



# Kaldi

## Release Notes

# Table of Contents

Chapter 1. Kaldi Overview.....	1
Chapter 2. Pulling A Container.....	2
Chapter 3. Running Kaldi.....	3
Chapter 4. Kaldi Release 24.02.....	5
Chapter 5. Kaldi Release 24.01.....	6
Chapter 6. Kaldi Release 23.12.....	7
Chapter 7. Kaldi Release 23.11.....	8
Chapter 8. Kaldi Release 23.10.....	13
Chapter 9. Kaldi Release 23.09.....	17
Chapter 10. Kaldi Release 23.08.....	21
Chapter 11. Kaldi Release 23.07.....	25
Chapter 12. Kaldi Release 23.06.....	29
Chapter 13. Kaldi Release 23.05.....	33
Chapter 14. Kaldi Release 23.04.....	37
Chapter 15. Kaldi Release 23.03.....	41
Chapter 16. Kaldi Release 23.02.....	45
Chapter 17. Kaldi Release 23.01.....	49
Chapter 18. Kaldi Release 22.12.....	53
Chapter 19. Kaldi Release 22.11.....	57
Chapter 20. Kaldi Release 22.10.....	61
Chapter 21. Kaldi Release 22.09.....	65
Chapter 22. Kaldi Release 22.08.....	69
Chapter 23. Kaldi Release 22.07.....	73
Chapter 24. Kaldi Release 22.06.....	77
Chapter 25. Kaldi Release 22.05.....	81
Chapter 26. Kaldi Release 22.04.....	85
Chapter 27. Kaldi Release 22.03.....	89
Chapter 28. Kaldi Release 22.02.....	93
Chapter 29. Kaldi Release 22.01.....	97

Chapter 30. Kaldi Release 21.12.....	101
Chapter 31. Kaldi Release 21.11.....	105
Chapter 32. Kaldi Release 21.10.....	109
Chapter 33. Kaldi Release 21.09.....	112
Chapter 34. Kaldi Release 21.08.....	115
Chapter 35. Kaldi Release 21.07.....	118
Chapter 36. Kaldi Release 21.06.....	121
Chapter 37. Kaldi Release 21.05.....	124
Chapter 38. Kaldi Release 21.04.....	127
Chapter 39. Kaldi Release 21.03.....	130
Chapter 40. Kaldi Release 21.02.....	133
Chapter 41. Kaldi Release 21.01.....	136
Chapter 42. Kaldi Release 20.12.....	137
Chapter 43. Kaldi Release 20.11.....	140
Chapter 44. Kaldi Release 20.10.....	143
Chapter 45. Kaldi Release 20.09.....	146
Chapter 46. Kaldi Release 20.08.....	149
Chapter 47. Kaldi Release 20.07.....	152
Chapter 48. Kaldi Release 20.06.....	155
Chapter 49. Kaldi Release 20.03.....	158
Chapter 50. Kaldi Release 20.02.....	160
Chapter 51. Kaldi Release 20.01.....	162
Chapter 52. Kaldi Release 19.12.....	164
Chapter 53. Kaldi Release 19.11.....	166
Chapter 54. Kaldi Release 19.10.....	168
Chapter 55. Kaldi Release 19.09.....	170
Chapter 56. Kaldi Release 19.08.....	172
Chapter 57. Kaldi Release 19.07.....	174
Chapter 58. Kaldi Release 19.06.....	176
Chapter 59. Kaldi Release 19.05.....	178
Chapter 60. Kaldi Release 19.04.....	180

Chapter 61. Kaldi Release 19.03..... 182

---

# Chapter 1. Kaldi Overview

The NVIDIA® Deep Learning SDK accelerates widely-used deep learning frameworks such as [Kaldi](#).

The Kaldi Speech Recognition Toolkit project began in 2009 at [Johns Hopkins University](#) with the intent of developing techniques to reduce the cost and time required to build speech recognition systems. While originally focused on ASR support for new languages and domains, the Kaldi project has steadily grown in size and capabilities and enables hundreds of researchers to participate in advancing the field. Now the de-facto speech recognition toolkit in the community, Kaldi helps enable speech services that are used by millions of people every day.

In the container, see the `/workspace/README.md` file for information about customizing your Kaldi image. For more information about Kaldi, including tutorials, documentation, and examples, see the [Kaldi Speech Recognition Toolkit](#).

This document describes the key features, software enhancements and improvements, known issues, and how to run this container.

---

# Chapter 2. Pulling A Container

## About this task

**Before** you can pull a container from the NGC container registry:

- ▶ Install Docker.
  - ▶ For NVIDIA DGX™ users, see [Preparing to use NVIDIA Containers Getting Started Guide](#).
  - ▶ For non-DGX users, see NVIDIA® GPU Cloud™ (NGC) container registry [installation documentation](#) based on your platform.
- ▶ Ensure that you have access and can log in to the NGC container registry.

Refer to [NGC Getting Started Guide](#) for more information.

The deep learning frameworks, the NGC Docker containers, and the deep learning framework containers are stored in the `nvcr.io/nvidia` repository.

---

# Chapter 3. Running Kaldi

## Before you begin

Before you can run an NGC deep learning framework container, your Docker environment must support NVIDIA GPUs. To run a container, issue the appropriate command as explained in [Running A Container](#) and specify the registry, repository, and tags.

## About this task

On a system with GPU support for NGC containers, when you run a container, the following occurs :

- ▶ The Docker engine loads the image into a container that runs the software.
- ▶ You define the container's runtime resources by including the additional flags and settings that are used with the command.

These flags and settings are described in [Running A Container](#).

- ▶ The GPUs are explicitly defined for the Docker<sup>®</sup> container, which defaults to all GPUs, but can be specified by using the `NVIDIA_VISIBLE_DEVICES` environment variable.

For more information, refer to the [nvidia-docker documentation](#).



**Note:** Starting in Docker 19.03, complete the steps below.

The method implemented in your system depends on the DGX OS version that you installed (for DGX systems), the NGC Cloud Image that was provided by a Cloud Service Provider, or the software that you installed to prepare to run NGC containers on TITAN PCs, Quadro PCs, or NVIDIA Virtual GPUs (vGPUs).

## Procedure

1. Issue the command for the applicable release of the container that you want.

The following command assumes that you want to pull the latest container.

```
docker pull nvcr.io/nvidia/kaldi:23.12-py3
```

2. Open a command prompt and paste the `pull` command.

Ensure that the pull process successfully completes before you proceed to step 3.

3. Run the container image.

- ▶ If you have **Docker 19.03 or later**, a typical command to launch the container is:

```
docker run --gpus all -it --rm -v local_dir:container_dir nvcr.io/nvidia/  
kaldi:<xx.xx>-py3
```

- ▶ If you have **Docker 19.02 or earlier**, a typical command to launch the container is:

```
nvidia-docker run -it --rm -v local_dir:container_dir nvcr.io/nvidia/kaldi:<xx.xx>-  
py3
```

---

# Chapter 4. Kaldi Release 24.02

There is no Kaldi container in DLFW release 24.02.

---

# Chapter 5. Kaldi Release 24.01

There is no Kaldi container in DLFW release 24.01.

---

## Chapter 6. Kaldi Release 23.12

There is no Kaldi container in DLFW release 23.12.

---

# Chapter 7. Kaldi Release 23.11

The NVIDIA container image for Kaldi, release 23.11 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 22.04](#) including [Python 3.10](#)
- ▶ [NVIDIA CUDA® 12.3.0](#)
- ▶ [NVIDIA cuBLAS 12.3.2.1](#)
- ▶ [NVIDIA cuDNN 8.9.6](#)
- ▶ [NVIDIA NCCL 2.19.3](#)
- ▶ [rdma-core 39.0](#)
- ▶ [Nsight Compute 2023.3.0.12](#)
- ▶ [Nsight Systems 2023.3.1.92](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.16 with UCX 1.15.0
- ▶ [NVIDIA TensorRT™ 8.6.1.6](#)
- ▶ [NVIDIA DALI® 1.31.0](#)

## Driver Requirements

Release 23.11 is based on [CUDA 12.3.0](#), which requires [NVIDIA Driver](#) release 545 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), 510.47 (or later R510), or 525.85 (or later R525), or 535.86 (or later R535) or 545.23 (or later R545).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, R460, and R520 drivers, which are not forward-compatible with CUDA 12.3. For a complete list of supported drivers, see the [CUDA](#)

[Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 23.11 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal™, NVIDIA Volta™, NVIDIA Turing™, NVIDIA Ampere architecture, NVIDIA Hopper™, and NVIDIA Ada Lovelace architecture families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 23.11 is based on [Kaldi 5.5](#) up to [Kaldi 26b9f648](#).
- ▶ Ubuntu 22.04 with October 2023 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
23.11	22.04	<a href="#">NVIDIA CUDA 12.3.0</a>	5.5 up to commit 26b9f648	<a href="#">TensorRT 8.6.1.6</a>
23.10		<a href="#">NVIDIA CUDA 12.2.2</a>		
<a href="#">23.09</a>		<a href="#">NVIDIA CUDA 12.2.1</a>		
<a href="#">23.08</a>				
<a href="#">23.07</a>		<a href="#">NVIDIA CUDA 12.1.1</a>		

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT		
<a href="#">23.06</a>	20.04					
<a href="#">23.05</a>				<a href="#">TensorRT 8.6.1.2</a>		
<a href="#">23.04</a>				<a href="#">NVIDIA CUDA 12.1.0</a>	<a href="#">TensorRT 8.6.1</a>	
<a href="#">23.03</a>				<a href="#">NVIDIA CUDA 12.0.1</a>	<a href="#">TensorRT 8.5.3</a>	
<a href="#">23.02</a>						
<a href="#">23.01</a>				<a href="#">NVIDIA CUDA 11.8.0</a>	<a href="#">TensorRT 8.5.2.2</a>	
<a href="#">22.12</a>					<a href="#">TensorRT 8.5.1</a>	
<a href="#">22.11</a>						
<a href="#">22.10</a>					<a href="#">TensorRT 8.5.0.12</a>	
<a href="#">22.09</a>						
<a href="#">22.08</a>				<a href="#">NVIDIA CUDA 11.7.1</a>	<a href="#">TensorRT 8.4.2.4</a>	
<a href="#">22.07</a>				<a href="#">NVIDIA CUDA 11.7 Update 1 Preview</a>	<a href="#">TensorRT 8.4.1</a>	
<a href="#">22.06</a>				<a href="#">NVIDIA CUDA 11.7.0</a>	<a href="#">TensorRT 8.2.5</a>	
<a href="#">22.05</a>						
<a href="#">22.04</a>				<a href="#">NVIDIA CUDA 11.6.2</a>	<a href="#">TensorRT 8.2.4.2</a>	
<a href="#">22.03</a>				<a href="#">NVIDIA CUDA 11.6.1</a>	<a href="#">TensorRT 8.2.3</a>	
<a href="#">22.02</a>				<a href="#">NVIDIA CUDA 11.6.0</a>	<a href="#">TensorRT 8.2.3</a>	
<a href="#">22.01</a>					<a href="#">TensorRT 8.2.2</a>	
<a href="#">21.12</a>				<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5 up to commit c6772ba4</a>	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>					<a href="#">5.5 up to commit c6772ba4</a>	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>	<a href="#">NVIDIA CUDA 11.4.2 with cuBLAS 11.6.5.2</a>	<a href="#">5.5 up to commit 2897564e</a>				
<a href="#">21.09</a>	<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5 up to commit 40573bb</a>	<a href="#">TensorRT 8.0.3</a>			
<a href="#">21.08</a>	<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5 up to commit 851a0a79</a>	<a href="#">TensorRT 8.0.1.6</a>			

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>				
<a href="#">19.09</a>			<a href="#">NVIDIA CUDA 10.1.243</a>	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

None.

---

# Chapter 8. Kaldi Release 23.10

The NVIDIA container image for Kaldi, release 23.10 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 22.04](#) including [Python 3.10](#)
- ▶ [NVIDIA CUDA® 12.2.2](#)
- ▶ [NVIDIA cuBLAS 12.2.5.6](#)
- ▶ [NVIDIA cuDNN 8.9.5](#)
- ▶ [NVIDIA NCCL 2.19.3](#)
- ▶ [rdma-core 39.0](#)
- ▶ [Nsight Compute 2023.2.1.3](#)
- ▶ [Nsight Systems 2023.3.1.92](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.16 with UCX 1.15.0
- ▶ [NVIDIA TensorRT™ 8.6.1.6](#)
- ▶ [NVIDIA DALI® 1.30.0](#)

## Driver Requirements

Release 23.10 is based on [CUDA 12.2.2](#), which requires [NVIDIA Driver](#) release 535 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), 510.47 (or later R510), or 525.85 (or later R525), or 535.86 (or later R535).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, R460, and R520 drivers, which are not forward-compatible with CUDA 12.2. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 23.10 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal™, NVIDIA Volta™, NVIDIA Turing™, NVIDIA Ampere architecture, NVIDIA Hopper™, and NVIDIA Ada Lovelace architecture families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 23.10 is based on [Kaldi 5.5](#) up to [Kaldi 26b9f648](#).
- ▶ Ubuntu 22.04 with September 2023 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
23.10	22.04	<a href="#">NVIDIA CUDA 12.2.2</a>	5.5 up to commit 26b9f648	<a href="#">TensorRT 8.6.1.6</a>
<a href="#">23.09</a>		<a href="#">NVIDIA CUDA 12.2.1</a>		
<a href="#">23.08</a>				
<a href="#">23.07</a>		<a href="#">NVIDIA CUDA 12.1.1</a>		
<a href="#">23.06</a>				
<a href="#">23.05</a>				<a href="#">TensorRT 8.6.1.2</a>
<a href="#">23.04</a>	20.04	<a href="#">NVIDIA CUDA 12.1.0</a>		<a href="#">TensorRT 8.6.1</a>
<a href="#">23.03</a>				<a href="#">TensorRT 8.5.3</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">23.02</a>		<a href="#">NVIDIA CUDA 12.0.1</a>		
<a href="#">23.01</a>				<a href="#">TensorRT 8.5.2.2</a>
<a href="#">22.12</a>		<a href="#">NVIDIA CUDA 11.8.0</a>		<a href="#">TensorRT 8.5.1</a>
<a href="#">22.11</a>				
<a href="#">22.10</a>				<a href="#">TensorRT 8.5.0.12</a>
<a href="#">22.09</a>				
<a href="#">22.08</a>		<a href="#">NVIDIA CUDA 11.7.1</a>		<a href="#">TensorRT 8.4.2.4</a>
<a href="#">22.07</a>		<a href="#">NVIDIA CUDA 11.7 Update 1 Preview</a>		<a href="#">TensorRT 8.4.1</a>
<a href="#">22.06</a>				<a href="#">TensorRT 8.2.5</a>
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.7.0</a>		
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>		<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>				<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2 with cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT	
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae		
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c		
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca		<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf		<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a		<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>	
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837		
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>	
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36		
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc		
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>	
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>	
<a href="#">20.02</a>			<a href="#">5.5</a>		
<a href="#">20.01</a>					
<a href="#">19.12</a>					<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>					
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>			
<a href="#">19.09</a>					
<a href="#">19.08</a>	<a href="#">TensorRT 5.1.5</a>				

## Known Issues

None.

---

# Chapter 9. Kaldi Release 23.09

The NVIDIA container image for Kaldi, release 23.09 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 22.04](#) including [Python 3.10](#)
- ▶ [NVIDIA CUDA® 12.2.1](#)
- ▶ [NVIDIA cuBLAS 12.2.5.6](#)
- ▶ [NVIDIA cuDNN 8.9.5](#)
- ▶ [NVIDIA NCCL 2.18.5](#)
- ▶ [rdma-core 39.0](#)
- ▶ [Nsight Compute 2023.2.1.3](#)
- ▶ [Nsight Systems 2023.3.1.92](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.16 with UCX 1.15.0
- ▶ [NVIDIA TensorRT™ 8.6.1.6](#)
- ▶ [NVIDIA DALI® 1.29.0](#)

## Driver Requirements

Release 23.09 is based on [CUDA 12.2.1](#), which requires [NVIDIA Driver](#) release 535 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), 510.47 (or later R510), or 525.85 (or later R525), or 535.86 (or later R535).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, R460, and R520 drivers, which are not forward-compatible with CUDA 12.2. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 23.09 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal™, NVIDIA Volta™, NVIDIA Turing™, NVIDIA Ampere architecture, NVIDIA Hopper™, and NVIDIA Ada Lovelace architecture families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 23.09 is based on [Kaldi 5.5](#) up to [Kaldi 26b9f648](#).
- ▶ Ubuntu 22.04 with August 2023 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
23.09	22.04	<a href="#">NVIDIA CUDA 12.2.1</a>	5.5 up to commit 26b9f648	<a href="#">TensorRT 8.6.1.6</a>
<a href="#">23.08</a>				
<a href="#">23.07</a>		<a href="#">NVIDIA CUDA 12.1.1</a>		<a href="#">TensorRT 8.6.1.2</a>
<a href="#">23.06</a>				
<a href="#">23.05</a>				
<a href="#">23.04</a>	20.04	<a href="#">NVIDIA CUDA 12.1.0</a>	<a href="#">TensorRT 8.6.1</a>	
<a href="#">23.03</a>				
<a href="#">23.02</a>		<a href="#">NVIDIA CUDA 12.0.1</a>		<a href="#">TensorRT 8.5.3</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">23.01</a>				<a href="#">TensorRT 8.5.2.2</a>
<a href="#">22.12</a>		<a href="#">NVIDIA CUDA 11.8.0</a>		<a href="#">TensorRT 8.5.1</a>
<a href="#">22.11</a>				
<a href="#">22.10</a>				<a href="#">TensorRT 8.5.0.12</a>
<a href="#">22.09</a>				
<a href="#">22.08</a>		<a href="#">NVIDIA CUDA 11.7.1</a>		<a href="#">TensorRT 8.4.2.4</a>
<a href="#">22.07</a>		<a href="#">NVIDIA CUDA 11.7 Update 1 Preview</a>		<a href="#">TensorRT 8.4.1</a>
<a href="#">22.06</a>				<a href="#">TensorRT 8.2.5</a>
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.7.0</a>		
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>		<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>				<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2</a> with <a href="#">cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c		
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>	
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>	
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>	
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>	
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837		
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>	
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36		
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc		
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>	
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>	
<a href="#">20.02</a>			<a href="#">5.5</a>		
<a href="#">20.01</a>					
<a href="#">19.12</a>					<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>					
<a href="#">19.10</a>			<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>					
<a href="#">19.08</a>					<a href="#">TensorRT 5.1.5</a>

## Known Issues

None.

---

# Chapter 10. Kaldi Release 23.08

The NVIDIA container image for Kaldi, release 23.08 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 22.04](#) including [Python 3.10](#)
- ▶ [NVIDIA CUDA<sup>®</sup> 12.2.1](#)
- ▶ [NVIDIA cuBLAS 12.2.5](#)
- ▶ [NVIDIA cuDNN 8.9.4](#)
- ▶ [NVIDIA NCCL 2.18.3](#)
- ▶ [rdma-core 39.0](#)
- ▶ [Nsight Compute 2023.2.1.3](#)
- ▶ [Nsight Systems 2023.2.3.1001](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.15 with UCX 1.15.0
- ▶ [NVIDIA TensorRT<sup>™</sup> 8.6.1.6](#)
- ▶ [NVIDIA DALI<sup>®</sup> 1.28.0](#)

## Driver Requirements

Release 23.08 is based on [CUDA 12.2.1](#), which requires [NVIDIA Driver](#) release 535 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), 510.47 (or later R510), or 525.85 (or later R525), or 535.86 (or later R535).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, R460, and R520 drivers, which are not forward-compatible with CUDA 12.2. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 23.08 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal™, NVIDIA Volta™, NVIDIA Turing™, NVIDIA Ampere architecture, NVIDIA Hopper™, and NVIDIA Ada Lovelace architecture families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 23.08 is based on [Kaldi 5.5](#) up to [Kaldi 26b9f648](#).
- ▶ Ubuntu 22.04 with July 2023 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
23.08	22.04	<a href="#">NVIDIA CUDA 12.2.1</a>	5.5 up to commit 26b9f648	<a href="#">TensorRT 8.6.1.6</a>
<a href="#">23.07</a>		<a href="#">NVIDIA CUDA 12.1.1</a>		
<a href="#">23.06</a>				
<a href="#">23.05</a>				<a href="#">TensorRT 8.6.1.2</a>
<a href="#">23.04</a>	20.04	<a href="#">NVIDIA CUDA 12.1.0</a>		<a href="#">TensorRT 8.6.1</a>
<a href="#">23.03</a>			<a href="#">TensorRT 8.5.3</a>	
<a href="#">23.02</a>		<a href="#">NVIDIA CUDA 12.0.1</a>		
<a href="#">23.01</a>			<a href="#">TensorRT 8.5.2.2</a>	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">22.12</a>		<a href="#">NVIDIA CUDA 11.8.0</a>		<a href="#">TensorRT 8.5.1</a>
<a href="#">22.11</a>				
<a href="#">22.10</a>				<a href="#">TensorRT 8.5.0.12</a>
<a href="#">22.09</a>				
<a href="#">22.08</a>		<a href="#">NVIDIA CUDA 11.7.1</a>		<a href="#">TensorRT 8.4.2.4</a>
<a href="#">22.07</a>		<a href="#">NVIDIA CUDA 11.7 Update 1 Preview</a>		<a href="#">TensorRT 8.4.1</a>
<a href="#">22.06</a>				<a href="#">TensorRT 8.2.5</a>
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.7.0</a>		
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>		<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>				<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2 with cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c		
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>	
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>	
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>	
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>	
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837		
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>	
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36		
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc		
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>	
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>	
<a href="#">20.02</a>			<a href="#">5.5</a>		
<a href="#">20.01</a>					
<a href="#">19.12</a>					<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>					
<a href="#">19.10</a>			<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>					
<a href="#">19.08</a>					<a href="#">TensorRT 5.1.5</a>

## Known Issues

None.

---

# Chapter 11. Kaldi Release 23.07

The NVIDIA container image for Kaldi, release 23.07 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 22.04](#) including [Python 3.10](#)
- ▶ [NVIDIA CUDA® 12.1.1](#)
- ▶ [NVIDIA cuBLAS 12.1.3.1](#)
- ▶ [NVIDIA cuDNN 8.9.3](#)
- ▶ [NVIDIA NCCL 2.18.3](#)
- ▶ [rdma-core 39.0](#)
- ▶ [Nsight Compute 2023.1.1.4](#)
- ▶ [Nsight Systems 2023.2.3.1001](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.15 with UCX 1.15.0
- ▶ [NVIDIA TensorRT™ 8.6.1.6](#) for x64 Linux
- ▶ [NVIDIA DALI® 1.27.0](#)

## Driver Requirements

Release 23.07 is based on [CUDA 12.1.1](#), which requires [NVIDIA Driver](#) release 530 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), 510.47 (or later R510), 515.65 (or later R515), 525.85 (or later R525), or 530.30 (or later R530).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, R460, and R520 drivers, which are not forward-compatible with CUDA 12.1. For a complete list of supported drivers, see the [CUDA](#)

[Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 23.07 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal™, NVIDIA Volta™, NVIDIA Turing™, NVIDIA Ampere architecture, NVIDIA Hopper™, and NVIDIA Ada Lovelace architecture families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 23.07 is based on [Kaldi 5.5](#) up to [Kaldi 26b9f648](#).
- ▶ Ubuntu 22.04 with June 2023 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
23.07	22.04	<a href="#">NVIDIA CUDA 12.1.1</a>	5.5 up to commit 26b9f648	<a href="#">TensorRT 8.6.1.6</a>
<a href="#">23.06</a>				
<a href="#">23.05</a>				<a href="#">TensorRT 8.6.1.2</a>
<a href="#">23.04</a>	20.04	<a href="#">NVIDIA CUDA 12.1.0</a>		<a href="#">TensorRT 8.6.1</a>
<a href="#">23.03</a>				<a href="#">TensorRT 8.5.3</a>
<a href="#">23.02</a>				
<a href="#">23.01</a>				<a href="#">TensorRT 8.5.2.2</a>
		<a href="#">NVIDIA CUDA 12.0.1</a>		

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">22.12</a>		<a href="#">NVIDIA CUDA 11.8.0</a>		<a href="#">TensorRT 8.5.1</a>
<a href="#">22.11</a>				
<a href="#">22.10</a>				<a href="#">TensorRT 8.5.0.12</a>
<a href="#">22.09</a>				
<a href="#">22.08</a>		<a href="#">NVIDIA CUDA 11.7.1</a>		<a href="#">TensorRT 8.4.2.4</a>
<a href="#">22.07</a>		<a href="#">NVIDIA CUDA 11.7 Update 1 Preview</a>		<a href="#">TensorRT 8.4.1</a>
<a href="#">22.06</a>				<a href="#">TensorRT 8.2.5</a>
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.7.0</a>		
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>		<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>				<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2 with cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c		
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>	
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>	
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>	
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>	
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837		
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>	
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36		
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc		
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>	
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>	
<a href="#">20.02</a>			<a href="#">5.5</a>		
<a href="#">20.01</a>					
<a href="#">19.12</a>					<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>					
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>			
<a href="#">19.09</a>					
<a href="#">19.08</a>					<a href="#">TensorRT 5.1.5</a>

## Known Issues

None.

---

# Chapter 12. Kaldi Release 23.06

The NVIDIA container image for Kaldi, release 23.06 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 22.04](#) including [Python 3.10](#)
- ▶ [NVIDIA CUDA® 12.1.1](#)
- ▶ [NVIDIA cuBLAS 12.1.3.1](#)
- ▶ [NVIDIA cuDNN 8.9.2](#)
- ▶ [NVIDIA NCCL 2.18.1](#)
- ▶ [rdma-core 39.0](#)
- ▶ [Nsight Compute 2023.1.1.4](#)
- ▶ [Nsight Systems 2023.2.3.1001](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.15 with UCX 1.15.0
- ▶ [NVIDIA TensorRT™ 8.6.1.6](#) for x64 Linux
- ▶ [NVIDIA DALI® 1.26.0](#)

## Driver Requirements

Release 23.06 is based on [CUDA 12.1.1](#), which requires [NVIDIA Driver](#) release 530 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), 510.47 (or later R510), 515.65 (or later R515), 525.85 (or later R525), or 530.30 (or later R530).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, R460, and R520 drivers, which are not forward-compatible with CUDA 12.1. For a complete list of supported drivers, see the [CUDA](#)

[Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 23.06 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal™, NVIDIA Volta™, NVIDIA Turing™, NVIDIA Ampere architecture, NVIDIA Hopper™, and NVIDIA Ada Lovelace architecture families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 23.06 is based on [Kaldi 5.5](#) up to [Kaldi 26b9f648](#).
- ▶ Ubuntu 22.04 with May 2023 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
23.06	22.04	<a href="#">NVIDIA CUDA 12.1.1</a>	5.5 up to commit 26b9f648	<a href="#">TensorRT 8.6.1.6</a>
<a href="#">23.05</a>				<a href="#">TensorRT 8.6.1.2</a>
<a href="#">23.04</a>	20.04	<a href="#">NVIDIA CUDA 12.1.0</a>		<a href="#">TensorRT 8.6.1</a>
<a href="#">23.03</a>				<a href="#">TensorRT 8.5.3</a>
<a href="#">23.02</a>				
<a href="#">23.01</a>				<a href="#">TensorRT 8.5.2.2</a>
		<a href="#">NVIDIA CUDA 12.0.1</a>		

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">22.12</a>		<a href="#">NVIDIA CUDA 11.8.0</a>		<a href="#">TensorRT 8.5.1</a>
<a href="#">22.11</a>				
<a href="#">22.10</a>				<a href="#">TensorRT 8.5.0.12</a>
<a href="#">22.09</a>				
<a href="#">22.08</a>		<a href="#">NVIDIA CUDA 11.7.1</a>		<a href="#">TensorRT 8.4.2.4</a>
<a href="#">22.07</a>		<a href="#">NVIDIA CUDA 11.7 Update 1 Preview</a>		<a href="#">TensorRT 8.4.1</a>
<a href="#">22.06</a>				<a href="#">TensorRT 8.2.5</a>
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.7.0</a>		
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>		<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>	<a href="#">TensorRT 8.2.3</a>	
<a href="#">22.01</a>			<a href="#">TensorRT 8.2.2</a>	
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5 up to commit c6772ba4</a>	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5 up to commit c6772ba4</a>	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2 with cuBLAS 11.6.5.2</a>	<a href="#">5.5 up to commit 2897564e</a>	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5 up to commit 40573bb</a>	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5 up to commit 851a0a79</a>	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5 up to commit 9bca2777</a>	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5 up to commit b1786aa5</a>	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5 up to commit e92970ae</a>	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c		
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>	
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>	
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>	
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>	
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837		
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>	
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36		
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc		
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>	
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>	
<a href="#">20.02</a>			<a href="#">5.5</a>		
<a href="#">20.01</a>					
<a href="#">19.12</a>					<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>					
<a href="#">19.10</a>			<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>					
<a href="#">19.08</a>					<a href="#">TensorRT 5.1.5</a>

## Known Issues

None.

---

# Chapter 13. Kaldi Release 23.05

The NVIDIA container image for Kaldi, release 23.05 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 22.04](#) including [Python 3.10](#)
- ▶ [NVIDIA CUDA® 12.1.1](#)
- ▶ [NVIDIA cuBLAS 12.1.3.1](#)
- ▶ [NVIDIA cuDNN 8.9.1.23](#)
- ▶ [NVIDIA NCCL 2.18.1](#)
- ▶ [rdma-core 36.0](#)
- ▶ [Nsight Compute 2023.1.1.4](#)
- ▶ [Nsight Systems 2023.2](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.14 with UCX 1.14.0
- ▶ [NVIDIA TensorRT™ 8.6.1.2](#) for x64 Linux
- ▶ [NVIDIA DALI® 1.25.0](#)

## Driver Requirements

Release 23.05 is based on [CUDA 12.1.1](#), which requires [NVIDIA Driver](#) release 530 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), 510.47 (or later R510), 515.65 (or later R515), 525.85 (or later R525), or 530.30 (or later R530).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, R460, and R520 drivers, which are not forward-compatible with CUDA 12.1. For a complete list of supported drivers, see the [CUDA](#)

[Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 23.05 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal™, NVIDIA Volta™, NVIDIA Turing™, NVIDIA Ampere architecture, NVIDIA Hopper™, and NVIDIA Ada Lovelace architecture families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 23.05 is based on [Kaldi 5.5](#) up to [Kaldi 26b9f648](#).
- ▶ Ubuntu 22.04 with April 2023 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
23.05	22.04	<a href="#">NVIDIA CUDA 12.1.1</a>	5.5 up to commit 26b9f648	<a href="#">TensorRT 8.6.1.2</a>
<a href="#">23.04</a>	20.04	<a href="#">NVIDIA CUDA 12.1.0</a>		<a href="#">TensorRT 8.6.1</a>
<a href="#">23.03</a>				<a href="#">TensorRT 8.5.3</a>
<a href="#">23.02</a>				
<a href="#">23.01</a>				<a href="#">NVIDIA CUDA 12.0.1</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">22.12</a>		<a href="#">NVIDIA CUDA 11.8.0</a>		<a href="#">TensorRT 8.5.1</a>
<a href="#">22.11</a>				
<a href="#">22.10</a>		<a href="#">TensorRT 8.5.0.12</a>		
<a href="#">22.09</a>				
<a href="#">22.08</a>		<a href="#">NVIDIA CUDA 11.7.1</a>		<a href="#">TensorRT 8.4.2.4</a>
<a href="#">22.07</a>		<a href="#">NVIDIA CUDA 11.7 Update 1 Preview</a>		<a href="#">TensorRT 8.4.1</a>
<a href="#">22.06</a>				<a href="#">TensorRT 8.2.5</a>
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.7.0</a>		
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>		<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>	<a href="#">TensorRT 8.2.3</a>	
<a href="#">22.01</a>			<a href="#">TensorRT 8.2.2</a>	
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5 up to commit c6772ba4</a>	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5 up to commit c6772ba4</a>	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2 with cuBLAS 11.6.5.2</a>	<a href="#">5.5 up to commit 2897564e</a>	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5 up to commit 40573bb</a>	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5 up to commit 851a0a79</a>	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5 up to commit 9bca2777</a>	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5 up to commit b1786aa5</a>	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5 up to commit e92970ae</a>	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c		
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>	
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>	
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>	
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>	
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837		
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>	
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36		
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc		
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>	
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>	
<a href="#">20.02</a>			<a href="#">5.5</a>		
<a href="#">20.01</a>					
<a href="#">19.12</a>					<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>					
<a href="#">19.10</a>			<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>					
<a href="#">19.08</a>					<a href="#">TensorRT 5.1.5</a>

## Known Issues

None.

---

# Chapter 14. Kaldi Release 23.04

The NVIDIA container image for Kaldi, release 23.04 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA<sup>®</sup> 12.1.0](#)
- ▶ [NVIDIA cuBLAS 12.1.3](#)
- ▶ [NVIDIA cuDNN 8.9.0](#)
- ▶ [NVIDIA NCCL 2.17.1](#)
- ▶ [rdma-core 36.0](#)
- ▶ [Nsight Compute 2023.1.0.15](#)
- ▶ [Nsight Systems 2023.1.1.127](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.13 with UCX 1.14.0
- ▶ [NVIDIA TensorRT<sup>™</sup> 8.6.1](#) for x64 Linux
- ▶ [NVIDIA DALI<sup>®</sup> 1.24.0](#)

## Driver Requirements

Release 23.04 is based on [CUDA 12.1.0](#), which requires [NVIDIA Driver](#) release 530 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), 510.47 (or later R510), 515.65 (or later R515), 525.85 (or later R525), or 530.30 (or later R530).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, R460, and R520 drivers, which are not forward-compatible with CUDA 12.1. For a complete list of supported drivers, see the [CUDA](#)

[Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 23.04 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal™, NVIDIA Volta™, NVIDIA Turing™, NVIDIA Ampere architecture, NVIDIA Hopper™, and NVIDIA Ada Lovelace architecture families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 23.04 is based on [Kaldi 5.5](#) up to [Kaldi 26b9f648](#).
- ▶ Ubuntu 20.04 with March 2023 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
23.04	20.04	<a href="#">NVIDIA CUDA 12.1.0</a>	5.5 up to commit 26b9f648	<a href="#">TensorRT 8.6.1</a>
<a href="#">23.03</a>				<a href="#">TensorRT 8.5.3</a>
<a href="#">23.02</a>		<a href="#">NVIDIA CUDA 12.0.1</a>		
<a href="#">23.01</a>				<a href="#">TensorRT 8.5.2.2</a>
<a href="#">22.12</a>		<a href="#">NVIDIA CUDA 11.8.0</a>		<a href="#">TensorRT 8.5.1</a>
<a href="#">22.11</a>				
<a href="#">22.10</a>				<a href="#">TensorRT 8.5.0.12</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">22.09</a>				
<a href="#">22.08</a>		<a href="#">NVIDIA CUDA 11.7.1</a>		<a href="#">TensorRT 8.4.2.4</a>
<a href="#">22.07</a>		<a href="#">NVIDIA CUDA 11.7 Update 1</a>		<a href="#">TensorRT 8.4.1</a>
<a href="#">22.06</a>		<a href="#">Preview</a>		<a href="#">TensorRT 8.2.5</a>
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.7.0</a>		
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>		<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>				<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5 up to commit c6772ba4</a>	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5 up to commit c6772ba4</a>	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2 with cuBLAS 11.6.5.2</a>	<a href="#">5.5 up to commit 2897564e</a>	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5 up to commit 40573bb</a>	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5 up to commit 851a0a79</a>	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5 up to commit 9bca2777</a>	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5 up to commit b1786aa5</a>	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5 up to commit e92970ae</a>	
<a href="#">21.04</a>			<a href="#">5.5 up to commit ec554a7c</a>	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5 up to commit 376281ca</a>	<a href="#">TensorRT 7.2.2.3</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				
<a href="#">19.11</a>				
<a href="#">19.10</a>	<a href="#">NVIDIA CUDA 10.1.243</a>			
<a href="#">19.09</a>				
<a href="#">19.08</a>			<a href="#">TensorRT 5.1.5</a>	

## Known Issues

None.

---

# Chapter 15. Kaldi Release 23.03

The NVIDIA container image for Kaldi, release 23.03 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA® 12.1.0](#)
- ▶ [NVIDIA cuBLAS from CUDA 12.1.0](#)
- ▶ [NVIDIA cuDNN 8.8.1.3](#)
- ▶ [NVIDIA NCCL 2.16.5](#)
- ▶ [rdma-core 36.0](#)
- ▶ [Nsight Compute 2023.1.0.15](#)
- ▶ [Nsight Systems 2023.1.1.127](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.13 with UCX 1.14.0
- ▶ [NVIDIA TensorRT™ 8.5.3](#) for x64 Linux
- ▶ [NVIDIA DALI® 1.23.0](#)

## Driver Requirements

Release 23.03 is based on [CUDA 12.1.0](#), which requires [NVIDIA Driver](#) release 530 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), 510.47 (or later R510), 515.65 (or later R515), 525.85 (or later R525), or 530.30 (or later R530).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, R460, and R520 drivers, which are not forward-compatible with CUDA 12.1. For a complete list of supported drivers, see the [CUDA](#)

[Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 23.03 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal, NVIDIA Volta™, NVIDIA Turing™, NVIDIA Ampere architecture, and NVIDIA Hopper™ architecture families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 23.03 is based on [Kaldi 5.5](#) up to [Kaldi 26b9f648](#).
- ▶ Ubuntu 20.04 with February 2023 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
23.03	20.04	<a href="#">NVIDIA CUDA 12.1.0</a>	5.5 up to commit 26b9f648	<a href="#">TensorRT 8.5.3</a>
<a href="#">23.02</a>		<a href="#">NVIDIA CUDA 12.0.1</a>		<a href="#">TensorRT 8.5.2.2</a>
<a href="#">23.01</a>		<a href="#">NVIDIA CUDA 11.8.0</a>		<a href="#">TensorRT 8.5.1</a>
<a href="#">22.12</a>				<a href="#">TensorRT 8.5.0.12</a>
<a href="#">22.11</a>				
<a href="#">22.10</a>				

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">22.09</a>				
<a href="#">22.08</a>		<a href="#">NVIDIA CUDA 11.7.1</a>		<a href="#">TensorRT 8.4.2.4</a>
<a href="#">22.07</a>		<a href="#">NVIDIA CUDA 11.7 Update 1 Preview</a>		<a href="#">TensorRT 8.4.1</a>
<a href="#">22.06</a>				<a href="#">TensorRT 8.2.5</a>
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.7.0</a>		
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>		<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>				<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2 with cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				
<a href="#">19.11</a>				
<a href="#">19.10</a>	<a href="#">NVIDIA CUDA 10.1.243</a>			
<a href="#">19.09</a>				
<a href="#">19.08</a>			<a href="#">TensorRT 5.1.5</a>	

## Known Issues

None.

---

# Chapter 16. Kaldi Release 23.02

The NVIDIA container image for Kaldi, release 23.02 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA® 12.0.1](#)
- ▶ [NVIDIA cuBLAS from CUDA 12.0.1](#)
- ▶ [NVIDIA cuDNN 8.7.0](#)
- ▶ [NVIDIA NCCL 2.16.5](#)
- ▶ [rdma-core 36.0](#)
- ▶ [Nsight Compute 2022.4.1.6](#)
- ▶ [Nsight Systems 2022.5.1.93](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.13 with UCX 1.14.0
- ▶ [NVIDIA TensorRT™ 8.5.3](#) for x64 Linux
- ▶ [NVIDIA DALI® 1.22.0](#)

## Driver Requirements

Release 23.02 is based on [CUDA 12.0.1](#), which requires [NVIDIA Driver](#) release 525 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), 510.47 (or later R510), 515.65 (or later R515), or 525.85 (or later R525).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, R460, and R520 drivers, which are not forward-compatible with CUDA 12.0. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 23.02 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal, NVIDIA Volta™, NVIDIA Turing™, NVIDIA Ampere architecture, and NVIDIA Hopper™ architecture families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 23.02 is based on [Kaldi 5.5](#) up to [Kaldi 26b9f648](#).
- ▶ Ubuntu 20.04 with January 2023 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
23.02	20.04	<a href="#">NVIDIA CUDA 12.0.1</a>	<a href="#">5.5</a> up to commit <a href="#">26b9f648</a>	<a href="#">TensorRT 8.5.3</a>
<a href="#">23.01</a>				<a href="#">TensorRT 8.5.2.2</a>
<a href="#">22.12</a>		<a href="#">NVIDIA CUDA 11.8.0</a>		<a href="#">TensorRT 8.5.1</a>
<a href="#">22.11</a>				
<a href="#">22.10</a>				<a href="#">TensorRT 8.5.0.12</a>
<a href="#">22.09</a>				
<a href="#">22.08</a>		<a href="#">NVIDIA CUDA 11.7.1</a>		<a href="#">TensorRT 8.4.2.4</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">22.07</a>		<a href="#">NVIDIA CUDA 11.7 Update 1 Preview</a>		<a href="#">TensorRT 8.4.1</a>
<a href="#">22.06</a>				<a href="#">TensorRT 8.2.5</a>
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.7.0</a>		
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>		<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>				<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5 up to commit c6772ba4</a>	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5 up to commit c6772ba4</a>	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2 with cuBLAS 11.6.5.2</a>	<a href="#">5.5 up to commit 2897564e</a>	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5 up to commit 40573bb</a>	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5 up to commit 851a0a79</a>	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5 up to commit 9bca2777</a>	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5 up to commit b1786aa5</a>	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5 up to commit e92970ae</a>	
<a href="#">21.04</a>			<a href="#">5.5 up to commit ec554a7c</a>	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5 up to commit 376281ca</a>	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5 up to commit b78fdcaf</a>	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5 up to commit 7905fe5a</a>	<a href="#">TensorRT 7.2.2</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

None.

---

# Chapter 17. Kaldi Release 23.01

The NVIDIA container image for Kaldi, release 23.01 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA® 12.0.1](#)
- ▶ [NVIDIA cuBLAS from CUDA 12.0.1](#)
- ▶ [NVIDIA cuDNN 8.7.0](#)
- ▶ [NVIDIA NCCL 2.16.5](#)
- ▶ [rdma-core 36.0](#)
- ▶ [Nsight Compute 2022.4.1.6](#)
- ▶ [Nsight Systems 2022.5.1](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.13 with UCX 1.14.0
- ▶ [NVIDIA TensorRT™ 8.5.2.2](#) for x64 Linux
- ▶ [NVIDIA DALI® 1.21.0](#)

## Driver Requirements

Release 23.01 is based on [CUDA 12.0.1](#), which requires [NVIDIA Driver](#) release 525 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), 510.47 (or later R510), 515.65 (or later R515), or 525.85 (or later R525).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, and R460 drivers, which are not forward-compatible with CUDA 12.0. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 23.01 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal, NVIDIA Volta™, NVIDIA Turing™, NVIDIA Ampere architecture, and NVIDIA Hopper™ architecture families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 23.01 is based on [Kaldi 5.5](#) up to [Kaldi 26b9f648](#).
- ▶ Ubuntu 20.04 with December 2022 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
23.01	20.04	<a href="#">NVIDIA CUDA 12.0.1</a>	5.5 up to commit 26b9f648	<a href="#">TensorRT 8.5.2.2</a>
<a href="#">22.12</a>		<a href="#">NVIDIA CUDA 11.8.0</a>		<a href="#">TensorRT 8.5.1</a>
<a href="#">22.11</a>				
<a href="#">22.10</a>				<a href="#">TensorRT 8.5.0.12</a>
<a href="#">22.09</a>				
<a href="#">22.08</a>		<a href="#">NVIDIA CUDA 11.7.1</a>		<a href="#">TensorRT 8.4.2.4</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">22.07</a>		<a href="#">NVIDIA CUDA 11.7 Update 1 Preview</a>		<a href="#">TensorRT 8.4.1</a>
<a href="#">22.06</a>				<a href="#">TensorRT 8.2.5</a>
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.7.0</a>		
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>		<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>				<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5 up to commit c6772ba4</a>	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5 up to commit c6772ba4</a>	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2 with cuBLAS 11.6.5.2</a>	<a href="#">5.5 up to commit 2897564e</a>	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5 up to commit 40573bb</a>	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5 up to commit 851a0a79</a>	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5 up to commit 9bca2777</a>	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5 up to commit b1786aa5</a>	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5 up to commit e92970ae</a>	
<a href="#">21.04</a>			<a href="#">5.5 up to commit ec554a7c</a>	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5 up to commit 376281ca</a>	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5 up to commit b78fdcaf</a>	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5 up to commit 7905fe5a</a>	<a href="#">TensorRT 7.2.2</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

None.

---

# Chapter 18. Kaldi Release 22.12

The NVIDIA container image for Kaldi, release 22.12 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA<sup>®</sup> 11.8.0](#)
- ▶ [NVIDIA cuBLAS 11.11.3.6](#)
- ▶ [NVIDIA cuDNN 8.7.0](#)
- ▶ [NVIDIA NCCL 2.15.5](#)
- ▶ [rdma-core 36.0](#)
- ▶ [Nsight Compute 2022.3.0.22](#)
- ▶ [Nsight Systems 2022.4.2.1](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.13 with UCX 1.14.0
- ▶ [NVIDIA TensorRT<sup>™</sup> 8.5.1](#) for x64 Linux
- ▶ [NVIDIA DALI<sup>®</sup> 1.20.0](#)

## Driver Requirements

Release 22.12 is based on [CUDA 11.8.0](#), which requires [NVIDIA Driver](#) release 520 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), 510.47 (or later R510), or 515.65 (or later R515).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, and R460 drivers, which are not forward-compatible with CUDA 11.8. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 22.12 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal, NVIDIA Volta™, NVIDIA Turing™, NVIDIA Ampere architecture, and NVIDIA Hopper™ architecture families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 22.12 is based on [Kaldi 5.5](#) up to [Kaldi 26b9f648](#).
- ▶ Ubuntu 20.04 with November 2022 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
22.12	20.04	<a href="#">NVIDIA CUDA 11.8.0</a>	<a href="#">5.5</a> up to commit 26b9f648	<a href="#">TensorRT 8.5.1</a>
<a href="#">22.11</a>				
<a href="#">22.10</a>				<a href="#">TensorRT 8.5.0.12</a>
<a href="#">22.09</a>				
<a href="#">22.08</a>		<a href="#">NVIDIA CUDA 11.7.1</a>		<a href="#">TensorRT 8.4.2.4</a>
<a href="#">22.07</a>		<a href="#">NVIDIA CUDA 11.7 Update 1</a>		<a href="#">TensorRT 8.4.1</a>
<a href="#">22.06</a>		<a href="#">11.7 Update 1 Preview</a>		<a href="#">TensorRT 8.2.5</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.7.0</a>		
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>		<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>				<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2 with cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

### Known Issues

None.

---

# Chapter 19. Kaldi Release 22.11

The NVIDIA container image for Kaldi, release 22.11 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA<sup>®</sup> 11.8.0](#)
- ▶ [NVIDIA cuBLAS 11.11.3.6](#)
- ▶ [NVIDIA cuDNN 8.7.0](#)
- ▶ [NVIDIA NCCL 2.15.5](#)
- ▶ [rdma-core 36.0](#)
- ▶ [Nsight Compute 2022.3.0.22](#)
- ▶ [Nsight Systems 2022.4.2.1](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.12.2tp1 with UCX 1.14.0
- ▶ [NVIDIA TensorRT<sup>™</sup> 8.5.1](#) for x64 Linux
- ▶ [NVIDIA DALI<sup>®</sup> 1.18.0](#)

## Driver Requirements

Release 22.11 is based on [CUDA 11.8.0](#), which requires [NVIDIA Driver](#) release 520 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), 510.47 (or later R510), or 515.65 (or later R515).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, and R460 drivers, which are not forward-compatible with CUDA 11.8. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 22.11 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal, NVIDIA Volta™, NVIDIA Turing™, NVIDIA Ampere architecture, and NVIDIA Hopper™ architecture families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 22.11 is based on [Kaldi 26b9f648](#).
- ▶ Ubuntu 20.04 with October 2022 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
22.11	20.04	<a href="#">NVIDIA CUDA 11.8.0</a>	<a href="#">5.5 up to commit 26b9f648</a>	<a href="#">TensorRT 8.5.1</a>
<a href="#">22.10</a>				<a href="#">TensorRT 8.5.0.12</a>
<a href="#">22.09</a>				
<a href="#">22.08</a>		<a href="#">NVIDIA CUDA 11.7.1</a>		<a href="#">TensorRT 8.4.2.4</a>
<a href="#">22.07</a>		<a href="#">NVIDIA CUDA 11.7 Update 1</a>		<a href="#">TensorRT 8.4.1</a>
<a href="#">22.06</a>		<a href="#">11.7 Update 1 Preview</a>		<a href="#">TensorRT 8.2.5</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.7.0</a>		
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>		<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>				<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5 up to commit c6772ba4</a>	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5 up to commit c6772ba4</a>	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2 with cuBLAS 11.6.5.2</a>	<a href="#">5.5 up to commit 2897564e</a>	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5 up to commit 40573bb</a>	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5 up to commit 851a0a79</a>	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5 up to commit 9bca2777</a>	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5 up to commit b1786aa5</a>	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5 up to commit e92970ae</a>	
<a href="#">21.04</a>			<a href="#">5.5 up to commit ec554a7c</a>	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5 up to commit 376281ca</a>	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5 up to commit b78fdcaf</a>	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5 up to commit 7905fe5a</a>	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5 up to commit edf6963c</a>	<a href="#">TensorRT 7.2.1</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

None.

---

# Chapter 20. Kaldi Release 22.10

The NVIDIA container image for Kaldi, release 22.10 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA<sup>®</sup> 11.8.0](#)
- ▶ [NVIDIA cuBLAS 11.11.3.6](#)
- ▶ [NVIDIA cuDNN 8.6.0.163](#)
- ▶ [NVIDIA NCCL 2.15.5](#)
- ▶ [rdma-core 36.0](#)
- ▶ [Nsight Compute 2022.3.0.22](#)
- ▶ [Nsight Systems 2022.4.2.1](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.12.2tp1 with UCX 1.14.0
- ▶ [NVIDIA TensorRT™ 8.5.0.12](#) for x64 Linux
- ▶ [NVIDIA DALI<sup>®</sup> 1.18.0](#)

## Driver Requirements

Release 22.10 is based on [CUDA 11.8.0](#), which requires [NVIDIA Driver](#) release 520 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), 510.47 (or later R510), or 515.65 (or later R515).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, and R460 drivers, which are not forward-compatible with CUDA 11.8. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 22.10 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal, NVIDIA Volta™, NVIDIA Turing™, NVIDIA Ampere architecture, and NVIDIA Hopper™ architecture families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 22.10 is based on [Kaldi 26b9f648](#).
- ▶ Ubuntu 20.04 with September 2022 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT	
22.10	20.04	<a href="#">NVIDIA CUDA 11.8.0</a>	5.5 up to commit 26b9f648	<a href="#">TensorRT 8.5.0.12</a>	
<a href="#">22.09</a>					
<a href="#">22.08</a>		<a href="#">NVIDIA CUDA 11.7.1</a>			<a href="#">TensorRT 8.4.2.4</a>
<a href="#">22.07</a>		<a href="#">NVIDIA CUDA 11.7 Update 1 Preview</a>			<a href="#">TensorRT 8.4.1</a>
<a href="#">22.06</a>					<a href="#">TensorRT 8.2.5</a>
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.7.0</a>			

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>		<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>				<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2 with cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>			<a href="#">NVIDIA CUDA 10.1.243</a>	
<a href="#">19.11</a>				
<a href="#">19.10</a>				
<a href="#">19.09</a>				
<a href="#">19.08</a>		<a href="#">TensorRT 5.1.5</a>		

## Known Issues

None.

---

# Chapter 21. Kaldi Release 22.09

The NVIDIA container image for Kaldi, release 22.09 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA® 11.8.0](#)
- ▶ [NVIDIA cuBLAS 11.11.3.6](#)
- ▶ [NVIDIA cuDNN 8.6.0.163](#)
- ▶ [NVIDIA NCCL 2.15.1](#)
- ▶ [rdma-core 36.0](#)
- ▶ [Nsight Compute 2022.3.0.22](#)
- ▶ [Nsight Systems 2022.3.1.43](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.12.1a0 with UCX 1.14.0
- ▶ [NVIDIA TensorRT™ 8.5.0.12](#) for x64 Linux
- ▶ [NVIDIA DALI® 1.17.0](#)

## Driver Requirements

Release 22.09 is based on [CUDA 11.8.0](#), which requires [NVIDIA Driver](#) release 520 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), 510.47 (or later R510), or 515.65 (or later R515).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, and R460 drivers, which are not forward-compatible with CUDA 11.8. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 22.09 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal, NVIDIA Volta™, NVIDIA Turing™, NVIDIA Ampere architecture, and NVIDIA Hopper™ architecture families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 22.09 is based on [Kaldi 26b9f648](#).
- ▶ Ubuntu 20.04 with August 2022 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">22.09</a>	20.04	<a href="#">NVIDIA CUDA 11.8.0</a>	5.5 up to commit 26b9f648	<a href="#">TensorRT 8.5.0.12</a>
<a href="#">22.08</a>		<a href="#">NVIDIA CUDA 11.7.1</a>		<a href="#">TensorRT 8.4.2.4</a>
<a href="#">22.07</a>		<a href="#">NVIDIA CUDA 11.7 Update 1</a>		<a href="#">TensorRT 8.4.1</a>
<a href="#">22.06</a>		<a href="#">Preview</a>		<a href="#">TensorRT 8.2.5</a>
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.7.0</a>		

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>		<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>				<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2 with cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>			<a href="#">NVIDIA CUDA 10.1.243</a>	
<a href="#">19.11</a>				
<a href="#">19.10</a>				
<a href="#">19.09</a>				
<a href="#">19.08</a>		<a href="#">TensorRT 5.1.5</a>		

## Known Issues

None.

---

# Chapter 22. Kaldi Release 22.08

The NVIDIA container image for Kaldi, release 22.08 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA<sup>®</sup> 11.7.1](#)
- ▶ [NVIDIA cuBLAS 11.10.3.66](#)
- ▶ [NVIDIA cuDNN 8.5.0.96](#)
- ▶ [NVIDIA NCCL 2.12.12](#)
- ▶ [rdma-core 36.0](#)
- ▶ [Nsight Compute 2022.1.1.2](#)
- ▶ [Nsight Systems 2022.1.3.18](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.10 with UCX 1.12.0
- ▶ [NVIDIA TensorRT<sup>™</sup> 8.4.2.4](#) for x64 Linux
- ▶ [NVIDIA DALI<sup>®</sup> 1.16.0](#)

## Driver Requirements

Release 22.08 is based on [CUDA 11.7.1](#), which requires [NVIDIA Driver](#) release 515 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), or 510.47 (or later R510).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, and R460 drivers, which are not forward-compatible with CUDA 11.7. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 22.08 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal, NVIDIA Volta™, NVIDIA Turing™, and NVIDIA Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 22.08 is based on [Kaldi 26b9f648](#).
- ▶ Ubuntu 20.04 with July 2022 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT	
<a href="#">22.08</a>	20.04	<a href="#">NVIDIA CUDA 11.7.1</a>		<a href="#">TensorRT 8.4.2.4</a>	
<a href="#">22.07</a>		<a href="#">NVIDIA CUDA 11.7 Update 1 Preview</a>		5.5 up to commit 26b9f648	<a href="#">TensorRT 8.4.1</a>
<a href="#">22.06</a>		<a href="#">NVIDIA CUDA 11.7.0</a>			<a href="#">TensorRT 8.2.5</a>
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.6.2</a>			
<a href="#">22.04</a>					<a href="#">TensorRT 8.2.4.2</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>				<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	5.5 up to commit c6772ba4	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			5.5 up to commit c6772ba4	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2 with cuBLAS 11.6.5.2</a>	5.5 up to commit 2897564e	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	5.5 up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	5.5 up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	5.5 up to commit 9bca2777	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	5.5 up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	5.5 up to commit e92970ae	
<a href="#">21.04</a>			5.5 up to commit ec554a7c	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	5.5 up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	5.5 up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	5.5 up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	5.5 up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			5.5 up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	5.5 up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

### Known Issues

None.

---

# Chapter 23. Kaldi Release 22.07

The NVIDIA container image for Kaldi, release 22.07 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA<sup>®</sup> 11.7 Update 1 Preview](#)
- ▶ [NVIDIA cuBLAS 11.10.3.66](#)
- ▶ [NVIDIA cuDNN 8.4.1](#)
- ▶ [NVIDIA NCCL 2.12.12](#)
- ▶ [rdma-core 36.0](#)
- ▶ [Nsight Compute 2022.1.1.2](#)
- ▶ [Nsight Systems 2022.1.3.3](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.10 with UCX 1.12.0
- ▶ [NVIDIA TensorRT<sup>™</sup> 8.4.1](#) for x64 Linux
- ▶ [NVIDIA DALI<sup>®</sup> 1.15.0](#)

## Driver Requirements

Release 22.07 is based on [CUDA 11.7 Update 1 Preview](#), which requires [NVIDIA Driver](#) release 515 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), or 510.47 (or later R510).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, and R460 drivers, which are not forward-compatible with CUDA 11.7. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 22.07 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal, NVIDIA Volta™, NVIDIA Turing™, and NVIDIA Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 22.07 is based on [Kaldi 26b9f648](#).
- ▶ Ubuntu 20.04 with June 2022 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT	
22.07	20.04	<a href="#">NVIDIA CUDA 11.7 Update 1 Preview</a>	5.5 up to commit 26b9f648	<a href="#">TensorRT 8.4.1</a>	
<a href="#">22.06</a>				<a href="#">TensorRT 8.2.5</a>	
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.7.0</a>			
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>			<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>			<a href="#">TensorRT 8.2.3</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT	
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>	
<a href="#">22.01</a>				<a href="#">TensorRT 8.2.2</a>	
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.2.1.8</a>	
<a href="#">21.11</a>				<a href="#">TensorRT 8.0.3.4</a>	
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2</a> with <a href="#">cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e		
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>	
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>	
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777		
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>	
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae		
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c		
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>	
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>	
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>	
<a href="#">20.11</a>		18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>				<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>			<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>				<a href="#">5.5</a> up to commit f88d5a36	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

None.

---

# Chapter 24. Kaldi Release 22.06

The NVIDIA container image for Kaldi, release 22.06 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA<sup>®</sup> 11.7 Update 1 Preview](#)
- ▶ [NVIDIA cuBLAS 11.10.3.66](#)
- ▶ [NVIDIA cuDNN 8.4.1](#)
- ▶ [NVIDIA NCCL 2.12.12](#)
- ▶ [rdma-core 36.0](#)
- ▶ [Nsight Compute 2022.1.1.2](#)
- ▶ [Nsight Systems 2022.1.3.3](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.10 with UCX 1.12.0
- ▶ [NVIDIA TensorRT<sup>™</sup> 8.2.5](#) for x64 Linux
- ▶ [NVIDIA DALI<sup>®</sup> 1.14.0](#)

## Driver Requirements

Release 22.06 is based on [CUDA 11.7 Update 1 Preview](#), which requires [NVIDIA Driver](#) release 515 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), or 510.47 (or later R510).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, and R460 drivers, which are not forward-compatible with CUDA 11.7. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 22.06 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal, NVIDIA Volta™, NVIDIA Turing™, and NVIDIA Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 22.06 is based on [Kaldi 26b9f648](#).
- ▶ Ubuntu 20.04 with May 2022 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
22.06	20.04	<a href="#">NVIDIA CUDA 11.7 Update 1 Preview</a>	5.5 up to commit 26b9f648	<a href="#">TensorRT 8.2.5</a>
<a href="#">22.05</a>		<a href="#">NVIDIA CUDA 11.7.0</a>		
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>		<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT	
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>	
<a href="#">22.01</a>				<a href="#">TensorRT 8.2.2</a>	
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.2.1.8</a>	
<a href="#">21.11</a>				<a href="#">TensorRT 8.0.3.4</a>	
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2 with cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e		
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>	
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>	
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777		
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>	
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae		
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c		
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>	
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>	
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>	
<a href="#">20.11</a>		18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>				<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>			<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>				<a href="#">5.5</a> up to commit f88d5a36	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

None.

---

# Chapter 25. Kaldi Release 22.05

The NVIDIA container image for Kaldi, release 22.05 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA<sup>®</sup> 11.7.0](#)
- ▶ [NVIDIA cuBLAS 11.10.1.25](#)
- ▶ [NVIDIA cuDNN 8.4.0.27](#)
- ▶ [NVIDIA NCCL 2.12.10](#) (optimized for [NVIDIA NVLink<sup>®</sup>](#))
- ▶ [rdma-core 36.0](#)
- ▶ [Nsight Compute 2022.1.1.2](#)
- ▶ [Nsight Systems 2022.1.3.3](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.10 with UCX 1.12.0
- ▶ [NVIDIA TensorRT<sup>™</sup> 8.2.5.1](#) for x64 Linux
- ▶ [NVIDIA DALI<sup>®</sup> 1.13.0](#)

## Driver Requirements

Release 22.05 is based on [CUDA 11.7](#), which requires [NVIDIA Driver](#) release 515 or later. However, if you are running on a data center GPU (for example, T4 or any other data center GPU), you can use NVIDIA driver release 450.51 (or later R450), 470.57 (or later R470), or 510.47 (or later R510).

The CUDA driver's compatibility package only supports particular drivers. Thus, users should upgrade from all R418, R440, and R460 drivers, which are not forward-compatible with CUDA 11.7. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 22.05 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal, NVIDIA Volta™, NVIDIA Turing™, and NVIDIA Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 22.05 is based on [Kaldi 26b9f648](#).
- ▶ Ubuntu 20.04 with April 2022 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">22.05</a>	20.04	<a href="#">NVIDIA CUDA 11.7.0</a>	5.5 up to commit 26b9f648	<a href="#">TensorRT 8.2.5</a>
<a href="#">22.04</a>		<a href="#">NVIDIA CUDA 11.6.2</a>		<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>		<a href="#">TensorRT 8.2.2</a>		

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2</a> with <a href="#">cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>		18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>	<a href="#">NVIDIA CUDA 11.0.3</a>		<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>	<a href="#">NVIDIA CUDA 11.0.194</a>		<a href="#">5.5</a> up to commit 67db30cc	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5 up to commit 2c7e78f</a>	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5 up to commit da71f301</a>	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>	<a href="#">5.5</a>			
<a href="#">20.01</a>				
<a href="#">19.12</a>			<a href="#">TensorRT 6.0.1</a>	
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>			<a href="#">TensorRT 5.1.5</a>	

## Known Issues

None.

---

# Chapter 26. Kaldi Release 22.04

The NVIDIA container image for Kaldi, release 22.04 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA<sup>®</sup> 11.6.2](#)
- ▶ [cuBLAS 11.9.3.115](#)
- ▶ [NVIDIA cuDNN 8.4.0.27](#)
- ▶ [NVIDIA NCCL 2.12.10](#) (optimized for [NVIDIA NVLink<sup>®</sup>](#))
- ▶ [rdma-core 36.0](#)
- ▶ [Nsight Compute 2022.1.1.2](#)
- ▶ [Nsight Systems 2022.2.1.31-5fe97ab](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.10 with UCX 1.12.0
- ▶ [NVIDIA TensorRT<sup>™</sup> 8.2.4.2](#) for x64 Linux
- ▶ [NVIDIA DALI<sup>®</sup> 1.12.0](#)
- ▶ RAPIDS 22.02

## Driver Requirements

Release 22.04 is based on [CUDA 11.6.2](#), which requires [NVIDIA Driver](#) release 510 or later. However, if you are running on a Data Center GPU (for example, T4 or any other Tesla board), use NVIDIA driver release 418.40 (or later R418), 440.33 (or later R440), 450.51 (or later R450), 460.27 (or later R460), or 470.57 (or later R470). The CUDA driver's compatibility package only supports specific drivers. For a complete list of supported drivers, see [CUDA Application Compatibility](#). For more information, see [CUDA Compatibility and Upgrades](#) and [NVIDIA CUDA and Drivers Support](#).

## GPU Requirements

Release 22.04 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal, NVIDIA Volta™, NVIDIA Turing™, and NVIDIA Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 22.04 is based on [Kaldi 26b9f648](#).
- ▶ Ubuntu 20.04 with March 2022 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">22.04</a>	20.04	<a href="#">NVIDIA CUDA 11.6.2</a>	<a href="#">5.5</a> up to commit 26b9f648	<a href="#">TensorRT 8.2.4.2</a>
<a href="#">22.03</a>		<a href="#">NVIDIA CUDA 11.6.1</a>	<a href="#">5.5</a> up to commit 26b9f648	<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>	<a href="#">5.5</a> up to commit 26b9f648	<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>			<a href="#">5.5</a> up to commit 26b9f648	<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.2.1.8</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">21.11</a>			<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2</a> with <a href="#">cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a> <a href="#">20.01</a>			<a href="#">5.5</a>	
<a href="#">19.12</a> <a href="#">19.11</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		<a href="#">TensorRT 6.0.1</a>
<a href="#">19.10</a> <a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

### Known Issues

None.

---

# Chapter 27. Kaldi Release 22.03

The NVIDIA container image for Kaldi, release 22.03 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in the `/opt/kaldi` directory.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA<sup>®</sup> 11.6.1](#)
- ▶ [cuBLAS 11.8.1.74](#)
- ▶ [NVIDIA cuDNN 8.3.3.40](#)
- ▶ [NVIDIA NCCL 2.12.9](#) (optimized for [NVIDIA NVLink<sup>®</sup>](#))
- ▶ [rdma-core 36.0](#)
- ▶ [Nsight Compute 2022.1.1.2](#)
- ▶ [Nsight Systems 2021.5.2.53](#)
- ▶ GDRCopy 2.3
- ▶ NVIDIA HPC-X 2.10 with UCX 1.12.0
- ▶ [NVIDIA TensorRT<sup>™</sup> 8.2.3](#) for x64 Linux
- ▶ [NVIDIA DALI<sup>®</sup> 1.11.1](#)
- ▶ RAPIDS 22.02

## Driver Requirements

Release 22.03 is based on [CUDA 11.6.1](#), which requires [NVIDIA Driver](#) release 510 or later. However, if you are running on a Data Center GPU (for example, T4 or any other Tesla board), use NVIDIA driver release 418.40 (or later R418), 440.33 (or later R440), 450.51 (or later R450), 460.27 (or later R460), or 470.57 (or later R470). The CUDA driver's compatibility package only supports specific drivers. For a complete list of supported drivers, see [CUDA Application Compatibility](#). For more information, see [CUDA Compatibility and Upgrades](#) and [NVIDIA CUDA and Drivers Support](#).

## GPU Requirements

Release 22.03 supports CUDA compute capability 6.0 and later. This corresponds to GPUs in the NVIDIA Pascal, NVIDIA Volta™, NVIDIA Turing™, and NVIDIA Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 22.03 is based on [Kaldi 26b9f648](#).
- ▶ Ubuntu 20.04 with February 2022 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh`, which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh`, which runs inference on the trained model and data.

Setting the `ONLINE=1` environment variable makes the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
22.03	20.04	<a href="#">NVIDIA CUDA 11.6.1</a>	<a href="#">5.5</a> up to commit 26b9f648	<a href="#">TensorRT 8.2.3</a>
<a href="#">22.02</a>		<a href="#">NVIDIA CUDA 11.6.0</a>	<a href="#">5.5</a> up to commit 26b9f648	<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>			<a href="#">5.5</a> up to commit 26b9f648	<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.0.3.4</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2</a> with <a href="#">cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA</a>		
<a href="#">19.09</a>		<a href="#">10.1.243</a>		
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

### Known Issues

None.

---

# Chapter 28. Kaldi Release 22.02

The NVIDIA container image for Kaldi, release 22.02 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA 11.6.0](#)
- ▶ [cuBLAS 11.8.1.74](#)
- ▶ [NVIDIA cuDNN 8.3.2.44](#)
- ▶ [NVIDIA NCCL 2.11.4](#) (optimized for [NVLink™](#))
- ▶ [rdma-core 36.0](#)
- ▶ NVIDIA HPC-X 2.10
- ▶ [OpenMPI 4.1.2rc4+](#)
- ▶ OpenUCX 1.12.0
- ▶ GDRCopy 2.3
- ▶ [Nsight Systems 2021.5.2.53](#)
- ▶ [TensorRT 8.2.3](#)
- ▶ SHARP 2.5
- ▶ [DALI 1.10](#)

## Driver Requirements

Release 22.02 is based on [NVIDIA CUDA 11.6.0](#), which requires [NVIDIA Driver](#) release 510 or later. However, if you are running on a Data Center GPU (for example, T4 or any other Tesla board), you may use NVIDIA driver release 418.40 (or later R418), 440.33 (or later R440), 450.51 (or later R450), 460.27 (or later R460), or 470.57 (or later R470). The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#) and [NVIDIA CUDA and Drivers Support](#).

## GPU Requirements

Release 22.02 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 22.02 is based on [Kaldi 26b9f648](#).
- ▶ Ubuntu 20.04 with January 2022 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
22.02	20.04	<a href="#">NVIDIA CUDA 11.6.0</a>		<a href="#">TensorRT 8.2.3</a>
<a href="#">22.01</a>			<a href="#">5.5</a> up to commit 26b9f648	<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2</a> with <a href="#">cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT	
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>	
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777		
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>	
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae		
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c		
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca		
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>	
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>	
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>	
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837		
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>	
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36		
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc		
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>	
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>	
<a href="#">20.02</a>			<a href="#">5.5</a>		
<a href="#">20.01</a>					
<a href="#">19.12</a>					<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>					

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

None.

---

# Chapter 29. Kaldi Release 22.01

The NVIDIA container image for Kaldi, release 22.01 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA 11.6.0](#)
- ▶ [cuBLAS 11.8.1.74](#)
- ▶ [NVIDIA cuDNN 8.3.2.44](#)
- ▶ [NVIDIA NCCL 2.11.4](#) (optimized for [NVLink™](#))
- ▶ [rdma-core 36.0](#)
- ▶ NVIDIA HPC-X 2.10
- ▶ [OpenMPI 4.1.2rc4+](#)
- ▶ OpenUCX 1.12.0
- ▶ GDRCopy 2.3
- ▶ [Nsight Systems 2021.5.2.53](#)
- ▶ [TensorRT 8.2.2](#)
- ▶ SHARP 2.5
- ▶ [DALI 1.9](#)

## Driver Requirements

Release 22.01 is based on [NVIDIA CUDA 11.6.0](#), which requires [NVIDIA Driver](#) release 510 or later. However, if you are running on a Data Center GPU (for example, T4 or any other Tesla board), you may use NVIDIA driver release 418.40 (or later R418), 440.33 (or later R440), 450.51 (or later R450), 460.27 (or later R460), or 470.57 (or later R470). The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#) and [NVIDIA CUDA and Drivers Support](#).

## GPU Requirements

Release 22.01 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 22.01 is based on [Kaldi 26b9f648](#).
- ▶ Ubuntu 20.04 with December 2021 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
22.01	20.04	<a href="#">NVIDIA CUDA 11.6.0</a>	<a href="#">5.5</a> up to commit 26b9f648	<a href="#">TensorRT 8.2.2</a>
<a href="#">21.12</a>		<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2</a> with <a href="#">cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT	
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>	
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777		
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>	
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae		
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c		
<a href="#">21.03</a>			<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>			<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>			<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>	
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837		
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>	
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36		
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc		
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>	
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>	
<a href="#">20.02</a>			<a href="#">5.5</a>		
<a href="#">20.01</a>					
<a href="#">19.12</a>					<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>					

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

None.

---

# Chapter 30. Kaldi Release 21.12

The NVIDIA container image for Kaldi, release 21.12 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA 11.5.0](#)
- ▶ [cuBLAS 11.7.3.1](#)
- ▶ [NVIDIA cuDNN 8.3.1.22](#)
- ▶ [NVIDIA NCCL 2.11.4](#) (optimized for [NVLink™](#))
- ▶ [rdma-core 36.0](#)
- ▶ [OpenMPI 4.1.1+](#)
- ▶ [OpenUCX 1.11.0rc1](#)
- ▶ [GDRCopy 2.3](#)
- ▶ [NVIDIA HPC-X 2.9](#)
- ▶ [Nsight Systems 2021.3.2.4](#)
- ▶ [TensorRT 8.2.1.8](#)
- ▶ [SHARP 2.5](#)
- ▶ [DALI 1.8](#)
- ▶ [DLProf 1.8.0](#)

## Driver Requirements

Release 21.12 is based on [NVIDIA CUDA 11.5.0](#), which requires [NVIDIA Driver](#) release 495 or later. However, if you are running on a Data Center GPU (for example, T4 or any other Tesla board), you may use NVIDIA driver release 418.40 (or later R418), 440.33 (or later R440), 450.51 (or later R450), 460.27 (or later R460), or 470.57 (or later R470). The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#) and [NVIDIA CUDA and Drivers Support](#).

## GPU Requirements

Release 21.12 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 21.12 is based on [Kaldi c6772ba4](#).
- ▶ Ubuntu 20.04 with November 2021 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
21.12	20.04	<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit	<a href="#">TensorRT 8.2.1.8</a>
<a href="#">21.11</a>			<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2</a> with <a href="#">cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				
<a href="#">19.11</a>				
<a href="#">19.10</a>				
<a href="#">19.09</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.08</a>			<a href="#">TensorRT 5.1.5</a>	

## Known Issues

None.

---

# Chapter 31. Kaldi Release 21.11

The NVIDIA container image for Kaldi, release 21.11 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA 11.5.0](#)
- ▶ [cuBLAS 11.7.3.1](#)
- ▶ [NVIDIA cuDNN 8.3.0.96](#)
- ▶ [NVIDIA NCCL 2.11.4](#) (optimized for [NVLink™](#))
- ▶ [rdma-core 36.0](#)
- ▶ [OpenMPI 4.1.1+](#)
- ▶ [OpenUCX 1.11.0rc1](#)
- ▶ [GDRCopy 2.3](#)
- ▶ [NVIDIA HPC-X 2.9](#)
- ▶ [Nsight Systems 2021.3.2.4](#)
- ▶ [TensorRT 8.0.3.4](#)
- ▶ [SHARP 2.5](#)
- ▶ [DALI 1.7](#)
- ▶ [DLProf 1.7.0](#)

## Driver Requirements

Release 21.11 is based on [NVIDIA CUDA 11.5.0](#), which requires [NVIDIA Driver](#) release 495 or later. However, if you are running on a Data Center GPU (for example, T4 or any other Tesla board), you may use NVIDIA driver release 418.40 (or later R418), 440.33 (or later R440), 450.51 (or later R450), 460.27 (or later R460), or 470.57 (or later R470). The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#) and [NVIDIA CUDA and Drivers Support](#).

## GPU Requirements

Release 21.11 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 21.11 is based on [Kaldi c6772ba4](#).
- ▶ Ubuntu 20.04 with October 2021 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">21.11</a>	20.04	<a href="#">NVIDIA CUDA 11.5.0</a>	<a href="#">5.5</a> up to commit c6772ba4	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.10</a>		<a href="#">NVIDIA CUDA 11.4.2</a> with <a href="#">cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e	
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				
<a href="#">19.11</a>				
<a href="#">19.10</a>				
<a href="#">19.09</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.08</a>			<a href="#">TensorRT 5.1.5</a>	

## Known Issues

None.

---

# Chapter 32. Kaldi Release 21.10

The NVIDIA container image for Kaldi, release 21.10 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA 11.4.2](#) with [cuBLAS 11.6.5.2](#)
- ▶ [NVIDIA cuDNN 8.2.4.15](#)
- ▶ [NVIDIA NCCL 2.11.4](#) (optimized for [NVLink™](#))
- ▶ [rdma-core 36.0](#)
- ▶ [OpenMPI 4.1.1+](#)
- ▶ [OpenUCX 1.11.0rc1](#)
- ▶ [GDRCopy 2.3](#)
- ▶ [NVIDIA HPC-X 2.9](#)
- ▶ [Nsight Systems 2021.3.2.4](#)
- ▶ [TensorRT 8.0.3.4](#)
- ▶ [SHARP 2.5](#)
- ▶ [DALI 1.6](#)
- ▶ [DLProf 1.6.0](#)

## Driver Requirements

Release 21.10 is based on [NVIDIA CUDA 11.4.2](#) with [cuBLAS 11.6.5.2](#), which requires [NVIDIA Driver](#) release 470 or later. However, if you are running on Data Center GPUs (formerly Tesla), for example, T4, you may use NVIDIA driver release 418.40 (or later R418), 440.33 (or later R440), 450.51 (or later R450), or 460.27 (or later R460). The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#) and [NVIDIA CUDA and Drivers Support](#).

## GPU Requirements

Release 21.10 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 21.10 is based on [Kaldi 2897564e](#).
- ▶ Ubuntu 20.04 with September 2021 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
21.10	20.04	<a href="#">NVIDIA CUDA 11.4.2</a> with <a href="#">cuBLAS 11.6.5.2</a>	<a href="#">5.5</a> up to commit 2897564e	<a href="#">TensorRT 8.0.3.4</a>
<a href="#">21.09</a>		<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT	
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae		
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c		
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca		<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf		<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a		<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>	
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837		
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>	
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36		
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc		
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>	
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>	
<a href="#">20.02</a>			<a href="#">5.5</a>		
<a href="#">20.01</a>					
<a href="#">19.12</a>					<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>					
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>			
<a href="#">19.09</a>					
<a href="#">19.08</a>	<a href="#">TensorRT 5.1.5</a>				

## Known Issues

None.

---

# Chapter 33. Kaldi Release 21.09

The NVIDIA container image for Kaldi, release 21.09 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA 11.4.2](#)
- ▶ [cuBLAS 11.6.1.51](#)
- ▶ [NVIDIA cuDNN 8.2.4.15](#)
- ▶ [NVIDIA NCCL 2.11.4](#) (optimized for [NVLink™](#))
- ▶ [rdma-core 36.0](#)
- ▶ [OpenMPI 4.1.1+](#)
- ▶ [OpenUCX 1.11.0rc1](#)
- ▶ [GDRCopy 2.3](#)
- ▶ [NVIDIA HPC-X 2.9](#)
- ▶ [Nsight Systems 2021.3.1.57](#)
- ▶ [TensorRT 8.0.3](#)

## Driver Requirements

Release 21.09 is based on [NVIDIA CUDA 11.4.2](#), which requires [NVIDIA Driver](#) release 470 or later. However, if you are running on Data Center GPUs (formerly Tesla), for example, T4, you may use NVIDIA driver release 418.40 (or later R418), 440.33 (or later R440), 450.51 (or later R450), or 460.27 (or later R460). The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#) and [NVIDIA CUDA and Drivers Support](#).

## GPU Requirements

Release 21.09 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 21.09 is based on [Kaldi 40573bb](#).
- ▶ Ubuntu 20.04 with August 2021 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
21.09	20.04	<a href="#">NVIDIA CUDA 11.4.2</a>	<a href="#">5.5</a> up to commit 40573bb	<a href="#">TensorRT 8.0.3</a>
<a href="#">21.08</a>		<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	
<a href="#">21.04</a>				<a href="#">5.5</a> up to commit ec554a7c

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>	
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>	
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>	
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>	
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837		
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>	
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36		
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc		
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>	
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>	
<a href="#">20.02</a>					<a href="#">5.5</a>
<a href="#">20.01</a>					
<a href="#">19.12</a>					<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>					
<a href="#">19.10</a>			<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>					
<a href="#">19.08</a>			<a href="#">TensorRT 5.1.5</a>		

## Known Issues

None.

---

# Chapter 34. Kaldi Release 21.08

The NVIDIA container image for Kaldi, release 21.08 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA 11.4.1](#)
- ▶ [cuBLAS 11.5.4](#)
- ▶ [NVIDIA cuDNN 8.2.2.26](#)
- ▶ [NVIDIA NCCL 2.10.3](#) (optimized for [NVLink™](#))
- ▶ [rdma-core 36.0](#)
- ▶ [OpenMPI 4.1.1+](#)
- ▶ [OpenUCX 1.11.0rc1](#)
- ▶ [GDRCopy 2.2](#)
- ▶ [NVIDIA HPC-X 2.9](#)
- ▶ [Nsight Systems 2021.2.4.12](#)
- ▶ [TensorRT 8.0.1.6](#)

## Driver Requirements

Release 21.08 is based on [NVIDIA CUDA 11.4.1](#), which requires [NVIDIA Driver](#) release 470 or later. However, if you are running on Data Center GPUs (formerly Tesla), for example, T4, you may use NVIDIA driver release 418.40 (or later R418), 440.33 (or later R440), 450.51 (or later R450), or 460.27 (or later R460). The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#) and [NVIDIA CUDA and Drivers Support](#).

## GPU Requirements

Release 21.08 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 21.08 is based on [Kaldi 851a0a79](#).
- ▶ Ubuntu 20.04 with July 2021 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
21.08	20.04	<a href="#">NVIDIA CUDA 11.4.1</a>	<a href="#">5.5</a> up to commit 851a0a79	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.07</a>		<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c	
<a href="#">21.03</a>			<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				
<a href="#">19.11</a>				
<a href="#">19.10</a>	<a href="#">NVIDIA CUDA 10.1.243</a>			
<a href="#">19.09</a>				
<a href="#">19.08</a>			<a href="#">TensorRT 5.1.5</a>	

## Known Issues

None.

---

# Chapter 35. Kaldi Release 21.07

The NVIDIA container image for Kaldi, release 21.07 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA 11.4.0](#)
- ▶ [cuBLAS 11.5.2.43](#)
- ▶ [NVIDIA cuDNN 8.2.2.26](#)
- ▶ [NVIDIA NCCL 2.10.3](#) (optimized for [NVLink™](#))
- ▶ [rdma-core 32.1](#)
- ▶ [OpenMPI 4.1.1rc1](#)
- ▶ [OpenUCX 1.10.1](#)
- ▶ [GDRCopy 2.2](#)
- ▶ [NVIDIA HPC-X 2.8.2rc3](#)
- ▶ [Nsight Compute 2021.1.0.18](#)
- ▶ [Nsight Systems 2021.2.4.12](#)
- ▶ [TensorRT 8.0.1.6](#)

## Driver Requirements

Release 21.07 is based on [NVIDIA CUDA 11.4.0](#), which requires [NVIDIA Driver](#) release 470 or later. However, if you are running on Data Center GPUs (formerly Tesla), for example, T4, you may use NVIDIA driver release 418.40 (or later R418), 440.33 (or later R440), 450.51 (or later R450), or 460.27 (or later R460). The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#) and [NVIDIA CUDA and Drivers Support](#).

## GPU Requirements

Release 21.07 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 21.07 is based on [Kaldi 9bca2777](#).
- ▶ Ubuntu 20.04 with June 2021 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
21.07	20.04	<a href="#">NVIDIA CUDA 11.4.0</a>	<a href="#">5.5</a> up to commit 9bca2777	<a href="#">TensorRT 8.0.1.6</a>
<a href="#">21.06</a>		<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5 up to commit 7905fe5a</a>	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5 up to commit edf6963c</a>	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5 up to commit 49dab837</a>	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5 up to commit 8372cc6a</a>	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5 up to commit f88d5a36</a>	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5 up to commit 67db30cc</a>	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5 up to commit 2c7e78f</a>	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5 up to commit da71f301</a>	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				
<a href="#">19.11</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>			<a href="#">TensorRT 5.1.5</a>	

## Known Issues

The 21.07 release includes libsystemd and libudev versions that have a known vulnerability that was discovered late in our QA process. See [CVE-2021-33910](#) for details. This will be fixed in the next release.

---

# Chapter 36. Kaldi Release 21.06

The NVIDIA container image for Kaldi, release 21.06 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA 11.3.1](#)
- ▶ [cuBLAS 11.5.1.109](#)
- ▶ [NVIDIA cuDNN 8.2.1](#)
- ▶ [NVIDIA NCCL 2.9.9](#) (optimized for [NVLink™](#))
- ▶ [rdma-core 32.1](#)
- ▶ [OpenMPI 4.1.1rc1](#)
- ▶ OpenUCX 1.10.1
- ▶ GDRCopy 2.2
- ▶ NVIDIA HPC-X 2.8.2rc3
- ▶ [Nsight Compute 2021.1.0.18](#)
- ▶ [Nsight Systems 2021.2.1.58](#)
- ▶ [TensorRT 7.2.3.4](#)

## Driver Requirements

Release 21.06 is based on [NVIDIA CUDA 11.3.1](#), which requires [NVIDIA Driver](#) release 465.19.01 or later. However, if you are running on Data Center GPUs (formerly Tesla), for example, T4, you may use NVIDIA driver release 418.40 (or later R418), 440.33 (or later R440), 450.51 (or later R450), or 460.27 (or later R460). The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#) and [NVIDIA CUDA and Drivers Support](#).

## GPU Requirements

Release 21.06 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 21.06 is based on [Kaldi b1786aa5](#).
- ▶ Ubuntu 20.04 with May 2021 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">21.06</a>	20.04	<a href="#">NVIDIA CUDA 11.3.1</a>	<a href="#">5.5</a> up to commit b1786aa5	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.05</a>		<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

There are no known issues in this release.

---

# Chapter 37. Kaldi Release 21.05

The NVIDIA container image for Kaldi, release 21.05 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA 11.3.0](#)
- ▶ [cuBLAS 11.5.1.101](#)
- ▶ [NVIDIA cuDNN 8.2.0.51](#)
- ▶ [NVIDIA NCCL 2.9.8](#) (optimized for [NVLink™](#))
- ▶ [rdma-core 32.1](#)
- ▶ [OpenMPI 4.1.1rc1](#)
- ▶ [OpenUCX 1.10.0](#)
- ▶ [GDRCopy 2.2](#)
- ▶ [NVIDIA HPC-X 2.8.2rc3](#)
- ▶ [Nsight Compute 2021.1.0.18](#)
- ▶ [Nsight Systems 2021.1.3.14](#)
- ▶ [TensorRT 7.2.3.4](#)

## Driver Requirements

Release 21.05 is based on [NVIDIA CUDA 11.3.0](#), which requires [NVIDIA Driver](#) release 465.19.01 or later. However, if you are running on Data Center GPUs (formerly Tesla), for example, T4, you may use NVIDIA driver release 418.40 (or later R418), 440.33 (or later R440), 450.51 (or later R450), or 460.27 (or later R460). The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#) and [NVIDIA CUDA and Drivers Support](#).

## GPU Requirements

Release 21.05 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 21.05 is based on [Kaldi e92970ae](#).
- ▶ Ubuntu 20.04 with April 2021 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
21.05	20.04	<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit e92970ae	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.04</a>			<a href="#">5.5</a> up to commit ec554a7c	
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

There are no known issues in this release.

---

# Chapter 38. Kaldi Release 21.04

The NVIDIA container image for Kaldi, release 21.04 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA 11.3.0](#)
- ▶ [cuBLAS 11.5.1.101](#)
- ▶ [NVIDIA cuDNN 8.2.0.41](#)
- ▶ [NVIDIA NCCL 2.9.6](#) (optimized for [NVLink™](#))
- ▶ [rdma-core 32.1](#)
- ▶ [OpenMPI 4.1.1rc1](#)
- ▶ [OpenUCX 1.10.0](#)
- ▶ [GDRCopy 2.2](#)
- ▶ [NVIDIA HPC-X 2.8.2rc3](#)
- ▶ [Nsight Compute 2021.1.0.18](#)
- ▶ [Nsight Systems 2021.1.3.14](#)
- ▶ [TensorRT 7.2.3.4](#)

## Driver Requirements

Release 21.04 is based on [NVIDIA CUDA 11.3.0](#), which requires [NVIDIA Driver](#) release 465.19.01 or later. However, if you are running on Data Center GPUs (formerly Tesla), for example, T4, you may use NVIDIA driver release 418.40 (or later R418), 440.33 (or later R440), 450.51 (or later R450), or 460.27 (or later R460). The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#) and [NVIDIA CUDA and Drivers Support](#).

## GPU Requirements

Release 21.04 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 21.04 is based on [Kaldi ec554a7c](#).
- ▶ Ubuntu 20.04 with March 2021 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">21.04</a>	20.04	<a href="#">NVIDIA CUDA 11.3.0</a>	<a href="#">5.5</a> up to commit <a href="#">ec554a7c</a>	<a href="#">TensorRT 7.2.3.4</a>
<a href="#">21.03</a>		<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit <a href="#">376281ca</a>	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit <a href="#">b78fdcaf</a>	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit <a href="#">7905fe5a</a>	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit <a href="#">edf6963c</a>	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit <a href="#">49dab837</a>	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				<a href="#">TensorRT 5.1.5</a>
<a href="#">19.08</a>				

## Known Issues

There are no known issues in this release.

---

# Chapter 39. Kaldi Release 21.03

The NVIDIA container image for Kaldi, release 21.03 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA 11.2.1](#) including [cuBLAS 11.4.1.1026](#)
- ▶ [NVIDIA cuDNN 8.1.1](#)
- ▶ [NVIDIA NCCL 2.8.4](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED 5.1](#)
- ▶ [OpenMPI 4.0.5](#)
- ▶ [Nsight Compute 2020.3.0.18](#)
- ▶ [Nsight Systems 2020.4.3.7](#)
- ▶ [TensorRT 7.2.2.3](#)

## Driver Requirements

Release 21.03 is based on [NVIDIA CUDA 11.2.1](#), which requires [NVIDIA Driver](#) release 460.32.03 or later. However, if you are running on Data Center GPUs (formerly Tesla), for example, T4, you may use NVIDIA driver release 418.40 (or later R418), 440.33 (or later R440), 450.51 (or later R450). The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#) and [NVIDIA CUDA and Drivers Support](#).

## GPU Requirements

Release 21.03 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 21.03 is based on [Kaldi 376281ca](#).
- ▶ The latest version of [NVIDIA CUDA 11.2.1](#) including [cuBLAS 11.4.1.1026](#)
- ▶ The latest version of [NVIDIA cuDNN 8.1.1](#)
- ▶ Ubuntu 20.04 with February 2021 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
21.03	20.04	<a href="#">NVIDIA CUDA 11.2.1</a>	<a href="#">5.5</a> up to commit 376281ca	<a href="#">TensorRT 7.2.2.3</a>
<a href="#">21.02</a>		<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit b78fdcaf	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

### Known Issues

There are no known issues in this release.

---

# Chapter 40. Kaldi Release 21.02

The NVIDIA container image for Kaldi, release 21.02 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA 11.2.0](#) including [cuBLAS 11.3.1](#)
- ▶ [NVIDIA cuDNN 8.1.0](#)
- ▶ [NVIDIA NCCL 2.8.4](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED 5.1](#)
- ▶ [OpenMPI 4.0.5](#)
- ▶ [Nsight Compute 2020.3.0.18](#)
- ▶ [Nsight Systems 2020.4.3.7](#)
- ▶ [TensorRT 7.2.2](#)

## Driver Requirements

Release 21.02 is based on [NVIDIA CUDA 11.2.0](#), which requires [NVIDIA Driver](#) release 460.27.04 or later. However, if you are running on Data Center GPUs (formerly Tesla), for example, T4, you may use NVIDIA driver release 418.40 (or later R418), 440.33 (or later R440), 450.51 (or later R450). The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#) and [NVIDIA CUDA and Drivers Support](#).

## GPU Requirements

Release 21.02 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 21.02 is based on [Kaldi b78fdcaf](#).
- ▶ The latest version of [NVIDIA CUDA 11.2.0](#) including [cuBLAS 11.3.1](#)
- ▶ The latest version of [NVIDIA cuDNN 8.1.0](#)
- ▶ The latest version of [NVIDIA NCCL 2.8.4](#) (optimized for [NVLink™](#))
- ▶ The latest version of [Nsight Compute 2020.3.0.18](#)
- ▶ The latest version of [Nsight Systems 2020.4.3.7](#)
- ▶ The latest version of [TensorRT 7.2.2.3+cuda11.1.0.024](#)
- ▶ Ubuntu 20.04 with January 2021 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
21.02	20.04	<a href="#">NVIDIA CUDA 11.2.0</a>	<a href="#">5.5</a> up to commit <a href="#">b78fdcaf</a>	<a href="#">TensorRT 7.2.2.3+cuda11.1.0.024</a>
<a href="#">20.12</a>		<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit <a href="#">7905fe5a</a>	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit <a href="#">edf6963c</a>	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit <a href="#">49dab837</a>	
<a href="#">20.09</a>			<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit <a href="#">8372cc6a</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

### Known Issues

There are no known issues in this release.

---

## Chapter 41. Kaldi Release 21.01

The NVIDIA container image release for Kaldi 21.01 has been canceled. The next release will be the 21.02 release which is expected to be released at the end of February.

---

# Chapter 42. Kaldi Release 20.12

The NVIDIA container image for Kaldi, release 20.12 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 20.04](#) including [Python 3.8](#)
- ▶ [NVIDIA CUDA 11.1.1](#) including [cuBLAS 11.3.0](#)
- ▶ [NVIDIA cuDNN 8.0.5](#)
- ▶ [NVIDIA NCCL 2.8.3](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED 5.1](#)
- ▶ [OpenMPI 4.0.5](#)
- ▶ [Nsight Compute 2020.2.1.8](#)
- ▶ [Nsight Systems 2020.3.4.32](#)
- ▶ [TensorRT 7.2.2](#)

## Driver Requirements

Release 20.12 is based on [NVIDIA CUDA 11.1.1](#), which requires [NVIDIA Driver](#) release 455 or later. However, if you are running on Tesla (for example, T4 or any other Tesla board), you may use NVIDIA driver release 418.xx, 440.30, or 450.xx. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 20.12 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 20.12 is based on [Kaldi 7905fe5a](#).
- ▶ The latest version of [NVIDIA CUDA 11.1.1](#) including [cuBLAS 11.3.0](#)
- ▶ The latest version of [NVIDIA cuDNN 8.0.5](#)
- ▶ The latest version of [NVIDIA NCCL 2.8.3](#) (optimized for [NVLink™](#))
- ▶ The latest version of [Nsight Compute 2020.2.1.8](#)
- ▶ The latest version of [TensorRT 7.2.2](#)
- ▶ Ubuntu 20.04 with November 2020 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
20.12	20.04	<a href="#">NVIDIA CUDA 11.1.1</a>	<a href="#">5.5</a> up to commit 7905fe5a	<a href="#">TensorRT 7.2.2</a>
<a href="#">20.11</a>	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>		<a href="#">5.5</a> up to commit f88d5a36		
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5 up to commit 2c7e78f</a>	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5 up to commit da71f301</a>	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>	<a href="#">5.5</a>			
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

There are no known issues in this release.

---

# Chapter 43. Kaldi Release 20.11

The NVIDIA container image for Kaldi, release 20.11 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 18.04](#) including [Python 3.6](#)
- ▶ [NVIDIA CUDA 11.1.0](#) including [cuBLAS 11.2.1](#)
- ▶ [NVIDIA cuDNN 8.0.4](#)
- ▶ [NVIDIA NCCL 2.8.2](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED 5.1](#)
- ▶ [OpenMPI 4.0.5](#)
- ▶ [Nsight Compute 2020.2.0.18](#)
- ▶ [Nsight Systems 2020.3.4.32](#)
- ▶ [TensorRT 7.2.1](#)

## Driver Requirements

Release 20.11 is based on [NVIDIA CUDA 11.1.0](#), which requires [NVIDIA Driver](#) release 455 or later. However, if you are running on Tesla (for example, T4 or any other Tesla board), you may use NVIDIA driver release 418.xx, 440.30, or 450.xx. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 20.11 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 20.11 is based on [Kaldi edf6963c](#).
- ▶ The latest version of [NVIDIA NCCL 2.8.2](#)
- ▶ Ubuntu 18.04 with October 2020 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
20.11	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit edf6963c	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.10</a>			<a href="#">5.5</a> up to commit 49dab837	
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA</a>		
<a href="#">19.09</a>		<a href="#">10.1.243</a>		
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

There are no known issues in this release.

---

# Chapter 44. Kaldi Release 20.10

The NVIDIA container image for Kaldi, release 20.10 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 18.04](#) including [Python 3.6](#)
- ▶ [NVIDIA CUDA 11.1.0](#) including [cuBLAS 11.2.1](#)
- ▶ [NVIDIA cuDNN 8.0.4](#)
- ▶ [NVIDIA NCCL 2.7.8](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED](#)
- ▶ [OpenMPI 3.1.6](#)
- ▶ [Nsight Compute 2020.2.0.18](#)
- ▶ [Nsight Systems 2020.3.4.32](#)
- ▶ [TensorRT 7.2.1](#)

## Driver Requirements

Release 20.10 is based on [NVIDIA CUDA 11.1.0](#), which requires [NVIDIA Driver](#) release 455 or later. However, if you are running on Tesla (for example, T4 or any other Tesla board), you may use NVIDIA driver release 418.xx, 440.30, or 450.xx. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 20.10 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 20.10 is based on [Kaldi 49dab837](#).
- ▶ The latest version of [NVIDIA CUDA 11.1.0](#)
- ▶ The latest version of [cuBLAS 11.2.1](#)
- ▶ The latest version of [NVIDIA cuDNN 8.0.4](#)
- ▶ The latest version of [Nsight Compute 2020.2.0.18](#)
- ▶ The latest version of [Nsight Systems 2020.3.4.32](#)
- ▶ The latest version of [TensorRT 7.2.1](#)
- ▶ Ubuntu 18.04 with September 2020 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
20.10	18.04	<a href="#">NVIDIA CUDA 11.1.0</a>	<a href="#">5.5</a> up to commit 49dab837	<a href="#">TensorRT 7.2.1</a>
<a href="#">20.09</a>		<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36	
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a> <a href="#">20.01</a>			<a href="#">5.5</a>	
<a href="#">19.12</a> <a href="#">19.11</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		<a href="#">TensorRT 6.0.1</a>
<a href="#">19.10</a> <a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

### Known Issues

There are no known issues in this release.

---

# Chapter 45. Kaldi Release 20.09

The NVIDIA container image for Kaldi, release 20.09 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 18.04](#) including [Python 3.6](#)
- ▶ [NVIDIA CUDA 11.0.3](#) including [cuBLAS 11.2.0](#)
- ▶ [NVIDIA cuDNN 8.0.4](#)
- ▶ [NVIDIA NCCL 2.7.8](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED](#)
- ▶ [OpenMPI 3.1.6](#)
- ▶ [Nsight Compute 2020.1.2.4](#)
- ▶ [Nsight Systems 2020.3.2.6](#)
- ▶ [TensorRT 7.1.3](#)

## Driver Requirements

Release 20.09 is based on [NVIDIA CUDA 11.0.3](#), which requires [NVIDIA Driver](#) release 450 or later. However, if you are running on Tesla (for example, T4 or any other Tesla board), you may use NVIDIA driver release 418.xx or 440.30. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 20.09 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 20.09 is based on [Kaldi 8372cc6a](#).
- ▶ The latest version of [NVIDIA cuDNN 8.0.4](#)
- ▶ Ubuntu 18.04 with August 2020 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT	
20.09	18.04	<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit 8372cc6a	<a href="#">TensorRT 7.1.3</a>	
<a href="#">20.08</a>			<a href="#">5.5</a> up to commit f88d5a36		
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc		
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>	
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>	
<a href="#">20.02</a>			<a href="#">5.5</a>		
<a href="#">20.01</a>					
<a href="#">19.12</a>					<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>					

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

There are no known issues in this release.

---

# Chapter 46. Kaldi Release 20.08

The NVIDIA container image for Kaldi, release 20.08 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 18.04](#) including [Python 3.6](#)
- ▶ [NVIDIA CUDA 11.0.3](#) including [cuBLAS 11.2.0](#)
- ▶ [NVIDIA cuDNN 8.0.2](#)
- ▶ [NVIDIA NCCL 2.7.8](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED](#)
- ▶ [OpenMPI 3.1.6](#)
- ▶ [Nsight Compute 2020.1.2.4](#)
- ▶ [Nsight Systems 2020.3.2.6](#)
- ▶ [TensorRT 7.1.3](#)

## Driver Requirements

Release 20.08 is based on [NVIDIA CUDA 11.0.3](#), which requires [NVIDIA Driver](#) release 450 or later. However, if you are running on Tesla (for example, T4 or any other Tesla board), you may use NVIDIA driver release 418.xx or 440.30. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 20.08 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 20.08 is based on [Kaldi 67db30cc](#).
- ▶ The latest version of [NVIDIA CUDA 11.0.3](#) including [cuBLAS 11.2.0](#)
- ▶ The latest version of [NVIDIA cuDNN 8.0.2](#)
- ▶ The latest version of [NVIDIA NCCL 2.7.8](#)
- ▶ The latest version of [Nsight Compute 2020.1.2.4](#)
- ▶ The latest version of [Nsight Systems 2020.3.2.6](#)
- ▶ Ubuntu 18.04 with July 2020 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
20.08	18.04	<a href="#">NVIDIA CUDA 11.0.3</a>	<a href="#">5.5</a> up to commit f88d5a36	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.07</a>		<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	
<a href="#">20.06</a>		<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>		<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">19.11</a>				
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

There are no known issues in this release.

---

# Chapter 47. Kaldi Release 20.07

The NVIDIA container image for Kaldi, release 20.07 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 18.04](#) including [Python 3.6](#)
- ▶ [NVIDIA CUDA 11.0.194](#) including [cuBLAS 11.1.0](#)
- ▶ [NVIDIA cuDNN 8.0.1](#)
- ▶ [NVIDIA NCCL 2.7.6](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED](#)
- ▶ [OpenMPI 3.1.6](#)
- ▶ [Nsight Compute 2020.1.1.8](#)
- ▶ [Nsight Systems 2020.3.2.6](#)
- ▶ [TensorRT 7.1.3](#)

## Driver Requirements

Release 20.07 is based on [NVIDIA CUDA 11.0.194](#), which requires [NVIDIA Driver](#) release 450 or later. However, if you are running on Tesla (for example, T4 or any other Tesla board), you may use NVIDIA driver release 418.xx or 440.30. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 20.07 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 20.07 is based on [Kaldi 67db30cc](#).
- ▶ The latest version of [NVIDIA CUDA 11.0.194](#) including [cuBLAS 11.1.0](#)
- ▶ The latest version of [NVIDIA NCCL 2.7.6](#)
- ▶ The latest version of [Nsight Compute 2020.1.1](#)
- ▶ The latest version of [Nsight Systems 2020.3.2](#)
- ▶ The latest version of [TensorRT 7.1.3](#)
- ▶ Ubuntu 18.04 with June 2020 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
20.07	18.04	<a href="#">NVIDIA CUDA 11.0.194</a>	<a href="#">5.5</a> up to commit 67db30cc	<a href="#">TensorRT 7.1.3</a>
<a href="#">20.06</a>				<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>			<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>				<a href="#">5.5</a>
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">19.10</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

There are no known issues in this release.

---

# Chapter 48. Kaldi Release 20.06

The NVIDIA container image for Kaldi, release 20.06 is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 18.04](#) including [Python 3.6](#)
- ▶ [NVIDIA CUDA 11.0.167](#) including [cuBLAS 11.1.0](#)
- ▶ [NVIDIA cuDNN 8.0.1](#)
- ▶ [NVIDIA NCCL 2.7.5](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED](#)
- ▶ [OpenMPI 3.1.6](#)
- ▶ [Nsight Compute 2020.1.0](#)
- ▶ [Nsight Systems 2020.2.5](#)
- ▶ [TensorRT 7.1.1](#)

## Driver Requirements

Release 20.06 is based on [NVIDIA CUDA 11.0.167](#), which requires [NVIDIA Driver](#) release 450 or later. However, if you are running on Tesla (for example, T4 or any other Tesla board), you may use NVIDIA driver release 418.xx or 440.30. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 20.06 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, Turing, and Ampere Architecture GPU families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 20.06 is based on [Kaldi 2c7e78f](#).
- ▶ The latest version of [NVIDIA CUDA 11.0.167](#) including [cuBLAS 11.1.0](#)
- ▶ The latest version of [NVIDIA cuDNN 8.0.1](#)
- ▶ The latest version of [NVIDIA NCCL 2.7.5](#)
- ▶ The latest version of [Nsight Compute 2020.1.0](#)
- ▶ The latest version of [Nsight Systems 2020.2.5](#)
- ▶ The latest version of [TensorRT 7.1.2](#)
- ▶ Ubuntu 18.04 with May 2020 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
20.06	18.04	<a href="#">NVIDIA CUDA 11.0.167</a>	<a href="#">5.5</a> up to commit 2c7e78f	<a href="#">TensorRT 7.1.2</a>
<a href="#">20.03</a>			<a href="#">5.5</a> up to commit da71f301	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>			<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>				
<a href="#">19.09</a>			<a href="#">NVIDIA CUDA 10.1.243</a>	

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

There are no known issues in this release.

---

# Chapter 49. Kaldi Release 20.03

The NVIDIA container image for Kaldi, release 20.03, is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 18.04](#) including [Python 3.6](#)
- ▶ [NVIDIA CUDA 10.2.89](#) including [cuBLAS 10.2.2.89](#)
- ▶ [NVIDIA cuDNN 7.6.5](#)
- ▶ [NVIDIA NCCL 2.6.3](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED](#)
- ▶ [OpenMPI 3.1.4](#)
- ▶ [Nsight Compute 2019.5.0](#)
- ▶ [Nsight Systems 2020.1.1](#)
- ▶ [TensorRT 7.0.0](#)

## Driver Requirements

Release 20.03 is based on [NVIDIA CUDA 10.2.89](#), which requires [NVIDIA Driver](#) release 440.33.01. However, if you are running on Tesla (for example, T4 or any other Tesla board), you may use NVIDIA driver release 396, 384.111+, 410, 418.xx or 440.30. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 20.03 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, and Turing families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 20.03 is based on [Kaldi da71f301](#).
- ▶ Ubuntu 18.04 with February 2020 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare_data.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` which runs inference on the trained model and data. Setting the environment variable `ONLINE=1` will make the benchmark script run in online mode.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
20.03	18.04	<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">da71f301</a>	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.02</a>	16.04		<a href="#">5.5</a>	
<a href="#">20.01</a>				
<a href="#">19.12</a>				<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>				
<a href="#">19.09</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		
<a href="#">19.08</a>				

## Known Issues

There are no known issues in this release.

---

# Chapter 50. Kaldi Release 20.02

The NVIDIA container image for Kaldi, release 20.02, is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 18.04](#) including [Python 3.6](#)
- ▶ [NVIDIA CUDA 10.2.89](#) including [cuBLAS 10.2.2.89](#)
- ▶ [NVIDIA cuDNN 7.6.5](#)
- ▶ [NVIDIA NCCL 2.5.6](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED](#)
- ▶ [OpenMPI 3.1.4](#)
- ▶ [Nsight Compute 2019.5.0](#)
- ▶ [Nsight Systems 2020.1.1](#)
- ▶ [TensorRT 7.0.0](#)

## Driver Requirements

Release 20.02 is based on [NVIDIA CUDA 10.2.89](#), which requires [NVIDIA Driver](#) release 440.33.01. However, if you are running on Tesla (for example, T4 or any other Tesla board), you may use NVIDIA driver release 396, 384.111+, 410, 418.xx or 440.30. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 20.02 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, and Turing families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 20.02 is based on [Kaldi 5.5](#).
- ▶ Latest version of [Nsight Systems 2020.1.1](#)
- ▶ Ubuntu 18.04 with January 2020 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` and `/workspace/nvidia-examples/librispeech/run_multigpu_benchmark.sh` which run inference on the trained model and data.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
20.02	18.04	<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a>	<a href="#">TensorRT 7.0.0</a>
<a href="#">20.01</a>	16.04			<a href="#">TensorRT 6.0.1</a>
<a href="#">19.12</a>				
<a href="#">19.11</a>				
<a href="#">19.10</a>				
<a href="#">19.09</a>		<a href="#">NVIDIA CUDA 10.1.243</a>		<a href="#">TensorRT 5.1.5</a>
<a href="#">19.08</a>				

## Known Issues

There are no known issues in this release.

---

# Chapter 51. Kaldi Release 20.01

The NVIDIA container image for Kaldi, release 20.01, is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 18.04](#) including [Python 3.6](#)
- ▶ [NVIDIA CUDA 10.2.89](#) including [cuBLAS 10.2.2.89](#)
- ▶ [NVIDIA cuDNN 7.6.5](#)
- ▶ [NVIDIA NCCL 2.5.6](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED](#)
- ▶ [OpenMPI 3.1.4](#)
- ▶ [Nsight Compute 2019.5.0](#)
- ▶ [Nsight Systems 2019.6.1](#)
- ▶ [TensorRT 7.0.0](#)

## Driver Requirements

Release 20.01 is based on [NVIDIA CUDA 10.2.89](#), which requires [NVIDIA Driver](#) release 440.33.01. However, if you are running on Tesla (for example, T4 or any other Tesla board), you may use NVIDIA driver release 396, 384.111+, 410, 418.xx or 440.30. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 20.01 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, and Turing families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 20.01 is based on [Kaldi 5.5](#).
  - ▶ Commit hash: ba92f608
  - ▶ This contains minor bug and performance fixes.
- ▶ Latest version of [TensorRT 7.0.0](#)
- ▶ Ubuntu 18.04 with December 2019 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` and `/workspace/nvidia-examples/librispeech/run_multigpu_benchmark.sh` which run inference on the trained model and data.

## NVIDIA Kaldi Container Versions

The following table shows what versions of Ubuntu, CUDA, Kaldi, and TensorRT are supported in each of the NVIDIA containers for Kaldi. For older container versions, refer to the [Frameworks Support Matrix](#).

Container Version	Ubuntu	CUDA Toolkit	Kaldi	TensorRT
20.01	18.04	<a href="#">NVIDIA CUDA 10.2.89</a>	<a href="#">5.5</a>	<a href="#">TensorRT 7.0.0</a>
<a href="#">19.12</a>	16.04			<a href="#">TensorRT 6.0.1</a>
<a href="#">19.11</a>				
<a href="#">19.10</a>				
<a href="#">19.09</a>				
<a href="#">19.08</a>				<a href="#">TensorRT 5.1.5</a>

## Known Issues

There are no known issues in this release.

---

# Chapter 52. Kaldi Release 19.12

The NVIDIA container image for Kaldi, release 19.12, is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 18.04](#) including [Python 3.6](#)
- ▶ [NVIDIA CUDA 10.2.89](#) including [cuBLAS 10.2.2.89](#)
- ▶ [NVIDIA cuDNN 7.6.5](#)
- ▶ [NVIDIA NCCL 2.5.6](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED](#)
- ▶ [OpenMPI 3.1.4](#)
- ▶ [Nsight Compute 2019.5.0](#)
- ▶ [Nsight Systems 2019.6.1](#)
- ▶ [TensorRT 6.0.1](#)

## Driver Requirements

Release 19.12 is based on [NVIDIA CUDA 10.2.89](#), which requires [NVIDIA Driver](#) release 440.30. However, if you are running on Tesla (for example, T4 or any other Tesla board), you may use NVIDIA driver release 396, 384.111+, 410, 418.xx or 440.30. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 19.12 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, and Turing families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 19.12 is based on [Kaldi 5.5](#).
- ▶ Latest version of [Nsight Systems 2019.6.1](#)
- ▶ Ubuntu 18.04 with November 2019 updates.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` and `/workspace/nvidia-examples/librispeech/run_multigpu_benchmark.sh` which run inference on the trained model and data.

## Known Issues

There are no known issues in this release.

---

# Chapter 53. Kaldi Release 19.11

The NVIDIA container image for Kaldi, release 19.11, is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 18.04](#) including [Python 3.6](#)
- ▶ [NVIDIA CUDA 10.2.89](#) including [cuBLAS 10.2.2.89](#)
- ▶ [NVIDIA cuDNN 7.6.5](#)
- ▶ [NVIDIA NCCL 2.5.6](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED](#)
- ▶ [OpenMPI 3.1.4](#)
- ▶ [Nsight Compute 2019.5.0](#)
- ▶ [Nsight Systems 2019.5.2](#)
- ▶ [TensorRT 6.0.1](#)

## Driver Requirements

Release 19.11 is based on [NVIDIA CUDA 10.2.89](#), which requires [NVIDIA Driver](#) release 440.30. However, if you are running on Tesla (for example, T4 or any other Tesla board), you may use NVIDIA driver release 396, 384.111+, 410 or 418.xx. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 19.11 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, and Turing families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 19.11 is based on [Kaldi 5.5](#).
- ▶ Latest version of [Python 3.6](#)
- ▶ Latest version of [NVIDIA CUDA 10.2.89](#) including [cuBLAS 10.2.2.89](#)
- ▶ Latest version of [NVIDIA cuDNN 7.6.5](#)
- ▶ Latest version of [NVIDIA NCCL 2.5.6](#)
- ▶ Latest version of [Nsight Compute 2019.5.0](#)
- ▶ Latest version of [Nsight Systems 2019.5.2](#)
- ▶ Pulled in the latest Kaldi which includes bug fixes for stability and accuracy.
- ▶ Ubuntu 18.04 with October 2019 updates

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` and `/workspace/nvidia-examples/librispeech/run_multigpu_benchmark.sh` which run inference on the trained model and data.

## Known Issues

There are no known issues in this release.

---

# Chapter 54. Kaldi Release 19.10

The NVIDIA container image for Kaldi, release 19.10, is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 18.04](#) including [Python 3.5](#)
- ▶ [NVIDIA CUDA 10.1.243](#) including [cuBLAS 10.2.1.243](#)
- ▶ [NVIDIA cuDNN 7.6.4](#)
- ▶ [NVIDIA NCCL 2.4.8](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED](#)
- ▶ [OpenMPI 3.1.4](#)
- ▶ [Nsight Compute 2019.4.0](#)
- ▶ [Nsight Systems 2019.5.1](#)
- ▶ [TensorRT 6.0.1](#)

## Driver Requirements

Release 19.10 is based on [NVIDIA CUDA 10.1.243](#), which requires [NVIDIA Driver](#) release 418.xx. However, if you are running on Tesla (for example, T4 or any other Tesla board), you may use NVIDIA driver release 396, 384.111+ or 410. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 19.10 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, and Turing families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 19.10 is based on [Kaldi 5.5](#).
- ▶ Latest version of [NVIDIA cuDNN 7.6.4](#)
- ▶ Latest version of [Nsight Systems 2019.5.1](#)
- ▶ Added an ASPIRE example in the `/workspace/nvidia-examples/aspire` directory.
- ▶ Ubuntu 18.04 with September 2019 updates

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` and `/workspace/nvidia-examples/librispeech/run_multigpu_benchmark.sh` which run inference on the trained model and data.

## Known Issues

There are no known issues in this release.

---

# Chapter 55. Kaldi Release 19.09

The NVIDIA container image for Kaldi, release 19.09, is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 18.04](#) including [Python 3.5](#)
- ▶ [NVIDIA CUDA 10.1.243](#) including [cuBLAS 10.2.1.243](#)
- ▶ [NVIDIA cuDNN 7.6.3](#)
- ▶ [NVIDIA NCCL 2.4.8](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED](#)
- ▶ [OpenMPI 3.1.4](#)
- ▶ [Nsight Compute 2019.4.0](#)
- ▶ [Nsight Systems 2019.4.2](#)
- ▶ [TensorRT 6.0.1](#)

## Driver Requirements

Release 19.09 is based on [NVIDIA CUDA 10.1.243](#), which requires [NVIDIA Driver](#) release 418.xx. However, if you are running on Tesla (for example, T4 or any other Tesla board), you may use NVIDIA driver release 396, 384.111+ or 410. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 19.09 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, and Turing families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 19.09 is based on [Kaldi 5.5](#).
- ▶ Latest version of [NVIDIA cuDNN 7.6.3](#)
- ▶ Latest version of [TensorRT 6.0.1](#)
- ▶ Latest versions of [Nsight Compute 2019.4.0](#) and [Nsight Systems 2019.4.2](#)
- ▶ Improved benchmark scripts, for example:
  - ▶ Unified `run_benchmark.sh` script so that it can now run single and multi-GPU.
  - ▶ Versioned output so that it is not overwritten.
  - ▶ Produced human readable transcriptions.
  - ▶ Added support for automatic segmentation.
  - ▶ Automatically create dataset when `wav.scp` does not exist.

See `NVREADME.md` for more details.

- ▶ Ubuntu 18.04 with August 2019 updates

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` and `/workspace/nvidia-examples/librispeech/run_multigpu_benchmark.sh` which run inference on the trained model and data.

## Known Issues

There are no known issues in this release.

---

# Chapter 56. Kaldi Release 19.08

The NVIDIA container image for Kaldi, release 19.08, is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 18.04](#) including [Python 3.5](#)
- ▶ [NVIDIA CUDA 10.1.243](#) including [cuBLAS 10.2.1.243](#)
- ▶ [NVIDIA cuDNN 7.6.2](#)
- ▶ [NVIDIA NCCL 2.4.8](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED +4.0](#)
- ▶ [OpenMPI 3.1.4](#)
- ▶ [Nsight Compute 10.1.168](#)
- ▶ [Nsight Systems 2019.3.7.9](#)
- ▶ [TensorRT 5.1.5](#)

## Driver Requirements

Release 19.08 is based on [NVIDIA CUDA 10.1.243](#), which requires [NVIDIA Driver](#) release 418.87. However, if you are running on Tesla (Tesla V100, Tesla P4, Tesla P40, or Tesla P100), you may use NVIDIA driver release 384.111+ or 410. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 19.08 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, and Turing families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 19.08 is based on [Kaldi 5.5](#).
- ▶ Latest version of [NVIDIA CUDA 10.1.243](#) including [cuBLAS 10.2.1.243](#)
- ▶ Latest version of [NVIDIA cuDNN 7.6.2](#)
- ▶ Latest version of [NVIDIA NCCL 2.4.8](#)
- ▶ Latest version of [Nsight Systems 2019.3.7.9](#)
- ▶ Latest version of [MLNX\\_OFED +4.0](#)
- ▶ Latest version of [OpenMPI 3.1.4](#)
- ▶ Ubuntu 18.04 with July 2019 updates

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` and `/workspace/nvidia-examples/librispeech/run_multigpu_benchmark.sh` which run inference on the trained model and data.

## Known Issues

There are no known issues in this release.

---

# Chapter 57. Kaldi Release 19.07

The NVIDIA container image for Kaldi, release 19.07, is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 18.04](#) including [Python 3.5](#)
- ▶ [NVIDIA CUDA 10.1.168](#) including [cuBLAS 10.2.0.168](#)
- ▶ [NVIDIA cuDNN 7.6.1](#)
- ▶ [NVIDIA NCCL 2.4.7](#) (optimized for [NVLink™](#))
- ▶ [MLNX\\_OFED +3.4](#)
- ▶ [OpenMPI 3.1.3](#)
- ▶ [TensorRT 5.1.5](#)

## Driver Requirements

Release 19.07 is based on [NVIDIA CUDA 10.1.168](#), which requires [NVIDIA Driver](#) release 418.67. However, if you are running on Tesla (Tesla V100, Tesla P4, Tesla P40, or Tesla P100), you may use NVIDIA driver release 384.111+ or 410. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 19.07 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, and Turing families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 19.07 is based on [Kaldi 5.5](#).
- ▶ Latest version of [NVIDIA cuDNN 7.6.1](#)
- ▶ Latest version of [MLNX\\_OFED +3.4](#)
- ▶ Latest version of [Ubuntu 18.04](#)
- ▶ Added option for GPU feature extraction.
- ▶ Performance improvements.

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` and `/workspace/nvidia-examples/librispeech/run_multigpu_benchmark.sh` which run inference on the trained model and data.

## Known Issues

There are no known issues in this release.

---

# Chapter 58. Kaldi Release 19.06

The NVIDIA container image for Kaldi, release 19.06, is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 16.04](#) including [Python 3.5](#)
- ▶ [NVIDIA CUDA 10.1.168](#) including [cuBLAS 10.2.0.168](#)
- ▶ [NVIDIA cuDNN 7.6.0](#)
- ▶ [NVIDIA NCCL 2.4.7](#) (optimized for [NVLink™](#))
- ▶ [OpenMPI 3.1.3](#)
- ▶ [TensorRT 5.1.5](#)

## Driver Requirements

Release 19.06 is based on [NVIDIA CUDA 10.1.168](#), which requires [NVIDIA Driver](#) release 418.xx. However, if you are running on Tesla (Tesla V100, Tesla P4, Tesla P40, or Tesla P100), you may use NVIDIA driver release 384.111+ or 410. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 19.06 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, and Turing families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 19.06 is based on [Kaldi 5.5](#).

- ▶ Latest version of [NVIDIA CUDA 10.1.168](#) including [cuBLAS 10.2.0.168](#)
- ▶ Latest version of [NVIDIA NCCL 2.4.7](#)
- ▶ Ubuntu 16.04 with May 2019 updates (see Announcements)

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` and `/workspace/nvidia-examples/librispeech/run_multigpu_benchmark.sh` which run inference on the trained model and data.

## Announcements

In the next release, we will no longer support [Ubuntu 16.04](#). Release 19.07 will instead support [Ubuntu 18.04](#).

## Known Issues

There are no known issues in this release.

---

# Chapter 59. Kaldi Release 19.05

The NVIDIA container image for Kaldi, release 19.05, is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 16.04](#) including [Python 3.5](#)
- ▶ [NVIDIA CUDA 10.1 Update 1](#) including [cuBLAS 10.1 Update 1](#)
- ▶ [NVIDIA cuDNN 7.6.0](#)
- ▶ [NVIDIA NCCL 2.4.6](#) (optimized for [NVLink™](#))
- ▶ [OpenMPI 3.1.3](#)
- ▶ [TensorRT 5.1.5](#)

## Driver Requirements

Release 19.05 is based on CUDA 10.1 Update 1, which requires [NVIDIA Driver](#) release 418.xx. However, if you are running on Tesla (Tesla V100, Tesla P4, Tesla P40, or Tesla P100), you may use NVIDIA driver release 384.111+ or 410. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 19.05 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, and Turing families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 19.05 is based on [Kaldi 5.5](#).

- ▶ Latest version of [NVIDIA CUDA 10.1 Update 1](#) including [cuBLAS 10.1 Update 1](#)
- ▶ Latest version of [NVIDIA cuDNN 7.6.0](#)
- ▶ Latest version of [TensorRT 5.1.5](#)
- ▶ Ubuntu 16.04 with April 2019 updates

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` and `/workspace/nvidia-examples/librispeech/run_multigpu_benchmark.sh` which run inference on the trained model and data.

## Known Issues

There are no known issues in this release.

---

# Chapter 60. Kaldi Release 19.04

The NVIDIA container image for Kaldi, release 19.04, is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 16.04](#) including [Python 3.5](#)
- ▶ [NVIDIA CUDA 10.1.105](#) including [cuBLAS 10.1.0.105](#)
- ▶ [NVIDIA cuDNN 7.5.0](#)
- ▶ [NVIDIA NCCL 2.4.6](#) (optimized for [NVLink™](#))
- ▶ [OpenMPI 3.1.3](#)
- ▶ [TensorRT 5.1.2](#)

## Driver Requirements

Release 19.04 is based on CUDA 10.1, which requires [NVIDIA Driver](#) release 418.xx.x +. However, if you are running on Tesla (Tesla V100, Tesla P4, Tesla P40, or Tesla P100), you may use NVIDIA driver release 384.111+ or 410. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 19.04 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, and Turing families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 19.04 is based on [Kaldi 5.5](#).

- ▶ Latest version of [NVIDIA NCCL 2.4.6](#)
- ▶ Latest version of [cuBLAS 10.1.0.105](#)
- ▶ Reduced memory usage and improved performance
- ▶ Improved scripts for running benchmarks
- ▶ Ubuntu 16.04 with March 2019 updates

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` and `/workspace/nvidia-examples/librispeech/run_multigpu_benchmark.sh` which run inference on the trained model and data.

## Known Issues

There are no known issues in this release.

---

# Chapter 61. Kaldi Release 19.03

The NVIDIA container image for Kaldi, release 19.03, is available on [NGC](#).

## Contents of the Kaldi container

This container image contains the complete source of the version of Kaldi in `/opt/kaldi`.

The container also includes the following:

- ▶ [Ubuntu 16.04](#) including [Python 3.5](#)
- ▶ [NVIDIA CUDA 10.1.105](#) including [cuBLAS 10.1.105](#)
- ▶ [NVIDIA cuDNN 7.5.0](#)
- ▶ [NVIDIA NCCL 2.4.3](#) (optimized for [NVLink™](#))
- ▶ [OpenMPI 3.1.3](#)
- ▶ [TensorRT 5.1.2](#)

## Driver Requirements

Release 19.03 is based on CUDA 10.1, which requires [NVIDIA Driver](#) release 418.xx+. However, if you are running on Tesla (Tesla V100, Tesla P4, Tesla P40, or Tesla P100), you may use NVIDIA driver release 384.111+ or 410. The CUDA driver's compatibility package only supports particular drivers. For a complete list of supported drivers, see the [CUDA Application Compatibility](#) topic. For more information, see [CUDA Compatibility and Upgrades](#).

## GPU Requirements

Release 19.03 supports CUDA compute capability 6.0 and higher. This corresponds to GPUs in the Pascal, Volta, and Turing families. Specifically, for a list of GPUs that this compute capability corresponds to, see [CUDA GPUs](#). For additional support details, see [Deep Learning Frameworks Support Matrix](#).

## Key Features and Enhancements

This Kaldi release includes the following key features and enhancements.

- ▶ [Kaldi](#) container image version 19.03 is based on [Kaldi 5.5](#).

- ▶ Latest version of [NVIDIA CUDA 10.1.105](#) including [cuBLAS 10.1.105](#)
- ▶ Latest version of [NVIDIA cuDNN 7.5.0](#)
- ▶ Latest version of [NVIDIA NCCL 2.4.3](#)
- ▶ Latest version of [OpenMPI 3.1.3](#)
- ▶ Latest version of [TensorRT 5.1.2](#)
- ▶ This is the first release for the optimized Kaldi container.
- ▶ Ubuntu 16.04 with February 2019 updates

## Packaged scripts

The Kaldi container comes with the following scripts:

- ▶ `/workspace/nvidia-examples/librispeech/prepare.sh` which downloads a trained model and data.
- ▶ `/workspace/nvidia-examples/librispeech/run_benchmark.sh` and `/workspace/nvidia-examples/librispeech/run_multigpu_benchmark.sh` which run inference on the trained model and data.

## Known Issues

- ▶ If using or upgrading to a 3-part-version driver, for example, a driver that takes the format of `xxx.yy.zz`, you will receive a `Failed to detect NVIDIA driver version.` message. This is due to a known bug in the entry point script's parsing of 3-part driver versions. This message is non-fatal and can be ignored. This will be fixed in the 19.04 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.

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 NON-INFRINGEMENT, 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.

## 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.

**Trademarks**

NVIDIA, the NVIDIA logo, and cuBLAS, CUDA, DALI, DGX, DGX-1, DGX-2, DGX Station, DLProf, Jetson, Kepler, Maxwell, NCCL, Nsight Compute, Nsight Systems, NvCaffe, PerfWorks, Pascal, SDK Manager, Tegra, TensorRT, Triton Inference Server, Tesla, TF-TRT, and Volta are trademarks and/or registered trademarks of NVIDIA Corporation in the U.S. and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

**Copyright**

© 2018-2026 NVIDIA Corporation & Affiliates. All rights reserved.

