

Data Management Basics at EuXFEL

Services, Storage, Accesses



Luis Maia & Robert Rosca
On behalf of EuXFEL Data Department

Hamburg, 26th January 2024

Agenda

Services

- Overview
- UPEX
- myMdC
- myLog
- DAMNIT
- Storage and Computing
- Maxwell
 - ▶ FastX (Remote Desktop)
 - ▶ JupyterHub
 - ▶ SSH

Remote Access to the Facility

- Experiment
- Data Analysis

Allocating Resources

- SLURM and Jobs
 - ▶ Sbatch
 - ▶ Salloc

- Data
- Software
- EuXFEL software
- Documentation

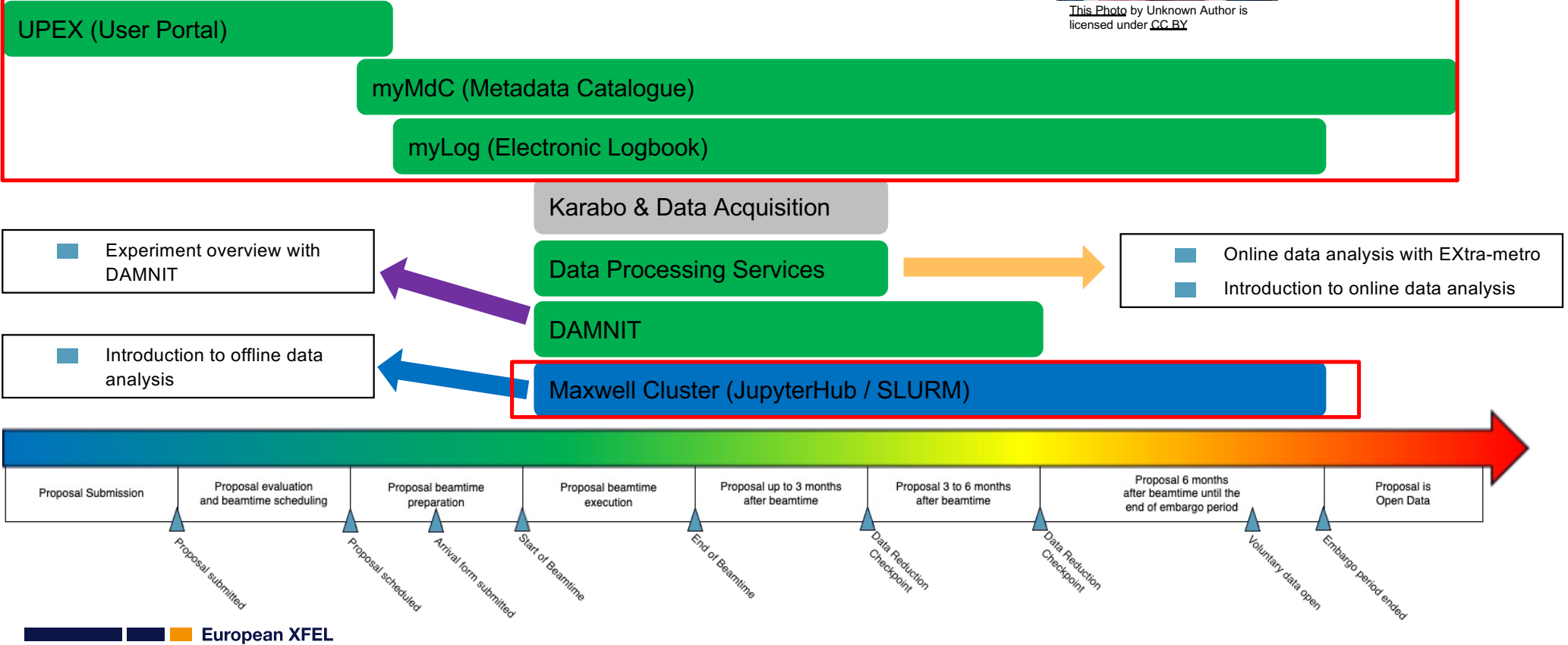
Resources

Overview

Services during Experiment Lifecycle



This Photo by Unknown Author is licensed under [CC BY](#)



Information Entry Points

<https://www.xfel.eu/>

<https://in.xfel.eu/>

European XFEL Services

Link is missing/broken? Contact it: support@xfel.eu

XFEL
xfel.eu | User Portal | Metadata Catalogue | Publications | Guest Portal

DESY
DESY.de | IT services | Library | XFEL Project Group

BeamStop Restaurant
Weekly menu
Catering offer, Catering order sheet, Reservation sheet

Campus
Campus map | Phone book

European XFEL

European XFEL

VISITORS CONTACTS CAREERS UPEX EN DE

FACILITY ORGANIZATION SCIENCE NEWS AND EVENTS **USERS**

Public INTERNAL IT (for IT staff only) Logout

Users

To find out about beamtime and how to prepare for your experiment, please refer to the information in this section.

The European XFEL is a research facility open to scientists worldwide. Beamtime is free of charge, but experiment proposals must go through a peer-review process. Proposals must be submitted through the User Portal to the European XFEL (UPEX) in the frame of specific calls for proposals. As an important condition for access, after the experiments, the authors must strive to publish the results in peer-reviewed journals. More information about our user publication policy can be found here.

The rules concerning the storage and access to the data collected during the experiments follow a specific policy that is comparable to best practices at other European large-scale user facilities. More details can be found here. If you do not find the information you are looking for in these pages, please contact us at usersoffice@xfel.eu.

- User Campus Access
- Beamtime
- User safety
- User guide
- Guest House
- Policies
- User office
- User events
- User Organization
- Industrial users
- Quick Links

Member of EIROforum

Subscribe to our newsletter

Follow us

UPEX Purpose (I)

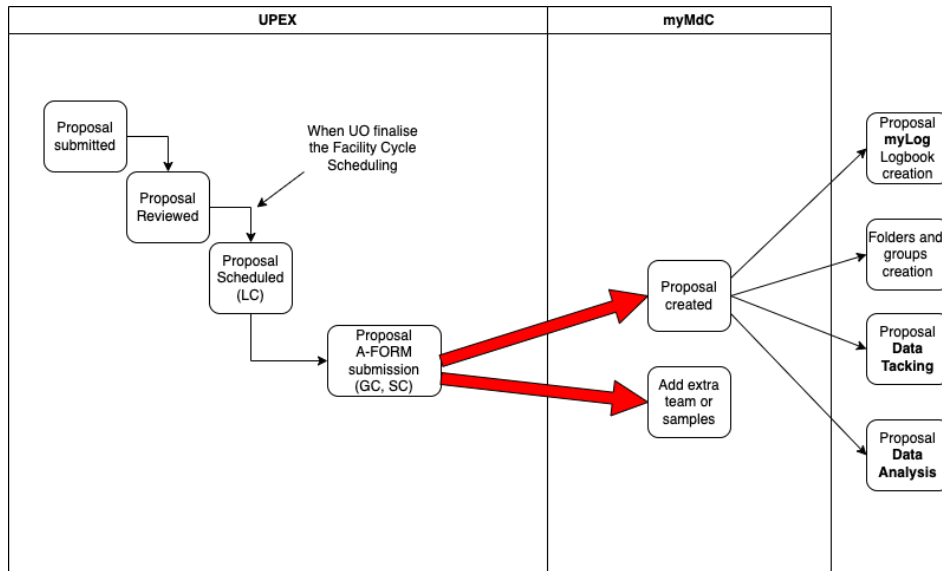
- Proposal submission
 - Proposal submission forms (for users and internal beamtime)
 - Proposal calls/cycles management
- Proposal review and scheduling
 - Feasibility checks
 - PRP reviews and final conclusions
 - Scheduling
- Experiment preparation
 - Experiment team and arrival data
 - Campus accounts setup
 - Export proposals to myMdC
- User management
 - Lightweight and Campus accounts
 - Communication
 - Legal documents

The screenshot shows a web browser window with the URL `in.xfel.eu/upex/auth/login`. The page header includes navigation links for `HOME`, `EUROPEAN XFEL`, and `LOG IN`, along with a timestamp: `UPEX time: 2024-01-20 13:47`. The main heading is `Login to User Portal to the European XFEL`. The login form contains the following fields and options:

- Username:** A text input field containing the text `maial`.
- Password:** A password input field with masked characters `.....`.
- Stay signed in:** A checkbox that is currently unchecked.

Below the form, there is a link to the `UPEX Privacy Policy` and a prominent orange `LOGIN` button. At the bottom of the form area, there are links for `Create new Account`, `Reset your Password`, and `Check your account`. A footer note states: `If you need to contact the User office, please email us at useroffice@xfel.eu`. A yellow banner at the very bottom of the page contains a cookie consent message: `This website uses cookies to ensure functionality. If you continue browsing the site, you are giving implied consent to the use of cookies on this website. See the UPEX Privacy Policy for details.`

UPEX accounts



Check your account status (using username or email): <https://in.xfel.eu/upex/password/check>

LIGHTWEIGHT account



Stored in
UPEX database

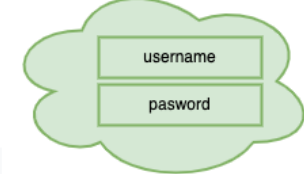
Created by
By a user on one's own

Managed by
UPEX

Access
UPEX only

Expiration
NO

CAMPUS account



Stored in
DESY database

Created by
By XFEL or DESY when required

Managed by
DESY tools:
Registry, IAM, LDAP
By UPEX via API

Access
Many XFEL and DESY services:
UPEX, myMdc, Maxwel, Trainings
.....

Expiration
Password expires regularly
and have to be changed in time
Account expires if a user has no active
proposals
Can be reactivated in future

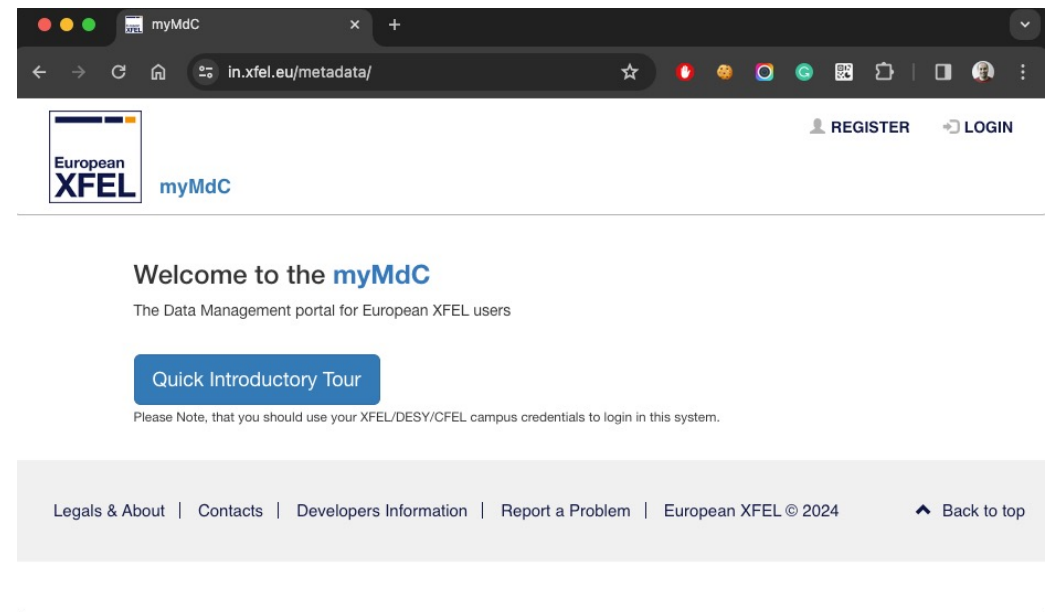
UPEX access

Both credentials are related to the same user record in UPEX

myMdC

- The Experimental Data Management portal and Metadata Catalogue for European XFEL users

■ <https://in.xfel.eu/metadata/>

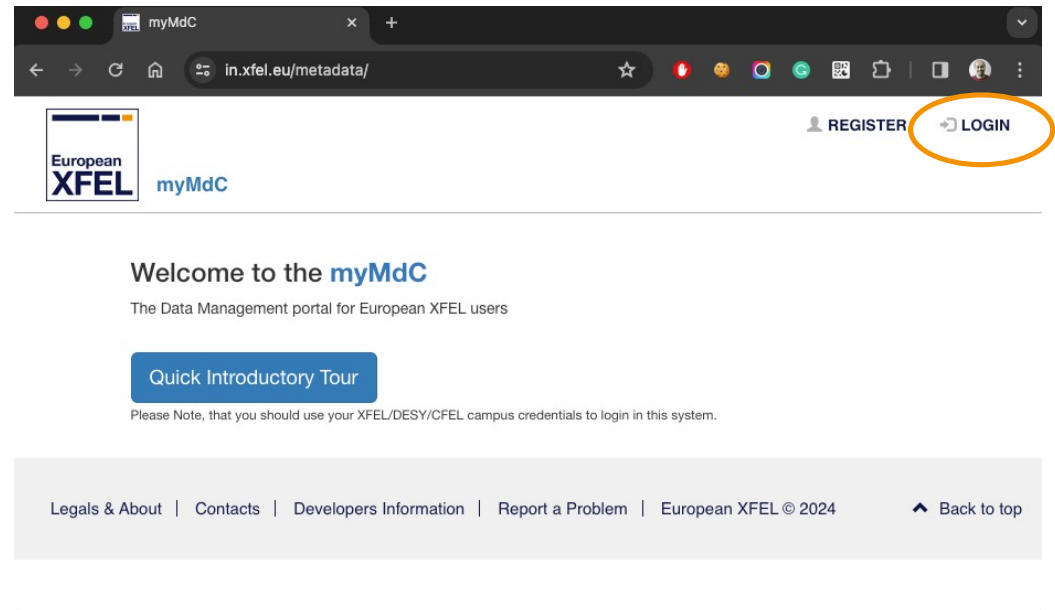


myMdC Purpose (I)

- Provide a means of storing, retrieving and query raw and run based data in an organized way
- Organize and manage data and metadata in a coherent way
- Assist on the execution of the Data Management policies, workflows and notifications
- Manage data and metadata authorization and responsibilities (especially during the Embargo period)
- Manage and restrict the GLOBAL unique entities used during data taking
 - Proposal number
 - Proposal path
 - Sample
 - Run type
 - Techniques

myMdc Introduction (I)

- Visit <https://in.xfel.eu/metadata/>
 - Log in with your EuXFEL user **Campus account**
 - Give the 'Quick Introduction Tour' a read



myMdC Introduction (II)

The screenshot shows the myMdC portal interface. At the top left is the European XFEL logo and the text 'myMdC'. The top navigation bar includes links for HOME, MY TOKENS, MY ACCOUNT, and LOGOUT. The main content area features a welcome message: 'Welcome to the myMdC' and 'The Data Management portal for European XFEL users'. A blue button labeled 'Quick Introductory Tour' is present, along with a note: 'Please Note, that you should use your XFEL/DESY/CFEL campus credentials to login in this system.' A navigation menu is open on the right side, listing 'Proposals', 'Data Source Groups', 'Techniques', 'Developers Information', and 'Legals & About'. The 'Proposals' link is circled in orange, and an orange arrow points from the text 'Go to Proposals' to it. Below the arrow, there are two sub-points: 'View all your proposals' and 'View Open Data proposals'. The footer contains links for 'Legals & About', 'Contacts', 'Developers Information', 'Report a Problem', 'European XFEL © 2023', and a 'Back to top' button.

Go to Proposals

- View all your proposals
- View Open Data proposals













myMdC Introduction (III)

The screenshot shows the myMdC web application interface. At the top, there is a navigation bar with links for HOME, ADMIN, USERS, MY TOKENS, MY ACCOUNT, and LOGOUT. Below this is the European XFEL logo and the text 'myMdC'. The main content area is titled 'Proposals' and features a search bar with the text '7000' entered. A table of proposals is displayed below the search bar, with columns for Number, Title, PI, Instrument, Status, Start date, E-Log, and Actions. The table contains four entries: 700003 (SCS example data), 700002 (FXE example data), 700001 (Detector Calibration Test Data), and 700000 (Example Data). Three orange arrows point to specific elements: one to the search bar, one to the 'E-Log' column header, and one to the '700001' entry in the table. A footer contains links for Legals & About, Contacts, Developers Information, Report a Problem, European XFEL © 2024, and a 'Back to top' button.

Search or order to find the desired Proposal

Jump to Logbook content

Jump to Proposal information

Number	Title	PI	Instrument	Status	Start date	E-Log	Actions
700003	SCS example data	Andreas Scherz	XMPL	Ready	2023-01-29 23:00:00 UTC		  
700002	FXE example data	Christopher Milne	XMPL	Ready	2021-09-27 07:41:14 UTC		 
700001	Detector Calibration Test Data	Steffen Hauf	XMPL	Ready	2019-01-19 22:00:00 UTC		  
700000	Example Data	Luca Gelisio	XMPL	Ready	2017-11-01 23:00:00 UTC		  

myMdc Introduction (IV)

Proposal "General" information

Proposal no. 700000

Status: 2024-01-19 02:21:50 SET **Runs: 30** **Calibrations: 9** **Team: 22** **Size: 8.45 TiB (9.29 TB)** **Files: 622** **Dark Calibrations: 0**

Proposal Tabs + Add another Runs Beamtime status -

General Public Information Runs Logbook Team Repositories (Beta) Calibration Constants Publication History

General

Proposal DOI [10.22003/XFEL.EU-DATA-700000-00](https://doi.org/10.22003/XFEL.EU-DATA-700000-00) DOI Published

Link to Globus

Proposal Information

#id: 30
Proposal Number: 700000
Name: p700000
Title: Example Data
URL: /metadata/proposals/30
Abstract: The European XFEL (EuXFEL) example data proposal contains experimental datasets from various original beam-times, currently covering the techniques of serial femtosecond crystallography (SFX), coherent diffraction imaging (single particle imaging, SPI), X-ray powder diffraction, small-angle X-ray scattering (SAXS) and X-ray photon correlation spectroscopy (XPCS).

Proposal Folder: p700000
Proposal Path: /gpfs/xfel/d/raw/XMPL/201750/p700000
Logbook Identifier: ---
Auto assign Run Quality: To be evaluated manually
[Auto assign Run Quality -](#)
Preferred data output: Raw
[Preferred data output -](#)

Proposal Last Run: 128
Instrument: Example Data
Instrument Cycle: XMPL: 201750 (starts: 2017-01-01 00:00:00 +0100, ends: 2019-12-31 23:59:59 +0100)
Principal Investigator: Luca Gelisio
Main Proposer: Fabio Dall'Antonia
Local Contact: ---
Start date: 2017-11-08 00:00:00 +0100
End date: 2017-12-31 23:59:59 +0100
Expected end of Embargo: 2018-01-01 00:00:00 +0100
Open data since: 2022-10-27 11:24:32 +0200
 open data?: Yes
No. preparation shifts: 0
No. delivered shifts: ---

Zoom to proposal team [Create Zoom meeting](#)

Proposal status **Beamtime status:** Ready
[Beamtime status -](#)

Important information:

- Proposal path
- Relevant Dates
- Data Output configurations

myMdC Introduction (V)

Proposal "General" information

Cloneable status: [Clone \(ready\)](#)
[Cloneable Status -](#)

Beamtime status: Ready
[Beamtime status -](#)

Start Beamtime date: ---
End Beamtime date: ---
Available: Yes
Proposal system: Example

Description: The European XFEL (EuXFEL) example data proposal contains experimental datasets from various original beam-times, currently covering the techniques of serial femtosecond crystallography (SFX), coherent diffraction imaging (single particle imaging, SPI), X-ray powder diffraction, small-angle X-ray scattering (SAXS) and X-ray photon correlation spectroscopy (XPCS).

Created by: System
Creation date: 2017-11-08 13:49:30 +0100
Modified by: Fabio Dall'Antonia
Last modification date: 2023-03-17 10:47:33 +0100

All Proposal Run types + New

- [e0001]: AGIPD
- [e0005]: Calibration - Dark HG
- [e0007]: Calibration - Dark LG
- [e0006]: Calibration - Dark MG
- [e0003]: Configuration Tests
- [e0014]: Dark
- [e0008]: Diffraction
- [e0010]: Diffraction data
- [e0004]: General
- [e0002]: Sample
- [e0016]: SAXS 500kHz // no pump laser
- [e0011]: scattering
- [e0015]: SFX Jetting
- [e0012]: Single Particle Diffraction
- [e0013]: Test DAQ
- [e0018]: XGM
- [e0017]: XGM calibration
- [e0009]: XPCS

All Proposal Samples

- Water
- Lysozyme (201804 small crystals)
- Lithium Titanate
- No Sample
- Lysozyme
- Silica 50nm
- Vycor
- Sucrose Solution 3% v/v
- Potassium hexacyanoferrate(II) trihydrate
- Cu foil
- Xenon
- 2-Co8_pt14_8fold - 30nm Pt cap
- 1-Co10_Pt_6fold
- Ni-20 MLs - b
- Ni75-11 MLs-b
- Bolometer

Proposal Technique Link

- coherent diffraction imaging
- serial femtosecond crystallography
- small angle x-ray scattering
- x-ray photon correlation spectroscopy
- x-ray powder diffraction

Back
Edit
Clone
+ Add another
Runs
Beamtime status -

myMdC Introduction (VI)

Proposal "Team" and "History"

Proposal no. 900054

STATUS: 2024-01-19 03:45:00 CET | Runs: 82 | Calibrations: 0 | Team: 0 | Size: 0 Bytes | Dark Calibrations: 0

Back Edit Clone Add another Runs Beamtime status

General Public Information Runs Logbook Team Repositories (Beta) Calibrations Constants Publication **History**

Experiment Team

Principal Investigator: Djelloul Boukhelef - (boukhele) as Data Manager & Participant
Main Proposer: Nasser Al-Qudami - (alqudami) as Data Manager & Participant
Team member 1: Luis Maia - (maia) as Participant
Team member 2: Jolanta Sztuk-Dambietz - (jsztuk) as Data Manager & Participant
Team member 3: Maurizio Manetti - (manettim) as Participant
Team member 4: Krzysztof Wrona - (wrona) as Data Manager & Participant
Team member 5: Peter Zalden - (zaldenp) as Data Manager & Participant
Team member 6: Janusz Malka - (jmal) as Data Manager & Participant
Team member 7: Juergen Hannappel - (hannappj) as Participant

Download: CSV JSON Blog format

Back Edit Clone Add another Runs Beamtime status

Team members

User	Data access	Lab Access
Luis Maia	Participant	<input type="radio"/> Yes <input checked="" type="radio"/> No
Djelloul Boukhelef	Data Manager & Participant	<input type="radio"/> Yes <input checked="" type="radio"/> No
Jolanta Sztuk-Dambietz	Data Manager & Participant	<input type="radio"/> Yes <input checked="" type="radio"/> No
Maurizio Manetti	Participant	<input type="radio"/> Yes <input checked="" type="radio"/> No
Krzysztof Wrona	Data Manager & Participant	<input type="radio"/> Yes <input checked="" type="radio"/> No

Proposal no. 900054

Back Edit Clone Add another Runs Delete Beamtime status

General Public Information Runs Team **History**

History

2019-02-14 15:27:59 +0100 – ## System ##		
Preferred data output	Processed	Raw
2019-01-29 18:23:55 +0100 – ## System ##		
Repository	---	XFEL_GPFS_ONLINE_SASE_2
2019-01-29 03:41:47 +0100 – Luis Maia		
Beamtime status	Active	Finished (Beamtime)
2019-01-23 16:24:10 +0100 – Janusz Malka		
Team changes	---	Juergen Hannappel (Participant)
2018-10-19 16:39:06 +0200 – Janusz Malka		
Beamtime status	Ready	Active
2018-09-27 16:38:46 +0200 – Janusz Malka		
Team changes	---	Janusz Malka (Data Manager & Participant)
2018-09-25 17:30:09 +0200 – Luis Maia		
Beamtime status	GPFS Online Configuration Done	Ready
2018-09-25 17:28:04 +0200 – Luis Maia		
Beamtime status	GPFS Offline Configuration Done	GPFS Online Configuration Done

myMdc Introduction (VII)

Proposal "Public information" (DOI) and "Publication" instructions

General Public Information Runs Logbook Team Repositories **Beta!** Calibration Constants Publication History

Proposal no. 900398

DOI: 10.22003/XFEL-EU-DATA-900398-00 Internal Only: DOI not searchable in Metadata

Proposal Number: 900398
Name: p900398
Title: ITLAB SASEd DAQ Logbook tests
Beamtime 1: 2024-01-31 00:00:00 +0100 - 2024-02-20 23:59:59 +0100

General Public Information Runs Logbook Team Repositories **Beta!** Calibration Constants **Publication** History

Proposal no. 700000

DOI: 10.22003/XFEL-EU-DATA-700000-00 DOI Published

Proposal Number: 700000
Name: p700000
Title: Example Data
Abstract: The European XFEL (EuXFEL) example data proposal contains experimental datasets from various original beam-times, currently femtosecond crystallography (SFX), coherent diffraction imaging (single particle imaging, SPI), X-ray powder diffraction, small-angle X-ray photon correlation spectroscopy (XPCS).
Beamtime 1: 2017-11-08 00:00:00 +0100 - 2017-12-31 23:59:59 +0100
Instrument: Example Data

European XFEL myMdc REGISTER LOGIN

Proposal no. 700000

DOI: 10.22003/XFEL-EU-DATA-700000-00
 Proposal Number: 700000
 Name: p700000
 Title: Example Data
 Abstract: The European XFEL (EuXFEL) example data proposal contains experimental datasets from various original beam-times, currently covering the techniques of serial femtosecond crystallography (SFX), coherent diffraction imaging (single particle imaging, SPI), X-ray powder diffraction, small-angle X-ray scattering (SAXS) and X-ray photon correlation spectroscopy (XPCS).
 Beamtime 1: 2017-11-08 00:00:00 +0100 - 2017-12-31 23:59:59 +0100
 Instrument: Example Data
 Instrument Cycle: 201750
 Principal Investigator: Luca Gelsio
 Main Proposer: Fabio Dall'Antonia
 Local Contact: ---
 Expected end of Embargo: 2018-01-01 00:00:00 +0100
 Open data since: 2022-10-27 11:24:32 +0200
 open data?: Yes

This proposal data is open

Would you like to get access to this proposal datasets?
 Please contact us through the open.data@xfel.eu email address.
 Thank you for visiting!

General Public Information Runs Logbook Team Repositories **Beta!** Calibration Constants Publication History

Publications important information

Acknowledgements
 The publications must include an acknowledgement of the help and support provided by the staff of European XFEL.
 Example for this proposal:
 "We acknowledge European XFEL in Schenefeld, Germany, for provision of X-ray free-electron laser beamtime at **Example Data** and would like to thank the staff for their assistance."

DOIs
 Publications related to experiments carried out at the European XFEL facility must cite the DOI of the beamtime data.
 Example for this proposal:
 "Data recorded for the experiment at the European XFEL are available at doi:10.22003/XFEL-EU-DATA-700000-00."

Guidelines
 For publication guidelines, please see the User Publication Policy.

Plans
 Information about any planned publication should be provided to the User Office (useroffice@xfel.eu) as soon as the paper is accepted by a journal.

myMdC Introduction (VIII)

- Proposal "Runs" View
- Run basic Information
- Data assessment defines calibration options and data long term preservation
- Calibration available actions

MENU ☰

Status: 2024-01-19 02:21:50 CET **Runs: 59** **Calibrations: 0** **Team: 22** **Size: 8.45 TiB (8.29 TB)** **Files: 022** **Dark Calibrations: 0**

Proposal no. 700000

Back Edit Clone + Add another **Runs** Beamtime status ▾

General Public Information **Runs** Logbook Team Repositories Beta Calibration Constants Publication History

Proposal Runs

Automatically assess new runs (after being closed by DAQ) as: **To be evaluated manually** ▾
 Automatically start run processing after migration: **No** ▾ (Note: Calibration service will not calibrate runs with run types assessed as "Darks" or "Test experiments" types)

Page auto refresh Off ▾

Run Number (alias)	Run type	Sample Name	Techniques	Start date	Run status	Data Assessment	Calibration	Run Comment	Edit
0128 (X-ray gas monitor measurement for photon diagnostics)	XGM	No Sample		2018-06-15 08:38:40 +0200	Closed	Good	<input type="checkbox"/>		
0127 (X-ray gas monitor calibration for photon diagnostics)	XGM calibration	Bolometer		2018-11-11 16:08:25 +0100	Closed	Good	<input type="checkbox"/>		
0038 (pnCCD detector calibration with copper foil, SQS instrument)	Test DAQ	Cu foil		2020-08-14 20:00:22 +0200	Closed	Good	<input type="checkbox"/>		
0037 (pnCCD detector calibration with copper foil, SQS instrument)	Test DAQ	Cu foil		2020-08-14 19:59:11 +0200	Closed	Good	<input type="checkbox"/>		
0036 (pnCCD detector calibration with copper foil, SQS instrument)	Test DAQ	Cu foil		2020-08-14 16:08:07 +0200	Closed	Good	<input type="checkbox"/>		
0035 (pnCCD detector calibration with copper foil, SQS instrument)	Test DAQ	Cu foil		2020-08-14 16:07:02 +0200	Closed	Good	<input type="checkbox"/>		
0034 (SPI on sucrose solution, AGIPD detector at SPB instrument)	Single Particle Diffraction	Sucrose Solution 3% v/v	coherent diffraction imaging	2021-06-01 02:25:08 +0200	Closed	Good			
0033 (SAXS on vycor sample, AGIPD detector at MID instrument)	scattering	Vycor	small angle x-ray scattering	2021-04-10 14:48:20 +0200	Closed	Good			
0031 (SFX on Hen egg-white lysozyme, AGIPD detector)	Diffraction data	Lysozyme	serial femtosecond crystallography	2021-04-15 10:48:26 +0200	Closed	Good			
0030 (SFX on Hen egg-white lysozyme, AGIPD detector)	Diffraction		serial femtosecond crystallography	2020-03-09 01:20:02 +0100	Closed	Good			
0029 (SFX on Hen egg-white lysozyme, AGIPD detector)	Diffraction		serial femtosecond crystallography	2020-03-09 01:07:51 +0100	Closed	Good			

Run Number

Run Type, Sample and Technique(s) associated to each run

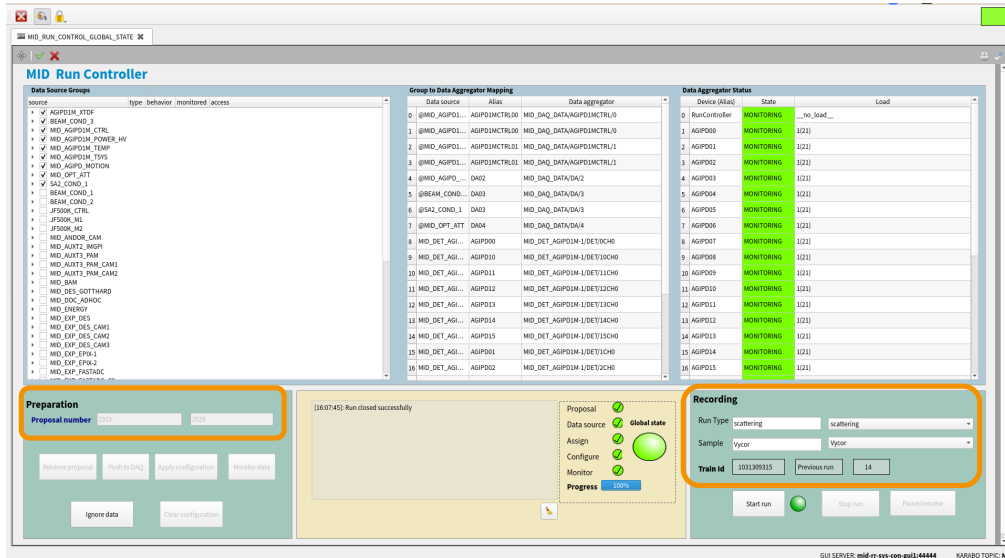
Data Output configurations

Run data quality assessment

Run calibration

myMdC Introduction (IX)

- myMdC integration with
 - DAQ
 - Data Migration service
 - Calibration Pipeline service



Proposal no. 002929

MENU

Back Edit Clone Add another Runs Delete Beamtime status

General Public Information Runs Team Repositories Calibration Constants History

Proposal Runs



Run Number (alias)	Run type	Sample Name	Start date	Run status	Data Assessment	Calibration	Run Comment	Edit
0020	scattering	Vycor	2021-04-08 18:12:14 +0200	Closed	Run Quality	<input type="checkbox"/>		
0019	scattering	Vycor	2021-04-08 18:10:29 +0200	Closed	Run Quality	<input type="checkbox"/>		
0018	scattering	Vycor	2021-04-08 18:08:30 +0200	Closed	Run Quality	<input type="checkbox"/>		
0017	scattering	Vycor	2021-04-08 18:06:44 +0200	Closed	Run Quality	<input type="checkbox"/>		
0016	scattering	Vycor	2021-04-08 18:01:40 +0200	Closed	Run Quality	<input type="checkbox"/>		
0015	scattering	Vycor	2021-04-08 17:59:05 +0200	Closed	Run Quality	<input type="checkbox"/>		
0014	scattering	Vycor	2021-04-08 16:07:20 +0200	Closed	Good (migrate data to Maxwell)		Good (migrate data to Maxwell)	
0013	scattering	Vycor	2021-04-08 16:05:18 +0200	Closed	Unclear (migrate data to Maxwell)		Unclear (migrate data to Maxwell)	
0012	scattering	Vycor	2021-04-08 16:04:04 +0200	Closed	Not interesting (data won't be migrated to Maxwell)		Not interesting (data won't be migrated to Maxwell)	
0016	scattering	Vycor	2021-04-08 18:01:40 +0200	Closed	Good	<input type="checkbox"/>		
0015	scattering	Vycor	2021-04-08 17:59:05 +0200	Closed	Good	<input type="checkbox"/>		
0014	scattering	Vycor	2021-04-08 16:07:20 +0200	Closed	Good	<input type="checkbox"/>		
0013	scattering	Vycor	2021-04-08 16:05:18 +0200	Closed	Good	<input type="checkbox"/>		
0012	scattering	Vycor	2021-04-08 16:04:04 +0200	Closed	Good	<input type="checkbox"/>		

Calibrated data available on the proposal /proc folder.

Start new calibration

Repeat previous calibration

myMdC Introduction (X)

Run "General" information

Proposal no. 700000

Back Edit Clone Add another Runs Beamtime status

General Public Information Runs Logbook Team Repositories Calibration Constants Publication History

Proposal Runs

- Automatically assess new runs (after being closed by DAO) as: To be evaluated manually
- Automatically start run processing after migration: No (Note: Calibration service will not calibrate runs with run types assessed as "marks" or "Test experiments" types)

Run Number (alias)	Run type	Sample Name	Techniques
0128 (X-ray gas monitor measurement for photon diagnostics)	XGM	No Sample	
0127 (X-ray gas monitor calibration for photon diagnostics)	XGM calibration	Beam stop	
0038 (pnCCD detector calibration with copper foil, SQS instrument)	Test DAQ	Cu foil	
0037 (pnCCD detector calibration with copper foil, SQS instrument)	Test DAQ	Cu foil	
0036 (pnCCD detector calibration with copper foil, SQS instrument)	Test DAQ	Cu foil	
0035 (pnCCD detector calibration with copper foil, SQS instrument)	Test DAQ	Cu foil	
0034 (SPI on sucrose solution, AGIPD detector at SPB instrument)	Single Particle Diffraction	Sucrose Solution 3% v/v	coherent diffraction imaging
0033 (SAXS on vycor sample, AGIPD detector at MID instrument)	scattering	Vycor	small angle x-ray scattering
0031 (SFX on Hen egg-white lysozyme, AGIPD detector)	Diffraction data	Lysozyme	serial femtosecond crystallography

Proposal no. 700000 > Run no. 34

Back Edit Clone Add another Delete

General Run > Slow Data Parameters Processing Reports Data Groups History

General

#id: 125468
 Proposal Number: 700000
 Run Number: 34
 Run Alias: SPI on sucrose solution, AGIPD detector at SPB instrument
 Run Folder: r0034
 Run type: Single Particle Diffraction
 Sample Name: Sucrose Solution 3% v/v
 Start date: 2021-06-01 02:25:08 +0200
 End date: 2021-06-01 02:29:57 +0200
 First Train: 1077459200
 Last Train: 1077462095
 Size: 3.83 TiB (4.21 TB)
 No. files: 201
 Available: Yes
 Run Quality: Good
 Run Status: Closed

Original Format: ---
 DAQ system message: ---

Migration requested at: 2021-06-01 02:30:14 +0200
 Migration begin at: 2021-06-01 02:30:28 +0200
 Migration end at: 2021-06-01 02:34:38 +0200

Calibration last request at: 2022-01-28 13:38:18 +0100
 No. Calibration Requests: 1
 Calibration last begin at: 2022-01-28 13:38:23 +0100
 Calibration last end at: 2022-01-28 14:59:38 +0100
 Calibrated Process data status: Available
 Calibration Pipeline auto Reply: COMPLETED

Description: REMARK: Run cloned from proposal 900201 run 0372
 Created by: Varun Singh
 Creation date: 2022-03-03 13:12:57 +0100
 Modified by: Fabio Dall'Antonia
 Last modification date: 2022-11-01 17:27:47 +0100

Runs Techniques [Link technique](#)

- coherent diffraction imaging

myMdC Introduction (XI)

Proposal "Calibration Constants" integration with

- CalCat
- Calibration Pipeline
- Calibration Report

Report

Id: 13
 Name: dark_2021-05-03_13:45:56_0984_0984_210211_122050.pdf
 File path: /gpfs/xfel/xfel_cal_cat/cal_cat_data/storedata/reports/CALCAB/HED_DET_AGP0500K03/dark/data/2021-05-03_13:45:56_0984_0984_210211_122050.pdf
 Download: [icon]
 Available: [icon]
 Description: ...
 Created by: Admin User
 Creation date: 2021-02-11T12:27:19-0100
 Modified by: Admin User
 Last modification date: 2021-02-11T12:27:19-0100

Associated Calibration Constant Versions

Display all versions

In the table below, Karabo DA and Detector of the associated Physical Detector Unit are shown as they were at the moment of the calibration constant version's creation date and time

Name	Deployed	Good quality	Start date	Validity start	Creation date	Calibration Constant	Physical Detector Unit	Detector
20210211_112807_akt=0	false	true	2020-09-23 13:26:59	2020-09-23 13:26:59	2021-02-11 13:26:59	AGPD-Type_BadPhaseDark_AGP0 Det#7(p30904)	AGPD_SV1_AGIPD11_M 1159 (AGIPD05)	HED_DET_AGP0500K03
20210211_112811_akt=0	false	true	2020-09-23 13:26:59	2020-09-23 13:26:59	2021-02-11 13:26:59	AGPD-Type_BadPhaseDark_AGP0 Det#8(p3127)	AGPD_SV1_AGIPD11_M 1179 (AGIPD05)	HED_DET_AGP0500K03
20210211_112811_akt=0	false	true	2020-09-23 13:26:59	2020-09-23 13:26:59	2021-02-11 13:26:59	AGPD-Type_BadPhaseDark_AGP0 Det#8(p3127)	AGPD_SV1_AGIPD11_M 1179OUT (AGIPD03)	HED_DET_AGP0500K03

Jungfrau Dark Image Characterization

European XFEL

Detector group
 Based on data sample: /gpfs/xfel/exp/SPB/202130p900204/raw
 Release : 3.1.0-boostv4
 March 18, 2021

Proposal no. 900204

Status: 2024-01-19 04:48:03 CET Runs: 188 Calibrations: 16 Team: 1 Size: 69.32 TiB (76.22 TB) Files: 6306 Dark Calibrations: 21

Back Edit Clone Add another Runs Beamtime status

General Public Information Runs Logbook Team Repositories **Beta!** Calibration Constants Publication History

Calibration Constants

Request Dark Run Calibration

Detector: SPB_DET_AGIPD1M-1

Detector Units:

- AGIPD00 (Q1M1)
- AGIPD01 (Q1M2)
- AGIPD02 (Q1M3)
- AGIPD03 (Q1M4)
- AGIPD04 (Q2M1)
- AGIPD05 (Q2M2)
- AGIPD06 (Q2M3)
- AGIPD07 (Q2M4)
- AGIPD08 (Q3M1)
- AGIPD09 (Q3M2)
- AGIPD10 (Q3M3)
- AGIPD11 (Q3M4)
- AGIPD12 (Q4M1)
- AGIPD13 (Q4M2)
- AGIPD14 (Q4M3)
- AGIPD15 (Q4M4)

Operation Mode: Three gains (ADAPTIVE_GAIN)

Operation Mode Description: Standard operation mode for AGIPD and Jungfrau detectors with 3 gain stages used. To be able to process calibration constants from dark data, the data should be taken for each gain stage in separate runs. Therefore, the run number for each gain stage has to be specified in the following order #1- high (HG/G0) #2-medium (MG/G1) #3-low (LG/G2). NOTE: This operation mode must be used for Adaptive and Fixed gain runs for AGIPD.

Run Number(s):

- Run #1: [input]
- Run #2: [input]
- Run #3: [input]

Description: [input]

[Request]

Dark runs requests

Status	Last Updated at	Detector	Detector Units	Operation Mode	External resources
Finished	2021-07-12 11:52:48 +0200	SPB_IRDA_JF4M	JNGFR01, JNGFR02, JNGFR03, JNGFR04, JNGFR05, JNGFR...	Adaptive Gain	166, 167, 168
Finished	2021-05-03 13:45:56 +0200	SPB_IRDA_JF4M	JNGFR01, JNGFR02, JNGFR03, JNGFR04, JNGFR05, JNGFR...	Adaptive Gain	141, 142, 143
Finished	2021-04-19 18:42:43 +0200	SPB_IRDA_JF4M	JNGFR01, JNGFR02, JNGFR03, JNGFR04, JNGFR05, JNGFR...	Adaptive Gain	126, 127, 128
Finished	2021-04-19 18:41:43 +0200	SPB_IRDA_JF4M	JNGFR01, JNGFR02, JNGFR03, JNGFR04, JNGFR05, JNGFR...	Adaptive Gain	115, 116, 117
Finished	2021-04-19 18:39:40 +0200	SPB_IRDA_JF4M	JNGFR01, JNGFR02, JNGFR03, JNGFR04, JNGFR05, JNGFR...	Adaptive Gain	111, 112, 113

myMdc Introduction (XII)

myMdc RESTful API

- Token-based Authentication
 - Oauth 2.0
- Role-Based Access Control
- Data Validation and Error Handling
- Pagination and Filtering
- Rate Limiting to prevent abuse
- Swagger/OpenAPI

➤ <https://in.xfel.eu/metadata/api-docs/>

Metadata-client python package

- myMdc official python package

➤ <https://pypi.org/project/metadata-client/>

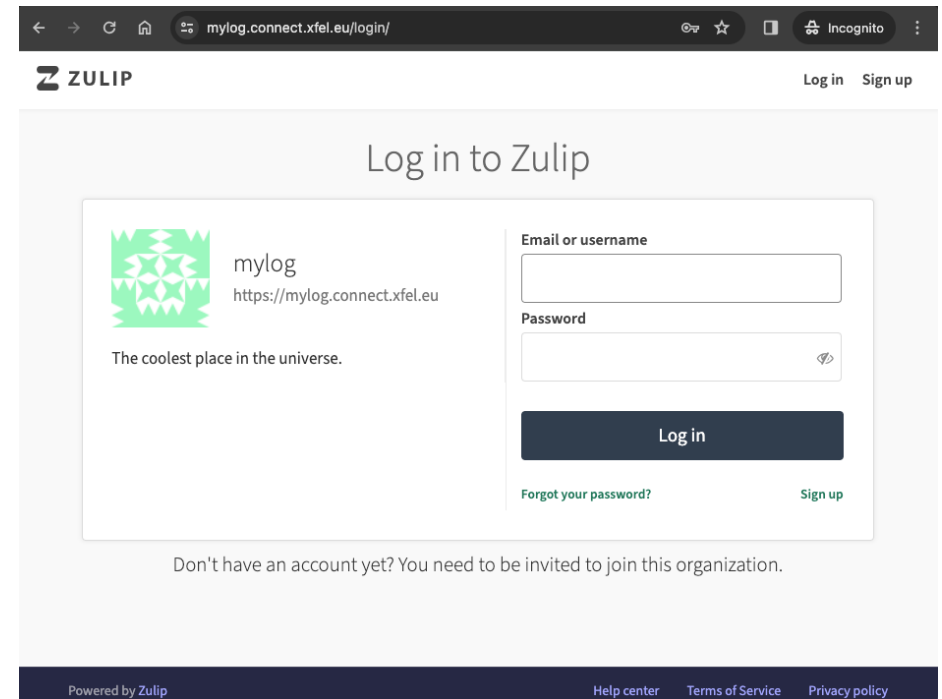
The image shows a screenshot of the myMdc web interface and a terminal window. The web interface displays the 'MY TOKENS' section, which includes an 'Application: Mdc token' card with fields for Application Id, Secret, Scopes, and Callback urls. There are 'Edit' and 'Destroy' buttons for this token. A sidebar on the right shows navigation options for 'Users', 'Proposals', and 'Data Source Groups'. Below this, a REST client interface is shown for the endpoint `/proposals/by_number/{number}`. The parameters section shows a required integer parameter 'number' with the value '900184'. The 'Execute' button is highlighted. The 'Responses' section shows a 200 status code and a detailed JSON response body for proposal 900184, including fields like 'id', 'number', 'title', 'abstract', 'proposal_folder', 'def_proposal_path', 'dot', 'url', 'instrument_id', 'instrument_cycle_id', 'principal_investigator_id', 'main_proposer_id', 'local_contact_id', 'oria_repository_id', 'last_run', 'begin_at', 'end_at', 'release_at', 'num_preparation_shifts', 'num_delivered_shifts', 'flag_public', 'flag_proposal_system', and 'flag_nonfinal_data_output'. The terminal window shows the execution of a Python script named 'maia.py' that uses the 'metadata-client' package to fetch proposal information for proposal number 900054.

myLog

myLog is the **new** European XFEL Electronic Logbook

<https://mylog.connect.xfel.eu/>

myLog is a Zulip based platform



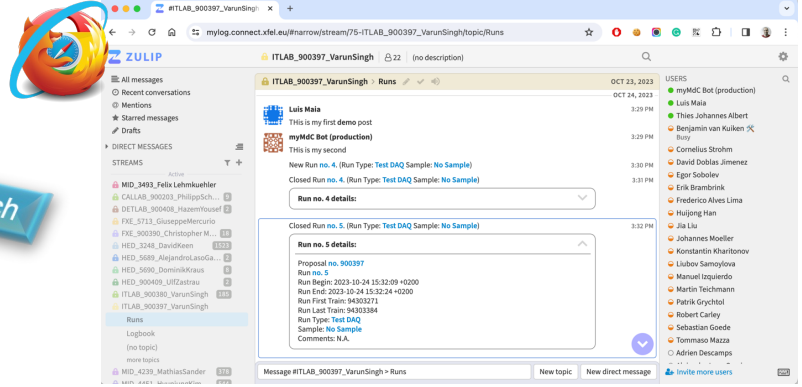
myLog Purpose (I)

- Provide a means to annotate experiment details during the beamtime, but also provides transparent sharing of data and a robust way of communication between scientists
- Foster communication and documentation between all involved parties
- Easy integration of new sources of data (e.g. Karabo, DAMNIT, etc.) to leverage automatic documentation of relevant metadata
- Flexible configuration and APIs allowing myMdC to automatically manage myLog configuration, membership and desired configuration of data sources

myLog Introduction (I)



Content search



Zulip - mylog

ITLAB_900397_VarunSingh (no description)

ITLAB_900397_VarunSingh > Runs

Luis Maia: This is my first demo post

myMdC Bot (production): This is my second

New Run no. 4. (Run Type: Test DAQ Sample: No Sample)

ITLAB_900397_VarunSingh > Logbook

Luis Maia: This run looks really good

ITLAB_900397_VarunSingh > Runs

myMdC Bot (production): Closed Run no. 4. (Run Type: Test DAQ Sample: No Sample)

Run no. 4 details:

Closed Run no. 5. (Run Type: Test DAQ Sample: No Sample)

Run no. 5 details:

Proposal no. 900397
Run no. 5
Run Begin: 2023-10-24 15:32:09 +0200
Run End: 2023-10-24 15:32:24 +0200
Run First Train: 94303271
Run Last Train: 94303384
Run Type: Test DAQ
Sample: No Sample
Comments: N.A.

Message actions

Message #ITLAB_900397_VarunSingh > Runs

New topic New direct message

Logbook name

Entry "topic" context

Entry Actions

Entry content

New entry

User available Logbooks (streams) and respective topics

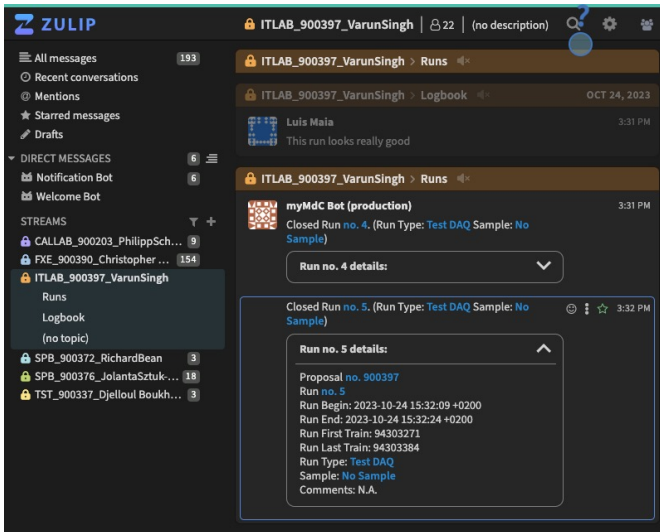
BACK TO STREAMS

MID_900389_AndersMads... 179

Filter topics

- Diamond Detectors 55
- 30keV 1
- ShortPulses 7
- PAM 1
- Test_ZYLA&NanoCube 2
- Levitation device 10
- Diamond Monochromator 18
- Laser 1
- Shift Summary 8
- Detector status 17
- Alignment 47
- Beam status 5
- Data Analysis 1
- Data runs&scans 3
- Problems 1
- Chat 1
- Sample status
- Darks
- Experimental Setup
- (no topic) 1

myLog Introduction (II)



Proposal no. 900397

Status: 2024-01-24 17:02:26 CET | Runs: 5 | Calibrations: 0 | Team: 21 | Size: 0 Bytes | Files: 0 | Dark Calibrations: 0

Back Edit Clone Add another Runs Beamtime status

General Public Information Runs Logbook Team Repositories (Beta) Calibration Constants Publication History

Logbook: ITLAB_900397_VarunSingh [Delete]

Search [x] [i] All Topics [Q] Sort Desc (latest post first) [Clear Filters]

Logbook name and link

Preview

Logbook "Preview" Tab

Content search

Entry actions and info

Runs

This run data is promising!

Hello

Closed Run no. 5. (Run Type: Test DAQ Sample: No Sample)

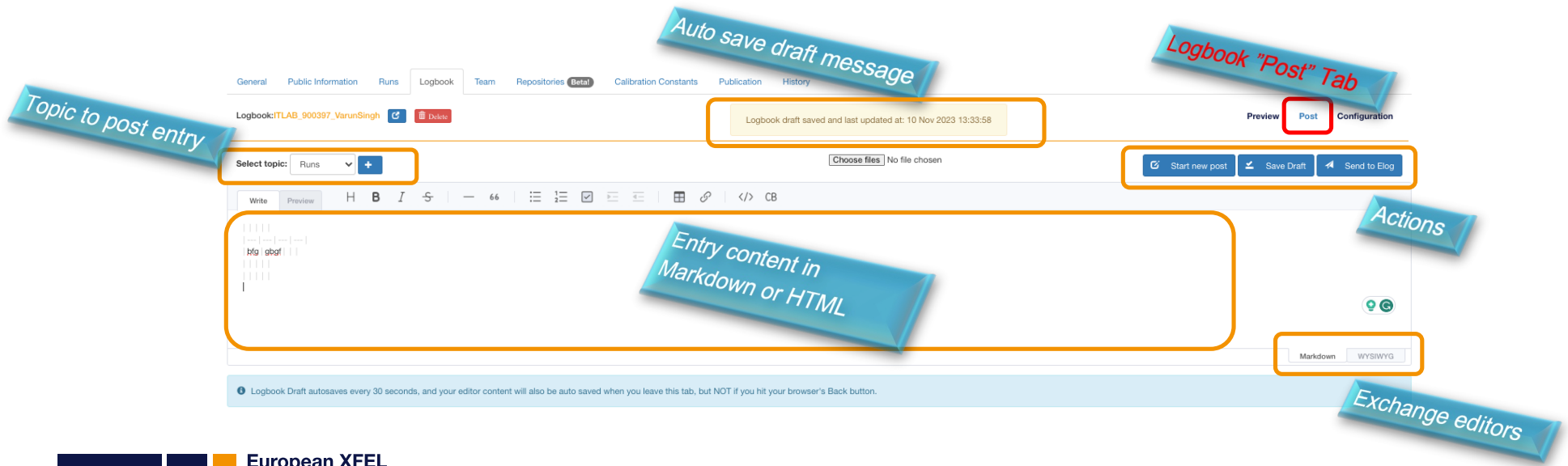
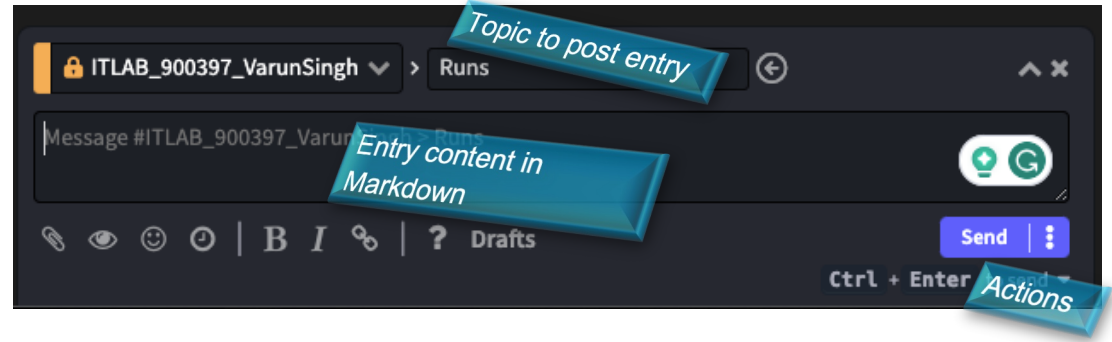
Run no. 5 details:

Proposal no. 900397
Run no. 5
Run Begin: 2023-10-24 15:32:09 +0200
Run End: 2023-10-24 15:32:24 +0200
Run First Train: 94303271
Run Last Train: 94303384
Run Type: Test DAQ
Sample: No Sample
Comments: N.A.

Entries in topic "Runs"

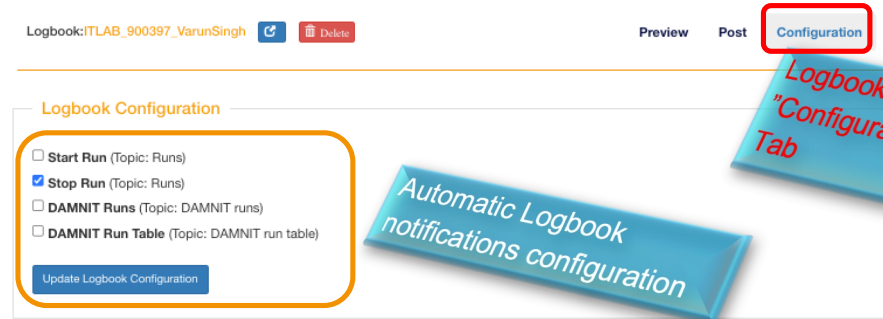
myLog Introduction (III)

- Posting an entry in the proposal logbook



myLog Introduction (IV)

- Configuration of **sources of metadata** that can post **automatically** to the logbook



MID_5397_ChristianGutt > Data runs&scans NOV 23, 2023

DAMNIT 5397 test 11:13 PM

Run	Timestamp	Comment	Pulses	Run type	XGM intensity [uJ]	Photon energy [keV]	Sample	Trains	Raw size [TB]
1	23:05:42 23/11/2023	test damnit	1	Test DAQ	1915.0964	10.0000	Eggyolk	427	0.6372

- DAMNIT BOT** automatic posting configured data into Logbook as soon as a new run is processed.
- DAMNIT BOT** automatic update entry in the Logbook upon reprocessing of the run.

ITLAB_900397_VarusSingh > Runs 3:31 PM

myMdc Bot (production)
Closed Run no. 4. (Run Type: Test DAQ Sample: No Sample)

Run no. 4 details:

Proposal no. 900397
Run no. 4
Run Begin: 2023-10-24 15:30:44 +0200
Run End: 2023-10-24 15:31:31 +0200
Run First Train: 94302419
Run Last Train: 94302849
Run Type: Test DAQ
Sample: No Sample
Comments: N.A.

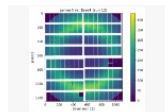
- myMdc Bot** automatic posting into Logbook information about each run start/stop as as soon as it starts/stops on the DAQ.

myLog Introduction (V)

- Sharing information from DAMNIT to myLog
- DAMNIT BOT exporting images

The screenshot shows the 'Data And Metadata Inspection Interactive Thing' interface. On the left, a table lists run data. On the right, a heatmap plot shows 'james3 vs. Event (run 7)'. A dialog box for posting to the logbook is open, showing the following details:

- Stream: ITLAB_900369_VarunSingh
- Topic: New_topic
- Title: (empty)
- Column selection: Select all (checked), Deselect all (unchecked)
- Selected items: Run, Timestamp, Comment, Run size (TB), Trains
- Unselected items: Pulse intensity, Pulses5, XGM intensity



DAMNIT BOT exporting tables

The screenshot shows the 'Data And Metadata Inspection Interactive Thing' interface. A table of run data is visible. A dialog box for posting to the logbook is open, showing the following details:

- Stream: ITLAB_900369_VarunSingh
- Topic: New_topic
- Title: Title input
- Column selection: Select all (checked), Deselect all (unchecked)
- Selected items: Run, Timestamp, Comment, Run size (TB), Trains
- Unselected items: Pulse intensity, Pulses5, XGM intensity

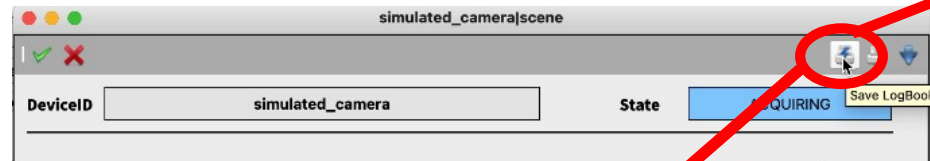
Title input

Run	Timestamp	Comment	Run size (TB)	Trains
5	21:46:46 31/08/2022	A comment	2.7173	3105
6	21:52:08 31/08/2022	Another comment	2.7330	3123
7	21:57:32 31/08/2022	Yet another comment	2.7182	3106

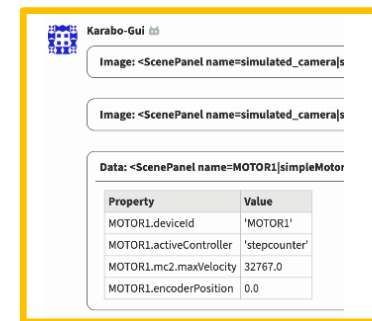
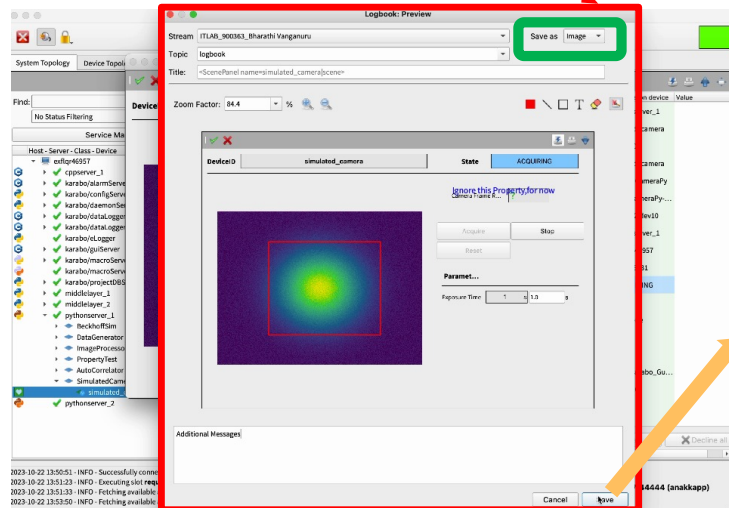
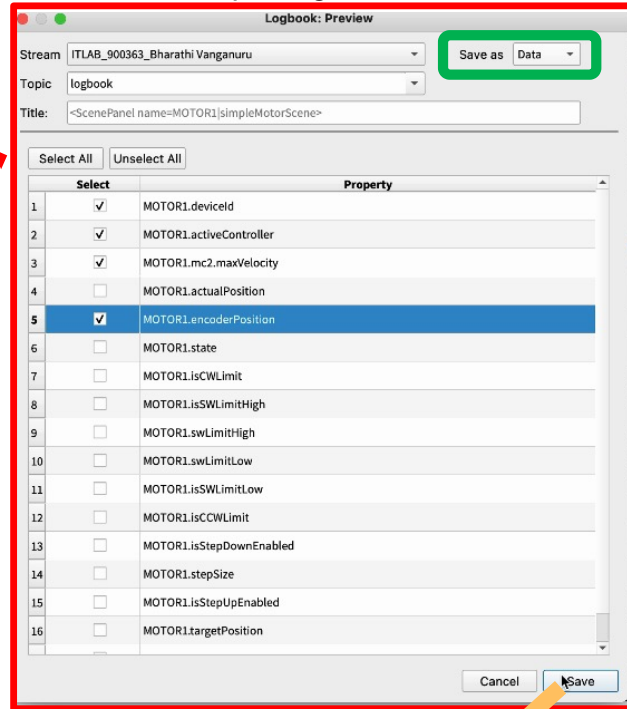
myLog Introduction (VI)

Sharing information from Karabo to myLog

Karabo-Gui BOT exporting images

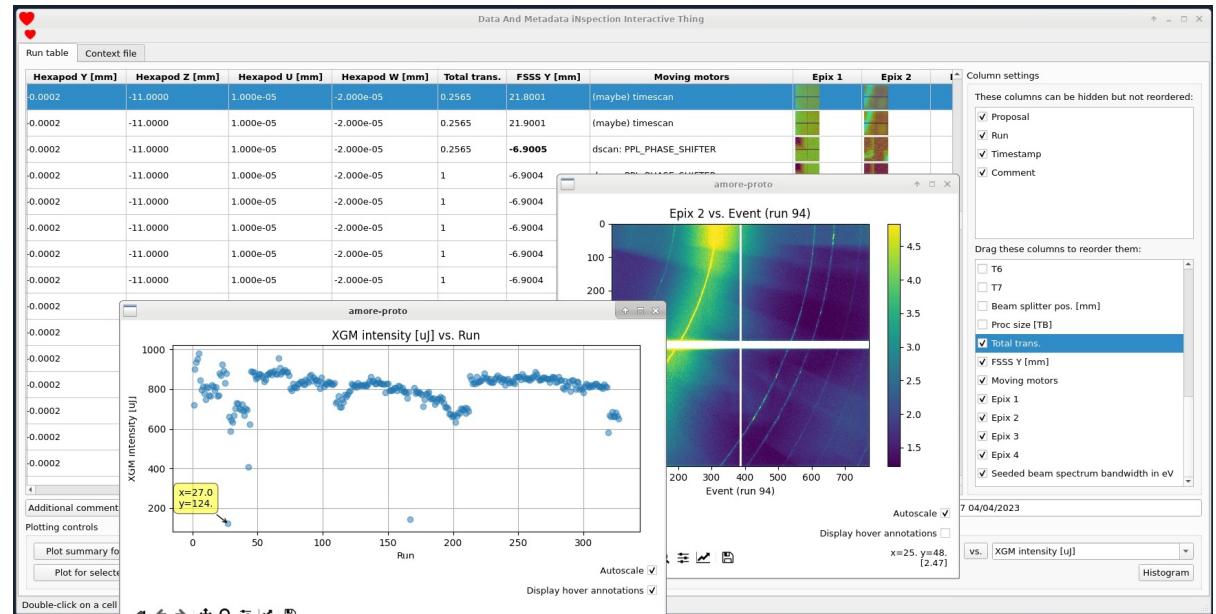


Karabo-Gui BOT exporting tables



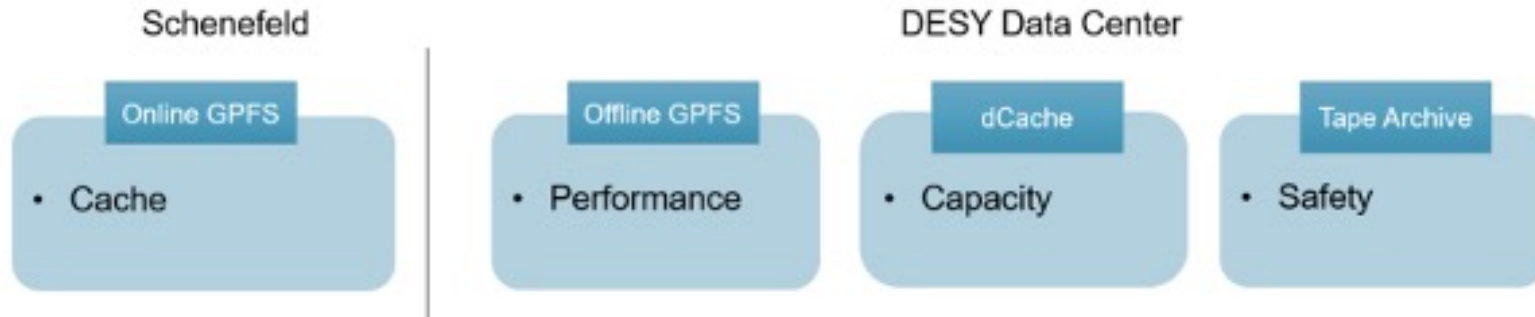
DAMNIT

- Dedicated session
- Experiment overview with DAMNIT



Storage and computing (I)

See for more details: <https://iopscience.iop.org/article/10.1088/1742-6596/898/6/062049>



- Data is collected at the instrument into the 'Online Storage'
 - Used for online data analysis/preview during/immediately after data acquisition

- If data is assessed as 'good' or 'unclear' it is moved to the the 'Offline Storage' accessible from Maxwell
 - Used for offline calibration and data analysis

- High capacity mass storage (dCache) accessible from Maxwell
 - Mid term storage for the raw data
 - Interface to the tape archive

Proposal no. 004316

Status: 2024-01-25 00:06:06 CET | Runs: 159 | Calibrations: 0 | Team: 03 | Size: 641.82 GiB (689.15 GB) | Files: 6870 | Dark Calibrations: 0

[Back](#) | [Edit](#) | [Clone](#) | [+ Add another](#) | [Runs](#) | [Beamline status](#) | [Repositories \(Beta\)](#) | [Calibration Constants](#) | [Publication](#) | [History](#)

Repositories

Request the Copy of Raw Data to available Remote repositories

No remote repositories are available.

Raw data availability per run and repository

Run Number (beam)	Data Group	Run Quality	XFEL GPFS Raw online for SASE 1	XFEL GPFS online for SASE 2	XFEL GPFS online for SASE 3	XFEL GPFS online Raw data in DESY-CC	dCache Raw Data in DESY-CC
0159	raw_e0001_n0159	Good					
0158	raw_e0001_n0158	Good					
0157	raw_e0001_n0157	Good					

Storage and computing (II)

- EuXFEL experiment data is stored under

```
/gpfs/exfel/exp/${INSTRUMENT}/${CYCLE}/p${PROPOSAL_ID}/
```

- Every proposal has these four sub-directories:

Storage	Quota	Permission	Comments
raw	None	Read	Fast accessible raw data
usr	5TB	Read/Write	User data, results - synced between online/offline storage
proc	None	Read	Facility processed data e.g. calibrated data
scratch	None	Read/Write	Temporary data

What is Maxwell?

- The Maxwell Cluster is a resource dedicated to parallel and multi-threaded application
 - The cluster is managed by SLURM scheduler.

- Maxwell Cluster is a shared resource that incorporates resources for Photon Science data analysis from CFEL, CSSB, Petra4, EuXFEL, etc.
 - EuXFEL owns ~50% of the Hardware (455/931 nodes)

European XFEL USER PORTAL TO THE EUROPEAN XFEL

HOME EUROPEAN XFEL PRINT LOG OUT

UPEX time: 2024-01-25 12:13

Welcome Mr. Luis Maia. [Log out](#)

Mailing lists

You can subscribe to mailing lists here just in one click. Your email luis.maia@xfel.eu will be used for the subscription.

- **computing@xfel.eu - European XFEL data analysis-related topics**
 - Subscribers to this mailing list will get email announcements about: updates to the European XFEL Scientific Data Policy; availability of the data analysis infrastructure (e.g. Maxwell Cluster, GPFS file system); availability of new software packages for data analysis; data calibration; any other topics relevant to the analysis of data collected at the European XFEL.
- **user.events@xfel.eu - European XFEL user event announcements**
 - Announcements of workshops and conferences Subscribers to this mailing list will receive email announcements about the annual European XFEL Users' Meeting, as well as workshops and conferences related to our instruments.
- **user.seminars@xfel.eu - European XFEL user seminar announcements**
 - Announcements of science seminars. Subscribers to this mailing list will receive email announcements related to (online) science seminars organized by European XFEL.

● - Already subscribed; ● - Not subscribed yet; ● - Subscription status unknown;

See for more details: <https://in.xfel.eu/upex/user/maillinglist>

The entire cluster is managed by SLURM scheduler. Different rules apply to different partitions. See [Partitions](#) and [Access](#) pages for details.

- To get started, please have a look at the [Getting Started](#) page!
- The [Compute infrastructure](#) page provides a list of currently available nodes & configurations.
- The [Partitions](#) page provides a quick overview of the nodes, capacities, features and limits of the individual partitions.

Read the documentation! It should cover at least the essentials. If you come across incorrect or outdated information: please let us know!

Contact

For any questions, problems, suggestions please contact our ticket system: maxwell.service@desy.de

All Announcements will be sent via maxwell-user@desy.de.

We strongly recommend that all maxwell-users subscribe to the mailing list. Only user of the maxcpu partition are automatically added.

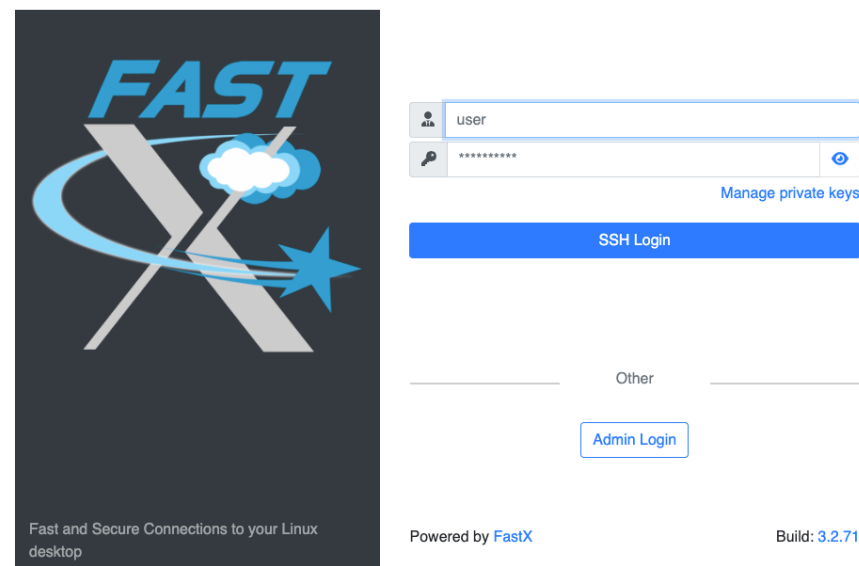
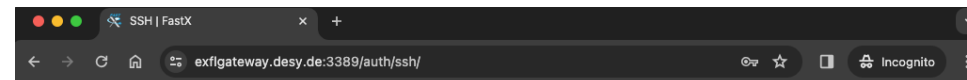
Subscribe: <https://lists.desy.de/sympa/info/maxwell-user>

See for more details: <https://confluence.desy.de/display/MXW/Compute+Infrastructure>

Accessing Maxwell – FastX (I)

■ Login at <https://max-exfl-display.desy.de:3389> with your **Campus account**

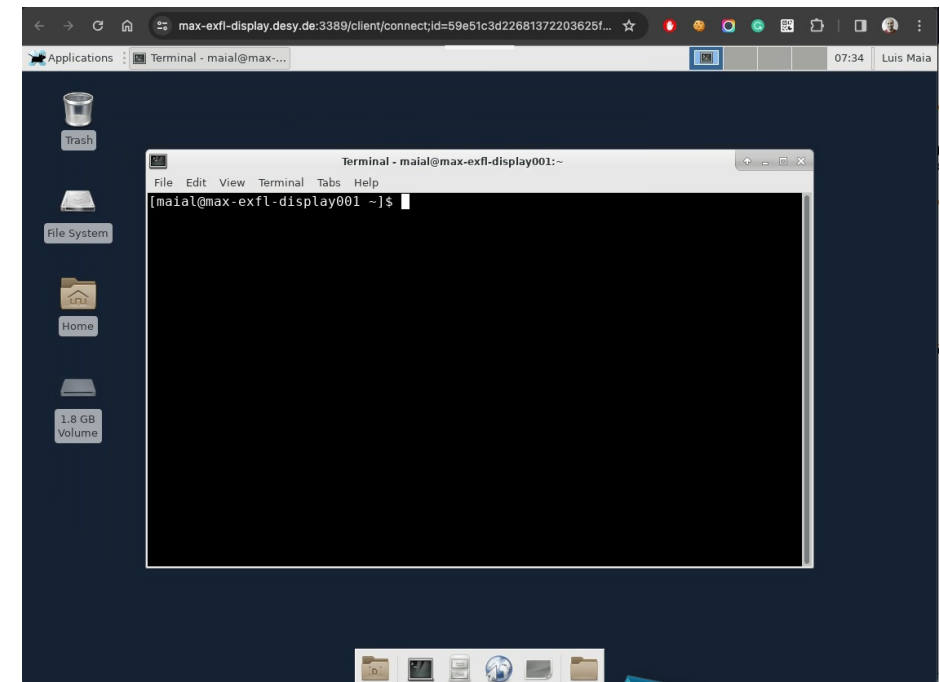
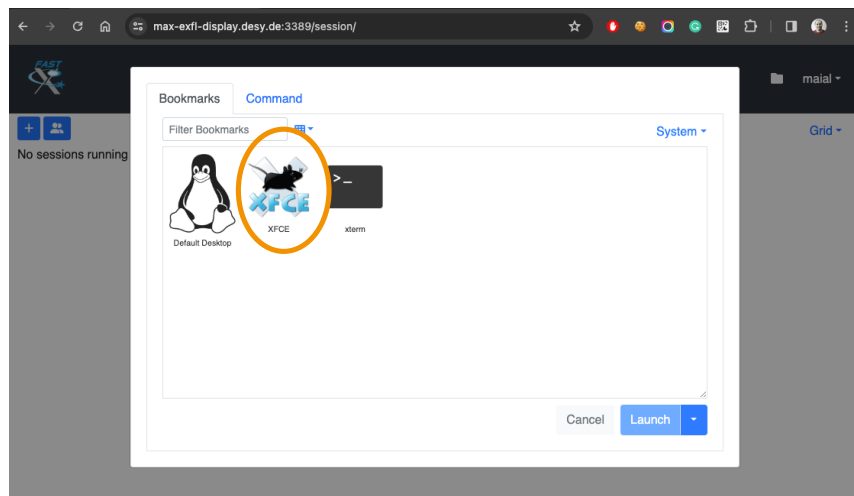
- Accessible from outside the DESY/EuXFEL network
 - ▶ **News: 2nd-factor authentication required from 30.01.2024!**
- 2 'shared node' used by multiple users
- Suitable for:
 - ▶ **Job submissions to SLURM**
 - **SBATCH** or **SALLOC**
 - ▶ Short/low intensity compute jobs
 - ▶ Code compilations (if it's not using all cores)
 - ▶ Programs requiring GUI/GPU acceleration
- Not suitable for:
 - ▶ Long-running/intense workloads like simulation or analysis. SLURM should be used instead.



See for more details: <https://confluence.desy.de/display/MXW/FastX+on+Display+nodes>

Accessing Maxwell – FastX (II)

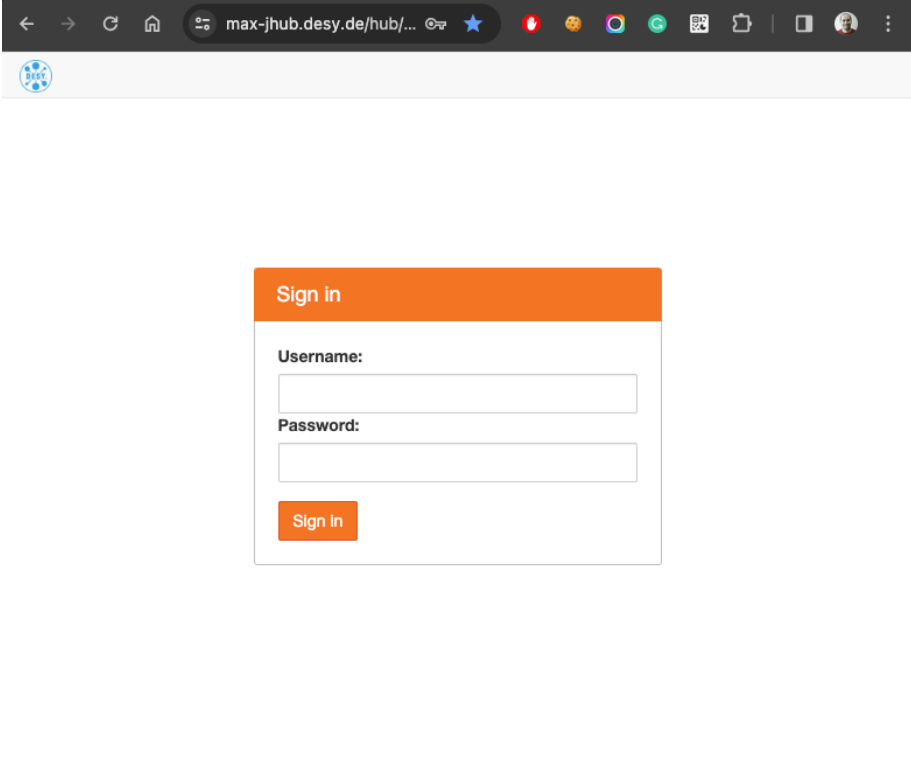
- After login
 - Create a new session
 - Select '**XFCE**' for a virtual desktop, or '**xterm**' for a terminal session
 - ▶ Session is kept, consider doing **kinitt**



Remote desktop session on Maxwell shared display node

Accessing Maxwell – JupyterHub (I)

- Login at <https://max-jhub.desy.de> with your **Campus account**
 - Accessible from outside the DESY/EuXFEL network
 - ▶ **News: 2nd-factor authentication required from 30.01.2024!**
 - Easiest way to run and access Jupyter on Maxwell
 - Provides graphical interface to allocating a node via slurm
 - When using shared nodes
 - ▶ Shared nodes have a max allocation time of 7 days
 - ▶ Dedicated nodes have a max allocation time of 8 hours
 - If a dedicated node is picked it can be suitable for intense computation
 - ▶ EuXFEL users should use partitions: `jhub`
 - Usage of `upex` and `allcpu` should be only used for intensive jobs



The screenshot shows a web browser window with the address bar displaying "max-jhub.desy.de/hub/...". The page content features a "Sign in" form with an orange header. The form includes two input fields: "Username:" and "Password:", each with a corresponding text box. Below the fields is an orange "Sign in" button.

See for more details: <https://confluence.desy.de/display/MXW/JupyterHub+on+Maxwell>
<https://rtd.xfel.eu/docs/data-analysis-user-documentation/en/latest/jhub/>

Accessing Maxwell – JupyterHub (II)

■ Select allocation options and start session

Maxwell Jupyter Job Options

Annotations: *Settings Help*, *Special Reservations*, *Job allocation options*, *Start session*

Form fields: Maxwell partitions (jhub), Python Version (Conda Python 3.9), Launch modus (JupyterLab), Number of Nodes (1), Constraints, Environment.

Start button

Node and GPU availability						
Partition	# nodes	# avail	# GPUs	# P100	# V100	# A100
jhub	8	69/240	0	0	0	0
allcpu	598	483	0	0	0	0
allgpu	188	102	102	49	22	27
cfel	19	19	5	0	1	4
csb	8	6	0	0	0	0
xfel	438	309	17	14	1	2
maxcpu	126	103	0	0	0	0
maxgpu	88	37	37	16	17	4
mpaj	4	3	3	0	0	3
petra4	62	56	0	0	0	0
ps	36	7	7	7	0	0
pex	26	10	6	6	0	0
uke	16	16	0	0	0	0
upex	438	309	17	14	1	2

last updated: 19/01/2024 07:48:01

max-jhub.desy.de/user/maia/lab

Launcher

Filter files by name

File Edit View Run Kernel Git Tabs Settings Help

Environment options:

- Python 3.9
- Astra-Toolbox
- Bash
- European XFEL Geometry Powder
- Holoviews
- noise2inverse
- Pennylane PennyLane
- pyFAI
- Pytorch
- Qiskit
- RAPIDS
- Spyder
- Tensorflow
- Tomopy
- European XFEL xfel (202302)
- European XFEL xfel (202401)
- European XFEL xfel (current)
- European XFEL xfel (Python 3.7)

Console

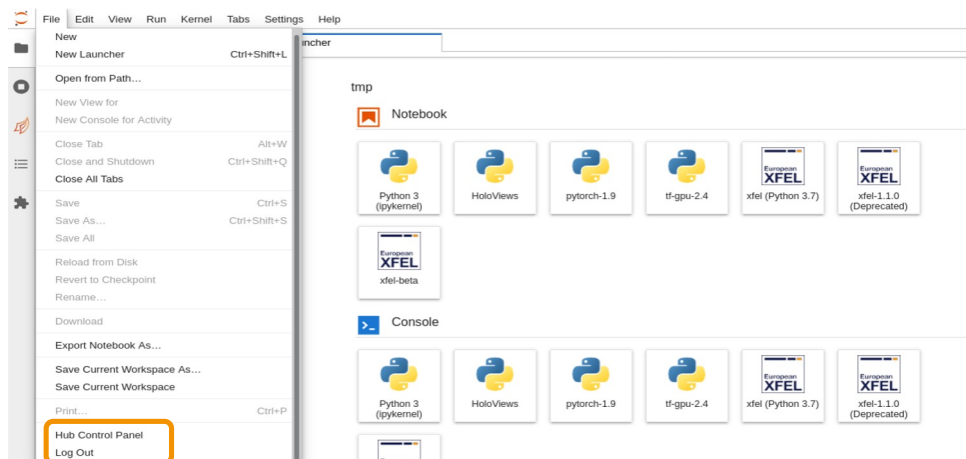
Python 3.9 Astra-Toolbox Bash European XFEL Geometry Holoviews

Simple 0 1 1 Mem: 197.43 MB Launcher

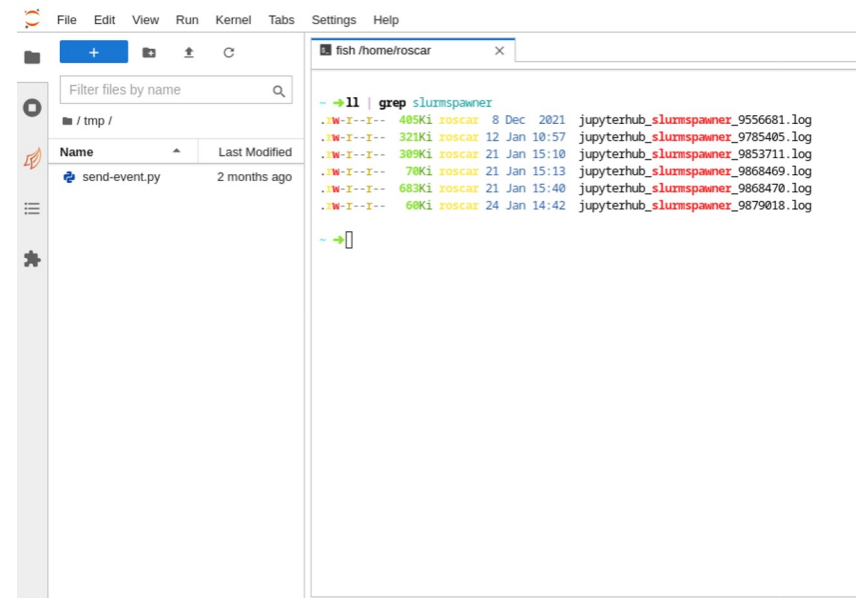
Annotation: *DA provided Python environment*

Accessing Maxwell – JupyterHub (III)

Logout and Troubleshooting



- Shut down your server on the **File > Hub Control Panel**
- Log Out option on menu **File**



Troubleshooting issues?

- The Jupyter server running in the background writes its logs to your home directory

Accessing Maxwell – SSH/Non-Graphical (I)

■ SSH to max-exfl-display.desy.de with your Campus account

▶ `ssh $USER@max-exfl-display.desy.de`

■ Accessible from outside the DESY/EuXFEL network

▶ **News: 2nd-factor authentication required from 30.01.2024!**

■ 'shared node' used by multiple users.

■ Suitable for:

▶ **Job submissions to SLURM**

• **SBATCH** or **SALLOC**

▶ Short/low intensity compute jobs

▶ Code compilations (if it's not using all cores)

▶ Programs requiring GUI/GPU acceleration

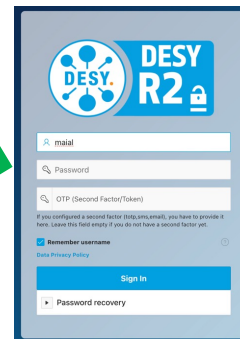
■ Not suitable for:

▶ Long-running/intense workloads like simulation or analysis. SLURM should be used instead.

```
[maial@exflqr47437 metadata_catalog % ssh maial@max-exfl-display.desy.de
Last login: Fri Jan 19 08:34:47 2024 from bl15-73-117.dsl.telepac.pt
[maial@max-exfl-display001 ~]$
[maial@max-exfl-display001 ~]$
[maial@max-exfl-display001 ~]$
```

2nd-factor authentication for users

- 2nd factor authentication mandatory on Maxwell from **30.01.2024** onwards
- Restriction applies to all scientific users
 - Documentation and instructions available at:
 - https://it.desy.de/services/mfa/naf_maxwell/index_eng.html (password protected)
 - Where can users setup their TOTP?
 - passwd.desy.de



More information on the “**News and data highlights**” session!

Agenda

Services

- Overview
- UPEX
- myMdC
- myLog
- DAMNIT
- Storage and Computing
- Maxwell
 - ▶ FastX (Remote Desktop)
 - ▶ JupyterHub
 - ▶ SSH

Remote Access to the Facility

- Experiment
- Data Analysis

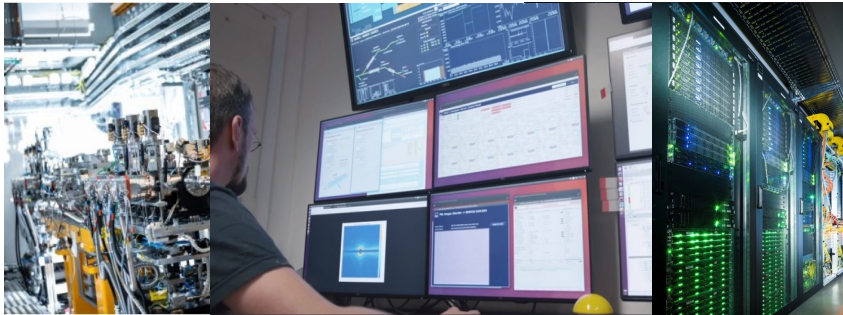
Allocating Resources

- SLURM and Jobs
 - ▶ Sbatch
 - ▶ Salloc

- Data
- Software
- EuXFEL software
- Documentation

Resources

Remote access to the facility



- Use cases:
- Data Analysis
- Experiments



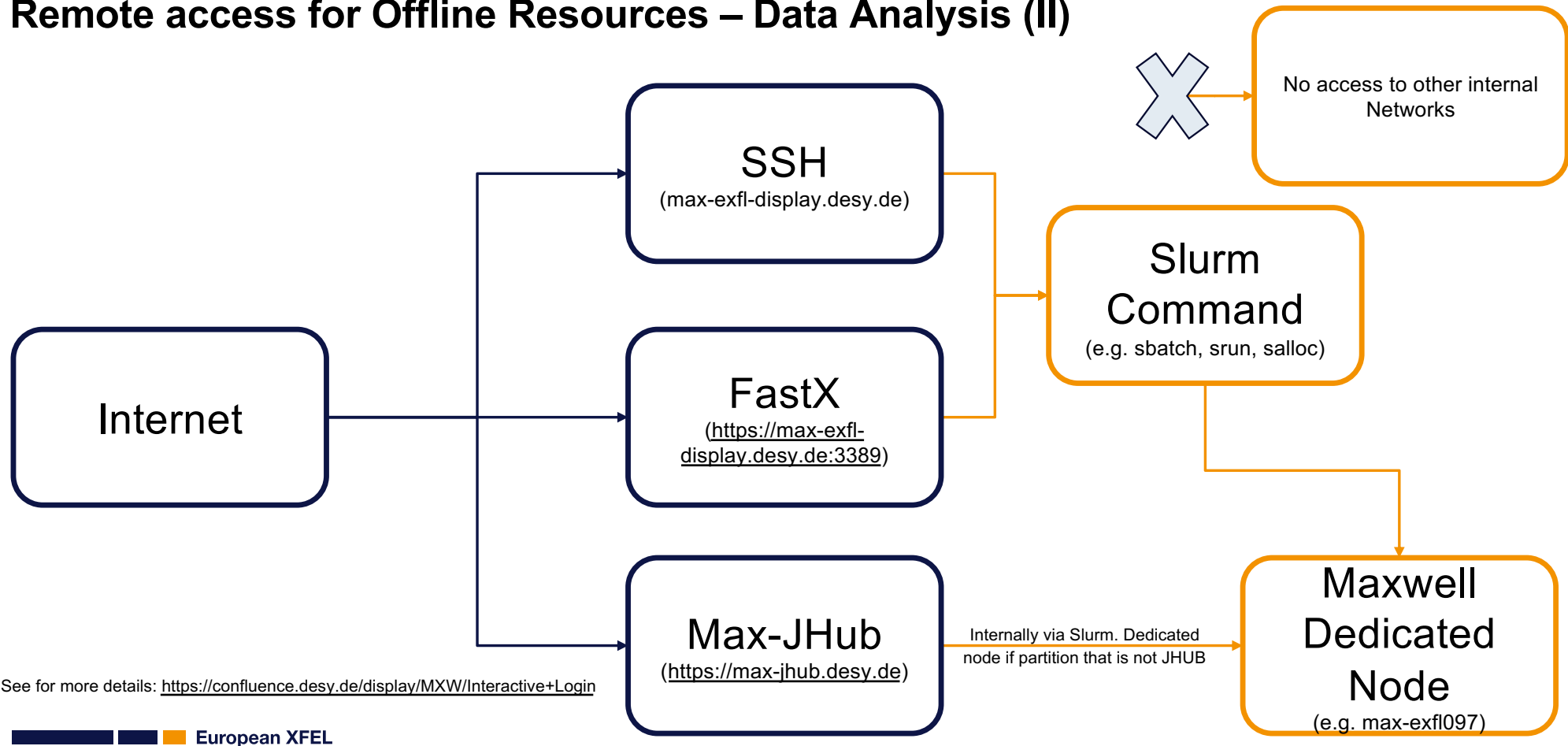
Remote access for Offline Resources – Data Analysis (I)

- Requirements to connect using FastX or SSH to Online Resources
 - **Membership of an experiment team**
 - ▶ Access is granted for periods of one year
 - From beamtime end until expected embargo period
 - ▶ Access is granted for new proposals after Arrival form submission

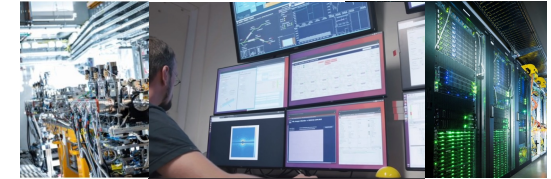
- Access is granted for:
 - Maxwell data analysis facility



Remote access for Offline Resources – Data Analysis (II)



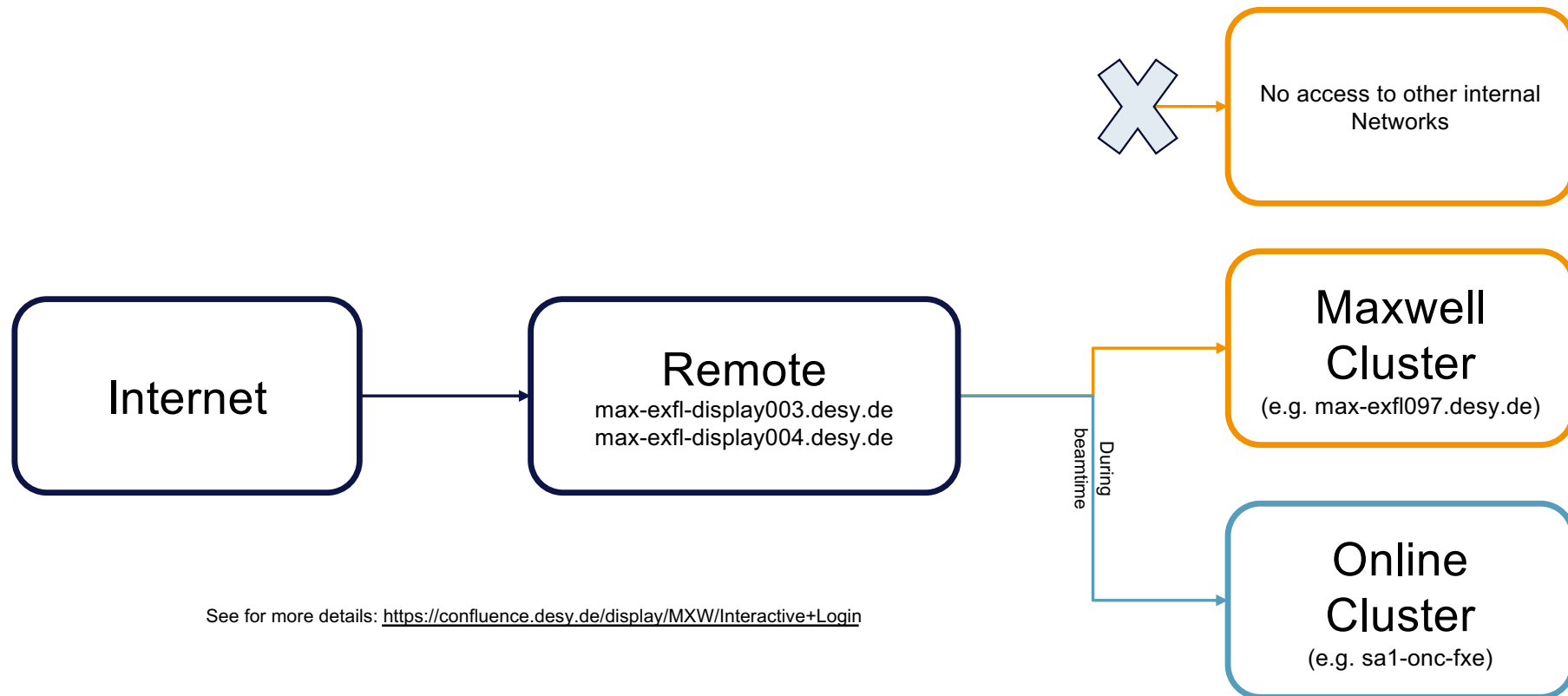
Remote access for Online Resources – experiment (I)



- Requirements to connect using FastX or SSH to Online Resources
 - **Membership of the active experiment team**
 - Login to beamtime dedicated gateway is **configured on a request**
 - ▶ from the PI/local contact or team member

- Access is granted for:
 - Maxwell data analysis facility
 - Specific services at the instrument Online Cluster
 - ▶ Location (Instrument) and time-related (beamtime)
 - ▶ Online analysis cluster with access to data
 - ▶ Karabo control system using a dedicated read-only Karabo GUI server
 - Alternatively, the proposal dedicated Zoom room can be used

Remote access for Online Resources – experiment (II)



See for more details: <https://confluence.desy.de/display/MXW/Interactive+Login>

Agenda

Services

- Overview
- UPEX
- myMdC
- myLog
- DAMNIT
- Storage and Computing
- Maxwell
 - ▶ FastX (Remote Desktop)
 - ▶ JupyterHub
 - ▶ SSH

Remote Access to the Facility

- Experiment
- Data Analysis

Allocating Resources

- SLURM and Jobs
 - ▶ Sbatch
 - ▶ Salloc

- Data
- Software
- EuXFEL software
- Documentation

Resources

Allocating Resources

Allocating Nodes and Submitting SLURM Jobs

SLURM and Jobs

- Maxwell uses the slurm scheduler
- When you connect via fastx/ssh you are on a shared node
- Shared nodes are not suitable for large loads
- You should use an allocated node for data analysis tasks
- Three main ways to run jobs:
 - sbatch - submit a job via a 'batch' script
 - salloc - allocate a node for interactive use
 - srun - submit a job via command line arguments

See for more details: <https://confluence.desy.de/display/MXW/Running+batch+jobs>
<https://rtd.xfel.eu/docs/data-analysis-user-documentation/en/latest/offline/#examples>

SLURM and Jobs - sbatch

- Submits a 'batch' (script) file to slurm
- Non-blocking - after submission you can close the SSH session or carry on with other tasks
- Slurm queues/allocates the requested resources
- Script is executed
- Once the script finishes or time expires, allocation is released
- Use cases:
 - Computationally intense work
 - Multi-node workflows
 - Short to long-length analysis - seconds to days
- Preferred way of running jobs on Maxwell

See for more details: <https://confluence.desy.de/display/MXW/Running+batch+jobs>
<https://rtd.xfel.eu/docs/data-analysis-user-documentation/en/latest/offline/#examples>

SLURM and Jobs - sbatch

Offline Analysis (Maxwell) Search Git

Examples

If you can define your job in a script, you can submit it like this:

```
sbatch -p upex -t 8:00:00 myscript.sh
```

- `-p` specifies the 'partition' to use. External users should use `upex`, while EuXFEL staff use `exfel`.
- `-t` specifies a time limit: `8:00:00` means 8 hours. If your job doesn't finish in this time, it will be killed. The default is 1 hour, and the maximum is 2 weeks.
- Your script should start with a 'shebang', a line like `#!/usr/bin/bash` pointing to the interpreter it should run in, e.g:

```
#!/usr/bin/bash

echo "Job started at $(date) on $(hostname)"

# To use the 'module' command, source this script first:
source /usr/share/Modules/init/bash
module load exfel exfel-python

python -c "print(9 * 6)"
```

SLURM and Jobs - salloc

- Allocates a node in the background which can have commands sent to it interactively
- Semi-blocking - new shell is spawned after allocation, exiting shell releases the allocation
- Slurm queues/allocates the requested node
- Slurm echos the node hostname
- Once the shell is exited or the time elapses, allocation is released
- Use cases:
 - Interactive development - executing srun on the allocation, or an interactive shell session
 - Medium-length analysis - minutes to hours
- Only recommended for short periods of interactive analysis/development
- salloc means that resources are blocked even when idle, wasting compute resources
- Only use when unavoidable - stick to srun/sbatch when possible

SLURM and Jobs - Misc

- If requested by local contact*, proposals can have a number of nodes reserved Maxwell:
 - Check your reservations with `scontrol show res`
 - Specify your reservation with `--reservation=upex_00PPPP` (where PPPP is the proposal number) or equivalent sbatch comment
 - Proposal reservation only available during beamtime
- Partitions available on Maxwell for slurm: <https://confluence.desy.de/display/MXW/Partitions>
- Check the list of useful commands: <https://confluence.desy.de/display/MXW/Useful+commands>
- Slurm quick-start guide: <https://slurm.schedmd.com/quickstart.html>

* soon per-proposal reservations will be specified in the DMP

Data

Where To Read Data From, What It Contains, and Where To Save Results

Data

- The Offline Analysis talk will go into more details on this
- EuXFEL data is saved as HDF5 files, with a specific internal and external structure
 - 'Internal' meaning the structure within the HDF5 files
 - 'External' meaning the name of the files and the directory they are in
 - Details in following talk and [User Documentation/Data Files](#)
- If you want to play around with the data and our tools before your beamtime to develop your data analysis look in `/gpfs/exfel/exp/XMPL/201750/`
- This contains open, example data, which is used in the tutorial notebooks for EXtra-data

Software

How to Use the Environment Module System to Load Software

Module System

- Environment Modules are used on Maxwell to allow modification of what software is available
- Lets you load different versions of software as required
- A large amount of software is provided by both DESY and EuXFEL:
 - DESY: <https://confluence.desy.de/display/IS/Alphabetical+List+of+Packages>
 - XFEL: <https://rtd.xfel.eu/docs/data-analysis-user-documentation/en/latest/software.html>
- Basics are:
 - `module avail` - list the available modules
 - `module load X` - load a module, and any of its dependencies
 - `module list` - list the loaded modules
 - `xwhich` - searches through modules for an executable (e.g. `xwhich python` lists modules that provide python)

Module System - EuXFEL Software

- To use the software provided by EuXFEL:
 - `module load xfel` - enables the EuXFEL group of modules
 - `module avail` - list all available modules, EuXFEL ones are under the category `/gpfs/exfel/sw/software/xfel_modules`
 - `xwhich` - now that EuXFEL modules are loaded, `xwhich` will search them for executables
- We provide an `xfel-python` module which loads a conda environment containing 'essential' packages
 - To improve reproducibility, a new version of this environment is created per-cycle
 - To load the current cycle environment run `module load xfel xfel-python`
 - At the start of a new cycle, the previous environment is no longer modified
 - Versions accessible with a forward slash, e.g. `xfel-python/202301`
- When using Max-JHub an `xfel` kernel will automatically be available which uses this environment

Module System - EuXFEL Software



EuXFEL Software - Docs, Specifications, and Versioning

- Specifications for software/environments provided by EuXFEL are stored in a repository on GitHub
 - URL: <https://github.com/European-XFEL/environments>

- Documentation page built from this repository
 - URL: <https://european-xfel.github.io/environments/>
 - Contains information on using our environments and creating your own
 - Shows what packages/versions are available

- Environments are defined per cycle:
 - Each cycle gets a new Conda environment
 - Major updates performed only when a new environment is created
 - Environment specifications and lock files stored in the git repo
 - Major changes and contents of environment available on docs pages (e.g. [202401 Environment](#))
 - [Zenodo record](#) available for environments to make citation easy - DOI [10.5281/zenodo.10548700](https://doi.org/10.5281/zenodo.10548700)

Agenda

Services

- Overview
- UPEX
- myMdC
- myLog
- DAMNIT
- Storage and Computing
- Maxwell
 - ▶ FastX (Remote Desktop)
 - ▶ JupyterHub
 - ▶ SSH

Remote Access to the Facility

- Experiment
- Data Analysis

Allocating Resources

- SLURM and Jobs
 - ▶ Sbatch
 - ▶ Salloc

- Data
- Software
- EuXFEL software
- Documentation

Resources

Resources

- EuXFEL and DESY, have large amounts of documentation on many key topics
- If anything is missing or if there's anything you'd like please contact us with suggestions
- EuXFEL Services - <https://in.xfel.eu/>
 - User Portal - <https://in.xfel.eu/upex/home/user>
 - MyMdC - <https://in.xfel.eu/metadata>
- EuXFEL Data Analysis Docs - <https://rtd.xfel.eu/docs/data-analysis-user-documentation/en/latest/>
 - EXtra-data - <https://extra-data.readthedocs.io/en/latest/>
 - EXtra-geom - <https://extra-geom.readthedocs.io/en/latest/>
 - EXtra-foam - <https://extra-foam.readthedocs.io/en/latest/>
 - EXtra-metro - <https://desy.de/~schmidtp/metropc-docs/>
- DESY Maxwell Docs - <https://confluence.desy.de/display/MXW/Documentation>
 - More in-depth intro (but from a few years ago, some details have changed) - https://indico.desy.de/event/20263/attachments/24956/31645/Using_Maxwell.pdf
 - DESY Computing How-Tos - <https://confluence.desy.de/display/IS/How-to+articles>

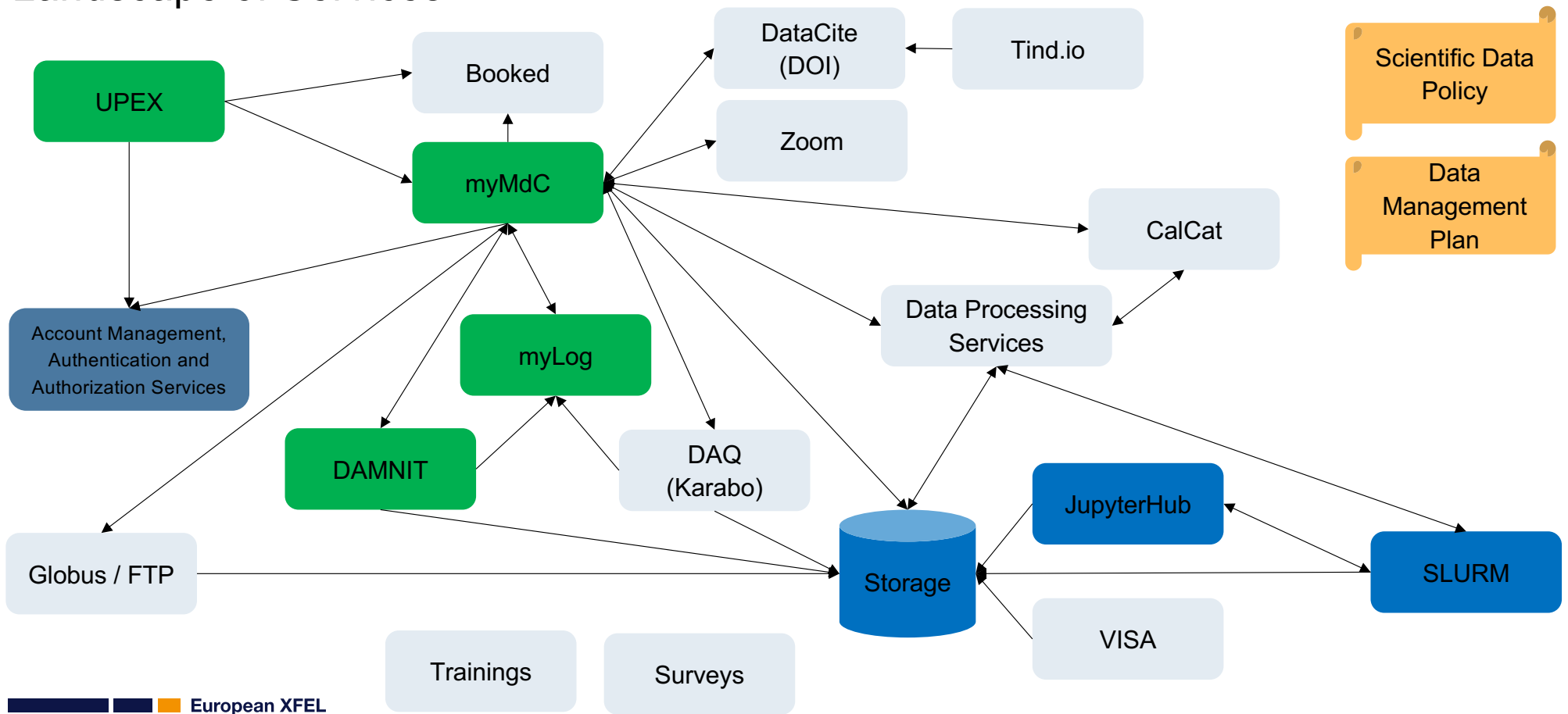
Thank you!

Questions?

Spare Slides

Overview (I)

Landscape of Services



myMdC Purpose (II)

- Main feature is to integrate and orchestrate different systems and services
 - Proposal authorisation, team administration
 - Proposals access to real data, data audit and reconciliation
 - Proposal Run types and Samples
 - Proposal workflow and notifications
 - Run management activities (quality assessment, calibration request, describe runs)
 - Repositories management
 - DataCite (DOI) management
 - Experimental techniques
 - Access data with Globus
 - Integration with myLog
 - Zoom for remote control room
 - Technical Integrations (e.g. RESTful)

myMdc Introduction (IV)

Proposal "General" information

Proposal no. 700000

Status: 2024-01-19 02:21:50 CET Runs Calibration Team Size 4.5 TB (0.25 TB) Files 628 Dark Calibration

Back Edit Clone + Add another Runs Beamtime status

General Public Information Runs Logbook Team Repositories **Beta!** Calibration Constants Publication History

General

id: 30
DOI: 10.22003/XFEL-EU-DATA-700000-00 DOI Published

Proposal Number: 700000
Name: p700000
Title: Example Data
URL: /metadata/proposals/30
Abstract: The European XFEL (EuXFEL) example data proposal contains experimental datasets from various original beam-times, currently covering the techniques of serial femtosecond crystallography (SFX), coherent diffraction imaging (single particle imaging, SPI), X-ray powder diffraction, small-angle X-ray scattering (SAXS) and X-ray photon correlation spectroscopy (XPCS).

Proposal Folder: p700000
Proposal Path: /gifs/xfel/draw/XMPL/201750/p700000
Logbook Identifier: ---
Auto assign Run Quality: To be evaluated manually
[Auto assign Run Quality](#)

Preferred data output: Raw [Preferred data output](#)

Proposal Last Run: 128
Instrument: Example Data
Instrument Cycle: XMPL: 201750 (starts: 2017-01-01 00:00:00 +0100, ends: 2019-12-31 23:59:59 +0100)
Principal Investigator: Luca Gelliso
Main Proposer: Fabio Dall'Antonia
Local Contact: ---
Start date: 2017-11-08 00:00:00 +0100
End date: 2017-12-31 23:59:59 +0100
Expected end of Embargo: 2018-01-01 00:00:00 +0100
Open data since: 2023-10-27 11:24:32 +0200
open data?: Yes
No. delivered shifts: 0

Beamtime: 2017-11-08 00:00:00 +0100 - 2017-12-31 23:59:59 +0100
[Zoom meeting](#) [Zoom to proposal team](#)
Cloneable Status: Simple Clone (Team) [Cloneable Status](#)
Beamtime status: Ready [Beamtime status](#)

Start Beamtime date: ---
End Beamtime date: ---
Available: Yes
Proposal system: Example
Description: The European XFEL (EuXFEL) example data proposal contains experimental datasets from various original beam-times, currently covering the techniques of serial femtosecond crystallography (SFX), coherent diffraction imaging (single particle imaging, SPI), X-ray powder diffraction, small-angle X-ray scattering (SAXS) and X-ray photon correlation spectroscopy (XPCS).

Created by: System
Creation date: 2017-11-08 13:49:30 +0100
Modified by: Fabio Dall'Antonia
Last modification date: 2023-03-17 10:47:33 +0100

All Proposal Run types [+ New](#)

- {e0001}: AGIPD
- {e0002}: Calibration - Dark HG
- {e0003}: Calibration - Dark LG
- {e0004}: Calibration - Dark MG
- {e0005}: Configuration Tests
- {e0006}: Dark
- {e0007}: Diffraction
- {e0008}: Diffraction data
- {e0009}: General
- {e0010}: Sample
- {e0011}: SAXS 500kHz // no pump laser
- {e0012}: scattering
- {e0013}: SFX Jetting
- {e0014}: Single Particle Diffraction
- {e0015}: Test DAQ
- {e0016}: XGM

All Proposal Samples

- Water
- Lysozyme (201804 small crystals)
- Lithium Titanate
- No Sample
- Lysozyme
- Silica 50nm
- Vycor
- Sucrose Solution 5% w/v
- Potassium hexacyanoferrate(II) trihydrate
- Cu foil
- Xenon
- 2-Co8_p14_8fold - 30nm Pt cap
- 1-Co10_Pt_6fold
- Ni-20 MLs - b
- Ni75-11 MLs-b
- Bolometer

Proposal Technique [Link](#)

- coherent diffraction imaging
- serial femtosecond crystallography
- small angle x-ray scattering
- x-ray photon correlation spectroscopy
- x-ray powder diffraction

Proposal Run Types

Back Edit Clone + Add another Beamtime status

Proposal Tabs

Proposal summary numbers

Link to Globus

Proposal DOI

Proposal Information

Important information:

- Proposal path
- Relevant Dates
- Data Output configurations

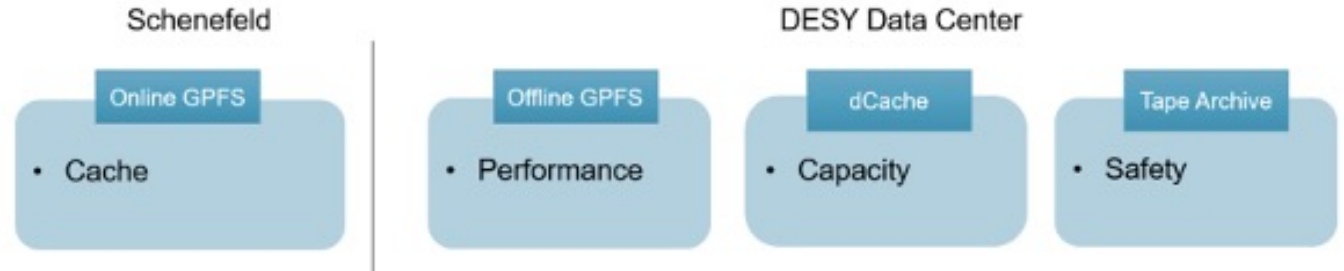
Zoom to proposal team

Proposal status

Proposal associated techniques

Proposal associated samples

Storage and computing



- Data is collected at the instrument into the ‘Online Storage’
 - Data on the ‘Online Storage’ is only accessible via the ‘Online Cluster’
 - Several nodes are dedicated for an active experiment
 - Used for online data analysis/preview during/immediately after data acquisition

- If data is assessed as ‘good’ or ‘unclear’ it is moved to the the ‘Offline Storage’ accessible from Maxwell
 - Used for offline calibration and data analysis
 - Hundreds of nodes within the Maxwell cluster

- High capacity mass storage (dCache) accessible from Maxwell
 - Mid term storage for the raw data
 - Interface to the tape archive

Proposal no. 004316

Status: 2024-01-25 00:06:06 CET Runs: 159 Calibrations: 0 Team: 03 Size: 041.82 GiB (889.15 GB) Files: 6870 Dark Calibrations: 0

Back Edit Clone + Add another Runs Beamline status

[General](#) [Public Information](#) [Runs](#) [Logbook](#) [Team](#) [Repositories Beta](#) [Calibration Constants](#) [Publication](#) [History](#)

Repositories

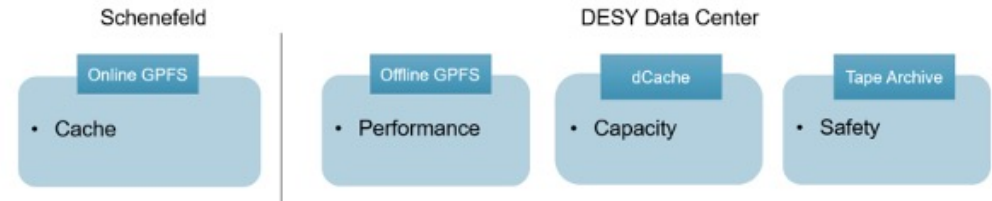
Request the Copy of Raw Data to available Remote repositories

No remote repositories are available.

Raw data availability per run and repository

Run Number (beam)	Data Group	Run Quality	XFEL GPFS Raw online for SASE 1	XFEL GPFS online for SASE 2	XFEL GPFS online for SASE 3	XFEL GPFS online Raw data in DESY-CC	dCache Raw Data in DESY-CC
0159	raw_e0001_n0159	Good					
0158	raw_e0001_n0158	Good					
0157	raw_e0001_n0157	Good					

Storage and computing



- Data is collected at the instrument into the 'Online Storage'
 - Data on the 'Online Storage' is only accessible via the 'Online Cluster'
 - Several nodes are dedicated for an active experiment
 - Used for online data analysis/preview during/immediately after data acquisition

- If data is assessed as 'good' or 'unclear' it is moved to the the 'Offline Storage' accessible from Maxwell
 - Used for offline calibration and data analysis
 - Hundreds of nodes within the Maxwell cluster

- High capacity mass storage (dCache) accessible from Maxwell
 - Mid term storage for the raw data
 - Interface to the tape archive

Node	IP	System	Storage	Access	Access	Access	Access	Access
maxwell01	193.50.130.1	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell02	193.50.130.2	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell03	193.50.130.3	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell04	193.50.130.4	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell05	193.50.130.5	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell06	193.50.130.6	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell07	193.50.130.7	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell08	193.50.130.8	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell09	193.50.130.9	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell10	193.50.130.10	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell11	193.50.130.11	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell12	193.50.130.12	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell13	193.50.130.13	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell14	193.50.130.14	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell15	193.50.130.15	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell16	193.50.130.16	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell17	193.50.130.17	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell18	193.50.130.18	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell19	193.50.130.19	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell20	193.50.130.20	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell21	193.50.130.21	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell22	193.50.130.22	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell23	193.50.130.23	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell24	193.50.130.24	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell25	193.50.130.25	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell26	193.50.130.26	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell27	193.50.130.27	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell28	193.50.130.28	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell29	193.50.130.29	CentOS	100TB	100TB	100TB	100TB	100TB	100TB
maxwell30	193.50.130.30	CentOS	100TB	100TB	100TB	100TB	100TB	100TB

What is Maxwell?

The Maxwell Cluster is the computing platform at DESY (Hamburg) for Photon Science data analysis, GPU accelerated computations (AI), High Performance Computing and scientific computing in general. The cluster serves myriads of applications and scientific fields.

Compute Hardware		Infiniband Hardware		Storage	
CPU+GPU nodes	798	root switches	6	GPFS exfel	~40 PB
Total number of cores with hyperthreading	61696	top switches	12	GPFS petra3	~20 PB
Total number of PHYSICAL cores	30898	leaf switches	42	BeeGFS desy	1.5 PB
Theoretical CPU peak performance	1074 TFlops	IB cables (#)	>1432	BeeGFS cssb	3.2 PB
Total RAM	420 TB	IB cables (length)	>7.6km		
GPU nodes	180				
Total number of GPUs	379				
Theoretical GPU peak performance	2330 TFlops				
Total peak performance	3404 TFlops ¹				