
42ND INTERNATIONAL SYMPOSIUM
ON LATTICE FIELD THEORY

Tata Institute of Fundamental Research
Mumbai, India

02-08 November, 2025

# MACHINES AND EXASCALE COMPUTING IN LQCD

## A review of current hardware trends in HPC

LATTICE 2025, MUMBAI – 06/11/2025 – S. KRIEG

JÜLICH
Forschungszentrum
Shaping Change

Member of the Helmholtz Association

# A BIT OF HISTORY

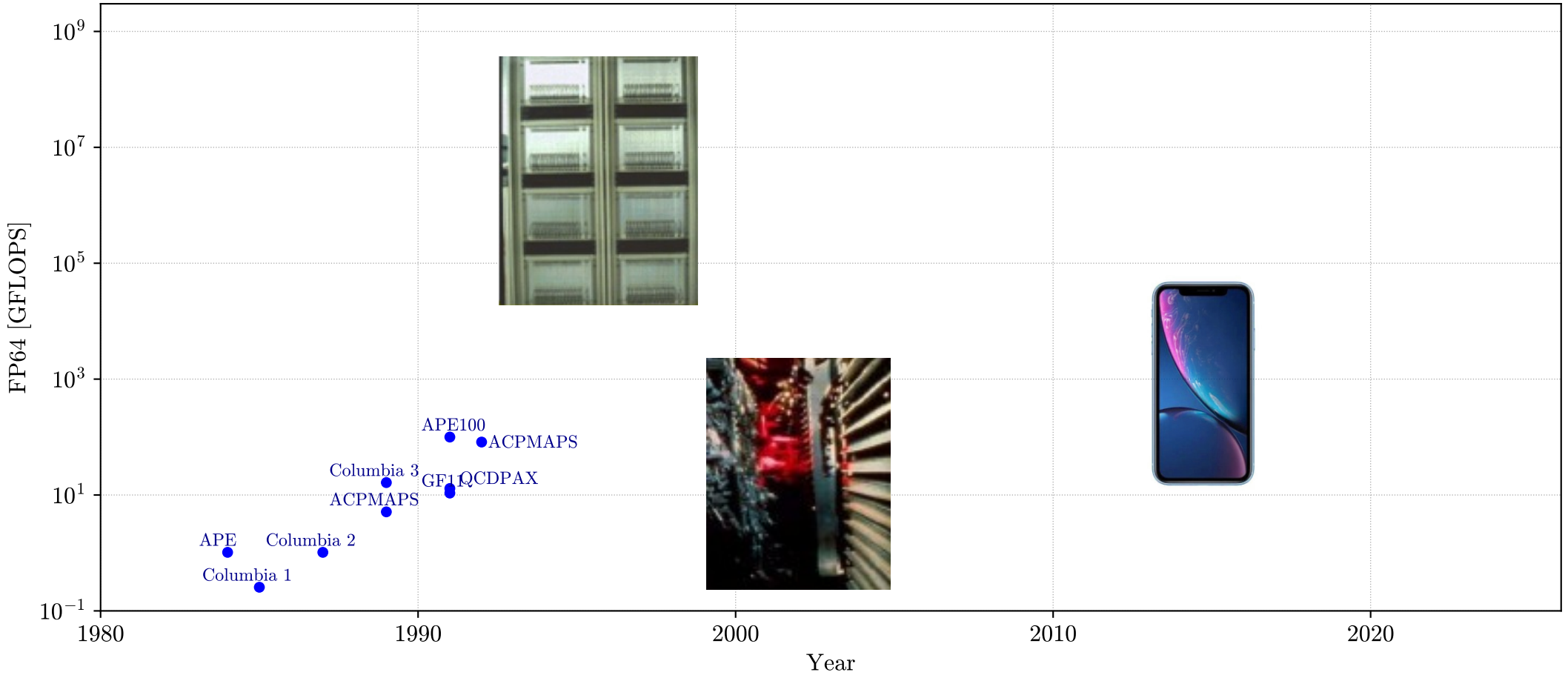
# SIMULATIONS OF LATTICE QCD

## A history of counting cycles

- Immense costs of LQCD calculations drove the development of HPC systems
- Lattice QCD researchers joined hardware vendors to develop hardware suitable for "the cause"
- Initially, only **quenched** calculations were possible
  → fermion determinant ignored → unquantifiable systematics
- Still, a complete quenched spectrum calculation took **20 years** and the fastest supercomputers of their time
- Status 1983: $m_N = 1000(150)$ MeV, $m_\rho = 800(150)$ MeV, $m_\Delta = 1300(150)$ MeV (incl. $SU(2)$, discrete approx. etc.)

# SIMULATIONS OF LATTICE QCD

## A history of counting cycles

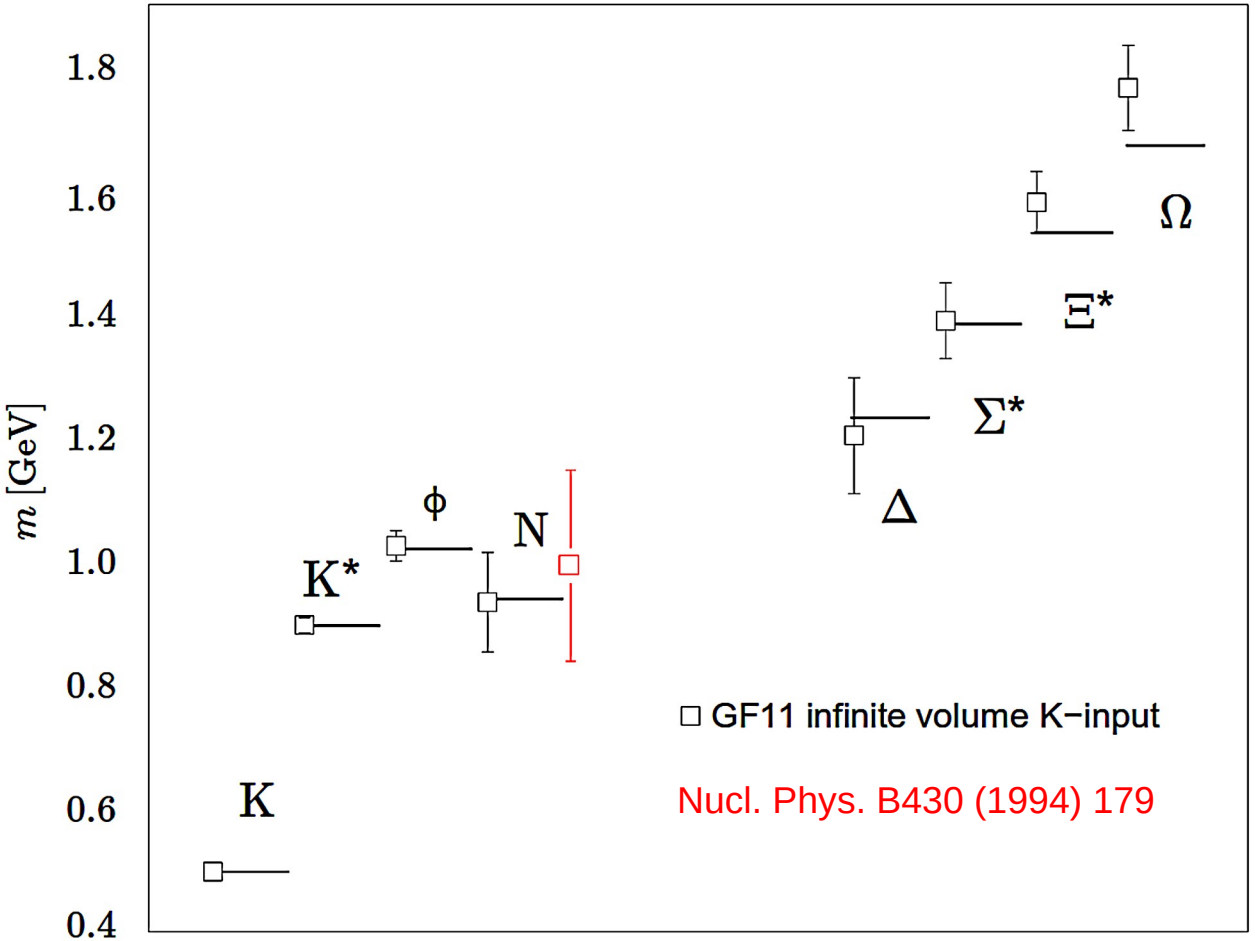





JÜLICH Forschungszentrum

# SIMULATIONS OF LATTICE QCD

## 10 years of progress: early (quenched) calculations



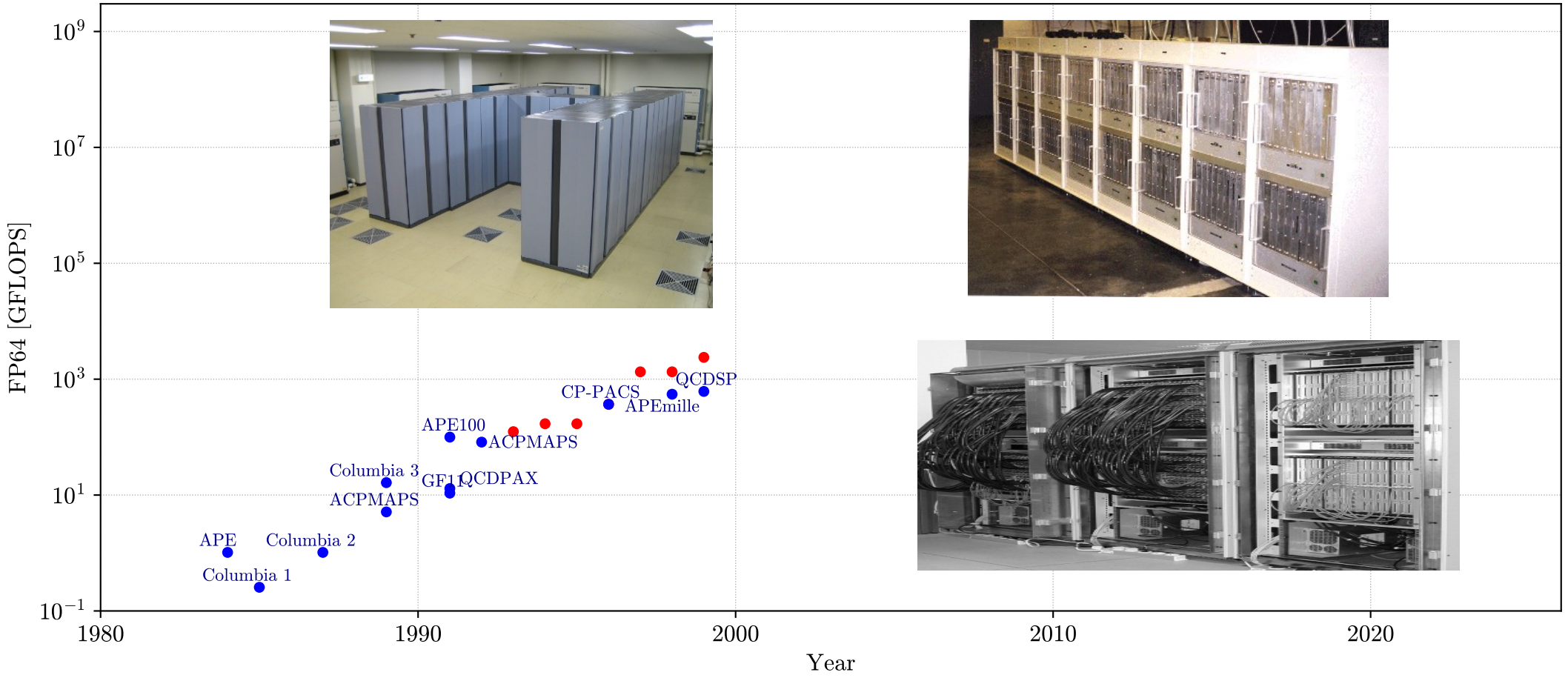

Adapted from
hep-lat/9904003



JÜLICH Forschungszentrum

# SIMULATIONS OF LATTICE QCD

## A history of counting cycles







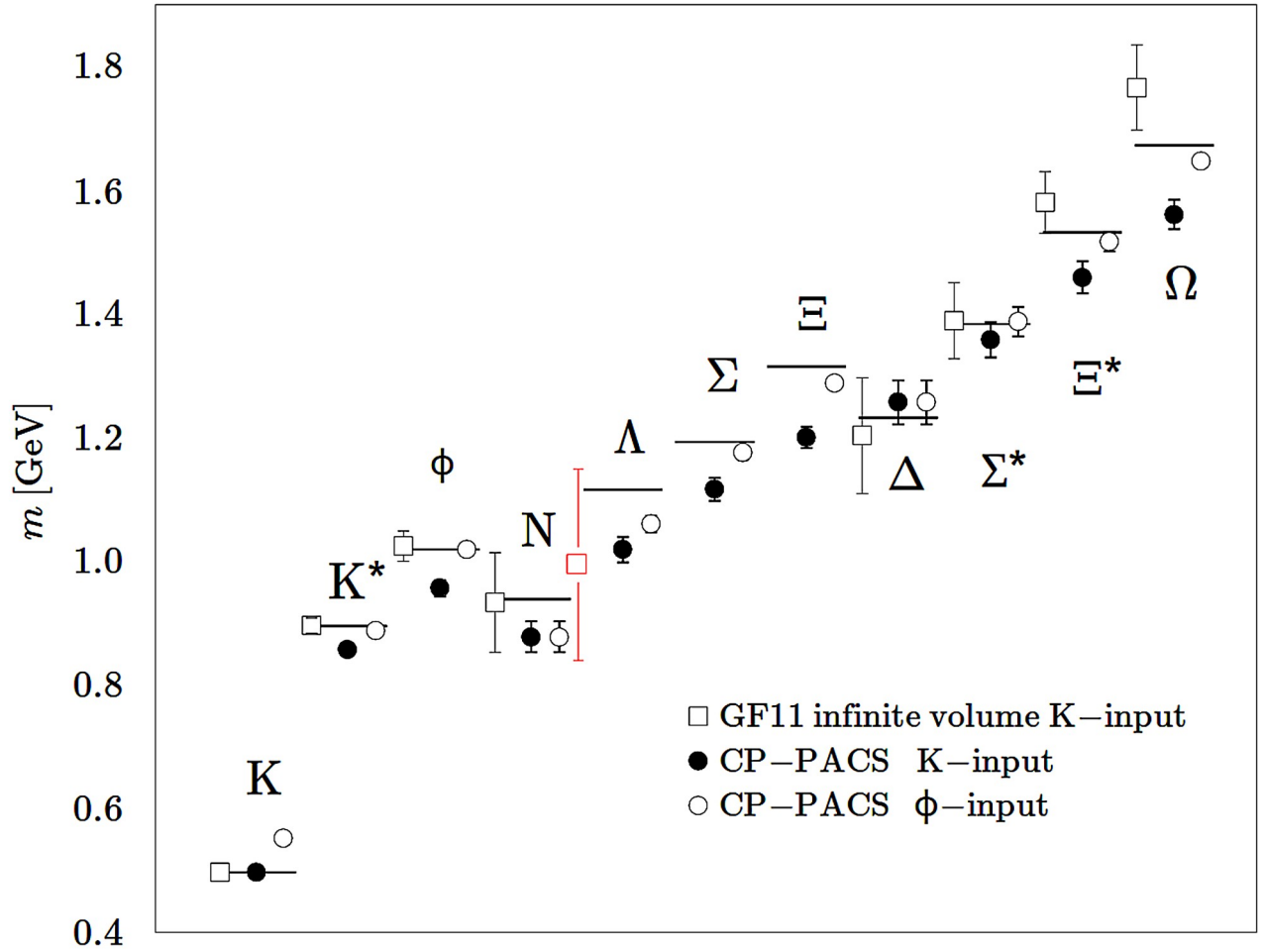





JÜLICH Forschungszentrum

# SIMULATIONS OF LATTICE QCD

## 20 years of progress: end of the quenched era




- Continuum limit
- $m_\pi = 300$ MeV
- Systematics are $\approx 10\%$

Phys. Rev. Lett. 84 (2000) 238

Adapted from hep-lat/9904003



JÜLICH Forschungszentrum

# SIMULATIONS OF LATTICE QCD

## The Berlin Wall

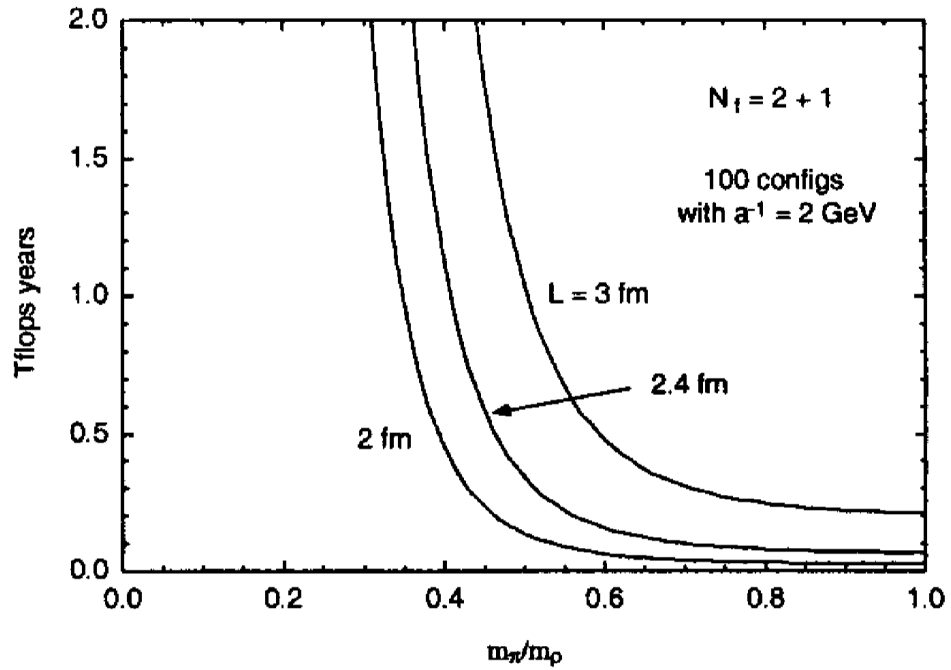



Figure 3. Cost of $N_f = 2+1$ QCD at $a^{-1} = 2$ GeV for 100 independent configurations.

Nucl. Phys. B (PS) 106–107 (2002) 195

Algorithmic changes were needed for physical simulations. Hardware performance was not enough

→ Optimized integrators
→ Multigrid solvers
→ Hasenbusch mass preconditioning
→ Multiple integration step-sizes



JÜLICH Forschungszentrum

# SIMULATIONS OF LATTICE QCD

## A history of counting cycles

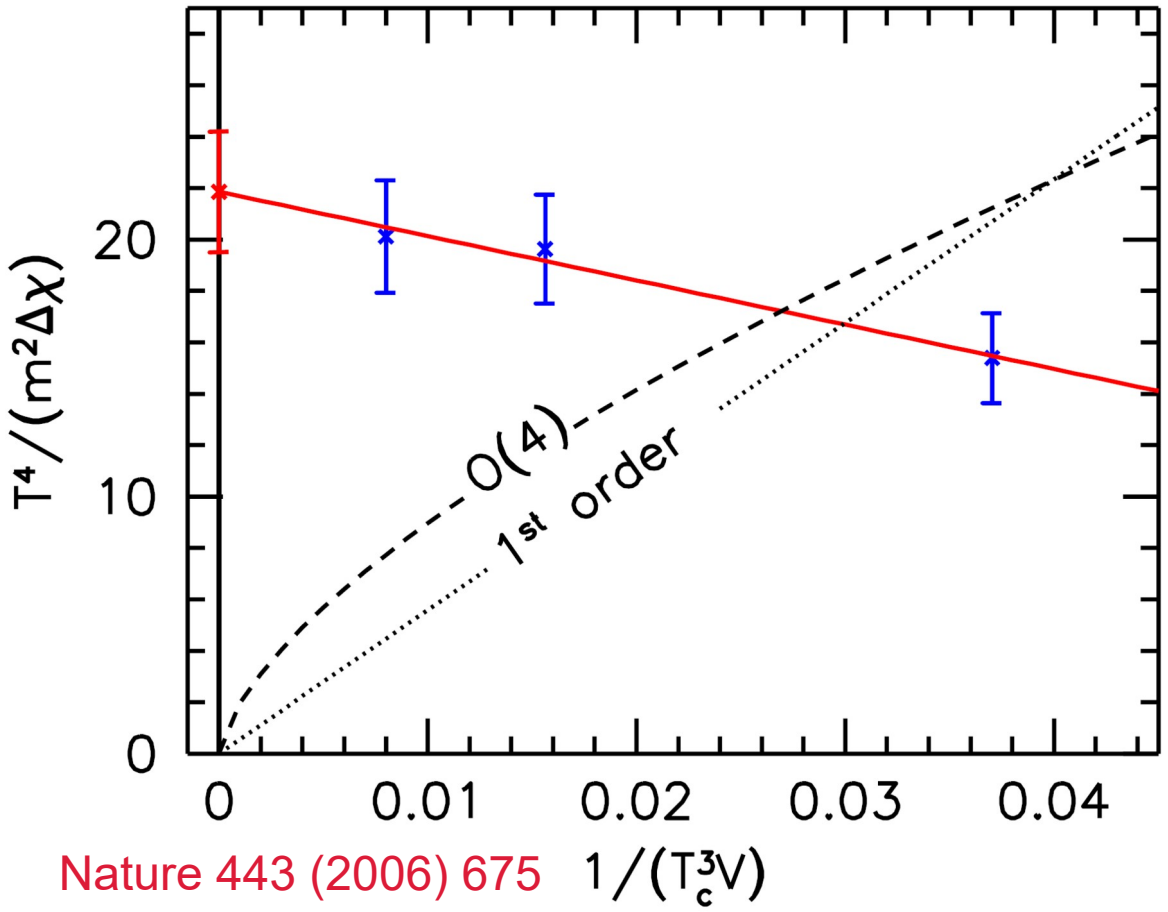




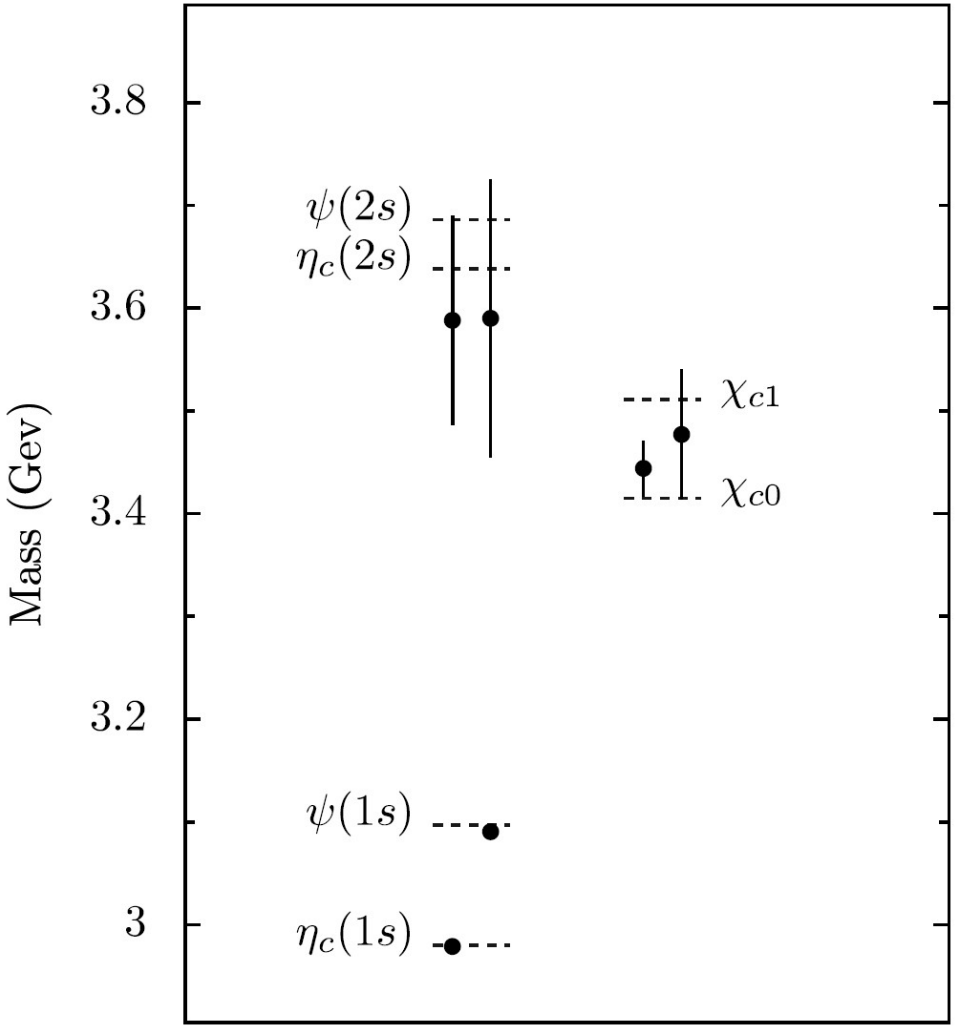

JÜLICH Forschungszentrum

# SIMULATIONS OF LATTICE QCD

## Era of unquenched simulations




Nature 443 (2006) 675




Phys. Rev. D 75 (2007) 054502



JÜLICH Forschungszentrum

# SIMULATIONS OF LATTICE QCD

## A history of counting cycles



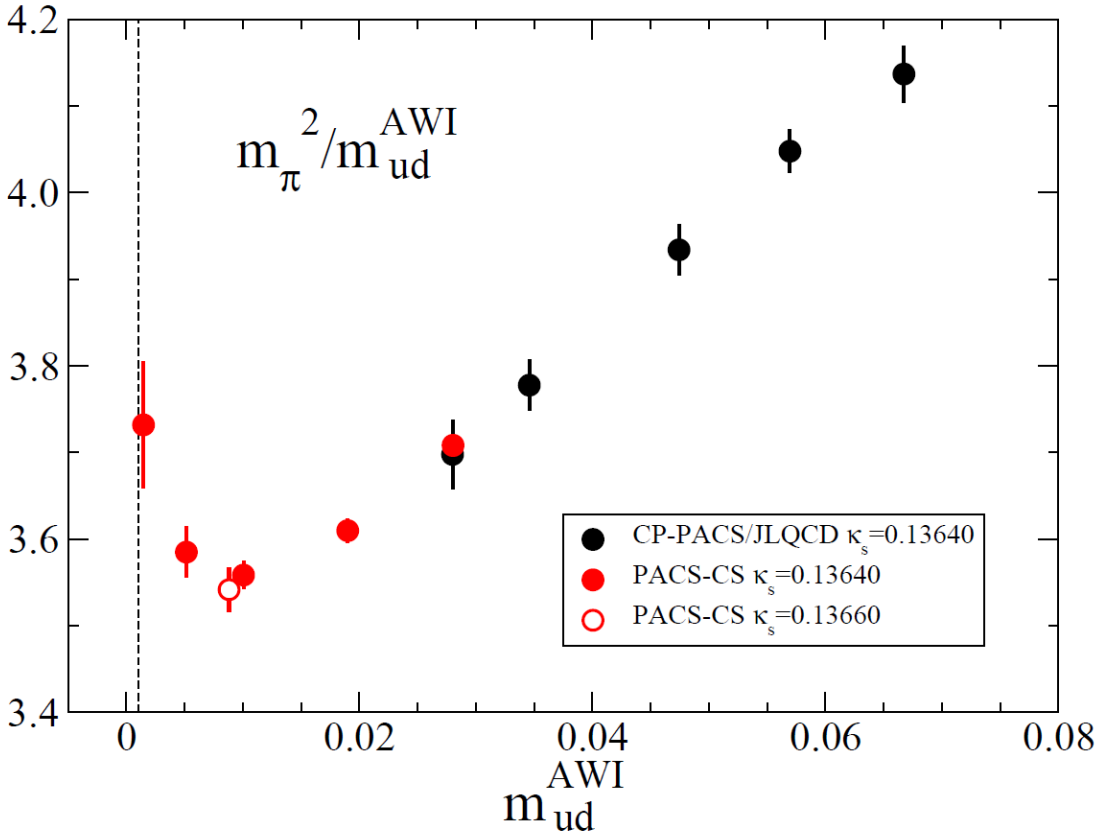




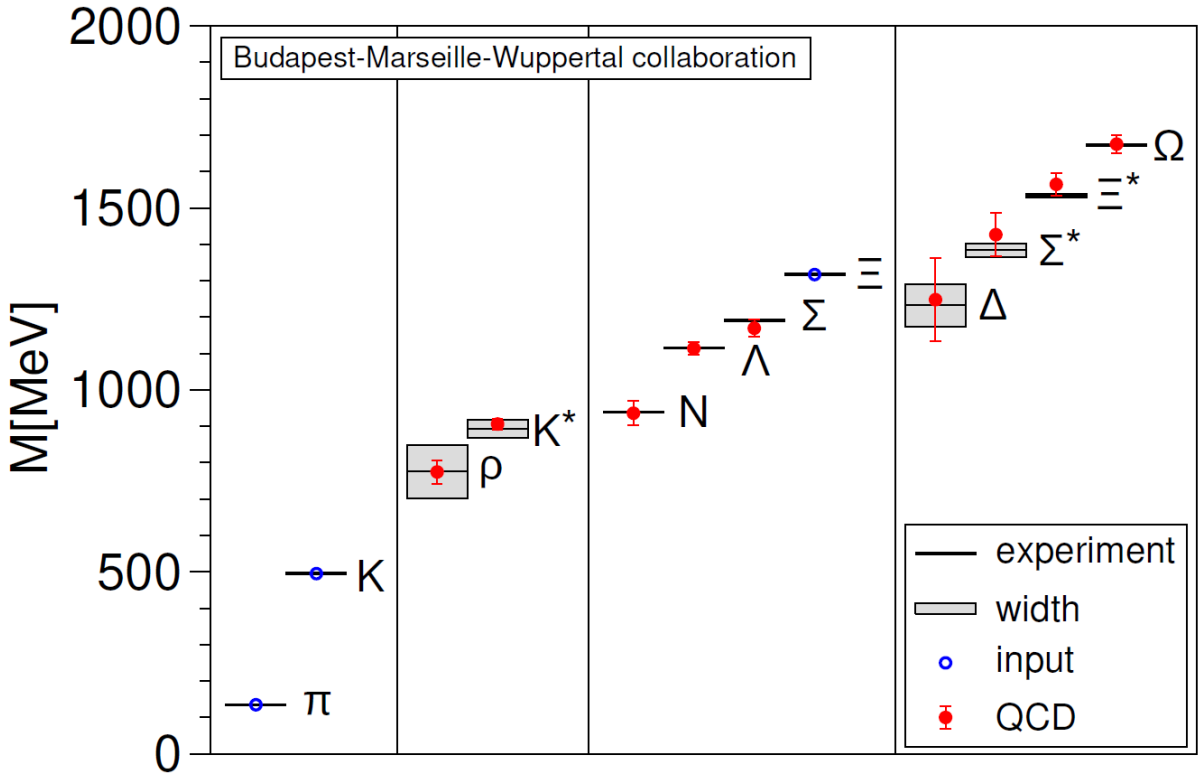
JÜLICH Forschungszentrum

# SIMULATIONS OF LATTICE QCD

## 30 years of progress: dynamical simulations, physical point, and spectrum




Phys. Rev. D 79 (2009) 034503




Science 322 (2008) 1224



JÜLICH Forschungszentrum

# SIMULATIONS OF LATTICE QCD

## A history of counting cycles



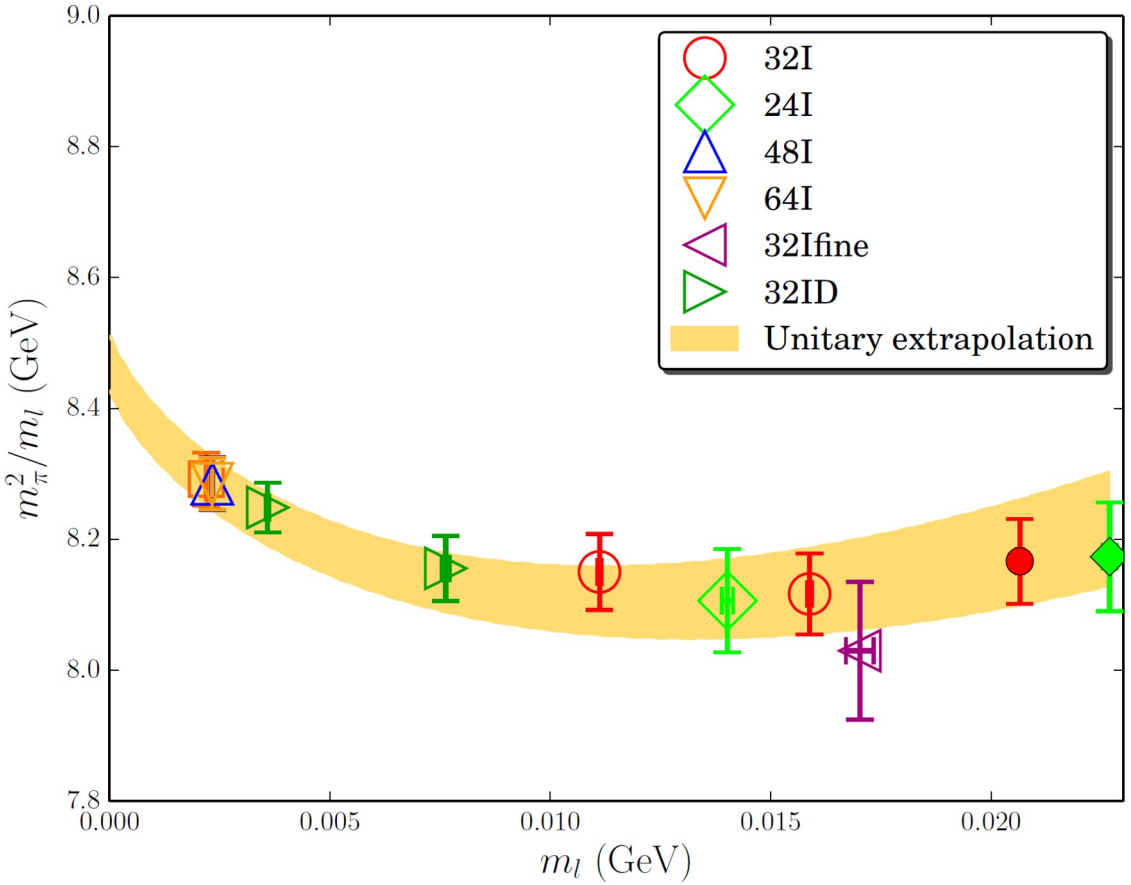



JÜLICH Forschungszentrum

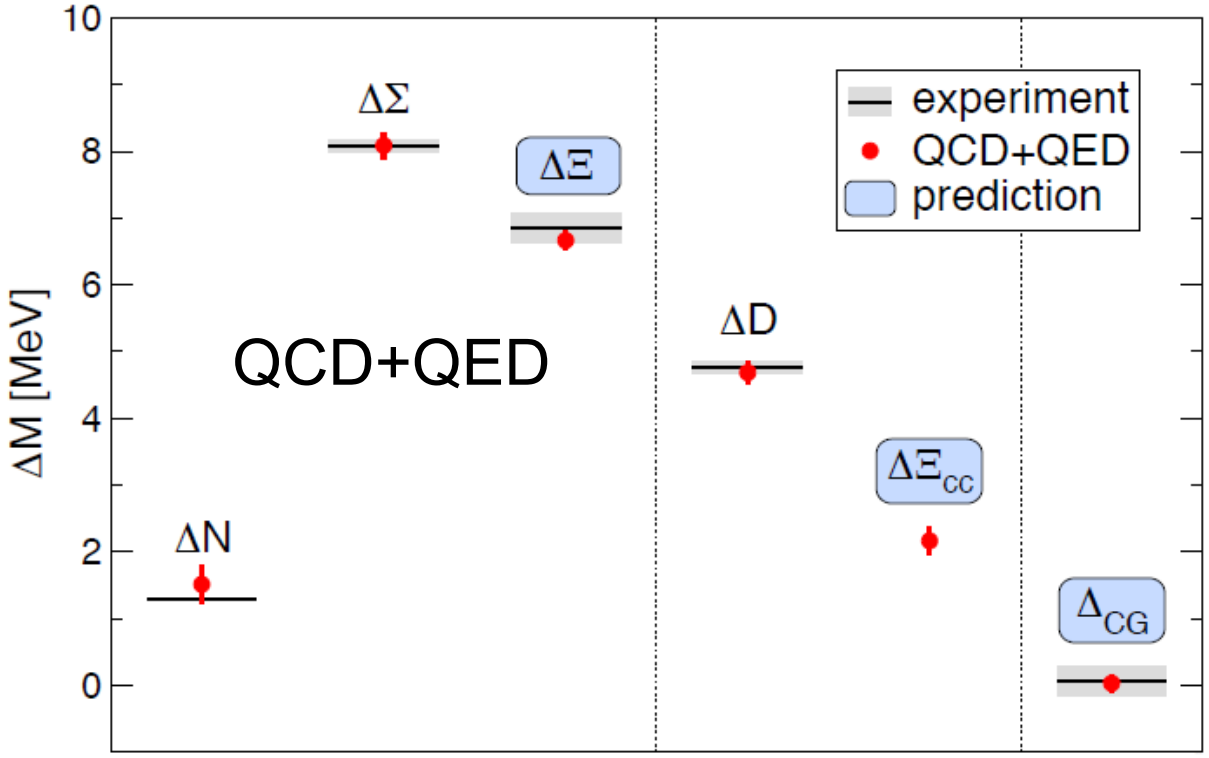
# SIMULATIONS OF LATTICE QCD

## New milestone: physical point with chiral fermions, dynamical QCD+QED

Domain wall QCD with physical quark masses




Phys. Rev. D 93 (2016) 074505



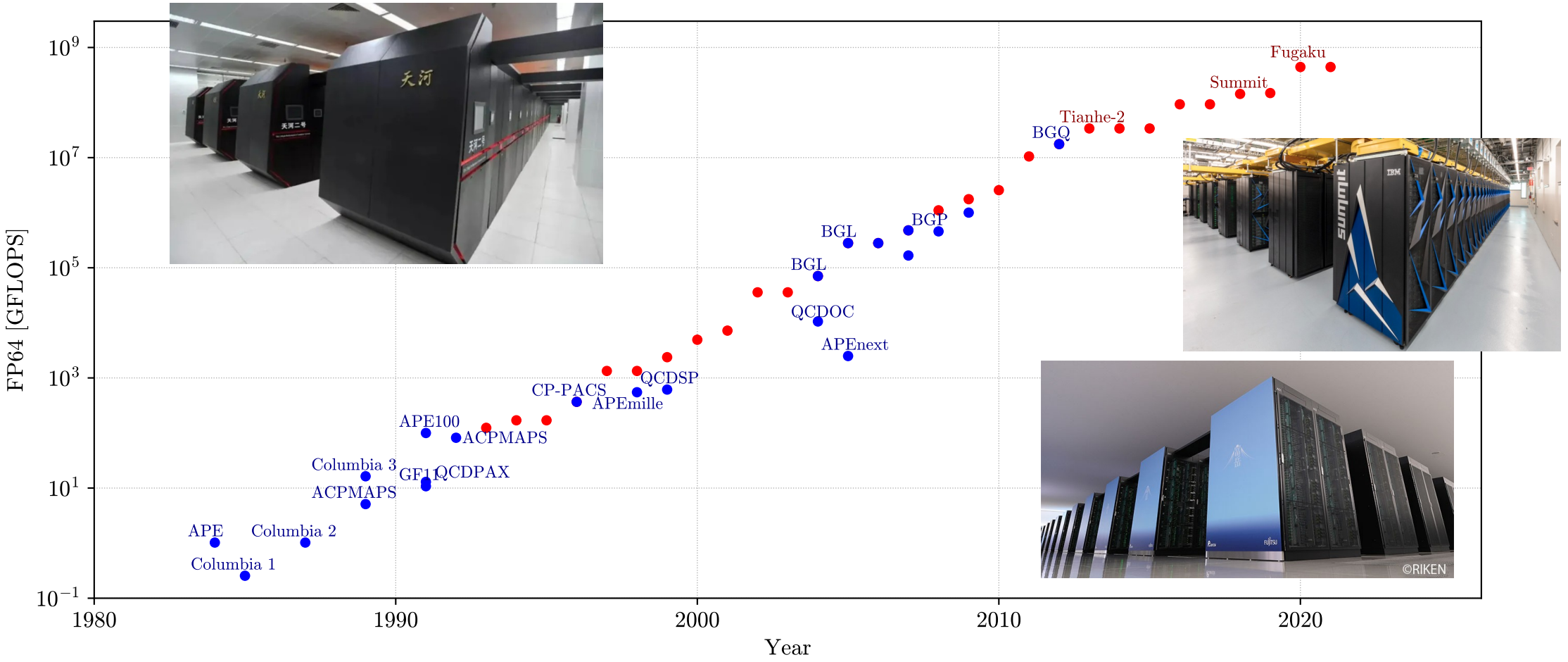

Science 347 (2015) 1452



JÜLICH Forschungszentrum

# SIMULATIONS OF LATTICE QCD

## The end of specialized hardware and rise of GPUs

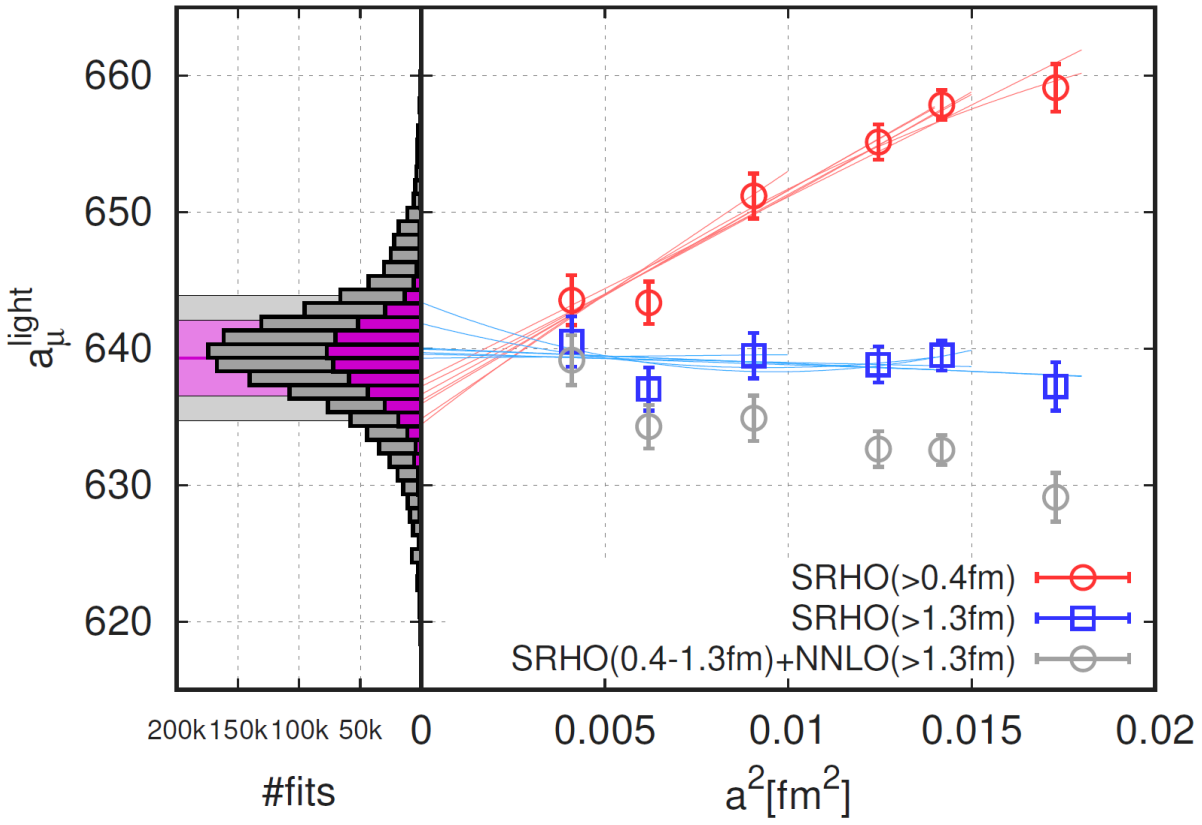




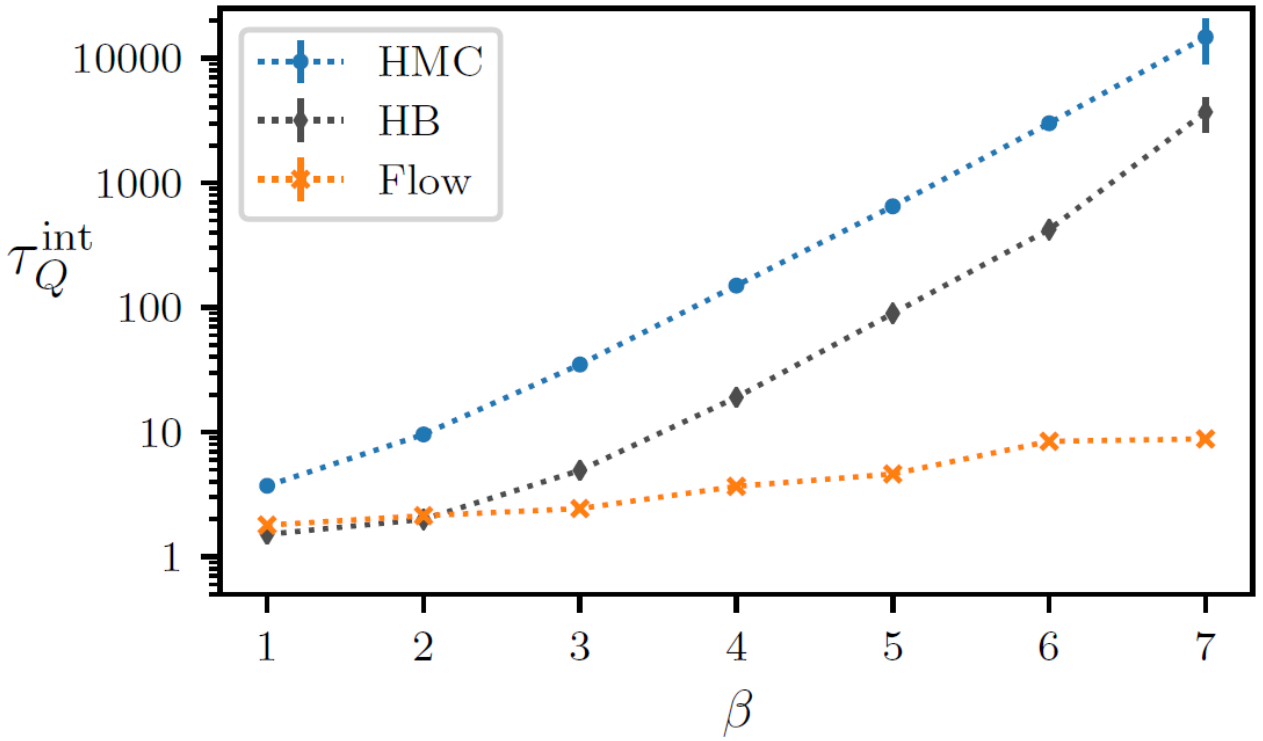

JÜLICH Forschungszentrum

# SIMULATIONS OF LATTICE QCD

## Precision physics and the arrival of AI




Nature 593 (2021) 51



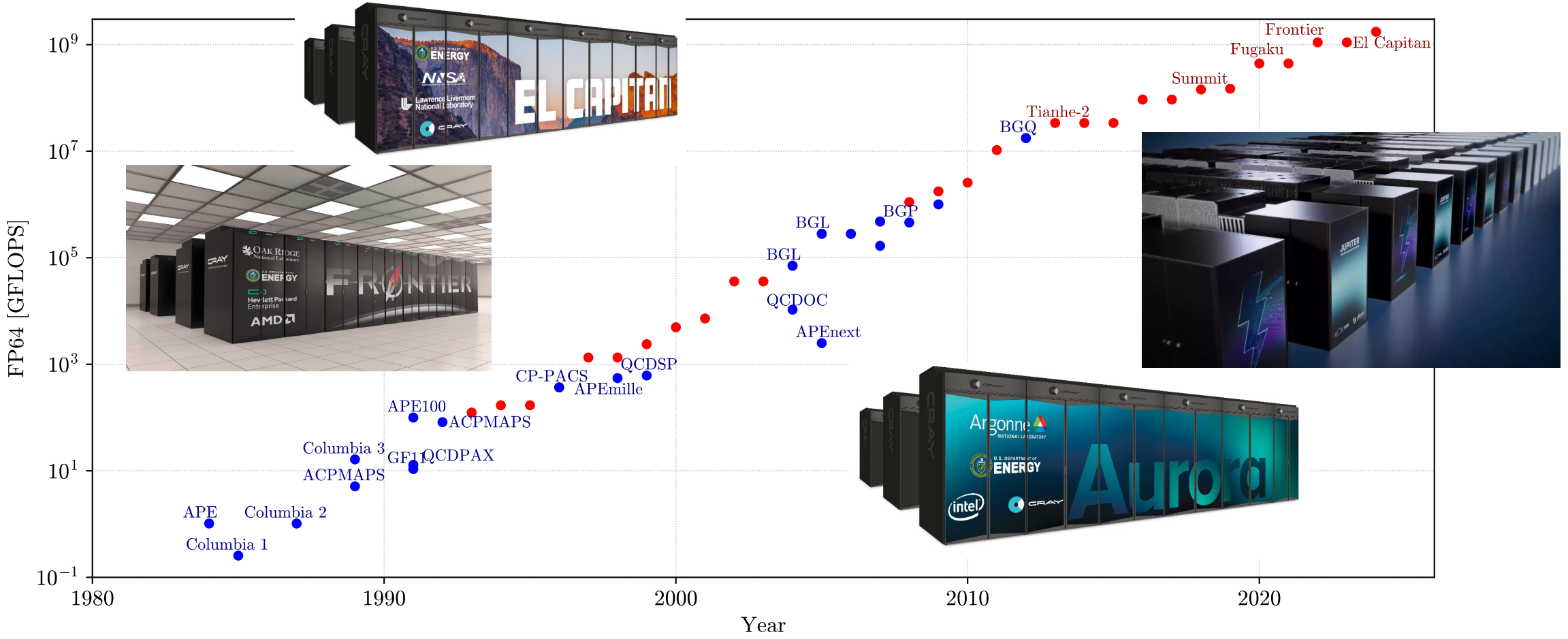

Phys. Rev. Lett. 125 (2020) 121601



JÜLICH Forschungszentrum

# SIMULATIONS OF LATTICE QCD

## The present: Exascale computing



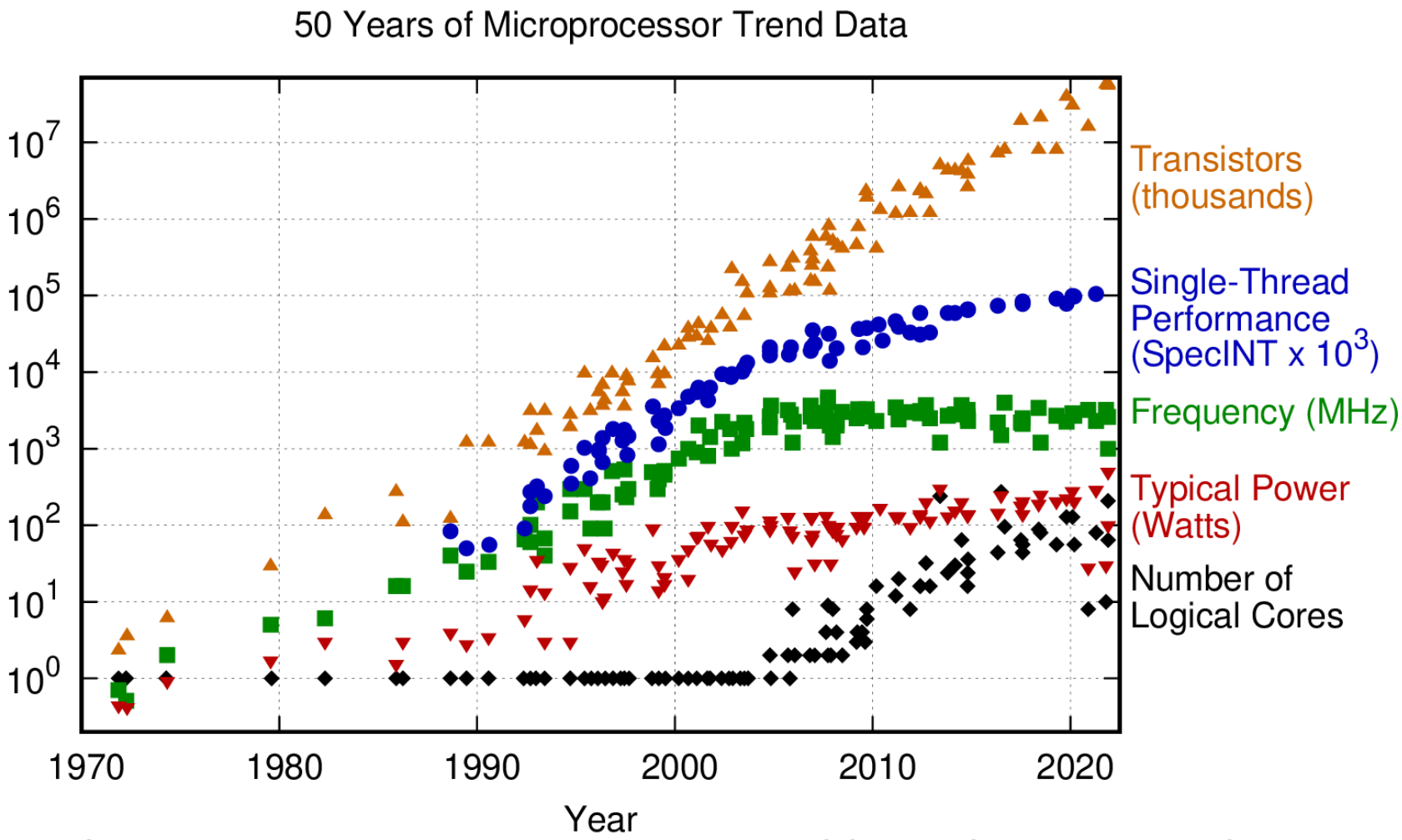



JÜLICH Forschungszentrum

# THE PRESENT

# MOORE'S LAW...

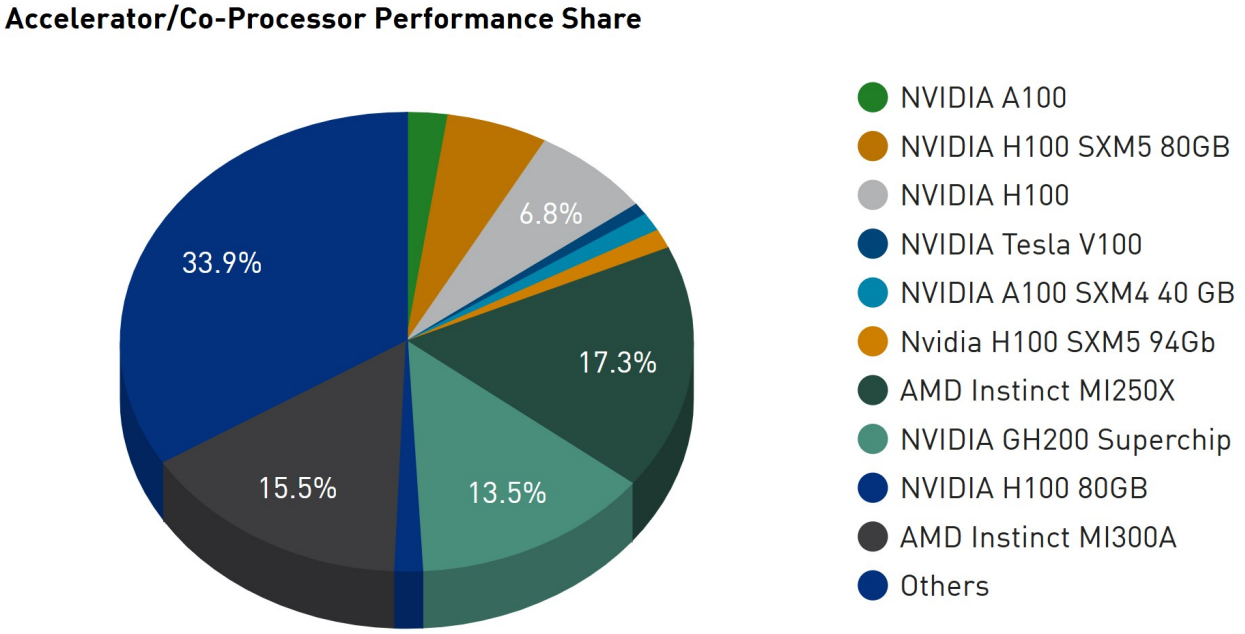
https://github.com/karlrupp/microprocessor-trend-data






JÜLICH Forschungszentrum

# THE LIST (JUNE ‘25)

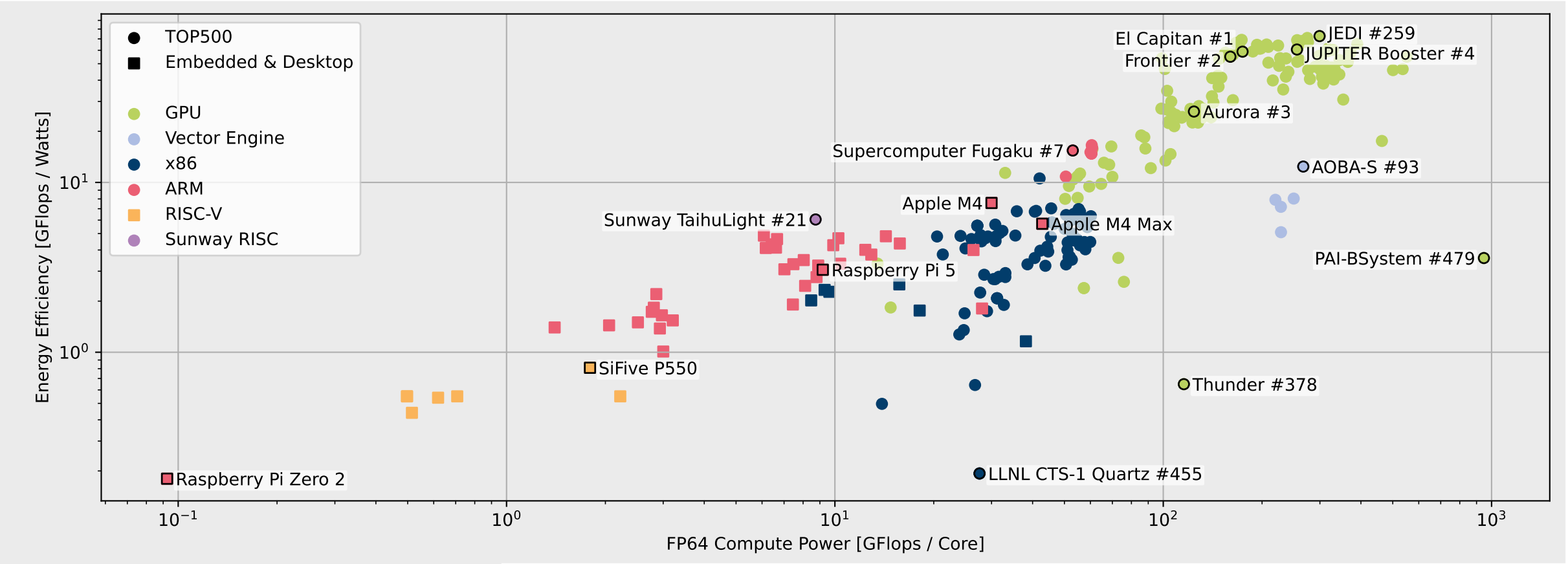






JÜLICH Forschungszentrum

# ENERGY EFFICIENCY OF VARIOUS ARCHITECTURES

## High-Performance Linpack Benchmark

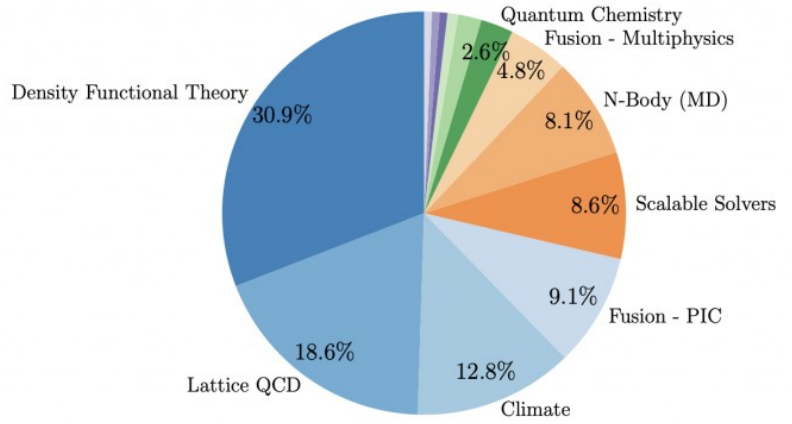



**Sources: top500.org (2025-06) | github.com/geerlingguy/top500-benchmark**

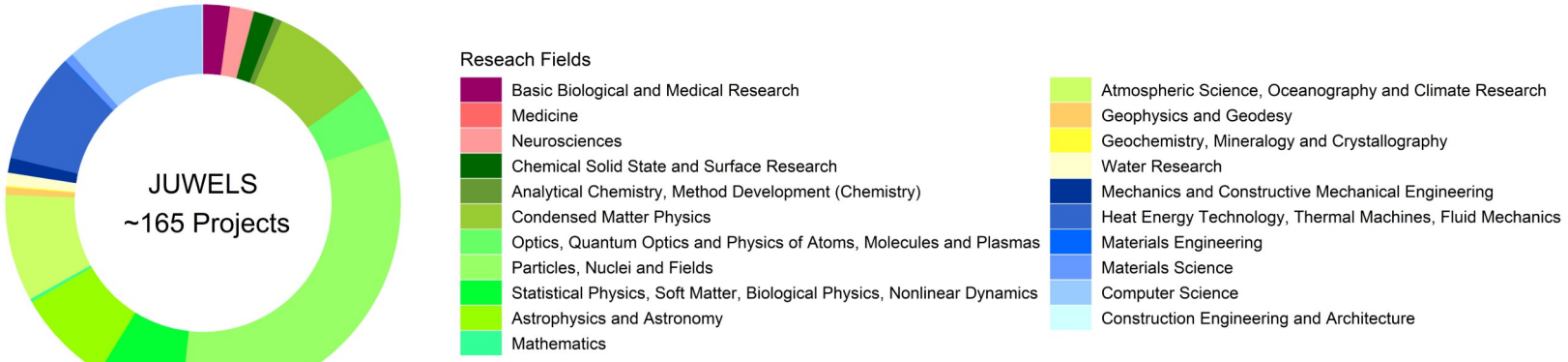
# SIMULATIONS OF LATTICE QCD

## From custom hardware to “regular” users

- LQCD runs on most HPC centers in the world
- Optimized codes for different classes of target hardware to improve performance
- Consumer of 10+% of public supercomputer cycles




[Perlmutter@NERSC 2025]




[JUWELS@JSC 2022]

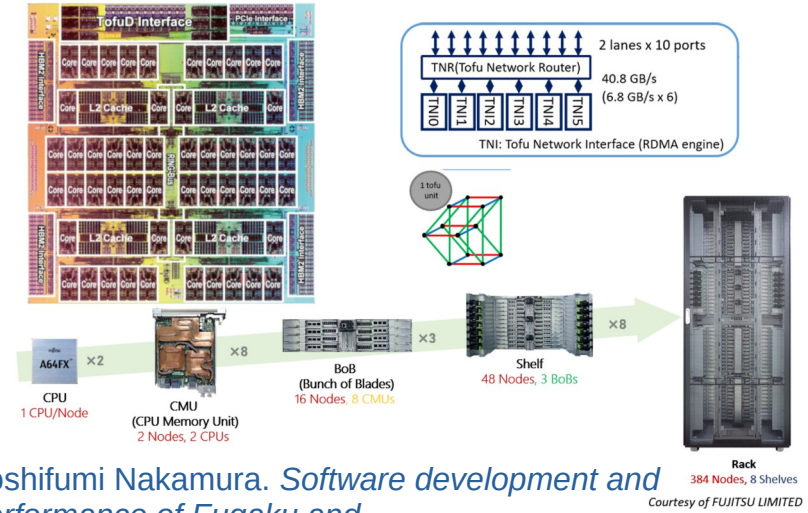


JÜLICH Forschungszentrum

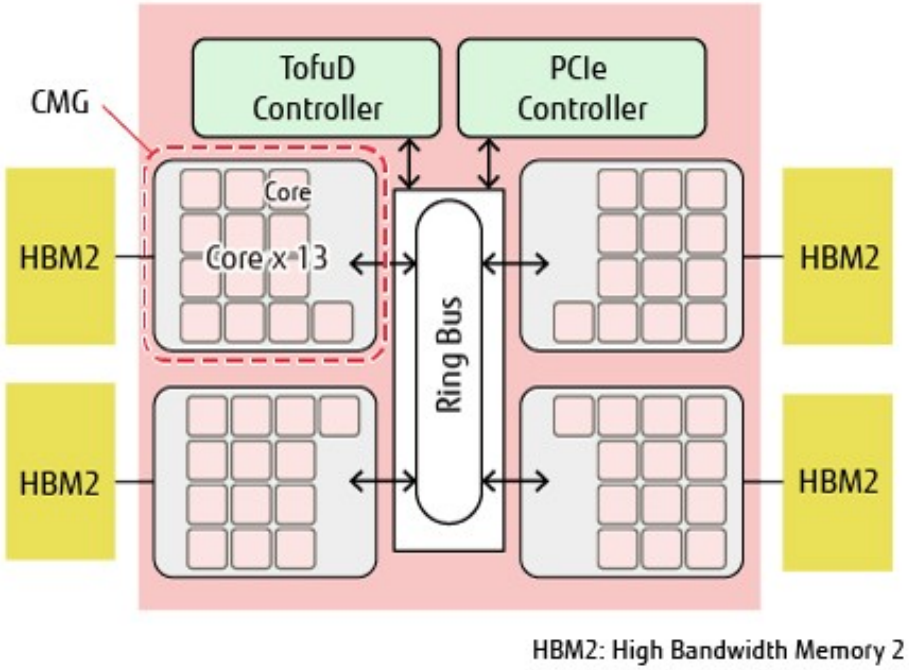
# PRE-EXASCALE SYSTEMS

# ARM A64FX

## Fugaku @ RIKEN - #7




- 158,976 nodes
  - 1x Fujitsu ARM A64FX
  - HBM2 32 GiB, 1024 GB/s
  - Tofu Interconnect D (28 Gbps x 2 lane x 10 port)
- 7,630,848 total cores
- 442.0 PFlop/s on HPL benchmark (#7 on Top500)
- 16.0 PFlop/s on HPCG benchmark (#2 on Top500)



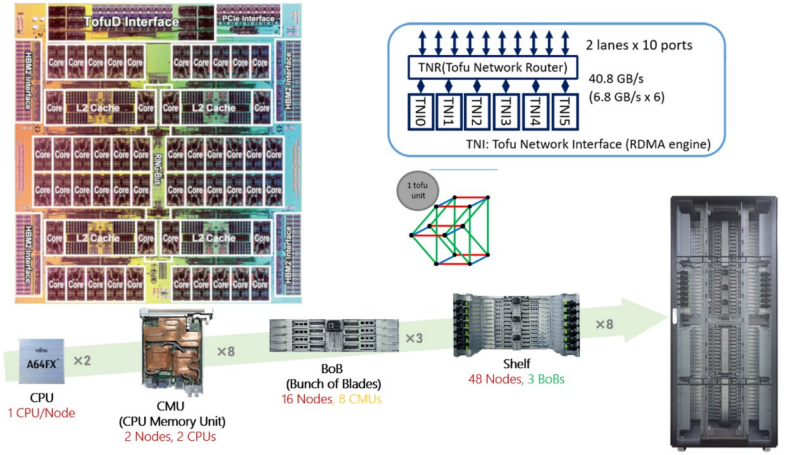

Yoshifumi Nakamura. *Software development and performance of Fugaku and ARM architectures.* PoS LATTICE2021 023 2022



JÜLICH Forschungszentrum

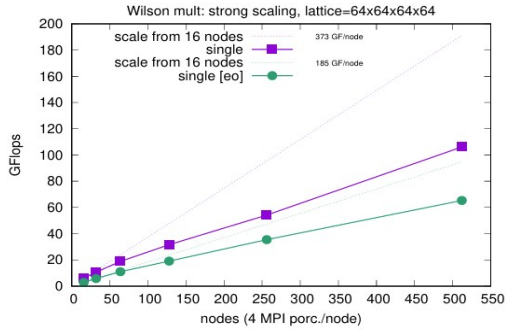
# ARM A64FX

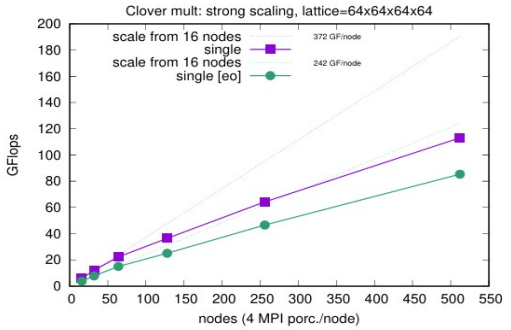
## Fugaku @ RIKEN - #7

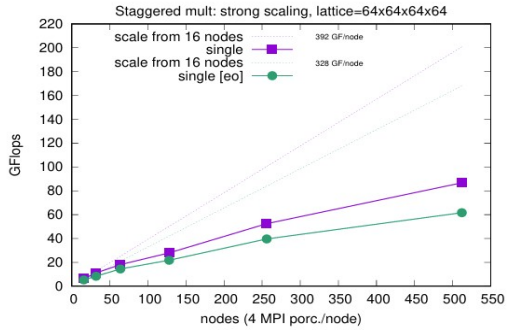


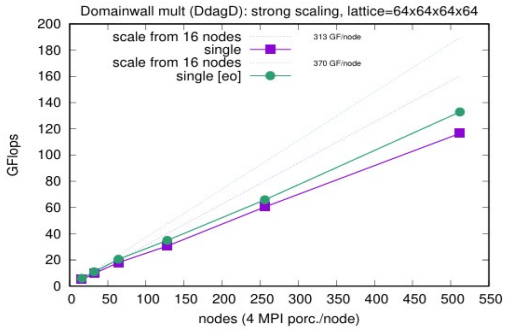

Yoshifumi Nakamura. *Software development and performance of Fugaku and ARM architectures.* PoS LATTICE2021 023 2022

- 158,976 nodes
  - 1x Fujitsu ARM A64FX
  - HBM2 32 GiB, 1024 GB/s
  - Tofu Interconnect D (28 Gbps x 2 lane x 10 port)
- 7,630,848 total cores
- 442.0 PFlop/s on HPL benchmark (#7 on Top500)
- 16.0 PFlop/s on HPCG benchmark (#2 on Top500)




Strong scaling of Wilson Clover, Staggered and Domain Wall operators on Fugaku with Bridge++2.0.

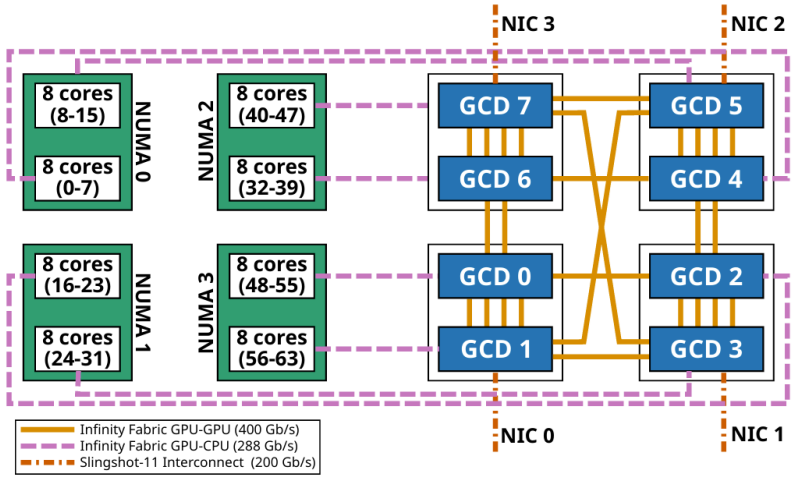
Aoyama et. al. *Bridge++2.0: Benchmark Results on supercomputer Fugaku. PoS LATTICE2022 284 2023*



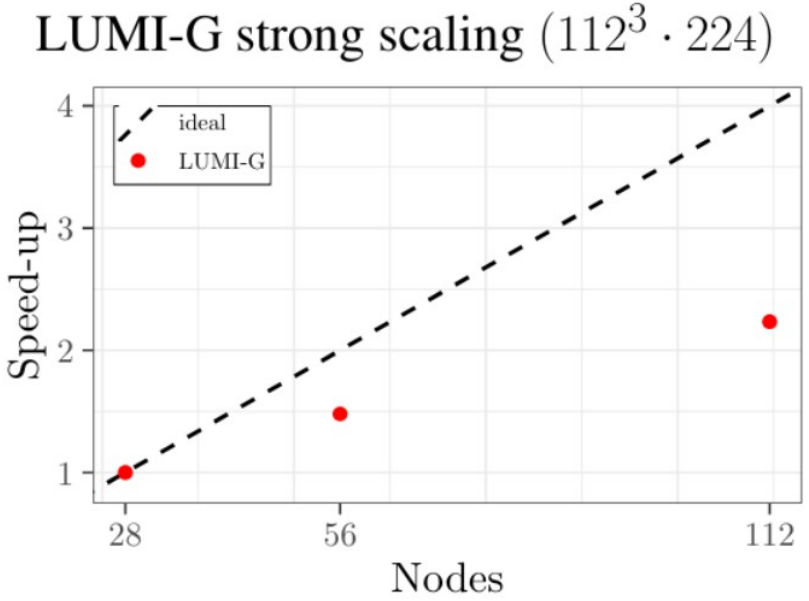
JÜLICH Forschungszentrum

# AMD MX250

## LUMI-G @ CSC - #9




- 2,978 GPU nodes
  - 4x AMD MI250X
  - 64-core AMD EPYC 7A53 "Trento" CPU
  - 512 GiB CPU memory
  - Infinity Fabric Interconnect
- 2,566,080 total accelerator cores
- 379.7 PFlop/s on HPL benchmark (#9 on Top500)
- 4.6 PFlop/s on HPCG benchmark (#5 on Top500)



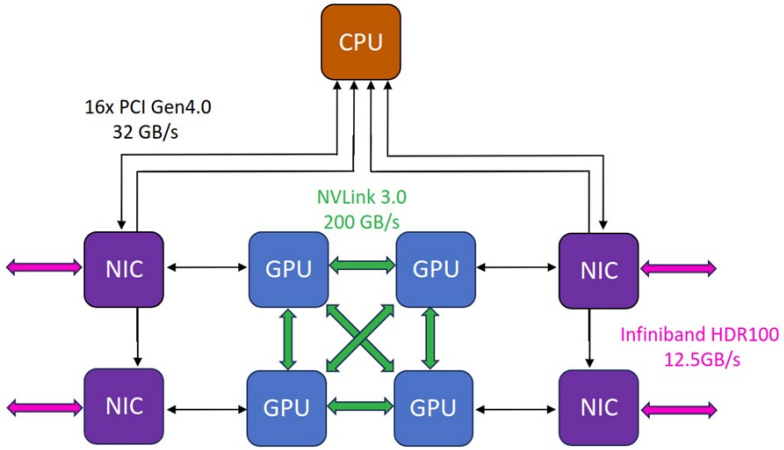

Strong scaling on LUMI-G with tmLQCD + QUDA. Extend Twisted Mass Collab. *Status of the ETMC ensemble generation effort.* LATTICE24



JÜLICH Forschungszentrum

# NVIDIA A100

## Leonardo Booster @ CINECA - #10


- 3,456 nodes with:
  - 4x NVIDIA A100 Custom GPUs (124 SM)
  - 1x Intel Ice Lake Xeon Platinum 8358 processor
  - 512 GiB DDR4 memory @ 3200 MHz
  - Quad-rail NVIDIA HDR100 Infiniband
- 1,714,176 total accelerator cores
- 241.2 PFlops/s on HPL benchmark (#10 on Top500)
- 3.11 Pflops/s on HPCG benchmark (#8 on Top500)

| Application name | Domain | Nodes | TTS | ETS |
|---|---|---|---|---|
| QuantumEspresso | Quantum Chemistry | 12 | 439 | 1.14 |
| MILC | Quantum Chromodynamics | 12 | 178 | 0.56 |
| SPECFEM3D | Solid Earth | 16 | 270 | 1.43 |
| PLUTO | Astrophysics | 32 | 2874 | 11.7 |

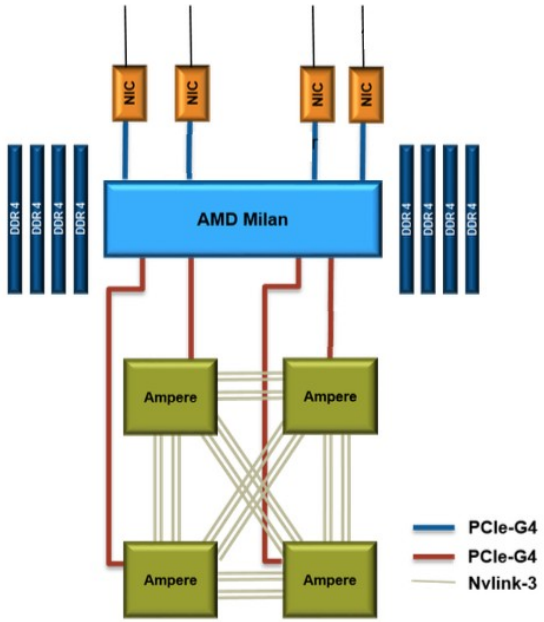
Time to solution (TTS) in seconds and energy to solution (ETS) in kWh of the MILC benchmark. Instrument Scientists, Journal of large-scale research facilities, 8, A186 (2024)



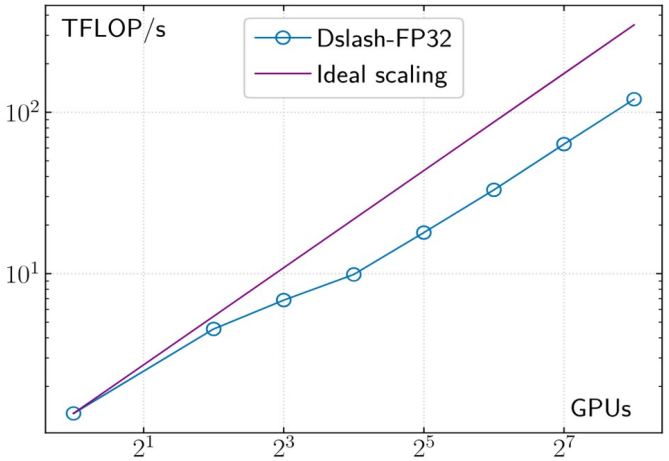
JÜLICH Forschungszentrum

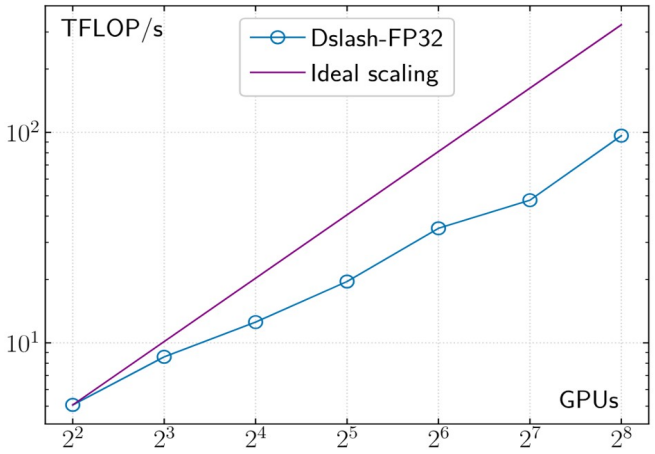
# NVIDIA A100

## Perlmutter @ NERSC - #25


- 1792 GPU nodes
  - 4x NVIDIA A100 40 GB (256 nodes with 80 GB)
  - 1x AMD EPYC 7763
  - 256 GB DDR4 DRAM
  - 4x HPE Slingshot 11
- 774,144 total accelerator cores
- 79.2 PFlop/s on HPL benchmark (#25 on Top500)
- 1.9 PFlop/s on HPCG benchmark (#9 on Top500)



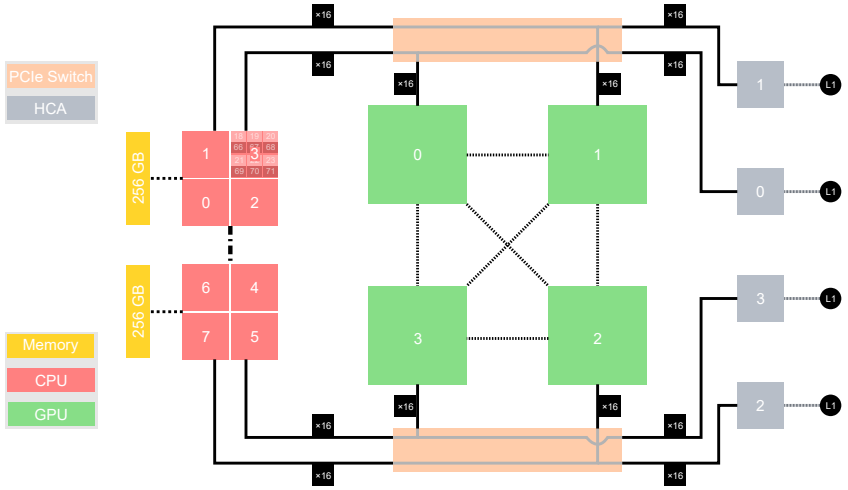

HISQ Dirac operator scaling on Perlmutter in SIMULATeQCD. HotQCD Collab. Computer Physics Communications 300 (2024) 109164 March 2024



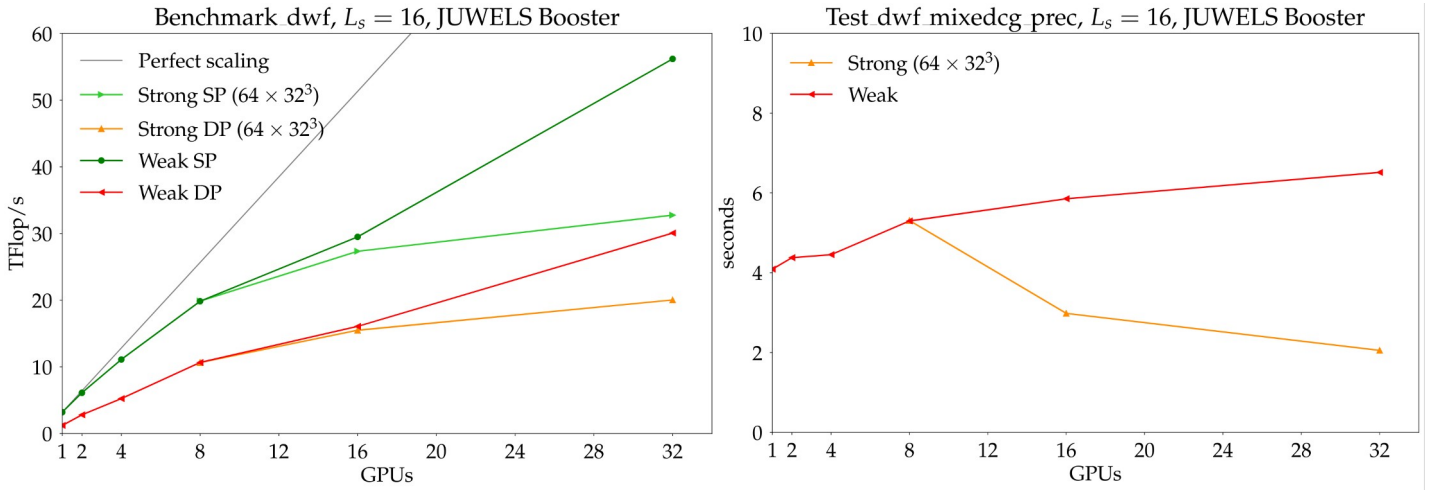
JÜLICH Forschungszentrum

# NVIDIA A100

## JUWELS Booster @ JSC - #43


- 936 GPU nodes
  - 4x NVIDIA A100 40 GB
  - AMD EPYC 7402 with 2 sockets and 24 cores per socket
  - 512 GB DDR4 DRAM
  - 4x Mellanox HDR200 InfiniBand ConnectX 6
- 404,352 total accelerator cores
- 44.1 PFlop/s on HPL benchmark (#43 on Top500)
- 1.3 PFlop/s on HPCG benchmark (#12 on Top500)

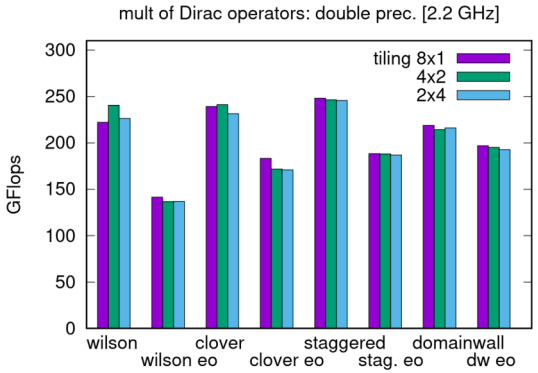


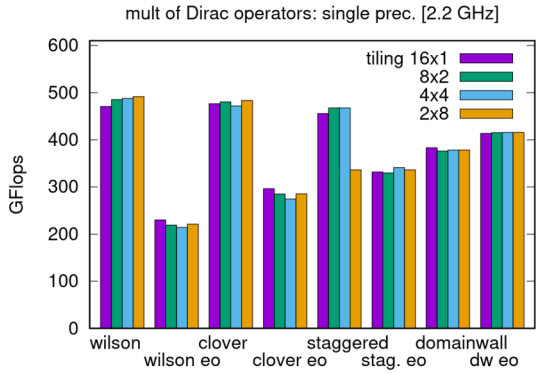

Scaling of Domain Wall fermions on JUWELS with Grid. PUNCH4NFDI.



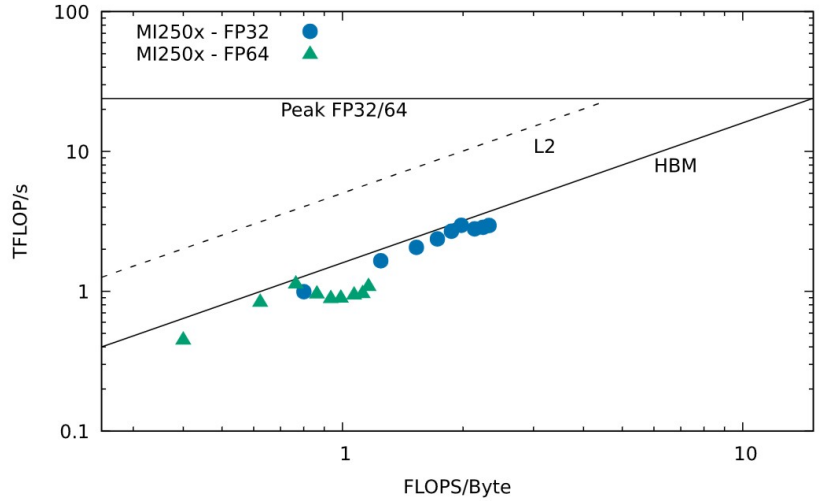
JÜLICH Forschungszentrum

# LATTICE BENCHMARKS



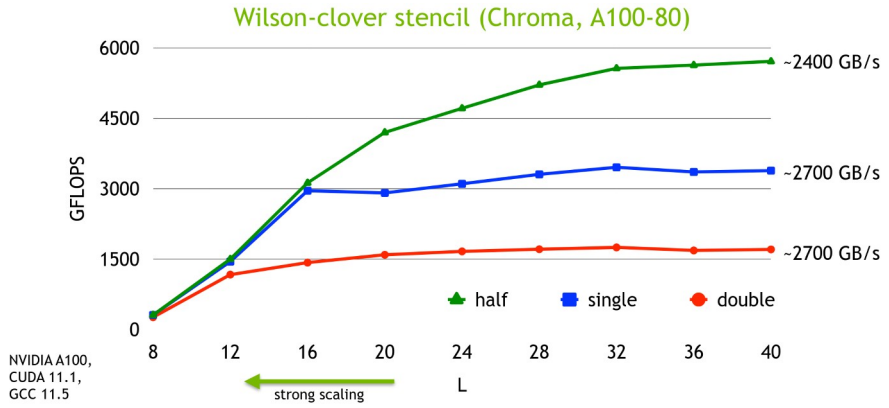




Performance of Dirac operator applications on a single ARM A64FX with Bridge+ +2.0. Aoyama et. al. *Bridge++2.0: Benchmark Results on supercomputer Fugaku. PoS LATTICE2022 284 2023*

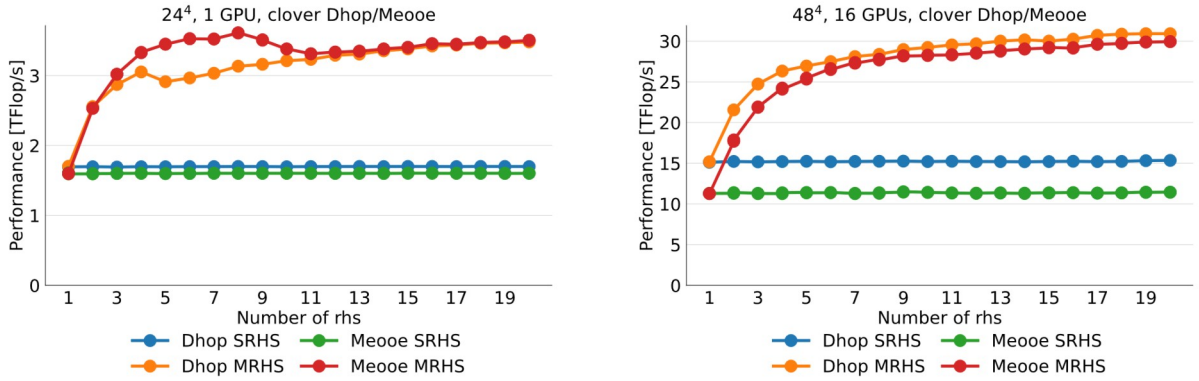



HotQCD Collab. Computer Physics Communications 300 (2024) 109164 March 2024

Roofline plot of the HISQ Dirac operator application with varying number of RHS vectors on a $32^4$ lattice on a single MI250x GCD with SIMULATeQCD.




Wilson-Clover performance on an NVIDIA A100 with Chroma + QUDA. Clark et. al. *Rearchitecting QUDA for multi-RHS computations. LATTICE2024*




Performance of the Wilson-Clover hopping term as a function of the number of RHS for volumes $24^4$ on a single GPU (left) and $48^4$ on 16 GPUs (right) with Grid. Richtmann et. al. *MRHS multigrid solver for Wilson-clover fermions*. PoS Lattice 2022 285 2023

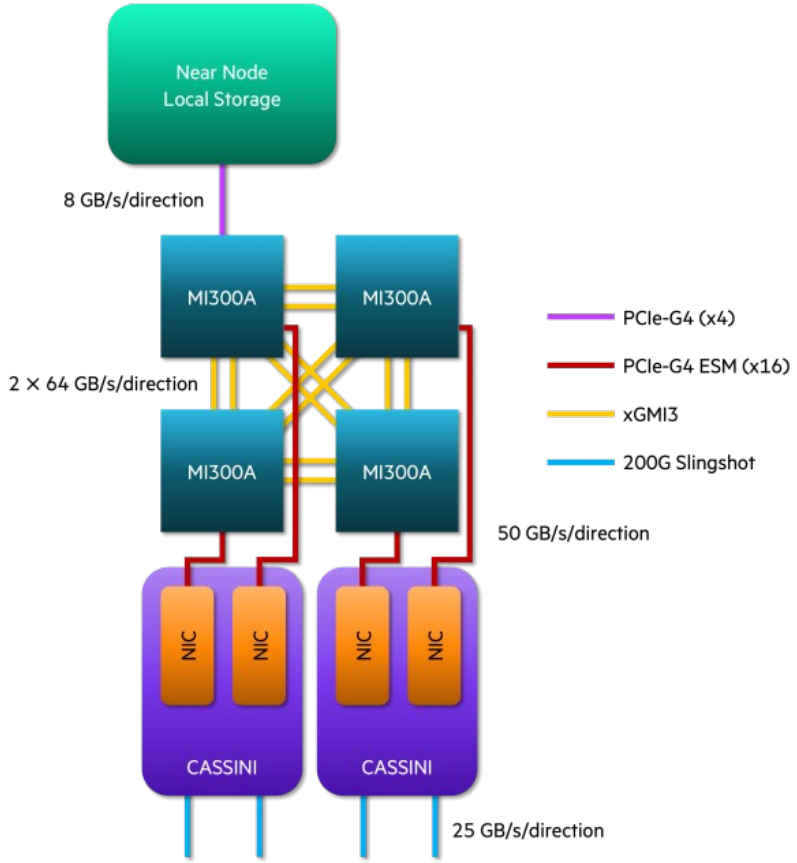


JÜLICH Forschungszentrum

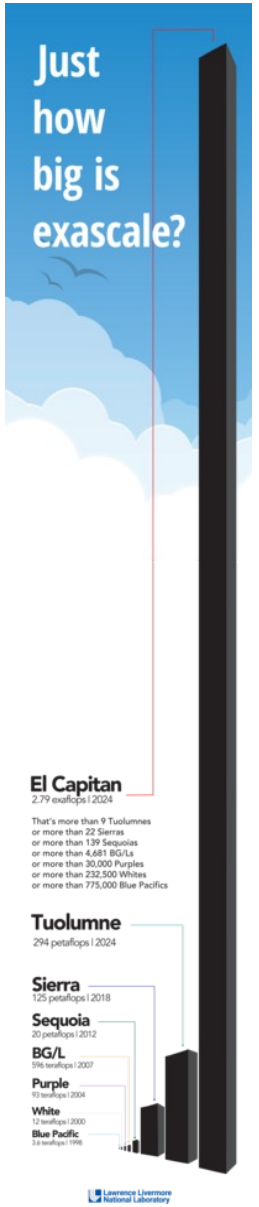
# EXASCALE SYSTEMS

# EL CAPITAN

## LLNL, USA




CPU: AMD EPYC “Genoa”
GPU: **AMD Instinct MI300A** (4 per node)
APU (CPU GPU bundle) 128 GB HBM3 each
Performance peak: 1.74 Eflops (29.6 MW)
Deployment: 2024
#1 in TOP500




JÜLICH Forschungszentrum

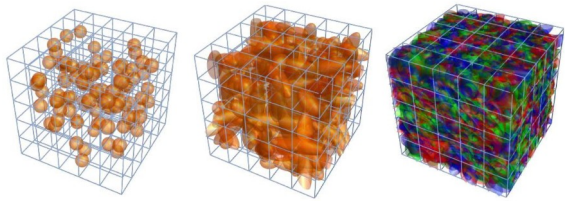
# FRONTIER

## ORNL, USA

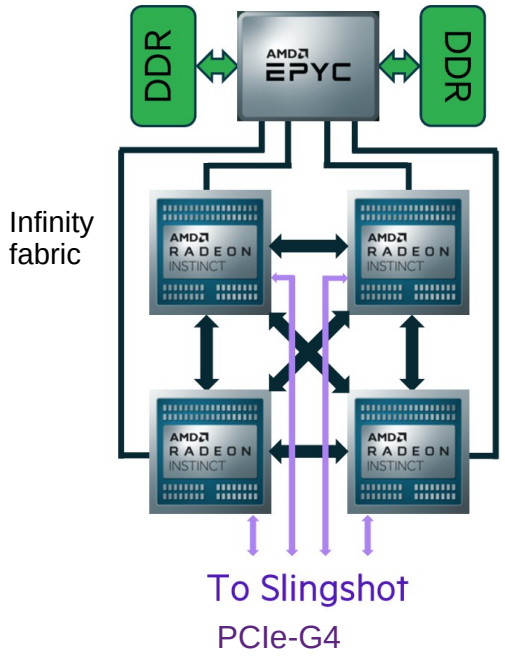
CPU: AMD EPYC “Trento”, 512 GB
GPU: **AMD MI250** (4 per node, 128 GB each)
Performance peak:  1.35 Eflops (24.6 MW)
Deployment: 2022
#2 in TOP500

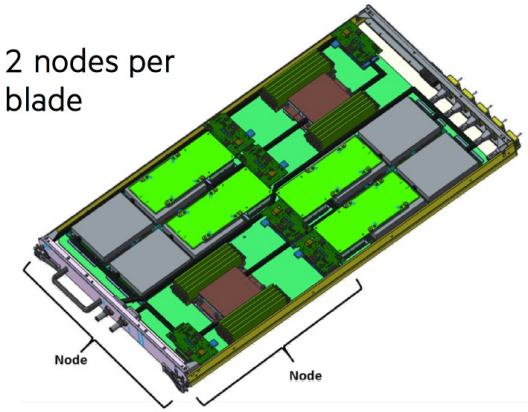

## QCD Constraints on Isospin-Dense Matter and the Nuclear Equation of State

Phys. Rev. Lett. 134, 011903 – Published 6 January, 2025



AMD GPU (ORNL)

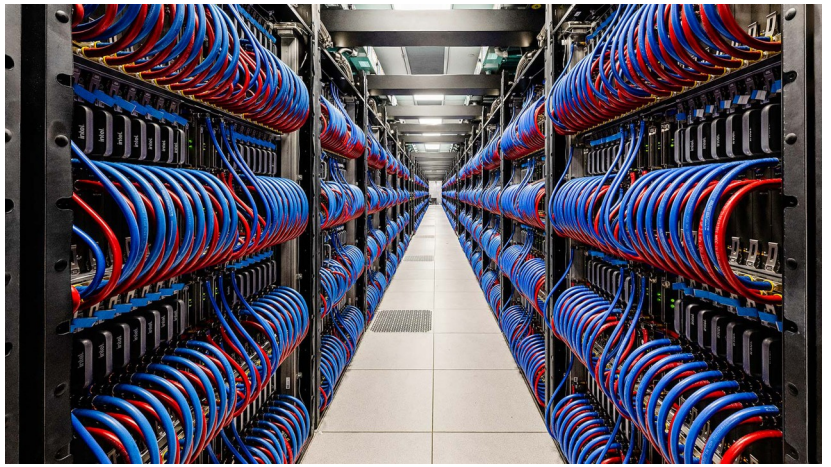


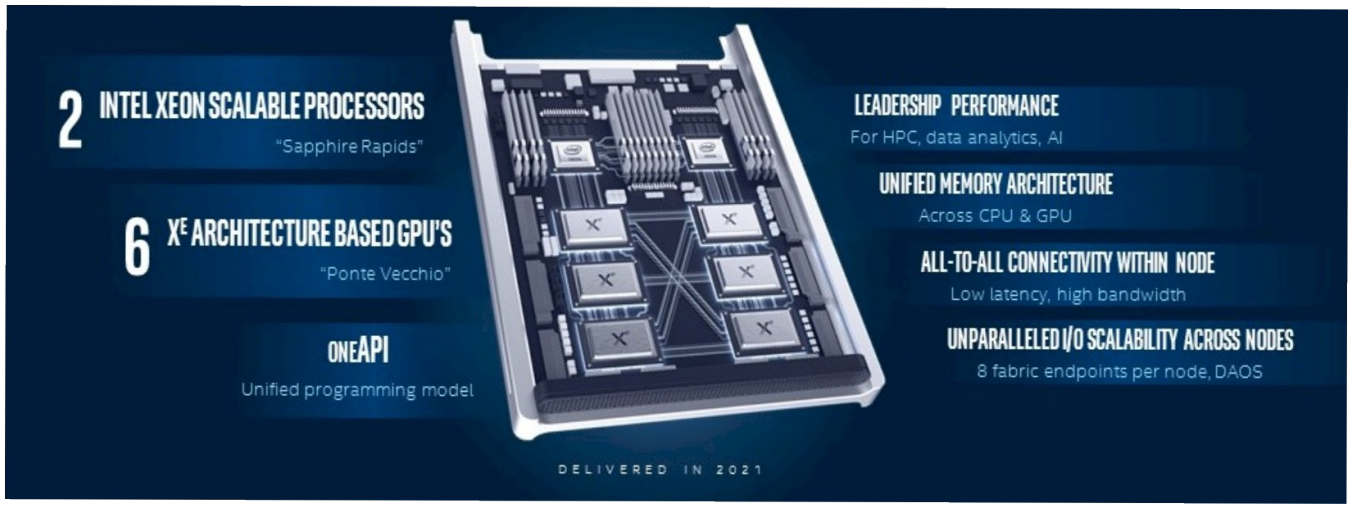



# AURORA

## ANL, USA

CPU: 2 x Intel Xeon Max, 128 GB HBM, 1024 GB DDR5
GPU: **Intel DC GPU Max** (6 per node, 128 GB HBM each)
Performance peak:  1.01 Eflops (38.7 MW)
Deployment: 2023
#3 in TOP500







CPU-CPU: UPI, CPU-GPU: PCIe-G5, GPU-GPU: Xe link, CPU-NIC: PCIe-G5 switched to G4

# AN INCREASINGLY DIVERSE LANDSCAPE


**JUPITER**@JSC (Germany, EU)
CPU: ARM Grace (GH200)
GPU: **H100** (GH200)
Performance peak: 0.79 Eflops
Deployment: 2025
#4 in TOP500

**Fugaku** @ Riken (Japan)
CPU: ARM
GPU: **none**
Performance Peak: 0.44 Eflops
Deployment: 2021
#7 in TOP500


**ALPS**@CSCS (Switzerland)
CPU: ARM Grace (GH200)
GPU: **H100** (GH200)
Performance peak: 0.43 Eflops
Deployment: 2024
#8 in TOP500


**LUMI**@CSC (Finland, EU)
CPU: AMD EPYC
GPU: **AMD MI250**
Performance peak: 0.38 Eflops
Deployment: 2023
#9 in TOP500

**Leonardo**@CINECA (Italy, EU)
CPU: Intel Xeon
GPU: **Nvidia A100**
Performance peak: 0.17 Eflops
Deployment: 2022
#10 in TOP500

June 2025: 4/10 top systems use Nvidia GPUs, 4/10 AMD GPUs, 1/10 Intel GPUs, 1/10 no GPUs
November 2022: 5/10 top systems use Nvidia GPUs, 3/10 no GPUs, 2/10 AMD GPUs, 0/10 Intel GPUs
November 2021: 7/10 top systems used Nvidia GPUs, 3/10 no GPUs, 0/10 AMD or Intel GPUs

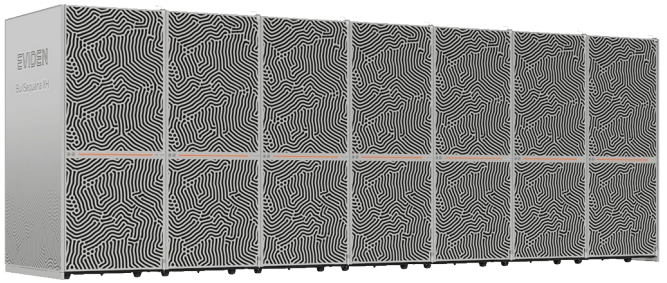
# JUPITER

# JUPITER

## The first exascale machine in Europe

EuroHPC Joint Undertaking

- ParTec/Eviden Supercomputer Consortium
- Implementing Modular Supercomputing Architecture
- JUPITER Booster: High scalability; 1 EFLOP/s HPL, >70 EFLOP/s FP8
- JUPITER Cluster: High versatility; 0.5 B/FLOP balance
- Network: 200/400 Gigabit NVIDIA Mellanox InfiniBand NDR
- Storage: 29 PB Flash IBM Storage Scale 6000
- **17 Megawatt** Linpack Power Consumption
- Direct Liquid Cooled (36 -> 4x degree) to enable heat-reuse



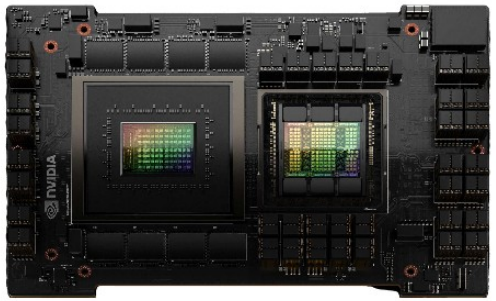
BullSequana XH3000; DLC

ParTec MODULAR SUPERCOMPUTING EVIDEN NVIDIA SIPEARL

# JUPITER MODULES

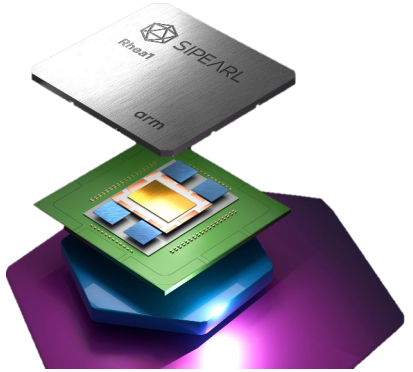
## JUPITER Booster

- ~125 Racks BullSequana XH3000
- Node design
  - ~6000 nodes
  - 4× NVIDIA CG1 per node
- CG1: NVIDIA 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

- ~14 Racks BullSequana XH3000
- Node design
  - ~1300 nodes
  - 2× SiPearl Rhea1 per node
- Rhea1
  - 80 Arm Neoverse V1 cores (2×256b SVE)
  - 256 GB DDR5, 64 GB HBM2e




**ExaFLASH:** 29PB (raw) NVMe, IBM SS6000

**ExaSTORE:** 308PB (raw) HDD, IBM SS6000

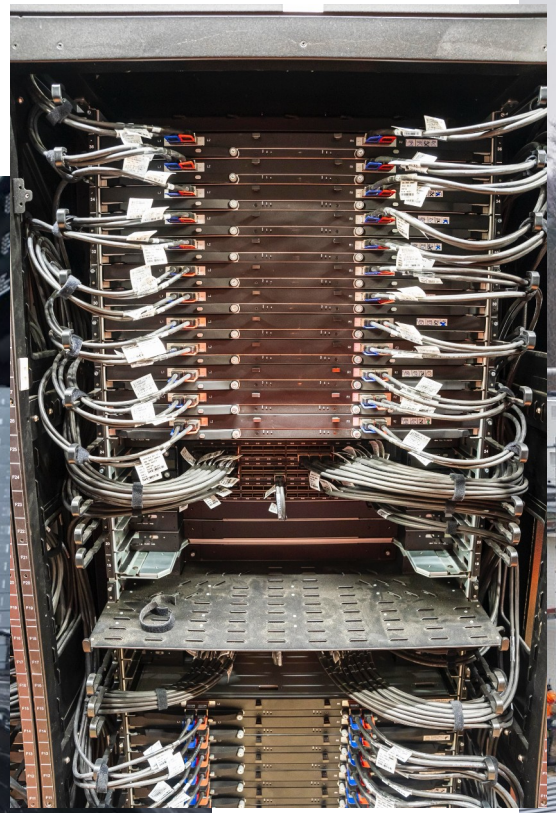
**ExaTAPE:** 370PB Tape, LTO9



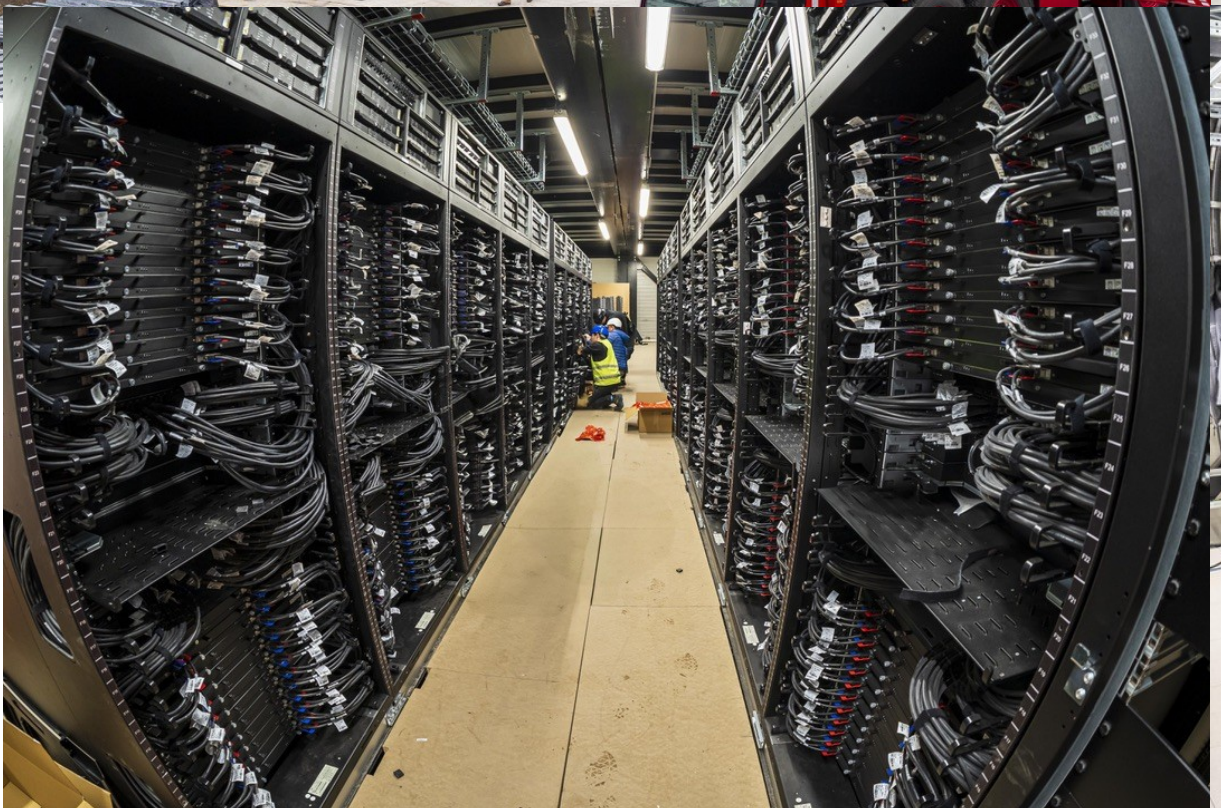
JÜLICH Forschungszentrum

# JUPITER ASCENDING

## Since January 2025

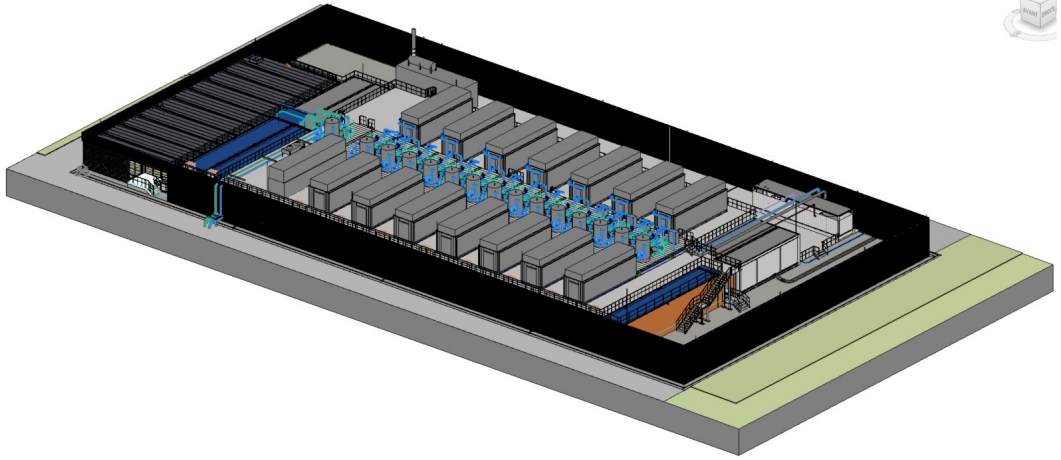



0221 / 956 86-0
www.COLONIA.info



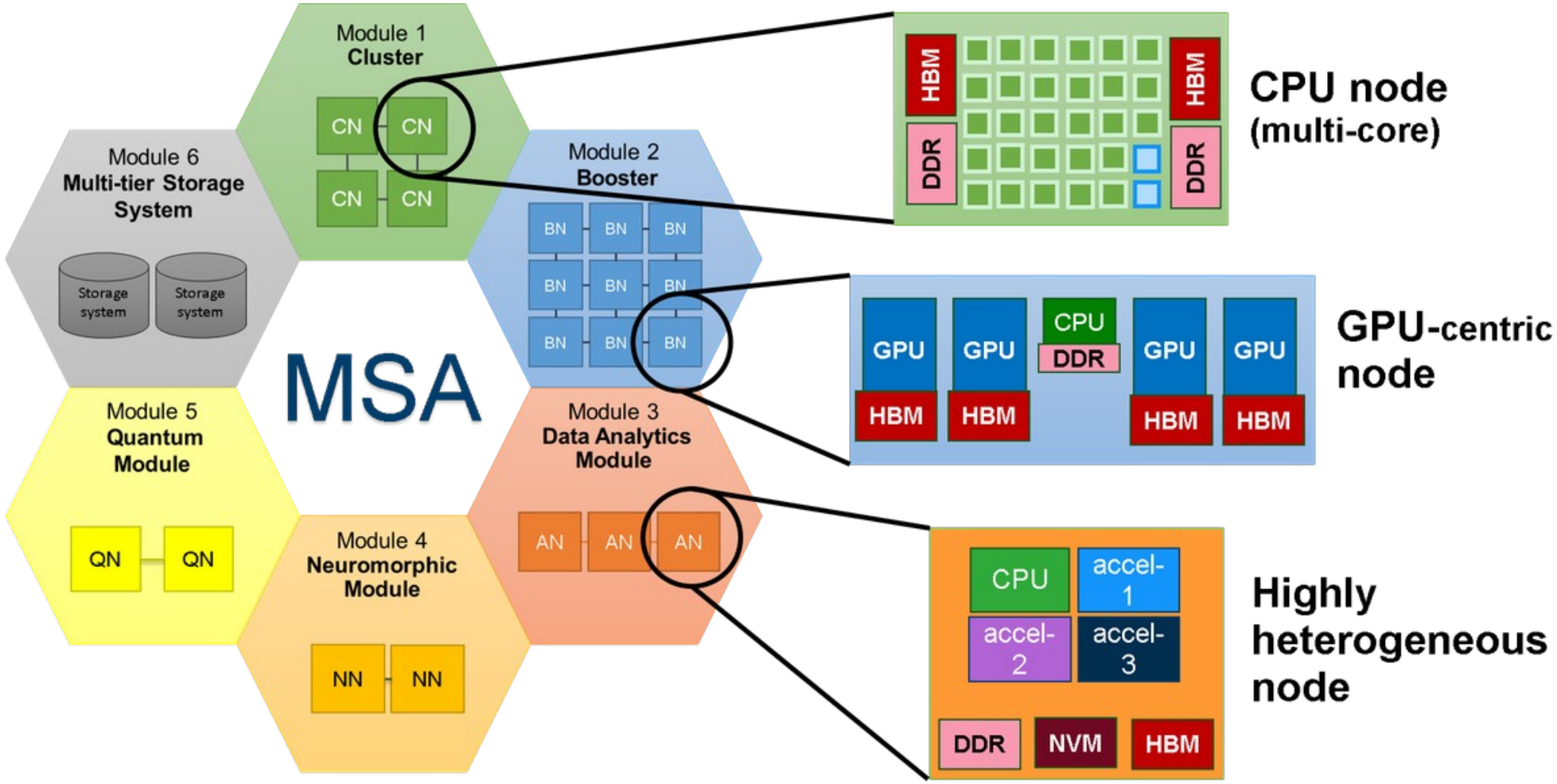
# MODULAR DATA CENTER FOR JUPITER

- Vendor: Eviden
- Area: ~2300m²
- 1x Datahall (Storage, Management)
- 7x IT Modules (20 Racks per module)
- UPS, Generator
- Entrance area
- Workshop, Warehouse
- 15x 2,5 Megawatt Power Stations



# GENERAL PURPOSE VS SPECIALIZED SYSTEMS

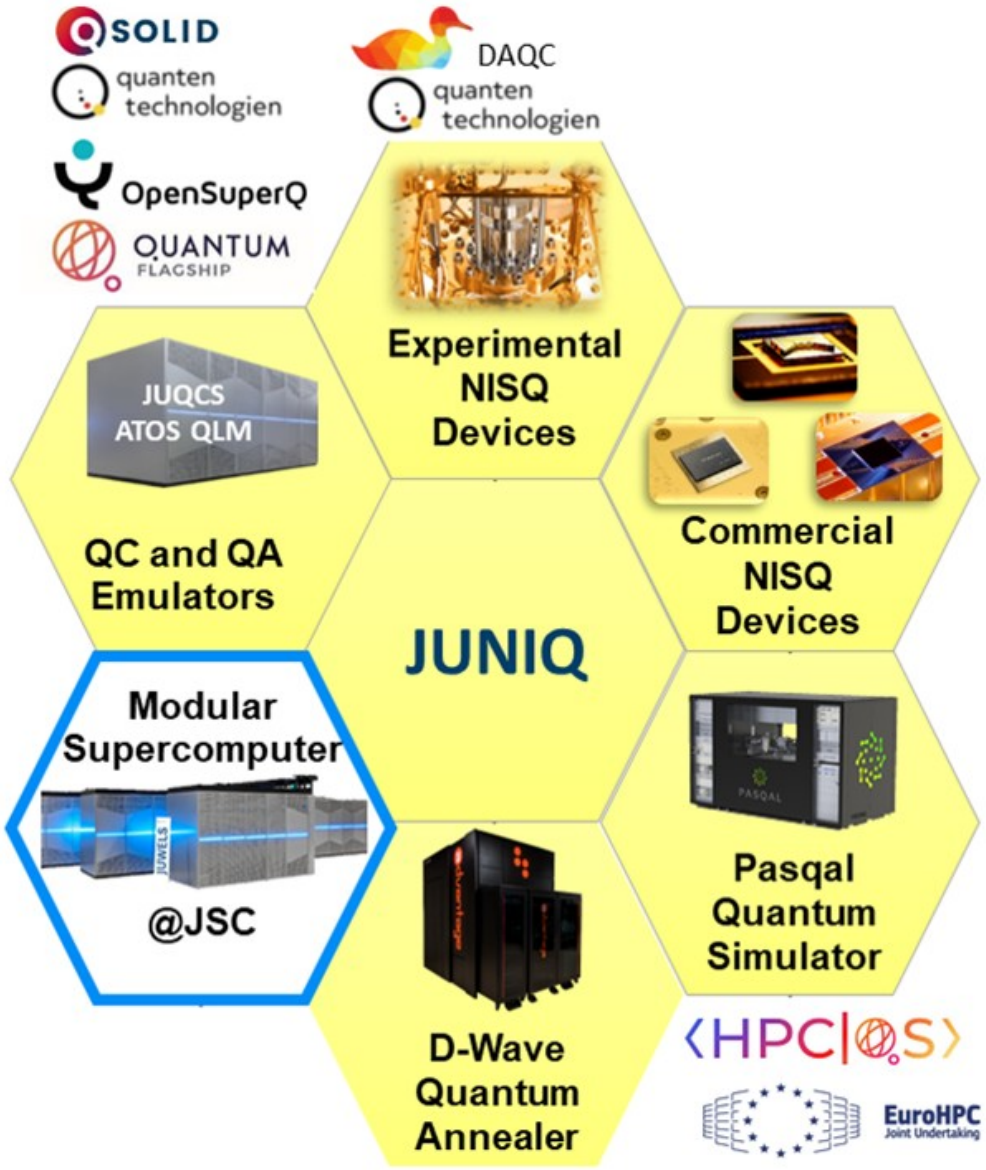
## Modular Supercomputing Architecture

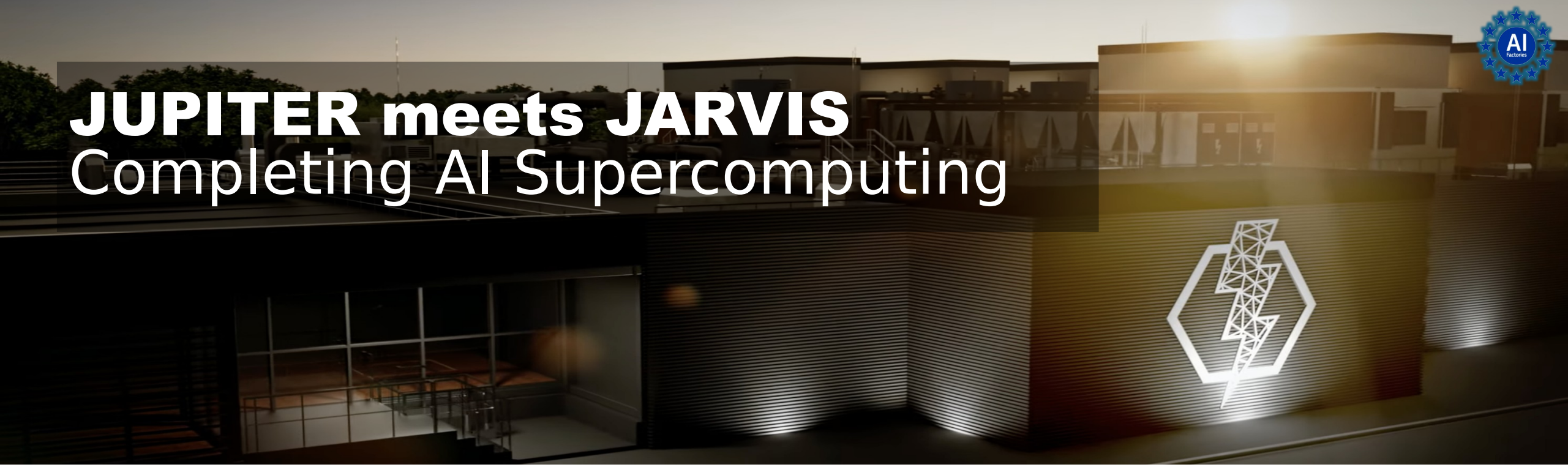


# JUNIQ USER FACILITY

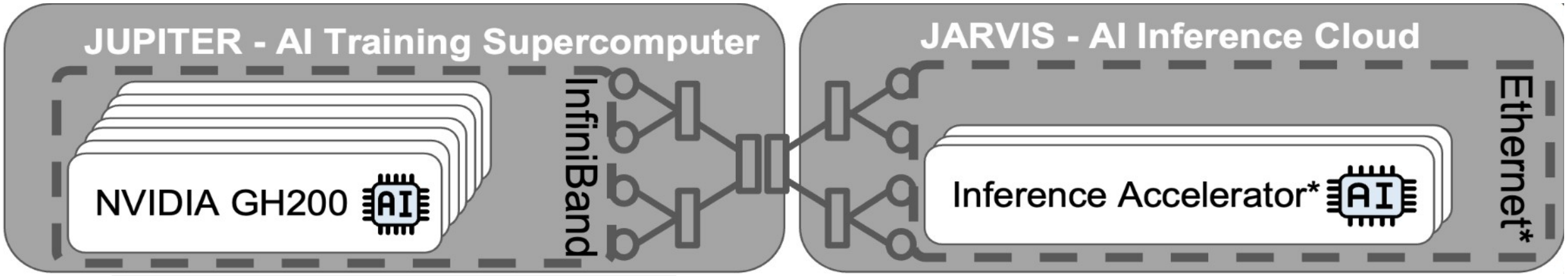
## Quantum computing

- JUNIQ provides access to QC resources
- Usual merit based procedure (proposals)
- Presently available
  - 5000 QBit D-Wave
  - Several smaller devices
- Upcoming
  - 25 QBit digital system
  - 100 QBit Pasqal system





# JUPITER meets JARVIS
## Completing AI Supercomputing




**40 ExaFLOP/s FP8 AI performance**

24 000 NVIDIA GH200 superchips

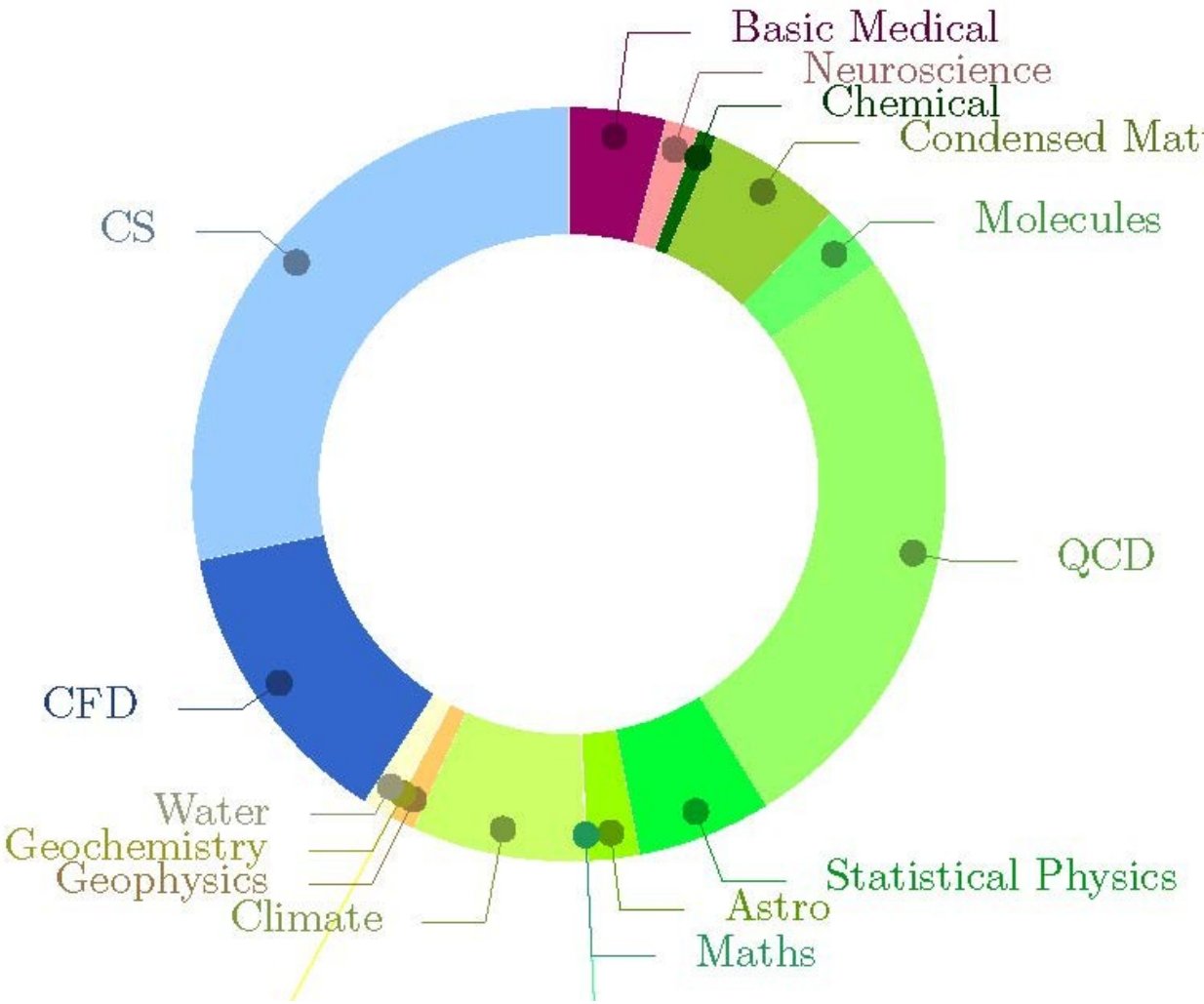
**Scalability, Sovereignty**

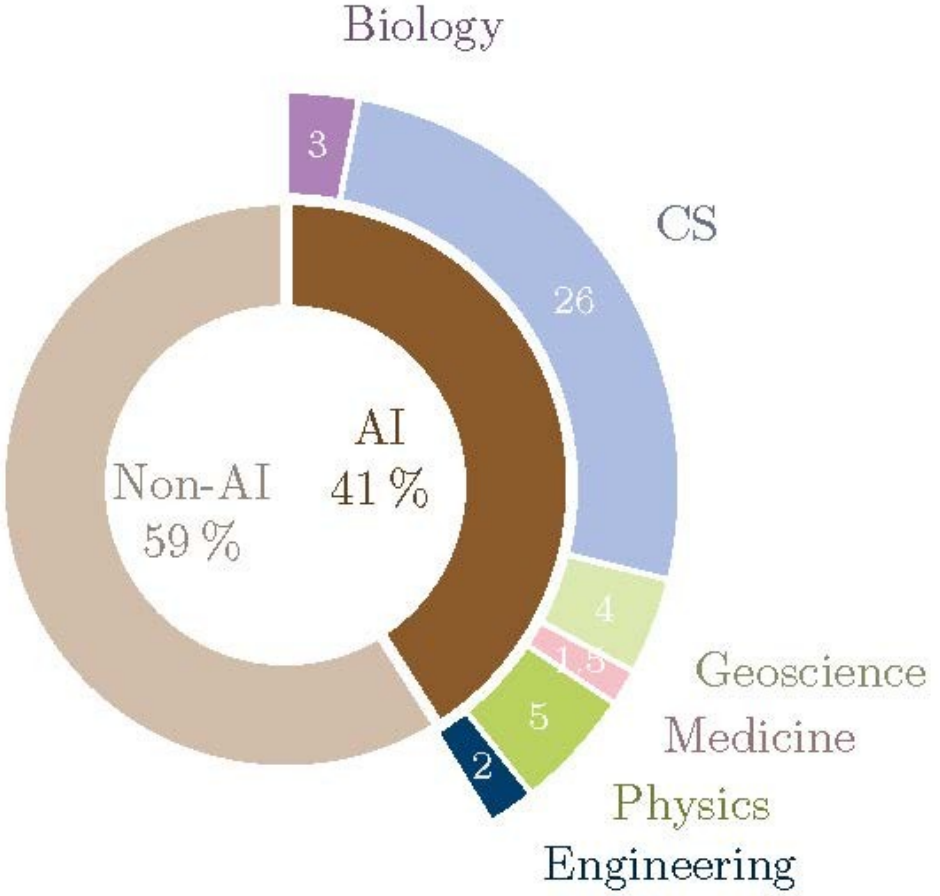
European (experimental) technology

JUPITER AI FACTORY

# BENCHMARKING

→ Use For Benchmarks



# APPLICATIONS FOR THE JUPITER PROCUREMENT

- Selection criteria
  - Current workload
  - Future workload
  - Relevance
  - Balance with other applications
    - Domains
    - Programming models
    - Programming languages
    - Profile
- High Scalability up to Exascale

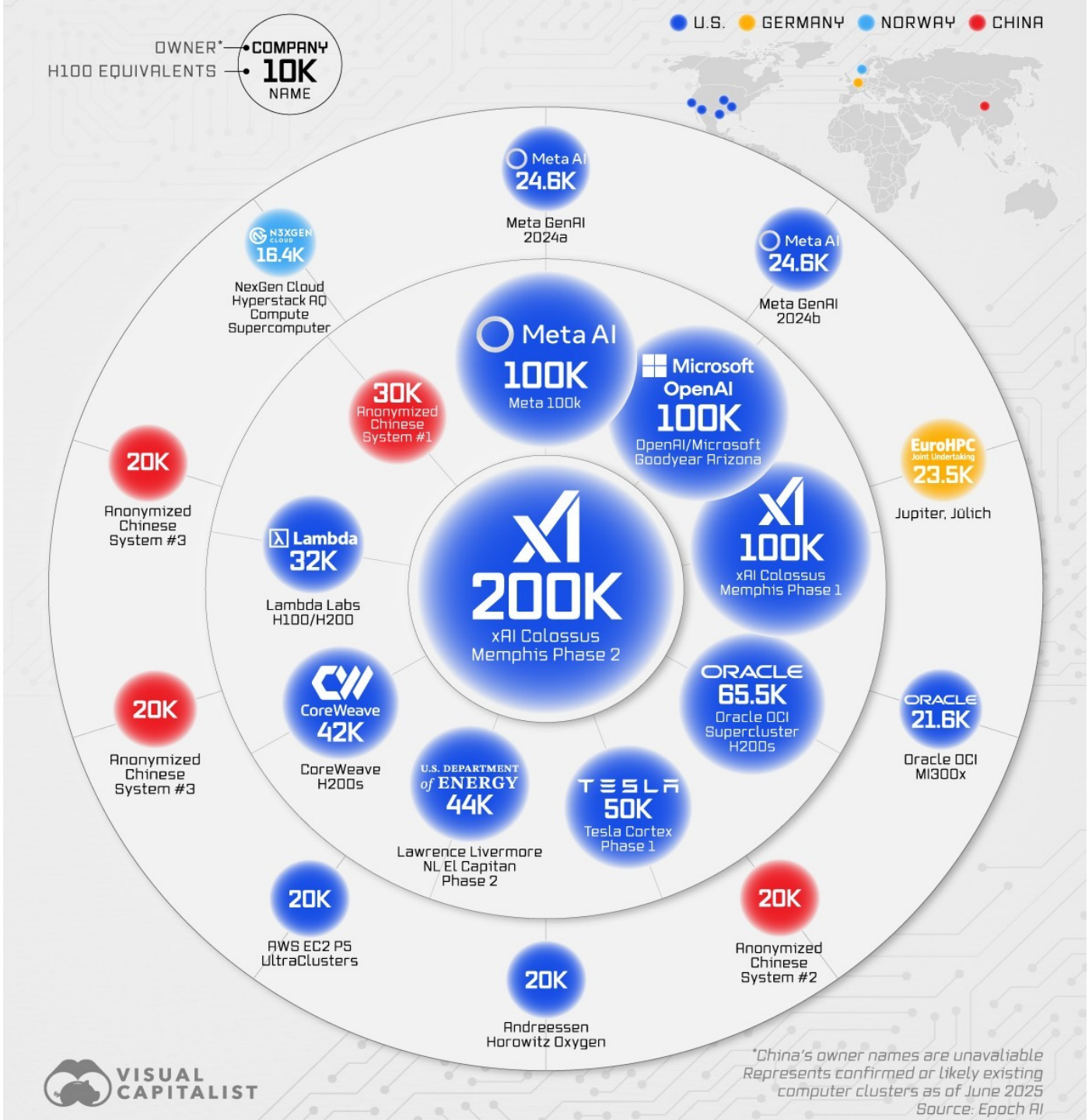
| | Booster | | | Cluster | MSA |
|---|---|---|---|---|---|
| Benchmark | GPU | GPU High-Scale | CPU | CPU | |
| Arbor | ✓ | ✓ | | | |
| Chroma | ✓ | ✓ | | | |
| Gromacs | ✓ | | | | |
| ICON | ✓ | | | | |
| JUQCS | ✓ | ✓ | | | ✓ |
| nekRS | ✓ | ✓ | | | |
| ParFlow | ✓ | | | | |
| PIConGPU | ✓ | ✓ | | | |
| Quantum ESPRESSO | ✓ | | | | |
| AI-MMoCLIP | ✓ | | | | |
| AI-NLP | ✓ | | | | |
| dynQCD | | | | ✓ | |
| NAStJA | | | | ✓ | |
| Graph500 | | | ✓ | | |
| HPCG | ✓ | | | ✓ | |
| HPL | ✓ | | | ✓ | |
| IOR | | | ✓ | ✓ | |
| LinkTest | | | ✓ | ✓ | ✓ |
| OSU | ✓ | | ✓ | ✓ | |
| STREAM | ✓ | | | ✓ | |

# THE FUTURE

# REALITY...

## ...sets in

AI takes over





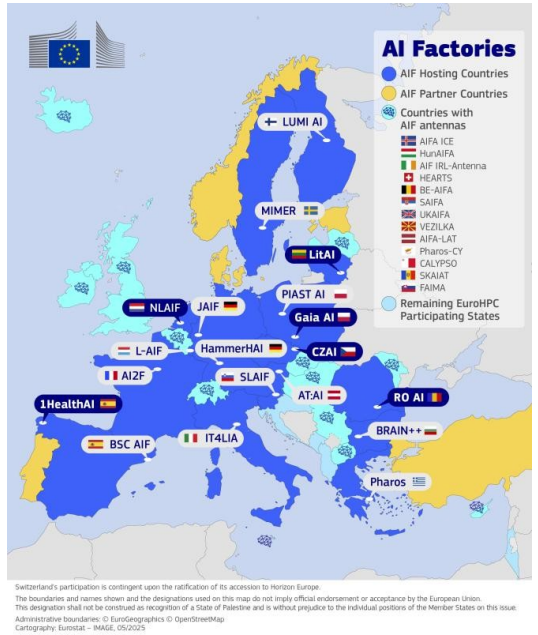

JÜLICH
Forschungszentrum

# HARDWARE TRENDS – AI RULES

## AI-Focused Projects

- Project Stargate, USA (OpenAI, Oracle, Softbank)
  - $100 billion investment in AI supercomputers

- 5 AI Gigafactories + 19 AI Factories, EU
  - €20 billion for large datacenters as part of €200 billion fund




- Mission, Vision (LANL), Solstice, Equinox (ANL), Lux (ORNL)
  - Upcoming DOE AI-specific systems



JÜLICH Forschungszentrum

# HARDWARE TRENDS

See talk by Mathias Wagner Tue 17:40

## Push to low precision HW

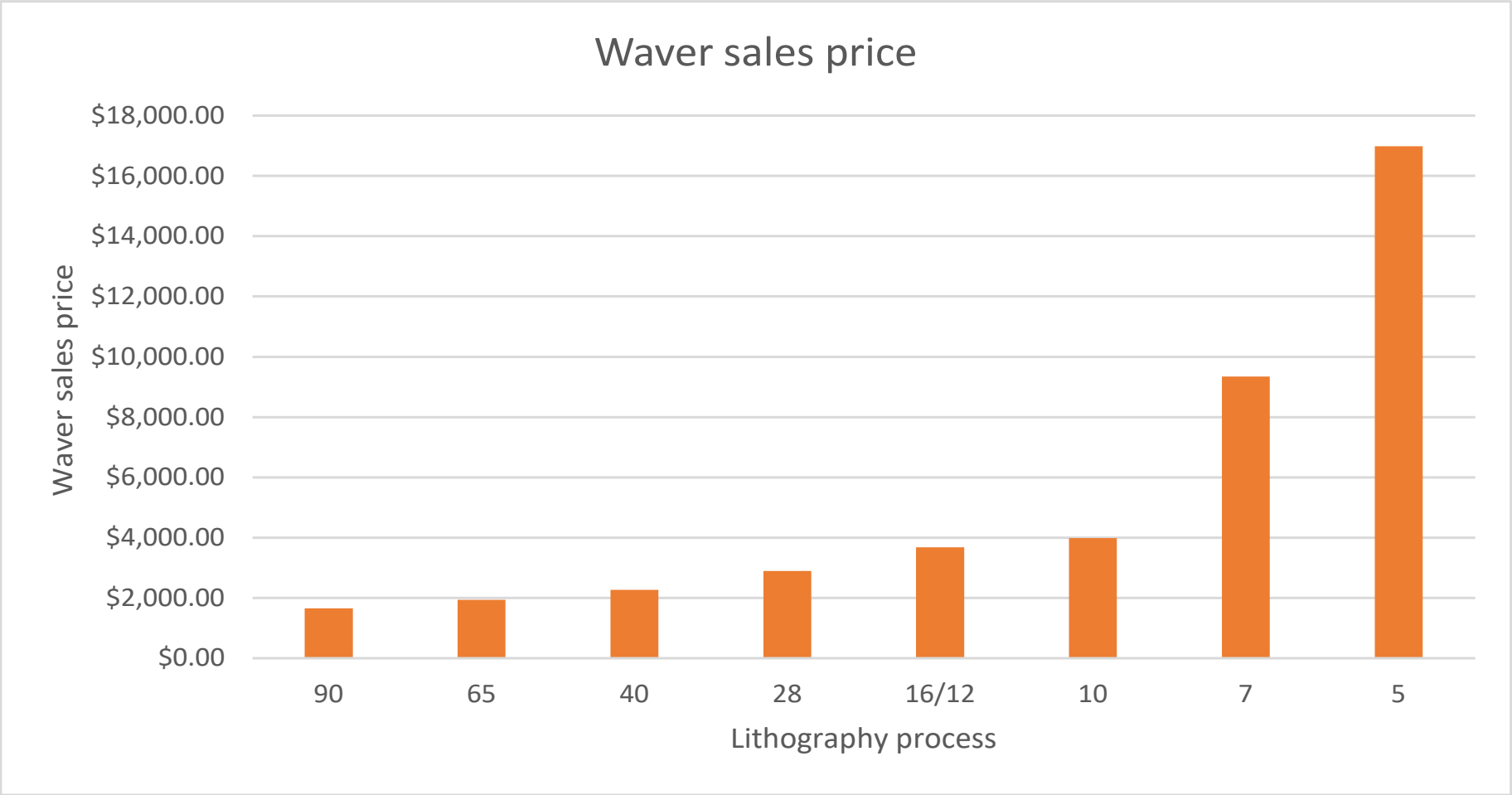
| Device | FP64 TC | TF32 TC (dense) | FP4 TC (dense) | GPU mem (GB) | Mem BW (TB/s) | CPU | CPU mem (GB) | Max TDP (W) |
|---|---|---|---|---|---|---|---|---|
| Vera Rubin | TBD | TBD | 100,000 | ≈576 (HBM4) | ≈26 (est.) | Vera (88c) | 2,000 (LPDDR5X) | yes |
| GB200 | 80 | 2,200 | 18,000 | 372 (HBM3e) | 16 | Grace (72c) | 480 (LPDDR5X) | 2,700 |
| B300 | 1.2 | 1,250 | 15,000 | 270 (HBM3e) | 8.0 | — | — | 1,400 |
| B200 | 40 | 1,100 | 9,000 | 192 (HBM3e) | 7.7 | — | — | 1,000 |
| B100 | 30 | 900 | 7,000 | 192 (HBM3e) | 8.0 | — | — | 700 |
| GH200 | 67 | 989 | — | 144 (HBM3e) | 48 | Grace (72c) | 480 (LPDDR5X) | 1000 |
| H100 PCIe | 51 | 756 | — | 80 (HBM2e) | 2.0 | — | — | 350 |

E2M1: −6, −4, −3, −2, −1.5, −1, −0.5, **−0**, **+0**, +0.5, +1, +1.5, +2, +3, +4, +6

# HARDWARE TRENDS

See talk by Mathias Wagner Tue 17:40

## Limits to growth




https://www.tomshardware.com/news/tsmcs-wafer-prices-revealed-300mm-wafer-at-5nm-is-nearly-dollar17000

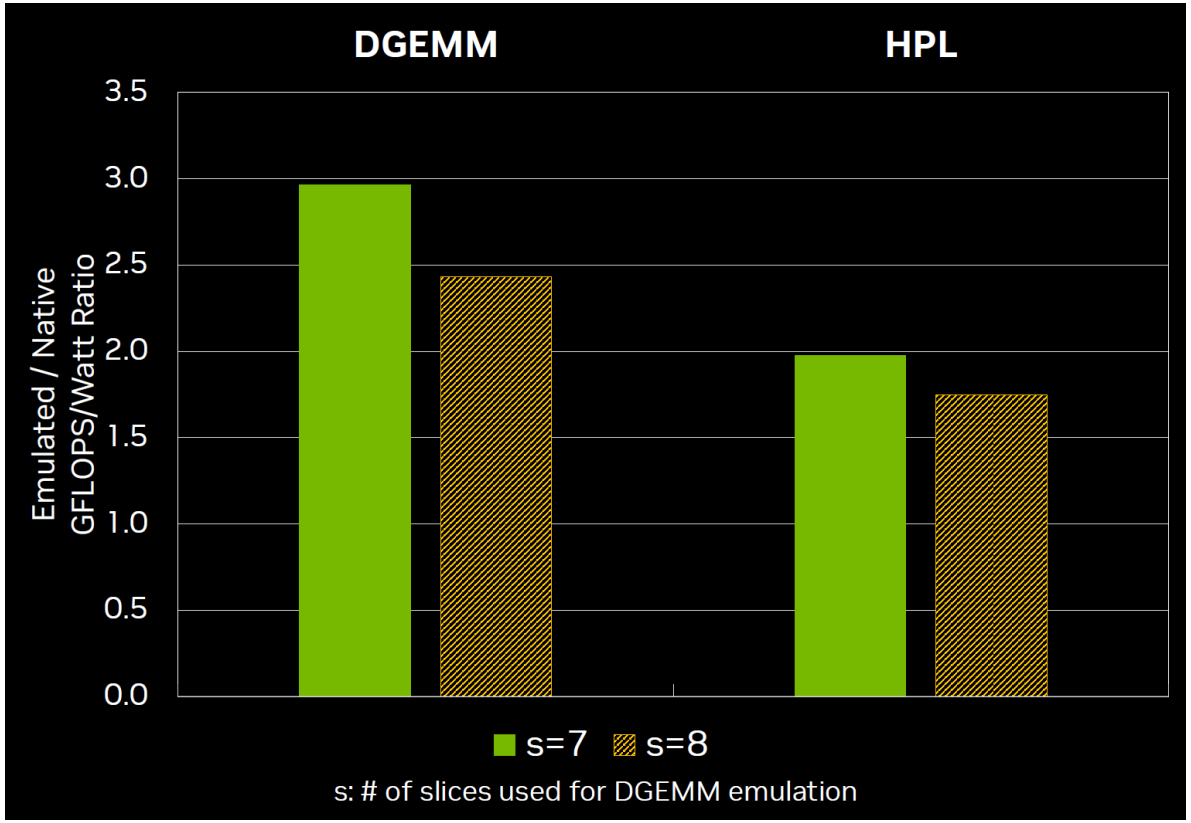
# HARDWARE TRENDS

See talk by Mathias Wagner Tue 17:40

## Utilizing low precision

| Precision | Operation | Energy per FLOP (Matrix Multiply) |
|---|---|---|
| FP64 | FMA | 2.5x |
| **FP32** | **FMA** | **1.0x** |
| FP16 | FMA | 0.5x |
| FP64 | Tensor Core MMA | 1.5x |
| FP16 | Tensor Core MMA | 0.12x |
| FP8 | Tensor Core MMA | 0.06x |
| INT8 | Tensor Core MMA | 0.04x |

FP64 emulation using INT8 Tensor Cores




Ootomo, Ozaki, Yokota, arXiv:2306.11975

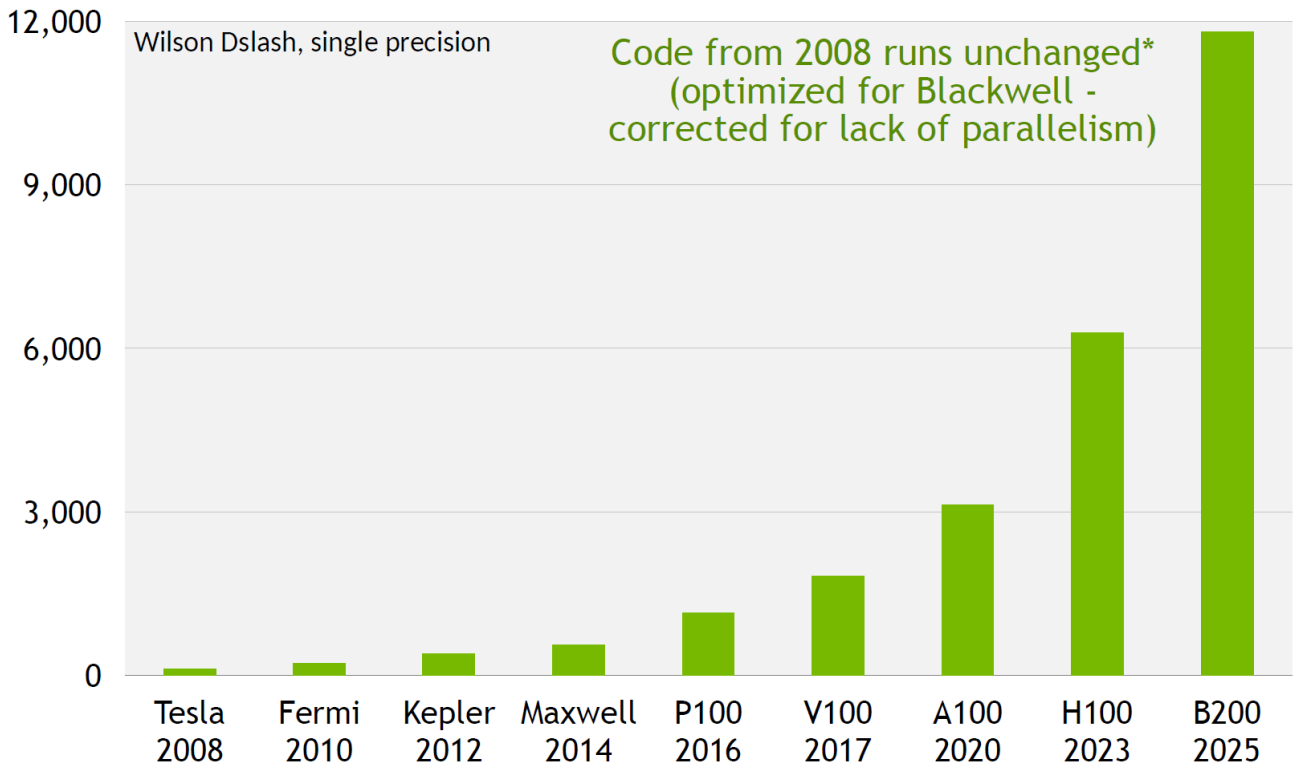


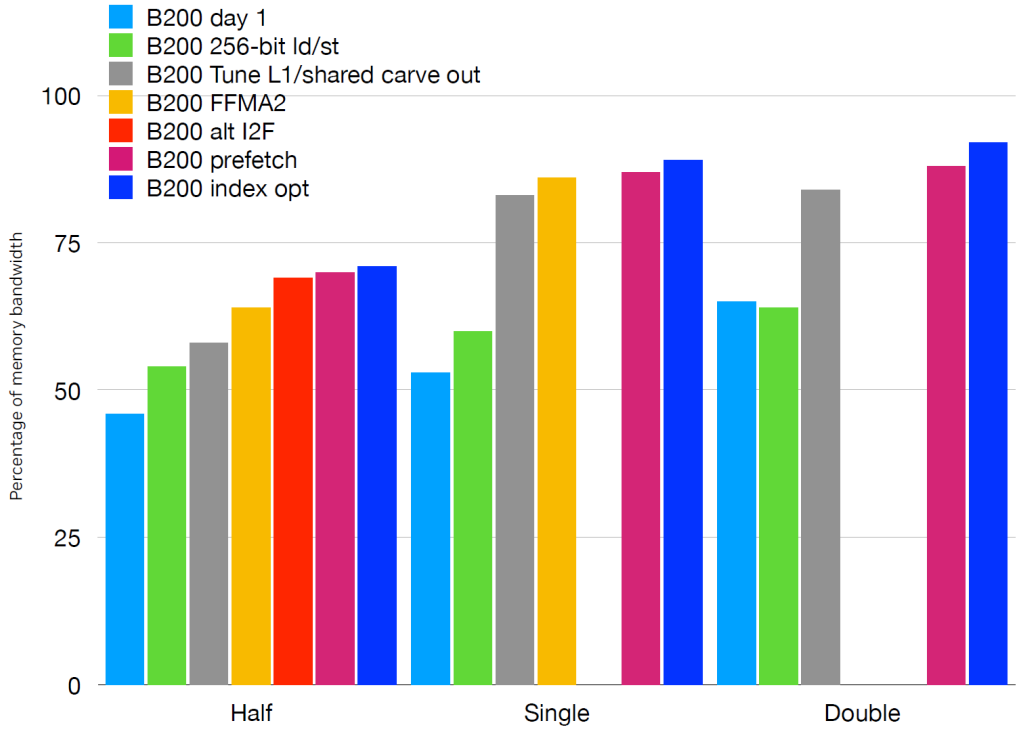
JÜLICH Forschungszentrum

# HARDWARE TRENDS

See talk by Mathias Wagner Tue 17:40

## LQCD performance on NVIDIA



# HARDWARE TRENDS

## AMD – AI again

| Device | FP64 | FP32 | FP4 (Matrix) | Mem (GB) | Mem BW (TB/s) | TDP / TBP (W) |
|---|---|---|---|---|---|---|
| MI250X (CDNA2, OAM) | 47.9 | 47.9 | - | 128 (HBM2e) | 3.2 | 560 |
| MI300X (CDNA3, OAM) | 81.7 | 163.4 | - | 192 (HBM3) | 5.3 | 750 |
| MI300A (CDNA3 APU, SH5) | 61.3 | 122.6 | - | 128 (HBM3) | 5.3 | 760 |
| MI325X (CDNA3, OAM) | 81.7 | 163.4 | - | 256 (HBM3e) | 6.0 | 1,000 |
| MI350X (CDNA4, OAM) | 72.1 | 144.2 | 9,200 | 288 (HBM3e) | 8.0 | 1,000 |
| MI355X (CDNA4, OAM) | 78.6 | 157.3 | 10,100 | 288 (HBM3e) | 8.0 | 1,400 |
| MI450 Series (CDNA “Next”) | TBD | TBD | 40,000 | 432 (HBM4) | 20 | TBD |

# HARDWARE TRENDS

## Upcoming “Traditional” HPC

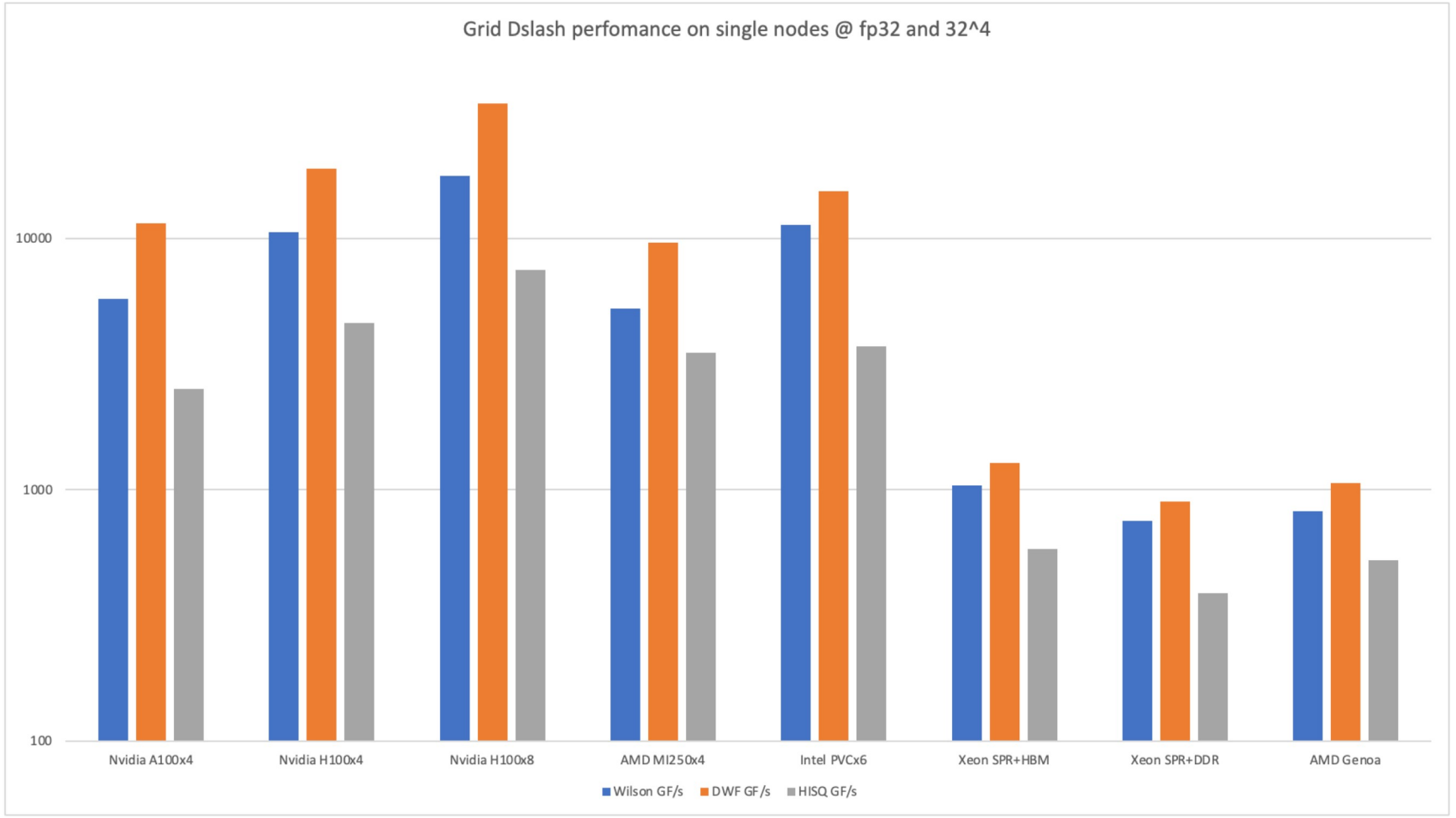
- Discovery (ORNL, USA) AMD EPYC “Venice” & AMD MI430X
- Alice Recocque (CEA, France) – unclear: tender out
- Blue Lion (LRZ, Germany) NVIDIA Vera Rubin
- Horizon (TACC, USA) NVIDIA?
- Doudna (NERSC, USA) NVIDIA Vera Rubin
- FugakuNEXT (RIKEN, Japan) Fujitsu/NVIDIA hybrid

→ Ask the panel!

# HARDWARE TRENDS

## Grid performance for NVIDIA, AMD & intel

Boyle & Lehner




Intel PVC: Aurora@ANL
AMD MI250: Frontier@ORNL,
Lumi-G@CSC
A100: Perlmutter@NERSC,
Booster@Juelich,
Tursa@Edinburgh
Leonard@Cineca
H100, AMD Genoa: SDCC@BNL

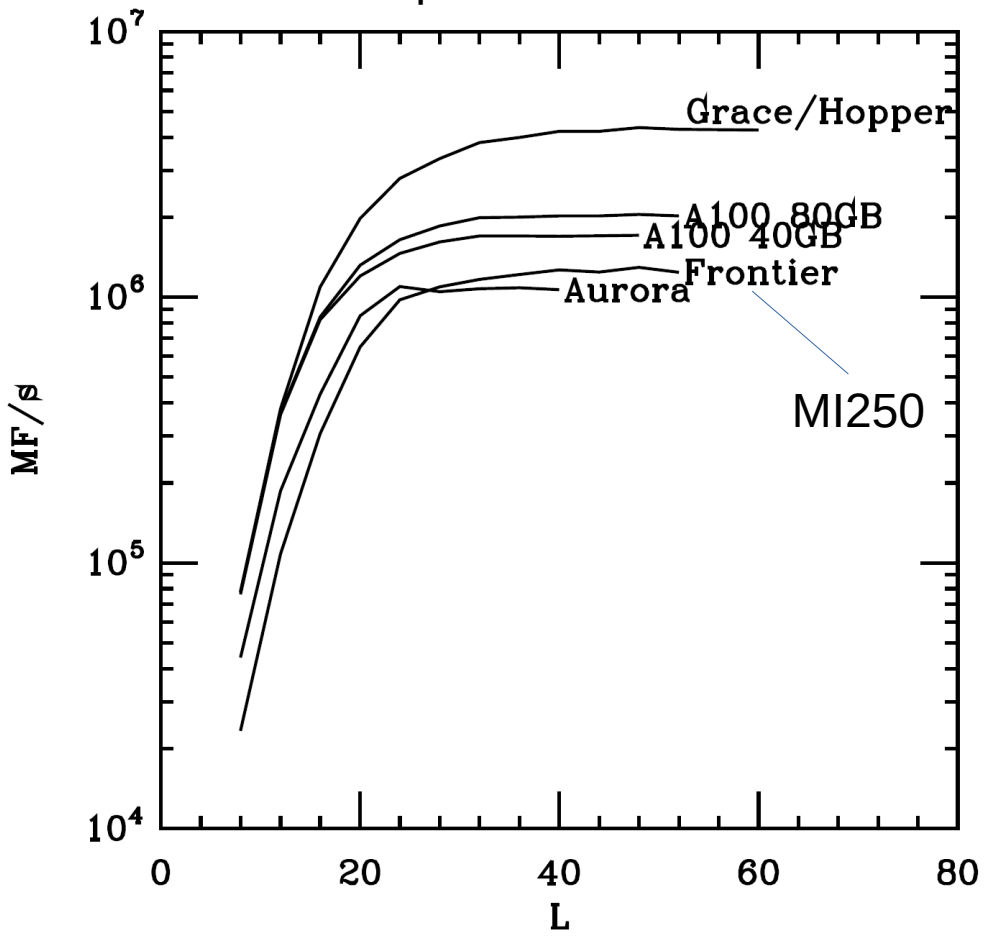


JÜLICH Forschungszentrum

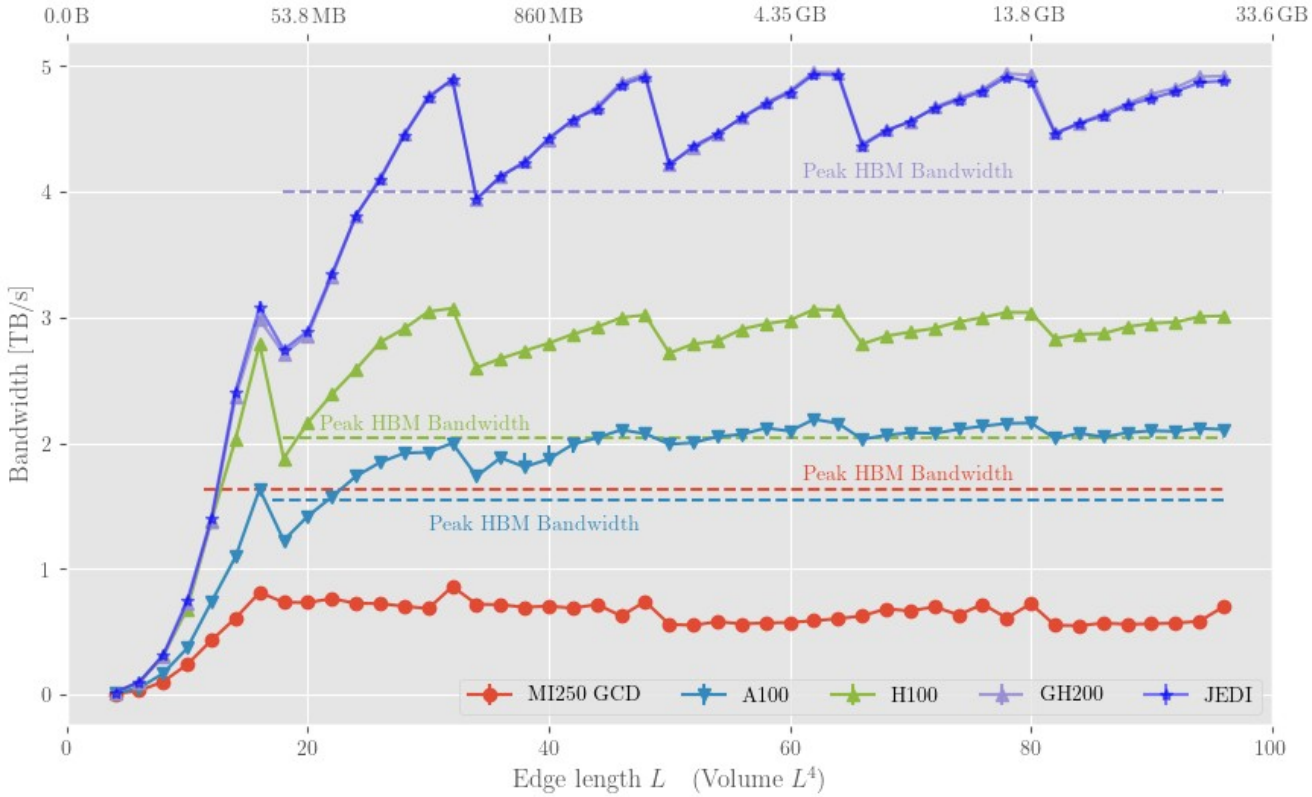
# HARDWARE TRENDS

## MILC performance for NVIDIA, AMD & intel

FP32 CG performance - S. Gottlieb




Kokkos staggered kernel – S. Schlepphorst



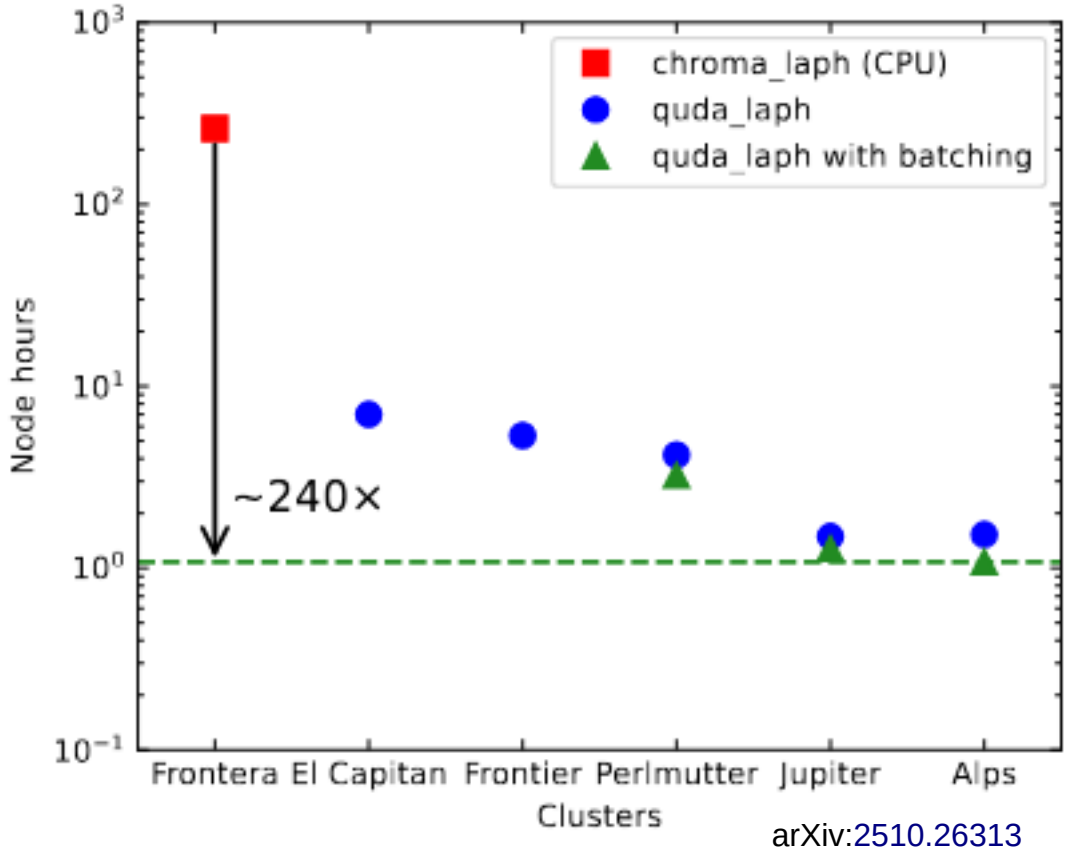



JÜLICH Forschungszentrum

# HARDWARE TRENDS

## .. and usage consequences

- Performance portability
  - Leverage libraries that can easily handle multiple architectures well without changes to algorithms (e.g. Kokkos)
- Task scheduling / job bundling
  - Increase the average size of nodes and improve workflows to play nicely with the schedulers on very large machines. Either in code or with external tools like Flux.
  - Maximize node utilization, don't leave CPUs idling
- Better data management workflows
  - With exascale computing comes exascale data
  - Hopefully upcoming ILDG extensions to new file formats and new data objects (propagators etc...)




arXiv:2510.26313



JÜLICH Forschungszentrum

# EXOTIC HARDWARE

# EXOTICS: RISC-V

## Free and open instruction set architecture (ISA)

### Embedded Systems

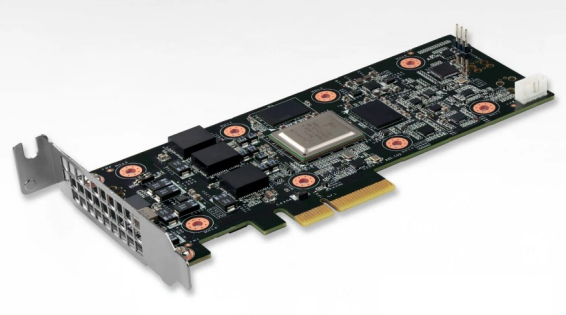
**Milk-V Jupiter**

”RISC-V PC for Everyone”

### AI Accelerators

**trenstorrent Wormhole**

”Networked AI”



**Axelera AI Metis AIPU**

”the best AI Processing Unit”

# EXOTICS: RISC-V

## Free and open instruction set architecture (ISA)

### Embedded & Desktop Systems

Octa-Core **AI-CPU** (RV64GCVB)

- 256-bit Vector support
- 5.66 Gflops @ FP64 HPL
- 10.6 Watts
- 2 TOPS for AI

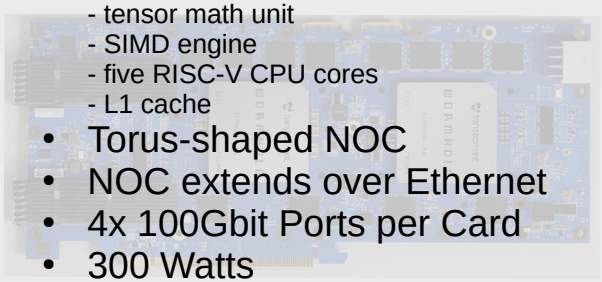
**Milk-V Jupiter**

”RISC-V PC for Everyone”

### AI Accelerators

**Scalable AI building blocks**

- Tensix Cores
  - tensor math unit
  - SIMD engine
  - five RISC-V CPU cores
  - L1 cache
- Torus-shaped NOC
- NOC extends over Ethernet
- 4x 100Gbit Ports per Card
- 300 Watts
- 466 TFlops @ FP8
- 132 TFlops @ FP16
- Supports FP32 output

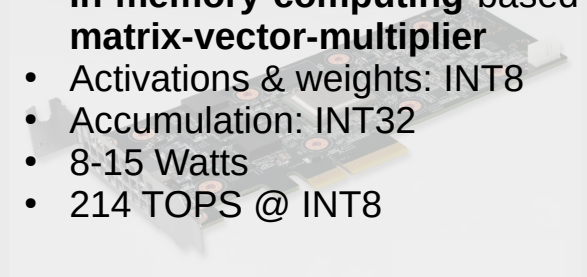


**tenstorrent Wormhole**

”networked AI”

For **AI Edge Application**

- 4 AI cores
- **In-memory-computing**-based **matrix-vector-multiplier**
- Activations & weights: INT8
- Accumulation: INT32
- 8-15 Watts
- 214 TOPS @ INT8



**Axelera AI Metis AIPU**

”the best AI Processing Unit”

# EXOTICS: RISC-V

## HPC hardware made in Europe

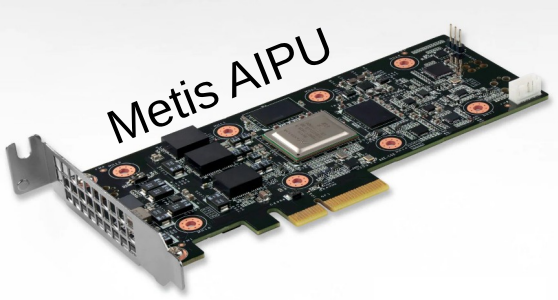
**dare**

**Combining open RISC-V architectures, chiplet technologies, and a co-designed software ecosystem for European digital autonomy in HPC and AI.**

**VEC**
**Vector Accelerator**

openchip
Enabling the Digital World

**AIPU**
**Inference Accelerator**




AXELERA
ARTIFICIAL INTELLIGENCE

**GPP**
**General Purpose CPU**

Codasip

**Applications**
**and Software**

BSC Barcelona Supercomputing Center
Centro Nacional de Supercomputación

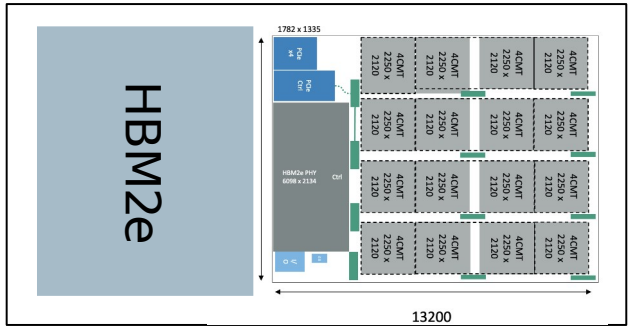
JÜLICH Forschungszentrum
JÜLICH SUPERCOMPUTING CENTRE

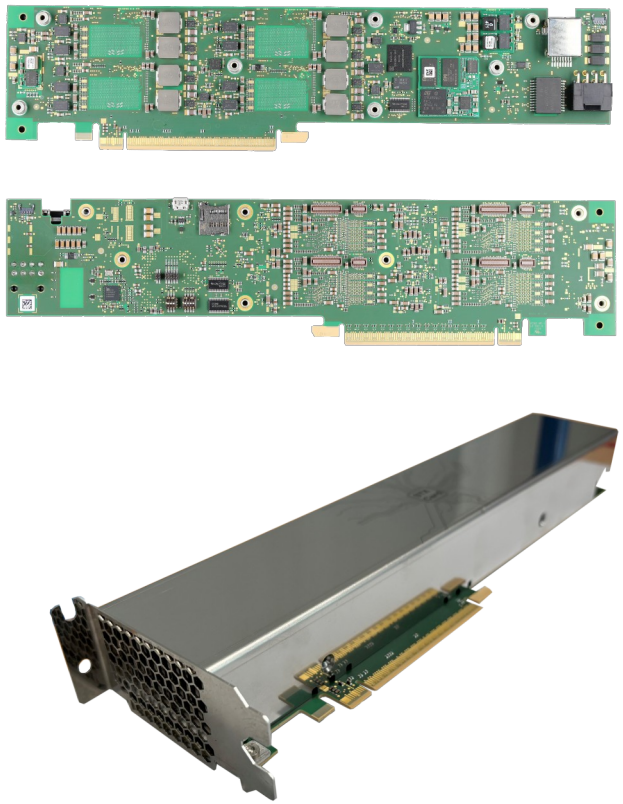
# EXOTICS: STENCIL ACCELERATORS

## Funded through EPI, EuPilot, STXDemo & STXMOD Projects

**STX Stencil Acclerator (Gen 1)**

- **RISC-V compute clusters with stencil extension**
- **High energy efficiency**



Fraunhofer ITWM

Fraunhofer IZM

Fraunhofer IIS

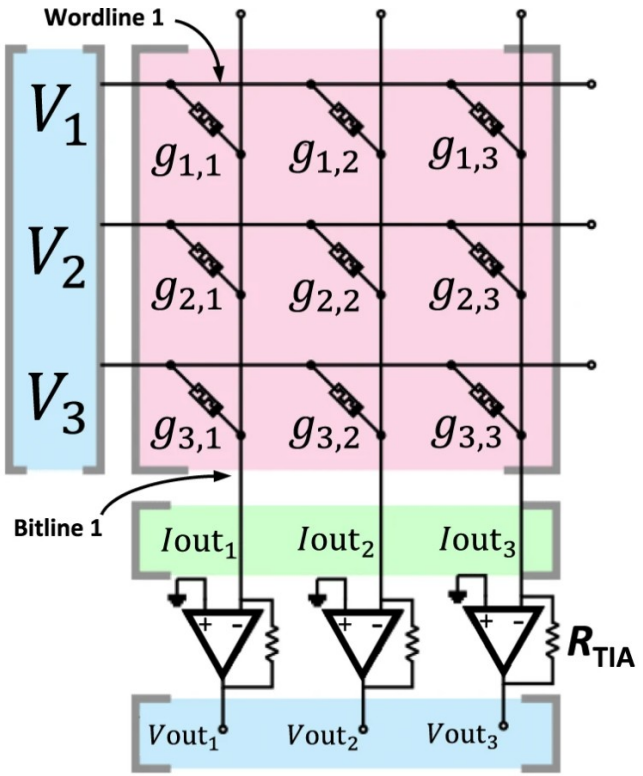
JÜLICH Forschungszentrum

RPTU Rheinland-Pfälzische Technische Universität Kaiserslautern Landau

Zollner

# EXOTICS: NEUROMORPHIC COMPUTING

- Brain inspired computing approach
- Memristor-based architectures
  - Low energy-cost and On-Chip-Memory
- Example of neuromorphic architectures:
  - Spiking (typically digital):
    - Mimic neuron switching behavior
    - Low-precision, pot. faster than biological timescale
    - Examples: SpiNNaker, Intel Loihi
  - Multilevel arrays:
    - Nanosecond timescales with higher precision
    - Only limited by noise and parasitics
- Potential for stochastic computing, true random number generation, and other emerging fields






JÜLICH Forschungszentrum

# EXOTICS: ANALOG PHOTONIC COMPUTING

- Energy efficient devices using photonic circuits for operations
- Laser pulses with different amplitudes and phases encode information
- 16-bit floating point precision
- No cooling required → increased computational density
- Very low power consumption
- Claim: “30 times more energy efficient”

A few demonstrators already being tested in HPC context, for example Q.ANT at JSC in Germany.


[Lightmatter Passage M1000]

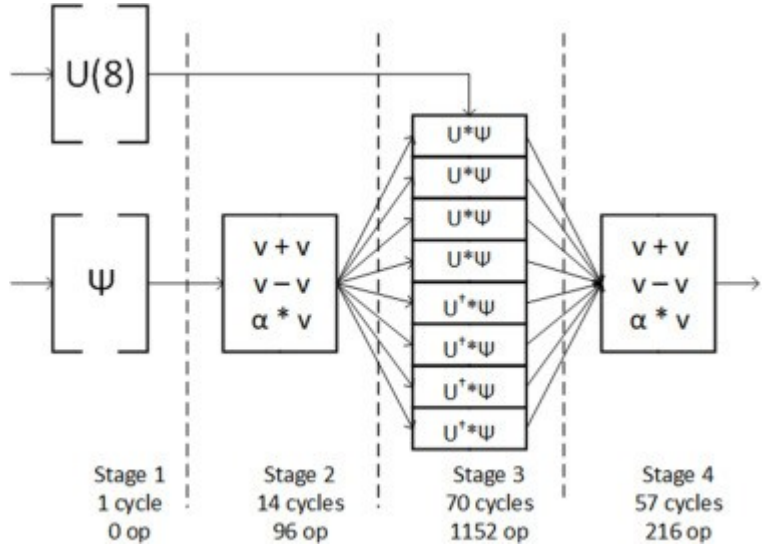

[Q.ANT Native Processing Server (NPS)]

See also talk by Timoteo Lee Tue 17:20

# EXOTICS: FIELD-PROGRAMMABLE GATE ARRAYS (FPGA)

- Custom hardware pipelines with logic gates

  → high throughput for specific algorithms
- Potential for low latency data flow, fine-grained parallelism
- Energy efficiency if tailored properly
- Memory bandwidth often the bottleneck
- FPGA development (HDL, HLS) is more involved than GPU programming
- Scaling to large lattices and multi-node systems remains non-trivial: communication, network interconnect, FPGA cluster design
- Wilson Dirac operator tests date back to 2005 [Callanan et al. 2005]




[Korcyl & Korcyl 2020]

[Intel/Altera Stratix 10]



JÜLICH Forschungszentrum

# QUANTUM COMPUTING

## A very broad landscape

### Several computing paradigms

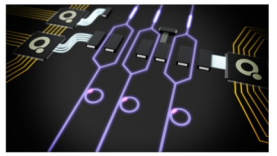

Quantum Annealer


Analog Quantum Computer

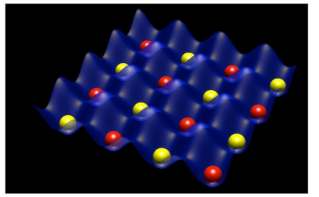
Universal Quantum Computer

### Several hardware technologies

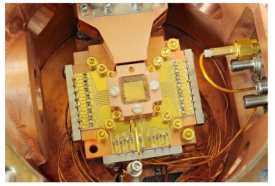
Superconducting



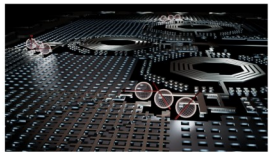
Photonic



Neutral Atoms



Trapped Ions



Quantum Dots

A very active field of research for lattice gauge theories.

→ See plenaries after the break & ask the panel!

# SUMMARY

- Pressure by needs of AI applications changes the vendor roadmaps
  - Mixed precision is nothing new
  - Bootstrapping from low precision shown to work (even integers)
- Energy efficiency will be a concern for scientific HPC applications (can’t afford a backyard nuclear power plant)
  - A window for “specialized” hardware may open again
  - What this means for software is yet unclear

Thank you!
Christoph Lehner
Peter Boyle
Kate Clark
Mathias Wagner
Steve Gottlieb
Tilo Wettig
Maria Lombardo
Jacob Finkenrath
Bartosz Kostrzewa
Marco Garofalo
Aniket Sen
Leon Hostetler
Giovanni Pederiva
Simon Schlepphorst
Travis Whyte
Eric Gregory