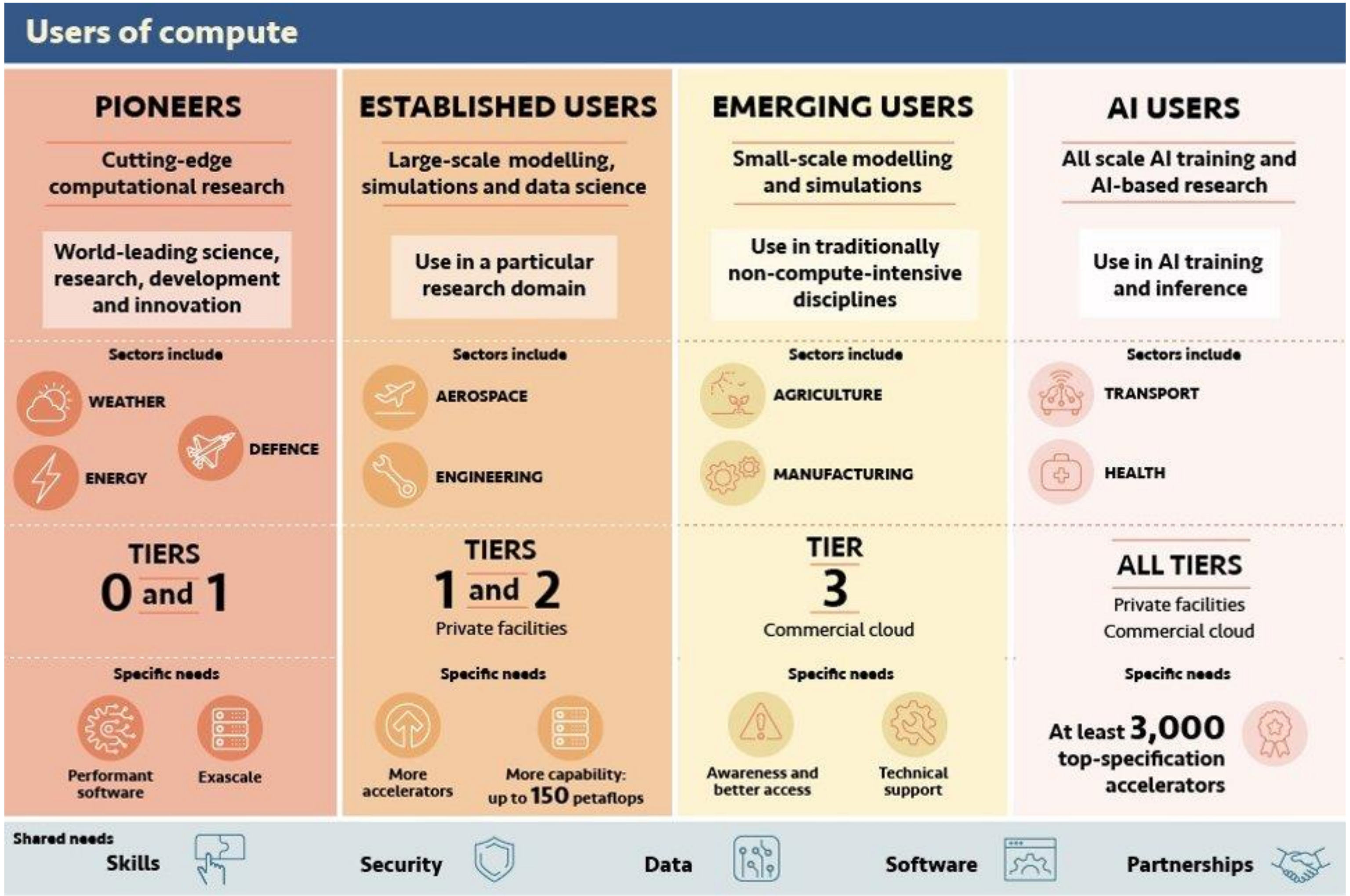
Isambard-AI: a leadership class supercomputer optimised specifically for Artificial Intelligence

Simon McIntosh-Smith
Sadaf R Alam
Christopher Woods
CUG May 2024

Outline

- Background and timeline
- Design specifications for an AI Research Resource (AI RR)
- Implementation details
- Future outlook

Design Specifications for AI Research Resource (RR)



Accessible to all users

Sustainable AI supercomputing

Accessible to all sectors

Performance for all tiers

Sustainable AI RR program

References: [Independent Review of The Future of Compute: Final report and recommendations, March 2023]; [National AI Strategy - AI Action Plan, July 2022]; [£300 million to launch first phase of new AI Research Resource]

University of BRISTOL

2016–July 2023

(ARM based Isambard 1, 2) Isambard 3 in 2023/4 (no data centre and Isambard PI Simon MS with a non-dedicated GW4 team)

Aug. 2023

UK govt feedback on Bristol's Isambard-AI proposal

Sep.–Oct. 2023

Isambard-AI procurement

Nov. 2023

AI Safety Summit

Dec. 2023

Isambard hiring started

Mar. 2024

Modular data centre (POD) and Isambard-AI phase 1 installed + a new team in ~3 months

Apr. 2024

Service configuration and hardening for AI users

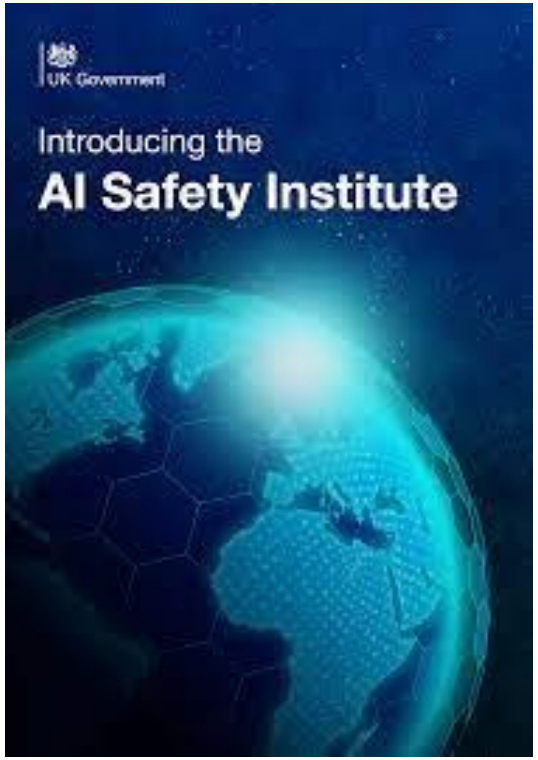
University of BRISTOL

What is Isambard-AI for?

- UKRI-funded AI research in the UK, e.g.:
  - Training large language models
  - Large-scale inference
  - Foundational AI research
  - AI safety and understanding
  - Hybrid AI + simulation workflows
  - Machine learning
- Research on Isambard-AI must have a strong AI component
- Accommodate GPU jobs at any scale
  - Interactivity via JupyterHub—single to 100s of GPUs
  - Long running jobs for large-scale training—10s to 1000s of GPUs

UK Research and Innovation

Department for Science, Innovation, & Technology



University of BRISTOL

Isambard 3 – Leading ARM based HPC since 2016

- Built on the success of Isambard 1 and 2 ARM based HPC supercomputing concept
  - Nvidia Grace ARM processors
  - 55,000 cores
- High speed network connectivity
  - Cray/HPE SlingShot 11 (made in Bristol)
- Air cooled racks i.e. not HPE Cray EX
- Nvidia and HPE/Cray programming environments for easy transition from Isambard 2

Nvidia ARM Grace Superchip

University of BRISTOL

Design points and constraints for Isambard-AI

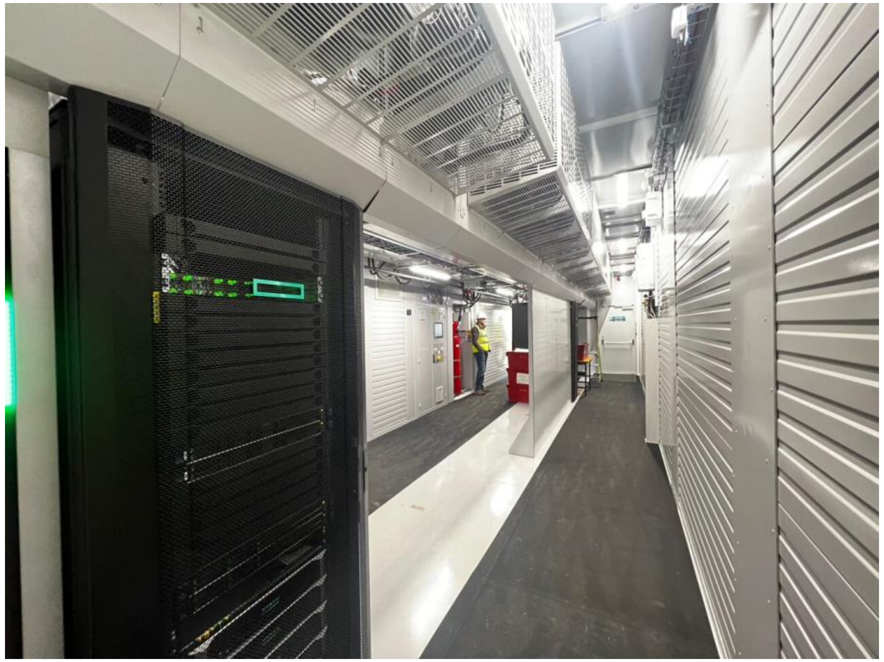
- Self Contained Units (SCUs) for rapid deployment and sustainability
- AI and performance optimised compute and network
- Physical space limits
- Direct liquid cooling
- Storage and data platforms for AI
- High performance, cloud-native software ecosystem for AI platforms
- Lowering accessibility barriers and federation
- Cybersecurity and RBAC (roles based access controls)

Please see paper for details

University of BRISTOL

Self Contained Units—Modular Data Centre (MDC)

- HPE Performance Optimised Data Centre (POD) solution
  - Rapid deployment
  - Customised for requirements e.g. DLC, security, power, cooling, etc.
  - Contains everything
    - IT modules
    - Power modules
    - Cooling modules
- Sustainability
  - Operational usage e.g. low PUE
  - NetZero emissions reporting
  - Fine grain and holistic telemetry
  - Recycle and reuse



Isambard Site on March 15, 2024

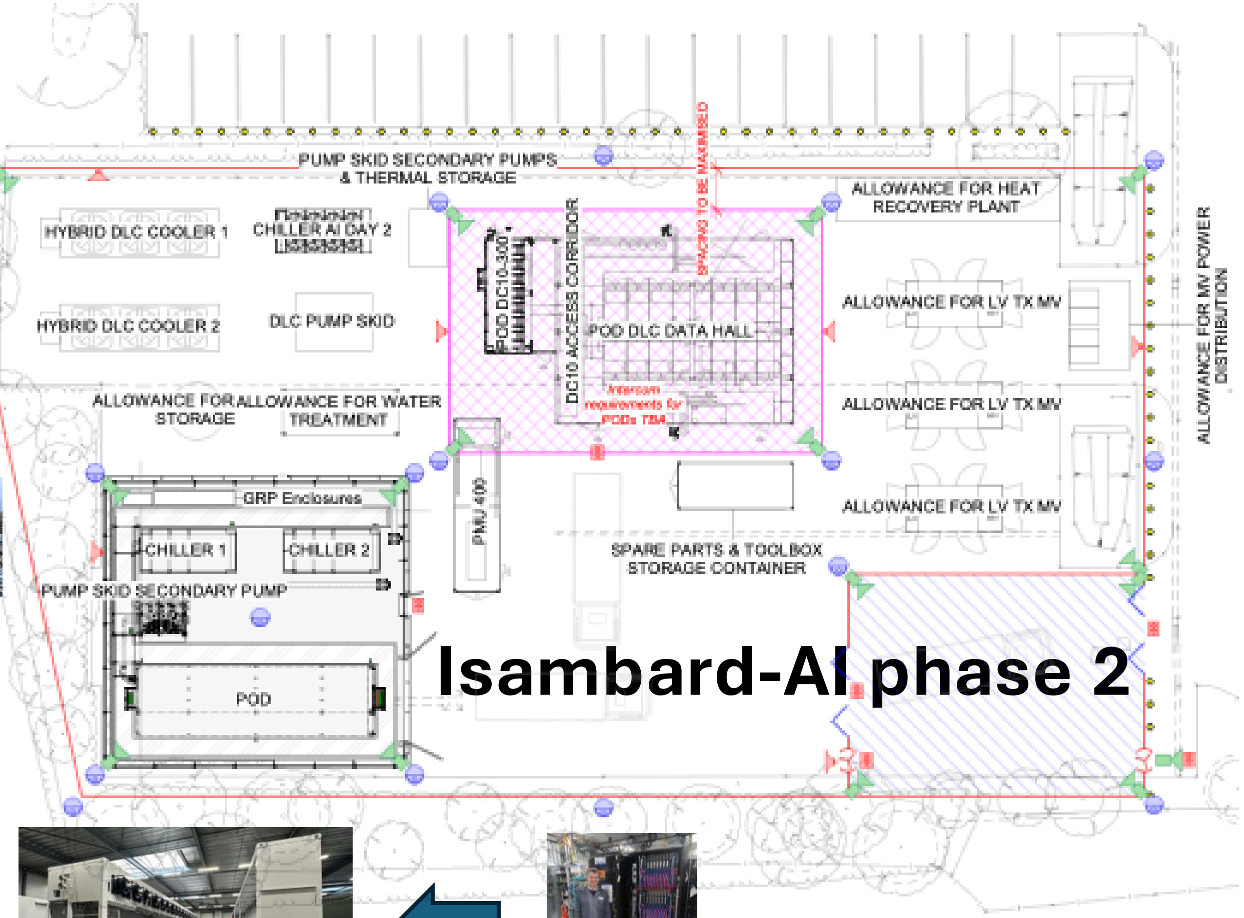
MDC craning (on a cloudy, windy day)

Chiller and pump skid

Utilities and network connection from NCC to MDC

University of BRISTOL

The Isambard Site



MDC before assembly

Isambard-AI phase 1 168 Grace-Hopper cabinet

Isambard-AI AIRR:

- Phase 1 arrived in March 2024
- Phase 2 arrives late summer (July)

- >£200M including Modular Data Centre, all cooling etc.
- >£300M total investment over 5 years
- 5MW
- Extremely energy efficient, PUE <1.1
- Almost entirely direct liquid cooling, plans to reuse the waste heat

University of BRISTOL

Sustainability as a Key Design Principle

- Optimisation targets
  - PUE = Power Usage Effectiveness
    - Target <1.1
  - CUE = Carbon Usage Effectiveness
    - Non-fossil fuel sources
  - Plan for heat reuse for nearby buildings and local district heat circuit in future
- Aligning with university of Bristol Net Zero and sustainability targets for 2030
  - Categorising emissions
    - Scope 1 (~0%), 2 (90%) and 3 (10%)—based on an average UK data of 0.2123 kg CO_2/kWh (IEA 2022 data)
  - Recycling 90% of components at the end of life in the UK

HPE EX Series DLC and GH 200

- HPE EX solution
  - Direct liquid cooling for high performance computing and networking
- 4-way Grace-Hopper superchip
  - NVLink-C2C also only uses 1.3 picojoules/bit transferred—5x more energy efficient than PCIe Gen 5

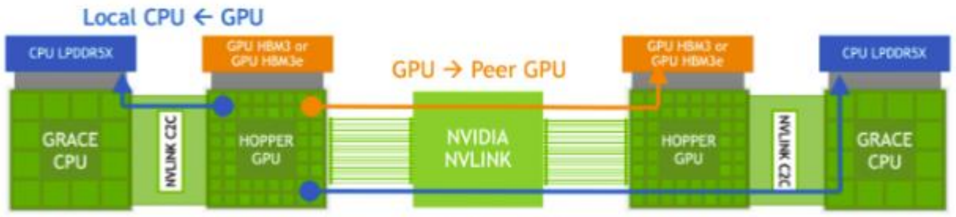


Figure 5. Memory Accesses across NVLink-connected Grace Hopper Superchips

Source: NVIDIA Grace Hopper Superchip Architecture Whitepaper

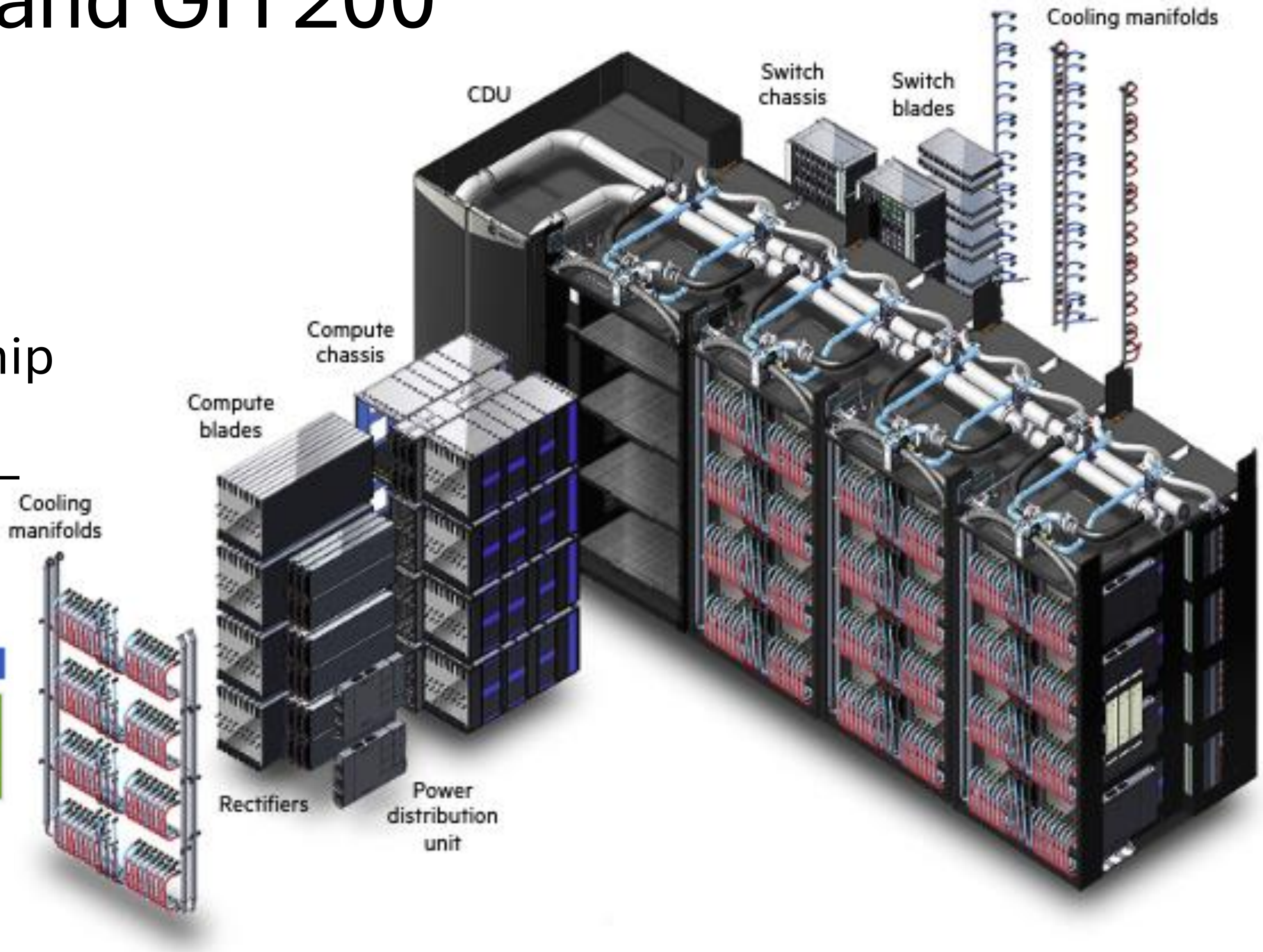
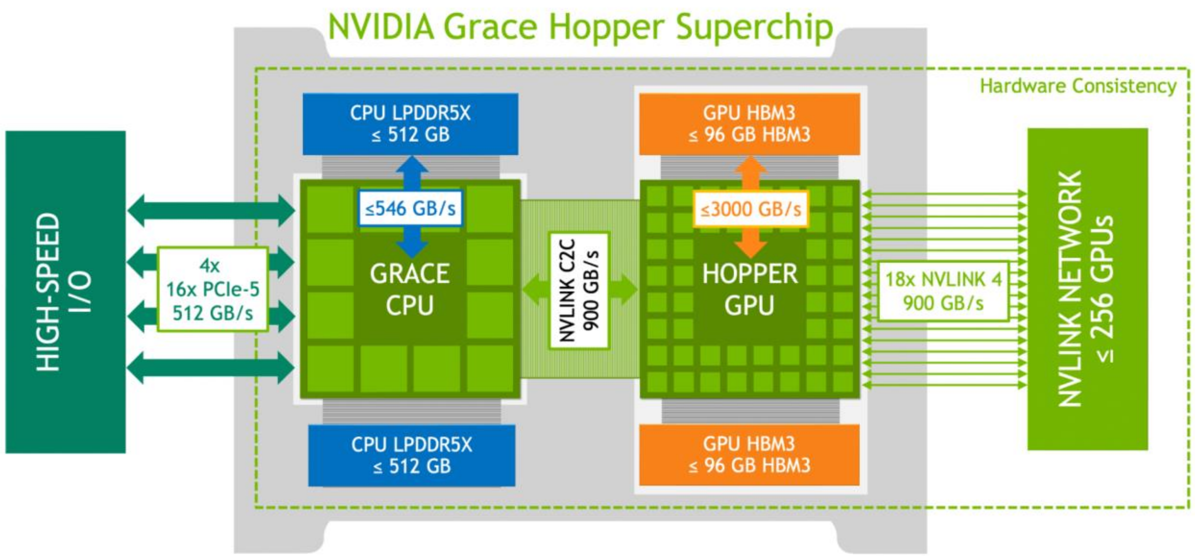


FIGURE 1. HPE Cray EX cabinet exploded view

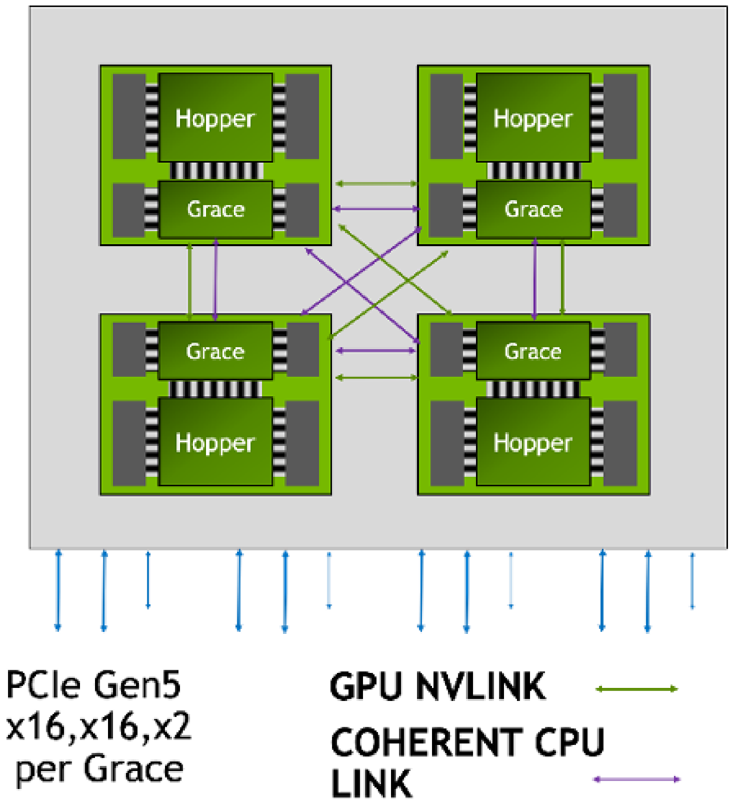
Source: HPE CRAY EX Liquid-Cooled Cabinet for Large Scale Systems brochure

University of BRISTOL

Grace-Hopper Superchip & HPE EX Compute Blade



Source: NVIDIA Grace Hopper Superchip Architecture Whitepaper



Source: HPE EX4000 Grace-Hopper blade

4 x Grace ARM CPUs
288 cores
512 GB Memory

4 x Hopper GPUs
~260 64-bit Tflops, ~16k 8-bit Tflops
384 GB High Bandwidth Memory

896 GB Memory Total
NVLink-C2C = 900 GB/s

Isambard-AI node = 4 x GH200
Injection bandwidth = 4 x 200 Gbps

University of BRISTOL

Memory Architecture of GH 200

A boon for developers and users

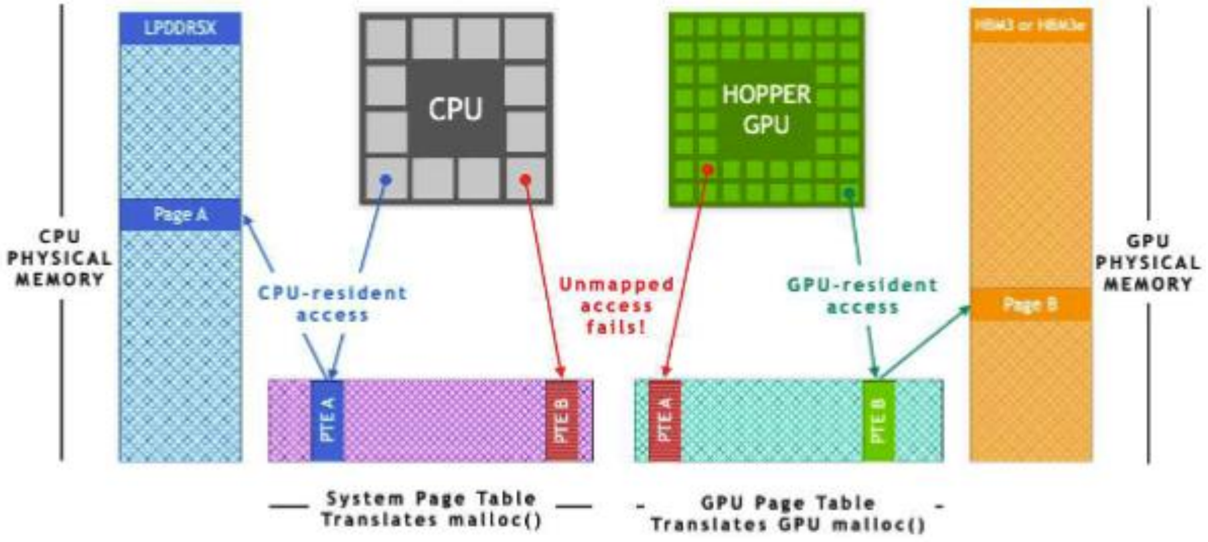


Figure 7. NVIDIA Hopper System with Disjoint Page Tables

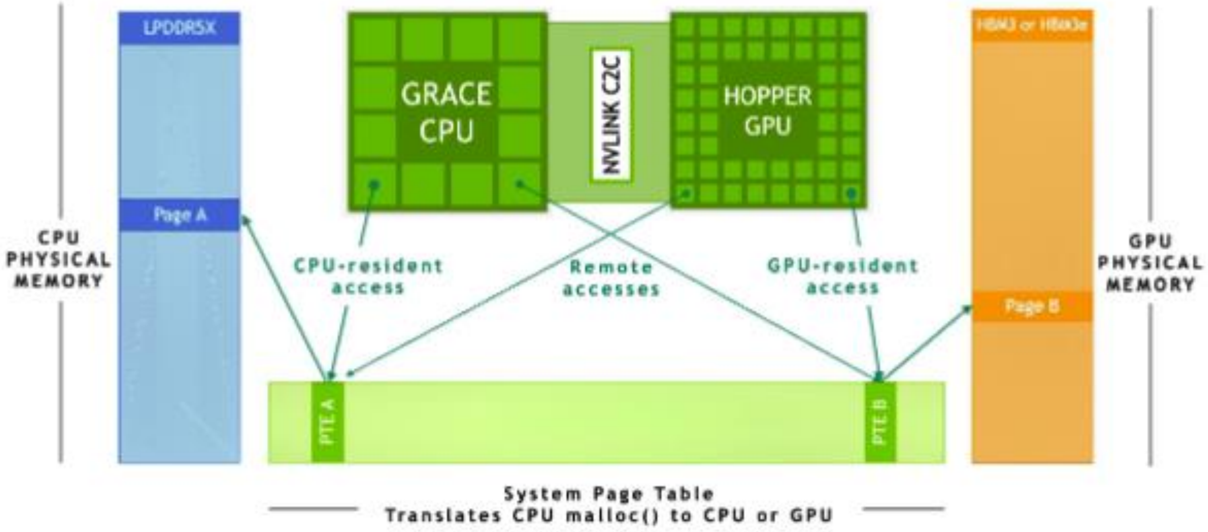


Figure 8. ATS in an NVIDIA Grace Hopper Superchip System

Source: NVIDIA Grace Hopper Superchip Architecture Whitepaper

In PCIe-connected x86+Hopper systems, the CPU and the GPU have independent per process page tables, and system allocated memory is not directly accessible from the GPU

Address Translation Service (ATS) enables the CPU and GPU to share a single per-process page table, enabling all CPU and GPU threads to access all system-allocated memory

HPE SlingShot High Speed Interconnect for AI & HPC

Ethernet	Slingshot	HPC Networks
Standards based / interoperable	Standards based / interoperable	Proprietary (single vendor)
Commodity technology	Commodity technology	Non-commodity technology
Converged network	Converged network	HPC interconnect only
Limited HPC features	Full set of HPC features	Full set of HPC features
High latency	Low latency	Low latency
Efficient for large payloads only	Efficient for small to large payloads	Efficient for small to large payloads
Limited scalability for HPC	Very scalable for HPC & Big Data	Very scalable for HPC & Big Data

Source: https://www.nextplatform.com/2019/08/16/how-cray-makes-ethernet-suited-for-hpc-and-ai-with-slingshot/

Liquid cooled interconnect (sustainability & scalability)

Example with 16-switch group

2 switches per chassis for single injection to 32 compute nodes (8 compute blades)

6 switches per cabinet for single injection to 256 compute nodes (64 compute blades)

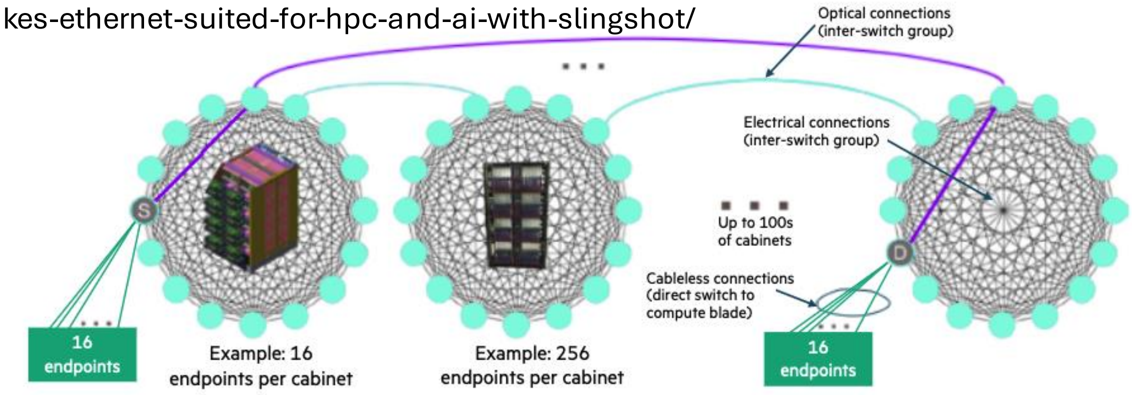
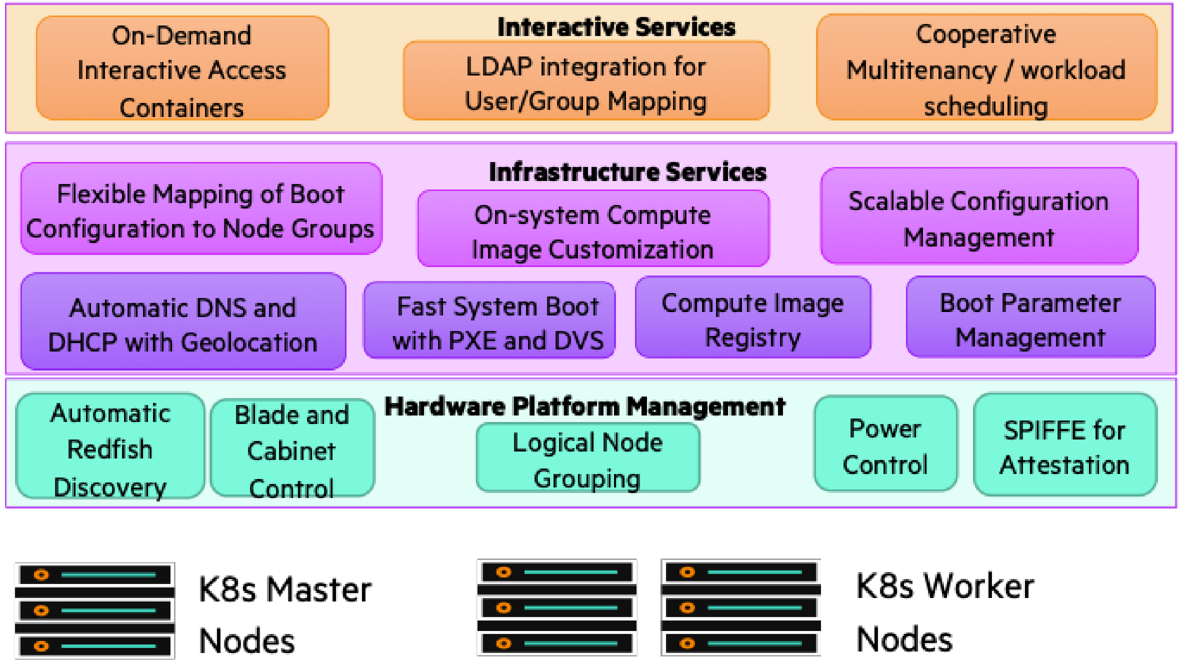


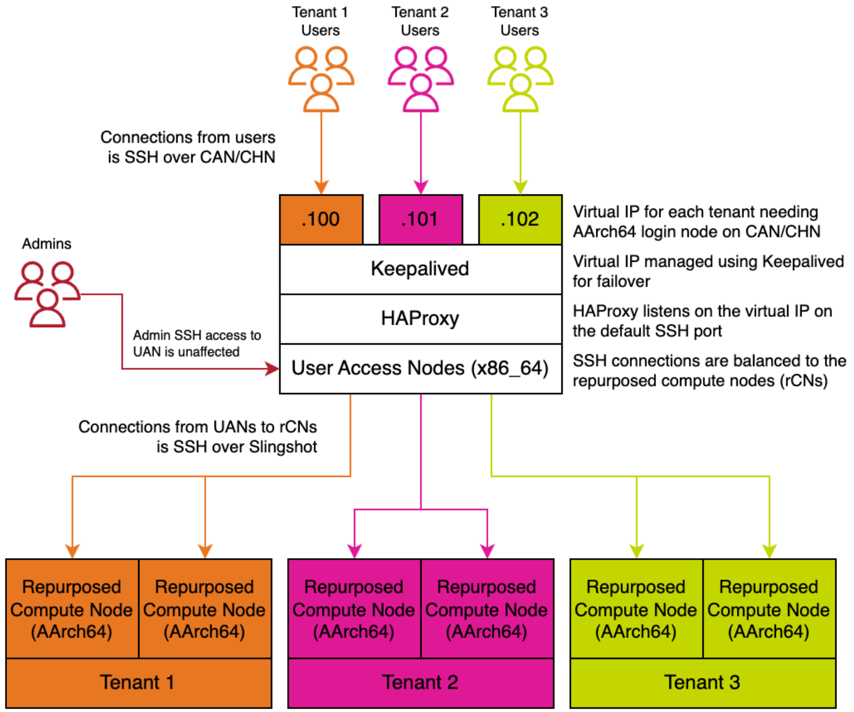
FIGURE 7. Example of Dragonfly topology in HPE Slingshot switches

Source: HPE CRAY EX Liquid-Cooled Cabinet for Large Scale Systems brochure

High performance, cloud-native software ecosystem for AI platforms



CSM provisioning system

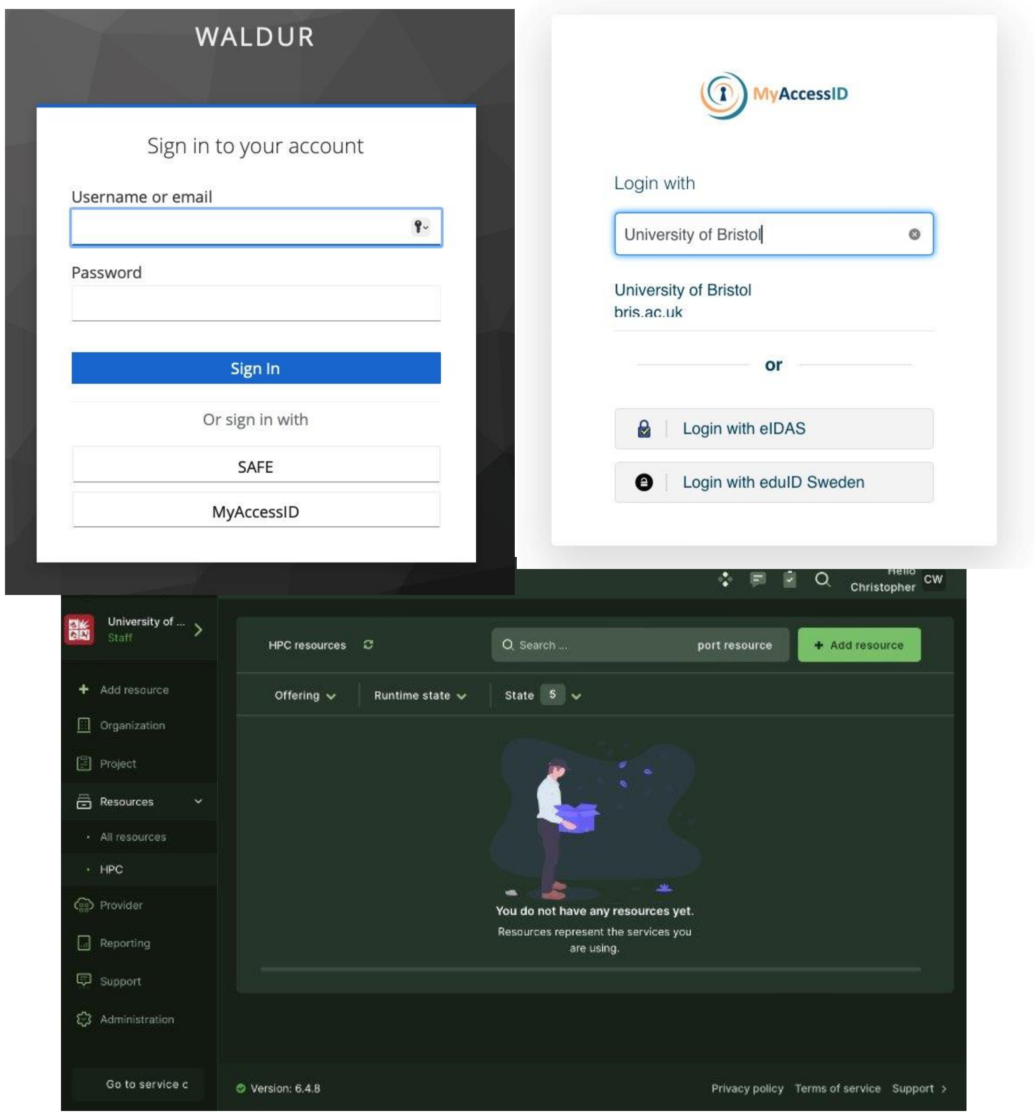


Multi-tenant setup

University of BRISTOL

Lowering Access Barriers

- Single sign-on access with federated academic and research credentials
  - MyAccessID common identity layer
  - Supported by JISC UK Access Management Federation (UKAMF)
  - Security via multi-factor authentication for different service layers (web & ssh)
- Self-service, cloud-native user and project management portal
- Single pane of glass for accessing all services on Isambard AI including accounting and reporting

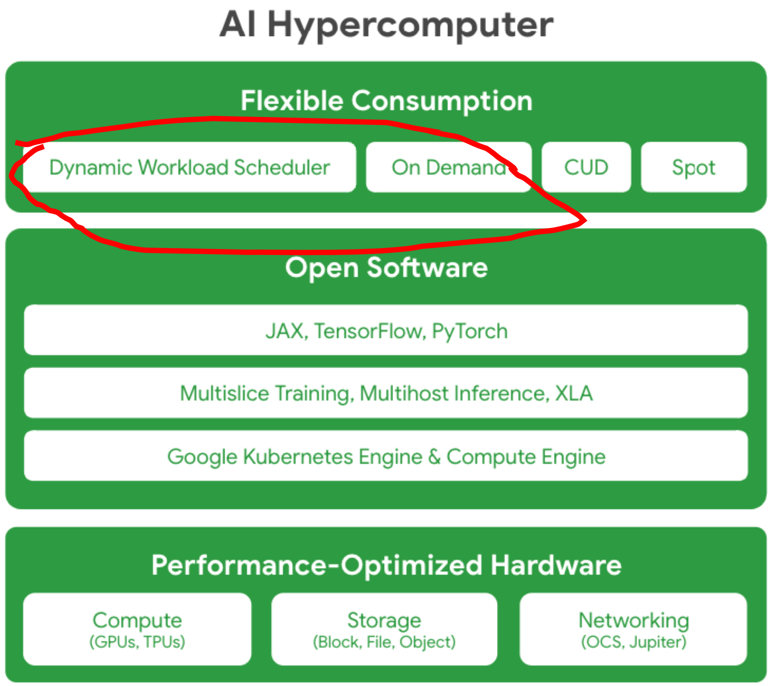


Source: Federated access to Isambard-AI - Requirements for attribute release by IDPs-- a joint presentation at March'24 HPC SIG by university of Bristol and JISC colleagues

University of BRISTOL

Improving User and Usage Productivity

- Interactive compute with JupyterHub on Isambard AI
  - Supporting NGC (Nvidia GPU Cloud) containers
- Supporting common usage patterns for AI
  1. Experimentation: shorter period with large capacity, interactivity and steering needs
  2. Training: longer periods, days to months on large capacities
  3. Fine tuning: shorter period with low capacity
- Leveraging Google's AI computing approaches
  - Flex Start with guaranteed completion aka HPC/slurm batch—with periodic checkpoint/restart (2)
  - Calendar with start/stop time aka HPC/slurm reservation—needs automation (1, 3)
    - Simpler to implement for 1 to 1000s of GPUs due to tightly integrated Isambard AI computer, high performance network & storage resources (unlike public cloud)

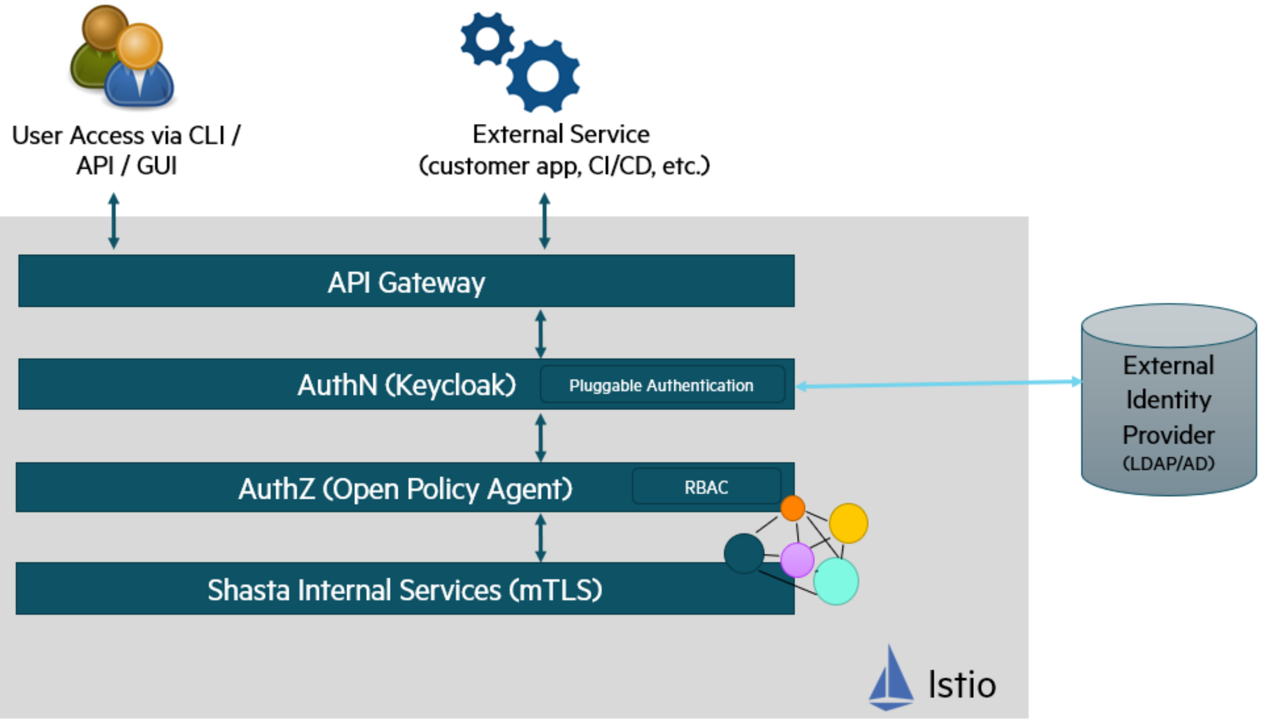


Source: Google AI Hypercomputer

University of BRISTOL

Cybersecurity and RBAC (roles based access controls)

- UK legislative requirements for a comprehensive view of AI RR available to academic and research communities for open research in AI safety
- Access models and threat models being defined
- CSM RBAC for API and CLI
    - TAPMS for multi-tenancy
    - SlingShot isolation and VNI tagging
    - VAST Software Defined Storage alongside ClusterStor



CSM Identity and Access Management (IAM)

University of BRISTOL

	Isambard-AI Phase 1	Isambard-AI phase 2	Isambard 3
Total nodes & superchips	42 nodes / 168 Nvidia GH200 superchips	1,320 nodes / 5,280 Nvidia GH200 superchips	384 nodes / 384 Nvidia Grace superchips
Compute per node	288 Grace cores, 4 H100 GPUs	288 Grace cores, 4 H100 GPUs	144 Arm Neoverse v2 cores
Memory per node	512 GB LPDDR + 384 GB HBM3	512 GB LPDDR5 + 384 GB HBM3	256 GB LPDDR5
Network injection per node	4 x 200 Gb/s Slingshot	4 x 200 Gb/s Slingshot	1 x 200 Gb/s Slingshot
Storage	~1 PB all-flash ClusterStor E1000 Lustre	~20.3 PB all-flash ClusterStor E1000 Lustre, ~3.56 PB VAS	~2.3 PB ClusterStor E1000 Lustre
Programming environment	Cray PE / NVIDIA HPC SDK / Arm compilers	Cray PE / NVIDIA HPC SDK / Arm compilers	Cray PE / NVIDIA HPC SDK / Arm compilers
Types of compute racks	Cray EX2500	Cray EX4000	Cray XD2000
Number of compute cabinets	1	12	6
MDC POD	DC20 (modified for EX2500)	DC10 with air cooled parts	DC20 air cooled

University of BRISTOL

Scaling out in two phases

Phase 1 (~0.7 8-bit AI Exaflops)

Arrived in March 2024 –in Isambard 3 MDC

1 x DLC EX2500 cabinet

42 blades (4-way Grace-Hopper)

168 GH superchips

12,096 Neoverse V2 Armv9 CPU cores

168 Hopper GPUs

21.5 TB CPU memory

16.1 TB high bandwidth GPU memory

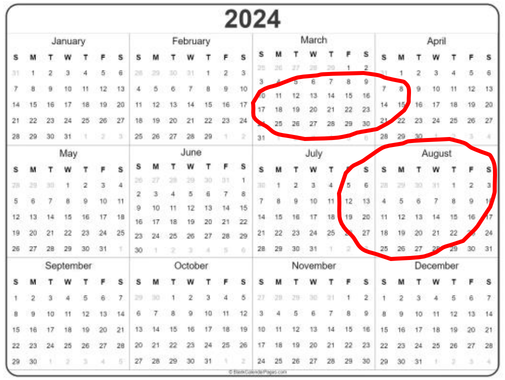
37.6 TB total memory

AI high performance storage

~1 PB all-flash ClusterStor Lustre

Phase 2 (~21 8-bit AI Exaflops)

Arriving Summer 2024 – new Isambard-AI MDC



12 x DLC EX4000 cabinets

660 blades (4-way Grace-Hopper)

5,280 GH superchips

380,160 Neoverse V2 Armv9 CPU cores

5,280 Hopper GPUs

675 TB CPU memory

506 TB high bandwidth GPU memory

1.18 PB total memory

AI high performance storage

~27 PB all-flash storage!

(~20 PB Lustre, ~7 PB software defined VAST)

University of BRISTOL

Thank you

Stay tuned!

University of BRISTOL