

1.1 PSNAP

1.1.1 Summary

P-SNAP [5] is a system benchmark for quantifying operating system interference or noise. It consists of a spin loop which is calibrated to take a given amount of time (typically 1ms). This loop is repeated for a number of iterations. The actual time each iteration takes is recorded.

1.1.2 TestCase

1) Preparation and Compilation

The code is available on PRACE SVN.

(<https://prace.osd.sara.nl/svn/trunk/pracewp74a/PABS/applications/psnap/>)

After downloading the code, provide other JuBE run xml files (same as the case of HPCC):

- Edit/add CURIE setup on compile.xml: Compilation options are the same as HPCC
- Create the running command file 'Nbench-Intel-Nehalem-CURIE.xml'
- Edit execute.xml

2) Parameters

P-SNAP should be run one psnap process per CPU (core) on each node. To obtain an accurate sample of the system noise a typical run would be for 1000000 repetitions.

- n <reps> number of repetitions, default: 100000
- w <reps> number of warm-up repetitions, default: 10% of the number of reps
- c <count> calibration count, default: perform a calibration to match granularity
- g <usecs> granularity of the test in microseconds, default: 1000

At the beginning of a run the program calibrates the timing loop to take <granularity> microseconds. As the calibration is running, the program outputs lines such as:

```
# count = 736377, time = 1001
```

At the beginning of each histogram a line appears with a '#' in the first column. The columns are MPI rank, correlation, total time, hostname, as follows:

```
# 1 0.009036 1002103 pal0
```

The correlation is computed between the time series measured on each process and the time series measured on MPI rank 0. If the noise is correlated in time across the system, the correlation should be close to one. The total time is the total amount of time taken to run the loops.

You can get a rough idea of the uniformity of the processors from the initial calibration. All of the procs should have a similar count. grep for "#rank" to examine the individual counts. The summary of the histograms can be combined by node name and a % slowdown can be calculated, a script like:

D7.4.2 Benchmarking and Performance Modelling on Tier-0 systems

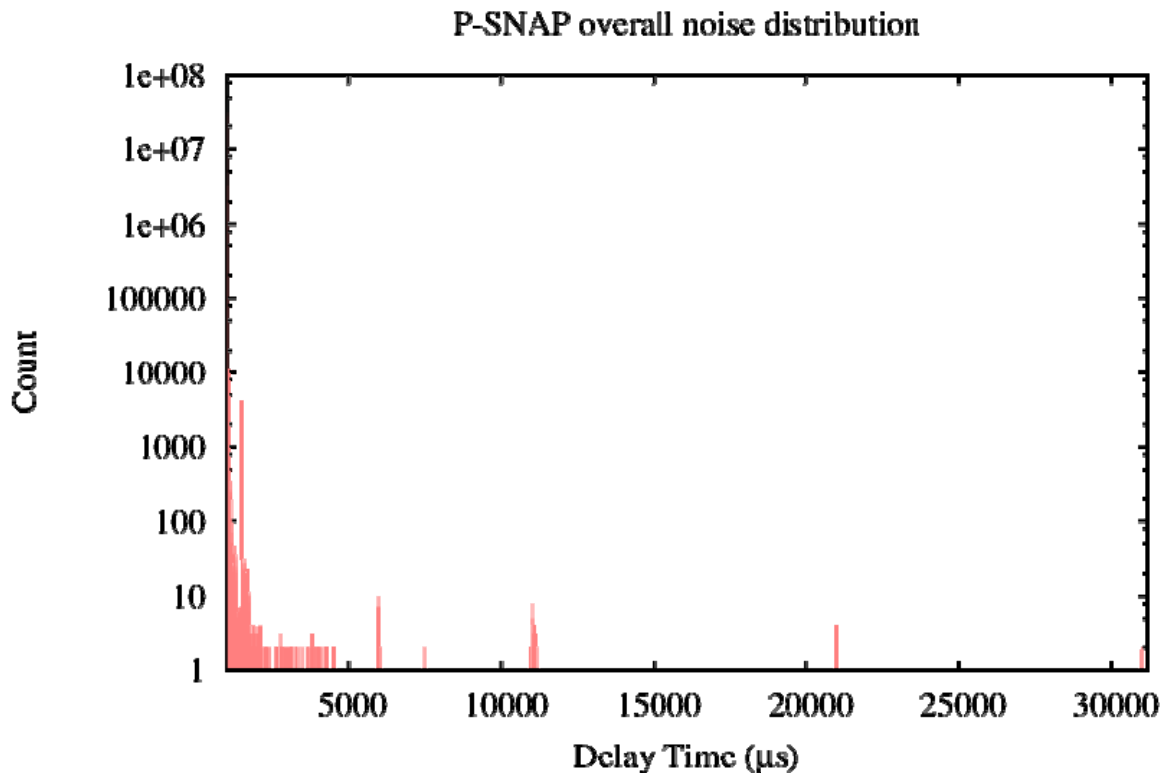
```
grep "#\ " psnap.out | awk ' {sum[$NF]+=$4} END {for (m in sum) {print m, sum[m], (1.0 - 400000000/sum[m])*100 } }' | sort -k2 # (assumes 4 ranks per OS, and psnap run with defaults)
```

1.1.3 Results and Analysis

The operating system noise is measured with the P-SNAP benchmark [ref to P-SNAP] which runs a calibrated task and then reports the actual elapsed time of the task. The difference between the expected calibrated time and the actual time is a measure of operating system noise.

The idea is to run PSNAP on *all* nodes. This was not possible for the person performing the benchmark, because he is an ordinary user with no possibilities to select *all* nodes for the jobs. Therefore, the histogram depicting noise per core over the *whole* system will not be present. Jobs with 1, 2, 4, 8 and 16 cores were run on node curie1257. The job with 32 cores was run on node curie1351.

Nodes	Tasks	Calibration Time (s)	Mean Run Time (s)	Standard Deviation of Run Time (s)	Operating System Noise (%)
1	1,2,4,8,16 & 32	1000.00	1001.22	25.49	0.12



D7.4.2 Benchmarking and Performance Modelling on Tier-0 systems

In detail (per case), results are as follows:

Number of Cores	System Correlation	Noise (%)	Runtime (S.D.)
1	1	0.0945496	0
2	0.006209	0.09027095	0.024538503
4	0.007913667	0.091571225	0.028555045
8	0.002753286	0.091499412	0.029876253
16	0.000754533	0.147450294	1.774546325
32	0.000944774	0.124535631	0.957559152

From the result, we detect that the correlation between cores (system-level noise which affect all processes) are very small. System noise is around 0.14% in the highest case. Standard deviation of runtime between processors (different from above result which describes the noise among individual iteration) is in 1 second level.

References

- [1] HPCC: <http://icl.cs.utk.edu/hpcc/>
- [2] Report on available performance analysis and benchmark tools, PRACE Preparatory Phase Deliverable D6.3.1, November 2008.
- [3] Technical Assessment Report of Prototype Systems, PRACE Preparatory Phase Deliverable D5.2, December 2009
- [4] JuBE framework, <http://www2.fz-juelich.de/jsc/jube/>
- [5] P-SNAP homepage, <http://www.ccs3.lanl.gov/pal/software/psnap/>
- [6] mixedMode, <http://www2.epcc.ed.ac.uk/~markb/mapiopenmpbench/intro.html>