

Seminar: Next-Generation Data Systems

Introductory session

Michalis Georgoulakis, Tobias Götz

Garching, October 14th, 2026



Meet the team



Michalis Georgoulakis



Tobias Götz



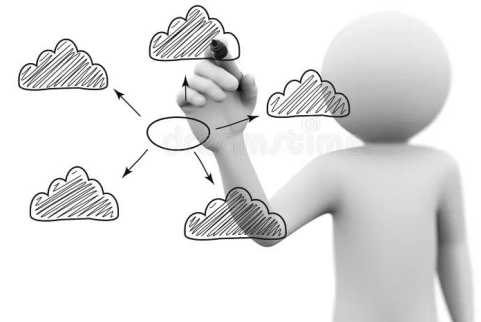
Prof. Dr. Jana Giceva

What this seminar is about

How modern data systems are built for the cloud and for new hardware

Three areas:

- Transactions (*OLTP*)
- Analytics (*OLAP*)
- Modern Hardware

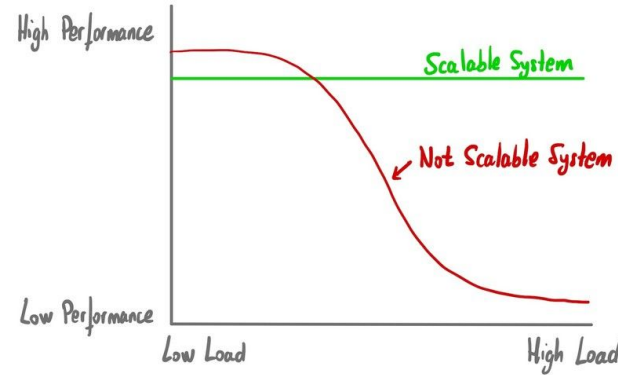


Scalability Matters

The cloud offers resources on demand:

- **Scale up**: a bigger machine; simple, but capped
- **Scale out**: more machines; no fixed cap, but needs coordination
- **Elasticity**: grow and shrink with the load, pay per use

A scalable system keeps its performance as the load grows.



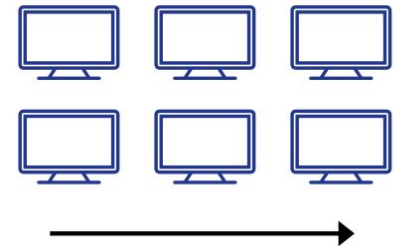
VERTICAL SCALING

Increase size of instance
(RAM, CPU etc.)

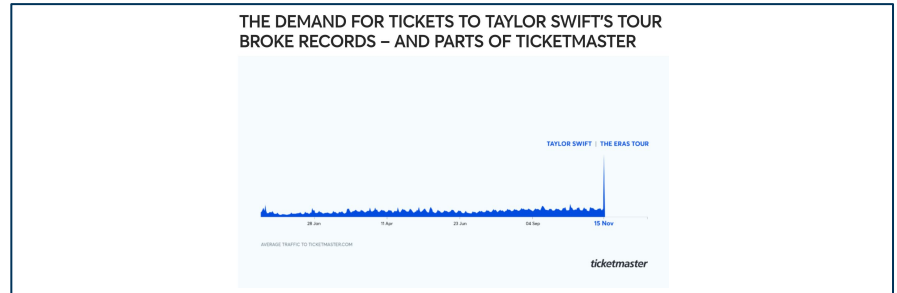
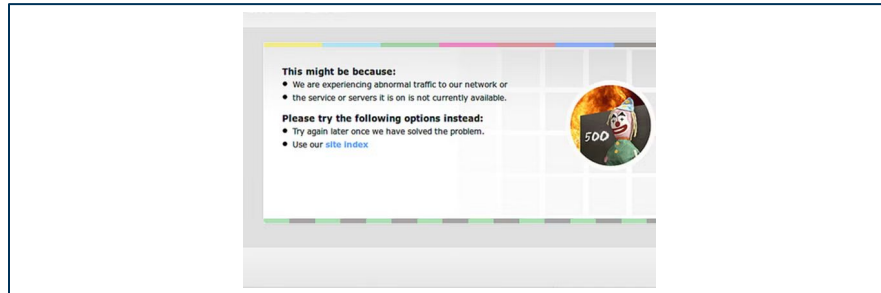
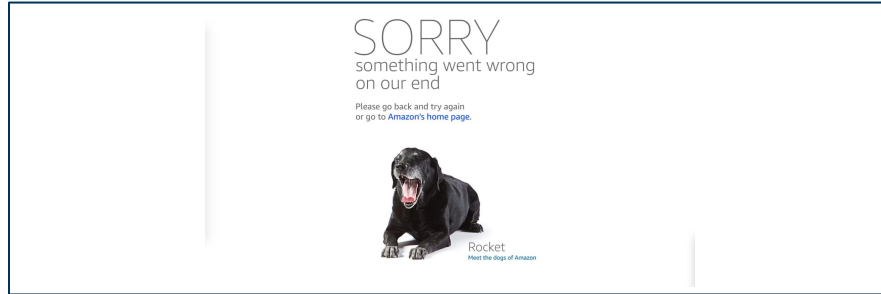


HORIZONTAL SCALING

(Add more instances)



When Systems Fail to Scale

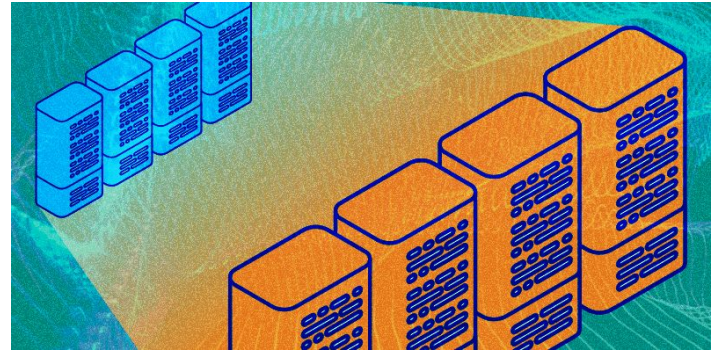


Real outages: demand spikes that the systems could not absorb.

Why Do Systems Fail to Scale?

Usually a combination of:

- **Fixed capacity:** one server hits its CPU, memory or I/O limit
- **Sudden surges:** resources cannot be added fast enough
- **Hidden bottlenecks:** one component, e.g. a single database writer, limits everything
- **Wasted work:** more hardware cannot fix inefficient code



In this seminar, we examine how data systems are rebuilt to scale.

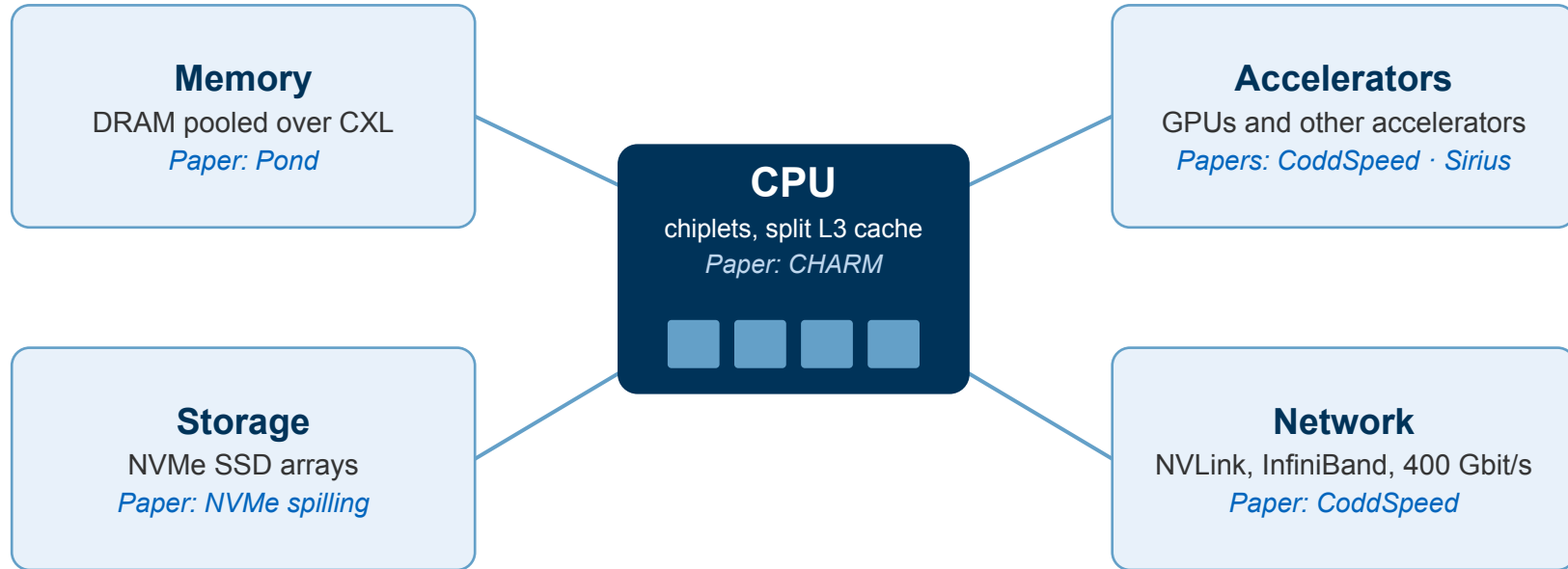
The hardware landscape is also shifting

Hardware is changing faster than the systems built on it:

- **CPUs:** core counts grow faster than memory bandwidth; chiplets split the cache
- **Memory:** DRAM is up to half of a cloud server's cost; CXL adds pooled, slower memory
- **Storage:** a single NVMe SSD reads over 10 GB/s; arrays come close to memory bandwidth
- **Accelerators:** GPUs are now plentiful thanks to AI, with up to 288 GB of memory each
- **Networks:** 100–400 Gbit/s links make disaggregated designs practical

Old assumptions (“disk is slow”, “memory is local”, “the CPU does everything”) no longer hold.

Modern Hardware: Where the Papers Fit



What you will learn

After the seminar, you can:

- **Understand:** explain a systems paper's problem, idea, design and evaluation
- **Read critically:** judge assumptions, trade-offs and evaluation
- **Connect:** relate a paper to other systems and reason about hypothetical use cases
- **Communicate:** present a complex system in 15 minutes and take a position
- **Lead a discussion:** moderate and ask relevant questions
- **Write critically:** write a structured review of a paper



What you will do

- You vote for your paper preferences and get assigned one.
- Present one paper (15 min)
- Answer questions from the audience and the supervisors
- Moderate the discussion of one other paper
- Take part in every session's discussions
- Write a critical review of your paper
- Have two feedback meetings with your supervisor (slides, report)



The Role of Supervisors

Each paper is supervised by a specific supervisor

- Can help you to **understand** the contents of the paper
- Might **discuss** some details of a paper with you
- Provides **feedback** to your submissions
- Will be **present during your presentation** and **join the discussion**
- Will **grade** your individual submissions

We are all nice; just write us an email or send a direct message.

Session Breakdown

Usually two papers per session, each:

- **Talk:** 15 min
- **Open discussion**, led by a student moderator: 10 min
- **Questions** from the supervisors: 10 min

Then a joint panel (15 min):

- Both presenters at the front
- Short outlook on the field by the supervisors
- Open discussion comparing the two papers and brainstorming

Grades

Individual submissions will be graded by the respective supervisor of the topic

Preliminary distribution for the final grade:

- Presentation (~60%)
- Moderation and Discussion Participation (~20%)
- Report (~20%)

Attendance during all presentations and discussions is mandatory to receive a passing grade!

In case of an urgent event, send us an email beforehand.

Presentation

15 minutes: a strict limit

Onsite in room 02.09.014, from 25.11.2026 (see the deadline overview for your date)

Presents and assesses the paper:

- Key contributions
- Ideas and concepts
- Experiments and results
- Your position (last slide): what convinces you, and what you doubt

You cannot cover everything in 15 minutes: choose what you think is most important.

A Good Presentation...

- ... should define a problem
- ... state the contributions
- ... use examples, figures, or more to explain the idea
- ... explain the most important experiments and their results
- ... finish with its main takeaways

We will talk more about presentations in the next weeks.

Feedback Meetings

Presentation

- Send draft slides 7 days before your talk
- Meeting with your supervisor ~5 days before: explain your structure, what you included

Report

- Send your report draft ~2 weeks after your talk; meeting with your supervisor in the following week
- Final report ~4 weeks after your talk

Both meetings: mandatory, 30 minutes, in person in Garching (remote in exceptional cases).

Report

Should contain:

1. Abstract-like summary of the paper
2. Critical review of the paper, responding to points from the discussion
3. ~1 page about your own research idea
 - a. Problem statement
 - b. Expected impact
 - c. Research questions
 - d. Approach

The report should be **≤ 4 pages** and formatted as for SIGMOD (template on the course website)

Writing hints are on [GitLab](#); request access [here](#)

Moderation

Minimum 10 minutes right after the respective presentation

You will moderate a different paper than you present

Acknowledge questions

- Order all questions and call all questioners
- Help the presenter and the participants to find an answer

Prepare some **discussion aspects** (e.g. might contain aspects from the review process)

- Novelty
- Significance (determine the impact of the paper)
- Depth, quality, and chosen experiments
- Presentation of the paper's contents



Discussion

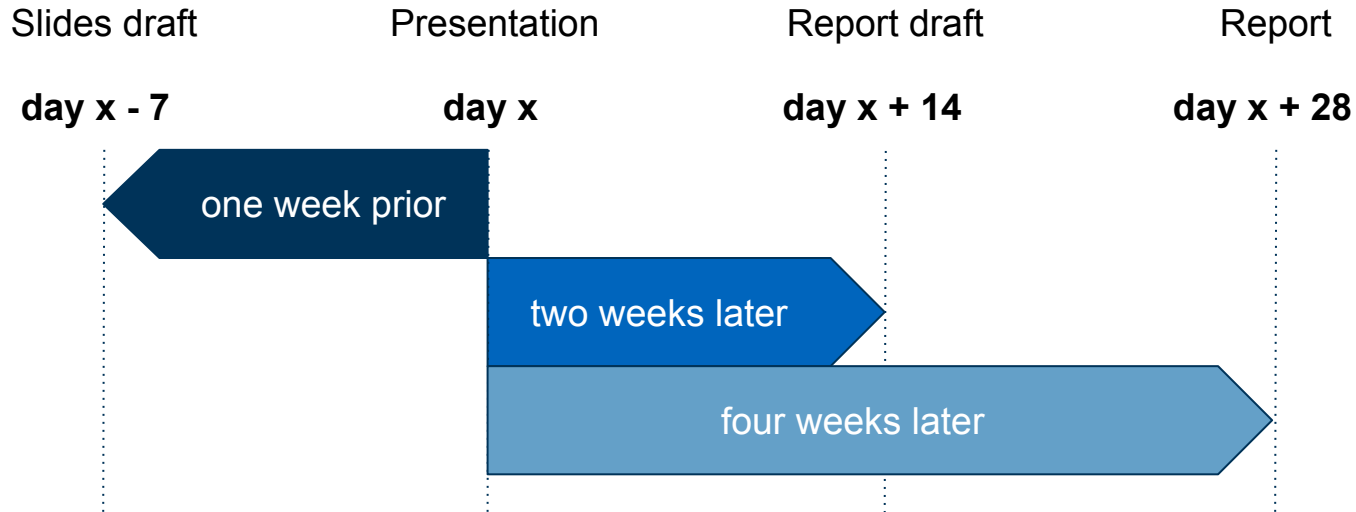
Everyone in this seminar is expected to read a paper before its presentation

- You can take notes of your ideas while you read
- Your thoughts, ideas, alternative designs, or input from your previous experience are valuable!
- Even if something is unclear (could be discussed as well)

Do not hesitate to ask questions to all participants or start a new discussion

Your participation in the discussions is part of your final grade

Deadlines



Send all submissions **to your supervisor via e-mail**.

Exact dates, adjusted around Christmas: see the overview on slide 24.

Matching Procedure

You will be assigned to two papers (not necessarily presented in the same session)

- one for your report
- the other one for moderation

Starting from **today** until **October 25th**, you can send us your **three most preferred papers** via email.

We will announce the assignments **the latest October 27th via E-mail**.

Use of AI Tools

Allowed

- Understanding background concepts and terminology
- Polishing the language of text you wrote yourself

Not allowed

- Generating your slides, your report text or your position

You are responsible for every claim on your slides and in your report, and should be ready to explain them in person.

Timeline

Date	Session Structure	Teaching Material
14.10.2026	Kick-off, papers handed out	Reading a paper
21.10.2026	Q&A on topics; preferences due Sun 25.10	Doing a presentation
28.10.2026	Topics assigned, instructions	Academic Writing
04.–18.11.2026	Preparation (no talks)	-
25.11.2026 – 03.02.2027	Presentation sessions (no session on 23.12.)	-

Deadlines - Overview

Presentation day	Slides Draft	Report Draft	Report Deadline
25.11.2026	18.11.2026	16.12.2026	13.01.2027
02.12.2026	25.11.2026	16.12.2026	13.01.2027
09.12.2026	02.12.2026	23.12.2026	13.01.2027
16.12.2026	09.12.2026	13.01.2027	27.01.2027
13.01.2027	16.12.2026	27.01.2027	10.02.2027
20.01.2027	13.01.2027	03.02.2027	17.02.2027
27.01.2027	20.01.2027	10.02.2027	25.02.2027
03.02.2027	27.01.2027	17.02.2027	03.03.2027

List of Papers (1) - Data Processing on Modern Hardware

GPUs:

- CoddSpeed: Hardware Accelerated Query Processing in Microsoft Fabric
- Sirius: Rethinking Analytical Processing in the GPU Era

Memory and CPUs:

- Pond: CXL-Based Memory Pooling Systems for Cloud Platforms
- CHARM: Chiplet Heterogeneity-Aware Runtime Mapping System

List of Papers (2) - Transactional Systems

Lake storage:

- Delta Lake: High-Performance ACID Table Storage over Cloud Object Stores
- Interoperable ACID Transactions for Open Table Formats

Disaggregation:

- Amazon Aurora: Design Considerations for High Throughput Cloud-Native Relational Databases
- Socrates: The New SQL Server in the Cloud

Geo-distribution:

- Aurora DSQL: Scalable, Multi-Region OLTP
- SunStorm: Geographically Distributed Transactions over Aurora-Style Systems

List of Papers (3) - Analytical Systems

Object storage:

- AnyBlob: Exploiting Cloud Object Storage for High-Performance Analytics
- Pruning in Snowflake: Working Smarter, Not Harder
- High-Performance Query Processing with NVMe Arrays: Spilling without Killing Performance

Compute pushdown:

- FlexPushdownDB: Hybrid Pushdown and Caching in a Cloud DBMS
- LiquidCache: Efficient Pushdown Caching for Cloud-Native Data Analytics

Elastic compute and serverless:

- POLARIS: The Distributed SQL Engine in Azure Synapse
- Skyrise: Exploiting Serverless Cloud Infrastructure for Elastic Data Processing

How to Read a Paper

Three passes:

- **1. Skim** (10 min)
- **2. Read** (1–2 h)
- **3. Reconstruct**

While reading, ask:

- Is the problem real? Who has it?
- What does the design give up?
- Are the baselines fair? What experiment is missing?
- What is shown, and what is only asserted?

Pass 1: The Bird's-Eye View

Goal: get a general idea about the paper (5–10 min)

- Read the title, abstract and introduction carefully
- Read the section headings, but nothing else
- Read the conclusions and glance over the references

Afterwards, you can answer the five Cs:

Category

What type of paper is it?

Context

Which papers is it related to?

Correctness

Do the assumptions seem valid?

Contributions

What are the main ones?

Clarity

Is it well written?

Source: [S. Keshav, How to Read a Paper, ACM SIGCOMM CCR, 2007](#)

Pass 2: Grasp the Content

Goal: understand the content (up to 1 hour)

- Read the whole paper, but skip details such as proofs
- Look closely at figures and graphs: are the axes labelled? Are there error bars? Is the comparison with the baseline fair?
- Note key points and questions in the margins
- Mark unread references for later

Afterwards: you can summarize the main idea, with its evidence, to someone else

Still lost? Read background material and come back, or move on to pass 3.

Source: [S. Keshav. How to Read a Paper. ACM SIGCOMM CCR. 2007](#)

Pass 3: Re-Implement It Virtually

Goal: understand the paper fully, as a reviewer would (4–5 hours)

- Re-create the work under the authors' assumptions, then compare with the paper
- Challenge every assumption, and ask how you would present each idea
- Note ideas for future work

Afterwards, you can:

- reconstruct the paper's structure from memory
- name its strong and weak points
- spot hidden assumptions, missing citations and weak experiments

This is the level we expect for the paper you present.

Source: [S. Keshav. How to Read a Paper. ACM SIGCOMM CCR. 2007](#)

Communication and Material

Officially via e-mail ([Michalis](#), [Tobias](#))

All material can be found on [GitLab](#); request access [here](#)



Questions?

