

Present and Future of GPU HPC in Hungary

National HPC development

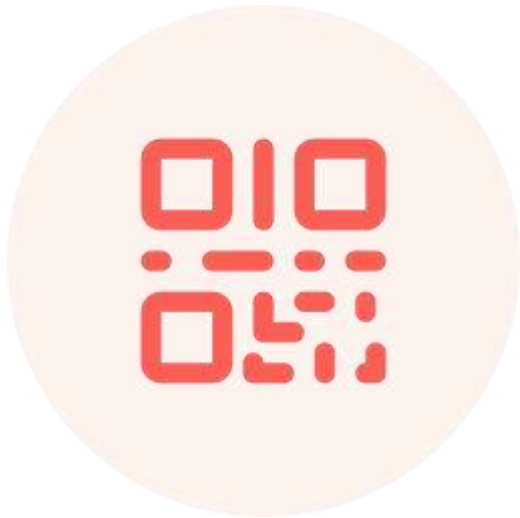
Zoltan Kiss, Head of HPC Dpt.

KIFÜ



HPC @hu
Kompetencia Központ

slido



**Join at slido.com
#2522236**

ⓘ Start presenting to display the joining instructions on this slide.

slido

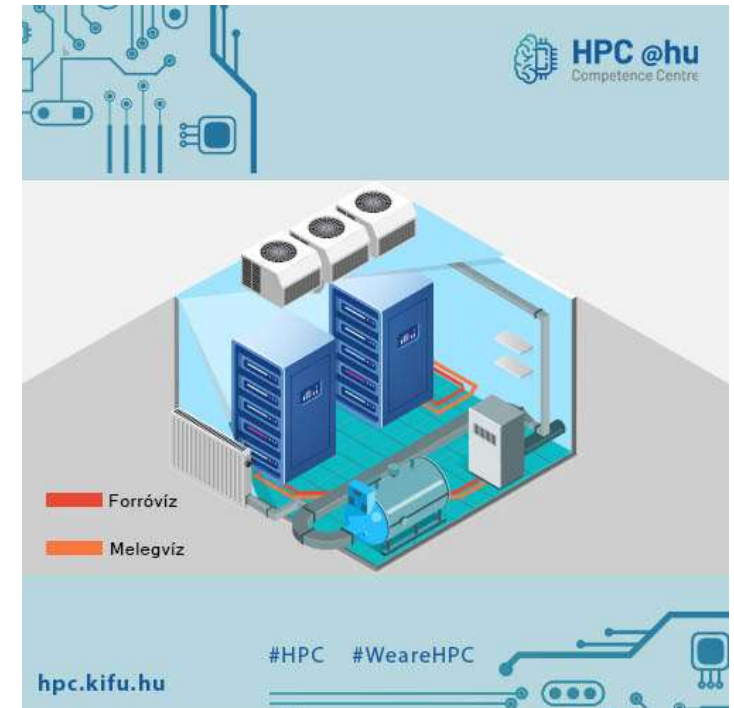


What ICT Infra are you missing the most?

ⓘ Start presenting to display the poll results on this slide.

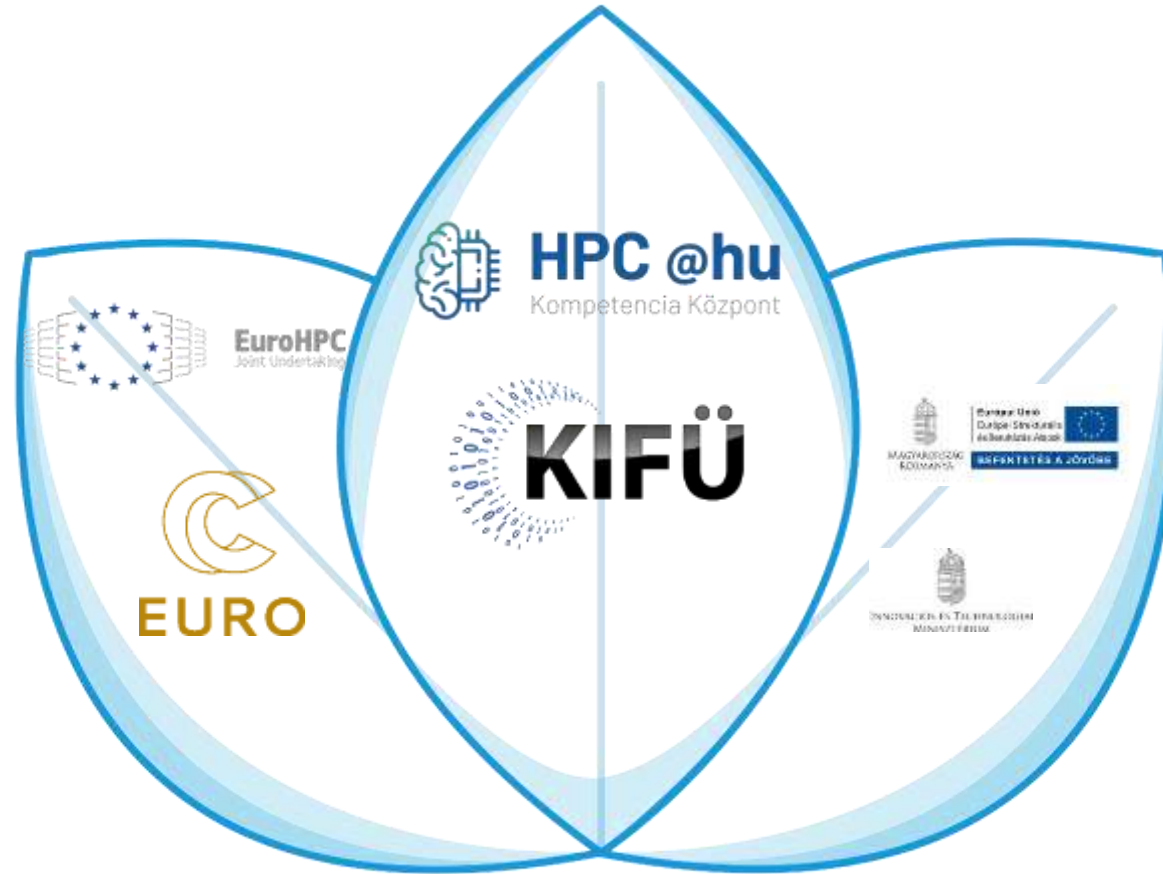
Contents

- 🧠 HPC Competence Centre
- 🧠 IT development
- 🧠 HPC infrastructure
- 🧠 Portals
- 🧠 EuroHPC
- 🧠 Trends and Input



HPC ecosystem

Competence Centre and
SME relations



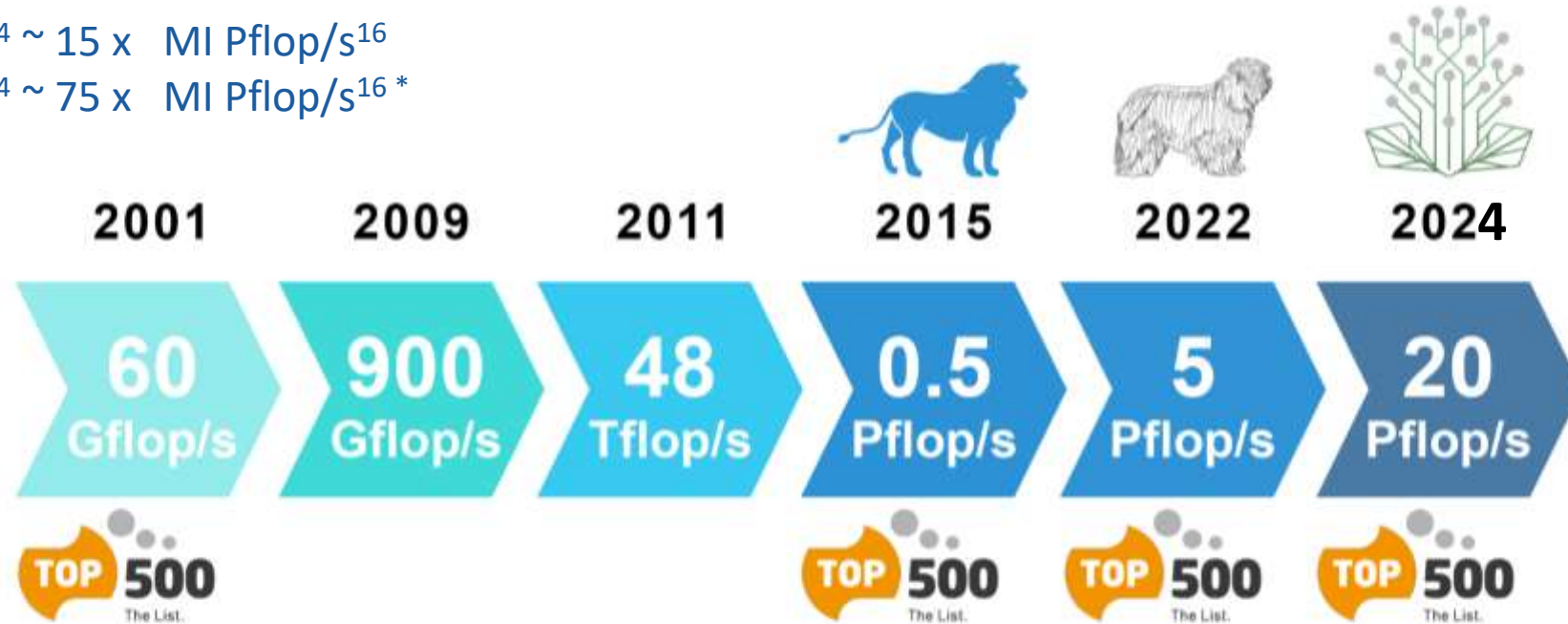
Infrastructure
development



HPC @hu
Competence Centre

20 years of Hungarian HPC

1 HPC Pflop/s⁶⁴ ~ 15 x MI Pflop/s¹⁶
5 HPC Pflop/s⁶⁴ ~ 75 x MI Pflop/s¹⁶*



* A100 values

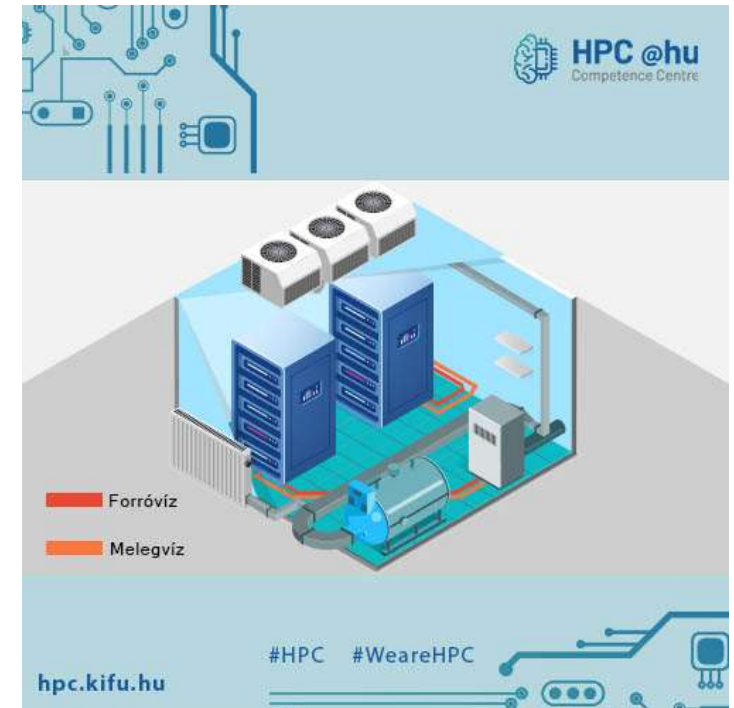


HPC @hu
Competence Centre

Infrastructure development

 Hardware

 Software



HPC @hu
Competence Centre

Green DC

🧠 Hot water used

🧠 90% air cooled

2022 Q3 – 5+ PF HPC

🧠 Efficiency

🧠 Extreme density

🧠 Award winning building designer

🧠 In Debrecen Campus



HPC @hu
Competence Centre

Komondor



HPC @hu
Competence Centre

CRAY EX EXASCALE INFRASTRUCTURE



1.6 ExaFlops (US)



2 ExaFlops (US)



550 Pflops (EU)



180 Pflops (US)



COMPUTING

70 Pflops (US) ~ Komondor specs



Hewlett Packard Enterprise



All Flash entry point (6U)
Up to **80/50 GB/sec** read/write
and **115 TB** capacity*



Disk entry point (10U):
15 GB/sec and
315 TB capacity*

All Flash base rack: > 1TB/sec and up to 4.5 PB capacity*

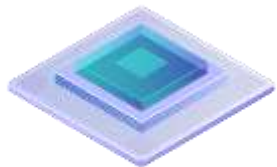
- Expansion rack: > 2 TB/sec and 4.6 PB*

Disk base rack: 90 GB/sec and 7.5 PB capacity*

- Expansion rack: 120 GB/sec and 10 PB*

*usable capacity

20 000+
CPU mag



CPU 0,7+
petaflops

CPU

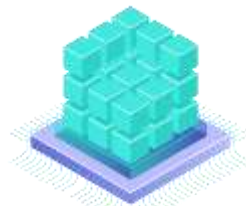
200+
GPU



GPU 4+
petaflops

GPU

Big Data 9+ TB
memory



BIG DATA

8GPU/node



MI

10 PB



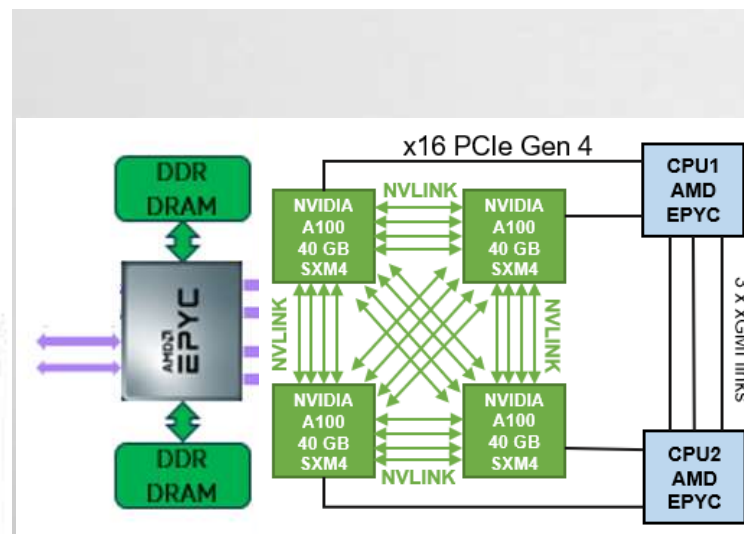
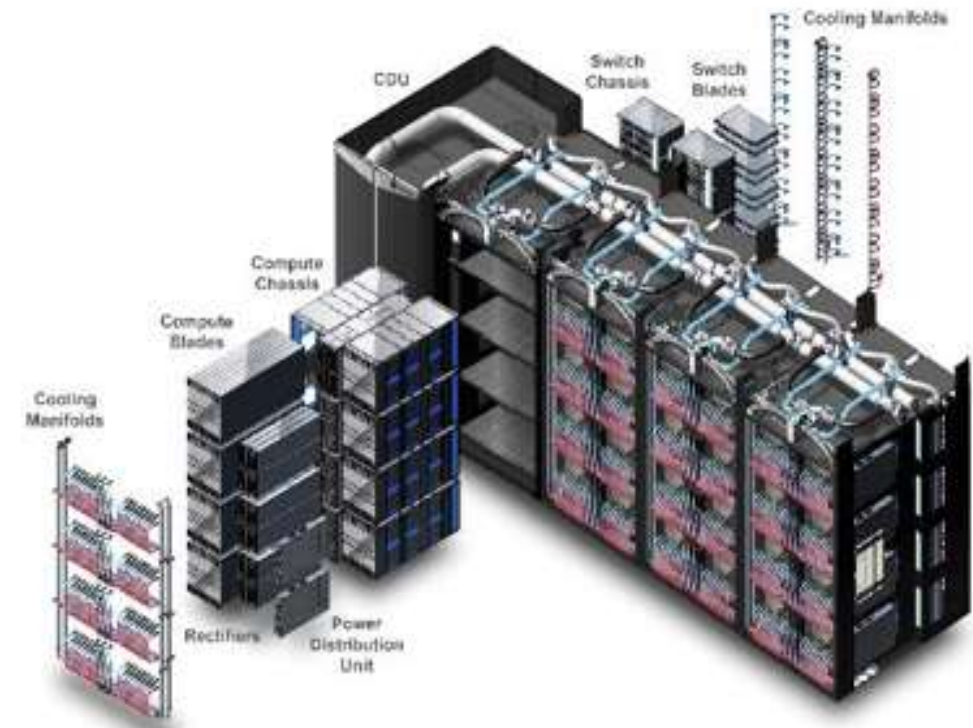
Tape archive

2 PB



Ultrafast

STORAGE





Internet



Tűzfal

Ethernet hálózat
100 Gbps külső kapcsolat



Login node

400 TB ultragyors tárolóegység
(300 Gbyte/s)



1,5 PB gyors tárolóegység
(rövidtávú tárolás)



10 PB hosszútávú archiválás



200 Gbps Slingshot
belső hálózat



CPU only partíció

140 node, egyenként 2 db 64 magos AMD EPYC™ 7763 (Milan) CPU és 256 GB RAM (összesen 17920 core), 200Gb/s Slingshot interconnect
Rpeak=0,7 petaflops



Gyorsított (GPU) partíció

50 node, egyenként 1 db 64 magos AMD EPYC™ 7763 (Milan) CPU és 256 GB RAM és 4 db NVIDIA A100 GPU (összesen 200 db GPU), 2x200Gb/s Slingshot interconnect
Rpeak= 4 petaflops



Mesterséges Intelligencia (AI) partíció

2 node (HPE Apollo 6500 Gen10 Plus), egyenként 2 db 64 magos AMD EPYC™ 7763 (Milan) CPU és 512 GB RAM és 8 db NVIDIA A100 GPU (összesen 16 db GPU), 2x200Gb/s Slingshot interconnect
Rpeak= 0,3 petaflops











Big Data (Data Analytics) partíció

1 node (SMP/NUMA), 12 db 18 magos Intel® Xeon® Gold 6254 CPU (összesen 216 core) és 9 TB RAM, 2 TB SSD, 2x200Gb/s Slingshot interconnect
Rpeak=21 teraflops



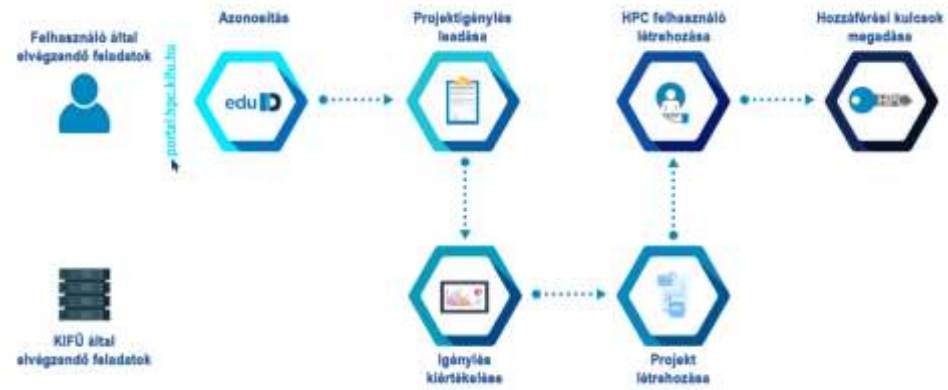
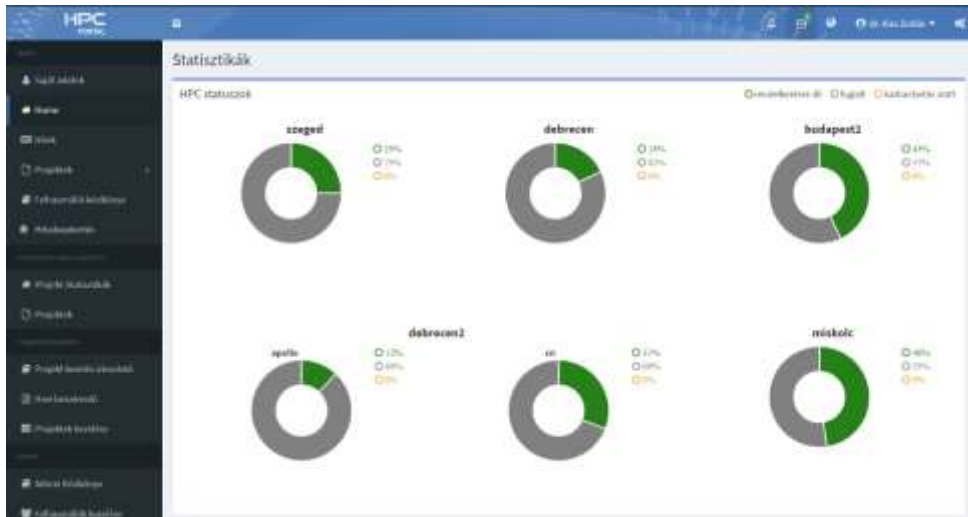
HPC@hu
Competence Centre

Software

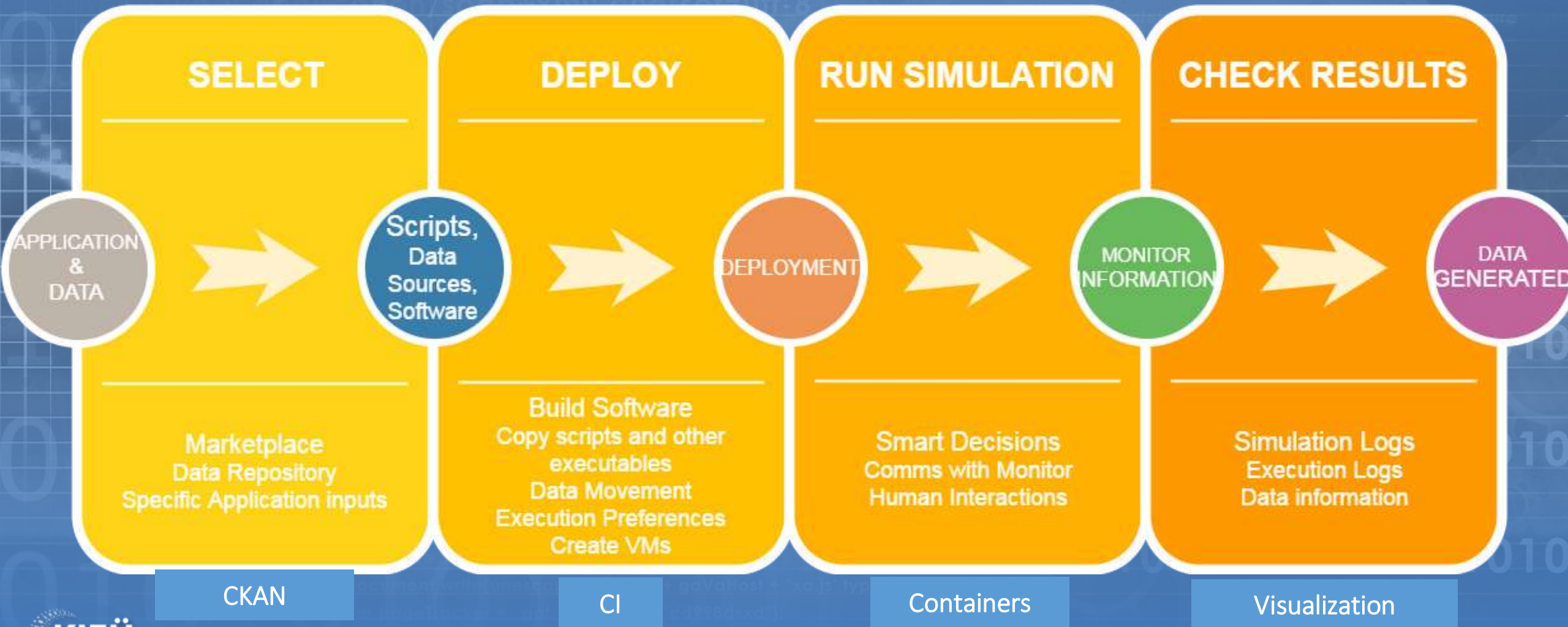
-  Ansys
-  Amber20
-  Amsterdam Density Functional (ADF)
-  CharMM
-  Gaussian 16
-  Terachem
-  Schrödinger
-  TotalView





HPC portal



Deploy workflow - example



EuroHPC

-  20 PF
-  CPU + GPU + Big Data + MI
-  Multilevel storage system
-  Visualization subsystem
-  Modular architecture
-  Quantum HPC integration
-  HU-DE (Juelich – TIM – KIFÜ) cooperation
-  Portal and software developments



HPC @hu
Competence Centre

What's in the oven?

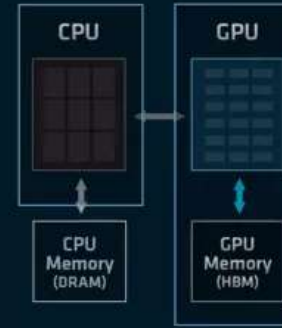
AMD INSTINCT™ MI300 THE WORLD'S FIRST DATA CENTER APU

- 4th Gen AMD Infinity Architecture: AMD CDNA™ 3 and EPYC™ CPU “Zen 4” Together
CPU and GPU cores share a unified on-package pool of memory
- Groundbreaking 3D Packaging
CPU | GPU | Cache | HBM
- Designed for Leadership Memory Bandwidth and Application Latency
- APU Architecture Designed for Power Savings Compared to Discrete Implementation

Available **2023**

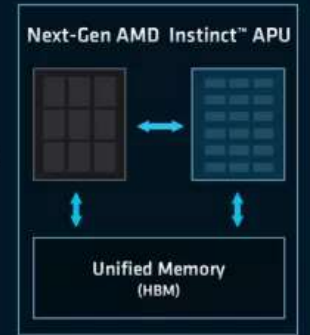
AMD CDNA™ 2 Coherent Memory Architecture

- Simplifies Programming
- Low Overhead 3rd Gen Infinity Interconnect
- Industry Standard Modular Design



AMD CDNA™ 3 Unified Memory APU Architecture

- Eliminates Redundant Memory Copies
- High-Efficiency 4th Gen AMD Infinity Architecture
- Low TCO with Unified Memory APU Package



>8X

Expected AI Training Performance
vs. MI250X

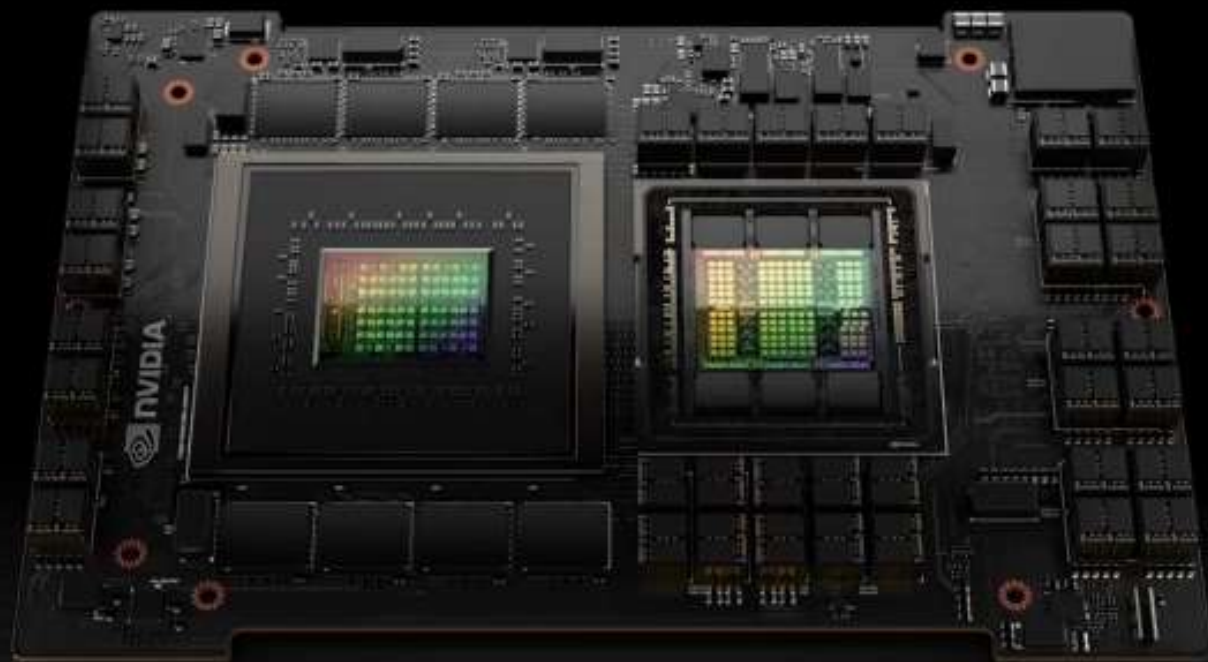
See Endnote MI300-03. Preliminary data and projections, subject to change.

HPC @hu
Competence Centre

What's in the oven?

ANNOUNCING
GRACE CPU SUPERCHIP

Superchips Connected by NVLink C2C
144 CPU Cores
1 TB/s LPDDR5X with ECC
396MB On-Chip Cache
SPECrate®2017_int_base Over 740 (est.)



ANNOUNCING
NVIDIA GRACE HOPPER

Grace Hopper Superchip
Densest NVIDIA Accelerated Computing System
New NVLink Chip-to-Chip Coherent Inference
900 GB/s

Nvidia Datacenter GPUs

Data by Nvidia, H100 specifications not finalized

GPU Performance	NVIDIA A100	NVIDIA H100 SXM5
Peak FP8 Tensor TFLOPS with FP16 Accumulate	NA	2000/4000
Peak FP8 Tensor TFLOPS with FP32 Accumulate	NA	2000/4000
Peak FP16 Tensor TFLOPS with FP16 Accumulate	312/624	1000/2000
Peak FP16 Tensor TFLOPS with FP32 Accumulate	312/624	1000/2000
Peak BF16 Tensor TFLOPS with FP32 Accumulate	312/624	1000/2000
Peak TF32 Tensor TFLOPS	156/312	500/1000
Peak FP64 Tensor TFLOPS	19.5	60
Peak INT8 Tensor TOPS	624/1248	2000/4000
Peak FP16 TFLOPS (non-Tensor)	78	120
Peak BF16 TFLOPS (non-Tensor)	39	120
Peak FP32 TFLOPS (non-Tensor)	19.5	60
Peak FP64 TFLOPS (non-Tensor)	9.7	30
Peak INT32 TOPS	19.5	30

slido



What chip are you using the most now?

ⓘ Start presenting to display the poll results on this slide.


slido



What chip would you see the most potential for your use case in the future? (multiple choice)

ⓘ Start presenting to display the poll results on this slide.

Thank you for your feedback!

 hpc.kifu.hu
 kiss.zoltan@kifu.gov.hu
 [@HPC.CC.hu](https://www.facebook.com/HPC.CC.hu)
 [@HPC_hu](https://twitter.com/HPC_hu)



HPC @hu
Kompetencia Központ

slido



Audience Q&A Session

ⓘ Start presenting to display the audience questions on this slide.