GROMACS - Bug #2504

gromacs 2018.1 doesn't run on KNL

05/17/2018 01:55 PM - Carlo Camilloni

Status:	Closed		
Priority:	Normal		
Assignee:	Roland Schulz		
Category:	analysis tools		
Target version:	2018.2		
Affected version - extra info:		**Difficulty:**	uncategorized
Affected version:	2018.1		

Description

I have tried to run a standard MD simulation on a KNL cluster but I get the following error:

```
===================================================================================
=   BAD TERMINATION OF ONE OF YOUR APPLICATION PROCESSES
=   PID 25280 RUNNING AT r065c04s03
=   EXIT CODE: 132
=   CLEANING UP REMAINING PROCESSES
=   YOU CAN IGNORE THE BELOW CLEANUP MESSAGES
===================================================================================

===================================================================================
=   BAD TERMINATION OF ONE OF YOUR APPLICATION PROCESSES
=   PID 25280 RUNNING AT r065c04s03
=   EXIT CODE: 4
=   CLEANING UP REMAINING PROCESSES
=   YOU CAN IGNORE THE BELOW CLEANUP MESSAGES
===================================================================================
   Intel(R) MPI Library troubleshooting guide:
      https://software.intel.com/node/561764
===================================================================================
```

The code is run as:

mpiexec -np 32 mdrun_knl -s topol0 -nb cpu -v -maxh 23.9 -nsteps -1 >& log

and is compiled with Intel 2017, the same happens with the intel 2018 and using the fftw instead of the mkl.

This is the log

md.2018.1.log:

```
ROMACS version:    2018.1
Precision:          single
Memory model:       64 bit
MPI library:        MPI
OpenMP support:     enabled (GMX_OPENMP_MAX_THREADS = 64)
GPU support:        disabled
SIMD instructions:  AVX_512_KNL
FFT library:        Intel MKL
RDTSCP usage:       enabled
TNG support:        enabled
Hwloc support:      hwloc-1.11.0
Tracing support:    disabled
Built on:           2018-05-17 11:33:17
Built by:           ccamillo@r000u06l01 [CMAKE]
Build OS/arch:      Linux 3.10.0-327.36.3.el7.x86_64 x86_64
Build CPU vendor:   Intel
Build CPU brand:    Intel(R) Xeon(R) CPU E5-2697 v4 @ 2.30GHz
```

```
Build CPU family:   6   Model: 79   Stepping: 1
Build CPU features: aes apic avx avx2 clfsh cmov cx8 cx16 f16c fma hle htt intel lahf mmx msr nons
top_tsc pcid pclmuldq pdcm pdpe1gb popcnt pse rdrnd rdtscp rtm sse2 sse3 sse4.1 sse4.2 ssse3 tdt x
2apic
C compiler:         /cineca/prod/opt/compilers/intel/pe-xe-2017/binary/bin/icc Intel 17.0.4.201704
11
C compiler flags:    -xMIC-AVX512   -mkl=sequential  -std=gnu99  -O3 -DNDEBUG -ip -funroll-all-loo
ps -alias-const -ansi-alias -no-prec-div -fimf-domain-exclusion=14 -qoverride-limits
C++ compiler:       /cineca/prod/opt/compilers/intel/pe-xe-2017/binary/bin/icpc Intel 17.0.4.20170
411
C++ compiler flags:  -xMIC-AVX512   -mkl=sequential  -std=c++11   -O3 -DNDEBUG -ip -funroll-all-lo
ops -alias-const -ansi-alias -no-prec-div -fimf-domain-exclusion=14 -qoverride-limits
```

(it stops here)

the same TPR with the same setup on the same cluster with gromacs 2016.5 compiled in the same way works well

md.2016.5.log:

```
GROMACS version:    2016.5
Precision:          single
Memory model:       64 bit
MPI library:        MPI
OpenMP support:     enabled (GMX_OPENMP_MAX_THREADS = 32)
GPU support:        disabled
SIMD instructions:  AVX_512_KNL
FFT library:        Intel MKL
RDTSCP usage:       enabled
TNG support:        enabled
Hwloc support:      hwloc-1.11.0
Tracing support:    disabled
Built on:           Thu May 17 12:17:27 CEST 2018
Built by:           ccamillo@r000u06l01 [CMAKE]
Build OS/arch:      Linux 3.10.0-327.36.3.el7.x86_64 x86_64
Build CPU vendor:   Intel
Build CPU brand:    Intel(R) Xeon(R) CPU E5-2697 v4 @ 2.30GHz
Build CPU family:   6   Model: 79   Stepping: 1
Build CPU features: aes apic avx avx2 clfsh cmov cx8 cx16 f16c fma hle htt lahf mmx msr nonstop_ts
c pcid pclmuldq pdcm pdpe1gb popcnt pse rdrnd rdtscp rtm sse2 sse3 sse4.1 sse4.2 ssse3 tdt x2apic
C compiler:         /cineca/prod/opt/compilers/intel/pe-xe-2017/binary/bin/icc Intel 17.0.4.201704
11
C compiler flags:    -xMIC-AVX512   -mkl=sequential  -std=gnu99  -O3 -DNDEBUG -ip -funroll-all-loo
ps -alias-const -ansi-alias
C++ compiler:       /cineca/prod/opt/compilers/intel/pe-xe-2017/binary/bin/icpc Intel 17.0.4.20170
411
C++ compiler flags:  -xMIC-AVX512   -mkl=sequential  -std=c++0x   -O3 -DNDEBUG -ip -funroll-all-lo
ops -alias-const -ansi-alias

Running on 1 node with total 68 cores, 272 logical cores
Hardware detected on host r065c06s01 (the node of MPI rank 0):
  CPU info:
    Vendor: Intel
    Brand:  Intel(R) Xeon Phi(TM) CPU 7250 @ 1.40GHz
    Family: 6   Model: 87   Stepping: 1
    Features: aes apic avx avx2 avx512f avx512pf avx512er avx512cd clfsh cmov cx8 cx16 f16c fma ht
t lahf mmx msr nonstop_tsc pclmuldq pdcm pdpe1gb popcnt pse rdrnd rdtscp sse2 sse3 sse4.1 sse4.2 s
sse3 tdt x2apic
    SIMD instructions most likely to fit this hardware: AVX_512_KNL
    SIMD instructions selected at GROMACS compile time: AVX_512_KNL

  Hardware topology: Basic
    Sockets, cores, and logical processors:
      Socket  0: [   0  68 136 204] [   1  69 137 205] [   2  70 138 206] [   3  71 139 207] [   4
  72 140 208] [   5  73 141 209] [   6  74 142 210] [   7  75 143 211] [   8  76 144 212] [   9  7
7 145 213] [  10  78 146 214] [  11  79 147 215] [  12  80 148 216] [  13  81 149 217] [  14  82 1
50 218] [  15  83 151 219] [  16  84 152 220] [  17  85 153 221] [  18  86 154 222] [  19  87 155
```

```
223] [  20  88 156 224] [  21  89 157 225] [  22  90 158 226] [  23  91 159 227] [  24  92 160 228
] [  25  93 161 229] [  26  94 162 230] [  27  95 163 231] [  28  96 164 232] [  29  97 165 233] [
  30  98 166 234] [  31  99 167 235] [  32 100 168 236] [  33 101 169 237] [  34 102 170 238] [  3
5 103 171 239] [  36 104 172 240] [  37 105 173 241] [  38 106 174 242] [  39 107 175 243] [  40 1
08 176 244] [  41 109 177 245] [  42 110 178 246] [  43 111 179 247] [  44 112 180 248] [  45 113
181 249] [  46 114 182 250] [  47 115 183 251] [  48 116 184 252] [  49 117 185 253] [  50 118 186
 254] [  51 119 187 255] [  52 120 188 256] [  53 121 189 257] [  54 122 190 258] [  55 123 191 25
9] [  56 124 192 260] [  57 125 193 261] [  58 126 194 262] [  59 127 195 263] [  60 128 196 264]
[  61 129 197 265] [  62 130 198 266] [  63 131 199 267] [  64 132 200 268] [  65 133 201 269] [
66 134 202 270] [  67 135 203 271]
```

Associated revisions

Revision f8b78130 - 05/21/2018 10:35 PM - Roland Schulz

Fix illegal instruction error on KNL

Fixes #2504

Change-Id: Ie2f55718f98d3dfbf3c312afa5141c77ead77a6d

History

#1 - 05/18/2018 12:48 AM - Mark Abraham

- Description updated

#2 - 05/18/2018 12:54 AM - Mark Abraham

Your .tpr works fine for me with intel 2018 compilers (+ mkl,openmpi,hwloc) on our dev-knl01 node.

The timing suggests that the module(?) for gromacs 2018.1 is failing to permit mdrun resolve hwloc library at run time. You might try ldd mdrun_knl to see if that is the case, or explictly loading an hwloc module (though I do not see such a module on marconi)

#3 - 05/18/2018 01:15 AM - Mark Abraham

However an otherwise identical build with icc 17.0.4.20170411 behaves very strangely, apparently starting 32 independent processes each of 272 threads. That's also consistent with Carlo's observations. I suggest Carlo use/request intel 2018 (I used 18.0.1.20171018)

#4 - 05/18/2018 02:35 AM - Roland Schulz

For me it's fine with 17.0.4. You might want to compile without HWLOC (-DGMX_HWLOC=no) to check whether that works. Also you should get an error before the MPI error. Please check whether there is any output before the MPI error and copy it here. BTW: You don't want to run with default number of OpenMP on KNL. You want to use "-ntomp 4".

#5 - 05/18/2018 03:07 PM - Carlo Camilloni

There is not much text before the mpi error

GROMACS: mdrun_knl, version 2018.1
Executable: /marconi/home/userexternal/ccamillo/opt/gromacs-2018.1/bin/mdrun_knl
Data prefix: /marconi/home/userexternal/ccamillo/opt/gromacs-2018.1
Working dir: /marconi_scratch/userexternal/ccamillo/test
Command line:
mdrun_knl -s topol0 -nb cpu -v -maxh 23.9 -nsteps -1

Back Off! I just backed up md.log to ./#md.log.9#

=== = BAD TERMINATION OF ONE OF YOUR APPLICATION PROCESSES = PID 25280 RUNNING AT r065c04s03 = EXIT CODE: 132 = CLEANING UP REMAINING PROCESSES = YOU CAN IGNORE THE BELOW CLEANUP MESSAGES
===

=== = BAD TERMINATION OF ONE OF YOUR APPLICATION PROCESSES = PID 25280 RUNNING AT r065c04s03 = EXIT CODE: 4 = CLEANING UP REMAINING PROCESSES = YOU CAN IGNORE THE BELOW CLEANUP MESSAGES
===
Intel(R) MPI Library troubleshooting guide:
https://software.intel.com/node/561764 ==

I will do some more test and report back

#6 - 05/18/2018 03:19 PM - Carlo Camilloni

Test 1:

(I have added hwloc in the module I am loading explicitly in the script and added explicitly -ntomp 2, but btw I am anyway setting the number of openmp thread in the slurm script, and indeed this is not an issue with GMX2016.5)

module load intel intelmpi mkl hwloc
mpiexec -np 32 mdrun_knl -s topol0 -nb cpu -v -maxh 23.9 -nsteps -1 -ntomp 2 >& log

Same error messages

Test 2:
Recompiled with HWLOC OFF

GROMACS version: 2018.1
Precision: single
Memory model: 64 bit
MPI library: MPI
OpenMP support: enabled (GMX_OPENMP_MAX_THREADS = 64)
GPU support: disabled
SIMD instructions: AVX_512_KNL
FFT library: Intel MKL
RDTSCP usage: enabled
TNG support: enabled
Hwloc support: disabled
Tracing support: disabled
Built on: 2018-05-17 11:33:17
Built by: ccamillo@r000u06l01 [CMAKE]
Build OS/arch: Linux 3.10.0-327.36.3.el7.x86_64 x86_64
Build CPU vendor: Intel
Build CPU brand: Intel(R) Xeon(R) CPU E5-2697 v4 @ 2.30GHz
Build CPU family: 6   Model: 79   Stepping: 1
Build CPU features: aes apic avx avx2 clfsh cmov cx8 cx16 f16c fma hle htt intel lahf mmx msr nonstop_tsc pcid pclmuldq pdcm pdpe1gb popcnt pse rdrnd rdtscp rtm sse2 sse3 sse4.1 sse4.2 ssse3 tdt x2apic
C compiler: /cineca/prod/opt/compilers/intel/pe-xe-2017/binary/bin/icc Intel 17.0.4.20170411
C compiler flags: -xMIC-AVX512   -mkl=sequential -std=gnu99 -O3 -DNDEBUG -ip -funroll-all-loops -alias-const -ansi-alias -no-prec-div -fimf-domain-exclusion=14 -qoverride-limits
C++ compiler: /cineca/prod/opt/compilers/intel/pe-xe-2017/binary/bin/icpc Intel 17.0.4.20170411
C++ compiler flags: -xMIC-AVX512   -mkl=sequential -std=c++11   -O3 -DNDEBUG -ip -funroll-all-loops -alias-const -ansi-alias -no-prec-div -fimf-domain-exclusion=14 -qoverride-limits

Same error message as before

I will test the Intel2018 available on Marconi

#7 - 05/18/2018 03:35 PM - Carlo Camilloni

Unfortunately it doesn't work also with the intel 2018 available

GROMACS version: 2018.1
Precision: single
Memory model: 64 bit
MPI library: MPI
OpenMP support: enabled (GMX_OPENMP_MAX_THREADS = 64)
GPU support: disabled
SIMD instructions: AVX_512_KNL
FFT library: Intel MKL
RDTSCP usage: enabled
TNG support: enabled
Hwloc support: disabled
Tracing support: disabled
Built on: 2018-05-18 13:21:43
Built by: ccamillo@r000u06l01 [CMAKE]
Build OS/arch: Linux 3.10.0-327.36.3.el7.x86_64 x86_64
Build CPU vendor: Intel
Build CPU brand: Intel(R) Xeon(R) CPU E5-2697 v4 @ 2.30GHz
Build CPU family: 6   Model: 79   Stepping: 1
Build CPU features: aes apic avx avx2 clfsh cmov cx8 cx16 f16c fma hle htt intel lahf mmx msr nonstop_tsc pcid pclmuldq pdcm pdpe1gb popcnt pse rdrnd rdtscp rtm sse2 sse3 sse4.1 sse4.2 ssse3 tdt x2apic
C compiler: /cineca/prod/opt/compilers/intel/pe-xe-2018/binary/bin/icc Intel 18.0.2.20180210
C compiler flags: -xMIC-AVX512   -mkl=sequential -std=gnu99 -O3 -DNDEBUG -ip -funroll-all-loops -alias-const -ansi-alias -no-prec-div -fimf-domain-exclusion=14 -qoverride-limits
C++ compiler: /cineca/prod/opt/compilers/intel/pe-xe-2018/binary/bin/icpc Intel 18.0.2.20180210
C++ compiler flags: -xMIC-AVX512   -mkl=sequential -std=c++11   -O3 -DNDEBUG -ip -funroll-all-loops -alias-const -ansi-alias -no-prec-div -fimf-domain-exclusion=14 -qoverride-limits

log:

Back Off! I just backed up md.log to ./#md.log.13#

=== = BAD TERMINATION OF ONE OF YOUR APPLICATION PROCESSES = PID 2883 RUNNING AT r065c01s03 = EXIT CODE: 132 = CLEANING UP REMAINING PROCESSES = YOU CAN IGNORE THE BELOW CLEANUP MESSAGES
===

=== = BAD TERMINATION OF ONE OF YOUR APPLICATION PROCESSES = PID 2883 RUNNING AT r065c01s03 = EXIT CODE: 4 = CLEANING UP REMAINING PROCESSES = YOU CAN IGNORE THE BELOW CLEANUP MESSAGES
===
Intel(R) MPI Library troubleshooting guide:
https://software.intel.com/node/561764 ==========

#8 - 05/19/2018 01:41 AM - Roland Schulz

It might be hard to help without access to the cluster given that we can't reproduce it on other machines.
Ideas:

- Try to run mdrun from an interactive slurm shell rather than starting it through a script. My hope is it shows more info in that case.
- Build with "-g" (CXXFLAGS and CFLAGS) and run inside debugger.

Have you asked the Marconi admins?

#9 - 05/21/2018 05:07 PM - Carlo Camilloni

I know, anyway for completeness

Compiled in debug mode it works

Compiled in RelWithDebInfo it doesn't and unfortunately the interactive mode doesn't give any additional information about the error

It is likely a compiler issue but a unfortunate one..

#10 - 05/21/2018 07:23 PM - Mark Abraham

Or trying using an Intel compiler with another MPI library.

#11 - 05/21/2018 09:52 PM - Carlo Camilloni

I don't know if it is of any help, but recompiling with openmpi and RelWithDebInfo and running it inside GDB I get

Tracepoint 1 at 0x63fc7b: file /marconi/home/userexternal/ccamillo/Codes/gromacs-2018.1/src/gromacs/hardware/identifyavx512fmaunits.cpp, line 244.

#12 - 05/21/2018 10:10 PM - Roland Schulz

Could you try the Release build and change the SIMD_AVX_512_CXX_SUPPORTED in the CMakeCache.txt file to 0?

#13 - 05/21/2018 10:13 PM - Roland Schulz

Could you also paste all (/much more) of your output from GDB? Did you set a tracepoint? In what context did you get that message?

#14 - 05/21/2018 10:19 PM - Carlo Camilloni

Changing SIMD_AVX_512_CXX_SUPPORTED in the CMakeCache.txt file to 0 makes it work!

About the gdb output it was not much more than what I reported:

Reading symbols from /marconi/home/userexternal/ccamillo/opt/gromacs-2018.1/bin/mdrun_knl...done.
(gdb) run
Starting program: /marconi/home/userexternal/ccamillo/opt/gromacs-2018.1/bin/mdrun_knl
[Thread debugging using libthread_db enabled]
Using host libthread_db library "/lib64/libthread_db.so.1".
Missing separate debuginfo for /cineca/prod/opt/compilers/intel/pe-xe-2017/binary/inspector/lib64/libstdc++.so.6
Missing separate debuginfo for /cineca/prod/opt/compilers/intel/pe-xe-2017/binary/inspector/lib64/libgcc_s.so.1
[New Thread 0x2aaab4359700 (LWP 38979)]
[New Thread 0x2aaab497d700 (LWP 39012)]
:-) GROMACS - mdrun_knl, 2018.1 (-:

```
GROMACS is written by:
[...]
Back Off! I just backed up md.log to ./#md.log.9#
```

Program received signal SIGILL, Illegal instruction.

checkDualAvx512FmaUnits () at /marconi/home/userexternal/ccamillo/Codes/gromacs-2018.1/src/gromacs/hardware/identifyavx512fmaunits.cpp:244
244   return (timeFmaAndShuf > 1.5 * timeFmaOnly);
Missing separate debuginfos, use: debuginfo-install glibc-2.17-157.el7_3.5.x86_64 infinipath-psm-3.3-0.g6f42cdb1bb8.2.el7.x86_64 libhfi1-0.5-27.el7.x86_64 libibumad-1.3.10.2-1.el7.x86_64 libibverbs-1.1.8-8.el7.x86_64 libipathverbs-1.3-2.el7.x86_64 libnl3-3.2.21-10.el7.x86_64 libpciaccess-0.13.4-2.el7.x86_64 libpsm2-10.2.235-1.x86_64 librdmacm-1.0.21-1.el7.x86_64 numactl-libs-2.0.9-6.el7_2.x86_64 opensm-libs-3.3.19-1.el7.x86_64 sssd-client-1.13.0-40.el7_2.12.x86_64

(gdb) trace
Tracepoint 1 at 0x63fc7b: file /marconi/home/userexternal/ccamillo/Codes/gromacs-2018.1/src/gromacs/hardware/identifyavx512fmaunits.cpp, line 244.

#15 - 05/21/2018 10:30 PM - Gerrit Code Review Bot

Gerrit received a related patchset '1' for Issue #2504.
Uploader: Roland Schulz (roland.schulz@intel.com)
Change-Id: gromacs~release-2018~Ie2f55718f98d3dfbf3c312afa5141c77ead77a6d
Gerrit URL: https://gerrit.gromacs.org/7926

#16 - 05/21/2018 10:31 PM - Roland Schulz

- Status changed from New to Fix uploaded

Could you please try the fix I uploaded to Gerrit?

#17 - 05/21/2018 10:33 PM - Roland Schulz

PS: Please undo the change to the CMakeCache.txt for the test.

#18 - 05/21/2018 10:50 PM - Carlo Camilloni

yes your fix (patch set 2) works

#19 - 05/21/2018 11:31 PM - Roland Schulz

Thanks. Until we release 2018.2 you can use the CMakeCache work-around. That shouldn't have any side-effects.

#20 - 05/21/2018 11:38 PM - Carlo Camilloni

Great, thanks!

#21 - 05/22/2018 10:45 AM - Roland Schulz

- Status changed from Fix uploaded to Resolved

Applied in changeset f8b78130e021a0fef1aaa2b1e39c73dbec48fcbb.

#22 - 05/22/2018 10:23 PM - Mark Abraham

- Category set to analysis tools

- Status changed from Resolved to Closed

- Assignee set to Roland Schulz

- Target version set to 2018.2

Files

topol0.tpr	1.31 MB	05/17/2018	Carlo Camilloni