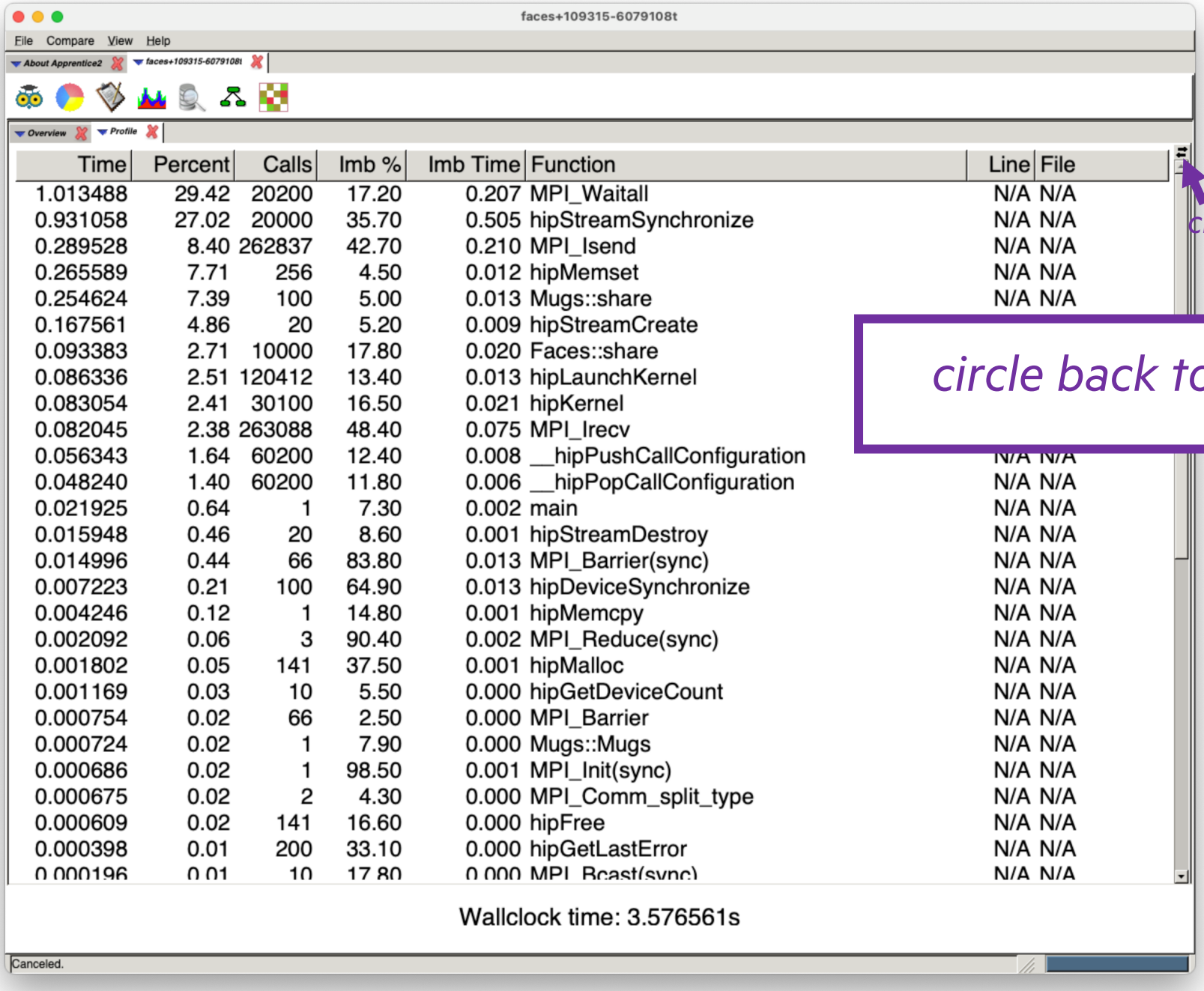

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
Function
Calls
Time
MPI_Waitall
hipStreamSynchronize
MPI_Isend
hipMemset
Mugs::share
hipStreamCreate
Faces::share
hipLaunchKernel
hipKernel
MPI_Irecv
__hipPushCallConfiguration
__hipPopCallConfiguration
main
hipStreamDestroy
MPI_Barrier(sync)
hipDeviceSynchronize
hipMemcpy
MPI_Reduce(sync)
hipMalloc
hipGetDeviceCount
MPI_Barrier
Mugs::Mugs
MPI_Init(sync)
MPI_Comm_split_type
hipFree
hipGetLastError
MPI_Bcast(sync)
__hipUnregisterFatBinary
2.6e+05
1.3e+05
0
0.51
1
Wallclock time: 3.576561s
Canceled.



click

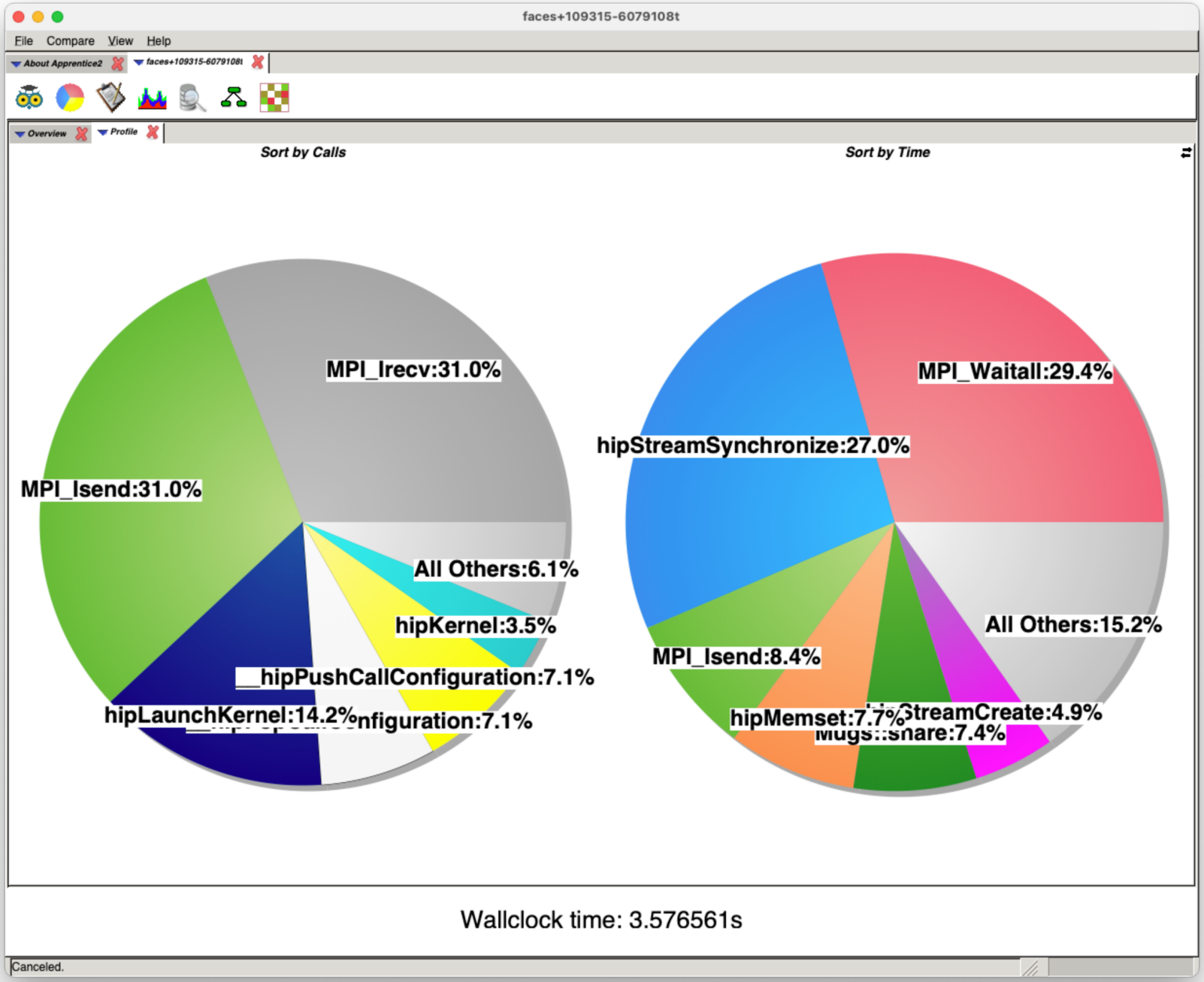
switch to sortable table

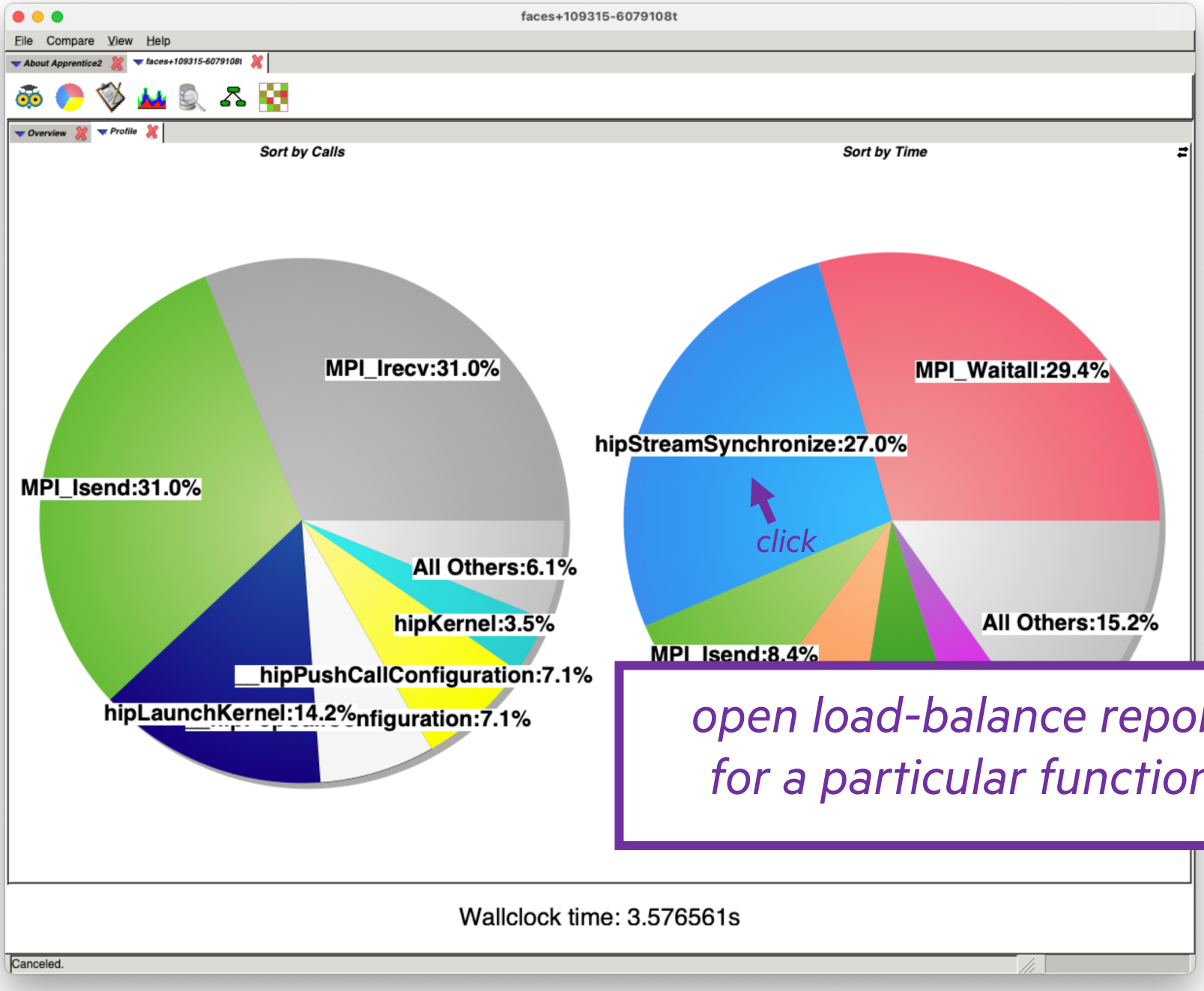

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
Time Percent Calls Imb % Imb Time Function Line File
1.013488 29.42 20200 17.20 0.207 MPI_Waitall N/A N/A
0.931058 27.02 20000 35.70 0.505 hipStreamSynchronize N/A N/A
0.289528 8.40 262837 42.70 0.210 MPI_Isend N/A N/A
0.265589 7.71 256 4.50 0.012 hipMemset N/A N/A
0.254624 7.39 100 5.00 0.013 Mugs::share N/A N/A
0.167561 4.86 20 5.20 0.009 hipStreamCreate N/A N/A
0.093383 2.71 10000 17.80 0.020 Faces::share N/A N/A
0.086336 2.51 120412 13.40 0.013 hipLaunchKernel N/A N/A
0.083054 2.41 30100 16.50 0.021 hipKernel N/A N/A
0.082045 2.38 263088 48.40 0.075 MPI_Irecv N/A N/A
0.056343 1.64 60200 12.40 0.008 __hipPushCallConfiguration N/A N/A
0.048240 1.40 60200 11.80 0.006 __hipPopCallConfiguration N/A N/A
0.021925 0.64 1 7.30 0.002 main N/A N/A
0.015948 0.46 20 8.60 0.001 hipStreamDestroy N/A N/A
0.014996 0.44 66 83.80 0.013 MPI_Barrier(sync) N/A N/A
0.007223 0.21 100 64.90 0.013 hipDeviceSynchronize N/A N/A
0.004246 0.12 1 14.80 0.001 hipMemcpy N/A N/A
0.002092 0.06 3 90.40 0.002 MPI_Reduce(sync) N/A N/A
0.001802 0.05 141 37.50 0.001 hipMalloc N/A N/A
0.001169 0.03 10 5.50 0.000 hipGetDeviceCount N/A N/A
0.000754 0.02 66 2.50 0.000 MPI_Barrier N/A N/A
0.000724 0.02 1 7.90 0.000 Mugs::Mugs N/A N/A
0.000686 0.02 1 98.50 0.001 MPI_Init(sync) N/A N/A
0.000675 0.02 2 4.30 0.000 MPI_Comm_split_type N/A N/A
0.000609 0.02 141 16.60 0.000 hipFree N/A N/A
0.000398 0.01 200 33.10 0.000 hipGetLastError N/A N/A
0.000196 0.01 10 17.80 0.000 MPI_Bcast(sync) N/A N/A
Wallclock time: 3.576561s
Canceled.



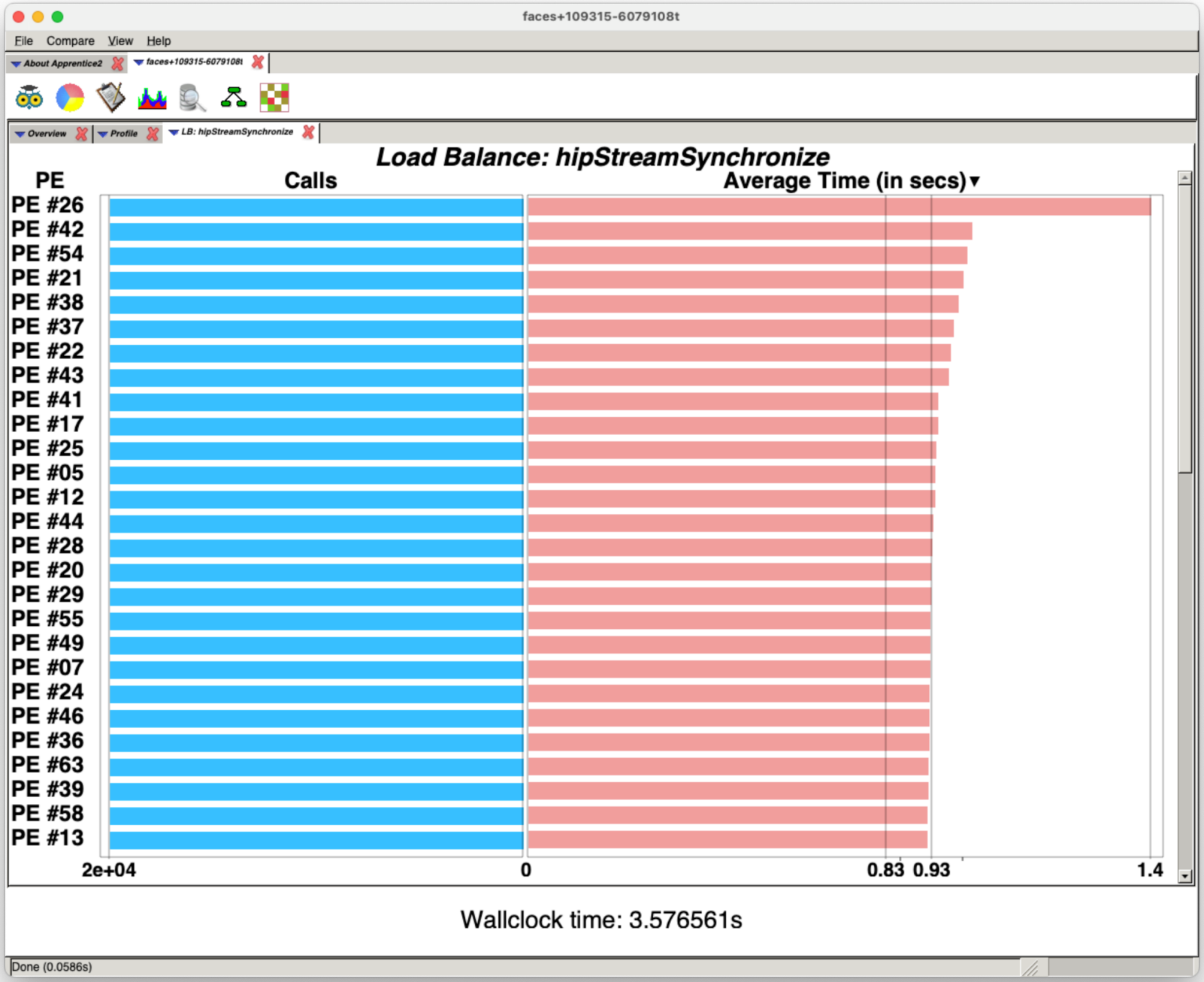
click

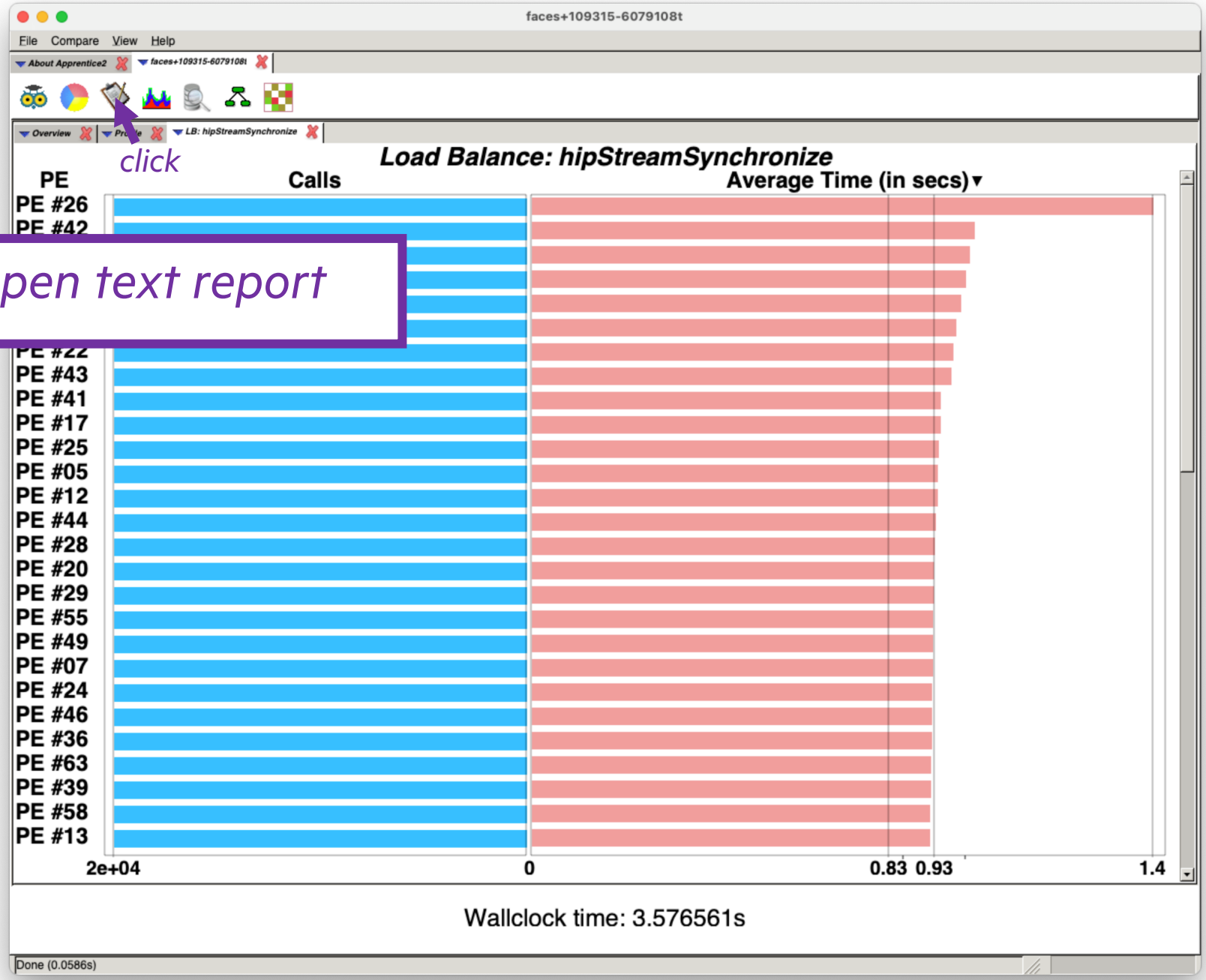
circle back to pie charts


faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
Sort by Calls
MPI_Irecv:31.0%
MPI_Isend:31.0%
All Others:6.1%
hipKernel:3.5%
__hipPushCallConfiguration:7.1%
hipLaunchKernel:14.2%
nfiguration:7.1%
Sort by Time
MPI_Waitall:29.4%
hipStreamSynchronize:27.0%
All Others:15.2%
MPI_Isend:8.4%
hipMemset:7.7%
StreamCreate:4.9%
Mugs::share:7.4%
Wallclock time: 3.576561s
Canceled.

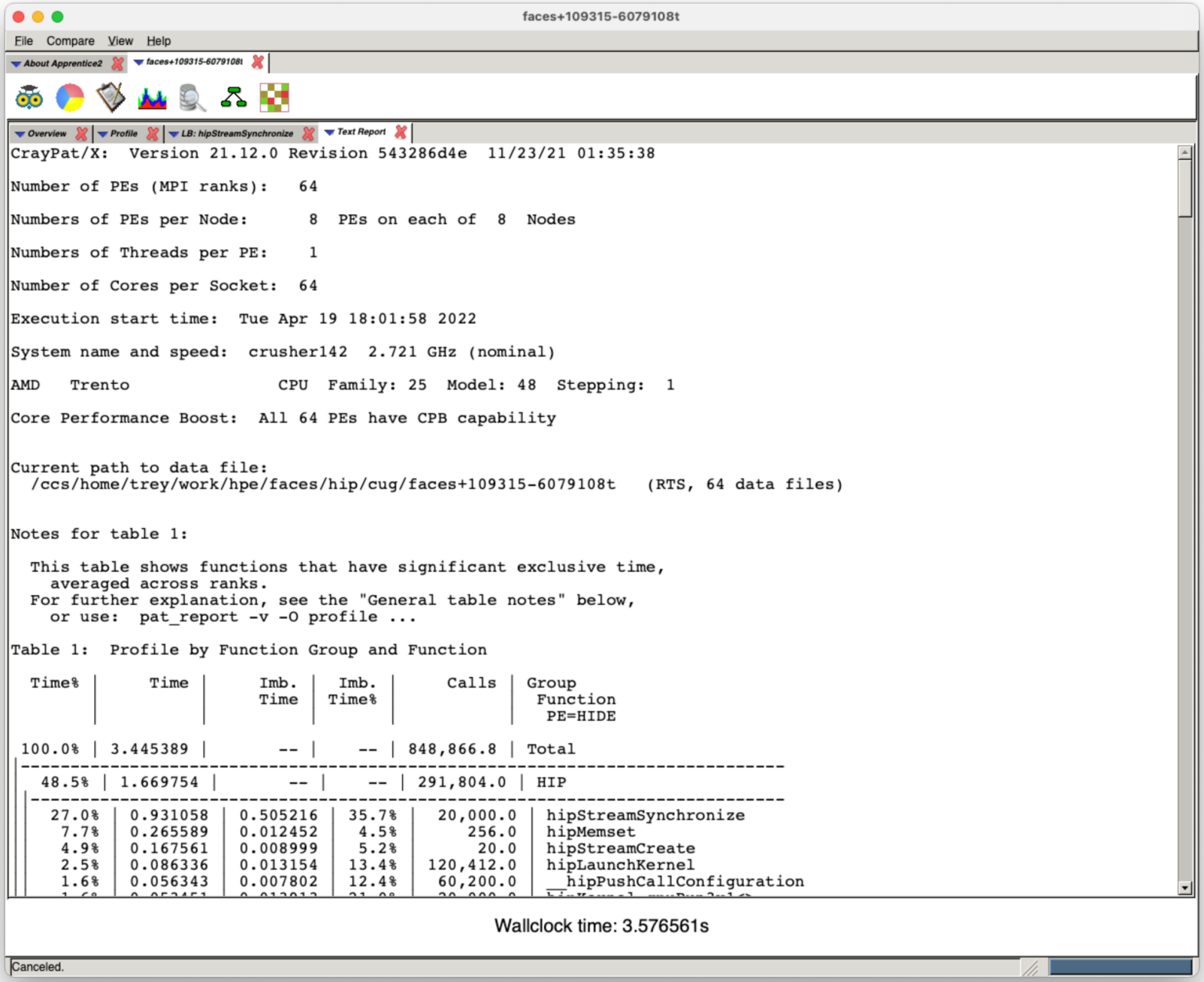


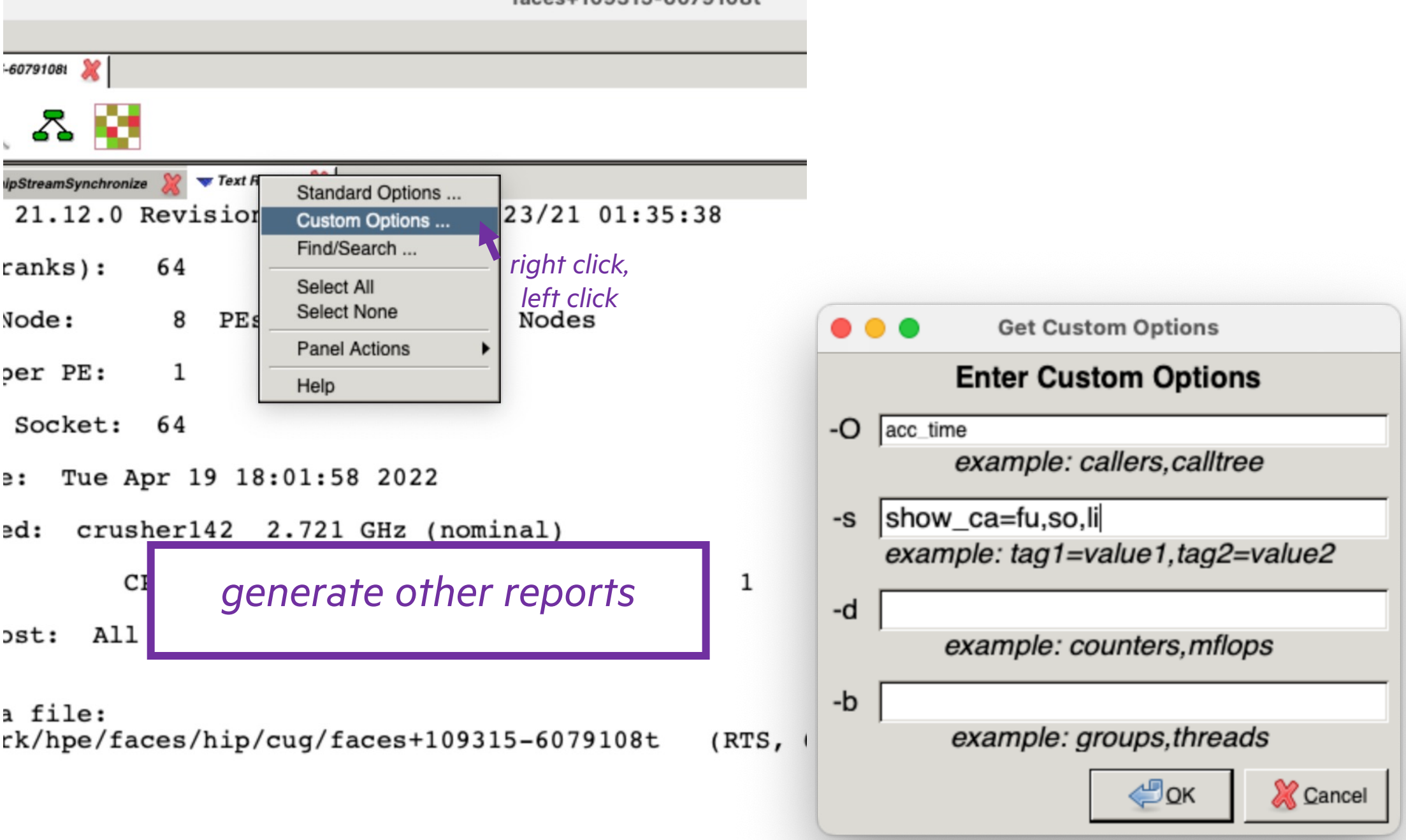
open load-balance report for a particular function

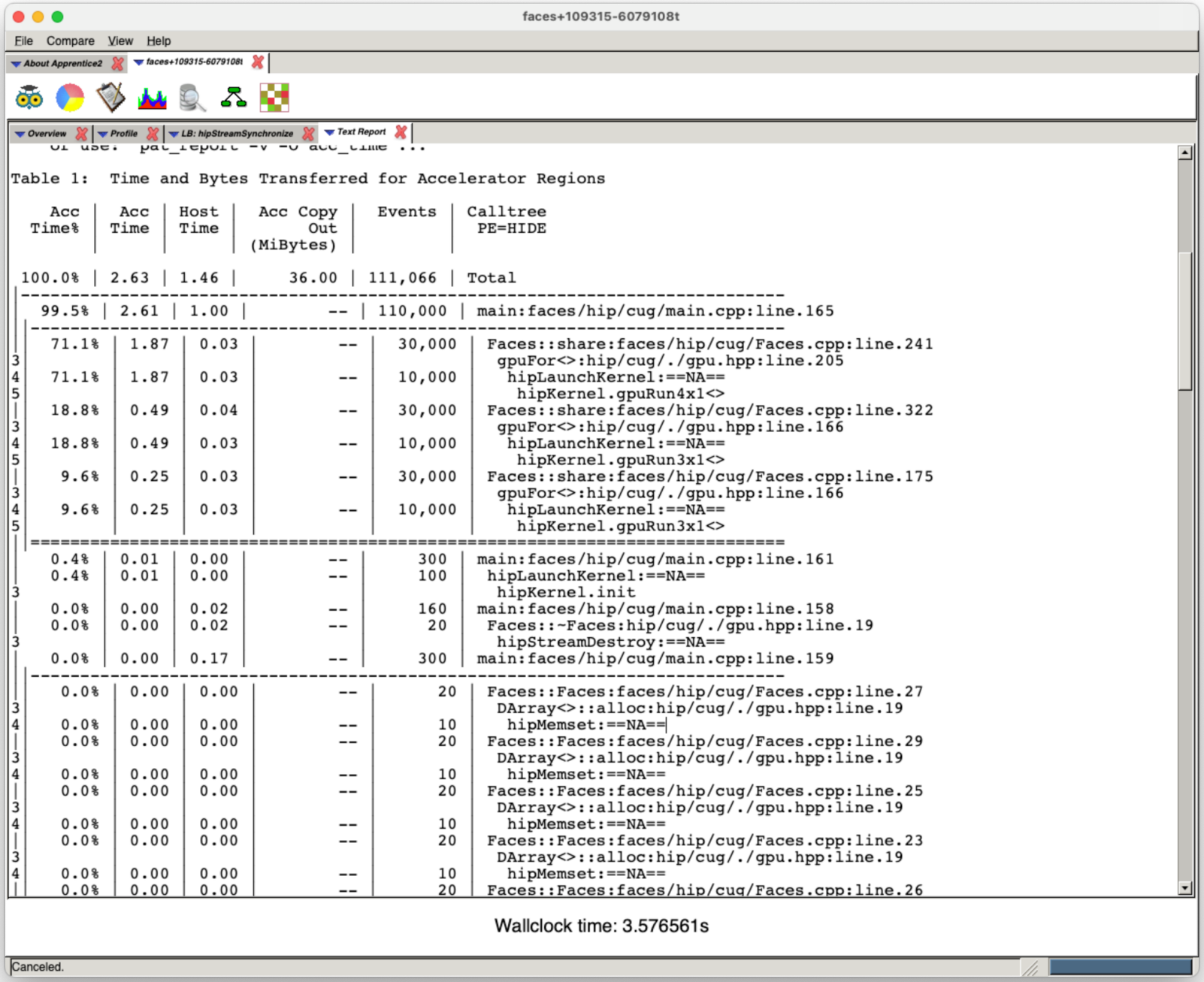

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
LB: hipStreamSynchronize
Load Balance: hipStreamSynchronize
PE
Calls
Average Time (in secs)
PE #26
PE #42
PE #54
PE #21
PE #38
PE #37
PE #22
PE #43
PE #41
PE #17
PE #25
PE #05
PE #12
PE #44
PE #28
PE #20
PE #29
PE #55
PE #49
PE #07
PE #24
PE #46
PE #36
PE #63
PE #39
PE #58
PE #13
2e+04
0
0.83
0.93
1.4
Wallclock time: 3.576561s
Done (0.0586s)

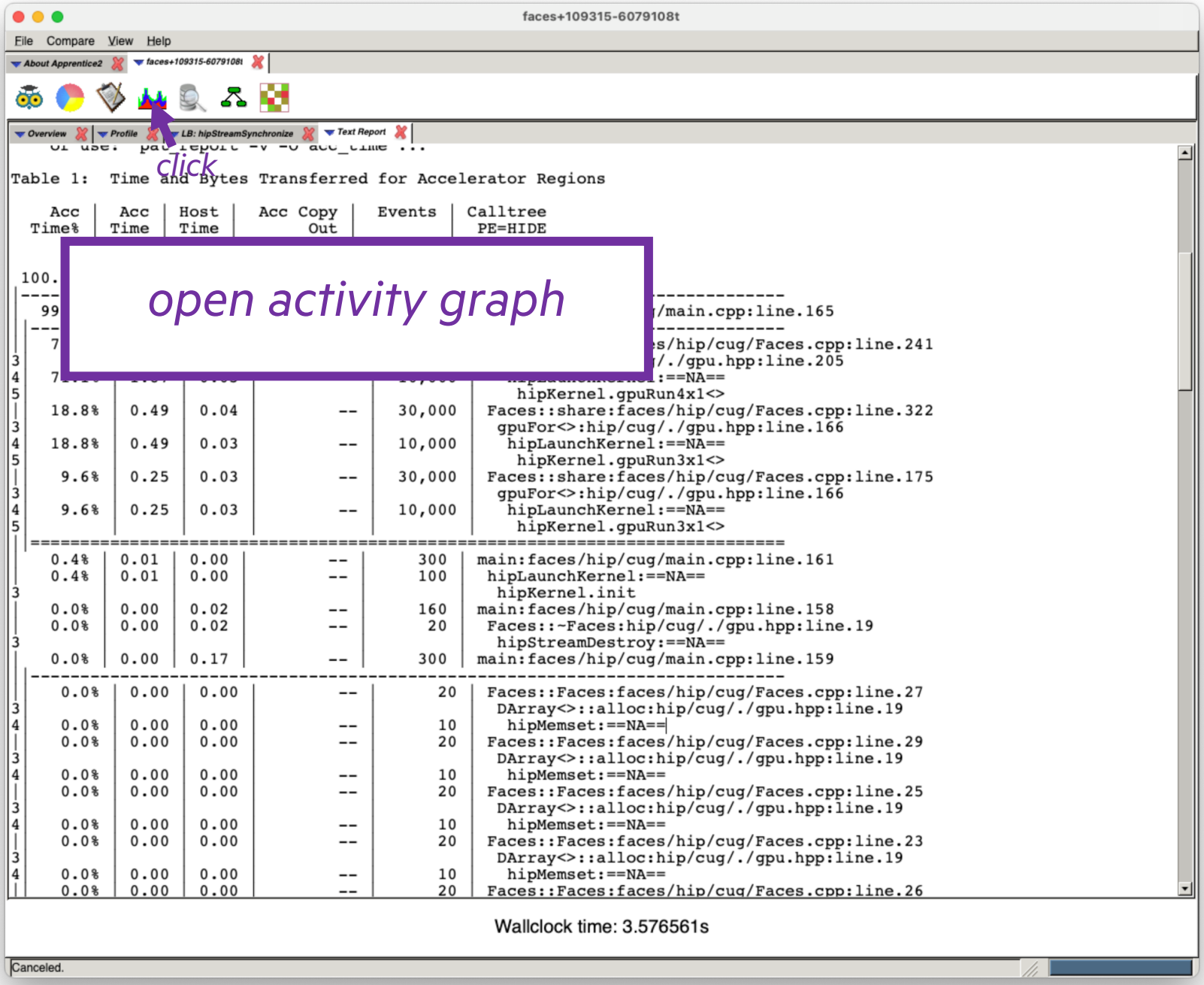


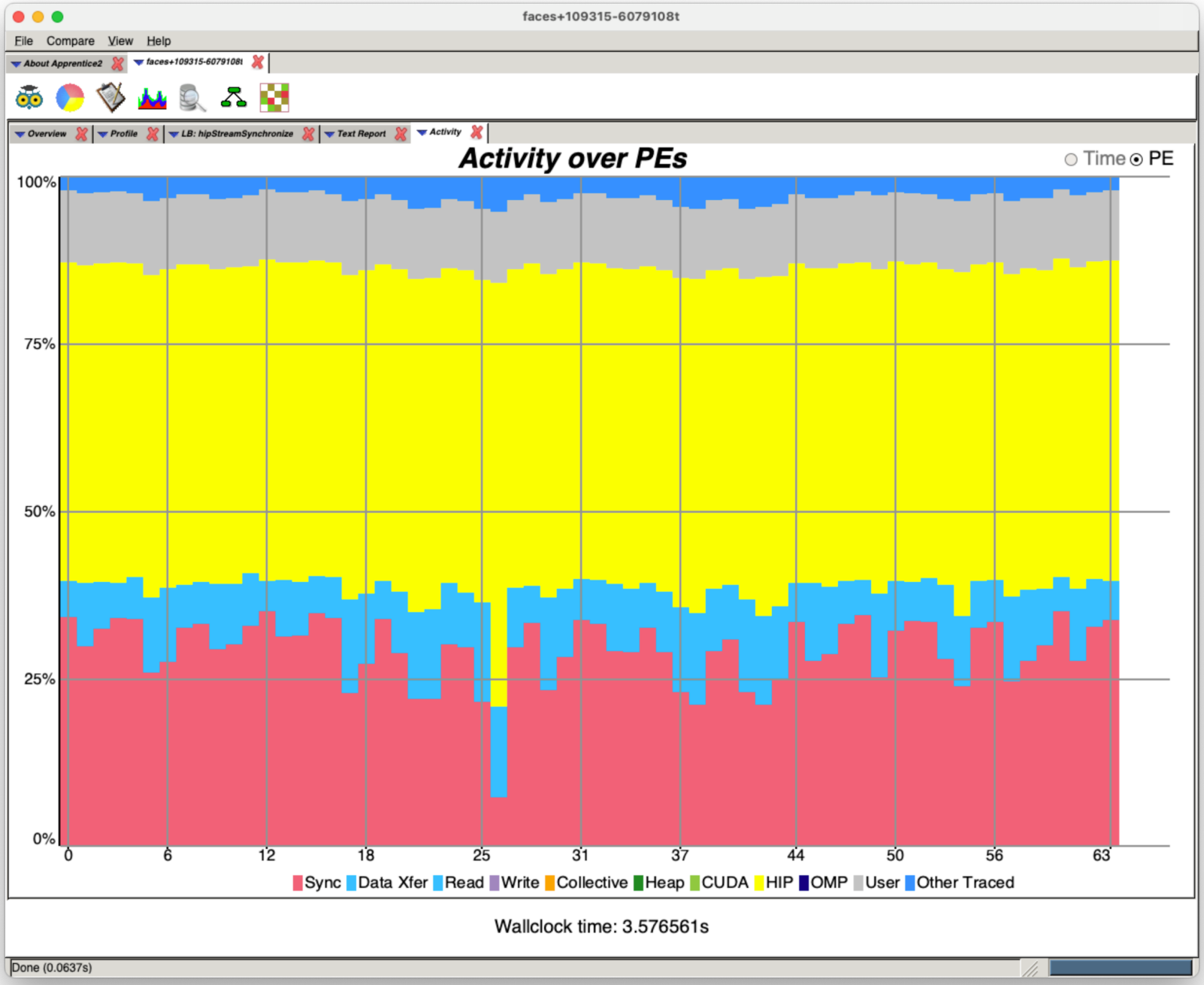
open text report

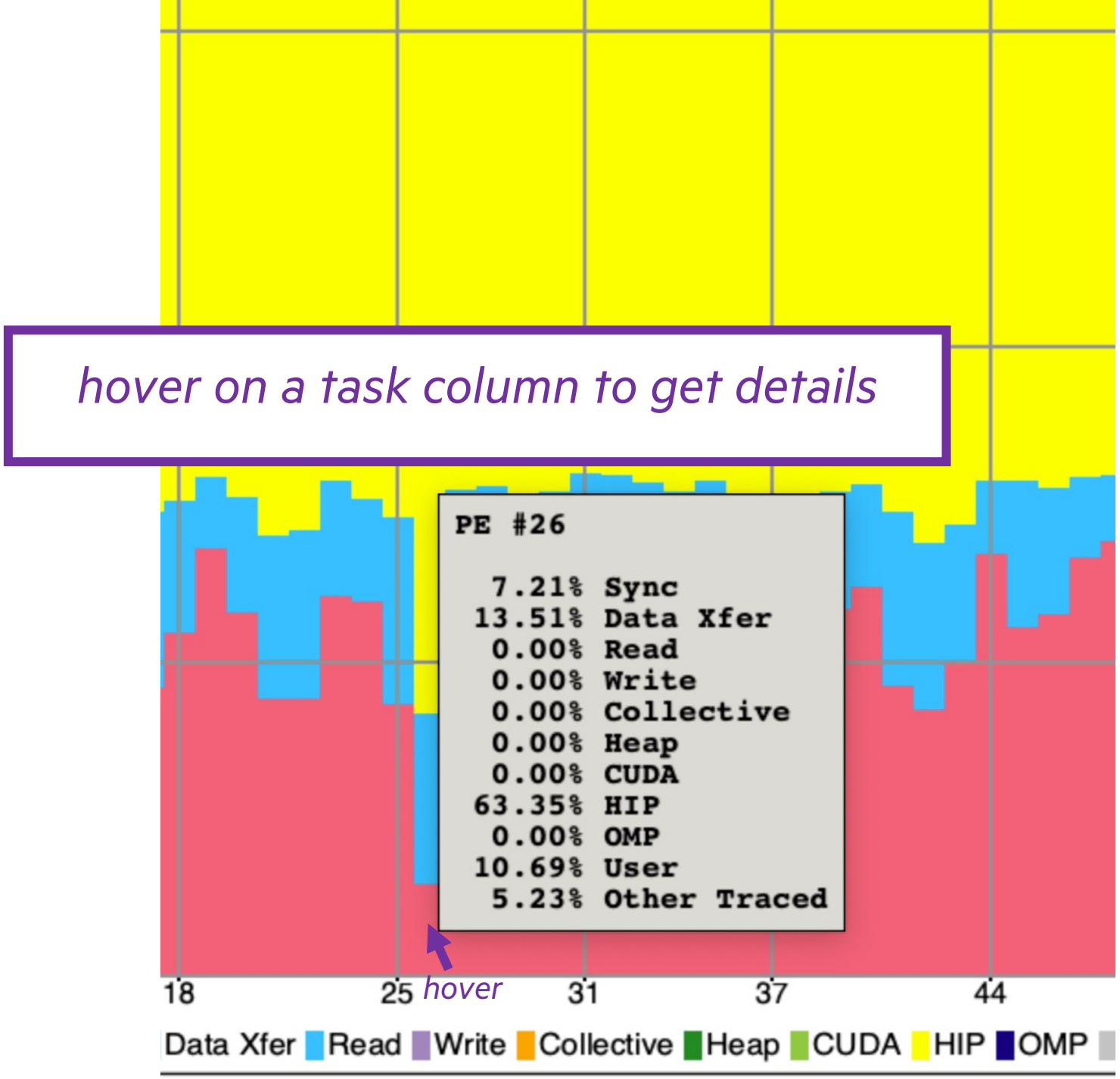

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
LB: hipStreamSynchronize
Text Report
CrayPat/X: Version 21.12.0 Revision 543286d4e 11/23/21 01:35:38
Number of PEs (MPI ranks): 64
Numbers of PEs per Node: 8 PEs on each of 8 Nodes
Numbers of Threads per PE: 1
Number of Cores per Socket: 64
Execution start time: Tue Apr 19 18:01:58 2022
System name and speed: crusher142 2.721 GHz (nominal)
AMD Trento CPU Family: 25 Model: 48 Stepping: 1
Core Performance Boost: All 64 PEs have CPB capability
Current path to data file:
/ccs/home/trey/work/hpe/faces/hip/cug/faces+109315-6079108t (RTS, 64 data files)
Notes for table 1:
This table shows functions that have significant exclusive time,
averaged across ranks.
For further explanation, see the "General table notes" below,
or use: pat_report -v -O profile ...
Table 1: Profile by Function Group and Function
Time% | Time | Imb. Time | Imb. Time% | Calls | Group Function PE=HIDE
100.0% | 3.445389 | -- | -- | 848,866.8 | Total
48.5% | 1.669754 | -- | -- | 291,804.0 | HIP
27.0% | 0.931058 | 0.505216 | 35.7% | 20,000.0 | hipStreamSynchronize
7.7% | 0.265589 | 0.012452 | 4.5% | 256.0 | hipMemset
4.9% | 0.167561 | 0.008999 | 5.2% | 20.0 | hipStreamCreate
2.5% | 0.086336 | 0.013154 | 13.4% | 120,412.0 | hipLaunchKernel
1.6% | 0.056343 | 0.007802 | 12.4% | 60,200.0 | __hipPushCallConfiguration
Wallclock time: 3.576561s
Canceled.


Standard Options ...
Custom Options ...
Find/Search ...
Select All
Select None
Panel Actions
Help
right click,
left click
21.12.0 Revisio
23/21 01:35:38
ranks): 64
Node: 8 PEs
Nodes
per PE: 1
Socket: 64
e: Tue Apr 19 18:01:58 2022
ed: crusher142 2.721 GHz (nominal)
generate other reports
1
ost: All
a file:
rk/hpe/faces/hip/cug/faces+109315-6079108t (RTS,
functions that have significant exclusive time,
ranks.
Get Custom Options
Enter Custom Options
-O acc_time
example: callers,calltree
-s show_ca=fu,so,li
example: tag1=value1,tag2=value2
-d
example: counters,mflops
-b
example: groups,threads
OK
Cancel

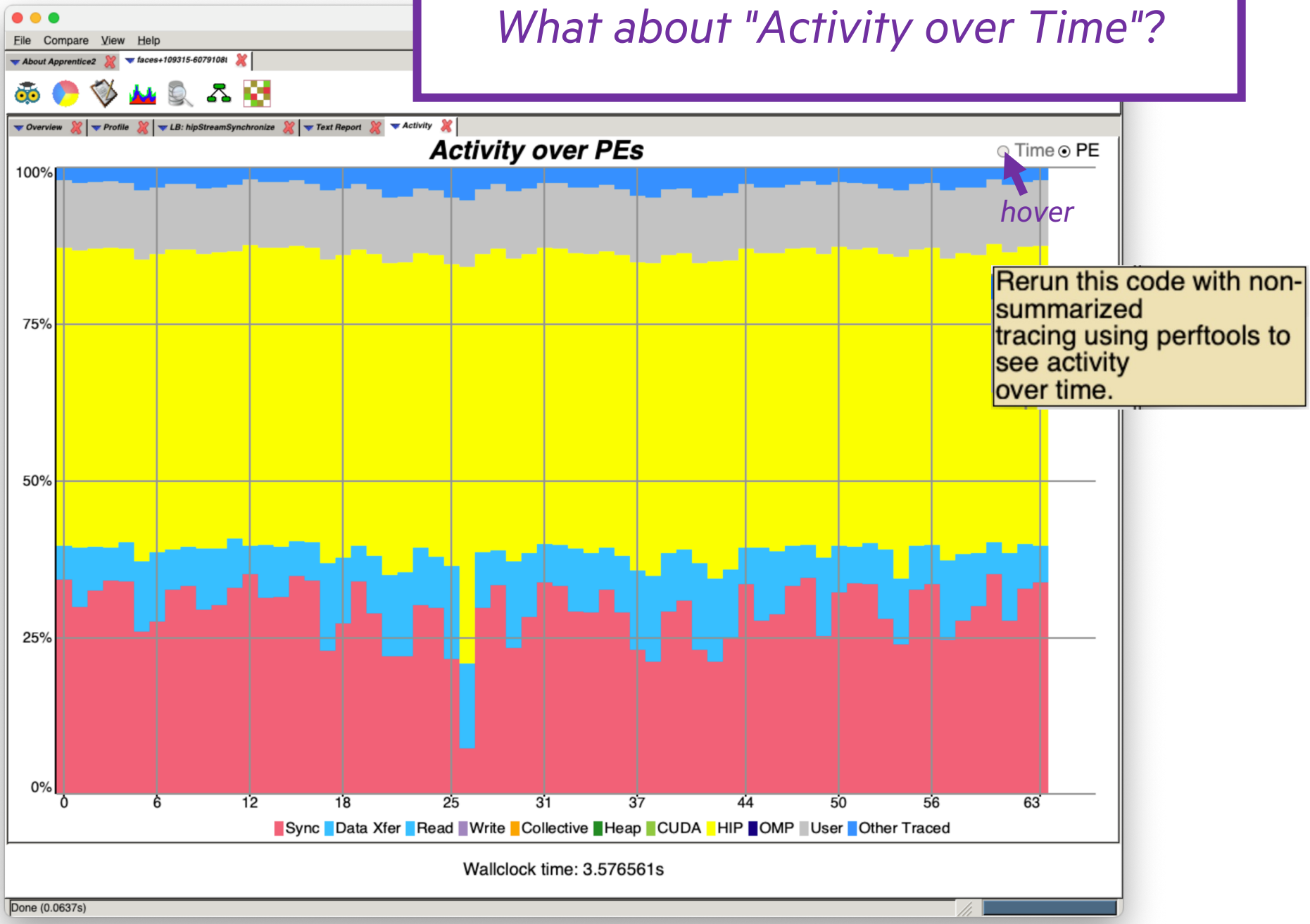

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
LB: hipStreamSynchronize
Text Report
Table 1: Time and Bytes Transferred for Accelerator Regions
Acc Time% | Acc Time | Host Time | Acc Copy Out (MiBytes) | Events | Calltree PE=HIDE
100.0% | 2.63 | 1.46 | 36.00 | 111,066 | Total
99.5% | 2.61 | 1.00 | -- | 110,000 | main:faces/hip/cug/main.cpp:line.165
71.1% | 1.87 | 0.03 | -- | 30,000 | Faces::share:faces/hip/cug/Faces.cpp:line.241
gpuFor<>:hip/cug/./gpu.hpp:line.205
71.1% | 1.87 | 0.03 | -- | 10,000 | hipLaunchKernel:==NA==
hipKernel.gpuRun4x1<>
18.8% | 0.49 | 0.04 | -- | 30,000 | Faces::share:faces/hip/cug/Faces.cpp:line.322
gpuFor<>:hip/cug/./gpu.hpp:line.166
18.8% | 0.49 | 0.03 | -- | 10,000 | hipLaunchKernel:==NA==
hipKernel.gpuRun3x1<>
9.6% | 0.25 | 0.03 | -- | 30,000 | Faces::share:faces/hip/cug/Faces.cpp:line.175
gpuFor<>:hip/cug/./gpu.hpp:line.166
9.6% | 0.25 | 0.03 | -- | 10,000 | hipLaunchKernel:==NA==
hipKernel.gpuRun3x1<>
0.4% | 0.01 | 0.00 | -- | 300 | main:faces/hip/cug/main.cpp:line.161
0.4% | 0.01 | 0.00 | -- | 100 | hipLaunchKernel:==NA==
hipKernel.init
0.0% | 0.00 | 0.02 | -- | 160 | main:faces/hip/cug/main.cpp:line.158
0.0% | 0.00 | 0.02 | -- | 20 | Faces::~Faces:hip/cug/./gpu.hpp:line.19
hipStreamDestroy:==NA==
0.0% | 0.00 | 0.17 | -- | 300 | main:faces/hip/cug/main.cpp:line.159
0.0% | 0.00 | 0.00 | -- | 20 | Faces::Faces:faces/hip/cug/Faces.cpp:line.27
DArray<>::alloc:hip/cug/./gpu.hpp:line.19
0.0% | 0.00 | 0.00 | -- | 10 | hipMemset:==NA==
0.0% | 0.00 | 0.00 | -- | 20 | Faces::Faces:faces/hip/cug/Faces.cpp:line.29
DArray<>::alloc:hip/cug/./gpu.hpp:line.19
0.0% | 0.00 | 0.00 | -- | 10 | hipMemset:==NA==
0.0% | 0.00 | 0.00 | -- | 20 | Faces::Faces:faces/hip/cug/Faces.cpp:line.25
DArray<>::alloc:hip/cug/./gpu.hpp:line.19
0.0% | 0.00 | 0.00 | -- | 10 | hipMemset:==NA==
0.0% | 0.00 | 0.00 | -- | 20 | Faces::Faces:faces/hip/cug/Faces.cpp:line.23
DArray<>::alloc:hip/cug/./gpu.hpp:line.19
0.0% | 0.00 | 0.00 | -- | 10 | hipMemset:==NA==
0.0% | 0.00 | 0.00 | -- | 20 | Faces::Faces:faces/hip/cug/Faces.cpp:line.26
Wallclock time: 3.576561s
Canceled.

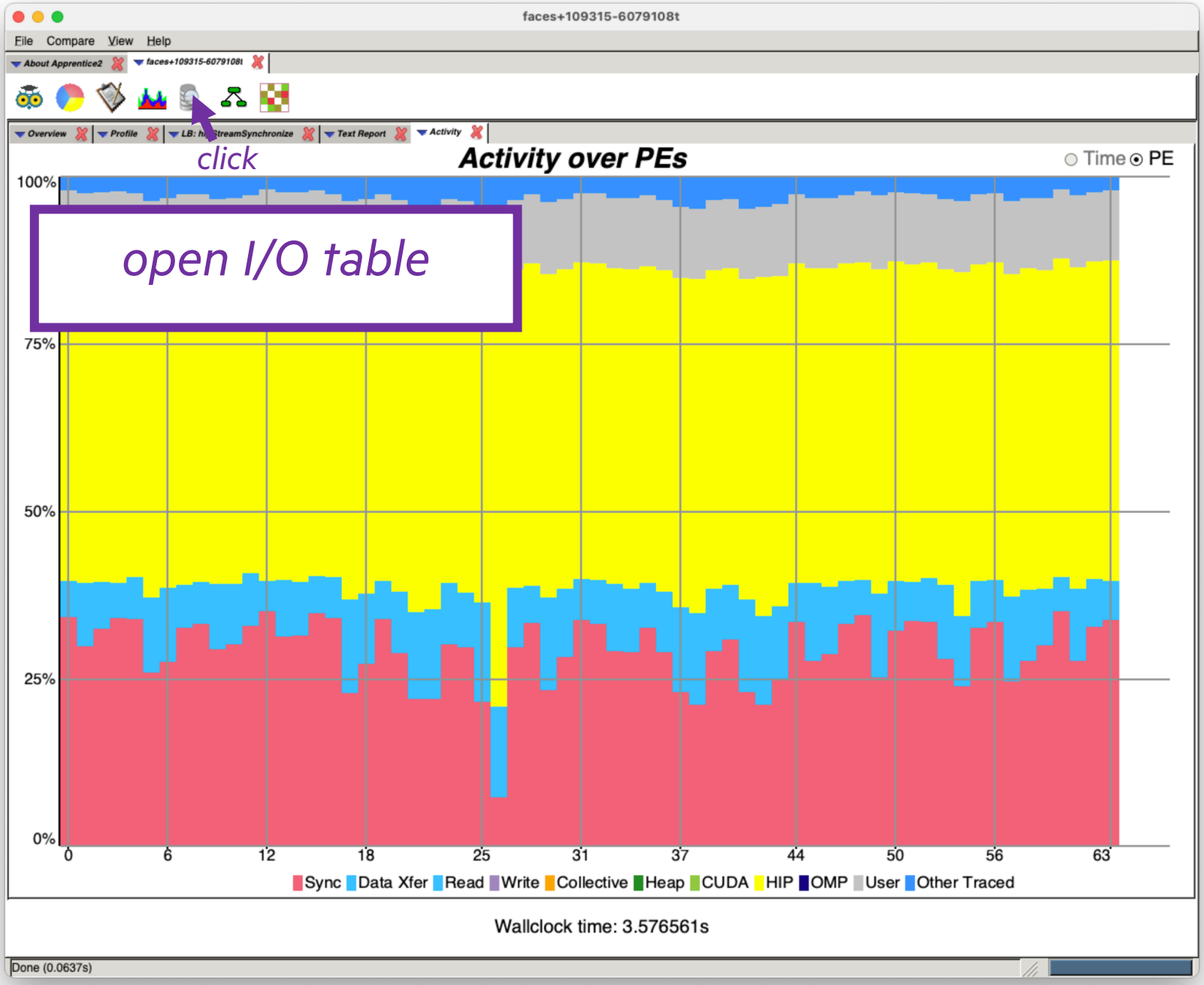

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
LB: hipStreamSynchronize
Text Report
click
open activity graph
Table 1: Time and Bytes Transferred for Accelerator Regions
Acc Time% | Acc Time | Host Time | Acc Copy Out | Events | Calltree PE=HIDE
18.8% | 0.49 | 0.04 | -- | 30,000 | Faces::share:faces/hip/cug/Faces.cpp:line.322
gpuFor<>:hip/cug/./gpu.hpp:line.166
18.8% | 0.49 | 0.03 | -- | 10,000 | hipLaunchKernel:==NA==
hipKernel.gpuRun3x1<>
9.6% | 0.25 | 0.03 | -- | 30,000 | Faces::share:faces/hip/cug/Faces.cpp:line.175
gpuFor<>:hip/cug/./gpu.hpp:line.166
9.6% | 0.25 | 0.03 | -- | 10,000 | hipLaunchKernel:==NA==
hipKernel.gpuRun3x1<>
0.4% | 0.01 | 0.00 | -- | 300 | main:faces/hip/cug/main.cpp:line.161
0.4% | 0.01 | 0.00 | -- | 100 | hipLaunchKernel:==NA==
hipKernel.init
0.0% | 0.00 | 0.02 | -- | 160 | main:faces/hip/cug/main.cpp:line.158
0.0% | 0.00 | 0.02 | -- | 20 | Faces::~Faces:hip/cug/./gpu.hpp:line.19
hipStreamDestroy:==NA==
0.0% | 0.00 | 0.17 | -- | 300 | main:faces/hip/cug/main.cpp:line.159
0.0% | 0.00 | 0.00 | -- | 20 | Faces::Faces:faces/hip/cug/Faces.cpp:line.27
DArray<>::alloc:hip/cug/./gpu.hpp:line.19
0.0% | 0.00 | 0.00 | -- | 10 | hipMemset:==NA==
0.0% | 0.00 | 0.00 | -- | 20 | Faces::Faces:faces/hip/cug/Faces.cpp:line.29
DArray<>::alloc:hip/cug/./gpu.hpp:line.19
0.0% | 0.00 | 0.00 | -- | 10 | hipMemset:==NA==
0.0% | 0.00 | 0.00 | -- | 20 | Faces::Faces:faces/hip/cug/Faces.cpp:line.25
DArray<>::alloc:hip/cug/./gpu.hpp:line.19
0.0% | 0.00 | 0.00 | -- | 10 | hipMemset:==NA==
0.0% | 0.00 | 0.00 | -- | 20 | Faces::Faces:faces/hip/cug/Faces.cpp:line.23
DArray<>::alloc:hip/cug/./gpu.hpp:line.19
0.0% | 0.00 | 0.00 | -- | 10 | hipMemset:==NA==
0.0% | 0.00 | 0.00 | -- | 20 | Faces::Faces:faces/hip/cug/Faces.cpp:line.26
Wallclock time: 3.576561s
Canceled.

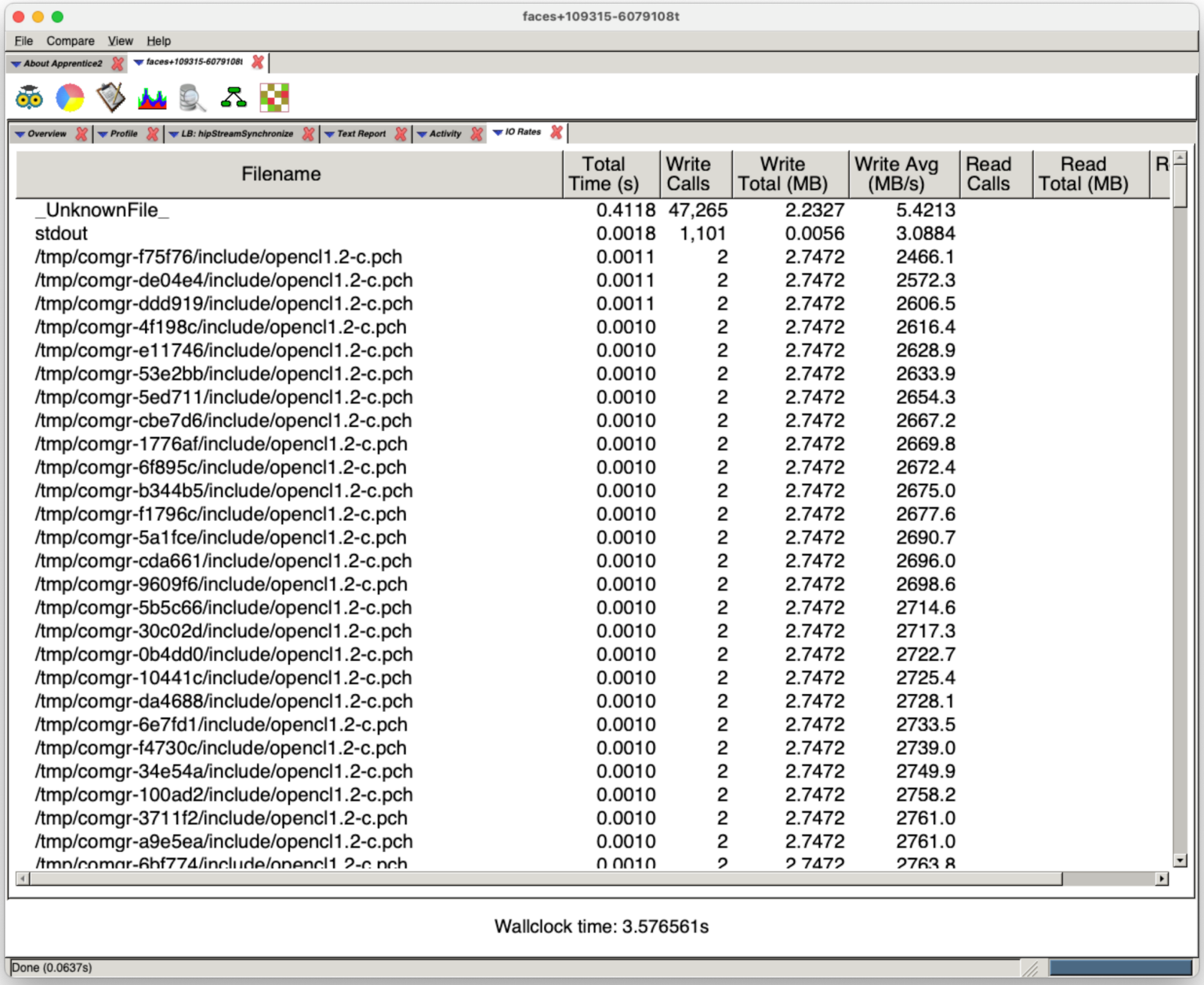

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
LB: hipStreamSynchronize
Text Report
Activity
Activity over PEs
Time
PE
100%
75%
50%
25%
0%
0 6 12 18 25 31 37 44 50 56 63
Sync Data Xfer Read Write Collective Heap CUDA HIP OMP User Other Traced
Wallclock time: 3.576561s
Done (0.0637s)

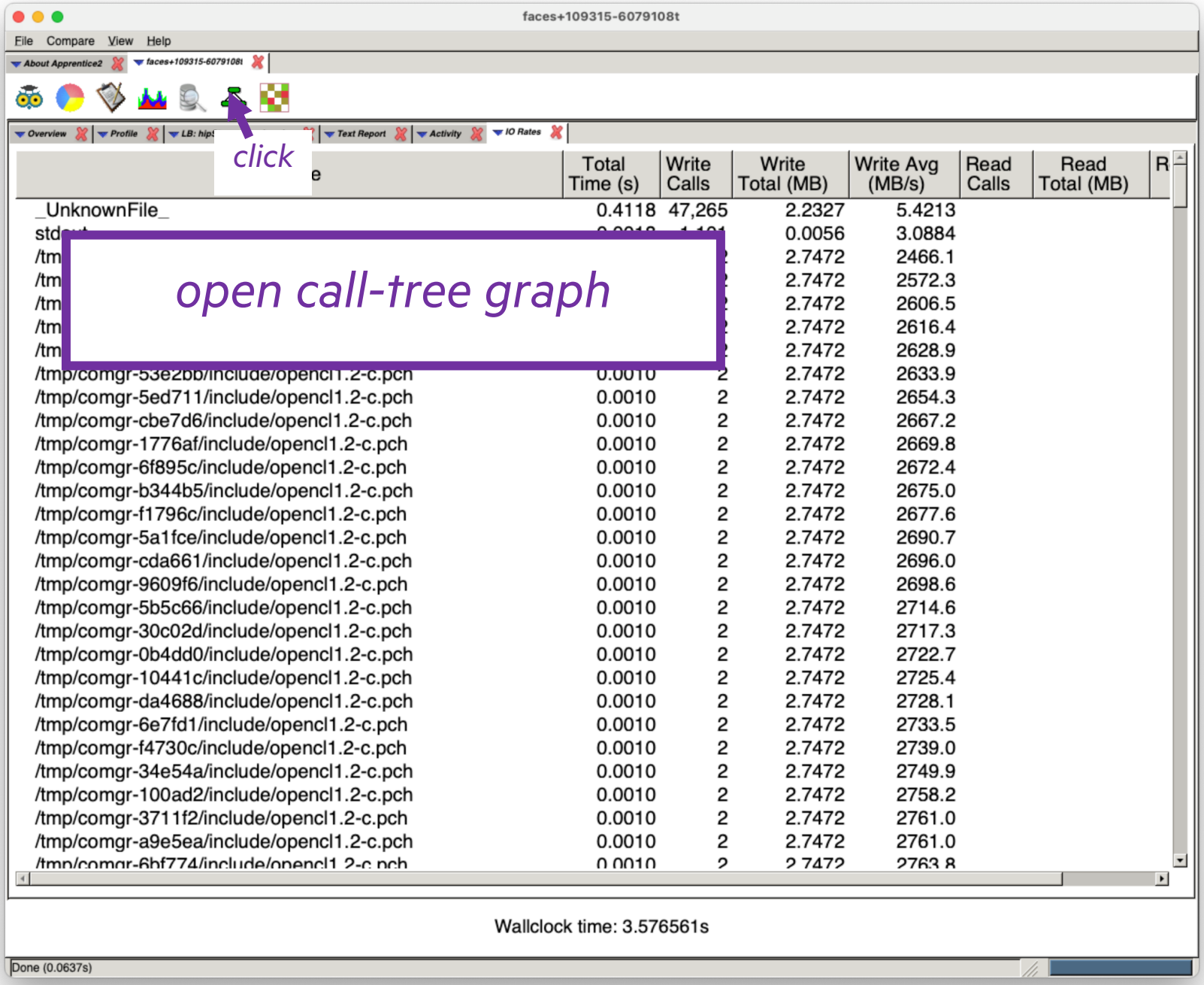

hover on a task column to get details
PE #26
7.21% Sync
13.51% Data Xfer
0.00% Read
0.00% Write
0.00% Collective
0.00% Heap
0.00% CUDA
63.35% HIP
0.00% OMP
10.69% User
5.23% Other Traced
hover
18
25
31
37
44
Data Xfer
Read
Write
Collective
Heap
CUDA
HIP
OMP

What about "Activity over Time"?

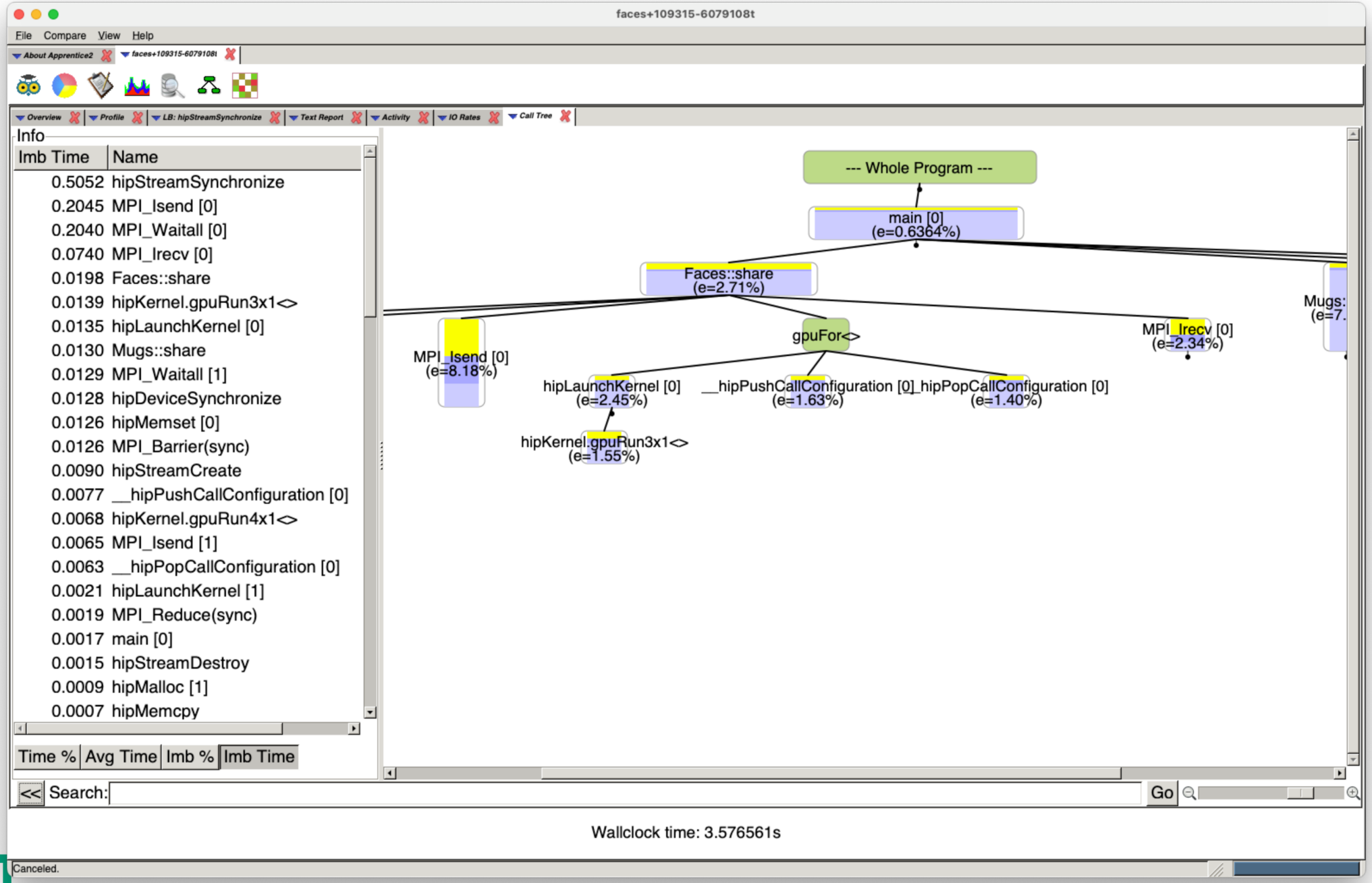


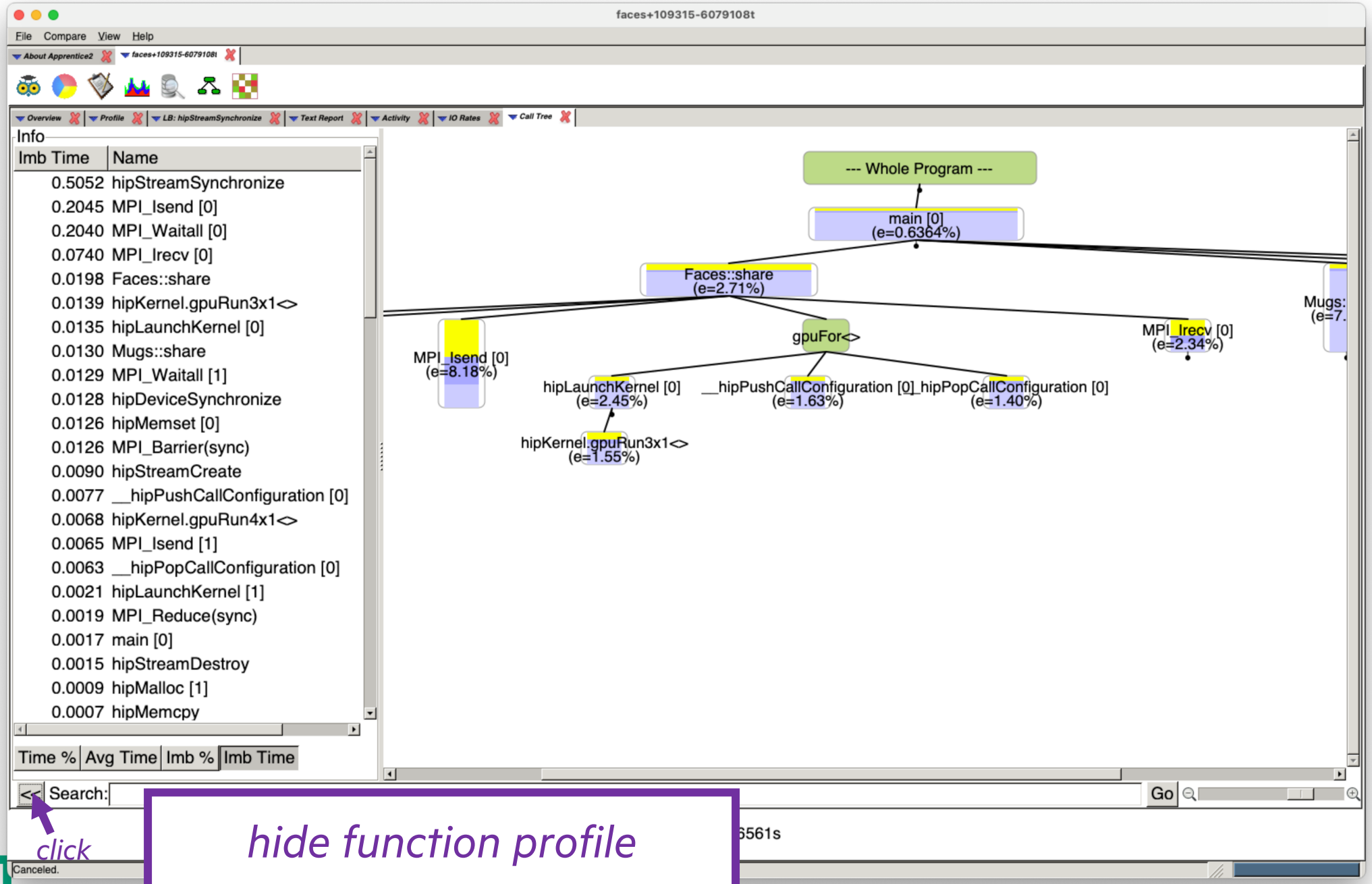

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
Text Report
Activity
click
open I/O table
Activity over PEs
Time PE
100%
75%
50%
25%
0%
0
6
12
18
25
31
37
44
50
56
63
Sync Data Xfer Read Write Collective Heap CUDA HIP OMP User Other Traced
Wallclock time: 3.576561s
Done (0.0637s)


faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
LB: hipStreamSynchronize
Text Report
Activity
IO Rates
Filename | Total Time (s) | Write Calls | Write Total (MB) | Write Avg (MB/s) | Read Calls | Read Total (MB) | R
UnknownFile 0.4118 47,265 2.2327 5.4213
stdout 0.0018 1,101 0.0056 3.0884
/tmp/comgr-f75f76/include/opencl1.2-c.pch 0.0011 2 2.7472 2466.1
/tmp/comgr-de04e4/include/opencl1.2-c.pch 0.0011 2 2.7472 2572.3
/tmp/comgr-ddd919/include/opencl1.2-c.pch 0.0011 2 2.7472 2606.5
/tmp/comgr-4f198c/include/opencl1.2-c.pch 0.0010 2 2.7472 2616.4
/tmp/comgr-e11746/include/opencl1.2-c.pch 0.0010 2 2.7472 2628.9
/tmp/comgr-53e2bb/include/opencl1.2-c.pch 0.0010 2 2.7472 2633.9
/tmp/comgr-5ed711/include/opencl1.2-c.pch 0.0010 2 2.7472 2654.3
/tmp/comgr-cbe7d6/include/opencl1.2-c.pch 0.0010 2 2.7472 2667.2
/tmp/comgr-1776af/include/opencl1.2-c.pch 0.0010 2 2.7472 2669.8
/tmp/comgr-6f895c/include/opencl1.2-c.pch 0.0010 2 2.7472 2672.4
/tmp/comgr-b344b5/include/opencl1.2-c.pch 0.0010 2 2.7472 2675.0
/tmp/comgr-f1796c/include/opencl1.2-c.pch 0.0010 2 2.7472 2677.6
/tmp/comgr-5a1fce/include/opencl1.2-c.pch 0.0010 2 2.7472 2690.7
/tmp/comgr-cda661/include/opencl1.2-c.pch 0.0010 2 2.7472 2696.0
/tmp/comgr-9609f6/include/opencl1.2-c.pch 0.0010 2 2.7472 2698.6
/tmp/comgr-5b5c66/include/opencl1.2-c.pch 0.0010 2 2.7472 2714.6
/tmp/comgr-30c02d/include/opencl1.2-c.pch 0.0010 2 2.7472 2717.3
/tmp/comgr-0b4dd0/include/opencl1.2-c.pch 0.0010 2 2.7472 2722.7
/tmp/comgr-10441c/include/opencl1.2-c.pch 0.0010 2 2.7472 2725.4
/tmp/comgr-da4688/include/opencl1.2-c.pch 0.0010 2 2.7472 2728.1
/tmp/comgr-6e7fd1/include/opencl1.2-c.pch 0.0010 2 2.7472 2733.5
/tmp/comgr-f4730c/include/opencl1.2-c.pch 0.0010 2 2.7472 2739.0
/tmp/comgr-34e54a/include/opencl1.2-c.pch 0.0010 2 2.7472 2749.9
/tmp/comgr-100ad2/include/opencl1.2-c.pch 0.0010 2 2.7472 2758.2
/tmp/comgr-3711f2/include/opencl1.2-c.pch 0.0010 2 2.7472 2761.0
/tmp/comgr-a9e5ea/include/opencl1.2-c.pch 0.0010 2 2.7472 2761.0
/tmp/comgr-6bf774/include/opencl1.2-c.pch 0.0010 2 2.7472 2763.8
Wallclock time: 3.576561s
Done (0.0637s)



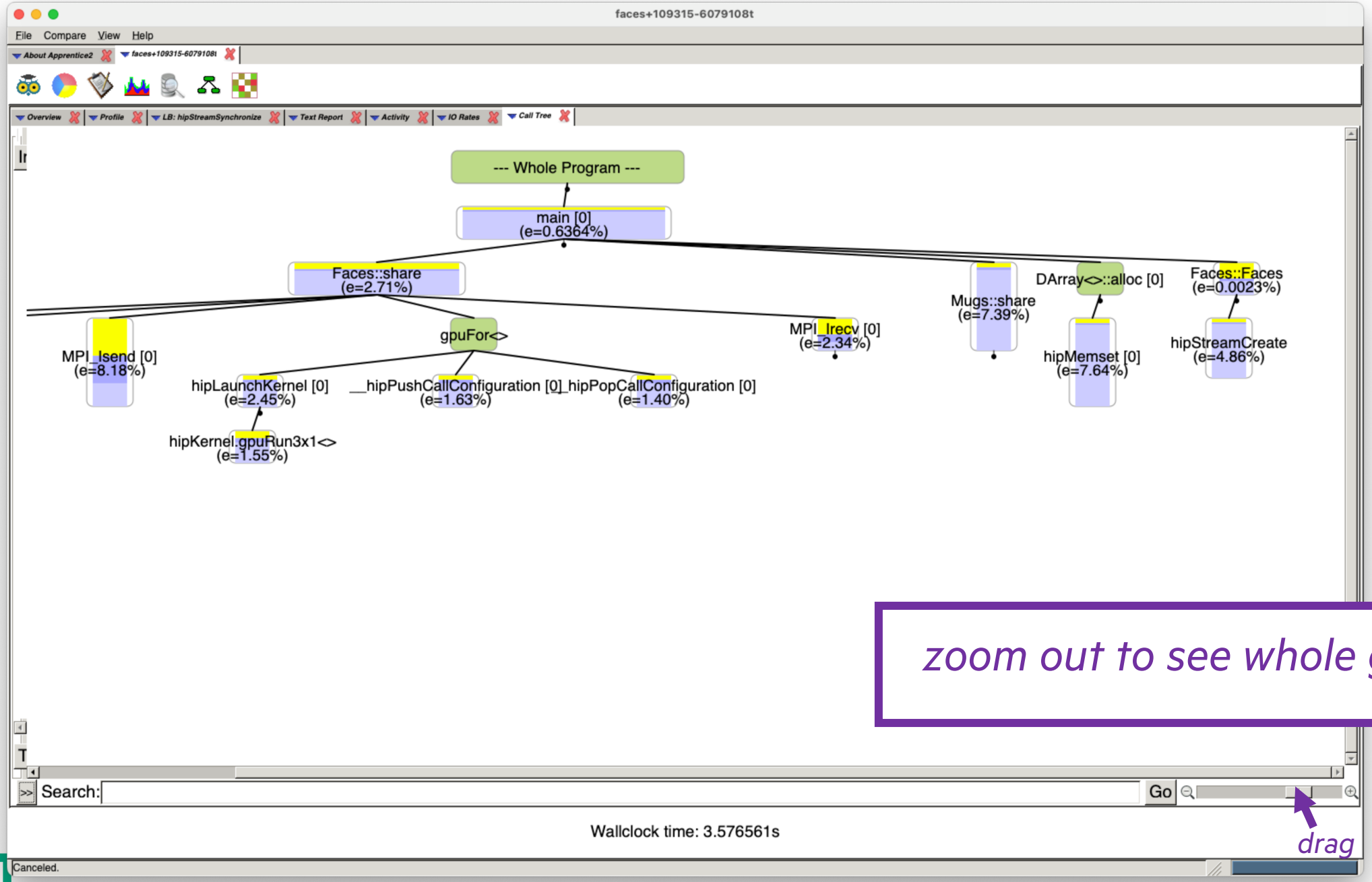
	Total Time (s)	Write Calls	Write Total (MB)	Write Avg (MB/s)	Read Calls	Read Total (MB)	R
UnknownFile	0.4118	47,265	2.2327	5.4213			
std			0.0056	3.0884			
/tm			2.7472	2466.1			
/tm			2.7472	2572.3			
/tm			2.7472	2606.5			
/tm			2.7472	2616.4			
/tm			2.7472	2628.9			
/tmp/comgr-53e2bb/include/opencl1.2-c.pch	0.0010	2	2.7472	2633.9			
/tmp/comgr-5ed711/include/opencl1.2-c.pch	0.0010	2	2.7472	2654.3			
/tmp/comgr-cbe7d6/include/opencl1.2-c.pch	0.0010	2	2.7472	2667.2			
/tmp/comgr-1776af/include/opencl1.2-c.pch	0.0010	2	2.7472	2669.8			
/tmp/comgr-6f895c/include/opencl1.2-c.pch	0.0010	2	2.7472	2672.4			
/tmp/comgr-b344b5/include/opencl1.2-c.pch	0.0010	2	2.7472	2675.0			
/tmp/comgr-f1796c/include/opencl1.2-c.pch	0.0010	2	2.7472	2677.6			
/tmp/comgr-5a1fce/include/opencl1.2-c.pch	0.0010	2	2.7472	2690.7			
/tmp/comgr-cda661/include/opencl1.2-c.pch	0.0010	2	2.7472	2696.0			
/tmp/comgr-9609f6/include/opencl1.2-c.pch	0.0010	2	2.7472	2698.6			
/tmp/comgr-5b5c66/include/opencl1.2-c.pch	0.0010	2	2.7472	2714.6			
/tmp/comgr-30c02d/include/opencl1.2-c.pch	0.0010	2	2.7472	2717.3			
/tmp/comgr-0b4dd0/include/opencl1.2-c.pch	0.0010	2	2.7472	2722.7			
/tmp/comgr-10441c/include/opencl1.2-c.pch	0.0010	2	2.7472	2725.4			
/tmp/comgr-da4688/include/opencl1.2-c.pch	0.0010	2	2.7472	2728.1			
/tmp/comgr-6e7fd1/include/opencl1.2-c.pch	0.0010	2	2.7472	2733.5			
/tmp/comgr-f4730c/include/opencl1.2-c.pch	0.0010	2	2.7472	2739.0			
/tmp/comgr-34e54a/include/opencl1.2-c.pch	0.0010	2	2.7472	2749.9			
/tmp/comgr-100ad2/include/opencl1.2-c.pch	0.0010	2	2.7472	2758.2			
/tmp/comgr-3711f2/include/opencl1.2-c.pch	0.0010	2	2.7472	2761.0			
/tmp/comgr-a9e5ea/include/opencl1.2-c.pch	0.0010	2	2.7472	2761.0			
/tmp/comgr-6bf774/include/opencl1.2-c.pch	0.0010	2	2.7472	2763.8			

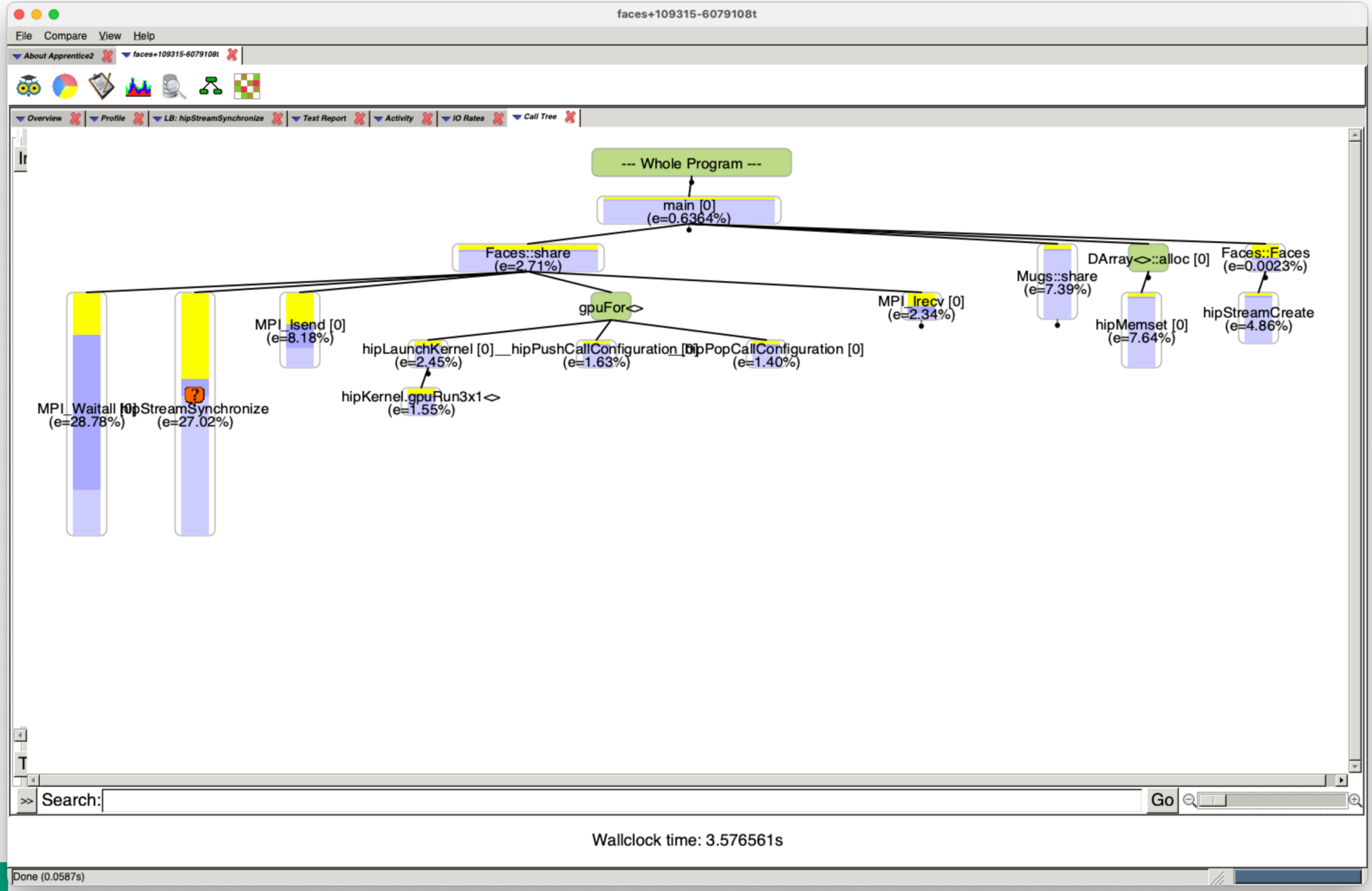

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
LB: hipStreamSynchronize
Text Report
Activity
IO Rates
Call Tree
Info
Imb Time Name
0.5052 hipStreamSynchronize
0.2045 MPI_Isend [0]
0.2040 MPI_Waitall [0]
0.0740 MPI_Irecv [0]
0.0198 Faces::share
0.0139 hipKernel.gpuRun3x1<>
0.0135 hipLaunchKernel [0]
0.0130 Mugs::share
0.0129 MPI_Waitall [1]
0.0128 hipDeviceSynchronize
0.0126 hipMemset [0]
0.0126 MPI_Barrier(sync)
0.0090 hipStreamCreate
0.0077 __hipPushCallConfiguration [0]
0.0068 hipKernel.gpuRun4x1<>
0.0065 MPI_Isend [1]
0.0063 __hipPopCallConfiguration [0]
0.0021 hipLaunchKernel [1]
0.0019 MPI_Reduce(sync)
0.0017 main [0]
0.0015 hipStreamDestroy
0.0009 hipMalloc [1]
0.0007 hipMemcpy
Time % Avg Time Imb % Imb Time
--- Whole Program ---
main [0] (e=0.6364%)
Faces::share (e=2.71%)
MPI_Isend [0] (e=8.18%)
gpuFor<>
MPI_Irecv [0] (e=2.34%)
Mugs: (e=7.
hipLaunchKernel [0] (e=2.45%)
__hipPushCallConfiguration [0]
(e=1.63%)
__hipPopCallConfiguration [0]
(e=1.40%)
hipKernel.gpuRun3x1<> (e=1.55%)
<< Search:
Go
Wallclock time: 3.576561s
Canceled.

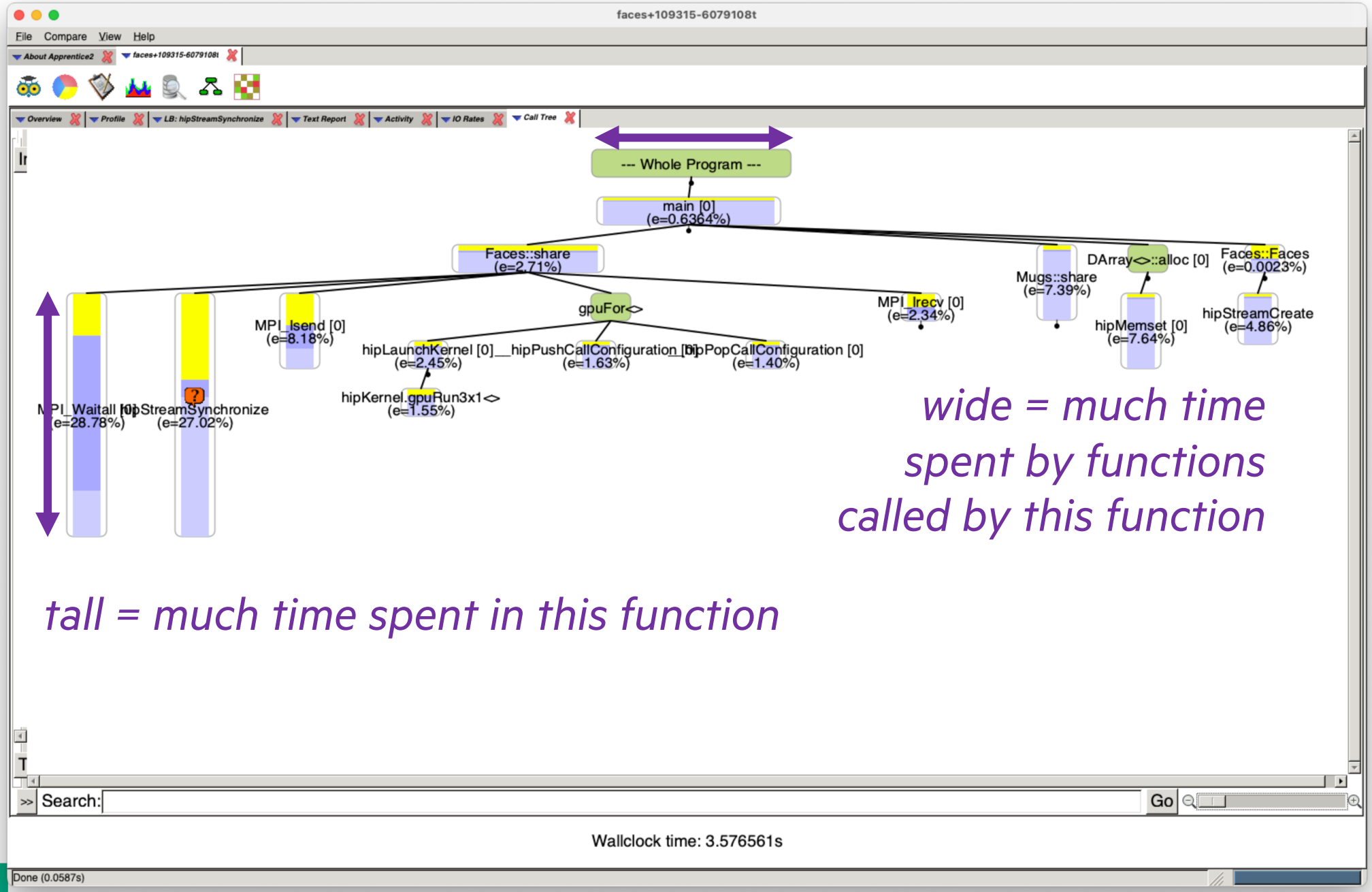


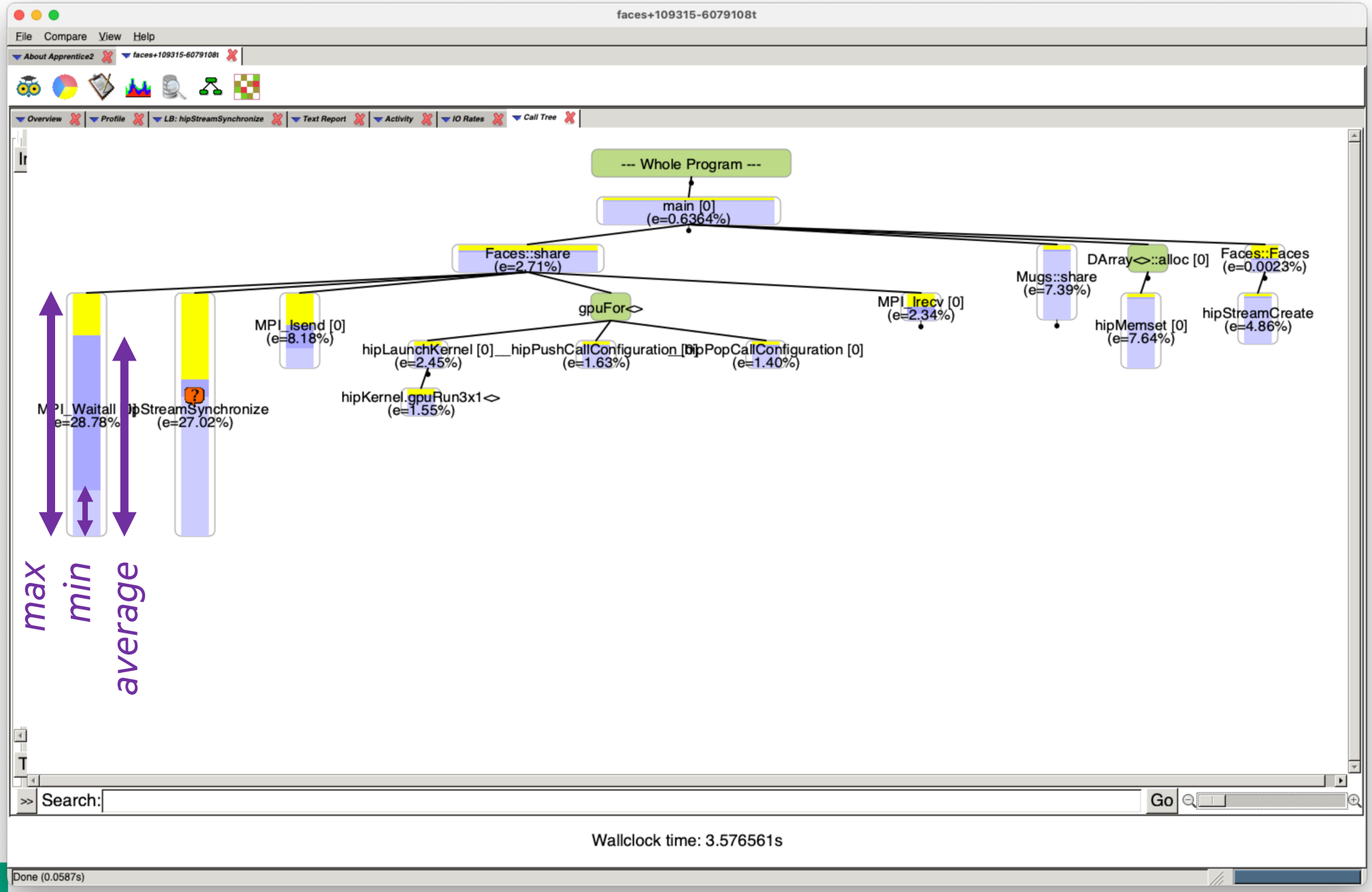
click

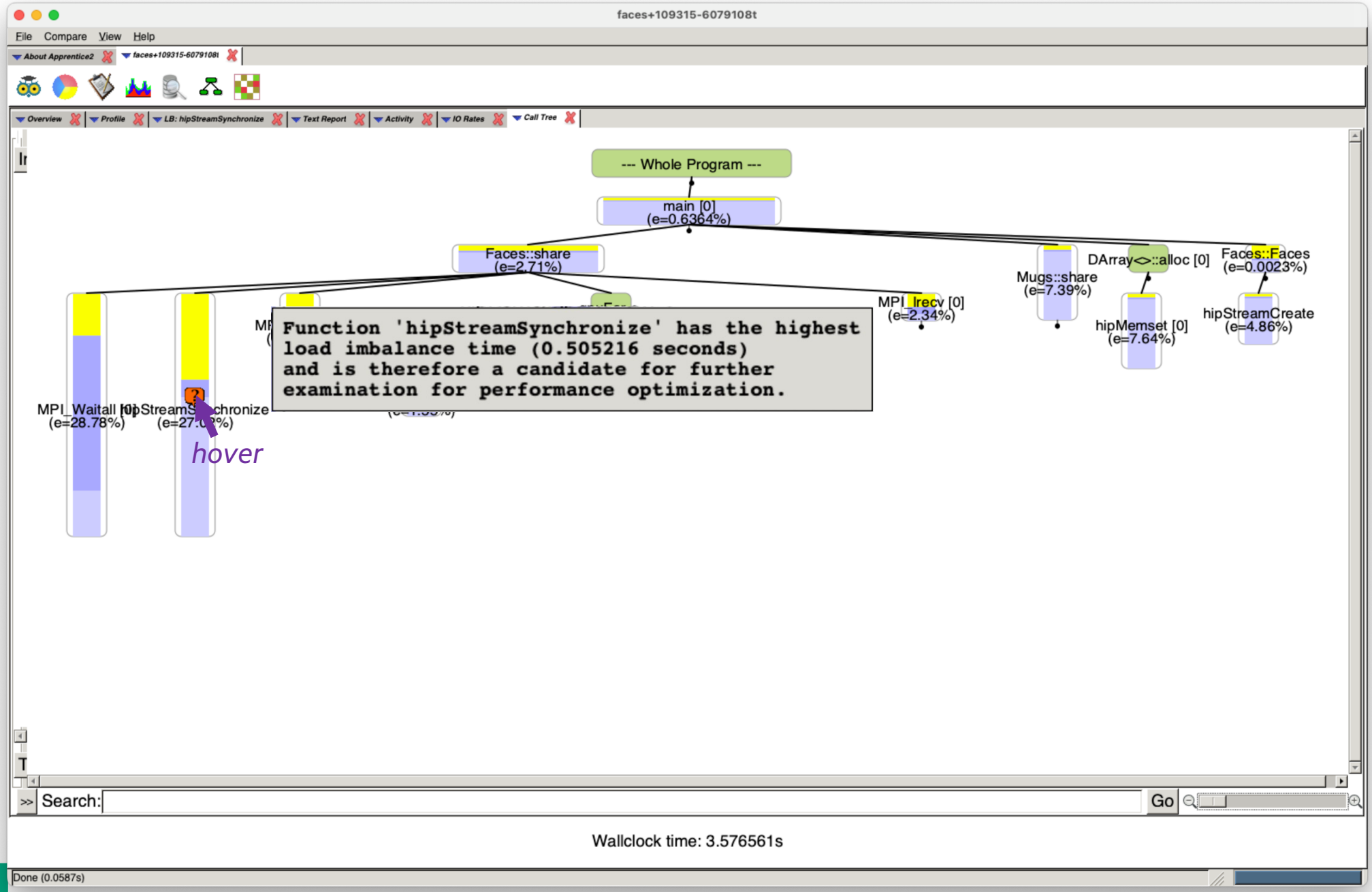
hide function profile

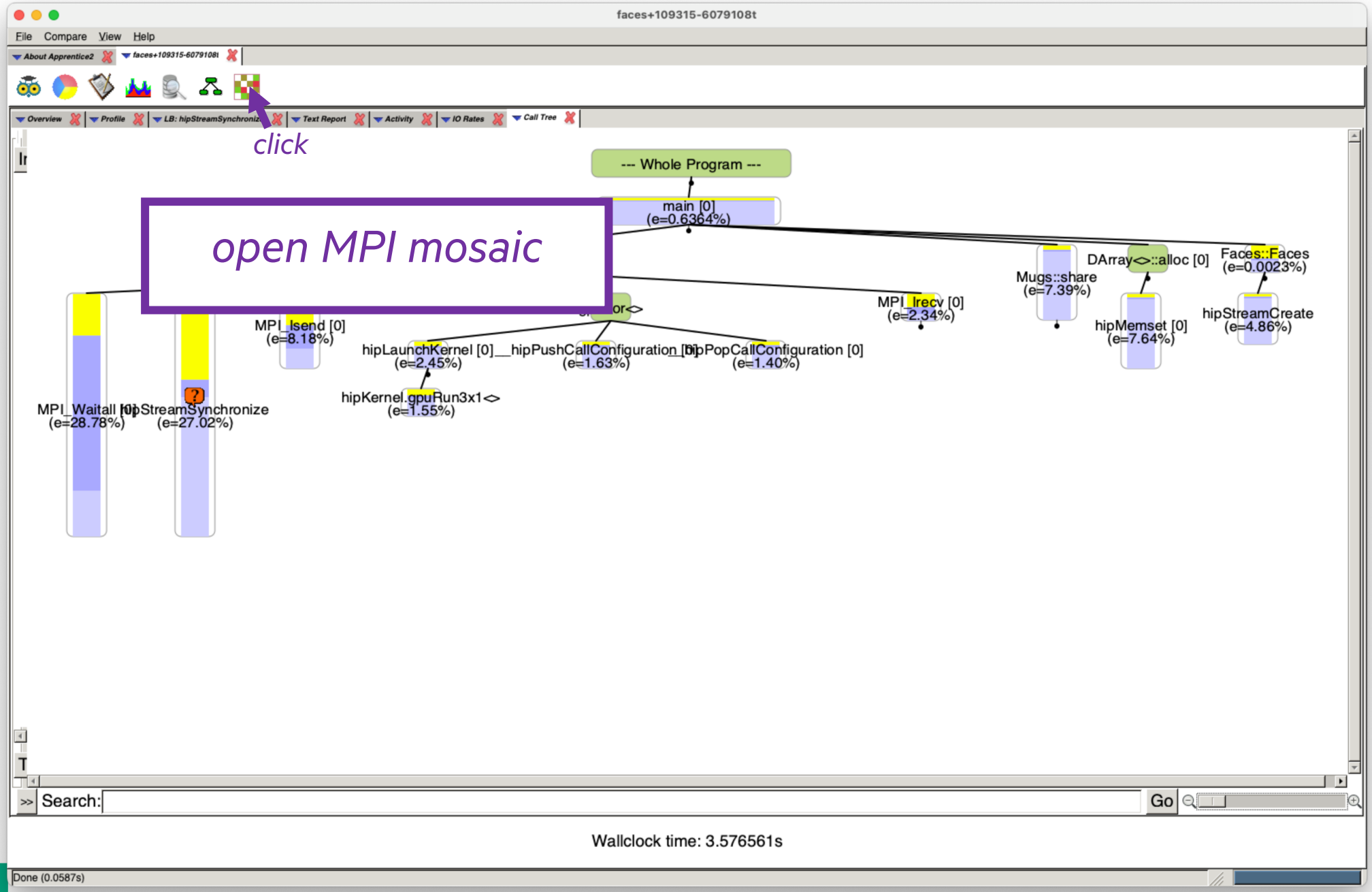

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
LB: hipStreamSynchronize
Text Report
Activity
IO Rates
Call Tree
--- Whole Program ---
main [0]
(e=0.6364%)
Faces::share
(e=2.71%)
MPI_Isend [0]
(e=8.18%)
gpuFor<>
hipLaunchKernel [0]
(e=2.45%)
hipKernel.gpuRun3x1<>
(e=1.55%)
__hipPushCallConfiguration [0]
(e=1.63%)
__hipPopCallConfiguration [0]
(e=1.40%)
MPI_Irecv [0]
(e=2.34%)
Mugs::share
(e=7.39%)
DArray<>::alloc [0]
hipMemset [0]
(e=7.64%)
Faces::Faces
(e=0.0023%)
hipStreamCreate
(e=4.86%)
Search:
Go
Wallclock time: 3.576561s
Canceled.
zoom out to see whole graph
drag

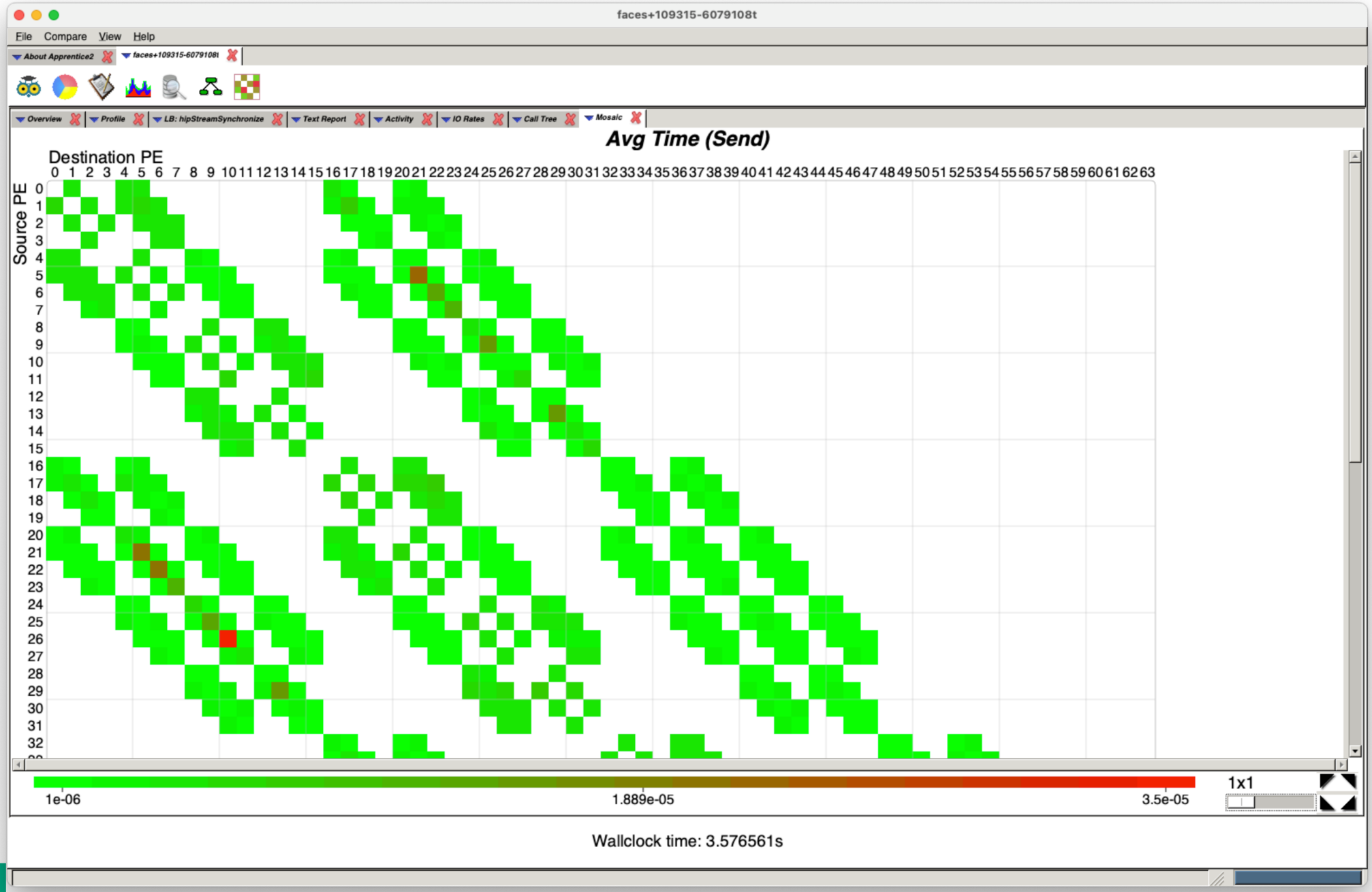

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
LB: hipStreamSynchronize
Text Report
Activity
IO Rates
Call Tree
--- Whole Program ---
main [0]
(e=0.6364%)
Faces::share
(e=2.71%)
Mugs::share
(e=7.39%)
DArray<>::alloc [0]
Faces::Faces
(e=0.0023%)
MPI_Irecv [0]
(e=2.34%)
hipMemset [0]
(e=7.64%)
hipStreamCreate
(e=4.86%)
MPI_Isend [0]
(e=8.18%)
gpuFor<>
hipLaunchKernel [0]
(e=2.45%)
__hipPushCallConfiguration [0]
(e=1.63%)
__hipPopCallConfiguration [0]
(e=1.40%)
hipKernel.gpuRun3x1<>
(e=1.55%)
MPI_Waitall [0]
(e=28.78%)
hipStreamSynchronize
(e=27.02%)
Search:
Go
Wallclock time: 3.576561s
Done (0.0587s)

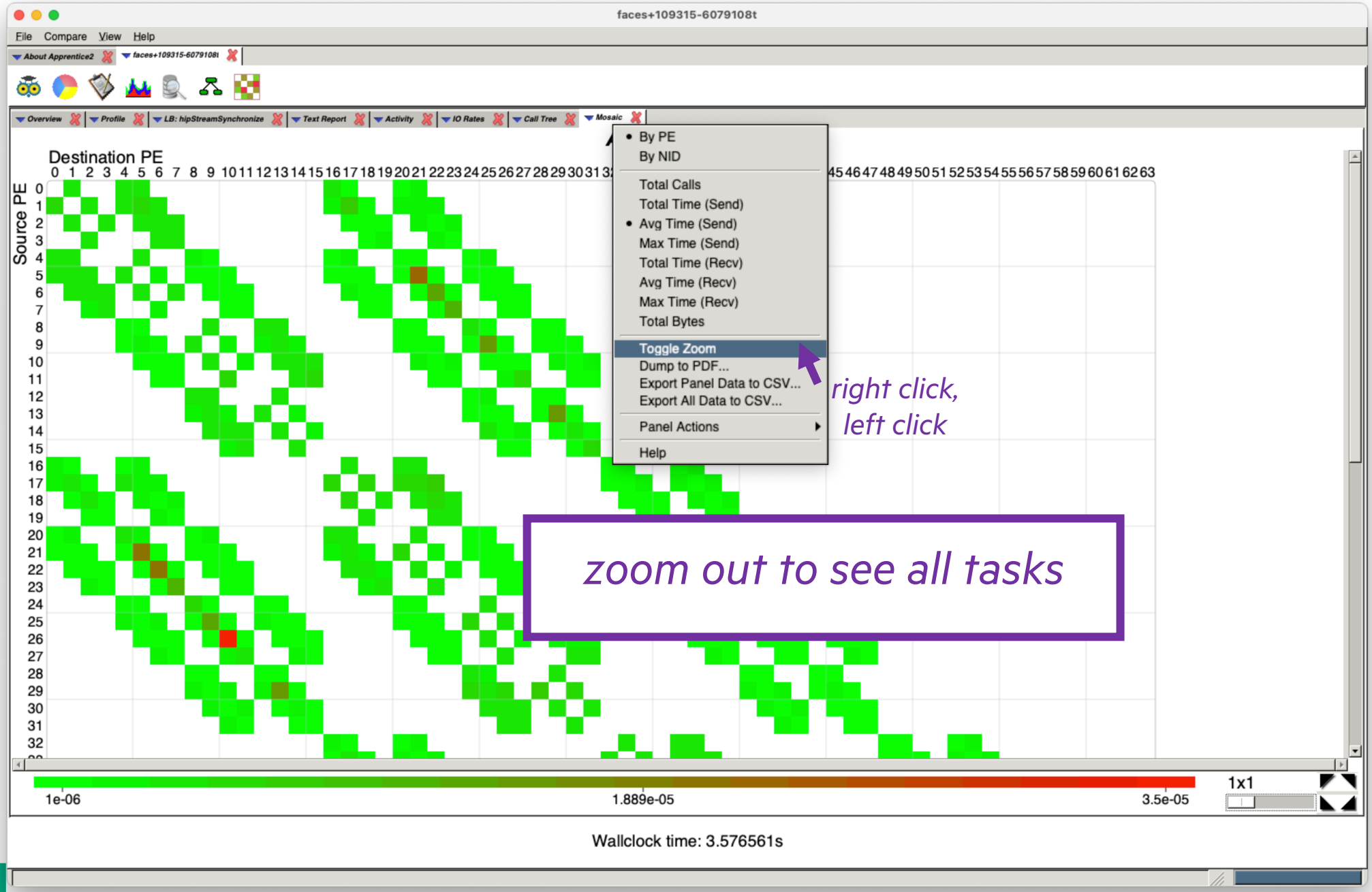

faces+109315-6079108t
File Compare View Help
Overview Profile LB: hipStreamSynchronize Text Report Activity IO Rates Call Tree
--- Whole Program ---
main [0] (e=0.6364%)
Faces::share (e=2.71%)
Mugs::share (e=7.39%)
DArray<>::alloc [0]
Faces::Faces (e=0.0023%)
MPI_Irecv [0] (e=2.34%)
MPI_Isend [0] (e=8.18%)
gpuFor<>
hipMemset [0] (e=7.64%)
hipStreamCreate (e=4.86%)
hipLaunchKernel [0] (e=2.45%)
__hipPushCallConfiguration [0] (e=1.63%)
__hipPopCallConfiguration [0] (e=1.40%)
hipKernel.gpuRun3x1<> (e=1.55%)
MPI_Waitall [0] (e=28.78%)
hipStreamSynchronize (e=27.02%)
wide = much time spent by functions called by this function
tall = much time spent in this function
Search:
Go
Wallclock time: 3.576561s
Done (0.0587s)

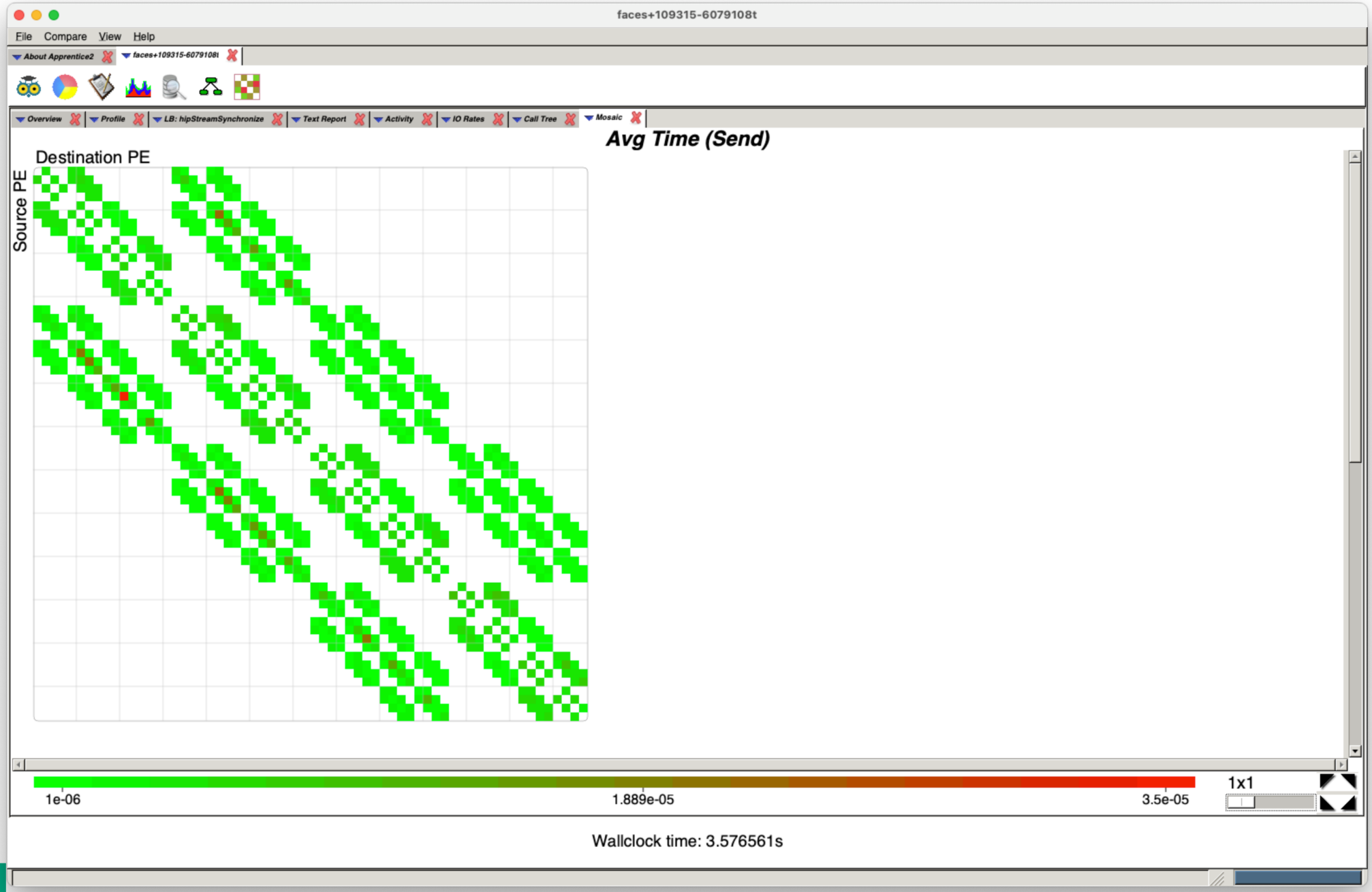

faces+109315-6079108t
File Compare View Help
Overview
Profile
LB: hipStreamSynchronize
Text Report
Activity
IO Rates
Call Tree
--- Whole Program ---
main [0]
(e=0.6364%)
Faces::share
(e=2.71%)
Mugs::share
(e=7.39%)
DArray<>::alloc [0]
Faces::Faces
(e=0.0023%)
MPI_Irecv [0]
(e=2.34%)
hipMemset [0]
(e=7.64%)
hipStreamCreate
(e=4.86%)
MPI_Isend [0]
(e=8.18%)
gpuFor<>
hipLaunchKernel [0]
(e=2.45%)
__hipPushCallConfiguration [0]
(e=1.63%)
__hipPopCallConfiguration [0]
(e=1.40%)
hipKernel.gpuRun3x1<>
(e=1.55%)
MPI_Waitall
e=28.78%
hipStreamSynchronize
(e=27.02%)
max
min
average
Search:
Go
Wallclock time: 3.576561s
Done (0.0587s)

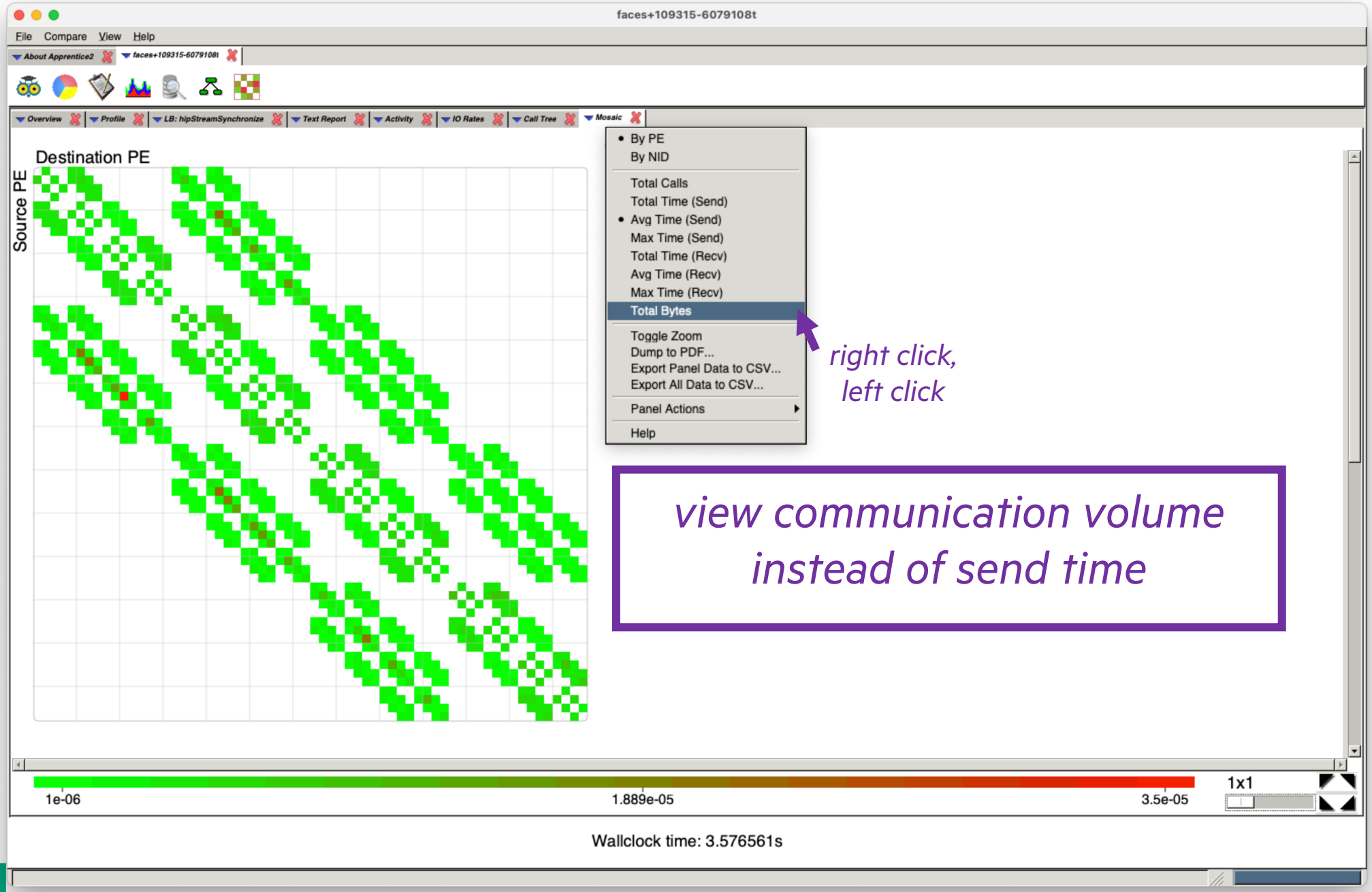

faces+109315-6079108t
File Compare View Help
Overview
Profile
LB: hipStreamSynchronize
Text Report
Activity
IO Rates
Call Tree
--- Whole Program ---
main [0]
(e=0.6364%)
Faces::share
(e=2.71%)
MPI_Irecv [0]
(e=2.34%)
Mugs::share
(e=7.39%)
DArray<>::alloc [0]
hipMemset [0]
(e=7.64%)
Faces::Faces
(e=0.0023%)
hipStreamCreate
(e=4.86%)
MPI_Waitall [0]
(e=28.78%)
Function 'hipStreamSynchronize' has the highest load imbalance time (0.505216 seconds) and is therefore a candidate for further examination for performance optimization.
hover
Search:
Go
Wallclock time: 3.576561s
Done (0.0587s)

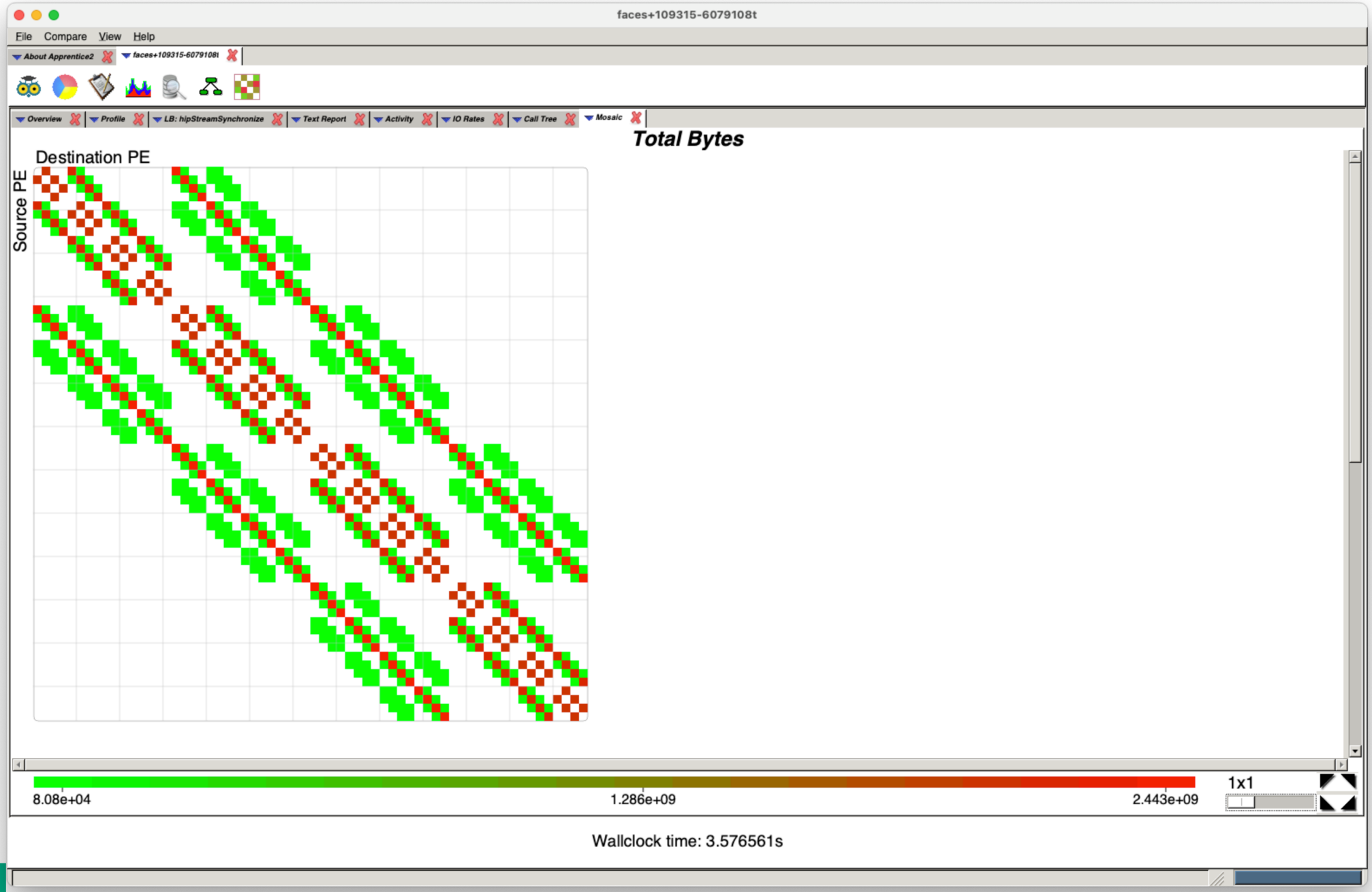

faces+109315-6079108t
File Compare View Help
click
open MPI mosaic
--- Whole Program ---
main [0]
(e=0.6364%)
MPI_Isend [0]
(e=8.18%)
hipLaunchKernel [0]
(e=2.45%)
hipPushCallConfiguration [0]
(e=1.63%)
hipPopCallConfiguration [0]
(e=1.40%)
hipKernel.gpuRun3x1<>
(e=1.55%)
MPI_Waitall [0]
(e=28.78%)
hipStreamSynchronize
(e=27.02%)
MPI_Irecv [0]
(e=2.34%)
Mugs::share
(e=7.39%)
DArray<>::alloc [0]
hipMemset [0]
(e=7.64%)
Faces::Faces
(e=0.0023%)
hipStreamCreate
(e=4.86%)
Search:
Go
Wallclock time: 3.576561s
Done (0.0587s)

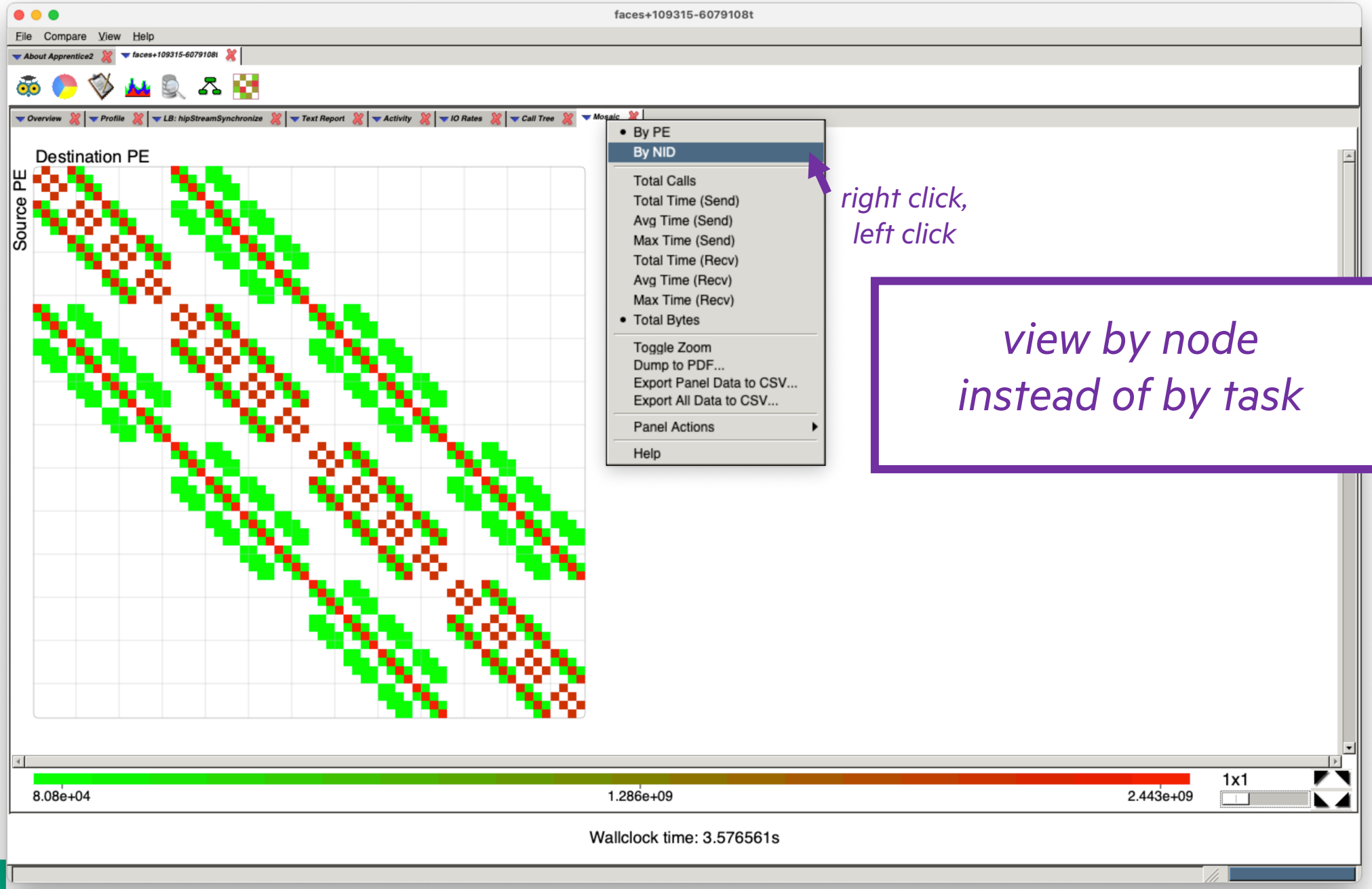

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
LB: hipStreamSynchronize
Text Report
Activity
IO Rates
Call Tree
Mosaic
Avg Time (Send)
Destination PE
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63
Source PE
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
1e-06
1.889e-05
3.5e-05
1x1
Wallclock time: 3.576561s

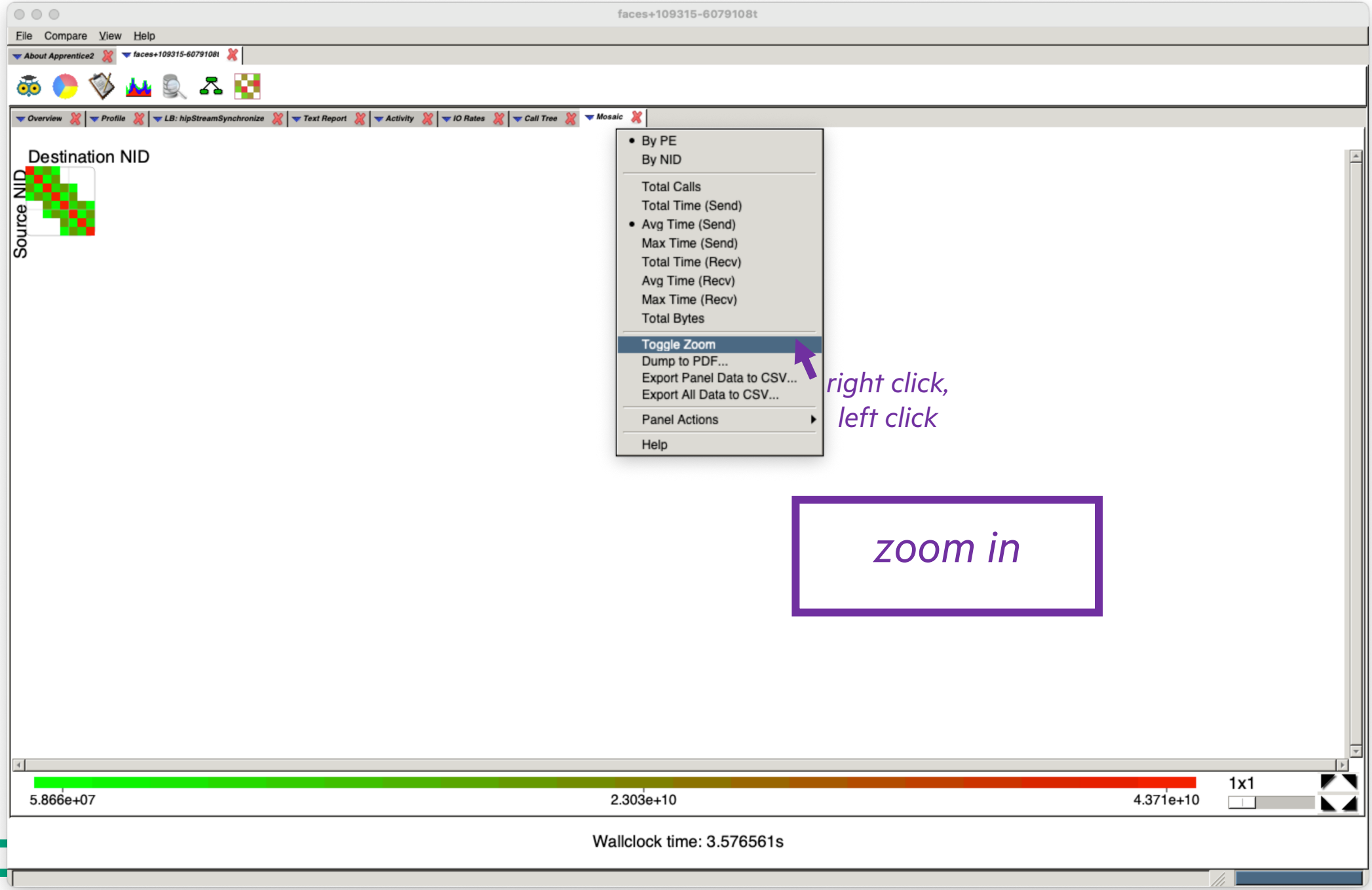

faces+109315-6079108t
File Compare View Help
Overview
Profile
LB: hipStreamSynchronize
Text Report
Activity
IO Rates
Call Tree
Mosaic
Destination PE
Source PE
By PE
By NID
Total Calls
Total Time (Send)
Avg Time (Send)
Max Time (Send)
Total Time (Recv)
Avg Time (Recv)
Max Time (Recv)
Total Bytes
Toggle Zoom
Dump to PDF...
Export Panel Data to CSV...
Export All Data to CSV...
Panel Actions
Help
right click,
left click
zoom out to see all tasks
1e-06
1.889e-05
3.5e-05
1x1
Wallclock time: 3.576561s

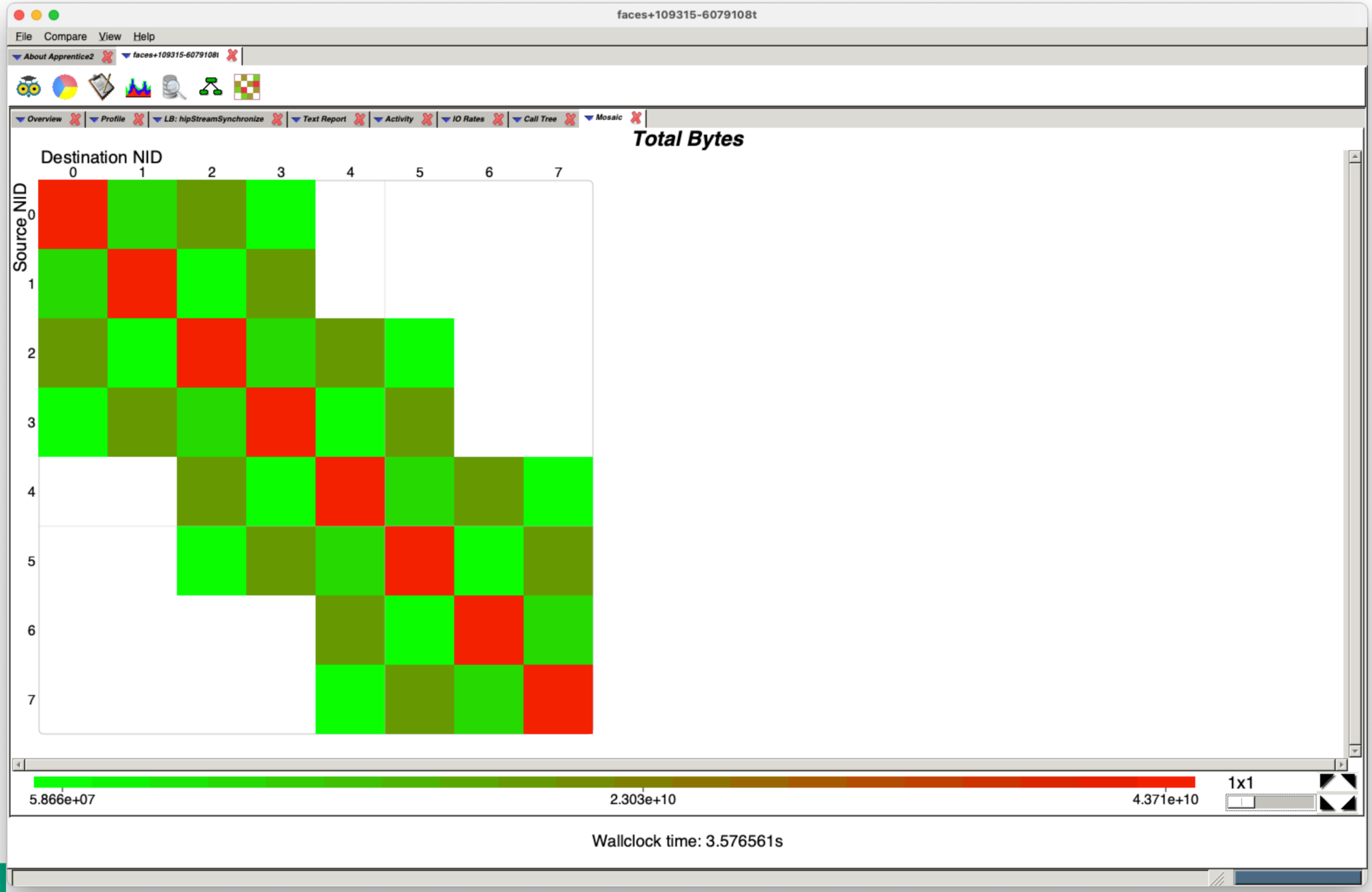

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
LB: hipStreamSynchronize
Text Report
Activity
IO Rates
Call Tree
Mosaic
Avg Time (Send)
Destination PE
Source PE
1e-06
1.889e-05
3.5e-05
1x1
Wallclock time: 3.576561s


faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
LB: hipStreamSynchronize
Text Report
Activity
IO Rates
Call Tree
Mosaic
Destination PE
Source PE
By PE
By NID
Total Calls
Total Time (Send)
Avg Time (Send)
Max Time (Send)
Total Time (Recv)
Avg Time (Recv)
Max Time (Recv)
Total Bytes
Toggle Zoom
Dump to PDF...
Export Panel Data to CSV...
Export All Data to CSV...
Panel Actions
Help
right click,
left click
view communication volume
instead of send time
1e-06
1.889e-05
3.5e-05
1x1
Wallclock time: 3.576561s


faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
LB: hipStreamSynchronize
Text Report
Activity
IO Rates
Call Tree
Mosaic
Total Bytes
Destination PE
Source PE
8.08e+04
1.286e+09
2.443e+09
1x1
Wallclock time: 3.576561s


faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
LB: hipStreamSynchronize
Text Report
Activity
IO Rates
Call Tree
Mosaic
Destination PE
Source PE
By PE
By NID
Total Calls
Total Time (Send)
Avg Time (Send)
Max Time (Send)
Total Time (Recv)
Avg Time (Recv)
Max Time (Recv)
Total Bytes
Toggle Zoom
Dump to PDF...
Export Panel Data to CSV...
Export All Data to CSV...
Panel Actions
Help
right click,
left click
view by node
instead of by task
8.08e+04
1.286e+09
2.443e+09
1x1
Wallclock time: 3.576561s

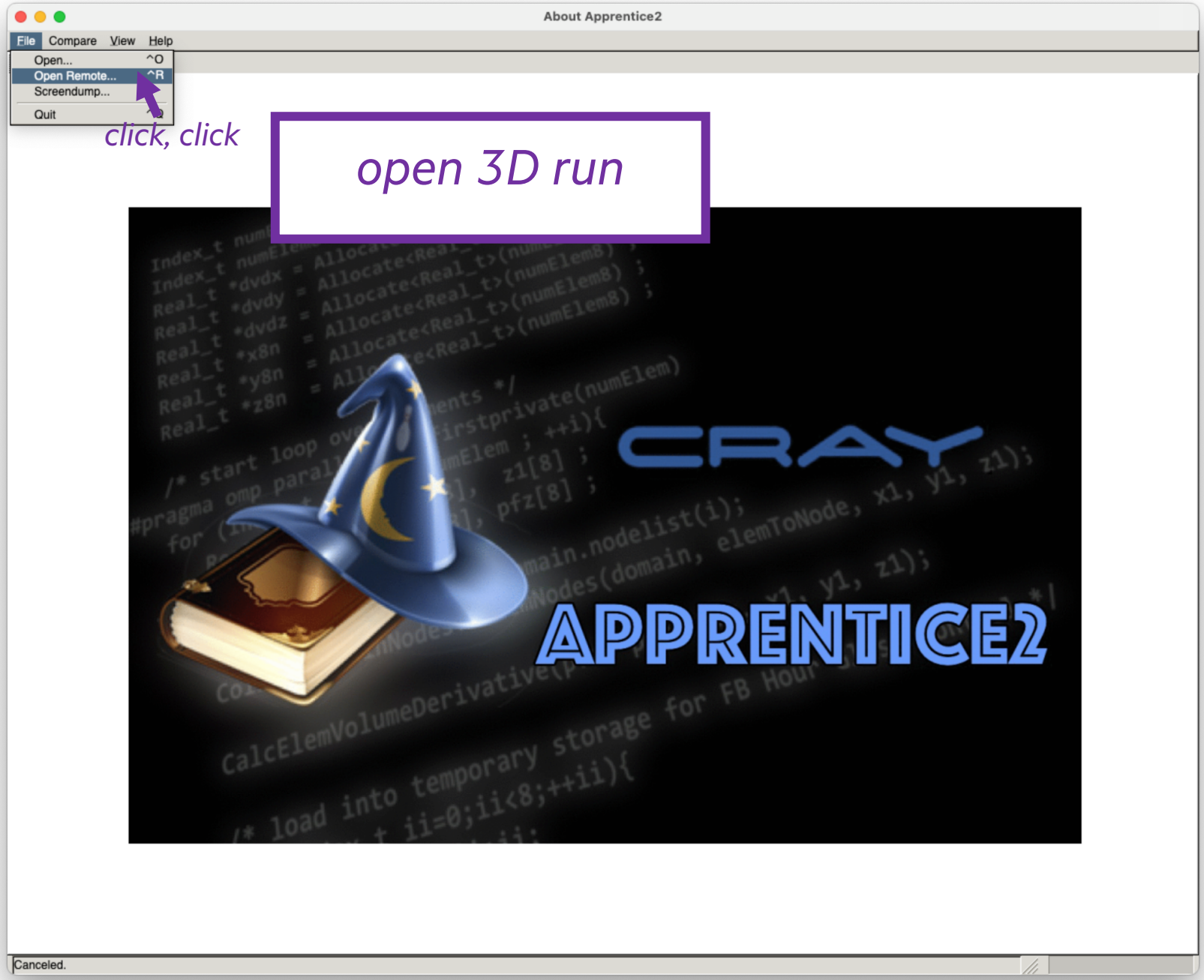

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
LB: hipStreamSynchronize
Text Report
Activity
IO Rates
Call Tree
Mosaic
Destination NID
Source NID
By PE
By NID
Total Calls
Total Time (Send)
Avg Time (Send)
Max Time (Send)
Total Time (Recv)
Avg Time (Recv)
Max Time (Recv)
Total Bytes
Toggle Zoom
Dump to PDF...
Export Panel Data to CSV...
Export All Data to CSV...
Panel Actions
Help
right click,
left click
zoom in
5.866e+07
2.303e+10
4.371e+10
1x1
Wallclock time: 3.576561s

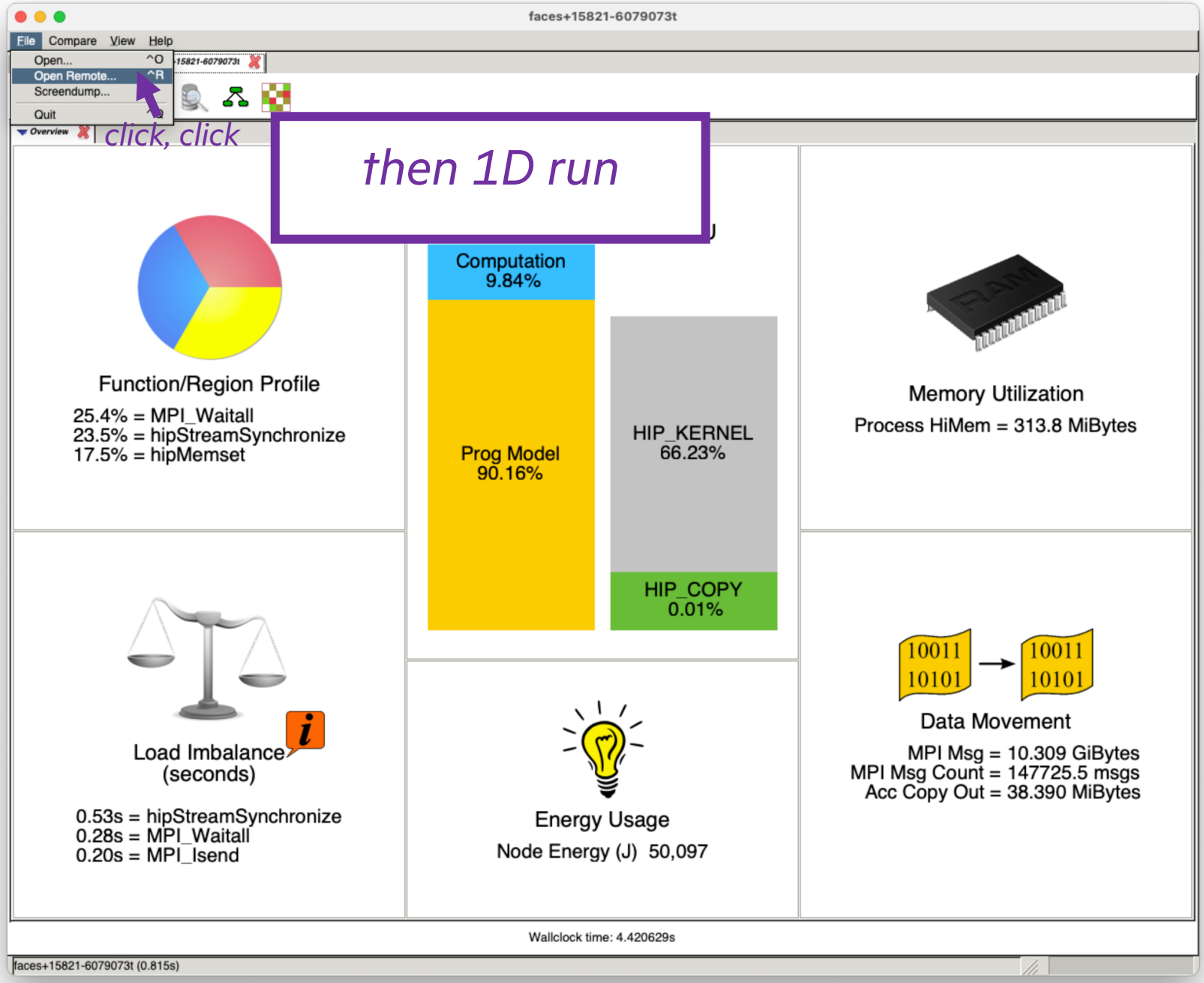

faces+109315-6079108t
File Compare View Help
About Apprentice2
faces+109315-6079108t
Overview
Profile
LB: hipStreamSynchronize
Text Report
Activity
IO Rates
Call Tree
Mosaic
Total Bytes
Destination NID
0 1 2 3 4 5 6 7
Source NID
0 1 2 3 4 5 6 7
5.866e+07
2.303e+10
4.371e+10
1x1
Wallclock time: 3.576561s

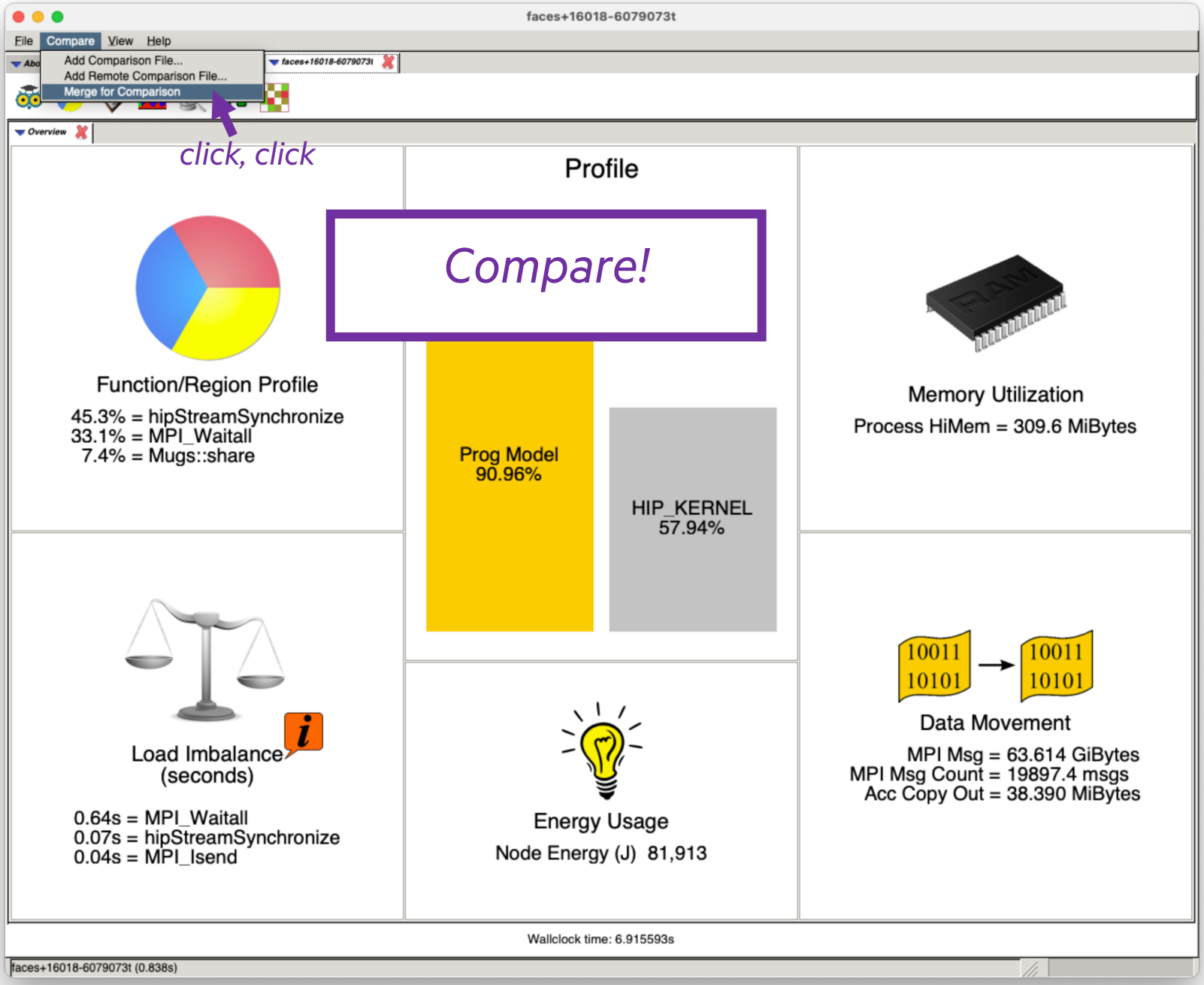
What does "Compare" do?

HIPCC RUN COMPARING 1D AND 3D DECOMPOSITIONS

```
salloc -N 8 -t 30:00
module load perftools-lite-gpu
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in3D.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
srun -l -u -t 5:00 -i in1D.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
```


About Apprentice2
File
Compare
View
Help
Open...
Open Remote...
Screendump...
Quit
click, click
open 3D run
CRAY
APPRENTICE2
Canceled.

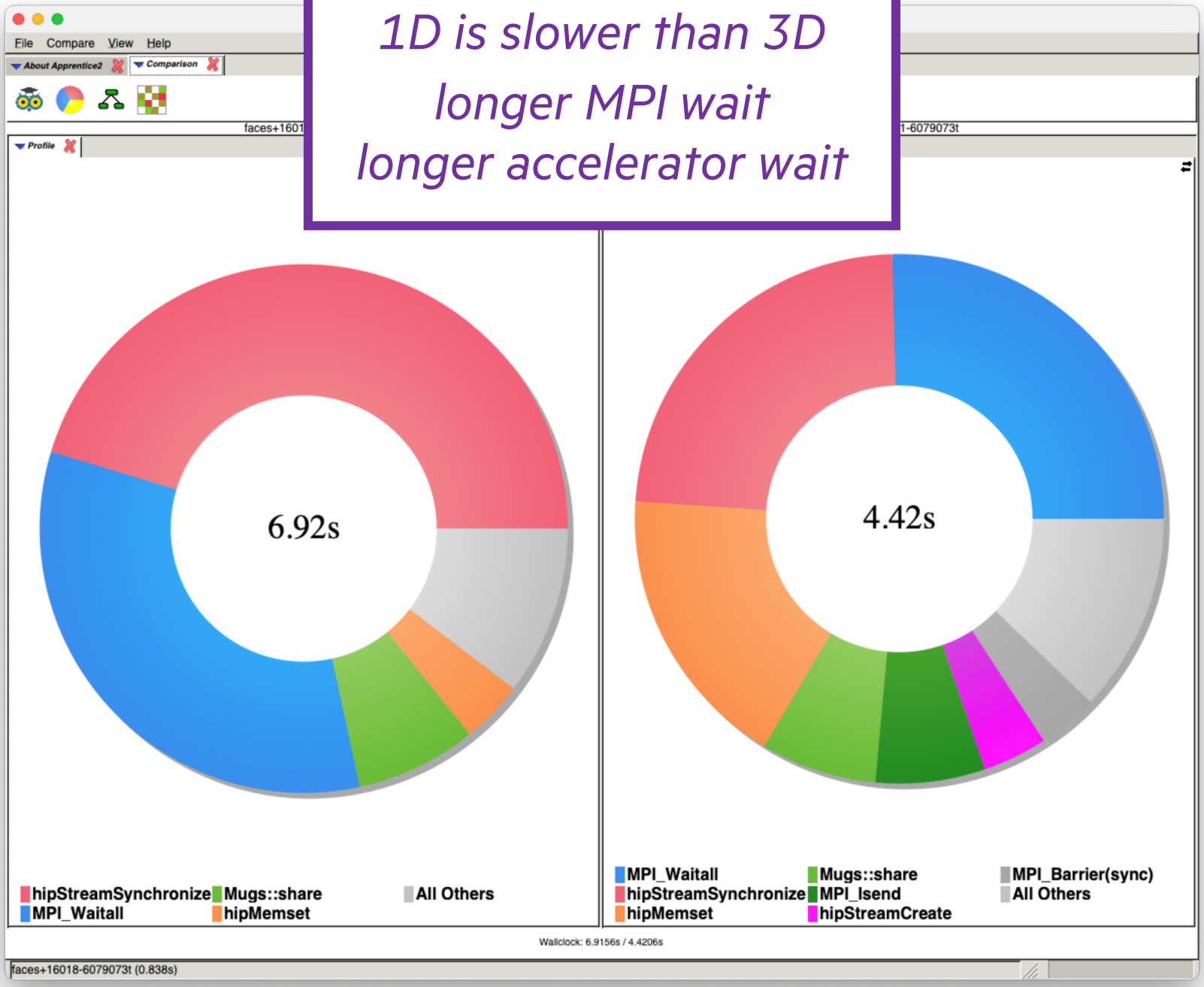

faces+15821-6079073t
File Compare View Help
Open... ^O
Open Remote... ^R
Screendump...
Quit
Overview
click, click
then 1D run
Function/Region Profile
25.4% = MPI_Waitall
23.5% = hipStreamSynchronize
17.5% = hipMemset
Computation
9.84%
Prog Model
90.16%
HIP_KERNEL
66.23%
HIP_COPY
0.01%
Memory Utilization
Process HiMem = 313.8 MiBytes
Load Imbalance
(seconds)
0.53s = hipStreamSynchronize
0.28s = MPI_Waitall
0.20s = MPI_Isend
Energy Usage
Node Energy (J) 50,097
10011
10101
10011
10101
Data Movement
MPI Msg = 10.309 GiBytes
MPI Msg Count = 147725.5 msgs
Acc Copy Out = 38.390 MiBytes
Wallclock time: 4.420629s
faces+15821-6079073t (0.815s)

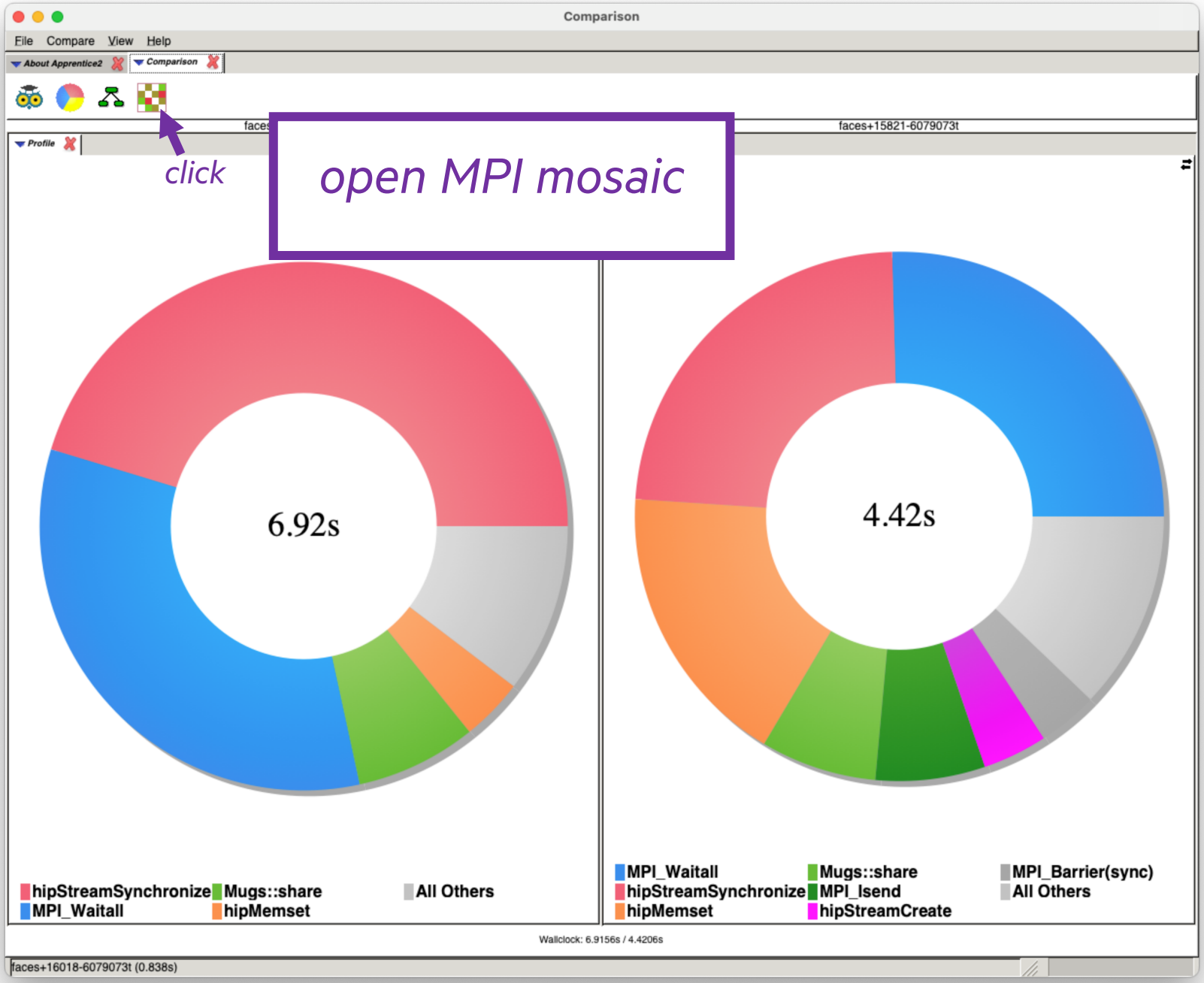

faces+16018-6079073t
File
Compare
View
Help
Add Comparison File...
Add Remote Comparison File...
Merge for Comparison
click, click
Overview
Profile
Compare!
Function/Region Profile
45.3% = hipStreamSynchronize
33.1% = MPI_Waitall
7.4% = Mugs::share
Prog Model
90.96%
HIP_KERNEL
57.94%
Memory Utilization
Process HiMem = 309.6 MiBytes
Load Imbalance
(seconds)
0.64s = MPI_Waitall
0.07s = hipStreamSynchronize
0.04s = MPI_Isend
Energy Usage
Node Energy (J) 81,913
10011
10101
10011
10101
Data Movement
MPI Msg = 63.614 GiBytes
MPI Msg Count = 19897.4 msgs
Acc Copy Out = 38.390 MiBytes
Wallclock time: 6.915593s
faces+16018-6079073t (0.838s)

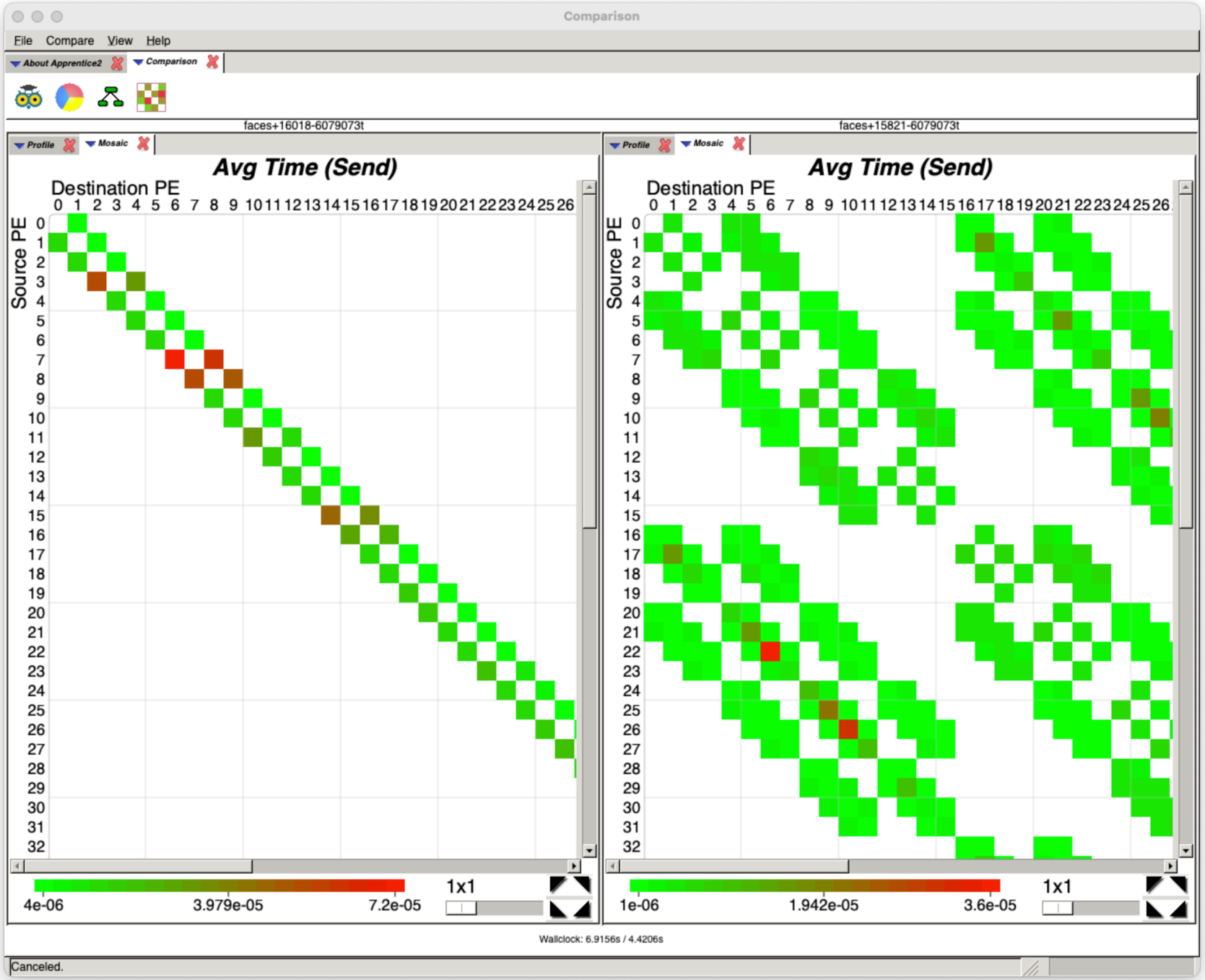
1D is slower than 3D

longer MPI wait

longer accelerator wait

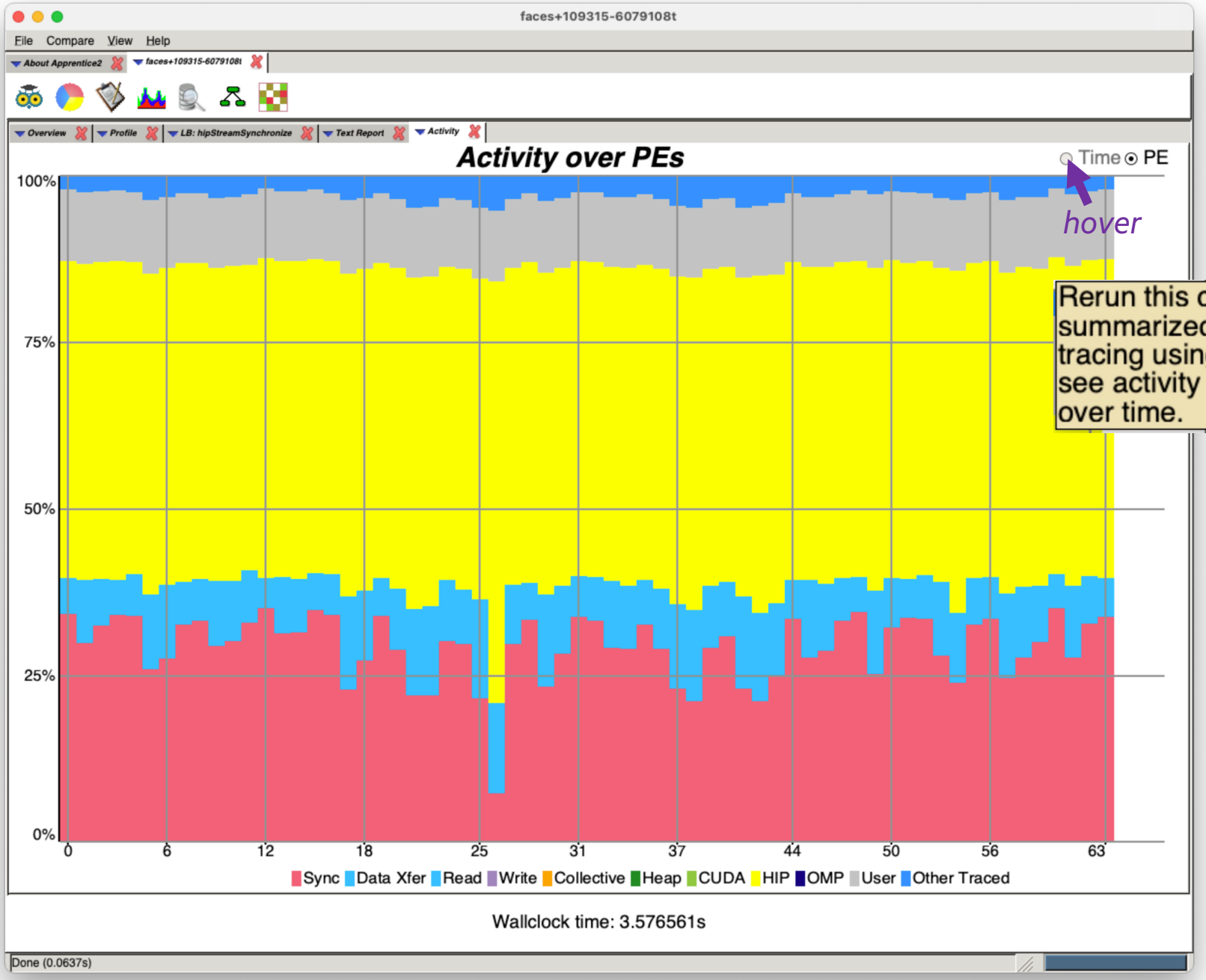



Comparison
File Compare View Help
About Apprentice2
Comparison
click
open MPI mosaic
faces
faces+15821-6079073t
Profile
6.92s
4.42s
hipStreamSynchronize
Mugs::share
All Others
MPI_Waitall
hipMemset
MPI_Waitall
Mugs::share
MPI_Barrier(sync)
hipStreamSynchronize
MPI_Isend
All Others
hipMemset
hipStreamCreate
Wallclock: 6.9156s / 4.4206s
faces+16018-6079073t (0.838s)


Comparison
File Compare View Help
About Apprentice2
Comparison
faces+16018-6079073t
faces+15821-6079073t
Profile
Mosaic
Avg Time (Send)
Destination PE
Source PE
4e-06
3.979e-05
7.2e-05
1e-06
1.942e-05
3.6e-05
1x1
Wallclock: 6.9156s / 4.4206s
Canceled.

PAT_RT_SUMMARY=0

or *Every Breath You Take*

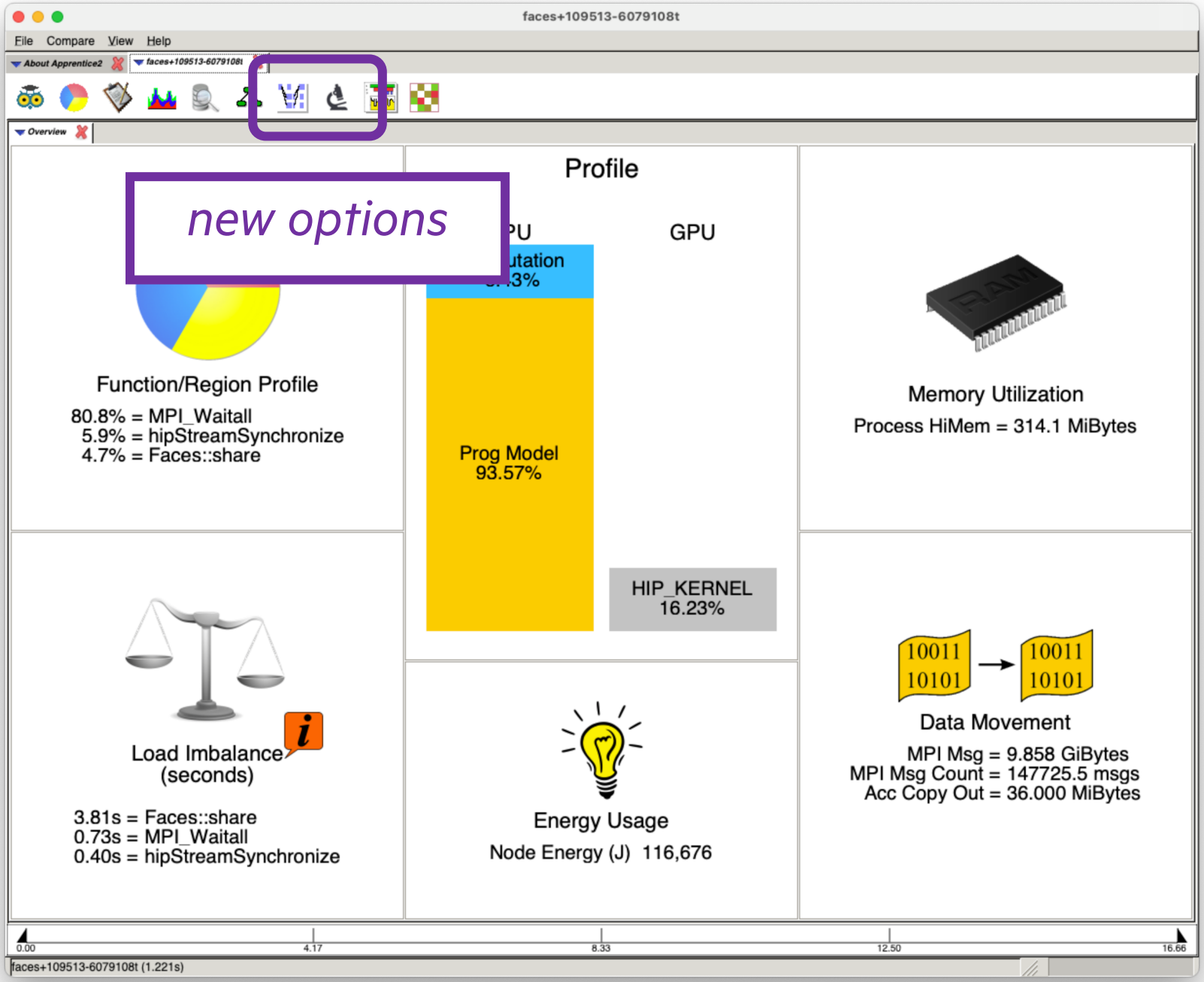


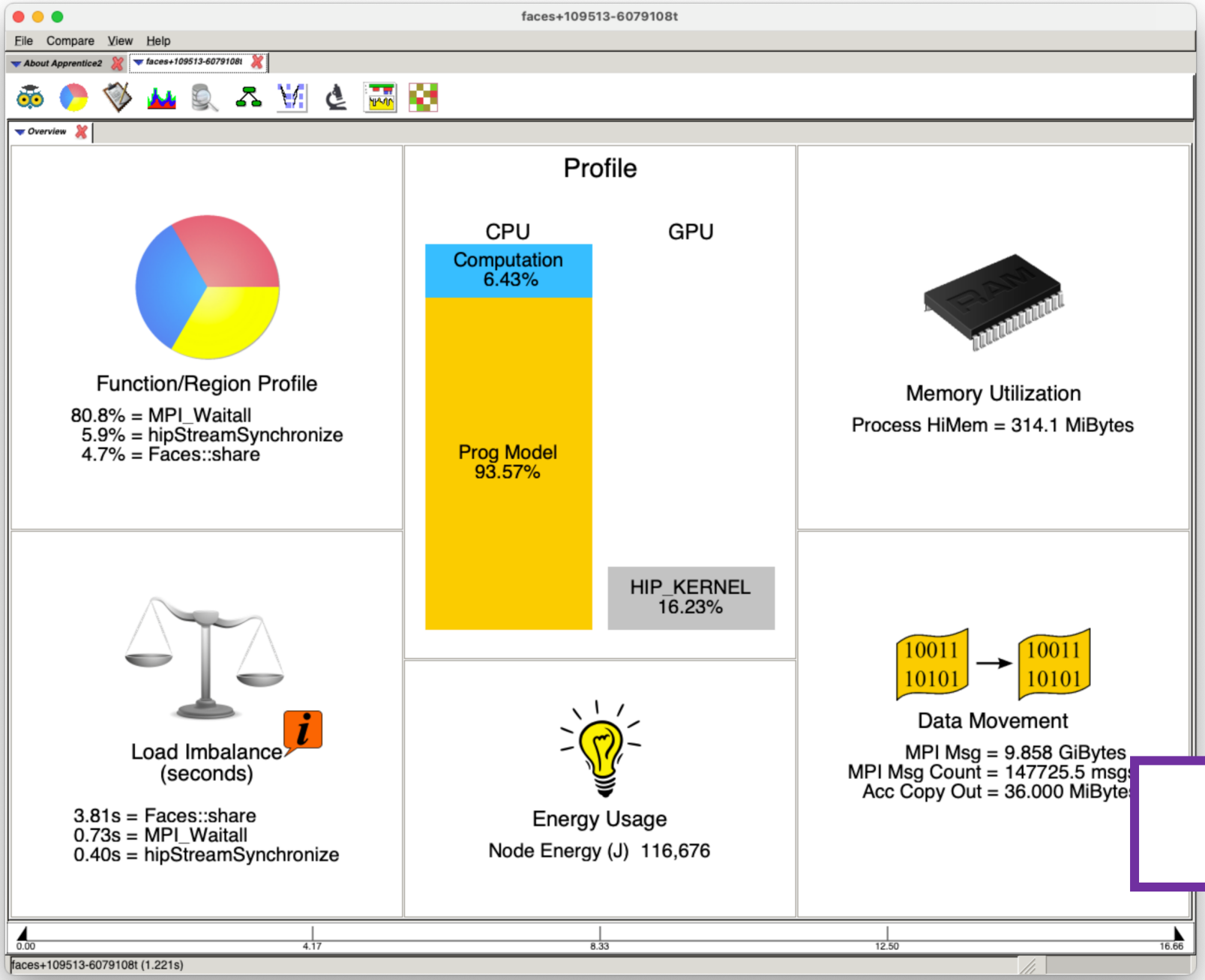
hover

Rerun this code with non-summarized tracing using perftools to see activity over time.

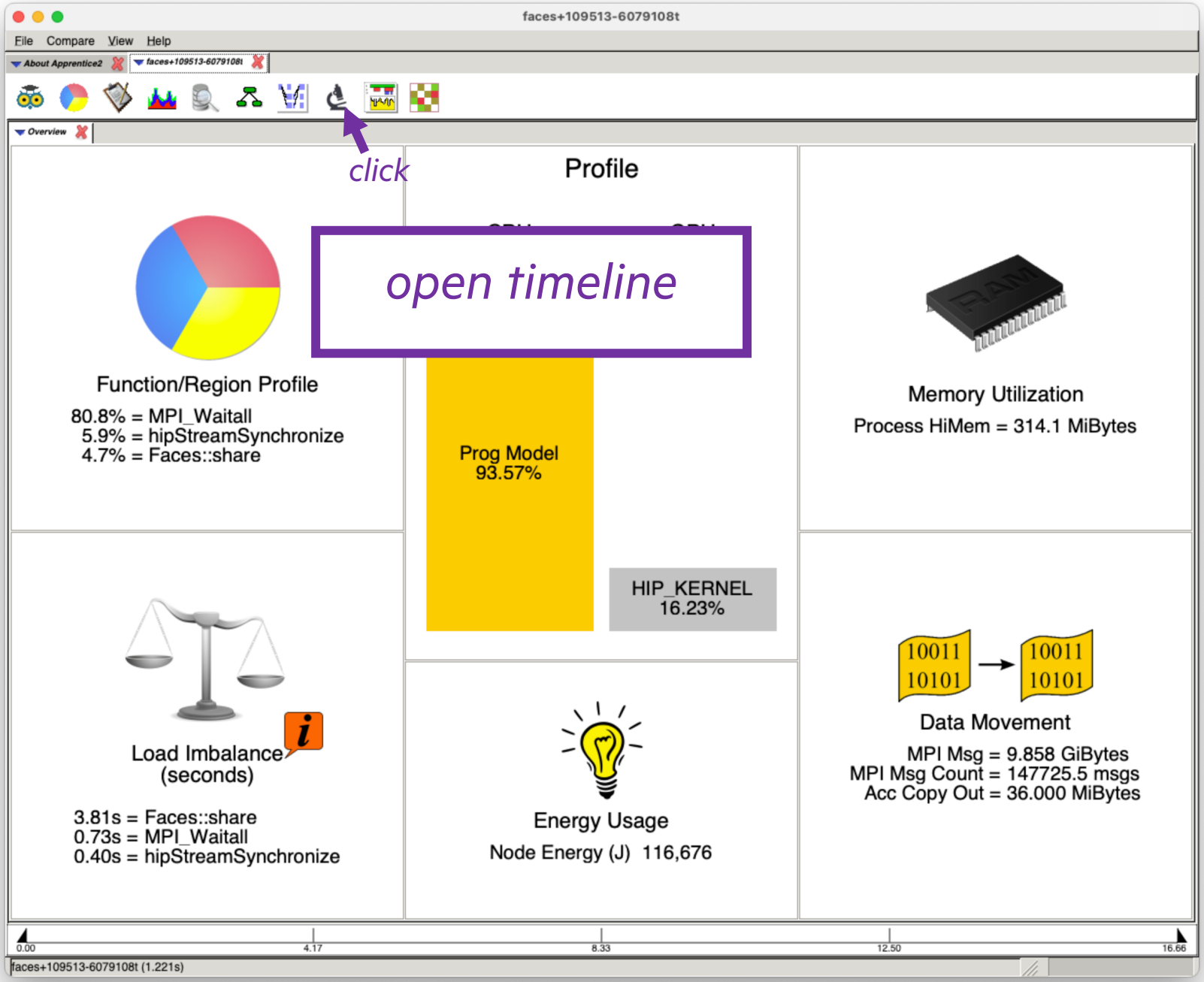
HIPCC RUN WITH NON-SUMMARIZED TRACING

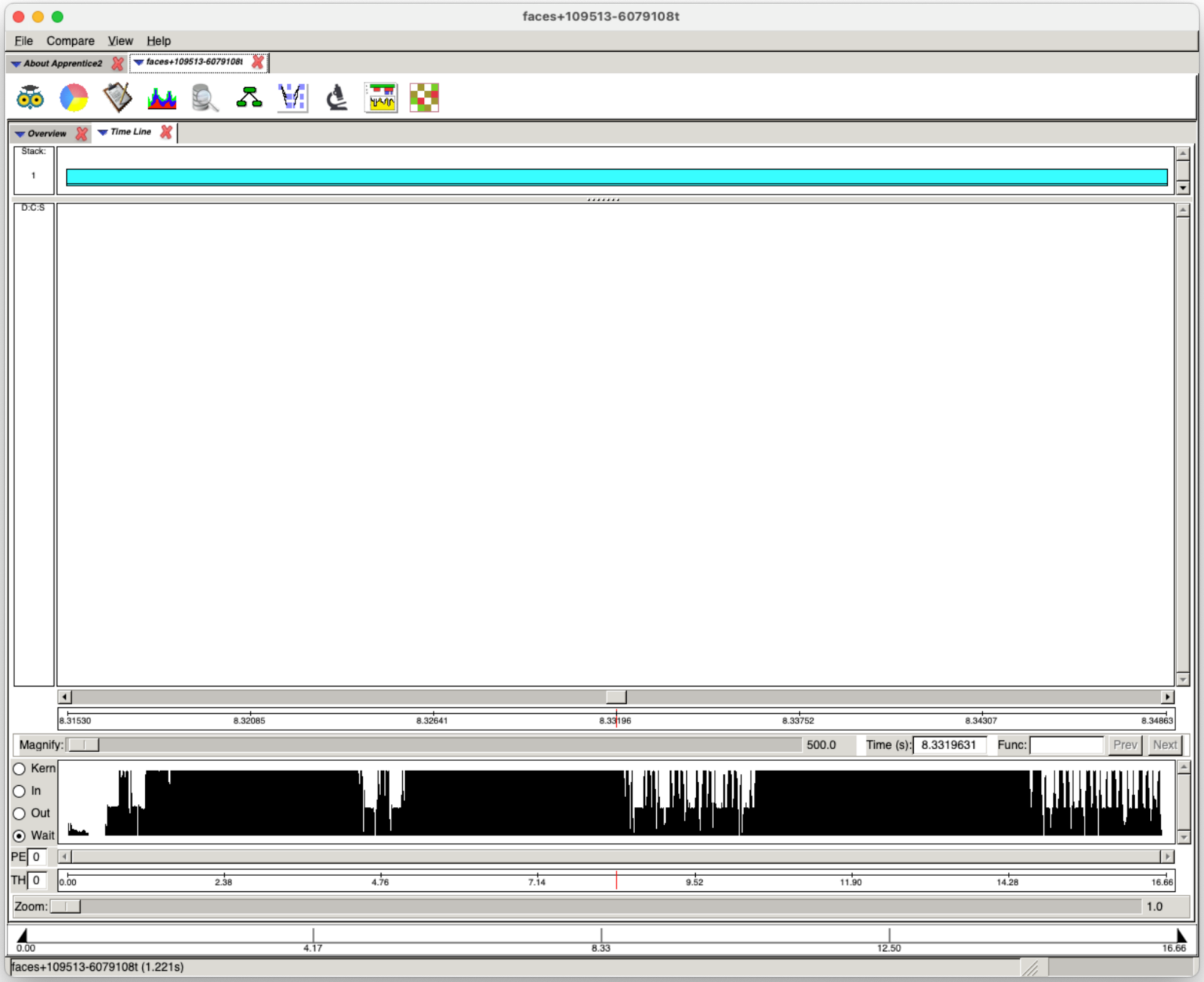
```
module load perftools-lite-gpu
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
export PAT_RT_SUMMARY=0
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces
```

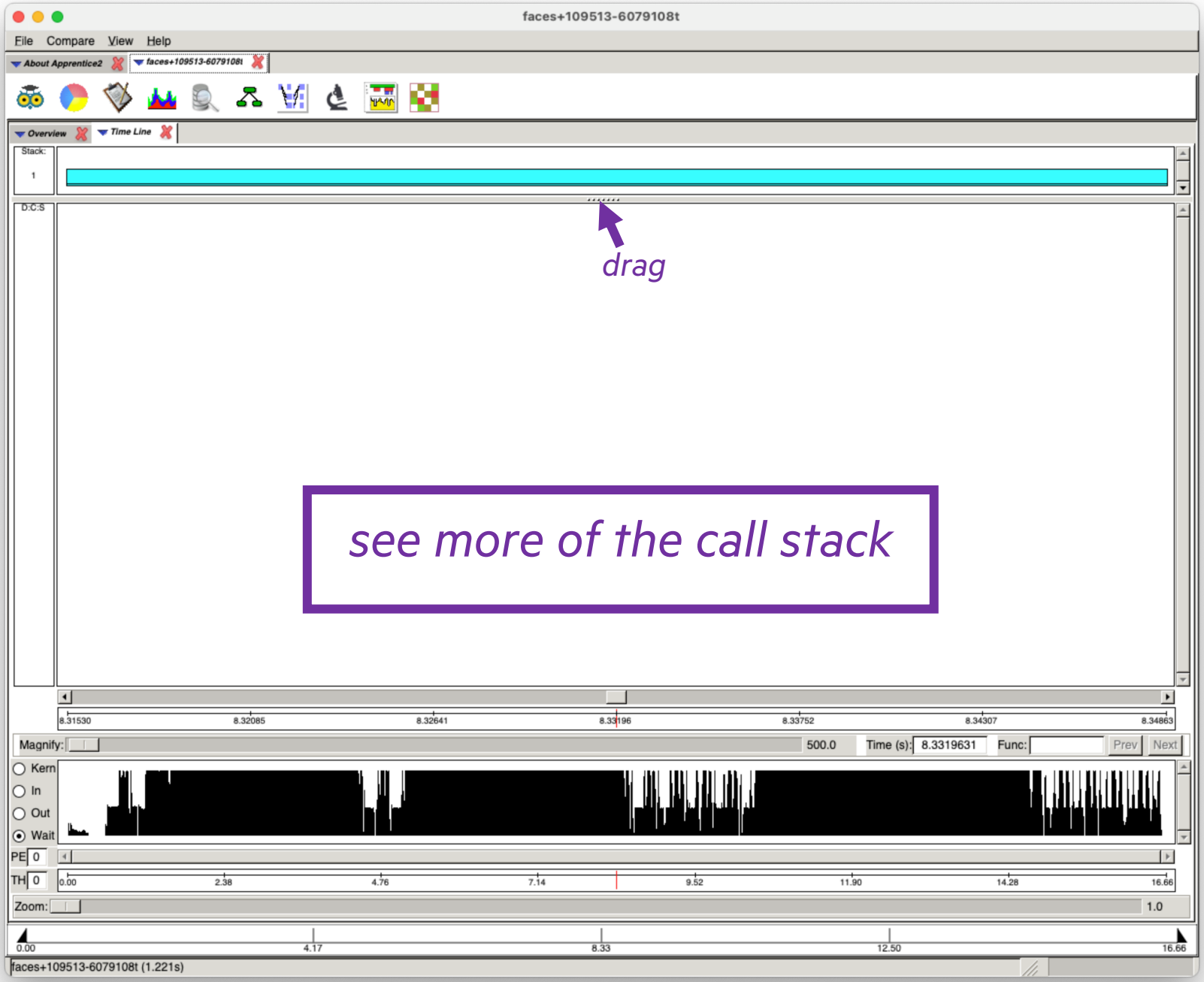

faces+109513-6079108t
File Compare View Help
About Apprentice2
faces+109513-6079108t
Overview
new options
Profile
GPU
Prog Model
93.57%
HIP_KERNEL
16.23%
Function/Region Profile
80.8% = MPI_Waitall
5.9% = hipStreamSynchronize
4.7% = Faces::share
Memory Utilization
Process HiMem = 314.1 MiBytes
Load Imbalance
(seconds)
3.81s = Faces::share
0.73s = MPI_Waitall
0.40s = hipStreamSynchronize
Energy Usage
Node Energy (J) 116,676
10011
10101
10011
10101
Data Movement
MPI Msg = 9.858 GiBytes
MPI Msg Count = 147725.5 msgs
Acc Copy Out = 36.000 MiBytes
0.00
4.17
8.33
12.50
16.66
faces+109513-6079108t (1.221s)

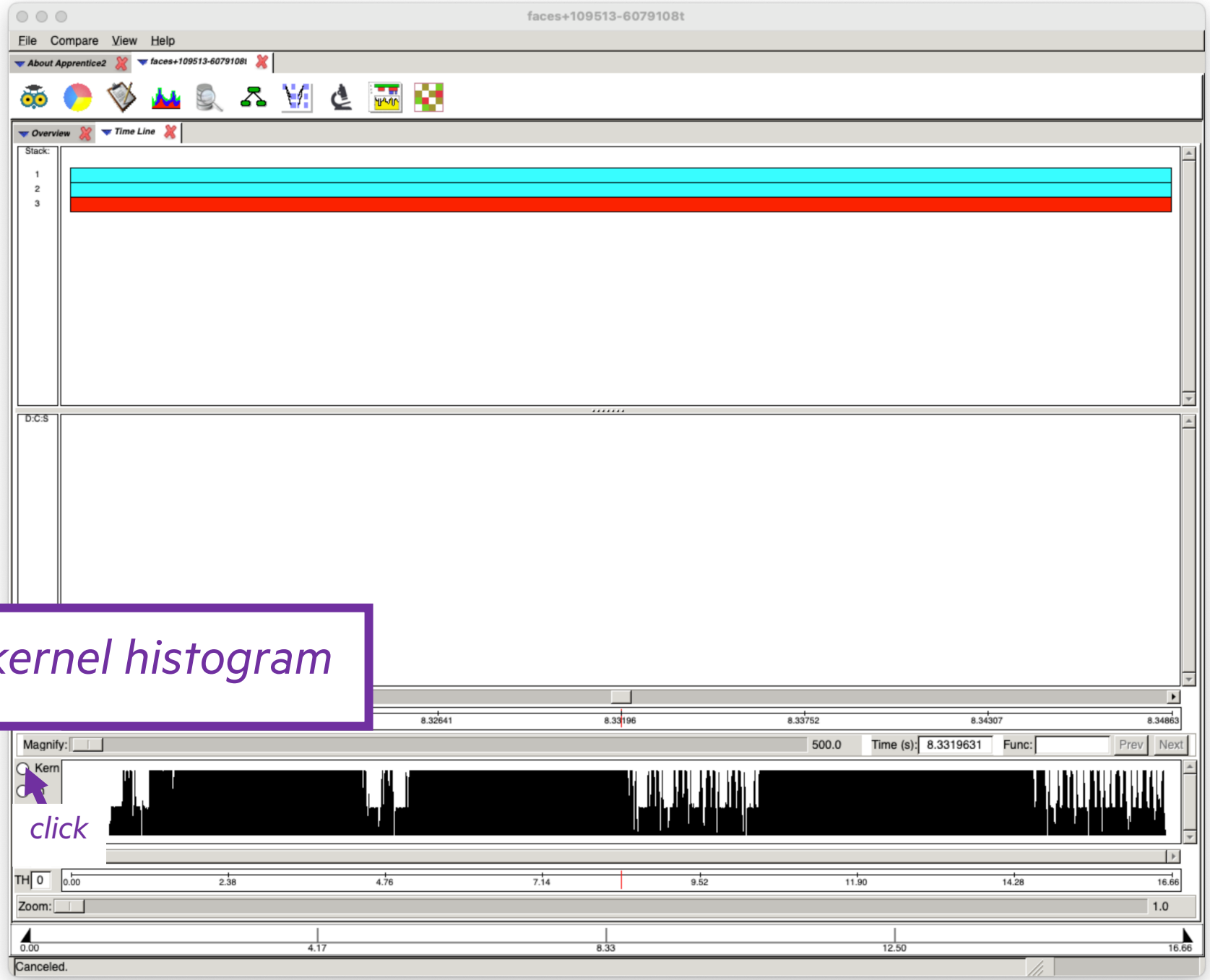

faces+109513-6079108t
File Compare View Help
About Apprentice2
faces+109513-6079108t
Overview
Profile
CPU
GPU
Computation
6.43%
Prog Model
93.57%
HIP_KERNEL
16.23%
Function/Region Profile
80.8% = MPI_Waitall
5.9% = hipStreamSynchronize
4.7% = Faces::share
Memory Utilization
Process HiMem = 314.1 MiBytes
Load Imbalance
(seconds)
3.81s = Faces::share
0.73s = MPI_Waitall
0.40s = hipStreamSynchronize
Energy Usage
Node Energy (J) 116,676
10011
10101
10011
10101
Data Movement
MPI Msg = 9.858 GiBytes
MPI Msg Count = 147725.5 msgs
Acc Copy Out = 36.000 MiBytes
0.00
4.17
8.33
12.50
16.66
faces+109513-6079108t (1.221s)

"calipers"

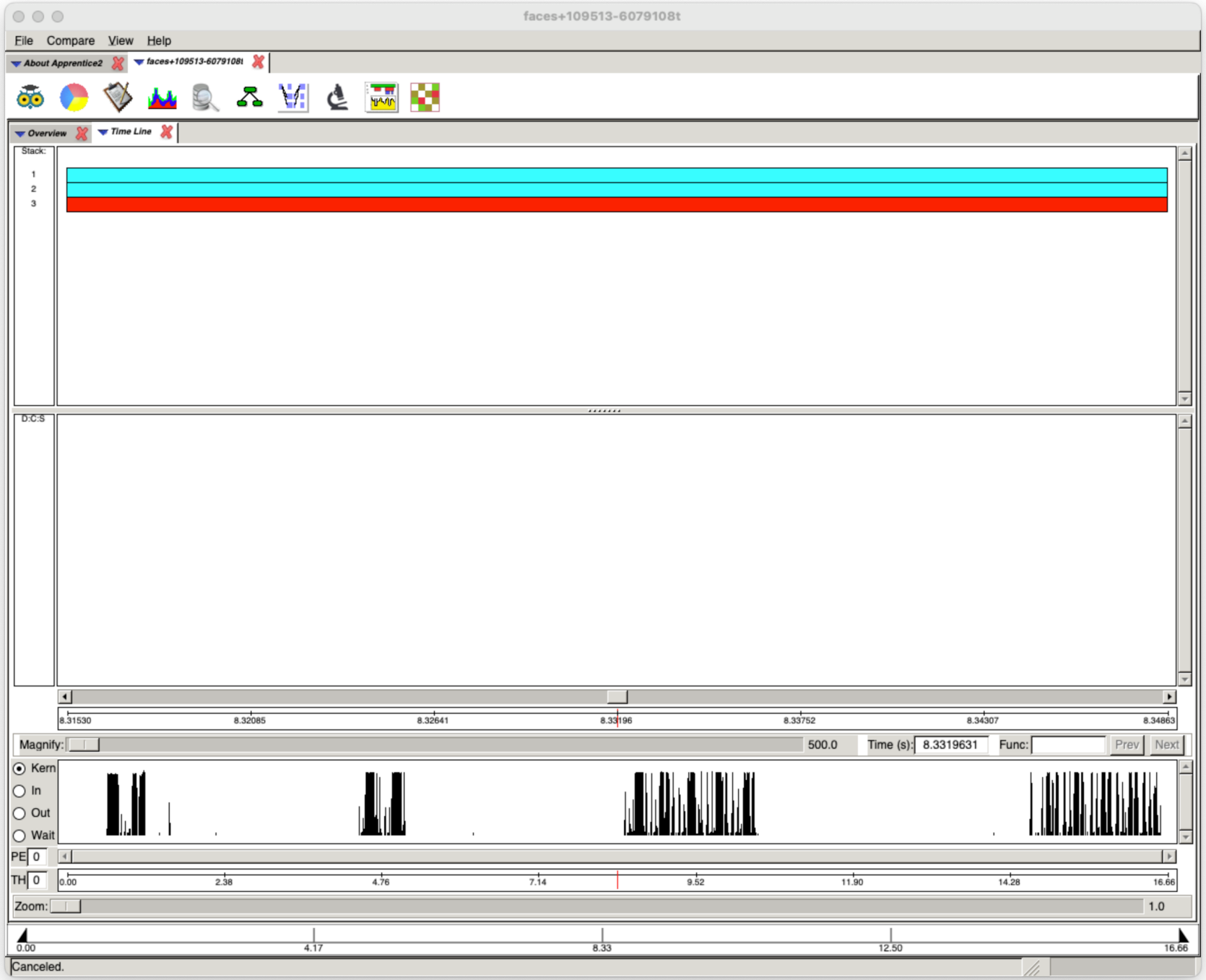

faces+109513-6079108t
File Compare View Help
About Apprentice2
faces+109513-6079108t
Overview
click
Profile
open timeline
Function/Region Profile
80.8% = MPI_Waitall
5.9% = hipStreamSynchronize
4.7% = Faces::share
Prog Model
93.57%
HIP_KERNEL
16.23%
Memory Utilization
Process HiMem = 314.1 MiBytes
Load Imbalance
(seconds)
3.81s = Faces::share
0.73s = MPI_Waitall
0.40s = hipStreamSynchronize
Energy Usage
Node Energy (J) 116,676
10011
10101
10011
10101
Data Movement
MPI Msg = 9.858 GiBytes
MPI Msg Count = 147725.5 msgs
Acc Copy Out = 36.000 MiBytes
0.00
4.17
8.33
12.50
16.66
faces+109513-6079108t (1.221s)


faces+109513-6079108t
File Compare View Help
About Apprentice2
faces+109513-6079108t
Overview
Time Line
Stack:
1
D:C:S
8.31530
8.32085
8.32641
8.33196
8.33752
8.34307
8.34863
Magnify:
500.0
Time (s):
8.3319631
Func:
Prev
Next
Kern
In
Out
Wait
PE
0
TH
0
0.00
2.38
4.76
7.14
9.52
11.90
14.28
16.66
Zoom:
1.0
0.00
4.17
8.33
12.50
16.66
faces+109513-6079108t (1.221s)


faces+109513-6079108t
File Compare View Help
About Apprentice2
faces+109513-6079108t
Overview
Time Line
Stack:
1
D:C:S
drag
see more of the call stack
8.31530
8.32085
8.32641
8.33196
8.33752
8.34307
8.34863
Magnify:
500.0
Time (s):
8.3319631
Func:
Prev
Next
Kern
In
Out
Wait
PE
0
TH
0
0.00
2.38
4.76
7.14
9.52
11.90
14.28
16.66
Zoom:
1.0
0.00
4.17
8.33
12.50
16.66
faces+109513-6079108t (1.221s)



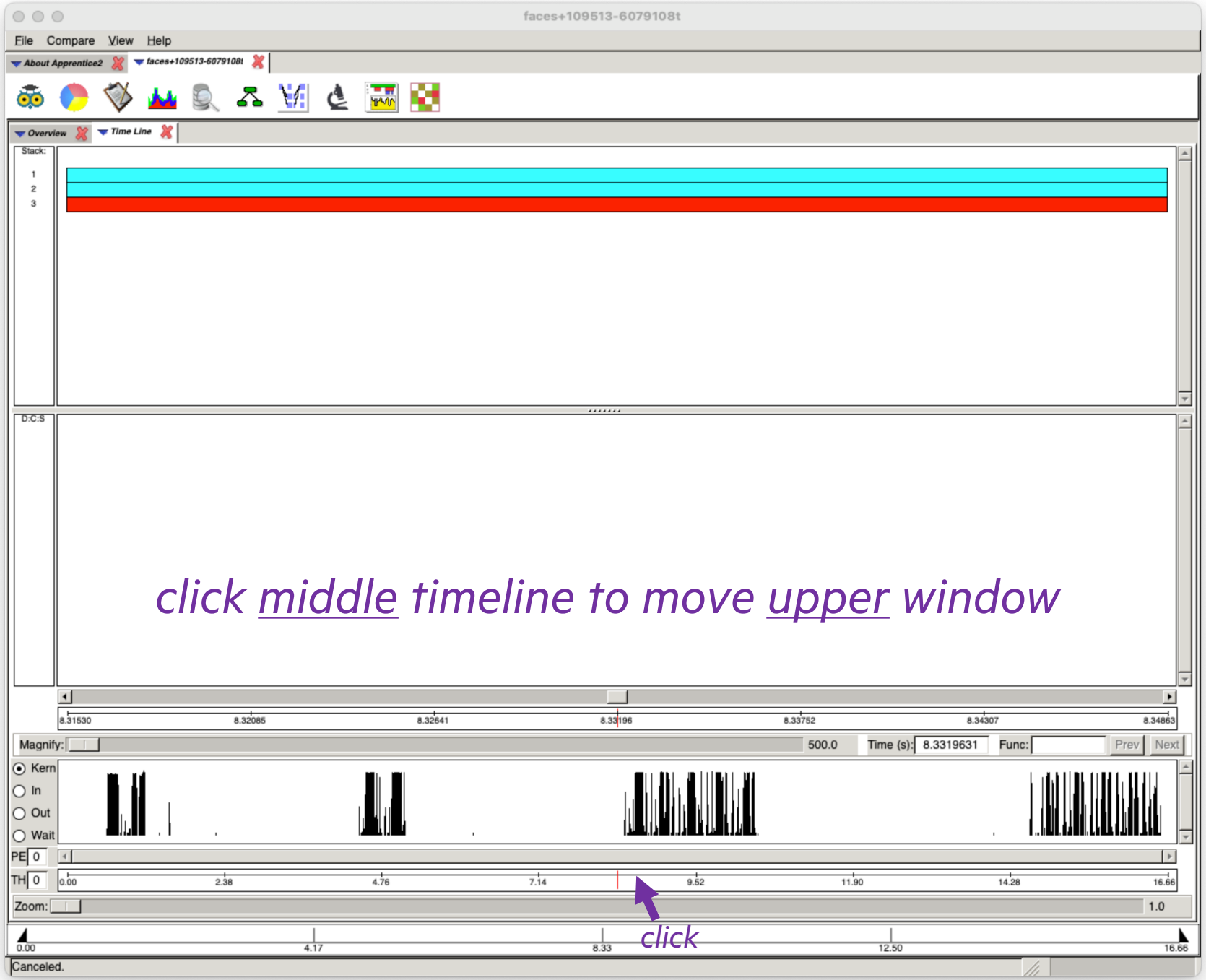
switch to kernel histogram

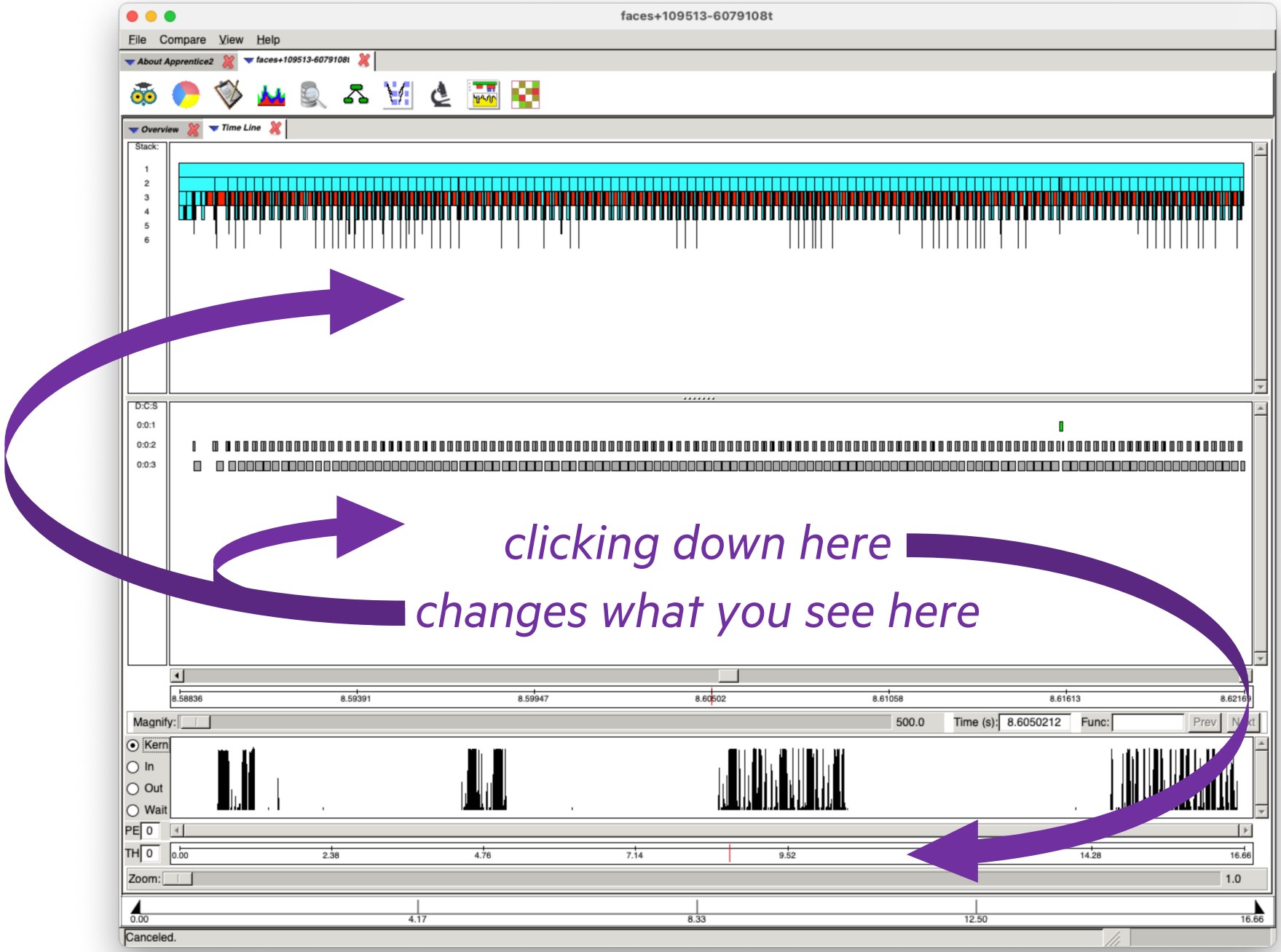


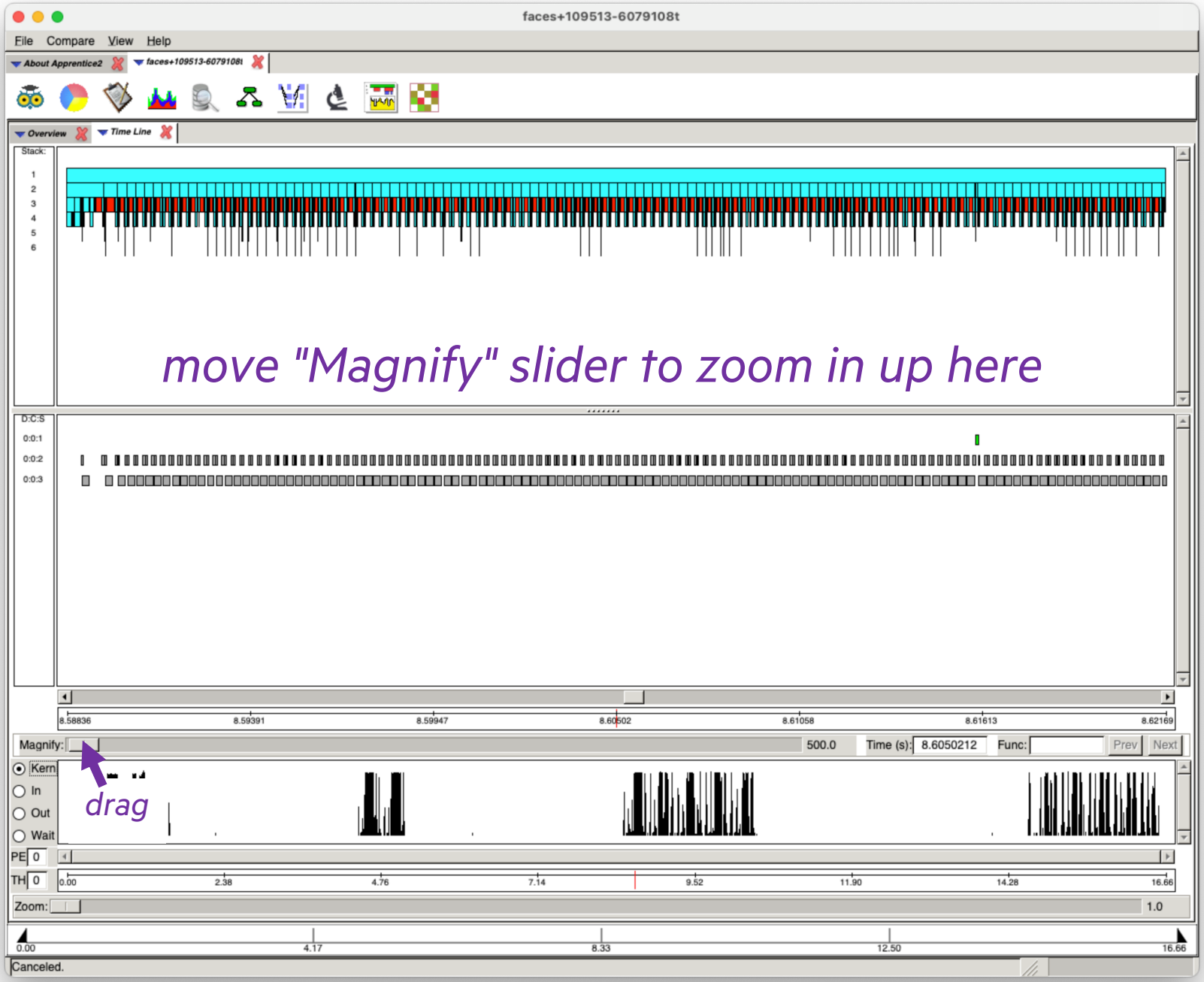
"magnify"

"zoom"

"calipers"

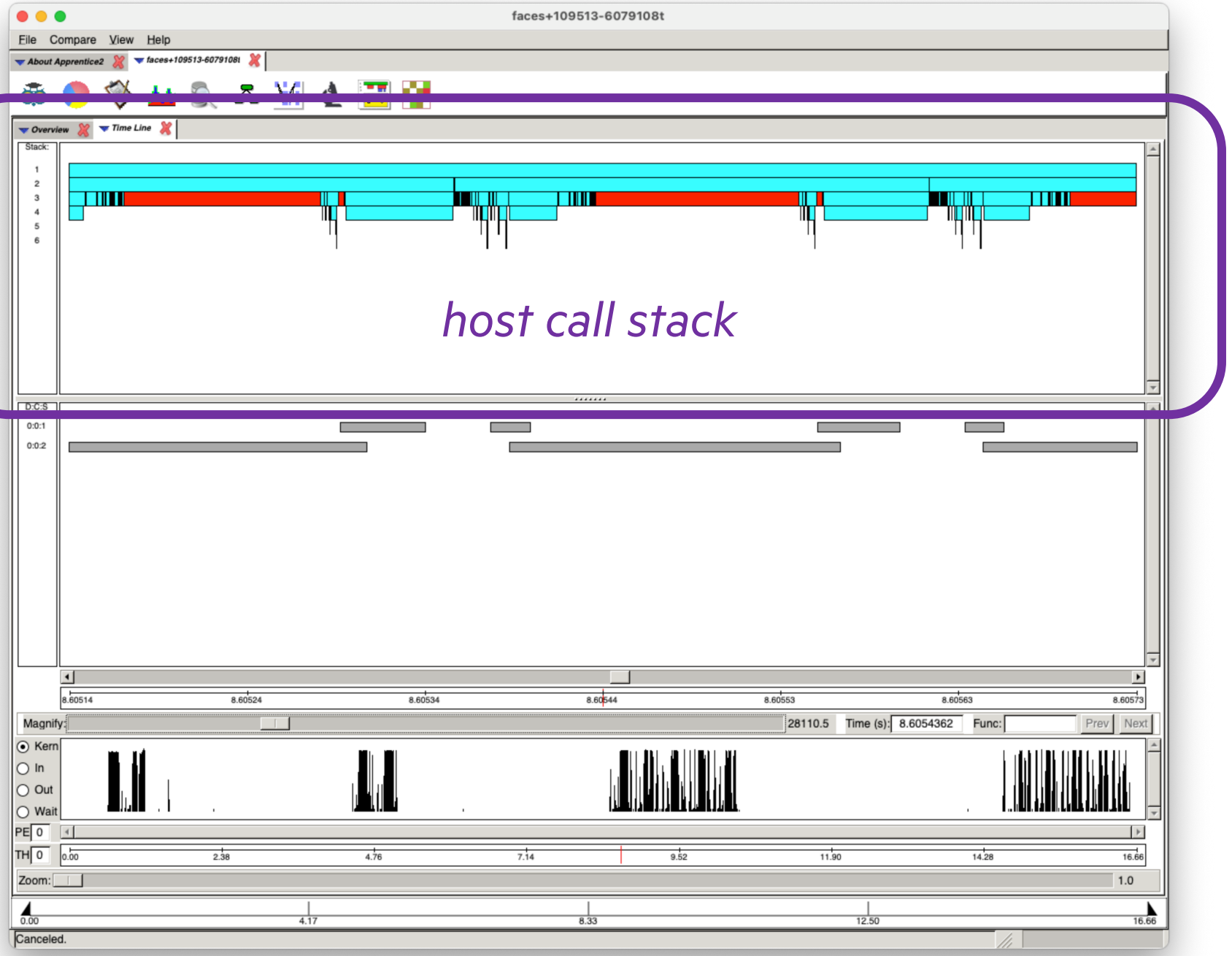

faces+109513-6079108t
File Compare View Help
About Apprentice2
faces+109513-6079108t
Overview
Time Line
Stack:
1
2
3
D:C:S
click middle timeline to move upper window
8.31530
8.32085
8.32641
8.33196
8.33752
8.34307
8.34863
Magnify:
500.0
Time (s):
8.3319631
Func:
Prev
Next
Kern
In
Out
Wait
PE
0
TH
0
0.00
2.38
4.76
7.14
9.52
11.90
14.28
16.66
Zoom:
1.0
click
0.00
4.17
8.33
12.50
16.66
Canceled.

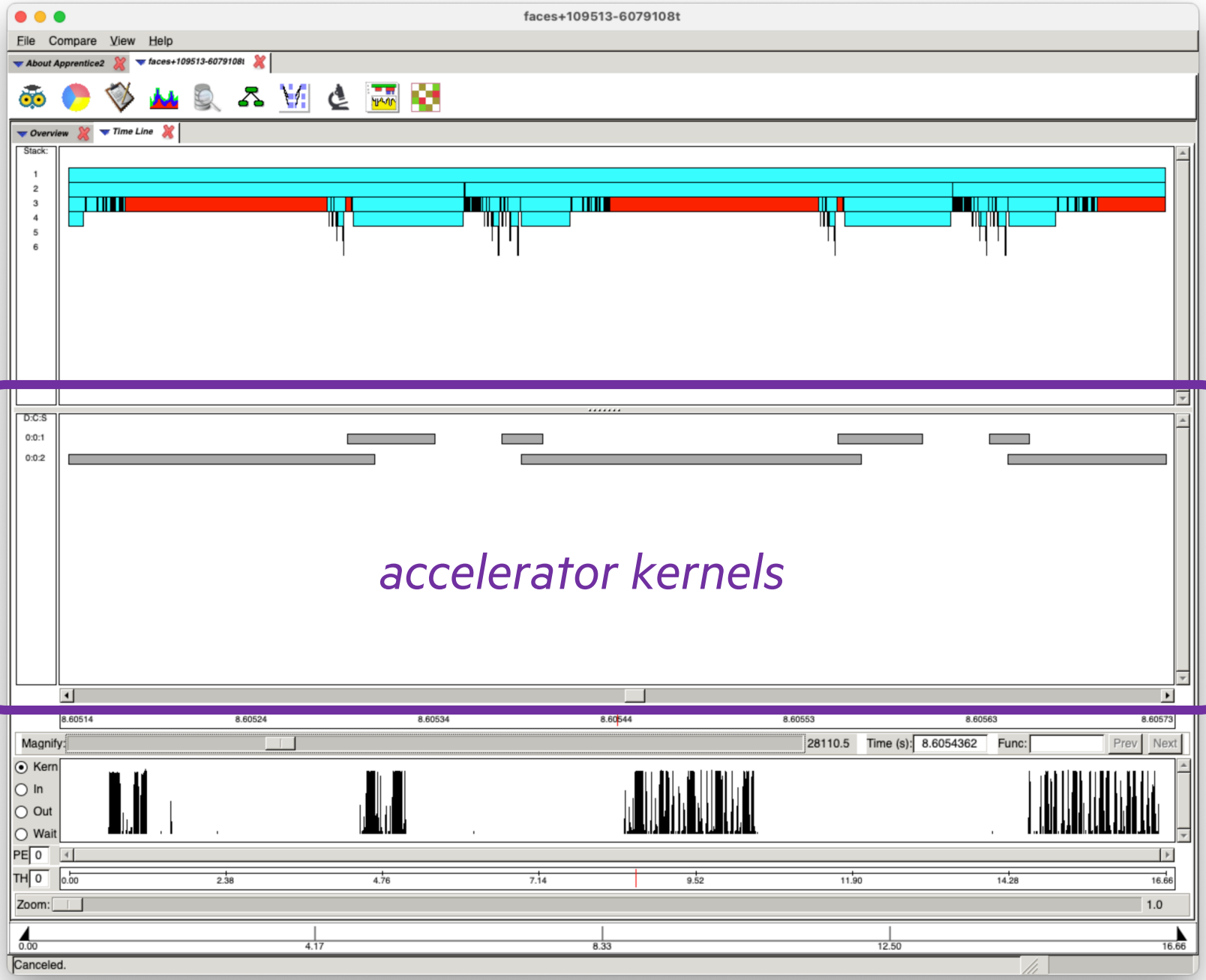

faces+109513-6079108t
File Compare View Help
About Apprentice2
faces+109513-6079108t
Overview
Time Line
Stack:
D:C:S
0:0:1
0:0:2
0:0:3
clicking down here
changes what you see here
Magnify:
500.0
Time (s):
8.6050212
Func:
Prev
Next
Kern
In
Out
Wait
PE
TH
Zoom:
1.0
Canceled.


faces+109513-6079108t
File Compare View Help
About Apprentice2
faces+109513-6079108t
Overview
Time Line
Stack:
move "Magnify" slider to zoom in up here
D:C:S
0:0:1
0:0:2
0:0:3
8.58836
8.59391
8.59947
8.60502
8.61058
8.61613
8.62169
Magnify:
500.0
Time (s):
8.6050212
Func:
Prev
Next
Kern
In
Out
Wait
drag
PE
TH
0.00
2.38
4.76
7.14
9.52
11.90
14.28
16.66
Zoom:
1.0
0.00
4.17
8.33
12.50
16.66
Canceled.

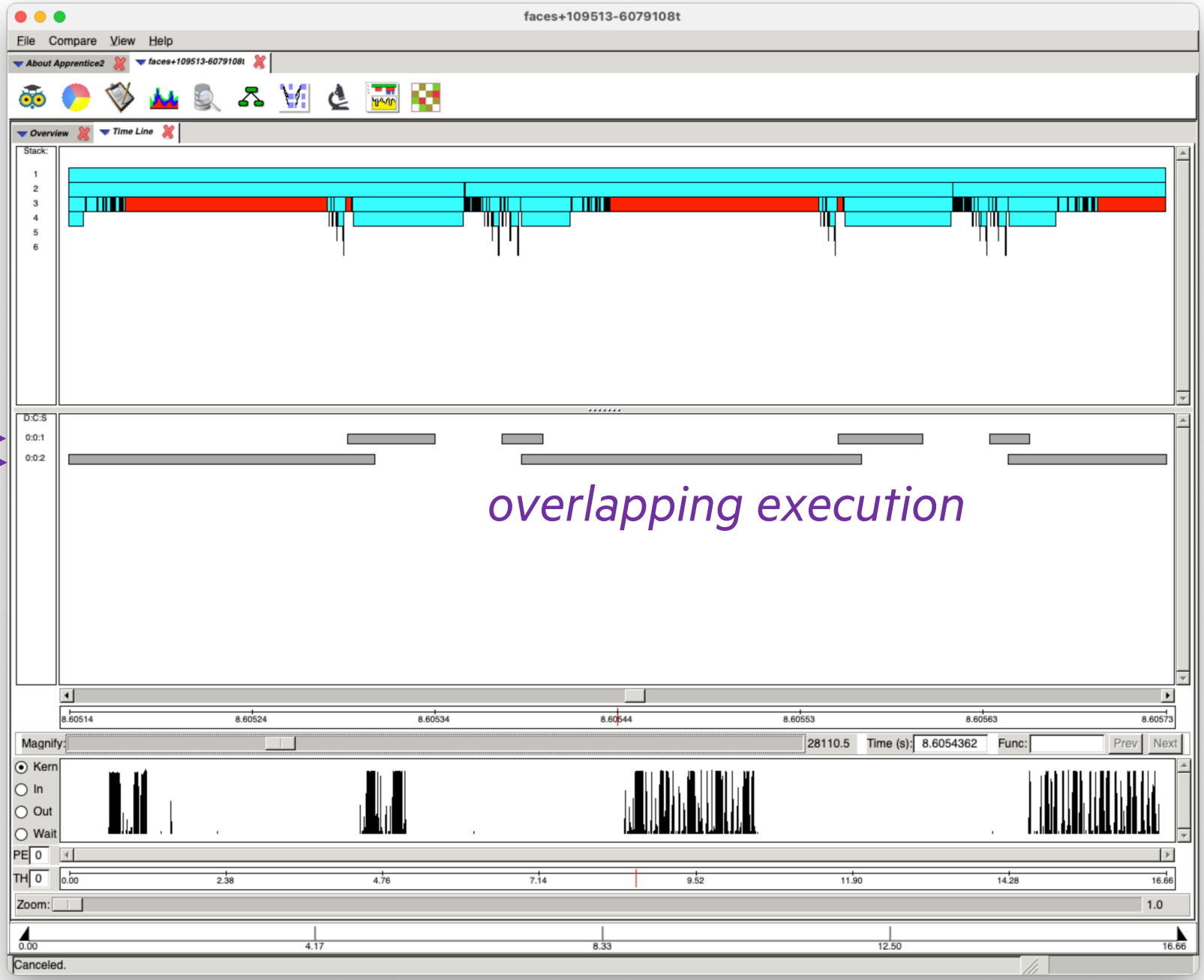
zoomed in

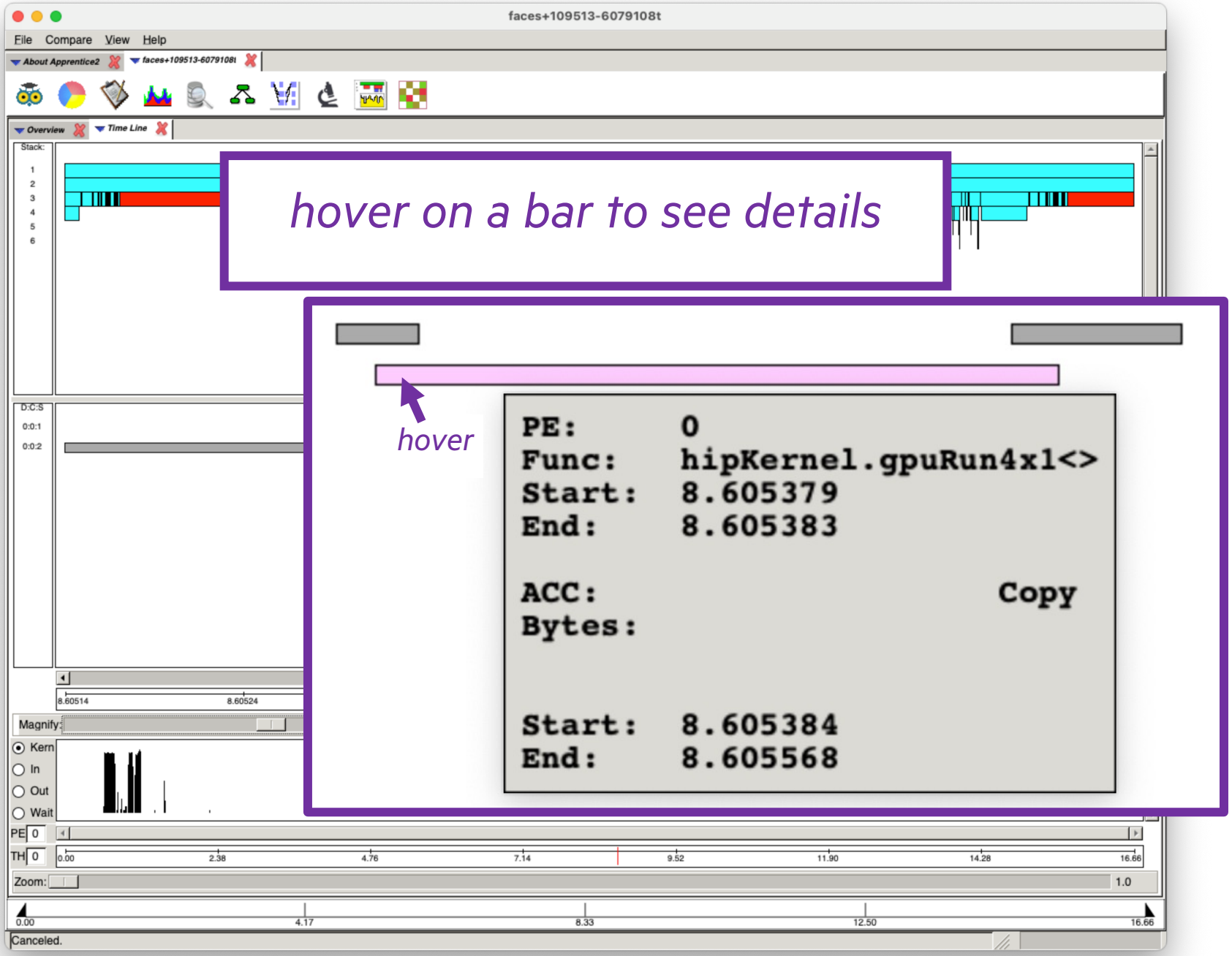
full timeline

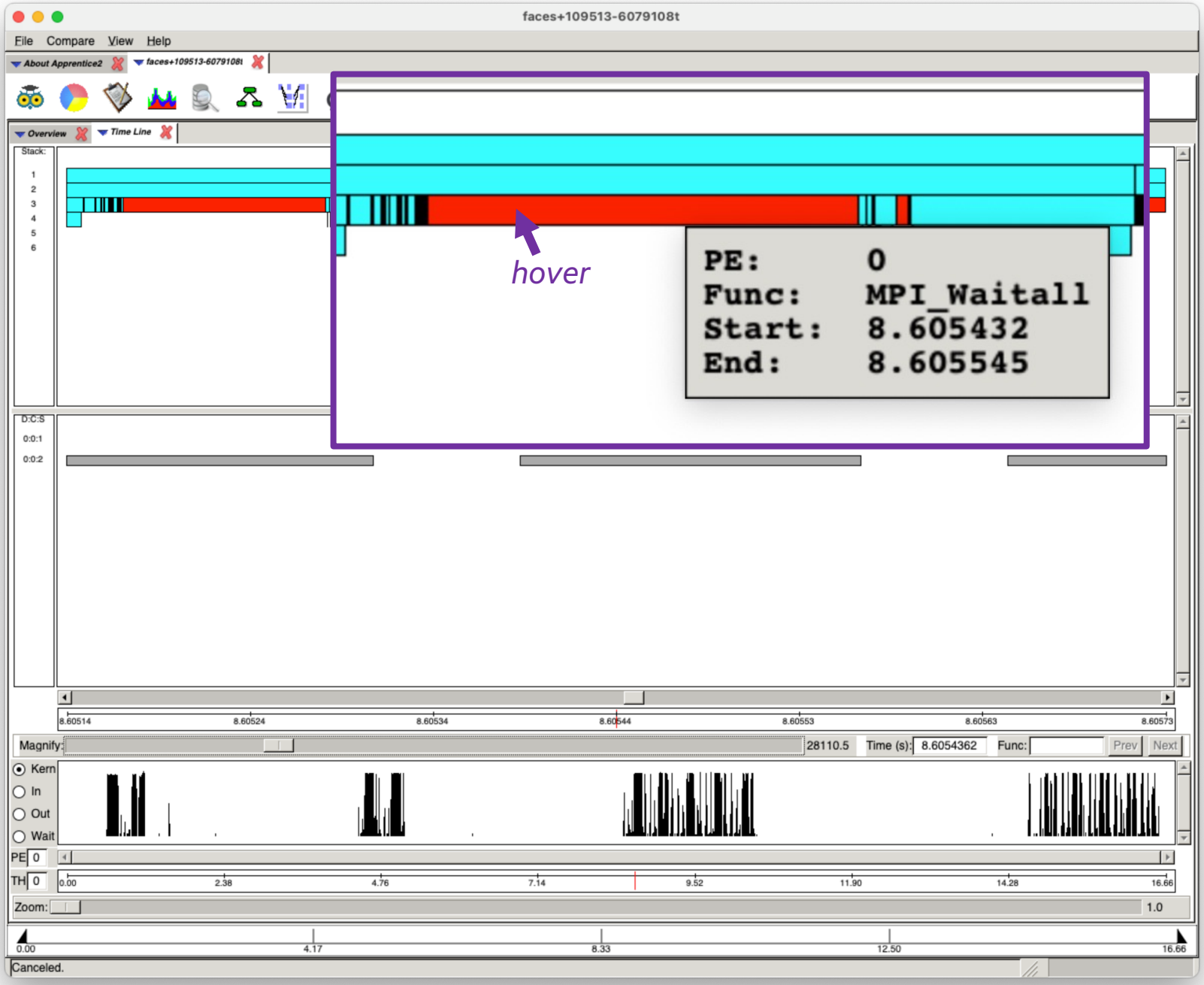

faces+109513-6079108t
File Compare View Help
About Apprentice2
faces+109513-6079108t
Overview
Time Line
Stack:
1
2
3
4
5
6
host call stack
D:C:S
0:0:1
0:0:2
8.60514
8.60524
8.60534
8.60544
8.60553
8.60563
8.60573
Magnify:
28110.5
Time (s):
8.6054362
Func:
Prev
Next
Kern
In
Out
Wait
PE
0
TH
0
0.00
2.38
4.76
7.14
9.52
11.90
14.28
16.66
Zoom:
1.0
0.00
4.17
8.33
12.50
16.66
Canceled.


faces+109513-6079108t
File Compare View Help
About Apprentice2
faces+109513-6079108t
Overview
Time Line
Stack:
1
2
3
4
5
6
D:C:S
0:0:1
0:0:2
accelerator kernels
8.60514
8.60524
8.60534
8.60544
8.60553
8.60563
8.60573
Magnify:
28110.5
Time (s):
8.6054362
Func:
Prev
Next
Kern
In
Out
Wait
PE
0
TH
0
0.00
2.38
4.76
7.14
9.52
11.90
14.28
16.66
Zoom:
1.0
0.00
4.17
8.33
12.50
16.66
Canceled.

different streams

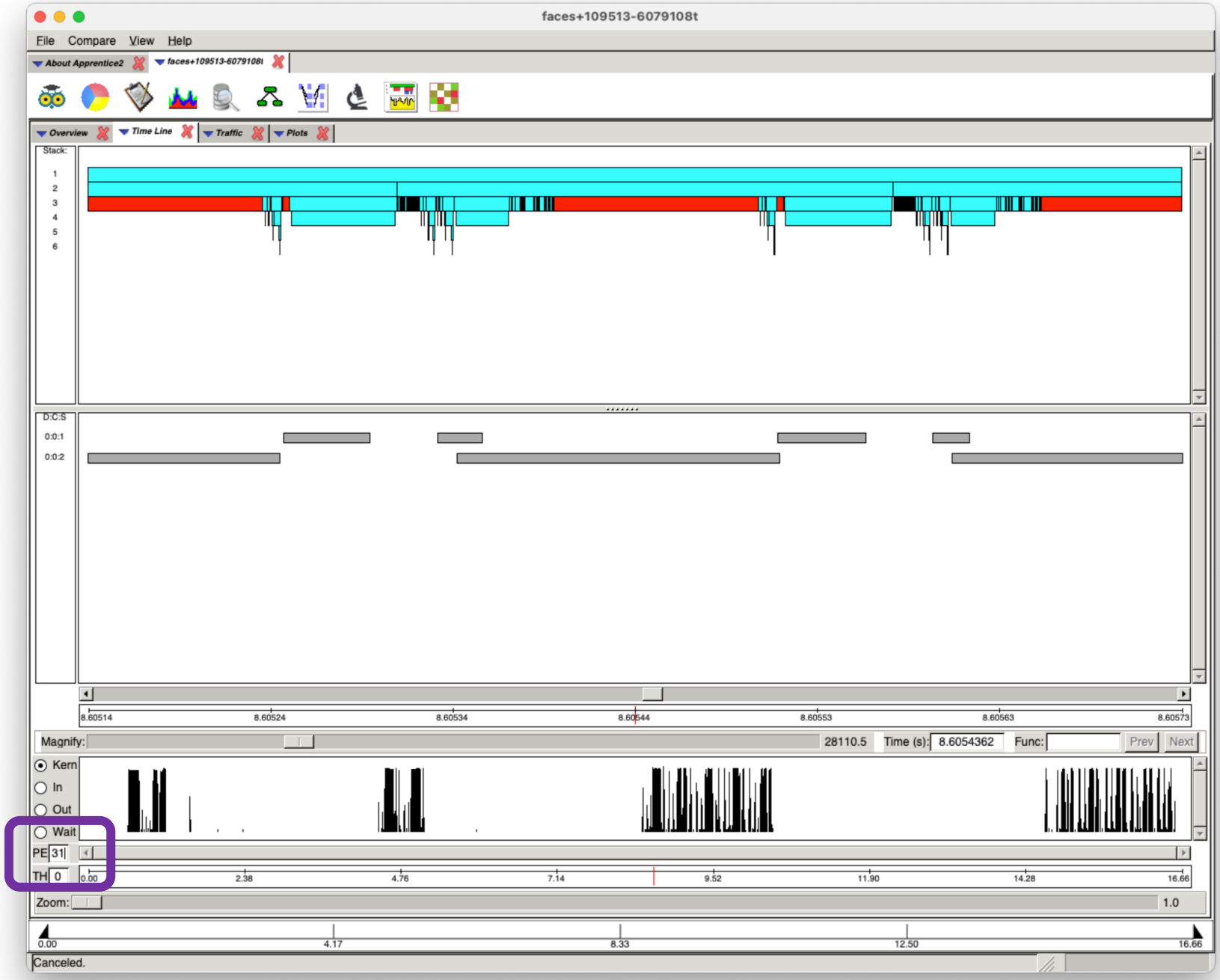


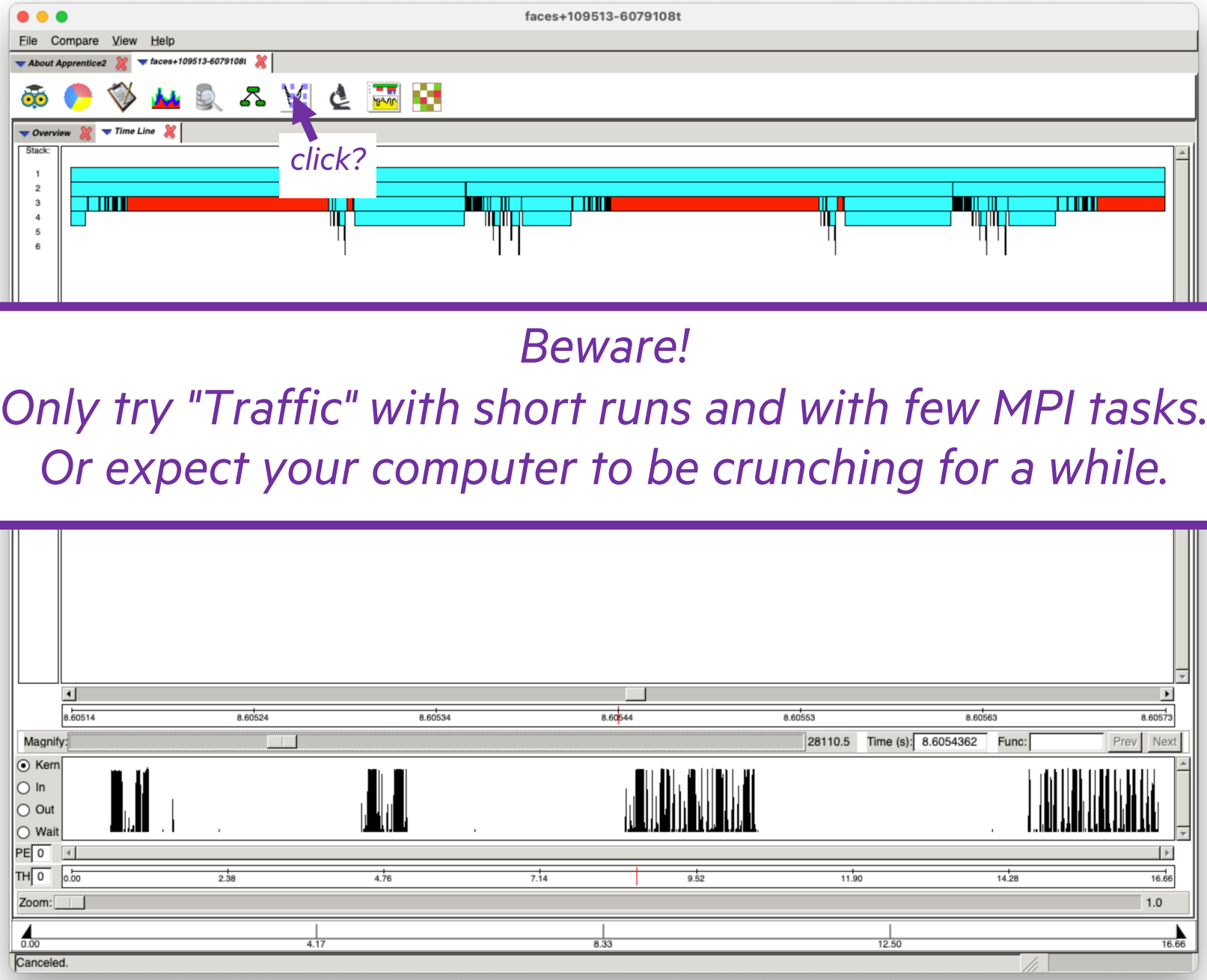

faces+109513-6079108t
File Compare View Help
About Apprentice2
faces+109513-6079108t
Overview
Time Line
Stack:
hover on a bar to see details
hover
PE: 0
Func: hipKernel.gpuRun4x1<>
Start: 8.605379
End: 8.605383
ACC: Copy
Bytes:
Start: 8.605384
End: 8.605568
Magnify:
Kern
In
Out
Wait
PE 0
TH 0
Zoom: 1.0
Canceled.


faces+109513-6079108t
File Compare View Help
About Apprentice2
faces+109513-6079108t
Overview
Time Line
Stack:
hover
PE: 0
Func: MPI_Waitall
Start: 8.605432
End: 8.605545
D:C:S
0:0:1
0:0:2
8.60514
8.60524
8.60534
8.60544
8.60553
8.60563
8.60573
Magnify:
28110.5
Time (s):
8.6054362
Func:
Prev
Next
Kern
In
Out
Wait
PE 0
TH 0
0.00
2.38
4.76
7.14
9.52
11.90
14.28
16.66
Zoom:
1.0
0.00
4.17
8.33
12.50
16.66
Canceled.

default task is MPI rank 0

change to another rank

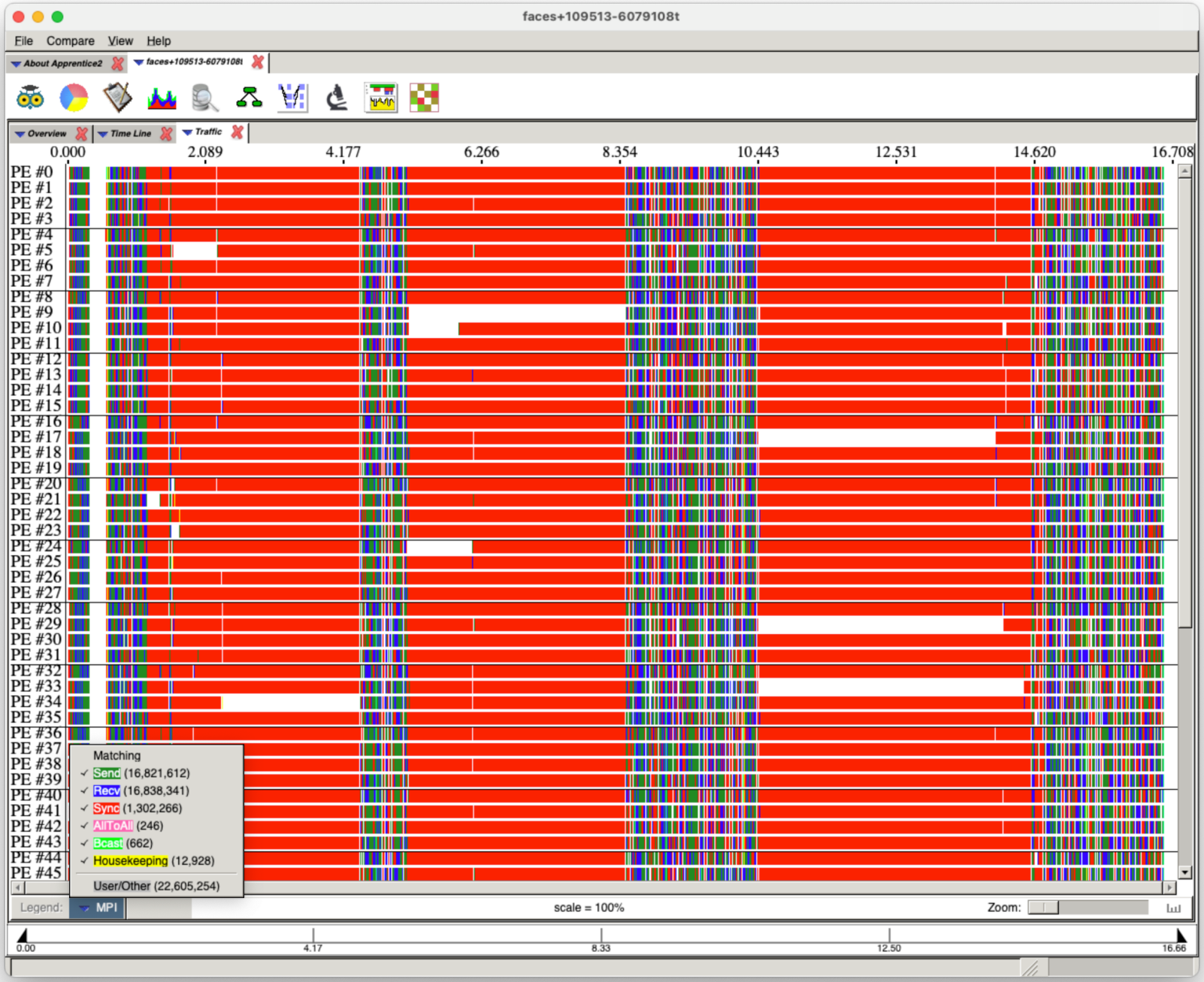


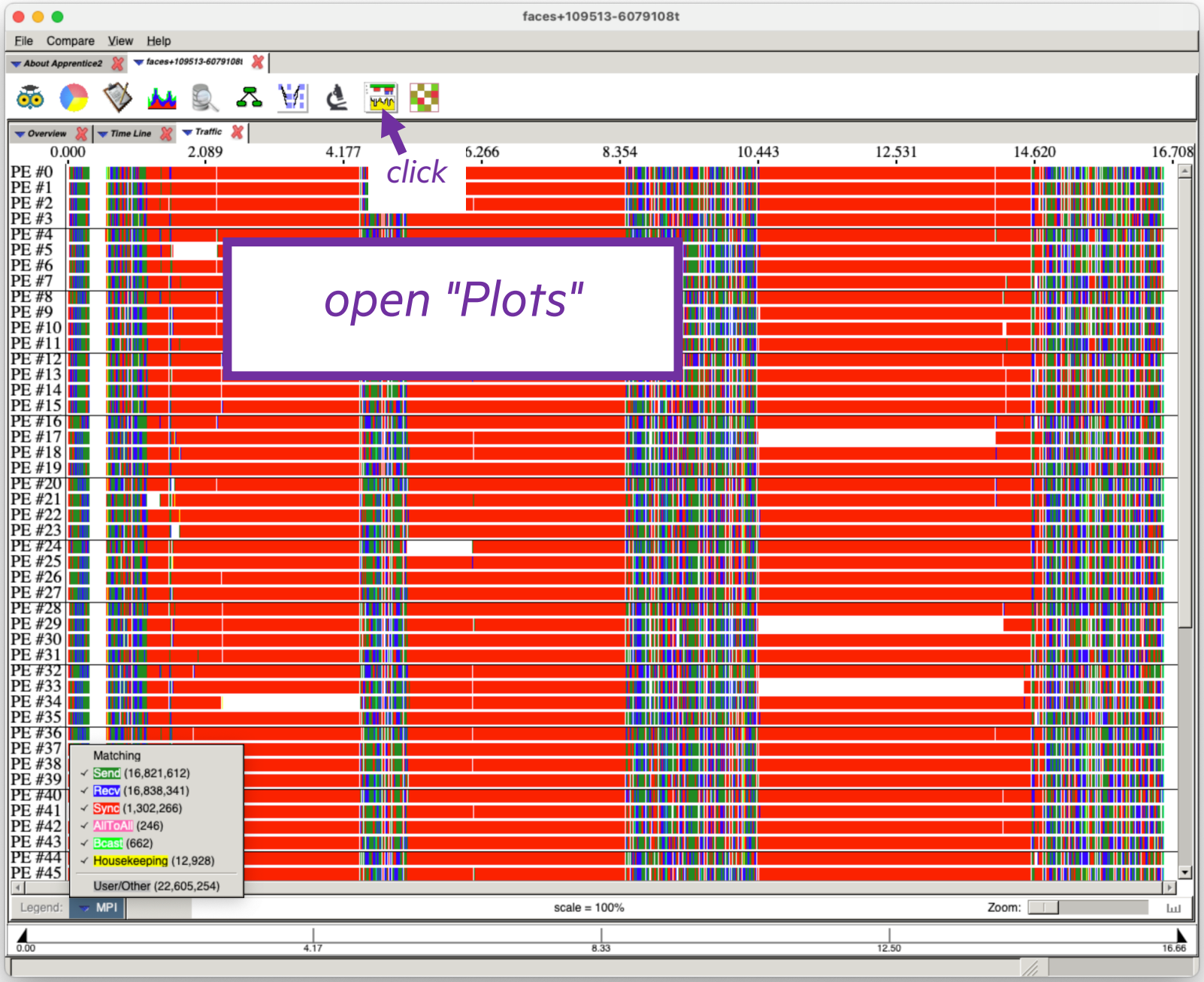


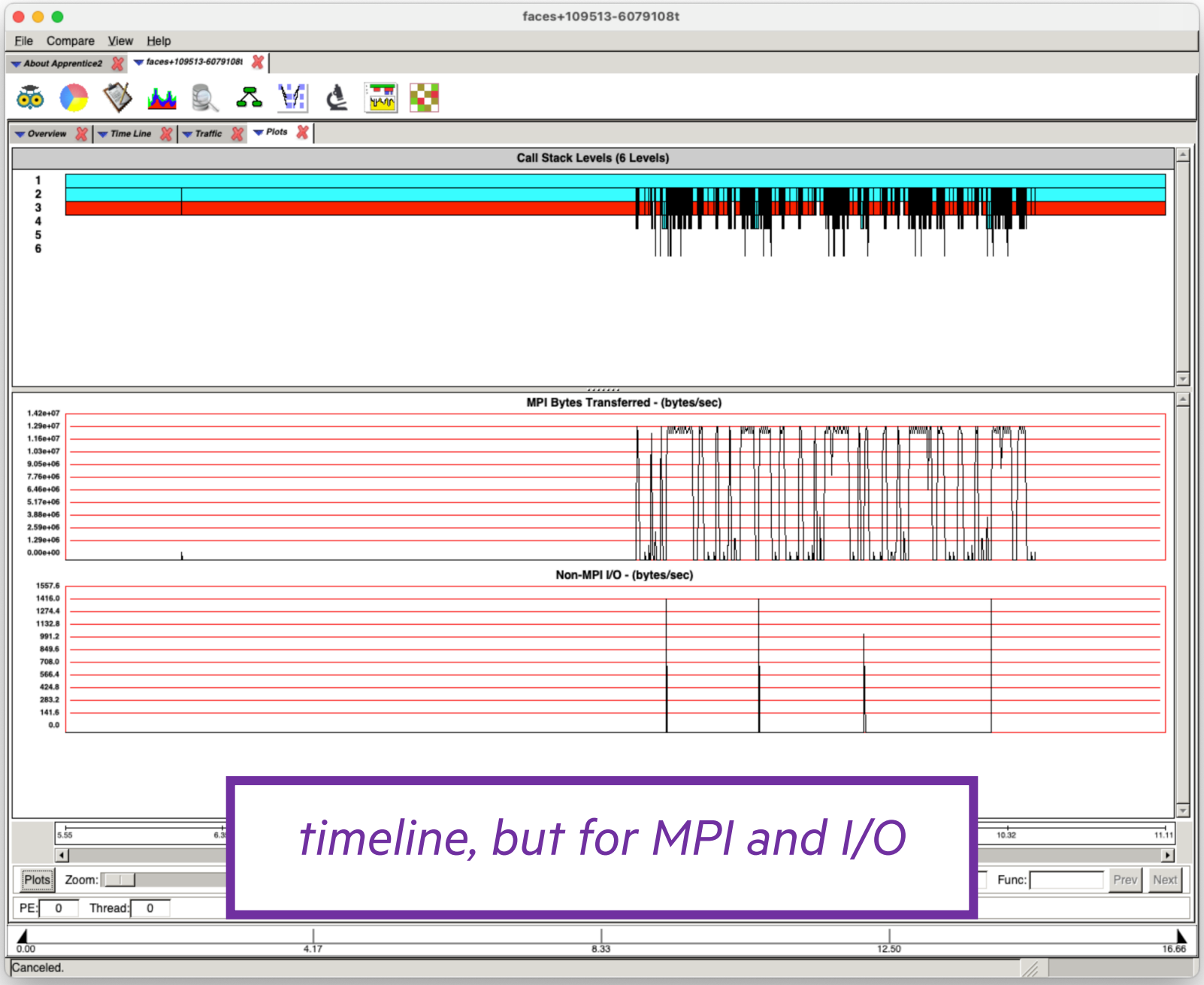
Beware!

Only try "Traffic" with short runs and with few MPI tasks.

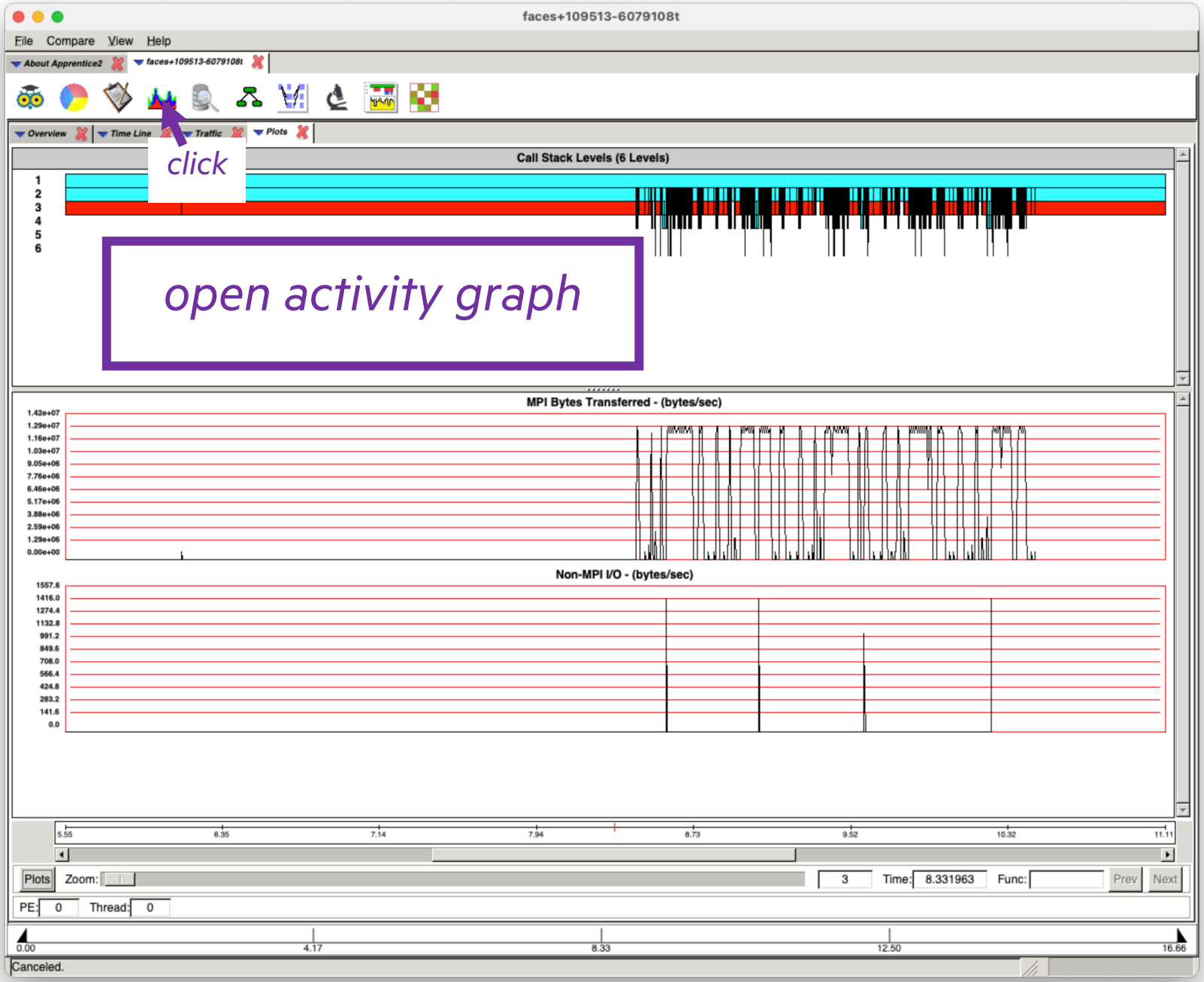
Or expect your computer to be crunching for a while.

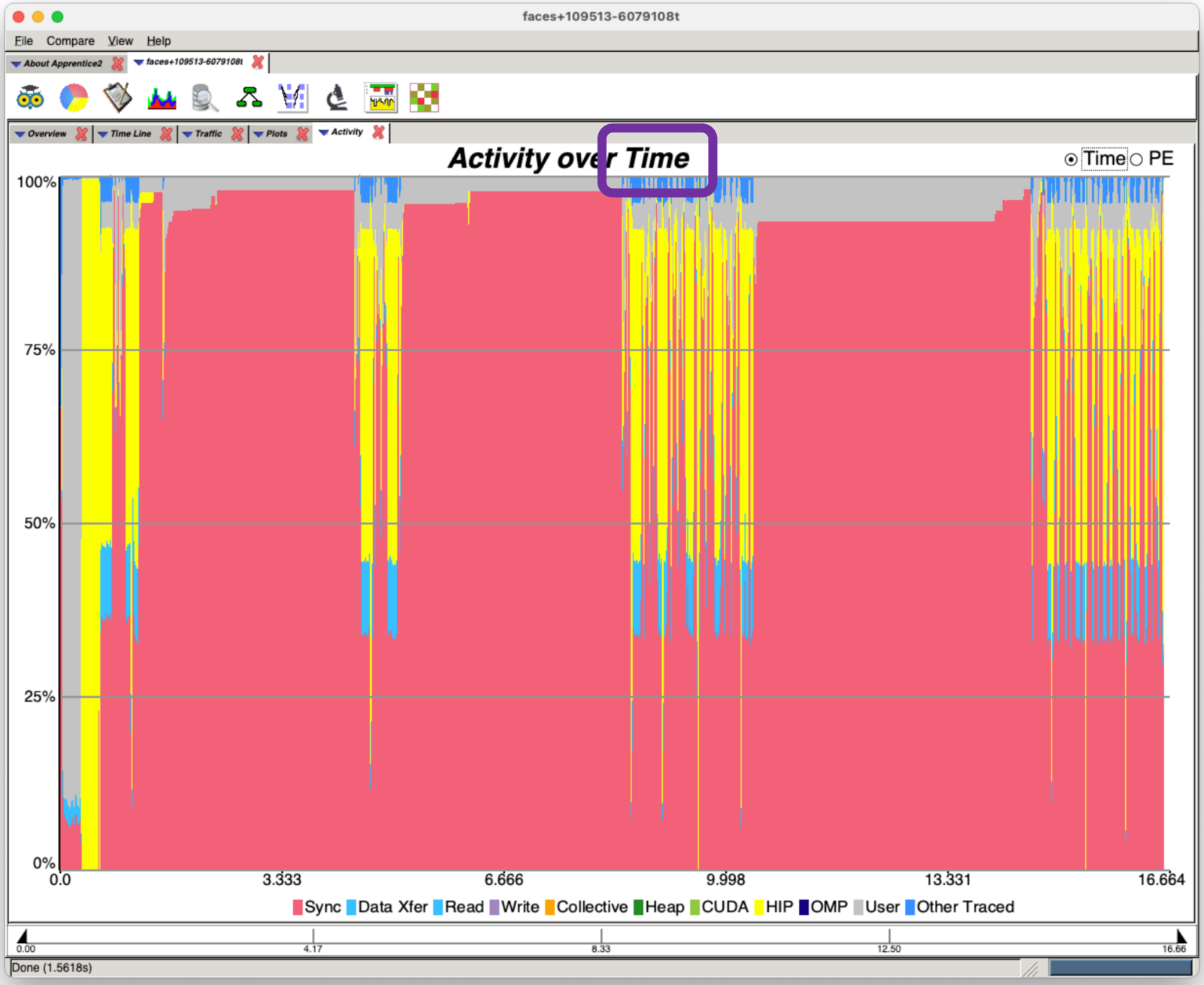

faces+109513-6079108t
File Compare View Help
About Apprentice2
faces+109513-6079108t
Overview
Time Line
Traffic
0.000
2.089
4.177
6.266
8.354
10.443
12.531
14.620
16.708
PE #0
PE #1
PE #2
PE #3
PE #4
PE #5
PE #6
PE #7
PE #8
PE #9
PE #10
PE #11
PE #12
PE #13
PE #14
PE #15
PE #16
PE #17
PE #18
PE #19
PE #20
PE #21
PE #22
PE #23
PE #24
PE #25
PE #26
PE #27
PE #28
PE #29
PE #30
PE #31
PE #32
PE #33
PE #34
PE #35
PE #36
PE #37
PE #38
PE #39
PE #40
PE #41
PE #42
PE #43
PE #44
PE #45
Matching
Send (16,821,612)
Recv (16,838,341)
Sync (1,302,266)
AllToAll (246)
Bcast (662)
Housekeeping (12,928)
User/Other (22,605,254)
Legend:
MPI
scale = 100%
Zoom:
0.00
4.17
8.33
12.50
16.66


faces+109513-6079108t
File Compare View Help
About Apprentice2
faces+109513-6079108t
Overview
Time Line
Traffic
0.000
2.089
4.177
6.266
8.354
10.443
12.531
14.620
16.708
click
open "Plots"
PE #0
PE #1
PE #2
PE #3
PE #4
PE #5
PE #6
PE #7
PE #8
PE #9
PE #10
PE #11
PE #12
PE #13
PE #14
PE #15
PE #16
PE #17
PE #18
PE #19
PE #20
PE #21
PE #22
PE #23
PE #24
PE #25
PE #26
PE #27
PE #28
PE #29
PE #30
PE #31
PE #32
PE #33
PE #34
PE #35
PE #36
PE #37
PE #38
PE #39
PE #40
PE #41
PE #42
PE #43
PE #44
PE #45
Matching
Send (16,821,612)
Recv (16,838,341)
Sync (1,302,266)
AllToAll (246)
Bcast (662)
Housekeeping (12,928)
User/Other (22,605,254)
Legend:
MPI
scale = 100%
Zoom:
0.00
4.17
8.33
12.50
16.66



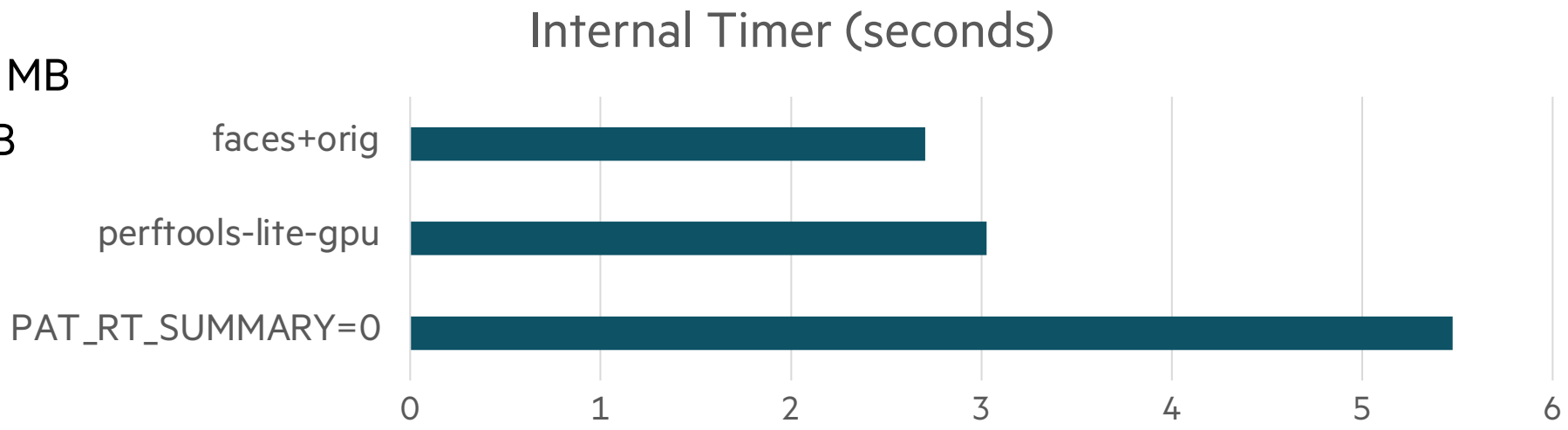
timeline, but for MPI and I/O

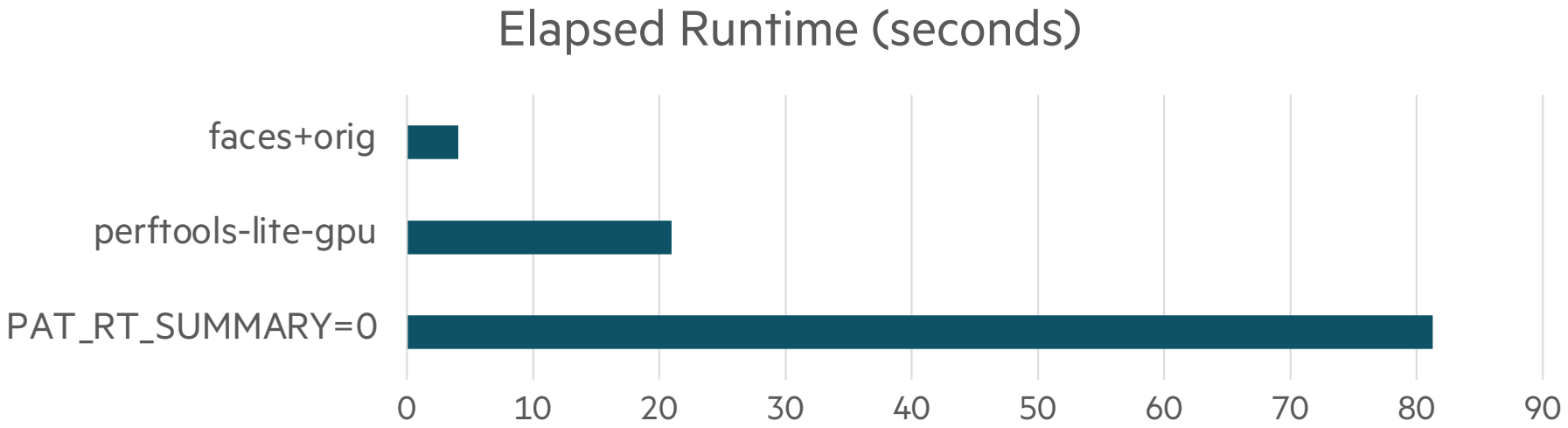

faces+109513-6079108t
File Compare View Help
About Apprentice2
faces+109513-6079108t
Overview
Time Line
Traffic
Plots
click
open activity graph
Call Stack Levels (6 Levels)
1
2
3
4
5
6
MPI Bytes Transferred - (bytes/sec)
1.42e+07
1.29e+07
1.16e+07
1.03e+07
9.05e+06
7.76e+06
6.46e+06
5.17e+06
3.88e+06
2.59e+06
1.29e+06
0.00e+00
Non-MPI I/O - (bytes/sec)
1557.6
1416.0
1274.4
1132.8
991.2
849.6
708.0
566.4
424.8
283.2
141.6
0.0
5.55
6.35
7.14
7.94
8.73
9.52
10.32
11.11
Plots
Zoom:
3
Time:
8.331963
Func:
Prev
Next
PE:
0
Thread:
0
0.00
4.17
8.33
12.50
16.66
Canceled.


faces+109513-6079108t
File Compare View Help
About Apprentice2
faces+109513-6079108t
Overview
Time Line
Traffic
Plots
Activity
Activity over Time
Time
PE
100%
75%
50%
25%
0%
0.0
3.333
6.666
9.998
13.331
16.664
Sync
Data Xfer
Read
Write
Collective
Heap
CUDA
HIP
OMP
User
Other Traced
0.00
4.17
8.33
12.50
16.66
Done (1.5618s)

OVERHEAD OF USING PERFTOOLS-LITE-GPU

- Original exe: 732 kB
- Instrumented exe: 2 MB
- Summary run: 35 MB
- Tracing run: 12 GB

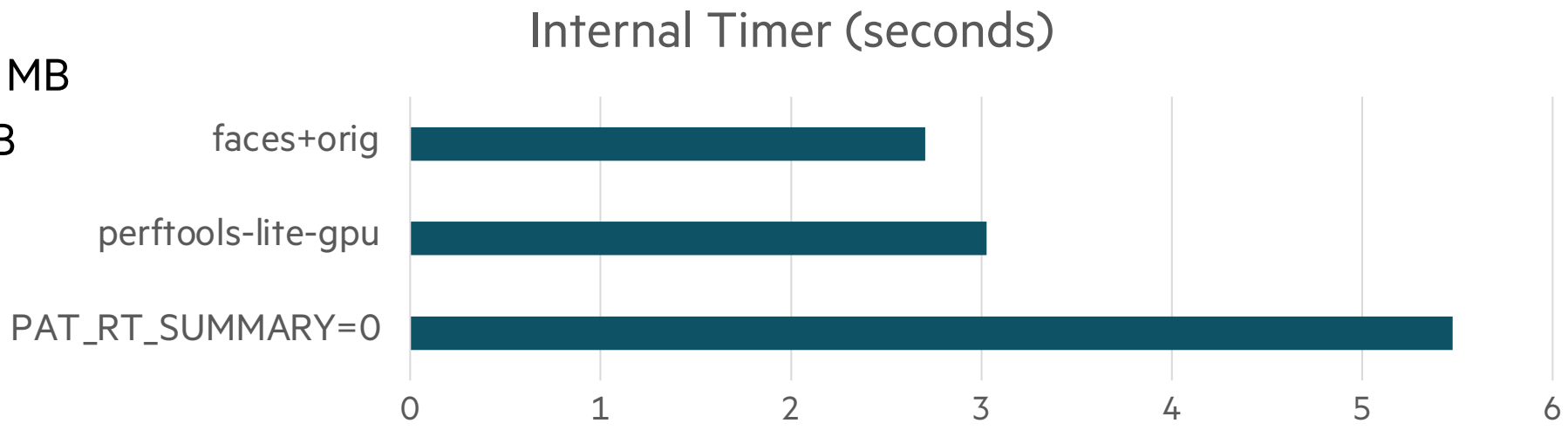


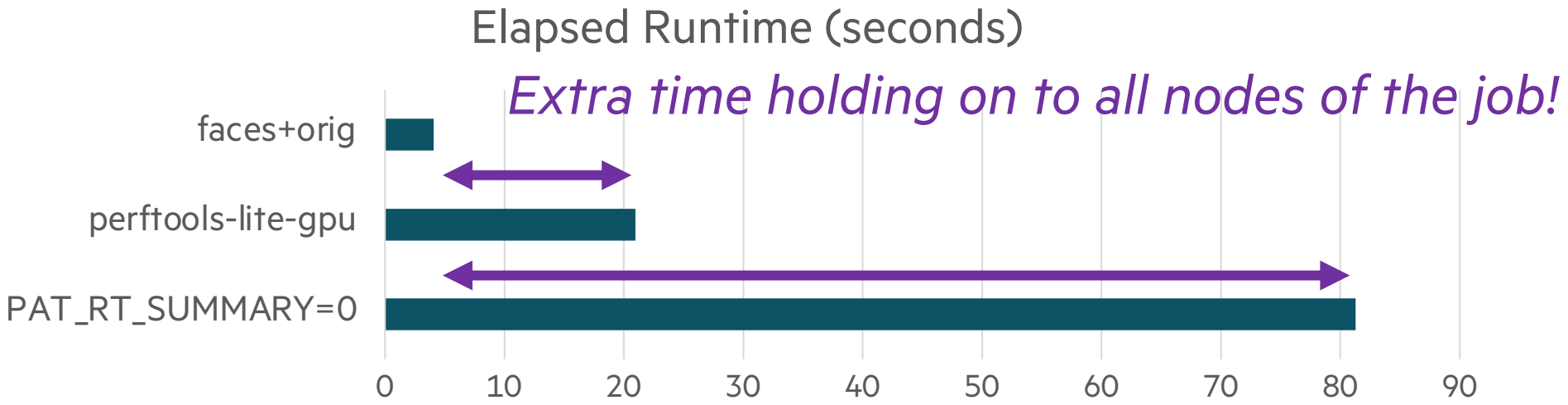


(stored in NFS home directory)

OVERHEAD OF USING PERFTOOLS-LITE-GPU

- Original exe: 732 kB
- Instrumented exe: 2 MB
- Summary run: 35 MB
- Tracing run: 12 GB





PAT_BUILD

or *Time in a Bottle*

PRGENV-AMD BUILD WITH PAT_BUILD

```
module load PrgEnv-amd
module load craype-accel-amd-gfx90a
module load rocm
module load perftools
export PATH="${PATH}:${ROCM_PATH}/llvm/bin"
export CXX='CC -x hip'
export CXXFLAGS='-ggdb -O3 -std=c++17 -Wall'
export LD='CC'
export LDFLAGS="${CXXFLAGS} -L${ROCM_PATH}/lib"
export LIBS='-lamdhip64'
make clean
make
pat_build -g hip,io,mpi -w -f faces
```

use explicit instrumentation instead of "lite-gpu"

PRGENV-AMD BUILD WITH PAT_BUILD

```
module load PrgEnv-amd
module load craype-accel-amd-gfx90a
module load rocm
module load perftools
export PATH="${PATH}:${ROCM_PATH}/llvm/bin"
export CXX='CC -x hip'
export CXXFLAGS='-ggdb -O3 -std=c++17 -Wall'
export LD='CC'
export LDFLAGS="${CXXFLAGS} -L${ROCM_PATH}/lib"
export LIBS='-lamdhip64'
make clean
make
pat_build -g hip,io,mpi -w -f faces
```

trace Hip, I/O, MPI, and all user functions

PRGENV-AMD BUILD WITH PAT_BUILD

```
module load PrgEnv-amd
module load craype-accel-amd-gfx90a
module load rocm
module load perftools
export PATH="${PATH}:${ROCM_PATH}/llvm/bin"
export CXX='CC -x hip'
export CXXFLAGS='-ggdb -O3 -std=c++17 -Wall'
export LD='CC'
export LDFLAGS="${CXXFLAGS} -L${ROCM_PATH}/lib"
export LIBS='-lamdhip64'
make clean
make
pat_build -g hip,io,mpi -w -f faces
```

overwrite <exe>+pat, *if it exists*

PRGENV-AMD RUN WITH PAT_BUILD

```
module load PrgEnv-amd
module load craype-accel-amd-gfx90a
module load rocm
module load perftools
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces+pat
```

PRGENV-AMD RUN WITH PAT_BUILD

```
module load PrgEnv-amd
module load craype-accel-amd-gfx90a
module load rocm
module load perftools
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces+pat
```

```
0: Experiment data directory written:
0: .../faces+pat+130209-6077886t
```

minimal extra output

HIPCC BUILD WITH PAT_BUILD

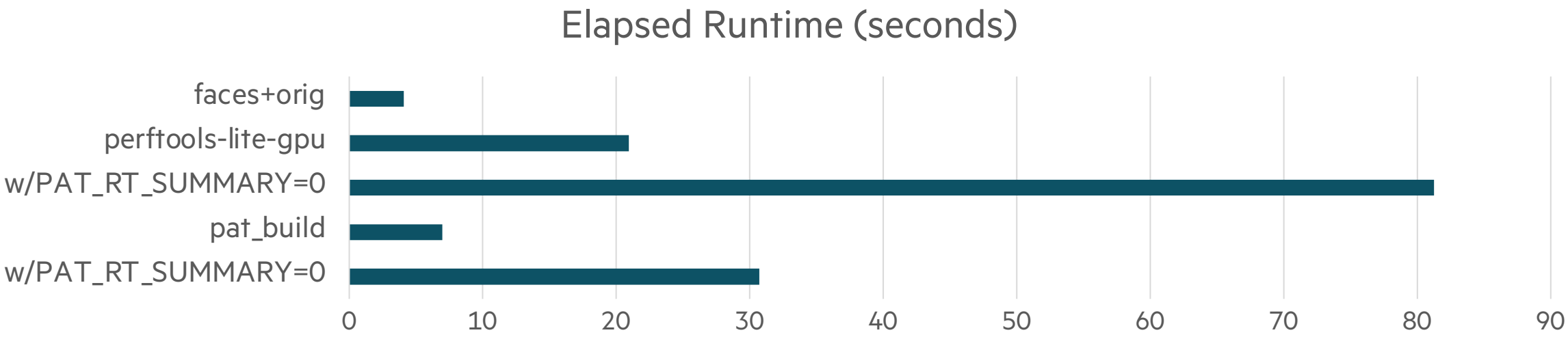
```
module load perftools
module load craype-accel-amd-gfx90a
module load rocm
export CXX='hipcc'
export CXXFLAGS="$(pat_opts include hipcc) \
  $(pat_opts pre_compile hipcc) -g -O3 -std=c++17 -Wall  \
  --offload-arch=gfx90a -I${CRAY_MPICH_DIR}/include \
  $(pat_opts post_compile hipcc)"
export LD='hipcc'
export LDFLAGS="$(pat_opts pre_link hipcc) ${CXXFLAGS} \
  -L${CRAY_MPICH_DIR}/lib ${PE_MPICH_GTL_DIR_amd_gfx908}"
export LIBS="-lmpi ${PE_MPICH_GTL_LIBS_amd_gfx908} \
  $(pat_opts post_link hipcc)"
make clean
make
pat_build -g hip,io,mpi -w -f faces
```

HIPCC RUN WITH PAT_BUILD

```
module load perftools
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest ./faces+pat
```

```
0: Experiment data directory written:
0: .../faces+pat+17135-6077957t
```

OVERHEAD OF PERFTOOLS-LITE-GPU VS. PAT_BUILD



much lower overhead by deferring report generation

PAT_BUILD AND APPRENTICE2

- *Apprentice2*: "Directory … contains no .ap2 files"
- Need to run any *pat_report* first, will generate *.ap2* files
  - Implicit report from *perftools-lite-gpu* does this automatically
- If a *pat_report* is running long, *control-C* could cause incomplete *.ap2* files
  - Delete incomplete files and run *pat_report* again to completion

```
rm -rf faces+pat+5133-6077958t/*ap2*
pat_report faces+pat+5133-6077958t
```

PERFTOOLS TAKEAWAYS

- Build with *PrgEnv-amd* or *hipcc*
- Instrument with *pat_build -g hip,io,mpi -w -f*
- Run with minimal overhead
- Generate reports
  - For overview, *pat_report*
  - For accelerator kernels, *pat_report -O acc_time -s show_ca=fu,so,li*
- Maybe run with full tracing, *PAT_RT_SUMMARY=0*
- Use *pat_report* to generate *.ap2* files
- Explore load imbalance and MPI/accelerator overlap with *Apprentice2*

ROCPROF

or *One Way or Another*

ROCPROF BUILD?

```
module load craype-accel-amd-gfx90a
module load rocm
export CXX='hipcc'
export CXXFLAGS="-ggdb -O3 -std=c++17 -Wall \
  --offload-arch=gfx90a -I${CRAY_MPICH_DIR}/include"
export LD='hipcc'
export LDFLAGS="${CXXFLAGS} -L${CRAY_MPICH_DIR}/lib \
  ${PE_MPICH_GTL_DIR_amd_gfx90a}"
export LIBS="-lmpi ${PE_MPICH_GTL_LIBS_amd_gfx90a}"
make clean
make
```

no changes, PrgEnv-amd *or* hipcc

ROCPROF RUN? *each MPI task writes its own trace output files*

```
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest bash -c \
  'rocprof -o ${SLURM_JOBID}-${SLURM_PROCID}.csv --hip-trace ./faces'
```

ROCPROF RUN?

```
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest bash -c \
  'rocprof -o ${SLURM_JOBID}-${SLURM_PROCID}.csv --hip-trace ./faces'
```

inline a wrapper script using bash -c

ROCPROF RUN?

```
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest bash -c \
  'rocprof -o ${SLURM_JOBID}-${SLURM_PROCID}.csv --hip-trace ./faces'
```

use Slurm environment variables to name each file differently for each MPI task

ROCPROF RUN?

```
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest bash -c \
 'rocprof -o ${SLURM_JOBID}-${SLURM_PROCID}.csv --hip-trace ./faces'
```

Single quotes!

needed to keep shell from immediately evaluating Slurm environment variables

ROCPROF RUN?

```
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest bash -c \
  'rocprof -o ${SLURM_JOBID}-${SLURM_PROCID}.csv --hip-trace ./faces'
```

trace Hip calls and kernels

ROCPROF: TOO MUCH TIME ON MY HANDS

- 1,385,532 extra lines of output
- 384 output files (6 per MPI task)
- 6.3 GB of output

RUN ROCPROF ON ONE OF THE TASKS

```
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest bash -c \
  'if [ ${SLURM_PROCID} -eq 0 ]; then rocprof --hip-trace ../faces; \
   else ../faces; fi'
```

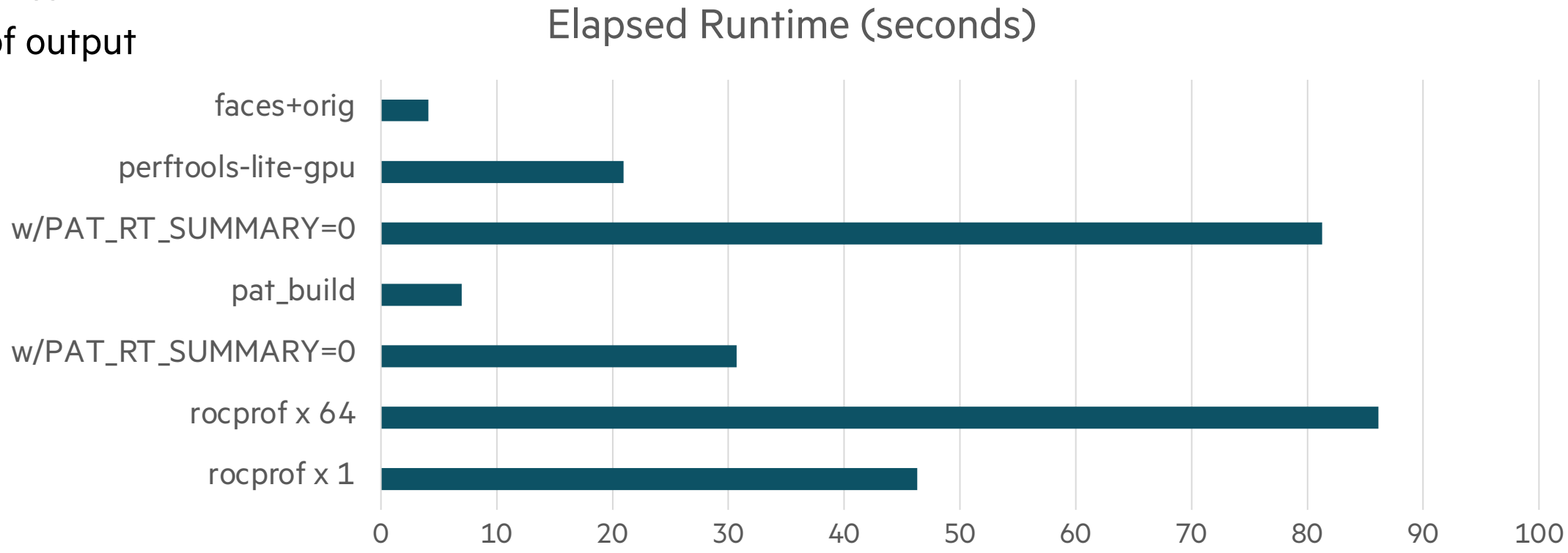
RUN ROCPROF ON ONE OF THE TASKS

```
module load craype-accel-amd-gfx90a
module load rocm
export MPICH_GPU_SUPPORT_ENABLED=1
srun -l -u -t 5:00 -i in.txt -n 64 -N 8 -c 8 \
  --gpus-per-node=8 --gpu-bind=closest bash -c \
  'if [ ${SLURM_PROCID} -eq 0 ]; then rocprof --hip-trace ../faces; \
   else ../faces; fi'
```

pick a task ID, here the root task

ROCPROF: JUST ONE LOOK

- 21,584 extra lines of output
- 6 output files
- 100 MB of output



ROCPROF OUTPUT FILES

```
155 results.copy_stats.csv
28M results.db
854 results.hip_stats.csv
72M results.json
764 results.stats.csv
18K results.sysinfo.txt
```

ROCPROF OUTPUT FILES

```
155 results.copy_stats.csv
28M results.db
854 results.hip_stats.csv
72M results.json
764 results.stats.csv
18K results.sysinfo.txt
```

CSV = comma-separated values

RESULTS.STATS.CSV

```
"Name","Calls","TotalDurationNs","AverageNs","Percentage"
void gpuRun4x1<Faces::share(DArray<double, 6>&)::{lambda(int, int, int,
int, int)#1}>(Faces::share(DArray<double, 6>&)::{lambda(int, int, int,
int, int)#1}, int, int, int, int, int,
int),10000,1804489434,180448,69.1574710360635 4
void gpuRun3x1<Faces::share(DArray<double, 6>&)::{lambda(int, int, int,
int)#2}>(Faces::share(DArray<double, 6>&)::{lambda(int, int, int, int)#2},
int, int, int, int),10000,493243107,49324,18.90365509676378
void gpuRun3x1<Faces::share(DArray<double, 6>&)::{lambda(int, int, int,
int)#1}>(Faces::share(DArray<double, 6>&)::{lambda(int, int, int, int)#1},
int, int, int, int),10000,298869535,29886,11.45424341139383
init(DArray<double, 6>),190,12645207,66553,0.4846304557788438 4
```

RESULTS.STATS.CSV

```
"Name","Calls","TotalDurationNs","AverageNs","Percentage"
void gpuRun4x1<Faces::share(DArray<double, 6>&)::{lambda(int, int, int,
int, int)#1}>(Faces::share(DArray<double, 6>&)::{lambda(int, int, int,
int, int)#1}, int, int, int, int, int,
int),10000,1804489434,180448,69.1574710360635 4
void gpuRun3x1<Faces::share(DArray<double, 6>&)::{lambda(int, int, int,
int)#2}>(Faces::share(DArray<double, 6>&)::{lambda(int, int, int, int)#2},
int, int, int, int),10000,493243107,49324,18.9036550967637 8
void gpuRun3x1<Faces::share(DArray<double, 6>&)::{lambda(int, int, int,
int)#1}>(Faces::share(DArray<double, 6>&)::{lambda(int, int, int, int)#1},
int, int, int, int),10000,298869535,29886,11.4542434113938 3
init(DArray<double, 6>),190,12645207,66553,0.484630455778843 84
```

not handsome, but handy

ROCPROF OUTPUT FILES

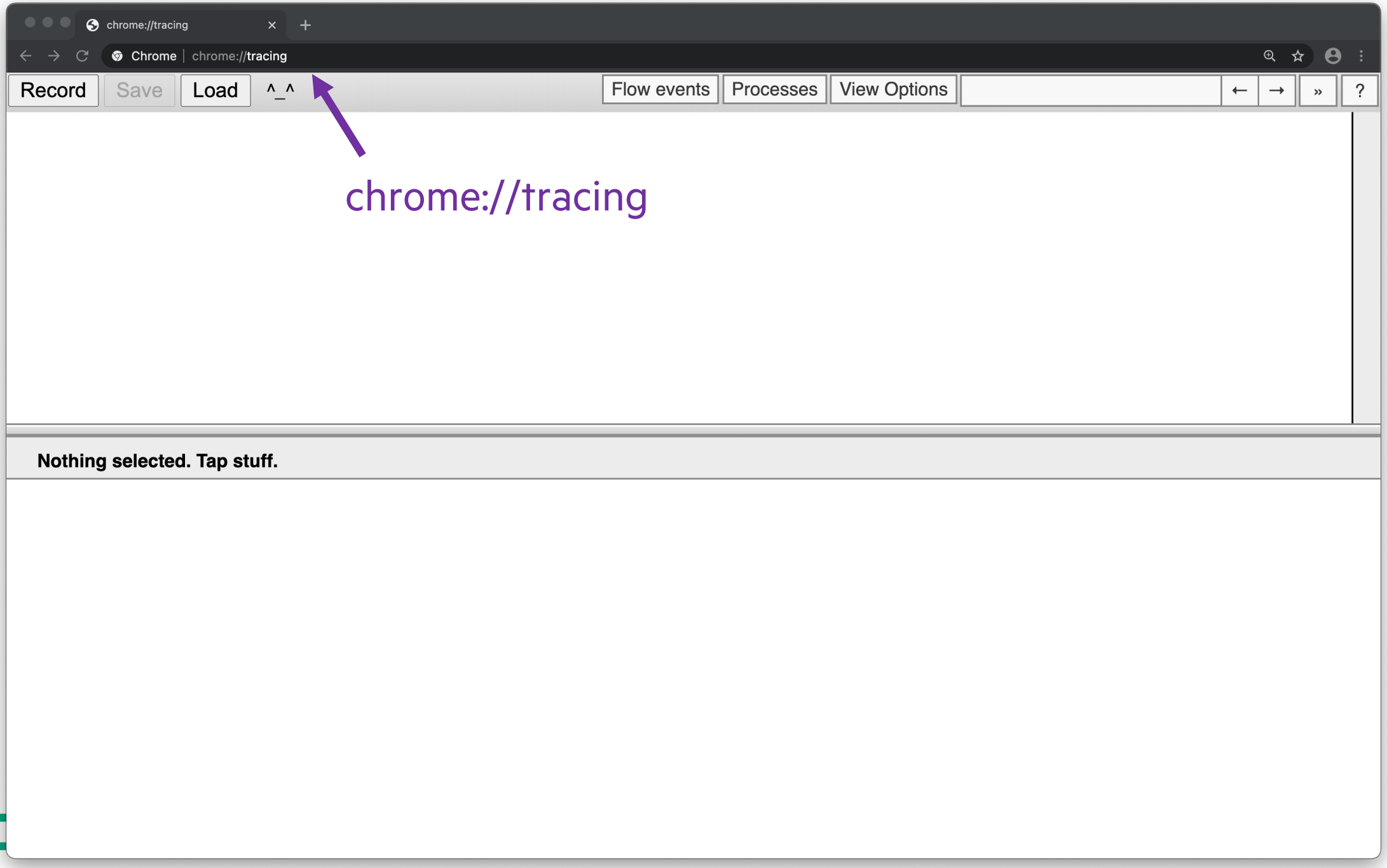
```
155 results.copy_stats.csv
28M results.db
854 results.hip_stats.csv
72M results.json
764 results.stats.csv
18K results.sysinfo.txt
```

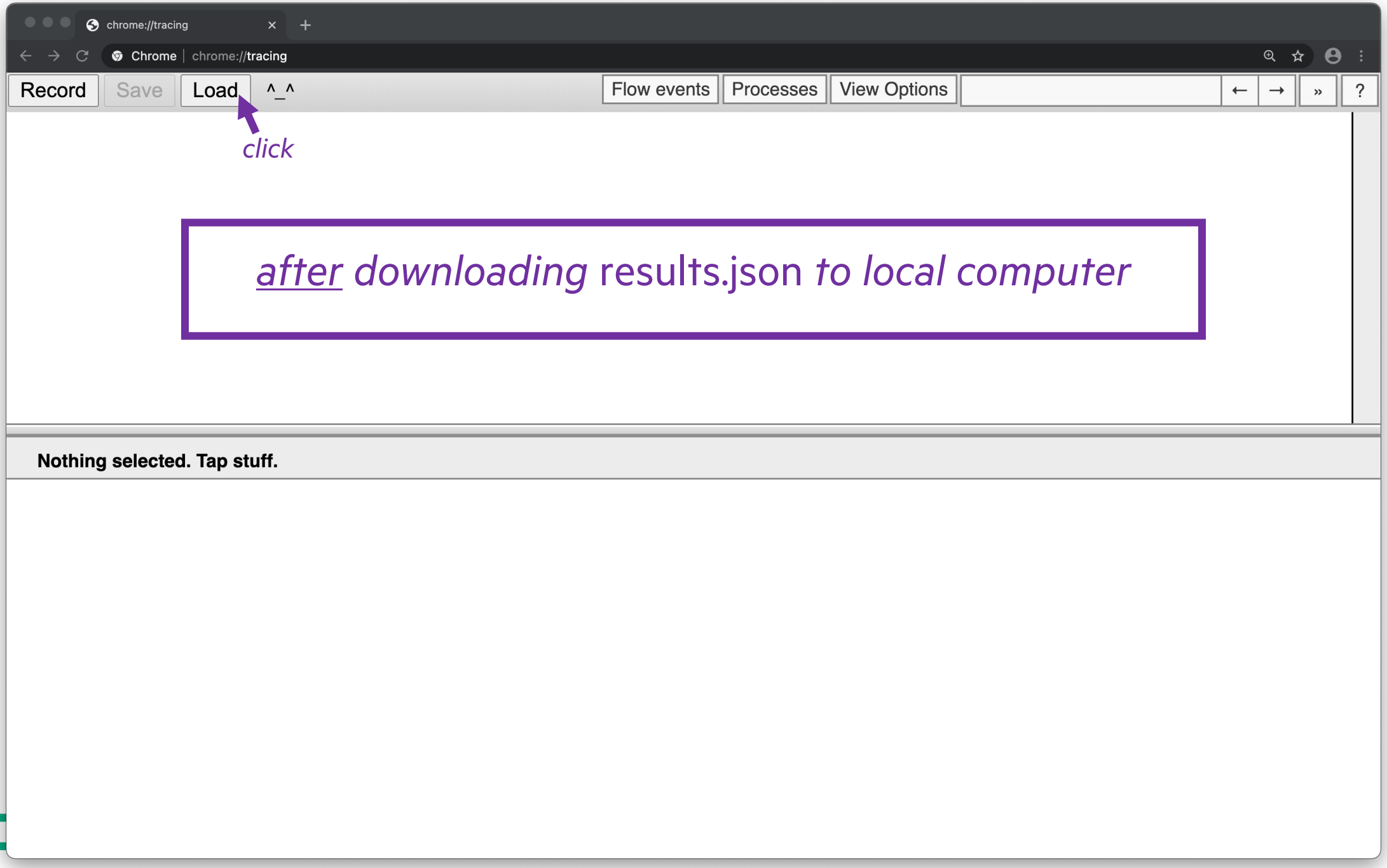
trace file

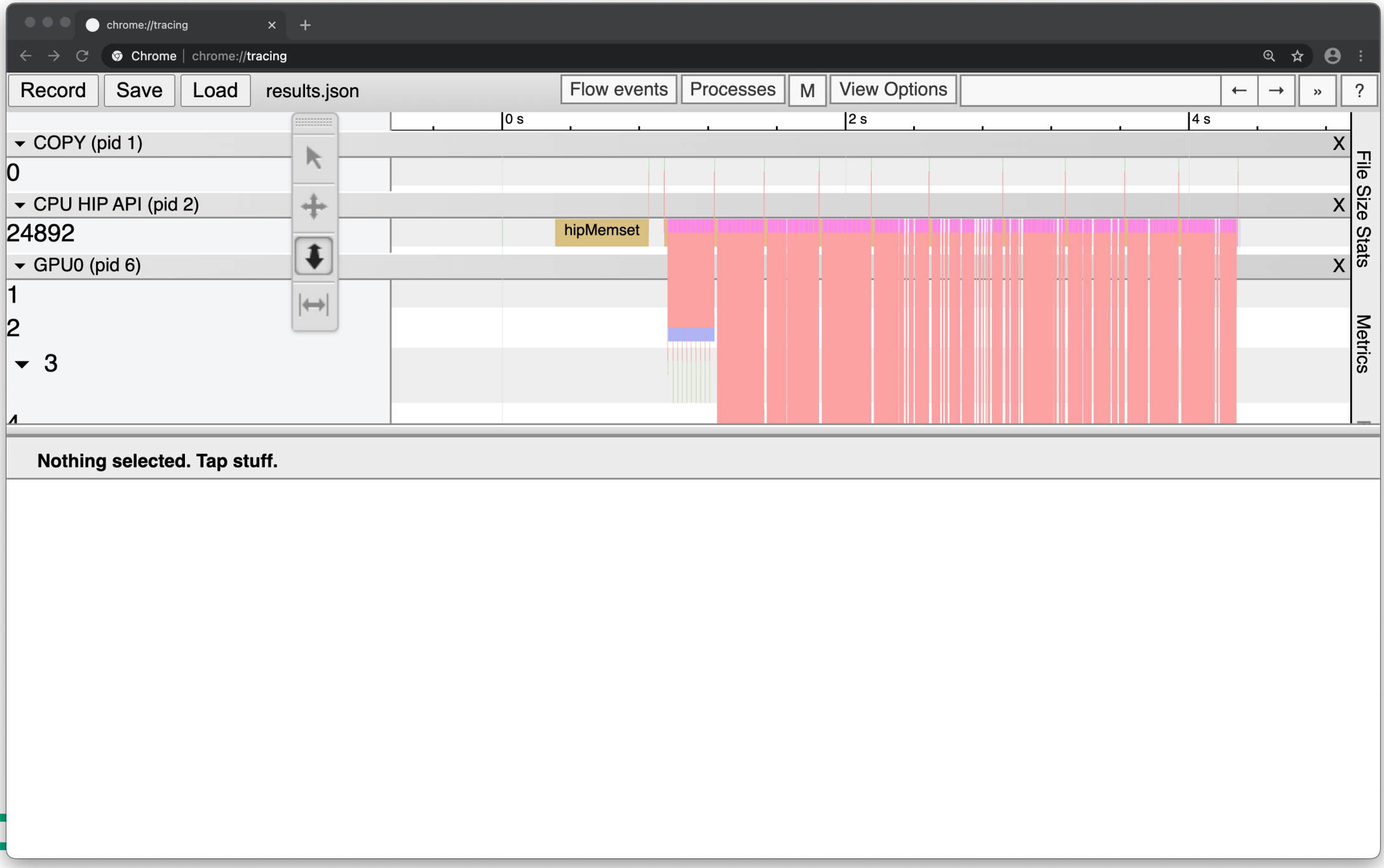
view with Chrome

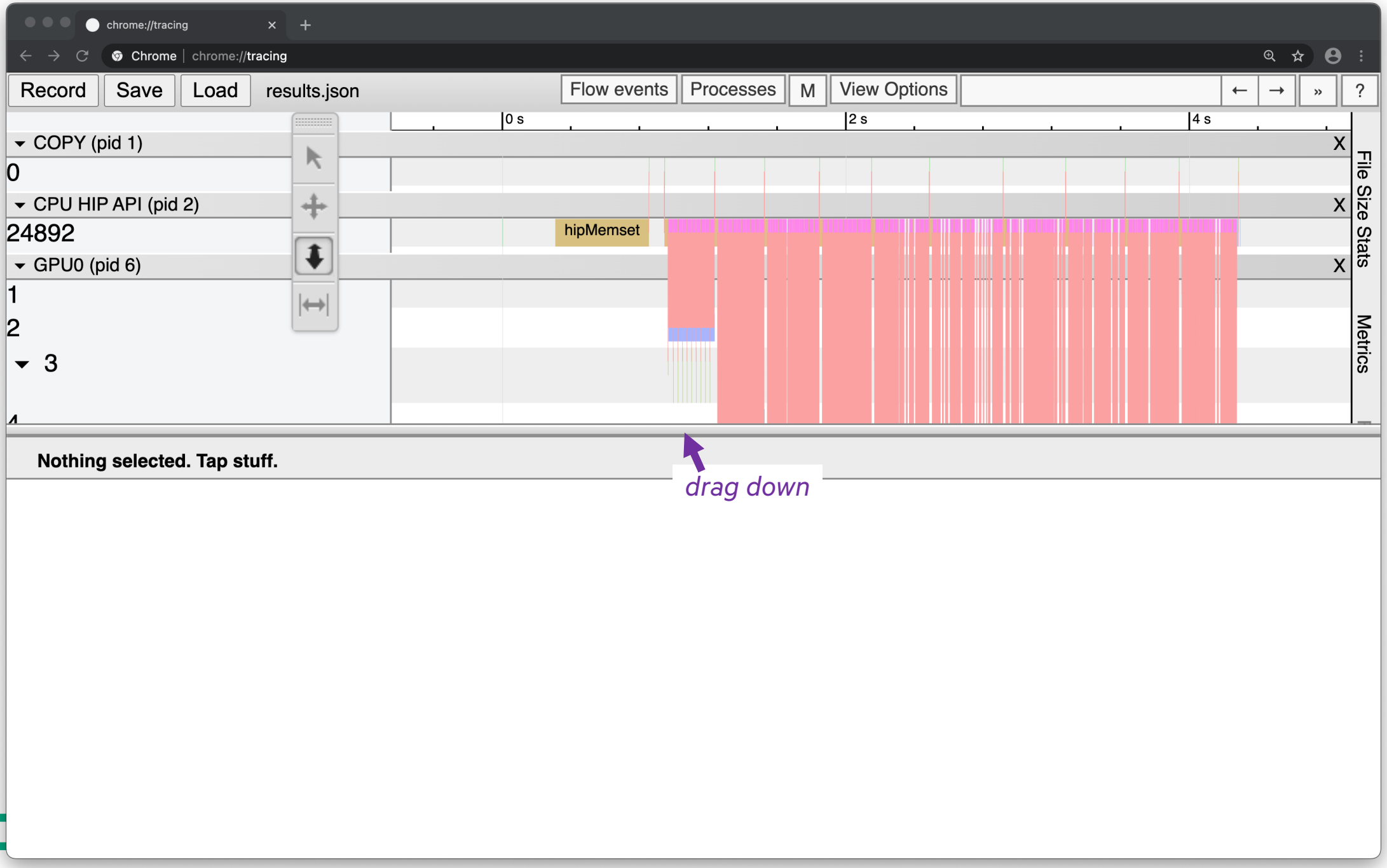
CHROME TRACING

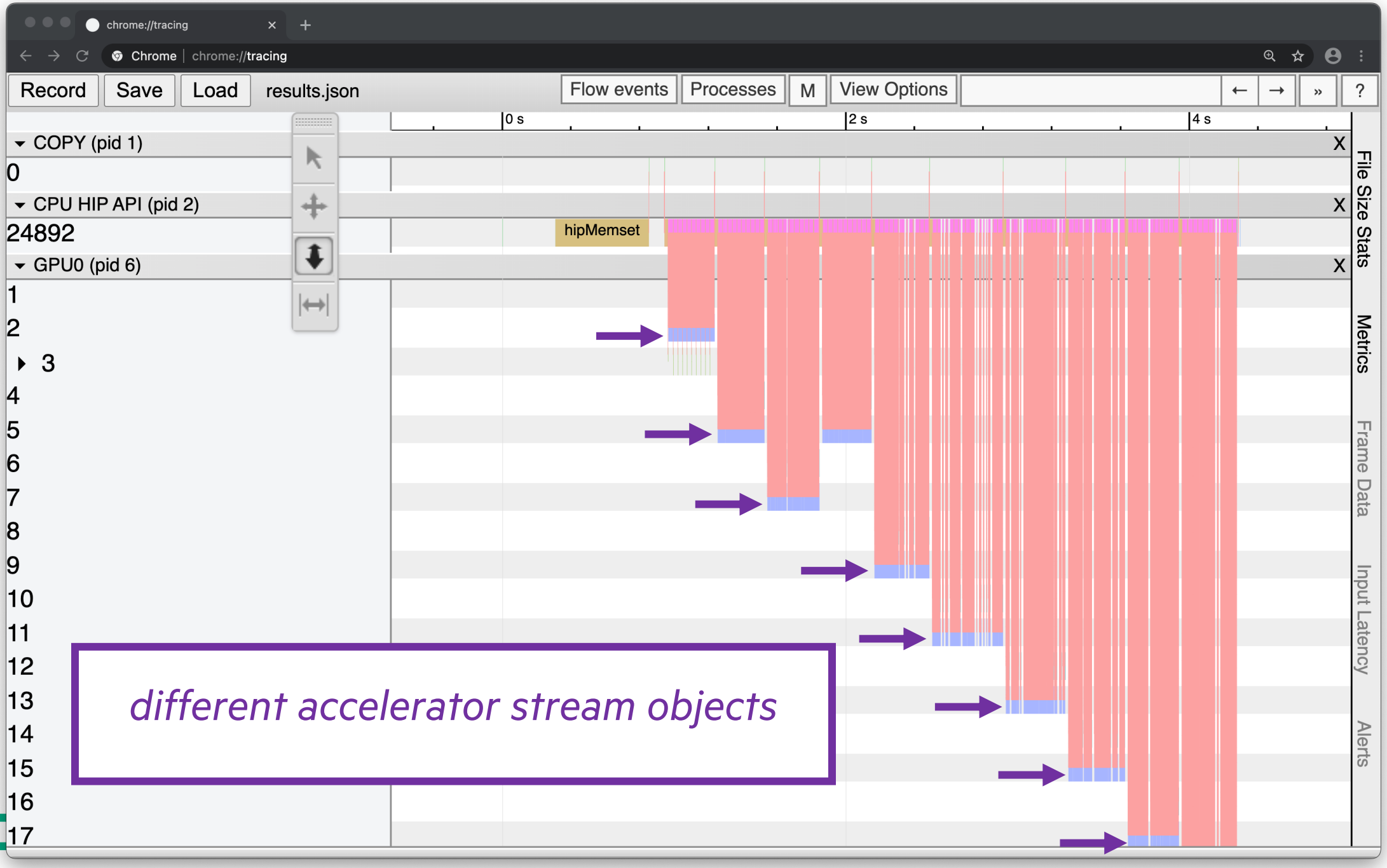
or *The Browser with All the Gifts*

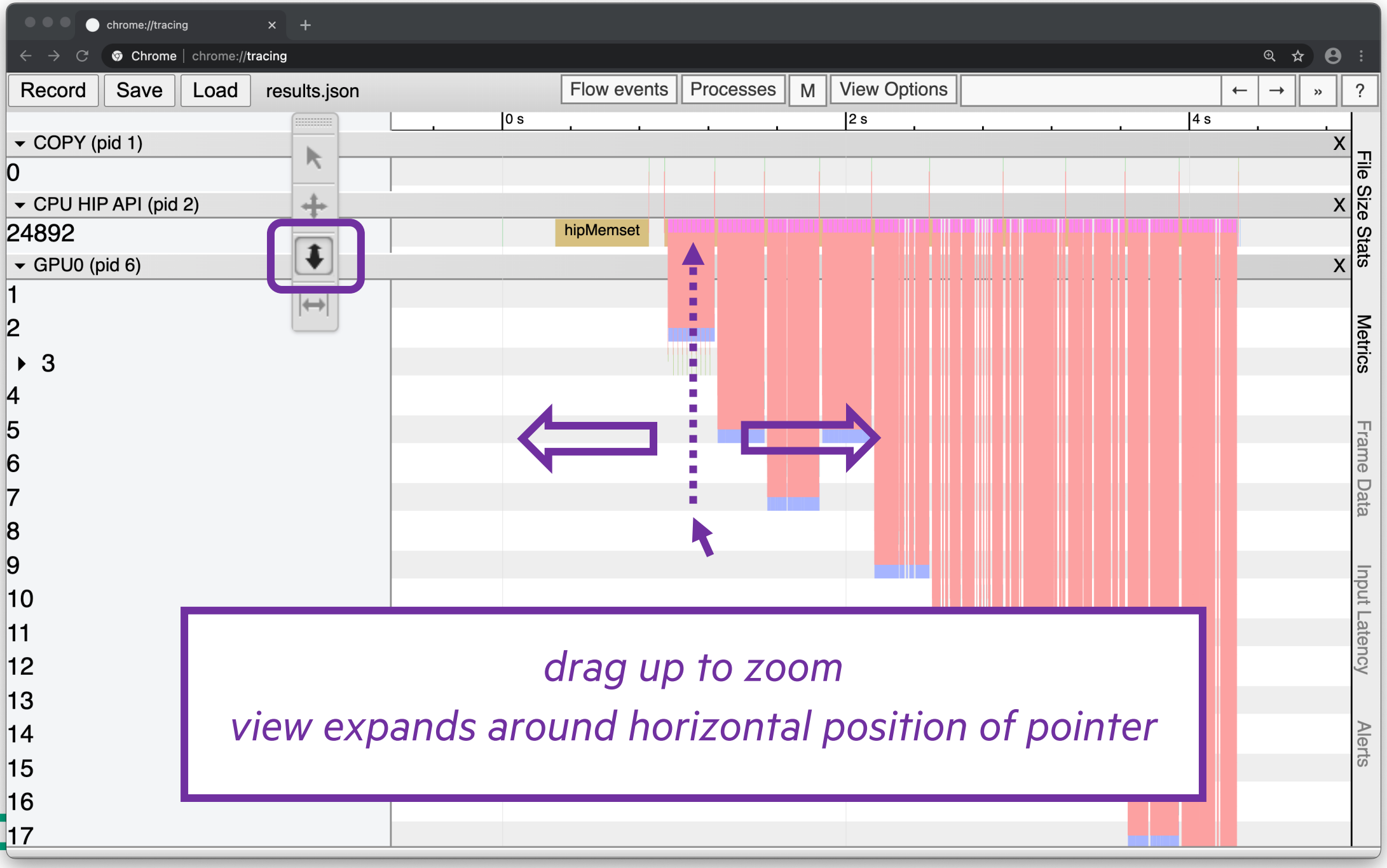

chrome://tracing
Chrome | chrome://tracing
Record
Save
Load
^_^
Flow events
Processes
View Options
chrome://tracing
Nothing selected. Tap stuff.

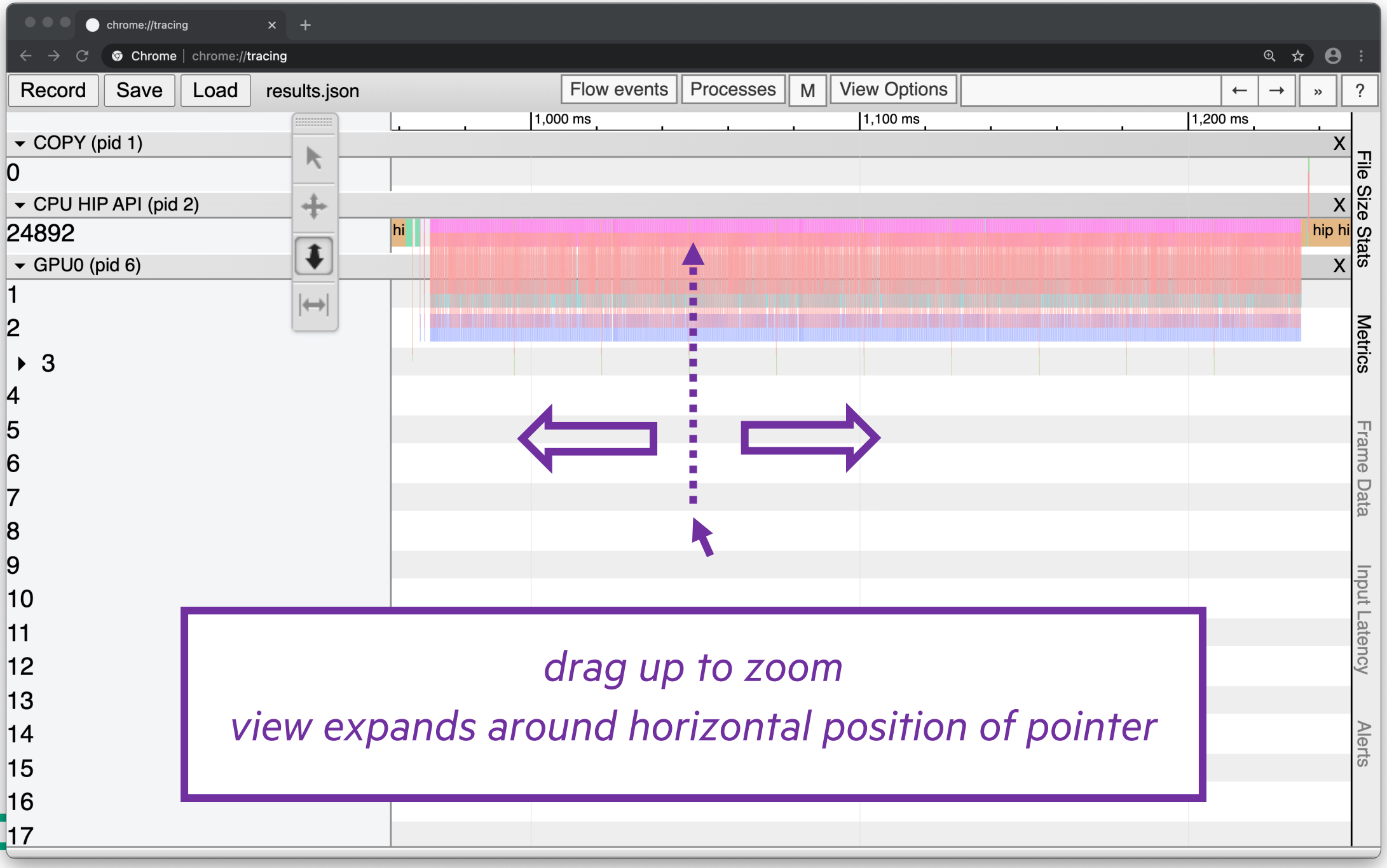

chrome://tracing
Chrome | chrome://tracing
Record
Save
Load
^_^
Flow events
Processes
View Options
click
after downloading results.json to local computer
Nothing selected. Tap stuff.

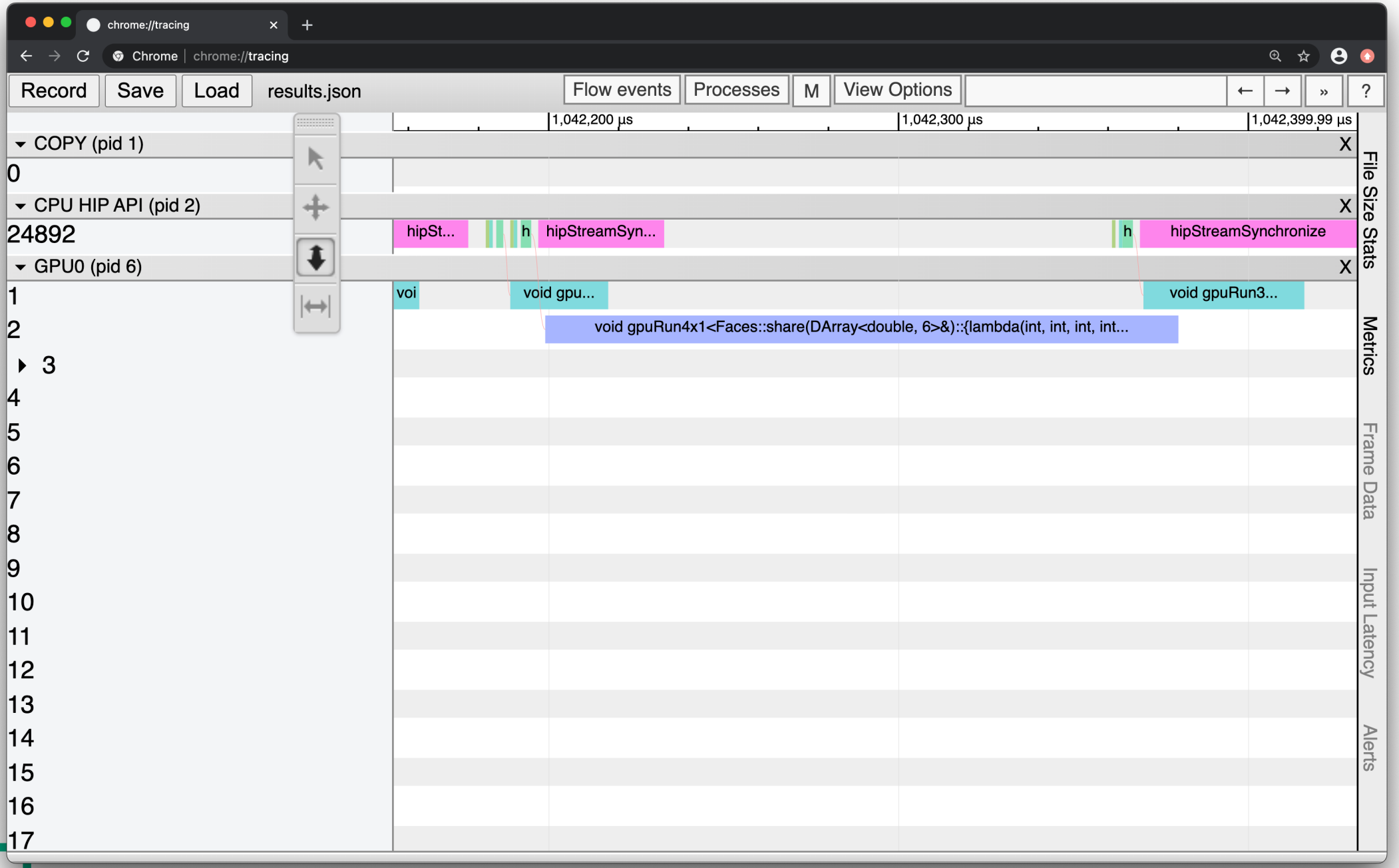

chrome://tracing
Chrome | chrome://tracing
Record
Save
Load
results.json
Flow events
Processes
M
View Options
0 s
2 s
4 s
COPY (pid 1)
0
CPU HIP API (pid 2)
24892
hipMemset
GPU0 (pid 6)
1
2
3
File Size Stats
Metrics
Nothing selected. Tap stuff.

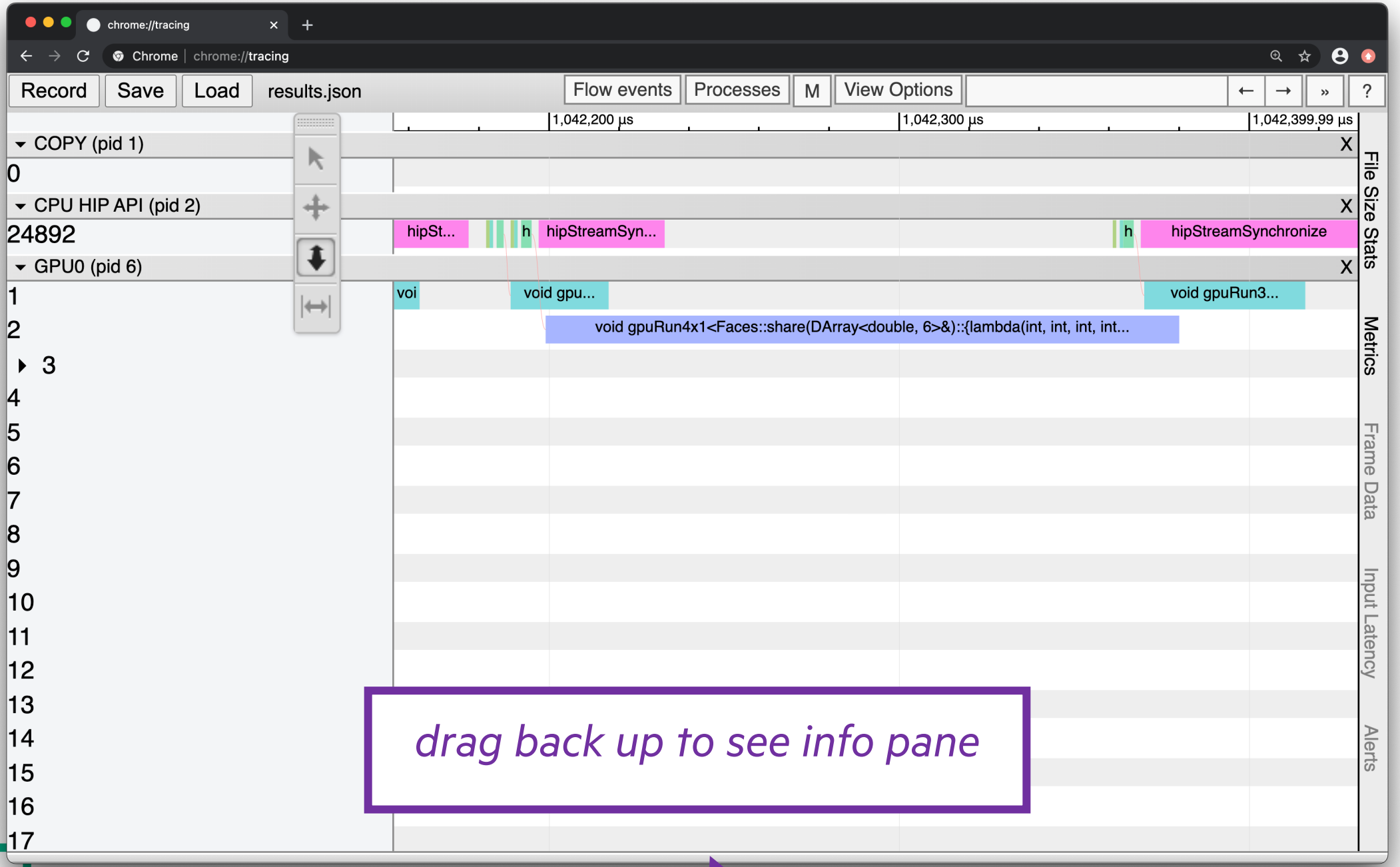

chrome://tracing
Chrome | chrome://tracing
Record
Save
Load
results.json
Flow events
Processes
M
View Options
0 s
2 s
4 s
COPY (pid 1)
0
CPU HIP API (pid 2)
24892
hipMemset
GPU0 (pid 6)
1
2
3
File Size Stats
Metrics
Nothing selected. Tap stuff.
drag down

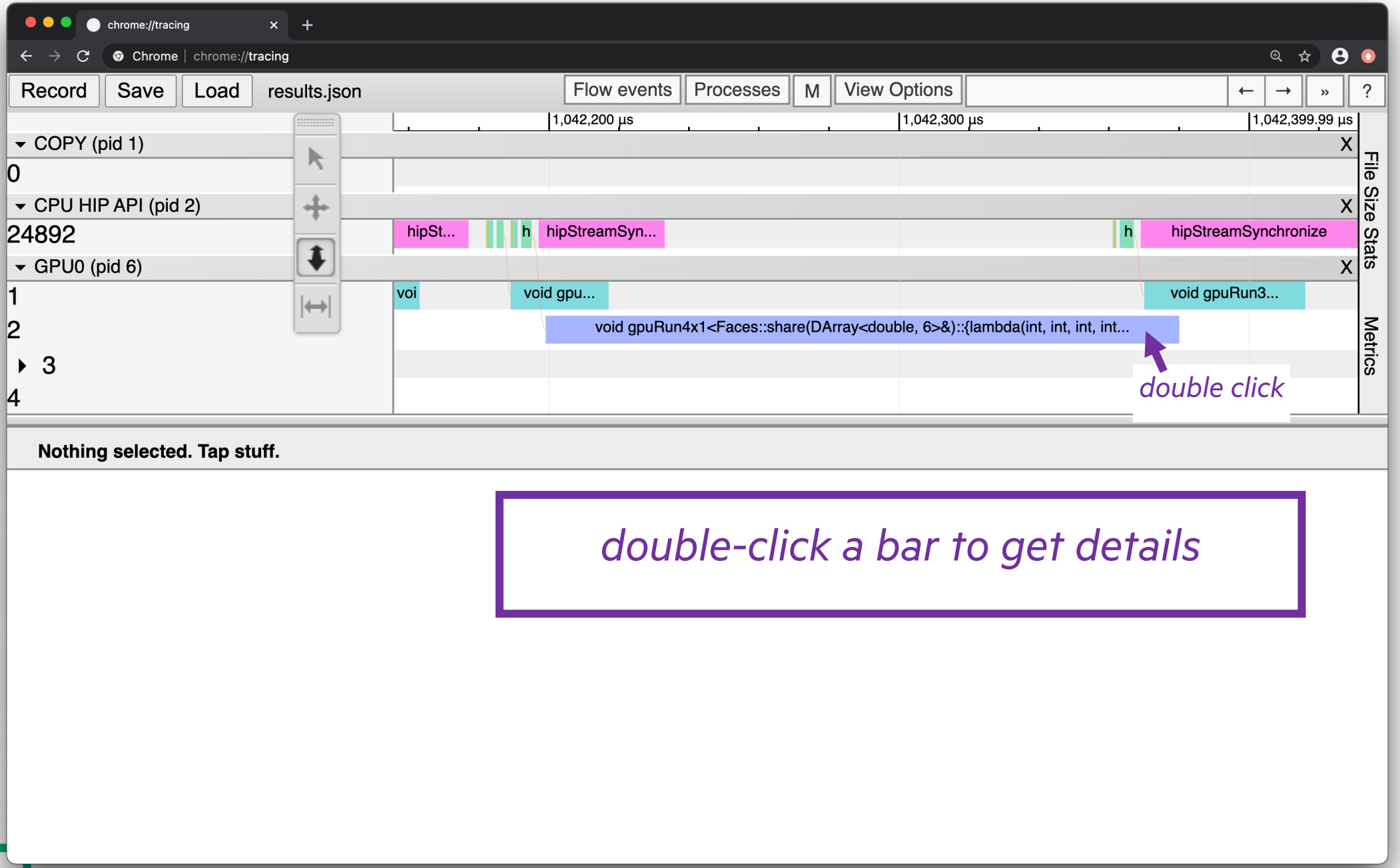

chrome://tracing
Chrome | chrome://tracing
Record
Save
Load
results.json
Flow events
Processes
M
View Options
0 s
2 s
4 s
COPY (pid 1)
0
CPU HIP API (pid 2)
24892
hipMemset
GPU0 (pid 6)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
File Size Stats
Metrics
Frame Data
Input Latency
Alerts
different accelerator stream objects

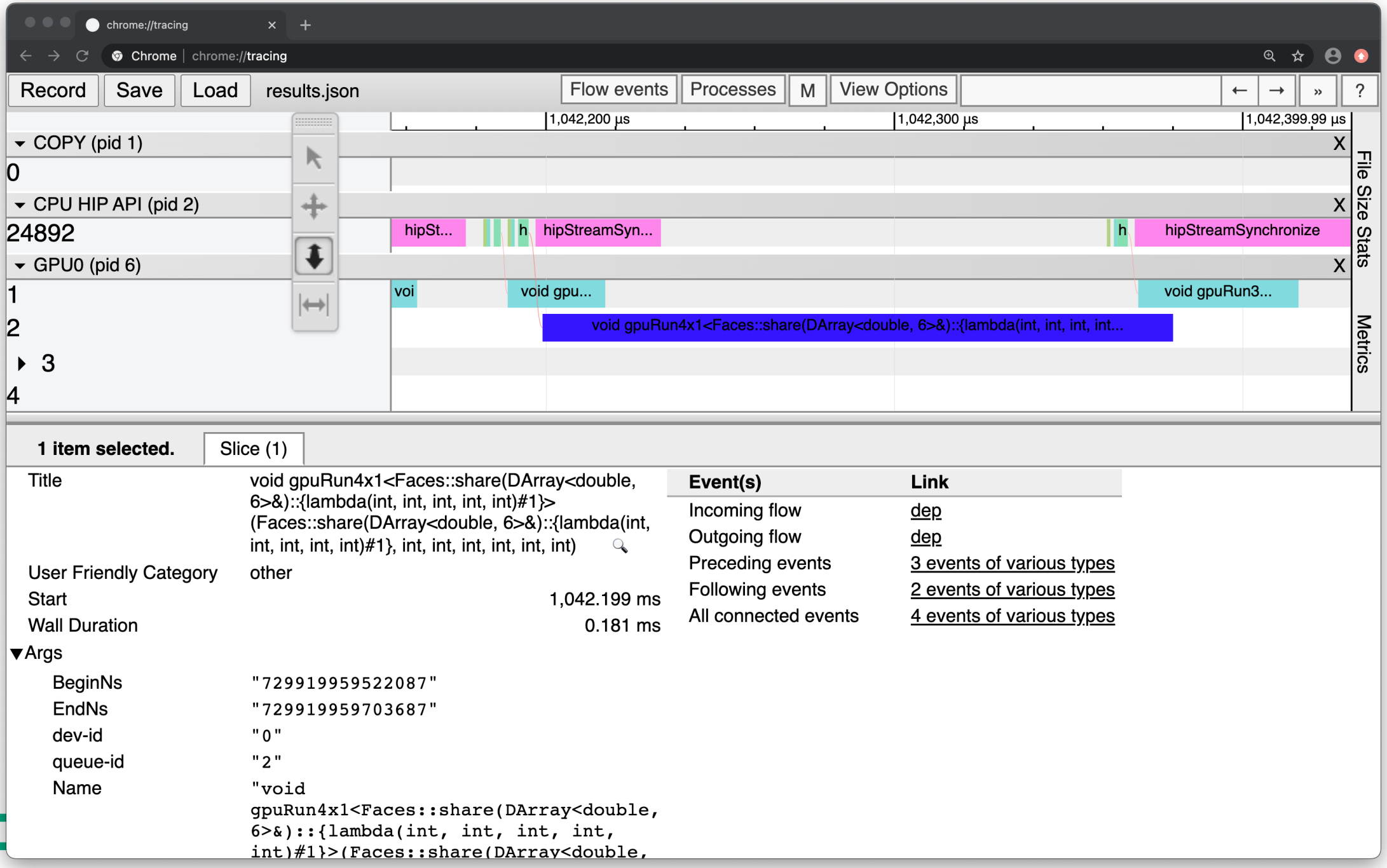

chrome://tracing
Chrome | chrome://tracing
Record
Save
Load
results.json
Flow events
Processes
M
View Options
0 s
2 s
4 s
COPY (pid 1)
0
CPU HIP API (pid 2)
24892
hipMemset
GPU0 (pid 6)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
File Size Stats
Metrics
Frame Data
Input Latency
Alerts
drag up to zoom
view expands around horizontal position of pointer


chrome://tracing
Chrome | chrome://tracing
Record
Save
Load
results.json
Flow events
Processes
M
View Options
1,000 ms
1,100 ms
1,200 ms
COPY (pid 1)
0
CPU HIP API (pid 2)
24892
hi
hip hi
GPU0 (pid 6)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
File Size Stats
Metrics
Frame Data
Input Latency
Alerts
drag up to zoom
view expands around horizontal position of pointer


chrome://tracing
Chrome | chrome://tracing
Record
Save
Load
results.json
Flow events
Processes
M
View Options
1,042,200 µs
1,042,300 µs
1,042,399.99 µs
COPY (pid 1)
0
CPU HIP API (pid 2)
24892
hipSt...
h
hipStreamSyn...
h
hipStreamSynchronize
GPU0 (pid 6)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
voi
void gpu...
void gpuRun3...
void gpuRun4x1<Faces::share(DArray<double, 6>&)::{lambda(int, int, int, int...
File Size Stats
Metrics
Frame Data
Input Latency
Alerts


chrome://tracing
Chrome | chrome://tracing
Record
Save
Load
results.json
Flow events
Processes
M
View Options
1,042,200 µs
1,042,300 µs
1,042,399.99 µs
COPY (pid 1)
0
CPU HIP API (pid 2)
24892
hipSt...
h
hipStreamSyn...
h
hipStreamSynchronize
GPU0 (pid 6)
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
voi
void gpu...
void gpuRun3...
void gpuRun4x1<Faces::share(DArray<double, 6>&)::{lambda(int, int, int, int...
File Size Stats
Metrics
Frame Data
Input Latency
Alerts
drag back up to see info pane
drag up


chrome://tracing
Chrome | chrome://tracing
Record
Save
Load
results.json
Flow events
Processes
M
View Options
1,042,200 µs
1,042,300 µs
1,042,399.99 µs
COPY (pid 1)
0
CPU HIP API (pid 2)
24892
hipSt...
h
hipStreamSyn...
h
hipStreamSynchronize
GPU0 (pid 6)
1
2
3
4
voi
void gpu...
void gpuRun3...
void gpuRun4x1<Faces::share(DArray<double, 6>&)::{lambda(int, int, int, int...
File Size Stats
Metrics
double click
Nothing selected. Tap stuff.
double-click a bar to get details


chrome://tracing
Chrome | chrome://tracing
Record
Save
Load
results.json
Flow events
Processes
M
View Options
1,042,200 µs
1,042,300 µs
1,042,399.99 µs
COPY (pid 1)
0
CPU HIP API (pid 2)
24892
hipSt...
h
hipStreamSyn...
h
hipStreamSynchronize
GPU0 (pid 6)
1
2
3
4
voi
void gpu...
void gpuRun3...
void gpuRun4x1<Faces::share(DArray<double, 6>&)::{lambda(int, int, int, int...
File Size Stats
Metrics
1 item selected.
Slice (1)
Title
void gpuRun4x1<Faces::share(DArray<double, 6>&)::{lambda(int, int, int, int, int)#1}> (Faces::share(DArray<double, 6>&)::{lambda(int, int, int, int, int)#1}, int, int, int, int, int, int)
User Friendly Category
other
Start
1,042.199 ms
Wall Duration
0.181 ms
Args
BeginNs
"729919959522087"
EndNs
"729919959703687"
dev-id
"0"
queue-id
"2"
Name
"void gpuRun4x1<Faces::share(DArray<double, 6>&)::{lambda(int, int, int, int, int)#1}>(Faces::share(DArray<double,
Event(s)
Link
Incoming flow
dep
Outgoing flow
dep
Preceding events
3 events of various types
Following events
2 events of various types
All connected events
4 events of various types

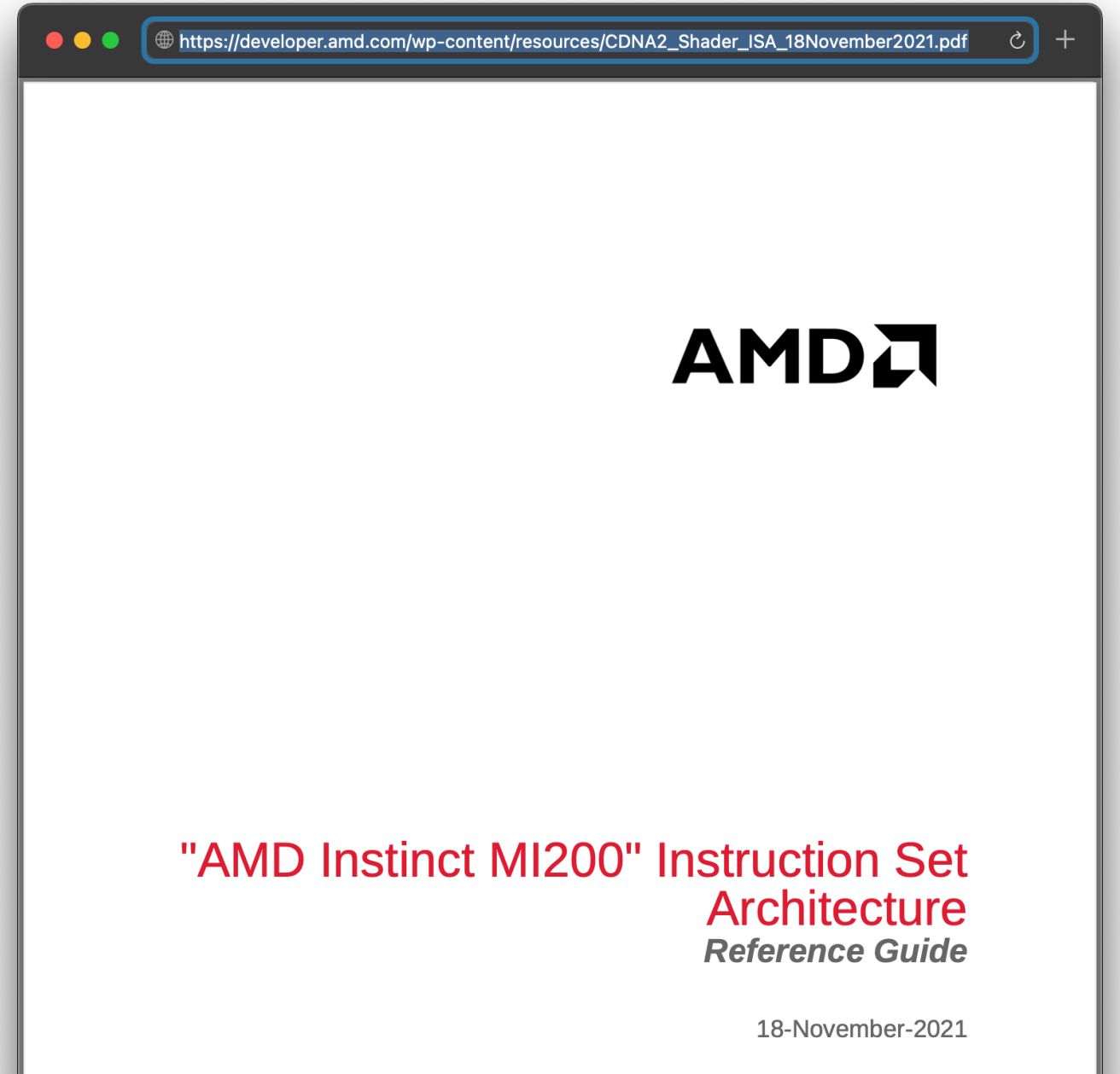
ROCPROF TAKEAWAYS

- Run *rocprof* on a small number of your MPI tasks
- No need to recompile
- Get a quick profile of kernels
- View traces with *Chrome*
  - See where host and accelerator are waiting for each other
  - See overlap among kernels
  - But no MPI events

COMPILE-TIME PERFORMANCE MEASUREMENT

or *How I Learned to Stop Profiling and Love Assembly*

REFERENCE FOR THE NEXT SLIDE



A SIMPLE MI250X WORD PROBLEM

- Each wavefront has 64 threads
- Each workgroup uses 1 to 1024 threads
- Each CU (Compute Unit) has 4 EUs (Excution Units, or SIMD Units)
- All the wavefronts of a workgroup run on the same CU
- Each EU can have up to 8 active wavefronts ⟶ "Occupancy"
- Each EU has 512 vector registers
- Based on this, …

A SIMPLE MI250X WORD PROBLEM

- Each wavefront has 64 threads
- Each workgroup uses 1 to 1024 threads
- Each CU (Compute Unit) has 4 EUs (Excution Units, or SIMD Units)
- All the wavefronts of a workgroup run on the same CU
- Each EU can have up to 8 active wavefronts ⟶ "Occupancy"
- Each EU has 512 vector registers
- Based on this, ***how did Odysseus navigate the Strait of Messina?***

https://upload.wikimedia.org/wikipedia/commons/1/16/Mosaïque_d%27Ulysse_et_les_sirènes.jpg

THE SCYLLA AND CHARYBDIS OF ACCELERATOR KERNELS

- Scylla: low occupancy
  - Use too many registers ⟶ limit the number of wavefronts active at a time
  - Devours accelerator resources
- Charybdis: spilling registers
  - Limit register use in a kernel
    ⟶ increase the number of wavefronts you can run at a time
    ⟶ spill registers
  - Every thread stores/loads register values to/from memory
  - Sucks down memory bandwidth

https://upload.wikimedia.org/wikipedia/commons/1/19/Denarius_Sextus_Pompeius-Scilla.jpg
https://upload.wikimedia.org/wikipedia/commons/6/6f/Naruto_Whirlpools_taken_4-21-2008.jpg

THE HOMERIC GALLEY OF ACCELERATOR KERNELS

```
__launch_bounds__(TMAX,WMIN) __global__ void kernel(...) { ... }
```

- **TMAX**
  - Upper bound on threads per workgroup
  - Defaults to hardware max (1024)
  - Set lower to let kernel use more registers
    - Avoid spills
    - May reduce occupancy

- **WMIN**
  - Lower bound on wavefronts per EU
  - Default to hardware min (1)
  - Set higher to limit number of kernel registers
    - Improve occupancy
    - May cause spills

NAVIGATING WITH --SAVE-TEMPS

```
export CXX='hipcc'
export CXXFLAGS="-ggdb -O3 -std=c++17 -Wall --save-temps \
  --offload-arch=gfx90a -I${CRAY_MPICH_DIR}/include"
```

Get many new files:

```
main.cc
main.cc-hip-amdgcn-amd-amdhsa.hipfb
main-hip-amdgcn-amd-amdhsa-gfx90a.bc
main-hip-amdgcn-amd-amdhsa-gfx90a.cui
main-hip-amdgcn-amd-amdhsa-gfx90a.o
main-hip-amdgcn-amd-amdhsa-gfx90a.out*
main-hip-amdgcn-amd-amdhsa-gfx90a.out.resolution.txt
main-hip-amdgcn-amd-amdhsa-gfx90a.s
main-host-x86_64-unknown-linux-gnu.bc
main-host-x86_64-unknown-linux-gnu.cui
main-host-x86_64-unknown-linux-gnu.s
main.o
```

NAVIGATING WITH --SAVE-TEMPS

```
export CXX='hipcc'
export CXXFLAGS="-ggdb -O3 -std=c++17 -Wall --save-temps \
  --offload-arch=gfx90a -I${CRAY_MPICH_DIR}/include"
```

Get many new files:

```
main.cc
main.cc-hip-amdgcn-amd-amdhsa.hipfb
main-hip-amdgcn-amd-amdhsa-gfx90a.bc
main-hip-amdgcn-amd-amdhsa-gfx90a.cui
main-hip-amdgcn-amd-amdhsa-gfx90a.o
main-hip-amdgcn-amd-amdhsa-gfx90a.out*
main-hip-amdgcn-amd-amdhsa-gfx90a.out.resolution.txt
main-hip-amdgcn-amd-amdhsa-gfx90a.s
main-host-x86_64-unknown-linux-gnu.bc
main-host-x86_64-unknown-linux-gnu.cui
main-host-x86_64-unknown-linux-gnu.s
main.o
```

the files you want end in gfx90a.s

```
$ less main-hip-amdgcn-amd-amdhsa-gfx90a.s
```

```
        .text
        .amdgcn_target "amdgcn-amd-amdhsa--gfx90a"
        .file   0 ".../Quicksilver/src" "main.cc" md5 0x308b971688524ce790709a3c97d11f63
        .file   1 "/opt/rocm-5.1.0/include/hip/amd_detail" "amd_hip_runtime.h"
        .file   2 "/usr/include/bits" "types.h"
        .file   3 "/usr/include/bits" "stdint-uintn.h"
        .file   4 "/opt/rocm-5.1.0/include/hip" "hip_runtime_api.h"
        .file   5 "/usr/lib64/gcc/x86_64-suse-linux/7/../../../../include/c++/7/ext" "concurrence.h"
        .file   6 "." "Device.hh"
        .file   7 "." "MC_Segment_Outcome.hh"
        .file   8 "." "MC_Tally_Event.hh"
        .file   9 "/opt/rocm-5.1.0/include/hip/amd_detail" "device_library_decls.h"
        .file   10 "." "MC_Facet_Adjacency.hh"
        .file   11 "." "NuclearData.hh"
        .file   12 "/opt/rocm-5.1.0/include/hip/amd_detail" "amd_hip_vector_types.h"
        .file   13 "." "MC_Vector.hh"
        .file   14 "." "MC_Particle.hh"
        .file   15 "." "DirectionCosine.hh"
        .file   16 "." "MC_Location.hh"
        .file   17 "/usr/lib64/gcc/x86_64-suse-linux/7/../../../../include/c++/7/bits" "stringfwd.h"
        .file   18 "." "MC_Nearest_Facet.hh"
        .file   19 "." "MC_Distance_To_Facet.hh"
        .file   20 "/usr/include" "stdlib.h"
        .file   21 "/usr/lib64/gcc/x86_64-suse-linux/7/../../../../include/c++/7/bits" "std_abs.h"
main-hip-amdgcn-amd-amdhsa-gfx90a.s lines 1-24/65725 0%
```

```
        .text
        .amdgcn_target "amdgcn-amd-amdhsa--gfx90a"
        .file   0 ".../Quicksilver/src" "main.cc" md5 0x308b971688524ce790709a3c97d11f63
        .file   1 "/opt/rocm-5.1.0/include/hip/amd_detail" "amd_hip_runtime.h"
        .file   2 "/usr/include/bits" "types.h"
        .file   3 "/usr/include/bits" "stdint-uintn.h"
        .file   4 "/opt/rocm-5.1.0/include/hip" "hip_runtime_api.h"
        .file   5 "/usr/lib64/gcc/x86_64-suse-linux/7/../../../../include/c++/7/ext" "concurrence.h"
        .file   6 "." "Device.hh"
        .file   7 "." "MC_Segment_Outcome.hh"
        .file   8 "." "MC_Tally_Event.hh"
        .file   9 "/opt/rocm-5.1.0/include/hip/amd_detail" "device_library_decls.h"
        .file   10 "." "MC_Facet_Adjacency.hh"
        .file   11 "." "NuclearData.hh"
        .file   12 "/opt/rocm-5.1.0/include/hip/amd_detail" "amd_hip_vector_types.h"
        .file   13 "." "MC_Vector.hh"
        .file   14 "." "MC_Particle.hh"
        .file   15 "." "DirectionCosine.hh"
        .file   16 "." "MC_Location.hh"
        .file   17 "/usr/lib64/gcc/x86_64-suse-linux/7/../../../../include/c++/7/bits" "stringfwd.h"
        .file   18 "." "MC_Nearest_Facet.hh"
        .file   19 "." "MC_Distance_To_Facet.hh"
        .file   20 "/usr/include" "stdlib.h"
        .file   21 "/usr/lib64/gcc/x86_64-suse-linux/7/../../../../include/c++/7/bits" "std_abs.h"
/CycleTrackingGuts
```

search for the kernel you want

```
 .section        .text._ZL17CycleTrackingGutsii6DeviceiPiP15MessageParticle,#alloc,#execinstr
        .globl  _ZL17CycleTrackingGutsii6DeviceiPiP15MessageParticle ; -- Begin function
_ZL17CycleTrackingGutsii6DeviceiPiP15MessageParticle
        .p2align        8
        .type   _ZL17CycleTrackingGutsii6DeviceiPiP15MessageParticle,@function
_ZL17CycleTrackingGutsii6DeviceiPiP15MessageParticle: ;
@_ZL17CycleTrackingGutsii6DeviceiPiP15MessageParticle
.Lfunc_begin0:
        .loc    69 34 0                         ; main.cc:34:0
        .cfi_sections .debug_frame
        .cfi_startproc
; %bb.0:
        .cfi_escape 0x0f, 0x03, 0x30, 0x36, 0xe1 ;
        .cfi_undefined 16
        s_load_dwordx2 s[0:1], s[8:9], 0x0
        s_load_dwordx16 s[12:27], s[8:9], 0x8
.Ltmp0:
        .loc    69 38 21 prologue_end           ; main.cc:38:21
        v_cmp_gt_u32_e64 s[40:41], 8, v0
        .loc    69 38 45 is_stmt 0              ; main.cc:38:45
        v_lshlrev_b32_e32 v3, 3, v0
        s_and_saveexec_b64 s[2:3], s[40:41]
        s_cbranch_execz .LBB0_2
; %bb.1:
        .loc    69 38 66                        ; main.cc:38:66
/Kernel Info
```

then search for Kernel Info

```
; Kernel info:
; codeLenInByte = 35624
; NumSgprs: 104
; NumVgprs: 110
; NumAgprs: 0
; TotalNumVgprs: 110
; ScratchSize: 0
; MemoryBound: 0
; FloatMode: 240
; IeeeMode: 1
; LDSByteSize: 72 bytes/workgroup (compile time only)
; SGPRBlocks: 12
; VGPRBlocks: 13
; NumSGPRsForWavesPerEU: 104
; NumVGPRsForWavesPerEU: 110
; AccumOffset: 112
; Occupancy: 4
; WaveLimiterHint : 1
; COMPUTE_PGM_RSRC2:SCRATCH_EN: 0
; COMPUTE_PGM_RSRC2:USER_SGPR: 10
; COMPUTE_PGM_RSRC2:TRAP_HANDLER: 0
; COMPUTE_PGM_RSRC2:TGID_X_EN: 1
; COMPUTE_PGM_RSRC2:TGID_Y_EN: 0
; COMPUTE_PGM_RSRC2:TGID_Z_EN: 0
; COMPUTE_PGM_RSRC2:TIDIG_COMP_CNT: 0
main-hip-amdgcn-amd-amdhsa-gfx90a.s lines 7557-7581/65725 8%
```

```
; Kernel info:
; codeLenInByte = 35624
; NumSgprs: 104
; NumVgprs: 110
; NumAgprs: 0
; TotalNumVgprs: 110
; ScratchSize: 0
; MemoryBound: 0
; FloatMode: 240
; IeeeMode: 1
; LDSByteSize: 72 bytes/workgroup (compile time only)
; SGPRBlocks: 12
; VGPRBlocks: 13
; NumSGPRsForWavesPerEU: 104
; NumVGPRsForWavesPerEU: 110
; AccumOffset: 112
; Occupancy: 4
; WaveLimiterHint : 1
; COMPUTE_PGM_RSRC2:SCRATCH_EN: 0
; COMPUTE_PGM_RSRC2:USER_SGPR: 10
; COMPUTE_PGM_RSRC2:TRAP_HANDLER: 0
; COMPUTE_PGM_RSRC2:TGID_X_EN: 1
; COMPUTE_PGM_RSRC2:TGID_Y_EN: 0
; COMPUTE_PGM_RSRC2:TGID_Z_EN: 0
; COMPUTE_PGM_RSRC2:TIDIG_COMP_CNT: 0
main-hip-amdgcn-amd-amdhsa-gfx90a.s lines 7557-7581/65725 8%
```

110 vector registers means occupancy of 512/110 = 4.65

occupancy of 4 means up to 512/4 = 128 vector registers

```
; Kernel info:
; codeLenInByte = 35624
; NumSgprs: 104
; NumVgprs: 110
; NumAgprs: 0
; TotalNumVgprs: 110
; ScratchSize: 0
; MemoryBound: 0
; FloatMode: 240
; IeeeMode: 1
; LDSByteSize: 72 bytes/workgroup (compile time only)
; SGPRBlocks: 12
; VGPRBlocks: 13
; NumSGPRsForWavesPerEU: 104
; NumVGPRsForWavesPerEU: 110
; AccumOffset: 112
; Occupancy: 4
; WaveLimiterHint : 1
; COMPUTE_PGM_RSRC2:SCRATCH_EN: 0
; COMPUTE_PGM_RSRC2:USER_SGPR: 10
; COMPUTE_PGM_RSRC2:TRAP_HANDLER: 0
; COMPUTE_PGM_RSRC2:TGID_X_EN: 1
; COMPUTE_PGM_RSRC2:TGID_Y_EN: 0
; COMPUTE_PGM_RSRC2:TGID_Z_EN: 0
; COMPUTE_PGM_RSRC2:TIDIG_COMP_CNT: 0
/spill_count
```

one more thing to search for

```
 .sgpr_spill_count: 202
    .symbol:          _ZL17CycleTrackingGutsii6DeviceiPiP15MessageParticle.kd
    .vgpr_count:      110
    .vgpr_spill_count: 0
    .wavefront_size: 64
amdhsa.target:   amdgcn-amd-amdhsa--gfx90a
amdhsa.version:
  - 1
  - 1
...

        .end_amdgpu_metadata
        .section        .debug_line
.Lline_table_start0:
~
~
~
~
~
~
~
~
~
~
~
main-hip-amdgcn-amd-amdhsa-gfx90a.s lines 65712-65725/65725 (END)
```

zero is good

CUT TO THE CHASE

```
$ grep -e GPRs -e Occupancy -e spill_count main-hip-amdgcn-amd-amdhsa-gfx90a.s
; NumSGPRsForWavesPerEU: 104
; NumVGPRsForWavesPerEU: 110
; Occupancy: 4
    .sgpr_spill_count: 202
    .vgpr_spill_count: 0
```

THE CHARYBDIS OPTION

Crank up the occupancy!

```
__global__ __launch_bounds__(1024,8)
static void CycleTrackingGuts( const int ipMin, int ipMax, Device device,
  const int maxCount, int *__restrict__ const sendCounts,
  MessageParticle *__restrict__ const sendParts)
{
  ...
```

```
$ grep -e GPRs -e Occupancy -e spill_count main-hip-amdgcn-amd-amdhsa-gfx90a.s
; NumSGPRsForWavesPerEU: 78
; NumVGPRsForWavesPerEU: 64
; Occupancy: 8
    .sgpr_spill_count: 248
    .vgpr_spill_count: 255
```

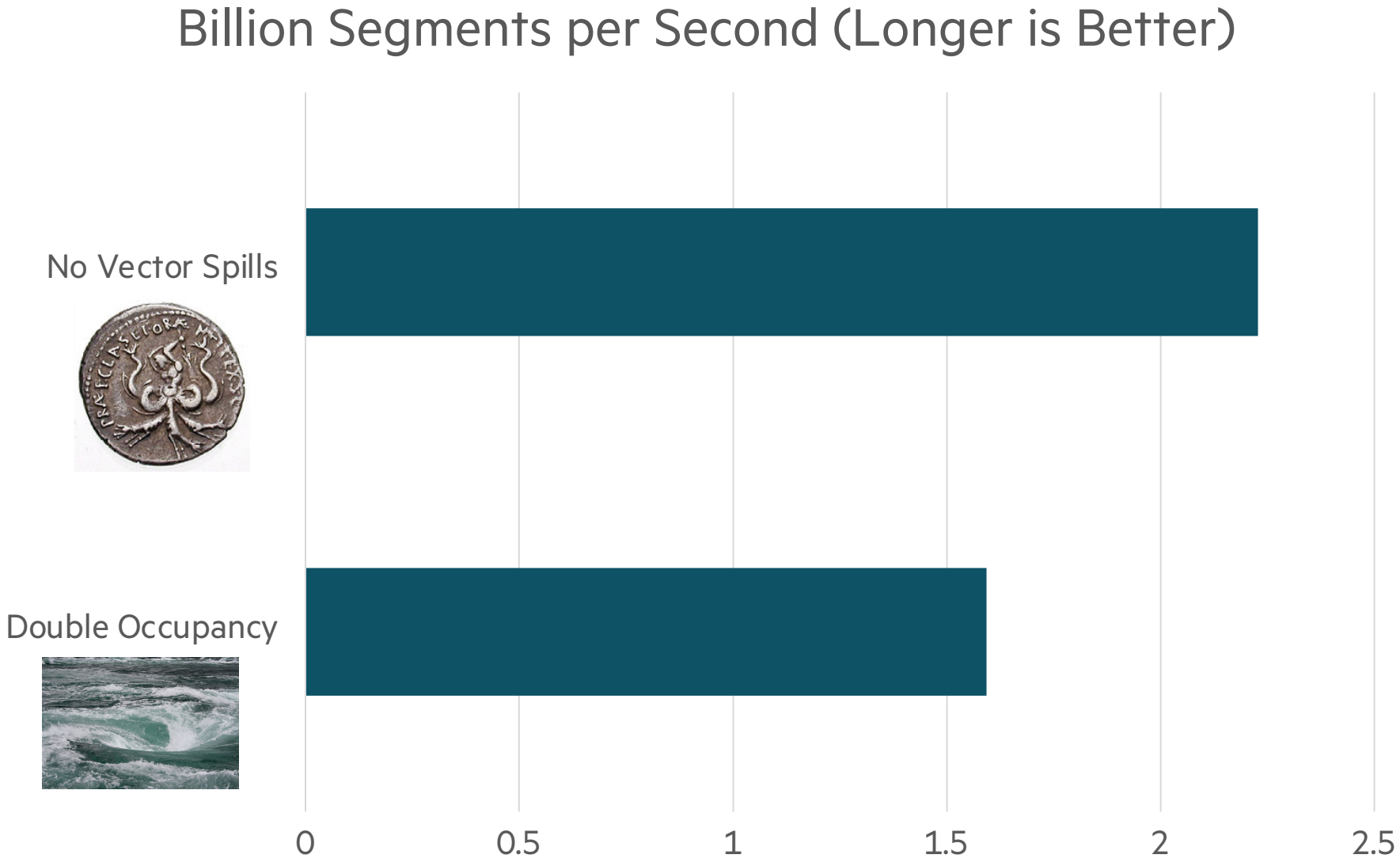
= 512/8

Success!

What, me worry?

SCYLLA VS. CHARYBDIS

- Quicksilver benchmark
- 2x2x2 grid
- Single node of Crusher
- Rocm 5.1.0
- 32000000 particles

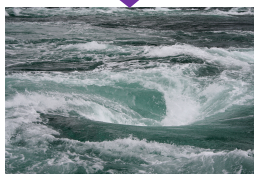


vector-register spills are often worse than lower occupancy

HOW TO WIDEN THE STRAIT OF MESSINA

or *How to Reduce Register Pressure*

- Avoid *assert* and *printf* in kernels
  - Use *if (…) abort();* instead of *assert(…);*
- Wait for compiler improvements
- Avoid arrays on the stack in device code
- *<looking for more strategies>*



AVOIDING ARRAYS ON THE STACK

- Look for this pattern
  - Small constant-size work arrays
    - Often size [3] for X, Y, Z directions
  - A small loop sets the values
  - Other small loops use the values
- Try to replace it
  - Fuse loops
  - Replace each array with a single scalar inside the loop
  - Reduce vector-register use!

```
static constexpr int N = 3;
double a[N], b[N];
for (int i = 0; i < N; i++) {
  a[i] = computeA(i);
}
for (int i = 0; i < N; i++) {
  b[i] = computeB(i);
}
...
double c = 0;
for (int i = 0; i < N; i++) {
  c += computeC(a[i],b[i])
}

double c = 0;
for (int i = 0; i < N; i++) {
  const double a = computeA(i);
  const double b = computeB(i);
  c += computeC(a,b);
}
```

COMPILER TAKEAWAYS

- If you know the maximum workgroup size *N* that a kernel will use, tell the compiler with *__launch_bounds__(N)*
- You probably don't want to force higher occupancy with the second argument of *__launch_bounds__*
- Generate annotated assembly code for kernels with compiler argument *--save-temps*
- Search assembly files for register use, occupancy, and register spills: *grep -e GPRs -e Occupancy -e spill_count *-gfx90a.s*
- Look out for low occupancy and nonzero vector spills
- Try to reduce register pressure in code

THE END IS NEAR

or *What's for lunch?*

WISHFUL-THINKING PIE-IN-THE-SKY FANTASY DREAM TOPICS FOR SOME UNKNOWN TIME IN THE FUTURE MAYBE

- OpenMP offload
- Hardware performance counters
- In-kernel profiling
- Compiler reports
- Roofline plots
- Performance-tuning success stories

MORE INFORMATION ON PERFORMANCE PROFILING

- HPE
  - *man perftools*
  - *pat_report -O -h*
  - https://support.hpe.com/ → search "performance analysis tools"
- Cool stuff from others (also consider these acknowledgements)
  - AMD ROCm Platform
    https://rocmdocs.amd.com/en/latest/
    https://rocmdocs.amd.com/en/latest/ROCm_Tools/ROCm-Tools.html
    https://rocmdocs.amd.com/en/latest/Programming_Guides/HIP-porting-guide.html
    rocprof --help
  - Google Chrome tracing
    https://sites.google.com/a/chromium.org/dev/developers/how-tos/trace-event-profiling-tool/frame-viewer

Hewlett Packard
Enterprise

DEBUGGING AND PERFORMANCE PROFILING ON HPE CRAY SUPERCOMPUTERS WITH AMD GPUS

Steve Abbott
stephen.abbott@hpe.com
Trey White
trey.white@hpe.com
Kostas Makrides
makrides@hpe.com