ARCHER2 User Forum at the Celebration of Science

epcc

archer2

THE UNIVERSITY of EDINBURGH

ARCHER2 User Forum

at the

Friday 8th March 2024 09:00 - 10:00

EPCC Staff will present information & answer questions about the ARCHER2 service.

ARCHER2 GPU Development Platform

ARCHER2 CSE Team, EPCC, The University of Edinburgh

support@epcc.ed.ac.uk

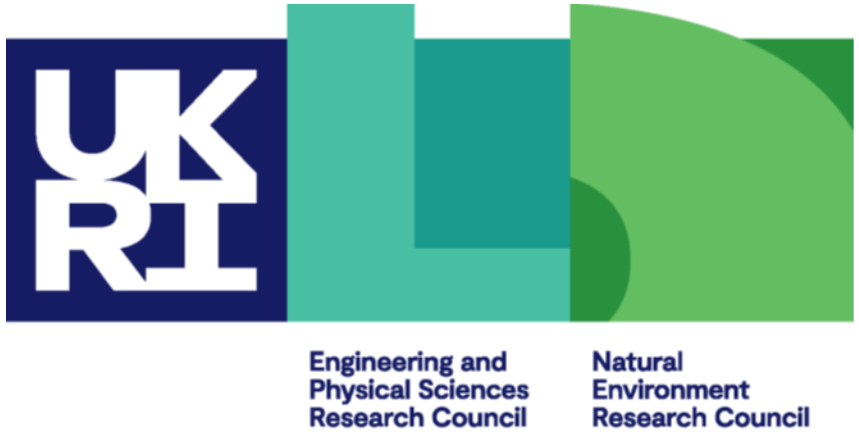
www.archer2.ac.uk

Outline

- GPU development platform overview
- GPU node hardware
- Scheduler configuration
- Thoughts for the discussion/coffee break
- Upcoming GPU Training
- GPU eCSE Call

ARCHER2 Partners

epcc



epcc

THE UNIVERSITY of EDINBURGH

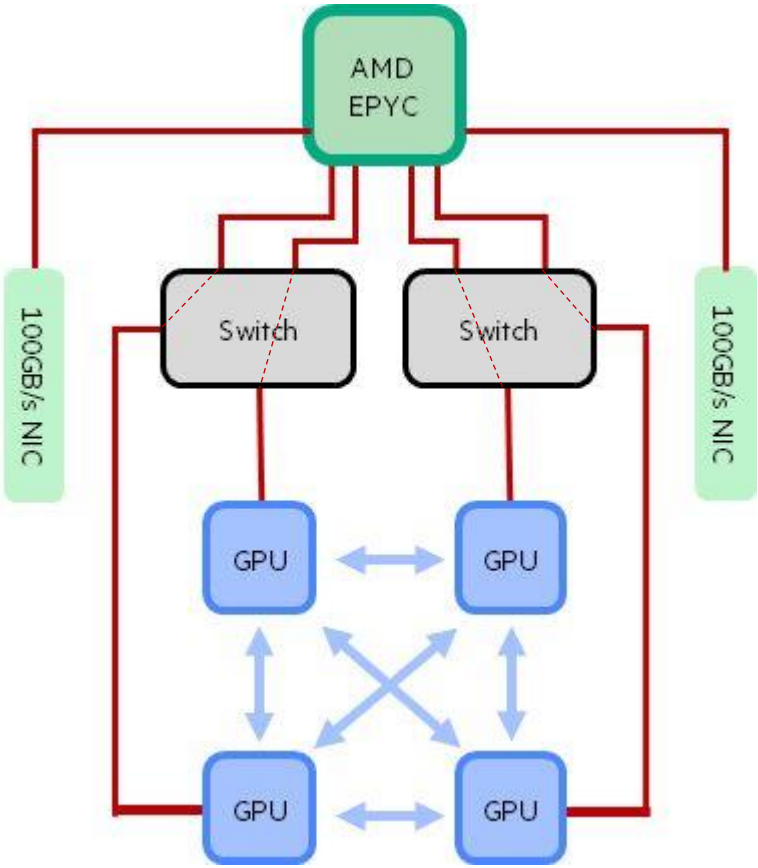
Hewlett Packard Enterprise

GPU Development Platform overview

- A small resource
  - 4 nodes, 16 GPU in total
- For development and testing, not production use.
- Integrated into ARCHER2 system
  - No separate login nodes, same scheduler
- Available to all ARCHER2 users with a positive CU budget
  - **No charge for GPU node use!**
- Documentation at: https://docs.archer2.ac.uk/user-guide/gpu/
  - Basic documentation available
  - Will be added to and updated over time

GPU node hardware

- Each node:
  - 1x AMD CPU
  - 4x AMD Instinct MI210 GPU
  - 2x single port 100 Gbps Slingshot interconnect
- /work and solid state (NVMe) file systems available
- Comms between GPUs on node “fast” (blue arrows)
- Comms between GPUs on different nodes “less fast” (red lines)
- Test yourself with `rocm-bandwidth-test`



Compiling

- Programming environment with GPU support available on:
  - Login nodes
  - GPU compute nodes (terminal access via Slurm)
  - Serial (data analysis nodes)
- Full details of offload methods compiler support in ARCHER2 documentation

Programming Environment	Description	Actual compilers called by `ftn`, `cc`, `CC`
`PrgEnv-amd`	AMD LLVM compilers	`amdflang`, `amdclang`, `amdclang++`
`PrgEnv-cray`	Cray compilers	`crayftn`, `craycc`, `crayCC`
`PrgEnv-gnu`	GNU compilers	`gfortran`, `gcc`, `g++`
`PrgEnv-gnu-amd`	hybrid	`gfortran`, `amdclang`, `amdclang++`
`PrgEnv-cray-amd`	hybrid	`crayftn`, `amdclang`, `amdclang++`

PrgEnv	Actual compiler	OpenMP Offload	HIP	OpenACC
`PrgEnv-amd`	`amdflang`	✅	❌	❌
`PrgEnv-amd`	`amdclang`	✅	❌	❌
`PrgEnv-amd`	`amdclang++`	✅	✅	❌
`PrgEnv-cray`	`crayftn`	✅	❌	✅
`PrgEnv-cray`	`craycc`	✅	❌	❌
`PrgEnv-cray`	`crayCC`	✅	✅	❌
`PrgEnv-gnu`	`gfortran`	❌	❌	❌
`PrgEnv-gnu`	`gcc`	❌	❌	❌
`PrgEnv-gnu`	`g++`	❌	❌	❌

Scheduler configuration

- Submit from ARCHER2 login nodes as for other job types
- Use your job script to select the number of GPU you require
  - `--gpus=N` option
  - Do not specify number of CPU cores or amount of memory
  - CPU cores and memory assigned pro rata based on number of GPU you request
  - 8 CPU cores and 128 GiB memory allocated per GPU requested
- Once you have your allocated resources, use `srun` to specify how many tasks (MPI processes) and threads you need and in what configuration

QoS	Max Nodes Per Job	Max Walltime	Jobs Queued	Jobs Running	Partition(s)	Notes
gpu-shd	1	1 hr	2	1	gpu	Nodes potentially shared with other users
gpu-exc	2	1 hr	2	1	gpu	Exclusive node access

Considerations for GPU enabled HPC

- Balance between performance, portability and productivity.
- This isn't all about the developers we need users to know how to effectively use new HPC services.
- What support do your communities need to be ready for GPU based HPC services?

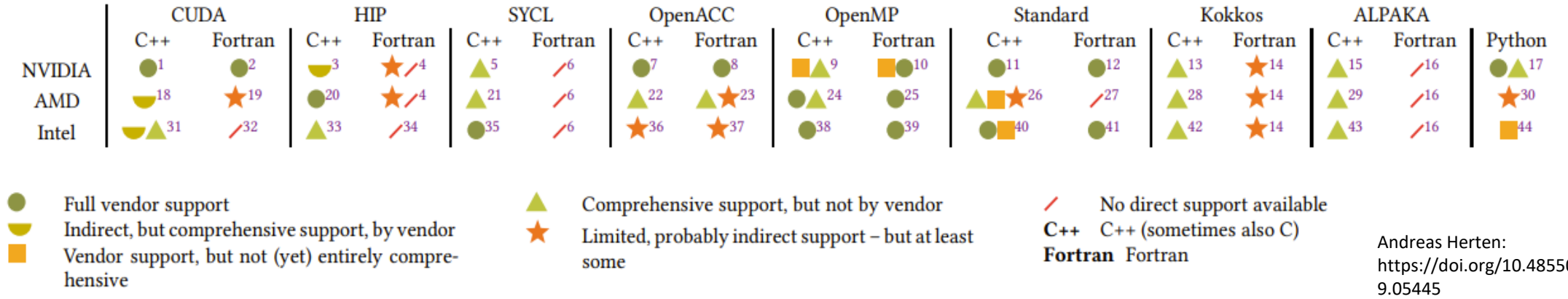
from HPE ARCHER2 GPU workshop slides

Considerations for GPU enabled HPC

- Balance between performance, portability and productivity.
- This isn't all about the developers we need users to know how to effectively use new HPC services.
- What support do your communities need to be ready for GPU based HPC services?

from HPE ARCHER2 GPU workshop slides

GPU Hardware-software landscape



Andreas Herten: https://doi.org/10.48550/arXiv.2309.05445

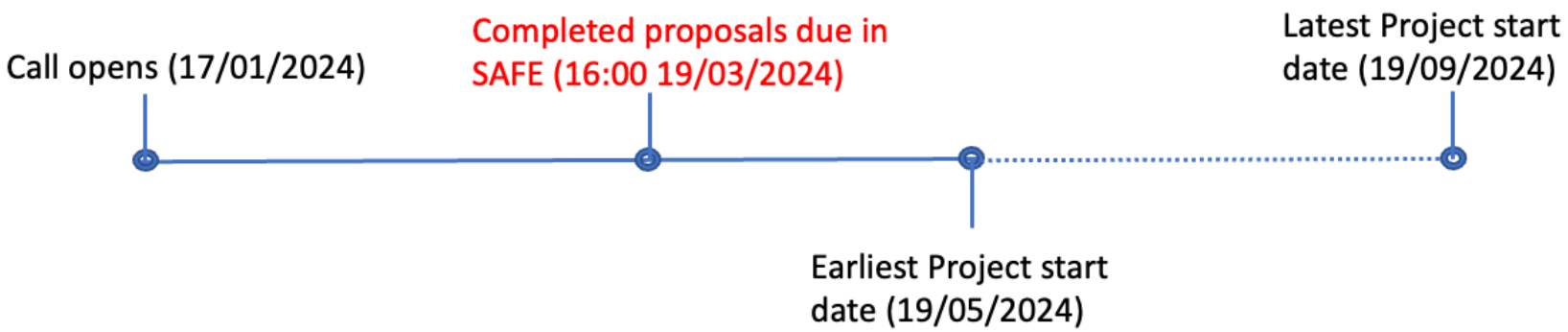
Questions for the discussion/coffee break:

1) What directions for porting to GPU is your community looking at?
2) What combinations of hardware and software have you tested?
3) What works and what doesn't work for your scientific codes?
4) If you have not started thinking about GPUs what would help you start?
5) Is your community's route to using GPU enabled HPC well defined?

Upcoming GPU training courses

- Overview of the ARCHER2 GPU Development Platform
  - 12 - 13 March 2024 09:00 - 17:00 and 09:00 - 15:30
  - Location: Online
  - Level: Advanced
  - Audience: Software Developers
- Introduction to GPU programming with HIP
  - 18 - 19 April 2024 09:30 - 16:00
  - Location: Online
  - Level: Intermediate
  - Audience: Software Developers

GPU eCSE Programme



- Programme of GPU eCSE software development calls
  - Expect to run 2-3 calls this year
  - 1st call open now
- Open to proposals to support research across UKRIs remit
- Up to 36 person months of effort available per call
  - Max duration 2 years
  - Flexibility of how effort spend (e.g. 1 person 50% for 2 years, 2 people 100% for 18 months, etc.)
  - Funding can be for RSE at PI's institution, RSE at third-party institution, member of ARCHER2 CSE team – or combination of the above

https://www.archer2.ac.uk/ecse/calls/

Summary

- GPU Development Platform available to all ARCHER2 users
  - Small resource – aimed at development and testing, not production research
- 4 nodes, each with 4 AMD Instinct MI210 GPU
- Shared node access, can request a single GPU
- Exclusive node access – maximum of 2 nodes per job
- Documentation and guidance are work in progress and will be expanded
- GPU-based eCSE call open – closing date 19 Mar 2024

https://docs.archer2.ac.uk/user-guide/gpu/

Capability Days

epcc

archer2

THE UNIVERSITY of EDINBURGH

ARCHER2 solid state storage – scratch file system

Kieran Leach, EPCC, The University of Edinburgh
k.leach@epcc.ed.ac.uk

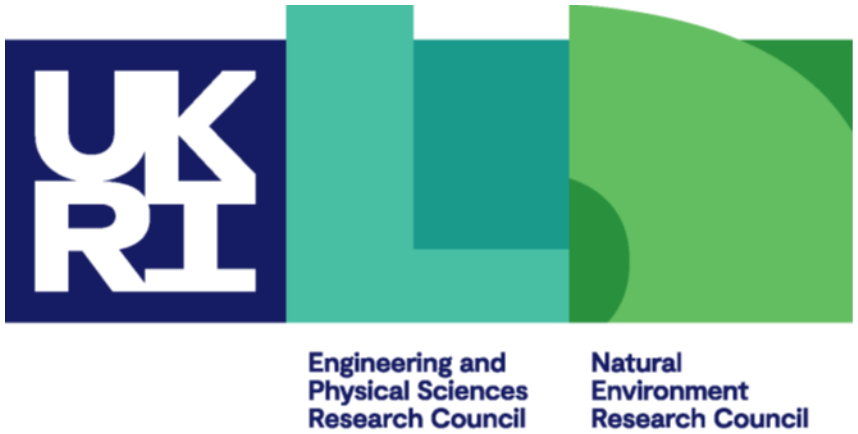
8 March 2024

www.archer2.ac.uk

Reusing this material

Partners

epcc



epcc

THE UNIVERSITY of EDINBURGH

Hewlett Packard Enterprise

Solid state (NVMe) storage technology

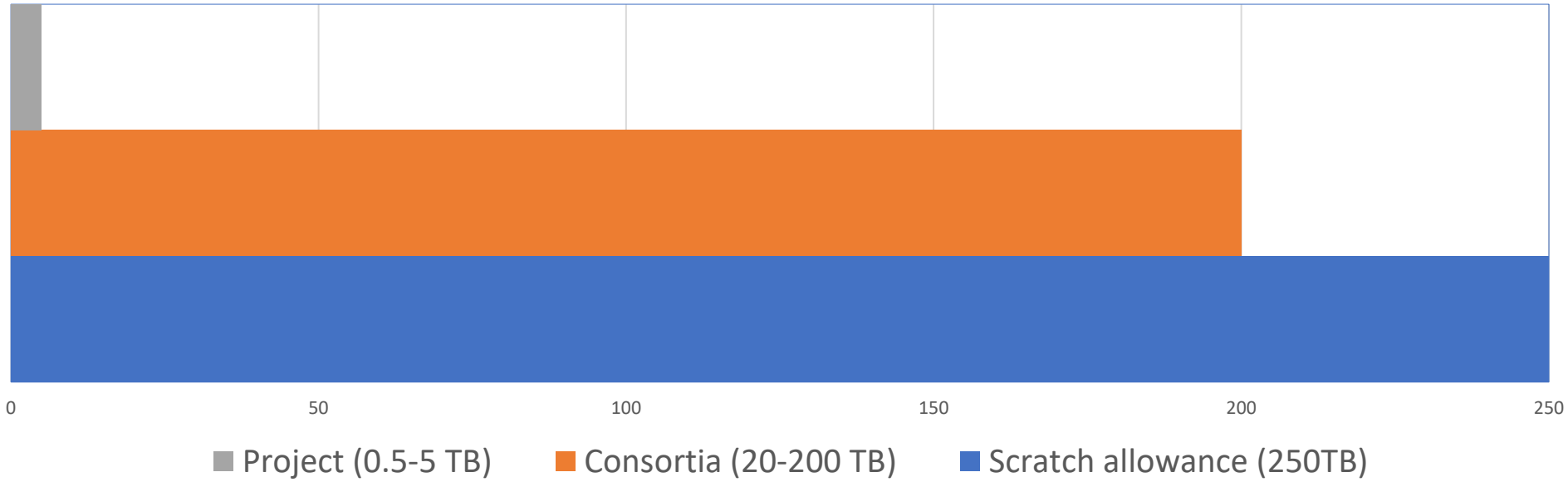
- Storage provided ClusterStor E1000F Lustre appliance
- 1.1PB of NVMe flash storage
- Same Lustre as the /work file systems
- Data storage across 10 storage servers on 240 x 6.4TB PCI4 NVMe SSD
- Metadata storage via independent metadata server on 24 x 1.6TB PCI4 NVMe SSD
- Total bandwidth of 8,400 Gbit/s to storage and metadata servers

Access and quotas

- Access to solid state storage available on request to service desk
- Increased performance compared to /work file systems for some IO patterns
  - If you use MPI-IO approaches to reading/writing data - this includes parallel HDF5 and parallel NetCDF - unlikely to see any performance improvements from using the solid state storage over the standard parallel Lustre file systems
- No traditional quotas in place for allocation
- Each project limited to 250TB to prevent swamping of file system
- Applied as a blanket quota across a project
  - Shared by all project members
- No ability for PIs to manage quotas on solid state storage
- This may be reviewed in future if necessary

/scratch-nvme

- For comparison:
  - Projects receive on the order of 500GB-5TB
  - Consortia receive on the order of 20TB-200TB



  - However…

Scratch file system setup (data purge policy)

- **All files older than 28 days are automatically deleted**
- This is to ensure that capacity is retained for all users, and to ensure that this space is used as genuine “scratch” temporary storage
- Deletion is based on the last access time so files which are in active use should not be impacted
- Deletion is triggered every 24 hours
- You can identify candidate files for deletion using:

```
find /mnt/lustre/a2fs-nvme/work/<project code> -atime +28 -type f -print
```

Summary

Summary

- 1.1 PB of solid state storage available
  - Mounted on login, data analysis and compute nodes
  - Lustre based
- Access to solid state storage available on request to service desk
- All projects have a blanket 250 TB quota
- Configured as a scratch file system – files not accessed in last 28 days are automatically deleted
  - Automated deletion runs daily
- May see improved performance over spinning disk Lustre file systems for some IO patterns
  - Unlikely to see improvement for MPI-IO, parallel NetCDF, parallel HDF5 use

https://docs.archer2.ac.uk/user-guide/data/#solid-state-nvme-file-system-scratch-storage

ARCHER2 Training

Juan F. R. Herrera, EPCC, The University of Edinburgh

ARCHER2 User Forum, 8th March 2024

www.archer2.ac.uk

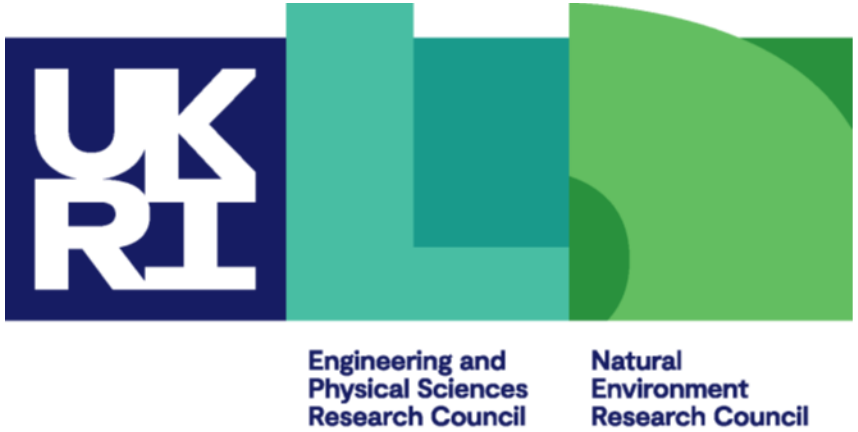
archer2

epcc

Reusing this material

Partners

epcc



epcc

THE UNIVERSITY of EDINBURGH

Hewlett Packard Enterprise

Contents

- Introduction
- ARCHER2 Training Programme 2023/24 draft
- Questions/discussion

Introduction

- 60+ days of interactive face-to-face and online courses and webinars.
- Course catalogue aimed to cover users' needs.
- Self-service courses also available:
  - MPI
  - OpenMP
  - GROMACS+CP2K
  - Intro to HPC
- Past course materials and recordings available on the ARCHER2 website.
- Feedback from:
  - Past course attendees.
  - ARCHER2 training forum.
  - ARCHER2 training panel.

ARCHER2 Training Programme 2024/25 draft

Topic	Number of days
Software Carpentry (#2), Data Carpentry (#2)	8
HPC Carpentry (#3)	6
Introduction and Advanced MPI	4
Introduction and Advanced OpenMP	4
Modern C++ (#2)	4
Introduction and Intermediate Fortran	4
Intro to Data Science and ML, AI, Data Analysis and Visualisation in Python	6
GPU training (#2)	4
Containers, Efficient Parallel IO, ReFrame, Single-Node Opt.	8
Training delivered by HPE	3
Webinars (#10)	10
	61

Questions/discussion

- What topics would you add/remove/keep?
- Is there any course that you (or your consortium) might be particularly interested in?
- Would you prefer to enrol in any of the below courses in a self-service format?

Topic	Number of days
Software Carpentry (#2), Data Carpentry (#2)	8
HPC Carpentry (#3)	6
Introduction and advanced MPI	4
Introduction and advanced OpenMP	4
Modern C++ (#2)	4
Introduction and Intermediate Fortran	4
Intro to Data Science and ML, AI, Data Analysis and Visualisation in Python	6
GPU training (#2)	4
Containers, Efficient Parallel IO, ReFrame, Single-Node Opt.	8
Training delivered by HPE	3
Webinars (#10)	10
	61

Questions?

epcc

archer2

THE UNIVERSITY of EDINBURGH