Parallel Processing
Lecture 5

Dr. Guy Tel-Zur

~~Version 11-2014~~
~~Version 05-2015~~
~~Version 04-2017~~
~~Version 04-2020~~
~~Version 11-2020~~
Version 11-2021

תכנית השיעור

- סיכום השיעורים הקודמים והמעבדה.
- נושא השיעור היום: המשך לימוד MPI + אלגוריתמים מקביליים:

- Partitioning and Divide and Conquer Strategies (slides4.ppt)
- Synchronous computations (slides6.ppt)
- Following:

Dr. Barry Wilkinson

- **UNIVERSITY OF NORTH CAROLINA AT CHARLOTTE**
  Department of Computer Science
  ITCS 4145/5145 Parallel Programming
  Spring 2009

Scalasca: a toolset for scalable performance analysis of large-scale applications

1. prepare application objects and executable for measurement:

```
scalasca -instrument <compile-or-link-command> # skin
```

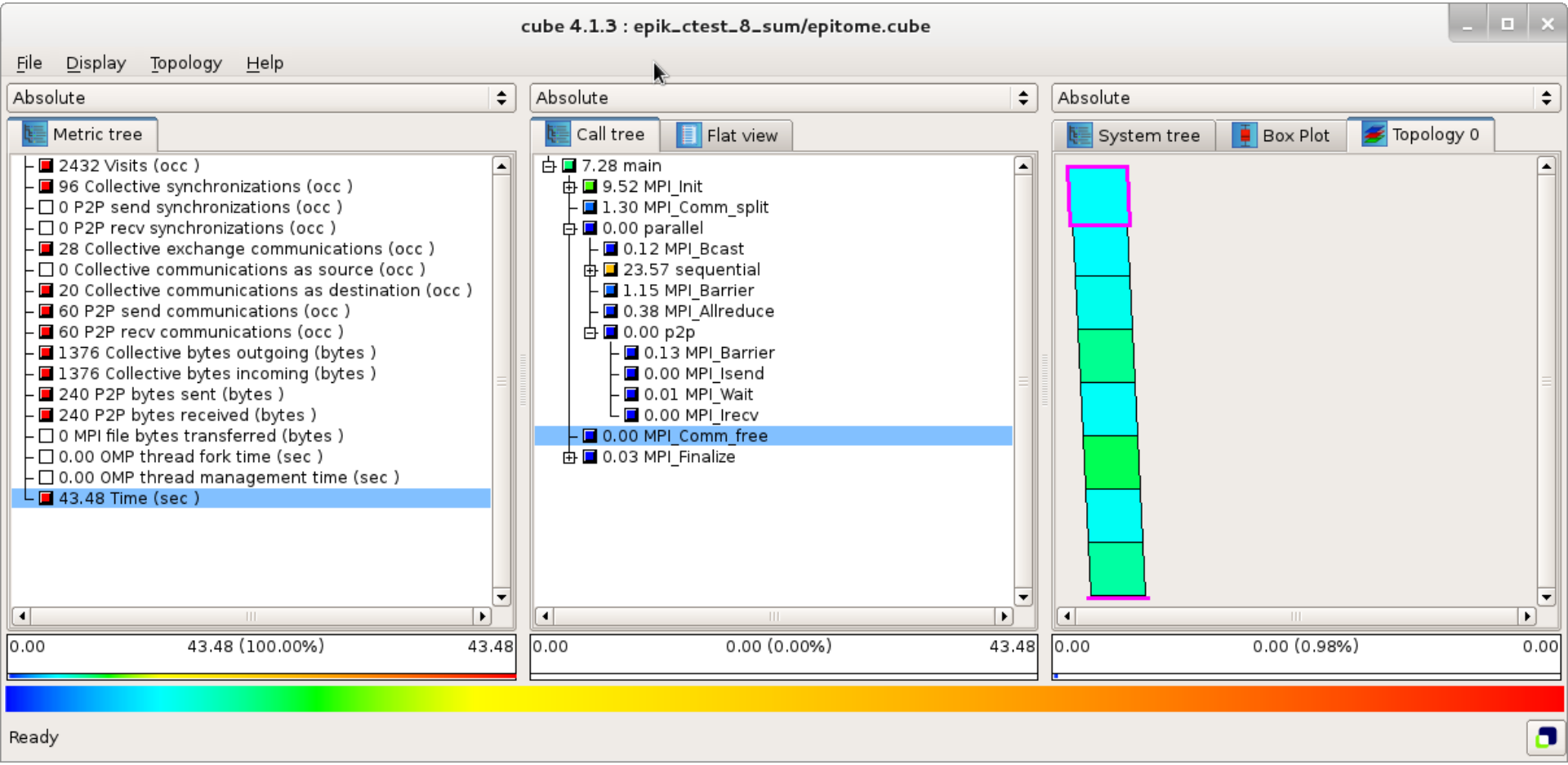
2. run application under control of measurement system:

```
scalasca -analyze <application-launch-command> # scan
```

3. interactively explore measurement analysis report:

```
scalasca -examine <experiment-archive|report> # square
```

...square

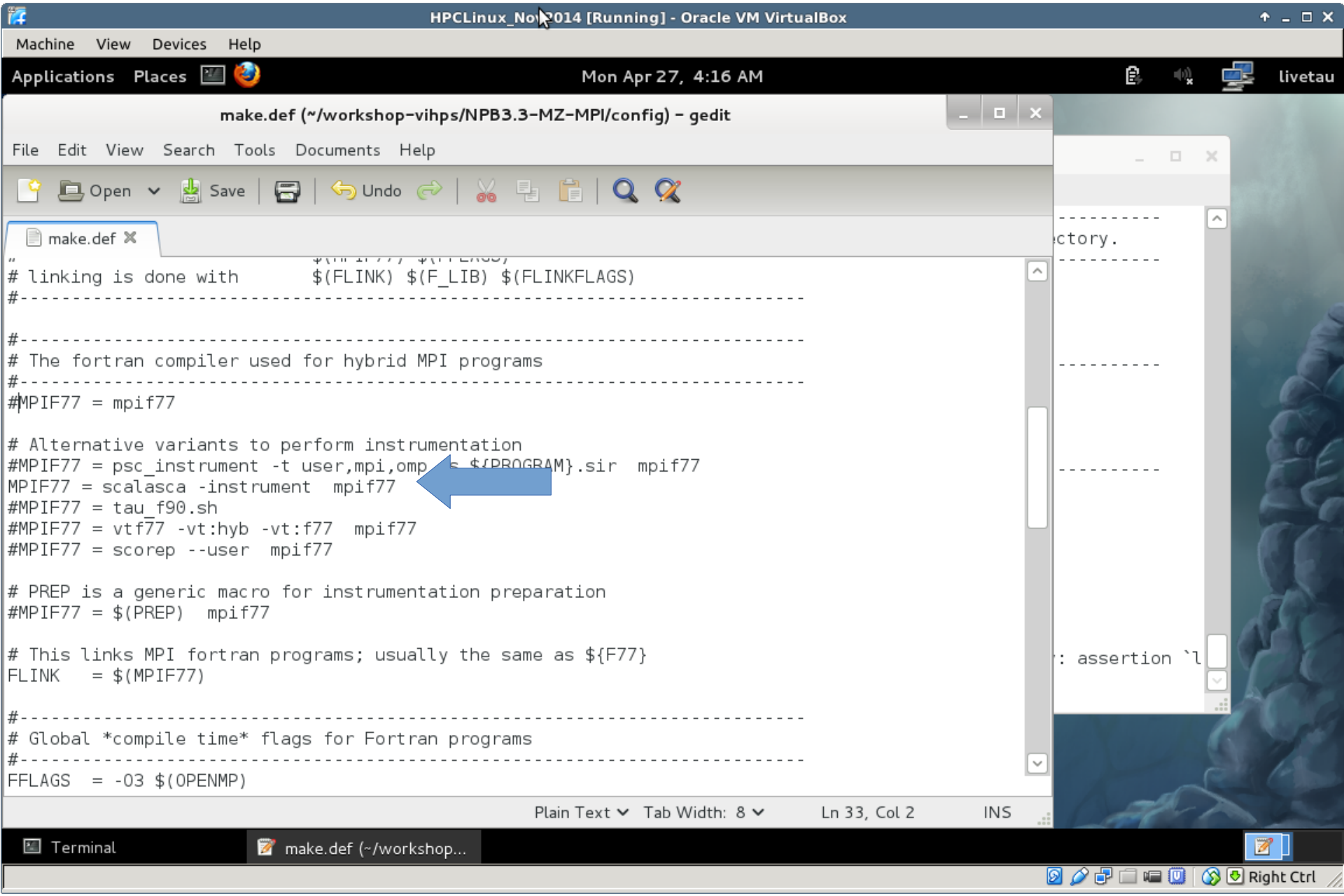

cube 4.1.3 : epik_ctest_8_sum/epitome.cube
File
Display
Topology
Help
Absolute
Absolute
Absolute
Metric tree
Call tree
Flat view
System tree
Box Plot
Topology 0
2432 Visits (occ)
96 Collective synchronizations (occ)
0 P2P send synchronizations (occ)
0 P2P recv synchronizations (occ)
28 Collective exchange communications (occ)
0 Collective communications as source (occ)
20 Collective communications as destination (occ)
60 P2P send communications (occ)
60 P2P recv communications (occ)
1376 Collective bytes outgoing (bytes)
1376 Collective bytes incoming (bytes)
240 P2P bytes sent (bytes)
240 P2P bytes received (bytes)
0 MPI file bytes transferred (bytes)
0.00 OMP thread fork time (sec)
0.00 OMP thread management time (sec)
43.48 Time (sec)
7.28 main
9.52 MPI_Init
1.30 MPI_Comm_split
0.00 parallel
0.12 MPI_Bcast
23.57 sequential
1.15 MPI_Barrier
0.38 MPI_Allreduce
0.00 p2p
0.13 MPI_Barrier
0.00 MPI_Isend
0.01 MPI_Wait
0.00 MPI_Irecv
0.00 MPI_Comm_free
0.03 MPI_Finalize
0.00
43.48 (100.00%)
43.48
0.00
0.00 (0.00%)
43.48
0.00
0.00 (0.98%)
0.00
Ready

Scalasca and TAU in action - a demo on the BGU_VM/HPC_VM

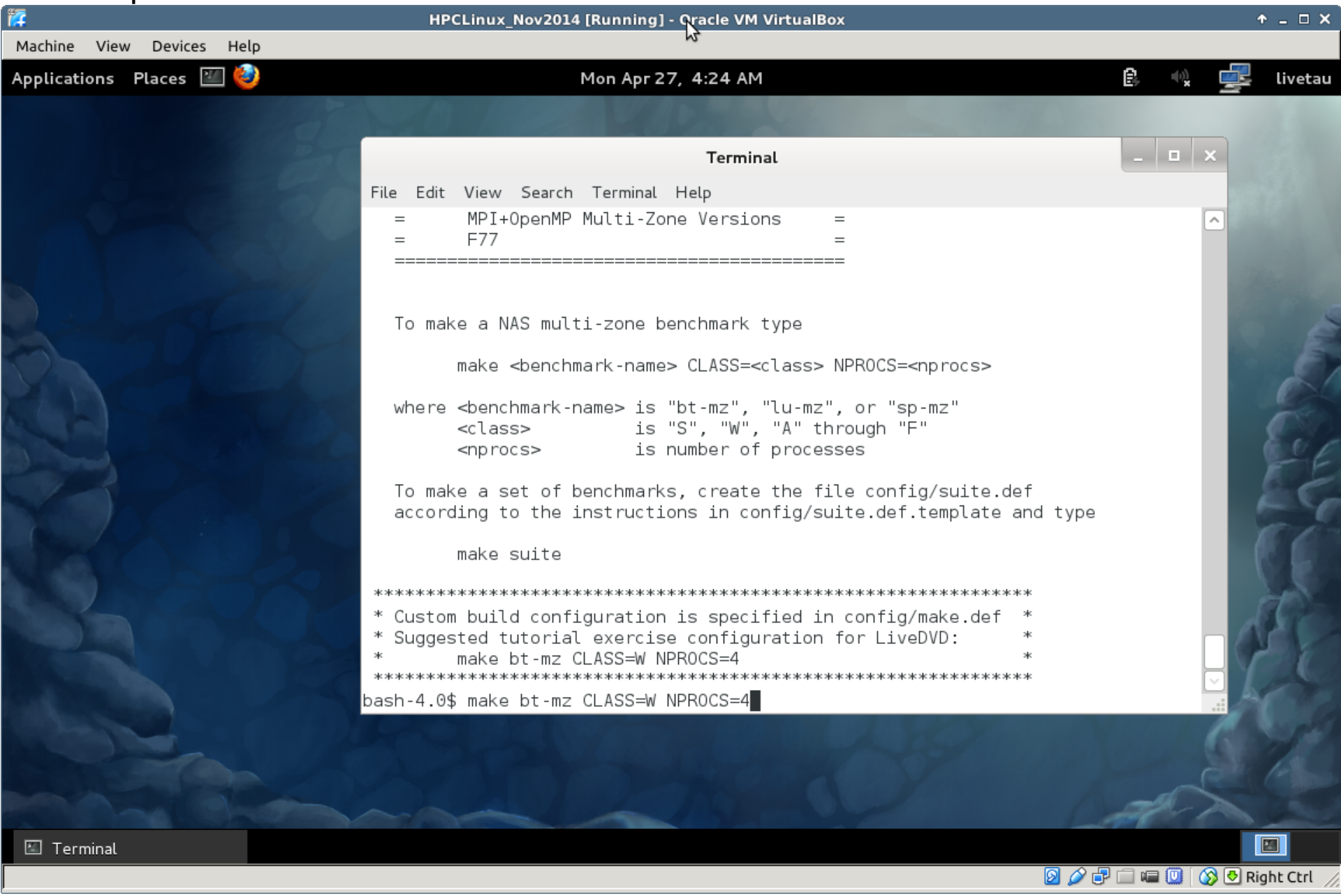
2 Demos:

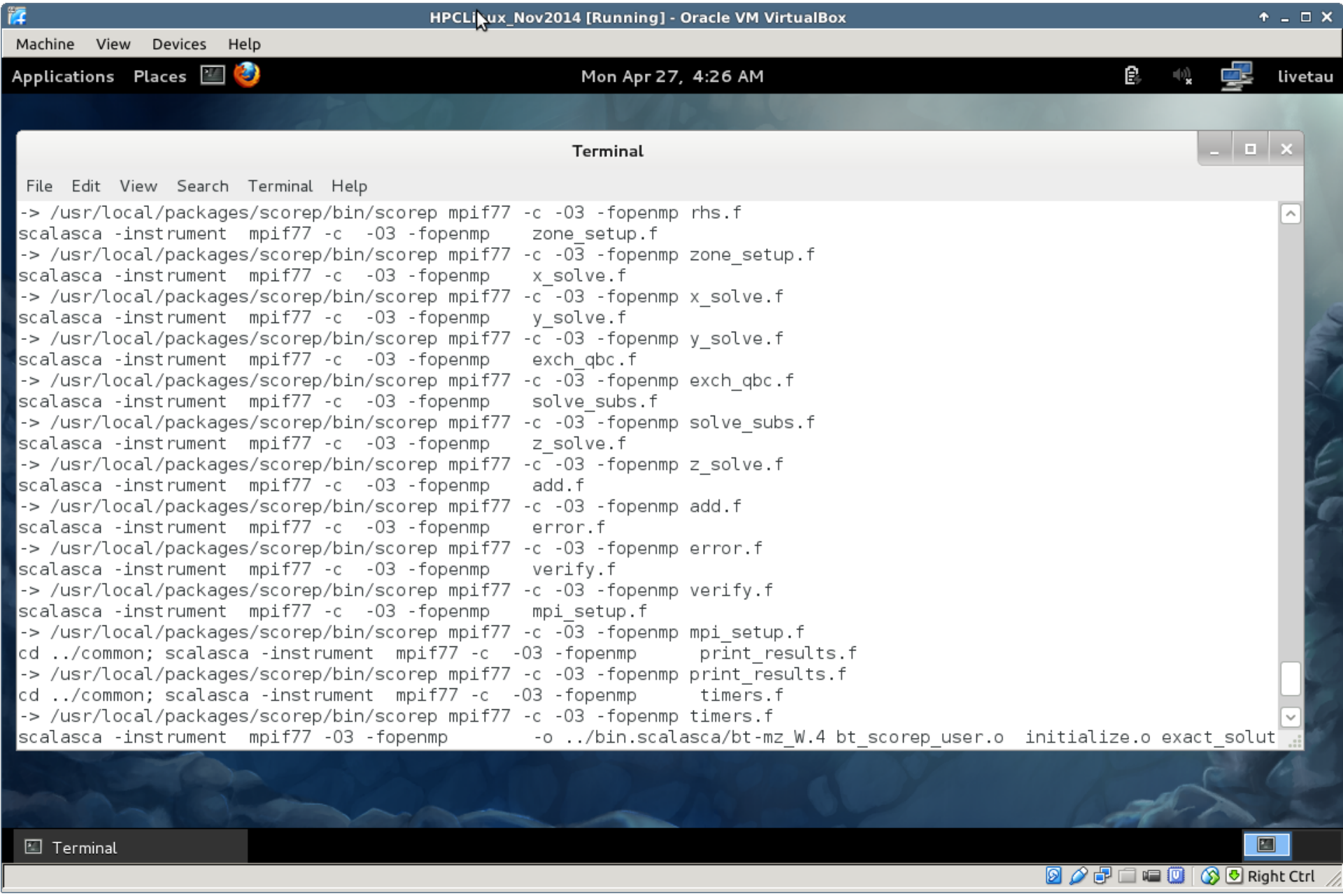
1. using scalasca
2. using TAU and jumpshot

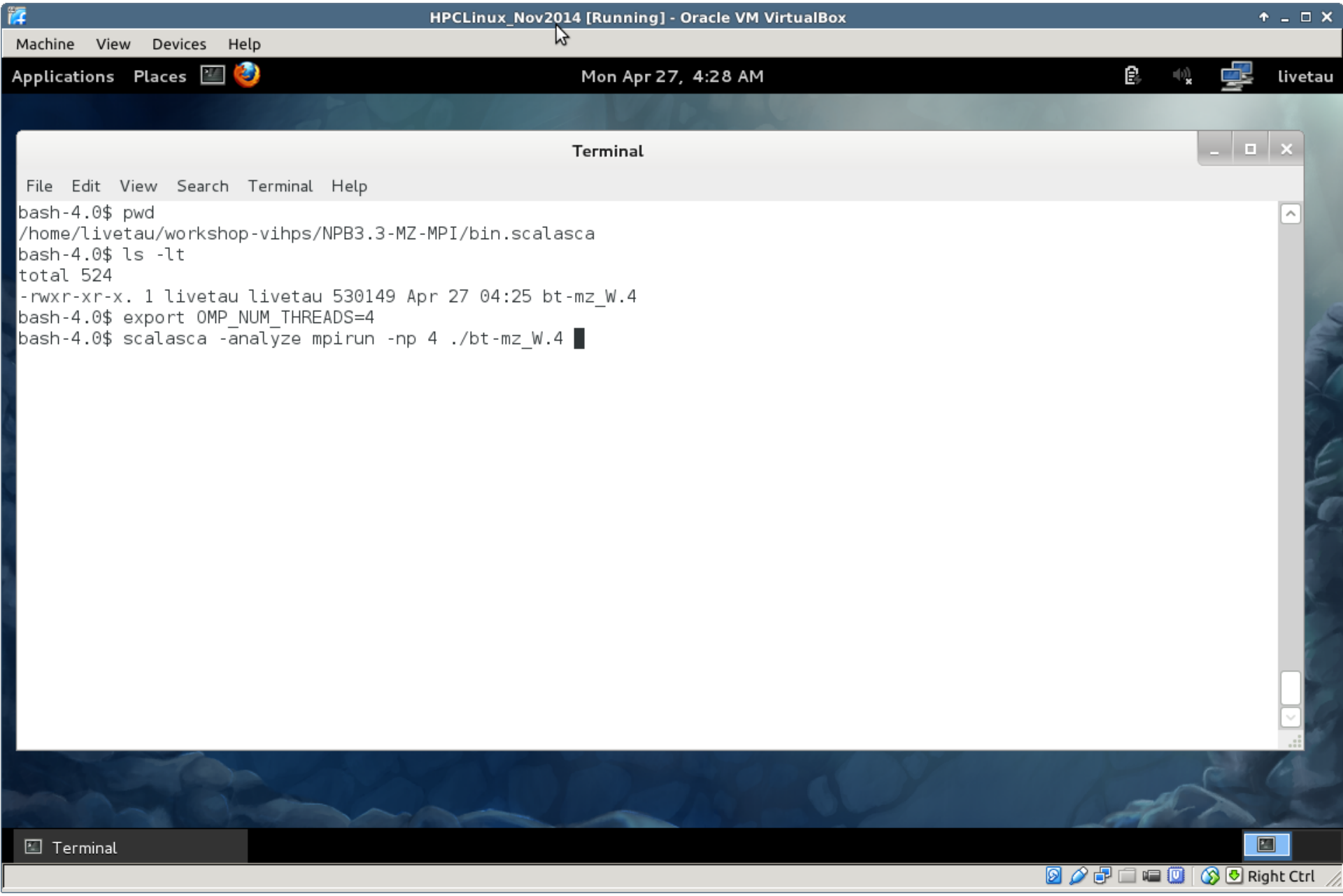
Demo #1: cd to folder: `cd /home/livetau/workshop-vihps/NPB3.3-MZ-MPI/config`

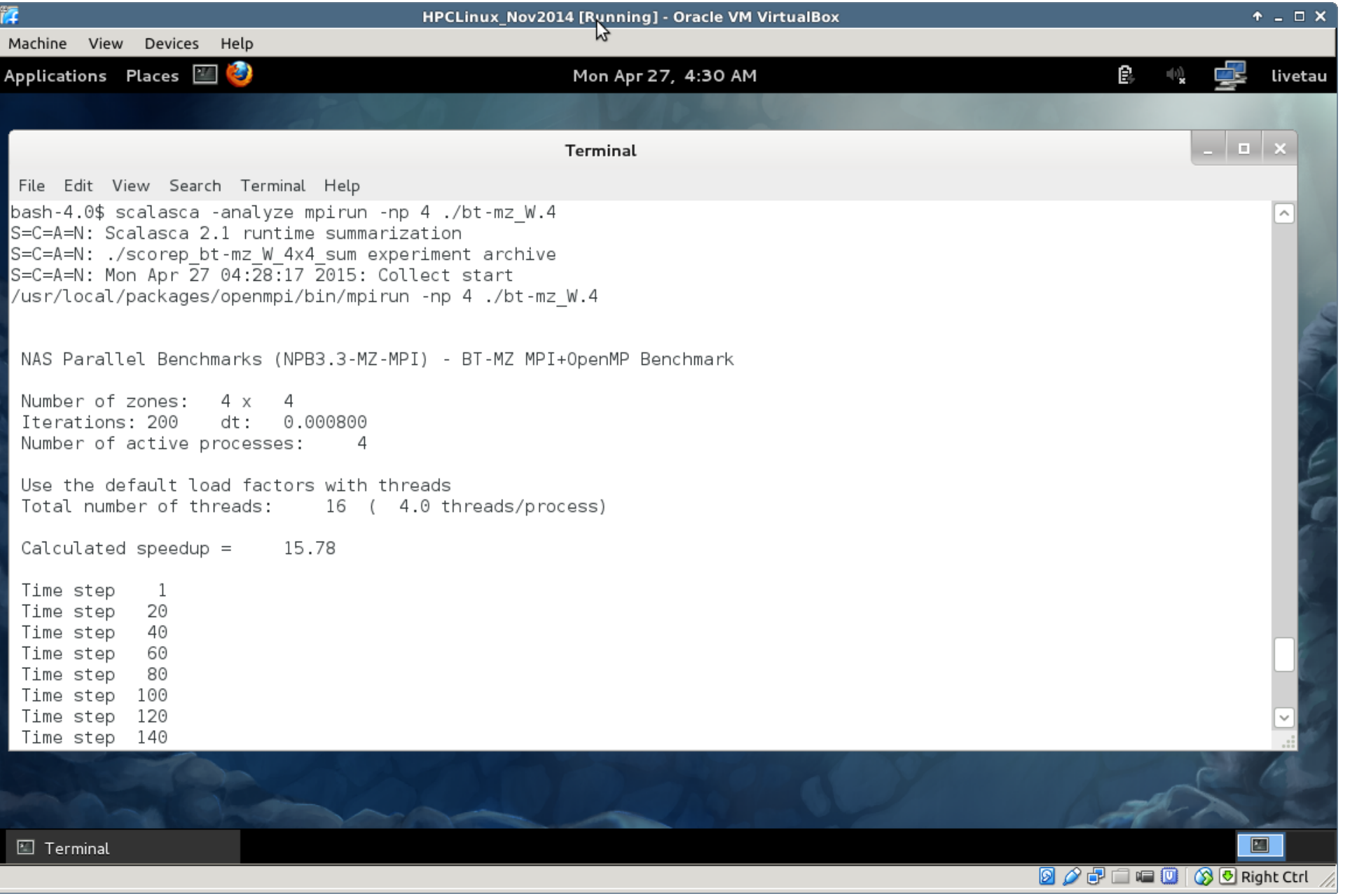


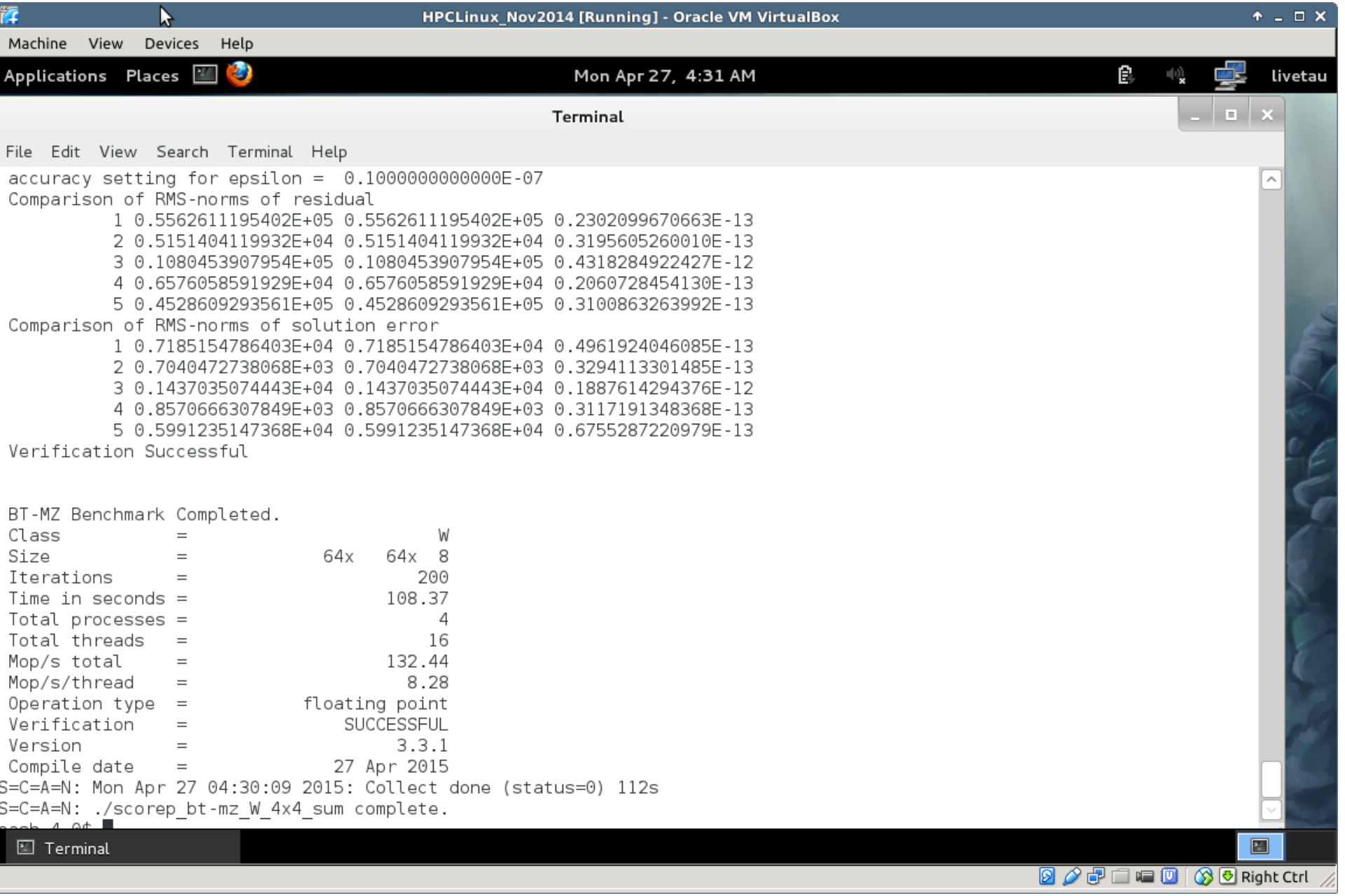
cd up one level: `cd ..`

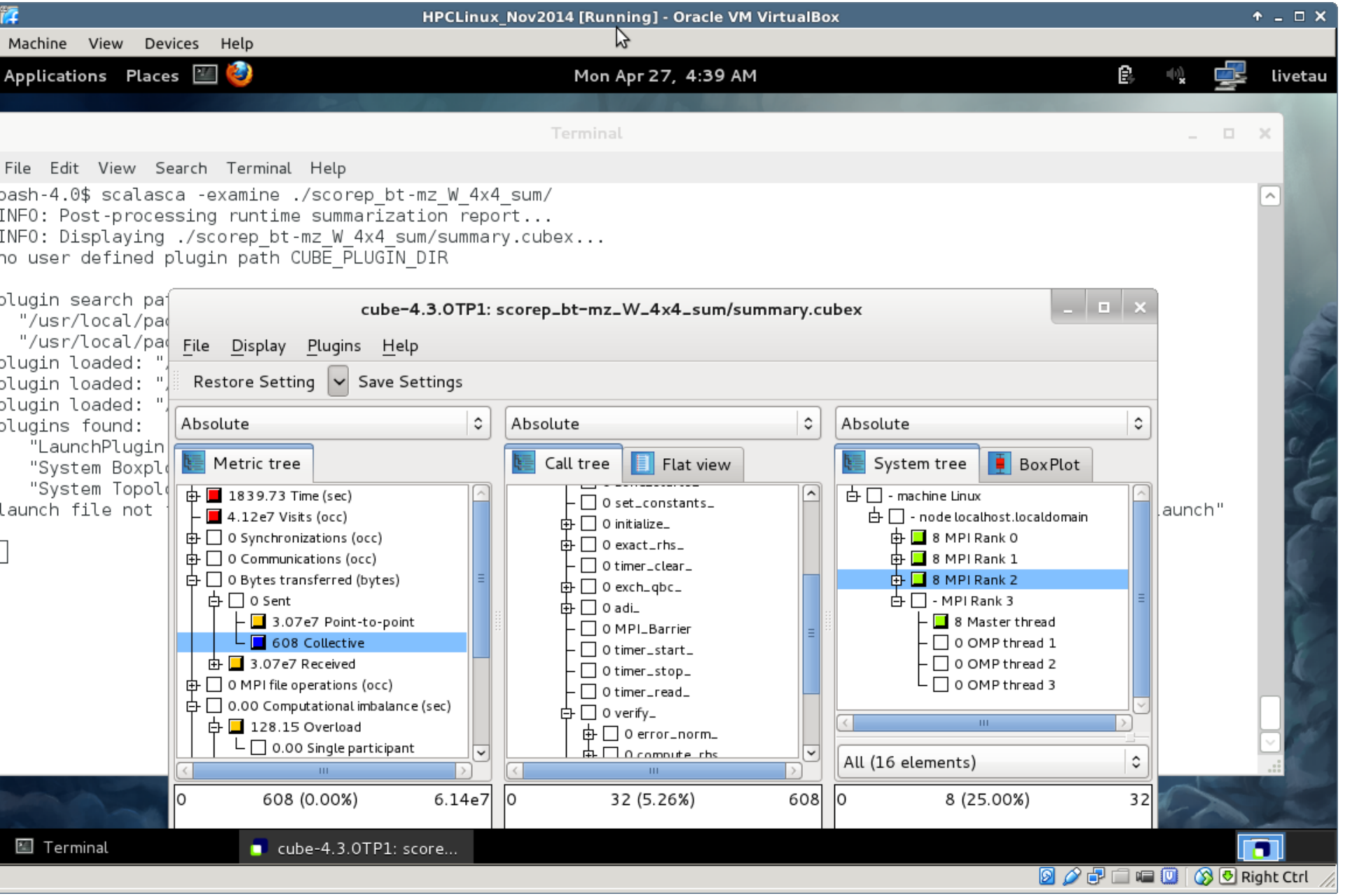


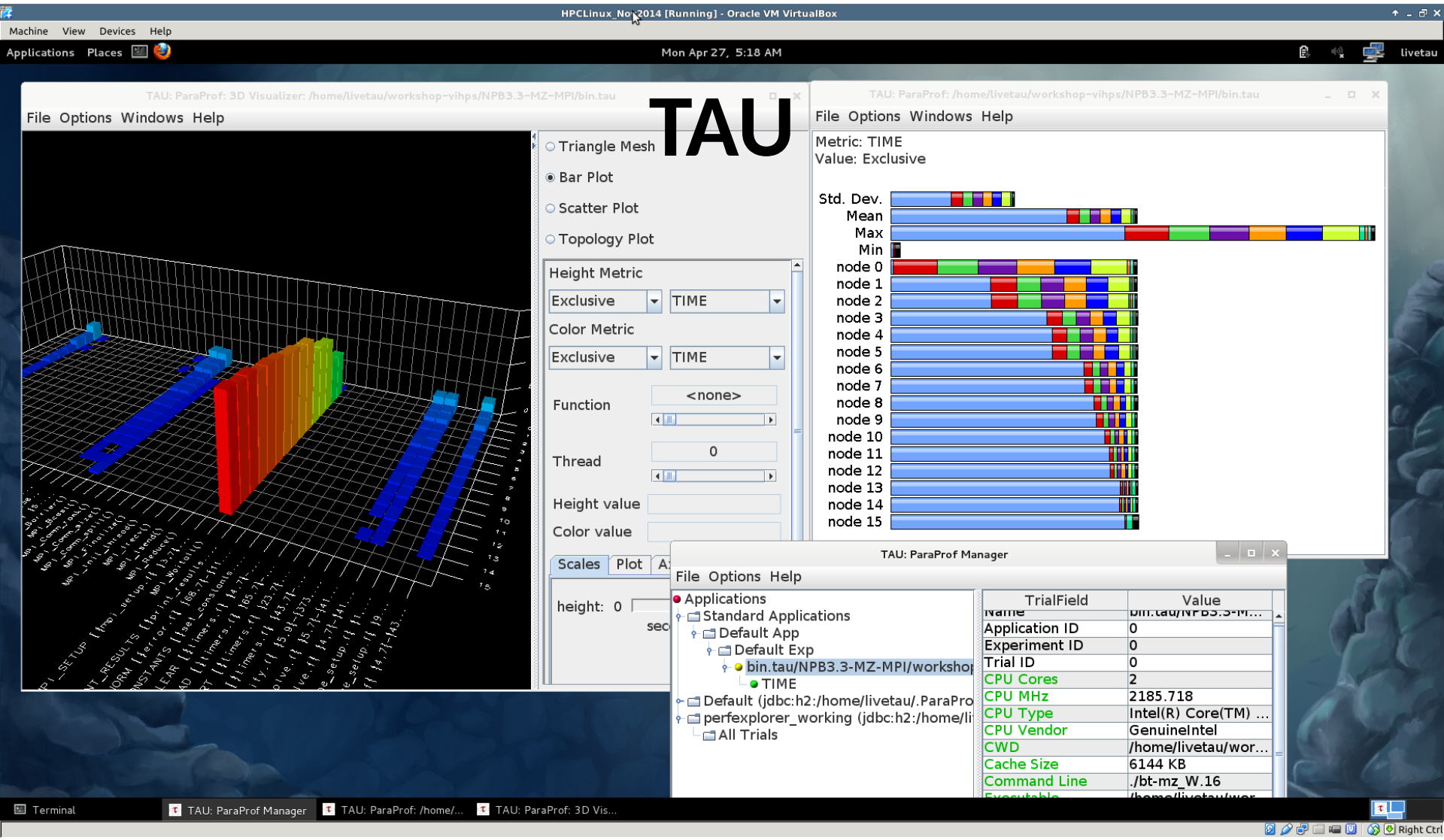

HPCLinux_Nov2014 [Running] - Oracle VM VirtualBox
Machine View Devices Help
Applications Places
Mon Apr 27, 4:26 AM
livetau
Terminal
File Edit View Search Terminal Help
-> /usr/local/packages/scorep/bin/scorep mpif77 -c -03 -fopenmp rhs.f
scalasca -instrument mpif77 -c -03 -fopenmp zone_setup.f
-> /usr/local/packages/scorep/bin/scorep mpif77 -c -03 -fopenmp zone_setup.f
scalasca -instrument mpif77 -c -03 -fopenmp x_solve.f
-> /usr/local/packages/scorep/bin/scorep mpif77 -c -03 -fopenmp x_solve.f
scalasca -instrument mpif77 -c -03 -fopenmp y_solve.f
-> /usr/local/packages/scorep/bin/scorep mpif77 -c -03 -fopenmp y_solve.f
scalasca -instrument mpif77 -c -03 -fopenmp exch_qbc.f
-> /usr/local/packages/scorep/bin/scorep mpif77 -c -03 -fopenmp exch_qbc.f
scalasca -instrument mpif77 -c -03 -fopenmp solve_subs.f
-> /usr/local/packages/scorep/bin/scorep mpif77 -c -03 -fopenmp solve_subs.f
scalasca -instrument mpif77 -c -03 -fopenmp z_solve.f
-> /usr/local/packages/scorep/bin/scorep mpif77 -c -03 -fopenmp z_solve.f
scalasca -instrument mpif77 -c -03 -fopenmp add.f
-> /usr/local/packages/scorep/bin/scorep mpif77 -c -03 -fopenmp add.f
scalasca -instrument mpif77 -c -03 -fopenmp error.f
-> /usr/local/packages/scorep/bin/scorep mpif77 -c -03 -fopenmp error.f
scalasca -instrument mpif77 -c -03 -fopenmp verify.f
-> /usr/local/packages/scorep/bin/scorep mpif77 -c -03 -fopenmp verify.f
scalasca -instrument mpif77 -c -03 -fopenmp mpi_setup.f
-> /usr/local/packages/scorep/bin/scorep mpif77 -c -03 -fopenmp mpi_setup.f
cd ../common; scalasca -instrument mpif77 -c -03 -fopenmp print_results.f
-> /usr/local/packages/scorep/bin/scorep mpif77 -c -03 -fopenmp print_results.f
cd ../common; scalasca -instrument mpif77 -c -03 -fopenmp timers.f
-> /usr/local/packages/scorep/bin/scorep mpif77 -c -03 -fopenmp timers.f
scalasca -instrument mpif77 -03 -fopenmp -o ../bin.scalasca/bt-mz_W.4 bt_scorep_user.o initialize.o exact_solut
Terminal
Right Ctrl


HPCLinux_Nov2014 [Running] - Oracle VM VirtualBox
Machine View Devices Help
Applications Places
Mon Apr 27, 4:28 AM
livetau
Terminal
File Edit View Search Terminal Help
bash-4.0$ pwd
/home/livetau/workshop-vihps/NPB3.3-MZ-MPI/bin.scalasca
bash-4.0$ ls -lt
total 524
-rwxr-xr-x. 1 livetau livetau 530149 Apr 27 04:25 bt-mz_W.4
bash-4.0$ export OMP_NUM_THREADS=4
bash-4.0$ scalasca -analyze mpirun -np 4 ./bt-mz_W.4
Terminal
Right Ctrl


HPCLinux_Nov2014 [Running] - Oracle VM VirtualBox
Machine View Devices Help
Applications Places
Mon Apr 27, 4:30 AM
livetau
Terminal
File Edit View Search Terminal Help
bash-4.0$ scalasca -analyze mpirun -np 4 ./bt-mz_W.4
S=C=A=N: Scalasca 2.1 runtime summarization
S=C=A=N: ./scorep_bt-mz_W_4x4_sum experiment archive
S=C=A=N: Mon Apr 27 04:28:17 2015: Collect start
/usr/local/packages/openmpi/bin/mpirun -np 4 ./bt-mz_W.4
NAS Parallel Benchmarks (NPB3.3-MZ-MPI) - BT-MZ MPI+OpenMP Benchmark
Number of zones: 4 x 4
Iterations: 200 dt: 0.000800
Number of active processes: 4
Use the default load factors with threads
Total number of threads: 16 (4.0 threads/process)
Calculated speedup = 15.78
Time step 1
Time step 20
Time step 40
Time step 60
Time step 80
Time step 100
Time step 120
Time step 140
Terminal
Right Ctrl


HPCLinux_Nov2014 [Running] - Oracle VM VirtualBox
Machine View Devices Help
Applications Places
Mon Apr 27, 4:31 AM
livetau
Terminal
File Edit View Search Terminal Help
accuracy setting for epsilon = 0.1000000000000E-07
Comparison of RMS-norms of residual
1 0.5562611195402E+05 0.5562611195402E+05 0.2302099670663E-13
2 0.5151404119932E+04 0.5151404119932E+04 0.3195605260010E-13
3 0.1080453907954E+05 0.1080453907954E+05 0.4318284922427E-12
4 0.6576058591929E+04 0.6576058591929E+04 0.2060728454130E-13
5 0.4528609293561E+05 0.4528609293561E+05 0.3100863263992E-13
Comparison of RMS-norms of solution error
1 0.7185154786403E+04 0.7185154786403E+04 0.4961924046085E-13
2 0.7040472738068E+03 0.7040472738068E+03 0.3294113301485E-13
3 0.1437035074443E+04 0.1437035074443E+04 0.1887614294376E-12
4 0.8570666307849E+03 0.8570666307849E+03 0.3117191348368E-13
5 0.5991235147368E+04 0.5991235147368E+04 0.6755287220979E-13
Verification Successful
BT-MZ Benchmark Completed.
Class = W
Size = 64x 64x 8
Iterations = 200
Time in seconds = 108.37
Total processes = 4
Total threads = 16
Mop/s total = 132.44
Mop/s/thread = 8.28
Operation type = floating point
Verification = SUCCESSFUL
Version = 3.3.1
Compile date = 27 Apr 2015
S=C=A=N: Mon Apr 27 04:30:09 2015: Collect done (status=0) 112s
S=C=A=N: ./scorep_bt-mz_W_4x4_sum complete.
Terminal
Right Ctrl


HPCLinux_Nov2014 [Running] - Oracle VM VirtualBox
Machine View Devices Help
Applications Places
Mon Apr 27, 4:39 AM
livetau
Terminal
File Edit View Search Terminal Help
bash-4.0$ scalasca -examine ./scorep_bt-mz_W_4x4_sum/
INFO: Post-processing runtime summarization report...
INFO: Displaying ./scorep_bt-mz_W_4x4_sum/summary.cubex...
no user defined plugin path CUBE_PLUGIN_DIR
cube-4.3.0TP1: scorep_bt-mz_W_4x4_sum/summary.cubex
File Display Plugins Help
Restore Setting
Save Settings
Absolute
Absolute
Absolute
Metric tree
Call tree
Flat view
System tree
BoxPlot
1839.73 Time (sec)
4.12e7 Visits (occ)
0 Synchronizations (occ)
0 Communications (occ)
0 Bytes transferred (bytes)
0 Sent
3.07e7 Point-to-point
608 Collective
3.07e7 Received
0 MPI file operations (occ)
0.00 Computational imbalance (sec)
128.15 Overload
0.00 Single participant
0 set_constants_
0 initialize_
0 exact_rhs_
0 timer_clear_
0 exch_qbc_
0 adi_
0 MPI_Barrier
0 timer_start_
0 timer_stop_
0 timer_read_
0 verify_
0 error_norm_
- machine Linux
- node localhost.localdomain
8 MPI Rank 0
8 MPI Rank 1
8 MPI Rank 2
- MPI Rank 3
8 Master thread
0 OMP thread 1
0 OMP thread 2
0 OMP thread 3
All (16 elements)
0 608 (0.00%) 6.14e7
0 32 (5.26%) 608
0 8 (25.00%) 32
Terminal
cube-4.3.0TP1: score...
Right Ctrl

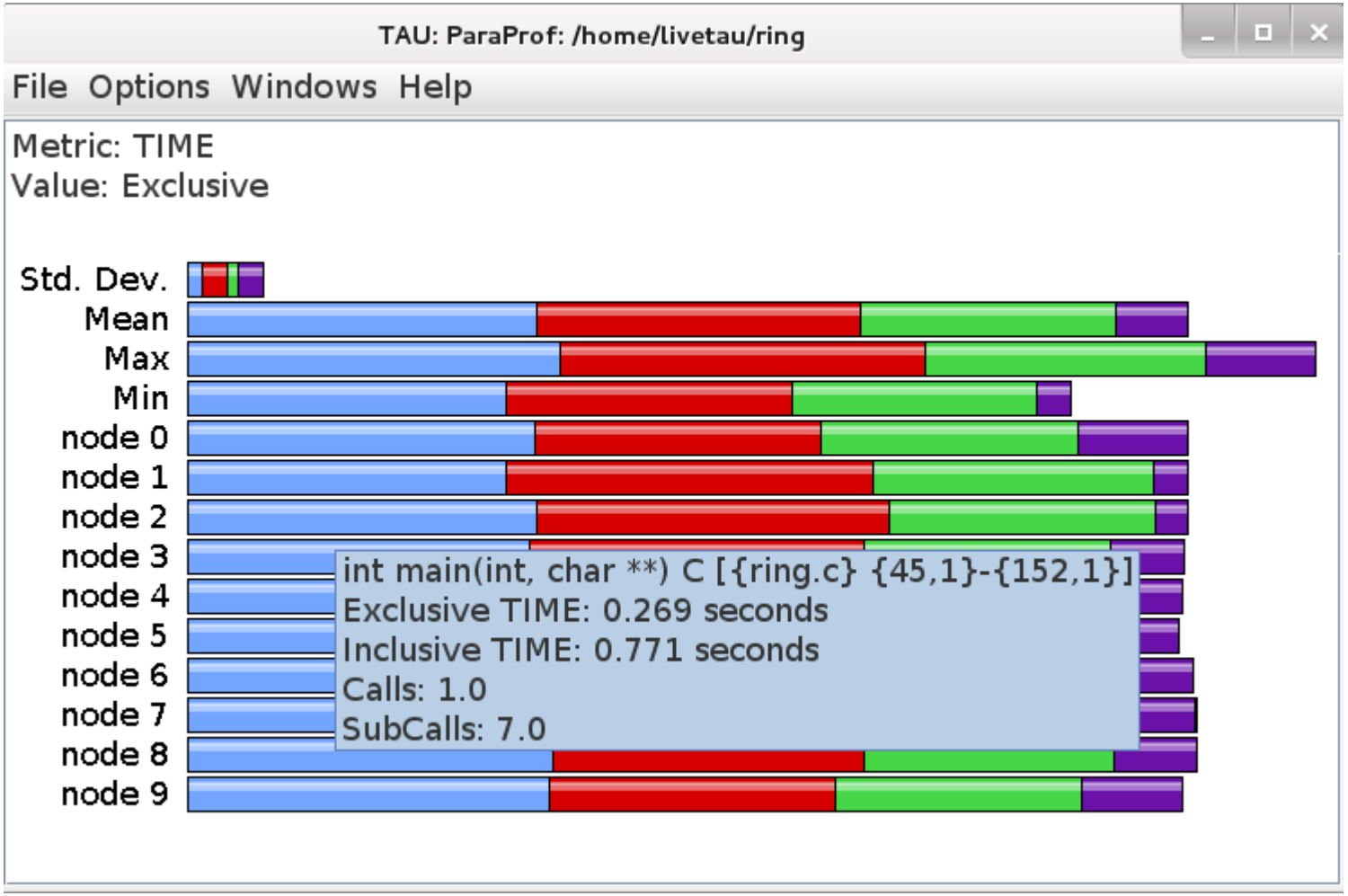

HPCLinux_Nov2014 [Running] - Oracle VM VirtualBox
Machine View Devices Help
Applications Places
Mon Apr 27, 5:18 AM
livetau
TAU
TAU: ParaProf: 3D Visualizer: /home/livetau/workshop-vihps/NPB3.3-MZ-MPI/bin.tau
File Options Windows Help
Triangle Mesh
Bar Plot
Scatter Plot
Topology Plot
Height Metric
Exclusive
TIME
Color Metric
Exclusive
TIME
Function
<none>
Thread
0
Height value
Color value
Scales
Plot
height: 0
TAU: ParaProf: /home/livetau/workshop-vihps/NPB3.3-MZ-MPI/bin.tau
File Options Windows Help
Metric: TIME
Value: Exclusive
Std. Dev.
Mean
Max
Min
node 0
node 1
node 2
node 3
node 4
node 5
node 6
node 7
node 8
node 9
node 10
node 11
node 12
node 13
node 14
node 15
TAU: ParaProf Manager
File Options Help
Applications
Standard Applications
Default App
Default Exp
bin.tau/NPB3.3-MZ-MPI/workshop
TIME
Default (jdbc:h2:/home/livetau/.ParaPro
perfexplorer_working (jdbc:h2:/home/li
All Trials
TrialField
Value
Application ID
0
Experiment ID
0
Trial ID
0
CPU Cores
2
CPU MHz
2185.718
CPU Type
Intel(R) Core(TM) ...
CPU Vendor
GenuineIntel
CWD
/home/livetau/wor...
Cache Size
6144 KB
Command Line
./bt-mz_W.16
Terminal
TAU: ParaProf Manager
TAU: ParaProf: /home/...
TAU: ParaProf: 3D Vis...
Right Ctrl

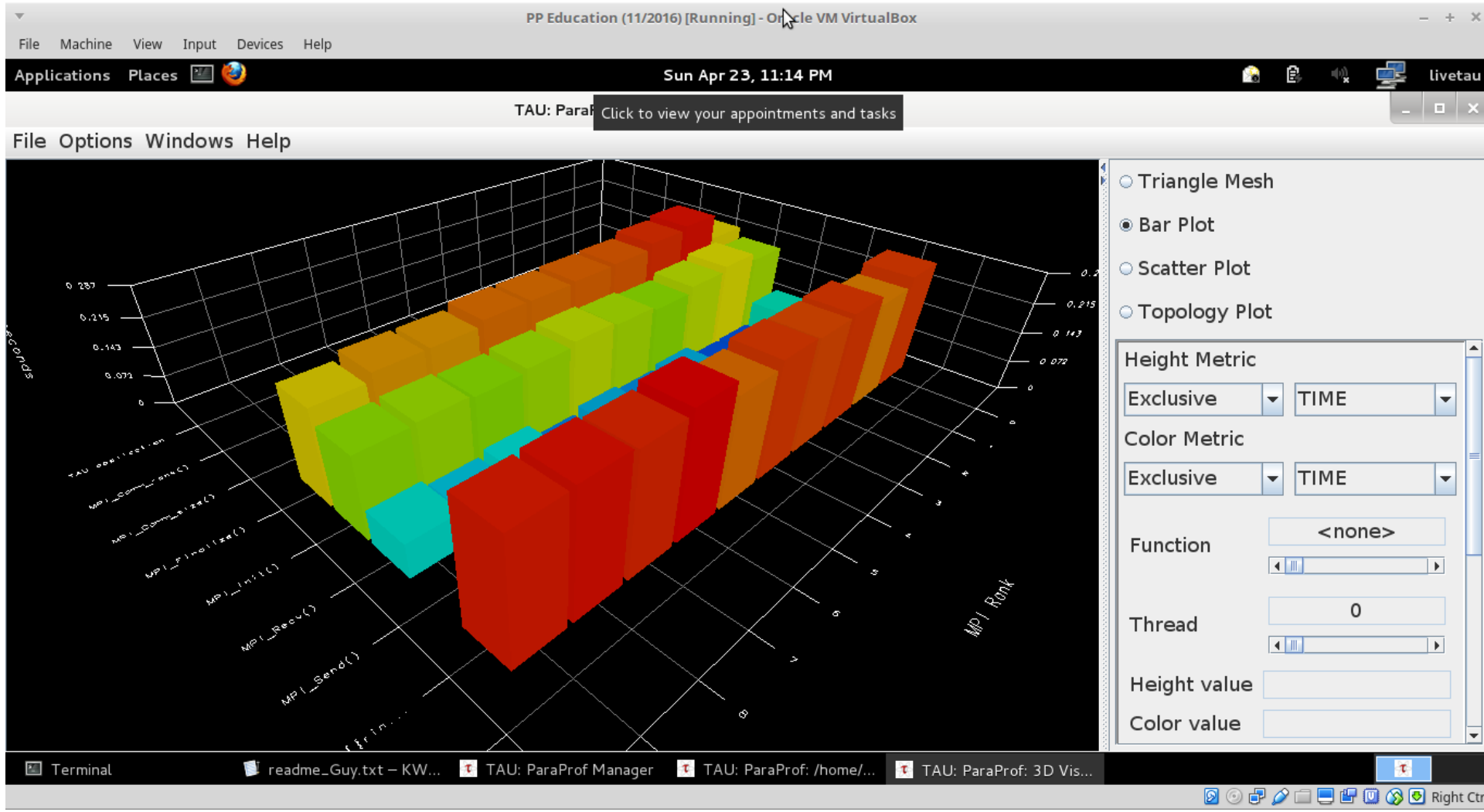
Demo #2: Using TAU and jumpshot in 11 steps

Demo program ~/ring/ring.c

~~step #1: `module load tau`~~
step #2: `export TAU_TRACE=1`
step #3: `export TAU_PROFILE=1`
step #4: `export TAU_COMM_MATRIX=1`
step #5: `export TAU_MAKEFILE=/usr/local/packages/tau-2.25.1/x86_64/Makefile.tau-papi-mpi-pdt` (מה שפורסם במודל)
step #6: compile: instead of `mpicc` type: `tau_cc.sh -o ring ./ring.c`
step #7: execute: `mpirun -n 10 ./ring`
step #8: view profiling: `paraprof`
step #9: merge log files: `tau_treemerge`
step #10: convert to slog format: `tau2slog2 tau.trc tau.edf -o tau.slog2`
step #11: view tracing in jumpshot: `jumpshot tau.slog2`

paraprof


TAU: ParaProf: /home/livetau/ring
File Options Windows Help
Metric: TIME
Value: Exclusive
Std. Dev.
Mean
Max
Min
node 0
node 1
node 2
node 3
node 4
node 5
node 6
node 7
node 8
node 9
int main(int, char **) C [{ring.c} {45,1}-{152,1}]
Exclusive TIME: 0.269 seconds
Inclusive TIME: 0.771 seconds
Calls: 1.0
SubCalls: 7.0


PP Education (11/2016) [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places
Sun Apr 23, 11:14 PM
livetau
TAU: ParaP
Click to view your appointments and tasks
File Options Windows Help
Triangle Mesh
Bar Plot
Scatter Plot
Topology Plot
Height Metric
Exclusive
TIME
Color Metric
Exclusive
TIME
Function
<none>
Thread
0
Height value
Color value
0.287
0.215
0.143
0.072
0
Seconds
MPI_Send()
MPI_Recv()
MPI_Init()
MPI_Finalize()
MPI_Comm_size()
MPI_Comm_rank()
TAU application
MPI Rank
Terminal
readme_Guy.txt – KW...
TAU: ParaProf Manager
TAU: ParaProf: /home/...
TAU: ParaProf: 3D Vis...
Right Ctr

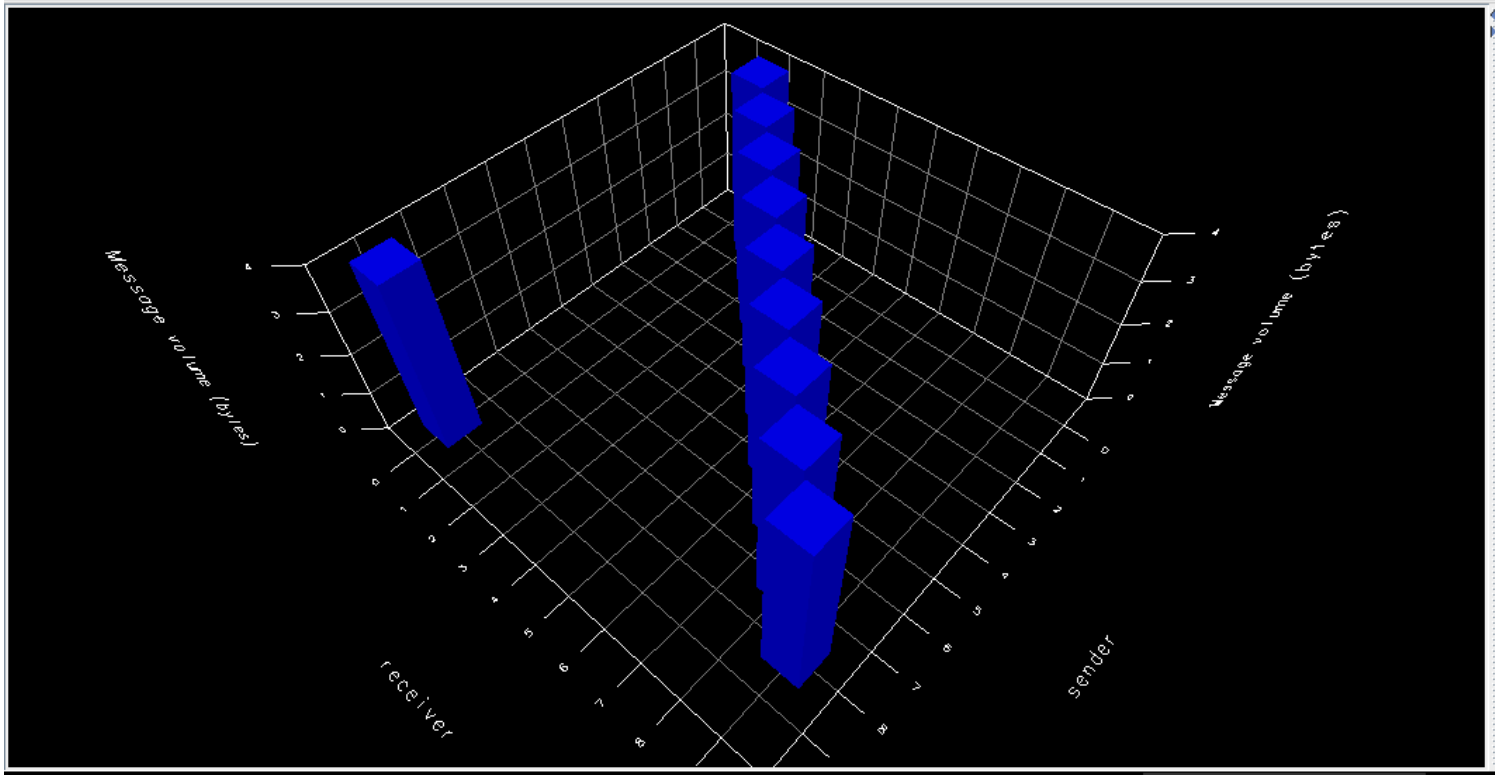
PP Education (11/2016) [Running] - Oracle VM VirtualBox

File Machine View Input Devices Help

Applications Places Sun Apr 23, 11:15 PM livetau

TAU: ParaProf: 3D Communication Matrix: /home/livetau/ring

File Windows Help



All Paths

Height Value:

Message volume (bytes)

Color Value:

Max message size (bytes)

Sender 0

Receiver 0

Scales Plot Axes Colo...

height: 0 4 bytes

color: 0 4 bytes

Terminal readme_Guy.txt – KW... TAU: ParaProf Manager TAU: ParaProf: /home/... TAU: ParaProf: 3D Co...

Right Ctr

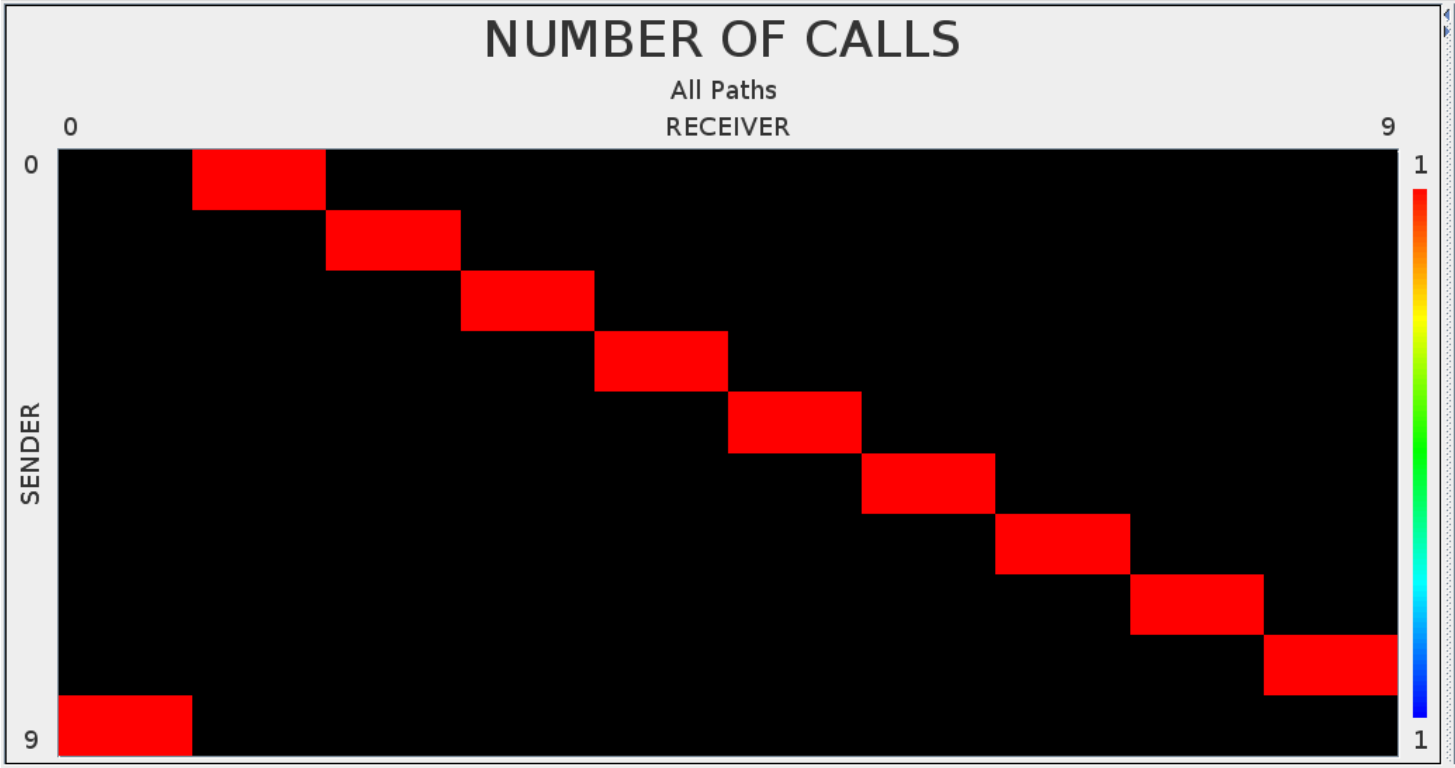
PP Education (11/2016) [Running] - Oracle VM VirtualBox

File Machine View Input Devices Help

Applications Places Sun Apr 23, 11:16 PM livetau

Message Size Heat Maps

File Windows Help



DISPLAY OPTIONS

Callpath:

All Paths

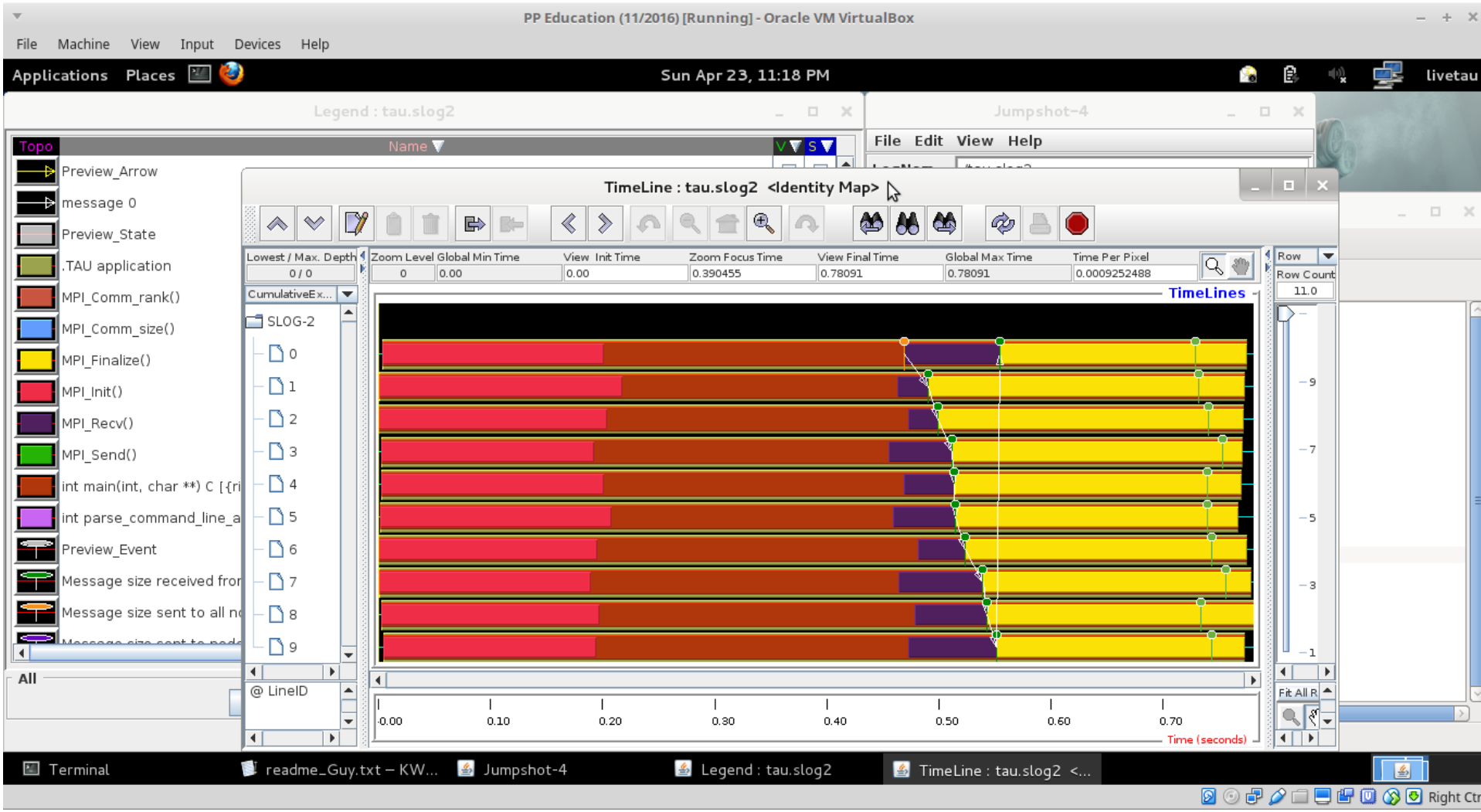
Dataset:

NUMBER OF CALLS

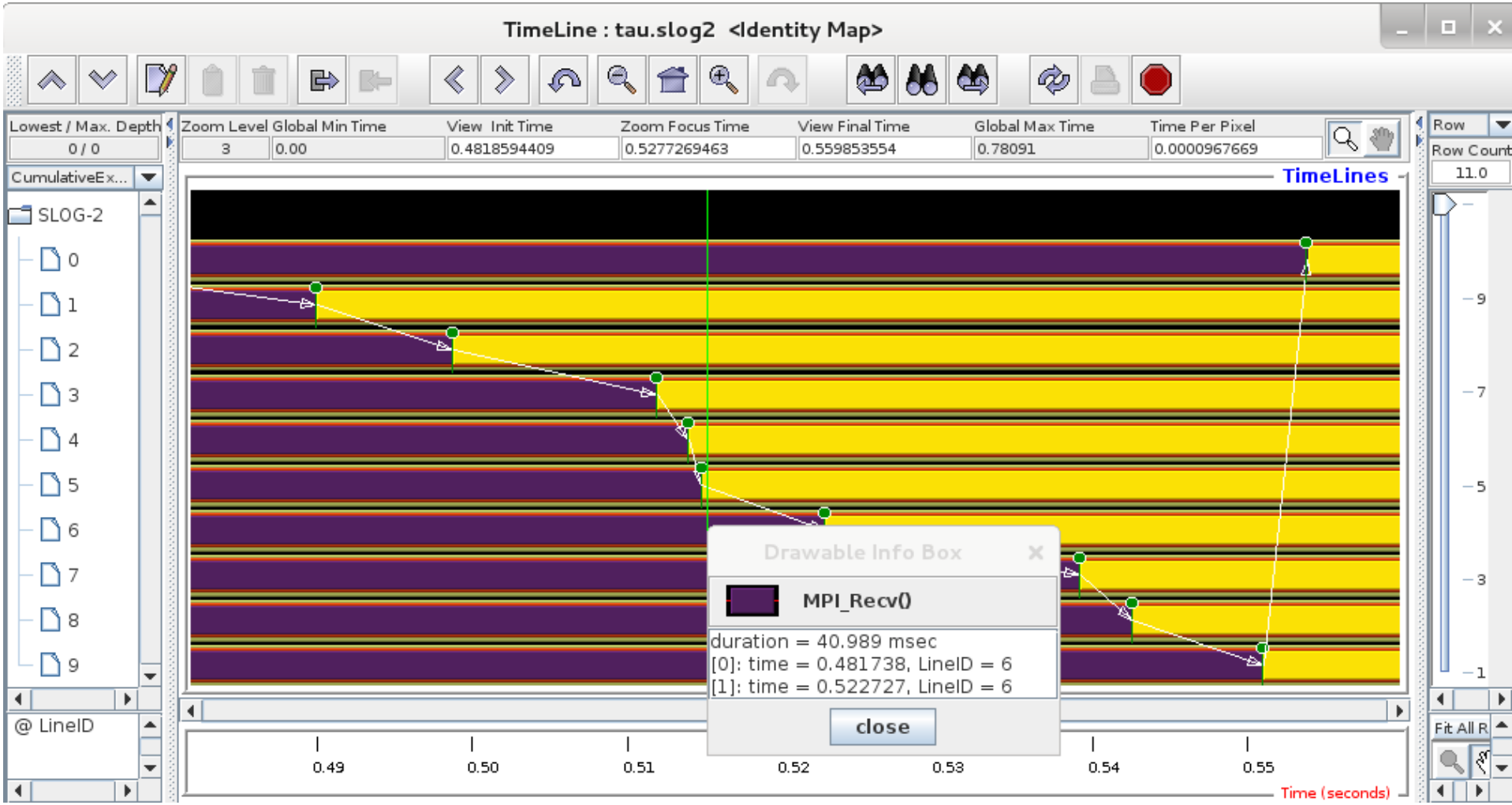
Terminal readme_Guy.txt – KW... TAU: ParaProf Manager TAU: ParaProf: /home/... Message Size Heat M...

Right Ctr

jumpshot


PP Education (11/2016) [Running] - Oracle VM VirtualBox
File Machine View Input Devices Help
Applications Places
Sun Apr 23, 11:18 PM
livetau
Legend : tau.slog2
Topo
Name
Preview_Arrow
message 0
Preview_State
.TAU application
MPI_Comm_rank()
MPI_Comm_size()
MPI_Finalize()
MPI_Init()
MPI_Recv()
MPI_Send()
int main(int, char **) C [{ri
int parse_command_line_a
Preview_Event
Message size received fror
Message size sent to all n
All
Jumpshot-4
File Edit View Help
TimeLine : tau.slog2 <Identity Map>
Lowest / Max. Depth
0 / 0
Zoom Level
Global Min Time
View Init Time
Zoom Focus Time
View Final Time
Global Max Time
Time Per Pixel
0
0.00
0.00
0.390455
0.78091
0.78091
0.0009252488
CumulativeEx...
SLOG-2
0
1
2
3
4
5
6
7
8
9
@ LineID
TimeLines
Row
Row Count
11.0
Fit All R
0.00
0.10
0.20
0.30
0.40
0.50
0.60
0.70
Time (seconds)
Terminal
readme_Guy.txt – KW...
Jumpshot-4
Legend : tau.slog2
TimeLine : tau.slog2 <...
Right Ctr

Zoom in


TimeLine : tau.slog2 <Identity Map>
Lowest / Max. Depth
0 / 0
Zoom Level
3
Global Min Time
0.00
View Init Time
0.4818594409
Zoom Focus Time
0.5277269463
View Final Time
0.559853554
Global Max Time
0.78091
Time Per Pixel
0.0000967669
Row
Row Count
11.0
TimeLines
CumulativeEx...
SLOG-2
0
1
2
3
4
5
6
7
8
9
@ LineID
Drawable Info Box
MPI_Recv()
duration = 40.989 msec
[0]: time = 0.481738, LineID = 6
[1]: time = 0.522727, LineID = 6
close
0.49
0.50
0.51
0.52
0.53
0.54
0.55
Time (seconds)
Fit All R

!

במקרה ונופלת ריצה שקומפלה עם הסקריפט tau_cc.sh
אך היא רצה בהצלחה ללא tau עם mpirun
נסו את התחליף הבא:

1) Compile with:
mpicc -g -o your_executable ./your_source.c

2) execute with:
mpirun -np N tau_exec -ebs ./your_executable

replace N with the desired number of MPI tasks

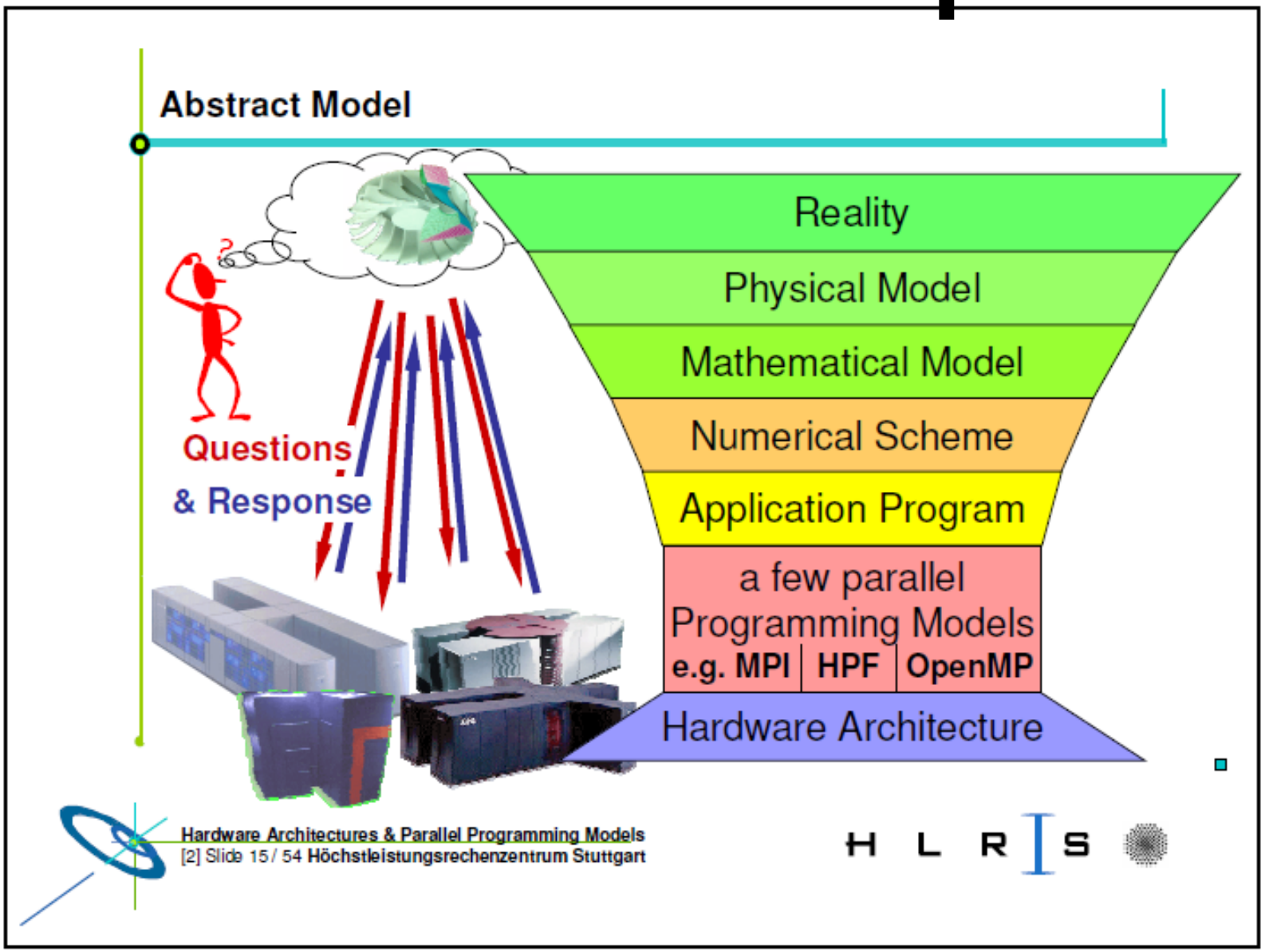
To sum up:

In this course we are using both **Jumpshot(*)** and **Scalasca** as our profiling and tracing tools

Both tools are installed on the BGU-VM & HPC_VM

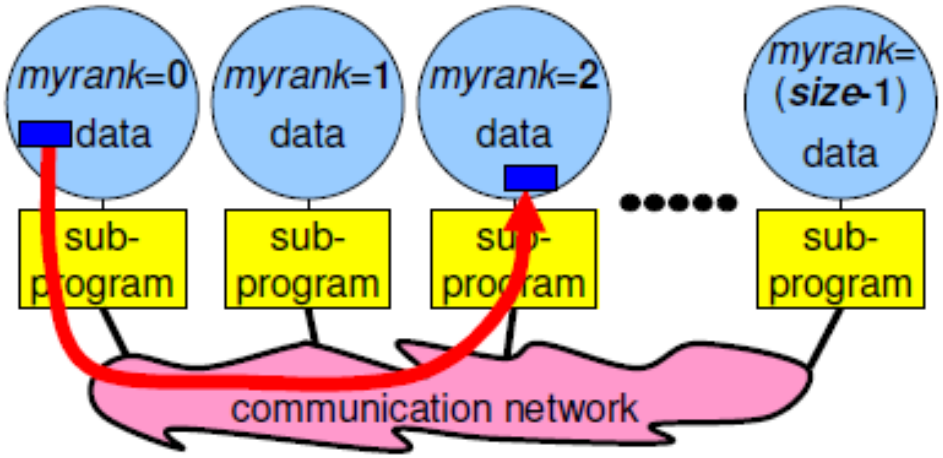
(*) Jumpshot is part of the TAU package

נקודת מבט מגבוה



Message Passing Program Paradigm – MPI, I.

- Each processor in a message passing program runs a ***sub-program***
  - written in a conventional sequential language, e.g., C or Fortran,
  - typically the same on each processor (SPMD)
- All work and data distribution is based on value of ***myrank***
  - returned by special library routine
- Communication via special send & receive routines (***message passing***)



$$\int_0^1 f(x)dx = \int_0^1 \frac{4}{1+x^2}dx = \pi .$$

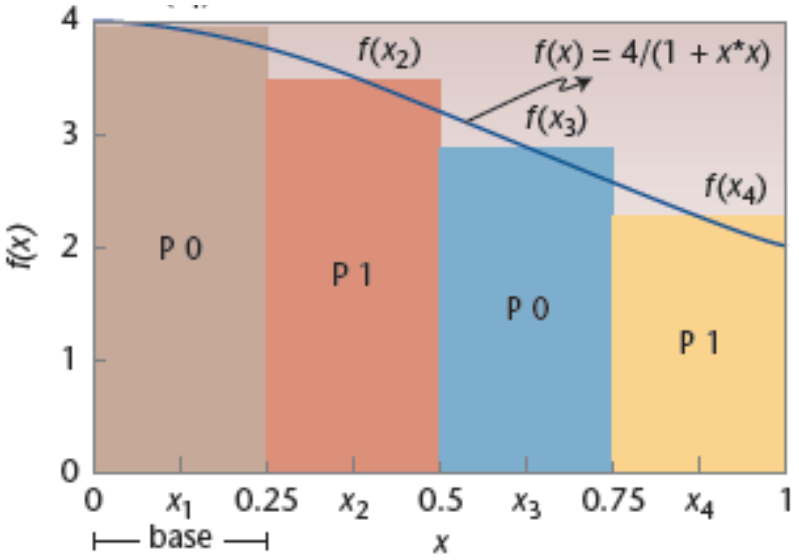


Figure 5. The `cpi.c` program's principle for computing π with multiple processes. Here, the approximate value of π is the summation of four rectangles' areas calculated by two processes.

Figures in this slide are taken from: **"Ch MPI: Interpretive Parallel Computing in C",** By Yu-Cheng Chou *Chung Yuan Christian University, Taiwan.* Stephen S. Nestinger, *Worcester Polytechnic Institute.* Harry H. Cheng, *University of California, Davis*

Computing in Science and Engineering, March/April 2010

cpi_mathematica.nb *

In[7]:=

Z:\cpi_mathematica.nb

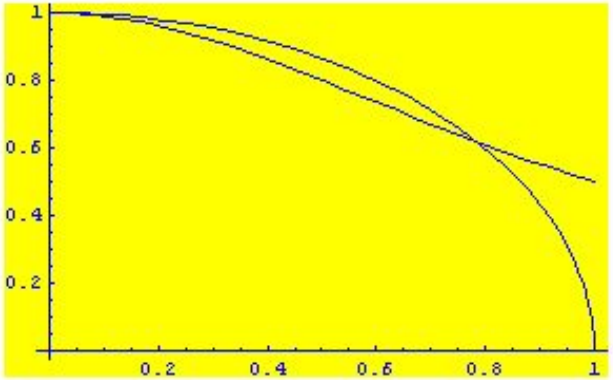
```
Integrate[Sqrt[1 - x^2], {x, 0, 1}]
```

Out[7]= $\frac{\pi}{4}$

```
Integrate[1 / (1 + x^2), {x, 0, 1}]
```

Out[8]= $\frac{\pi}{4}$

```
Plot[{Sqrt[1 - x^2], 1 / (1 + x^2)}, {x, 0, 1},
 Background → RGBColor[1, 1, 0]]
```



Out[27]= - Graphics -

100%

The Sage Notebook
Version 4.2

gtelzur | Toggle | Home | Published | Log | Settings | Re

PI Integration for PP course

last edited on November 08, 2009 03:02 AM by gtelzur

Save

File... Action... Data... sage Typeset

Print Worksheet Edit Text

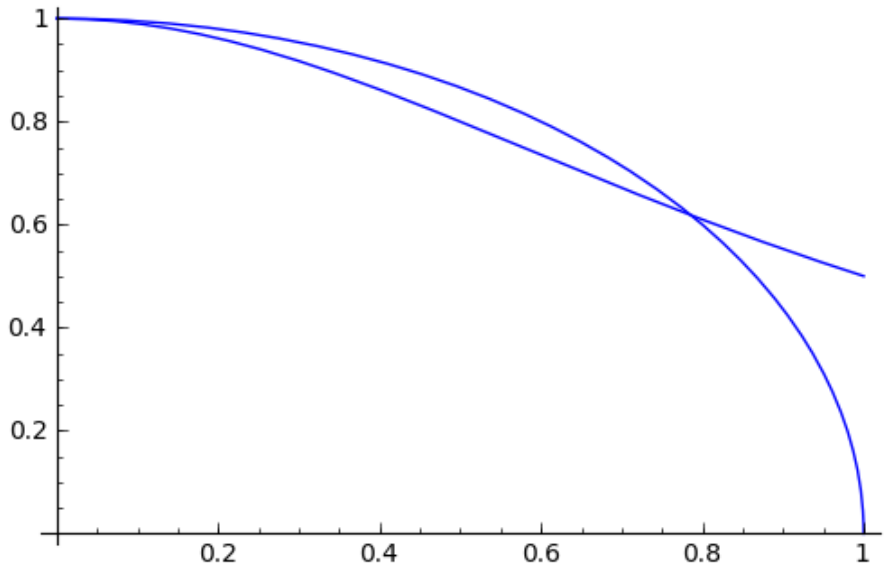
```
integral(sqrt(1-x^2),x,0,1);
```

```
    1/4*pi
```

```
integral(1/(1+x^2),x,0,1);
```

```
    1/4*pi
```

```
plot([sqrt(1-x^2),1/(1+x^2)],0,1)
```



http://www.sagenb.org

evaluate

Find: calc Next Previous Highlight all Match case Reached end of page, continued from top

That's it!