JUPITER

2024-10-07 I ANDREAS HERTEN | JÜLICH SUPERCOMPUTING CENTRE

Member of the Helmholtz Association

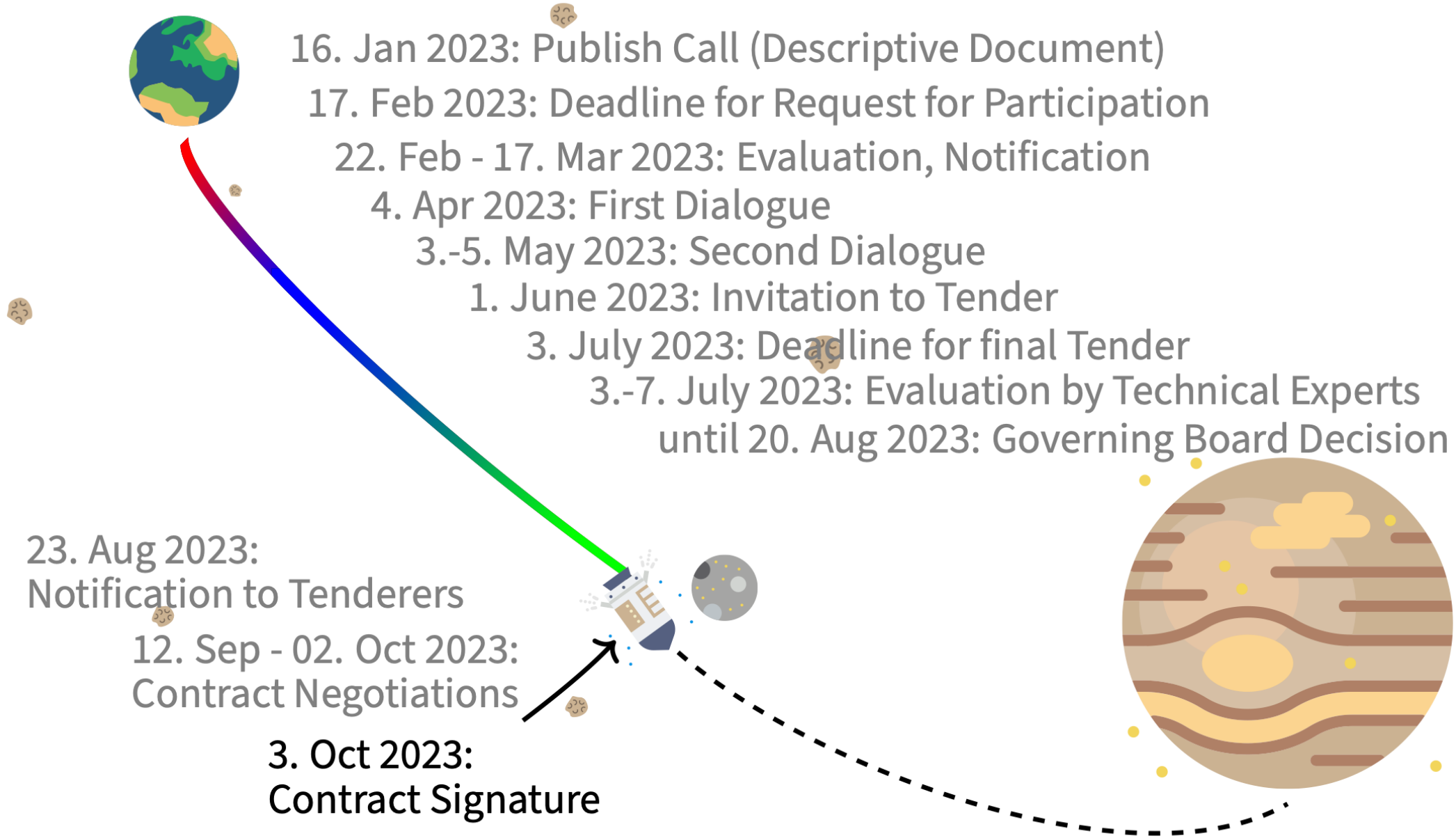
EuroHPC Joint Undertaking

Bundesministerium für Bildung und Forschung

Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen

GCS Gauss Centre for Supercomputing

JÜLICH Forschungszentrum Shaping Change


16. Jan 2023: Publish Call (Descriptive Document)
17. Feb 2023: Deadline for Request for Participation
22. Feb - 17. Mar 2023: Evaluation, Notification
4. Apr 2023: First Dialogue
3.-5. May 2023: Second Dialogue
1. June 2023: Invitation to Tender
3. July 2023: Deadline for final Tender
3.-7. July 2023: Evaluation by Technical Experts
until 20. Aug 2023: Governing Board Decision
23. Aug 2023:
Notification to Tenderers
12. Sep - 02. Oct 2023:
Contract Negotiations
3. Oct 2023:
Contract Signature

JUPITER

- 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 MW Linpack Power Consumption
- Direct Liquid Cooled to enable heat-reuse

BullSequana XH3000: DLC

ParTec MODULAR SUPERCOMPUTING EVIDEN NVIDIA SIPEARL

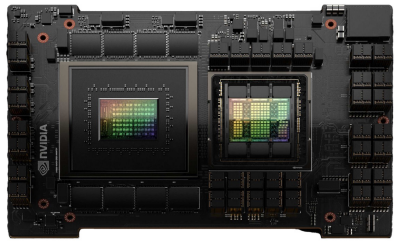
JÜLICH Forschungszentrum

SYSTEM DETAILS

JUPITER MODULES

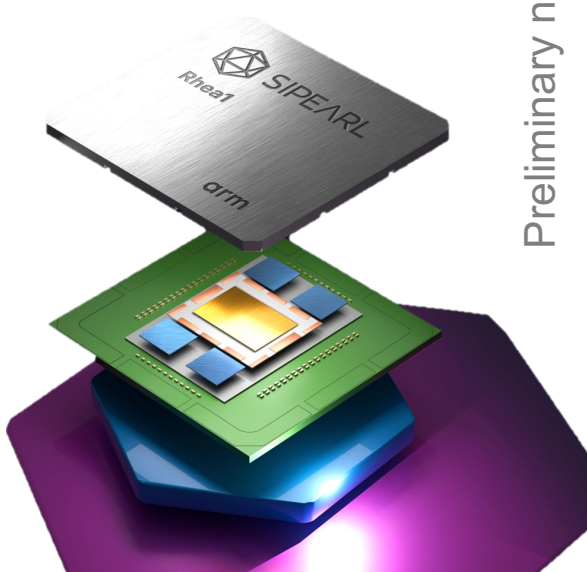
JUPITER Booster

- ~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



JUPITER – BOOSTER COMPUTE NODE ARCHITECTURE

Node Specs

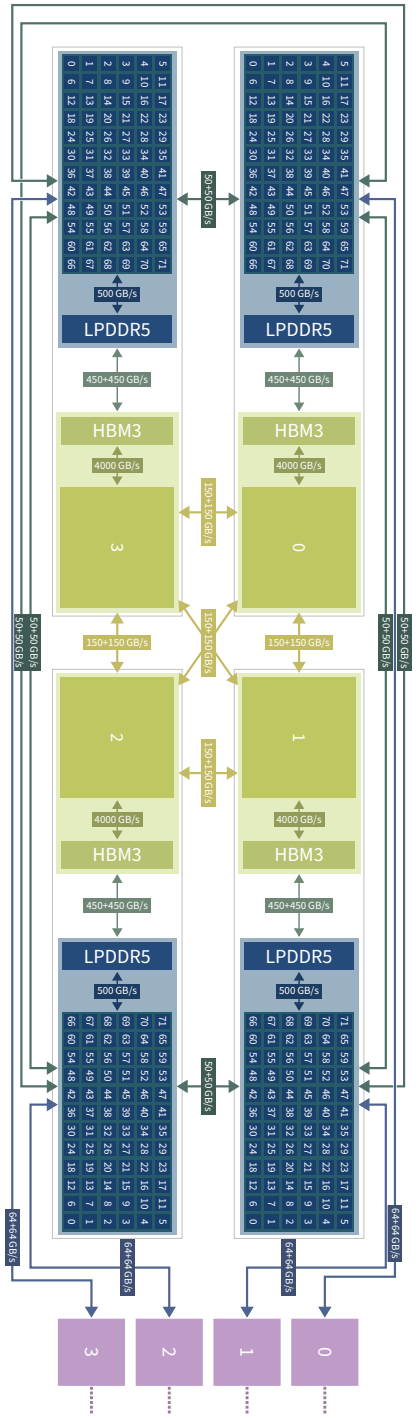
- 4× NVIDIA Grace-Hopper in SXM5 Board (4× 680W)
- 4× NVIDIA InfiniBand NDR200
- 480 GB LPDDR5X / 360 GB HBM3 (usable)
- NVLink 4
  - GPU-GPU 150 GB/s per dir, CPU-GPU 450 GB/s per dir, CPU-CPU 100 GB/s per dir
- CG4 Motherboard (4× CG1 GH module + 4× CX7 HCA assembly)
  - All NVIDIA, except the BMC

CPU Specs

- **ARM Neoverse V2**
  - SVE2/NEON (4x 128 bit vector op)
- 72 cores @ ~2.4GHz (~3.2 GHz turbo)
- 120 GB LPDDR5X (8 channels)
  - *≥450 GB/s*
  - *~150 ns latency*

GPU Specs

- **H100**
  - ~50 TFLOP/s (HPL single GPU)
- 96 GB HBM3
  - 4000 GB/s
  - ~450 ns latency

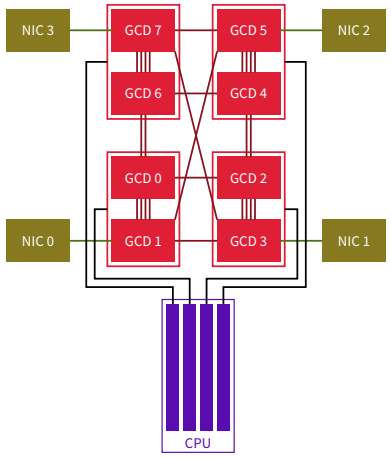


NODE COMPARISON

- JUWELS Booster: 2× CPU, 4× GPU, 4× IB
- JUPITER Booster: 4× CPU+GPU, 4× IB
- Frontier: 1× CPU, 4×(2× GPU), 4× Slingshot
- Aurora: 2× CPU, 6× GPU, 8× Slingshot
- El Capitan: 4× APU

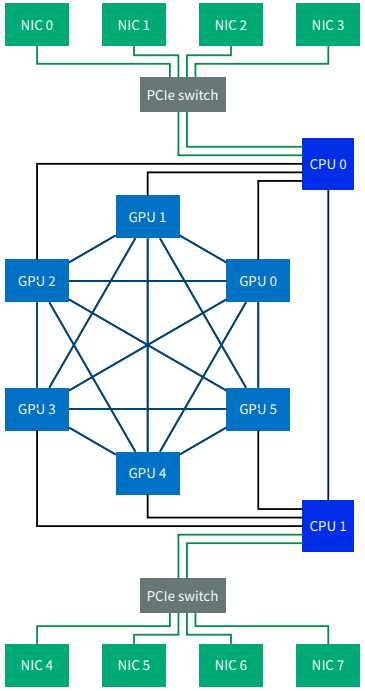
ORNL

Frontier



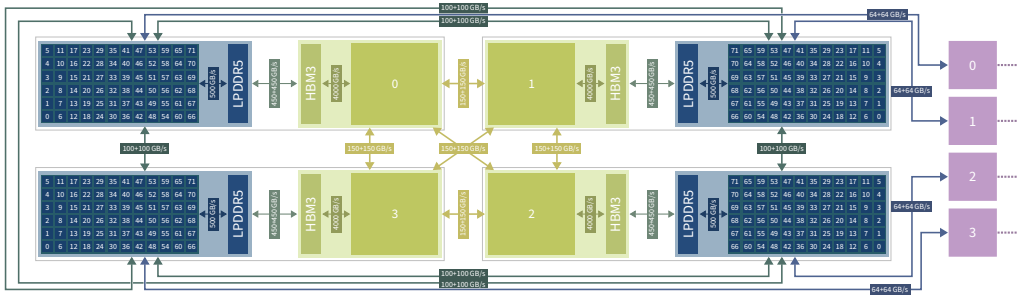
Argonne

Aurora

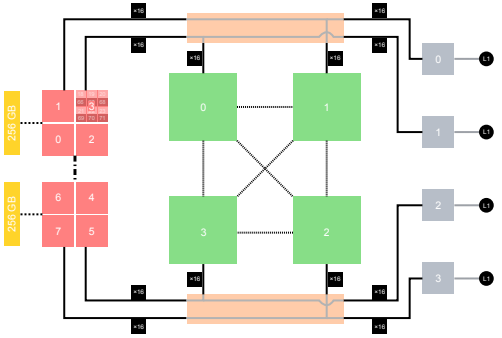


JSC

JUPITER Booster



JUWELS Booster



JÜLICH Forschungszentrum

JUPITER – CLUSTER COMPUTE NODE ARCHITECTURE

Node Specs

- 2× SiPearl Rhea1
- 1× NVIDIA InfiniBand NDR200
- 512 GB DDR5 (36 nodes with 1024 GB)
- CCIX

CPU Specs

- ARM Neoverse V1 Zeus
  - 2 x 256 SVE per core
- 2.5 GHz (~3.0 GHz turbo)
- 64 GB HBM2e per Socket
  - 1.64 TB/s
- 256 GB DDR5
- PCIe Gen5

JUPITER – STORAGE (SCRATCH)

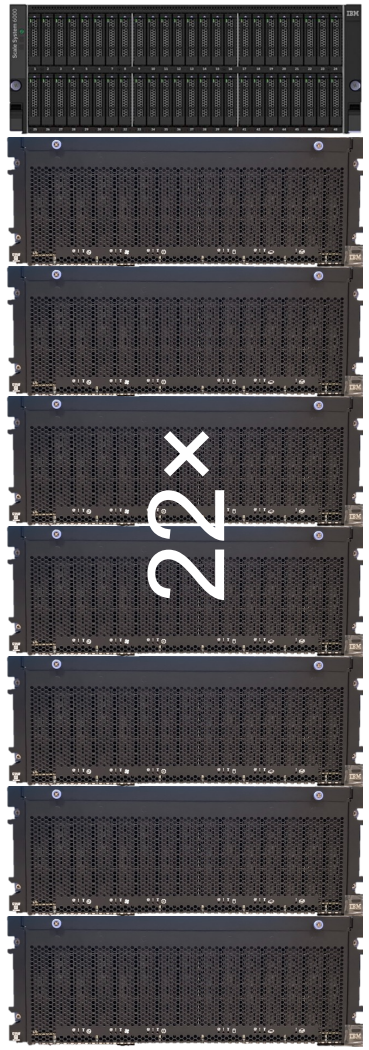
- Gross Capacity: 29 PB; Net Capacity: 21 PB
- Bandwidth: 2.1 TB/s Write, 3.1 TB/s Read
- 20× IBM SSS6000 Building Blocks (40 servers)
  - 2× NDR400 per server
  - 48× 30 TB NVMe drives per block
  - IBM Storage Scale (aka Spectrum Scale/GPFS)
- Manager and Datamover Nodes
- Exclusive for JUPITER
  - Integrated into InfiniBand fabric

Preliminary numbers, might change during installation

JUPITER – STORAGE (EXASTORE)

In kind contribution from JSC, not part of the JUPITER procurement

- Gross Capacity: 308 PB; Net Capacity: 210 PB
- Bandwidth: 1.1 TB/s Write, 1.4 TB/s Read
- 22× IBM SSS6000 Building Blocks (44 servers)
  - 2× NDR200 per server
  - 7× JBOD enclosures, each with 91x 22 TB Spinning Disks per block
  - IBM Storage Scale (aka Spectrum Scale/GPFS)
- Manager and Datamover Nodes
- Exclusive for JUPITER
  - Integrated into InfiniBand fabric



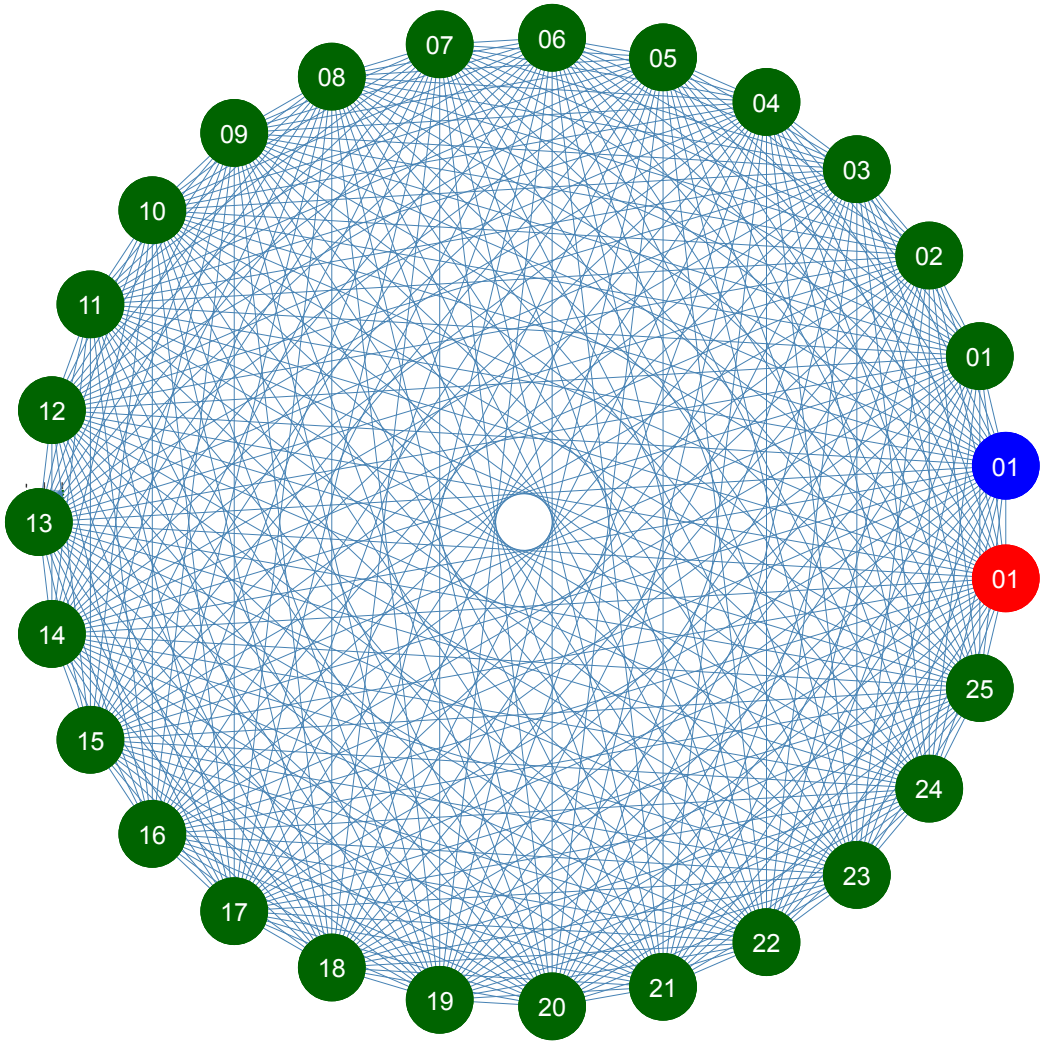
Preliminary numbers, might change during installation

JUPITER – INTERCONNECT

One Network to Rule Them All

EVIDEN
an atos business

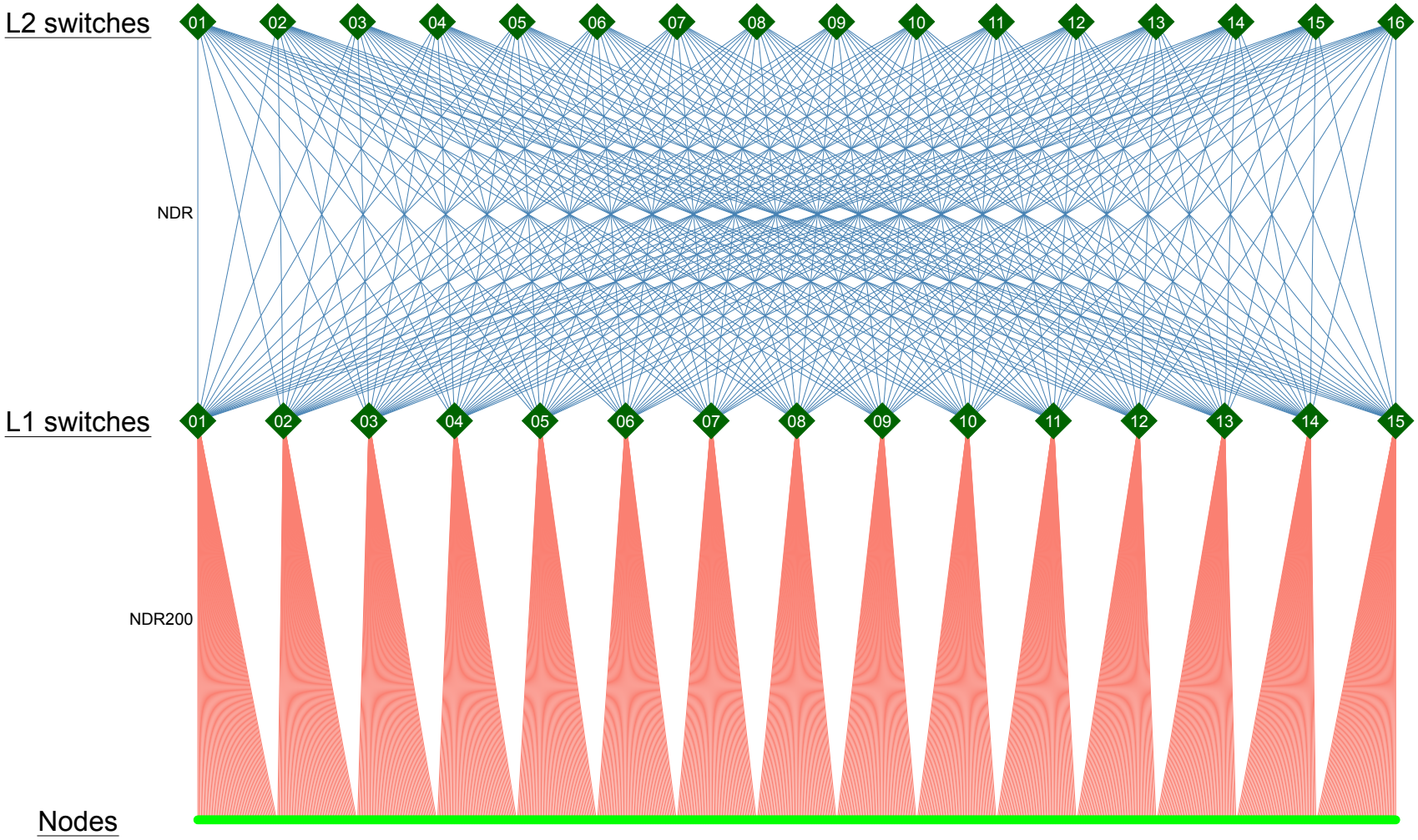
NVIDIA



JÜLICH
Forschungszentrum

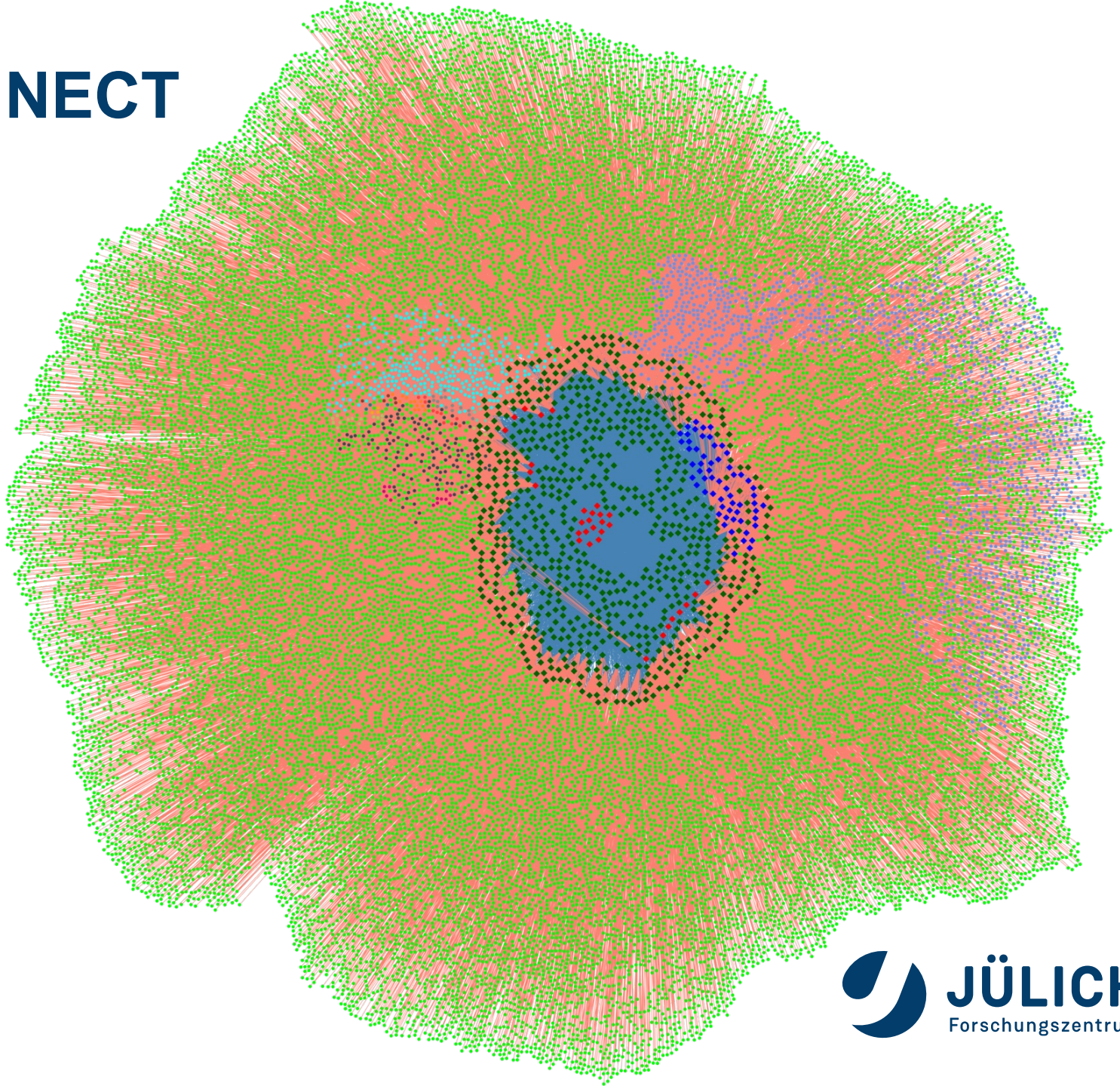
JUPITER – INTERCONNECT

One Network to Rule Them All



JUPITER – INTERCONNECT

One Network to Rule Them All



- Cluster module switch
- Booster module switch
- Booster node
- Flash storage node
- Misc node
- NDR link
- Admin module switch
- Cluster node
- Gateway node
- Storage node
- NDR200 link

STATUS

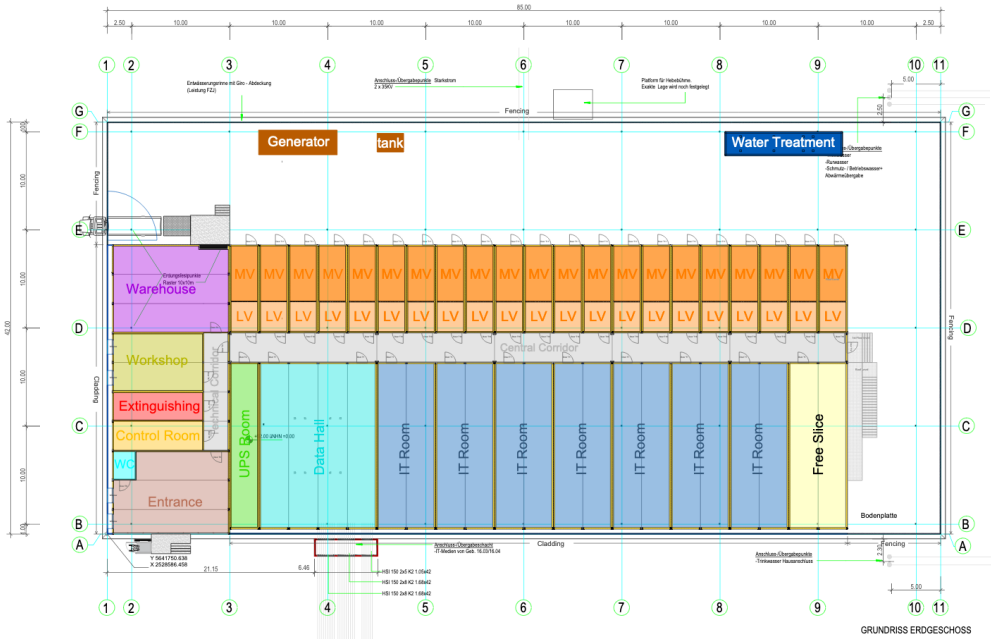
POWER TRANSFORMER SUBSTATION AND LINES

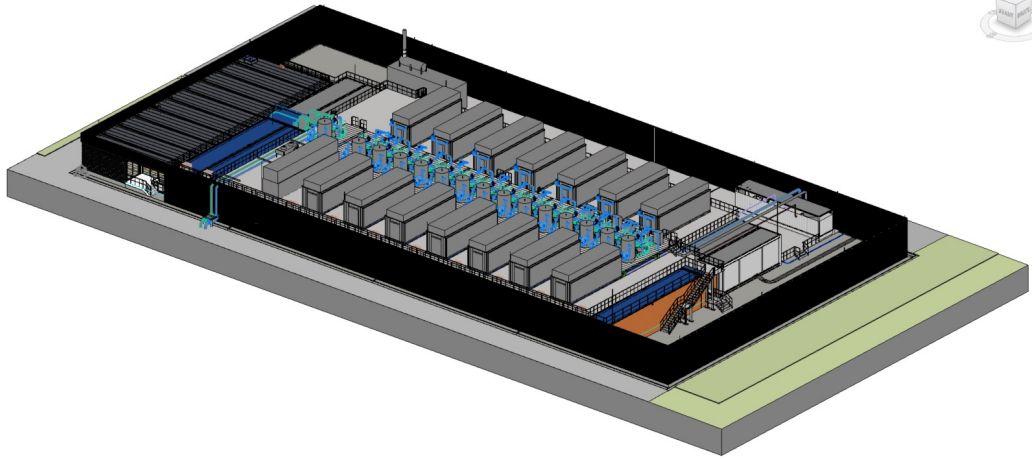
Upgrade of transformers 110 kV / 35 kV from 2 x 40 MVA to 2 x 60-80 MVA and upgrade 110kV power line

MODULAR DATA CENTER FOR JUPITER

EVIDEN

- Vendor: **Eviden**
- Area: ~2300 m²
- 1× Datahall (storage, management)
- 7× IT modules (20 racks per module)
- UPS, generator
- Entrance area
- Workshop, warehouse
- 15 × 2.5 MW power stations





JÜLICH Forschungszentrum

MODULAR DATA CENTER FOR JUPITER

EVIDEN

CONCRETE FOUNDATION

CONCRETE FOUNDATION

Construction of concrete slab 85 m x 42 m x 0.5 m


REMONDIS

MDC SHIPMENT START

10./11.9.2024

JUPITER INSTALLATION IN ANGERS (EVIDEN FACTORY)

- 10 XH3000 racks, 480 nodes
- Hardware tests and benchmark preparation
- JUPITER Management Server preparation
- **Afterwards**
  - Integration into containers
  - Shipment to Jülich
  - “Plug in and run”

APPLICATIONS

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	✓			✓	

GH200 TEST NODES

- GH200 Prototype
- 2× Grace-Hopper superchips
  - 1 Grace CPU (72 cores), 480 GB LPDDR5X RAM
  - 1 H100 GPU
  - TDP 700-1000 W
- Slightly different variant compared to JUPITER node design

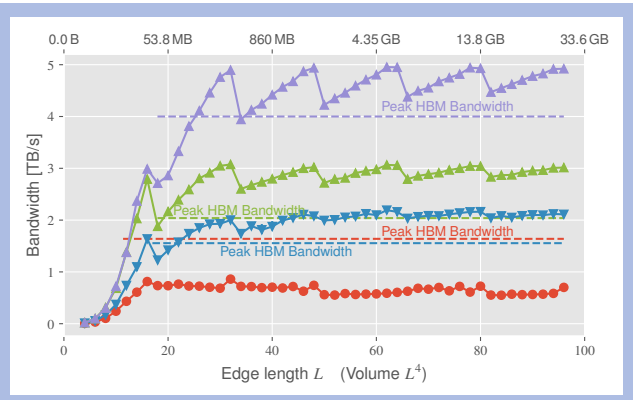
ENABLEMENT: JEDI, JUREAP

- 🖌️ JEDI: JUPITER test system
  - 48 nodes; JUPITER design

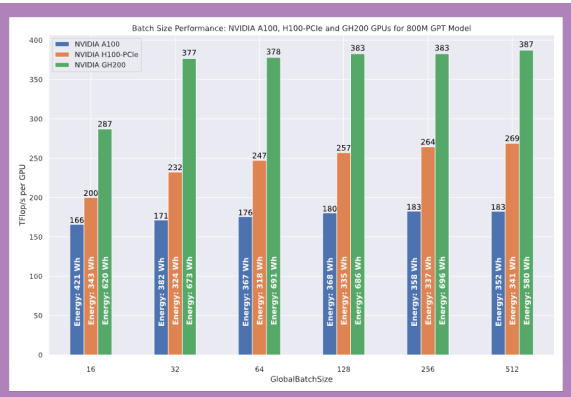
🎉 Top 1 Green500!

- Usage
  - System management preparations
  - Application porting
  - JUREAP; Research and Early Access Program

FIRST GPU EXPERIENCES *(H100)*



LQCD benchmark: Great mem utilization



LLM benchmark: >2× vs A100

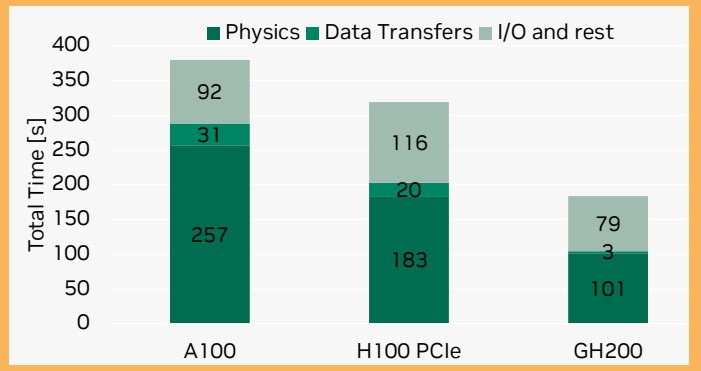
ChASE: >2× vs. A100 across all solvers

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

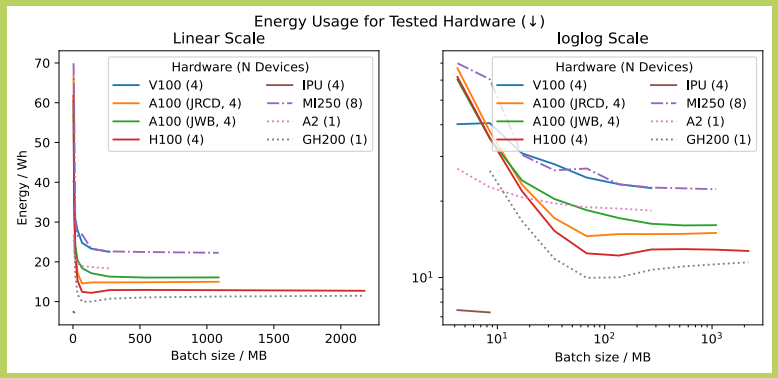
nekRS: 2.1× vs. A100 for RBC benchmark

Arbor: 1.97× vs. A100 for Busyring benchmark

JUQCS: 2.6× vs. A100 for 31 Qubits



MPTRAC: >2× vs A100



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

JUREAP

Seeding Exascale in Europe!

jureap@fz-juelich.de • https://www.fz-juelich.de/en/ias/jsc/jupiter/jureap

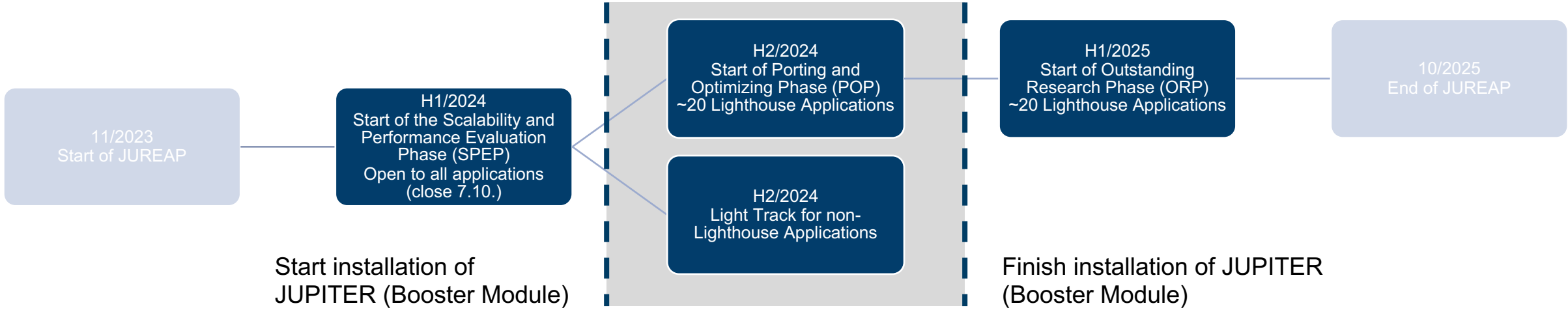
JUPITER Research and Early Access Program

JÜLICH Forschungszentrum

OVERVIEW

Timeline

Current state:
- GCS Exascale Pioneer Call (close 28.10.)
- JUREAP Exascale Certificates
- Focus Topics, Big Days



Phase 1: Scalability and Performance Evaluation Phase (SPEP)
Phase 2: Porting and Optimizing Phase (POP)
Phase 3: Outstanding Research Phase (ORP)

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

JOINING FORCES

Ministerium für
Kultur und Wissenschaft
des Landes Nordrhein-Westfalen

fz-juelich.de/jupiter