GPU Accelerated Computing with Python

Python is one of the most popular programming languages today for science, engineering, data analytics and deep learning applications. However, as an interpreted language, it has been considered too slow for high-performance computing. That has changed with CUDA Python from Continuum Analytics.

With CUDA Python, using the Numba (https://docs.continuum.io/numbapro/quickstart) Python compiler, you get the best of both worlds: rapid iterative development with Python combined with the speed of a compiled language targeting both CPUs and NVIDIA GPUs.

1

SETUP CUDA PYTHON

To run CUDA Python, you will need the CUDA Toolkit (/cuda-toolkit) installed on a system with CUDA capable GPUs.

If you do not have a CUDA-capable GPU (/cuda-gpus), you can access one of the thousands of GPUs available from cloud service providers including Amazon AWS, Microsoft Azure and IBM SoftLayer. The NVIDIA-maintained CUDA Amazon Machine Image (https://aws.amazon.com/marketplace/pp/B01LZMLK1K) (AMI) on AWS, for example, comes pre-installed with CUDA and is available for use today.

NVIDIA AMIs on AWS (https://aws.amazon.com/marketplace/pp/B01LZMLK1K)

Use this guide (https://developer.nvidia.com/how-to-cuda-c-cpp) for easy steps to install CUDA. To setup CUDA Python, first install the Anaconda python distribution (https://www.continuum.io/downloads). Then install the latest version of the Numba package (https://docs.continuum.io/anaconda/packages/pkg-docs). You can find detailed installation instructions (https://docs.continuum.io/docs_oss/numba/user/installing) in the Numba documentation.

Or, watch the short video below and follow along.

2

YOUR FIRST CUDA PYTHON PROGRAM

You are now ready for your first python program on the GPU. The video below walks through a simple example that adds two vectors for you to follow along.

If you are new to Python, explore the beginner section (https://www.python.org/about/gettingstarted/) of the Python website for some excellent getting started resources. The blog, An Even Easier Introduction to CUDA (https://devblogs.nvidia.com/parallelforall/even-easier-introduction-cuda/), introduces key CUDA concepts through simple examples.

In the Numba documentation you will find information about how to vectorize functions (https://docs.continuum.io/docs_oss/numba/cuda/ufunc) to accelerate them automatically as well as how to write CUDA code in Python (https://docs.continuum.io/docs_oss/numba/cuda/kernels). Download and execute Jupyter Notebooks (https://jupyter.org/) for the Mandelbrot (http://nbviewer.ipython.org/f5707335f40af9463c43) and Monte Carlo Option Pricer (http://nbviewer.ipython.org/835a8ca39ced77fe751d) examples on your local machine.

3

PRACTICE

Check out Numbas github repository (https://github.com/numba/numba/tree/master/examples) for additional examples to practice.

NVIDIA also provides hands-on training through a collection of self-paced labs . The labs guide you step-by-step through editing and execution of code, and even interaction with visual tools is all woven together into a simple immersive experience. Practice the techniques you learned in the materials above through hands-on labs.

Hands-on Labs (https://nvlabs.qwiklab.com/quests/6?locale=en)

For a more formal,instructor-led introduction to CUDA, explore the Introduction to Parallel Programming on UDACITY. The course covers a series of image processing algorithms such as you might find in Photoshop or Instagram. You'll be able to program and run your assignments on high-end GPUs, even if you don't have one yourself.

UDACITY COURSE (https://eu.udacity.com/course/intro-to-parallel-programming--cs344)

Availability

The Numba package (https://numba.pydata.org/) is available as a Continuum Analytics (https://www.continuum.io/) sponsored open-source project.

The CUDA Toolkit (/cuda-toolkit) is a free download from NVIDIA and is supported on Windows, Mac, and most standard Linux distributions.

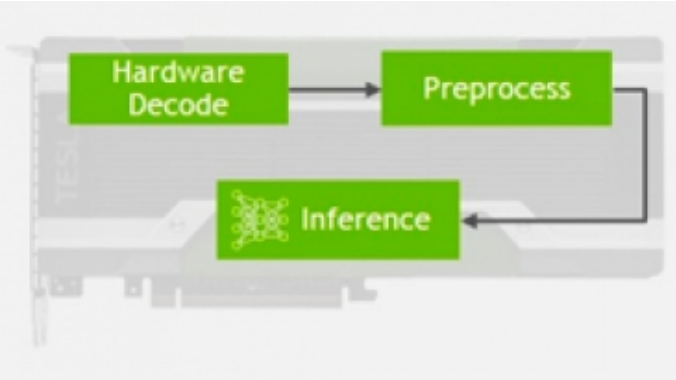
So, now youre ready to deploy your application?

Register today (http://bit.ly/RegDevPages) **for free access to NVIDIA TESLA GPUs in the cloud**.

LATEST NEWS

Download DeepStream SDK 2.0 Today to Develop Scalable Video Analytics Applications

Jun 19, 2018



NVIDIA has released the DeepStream Software Development Kit (SDK) 2.0 for Tesla GPUs, which is a key part of the NVIDIA Metropolis platform.

Read more (/download-deepstream-sdk-20-today-develop-scalable-video-analytics-applications)

A Trio of New Nsight Tools That Empower Developers to Fully Optimize their CPU and GPU Performance

May 31, 2018

Three big NVIDIA Nsight releases on the same day! NSight Systems is a brand new optimization tool; Nsight Visual Studio Edition 5.6 extends support to Volta GPUs and Win10 RS4; and NSight GRAPHICS 1.2 replaces the current Linux Graphics Debugger.

Read more (/trio-new-nsight-tools-empower-developers-fully-optimize-their-cpu-and-gpu-performance)

CUDA 9.2 Now Available

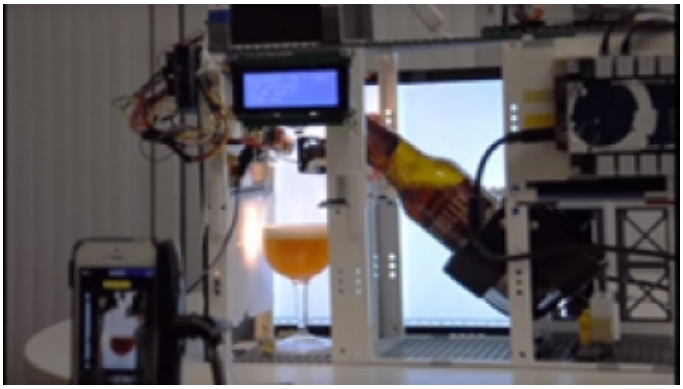
May 21, 2018

CUDA 9.2 includes updates to libraries, a new library for accelerating custom linear-algebra algorithms, and lower kernel launch latency.

Read more (/cuda-92-now-available)

Drink up! Beer Tasting Robot Uses AI to Assess Quality

Apr 6, 2018



Can a beer tasting robot do a better job than humans in judging a beer? Researchers in Australia developed a robot that uses machine learning to assess the quality of the beer.

Read more (/drink-beer-tasting-robot-uses-ai-assess-quality)

All Latest News > (/category/tags/cuda)

BLOGS: PARALLEL FORALL

Coffee Break Series: NVIDIA Highlights

Jul 23, 2018

NVIDIA Highlights enables automatic video capture of key moments, clutch kills, and match-winning plays, ensuring gamers' best gaming moments are always saved without requiring player intervention.
Read more (/coffee-break-series-nvidia-highlights)

Neural Machine Translation Inference with TensorRT 4

Jul 18, 2018

Neural machine translation exists across a wide variety consumer applications, including web sites, road signs, generating subtitles in foreign languages, and more.
Read more (/neural-machine-translation-inference-tensorrt-4)

Using OpenACC to Port Solar Storm Modeling Code to GPUs

Jul 16, 2018

Solar storms consist of massive explosions on the Sun that can release the energy of over 2 billion megatons of TNT in the form of solar flares and Coronal Mass Ejections (CMEs).
Read more (/using-openacc-port-solar-storm-modeling-code-gpus)

Hacking Ansel to Slash VR Rendering Times

Jul 12, 2018

Warrior9 VR team members started working on The PhoenIX – a sci-fi animated series in virtual reality (VR) — two years ago.
Read more (/%C2%A0hacking-ansel-slash-vr-rendering-times)

All Blog Posts > (/taxonomy/term/700)

COMPUTEWORKS (/COMPUTEWORKS)

GAMEWORKS (/GAMEWORKS)

JETPACK (/EMBEDDED-COMPUTING)

DESIGNWORKS (/DESIGNWORKS)

DRIVE (HTTPS://DEVELOPER.NVIDIA.COM/DRIVE)

GET STARTED

About CUDA (/about-cuda)
Parallel Computing (/accelerated-computing-training)
CUDA Toolkit (/cuda-toolkit)
CUDACasts (http://www.youtube.com/playlist?list=PL5B692fm6--vScfBaxgY89IRWFzDt0Khm)

LEARN MORE

Training and Courseware (/cuda-education-training)
Tools and Ecosystem (/tools-ecosystem)
Academic Collaboration (/academia)
Documentation (http://docs.nvidia.com/cuda/index.html)

GET INVOLVED

Forums (https://devtalk.nvidia.com/)
Developer Blog (http://devblogs.nvidia.com/parallelforall/)
Contact Us (/contact)