

# **JUPITER - FROM LAUNCH TO LANDING**

**Accelerated Data Analytics and Computing Symposium (ADAC17)** 

2025-09-03 | JÜLICH SUPERCOMPUTING CENTRE | BENEDIKT VON ST. VIETH









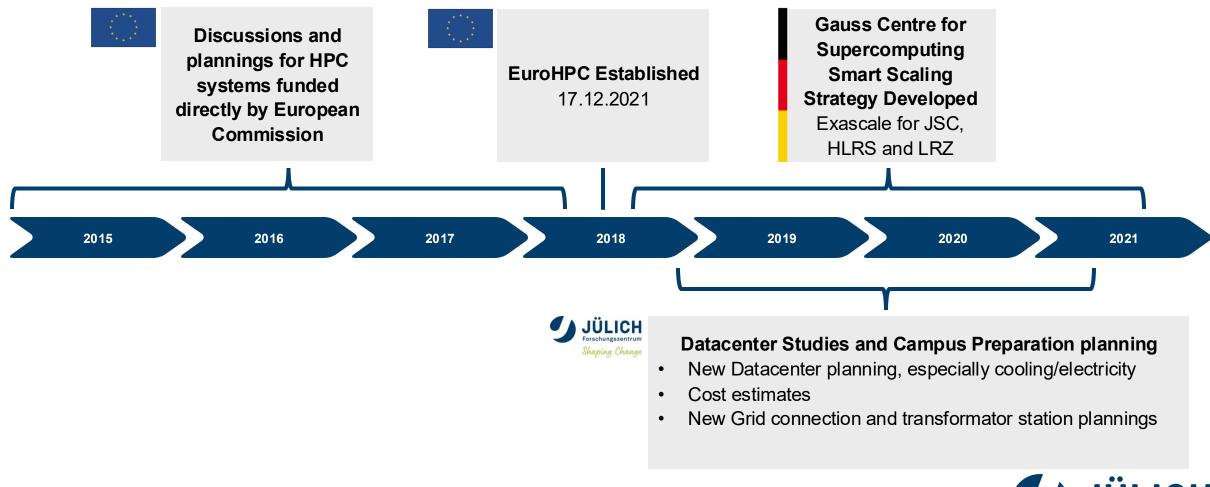




# A LONG TIME AGO

BECOMING A HOSTING ENTITY AND PROCURING JUPITER

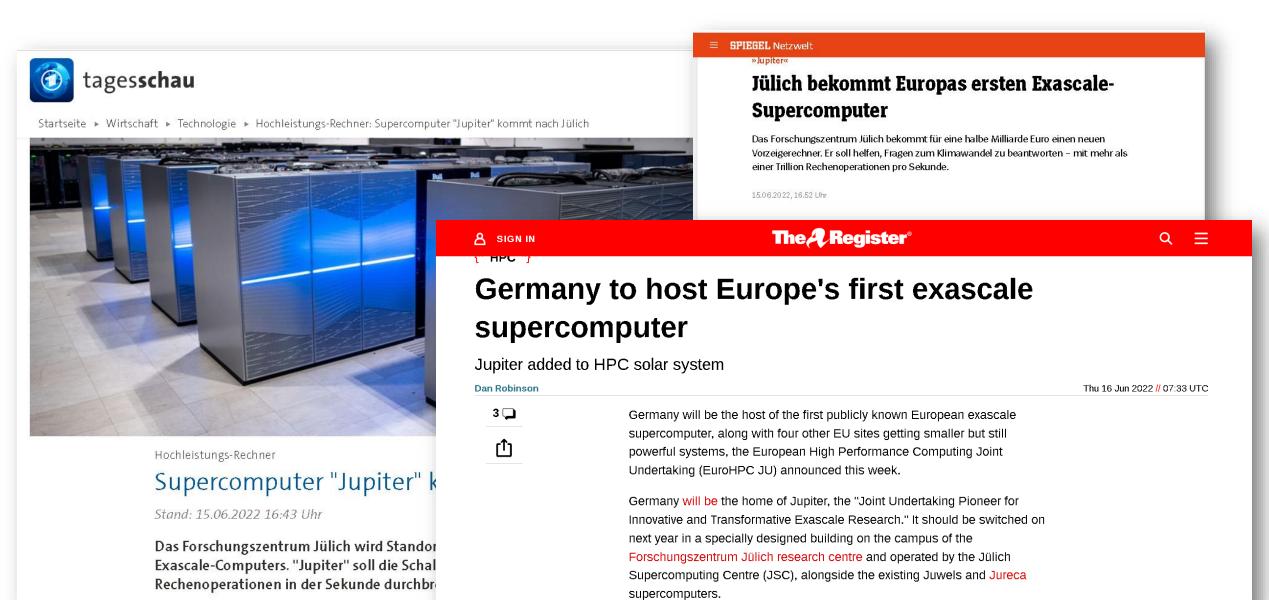
# **EUROHPC AND SMART SCALING@GCS**





# JUPITER - HOSTING ENTITY DECISION

15.06.2022



# LAYING THE FOUNDATION FOR TAKEOFF (2022)

- Numerous calls/meetings/discussions
- Datacenter to Modular HPC Datacenter decision
- Preparation of the Descriptive Document to start procurement

Hosting Entity
Decision
15.06.2022

Member of the Helmholtz Association

Page 5

October November December

November December

December

Hosting Agreement
14.12.2022

Forschungszentrum

# PREPARING FOR LAUNCH

- Mission planning
  - Preparing descriptions, conditions, requirements, evaluation
  - Regular meetings
  - Started already early in year
  - Location: Earth
- Target: JUPITER
  - Booster
  - Cluster
  - Storage
  - (Machine Hall)







# Ready for take off





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

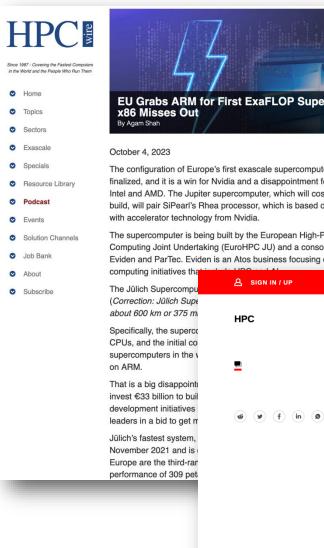
> 12. Sep - 02. Oct 2023: Contract Negotiations

> > 3. Oct 2023: Contract Signature

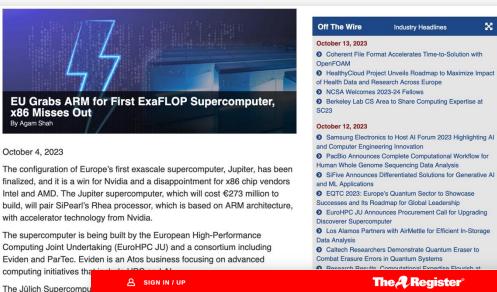


# JUPITER CONTRACT ANNOUNCEMENT

3.10.2023



Member of the Helmholtz Asso





# Atos subsidiary Eviden scores contract win in **Europe's first exascale system**

×

\$526M Jupiter set to rule EU's tech orbit by 2024

Wed 4 Oct 2023 // 16:45 UTC Dan Robinson

The EU's supercomputing initiative, the European High Performance Computing Joint Undertaking (EuroHPC JU), has awarded a procurement contract for Europe's first exascale system, with installation due to start in early 2024.

Known as Jupiter (Joint Undertaking Pioneer for Innovative and Transformative Exascale Research), the system was announced last year followed by a call for tender in January of this year.

EuroHPC JU said the procurement contract for Jupiter was awarded to a consortium comprising of Eviden, the professional services side of French IT giant Atos, and ParTec, a German supercomputing hardware company.

The project is expected to have a total cost of €273 million (\$287 million) covering the build, delivery, installation, and maintenance of Jupiter, according to the EuroHPC.

However, Eviden put the overall project cost at €500 million (\$526 million), saying that this is the figure for the entire project, including the system manufacturing and its

me details are emerging on Europe's first exascale system, codenamed "Jupiter" and to be installed at the ich Supercomputing Center in Germany in 2024. There has been a lot of speculation about what Jupiter ll include for its compute engines and networking and who will build and maintain the system. We now ow some of this and can infer some more from the statements that were made by the organizations ticipating in the Jupiter effort.

une 2022, the Forschungszentrum Jülich in Germany, which has played host to many supercomputers ice it was founded in 1987, was chosen to host the first of three European exascale-class puters to be funded through the EuroHPC Joint Undertaking and through the European national d state governments countries who are essentially paying to make sure these HPC and AI clusters are ere they want them. With Germany having the largest economy in Europe and being a heavy user of thanks to its manufacturing focus, Jülich was the obvious place to park the first machine in Europe to ak the exaflops barrier.

at barrier is as much an economic one as it is a technical one. The six-year budget for Jupiter weighs in at 00 million, which is around \$526.1 million at current exchange rates between the US dollar and the ropean euro. That is in the same ballpark price as what the "Frontier" exascale machine at Oak Ridge tional Laboratory and the "El Capitan" machine that is being installed right now at Lawrence Livermore onal Laboratory – both of which are based on a combination of AMD CPUs and GPUs and Hewlett ckard Enterprise's Slingshot variant of Ethernet with HPE as the prime contractor.

erybody knows that Jupiter was going to use SiPearl's first generation Arm processor based on the erse "Zeus" VI core from Arm Ltd, which is codenamed "Rhea" by SiPearl and which is appropriate



# JUPITER ARCHITECTURE





# **JUPITER MODULES**

# **JUPITER Booster (~1ExaFLOP/s)**

~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)



- ~15 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



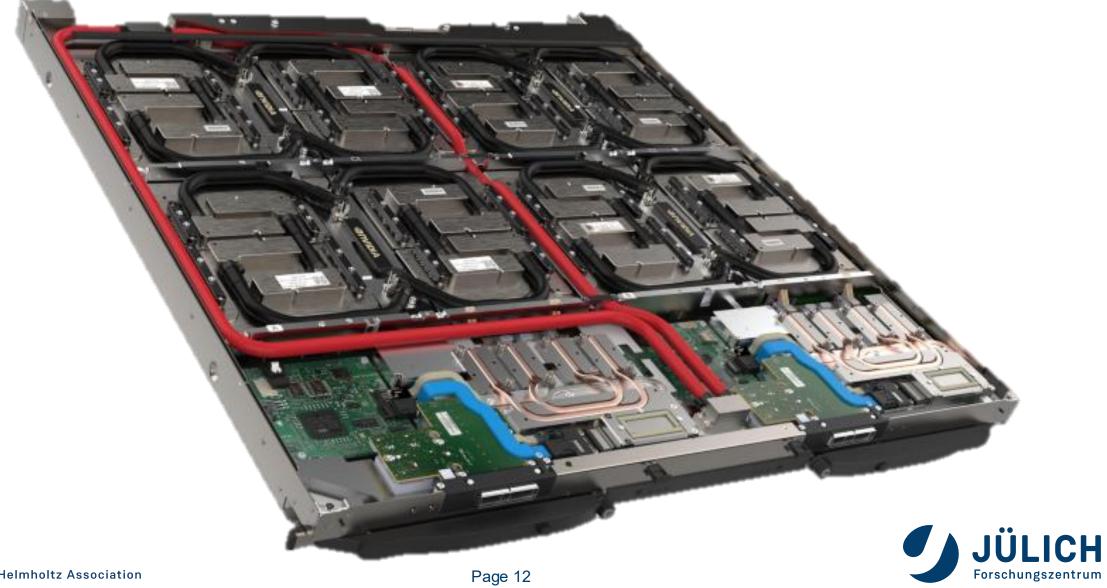
ExaFLASH: 29PB (raw) NVMe, IBM SSS6000

ExaSTORE: 308PB (raw) HDD\*, IBM SSS6000

**ExaTAPE: 370PB Tape\*, LTO9** 



# JUPITER BOOSTER BLADE OVERVIEW



Member of the Helmholtz Association

# JUPITER - BOOSTER COMPUTE NODE ARCHITECTURE

• 4× NVIDIA Grace-Hopper in SXM5 Board (4× 680W)

**Node Specs** 

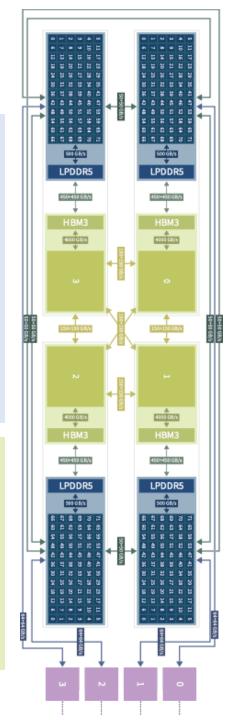
**GPU Specs** 

- 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
- ARM Neoverse V2

**CPU Specs** 

- 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

- H100
  - 47.5 TFLOP/s (HPL Rmax single GPU)
- 90 GB HBM3
  - ≥3600 GB/s
  - ~450 ns latency

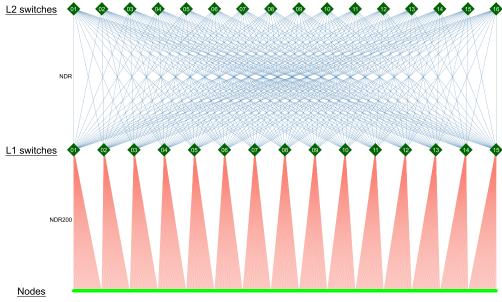


# JUPITER - INTERCONNECT

## One Network to Rule Them All

- NVIDIA Mellanox InfiniBand NDR/NDR200
  - NVIDIA Quantum-2 switches
  - NVIDIA Connect-X7 HCAs
- Dragonfly+ topology
  - 27 Dragonfly groups
  - Within each group: full fat tree
- 51000 links, 102000 logical ports, 25400 endpoints, 867 switches
- Adaptive Routing
- In-network processing on switch level (SHARPv3), tentatively









# **SYSTEM MANAGEMENT**

# JUPITER MANAGEMENT STACK

# 3 main pillars/actors

# **SMC** xScale

Core part of the stack. Vast majority of components come from here.

Developed by Eviden

Heavily based on open source and cloud technologies



# JUPITER MANAGEMENT STACK

# 3 main pillars/actors

SMC xScale	ParaStation
Core part of the stack. Vast majority of components come from here.	Enhancement of the core
Developed by Eviden	Developed by ParTec
Heavily based on open source and cloud technologies	Integrates ParTec tools in SMCx to streamline their support workflows



# JUPITER MANAGEMENT STACK



# 3 main pillars/actors

SMC xScale	ParaStation	xOPS
Core part of the stack. Vast majority of components come from here.	Enhancement of the core	Enhancement of the core
Developed by Eviden	Developed by ParTec	Developed by JSC
Heavily based on open source and cloud technologies	Integrates ParTec tools in SMCx to streamline their support workflows	Extensive set of Ansible roles for HPC, targeting JSC's requirements and needs



# JUPITER MANAGEMENT STACK – KEY AREAS



Forschungszentrum

	Technology	Challenges	Provider
Operating System	Linux	Security Performance Stability HW support	
Management Storage	Ceph	Multi-use Scalable Performance	IBM
Management Plane	Kubernetes	Scalable 0 downtime Flexible Open	kubernetes
Configuration Management	Ansible	Standard Easy to extend Open	
Boot Image(s) Management	ImageBuilder	ARM / x86 Tracking support Integration	
Container(s)	UBI Universal Binary Images	Standard Security Consistency	
TA JULICH			

# JUPITER MANAGEMENT STACK – KEY AREAS



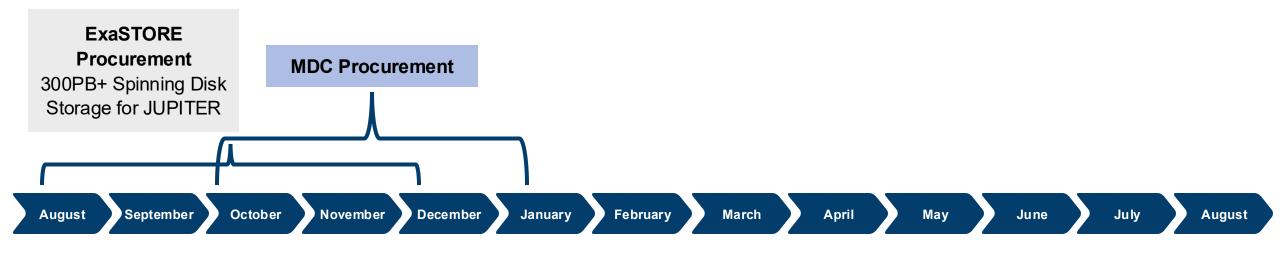
Forschungszentrum

	Technology	Challenges	Provider
Resource Manager	Slurm	Scalable Known API	SCHED ParaStation The Slurm Company ParaSturm
Parallel Storage	Storage Scale System (GPFS)	Performance Scalable Data security	IBM
MPI Runtime	Message Passing Interface	Stable GPU-support Performance Bug-free	ParaStation MPI Open MPI
GPU Support	CUDA HPC SDK	Memory Performance management Integration	NVIDIA.
Monitoring & Logging	Prometheus + Thanos Syslog + Fluentd	Usable Scalable Handle data storm	Prometheus Thanos fluentd
Reference Database	Data Center Information Management	Automation API Coherent	netbox

# THE PREPARATION PHASE

# 2023/2024 - PREPARING CAMPUS, SLAB, STORAGE, JEDI

... it is not only about waiting for JUPITER





# MODULAR DATA CENTER FOR JUPITER



Vendor: Eviden

Area: ~2300m²

• 1x Datahall (Storage, Management)

• 7x IT Modules (20 Racks per module)

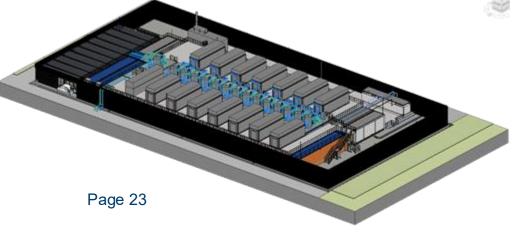
UPS, Generator

Entrance area

Workshop, Warehouse

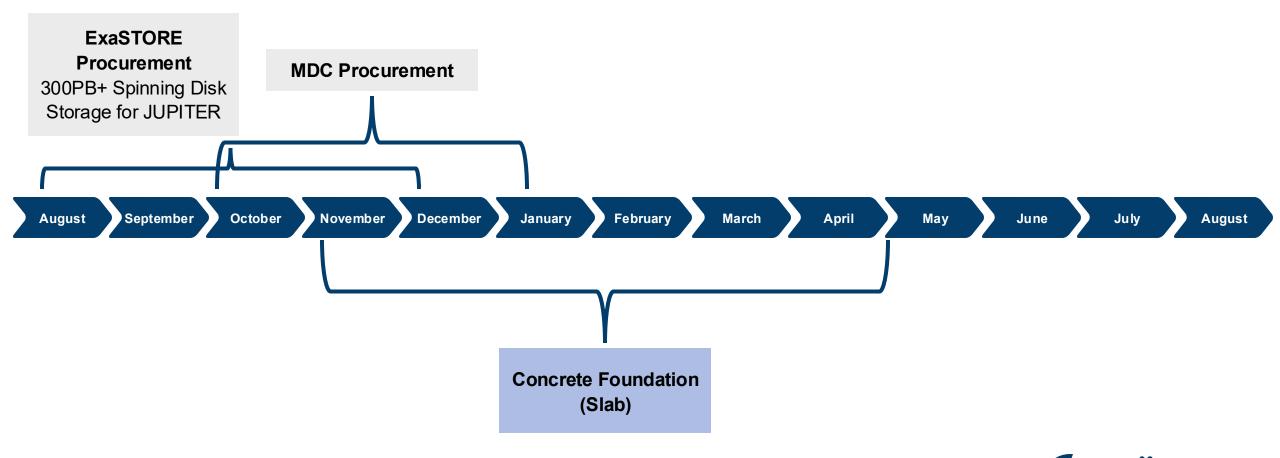
15x 2,5 Megawatt Power Stations





# 2023/2024 - PREPARING CAMPUS, SLAB, STORAGE, JEDI

... it is not only about waiting for JUPITER





# **CONCRETE FOUNDATION**





# **CONCRETE FOUNDATION**

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







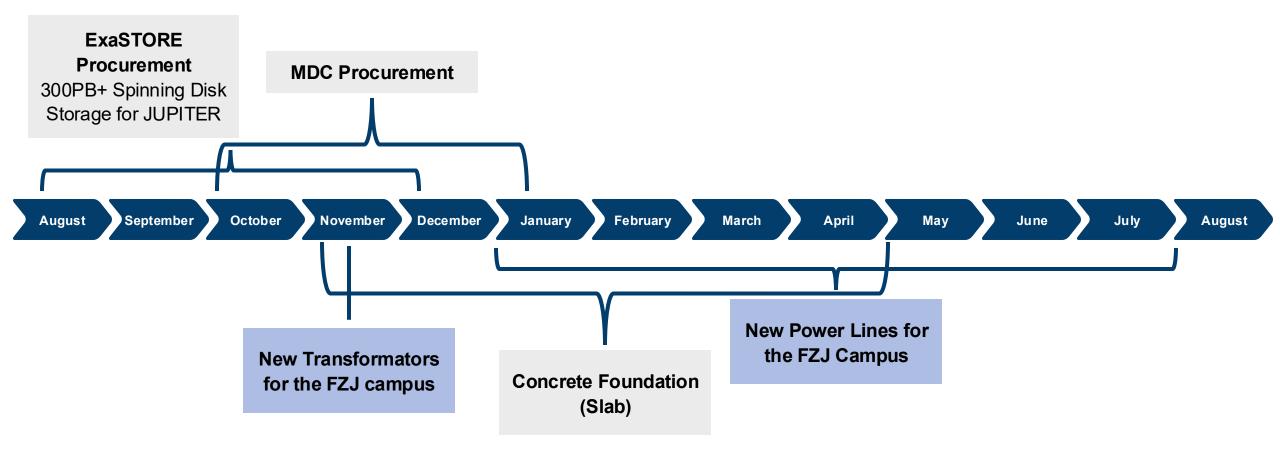






# 2023/2024 - PREPARING CAMPUS, SLAB, STORAGE, JEDI

... it is not only about waiting for JUPITER





# 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







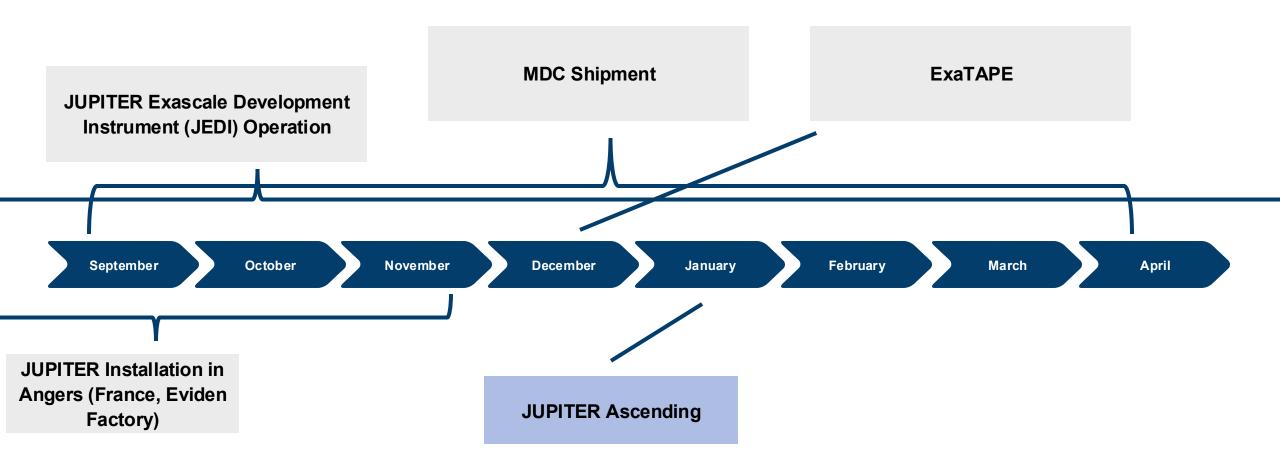




# RECENT DEVELOPMENTS

# 2024/2025 - MDC AND JUPITER INSTALLATION

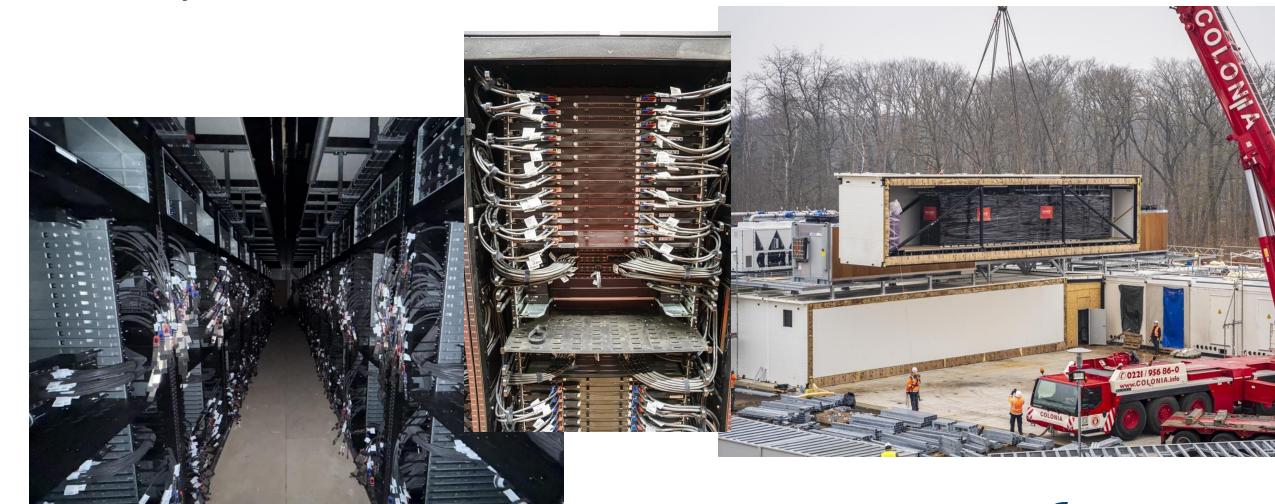
Always in motion is the future...



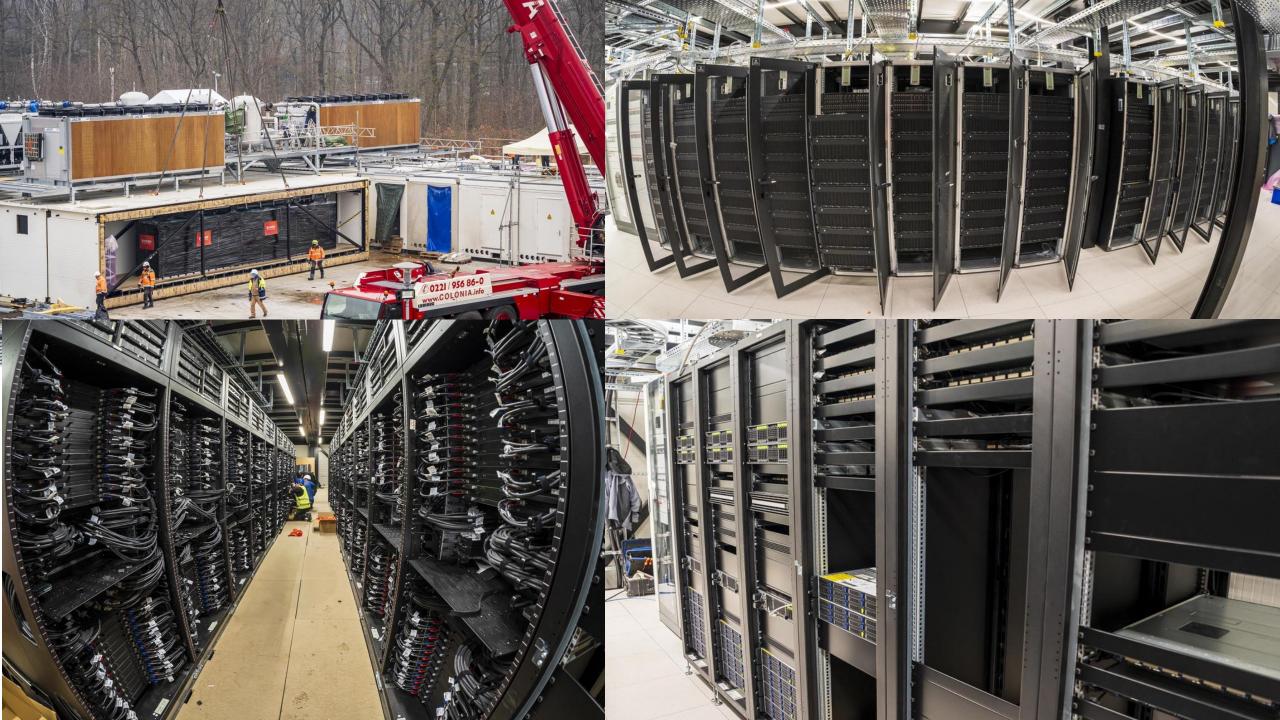


# **JUPITER ASCENDING**

**Since January 2025** 

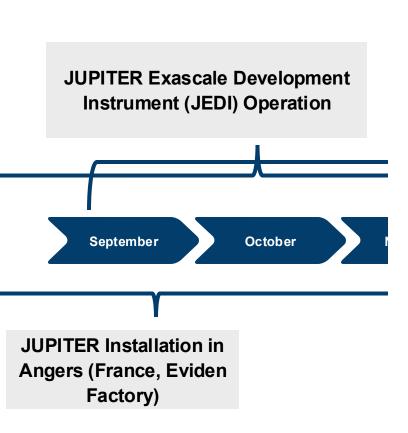






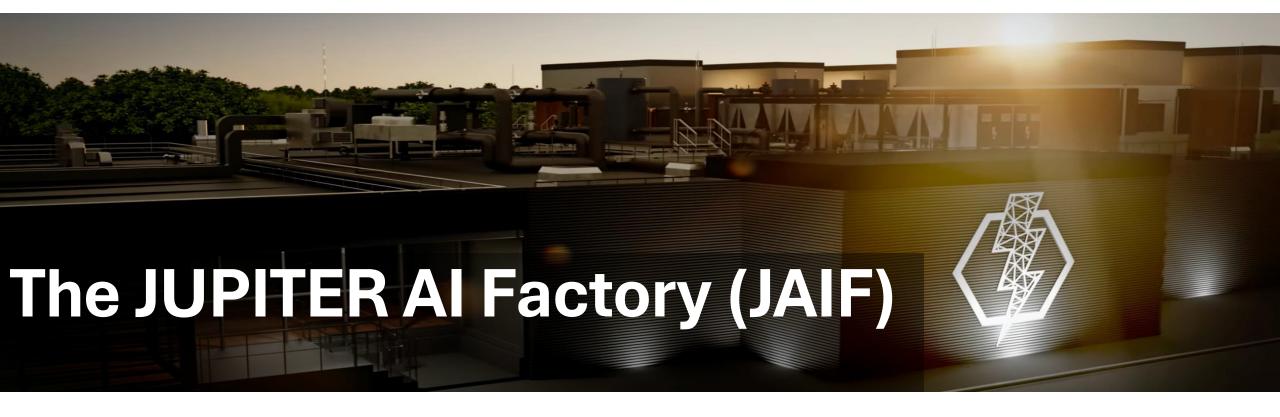
# 2024/2025 - MDC AND JUPITER INSTALLATION

Always in motion is the future.















Fraunhofer hessian.Al



**Associated** partners:

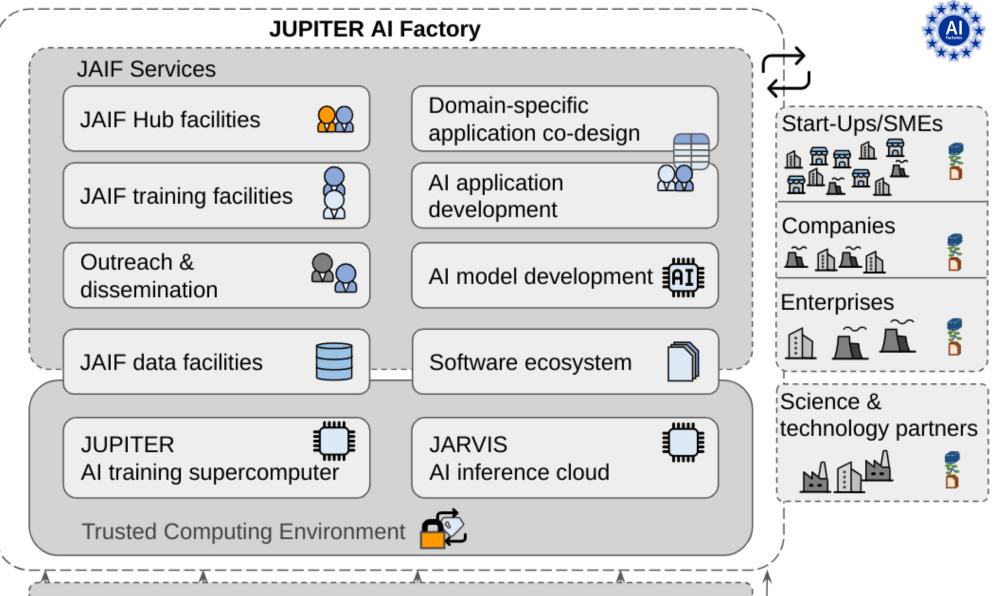








## THE **JAIF** ONE STOP SHOP



National and European networks and associates

Public data sources

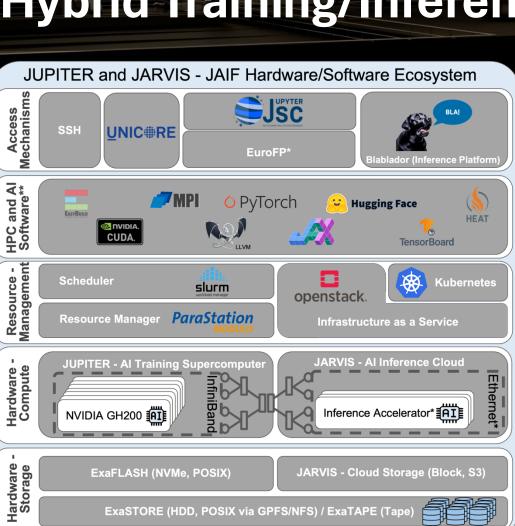
Proprietary data sources

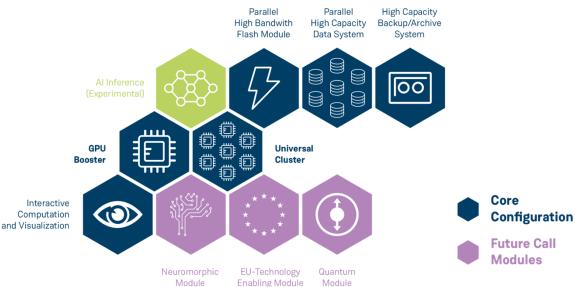
Public AI models and ML tools

University talent pool

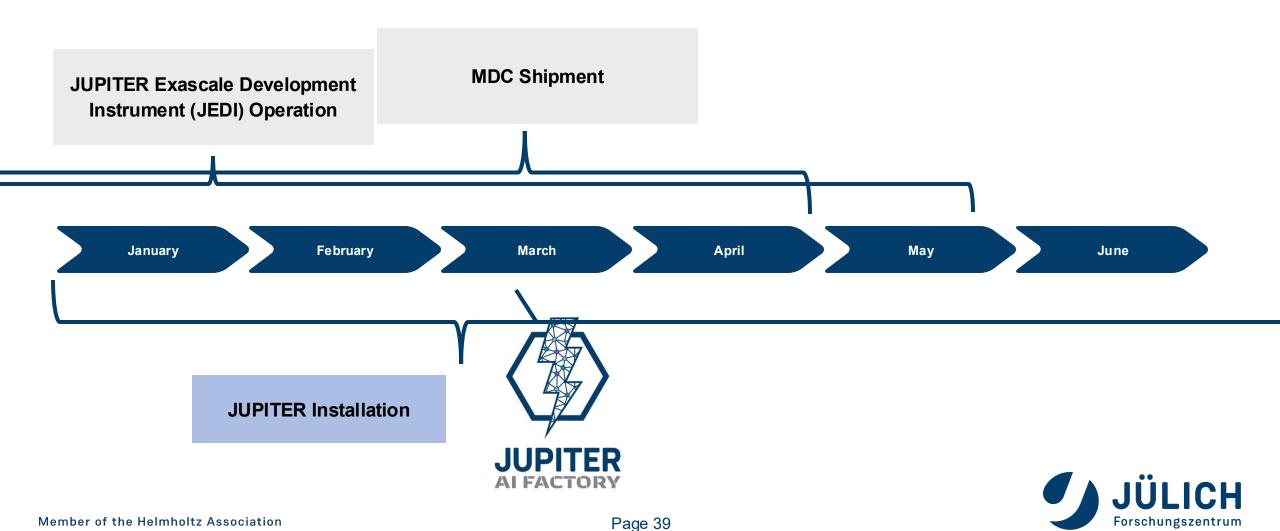
Company staff

# Modular JUPITER Hybrid Training/Inference AI System





#### 2024/2025 - MDC AND JUPITER INSTALLATION



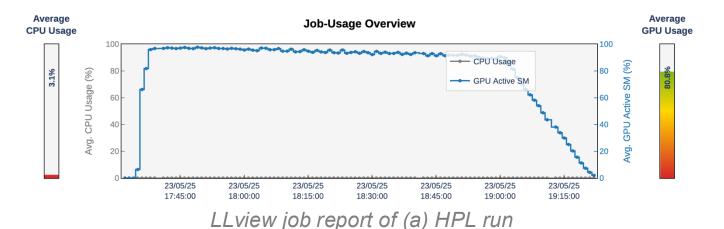




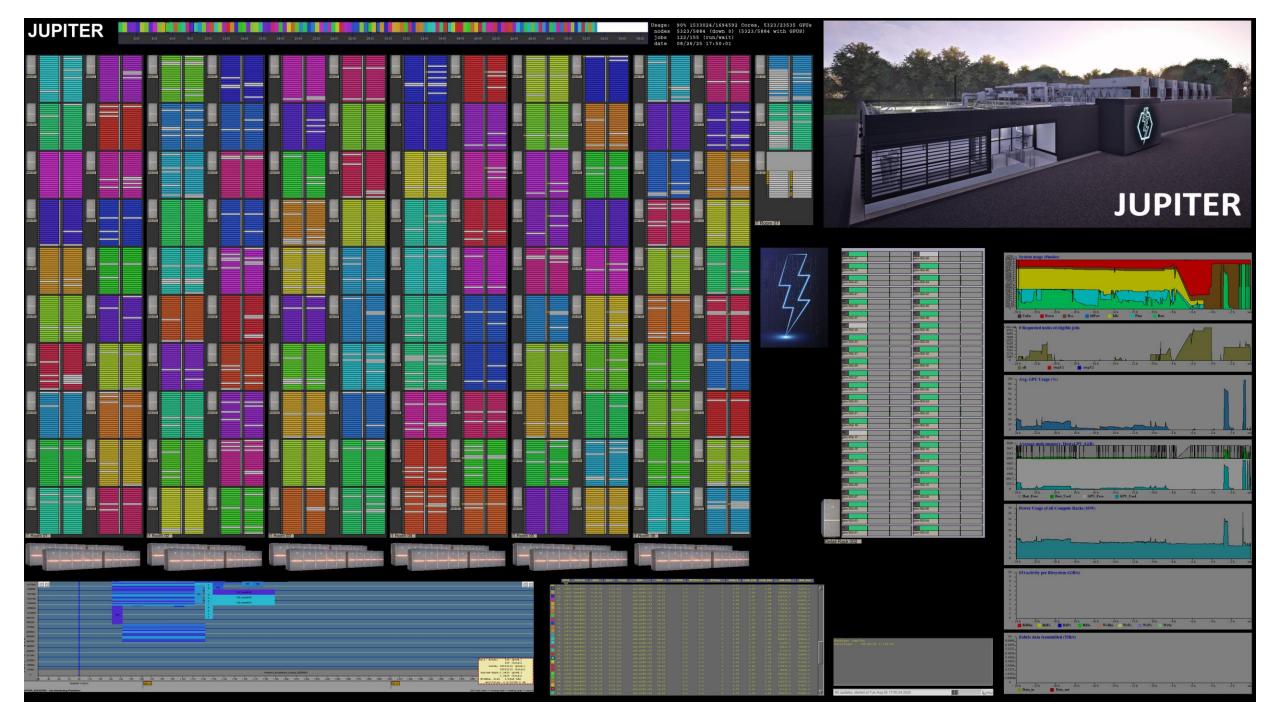


#### **TOP500**

- Entry to TOP500 June 2025 as #4 world, #1 EU
- 793 PFLOP/s HPL of 930 PFLOP/s th. peak
- Achieved with 4650 nodes







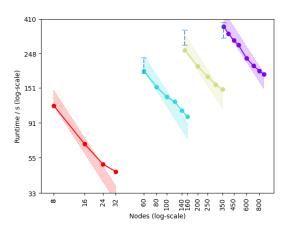
#### **POPULATING JUPITER**

#### **Current Status and Next Steps**

- JUPITER Research and Early Access Phase ongoing
  - > 100 Applications, 33 Lighthouses
- GCS AI Competition access started
- (Acceptance) Benchmarking running
  - TCO Application mix, High-Scaling
  - Storage Acceptance
- High-Performance Linpack
  - ... otherwise it would be too easy
- JUPITER Inauguration on 5.9.2025
- Autumn/winter: EuroHPC/GCS calls









#### **QUESTIONS?!**





### **JOINING FORCES**







Ministerium für Kultur und Wissenschaft des Landes Nordrhein-Westfalen















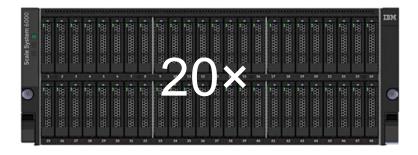


fz-juelich.de/jupiter

#### 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



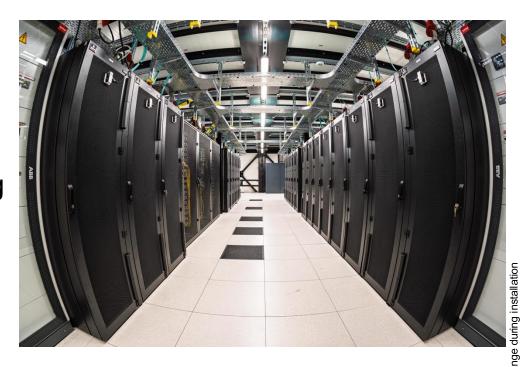
#### JUPITER – STORAGE (EXASTORE)





TCO 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 Building Block (14014 disks)
  - IBM Storage Scale (aka Spectrum Scale/GPFS)
- Manager and Datamover Nodes
- Exclusive for JUPITER
  - Integrated into InfiniBand fabric





#### JUPITER – STORAGE (EXATAPE)



TCO contribution from JSC, not part of the JUPITER procurement

- 369 Petabyte Tape Capacity
- Procurement Q3/Q4 2024, 2M€ Invest
- 2 x IBM TS4500 with LTO9
- TS1 in 16.4
  - 11 Frames
  - 15x LTO 9 tape drives
  - 10240 LTO9 media
- TS2 in 16.3
  - 11 Frames
  - 15x LTO9 tape drives
- 10260 LT09 media



