March 12th 2017

Bastille – Phase 1: Mantid Project plan

Miguel A. González, Computing for Science

This document is the third version of the document describing the project organisation for Bastille Phase 1. The previous version was written by Mark Johnson in March 2016, when he was both the project leader and the leader of the Computing for Science (CS) group, and contained a joined vision describing the project and also the status and possible evolution of the CS group. Taking into account the changes during last year – the project leader is now Miguel A. González and the CS is joining the SCI group under the direction of Paolo Mutti – it has been decided to revise thoroughly the project plan in order to decouple the project work and the remaining CS tasks. Additionally the document now focuses only on the activities for which extra resources have been demanded, so the references to NSXTool and Lamp have been removed. Nevertheless there is no major change in the scope of the project. The main work remains to implement Mantid at ILL on approximately 20 instruments and a live data analysis server for linking data collection with live data reduction and the project extends over three years (May 2016 – April 2019). This updated document also introduces some new work concerning the Mantid framework and a revised time planning that takes into account the experience gained during the first 10 months of the project. The current status of the project is briefly summarized in the Milestones section. Additionally, in order to account for the changes in the project plan described above a new risk has been added concerning the maintenance of Mantid once the project ends. Supporting documents written before the project start can be found on the ILL intranet* and all the subsequent related project documents are on GitHub†.

Document versions:

Version	Date	Author	Comments
1	8 November 2015	Mark Johnson	Original version
2	25 March 2016	Mark Johnson	Risk register update
3	12 March 2017	Miguel A. González	New project structure

Recipients: Project team, CS, Endurance review panel, Charles Dewhurst

	Written	Verified	Approved
Name	M. A. González	P. Mutti, I. Bush	PMC
Signature			

* http://intranet.ill.eu/divisions/science-ds/endurance-programme/bastille/
† https://github.com/mantidproject/documents/tree/master/Project-Management/ILL

Contents

GENERAL / CONTEXT

Data reduction and analysis is a key component in the production of scientific results. The current situation for data reduction and analysis could be regarded as adequate, but several areas for improvement had already been noted in a user survey performed in 2009 and it has been increasingly recognized the necessity to renew and modernise the data treatment software at the ILL. Further consideration of the limited possibilities for extending the only general data reduction package used at the ILL (LAMP) due both to the age of the code (>20 years) and the retirement in 2017 of its main author and maintainer (Didier Richard) lead to propose the BASTILLE (Better Analysis Software to Treat ILL Experiments) project. Its main goal is to provide a coherent approach to data reduction and analysis at ILL, developing a common framework maintained by a team of developers and therefore not relying in a single person. Taking into account that the same vision is shared by other neutron facilities and also the benefits of collaborating inside the SINE 2020 project, it was finally decided to join the Mantid initiative shared by ISIS, SNS, and ESS. Thus the main objective of BASTILLE Phase 1 consists in the implementation of Mantid in all the relevant ILL instruments.

FUNCTIONAL DESCRIPTION

Overview

The Phase 1 of Bastille is focused into Mantid, which by the end of the project will be fully adapted to the relevant ILL instruments and will accomplish the following purposes:

1. Mantid will allow to visualize, reduce and in some cases analyse the data acquired in the following instruments[‡]:

> **Backscattering:** IN16B, (IN13)
>
> **Time-of-flight spectrometers:** IN4C, IN5B, IN6, (BRISP), (Panther), (Ramses)
>
> **SANS:** D11, D22, D33
>
> **Reflectometers**: D17, FIGARO, (SuperADAM)
>
> **Powder diffractometers:** D1B, D2B, D20, (XtremeD)
>
> **Liquid and amorphous diffractometers:** D4
>
> **Polarized diffractometers and spectrometer:** D7
>
> **Others:** LAGRANGE, D16, SALSA, (D50)

2. Mantid will be able to deal both with the standard data and the event mode data that could be produced by all the instruments above. The latter are now used occasionally in some instruments and it is expected that their use will increase in the future, at least for certain types of experiments.

3. The interaction between Mantid and Nomad will be explored, in order to provide a live data analysis tool. In this context, methods to perform automatically the data reduction will be developed.[§]

Details

The specific work to be done in implementing Mantid on the ILL instruments can be broken down into different types:

Data loaders – Mantid requires loaders for each technique and specific Instrument Definition Files (IDF) for each instrument. Those describe the instrument/detector geometry and other meta-data required by the data reduction algorithms. Almost all of the above-listed instruments produce one data set/numor per instrument configuration (e.g. detector position). The instrument configuration can change between numors, in which case a parameterised, base IDF is used to create the corresponding IDF on-the-fly using metadata from the data file. The two exceptions are D2b (one numor contains data from 25 detector positions) and D7 (one numor contains data corresponding to several polarisation states). The project must provide also tests for loaders for SCI to ensure that any modifications to data files that they make do not affect the subsequent data reduction.

[‡] Instruments in parentheses will be considered only by the end of the project, depending on the status at that moment of both the BASTILLE project and the CRG or instrument project.

[§] It should be noted that a fully working live data reduction environment would probably require additional efforts needing a phase 2 of BASTILLE.

Data format – Mantid will read NEXus/HDF files produced by the present instruments' configurations, as well as future files reflecting instrument changes or upgrades. Close collaboration with SCI is in place in order to ensure that NeXus/HDF data are produced as required, and that changes in the file format do not affect the Mantid loaders. However in the current project is not planned that Mantid reads the old ASCII format or past instrument data.

Data types – Data files normally contain a *single* data set. In *kinetic* (time-resolved) measurements, it is common to store all the time-slices of data in a single file. Treating such a kinetic data file is a simple loop over the treatment of data files containing a single data set. *Event mode* data storage (every detected neutron is stored individually) is beginning to be occasionally used for measurements in which a non-periodic, external perturbation is applied. Recent developments in LAMP allow this data to be time-sliced in an optimal way during data reduction. IN5 should use event mode data storage for single crystal measurements to allow continuous sample rotation and continuous streaming of data which would allow the 4D scattering function to be constructed in real-time and avoid the significant time actually lost in writing data files from short measurements corresponding to specific crystal orientations. At present, event mode data have been handled either directly by the users who have developed their own tools (cases in D17 and Figaro) or in an *ad hoc* manner in Lamp for several experiments performed in D22 and SALSA. Mantid will provide a general solution for all the instruments above, allowing to manipulate and bin the stored events in the desired way and then performing the data reduction on them.

Algorithms – These are required for instrument calibration (e.g. detector calibration) and data reduction. In most cases, algorithms exist but they need to be understood, implemented and tested with ILL data and when necessary modified to meet ILL needs. Furthermore some new algorithms will have to be developed for some types of instruments or methods which do not exist yet in Mantid. They will be identified along the project, but current known examples include the development of the algorithms to handle elastic and inelastic fixed window scans in IN16 or the coherent method recently implemented for some special cases in D17 and Figaro.

Data reduction scripts – Python scripts allow sequences of algorithms to be applied for complete data reduction. They can be standard scripts or customised by users. Using a Python template, they can be presented in simple interfaces. Efficient scripts to produce fully reduced data for all the instruments above will be provided. Nomad will be able to use these scripts to perform the live analysis.

Simple interfaces and Graphical User Interfaces (GUI's) – High-level GUI's are essential for efficient data reduction on high throughput instruments and for occasional users, and require a significant development and support effort compared to the underlying algorithms. It is planned to develop GUI's for TOF-BS, SANS and reflectometry techniques, and it is likely that additional interfaces for the other techniques/instruments cited above will also be requested. Such GUI's exist in Mantid and they will be used as a starting point, but it is anticipated that they will have to be adapted or fully redesigned in order to respond to the needs and habits of ILL scientists and users.

Framework – The Mantid framework is well-established and in the first versions of the document the only expected work consisted in doing some adaptations to this framework in order to handle the scanning instruments (D2B, D4, D16, D7). However in the last months two relatively large projects related to the Mantid framework (Mantid 4 and Instrument 2.0) have been put forward. The main effort in these projects will be done by the other Mantid partners (ISIS and SNS for Mantid 4 and ESS for Instrument 2.0), but as a member of the collaboration the ILL is also expected to contribute to them. The nature of the projects and the expected benefits are briefly described here.

- ***Mantid* 4:** The main focus of this work is to create a new graphical user interface for Mantid, including improvements to plotting, consistency in algorithm naming, and new interfaces in the Python API. Migrating completely to Python 3 is also considered. The plotting has already been identified by some ILL scientists as an area in need of improvement, so this will be a focus of the ILL contribution. This project is foreseen to start on May 2017 and to last 1 year. The effort requested is 5 PY, which will be shared by ISIS (2 PY), SNS (2 PY), ESS (0.5 PY) and ILL (0.5 PY).

- ***Instrument 2.0:*** The current instrument geometry matched the initial requirements for Mantid at ISIS, but the increase of the performance demands over the years, the increasing complexity of the experimental setup, and the complexity of the current syntax for describing instruments ask for a new design. This is a critical point for the ESS, as the anticipated demands for live-reduction there could be hard to satisfy within the current framework. The new geometry design gives more flexibility, and the design includes support for scanning instruments. This part of the work has been implemented early to support the scanning instrument requirements at the ILL. The work on the changes to the instrument geometry started on January 2017 and the ILL contribution (0.5 PY) will be finished by June 2017. The remaining effort (2 PY) is provided by the ESS.

PROJECT DESCRIPTION

Project team

Name	Hierarchical responsible	Division/Service	Responsibility
Miguel A. González	P. Mutti	DS/CS	Scientific project leader
Ian Bush	No ILL	---	Technical project leader
Verena Reimund	P. Mutti	DS/CS	Software developer
Antti Soininen	P. Mutti	DS/CS	Software developer
Gagik Vardanyan	P. Mutti	DS/CS	Software developer

Roles:

1. Scientific leader

Must be competent in the scientific aspects of the code. He/she will

- Help *develop* scientific algorithms and scripts in the framework provided by the technical project leader.
- Ensure the effective interaction with the scientist clients to deliver software with the best performance and ergonomy.
- Work closely with the technical project leader in order to ensure that the necessary developments are undertaken.
- Coordinate *support* for new software as it is deployed.

2. Technical leader

Must be technically competent in all aspects of the code (languages – C++ and Python - and tools) and have a thorough understanding of the structure and details of the whole project. He/she will

- Have a software engineering background and be experienced in software development.
- Lead the software development team of 3 people.
- Contribute also as software developer.

During the first two years of the project, this role will be subcontracted to Tessella, in order to profit from the experience of this company in the development of Mantid. The specific tasks requested in the contract signed with Tessella are:

- Train and organize the team of developers, helping them to identify the key members of the Mantid international collaboration and the standard working practice and methods to contribute to it.
- Provide a solution to treat simple scanning instruments in Mantid. It has been decided that this should be part of a wider collaboration with ESS and ISIS to ensure that an efficient long term solution is implemented.

- Setup and run a build server to ensure compatibility with the operating systems in use at the ILL.
- Undertake the required work to ensure that Mantid's framework supports specific ILL requirements, such as handling the merging or runs, the use of BLOSC compression in NeXus files, and the ILL format for event-mode files.
- Write annual reports describing the status of the project.
- Organize user training at the ILL to facilitate the adoption of Mantid by ILL scientists and user.
- Perform a dedicated technical training to the future technical project leader.

3. Software developers

A team of 3 developers will be formed and led by the Technical Project Leader. They will proficient in the required software languages (Python and C++) and after the appropriate training will

- Understand the Mantid framework and the development methods and tools required to contribute to the project.
- Develop the loaders, algorithms, scripts and interfaces required for each technique or instrument.
- Write appropriate technical and user documentation.

Scientific advisors

Name	Hierarchical responsible	Division/Service	Role
Björn Fak	B. Farago	DS/Spectroscopy	TOF advisor
Tilo Seydel	B. Farago	DS/Spectroscopy	Backscattering advisor
Thomas Saerbeck	G. Fragneto	DS/LSS	Reflectivity advisor
Dirk Honecker	G. Fragneto	DS/LSS	SANS advisor
Thomas Hansen	T. Fernández-Díaz	DS/DIF	Powder diffraction advisor
Gabriel Cuello	T. Fernández-Díaz	DS/DIF	Powder and liquid diff.
Andrew Wildes	T. Fernández-Díaz	DS/DIF	D7 advisor
Thilo Pirling	T. Fernández-Díaz	DS/DIF	Salsa advisor
Mónica Jiménez-Ruiz	B. Farago	DS/Spectroscopy	Lagrange advisor
Viviana Cristiglio	G. Fragneto	DS/LSS	D16 advisor

On behalf of all scientists and users, the scientific advisors will

- Define the requirements for the software in terms of functionality, performance and ergonomy.
- Accept software developed or give advice on the necessary improvements.

SCI

Regular contacts and collaboration with the SCI group and specially with the Nomad team are foreseen. Since February 2017, Paolo Mutti, Franck Cecillon, and Yannick Le Goc attend regularly the weekly Mantid meetings and the initial contacts to discuss the definition and implementation of the NeXus files for diffraction instruments have evolved into a close collaboration. The cooperation between the Mantid and Nomad teams will increase along the project and during the last phase the connection will be further reinforced in order to complete the tasks related to the event mode files and the automatic and live data reduction.

Workpackages

1. Framework

These update of the project plan includes two new tasks related to the Mantid framework which were not present in the original one, as the decision to launch both projects was taken by the Mantid steering committee only last year. They will be carried out jointly by all the Mantid partners and the ILL contribution is part of the engagement in the Mantid collaboration. In particular the Instrument 2.0 project is foreseen to facilitate largely the implementation of scanning instruments into Mantid.

1.1 Mantid version 4.0

This is a major revision of Mantid that will include a number of desired changes breaking backward compatibility in several aspects. There will be major changes to the user interface and the plotting tools will be improved, changing to use Matplotlib as plotting library. It is also foreseen to complete the migration from Python 2 to Python 3. The existing algorithms will be revised and sanitised, removing unused or deprecated algorithms and adopting consisting naming rules. Finally a new Python API will replace the old one. This project is foreseen to start on May 2017 and be completed within 1 year. The effort requested is 5 PY, which will be shared by ISIS (2 PY), SNS (2 PY), ESS (0.5 PY) and ILL (0.5 PY). The ILL contribution will be provided initially by the technical project leader, who will participate in the design of the workbench. In a second step, one of the developers will contribute to the 1D and 2D plotting, where some improvements have already requested by ILL scientists.

1.2 Instrument 2.0 and scanning instruments

This is a full redefinition of the instrument geometry scheme implemented in Mantid, originally started to response to the foreseen ESS demands for live-reduction of very large data sets. A new geometry design giving more flexibility is being worked out. The latter provides a much more convenient framework to implement scanning instruments by decoupling the detector geometry from the detector information (position and time stamp). The technical project leader has participated in the new design and he will use it to implement the simple ILL scanning instruments (D2B, D4, D7, D16).

2. Techniques / Instruments

The following workpackages describe the work foreseen to implement Mantid for the 17 instruments listed above. The inclusion of the additional 7 instruments appearing in parenthesis in the list will be considered along the project, depending on the status of the different CRG or projects. In most cases, this implementation should be relatively straightforward, as most of them have the same characteristics of some of the first 17 instruments.

2.1 Backscattering

Mantid will treat the three types of acquisition modes available in IN16B (quasielastic spectra and elastic and inelastic fixed window scans), providing both a user interface and a scriptable workflow. Furthermore a common 'spreadsheet'-like interface for TOF and BS, allowing to handle efficiently large data sets, will be developed.

2.2 TOF

Mantid will be able to reduce data from IN4, IN5, and IN6. As for BS, appropriate interfaces and scripts facilitating the passage from raw data to physical quantities, notably $S(Q,\omega)$ and $G(\omega)$ will be provided.

2.3 Powder diffraction

Mantid will handle D1B, D20, and D2B data. D1B and D20 will be completed first, while D2B will be the first scanning instrument to be implemented once the needed work in the Instrument 2.0 geometry is completed. It will be possible to load multiple files, calibrate them, plot them as a function of a given parameter (e.g. thermodiffractograms), and export the data to multiple Rietveld programs. It will also deal with all kind of standard scans (e.g. texture measurements).

2.4 Reflectometry

The initial effort will be concentrated in TOF reflectometry in both D17 and Figaro. Monochromatic reflectometry could be considered in a second stage. An easy to use 'spreadsheet'-like interface facilitating the treatment of large data sets will be delivered. Additionally during the project we will explore the possibility of creating automatic procedures to increase the efficiency.

2.5 SANS

A common interface similar to the existing ones (LAMP, GRASP) which are highly appreciated by the ILL scientists and users will be created to treat D11, D22 and D33 (in both standard and TOF modes) data. Automatic procedures will also be explored.

2.6 D4

This is a scanning instrument. Data can be acquired using very different schemes and need to be regrouped before applying a series of specific corrections (calibration, background subtraction, inelastic corrections) giving finally S(Q) in absolute units. Typically the latter is then Fourier transformed to produce a radial distribution function in real space. Mantid should be able to perform all the required operations.

2.7 D7

This is a scanning instrument. It can be used as a polarized diffractometer and spectrometer. Correction procedures are similar to those employed in both techniques, but additional steps are needed to properly handle the polarization. Implementation in Mantid will be done in close collaboration with the instrument responsible.

2.8 Lagrange

The specific needs for this vibrational spectrometer will be discussed with the instrument responsible. No particular difficulties are anticipated.

2.9 Salsa

The specific needs for this vibrational spectrometer will be discussed with the instrument responsible. No particular difficulties are anticipated.

2.10 D16

This instrument can work either as a SANS diffractometer, a reflectometer or a wide-angle diffractometer (D4-like). Once Mantid is able to handle the three techniques, the implementation of D16 in Mantid will benefit from the previous work and is expected to be relatively straightforward. Specific necessities will be discussed with the instrument responsible.

3. Event mode data

This acquisition mode can be extremely helpful when performing experiments whose kinetics are not known beforehand, as the user can store the information of each recorded neutron (time and position in detector) and try different binning schemes during the data analysis. It could also help to reduce acquisition times in single crystal measurements in IN5, allowing a continuous sample rotation and continuous data streaming that would allow the 4D scattering function to be constructed in real-time and avoid the significant time actually lost in writing data files from short measurements corresponding to specific crystal orientations. This mode is already regularly used in D22 and other instruments (e.g. Salsa, D17, Figaro) have used it occasionally. It is also used in other neutron facilities, notably SNS, so Mantid can already handle event mode data. But some additional work will be needed to read the event mode format used at ILL and provide the appropriate tools facilitating the exploration by the user of the event data and of different binning alternatives.

4. Live data analysis

Mantid should be able to interact fully with Nomad, allowing data analysis to be performed directly during the data acquisition. This will improve the efficiency by allowing the user to take rapid decisions on the basis of corrected data. Furthermore Nomad could use the output of the Mantid analysis (e.g. the statistical noise of the final result) to control the acquisition workflow.

TOTAL COSTING

The only costs of the project are related to the recruitment of the software development team. The nominal cost for 3 developers for 3 years at 75 k€/person/year is 675 k€. And the contribution from Tessella amounts to about 360 k€ (175 k€ for year 1 and 185 k€ for year 2), making a total cost of 1035 k€ distributed over three years. As the SINE 2020 software work-package provides 150 k€ of funding to complete the task of integrating Mantid in continuous sources, the requested ILL funding is 885 k€.

Spending profile

2016	2017	2018	2019
267 k€	406 k€	287 k€	75 k€

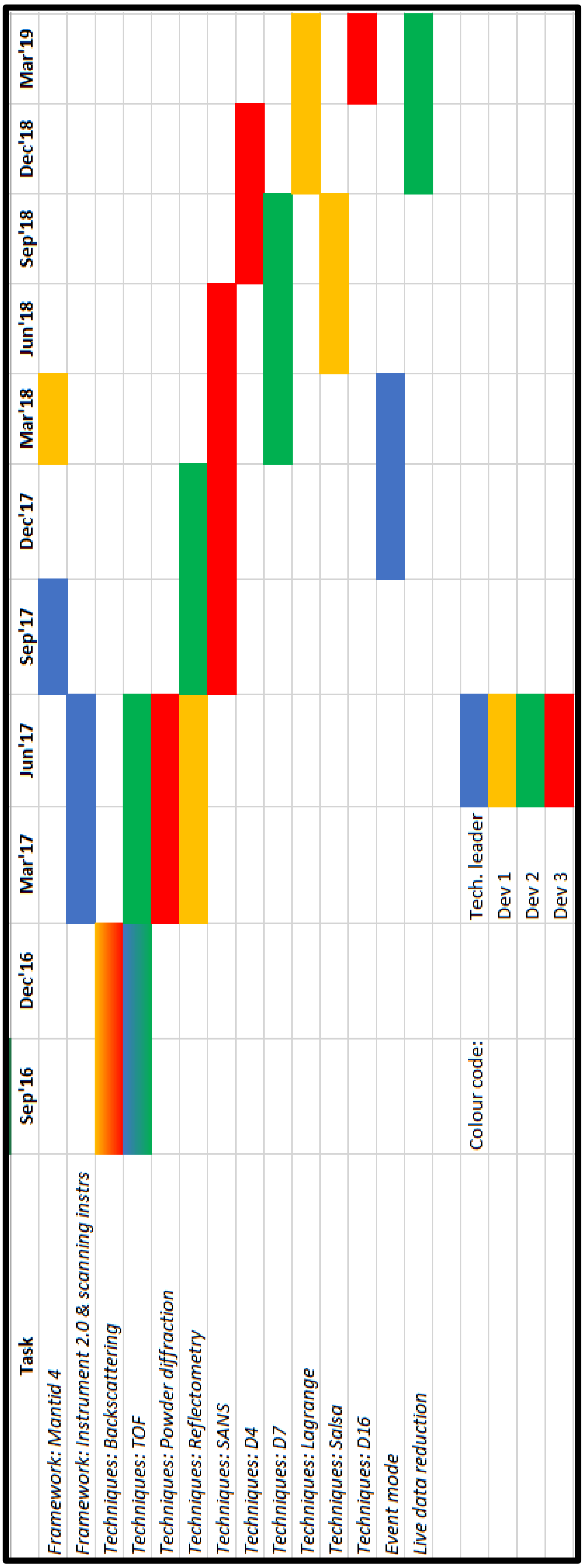
MILESTONES

Date	Milestone
December 2016	Mantid deployed in IN16B. User tests during 1st cycle in 2017.
31 March 2017	Demonstration to TOF scientists. Deploy Mantid in IN4 / IN5 / IN6.
September 2017	Mantid deployed in D1B and D20.
December 2017	Mantid deployed in D2B, D17 and Figaro.
March 2018	ILL event mode data handling.
June 2018	Sans interface ready. Deploy Mantid in D11 / D22 / D33.
December 2018	Mantid deployed in D4, D7 and Salsa.
May 2019	Mantid deployed in Lagrange and D16.
May 2019	Live-data analysis server prototype.

Current status (March 2017)

- Mantid deployed in IN16B and tested during 1st reactor cycle in 2017.
- Work for TOF instruments nearly completed. A demonstration and training session with BS and TOF instrument scientists will be organized on March 31st.
- Common TOF-BS interface to treat multiple data sets under development.
- Instrument Definition Files for D17, Figaro, D1B, D20, D2B, and D4 created.
- Loaders for D17, Figaro and D20 ready.
- Work on Instrument 2.0 and scanning instruments initiated.

PLANNING


Task
Sep'16
Dec'16
Mar'17
Jun'17
Sep'17
Dec'17
Mar'18
Jun'18
Sep'18
Dec'18
Mar'19
Framework: Mantid 4
Framework: Instrument 2.0 & scanning instrs
Techniques: Backscattering
Techniques: TOF
Techniques: Powder diffraction
Techniques: Reflectometry
Techniques: SANS
Techniques: D4
Techniques: D7
Techniques: Lagrange
Techniques: Salsa
Techniques: D16
Event mode
Live data reduction
Colour code:
Tech. leader
Dev 1
Dev 2
Dev 3

RISK ANALYSIS

Mantid is not delivered on time. Risk factor: 15 / 25

Impact: 5. Occurrence: 3.

Mitigation: Recruit and start the project as soon as possible. Use agile (and related) development methods, which are to be implemented with the help of TESSELLA, to ensure that the project remains on schedule. Adapt project milestones and deliverables to known external factors e.g. limited support for LAMP after retirement of Didier Richard.

Mantid is not supported or maintained adequately after the end of the project. Risk factor: 15 / 25

Impact: 5. Occurrence: 3.

Mitigation: Anticipate this risk before 2019 and recruit sufficient permanent staff with the technical experience needed to maintain the complex Mantid framework.

Insufficient interaction with ILL scientists and users. Risk Factor: 12/25

Impact: 4. Occurrence: 3.

Mitigation: Regular meetings with scientists to communicate project progress and demonstrate new software. This is based on agile (and related) development methods which are to be implemented with the help of TESSELLA.

Poor adoption of Mantid at ILL. Risk Factor: 12/25

Impact: 4. Occurrence: 3.

Mitigation: Communicate with scientists and users to ensure that Mantid meets requirements. Reduce support for other, older software. Ensure adequate and appropriate staffing in CS to support Mantid (and other CS activities) in the long term.

Insuficient interaction with Mantid project and team. Risk factor: 8/25

Impact: 4. Occurrence: 2.

Mitigation: Employ Mantid/TESSELLA project person at ILL for Years 1 and 2. Attend annual Mantid developer meetings. Verify and validate that effective communication with Mantid project and team is in place.

Mantid does not satisfy the technical requirements at ILL. Risk Factor: 9/25

Impact: 3. Occurrence: 3.

Mitigation: The MANTID framework must be adapted to handle efficiently simple scanning instruments. Make this a deliverable of the TESSELLA involvement.

Interaction with other ILL groups (SCI and SI). Risk Factor: 6/25

Impact: 3. Occurrence: 2.

Mitigation: Involve SCI and SI in the final stages of project preparation and ensure effective communication during the project.