GPU Programming Workshop

3 – 6 February 2025

Workshop material:
https://tinyurl.com/hdli1w24

Overview

- The workshop is co-organised by Leibniz Supercomputing Centre (LRZ), Erlangen National High Performance Computing Center (NHR@FAU) and NVIDIA Deep Learning Institute (DLI).

- NVIDIA Deep Learning Institute (DLI) offers hands-on training for developers, data scientists, and researchers looking to solve challenging problems with deep learning.

- The workshop combines lectures about Fundamentals of Accelerated Computing with OpenACC, CUDA C/C++ and Python on a single GPU with a lecture about Accelerating CUDA C++ Applications with Multiple GPUs.

- The lectures are interleaved with many hands-on sessions using Jupyter Notebooks. The exercises will be done on a fully configured GPU-accelerated workstation in the cloud.

DEEP LEARNING INSTITUTE

DLI Mission: Help the world to solve the most challenging problems using AI and deep learning

We help developers, data scientists and engineers to get started in architecting, optimizing, and deploying neural networks to solve real-world problems in diverse industries such as autonomous vehicles, healthcare, robotics, media & entertainment and game development.
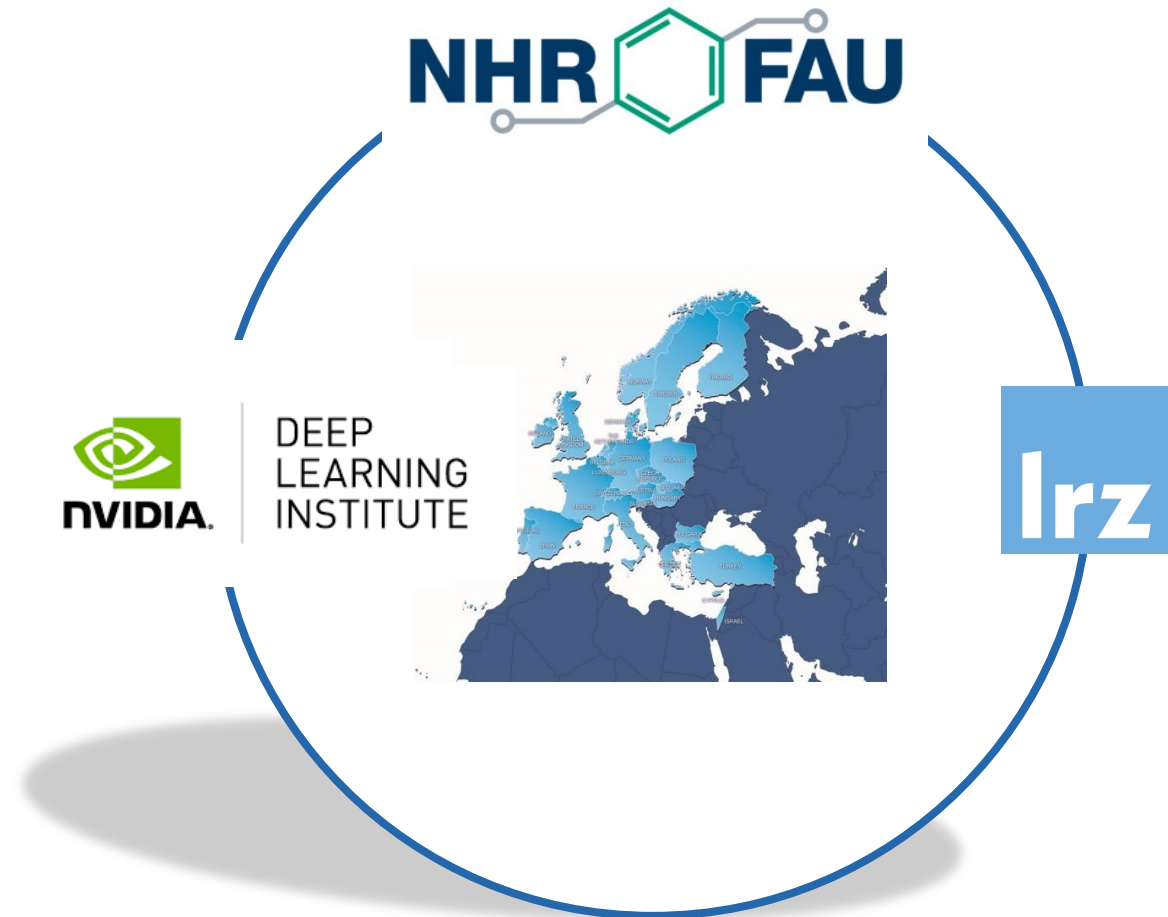
Lecturers

- **Lecturers**:
  - Dr. Momme Allalen (LRZ)
  - Dr. Sebastian Kuckuk (NHR@FAU)
  - Dr. Volker Weinberg (LRZ)

All instructors are NVIDIA certified University Ambassadors.

1st day:
Fundamentals of Accelerated Computing with OpenACC

- Learning Objectives:
  - Profile and optimize your CPU-only applications to identify hot spots for acceleration.
  - Use OpenACC directives to GPU-accelerate your codebase.
  - Optimize data movement between the CPU and GPU accelerator.

2nd day:
Fundamentals of Accelerated Computing with CUDA C/C++

- This lecture teaches the fundamental tools and techniques for accelerating C/C++ applications to run on massively parallel GPUs with CUDA.

- You'll learn how to write code, configure code parallelisation with CUDA, optimise memory migration between the CPU and GPU accelerator, and implement the workflow that you've learned on a new task—accelerating a fully functional, but CPU-only, particle simulator for observable massive performance gains.

- At the end of the lecture, you will be able to create new GPU-accelerated applications on your own.

3rd day:
Fundamentals of Accelerated Computing with CUDA Python

- This lecture explores how to use Numba — the just-in-time, type-specialising Python function compiler — to accelerate Python programs to run on massively parallel NVIDIA GPUs. You'll learn how to:

- Use Numba to compile CUDA kernels from NumPy universal functions (ufuncs)

- Use Numba to create and launch custom CUDA kernels

- Apply key GPU memory management techniques

- Upon completion, you'll be able to use Numba to compile and launch CUDA kernels to accelerate your Python applications on NVIDIA GPUs.

4th day:
Accelerating CUDA C++ Applications with Multiple GPUs

- Learning Objectives
  - Use concurrent CUDA streams to overlap memory transfers with GPU computation

  - Utilize all available GPUs on a single node to scale workloads across all available GPUs

  - Combine the use of copy/compute overlap with multiple GPUs

  - Rely on the NVIDIA Nsight™ Systems Visual Profiler timeline to observe improvement opportunities and the impact of the techniques covered in the workshop

General Agenda

- 10:00-12:00 Lecture / Hands-on

- **12:00-13:00 Lunch Break**

- 13:00-15:00 Lecture / Hands-on

- **15:00-15:30 Coffee Break**

- 15:30-16:45 Lecture / Hands-on

- 16:45-17:00 Q&A, Final Remarks

All times are in CET

Workshop Webpage

- All slides will be made available during the workshop under:

- https://tinyurl.com/hdli1w24

- Further information on:

    - Agenda
    - Training Setup
    - Slides
    - Documentation

Training Setup

- To get started, follow these steps:
- Create an NVIDIA Developer account at https://learn.nvidia.com/join Select "Log in with my NVIDIA Account" and then '"Create Account".

- If you use your own laptop, make sure that WebSockets works for you:
  Test your Laptop at http://websocketstest.com
  - Under ENVIRONMENT, confirm that '"WebSockets" is checked yes.
  - Under WEBSOCKETS (PORT 80]. confirm that "Data Receive", "Send", and "Echo Test" are checked yes.
  - If there are issues with WebSockets, try updating your browser.
    We recommend Chrome or Firefox for an optimal performance.

- Visit https://learn.nvidia.com/dli-event and enter the event code provided by the instructor.
- You're ready to get started.

Course Datasheets by NVIDIA

- https://developer.nvidia.com/dli/getready

- **Prepare For Your NVIDIA DLI Training**

- **Course Datasheets**
  - Please review the workshop datasheet, which includes prerequisites, agenda, suggested material, and resources for continued learning.
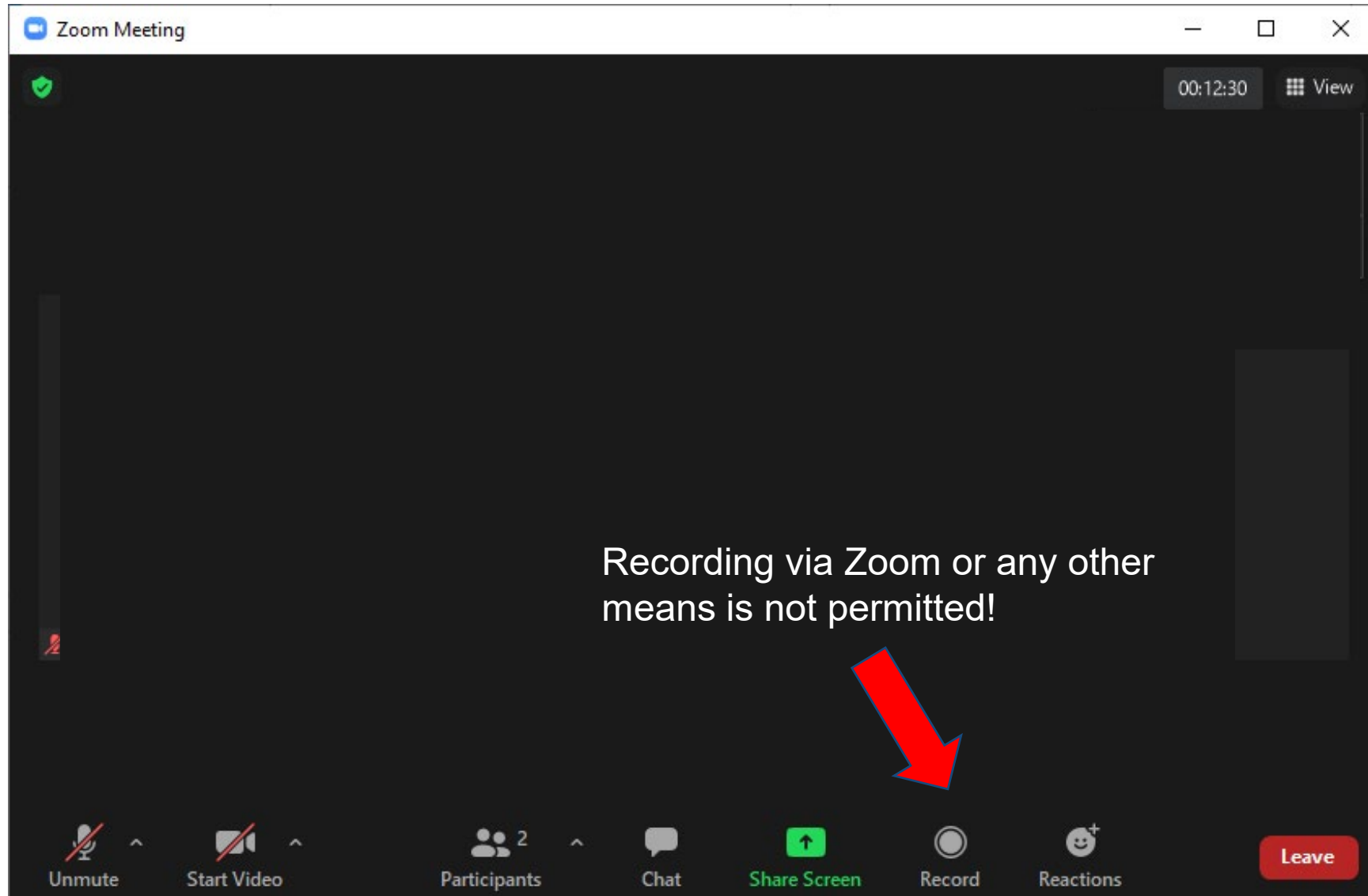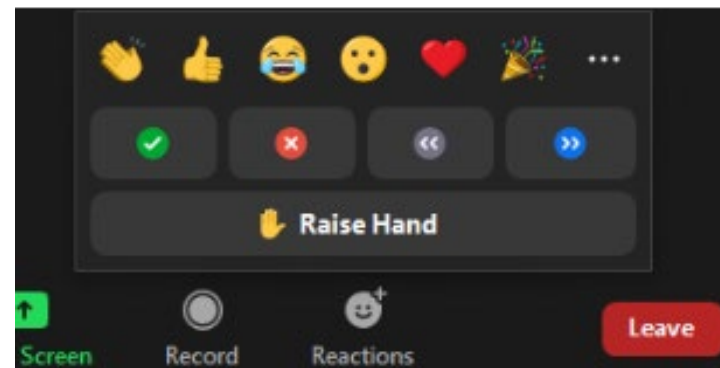
Participants List

Kindly use "<first name> <last name> (<institute>)" as your screenname.

.

Otherwise you will not receive a certificate of attendance after the course.

Zoom Q&A

- Please **raise your hand** if you have questions (of general interest).
- You can also use **chat window** to ask questions.
- If you do not mind, please **show your video when asking questions** to make this course as interactive as possible.
- **Push to Talk:** The Push to Talk feature allows you to remain muted throughout the Zoom meeting and only if you hold down the spacebar you will be unmuted.

- **Instant Feedback:**

And now …

Enjoy the workshop!