Gromacs - Bug #1572
Incorrect PME energies and forces with high numbers of OpenMP threads
08/02/2014 10:46 AM - Berk Hess

<table>
<thead>
<tr>
<th>Status:</th>
<th>Closed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priority:</td>
<td>High</td>
</tr>
<tr>
<td>Assignee:</td>
<td>Berk Hess</td>
</tr>
<tr>
<td>Category:</td>
<td>mdrun</td>
</tr>
<tr>
<td>Target version:</td>
<td>5.0.5</td>
</tr>
<tr>
<td>Affected version - extra info:</td>
<td>4.6</td>
</tr>
</tbody>
</table>

**Description**

With combined MPI and very high OpenMP parallelization the PME energy and forces can be incorrect and there are data races between threads. This only happens with PME grid spreading subdivision along y over 3 or more threads with the y thread-local grid size smaller than pme_order-1. For production runs this situation will likely never occur. The only case where this happens is a relatively small PME grid with a very high OpenMP thread count containing relatively large prime factors, e.g. 20^3 grid with 3x2x1 domain decomposition and 25 OpenMP threads (which is an order of magnitude beyond the scaling limit).

**Related issues:**

- Related to Gromacs - Bug #1578: PME incorrect with MPI+OpenMP and multiple MP... Closed 08/15/2014
- Related to Gromacs - Bug #1388: Data race in PME with large prime number of t... Closed 11/25/2013

**Associated revisions**

**Revision 27189bba - 08/02/2014 01:55 PM - Berk Hess**

Fixed PME bug with high OpenMP thread count

PME energies and forces could be incorrect with combined MPI+OpenMP parallelization. This would, only, happen when pmegrids->nthread_comm[YY] >> 2, which can only occur with high OpenMP thread count with multiple large prime factors.

It's unlikely that this issue affected production runs.

Fixes #1572.

Change-id: l03b38c279c8f8ab2e111dad0976edad88b3ea93b

**Revision 6ba80a26 - 08/28/2014 04:08 PM - Berk Hess**

Fixed two PME issues with MPI+OpenMP

Change 272736bc partially fixed #1388, but broke the more general case of multiple MPI communication pulses in PME. Change 272736bc incorrectly changed tx1 and ty1. This change has been reverted.

Change 27189bba fixed the incorrect PME grid reduction with multiple thread grid overlap in y. But it broke the, much more common, case where the y-size of the PME grid is not divisible by the domains in y. This change, incorrectly, changed buf_my.

Now buf_my is set to the correct value, which solves both issues.

Fixes #1578.

Refs #1388 and #1572.
Re-fixed PME bug with high OpenMP thread count

PME energies and forces could be incorrect with combined MPI+OpenMP parallelization. This would, only, happen when
pmegrids->nthread_comm[YY] >= 2, which can only occur with high OpenMP thread count with multiple prime factors that are large wrt the grid.
It's unlikely that this issue affected production runs.
This bug was fixed in 27189bba, but 6ba80a26 broke it again.

Fixes #1572.

Change-Id: lc01bed4193062f8ca85fbcb6bf347f2ef0de909f

History

#1 - 08/02/2014 10:51 AM - Gerrit Code Review Bot
Gerrit received a related patchset '1' for Issue #1572.
Uploader: Berk Hess (hess@kth.se)
Change-Id: f03b38c279e889b6e1111d0976edad88b3ea93b
Gerrit URL: https://gerrit.gromacs.org/3847

#2 - 08/04/2014 01:24 PM - Berk Hess
- Status changed from New to Fix uploaded

#3 - 08/10/2014 08:20 PM - Roland Schulz
- Status changed from Fix uploaded to Closed

#4 - 08/15/2014 11:19 AM - Gerrit Code Review Bot
Gerrit received a related patchset '1' for Issue #1572.
Uploader: Berk Hess (hess@kth.se)
Change-Id: ld2d7d013a3b8cdc04eda1fb026567088a38ec81f
Gerrit URL: https://gerrit.gromacs.org/3896

#5 - 08/18/2014 03:05 AM - Gerrit Code Review Bot
Gerrit received a related DRAFT patchset '1' for Issue #1572.
Uploader: Szilárd Páll (pall.szilard@gmail.com)
Change-Id: i3649294a143bb744a2e26fd1d9dbf87dea421ca
Gerrit URL: https://gerrit.gromacs.org/3905
This wasn't fully solved. The rnbxn-free-energy test run with "-nt 8 -ntomp 16" and commit 50e7c6408e45f (merge of master and release-5-0) gives:

```
WARNING: ThreadSanitizer: data race (pid=12745)
Write of size 4 at 0x7d8800067e7c by thread T26:
  #0 reduce_threadgrid_overlap ../src/gromacs/ewald/pme.c:4071 (libgromacs.so.1+0x000000270bf4)
  #1 gomp_thread_start ../../../libgomp/team.c:117 (libgomp.so.1+0x00000000ccc9)
Previous write of size 4 at 0x7d8800067e7c by thread T83:
  #0 reduce_threadgrid_overlap ../src/gromacs/ewald/pme.c:4071 (libgromacs.so.1+0x000000270bf4)
  #1 gomp_thread_start ../../../libgomp/team.c:117 (libgomp.so.1+0x00000000ccc9)
Location is heap block of size 5152 at 0x7d8800067800 allocated by thread T3:
  #0 calloc <null>:0 (libtsan.so.0+0x00000004c571)
  #1 save_calloc ../src/gromacs/utility/asmalloc.c:179 (libgromacs.so.1+0x0000002e8b6c)
  #2 init_overlap_comm ../src/gromacs/ewald/pme.c:3234 (libgromacs.so.1+0x000000225dfb2)
  #3 gmx_pme_init ../src/gromacs/ewald/pme.c:3574 (libgromacs.so.1+0x00000027ba12)
  #4 mdrunner ../src/programs/mdrun/runner.cpp:1644 (gmx+0x000000024a27)
  #5 mdrunner_start_fn ../src/programs/mdrun/runner.cpp:187 (gmx+0x000000025e67)
  #6 tMPI_Thread Starter ../src/external/thread_mpi/src/tmpi_init.c:397 (libgromacs.so.1+0x000000f89409)
  #7 tMPI_Thread Starter ../src/external/thread_mpi/src/tfthreads.c:234 (libgromacs.so.1+0x000000f7d28)
  #8 <null> <null>:0 (libtsan.so.0+0x000000035609)
```

This specific version isn't necessary. A few other master versions give the same error. If I revert commit 6ba80a267eaf2 (the 2nd patch) then it is OK.

---

#7 - 12/06/2014 01:21 AM - Roland Schulz
- Related to Bug #1578: PME incorrect with MPI+OpenMP and multiple MPI communication pulses added

#8 - 12/06/2014 01:21 AM - Roland Schulz
- Related to Bug #1388: Data race in PME with large prime number of threads added

#9 - 12/06/2014 01:28 AM - Roland Schulz
PS: If I also revert 27189bba (1st patch) then it is still OK. If I revert also 272736bc then it isn't OK. So the 6ba80a267eaf2 breaks the fix in 272736bc.

#10 - 12/09/2014 04:16 PM - Mark Abraham
- Target version changed from 5.0.3 to 5.0.4
#11 - 12/15/2014 09:45 PM - Mark Abraham
- Target version changed from 5.0.4 to 5.0.5

Sorry haven't had time to look into this, and won't have time until January.

#12 - 02/05/2015 04:41 AM - Roland Schulz
- Priority changed from Normal to High

Increasing priority: It produces incorrect results. It doesn't cause a crash. It affects the regression tests.

#13 - 02/05/2015 09:25 PM - Berk Hess
- File pme.c added

I think both min operations in the final change should be combined for the communication buffer. The attached pme.c fixes this issue. But we should check thoroughly that it doesn't brake anything else. I'm currently too tired to think this through carefully.

#14 - 02/05/2015 09:26 PM - Berk Hess
- Status changed from Accepted to In Progress

#15 - 02/06/2015 02:07 PM - Gerrit Code Review Bot

Gerrit received a related patchset '1' for Issue #1572.
Uploader: Berk Hess (hess@kth.se)
Change-Id: Ic01bed4193062f8ca885fcb6bf347f2ef0de90f
Gerrit URL: https://gerrit.gromacs.org/4440

#16 - 04/09/2015 10:47 AM - Mark Abraham
- Status changed from In Progress to Resolved

I think this is fixed and merged through to all branches now?

#17 - 05/06/2015 04:50 PM - Mark Abraham
- Status changed from Resolved to Closed

No known work to do, closing

Files

<table>
<thead>
<tr>
<th>File</th>
<th>Size</th>
<th>Date</th>
<th>User</th>
</tr>
</thead>
<tbody>
<tr>
<td>pme.c</td>
<td>170 KB</td>
<td>02/05/2015</td>
<td>Berk Hess</td>
</tr>
</tbody>
</table>