



# Isolating Runtime Faults with Callstack Debugging using TAU

**John C. Linford<sup>1</sup>, Sameer Shende<sup>1</sup>, Allen D. Malony<sup>1</sup>,  
Andrew Wissink<sup>2</sup>, Stephen Adamec<sup>3</sup>**

<sup>1</sup>ParaTools, Inc.  
Eugene, OR, USA, 21093

<sup>2</sup>Ames Research Center  
Moffett Field, CA, USA, 90435

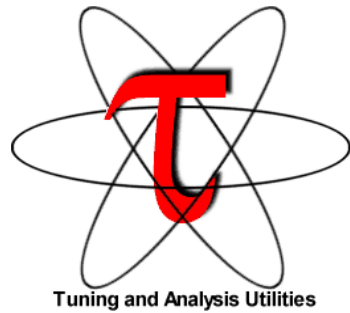
<sup>3</sup>University of Alabama at Birmingham  
Birmingham, AL, USA, 35209



# Outline

- **Brief overview of TAU**
- **Multi-language callstack debugging with TAU**
- **Design and implementation**
- **Examples: CREATE-AV Helios and Kestrel**
- **Summary and conclusions**

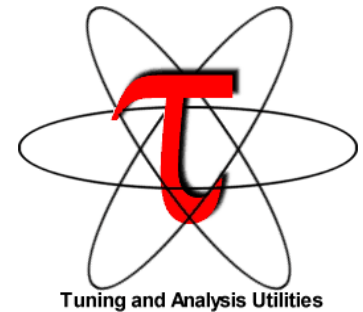
# Brief overview of TAU



# TAU is a performance evaluation tool

- **TAU supports parallel profiling and tracing**
  - Profiling: *how much time* was spent in each routine
  - Tracing: *when* the events take place in each process
- **TAU can measure hardware performance counters**
- **TAU can automatically instrument your source code**
  - Routines, loops, I/O, memory, phases, etc.
- **TAU runs on all HPC platforms and it is free**
- **TAU has instrumentation, measurement and analysis tools**
  - ParaProf, PerfExplorer, Jumpshot, etc.
- **TAU has performance database technology (TAUdb)**

# For more information



**TAU Website:** <http://tau.uoregon.edu/>

- Software download
- Release notes
- Documentation

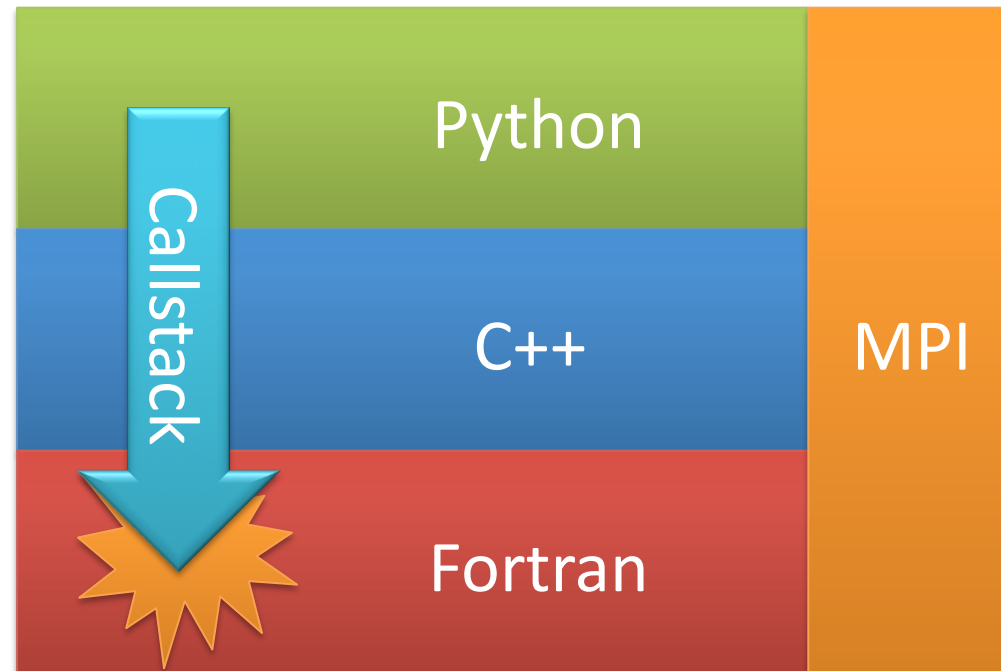
**TAU LiveDVD:** <http://www.hpclinux.com/>

- Boot up on your laptop or desktop
- Includes TAU and variety of other packages
- Include documentation and tutorial slides



# **Multi-language callstack debugging with TAU**

# Segfault! What do you do?





# Debugging challenges

- **Execution text output rarely sufficient**
- **Core files aren't much help for 10k processes**
- **What if the fault occurred in a DSO?**
- **Most debuggers are monolingual**
- **Developers need to reproduce the crash to fix it, but program inputs are sensitive or proprietary**





## TAU has already solved most of these

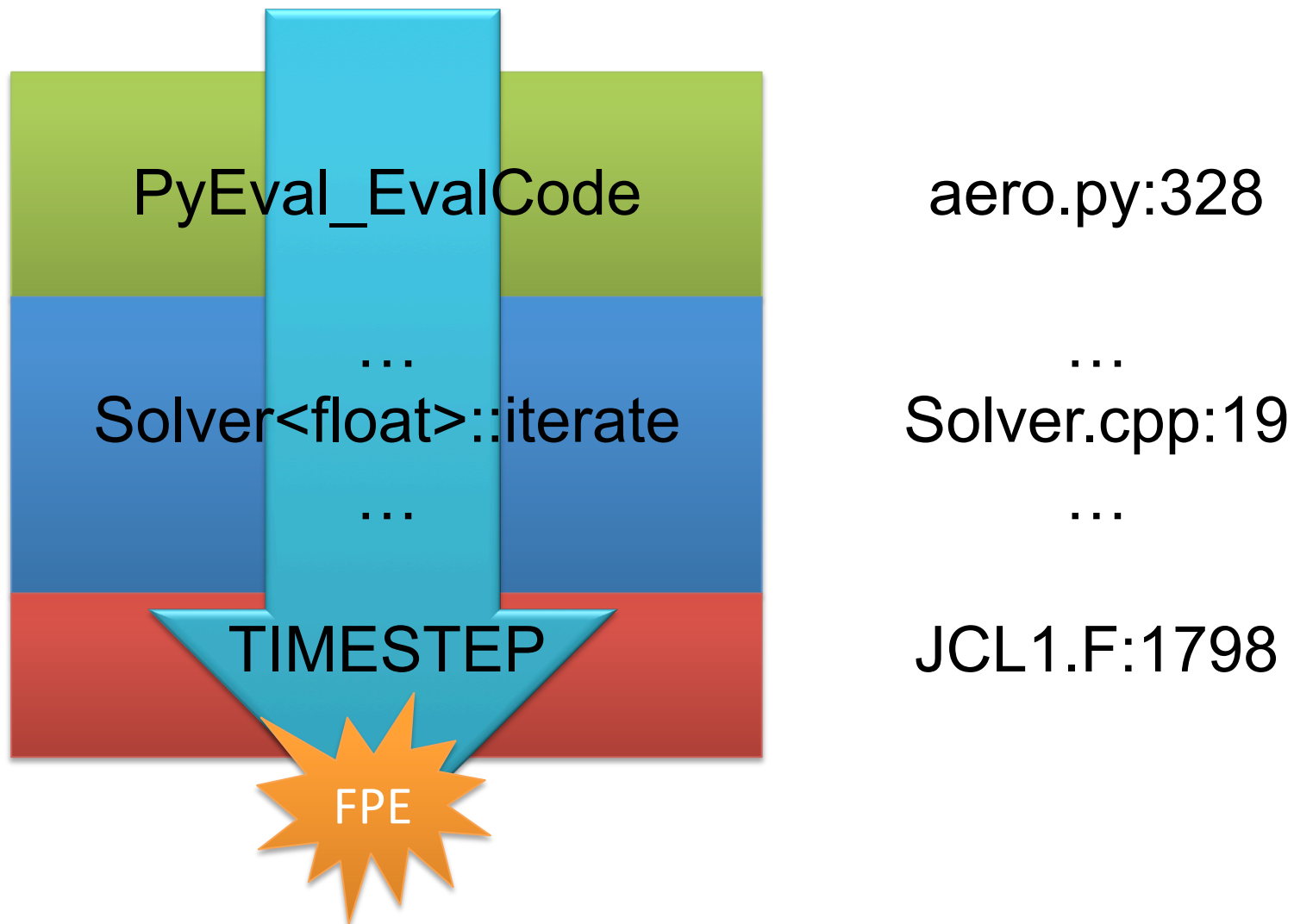
- Rich performance information
- Scales to 100k processes and beyond
- Highly efficient packed profile format
- Maintains and updates address maps for DSOs
- C, C++, CUDA, Fortran, UPC, Python, Java, etc.
- With *tau\_exec*, recompile **not required**



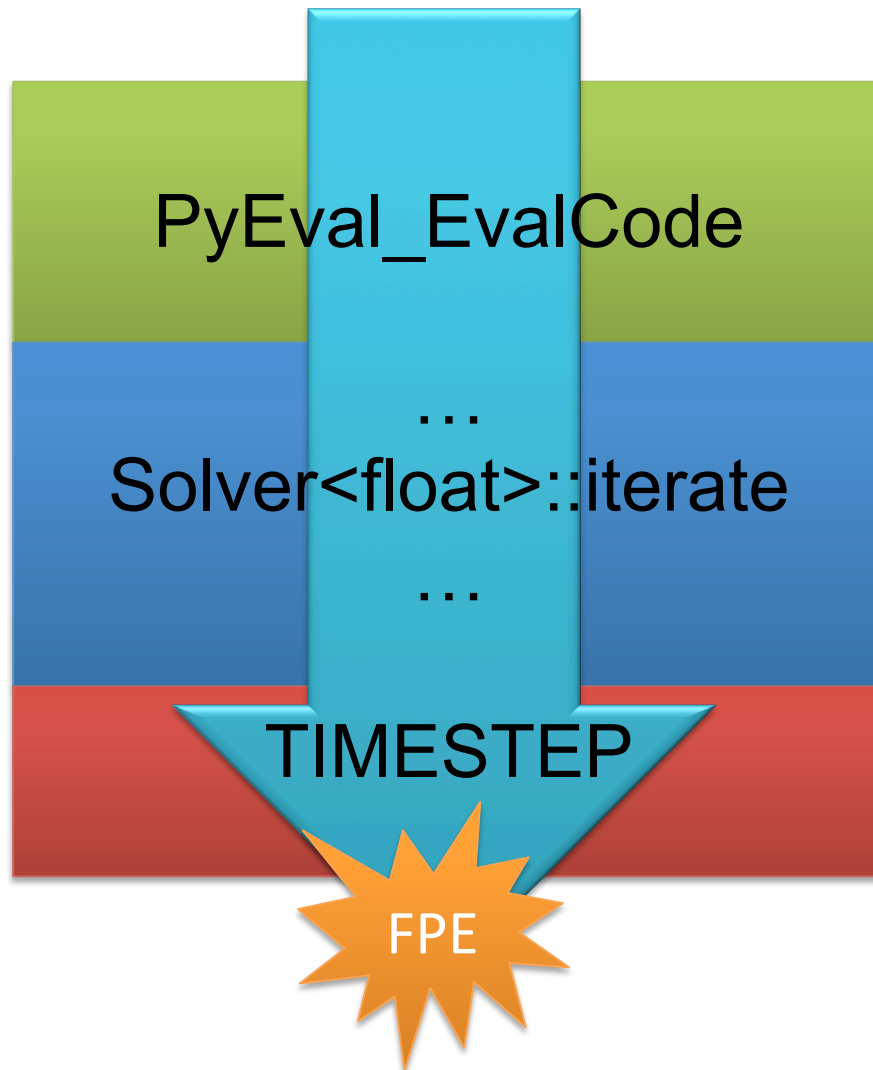
# TAU callstack debugging has two goals

- **Assist in debugging multi-language applications**
  - Unwind the callstack across C/C++, Fortran, Python, Java, UPC, etc.
  - Explore application performance at all levels
- **Close the loop with developers for more rapid turnaround of bug fixes**
  - Compact, portable, informative reports
  - Independent of sensitive or proprietary inputs

# Callstack unwinding is the key to debugging



# Performance information is retained



Started on rank 13

...

MPI\_Recv read 451k

...

Allocated 3298k heap

954s runtime



# The debugger can answer these questions

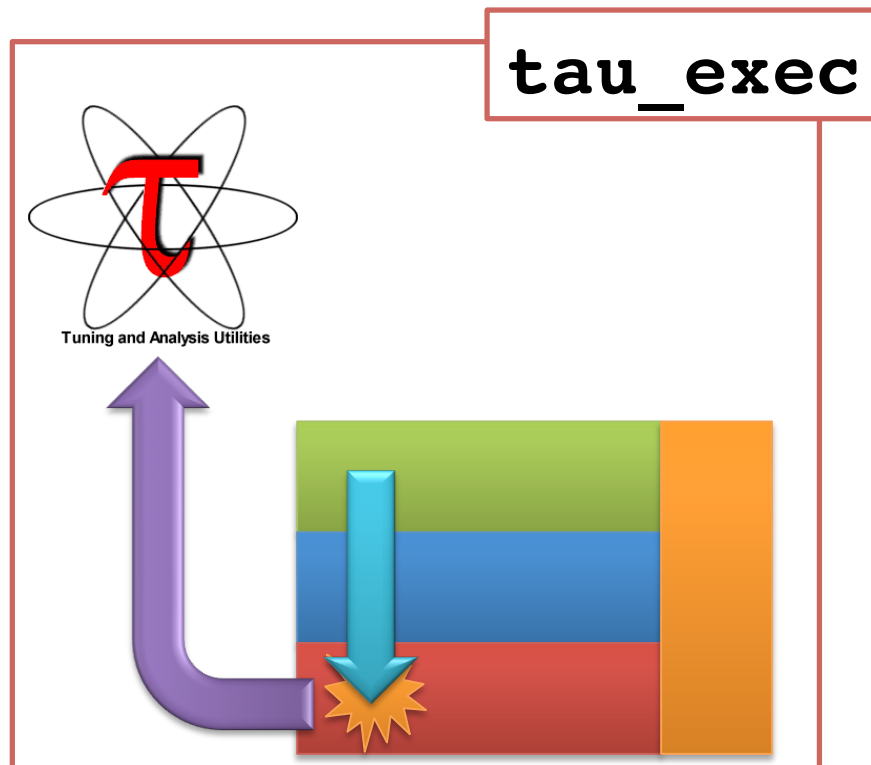
- 1. Where and when did the program fail?**
- 2. What was the nature of the fault?**
- 3. What was the application's heap memory utilization?**
- 4. Where there any memory leaks in the application?**
- 5. What were the application's performance characteristics?**
- 6. How much time did the application spend in I/O and communication operations?**



# **Design and Implementation**

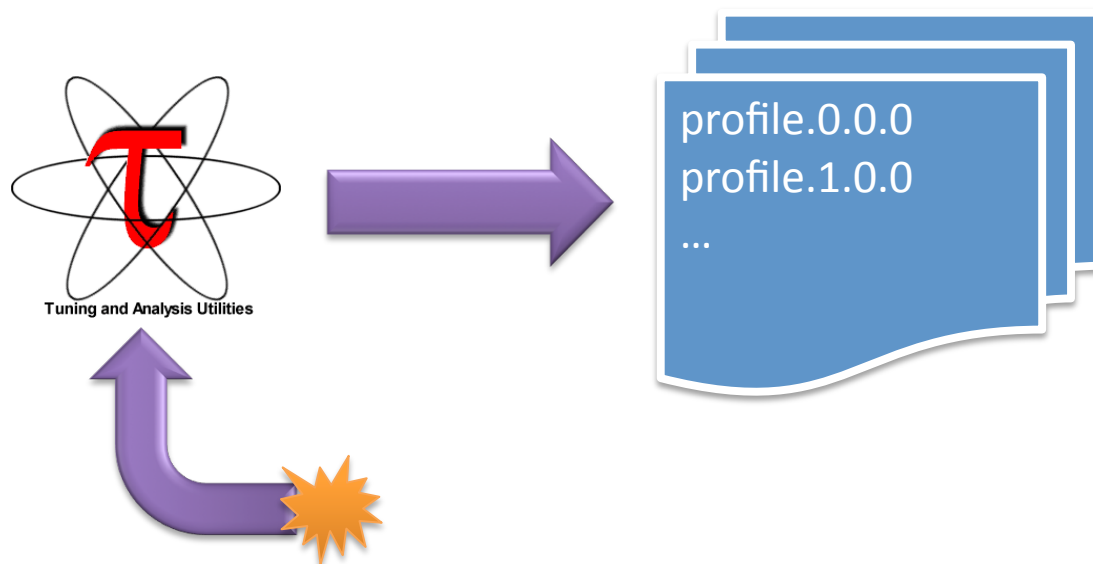
# Signal handler intercepts the fault signal

- `export TAU_TRACK_SIGNALS=1`
- `tau_exec` registers a signal handler
- Error signal triggers callstack unwind



# TAU unwinds the callstack of each thread

- **GLIBC backtrace API and GNU binutils determine routine name, file name, source line number**
- **A profile file is created for each thread**








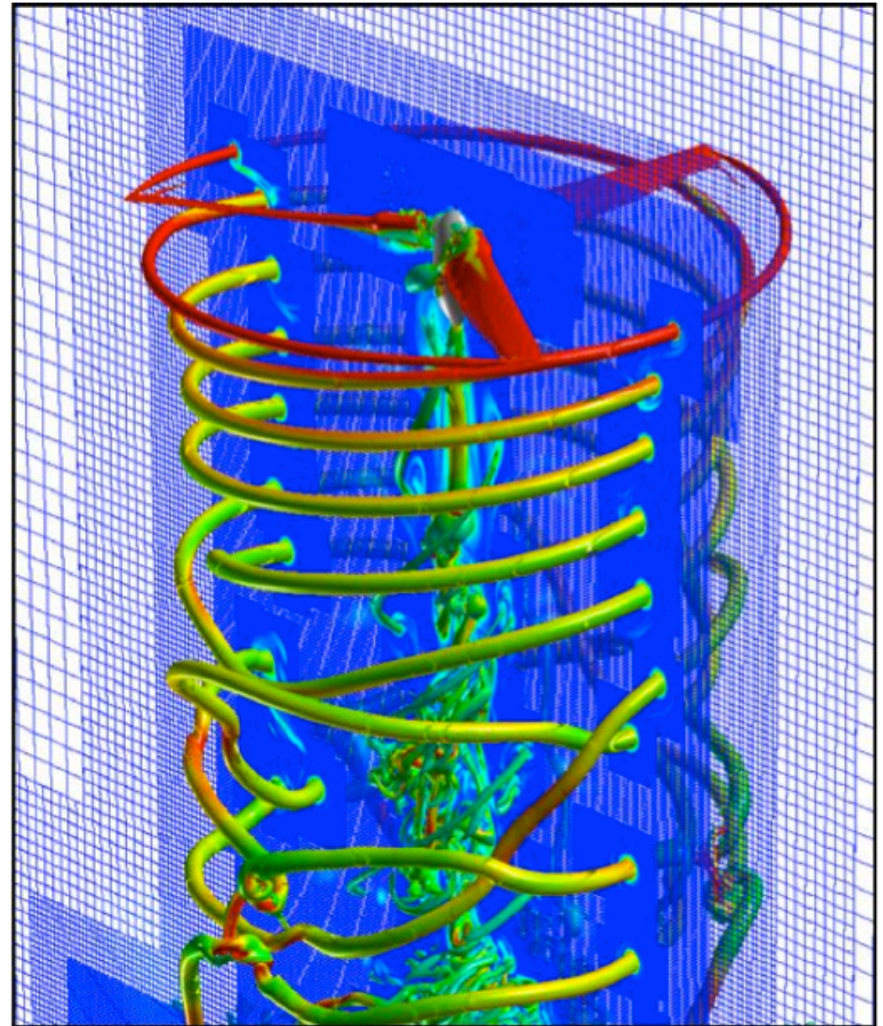
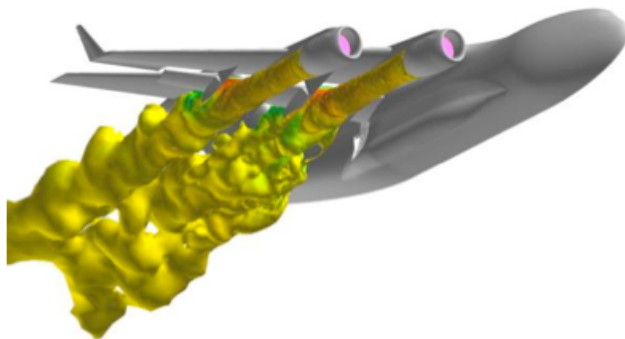
## TAU orchestrates a graceful shutdown

- **Profile files containing diagnostic information are created for all threads, healthy or otherwise**
- **Healthy threads exit gracefully**

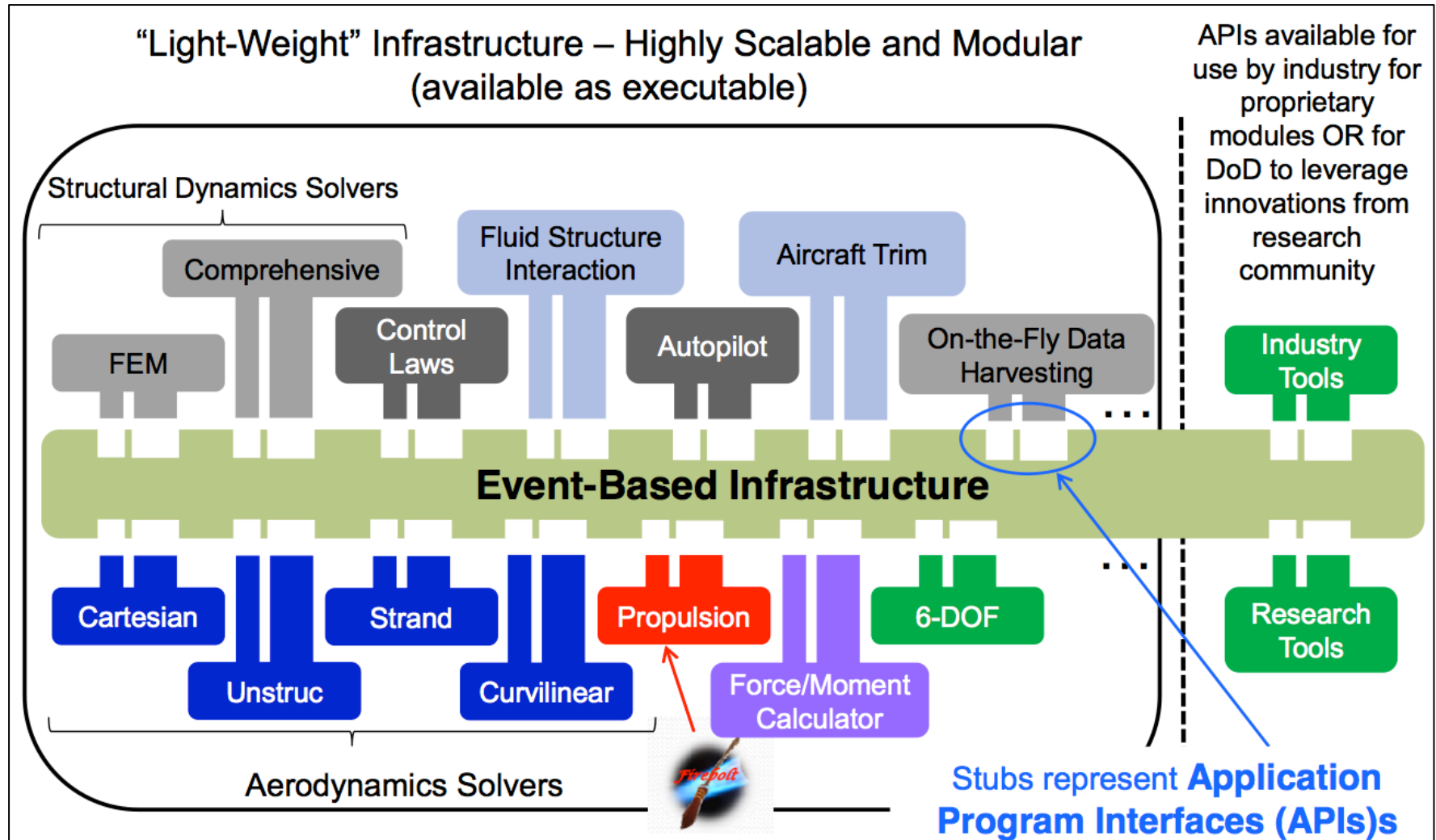


**Examples:**  
**CREATE-AV Helios and Kestrel**

# CREATE-AV Helios and Kestrel



# Helios and Kestrel software architecture



# Build with debugging symbols (-g) for a more informative backtrace

```
jlinford — jlinford@m0607:~/py-c++-f90-create — ssh — 100x24
[jlinford@m0607 py-c++-f90-create]$ make
mpicxx -fPIC -fpermissive -g -c -o SAMINT.o SAMINT.C
SAMINT.C: In static member function 'static void SAMINT::timestep(double, double)':
SAMINT.C:77: warning: division by zero in '4 / 0'
ifort -g -c fortmthds.f
mpicxx -fPIC -fpermissive -g -c pycintfc.C -I/mnt/cfs/pkgs/PTOOLS/pkgs/ptoolsrte-0.55/linux-redhat5.6-gnu-x86_64/Python-2.7.2/include/python2.7 -I/mnt/cfs/pkgs/PTOOLS/pkgs/ptoolsrte-0.55/linux-redhat5.6-gnu-x86_64/numpy-1.6.1/lib/python2.7/site-packages/numpy/core/include
cd swig; make all
make[1]: Entering directory `/mnt/home/jlinford/py-c++-f90-create/swig'
swig -python -I../ -I/mnt/cfs/pkgs/PTOOLS/pkgs/ptoolsrte-0.55/packages/swig-2.0.4/share/swig/2.0.4 -I/mnt/cfs/pkgs/PTOOLS/pkgs/ptoolsrte-0.55/packages/swig-2.0.4/share/swig/2.0.4/python samint.i
../pyGlobals.h:14: Warning 454: Setting a pointer/reference variable may leak memory.
../pyGlobals.h:15: Warning 454: Setting a pointer/reference variable may leak memory.
/bin/cp samint.py samint_wrap.c ../
make[1]: Leaving directory `/mnt/home/jlinford/py-c++-f90-create/swig'
mpicxx -fPIC -fpermissive -g -c samint_wrap.c -I/mnt/cfs/pkgs/PTOOLS/pkgs/ptoolsrte-0.55/linux-redhat5.6-gnu-x86_64/Python-2.7.2/include/python2.7 -I/mnt/cfs/pkgs/PTOOLS/pkgs/ptoolsrte-0.55/linux-redhat5.6-gnu-x86_64/numpy-1.6.1/lib/python2.7/site-packages/numpy/core/include
mpicxx -fPIC -fpermissive -g -c linkcheck.C
mpicxx -fPIC -shared pycintfc.o samint_wrap.o SAMINT.o fortmthds.o \
-o _samint.so
[jlinford@m0607 py-c++-f90-create]$
```

# Create a wrapper file to see Python events

wrapper.py

```
import tau

def OurMain():
    import samarcrun

tau.run('OurMain()')
```

# Set environment variables and run with *tau\_exec*

## Normal Execution

---

```
mpirun -np {n} pyMPI ./samarcrun.py
```

## Debugging with TAU

---

```
export TAU_TRACK_SIGNALS=1
```

```
export TAU_CALLPATH_DEPTH=100
```

```
mpirun -np {n} tau_exec -T python pyMPI wrapper.py
```

## Debugging with TAU + I/O and Memory Tracking

---

```
mpirun -np {n} tau_exec -T python pyMPI -io -memory \  
wrapper.py
```

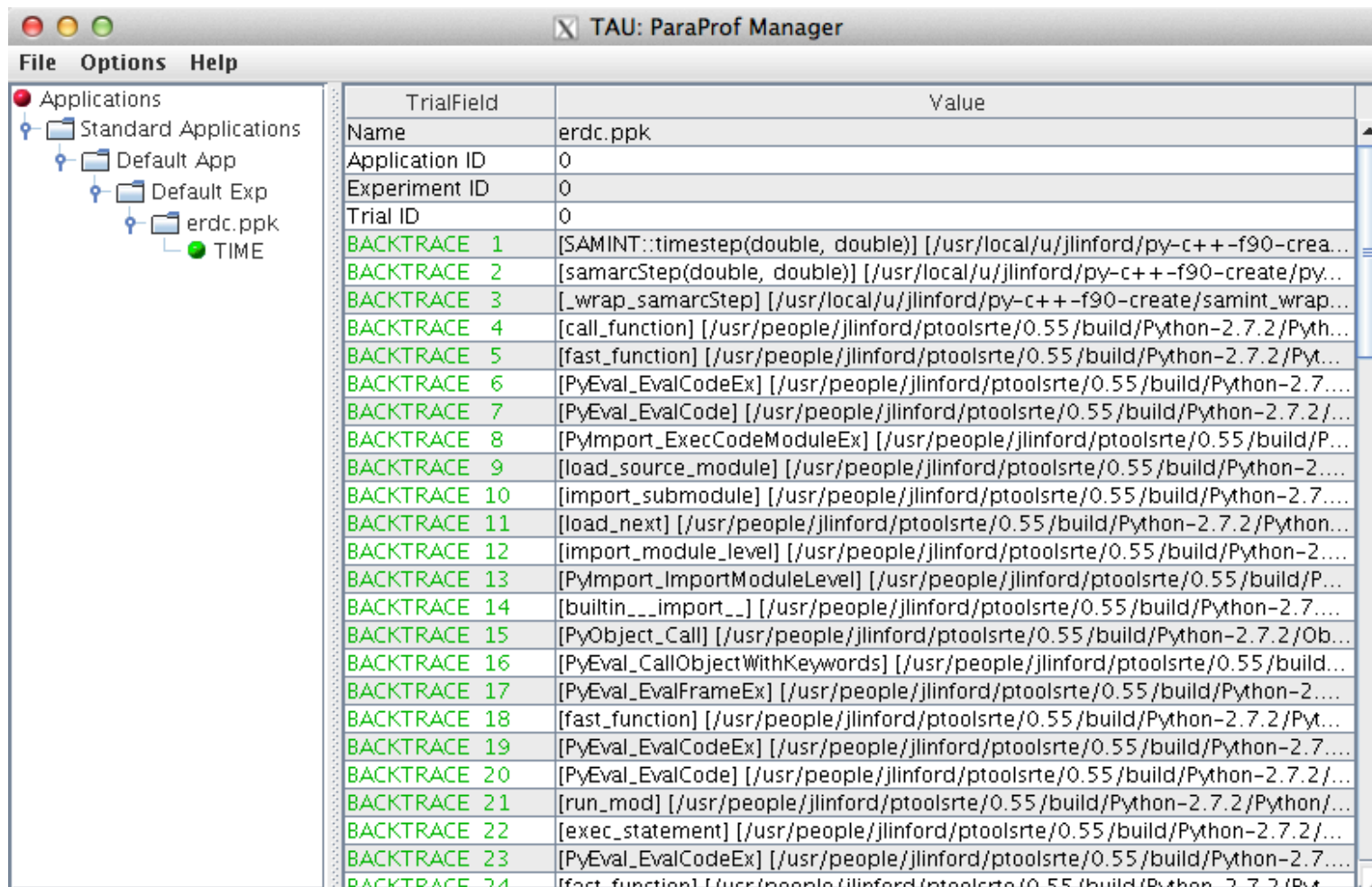
# TAU generates profile data at time of failure

```
        step: 1      time: 0.0
SAMINT::timestep()
printing output
SAMINT::getGlobalNumberPatches()
SAMINT::getLocalNumberPatches()
SAMINT::writePlotData()
-----
        step: 1      time: 0.0
SAMINT::timestep()
TAU: Caught signal 8 (Floating point exception), dumping profile with stack trace: [rank=0, pid=2422
6, tid=0]...
TAU: Caught signal 8 (Floating point exception), dumping profile with stack trace: [rank=1, pid=2422
5, tid=0]...
TAU: Caught signal 8 (Floating point exception), dumping profile with stack trace: [rank=2, pid=2422
4, tid=0]...
TAU: Caught signal 8 (Floating point exception), dumping profile with stack trace: [rank=3, pid=2422
3, tid=0]...

mpirun has exited due to process rank 1 with PID 24100 on
node m0607.mana exiting without calling "finalize". This may
have caused other processes in the application to be
terminated by signals sent by mpirun (as reported here).
-----
[jlinford@m0607 example]$
```



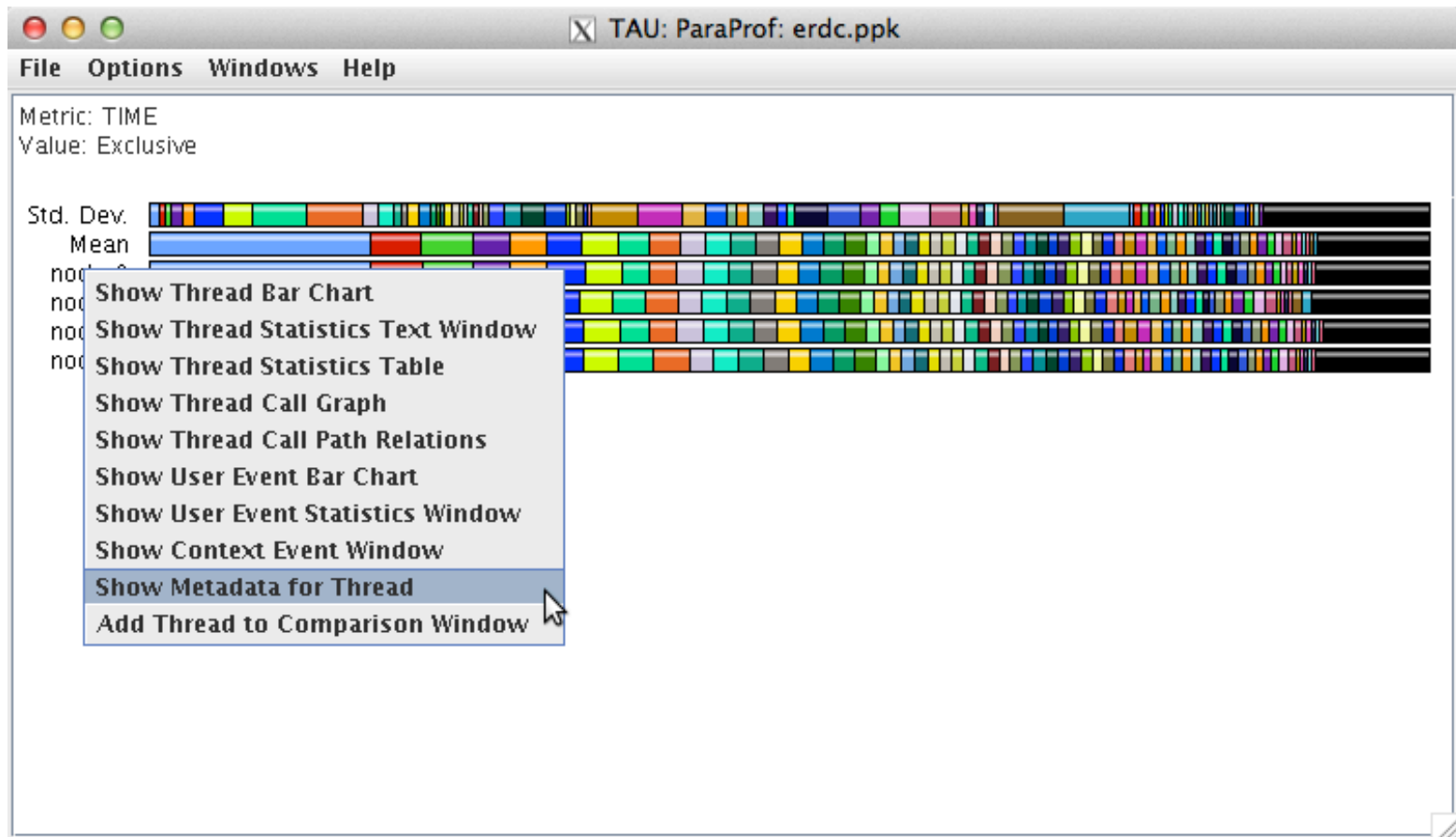
# Use ParaProf to explore the profile data



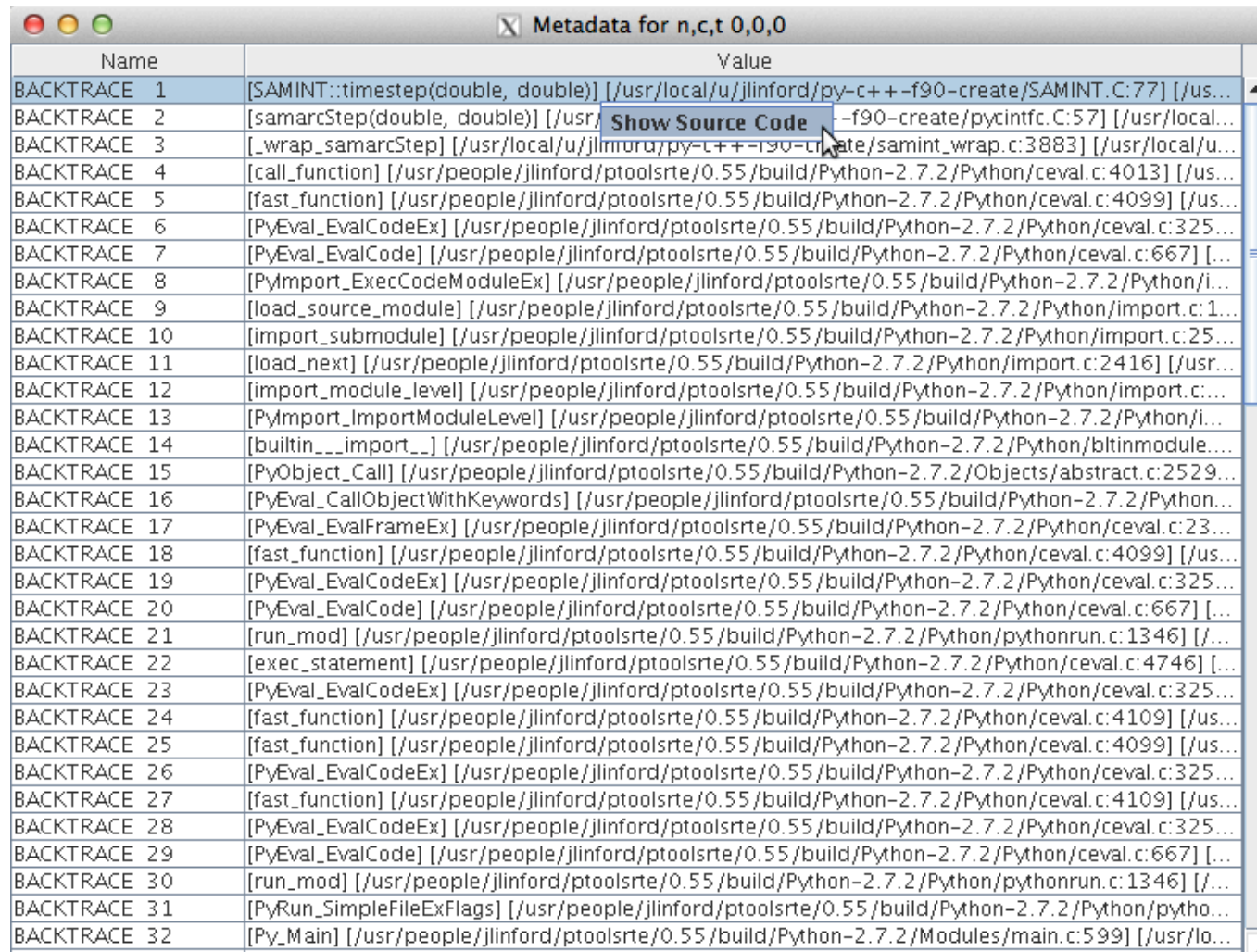
The screenshot shows the TAU: ParaProf Manager application window. The title bar reads "TAU: ParaProf Manager". The menu bar includes "File", "Options", and "Help". On the left, a tree view shows the application hierarchy: Applications > Standard Applications > Default App > Default Exp > erdc.ppk > TIME. The main area is a table with two columns: "TrialField" and "Value".

| TrialField     | Value   |
|----------------|---|
| Name           | erdc.ppk  |
| Application ID | 0   |
| Experiment ID  | 0   |
| Trial ID       | 0   |
| BACKTRACE 1    | [SAMINT::timestep(double, double)] [/usr/local/u/jlinford/py-c++-f90-crea...    |
| BACKTRACE 2    | [samarcStep(double, double)] [/usr/local/u/jlinford/py-c++-f90-create/py...     |
| BACKTRACE 3    | [_wrap_samarcStep] [/usr/local/u/jlinford/py-c++-f90-create/samint_wrap...      |
| BACKTRACE 4    | [call_function] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Pyth... |
| BACKTRACE 5    | [fast_function] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Pyt...  |
| BACKTRACE 6    | [PyEval_EvalCodeEx] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7....   |
| BACKTRACE 7    | [PyEval_EvalCode] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/...   |
| BACKTRACE 8    | [PyImport_ExecCodeModuleEx] [/usr/people/jlinford/ptoolsrte/0.55/build/P...     |
| BACKTRACE 9    | [load_source_module] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2....    |
| BACKTRACE 10   | [import_submodule] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7....    |
| BACKTRACE 11   | [load_next] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python...   |
| BACKTRACE 12   | [import_module_level] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2....   |
| BACKTRACE 13   | [PyImport_ImportModuleLevel] [/usr/people/jlinford/ptoolsrte/0.55/build/P...    |
| BACKTRACE 14   | [builtin__import_] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7....    |
| BACKTRACE 15   | [PyObject_Call] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Ob...   |
| BACKTRACE 16   | [PyEval_CallObjectWithKeywords] [/usr/people/jlinford/ptoolsrte/0.55/build...   |
| BACKTRACE 17   | [PyEval_EvalFrameEx] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2....    |
| BACKTRACE 18   | [fast_function] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Pyt...  |
| BACKTRACE 19   | [PyEval_EvalCodeEx] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7....   |
| BACKTRACE 20   | [PyEval_EvalCode] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/...   |
| BACKTRACE 21   | [run_mod] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/...    |
| BACKTRACE 22   | [exec_statement] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/...    |
| BACKTRACE 23   | [PyEval_EvalCodeEx] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7....   |
| BACKTRACE 24   | [fact_function] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Pyt...  |

# Right-click the thread you want to explore

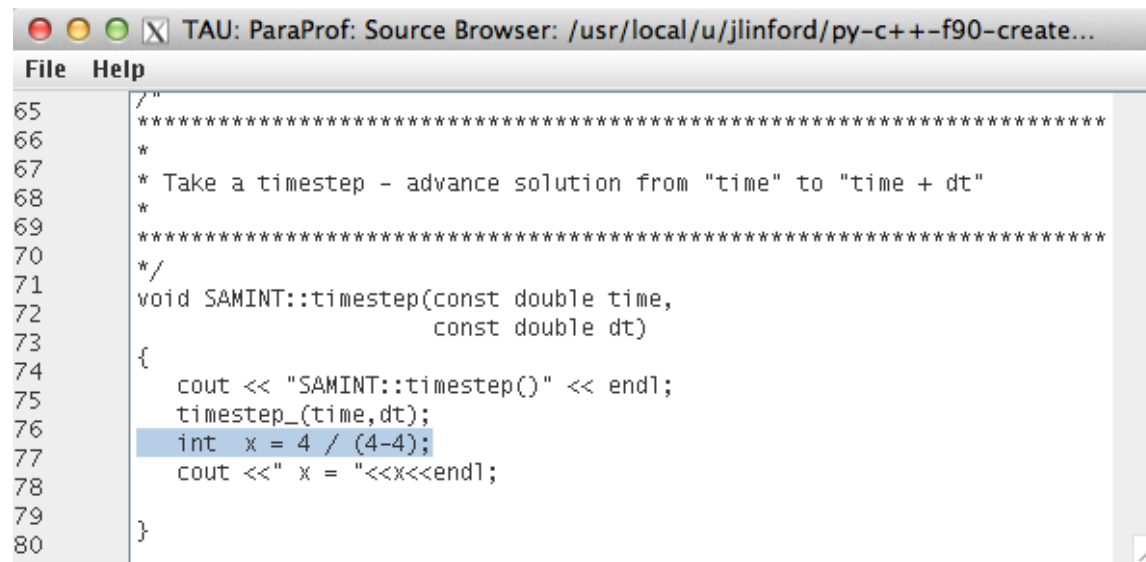


# Use the Metadata window to locate the source line that caused the error



| Name         | Value  |
|--------------|--|
| BACKTRACE 1  | [SAMINT::timestep(double, double)] [/usr/local/u/jlinford/py-c++-f90-create/SAMINT.C:77] [/usr/lo... |
| BACKTRACE 2  | [samarcStep(double, double)] [/usr/lo... --f90-create/pycintfc.C:57] [/usr/local/...                 |
| BACKTRACE 3  | [_wrap_samarcStep] [/usr/local/u/jlinford/py-c++-f90-create/samint_wrap.c:3883] [/usr/local/u...     |
| BACKTRACE 4  | [call_function] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/ceval.c:4013] [/us... |
| BACKTRACE 5  | [fast_function] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/ceval.c:4099] [/us... |
| BACKTRACE 6  | [PyEval_EvalCodeEx] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/ceval.c:325...    |
| BACKTRACE 7  | [PyEval_EvalCode] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/ceval.c:667] [...]  |
| BACKTRACE 8  | [PyImport_ExecCodeModuleEx] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/i...      |
| BACKTRACE 9  | [load_source_module] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/import.c:1...    |
| BACKTRACE 10 | [import_submodule] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/import.c:25...     |
| BACKTRACE 11 | [load_next] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/import.c:2416] [/usr/...  |
| BACKTRACE 12 | [import_module_level] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/import.c:...    |
| BACKTRACE 13 | [PyImport_ImportModuleLevel] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/i...     |
| BACKTRACE 14 | [builtin___import_] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/bltinmodule...    |
| BACKTRACE 15 | [PyObject_Call] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Objects/abstract.c:2529...   |
| BACKTRACE 16 | [PyEval_CallObjectWithKeywords] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/...   |
| BACKTRACE 17 | [PyEval_EvalFrameEx] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/ceval.c:23...    |
| BACKTRACE 18 | [fast_function] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/ceval.c:4099] [/us... |
| BACKTRACE 19 | [PyEval_EvalCodeEx] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/ceval.c:325...    |
| BACKTRACE 20 | [PyEval_EvalCode] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/ceval.c:667] [...]  |
| BACKTRACE 21 | [run_mod] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/pythonrun.c:1346] [/...     |
| BACKTRACE 22 | [exec_statement] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/ceval.c:4746] [...]  |
| BACKTRACE 23 | [PyEval_EvalCodeEx] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/ceval.c:325...    |
| BACKTRACE 24 | [fast_function] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/ceval.c:4109] [/us... |
| BACKTRACE 25 | [fast_function] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/ceval.c:4099] [/us... |
| BACKTRACE 26 | [PyEval_EvalCodeEx] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/ceval.c:325...    |
| BACKTRACE 27 | [fast_function] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/ceval.c:4109] [/us... |
| BACKTRACE 28 | [PyEval_EvalCodeEx] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/ceval.c:325...    |
| BACKTRACE 29 | [PyEval_EvalCode] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/ceval.c:667] [...]  |
| BACKTRACE 30 | [run_mod] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/pythonrun.c:1346] [/...     |
| BACKTRACE 31 | [PyRun_SimpleFileExFlags] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Python/pytho...    |
| BACKTRACE 32 | [Py_Main] [/usr/people/jlinford/ptoolsrte/0.55/build/Python-2.7.2/Modules/main.c:599] [/usr/lo...    |

# ParaProf highlights the erroneous line

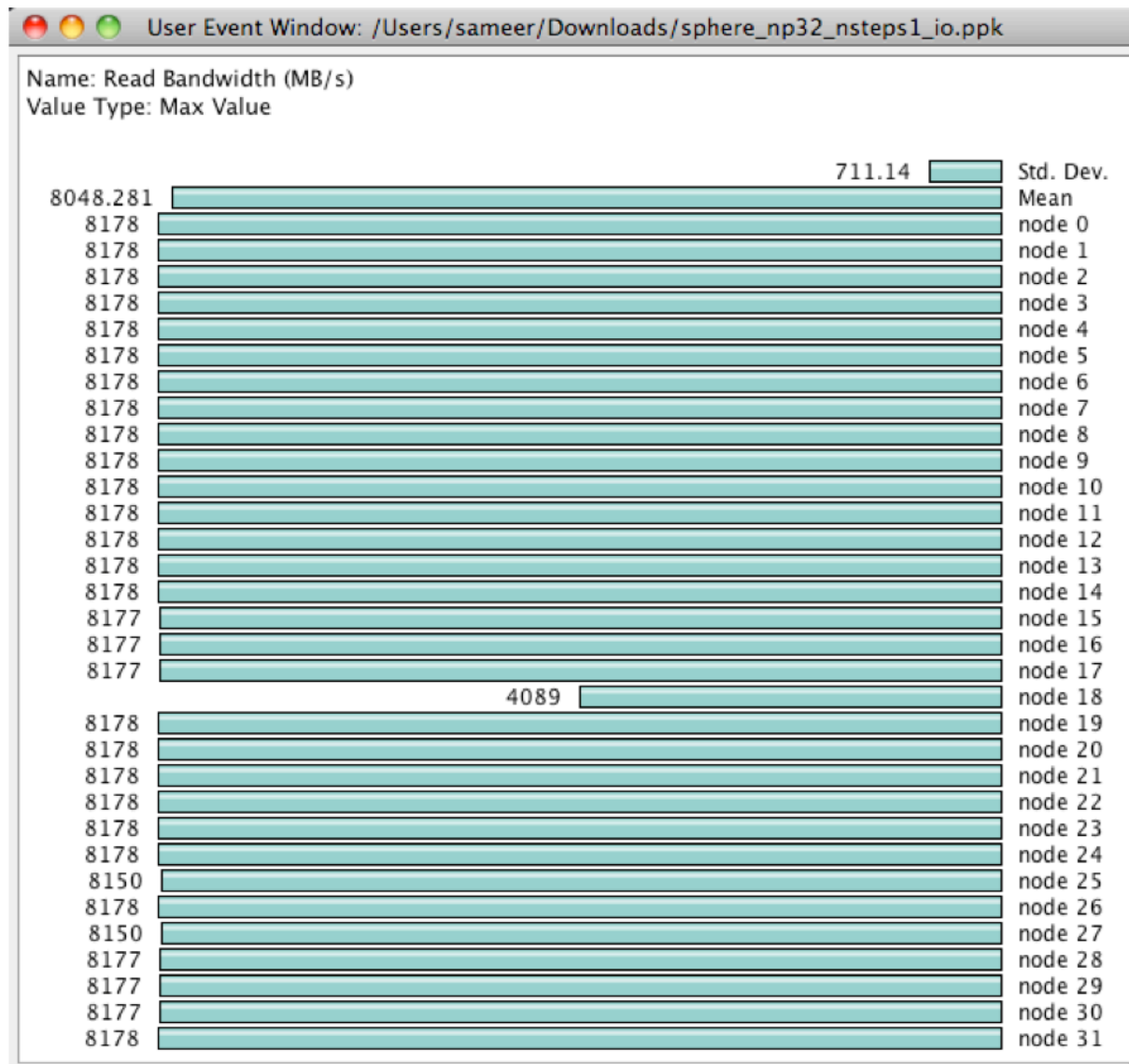


The screenshot shows a window titled "TAU: ParaProf: Source Browser: /usr/local/u/jlinford/py-c++-f90-create...". The window contains a code editor with the following C++ code:

```
65  /*  
66  *  
67  * Take a timestep - advance solution from "time" to "time + dt"  
68  *  
69  *  
70  */  
71  void SAMINT::timestep(const double time,  
72                      const double dt)  
73  {  
74      cout << "SAMINT::timestep()" << endl;  
75      timestep_(time,dt);  
76      int x = 4 / (4-4);  
77      cout << " x = " << x << endl;  
78  }  
79  
80
```

The line `int x = 4 / (4-4);` on line 76 is highlighted in blue, indicating it is the erroneous line.

# Peak read bandwidth in Helios



# A segmentation fault in Kestrel with memory and I/O diagnostics

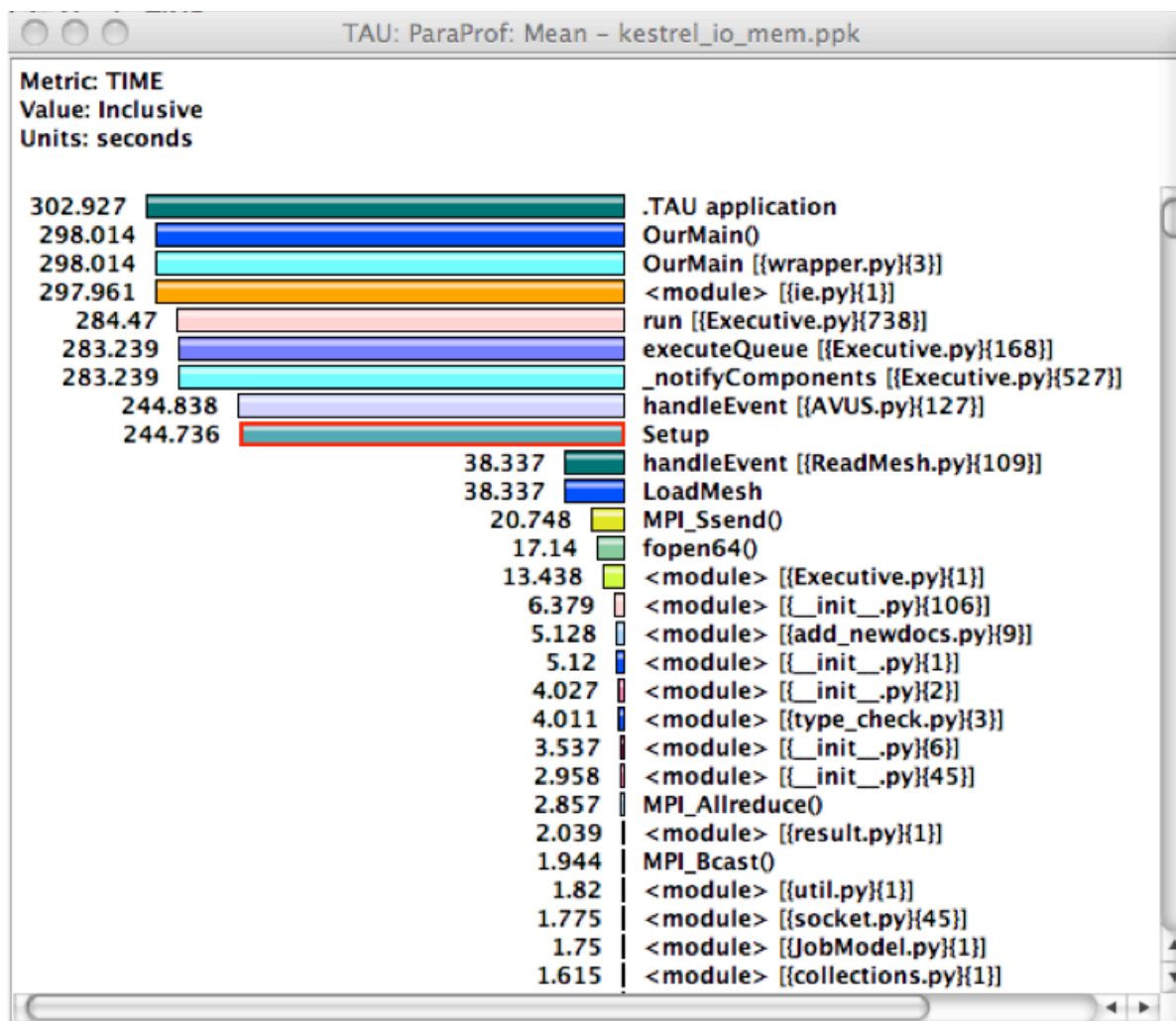
TAU: ParaProf: Mean Context Events - kestrel\_error\_8p.ppk

| Name  | Total          | NumSamples | MaxValue   | MinValue | MeanValue     | Std. Dev.     |
|---|----------------|------------|------------|----------|---------------|---------------|
| .TAU application                            |                |            |            |          |               |               |
| OurMain()                                   |                |            |            |          |               |               |
| OurMain [{{wrapper.py}}{3}]                 |                |            |            |          |               |               |
| <module> [{{ie.py}}{1}]                     |                |            |            |          |               |               |
| run [{{Executive.py}}{738}]                 |                |            |            |          |               |               |
| executeQueue [{{Executive.py}}{168}]        |                |            |            |          |               |               |
| _notifyComponents [{{Executive.py}}{527}]   |                |            |            |          |               |               |
| handleEvent [{{ForcesMomentsCalc.py}}{105}] |                |            |            |          |               |               |
| Initialize                                  |                |            |            |          |               |               |
| TAU_SIGNAL (Segmentation fault)             | 1              | 1          | 1          | 1        | 1             | 0             |
| Message size for all-gather                 | 10,621,936     | 38         | 10,318,464 | 4        | 279,524.632   | 1,651,028.694 |
| Message size for all-reduce                 | 13,700         | 597        | 1,008      | 4        | 22.948        | 42.576        |
| Message size for all-to-all                 | 152            | 19         | 8          | 8        | 8             | 0             |
| Message size for broadcast                  | 144,068,872    | 204        | 56,427,296 | 3        | 706,219.961   | 4,746,523.542 |
| Message size for gather                     | 41,492,662.361 | 10.375     | 14,167,198 | 4        | 3,999,292.758 | 4,242,989.257 |
| Message size for reduce                     | 92             | 22         | 8          | 4        | 4.182         | 0.833         |
| Message size for scatter                    | 621,421        | 12         | 621,377    | 4        | 51,785.083    | 171,738.425   |
| TAU_SIGNAL (Segmentation fault)             | 1              | 1          | 1          | 1        | 1             | 0             |

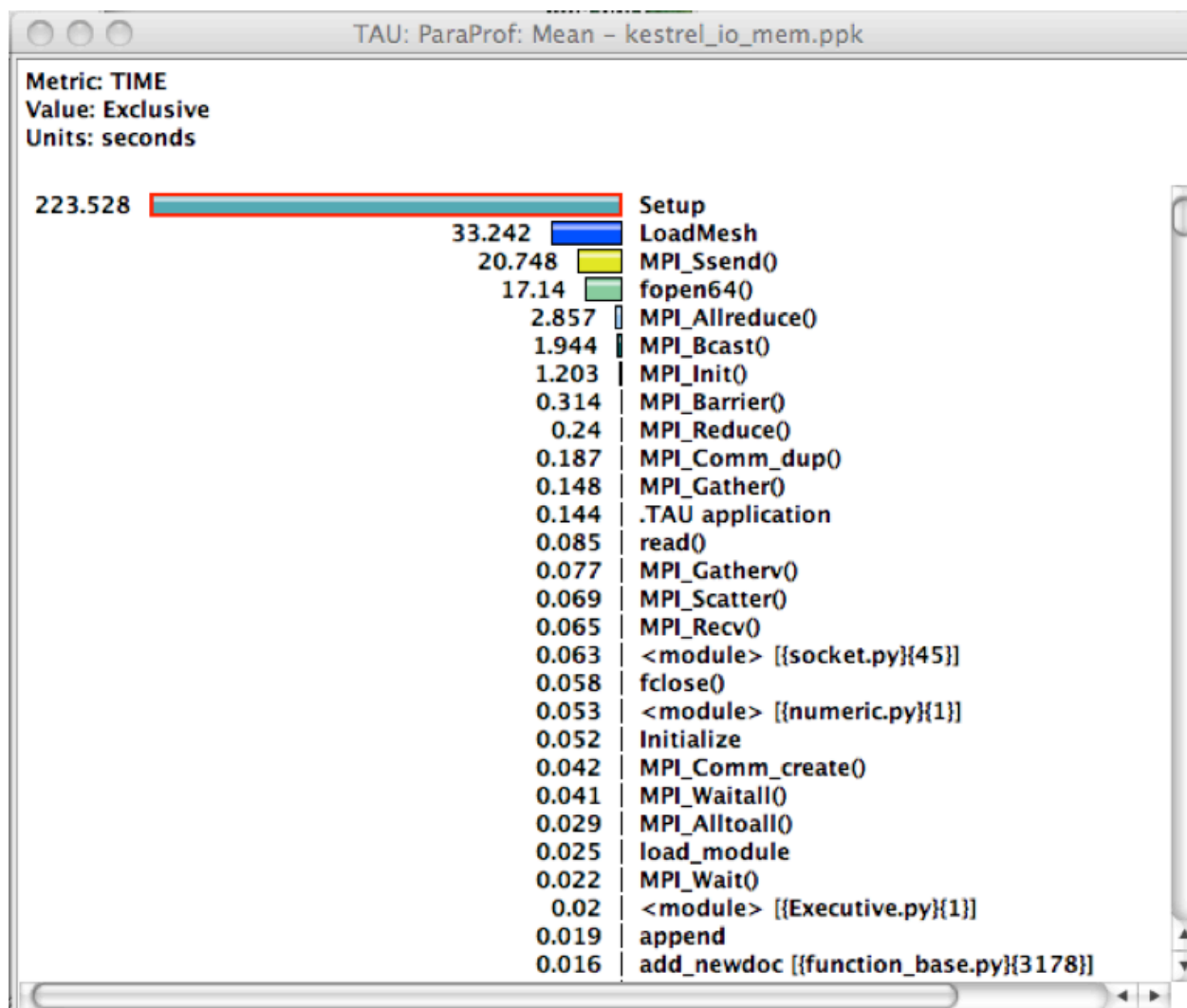
TAU: ParaProf: Context Events for: n,c,t 0,0,0 - kestrel\_io\_mem.ppk

| Name                                     | Total         | NumSamples    | MaxValue    | MinValue    | MeanValue  | Std. Dev.   |
|--|---------------|---------------|-------------|-------------|------------|-------------|
| .TAU application                         |               |               |             |             |            |             |
| Memory Utilization (heap, in KB)         | 47,690,869    | 1,319,090.592 | 0.023       | 275,509.524 | 38,071.609 |             |
| malloc size (bytes)                      | 3,012,988,071 | 23,869,030    | 496,523,280 | 1           | 126.23     | 113,322.074 |
| free size (bytes)                        | 2,248,780,956 | 23,821,839    | 496,523,280 | 1           | 94.4       | 108,796.005 |
| Bytes Read                               | 158,530,329   | 19,180        | 1,126,076   | 2           | 8,265.398  | 9,607.032   |
| Read Bandwidth (MB/s)                    | 19,156        | 10,982        | 0.051       | 1,209.371   | 277.131    |             |
| Bytes Read <file=/mnt/cfs/pkgs/create/av | 151,920,360   | 18,545        | 8,192       | 8,172       | 8,191.985  | 0.543       |
| Read Bandwidth (MB/s) <file=/mnt/cfs/pk  | 18,545        | 2,048         | 327.68      | 1,238.092   | 153.524    |             |
| Message size for all-reduce              | 13,700        | 597           | 1,008       | 4           | 22.948     | 42.576      |
| Bytes Written                            | 64,053        | 238           | 4,161       | 2           | 269.13     | 973.394     |
| Write Bandwidth (MB/s)                   | 212           | 1,040.25      | 0.085       | 46.81       | 174.011    |             |

# Inclusive time spent in Kestrel code regions



# Exclusive time spent in Kestrel code regions







# Summary and conclusions

1. TAU callstack debugging isolates errors in multi-language HPC software by intercepting signals at runtime
2. Run codes with *tau\_exec* to register the TAU signal handler, create profile files, and shutdown gracefully
3. The profile can be sent to developers when sensitive or proprietary inputs cannot be provided
4. Developers use ParaProf to analyze the fault location and runtime performance data in the profiles
5. Memory use, IO, and runtime performance are recorded
6. No recompilation necessary



# Acknowledgments

**This work was supported by the DoD High Performance Computing Modernization Program (HPCMP) User Productivity Enhancement, Technology Transfer and Training (PETTT) program and through support provided by the DoD HPCMO to the HIARMS Institute and the CREATE program.**