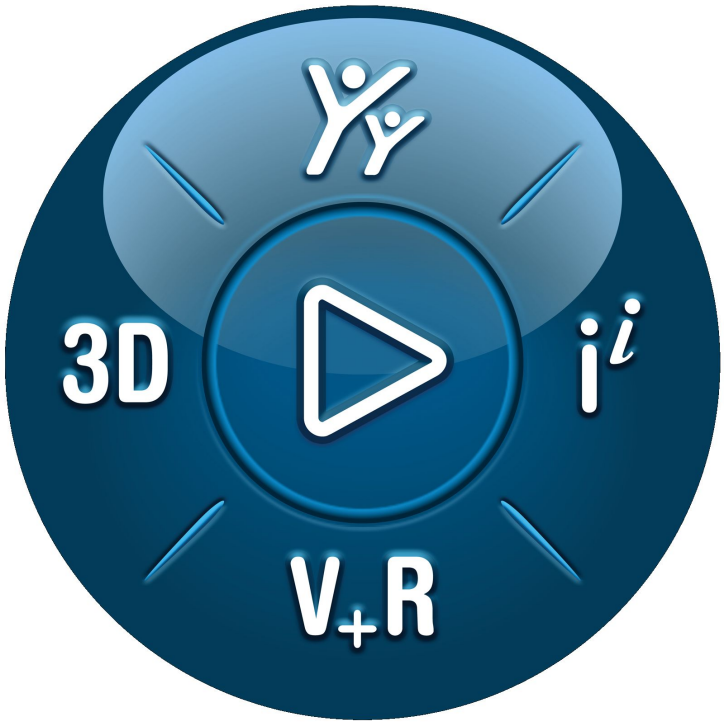
CST Studio Suite® 2022

GPU Computing Guide



3DEXPERIENCE®

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

Copyright 1998-2021 Dassault Systemes Deutschland GmbH.
CST Studio Suite is a Dassault Systèmes product.
All rights reserved.

DASSAULT SYSTEMES | The 3DEXPERIENCE® Company

Contents

1 Nomenclature **4**

2 Supported Hardware **5**
2.1 Supported Solvers and Features for NVIDIA GPUs 5
2.2 Supported Solvers and Features for AMD GPUs 6

3 Operating System Support **6**

4 Licensing **6**

5 Switch On GPU Computing **6**
5.1 Interactive Simulations 6
5.2 Simulations in Batch Mode 7

6 List of supported NVIDIA GPU hardware for CST Studio Suite 2022 **7**

7 Unsupported NVIDIA Hardware **11**

8 NVIDIA Drivers Download and Installation **12**
8.1 GPU Driver Installation 12
8.2 Verifying Correct Installation of GPU Hardware and Drivers 15
8.3 Uninstalling NVIDIA Drivers 16

9 NVIDIA Usage Guidelines **17**
9.1 The Error Correction Code (ECC) Feature 17
9.2 Tesla Compute Cluster (TCC) Mode 19
9.3 Exclusive Compute Mode 20
9.4 Display Link 20
9.5 Combined MPI Computing and GPU Computing 20
9.6 Service User 20
9.7 GPU Computing using Windows Remote Desktop (RDP) 20
9.8 Running Multiple Simulations at the Same Time 21
9.9 Video Card Drivers 21
9.10 Operating Conditions 21
9.11 Latest CST Service Pack 21
9.12 GPU Monitoring/Utilization 22
9.13 Select Subset of Available GPU Cards 23

10 NVIDIA GPU Boost **24**

11 NVIDIA Troubleshooting Tips **27**

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

12 List of supported AMD GPU hardware for CST Studio Suite 2022 **29**
12.1 Supported GPUs: General information 30
12.2 Supported GPUs: Specifications 30

13 Unsupported AMD Hardware **31**

14 AMD Drivers Download and Installation **31**
14.1 GPU Driver Installation 31

15 History of Changes **32**

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

1 Nomenclature

The following section explains the nomenclature used in this document.

`command`	Commands you have to enter either on a command prompt (cmd on MS Windows or your favorite shell on Linux) are typeset using typewriter fonts.
`<...>`	Within commands the sections you should replace according to your environment are enclosed in "`<...>`". For example "`<CST_DIR>`" should be replaced by the directory where you have installed CST Studio Suite (e.g. "`c:\Program Files\CST Studio Suite`").

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

2 Supported Hardware

CST Studio Suite supports Hardware Acceleration for various Solvers and GPUs. In combination with NVIDIA GPUs many different kind of Solvers and GPUs are supported. Please check the tables below for more details.
Starting with CST Studio Suite 2021 also selected AMD GPUs are supported to accelerate the Time Domain Solver only. For more information please contact our local 3DS/SIMULIA support team (https://www.3ds.com/products-services/simulia/locations).

CST Studio Suite currently supports up to 16 GPU devices in a single host system, meaning each number of GPU devices between 1 and 16 is supported.[1]

2.1 Supported Solvers and Features for NVIDIA GPUs

- Time Domain Solver (T-HEX-solver and TLM-solver)
- Integral Equation Solver (direct solver and MLFMM only)
  - GPUs with good double precision performance required
- Multilayer solver (M-solver)
  - GPUs with good double precision performance required
- Particle-In-Cell (PIC-solver)
  - Modulation of External Fields not supported
  - Open Boundaries not supported
- Electrostatic Particle-In-Cell (Es-PIC-solver)
  - Multi-GPU not supported
  - Field-dependent Particle Sources and Particle Interfaces not supported
  - Secondary Emission from Solids not supported
  - Monte-Carlo Collision Models: Excitation and Scattering not supported
  - Particle 2D and Current Density Monitors not supported
  - Sheet Transparency for Particles not supportedT
  - Particle Losses on Solids not supported
  - Periodic Boundary Conditions not supported
- Conjugate Heat Transfer Solver (CHT-solver)

[1] It is strongly recommended to contact CST before purchasing a system with more than four GPU cards to ensure that the hardware is working properly and is configured correctly for CST Studio Suite.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

- Asymptotic Solver (A-solver)
  - GPUs with good double precision performance required
  - On Windows TCC mode is required

2.2 Supported Solvers and Features for AMD GPUs

- Transient HEX Solver (T-HEX-solver)

3 Operating System Support

CST Studio Suite is continuously tested on different operating systems. For a list of supported operating systems please refer to

https://updates.cst.com/downloads/CST-OS-Support.pdf

In general, GPU computing can be used on any of the supported operating systems.

4 Licensing

The GPU computing feature is licensed either by Acceleration Tokens with CST Studio Suite License Model or by SimUnit tokens or credits with the SIMULIA Unified License model. I.e. your license must contain a sufficient amount of Acceleration Tokens or SimUnit tokens/credits, depending on the used hardware configuration and license model, if you want to accelerate your simulations by using a GPU. Please contact your Dassault Systèmes sales representative for further information.

5 Switch On GPU Computing

5.1 Interactive Simulations

GPU Computing needs to be enabled via the acceleration dialog box before running a simulation. To turn on GPU Computing:

1. Open the dialog of the solver.
2. Click on the "Acceleration" button.
3. Switch on "Hardware acceleration" and specify how many GPU devices should be used for this simulation. The specification of the number of devices is per solver (e.g. if DC is used). Please note that the maximum number of GPU devices available for a simulation depends upon the number of tokens in your license.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022



5.2 Simulations in Batch Mode

If you start your simulations in batch mode (e.g. via an external job queuing system) there is a command line switch (`-withgpu`) which can be used to switch on the GPU Computing feature. The command line switch can be used as follows:[2]

In Windows:

```
"<CST_INSTALL_DIR>/CST Design Environment.exe" -m -r -withgpu=<NUMBER_OF_GPUs> "<FULL_PATH_TO_CST_FILE>"
```

In Linux:

```
"<CST_INSTALL_DIR>/cst_design_environment" -m -r -withgpu=<NUMBER_OF_GPUs> "<FULL_PATH_TO_CST_FILE>"
```

6 List of supported NVIDIA GPU hardware for CST Studio Suite 2022

The following tables contain some basic information about NVIDIA GPUs currently supported by the GPU Computing feature of CST Studio Suite, as well as the requirements for the host system equipped with the hardware. To ensure compatibility of GPU hardware and host system please check

https://www.nvidia.com/object/tesla-qualified-servers.html

- Please note that a 64 bit computer architecture is required for GPU Computing.
- CST Studio Suite officially supports the NVIDIA Tesla and Quadro cards listed in the table below. That means that these GPUs are well tested and validated with CST software and you can contact CST support in case you run into any problems.

[2] This example shows the batch mode simulation for the transient solver (`-m -r`). To learn more about the command line switches used by CST Studio Suite please refer to the online help documentation in the section 'General Features', subsection 'Command Line Options'.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

- Please note that cards of different generations (e.g. "Ampere" and "Volta") can't be combined in a single host system for GPU Computing.
- Platform = Servers: These GPUs are only available with a passive cooling system which only provides sufficient cooling if it's used in combination with additional fans. These fans are usually available for server chassis only!
- Platform = Workstations: These GPUs provide active cooling, so they are suitable for workstation computer chassis as well.
- GPUs with reduced FP64 performance (e.g. GPUs where the FP64 performance is less than half the speed of the FP32 performance) are not well suited for solvers which require a high FP64 performance. Pease check chapter 2 Supported Hardware for affected solvers.
- GPUs of Kepler (sm_35 and sm_37) and Maxwell (sm_50) generation are marked as deprecated; they might not be supported in newer Nvidia driver and also newer MWS releases.
- Memory and problem size: For the Time Domain Solver (FIT) the rule of thumb is "available memory divided by 100" (e.g. a GPU with 16 GB of Memory will be sufficient for a simulation with 160 million meshcells. The memory consumption is dependent on the used features, so the maximum simulation size might be more or less than that).

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

Device Name	Generation	Platform	Min. CST Version
Tesla A100-SXM4-80GB	Ampere	Server	2021 SP5
Tesla A100-PCIE-80GB	Ampere	Server	2021 SP5
Tesla A30	Ampere	Server	2021 SP5
Tesla A10 [3]	Ampere	Server	2021 SP5
Tesla A16 [3]	Ampere	Server	2021 SP5
Tesla A40 [3]	Ampere	Server	2021 SP5
RTX A6000 [3]	Ampere	Workstation	2021 SP5
RTX A5000 [3]	Ampere	Workstation	2021 SP5
RTX A4000 [3]	Ampere	Workstation	2021 SP5
A100-SXM4-40GB	Ampere	Server	2021 release
A100-PG509-200	Ampere	Server	2021 release
A100-PCIE-40GB	Ampere	Server	2021 release
Quadro RTX 8000 [3]	Turing	Workstation	2019 SP6
Quadro RTX 6000 [3]	Turing	Workstation	2019 SP6
Quadro RTX 5000 [3]	Turing	Workstation	2019 SP6
Quadro RTX 4000 [3]	Turing	Workstation	2019 SP6
Tesla T4 [3]	Turing	Server	2021
Quadro GV100	Volta	Workstation	2018 SP6
Tesla V100-SXM2-32GB (Chip)	Volta	Server	2018 SP6
Tesla V100-PCIE-32GB	Volta	Server	2018 SP6
Tesla V100-SXM2-16GB (Chip)	Volta	Server	2018 SP1
Tesla V100-PCIE-16GB	Volta	Server	2018 SP1
Tesla P100-SXM2 (Chip)	Pascal	Server	2017 release
Tesla P100-PCIE-16GB	Pascal	Server	2017 release
Tesla P100 16GB	Pascal	Server	2017 release
Tesla P100-PCIE-12GB	Pascal	Servers	2017 SP2
Quadro P6000 [3]	Pascal	Workstation	2017 SP 2
Quadro GP100	Pascal	Workstation	2017 SP2
Tesla P40 [3]	Pascal	Server	2018 release
Tesla P4 [3]	Pascal	Server	2018 release
Tesla M60 [3] [4]	Maxwell	Server/Workst.	2016 SP4
Tesla M40 [3] [4]	Maxwell	Server	2016 SP4
Quadro M6000 24GB [3] [4]	Maxwell	Workstation	2016 SP4
Quadro M6000 [3] [4]	Maxwell	Workstation	2015 SP4
Tesla K80 [4]	Kepler	Server	2014 SP6
Tesla K40 m/c/s/st/d/t [4]	Kepler	Server/Workst.	2013 SP5
Quadro K6000 [4]	Kepler	Workstation	2013 SP4
Tesla K20 m/c/s/X [4]	Kepler	Server/Workst.	2013 release

[3] **Important:** GPU device with reduced double precision performance; please check chapter 2 Supported Hardware for affected solvers.

[4] **Important:** This GPU is marked as deprecated and it might not be supported in upcoming releases.

Device Name	Memory (GB)	Bandwidth (GB/s)	FP32 (TFlops)	FP64 (TFlops)
Tesla A100-SXM4-80GB	80	2000	19.5	9.7
Tesla A100-PCIE-80GB	80	2000	19.5	9.7
Tesla A30	24	933	10.3	5.3
Tesla A10	24	600	31	1
Tesla A16	16 x4	232 x4	8.6 x4	0.271 x4
Tesla A40	48	695	37	1.1
RTX A6000	48	768	40	1.25
RTX A5000	24	768	27.7	0.9
RTX A4000	16	448	19.2	0.6
Tesla A100-SXM4-40GB	40	1550	19.5	9.7
Tesla A100-PG509-200	40	1550	19.5	9.7
Tesla A100-PCIE-40GB	40	1550	19.5	9.7
Quadro RTX 8000	48	672	16	0.5
Quadro RTX 6000	24	672	16	0.5
Quadro RTX 5000	16	448	11	0.35
Quadro RTX 4000	8	416	7.1	0.22
Tesla T4	16	320	8	0.25
Quadro GV100	32	900	14	7
Tesla V100-SXM2-32GB	32	900	15	7.5
Tesla V100-PCIE-32GB	32	900	14	7
Tesla V100-SXM2-16GB	16	900	15	7.5
Tesla V100-PCIE-16GB	16	900	14	7
Tesla P100-SXM2	16	732	10.6	5.3
Tesla P100-PCIE-16GB	16	732	9.3	4.7
Tesla P100 16GB	16	732	9.3	4.7
Tesla P100-PCIE-12GB	12	549	9.3	4.7
Quadro P6000	24	432	12	0.2
Quadro GP100	16	720	10.3	5.2
Tesla P40	24	694	11.8	0.367
Tesla P4	8	192	5.7	0.2
Tesla M60	8	160	4.8	0.150
Tesla M40	12	288	6.8	0.213
Quadro M6000 24GB	24	317	6.8	0.213
Quadro M6000	12	317	6.8	0.213
Tesla K80	12 x2	240 x2	4.1 x2	1.37 x2
Tesla K40 m/c/s/st/d/t	12	288	5	1.7
Quadro K6000	12	288	5	1.7
Tesla K20 m/c/s/X	5	208	3.5	1.1

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

7 Unsupported NVIDIA Hardware

If you have an NVIDIA GPU card which is not supported (please see the list of supported cards in the previous section), but fulfills the requirements below, you may enable it for simulations by following these instructions. Note that using unsupported and untested hardware is not recommended. CST will not provide any support for any problems resulting from using unsupported GPU cards and CST will not run any tests on these GPUs.

The NVIDIA GPUs that fulfill the following requirements may be enabled for simulations:

- capable of running CUDA 11.2 code
- at least 4 GB of free memory

To enable such GPUs, set the environment variable `CST_HWACC_ALLOW_UNVERIFIED_HARDWARE = 1`. Then all eligible installed and visible NVIDIA GPUs will be considered for computation. You can limit the visible devices by using the environment variable `CUDA_VISIBLE_DEVICES`.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

8 NVIDIA Drivers Download and Installation

An appropriate driver is required in order to use the GPU hardware. Please download the driver appropriate to your GPU hardware and operating system from the NVIDIA website. The driver versions listed below are verified for use with our software. Other driver versions provided by NVIDIA might also work but it is highly recommended to use the versions verified by CST.
We recommend the following driver versions for all supported GPU cards:

Windows: Version 462.31

Linux: Version 460.73.01

The recommended drivers will work for all supported Tesla and Quadro GPUs. For Quadro GPUs: Please check the NVIDIA Driver website for a similar Tesla GPU (same generation, e.g. Pascal or Volta) to be able to find the recommended driver.

8.1 GPU Driver Installation

8.1.1 Installation on Windows

After you have downloaded the installer executable please start the installation procedure by double clicking on the installer executable. After a quick series of pop-up windows, the NVIDIA InstallShield Wizard will appear. Press the "Next" button and driver installation will begin (The screen may turn black momentarily.). You may receive a message indicating that the hardware has not passed Windows logo testing. In case you get this warning select "Continue Anyway".
If you are updating from a previously installed NVIDIA driver, it's recommended to select "clean installation" in the NVIDIA Installshield Wizard. This will remove the current driver prior to installing the new driver.
The "Wizard Complete" window will appear as soon as the installation has finished. Select "Yes, I want to restart my computer now" and click the "Finish" button.
It is recommended that you run the HWAccDiagnostics tool after the installation to confirm that the driver has been successfully installed. Please use HWAccDiagnostics_AMD64.exe which can be found in the AMD64 directory of the installation folder.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

8.1.2 Installation on Linux

1. Login on the Linux machine as root.

2. Make sure that the adapter has been recognized by the system using the command
   `/sbin/lspci | grep -i nvidia`
   If you do not see any settings try to update the PCI hardware database of your system using the command
   `/sbin/update-pciids`

3. Stop the X-Server by running in a terminal the command (You may skip this step if you are working on a system without X-server)

   `systemctl isolate multi-user.target`
   (on systems using Systemd)

   `init 3`
   (on systems using SysVinit)

4. Install the NVIDIA graphics driver. Follow the instructions of the setup script. In most cases the installer needs to compile a specific kernel module. If this is the case the gcc compiler and Linux kernel headers need to be available on the machine.

5. Restart the X-server by running the command (You may skip this step if you are working on a system without X-server)

   `systemctl isolate graphical.target`
   (on systems using Systemd)

   `init 5`
   (on systems using SysVinit)

   Note: In case you're using the CST Distributed Computing system and a DC Solver Server is running on the machine where you just installed the driver you need to restart the DC Solver Server as otherwise the GPUs cannot be detected properly.

   Note: The OpenGL libraries should not be installed on a system which has no rendering capabilities (like a pure DC Solver Server or a pure cluster node). This can be accomplished by starting the NVIDIA installer using the option `"-no-opengl-files"`.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

6. You may skip this step if a X-server is installed on your system **and** you are using a NVIDIA graphics adapter (in addition to the GPU Computing devices) in your system. If no X-server is installed on your machine or you don't have an additional NVIDIA graphics adapter, the NVIDIA kernel module will not be loaded automatically. Additionally, the device files for the GPUs will not be generated automatically. The following commands will perform the necessary steps to use the hardware for GPU Computing. It is recommended to append this code to your `rc.local` file such that it is executed automatically during system start.

```
# Load nvidia kernel module
modprobe nvidia

if [ "$?" -eq 0 ]; then

  # Count the number of NVIDIA controllers found.
  N3D=$(/sbin/lspci | grep -i nvidia | grep "3D controller" | wc -l)
  NVGA=$(/sbin/lspci | grep -i nvidia | grep "VGA compatible controller" | wc -l)

  N=$(expr $N3D + $NVGA - 1)
  for i in $(seq 0 $N); do
    mknod -m 666 /dev/nvidia$i c 195 $i;
  done

  mknod -m 666 /dev/nvidiactl c 195 255

fi
```

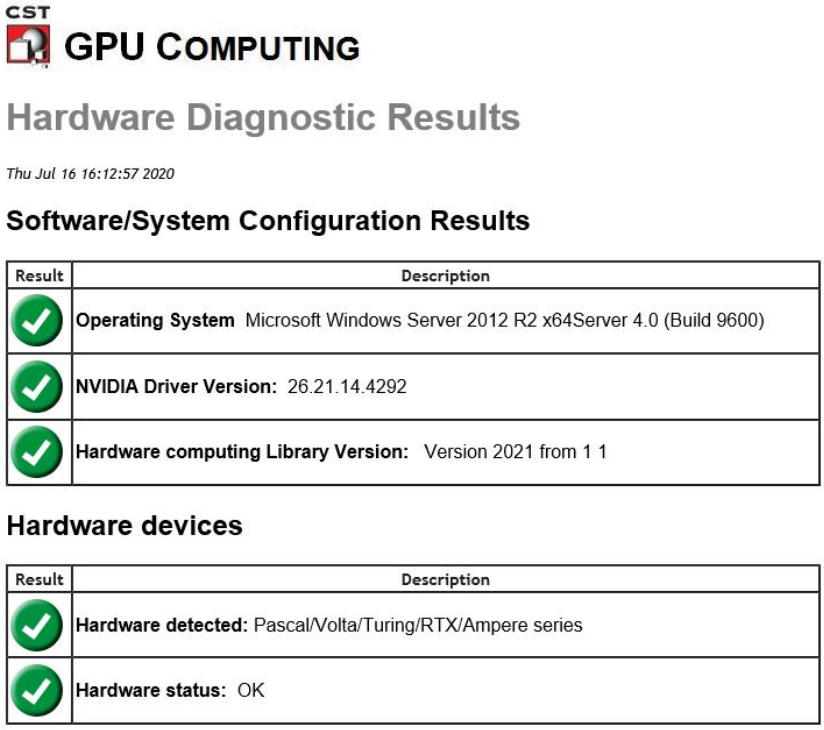
Please note:

- If you encounter problems during restart of the X-server please check chapter 8 "Common Problems" in the file README.txt located at `/usr/share/doc/NVIDIA_GLX-1.0`. Please also consider removing existing sound cards or deactivating onboard sound in the BIOS. Furthermore, make sure you are running the latest BIOS version.
- After installation, if the X system reports an error like `no screen found`, please check Xorg log files in `/var/log`. Open the log files in an editor and search for `"PCI"`. According to the number of hardware cards in your system you will find entries of the following form: `PCI: (0@7:0:0)`. In `/etc/X11`, open the file `xorg.conf` in an editor and search for `"nvidia"`. After the line `BoardName "Quadro M6000"` (or whatever card you are using) insert a new line that reads `BusID "PCI:7:0:0"` according to the entries found in the log files before. Save and close the `xorg.conf` file and type `startx`. If X still refuses to start, try the other entries found in the Xorg log files.
- You need the installation script to uninstall the driver. Thus, if you want to be able to uninstall the NVIDIA software you need to keep the installer script.
- Be aware of the fact that you need to reinstall the NVIDIA drivers if your kernel is updated as the installer needs to compile a new kernel module in this case.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

8.2 Verifying Correct Installation of GPU Hardware and Drivers

As a final test to verify that the GPU hardware has been correctly installed, the following test can be executed: Log in to the machine and execute the HWAccDiagnostics_AMD64 program found in the AMD64 subfolder of your CST installation (Windows) or in the folder LinuxAMD64 on a Linux system. The macro "Check GPU Computing Setup" in the Solver macros performs exactly this check.The output of the tool should look similar to the following picture if the installation was successful.


CST GPU COMPUTING

Hardware Diagnostic Results

Thu Jul 16 16:12:57 2020

Software/System Configuration Results

Result	Description
✓	**Operating System** Microsoft Windows Server 2012 R2 x64Server 4.0 (Build 9600)
✓	**NVIDIA Driver Version:** 26.21.14.4292
✓	**Hardware computing Library Version:** Version 2021 from 1 1

Hardware devices

Result	Description
✓	**Hardware detected:** Pascal/Volta/Turing/RTX/Ampere series
✓	**Hardware status:** OK

Figure 1: Output of HWAccDiagnostics_AMD64.exe tool.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

8.3 Uninstalling NVIDIA Drivers

8.3.1 Uninstall Procedure on MS Windows

To uninstall NVIDIA drivers, select "NVIDIA Drivers" from the "App & features" list and press the "Uninstall" button (see fig. 2). After the uninstall process has finished you will be prompted to reboot.

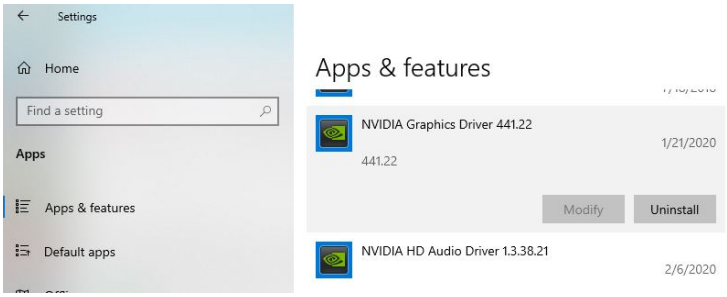


Figure 2: "App & features" dialog on Windows

8.3.2 Uninstall Procedure on Linux

Start the installer with the "`-uninstall`" option. This requires root permissions.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

9 NVIDIA Usage Guidelines

9.1 The Error Correction Code (ECC) Feature

ECC can detect and eventually correct problems caused by faulty GPU memory. Such GPU memory errors typically cause unstable simulations. However, this feature deteriorates the performance of older GPU hardware (all cards of the Fermi, Kepler, and Maxwell series are affected). Therefore, we recommend disabling the feature. If simulations running on GPU hardware become unstable it is recommended to enable ECC temporarily as a diagnostic tool to determine whether the problems are caused by a GPU memory defect. Please also refer to section 11.
The latest NVIDIA GPU hardware (Pascal) has native ECC support with no performance overhead. For those GPUs ECC can't be switched off.
The ECC feature can be managed by using either the NVIDIA Control Panel or the command line tool `nvidia-smi`. Please note, that on Windows 7, Windows Server 2008 R2, and newer version of Windows the following commands have to be run as administrator.

9.1.1 Managing the ECC Feature via Command Line

This procedure works on all supported versions of Windows and on all supported Linux distributions.

The command requires administrator privileges on Windows and root privileges on Linux, respectively.

1. Locate the file nvidia-smi. This file is typically found in `"c:\Program Files\NVIDIA Corporation\NVSMI"` or in `/usr/bin` on Linux.
2. Open up a command prompt/terminal window and navigate to this folder.
3. Execute the following command:
   `nvidia-smi -L`
4. Please note down how many GPUs are found.
5. To disable ECC: Please execute the following command for each of the GPUs:
   `nvidia-smi -i <number_of_the_GPU_card> -e 0`
6. To enable ECC: Please execute the following command for each of the GPUs:
   `nvidia-smi -i <number_of_the_GPU_card> -e 1`
7. Reboot.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

9.1.2 Managing the ECC Feature via NVIDIA Control Panel

This procedure works on all versions of Windows.

1. Start the Control Panel via the Windows start menu.
2. Start the NVIDIA Control Panel.
3. Search for the term "ECC State" in the navigation tree of the dialog and open the "ECC State" page of the dialog by clicking on the tree item.
4. Disable or enable the ECC feature for all Tesla devices (see fig. 3).

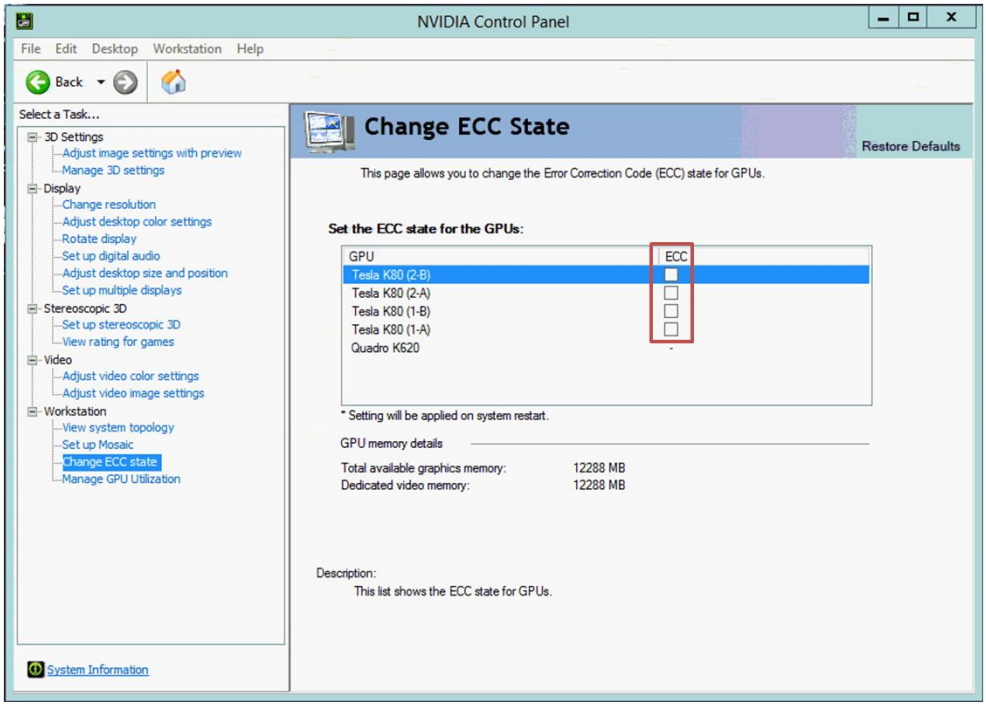


Figure 3: Switch off the ECC feature for all Tesla cards.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

9.2 Tesla Compute Cluster (TCC) Mode (Windows only)

When available, the GPUs have to operate in TCC mode. **Please note that the TCC mode is currently required for A-solver GPU computing.**

9.2.1 Enable the TCC Mode

When available, the GPUs have to operate in TCC mode[5]. Please enable the mode, if not yet enabled.
Please note that the following commands require administrator privileges.

1. Locate the file nvidia-smi. This file is typically found in "c:\Program Files\NVIDIA Corporation\NVSMI".
2. Open up a command prompt and navigate to this folder.
3. Execute the following command:
   `nvidia-smi -L`
4. Please note down how many GPUs are found.
5. For each of the GPUs, please execute the following command:
   `nvidia-smi -i <number_of_the_GPU_card> -dm 1`
6. Reboot.

9.2.2 Disabling the TCC Mode

If available, this feature should always be enabled. However, under certain circumstances you may need to disable this mode.
Please note that the following commands require administrator privileges.

1. Locate the file nvidia-smi. This file is typically found in "c:\Program Files\NVIDIA Corporation\NVSMI".
2. Open up a command prompt and navigate to this folder.
3. Execute the following command:
   `nvidia-smi -L`
4. Please note down how many GPUs are found.
5. For each of the GPUs, please execute the following command:
   `nvidia-smi -i <number_of_the_GPU_card> -dm 0`
6. Reboot.

[5] The TCC Mode is available on all Tesla and on most Quadro cards. This mode is not available for Quadro cards which are connected to a display/monitor.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

9.3 Exclusive Compute Mode

Nvidia GPUs allow to set different compute modes which indicates whether individual or multiple compute applications may run on the GPU.

- "Default" means multiple contexts are allowed per device.
- "Exclusive Process" means only one context is allowed per device, usable from multiple threads at a time.

In case of "Exclusive Process" the user has to make sure that no other process is running on the GPU before he starts a GPU solver run using CST Studio Suite.

9.4 Display Link

Some cards of the Tesla series provide a display link to plug in a monitor. Using this display link has the following implications:

- The TCC Mode of the card cannot be used. This deteriorates the performance.
- GPU Computing can't be used in a remote desktop session.

Because of these limitations we recommend using an additional graphics adapter for the graphics output, or if available, an onboard graphics chipset.

9.5 Combined MPI Computing and GPU Computing (Windows only)

For combined MPI Computing and GPU Computing the TCC mode of the GPU hardware must be enabled (see 9.2).

9.6 Service User (Windows only)

If you are using GPU Computing via the CST Distributed Computing system and your DC Solver Server runs on Windows then the DC Solver Server service must be started using the Local System account (see fig. 4). The CST Studio Suite installer installs the service by default using the correct account.

9.7 GPU Computing using Windows Remote Desktop

For users with a LAN license, GPU Computing using RDP can be used in combination with Tesla or Quadro GPU cards as long as there is no monitor connected to the GPU.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

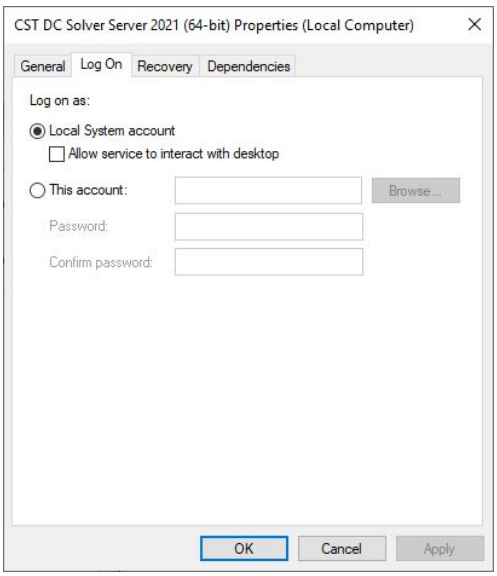


Figure 4: Local System Account.

9.8 Running Multiple Simulations at the Same Time

Running multiple simulations in parallel on **the same** GPU card will deteriorate the performance. Therefore we recommend to run just one simulation at a time. If you have a system with multiple GPU cards and would like to assign simulations to specific GPU cards please refer to section 9.13.

9.9 Video Card Drivers

Please use only the drivers recommended in this document or by the hardware diagnostics tool (See section 8.2). They have been tested for compatibility with CST products.

9.10 Operating Conditions

CST recommends that GPU Computing is operated in a well ventilated temperature controlled area. For more information, please contact your hardware vendor.

9.11 Latest CST Service Pack

Download and install the latest CST Service Pack prior to running a simulation or HWAccDiagnostics.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

9.12 GPU Monitoring/Utilization

Locate the file nvidia-smi. This file is typically found in "`c:\Program Files\NVIDIA Corporation\NVSMI`" on Windows or in `/usr/bin` on Linux. If you start this tool with the command line switch `-l` or `-loop` it will show the utilization and other interesting information such as the temperatures of the GPU cards. The `-l` option makes sure that the tool runs in a loop such that the information gets updated every couple seconds. For more options please run `nvidia-smi -h`. If you want to check the GPU utilization only, you can also run the graphical tool NvGpuUtilization (Windows only). This file is typically found in "`c:\Program Files\NVIDIA Corporation\Control Panel Client`".

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

9.13 Select Subset of Available GPU Cards

If you have multiple GPU cards supported for GPU computing in the same machine you may want to specify the cards visible to the CST software such that your simulations are only started on a subset of the cards. This can be accomplished in two different ways.

9.13.1 Environment Variable `CUDA_VISIBLE_DEVICES`

The environment variable `CUDA_VISIBLE_DEVICES` which contains a comma separated list of GPU IDs will force a process (such as a CST solver) to use the specified subset of GPU cards only).[6] If this variable is set in the environment of the CST software or globally on your system the simulation will be started on the cards listed in the `CUDA_VISIBLE_DEVICES` list only.
Example: Open a shell (cmd on Windows, or bash on Linux) and enter

- `set CUDA_VISIBLE_DEVICES=0`

on Windows or

- `export CUDA_VISIBLE_DEVICES=0`

on Linux to bind all CST solver processes started from this shell to the GPU with ID 0. To make the setting persistent for all CST instances started on the system you may add the variable to the global system environment variables.

9.13.2 Distributed Computing

The CST Distributed Computing (DC) system can be used to assign the GPU cards of a multi-GPU system to different DC Solver Servers. The solver processes executed by a certain DC Solver Server will only be able to access the GPU cards assigned to this Solver Server (see fig. 5). Please refer to the online help documents of CST Studio Suite (section "Simulation Acceleration", subsection "Distributed Computing") to learn more about the setup and configuration of the DC system.

[6] Execute the command `nvidia-smi -L` to get the GPU IDs of your cards.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

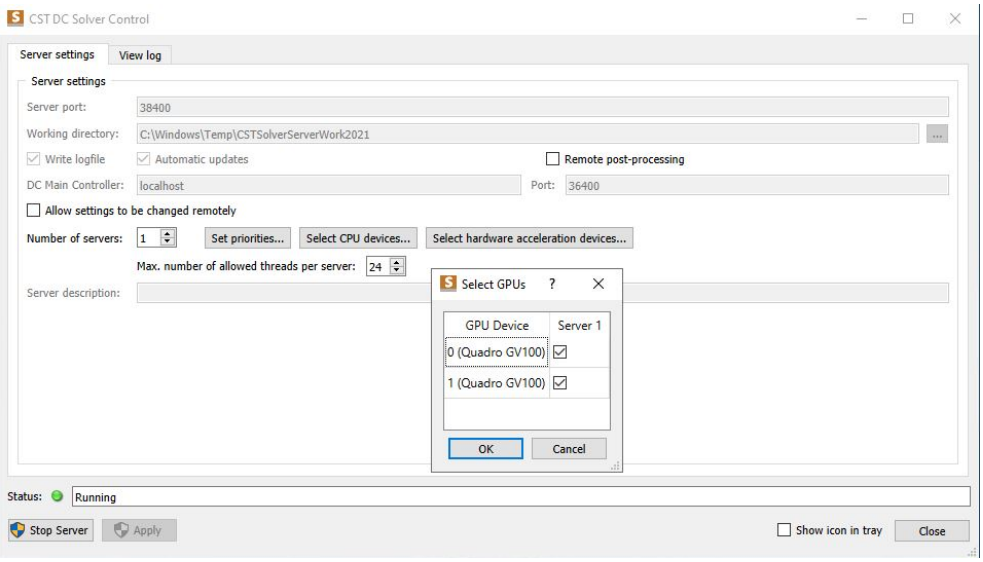


Figure 5: Assignment of GPUs to specific DC Solver Servers.

10 NVIDIA GPU Boost

NVIDIA GPU Boost™is a feature available on the recent NVIDIA Tesla products.
This feature takes advantage of any power and thermal headroom in order to boost performance by increasing the GPU core and memory clock rates. The Tesla GPUs are designed with a specific Thermal Design Power (TDP). Frequently HPC workloads do not come close to reaching this power limit, and therefore have power headroom. A performance improvement can be expected when using the GPU Boost feature on the CST solvers. The Tesla GPUs come with a "Base clock" and several "Boost Clocks" which may be manually selected for compute intensive workloads with available power headroom. The Tesla GPUs give full control to end-users to select one of the core clock frequencies via the NVIDIA System Management Interface (nvidia-smi). For the K40 card, figuring out the right boost clock setting may require some experimentation to see what boost clock works best for a specific workload. NVIDIA GPU Boost on the Tesla K80 is enabled by default and dynamically selects the appropriate GPU clock based on the power headroom.
The GPU Boost feature can be employed by using either the NVIDIA Control Panel or the command line tool nvidia-smi. The nvidia-smi file is typically found in "`c:\Program Files\NVIDIA Corporation\NVSMI`" in Microsoft Windows or `/usr/bin` in Linux.
The following are common commands for setting the GPU Boost feature and checking GPU performance.
To display the current application clock in use execute the following command:

```
nvidia-smi -q -d CLOCK
```

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

```
Default Applications Clocks
    Graphics                    : 745 MHz
    Memory                      : 3004 MHz
Max Clocks
    Graphics                    : 875 MHz
    SM                          : 875 MHz
    Memory                      : 3004 MHz
```

Before making any changes to the clocks, the GPU should be set to Persistence Mode. Persistence mode ensures that the driver stays loaded and does not revert back to the default clock once the application is complete and no CUDA or X applications are running on the GPU.
To enable persistence mode use the following command:

```
nvidia-smi -pm 1
```

To view the clocks that are supported by the Tesla board:

```
nvidia-smi -q -d SUPPORTED_CLOCKS
```

```
GPU 0000:02:00.0
    Supported Clocks
        Memory                  : 3004 MHz
            Graphics            : 875 MHz
            Graphics            : 810 MHz
            Graphics            : 745 MHz
            Graphics            : 666 MHz
        Memory                  : 324 MHz
            Graphics            : 324 MHz
```

Please note that the supported graphics clock rates are tied to a specific memory clock rate so when setting application clocks you must set both the memory clock and the graphics clock[7]. Do this using the `-ac <MEM clock, Graphics clock>` command line option.

```
nvidia-smi -ac 3004,875
```

```
Applications clocks set to "(MEM 3004, SM 875)" for GPU 0000:02:00.0
Applications clocks set to "(MEM 3004, SM 875)" for GPU 0000:03:00.0
Applications clocks set to "(MEM 3004, SM 875)" for GPU 0000:83:00.0
Applications clocks set to "(MEM 3004, SM 875)" for GPU 0000:84:00.0
```

Execute the following command to reset the application clocks back to default settings.

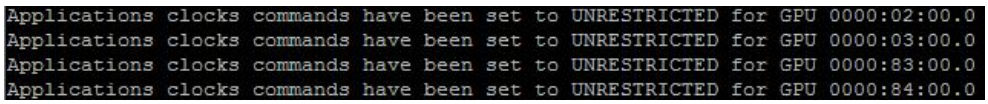
```
nvidia-smi -rac
```

Changing application clocks requires administrative privileges. However, a system administrator can remove this requirement to allow non-admin users to change application clocks by setting the application clock permissions to 'UNRESTRICTED' using the following command:

```
nvidia-smi -acp UNRESTRICTED
```

[7]The memory clock should remain at 3 GHz for the Tesla K40.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022



Please be aware that the application clock setting is a recommendation. If the GPU cannot safely run at the selected clocks, for example due to thermal or power reasons, it will automatically lower the clocks to a safe clock frequency. You can check whether this has occurred by typing the following command while the GPU is active:

```
nvidia-smi -a -d PERFORMANCE
```

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

11 NVIDIA Troubleshooting Tips

The following troubleshooting tips may help if you experience problems.

- If you experience problems during the installation of the NVIDIA driver on the Windows operating system please try to boot Windows in "safe mode" and retry the driver installation.
- NVIDIA DGX A100 and NVIDIA HGX A100 8-GPU server-systems: "cudaGetDeviceCount returned errorcode 802: system not yet initialized"
  Please make sure nVidia fabric manager is installed and activated.
  - install nvidia DCGM.
  - terminate the nv-hostengine* first in order to enable fabric manager.
  - sudo nv-hostengine -t
  - Enable Fabric manager: service nvidia-fabricmanager start
  - document for reference: https://docs.nvidia.com/datacenter/dcgm/latest/dcgm-user-guide/overview.html

  This solution is taken from: https://www.supermicro.com/support/faqs/faq.cfm?faq=31029
- If you have a multi-GPU setup (4 or 8 GPUs) and you encounter an "out-of-memory" problem please set the environment variable
  `CUDA_DEVICE_MAX_CONNECTIONS=1`.
  In case your host system has at least 512 GB of RAM please also check out the following website: GPU addressing capabilities.
- Please note that CST Studio Suite cannot run on GPU devices when they are in "exclusive mode". Please refer to section 9.3 on how to disable this mode.
- If you are using an external GPU device ensure that the PCI connector cable is securely fastened to the host interface card.
- Uninstall video drivers for your existing graphics adapter prior to installing the new graphics adapter.
- Make sure the latest motherboard BIOS is installed on your machine. Please contact your hardware vendor support for more information about the correct BIOS version for your machine.
- Use the HWAccDiagnostics tool to find out whether your GPU hardware and your driver is correctly recognized.
- GPU temperatures are crucial for the performance and overheating of GPU devices can lead to hardware failures. Please refer to section 9.12 for details.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

- A faulty GPU device can be responsible for seemingly random solver crashes. To ensure that your GPU is working properly please run tests provided by the HWAccDiagnostics tool found in the CST installation. Examples of usage:
  - `HWAccDiagnostics_AMD64 -runstresstest -duration=2000 -percentage=99` will run a memory test on all GPUs one by one (helps to identify GPU hardware problems usually related to a specific GPU).
  - `HWAccDiagnostics_AMD64 -runstresstest2 -duration=2000 -percentage=99` will run a simulation on all GPUs concurrently first, followed by running the same simulation on each GPU seperately (helps to identify thermal issues).
  - `HWAccDiagnostics_AMD64 -runstresstest2 -duration=2000 -percentage=99 -deviedID=0` will run the same simulation as before on device ID 0 only (helps to verify problems usually related to a specific GPU).
  - Please run `HWAccDiagnostics_AMD64 -h` to see all possible options.
- In case simulations are getting unstable when running on the GPU it's recommended to check the GPU memory by switching on the ECC feature on the GPU (see 9.1).
- If a GPU is not recognized during the installation please check if `Memory Mapped I/O above 4GB` is enabled in your bios settings.
- CUDA error 11: If you encounter an error message similar to
  `CUDA error 11: invalid argument, dev: 1`
  you can usually fix this problem by using CUDA_VISIBLE_DEVICES so that only devices are visble for CUDA which will be used for simulations. Please refer to section 9.13 for details.
- TCC mode: GPUs running in WDDM instead of TCC mode might eventually fail during memory allocations. It is highly recommended to put all GPUs in TCC mode to avoid these kind of problems (see 9.2).
- Please execute the nvidia-smi tool found in
  `"c:\Program Files\NVIDIA Corporation\NVSMI"` on Windows and in `"/usr/bin"` on Linux in order to find out whether the GPUs are correctly recognized by the GPU driver.

If you need further assistance please contact your local 3DS/SIMULIA support team (https://www.3ds.com/products-services/simulia/locations).

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

12 List of supported AMD GPU hardware for CST Studio Suite 2022

The following tables contain some basic information about AMD GPUs currently supported by the GPU Computing feature of CST Studio Suite, as well as the requirements for the host system equipped with the hardware.

- Please note that currently only the Time Domain Solver (FIT) supports AMD GPUs.
- Please note that a 64 bit computer architecture is required for GPU Computing.
- CST Studio Suite officially supports the AMD GPUs listed in the table below. That means that these GPUs are tested and validated with CST software and you can contact CST support in case you run into any problems.
- Please do not combine different GPUs in a single host system for GPU Computing.
- Platform = Servers: These GPUs are only available with a passive cooling system which only provides sufficient cooling if it's used in combination with additional fans. These fans are usually available for server chassis only!
- Platform = Workstations: These GPUs provide active cooling, so they are suitable for workstation computer chassis as well.
- GPUs with reduced FP64 performance (e.g. GPUs where the FP64 performance is less than half the speed of the FP32 performance) are not well suited for solvers which require a high FP64 performance. Pease check chapter 2 Supported Hardware for affected solvers.
- Memory and problem size: For the Time Domain Solver (FIT) the rule of thumb is "available memory divided by 100" (e.g. a GPU with 16 GB of Memory will be sufficient for a simulation with 160 million meshcells. The memory consumption is dependent on the used features, so the maximum simulation size might be more or less than that).

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

12.1 Supported GPUs: General information

Device Name	Generation	Platform	Min. CST Version
Radeon VII	Radeon	Workstation	2022 SP1
Radeon VII pro	Radeon	Workstation	2022 SP1
Instinct MI50	Radeon	Server	2022 SP1
Instinct MI100	Radeon	Server	2022 SP1
WX 9100 [3]	Radeon	Workstation	2022

12.2 Supported GPUs: Specifications

Device Name	Memory (GB)	Bandwidth (GB/s)	FP32 (TFlops)	FP64 (TFlops)
Radeon VII	16	1024	13.44	3.36
Radeon VII pro	16	1024	13.06	6.528
Instinct MI50	16	1024	13.41	6.705
Instinct MI100	32	1229	23.07	11.54
WX 9100	16	483	12.29	0.8

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

13 Unsupported AMD Hardware

If you have an AMD GPU card which is not supported (please see the list of supported cards in the previous section), but fulfills the requirements below, you may enable it for simulations by following these instructions. Note that using unsupported and untested hardware is not recommended. CST will not provide any support for any problems resulting from using unsupported GPU cards and CST will not run any tests on these GPUs.

The AMD GPUs that fulfill the following requirements may be enabled for simulations:

- at least 4 GB of free memory

To enable such GPUs, set the environment variable `CST_HWACC_ALLOW_UNVERIFIED_HARDWARE = 1`; after that all eligible installed and visible AMD GPUs will be considered for computation.

14 AMD Drivers Download and Installation

An appropriate driver is required in order to use the GPU hardware. Please download the driver appropriate to your GPU hardware and operating system from the AMD website. The driver versions listed below are verified for use with our software. Other driver versions provided by AMD might also work but it is highly recommended to use the versions verified by CST.
We recommend the following driver versions for all supported AMD GPU cards:

Windows: **Version 21.Q2.1**

Linux: **Version 21.Q2**

The recommended drivers will work for all supported AMD GPUs.

14.1 GPU Driver Installation

Please follow the instructions on the AMD website for the various supported OS systems.

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022

15 History of Changes

The following changes have been applied to the document in the past.

Date	Description
July 19 2021	initial version of this document

3DS.COM/SIMULIA © Dassault Systèmes GPU Computing Guide 2022