



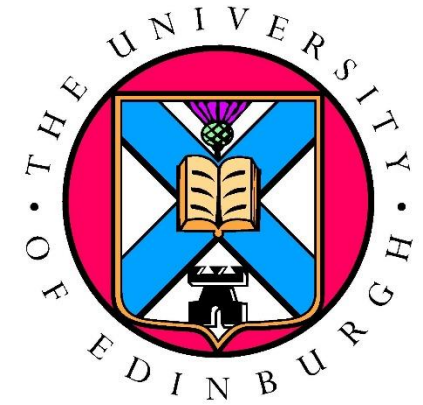
UK Research  
and Innovation

# The UK Exascale Supercomputer Project

Professor Mark Parsons  
EPSRC Director of Research Computing

10<sup>th</sup> December 2021





# THE UK EXASCALE SUPERCOMPUTER PROJECT

---

Professor Mark Parsons

EPCC Director

Dean of Research Computing







# History of the project

- In 2017 establishment of EuroHPC was announced at 60<sup>th</sup> Anniversary of Treaty of Rome celebrations in Rome
- Towards end of 2018, UK declined to join EuroHPC and relinquished its “observer” status on EuroHPC Governing Board
- Exascale Project Working Group set up in late 2018 to develop Outline Business Case for Government
  - Draft OBC first completed in late 2019
  - In parallel Supercomputing Science Case completed and published
- **Since 2020 has moved into UKRI as a cross-Research Council development project within DRI Programme**

# Exascale Requirements from Government

- System should support both **traditional Modelling & Simulation** and **Artificial Intelligence / Deep Learning** applications
  - Technology choices may be impacted by this
  - But future technologies blur the distinction
- System should support both **scientific user communities** and **industry users**
  - A greater focus is proposed with regard to industry use for research
  - Pay-per-use production access will be supported
  - Specific support for SMEs
- System should be **operational around time of EU systems - 2024**

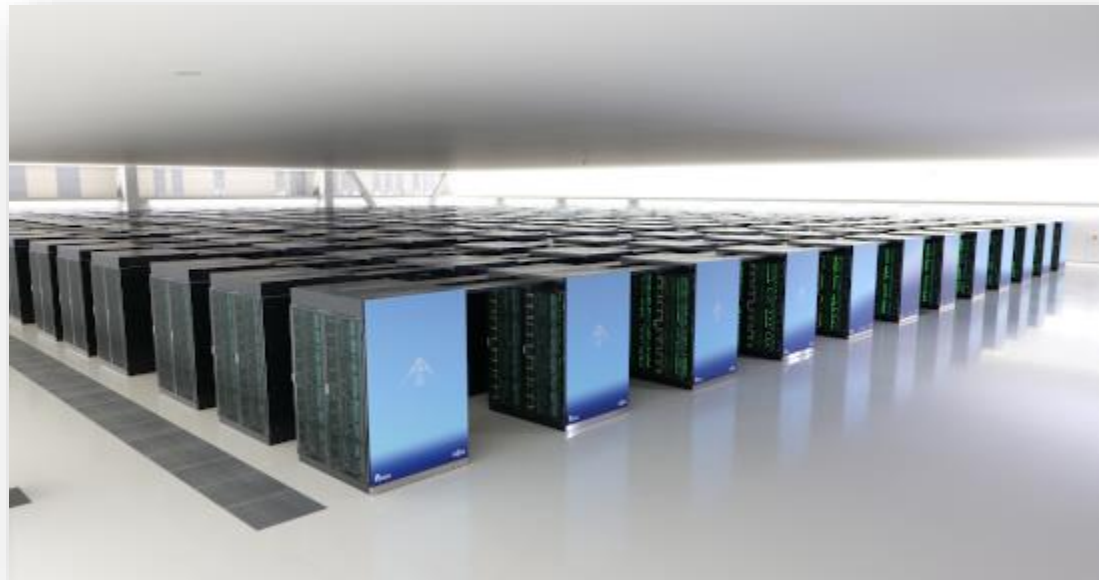
# The Exascale era – worldwide progress

| Country or Region  | Timescale        | Detail   |
|--|------------------|--|
| Japan     | 2020             | Fugaku : based on Fujitsu A64FX Arm processors   |
| China     | 2021             | Two systems in operation - next generation Sunway and Tianhe 3 system. Third system delayed.               |
| USA       | 2021<br>2022     | Frontier : based on AMD EPYC CPU + AMD GPU<br>Aurora : Intel Sapphire Rapids CPU + Intel Ponte Vecchio GPU |
| Europe  | 2021/2<br>2023/4 | Pre-Exascale systems in Finland / Italy + possibly Spain<br>Two Exascale systems in 2024                   |

41 million cores!

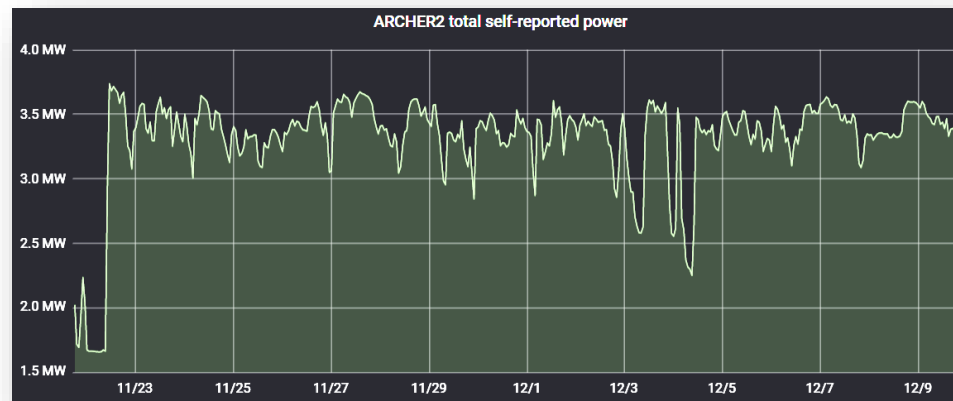
# ... Fugaku wears the crown

- Fugaku became the world's fastest supercomputer in June 2020 with a cores-only approach based on the Fujitsu A64FX Arm CPU
- Processor developed in long-term co-design (10 years) with Japanese computational science community led by Riken CCS
- 7,630,848 Arm CPU cores
- $R_{\text{peak}} = 573.2$  Petaflop/s
- $R_{\text{max}} = 442.0$  Petaflop/s
- Power = 29.9 MW
- Single precision  $> 1$  Exaflop

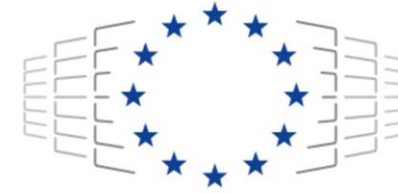


# ... and ARCHER2 is finally here

- The 23 cabinet system finally opened for all users on 22<sup>nd</sup> November
- Very difficult 18 months
- Performance of the system is now good – we hope users agree
- Busy from Day 1 – and has remained busy



# Exascale in the EU



**EuroHPC**  
Joint Undertaking

- EuroHPC Joint Undertaking established to co-fund Pre-Exascale and Exascale systems with Member States
  - Long-term plan – including development of EU processor by SiPEARL
  - Funding of €7billion from 2021-2027
- Three sites chosen for pre-Exascale systems in 2019 – Finland (CSC), Italy (CINECA) and Spain (BSC)
- Two pre-Exascale systems procured for Finland and Italy
  - Spanish procurement is being re-run
- Exascale systems planned for 2024/25
  - Hosting locations likely to be Germany and France



# Recent EuroHPC announcements

- Finland (CSC) is hosting Lumi
  - 375 Petaflops (HPL) / 550 Petaflops (Peak)
  - €145 million
  - Supplied by HPE
  - AMD EPYC CPUs + AMD GPUs
- Italy (CINECA) will host Leonardo
  - 249 Petaflops (HPL) / 324 Petaflops (Peak)
  - €120 million
  - Supplied by ATOS
  - Intel Icelake CPUs + NVIDIA A100 GPUs



First racks  
installed



Delayed

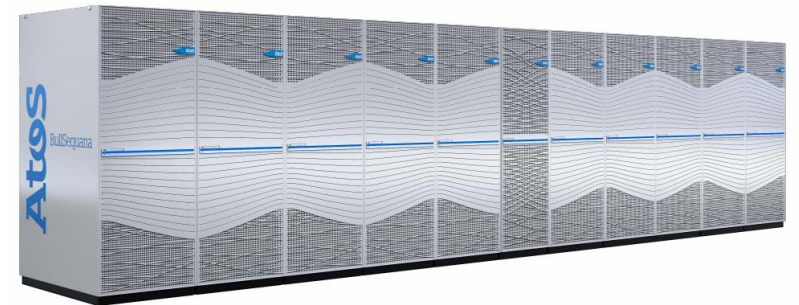
# More detail on Lumi

- HPE Cray EX system
  - Same platform as ARCHER2
- GPU partition
  - 2,560 nodes – 1 AMD Trento CPU + 4x AMD MI250X GPUs
  - 10,240 GPUs and 16,384 cores
- CPU partition
  - 1,536 nodes – 2x AMD Trento CPUs
  - 196,608 cores
- 375 PFlops (HPL) / 550 PFlops (Peak)



# More detail on Leonardo

- ATOS BullSequana system
  - Two partitions – “Booster” and “Data Centric”
- GPU partition (Booster) – 3,456 nodes
  - 221,184 cores - Intel Icelake CPUs
  - 13,824 NVIDIA A100 GPUs
- CPU partition (Data Centric) – 1,536 nodes
  - 79,872 cores - Intel Sapphire Rapids CPUs
  - Local NVM (DCPMM?) for data analysis
- 249 PFlops (HPL) / 324 PFlops (Peak)



# Scientific impact

**Anton 3**

David E. Shaw,<sup>1</sup> J. Adam Butts, Tim Michael Fenn, Christop Justin Gallingsrud, Lev Iserovich, Bryan L. Richard McGowen, Jon L. Petricolas, G Jochen Spengler,<sup>2</sup> Tamas Stan

**Abstract**—Anton 3 supercomputers special molecules relevant to (PIC) whole-volum new Sunway magnetized billion grids, 201.1 PFLOI achieving 29 unprecedenti fully kinetic Experimenta designed on Reactor (CF) be investigat problems an plasma direc PIC method Index Ter methods, het

I. JUSTIFICATIO  
Anton 3 sets new speed of all-atom biom fastest general-purpos 120 times faster on a 2 512-node machine is ribosome system.

II. PER

|                           |
|---------------------------|
| Category of achievement   |
| Type of method used       |
| Results reported on basis |
| Precision reported        |
| System scale              |
| Measurement mechanism     |

<sup>1</sup> David E. Shaw is also affili E-mail correspondence: Da  
<sup>2</sup> See Acknowledgements s

XXX-X-XXXX-XXXX-X

**Synthetic Whole-Volume Simulation**

Jianguo Zhang,<sup>1</sup> Ziyu Zhang,<sup>1</sup> Yong (Ale) Jia,<sup>1</sup>

**Abstract**—We develop a high-precision quantum circuit Our major innovation and a path-optimiz ity and compute den that scales to about 4 multiplication design a wide range of tem precision scheme to ulator effectively ex- clude the 10x10(qubi performance of 1.2 E precision) as a new mi cuts; and reduces the to 304 seconds, from

ACM Reference Form  
Yong (Alexander) Liu<sup>1,2</sup>, Fu<sup>2</sup>, Yuling Yang<sup>1,2</sup>, Jia Peng<sup>2</sup>, Huotong Chen<sup>1,2</sup>, Dexun Chen<sup>2,3</sup>. 2021. C Real-Time Simulation of Supercomputer. In SC21 Computing, Networking Louis, MO. ACM, New Y

I. JUSTIFICATIO  
We empl preserving f coordinates and a sustain nodes of tl whole-volum 2.57 × 10<sup>10</sup>

Category of Achievement  
Type of method used  
Precision reported  
System scale  
Measurement mechanism

<sup>1</sup> Correspond  
han@ustc.edu

**Closing the Simulation**

Yong (Ale) Jia,<sup>1</sup>

**Abstract**—We develop a high-precision quantum circuit Our major innovation and a path-optimiz ity and compute den that scales to about 4 multiplication design a wide range of tem precision scheme to ulator effectively ex- clude the 10x10(qubi performance of 1.2 E precision) as a new mi cuts; and reduces the to 304 seconds, from

ACM Reference Form  
Yong (Alexander) Liu<sup>1,2</sup>, Fu<sup>2</sup>, Yuling Yang<sup>1,2</sup>, Jia Peng<sup>2</sup>, Huotong Chen<sup>1,2</sup>, Dexun Chen<sup>2,3</sup>. 2021. C Real-Time Simulation of Supercomputer. In SC21 Computing, Networking Louis, MO. ACM, New Y

I. JUSTIFICATIO  
We empl preserving f coordinates and a sustain nodes of tl whole-volum 2.57 × 10<sup>10</sup>

Category of Achievement  
Type of method used  
Precision reported  
System scale  
Measurement mechanism

<sup>1</sup> Correspond  
han@ustc.edu

**A 400 trillion Supercomputer Relic Network**

Jianguo Zhang,<sup>1</sup> Ziyu Zhang,<sup>1</sup> Yong (Ale) Jia,<sup>1</sup>

**Abstract**—We develop a high-precision quantum circuit Our major innovation and a path-optimiz ity and compute den that scales to about 4 multiplication design a wide range of tem precision scheme to ulator effectively ex- clude the 10x10(qubi performance of 1.2 E precision) as a new mi cuts; and reduces the to 304 seconds, from

ACM Reference Form  
Yong (Alexander) Liu<sup>1,2</sup>, Fu<sup>2</sup>, Yuling Yang<sup>1,2</sup>, Jia Peng<sup>2</sup>, Huotong Chen<sup>1,2</sup>, Dexun Chen<sup>2,3</sup>. 2021. C Real-Time Simulation of Supercomputer. In SC21 Computing, Networking Louis, MO. ACM, New Y

I. JUSTIFICATIO  
We empl preserving f coordinates and a sustain nodes of tl whole-volum 2.57 × 10<sup>10</sup>

Category of Achievement  
Type of method used  
Precision reported  
System scale  
Measurement mechanism

<sup>1</sup> Correspond  
han@ustc.edu

**Extreme-Scale Simulation**

Jianguo Zhang,<sup>1</sup> Ziyu Zhang,<sup>1</sup> Yong (Ale) Jia,<sup>1</sup>

**Abstract**—We develop a high-precision quantum circuit Our major innovation and a path-optimiz ity and compute den that scales to about 4 multiplication design a wide range of tem precision scheme to ulator effectively ex- clude the 10x10(qubi performance of 1.2 E precision) as a new mi cuts; and reduces the to 304 seconds, from

ACM Reference Form  
Yong (Alexander) Liu<sup>1,2</sup>, Fu<sup>2</sup>, Yuling Yang<sup>1,2</sup>, Jia Peng<sup>2</sup>, Huotong Chen<sup>1,2</sup>, Dexun Chen<sup>2,3</sup>. 2021. C Real-Time Simulation of Supercomputer. In SC21 Computing, Networking Louis, MO. ACM, New Y

I. JUSTIFICATIO  
We empl preserving f coordinates and a sustain nodes of tl whole-volum 2.57 × 10<sup>10</sup>

Category of Achievement  
Type of method used  
Precision reported  
System scale  
Measurement mechanism

<sup>1</sup> Correspond  
han@ustc.edu

**Digital transformation of droplet/aerosol infection risk assessment realized on "Fugaku" for the fight against COVID-19**

Makoto Tsubokura,<sup>1</sup> Keiji Onishi<sup>1\*</sup>, and

Supercomputing '21, November 14–19, 2021  
XXXX-1-14  
©The Author(s) 2021  
Reprints and permission:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/16865805211011111  
www.sagepub.com/  
SAGE

Time: 2.90s

Radius  
1.2e-04  
5e-5  
2e-5  
1e-5  
5e-6  
2e-6  
1e-6  
4.5e-07

formation of epidemiology in allowing first time, but also transformed the ssion risks in multitudes of societal sized out of a combination of a new simulations to scale massively with extremely rapid time-to-solution due ns in minutes not week, attaining true st 1.5 years on Fugaku, cumulatively be media as well as becoming official

ed boundary method, Dirty CAD,

t and/or on shaky scientific grounds, coming from seemingly authoritative disrupted the socio-economic activities One might still recall that, in the early c, even institutions such as the WHO CDC gave somewhat skeptical views s, which might have misdirected the s causing pandemic to worsen.

O declared the COVID-19 pandemic Ministry of Education, Culture, Sports, logy (MEXT) and RIKEN Center science (R-CCS), jointly announced a ploit the computational capability of computer, which was still in the early n, to combat COVID-19. As Fugaku the COVID-19 applications would with both resources and support from he design and manufacturing partner; he project will be granted computing e entire dominance of a top-tier class quivalent to tens of millions of node

ational Science (R-CCS), Japan

<sup>1</sup>Tokyo Institute of Technology, Japan  
<sup>\*</sup>All authors are listed in alphabetical order by surnames.

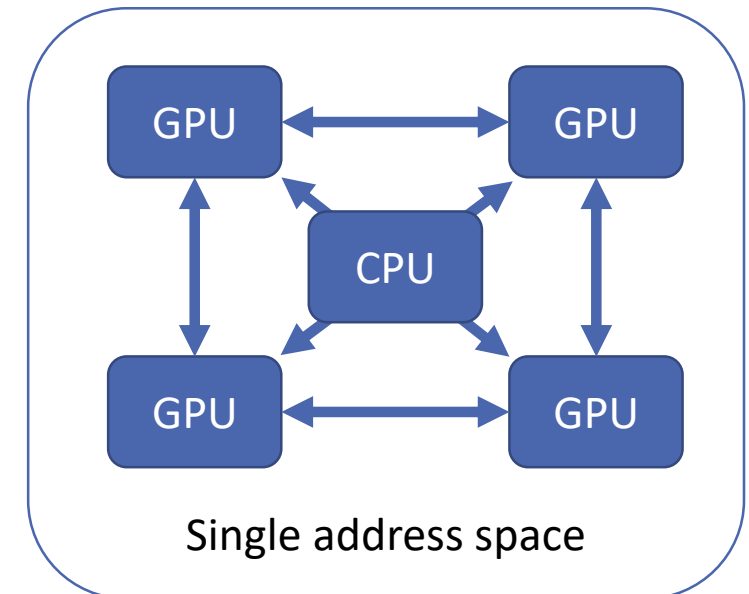
Corresponding author:  
Makoto Tsubokura, Complex Phenomena Unified Simulation Research Team, RIKEN Center for Computational Science (R-CCS), Kobe, Japan.  
Email: mtsubo@riken.jp

Prepared using sagej.cls [Version: 2017/01/17 v1.20]

Provide the capability and scientists will use it

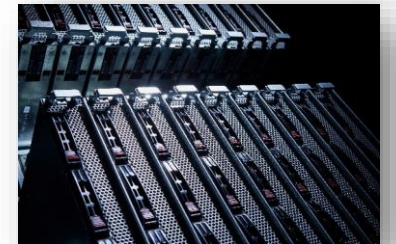
# Technology –recent Exascale vendor briefings

- Memory is changing
  - Many Exascale blades include HBM
    - Some designs have no DRAM at all
  - But recently LPDDR5 is being mentioned more
- Four-way competition for CPUs and/or GPUs
  - Intel versus AMD versus Arm versus NVIDIA
- GPUs market is broadening
  - AMD is strongly competing with NVIDIA
- Cabinet energy densities are rocketing
  - Today's 80-100KW cabinets will be eclipsed by cabinets at 300KW+
- Multicore CPUs are also getting AI Deep Learning features



# General design principles for UK Exascale Project

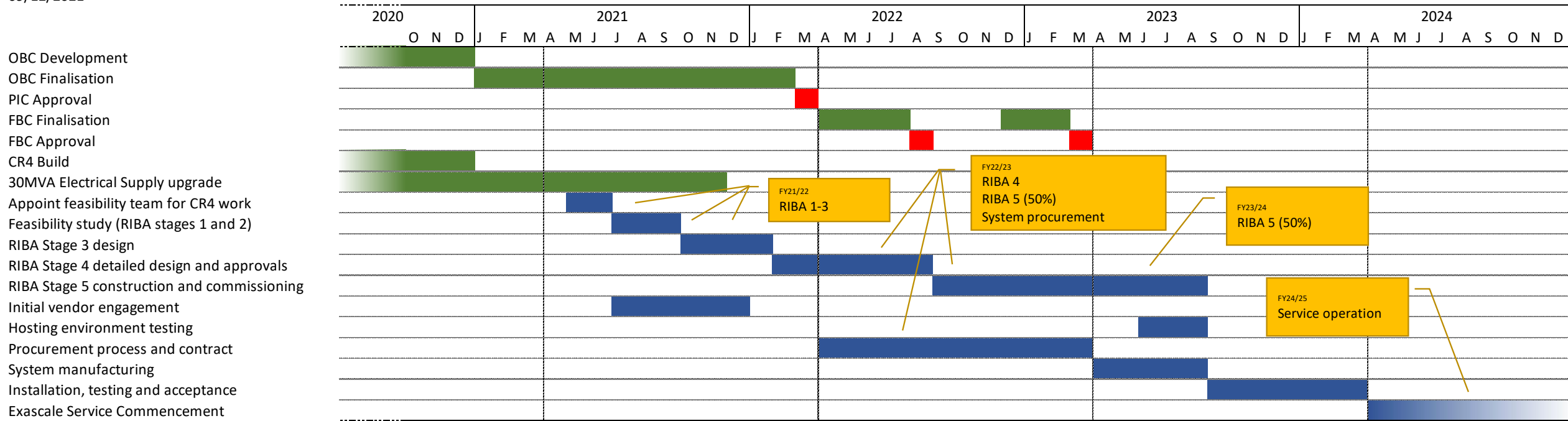
- 25MW system + 5MW support and cooling
- Single tightly coupled system
- Main compute power from GPU partition
  - Target 1 Exaflop/s  $R_{MAX}$
- Remainder of space or power budget for CPU partition
  - Designed to provide attractive powerful resource for non-accelerated codes as they transition
- Large Software Programme envisaged
  - Multiple activities – Grand Challenge based to eCSE type activities
  - Lots of requirements gathering / consultation to do



# Project timeline

## UK Exascale Supercomputer Timeline

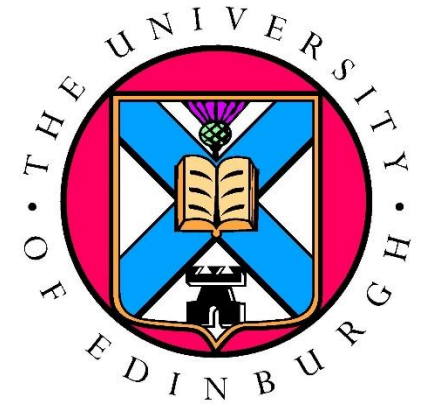
09/12/2021



Activities already funded by University of Edinburgh / Government and underway  
 Activities funded via Exascale Project (not all yet funded)  
 Latest date for key decision points

**Entirely dependent on funding and UKRI prioritisation**





# SYSTEM HOSTING AND OUTLINE DESIGNS

---

Professor Mark Parsons

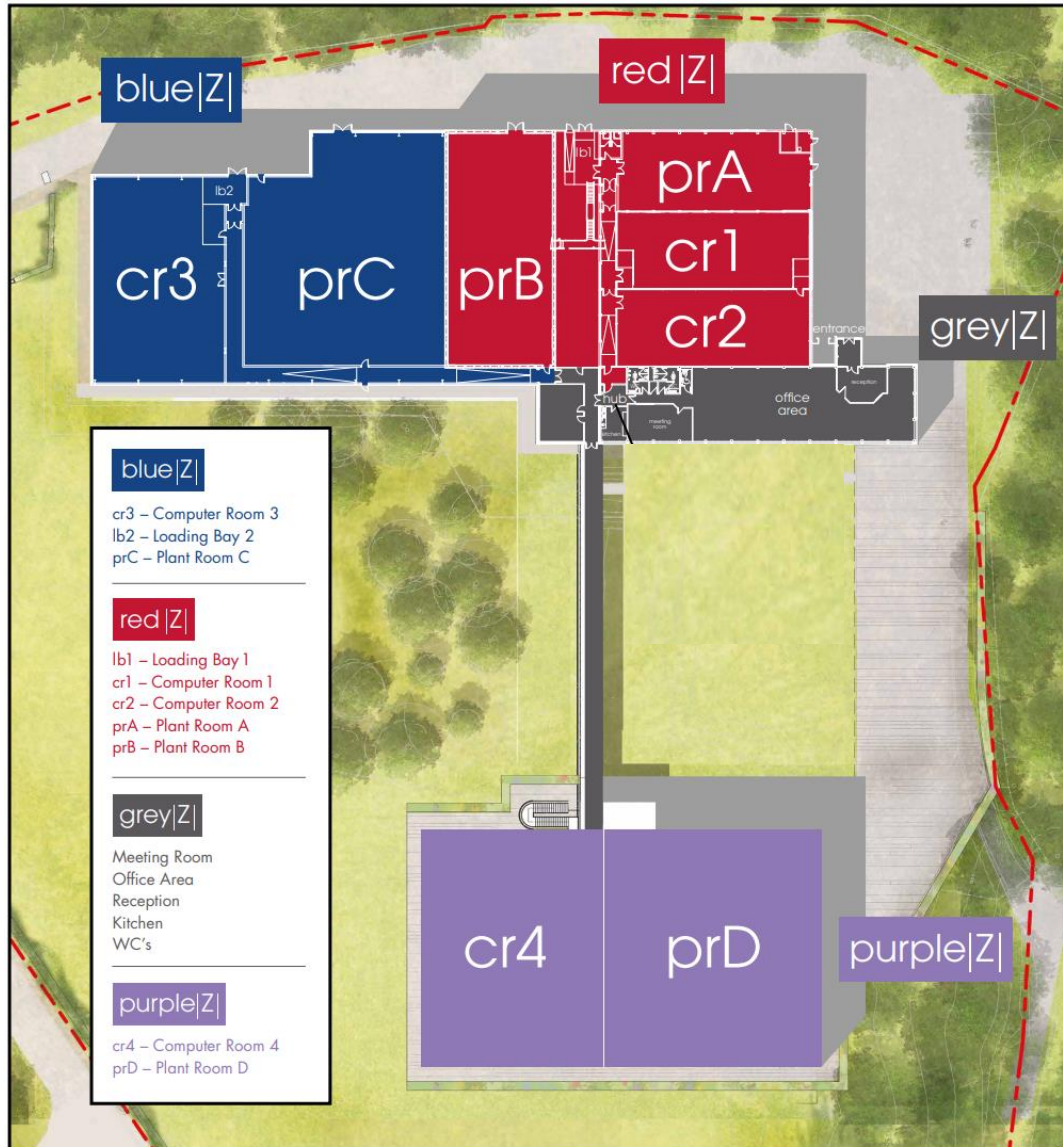
EPCC Director

Dean of Research Computing





# EPCC's Advanced Computing Facility Data Centre



- Plant Room A and Computer Rooms 1&2 date back to 1970s
- Plant Room B added for HECToR
- Computer Room 3 and Plant Room C added for ARCHER – 4MW capability
- Computer Room 4 and Plant Room D added in 2020 – current configuration 6MW



## Computer Room 4

£20m – CR 4 + PR D

£8.6m – 30MVA additional power

Space for 270 standard racks

Opened Dec 2020



# Preparing the ACE facility

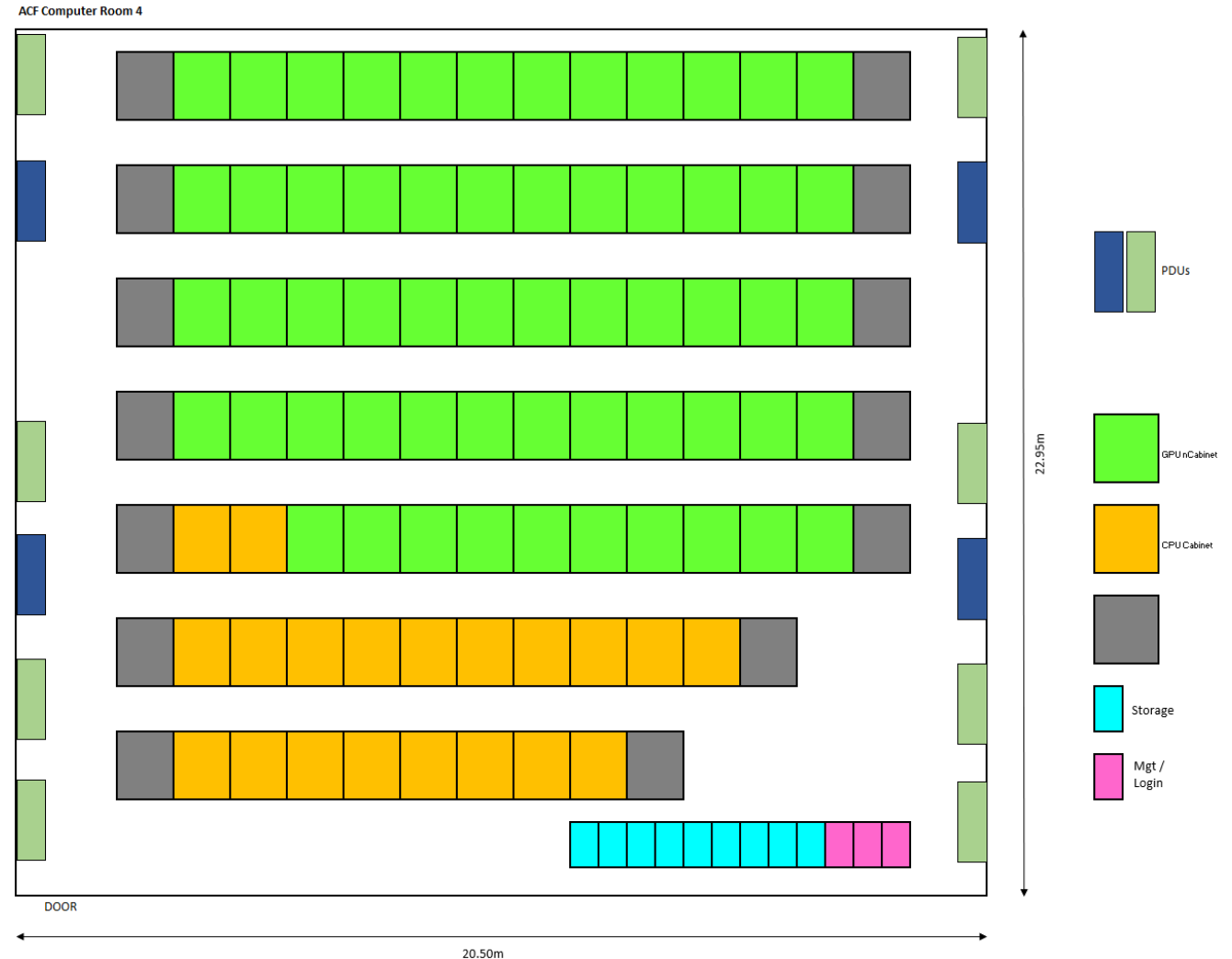
- EPCC's
- Very complex
- **New £20m**
- **New £9m**
- by Easter
- CR4 is a
- electrical
- Process
- Edinburgh
- **Preparations must be complete by Q2 2023**



to  
es  
4  
riv  
m  
E  
ys  
dy  
by 2024  
be live  
ical and  
y of  
FI

# Example from RFI responses (obfuscated)

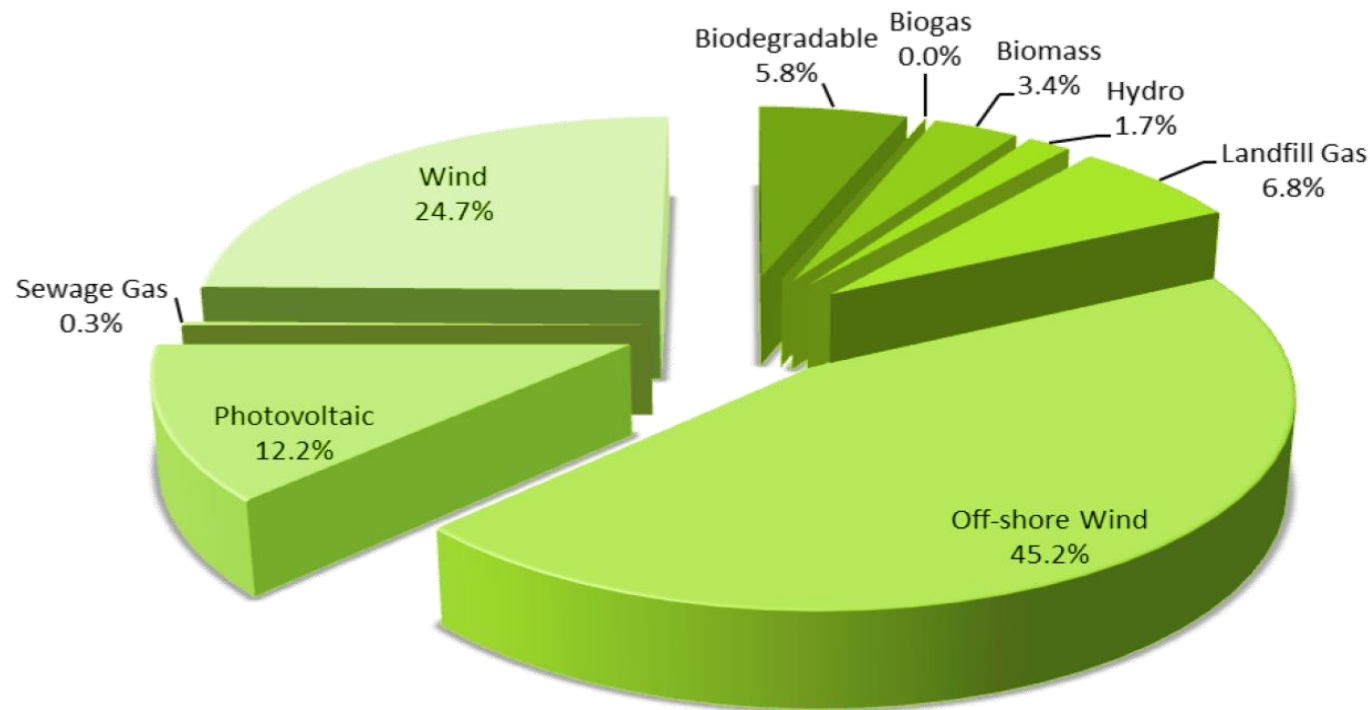
- Vendors asked to produce designs up to 25MVA
- Combination of
  - 1 ExaFlop HPL  $R_{max}$
  - Cores-only partition
- Dual approach provides route from cores-only world to accelerated world
- GPU Partition
  - **24,000 GPUs**
  - **380,000 cores**
  - 60 racks to reach 1 ExaFlop HPL – 19MVA power
- CPU Partition
  - **1,000,000 cores**
  - 20 racks – 6MVA power (limit reached)
- Plus
  - 100PB storage system
  - Login and service nodes



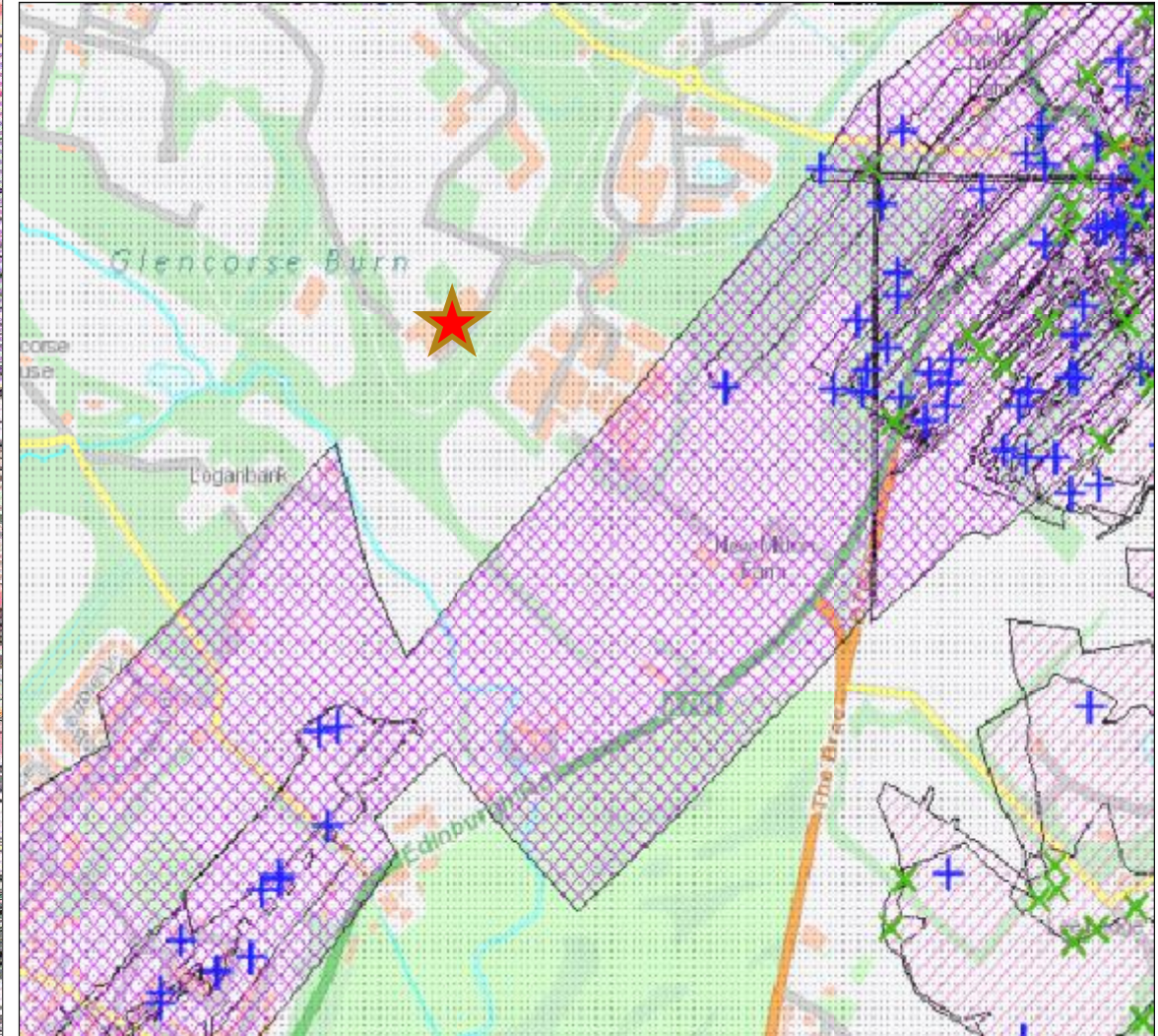
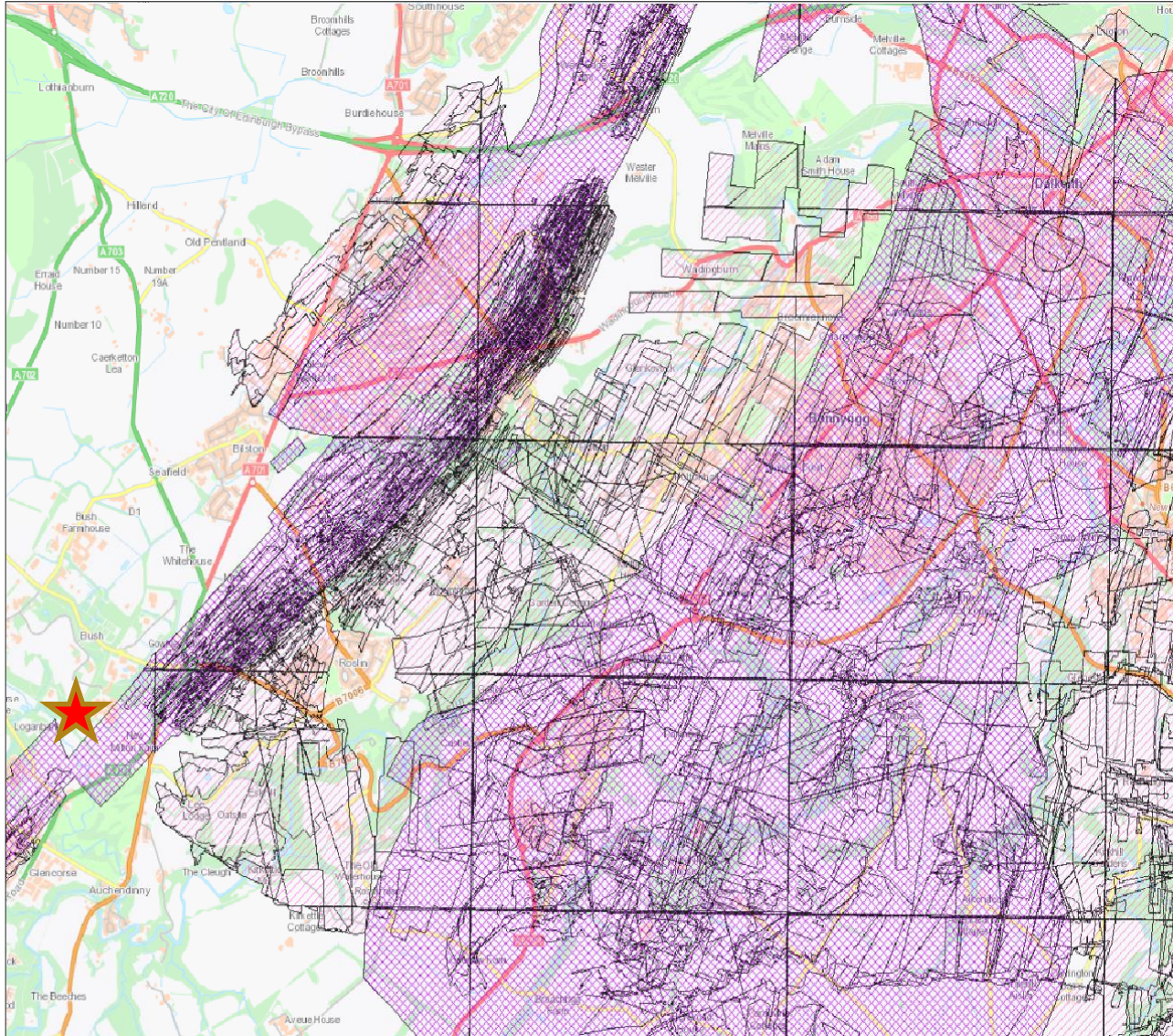
ACF CR4

## Aim for Net Zero - 100% Renewable Energy

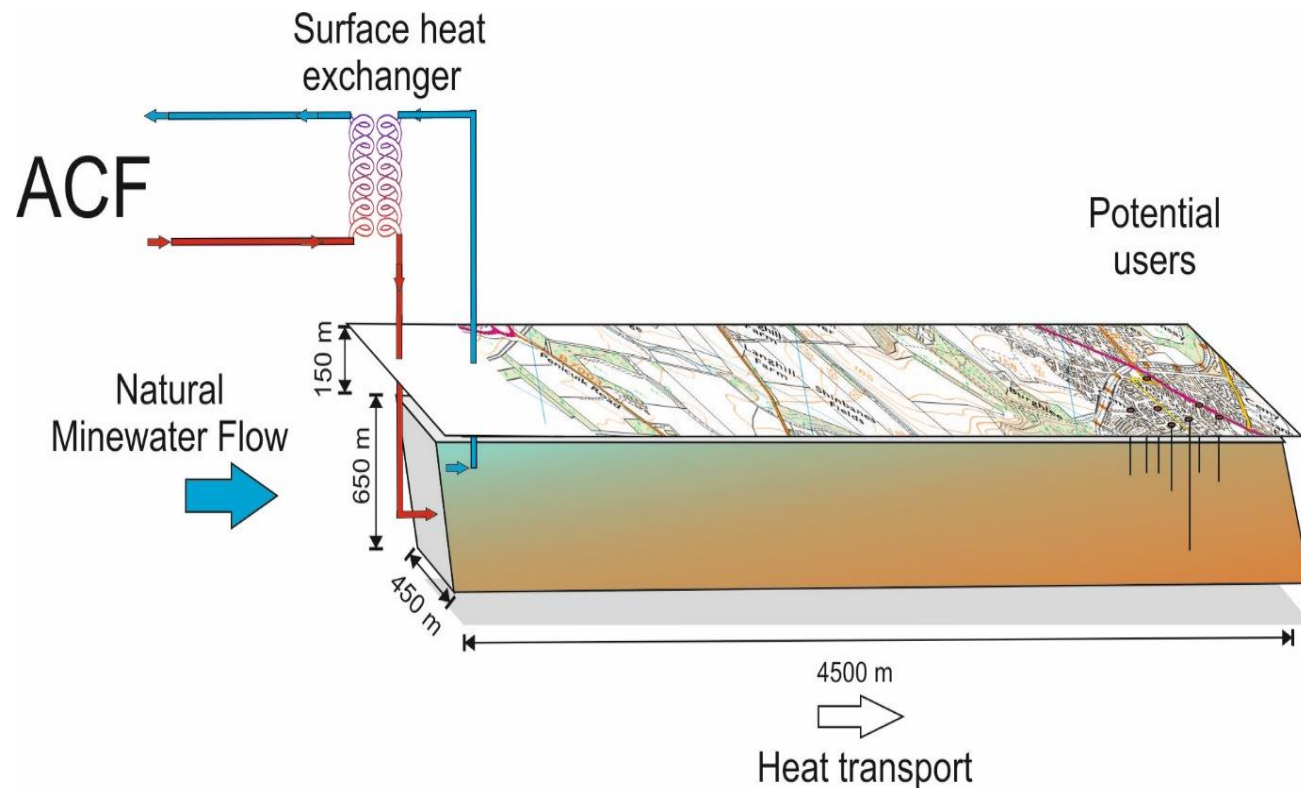
- The University of Edinburgh is part of the Scottish Public Procurement contract for electricity
- We choose the 100% renewable energy option



- The ACF consumed 24.46 GWhrs in FY2018/19 ...
- With ARCHER 2 this will rise to ~50 GWhrs per annum

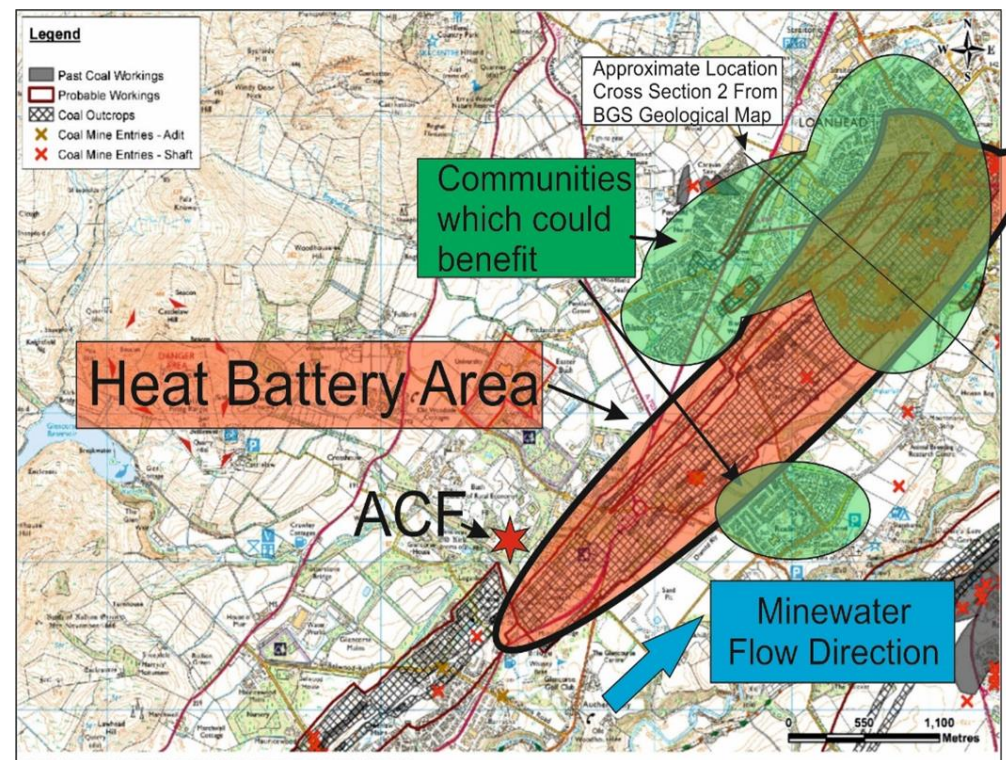


# Aiming for better than Net Zero



Bilston Glen Colliery, 670m, 15.0C, Minewater  
 Monktonhall, 866m, 25.5C, Rock  
 Lady Victoria, 768m, 18C, Minewater

- Detailed feasibility study now completed to use hot water to heat abandoned mine workings
- Will create geothermal heat battery for us by homes, public and commercial buildings
- Battery will extend into South Edinburgh



# Conclusion

- Delivering an Exascale capability will allow the UK's computational science community to compete with their international peers
- A true demonstration of the UK as a Science & Technology superpower
- But ...
  - There is no guarantee funding will be made available
  - Timescales can easily slip
  - As many current projects are showing, these very large systems are not easy to procure, install or operate
- ... however, if we don't try we'll never succeed!