NVIDIA TensorRT

Support Matrix | NVIDIA Docs

Table of Contents

Chapter 1. Features for Platforms and Software

This section lists the supported NVIDIA® TensorRT™ features based on which platform and software.

Table 1.   List of Supported Features per Platform

	Linux x86-64	Windows x64	Linux ppc64le	Linux AArch64
	8.5.x	8.5.x	8.5.x	8.5.x
Supported NVIDIA CUDA® versions	11.8 11.7 update 1[1] 11.6 update 2[2] 11.5 update 2[3] 11.4 update 4[4] 11.3 update 1[5]	11.8 11.7 update 1[9] 11.6 update 2[10] 11.5 update 2[11] 11.4 update 4[12] 11.3 update 1[13]	11.8	11.8 11.4

---

[1] These CUDA versions are supported using a single build, built with CUDA toolkit 11.8. It is compatible with all CUDA 11.x versions and only requires driver 450.x.

[9] These CUDA versions are supported using a single build, built with CUDA toolkit 11.8. It is compatible with all CUDA 11.x versions and only requires driver 450.x.

[2] These CUDA versions are supported using a single build, built with CUDA toolkit 11.8. It is compatible with all CUDA 11.x versions and only requires driver 450.x.

[10] These CUDA versions are supported using a single build, built with CUDA toolkit 11.8. It is compatible with all CUDA 11.x versions and only requires driver 450.x.

[3] These CUDA versions are supported using a single build, built with CUDA toolkit 11.8. It is compatible with all CUDA 11.x versions and only requires driver 450.x.

[11] These CUDA versions are supported using a single build, built with CUDA toolkit 11.8. It is compatible with all CUDA 11.x versions and only requires driver 450.x.

[4] These CUDA versions are supported using a single build, built with CUDA toolkit 11.8. It is compatible with all CUDA 11.x versions and only requires driver 450.x.

[12] These CUDA versions are supported using a single build, built with CUDA toolkit 11.8. It is compatible with all CUDA 11.x versions and only requires driver 450.x.

[5] These CUDA versions are supported using a single build, built with CUDA toolkit 11.8. It is compatible with all CUDA 11.x versions and only requires driver 450.x.

[13] These CUDA versions are supported using a single build, built with CUDA toolkit 11.8. It is compatible with all CUDA 11.x versions and only requires driver 450.x.

	Linux x86-64	Windows x64	Linux ppc64le	Linux AArch64
	8.5.x	8.5.x	8.5.x	8.5.x
	11.2 update 2[6]	11.2 update 2[14]		
	11.1 update 1[7]	11.1 update 1[15]		
	11.0 update 1[8]	11.0 update 1[16]		
	10.2	10.2		
Supported cuBLAS versions	11.11.3.6	11.11.3.6	11.11.3.6	11.11.3.6
	11.10.3.66	11.10.3.66		11.6.5.2
	11.9.2.110	11.9.2.110		
	11.7.4.6	11.7.4.6		
	11.6.5.2	11.6.5.2		
	11.5.1.109	11.5.1.109		
	11.4.1.1043	11.4.1.1043		
	11.3.0.106	11.3.0.106		
	11.2.0.252	11.2.0.252		
	10.2.3.254	10.2.3.254		
Supported cuDNN versions	cuDNN 8.6.0	cuDNN 8.6.0	cuDNN 8.6.0	cuDNN 8.6.0
TensorRT Python API	Yes	Yes	Yes	Yes
NvUffParser	Yes	Yes	Yes	Yes
NvOnnxParser	Yes	Yes	Yes	Yes
Loops	Yes	Yes	Yes	Yes

> **Note:**
>
> ▶ Serialized engines are not portable across platforms or TensorRT versions.

---

[6] These CUDA versions are supported using a single build, built with CUDA toolkit 11.8. It is compatible with all CUDA 11.x versions and only requires driver 450.x.

[14] These CUDA versions are supported using a single build, built with CUDA toolkit 11.8. It is compatible with all CUDA 11.x versions and only requires driver 450.x.

[7] These CUDA versions are supported using a single build, built with CUDA toolkit 11.8. It is compatible with all CUDA 11.x versions and only requires driver 450.x.

[15] These CUDA versions are supported using a single build, built with CUDA toolkit 11.8. It is compatible with all CUDA 11.x versions and only requires driver 450.x.

[8] These CUDA versions are supported using a single build, built with CUDA toolkit 11.8. It is compatible with all CUDA 11.x versions and only requires driver 450.x.

[16] These CUDA versions are supported using a single build, built with CUDA toolkit 11.8. It is compatible with all CUDA 11.x versions and only requires driver 450.x.

▶ Refer to the minimum compatible driver versions in the *NVIDIA CUDA Release Notes* for specific NVIDIA Driver versions.

Chapter 2. Hardware and Precision

The following table lists NVIDIA hardware and which precision modes that each hardware supports. TensorRT supports all NVIDIA hardware with capability SM 5.0 or higher. It also lists the availability of DLA on this hardware. Refer to the following tables for the specifics.

> 💬 **Note:** Support for CUDA compute capability version 3.0 has been removed. Support for CUDA compute capability versions below 5.0 may be removed in a future release and is now deprecated.

Table 2.   Supported Hardware

CUDA Compute Capability	Example Device	TF32	FP32	FP16	INT8	FP16 Tensor Cores	INT8 Tensor Cores	DLA
9.0	NVIDIA H100	Yes	Yes	Yes	Yes	Yes	Yes	No
8.9	NVIDIA RTX 4090	Yes	Yes	Yes	Yes	Yes	Yes	No
8.7	NVIDIA DRIVE AGX Orin™	Yes	Yes	Yes	Yes	Yes	Yes	Yes
8.6	NVIDIA A10	Yes	Yes	Yes	Yes	Yes	Yes	No
8.0	NVIDIA A100/ GA100 GPU	Yes	Yes	Yes	Yes	Yes	Yes	No
7.5	NVIDIA T4	No	Yes	Yes	Yes	Yes	Yes	No

CUDA Compute Capabilit	Example Device	TF32	FP32	FP16	INT8	FP16 Tensor Cores	INT8 Tensor Cores	DLA
7.2	Jetson AGX Xavier	No	Yes	Yes	Yes	Yes	Yes	Yes
7.0	NVIDIA V100	No	Yes	Yes	Yes	Yes	No	No
6.1	NVIDIA P4	No	Yes	Yes	Yes	No	No	No
6.0	NVIDIA P100	No	Yes	Yes	No	No	No	No
5.2	NVIDIA M4	No	Yes	No	No	No	No	No
5.0	Quadro K2200	No	Yes	No	No	No	No	No

Deprecated Hardware

Table 3.   List of Supported Precision Mode per Hardware

CUDA Compute Capability	Example Device	FP32	FP16	INT8	FP16 Tensor Cores	INT8 Tensor Cores	DLA
3.7	NVIDIA K80	Yes	No	No	No	No	No
3.5	NVIDIA K40	Yes	No	No	No	No	No

Removed Hardware

Table 4.   List of Supported Precision Mode per Hardware

CUDA Compute Capability	Example Device	FP32	FP16	INT8	FP16 Tensor Cores	INT8 Tensor Cores	DLA
3.0	NVIDIA K10	Yes	No	No	No	No	No

Chapter 3. Compute Capability Per Platform

The section lists the supported compute capability based on platform.

Table 5.　　　Compute Capability per Platform

Platform	Compute capability
Linux x86-64	3.5, 3.7, 5.0, 5.2, 6.0, 6.1, 7.0, 7.5, 8.0[17], 8.6[18], 8.9[19], 9.0[20]
Windows 10 x64	3.5, 3.7, 5.0, 5.2, 6.0, 6.1, 7.0, 7.5, 8.0[21], 8.6[22], 8.9[23], 9.0[24]
CentOS 8.5 ppc64le	7.0, 7.5, 8.0, 8.6, 9.0
Ubuntu 20.04 SBSA	7.0, 7.5, 8.0, 8.6, 9.0
NVIDIA JetPack AArch64	7.2, 8.7

---

[17] Requires CUDA Toolkit 11.0 or newer and a TensorRT CUDA 11.x build.
[18] Requires CUDA Toolkit 11.1 or newer and a TensorRT CUDA 11.x build.
[19] Requires CUDA Toolkit 11.8 or newer and a TensorRT CUDA 11.x build.
[20] Requires CUDA Toolkit 11.8 or newer and a TensorRT CUDA 11.x build.
[21] Requires CUDA Toolkit 11.0 or newer and a TensorRT CUDA 11.x build.
[22] Requires CUDA Toolkit 11.1 or newer and a TensorRT CUDA 11.x build.
[23] Requires CUDA Toolkit 11.8 or newer and a TensorRT CUDA 11.x build.
[24] Requires CUDA Toolkit 11.8 or newer and a TensorRT CUDA 11.x build.

Chapter 4. Software Versions Per Platform

The section lists the supported software versions based on platform.

Table 6.   List of Supported Platforms per Software Version

Platform	Compiler version	Python version
Ubuntu 18.04 x86-64	gcc 8.3.1	3.6
Ubuntu 20.04 x86-64	gcc 8.3.1	3.8
Ubuntu 22.04 x86-64	gcc 8.3.1	3.10
CentOS 7.9 x86-64	gcc 8.3.1	3.6
CentOS 8.5 x86-64	gcc 8.3.1	3.8
SLES 15 x86-64	gcc 8.3.1	Not Applicable
Windows 10 x64	MSVC 2017u8	Not Applicable
CentOS 8.5 ppc64le	Clang 14.0.6	3.8
Ubuntu 20.04 SBSA	gcc 8.4.0	3.8
NVIDIA JetPack AArch64	gcc 9.4.0	3.8

> **Note:** Python versions supported when using Debian or RPM packages. When using Python wheel files, versions 3.6, 3.7, 3.8, 3.9, and 3.10 are supported.

Chapter 5.   ONNX Operator Support

The ONNX operator support list for TensorRT can be found [here](#).

NVIDIA Corporation | 2788 San Tomas Expressway, Santa Clara, CA 95051
http://www.nvidia.com

Trademarks

NVIDIA, the NVIDIA logo, and BlueField, CUDA, DALI, DRIVE, Hopper, JetPack, Jetson AGX Xavier, Jetson Nano, Maxwell, NGC, Nsight, Orin, Pascal, Quadro, Tegra, TensorRT, Triton, Turing and Volta are trademarks and/or registered trademarks of NVIDIA Corporation in the United States and other countries. Other company and product names may be trademarks of the respective companies with which they are associated.

Copyright