

GROMACS - Feature #1086

implement the two frequently used thread-pinning strategies

12/19/2012 07:56 PM - Erik Lindahl

Status:	Closed
Priority:	Low
Assignee:	
Category:	mdrun
Target version:	4.6
Difficulty:	uncategorized
Description	
Linux default HT cores at the end, Win interleaved) and enable swithching between them (through a hidden command line option?) @runner.c:set_cpu_affinity()	

Associated revisions

Revision a1bd375a - 01/11/2013 07:05 PM - Erik Lindahl

Added basic CPU topology information to cpuid code

We can now detect the locality of hardware threads, cores, and packages for Intel and AMD CPUs under Linux and Windows. In particular, this provides an array with locality order for logical processors that can be used to optimize placement. Refs #1086, #1101.

Change-Id: I3f7985b1b67729376918c5a135b9157a9086235e

History

#1 - 01/02/2013 07:40 PM - Szilárd Páll

- *Category set to mdrun*

A bit more details are needed for this, i think, in case if someone could pitch in with the code.

With Intel HT, which we can partially detect now for the (probably) most frequent case, when nrankx nthreads = ncpus, there are two CPU layouts that we know of:

- physical cores first, HT siblings after on Linux (on a 4-core CPU 0123 0123);
- interleaved on WInows (on a 4-core CPU 00112233).

With NUMA hardware, especially when using OpenMP (which is the default with up to 16 threads/process on Intel) correct pinning is crucial. The current implementation uses interleaved HT pinning by default which will only work on Linux with kernels configured the standard way, but will be incorrect on Windows and on Linux with custom kernel configuration.

The suggested solution is to implement both HT pinning layouts and use the appropriate ones on different platforms and facilitate selecting the layout by e.g an advanced environment variable.

#2 - 01/07/2013 01:08 PM - Mark Abraham

Target this to 4.6.1?

#3 - 01/09/2013 12:10 AM - Szilárd Páll

- *Status changed from New to In Progress*

Mark Abraham wrote:

Target this to 4.6.1?

AFAIK Erik is working on it, so I'll change the state.

#4 - 01/09/2013 02:14 PM - Erik Lindahl

Have a look at <https://gerrit.gromacs.org/#/c/2000/>.

There is now a routine to get basic CPU topology information, including a locality-sorted list of processors, from the cpuid code. Currently this works for Intel and AMD processors under Linux and Windows (not Mac OS X, since we cannot to thread pinning there). The topology is only complicated when we have to detect it on the fly, so if there are other systems (BlueGene?) where we have a known static order of things it should only be 5 minutes of work to add an #ifdef for that case.

#5 - 01/09/2013 02:51 PM - Mark Abraham

BlueGeneQ is very probably pre-pinned, but at least for group kernels (where I understand OpenMP is not useful for PP processes) I expect BlueGeneQ will require an MPI process per thread.

With Verlet kernels we will have more scope to use OpenMP on BlueGeneQ, because both PME and PP nodes can use it.

#6 - 01/18/2013 03:57 PM - Erik Lindahl

- *Status changed from In Progress to Closed*

Fixed by gerrit 2051.