The Carpentries and HPC

CarpentryCon sessions

- Breakout
  - Alan O'Cais, Juelich Supercomputing Centre
  - Daniel G. A. Smith, The Molecular Sciences Software Institute
  - Andrew Turner, EPCC (Report)
- Workshop
  - Peter Steinbach, Scionics Computer Innovation
  - Christina Koch, University of Wisconsin

Keeping track of your opinions

- We will use Socrative to gather your opinions and help drive the discussion
  - https://b.socrative.com/student
  - Room: ALAN1854

What's been done?

- Birds-of-a-Feather session at SC17
  - Report
- Two novice lessons are "in the wild"
  - HPC-in-a-day
  - HPC intro

BoF Learner Profile

- Vague idea of what HPC is but not sure how this translates to their research.
- Lack of experience with Linux, command line, text editors, batch systems etc.
- Lack of knowledge about what HPC systems are: how are they put together, how do they enable faster/larger calculations and how are resources shared?
- Want to know how to support themselves when things don't work and keen to learn more.

Prerequisites

- **BoF** *Basic Linux shell use and scripting*
- **hpc-intro** *"There are no real prerequisites for this lesson, but prior programming and/or command line experience will be helpful."*
- **hpc-in-a-day** *"If you have already written small programs with a language of your choice and know the difference between a 'variable' and a 'function' and obtain a minimal knowledge of using the UNIX command line."*

Topics

Topic	hpc-intro	hpc-in-a-day
Why use HPC?	00:00 What is HPC?	00:00 Taking the space shuttle
Login, interactive access; transferring data	00:20 Connecting to the cluster 03:35 Working on a Cluster 06:00 Transferring Files	00:00 Taking the space shuttle
Understanding HPC jargon	Covered across episodes	Covered across episodes
Basic understanding of HPC architectures	03:35 Working on a cluster	01:20 Batch systems and schedulers 101 05:25 Bonus session: Distributing computations among computers
File systems on HPC systems	00:40 Moving around and looking at things 01:00 Writing and reading files	00:50 Navigating files and directories 02:35 Working with the shared file system

In-terminal text editors	01:00 Writing and reading files	01:20 Batch systems and schedulers 101 (mentioned)
Modules and Environment	05:15 Accessing software	(Not covered)
The batch system	04:00 Scheduling jobs 06:40 Using resources effectively	01:20 Batch systems and schedulers 101 02:05 Working with the scheduler
Shared system etiquette	07:05 Using resources effectively	
Troubleshooting strategies	(Not covered)	02:05 Working with the scheduler

Other Topics		
Command line and shell scripting	01:45 Wildcards and pipes 02:40 Scripts, variables and loops	(Not covered)
Performance and profiling	(Not covered)	03:05 Estimation of Pi for Pedestrians
Parallel concepts and MPI programming	(Not covered)	03:55 Parallel Estimation of Pi for Pedestrians 04:40 Searching for Pi 05:25 Bonus session: Distributing computations among computers

Portability

- What aspects affect the portability of lessons
  - Access (keys, passwords, ...)
  - Schedulers (queues, limits)
  - File systems (number, type)
  - Module environment (module names)
- Both lessons have made efforts to make the material generic
- HPC-in-a-day uses settings in the "_config.yml" file for Jekyll functionality to customise the lesson
  - File system names, scheduler, code snippets

Thanks!

Questions?