

GROMACS - Bug #3159

Task # 3370 (New): Further improvements to GPU Buffer Ops and Comms

eliminate regression due to moving `gmx_pme_send_coordinates()`

10/17/2019 02:57 PM - Szilárd Páll

Status: Closed	
Priority: High	
Assignee: Alan Gray	
Category: mdrun	
Target version: 2020.1	
Affected version - extra info:	Difficulty: uncategorized
Affected version: 2020-beta3	

Description

`gmx_pme_send_coordinates()` is moved past the x H2D which will lead to regressions on the default code-path.

This needs to be made conditional and done only

i) eliminate this in favor of direct CPU->GPU copy.

ii) if there is proof that there is benefit from doing a H2D followed by direct GPU->GPU is beneficial (which may be the case if we have PCIe between CPU<->GPU but NVLink between GPU<->GPU) -- side-note: that's why perhaps we should use `nccl`.

Associated revisions

Revision 5b594f3b - 10/19/2019 12:52 AM - Alan Gray

GPU Receive for PME/PP GPU Force Communications

This change extends the PME/PP GPU force communication functionality to allow the force buffer to be received direct to GPU memory on the PP task.

Implements part of #2817

Refs #3158 #3159

Change-Id: I5b1cff1846c7c3bd966b6bf9c0af72769600ef18

Revision c5595a8e - 10/21/2019 11:32 AM - Alan Gray

GPU Coordinate PME/PP Communications

Extends `PmePpCommGpu` class to provide PP-side support for coordinate transfers from either GPU or CPU to PME task, and adds new `PmeCoordinateReceiverGpu` class to receive coordinate data directly to the GPU on the PME task.

Implements part of #2817

Refs TODOs #3157 #3158 #3159

Change-Id: Iefa2bdfd9813282ad8b07feeb7691f16880e61a2

Revision 1953fd66 - 12/18/2019 08:29 PM - Alan Gray

Fix conditional on assertion for combining GPU Update and GPU coordinate send

Fixes a problem introduced in 9a3c2ce312c1e04634d636655388a8aaad53c8d1 which unintentionally blocks any use of GPU update with PME-PP Communication because `!stepWork.doNeighborSearch` is missing from the condition on the assertion.

Related to #3159 and #3160 which, when implemented, will remove the restriction on coordinate send originating from the GPU (on non search steps).

Change-Id: Ie9b1102386b2c1bc927c09e4d8ead2bd2321320d

Revision ddbaf0b9 - 03/03/2020 11:40 AM - Alan Gray

Allow PME coordinate send before H2D coordinate transfer

The introduction of the GPU PME-PP communication functionality had the side effect of delaying the PME coordinate send until after the H2D coordinate transfer, even on the default code path. This patch allows the PME transfer to occur in its original location when the send is not originating from GPU memory. This is a lightweight solution, without any new functionality, suitable for the release branch. (There will be a more comprehensive change in the master branch which also extends the GPU PME-PP communication functionality.)

Implements #3159

Change-Id: Ic30c154e04bb4c2846bbad3de603f879a71b9133

History

#1 - 11/01/2019 11:26 AM - Alan Gray

- Status changed from New to In Progress

- Target version changed from 2020 to 2020-beta3

Direct CPU->GPU copy already works in master branch by setting sendCoordinatesFromGpu=true.

With NVLINK on DGX, it is definitely better to use existing mechanism otherwise there are two H2D PCI transfers (to PP and PME) on the same bus. For non-NVLINK, yes it should be better to go direct CPU->GPU. We can choose path based on whether peer access is enabled (and possibly later extend to instead query architecture using NVML which would give more detail). Working on this at the moment.

#2 - 12/02/2019 10:02 AM - Paul Bauer

what is the status of this?

#3 - 12/02/2019 11:20 AM - Alan Gray

- Target version changed from 2020-beta3 to 2021-infrastructure-stable

Bumping to 2021

#4 - 12/02/2019 01:54 PM - Szilárd Páll

Alan Gray wrote:

Bumping to 2021

This is a regression in 2020 default code-path as a side-effect of new code, so IMO it needs to be fixed in 2020.

#5 - 12/02/2019 01:59 PM - Alan Gray

- Target version changed from 2021-infrastructure-stable to 2020-rc1

changing target to 2020-RC1

#6 - 12/17/2019 01:59 PM - Szilárd Páll

What is the status of this? I still think this is high prio and we should not leave this unresolved in the release.

#7 - 12/17/2019 02:11 PM - Alan Gray

What is the status of this? I still think this is high prio and we should not leave this unresolved in the release.

It is addressed in

<https://gerrit.gromacs.org/c/gromacs/+/14238>

I moved it to master after a request from Paul, but can move it back to release-2020 if Paul has no objections?

It currently still has an issue working correctly when update is on GPU. Under master, I therefore put it on hold until the release branch is merged. But if it goes back to the release branch it is easier to work on it and fix the problem.

#8 - 12/18/2019 10:36 AM - Alan Gray

- Status changed from In Progress to Fix uploaded

#9 - 12/18/2019 08:11 PM - Szilárd Páll

- *Tracker changed from Task to Bug*
- *Affected version set to 2020-beta3*

This is a bug in beta3so I think the 2020-rc1 target is valid -- we agreed to not have the new experimental features introduce regressions, and this is clearly a case of regression.

I agree that the proposed solution goes beyond just fixing the bug.

However, I am not sure there is a net positive value in avoiding the slight risk of destabilizing the experimental features (i.e. PP-PME direct GPU comm) especially when the upside seems significant to the rest of the code: i) fixing the regression ii) simpler and more maintainable code on the long run.

#10 - 12/20/2019 08:28 AM - Paul Bauer

- *Target version changed from 2020-rc1 to 2020*

@Szilard and @Alan if this is a regression it needs fixing!

#11 - 12/20/2019 01:05 PM - Alan Gray

It is fixed in <https://gerrit.gromacs.org/c/gromacs/+14238> which I believe is now working correctly.

#12 - 12/27/2019 04:15 PM - Paul Bauer

- *Target version changed from 2020 to 2020.1*

Alan Gray wrote:

It is fixed in <https://gerrit.gromacs.org/c/gromacs/+14238> which I believe is now working correctly.

Yes, but this included feature work that should not have been targeted at 2020.
I retargeted this at 2020.1 for just fixing the regression!

#13 - 01/06/2020 03:15 PM - Alan Gray

Yes, but this included feature work that should not have been targeted at 2020. I retargeted this at 2020.1 for just fixing the regression!

I've uploaded a new lightweight fix, with no new functionality, at <https://gerrit.gromacs.org/c/gromacs/+15200>

#14 - 02/07/2020 01:03 PM - Alan Gray

- *Status changed from Fix uploaded to Closed*

Moved to umbrella task <https://redmine.gromacs.org/issues/3370>

#15 - 02/14/2020 11:31 AM - Alan Gray

- *Status changed from Closed to Fix uploaded*
- *Parent task changed from #2891 to #3370*

Re-opening and moving to subtask of [#3370](https://redmine.gromacs.org/issues/3370), so we don't lose the discussion.

#16 - 02/27/2020 11:06 AM - Paul Bauer

is still an issue or can we close the bug?

#17 - 02/27/2020 12:08 PM - Alan Gray

is still an issue or can we close the bug?

Requires <https://gerrit.gromacs.org/c/gromacs/+15200> to be merged.

#18 - 03/03/2020 11:51 AM - Paul Bauer

- Status changed from Fix uploaded to Resolved

#19 - 03/03/2020 11:51 AM - Paul Bauer

- Status changed from Resolved to Closed