NVIDIA DATA LOADING LIBRARY (DALI)

RN-09096-001 _v01 | April 2020

Release Notes

TABLE OF CONTENTS

Chapter 1.
DALI OVERVIEW

NVIDIA® Data Loading Library™ (DALI) accelerates the preprocessing of input data for deep learning applications. By offloading augmentations onto GPUs, DALI addresses performance bottlenecks in today's computer vision deep learning applications that include complex, multi-stage data augmentation steps. With DALI beta release, deep learning researchers can scale training performance on image classification models such as ResNet-50 with MXNet, TensorFlow, and PyTorch across Amazon Web Services P3 8 GPU instances or DGX-1 systems with Volta architecture. Framework developers will have less duplication due to better code reuse and maintainability.

DALI offers both performance and flexibility of accelerating different data pipelines (graphs that can have multiple outputs and inputs), as a single library, that can be easily integrated into different deep learning training and inference applications.

Chapter 2.
DALI RELEASE 0.21.0 BETA

The DALI 0.21.0 is a beta release, therefore, all features, functionality, and performance will likely be limited.

Using DALI 0.21.0 Beta

To upgrade to DALI 0.21.0 beta from an older version of DALI, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

> The internal DALI C++ API used for operator's implementation, and the C++ API that enables using DALI as a library from native code, is not yet officially supported. Hence these APIs may change in the next release without advance notice.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Introduced experimental Functional API (1598)
    - Operators can be used directly with a single call, no need to create an instance with a constructor.
    - ◦DALI pipeline can be used in Context Manager.
    - ◦There is no need to subclass Pipeline.
- Simplified usage of **`ExternalSource`** (#1598, #1832) - it accepts callbacks or generators as a parameter.
- Added Python 3.8 build and support (#1782.
- Allowed seed to be set for serialized pipeline (#1844)
- New operators:
    - **`ToDecibels`** GPU operator (#1837)
    - One hot encoding CPU operator (##1807)

Fixed Issues

This DALI release includes the following fixes.

- Fixed transpose bugs - degenerate dims and non-uniform batch on GPU (#1817)
- Fixed a leak of the last created DALI pipeline instance (#1845)
- Make SSDRandomCrop calculate crop window in double precision (#1848)

Deprecated Features

- CUDA 9 support will soon end in a future release.

Known Issues

- The video loader operator requires that the key frames occur at a minimum every 10 to 15 frames of the video stream. If the key frames occur at a lesser frequency, then the returned frames may be out of sync.
- The DALI TensorFlow plugin may not be compatible with TensorFlow versions 1.15.0 and/or later. If the user wants to use DALI with the TensorFlow version which doesn't have prebuilt plugin binary shipped with DALI it requires the gcc compiler that matches the one used to build TensorFlow (gcc 4.8.4 or gcc, 4.8.5 or 5.4, depending on the particular version) is present on the system.
- Due to some known issues with meltdown/spectra mitigations and DALI, DALI shows best performance when running in Docker with escalated privileges, for example:
    - **`privileged=yes`** in Extra Settings for AWS data points
    - **`--privileged`** or **`--security-opt seccomp=unconfined`** for bare Docker

Chapter 3.
DALI RELEASE 0.20.0 BETA

The DALI 0.20.0 is a beta release, therefore, all features, functionality, and performance will likely be limited.

Using DALI 0.20.0 Beta

To upgrade to DALI 0.20.0 beta from an older version of DALI, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

> The internal DALI C++ API used for operator's implementation, and the C++ API that enables using DALI as a library from native code, is not yet officially supported. Hence these APIs may change in the next release without advance notice.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Added operators:
  - Spectrogram for GPU (#1786)
  - MelFilterBank for GPU (#1796)
- Allow align-only behavior in Pad operator by treating shape argument as minimum shape (#1764)
- Added **`data_ptr`** method to Tensor and TensorList (#1773) - it enables **`__array_interface__`** and **`__cuda_array_interface__`** support.
- Extended shape support in DALI Dataset for TensorFlow (#1723)
- Documentation improvements: layouts, Python API.
- Added Gluon iterator plugin (#1683)

Fixed Issues

This DALI release includes the following fixes.

- Implemented several fixes in the documentation and build system.
- Fixed precision loss in **`CropWindowGenerator`** (#1735)
- Fixed **`BBFlip`** stream synchronization issue (#1738)

Deprecated Features

- CUDA 9 support will soon end in a future release

Known Issues

- The video loader operator requires that the key frames occur at a minimum every 10 to 15 frames of the video stream. If the key frames occur at a lesser frequency, then the returned frames may be out of sync.
- The DALI TensorFlow plugin may not be compatible with TensorFlow versions 1.15.0 and/or later. If the user wants to use DALI with the TensorFlow version which doesn't have prebuilt plugin binary shipped with DALI it requires the gcc compiler that matches the one used to build TensorFlow (gcc 4.8.4 or gcc, 4.8.5 or 5.4, depending on the particular version) is present on the system.
- Due to some known issues with meltdown/spectra mitigations and DALI, DALI shows best performance when running in Docker with escalated privileges, for example:
    - **`privileged=yes`** in Extra Settings for AWS data points
    - **`--privileged`** or **`--security-opt seccomp=unconfined`** for bare Docker

Chapter 4. DALI RELEASE 0.19.0 BETA

The DALI 0.19.0 is a beta release, therefore, all features, functionality, and performance will likely be limited.

Using DALI 0.19.0 Beta

To upgrade to DALI 0.19.0 beta from an older version of DALI, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

> The internal DALI C++ API used for operator's implementation, and the C++ API that enables using DALI as a library from native code, is not yet officially supported. Hence these APIs may change in the next release without advance notice.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Reduced the size of the underlying allocation for non-pinned host buffers when the new size is less than 90% of current allocation size. Added several options for memory usage configuration (#1712, #1719) - up to 50% reduction in some cases.
- New Operators:
  - Erase CPU (#1609)
  - Transpose CPU (#1677)
  - Audio non silence CPU (#1701)
  - Constant CPU & GPU (#1699)
  - Normalize (sample/batch wide) (#1670)
- Improved Operators:
  - Reshape (#1634, #1691, #1714)
- Python operators for GPU with examples (#1655, #1681, #1685, #1715)

- Mitigate the OS file-mhax limit in the **`VideoReader`** (#1659)
- Support for array subscript operator **`[]`** for TensorList in Python API (#1682)

Fixed Issues

This DALI release includes the following fixes.

- Updated examples with COCO data set and fix reader behavior for padding (#1557)
- Fixed formatting of the example in the FW iterators docs (#1649)
- Fixed torch stream initialization in TorchPythonFunction (#1681)
- Fixed multi-channel fill value check in Erase operator (#1675)
- Added synchronization to read/write operations in image decoder cache (#1702)
- Fixed Buffer linkage and Reshape bug (#1714)
- Fixed Pad operator bug (#1713)

Deprecated Features

- CUDA 9 support will end in several releases (#1684)
- Access to Tensors of **`TensorListCPU`** and **`TensorListGPU`** with **`at`** was replaced by array subscript operator. (#1682)

Known Issues

- The video loader operator requires that the key frames occur at a minimum every 10 to 15 frames of the video stream. If the key frames occur at a lesser frequency, then the returned frames may be out of sync.
- The DALI TensorFlow plugin may not be compatible with TensorFlow versions 1.15.0 and/or later. If the user wants to use DALI with the TensorFlow version which doesn't have prebuilt plugin binary shipped with DALI it requires the gcc compiler that matches the one used to build TensorFlow (gcc 4.8.4 or gcc, 4.8.5 or 5.4, depending on the particular version) is present on the system.
- Due to some known issues with meltdown/spectra mitigations and DALI, DALI shows best performance when running in Docker with escalated privileges, for example:
  - **`privileged=yes`** in Extra Settings for AWS data points
  - **`--privileged`** or **`--security-opt seccomp=unconfined`** for bare Docker

Chapter 5. DALI RELEASE 0.18.0 BETA

The DALI 0.18.0 is a beta release, therefore, all features, functionality, and performance will likely be limited.

Using DALI 0.18.0 Beta

To upgrade to DALI 0.18.0 beta from an older version of DALI, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

> The internal DALI C++ API used for operator's implementation, and the C++ API that enables using DALI as a library from native code, is not yet officially supported. Hence these APIs may change in the next release without advance notice.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Documentation improvements: `explicit __call__ operator` docstring, arithmetic expressions, and formatting.
- New audio operators functionality:
  - Audio decoder extended with resampling and downmixing (#1582)
  - MFCC CPU Operator (#1577)
- New signal processing kernels:
  - Extract windows GPU kernel (#1538)
  - Added DCT 1D CPU kernel (#1569)
- Arithmetic expressions improvements:
  - Added comparison operators and bool handling in arithmetic ops (#1541)
  - Added bitwise operators: `&, |, ^` (#1594)

  - Added support for Unary Ops: + and - (#1392)
- Video support improvements:
  - Improved support for labels in VideoReader - now the user can provide labels for selected frame ranges as well as obtain frame numbers and timestamps. (#1500)
- Added an ability to return duplicate outputs from the DALI pipeline. (#1556)
- Removed the need to have GPU available when DALI is just imported. (#1601)

Fixed Issues

This DALI release includes the following fixes.

- Fixed **`ToContiguousXXX`** for more than 2 inputs. (#1572)
- Fixed a race condition in **`GetGPUAllocator`**. (#1575)
- Used a different stream base for different videos so it is possible to decode videos with a different time base at the same time. (#1592)
- Fixed **`Transpose`** operator for batch size 1 as well as 1 channel images. (#1624)

Breaking Changes

- Python 2.7 is no longer available. To stay up-to-date with DALI, upgrade to Python 3.5 or later.

Known Issues

- The video loader operator requires that the key frames occur at a minimum every 10 to 15 frames of the video stream. If the key frames occur at a lesser frequency, then the returned frames may be out of sync.
- The DALI TensorFlow plugin may not be compatible with TensorFlow versions 1.15.0 and/or later. If the user wants to use DALI with the TensorFlow version which doesn’t have prebuilt plugin binary shipped with DALI it requires the gcc compiler that matches the one used to build TensorFlow (gcc 4.8.4 or gcc, 4.8.5 or 5.4, depending on the particular version) is present on the system.
- Due to some known issues with meltdown/spectra mitigations and DALI, DALI shows best performance when running in Docker with escalated privileges, for example:
  - **`privileged=yes`** in Extra Settings for AWS data points
  - **`--privileged`** or **`--security-opt seccomp=unconfined`** for bare Docker

Chapter 6. DALI RELEASE 0.17.0 BETA

The DALI 0.17.0 is a beta release, therefore, all features, functionality, and performance will likely be limited.

Using DALI 0.17.0 Beta

To upgrade to DALI 0.17.0 beta from an older version of DALI, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

> The internal DALI C++ API used for operator's implementation, and the C++ API that enables using DALI as a library from native code, is not yet officially supported. Hence these APIs may change in the next release without advance notice.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- New operators:
  - **`NormalDistribution`** (#1529)
- New operators (CPU backend) with accompanying kernels for audio and signal processing:
  - **`AudioDecoder`** (#1481)
  - **`PowerSpectrum`** (#1460)
  - **`Spectrogram`** (#1468)
  - **`Preemphasis`** (#1515)
  - **`ToDecibels`** (#1518)
  - **`MelFilterBank`** (#1535)
- Improvements to existing operators:

  - Sequence and 3D support in **`Flip`** operator (#1439)
  - Warp 3D (#1366, #1442)
  - Improve the accuracy of 3D rotation (#1466)
  - Various fixes and improvements to **`ColorTwist`**, **`BrightnessContrast`** and **`HSV`** operators
- Added integration with PaddlePaddle.
- Added Jupyter Notebooks with example usage of arithmetic ops (#1438)

Fixed Issues

This DALI release includes the following fixes.

- Fixed race between consecutive invocations of Mixed-stage. (#1493)
- Fixed **`ExternalSource`** for GPU (#1452)
- Fixed stream usage in **`HSV`** and **`BrighnessContrast`**. (#1566)
- Added wait for thread pool to finish work in **`BrightnessConstrast`** (#1549)

Breaking Changes

- The **`asCPU`** method is no longer available and has been replaced with **`as_cpu`**.

Deprecated Features

- DALI 0.17 is the last official release for Python 2.7, which reaches the end of life on January 1st, 2020. To stay up to date with DALI, upgrade to Python 3.5 or later.
- **`ColorTwist`** operator was deprecated and replaced by **`BrightnessContrast`** and **`HSV`** operators cleanup (#1532)

Known Issues

- The video loader operator requires that the key frames occur at a minimum every 10 to 15 frames of the video stream. If the key frames occur at a lesser frequency, then the returned frames may be out of sync.
- The DALI TensorFlow plugin may not be compatible with TensorFlow versions 1.15.0 and/or later. If the user wants to use DALI with the TensorFlow version which doesn't have prebuilt plugin binary shipped with DALI it requires the gcc compiler that matches the one used to build TensorFlow (gcc 4.8.4 or gcc, 4.8.5 or 5.4, depending on the particular version) is present on the system.
- Due to some known issues with meltdown/spectra mitigations and DALI, DALI shows best performance when running in Docker with escalated privileges, for example:
  - **`privileged=yes`** in Extra Settings for AWS data points
  - **`--privileged`** or **`--security-opt seccomp=unconfined`** for bare Docker

Chapter 7. DALI RELEASE 0.16.0 BETA

The DALI 0.16.0 is a beta release, therefore, all features, functionality, and performance will likely be limited.

Using DALI 0.16.0 Beta

To upgrade to DALI 0.16.0 beta from an older version of DALI, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

> The internal DALI C++ API used for operators implementation, and the C++ API that enables using DALI as a library from native code, are not yet officially supported. Hence these APIs may change in the next release without advance notice.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- SupportOps are removed and CPU operators can now be used in their place (#1423)
  - Allows for more generic shapes to be provided as argument inputs
  - Allows to return the data previously generated by support Op (like **`ops.Uniform`**) from pipeline
  - Internally Argument Inputs accept TensorVector instead of Tensor.
- Extended support for TensorFlow dataset API:
  - GPU (#1354) and initial multi-GPU (#1382) support
  - TensorFlow 1.15 and 2.0 support for TensorFlow dataset (#1395)
- New operators:
  - HSV manipulation operator for GPU and CPU (#1338)
  - GPU **`dltensor`** operator (#1261)

    - Pad operator (#1180)
- Initial support for arithmetic operators (#1322, #1355, #1426, #1449):
    - Regular Python arithmetic expressions inside DALI pipeline definition
    - Supported binary operators: +, -, *, /, //
- Improvements to current operators:
    - Support float16 in Cast GPU operator (#1368)
    - Rotate operator using a new Warp kernel (#1403)
    - Optimized implementation for **`WarpAffine`** (#1387)
    - Added **`additional_decode_surfaces`** parameter to videoreader (#1393)
- New and updated examples:
    - **`BrightnessContrast`** operator example (#1414)
    - New warp example (#1158)
- Build system improvements for architectures other than x86
- Operators moved to separate library (#1380, #1384, #1406)

Fixed Issues

This DALI release includes the following fixes.

- Fixed DALI TensorFlow plugin CXX11 ABI issue (#1361)
- Fixed DALI TensorFlow installation for TensorFlow 2.0 (#1386)

Breaking Changes

- Support operators were removed and their usage replaced with CPU operators. Old support operators were converted to CPU operators.
- DALI operators were moved into a separate library to detach them from the pipeline. It is linked into the target DALI library.

Deprecated Features

- Removed the following deprecated operators:
    - **`HostDecoder`** and **`nvJPEGDecoder`** (generic **`ImageDecoder`** should be used instead) (#1398)
    - **`NormalizePermute`** (**`CropMirrorNormalize`** should be used instead) (#1402)
    - DALI 0.17 is the last official release for Python 2.7, which reaches the end of life on January 1st, 2020. To stay up to date with DALI, upgrade to Python 3.5 or later.
    - **`Crop`**, **`CropMirrorNormalize`** and **`Slice`** operator possible output types are limited to one of **`uint8_t, int16_t, uint16_t, int32_t, float, float16`** or passing through the input type (#1418).

Known Issues

- The video loader operator requires that the key frames occur at a minimum every 10 to 15 frames of the video stream. If the key frames occur at a lesser frequency, then the returned frames may be out of sync.
- DALI TensorFlow plugin may not be compatible with TensorFlow versions 1.15.0 and/or later. If the user wants to use DALI with TensorFlow version which doesn't have prebuilt plugin binary shipped with DALI it requires the gcc compiler that matches the one used to build TensorFlow (gcc 4.8.4 or gcc, 4.8.5 or 5.4, depending on the particular version) is present on the system.
- Due to some known issues with meltdown/spectra mitigations and DALI, DALI shows best performance when run in Docker with escalated privileges, for example:
  - **`privileged=yes`** in Extra Settings for AWS data points
  - **`--privileged`** or **`--security-opt seccomp=unconfined`** for bare Docker

Chapter 8.
DALI RELEASE 0.15.0 BETA

The DALI 0.15.0 is a beta release. Hence, for all the features, the functionality and performance will likely be limited.

Using DALI 0.15.0 Beta

To upgrade to DALI 0.15.0 beta from an older version of DALI, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

> The internal DALI C++ API used for operators implementation, and the C++ API that enables using DALI as a library from native code, are not yet officially supported. Hence these APIs may change in the next release without advance notice.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- New operators and extended functionality of the following:
  - CPU **`DLTensor`** operator (#1233)
  - Added Lookup table operator (#1251)
  - **`BrightnessContrast`** operator (#1188)
  - Added **`CropMirrorNormalize`** 3D support (#1326)
  - Added **`Reshape`** operator (#1327)
- New kernels:
  - Linear transformation GPU kernel (#1262)
  - Linear Transformation kernel for CPU (#1300)
- Video reader improvements:
  - Improved heuristic for variable frame rate detection (#1242)

  - Added support for decoding multiple resolution videos in the same pipeline (#1144)
- Tensor layout (#1237) and Layout refactor (#1250) - new, domain agnostic, flexible description of a tensor layout that is compatible with string representation.
- Multi channel image decoding - supporting more than 3 channels:
  - **`ImageDecoder libtiff`** implementation (#1264)
  - Handle **`nchannels>3`** in **`ImageDecoder`** (#1285)
  - Parse number of channels in **`PNGImage::PeekShape`** (#1288)
  - **`TiffImage::PeekShapeImpl`** parse and return number of channels (#1304)
  - Add parsing of number of channels in **`JpegImage::PeekShapeImpl`** (#1306)
- DALI TensorFlow plugin installation improvements.

Fixed Issues

This DALI release includes the following fixes.

- Fixed **`Transpose`** operator when data shape with dimension of size 1 (#1244)
- Fixed problems with seeking when stream **`start_time`** is **`!= 0`**. (#1287)
- Suppressed warning when **`FileReader`** encounters dot and dot-dot entries (#1318)
- Fixed **`last_batch_padded`** docs (#1314)
- Fixed DALI TensorFlow plugin CXX11 ABI issue (#1361)

Breaking Changes

- Changed **`Outputs`** and **`SharedOuputs`** return type to tuple (#1243)

Known Issues

- The new video reader operator requires NVIDIA VIDEO CODEC SDK support in the platform. Prior to 19.01, the NVIDIA GPU Cloud (NGC) optimized containers lack this functionality in the default configuration. To enable the functionality, run the container with the **`video`** capability enabled, as shown below:

```
-e "NVIDIA_DRIVER_CAPABILITIES=compute,utility,video"
```

- The video loader operator requires that the key frames occur at a minimum every 10 to 15 frames of the video stream. If the key frames occur at a lesser frequency, then the returned frames may be out of sync.
- DALI TensorFlow plugin may not be compatible with TensorFlow versions 1.15.0 and/or later. If the user wants to use DALI with TensorFlow version which doesn't have prebuilt plugin binary shipped with DALI it requires the gcc compiler that matches the one used to build TensorFlow (gcc 4.8.4 or gcc, 4.8.5 or 5.4, depending on the particular version) is present on the system.

Chapter 9.
DALI RELEASE 0.14.0 BETA

Using DALI 0.14.0 Beta

The DALI 0.14.0 can be used with the 19.09 NVIDIA GPU Cloud (NGC) optimized container for MXNet, PyTorch, and TensorFlow.

To upgrade to DALI 0.14.0 beta from an older version of DALI, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

> The internal DALI C++ API used for operators implementation, and the C++ API that enables using DALI as a library from native code, are not yet officially supported. Hence these APIs may change in the next release without advance notice.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Added an example for using DALI with multiple GPUs.
- Added a **`Shapes`** operator which returns the input shape as a tensor.
- Added conda recipes for DALI used for pre-built packages in Watson Machine Learning Community Edition (IBM Power and x86 systems).
- Extended the **`ExternalSource`** operator so it can hold more than one batch of inputs at a time.

Fixed Issues

This DALI release includes the following fixes.

- Fixed FP16 bug from #1129 and added FP16 test case. (#1160)
- Fixed framework iterators behavior when **`iter_setup`** raises **`StopIteration`**. (#1136)

- Fixed **nvjpeg** legacy API. (#1179)
- Fixed conversions to INT64 and UINT64. (#1205)
- Fixed DALI TensorFlow install for conda environments. (#1214)

Breaking Changes

- Extended external source operator capacity (#1127) - it now requires input to be set for every iteration.
- Adjusted **Operator::Run** to take reference instead of pointer (#1168) (C++ Backend API).

Deprecated Features

- Python **.cpu()** function for **EdgeReference** was removed to reduce confusion. (#1181)

Known Issues

- The new video reader operator requires NVIDIA VIDEO CODEC SDK support in the platform. Prior to 19.01, the NVIDIA GPU Cloud (NGC) optimized containers lack this functionality in the default configuration. To enable the functionality, run the container with the "video" capability enabled, as shown below:

  ```
  -e "NVIDIA_DRIVER_CAPABILITIES=compute,utility,video"
  ```
- The video loader operator requires that the key frames occur at a minimum every 10 to 15 frames of the video stream. If the key frames occur at a lesser frequency, then the returned frames may be out of sync.
- DALI TensorFlow plugin may be not compatible with TensorFlow 1.14.0 release. The DALI TensorFlow plugin requires that the **gcc** compiler that matches the one used to build TensorFlow (**gcc** 4.8.4 or **gcc** 4.8.5, depending on the particular version) be present on the system.

Chapter 10. DALI RELEASE 0.13.0 BETA

The DALI 0.13.0 is a beta release. Hence, for all the features, the functionality and performance will likely be limited.

Using DALI 0.13.0 Beta

The DALI 0.13.0 can be used with the 19.08 NVIDIA GPU Cloud (NGC) optimized container for MXNet, PyTorch, and TensorFlow. Also, the 19.08 container will be shipped with DALI 0.13.0.

To upgrade to DALI 0.13.0 beta from an older version of DALI, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

> The internal DALI C++ API used for operators implementation, and the C++ API that enables using DALI as a library from native code, are not yet officially supported. Hence these APIs may change in the next release without advance notice.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Added new:
  - Fast coco reader
  - TorchPythonFunction operator
  - Sink operators
- Reworked how the reader pick samples from the shuffling buffer.
- Used DALI_extra for test data.
- Added checks to see if Python API is not mixed between simple, scheduled and iterator APIs.
- Added support for reading video files with labels using **`file_list`** argument.

Fixed Issues

This DALI release includes the following fixes.

- Restored support of `use_batched_decode` argument in `nvJPEGDecoder` operator (only for legacy nvJPEGDecoder implementation).
- Fixed FP16 support in DALI TensorFlow plugin.
- Fixed Python operator with side effects.
- Fixed a race condition in async pipeline executor.
- Disabled `video_reader_op` test when NVDEC is disabled.
- Fixed sampling of chroma in the VideoReader operation.
- Fix detection pipeline example.

Breaking Changes

- Reader sampling from shuffling buffer was adjusted. Now samples are not mixed between epochs.

Deprecated Features

- Deprecated `NormalizePermute` in favor of `CropMirrorNormalize`
- Multiple Input Sets handling was removed from backend and is only Python level syntactic sugar.

Known Issues

- The new video reader operator requires NVIDIA VIDEO CODEC SDK support in the platform. Prior to 19.01, the NVIDIA GPU Cloud (NGC) optimized containers lack this functionality in the default configuration. To enable the functionality, run the container with the "video" capability enabled, as shown below:

  ```
  -e "NVIDIA_DRIVER_CAPABILITIES=compute,utility,video"
  ```
- The video loader operator requires that the key frames occur at a minimum every 10 to 15 frames of the video stream. If the key frames occur at a lesser frequency, then the returned frames may be out of sync.
- DALI TensorFlow plugin may be not compatible with TensorFlow 1.14.0 release. The DALI TensorFlow plugin requires that the `gcc` compiler that matches the one used to build TensorFlow (`gcc` 4.8.4 or `gcc` 4.8.5, depending on the particular version) be present on the system.

Chapter 11. DALI RELEASE 0.12.0 BETA

The DALI 0.12.0 is a beta release. Hence, for all the features, the functionality and performance will likely be limited.

Using DALI 0.12.0 Beta

The DALI 0.12.0 can be used with the 19.08 NVIDIA GPU Cloud (NGC) optimized container for MXNet, PyTorch, and TensorFlow. Also, the 19.08 container will be shipped with DALI 0.12.0.

To upgrade to DALI 0.12.0 beta from an older version of DALI, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

> The internal DALI C++ API used for operators implementation, and the C++ API that enables using DALI as a library from native code, are not yet officially supported. Hence these APIs may change in the next release without advance notice.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Added a new optimized implementation for CropMirrorNormalize for CPU and GPU.
- Added the ability to read video files together with labels.

Fixed Issues

This DALI release includes the following fixes.

- Resumed shipping the prebuilt version of the TensorFlow plugin as a part of **`nvidia-dali-tf-plugin`** (as a result of fixing a compatibility issue). With this fix, the user must first install **`nvidia-dali-tf-plugin`** package. When this package is installed, one of the prebuilt binaries of the TensorFlow plugin is selected. See Binary Installation.

Breaking Changes

- Source code build now requires the compiler be at least C++14 capable.

Deprecated Features

- Deprecated the following as noted below:
    - **`_run`**: use **`schedule_run`** instead
    - **`_share_outputs`**: use **`share_outputs`** instead, and
    - **`_release_outputs`**: use **`release_outputs`** instead.
- Replaced **`HostDecoder`** and **`nvJPEGDecoder`** with generic ImageDecoder. **`ImageDecoder`** is the recommended function for the image decoding, and the replaced APIs **`HostDecoder`** and **`nvJPEGDecoder`** will be removed in the future.
- **`NormalizePermute`** is replaced by the new CropMirrorNormalize.

Known Issues

- The new video reader operator requires NVIDIA VIDEO CODEC SDK support in the platform. Prior to 19.01, the NVIDIA GPU Cloud (NGC) optimized containers lack this functionality in the default configuration. To enable the functionality, run the container with the "video" capability enabled, as shown below:

```
-e "NVIDIA_DRIVER_CAPABILITIES=compute,utility,video"
```

- The video loader operator requires that the key frames occur at a minimum every 10 to 15 frames of the video stream. If the key frames occur at a lesser frequency, then the returned frames may be out of sync.
- DALI TensorFlow plugin may be not compatible with TensorFlow 1.14.0 release. The DALI TensorFlow plugin requires that the **`gcc`** compiler that matches the one used to build TensorFlow (**`gcc`** 4.8.4 or **`gcc`** 4.8.5, depending on the particular version) be present on the system.

Chapter 12. DALI RELEASE 0.11.0 BETA

The DALI 0.11.0 is a beta release. Hence, for all the features, the functionality and performance will likely be limited.

Using DALI 0.11.0 Beta

The DALI 0.11.0 can be used with the 19.07 NVIDIA GPU Cloud (NGC) optimized container for MXNet, PyTorch, and TensorFlow. Also, the 19.07 container will be shipped with DALI 0.11.0.

To upgrade to DALI 0.11.0 beta from an older version of DALI, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

> The internal DALI C++ API used for operators implementation, and the C++ API that enables using DALI as a library from native code, are not yet officially supported. Hence these APIs may change in the next release without advance notice.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Added the ability to provide more than one input to, and return more than one output from, a Python-based operator.
- Extended the bounding box encoder for SSD to return offsets also. See nvidia.dali.ops.BoxEncoder.
- Added the ability to build DALI by mounting the source code to the Docker so consecutive rebuilds are much faster.
- Added experimental support for **`aarch64`** (ARM) platform. Note that this support is added only for the native part—Python is not supported yet.
- Re-implemented the flip operator to increase its performance. See nvidia.dali.ops.Flip.

- Improved the performance of nvJPEG Decoder with new internal API to match the previous implementation. See nvidia.dali.ops.nvJPEGDecoder.

Fixed Issues

This DALI release includes the following fixes.

- Fixed an issue wherein loading a plugin could break other operators that were already registered.
- Fixed an interoperability issue with PyCUDA. Now DALI will not interfere with PyCUDA's CUDA context management.

Breaking API Changes

- CPU operators have moved from per-sample processing (pipeline process sample after sample, all the way through the pipeline) to batch-procession (all samples are processed by the first operator before moving to the next operator). This may result in a small performance degradation for some use cases. However, in the long term it will make some currently unavailable optimizations possible, for example: operations that need to view the whole batch during the processing (like random sample blending inside a batch).
- **CropCastPermute** is removed. CropMirrorNormalize should be used instead (with the default values for normalization).

Deprecated Features

- Removed the prebuilt version of TensorFlow plugin for DALI. Now it is always necessary to install a separate **nvidia-dali-tf-plugin** package. See Binary Installation.

Known Issues

- The new video reader operator requires NVIDIA VIDEO CODEC SDK support in the platform. Prior to 19.01, the NVIDIA GPU Cloud (NGC) optimized containers lack this functionality in the default configuration. To enable the functionality, run the container with the "video" capability enabled, as shown below:

  ```
  -e "NVIDIA_DRIVER_CAPABILITIES=compute,utility,video"
  ```
- The video loader operator requires that the key frames occur at a minimum every 10 to 15 frames of the video stream. If the key frames occur at a lesser frequency, then the returned frames may be out of sync.
- DALI TensorFlow plugin may be not compatible with TensorFlow 1.14.0 release. The DALI TensorFlow plugin requires that the **gcc** compiler that matches the one used to build TensorFlow (**gcc** 4.8.4 or **gcc** 4.8.5, depending on the particular version) be present on the system.

Chapter 13. DALI RELEASE 0.10.0 BETA

The DALI 0.10.0 is a beta release. Hence, for all the features, the functionality and performance will likely be limited.

Using DALI 0.10.0 Beta

The DALI 0.10.0 can be used with the 19.06 NVIDIA GPU Cloud (NGC) optimized container for MXNet, PyTorch, and TensorFlow. Also, the 19.06 container will be shipped with DALI 0.10.0.

To upgrade to DALI 0.10.0 beta from an older version of DALI, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

> The internal DALI C++ API used for operators implementation, and the C++ API that enables using DALI as a library from native code, are not yet officially supported. Hence these APIs may change in the next release without advance notice.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Reduced peak memory consumption. DALI tends to do new allocation before releasing the old memory during buffer resize. As it does not copy the old memory content, the old memory can be freed before allocating the new memory.
- Started publishing DALI nightly builds for CUDA 9 and CUDA 10, and weekly for CUDA 10.
- Added Python function operator. Now the user can create a Python-based operator that accepts one input and produces one output.

Breaking API Changes

- None.

Deprecated Features

- None.

Known Issues

- The new video reader operator requires NVIDIA VIDEO CODEC SDK support in the platform. Prior to 19.01, the NVIDIA GPU Cloud (NGC) optimized containers lack this functionality in the default configuration. To enable the functionality, run the container with the "video" capability enabled, as shown below:

  ```
  -e "NVIDIA_DRIVER_CAPABILITIES=compute,utility,video"
  ```
- The video loader operator requires that the key frames occur at a minimum every 10 to 15 frames of the video stream. If the key frames occur at a lesser frequency, then the returned frames may be out of sync.

Chapter 14. DALI RELEASE 0.9.1 BETA

The DALI 0.9.1 is a beta release. Hence, for all the features, the functionality and performance will likely be limited.

Using DALI 0.9.1 Beta

The DALI 0.9.1 can be used with the 19.05 NVIDIA GPU Cloud (NGC) optimized container for MXNet, PyTorch, and TensorFlow. Also, the 19.05 container will be shipped with DALI 0.9.1.

To upgrade to DALI 0.9.1 beta from an older version of DALI, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

> The internal DALI C++ API used for operators implementation, and the C++ API that enables using DALI as a library from native code, are not yet officially supported. Hence these APIs may change in the next release without advance notice.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Added Optical Flow example.
- Added full support for .pnm (.ppm/.pgm/.pbm) (in case of nvJPEG, fallback to CPU is done).
- Added the ability to do the lazy initialization of DALI. It is now possible to separate touching the data during the actual run, and at the time of all the necessary allocations during the build of the pipeline.
- Reduced the resize operator scratch buffer size by processing the data in mini batches.
- Added the ability to set the DALI CUDA streams priority.
- Fixed the sync issue in the DALI Python iterators.

- Fixed the initialization of CUDA context on the default device during pipeline creation.

Breaking API Changes

- Internal Python pipeline API has changed. If any function **`_*`** was used, then that function should be updated to reflect the new semantic.

Deprecated Features

- None.

Known Issues

- The new video reader operator requires NVIDIA VIDEO CODEC SDK support in the platform. Prior to 19.01, the NVIDIA GPU Cloud (NGC) optimized containers lack this functionality in the default configuration. To enable the functionality, run the container with the "video" capability enabled, as shown below:

  ```
  -e "NVIDIA_DRIVER_CAPABILITIES=compute,utility,video"
  ```
- The video loader operator requires that the key frames occur at a minimum every 10 to 15 frames of the video stream. If the key frames occur at a lesser frequency, then the returned frames may be out of sync.

Chapter 15. DALI RELEASE 0.8.1 BETA

The DALI 0.8.1 is a beta release. Hence, for all the features, the functionality and performance will likely be limited.

Using DALI 0.8.1 Beta

The DALI 0.8.1 can be used with the 19.04 NVIDIA GPU Cloud (NGC) optimized container for MXNet, PyTorch, and TensorFlow. Also, the 19.05 container will be shipped with DALI 0.9.

To upgrade to DALI 0.8.1 beta from an older version of DALI, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

> The internal DALI C++ API used for operators implementation, and the C++ API that enables using DALI as a library from native code, are not yet officially supported. Hence these APIs may change in the next release without advance notice.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Fixed nvJPEGDecoder cache when using the new nvJPEG decoupled API.

Breaking API Changes

- None.

Deprecated Features

- None.

Known Issues

- The new video reader operator requires NVIDIA VIDEO CODEC SDK support in the platform. Prior to 19.01, the NVIDIA GPU Cloud (NGC) optimized containers lack this functionality in the default configuration. To enable the functionality, run the container with the "video" capability enabled, as below:

  ```
  -e "NVIDIA_DRIVER_CAPABILITIES=compute,utility,video"
  ```
- The video loader operator requires that the key frames occur at a minimum every 10 to 15 frames of the video stream. If the key frames occur at a lesser frequency, then the returned frames may be out of sync.

Chapter 16.
DALI RELEASE 0.8 BETA

The DALI 0.8 is a beta release. Hence, for all the features, the functionality and performance will likely be limited.

Using DALI 0.8 Beta

The DALI 0.8 can be used with the 19.03 NVIDIA GPU Cloud (NGC) optimized container for MXNet, PyTorch, and TensorFlow. Also, the 19.04 container will be shipped with DALI 0.8.

To upgrade to DALI 0.8 beta from an older version of DALI, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

> The internal DALI C++ API used for operators implementation, and the C++ API that enables using DALI as a library from native code, are not yet officially supported. Hence these APIs may change in the next release without advance notice.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Provides the CUDA 10-based DALI prebuilt binary.
- Added the customizable and decoupled executor stages. The CPU and GPU prefetch queues can now have different depths.
- The nvJPEG decoder now supports ROI-based decoding.
- The nvJPEG decoder now has split variants. The CPU and GPU part can be concurrently executed and independently prefetched.
- Improved the resize operator to support bilinear and triangular resampling. This improves the quality during aggressive down-scaling.
- Added the ability to return sparse tensor on CPU for TF DALI op.

- Fixed bugs in the detection pipeline operators. These bugs might have lead to inferior training accuracy.

Breaking API Changes

- None.

Known Issues

- The new video reader operator requires NVIDIA VIDEO CODEC SDK support in the platform. Prior to 19.01, the NVIDIA GPU Cloud (NGC) optimized containers lack this functionality in the default configuration. To enable the functionality, run the container with the "video" capability enabled, as below:

  ```
  -e "NVIDIA_DRIVER_CAPABILITIES=compute,utility,video"
  ```
- The video loader operator requires that the key frames occur at a minimum every 10 to 15 frames of the video stream. If the key frames occur at a lesser frequency, then the returned frames may be out of sync.

Chapter 17.
DALI RELEASE 0.7 BETA

The DALI 0.7 is a beta release. Hence, for all the features, the functionality and performance will likely be limited.

Using DALI 0.7 Beta

The DALI 0.7 can be used with the 19.03 NVIDIA GPU Cloud (NGC) optimized container for MXNet, PyTorch, and TensorFlow.

To upgrade to DALI 0.7 beta from an older version of DALI, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

> The internal DALI C++ API used for operators implementation, and the C++ API that enables using DALI as a library from native code, are not yet officially supported. Hence these APIs may change in the next release without advance notice.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Added new operators:
  - HostDecoderRandomCrop
  - Element Extract for sequences
- Added the API stub for incoming Optical Flow video operator.
- Introduced Python 3.7 based **`whl`** build.
- Added Docker-based "one click" build script.
- Added the ability to fine-tune the internal DALI buffers using per-operator presize hints to stage the output queues.
- Added **`step`**, **`stride`** and **`shuffling`** in **`SequenceReader`**.
- Added a new, internal test API.

- Added the ability to custom handle CPU affinity using the DALI_AFFINITY_MASK environment variable.

Breaking API Changes

- None.

Deprecated Features

- Dropped the Python 3.4 based **`whl`** build.

Known Issues

- The new video reader operator requires NVIDIA VIDEO CODEC SDK support in the platform. The NVIDIA GPU Cloud (NGC) optimized containers lack this functionality in the default configuration prior to the 19.01 version. To enable the functionality, run the container with the "video" capability enabled, as below:

  ```
  -e "NVIDIA_DRIVER_CAPABILITIES=compute,utility,video"
  ```
- There is no clear distinction in the documentation between the operators supporting video sequences and the operators supporting images.

Chapter 18. DALI RELEASE 0.6.1 BETA

The DALI 0.6.1 is a beta release. Hence, for all the features, the functionality and performance will likely be limited.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Added the DALI TensorFlow plugin to improve out-of-the-box forward compatibility.
- Added a new package **`nvidia-dali-tf-plugin`** that allows user to compile DALI TensorFlow plugin using the locally installed version.
- Added the below examples:
  - VideoReader
  - PyTorch and MXNet examples with various readers
- Added a new operator variant **`gpu box encoder`**.
- Added a new JSON parser—now COCO reader start up time is significantly faster (~10x).
- Added the ability to automatically reset MXNet and PyTorch Python iterators.
- Enhanced the Video reader to link dynamically with the CUDA Driver API and **`nvcuvid`**.
- Added the ability to deliver exactly 1 epoch from **`DALIGenericIterator`** (PyTorch and MXNet).

Breaking API Changes

- PyTorch iterator returns the exact number of samples per epoch, so the final batch could be smaller if epoch size is not divisible by the batch size. To keep the old behaviour when data is wrapped up, use **`stop_at_epoch`** argument.

Using DALI 0.6.1 Beta

The 19.01 NVIDIA GPU Cloud (NGC) optimized container for MXNet, PyTorch, and TensorFlow includes an older version of DALI.

To upgrade to DALI 0.6.1 beta, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

Known Issues

- The New Video reader operator requires NVIDIA VIDEO CODEC SDK support in the platform. The NVIDIA GPU Cloud (NGC) optimized containers lack this functionality in the default configuration prior to the 19.01 version, resulting in problem with loading **`libnvcuvid.so`**. To enable it, run the container with the 'video' capability enabled, as below:

  ```
  -e "NVIDIA_DRIVER_CAPABILITIES=compute,utility,video"
  ```

Chapter 19. DALI RELEASE 0.6 BETA

The DALI 0.6 is a beta release. Hence, for all the features, the functionality and performance will likely be limited.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Added the CPU variant of DALI TensorFlow operator. Now the DALI pipeline can be scheduled on the CPU also.
- Added the TensorFlow operator API to allow defining an arbitrary number of outputs and their types.
- Added full DALI detection pipeline sample for SSD.
- Added clear mapping between operators and the supported device types in the API documentation.
- Created initial support for the video-based pipeline - NVVL functionality is available as VideoReader.
- Added the below new operators:
  - **`BoxEncoder`**
  - **`VideoReader`**
  - **`SequenceCrop`**
  - **`SequenceReader`**
- Added new variants of the below existing operators:
  - For the GPU:
    - **`RandomBBoxCrop`**
    - **`BbFlip`**
    - **`Slice`**
  - For the CPU:
    - **`CropMirrorNormalize`**, and

  - **`BBoxPaste`**

Breaking API Changes

- DALI TensorFlow operator has a new API - check the examples for the reference.
- PyTorch and MXNet python iterators API have changed - check the examples for the reference.

Using DALI 0.6 Beta

The 19.01 NVIDIA GPU Cloud (NGC) optimized container for MXNet, PyTorch, and TensorFlow includes an older version of DALI. To upgrade to DALI 0.6 beta, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

Known Issues

- The New Video reader operator requires NVIDIA VIDEO CODEC SDK support in the platform. The NVIDIA GPU Cloud (NGC) optimized containers lack this functionality in the default configuration prior to the 19.01 version, resulting in problem with loading **`libnvcuvid.so`**. To enable it, run the container with the 'video' capability enabled, as below:

  ```
  -e "NVIDIA_DRIVER_CAPABILITIES=compute,utility,video"
  ```

Chapter 20.
DALI RELEASE 0.5 BETA

This DALI 0.5 release is a beta release.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Fixed rare hangs during DALI pipeline destruction.
- Added ResNet50 example with TensorFlow.
- Added more verbose error messages to HostDecoder and nvJpegDecoder.
- Added pipelined execution completion callback setter.
- Added support for Tiff images.

Breaking API Changes

- Random seed type changed from INT to INT64, therefore, serialized pipelines from versions prior to 0.5 are not compatible with the current DALI version.

Using DALI 0.5 Beta

The 18.11 NVIDIA GPU Cloud (NGC) optimized container for MXNet, PyTorch, and TensorFlow, includes an older version of DALI. To upgrade to DALI 0.5 beta, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

Known Issues

- This is a beta release. Hence, for all the features, the functionality and performance will likely be limited.

Chapter 21. DALI RELEASE 0.4.1 BETA

This DALI 0.4.1 release is a beta release.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Added TensorFlow 1.11 and 1.12 compatibility.
- Adjusted PyTorch example to use new nvJPEG API.
- Reduced DALI memory footprint.
- Fixed lack of ability to customize pipeline processing length.

Using DALI 0.4.1 Beta

The 18.10 NVIDIA GPU Cloud (NGC) optimized container for MXNet, PyTorch, and TensorFlow, includes an older version of DALI. To upgrade to DALI 0.4.1 beta, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

Known Issues

- The DALI integrated ResNet-50 samples in the 18.10 NGC TensorFlow and PyTorch containers may result in lower than expected performance results. We are working to address the issue in the next release.
- This is a beta release. All features are expected to be available, however, some aspects of functionality and performance will likely be limited compared to a non-beta release.

Chapter 22. DALI RELEASE 0.4 BETA

This DALI 0.4 release is a beta release.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Added basic operators for detection (SSD), such as COCO dataset reader (**`COCOReader`**), random crop operator with bounding boxes (**`SSDRandomCrop`**), and flip operator for bounding boxes (**`BbFlip`**).
- Added CPU versions of **`Crop`**/**`CropCastPermute`** operators.
- Added a random **`Paste`** operator.
- Upgraded OpenCV to 3.4.3.
- Upgraded MXNet to 1.3.0.
- Fixed parsing of JPEG headers by the Host Decoder known issue. The Host Decoder now handles all images from the ImageNet dataset.
- Added fine grained control over output buffers in the pipeline
- Updated to nvJPEG 0.2.0

Breaking API Changes

The pipeline constructor signature has changed; an additional argument (**`prefetch_queue_depth)`** was added to allow defining depth of the prefetch queue at runtime.

Using DALI 0.4 Beta

The 18.09 NVIDIA GPU Cloud (NGC) optimized container for MXNet, PyTorch, and TensorFlow, includes an older version of DALI. To upgrade to DALI 0.4 beta, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

Deprecated Features

- DALI 0.4 is not compatible with TensorFlow 1.11. This will be addressed in the next release.

Known Issues

- The DALI integrated ResNet-50 samples in the 18.10 NGC TensorFlow and PyTorch containers may result in lower than expected performance results. We are working to address the issue in the next release.
- This is a beta release. All features are expected to be available, however, some aspects of functionality and performance will likely be limited compared to a non-beta release.

Chapter 23. DALI RELEASE 0.3 BETA

This DALI 0.3 release is a beta release.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Updated PyTorch ResNet-50 example to obtain expected accuracy (Top1 76%).
- Introduced CPU variant of resize operator and added stand-alone flip operator.
- Added support for DALI to work with float16 data passed from Python.
- Added fallback to CPU for BMP images.
- Fixed training accuracy with TFRecord reader

Using DALI 0.3 Beta

The 18.08 NVIDIA GPU Cloud (NGC) optimized container for MXNet, PyTorch, and TensorFlow, includes an older version of DALI. To upgrade to DALI 0.2 beta, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

Known Issues

- HosDecoder cannot handle all jpeg files from ImageNet dataset. This will be fixed in the next release.
- On file systems where the directory entries are not stored in any order, the File Reader may assign different labels to the training and validation folders with the same name. This will be fixed in the next release.

- The DALI integrated ResNet-50 samples in the 18.09 NGC TensorFlow and PyTorch containers may result in lower than expected performance results. We are working to address the issue in the next release.
- This is a beta release. All features are expected to be available, however, some aspects of functionality and performance will likely be limited compared to a non-beta release.

Chapter 24. DALI RELEASE 0.2 BETA

This DALI 0.2 release is a beta release.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Added Sphinx based documentation that is in sync with the code on GitHub. For more information, see DALI Master Branch User Guide.
- Build system has been refined, common errors have meaningful messages, improved localization and version detection of key dependency packages, such as, nvJPEG, libturbo-jpeg, and LMDB.
- Added **`Unfused Crop`** and **`CropCastPermute`** operators.
- Added improvements for TensorFlow plugin (polymorphism and shape argument for the output).
- Expanded examples of TensorFlow working with different readers, such as, MXNetReader, FileReader, and TFRecordReader.
- Updated nvJPEG to 0.1.4
- Added fallback to host decoder when image is not JPEG but PNG instead. For example, **`n02105855_2933.JPEG`** from ImageNet.

Breaking API Changes

- The API for the **`Resize`** operator changed to match other similar operators like **`ResizeCropMirror.`**
- The API for the TensorFlow plugin changed to allow specifying the whole shape of the tensor instead of **`N`**, **`H`**, and **`W`** separately; which enables handling both **`NCHW`** and **`NHWC`** outputs.

- The type of labels produced by the TensorFlow plugin have changed. In DALI version 0.1.2, it was always **`tf.float32`**. In this release, a new optional parameter called **`label_type`** is introduced to the TensorFlow plugin to control the type of label. The default value for **`label_type`** is **`tf.int64`** to better align with the label type in TFRecord.

Using DALI 0.2 Beta

The 18.08 NVIDIA GPU Cloud (NGC) optimized container for MXNet, PyTorch, and TensorFlow, includes an older version of DALI. To upgrade to DALI 0.2 beta, follow the installation instructions in the DALI Quick Start Guide.

Refer to the DALI Developer Guide for usage details.

Known Issues

- This is a beta release, therefore, not all functionality is fully supported and working. This beta release is meant for testing and research.
- The DALI integrated ResNet-50 samples in the 18.08 NGC TensorFlow and PyTorch containers have lower than expected accuracy and performance results. We are working to address the issue in the next release.
- This is a beta release. All features are expected to be available, however, some aspects of functionality and performance will likely be limited compared to a non-beta release.

Chapter 25. DALI RELEASE 0.1.2 BETA

This DALI 0.1.2 release is a beta release.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

- Fixed compatibility with TensorFlow 1.9.
- Updated to nvJPEG v0.1.2 to fix batched decoding when a batch contains both gray-scale and color images.
- Added Tensorflow 1.7 support.
- Improved overlap support when using DALI with multi-GPU in MXNet and PyTorch.

Using DALI 0.1.2 Beta

The 18.07 optimized container for MXNet, PyTorch, and TensorFlow, includes an older version of DALI. To upgrade to DALI 0.2 beta, follow the installation instructions in the DALI Quick Start Guide.

Refer to the the DALI Developer Guide for usage details.

Known Issues

- This is a beta release, therefore, not all functionality is fully supported and working. This beta release is meant for testing and research.
- This is a beta release. All features are expected to be available, however, some aspects of functionality and performance will likely be limited compared to a non-beta release.

Chapter 26.
DALI RELEASE 0.1.1 BETA

This is the first release of DALI. This DALI 0.1.1 release is a beta release.

Key Features and Enhancements

This DALI release includes the following key features and enhancements.

Performance

On dense GPU systems, deep learning applications can be significantly bottlenecked on the CPU, limiting the overall performance and scalability of training and inference tasks. DALI enables offloading key deep learning augmentation steps on to GPUs, alleviating CPU bottleneck on the deep learning preprocessing pipelines. This results in out-of-box performance of overall training workflow and efficient utilization of multi-GPU resources on the system.

Drop-in Integration

DALI comes with built-in plugins for key frameworks such as MXNet, TensorFlow, and PyTorch. This enables automatic integration with frameworks so that researchers and developers can get up and running with DALI easily and quickly.

Flexibility

DALI supports multiple input data formats that are commonly used in computer vision deep learning applications, for example, JPEG images, raw formats, Lightning Memory-Mapped Database (LMDB), RecordIO and TFRecord. The flexibility of input data formats allows portability of training workflows across different frameworks and models, and helps to avoid intermediate data conversion steps. DALI enables better code reuse and maintainability with optimized building blocks and support for different data formats.

Using DALI 0.1.1 Beta

Ensure you are familiar with the following notes when using this release.

- To install DALI, see the DALI Quick Start Guide.

> If you are using the 18.07 NGC optimized container for MXNet, PyTorch, or TensorFlow, you do not need to reinstall DALI. DALI now comes included in the container. Instead, start with the Getting Started Tutorial.

- To interact with the code via GitHub, see the Getting Started Tutorial.
- To learn how to define, build, and run a DALI pipeline, see the DALI Developer Guide.

Known Issues

- This is a beta release, therefore, not all functionality is fully supported and working. This beta release is meant for testing and research.
- This is a beta release. All features are expected to be available, however, some aspects of functionality and performance will likely be limited compared to a non-beta release.

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 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.

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.

VESA DisplayPort

HDMI

OpenCL

Trademarks

Copyright