

Seminar course

# Accelerated Computing Systems

(aka “acc-systems”)

Preliminary meeting

<https://dse.in.tum.de/>

Dr. Atsushi Koshiba, Harshavardhan Unnibhav, Jiyang Chen,  
Sebastian Reimers, Patrick Sebanic,  
Prof. Pramod Bhatotia



## Chair of Distributed Systems & Operating Systems

<https://dse.in.tum.de/team/>



**Dr. Atsushi Koshiba**

Postdoc



**Harshavardhan Unnibhavi**

PhD student



**Jiyang Chen**

PhD student



**Sebastian Reimers**

PhD student



**Patrick Sebanic**

PhD student



**Communication:**

Join us with TUM email address (@tum.de)

[ls1-courses-tum.slack.com](https://ls1-courses-tum.slack.com)

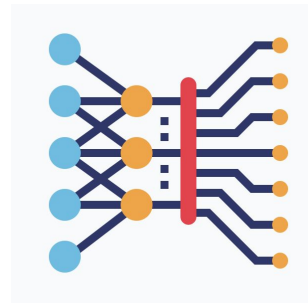
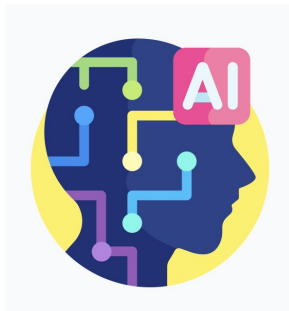
[#ws-23-acc-systems](#)

<https://github.com/TUM-DSE/seminars/>

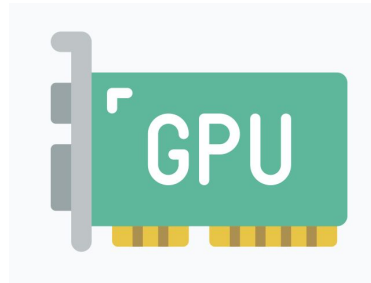
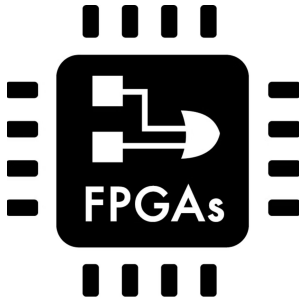
Context

# The need for high performance computing

- The rise of AI is powered by **large-scale data-driven learning**
  - To meet the computational requirements of these modern workloads, we need **high-performance computing**
- **Only CPU-centric computing is still limiting**
  - We need large numbers of high performance cores!
  - Led to the rise of accelerators for compute-intensive, data-intensive tasks



- **Accelerated computing systems** offer massively parallel cores
  - Graphic processing units (GPUs)
  - Field processing gate arrays (FPGAs)
  - Tensor processing units (TPUs) or specialized AI accelerators

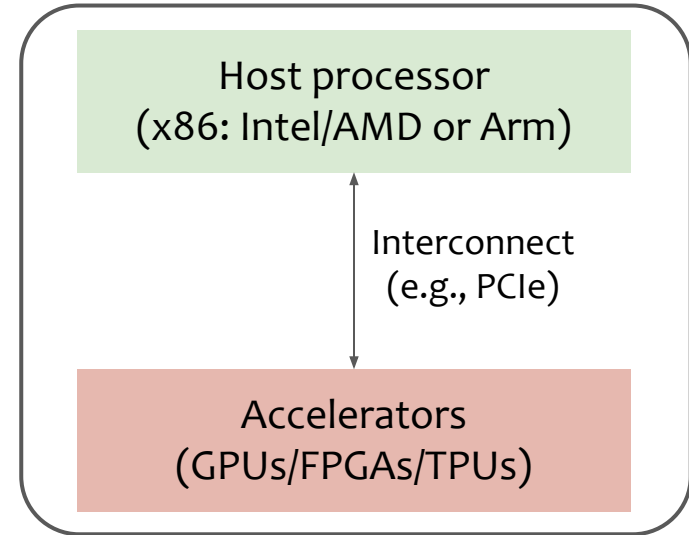


TPUs or  
AI-specialized accelerators

# Accelerated computing systems

- **Accelerated computing** offloads the data-intensive parts of an application
  - while leaving the control functionality to be processed on the CPU
- Having separate types of accelerators is known as **heterogeneous computing**
  - There are multiple types of compute resources available for the application to utilize

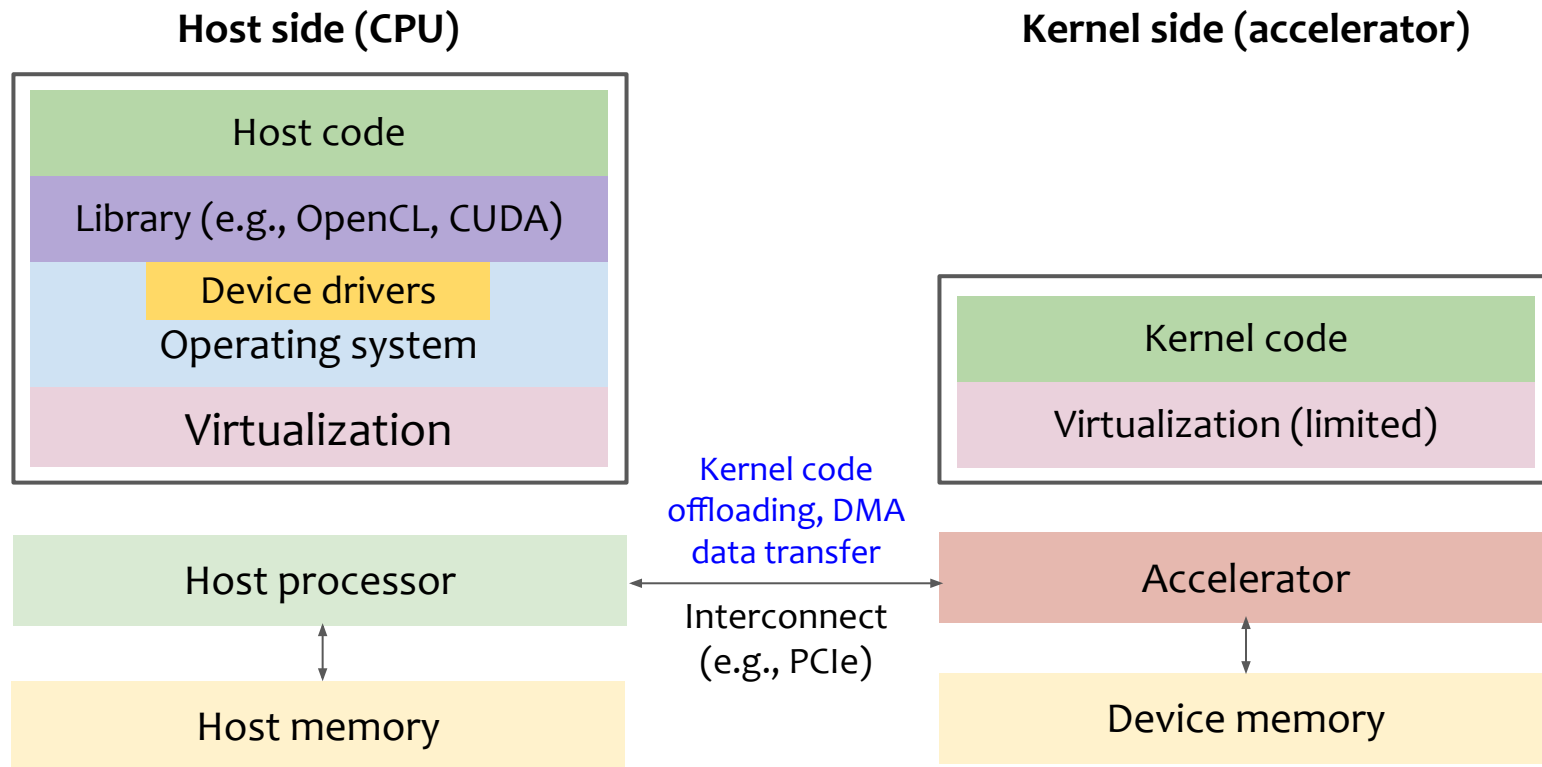
Host code runs on a host processor  
(e.g., filesystem, networking)



**A server architecture w/ accelerators**

Compute-intensive “**kernel**” code  
runs on the accelerators  
(e.g., training an AI model)

# System stack





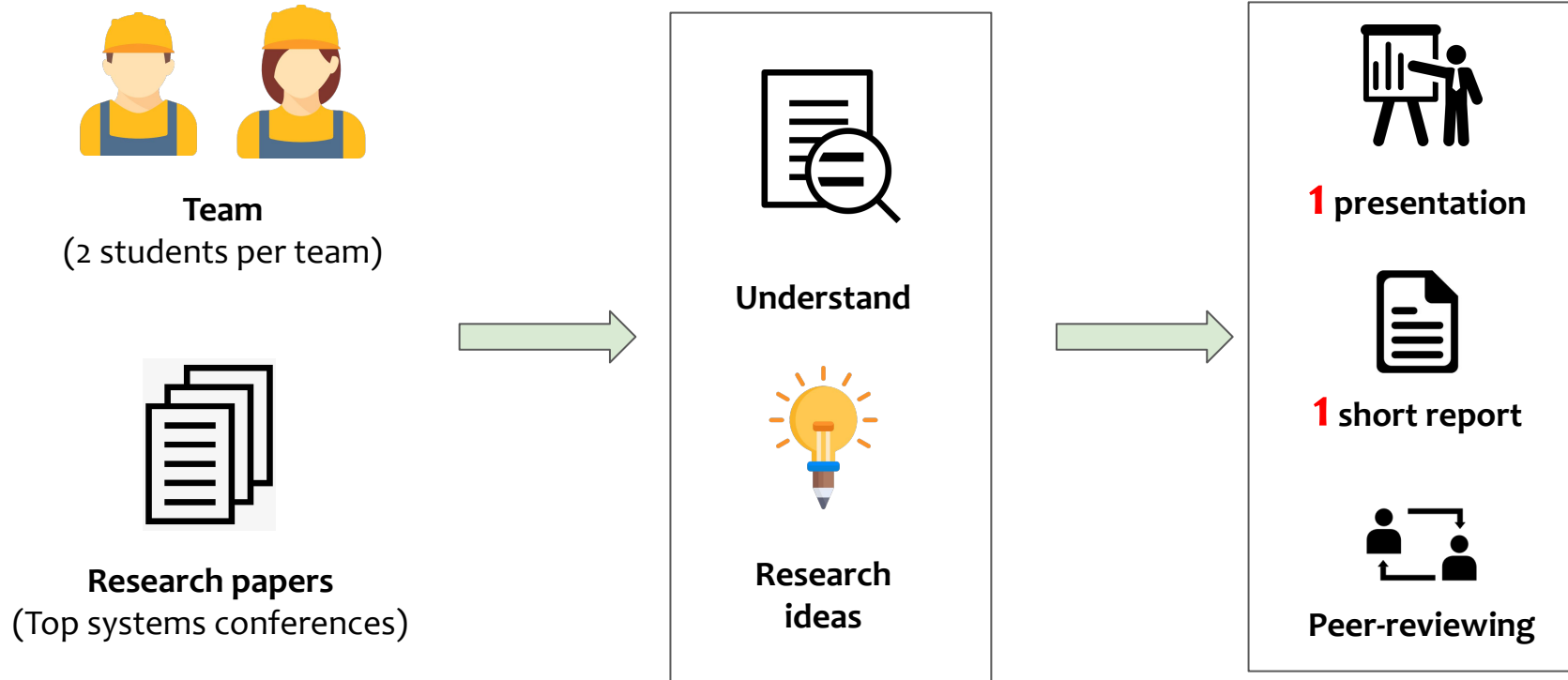
# Tentative topics

Papers from top systems conferences: ASPLOS, OSDI, USENIX ATC, EuroSys, ISCA, and MICRO

Tentative topics
Virtualization for accelerators
Security for accelerators
Heterogeneous task scheduling
SmartNICs
Software-hardware co-design for accelerators
Hardware/OS support for heterogeneous computing
Near-data processing
Resource disaggregation
...

Format

# Bird's eyes view



## Phase I

Kick-off



## Phase II: Understand & explore

Understand



Presentation

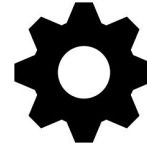


## Phase III: Research

Design



Implement  
( Bonus)

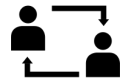


## Phase IV: Report & review

Report



Peer-review



# Phase I: Kick-off meeting



**Format and motivation**  
(all participants meeting)



**Team formation**  
(2 students per team)



**Paper selection**  
(Top systems conferences)



**The first week**

## **NOTE**

1. A list of papers will be provided for FCFS bidding
2. Paper presentation guidelines will be provided for the next phase

# Phase II: Understand & explore



## Understand the paper(s)

### Focus

1. **Understand** the paper and related work
2. Also **explore** a “laundry list” of research ideas/directions



## Paper presentation

### Focus

1. Explain the work/related work (“**why?**” and “**how?**”)
2. Explain and discuss all possible research directions
3. Pick a research direction



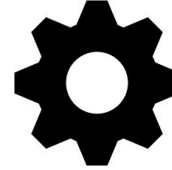
# Phase III: Research



## Research work

### **Focus:**

Indepth research work to nail-down the problem and detailed approach to solve it!



## Research prototype

### **Bonus: (Optional)**

**“Build the system to solve it!”** and show us the working idea and associated results



# Phase IV: Report & review



## Report

### Focus

Prepare a single “short & sweet” report summarizing

- (a) Paper
- (b) Research work



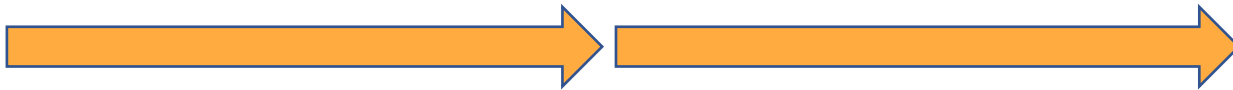
## Peer-review

### Focus

Give constructive (positive and critical) feedback for

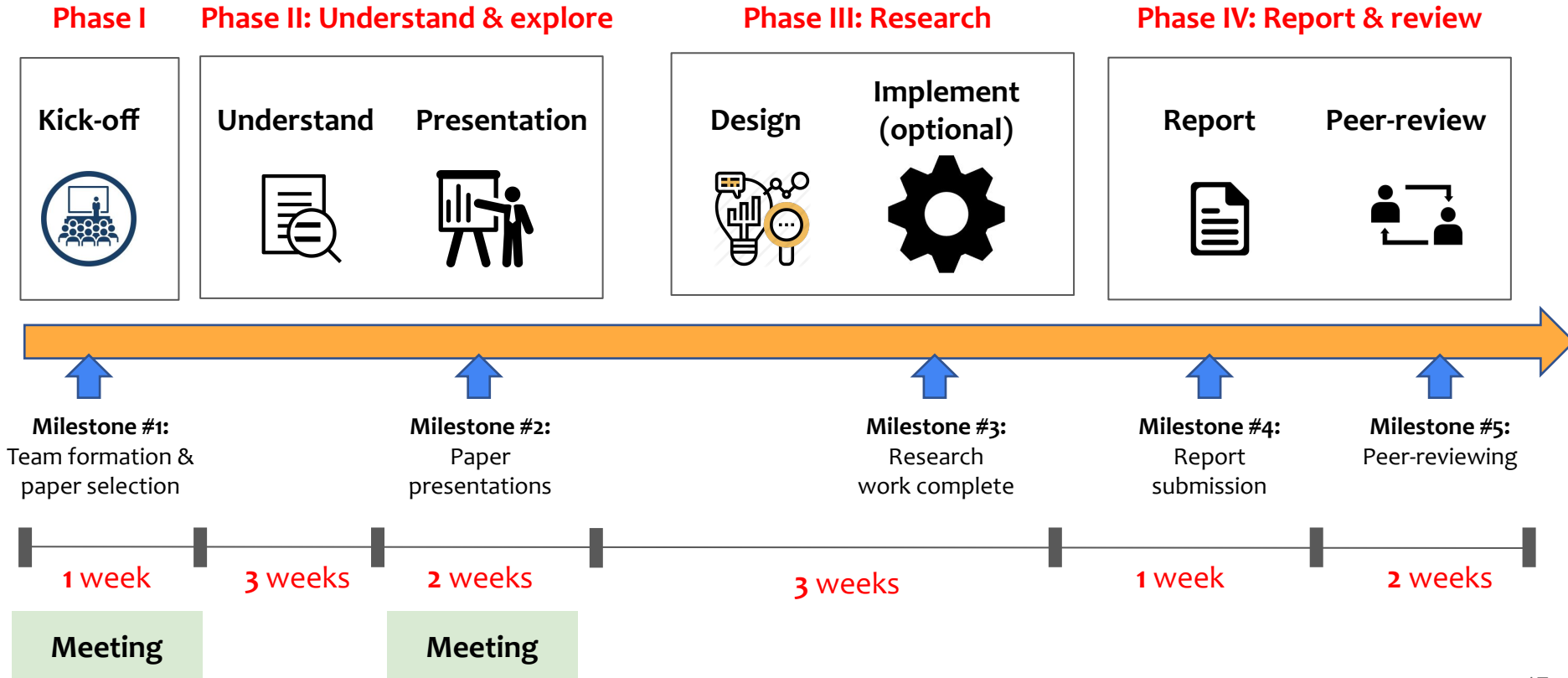
- (a) Paper summary
- (b) Research work

# END.





# Overall timeline



- Format
  - Team-based seminar course (2 students per team)
- Communication
  - Slack for announcements and information sharing
  - Hotcrp for report submission and peer-reviewing
- Meetings (**in-person, attendance is compulsory**)
  - **Meeting #1:** Kick-off
  - **Meeting #2:** Paper presentation

# Learning goals

- Learn about the cutting-edge research in computer systems
- Promote critical thinking
- Cultivate an environment for innovation
  - To push the boundaries by advancing the state-of-the-art
- Improve scientific skills
  - Presentation
  - Writing
  - Communication: discussion and arguing
  - Mentorship: giving feedback and moderating discussion
- Encourage system building and evaluation
  - Learn by building, breaking, and benchmarking systems
- Importantly, to have fun!

- University plagiarism policy
  - <https://www.in.tum.de/en/current-students/administrative-matters/student-code-of-conduct/>
- Decorum
  - Promote freedom of thoughts and open exchange of ideas
  - Cultivate dignity, understanding and mutual respect, and embrace diversity
  - Racism and bullying will not be tolerated

# Interested?

## Matching platform

Welcome to the Matching platform [matching.in.tum.de/](https://matching.in.tum.de/)!

Dear students,

we changed the name of the course "Seminar: Recent advances in Computer Systems", for consistency reasons.  
The new name are "Seminar: Hot Topics in Computer Systems", now.

Login with your TUM identifier.

 TUM login

Login for exchange students  
(without TUM identifier)

 Exchange student login

Any questions? Visit the FAQs!

 FAQs

### Sign up on the TUM matching platform

# Contacts

- Dr. Atsushi Koshiba
  - [atsushi.koshiba@tum.de](mailto:atsushi.koshiba@tum.de)
- **All seminar-related info:** <https://github.com/TUM-DSE/seminars>



**Workspace:** <http://ls1-courses-tum.slack.com/>

**Channel:** [#ws-23-acc-systems](#)

Join us with TUM email address (@tum.de)