ARCHER2 for Data Scientists

Practical 2: Filesystem performance

Adrian Jackson

1 Introduction

In this exercise we are going to investigate filesystem performance on ARCHER2.

2 Filesystem performance

To get started move into the `IO` directory in the course repository on your `work` filesystem on ARCHER2.

```
cd archer2-data-science/exercises/IO
```

In there, there is a archive file called `IOR.tar.gz`. You should be able unpack both with the commands:

`tar xf IOR.tar.gz`.

For IOR you need to go into the IOR directory and type:

```
make mpiio
```

Then you can run a Lustre IOR benchmark using:

```
sbatch lustre_ior.sh
```

The purpose of this practical is to run IOR and investigate the performance of two different configurations of the IOR benchmark, IOR easy (which creates a file per process for I/O) and IOR hard (which creates one shared file for all processes).

IOR is a bandwidth benchmark for reading and writing files. It should provide a good estimate of how quickly data can be written to and retrieved from a filesystem for a given configuration (number of files, size of file, size of I/O operations, etc…).

The current batch script will run both benchmarks (easy and hard) using the default Lustre configuration on ARCHER2 using a single node (note this will not provide an indication of the total I/O bandwidth of ARCHER2, for that you need to use many more than a single node, but it will provide a reasonable indiciation of the bandwidth of a single node). Run the batch script as outlined above and see what results are produced.

Once you have done that, you can alter the batch script to change the default Lustre striping for both benchmarks. Add the following command after the `mkdir` command for both benchmarks in the batch script to see what impact striping has on performance:

```
lfs setstripe -c -1 data
```

Note, it is important that this command is run after the `mkdir` command in the script but before the `srun`. That way it will configure Lustre for any new files produced by IOR.

What impact does Lustre striping have on I/O bandwidth for easy (single file per worker) and hard (single file for all workers) configurations?