LEAPS-IT Working Group 3 Workshop on "IT Solutions to address Operational Challenges during Covid-19"

Workflow at DESY Photon Science and selected actions taken at PETRA III to the pandemic

A. Rothkirch

DESY Photon Science, FS-EC

and many colleagues from FS-EC, DESY Central IT and the beamlines

HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

DESY.

Outline/brief overview

- **Experiment/data life cycle**

- **Digital User Office**

- **Managed storage system - Start/Stop a Beamtime**

- **gamma-portal**

- **Data access and analysis environment**

- **Response/changes to pandemic (selection)**

- **Remote control at P11**

PETRA III as of May 6
- *No new users since March 13th 2020*
- *Stand-by of PETRA III at noon on March 20th, 2020 with ability to restart operation within ~24 h for fast-track access Corona-relevant experiments*
- *Corona-relevant experiments done at P03, P10, P11, P13, P14 since March 30*
- PETRA III operation resumed on May 04 in reduced mode (german users and mail-in experiments)

HELMHOLTZ RESEARCH FOR GRAND CHALLENGES

DESY.

Experiment/data life cycle

Apply for an experiment

Experiment preparation

- Integrate brought in equipment

Start of the experiment

- access to the storage space
- access for functional account & users
- configure exports and endpoints

Data acquisition

- 0D, 1D, 2D, 3D data
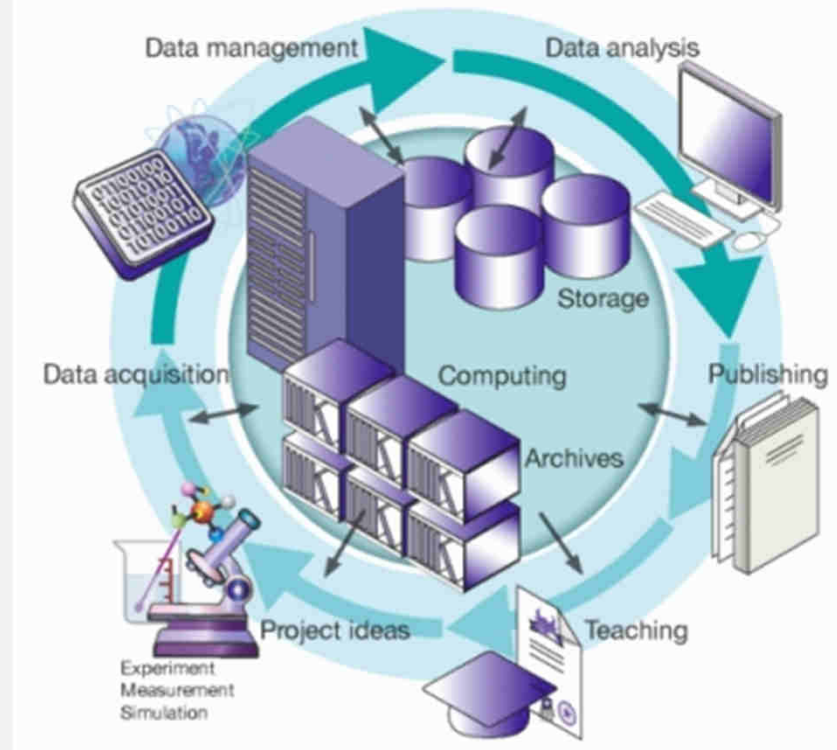- variety of formats and sizes

Activities during the experiment

End of the experiment

- Data not accessible for next user group

Data access past the experiment
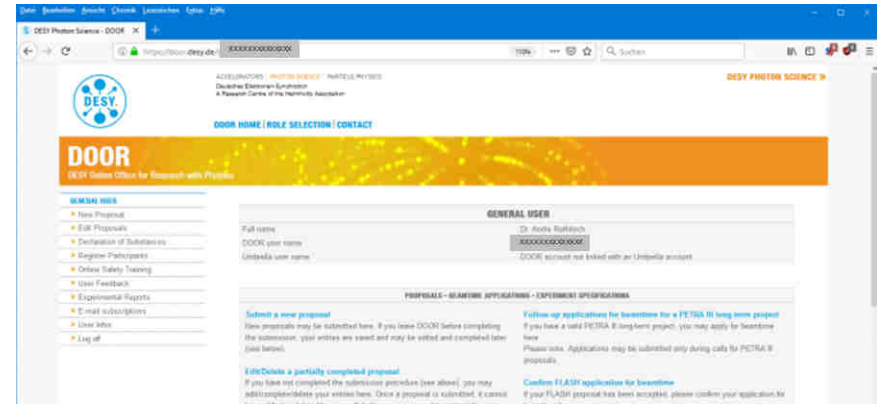
- Offline analysis on- and off-site

Data archival

Digital User Office

J.P. Kurz (EC), D. Unger (PS), U. Lindemann (IT)

The Digital User Office DOOR facilitates

- Proposal submission

- Peer reviews

- Beamtime scheduling

- Declaration of substances/
  List of participants

- Miscellaneous administrative
  tasks.

- DOOR is based on DUO (PSI).
  It is a common activity between
  the FS department and central
  IT

- Generation of unique ID per BT
  "**Beamtime Application ID**"

The DOOR user portal
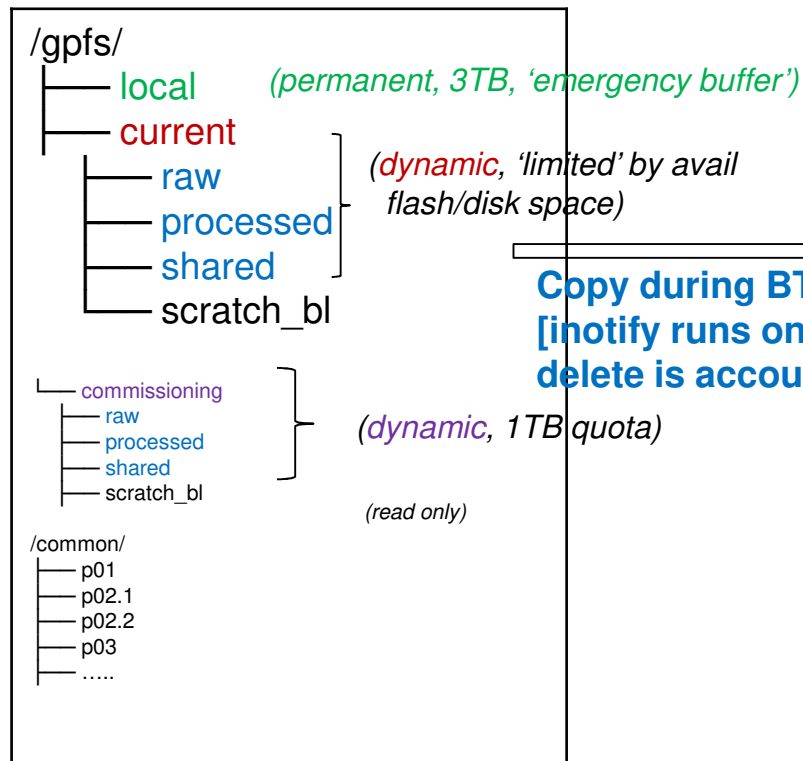
Start/Stop a Beamtime
Predefined directory toplevel structure and rules/constraints

[psana001.desy.de]$ **startBeamtime** *--beamtimeId <beamtimeID> --beamline <beamline>*

Temporary storage (**on IPs**)

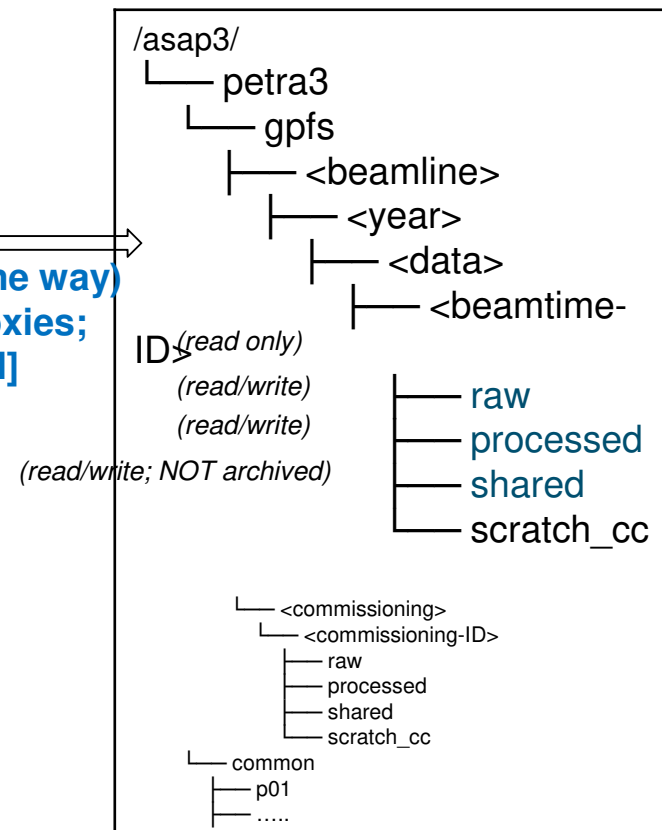**Limited to Beamline & beamtime ("BL-FS")**

GPFS with NFS 3 / SMB

```
/gpfs/
├── local          (permanent, 3TB, 'emergency buffer')
├── current         (dynamic, 'limited' by avail
│   ├── raw          flash/disk space)
│   ├── processed
│   ├── shared
│   └── scratch_bl

└── commissioning    (dynamic, 1TB quota)
    ├── raw
    ├── processed
    ├── shared
    └── scratch_bl

                     (read only)
/common/
├── p01
├── p02.1
├── p02.2
├── p03
└── .....
```

**Copy during BT (one way)
[inotify runs on proxies;
delete is accounted]**

Permanent storage (**on ACLs**)

"GPFS Core file system"

*GPFS - native on Analysis nodes
      - or by NFS 4 mapping / SMB (office)*

```
/asap3/
└── petra3
    └── gpfs
        ├── <beamline>
        │   └── <year>
        │       └── <data>
        │           └── <beamtime-
ID>  (read only)
         (read/write)
         (read/write)

   (read/write; NOT archived)
                            ├── raw
                            ├── processed
                            ├── shared
                            └── scratch_cc

        └── <commissioning>
            └── <commissioning-ID>
                ├── raw
                ├── processed
                ├── shared
                └── scratch_cc
    └── common
        ├── p01
        └── .....
```
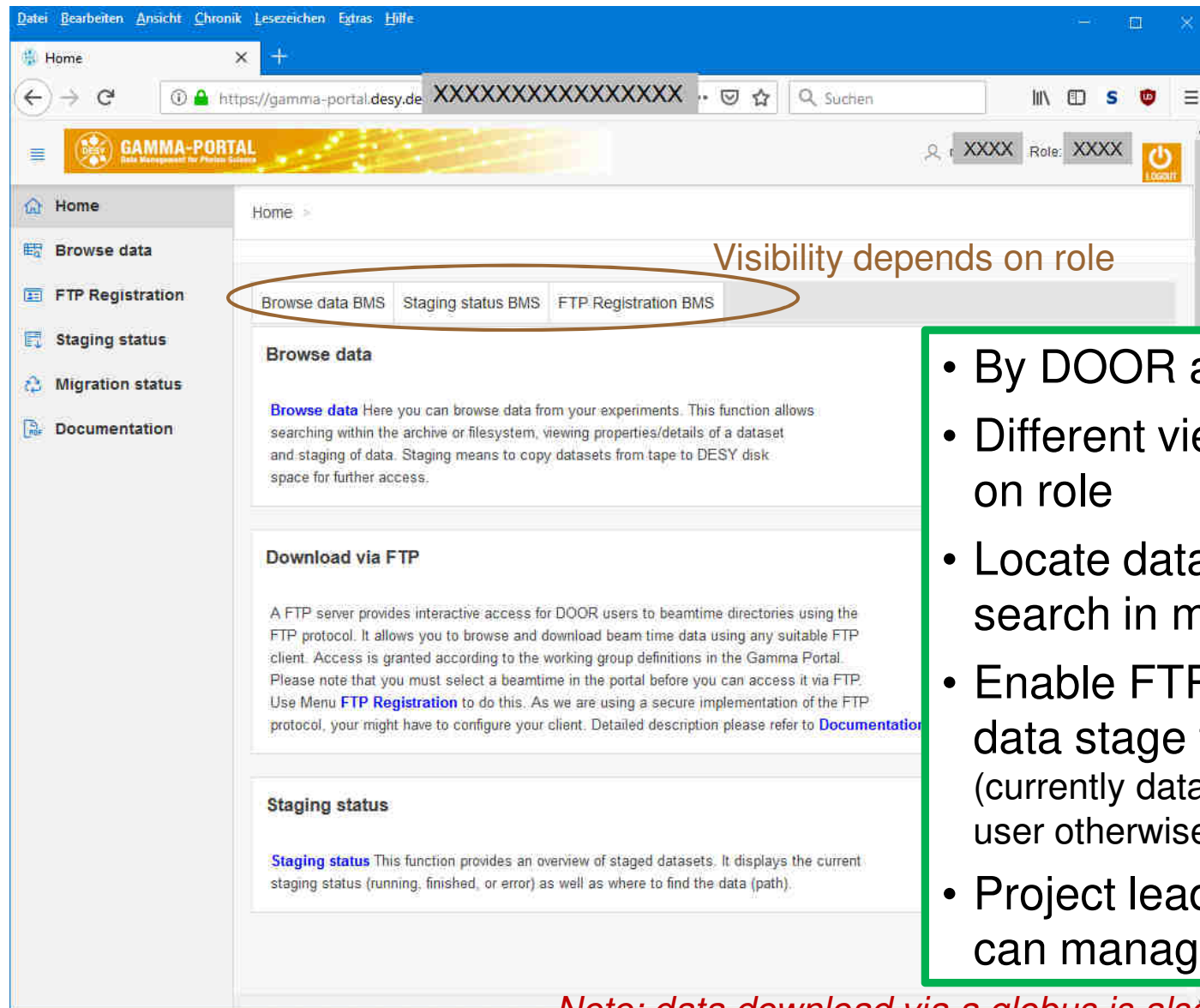
WGS

dCache

Portal

Gamma-Portal (htts://gamma-portal.desy.de)
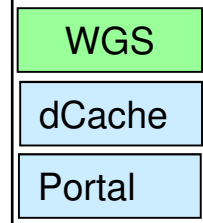
- By DOOR account
- Different views depending on role
- Locate data by keyword search in metadata
- Enable FTP download & data stage from dCache (currently data before 2015 by user otherwise via FS-EC)
- Project leader & applicant can manage ACLs

Note: data download via a globus is also possible and implies scientific account

Data access and analysis environment
(located @ Computer Center)

- Invention of Scientific accounts (i.e. DESY accounts for external users ('external' ≠ industry/commercial) with own namespace 'psx'

- Provision of interactive resources max-fsc/max-fsg

- Creation of specific batch resources for PS managed by SLURM
  - Slurm partition ps (inhouse)
  - Slurm partition psx (external [non-commercial] users or use cases)

- Invention of display-servers for processing involving GUI

- Remote access (firewall/tunnel or Web-Browser)

Batch (Maxwell cluster)
- *Exclusive resources usage for jobs managed by SLURM*
- *Efficient resource usage (batch queue, resource definitions, optimize costs etc.)*
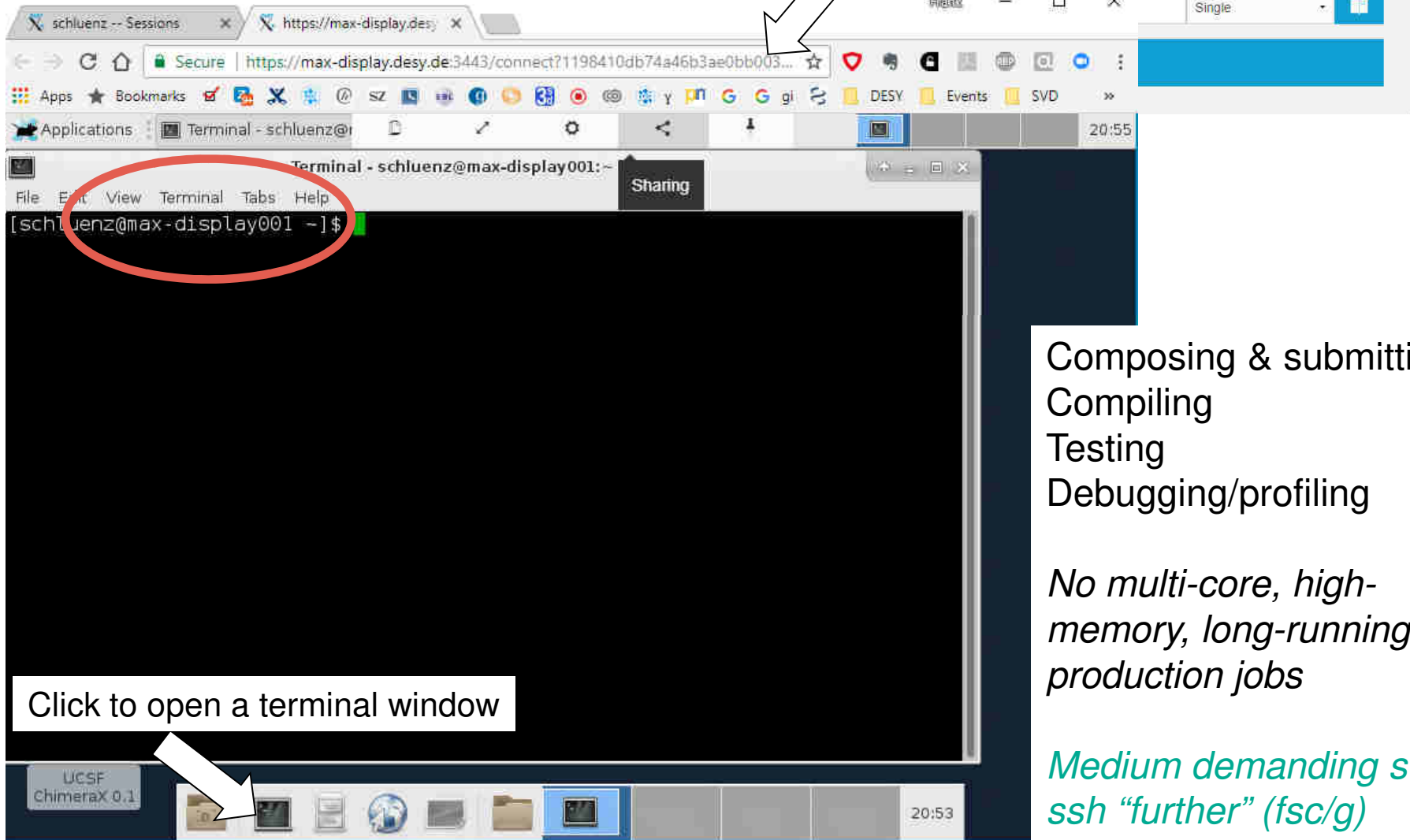- *Homogeneous/common environment for 'all groups', e.g. rules, IB, GPFS …*

Getting Started

Submission / Login nodes – max-display [from everywhere by Browser]

Getting Started

Submission / Login nodes

Click to open a terminal window
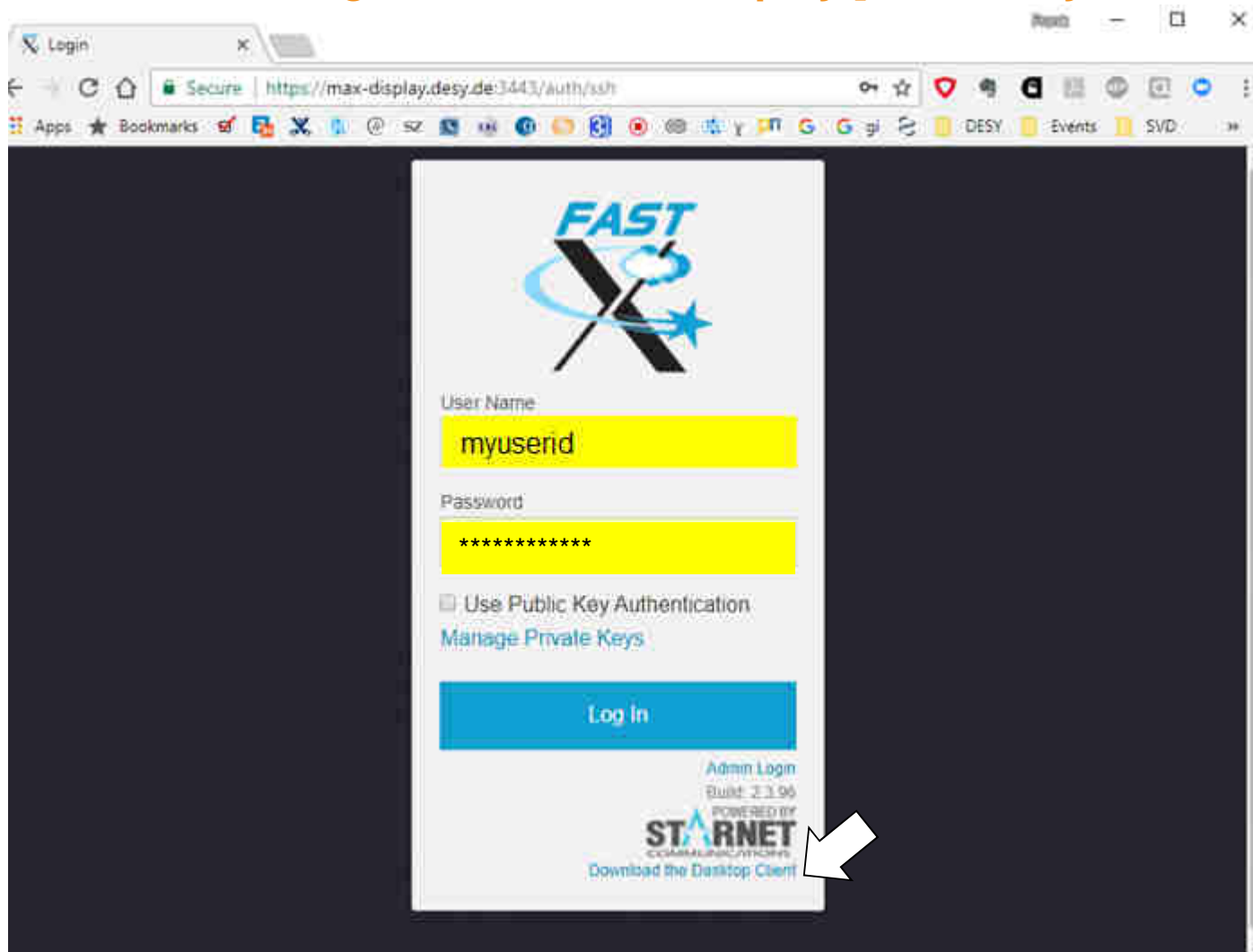
Composing & submitting
Compiling
Testing
Debugging/profiling

No multi-core, high-memory, long-running production jobs

Medium demanding stuff ssh "further" (fsc/g)

Response/changes to pandemic (selection)

- Pushing implementation of remote control at P11, handover to/allow (selected) users
- Mail-in offered by beamlines which never did it before (9 beamlines in total)
- Likely no increase in unattended operation of experiments (depending on what's meant, it's already daily business at the BIs [and some have to run supervised/attended due to e.g. safety or instrument complexity])

Some future plans in response to COVID 19

- Improvements / further development of remote control & communication at P11
- Implementation/adaption of remote concept at other beamlines, P02.1 next/to come soon
- Increase amount of Mail-in beamtimes (more an „orga"-challenge)

Remote access/control P11

Prerequisites

What was wanted and what was needed

P11, FS-EC and IT

- Beamline control through one user exclusively

- Do not expose any functionality which is not needed

- Eliminate abuse

- Ability to monitor the user actions

- Take back control at any time

- No special hardware, software or OS dependencies for the users

- Possibility to forward control to between legitimized users

D5 and User Office

- Authentication against a personalized account

- Make remote access dependent on safety instructions, Door account, etc.

- Introduce additional safety instructions for remote users

- Log remote access to database for statistical purposes

- Remote access has to be requested in DOOR
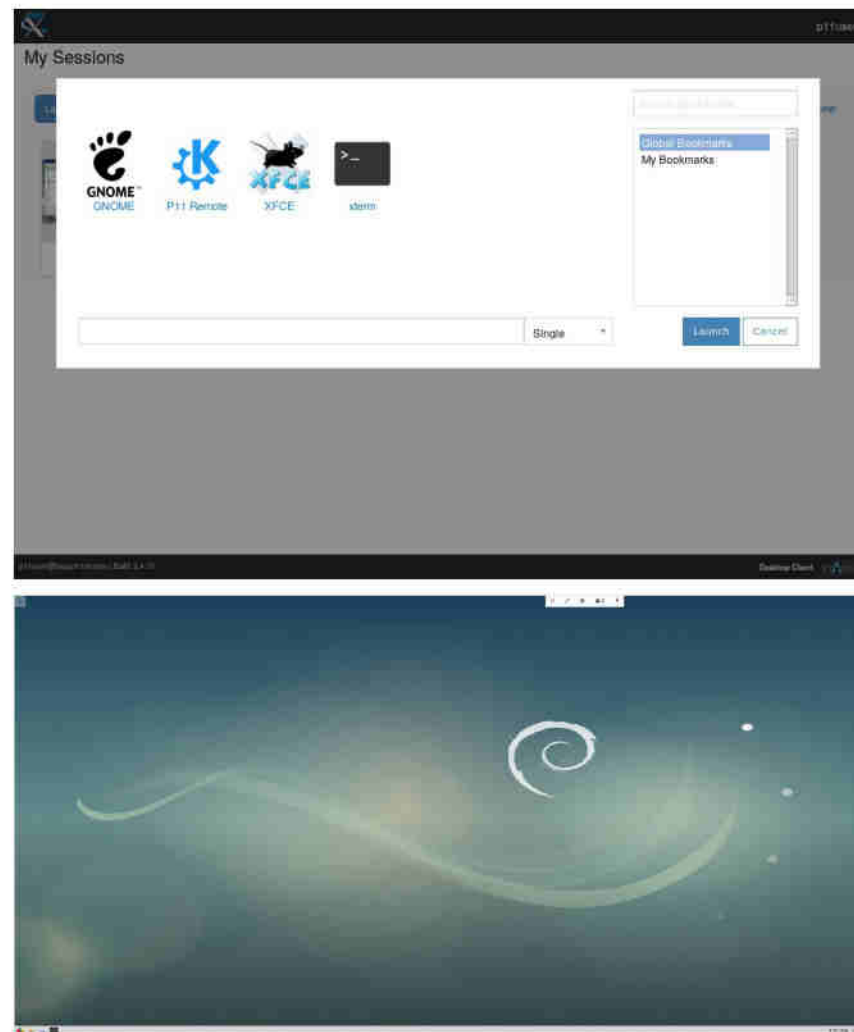
Technical Implementation

Make it working

FastX

- Commercial software, license available

- Already in use for Maxwell access

- X server in a browser

- Sessions can be shared

- Running on a dedicated host

Tailored X session

- Kiosk mode KDE

- Experiment control GUI

- Beam position monitor / feedback

- Browser (results, cameras, wiki)

- No terminal!

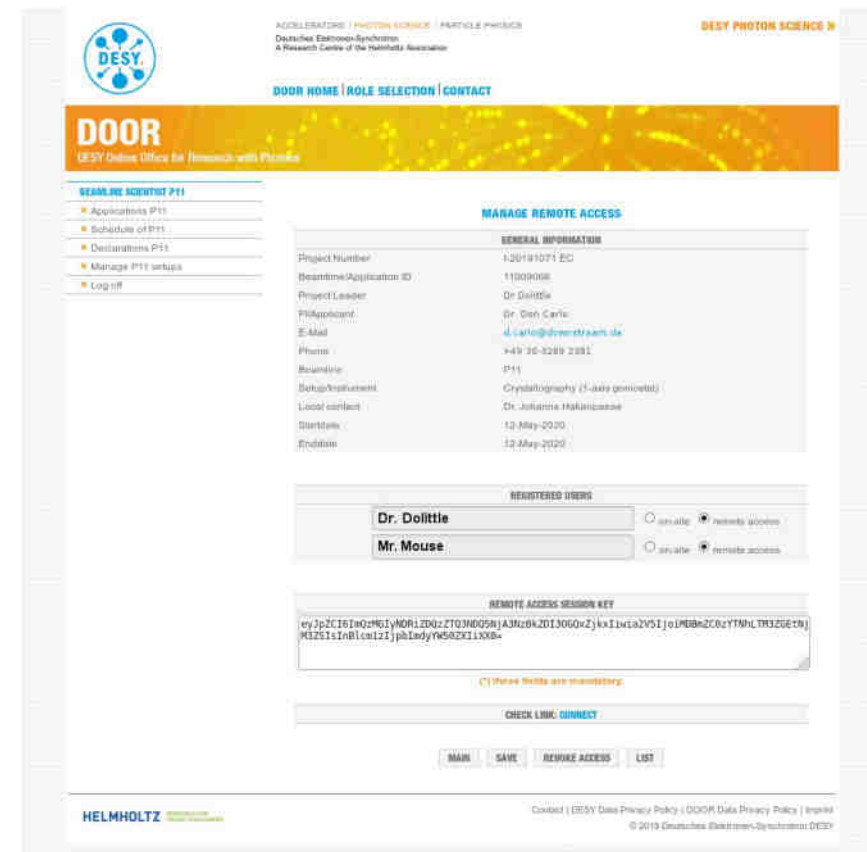# Formal Arrangements

Make it legal

Safety Requirements

- Currently as onsite, but shutter permission
- One user with onsite experience
- Additional online instructions in the near future

Door

- Grant online access per beamtime and user
- Save FastX session key
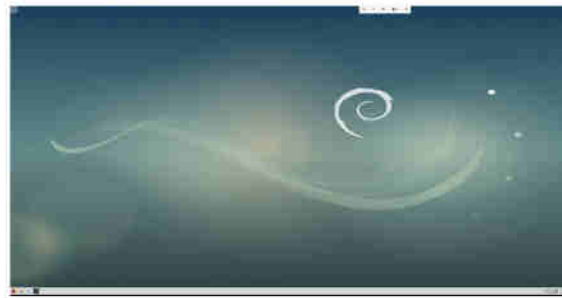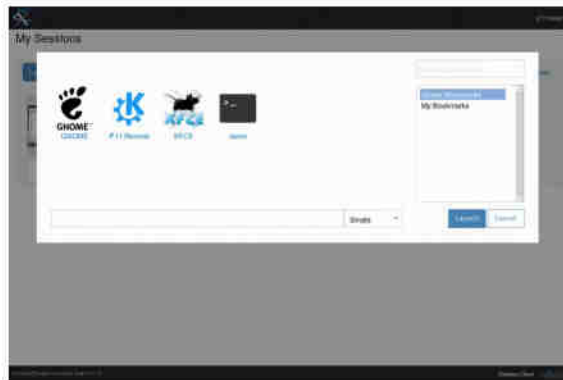
Proxy Node

- Authenticate against Door

Handling

How things work from the staff perspective

Start

- Prepare beamline, load samples, etc.

- Open beamtime

- Open FastX session and share it

- Save shared session key in door

- Contact users by video chat or phone

Stop

- Close FastX session
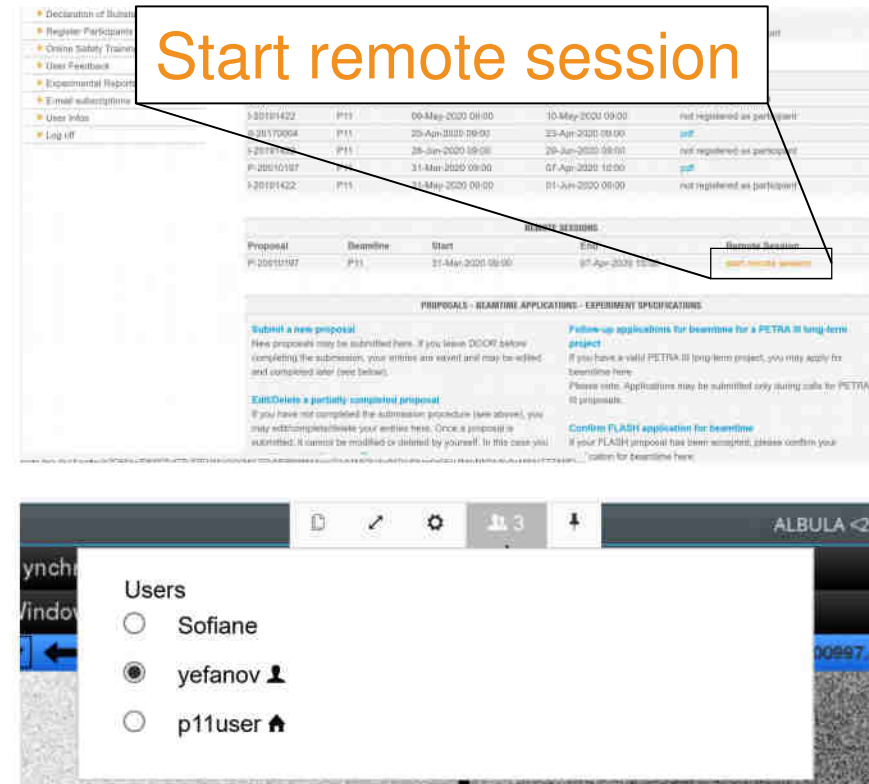
- Unload samples

- ...

Handling

How things work from the users perspective

Start

- Log in to Door, it presents you a link to the remote session

- Take control
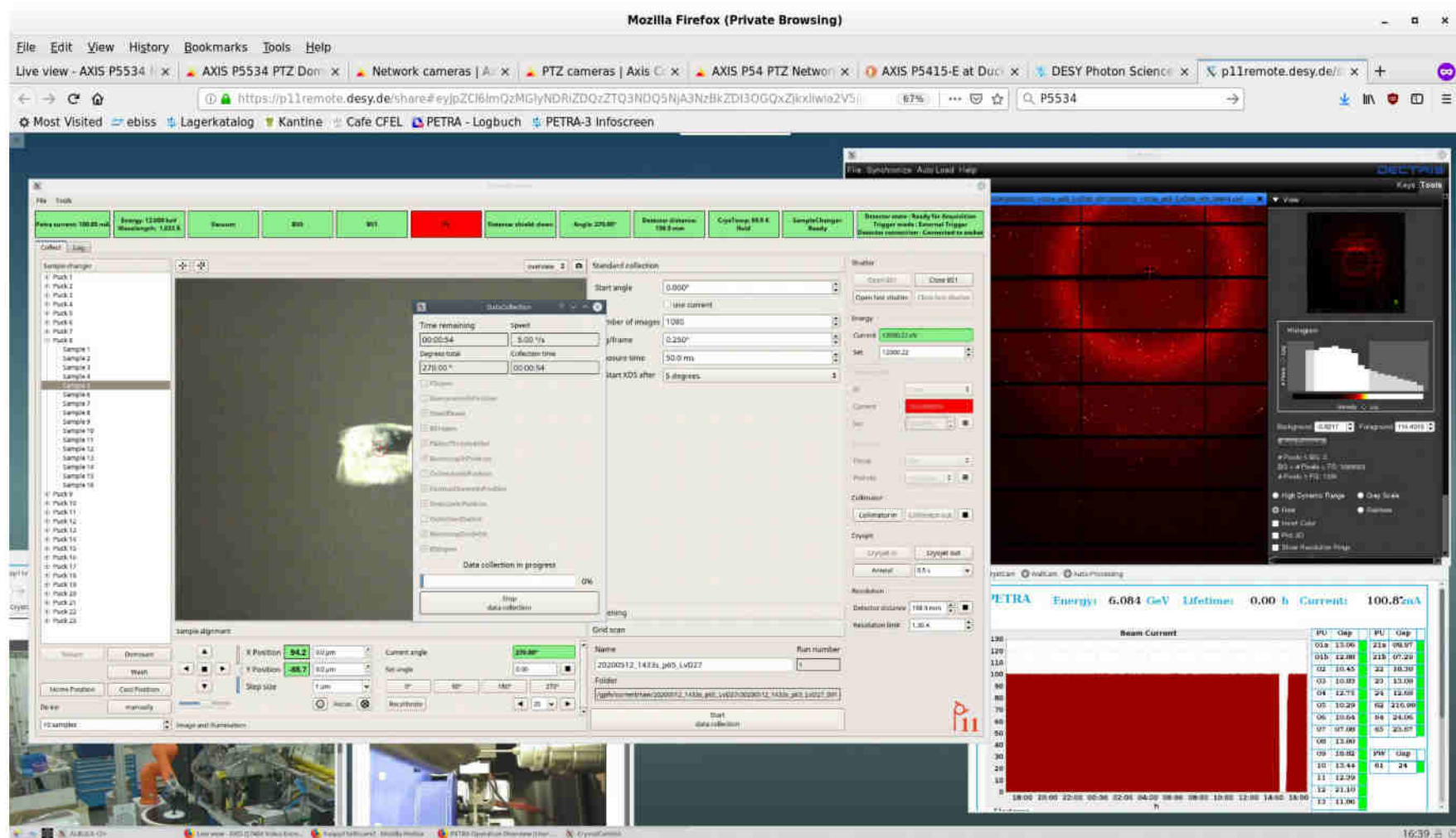
- Adjust screen size

- Go on like normal

Stop

- Log out from Door

- Close browser window
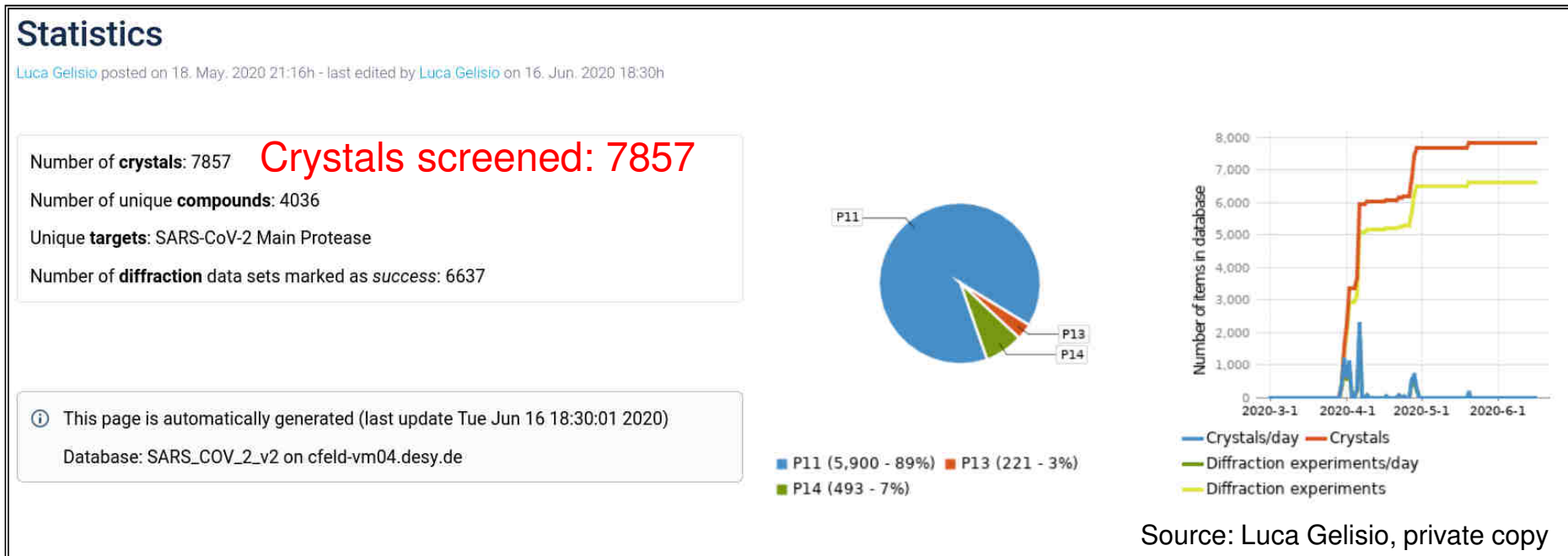

Start remote session

Handling

What it looks like

First Experiences – *it works :-)*

- Stable – first session was running for more than a week in large collaboration
Compound target screening of the Covid-19 main protease, A. Meents et al. (DESY, CFEL, Uni HH [CUI], EMBL, TU HH, CIT, SLAC, DLS, ….)

Source: Luca Gelisio, private copy

- Responsiveness depends on the network connection
- Screen size is an issue
- Communication is another
- Handover of control between users during the experiment is a big advantage
- Remove possibilities to break the interlock from GUI

Environment

Other things to think about

- Sample shipping, delivery, handling, loading, ...

- Coordinate which sample is loaded where

- How to deal with new users, how to train them in remote control?

- There might be problems coming from the connection or the user computer

- Communication between local contact and controlling user

- Make people locally aware of ongoing remote activities

Backup slide

Maxwell cluster batch resource (by IT)

What is the Maxwell cluster?

- **A large number of powerful computers (named max-<something>)**
  - All connected through a fast low latency network (56Gb EDR/FDR)
  - All connected to Petra3 GPFS storage (and CFEL, EXFEL, CSSB storage)
  - All connected to dCache ("on demand")
  - All equipped with 256GB up to 1.5TB of memory per node
  - Quite a number of nodes with 1-4 Nvidia P100 GPUs
  - Lots of software pre-installed

- **Main purpose**
  - High Performance Computing
  - Offline Data Analysis
  - Simulations of all kind
  - Remote Visualization
  - Any application which can make use of the special features of Maxwell!

E.g. Ansys, Comsol, Fdmnes (MPI version), Matlab, OpenFOAM, Orca, Quantum espresso, Tensorflow, Xds, Xmimsim, XRT .....

E.g. Conuss less well suited (single threaded/few mem. ....)

- **All jobs are scheduled by the SLURM scheduler (via submission hosts)!**
  - Usually jobs don't have to wait very long
  - But it depends on the jobs requirements
  - and there is no VIP fast lane …

- max-p3a* aka max-fsc, max-fsg, desy-ps-cpu, desy-ps-gpu are NOT part of SLURM

Overview