Multidir simulations can stop at different times when using mdrun -maxh

I am running mdrun -multi dir-{1..4} (together with awh-share-multisim, but I think it's irrelevant). I'm running on a cluster, submitting many jobs with -maxh 0.25 -cpi. At some point I get (see attached slurm-last-ok.out)

```
Step 57822720: Run time exceeded 0.247 hours, will terminate the run
Step 57822760: Run time exceeded 0.247 hours, will terminate the run
Step 57822760: Run time exceeded 0.247 hours, will terminate the run
Step 57822760: Run time exceeded 0.247 hours, will terminate the run
```

At the next restart I get a fatal error:

```
simulation part is not equal for all subsystems
subsystem 0: 52
subsystem 1: 51
subsystem 2: 51
subsystem 3: 51
```

For more information and tips for troubleshooting, please check the GROMACS website at http://www.gromacs.org/Documentation/Errors

see slurm-first-error.out. What is the intended behavior is here? I think it should either be ok that the different simulations start at different steps or they should never terminate at different steps.

Related issues:
Related to GROMACS - Bug #2447: "gmx mdrun -maxh" not working with AWH Closed
Related to GROMACS - Bug #3448: GMX 2020.1 - Multidir simulations can stop at... New

Associated revisions
Revision 71fd5b5b - 03/13/2018 02:21 PM - Berk Hess
Fix AWH ensemble checkpointing

AWH multi-runs with shared bias could terminate at different steps leading to checkpointing from which continuation was impossible.
Fixes #2440.

Change-Id: I728c6aea9d030fcafad04efc6082309dabe3408bd

Revision b1d72b28 - 03/14/2018 02:09 PM - Berk Hess
Fix mdrun signalling with AWH bias sharing

The recent commit 717fd5b8, which should have fixed issue #2440, disabled checkpointing and termination by disabled signalling between AWH runs that share a bias.

Fixes #2447
Refs #2440

Change-Id: lc80ea07cfa1bde1b31fa9d660f00492b794b000f

Revision e4dea12 - 03/14/2018 06:33 PM - Berk Hess
Improve inter-simulation signaling

Use a single boolean for controlling inter-simulation signaling. This would have prevented issue #2447.
Introduce an inter-simulating signaling frequency. This both simplifies the code and significantly reduces communication overhead for many simulations and/or many ranks per simulation.
Correct the termination step print with replica exchange.

Refs #2440 and #2447

Change-Id: l8ed50c5b11564d9a967bcac15cfc2350e5e0caea

Revision 0fa62d3e - 02/21/2020 02:00 PM - Berk Hess
Add fatal error when multisim runs sharing state have different init_step

Refs #2440
Fixes #3990

Change-Id: l052cd53dc9517a9df5363e52720ee5a80ea65c0

Revision ade8e131 - 02/25/2020 06:43 PM - Magnus Lundborg
Fix out of sync checkpoint files in simulations sharing state

When multidir simulations share the state, the checkpoint files of the different simulations should all be from the same step. To ensure this, MPI barriers have been added before renaming the checkpoint files from their temporary to their final names. So now the contents can never be out of sync. In the worst, and rather unlikely, case that something going wrong during renaming, some checkpoint files could have temporary and some final names.

Refs #2440.

Change-Id: l88088abb7726a36dbf9a9db2fa2eb4a46c3bf2cd7

History

#1 - 03/12/2018 02:27 PM - Mark Abraham
Possibly same issue as #2233

#2 - 03/12/2018 04:45 PM - Gerrit Code Review Bot
Gerrit received a related patchset '1' for Issue #2440.
Uploader: Berk Hess (hess@kth.se)
Change-Id: gromacs~release-2018~I728c6aea9d030fcafad04efc6082309dabe3408bd
Gerrit URL: https://gerrit.gromacs.org/7664

#3 - 03/12/2018 04:46 PM - Berk Hess
- Status changed from New to Fix uploaded
- Assignee set to Berk Hess
- Target version set to 2018.1
#4 - 03/13/2018 02:30 PM - Berk Hess
- Status changed from Fix uploaded to Resolved

Applied in changeset 717d65b8c0b47ab8bf70cfbd189e52a452db543.

#5 - 03/13/2018 10:59 PM - Viveca Lindahl
- Related to Bug #2447: "gmx mdrun -maxh" not working with AWH added

#6 - 03/14/2018 09:38 AM - Gerrit Code Review Bot
Gerrit received a related patchset '1' for Issue #2440.
Uploader: Berk Hess (hess@kth.se)
Change-id: gromacs~release-2018~fc80ea07cfa1bde1b31fa9d660f00492b794b000f
Gerrit URL: https://gerrit.gromacs.org/7675

#7 - 03/14/2018 12:05 PM - Gerrit Code Review Bot
Gerrit received a related patchset '1' for Issue #2440.
Uploader: Berk Hess (hess@kth.se)
Change-id: gromacs~release-2018~18ed50c5b11564d9a967bcacb15c4fcd3be05caea
Gerrit URL: https://gerrit.gromacs.org/7679

#8 - 03/17/2018 12:16 AM - Mark Abraham
- Status changed from Resolved to Closed

#9 - 02/15/2020 04:56 PM - Magnus Lundborg
- Status changed from Closed to Accepted
- Target version changed from 2018.1 to 2019.6
- Affected version changed from 2018.1 to 2019.5

This still happens in 2019.5, so I am opening this again. I am not sure if it is related to -maxh or just when jobs get terminated, e.g., by a slurm timeout. I have only suffered from it when running awh, but since that is the only way I am using -multidir simulations I cannot say if it is related to awh or not.

#10 - 02/20/2020 08:04 AM - Magnus Lundborg
- Target version changed from 2019.6 to 2020.2

#11 - 02/20/2020 08:23 AM - Magnus Lundborg
- Target version changed from 2020.2 to 2020.1

#12 - 02/20/2020 10:13 PM - Daniel Kozuch
Magnus Lundborg wrote:

This still happens in 2019.5, so I am opening this again. I am not sure if it is related to -maxh or just when jobs get terminated, e.g., by a slurm timeout. I have only suffered from it when running awh, but since that is the only way I am using -multidir simulations I cannot say if it is related to awh or not.

Apologies if this is not the right place to post this - but I can confirm that the issue appears for 2019.5 during normal replica exchange without -maxh (slurm terminated)

#13 - 02/21/2020 09:15 AM - Mark Abraham
Fundamentally there's no promise of coherency between MPI operations and filesystem state, which is particularly problematic when multiple output files are involved. The best one can do is try to read the output file to observe its state is consistent, and communicate about that. But that's still not a promise that the file was flushed to physical disk

#14 - 02/27/2020 11:09 AM - Paul Bauer
- Status changed from Accepted to Resolved

the main issue here should be fixed with the related patches, but it again points to it that more work is needed here

#15 - 02/28/2020 09:13 AM - Paul Bauer
- Status changed from Resolved to Closed

04/04/2020
- Related to Bug #3448: GMX 2020.1 - Multidir simulations can stop at different times when killed by job manager added

Files

<table>
<thead>
<tr>
<th>File Name</th>
<th>Size</th>
<th>Date</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>slurm-first-error.out</td>
<td>3.35 KB</td>
<td>03/10/2018</td>
<td>Viveca Lindahl</td>
</tr>
<tr>
<td>slurm-last-ok.out</td>
<td>19.9 KB</td>
<td>03/10/2018</td>
<td>Viveca Lindahl</td>
</tr>
<tr>
<td>md-0.log</td>
<td>495 KB</td>
<td>03/10/2018</td>
<td>Viveca Lindahl</td>
</tr>
</tbody>
</table>