



Riding the Wave or Paradigm Shift?

EXASCALE COMPUTERS TO TRAIN FOUNDATIONAL MODELS

THOMAS LIPPERT

DECEMBER 4, 2024 | JSC END OF YEAR COLLOQUIUM



Program

11.00-11.10 h *Thomas Lippert*
Welcome on Board

11.10-11.30 h *Thomas Lippert*
Riding the Wave?

11.30-11.50 h *Alexander Strube*
→



11.50-12.10 h *Georgia Psychou*
Setting Sail

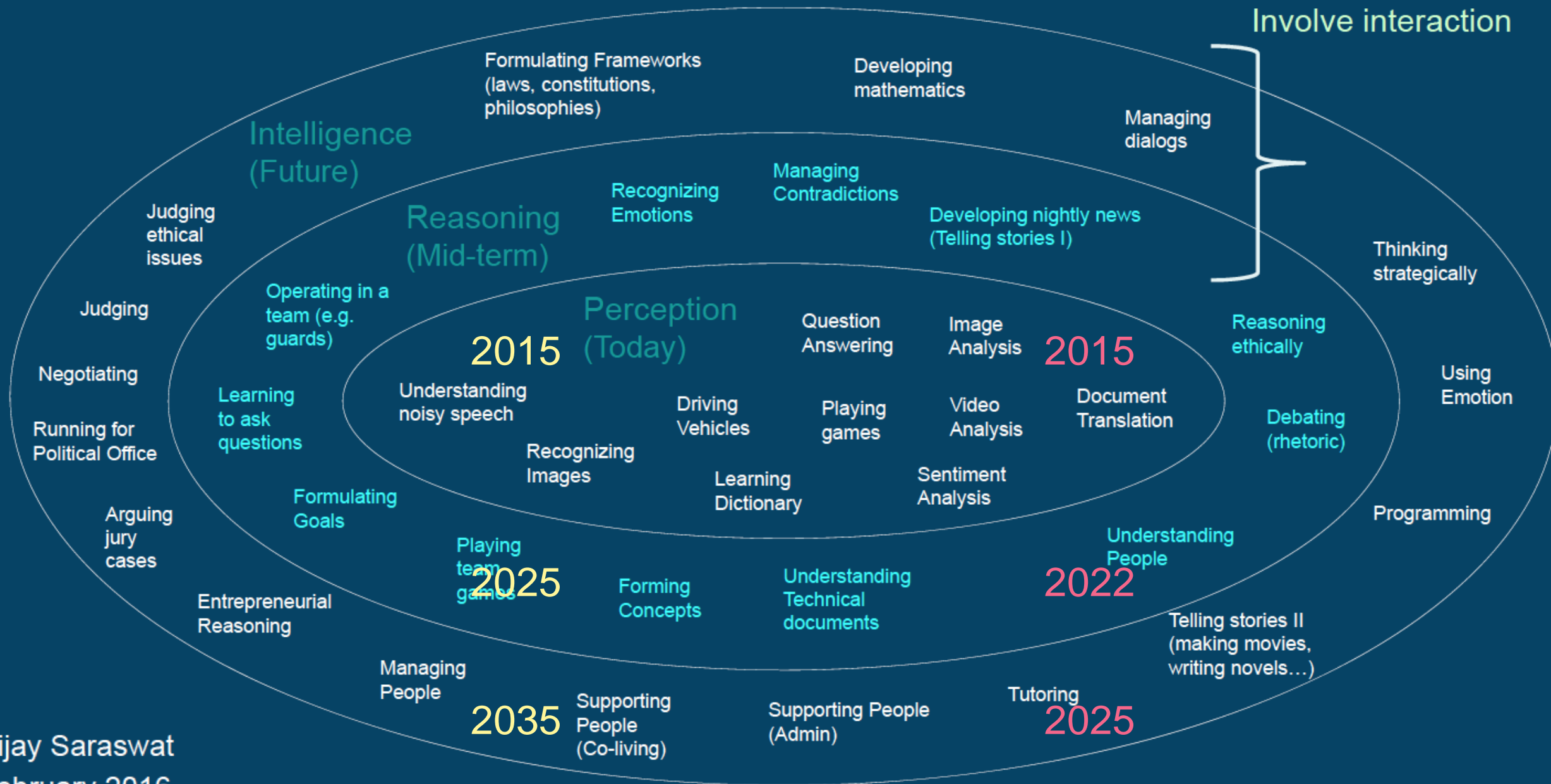
12.10-12.30 h *David Rabanus*
Into the Cold

12.30-12.50 h *Salem El Sayed*
Against all Odds

12.50-13.00 h *Thomas Lippert*
Last Words

13.00-15.00 h *Get-together with poster session*
December 4, 2024 Thomas Lippert

Cognitive Task Landscape



FOUNDATION MODELS

Since ChatGPT ...

- Pre-trained on massive datasets for various tasks
- Transformative for scientific research and industrial applications

PARADIGM SHIFT IN HIGH-PERFORMANCE COMPUTING ?

- Foundation models demand massive computational resources
- HPC evolving to support AI training (alongside traditional simulations?)
- Challenges include energy and data sovereignty

HIGHLIGHT LAION

8B Image data set in Europa



Stable Diffusion 1.5,
trained on JUWELS
using the LAION-5B data set

Prompt:

“Painting by Picasso’s hand a cubistic elephant”



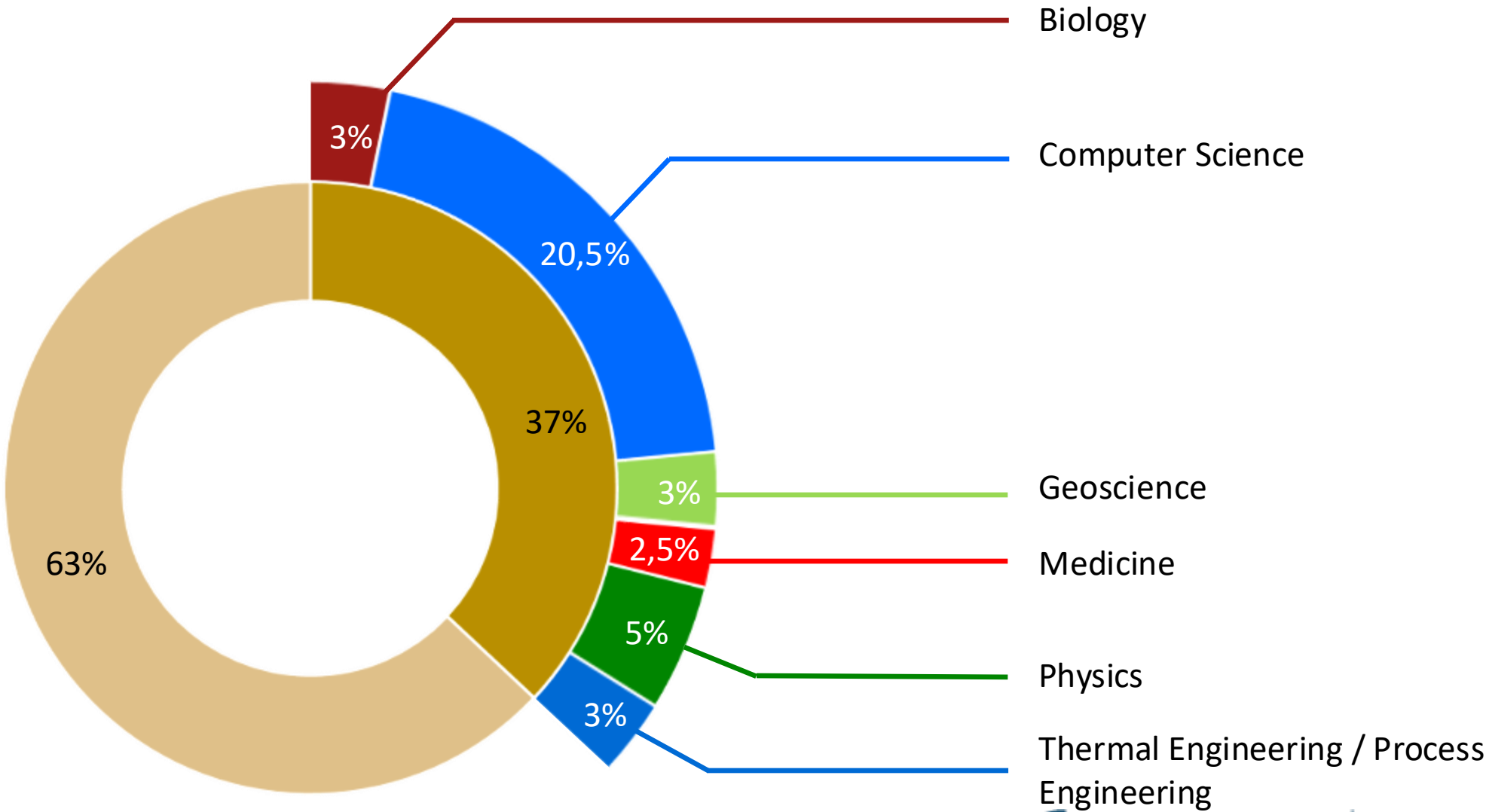
JJ

Outstanding NeurIPS 2022 award

JUWELS – STILL RANK 33 IN TOP500 11/2024

Resource Shares* of GCS/NIC, ESM, and VSR Projects

■ AI-related
■ Not AI-related





*) allocated resource shares (>1%) based on EFLOP on JUWELS (Cluster+Booster) and JURECA-DC (CPU+GPU)

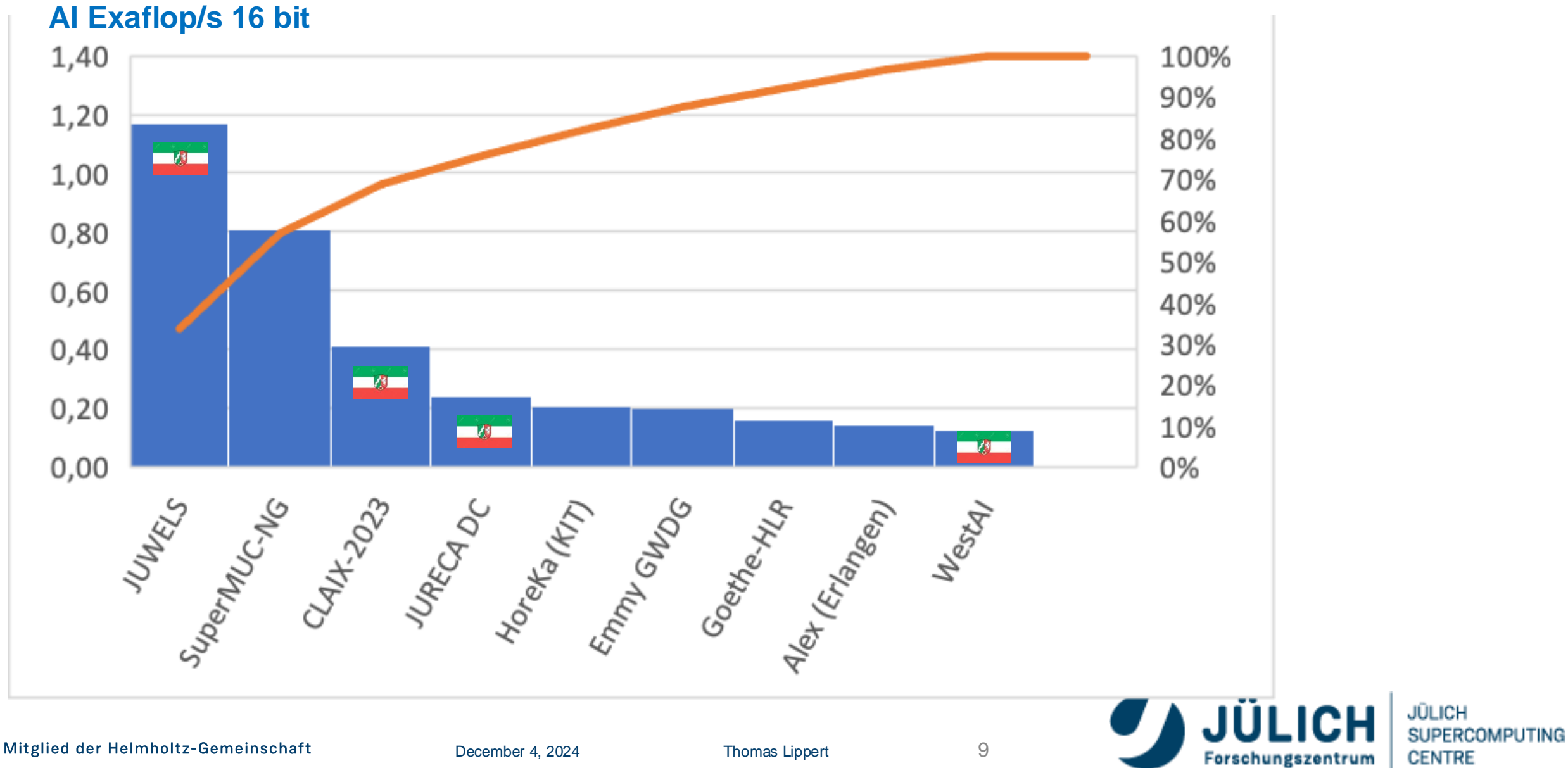
SUPERCOMPUTING: GERMANY AND EU

A developing country?

- Germany lacks hyperscaler-level AI infrastructure
- JUWELS most capable system for AI in Germany
- Europe catching up through EuroHPC JU and national engagement
- Key for Germany and Europe to secure scientific and technological leadership

Rank	Country	# of computers in TOP500 ranking	Maximum performance (TFlops)
1	 US	172	6,475,559
2	 Japan	34	940,710
3	 Italy	13	837,788
4	 Switzerland	5	473,517
5	 Germany	41	405,341
6	 Finland	3	391,388
7	 China	63	319,062
8	 France	24	298,086
9	 Spain	3	221,873
10	 S. Korea	13	213,091
11	 Taiwan	7	103,541
12	 Netherlands	10	98,364
13	 Saudi Arabia	7	96,036

SUPERCOMPUTERS SUITABLE AS AI-TRAINERS IN D – 2024



Enter

JUPITER

The Arrival of
Exascale in Europe



ParTec
MODULAR SUPERCOMPUTING

EVIDEN

fz-juelich.de/jupiter | [#exa_jupiter](https://twitter.com/exa_jupiter)



EuroHPC
Joint Undertaking



Federal Ministry
of Education
and Research

Seite 10

Ministry of Culture and Science
of the State of
North Rhine-Westphalia



GCS
Gauss Centre for Supercomputing

A large supercomputer system, the JEDI, is shown in a data center. The system consists of several tall, dark blue server racks. The first rack on the left has the word "EVIDEN" and "BullSequana XH" printed on it. The second rack features a complex, white, wavy pattern. A bright blue light beam passes through the racks, and a red light beam is visible on the second rack. The floor is made of large, light-colored tiles, and the ceiling has a complex metal structure with lights.


JEDI

JUPITER EXASCALE DEVELOPMENT INSTRUMENT



Rank	TOP500 Rank	System	Cores	Rmax (PFlop/s)	Power (kW)	Energy Efficiency (GFlops/watts)
1	222	JEDI - BullSequana XH3000, Grace Hopper Superchip 72C 3GHz, NVIDIA GH200 Superchip, Quad-Rail NVIDIA InfiniBand NDR200, ParTec/EVIDEN EuroHPC/FZJ Germany	19,584	4.50	67	72.733

JUNE 2024





JETI

Jupiter
Exascale
Transition
Instrument

Rank 18

83 PF/s

Rank 6 in
Green500

Mitglied der Helmholtz-Ge

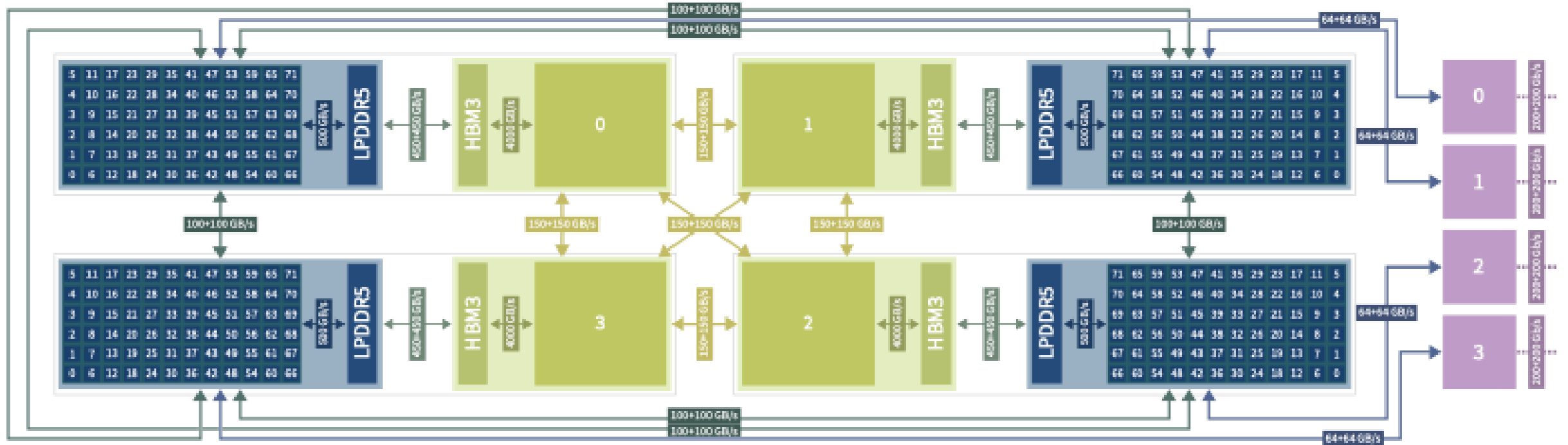


JUPITER: EUROPE'S FIRST EXASCALE SYSTEM

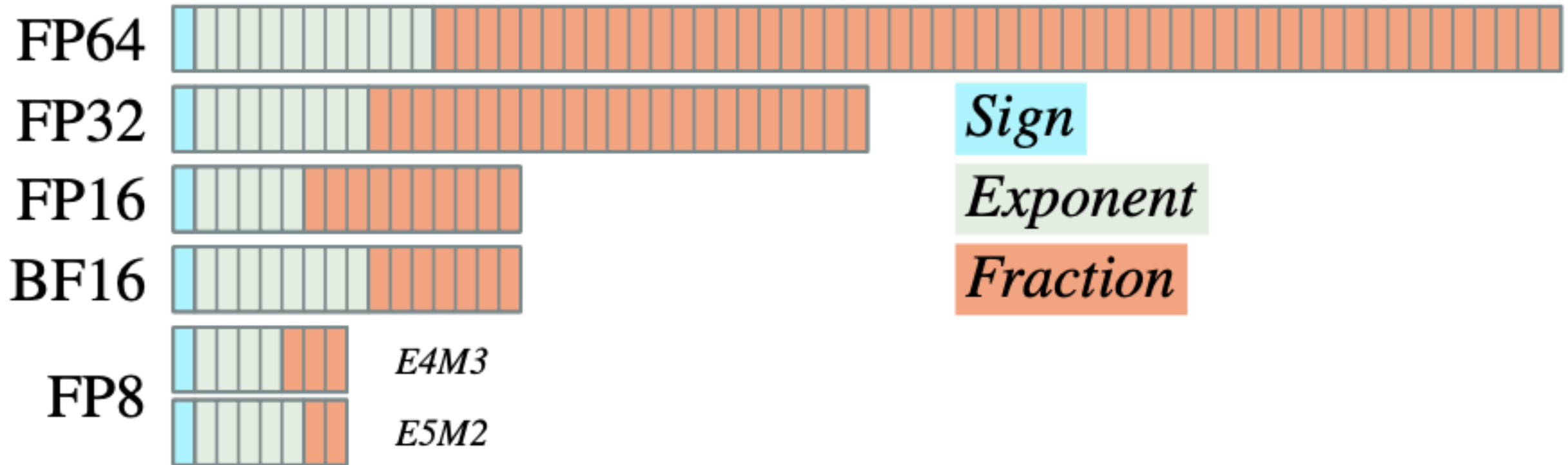
A Paradigm Shift in Itself – Germany will be on rank wto in total power worldwide

- Balances compute, network, and energy efficiency for versatile apps
- Designed to address the AI training and simulation gap
- Capable of 40 to 80 AI exaFLOP/s at 8 bit
- 24,000 NVIDIA GH200 superchips with high interconnect bandwidth.
- DragonFly+ topology for scalability.
- Storage includes 29 PB NVMe systems, 310 PB HDDs and +700 PB tape capacity

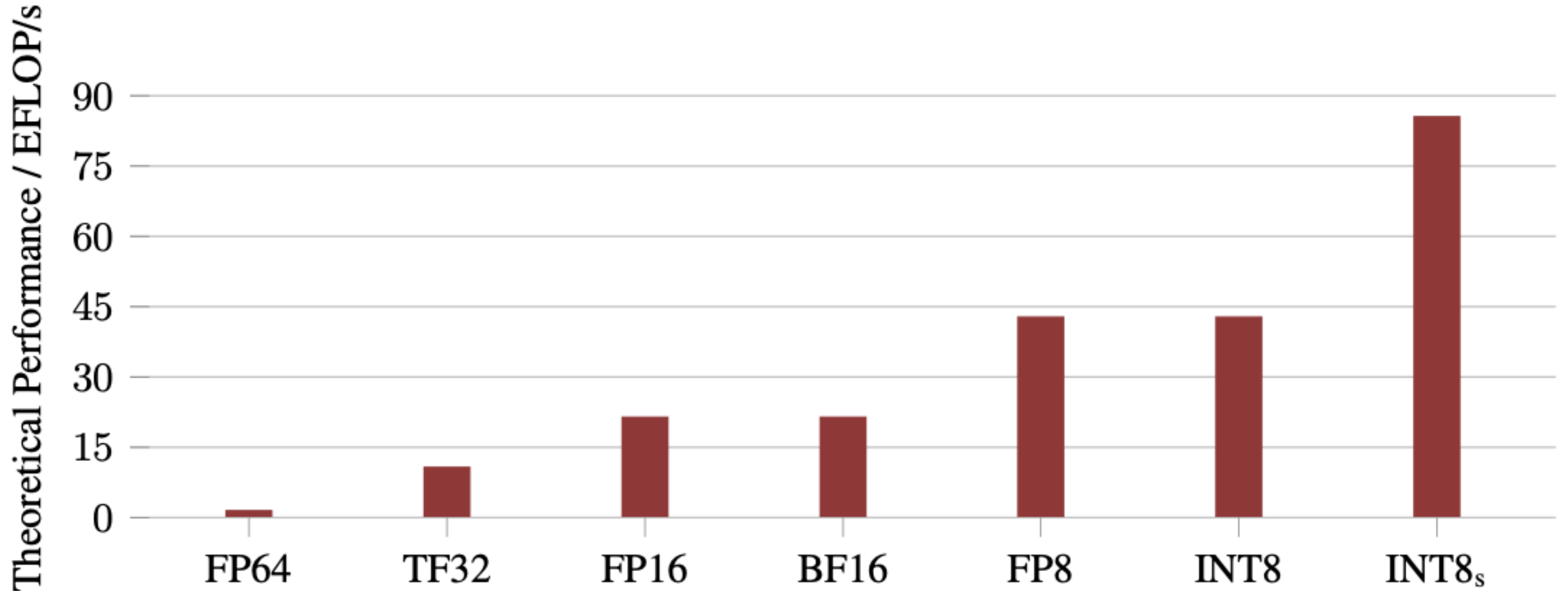
JUPITER BOOSTER COMPUTE NODE DESIGN



ENCODING OF PRECISION



AI PERFORMANCE



JUPITER

Installation in progress

- Logistic containers: September
- Energy containers: Begin October
- Data Centre: Mid October
- HPC-Systems: Mid December
- Begin of operation: End of Jan.
 - JUREAP on JUPITER starts
- April: Full system
- May: LINPACK
- June: Begin of full operation
 - KI Competition
 - 1st regular call





A PARADIGM SHIFT – NOT RIDING THE WAVE



A blissful, pleasant, and relaxing Christmas break to you and your families...

FP7 - H2020
DEEP
Projects

2011



2020

EuroHPC

SEA
Projects

2021

<HPC|@S>

2022

...and a Happy
and Healthy
New Year!

EU PEX

2023

2024

2025

