Industrial Year Report

Software Developer at STFC, Rutherford Appleton Laboratory

Samuel Jackson
University Of Aberystwyth
slj11@aber.ac.uk

August 25, 2014

Contents

1 Introduction

This report details the industrial placement undertaken by the author as part of a year long placement in industry as part of the Software Engineering MEng course at Aberystwyth university. The position of the placement was a junior role as part of the Mantid data analysis framework development team based at the ISIS facility at Rutherford Appleton Laboratory in Harwell, Oxfordshire, owned by the Science and Technologies Facilities Council (STFC) for the duration of a one year contract of employment. This position was then extended by an additional two months to continue with existing and new project commitments.

ISIS is a world leading neutron and muon scattering facility. The facility operates a 800 MeV proton synchrotron which acts as a source of neutrons for both neutron and muon spectroscopy [1]. Neutron spectroscopy is used to probe the structure and dynamics of materials at the fundamental level.

The Mantid project aims to provide a single, unified application for the analysis of neutron and muon scattering facilities such as ISIS [2]. The project is primarily developed by two teams of developers, one based at ISIS and one at the Spallation Neutron Source (SNS) at Oakridge laboratory in Tennessee, USA.

2 Organisational Environment

2.1 Organisation of STFC

The Science and Technologies Facilities Council is a UK government funded body that carries out a wide variety of scientific research across a multitude of disciplines including particle physics, nuclear physics, space science and engineering, medical and biological sciences, and computational science.

While the organisation is funded by the government, it is classified as a non-governmental body which acts as an umbrella organisation for an array of facilities based across the UK. These include (but are not limited to) the central laser facility, diamond light source (which is a publically limited company of which STFC holds an 86% share), ISIS neutron source and RAL space based at Rutherford Appleton Laboratory in Oxfordshire, the Daresbury Laboratory located in Cheshire and the Chilbolton Observatory based in Hampshire.

The organisation has its head office located at Polaris House, Swindon, Wiltshire and is headed by the Chief Executive John Womersley. The purpose of STFC is to control the general overall management of the facilities under its control, in particular it is responsible for allocating budgetary and staffing allowances and liaising between government departments, particularly the department for business, innovation and skills.

2.2 Organisation of ISIS Neutron Source

ISIS Neutron source is a project that is owned, operated and funded by the STFC. The organisational hierarchy of ISIS is headed by the director Prof. Robert McGreevy and several division heads for individual functional areas within ISIS such as the diffraction, spectroscopy and support, experimental operations, instrumentation, design, and accelerator divisions.

ISIS is also split into a number of research, operations, and experimental support groups. The computing group, of which the Mantid project is a part of, falls into the experimental support group and is technically a part of the scientific computing group which also includes the ICAT data catalogue which manages the data collected from sample runs on the instruments for future analysis.

While this description provides an overview of how the staffing of ISIS can be divided, in practice there tends to be a lot of cross over between sections of the organisation depending on a employees skills and responsibilities. For example, while I was employed as part of the scientific computing group, my line

manager and senior manager were both members of the molecular spectroscopy research group whose interests I was responsible for within Mantid.

2.3 Organisation of the Mantid Project Development Team

The Mantid development team in the UK is a subset of the scientific computing group of ISIS. The Mantid group is headed by a single project manager (Nick Draper) who is based at ISIS where the project first started and is responsible for the overall management and direction of the project. The project is split into two teams, one based at ISIS and one at the SNS in Oakridge, Tennessee. Both teams consist of a single lead developer and several senior developers who oversee the major technical developments and help to guide and manage the rest of the development team. Both teams in the UK and the US also have there own manager, but the project manager based at ISIS is in overall control. Within Mantid, the project manager and the majority of the senior developers are actually contractors from Tessella Ltd, based in Abingdon, Oxfordshire. The rest of the development team are directly employed by ISIS, or in the case of the Americans, by the SNS.

While the author was primarily situated within the development team based at ISIS, many developers within the team are generally also attached to a specific scientific group at the facility. In the author's case this was the ISIS Molecular Spectroscopy group (MSG) [3] which specialises in using indirect geometry spectrometers to preform quasi-elastic, inelastic, and deep inelastic (also know as Compton) neutron scattering [4, 5] for condensed matter science and thus much of the work done over the course of the year was carried out in relation to the needs and requirements of the MSG.

3 Technical and Application Environments

The development team based in the UK consisted of a single office located within the main office building of the ISIS facility. This office consisted of approximately 13 workstations for use by the team. The number of machines varied throughout the course of the year depending on the level of staffing available for the project. Each of the machines were reasonably powerful 64-bit Dell workstations with between 8-16 Gb of RAM and 8-16 core intel i7 processors. Typical hard drive space for the machines was between 512 Gb to 1 Tb with the majority still being disk drives, but some of the newer machines had flash based storage.

The operating system that each machine ran was completely left to the preference of the developer, but it was recommended that the developer run one of the operating systems supported by Mantid for obvious reasons. In practice this meant that there was a good variety of developers using different platforms. The author chose to run Ubuntu 12.13 as his operating system of choice for the majority of development work, with a dual parititon running Windows 7 which could be swapped to when circumstances required. Other operating systems used by developers in the team included Windows 8, Mac OSX Mountain Lion and Mavericks, Red Hat Enterprise Linux 6, and Fedora 20.

Apart from the workstations, the development also had a collection of Jenkins build servers in order to support a continuous integration and testing workflow in conjunction with the Gitflow workflow [6]. The build servers and jointly located at both ISIS and the SNS. At the start of the placement, the build servers for ISIS and the SNS were completely separate and located at different web address. Each individual build was run as a single job on the Jenkins build servers. At the beginning of the current year this was changed so that the servers were located at the same web address and the organisation of the build servers were changed to make use of matrix builds. This is where multiple builds are are kicked off at the same time under a single umbrella job. For example the development branch matrix build would build the project and run the unit test on each officially supported OS.

Like the choice of operating system, the development software used by the team was flexible and open to developer preference. The project is built using the CMake build system on all supported platforms. On windows platforms the only supported compiler was Visual Studio 2012 or 2014 and most developers

either chose to use the Visual Studio IDE, the Qt Creator IDE, or the Eclipse IDE. On Mac the Intel C++ compiler is used and typical IDEs are XCode, Eclipse, or Qt Creator. On Linux distributions the GNU compiler is the main supported compiler, with either Eclipse or Qt Creator used as the IDE for development. Many Linux developers are also content to just use the make command to build the project from the command line. This approach is often used in conjunction with lightweight editors such as vim or sublime text.

Unit tests are optionally built along side the project using a separate build target generated by CMake using the CxxTest unit testing framework. System tests are written in python and make use of a collection of custom scripts loosely based on the unittest python module and makes use of the Mantid applications python API. Debugging software used typically makes use of Visual Studio on Windows, XCode on Mac and GDB on Linux distributions.

The Mantid application make use of data files produced directly from neutron spectrometers. These files are collected on a collection of servers based by ISIS and owned by the scientific computing department and provide both the instrument scientists and visiting scientists with access to the their data. The development team also has direct access to all of the data generated from the instruments. This are available through nextwork drives. On Windows operating systems access is provided though the in-built network drive capabilities. On Mac and Linux access is obtained through using Samba software in conjunction with the SMB protocol. Copies of actual instrument data are frequently used as part of test scripts, especially in the case where the data required for the test cannot be easily simulated programatically.

4 Description of Job Role and Work Done

As mentioned in section Introduction, the majority of my year was spent attached to the Molecular Spectroscopy Group (MSG) at ISIS. My role in the development team was to satisfy the data analysis requirements of the MSG within Mantid. This included involvement in every part of the development cycle, from gathering requirements from the users (the instrument scientists in the MSG) through to implementation of requested features, to testing and maintenece/bug fixing.

5 Brief Introduction to Neutron Scattering

In order to fully understand the work done, a small amount of background knowledge of the techniques used in neutron scattering experiments are required along with some definitions of key terms. Broadly speaking, neutron scattering can be split into two categories: elastic and inelastic. elastic scattering is where the final energy of a scattered neutron is equal to the energy of the incident neutron, i.e. there is no transfer of energy to or from the sample. Inelastic scattering (which is what is principally measured by instruments belonging to the MSG) is the more complex case where the energy of the incident neutron and the scattered neutron are not equal, i.e. there is a transfer of energy to or from the sample. From this transfer in energy and from known parameters of the instrument an instrument independent scattering function can be defined usually denoted as $S(Q, \omega)$, where Q is the momentum transfer and ω is energy transfer.

6 Critical Evaluation of Placement

References

1. Source. IN. 2014. How iSIS works

2. Project M. 2013. Mantid: Manipulation and analysis toolkit for instrument data

3. Source IN. 2014. Molecular spectroscopy group

4. Sivia DS. 2011. *Elementary scattering theory for x-ray and neutron users*. OUP Oxford. pp. ed.

5. Fernandez-Alonso F, Price DL. 2013. *Neutron scattering fundermentals*. Academic Press. pp. 1st ed.

6. Project M. 2014. Git workflow