# JUPITER – FROM AN ARM PERSPECTIVE

## ISC2024 Arm BoF

2024-05-14 I ANDREAS HERTEN, JÜLICH SUPERCOMPUTING CENTRE

Member of the Helmholtz Association

EuroHPC Joint Undertaking

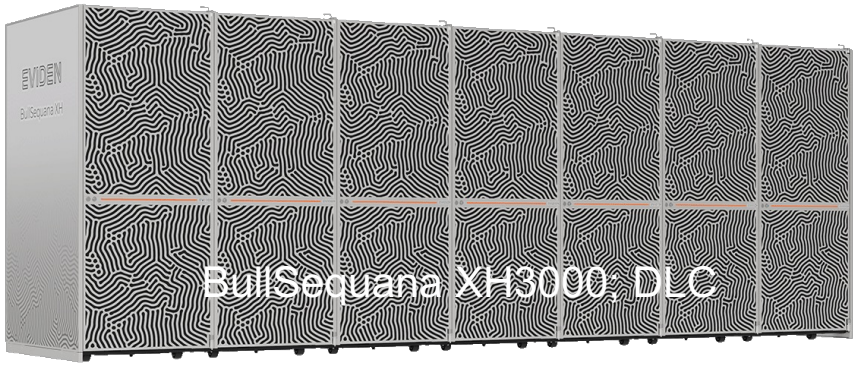
Bundesministerium für Bildung und Forschung

Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen

GCS Gauss Centre for Supercomputing

JÜLICH Forschungszentrum Shaping Change

# DISCOVERING JUPITER

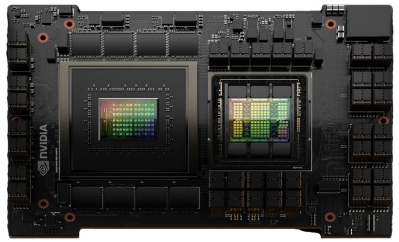



- First Exascale system in Europe (HPL); modular system
- Procured/funded by: EuroHPC JU, BMBF/NRW-MKW
  - Contract signed end of 2023
  - Installation starting soon
- JUPITER Booster: High scalablibty; 1 EFLOP/s HPL, >70 EFLOP/s FP8
  JUPITER Cluster: High versatility; 0.5 B/FLOP balance
- Network: InfiniBand NDR; Storage: 20 PB NVMe, 200 PB HDD
- Deployed in Modular Datacenter
- Building on: MSA (JUWELS); DEEP, EPI; ThunderX2, Ampere; …
- About 1.936.000 Arm cores

# JUPITER MODULES

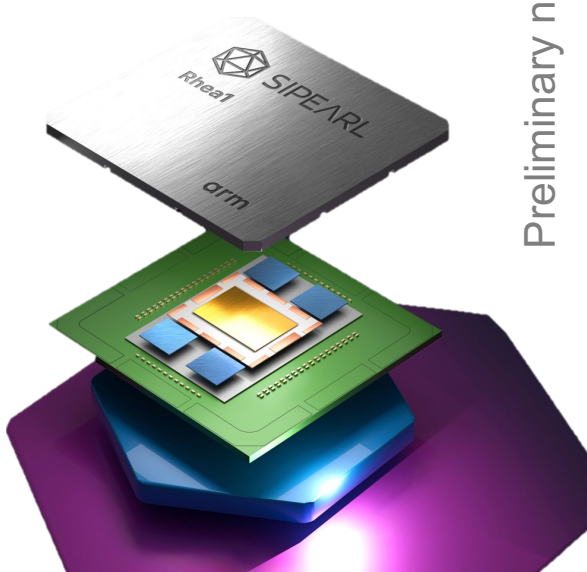
## JUPITER Booster

- Node design
  - ~6000 nodes
  - 4× NVIDIA CG1 per node
- CG1: Grace-Hopper
  - 72 Arm Neoverse V2 cores (4×128b SVE2); 120 GB LPDDR5
  - H100 (132 SMs); 96 GB HBM3
  - NVLink C2C (900 GB/s)



## JUPITER Cluster

- Node design
  - ~1300 nodes
  - 2× SiPearl Rhea1 per node
- Rhea1
  - 80 Arm Neoverse V1 cores (2×256b SVE)
  - 256 GB DDR5, 64 GB HBM2e




Preliminary numbers, might change during installation

# ENABLEMENT: JEDI, JUREAP

- 🖌️ JEDI: JUPITER test system
  - 48 nodes, *currently 24;* JUPITER design

🎉 Top 1 Green500!

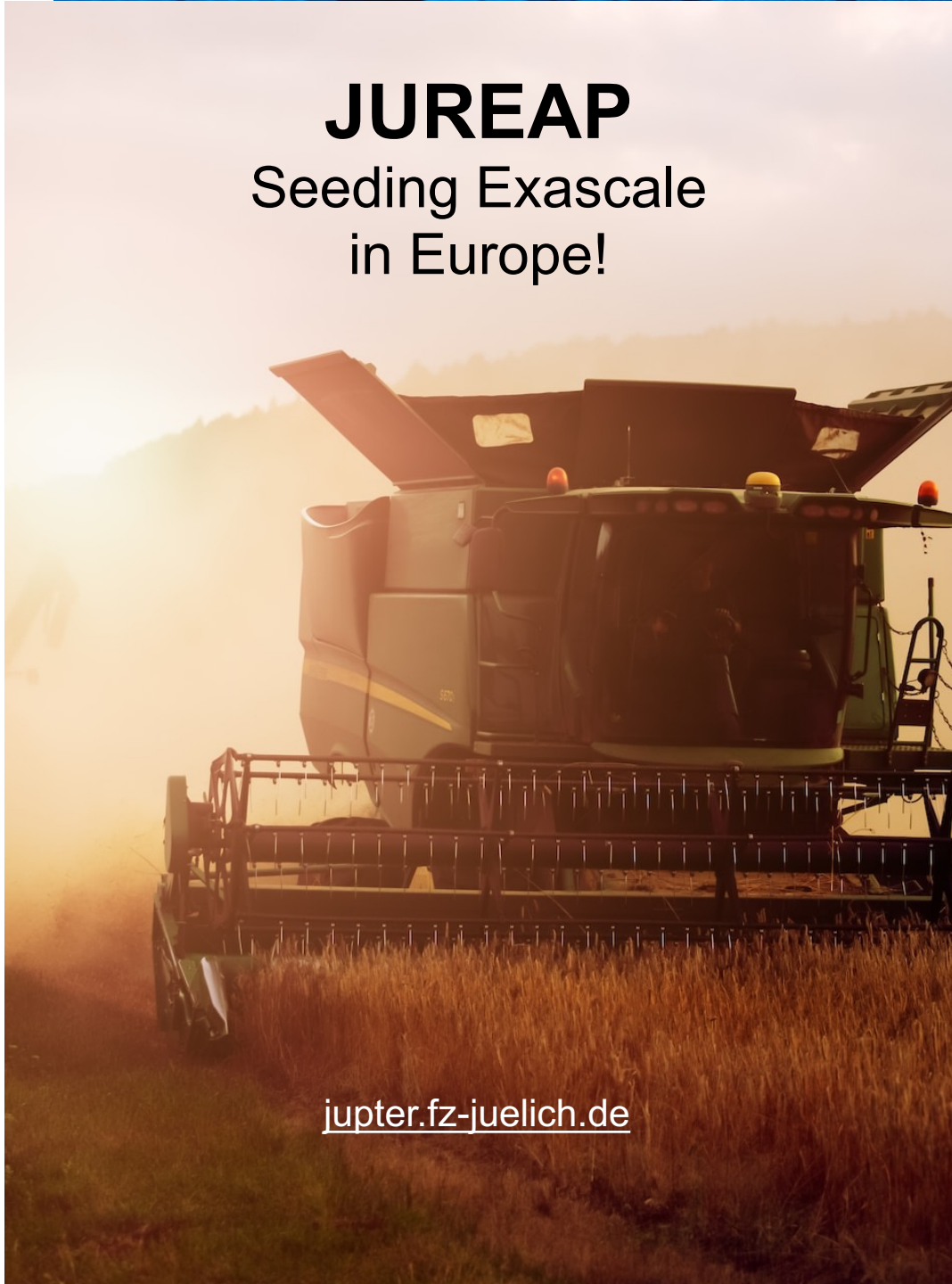
- Usage
  - System management preparations
  - Application porting
  - JUREAP; Research and Early Access Program
    - Currently: Call for Participation
    - Soon: Invitation for Proposals
- Also: GH200 COTS test nodes for some months
  (1 GH, 480 GB LPDDR5, 1000 W TDP)

# ENABLEMENT: JEDI, JUREAP

- 🖌️ JEDI: JUPITER test system
  - 48 nodes, *currently 24;* JUPITER design

🎉 Top 1 Green500!

- Usage
  - System management preparations
  - Application porting
  - JUREAP; Research and Early Access Program
    - Currently: Call for Participation
    - Soon: Invitation for Proposals
- Also: GH200 COTS test nodes for some months
  (1 GH, 480 GB LPDDR5, 1000 W TDP)



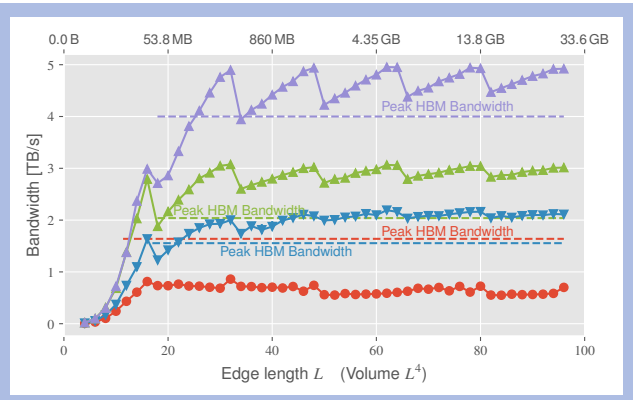
# ENABLEMENT: JEDI, JUREAP

- 🖌️ JEDI: JUPITER test system
  - 48 nodes, *currently 24;* JUPITER design

🎉 Top 1 Green500!

- Usage
  - System management preparations
  - Application porting
  - JUREAP; Research and Early Access Program
    - Currently: Call for Participation
    - Soon: Invitation for Proposals
- Also: GH200 COTS test nodes for some months
  (1 GH, 480 GB LPDDR5, 1000 W TDP)

# FIRST GPU EXPERIENCES *(H100)*




**LQCD** benchmark: Great mem utilization

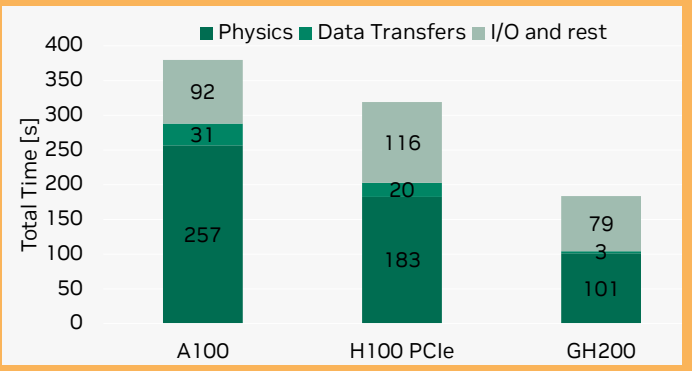
**ChASE**: >2× vs. A100 across all solvers

**ICON**: 1.6× vs. A100 in first benchmark (R2B4)

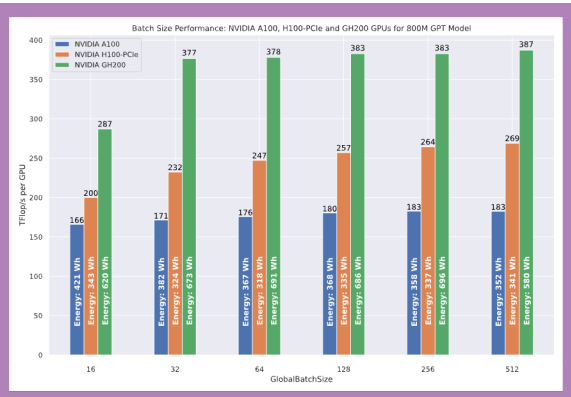
**nekRS**: 2.1× vs. A100 for RBC benchmark

**Arbor**: 1.97× vs. A100 for Busyring benchmark

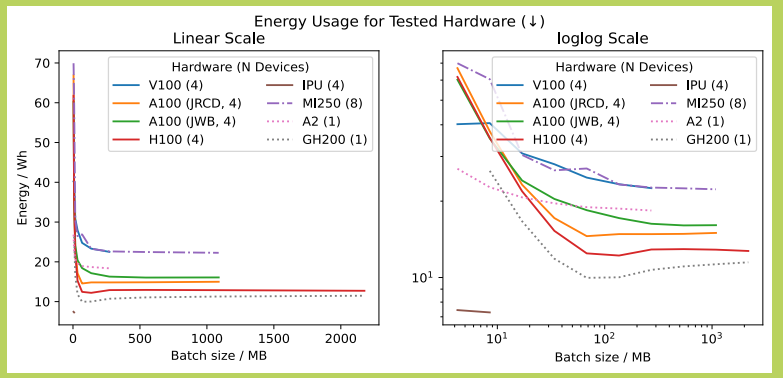
**JUQCS**: 2.6× vs. A100 for 31 Qubits




**MPTRAC**: >2× vs A100




**LLM** benchmark: >2× vs A100




**MAELSTROM** AP1: >6× energy efficiency vs. A100

# FIRST CPU INVESTIGATIONS *(GRACE)*

- Focus mostly on GPU currently
- Some first results on Grace hardware

→ Very competitive performance, especially wrt TDP (but still early)

**DynQCD**: 1.5× vs. EPYC Rome 7742 (2×64 cores)
- Best: Grace-Clang, ACfL
- Slightly worse: GCC
- Investigating FMLA instructions
- *(Auto-Vectorization works well!)*

**NAStJA**:
- 2.3× vs. EPYC Rome 7402 (2×24 cores)
- 5.6× vs Intel Skylake 8168 (2×24 cores)

**JUQCS**: 1.35× vs. EPYC Rome 7402 for 31 Qubits (2×24 cores)

**FLEUR**:
- 1.2× vs. Intel Skylake 8168 (2×24 cores, 400 W TDP)
- 0.8× vs. EPYC Rome 7742 (2×64 cores, 450 W TDP)
- 1.5× vs. Intel SPR-HBM (2×32 cores, 700 W TDP)

MAX DRIVING THE EXASCALE TRANSITION

# JUPITER

## The Arrival of Exascale in Europe

fz-juelich.de/jupiter | #exa_jupiter

EuroHPC Joint Undertaking

Federal Ministry of Education and Research

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

GCS Gauss Centre for Supercomputing

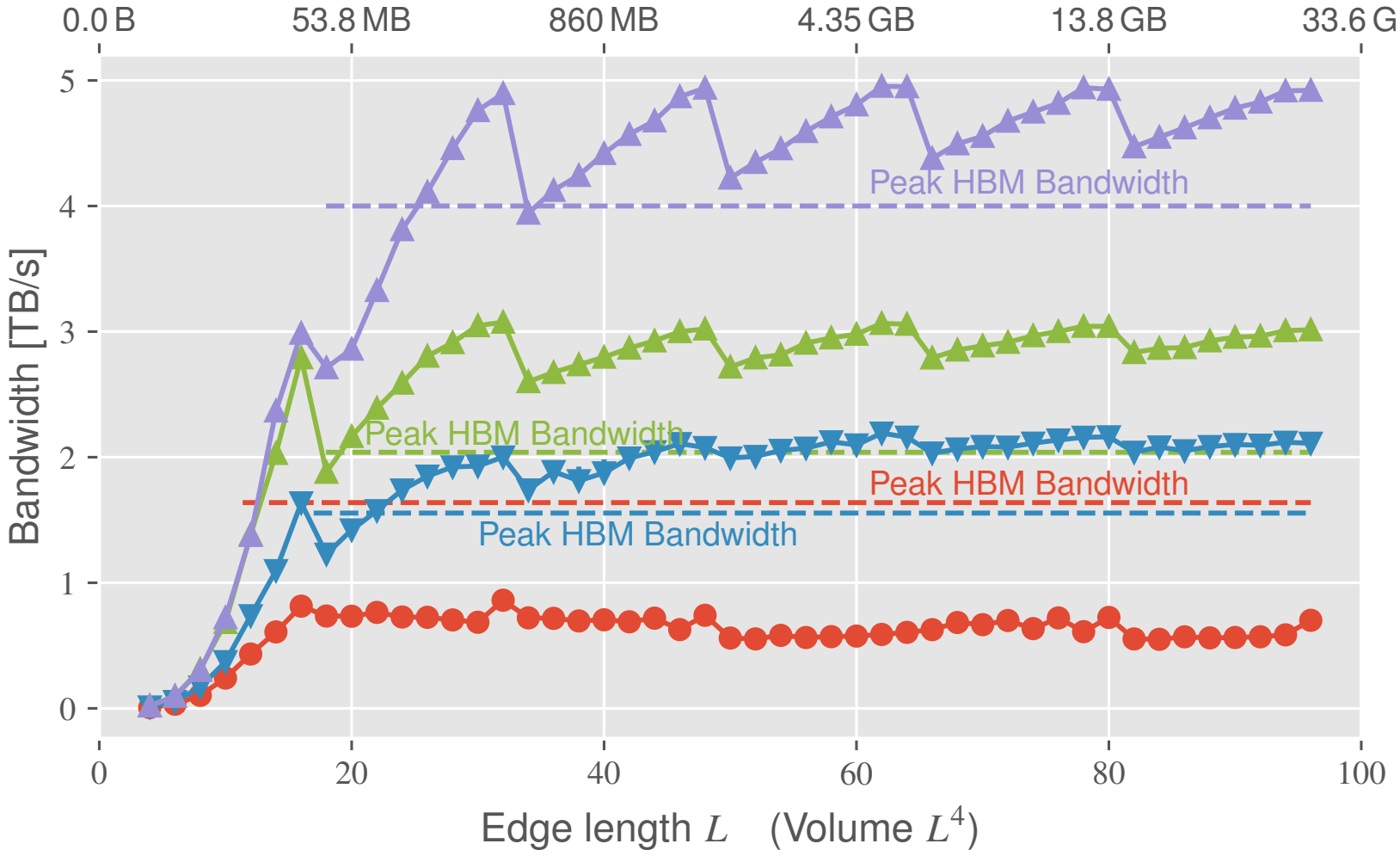
# RESOURCES USED

- Page 2 (JUPITER Overview): Renderings by Eviden
- Page 3 (JUPITER Booster/Cluster): Grace-Hopper picture by NVIDIA; SiPearl Rhea picture by SiPearls
- Page 4 (JEDI/JUREAP): JEDI picture by FZJ-JSC (S. Habbinga/H. Zilken), Harvester image by Robert Wiedeman on Unsplash , GH prototype picture by FZJ-JSC (S. Achilles); Lightsaber icon by Fourup Company on Noun Project
- Page 5 (GPU results): See next slides for details
- Page 6 (CPU results): All results by FZJ-JSC; DynQCD by S. Nassyr, NAStJA by A. Dabah, JUQCS by H. d. Raedt, FLEUR by K. Haghighi Mood

# KOKKOS LQCD BENCHMARK

**By Simon Schlepphorst / FZJ-JSC**

- Simplified Staggered Fermion Dirac Operator
- MI250
- A100
- H100 (PCIe)
- H100 (GH200)

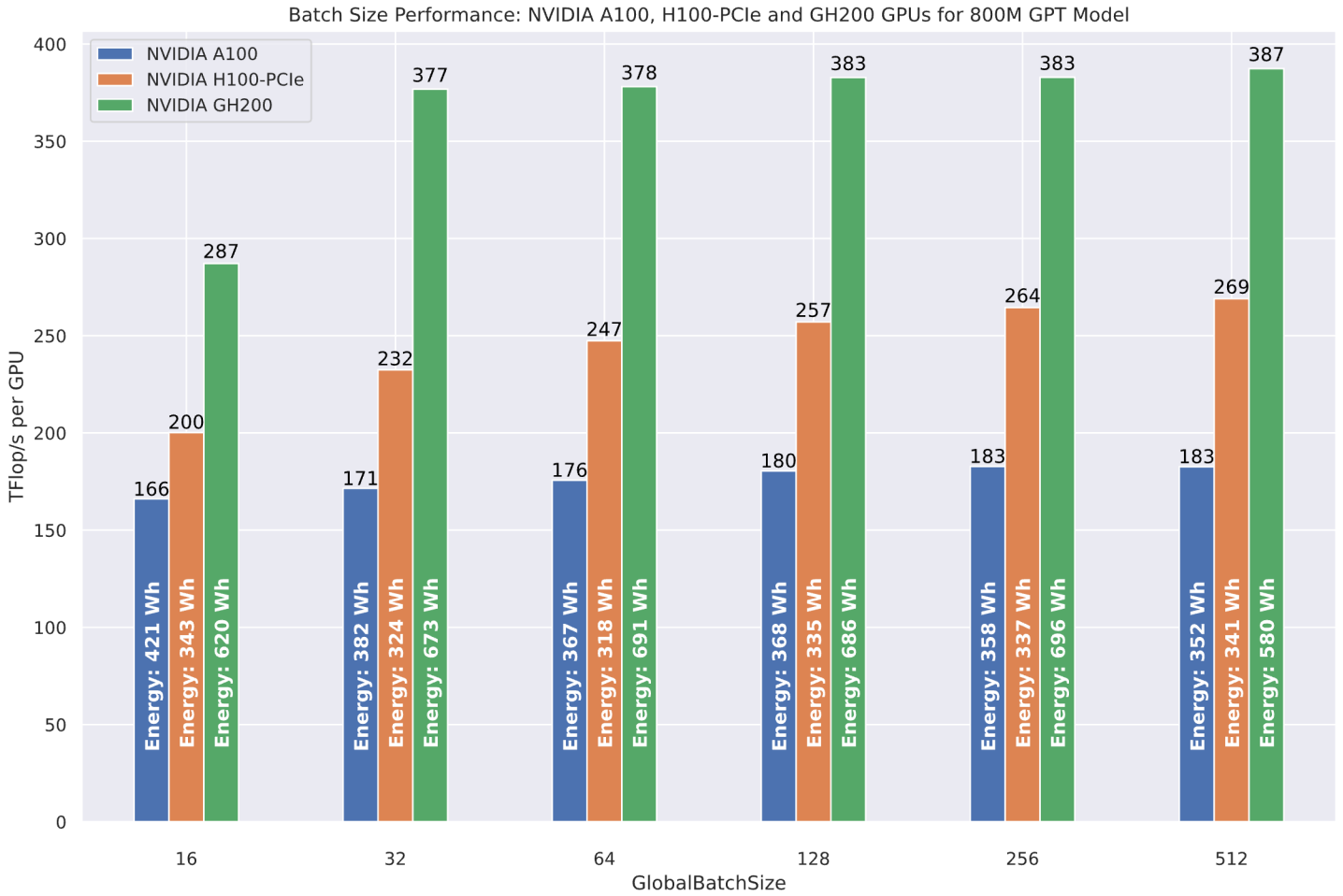



https://gitlab.jsc.fz-juelich.de/sdl-nqft/kokkos-lqcd-benchmarks

# MEGATRON-LM

**By Chelsea John / FZJ-JSC**

- OpenGPT-X: BMWK project for LLM with EU languages, Open
- 1 node benchmark: 800M GPT Model
- Increasing Batchsizes
- Excellent performance on H100, especially GH200 variant (HBM3, TDP)

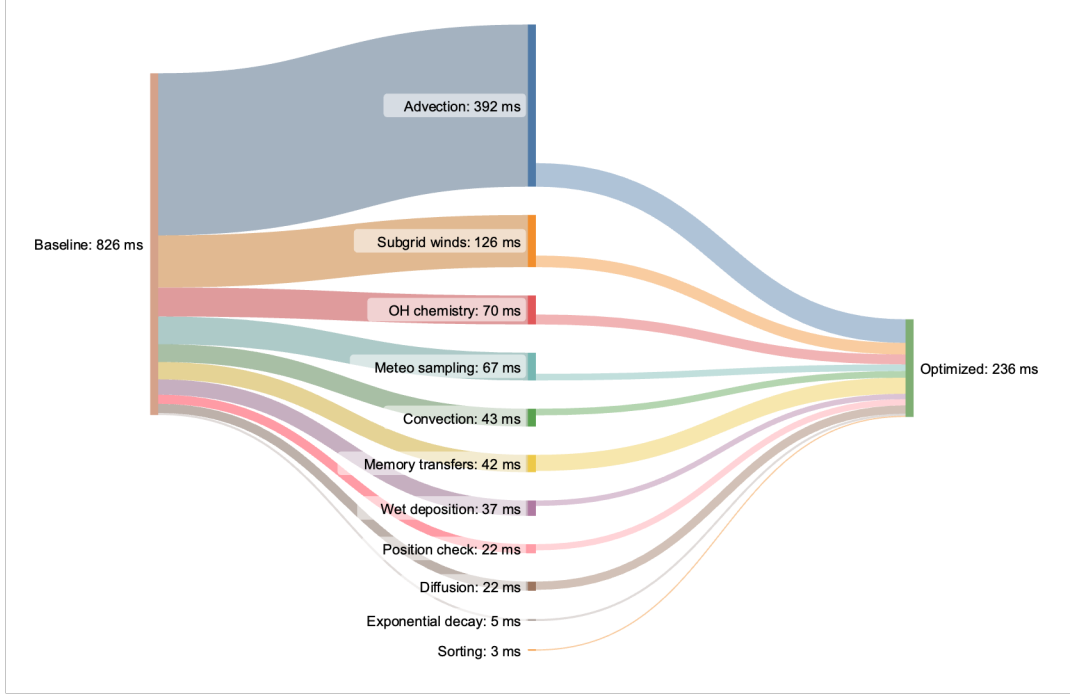


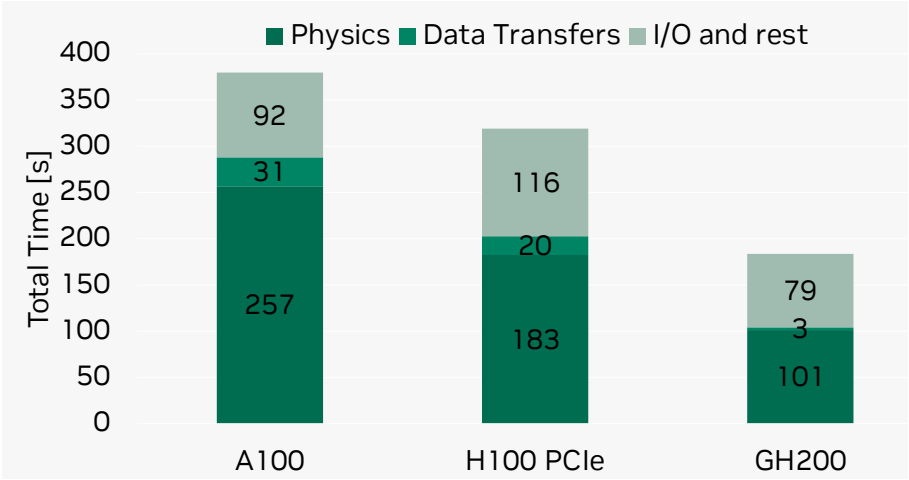

https://github.com/OpenGPTX/Megatron-LM/

# MPTRAC

## By Lars Hoffmann / FZJ-JSC

- Lagrangian particle dispersion model: atmospheric transport processes (troposphere/stratosphere)
  → volcanic emissions
- Continuously optimized for GPUs
  Recently: Significant speedup on A100
- First test on GH200



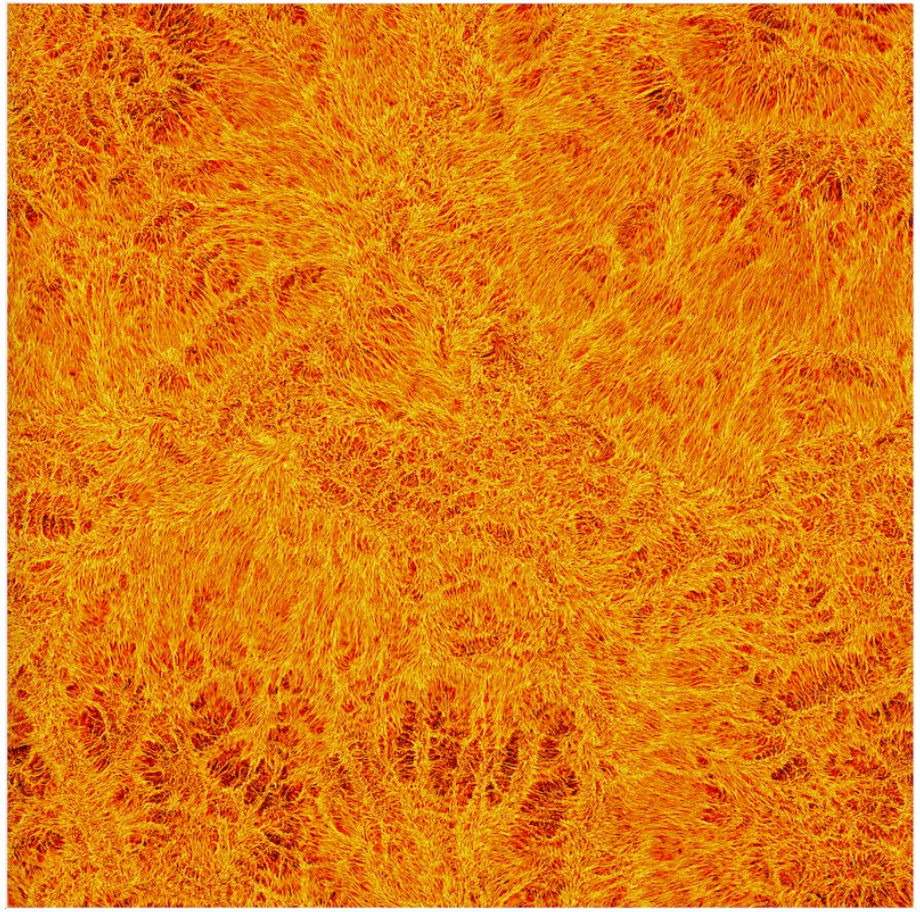




*See also GTC talk by Mathias Wagner*

# OTHER GH200 RESULTS

## JUWELS Booster A100 vs. GH200

- ICON:
  - R2B4 Benchmark: 563 s vs. 343 s → 1.62×
  - By Manoel Römmer / FZJ-JSC
- nekRS
  - Rayleigh–Bénard Convection: 2.16×
  - By Mathis Bode / FZJ-JSC
- Arbor
  - Busyring Benchmark: 330 s vs 167 s → 1.97×
  - By Thorsten Hater / FZJ-JSC

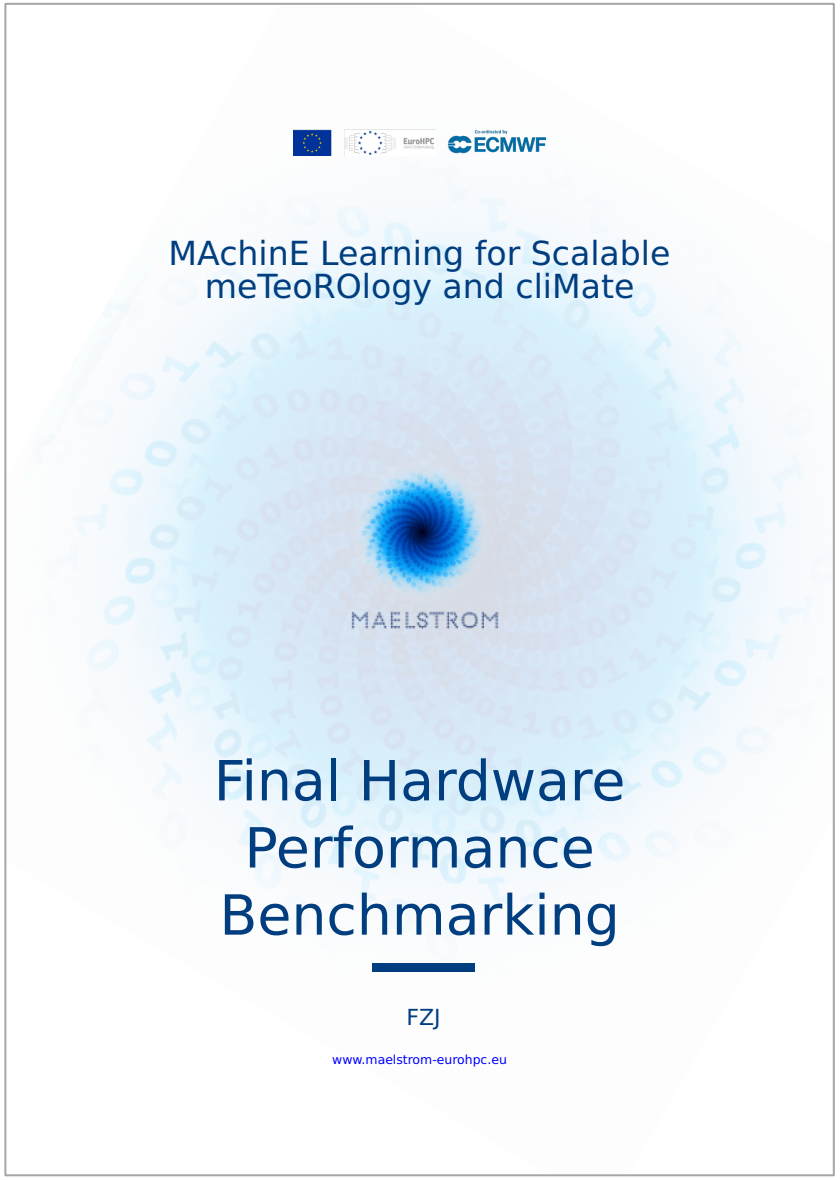


RBC Visualization (M. Bode)

# MAELSTROM TESTS

- MEALSTROM: EuroHPC-JU-funded project, ended in March 2024
- Enablement/optimization of 6 weather & climate applications using ML
- D3.7 (Feb 2024): Benchmarking of final applications on diverse hardware

→ www.maelstrom-eurohpc.eu



# MAELSTROM TESTS

## AP1: MetNor