High Performance Computing: EPSRC Service Specification

Please find specific information about each service below. This will help guide your application and determine which service to apply to.

The indicative level of computational resource available through this call is accurate as of the date of publication (see charge log). Many of the Tier2 services are currently in the process of installing new compute hardware and so the

ARCHER2

Service details

Service Contact Details	support@archer2.ac.uk
Service Webpage	https://www.archer2.ac.uk/

Hardware and Technical specifications

System name	ARCHER2
Compute nodes	5,860 compute nodes, each with dual AMD Rome 64 core CPUs at 2.2GHz, for 748,544 cores in total and 1.57 PBytes of total system memory
Processors	
Interconnect	Cray Slingshot
Storage	14.5 PBytes of Lustre work storage in 4 file systems
Software available	https://www.archer2.ac.uk/about/hardware.html
Additional information on hardware available	https://www.archer2.ac.uk/about/hardware.html
Use cases particularly suited to this Service	Large capacity jobs

Resources available through this call

Unit(s) of Allocation	ARCHER2 allocates its compute resource in ARCHER2 Compute Units (CU). Please note 1 node hour on ARCHER2 costs 1 CU, unless jobs are submitted in low priority queues where a discount applies.
Indicative level of computational resource available through this call	Up to 3.4 MCUs, 10% of EPSRC's ARCHER2 compute
Indicative sizes of previously successful applications (not a restriction)	Access to HPC Call: Autumn 2022 Ranged from 12 kCUs – 611 kCUs
% compute allocated to EPSRC mechanisms (including but not limited to this call)	~77- 83%, this is the total % of ARCHER2 EPSRC can utilise each year i.e. EPSRC's ARCHER2 compute.
Storage available	Flexible with justification

Requirements on applications for the service

Project length restrictions over and above those in the call	1 year
Maximum and Minimum requests	Users must request more than 4000 CU. Users who want less can use the Pump-priming access route to ARCHER2, see https://www.archer2.ac.uk/support-access/access.html.

Cirrus

Service details

Service Contact Details	support@cirrus.ac.uk
Service Webpage	https://www.cirrus.ac.uk/

Hardware and Technical specifications

System name	Cirrus
Compute nodes	280 dual CPU compute nodes and 2 quad GPU nodes 144 NVIDIA V100 GPUs and an accompanying fast storage layer.
Processors	Cirrus standard compute nodes each contain two 2.1 GHz, 18-core Intel Xeon E5-2695 (Broadwell) series processors. Each of the cores in these processors support 2 hardware threads (Hyperthreads), which are enabled by default. The standard compute nodes on Cirrus have 256 GB of memory shared between the two processors. The Cirrus GPU compute nodes each contain two 2.4 GHz, 20-core Intel Xeon Gold 6148 (Skylake) series processers. Each of the cores in these processors support 2 hardware threads (Hyperthreads), which are enabled by default. The nodes also each contain four NVIDIA Tesla V100-PCIE-16GB (Volta) GPU accelerators connected to the host processors and each other via PCIe.

	Cirrus Phase II added a GPU node upgrade with a further 36 'Plainfield' blades (single GPU node with two Intel processors and four GPU's)
	Nlong with the necessary power supplies, EDR IB switches and cables. These blades are similar to the two in Cirrus Phase I, except that they have Intel 'CascadeLake' processors (6248), 2933 MHz memory and use EDR IB mezzanine cards and EDR IB switches.
	Each GPU node have four NVIDIA V100's (16GB) for a total of 144 GPU's.
Interconnect	FDR Infiniband Hypercube
Storage	/home: CEPH filesystem for /home
	/work: Lustre file system (406 TB) (will shortly be replaced with 1PB E1000)
	HPE XFS/RPOOL with NVMe devices for fast storage
Software available	http://www.cirrus.ac.uk/about/Cirrus_Service_Component_Catalogue.pdf
Additional information on hardware available	http://www.cirrus.ac.uk/about/hardware.html
Use cases particularly suited to this Service	Good mixture of use cases including CPU and GPU

Resources available through this call

Unit(s) of Allocation	CPU core hours (CPUh), GPU core hours (GPUh)
Indicative level of computational resource available through this call	20.8 MCPUh
	500 kGPUh
Indicative sizes of previously successful applications (not a restriction)	Autumn 2022 Call:
	CPU requests ranged from 2 KCPUh – 6.9 MCPUhs
	GPU requests ranged from 31 KGPUhs – 96 KGPUhs
% compute allocated to EPSRC mechanisms (including but not limited to this call)	70%
Storage available	Defaults are /home 100GiB and /work 250GiB, can allocate more space if justified

Requirements on applications for the service

Project length restrictions over and above those in the call	None
Maximum and Minimum requests	Flexible with justification

CSD3

Service details

Service Contact Details	Stuart Rankin sjr20@cam.ac.uk

Service Webpage	https://www.hpc.cam.ac.uk/high-performance-computing
Service Reference	CSD3

Hardware and Technical specifications

System name	CSD3
Compute nodes	Dell PowerEdge C6420, 56 cores, 192GiB RAM (Cascade Lake CPU)
	and Dell PowerEdge C6520, 76 cores, 256GiB RAM (Ice Lake CPU).
	Dell PowerEdge XE8545, 128 cores (AMD EPYC 7763), 1TiB RAM (GPU).
Processors	Total 10492 CPU cores from a pool containing:
	Intel Xeon Platinum 8276 @ 2.20GHz (Cascade Lake)
	Intel(R) Xeon(R) Platinum 8368Q CPU @ 2.60GHz (Ice Lake).
	Total 62 GPUs (NVIDIA A100 SXM4 80G).
Interconnect	Mellanox HDR100 InfiniBand (Cascade Lake CPU).
	Mellanox HDR200 InfiniBand (Ice Lake CPU).
	Dual-rail Mellanox HDRD200 InfiniBand (GPU).
Storage	2133TB Lustre filesystem.
Software available	GCC and Intel (on CPU nodes) and NVIDIA (on GPU nodes) development environments and various packages, see https://docs.hpc.cam.ac.uk/hpc/ for more information.
Additional information on hardware available	
	https://docs.hpc.cam.ac.uk/hpc/
Use cases particularly suited to this Service	MPI-parallel multinode jobs, shared memory single node jobs, CUDA-accelerated simulations on GPU, machine learning.

Resources available through this call

Unit(s) of Allocation	CPU core hours (CPUh), GPU core hours (GPUh)
Indicative level of computational resource available through this call	22,977,046 CPUh per quarter
	111,155 GPUh per quarter
Indicative sizes of previously successful applications (not a restriction)	1M CPUh per quarter.
	10K GPUh per quarter.
% compute allocated to EPSRC mechanisms (including but not limited to this call)	80
Policies	36 hours maximum wallclock time per job. 4256 CPUs and 64 GPUs maximum active at one time per user. - All proprietary licenses for user software must be provided by the applicant. - All data must be removed at the end of the allocation period.

Requirements on applications for the service

Project length restrictions over and above those in the call	None
Maximum and Minimum requests	No formal limits

Isambard 3

Isambard HPC resources for this opportunity have yet to be confirmed.

Baskerville

Service details

Service Contact Details	baskerville-tier2@contacts.bham.ac.uk
Service Webpage	www.baskerville.ac.uk

Hardware and Technical specifications

System name	Baskerville
Compute nodes	52 Lenovo SD650-N V2 liquid cooled compute nodes
Processors	- All nodes host 2x Intel® Xeon® Platinum 8360Y CPUs, each with 36 cores at 2.4GHz (with boost to 3.5GHz) 46 nodes host 4x NVIDIA A100 40GB GPUs 6 nodes host 4x NVIDIA A100 80GB GPGPUs
Interconnect	All nodes are connected by 1x 25GbE NVIDIA® Mellanox® (on-planar ConnectX-4 port) 1x HDR (200Gbps) NVIDIA Mellanox Infiniband port (ConnectX-6 PCIe gen4 adapter)
Storage	5PB of HDD and 0.5PB of SSD are available. The storage systems are Lenovo DSS-G running IBM® Spectrum Scale
Software available	A full list of installed software is available at https://apps.baskerville.ac.uk/applications/
Additional information on hardware available	N/A
Use cases particularly suited to this Service	Baskerville is designed primarily for GPU-accelerated workloads and is suitable for both single-GPU and multi-GPU jobs, with parallelisation up to 8 nodes (32 GPUs) supported as standard. Workloads that are predominantly CPU-focussed will not be accepted.

Resources available through this call

Unit(s) of Allocation	GPU core hours (GPUh)
Indicative level of computational resource available through this call	29000 GPUh available for 2024 allocation period. 180000 GPUh available for 2025 allocation period.

Indicative sizes of previously successful applications (not a restriction)	Previous successful applications have been awarded between 5000 and 150,000 GPUh in a six-month period.
% compute allocated to EPSRC mechanisms (including but not limited to this call)	40%
Storage available	Up to 500TB of storage can be requested. Globus is available to support parallel data transfer.

Requirements on applications for the service

Project length restrictions over and above those in the call	None
Maximum and Minimum requests	No formal limits

Sulis

Service details

Service Contact Details	sulis@warwick.ac.uk
Service Webpage	https://sulis.ac.uk
Service Reference	https://sulis-hpc.github.io

Hardware and Technical specifications

System name	Sulis
Compute nodes	Dell PowerEdge R6525 compute nodes each with 128 cores per node; 512 GB DDR4-3200 RAM per node
Processors	2 x AMD EPYC 7742 2.25 GHz 64-core processors per node Type 1 GPU nodes contain 3x Nvidia A100 40 GB (PCIe variant) Type 2 GPU nodes contain 3x Nvidia L40 48GB
Interconnect	Mellanox ConnectX-6 HDR100 (100 Gbit/s) InfiniBand
Storage	2 PB IBM Spectrum Scale (GPFS)
Software available	Large suite of compilers/libraries available via environment modules, Python and R packages and end-user applications. Others installed on request. Containers are supported via Singularity.
Additional information on hardware available	Sulis has 4 high memory (1 TB RAM) and 3 very high memory (4 TB of RAM) nodes available for pre- and post-processing of data, memory intensive analytics etc.
Use cases particularly suited to this Service	Sulis is particularly suited for high throughput and ensemble computing workloads consisting of many small (workstation scale) calculations that can be bundled up into a smaller number of job submissions. See sulis-hpc.github.io/advanced/ensemble for examples.

Sulis is less suited for large calculations that require high speed low-latency communication between GPUs or CPU nodes, e.g. high fidelity CFD simulations or data processing that relies on tightly coupled GPUs.

Resources available through this call

Unit(s) of Allocation	CPU core hours (CPUh), GPU core hours (GPUh)
Indicative level of computational resource available through this call	25M CPUh 104k A100 GPUh and 98k L40 GPUh
Indicative sizes of previously successful applications (not a restriction)	Typically 1-5 million CPUh or 15-25k GPUh over 12 months.
% compute allocated to EPSRC mechanisms (including but not limited to this call)	30%
Storage available	Default 2TB per user. Larger allocations can be accommodated on request. An SSD scratch partition is available for workloads performing intensive and non-sequential IO operations.

Requirements on applications for the service

Project length restrictions over and above those in the call	None
Maximum and Minimum requests	No formal limits. Projects needing less than 100,000 CPUh or 1,000 GPUh should request exploratory access via sulis.ac.uk rather than applying via this call.

JADE

JADE HPC resources are not available for this Access 2 HPC opportunity.

MMM Hub

The MMM Hub operates a different allocation mechanism to the other Tier 2 Centres. Access is only available via membership of two HEC consortia and not via this call. For more information see: https://mmmhub.ac.uk/young/.

NI-HPC (Kelvin-2)

Service details

Service Contact Details	v.purnell@qub.ac.uk
Service Webpage	www.ni-hpc.ac.uk

Hardware and Technical specifications

System name	Kelvin-2
Compute nodes	Standard

- 60x Dell PowerEdge R6525 with 768GB RAM
- 28x Dell PowerEdge R6525 with 1TB RAM

X-series compute:
- 2x Dell PowerEdge R6525 with 1TB RAM with dual 7773X CPU series

Hi-memory
- 6x Dell PowerEdge R6525 with 2TB RAM

GPU
- 8 x Dell DSS8440 (each with 2x Intel Xeon Platinum 8168 24 Core CPU). Provides 32x NVIDIA Tesla v100 32GB. One node is run in MIG mode with 27 instances per GPU.
- 4 x XE8545 (each with 4x NVIDIA A100 80GB) AMD Rome 2x64core 7702

Processors	AMD Rome 2x64core 7702
Interconnect	Mellanox EDR infiniband
Storage	2PB usable lustre for scratch storage Metadata Servers: Dell R640 Metadata Targets: Dell Powervault ME2024 with 1TB SSD Object Storage servers: Dell Powervault ME4084
Software available	Centos 7.7 (planning to move to Centos 9 stream in 2024) Lustre scratch file system Alces flight cluster manager
Additional information on hardware available	
Use cases particularly suited to this Service	Priority areas from original bid focussed on: computational neuroscience, advanced chemistry, innovative drug delivery, precision medicine, food fingerprinting and hydrogen deflagration Service is general purpose machine so can deliver on a wide range of use cases.

Resources available through this call

Unit(s) of Allocation	CPU core hours (CPUh), GPU core hours (GPUh)
Indicative level of computational resource available through this call	8.5M CPUh per quarter 33,000 GPUh per quarter
Indicative sizes of previously successful applications (not a restriction)	2M CPUh per quarter 10,000 GPUh per quarter
% compute allocated to EPSRC mechanisms (including but not limited to this call)	35
Storage available	2PB shared scratch

Requirements on applications for the service

Project length restrictions over and above those in the call	None
Maximum and Minimum requests	Max number of cores per job: 1344
	Min number of cores per job: 600
	GPUs per job max: 4
	GPUs per job min: 1

Northern Intensive Computing Environment (Bede)

Service details

Service Contact Details	Rebecca Appleby (Rebecca.appleby@durham.ac.uk, +44(0) 191 33 4250
Service Webpage	https://n8cir.org.uk/bede/

Hardware and Technical specifications

System name	bede.dur.ac.uk
Compute nodes	32x IBM AC922 with 0.5TB and 4x32GB V100 GPU,
	4x IBM IC922 with 256GB and 4xT4 GPU,
	3x Grace Hopper 480GB LPDDR5X RAM
Processors	AC922: 2x16core 2.7Ghz Power 9.
	IC922: 2x20core 2.9Ghz Power 9.
	Grace Hopper: Grace 72-core aarch64 CPU 3.483Ghz, Hopper H100 96GB
Interconnect	Mellanox EDR
Storage	2Pb, 10GB/s Lustre filesystem for running jobs.
Software available	https://bede-documentation.readthedocs.io/en/latest/software/index.html
Additional information on hardware available	https://bede-documentation.readthedocs.io/en/latest/hardware/index.html
Use cases particularly suited to this Service	Accelerated computing that requires more than what can be provided by a single accelerator: Extending accelerator memory into CPU (memory coherence); using multiple accelerators per node & using multiple accelerators across nodes. Both simulation and AI/ML workloads.

Resources available through this call

Unit(s) of Allocation	GPU core hours (GPUh)
Indicative level of computational resource available through this call	160,000 GPUh per year.
	Access to Grace-Hopper seed units for evaluation

Indicative sizes of previously successful applications	14,000 GPUh over 12 months
(not a restriction)	
% compute allocated to EPSRC mechanisms (including but not limited to this call)	38%
Storage available	N/A

Requirements on applications for the service

Project length restrictions over and above those in the call	None
Maximum and Minimum requests	40,000 GPUh / year maximum

Change log

Name	Date	Version	Change
Richard Bailey	01/03/2023	1	Initial version
Richard Bailey	27//09/2023	2	Updated for Autumn 2023 call
Richard Bailey	23/02/2024	3	Updated for 2024 call