Ferramentas de Avaliação de Desempenho

Roteiro

Scalasca

https://www.scalasca.org

Visualizando: CubeGUI

- O CubeGUI pode ser baixado e instalado para visualizar os resultados obtidos com o Scalasca
- `https://www.scalasca.org/scalasca/software/cube-4.x/download.html`
- Binários prontos em "Supplementary packages for download (Comfort zone)"
- Resultados previamente obtidos no SDumont estão no arquivo **`profiling_scalasca_sequana.zip`** do repositório no GitHub:

```
git clone https://github.com/robertopsouto/ESD2024.git
ESD2024/sdbase/profiling_scalasca_sequana.zip
```

NPB: estudo de caso

```
$ cd profiling/NUMNODES-1/scorep_bt-mz_W_sum_MPI-1_OMP-1_JOBID-10407415
$ cube summary.cubex
```

Estudo de caso

NPB: perfil de desempenho

```
scorep_bt-mz_S_sum_MPI-1_OMP-1_JOBID-437607/
   profile.cubex
   summary.cubex
   scorep.score
   scorep.cfg
   scorep.log
   slurm-437607.out
```

Estudo de caso

NPB: perfil de desempenho

```
scorep_bt-mz_S_sum_MPI-1_OMP-1_JOBID-437607/
   profile.cubex     --> análise básica, a partir de dados coletados durante execução
   summary.cubex     --> análise mais detalhada
   scorep.score      --> relatório formato texto com a análise
   scorep.cfg        --> configuração da coleta de dados
   scorep.log        --> output da aplicação
   slurm-437607.out --> output do SLURM
```

Estudo de caso

NPB: perfil de desempenho

```
scorep_bt-mz_S_sum_MPI-1_OMP-1_JOBID-437607/
   profile.cubex
   summary.cubex
   scorep.score
   scorep.cfg
   scorep.log
   slurm-437607.out
```

Estudo de caso

```
$ cat slurm-437607.out

Cluster configuration:
===
Partition:  treinamento
Number of nodes:  1
Number of MPI processes:  1  ( 1  nodes)
Number of MPI processes per node:  1
Number of threads per MPI process:  1
NPB Benchmark:  bt-mz
Bechmark class problem:  S
scalasca 2.4 for GNU OpenMPI loaded
Compiled with openMPI 2.0.4.2 and GNU compilers Red Hat 4.8.5-36
S=C=A=N: Scalasca 2.4 runtime summarization
S=C=A=N: ./scorep_bt-mz_1x1_sum experiment archive
S=C=A=N: Tue Jan 28 14:17:02 2020: Collect start
/usr/bin/srun --resv-ports -n 1 /scratch/treinamento/professor/MC1-I/tools/scalasca/NPB3.3.1-MZ/NP
[1580231823.240921] [sdumont5000:73441:0]         mxm.c:196  MXM  WARN  The 'ulimit -s' on the sys
[1580231823.242634] [sdumont5000:73441:0]         mxm.c:196  MXM  WARN  The 'ulimit -s' on the sys


 NAS Parallel Benchmarks (NPB3.3-MZ-MPI) - BT-MZ MPI+OpenMP Benchmark

 Number of zones:   2 x   2
 Iterations:  60    dt:   0.010000
 Number of active processes:     1


 Use the default load factors with threads
```

Estudo de caso (cont.)

```
Total number of threads:      1   (  1.0 threads/process)

Calculated speedup =      1.00

Time step    1
Time step   20
Time step   40
Time step   60
Verification being performed for class S
accuracy setting for epsilon =  0.1000000000000E-07
Comparison of RMS-norms of residual
          1 0.1047687395830E+04 0.1047687395830E+04 0.1751386499571E-12
          2 0.9419911314792E+02 0.9419911314792E+02 0.1478425555772E-13
          3 0.2124737403068E+03 0.2124737403068E+03 0.9002435039286E-13
          4 0.1422173591794E+03 0.1422173591794E+03 0.3089634277625E-12
          5 0.1135441572375E+04 0.1135441572375E+04 0.3103895484466E-13
Comparison of RMS-norms of solution error
          1 0.1775416062982E+03 0.1775416062982E+03 0.1922618237923E-12
          2 0.1875540250835E+02 0.1875540250835E+02 0.1558955269742E-12
          3 0.3863334844506E+02 0.3863334844506E+02 0.1105356386074E-12
          4 0.2634713890362E+02 0.2634713890362E+02 0.3991337551951E-13
          5 0.1965566269675E+03 0.1965566269675E+03 0.2336704854379E-12
Verification Successful


BT-MZ Benchmark Completed.
Class           =                        S
Size            =             24x   24x  6
Iterations      =                       60
Time in seconds =                     0.35
```

Estudo de caso (cont.)

```
 Total processes =                      1
 Total threads   =                      1
 Mop/s total     =                1093.41
 Mop/s/thread    =                1093.41
 Operation type  =          floating point
 Verification    =              SUCCESSFUL
 Version         =                   3.3.1
 Compile date    =             21 Jan 2020

 Compile options:
    F77          = scalasca -instrument mpif77
    FLINK        = $(F77)
    F_LIB        = (none)
    F_INC        = (none)
    FFLAGS       = -O3 -fopenmp
    FLINKFLAGS   = $(FFLAGS)
    RAND         = (none)


 Please send all errors/feedbacks to:

 NPB Development Team
 npb@nas.nasa.gov


S=C=A=N: Tue Jan 28 14:17:03 2020: Collect done (status=0) 1s
S=C=A=N: ./scorep_bt-mz_1x1_sum complete.
INFO: Post-processing runtime summarization report...
/opt/bullxde/utils/scalasca/openmpi-gnu/scorep/bin/scorep-score -r ./scorep_bt-mz_S_sum_MPI-1_OMP-
INFO: Score report written to ./scorep_bt-mz_S_sum_MPI-1_OMP-1_JOBID-437607/scorep.score
```

Estudo de caso (cont.)

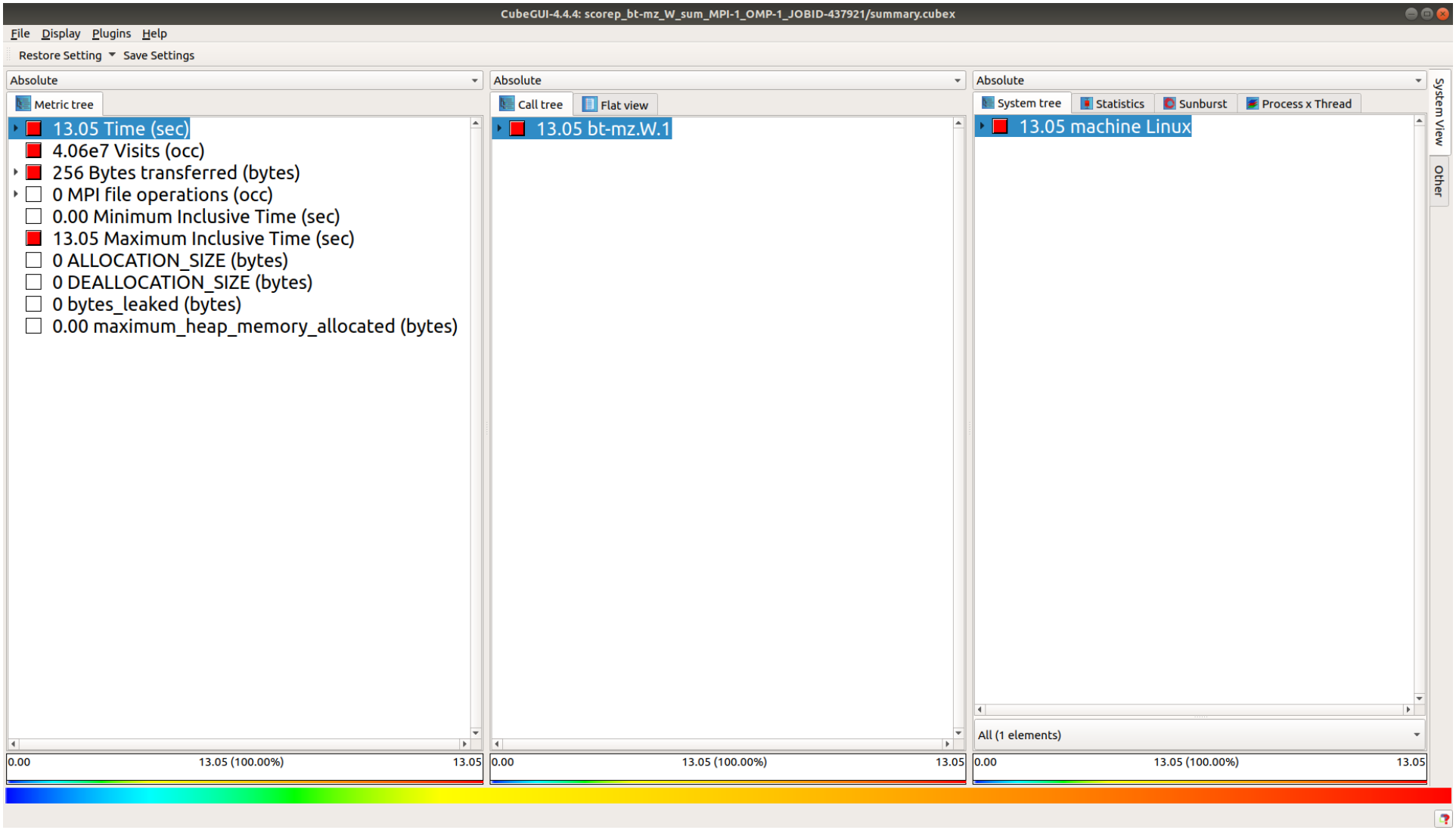
Estudo de caso

NPB: perfil de desempenho

```
scorep_bt-mz_W_sum_MPI-1_OMP-1_JOBID-437632/
   profile.cubex
   summary.cubex
   scorep.score
   scorep.cfg
   scorep.log
   slurm-437632.out
```

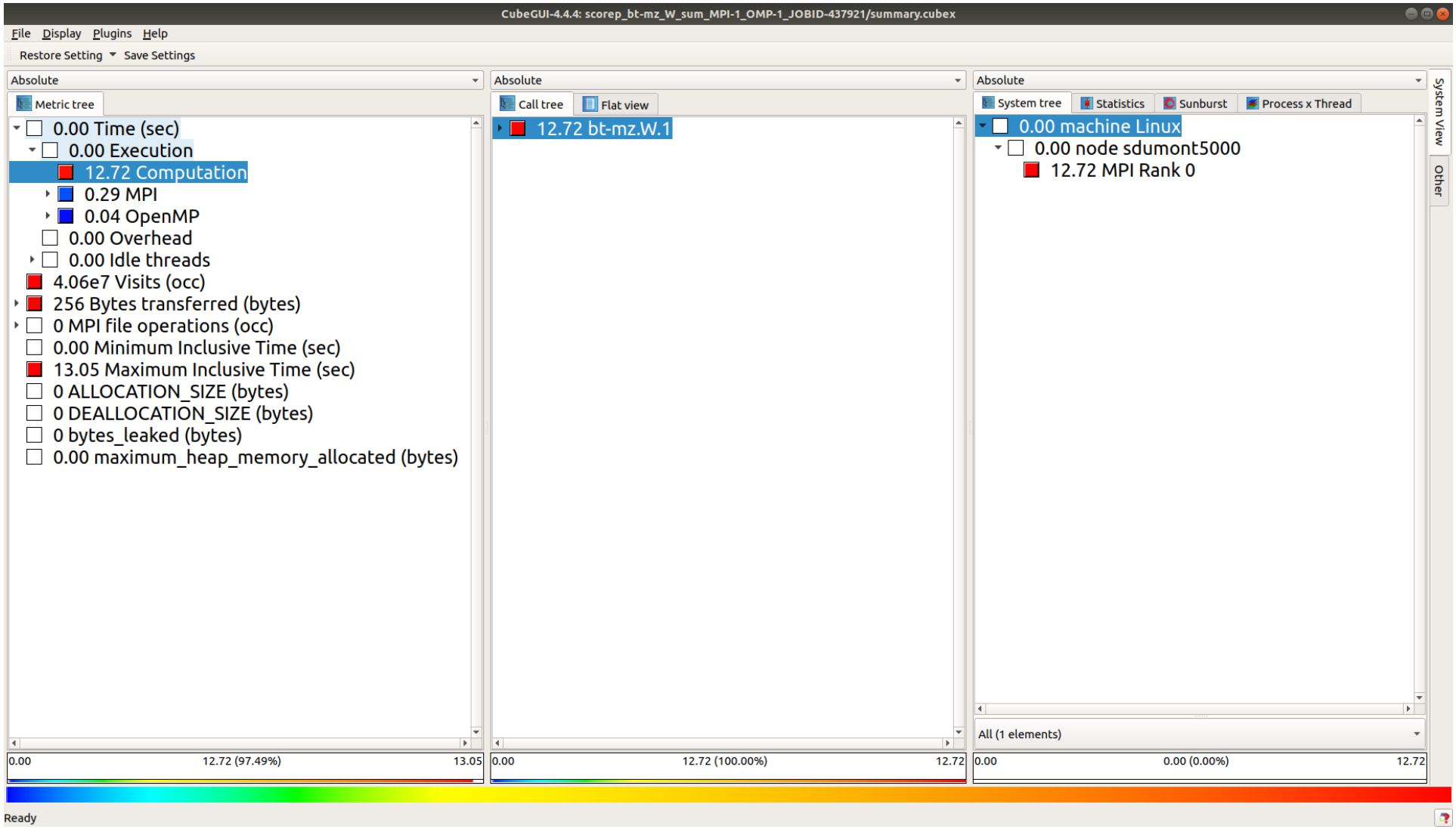
Visualizando: CubeGUI

`-nodes=1 -ntasks=1`



Visualizando: CubeGUI

-nodes=1 -ntasks=1



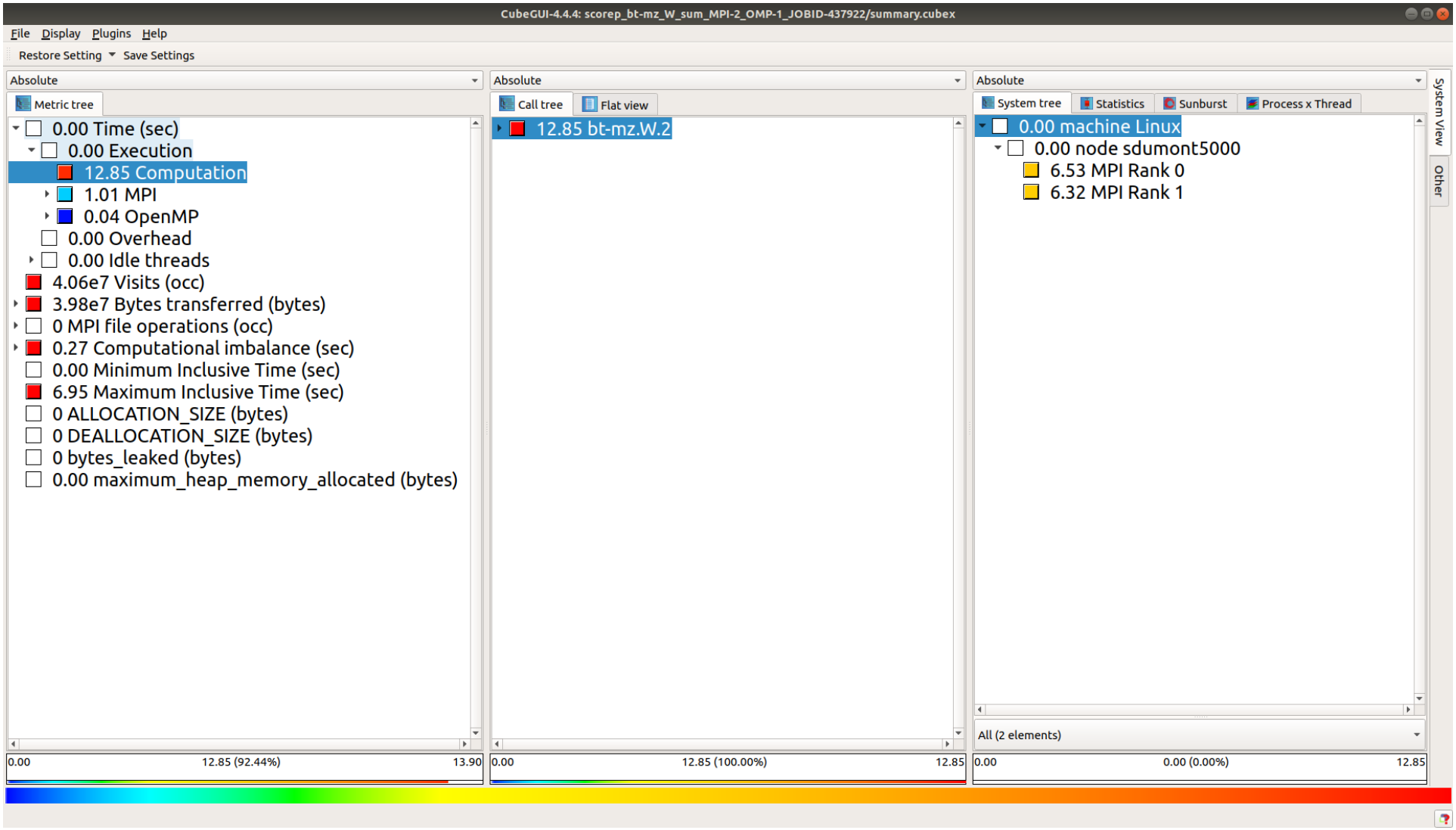
Visualizando: CubeGUI

divisão do tempo de processamento

- Tempo de computação: 12.72s
- Tempo de MPI: 0.29s
- Tempo de OpenMP: 0.04s

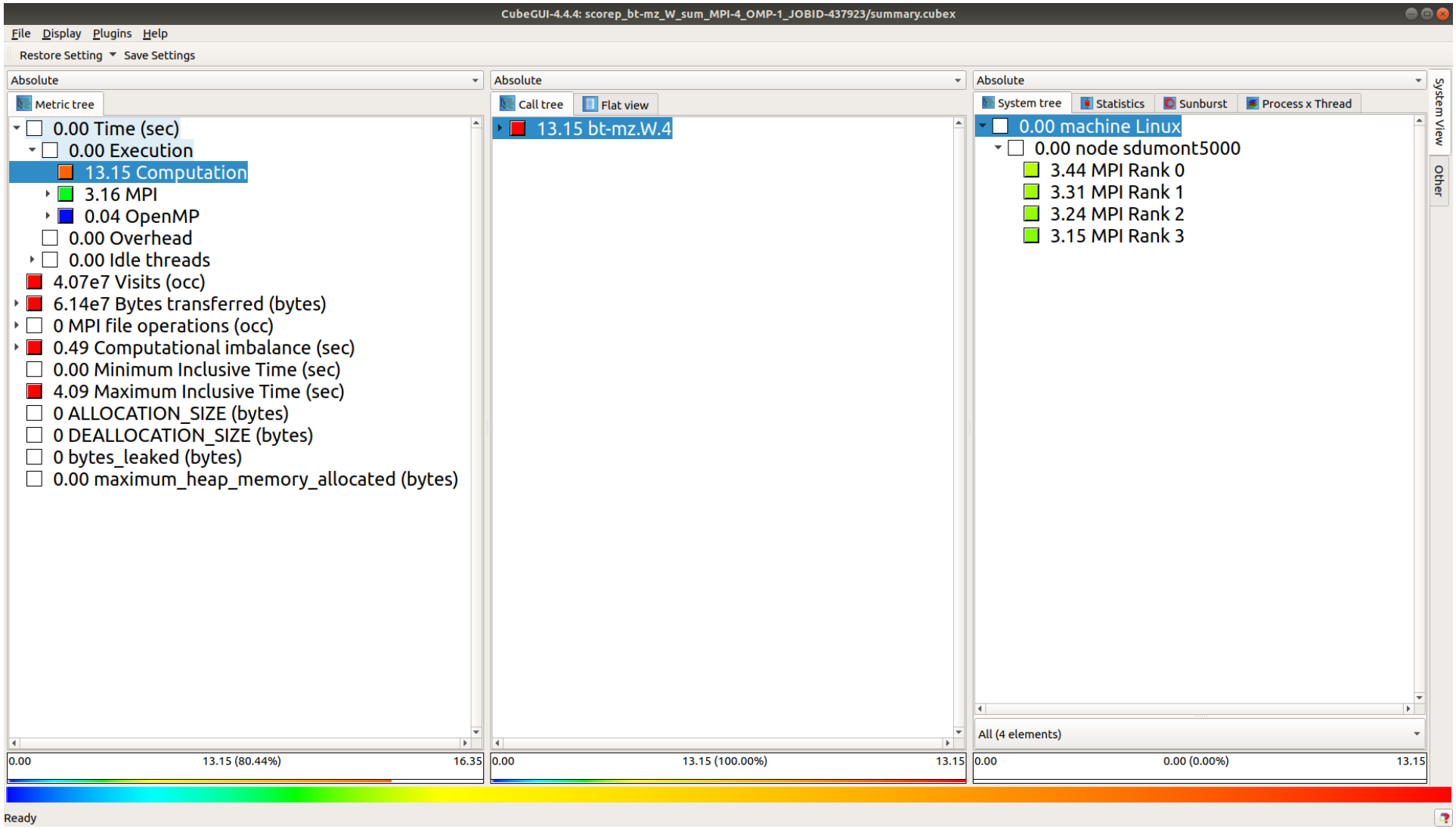
Visualizando: CubeGUI

-nodes=1 -ntasks=2



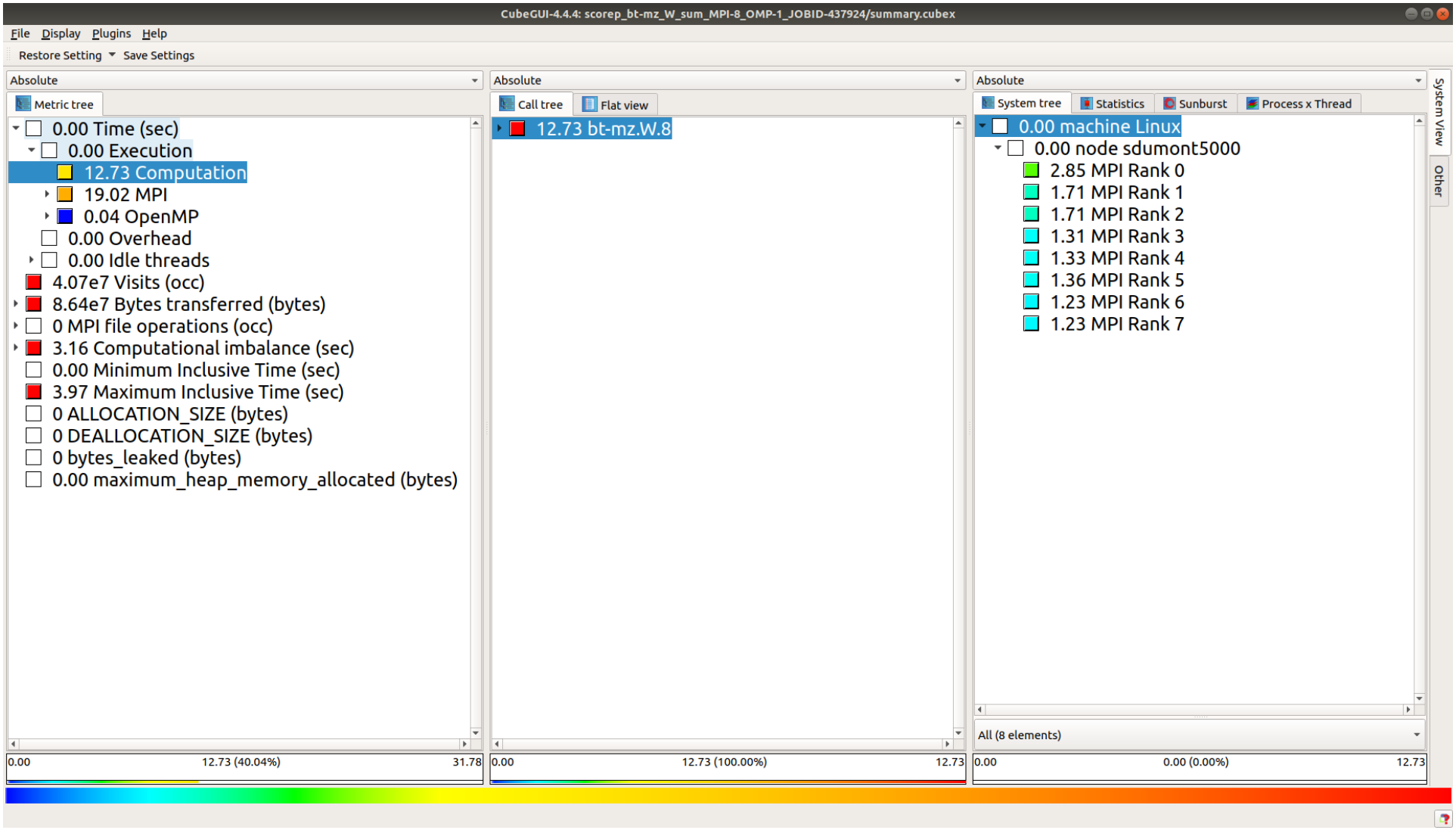
Visualizando: CubeGUI

-nodes=1 -ntasks=4



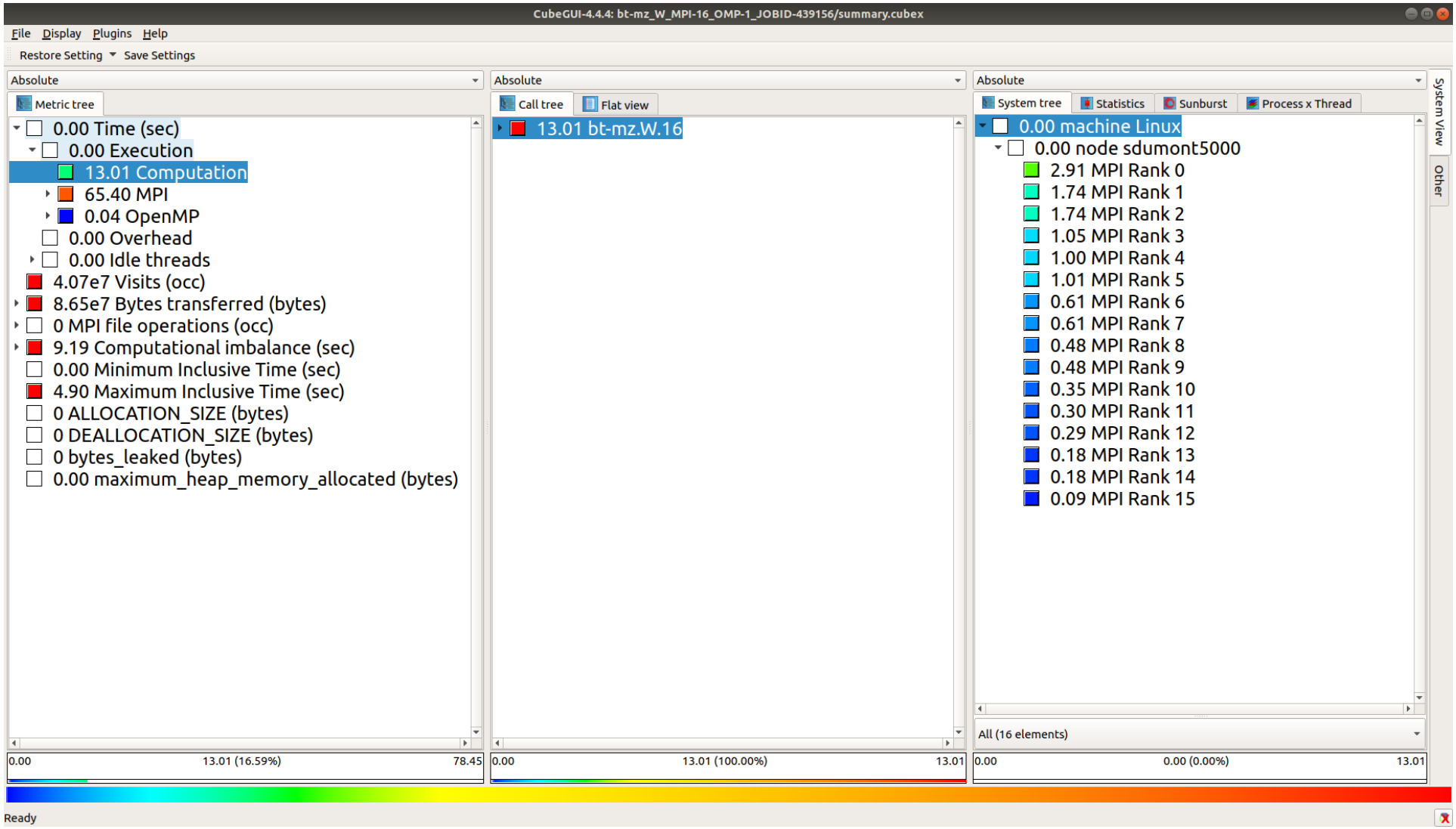
Visualizando: CubeGUI

-nodes=1 -ntasks=8



Visualizando: CubeGUI

-nodes=1 -ntasks=16



BT-MZ *benchmark*: divisão de domínio

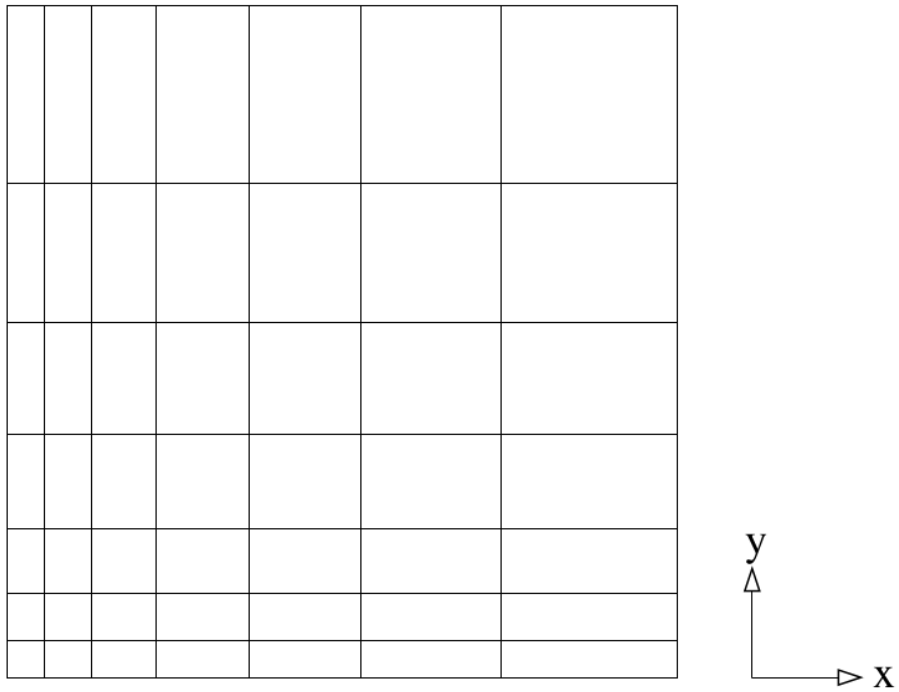


Figure 3: Example of uneven mesh tiling (horizontal cut through mesh system) for the BT-MZ benchmark.

Scalasca

Definição

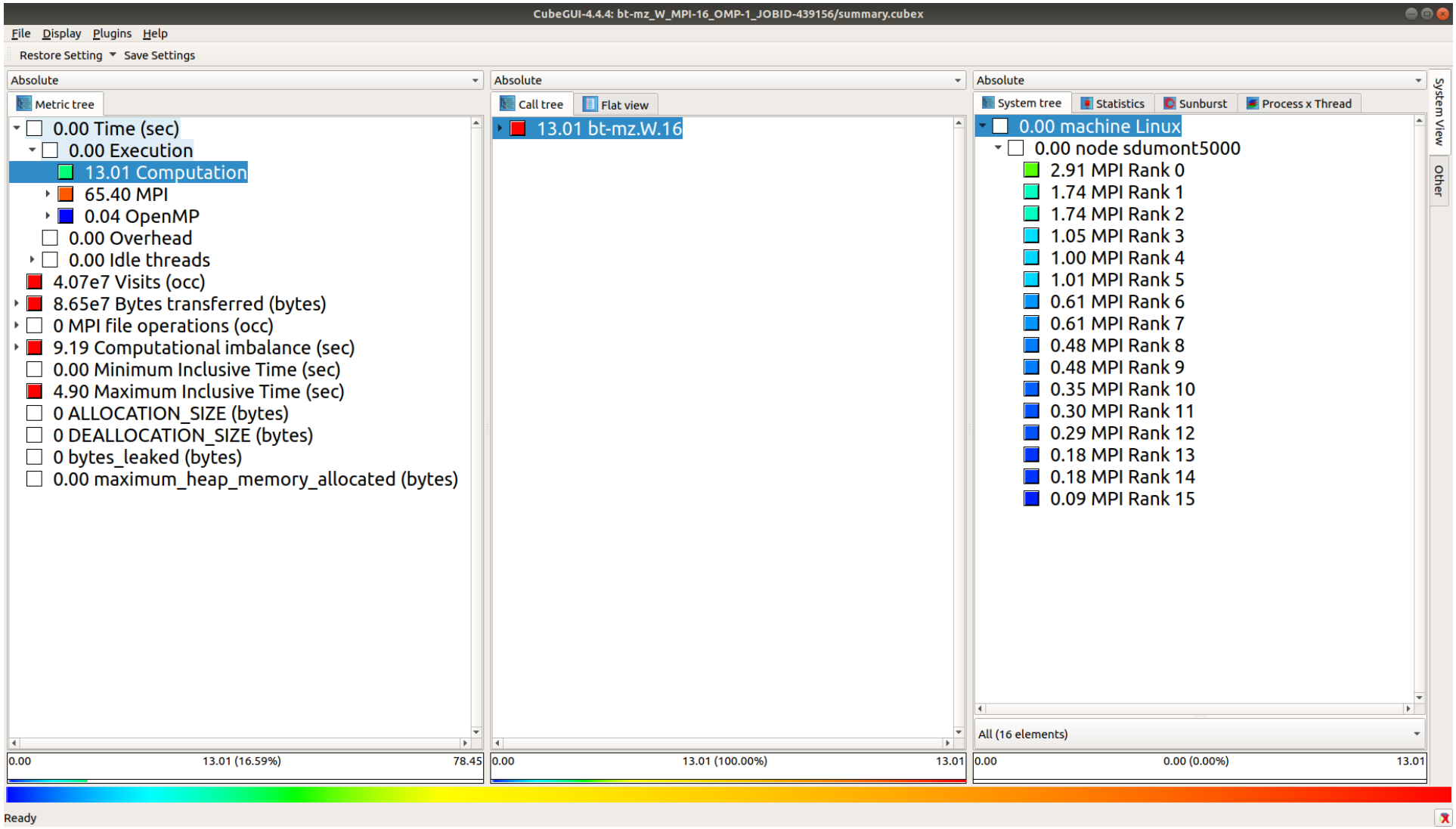
Balanço de carga de computação (LB):

$$LB = \frac{avg(tcomp)}{max(tcomp)}$$

FONTE: **`https://pop-coe.eu`**

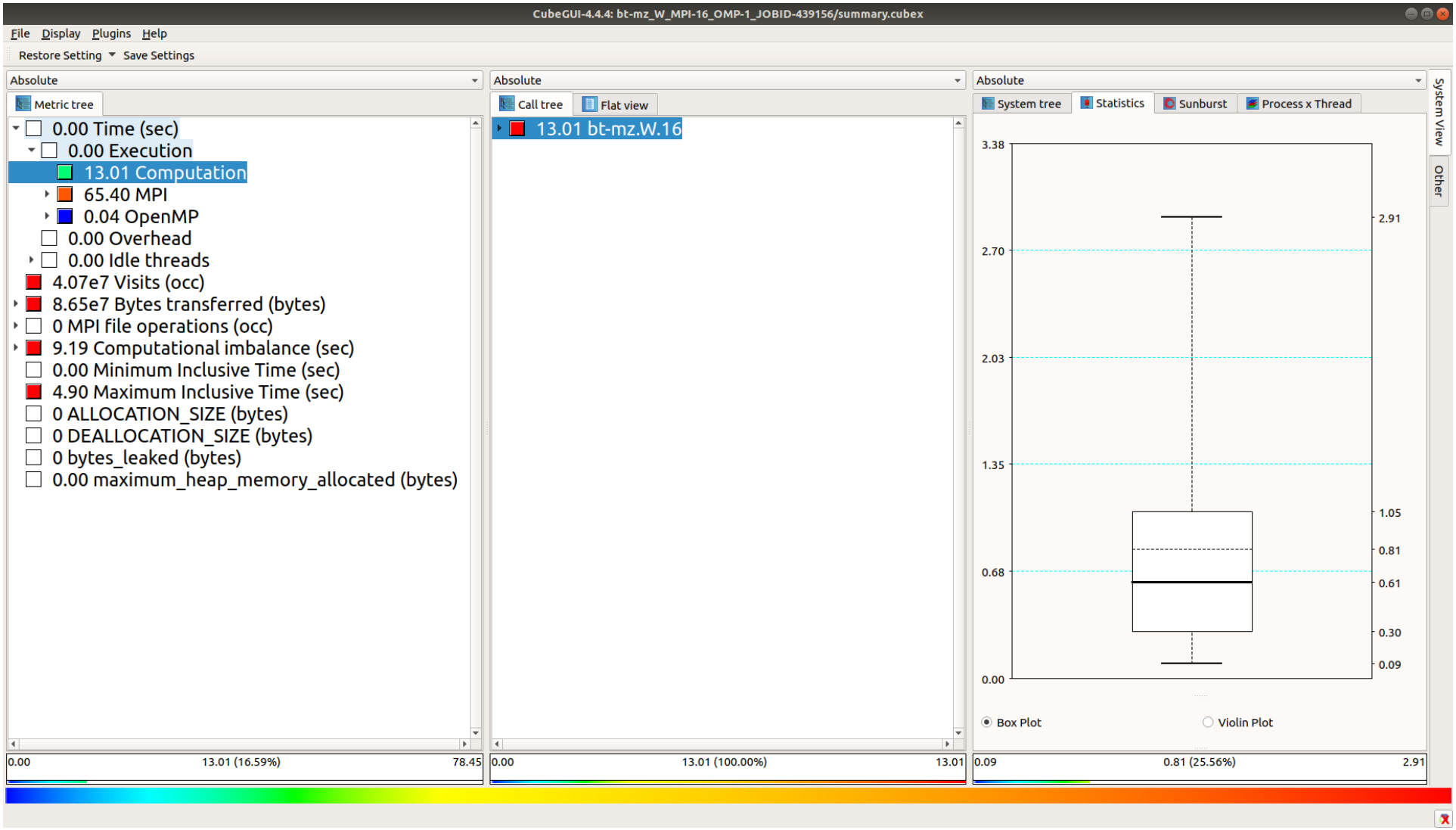
Visualizando: CubeGUI

-nodes=1 -ntasks=16



Visualizando: CubeGUI

-nodes=1 -ntasks=16



Scalasca

Cálculo

Balanço de carga de computação (LB)

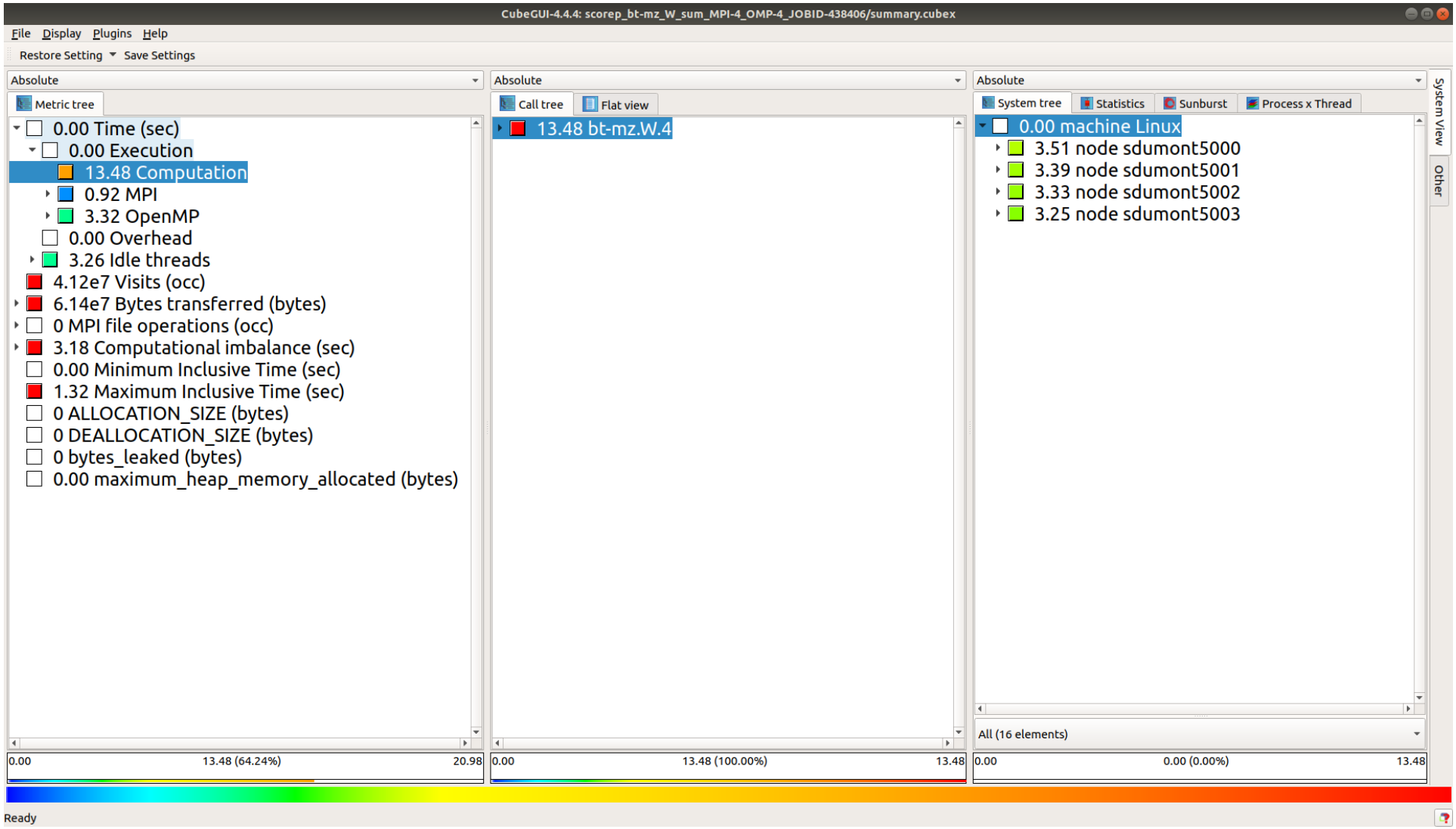
$$LB = \frac{avg(tcomp)}{max(tcomp)}$$

$$LB = \frac{0.81}{2.91}$$

$$LB = 0.28$$

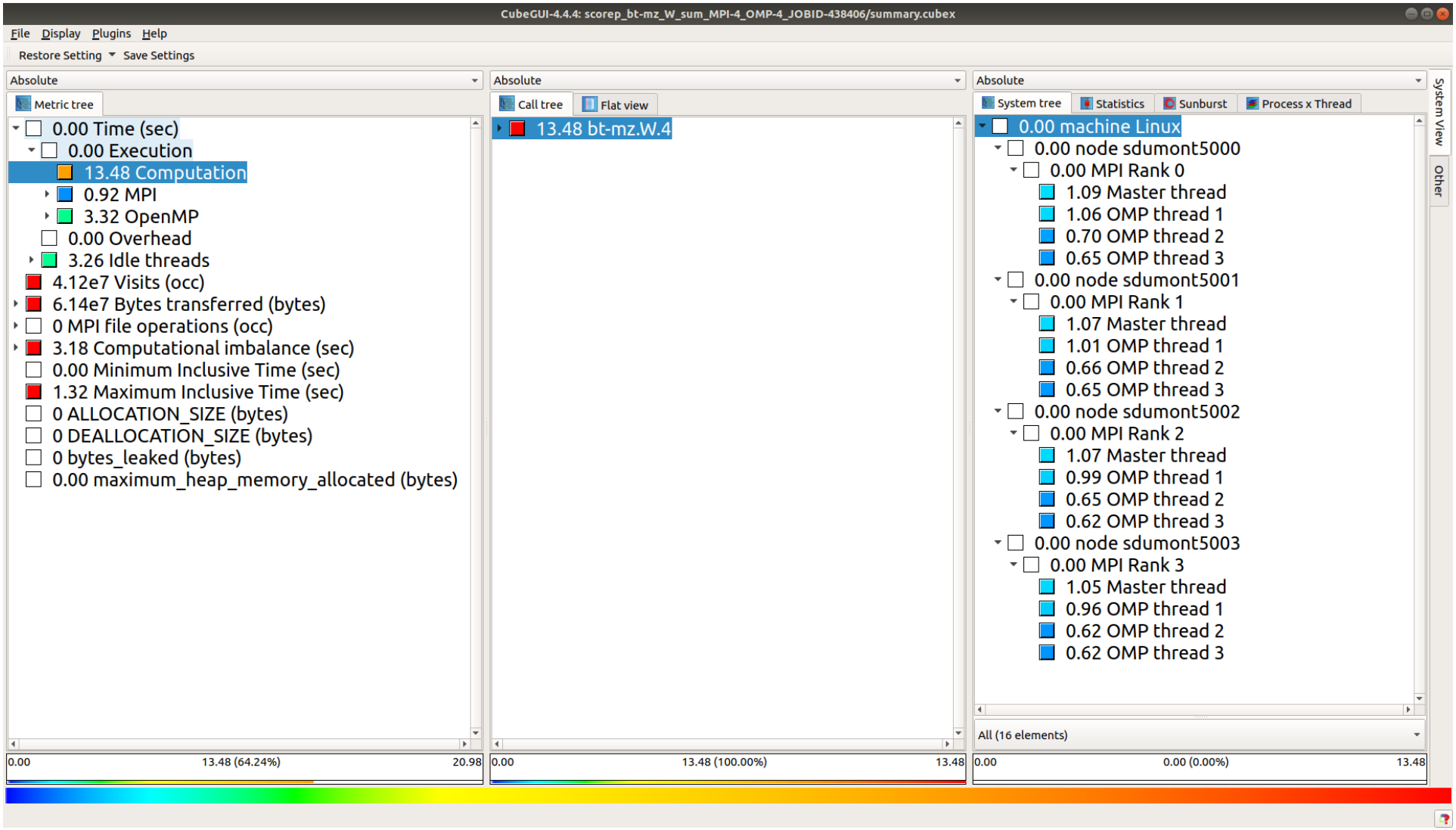
Visualizando: CubeGUI

`-nodes=4 -ntasks=4 -cpus-per-task=4`



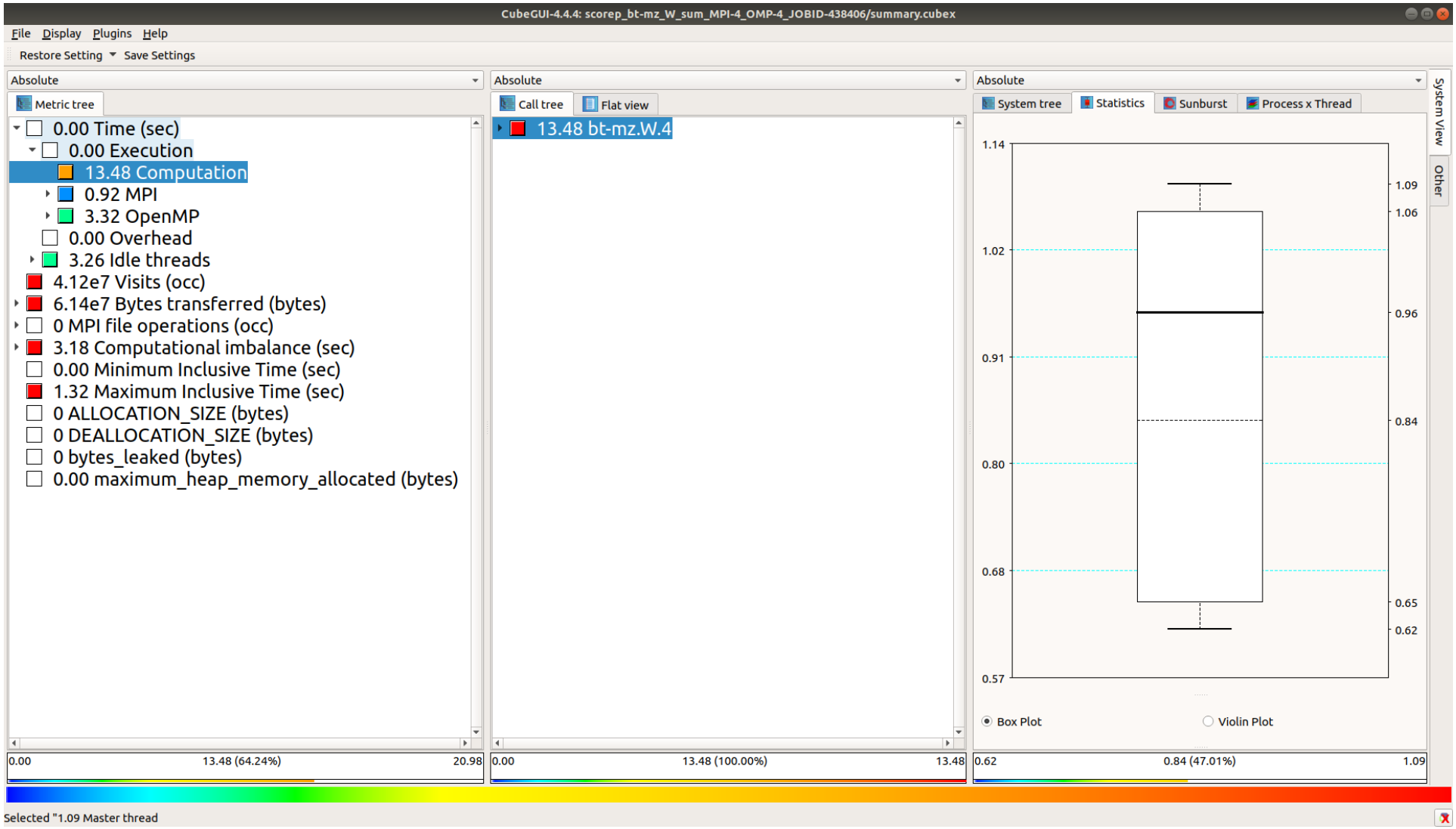
Visualizando: CubeGUI

`-nodes=4 -ntasks=4 -cpus-per-task=4`



Visualizando: CubeGUI

`-nodes=4 -ntasks=4 -cpus-per-task=4`



Scalasca

Cálculo

Balanço de carga de computação (LB)

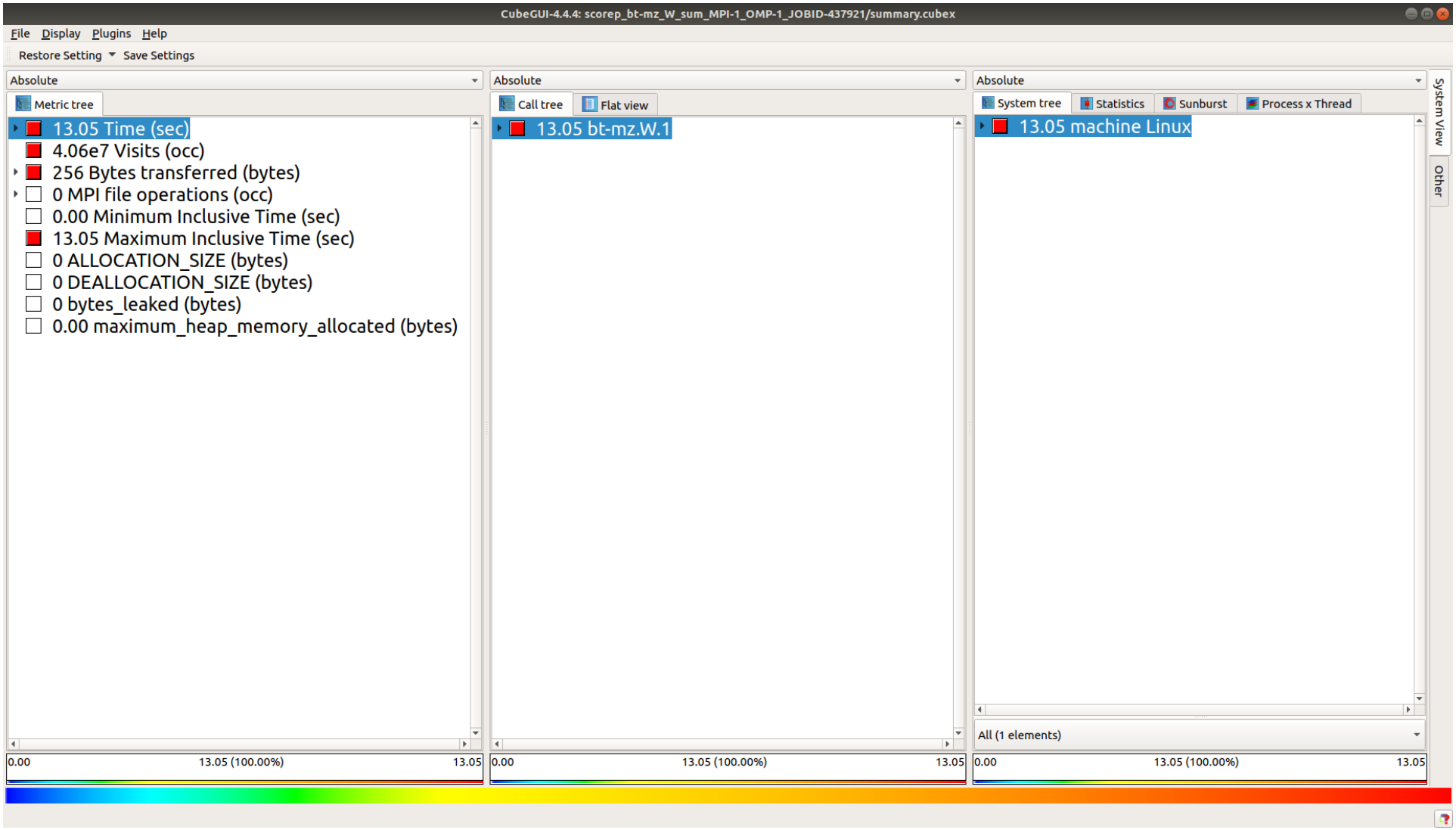
$$LB = \frac{avg(tcomp)}{max(tcomp)}$$

$$LB = \frac{0.84}{1.09}$$

$$LB = 0.77$$

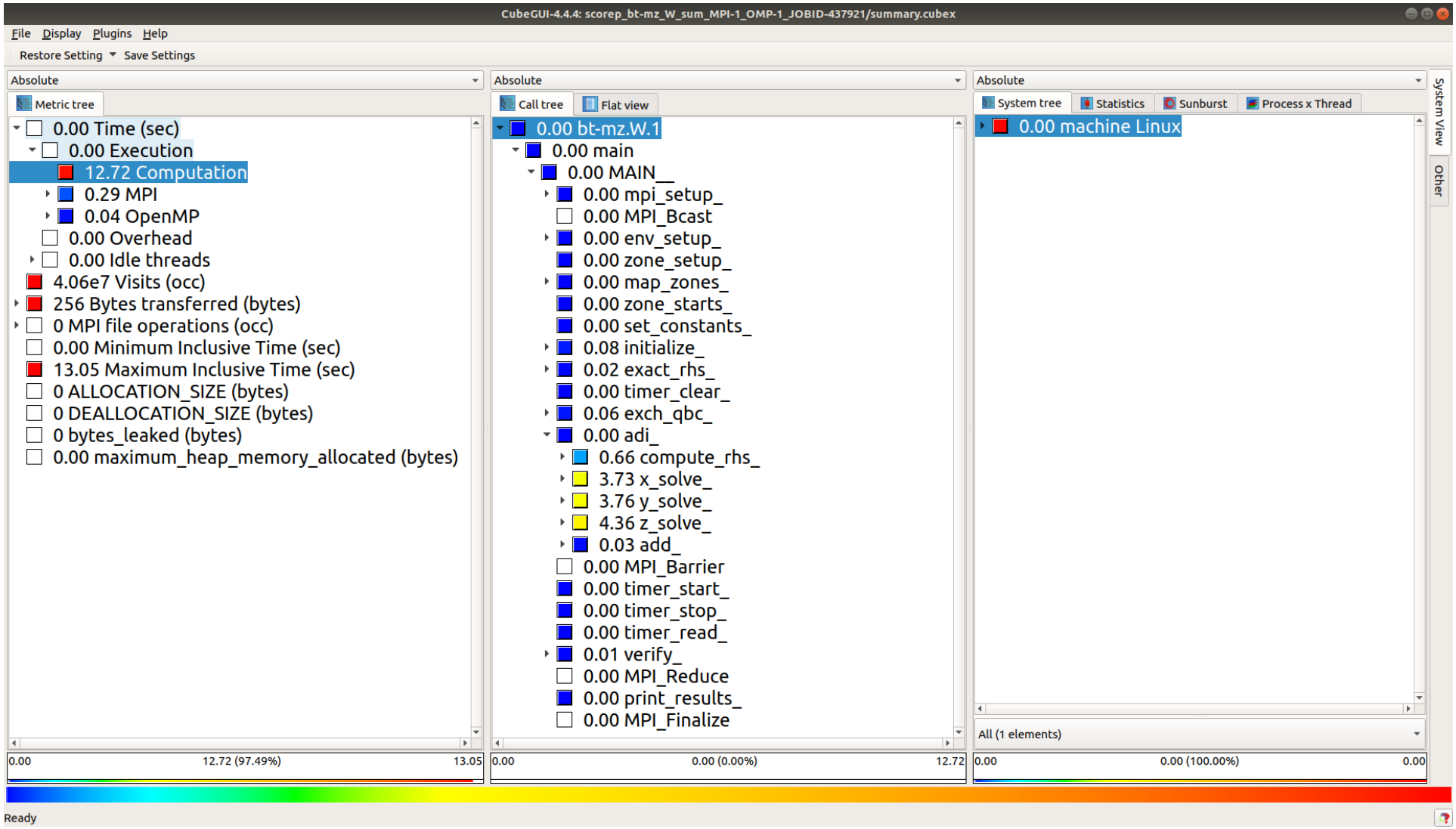
Visualizando: CubeGUI

`-nodes=1 -ntasks=1` / **Absolute**



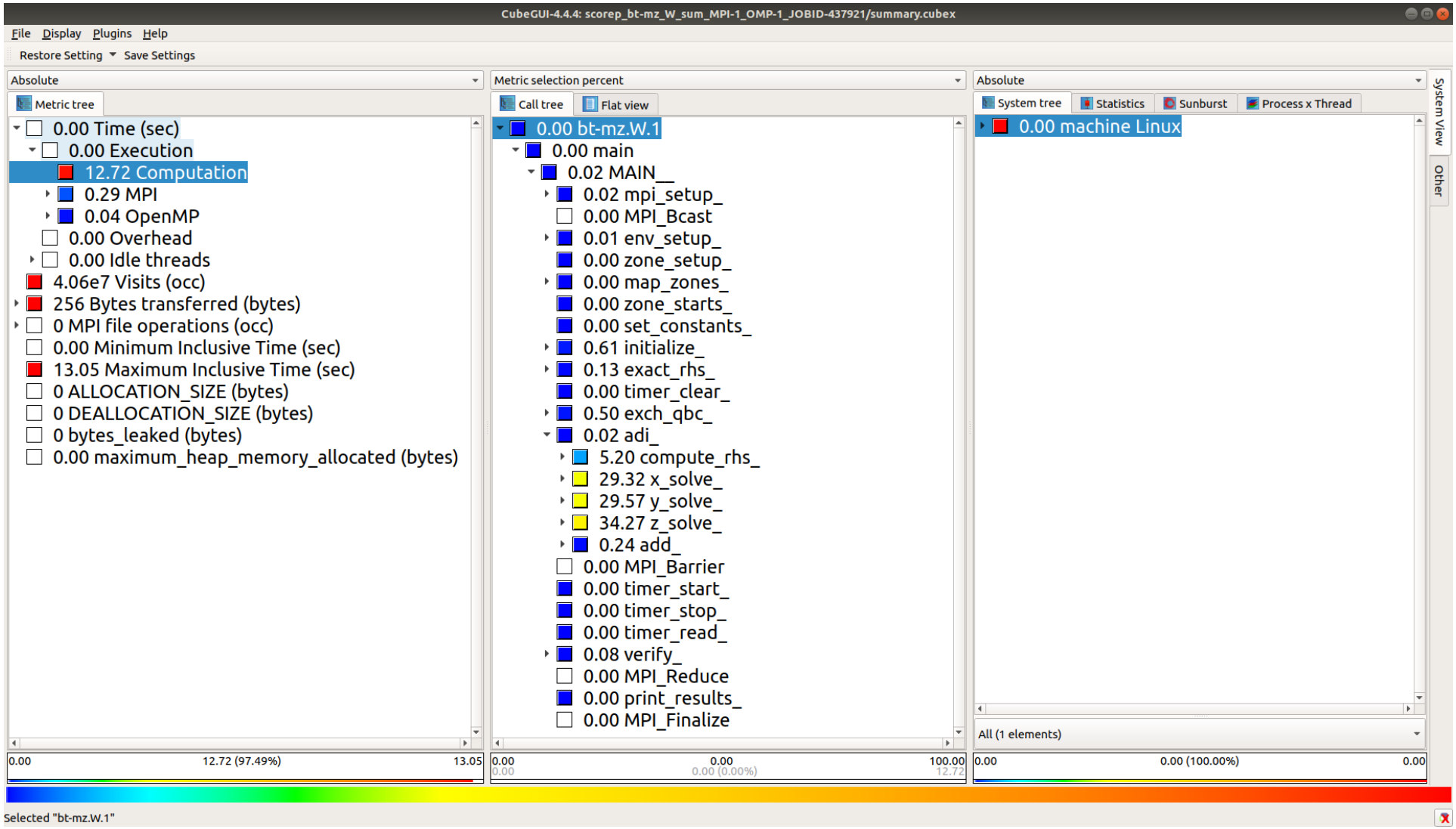
Visualizando: CubeGUI

`-nodes=1 -ntasks=1` / **Absolute**



Visualizando: CubeGUI

`-nodes=1 -ntasks=1` / **Metric Own percent**



Visualizando: CubeGUI

3 hotspots de computação

- **`x_solve`**: 3.73s (29.32%)
- **`y_solve`**: 3.76s (29.57%)
- **`z_solve`**: 4.36s (34.27%)

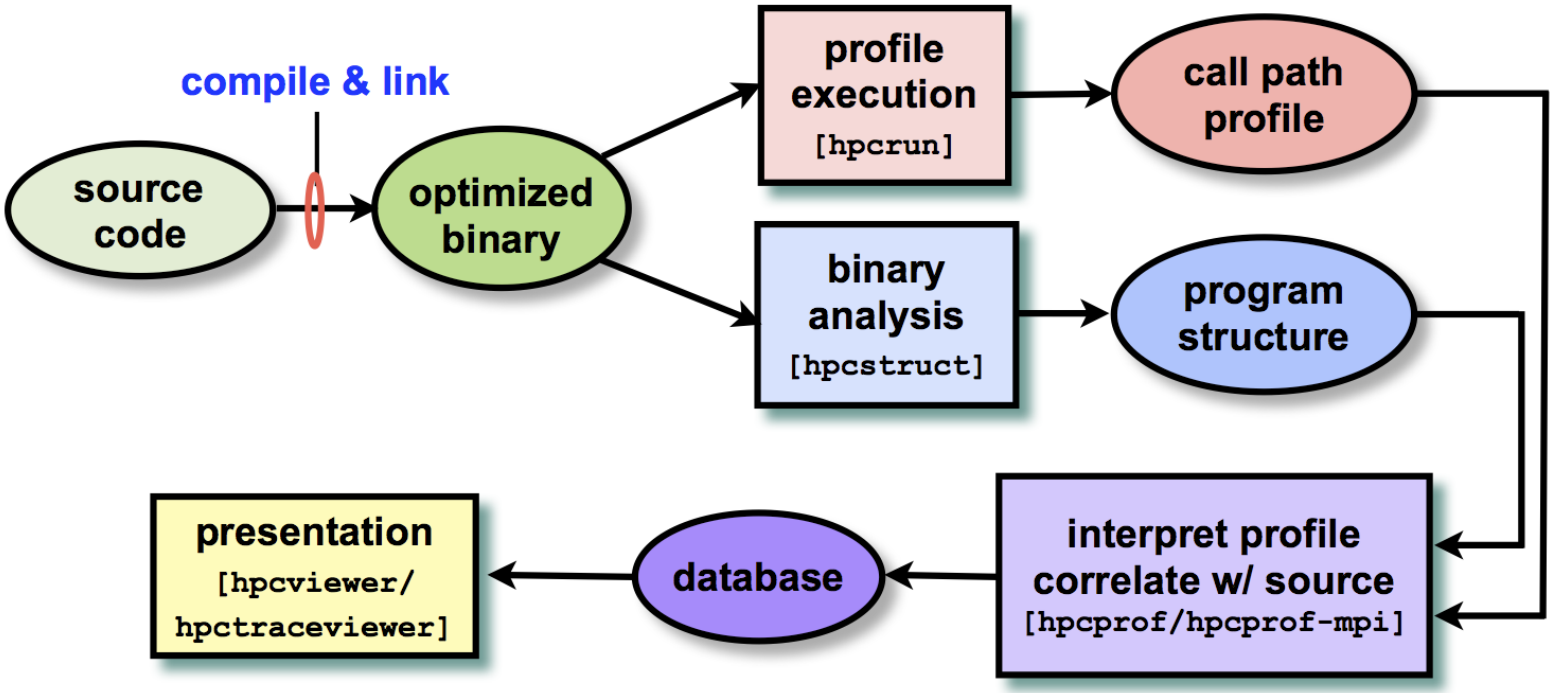
Roteiro

1. Scalasca, Score-p, Cube

2. hpctoolkit

HPCToolkit

http://hpctoolkit.org



Visualizando no **hpcviewer**

- O **hpcviewer** pode ser baixado e instalado para visualizar os resultados obtidos com o HPCToolkit
- `http://hpctoolkit.org/download.html`
- Binários prontos em "Binary Releases of HPCToolkit's hpcviewer Graphical User Interface"
- Resultados previamente obtidos no SDumont estão no arquivo **profiling_hpctoolkit_sdbase.zip** do repositório no GitHub:

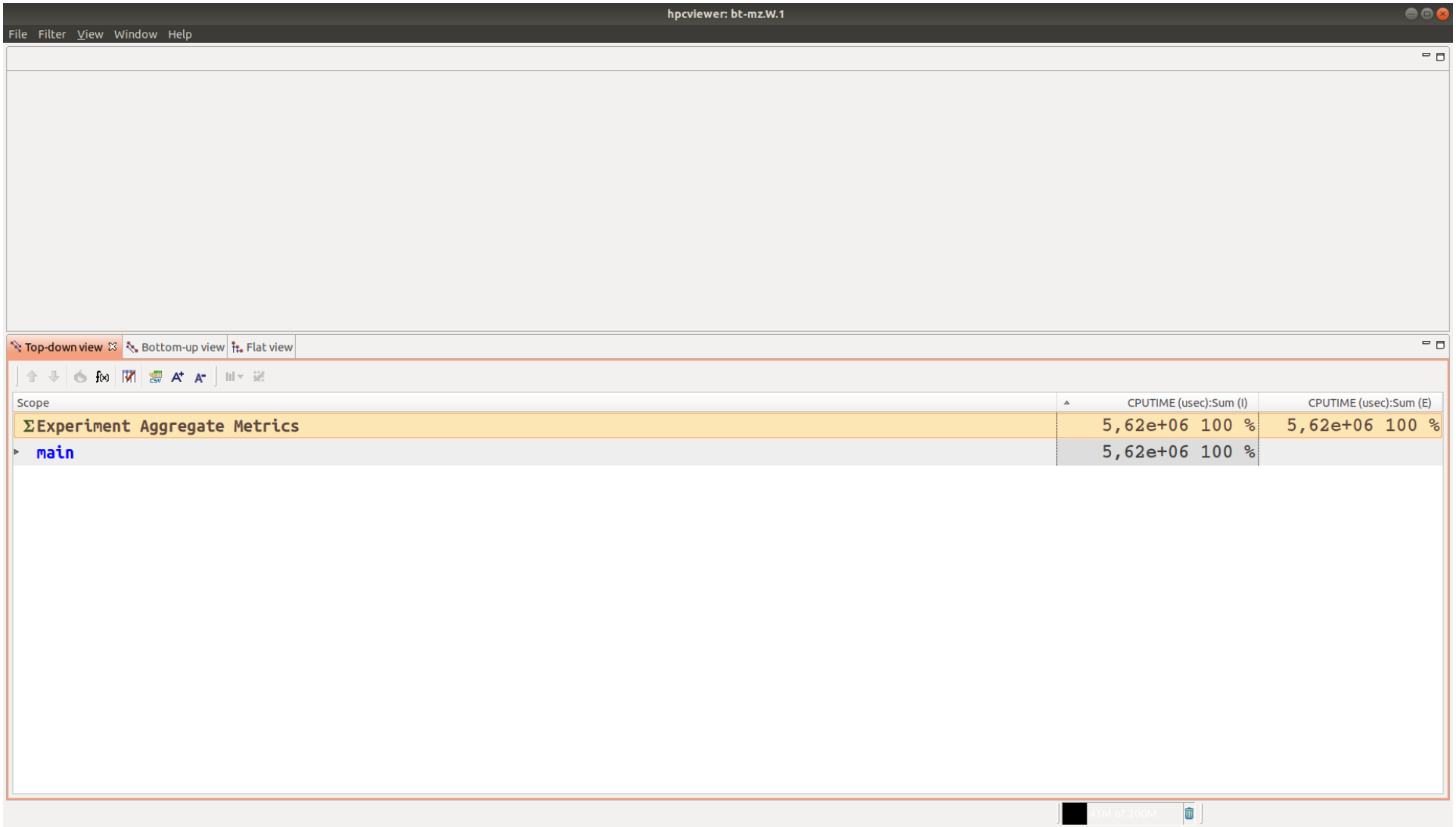
```
git clone https://github.com/robertopsouto/ESD2024.git
ESD2024/sdbase/profiling_hpctoolkit_sequana.zip
```

NPB: estudo de caso

```
$ cd ESD2024/sdbase/profiling/hpctoolkit/NUMNODES-1/bt-mz_A_MPI-1_OMP-1_JO
$ hpcviewer hpctoolkit-bt-mz.A.1-database-10756696
```

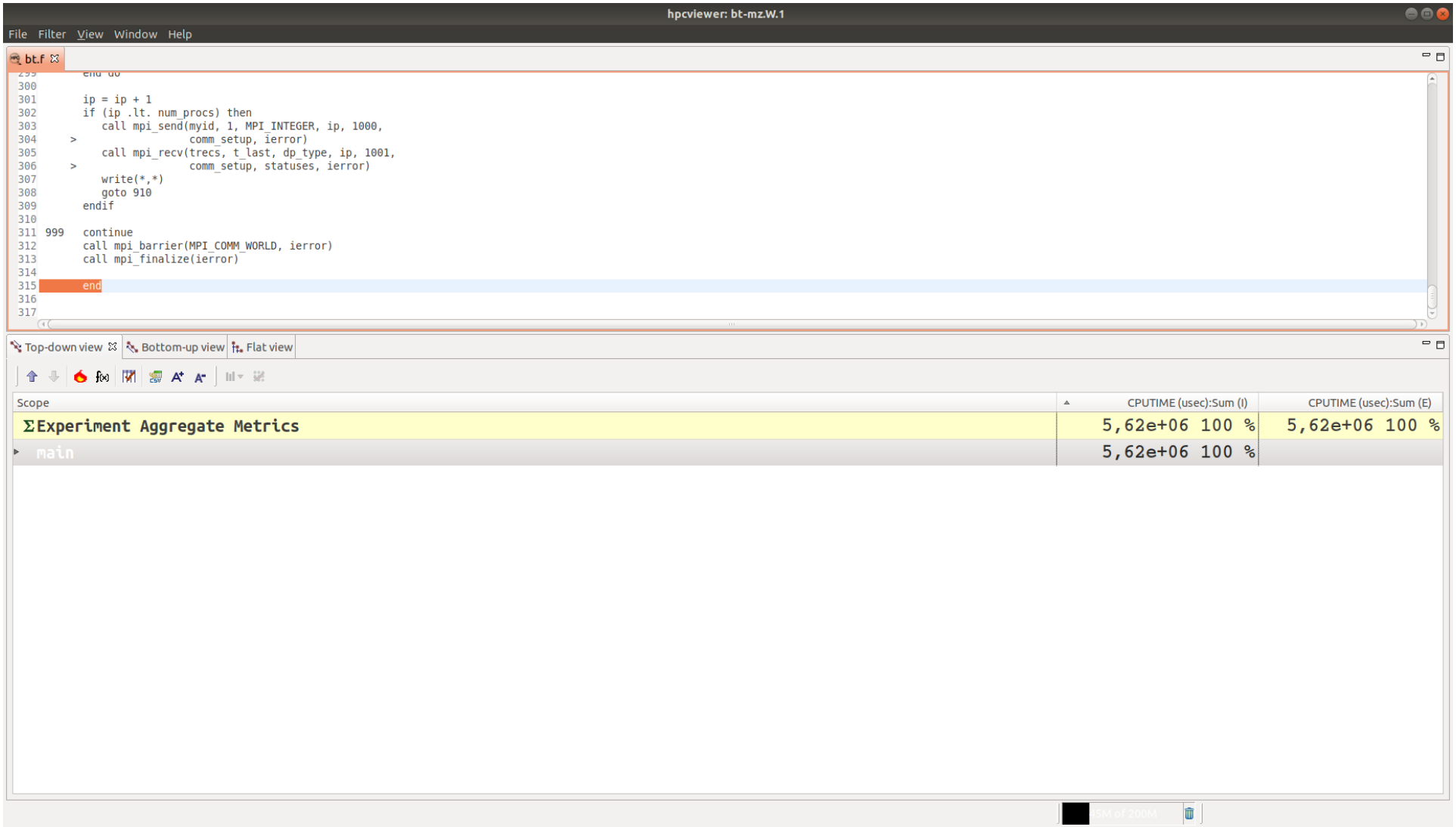
Visualizando no `hpcviewer`

`-nodes=1 -ntasks=1 / Top-down view`



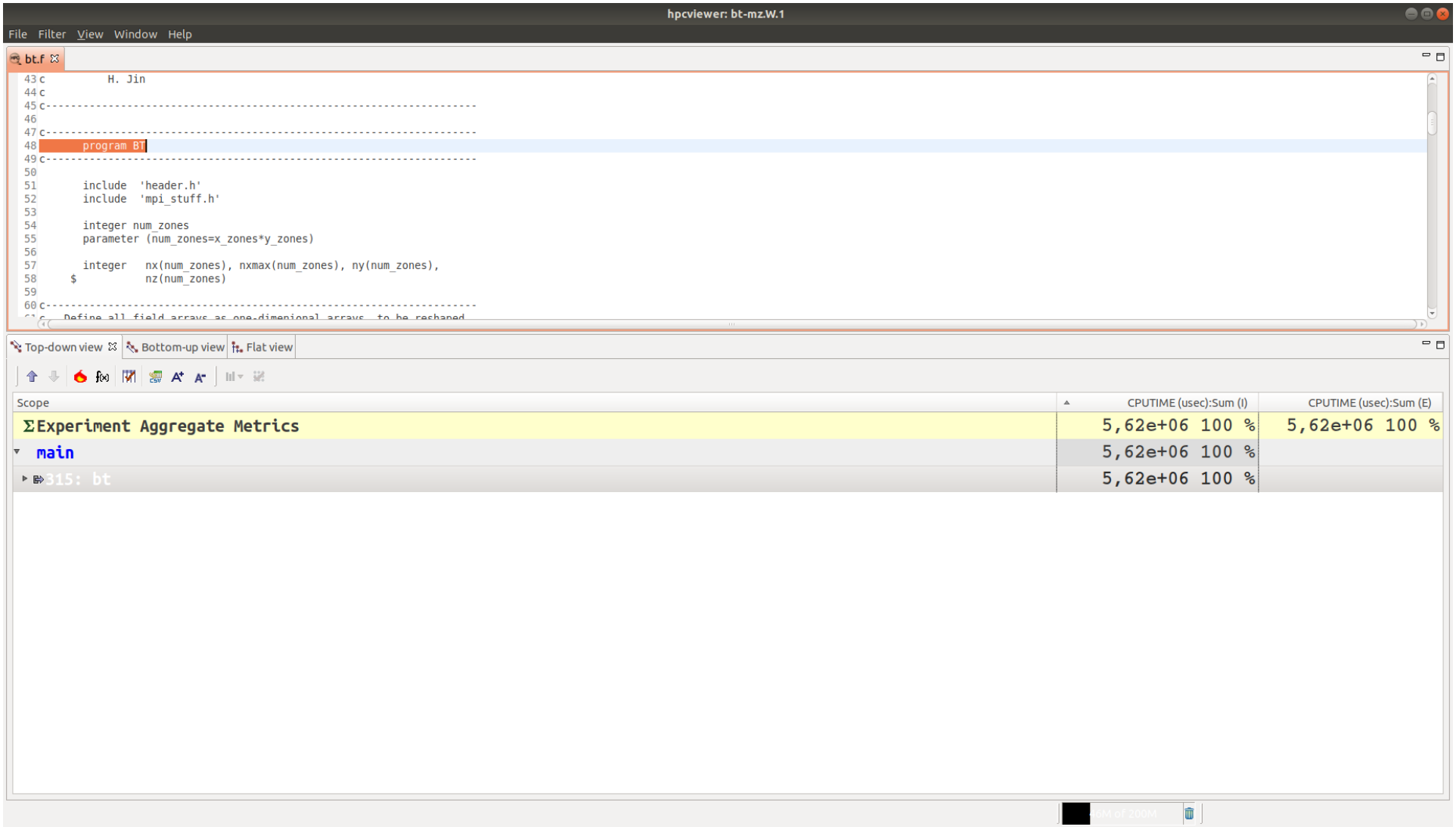
Visualizando no `hpcviewer`

`-nodes=1 -ntasks=1 / Top-down view`



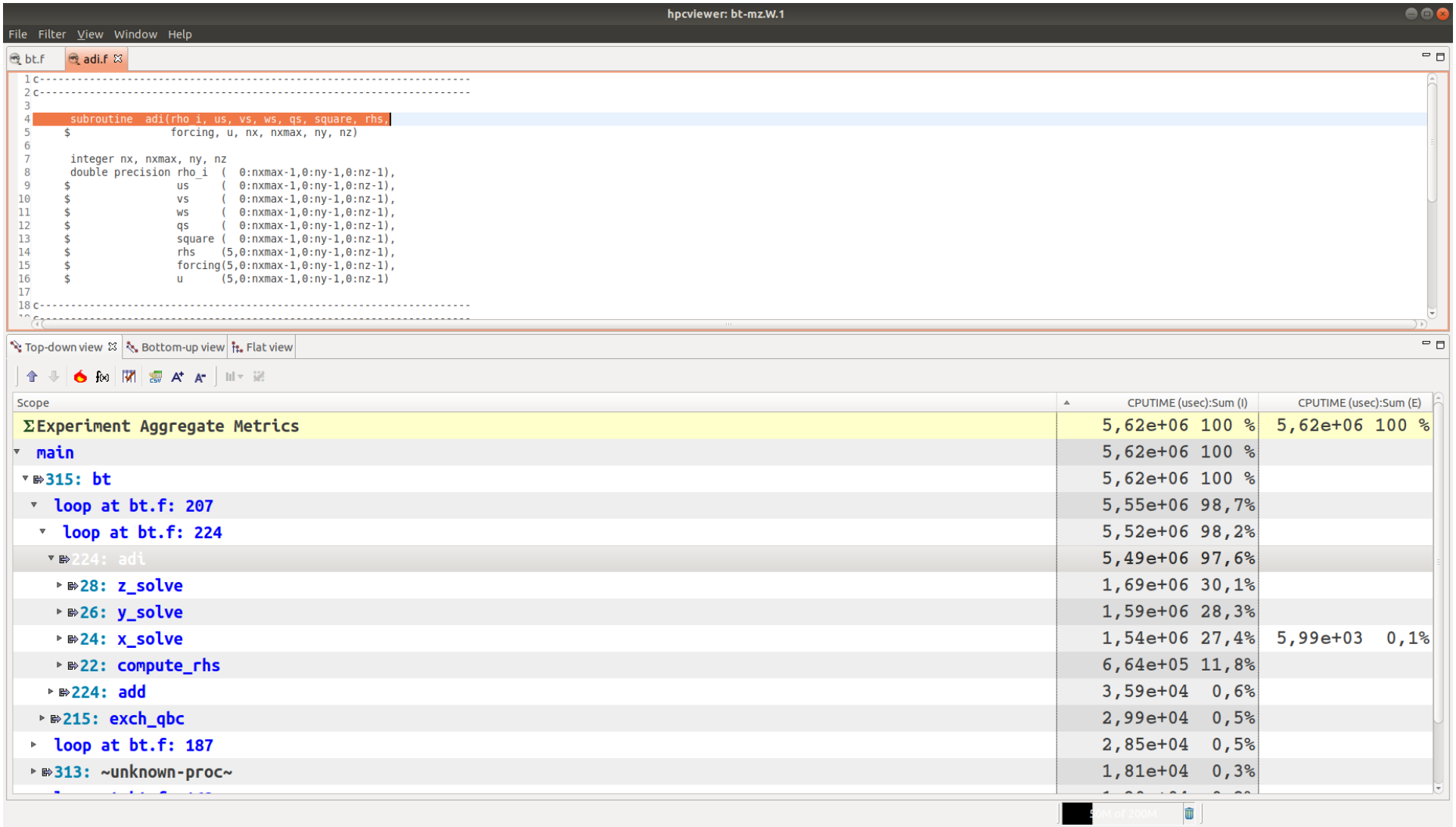
Visualizando no **hpcviewer**

-nodes=1 -ntasks=1 / Top-down view



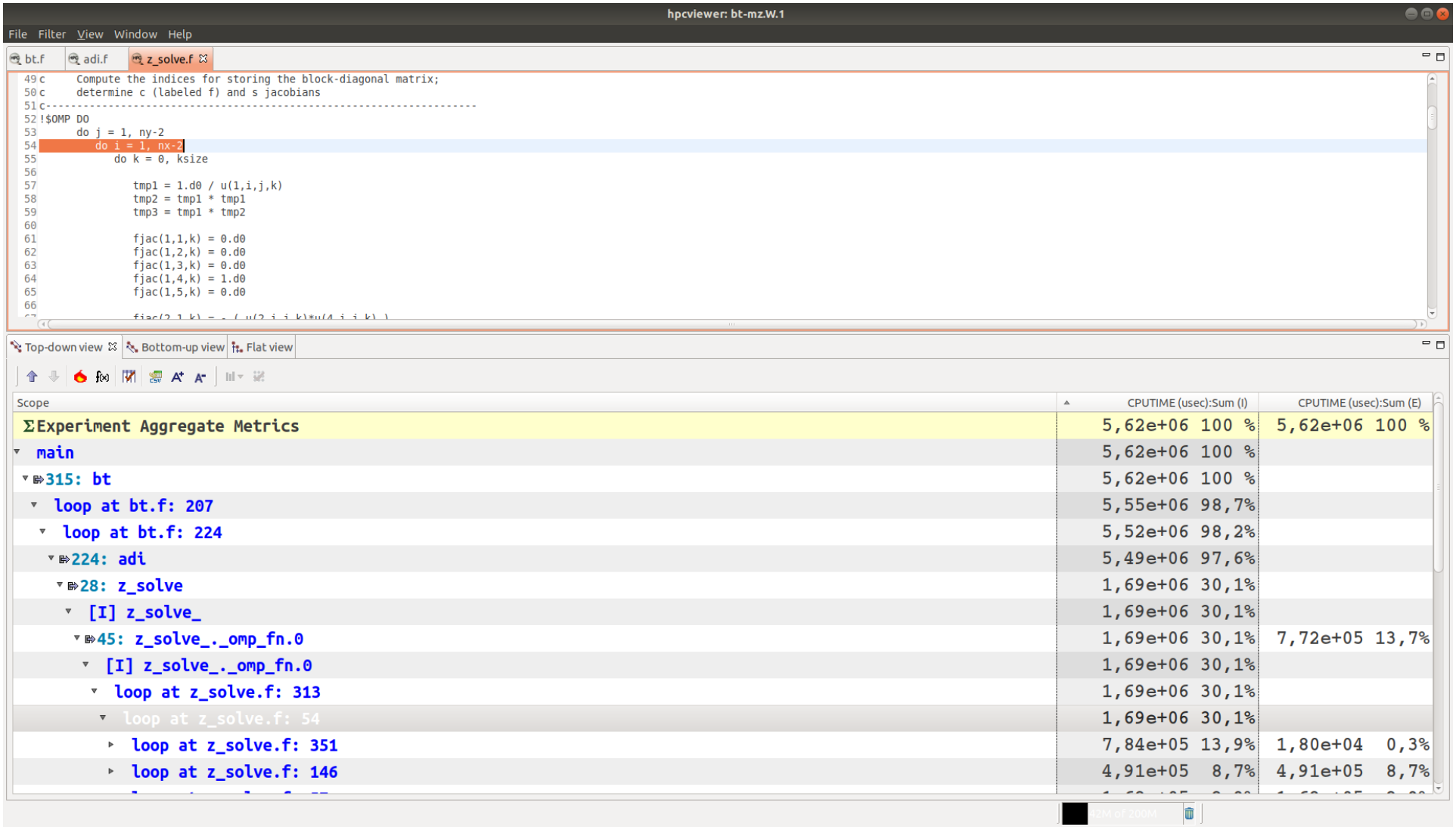
Visualizando no **hpcviewer**

-nodes=1 -ntasks=1 / Top-down view



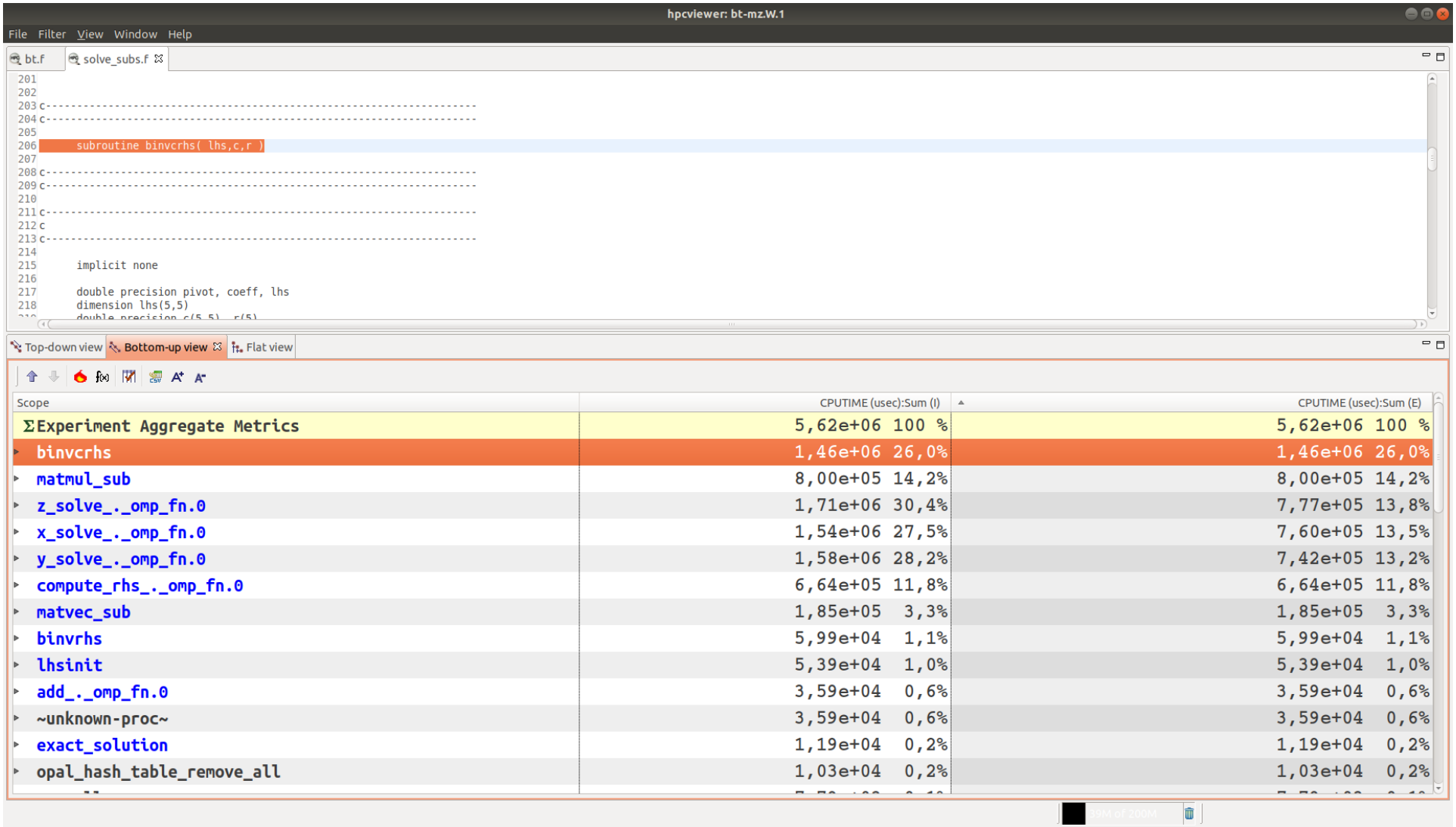
Visualizando no **hpcviewer**

-nodes=1 -ntasks=1 / Top-down view



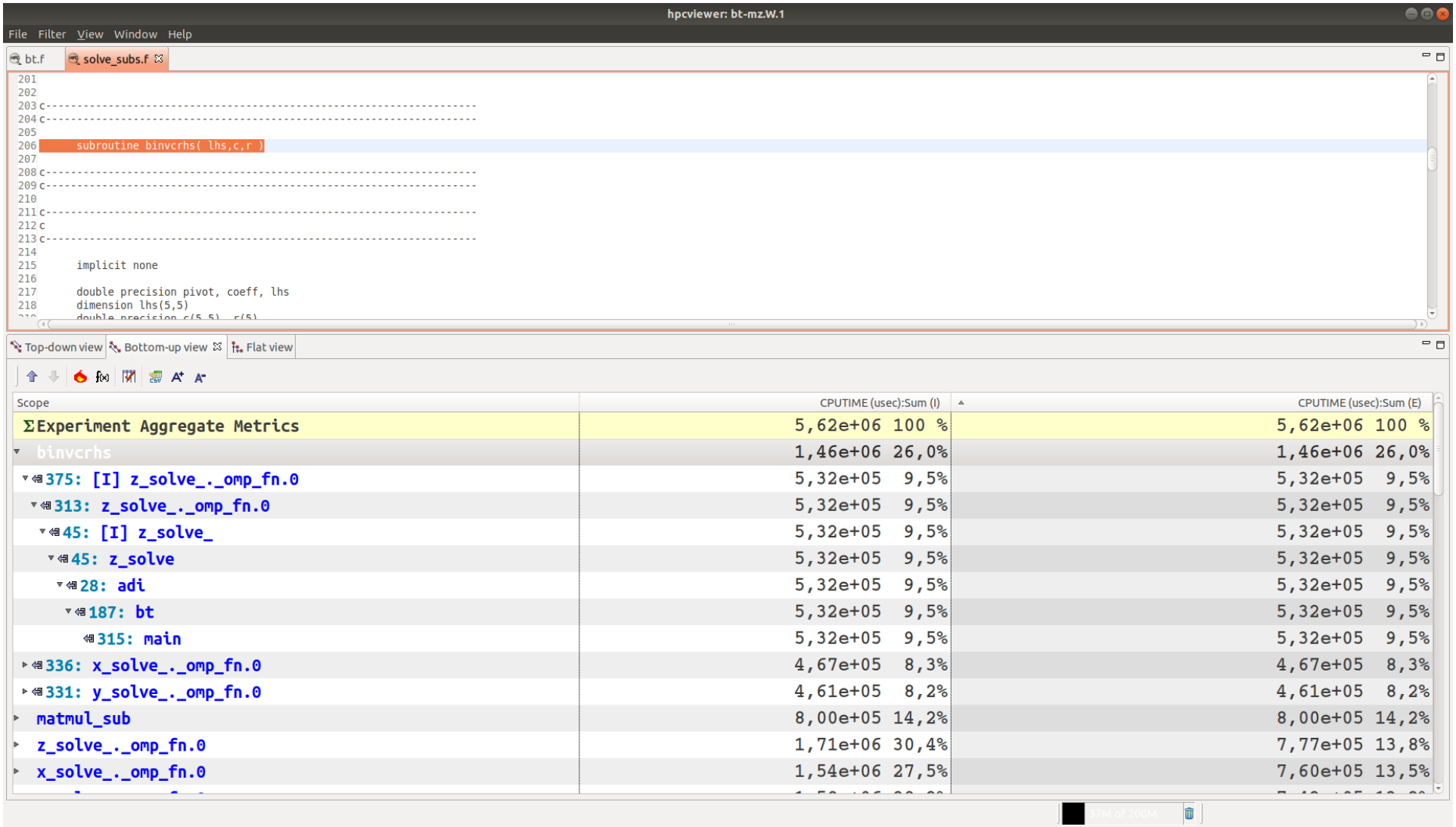
Visualizando no `hpcviewer`

`-nodes=1 -ntasks=1 / Bottom-up view`



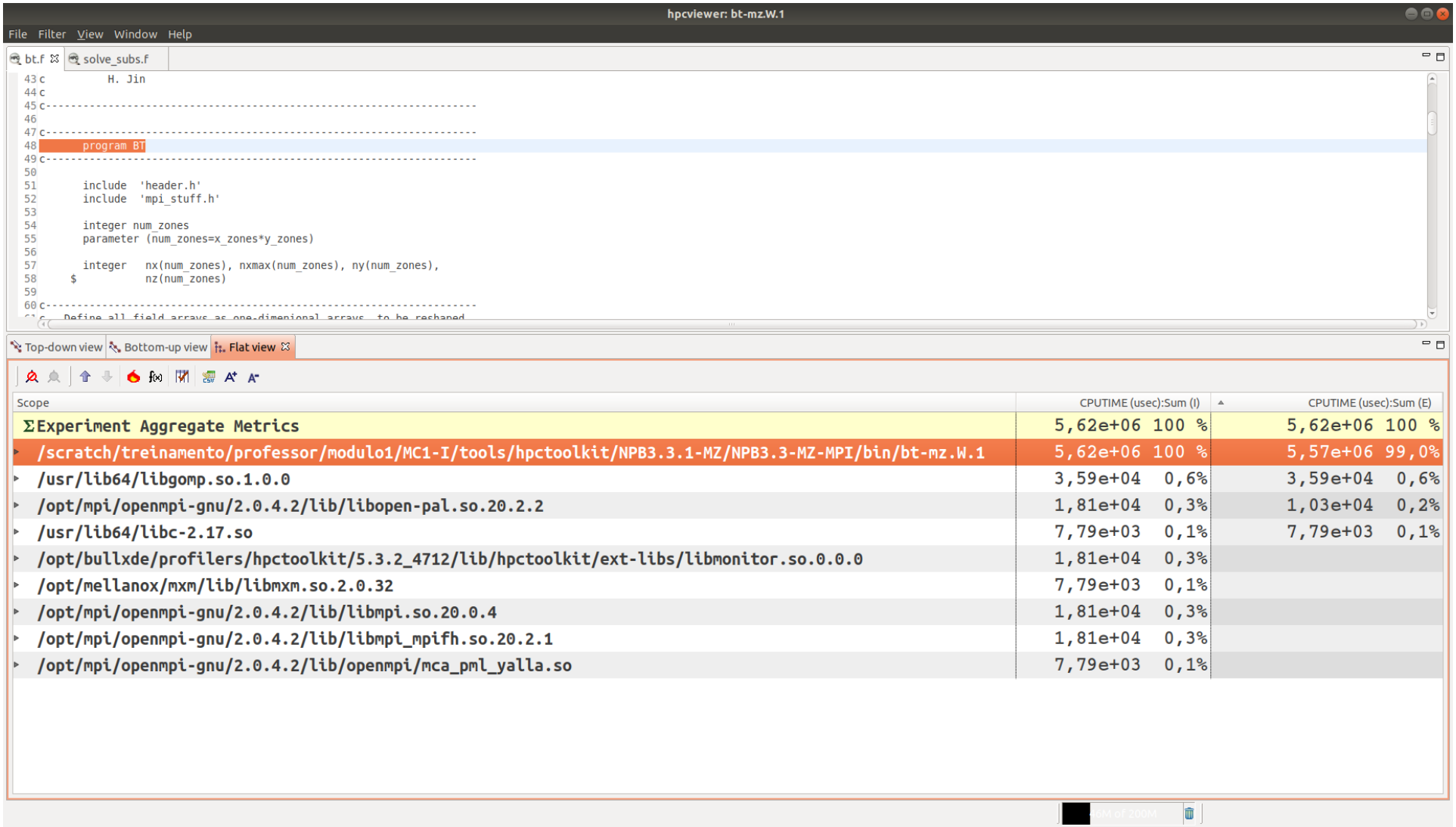
Visualizando no **hpcviewer**

-nodes=1 -ntasks=1 / Bottom-up view



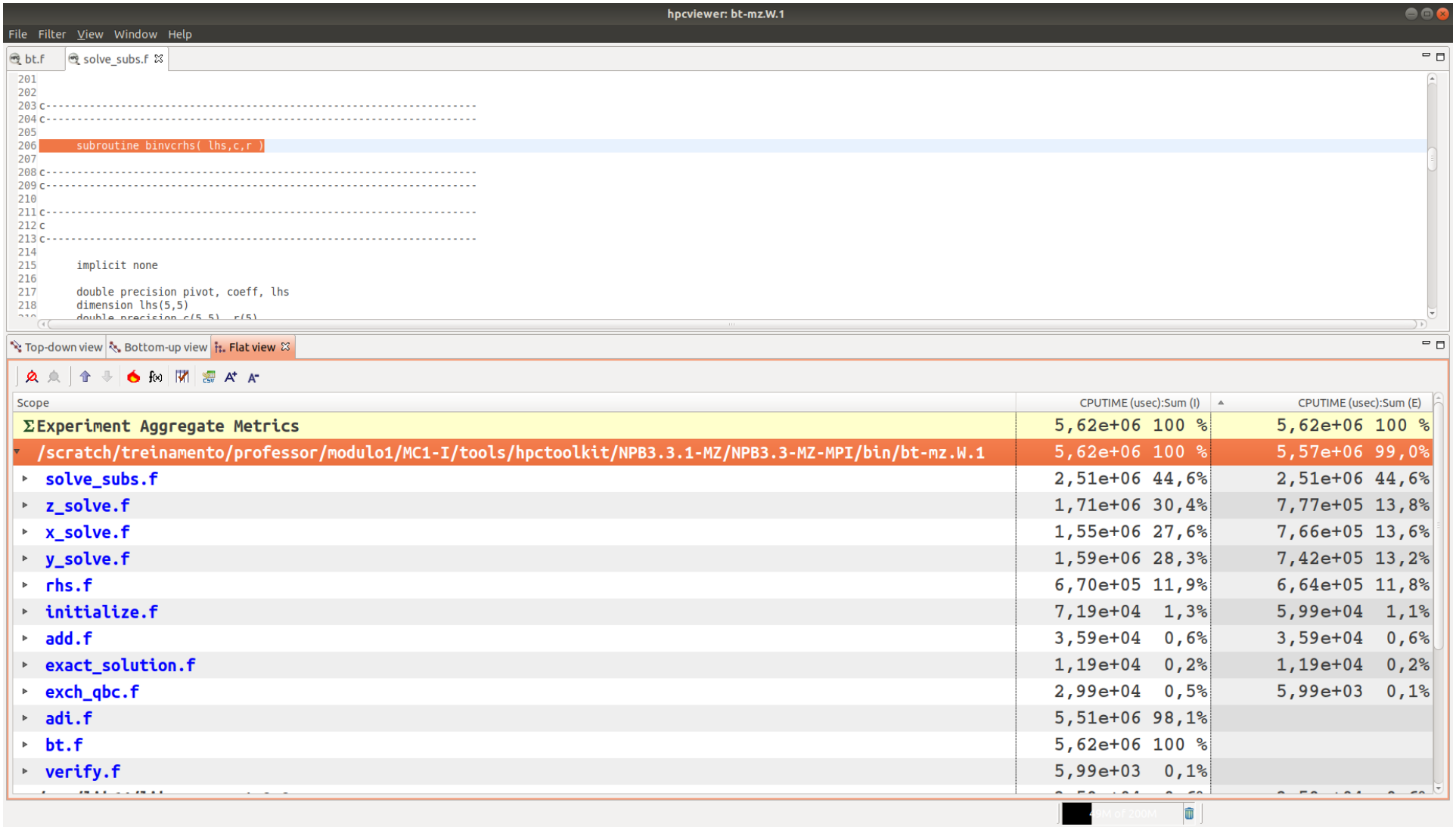
Visualizando no `hpcviewer`

`-nodes=1 -ntasks=1 / Flat view`



Visualizando no **hpcviewer**

-nodes=1 -ntasks=1 / Flat view



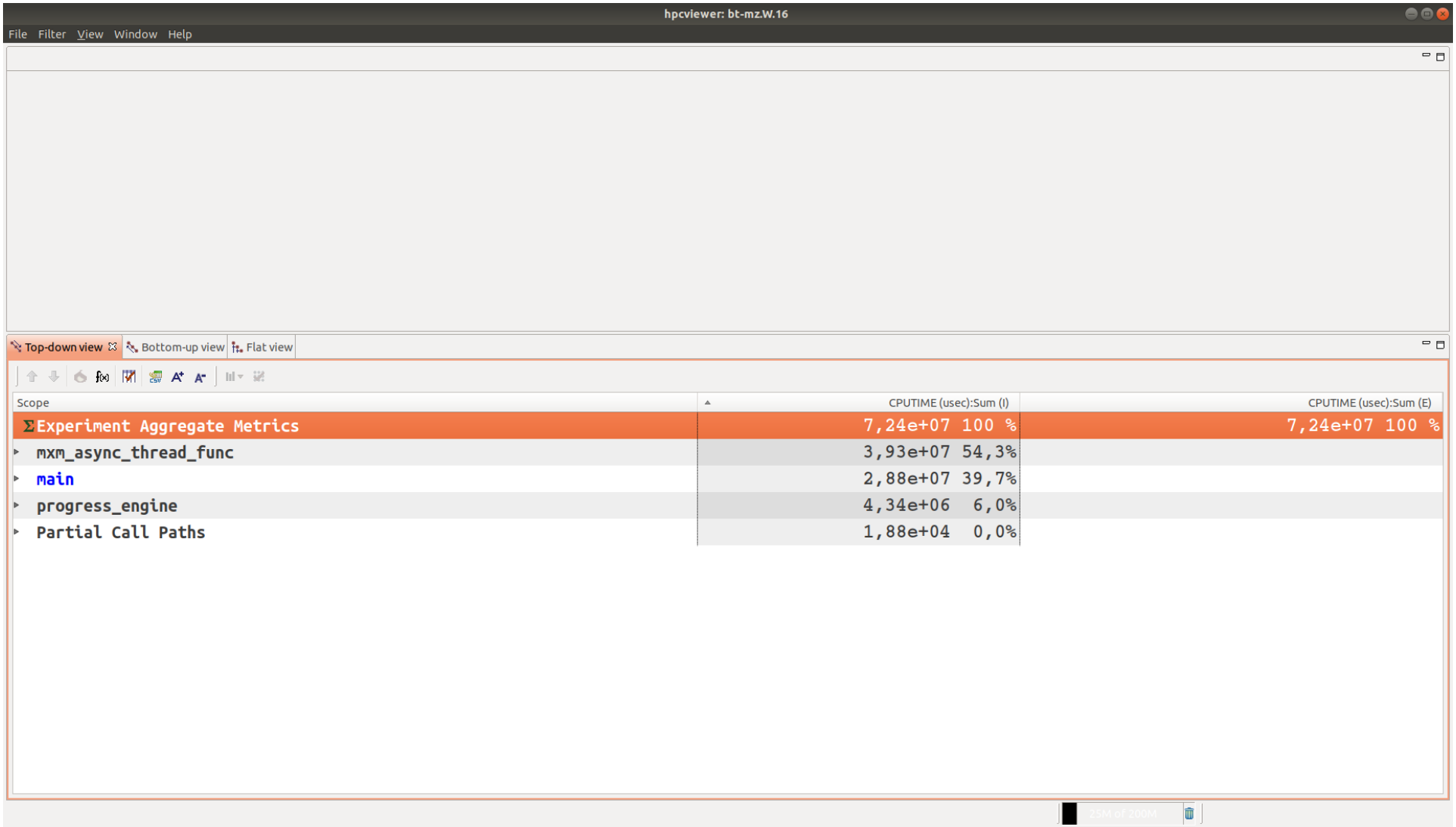
Visualizando no `hpcviewer`

NPB: estudo de caso

```
$ cd profiling/hpctoolkit/NUMNODES-1/bt-mz_W_MPI-16_OMP-1_JOBID-439032
$ hpcviewer hpctoolkit-bt-mz.W.16-database-439032
```

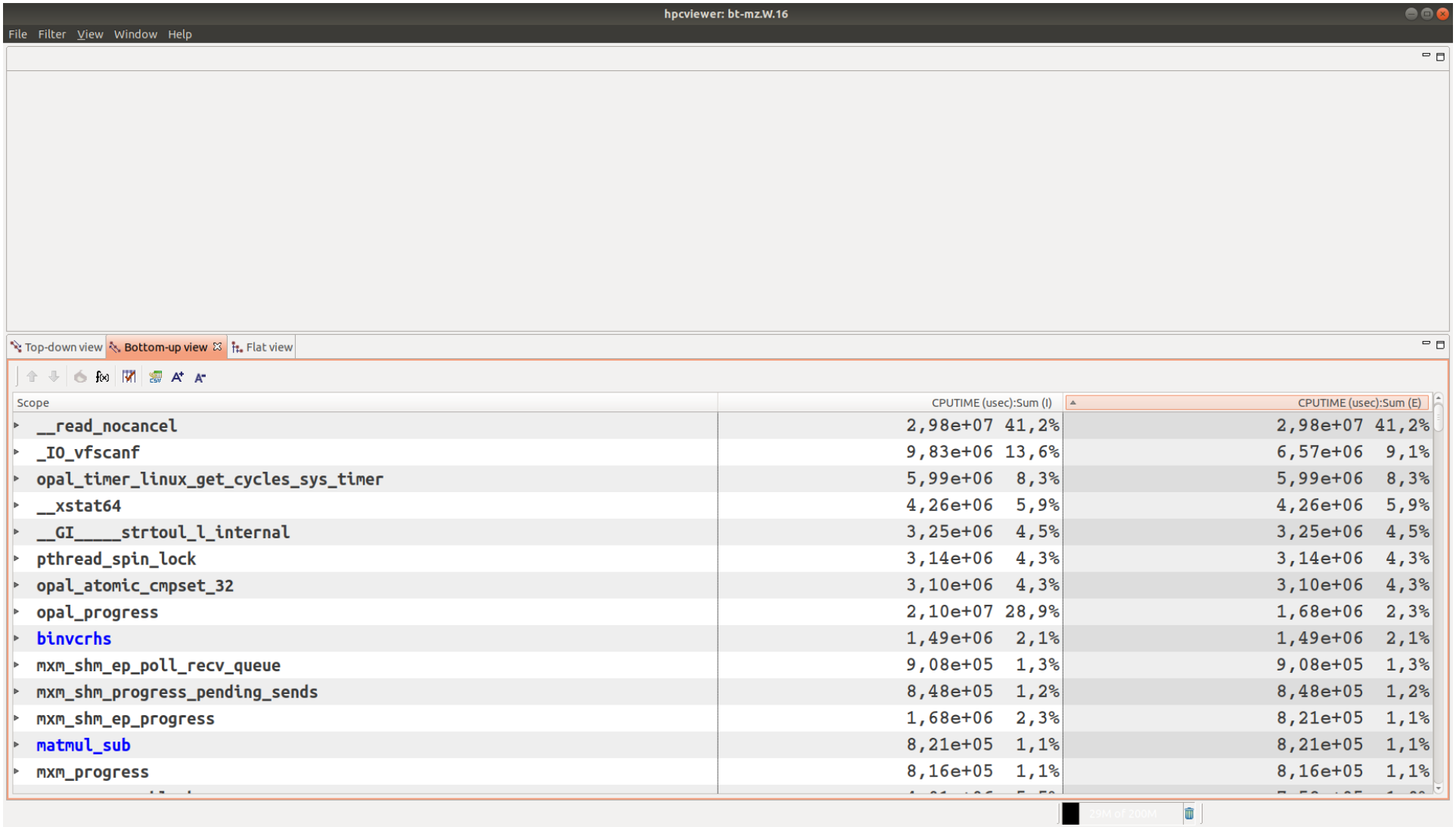
Visualizando no **hpcviewer**

-nodes=1 -ntasks=16 / Top-down view



Visualizando no **hpcviewer**

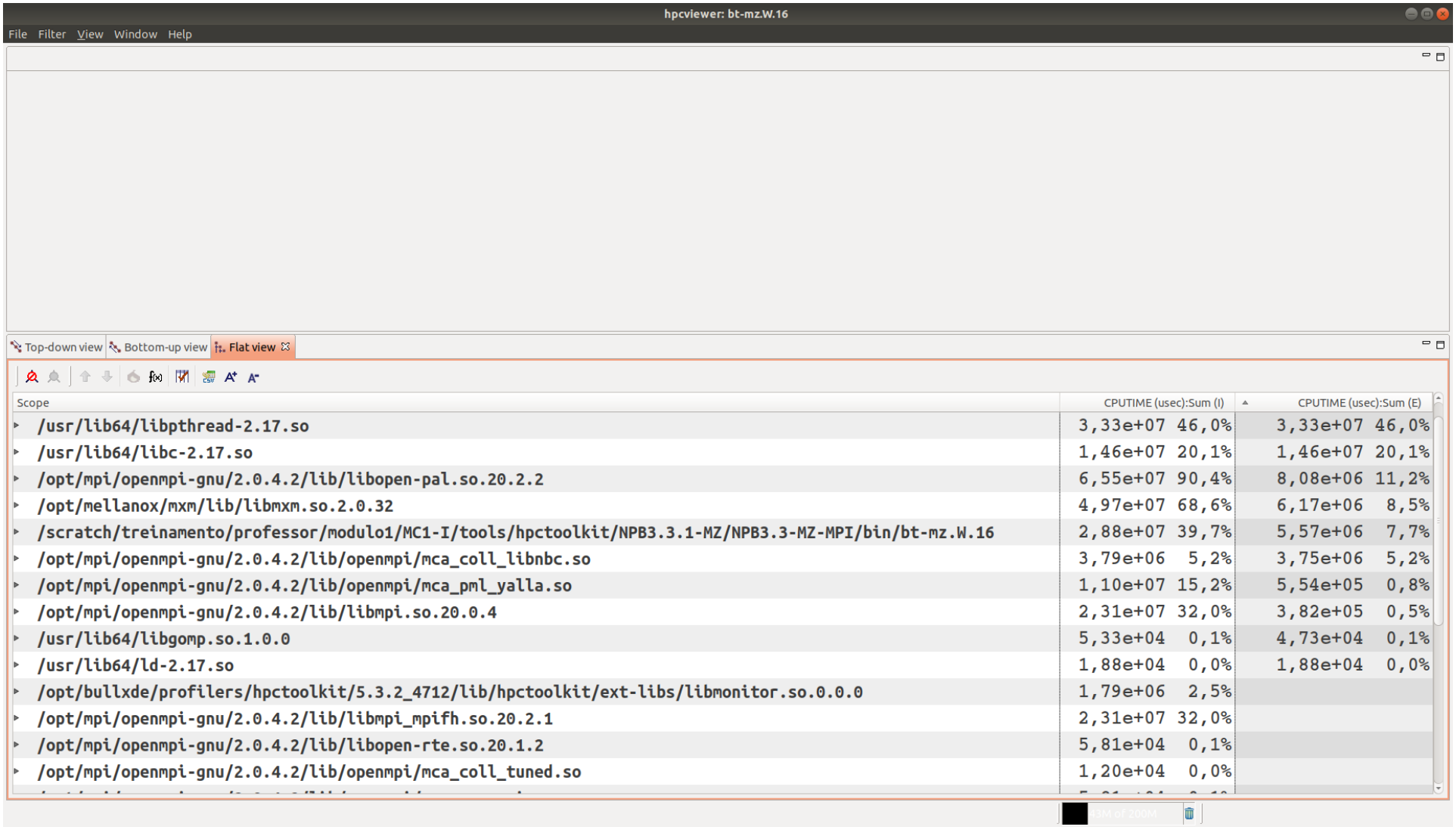
-nodes=1 -ntasks=16 / Bottom-up view



Scope	CPUTIME (usec):Sum (I)	CPUTIME (usec):Sum (E)
__read_nocancel	2,98e+07 41,2%	2,98e+07 41,2%
_IO_vfscanf	9,83e+06 13,6%	6,57e+06 9,1%
opal_timer_linux_get_cycles_sys_timer	5,99e+06 8,3%	5,99e+06 8,3%
__xstat64	4,26e+06 5,9%	4,26e+06 5,9%
__GI_____strtoul_l_internal	3,25e+06 4,5%	3,25e+06 4,5%
pthread_spin_lock	3,14e+06 4,3%	3,14e+06 4,3%
opal_atomic_cmpset_32	3,10e+06 4,3%	3,10e+06 4,3%
opal_progress	2,10e+07 28,9%	1,68e+06 2,3%
binvcrhs	1,49e+06 2,1%	1,49e+06 2,1%
mxm_shm_ep_poll_recv_queue	9,08e+05 1,3%	9,08e+05 1,3%
mxm_shm_progress_pending_sends	8,48e+05 1,2%	8,48e+05 1,2%
mxm_shm_ep_progress	1,68e+06 2,3%	8,21e+05 1,1%
matmul_sub	8,21e+05 1,1%	8,21e+05 1,1%
mxm_progress	8,16e+05 1,1%	8,16e+05 1,1%

Visualizando no `hpcviewer`

`-nodes=1 -ntasks=16 / Flat view`



Scope	CPUTIME (usec):Sum (I)	CPUTIME (usec):Sum (E)
/usr/lib64/libpthread-2.17.so	3,33e+07 46,0%	3,33e+07 46,0%
/usr/lib64/libc-2.17.so	1,46e+07 20,1%	1,46e+07 20,1%
/opt/mpi/openmpi-gnu/2.0.4.2/lib/libopen-pal.so.20.2.2	6,55e+07 90,4%	8,08e+06 11,2%
/opt/mellanox/mxm/lib/libmxm.so.2.0.32	4,97e+07 68,6%	6,17e+06 8,5%
/scratch/treinamento/professor/modulo1/MC1-I/tools/hpctoolkit/NPB3.3.1-MZ/NPB3.3-MZ-MPI/bin/bt-mz.W.16	2,88e+07 39,7%	5,57e+06 7,7%
/opt/mpi/openmpi-gnu/2.0.4.2/lib/openmpi/mca_coll_libnbc.so	3,79e+06 5,2%	3,75e+06 5,2%
/opt/mpi/openmpi-gnu/2.0.4.2/lib/openmpi/mca_pml_yalla.so	1,10e+07 15,2%	5,54e+05 0,8%
/opt/mpi/openmpi-gnu/2.0.4.2/lib/libmpi.so.20.0.4	2,31e+07 32,0%	3,82e+05 0,5%
/usr/lib64/libgomp.so.1.0.0	5,33e+04 0,1%	4,73e+04 0,1%
/usr/lib64/ld-2.17.so	1,88e+04 0,0%	1,88e+04 0,0%
/opt/bullxde/profilers/hpctoolkit/5.3.2_4712/lib/hpctoolkit/ext-libs/libmonitor.so.0.0.0	1,79e+06 2,5%	
/opt/mpi/openmpi-gnu/2.0.4.2/lib/libmpi_mpifh.so.20.2.1	2,31e+07 32,0%	
/opt/mpi/openmpi-gnu/2.0.4.2/lib/libopen-rte.so.20.1.2	5,81e+04 0,1%	
/opt/mpi/openmpi-gnu/2.0.4.2/lib/openmpi/mca_coll_tuned.so	1,20e+04 0,0%	

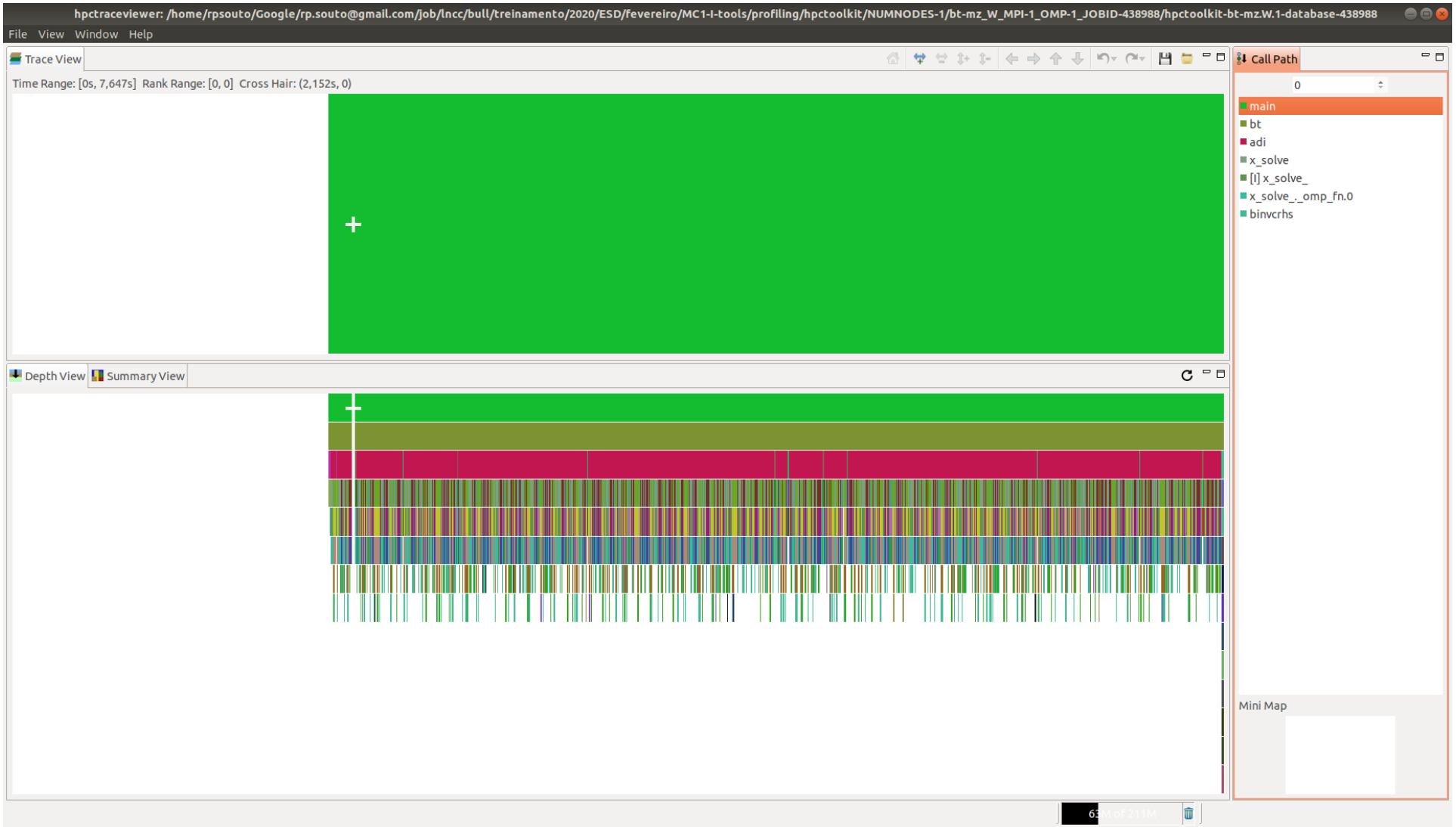
Visualizando no `hpctraceview`

NPB: estudo de caso

```
$ cd profiling/hpctoolkit/NUMNODES-1/bt-mz_W_MPI-1_OMP-1_JOBID-438988
$ hpctraceview hpctoolkit-bt-mz.W.1-database-438988/
```

Visualizando no `hpctraceview`

`-nodes=1 -ntasks=1`



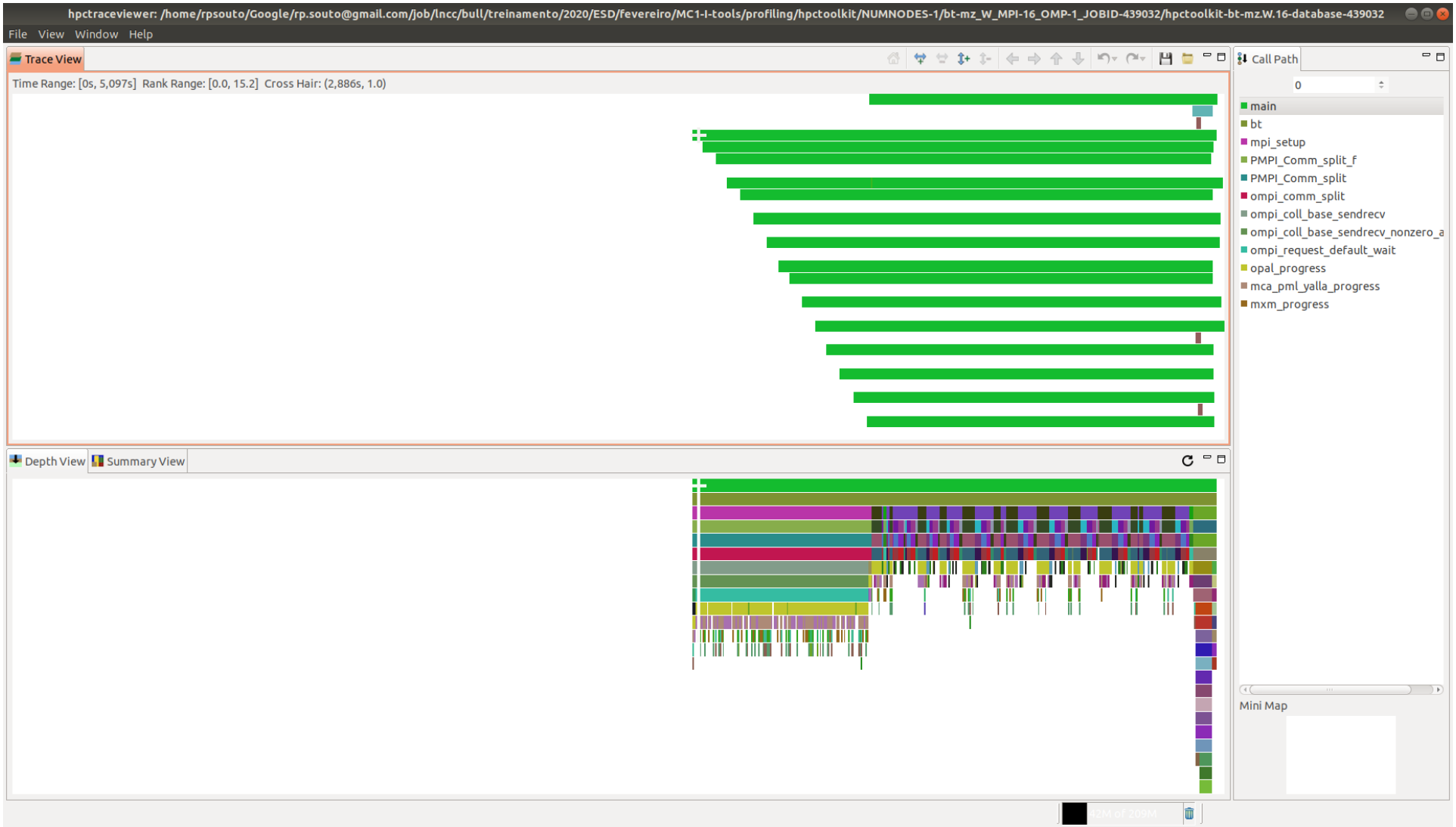
Visualizando no `hpctraceview`

NPB: estudo de caso

```
$ cd profiling/hpctoolkit/NUMNODES-1/bt-mz_W_MPI-16_OMP-1_JOBID-439032
$ hpctraceview hpctoolkit-bt-mz.W.16-database-439032
```

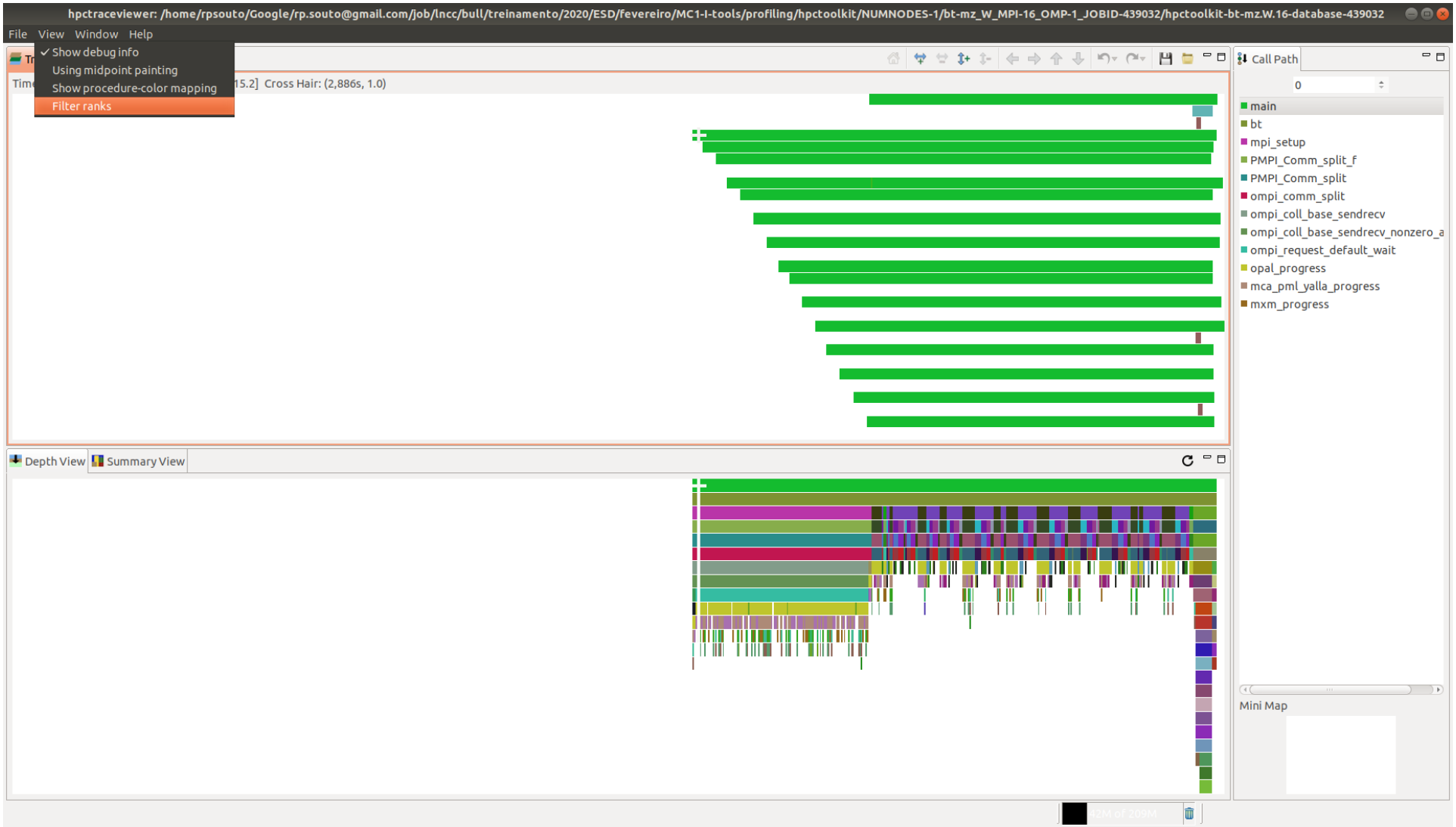
Visualizando no `hpctraceview`

`-nodes=1 -ntasks=16`



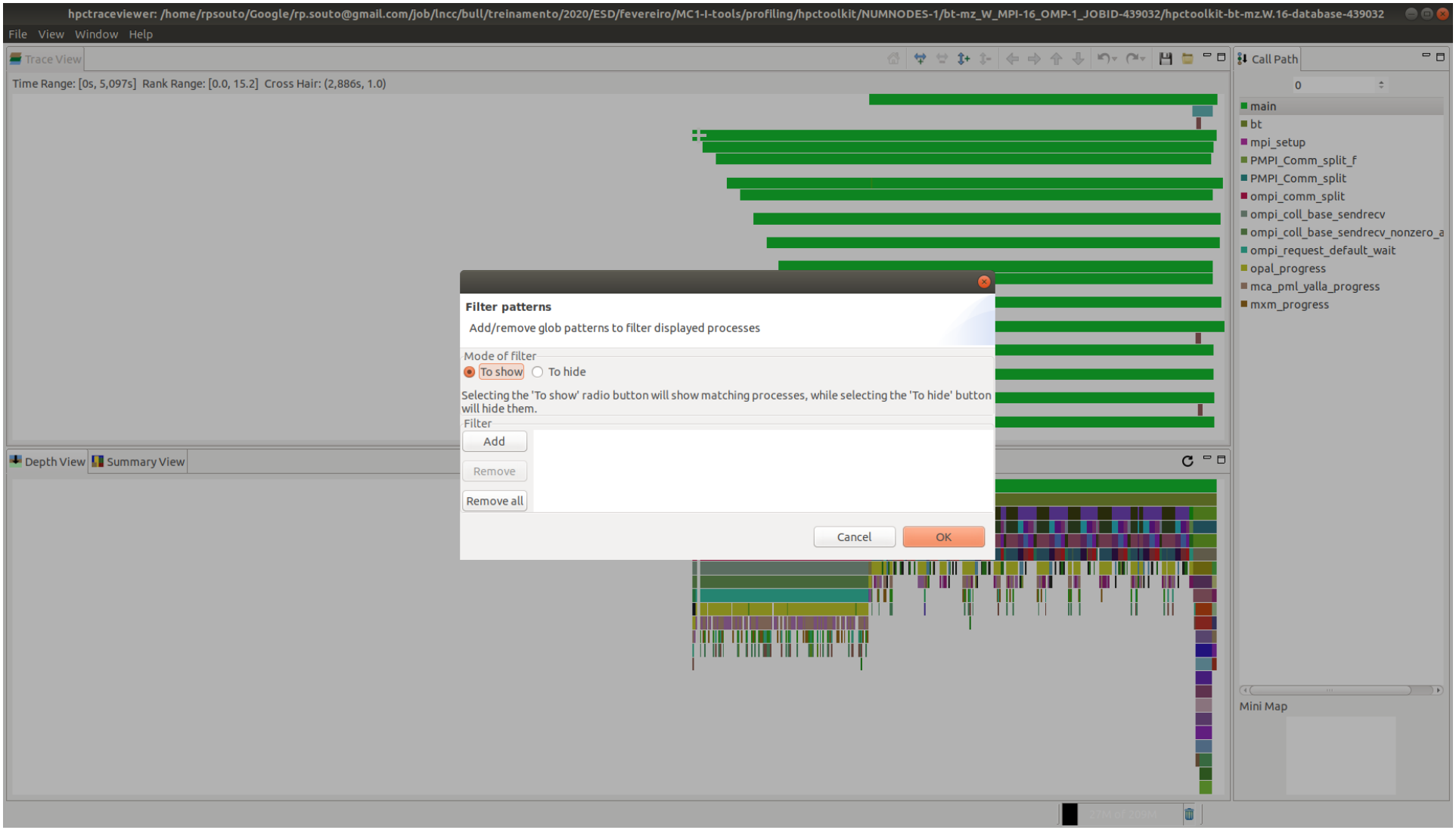
Visualizando no `hpctraceview`

`-nodes=1 -ntasks=16`



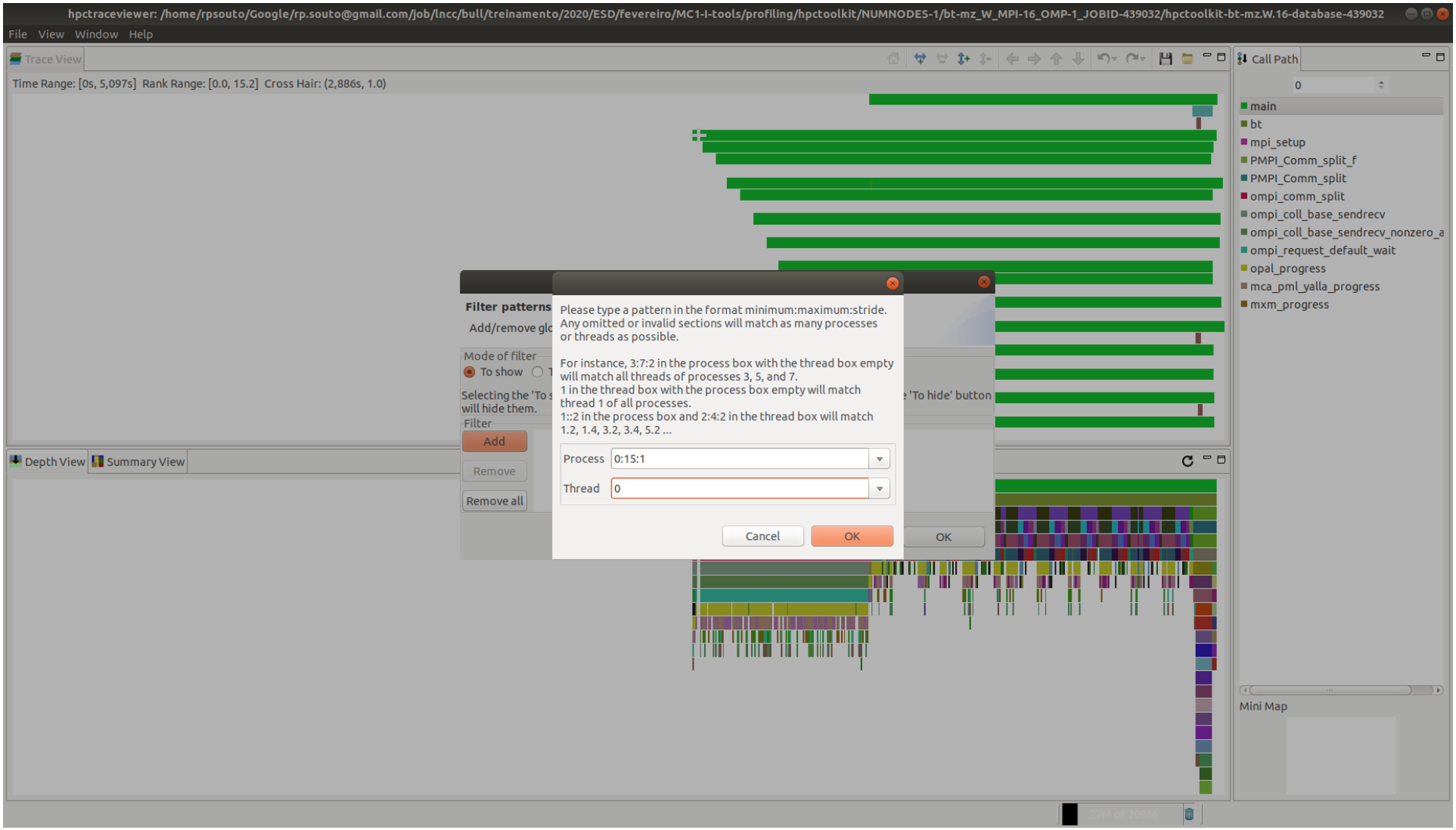
Visualizando no `hpctraceview`

`-nodes=1 -ntasks=16`



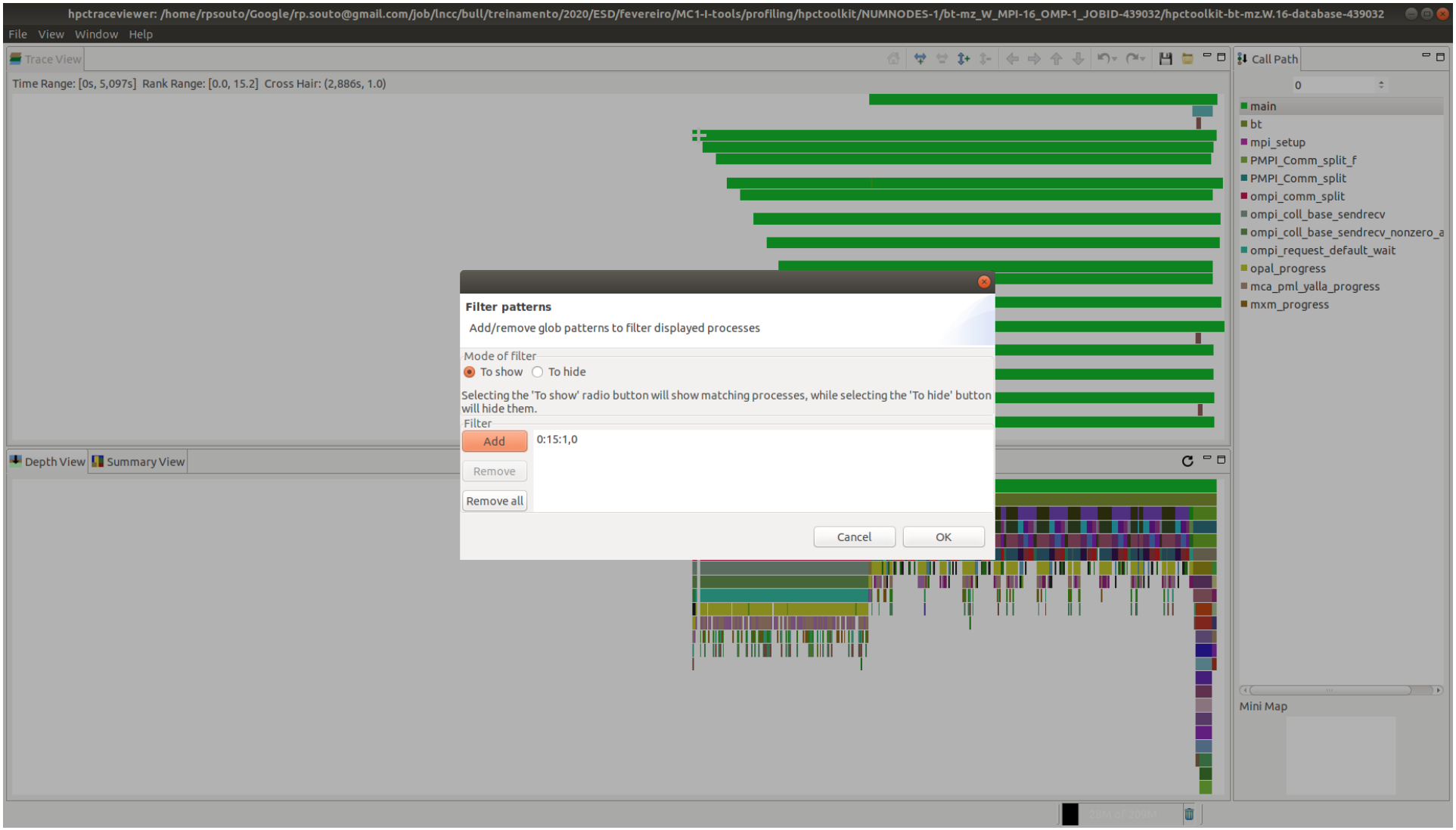
Visualizando no `hpctraceview`

`-nodes=1 -ntasks=16`



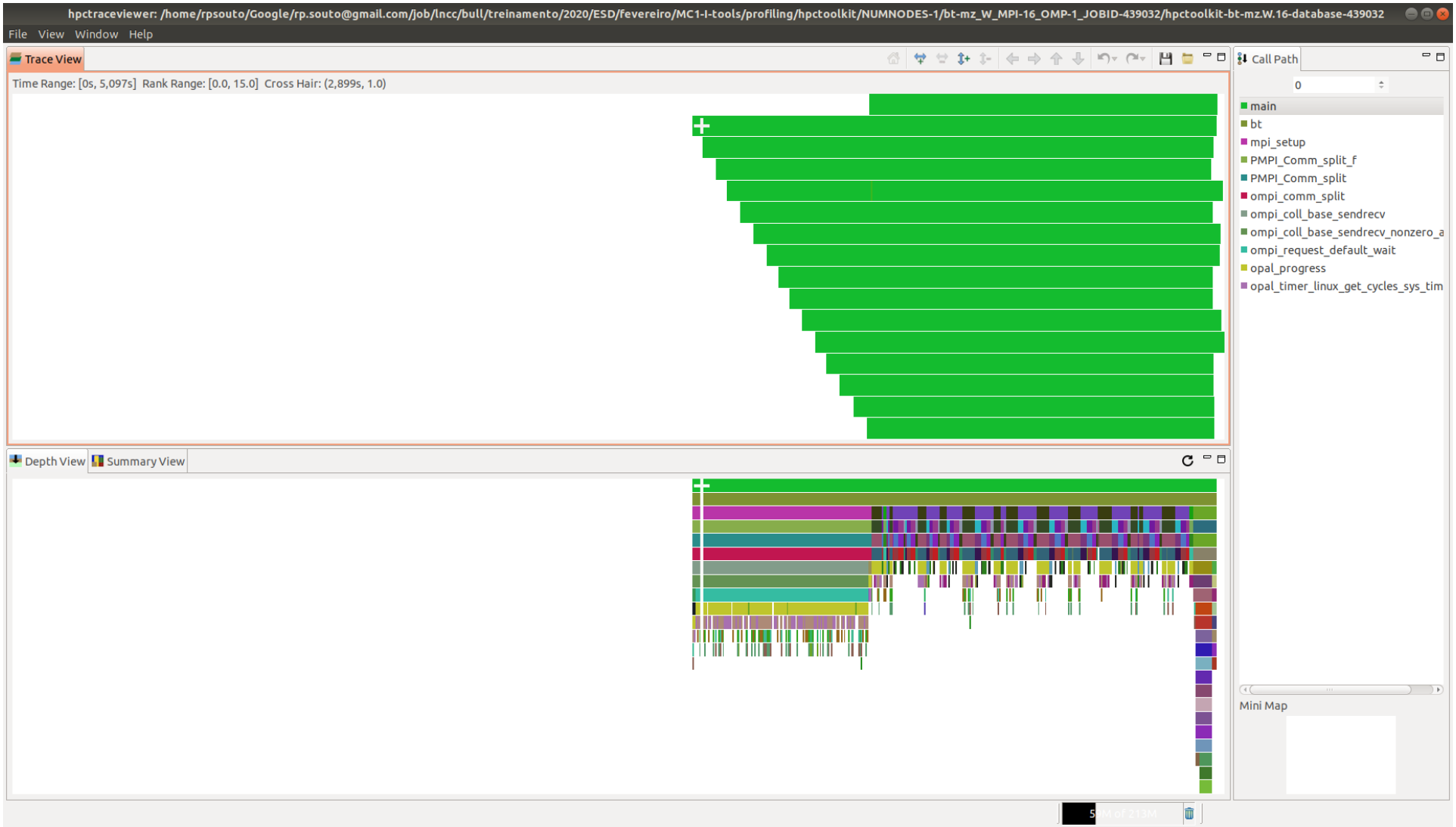
Visualizando no `hpctraceview`

`-nodes=1 -ntasks=16`



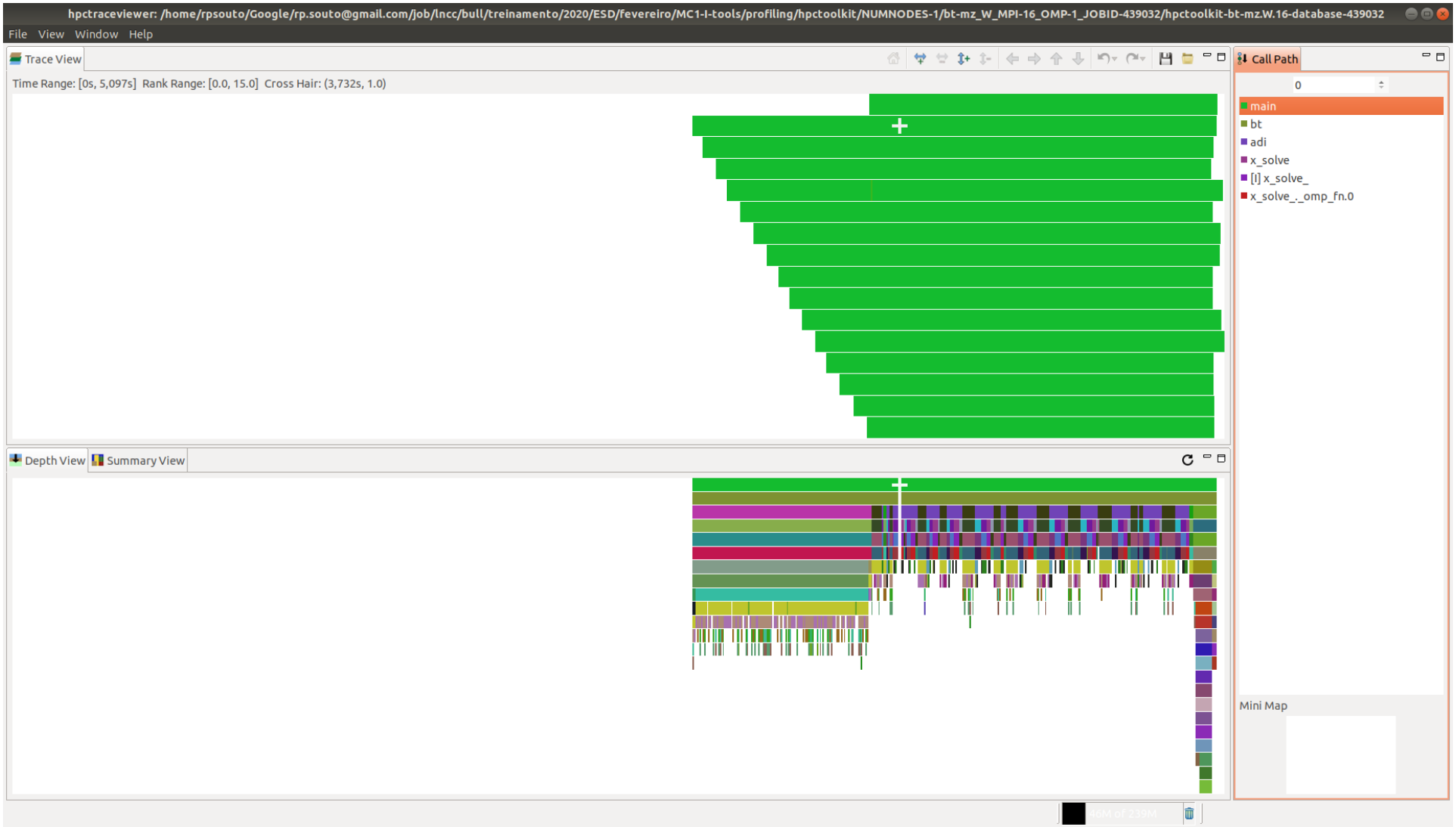
Visualizando no `hpctraceview`

`-nodes=1 -ntasks=16`



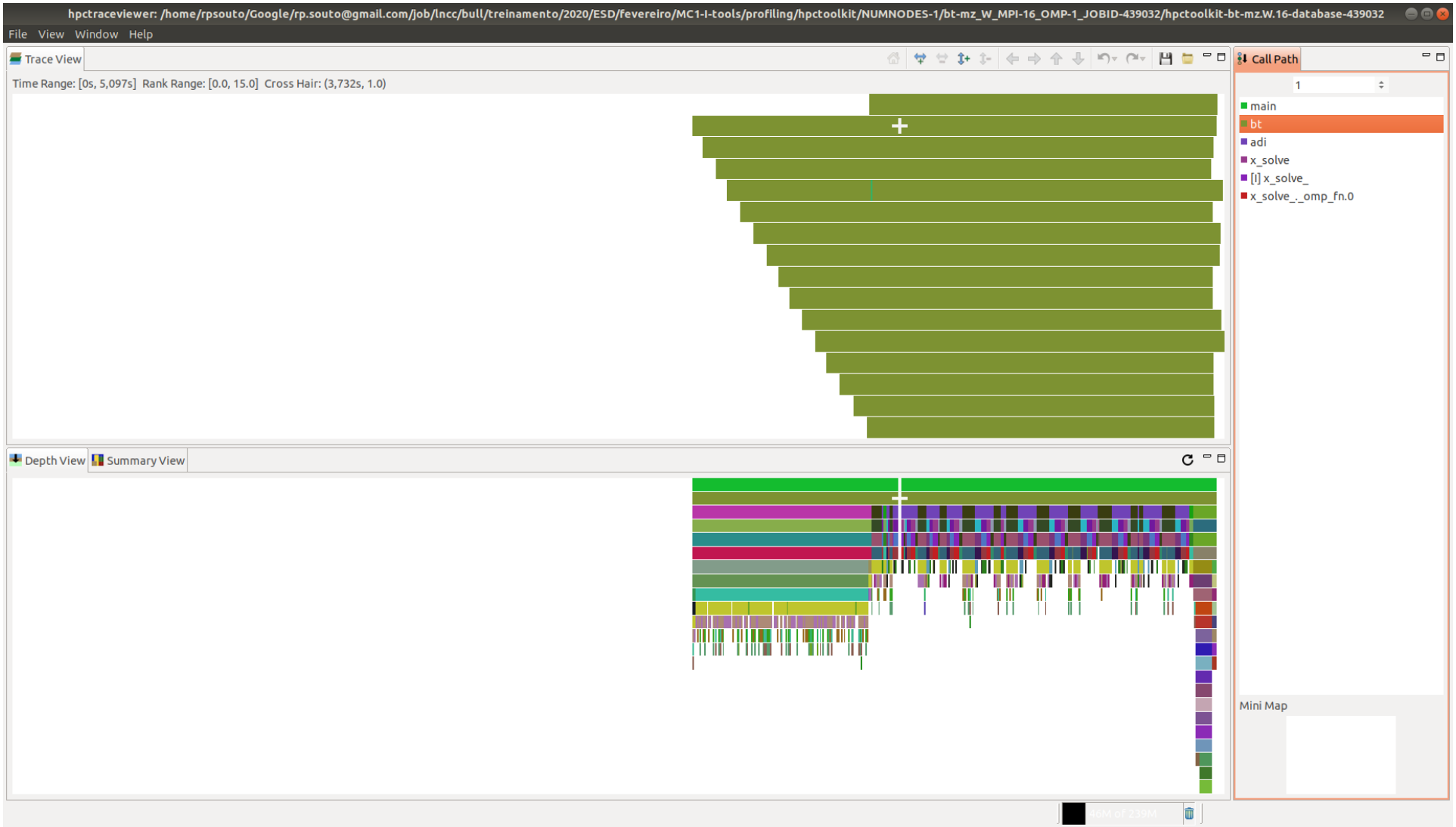
Visualizando no **hpctraceview**

-nodes=1 -ntasks=16 – função **main**



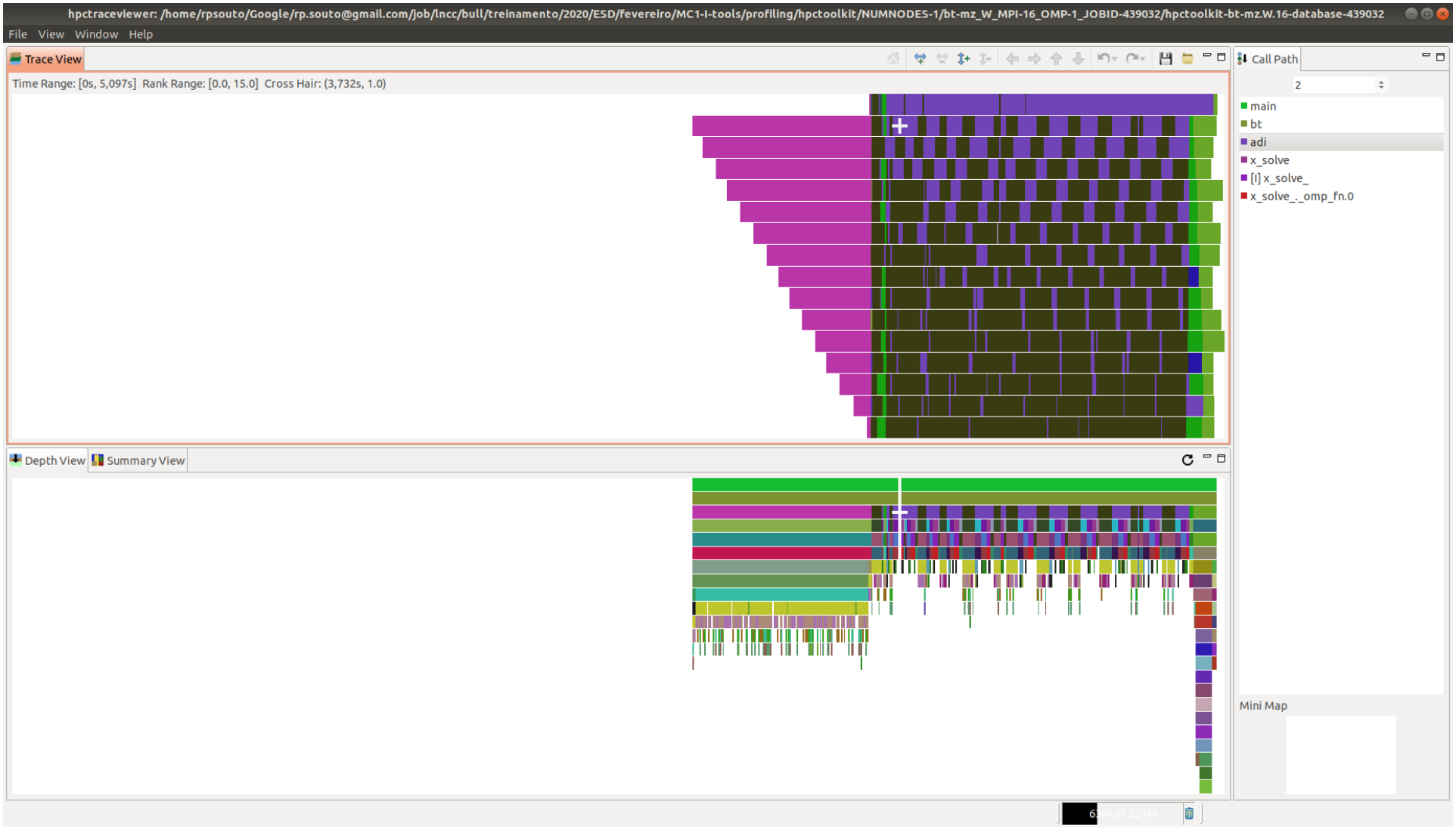
Visualizando no **hpctraceview**

-nodes=1 -ntasks=16 – função **bt**



Visualizando no **hpctraceview**

-nodes=1 -ntasks=16 – função **adi**



Visualizando no **hpctraceview**

-nodes=1 -ntasks=16 – função **xsolve**