Barcelona Supercomputing Center
Centro Nacional de Supercomputación

EXCELENCIA SEVERO OCHOA

It's the Memory, Stupid!

Petar Radojkovic

Barcelona Supercomputing Center

25th of February 2026.

Before we start

- **The course will be recorded**

- **We will do a “family photo” just before the lunch**
  - It would be nice if you could make it, but it is not obligatory

  - If you do not want to appear on the photo, you can still have a lunch with us …
    … but no desert! ;-)

- **Please follow the instructions of the organizers**
  - People in the blue BSC t-shirts

Wifi

- **eduroam**

- Network: **BSC_TC_Memory_usage_2026**
  - Password: **Memory26!**

Today will be an interesting day!

- We were working hard to:
  - Debunk some common misconceptions:
    "Long DRAM latency continues to be a critical performance bottleneck in modern systems"
  - Demystify memory system performance
    - Abstract & Esoteric Tangible

- I am excited
  - Very interesting lecturers
  - Very interesting audience
    - People who got their PhDs in 1990s (educated guess) and people who got the birth certificates in 2000s
    - Difficult to make: *"One size fits all"* course

Fits all: Coffee break (30 mins)

+ Lunch (1:30h)

+ Networking event (2h)

- **Hunt the lecturers!**
  - P.S. I am not a lecturer, just giving the intro …
- 60 people in the room **You can talk to 59 people**
- @zoom:
  - **Big apology** to people who wanted to join us in person, but we could not host them

El 14/07/2025 a las 10:39, Petar Radojkovic escribió:

Hello Jana,

26th (Thu) or 27th (Fri) of Feb 2026 are good.
Maybe Friday is better because of the availability of people. Any suggestions from your side?

Regarding the format of the course:

1. Title: It's the Memory, Stupid!

2. (Tentative) Schedule:
The plan is to have 2/3/4 hours of lectures and discussions followed by (simple) hands-on session.
We also plan to provide some reading material before and during the course.

Regarding the breaks:
- One coffee break
- One lunch break

Can we cover the expenses of these breaks?
In campus we can have a lunch for 8eur per person. This would be a great opportunity for an informal discussion with attende
If needed, I can look for the projects that could co-fund this.

3. Format: In person plus recording
I would like to have it in-person plus recording. (I would like to avoid zoom.)
Can we have a good recording: e.g. two (fixed) cameras + a good mike ("pinganillo") for the presenter?

The idea would be to do some (simple) post-editing and upload this course.
I hope that this can give some visibility to the BSC Memory team.

4. Expected audience & Room
I would expect around 20 people.
E.g. C6 E101 or C6 E106 would work for us.

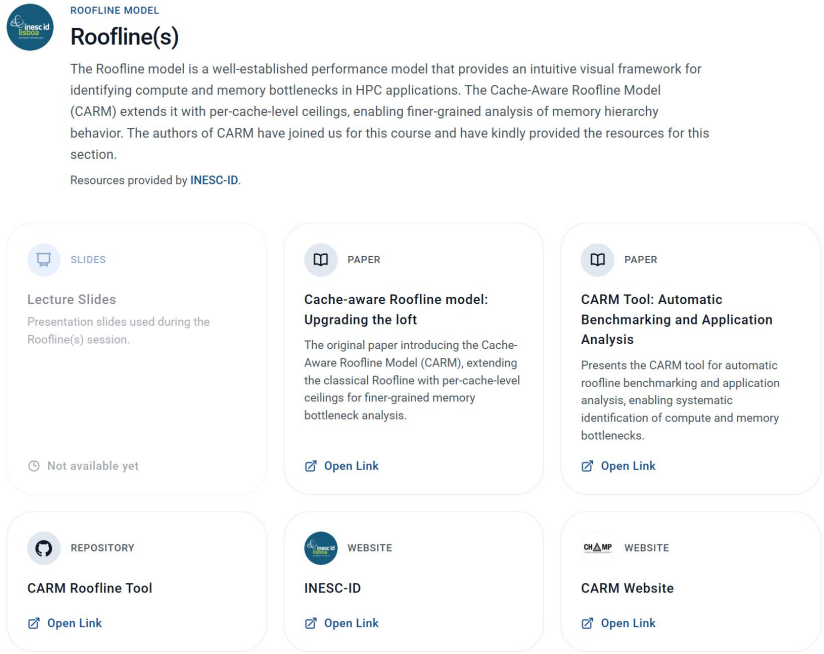
5. Next steps:
Please let me know if you need anything else from my side.
I am also more than happy to discuss the course logistics with the BSC Education.
You have a lot of experience with courses. I hope that you can help us to deliver this well.

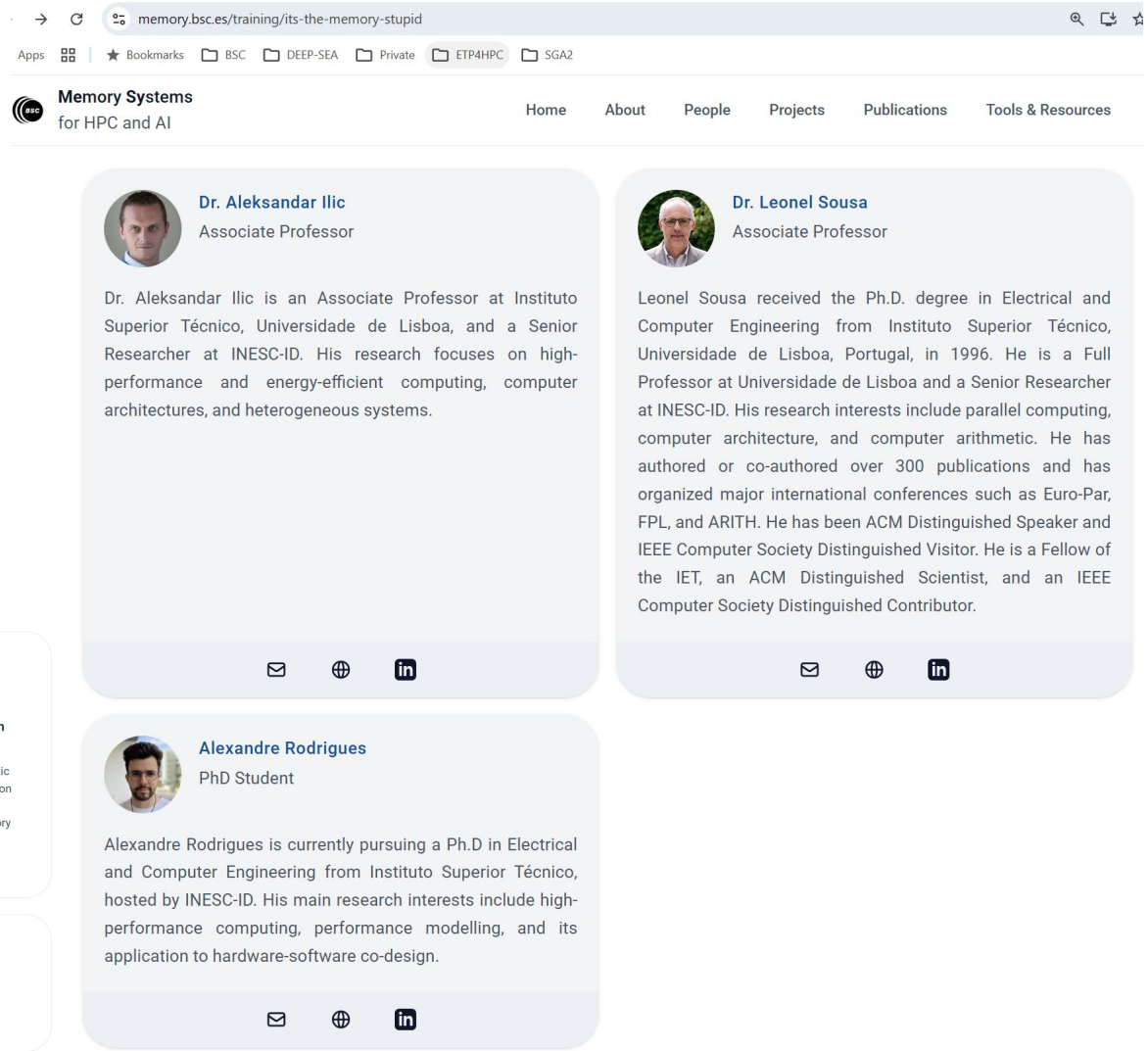
Thanks!
Petar

End of a session Just a beginning of the journey

- Pointers to the related material:
  - *“Tell me more …”* You know where to start!

- Contact ***Do not underestimate the value of this!***
  - *“I have a question …”*
  - *“I want to do a PhD …”*





Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Just a beginning of the journey

- Planning of this course started some time ago

- Now, we are thinking about future trainings
  - **Your opinion is important!**

[Q1] My primary role is:
a. Student
b. Research engineer
c. Researcher
d. Professor

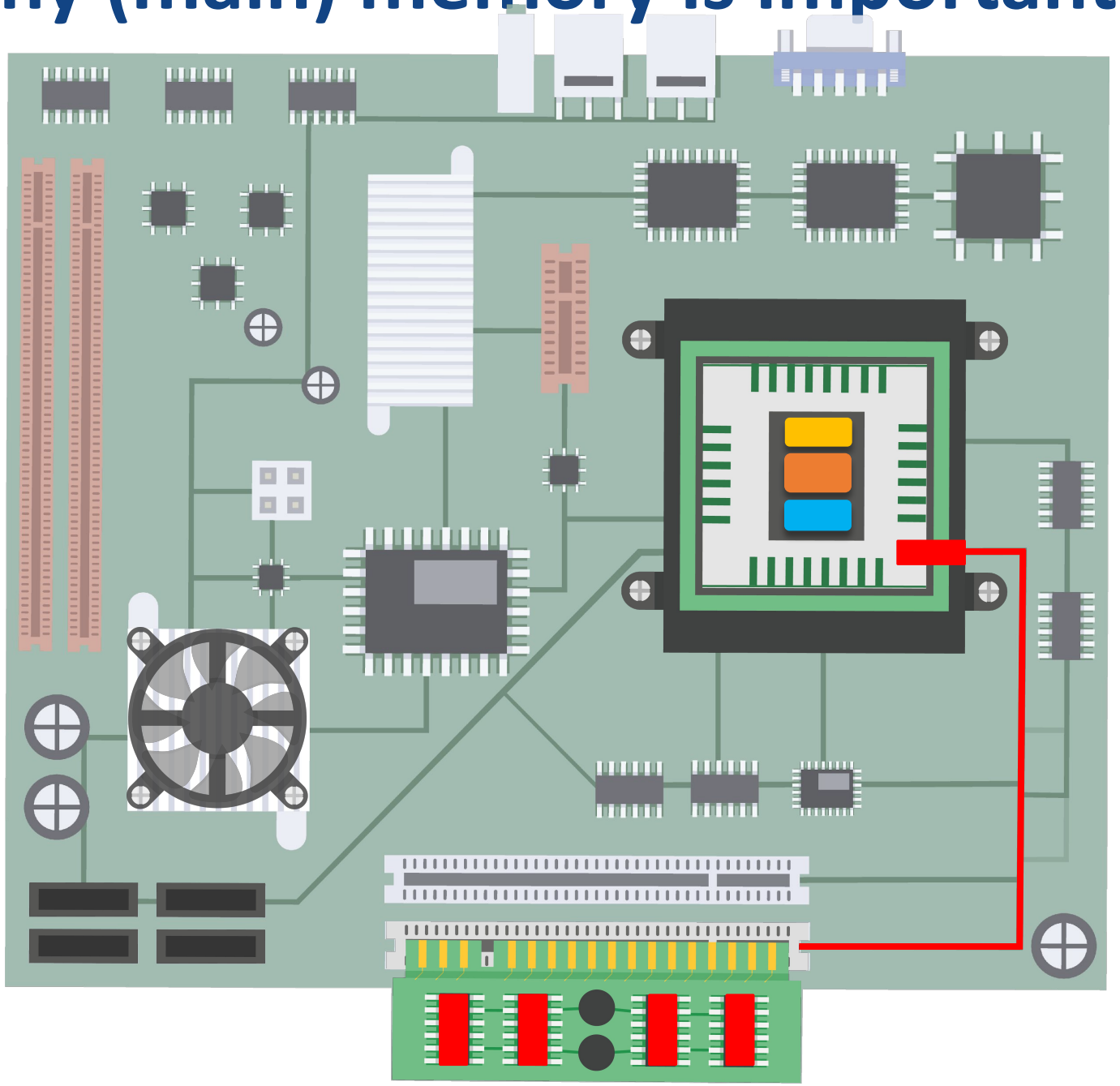
[Q2] I am:
a. HPC application user
b. HPC application developer
c. HPC system admin
d. None of the above

[Q3] How familiar were you with the Roofline analysis before today?
a. Not at all
b. I have heard of it, but never actually used it
c. I used it a couple of times
d. I use it regularly

[Q4] The Roofline presentation was:
a. Easy --- I already knew most of it
b. About right --- I understood most of it, I learned new things
c. Too complex --- I got lost in some explanations

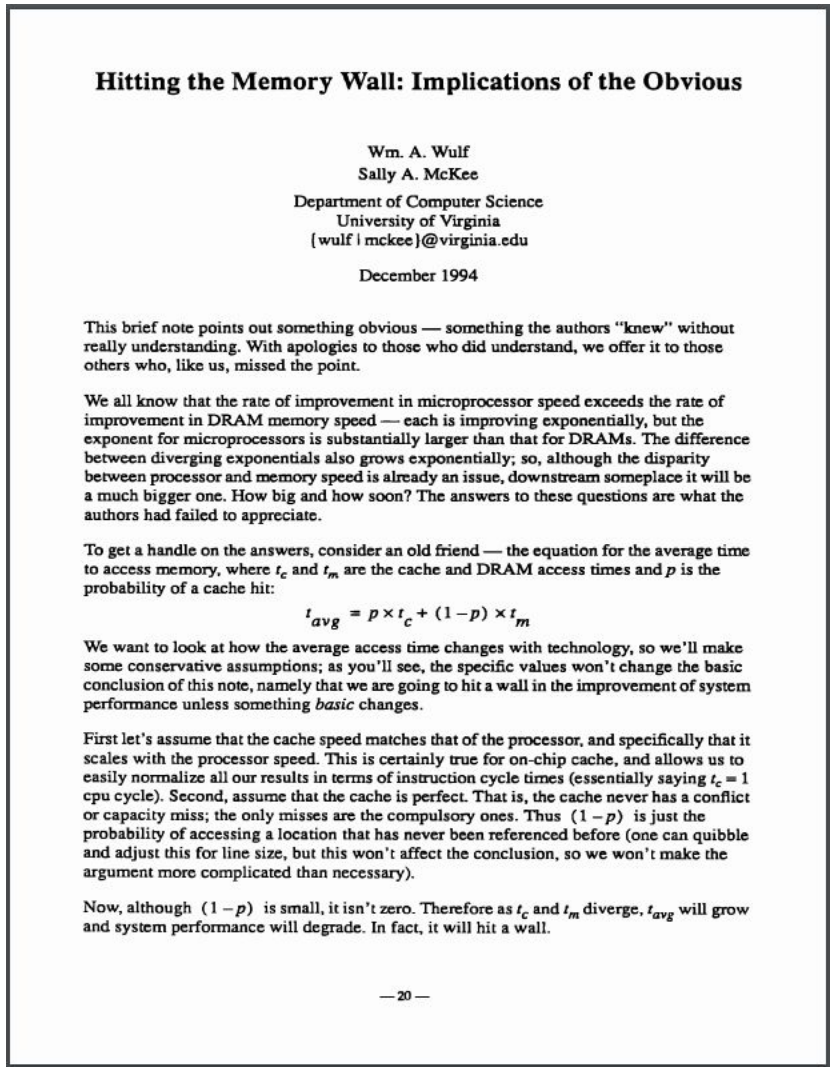
[Q5] I would like a follow-up Roofline training with more details and use cases
a. Yes
b. No
c. Not sure yet

Why (main) memory is important?



- Int add
- FP Div
- L1 Access
- Main memory

Who cares! Caches have 99% hit rate!



Hitting the Memory Wall: Implications of the Obvious

Wm. A. Wulf
Sally A. McKee

Department of Computer Science
University of Virginia
{wulf | mckee}@virginia.edu

December 1994

This brief note points out something obvious — something the authors "knew" without really understanding. With apologies to those who did understand, we offer it to those others who, like us, missed the point.

We all know that the rate of improvement in microprocessor speed exceeds the rate of improvement in DRAM memory speed — each is improving exponentially, but the exponent for microprocessors is substantially larger than that for DRAMs. The difference between diverging exponentials also grows exponentially; so, although the disparity between processor and memory speed is already an issue, downstream someplace it will be a much bigger one. How big and how soon? The answers to these questions are what the authors had failed to appreciate.

To get a handle on the answers, consider an old friend — the equation for the average time to access memory, where t_c and t_m are the cache and DRAM access times and p is the probability of a cache hit:

$$t_{avg} = p \times t_c + (1-p) \times t_m$$

We want to look at how the average access time changes with technology, so we'll make some conservative assumptions; as you'll see, the specific values won't change the basic conclusion of this note, namely that we are going to hit a wall in the improvement of system performance unless something *basic* changes.

First let's assume that the cache speed matches that of the processor, and specifically that it scales with the processor speed. This is certainly true for on-chip cache, and allows us to easily normalize all our results in terms of instruction cycle times (essentially saying $t_c = 1$ cpu cycle). Second, assume that the cache is perfect. That is, the cache never has a conflict or capacity miss; the only misses are the compulsory ones. Thus $(1-p)$ is just the probability of accessing a location that has never been referenced before (one can quibble and adjust this for line size, but this won't affect the conclusion, so we won't make the argument more complicated than necessary).

Now, although $(1-p)$ is small, it isn't zero. Therefore as t_c and t_m diverge, t_{avg} will grow and system performance will degrade. In fact, it will hit a wall.

— 20 —

In 1995, Wulf and McKee published a four-page note entitled "Hitting the Memory Wall: Implications of the Obvious" in the (un-refereed) ACM SIGARCH *Computing Architecture News* [27]. The motivation was simple: at the time, researchers were so focused on improving cache designs and developing other latency-tolerance techniques that the computer architecture community largely ignored main memory systems. The article projected the performance impact of the increasing speed gap between processors and memory. The study predicted that if the trends held, even with cache hit rates above 99%, relative memory latencies would soon be so large that the processor would essentially always be waiting for memory — which amounts to "hitting the wall".

Assignment:
Read this paper.
READ IT CAREFULLY

Wm. A. Wulf and S. A. McKee. *Hitting the memory wall: Implications of the obvious.* Computer architecture news, 1995.

Actually, read all three of them

- Wulf and McKee. **“Hitting the memory wall: Implications of the obvious”.** Computer Architecture News. 1995.

- McKee. **“Reflections on the Memory Wall”.** International Conference on Computing Frontiers. 2004.

- Radulovic et al., **“Another Trip to the Wall: How Much Will Stacked DRAM Benefit HPC?”**. International Symposium on Memory Systems (MEMSYS). 2015.

2017-2025:
Introduction to heterogeneous memory usage

Assignment

- Rules
  - Individual task: Just think about this and write down the answer
  - 3 min
- Task
  - One high-end server (e.g. Intel) with DDR4/DDR5
  - Whatever benchmark you select
  - What is your expected performance improvement if you add Optane to this platform?
    - Not at the end of the day; today is probably your first date with Optane.
    - After 3-6 months of optimizations.
- Write down
  - Benchmark (suite)
  - Expected performance improvement

In this session, we will think about this problem

"If I had an hour to solve a problem,
I would spend 55 minutes thinking about the problem
and 5 minutes thinking about solutions."

"Make everything as simple as possible, but not simpler"

- If you think in terms:
  - *"High-performance memory"* vs. *"Low-performance memory"*
  - *"Big & slow"* vs. *"Small & fast"* memory

  This is too simple

- **If you want to work in area of memory systems**
  - Adoption of new memory devices
  - Heterogeneous memory systems
  - Performance analysis
    - Memory devices
    - Applications
    - …

  The course today is "as simple as possible" [!]
    - At least 80% of the course today

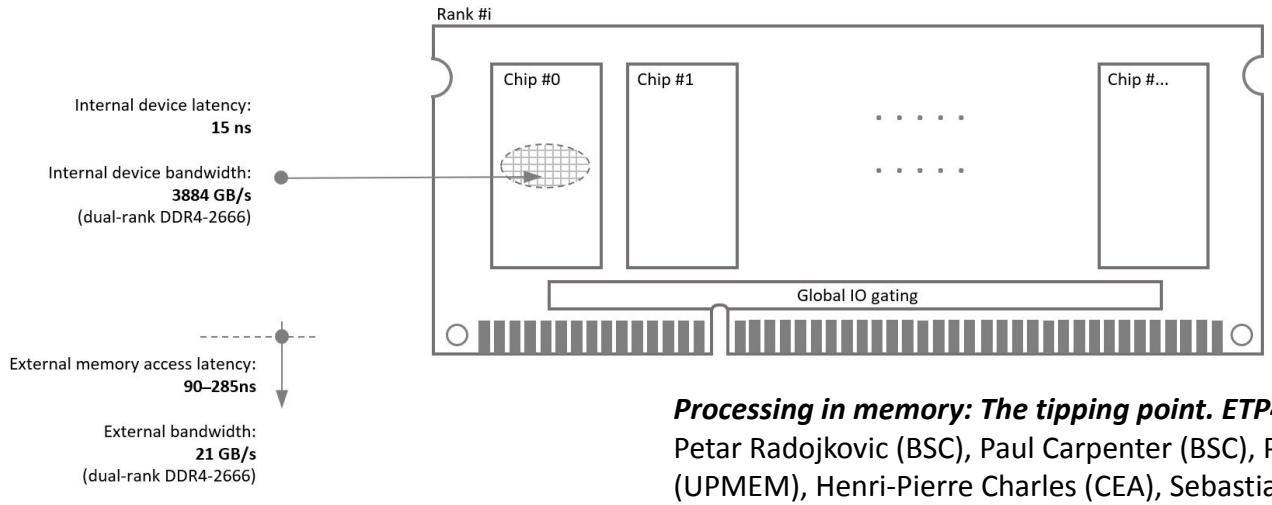
Brutal misconception:

"DRAM latency is constant for decades …
… Therefore, long DRAM latency continues to be a critical performance bottleneck in modern systems."

- **The truth is a bit more complex:**

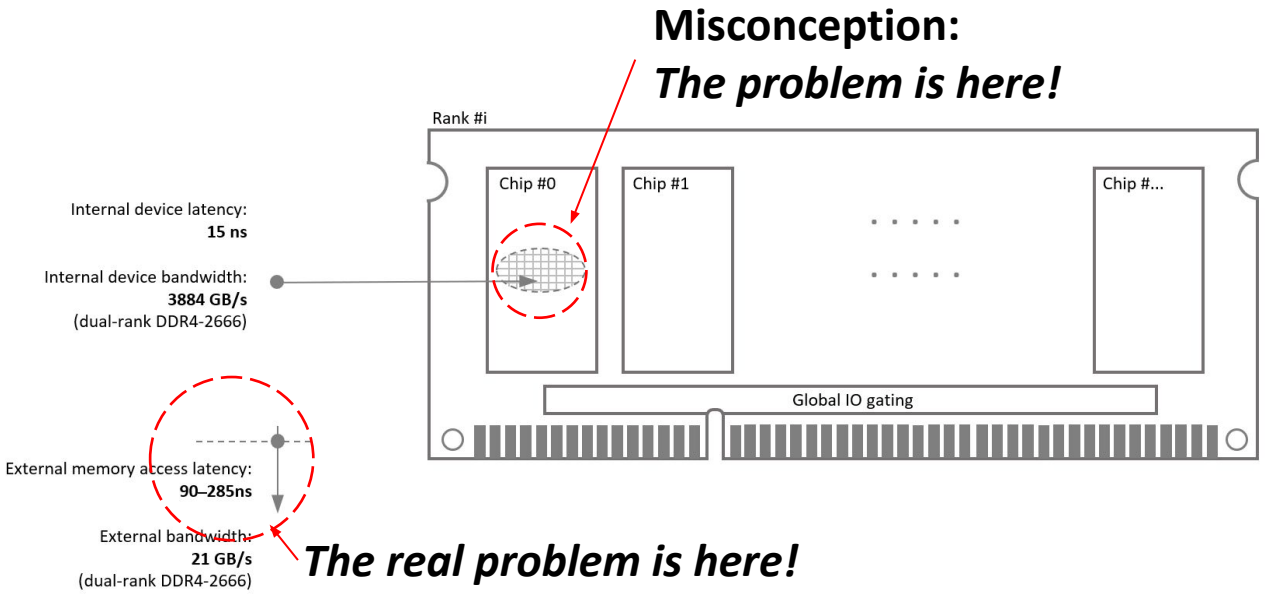
The high latency and low bandwidth of main memory is often thought to be a consequence of the memory technology, e.g. the timing parameters associated to the memory cells or cell-arrays. But this is wrong. In fact, most of the main memory access latency does not come from the time required to access the memory cells, but from the overall complexity of the memory system in both hardware (e.g. CPU cache hierarchy) and software (e.g. virtual memory management).

Internal memory latency in current DDR3 and DDR4 devices could be as low as 15 ns [6, 7]. This is the time required to read data from the memory array to the sense amplifiers, defined by the JEDEC standard as the Activate to internal Read delay (tRCD) timing parameter. The **external memory access latency**, the main memory latency perceived by the user, is the time required for a load instruction to access data in main memory. This timing is platform-specific as it includes the time spent in the CPU load/store queues, cache hierarchy, network-on-chip and on-chip memory controller. In current architectures the minimum external memory access latency, corresponding to the simplest case of a single memory load in an idle system, starts at 90 ns [8], exceeding the internal device latency by **6x**.



Processing in memory: The tipping point. ETP4HPC white paper. 2021.
Petar Radojkovic (BSC), Paul Carpenter (BSC), Pouya Esmaili-Dokht (BSC), Rémy Cimadomo (UPMEM), Henri-Pierre Charles (CEA), Sebastian Abu (IBM), Paolo Amato (Micron)

"DRAM latency is constant for decades …
… Long DRAM latency continues to be a critical performance bottleneck in modern systems"



And 15ns of the "DRAM cell" latency is irrelevant.

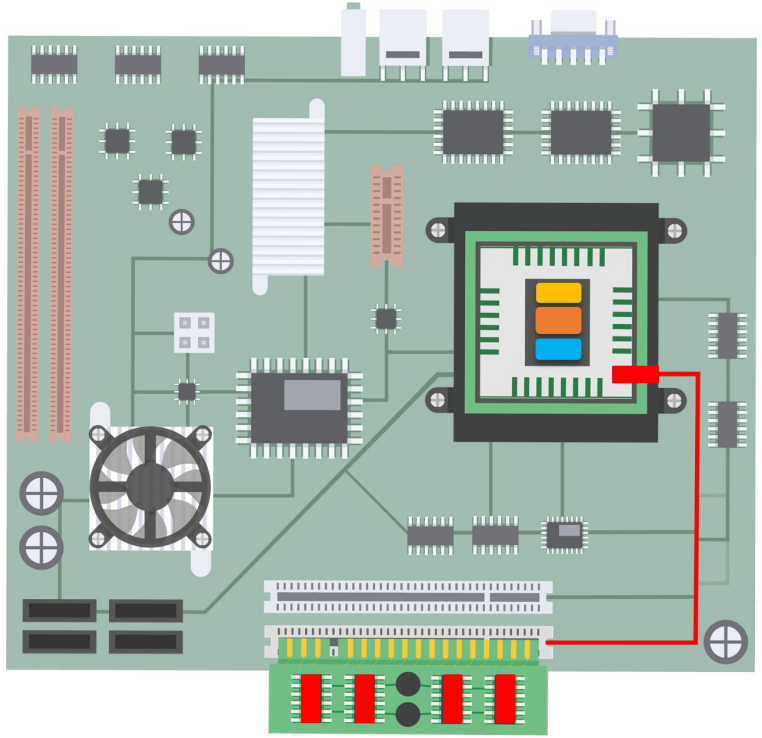
DRAM latency is constant for decades

- Serious side affect: *“There is nothing I can do about it …”*
  - Learned helplessness

- **You could not be more wrong!**
  - Memory access latency that matters (load-to-use latency) ranges between **100ns and 500ns** in **the servers which you use**! We will show you this today

There is so much you could do about it!

Step 1: Understand it!



Next: Memory bandwidth

- **Brain teaser 1:**
  - The memory system bandwidth (from the data-sheets) is 100GB/s
  - Your application uses 70GB/s
  - Would you benefit from more bandwidth?

- **Brain teaser 2:**
  - The memory system bandwidth (from the data-sheets) is 100GB/s
  - You run a bandwidth-hungry benchmark (e.g. STREAM) and you measure 75GB/s
  - Your application uses 70GB/s
  - Would you benefit from more bandwidth?

SCHEDULE

Course Agenda

View all materials

Part 1 10:00h – 10:45h

Introduction

Dr. Petar Radojkovic – BSC

Part 2 10:45h – 16:00h

Performance Analysis: Memory Systems & Applications

10:45h – 11:30h · Roofline(s)
Dr. Aleksandar Ilic — INESC-ID — Materials

11:30h – 12:00h · Coffee Break
Location: Vertex garden

12:00h – 12:45h · Memory System Performance is Messy
Pouya Esmaili-Dokht — BSC, Victor Xirau — BSC — Materials

12:45h – 13:15h · PROFET: No-stress Performance Prediction
Mariana Carmin — BSC — Materials

13:15h – 14:45h · Lunch Break
Location: Vertex garden

14:45h – 15:45h · TopDown Microarchitecture Analysis
Dr. Harald Servat — Intel, Victor Xirau — BSC — Materials

15:45h – 16:00h · Short Break

Part 3 16:00h – 17:00h

Advanced Memory Systems

16:00h – 16:30h · Heterogeneous Memory Systems
Dr. Toni Peña — BSC — Materials

16:30h – 17:00h · Panel Discussion
Dr. Aleksandar Ilic — INESC-ID, Dr. Harald Servat — Intel, Dr. Toni Peña — BSC, Dr. Petar Radojkovic — BSC

Part 4 17:00h – 19:00h

Networking event

Location: Vertex garden

What will you get today?

(apart from the WiFi & Coffee)

Outline

- General ideas / Fundamentals
- System profiling
- Application profiling
- Putting it all together: System & Application profiling
- Demo / Hands-on [Advanced level]
- Explore more

It's the Memory, Stupid!

SCHEDULE

Course Agenda

View all materials

Part 1 10:00h – 10:45h

Introduction

Dr. Petar Radojkovic – BSC

Part 2 10:45h – 16:00h

Performance Analysis: Memory Systems & Applications

10:45h – 11:30h · Roofline(s)
Dr. Aleksandar Ilic — INESC-ID — Materials

11:30h – 12:00h · Coffee Break
Location: Vertex garden

12:00h – 12:45h · Memory System Performance is Messy
Pouya Esmaili-Dokht — BSC, Victor Xirau — BSC — Materials

12:45h – 13:15h · PROFET: No-stress Performance Prediction
Mariana Carmin — BSC — Materials

13:15h – 14:45h · Lunch Break
Location: Vertex garden

14:45h – 15:45h · TopDown Microarchitecture Analysis
Dr. Harald Servat — Intel, Victor Xirau — BSC — Materials

15:45h – 16:00h · Short Break

Part 3 16:00h – 17:00h

Advanced Memory Systems

16:00h – 16:30h · Heterogeneous Memory Systems
Dr. Toni Peña — BSC — Materials

16:30h – 17:00h · Panel Discussion
Dr. Aleksandar Ilic — INESC-ID, Dr. Harald Servat — Intel, Dr. Toni Peña — BSC, Dr. Petar Radojkovic — BSC

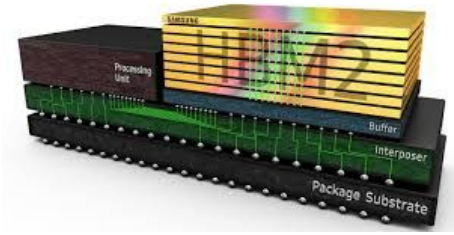
Part 4 17:00h – 19:00h

Networking event

Location: Vertex garden

We will tackle important questions

- **Q1: Which memory system should I select for my application?**



CXL Compute Express Link®

- **Q2: Can I optimize my app to better use a given memory system?**
  - Where is my performance problem?
  - Is there anything I can do about it?

It's the Memory, Stupid!

SCHEDULE

Course Agenda

View all materials

Part 1 10:00h – 10:45h

Introduction

Dr. Petar Radojkovic – BSC

Part 2 10:45h – 16:00h

Performance Analysis: Memory Systems & Applications

10:45h – 11:30h · Roofline(s)
Dr. Aleksandar Ilic — INESC-ID — Materials

11:30h – 12:00h · Coffee Break
Location: Vertex garden

12:00h – 12:45h · Memory System Performance is Messy
Pouya Esmaili-Dokht — BSC, Victor Xirau — BSC — Materials

12:45h – 13:15h · PROFET: No-stress Performance Prediction
Mariana Carmin — BSC — Materials

13:15h – 14:45h · Lunch Break
Location: Vertex garden

14:45h – 15:45h · TopDown Microarchitecture Analysis
Dr. Harald Servat — Intel, Victor Xirau — BSC — Materials

15:45h – 16:00h · Short Break

Part 3 16:00h – 17:00h

Advanced Memory Systems

16:00h – 16:30h · Heterogeneous Memory Systems
Dr. Toni Peña — BSC — Materials

16:30h – 17:00h · Panel Discussion
Dr. Aleksandar Ilic — INESC-ID, Dr. Harald Servat — Intel, Dr. Toni Peña — BSC, Dr. Petar Radojkovic — BSC

Part 4 17:00h – 19:00h

Networking event

Location: Vertex garden

The best memory-related tools presented by (one of) the best experts

- Three different views to similar problems
  - In the same day, in the same room, without a bloodshed (we will see about that …)

Thank you!

Universitat Politecnica de Catalunya – BarcelonaTech

The Facultat d'Informàtica de Barcelona (FIB) is one of the top 10 computer science schools in Europe. We have the resources, professors and experience to offer a wide range of courses.

DARE – The Digital Autonomy with RISC-V in Europe

The DARE project aims to boost European digital transformation by developing novel computing and AI technologies

EVITA – The EuroHPC Virtual Training Academy

EVITA is building a high-quality training framework for High-Performance Computing (HPC) and emerging technologies across Europe. EVITA connects leading institutions to deliver recognised qualifications and modular learning content that supports academia, industry, and innovation. Together, we're shaping the future of HPC education.

Barcelona Zettascale Lab

At the Barcelona Zettascala Lab, our mission is to position Europe as a global leader in high-performance computing (HPC). Through innovation in advanced chip design, we aim to strengthen the continent's technological sovereignty and foster a robust ecosystem encompassing both hardware and software.

La Salle Barcelona – Ramon Llull University

International university campus offering studies in ICT Engineering, Architecture, Business, Digital Arts, Animation, Philosophy and Humanities and Health Engineering. Technology, artificial intelligence and innovation to face the challenges of today´s society.

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

Thank you!

Dr. Harald Servat
Software Enabling and Optimization Engineer

Alexandre Rodrigues
PhD Student

Victor Xirau
Research Engineer
Memory Systems for HPC and AI

Dr. Aleksandar Ilic
Associate Professor

Dr. Leonel Sousa
Associate Professor

Javier Beiro
Research Engineer
Memory Systems for HPC and AI

Dr. Eduard Ayguadé
Department Director
Computer Sciences

Dr. Petar Radojkovic
Group Leader
Memory Systems for HPC and AI

Pau Diaz
Research Engineer
Memory Systems for HPC and AI

Dr. Xavier Martorell
Group Leader
Parallel Programming Models

Dr. Toni Peña
Group Leader
Accelerators and Communications for HPC

Dr. Maria-Ribera Sancho
Head of Education
Education & Training

Pouya Esmaili-Dokht
Research Engineer
Memory Systems for HPC and AI

Mariana Carmin
Research Engineer
Memory Systems for HPC and AI

Jana Arabí Gracia
Organisation Coordination and Suppo
Education & Training

Thank you!

MIQUEL MORETO
HPC ARCHITECTURE RESEARCH AREA DIRECTOR

Computer Sciences - High Performance Domain-Specific Architectures

OSMAN SABRI UNSAL UDER
COMPUTER ARCHITECTURE FOR PARALLEL PARADIGMS GROUP MANAGER

Computer Sciences - Computer Architecture for Parallel Paradigms

JANINE MARIE GEHRIG LUX
SENIOR SCIENCE DISSEMINATION OFFICER

Project Management, Innovation & Dissemination - Dissemination

ROMANA KONJEVOD
PROJECT DISSEMINATION OFFICER

Project Management, Innovation & Dissemination - Dissemination

Ariadna Rodriguez Hernandez

Position: JUNIOR DISSEMINATION PROJECT OFFICER
Project Management, Innovation & Dissemination - Dissemination
email: ariadna.rodriguez@bsc.es


Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación
EXCELENCIA
SEVERO
OCHOA
Thank you!

mess@bsc.es

Your opinion is important

- https://form.typeform.com/to/wD0ngb1r

SCHEDULE

Course Agenda

View all materials

Part 1 10:00h – 10:45h

Introduction

Dr. Petar Radojkovic – BSC

Part 2 10:45h – 16:00h

Performance Analysis: Memory Systems & Applications

10:45h – 11:30h · Roofline(s) — Materials
Dr. Aleksandar Ilic — INESC-ID

11:30h – 12:00h · Coffee Break
Location: Vertex garden

12:00h – 12:45h · Memory System Performance is Messy — Materials
Pouya Esmaili-Dokht — BSC, Victor Xirau — BSC

12:45h – 13:15h · PROFET: No-stress Performance Prediction — Materials
Mariana Carmin — BSC

13:15h – 14:45h · Lunch Break
Location: Vertex garden

14:45h – 15:45h · TopDown Microarchitecture Analysis — Materials
Dr. Harald Servat — Intel, Victor Xirau — BSC

15:45h – 16:00h · Short Break

Part 3 16:00h – 17:00h

Advanced Memory Systems

16:00h – 16:30h · Heterogeneous Memory Systems — Materials
Dr. Toni Peña — BSC

16:30h – 17:00h · Panel Discussion
Dr. Aleksandar Ilic — INESC-ID, Dr. Harald Servat — Intel, Dr. Toni Peña — BSC, Dr. Petar Radojkovic — BSC

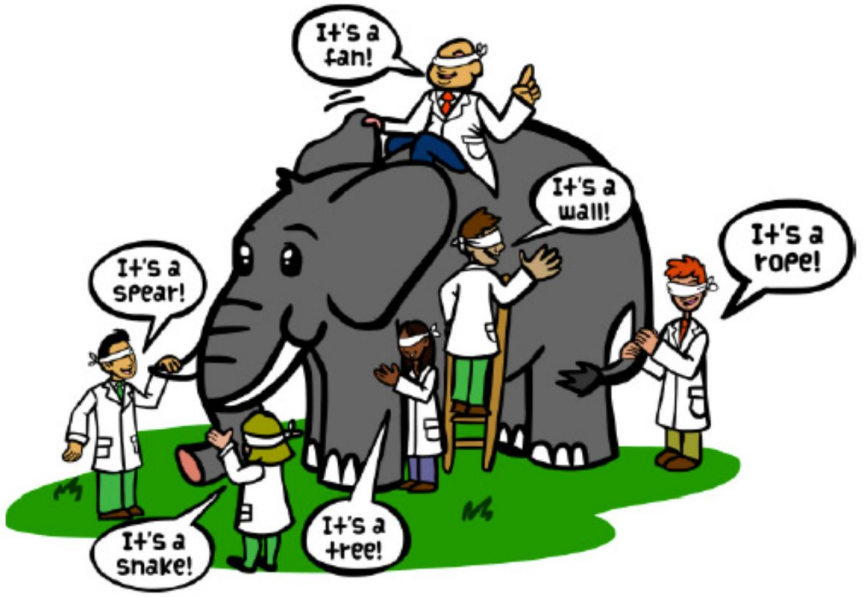
Part 4 17:00h – 19:00h

Networking event

Location: Vertex garden

The best memory-related tools presented by (one of) the best experts

… but maybe we are still just a bunch of blind people touching an elephant



The “Blind Men and the Elephant” tale illustrates how perception is based on what a person is able to see or touch. In the story, six blind men touch an elephant. Although each man touches the same animal, his determination of the elephant is based only what he is able to perceive. The tale warns the reader that preconceived notions and perceptions can lead to misinterpretation.