



JSC [HPC] SYSTEMS

JUWELS, JURECA-DC and JUSUF

21.05.2024 | D. ALVAREZ, S. ACHILLES

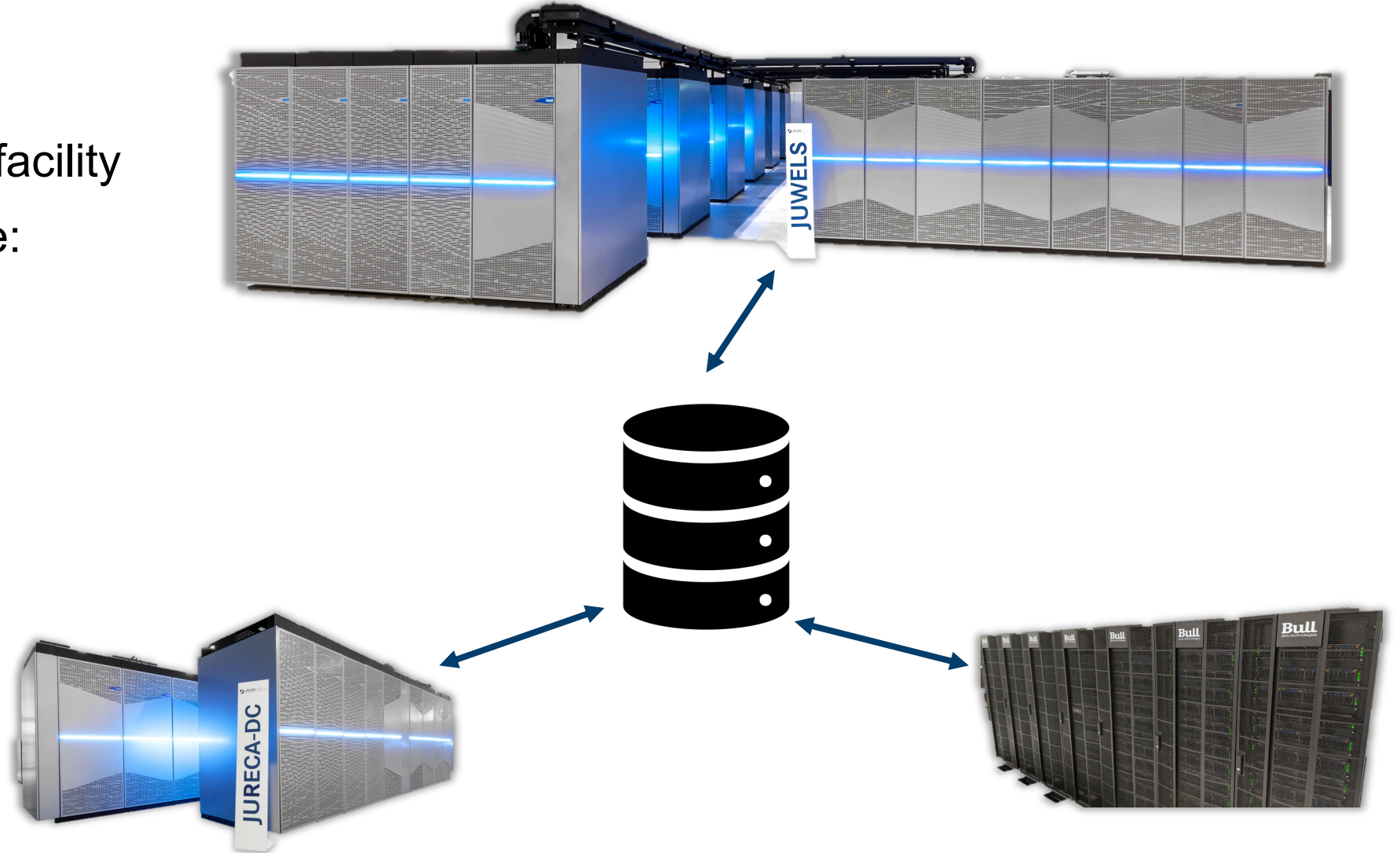
JSC [HPC] SYSTEMS

- JSC is a multi-system facility



JSC [HPC] SYSTEMS

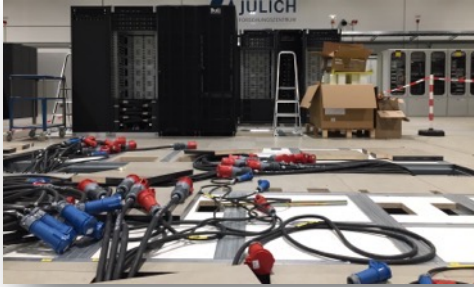
- JSC is a multi-system facility
- Main HPC systems are:
 - JUWELS
 - JURECA-DC
 - JUSUF
- Shared storage!
- Different talk



BRIEF JUWELS TIMELINE



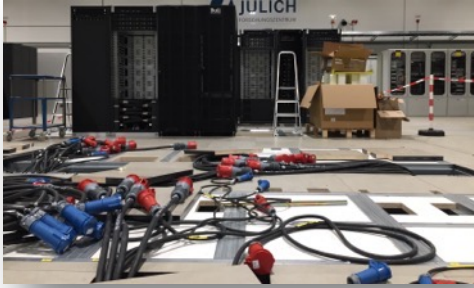
BRIEF JUWELS TIMELINE



JUWELS Cluster
installation begins



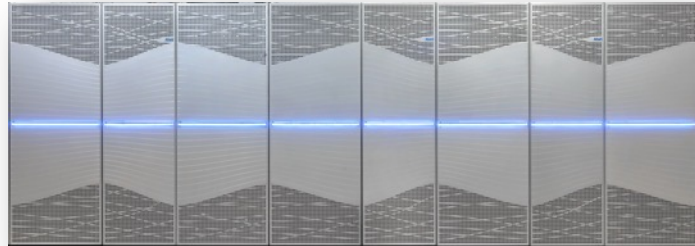
BRIEF JUWELS TIMELINE



JUWELS Cluster
installation begins



JUWELS Cluster
enters production



BRIEF JUWELS



CERTIFICATE

JUWELS Module 1 - Bull Sequana X1000, Xeon Platinum 8168 24C 2.7GHz,
Mellanox EDR InfiniBand/ParTec ParaStation ClusterSuite
Forschungszentrum Juelich (FZJ), Germany


is ranked

No. 127

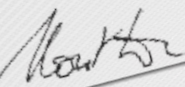
among the World's TOP500 Supercomputers
with **6.18 PFlop/s Linpack Performance**

in the 62nd TOP500 List published at the SC23
Conference on November 14, 2023.

Congratulations from the TOP500 Editors


Erich Strohmaier
NERSC/Berkeley Lab


Jack Dongarra
University of Tennessee


Horst Simon
NERSC/Berkeley Lab


Martin Meuer
Prometeus



JUWELS Module 1 - Bull Sequana X1000, Xeon Platinum 8168 24C 2.7GHz,
Mellanox EDR InfiniBand/ParTec ParaStation ClusterSuite
Forschungszentrum Juelich (FZJ), Germany

is ranked
No. 120

among the World's TOP500 Supercomputers
with **4.539 GFlops/watts Performance**


in the Green500 List published at the SC23
Conference on November 14, 2023.

Congratulations from the Green500 Editors


Wu-chun Feng
Virginia Tech

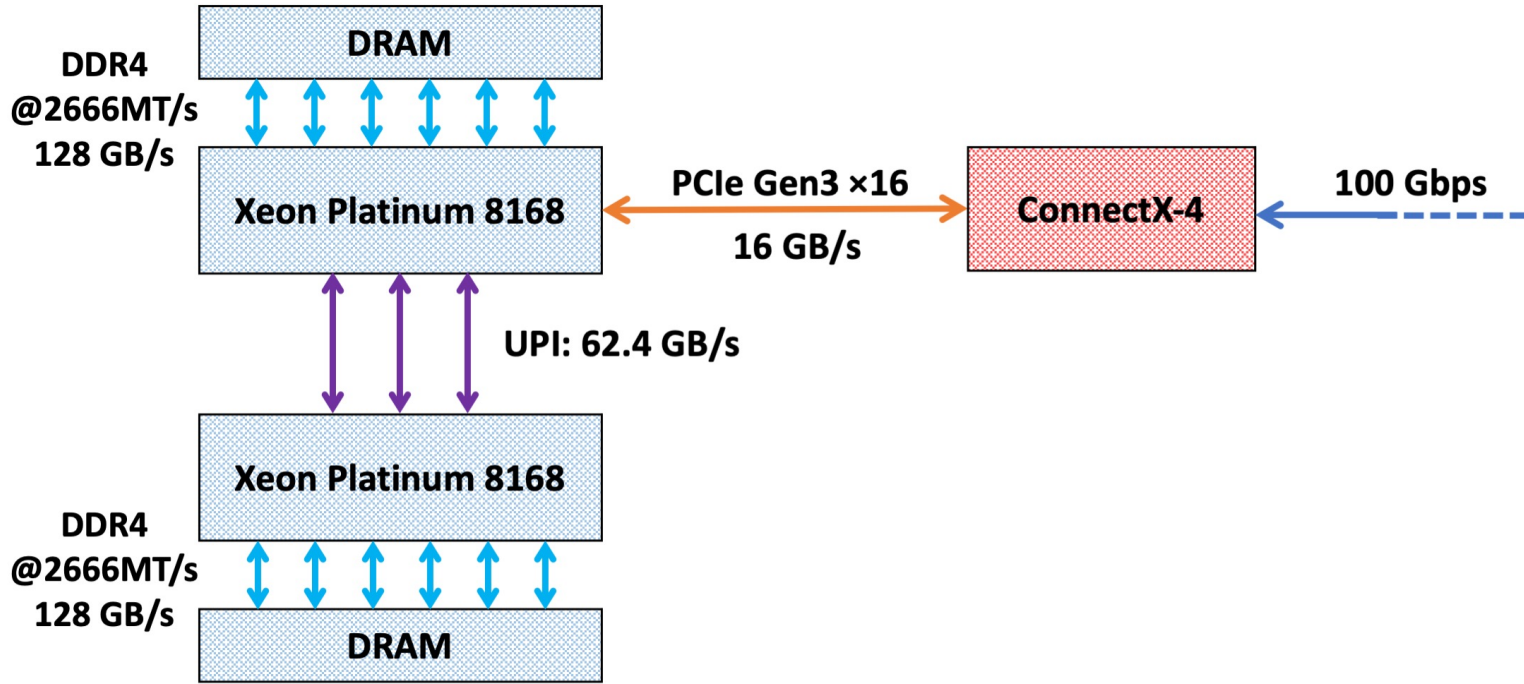

Kirk Cameron
Virginia Tech

JUWELS CLUSTER NODES





- 2511 compute nodes **Atos**
 - 2× 24-core Intel Xeon Platinum 8168 **intel**
 - 2x 6 memory channels
 - 2x 48 GB DDR4 @ 2.666 GHz
 - 240 nodes with 2x 96 GB DDR4 @ 2.666 GHz
 - PCIe Gen3
 - 1x EDR InfiniBand adapter (100Gbps) 



JUWELS CLUSTER NODES

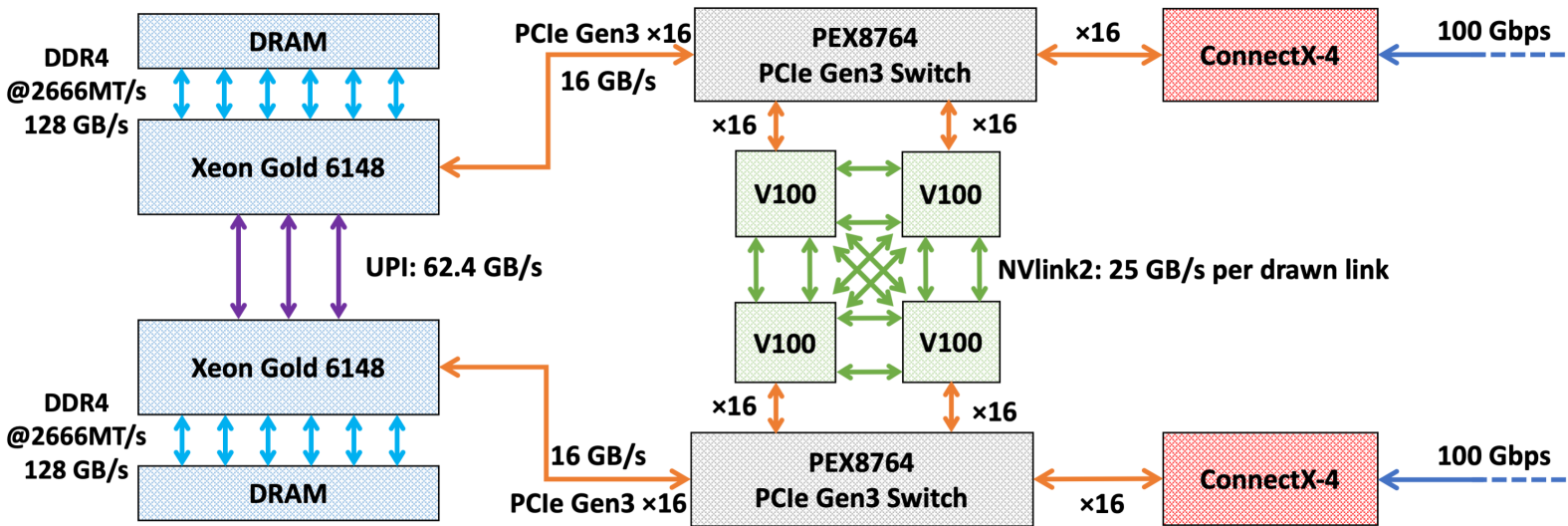


JUWELS CLUSTER GPU NODES

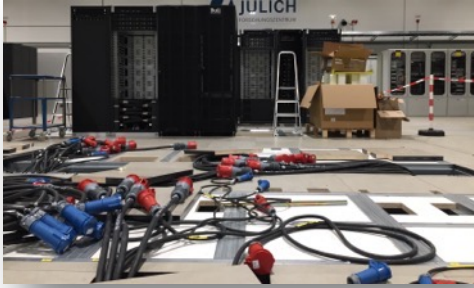
- 56 compute nodes 
 - 2x 20-core Intel Xeon Gold 6148 
 - 2x 6 memory channels
 - 2x 96 GB DDR4 @ 2.666 GHz
 - PCIe Gen3
 - PCIe Switch
 - 4x Nvidia V100 GPUs 
 - 7.8 TF/s peak
 - 16 GB HBM2
 - 900 GB/s memory performance
 - NVLink2 full mesh
 - 2 links (100GB/s bidir) between GPU pairs
 - PCIe Gen3 x16 (32 GB/s bidir)
 - 2x EDR InfiniBand adapter (100 Gbps) 



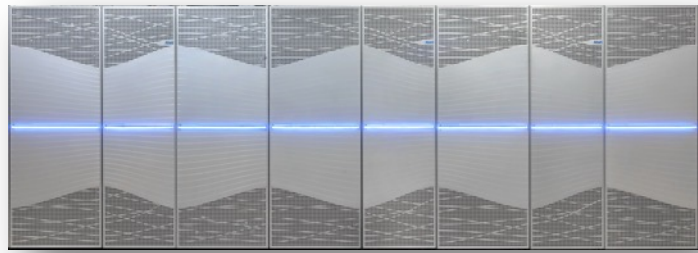
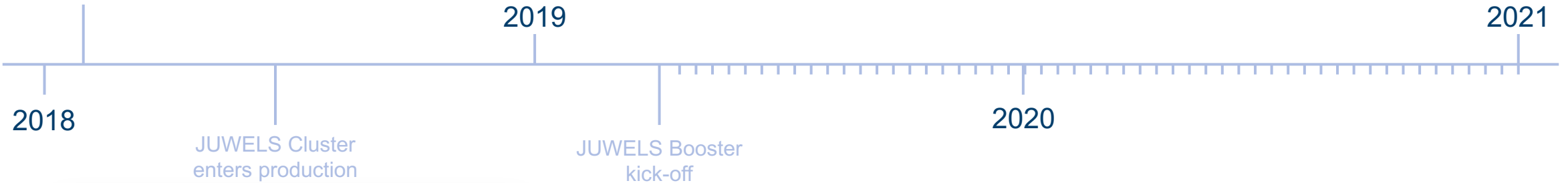
JUWELS CLUSTER GPU NODES



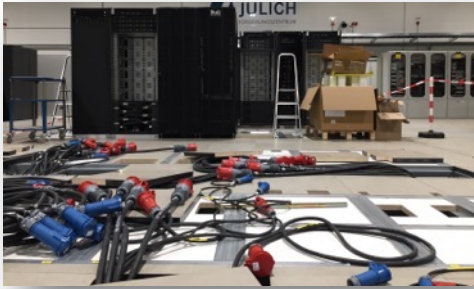
BRIEF JUWELS TIMELINE



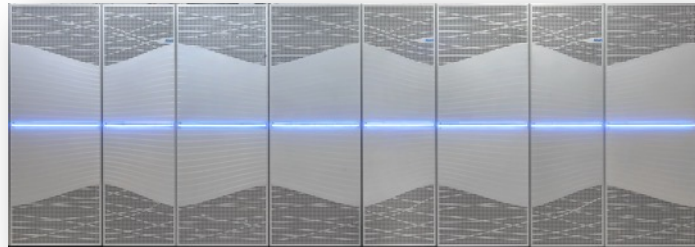
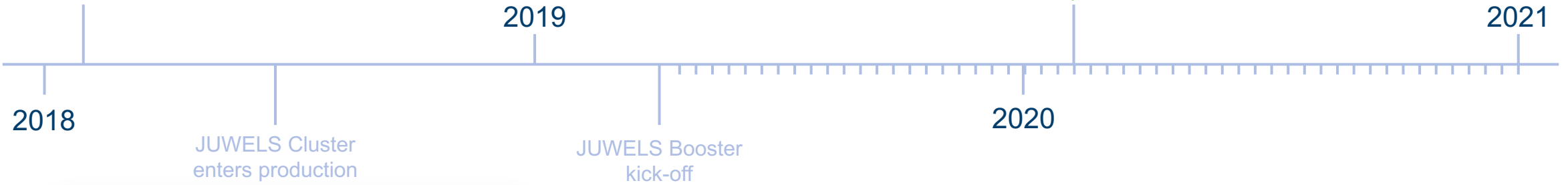
JUWELS Cluster installation begins



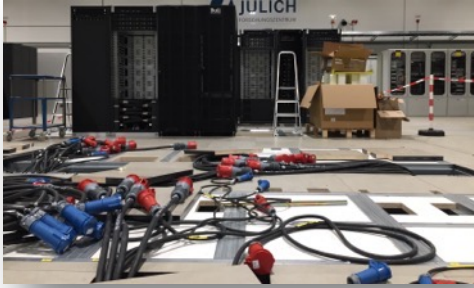
BRIEF JUWELS TIMELINE



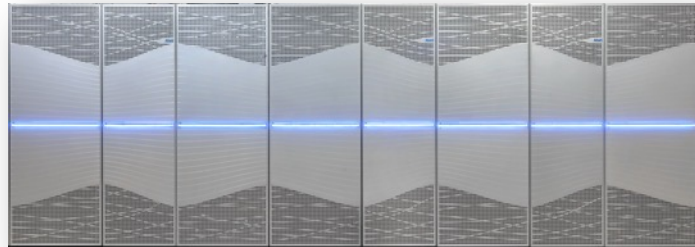
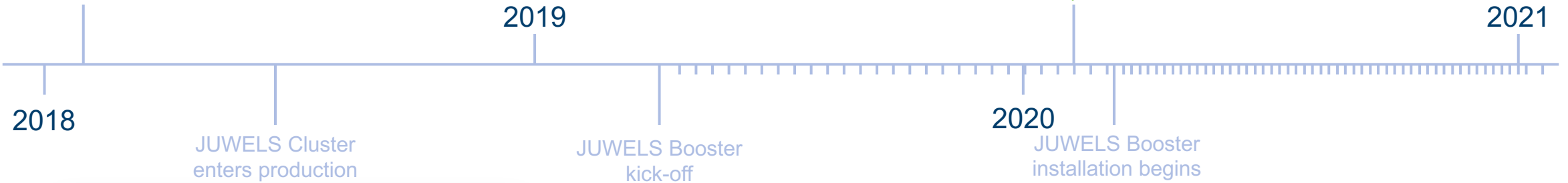
JUWELS Cluster installation begins



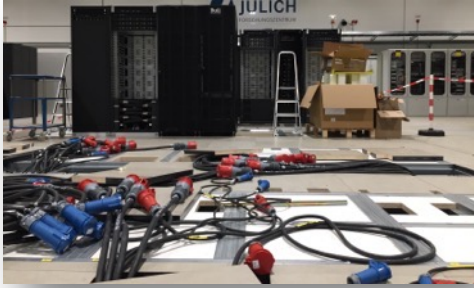
BRIEF JUWELS TIMELINE



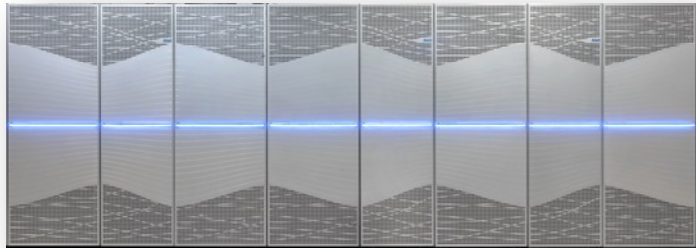
JUWELS Cluster installation begins



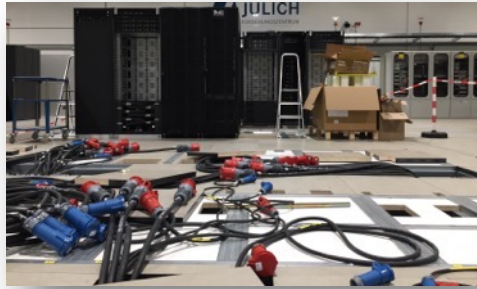
BRIEF JUWELS TIMELINE



JUWELS Cluster installation begins



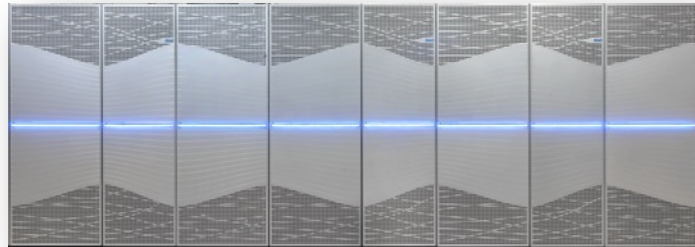
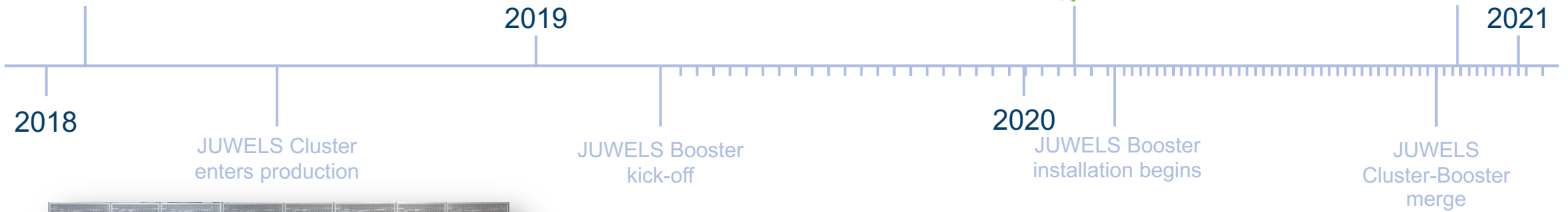
BRIEF JUWELS TIMELINE



JUWELS Cluster installation begins



JUWELS Booster enters production



BRIEF JUWELS TIME



TOP 500 CERTIFICATE

JUWELS Booster Module - Bull Sequana XH2000, AMD EPYC 7402 24C 2.8GHz, NVIDIA A100, Mellanox HDR InfiniBand/ParTec ParaStation ClusterSuite
Forschungszentrum Jülich (FZJ), Germany

JUWELS Booster
#1 in TOP500 Europe (11/2020), #7 WW
#1 in Green500 among the top 100 in HPL
#5 HPCG500
#4 HPL-AI

among the World's TOP500 Supercomputers
with 44.12 PFlop/s Linpack Performance
in the 62nd TOP500 List published at the SC23
Conference on November 14, 2023
Congratulations from the TOP500 Editors

Jack Dongarra
University of Tennessee

Horst Simon
NERSC/Berkeley Lab

Martin Meuer
Prometeus

The GREEN 500 CERTIFICATE



JUWELS Booster Module - Bull Sequana XH2000, AMD EPYC 7402 24C 2.8GHz, NVIDIA A100, Mellanox HDR InfiniBand/ParTec ParaStation ClusterSuite
Forschungszentrum Jülich (FZJ), Germany

is ranked **No. 31**
among the World's TOP500 Supercomputers
with 25.008 GFlops/watts Performance
in the Green500 List published at the SC23
Conference on November 14, 2023.
Congratulations from the Green500 Editors

Wu-chun Feng
Virginia Tech

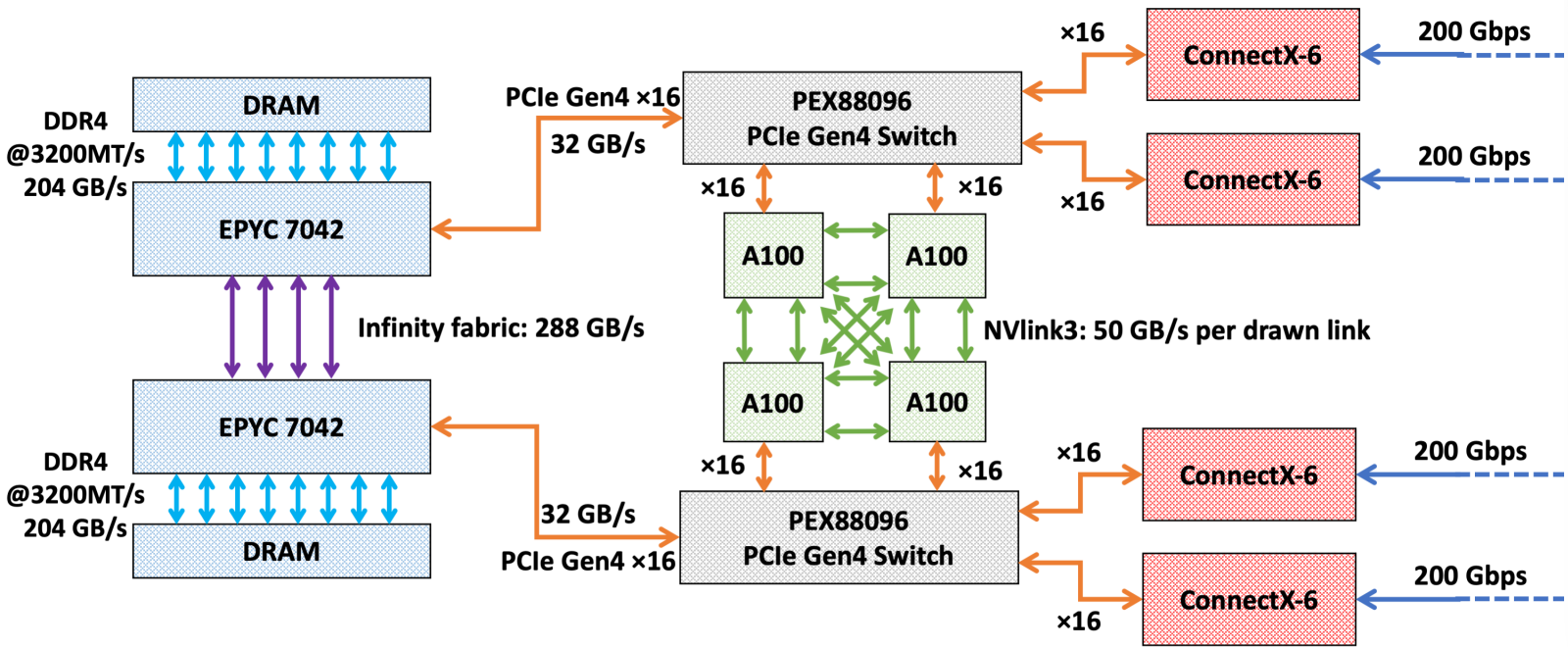
Kirk Cameron
Virginia Tech

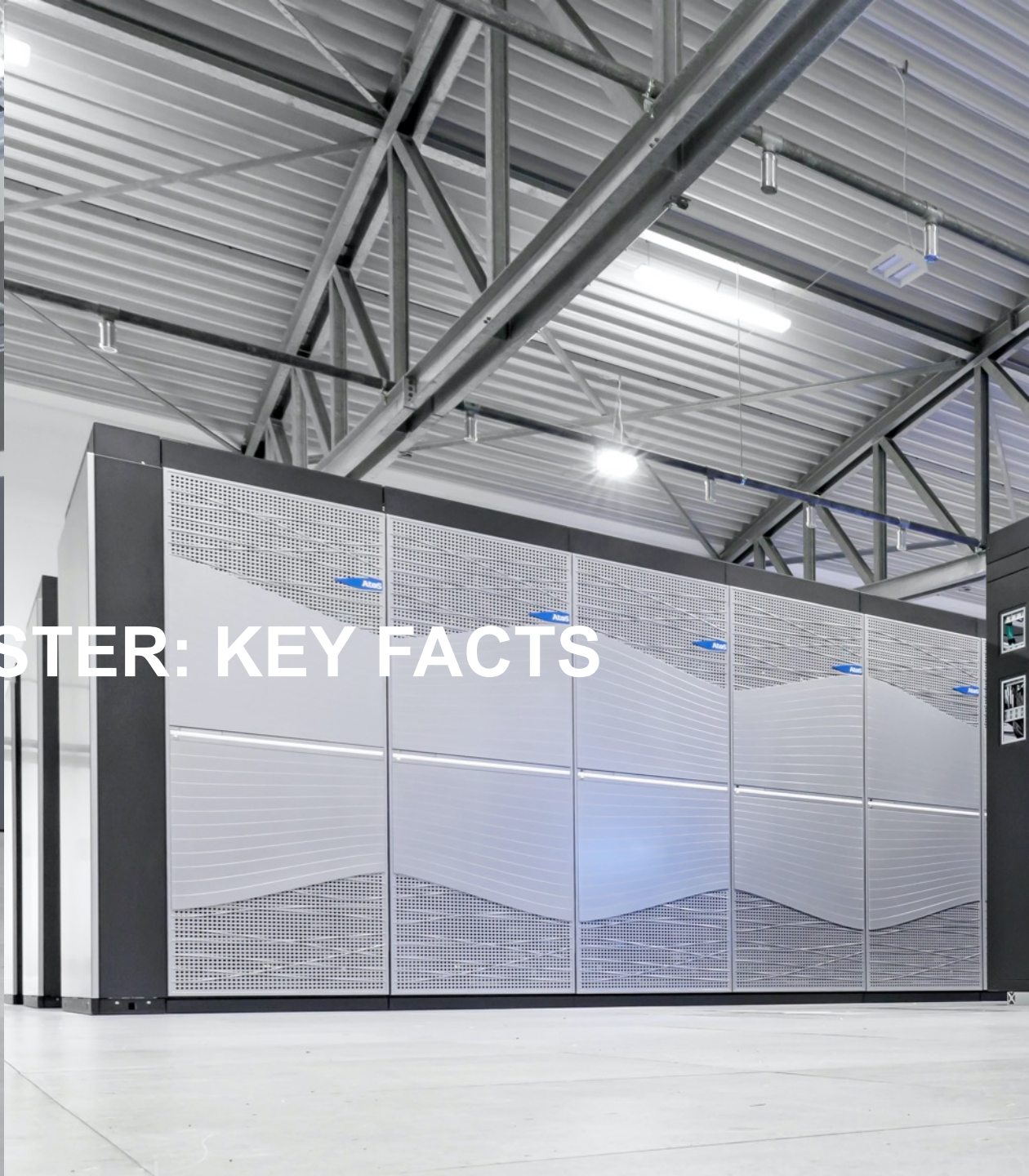
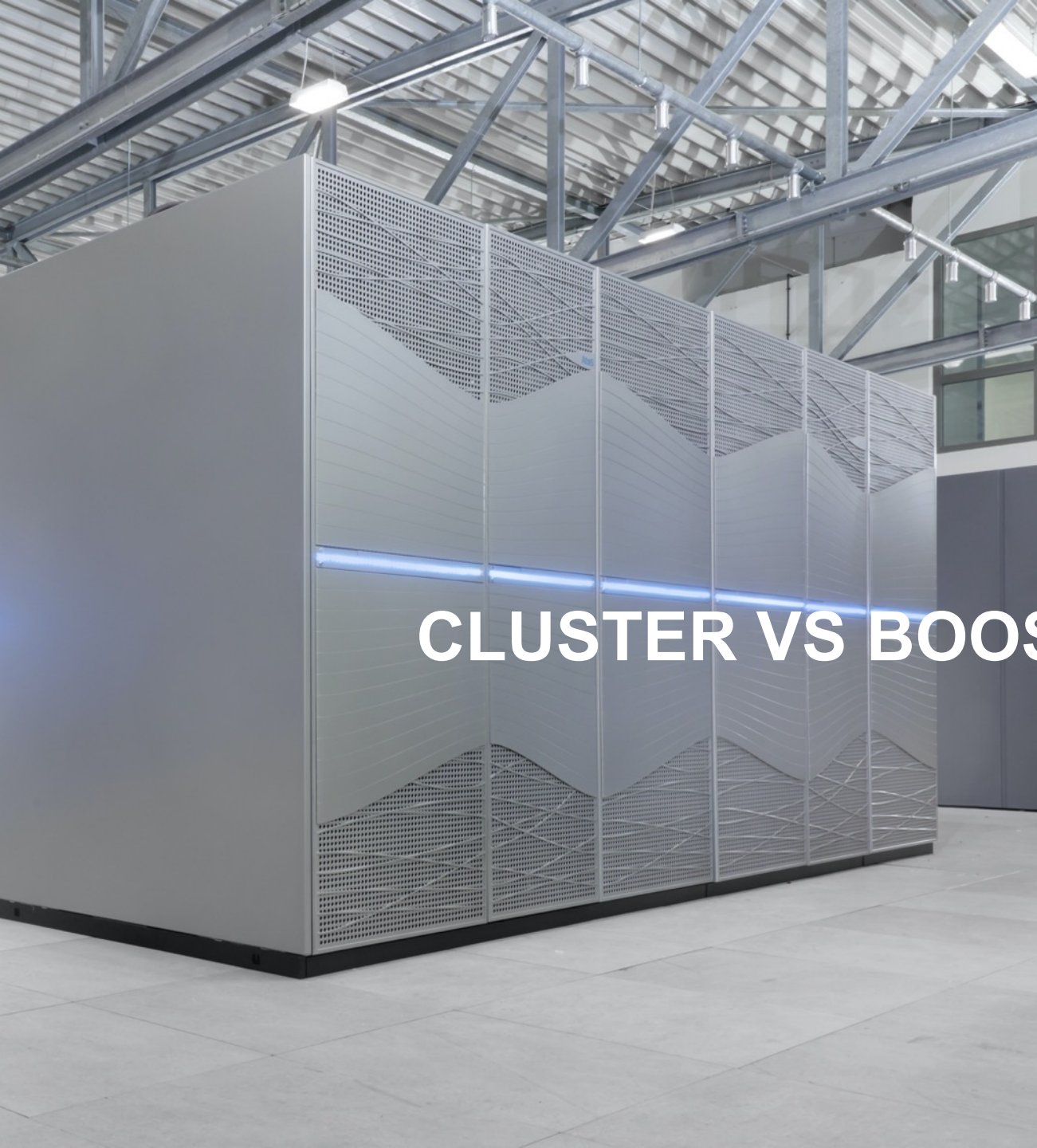
JUWELS BOOSTER NODES

- 936 compute nodes **Atos**
 - 2x 24-core AMD Epyc 7402 Rome CPUs **AMD**
 - 2x 8 memory channels
 - 2x 256 GB DDR4 @ 3.2GHz
 - 2x 4 NUMA domains
 - 96 PCIe Gen4 lanes
 - 512 GB DDR memory
 - **4x Nvidia A100 GPUs** 
 - 9.7 / 19.5 TF/s peak
 - 40 GB HBM2
 - 1.5 TB/s memory performance
 - NVLink3 full mesh
 - 4 links (200GB/s) between GPU pairs
 - PCIe Gen4 x32 (64 GB/s)
 - **4x HDR200 InfiniBand adapter (1 per GPU)** 



JUWELS BOOSTER NODES





CLUSTER VS BOOSTER: KEY FACTS

CLUSTER VS BOOSTER –NODE VIEW– (1/2)

JUWELS Cluster (w/o GPU nodes)

JUWELS Booster

Processors	Intel	-	AMD	Processors
Cores	48	x1	48	Cores
Vector width (CPU)	512	x0.5	256	Vector width (CPU)
Memory (main)	96/192 GB	x5.33/2.66	512 GB	Memory (main)
Memory BW (main)	256 GB/s	x1.59	408 GB/s	Memory BW (main)
GPUs	0	xNaN	4	GPUs
Memory (GPU)	0	xNaN	160 GB	Memory (GPU)
Memory BW (GPU)	0	xNaN	6 TB/s	Memory BW (GPU)
HCA	1	x4	4	HCA
Link BW	100 Gbps	x2	200 Gbps	Link BW
Network BW	100 Gbps	x8	800 Gbps	Network BW
TFLOPs	4.15	x18.8	78	TFLOPs (GPUs)

CLUSTER VS BOOSTER –GLOBAL VIEW– (2/2)

JUWELS Cluster (w/o GPU nodes)

Peak performance	10.6 PF
Concurrency	240 K
Total memory	96 TB
Total memory BW	0.6 PB/s
Gb per TF	24.1
Injection BW	251 Tb/s
Topology	Prun. FT
Global network bandwidth	63 Tb/s
Routing	Detem.

x6.88

x216

x6.5

x9.3

x0.42

x2.98

-

x3.17

-

JUWELS Booster

73 PF

»52 M

629 TB

5.6 PB/s

10.3

749 Tb/s

DF+

200 Tb/s

Adaptive

Peak performance

Concurrency

Total memory

Total memory BW

Gb per TF

Injection BW

Topology

Global network bandwidth

Routing

JUWELS CLUSTER LOGIN NODES

- 9 + 2 standard login nodes
 - 2× 20-core Intel Xeon Gold 6148
 - 756 GB DDR4 @ 2.666 GHz
 - 100 GigE external network
- 4 visualization nodes
 - 2× 20-core Intel Xeon Gold 6148
 - 756 GB DDR4 @ 2.666 GHz
 - 100 GigE external network
 - **1x Nvidia P100 GPU**
 - **Different compute capabilities than in compute nodes!**
- Used for:
 - Compile/submit jobs
 - **Careful with `make -j`!**
 - **Small** pre- and post-processing/visualization
 - **Shared nodes!**



JUWELS BOOSTER LOGIN NODES

- 4 login nodes
 - 2× 24-core AMD Epyc 7402 Rome CPUs
 - 512 GB DDR4 @ 3.2 GHz
 - 100 GigE external network
 - **No GPUs!**
- Used for:
 - Compile/submit jobs
 - **Careful with `make -j` !**
 - **Small** pre- and post-processing/visualization
 - **Shared nodes!**



JURECA-DC

DC = Data Centric

- Intended for mixed capacity and capability workloads
- Designed with big-data science needs in mind



JURECA-DC

DC = Data Centric

TOP 500 CERTIFICATE

The List.

JURECA Data Centric Module - BullSequana XH2000, AMD EPYC 7742 64C 2.25GHz,
NVIDIA A100 40GB, Mellanox HDR InfiniBand/ParTec ParaStation ClusterSuite
Forschungszentrum Juelich (FZJ), Germany

is ranked

No. 82

among the World's TOP500 Supercomputers
with **9.33 PFlop/s Linpack Performance**
in the 62nd TOP500 List published at the SC23
Conference on November 14, 2023.

Congratulations from the TOP500 Editors


Erich Strohmaier
NERSC/Berkeley Lab


Jack Dongarra
University of Tennessee


Horst Simon
NERSC/Berkeley Lab


Martin Meuer
Prometeus

The
GREEN
500

CERTIFICATE

JURECA Data Centric Module - BullSequana XH2000, AMD EPYC 7742 64C 2.25GHz,
NVIDIA A100 40GB, Mellanox HDR InfiniBand/ParTec ParaStation ClusterSuite
Forschungszentrum Juelich (FZJ), Germany

is ranked

No. 32


among the World's TOP500 Supercomputers
with **24.291 GFlops/watts Performance**
in the Green500 List published at the SC23
Conference on November 14, 2023.

Congratulations from the Green500 Editors


Wu-chun Feng
Virginia Tech






Kirk Cameron
Virginia Tech

JURECA-DC CPU NODES

- 576 compute nodes **Atos**
 - 2x **64-core** AMD Epyc 7742 Rome CPUs **AMD**
 - 2x 8 memory channels
 - 2x 256 GB DDR4 @ 3.2 GHz
 - 96 nodes with 2x 512 GB DDR4 @ 3.2 GHz
 - 2x 4 NUMA domains
 - PCIe Gen4
 - 1x HDR100 InfiniBand adapter (100Gbps) 



JURECA-DC GPU NODES

- 192 compute nodes 
 - 2x **64-core** AMD Epyc 7742 Rome CPUs 
 - 2x 8 memory channels
 - 2x 256 GB DDR4 @ 3.2GHz
 - 96 PCIe Gen4 lanes
 - 512 GB DDR memory
 - **4x** Nvidia A100 GPUs 
 - 9.7 / 19.5 TF/s peak
 - 40 GB HBM2
 - 1.5 TB/s memory performance
 - NVLink3 full mesh
 - 4 links (200GB/s) between GPU pairs
 - PCIe Gen4 x32 (64 GB/s)
 - **2x** HDR200 InfiniBand adapter (1 per GPU) 



JURECA-DC LOGIN NODES

- 12 login nodes
 - 2× 64-core AMD Epyc 7742 Rome CPUs
 - 1024 GB DDR4 @ 3.2 GHz
 - 100 GigE external network
 - 2x Nvidia RTX8000 GPUs
 - Different compute capabilities than in compute nodes!
- Used for:
 - Compile/submit jobs
 - Careful with `make -j` !
 - Small pre- and post-processing/visualization
 - Shared nodes!



JURECA-DC PROTOTYPE/TEST/NEW NODES

- 2x MI250X nodes
 - 2x 24-core AMD Epyc 7443 Milan CPUs
 - 512 GB DDR4 @ 3.2 GHz
 - 2x HDR200 InfiniBand adapter
 - 4x AMD MI250X GPUs
- 2x NVIDIA ARM HPC DevKit nodes
 - 1x Ampere Altra Q80-30
 - 512 GB DDR4 @ 3.2 GHz
 - 2x HDR200 InfiniBand adapter
 - 2x NVIDIA A100 GPUs
- 1x Graphcore IPU-M2000 node
 - 4x GC200 IPUs

JURECA-DC PROTOTYPE/TEST/NEW NODES

- 1x Sapphire Rapids + NVIDIA H100 node
 - 2x 36-core Intel Xeon Platinum 8452Y CPUs
 - 512 GB DDR5 @ 4.8 GHz
 - 4x NVIDIA H100 GPUs (PCIe/350W/80GB)
 - 1x BlueField-2 InfiniBand adapter

- 2x Grace-Hopper nodes
 - 1x Grace-Hopper Superchip
 - 72 ARM Neoverse V2 cores
 - 480 GB LPDDR5X (Grace)
 - 90 GB HBM3 (H100)
 - 1x HDR200 InfiniBand adapter



- 16x Sapphire Rapids + NVIDIA 4xH100 nodes
 - 2x 32-core Intel Xeon Platinum 8462Y CPUs
 - 512 GB DDR5 @ 4.8 GHz
 - 4x NVIDIA H100 GPUs (SXM5/700W/90 GB)
 - 2x NDR400 InfiniBand adapters

JUSUF

- Serves the ICEI project (Interactive Computing E-Infrastructure for the Human Brain Project)
- Contains 2 partitions
 - HPC
 - Cloud
- Air-cooled, less dense than other systems



JUSUF HPC PARTITION

- 124 compute nodes **Atos**
 - 2x **64-core** AMD Epyc 7742 Rome CPUs **AMD**
 - 2x 8 memory channels
 - 2x 128 GB DDR4 @ 3.2 GHz
 - 2x 4 NUMA domains
 - PCIe Gen4
 - 1x HDR100 InfiniBand adapter (100Gbps)
 - 1x 40 GbE adapter (for storage)
 - **1TB NVMe local scratch**
- 49 GPU nodes **Atos**
 - Same config as CPU nodes. Additionally:
 - 1x Nvidia V100 GPUs 
 - 7.8 TF/s peak
 - 16 GB HBM2
 - 900 GB/s memory performance
 - PCIe Gen3 x16 (32 GB/s bidir)





FURTHER INFORMATION

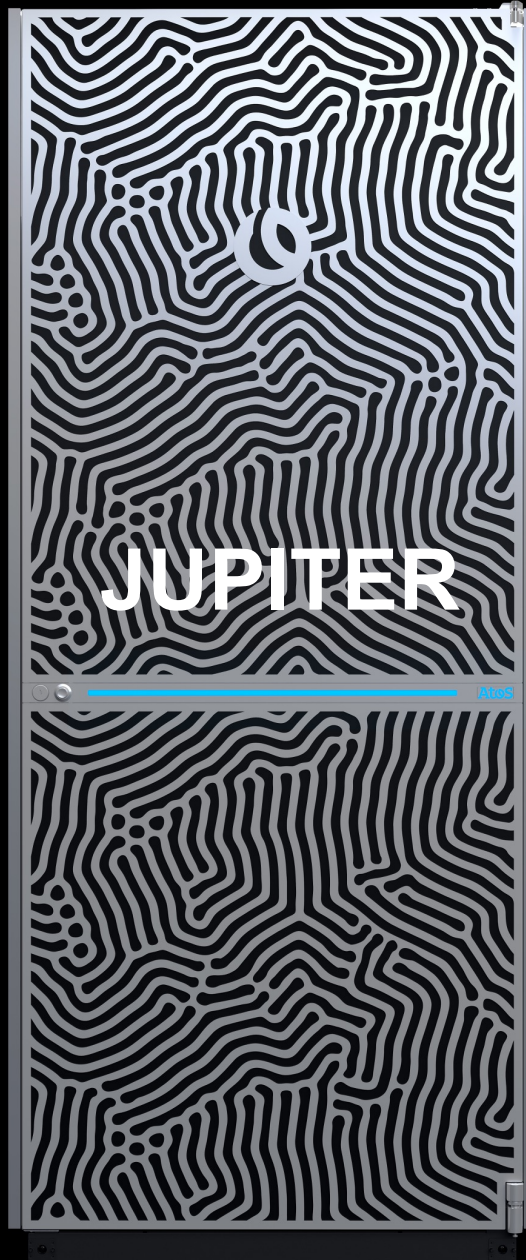
MAINTENANCE HANDLING

- JSC systems go on maintenance for any of the following reasons:
 - JUST (storage cluster) needs maintenance
 - Compute node updates (OS and/or FW and/or configuration changes)
 - Admin node updates (OS and/or FW and/or configuration changes)
 - Emergencies
- Frequency
 - Depends on pending issues
 - Typically decreases as system ages
- Days and duration
 - Typically on Tuesdays
 - Whole working day
 - Announced with at least 1 week in advance
- Communicated through **MOTD** and **status page**

IMPORTANT LINKS

- Status page:
 - <https://status.jsc.fz-juelich.de/>
- General system information
 - <https://go.fzj.de/JUWELS>
 - <https://go.fzj.de/juwels-known-issues>
 - <https://go.fzj.de/JURECA>
 - <https://go.fzj.de/jureca-known-issues>
 - <https://go.fzj.de/JUSUF>
 - <https://go.fzj.de/jusuf-known-issues>
- User documentation:
 - <https://apps.fz-juelich.de/jsc/hps/juwels/index.html>
 - <https://apps.fz-juelich.de/jsc/hps/jureca/index.html>
 - <https://apps.fz-juelich.de/jsc/hps/jusuf/index.html>
- Job reporting:
 - <https://go.fzj.de/llview-juwels>
 - <https://go.fzj.de/llview-juwelsbooster>
 - <https://go.fzj.de/llview-jureca>
- User support at FZJ
 - sc@fz-juelich.de
 - Phone: 02461 61-2828

1
MORE THING

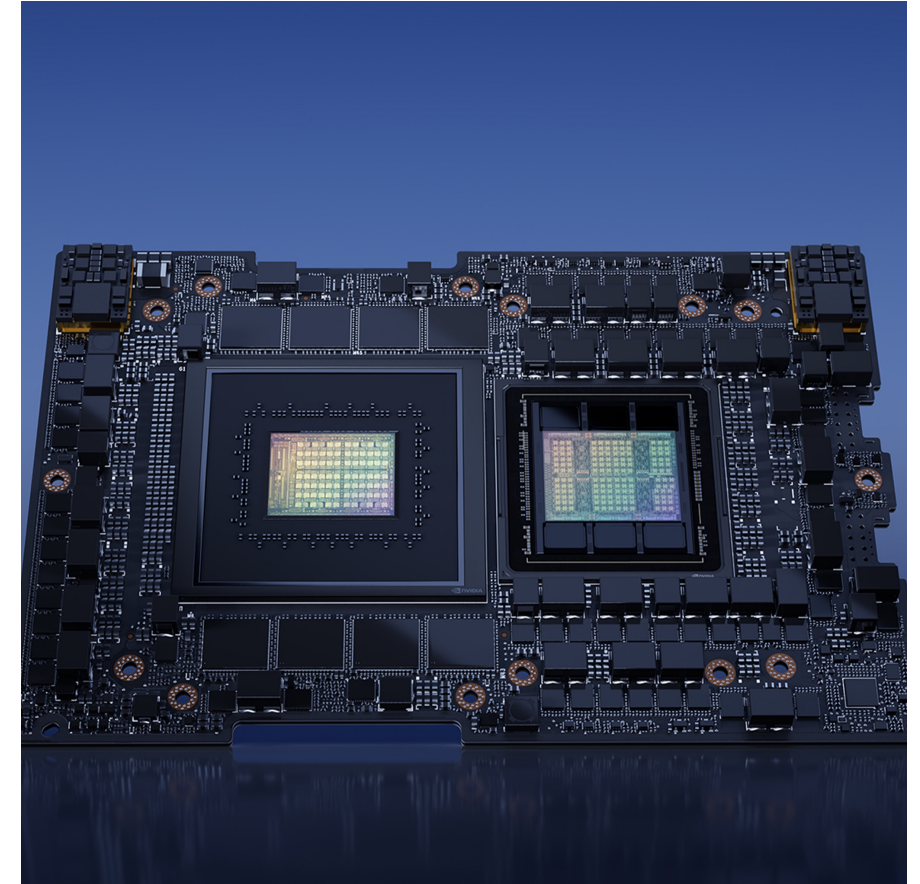


JUPITER – THE BOOSTER

Highly-Scalable Module for HPC and AI workloads

- 1 ExaFLOP/s (FP64, HPL)
- NVIDIA Grace-Hopper CG1
 - ~5900 compute nodes
 - 4x CG1 chips per compute node
- NVIDIA Mellanox NDR
 - 4 NDR200 NICs per compute node
- BullSequana XH3000
 - Direct Liquid Cooled blades
 - 2 compute node per blade

EVIDEN
an atos business



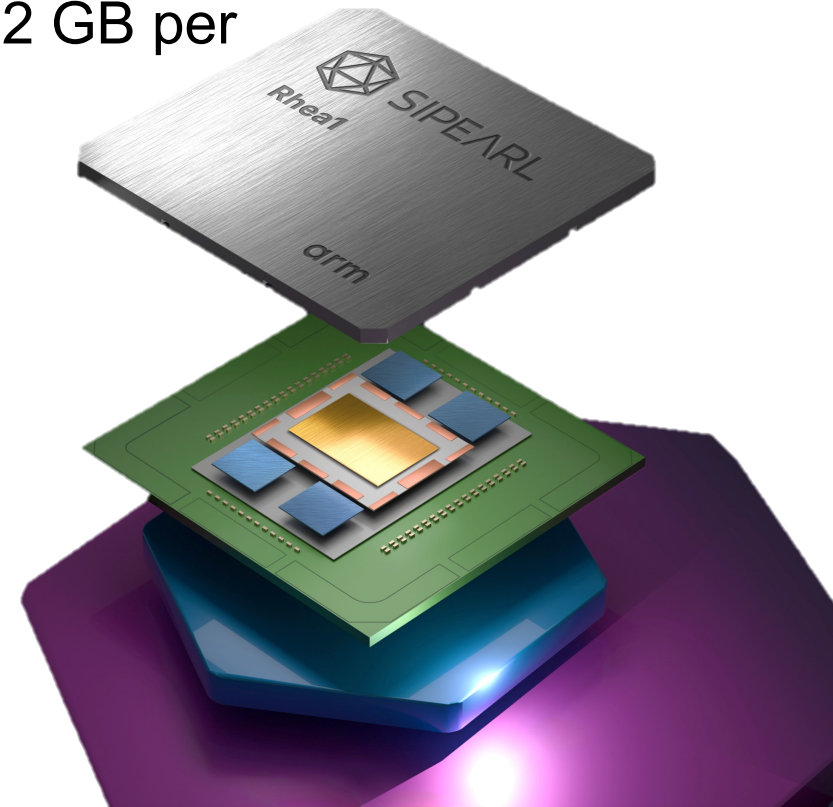
JUPITER – THE CLUSTER

General-Purpose Module for Mixed Workloads

- >5 PetaFLOP/s (FP64, HPL)
- SiPearl Rhea1
 - ~1340 compute nodes
 - 2× CPUs per node
- NVIDIA Mellanox NDR
 - 1× NDR200 NICs per compute node
- BullSequana XH3000
 - Direct Liquid Cooled blades
 - 3× compute nodes per blade



- 80 Neoverse V1 cores
 - 2× 256 SVE each
- 64 GB HBM (128 GB per node)
- 256 GB DDR5 (512 GB per node)



JUWELS VS. JUPITER

	JUWELS	JUPITER
Cluster	CPU: Intel Xeon Platinum 8168 GPU: NVIDIA V100 Peak: 10 PFlop/s	CPU: SiPearl Rhea1 GPU: none Mem. Bandwidth: 0,51 Byte/Flop
Booster	CPU: 2* AMD Epyc Rome GPU: 4× NVIDIA A100 GPUs Peak: 73 PFlop/s	CPU: 4* NVIDIA Grace GPU: 4* NVIDIA Hopper Peak: >1 EFlop/s
Network topology	Fat tree and DragonFly+	DragonFly+
System access	GCS or PRACE proposals	GCS and EuroHPC JU proposals
User support	HLST, SDL, ATML, training courses, targeted early access program	same



FIRST PUBLIC ACHIEVEMENTS

Copyright: — Forschungszentrum *Jülich* / Ralf-Uwe Limbach

The logo features the word "TOP" in white on an orange speech bubble, followed by "500" in large black font, and "CERTIFICATE" in a smaller grey font. Below "500" is the text "The List." in a smaller grey font.

TOP 500 CERTIFICATE
The List.

**JEDI - BullSequana XH3000, Grace Hopper Superchip 72C 3GHz, NVIDIA GH200 Superchip,
Quad-Rail NVIDIA InfiniBand NDR200**

EuroHPC/FZJ, Germany

is ranked

No. 189

among the World's TOP500 Supercomputers

with 4.50 PFlop/s Linpack Performance

in the 63rd TOP500 List published at the ISC24

Conference on June 01, 2024.


Congratulations from the TOP500 Editors



Erich Strohmaier
NERSC/Berkeley Lab



Jack Dongarra
University of Tennessee



Horst Simon
NERSC/Berkeley Lab



Martin Meuer
Prometeus



CERTIFICATE

JEDI - BullSequana XH3000, Grace Hopper Superchip 72C 3GHz, NVIDIA GH200 Superchip,
Quad-Rail NVIDIA InfiniBand NDR200

EuroHPC/FZJ, Germany

is ranked

No. 1

among the World's TOP500 Supercomputers

with 72.733 GFlops/watts Performance

in the Green500 List published at the ISC24

Conference on June 01, 2024.

Congratulations from the Green500 Editors

A handwritten signature in black ink, appearing to read 'Wu-chun Feng'.

Wu-chun Feng
Virginia Tech

A handwritten signature in black ink, appearing to read 'Kirk Cameron'.

Kirk Cameron
Virginia Tech



CERTIFICATE

JEDI - BullSequana XH3000, Grace Hopper Superchip 72C 3GHz, NVIDIA GH200 Superchip,
Quad-Rail NVIDIA InfiniBand NDR200

EuroHPC/FZJ, Germany

is ranked

No. 1

among the World's TOP500 Supercomputers

with 72.733 GFlops/watts Performance

in the Green500 List published at the ISC24

Conference on June 01, 2024.

Congratulations from the Green500 Editors

A handwritten signature in black ink, appearing to read 'Wu-chun Feng'.

Wu-chun Feng
Virginia Tech

A handwritten signature in black ink, appearing to read 'Kirk Cameron'.

Kirk Cameron
Virginia Tech

- 1 Rack 50% populated
 - 12 Blades
 - 24 Nodes

More details on the
Green500 BoF
tomorrow

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

JUPITER

The Arrival of
Exascale in Europe

fz-juelich.de/jupiter | [#exa_jupiter](https://twitter.com/#!/exa_jupiter)



Funding Agencies:



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

