

Sharpen Exercise

Using HPC resources and running parallel applications

Andrew Turner, Dominic Sloan-Murphy, David Henty, Adrian Jackson

March 26, 2015

Contents

| | | |
|----------|---|----------|
| 1 | Aims | 1 |
| 2 | Introduction | 2 |
| 3 | Instructions | 2 |
| 3.1 | Log into ARCHER frontend nodes and run commands | 2 |
| 3.1.1 | Procedure for Mac and Linux users | 2 |
| 3.1.2 | Procedure for Windows users | 2 |
| 3.2 | Running commands | 3 |
| 3.3 | Download and extract the exercise files | 3 |
| 3.4 | Using the Emacs text editor | 4 |
| 3.5 | Useful commands for examining files | 5 |
| 3.6 | Compile the source code to produce an executable file | 5 |
| 3.7 | Running a serial job | 6 |
| 3.8 | Viewing the images | 6 |
| 3.9 | Running on the compute nodes | 7 |
| 3.9.1 | Using PBS job scripts | 7 |
| 3.9.2 | Submitting scripts to PBS | 7 |
| 3.9.3 | Monitoring/deleting your batch job | 8 |
| 3.9.4 | Finding the output | 8 |
| 3.9.5 | Extra exercise | 8 |
| 3.10 | Running a parallel job on the compute nodes | 9 |
| 4 | Parallel Performance | 9 |
| 4.1 | Amdahl's law | 9 |
| 4.2 | OpenMP code | 9 |

1 Aims

The aim of this exercise is to get you used to logging into an HPC resource, using the command line and an editor to manipulate files, and using the batch submission system. We will be using ARCHER for this exercise. ARCHER is the UK national HPC service, and is a Cray XC30 system with a total of 118,010 cores (4920 nodes).

You can find more details on ARCHER and how to use it in the User Guide at:

- <http://www.archer.ac.uk/documentation/user-guide/>

2 Introduction

In this exercise you will run a simple program, in serial and parallel, to sharpen the provided image.

Using your provided guest account, you will:

1. log onto the ARCHER frontend nodes;
2. copy the source code from a central location to your account;
3. unpack the source code archive;
4. compile the source code to produce an executable file;
5. run a serial job on the login node;
6. submit a serial job to the compute nodes using the PBS batch system;
7. submit a parallel job using the PBS batch system;
8. run the parallel executable on using an increasing number of cores and examine the performance improvement.

Demonstrators will be on hand to help you as required. Please do ask questions if you do not understand anything in the instructions - this is what the demonstrators are here for.

3 Instructions

3.1 Log into ARCHER frontend nodes and run commands

You should have been given a guest account ID – referred to generically here as `guestXX` – and password. If you have not, please contact a demonstrator.

3.1.1 Procedure for Mac and Linux users

Open a command line *Terminal* and enter the following command:

```
user@laptop$ ssh -X guestXX@login.archer.ac.uk
Password:
```

you should be prompted to enter your password.

3.1.2 Procedure for Windows users

Windows does not generally have SSH installed by default so some extra work is required. You need to download and install a SSH client application - PuTTY is a good choice:

- <http://www.chiark.greenend.org.uk/~sgtatham/putty/>.

When you start PuTTY you should be able to enter the ARCHER login address (`login.archer.ac.uk`). When you connect you will be prompted for your username and password.

By default, PuTTY does not send graphics back to your local machine. We will need this later for viewing the sharpened image, so you should “Enable X11 Forwarding” which is an option in the “Category” menu under “Connection -> SSH -> X11”. You will need to do this each time before you log in with PuTTY.

3.2 Running commands

You can list the directories and files available by using the *ls* (LiSt) command:

```
guestXX@archer:~> ls
bin  work
```

NB: The first time you do this there will be no files or directories so 'ls' will return with an empty line.

You can modify the behaviour of commands by adding options. Options are usually letters or words preceded by '-'. For example, to see more details of the files and directories available you can add the '-l' (l for long) option to *ls*:

```
guestXX@archer:~> ls -l
total 8
drwxr-sr-x 2 user  z01 4096 Nov 13 14:47 bin
drwxr-sr-x 2 user  z01 4096 Nov 13 14:47 work
```

If you want a description of a particular command and the options available you can access this using the *man* (MANual) command. For example, to show more information on *ls*:

```
guestXX@archer:~> man ls
Man: find all matching manual pages
* ls (1)
  ls (1p)
Man: What manual page do you want?
Man:
```

In the manual, use the spacebar to move down, 'u' to move up, and 'q' to quit to the command line.

3.3 Download and extract the exercise files

Firstly, change directory to make sure you are on the "/work" filesystem on ARCHER.

```
guestXX@archer:~> cd /work/y14/y14/guestXX/
```

/work is a high performance parallel file system that can be accessed by both the frontend and compute nodes. **All jobs on ARCHER should be run from the /work filesystem.** ARCHER compute nodes cannot access the /home filesystem at all. Any jobs attempting to use /home will fail with an error.

Use *wget* (on ARCHER) to get the exercise files archive from the EPCC webserver. Material for the courses is stored in a standard location of the form:

```
http://www.archer.ac.uk/training/course-material/YYYY/MM/CourseName_Location/.
```

If you access the material with a web browser it will be downloaded to your laptop, which is not very useful if we want the files on ARCHER. It is easier to copy them straight from the website to ARCHER:

```
guestXX@archer:~> wget http://www.archer.ac.uk/.../Exercises/sharpen.tar.gz
--2014-06-27 16:15:42-- http://www.archer.ac.uk/training/course-material/2014/06/Int
Resolving www.archer.ac.uk... 193.62.216.12
Connecting to www.archer.ac.uk|193.62.216.12|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 1754173 (1.7M) [application/x-gzip]
Saving to: `sharpen.tar.gz'
```

```
100%[=====>] 1,754,173  --.-K/s  in 0.02s
```

```
2014-06-27 16:15:42 (107 MB/s) - `sharpen.tar.gz.1' saved [1754173/1754173]
```

To unpack the archive:

```
guestXX@archer:~> tar -xzvf sharpen.tar.gz
sharpen/C-SER/
sharpen/C-SER/filter.c
...
--snip--
...
sharpen/F-OMP/dosharpen.f90
sharpen/F-OMP/Makefile
sharpen/F-OMP/fuzzy.pgm
```

If you are interested in the C examples move to the *C-SER* subdirectory; for Fortran, move to *F-SER*. For example:

```
guestXX@archer:~> cd sharpen/C-SER
```

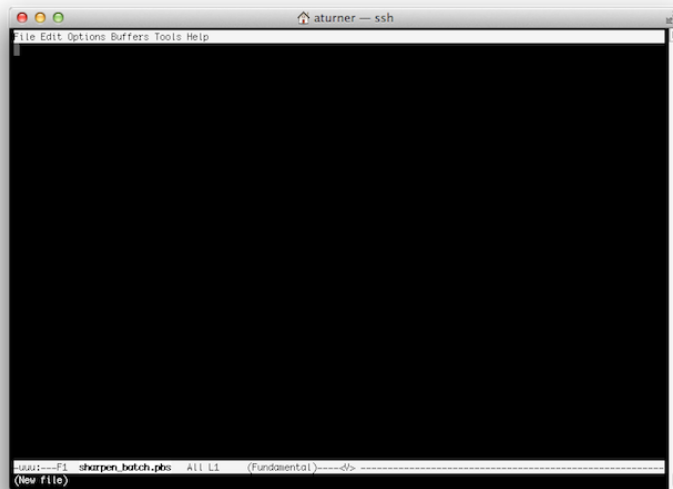
3.4 Using the Emacs text editor

Running interactive graphical applications on ARCHER can be very slow. It is therefore best to use Emacs will be used in *in-terminal* mode. In this mode you can edit the file as usual but you must use keyboard shortcuts to run operations such as “save file” (remember, there are no menus that can be accessed using a mouse).

Start Emacs with the *emacs -nw* command and the name of the file you wish to edit or create. For example:

```
guestXX@archer:~> emacs -nw sharpen.pbs
```

The terminal will change to show that you are now inside the Emacs text editor:



Typing will insert text as you would expect and backspace will delete text. You use special key sequences (involving the Ctrl and Alt buttons) to save files, exit Emacs and so on.

Files can be saved using the sequence “Ctrl-x Ctrl-s” (usually abbreviated in Emacs documentation to “C-x C-s”). You should see the following briefly appear in the line at the bottom of the window (the minibuffer in Emacs-speak):

```
Wrote ./sharpen.pbs
```

To exit Emacs and return to the command line use the sequence “C-x C-c”. If you have changes in the file that have not yet been saved Emacs will prompt you (in the minibuffer) to ask if you want to save the changes or not.

Although you could edit files on your local machine using whichever windowed text editor you prefer it is useful to know enough to use an in-terminal editor as there will be times where you want to perform a quick edit that does not justify the hassle of editing and re-uploading.

3.5 Useful commands for examining files

There are a couple of commands that are useful for displaying the contents of plain text files on the command line that you can use to examine the contents of a file without having to open in Emacs (if you want to edit a file then you will need to use Emacs). The commands are *cat* and *less*. *cat* simply prints the contents of the file to the terminal window and returns to the command line. For example:

```
guestXX@archer:~> cat sharpen.pbs
...
```

This is fine for small files where the text fits in a single terminal window. For longer files you can use the *less* command: *less* gives you the ability to scroll up and down in the specified file. For example:

```
guestXX@archer:~> less Makefile
```

Once in *less* you can use the spacebar to scroll down and ‘u’ to scroll up. When you have finished examining the file you can use ‘q’ to exit *less* and return to the command line.

This program takes a fuzzy image and uses a simple algorithm to sharpen the image. A very basic parallel version of the algorithm has been implemented which we will use in this exercise. There are a number of versions of the sharpen program available:

C-SER Serial C version

F-SER Serial Fortran version

C-MPI Parallel C version using MPI

F-MPI Parallel Fortran version using MPI

C-OMP Parallel C version using OpenMP

F-OMP Parallel Fortran version using OpenMP

3.6 Compile the source code to produce an executable file

We will first compile the serial version of the code for our example.

```
guestXX@archer:~> ls
cio.c          filter.c      Makefile     sharpen.h    utilities.c
dosharpen.c   fuzzy.pgm    sharpen.c    sharpen.pbs  utilities.h
```

```
guestXX@archer:~> make
cc -g -DC_SERIAL_PRACTICAL -c sharpen.c
cc -g -DC_SERIAL_PRACTICAL -c dosharpen.c
cc -g -DC_SERIAL_PRACTICAL -c filter.c
cc -g -DC_SERIAL_PRACTICAL -c cio.c
cc -g -DC_SERIAL_PRACTICAL -c utilities.c
cc -g -DC_SERIAL_PRACTICAL -o sharpen sharpen.o dosharpen.o filter.o cio.o utilities.o
```

```
guestXX@archer:~>
```

This should produce an executable file called *sharpen* which we will run on ARCHER.

For the Fortran version, the process is exactly the same as above, except you should move to the *F-SER* subdirectory and build the program there:

```
guestXX@archer:~> cd sharpen/F-SER
guestXX@archer:~> make
...
```

As before, this should produce a *sharpen* executable.

Don't worry about the C file *utilities.c* – it is just providing an easy method for printing out various information about the program at run time, and it is most easily implemented in C.

3.7 Running a serial job

You can run this serial program directly on the login nodes, e.g.:

```
guestXX@archer:~> ./sharpen
Image sharpening code running in serial

Input file is: fuzzy.pgm
Image size is 564 x 770

Using a filter of size 17 x 17

Reading image file: fuzzy.pgm
... done

Starting calculation ...
On core 0-31
... finished

Writing output file: sharpened.pgm

... done

Calculation time was 5.579000 seconds
Overall run time was 5.671895 seconds
```

3.8 Viewing the images

To see the effect of the sharpening algorithm, you can view the images using the *display* program from the ImageMagick suite. You must load a module to access *display* – note that you only need to load this module once when you first log in.

```
guestXX@archer:~> module load imagemagick
guestXX@archer:~> display fuzzy.pgm
guestXX@archer:~> display sharpened.pgm
```

Type “q” in the image window to close the program.

To view the image you will need an X window client installed. Linux or Mac systems will generally have such a program available, but Windows does not provide X windows functionality by default. There are many X window systems available to install on Windows; we recommend Xming available at:

- <http://sourceforge.net/projects/xming/>

3.9 Running on the compute nodes

As with other HPC systems, use of the compute nodes on ARCHER is mediated by the PBS job submission system. This is used to ensure that all users get access to their fair share of resources, to make sure that the machine is as efficiently used as possible and to allow users to run jobs without having to be physically logged in.

Whilst it is possible to run interactive jobs (jobs where you log directly into the backend nodes on ARCHER and run your executable there) on ARCHER, and they are useful for debugging and development, they are not ideal for running long and/or large numbers of production jobs as you need to be physically interacting with the system to use them.

The solution to this, and the method that users generally use to run jobs on systems like ARCHER, is to run in *batch* mode. In this case you put the commands you wish to run in a file (called a job script) and the system executes the commands in sequence for you with no need for you to be interacting.

3.9.1 Using PBS job scripts

We will first run the same serial program on the compute nodes. Look at the batch script:

```
guestXX@archer:~> emacs sharpen.pbs
```

The first line specifies which *shell* to use to interpret the commands we include in the script. Here we use the Bourne Again SHell (bash) which is the default on most modern systems. The `-login` option tells the shell to behave as if it was an interactive shell.

The line `-l select=[nodes]` is used to request the total number of compute nodes required for your job (1 in the example above).

The `#PBS` lines provide options to the job submission system where `"-l select"` specifies that we want to reserve 1 compute node for our job - the minimum job size on ARCHER is 1 node (24 cores); the `"-l walltime=00:01:00"` sets the maximum job length to 1 minute; `"-A y14"` sets the budget to charge the job to "y14"; `"-N sharpen"` sets the job name to "sharpen".

The remaining lines are the commands to be executed in the job. Here we have a comment beginning with `"#"`, a directory change to `$PBS_O_WORKDIR` (an environment variable that specifies the directory the job was submitted from) and the `aprun` command (this command tells the system to run the jobs on the compute nodes rather than the frontend nodes).

Jobs can only be run on the compute nodes with the `aprun` parallel job launcher. The line

```
aprun -n 1 ./sharpen
```

runs a single copy of our serial executable.

3.9.2 Submitting scripts to PBS

Simply use the `qsub` command with the reservation ID (i.e. (replace `<resID>` in the command below with the reservation ID provided by the trainer) and the job submission script name:

```
guestXX@archer:~> qsub -q <resID> sharpen.pbs
58306.sdb
```

The `jobID` returned from the `qsub` command is used as part of the names of the output files discussed below and also when you want to delete the job (for example, you have submitted the job by mistake).

3.9.3 Monitoring/deleting your batch job

The PBS command *qstat* can be used to examine the batch queues and see if your job is queued, running or complete. *qstat* on its own will list all the jobs on ARCHER (usually hundreds) so you can use the “-u \$USER” option to only show your jobs:

```
guestXX@archer:~> qstat -u $USER
```

```
sdb:
```

| Job ID | Username | Queue | Jobname | SessID | NDS | TSK | Req'd Memory | Req'd Time | Elap S | Time |
|-----------|----------|----------|---------|--------|-----|-----|-----------------|---------------|-----------|------|
| 58306.sdb | guest01 | standard | sharpen | -- | 1 | 24 | -- | 00:01 | Q | -- |

“Q” means the job is queued, “R” that it is running and “E” that it has recently completed; if you do not see your job, it usually means that it has completed.

If you want to delete a job, you can use the *qdel* command with the jobID. For example:

```
guestXX@archer:~> qdel 58306.sdb
```

3.9.4 Finding the output

The job submission system places the output from your job into two files: <job name>.o<jobID> and <job name>.e<jobID> (note that the files are only produced on completion of the job). The sharpen.o<jobID> file contains the output from your job sharpen.e<jobID> contains any errors.

```
guestXX@archer:~> cat sharpen.o58306
```

```
Image sharpening code running in serial
```

```
Input file is: fuzzy.pgm
```

```
Image size is 564 x 770
```

```
Using a filter of size 17 x 17
```

```
Reading image file: fuzzy.pgm
```

```
... done
```

```
Starting calculation ...
```

```
On core 0
```

```
... finished
```

```
Writing output file: sharpened.pgm
```

```
... done
```

```
Calculation time was 5.400482 seconds
```

```
Overall run time was 5.496556 seconds
```

```
Application 166570 resources: utime ~6s, stime ~0s, Rss ~18024, inblocks ~11176, outb...
```

3.9.5 Extra exercise

You could try running more than one copy of the serial program by changing the “-n” argument to “aprun”. What do you observe?

3.10 Running a parallel job on the compute nodes

Repeat the same procedure as above but use the parallel MPI version of the code in *C-MPI* or *F-MPI*. By default, the batch script is set up to run on 4 cores. You should see that the parallel code is faster than the serial one.

4 Parallel Performance

4.1 Amdahl's law

If you examine the log file you will see that it contains two timings: the total time taken by the entire program (including IO) and the time taken solely by the calculation. The image input and output is not parallelised so this is a serial overhead, performed by a single processor. The calculation part is, in theory, perfectly parallel (each processor operates on different parts of the image) so this should get faster on more cores.

You should do a number of runs and fill in Table 1: the IO time is the difference between the calculation time and the overall run time; the total CPU time is the calculation time multiplied by the number of cores. Note that if you run on more than 24 cores then you will need to request more than one node with the "select=" option to PBS.

Look at your results – do they make sense? Given the structure of the code, you would expect the IO time to be roughly constant, and the performance of the calculation to increase linearly with the number of cores: this would give a roughly constant figure for the total CPU time. Is this what you observe?

The parallel speedup should closely follow Amdahl's law as the calculation splits cleanly into a purely serial phase (IO) and a perfectly parallel phase (calculation).

4.2 OpenMP code

If you are interested, you can try running the OpenMP code in *C-OMP* or *F-OMP*.

Note that to change the number of cores that the program uses you must set the environment variable `OMP_NUM_THREADS`. This is already done for you in the PBS script, although you will need to change the actual number yourself.

Note that you can run using multiple threads on the login nodes – you must set `OMP_NUM_THREADS` before you run, e.g.:

```
guestXX@archer:~> export OMP_NUM_THREADS=3
guestXX@archer:~> ./sharpen
```

```
Image sharpening code running on 3 thread(s)
```

```
...
```

| # Cores | Overall run time | Calculation time | IO time | Total CPU time |
|---------|------------------|------------------|---------|----------------|
| 1 | | | | |
| 2 | | | | |
| 4 | | | | |
| 7 | | | | |
| 10 | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| # Cores | Overall run time | Calculation time | IO time | Total CPU time |
| 1 | | | | |
| 2 | | | | |
| 4 | | | | |
| 7 | | | | |
| 10 | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| # Cores | Overall run time | Calculation time | IO time | Total CPU time |
| 1 | | | | |
| 2 | | | | |
| 4 | | | | |
| 7 | | | | |
| 10 | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |

Table 1: Time taken by parallel image processing code