

Heterogeneous computing with performance modelling

Course info and setup

Mirko Myllykoski
mirkom@cs.umu.se

Department of Computing Science and HPC2N
Umeå University

4-5. November 2020



UMEÅ
UNIVERSITY



SNIC



HPC2N



People



Mirko Myllykoski
Instructor

Dr, Senior research engineer
Department of Computing Science
and HPC2N, Umeå University



Birgitte Brydso
Helper

MSc, IT systems administrator
HPC2N, Umeå University



Pedro Ojeda May
Helper

Dr, Senior research engineer
HPC2N, Umeå University

Organizations



High Performance Computing Center North (HPC2N) is a national center for Scientific and Parallel Computing.



HPC2N is a part of the Swedish National Infrastructure for Computing (SNIC).



HPC2N is located at Umeå University (Sweden) and closely connected to Department of Computing Science.



This HPC2N course is part of the PRACE Training courses (Partnership for Advanced Computing in Europe).

Course description

Modern high core count CPUs and accelerators, such as **GPUs**, have been shown to provide significant performance benefits in many different applications. However, for a novice, or even for a moderately experienced scientist or programmer, it is not always clear **which applications could potentially benefit from these resources and which do not**. For example, a Nvidia V100 GPU can perform artificial intelligence (AI) related computations in a fraction of the time it takes a regular CPU to perform the same computations but ill-informed OpenACC compiler pragma can actually make a code run slower.

Course description (continuation)

Questions to answer:

- ▶ Why is this?
- ▶ When should one invest time in parallelization and/or GPU-acceleration?
- ▶ How much speedup can be expected with a given application?

Purpose:

- ▶ The main goal of this one day course is to *start answering these questions*.
- ▶ The course also
 - ▶ covers the *basics of GPU programming* and
 - ▶ aims to provide the necessary *information for avoiding the most common pitfalls*.

Requirements: The course does not require any existing GPU programming knowledge but basic understanding of the C language and parallel programming are required for the hands-ons.

Course outline (Wednesday)

Day 1:

09:00 - 09:05 Course information

09:05 - 09:30 Introduction to HPC2N (Birgitte)

09:30 - 09:45 First hands-on

09:45 - 10:00 Coffee break (15 minutes)

10:00 - 12:00 GPU programming basic + hands-ons

12:00 - 13:00 Lunch break (60 min)

13:00 - 14:45 More GPU programming + hands-ons

14:45 - 15:15 Coffee break (30 minutes)

15:15 - 17:00 Hands-ons (more time for those who are beginners)

Course outline (Thursday)

Day 2:

09:00 - 10:15 Performance modelling

10:15 - 10:45 Coffee break (30 minutes)

10:45 - 12:00 Hands-ons

12:00 - 13:00 Lunch break (60 min)

13:00 - 14:45 Advanced GPU programming + hands-ons

14:45 - 15:15 Coffee break (30 minutes)

15:15 - 17:00 Hands-ons (more time for those who are beginners)

Things to know

- ▶ The course consist from a series of lectures and hands-on sessions.
 - ▶ During the hands-ons sessions, we may divide people into smaller groups and assign a breakout room and an instructor/helper for each group.
- ▶ **Two** Zoom sessions:
 - Lecture For lectures, webinar, **recorded**.
 - Hans-ons For hands-ons, regular meeting, **not** recorded.
- ▶ You should have received this information by email.
- ▶ You should have also received links to
 - ▶ a Questions & Answers page (QA) and
 - ▶ an information summary page.
- ▶ Materials: https://git.cs.umu.se/mirkom/gpu_course/

Hands-ons

- ▶ One hands-on under `hands-ons/0.setup`:
 - 1.compiling** Learn how to load the necessary modules.
Learn how to compile CUDA code.
Learn how to place jobs to the batch queue.
Learn how to use the course reservations.