FUTURE TRENDS IN HPC (AND AI)

Is the future harmony or discord?

Professor Mark Parsons

EPCC Director
Dean of Research Computing

Introduction

- To understand the future we have to understand today

- The AI hype curve is having many unintended consequences

- This talk will look at

  - The current National HPC Service in the UK – ARCHER2

  - A key development that complements it – the Edinburgh International Data Facility

  - Discuss if HPC and AI are converging … or diverging

  - Consider what has gone wrong with the UK Exascale project … and what we can learn
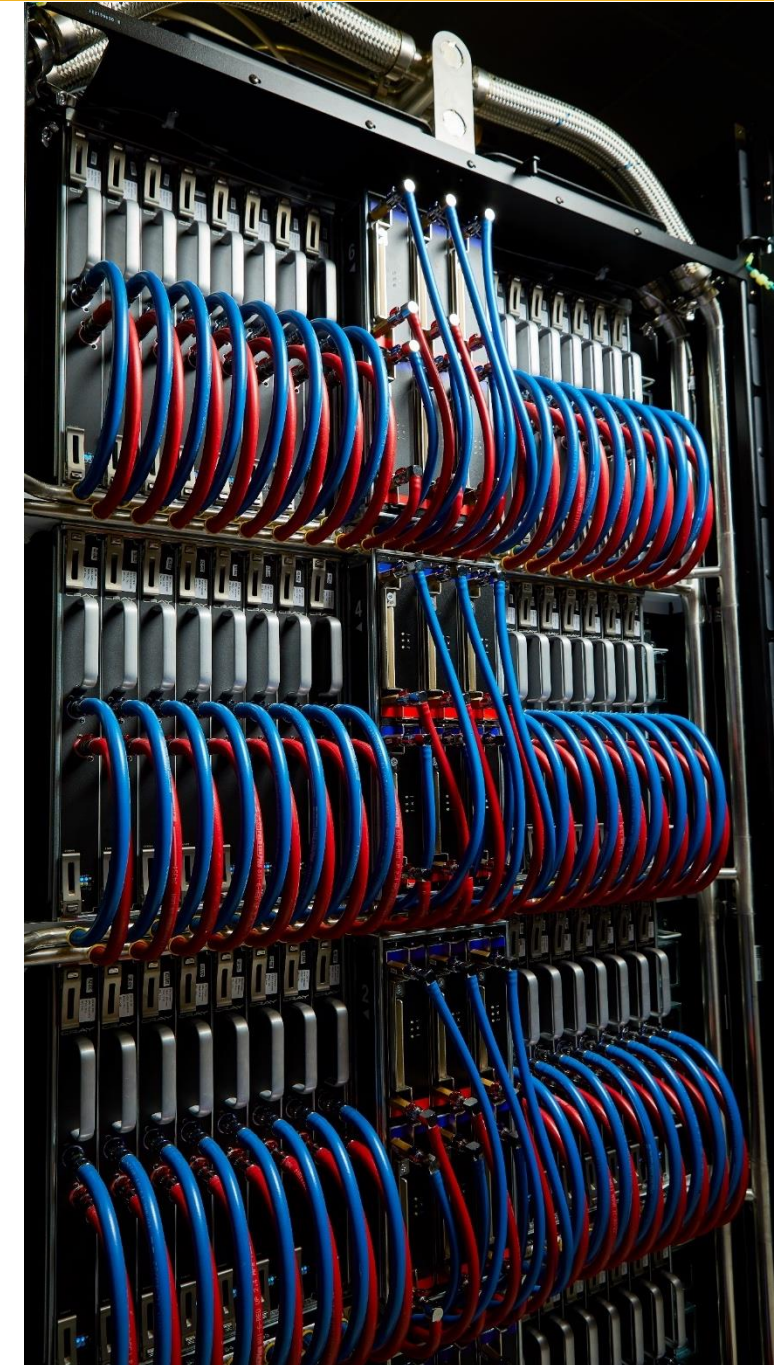
epcc

EPCC in 2024

49,740 Students
18,801 Staff
(12,394 FTE)

- Part of the University of Edinburgh

- Established in 1990 now with ~150 staff and ~110 students

- UK HPC National Service provider

- ~ £150m of HPC and data science services

- Wide range of research activities from **Supercomputing** to **AI** to **Data Science**

- Hosts **Edinburgh International Data Facility** as part of Data Driven Innovation programme

- Chosen site for **UK Exascale** system

ARCHER2

- HPE Cray EX Supercomputer – Serial no. 1!

- 5,860 compute nodes (750,080 CPU compute cores)

- HPE Slingshot 10 interconnect

- Compute nodes:
  - Dual AMD EPYCTM 7742 Processors, 64 cores, 2.25 GHz
  - 256 GiB / 512 GiB memory per node
  - Two 100 Gbps HPE Slingshot interfaces per node

- 16 PB storage – ClusterStor Lustre

- No 22 in Top 500 Nov 2021 – 19.5 Petaflop/s HPL

- Busy from Day 1 – and has remained busy

- Over 4,992 users since opening

ARCHER2

- HPE Cray EX Supercomputer – Serial no. 1!

- 5,860 compute nodes (750,080 CPU compute cores)

- HPE Slingshot 10 interconnect

- Compute nodes:
  - Dual AMD EPYCTM 7742 Processor
  - 256 GiB / 512 GiB memory per node
  - Two 100 Gbps HPE Slingshot interfaces per node

- 16 PB storage – ClusterStor Lustre

- No 22 in Top 500 Nov 2021 – 19.5 Petaflop/s HPL

- Busy from Day 1 – and has remained busy
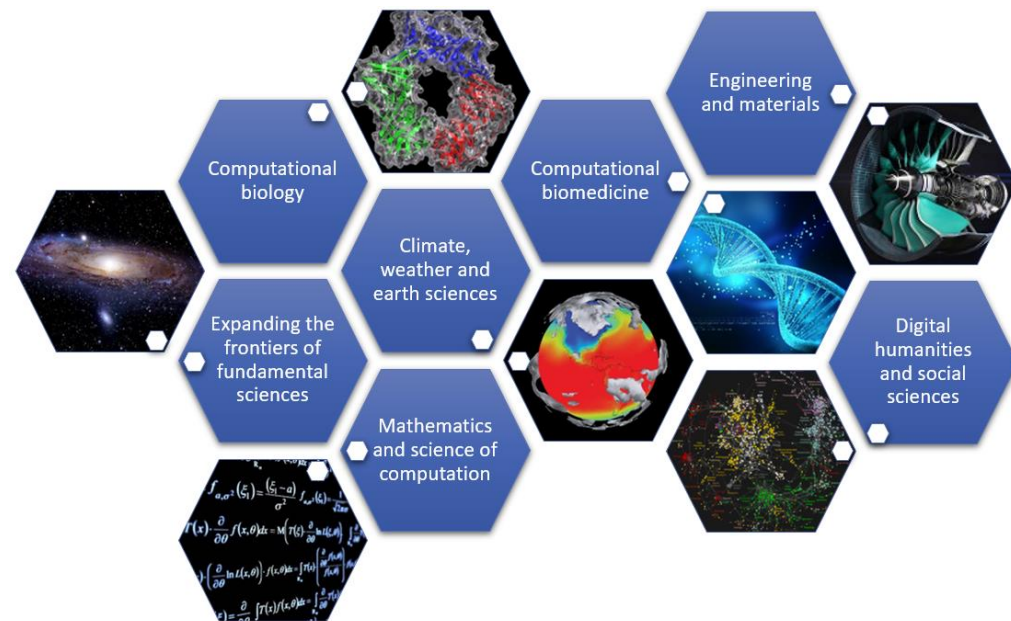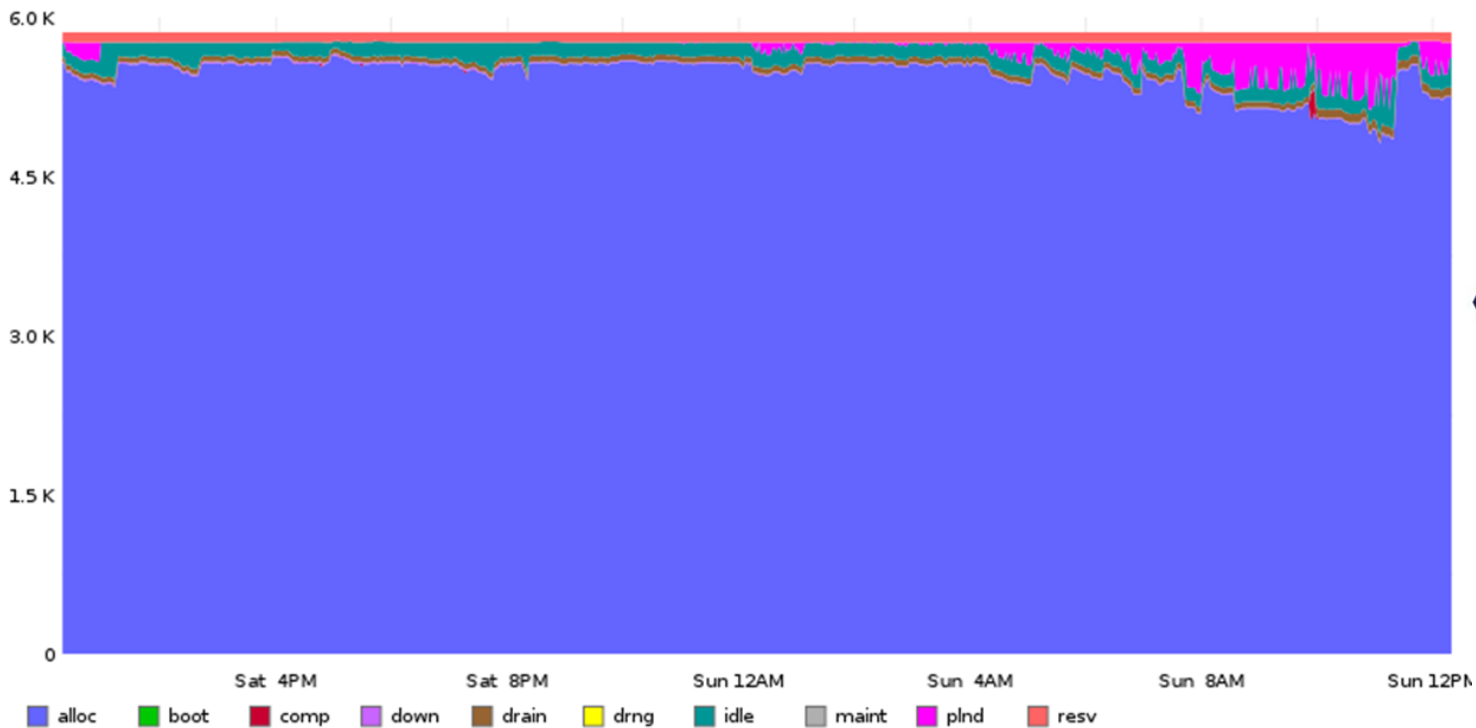
- Over 4,992 users since opening
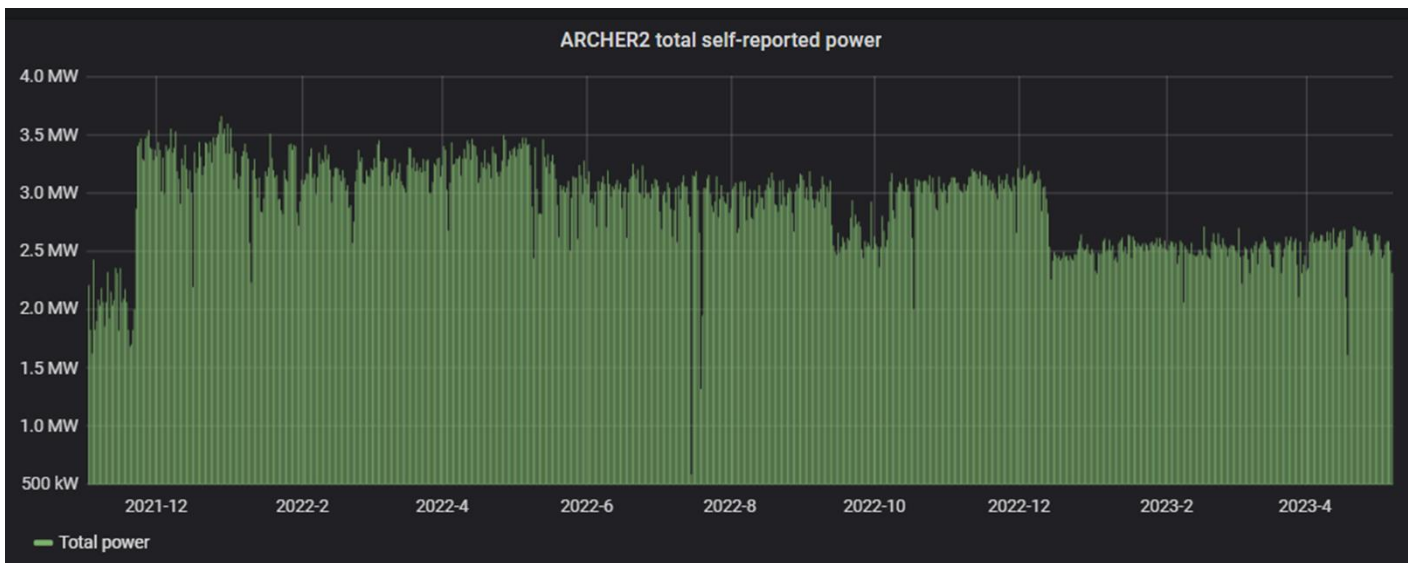
6th largest cores-only machine in the world

Already getting old … 49th in the world in June 2024
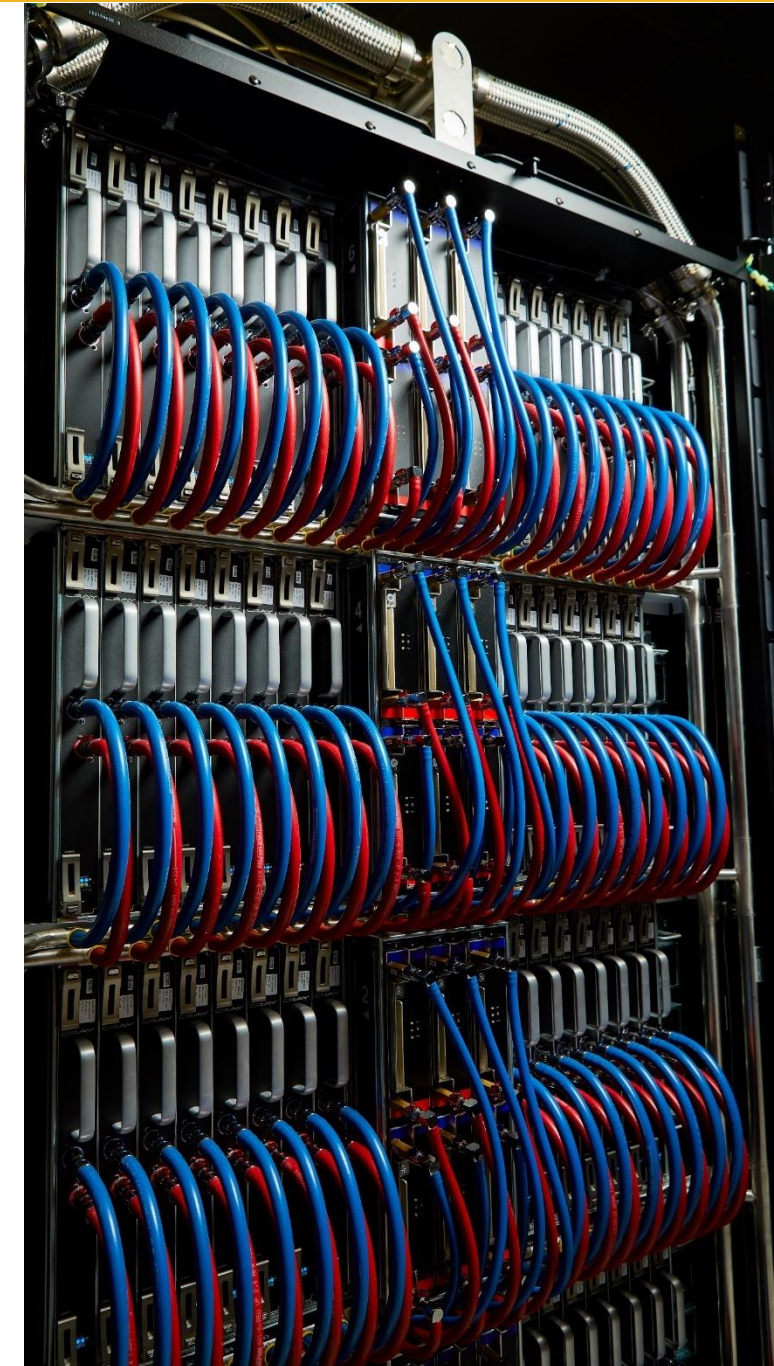
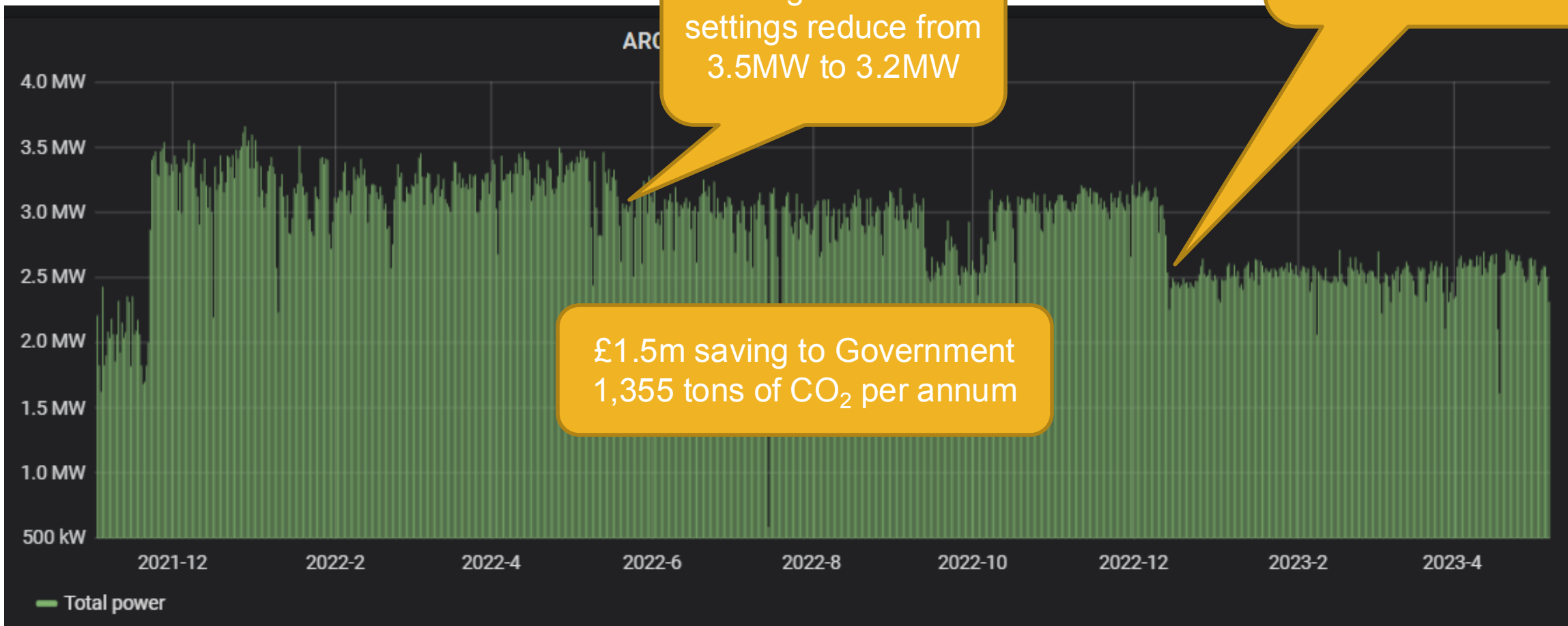# ARCHER2 is full of science …

This is the most consistently busy system EPCC has ever brought into operation

ARCHER2 total self-reported power

ARCHER2 continuous improvement in total power usage

Changes to node settings reduce from 3.5MW to 3.2MW

CPU frequency from 2.2GHz to 2.0GHz reduces to 2.7MW

£1.5m saving to Government 1,355 tons of CO_2 per annum

epcc

Edinburgh's Data Driven Innovation programme

- Capitalise on our academic excellence in Data Driven Innovation

- Create a trusted public-private-third sector partnership

- Unlock **economic opportunities** worth **£5 billion+**

- **Train 100,000 people** in data technologies

- Develop an underpinning infrastructure – the **Edinburgh International Data Facility** – a £100m data facility
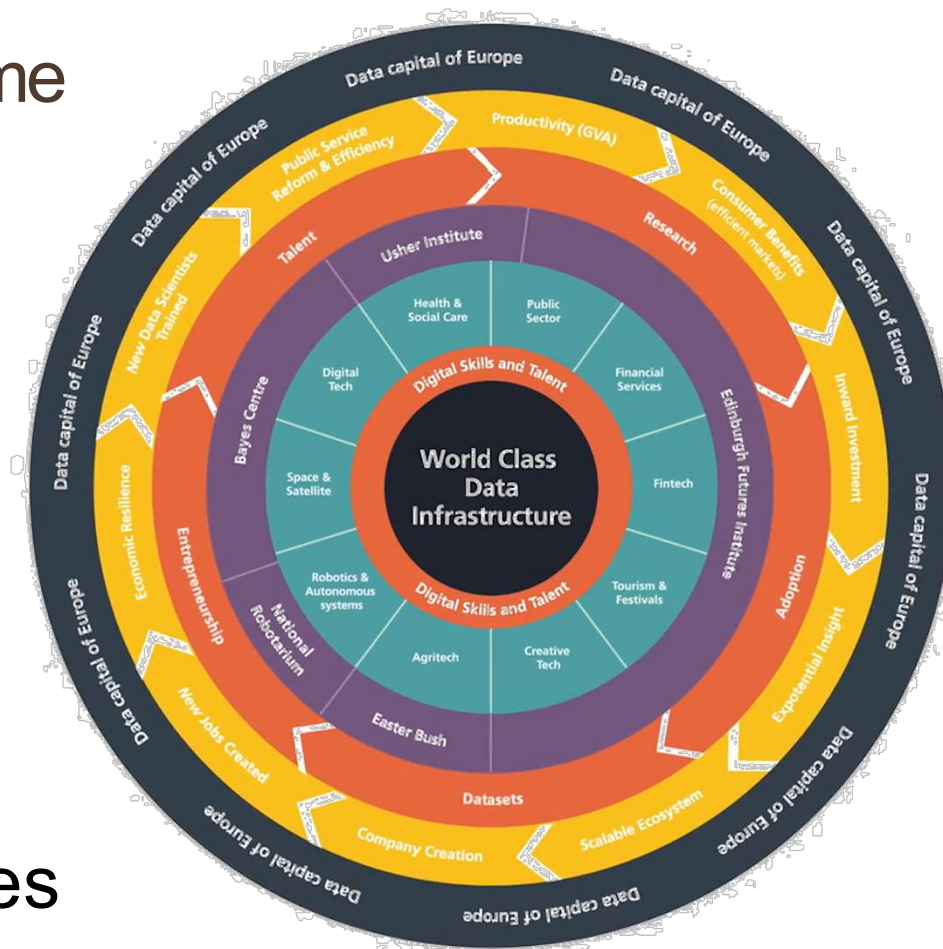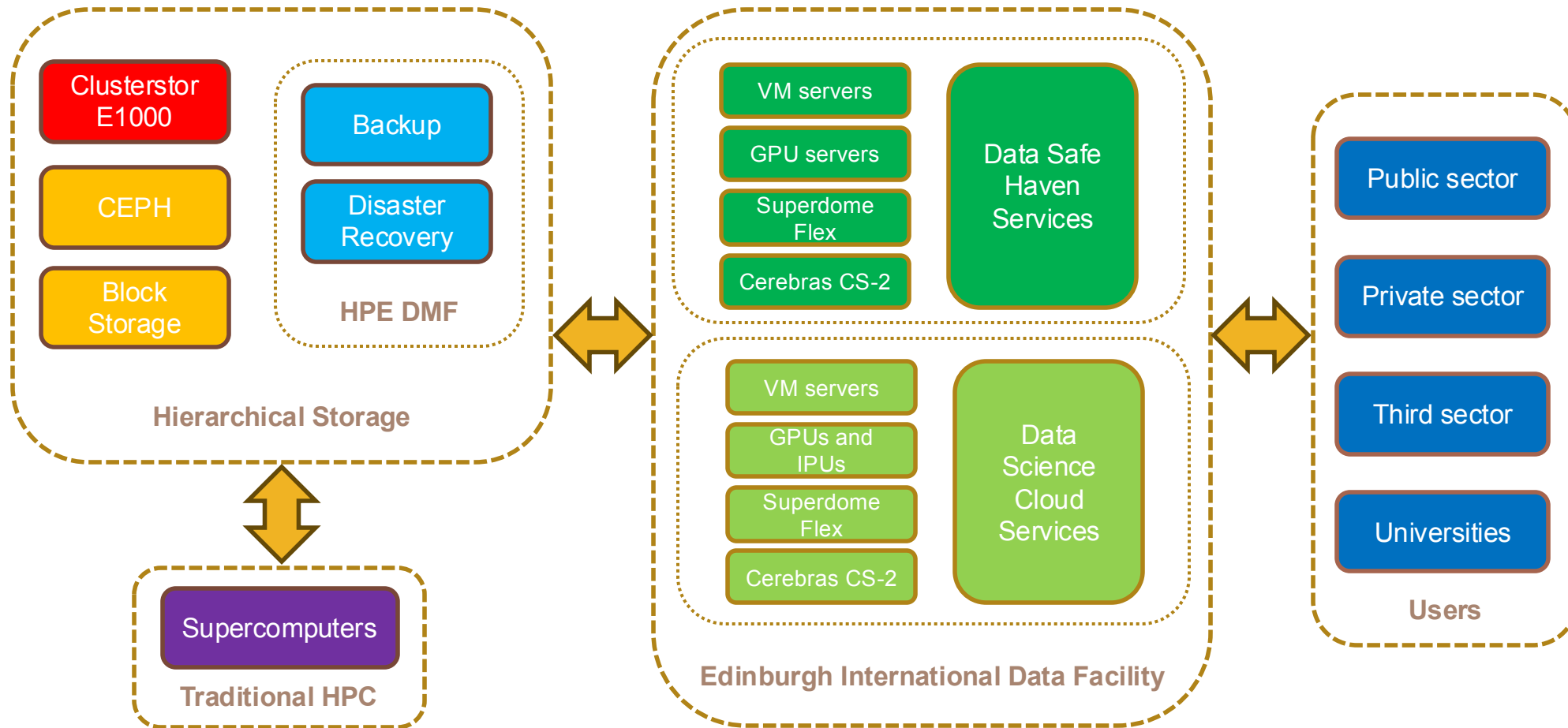
Designed for Data Science and AI projects and partnerships
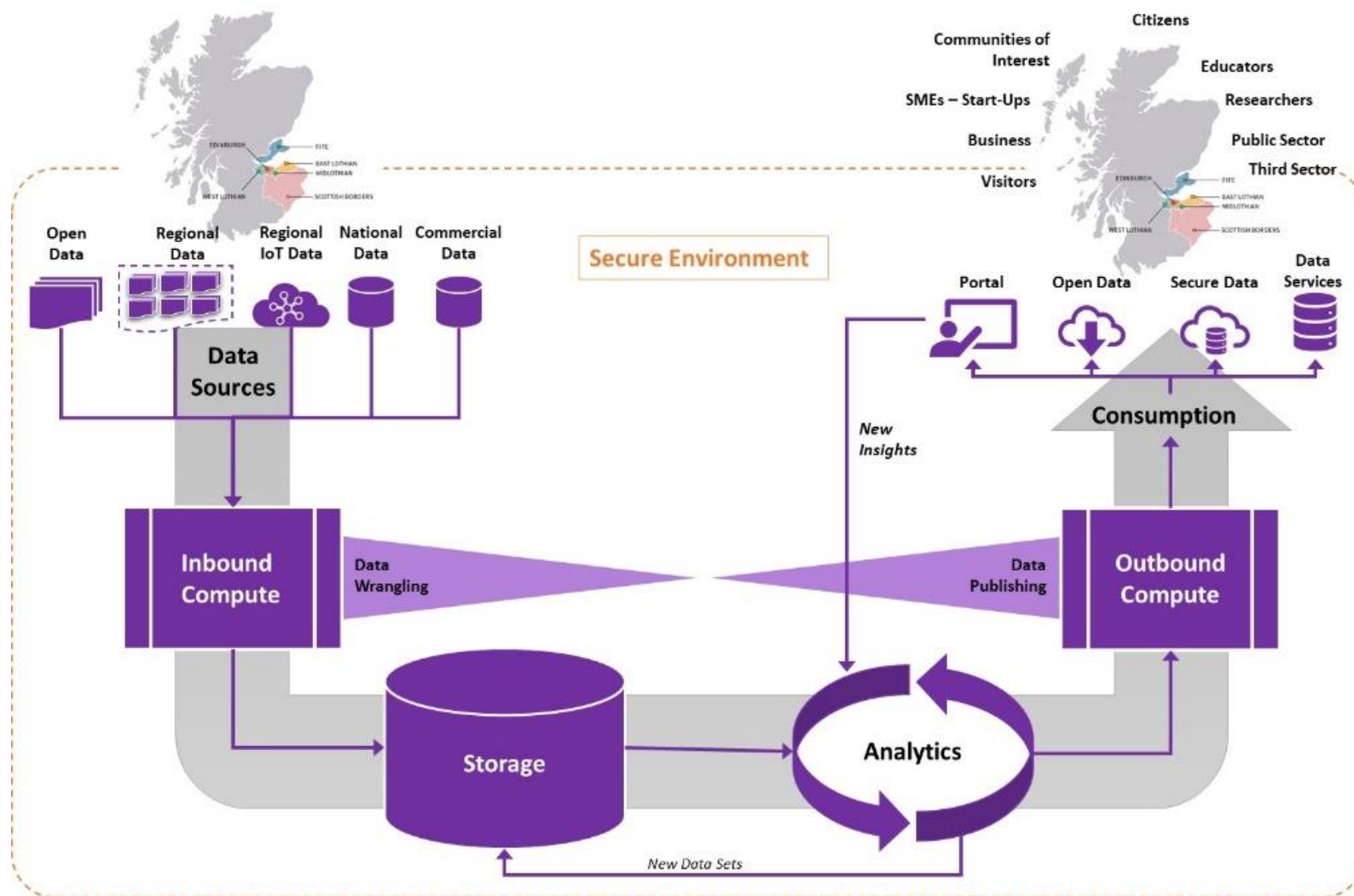
- Create a trusted **public-private-third sector** partnership

- Unlock economic opportunities worth **£5 billion+**

- Train **100,000 people** in data technologies

Edinburgh International Data Facility

Hierarchical Storage
- Clusterstor E1000
- CEPH
- Block Storage

HPE DMF
- Backup
- Disaster Recovery

Traditional HPC
- Supercomputers

Edinburgh International Data Facility
- VM servers
- GPU servers
- Superdome Flex
- Cerebras CS-2
- Data Safe Haven Services
- VM servers
- GPUs and IPUs
- Superdome Flex
- Cerebras CS-2
- Data Science Cloud Services

Users
- Public sector
- Private sector
- Third sector
- Universities

epcc

Beyond GPUs for AI …

- EPCC installed Europe's first Cerebras CS-1 in March 2021

- Then upgraded to two CS-2s
  - 850,000 AI cores for sparse linear algebra
  - 40GB on-chip SRAM – 20PB/s memory bandwidth
  - 220 Pb/s internal interconnect
  - 1.2Tb/s ethernet connection to host
  - AI focussed software stack

- PyTorch and Tensorflow etc supported

- Includes SwarmX / MemoryX cluster

21 cm
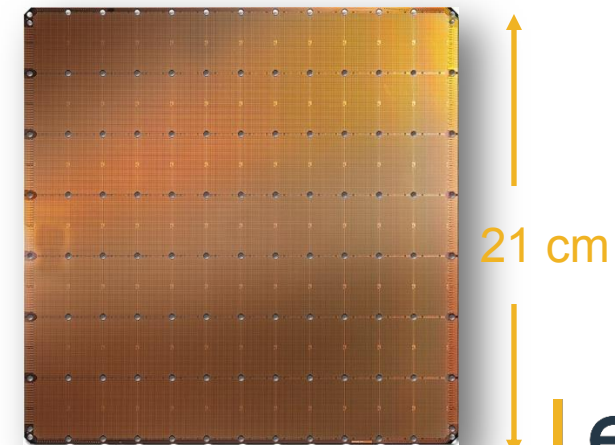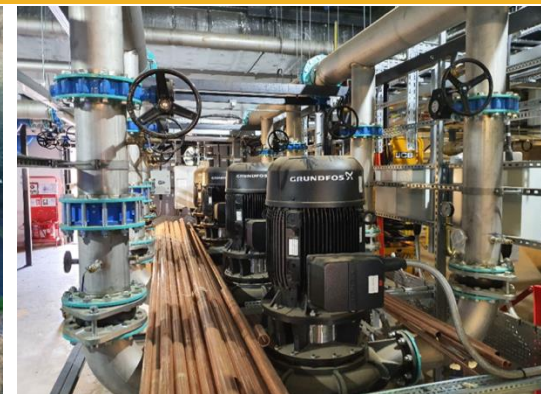
Computer Room 4

£20m – CR 4 + PR D
£8.6m – 30MVA additional power
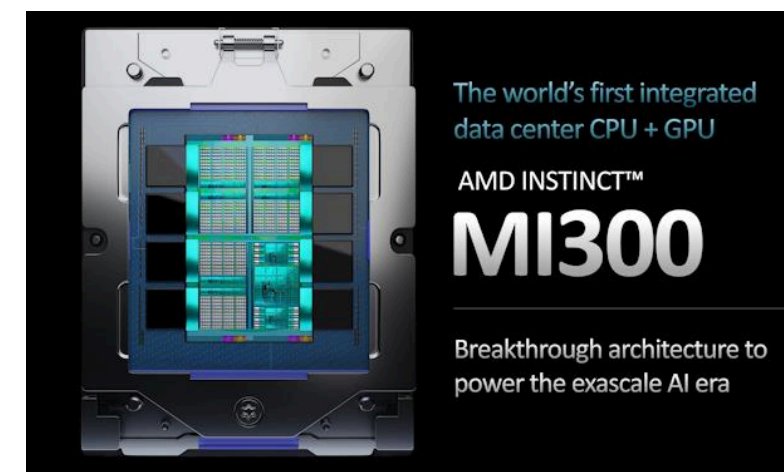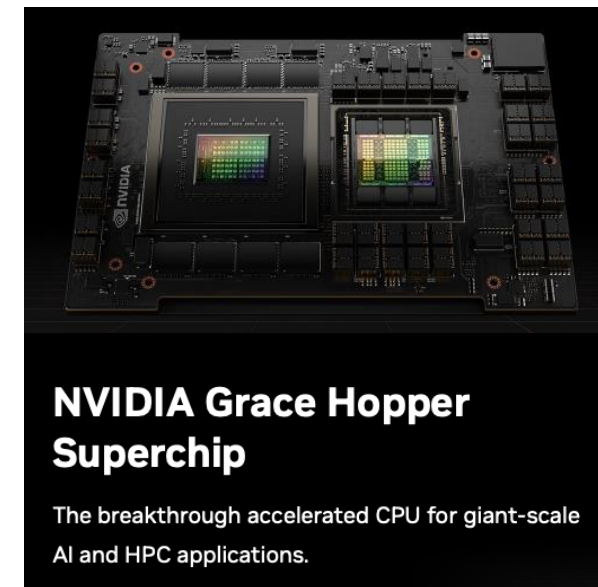Space for 270 standard racks

Opened Dec 2020

6MW of cooling water and power distribution in CR4

Exascale supercomputing in 2025

- Systems constructed from 6,000 nodes (servers)

- Each server has four GPUs and normal CPU cores

- Total of 24,000 GPUs in total

- Exascale supercomputers excel at **modelling and simulation** … but they also excel at **AI**

- All this comes at an electricity cost – at full speed an Exascale system uses approx. **22 Megawatts**

NVIDIA Grace Hopper Superchip

The breakthrough accelerated CPU for giant-scale AI and HPC applications.

The world's first integrated data center CPU + GPU

AMD INSTINCT™

MI300

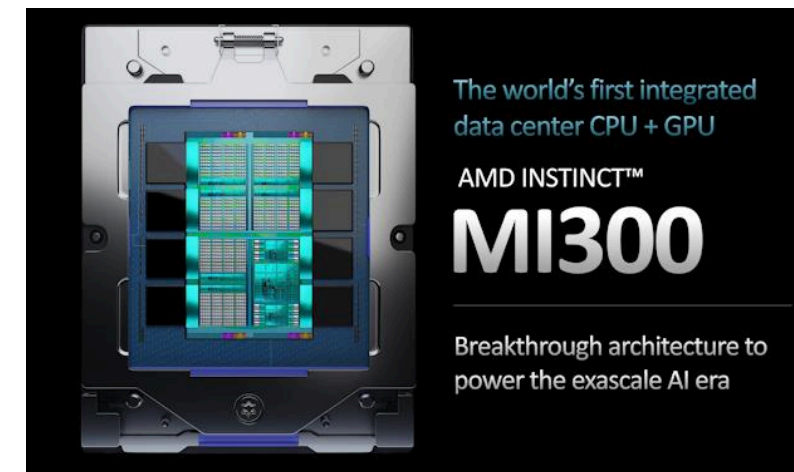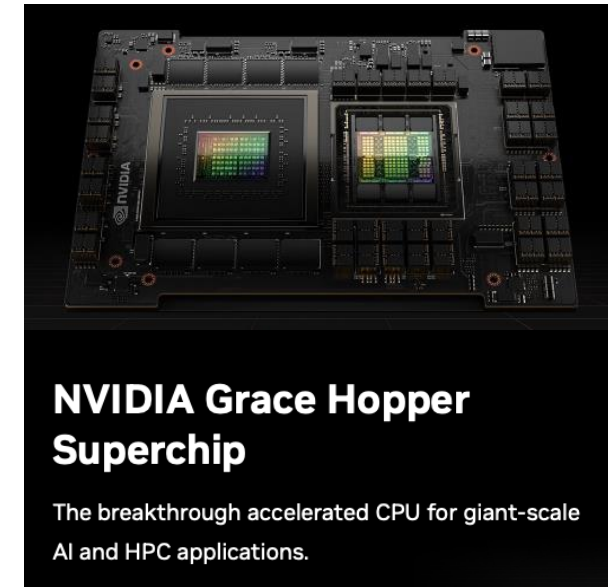Breakthrough architecture to power the exascale AI era

| epcc |

Exascale supercomputing in 2025

- Systems constructed from 6,000 nodes (servers)

- Each server has four GPU

- Total of 24,000 GPUs in t

- Exascale supercomputer
  simulation ... but they al

UK Government announced in 2023 that UK Exascale system will be hosted by EPCC at the ACF Data Centre

- All this comes at an electricity cost – at full speed an Exascale system uses approx. **22 Megawatts**

NVIDIA Grace Hopper Superchip

The breakthrough accelerated CPU for giant-scale AI and HPC applications.

The world's first integrated data center CPU + GPU

AMD INSTINCT™

MI300

Breakthrough architecture to power the exascale AI era

epcc

A modern Exascale supercomputer node

- 4 combined CPU+GPU modules (e.g. NVIDIA GH200 or AMD MI300A)

- Each CPU is multi-core – up to 300 cores per server

- Shared coherent memory – all data visible to all processors

- Connected by very fast internal network

- 4 network links to high performance interconnect

- **An Exascale supercomputer will have approx. 6,000 nodes**

4 external network links

CPU GPU CPU GPU
CPU GPU CPU GPU

Internal network

Single address space – shared memory

epcc
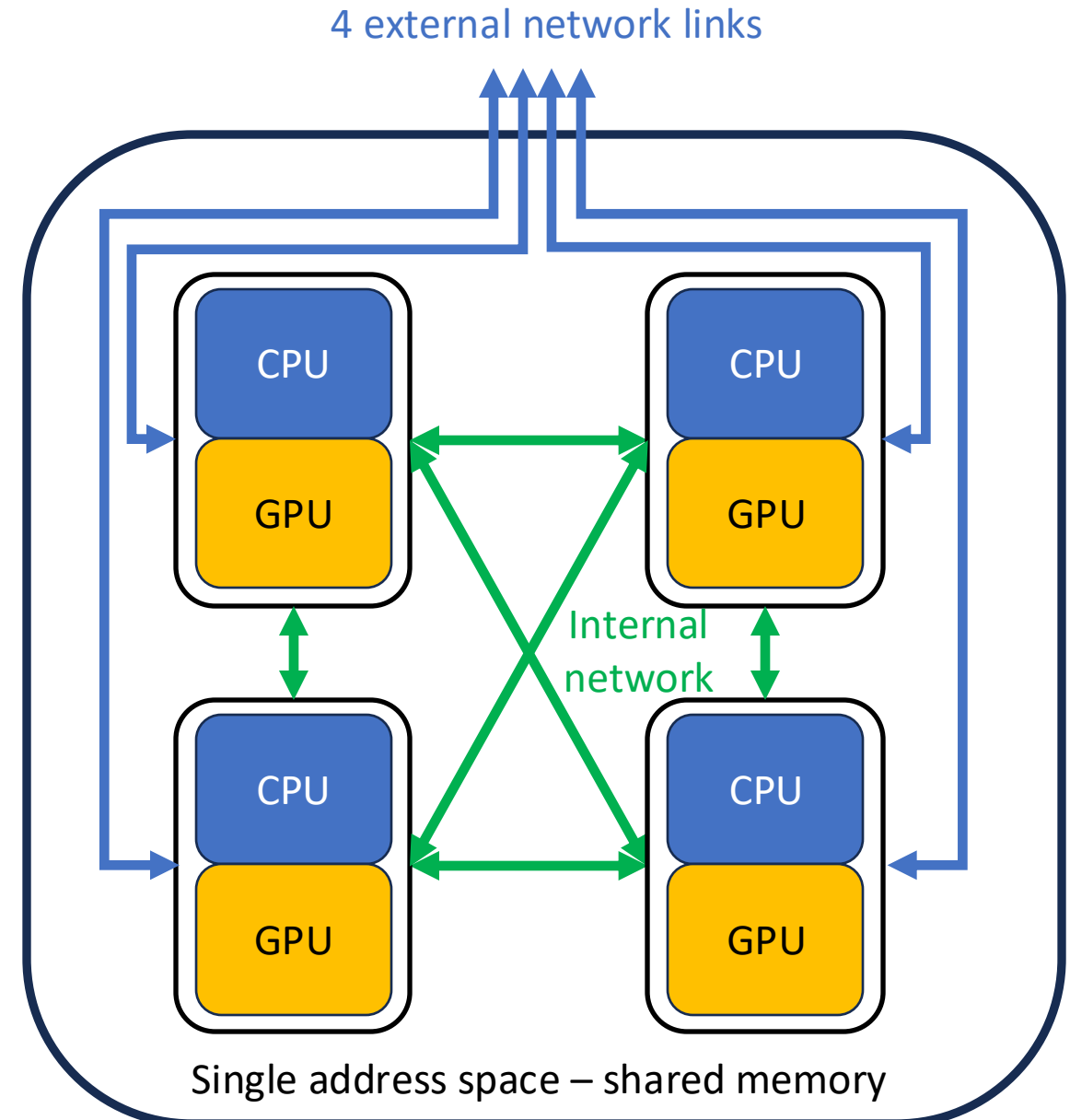
A modern 8-way GPU server for AI

- State-of-the-art GPU servers for AI have 8 GPUs (e.g. NVIDIA H100s)

- Very fast internal network

- Either one or two CPUs to control the GPUs

- 8-14 CPU cores per GPU

- External network is fast Ethernet or a high-performance interconnect (e.g. Infiniband)

- **AI clusters made up of multiple servers – e.g. 625 servers for a large 5,000 GPU system**
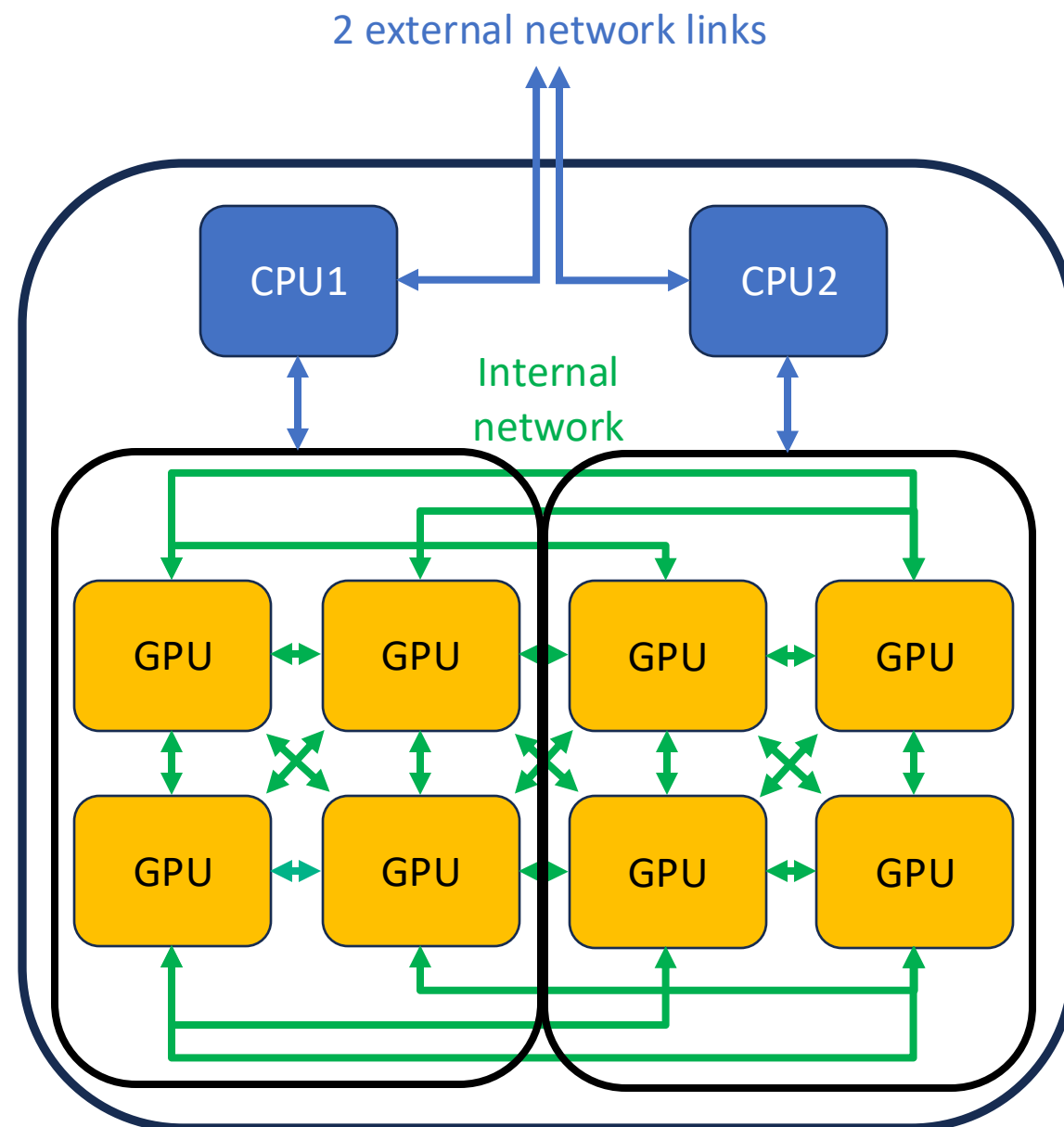
2 external network links

CPU1

CPU2

Internal network

GPU   GPU   GPU   GPU

GPU   GPU   GPU   GPU

epcc

Comparing AI and Exascale systems

Similarities

- Both built from CPUs and GPUs

- Multiple servers (nodes) networked together

- GPUs for AI and Exascale are (almost) identical

- Both have also have CPUs (but use them differently)

- Both can run similar workloads … but
  - AI specialises in AI workloads
  - Exascale specialises in modelling and simulation workloads

- Large-scale AI and Exascale have similar data storage needs

Differences

- Largest AI training problems scale to approx. 3-4,000 GPUs

- But majority of AI projects use a single 8-way server

- Largest Exascale projects scale to 24,000 GPUs – spread of job sizes in reality

- AI focussed on rich software environment – virtualisation / containerisation / Kubernetes etc

- Exascale more traditional software environment for simulation

- In **Exascale** the **CPU takes part in the calculation** with the GPUs

- In **AI** the **CPU is the "conductor of the orchestra"** of GPUs

- Exascale has a more expensive, complex external network between servers

epcc

Comparing AI and Exascale systems

Similarities

- Both built from CPUs and GPUs

- Multiple servers (nodes) networked together

- GPUs for AI and Exascale are (alm~~~~

- Both have also have CPUs (but us~~~~

- Both can run similar workloads …~~~~
  - AI specialises in AI workloads
  - Exascale specialises in modelli~~~~ workloads

- Large-scale AI and Exascale have similar data storage needs

Differences

- Largest AI training problems scale to approx. 3-4,000 GPUs

- But majority of AI projects use a single 8-way server

- ~~~~le projects scale to 24,000 GPUs – spread of ~~~~

- ~~~~ch software environment – virtualisation / ~~~~ Kubernetes etc

- ~~~~aditional software environment for simulation

- ~~~~e **CPU takes part in the calculation** with the GPUs

- In **AI** the **CPU is the "conductor of the orchestra"** of GPUs

- Exascale has a more expensive, complex external network between servers

These differences are going to get bigger rather than smaller over the next 3-4 years

epcc

Landing a £800m+ project takes time

CRESTA Project

2011 - 2014

EPCC's first EU-funded Exascale project

EuroHPC established

March 2017

Subsequently UK decides not to participate

Exascale Project Working Group

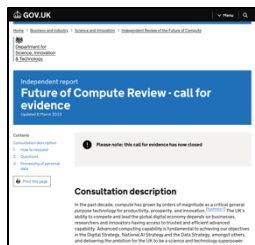March 2019

EPWG starts work on business case to Government

Exascale Project Working Group

August 2019

First business case draft completed

2011

2017

2019

Future of Compute Review

June 2022

Review announced by Government

ARCHER2 opening

February 2022

Prime Minister Johnson opens ARCHER2

Large-scale Computing Review

September 2021

Government Office of Science publishes report

2022

2021

Future of Compute Review

March 2023

Review published recommends Exascale

Budget 2023

March 2023

Chancellor announces £900m investment in Exascale and AI

UKRI formal approvals begin

March 2023

Business Case revision and procurement preparations

Exascale hosting site

October 2023

Edinburgh confirmed as hosting site

Today

August 2024

UK Government "shelves" UK Exascale

2023

2024

epcc

CRESTA Project

2011 - 2014

EPCC's first EU-funded Exascale project

EuroHPC established

March 2017

Subsequently UK decides not to participate

Exascale Project Working Group

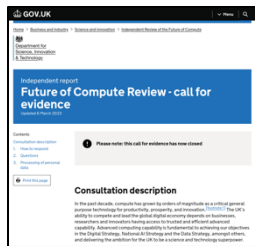March 2019

EPWG starts work on business case to Government

Exascale Project Working Group

August 2019

First business case draft completed

2011

2017

2019

Landing a £800m+ project takes time

Future of Compute Review

June 2022

Review announced by Government

AI is harming traditional HPC because vendors are telling politicians – "AI can do it all"

Large-scale Computing Review

September 2021

Government Office of Science publishes report

2021

Future of Compute Review

March 2023

Review published recommends Exascale

Budget 2023

March 2023

Chancellor announces £900m investment in Exascale and AI

UKRI formal approvals begin

March 2023

Business Case revision and procurement preparations

Exascale hosting site

October 2023

Edinburgh confirmed as hosting site

Today

August 2024

UK Government "shelves" UK Exascale

2023

2024

Fugaku – started with Nine Priority Issues

Achievement of a society that provides health and longevity

01 Innovative drug discovery infrastructure through functional control of biomolecular systems

02 Integrated computational life science to support personalized and preventive medicine

Disaster prevention and global climate problems

03 Development of integrated simulation systems for hazards and disasters induced by earthquakes and tsunamis

04 Advancement of meteorological and global environmental predictions utilizing observational "Big Data"

Energy problems

05 Development of new fundamental technologies for high-efficiency energy creation, conversion/storage and use

06 Accelerated development of innovative clean energy systems

Enhancement of industrial competitiveness

07 Creation of new functional devices and high-performance materials to support next-generation industries (CDMSI)
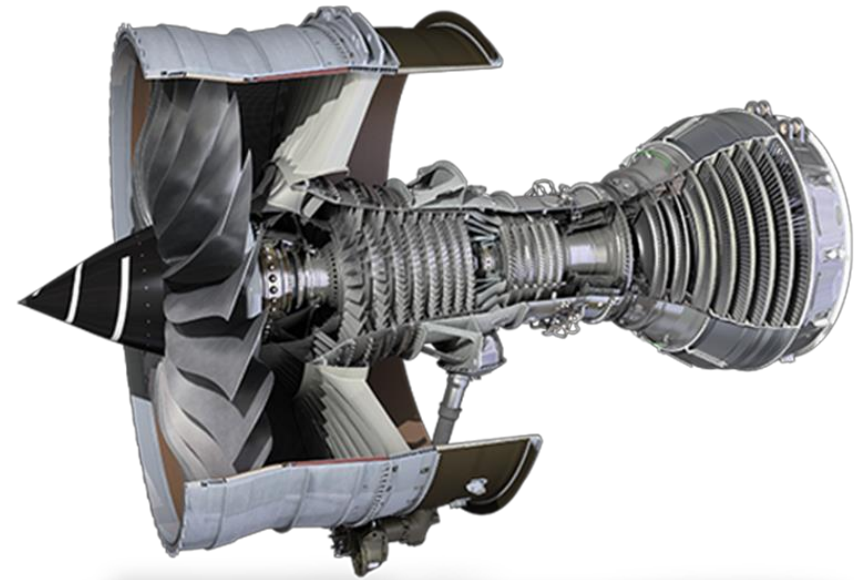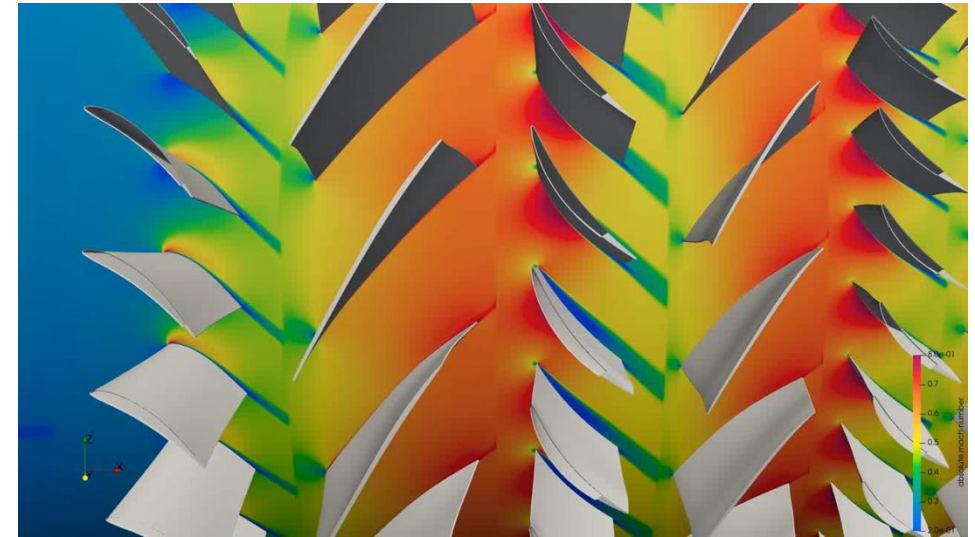
08 Development of innovative design and production processes that lead the way for the manufacturing industry in the near future

Development of basic science

09 Elucidation of the fundamental laws and evolution of the universe

epcc

Rolls Royce ASIMOV Project

- 5-year programme (10 years in total)

- World's first high-fidelity simulation of a gas turbine engine in operation

- Structure / Thermodynamics / Fluid dynamics / Electromagnetics

- A **TRILLION** degrees of freedom

- Towards virtual certification of new engine designs

- An engineering challenge designed for the Exascale era …

- EPSRC Prosperity Partnership with
  - **Rolls Royce**, **Edinburgh**, Warwick, Oxford, Cambridge, Bristol, Zenotech and CFMS

epcc

Thank you