



NVDEC Application Note

Release 13.0

NVIDIA Corporation

Dec 16, 2025

Contents

1	Introduction	1
2	NVDEC Capabilities	3
3	NVDEC Performance	5
4	Programming NVDEC	9
5	FFmpeg Support	11

Chapter 1. Introduction

NVIDIA GPUs and NVIDIA Jetson platforms contain a hardware-based decoder (referred to as NVDEC in this document) which provides fully accelerated hardware-based video decoding for several popular codecs. With complete decoding offloaded to NVDEC, the graphics engine and CPU are free for other operations.

NVDEC supports much faster than real-time decoding which makes it suitable for transcoding scenarios in addition to video playback.

The hardware capabilities available in NVDEC are exposed through APIs referred to as NVDECODE APIs in this document. This document provides information about the capabilities of the NVDEC engine and the features exposed through NVDECODE APIs.

Chapter 2. NVDEC Capabilities

At a high level, *NVDEC Hardware Capabilities* summarizes the capabilities of the NVDEC engine exposed through NVDEC CODE APIs.

Table 1: NVDEC Hardware Capabilities

Hardware Features	1 st Gen Maxwell GPUs	2 nd Gen Maxwell GPUs	Pascal GPUs	Volta GPUs	Turing/ GA100/ Hopper GPUs	GA10x ³ and ADA GPUs	Blackwell GPUs/ Jetson Thor
VC1 Simple, Main & Advanced profiles	Y	Y	Y	Y	Y	Y	Y
MPEG4 Simple and Advanced Simple Profiles	Y	Y	Y	Y	Y	Y	Y
MPEG2 Simple & Main profiles	Y	Y	Y	Y	Y	Y	Y
H.264 Baseline, Main, High Profiles	Y	Y	Y	Y	Y	Y	Y
H.264 8192x8192 De-coding support	N	N	N	N	N	N	Y
H264 High10/High422 profiles	N	N	N	N	N	N	Y
VP8	N	Y	Y ¹	Y	Y	Y	Y

continues on next page

Table 1 – continued from previous page

Hardware Features	1 st Gen Maxwell GPUs	2 nd Gen Maxwell GPUs	Pascal GPUs	Volta GPUs	Turing/ GA100/ Hopper GPUs	GA10x ³ and ADA GPUs	Blackwell GPUs/ Jetson Thor
HEVC Main and Main 10 Profile ¹	N	Y ¹	Y	Y	Y	Y	Y
HEVC 444 decoding	N	N	N	N	Y	Y	Y
HEVC main 422 10/12 profiles	N	N	N	N	N	N	Y
HEVC 8192x8192 De-coding support	N	N	Y ¹	Y	Y	Y	Y
VP9 8192x8192 Decoding support	N	N	Y ¹	Y	Y	Y	Y
VP9 Profile 0 ¹	N	Y ¹	Y	Y	Y	Y	Y
Multiple NVDECs ²	N	N	N	N	Y	Y	Y
AV1 Main Profile decoding	N	N	N	N	N	Y	Y

► **Y:** Supported, **N:** Unsupported

► ¹: Present in select GPUs

► ²: Present in select GPUs

► ³: GA10x GPUs include all GPUs based on Ampere architecture except GA100

Chapter 3. NVDEC Performance

NVDEC natively supports multiple hardware decoding contexts with negligible context-switching penalty. As a result, subject to the hardware performance limit and available memory, an application can decode multiple videos simultaneously.

The hardware and software maintain the context for each decoding session, allowing many simultaneous decoding sessions to run in parallel with minimal context switch penalty. *NVDEC Decoding Performance* provides indicative data of the decoding performance of NVDEC in GPUs based on Maxwell, Pascal, Turing, Ampere and Blackwell architectures as well as Jetson platforms based on Thor for AV1, HEVC, VP9, and H.264 encoded bitstreams. The performance varies across GPU classes (e.g. Quadro, Tesla), and scales (almost) linearly with the clock speeds for each hardware.

Table 1: NVDEC Decoding Performance (Indicative)

GPU Architecture/ Jetson Platform	Codec	Performance in frames/second
Pascal	H.264	694
	VP9	846
	HEVC	810
	HEVC Main10	789
Turing	H.264	771
	VP9	932
	VP9 10 bit	925
	HEVC	1316
	HEVC Main10	1158
Ampere	H.264	748
	VP9	1075
	VP9 10 bit	1120
	HEVC	1415
	HEVC Main10	1299
	AV1	790

continues on next page

Table 1 – continued from previous page

GPU Architecture/ Jetson Platform	Codec	Performance in frames/second
Ada	H.264	903
	VP9	1290
	VP9 10 bit	1342
	HEVC	1641
	HEVC Main10	1520
	AV1	1018
Blackwell	H.264	2172
	VP9	1445
	VP9 10 bit	1498
	HEVC	1872
	HEVC Main10	1818
	AV1	1119
Jetson Thor	H.264	1434
	VP9	1019
	VP9 10 bit	1016
	HEVC	1293
	HEVC Main10	1130
	AV1	794

- All the measurement is done on the highest video clocks as reported by nvidia-smi (i.e. 1544 MHz, 1860 MHz, 1665 MHz, 2160 MHz, 2362 MHz for Pascal, Turing, Ampere, Ada, and Blackwell GPUs respectively and 1691 MHz for Jetson Thor platform). The performance should scale according to the video clocks as reported by nvidia-smi on target GPU. Information on nvidia-smi can be found at <https://developer.nvidia.com/nvidia-system-management-interface>.
- Resolution/Input format: 1920x1080/YUV 4:2:0
- Software: Windows 11/Jetson Linux 38.2 for GPUs/Jetson platform, Video Codec SDK v13.0
- Hopper and GA100 GPUs contain NVDEC with same architecture as Turing. As a result, the decoding performance on Hopper and GA100 GPUs is same as that of Turing GPUs, scaled by the clock speed. To view the clocks available on your GPU, please use the tool nvidia-smi included with the NVIDIA driver.

While first-generation Maxwell GPUs had one NVDEC engine per chip, certain variants of the second-generation Maxwell, Pascal, Volta and Ada GPUs have two/three NVDEC engines per chip. This increases the aggregate decoder performance of the GPU. NVIDIA driver takes care of load balancing among multiple NVDEC engines on the chip, so that applications don't require any special code to take advantage of multiple decoders and automatically benefit from higher decoder capacity on higher-end GPU hardware. The decode performance listed in *NVDEC Decoding Performance* is given per NVDEC

engine. Thus, if a Quadro or Tesla GPU has 2 NVDECs, multiply the corresponding number in *NVDEC Decoding Performance* by the number of NVDECs per chip to get aggregate maximum performance (applicable only when running multiple simultaneous decode sessions). Note that performance with a single decoding session cannot exceed performance per NVDEC, regardless of the number of NVDECs present on the GPU. All GeForce products consist of a single NVDEC.

Chapter 4. Programming NVDEC

The NVDEC API provides access to the video decoding features of NVDEC described in the previous chapters. The API includes implementations of commonly used post-processing operations such as scaling, cropping, aspect ratio conversion, deinterlacing and color space conversion to the decoded output.

Refer to the SDK release notes for information regarding the required driver version.

Various capabilities of NVDEC are exposed to the application software via the NVIDIA proprietary application programming interface (NVDEC APIs). Refer to the Video Decoder Programming guide for details on using these APIs.

For a complete list of GPUs supporting hardware accelerated decoding refer to <https://developer.nvidia.com/nvidia-video-codec-sdk>.

Chapter 5. FFmpeg Support

FFmpeg is the most popular multimedia transcoding tool used extensively for video and audio transcoding. Note that FFmpeg is open-source project and its usage is governed by specific licenses and terms and conditions.

The video hardware accelerators in NVIDIA GPUs can be effectively used with FFmpeg to significantly speed up the video decoding, encoding and end-to-end transcoding at very high performance.

Note

The Video Codec SDK v13.0 does not support using video hardware accelerators with FFmpeg on Jetson platforms.

Notices

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 acknowledgment, 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.

Trademarks

NVIDIA, the NVIDIA logo, and cuBLAS, CUDA, CUDA Toolkit, cuDNN, DALI, DIGITS, DGX, DGX-1, DGX-2, DGX Station, DLProf, GPU, Jetson, Kepler, Maxwell, NCCL, Nsight Compute, Nsight Systems, NVcaffe, NVIDIA Deep Learning SDK, NVIDIA Developer Program, NVIDIA GPU Cloud, NVLink, NVSHMEM, PerfWorks, Pascal, SDK Manager, Tegra, TensorRT, TensorRT Inference Server, Tesla, TF-TRT, Triton Inference Server, 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

©2022-2025, NVIDIA Corporation & affiliates. All rights reserved