UNIVERSITAT JAUME•I

Test at University Jaume I using a small cluster of Bull servers

Rafael Mayo

HPC&A
High Performance Computing and Architectures
http://www.hpca.uji.es

Experimentation system

Small cluster of Bull NovaScale R423 E2 servers composed of:

- 9 servers with 2 intel Xeon E5520 quad-core processors at 2,27 GHz.
  - Total: 9 x 2 x 4 = 72 cores
  - Theoretical peak: 72 x 4 x 2,27 = 643,76 Gflops
- Each server equiped with a Tesla C2050 GPU
- Infiniband QDR with a Mellanox MTS3600 switch

Experimentation system

Power measurement subsystem:

- APC 8653 PDU with per socket measurement with 1 sample per second
- pmlib: Power measurement library developed at our group:
  - pm_init counter
  - pm_start_counter
  - pm_stop_counter
  - pm_get_counter

Experimentation system

HPL distribution in MKL library.

```
pm_start_counter(&pm_counter);
system("sleep 10");
system("./runme_intel64");
system("sleep 10");
pm_stop_counter(&pm_counter);
```

Begin power measurement
10 sec. before HPL
HPL
10 sec. after HPL
End power measurement

Experimentation system

HPL distribution in MKL library.

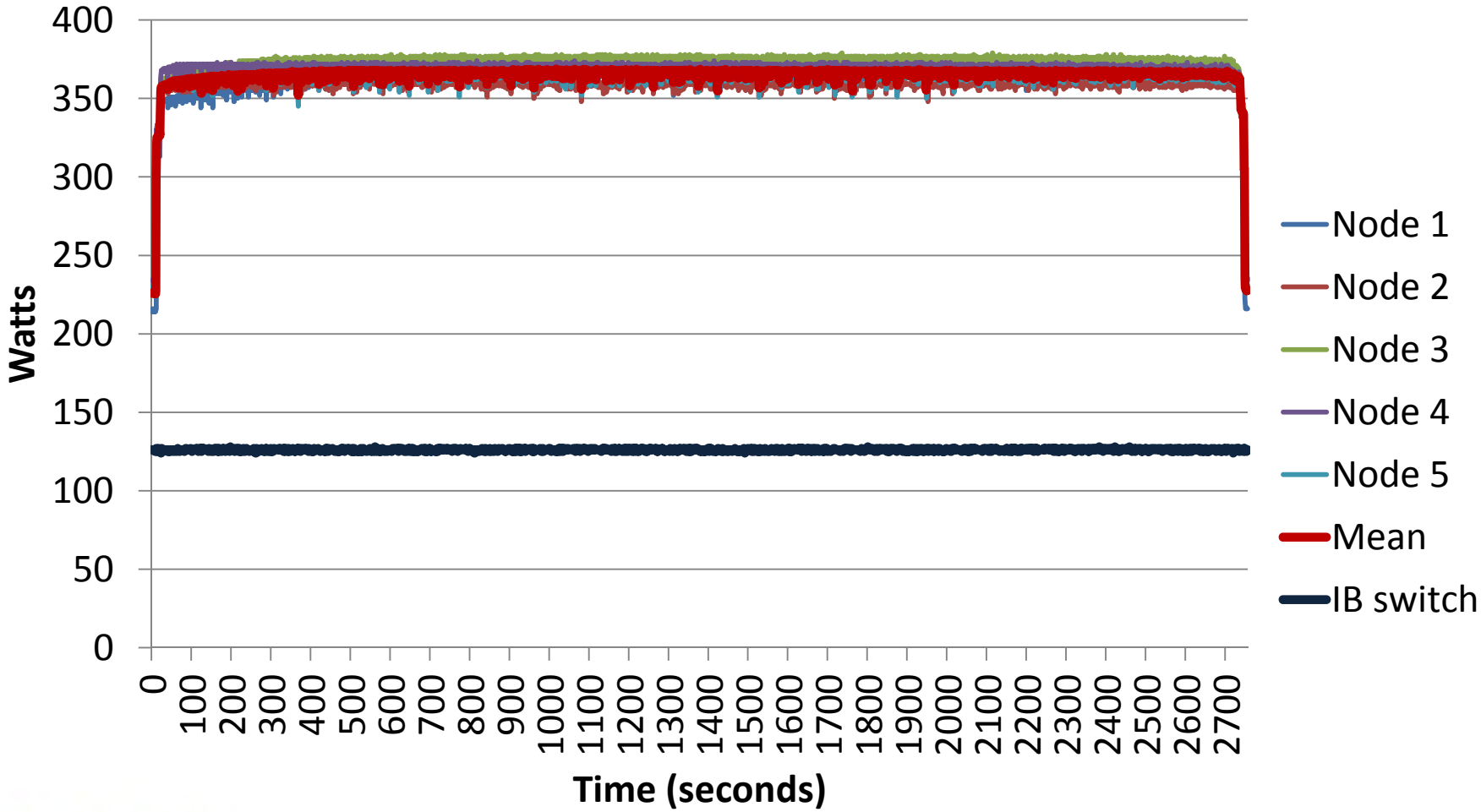
```
================================================================================
T/V                N    NB     P     Q               Time                 Gflops
--------------------------------------------------------------------------------
WR11R2C2      140000   168     8     9            3004.06            6.08964e+02
--------------------------------------------------------------------------------
||Ax-b||_oo/(eps*(||A||_oo*||x||_oo+||b||_oo)*N)=        0.0022672 ...... PASSED
================================================================================

Finished      1 tests with the following results:
              1 tests completed and passed residual checks,
              0 tests completed and failed residual checks,
              0 tests skipped because of illegal input values.
--------------------------------------------------------------------------------

End of Tests.
```

608,94 over 643,76 (94% of peak performance)

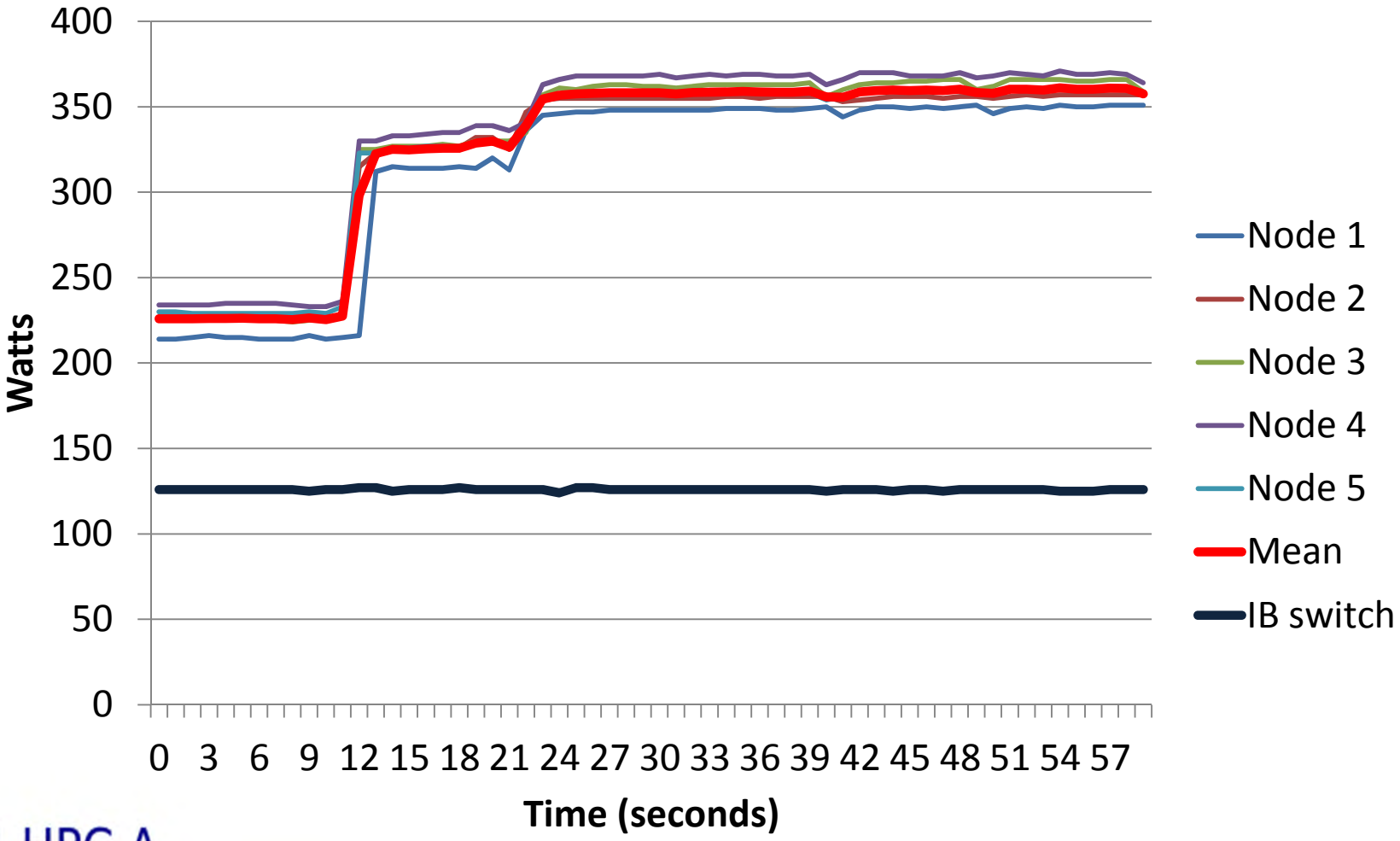
Measurements

- Connected to the PDU we have
  - 5 of the 9 nodes
  - The IB switch
- So, we do the work for the level 2 of the propossal.

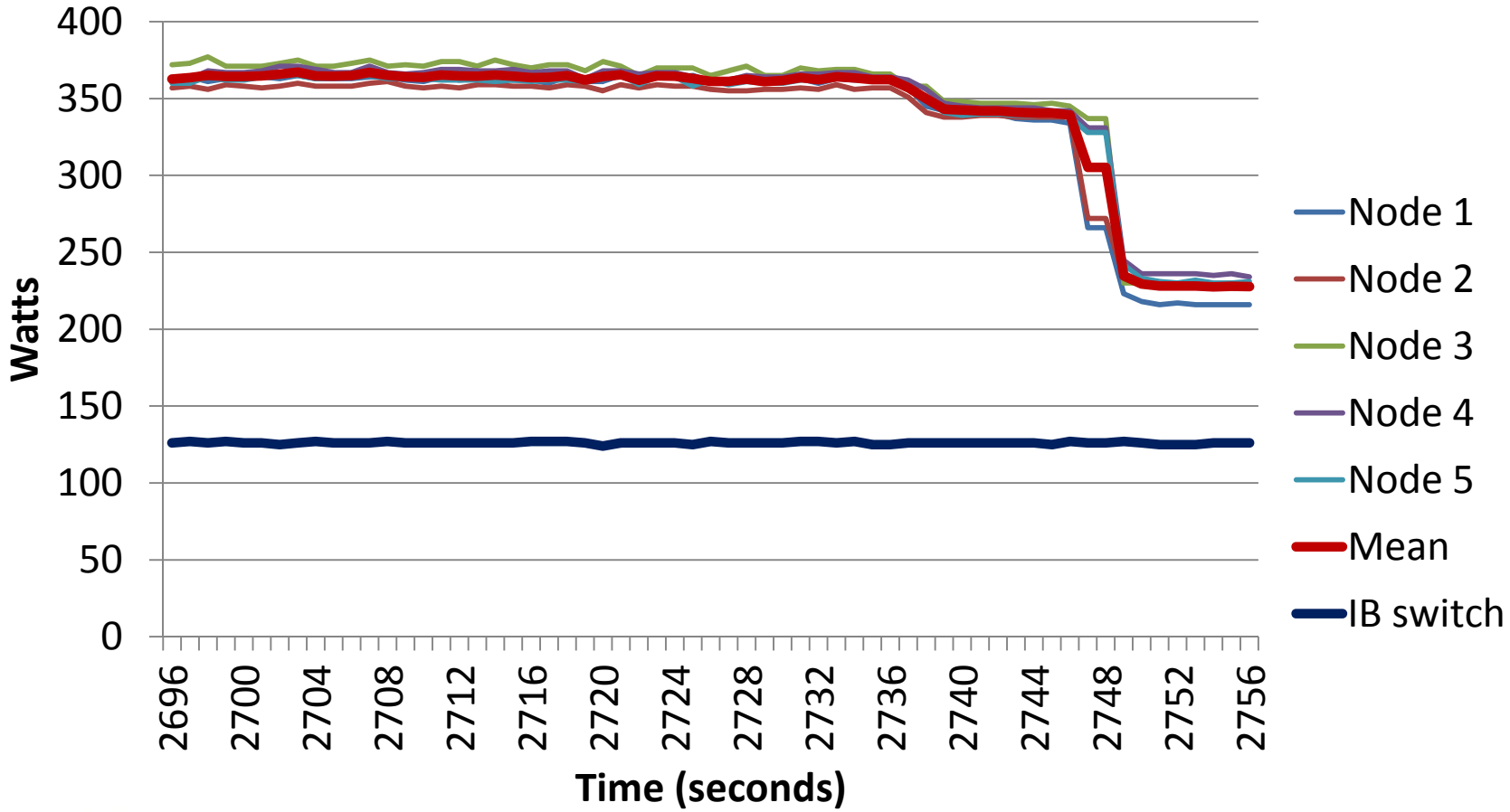
Measurements



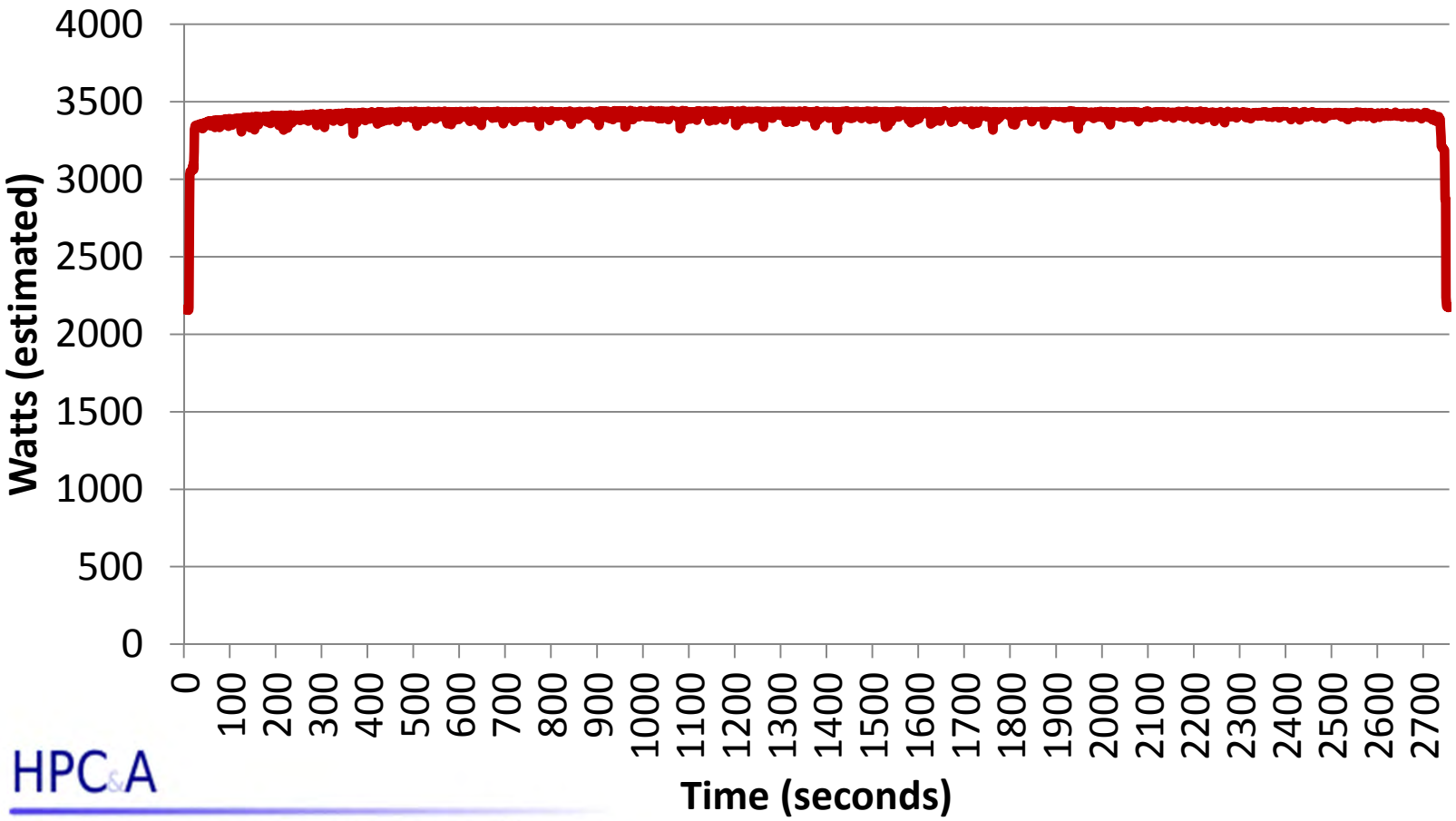
Detail of the first minute



Detail of the last minute



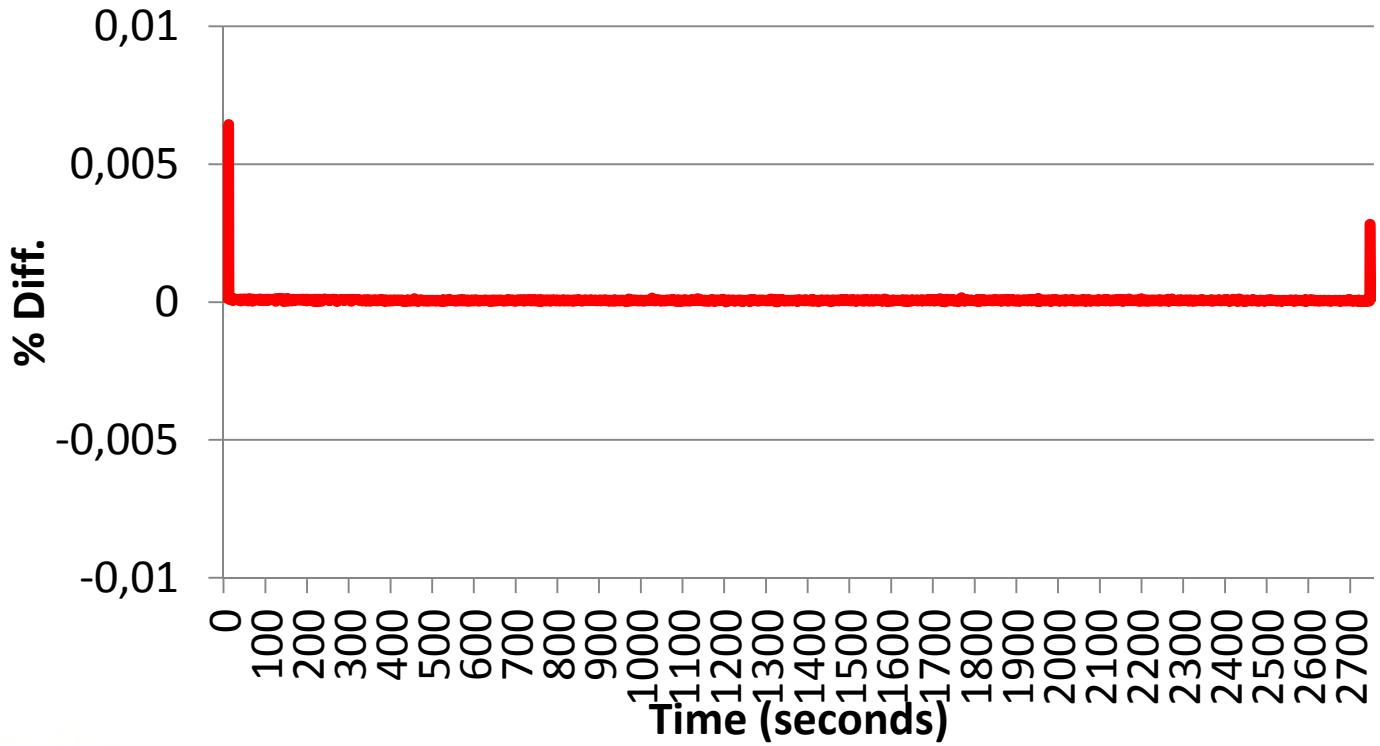
Estimated power consumption of the full system



Estimated power consumption of the full system

W(t) = 9 x Mean(t) + IB Switch (t)

W(t) = $\sum \mathrm{node_i}(t)$ x 4* Mean(t) + IB Switch (t)



Measurements File

Output of our measurement library

out.txt: Bloc de notas

Archivo Edición Formato Ver Ayuda

e	Line 1	Line 2	Line 6	Line 20	Line 21	Line 22	Sum
00000	214.000000	227.000000	126.000000	225.000000	234.000000	230.000000	1256.000000
04963	214.000000	227.000000	126.000000	225.000000	234.000000	230.000000	1256.000000
09927	215.000000	227.000000	126.000000	225.000000	234.000000	229.000000	1256.000000
14890	216.000000	227.000000	126.000000	225.000000	234.000000	229.000000	1257.000000
19854	215.000000	227.000000	126.000000	225.000000	235.000000	229.000000	1257.000000
24817	215.000000	228.000000	126.000000	225.000000	235.000000	229.000000	1258.000000
29780	214.000000	227.000000	126.000000	225.000000	235.000000	229.000000	1256.000000
34744	214.000000	227.000000	126.000000	225.000000	235.000000	229.000000	1256.000000
39707	214.000000	227.000000	126.000000	224.000000	234.000000	229.000000	1254.000000
44671	216.000000	228.000000	125.000000	225.000000	233.000000	230.000000	1257.000000
049634	214.000000	227.000000	126.000000	225.000000	233.000000	229.000000	1254.000000
154597	215.000000	227.000000	126.000000	227.000000	236.000000	233.000000	1264.000000
259561	216.000000	315.000000	127.000000	325.000000	330.000000	323.000000	1636.000000
364524	312.000000	323.000000	127.000000	325.000000	330.000000	323.000000	1740.000000
469487	315.000000	325.000000	125.000000	327.000000	333.000000	326.000000	1751.000000
574451	314.000000	324.000000	126.000000	327.000000	333.000000	326.000000	1750.000000
679415	314.000000	325.000000	126.000000	327.000000	334.000000	327.000000	1753.000000
784378	314.000000	325.000000	126.000000	328.000000	335.000000	327.000000	1755.000000
889341	315.000000	326.000000	127.000000	327.000000	335.000000	326.000000	1756.000000
994305	314.000000	332.000000	126.000000	330.000000	339.000000	329.000000	1770.000000
099268	320.000000	332.000000	126.000000	330.000000	339.000000	329.000000	1776.000000
204231	313.000000	325.000000	126.000000	330.000000	336.000000	328.000000	1758.000000
309195	336.000000	347.000000	126.000000	335.000000	341.000000	336.000000	1821.000000
414158	345.000000	354.000000	126.000000	357.000000	363.000000	354.000000	1899.000000
519121	346.000000	355.000000	124.000000	361.000000	366.000000	356.000000	1908.000000
624084	347.000000	355.000000	127.000000	360.000000	368.000000	358.000000	1915.000000
729048	347.000000	355.000000	127.000000	362.000000	368.000000	358.000000	1917.000000
834011	348.000000	355.000000	126.000000	363.000000	368.000000	357.000000	1917.000000
938974	348.000000	355.000000	126.000000	363.000000	368.000000	357.000000	1917.000000
043938	348.000000	355.000000	126.000000	362.000000	368.000000	358.000000	1917.000000
148903	348.000000	355.000000	126.000000	362.000000	369.000000	357.000000	1917.000000
253868	348.000000	355.000000	126.000000	361.000000	367.000000	357.000000	1914.000000
358833	348.000000	355.000000	126.000000	362.000000	368.000000	359.000000	1918.000000

Suggestions

Should there be a specification?

A proposal: CSV file

- First line some description and the mesurement unit.
- Next lines the values

```
Time (sec.);Node 1 (W); ….; Node n (W); IB switch (W);….
1;234,34;…;237,21;138,03;…
2;233,21;…;236,35;137,93;…
```

Suggestions

What happens with heterogeneous systems?

W(t) = 9 x Mean(t) + IB Switch (t)

W(t) = $\sum \mathrm{node_i(t)}$ x 4* Mean(t) + IB Switch (t)

The expressions are valid only for homogeneous systems

A proposal: For L2 at least one element (node) of each type must be measured

UNIVERSITAT JAUME·I

Thanks for your attention!!!

HPC&A

High Performance Computing and Architectures

http://www.hpca.uji.es