NVIDIA VIDEO CODEC SDK - DECODER

vNVDEC_DA-08097-001_v08 | June 2020

Application Note

TABLE OF CONTENTS

Chapter 1. NVIDIA HARDWARE VIDEO DECODER

1.1. Introduction

NVIDIA GPUs 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. The current document highlights *only* the changes in the current video codec SDK package with respect to the previous SDK packages. To know about the features exposed in earlier SDKs please refer to the earlier SDK package(s).

1.2. NVDEC Capabilities

At a high level, Table 1 summarizes the capabilities of the NVDEC engine exposed through NVDECODE APIs, and What's new summarize the features exposed through NVDECODE APIs in Video Codec SDK 10.0.

Table 1 NVDEC Hardware Capabilities

Hardware Features	Kepler GPUs	1st Gen Maxwell GPUs	2nd Gen Maxwell GPUs	Pascal GPUs	Volta GPUs	Turing/ GA100 GPUs
VC1 Simple, Main & Advanced profiles	Y	Y	Y	Y	Y	Y

Hardware Features	Kepler GPUs	1st Gen Maxwell GPUs	2nd Gen Maxwell GPUs	Pascal GPUs	Volta GPUs	Turing/ GA100 GPUs
MPEG4 Simple and Advanced Simple Profiles	Y	Y	Y	Y	Y	Y
MPEG2 Simple & Main profiles	Y	Y	Y	Y	Y	Y
H.264 Baseline, Main, High Profiles	Y	Y	Y	Y	Y	Y
VP8*	N	N	Y	Y	Y	Y
HEVC Main Profile*	N	N	Y	Y	Y	Y
VP9 Profile 0*	N	N	Y	Y	Y	Y
8192x8192** Decoding support (HEVC&VP9 only)	N	N	N	Y	Y	Y
Multiple NVDECs***	N	N	N	N	N	Y
HEVC 444 decoding	N	N	N	N	N	Y

- **Y**: Supported, **N**: Unsupported
- * : Present in select Maxwell second generation GPUs, all Pascal, Volta, Turing and GA100 GPUs
- ** : Present in select Pascal, all Volta, all Turing GPUs and GA100
- ***: Present in select Turing GPUs and GA100

1.3. What's new

SDK 10.0 adds support for GA100 and enhanced NVDEC throughput brought by GA100.

1.4. 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. Table 2 provides indicative data of the decoding performance of NVDEC across Kepler, Maxwell, Pascal, Turing and GA100 GPUs for 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 2 NVDEC decoding performance (indicative)

GPU Architecture	Codec	Performance in frames/second
Second generation Maxwell(M2000)	H.264	427
	VP9	528
	HEVC	514
	HEVC Main10	454
Pascal(P2000)	H.264	647
	VP9	825
	VP9 10 bit	802
	HEVC	793
	HEVC Main10	775
Turing (RTX8000)	H.264	689
	VP9	874
	VP9 10 bit	878
	HEVC	1234
	HEVC Main10	1145

- All the measurement is done on the highest video clocks as reported by nvidia-smi (i.e. 1129 MHz, 1683 MHz, 1755 MHz for M2000, P2000 and RTX8000 respectively). The performance should scale according to the video clocks as reported by nvidia-smi for other GPUs of every individual family. 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 10, Video Codec SDK 10.0, NVIDIA display driver: 445.87
- GA100 GPUs contain NVDEC with same architecture as Turing. As a result, the decoding performance on 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 Kepler, Maxwell, Pascal, and Volta generation GPUs had one NVDEC engine per chip, some Quadro and Tesla boards based on Turing architecture have multiple NVDEC engines per chip and GA100 has 5 NVDECs. This increases the aggregate decoding throughput of the GPU. The NVIDIA driver takes care of load balancing among multiple NVDEC engines on the chip so that applications don't require 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 Table 2 is given per NVDEC engine. Thus, if a Quadro or Tesla GPU has 2 NVDECs, multiply the corresponding number in Table 2 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.

1.5. Programming NVDEC

Video Codec SDK 10.0 is supported on R445 (Windows) and R450 (Linux) drivers and above. 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 (NVDECODE 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.

1.6. FFmpeg Support

FFmpeg is the most popular multimedia transcoding tool used extensively for video and audio transcoding.

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 that FFmpeg is open-source project and its usage is governed by specific licenses and terms and conditions.