TAU Commander Hands On Exercises

Dr. John C. Linford
ParaTools, Inc.
jlinford@paratools.com

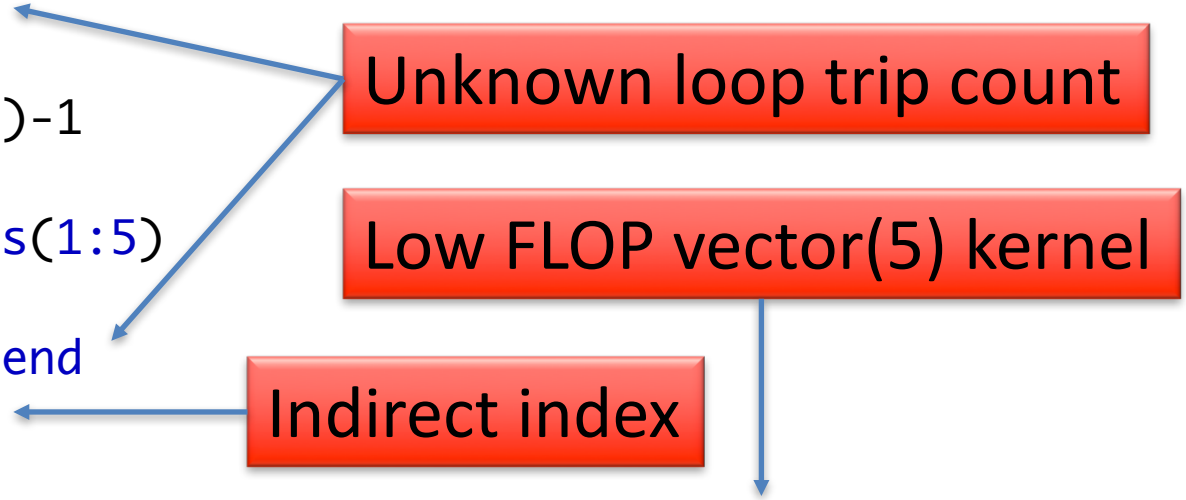
Dr. John Wohlbier
Engility Corp.
John.Wohlbier.ctr@engilitycorp.com

KNL Workshop

25 April 2017, ARL, Aberdeen, MD

MINIAPP 1 Code Structure

```
do sweep = 1, n_sweeps
  do color = sweep_start, sweep_end, sweep_stride
    do ipass = 1, 2
      start = color_indices(1,color)
      end = color_boundary_end(color)

      do n = start, end
        istart = iam(n)
        iend   = iam(n+1)-1

        f(1:5) = (+/-)res(1:5)

        do j = istart, iend
          icol = jam(j)
          do i = 1, 5
            f(1:5) = f(1:5) - a_off(1:5,i,j)*dq(i,icol)
          end do
```



MINIAPP 1 Kernel Unrolled

```
do j = istart,iend
  icol = jam(j)
  f1 = f1   - a_off(1,1,j)*dq(1,icol)
  f2 = f2   - a_off(2,1,j)*dq(1,icol)
  f3 = f3   - a_off(3,1,j)*dq(1,icol)
  f4 = f4   - a_off(4,1,j)*dq(1,icol)
  f5 = f5   - a_off(5,1,j)*dq(1,icol)
  f1 = f1   - a_off(1,2,j)*dq(2,icol)
  f2 = f2   - a_off(2,2,j)*dq(2,icol)
  f3 = f3   - a_off(3,2,j)*dq(2,icol)
  f4 = f4   - a_off(4,2,j)*dq(2,icol)
  f5 = f5   - a_off(5,2,j)*dq(2,icol)
  f1 = f1   - a_off(1,3,j)*dq(3,icol)
  f2 = f2   - a_off(2,3,j)*dq(3,icol)
  f3 = f3   - a_off(3,3,j)*dq(3,icol)
  f4 = f4   - a_off(4,3,j)*dq(3,icol)
  f5 = f5   - a_off(5,3,j)*dq(3,icol)
  f1 = f1   - a_off(1,4,j)*dq(4,icol)
  f2 = f2   - a_off(2,4,j)*dq(4,icol)
  f3 = f3   - a_off(3,4,j)*dq(4,icol)
  f4 = f4   - a_off(4,4,j)*dq(4,icol)
  f5 = f5   - a_off(5,4,j)*dq(4,icol)
  f1 = f1   - a_off(1,5,j)*dq(5,icol)
  f2 = f2   - a_off(2,5,j)*dq(5,icol)
  f3 = f3   - a_off(3,5,j)*dq(5,icol)
  f4 = f4   - a_off(4,5,j)*dq(5,icol)
  f5 = f5   - a_off(5,5,j)*dq(5,icol)
end do
```

```
do j = istart, iend
  icol = jam(j)
  do i = 1, 5
    f(1:5) = f(1:5) - a_off(1:5,i,j)*dq(i,icol)
  end do
```

- 56 Loads
- 26 Stores
- 50 FP-ops
  - Fused to 25
- ~0.17 FP-ops / byte
  - Fused: 0.083 FP-ops / byte

ARL's KNL System

Utility server as a SSH gateway

```
$ ssh -Y -K us.arl.hpc.mil
$ ssh -Y knl01
```

Show available SLURM partitions

```
$ sinfo
```

Start an interactive job

```
$ srun --pty -p pettt-qf $SHELL
```

ARL's KNL System

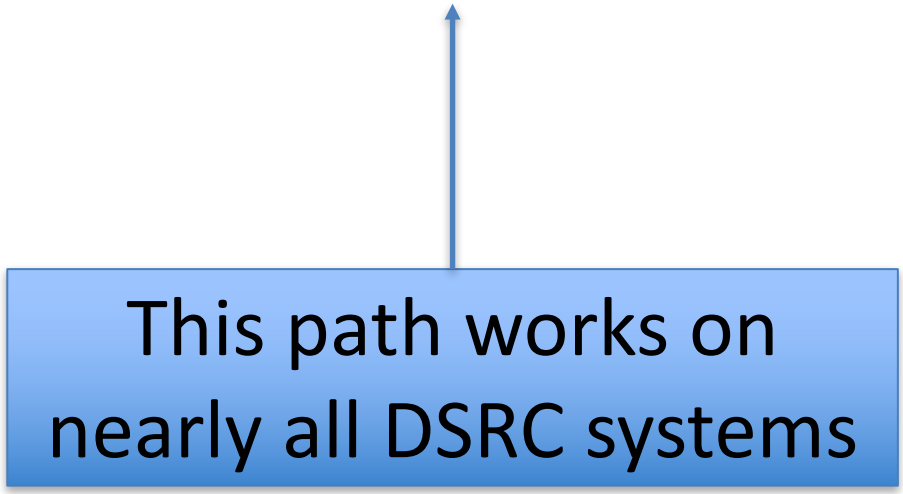
Submit a batch job

```
$ sbatch script.sh
```

```
#!/bin/bash
#SBATCH --partition=quadrant-flat

cd $HOME/FUN3D_Miniapp1
module load intel intelmpi
numactl -m 1 ./point_solve
```

TAU Commander on ARL's KNL

```
$ module load intel intelmpi

$ export PATH=\
    $PET_HOME/pkgs/taucmdr-latest/bin:$PATH
```



TAU Commander on Thunder

```
$ ssh -Y thunder.afrl.hpc.mil

$ export PATH=\
    $PET_HOME/pkgs/taucmdr-latest/bin:$PATH
```

Getting Started with TAU Commander

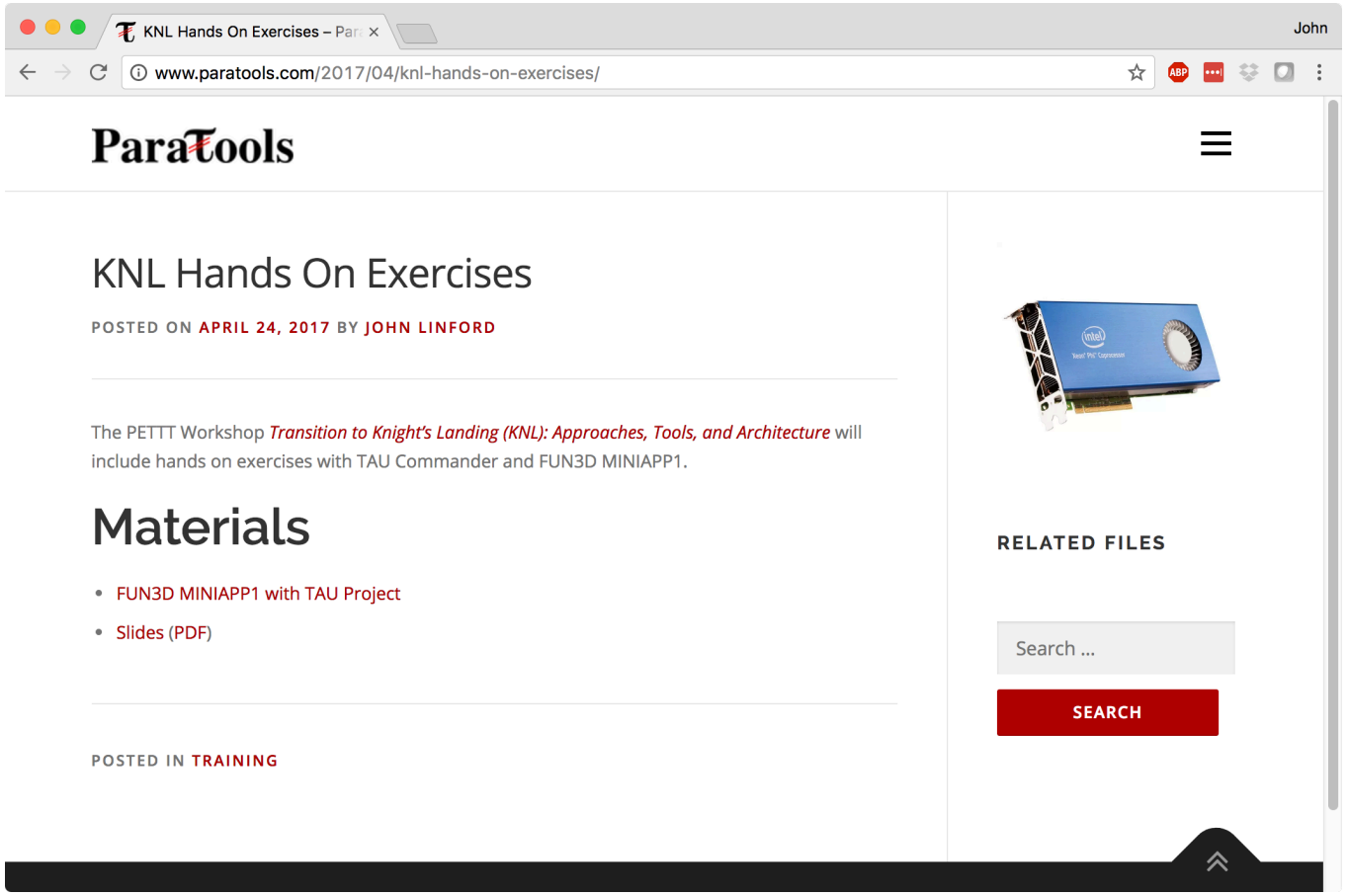
- **tau init** [options | --help]
- **tau** <<your compiler>> foo.c
  - e.g. tau mpif90 foo.f90
- **tau** mpirun -np 8 ./a.out

Online Help

- **tau --help**
- **tau help** <something>

Workshop Materials

http://www.paratools.com/knl-hands-on-exercises/



Initialize the TAU Project

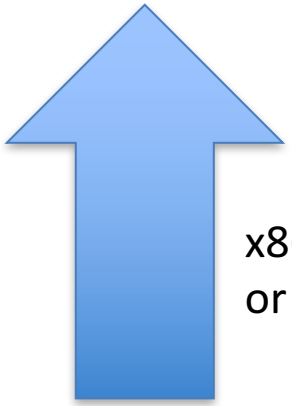
Workshop materials on KNL:

```
$ tar xzf $PET_HOME/pkgs/FUN3D_Miniapp1.tgz
```

Download from paratools.com if on another system

```
$ cd FUN3D_Miniapp1/
```

```
$ tau initialize --arch KNL
```



x86_64 on Thunder,
or don't use --arch

Default TAU Project for Miniapp 1

Desktop — jlinford@abutill-0001:~ — ssh -F ~/.ssh/hpcmp_config -Y -K us.arl.hpc.mil — 157×63

```
-bash-4.2$ tau init --arch KNL --compilers Intel
[TAU] Intel MPI C++ compiler '/work1/compiler/impi/2017.2.174/intel64/bin/mpiicpc' wraps
[TAU]     '/work1/compiler/compilers_and_libraries_2017.2.174/linux/bin/intel64/icpc'
[TAU] Intel MPI C compiler '/work1/compiler/impi/2017.2.174/intel64/bin/mpiicc' wraps
[TAU]     '/work1/compiler/compilers_and_libraries_2017.2.174/linux/bin/intel64/icc'
[TAU] Intel MPI Fortran compiler '/work1/compiler/impi/2017.2.174/intel64/bin/mpiifort' wraps
[TAU]     '/work1/compiler/compilers_and_libraries_2017.2.174/linux/bin/intel64/ifort'
[TAU] Created a new project named 'FUN3D_Miniapp1'.
[TAU] Added application 'FUN3D_Miniapp1' to project configuration 'FUN3D_Miniapp1'.
[TAU] Added target 'knl-login01' to project configuration 'FUN3D_Miniapp1'.
[TAU] Added measurement 'sample' to project configuration 'FUN3D_Miniapp1'.
[TAU] Added measurement 'profile' to project configuration 'FUN3D_Miniapp1'.
[TAU] Added measurement 'trace' to project configuration 'FUN3D_Miniapp1'.
[TAU] Created a new experiment 'knl-login01-FUN3D_Miniapp1-sample'
[TAU] Selected experiment 'knl-login01-FUN3D_Miniapp1-sample'.

== Project Configuration (/work1/people/jlinford/FUN3D_Miniapp1/.tau/project.json) ==========

+---------------+---------------+---------------+-------------------+---------------+
|      Name     |    Targets    |  Applications |    Measurements   | # Experiments |
+===============+===============+===============+===================+===============+
| FUN3D_Miniapp1 | knl-login01 | FUN3D_Miniapp1 | sample, profile, trace |       1       |
+---------------+---------------+---------------+-------------------+---------------+

== Targets in project 'FUN3D_Miniapp1' ==========

+-------------+---------+-----------+----------------+---------------+-----------------+
|     Name    | Host OS | Host Arch | Host Compilers | MPI Compilers | SHMEM Compilers |
+=============+=========+===========+================+===============+=================+
| knl-login01 |  Linux  |    KNL    |      Intel     |     Intel     |       None      |
+-------------+---------+-----------+----------------+---------------+-----------------+

== Applications in project 'FUN3D_Miniapp1' ==========

+---------------+---------+--------+----------+-----+-----+------+--------+-------+-----+
|      Name     | Linkage | OpenMP | Pthreads | TBB | MPI | CUDA | OpenCL | SHMEM | MPC |
+===============+=========+========+==========+=====+=====+======+========+=======+=====+
| FUN3D_Miniapp1 | dynamic |   No   |    No    |  No |  No |  No  |   No   |   No  |  No |
+---------------+---------+--------+----------+-----+-----+------+--------+-------+-----+

== Measurements in project 'FUN3D_Miniapp1' ==========

+---------+---------+-------+--------+--------------+---------------+--------+------+-----+-----+--------+
|   Name  | Profile | Trace | Sample | Source Inst. | Compiler Inst. | OpenMP | CUDA | I/O | MPI | SHMEM  |
+=========+=========+=======+========+==============+===============+========+======+=====+=====+========+
|  sample |   tau   |  none |   Yes  |     never    |     never     | ignore |  No  |  No |  No |   No   |
+---------+---------+-------+--------+--------------+---------------+--------+------+-----+-----+--------+
| profile |   tau   |  none |   No   |   automatic  |    fallback   | ignore |  No  |  No |  No |   No   |
+---------+---------+-------+--------+--------------+---------------+--------+------+-----+-----+--------+
|  trace  |  none   | slog2 |   No   |   automatic  |    fallback   | ignore |  No  |  No |  No |   No   |
+---------+---------+-------+--------+--------------+---------------+--------+------+-----+-----+--------+

== Experiments in project 'FUN3D_Miniapp1' ==========

+-------------------------------+--------+-----------+-------------+---------------+-------------+-------------------------+
|              Name             | Trials | Data Size |    Target   |  Application  | Measurement |       TAU Makefile      |
+===============================+========+===========+=============+===============+=============+=========================+
| knl-login01-FUN3D_Miniapp1-sample |   0    |   0.0B    | knl-login01 | FUN3D_Miniapp1 |   sample    | Makefile.tau-icpc-pthread |
+-------------------------------+--------+-----------+-------------+---------------+-------------+-------------------------+

Selected Experiment: knl-login01-FUN3D_Miniapp1-sample

-bash-4.2$
```

Modify project if needed

```
$ tau target --help
usage: tau target <subcommand> [arguments]

Create and manage target configurations.

Positional Arguments:
  <subcommand>  See 'subcommands' below.
  [arguments]   Arguments to be passed to <subcommand>.

Optional Arguments:
  -h, --help    Show this help message and exit.

Subcommands:
  copy            Copy and modify target configurations.
  create          Create target configurations.
  delete          Delete target configurations.
  edit            Modify target configurations.
  list            Show target configuration data.
  metrics         Show metrics available on this target.

See 'tau target <subcommand> --help' for more information on <subcommand>.
```

```
$ tau target edit knl-login01 --new-name knl
```

Use `tau` to compile

```
$ cd baseline
$ vi Makefile

  1 # Fortran Compiler
  2 FC = tau ifort
  3 #FC = mpif90
  4
...
 16
 17 # Program name(s)
 18 PROGRAMS = point_solve
 19
 20 .PHONY: all clean run
 21
 22 all: $(PROGRAMS)
 23
 24 run: all
 25   tau ./point_solve
 26
```

Prepend `tau` command

Prepend `tau` command

Use `tau` to compile

```
$ make
```

```
Desktop — jlinford@abutill-0002:/app/pet/pkgs — ssh -F ~/.ssh/hpcmp_config us.arl.hpc.mil — 169×26
tau ifort  -convert big_endian -O3 -ip -align array64byte -fno-alias -g -traceback -qopt-report=5 -std03 -xMIC-AVX512 -qopenmp  -c lmpi.F90
[TAU] TAU_MAKEFILE=/work1/cta/pet/pkgs/taucmdr-latest/system/tau/21ad983190f15c2428aa364beed419feaad96f0b/x86_64/lib/Makefile.tau-icpc-pthread
[TAU] TAU_OPTIONS=-optNoCompInst -optLinkOnly -optQuiet
[TAU] /work1/compiler/compilers_and_libraries_2017.2.174/linux/bin/intel64/ifort -g -convert big_endian -O3 -ip -align array64byte -fno-alias -g -traceback
[TAU]     -qopt-report=5 -std03 -xMIC-AVX512 -qopenmp -c lmpi.F90
ifort: remark #10397: optimization reports are generated in *.optrpt files in the output location
tau ifort  -convert big_endian -O3 -ip -align array64byte -fno-alias -g -traceback -qopt-report=5 -std03 -xMIC-AVX512 -qopenmp  -c lmpi_app.F90
[TAU] TAU_MAKEFILE=/work1/cta/pet/pkgs/taucmdr-latest/system/tau/21ad983190f15c2428aa364beed419feaad96f0b/x86_64/lib/Makefile.tau-icpc-pthread
[TAU] TAU_OPTIONS=-optNoCompInst -optLinkOnly -optQuiet
[TAU] /work1/compiler/compilers_and_libraries_2017.2.174/linux/bin/intel64/ifort -g -convert big_endian -O3 -ip -align array64byte -fno-alias -g -traceback
[TAU]     -qopt-report=5 -std03 -xMIC-AVX512 -qopenmp -c lmpi_app.F90
ifort: remark #10397: optimization reports are generated in *.optrpt files in the output location
tau ifort -convert big_endian -O3 -ip -align array64byte -fno-alias -g -traceback -qopt-report=5 -std03 -xMIC-AVX512 -qopenmp  -c point_solve.f90
[TAU] TAU_MAKEFILE=/work1/cta/pet/pkgs/taucmdr-latest/system/tau/21ad983190f15c2428aa364beed419feaad96f0b/x86_64/lib/Makefile.tau-icpc-pthread
[TAU] TAU_OPTIONS=-optNoCompInst -optLinkOnly -optQuiet
[TAU] /work1/compiler/compilers_and_libraries_2017.2.174/linux/bin/intel64/ifort -g -convert big_endian -O3 -ip -align array64byte -fno-alias -g -traceback
[TAU]     -qopt-report=5 -std03 -xMIC-AVX512 -qopenmp -c point_solve.f90
ifort: remark #10397: optimization reports are generated in *.optrpt files in the output location
tau ifort -convert big_endian -O3 -ip -align array64byte -fno-alias -g -traceback -qopt-report=5 -std03 -xMIC-AVX512 -qopenmp  -o point_solve kinddefs.o sort.o interp_de
fs.o allocations.o point_solve.o lmpi_app.o lmpi.o system_extensions.o
[TAU] TAU_MAKEFILE=/work1/cta/pet/pkgs/taucmdr-latest/system/tau/21ad983190f15c2428aa364beed419feaad96f0b/x86_64/lib/Makefile.tau-icpc-pthread
[TAU] TAU_OPTIONS=-optNoCompInst -optLinkOnly -optQuiet
[TAU] /work1/compiler/compilers_and_libraries_2017.2.174/linux/bin/intel64/ifort -g -convert big_endian -O3 -ip -align array64byte -fno-alias -g -traceback
[TAU]     -qopt-report=5 -std03 -xMIC-AVX512 -qopenmp -o point_solve kinddefs.o sort.o interp_defs.o allocations.o point_solve.o lmpi_app.o lmpi.o system_extensions.o
ifort: remark #10397: optimization reports are generated in *.optrpt files in the output location
-bash-4.2$
```

TAU Commander constructs a new command line to match the selected experiment.

- May replace compiler commands with TAU's compiler wrapper scripts.
- May set environment variables, parse configuration files, etc.
- If no changes are required then nothing is changed.

Use `tau` to run

```
$ pwd
~/FUN3D_Miniapp1/baseline
```

```
$ srun --pty -p pettt-qf $SHELL
```

Or, on Thunder

```
$ qsub -A <proj> -l ncpus=36 -q debug \
       -l walltime=00:30:00 -I
```

```
$ tau ./point_solve
```

Same on all platforms: X86_64, KNL, PowerPC…

Miniapp 1 Output with TAU

```
Desktop — jlinford@abutill-0002:/app/pet/pkgs — ssh -F ~/.ssh/hpcmp_config us.arl.hpc.mil — 122×47
-bash-4.2$ srun --pty -p pettt-qf $SHELL
bash-4.2$ tau ./point_solve
[TAU]
[TAU] == BEGIN Experiment at 2017-04-24 22:13:24.757560 ========================================================================
[TAU]
[TAU] TAU_CALLPATH=1
[TAU] TAU_CALLPATH_DEPTH=100
[TAU] TAU_COMM_MATRIX=0
[TAU] TAU_METRICS=TIME,
[TAU] TAU_PROFILE=1
[TAU] TAU_SAMPLING=1
[TAU] TAU_THROTTLE=1
[TAU] TAU_THROTTLE_NUMCALLS=100000
[TAU] TAU_THROTTLE_PERCALL=10
[TAU] TAU_TRACE=0
[TAU] TAU_TRACK_HEAP=0
[TAU] TAU_VERBOSE=0
[TAU] tau_exec -T serial,icpc -ebs ./point_solve
 Loading data...
0 Number of block 5x5 equations in data file: 1123718
 Done loading data...
 Solving Ax=b...
 Sweep seconds on master =    1.130100
 Sweep seconds on master =    1.095300
 Sweep seconds on master =    1.090900
 Sweep seconds on master =    1.091000
 Sweep seconds on master =    1.090800
 Sweep seconds on master =    1.091000
 Sweep seconds on master =    1.091100
 Sweep seconds on master =    1.090600
 Sweep seconds on master =    1.090900
 Sweep seconds on master =    1.090700
 Sweep seconds on master =    1.090800
 Sweep seconds on master =    1.090600
 Sweep seconds on master =    1.091100
 Sweep seconds on master =    1.090900
 Sweep seconds on master =    1.090600
 Total seconds taken on master =    16.40840
 Test passed.
[TAU]
[TAU] == END Experiment at 2017-04-24 22:13:47.025716 ==========================================================================
[TAU]
[TAU] Trial 0 produced 1 profile files.
bash-4.2$
```

Tracks experiment metadata

Sets appropriate environment variables

Stores generated data in a performance database.

Use `tau` to view data

```
$ tau show
```

No java on compute nodes

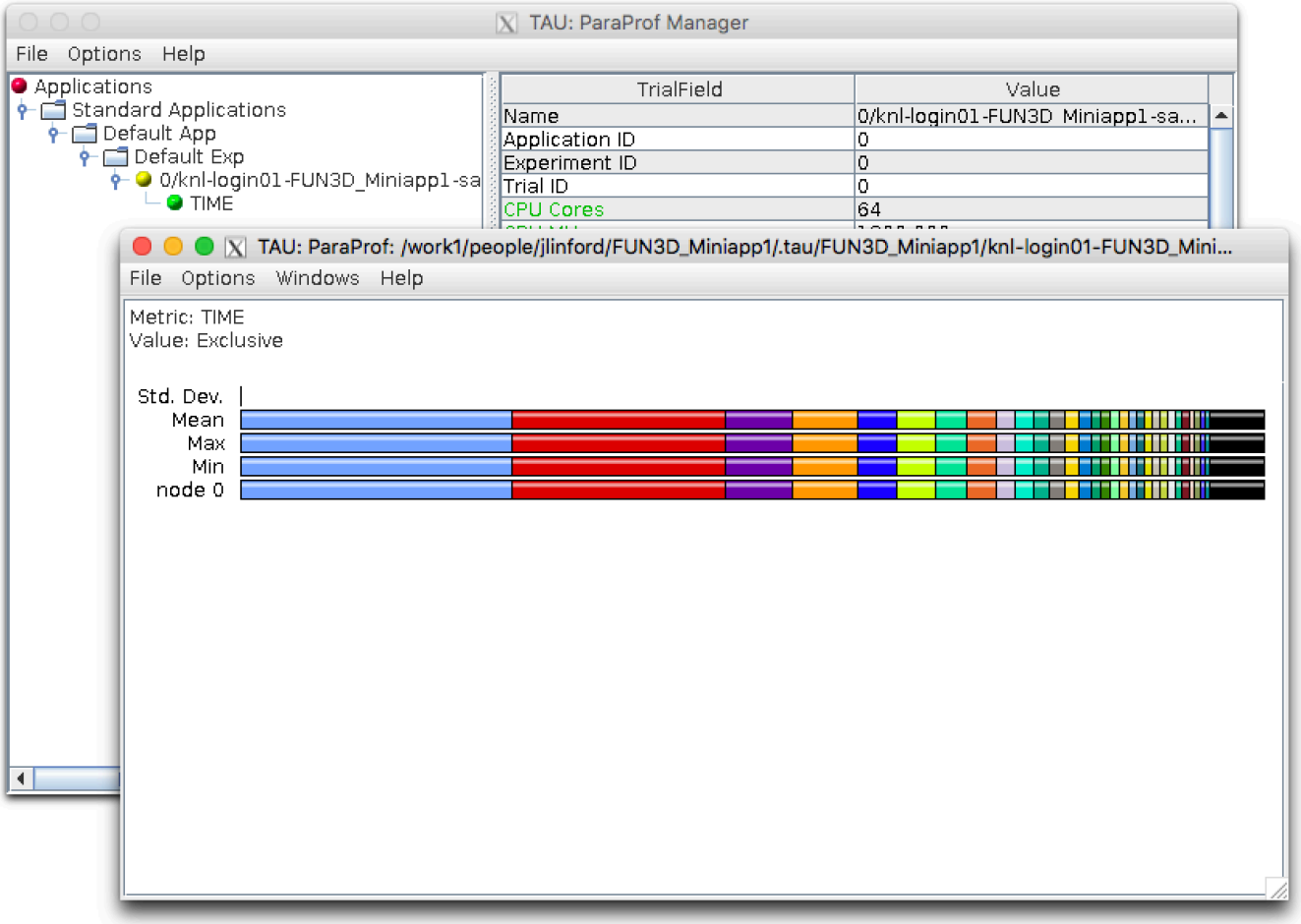
Desktop — jlinford@abutill-0002:/app/pet/pkgs — ssh -F ~/.ssh/hpcmp_config us.arl.hpc.mil — 157×47

```
bash-4.2$ tau show
which: no java in (/work1/cta/pet/pkgs/taucmdr-latest/system/tau/21ad983190f15c2428aa364beed419feaad96f0b/x86_64/bin:/work1/compiler/impi/2017.2.174/intel64/
bin:/work1/compiler/vtune_amplifier_xe/bin64:/work1/compiler/debugger_2017/gdb/intel64/bin:/work1/compiler/compilers_and_libraries_2017.2.174/linux/bin/intel
64:/usr/cta/pet/pkgs/taucmdr-latest/bin:/usr/lib64/qt-3.3/bin:/usr/sbin:/usr/bsd:/sbin:/usr/bin:/bin:/usr/bin/X11:/usr/krb5/bin:/usr/brl/bin:/usr/local/sbin)
readlink: missing operand
Try 'readlink --help' for more information.
dirname: missing operand
Try 'dirname --help' for more information.
/work1/cta/pet/pkgs/taucmdr-latest/system/tau/21ad983190f15c2428aa364beed419feaad96f0b/x86_64/bin/paraprof: line 166: java: command not found
[TAU] ********************************************************************************************************************************
[TAU]
[TAU] WARNING
[TAU]
[TAU] /work1/cta/pet/pkgs/taucmdr-latest/system/tau/21ad983190f15c2428aa364beed419feaad96f0b/x86_64/bin/paraprof failed
[TAU]
[TAU] ********************************************************************************************************************************
Reading Profile files in profile.*

NODE 0;CONTEXT 0;THREAD 0:
---------------------------------------------------------------------------------------
%Time    Exclusive    Inclusive       #Call      #Subrs  Inclusive Name
              msec   total msec                          usec/call
---------------------------------------------------------------------------------------
100.0       20,613       20,613           1           0   20613485 .TAU application
 99.3            0       20,460        2046           0      10000 .TAU application => [CONTEXT] .TAU application
 99.3            0       20,460        2046           0      10000 [CONTEXT] .TAU application
 78.9       16,270       16,270        1627           0      10000 .TAU application => [CONTEXT] .TAU application => [SAMPLE] solve_module_mp_point_solve_5_
 78.9       16,270       16,270        1627           0      10000 [SAMPLE] solve_module_mp_point_solve_5_
 14.2        2,921        2,921         292           0      10006 .TAU application => [CONTEXT] .TAU application => [SAMPLE] __read_nocancel
 14.2        2,921        2,921         292           0      10006 [SAMPLE] __read_nocancel
  5.9        1,208        1,208         121           0       9987 .TAU application => [CONTEXT] .TAU application => [SAMPLE] for_read_seq_xmit
  5.9        1,208        1,208         121           0       9987 [SAMPLE] for_read_seq_xmit
  0.2           39           39           4           0       9996 .TAU application => [CONTEXT] .TAU application => [SAMPLE] MAIN__
  0.2           39           39           4           0       9996 [SAMPLE] MAIN__
  0.0           10           10           1           0      10280 .TAU application => [CONTEXT] .TAU application => [SAMPLE] __kmp_ignore_mppbeg
  0.0           10           10           1           0      10280 [SAMPLE] __kmp_ignore_mppbeg
  0.0            9            9           1           0       9681 .TAU application => [CONTEXT] .TAU application => [SAMPLE] for__read_input
  0.0            9            9           1           0       9681 [SAMPLE] for__read_input
bash-4.2$
```

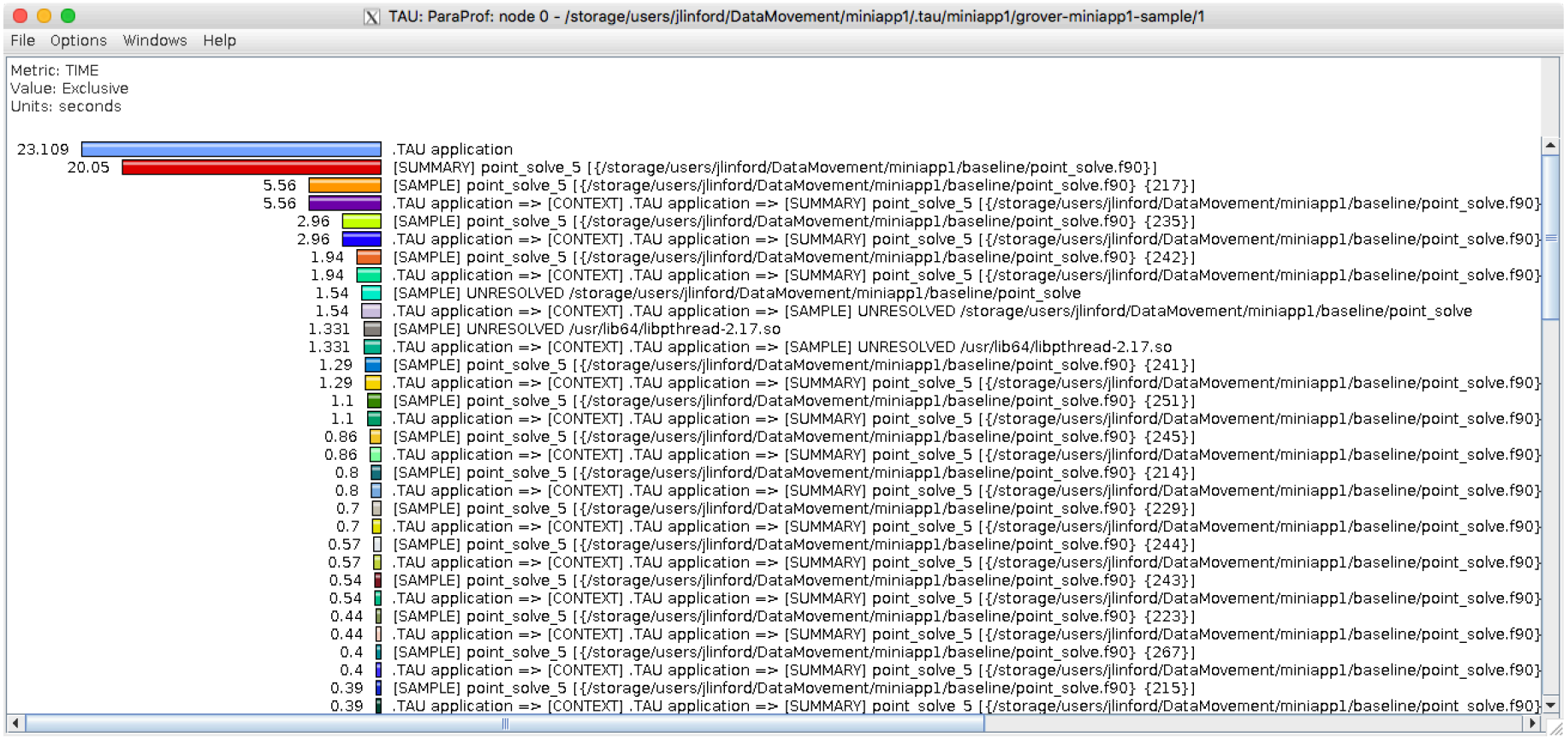
Fall back to console output

View profile from the head node

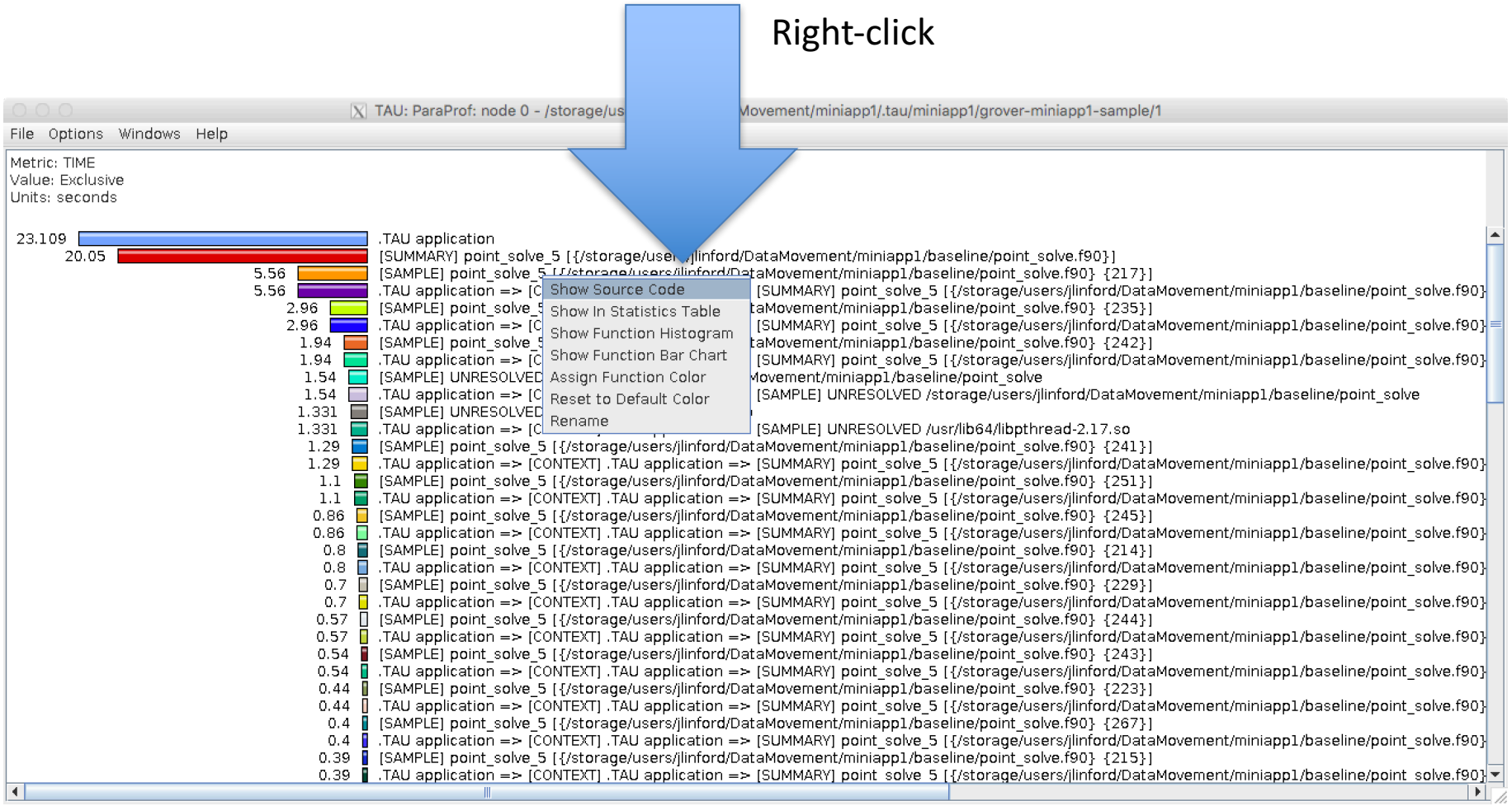
```
$ tau show
```



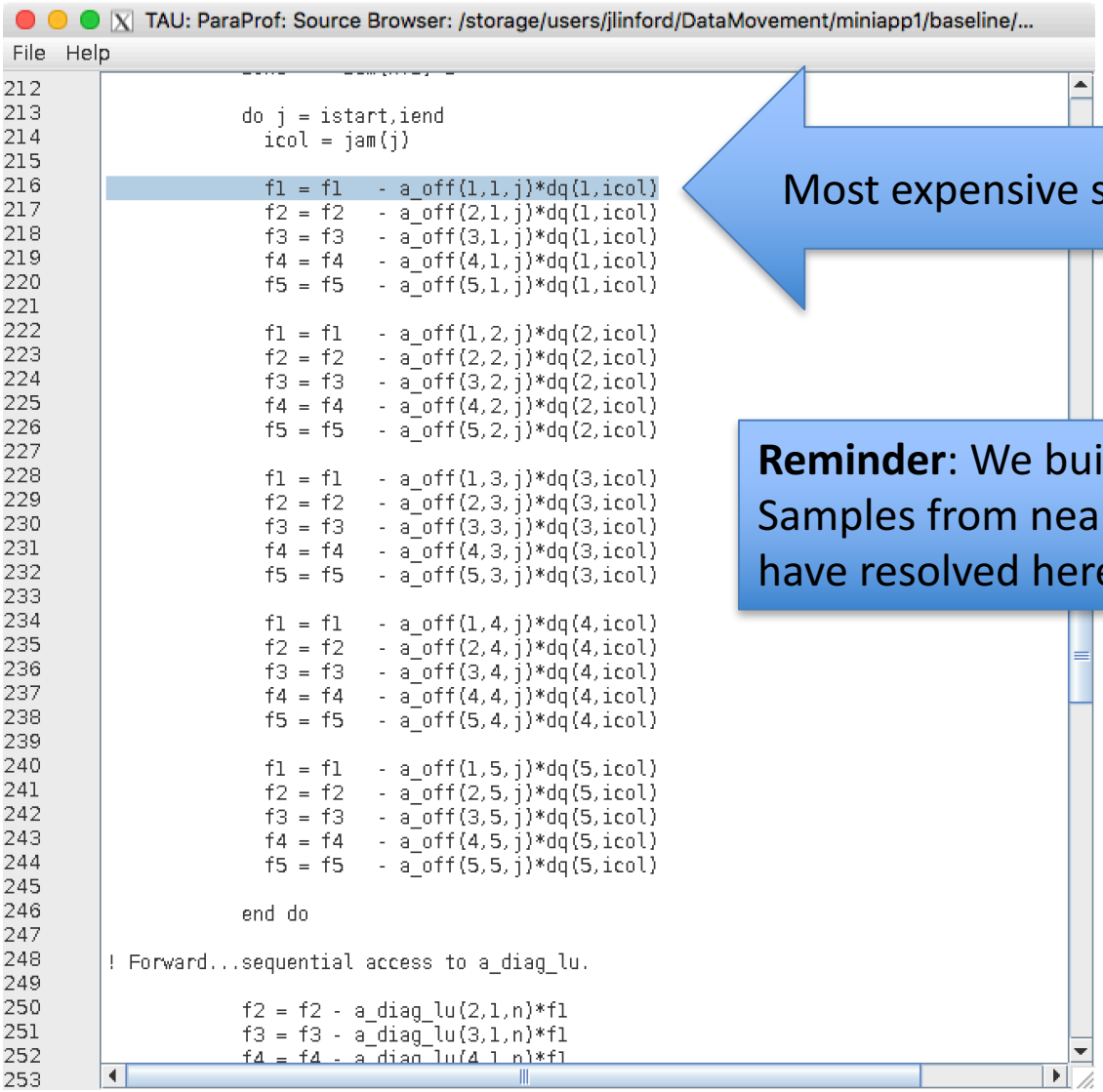
Node 0 Exclusive Time Profile



View Source Code

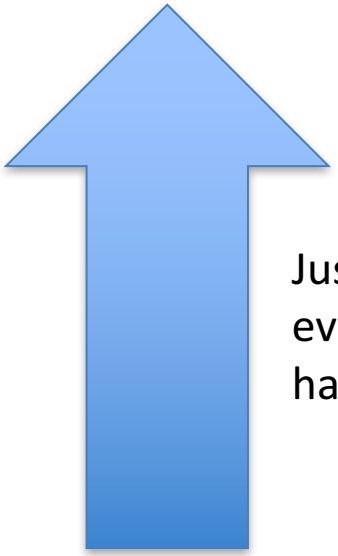


View Source Code



How to find most expensive line of code

1. **tau** initialize
2. **tau** ifort *.f90 -o foo
3. **tau** ./foo
4. **tau** show



Just put `**tau**` in front of everything and see what happens.

- This works on any supported system, even if TAU is not installed or has not been configured appropriately.
- TAU and all its dependencies will be downloaded and installed if required.

Use MCDRAM via numactl -m

```
$ tau app copy FUN3D_Miniapp1 miniapp1.MCDRAM
[TAU] Added application 'miniapp1.MCDRAM' to project configuration 'miniapp1'.

$ tau select miniapp1.MCDRAM sample
[TAU] Created a new experiment named 'grover-miniapp1.MCDRAM-sample'.
[TAU] Selected experiment 'grover-miniapp1.MCDRAM-sample'.

$ numactl -m 1 tau ./point_solve
```

Hardware Performance Counters

Execute these commands on a KNL node:

```
$ srun --pty -p pettt-qf $SHELL

$ tau target metrics knl
```

"knl" is the target configuration name. Yours may be different.

`tau target metrics`

Desktop — jlinford@abutill-0001:~ — ssh -F ~/.ssh/hpcmp_config -Y -K us.arl.hpc.mil — 157×51

```
bash-4.2$ tau target metrics knl
== PAPI Preset Metrics on Target 'knl' ==========================================================================================

      Name     |                                 Description
===============+==================================================================================================================
   PAPI_BR_CN  | Conditional branch instructions.
  PAPI_BR_INS  | Branch instructions.
  PAPI_BR_MSP  | Conditional branch instructions mispredicted.
  PAPI_BR_NTK  | Conditional branch instructions not taken.
  PAPI_BR_TKN  | Conditional branch instructions taken.
  PAPI_BR_UCN  | Unconditional branch instructions.
  PAPI_L1_DCA  | Level 1 data cache accesses.
  PAPI_L1_DCM  | Level 1 data cache misses.
  PAPI_L1_ICA  | Level 1 instruction cache accesses.
  PAPI_L1_ICH  | Level 1 instruction cache hits.
  PAPI_L1_ICM  | Level 1 instruction cache misses.
  PAPI_L1_LDM  | Level 1 load misses.
  PAPI_L1_TCM  | Level 1 cache misses.
  PAPI_L2_LDM  | Level 2 load misses.
  PAPI_L2_TCA  | Level 2 total cache accesses.
  PAPI_L2_TCH  | Level 2 total cache hits.
  PAPI_L2_TCM  | Level 2 cache misses.
  PAPI_LD_INS  | Load instructions.
 PAPI_LST_INS  | Load/store instructions completed.
 PAPI_REF_CYC  | Reference clock cycles.
 PAPI_RES_STL  | Cycles stalled on any resource.
  PAPI_SR_INS  | Store instructions.
 PAPI_STL_ICY  | Cycles with no instruction issue.
  PAPI_TLB_DM  | Data translation lookaside buffer misses.
 PAPI_TOT_CYC  | Total cycles.
 PAPI_TOT_INS  | Instructions completed.

== TAU Metrics on Target 'knl' ==================================================================================================

        Name           |                                 Description
=======================+==========================================================================================================
       CLOCK_GET_TIME  | Wall clock that calls clock_gettime.
             CPU_TIME  | CPU timer that calls getrusage.
      GET_TIME_OF_DAY  | Wall clock that calls gettimeofday.
         LINUX_TIMERS  | Linux high resolution wall clock.
        LOGICAL_CLOCK  | Logical clock that increments on each request.
         MEMORY_DELTA  | Instantaneous resident set size (RSS)
            PAPI_TIME  | Alias for P_WALL_CLOCK_TIME.  Wall clock that calls PAPI_get_real_usec.
    PAPI_VIRTUAL_TIME  | Alias for P_VIRTUAL_TIME.  PAPI virtual clock that calls PAPI_get_virt_usec.
       P_VIRTUAL_TIME  | PAPI virtual clock that calls PAPI_get_virt_usec.
    P_WALL_CLOCK_TIME  | Wall clock that calls PAPI_get_real_usec.
 TAU_MPI_MESSAGE_SIZE  | Running sum of all MPI messsage sizes.
                 TIME  | Alias for GET_TIME_OF_DAY. Wall clock that calls gettimeofday.
           USER_CLOCK  | User-defined clock. Implement 'void metric_write_userClock(int tid, double value)'  to set the clock value.

bash-4.2$
```

Measuring PAPI Counters

```
$ tau measurement copy sample sample.papi \
      --metrics TIME PAPI_L1_DCM PAPI_L2_TCM
```

Space-separated list of metrics

```
$ tau select sample.papi miniapp1.MCDRAM
[TAU] Created a new experiment 'knl-FUN3D_Miniapp1-sample.papi'
[TAU] Selected experiment 'knl-FUN3D_Miniapp1-sample.papi'.
[TAU] Application rebuild required:
[TAU]   - metrics changed from [TIME] to [TIME, PAPI_L1_DCM, PAPI_L2_TCM]
```

TAU Commander advises when application should be rebuilt.

PAPI Metric Compatibility Checks

jlinford — jlinford@grover:~/DataMovement/miniapp1/baseline — ssh grover.nic.uoregon.edu — 148×30

```
jlinford@grover ~/DataMovement/miniapp1/baseline $ tau meas copy sample sample.papi --metrics TIME PAPI_L1_DCM PAPI_L1_DCA PAPI_L2_TCM PAPI_L2_TCA
[TAU] Added measurement 'sample.papi' to project configuration 'miniapp1'.
jlinford@grover ~/DataMovement/miniapp1/baseline $ tau sel sample.papi
[TAU] Created a new experiment named 'grover-miniapp1-sample.papi'.
[TAU] XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
[TAU]
[TAU] CRITICAL
[TAU]
[TAU] PAPI metrics [PAPI_L1_DCM, PAPI_L1_DCA, PAPI_L2_TCM, PAPI_L2_TCA] are not compatible on this target.
[TAU]
[TAU] Hints:
[TAU]   * Use papi_avail to check metric availability.
[TAU]   * Spread the desired metrics over multiple measurements.
[TAU]   * Choose fewer metrics.
[TAU]
[TAU] TAU cannot proceed with the given inputs.
[TAU] Please check the selected configuration for errors or contact <support@paratools.com> for assistance.
[TAU]
[TAU] XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
jlinford@grover ~/DataMovement/miniapp1/baseline $
```

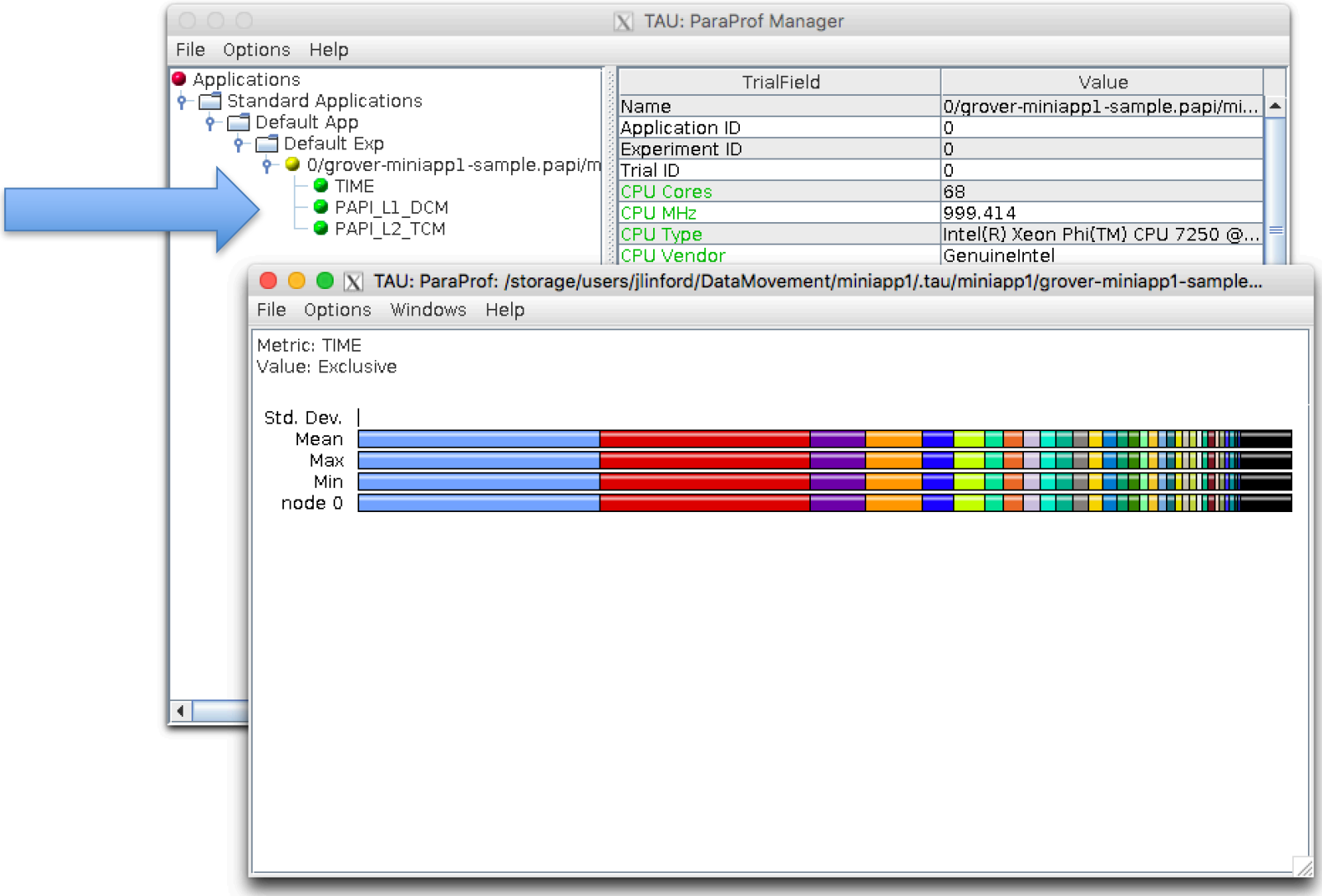
Internally uses papi_event_chooser to check metric compatibility.

Run exactly as before: `tau ./point_solve`

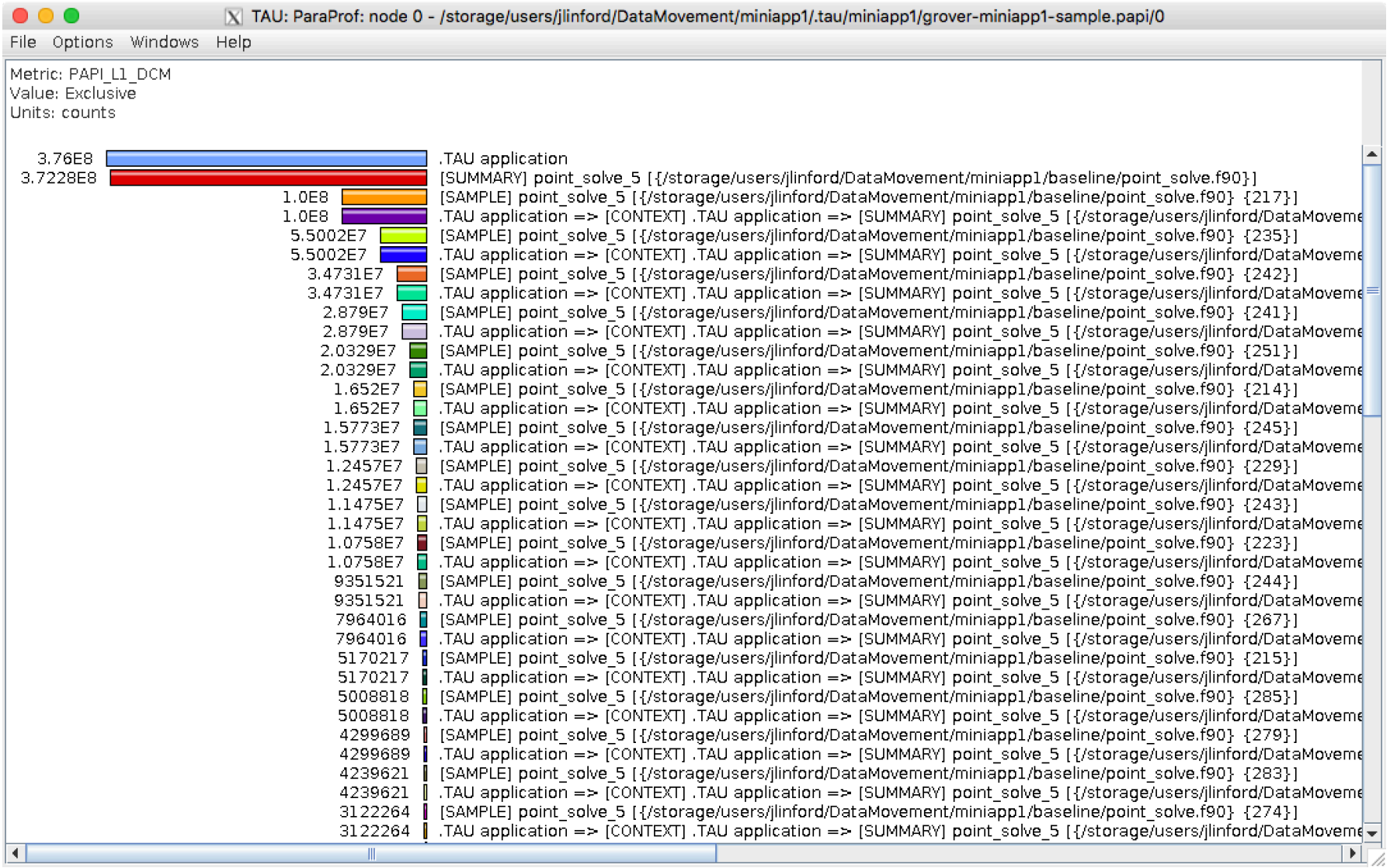
```
$ make run
tau ./point_solve
[TAU]
[TAU] == BEGIN Experiment at 2016-12-05 18:59:55.777171 ============================================================
[TAU]
[TAU] TAU_CALLPATH=1
[TAU] TAU_CALLPATH_DEPTH=100
[TAU] TAU_COMM_MATRIX=0
[TAU] TAU_METRICS=TIME,PAPI_L1_DCM,PAPI_L2_TCM
[TAU] TAU_PROFILE=1
[TAU] TAU_SAMPLING=1
[TAU] TAU_THROTTLE=1
[TAU] TAU_THROTTLE_NUMCALLS=100000
[TAU] TAU_THROTTLE_PERCALL=10
[TAU] TAU_TRACE=0
[TAU] TAU_TRACK_HEAP=0
[TAU] TAU_VERBOSE=0
[TAU] tau_exec -T serial,papi,icpc -ebs ./point_solve
 Loading data...
0 Number of block 5x5 equations in data file: 1123718
 Done loading data...
 Solving Ax=b...
 Sweep seconds on master =    1.369700
 Sweep seconds on master =    1.356500
 Sweep seconds on master =    1.358500
 Sweep seconds on master =    1.353100
 Sweep seconds on master =    1.347900
 Sweep seconds on master =    1.348200
 Sweep seconds on master =    1.348400
 Sweep seconds on master =    1.345000
 Sweep seconds on master =    1.350200
 Sweep seconds on master =    1.344700
 Sweep seconds on master =    1.344700
 Sweep seconds on master =    1.344000
 Sweep seconds on master =    1.343200
 Sweep seconds on master =    1.343400
 Sweep seconds on master =    1.343500
 Total seconds taken on master =    20.24260
 Test passed.
[TAU]
[TAU] == END Experiment at 2016-12-05 19:00:19.648510 ==============================================================
[TAU]
[TAU] Trial 0 produced 3 profile files.
```

Trial produced one profile for each metric.

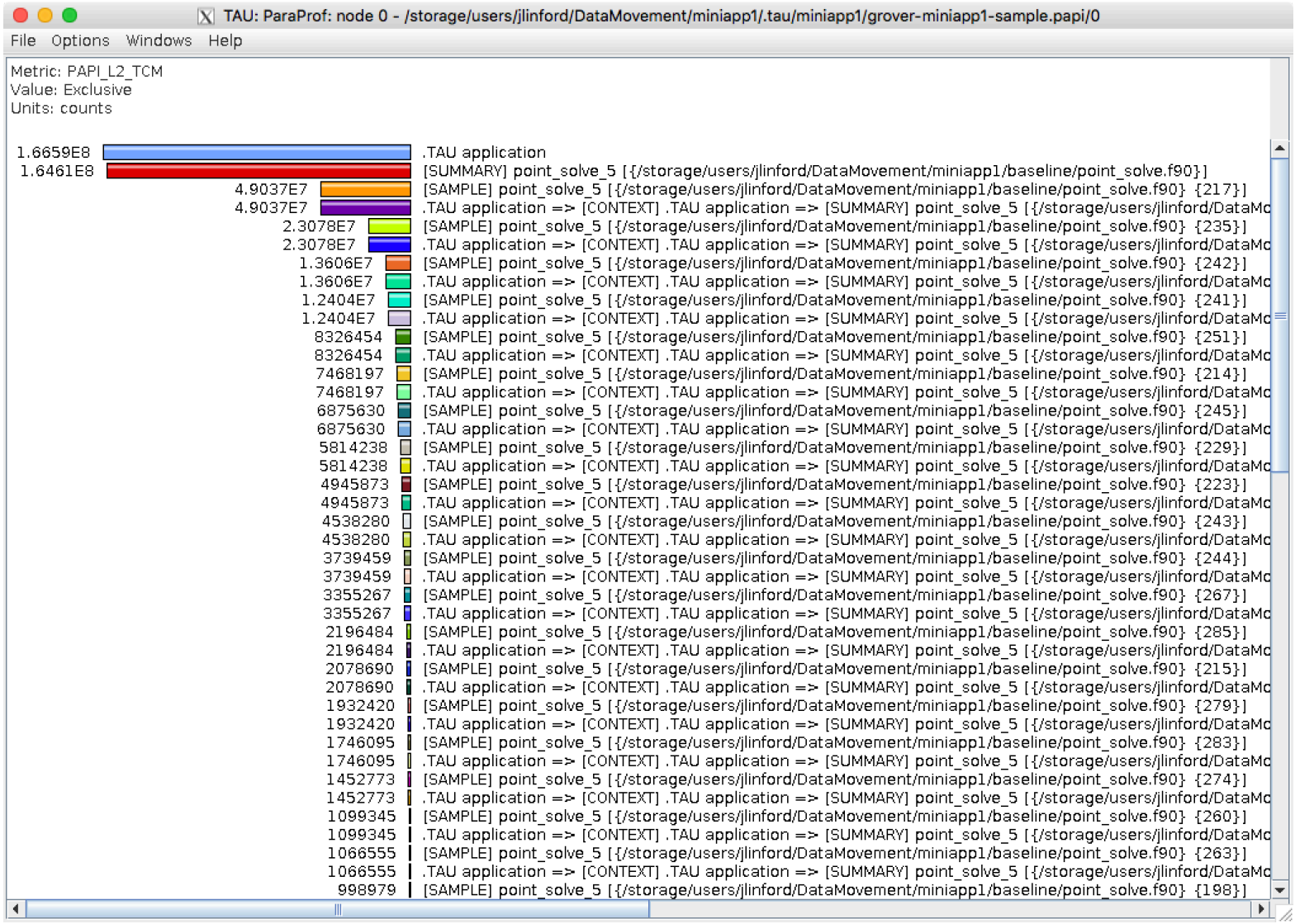
View profiles: `tau show`



L1 Data Cache Misses



L2 Total Cache Misses



Line with the most L1/L2 misses

TAU: ParaProf: Source Browser: /storage/users/jlinford/DataMovement/miniapp1/bas...

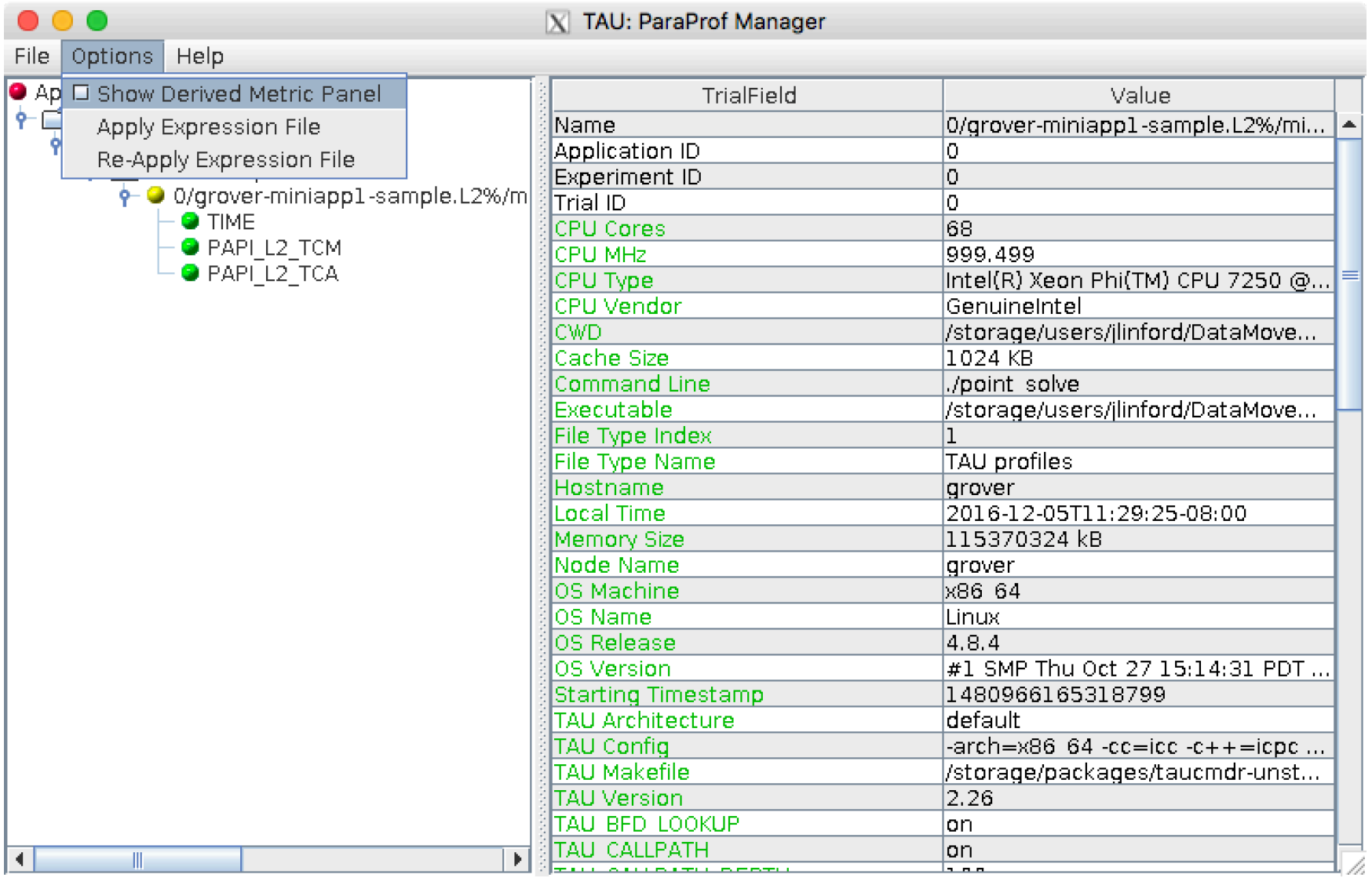
File Help

```
208             end if
209
210             istart = iam(n)
211             iend   = iam(n+1)-1
212
213             do j = istart,iend
214               icol = jam(j)
215
216               f1 = f1   - a_off(1,1,j)*dq(1,icol)
217               f2 = f2   - a_off(2,1,j)*dq(1,icol)
218               f3 = f3   - a_off(3,1,j)*dq(1,icol)
219               f4 = f4   - a_off(4,1,j)*dq(1,icol)
220               f5 = f5   - a_off(5,1,j)*dq(1,icol)
221
222               f1 = f1   - a_off(1,2,j)*dq(2,icol)
223               f2 = f2   - a_off(2,2,j)*dq(2,icol)
224               f3 = f3   - a_off(3,2,j)*dq(2,icol)
225               f4 = f4   - a_off(4,2,j)*dq(2,icol)
226               f5 = f5   - a_off(5,2,j)*dq(2,icol)
227
228               f1 = f1   - a_off(1,3,j)*dq(3,icol)
229               f2 = f2   - a_off(2,3,j)*dq(3,icol)
230               f3 = f3   - a_off(3,3,j)*dq(3,icol)
231               f4 = f4   - a_off(4,3,j)*dq(3,icol)
232               f5 = f5   - a_off(5,3,j)*dq(3,icol)
233
234               f1 = f1   - a_off(1,4,j)*dq(4,icol)
235               f2 = f2   - a_off(2,4,j)*dq(4,icol)
236               f3 = f3   - a_off(3,4,j)*dq(4,icol)
237               f4 = f4   - a_off(4,4,j)*dq(4,icol)
238               f5 = f5   - a_off(5,4,j)*dq(4,icol)
239
240               f1 = f1   - a_off(1,5,j)*dq(5,icol)
241               f2 = f2   - a_off(2,5,j)*dq(5,icol)
242               f3 = f3   - a_off(3,5,j)*dq(5,icol)
243               f4 = f4   - a_off(4,5,j)*dq(5,icol)
244               f5 = f5   - a_off(5,5,j)*dq(5,icol)
245
246             end do
247
248   ! Forward...sequential access to a diag lu.
249
```

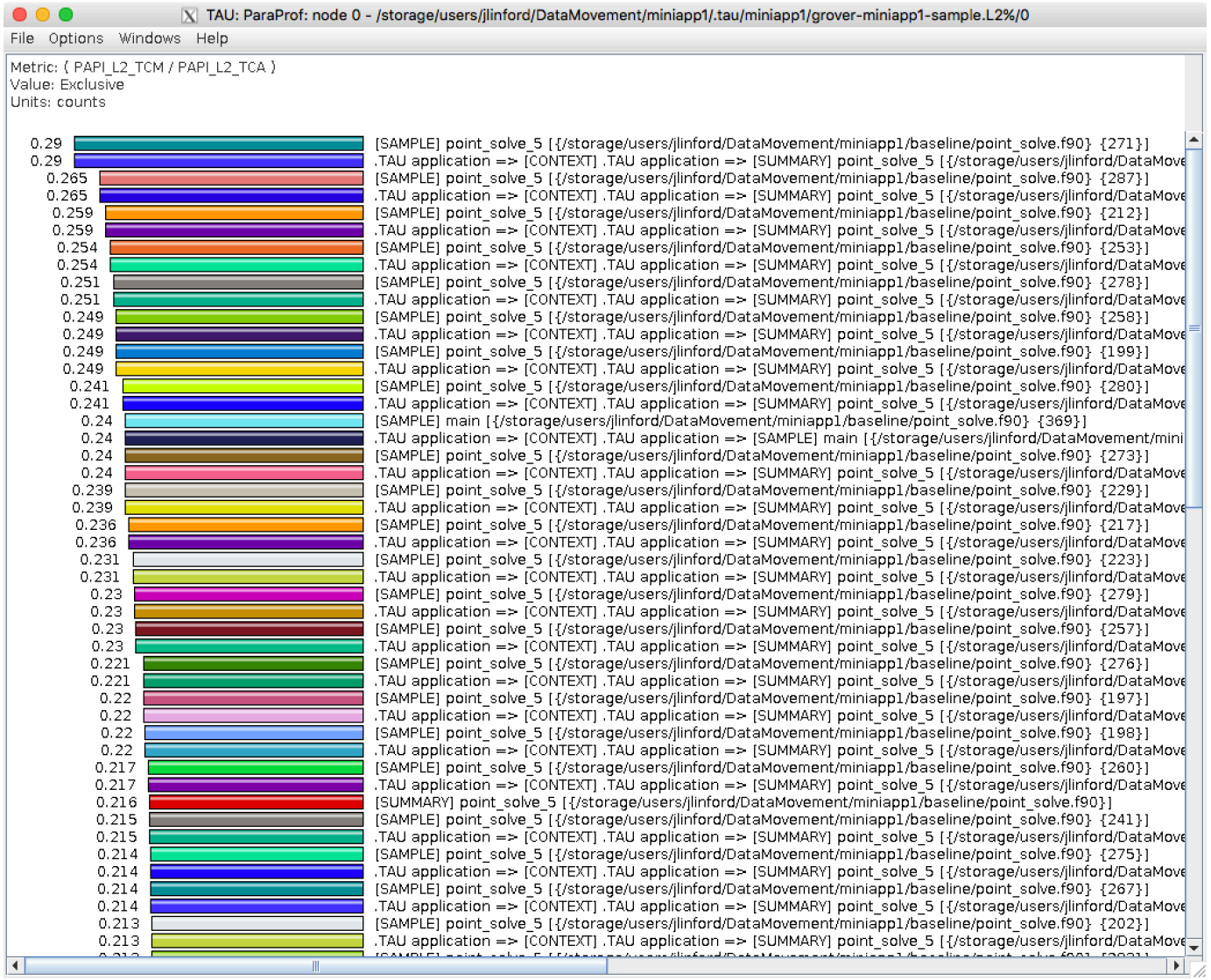
What percent of L2 accesses are misses?

```
$ tau meas copy sample.papi "sample.L2%" --metrics TIME PAPI_L2_TCM PAPI_L2_TCA
[TAU] Added measurement 'sample.L2%' to project configuration 'miniapp1'.

$ tau sel sample.L2%
[TAU] Created a new experiment named 'grover-miniapp1-sample.L2%'.
[TAU] Selected experiment 'grover-miniapp1-sample.L2%'.

$ make run
tau ./point_solve
[TAU]
[TAU] == BEGIN Experiment at 2016-12-05 19:29:24.677341 ===============
[TAU]
[TAU] TAU_CALLPATH=1
[TAU] TAU_CALLPATH_DEPTH=100
[TAU] TAU_COMM_MATRIX=0
[TAU] TAU_METRICS=TIME,PAPI_L2_TCM,PAPI_L2_TCA
```

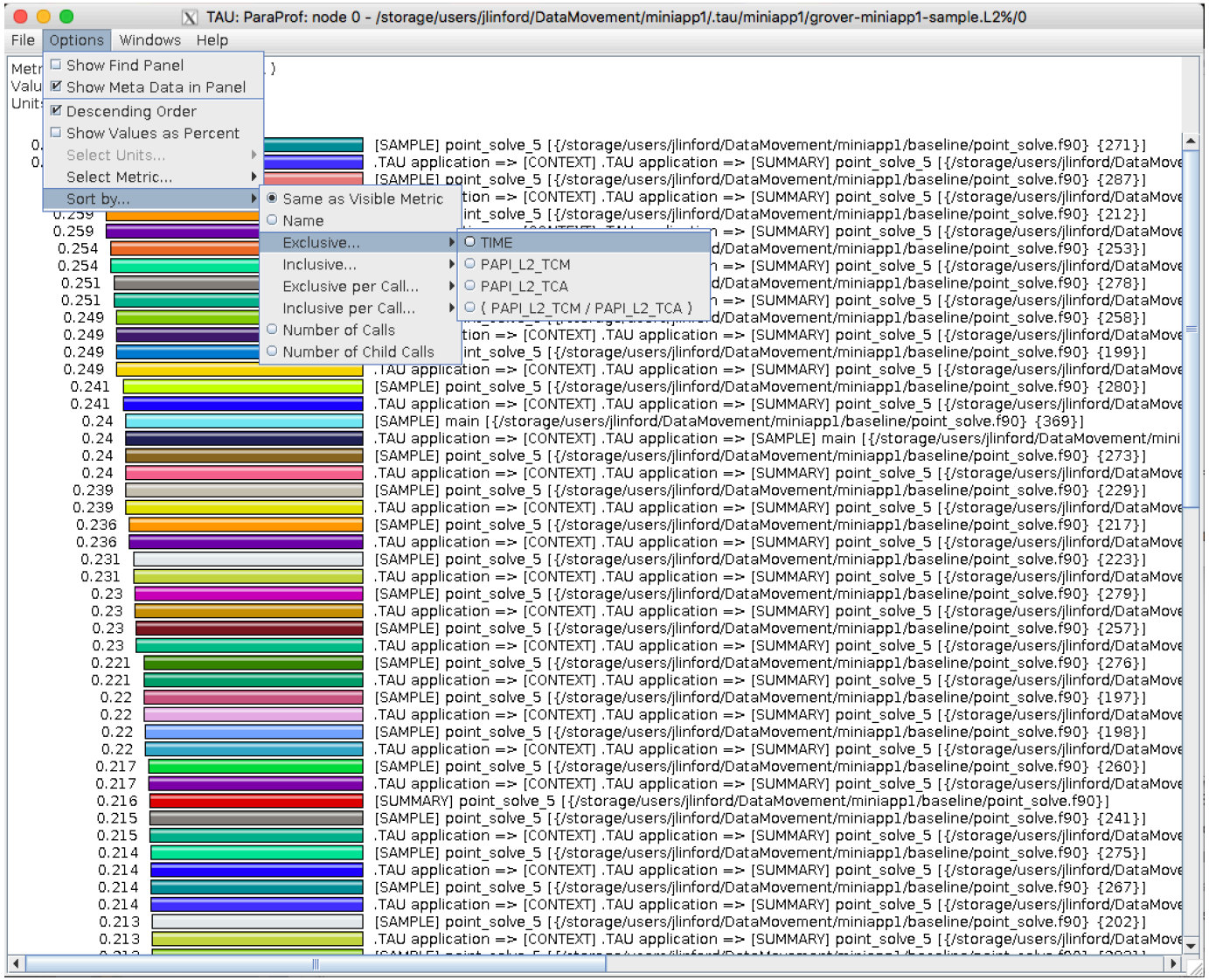
Create a new derived metric



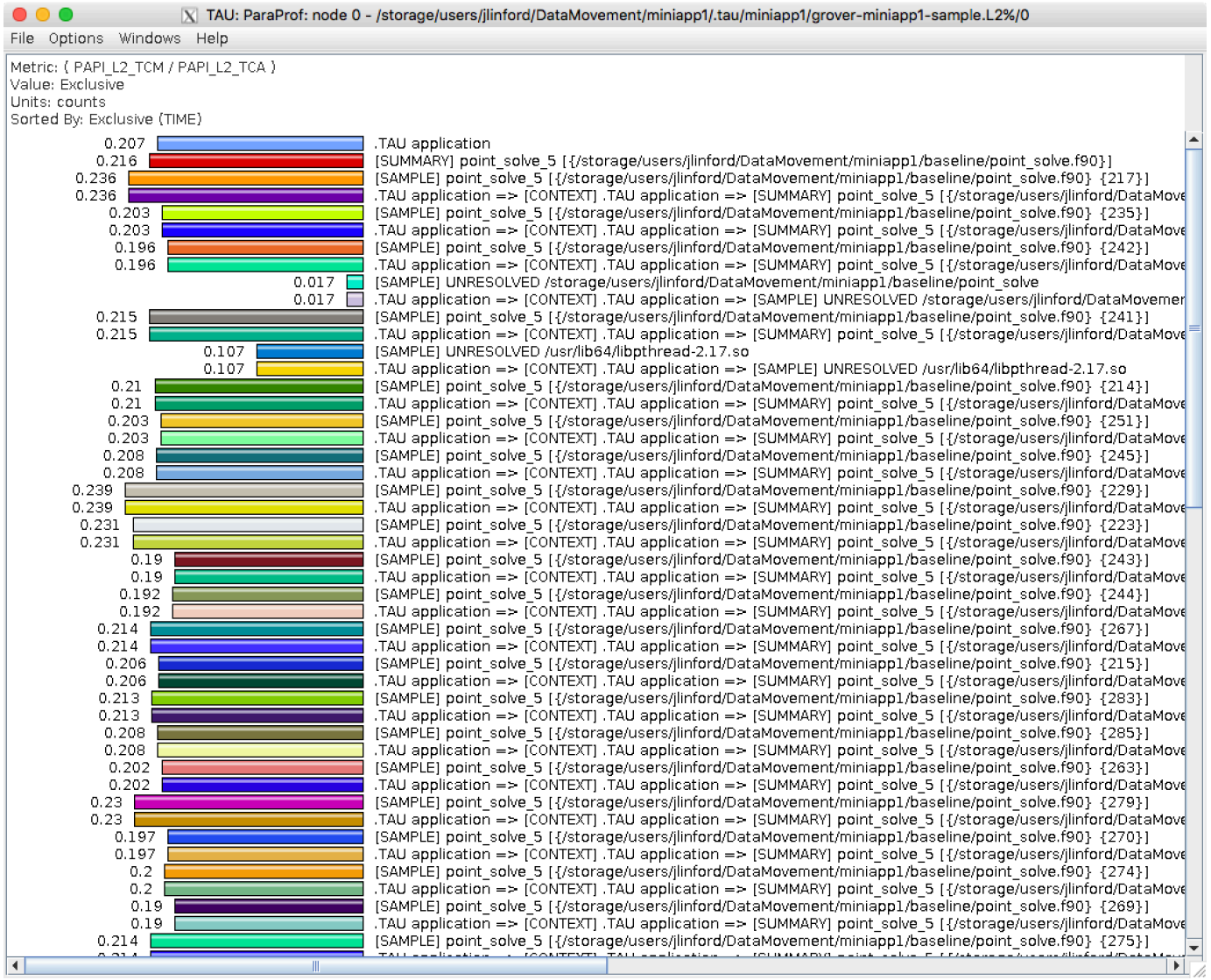
At worst, 29% of L2 fetches miss



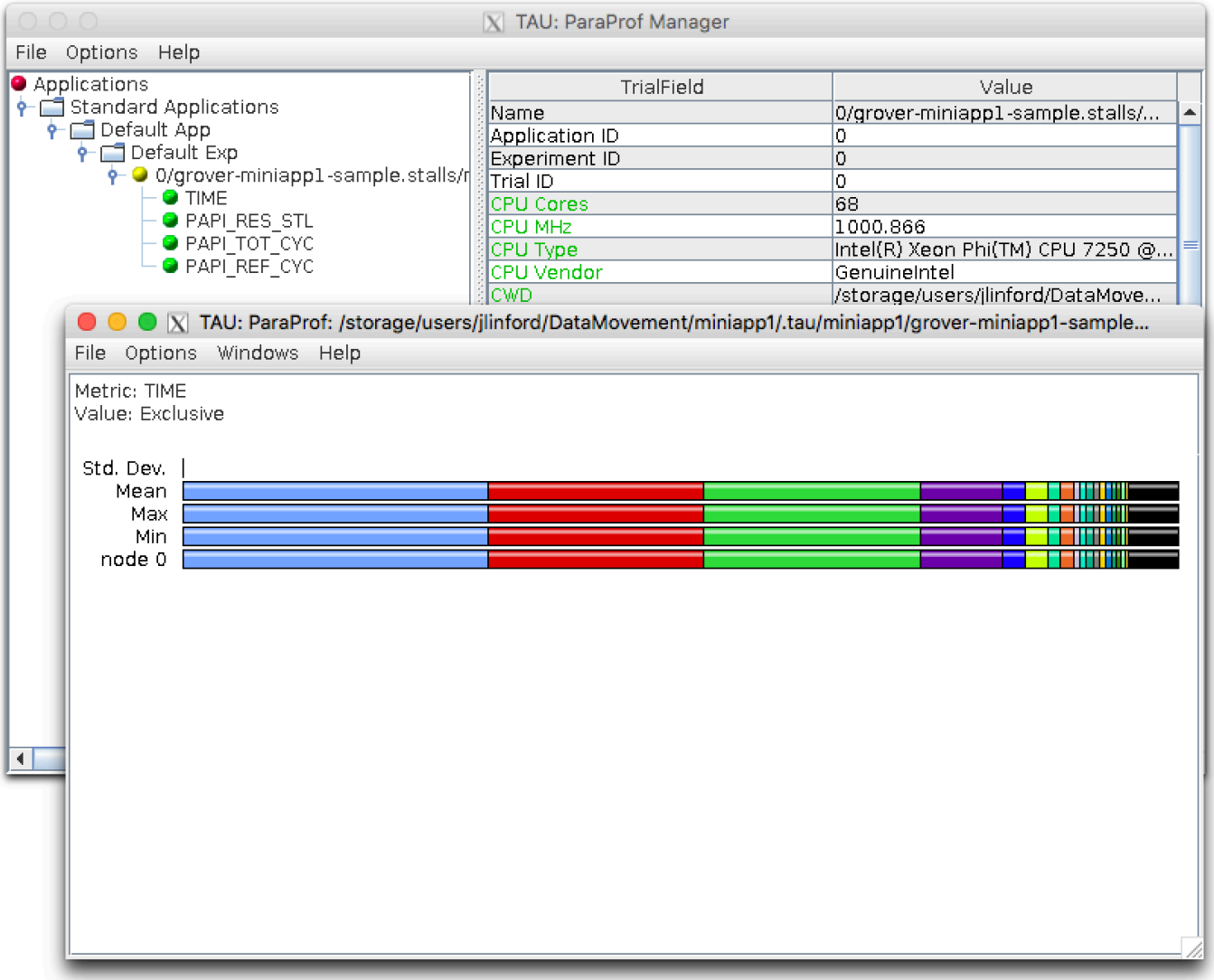
Sort by exclusive time



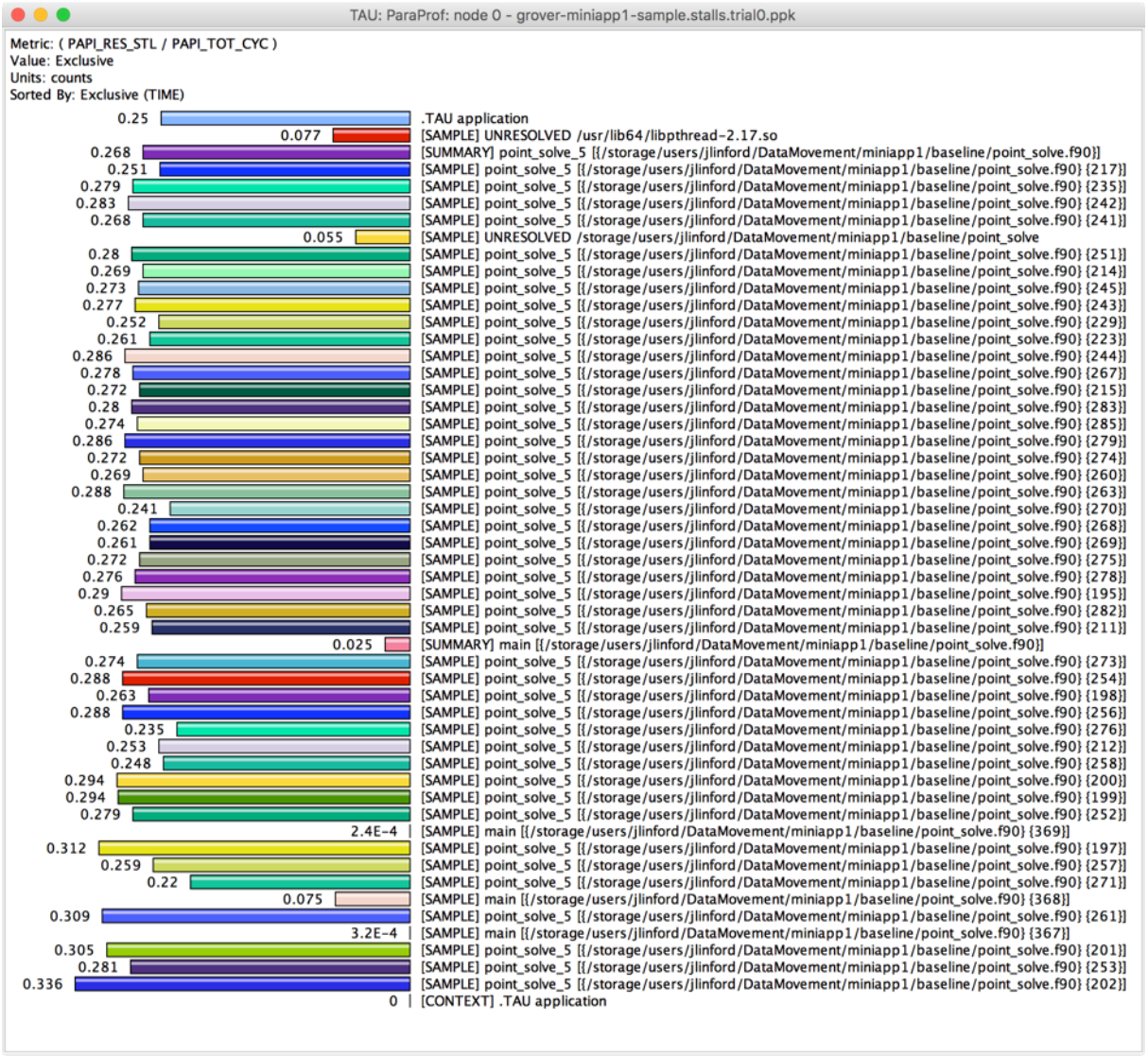
About 21% L2 fetches miss in kernel



% Cycles Stalled Waiting for Memory



% Cycles Stalled Waiting for Memory



OpenMP Parallelization

```
$ cd ~/FUN3D_Miniapp1/openmp

!$omp parallel default(shared)
do sweep = 1, n_sweeps
  do color = sweep_start, sweep_end, sweep_stride
    do ipass = 1, 2
      start = color_indices(1,color)
      end = color_boundary_end(color)

      !$omp do private(f1,f2,f3,f4,f5,n,j,icol,istart,iend) schedule(auto)
      do n = start, end
        istart = iam(n)
        iend   = iam(n+1)-1

        f(1:5) = (+/-)res(1:5)

        do j = istart, iend
          icol = jam(j)
          do i = 1, 5
            f(1:5) = f(1:5) - a_off(1:5,i,j)*dq(i,icol)
          end do
```

Create a new OpenMP Application Config

```
$ cd ~/FUN3D_Miniapp1/openmp

# Edit Makefile as before

$ tau app copy FUN3D_Miniapp1 miniapp1.openmp --openmp
[TAU] Added application 'miniapp1.openmp' to project configuration 'miniapp1'.

$ tau select miniapp1.openmp sample
[TAU] Selected experiment 'grover-miniapp1.openmp-sample'.
[TAU] Application rebuild required:
[TAU]   - openmp changed from False to True
```

Compile and run exactly as before

```
Desktop — jlinford@abutill-0001:~ — ssh -F ~/.ssh/hpcmp_config -Y -K us.arl.hpc.mil — 157×45
[-bash-4.2$ srun --pty -p pettt-qf $SHELL
[bash-4.2$ make run
tau ./point_solve
[TAU]
[TAU] == BEGIN Experiment at 2017-04-24 23:10:04.578666 ==========================================================================================================
[TAU]
[TAU] TAU_CALLPATH=1
[TAU] TAU_CALLPATH_DEPTH=100
[TAU] TAU_COMM_MATRIX=0
[TAU] TAU_METRICS=TIME,
[TAU] TAU_PROFILE=1
[TAU] TAU_SAMPLING=1
[TAU] TAU_THROTTLE=1
[TAU] TAU_THROTTLE_NUMCALLS=100000
[TAU] TAU_THROTTLE_PERCALL=10
[TAU] TAU_TRACE=0
[TAU] TAU_TRACK_HEAP=0
[TAU] TAU_VERBOSE=0
[TAU] tau_exec -T serial,pthread,icpc -ebs ./point_solve
 Loading data...
0 Number of block 5x5 equations in data file: 1123718
 Done loading data...
 Solving Ax=b...
 Sweep msec on master =   0.263401031494141
 Sweep msec on master =   3.889513015747070E-002
 Sweep msec on master =   3.510093688964844E-002
 Sweep msec on master =   3.503799438476562E-002
 Sweep msec on master =   6.075096130371094E-002
 Sweep msec on master =   7.602715492248535E-002
 Sweep msec on master =   7.072806358337402E-002
 Sweep msec on master =   7.519698143005371E-002
 Sweep msec on master =   7.883405685424805E-002
 Sweep msec on master =   6.817889213562012E-002
 Sweep msec on master =   6.178212165832520E-002
 Sweep msec on master =   5.966305732727051E-002
 Sweep msec on master =   7.982397079467773E-002
 Sweep msec on master =   7.002496719360352E-002
 Sweep msec on master =   7.700800895690918E-002
 Total msec taken on master =    1.15544199943542
 Test passed.
[TAU]
[TAU] == END Experiment at 2017-04-24 23:10:14.065383 ============================================================================================================
[TAU]
[TAU] Trial 0 produced 256 profile files.
bash-4.2$
```

14.2x faster with OpenMP

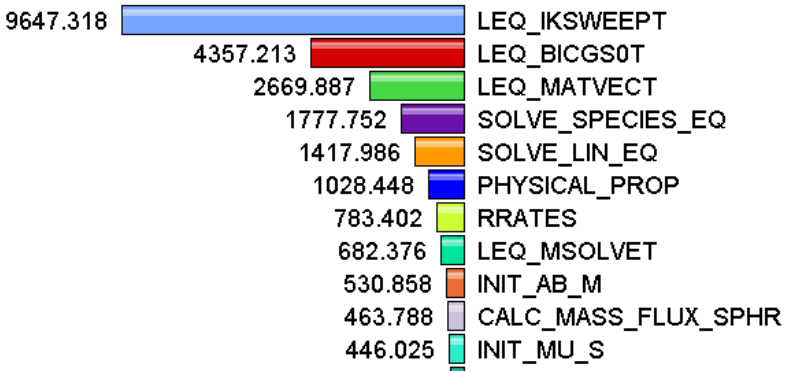
One profile per thread

KNL Hands On Exercises

REFERENCE

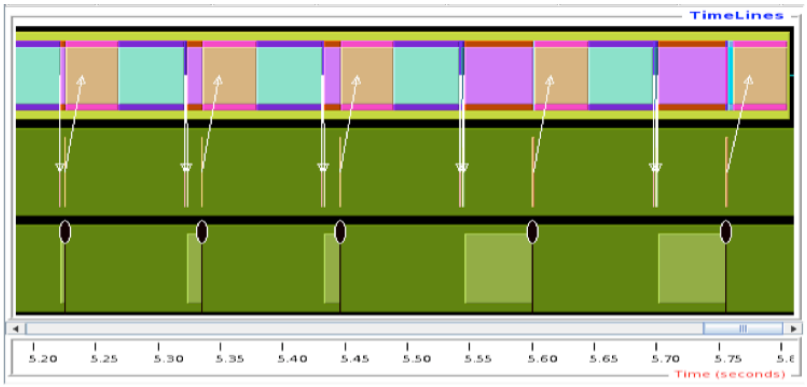
Measurement Approaches

Profiling



Shows **how much** time was spent in each routine

Tracing



Shows **when** events take place on a timeline

Types of Performance Profiles

- *Flat* profiles
  - Metric (e.g., time) spent in an event
  - Exclusive/inclusive, # of calls, child calls, …
- *Callpath* profiles
  - Time spent along a calling path (edges in callgraph)
  - "*main=> f1 => f2 => MPI_Send*"
- *Phase* profiles
  - Flat profiles under a phase (nested phases allowed)
  - Default "main" phase
  - Supports static or dynamic (e.g. per-iteration) phases

Direct Observation Events

- Interval events (begin/end events)
  - Measures exclusive & inclusive durations between events
  - Metrics monotonically increase
  - Example: Wall-clock timer
- Atomic events (trigger with data value)
  - Used to capture performance data state
  - Shows extent of variation of triggered values (min/max/mean)
  - Example: heap memory consumed at a particular point

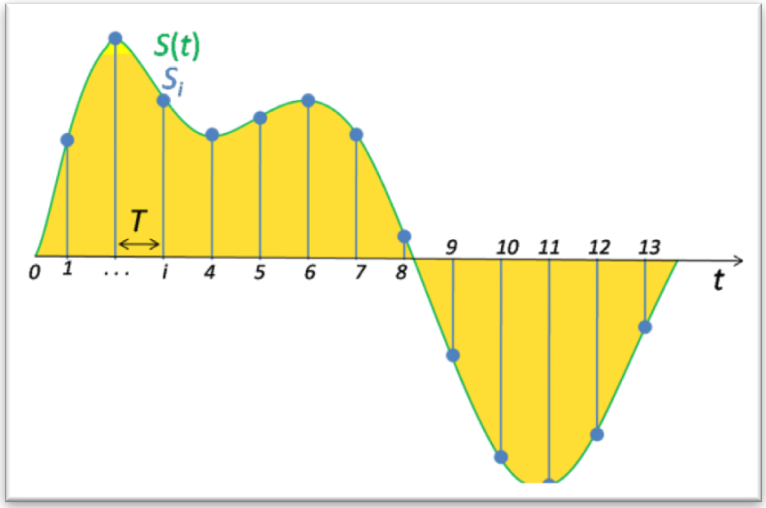
Direct vs. Indirect Measurement

Direct via Probes

```
call TAU_START('potential')
// code
call TAU_STOP('potential')
```

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

Indirect via Sampling



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

Inclusive vs. Exclusive Measurements

- Exclusive measurements for region only
- Inclusive measurements includes child regions