



# “Programming Massively Parallel Processors” Book and GPU Teaching Kit: New 3rd Edition

Wen-Mei Hwu (University of Illinois) and Joe Bungo (NVIDIA)  
Supercomputing Conference 2016, Salt Lake City, Utah

# AGENDA

Textbook Overview

New 3<sup>rd</sup> edition content

UIUC GPU Activities

NVIDIA GPU Educators Program and Teaching Kits

# PROGRAMMING MASSIVELY PARALLEL PROCESSORS

3<sup>rd</sup> Edition Textbook

Learn to program heterogeneous parallel computing systems

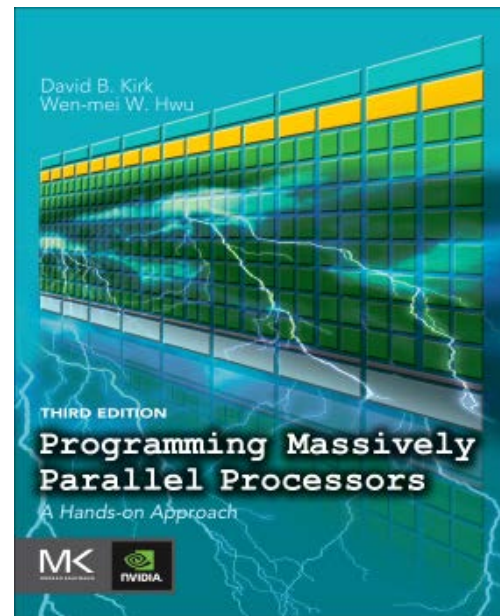
- High performance and energy-efficiency
- Functionality and maintainability
- Scalability across future generations

Technical subjects

- Parallel programming API, tools and techniques
- Principles and patterns of parallel algorithms
- Application case studies
- Processor architecture features and constraints

Free Electronic Copy w/Teaching Kit for Educators!

<https://developer.nvidia.com/educators>



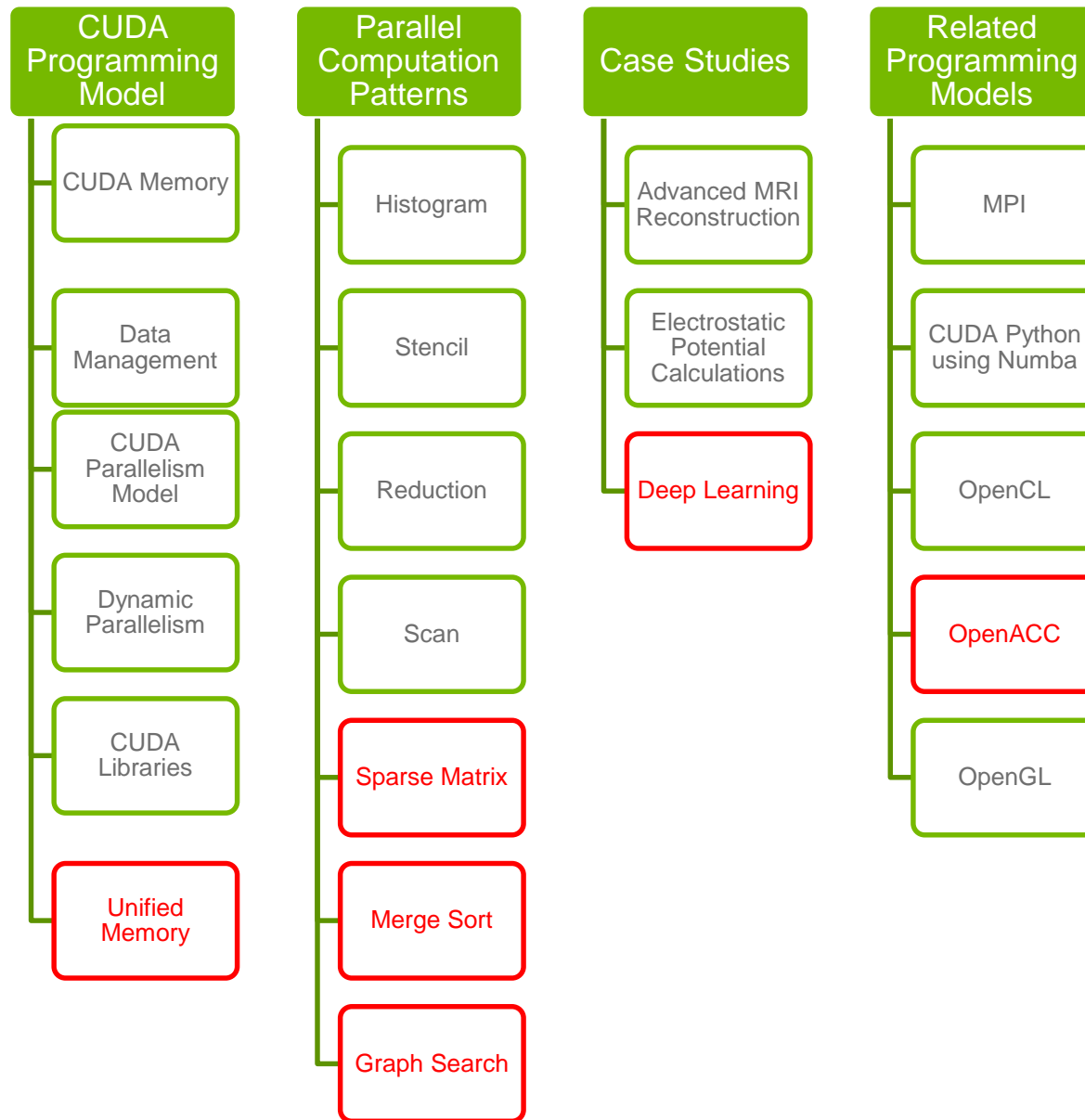
**PMPP  
Textbook**

**CUDA  
Programming  
Model**

**Parallel  
Computation  
Patterns**

**Case  
Studies**

**Related  
Programming  
Models**



## Parallel patterns: sparse matrix computation

An introduction to data compression and regularization

## 10

## Application case study— machine learning

## 16

Boris Ginsburg

## CHAPTER OUTLINE

- 10.1 Background ...
- 10.2 Parallel SpMV
- 10.3 Padding and Tiling
- 10.4 Using a Hybrid Approach
- 10.5 Sorting and Partitioning
- 10.6 Summary .....
- 10.7 Exercises.....
- References .....

Parallel patterns: merge sort  
An introduction to tiling with dynamic input data identification

## 11

Li-Wen Chang and Jie Lv

## API

- 1 Background
- 2 Core
- 3 Context
- 4 Runtime
- 5 Compiler
- 6 Examples

## Parallel programming with OpenACC

## 19

Jeff Larkin

## Parallel patterns: graph search

## 12

Juan Gómez-Luna and Izzat El Hajj

## CHAPTER OUTLINE

- 11.1 Background .....
- 11.2 A Sequential Merge Sort
- 11.3 A Parallelization Approach
- 11.4 Co-Rank Function Improvement
- 11.5 A Basic Parallel Merge Sort
- 11.6 A Tiled Merge Kernel
- 11.7 A Circular-Buffer Merge Sort
- 11.8 Summary .....
- 11.9 Exercises.....
- Reference.....

## CHAPTER OUTLINE

12.1 Background .....	258
12.2 Breadth-First Search .....	260
12.3 A Sequential BFS Function.....	262
12.4 A Parallel BFS Function .....	265
12.5 Optimizations.....	270
Memory Bandwidth .....	270
Hierarchical Queues.....	271
Kernel Launch Overhead .....	272
Load Balance .....	273
12.6 Summary .....	273
12.7 Exercises.....	273
References .....	274

## CHAPTER OUTLINE

- 19.1 The OpenACC
- 19.2 OpenACC Directives
- 19.3 OpenACC by Example
- The OpenACC
- The OpenACC
- Comparison
- OpenACC Directives
- OpenACC Loop
- OpenACC Reduction
- Asynchronous
- 19.4 Comparing OpenACC

## More on CUDA and graphics processing unit computing

## 20

Mark Harris and Isaac Gelado

## CHAPTER OUTLINE

20.1 Model of Host/Device Interaction .....	444
20.2 Kernel Execution Control .....	449
20.3 Memory Bandwidth and Compute Throughput.....	451
20.4 Programming Environment .....	453
20.5 Future Outlook .....	455
References .....	456

# UIUC Activities

## GPU Computing

### UIUC ECE408/CS483

Semester calendar, 15 weeks

Uses 18 modules

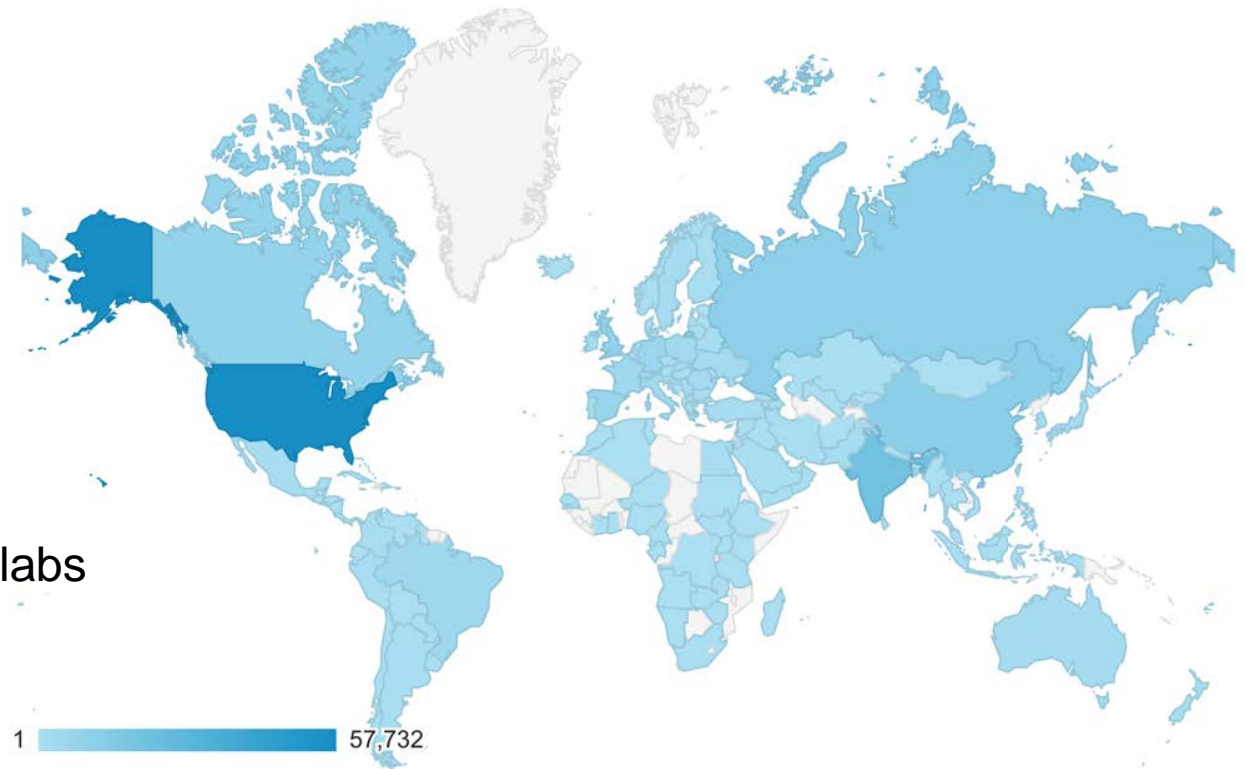
Lecture slides, quizzes, labs

### Coursera HPP

7 weeks

Uses 10 modules

Lecture videos, lecture slides, quizzes, labs



 **ECE408/CS483**

Around 100 students from UIUC

 **CS598 HK**

Around 80 students for UIUC and collaborating institutions

 **Summer School**

Around 100 students from all over the world

 **Coursera HPP**

Around 20,000 students worldwide



# WebGPU.com

## A System for Online GPU Development

- An online IDE for GPU development
- Used intensively at UIUC for the past 4 years
- Essential tool for the Coursera courses offered as well as the introductory and advanced teaching courses at UIUC and several other Universities
- Over 15,000 registered users

Published at EduPar 2016 "*WebGPU: A Scalable Online Development Platform for GPU Programming Courses*" - A. Dakkak, C. Pearson and W. Hwu



# GPU Teaching Kit for Accelerated Computing

Available to Instructors Now!

[developer.nvidia.com/educators](https://developer.nvidia.com/educators)

# NVIDIA GPU EDUCATORS PROGRAM

## Advancing STEM Education with GPU Accelerated Computing

"What an amazing resource for educators in GPU computing! The GPU Teaching Kit has a wealth of resources that allow both experienced and new teachers in parallel computing easily incorporate GPUs into their current course or design an entirely new course."

*Prof. John Owens, UC-Davis*

"The GPU teaching kit covers all aspects of GPU based programming.. the epitome for educators who want to float a course on heterogeneous computing using graphics processors as accelerators."

*Dr. Tajendra Singh, UCLA*

"Teaching resources such as these will be invaluable in helping the next generation of scientists and engineers know how to fully harness the capability of this exciting technology."

*Dr. Alan Gray, University of Edinburgh*

"The Teaching Kit covers all the needed content of a GPU/computing course.. The projects and quiz designs are handy, saving a lot of time and effort. Moreover, the whole structure is well organized to lead students step by step in CUDA programming. I highly recommend integrating it into a related syllabus."

*Dr. Bin Zhou, University of Science and Technology of China*

# FLAGSHIP OFFERING: GPU TEACHING KITS

## Breaking the Barriers to GPU Education in Academia

Co-develop with academic partners

Comprehensive teaching materials

Lecture slides and notes

Lecture videos

Hands-on labs/solutions

Larger coding projects/solutions

Quiz/exam questions/solution

GPU compute resource

Software tools

Textbooks and/or e-books

The collage features three main components:

- Slide 1:** Titled "blockIdx and threadIdx", it lists:
  - Each thread uses indices to decide what data to work on
    - blockIdx: 1D, 2D, or 3D (CUDA 4.0)
    - threadIdx: 1D, 2D, or 3D
  - Simplifies memory addressing when processing multidimensional data
    - Image processing
    - Solving PDEs on volumes
    - ...
- Slide 2:** Titled "Module 3 Lab - CUDA Image Color to Grayscale", it shows code for a CUDA kernel:

```
for (i = 0; i < height; i++)
for (j = 0; j < width; j++)
{
    float channel[3];
    * = colorImage(i, j);
    * = colorImage(i, j + 1);
    * = colorImage(i, j + 2);
    * = (float) (channel[0] + channel[1] + channel[2]) / 3;
}
end
```
- Bitbucket Repository:** Overview for "GPU Teaching Kit - Accelerated Computing Labs". It shows 2 branches, 0 tags, 0 forks, and 2 watchers. The "Software Requirements" section states: "Caution: You must have an NVIDIA CUDA Capable GPU to use the compiled binaries. The CUDA Toolkit can be downloaded from the CUDA Download page; instructions on how to install the CUDA Toolkit are available in the CUDA Start page; installation guides and the list of supported C compilers for Windows, Linux, and OSX are also found in the CUDA Toolkit Documentation Page." The "Compiling and Running Labs" section provides instructions on cloning the repository and building the labs.

# FLAGSHIP OFFERING: GPU TEACHING KITS

Breaking the Barriers to GPU Education in Academia

Different kits for different courses

- Accelerated/parallel computing
- Robotics
- Machine/deep learning (Coming Soon!)
- Virtual Reality
- Computer vision
- Computer architecture
- Computational domain sciences
- Etc.

Localizations/translations in progress

The collage displays various components of the GPU Teaching Kit. At the top, there's a slide from NVIDIA and the University of Illinois titled 'GPU Teaching Kit' and 'Lecture 2.3 - Introduction'. Below it, a slide titled 'blockIdx and threadIdx' explains that each thread uses indices to decide what data to work on, listing 'blockIdx: 1D, 2D, or 3D (CUDA 4.0)' and 'threadIdx: 1D, 2D, or 3D'. It also notes that this 'Simplifies memory addressing when processing multidimensional data' and lists applications like 'Image processing' and 'Solving PDEs on volumes'. A diagram shows a 3D grid of blocks with labels like 'Block (0,0)', 'Block (0,1)', 'Block (1,0)', and 'Block (1,1)'. Another slide is titled 'Module 3 Lab - CUDA Image Color to Grayscale'. The bottom part of the collage shows a Bitbucket repository page for 'GPU Teaching Kit - Accelerated Computing Labs', including an 'Overview' section with statistics (2 branches, 0 tags, 0 forks, 2 watchers) and a 'Software Requirements' section.

# OTHER PROGRAM OFFERINGS

## Collaborative Opportunities and Supporting Expertise

Instructor workshops, conferences, sponsorships and exhibits

Enablement web pages

Getting started guides/videos

Email updates

Feedback and enhancement requests



# GPU Teaching Kit for Accelerated Computing

Available to Instructors Now!

[developer.nvidia.com/educators](https://developer.nvidia.com/educators)

# GPU TECHNOLOGY CONFERENCE

May 8 - 11, 2017 | Silicon Valley | #GTC17  
[www.gputechconf.com](http://www.gputechconf.com)



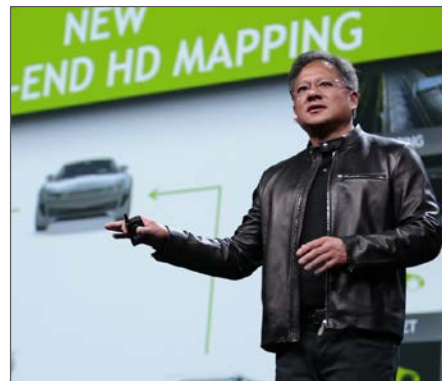
## CONNECT

Connect with technology experts from NVIDIA and other leading organizations



## LEARN

Gain insight and valuable hands-on training through hundreds of sessions and research posters



## DISCOVER

See how GPU technologies are creating amazing breakthroughs in important fields such as deep learning



## INNOVATE

Hear about disruptive innovations as early-stage companies and startups present their work

JOIN THE ACTION! APPLY BY NOV. 20 TO PRESENT A GTC 2017 TALK OR LAB AT [WWW.GPUTECHCONF.COM](http://WWW.GPUTECHCONF.COM)

Don't miss the world's most important event for GPU developers  
May 8 - 11, 2017 in Silicon Valley





# Meet for further discussion!

Tues. Nov. 15<sup>th</sup>, 3-4 PM or Wed. Nov. 16<sup>th</sup>, 10-11 AM  
Peery Hotel, Room Jasper, 110 West Broadway

## GPU Teaching Kit – Accelerated Computing

Available to Instructors Now!

[developer.nvidia.com/educators](https://developer.nvidia.com/educators)