# EXDCI Case Studies: HPC talent generation

Eilidh Troup: Applications Consultant at EPCC

Eilidh Troup is an Applications Consultant at Edinburgh Parallel Computing Centre (EPCC). She has a background in genetics, but discovered a passion for programming during a computational project in her final year at university, and is delighted to be able to combine these two interests at EPCC, working on HPC projects with a biological focus. Outside work, she loves cycling and reading.
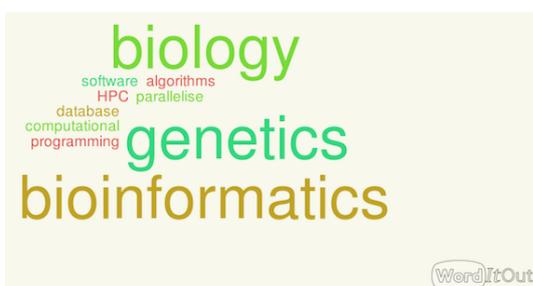
EXDCI is a European-funded project led by PRACE and ETP4HPC, the two most significant HPC bodies in Europe. It aims to co-ordinate the development and implementation of a common strategy for the European HPC ecosystem by supporting road-mapping, strategy-making and performance-monitoring activities.

This is one of a series of case studies designed to demonstrate the range of interesting careers in High Performance Computing (HPC). More case studies are available at http://www.exdci.eu/CaseStudies/

Eilidh, tell us a bit about yourself.

I work as an Applications Consultant at EPCC at the University of Edinburgh. My first degree was in genetics at the University of Glasgow, so I particularly enjoy the opportunities that I get in EPCC to work on projects with a biological focus.

How did your career path lead you to where you are now?

I loved genetics but decided lab work wasn't for me. In my final year I did a computational project and enjoyed it greatly, so then did an MSc in IT, followed by a research MSc in bioinformatics. I worked on a computing project for the Institute of Neuroscience and Psychology for a while, then got a job with a pensions company. I started working in HPC when a friend recommended working at EPCC, and I applied for a chance to do something a bit different from regular programming.

What was your first experience of HPC? Did that change the course of your career?

I worked at EPCC for a few years before I asked to work on an HPC project I liked the sound of – SPRINT (http://r-sprint.org/), a parallel version of R. It was challenging to learn R, brush up on C, and figure out how to use the supercomputer all at the same time, but the documentation was good and all of the more experienced HPC programmers were very helpful and supportive. It was a really nice project because SPRINT makes it easy for people who only know R to be able to use HPC. SPRINT takes R functions and provides a parallel version which you can call in just the same way, without needing to get into tricky parallel stuff as a user.

So what does your job involve?

I'm currently working with the School of Biological Sciences at Edinburgh University, setting up a database for their experimental results. Having a biology background makes it easier to work with biologists and understand what they are doing.

I have also been working on a really interesting project with a company which breeds animals. We are looking at the genetics pedigree, and that code is not traditional HPC, but is based on a Hadoop cluster with a spark interface on top. We have some interesting ideas for how to put the data into the database with some data structures which are right for that specific problem.

What are the exciting aspects of working in an HPC-related career?

The best thing is when users are happy with what I have done; when it lets them do something that they couldn't before. That's a really nice feeling.

A couple of years ago I did a piece of work for the biologists, who were studying circadian rhythm – the day and night cycle of plants – and different mutations that can give them different clock lengths. There was a bit of Fortran code that would fit the periodic curve to the gene expression, but it was a bit buggy and some bits of data would just crash when it tried to run. I rewrote it in Java and they were really pleased because it gave the same results and it didn't crash with the new data set that they had. It's really nice when you get a complete piece of work that somebody can just use like that.

I have always really liked working on algorithms, and I like the interesting ideas of different data structures, so that made HPC appeal to me because you can do more unusual stuff with the data structures and algorithms.

Are there any challenging aspects?

I think the big challenge for me moving to HPC work was that I had not done much programming in C before. Some things were not exactly HPC issues, but were definitely hurdles to get past, and these can be hard for your confidence at the beginning. I had already picked up working from the command line and I was all ready to get stuck into the conceptual issues of HPC, such as message passing and optimisation, but it was really some of the very early things which were hard for me. For example, I had some problems with the compilers, as things are set up differently on the HPC systems.

Where do you see your career leading you next?

Coming from a biology background, I was drawn to this job as a chance to work with scientists, and I'm happy working on various biology-related computing projects here. The science is always interesting, and there are plenty of new coding challenges as the amount of data generated in biology is increasing so rapidly.

What is your outlook on the use of HPC in your field?

This is an exciting time for HPC and biology. Physicists have been using HPC for years, and now biologists are beginning to need that kind of computing power too.
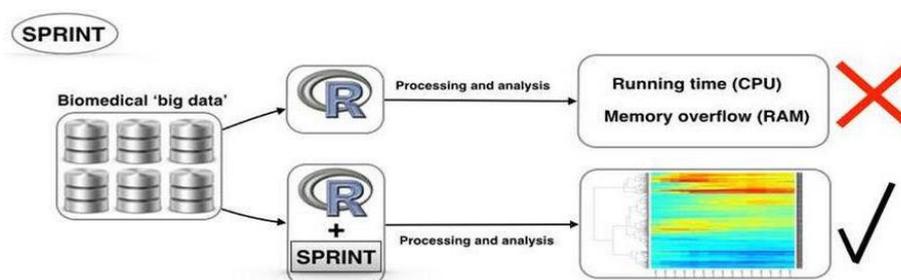
SPRINT helps biomedical researchers use HPC without needing to learn how to parallelise code