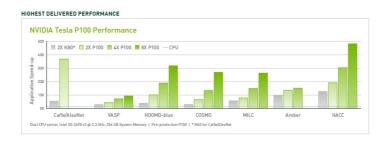
CUDA® is a parallel computing platform and programming model developed by NVIDIA for general computing on graphical processing units (GPUs). With CUDA, developers are able to dramatically speed up computing applications by harnessing the power of GPUs.

In GPU-accelerated applications, the sequential part of the workload runs on the CPU – which is optimized for single-threaded performance – while the compute intensive portion of the application runs on thousands of GPU cores in parallel. When using CUDA, developers program in popular languages such as C, C++, Fortran, Python and MATLAB and express parallelism through extensions in the form of a few basic keywords.

The CUDA Toolkit from NVIDIA provides everything you need to develop GPU-accelerated applications. The CUDA Toolkit includes GPU-accelerated libraries, a compiler, development tools and the CUDA runtime.



Thousands of applications developed with CUDA have been deployed to GPUs in embedded systems, work stations, datacenters and in the cloud.

















See More Applications

CUDA serves as a common platform across all NVIDIA GPU families so you can deploy and scale your application across GPU configurations.









The first GPUs were designed as graphics accelerators, becoming more programmable over the 90s, culminating in NVIDIA's first GPU in 1999. Researchers and scientists rapidly began to apply the excellent floating point performance of this GPU for general purpose computing. In 2003, a team of researchers led by lan Buck unveiled Brook, the first widely adopted programming model to extend C with data-parallel constructs. Ian Buck later joined NVIDIA and led the launch of CUDA in 2006, the world's first solution for general-computing on GPUs.

Since its inception, the CUDA ecosystem has grown rapidly to include software development tools, services and partner-based solutions. The CUDA Toolkit includes libraries, debugging and optimization tools, a compiler and a runtime library to deploy your application. You'll also find code samples, programming guides, user manuals, API references and other documentation to help you get started.

Libraries





Math Library



nvGRAPH





cuFFT



NCCL

Tools and Integrations







OpenACC



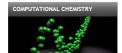




See More Libraries

See More Tools

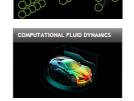
CUDA accelerates applications across a wide range of domains from image processing, to deep learning, numerical analytics and computational science.













More Applications

Get started with CUDA by downloading the CUDA Toolkit and exploring introductory resources including videos, code samples, hands-on labs and webinars.

Download Now>

Get Started with CUDA:

Accelerated Computing News

Read mare

Accelerated Computing - May 25 2018

GPU-accelerated Supercomputer Helps Automatically Detect Skin Cancer

Researchers at the University of Queensland in Australia are using GPU-accelerated supercomputers and deep learning to diagnose skin cancer from histology slides with the same accuracy as a trained pathologist.

"Pathologists do an incredible job

Read more



Features - May 25 2018

Google Researchers Use AI to Bring Still Photos to Life

Researchers from Google developed a deep learningbased system that can create short video clips from still images shot on stereo cameras, VR cameras, and dual lens cameras, such as an iPhone 7 or X.

Read more



Artificial Intelligence - May 25 2018

Al Helps Map Every Building in the U.S.

Researchers at the Oak Ridge National Laboratory in Tennessee developed a deep learning-based system that can map every building in the contiguous United States from satellite imagery.

Read more



Game Development - May 24 2018

Helping New Players Join the Fortnite Battle with NVIDIA Highlights

THE SWEET TASTE OF VICTORY #1 Victory Royale.
When this phrase splashes onto your screen, you know you've arrived

Read more

Parallel ForAll Blog

Accelerated Computing - May 17 2018 Making Containers Easier with HPC Container Maker

Today's groundbreaking scientific discoveries are taking place in high performance computing (HPC) data

Read more

Accelerated Computing - May 07 2018

Volta Tensor Core GPU Achieves New Al Performance Milestones

Artificial intelligence powered by deep learning now solves challenges once thought impossible, such as computers understanding and conversing in natural speech and autonomous driving.

Read more

Accelerated Computing - May 02 2018 NWL Accelerates Machine Learning on Video Datasets

Loading data onto GPUs for training has historically been a minor issue for most deep learning practitioners. Data read from a local spinning hard drive or NAS device would be preprocessed on the CPU, then shipped to the GPU for training.

Read more

→

Game Development - Apr 25 2018 NVIDIA SMP Assist API for VR Programming

The NVIDIA SMP (simultaneous multi-projection) Assist NVAPI driver extension is a simple method for integrating Multi-Res Shading and Lens-Matched Shading into a VR application.

Read mon

NVIDIA DEVELOPER

COMPUTEWORKS

GAMEWORKS

JETPACK

DESIGNWORKS

GET STARTED

About CUDA
Parallel Computin
CUDA Toolkit
CUDACasts

LEARN MORE

Training and Coursewar Tools and Ecosystem Academic Collaboration Documentation

GET INVOLVED

Forums

Developer Blog

Contact Us

Conversely © 2018 NVIDIA Corporation

al Information

Privacy Policy