NVIDIA VIDEO CODEC SDK - DECODER

NVDEC_DA-08097-001_v07| Aug 2019

Application Note

DOCUMENT CHANGE HISTORY

NVDEC_DA-08097-001_v07

Version	Date	Authors	Description of Change	Highlight
01	June 10, 2016	SM	Initial Release	Support for VP8&VP9 Decoding
02	Nov 15, 2016	SM	Updated for Video Codec SDK 7.1	
03	Feb 15, 2017	SM	Updated for Video Codec SDK 8.0	10&12-bit decoding
04	Jan 10, 2018	SM	Updated for Video Codec SDK 8.1	Sample applications based on re-usable classes
05	April 10, 2018	SM	Updated for Video Codec SDK 8.2	NVDECODE APIs added for decode status reporting and reconfiguration of the decoder
06	Jan 10, 2019	SM	Updated for Video Codec SDK 9.0	Support for Turing class GPUs
07	Aug 10, 2019	SM	Updated for Video Codec SDK 9.1	

TABLE OF CONTENTS

LIST OF TABLES

NVIDIA HARDWARE VIDEO DECODER

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).

2. NVDEC CAPABILITIES

At a high level, Table 1 summarizes the capabilities of the NVDEC engine exposed through NVDECODE APIs, and Table 2 and Table 3 summarize the features exposed through NVDECODE APIs in Video Codec SDK 9.0 and Video Codec SDK 9.1 respectively.

Table 1. NVDEC Hardware Capabilities

Hardware Features	Kepler GPUs	1st Gen Maxwell GPUs	2nd Gen Maxwell GPUs	Pascal GPUs	Volta GPUs	Turing GPUs
VC1 Simple, Main & Advanced profiles	✓	✓	✓	✓	✓	✓
MPEG4 Simple and Advanced Simple Profiles	✓	✓	✓	✓	✓	✓
MPEG2 Simple & Main profiles	✓	✓	✓	✓	✓	✓
H.264 Baseline, Main, High Profiles	✓	✓	✓	✓	✓	✓
VP8*	×	×	✓	✓	✓	✓
HEVC Main Profile*	×	×	✓	✓	✓	✓
VP9 Profile 0*	×	×	✓	✓	✓	✓
8192x8192** Decoding support (HEVC&VP9 only)	×	×	×	✓	✓	✓
Multiple NVDECs***	×	×	×	×	×	✓
HEVC 444 decoding	×	×	×	×	×	✓

* : Present in select Maxwell second generation GPUs, all Pascal, Volta and Turing GPUs

** : Present in select Pascal, all Volta and all Turing GPUs

***: Present in select Turing GPUs

3. WHAT'S NEW IN SDK 9.0 AND SDK 9.1

Table 2. What is new in SDK 9.0

Features	Description
Support for multiple NVDEC engines	Select Quadro and Tesla Turing GPUs have 2 or 3 NVDECs for higher video decoding throughput. The driver takes care of load balancing among multiple NVDECs on the chip. Applications don't need any changes for getting the benefit of additional NVDECs.
HEVC YUV 444 decoding	Decoding support for 8/10/12 bit YUV444 content was added in Turing GPUs. SDK 9.0 adds support for such bitstreams. YUV444 is the preferred format where preserving chroma fidelity is important.

Table 3. What is new in SDK 9.1

Features	Description
Enhancements to decode capability API	Support has been added to the NVDECODE API to report the output surface formats supported for the queried set of capabilities.
Memory optimization in sample applications	The sample applications in the SDK package have been optimized for video memory consumption.

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 4 provides indicative data of the decoding performance of NVDEC across Kepler, Maxwell, Pascal, and Turing 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.

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. 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 4 is given per NVDEC engine. Thus, if a Quadro or Tesla GPU has 2 NVDECs, multiply the corresponding number in Table 4 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.

Table 4. NVDEC decoding performance (indicative)

GPU Architecture	Codec	Performance in frames/second
Kepler(K2000)	H.264	141
Second generation Maxwell(M2000)	H.264	427
	VP9	532
	HEVC	515
	HEVC Main10	455
Pascal(P2000)	H.264	648
	VP9	820
	VP9 10 bit	797
	HEVC	789
	HEVC Main10	763
Turing (RTX8000)	H.264	690
	VP9	856
	VP9 10 bit	868
	HEVC	1261
	HEVC Main10	1157

- All the measurement is done on the highest video clocks as reported by nvidia-smi (i.e. 540 MHz, 1129 MHz, 1683 MHz, 1755 MHz for K2000, 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 here.
- Resolution/Input format: 1920x1080/YUV 4:2:0
- Software: Windows 10, Video Codec SDK 9.1, NVIDIA display driver: 436.15
- The encoding performance on Volta GPUs scales up with the performance numbers on Pascal GPUs in proportion to the highest video clocks as reported by nvidia-smi.
- Please note, some of the numbers may look slightly different from the earlier SDKs as the content used for evaluation is different.

5. PROGRAMMING NVDEC

Video Codec SDK 9.0 and Video Codec SDK 9.1 are supported on R418 and R435 drivers and above respectively. 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.

6. FFMPEG AND LIBAV SUPPORT

FFmpeg and Libav are the most popular multimedia transcoding tools used extensively for video and audio transcoding.

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

Note that FFmpeg and Libav are open-source projects and their usage is governed by specific licenses and terms and conditions for each of these projects.

Notice

Trademarks

Copyright