enable GPU peer to peer access
09/12/2019 04:24 PM - Sziliárd Páll

Status: Closed
Priority: Normal
Assignee: mdrun
Category: mdrun
Target version: 2020
Difficulty: uncategorized

Description
For efficient direct GPU communications peer to peer access between GPUs in the run should be enabled.

This functionality should however be implemented such that all/most errors are handled explicitly and the function only aborts the run if a to be fatal error is detected, otherwise, as it is only a performance concern the run should continue.

Related: current working assumption is that even if peer access is not enabled direct copy should not be sower than staged copy, but as we are not sure, we might want to consider disabling the GPU direct copy if enabling peer access fails.

Related issues:
Related to GROMACS - Feature #2890: GPU Halo Exchange
Related to GROMACS - Feature #2891: PME/PP GPU communications

Associated revisions
Revision 643e75da - 10/23/2019 08:47 AM - Alan Gray
Enable GPU Peer Access in GPU Utilities

When using the new GPU communication features, enabling peer access between pairs of GPUs (where supported) will allow peer-to-peer communications. In this patch the CUDA code to enable peer access is introduced into central GPU utilities and called from do_md.

Implements #3087

Change-Id: If668366b76d49f7b624eedb501f8af19135c4386

History
#1 - 09/12/2019 04:24 PM - Sziliárd Páll
- Related to Feature #2890: GPU Halo Exchange added

#2 - 09/12/2019 04:24 PM - Sziliárd Páll
- Related to Feature #2891: PME/PP GPU communications added

#3 - 10/24/2019 10:59 AM - Alan Gray
- Status changed from New to Closed