NSIGHT COMPUTE

v2019.1.1 | February 2019

User Manual

TABLE OF CONTENTS

LIST OF TABLES

Chapter 1.
INTRODUCTION

1.1. Overview

This document is a user guide to the next-generation NVIDIA Nsight Compute profiling tools. NVIDIA Nsight Compute is an interactive kernel profiler for CUDA applications. It provides detailed performance metrics and API debugging via a user interface and command line tool. In addition, its baseline feature allows users to compare results within the tool. NVIDIA Nsight Compute provides a customizable and data-driven user interface and metric collection and can be extended with analysis scripts for post-processing results.

Important Features

- ▶ Interactive kernel profiler and API debugger
- ▶ Graphical profile report
- ▶ Result comparison across one or multiple reports within the tool
- ▶ Fast Data Collection
- ▶ UI and Command Line interface
- ▶ Fully customizable reports and analysis rules

Chapter 2.
QUICKSTART

The following sections provide brief step-by-step guides of how to setup and run NVIDIA Nsight Compute to collect profile information. All directories are relative to the base directory of NVIDIA Nsight Compute, unless specified otherwise.

The UI executable is called nv-nsight-cu. A shortcut with this name is located in the base directory of the NVIDIA Nsight Compute installation. The actual executable is located in the folder **host\windows-desktop-win7-x64** on Windows or **host/linux-desktop-glibc_2_11_3-x64** on Linux. By default, NVIDIA Nsight Compute is installed in **/usr/local/cuda-<cuda-version>/NsightCompute-<version>** on Linux and in **C:\Program Files\NVIDIA Corporation\Nsight Compute <version>** Windows.

After starting NVIDIA Nsight Compute, by default the *Welcome Page* is opened. It provides links to recently opened reports and projects as well as quick access to the Connection Dialog, and the Projects dialogs. To immediately start a profile run, select *Continue* under *Quick Launch*. See Environment on how to change the start-up action.
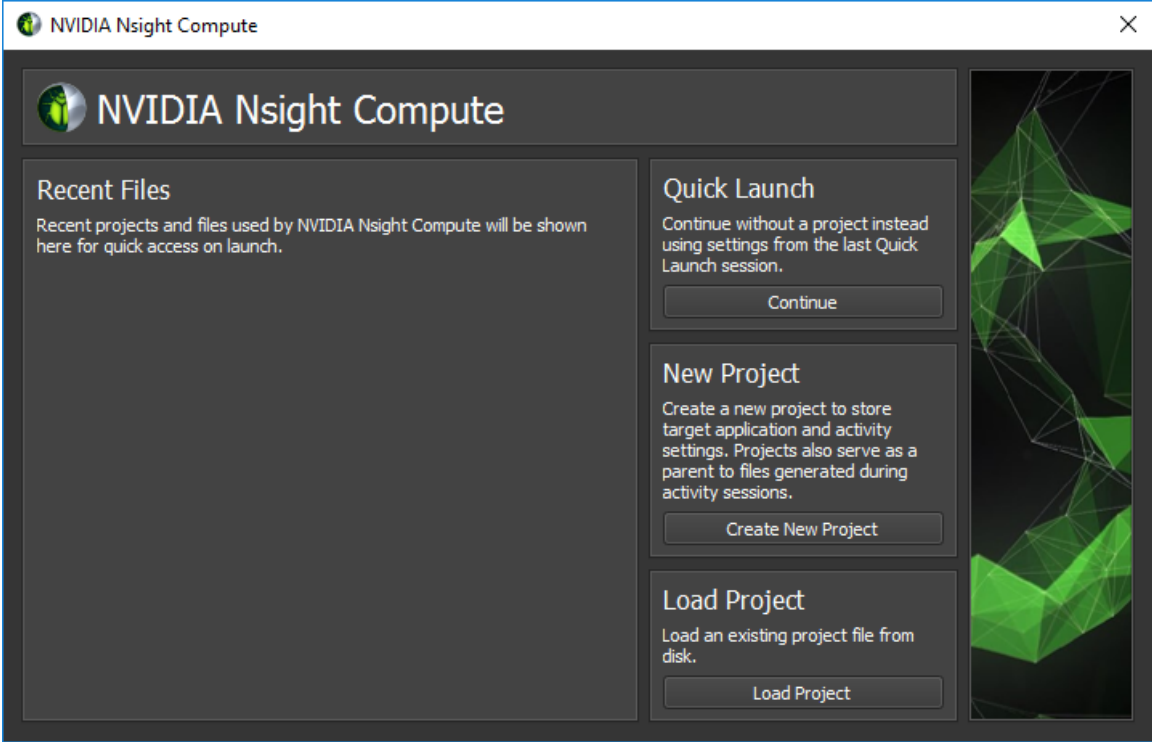
## 2.1. Interactive Profile Activity

1. **Launch the target application from NVIDIA Nsight Compute**

   When starting NVIDIA Nsight Compute, the *Welcome Page* will appear. Click on *Quick Launch* to open the *Connection* dialog. If the *Connection* dialog doesn't appear, you can open it using the *Connect* button from the main toolbar, as long as you are not currently connected. Select your target platform on the left-hand side and your connection target (machine) from the *Connection* drop down. If you have your local target platform selected, `localhost` will become available as a connection. Use + to add a new connection target. Then, fill in the launch details and select Launch. In the *Activity* panel, select the Interactive Profile activity to initiate a session that allows controlling the execution of the target application and selecting the kernels of interest interactively. Press *Launch* to start the session.
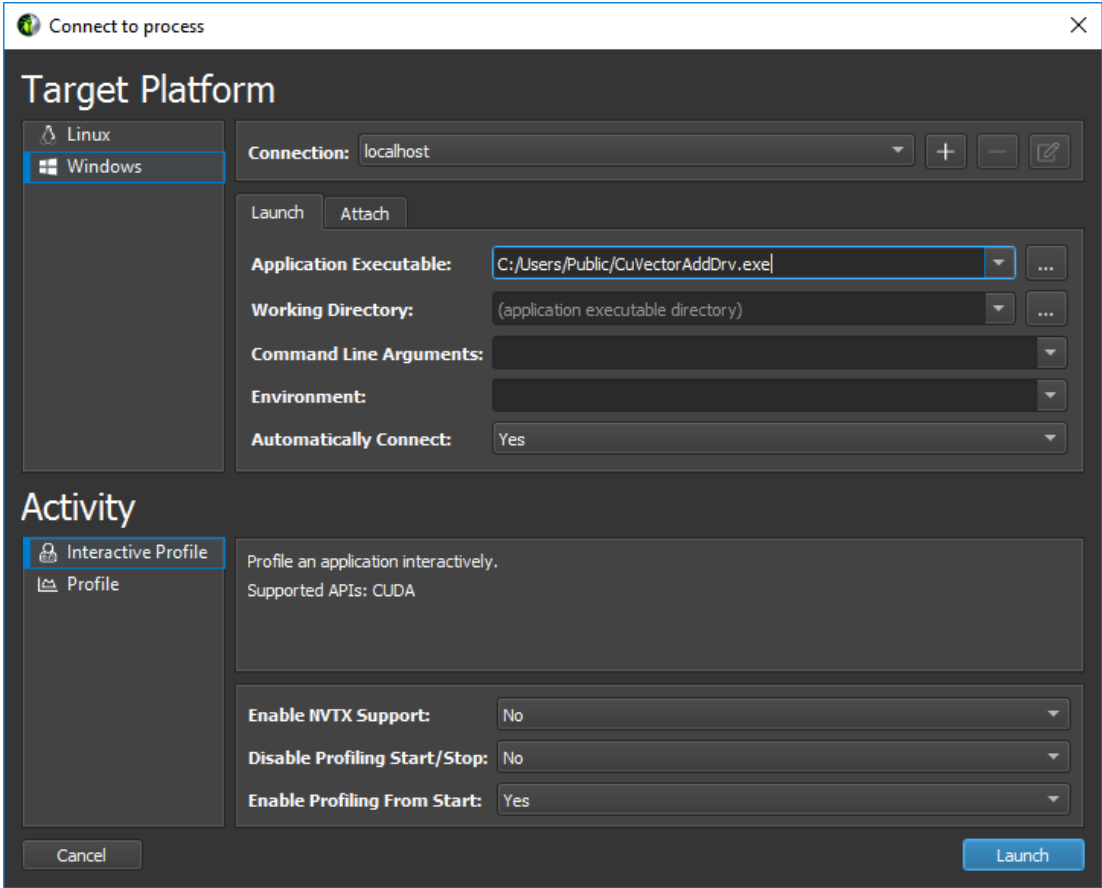
2. **Launch the target application with tools instrumentation from the command line**

   The nv-nsight-cu-cli can act as a simple wrapper that forces the target application to load the necessary libraries for tools instrumentation. The parameter **--mode=launch** specifies that the target application should be launched and suspended before the first instrumented API call. That way the application waits until we connect with the UI.

   ```
   $ nv-nsight-cu-cli --mode=launch CuVectorAddDrv.exe
   ```

3. **Launch NVIDIA Nsight Compute and connect to target application**
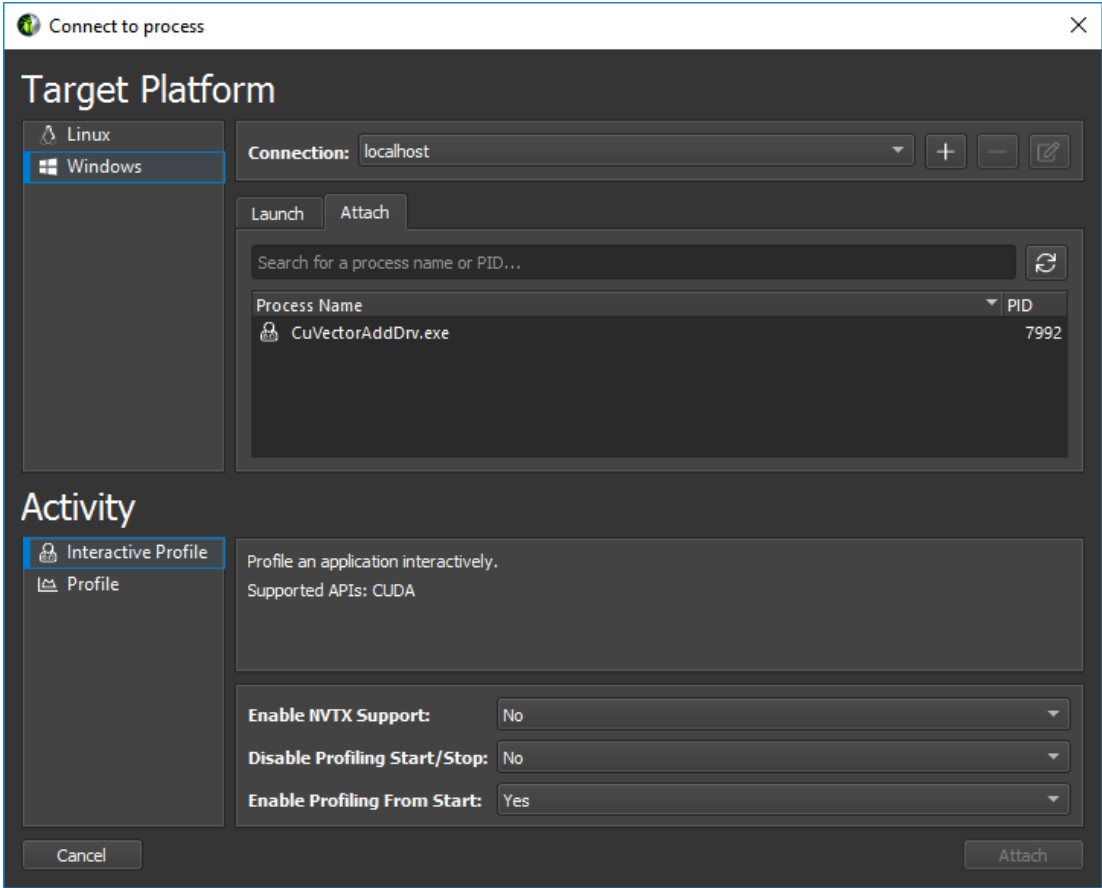
Select the target machine at the top of the dialog to connect and update the list of attachable applications. By default, *localhost* is pre-selected if the target matches your current local platform. Select the *Attach* tab and the target application of interest and press *Attach*. Once connected, the layout of NVIDIA Nsight Compute changes into stepping mode that allows you to control the execution of any calls into the instrumented API. When connected, the *API Stream* window indicates that the target application waits before the very first API call.
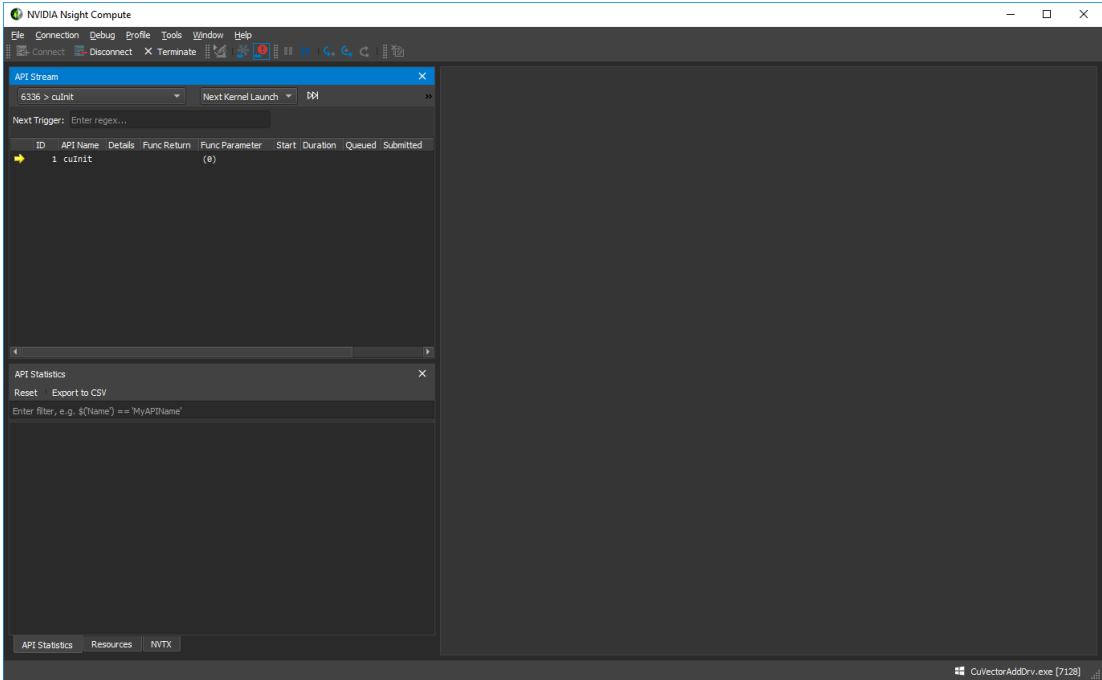
4. **Control application execution**

   Use the *API Stream* window to step the calls into the instrumented API. The dropdown at the top allows switching between different CPU threads of the application. *Step In* (F11), *Step Over* (F10), and *Step Out* (Shift + F11) are available from the *Debug* menu or the corresponding toolbar buttons. While stepping, function return values and function parameters are captured.
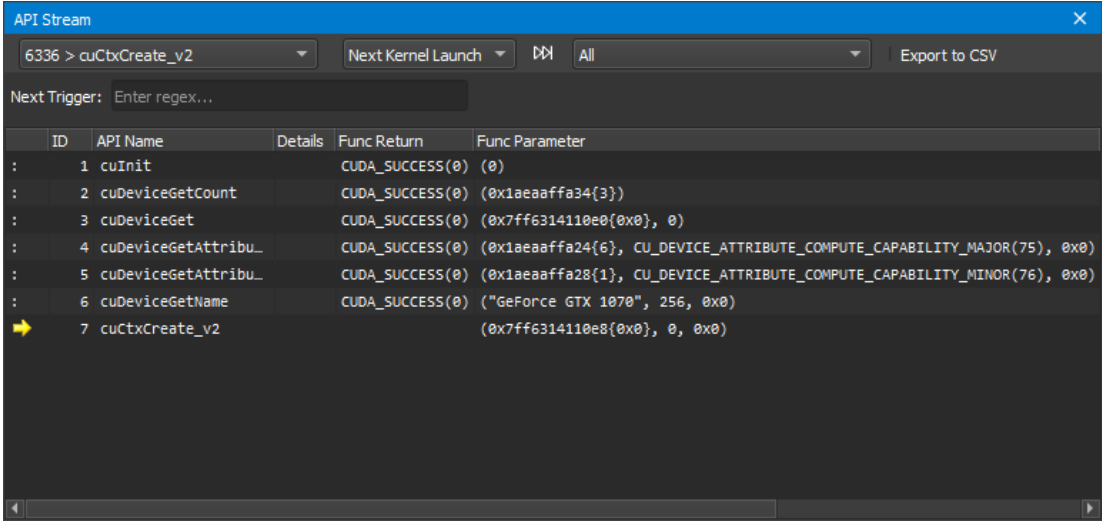
   Use *Resume* (F5) and *Pause* to allow the program to run freely. Freeze control is available to define the behavior of threads currently not in focus, i.e. selected in the thread drop down. By default, the *API Stream* stops on any API call that returns an error code. This can be toggled in the *Debug* menu by *Break On API Error*.

5. **Isolate a kernel launch**

To quickly isolate a kernel launch for profiling, use the *Next API Launch* button in the toolbar of the *API Stream* window to jump to the next kernel launch. The execution will stop before the kernel launch is executed.
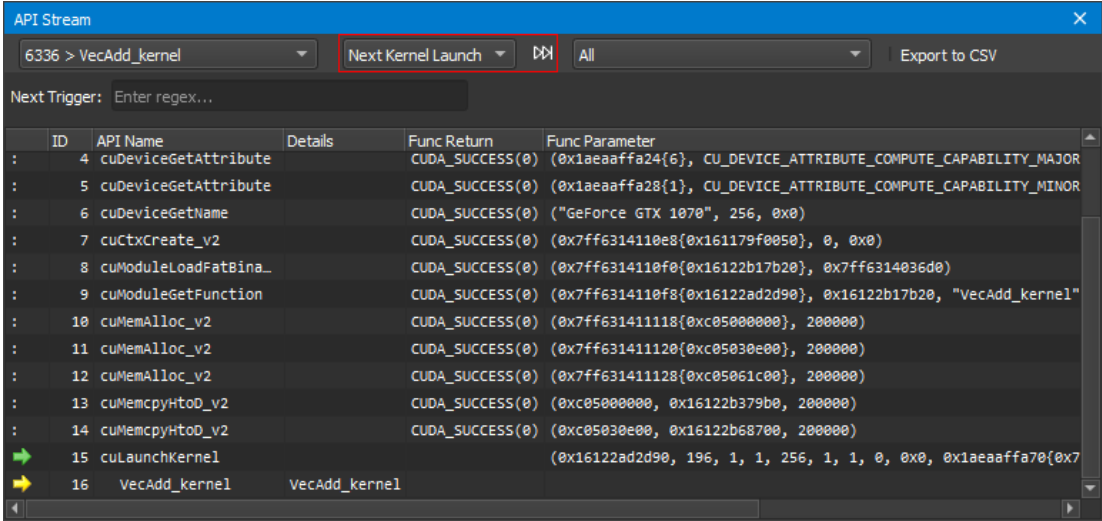
6. **Profile a kernel launch**

Once the execution of the target application is suspended at a kernel launch, additional actions become available in the UI. These actions are either available from the menu or from the toolbar. Please note that the actions are disabled, if the API stream is not at a qualifying state (not at a kernel launch or launching on an unsupported GPU). To profile, press *Profile* and wait until the result is shown in the Profiler Report. Profiling progress is reported in the lower right corner status bar.

Instead of manually selecting *Profile*, it is also possible to enable *Auto Profile* from the *Profile* menu. If enabled, each kernel matching the current kernel filter (if any) will be profiled using the current section configuration. This is especially useful if an application is to be profiled unattended, or the number of kernel launches to be profiled is very large. Sections can be enabled or disabled using the Sections/Rules Info tool window.

2.2. Non-Interactive Profile Activity

1. **Launch the target application from NVIDIA Nsight Compute**

When starting NVIDIA Nsight Compute, the *Welcome Page* will appear. Click on *Quick Launch* to open the *Connection* dialog. If the *Connection* dialog doesn't appear, you can open it using the *Connect* button from the main toolbar, as long as you are not currently connected. Select your target platform on the left-hand side and your localhost from the *Connection* drop down. Then, fill in the launch details and select *Launch*. In the *Activity* panel, select the *Profile* activity to initiate a session that pre-configures the profile session and launches the command line profiler to collect the data. Provide the *Output File* name to enable starting the session with the *Launch* button.
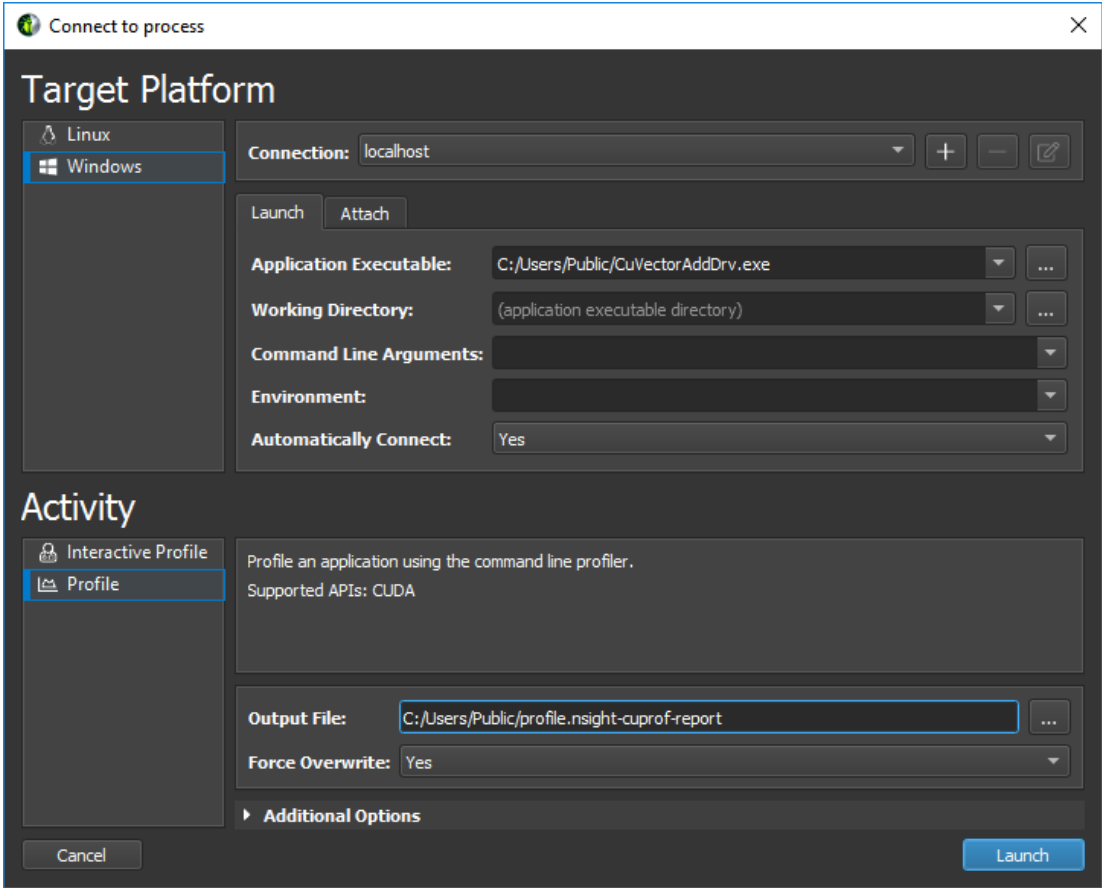
2. **Additional Launch Options**

   The expander at the bottom of the activity pane can be used to show additional configuration options. For more details on these options see the Command Line Options of the command line profiler. The options are grouped into tabs: The *Filter* tab exposes the options to specify which kernels should be profiled. Options include the kernel regex filter, the number of launches to skip, and the total number of launch to profile. The *Section* tab allows to select which sections should be collected for each kernel launch of interest. The *Other* tab includes the option to collect NVTX information or custom metrics via the **--metrics** option.
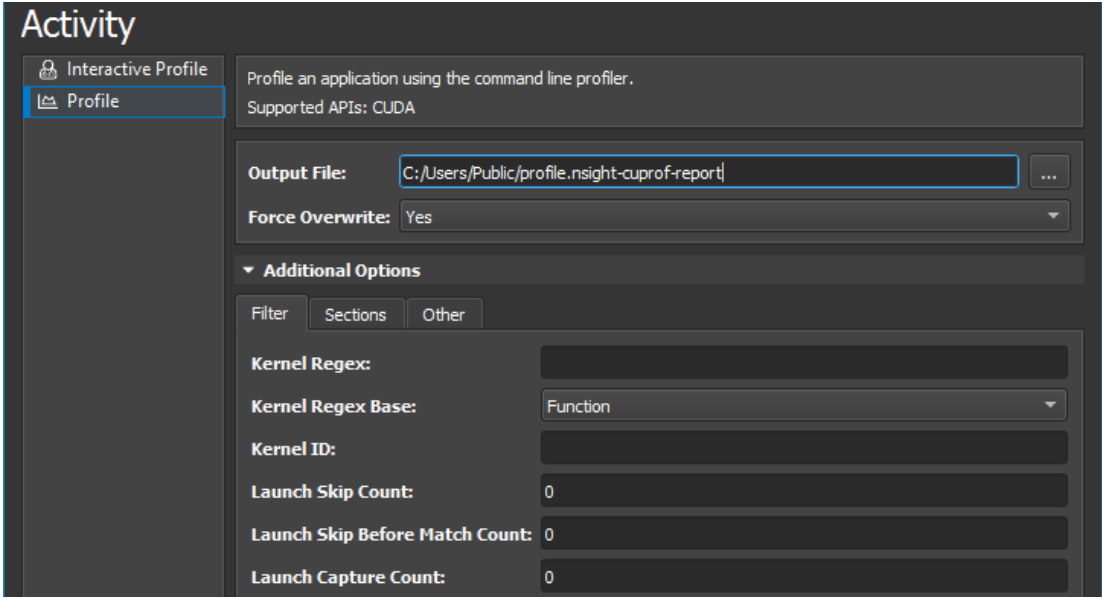
2.3. Navigate the Report

1. **Navigate the report**

   The profile report comes up by default on the *Details* page. You can switch between different Report Pages of the report with the dropdown labeled Page on the top-left of the report. A report can contain any number of results from kernel launches. The *Launch* dropdown allows switching between the different results in a report.
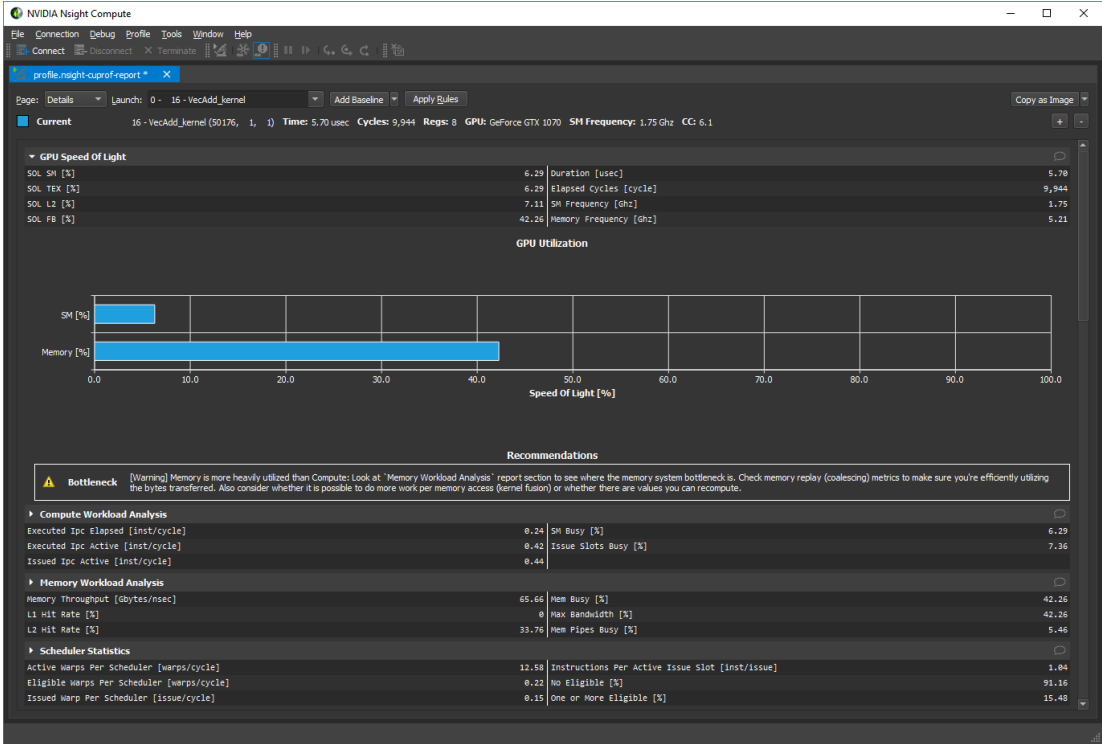
2. **Diffing multiple results**

On the *Details page,* press the button *Add Baseline* to promote the current result in focus to become the baseline all other results from this report and any other report opened in the same instance of NVIDIA Nsight Compute gets compared to. If a baseline is set, every element on the Details page shows two values: The current value of the result in focus and the corresponding value of the baseline or the percentage of change from the corresponding baseline value.
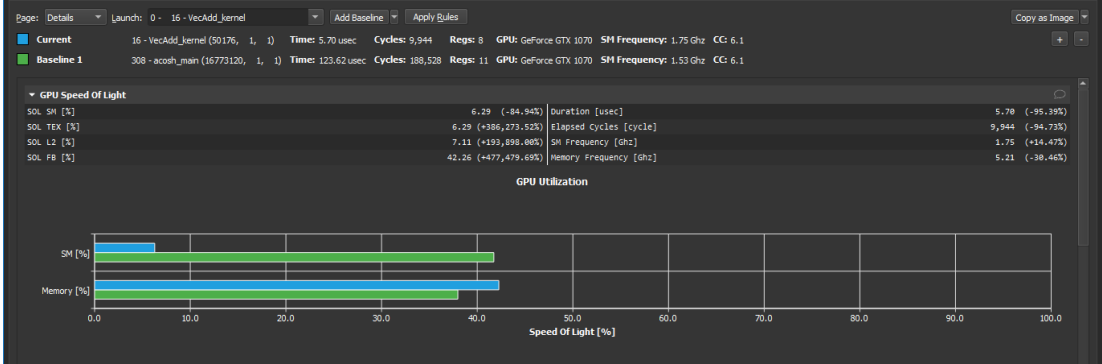
Use the *Clear Baselines* entry from the dropdown button, the Profile menu or the corresponding toolbar button to remove all baselines. For more information see Baselines.

3. **Executing rules**

On the *Details* page some sections may provide rules. Press the *Apply button* to execute an individual rule. The *Apply Rules* button on the top executes all available rules for the current result in focus. Rules can be user-defined too. For more information see the *Customization Guide.*

Chapter 3.
CONNECTION DIALOG

Use the *Connection Dialog* to launch and attach to applications on your local and remote platforms. Start by selecting the *Target Platform* for profiling. By default (and if supported) your local platform will be selected. Select the platform on which you would like to start the target application or connect to a running process.
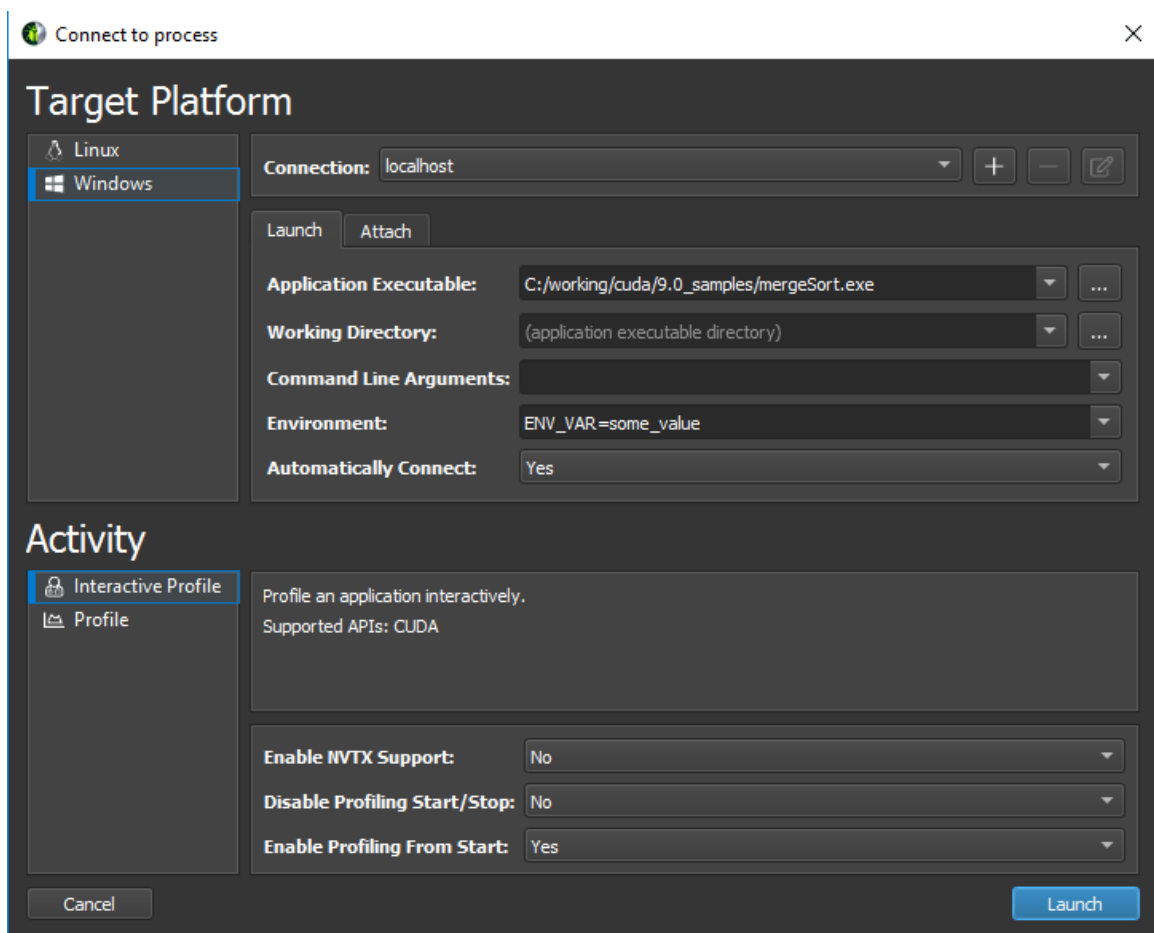
When using a remote platform, you will be asked to select or create a *Connection* in the top drop down. To create a new connection, select + and enter your connection details.

When using the local platform, *localhost* will be selected as the default and no further connection settings are required. You can still create or select a remote connection, if profiling will be on a remote system of the same platform.

Depending on your target platform, select either *Launch* or *Remote Launch* to launch an application for profiling on the target. Note that *Remote Launch* will only be available if supported on the target platform. Fill in launch details for the application, such as the executable, working directory, command line arguments etc.

Select *Attach* to attach the profiler to an application already running on the target platform. This application must have been started using another NVIDIA Nsight Compute CLI instance. The list will show all application processes running on the target system which can be attached. Select the refresh button to re-create this list.

Finally, select the *Activity* to be run on the target for the launched or attached application. Note that not all activities are necessarily compatible with all targets and connection options. Currently, the following activities exist:

▶ Interactive Profile Activity
▶ Non-Interactive Profile Activity

3.1. Remote Connections

Remote devices that support SSH can also be configured as a target in the *Connection Dialog*. To configure a remote device, ensure an SSH-capable *Target Platform* is selected, then press the + button. The following configuration dialog will be presented.

In this dialog, enter the following information:

▶ **IP/Host Name:** The IP address or host name of the target device.
▶ **User Name:** The user name to be used for the SSH connection.

▸ **Password:** The user password to be used for the SSH connection.

▸ **Port:** The port to be used for the SSH connection. (The default value is 22.)

▸ **Deployment Directory:** The directory to use on the target device to deploy supporting files. The specified user must have write permissions to this location.

When all information is entered, click the *Add* button to make use of this new connection.

When a remote connection is selected in the *Connection Dialog*, the *Application Executable* file browser will browse the remote file system using the configured SSH connection, allowing the user to select the target application on the remote device.

When an activity is launched on a remote device, the following steps are taken:

1. The command line profiler and supporting files are copied into the *Deployment Directory* on the the remote device. (Only files that do not exist or are out of date are copied.)

2. The *Application Executable* is executed on the remote device.

   ▸ For *Interactive Profile* activities, a connection is established to the remote application and the profiling session begins.

   ▸ For *Non-Interactive Profile* activities, the remote application is executed under the command line profiler and the specified report file is generated.

3. For non-interactive profiling activities, the generated report file is copied back to the host, and opened.

The progress of each of these steps is presented in the *Progress Log*.
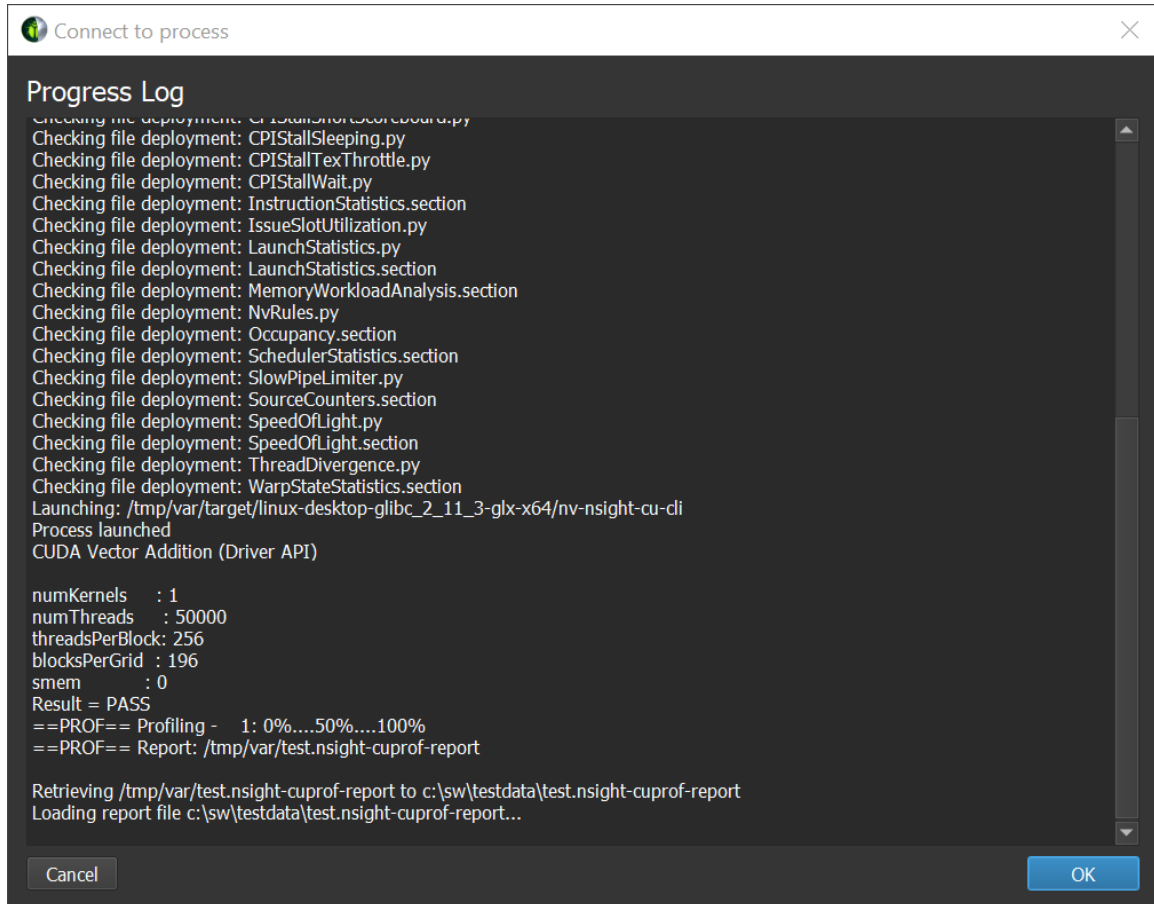
Note that once either activity type has been launched remotely, the tools necessary for further profiling sessions can be found in the *Deployment Directory* on the remote device.

Chapter 4.
MAIN MENU AND TOOLBAR

Information on the main menu and toolbar.
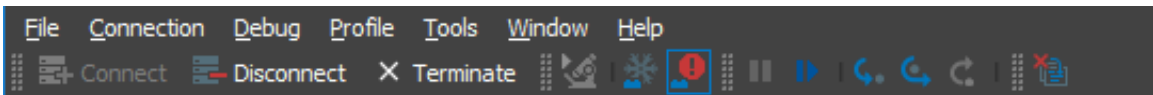
4.1. Main Menu

▸ File

  ▸ **New Project** Create new profiling Projects with the New Project Dialog
  ▸ **Open Project** Open an existing profiling project
  ▸ **Recent Projects** Open an existing profiling project from the list of recently used projects
  ▸ **Save Project** Save the current profiling project
  ▸ **Save Project As** Save the current profiling project with a new filename
  ▸ **Close Project** Close the current profiling project

▸ Connection

  ▸ **Connect** Open the Connection Dialog to launch or attach to a target application. Disabled when already connected.
  ▸ **Disconnect** Disconnect from the current target application, allows the application to continue normally and potentially re-attach.
  ▸ **Terminate** Disconnect from and terminate the current target application immediately.

▸ Debug

  ▸ **Pause** Pause the target application at the next intercepted API call or launch.
  ▸ **Resume** Resume the target application.
  ▸ **Step In** Step into the current API call or launch to the next nested call, if any, or the subsequent API call, otherwise.
  ▸ **Step Over** Step over the current API call or launch and suspend at the next, non-nested API call or launch.

- ▸ **Step Out** Step out of the current nested API call or launch to the next, non-parent API call or launch one level above.
- ▸ **Next API Launch** See API Stream
- ▸ **Next API Trigger** See API Stream
- ▸ **Freeze API**

  When disabled, all CPU threads are enabled and continue to run during stepping or resume, and all threads stop as soon as at least one thread arrives at the next API call or launch. This also means that during stepping or resume the currently selected thread might change as the old selected thread makes no forward progress and the API Stream automatically switches to the thread with a new API call or launch. When enabled, only the currently selected CPU thread is enabled. All other threads are disabled and blocked.

  Stepping now completes if the current thread arrives at the next API call or launch. The selected thread never changes. However, if the selected thread does not call any further API calls or waits at a barrier for another thread to make progress, stepping may not complete and hang indefinitely. In this case, pause, select another thread, and continue stepping until the original thread is unblocked. In this mode, only the selected thread will ever make forward progress.
- ▸ **Break On API Error** When enabled, during resume or stepping, execution is suspended as soon as an API call returns an error code.
- ▸ **API Statistics** Opens the API Statistics tool window
- ▸ **API Stream** Opens the API Stream tool window
- ▸ **Resources** Opens the Resources tool window
- ▸ **NVTX** Opens the NVTX tool window
- ▸ Profile

  - ▸ **Clear Baselines** Clear all current baselines.
  - ▸ **Auto Profile** Enable or disable auto profiling. If enabled, each kernel matching the current kernel filter (if any) will be profiled using the current section configuration.
  - ▸ **Section/Rules Info** Opens the Sections/Rules Info tool window.
- ▸ Tools

  - ▸ **Profile Kernel** When suspended at a kernel launch, select the profile using the current configuration.
  - ▸ **Project Explorer** Opens the Project Explorer tool window.
  - ▸ **Output Messages** Opens the Output Messages tool window.
  - ▸ **Options** Opens the Options dialog.
- ▸ Window

  - ▸ **Show Welcome Page** Opens the Welcome Page.
- ▸ Help

  - ▸ **Documentation** Opens the latest documentation for NVIDIA Nsight Compute online.

- ▶ **Documentation (local)** Opens the local HTML documentation for NVIDIA Nsight Compute that has shipped with the tool.
- ▶ **Reset Application Data** Reset all NVIDIA Nsight Compute configuration data saved on disk, including option settings, default paths, recent project references etc. This will not delete saved reports.
- ▶ **About** Opens the About dialog with information about the version of NVIDIA Nsight Compute.

4.2. Main Toolbar
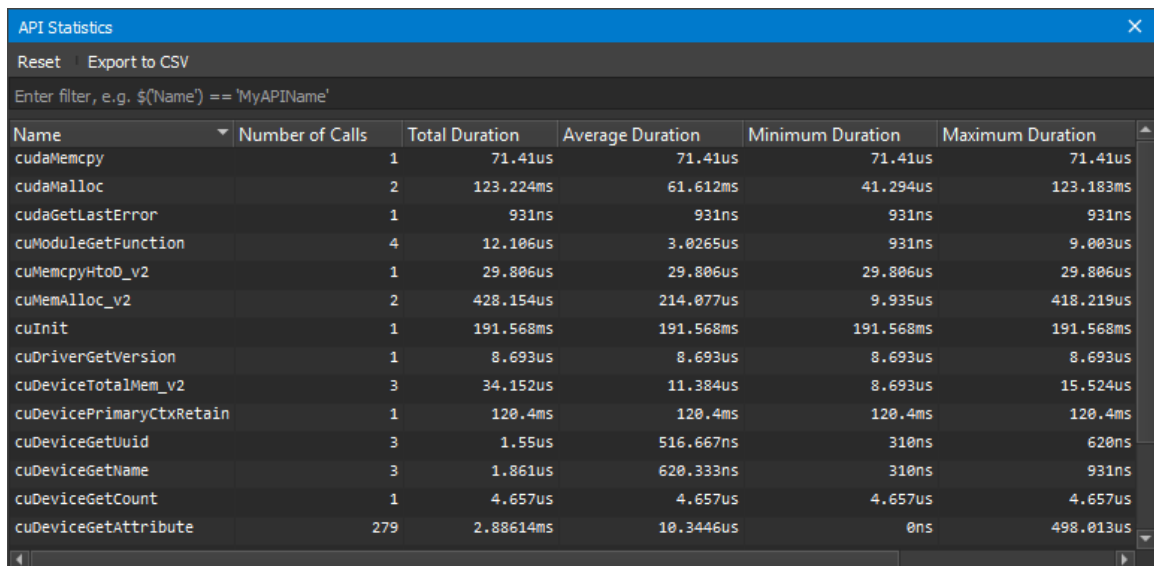
The main toolbar shows commonly used operations from the main menu. See Main Menu for their description.

Chapter 5.
TOOL WINDOWS

5.1. API Statistics

The *API Statistics* window is available when NVIDIA Nsight Compute is connected to a target application. It opens by default as soon as the connection is established. It can be re-opened using *Debug > API Statistics* from the main menu.

Whenever the target application is suspended, it shows a summary of tracked API calls with some statistical information, such as the number of calls, their total, average, minimum and maximum duration. Note that this view cannot be used as a replacement for Nsight Systems when trying to optimize CPU performance of your application.

The *Reset* button deletes all statistics collected to the current point and starts a new collection. Use the *Export to CSV* button to export the current statistics to a CSV file.

5.2. API Stream

The API Stream window is available when NVIDIA Nsight Compute is connected to a target application. It opens by default as soon as the connection is established. It can be re-opened using *Debug > API Stream* from the main menu.
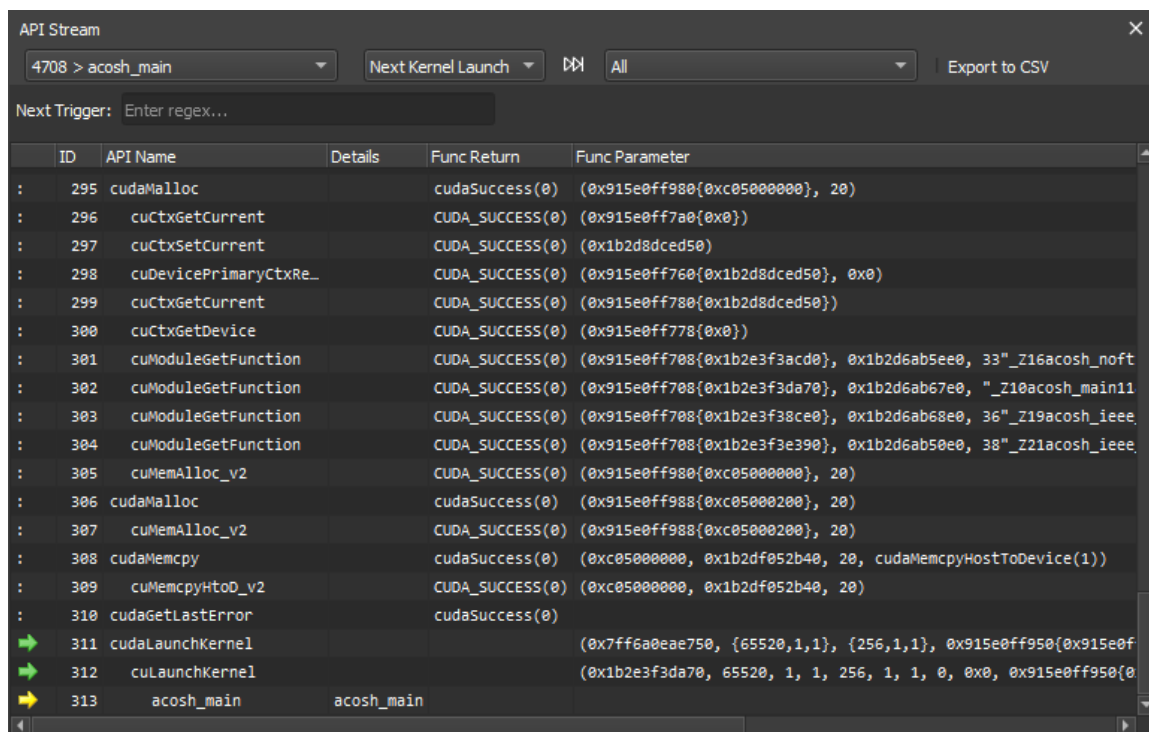
Whenever the target application is suspended, the window shows the history of API calls and traced kernel launches. The currently suspended API call or kernel launch (activity) is marked with a yellow arrow. If the suspension is at a subcall, the parent call is marked with a green arrow. The API call or kernel is suspended before being executed.

For each activity, further information is shown such as the kernel name or the function parameters (*Func Parameters*) and return value (*Func Return*). Note that the function return value will only become available once you step out or over the API call.

Use the *Current Thread* dropdown to switch between the active threads. The dropdown shows the thread ID followed by the current API name. One of several options can be chosen in the trigger dropdown, which are executed by the adjacent >> button. *Next Kernel Launch* resumes execution until the next kernel launch is found in any enabled thread. *Next API Call* resumes execution until the next kernel launch matching *Kernel Trigger* is found in any enabled thread. *Next Range Start* resumes execution until the next start of an active profiler range is found. Profiler ranges are defined by using the **cu(da)ProfilerStart/Stop** API calls. *Next Range Stop* resumes execution until the next stop of an active profiler range is found. The *API Level* dropdown changes which

API levels are shown in the stream. The *Export to CSV* button exports the currently visible stream to a CSV file.

5.3. NVTX

The *NVTX* window is available when NVIDIA Nsight Compute is connected to a target application. If closed, it can be re-opened using *Debug > NVTX* from the main menu. Whenever the target application is suspended, the window shows the state of all active NVTX domains and ranges in the currently selected thread. Note that NVTX information is only tracked if the launching command line profiler instance was started with `--nvtx` or NVTX was enabled in the NVIDIA Nsight Compute launch dialog.

NVTX					✕
11312		Category	Color	API Call ID	Paylo
▾ A					
▾ Start/End Ranges					
	<0, Outer>	My Category	#000000		
	<2, Inner>	My Category	#000000	340	

Use the *Current Thread* dropdown in the API Stream window to change the currently selected thread. NVIDIA Nsight Compute supports NVTX named resources, such as threads, CUDA devices, CUDA contexts, etc. If a resource is named using NVTX, the appropriate UI elements will be updated.
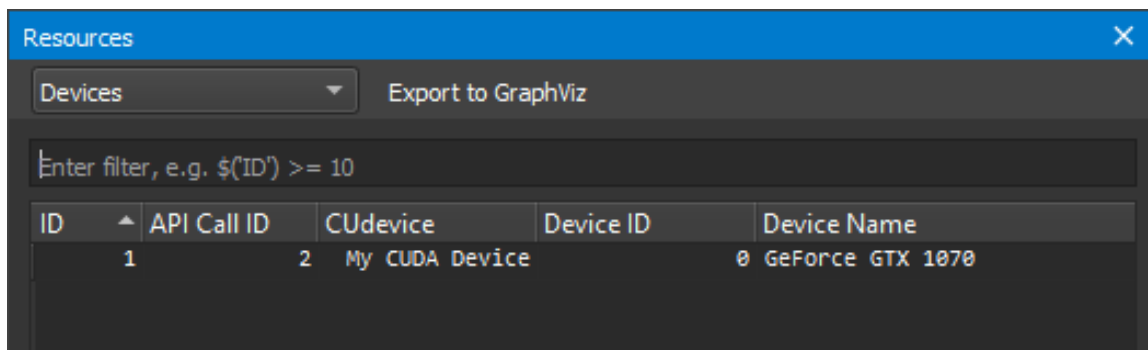
5.4. Resources

The *Resources* window is available when NVIDIA Nsight Compute is connected to a target application. It shows information about the currently known resources, such as CUDA devices, CUDA streams or kernels. The window is updated every time the target application is suspended. If closed, it can be re-opened using *Debug > Resources* from the main menu.
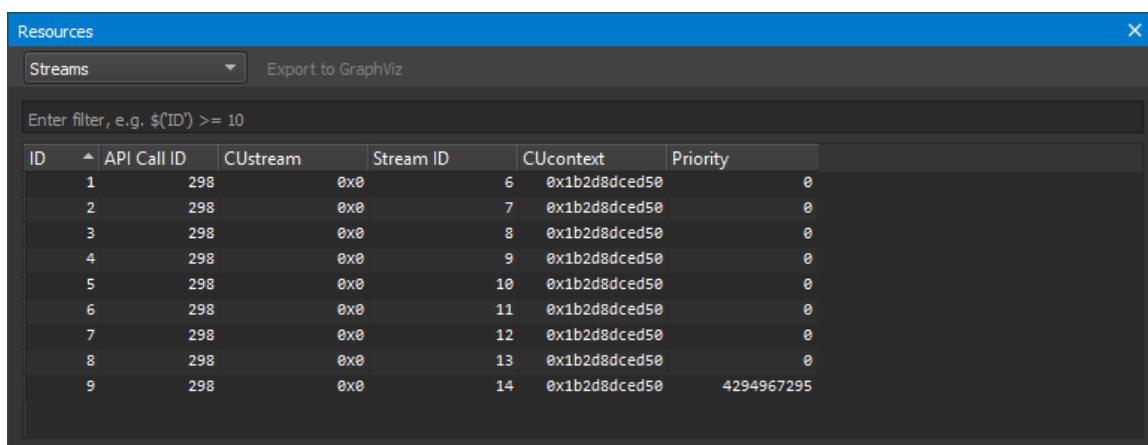
Using the dropdown on the top, different views can be selected, where each view is specific to one kind of resource (context, stream, kernel, …). The *Filter* edit allows you to create filter expressions using the column headers of the currently selected resource.

The resource table shows all information for each resource instance. Each instance has a unique ID, the *API Call ID* when this resource was created, its handle, associated handles, and further parameters. When a resource is destroyed, it is removed from its table.

5.5. Sections/Rules Info

The *Sections/Rules Info* window can be opened from the main menu using *Profile > Sections/Rules Info*. It tracks all sections and rules currently loaded in NVIDIA Nsight Compute, independent from a specific connection or report. The directory to load those files from can be configured in the Profile options dialog. It is used to inspect available

sections and rules, as well as to configure which sections should be collected and which rules should be applied.

Whenever a kernel is profiled manually, or when auto-profiling is enabled, only sections enabled in this window are collected. Similarly, whenever rules are applied, only rules enabled in this window are active.

The enabled states of sections and rules are persisted across NVIDIA Nsight Compute launches. The *Reload* button reloads all sections and rules from disk again. If a new section or rule is found, it will be enabled if possible. If any errors occur while loading a rule, they will be listed in an extra entry with a warning icon and a description of the error.

Use the *Enable All* and *Disable All* checkboxes to enable or disable all sections and rules at once. The Filter text box can be used to filter what is currently shown in the window. It does not alter enablement of any entry.

The table shows sections and rules with their enablement status, their relationship and further parameters, such as associated metrics or the original file on disk. Rules associated with a section are shown as children of their section entry. Rules independent of any section are shown under an additional *Independent Rules* entry.

Double-clicking an entry in the table's *Filename* column opens this file as a document. It can be edited and saved directly in NVIDIA Nsight Compute. After editing the file, *Reload* must be selected to apply those changes.

See Sections and Rules for the list of default sections for NVIDIA Nsight Compute.

Chapter 6.
PROFILER REPORT

The profiler report contains all the information collected during profiling for each kernel launch. In the user interface, it consists of a header with general information, as well as controls to switch between report pages or individual collected launches. By default, the report starts with the *Details* page selected.

6.1. Header

The *Page* dropdown can be used to switch between the available report pages, which are explained in detail in the next section.
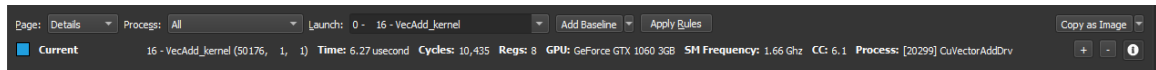
The *Process* dropdown filters which kernel launches are available in the *Launch* dropdown. Select a single process to see only launches from this process. Selecting *All* shows all instances.

The *Launch* dropdown can be used to switch between all collected kernel launches. The information displayed in each page commonly represents the selected launch instance. On some pages (e.g. *Raw*), information for all launches is shown and the selected instance is highlighted. You can type in this dropdown to quickly filter and find a kernel launch.

The *Add Baseline* button promotes the current result in focus to become the baseline of all other results from this report and any other report opened in the same instance of NVIDIA Nsight Compute. Select the arrow dropdown to access the *Clear Baselines* button, which removes all currently active baselines.

The *Apply Rules* button applies all rules available for this report. If rules had been applied previously, those results will be replaced. By default, rules are applied immediately once the kernel launch has been profiled. This can be changed in the options under *Tools > Options > Profile > Report UI > Apply Applicable Rules Automatically*.

On the right-hand side, multiple export buttons are available via an arrow dropdown. Each provides some export functionality to store the currently focused page and launch instance to another format, e.g. as an image or CSV file.

Information about the selected kernel is shown as *Current*. [+] and [-] buttons can be used to show or hide the section body content. The info toggle button *i* changes the section description's visibility.

6.2. Report Pages

Use the *Page* dropdown in the header to switch between the report pages.

6.2.1. Session Page

This *Session* page contains basic information about the report and the machine, as well as device attributes of all devices for which launches were profiled. When switching between launch instances, the respective device attributes are highlighted.

6.2.2. Summary Page

The *Summary* page shows a list of all collected results in this report, with selected important summary metrics. It gives you a quick comparison overview across all profiled kernel launches. You can transpose the table of kernels and metrics by using the *Transpose* button.

6.2.3. Details Page

The *Details* page is the main page for all metric data collected during a kernel launch. The page is split into individual sections. Each section consists of a header table and an optional body that can be expanded. The sections are completely user defined and can be changed easily by updating their respective files. For more information on customizing sections, see the *Customization Guide*. For a list of sections shipped with NVIDIA Nsight Compute, see Sections and Rules.

You can add or edit comments in each section of the *Details* view by clicking on the comment button (speech bubble). The comment icon will be highlighted in sections that contain a comment. Comments are persisted in the report and are summarized in the Comments Page.

▼ GPU Speed Of Light		
SOL SM [%]	6.29	Duration [usec]	5.70
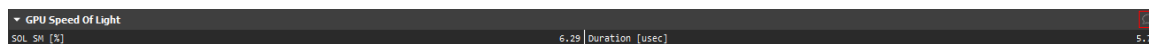
6.2.4. Source Page

The *Source* page correlates assembly (SASS) with high-level code and PTX. In addition, it displays metrics that can be correlated with source code.
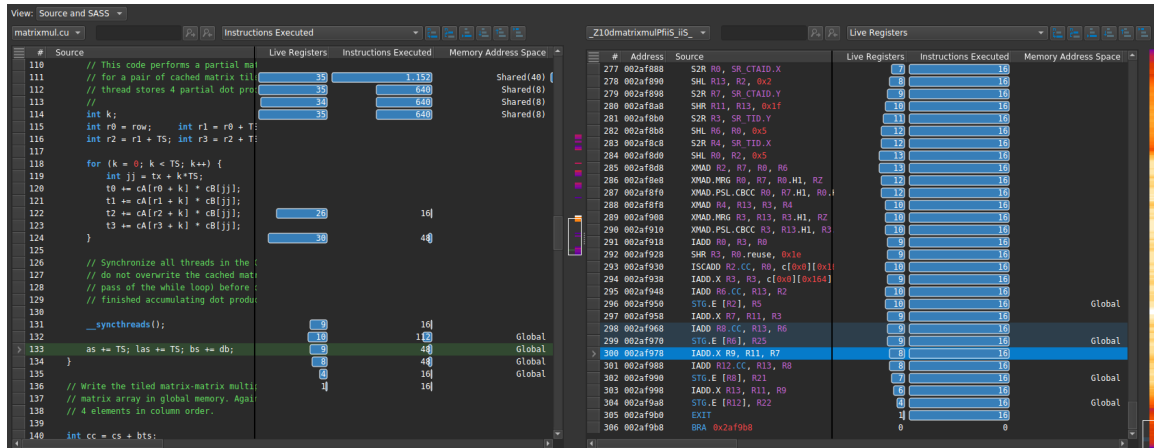
The *View* dropdown can be used to select different code (correlation) options. This includes SASS, PTX and Source (CUDA-C), as well as their combinations. Which options are available depends on the source information embedded into the executable.

The heatmap on the right-hand side of each view can be used to quickly identify locations with high metric values of the currently selected metric in the dropdown. The heatmap uses a black-body radiation color scale where black denotes the lowest mapped value and white the highest, respectively. The current scale is shown when clicking and holding the heatmap with the right mouse button.

If a view contains multiple source files or functions, [+] and [-] buttons are shown. These can be used to expand or collapse the view, thereby showing or hiding the file or function content except for its header. If collapsed, all metrics are shown aggregated to provide a quick overview.

Views allow you to fix columns to not move out of view when scrolling horizontally. By default, the *Source* column is fixed to the left, enabling easy inspection of all metrics correlated to a source line. To change fixing of columns, select the triangle icon in the respective column header.

Pre-Defined Source Metrics

▸ **Live Registers**

Number of registers that need to be kept valid by the compiler. A high value indicates that many registers are required at this code location, potentially increasing the register pressure and the maximum number of register required by the kernel.

▸ **Sampling Data (All)**

The number of samples from the Statistical Sampler at this program location.

▸ **Sampling Data (No Issue)**

The number of samples from the Statistical Sampler at this program location on cycles the warp scheduler issued no instructions. Note that *(No Issue)* samples may

be taken on a different profiling pass than *(All)* samples mentioned above, so their values do not strictly correlate.

This metric is only available on devices with compute capability 7.0 or higher.

▸ **(Predicated-On Thread) Instructions Executed**

Number of times the source (instruction) was executed (and any potential predicate was true, respectively).

▸ **Information on Memory Operation**

This includes *Memory Address Space*, *Memory Access Operation*, *Memory Access Size*, *Memory L1 Transactions Shared*, *Memory L2 Transactions Global* and *Memory L2 Transactions Local*.

▸ **Individual Sampling Data Metrics**

All *stall_** metrics show the information combined in *Sampling Data* individually. See Statistical Sampler for their description.

▸ See the *Customization Guide* on how to add additional metrics for this view.

6.2.5. Comments Page

The *Comments* page aggregates all section comments in a single view and allows the user to edit those comments on any launch instance or section, as well as on the overall report. Comments are persisted with the report. If a section comment is added, the comment icon of the respective section in the Details Page will be highlighted.

6.2.6. Raw Page

The *Raw* page shows a list of all collected metrics with their units per profiled kernel launch. It can be exported, for example, to CSV format for further analysis. The page features a filter edit to quickly find specific metrics. You can transpose the table of kernels and metrics by using the *Transpose* button.

6.3. Metrics and Units

Numeric metric values are shown in various places in the report, including the header and tables and charts on most pages. NVIDIA Nsight Compute supports various ways to display those metrics and their values.

When available and applicable to the UI component, metrics are shown along with their unit. This is to make it apparent if a metric represents cycles, threads, bytes/s, and so on. The unit will normally be shown in rectangular brackets, e.g. `Metric Name [bytes] 128`.

By default, units are scaled automatically so that metric values are shown with a reasonable order of magnitude. Units are scaled using their SI-factors, i.e. byte-based units are scaled using a factor of 1000 and the prefixes K, M, G, etc. Time-based units are also scaled using a factor of 1000, with the prefixes n, u and m. This scaling can be disabled in the Profile options.

Metrics which could not be collected are shown as `n/a` and assigned a warning icon. If the metric floating point value is out of the regular range (i.e. `nan` (Not a number) or `inf` (infinite)), they are also assigned a warning icon. The exception are metrics for which these values are expected and which are white-listed internally.

Chapter 7.
BASELINES

NVIDIA Nsight Compute supports diffing collected results across one or multiple reports using Baselines. Each result in any report can be promoted to a baseline. This causes metric values from all results in all reports to show the difference to the baseline. If multiple baselines are selected simultaneously, metric values are compared to the average across all current baselines. Note that currently, baselines are not stored with a report and are only available as long as the same NVIDIA Nsight Compute instance is open.
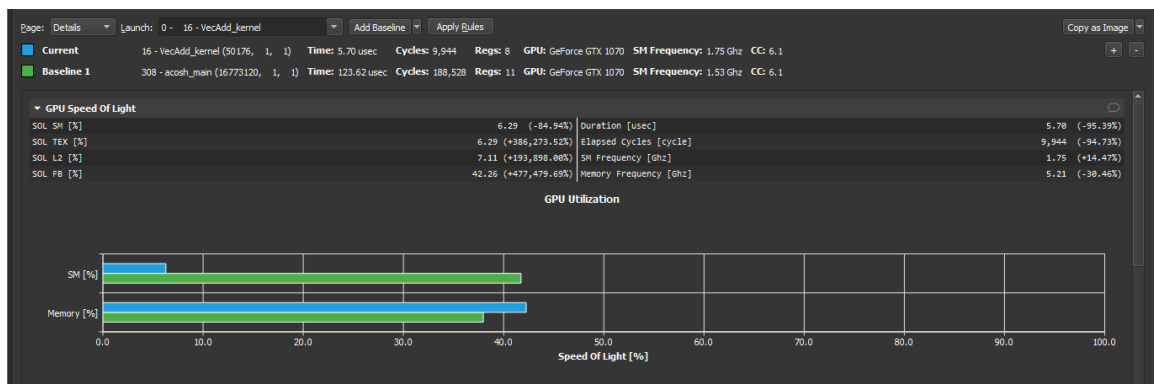
Select *Add Baseline* to promote the current result in focus to become a baseline. If a baseline is set, most metrics on the Details Page, Raw Page and Summary Page show two values: the current value of the result in focus, and the corresponding value of the baseline or the percentage of change from the corresponding baseline value. Use the *Clear Baselines* entry from the dropdown button, the Profile menu, or the corresponding toolbar button to remove all baselines.

If multiple baselines are selected, each metric will show the following notation:

```
<focus value> (<difference to baselines average [%]>, z=<standard score>@<number
 of values>)
```

The standard score is the difference between the current value and the average across all baselines, normalized by the standard deviation. If the number of metric values contributing to the standard score equals the number of results (current and all baselines), the @<number of values> notation is omitted.
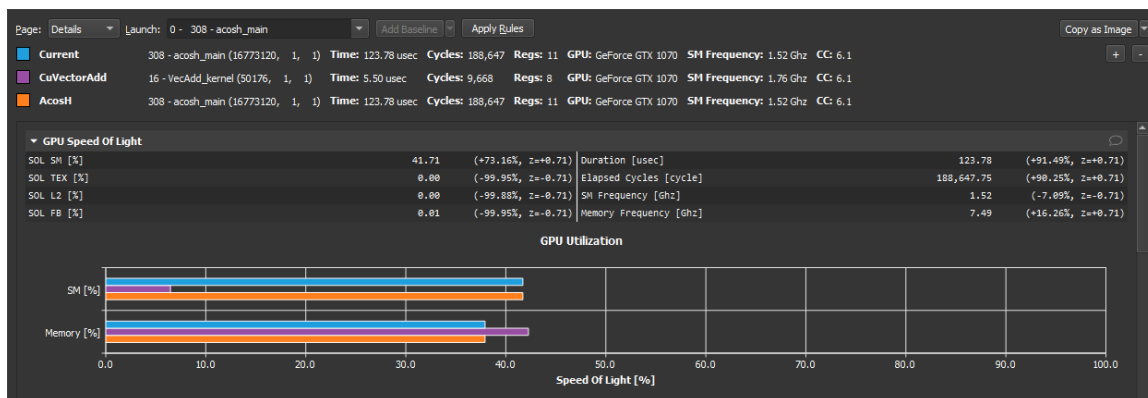
Hovering the mouse over a baseline name allows the user to edit the displayed name. Hovering over the baseline color icon allows the user to remove this specific baseline from the list.

Chapter 8.
OPTIONS

NVIDIA Nsight Compute options can be accessed via the main menu under *Tools > Options*. All options are persisted on disk and available the next time NVIDIA Nsight Compute is launched. When an option is changed from its default setting, its label will become bold. You can use the *Restore Defaults* button to restore all options to their default values.
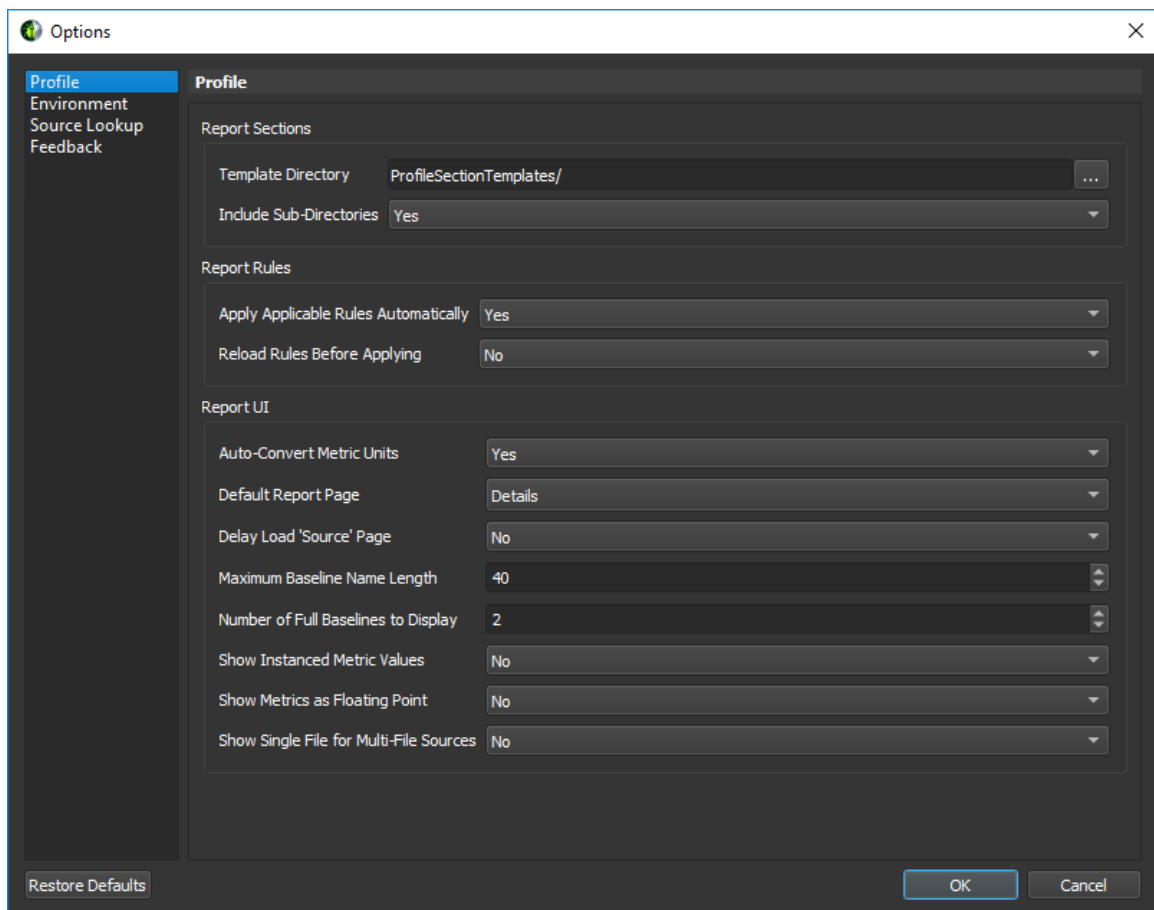
8.1. Profile

Table 1   NVIDIA Nsight Compute Profile Options

Name	Description	Values
Sections Directory	Directory from which to import section files and rules. Relative paths are with respect to the NVIDIA Nsight Compute installation directory.	
Include Sub-Directories	Recursively include section files and rules from sub-directories.	Yes (Default)/No
Apply Applicable Rules Automatically	Automatically apply active and applicable rules.	Yes (Default)/No
Reload Rules Before Applying	Force a rule reload before applying the rule to ensure changes in the rule script are recognized.	Yes/No (Default)
Auto-Convert Metric Units	Auto-adjust displayed metric units and values (e.g. Bytes to KBytes).	Yes (Default)/No
Default Report Page	The report page to show when a report is generated or opened.	▸ Session ▸ Summary ▸ Details (Default) ▸ Source ▸ Comments ▸ Raw ▸ Nvtx
Delay Load Source Page	Delays loading the content of the Source report page until the page becomes visible. Avoids processing costs and memory overhead until the page is opened.	Yes (Default)/No
Maximum Baseline Name Length	The maximum length of baseline names.	1..N
Number of Full Baselines to Display	Number of baselines to display in the report header with all details in addition to the current result.	0..N
Show Instanced Metric Values	Show the individual values of instanced metrics in tables.	Yes/No (Default)
Show Metrics As Floating Point	Show all numeric metrics as floating-point numbers.	Yes/No (Default)
Show Single File For Multi-File Sources	Shows a single file in each Source page view, even for multi-file sources.	Yes/No (Default)

8.2. Environment

Table 2   NVIDIA Nsight Compute Environment Options

Name	Description	Values
Color Theme	The currently selected UI color theme.	‣ Dark (Default) ‣ Light
Default Document Folder	Directory where documents unassociated with a project will be saved.	
At Startup	What to do when NVIDIA Nsight Compute is launched.	‣ Show welcome page (Default) ‣ Show quick launch dialog ‣ Load last project/Show empty environment

8.3. Source Lookup

Table 3   NVIDIA Nsight Compute Source Lookup Options

Name	Description	Values
Program Source Locations	Set program source search paths. These paths are used to resolve CUDA-C source files on the Source page if the respective file cannot be found in its original location.	

Chapter 9.
PROJECTS

NVIDIA Nsight Compute uses *Project Files* to group and organize profiling reports. At any given time, only one project can be open in NVIDIA Nsight Compute. Collected reports are automatically assigned to the current project. Reports stored on disk can be assigned to a project at any time. In addition to profiling reports, related files such as notes or source code can be associated with the project for future reference.

Note that only references to reports or other files are saved in the project file. Those references can become invalid, for example when associated files are deleted, removed or not available on the current system, in case the project file was moved itself.

NVIDIA Nsight Compute uses the `nsight-cuproj` file extension for project files.

9.1. Project Dialogs

▸ **New Project**

Creates a new project. The project must be given a name, which will also be used for the project file. You can select the location where the project file should be saved on disk. Select whether a new directory with the project name should be created in that location.

9.2. Project Explorer

The *Project Explorer* window allows you to inspect and manage the current project. It shows the project name as well as all *Items* (profile reports and other files) associated with it. Right-click on any entry to see further actions, such as adding, removing or grouping items. Type in the *Search project* toolbar at the top to filter the currently shown entries.
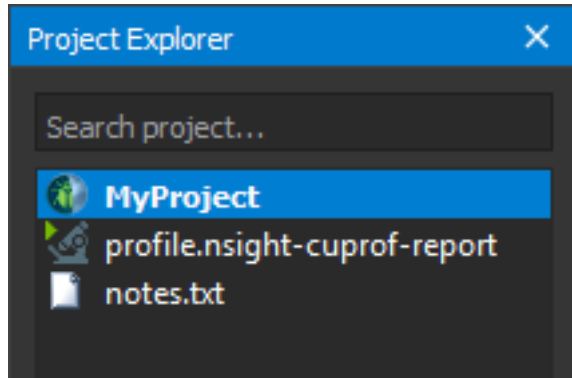
Appendix A.
STATISTICAL SAMPLER

NVIDIA Nsight Compute supports periodic sampling of the warp program counter and warp scheduler state on desktop devices of compute capability 6.1 and above.

At a fixed interval of cycles, the sampler in each streaming multiprocessor selects an active warp and outputs the program counter and the warp scheduler state. The tool selects the minimum interval for the device. On small devices, this can be every 32 cycles. On larger chips with more multiprocessors, this may be 2048 cycles. The sampler selects a random active warp. On the same cycle the scheduler may select a different warp to issue.

Table 4 Warp Scheduler States

State	Hardware Support	Description
Allocation Stall	5.2-6.1	Warp was stalled waiting for a branch to resolve, waiting for all memory operations to retire, or waiting to be allocated to the micro-scheduler.
Barrier	5.2+	Warp was stalled waiting for sibling warps at a CTA barrier.
Dispatch Stall	5.2+	Warp was stalled waiting on a dispatch stall.
Drain	5.2+	Warp was stalled waiting after EXIT for all memory instructions to complete so that warp resources can be freed.
IMC Miss	5.2+	Warp was stalled waiting for an immediate constant cache (IMC) miss.
LG Throttle	7.0+	Warp was stalled waiting for the L1 instruction queue for local and global (LG) memory operations to be not full.
Long Scoreboard	5.2+	Warp was stalled waiting for a scoreboard dependency on a L1TEX (local, global, surface, tex) operation.
Math Pipe Throttle	5.2+	Warp was stalled waiting for the execution pipe to be available.
Membar	5.2+	Warp was stalled waiting on a memory barrier.
MIO Throttle	5.2+	Warp was stalled waiting for the MIO (memory input/output) instruction queue to be not full.

State	Hardware Support	Description
Misc	5.2+	Warp was stalled for a miscellaneous hardware reason.
No Instructions	5.2+	Warp was stalled waiting to be selected to fetch an instruction or waiting on an instruction cache miss.
Not Selected	5.2+	Warp was stalled waiting for the microscheduler to select the warp to issue.
Selected	5.2+	Warp was selected by the microscheduler and issued an instruction.
Short Scoreboard	5.2+	Warp was stalled waiting for a scoreboard dependency on a MIO (memory input/output) operation (not to L1TEX).
Sleeping	7.0+	Warp was stalled due to all threads in the warp being in the blocked, yielded, or sleep state.
Tex Throttle	5.2+	Warp was stalled waiting for the L1 instruction queue for texture operations to be not full.
Wait	5.2+	Warp was stalled waiting on a fixed latency execution dependency.

Appendix B.
SECTIONS AND RULES

Table 5 Default Sections

Identifier	Filename	Description
ComputeWorkloadAnalysis	ComputeWorkloadAnalysis	Detailed analysis of the compute resources of the streaming multiprocessors (SM), including the achieved instructions per clock (IPC) and the utilization of each available pipeline. Pipelines with very high utilization might limit the overall performance.
ComputeWorkloadAnalysis	ComputeWorkloadAnalysis	Detailed analysis of the compute resources of the streaming multiprocessors (SM), including the achieved instructions per clock (IPC) and the utilization of each available pipeline. Pipelines with very high utilization might limit the overall performance.
InstructionStats	InstructionStatistics.section	Statistics of the executed low-level assembly instructions (SASS). The instruction mix provides insight into the types and frequency of the executed instructions. A narrow mix of instruction types implies a dependency on few instruction pipelines, while others remain unused. Using multiple pipelines allows hiding latencies and enables parallel execution.
LaunchStats	LaunchStatistics.section	Summary of the configuration used to launch the kernel. The launch configuration defines the size of the kernel grid, the division of the grid into blocks, and the GPU resources needed to execute the kernel. Choosing an efficient launch configuration maximizes device utilization.
MemoryWorkloadAnalysis	MemoryWorkloadAnalysis	Detailed analysis of the memory resources of the GPU. Memory can become a limiting factor for the overall kernel performance when fully utilizing the involved hardware units (Mem Busy), exhausting the available communication bandwidth between those units (Max Bandwidth), or by reaching the maximum throughput of issuing memory instructions (Mem Pipes Busy). Depending on the limiting factor, the memory chart and tables allow to identify the exact bottleneck in the memory system.
Occupancy	Occupancy.section	Occupancy is the ratio of the number of active warps per multiprocessor to the maximum number of possible active warps. Another way to view occupancy is the percentage

Identifier	Filename	Description
		of the hardware's ability to process warps that is actively in use. Higher occupancy does not always result in higher performance, however, low occupancy always reduces the ability to hide latencies, resulting in overall performance degradation. Large discrepancies between the theoretical and the achieved occupancy during execution typically indicates highly imbalanced workloads.
SchedulerStats	SchedulerStatistics.section	Summary of the activity of the schedulers issuing instructions. Each scheduler maintains a pool of warps that it can issue instructions for. The upper bound of warps in the pool (Theoretical Warps) is limited by the launch configuration. On every cycle each scheduler checks the state of the allocated warps in the pool (Active Warps). Active warps that are not stalled (Eligible Warps) are ready to issue their next instruction. From the set of eligible warps the scheduler selects a single warp from which to issue one or more instructions (Issued Warp). On cycles with no eligible warps, the issue slot is skipped and no instruction is issued. Having many skipped issue slots indicates poor latency hiding.
SourceCounters	SourceCounters.section	Source metrics, including PC sampling information.
SpeedOfLight	SpeedOfLight.section	High-level overview of the utilization for compute and memory resources of the GPU. For each unit, the Speed Of Light (SOL) reports the achieved percentage of utilization with respect to the theoretical maximum.
WarpStateStats	WarpStateStatistics.section	Analysis of the states in which all warps spent cycles during the kernel execution. The warp states describe a warp's readiness or inability to issue its next instruction. The warp cycles per instruction define the latency between two consecutive instructions. The higher the value, the more warp parallelism is required to hide this latency. For each warp state, the chart shows the average number of cycles spent in that state per issued instruction. Stalls are not always impacting the overall performance nor are they completely avoidable. Only focus on stall reasons if the schedulers fail to issue every cycle.

Appendix C.
COPYRIGHT AND LICENSES

NVIDIA CORPORATION

NVIDIA SOFTWARE LICENSE AGREEMENT

IMPORTANT - READ BEFORE DOWNLOADING, INSTALLING, COPYING OR USING THE LICENSED SOFTWARE

This Software License Agreement ("SLA"), made and entered into as of the time and date of click through action ("Effective Date"), is a legal agreement between you and NVIDIA Corporation ("NVIDIA") and governs the use of the NVIDIA computer software and the documentation made available for use with such NVIDIA software. By downloading, installing, copying, or otherwise using the NVIDIA software and/or documentation, you agree to be bound by the terms of this SLA. If you do not agree to the terms of this SLA, do not download, install, copy or use the NVIDIA software or documentation. IF YOU ARE ENTERING INTO THIS SLA ON BEHALF OF A COMPANY OR OTHER LEGAL ENTITY, YOU REPRESENT THAT YOU HAVE THE LEGAL AUTHORITY TO BIND THE ENTITY TO THIS SLA, IN WHICH CASE "YOU" WILL MEAN THE ENTITY YOU REPRESENT. IF YOU DON'T HAVE SUCH AUTHORITY, OR IF YOU DON'T ACCEPT ALL THE TERMS AND CONDITIONS OF THIS SLA, THEN NVIDIA DOES NOT AGREE TO LICENSE THE LICENSED SOFTWARE TO YOU, AND YOU MAY NOT DOWNLOAD, INSTALL, COPY OR USE IT.

1. LICENSE.

1.1 <u>License Grant.</u> Subject to the terms of the AGREEMENT, NVIDIA hereby grants you a non-exclusive, non-transferable license, without the right to sublicense (except as expressly set forth in a Supplement), during the applicable license term unless earlier terminated as provided below, to have Authorized Users install and use the Software, including modifications (if expressly permitted in a Supplement), in accordance with the Documentation. You are only licensed to activate and use Licensed Software for which you a have a valid license, even if during the download or installation you are presented with other product options. No Orders are binding on NVIDIA until accepted by NVIDIA. Your Orders are subject to the AGREEMENT.

<u>SLA Supplements:</u> Certain Licensed Software licensed under this SLA may be subject to additional terms and conditions that will be presented to you in a Supplement for acceptance prior to the delivery of such Licensed Software under this SLA and the

applicable Supplement. Licensed Software will only be delivered to you upon your acceptance of all applicable terms.

1.2 <u>Limited Purpose Licenses.</u> If your license is provided for one of the purposes indicated below, then notwithstanding contrary terms in Section 1.1 or in a Supplement, such licenses are for internal use and do not include any right or license to sub-license and distribute the Licensed Software or its output in any way in any public release, however limited, and/or in any manner that provides third parties with use of or access to the Licensed Software or its functionality or output, including (but not limited to) external alpha or beta testing or development phases. Further:

(i) <u>Evaluation License.</u> You may use evaluation licenses solely for your internal evaluation of the Licensed Software for broader adoption within your Enterprise or in connection with a NVIDIA product purchase decision, and such licenses have an expiration date as indicated by NVIDIA in its sole discretion (or ninety days from the date of download if no other duration is indicated).

(ii) <u>Educational/Academic License.</u> You may use educational/academic licenses solely for educational purposes and all users must be enrolled or employed by an academic institution. If you do not meet NVIDIA's academic program requirements for educational institutions, you have no rights under this license.

(iii) <u>Test/Development License.</u> You may use test/development licenses solely for your internal development, testing and/or debugging of your software applications or for interoperability testing with the Licensed Software, and such licenses have an expiration date as indicated by NVIDIA in its sole discretion (or one year from the date of download if no other duration is indicated).

1.3 <u>Pre-Release Licenses.</u> With respect to alpha, beta, preview, and other pre-release Software and Documentation ("Pre-Release Licensed Software") delivered to you under the AGREEMENT you acknowledge and agree that such Pre-Release Licensed Software (i) may not be fully functional, may contain errors or design flaws, and may have reduced or different security, privacy, accessibility, availability, and reliability standards relative to commercially provided NVIDIA software and documentation, and (ii) use of such Pre-Release Licensed Software may result in unexpected results, loss of data, project delays or other unpredictable damage or loss. THEREFORE, PRE-RELEASE LICENSED SOFTWARE IS NOT INTENDED FOR USE, AND SHOULD NOT BE USED, IN PRODUCTION OR BUSINESS-CRITICAL SYSTEMS. NVIDIA has no obligation to make available a commercial version of any Pre-Release Licensed Software and NVIDIA has the right to abandon development of Pre-Release Licensed Software at any time without liability.

1.4 <u>Enterprise and Contractor Usage.</u> You may allow your Enterprise employees and Contractors to access and use the Licensed Software pursuant to the terms of the AGREEMENT solely to perform work on your behalf, provided further that with respect to Contractors: (i) you obtain a written agreement from each Contractor which contains terms and obligations with respect to access to and use of Licensed Software no less protective of NVIDIA than those set forth in the AGREEMENT, and (ii) such Contractor's access and use expressly excludes any sublicensing or distribution rights for the Licensed Software. You are responsible for the compliance with the terms and conditions of the AGREEMENT by your Enterprise and Contractors. Any act or omission that, if committed by you, would constitute a breach of the AGREEMENT shall

be deemed to constitute a breach of the AGREEMENT if committed by your Enterprise or Contractors.

1.5 <u>Services.</u> Except as expressly indicated in an Order, NVIDIA is under no obligation to provide support for the Licensed Software or to provide any patches, maintenance, updates or upgrades under the AGREEMENT. Unless patches, maintenance, updates or upgrades are provided with their separate governing terms and conditions, they constitute Licensed Software licensed to you under the AGREEMENT.

2. LIMITATIONS.

2.1 <u>License Restrictions.</u> Except as expressly authorized in the AGREEMENT, you agree that you will not (nor authorize third parties to): (i) copy and use Software that was licensed to you for use in one or more NVIDIA hardware products in other unlicensed products (provided that copies solely for backup purposes are allowed); (ii) reverse engineer, decompile, disassemble (except to the extent applicable laws specifically require that such activities be permitted) or attempt to derive the source code, underlying ideas, algorithm or structure of Software provided to you in object code form; (iii) sell, transfer, assign, distribute, rent, loan, lease, sublicense or otherwise make available the Licensed Software or its functionality to third parties (a) as an application services provider or service bureau, (b) by operating hosted/virtual system environments, (c) by hosting, time sharing or providing any other type of services, or (d) otherwise by means of the internet; (iv) modify, translate or otherwise create any derivative works of any Licensed Software; (v) remove, alter, cover or obscure any proprietary notice that appears on or with the Licensed Software or any copies thereof; (vi) use the Licensed Software, or allow its use, transfer, transmission or export in violation of any applicable export control laws, rules or regulations; (vii) distribute, permit access to, or sublicense the Licensed Software as a stand-alone product; (viii) bypass, disable, circumvent or remove any form of copy protection, encryption, security or digital rights management or authentication mechanism used by NVIDIA in connection with the Licensed Software, or use the Licensed Software together with any authorization code, serial number, or other copy protection device not supplied by NVIDIA directly or through an authorized reseller; (ix) use the Licensed Software for the purpose of developing competing products or technologies or assisting a third party in such activities; (x) use the Licensed Software with any system or application where the use or failure of such system or application can reasonably be expected to threaten or result in personal injury, death, or catastrophic loss including, without limitation, use in connection with any nuclear, avionics, navigation, military, medical, life support or other life critical application ("Critical Applications"), unless the parties have entered into a Critical Applications agreement; (xi) distribute any modification or derivative work you make to the Licensed Software under or by reference to the same name as used by NVIDIA; or (xii) use the Licensed Software in any manner that would cause the Licensed Software to become subject to an Excluded License. Nothing in the AGREEMENT shall be construed to give you a right to use, or otherwise obtain access to, any source code from which the Software or any portion thereof is compiled or interpreted. You acknowledge that NVIDIA does not design, test, manufacture or certify the Licensed Software for use in the context of a Critical Application and NVIDIA shall not be liable to you or any third party, in whole or in part, for any claims or damages arising from such use. You agree to defend, indemnify and hold harmless NVIDIA and its Affiliates, and their respective employees, contractors, agents, officers and directors, from and

against any and all claims, damages, obligations, losses, liabilities, costs or debt, fines, restitutions and expenses (including but not limited to attorney's fees and costs incident to establishing the right of indemnification) arising out of or related to you and your Enterprise, and their respective employees, contractors, agents, distributors, resellers, end users, officers and directors use of Licensed Software outside of the scope of the AGREEMENT or any other breach of the terms of the AGREEMENT.

2.2 <u>Third Party License Obligations.</u> The Licensed Software may come bundled with, or otherwise include or be distributed with, third party software licensed by an NVIDIA supplier and/or open source software provided under an open source license (collectively, "Third Party Software"). Notwithstanding anything to the contrary herein, Third Party Software is licensed to you subject to the terms and conditions of the software license agreement accompanying such Third Party Software whether in the form of a discrete agreement, click-through license, or electronic license terms accepted at the time of installation and any additional terms or agreements provided by the third party licensor ("Third Party License Terms"). Use of the Third Party Software by you shall be governed by such Third Party License Terms, or if no Third Party License Terms apply, then the Third Party Software is provided to you as-is, without support or warranty or indemnity obligations, for use in or with the Licensed Software and not otherwise used separately. Copyright to Third Party Software is held by the copyright holders indicated in the Third Party License Terms.

<u>Audio/Video Encoders and Decoders.</u> You acknowledge and agree that it is your sole responsibility to obtain any additional third party licenses required to make, have made, use, have used, sell, import, and offer for sale your products or services that include or incorporate any Third Party Software and content relating to audio and/or video encoders and decoders from, including but not limited to, Microsoft, Thomson, Fraunhofer IIS, Sisvel S.p.A., MPEG-LA, and Coding Technologies as NVIDIA does not grant to you under the AGREEMENT any necessary patent or other rights with respect to audio and/or video encoders and decoders.

2.3 <u>Limited Rights.</u> Your rights in the Licensed Software are limited to those expressly granted under the AGREEMENT and no other licenses are granted whether by implication, estoppel or otherwise. NVIDIA reserves all rights, title and interest in and to the Licensed Software not expressly granted under the AGREEMENT.

3. CONFIDENTIALITY. Neither party will use the other party's Confidential Information, except as necessary for the performance of the AGREEMENT, nor will either party disclose such Confidential Information to any third party, except to personnel of NVIDIA and its Affiliates, you, your Enterprise, your Enterprise Contractors, and each party's legal and financial advisors that have a need to know such Confidential Information for the performance of the AGREEMENT, provided that each such personnel, employee and Contractors are subject to a written agreement that includes confidentiality obligations consistent with those set forth herein. Each party will use all reasonable efforts to maintain the confidentiality of all of the other party's Confidential Information in its possession or control, but in no event less than the efforts that it ordinarily uses with respect to its own Confidential Information of similar nature and importance. The foregoing obligations will not restrict either party from disclosing the other party's Confidential Information or the terms and conditions of the AGREEMENT as required under applicable securities regulations or pursuant to the order or requirement of a court, administrative agency, or other governmental body,

provided that the party required to make such disclosure (i) gives reasonable notice to the other party to enable it to contest such order or requirement prior to its disclosure (whether through protective orders or otherwise), (ii) uses reasonable effort to obtain confidential treatment or similar protection to the fullest extent possible to avoid such public disclosure, and (iii) discloses only the minimum amount of information necessary to comply with such requirements.

NVIDIA Confidential Information under the AGREEMENT includes output from Licensed Software developer tools identified as "Pro" versions, where the output reveals functionality or performance data pertinent to NVIDIA hardware or software products.

4. OWNERSHIP. You are not obligated to disclose to NVIDIA any modifications that you, your Enterprise or your Contractors make to the Licensed Software as permitted under the AGREEMENT. As between the parties, all modifications are owned by NVIDIA and licensed to you under the AGREEMENT unless otherwise expressly provided in a Supplement. The Licensed Software and all modifications owned by NVIDIA, and the respective Intellectual Property Rights therein, are and will remain the sole and exclusive property of NVIDIA or its licensors. You shall not engage in any act or omission that would impair NVIDIA's and/or its licensors' Intellectual Property Rights in the Licensed Software or any other materials, information, processes or subject matter proprietary to NVIDIA. NVIDIA's licensors are intended third party beneficiaries with the right to enforce provisions of the AGREEMENT with respect to their Confidential Information and/or Intellectual Property Rights.

5. FEEDBACK. You may, but you are not obligated, to provide Feedback to NVIDIA. You hereby grant NVIDIA and its Affiliates a perpetual, non-exclusive, worldwide, irrevocable license to use, reproduce, modify, license, sublicense (through multiple tiers of sublicensees), distribute (through multiple tiers of distributors) and otherwise commercialize any Feedback that you voluntarily provide without the payment of any royalties or fees to you. NVIDIA has no obligation to respond to Feedback or to incorporate Feedback into the Licensed Software.

6. NO WARRANTIES. THE LICENSED SOFTWARE AND ANY CONFIDENTIAL INFORMATION AND/OR SERVICES ARE PROVIDED BY NVIDIA "AS IS" AND "WITH ALL FAULTS," AND NVIDIA AND ITS AFFILIATES EXPRESSLY DISCLAIM ALL WARRANTIES OF ANY KIND OR NATURE, WHETHER EXPRESS, IMPLIED OR STATUTORY, INCLUDING, BUT NOT LIMITED TO, ANY WARRANTIES OF OPERABILITY, CONDITION, VALUE, ACCURACY OF DATA, OR QUALITY, AS WELL AS ANY WARRANTIES OF MERCHANTABILITY, SYSTEM INTEGRATION, WORKMANSHIP, SUITABILITY, FITNESS FOR A PARTICULAR PURPOSE, TITLE, NON-INFRINGEMENT, OR THE ABSENCE OF ANY DEFECTS THEREIN, WHETHER LATENT OR PATENT. NO WARRANTY IS MADE ON THE BASIS OF TRADE USAGE, COURSE OF DEALING OR COURSE OF TRADE. WITHOUT LIMITING THE FOREGOING, NVIDIA AND ITS AFFILIATES DO NOT WARRANT THAT THE LICENSED SOFTWARE OR ANY CONFIDENTIAL INFORMATION AND/OR SERVICES PROVIDED UNDER THE AGREEMENT WILL MEET YOUR REQUIREMENTS OR THAT THE OPERATION THEREOF WILL BE UNINTERRUPTED OR ERROR-FREE, OR THAT ALL ERRORS WILL BE CORRECTED.

7. LIMITATION OF LIABILITY. TO THE MAXIMUM EXTENT PERMITTED BY LAW, NVIDIA AND ITS AFFILIATES SHALL NOT BE LIABLE FOR ANY SPECIAL,

INCIDENTAL, PUNITIVE OR CONSEQUENTIAL DAMAGES, OR ANY LOST PROFITS, LOSS OF USE, LOSS OF DATA OR LOSS OF GOODWILL, OR THE COSTS OF PROCURING SUBSTITUTE PRODUCTS, ARISING OUT OF OR IN CONNECTION WITH THE AGREEMENT OR THE USE OR PERFORMANCE OF THE LICENSED SOFTWARE AND ANY CONFIDENTIAL INFORMATION AND/OR SERVICES PROVIDED UNDER THE AGREEMENT, WHETHER SUCH LIABILITY ARISES FROM ANY CLAIM BASED UPON BREACH OF CONTRACT, BREACH OF WARRANTY, TORT (INCLUDING NEGLIGENCE), PRODUCT LIABILITY OR ANY OTHER CAUSE OF ACTION OR THEORY OF LIABILITY. IN NO EVENT WILL NVIDIA'S AND ITS AFFILIATES TOTAL CUMULATIVE LIABILITY UNDER OR ARISING OUT OF THE AGREEMENT EXCEED THE NET AMOUNTS RECEIVED BY NVIDIA OR ITS AFFILIATES FOR YOUR USE OF THE PARTICULAR LICENSED SOFTWARE DURING THE TWELVE (12) MONTHS BEFORE THE LIABILITY AROSE or up to US $10.00 if you acquired the Licensed Software for no charge). THE NATURE OF THE LIABILITY, THE NUMBER OF CLAIMS OR SUITS OR THE NUMBER OF PARTIES WITHIN YOUR ENTERPRISE THAT ACCEPTED THE TERMS OF THE AGREEMENT SHALL NOT ENLARGE OR EXTEND THIS LIMIT. THE FOREGOING LIMITATIONS SHALL APPLY REGARDLESS OF WHETHER NVIDIA, ITS AFFILIATES OR ITS LICENSORS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES AND REGARDLESS OF WHETHER ANY REMEDY FAILS ITS ESSENTIAL PURPOSE. YOU ACKNOWLEDGE THAT NVIDIA'S OBLIGATIONS UNDER THE AGREEMENT ARE FOR THE BENEFIT OF YOU ONLY. The disclaimers, exclusions and limitations of liability set forth in the AGREEMENT form an essential basis of the bargain between the parties, and, absent any such disclaimers, exclusions or limitations of liability, the provisions of the AGREEMENT, including, without limitation, the economic terms, would be substantially different.

8. TERM AND TERMINATION.

8.1 <u>AGREEMENT, Licenses and Services.</u> This SLA shall become effective upon the Effective Date, each Supplement upon their acceptance, and both this SLA and Supplements shall continue in effect until your last access or use of the Licensed Software and/or services hereunder, unless earlier terminated as provided in this "Term and Termination" section. Each Licensed Software license ends at the earlier of (a) the expiration of the applicable license term, or (b) termination of such license or the AGREEMENT. Each service ends at the earlier of (x) the expiration of the applicable service term, (y) termination of such service or the AGREEMENT, or (z) expiration or termination of the associated license and no credit or refund will be provided upon the expiration or termination of the associated license for any service fees paid.

8.2 <u>Termination and Effect of Expiration or Termination.</u> NVIDIA may terminate the AGREEMENT in whole or in part: (i) if you breach any term of the AGREEMENT and fail to cure such breach within thirty (30) days following notice thereof from NVIDIA (or immediately if you violate NVIDIA's Intellectual Property Rights); (ii) if you become the subject of a voluntary or involuntary petition in bankruptcy or any proceeding relating to insolvency, receivership, liquidation or composition for the benefit of creditors, if that petition or proceeding is not dismissed with prejudice within sixty (60) days after filing, or if you cease to do business; or (iii) if you commence or participate in any legal proceeding against NVIDIA, with respect to the Licensed Software that is the subject of the proceeding during the pendency of such legal proceeding. If you or your authorized

NVIDIA reseller fail to pay license fees or service fees when due then NVIDIA may, in its sole discretion, suspend or terminate your license grants, services and any other rights provided under the AGREEMENT for the affected Licensed Software, in addition to any other remedies NVIDIA may have at law or equity. Upon any expiration or termination of the AGREEMENT, a license or a service provided hereunder, (a) any amounts owed to NVIDIA become immediately due and payable, (b) you must promptly discontinue use of the affected Licensed Software and/or service, and (c) you must promptly destroy or return to NVIDIA all copies of the affected Licensed Software and all portions thereof in your possession or control, and each party will promptly destroy or return to the other all of the other party's Confidential Information within its possession or control. Upon written request, you will certify in writing that you have complied with your obligations under this section. Upon expiration or termination of the AGREEMENT all provisions survive except for the license grant provisions.

9. CONSENT TO COLLECTION AND USE OF INFORMATION.

You hereby agree and acknowledge that the Software may access and collect data about your Enterprise computer systems as well as configures the systems in order to (a) properly optimize such systems for use with the Software, (b) deliver content through the Software, (c) improve NVIDIA products and services, and (d) deliver marketing communications. Data collected by the Software includes, but is not limited to, system (i) hardware configuration and ID, (ii) operating system and driver configuration, (iii) installed applications, (iv) applications settings, performance, and usage data, and (iv) usage metrics of the Software. To the extent that you use the Software, you hereby consent to all of the foregoing, and represent and warrant that you have the right to grant such consent. In addition, you agree that you are solely responsible for maintaining appropriate data backups and system restore points for your Enterprise systems, and that NVIDIA will have no responsibility for any damage or loss to such systems (including loss of data or access) arising from or relating to (a) any changes to the configuration, application settings, environment variables, registry, drivers, BIOS, or other attributes of the systems (or any part of such systems) initiated through the Software; or b) installation of any Software or third party software patches initiated through the Software. In certain systems you may change your system update preferences by unchecking "Automatically check for updates" in the "Preferences" tab of the control panel for the Software.

In connection with the receipt of the Licensed Software or services you may receive access to links to third party websites and services and the availability of those links does not imply any endorsement by NVIDIA. NVIDIA encourages you to review the privacy statements on those sites and services that you choose to visit so that you can understand how they may collect, use and share personal information of individuals. NVIDIA is not responsible or liable for: (i) the availability or accuracy of such links; or (ii) the products, services or information available on or through such links; or (iii) the privacy statements or practices of sites and services controlled by other companies or organizations.

To the extent that you or members of your Enterprise provide to NVIDIA during registration or otherwise personal data, you acknowledge that such information will be collected, used and disclosed by NVIDIA in accordance with NVIDIA's privacy policy, available at URL http://www.nvidia.com/object/privacy_policy.html.

10. GENERAL.

This SLA, any Supplements incorporated hereto, and Orders constitute the entire agreement of the parties with respect to the subject matter hereto and supersede all prior negotiations, conversations, or discussions between the parties relating to the subject matter hereto, oral or written, and all past dealings or industry custom. Any additional and/or conflicting terms and conditions on purchase order(s) or any other documents issued by you are null, void, and invalid. Any amendment or waiver under the AGREEMENT must be in writing and signed by representatives of both parties.

The AGREEMENT and the rights and obligations thereunder may not be assigned by you, in whole or in part, including by merger, consolidation, dissolution, operation of law, or any other manner, without written consent of NVIDIA, and any purported assignment in violation of this provision shall be void and of no effect. NVIDIA may assign, delegate or transfer the AGREEMENT and its rights and obligations hereunder, and if to a non-Affiliate you will be notified.

Each party acknowledges and agrees that the other is an independent contractor in the performance of the AGREEMENT, and each party is solely responsible for all of its employees, agents, contractors, and labor costs and expenses arising in connection therewith. The parties are not partners, joint ventures or otherwise affiliated, and neither has any authority to make any statements, representations or commitments of any kind to bind the other party without prior written consent.

Neither party will be responsible for any failure or delay in its performance under the AGREEMENT (except for any payment obligations) to the extent due to causes beyond its reasonable control for so long as such force majeure event continues in effect.

The AGREEMENT will be governed by and construed under the laws of the State of Delaware and the United States without regard to the conflicts of law provisions thereof and without regard to the United Nations Convention on Contracts for the International Sale of Goods. The parties consent to the personal jurisdiction of the federal and state courts located in Santa Clara County, California. You acknowledge and agree that a breach of any of your promises or agreements contained in the AGREEMENT may result in irreparable and continuing injury to NVIDIA for which monetary damages may not be an adequate remedy and therefore NVIDIA is entitled to seek injunctive relief as well as such other and further relief as may be appropriate. If any court of competent jurisdiction determines that any provision of the AGREEMENT is illegal, invalid or unenforceable, the remaining provisions will remain in full force and effect. Unless otherwise specified, remedies are cumulative.

The Licensed Software has been developed entirely at private expense and is "commercial items" consisting of "commercial computer software" and "commercial computer software documentation" provided with RESTRICTED RIGHTS. Use, duplication or disclosure by the U.S. Government or a U.S. Government subcontractor is subject to the restrictions set forth in the AGREEMENT pursuant to DFARS 227.7202-3(a) or as set forth in subparagraphs (c)(1) and (2) of the Commercial Computer Software - Restricted Rights clause at FAR 52.227-19, as applicable. Contractor/manufacturer is NVIDIA, 2788 San Tomas Expressway, Santa Clara, CA 95051.

You acknowledge that the Licensed Software described under the AGREEMENT is subject to export control under the U.S. Export Administration Regulations (EAR) and

economic sanctions regulations administered by the U.S. Department of Treasury's Office of Foreign Assets Control (OFAC). Therefore, you may not export, reexport or transfer in-country the Licensed Software without first obtaining any license or other approval that may be required by BIS and/or OFAC. You are responsible for any violation of the U.S. or other applicable export control or economic sanctions laws, regulations and requirements related to the Licensed Software. By accepting this SLA, you confirm that you are not a resident or citizen of any country currently embargoed by the U.S. and that you are not otherwise prohibited from receiving the Licensed Software.

Any notice delivered by NVIDIA to you under the AGREEMENT will be delivered via mail, email or fax. Please direct your legal notices or other correspondence to NVIDIA Corporation, 2788 San Tomas Expressway, Santa Clara, California 95051, United States of America, Attention: Legal Department.

GLOSSARY OF TERMS

Certain capitalized terms, if not otherwise defined elsewhere in this SLA, shall have the meanings set forth below:

a. "Affiliate" means any legal entity that Owns, is Owned by, or is commonly Owned with a party. "Own" means having more than 50% ownership or the right to direct the management of the entity.

b. "AGREEMENT" means this SLA and all associated Supplements entered by the parties referencing this SLA.

c. "Authorized Users" means your Enterprise individual employees and any of your Enterprise's Contractors, subject to the terms of the "Enterprise and Contractors Usage" section.

d. "Confidential Information" means the Licensed Software (unless made publicly available by NVIDIA without confidentiality obligations), and any NVIDIA business, marketing, pricing, research and development, know-how, technical, scientific, financial status, proposed new products or other information disclosed by NVIDIA to you which, at the time of disclosure, is designated in writing as confidential or proprietary (or like written designation), or orally identified as confidential or proprietary or is otherwise reasonably identifiable by parties exercising reasonable business judgment, as confidential. Confidential Information does not and will not include information that: (i) is or becomes generally known to the public through no fault of or breach of the AGREEMENT by the receiving party; (ii) is rightfully known by the receiving party at the time of disclosure without an obligation of confidentiality; (iii) is independently developed by the receiving party without use of the disclosing party's Confidential Information; or (iv) is rightfully obtained by the receiving party from a third party without restriction on use or disclosure.

e. "Contractor" means an individual who works primarily for your Enterprise on a contractor basis from your secure network.

f. "Documentation" means the NVIDIA documentation made available for use with the Software, including (without limitation) user manuals, datasheets, operations instructions, installation guides, release notes and other materials provided to you under the AGREEMENT.

g. "Enterprise" means you or any company or legal entity for which you accepted the terms of this SLA, and their subsidiaries of which your company or legal entity owns more than fifty percent (50%) of the issued and outstanding equity.

h. "Excluded License" includes, without limitation, a software license that requires as a condition of use, modification, and/or distribution that software be (i) disclosed or distributed in source code form; (ii) licensed for the purpose of making derivative works; or (iii) redistributable at no charge.

i. "Feedback" means any and all suggestions, feature requests, comments or other feedback regarding the Licensed Software, including possible enhancements or modifications thereto.

j. "Intellectual Property Rights" means all patent, copyright, trademark, trade secret, trade dress, trade names, utility models, mask work, moral rights, rights of attribution or integrity service marks, master recording and music publishing rights, performance rights, author's rights, database rights, registered design rights and any applications for the protection or registration of these rights, or other intellectual or industrial property rights or proprietary rights, howsoever arising and in whatever media, whether now known or hereafter devised, whether or not registered, (including all claims and causes of action for infringement, misappropriation or violation and all rights in any registrations and renewals), worldwide and whether existing now or in the future.

k. "Licensed Software" means Software, Documentation and all modifications owned by NVIDIA.

l. "Order" means a purchase order issued by you, a signed purchase agreement with you, or other ordering document issued by you to NVIDIA or a NVIDIA authorized reseller (including any on-line acceptance process) that references and incorporates the AGREEMENT and is accepted by NVIDIA.

m. "Software" means the NVIDIA software programs licensed to you under the AGREEMENT including, without limitation, libraries, sample code, utility programs and programming code.

n. "Supplement" means the additional terms and conditions beyond those stated in this SLA that apply to certain Licensed Software licensed hereunder.

MICROSOFT DETOURS

Microsoft Detours is used under the Professional license (http://research.microsoft.com/ en-us/projects/detours/).

NVIDIA agrees to include in all copies of the NVIDIA Applications a proprietary rights notice that includes a reference to Microsoft software being included in such applications. NVIDIA shall not remove or obscure, but shall retain in the Software, any copyright, trademark, or patent notices that appear in the Software.

BREAKPAD

Copyright 2006, Google Inc. All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

▶ Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

▶ Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

▶ Neither the name of Google Inc. nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

PROTOBUF

Copyright 2014, Google Inc. All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

▶ Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.

▶ Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

▶ Neither the name of Google Inc. nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF

THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

Code generated by the Protocol Buffer compiler is owned by the owner of the input file used when generating it. This code is not standalone and requires a support library to be linked with it. This support library is itself covered by the above license.

QTSINGLEAPPLICATION

Copyright 2013 Digia Plc and/or its subsidiary(-ies).

Contact: http://www.qt-project.org/legal

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

- Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
- Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.
- Neither the name of Digia Plc and its Subsidiary(-ies) nor the names of its contributors may be used to endorse or promote products derived from this software without specific prior written permission.

THIS SOFTWARE IS PROVIDED BY THE COPYRIGHT HOLDERS AND CONTRIBUTORS "AS IS" AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE COPYRIGHT OWNER OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

FONT AWESOME

Copyright 2018 FontIcons, Inc.

Contact: https://fontawesome.com/license/free

Some of the icons in NVIDIA Nsight Compute are provided by Font Awesome under the Font Awesome Free License.