NVIDIA GPUDirect Storage

Release Notes
Table of Contents

Chapter 1. Introduction........................................................................................................ 1
Chapter 2. New Features..................................................................................................... 2
Chapter 3. MLNX_OFED and Filesystem Requirements..................................................... 3
Chapter 4. Improvements.....................................................................................................4
Chapter 5. Included Packages............................................................................................. 5
Chapter 6. Minor Updates and Bug Fixes.............................................................................6
Chapter 7. Known Issues......................................................................................................7
Chapter 8. Known Limitations..............................................................................................8
Chapter 9. Deprecations.......................................................................................................9
Chapter 1. Introduction

Release information for NVIDIA® GPUDirect® Storage (GDS) for developers and users.

GDS is the newest addition to the GPUDirect family. GDS enables a direct data path for direct memory access (DMA) transfers between GPU memory and storage, which avoids a bounce buffer through the CPU. This direct path increases system bandwidth and decreases the latency and utilization load on the CPU.

GDS is enabled on distributed filesystems like DDN Exascaler®, WekaFS, and VAST. GDS documents and online resources provide additional context for the optimal use of and understanding of GPUDirect Storage.

Refer to the following guides for more information about GDS:

‣ GPUDirect Storage Design Guide
‣ GPUDirect Storage Overview Guide
‣ cuFile API Reference Guide
‣ GPUDirect Storage Best Practices Guide
‣ GPUDirect Storage Installation and Troubleshooting Guide
‣ GPUDirect Storage O_DIRECT Requirements Guide

To learn more about GDS, refer to the following blogs:

‣ GPUDirect Storage: A Direct Path Between Storage and GPU Memory
‣ The Magnum IO series.
Chapter 2. New Features

The following features have been added since the 0.9 release:

- Compatibility with POSIX IO is enabled by default
- Alpha level support for RHEL 8.3
- GDS is available as Technical preview for DGX OS
- Support for MLNX_OFED 5.3 for NVMe and NVMeOF
- Support for Excelero™ NVMesh devices
- Support for ScaleFlux computational storage
- Integration with DALI and Pytorch
- Experimental RAPIDS integration for cuDF, unoptimized, reads only
Chapter 3. MLNX_OFED and Filesystem Requirements

Here are the MLNX_OFED and filesystem requirements for GDS:

- Ubuntu 18.04 and 20.04, RHEL 8.3
- MLNX_OFED 5.1-2.5.8.0 and later, which supports NVMe NVMeoF, NFSoRDMA (VAST) on Linux kernel 4.15.x and 5.4.x

You need to install MLNX_OFED **before** you install GDS. Refer to Installing GPUDirect Storage for more information about installing MLNX_OFED.

- The following distributed filesystems:
  - WekaFS 3.8.0
  - DDN Exascaler 5.2
  - VAST 3.4

- Block/other file systems supported:
  - ScaleFlux CSD
  - NVMeMesh
  - PavillionData
Chapter 4. Improvements

The following improvements have been made to GDS since the 0.9.1 release.

- Added Dynamic Routing for distributed file systems
- Added Static library support
- Added RHEL8.3 support
- Made Compatible Mode default for GDS
- NUMA-based PCIe peer affinity computation in the nvidia_fs driver
- Minor bug fixes to nvidia_fs driver to improve resiliency and supportability
- Added stats for dynamic routing in the gds_stats tool
- Minor bug fixes in GDS tools to improve resiliency and supportability
- Minor enhancements and bug fixes in libcufile to improve resiliency and supportability
Chapter 5. Included Packages

The GDS package contains the following Debian packages:

- gds-tools-10-1_0.95.0.86-1_amd64.deb
- libcufile-10-1_0.95.0.86-1_amd64.deb
- libcufile-dev-10-1_0.95.0.86-1_amd64.deb
- nvidia-fs_2.6.86-1_amd64.deb
- nvidia-fs-dkms_2.6.86-1_amd64.deb
- nvidia-gds-10-1_10.1.20210407-1_amd64.deb
- nvidia-gds_10.1.20210407-1_amd64.deb

Note: Each component has a README file. For example, for gds-tools, the README file is in the `/usr/local/CUDA-X.Y/gds/tools/` directory.
Chapter 6. Minor Updates and Bug Fixes

The following minor updates and bug fixes were made after version 0.9.1:

- Improved cuFile library cleanup and error reporting for internal cuda errors.
- Fixed handling of cuFile IO operations for platforms that do not have NUMA information.
- Removed dependency on application CUDA context for internal GDS buffer allocations.
- Fixed nvidia-fs driver bug in reading sparse files with more than 768 holes in a single IO read.
- Fixed nvidia-fs makefile to handle MOFED src paths correctly for different distros.
- Fixed handling of IO completion checks when called in interrupt mode.
- Fixed incorrect error handling of NFS preads on GDS allocated internal buffers.
- Fixed Python script errors in gdscheck.py and python3 is now the default interpreter.
- Fixed gdscheck verification mode when buffers are not aligned.
- Fixed GDS packaging to reduce dependencies on libcuda and libjson packages.
Chapter 7. Known Issues

This section provides information about the known issues in this release of GDS.

- `nvidia_p2p_get_pages` performance has severely regressed in NVIDIA driver 440.33.01 compared to 418.116.00 in DGX-2.
- For Lustre filesystem:
  - with stripe count > 1, `cuFileRead` and `cuFileWrite` do not work with poll mode enabled for versions older than 2.12.5_ddn10.
  - with 2.12.5_ddn10, any reads beyond EOF causes a `BUG_ON` inside nvidia-fs.
- RHEL8.3 does not have default udev rules for detecting RAID members, which disables GDS on RAID volumes. Please refer to the section Adding udev Rules for RAID Volumes in the GPUDirect Storage Installation and Troubleshooting Guide.
- The `nfs-rdma` module in MLNX_OFED 5.3-1.0.0.1 does not compile. Expected to be fixed in an upcoming releases of MLNX_OFED 5.3.
- MLNX_OFED 5.3 has been tested with limited file systems (Ext4, DDN ExaScaler).
- CUDA 11.3 has been tested with CUDA driver versions 450.x, and 460.x. It is not expected to work with the default driver version (465 and above) that comes with CUDA toolkit 11.3.
- The `gdsio` tool takes only numeric values for "-a" and "-n" fields.
Chapter 8. Known Limitations

This section provides information about the known limitations in this release of GDS.

- For Lustre, checksum is disabled in the read/write IO path.
- For Weka, checksum is disabled in the read/write IO path.
- There is no per GPU configuration for cache and BAR memory usage.
- cuFile configuration is decided at application load time.
- cuFile APIs are not supported with applications using the `fork()` system call.
- There is no command or API to purge GDS internal caches without calling the `cuFileDriverClose` API.
- MLNX_OFED 5.3 has been tested with following file systems: Ext4, DDN ExaScaler.
- `max_direct_io_size_kb` in `cufile.json` should be multiples of 64K.
Chapter 9. Deprecations

This section provides information about the APIs that have been deprecated.
No APIs have been deprecated since version 0.9.
Notice

This document is provided for information purposes only and shall not be regarded as a warranty of a certain functionality, condition, or quality of a product. NVIDIA Corporation ("NVIDIA") makes no representations or warranties, expressed or implied, as to the accuracy or completeness of the information contained in this document and assumes no responsibility for any errors contained herein. NVIDIA shall have no liability for the consequences or use of such information or for any infringement of patents or other rights of third parties that may result from its use. This document is not a commitment to develop, release, or deliver any Material (defined below), code, or functionality.

NVIDIA reserves the right to make corrections, modifications, enhancements, improvements, and any other changes to this document, at any time without notice.

Customer should obtain the latest relevant information before placing orders and should verify that such information is current and complete.

NVIDIA products are sold subject to the NVIDIA standard terms and conditions of sale supplied at the time of order acknowledgement, unless otherwise agreed in an individual sales agreement signed by authorized representatives of NVIDIA and customer ("Terms of Sale"). NVIDIA hereby expressly objects to applying any customer general terms and conditions with regards to the purchase of the NVIDIA product referenced in this document. No contractual obligations are formed either directly or indirectly by this document.

NVIDIA reserves the right to make corrections, modifications, enhancements, improvements, and any other changes to this document, at any time without notice.

NVIDIA products are not designed, authorized, or warranted to be suitable for use in medical, military, aircraft, space, or life support equipment, nor in applications where failure or malfunction of the NVIDIA product can reasonably be expected to result in personal injury, death, or property or environmental damage. NVIDIA accepts no liability for inclusion and/or use of NVIDIA products in such equipment or applications and therefore such inclusion and/or use is at customer’s own risk.

NVIDIA makes no representation or warranty that products based on this document will be suitable for any specified use. Testing of all parameters of each product is not necessarily performed by NVIDIA. It is customer’s sole responsibility to evaluate and determine the applicability of any information contained in this document, ensure the product is suitable and fit for the application planned by customer, and perform the necessary testing for the application in order to avoid a default of the application or the product. Weaknesses in customer’s product designs may affect the quality and reliability of the NVIDIA product and may result in additional or different conditions and/or requirements beyond those contained in this document. NVIDIA accepts no liability related to any default, damage, costs, or problem which may be based on or attributable to: (i) the use of the NVIDIA product in any manner that is contrary to this document or (ii) customer product designs.

No license, either expressed or implied, is granted under any NVIDIA patent right, copyright, or other NVIDIA intellectual property right under this document. Information published by NVIDIA regarding third-party products or services does not constitute a license from NVIDIA to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property rights of the third party, or a license from NVIDIA under the patents or other intellectual property rights of NVIDIA.

Reproduction of information in this document is permissible only if approved in advance by NVIDIA in writing, reproduced without alteration and in full compliance with all applicable export laws and regulations, and accompanied by all associated conditions, limitations, and notices.

THIS DOCUMENT AND ALL NVIDIA DESIGN SPECIFICATIONS, REFERENCE BOARDS, FILES, DRAWINGS, DIAGNOSTICS, LISTS, AND OTHER DOCUMENTS (TOGETHER AND SEPARATELY, “MATERIALS”) ARE BEING PROVIDED “AS IS.” NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE MATERIALS, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE. TO THE EXTENT NOT PROHIBITED BY LAW, IN NO EVENT WILL NVIDIA BE LIABLE FOR ANY DAMAGES, INCLUDING WITHOUT LIMITATION ANY DIRECT, INDIRECT, SPECIAL, INCIDENTAL, PUNITIVE, OR CONSEQUENTIAL DAMAGES, HOWEVER CAUSED AND REGARDLESS OF THE THEORY OF LIABILITY, ARISING OUT OF ANY USE OF THIS DOCUMENT, EVEN IF NVIDIA HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. NOTWITHSTANDING ANY DAMAGES THAT CUSTOMER MIGHT INCUR FOR ANY REASON WHATSOEVER, NVIDIA’S AGGREGATE AND CUMULATIVE LIABILITY TOWARDS CUSTOMER FOR THE PRODUCTS DESCRIBED HEREIN SHALL BE LIMITED IN ACCORDANCE WITH THE TERMS OF SALE FOR THE PRODUCT.
No license, either expressed or implied, is granted under any NVIDIA patent right, copyright, or other NVIDIA intellectual property right under this document. Information published by NVIDIA regarding third-party products or services does not constitute a license from NVIDIA to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property rights of the third party, or a license from NVIDIA under the patents or other intellectual property rights of NVIDIA.

Reproduction of information in this document is permissible only if approved in advance by NVIDIA in writing, reproduced without alteration and in full compliance with all applicable export laws and regulations, and accompanied by all associated conditions, limitations, and notices.

THIS DOCUMENT AND ALL NVIDIA DESIGN SPECIFICATIONS, REFERENCE BOARDS, FILES, DRAWINGS, DIAGNOSTICS, LISTS, AND OTHER DOCUMENTS (TOGETHER AND SEPARATELY, "MATERIALS") ARE BEING PROVIDED "AS IS." NVIDIA MAKES NO WARRANTIES, EXPRESSED, IMPLIED, STATUTORY, OR OTHERWISE WITH RESPECT TO THE MATERIALS, AND EXPRESSLY DISCLAIMS ALL IMPLIED WARRANTIES OF NONINFRINGEMENT, MERCHANTABILITY, AND FITNESS FOR A PARTICULAR PURPOSE. TO THE EXTENT NOT PROHIBITED BY LAW, IN NO EVENT WILL NVIDIA BE LIABLE FOR ANY DAMAGES, INCLUDING WITHOUT LIMITATION ANY DIRECT, INDIRECT, SPECIAL, INCIDENTAL, PUNITIVE, OR CONSEQUENTIAL DAMAGES, HOWEVER CAUSED AND REGARDLESS OF THE THEORY OF LIABILITY, ARISING OUT OF ANY USE OF THIS DOCUMENT, EVEN IF NVIDIA HAS BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES. NOTWITHSTANDING ANY DAMAGES THAT CUSTOMER MIGHT INCUR FOR ANY REASON WHATSOEVER, NVIDIA'S AGGREGATE AND ACCUMULATIVE LIABILITY TOWARDS CUSTOMER FOR THE PRODUCTS DESCRIBED HEREIN SHALL BE LIMITED IN ACCORDANCE WITH THE TERMS OF SALE FOR THE PRODUCT.

VESA DisplayPort

DisplayPort and DisplayPort Compliance Logo, DisplayPort Compliance Logo for Dual-mode Sources, and DisplayPort Compliance Logo for Active Cables are trademarks owned by the Video Electronics Standards Association in the United States and other countries.

HDMI

HDMI, the HDMI logo, and High-Definition Multimedia Interface are trademarks or registered trademarks of HDMI Licensing LLC.

OpenCL

OpenCL is a trademark of Apple Inc. used under license to the Khronos Group Inc.