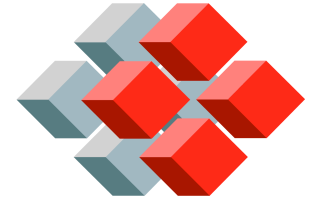


ETH

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

CSCS

Swiss National Supercomputing Centre



Parallel Profiling using TAU on Rosa

Introduction to Rosa Course
Scientific Computing Group



Outline

- TAU features
- Location and supported configurations on Rosa
- Using TAU with your application
 - Method 1
 - Method 2
 - Method 3
- Analysis tools
- References



TAU Features

- Typical options
 - Flat MPI and callgraph profiling
 - Hardware counter data collection
 - OpenMP & pthread profiling
 - MPI tracing
 - Memory profiling
- Characteristics features
 - Auto-instrumentation utility (PDT)
 - Custom configurations
 - Interoperability with other tools

Location and available configurations

- Available in /apps/tau/2.18.2p2 for:
 - PGI
 - Pathscale
 - GNU
- Available configurations (note further configurations could be easily built if needed):
 - MPI, callpath/tree and hardware counters profile
 - Memory usage and headroom profile
 - Pthread and OpenMP profile



Instrumenting an Application

- Method 1:
 - Modify makefile/config script
 - module load tau (sets TAU_ROOT)
- Method 2:
 - Define environment variables (TAU_MAKEFILE and PATH)
 - Call tau compiler scripts (tau_cc and tauf90)
- Method 3:
 - No recompilation, only relinking (no auto-instrumentation)

Method 1

- Users can include the TAU auto-instrumentation capability by including a script into their makefile or config scripts
- At runtime, could invoke options such as hardware counters, and detailed MPI profiles

```
> module load tau  
> vi makefile
```

```
include $(TAU_ROOT)/lib/Makefile.tau-papi-mpi-pdt-pgi
```

```
CC = $(TAU_COMPILER) CC  
LINK = $(TAU_COMPILER) CC  
LINKFLAGS = -O -L$(LIBS)
```

Method 2

- > `module load tau`
- > `export TAU_MAKEFILE=$TAU_ROOT/lib/Makefile.tau-papi-mpi-pdt-pgi`
- > `export PATH=$PATH:$TAU_ROOT/bin`

- > `tau_cc.sh -o hello hello.c`

Debug: Parsing with PDT Parser

```
Executing> /apps/tau/2.18.2p2/pgi_805/pdtoolkit-3.14.1/craycnl/bin/  
  cparse hello.c -I/apps/tau/2.18.2p2/pgi_805/tau-2.18.2p2/include  
  -DPROFILING_ON -DTAU_PAPI -I/opt/xt-tools/papi/3.6.2/v23/linux/  
  src -I/opt/xt-tools/papi/3.6.2/v23/linux/include -  
  DTAU_DOT_H_LESS_HEADERS -DTAU_MPI -DTAU_MPI_THREADED -  
  DTAU_LINUX_TIMERS -DTAU_MPIGREQUEST -DTAU_MPIDATAREP -  
  DTAU_MPIERRHANDLER -DTAU_MPIATTRFUNCTION -DTAU_MPITYPEEX -  
  DTAU_MPIADDEROR -DTAU_LARGEFILE -D_LARGEFILE64_SOURCE -DTAU_BFD  
  -DTAU_MPIFILE -DHAVE_GNU_DEMANGLE -I/opt/mpt/default/xt/mpich2-  
  pgi/include
```

Method 3

- A simple relink would only instrument library functions (for example, MPI) or functions instrumented with the TAU API:
 - TAU_PROFILE
 - TAU_PROFILE_SET_NODE
 - TAU_PROFILE_START
 - TAU_PROFILE_STOP
 - TAU_PROFILE_TIMER

```
> cc -o test test.c -L$TAU_ROOT/lib -lTauMpi-papi-mpi-pdt-pgi  
-lpthread -lrt -lmpichcxx -lmpich -lrt -ltau-papi-mpi-pdt-pgi  
-lstd -lC
```




Analysis tools

- After a successful run, output files are generated:
 - profile.x.x.x files
 - When PAPI counters are selected, output files are in MULTI___PAPI_COUNTER
- pprof—text based analysis tools
- paraprof—graphical tool



Offline Analysis

- Profile data collected on Rosa could be viewed on any system running JVM
- Steps taken to run paraprof offline on a personal system:
 1. Download TAU (<http://www.cs.uoregon.edu/research/tau/home.php>)
 2. Untar the file
 3. Type ./configure in tau_verno directory. This will generate paraprof in x86_64/bin or apple/bin
 4. Invoke paraprof after copying over the profile files from Rosa

Output formats (text and GUI)—Short demo to follow

The screenshot displays the TAU ParaProf Manager interface with several analysis windows:

- Load balance:** A horizontal bar chart showing the distribution of work across 51 nodes.
- Cost distribution:** A 3D surface plot showing the cost distribution across the node grid.
- COMM Outliers:** A heatmap titled "NUMBER OF CALLS" showing communication patterns between nodes.
- Quantitative evaluation:** A table of user events with columns for NumSamples, MaxValue, MinValue, Mean, and Message description.
- Callgraph bottleneck:** A network graph showing communication links between nodes, with a callout for a bottleneck.
- Config details:** A list of system and application parameters.

USER EVENTS Profile :NODE 0, CONTEXT	NumSamples	MaxValue	MinValue	Mean	Message
her	14	36	2	36	0 Message size for all-gat
uce	672	48	4	25.04	21.02 Message size for all-red
st	523	5.249E+04	2	1.203E+04	1.712E+04 Message size for broadca
om all nodes	1.176E+04	1.324E+05	0	7109	1.315E+04 Message size received fr
om all nodes	1.176E+04	1.349E+05	0	5759	1.276E+04 Message size sent to all
e 1	1961	1.349E+05	0	1.099E+04	1.985E+04 Message size sent to nod
e 1 : void LAMMPS_NS::Comm::borders() => MPI_Send()	12	1.349E+05	1.332E+05	1.34E+05	464.1 Message size sent to nod
e 1 : void LAMMPS_NS::Comm::borders() => MPI_Sendrecv()	12	4	4	4	0 Message size sent to nod
e 1 : void LAMMPS_NS::Comm::borders() => MPI_Sendrecv()	101	5.059E+04	4.997E+04	5.027E+04	158.1 Message size sent to nod

Name	TrialField	Value
Name	pro2/tmp/alam/xt-home/home/gpfs/	
Application ID	0	
Experiment ID	0	
Trial ID	0	
CPU Cores	4	
CPU MHz	2400.000	
CPU Type	Quad-Core AMD Opteron(tm) Process...	
CPU Vendor	AuthenticAMD	
CWD	/lus/scratch/alam/lammps-29Jun09/...	
Cache Size	512 kB	
Executable	/var/spool/alps/45750/lmp_xt5	
Hostname	nid01985	
Local Time	2009-06-30T09:44:13+02:00	
MPI Processor Name	nid01985	
Memory Size	16385988 kB	
Node Name	nid01985	
OS Machine	x86_64	
OS Name	Linux	
OS Release	2.6.16.60-0.33.1.0102.4270.2.2.2...	
OS Version	#1 SMP Thu Apr 23 11:54:49 PDT 2...	
Starting Timestamp	1246347849108323	
TAU Architecture	craycnl	
TAU Config	-arch=craycnl -pdt=/apps/tau/2.18...	
TAU Version	tau-2.18.2p2	
TAU_CALLPATH	on	
TAU_CALLPATH_DEPTH	2	
TAU_COMM_MATRIX	off	
TAU_COMPENSATE	off	



Reference

- TAU website (<http://www.cs.uoregon.edu/research/tau/home.php>)
 - Documentation
 - Tutorial
 - Downloads (offline analysis utilities)
- CSCS Rosa website (<http://www.cscs.ch/Performance-TAU.258.0.html>)
 - Module setup
 - Usage information