GROMACS - Bug #2172

EM run does give valid reason for not printing performance report

05/08/2017 09:58 PM - Mark Abraham

Status: Closed
Priority: Normal
Assignee: Mark Abraham
Category: mdrun
Target version: 2016.5
Affected version - extra info:

Affected version: 2016.3

Description

Log file snippet

<table>
<thead>
<tr>
<th>Step</th>
<th>Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>1025</td>
<td>1025.00000</td>
</tr>
</tbody>
</table>

Energies (kJ/mol)

<table>
<thead>
<tr>
<th>Bond</th>
<th>U-B</th>
<th>Proper Dih.</th>
<th>Improper Dih.</th>
<th>LJ-14</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.33152e+02</td>
<td>1.36263e+04</td>
<td>2.69669e+04</td>
<td>1.29206e+02</td>
<td>3.36597e+03</td>
</tr>
<tr>
<td>Coulomb-14</td>
<td>1.34164e+04</td>
<td>-2.90203e+05</td>
<td>3.82327e+03</td>
<td>2.48443e+02</td>
</tr>
<tr>
<td>Dih. Rest.</td>
<td>-2.82452e+05</td>
<td>-8.07096e+02</td>
<td>1.49630e-10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-5.47705e+04</td>
<td>1.34164e+04</td>
<td>-2.90203e+05</td>
<td>3.82327e+03</td>
</tr>
<tr>
<td></td>
<td>1.19652e+01</td>
<td>-2.82452e+05</td>
<td>-8.07096e+02</td>
<td>1.49630e-10</td>
</tr>
</tbody>
</table>

Steepest Descents converged to Fmax < 1000 in 1028 steps
Potential Energy = -2.82538847757876e+05
Maximum force = 8.87247960626888e+02 on atom 14990
Norm of force = 2.38207577988486e+01

Simulation ended prematurely, no performance report will be written.

This is a valid end of the simulation, but I forget whether the intent is to make a performance report.

Associated revisions

Revision 8f179303 - 12/14/2017 04:25 PM - Mark Abraham

Avoid confusing message at end of non-dynamical runs

EM, TPI, NM, etc. are not targets for performance optimization so we will not write performance reports. This commit fixes and oversight whereby we would warn a user when the lack of performance report is normal and expected.

Fixes #2172

Change-Id: i1097304d79701be748612510572382729f7126be
History
#1 - 05/31/2017 08:26 PM - Szilárd Páll

Mark Abraham wrote:

This is a valid end of the simulation, but I forget whether the intent is to make a performance report.

My brief tests show that even a vanilla EM run (no counter resetting) will claim "premature" exit and won't report. I can't think of a reason why we'd not want to report performance of an EM run.

#2 - 05/31/2017 10:50 PM - Mark Abraham

Szilárd Páll wrote:

Mark Abraham wrote:

This is a valid end of the simulation, but I forget whether the intent is to make a performance report.

My brief tests show that even a vanilla EM run (no counter resetting) will claim "premature" exit and won't report.

Counter reset is only implemented for do_md, and each EM has a separate "integrator" implementation, so that aspect doesn't matter. We just didn't think about EM when we did some refactoring some time.

I can't think of a reason why we'd not want to report performance of an EM run.

I don't think there is a good reason to report on it, because there's very little that a user would want to change (turn off PME, reduce cutoffs, and choose a different EM algorithm are all I can think of), and it doesn't take enough time for us to worry about it as a performance optimization target. But making behaviour consistent across "integrators" would also be a reasonable choice.

I think we should just fix the logic so that the absence of a performance report is normal and expected for many of the "integrators." Or at least have the error message not refer to "simulation," which is not really applicable to a non-dynamical calculation.

#3 - 06/01/2017 01:16 AM - Szilárd Páll

I tend to think users will be interested in the performance of em runs, not because settings can be tweaked but because the decision whether to use their local workstation, a faster lab box or a remote cluster is relevant. Hence, reporting some kind of metric of how quick is a minimization should be useful especially as a simulation system will often be set up multiple times. Knowing that minimization takes longer than a lunch break is useful, but it's not a great metric. I do realize "ns/day" is not the ideal metric either, so perhaps time/step (or steps/unit of time) is better in this case -- although to avoid confusion, using a familiar (though incorrect metric) may be better.

I also realize that I'm not the best person to judge what matters to users, so it could be better to ask around.

#4 - 06/01/2017 03:53 PM - Mark Abraham

Szilárd Páll wrote:

I tend to think users will be interested in the performance of em runs, not because settings can be tweaked but because the decision whether to use their local workstation, a faster lab box or a remote cluster is relevant. Hence, reporting some kind of metric of how quick is a minimization should be useful especially as a simulation system will often be set up multiple times. Knowing that minimization takes longer than a lunch break is useful, but it's not a great metric. I do realize "ns/day" is not the ideal metric either, so perhaps time/step (or steps/unit of time) is better in this case -- although to avoid confusion, using a familiar (though incorrect metric) may be better.

I've never seen an EM that takes more than a few minutes, so can't imagine someone planning a simulation to run for hours/days/months would care. We can ask at group meeting whether anybody has ever run an EM that took an amount of time that they cared about, and whether they considered doing anything about it.

#5 - 09/12/2017 11:48 AM - Mark Abraham

- Target version changed from 2016.4 to 2016.5

There has been some discussion on gmx-users in Sep 2017 that a user had a workflow where they noticed that GPU acceleration of EM was not useful, but there wasn't anything to suggest we should do a significant amount of work for EM performance.

#6 - 12/12/2017 01:30 AM - Erik Lindahl

- Status changed from New to Rejected

The EM code has completely different bottlenecks than our usual runs, and much of our normal performance analysis is not useful (such as performance in ns/day, etc.).
Most EM runs tend to stop because things have converged to machine precision rather long before the target number of steps is reached, and with -v output we anyway write a ton of other stuff which means the performance can't even be visible.

This is not a bug in the sense that we don't intend to write performance reports for EM runs. If somebody feels they both have the time to do it and want to prioritize it they can go right ahead and update all the performance code, but it's not a bug the entire team should invest efforts in fixing, IMHO.

#7 - 12/12/2017 03:52 AM - Mark Abraham
- Status changed from Rejected to Accepted

Bug is present, inasmuch as EM reports

"Simulation ended prematurely, no performance report will be written."

as a result of refactoring end-of-run cleanup also for simulations where that message is appropriate.

EM is not a simulation, EM generally didn't end prematurely when this message is written, and we don't intend to prioritise reporting on or optimizing its performance, so we should just write nothing rather than worry users that their EM is somehow invalid.

#8 - 12/12/2017 11:47 AM - Gerrit Code Review Bot
Gerrit received a related patchset '1' for Issue #2172.
Uploader: Mark Abraham (mark.j.abraham@gmail.com)
Change-Id: gromacs~release-2016~I1097304d79701be7486125105723827297f726be
Gerrit URL: https://gerrit.gromacs.org/7333

#9 - 12/12/2017 11:48 AM - Mark Abraham
- Status changed from Accepted to Fix uploaded

#10 - 12/14/2017 01:48 PM - Gerrit Code Review Bot
Gerrit received a related patchset '1' for Issue #2172.
Uploader: Mark Abraham (mark.j.abraham@gmail.com)
Change-Id: gromacs~release-2018~I1097304d79701be7486125105723827297f726be
Gerrit URL: https://gerrit.gromacs.org/7350

#11 - 12/15/2017 07:33 AM - Mark Abraham
- Status changed from Fix uploaded to Resolved

Applied in changeset 8f179303044a5ae3ccd9d042e98906c1c22bb741.

#12 - 12/17/2017 07:12 AM - Erik Lindahl
- Status changed from Resolved to Closed