Introduction

Computer Architecture

UTEC
UNIVERSIDAD DE INGENIERÍA Y TECNOLOGÍA

CS3051 - 2025I
PROF.: JGONZALEZ@UTEC.EDU.PE
SRC: HARRIS, HARRIS - DIGITAL DESIGN AND COMPUTER ARCHITECTURE

Dr. Jorge Luis Gonzalez Reaño
jgonzalez@utec.edu.pe
Assistant Professor
Department of Computer Science

He obtained his BS degree in Electronic Engineering from the Universidad Privada Antenor Orrego, MSc in Microelectronics from the University of São Paulo in 2013, and a PhD in Computer Science from the University of Campinas in 2021. Currently he is a research collaborator at the Lightwave Research Laboratory of Columbia University, NY. His current research interests are in computer architecture, optical interconnects, memory systems and intra-chip traffic.

Executive Summary

- **Motivation:** Computer Architecture studies and proposes novel system architectures using both low- and high-level analysis.
- **Problem:** Computing systems technology is continuously evolving and software trends are constantly changing.
- **Overview:**

 - CS3051 course logistics.

 - Computer Architecture Introduction and show abstraction levels.

 - Review of data number representation and introduce the fundamentals of digital logic.

- **Conclusion:** Computer Architecture provides insight of the processor design paradigm to propose next-gen designs and allows the designer to create efficient software.

Outline

Course Logistics

Introduction

Design Abstraction

Data Representation

Computer Architecture Perspective

Conclusions

CS5051 Computer Architecture

- **Introduces** fundamental **concepts of modern processor design** using a **theoretical and practical approach**.
- **Objectives:**

  - Show the details of how the processor works.

  - Explain the interaction between hardware and software.

  - Model a microprocessor with a limited set of instructions.
- **Content:**

  Distributed in two modules:

  - First module, from Week 1 to Week 8

  - Second module, from Week 9 to Week 16

Evaluation

- For more details, please refer to the course Syllabi in Canvas or:

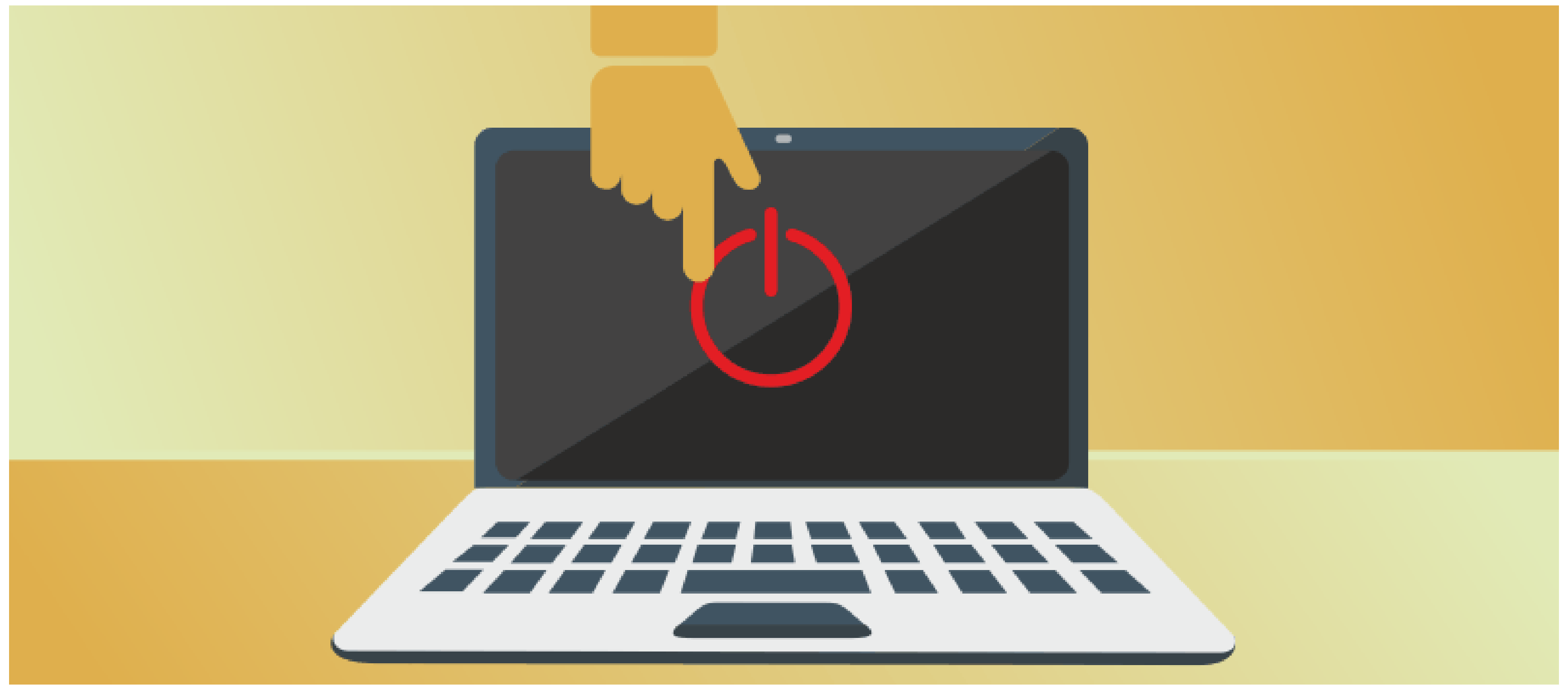
<table>
<tr><td rowspan="4">EVALUATION

*only if the student pass both theory and laboratory parts of the course</td><td>THEORY</td><td>LABORATORY</td></tr>
<tr><td>Exams (25%) (E1)
Exams (25%) (E2)</td><td>2 Lab Tests (20%) (PC)
1 Project (10%)(P)</td></tr>
<tr><td>50%</td><td>50%</td></tr>
<tr><td colspan="2">100%</td></tr>
</table>

Lab: sessions consist on complete lab guides and challenges for points on PC1, PC2, E1 and E2 (up to 3 points).

UTEC UNIVERSIDAD DE INGENIERÍA Y TECNOLOGÍA

Recall: Clasroom



- Do not use cellphones or laptops. **Follow instructor guidelines.**
- **Be on time.**

Important Rules

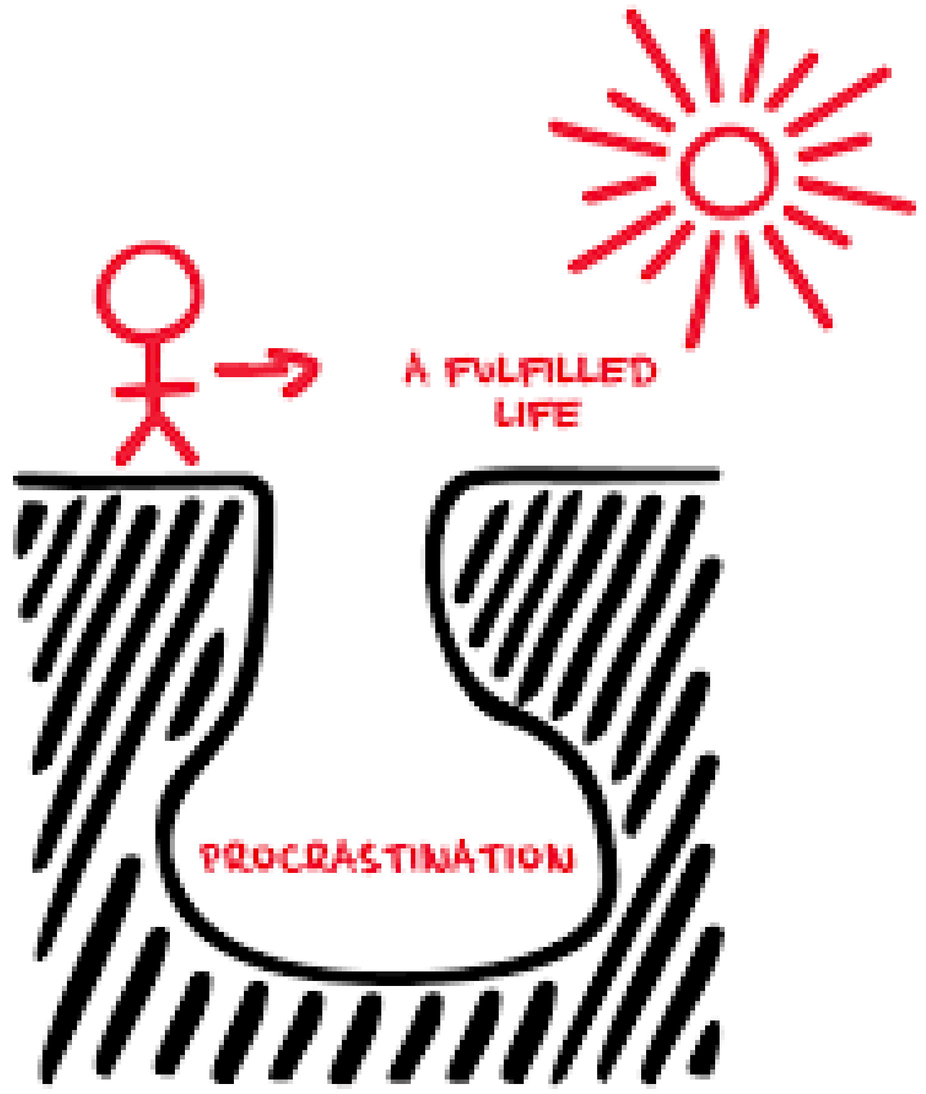
- UTEC rules:
- https://app.utec.edu.pe/sites/default/files/pdf/reglamento_de_disciplina_de_los_estudiantes_2022.pdf
- https://docs.google.com/presentation/d/1RFQFLW5zdmqS-OvsifBF5L8Mi6Vl4DwL9GjkOB3GHaw/edit#slide=id.g11d5e4a44d5_1_76

- **Do not:**
  - **Publish your solution repos online or share with other students.**
  - **Use partial or entire solutions and code implementations** from: a) online repositories, c) or other students (including those who have already taken the course).
    - **In doubt, ask the instructors and TA.**
- **Do:**
  - Discuss ideas and problems with other students
  - Ask the TA and instructor.
  - **Check** Lecturas Sugeridas file in Canvas. Solve final chapter exercises

Logistics

- We recommend students to attend all classes.
  - Work hard! Be honest and gentle!
- We encourage students to formulate questions.
  - Do not be afraid to ask!

- **Ask for help:**
  - Please contact bienestarestudiantil@utec.edu.pe in case you fight procrastination, digital addiction, etc. or need counseling.



UTEC UNIVERSIDAD DE INGENIERÍA Y TECNOLOGÍA

Logistics

- **All course communication via Canvas. Use email for emergencies only.**
- Do not send direct messages, use the server channels. Do not be afraid to ask ☺
  - **Arch** forum: do not send code, solutions, or illegal items.
  - https://piazza.com/utec.edu.pe/fall2025/cs3051

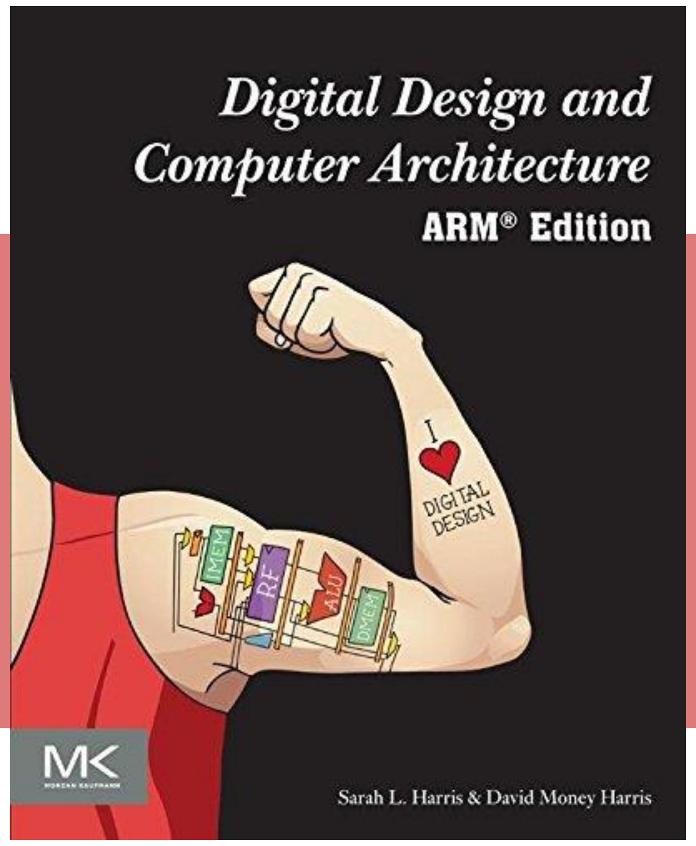
- **TA: Marcelo Chincha marcelo.chincha@utec.edu.pe**
  - **Attention hours (virtual):** TBD, 2 hours weekly
- **TA: Lucas carranza lucas.carranza@utec.edu.pe**
  - **Attention hours (virtual):** TBD, 2 hours weekly
- Gives hints, FPGA testing, solves exercises provided by instructor, makes jokes, recommends anime.
  - Talk with your TA ☺

Al ways remember!

1 You can do it!

2 Work hard!

Books



Harris, S., & Harris, D. (2015). *Digital Design and Computer Architecture: ARM Edition*. Morgan Kaufmann.

LaMeres, B. J. (2019). *Quick Start Guide to Verilog* (pp. 13-22). Springer, Cham.

UTEC
UNIVERSIDAD DE INGENIERÍA Y TECNOLOGÍA

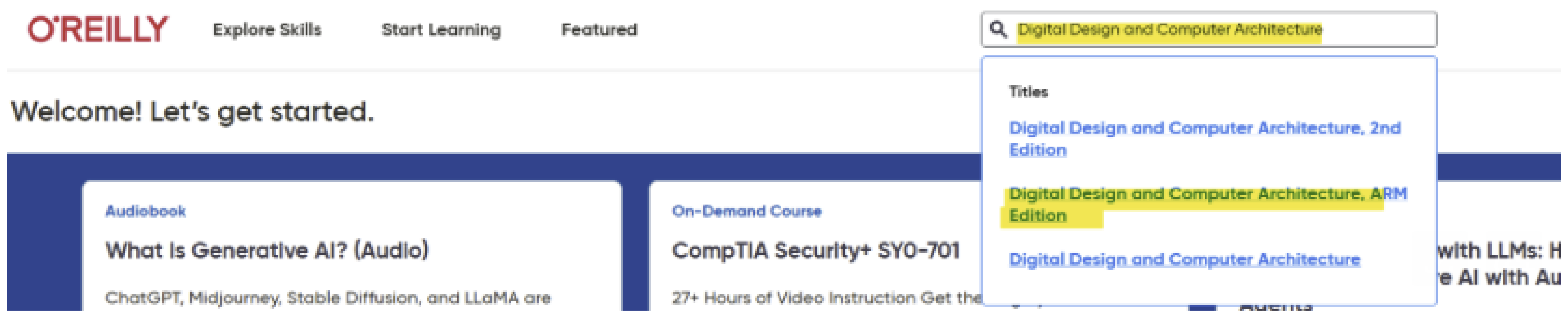
Ebook: How to access?

- https://ci.utec.edu.pe

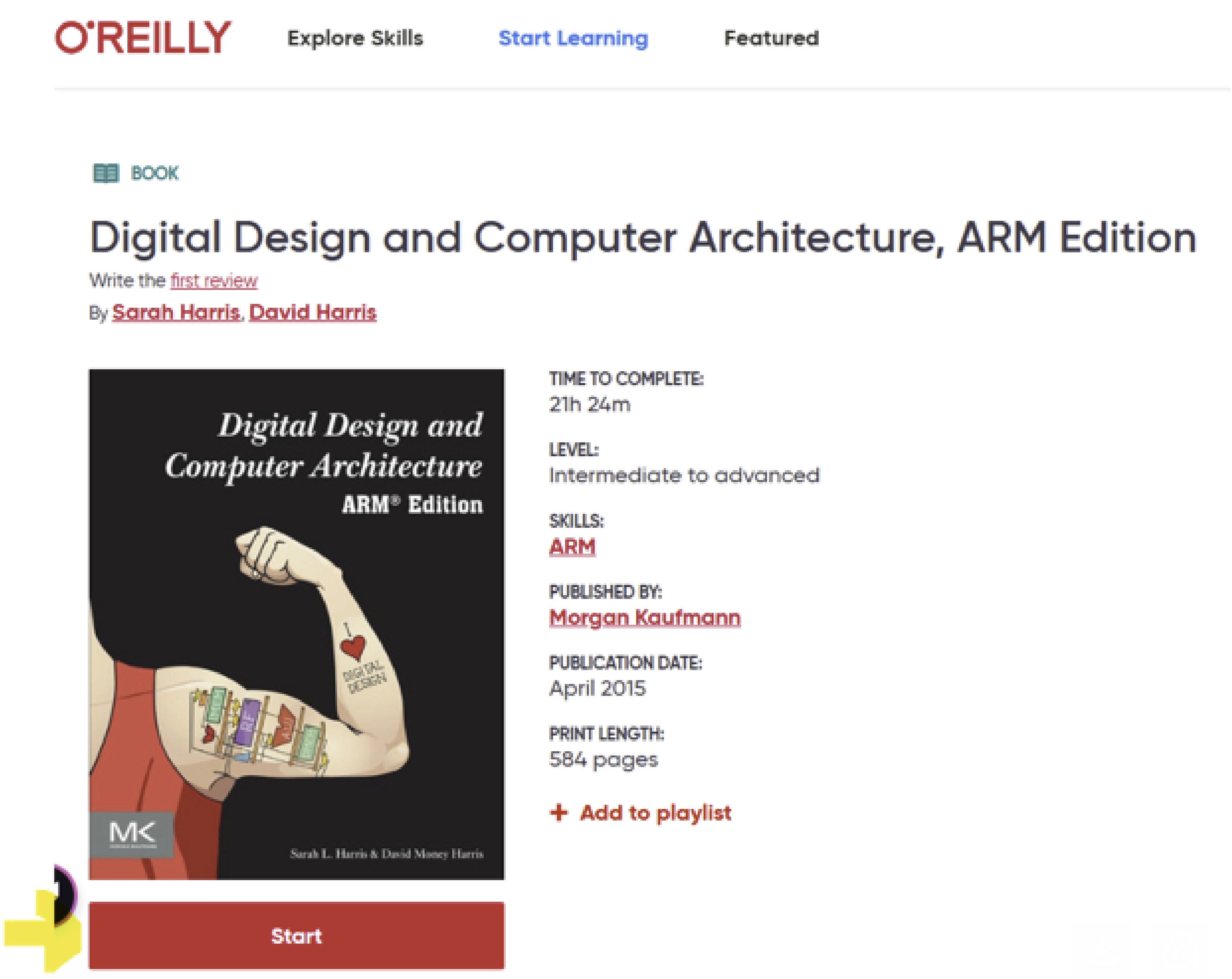
1

Para acceder al libro electrónico le recomendamos seguir los siguientes pasos:

1. Ingrese a la web de la **biblioteca** desde el navegador **Chrome** o **Edge**.
2. Clic en la opción **MyLOFT** (parte superior de la pág. web).
3. Ingrese con su correo UTEC.
4. Active la extensión de MyLOFT y ubique el logo de **O'Reilly.**
5. En la casilla de búsqueda, coloque el título del libro y seleccione haciendo clic en el título.



6. Clic en **"start"** para comenzar a leer el libro.



2

Outline

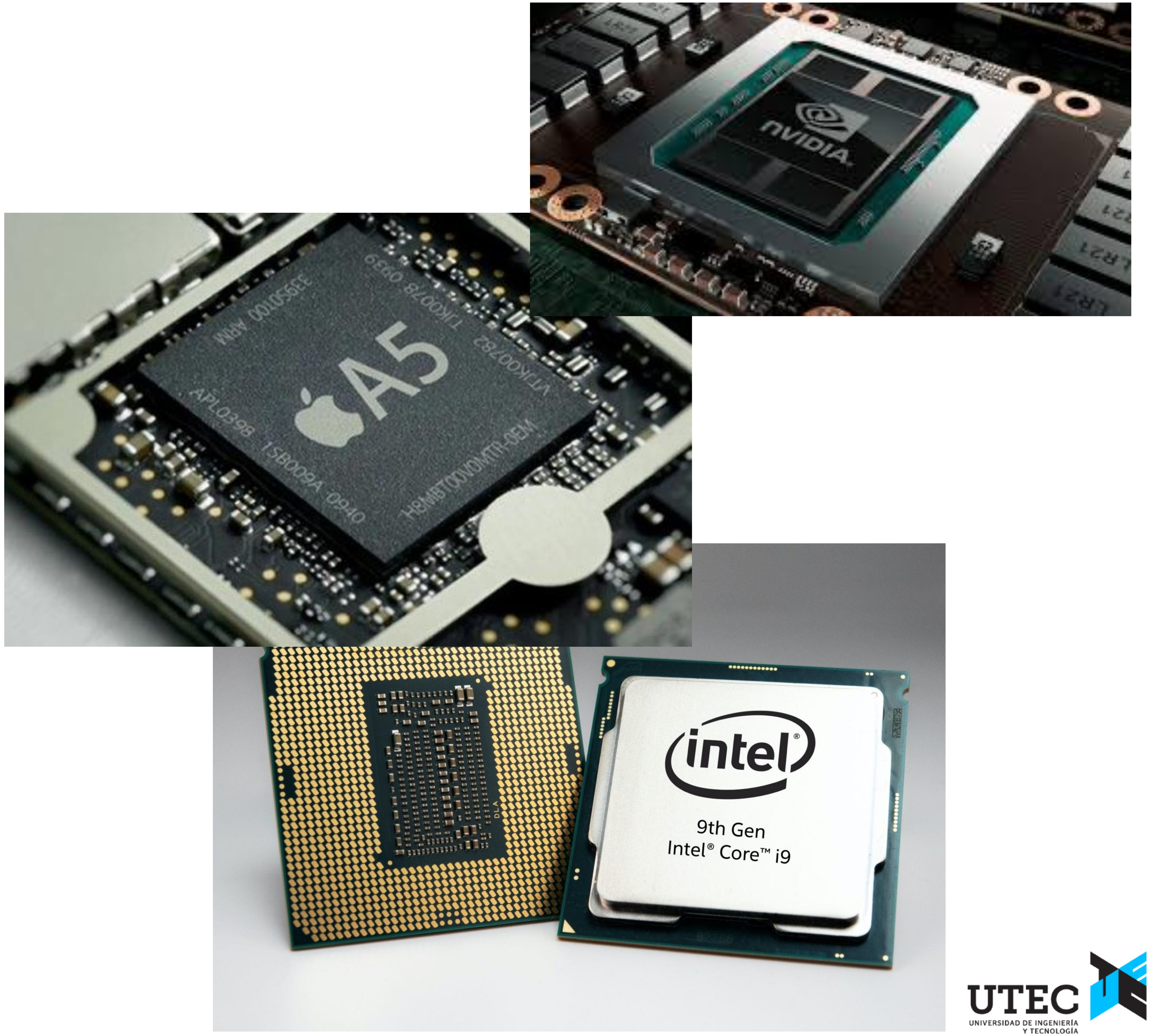
- Course Logistics
- Introduction
- Design Abstraction
- Data Representation
- Logic Gates
- Conclusions

Microprocessor Design

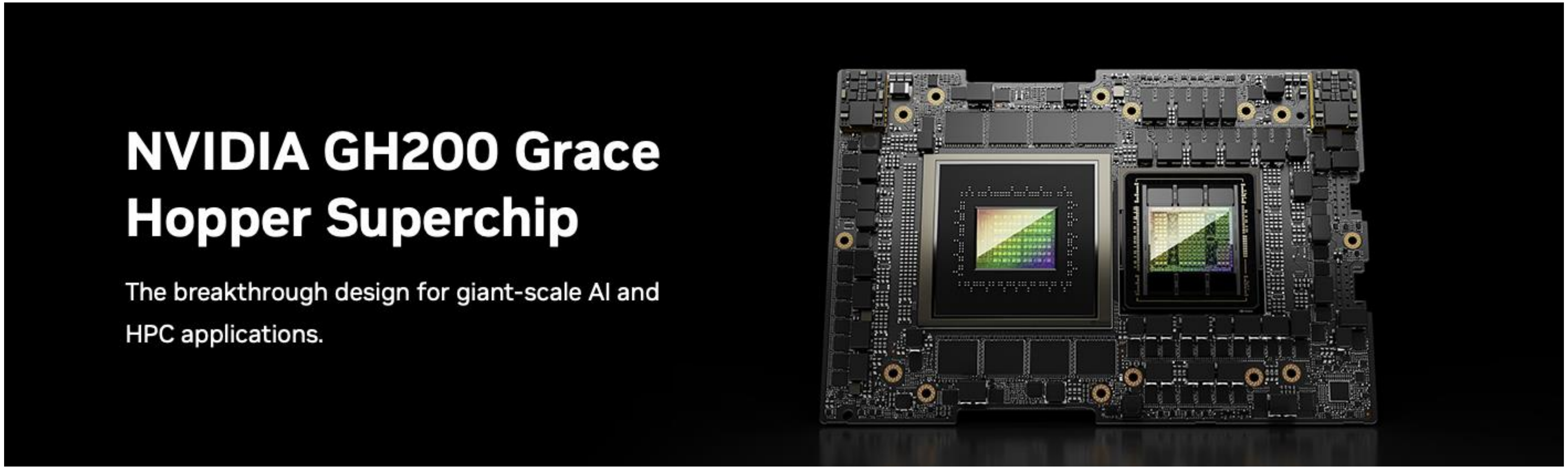
https://www.youtube.com/watch?v=_VMYPLXnd7E

Evolution of Computers

- **Microprocessors** are present in our daily activities in different formats.
- Progress in computer technology
  - Underpinned by **Moore's Law**
- Makes **novel applications feasible**
  - Computers in automobiles
  - Cell phones
  - Human genome project
  - World Wide Web



Today on news



SYSTEMS

2

AMD claims Nvidia's Grace CPU Superchip, Arm are no match for its Epyc Zen 4 cores

But does it matter when all Grace needs to is to babysit GPUs?

Tobias Mann Tue 23 Jul 2024 // 07:26 UTC

COMMENT AMD has claimed its current datacenter silicon is already more than twice as fast, and up to 2.75 times more efficient, than Nvidia's Grace CPU Superchips.

The chip design firm's assertions came after its own testing, published last week, in which it considered Nvidia's 2022 Grace CPU Superchip.

That product combines a pair of CPU dies packing 72 Arm Neoverse V2 cores apiece, connects them with a 900GB/sec NVLink chip-to-chip interconnect, and backs that with

Today on science

nature reviews electrical engineering

Explore content About the journal Publish with us Subscribe

nature > nature reviews electrical engineering > review articles > article

Review Article | Published: 11 January 2024

High-speed emerging memories for AI hardware accelerators

Anni Lu, Junmo Lee, Tae-Hyeon Kim, Muhammed Ahosan Ul Karim, Rebecca Sejung Park, Harsono Simka & Shimeng Yu

Nature Reviews Electrical Engineering **1**, 24–34 (2024)

Journals & Magazines > Journal of Lightwave Technology > Volume: 42 Issue: 22

Photonic-Electronic Integrated Circuits for High-Performance Computing and AI Accelerators

Publisher: IEEE Cite This PDF

Shupeng Ning ; Hanqing Zhu ; Chenghao Feng ; Jiaqi Gu ; Zhixing Jiang ; Zhoufeng Ying All Authors

Abstract:
In recent decades, the demand for computational power has surged, particularly with the rapid expansion of artificial intelligence (AI). As we navigate the post-Moore's law era, the limitations of traditional electrical digital computing, including process bottlenecks and power consumption issues, are propelling the search for alternative computing paradigms. Among various emerging technologies, integrated photonics stands out as a promising solution for next-generation high-performance

nature photonics

Explore content About the journal Publish with us

nature > nature photonics > articles > article

Article | Open access | Published: 21 March 2025

Three-dimensional photonic integration for ultra-low-energy, high-bandwidth interchip data links

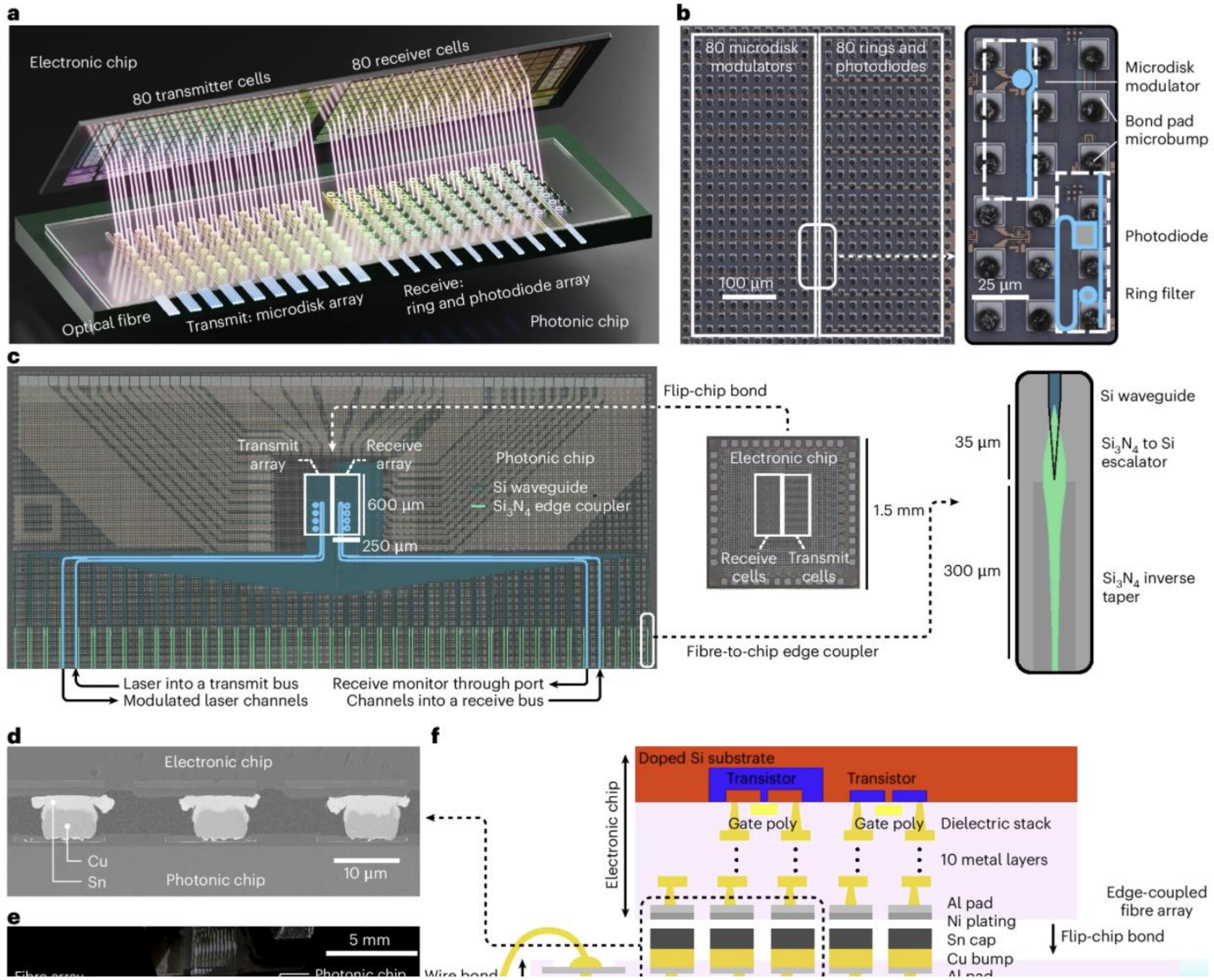
Stuart Daudlin, Anthony Rizzo, Sunwoo Lee, Devesh Khilwani, Christine Ou, Songli Wang, Asher Novick, Vignesh Gopal, Michael Cullen, Robert Parsons, Kaylx Jang, Alyosha Molnar & Keren Bergman

Nature Photonics (2025) | Cite this article

Abstract

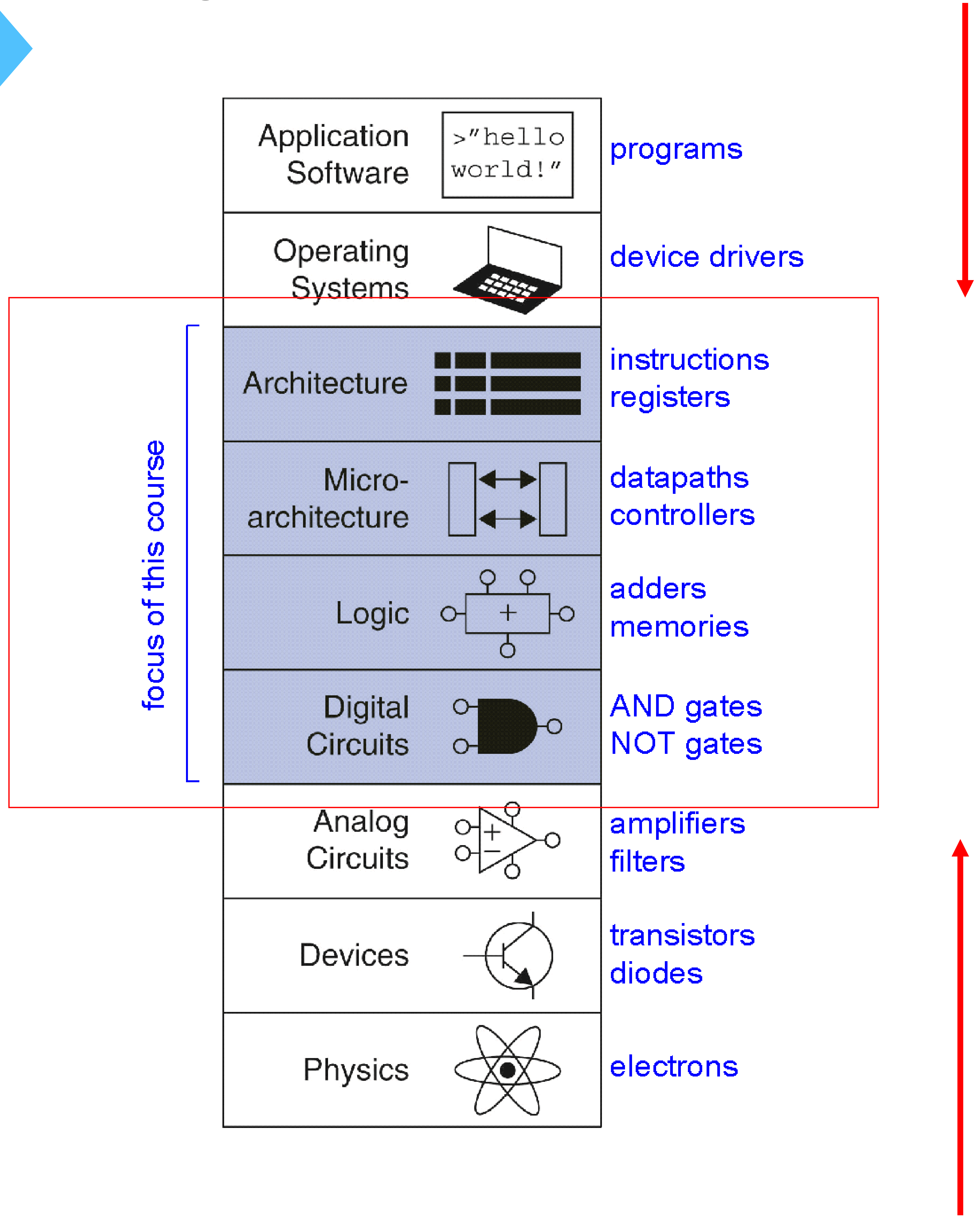
Artificial intelligence (AI) hardware is positioned to unlock revolutionary computational abilities by leveraging vast distributed networks of advanced semiconductor chips. However, a barrier for AI scaling is the disproportionately high energy and chip area required to transmit data between the chips. Here we present a solution to this long-standing overhead through dense three-dimensional (3D) integration of photonics and electronics. With 80 photonic transmitters and receivers occupying a combined chip footprint of only 0.3 mm^2, our platform achieves an order-of-magnitude-greater number of 3D-integrated channels than prior demonstrations. This enables both high-bandwidth (800 Gb s^{-1}) and highly efficient, dense (5.3 Tb s^{-1} mm^{-2}) 3D channels. The transceiver energy efficiency is showcased by a state-of-the-art 50 fJ and 70 fJ per communicated bit from the transmitter and receiver front ends, respectively, operating at 10 Gb s^{-1} per channel. Furthermore, the design is compatible with commercial complementary metal–oxide–semiconductor foundries fabrication on 300-

From: Three-dimensional photonic integration for ultra-low-energy, high-bandwidth interchip data links



Outline

Computer Architecture: Abstraction for Innovation



- **Computer Architecture** studies the actual processor paradigms to propose new architectures for future systems.

- In design, **higher abstraction** reduces the design complexity.
- From the abstraction levels, we will focus on the architectural levels.

UTEC UNIVERSIDAD DE INGENIERÍA Y TECNOLOGÍA

Outline

UTEC
UNIVERSIDAD DE INGENIERÍA Y TECNOLOGÍA

Binary Numbers Representation

- Bits

10010110

most significant bit

least significant bit

- Bytes & Nibbles

byte

10010110

nibble

- Bytes

CEBF9AD7

most significant byte

least significant byte

N-bit binary number:

- How many values? **2^N**
- Range: **$[0, 2^N - 1]$**
- Example: 3-digit binary number:
  - $2^3 = 8$ possible values
  - Range: $[0, 7] = [000_2$ to $111_2]$

UTEC
UNIVERSIDAD DE INGENIERÍA Y TECNOLOGÍA

Convert a Number from Decimal to Binary

- **Method 1:** Find the largest power of 2 that fits, subtract and repeat

Example:

53_{10}	$2^5 = 32$, then pos. 5 (1)
$53-32 = 21$	$2^4 = 16$, then pos. 4 (1)
$21-16 = 5$	$2^2 = 4$, then pos. 2 (1)
$5-4 = 1$	$2^0 = 1$, then pos. 0 (1)
	$\mathbf{= 110101_2}$

Convert a Number from Decimal to Binary

- **Method 1:** Find the largest power of 2 that fits, subtract and repeat

Example:

53_{10}	2^5 =32, then pos. 5 (1)
53-32 = 21	2^4 =16, then pos. 4 (1)
21-16 = 5	2^2 =4, then pos. 2 (1)
5-4 = 1	2^0=1, then pos. 0 (1)
	= 110101_2

- **Method 2:** Repeatedly divide by 2, remainder goes in next most significant bit

Example:

53_{10} =	53/2 = 26 R1	
	26/2 = 13 R0	
	13/2 = 6 R1	
	6/2 = 3 R0	
	3/2 = 1 R1	
	1/2 = 0 R1	**= 110101_2**

Binary Arithmetic

- Add the following 4-bit binary numbers

$$\begin{array}{r} 1 \\ 1001 \\ +\ \ 0101 \\ \hline 1110 \end{array}$$

- Add the following 4-bit binary numbers

$$\begin{array}{r} 111 \\ 1011 \\ +\ \ 0110 \\ \hline 10001 \end{array}$$

- **Binary addition:** similar to decimal addition.

- Digital systems (e.g., processors) operate on a fixed number of bits.

- Overflow: the result is too big to fit in the available number of bits.

Overflow detection (sign)

- No overflow when adding a positive and a negative number.
- No overflow when signs are the same for subtraction
- Overflow **occurs** when the value affects the sign:
  - When adding two positives yields a negative
  - or, adding two negatives gives a positive
  - or, subtract a negative from a positive and get a negative.
  - or, subtract a positive from a negative and get a positive.

Signed-magnitude Binary Numbers

- 1 sign bit, *N*-1 magnitude bits
- Sign bit is the most significant (left-most) bit
  - –**Positive** number: sign bit = **0**
  - –**Negative** number: sign bit = **1**

$A: \{a_{N-1}, a_{N-2}, \ldots a_2, a_1, a_0\}$ $$A = (-1)^{a_{N-1}} \sum_{i=0}^{N-2} a_i \, 2^i$$

- Example: 4-bit sign-mag representations of ± 6:

+6 = **0110**

- 6 = **1110**

- Range of an *N*-bit sign-magnitude number:

$[-(2^{N-1}-1), 2^{N-1}-1]$

Signed-magnitude Binary Numbers

- 1 sign bit, N-1 magnitude bits
- Sign bit is the most significant (left-most) bit
  - –**Positive** number: sign bit = **0**
  - –**Negative** number: sign bit = **1**

$A: \{a_{N-1}, a_{N-2}, \ldots a_2, a_1, a_0\}$

$$A = (-1)^{a_{N-1}} \sum_{i=0}^{N-2} a_i\, 2^i$$

- Example: **4**-bit sign-mag representations of ± 6:

+6 = **0110**

- 6 = **1110**

- Range of an N-bit sign-magnitude number:

$[-(2^{N-1}-1), 2^{N-1}-1]$

Problems:

- Addition doesn't work, for example -6 + 6:

```
  1110
+ 0110
 10100 (wrong!)
```

- Two representations of 0 (± 0):

```
1000
0000
```

Two's Complement Numbers

- Solve the problems with sign-magnitude numbers.
- msb has value of -2^{N-1}

$$A = a_{N-1}(-2^{N-1}) + \sum_{i=0}^{N-2} a_i \, 2^i$$

- Example:
  - Most positive 4-bit number: **0111**
  - Most negative 4-bit number: **1000**
- The most significant bit still indicates the sign:

  1 = negative, and 0 = positive
- Range of an *N*-bit two's complement number:

  $[-(2^{N-1}), 2^{N-1}-1]$

Two's Complement Inversion

- Flips **the sign** of a two's complement number
- **Method:**
  1. Invert the bits
  2. Add 1
- **Example:** Flip the sign of $3_{10} = 0011_2$
  1. **1100**
  2. **+ 1**

  $1101 = -3_{10}$

- **Example:** Two's complement of $6_{10} = 0110_2$
  1. **1001**
  2. **+ 1**

  $1010_2 = -6_{10}$

- **Example:** What is the decimal value of the two's complement number 1001_2?
  1. **0110**
  2. **+ 1**

  $0111_2 = 7_{10}$, so $1001_2 = -7_{10}$

Addition with Two's complement

- Similar to previously discussed binary addition
- **Example:** Add 6 + (-6) using two's complement numbers

$$\begin{array}{r} 111 \\ 0110 \\ +\quad 1010 \\ \hline 10000 \end{array}$$

- **Example:** Add -2 + 3 using two's complement numbers

$$\begin{array}{r} 111 \\ 1110 \\ +\quad 0011 \\ \hline 10001 \end{array}$$

Number extension

Extend number from *N* to *M* bits (*M* > *N*)

Sign-extension:

- Sign bit copied to msb's
- Number value is same

- **Example 1:**
  - 4-bit representation of 3 = **0**011
  - 8-bit sign-extended value: **0000**0011
- **Example 2:**
  - 4-bit representation of -5 = **1**011
  - 8-bit sign-extended value: **1111**1011

Zero-extension:

- Zeros copied to msb's
- Value changes for negative numbers

- **Example 1:**
  - 4-bit value = 0011 = 3_{10}
  - 8-bit zero-extended value: **0000**0011 = 3_{10}
- **Example 2:**
  - 4-bit value = 1011 = -5_{10}
  - 8-bit zero-extended value: **0000**1011 = 11_{10}

Binary Representation Summary

Number System	Range
Unsigned	$[0, 2^N-1]$
Sign-magnitude	$[-(2^{N-1}-1), 2^{N-1}-1]$
Two's Complement	$[-2^{N-1}, 2^{N-1}-1]$

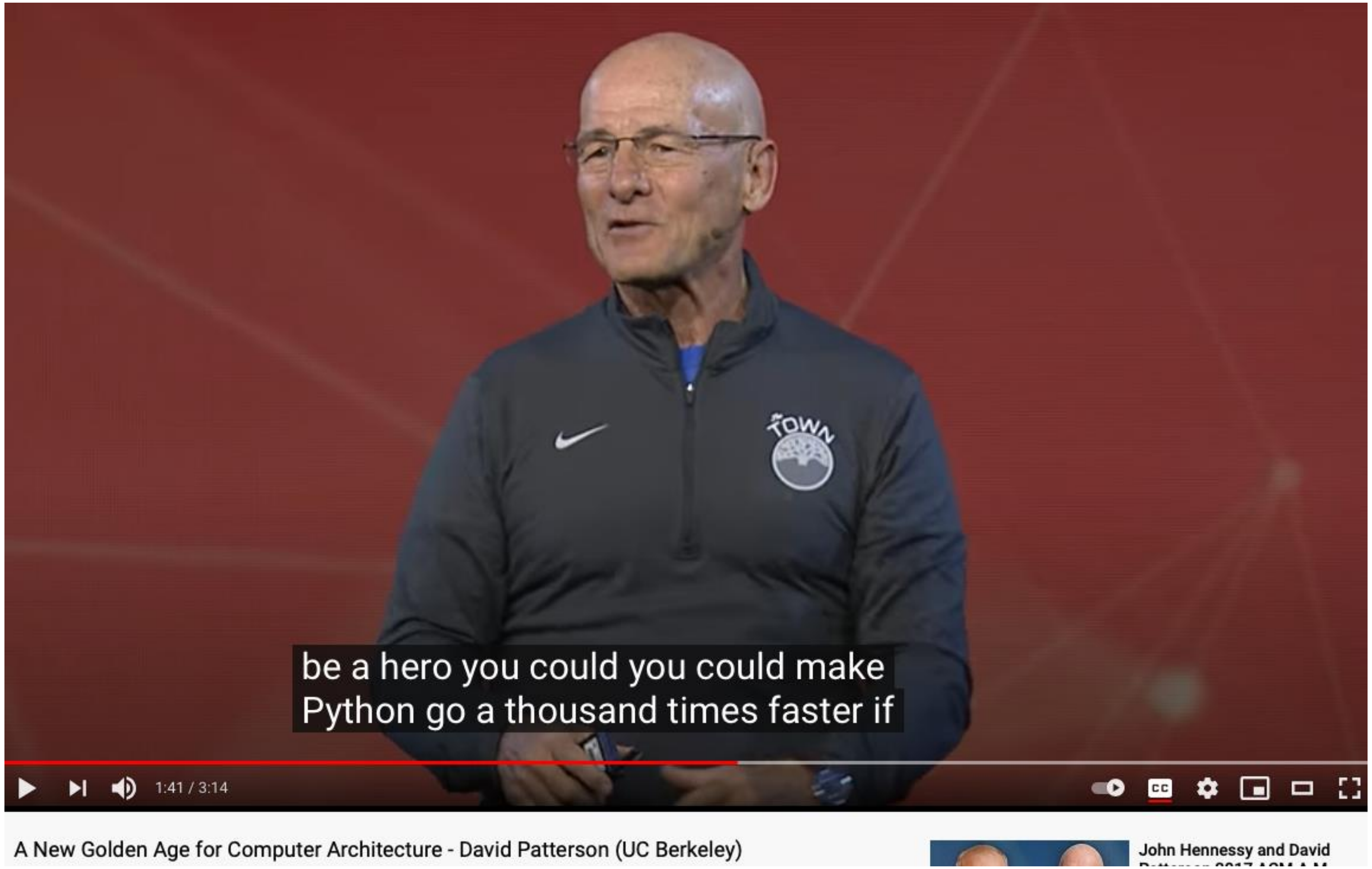
Example, 4-bit representation:

	Unsigned	Two's Complement	Sign/Magnitude
-8		1000	
-7		1001	1111
-6		1010	1110
-5		1011	1101
-4		1100	1100
-3		1101	1011
-2		1110	1010
-1		1111	1001
0	0000	0000	0000 1000
1	0001	0001	0001
2	0010	0010	0010
3	0011	0011	0011
4	0100	0100	0100
5	0101	0101	0101
6	0110	0110	0110
7	0111	0111	0111
8	1000		
9	1001		
10	1010		
11	1011		
12	1100		
13	1101		
14	1110		
15	1111		

Outline

New Golden Age for Computer Architecture

- https://www.youtube.com/watch?v=c03Z0Ms8pKg



David Patterson

- Computer Architect
- Turing Award 2017

UTEC
UNIVERSIDAD DE INGENIERÍA Y TECNOLOGÍA

Outline

Summary

- **Computer Architecture is a very important and active area in Computer Science.**
- Data is represented as fixed size values inside the computing system.
- We reviewed binary number representation.
  - Related to transistor operation (next lab).

We conclude that using the previous concepts, **we can start our design journey!**

Introduction

Computer Architecture

UTEC
UNIVERSIDAD DE INGENIERÍA Y TECNOLOGÍA

CS3051- 2025I
PROF: JGONZALEZ@UTEC.EDU.PE