Operational challenges during COVID19

Hans Fangohr

Data Analysis Group

Data Unit

European XFEL

Germany

LEAPS-IT Workshop

Lund (Zoom)

18 June 2020

Outline

■ Facility status

■ Remote data analysis
    ■ offline
    ■ online

■ Remote control

■ Citizen science project: Tracking COVID19 infections and deaths in your region (OSCOVIDA)

■ Communication: instant messaging

■ Disclaimer: views my own

European XFEL

COVID-19 reduced operation

■ European XFEL went into reduced operation with minimal staff on site by on 16 March

■ Extended use of home office
  ■ Software work
  ■ Data analysis
  ■ Infrastructure development
  ■ Publications

■ Restart with < 30% staff on site started 4 May

■ No **regular** user operation this year
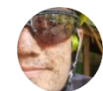
European XFEL

Hans Fangohr
@ProfCompMod

#homeoffice - Data Analysis team at @EuropeanXFEL working from home to #flattenthecurve . A mixture of instant chat, email, video conferencing, GitHub and Gitlab issues help staying in touch. Seen one cat on screen during video meetings so far. @AcademicChatter #coronavirus
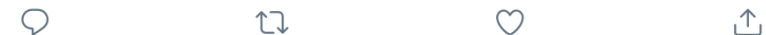

10:24 PM · Mar 17, 2020 · Tweetbot for Mac

1 Retweet   15 Likes

Alun 🚜🌾🏄⛵💻🏴󠁧󠁢󠁷󠁬󠁳󠁿 @AlunAshton · Mar 18
Replying to @ProfCompMod @EuropeanXFEL and @AcademicChatter
The science IT section @psich_en are 'virtually' the same 😹, but the cat invasion started Monday!

Should review the policy of pets in the workplace! (Said the cat 😉)

On site

- COVID-19 Safety Guidelines and training for work on site
    - Increased hygiene and distancing measures in place
    - Occupancy restrictions for all European XFEL rooms
        - ▶ Only 1 person per office
        - ▶ Maximum occupancy given at door
        - ▶ Includes experiment hutches

- Staff members only allowed on site with permission of Group Leader
    - Home office preferred if possible

European XFEL

XFEL.EU SR-2020-002-01.0

SAFETY REGULATION

Hygiene Guidelines for Safe Operation During the COVID-19 Epidemic

2 Persons max.

European XFEL

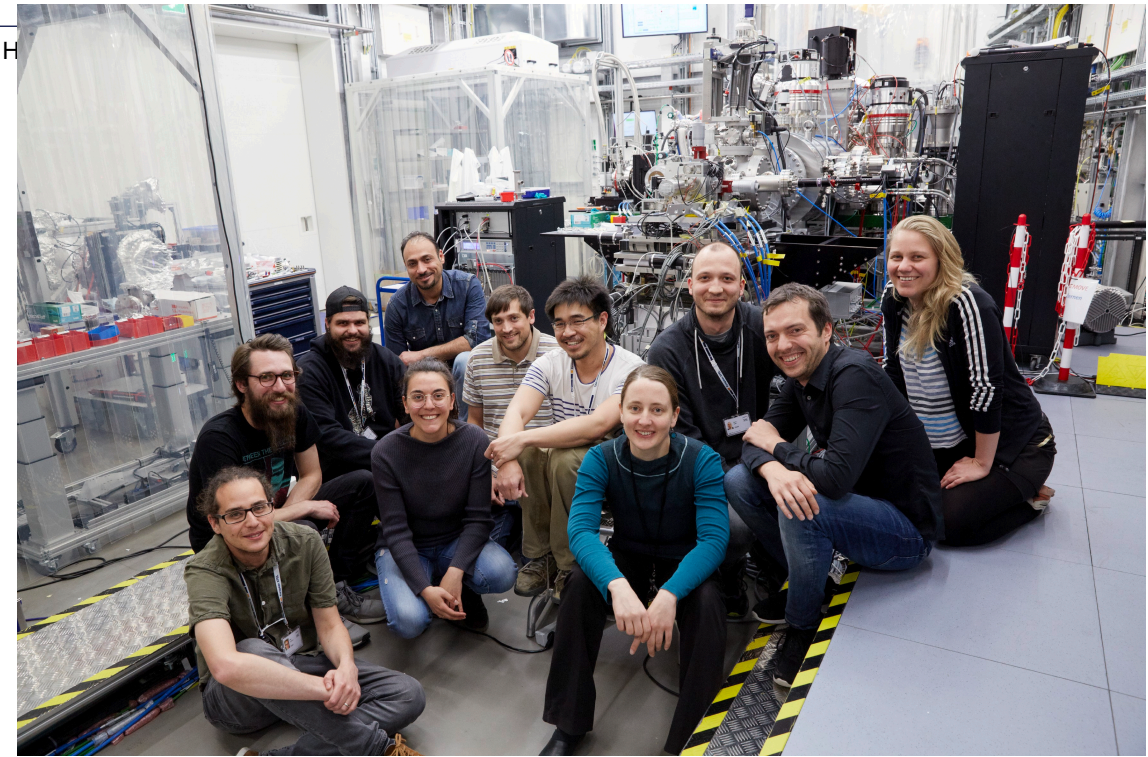# Remote experiments

- 🟧 Issues
  - 🟦 Can users come to the facility?
  - 🟦 What about social distancing / PP during experiment?
    - ► COVID19 regulation: at most 2 people in hutch

- 🟧 "Unattended operation of experiments"
  - 🟦 impossible
  - 🟦 people in hutch during experiment ~10

- 🟧 Remote control of experiments
  - 🟦 facility staff acting on behalf of users
  - 🟦 user representatives may join them
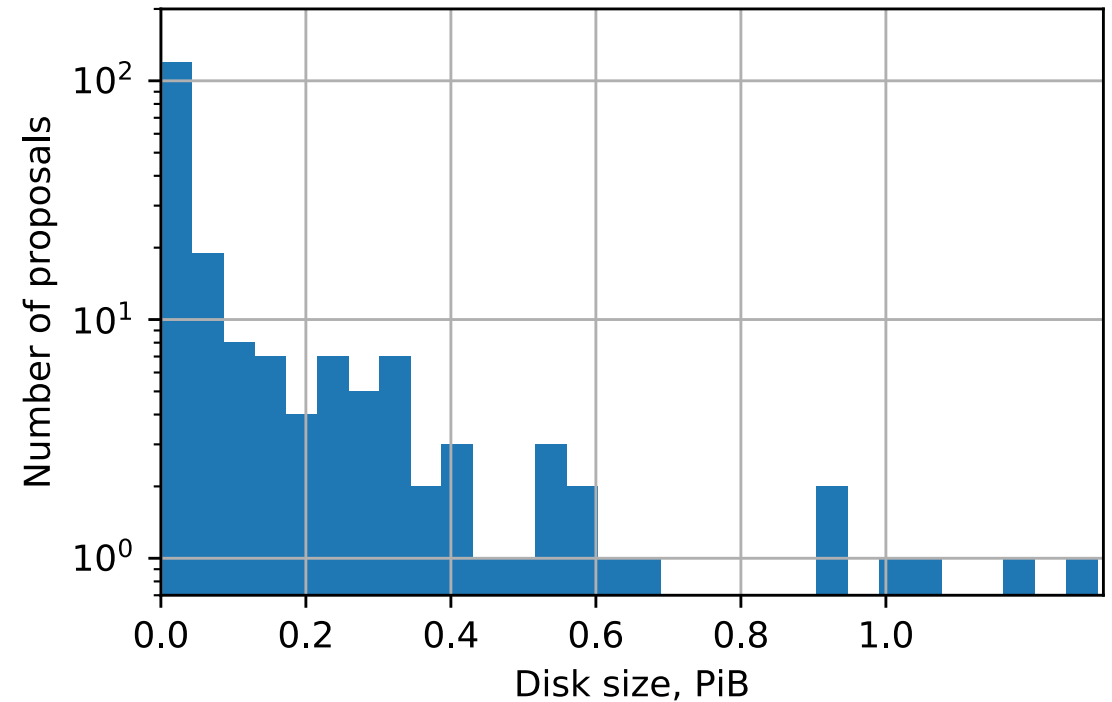 - 🟦 Other users *connected remotely*

European XFEL

Remote offline data analysis

- Data set size
  - ~100 TB not unusual per experiment
  - \> 1 PB is possible
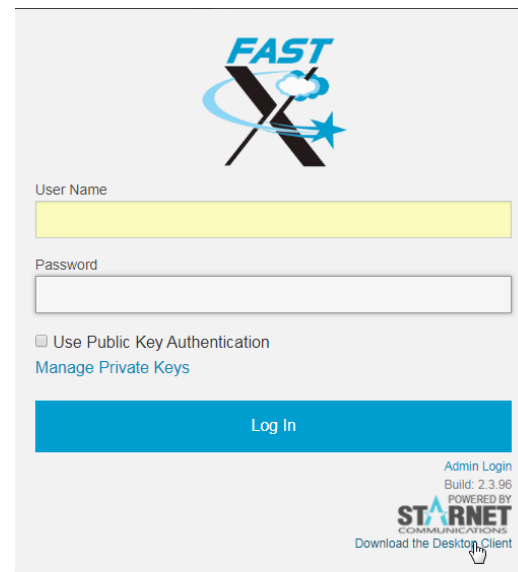   - ▶ Most offline data analysis is remote already

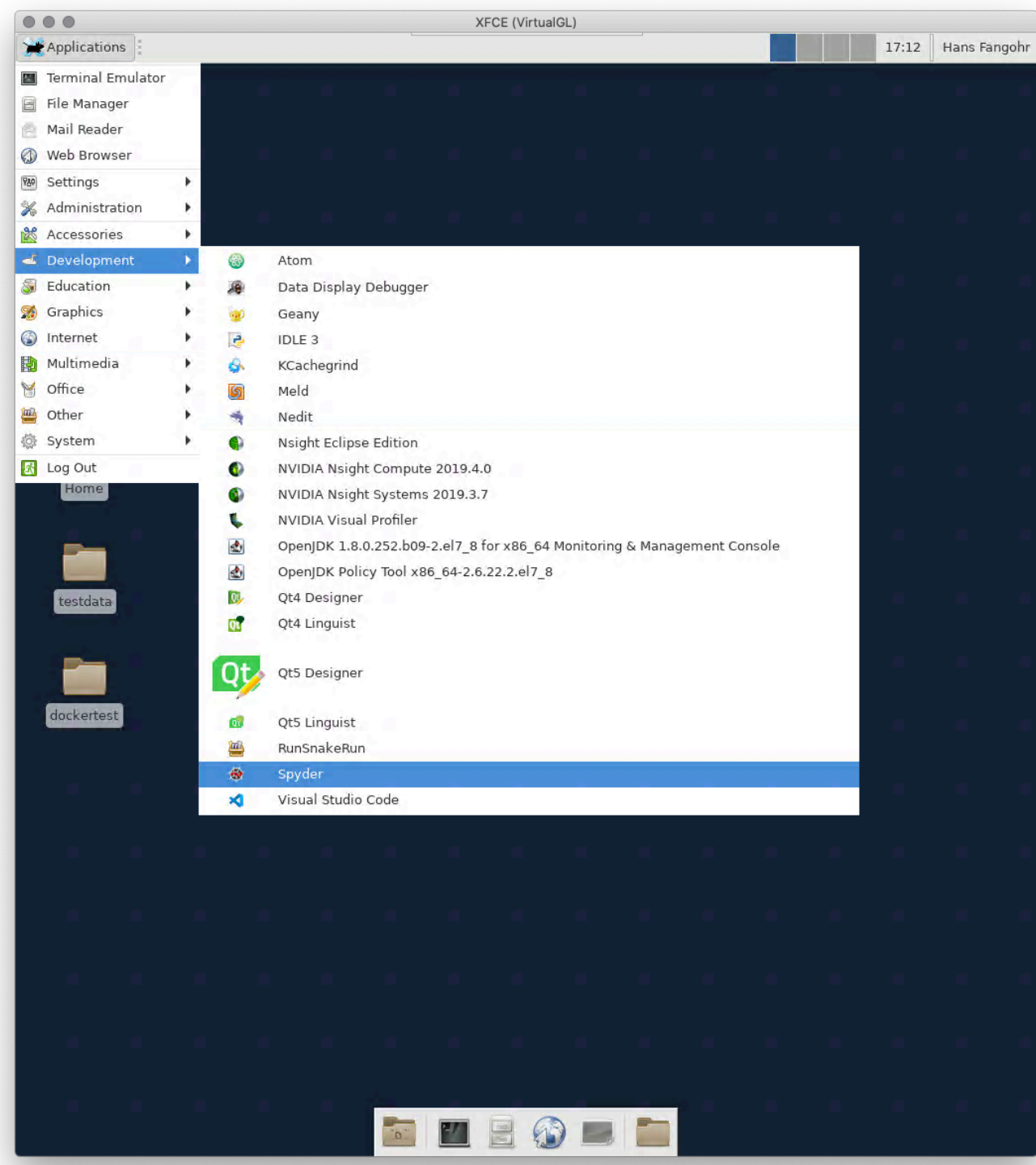- Remote offline data analysis
  - `ssh -X`
  - FastX
  - JupyterHub

FastX

■ Remote Desktop (VNC, NoMachine, …)

■ Graphical user interface (including window manager) made available to user via dedicated app or web browser

■ Better performance than X-forwarding

European XFEL

JupyterHub: use HPC resources through Jupyter Notebook from remote

■ JupyterHub
  ■ users connect through browser and `https`
  ■ serve notebooks on facility hardware
  ■ use existing authentication systems
  ■ connect to users' file storage

■ Popular with users:
  ■ browser and OS of choice
  ■ works locally and remotely the same

■ Orthogonal aspects:
  ■ Good for reproducibility
  ■ Cannot provide for all use cases (GUIs, …)

European XFEL

Maxwell Jupyter Job Options

Maxwell partitionsⓘ dedicated node on ALL partition (preemption) ⌄

Choice of GPUⓘ none ⌄
Note: For partitions without GPUs (or choice of GPUs) the GPU selection will be set to 'none'

Constraintsⓘ V100&INTEL
Note: This will override GPU selections!

Number of Nodesⓘ 4 ⇕
Note: Number of nodes will be set to 1 on shared jhub partition!

Job durationⓘ 1 hour ⌄
Note: on the shared Jupyter partition (jhub) the time limit is always 7 days!

Launch modusⓘ Launch JupyterLAB ⌄

Remote Notebookⓘ PyFAI Tutorials (Collection) ⌄

Node and GPU availability					
Partition	# nodes	# avail	# GPUs avail	# P100 avail	# V100 avail
jhub	3	3	0	0	0
maxwell	61	46	0	0	0
maxgpu	19	12	12	1	10
all	327	188	0	0	0
allgpu	88	67	67	48	10

Spawn

Use case: notebooks as recipes

🟧 Pre-populate notebook with cells to carry out a particular type of data analysis

  🟦 provide a directory full of such recipes to users

  🟦 users execute cells during beamtime and later

🟧 Convenient compromise between

  🟦 static recipe (=script)

  🟦 interactive exploration

```
In [3]: import ToolBox as tb
```

X-ray Absorption Spectroscopy

Step 1: Load data and align them by train id and pulse id

```
In [4]: proposalNB = 900074
        semesterNB = 201930
        runNB = 487
        topic = 'SCS'
        fields = ["SCS_photonFlux", "SCS_XGM", "MCP3apd", "nrj"]
        run = tb.load(fields, runNB, proposalNB, semesterNB, topic,
                      validate=True, display=False)
        nrun = tb.matchXgmTimPulseId(run)

        Checking run directory: /gpfs/exfel/exp/SCS/201930/p900074/raw
        /r0487/
        No problems found
```

Step 2: check the pulse integration window
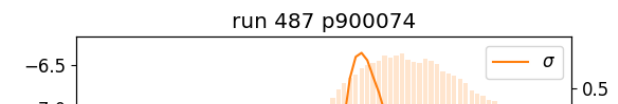
```
In [5]: tb.checkTimApdWindow(nrun, mcp=3)

        no raw data for MCP3. Loading trace from MCP3
```
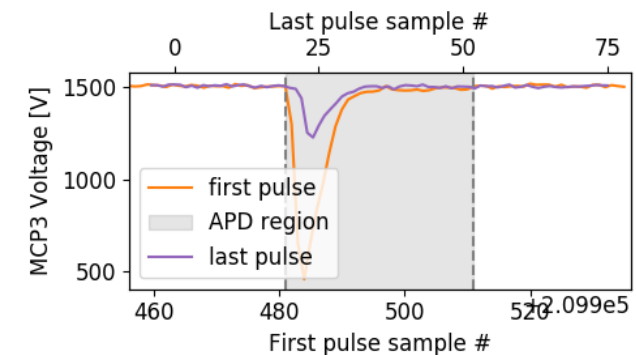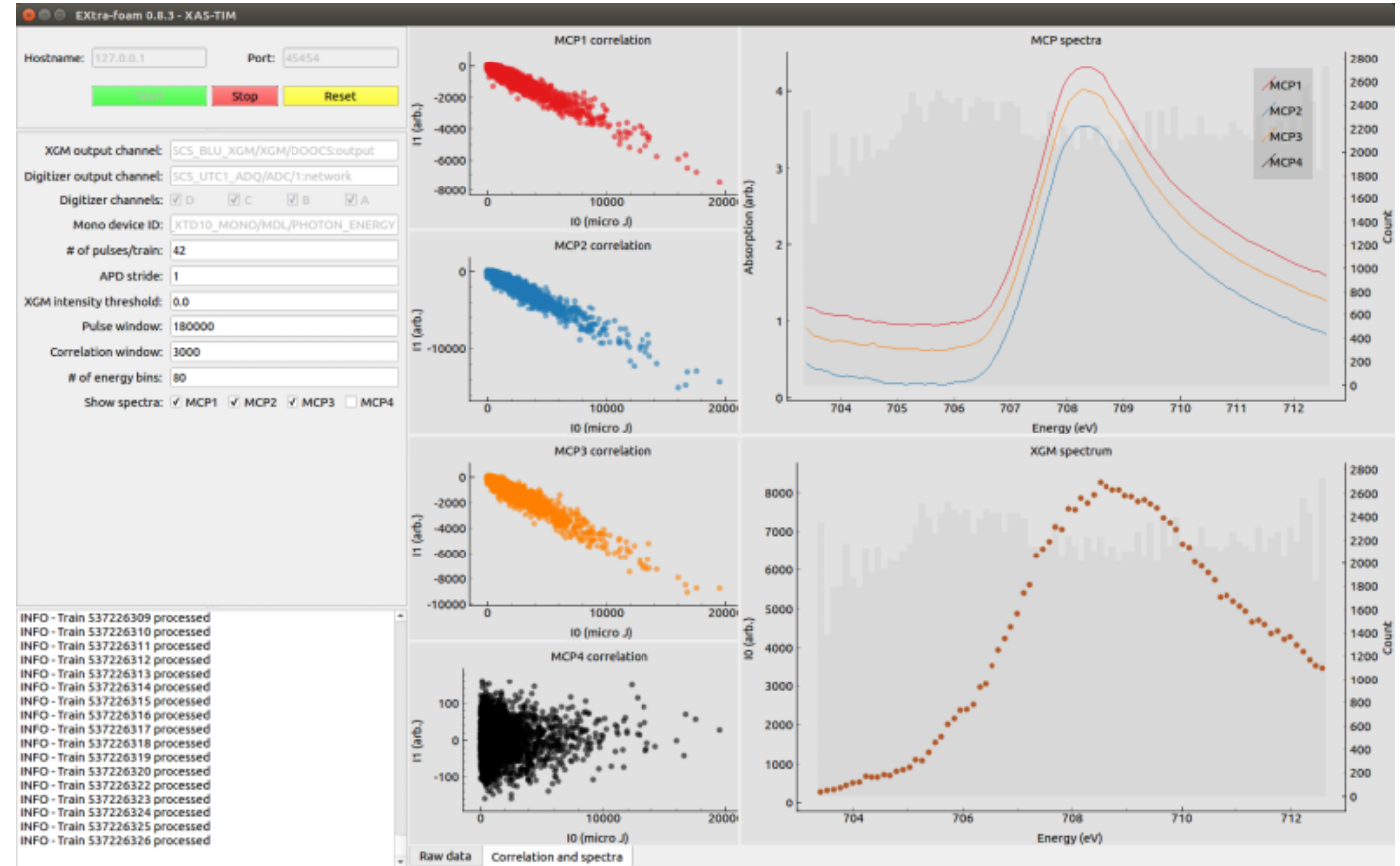

Step 3: bin the data and plot the XAS spectrum

```
In [6]: nrj = np.linspace(nrun.nrj.min(), nrun.nrj.max(), 80)
        xas = tb.xas(nrun, nrj, plot=True)
```
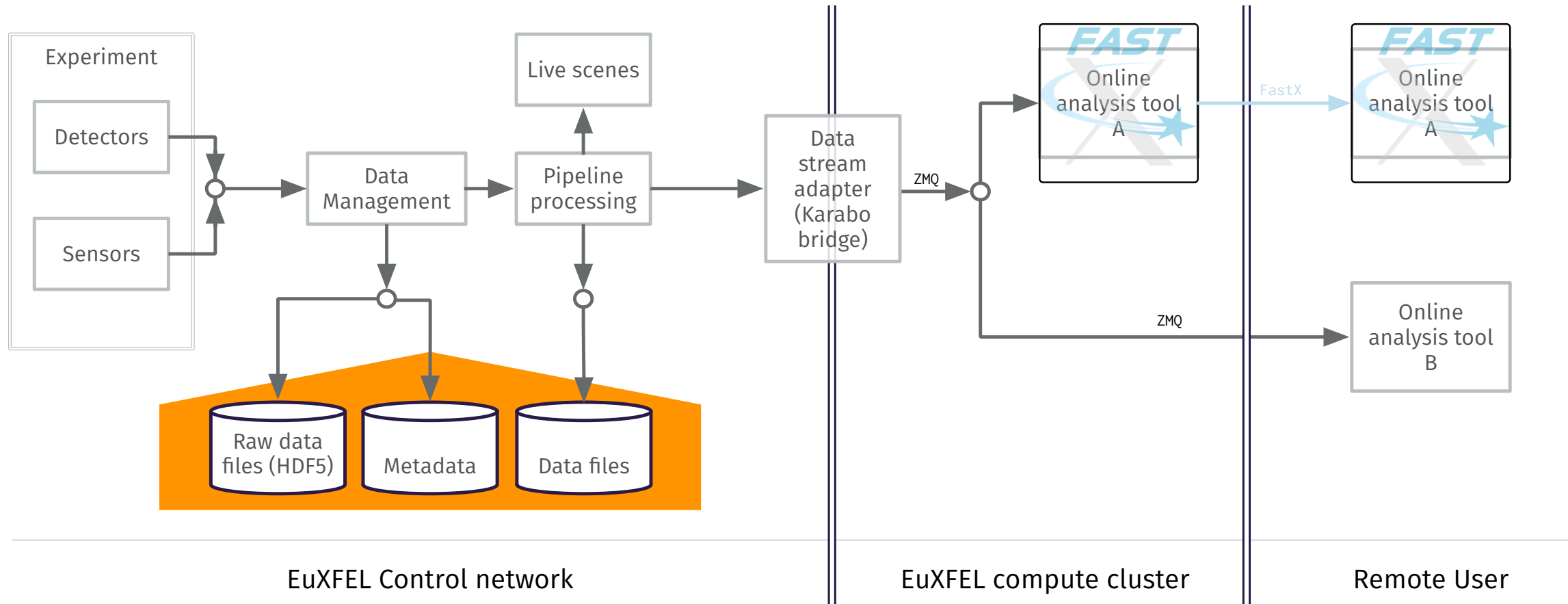
Online data analysis

■ Fast feedback during experiment

Remote Data Analysis options

Remote control

- Can users control hardware from remote?
    - ▶ No, safety regulations.

- Users need to participate in experiment
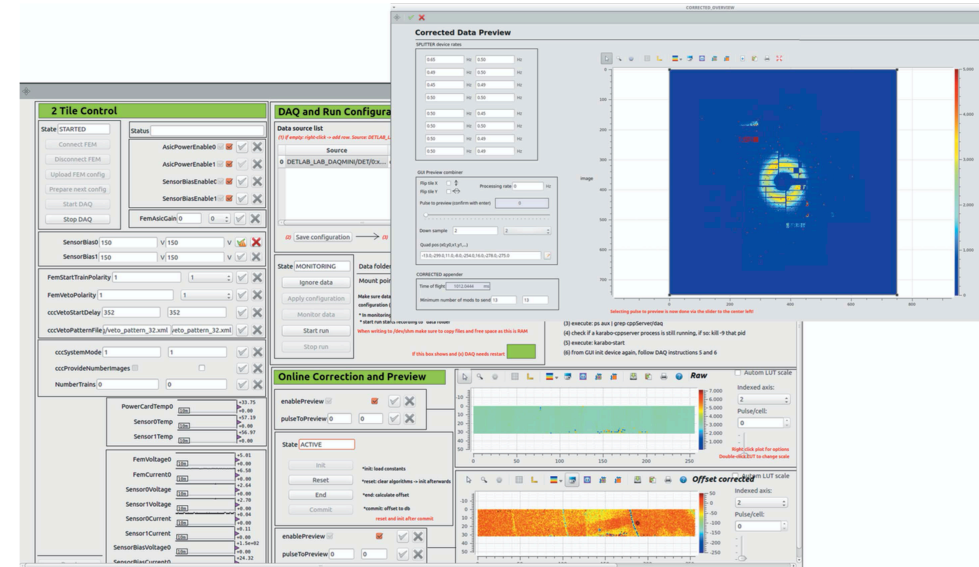   - ▶ Read-only view of control panels.

Remote control

🟧 Similar options as for remote data analysis

🟦 Either have read-only control panels and forward display via FastX (or other remote desktop software)

🟦 Or provide the software (here Karabo GUI client) to run on their remote hardware, and offer Karabo-specific data stream to connect to

► Preferred, as data can be interrogated interactively (zoom, pan, etc)

🟧 Must be *read-only.*

European XFEL

```
In [2]: overview("Italy");
```

Open Science COVID19 Analysis (OSCOVIDA) project

■ Photon and Neutron Open Science Cloud (PaNOSC) activity

■ Dashboard with time series analysis
  ■ infections and deaths
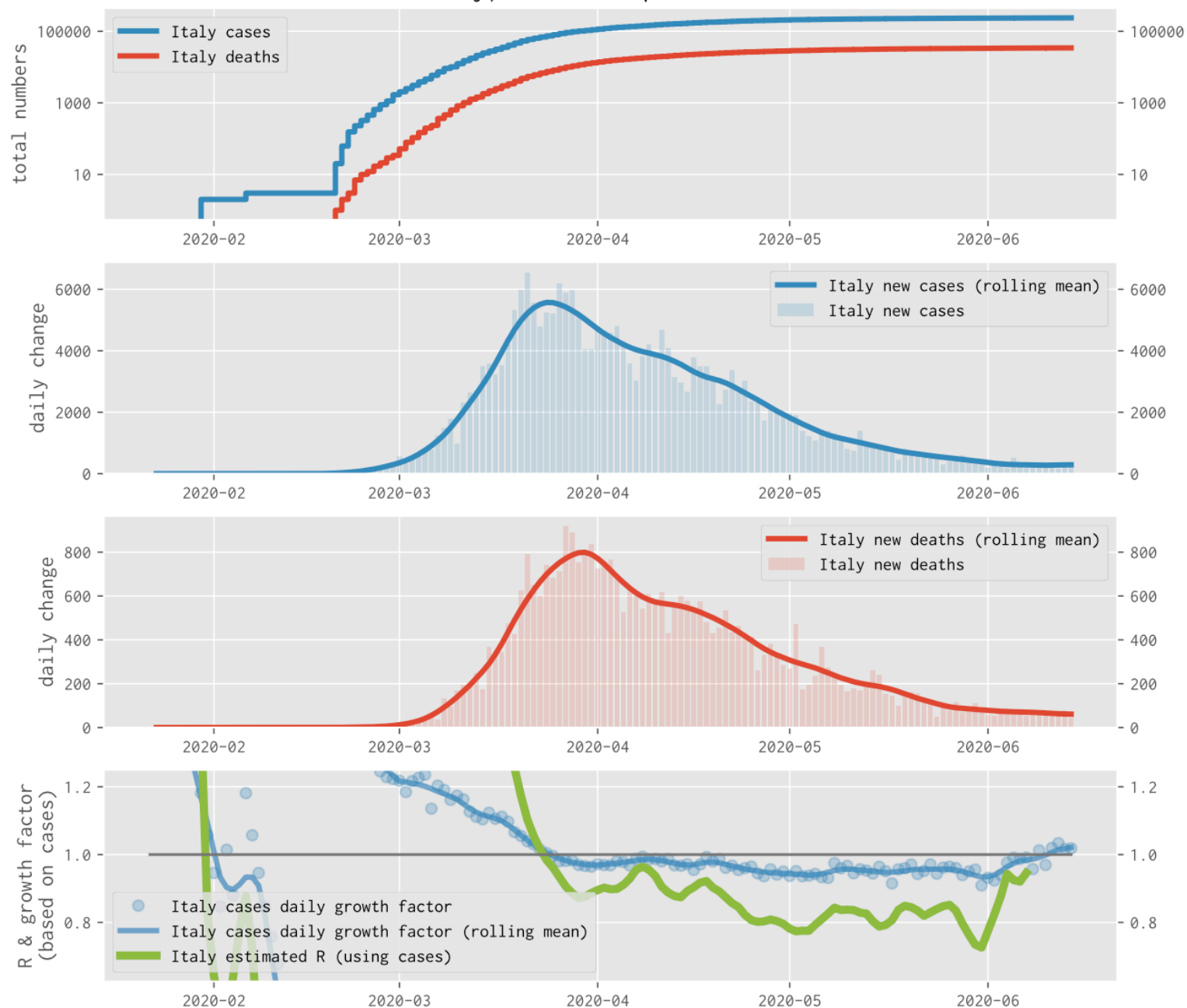 ■ updated daily

■ Data sets for
  ■ All countries (Johns Hopkins)
  ■ US states (Johns Hopkins)
  ■ 412 districts in Germany (Robert Koch Institute)

■ https://oscovida.github.io

European XFEL

Overview Italy, last data point from 2020-06-14

Open Science COVID19 Analysis (OSCOVIDA) project

- **Motivation**
  - put media reports into context
  - motivate social distancing
  - track pandemic & new outbreaks

- **Open Science & citizen science**
  - Data sources are public
  - Processing code is public

- **Re-producible and re-usable**
  - Plots created in Jupyter Notebooks
  - Can re-execute each notebook in the cloud (Binder)
  - Can re-use and extend analysis

- **Contributions welcome**

European XFEL

OSCOVIDA: Open Science COVID Analysis

Home   All regions   Countries   Germany   US   Articles   Analysis   About

Tracking plots: Countries of the world

Show 25 entries   Search:

Location	Total cases	Total deaths	New cases last week
Brazil	867,624	43,332	175,866
US	2,094,058	115,732	149,749
India	320,922	9,195	63,436
Russia	528,267	6,938	61,194
Pakistan	144,478	2,729	40,807
Chile	174,293	3,323	40,143
Peru	229,736	6,688	33,221
Mexico	146,837	17,141	29,734
Saudi Arabia	127,541	972	25,627
South Africa	70,038	1,480	21,753
Bangladesh	87,520	1,171	21,751
Iran	187,427	8,837	15,638
Qatar	79,602	73	10,812

Communication – instant messaging / chat tools

- 🟧 Mixture between email and phone call
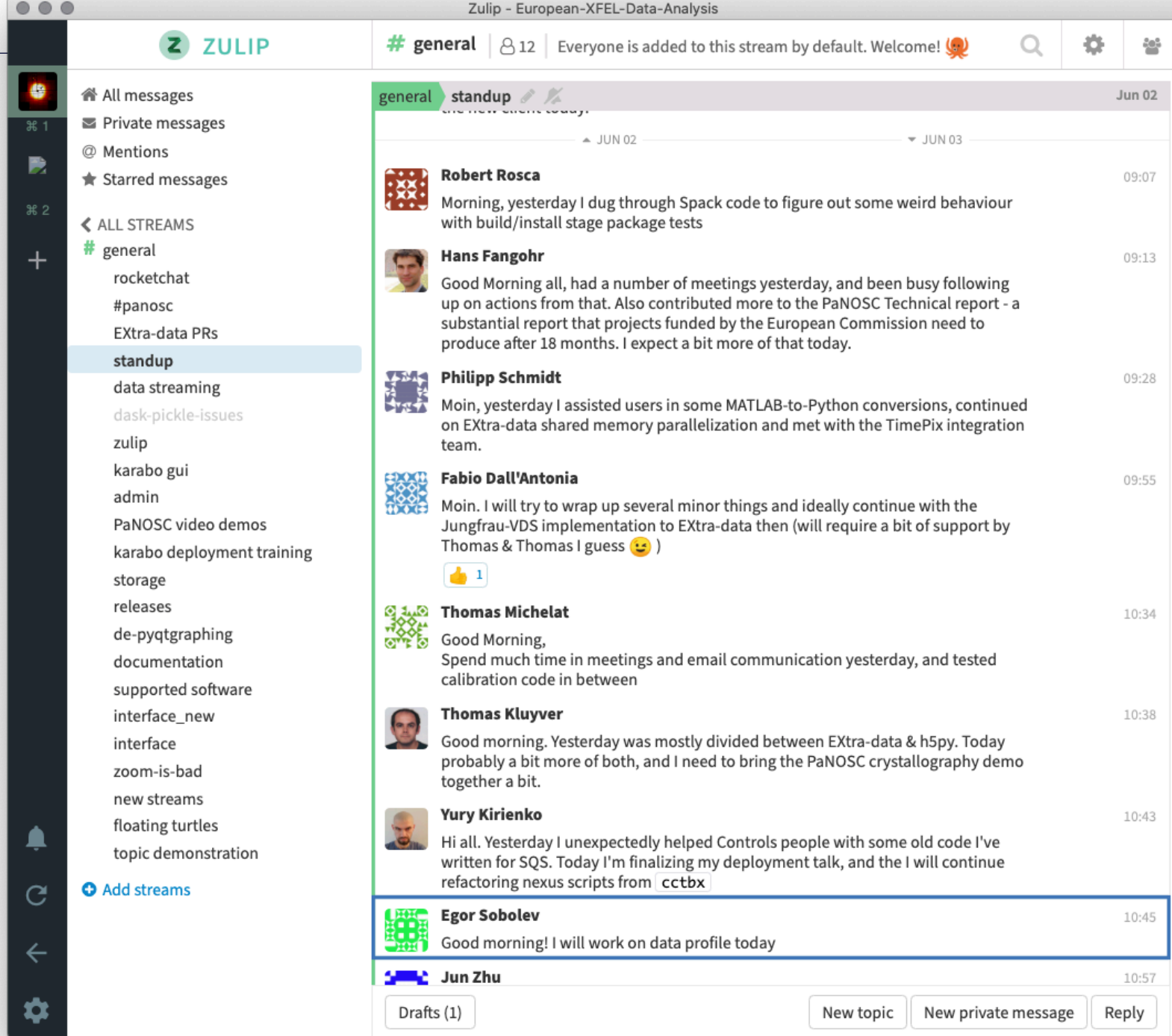 - 🟦 interactive
  - 🟦 but asynchronous

- 🟧 Tools tried
  - 🟦 Slack, RocketChat, Zulip, Mattermost, Zoom(chat), [IRC in the past]

- 🟧 Example:
  - 🟦 Data Analysis group at EuXFEL

European XFEL

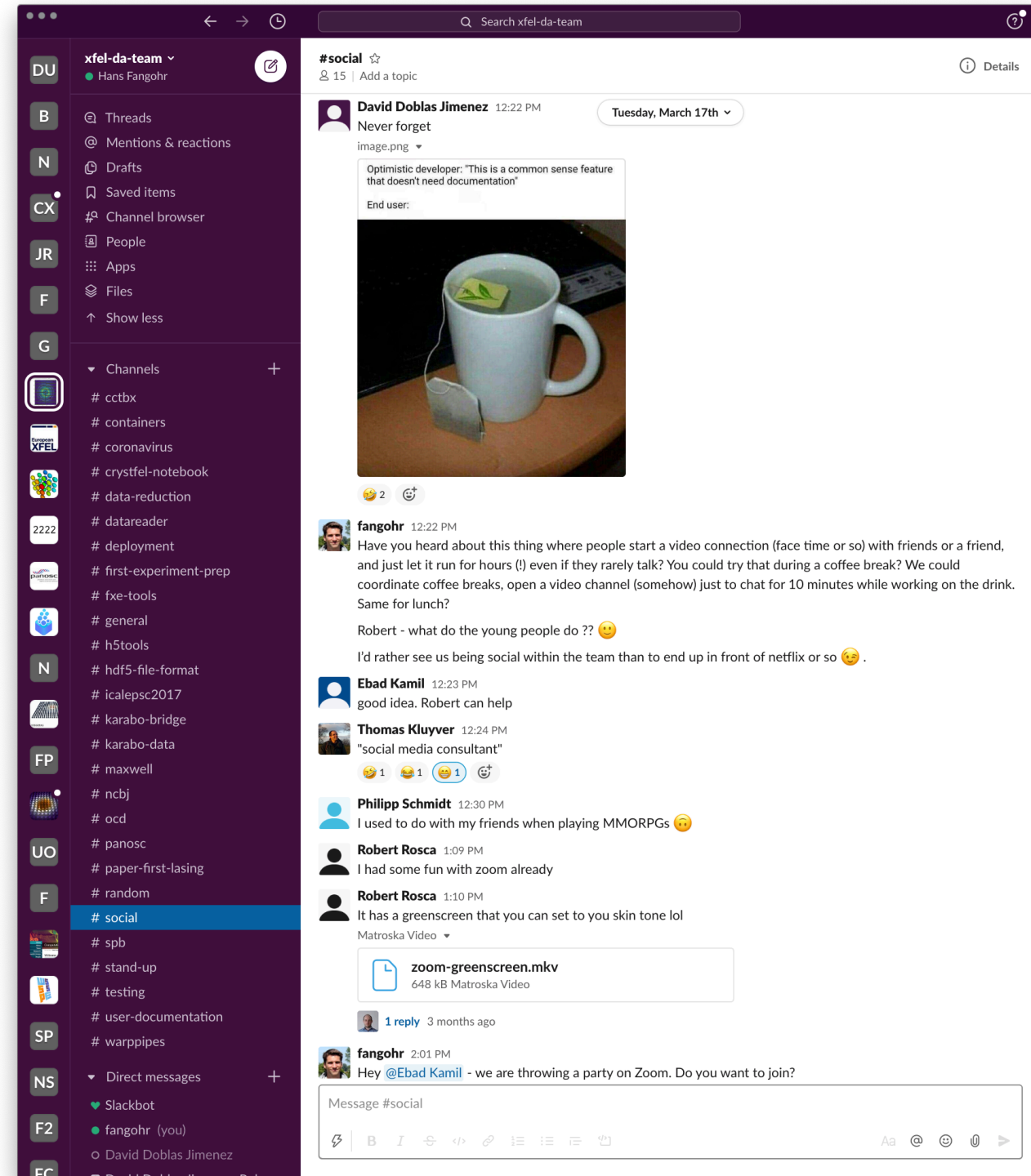# Communication – instant messaging / chat tools

■ High potential for communication improvements:
- ■ Within-team
- ■ Within-facility
- ■ Remote-teams
  - ► Social glue

- ■ For user groups
  - ► Meta data capture

■ Recent insight:
- ■ Zulip has interesting concept of "topics" within "channels/streams".

European XFEL

Thank you for your attention

European XFEL