Navigating the Diamond Cluster

Felicity Bertram
MX Support Scientist
mx-usersupport@diamond.ac.uk
Felicity.Bertram@diamond.ac.uk

Background

Before Covid:

- 50% of Diamond Users are MX
- 55% of Experimental Sessions are MX
- More than 75% of MX visits were remote

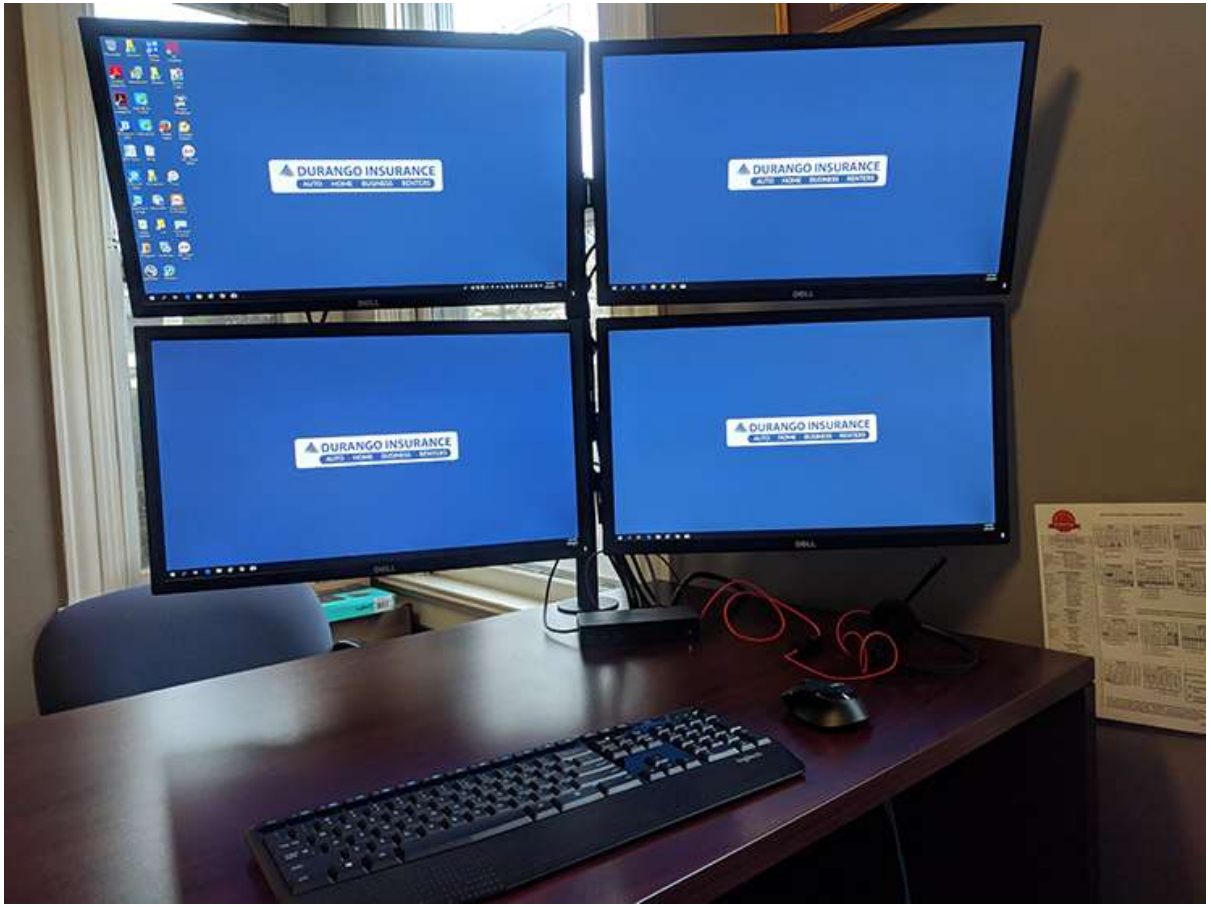
Since then added UDC to our offering

- UDC is “Unattended Data Collection” which doesn’t need any User interaction once they have put their data collection parameters into ISPyB, before shipping their crystals

- But now we don’t have Users on site, it’s harder to interact with them and train them – hence doing courses like this.

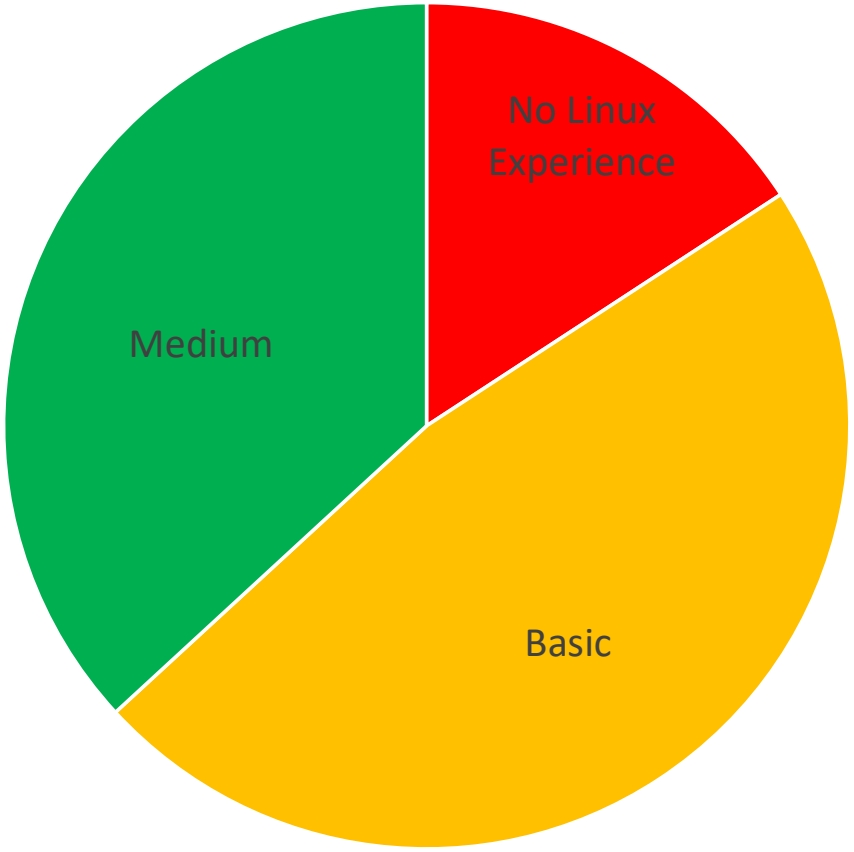
- Therefore, please take the opportunity to network with members of the Diamond team

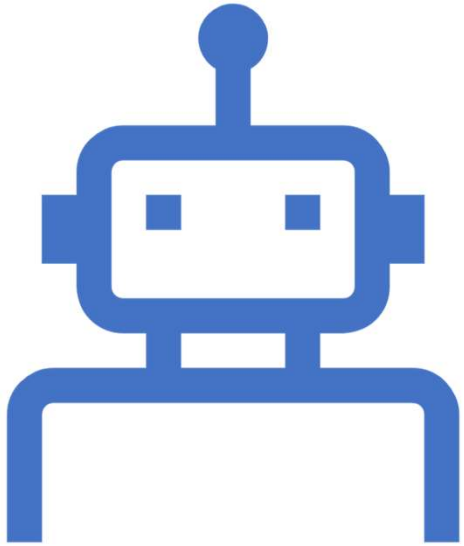
Lots of monitors to control a beamline…



- GDA
- ISPyB
- Diffraction Image Viewer
  - e.g. ADXV
- Machine Status Page

The Audience…





```
tar xvf DPF3_247398.tar

dials.import x247398/t1.0*.img.bz2

dials.show imported.expt

dials.import x247398/t1.0*.img.bz2 geometry.goniometer.axes=-1,0,0

dials.image_viewer imported.expt

dials.find_spots imported.expt

dials.reciprocal_lattice_viewer imported.expt strong.refl

dials.index imported.expt strong.refl
```

Multiple ways to get to the same destination…

You have to be precise

Fed IDs and Passwords

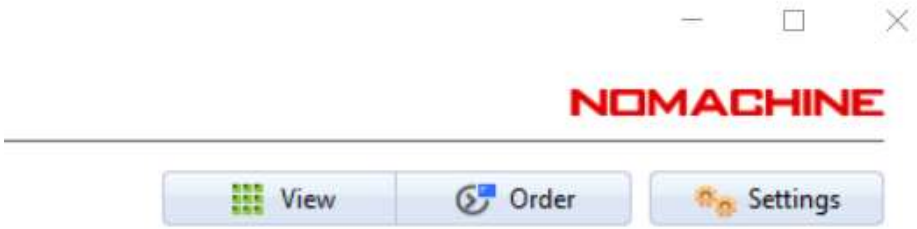
- You all have a Fed ID and password.
- If needed, Diamond staff can tell you what your Fed ID is, but not what your password is
- We can only request that a password is reset – and that is within UK office hours
- You will use this to log in to NX sessions (this talk) and ISPyB sessions – so make sure it works

NoMachine – Enterprise Client
→ your route in to Diamond

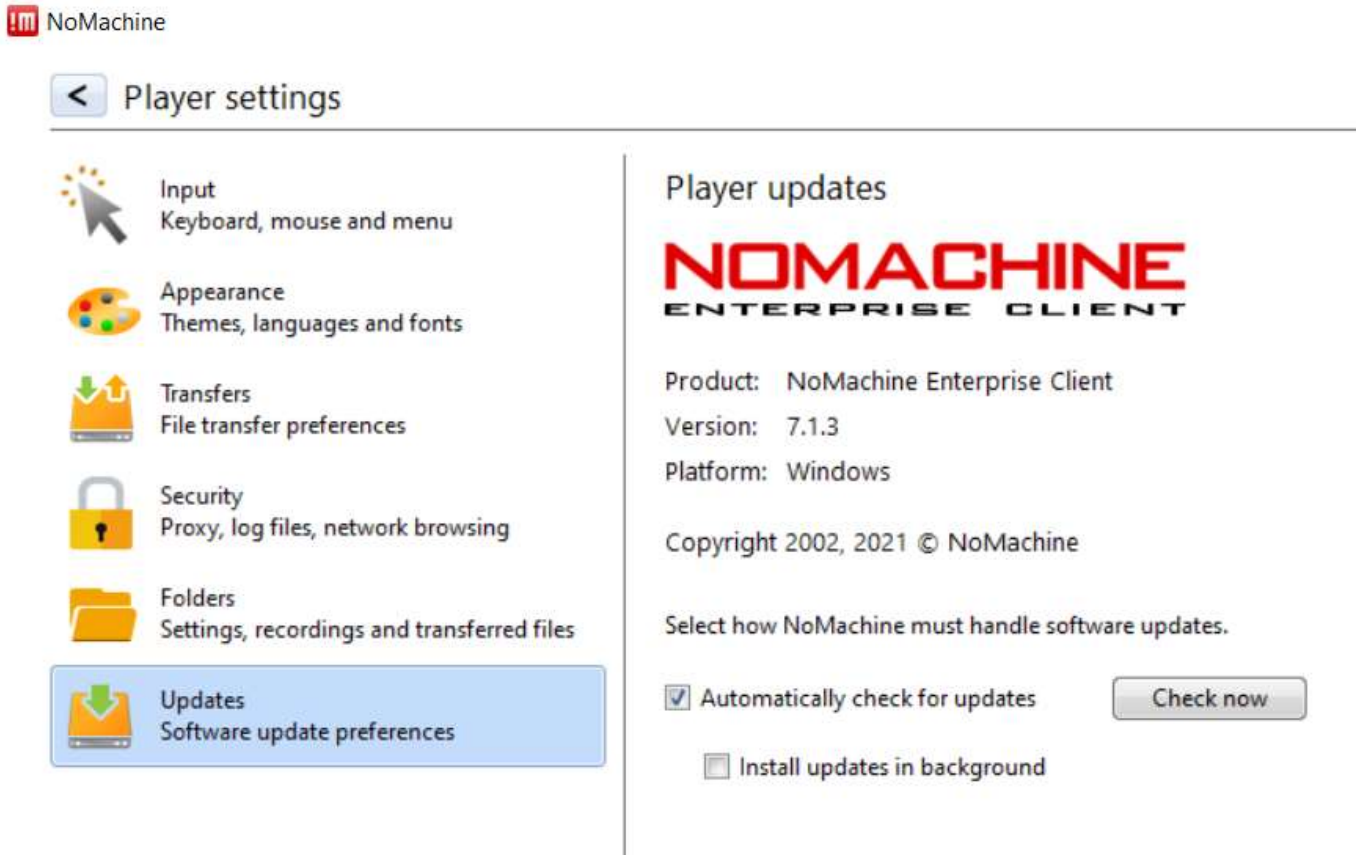
- Download to your machine - https://www.nomachine.com/product&p=NoMachine%20Enterprise%20Client

  OR

- Make sure it is the latest version in Settings



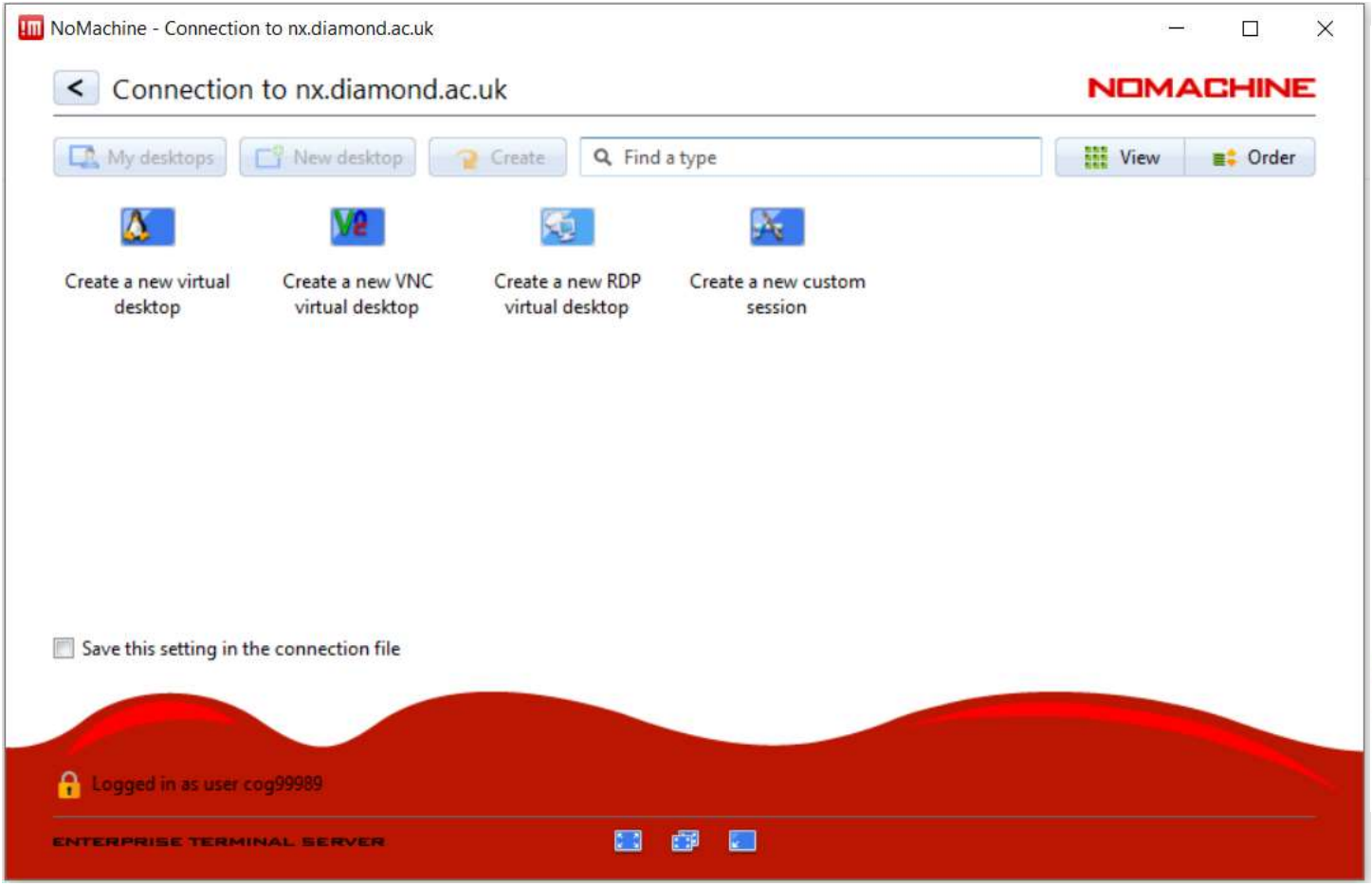
Checking for updates in settings



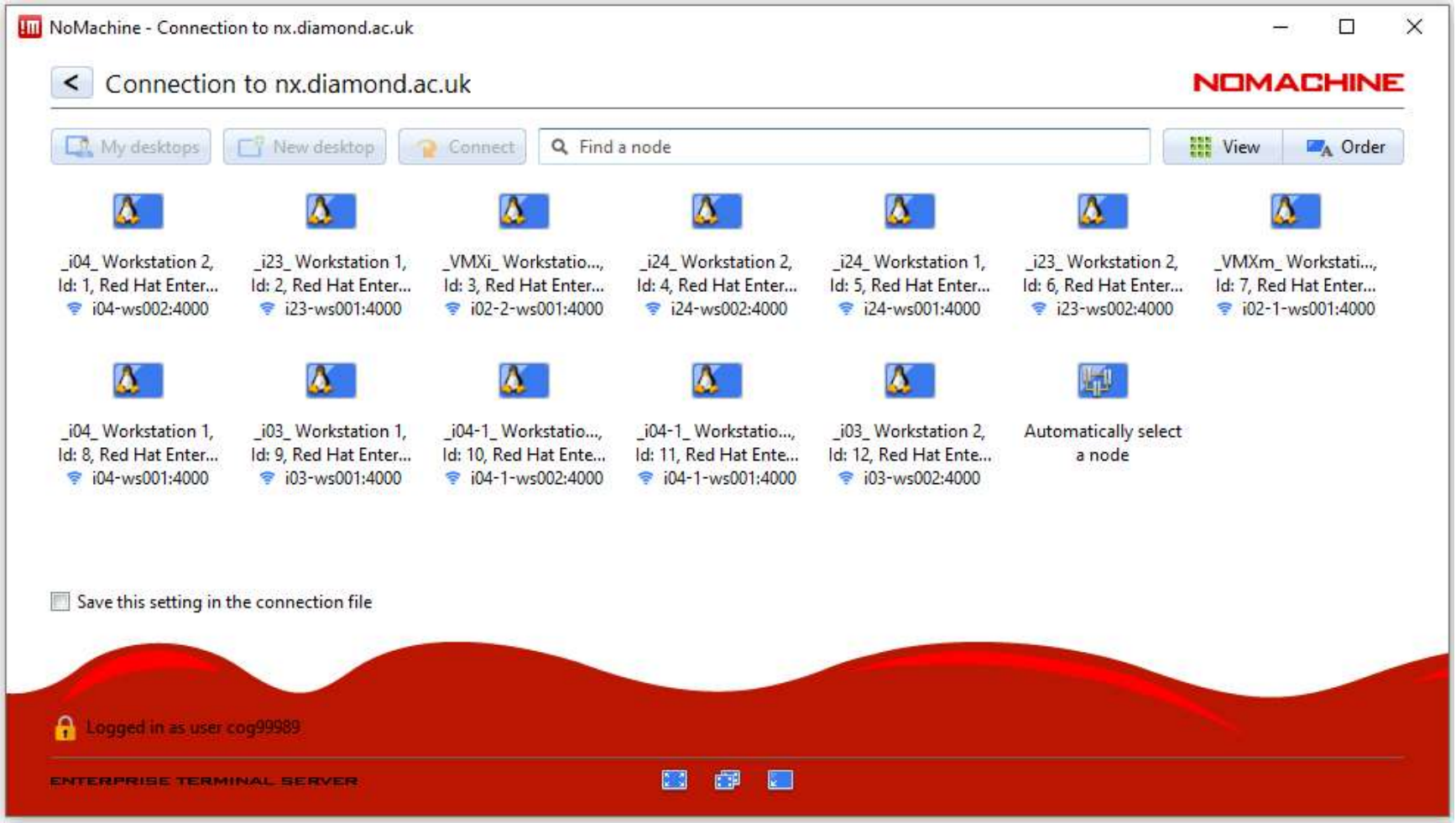
First Time Users/Computers

- Create a new connection using the protocol NX
- Set hostname to nx.diamond.ac.uk (ensure port is set to 4000)
- Use password authentication
- Don’t use a proxy
- Give a name to the new connection, e.g. "Connection to nx.diamond.ac.uk"

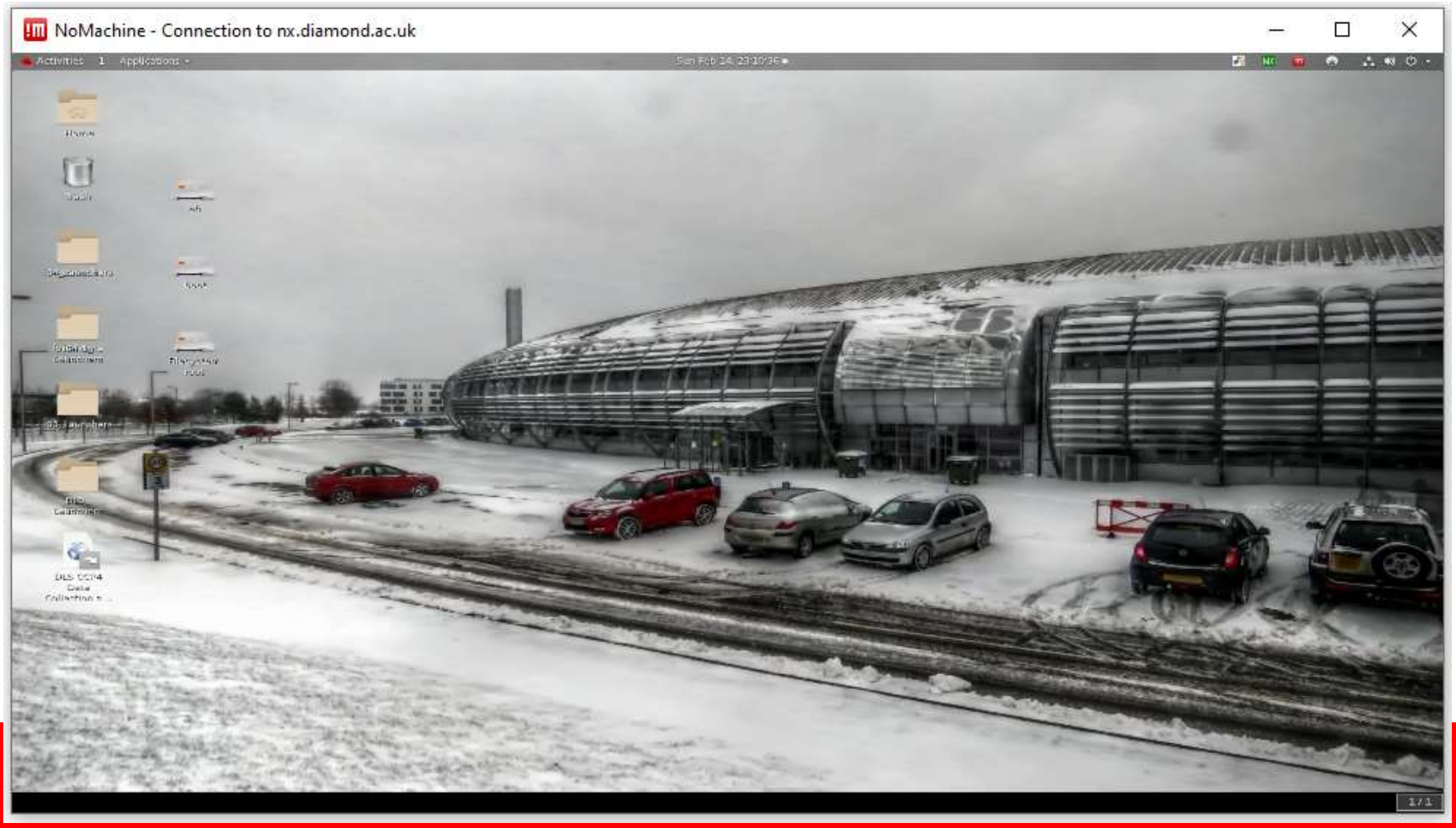
Create New Virtual Desktop



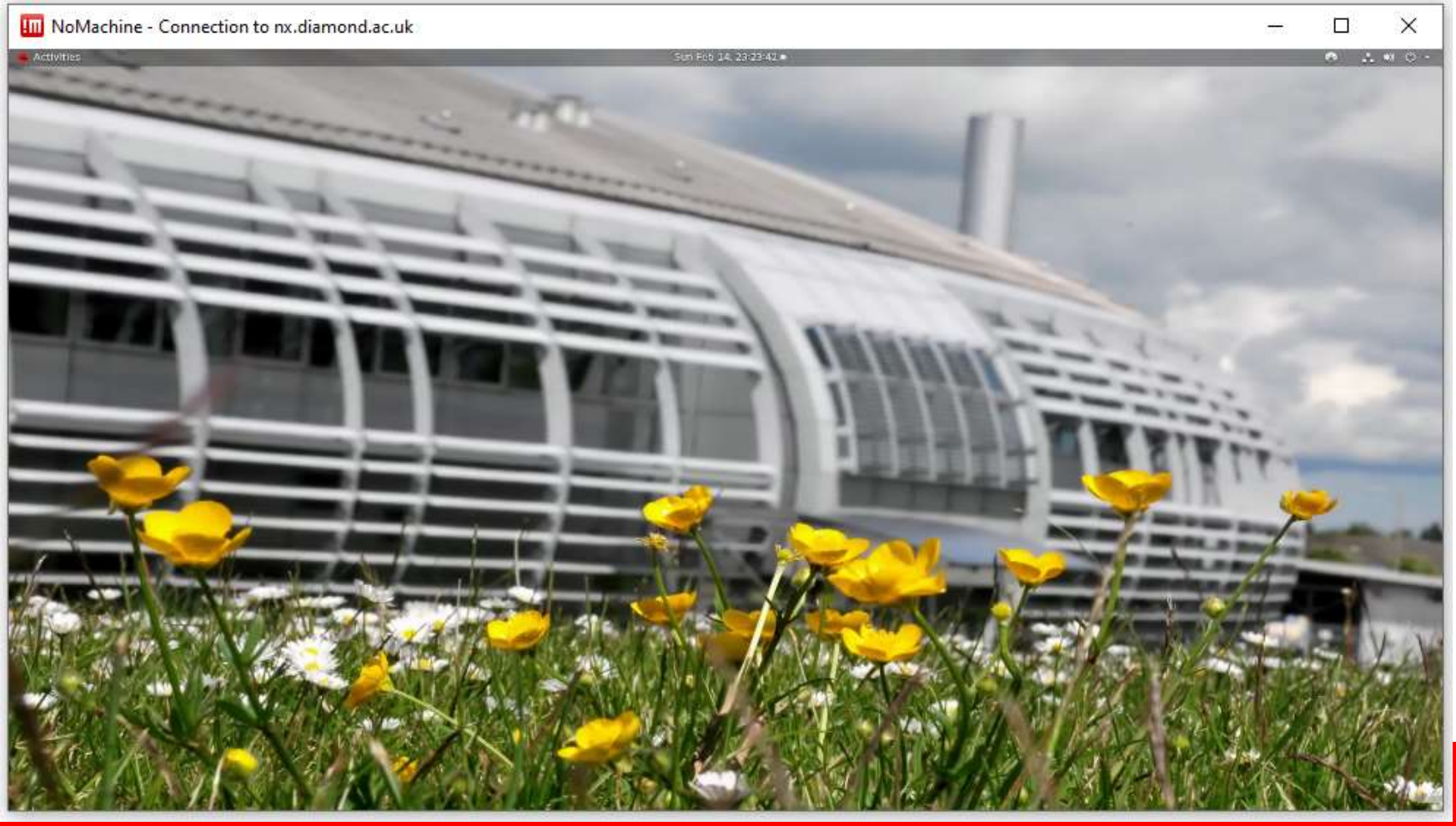
Connecting to NoMachine



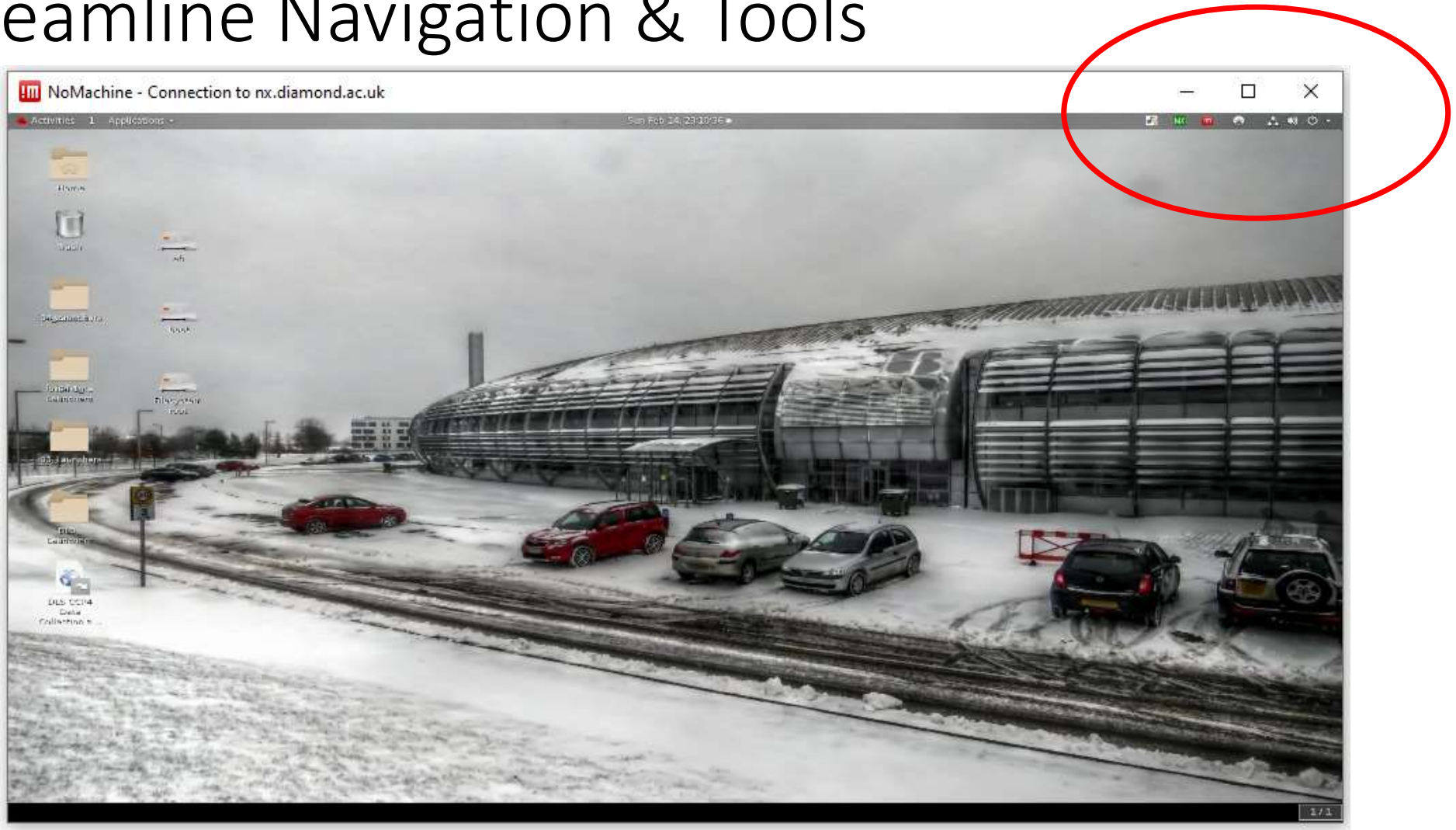
Connecting to a Beamline – only for a visit



Automatically select a node – all data processing

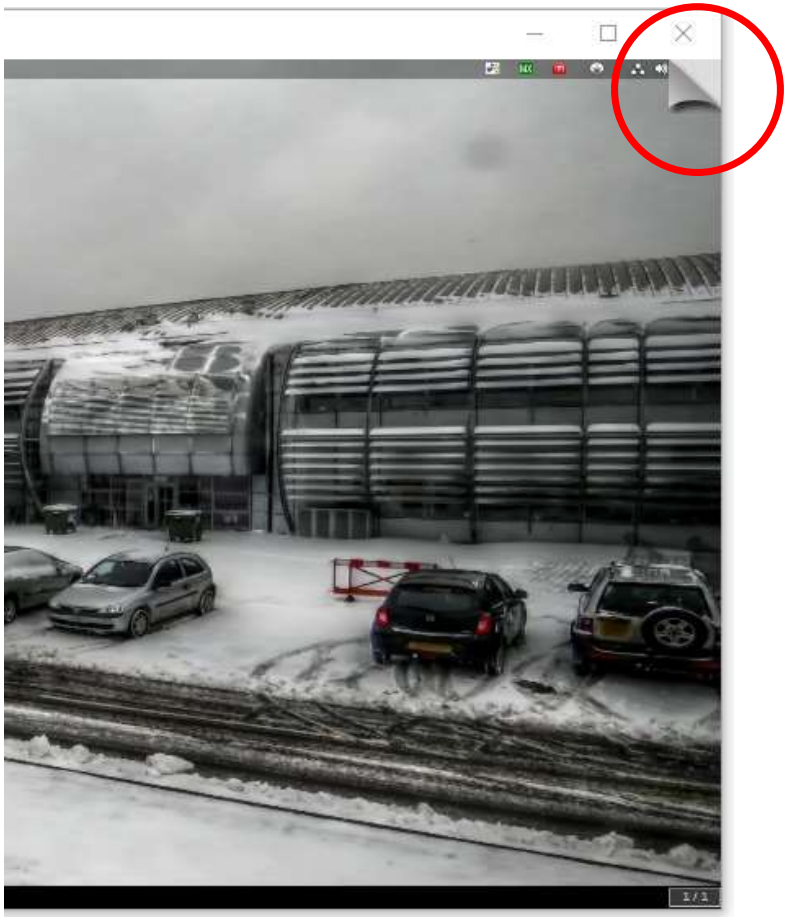


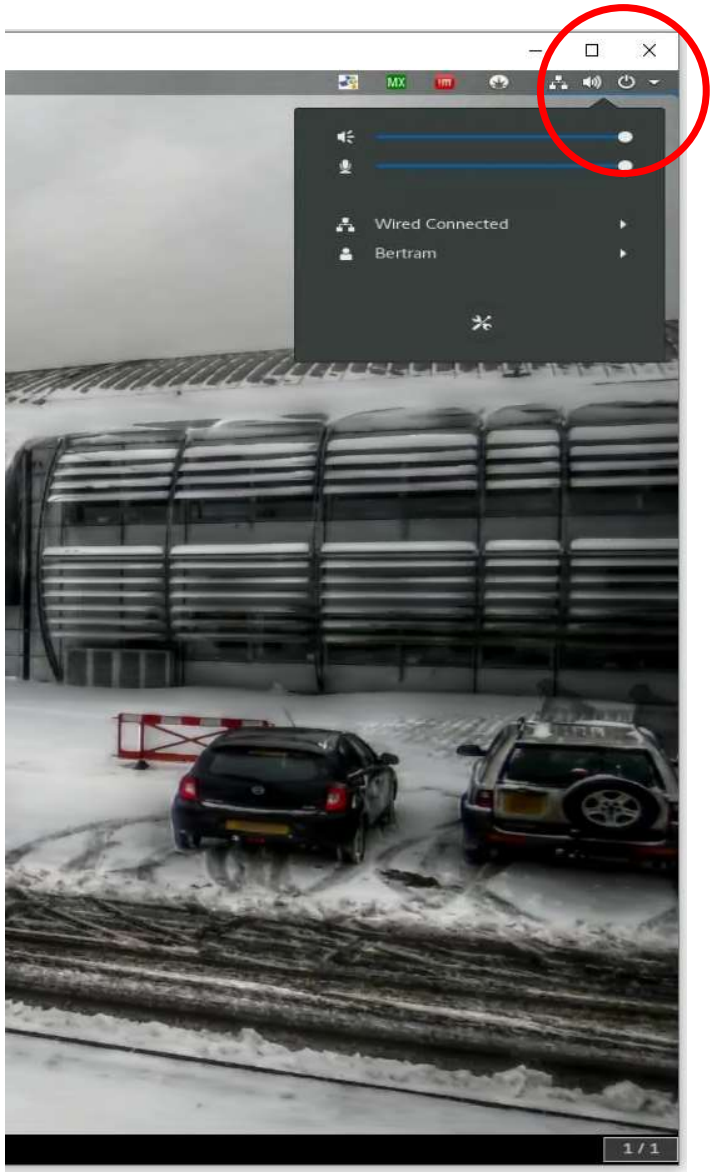
Beamline Navigation & Tools



Logging Off

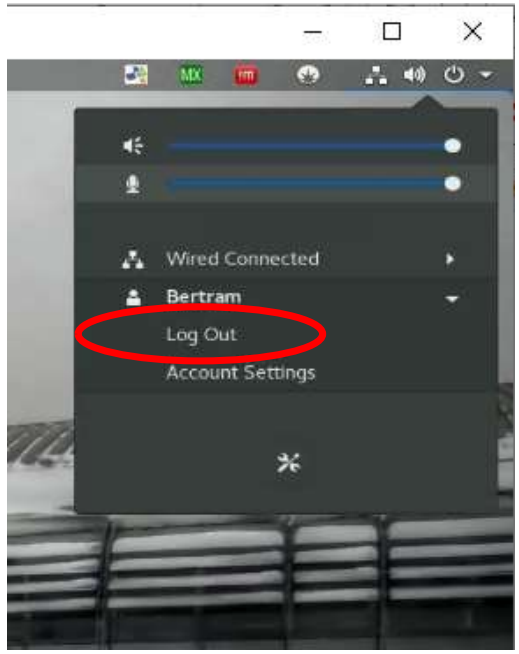
NX Settings – “corner peel”



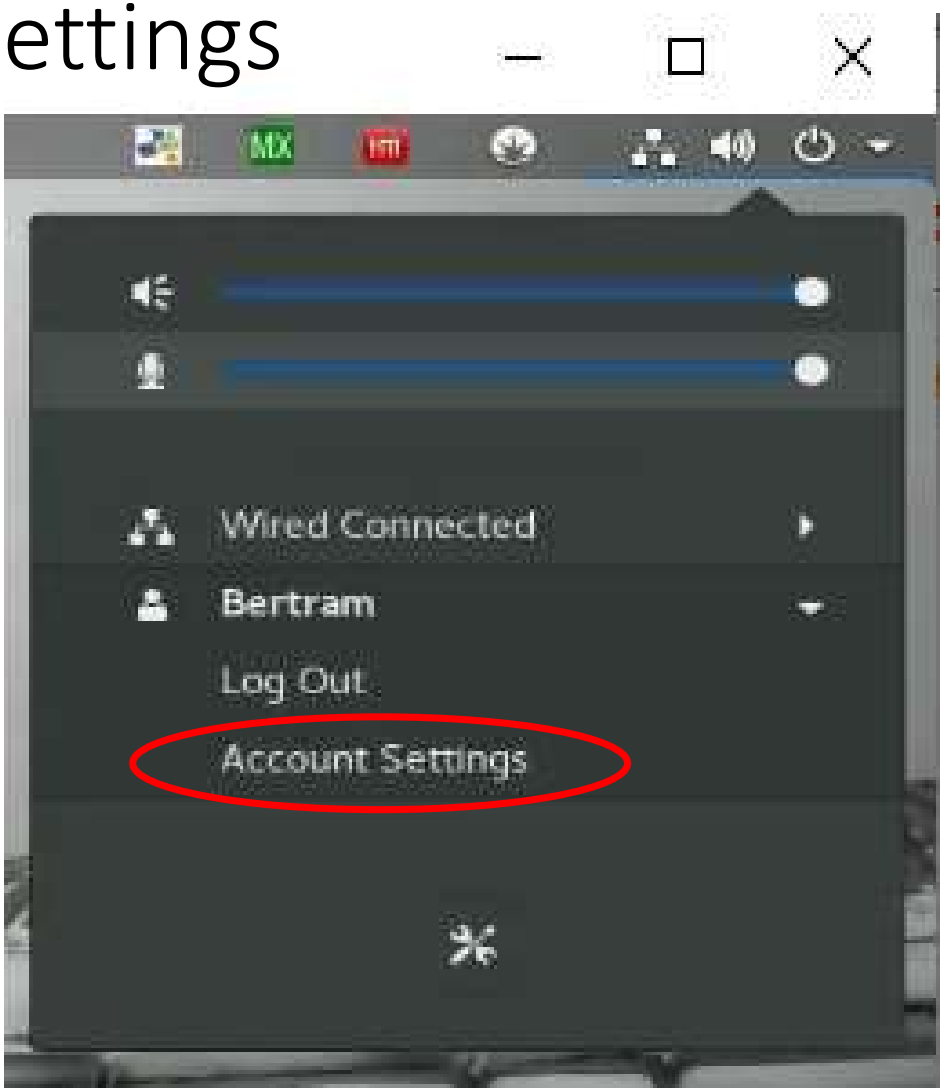


Log off:

- power button
- then click your name
- Then click “Log Out”

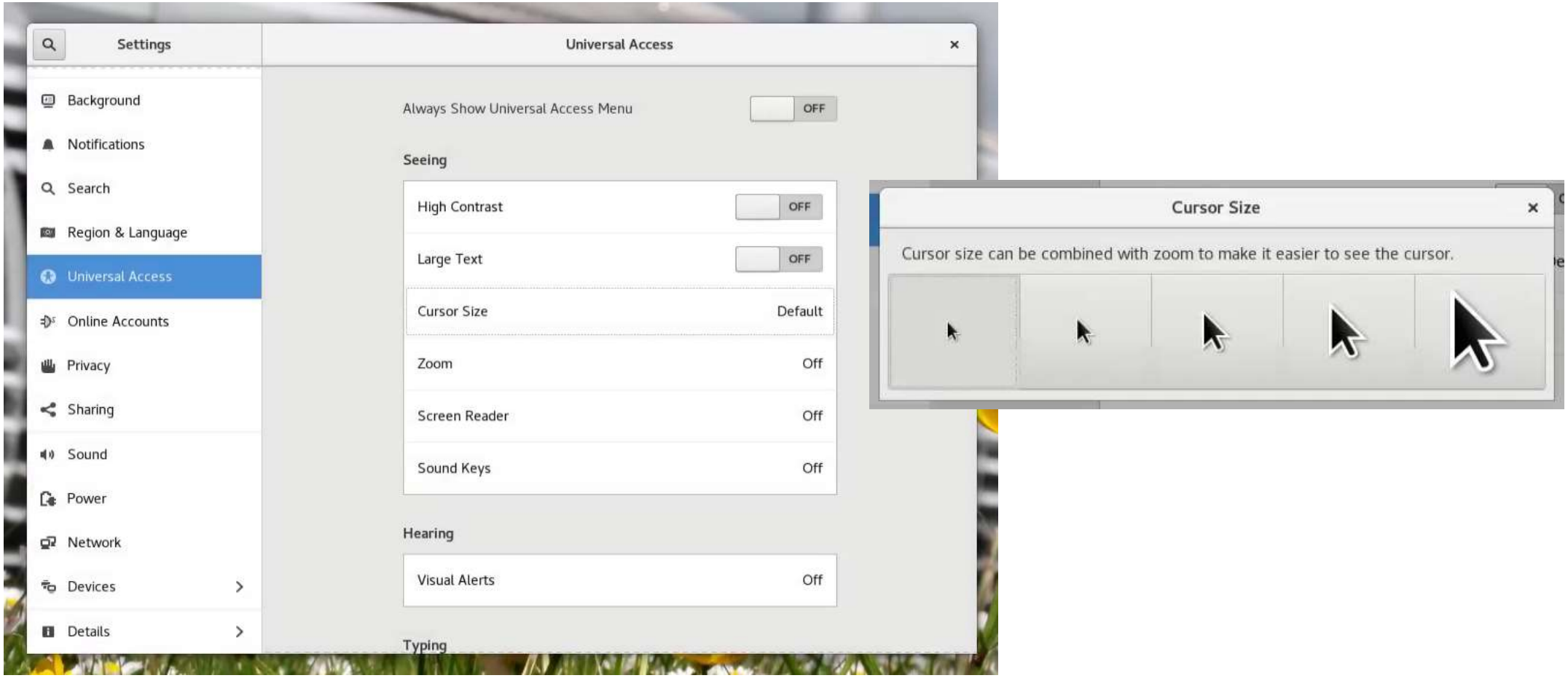


Account Settings



Universal Access

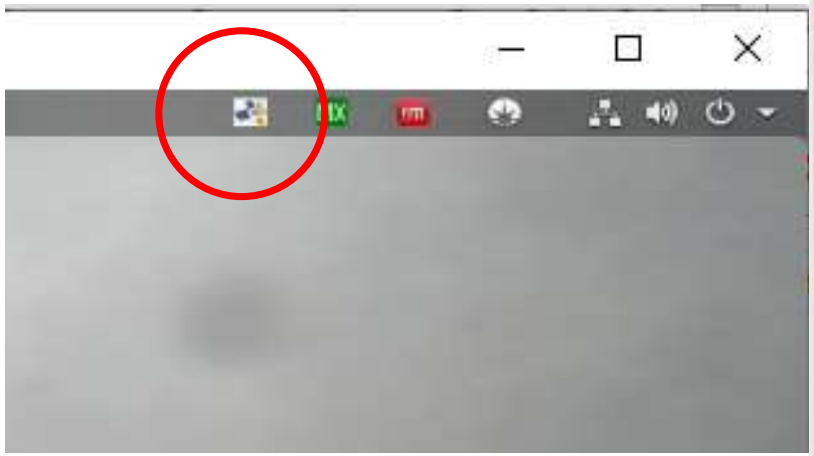
- use to increase the size of the cursor

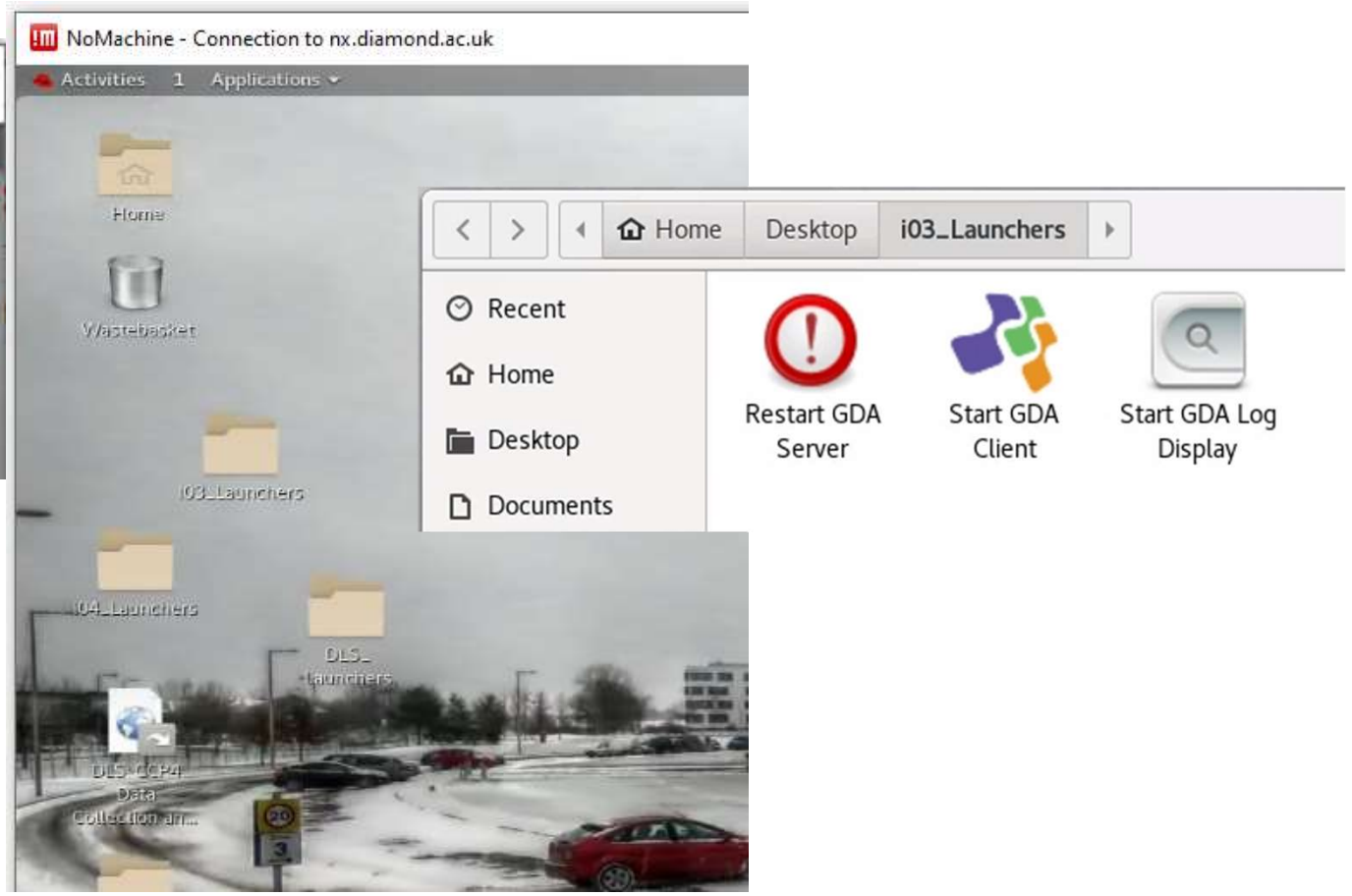


Nomachine Display Options

- “Enable viewport display” = Scrolling window (annoying)
- **“Scale to window” = Fits in window, everything shrinks when window shrinks**
- “Resize remote display” = Fits in window, icons remain large if window shrinks
- “Fullscreen” = window is maximised on current monitor
- “Fullscreen on all monitors” = windows are maximised on all monitors
- “Iconise” = minimise (ctrl + alt + M also works)
- “Settings”

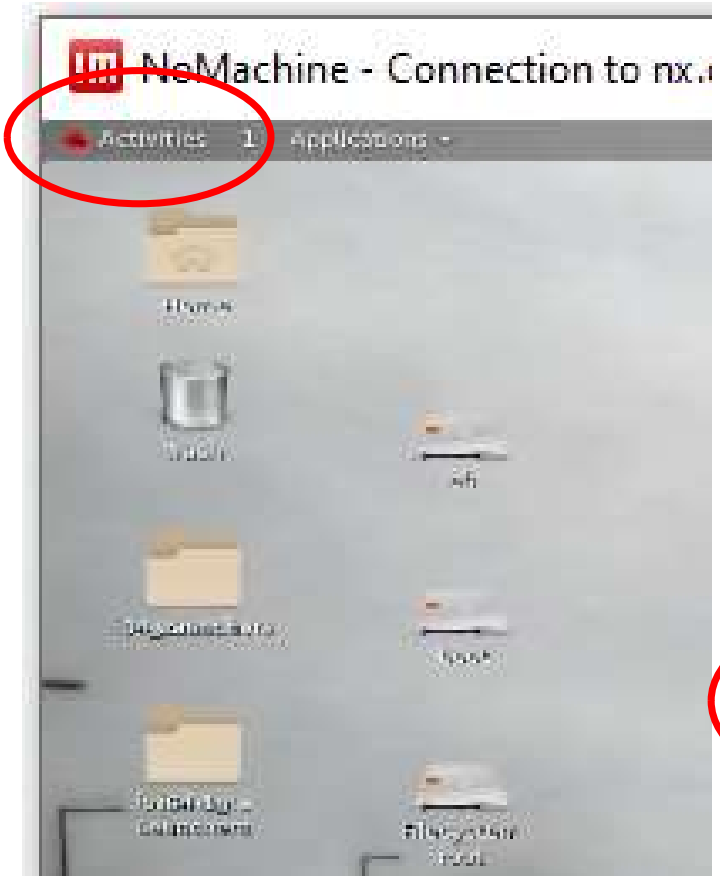
Launch GDA (Generic Data Acquisition)
- used to control beamlines and collect data

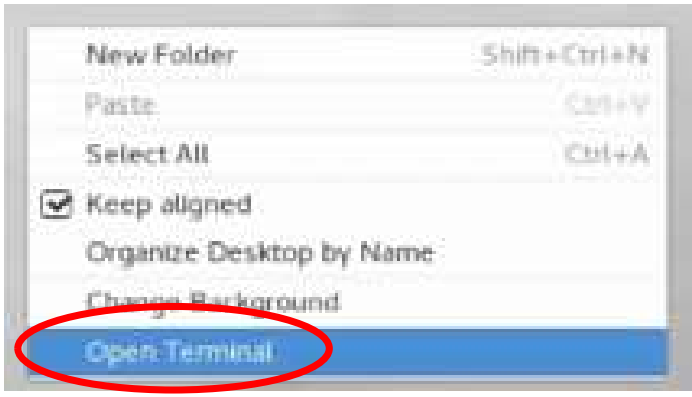




software for science

Launch GDA from a terminal





Start a terminal:

- Activities → Terminal
- Right-click → Open Terminal

In a terminal:

- **gdaclient – start GDA client**
- gdalog – start GDA log display
- Gdaservers – restart GDA servers

This Course - Identifiers

- All visits at Diamond happen under a Proposal – the Proposal for this course is **MX39148**
- You may also bring data from past visits – instructions from Marco
- There will be visits for this course, such as for each of the beamline days (one for I03, one for I04 and one for I04-1), and to give you access to the tutorials etc., which will be numbered -1, -2, -3, -4…

This Course - Identifiers

- mx37045-1 – today's visit for you to practice in & check access (I03)
- mx37045-2 – long visit (6th Nov – 6th Dec) to cover the whole course, with "data access only" for the participants to ensure access
- mx37045-3 – 27th Nov = beamline day on I03
- mx37045-4 – 27th Nov = beamline day on I04
- mx37045-5 – 27th Nov = beamline day on I04-1
- mx37045-6 – 27th Nov = beamline day on I24
- mx37045-7 – 27th Nov = beamline day on I23
- mx37045-8 – 27th Nov = beamline day on VMXm
- Visits for VMXi are actually crystallisation trays, not time-periods so these will be created for this

Data Collection Setup in GDA

- By default, data is stored in folders representing visits, with this structure:
- /dls/{beamline}/data/{year}/{visit number}
  - E.g. /dls/i03/data/2024/MX39148-2

- Then within that the default is:
  - /{proteinacronym}/{samplename}_{run_no}_{image_no}.cbf

- For large visits with multiple Users, or if you have multiple proteins, you can add YourName/ to the left of the folder field
  - E.g. /dls/i03/data/2021/MX28204-5/Felicity/MyProtein/ProteinWT_1_3600.cbf

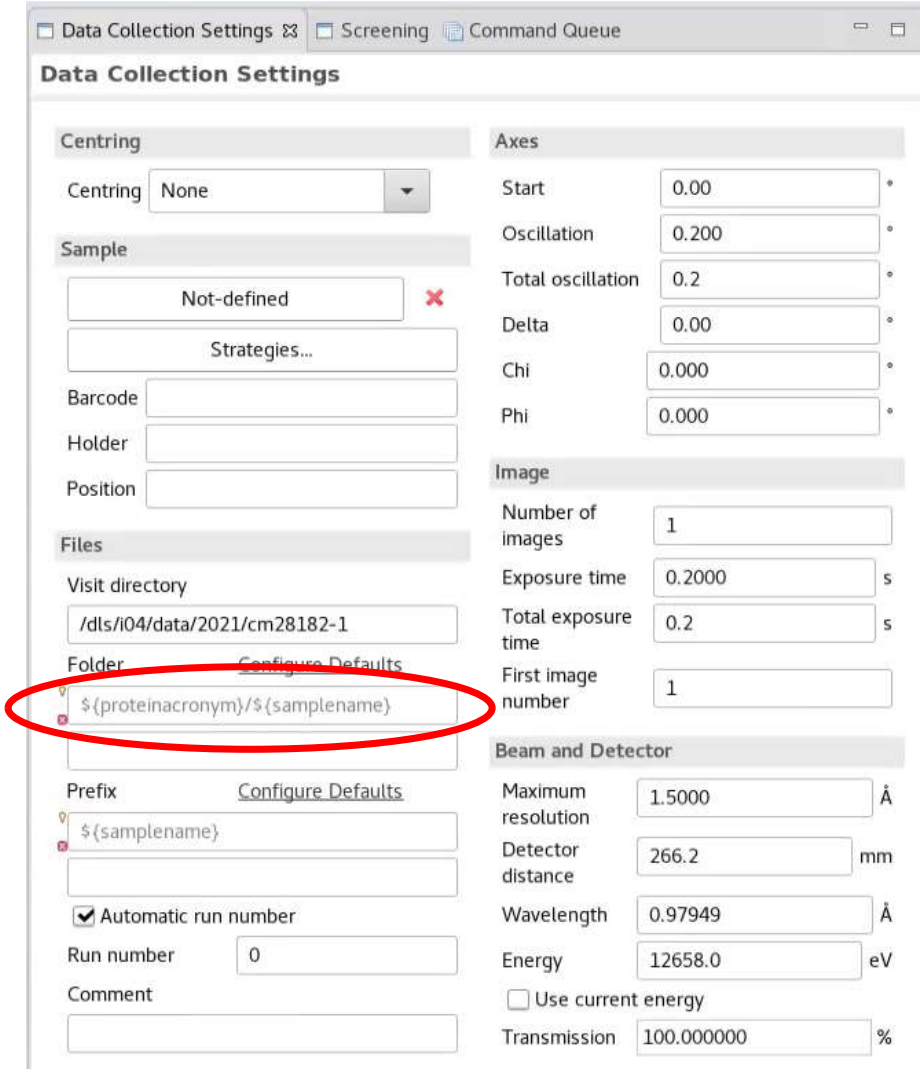
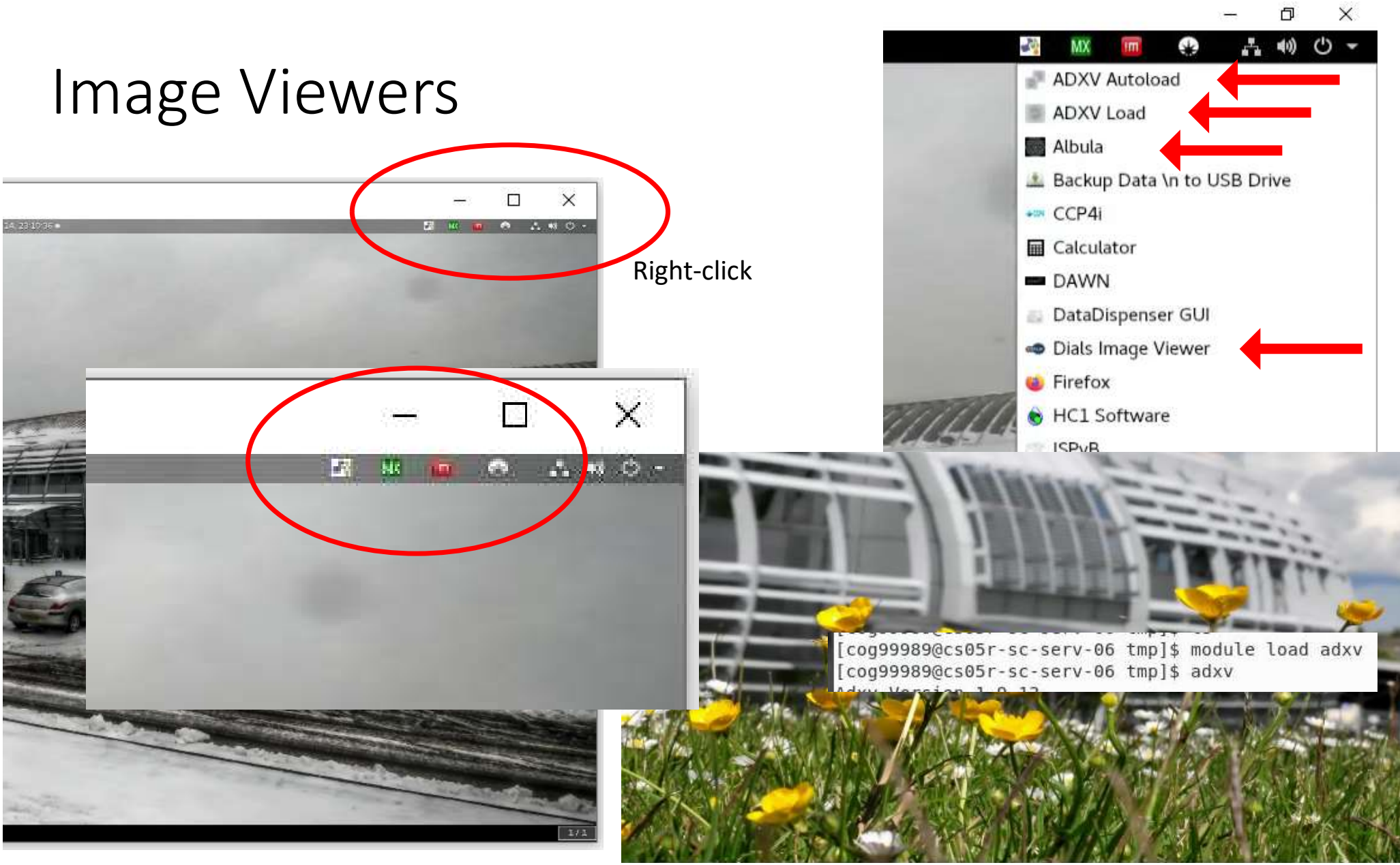
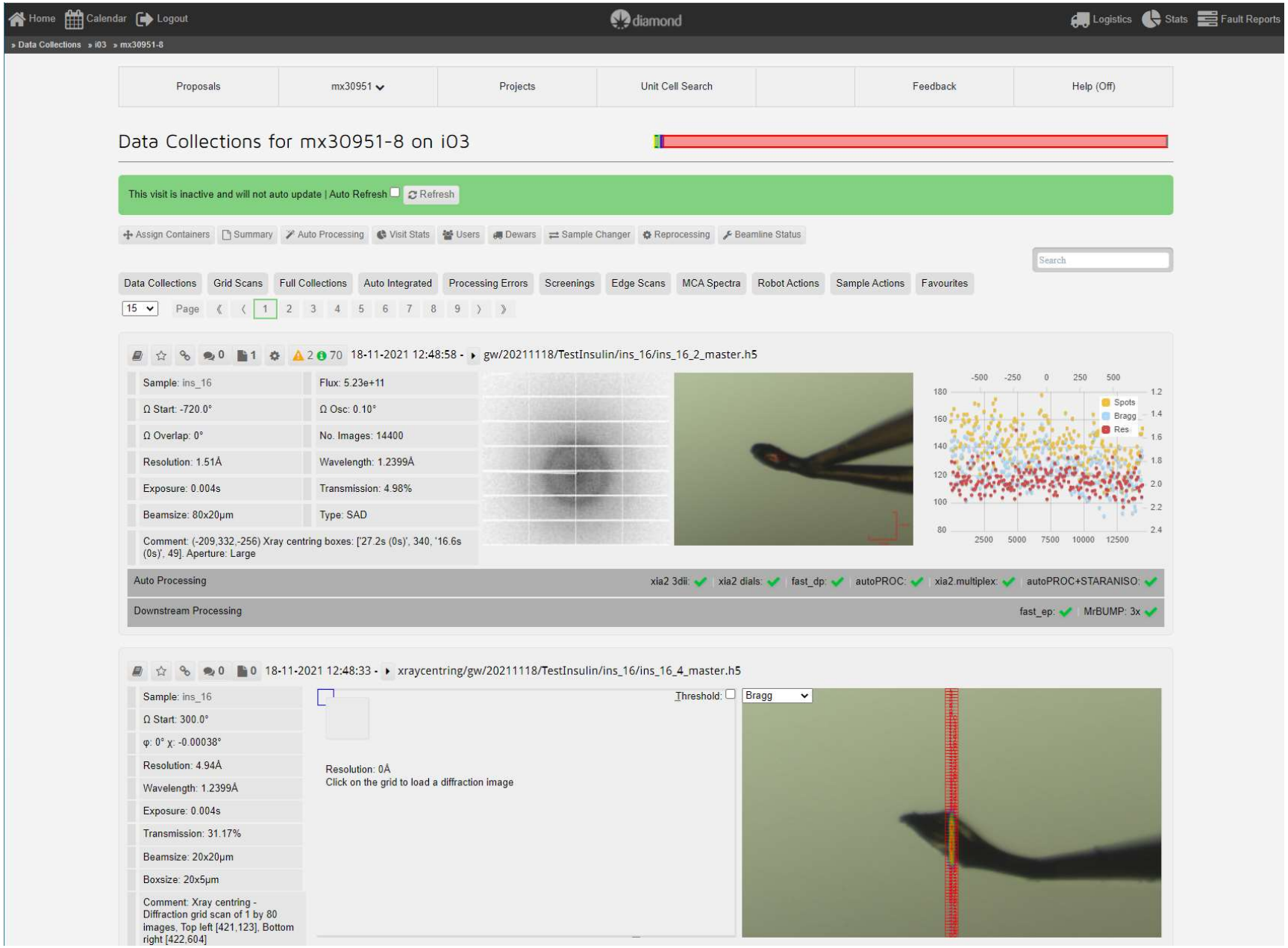


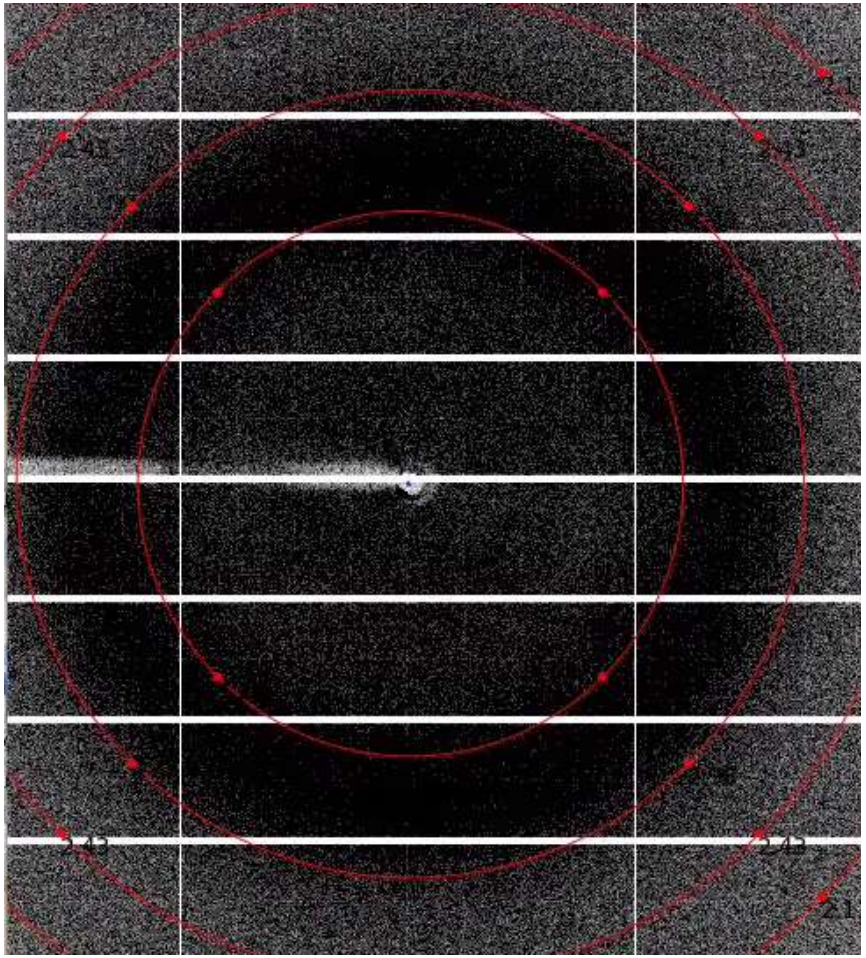
Image Viewers



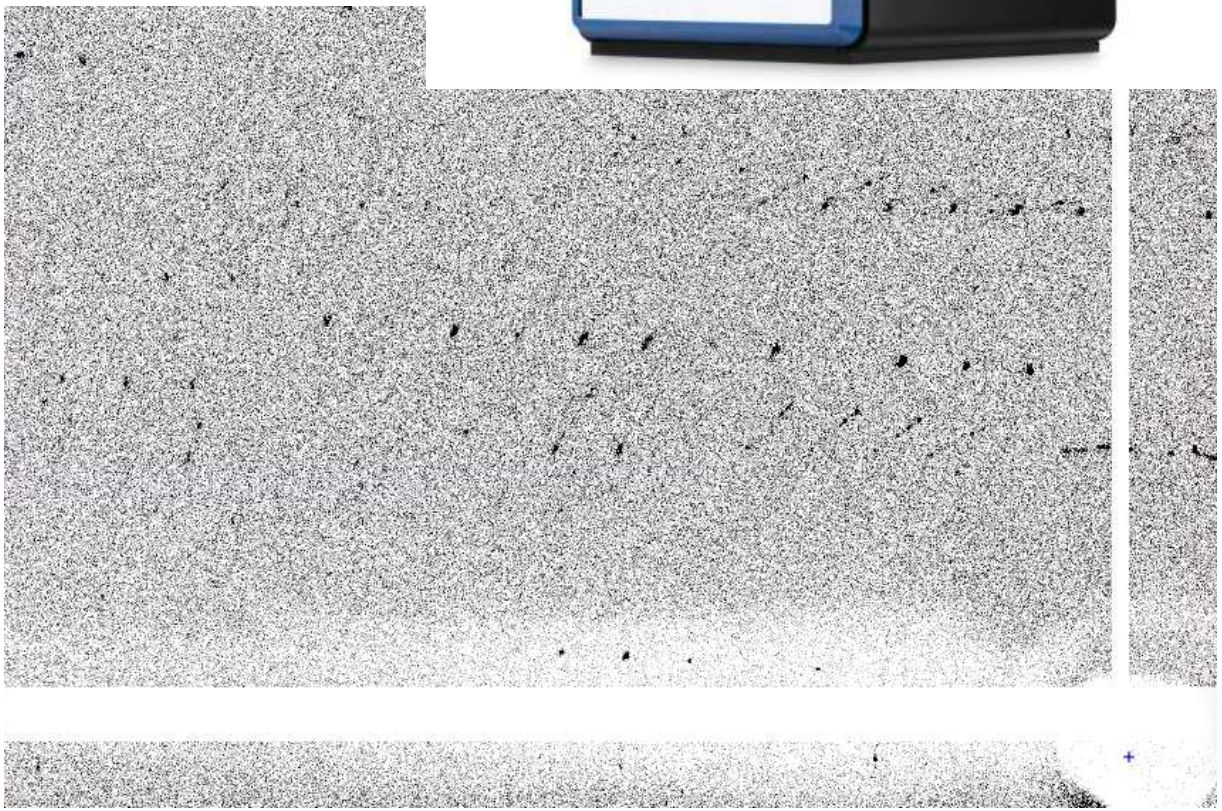
ISPyB View

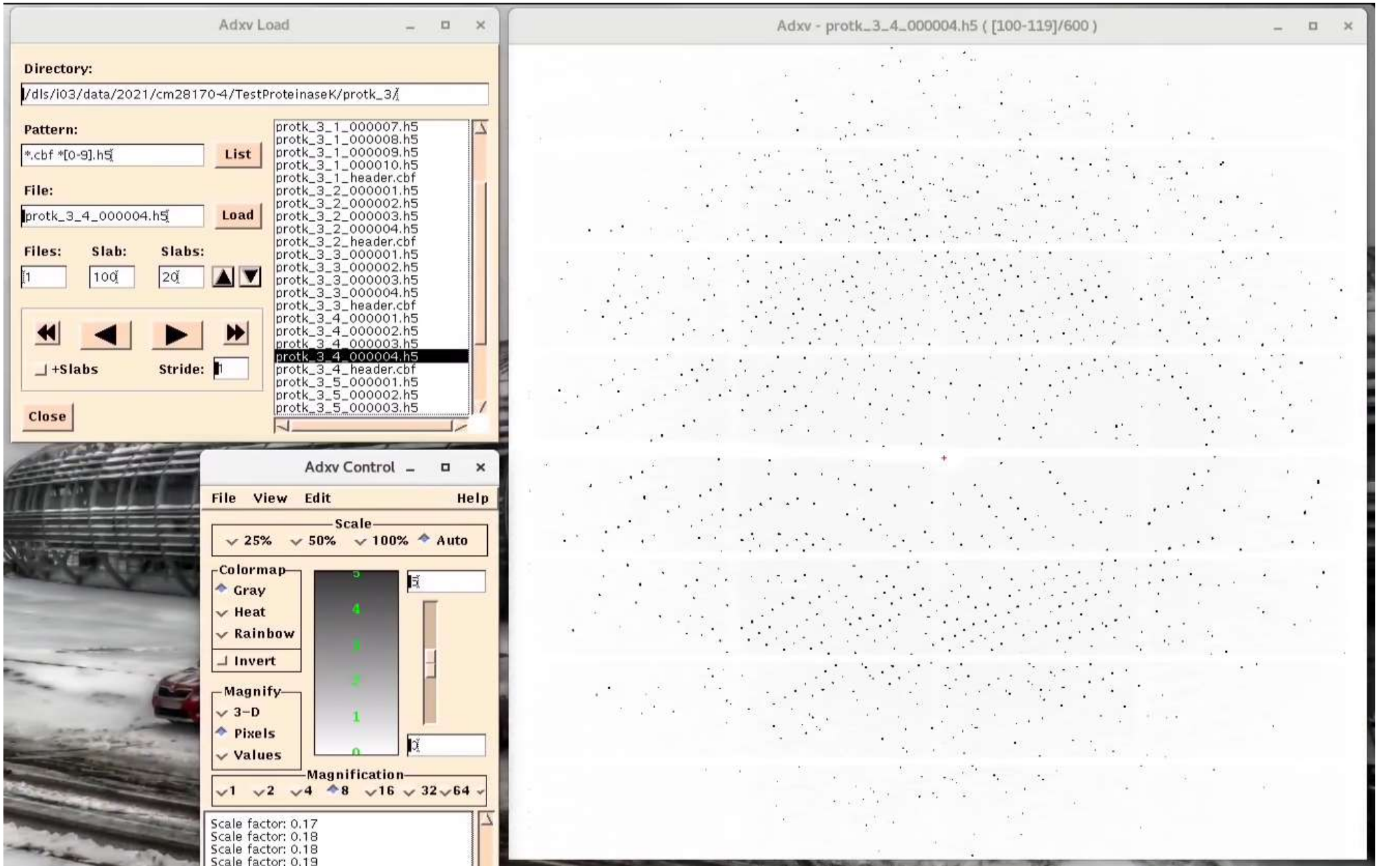

Home
Calendar
Logout
diamond
Logistics
Stats
Fault Reports
» Data Collections » i03 » mx30951-8
Proposals
mx30951
Projects
Unit Cell Search
Feedback
Help (Off)
Data Collections for mx30951-8 on i03
This visit is inactive and will not auto update | Auto Refresh
Refresh
Assign Containers
Summary
Auto Processing
Visit Stats
Users
Dewars
Sample Changer
Reprocessing
Beamline Status
Search
Data Collections
Grid Scans
Full Collections
Auto Integrated
Processing Errors
Screenings
Edge Scans
MCA Spectra
Robot Actions
Sample Actions
Favourites
18-11-2021 12:48:58 - gw/20211118/TestInsulin/ins_16/ins_16_2_master.h5
Sample: ins_16
Flux: 5.23e+11
Ω Start: -720.0°
Ω Osc: 0.10°
Ω Overlap: 0°
No. Images: 14400
Resolution: 1.51Å
Wavelength: 1.2399Å
Exposure: 0.004s
Transmission: 4.98%
Beamsize: 80x20µm
Type: SAD
Comment: (-209,332,-256) Xray centring boxes: ['27.2s (0s)', 340, '16.6s (0s)', 49]. Aperture: Large
Spots
Bragg
Res
Auto Processing
xia2 3dii:
xia2 dials:
fast_dp:
autoPROC:
xia2.multiplex:
autoPROC+STARANISO:
Downstream Processing
fast_ep:
MrBUMP: 3x
18-11-2021 12:48:33 - xraycentring/gw/20211118/TestInsulin/ins_16/ins_16_4_master.h5
Sample: ins_16
Ω Start: 300.0°
φ: 0° χ: -0.00038°
Resolution: 4.94Å
Wavelength: 1.2399Å
Exposure: 0.004s
Transmission: 31.17%
Beamsize: 20x20µm
Boxsize: 20x5µm
Comment: Xray centring - Diffraction grid scan of 1 by 80 images, Top left [421,123], Bottom right [422,604]
Threshold:
Bragg
Resolution: 0Å
Click on the grid to load a diffraction image

Dials Image Viewer or ADXV

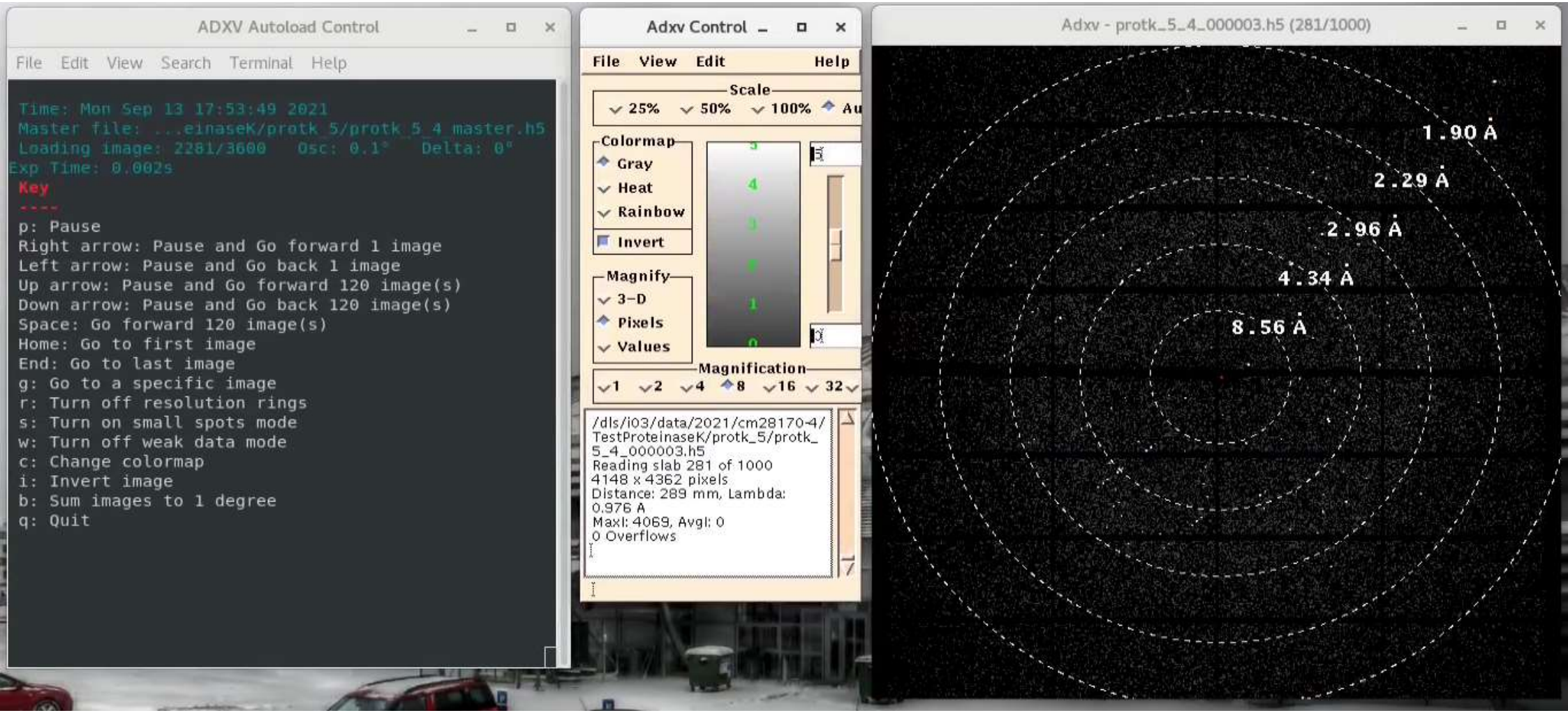


Not ISPyB



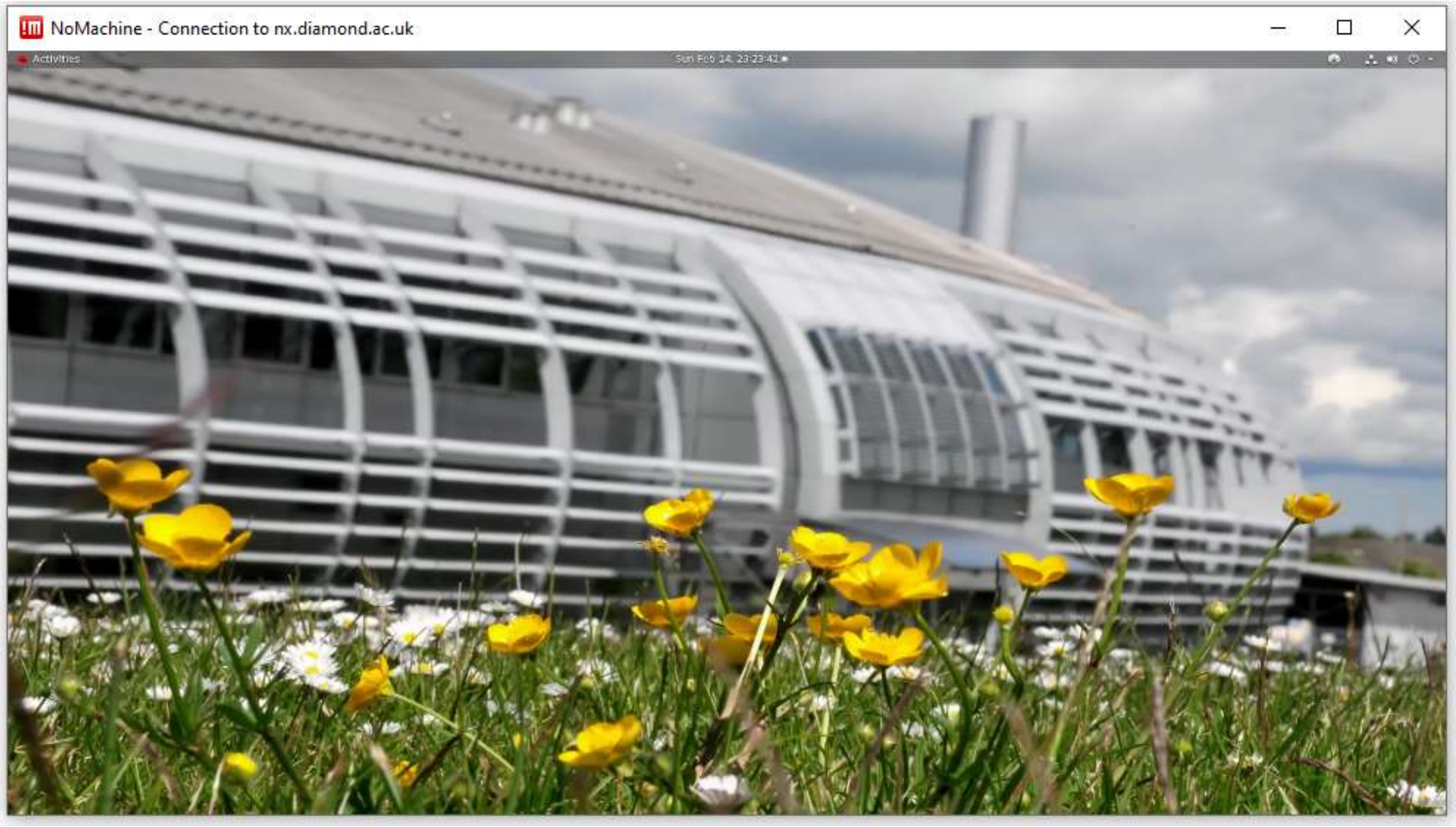

Adxv Load
Directory:
/dls/i03/data/2021/cm28170-4/TestProteinaseK/protk_3/
Pattern:
*.cbf *[0-9].h5
List
File:
protk_3_4_000004.h5
Load
Files:
Slab:
Slabs:
1
100
20
+Slabs
Stride:
1
Close
protk_3_1_000007.h5
protk_3_1_000008.h5
protk_3_1_000009.h5
protk_3_1_000010.h5
protk_3_1_header.cbf
protk_3_2_000001.h5
protk_3_2_000002.h5
protk_3_2_000003.h5
protk_3_2_000004.h5
protk_3_2_header.cbf
protk_3_3_000001.h5
protk_3_3_000002.h5
protk_3_3_000003.h5
protk_3_3_000004.h5
protk_3_3_header.cbf
protk_3_4_000001.h5
protk_3_4_000002.h5
protk_3_4_000003.h5
protk_3_4_000004.h5
protk_3_4_header.cbf
protk_3_5_000001.h5
protk_3_5_000002.h5
protk_3_5_000003.h5
Adxv - protk_3_4_000004.h5 ([100-119]/600)
Adxv Control
File
View
Edit
Help
Scale
25%
50%
100%
Auto
Colormap
Gray
Heat
Rainbow
Invert
Magnify
3-D
Pixels
Values
5
4
3
2
1
0
Magnification
1
2
4
8
16
32
64
Scale factor: 0.17
Scale factor: 0.18
Scale factor: 0.18
Scale factor: 0.19

ADXV Autoload



Remote Data Processing

Remember: Work here, not on a beamline workstation



DON’T PROCESS DATA IN YOUR HOME DIRECTORY – IT IS TINY

WHEN IT GETS FULL YOU WON’T BE ABLE TO LOG IN TO DELETE THE DATA THAT IS STOPPING YOU LOGGING IN AND IT CAUSES PROBLEMS

Error: Cannot create session directory: /home/abc12345/.nx/node/C-cs05r-sc-serv-08.diamond.ac.uk-1023-CC896785C4A9B79013D614E113768B6B

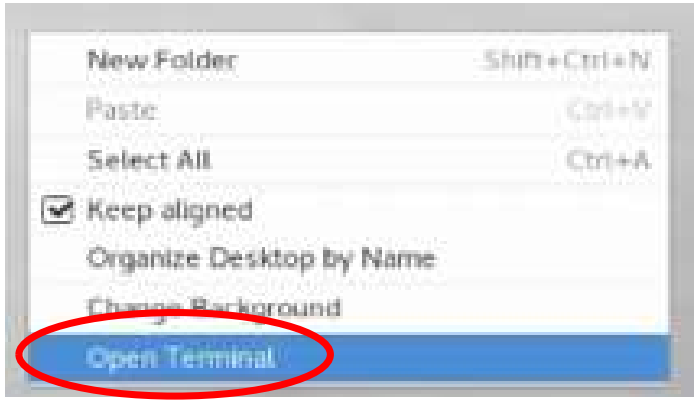
Error is: Disk quota exceeded

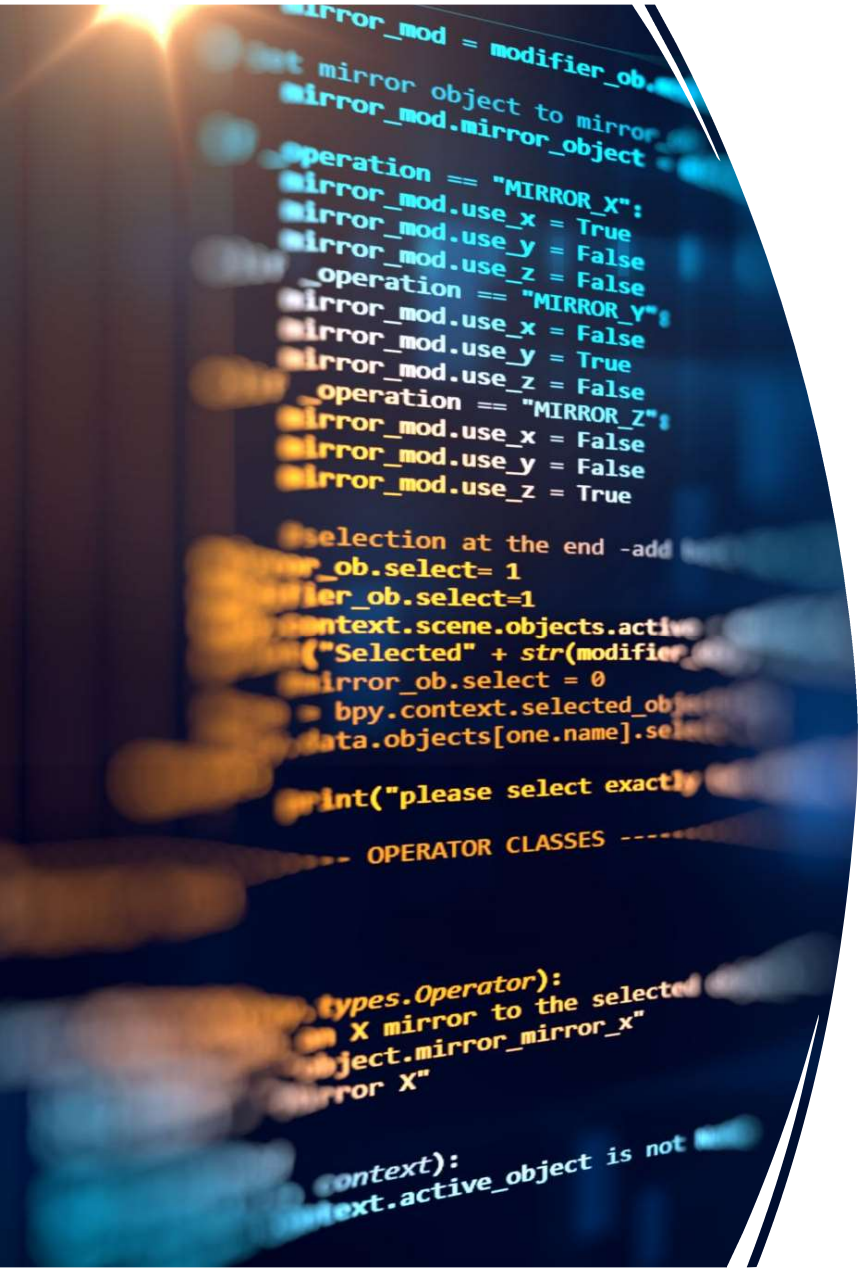
Launch a Terminal

Start a terminal:

- Activities → Terminal
- Right-click anywhere (on beamline)

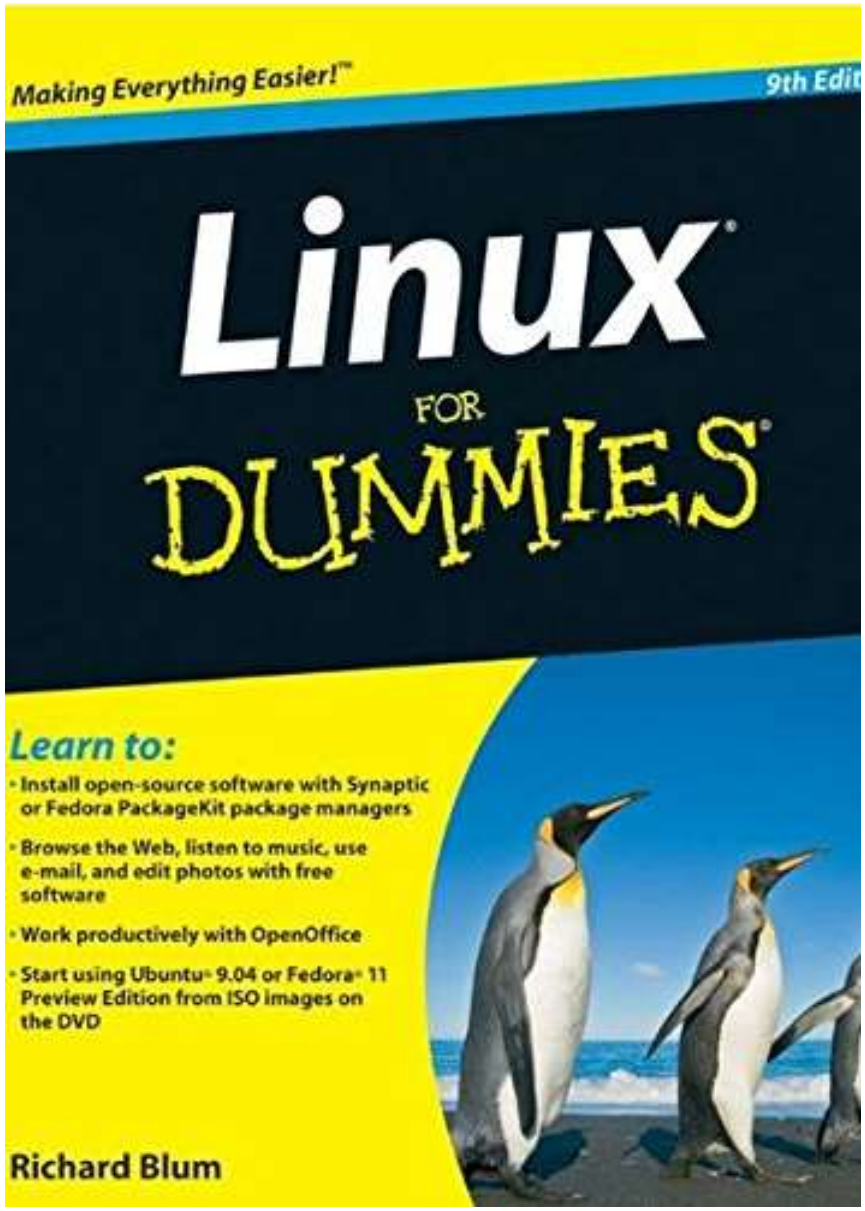
  → Open Terminal





Using the Command Line

- The Command line is more powerful than a GUI, and there will always be a Command line (not everything is packaged into a GUI)
- You can build on basic Command line work with scripts to run tasks more efficiently
- BUT get one character wrong and it won’t work…



redhat
Linux

udemy

“Folders” = “Directories”

Linux Cheat-Sheet – on Slack

- man – gives manual
  for command

- Grab someone for a tutorial if you need it

Linux cheat sheet and Diamond working directory

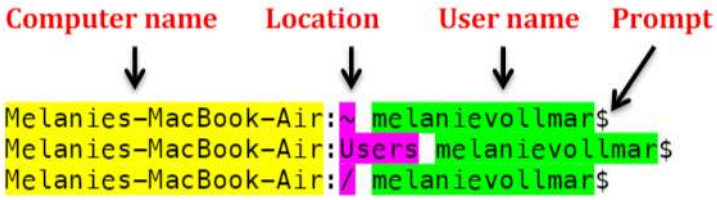
NOTE: do NOT use spaces in file names

NOTE: commands are case sensitive

NOTE: Do not work in your home directory: /home/my_fedid !!!

NOTE: Location in file system

module load ccp4-workshop



Command	options	What it does
Help		
man	ls	Loads manual page for ls command
Navigation		
pwd		Print working directory → tells you where you are; which folder you are in
cd		Change directory; and gets

Linux - Continued

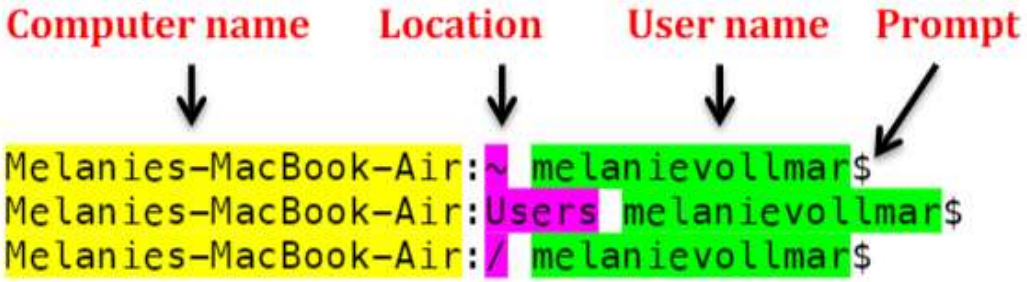
NOTE: do NOT use spaces in file names

NOTE: commands are case sensitive

NOTE: Do not work in your home directory: /home/my_fedid !!!

NOTE: Location in file system

module load ccp4-workshop



Linux man = Manuals

- Example **man ls**
- **Enter** = move down 1 line
- **Space** = move down 1 page
- **g** = move to the top of the page
- **G** = move to the bottom of the page (one of the few times you use a capital)
- **Q** = quit

Linux tree

- tree –d
- tree

cog99989@cs05r-sc-serv-07:/dls/i04/data/2021/cm28182-1/TestInsulin

File Edit View Search Terminal Help

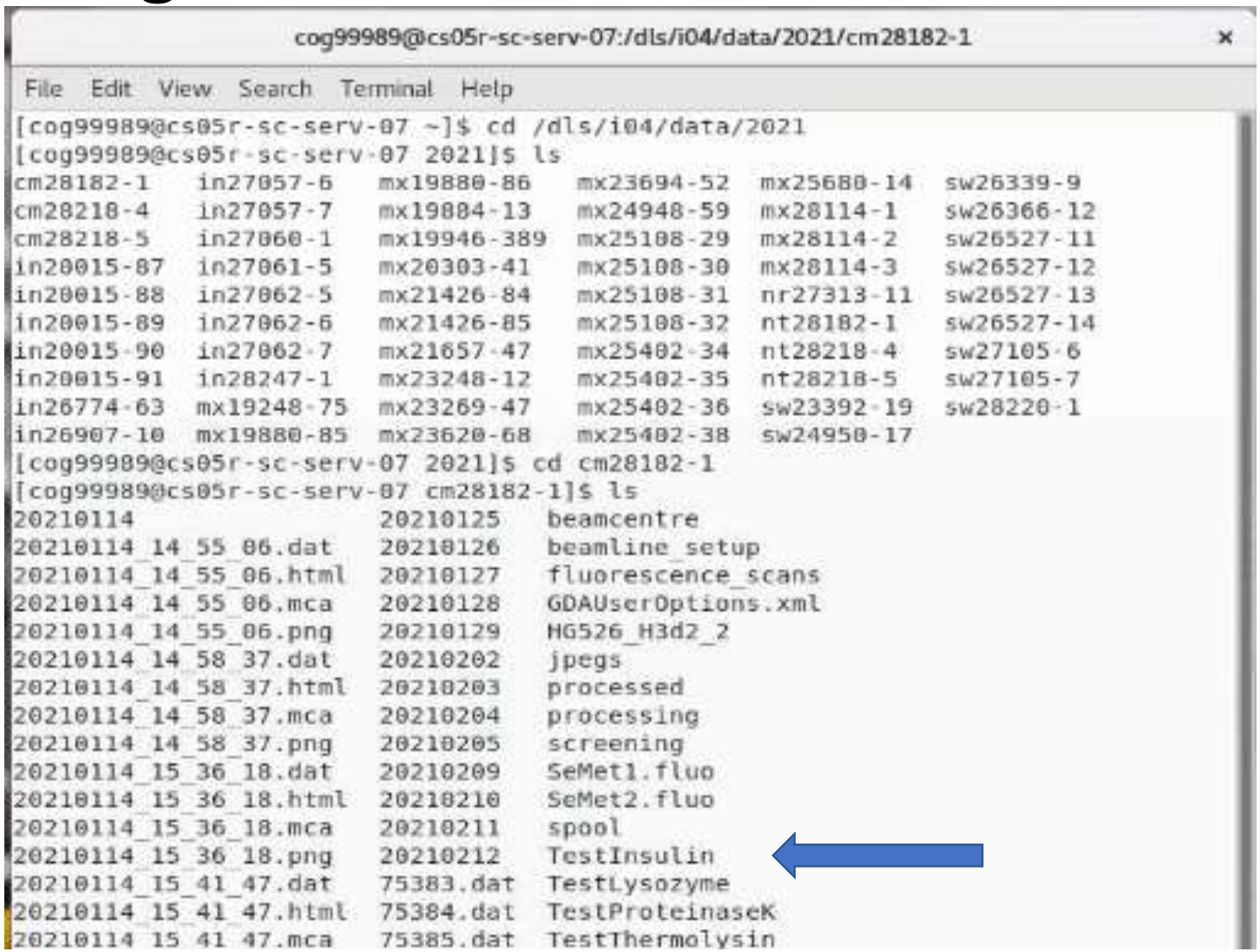
```
[cog99989@cs05r-sc-serv-07 TestInsulin]$ tree -d
.
├── ZincInsulinB3
├── ZincInsulinB4
└── ZincInsulinB5

3 directories
[cog99989@cs05r-sc-serv-07 TestInsulin]$ tree
.
├── ZincInsulinB3
│   ├── ZincInsulinB3_1_000001.h5
│   ├── ZincInsulinB3_1_000002.h5
│   ├── ZincInsulinB3_1_000003.h5
│   ├── ZincInsulinB3_1_000004.h5
│   ├── ZincInsulinB3_1_header.cbf
│   ├── ZincInsulinB3_1_master.h5
│   ├── ZincInsulinB3_1_meta.h5
│   ├── ZincInsulinB3_1.nxs
│   ├── ZincInsulinB3_2_000001.h5
│   ├── ZincInsulinB3_2_000002.h5
│   ├── ZincInsulinB3_2_000003.h5
│   └── ZincInsulinB3_2_000004.h5
```

Linux pwd

- pwd = Print Working Directory
- By default you start in your home directory **BUT DON'T WORK IN IT** (apart from the basic Linux tutorial)
- cd /dls/{beamline}…

Navigating to data



Deleting files – rm (ReMove)

- Proceed with caution
  - there is no recycle bin or undo
- rm file.img
  - Command will delete file.img

- You can also inadvertently delete files by over-writing them with a new file with the same name

Permissions

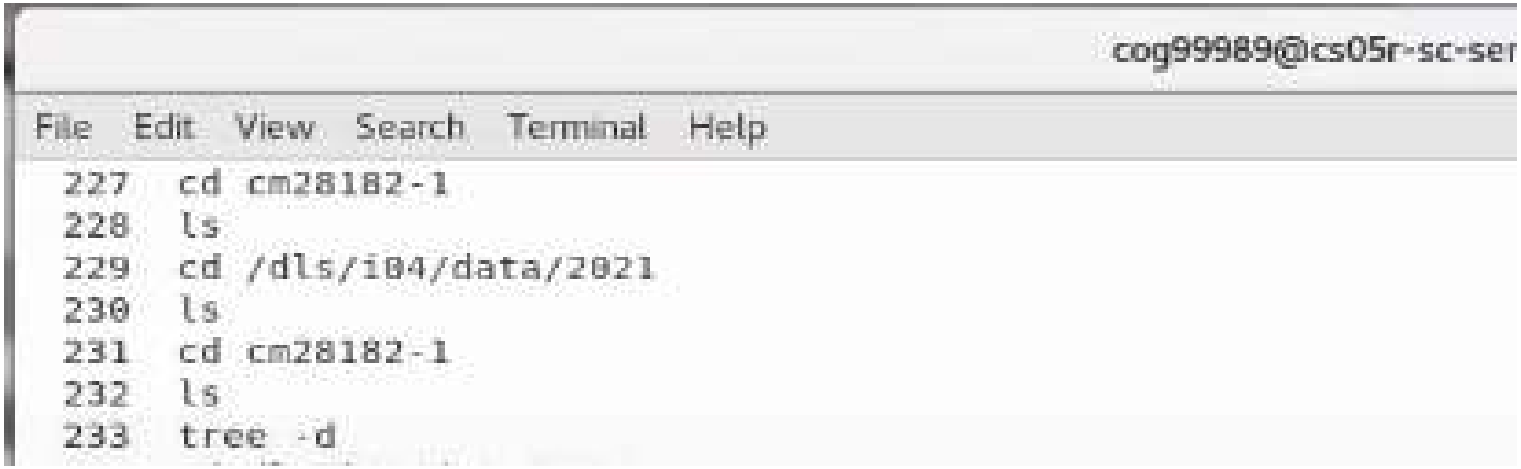
- ls –l

```
-rw-rw-r--. 1 cog99989 cog99989     74 Sep 18  2020 bashrc_local
drwxrwxr-x. 4 cog99989 cog99989   4096 Sep 23 15:04 cs-studio
drwxr-xr-x. 3 cog99989 cog99989   4096 Dec  4  2019 DepositFiles
drwxr-xr-x. 7 cog99989 cog99989   4096 Nov 21 09:25 Desktop
drwxr-xr-x. 3 cog99989 cog99989   4096 Aug 18 10:10 Documents
drwxr-xr-x. 2 cog99989 cog99989   4096 Aug 11 14:33 Downloads
drwxrwxr-x. 5 cog99989 cog99989   4096 Mar 10  2020 eclipse-workspace
```

- First character = file (-) or directory (d)
- Next 3 characters – permission of the owners
  - r = read permission
  - w = write permission
  - x = execute permission
- Characters 5-7 – permission of the group
- Characters 8-10 – permission of others

History

- history = gives history of most recent commands
  - Great if someone has shown you something, and you have forgotten what process they followed to do it!



Linux Tutorial – Navigating the Diamond Cluster

AIMS

The aim of this tutorial is to get a novice Linux User up to speed at navigating the Diamond cluster and getting started with data processing. Feel free to skip over anything that you already know and remember that there is often more than one way to achieve the same process.

Remember that commands are case-sensitive and are almost always in lower-case. In this tutorial, anything that is for you to type is written in bold such as the **tree** command here.

Please experiment! (we are scientists after all) – if you accidentally set off something that takes a long time, all you need to do is **Ctrl + C** and that will halt the current command. If you can suggest ways to improve this tutorial, please don't hesitate to contact me – felicity.bertram@diamond.ac.uk

TO START

Once in an NX session, open a terminal, either by:

- Clicking on "Activities" in the top left corner of the screen, and then clicking on the Terminal icon which is the 5th option down (a black box with a >_ in it – hover over an icon and it will tell you what it is
- right-clicking anywhere on the background screen, and selecting "Open Terminal" from the menu that appears – though this option may only work on beamline computers

See my slides from my presentation for more information about how you do this.

FINDING YOUR DATA

To find your data, you need to navigate away from your home directory. To do this, you will need to use the "Change Directory" command, which is:

cd

In its simplest form, you can just navigate to where your data is, in this instance by typing:

cd /dls/i04-1/data/2021/mx29507-1 (note there is a space between the cd and the file path)

However, you will come across much more complex paths than this, and if you type one character wrongly, it won't work, and will return a "No such file or directory" message. Therefore, you want to be able to explore folder by folder. To move up one folder, in this instance from /home/{your Fed ID} you need to type:

cd .. (once again there is a space between the cd and the ..)

If you then use the **pwd** command again, it will tell you that you are now in /home (go on, try it). If you want to see what is in the /home folder then you can use the **ls** command once again (in this instance it will list all Fed IDs currently using that node). We want to go up again, so type **cd ..** once again. This should mean that when you use the **pwd** once again, the location is simply / This is also called "root" as it is the top of the directories.

Remembering that the path for any data collected at Diamond is /dls/{beamline}/data/{year}/{visit number} you now want to navigate into /dls. To do this type:

cd dls (cd /dls also works)

pwd → this will show that you are in /dls now, not /

ls → this will show what folders there are in /dls – most are a list of lots of beamlines, not just MX

AUTOCOMPLETE

To set yourself up for the next step, type **cd ..** to take you back to /dls/i04-1/data/2021

When you are navigating, if you start to type an option, you can do a few letters, and then press tab to autocomplete. So if you start typing:

cd mx29 and then press **[Tab]**

It will start to autocomplete – but I've set this up as a trick question as there are two proposals in that folder which start *mx29* – they are *mx29502-2* (not yours) and *mx29507-1* (yours). You can see these in the folder by using the **ls** command. Therefore the autocomplete will only take you as far as **cd mx2950** so if you hit **[Enter]** after that

CLEAR

If your terminal becomes clogged up with all this practicing, then you can wipe the slate clean with:

clear

You will find everything has vanished, and you can start again. But not from the beginning, as if you do **pwd** again, you will find that you are still in the folder that you had just navigated to. You just don't have all the mess of switching around.

HISTORY

This command is particularly useful if someone has just shown you how to do something, and you need more time to make notes on it. Or if you have just used the **clear** command and want to check back on what you just did. Typing the command:

history

Produces a list of the history of commands on that session. Note that if you are working hard, you'll end up with a very long list of commands. Therefore, you can use:

history 10

This gives you the most recent 10 commands (you can use any number here, where I have used 10)

Retrieve your data and take it home

- Better to download it within 40 days

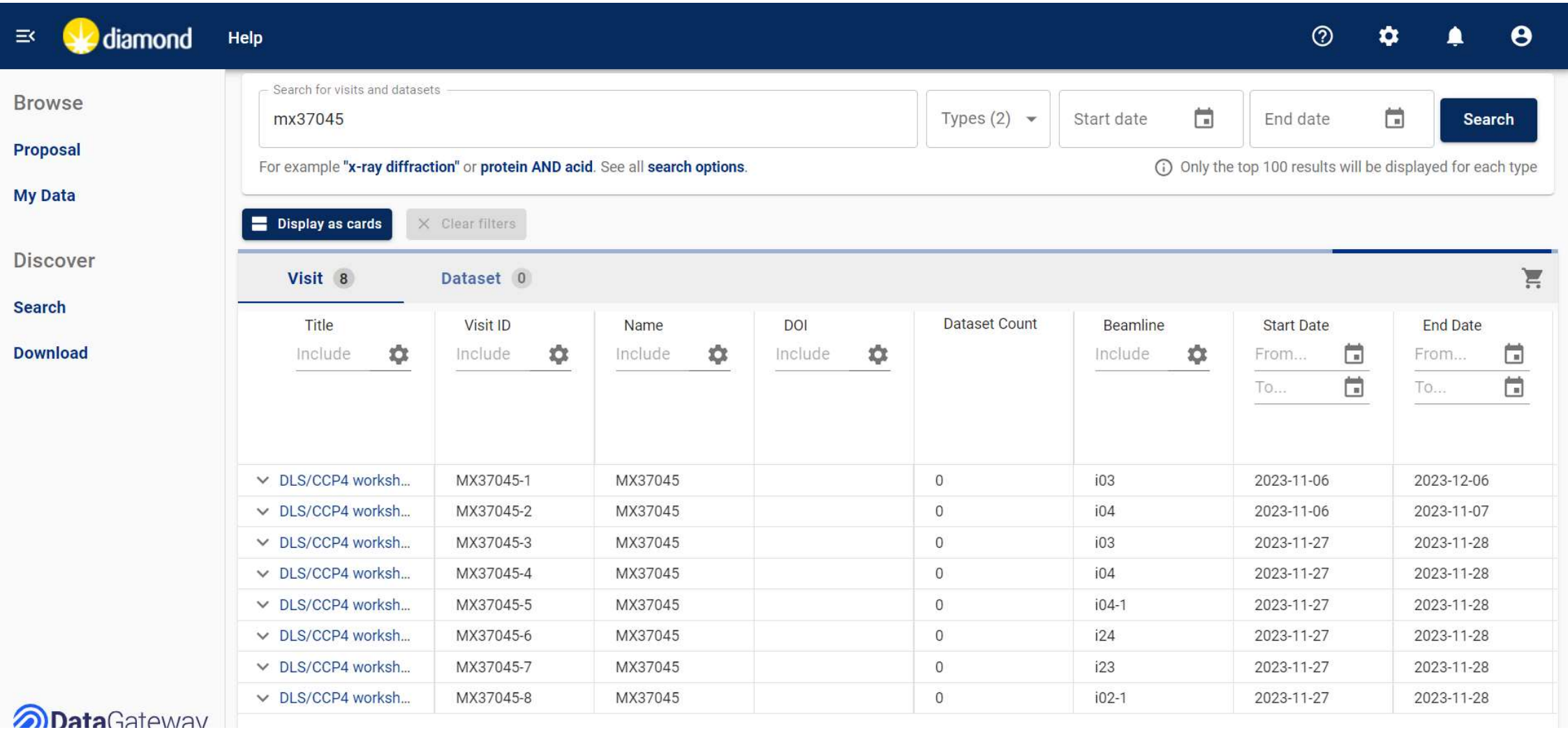
- **Globus** – main option, datasets over 20 GB
- **FTP** – File Transfer Protocol, only for datasets under 20 GB
- **SFTP** – Secure File Transfer Protocol, if your institution firewall blocks FTP, only for datasets under 20 GB
- **rsynch** - Good for Unix-like systems like Linux - only recommended for experienced rsync users

- https://www.diamond.ac.uk/Users/Experiment-at-Diamond/IT-User-Guide/Not-at-DLS/Retrieve-data.html

Retrieve your data and take it home

- After 40 days…
- Your data will now be in the archive and need to be restored from tape
- Download files using Diamond's TopCAT software
- https://www.diamond.ac.uk/Users/Experiment-at-Diamond/IT-User-Guide/Not-at-DLS/Retrieve-data.html
- Note that files will be found here instead of the normal link:
  - /dls/staging/dls/$instrument/data/$year/$visit
  - Not /dls/{beamline}/data/{year}/{visit number}

Globus


diamond
Help
Browse
Proposal
My Data
Discover
Search
Download
DataGateway
Search for visits and datasets
mx37045
Types (2)
Start date
End date
Search
For example "x-ray diffraction" or protein AND acid. See all search options.
Only the top 100 results will be displayed for each type
Display as cards
Clear filters
Visit 8
Dataset 0
Title
Visit ID
Name
DOI
Dataset Count
Beamline
Start Date
End Date
Include
From...
To...
DLS/CCP4 worksh... MX37045-1 MX37045 0 i03 2023-11-06 2023-12-06
DLS/CCP4 worksh... MX37045-2 MX37045 0 i04 2023-11-06 2023-11-07
DLS/CCP4 worksh... MX37045-3 MX37045 0 i03 2023-11-27 2023-11-28
DLS/CCP4 worksh... MX37045-4 MX37045 0 i04 2023-11-27 2023-11-28
DLS/CCP4 worksh... MX37045-5 MX37045 0 i04-1 2023-11-27 2023-11-28
DLS/CCP4 worksh... MX37045-6 MX37045 0 i24 2023-11-27 2023-11-28
DLS/CCP4 worksh... MX37045-7 MX37045 0 i23 2023-11-27 2023-11-28
DLS/CCP4 worksh... MX37045-8 MX37045 0 i02-1 2023-11-27 2023-11-28

Any questions…?