IMPI Benchmark (ping-pong) to measure latency among MPI processes assigned on different cores

Nicole Orzan

February 11, 2018

The aim of today's exercise was to run intel MPI ping-pong benchmark among to processors within a node and try to estimate latency and bandwidth. I worked on node cn07-33.
To execute the IMPI benchmark, first of all I had to load the following module:

```
module load impi-trial/5.0.1.035
```

There are 2 ways we can measure the latency: inside the same socket or between two different sockets. Do do this we can use the "hwloc" command, which permits to bind a processes to a given CPU set.
To measure the latency inside the same socket we have to specify the numbers of 2 cores from the same socket, while to measure it between two sockets we have to specify the numbers of 2 cores from different sockets.
To know which cores are in socket 0 and which are in socket 1 we can use the command numactl –hardware. In this case I obtained the result:

```
available: 2 nodes (0-1)
node 0 cpus: 0 1 2 3 4 5 6 7 8 9
node 0 size: 163811 MB
node 0 free: 160034 MB
node 1 cpus: 10 11 12 13 14 15 16 17 18 19
node 1 size: 163840 MB
node 1 free: 159807 MB
node distances:
node   0   1
  0:  10  11
  1:  11  10
```

So for the latency inside socket 0 I used the command:

```
mpirun -np 2 hwloc-bind core:0 core:5 /u/shared/programs/x86_64/
intel/impi_5.0.1/bin64/IMB-MPI1 PingPong
```

While for the latency between two sockets I had to change the values of the cores:

```
hwloc-bind core:0 core:13
```

We can also specify the number of iterations to use writing "-iter num". Case cores 1-3:

```
#------------------------------------------------------------
# Benchmarking PingPing
# #processes = 2
#------------------------------------------------------------
       #bytes #repetitions      t[usec]   Mbytes/sec
            0         1000         0.37         0.00
            1         1000         0.41         2.34
            2         1000         0.48         3.96
            4         1000         0.40         9.42
            8         1000         0.41        18.70
           16         1000         0.44        34.76
           32         1000         0.43        70.84
```

```
            64          1000          0.41        150.68
           128          1000          0.46        263.65
           256          1000          0.47        521.65
           512          1000          0.55        892.76
          1024          1000          0.63       1540.43
          2048          1000          0.73       2671.89
          4096          1000          1.06       3677.67
          8192          1000          1.64       4775.28
         16384          1000          2.91       5362.13
         32768          1000          5.03       6207.82
         65536           640         15.65       3994.39
        131072           320         28.59       4371.57
        262144           160         54.07       4623.22
        524288            80        100.82       4959.13
       1048576            40        193.27       5174.00
       2097152            20        370.65       5395.99
       4194304            10        720.41       5552.43
```

Case cores 0-5:

```
#---------------------------------------------------
# Benchmarking PingPing
# #processes = 2
#---------------------------------------------------
       #bytes #repetitions      t[usec]   Mbytes/sec
            0          1000         0.33         0.00
            1          1000         0.39         2.46
            2          1000         0.42         4.52
            4          1000         0.41         9.32
            8          1000         0.39        19.36
           16          1000         0.40        38.05
           32          1000         0.47        64.26
           64          1000         0.47       128.77
          128          1000         0.51       240.26
          256          1000         0.50       492.07
          512          1000         0.57       850.85
         1024          1000         0.65      1502.02
         2048          1000         0.77      2523.72
         4096          1000         1.07      3647.37
         8192          1000         1.60      4879.09
        16384          1000         2.75      5677.55
        32768          1000         5.65      5532.80
        65536           640        14.77      4232.82
       131072           320        28.10      4448.30
       262144           160        53.26      4693.72
       524288            80        98.56      5072.94
      1048576            40       189.15      5286.83
      2097152            20       369.00      5420.05
      4194304            10       697.30      5736.39
```

Case cores 0-13:

```
#---------------------------------------------------
# Benchmarking PingPing
# #processes = 2
#---------------------------------------------------
       #bytes #repetitions      t[usec]   Mbytes/sec
            0          1000         0.57         0.00
            1          1000         0.59         1.63
            2          1000         0.59         3.25
            4          1000         0.64         6.00
```

8	1000	0.61	12.47
16	1000	0.60	25.60
32	1000	0.74	41.36
64	1000	0.84	72.40
128	1000	0.85	144.14
256	1000	0.83	295.53
512	1000	1.07	455.01
1024	1000	1.26	778.12
2048	1000	1.58	1235.41
4096	1000	2.43	1606.90
8192	1000	3.72	2100.11
16384	1000	6.45	2422.06
32768	1000	11.70	2671.18
65536	640	23.90	2614.74
131072	320	44.30	2821.69
262144	160	80.22	3116.47
524288	80	159.40	3136.81
1048576	40	317.85	3146.10
2097152	20	633.86	3155.27
4194304	10	1269.10	3151.83

We can clearly say that the latency is higher when we bind the process to work on 2 different sockets. Computing a mean using the lowest bytes values we obtain that latency between two cores inside in the same socket is 0.39 ± 0.31 μs, while between different sockets is 0.60 ± 0.23 μs.