Logistics and Assignments

Stefano Markidis

Instructors

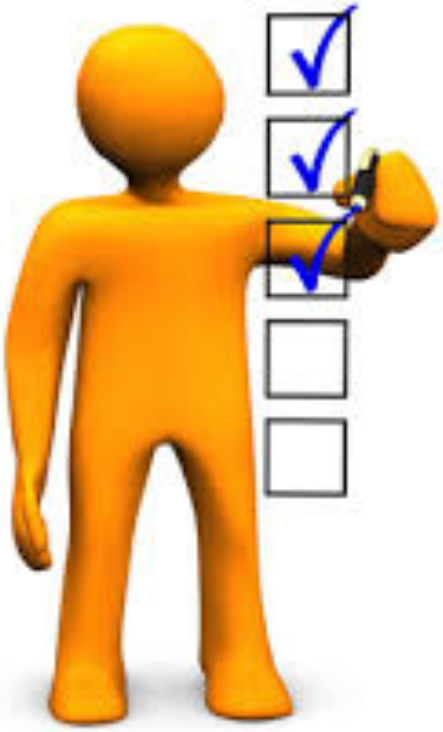
- Stefano Markidis
- Steven W.D. Chien
- Nikolaos Vassardanis

Examiner

- Erwin Laure (erwinl@kth.se)

Prerequisites

- Basic knowledge of **C** or **C++,** compiling codes and Linux commands (Lab workstations)
- Basic knowledge of computer architecture
  - We have one compulsory non-graded quiz to self-assess your knowledge of such topics (https://kth.instructure.com/courses/12406/quizzes/12740?module_item_id=158607).



Course Description

This course provides a broad **introduction to GPU programming with emphasis on application development**.

The course is divided in four modules:

1. GPU architecture
2. CUDA
    1. Basic Concepts
    2. Advanced Concepts
3. Programming Computation with OpenACC
4. Project work

The course will possibly also discuss additional selected topics posting papers and links (CUDA libraries or new features or other frameworks for GPU programming)

Course Format

We are going to use the **flipped classroom** approach:

1. **Study the lecture material at home**: watch video lectures, read articles and additional material, self-assess your knowledge with practice quizzes, …
2. **Do the assignments in class**, in this case lab session, with other students and instructors.

"This isn't what I imagined when they said 'flipped classroom'!"

Interaction and Feedback

- Use Canvas discussion page for posting **questions**, **doubts** and **problems**
  - Instructors will try to reply ASAP
  - Other students can reply also
- It is OK to post part of code and snippet codes
- Avoid to post solutions of the problems in the discussion webpages

Course Activities

Activity	Compulsory	Graded / Non-Graded
Video Lectures	Compulsory	Non-Graded
Quizzes	Compulsory	P/F (Need to have 80%)
Readings	Recommended	Non-Graded
Lab attendance	Highly Recommended	Non-Graded
4 Assignments	Compulsory	Graded: P/F
Project Work	Compulsory	Graded: A-F

Agenda

Date and time	Room	Topic
Tue 10/29 10 – 12	D33	Presentation of the course - Lecture
Thu 31/10 13 – 16	5V4 Magenta	Lab1 – Connect to Tegner – Work on Assign.1
Tue 11/5 9 – 12	5V4 Magenta	Work on Assign. 2 (Basic CUDA)
Thu 11/7 14 - 17	4V6 Brun	Work on Assign. 2 (Basic CUDA)
Tue 11/12 9-12	5V4 Magenta	Work on Assign. 3 (Advanced CUDA)
Thu 11/14 13 – 16	5V4 Magenta	Work on Assign. 3 (Advanced CUDA)
Thu 11/21 13 – 16	5V4 Magenta	Work on Assign. 3 (Advanced CUDA)
Mon 11/28 13 – 16	5V4 Magenta	Work on Assign. 4 (OpenACC)
Tue 12/3 10 – 12 Tue 12/6 13 – 15	E53 E31	Project Description + Discussion Q&A about final project
Thu 12/5 13 – 16 Tue 12/10 9 – 12	5V4 Magenta	Project Work

Assignments

Submission Deadline	Individual/Team	Grading	Topic
Sun 11/3 11.49 PM CET	Group of 2	P/F	GPU Architecture
Sun 11/10 11.49 PM CET	Group of 2	P/F	Basic CUDA
Sun 11/24 11.49 PM CET	Group of 2	P/F	Advanced CUDA *Need to use Tegner for MultiGPU exercise*
Sunu 12/1 11.49 PM CET	Group of 2	P/F	OpenACC – *Need to use Tegner*
Wed 12/8 11.49 PM CET	Group of 2	P/F	Project design document
Tue 1/14 11.49 PM CET	Group of 2	A-F	Final project report

- It is your choice if you want to use your laptop GPU, lab computer GPU or Tegner GPU
  - OpenACC assignment and multiGPU exercise only on Tegner
- Assignments must be uploaded to Canvas as *.pdf* files. Code has to be made available via a public GitHub repository (link in the report).
- **Late submission:** we will assign you a paper to read and ask you to write a 500 words summary

Computer Labs

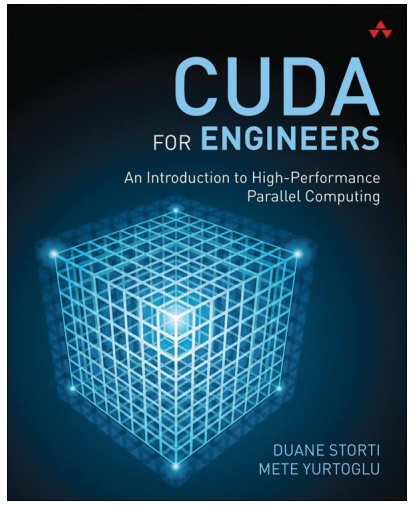
- Computer Lab workstations have Nvidia GPUs and CUDA installed
    - You can use your computer GPU if you have an NVIDIA GPU
- **Attendance** of labs sessions is **highly recommended**.
- Use the lab and discussion sessions to work with other students and instructors on the assignments

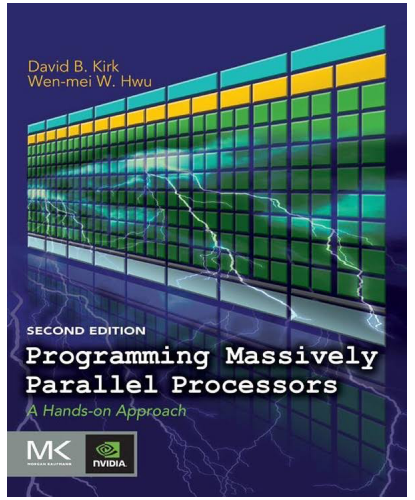
Course Material

- **Slides, papers, video lectures and course material** to complete the assignments will be **progressively posted in Canvas** (https://kth.instructure.com/courses/12406).
- **Text books**:
  - *CUDA for Engineers* by Duane Storti and Mete Yurtoglu

  (https://learning.oreilly.com/library/view/cuda-for-engineers/9780134177540/?ar)
  - *Programming Massively Parallel Processors* by David Kirk and Wen-mei W. Hwu

  (https://www.sciencedirect.com/book/9780124159921/programming-massively-parallel-processors)





Course Final Project

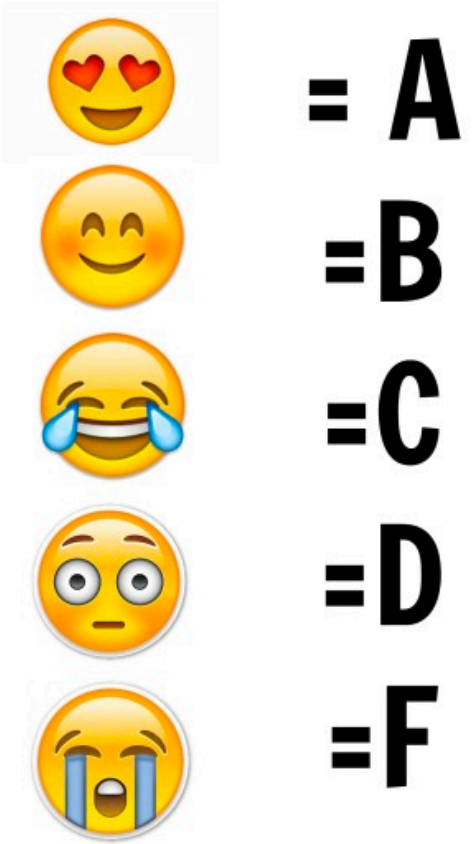
- Take a real-world code and port part of it to GPUs in two weeks work
  - iPIC3D (https://ipic3d.github.io/)
  - Detailed description of the project in the last module of the course.
  - Maximum 3-4 weeks of work.
- Project work in group of two members:
  - Design document (maximum one page) P/F
  - Report (maximum eight pages) A-F

Grading

In order to pass the course:

- Submit **four assignments** (GPU architecture, CUDA and OpenACC) with P/F
- Complete **an individual project course.**The project work requires
  1. **Project design document** (One page)
  2. **Final report** (max 8 pages).

The grade will be determined by the overall quality of the project report and difficulty and originality of the implemented solution (grading criteria will be posted in Canvas).



Resources

- All CS computer labs have workstation with NVIDIA GPUs that you can use for completing your assignments
- You can use your own GPU for completing your assignments and project
  - Need to take care of CUDA installation on your machine (if not installed already)
- You also have access to the Tegner multi-GPU supercomputer
  - You need to learn how to run jobs on supercomputers
    - We will cover this topic in the first lab on Thursday

This week … *What are GPUs?* module

- Watch the video lectures, read slides and the paper from the module **What are GPUs?**
- Try to do the bandwidth measurement (see second exercise in the first assignment) in the computer lab or on your NVIDIA GPU if you have one
  - On Thursday lab, we can try to solve problems
- Complete and submit to Canvas the group assignment

Questions?