Introduction to ARCHER

Outline of course

Reusing this material

archer

epcc

THE UNIVERSITY OF EDINBURGH

Course Parameters

- Pre-requisites
  - Familiarity with parallel programming is assumed
  - E.g. you have previous experience on
    - other HPC systems
    - compute clusters
  - Or have attended the “Hands-on Introduction to HPC” course

Learning Outcomes

- On completion of this course attendees should be able to:
  - Understand the ARCHER hardware environment.
  - Compile and run parallel programs on ARCHER.
  - Port applications to ARCHER.
  - Undertake the ARCHER driving test
    - http://www.archer.ac.uk/training/course-material/online/driving_test.php
    - on successful completion of the test, you can apply for an account an 80,000 core-hours (1,200 “kAUs”) on ARCHER for a 12-month period

ARCHER Service

Overview and Introduction

ARCHER Partners

- EPSRC
  - Managing partner on behalf of RCUK
- Cray
  - Hardware provider
- EPCC
  - Service Provision (SP) – Systems, Helpdesk, Administration, Overall Management (also input from STFC Daresbury Laboratory)
  - Computational Science and Engineering (CSE) – In-depth support, training, embedded CSE (eCSE) funding calls
  - Hosting of hardware – datacentre, infrastructure, etc.

EPCC's Advanced Computing Facility

archer

epcc

THE UNIVERSITY OF EDINBURGH

ARCHER in a nutshell

- UK National Supercomputing Service
  - £43 million 4-year project from 2013
- Cray XC30 Hardware
  - Nodes based on 2×Intel Ivy Bridge 12-core processors
  - 64GB (or 128GB) memory per node
  - 4920 nodes in total (118080 cores)
  - Linked by Cray Aries interconnect (dragonfly topology)
- Cray Application Development Environment
  - Cray, Intel, GNU Compilers (all support OpenMP)
  - Cray Parallel Libraries (including optimised MPI)
  - DDT Debugger, Cray Performance Analysis Tools

epcc

Storage

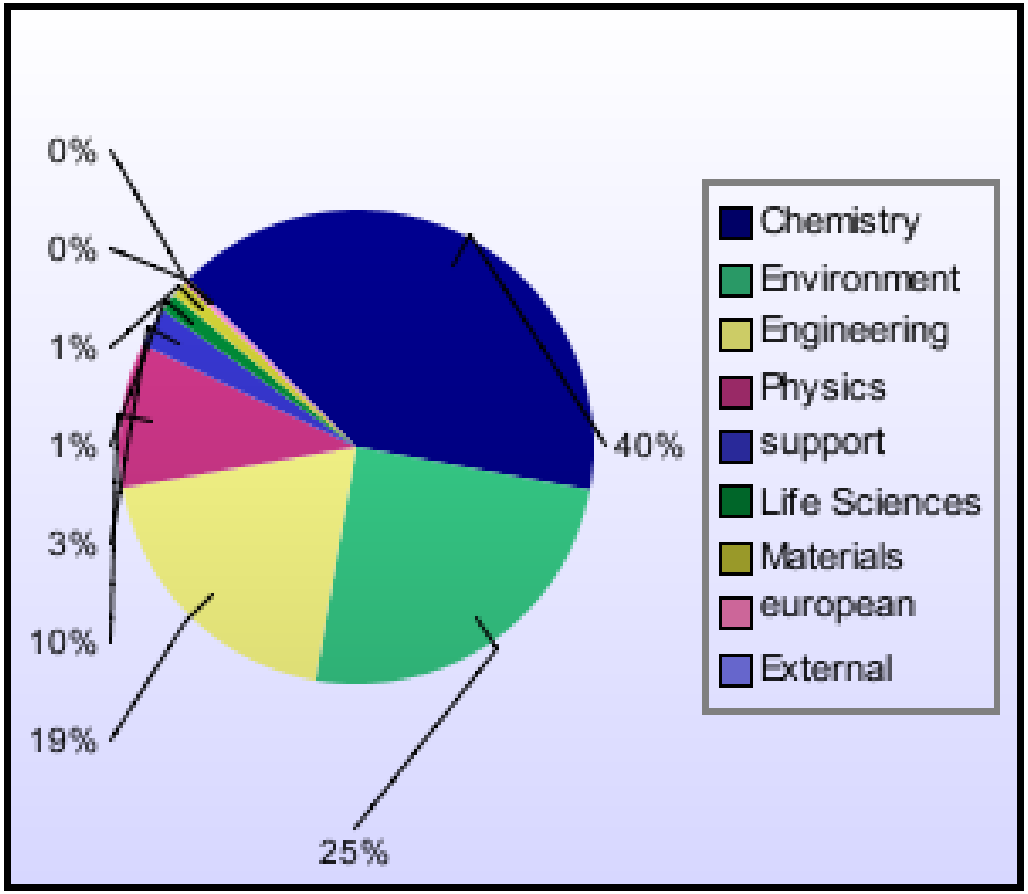
- /home – NFS mounted, not accessible on compute nodes
  - For source code and critical files
  - **Backed up**
  - > 200 TB total
- /work – Lustre, accessible on all nodes
  - High-performance parallel filesystem
  - **Not backed-up**
  - > 4PB total
- RDF – GPFS, not accessible on compute nodes
  - Research Data Facility
  - **Long term data storage**

Data Management

- Most of your online storage is in /work/
  - fast but not backed up!
  - you **must not** rely on /work/ for storage of critical data
  - you should copy elsewhere, e.g. the RDF

- For advice and instructions
  - http://www.archer.ac.uk/documentation/data-management/

What is it used for?



Summary

- ARCHER is a Cray XC30
  - It uses standard Intel processors
    - 2 processors per node, 24 cores per shared-memory Linux node
    - 64 GB memory on the majority of nodes (some have 128 GB)
    - Nodes similar to many other HPC systems
  - Cray ARIES switch
    - High performance, optimised for large parallel jobs
    - Standard usage but can get very good performance
  - Large storage and high performance filesystem
    - 4 PB high performance filesystem; 200 TB home space
  - Intel, GNU, and Cray compilers
    - Lots of standard scientific packages, libraries, and software installed
  - Compute nodes accessed via PBS batch system