

# THE MAX IV LABORATORY SCIENTIFIC DATA MANAGEMENT

V. Hardion, A. Barsek, F. Bolmsten, Y. Cerenius, F. Hennies,  
 K. Larsson, D. Spruce, J. Brudvik, Z. Matej  
 MAX IV Laboratory, Lund University, Lund, Sweden

## Abstract

The Scientific Data Management is a key aspect of the IT system of a user research facility like the MAX IV Laboratory. By definition, this system handles data produced by the experimental user of such a facility. It could be perceived as easy as using an external hard drive to store the experimental data to carry back to the home institute for analysis. But on the other hand the "data" can be seen as more than just a file in a directory and the "management" not only a copy operation. Simplicity and a good User Experience vs security/authentication and reliability are among the main challenges of this project along with all the mindset changes. This article will explain all the concepts and the basic roll-out of the system at the MAX IV Laboratory for the first users and the features anticipated in the future (Fig. 1).

## DO RESEARCHERS NEED SUPPORT TO MANAGE THEIR DATA?

In short yes, there is a growing consensus spanning the European commission, research infrastructures, higher education and research groups that services need to be improved for researchers. While being owners of the data, they are not able to dedicate enough focus to data management while the amount of data generated is rapidly growing. Within facilities such as MAX IV Laboratory, the handling of users is split into many groups which handle different parts of the workflow. This results in visiting researchers needing to access many internal tools and systems as they follow the facility experimental visit workflow from proposing to do an experiment, planning a visit, following safety procedures and using the beam line computer systems, data storage and computation and finally obtaining data and results. Without any internal integration of these systems, the user is normally obliged to carry information (login details, experimental context and history of the events) between different systems, in their heads, on paper or in their own electronic log formats. They must try to keep a track of the relevant information needed at each step and be able to take what is needed away from the facility. The Scientific Data Management (SDM) is designed to create links across all systems which permits each component to dialogue information without the researcher having to do it. The result is that each system can implement a smoother, less detailed workflow, increasing effectiveness while more easily allowing more complexity to be added, such as meta-data acquisition during the experiment, automated data processing workflows and even automated migration to destined remote data centres. It is important that the SDM workflows appear simple and reliable to build trust in using them, instead of the current

trust in using portable physical disks as a way of transferring data which makes it virtually untrackable once it leaves the facility. In addition, the new high data volume experiments are too much data for this method to be feasible.

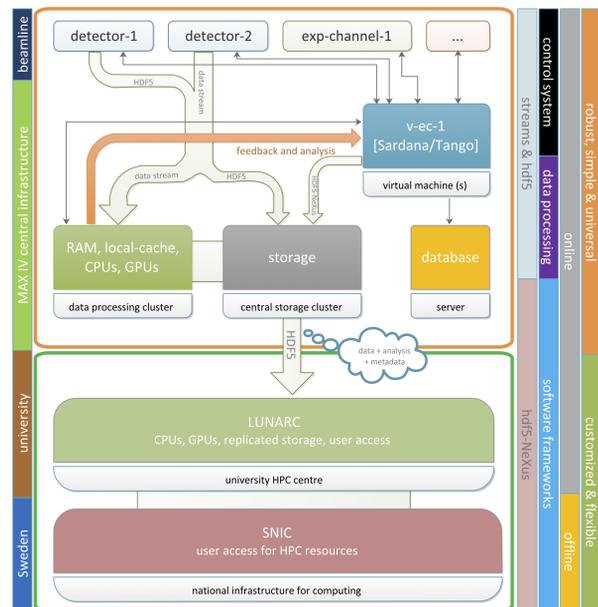


Figure 1: The MAX IV imaging concept. [1]

## SCIENTIFIC DATA MANAGEMENT

### History

The project started in 2013 with a White Paper trying to describe the need for managing the data at the new MAX IV Laboratory. This was distributed for feedback, both internally and at the MAX-lab 2013 user meeting. In order to address the resource issue, including both hardware and personnel, a collaboration was initiated with Lund University Centre for Scientific Computing (LUNARC [2]) after a meeting with representatives from the Swedish National Infrastructure for Computing (SNIC [3]).

To verify the content of the White Paper, a prototype project was launched in early 2014 and was finished by summer the same year. The prototype tried to address all aspects and provide more information for the continuation of the project. In particular, it provided valuable insight in terms of technical requirements, staffing needs and hardware costs. The prototype report has been published in 2016 [4]. In late summer 2014, MAX IV was asked by SNIC (on behalf of the the Swedish Research Council (VR)) to provide a report on our future e-infrastructre needs. Using the information from the prototype project and with new estimates from the

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2017). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

beamline projects, a report was submitted and published by VR mid 2015 [5].

One topic that has influenced the project is the Data Policy. This document governs the Scientific Data Management project in terms of what services to provide and how they shall be used. The first draft policy was made during the prototyping project and was then presented to the MAX IV Group Managers, Directors and Board. A workshop with MAX IV management was held in January 2015 finalising the draft policy taking into account the MAX IV budget and the Open Access guidelines from VR. This draft policy was included in a MAX IV Data Plan and submitted to VR in September 2015 for feedback.

### *Data Policy*

The MAX IV Data Policy [6] defines the responsibility of MAX IV and its users pertaining to the data collected at MAX IV. In terms of ownership, it is the Principal Investigator (PI) of the experiment who owns the data and is responsible for long term archiving and Open Access. MAX IV will retain the data on its own storage for 3 months, to facilitate the safe transfer of data to the PI's home institute. MAX IV will provide the means and expertise to assess and reduce the collected data as a part of the user service. The MAX IV Data Policy will comply with Swedish intellectual property rights and try to contribute to the challenge of Open Data Access. An important aspect is capturing meta-data connected to an experiment and making it available to the PI.

If any MAX IV Laboratory data policy is implemented, there will be no clear ownership and responsibility for data collected at MAX IV. The effects of not funding of the MAX IV data plan means that there can be no preparation for long term archiving and open access. It would be much more difficult and costly to do this at a later stage while in the mean time various user groups may have spent part of their funding money on their own local solutions. Furthermore, the new beam lines will not have access to the required e-infrastructure and will not be able to collect data where the full potential of MAX IV is used.

### *Scope*

The Scientific Data Management is a key component of MAX IV Laboratory. The scope is to establish a common infrastructure for collecting, accessing and analyse data collected at the MAX IV beamlines, including:

- Hardware for storing data (storage cluster).
- Hardware for analysis (computing cluster).
- Analysis environment - a desktop solution running on top of the hardware. It has remote access and an environment for installing and manage all software.
- Central user authentication and authorisation service, to be used by all services including the ones out of scope (see below).
- Data access (e.g. SFTP, Web portal, etc).

The initial budget aims only for data generation although an important part of the complete solution concerns the all lifecycle, in particular:

- The User Office - not included in the initial scope however has to be integrated in the SDM solution since it will be responsible for the user account management.
- Scientific analysis software - the SDM project will only deliver a platform where analysis software can be installed, it does not include the actual software.
- Data acquisition software - may be different in every beamline and the SDM project will provide the necessary connections so that the data acquisition software can acquire and store data at the right place and with the correct permissions.

### *Project Management*

The implementation of the SDM is under the responsibility of the all group of Control and IT infrastructure (KITS). The main stakeholders are the User Office and the beamlines staff. However collaboration with other stakeholders is important, this includes user office, scientific division, experimental stations, universities, LUNARC, SNIC, SUNET, respective non-Swedish institutions providing e-infrastructure for experimental stations funded in MAX IV, user community and other light sources.

### *Second MAX IV Beamline Prototype*

This project started in November 2015 and the objective was to create a generic beamline setup integrated with all the services within the SDM project. Team members from Controls, IT infrastructure and Information Management have been participating in this. The beamline prototype helped us to identify the basic components and design an architecture for the MAX IV SDM.

## **ARCHITECTURE**

With only the focus on the data generation we can identify three use cases to define the architecture:

- The proposal users can execute a data collection from the beamline.
- The data are stored in the proposal storage.
- The proposal user can retrieve the proposal data.

The relationship between the data and the experimenter is the main recurrent "theme" thus the importance to keep track of the "participant" of the experiment during the whole beamtime. It covers technically the proposal approval process, the experimentation in the beam lines, the data collection and on-line data processing and the data portal. As the source of the accepted proposal is given by the Digital User Office (DUO), it is full part of this architecture.

### *Directory*

The user directory is the principal component of the architecture in order to have a single point of truth in the system regarding authentication and authorization. Microsoft Active Directory have been chosen mainly to keep the compatibility with the Microsoft Windows OS, although Linux represents more than 90% of the Control System infrastructure, in order to be compatible with the detector coming with Windows

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2017). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

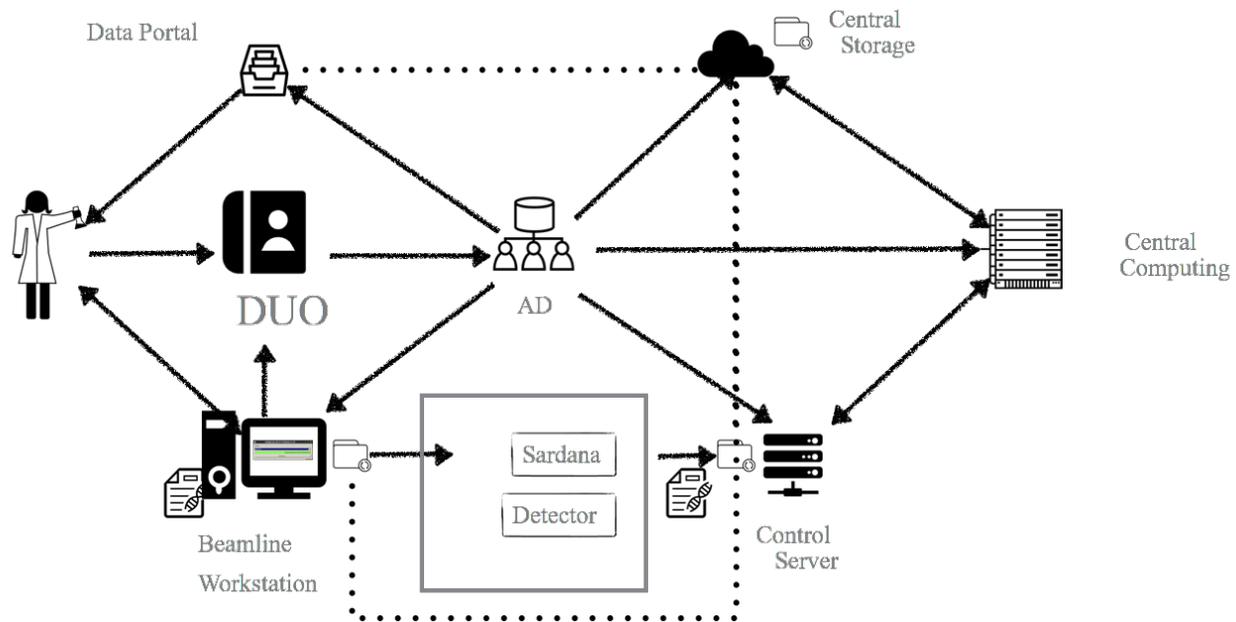


Figure 2: The architecture of the SDM system showing all components involved. The arrows give the consumer of the information data

controller or control/analysis software. Fortunately it comes with an LDAP implementation as the main communication channel for the SDM component (*software and OS*). The Fig. 2 make visible the centrality of the directory in the system.

### User Office

The User Office software is the main source of information regarding the users of the facility as well as the proposals and the visits (beam shift associated with users in this system). The choice of the DUO program made by PSI has been approved after a test period while the old/previous MAX-lab facility was operating. The SOAP API coming by default has been replaced by a simple REST API allowing to access all the structure of the database. This development was done during the process of adaption to the MAX IV user management.

A cron job crawling the DUO database runs continuously to supply the Active Directory with groups and users associated with the accepted proposals. In order to increase the security of the system this job runs only against an intranet version of DUO which by the way allows only the MAX IV staff to accept the proposals.

### Central Storage

Our current strategy of data management is "file oriented" thus the central storage is a critical component. Not only it has the responsibility to store the data in a secured and reliable way but it has to distribute the data between all the producers and consumers of the system e.g between the detector controller and the compute cluster. A first central storage system, so called "Bulk storage" has been installed for the first beamlines and was operational as a main central

storage in 2016–2017. Its main role consists of the online data storage i.e write/read of data during the beam time while secondary to keep the data for 3 months. This storage was replaced recently by a high performance flash-based IBM Spectrum Scale system. The storage cluster is based on GPFS offering good enough performance for the first experiments in MAX IV with the expert user i.e 1 GB/s (16 GB/s after upgrade) for the native protocol with a capacity of 200 TB. The client connects with NFS, Samba or GPFS. The later has been used to connect a buffer server to the Eiger 16M detector in order to offload the data as soon as possible. Low level protocol like InfiniBand can be used as well within the internal cluster infrastructure.

Another storage cluster is located off-site in the building of LUNARC computing centre that is part of the Lund University. The primary role of this site is to provide data replication and later, secondary for the access to offline computing resource.

### Control System Infrastructure

All computer involved in the beamlines and accelerator operation are connected to the Active Directory in order to use the OS authentication system. The main network storage system is mounted at the same location in order to access the data file. The structure of the directories allows fine authorization per beamlines, proposals, visits and group of users (*see Data Model*). The Linux system, mainly CentOS 7 and 6, is configured to use NSCD/NSLCD accessing Active Directory through LDAP. An attempt has been made to use SSSD. Apart the complexity of handling the configuration in our configuration management [7] the reliability and the compatibility with previous version of CentOS 7 was problematic. On the other hand SSSD would allow the

use of kerberos key for single sign on. By default the data storage are mounted with the NFS protocol on Linux and Samba (server side implementation) on Windows. For more advance usage GPFS can be use to increase the performance with the counter part to include the computer in the GPFS system (BioMAX [8]).

### Control System Software

The main interface for the users is the workstation located in the control room where the user interfaces of the control system is installed. A program allows the user, identified by the login of the OS, to choose their proposal/visit to run their experiments. This action will transmit the necessary information on the control system in order to save the data in the correct folder, accessible only by the group of the current users associated with this proposals/visits. The piece of software so-called "path-fixer" is a standard Tango device which is able to validate the current user and proposal against DUO. Once set correctly it gives to the control system the path where to store the data. One of its feature allows to propagate automatically this path to any Tango device which saves the data on disk e.g. Lima [9], Sardana [10]. Other software can have the same functionality (MxCube [11]) by integrating the SDM library.

The prerequisite for an equipment to integrate the SDM system is to be able to mount NFS or Samba, to connect Active Directory and to run Tango. The system is also compliant with a remote user access required for some experimental stations (e.g. [8]).

### Compute Cluster

The compute cluster is a user single point for experimental data access, visualization, processing and analysis. The system is maintained in a close collaboration with LUNARC and so it is very similar to other Swedish computing centres even though it is incomparable in its capacity, scaled mainly for the on-line data processing. However user can be familiar with slurm workload manager and job scheduler, modular software installation, OpenMPI protocol for multinode CPU applications or GPU accelerated frameworks as arrayfire [12] or bohrium [13]. Commercial software utilizing parallel environment as Matlab is available. ThinLinc remote desktop supports 3D hardware accelerated graphics. The idea is to keep the analysis environment similar to SNIC sites that academic users can benefit from the ecosystem even after the experiment.

### Scientific Software

Scientific software in large scale X-ray light sources is usually very versatile concerning data access and data formats. However it can still profit from the presence of a proper SDM concept. For the user the most important is the availability of robust and effective tools for experimental data visualization and analysis. These are built on top of the rich and modular software stack in the computing cluster and well maintained software distributions in the beamline workstations. For the developer data formats are crucial.

HDF5 is a MAX IV standard data container with NeXus format convention as an option, preferred where possible.

In the commissioning and initial operation phases an utilization of recent developments from other facilities proved to be very successful. In particular this is valid for the SILX [14] - a library of GUI widgets for scientific data visualization developed in ESRF. From the in-house developments it is worth to mention the HDF5-web-GUI (Fig. 3), based on HDF5 Server REST API, javascript, jquery and popular *plotly* [15] graphical plotting backend. This is very suitable for remote visualization of data saved in the native HDF5 format without any need of dedicated desktop applications installed, mounting of remote file systems or transferring raw data to the workstation. In addition mobile devices are supported. The HDF5-web-GUI is suitable for detector data acquisition units as well as infrastructure data portals. Latter beside all thanks to (optional) user-id and Central Authentication Service integration.

### Data Portal/Access

The Data Portal is the front end for the users to access their data. At MAX IV several possibilities of data access have been implemented. During the beamtime the users have access with no other latency than that present in the data transfer. As a standard the beamline acquisition system has 10–20 Gbit/s inter-connection to the *fast data* network integrated with storage and the compute cluster. The compute cluster, accessible by users via remote desktop, has Infiniband connection to storage while the workstations at the beamlines have 1 Gbit/s links. All the MAX IV resources are accessible remotely via VPN.

Data web portal can profit from the HDF5-web-GUI interface for accessing and visualising data from the central storage, mainly thanks to the data deployment guidelines described in SDM and the integration with Central Authentication Service and the Active Directory.

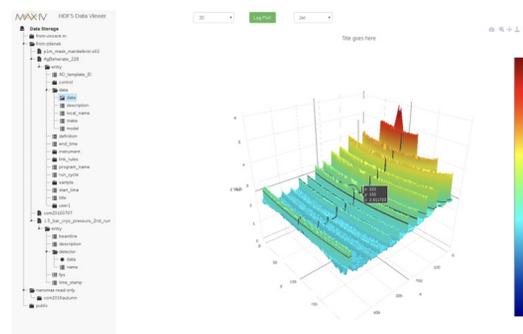


Figure 3: The HDF5 viewer web application can render data sets with different visualizations.

The Globus network [16] is another service proposed for the users. Data can be transferred asynchronously between the source and destination according to the user's order. This makes it easy to transfer data to users home institutes compute, storage clusters, desktops or laptops. This choice has been driven by the user experience from other X-ray light

Content from this work may be used under the terms of the CC BY 3.0 licence (© 2017). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.

facilities, straightforward integration with MAX IV systems, Word-wide spread and popularity of the ecosystem. External access services as Globus again profit from the AD and DUO integration.

### External Stakeholders

The internal MAX IV SDM described mainly here is the basis for collaboration with external stakeholders that can take wheel for the access to storage and computing resources after the user's experiment. This is in the first level the Lund University computing centre (LUNARC), Swedish computing infrastructure (SNIC) and relevant institutions for the other Nordic beam lines (DanMAX, FinEstBEAMS) in MAX IV. The system currently prototyped for the possible exchange of data between MAX IV and them is iRODS - Open Source Data Management Software [17].

## DATA MODEL

All the computers and servers included in the SDM deployment should respect this standard model especially the standard path structure. In case of several mounting points corresponding to different configuration of the file system and protocol used, the mounting configuration should be `"/(protocol)/(fileset)/"`, serving the standard path structure after `"/data"` with a local symlink to `"/data"` from the default mounting point.

Table 1: Group Definition

group	member of	Comment
Visitor		primary group for any visitor
Staff (MAX-Lab)		primary group for any MAX IV staff
<proposal>-group	Visitor	group all visitors associated to the proposal
<beamline>	<proposal>-group	group all beamline staff, automatically associated to the <proposal> group for maintenance purpose

The groups definition (Table 1) are based on these facts:

- Each user category has its own ID range.
- The same user can be part of several proposals groups.
- The *beamline-staff* group is a member of each proposal group.
- The beamline staff people are indirectly part of all proposal groups via their *beamline-staff* group.
- The beamline technical accounts are indirectly part of the proposal group via their *beamline-staff* group.

The user in the directory (Table 2) are based on these definition:

Table 2: User Definition

User	Primary Group	Proposal Group	Proposal Group member
<proposal>	Visitor	<proposal>-group	DUO
<user>	Visitor	<proposal>-group	DUO
<maxiv account>	Staff	no	<beamline>
<beamline>-user	Staff	no	<beamline> can have too much rights
<beamline>-service	Staff	no	<beamline> can have too much rights but limited if no login

- *<proposal>* is a group proposal account, mainly used to share the same session on the beamline computers.
- *<user>* is a visitor account, can be used for all case and to read the data outside of the beamline.
- *<maxiv account>* is a staff account, can be used for own research, commissioning, maintenance of the visitor data.
- *<beamline>-user* is a generic beamline account, mainly used to share the same session on the beamline computers for the staff.
- *<beamline>-service* is a generic beamline account without login, mainly used by the software which create the data files; daemon service not intended to open a session.

Table 3: Ownership and Access

path	/data/type	/beamline	/proposal	/visit/raw
owner	root	beamline-services	beamline-services	beamline-services
group	root	beamline-staff	proposal-group	proposal-group
posix	2755	2775 or 2755	2750	2770
	static	static	dynamic	dynamic

Precise ownership and access structure are described in Table 3. The standard path variable are based on these facts:

- The *type* includes visitor (bread and butter academic users), staff (MAX IV staff) and proprietary (mostly industrial users with a different need for data security).
- The *visit* should begin with the date when the first shift starts. In addition to */raw* there will be a */process* folder at the same level with full access, depending on the beamline preferences.

Information comes from several integrated sources:

- DUO is considered as the primary source for the visitors accounts
- Active Directory is considered as the central point of authentication outside of DUO
- The beamlines groups and beamlines staffs account are statically created

## Security

Full Security is not part of the initial implementation and rely on the IT infrastructure. The Tango control system used at MAX IV does not implement any strong security as it does not convey the authentication to the Tango device which writes the data on disk. But in any case the user can not read data from a proposal whose is not part of thanks to the data model.

## LESSONS LEARNED

The SDM project covers a large spectrum of functionalities, domain and expertise. It becomes apparent late in the project that all parties are strongly linked each other. The system became quickly unusable when one of the component was absent or failing, impacting directly the trust of the users to this system. In a way this is relevant of what the different stakeholders had to do to handles these transition themselves.

The project also suffers from several delays from the begin of project. The decision to go ahead with this project was taken late and with very few resource allocated. Different components were affected by this delay like the central storage installation, making hard to configure and test the system incrementally while the experimental stations were in operation. For example we had to accommodate the storage by swap of "proposal" and "beamline" field of the standard to simplify the creation of the root directories.

Several modification in the model was made during the development.

- Raw directory read only for user: It was agreed that different beam lines can have different preferences on restricting their users. Users may prefer to save their configuration files, additional material, notes etc. together with their data.
- Proprietary user may require to restrict access and members of the proposal-group according to its own requirements on its own risk (including malfunctions in the control system and limitations in local support).
- The generic beam line account broke the foundation of the SDM model since it was not mandatory to login to access the data. It questions the security aspect as well.
- The model of proposals and associated experiments has been refined several times within DUO. Its different implementations were not understood by all the stakeholders and the developers, i.e difference between a "visit" and "shift". It has been compound by the lack of common vocabulary, i.e user visit not registered considered as an expert but expected by the beam line.

- The path chooses by the user may contains more level like sample information which is not handle by the path-fixer. In overall the path fixer is only use to know the root path of the proposal data directory. Hereby we need a good sample management.

It could not even benefit of the maintenance breaks as the focus were on the stabilisation of the basic functionality. Furthermore the login of the user was not always known by the beam line staff, making hard for them to prepare the environment (desktop, profile) in advance.

## CONCLUSION

All the SDM components are currently in place and connected each other following the architecture described above in this article such as the beam lines and the accelerators can use the central storage. Regarding the User management the system has been more or less successful tested with the first call of proposals between March and June 2017 on the NanoMAX beamline [18]. Only at the end of this period the actual data of the users was stored in the correct place, otherwise the fall back was the staff directory. The next development of the SDM will mainly concern the upgrade of the overall system with fine management of the user and more diagnostic in order for the staff to automatically switch to this system.

## REFERENCES

- [1] Z. Matěj *et al.*, "The MAX IV imaging concept", *Advanced Structural and Chemical Imaging*, 2 (2016), <https://doi.org/10.1186/s40679-016-0029-7>
- [2] LUNARC website, <http://www.lunarc.lu.se>
- [3] SNIC website, <http://www.snic.vr.se>
- [4] P. Sjöblom *et al.*, "Motion control system of MAX IV Laboratory soft x-ray beamlines", *AIP Conference Proceedings*, 1741 (2016), 030045, <https://doi.org/10.1063/1.4952868>
- [5] Swedish Research Council (2015), "Survey of e-infrastructure needs for eight Large Infrastructures – Report from SNIC to the Swedish Research Council": [https://publikationer.vr.se/wp-content/uploads/2015/06/VR1536\\_Survey-of-e-infrastructure.pdf](https://publikationer.vr.se/wp-content/uploads/2015/06/VR1536_Survey-of-e-infrastructure.pdf)
- [6] "MAX IV Data Policy": <https://www.maxiv.lu.se/users/user-policies/>
- [7] V. Hardion *et al.*, "Configuration Management of the control system", THPPC013, *ICALEPCS'13*, San Francisco, USA, (2013).
- [8] M.Eguiraun *et al.*, "MAX IV BIOMAX Beamline Control System: From Commissioning Into User Operation", TUMPL08, presented at ICALEPCS 2017, Barcelona, Spain, 2017, this conference.
- [9] A. Homs *et al.*, "The LIMA Project Update", FRCOAAB08, *ICALEPCS'13*, San Francisco, USA, (2013).
- [10] "Sardana control user environment", <http://sardana-controls.org>

- Content from this work may be used under the terms of the CC BY 3.0 licence (© 2017). Any distribution of this work must maintain attribution to the author(s), title of the work, publisher, and DOI.
- [11] U. Mueller *et al.*, “MXCuBE3: A New Era of MX-Beamline Control Begins”, *Synchrotron Radiation News*, 30 (2017), 22-27, <https://10.1080/08940886.2017.1267564>
  - [12] “ArrayFire - a general purpose GPU library”: <https://arrayfire.com>
  - [13] “Bohrium - automatic acceleration of array operations in Python/NumPy, C, and C++”: <http://www.bh107.org/>
  - [14] “silx - ScIentific Library for eXperimentalists”: <https://doi.org/10.5281/zenodo.576042>
  - [15] “Plotly - an online data analytics and visualization tool”: <https://plot.ly/>
  - [16] “Globus - Research data management simplified”: <https://www.globus.org/>
  - [17] “iRODS - Open Source Data Management Software”: <https://irods.org/>
  - [18] P.Bell *et al.*, “Control and data acquisition using TANGO and Sardana at the NanMAX beamline at MAX IV”, TUPHA197, presented at ICALEPCS 2017, Barcelona, Spain, 2017, this conference.