Australian Synchrotron
Turning bright ideas into brilliant outcomes

eResearch Tools for the Australian Synchrotron User Community

Uli Felzmann
Australian Synchrotron

ICALEPCS 2013 – San Francisco

AUSTRALIAN SYNCHROTRON

Located at 800 Blackburn Road, Clayton, Victoria (Melbourne)

2015 ICALEPCS
melbourne • australia

Supported by
Australian Government
State Government Victoria

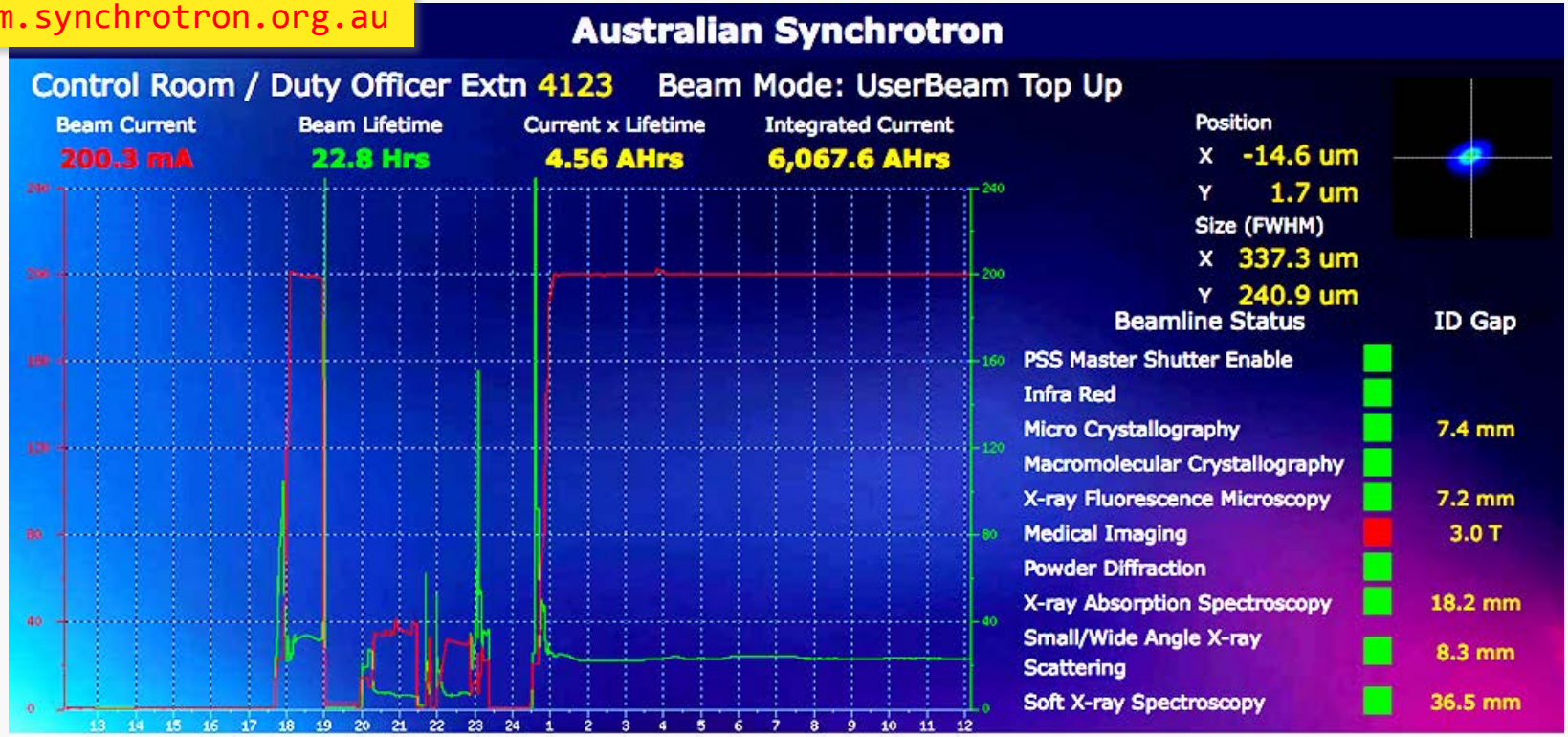
AUSTRALIAN SYNCHROTRON

- Australia's largest scientific user facility
- Opened for business 2007
- Currently 132 staff (~127 FTEs)
- Supports a broad range of research, with applications in sectors from medicine and nanotechnology to manufacturing and mineral exploration
- Involved in the generation of more than 1000 pre-reviewed publications in refereed journals (average impact factor 4.71)
- Since 2012 operated by ANSTO (www.ansto.gov.au) under the name **Synchrotron Light Source Australia** (SLSA)

ACCELERATOR SCIENCE

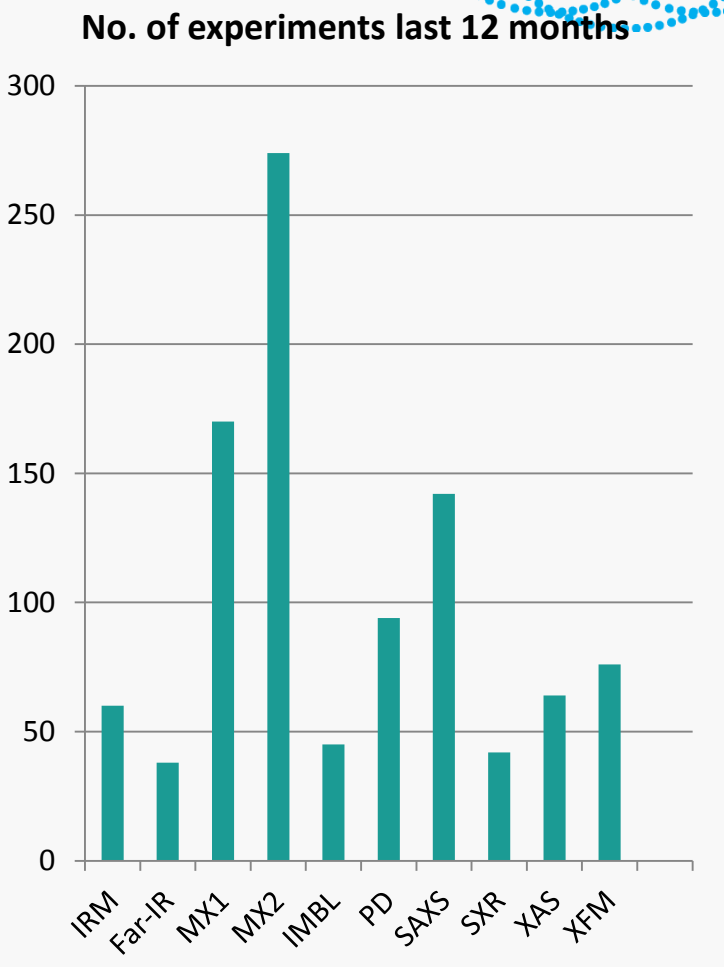
- 3rd generation light-source: 100 MeV linac, 100 MeV to 3 GeV booster, 3 GeV storage ring (216 m circumference)
- Since May 2012, storage ring is topped up to 200 mA every ~ 3.5 minutes
- 5000 h/year of stored beam since 2007, availability > 99% in last 3 years

fsm.synchrotron.org.au



BEAMLINES

- 9 experimental stations (beamlines) with 10 instruments in total
  - Imaging and medical beamline (IMBL)
  - Infrared microspectroscopy (IRM)
  - Far-infrared and high resolution FTIR (IRHR)
  - Macromolecular crystallography (MX1)
  - Micro Crystallography (MX2)
  - Powder diffraction (PD)
  - Small/Wide Angle X-ray Scattering (SAXS)
  - Soft X-ray spectroscopy (SXR)
  - X-ray absorption spectroscopy (XAS)
  - X-ray fluorescence microprobe (XFM)



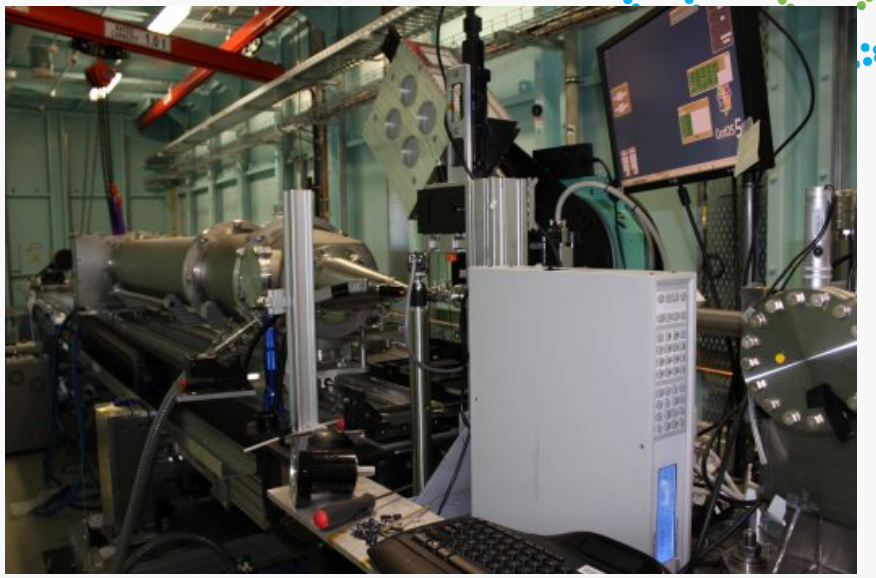
HIGH PERFORMANCE COMPUTING

- Multi-modal Australian ScienceS Imaging and Visualisation Environment (www.massive,org.au)
- Unique Australian HPC facility with a focus on fast data reduction and large-scale visualisation
- Two high performance computing facilities, located at the Australian Synchrotron and Monash University
- MASSIVE1 cluster, located at the Australian Synchrotron:
  - 42 nodes with 12 cores per node running at 2.66GHz (504 CPU-cores total)
  - 48 GB RAM per node (2,016 GB RAM total)
  - 2 nVidia M2070 GPUs with 6GB GDDR5 per node (84 GPUs total)
  - 150 TB of fast access parallel file system (GPFS)
  - 4x QDR Infiniband Interconnect
  - Windows 2008 HPC + CentOS 5.x
  - Real-time + post-experiment data analysis
  - Available to all Australian Synchrotron users

eRESEARCH TOOLS FOR THE AUSTRALIAN SYNCHROTRON RESEARCH COMMUNITY

- Currently concluding an 18-month programme of software development to increase users' ability to achieve scientific outcomes at the Australian Synchrotron
- Funded by NeCTAR (http://www.nectar.org.au/)
  - Australian Government project conducted as part of the Super Science initiative
  - National eResearch Collaboration Tools and Resources
  - 4 programs, all aiming at building (ICT) research infrastructure (2012 - 2014) to increase scientific output
    (ICT = Information and communication technologies)
  - Providing a national Research Cloud for Australian researchers
    - Federated approach, 8 nodes around Australia
    - Based on OpenStack
    - 2nd biggest research cloud in the world
- Technique-specific workflows providing automatic data-processing
- Extending and enhancing remote data analysis capabilities
- 4 dedicated software developers embedded in beamline teams

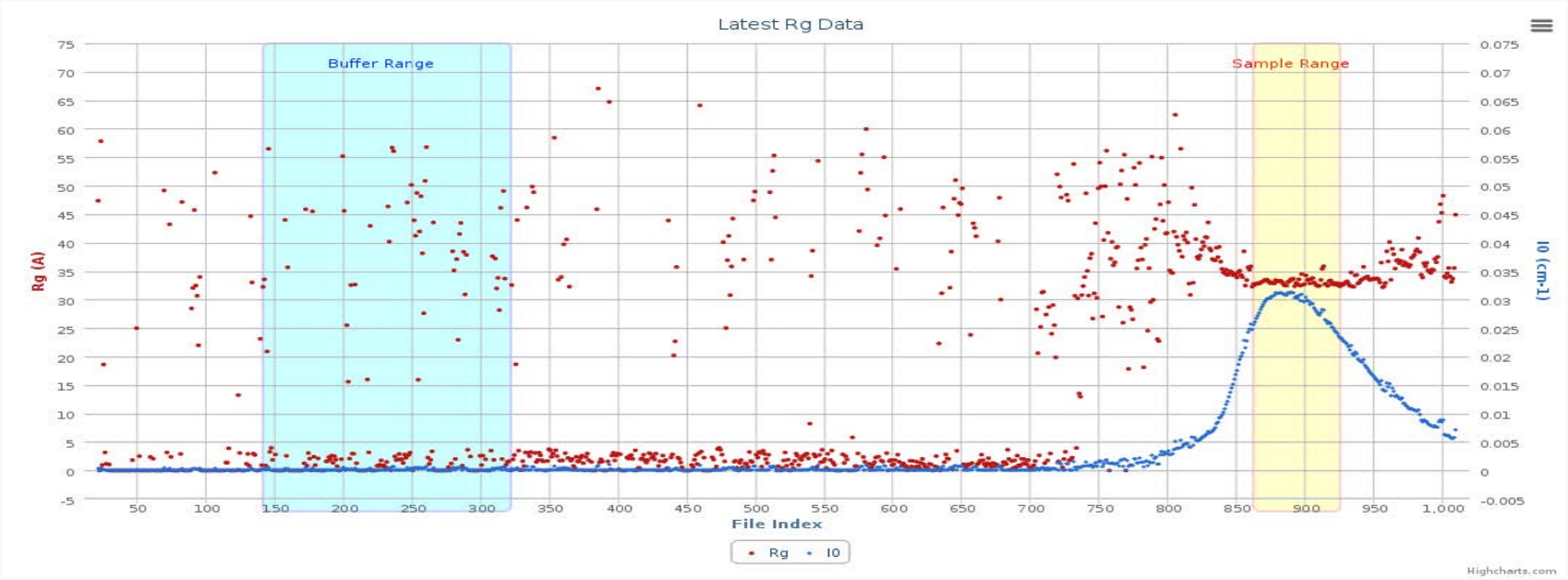
SMALL AND WIDE ANGLE SCATTERING (SAXS/WAXS)



- Highly flexible undulator beamline used for x-ray scattering analysis of a diverse range of solids, fluids and surfaces across a wide range of research fields
- Two Pilatus detectors are in operation for SAXS (1M detector) and WAXS (200k detector)
- May be run concurrently with excellent dynamic range, low noise and short exposures with up to 30 and 150 frames per second
- Data collection and reduction is normally performed using the ScatterBrain IDL program developed in house by the beamline team
  www.synchrotron.org.au/index.php/aussyncbeamlines/saxswaxs/software-saxswaxs

AUTOMATED REAL-TIME SEC-SAXS

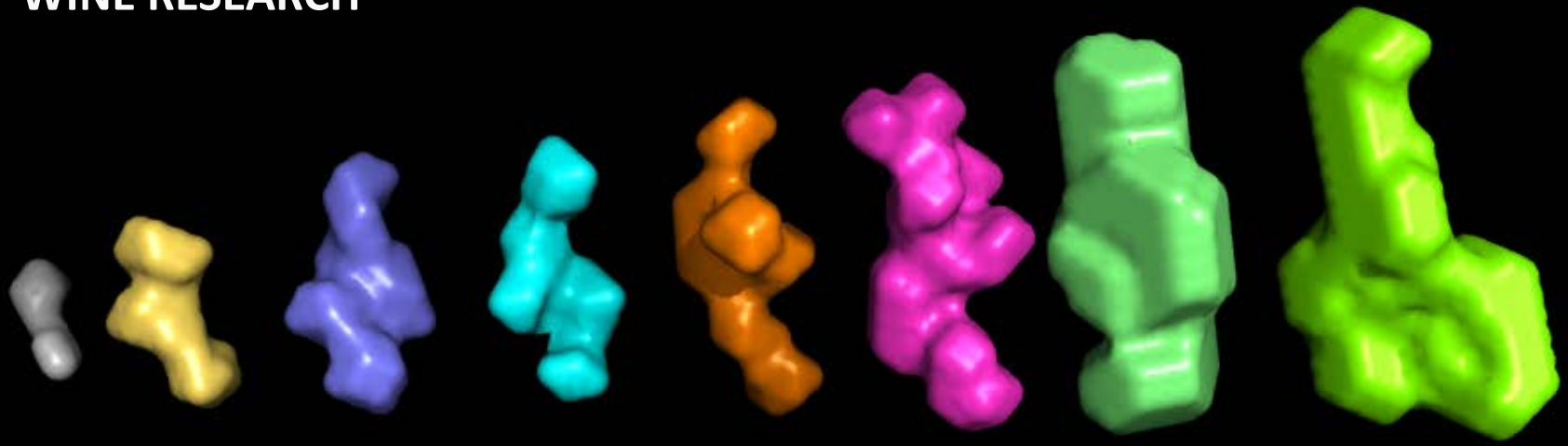
- Automated data processing and analysis pipeline has been developed for solution scattering in order to deal with the large numbers of samples
- Average from first 20 patterns (blanks) is subtracted from rolling 5-pattern averages
- Guinier analysis is performed to obtain the intensity at zero q (I0) and radius of gyration (Rg) and results are displayed graphically in a browser in real time
- At the end of an SEC run, Users can graphically redefine the analysis
- Straightforward to obtain reprocessed subtracted individual and averaged elution peak patterns



AUTOMATED REAL-TIME SEC-SAXS

- Ab initio shape determination with the ATSAS data analysis software (http://www.embl-hamburg.de/biosaxs/software.html)
- Providing real-time results during an experiment
- Easing the otherwise laborious task of post-experiment processing large data sets.

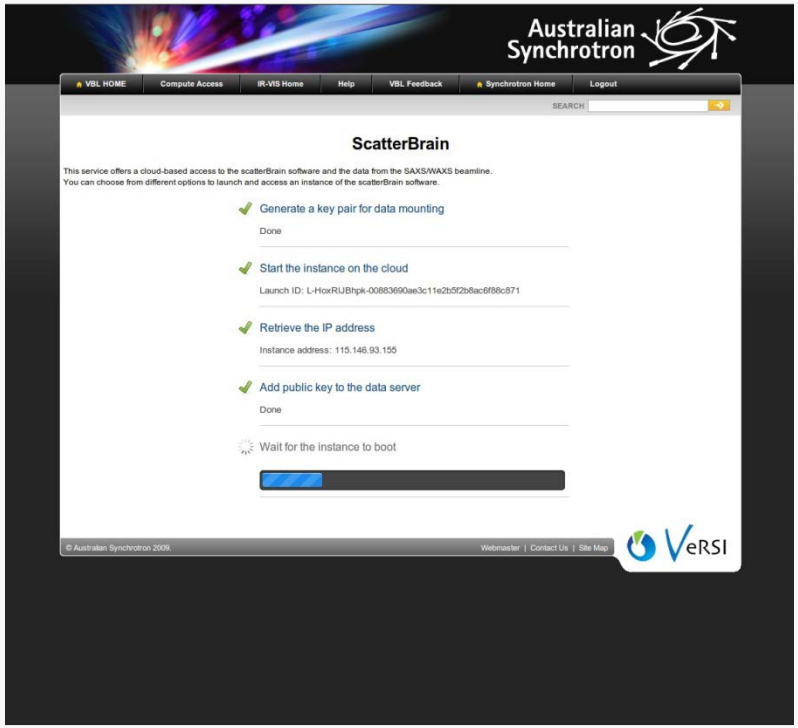
WINE RESEARCH



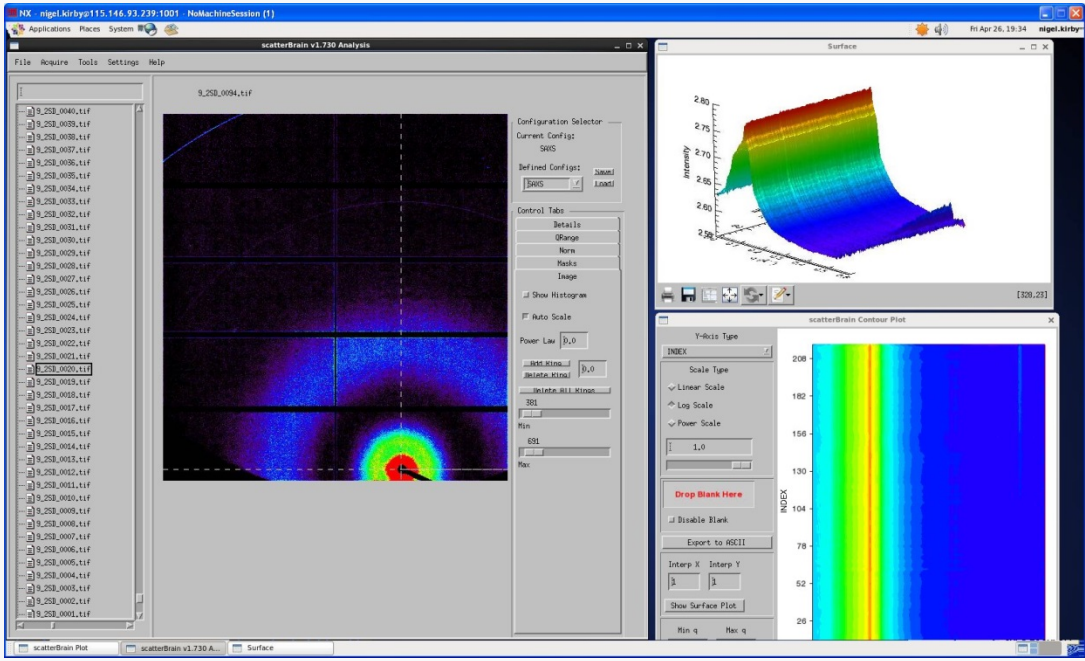
8 different tannin fractions from monomer, dimer ...continuously up to 8-mer (i.e. tannins containing 1,2,3,...8 subunits)

SAXS REMOTE DESKTOP

scatterCloud



Time series for an in-situ thermal treatment experiment showing development of an additional phase

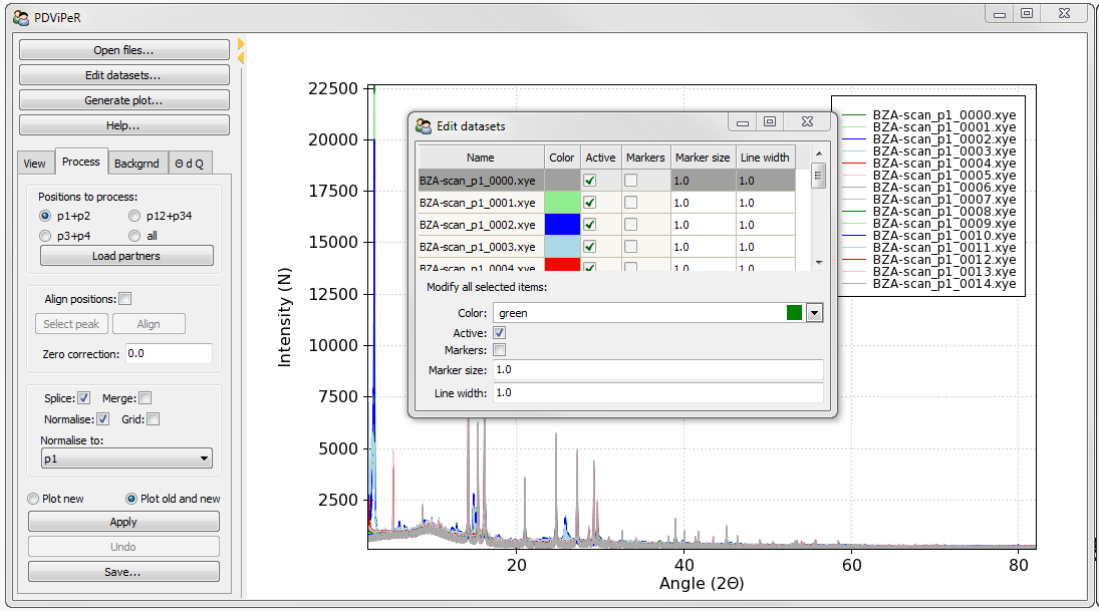


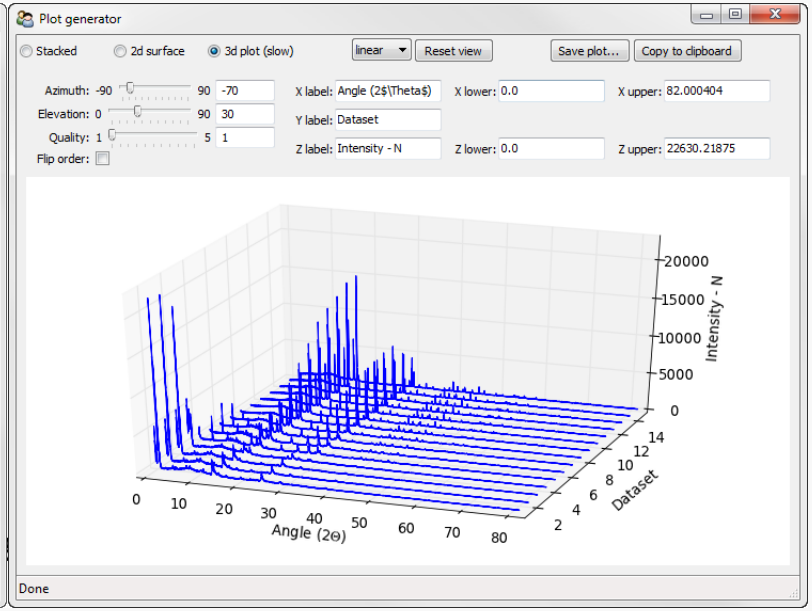
- Dynamic provisioning of Virtual Machines (VMs) on the NeCTAR Research Cloud (EC2) for SAXS data reduction hosting ScatterBrain, PyMol and ATSAS
- Linked to AS experimental data-store (via SFTP)
- Access either via NoMachine (web access or client) or within the browser using HTML5
- Still proto-type state, 43 sessions from 7 individual users in August 2013
- To be deployed for other beamlines

Supported by
Australian Government
State Government Victoria

POWDER DIFFRACTION VISUALISATION PROCESSING AND REPORTING TOOL (PDViPeR)

- Process data from the Mythen microstrip detector (http://www.synchrotron.org.au/pdviper)
- Dataset unit conversion, unit rescaling, normalisation, regridding, dataset pair alignment and merging.
- Automated peak and background fitting including background removal
- Generates high-quality image output for publications (4 plot types, vector and raster-based)
- Packaged, cross-platform, based on the Enthought Canopy python analysis framework https://www.enthought.com/products/canopy/
- 285 downloads since our first release in September 2013
- Source code: https://github.com/AustralianSynchrotron/pdviper





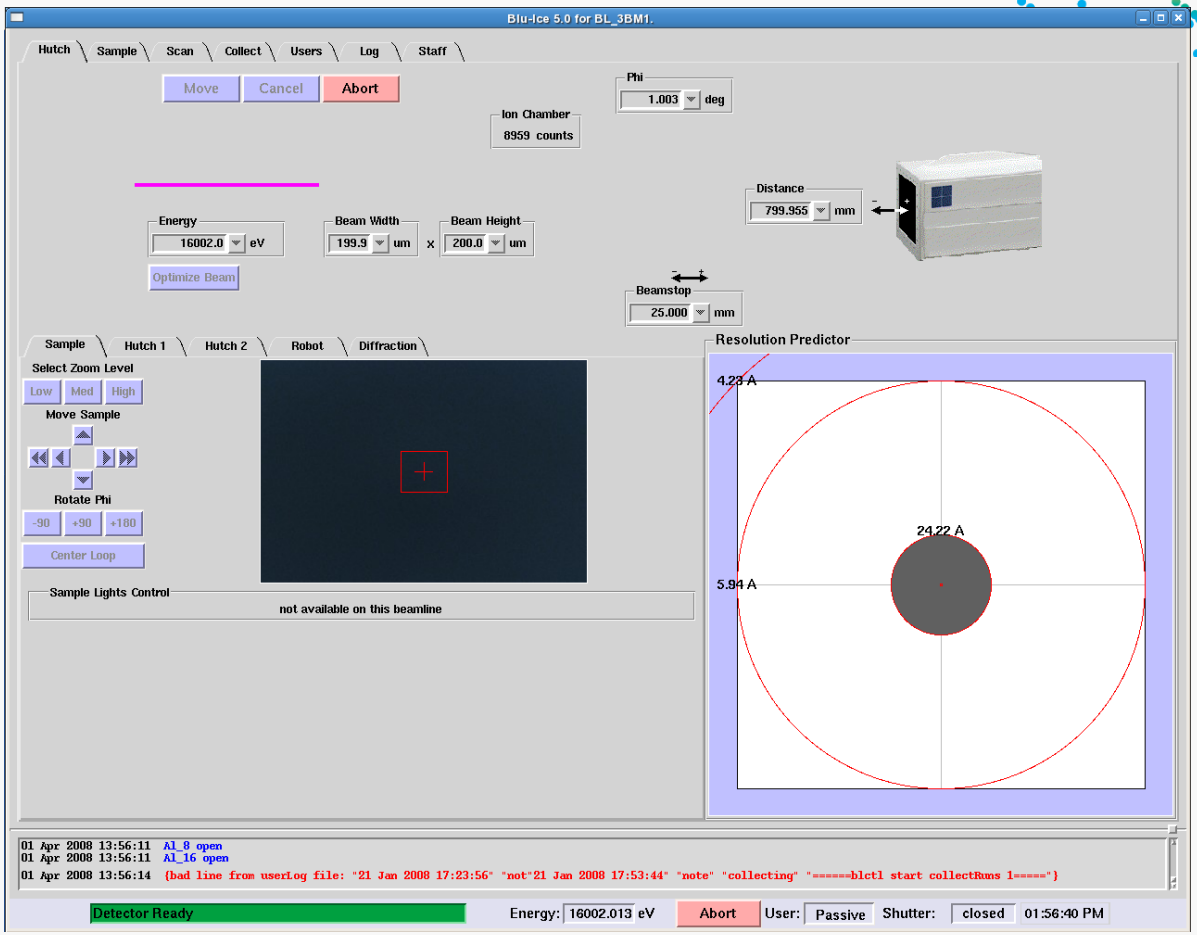
MACROMOLECULAR & MICRO-CRYSTALLOGRAPHY (MX1/2)

- MX1 is on a bending magnet (BM) and MX 2 is on an in-vacuum undulator
- The BM is a lower intensity source, with a larger beam (typical spot size 115 x 130 μm)
- The in-vacuum undulator provides a high-flux density spot (typical size 25 μm high x 30 μm wide)
- High-throughput: equipped with sample-mounting robots (each robot cassettes takes 96 samples, space for three)
- Remote access (NoMachine): Users can send their preloaded samples and then log in and control the beamlines remotely
- Highly automated: Workflow for data processing and protein structure determination from diffraction images

BEAMLINE GUI

Blu-Ice

- Remote access via NoMachine (NX)
- Excitation scans
- MAD scans
- Robot sample mounting
- Sample centring
- Energy change control
- Detector distance control
- Cryo temperature control (MX1)



Java or pure web-based version planned (2014).

TOOLS FOR THE STRUCTURAL BIOLOGY COMMUNITY

- Workflow for automatic data processing and protein structure determination from MX diffraction images (close to real-time)
- Locally and remotely accessible via web front-end (https://aswebmx.synchrotron.org.au/processing)

Processing Results | **Processing**

Processing

Sample	EPN	Type	Resolution	Space Group	Unit Cell	No. Frames	Directory	Status
C03_1	6863	Dataset	2.05	P212121	44.31,52.44,131.1,90,90,90	181	/data/frames/6863/kakudas/C03	Success
C03_0	6863	Indexing	3.64	P2	52.81,44.36,131.08,90,91.58,90	1	/data/frames/6863/kakudas/C03	Success
C03_0	6863	Indexing	1.97	P222	44.21,52.43,132.44,90,90,90	1	/data/frames/6863/kakudas/C03	Success
A16_1	6863	Dataset	2.03	P21212	53.25,65.64,70.16,90,90,90	180	/data/frames/6863/kakudas/A16	Success
A16_0	6863	Indexing	1.93	P2	54.17,65.67,70.33,90,88.57,90	1	/data/frames/6863/kakudas/A16	Success
A15_1	6863	Dataset	2.05	P212121	56.3,66.22,71.95,90,90,90	180	/data/frames/6863/kakudas/A15	Success
A15_0	6863	Indexing	2.07	P222	56.3,66.63,72.09,90,90,90	1	/data/frames/6863/kakudas/A15	Success
A15_0	6863	Indexing	2.31	P222	56.47,65.77,72.15,90,90,90	1	/data/frames/6863/kakudas/A15	Success
A14_1	6863	Dataset	1.94	P212121	54.31,66.06,70.56,90,90,90	180	/data/frames/6863/kakudas/A14	Success
A14_0	6863	Indexing	1.89	P222	54.35,66.26,70.66,90,90,90	1	/data/frames/6863/kakudas/A14	Success
A13_0	6863	Indexing			,,,,,	1	/data/frames/6863/kakudas/A13	Failed
A12_1	6863	Dataset	1.92	P212121	53.28,65.94,70.68,90,90,90	180	/data/frames/6863/kakudas/A12	Success

AUTO PROCESSING

Processing Results

Processing

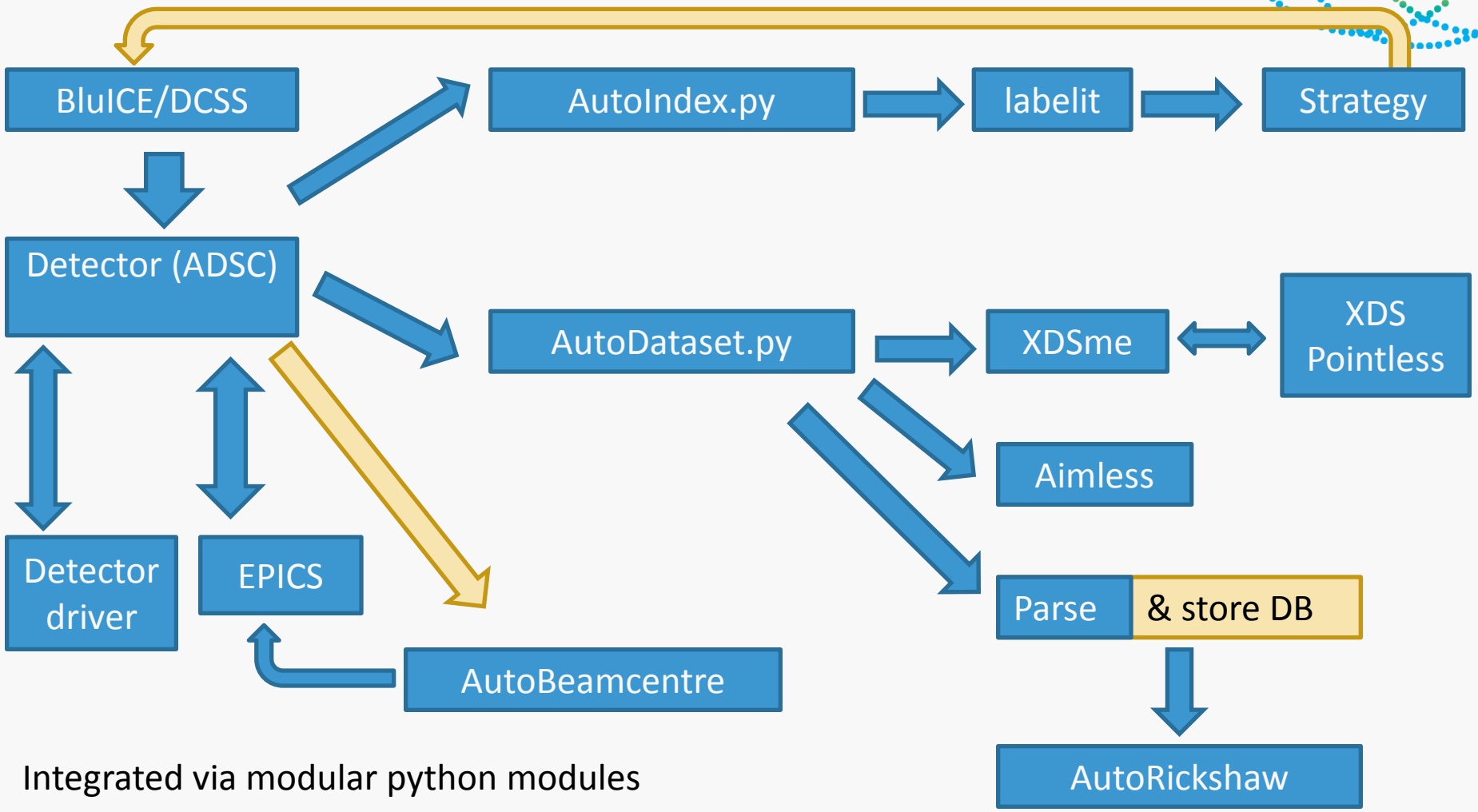
Sample	EPN	Type	Status
C03_1	6863	Datase	Success
C03_0	6863	Indexir	Success
C03_0	6863	Indexir	Success
A16_1	6863	Datase	Success
A16_0	6863	Indexir	Success
A15_1	6863	Datase	Success
A15_0	6863	Indexir	Success
A15_0	6863	Indexir	Success
A14_1	6863	Datase	Success
A14_0	6863	Indexir	Success
A13_0	6863	Indexir	Failed
A12_1	6863	Datase	Success
A12_0	6863	Indexir	Success
A12_0	6863	Indexir	Success
A09_1	6863	Datase	Success

C03_1 - dataset

Epn	6863
Started At	2013-03-27 18:57:41 EST
Status	Success
Sample	C03_1
Directory	/data/frames/6863/kakudas/C03
No Frames	181
Last Frame	/data/frames/6863/kakudas/C03/C03_1_183.img
Resolution	2.05
Space Group	P212121
Unit Cell	[44.31, 52.44, 131.1, 90.0, 90.0, 90.0]
Average Mosaicity	0.22

	Overall	Low	High
Low Resolution Limit	48.69	48.69	2.1
High Resolution Limit	2.05	8.79	2.05
Completeness	99.6	99.6	95.6
I/Sigma	10.2	29.5	2.6
Rmerge	0.156	0.031	0.581
Rpim(I)	0.067	0.014	0.254
Multiplicity	7.1	5.9	7.0
Anomalous Completeness	98.9	100.0	93.7
Anomalous Correlation	0.062	0.069	0.116

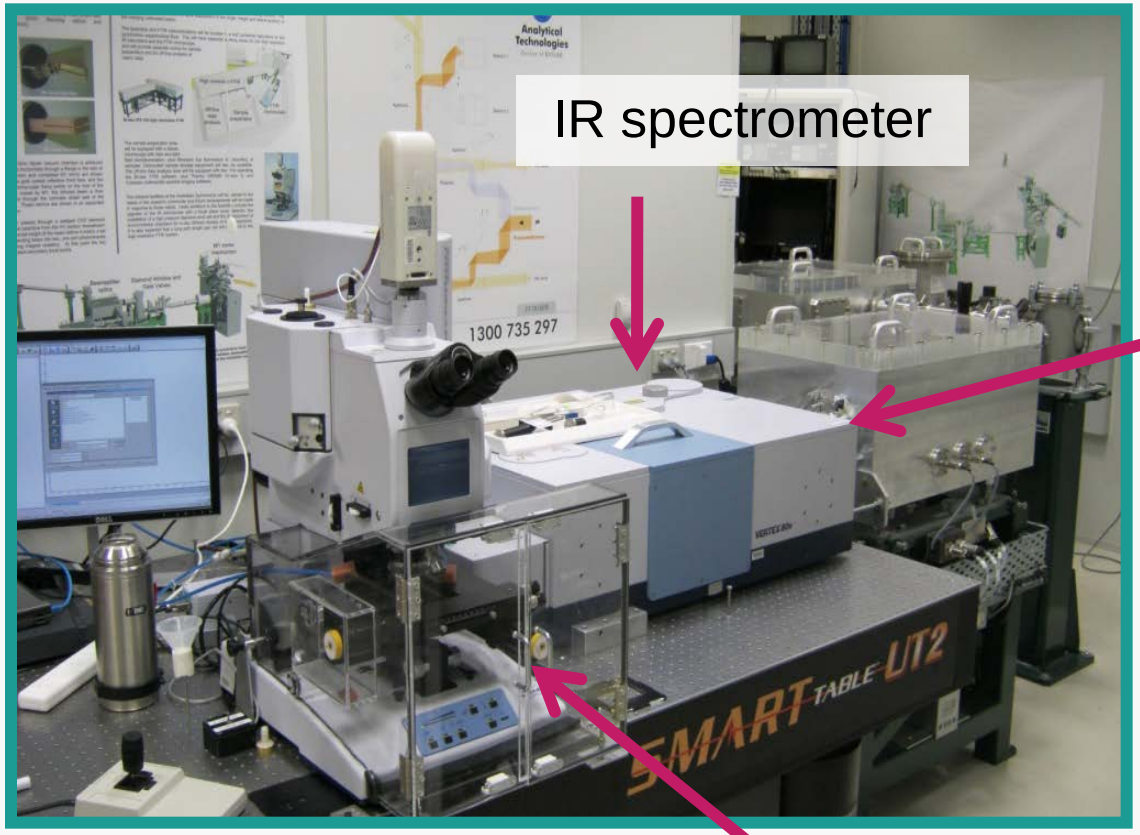
AUTO PROCESSING



Integrated via modular python modules

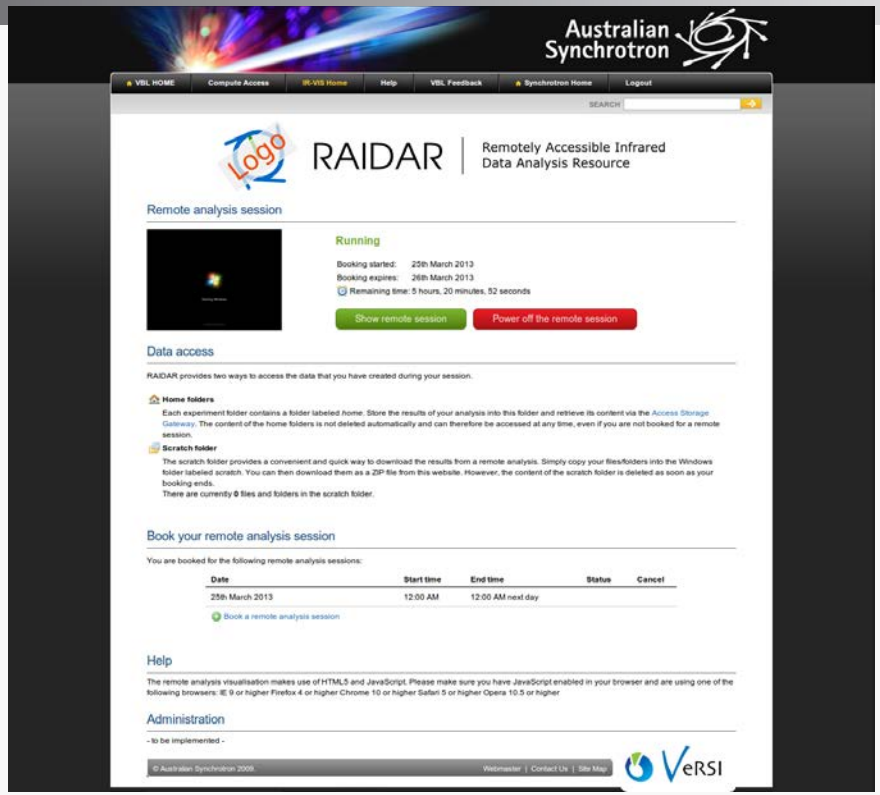
INFRA RED MICROSCOPY (IRM)

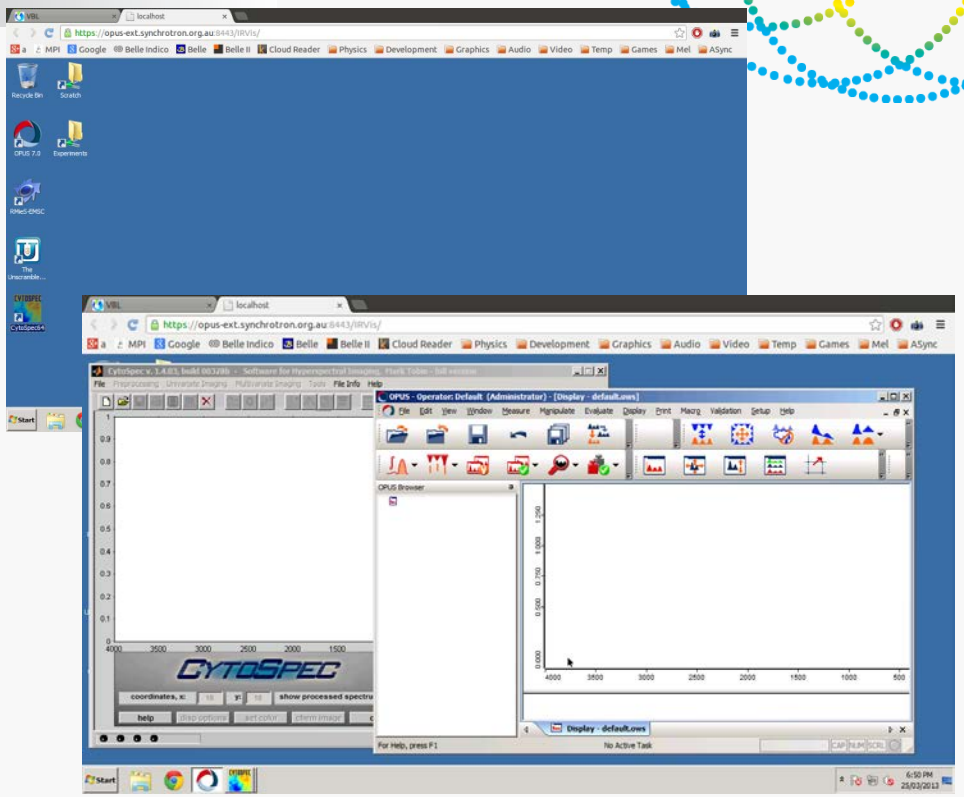
- By coupling an infrared spectrometer to a specialised microscope, IR data can be collected from microscopic samples



- SR source, single-point MCT detector
- **High spatial resolution, down to a few microns**
- Mid-IR range, 750-3850 cm^{-1}

RAIDAR | Remotely Accessible Infrared Data Analysis Resource

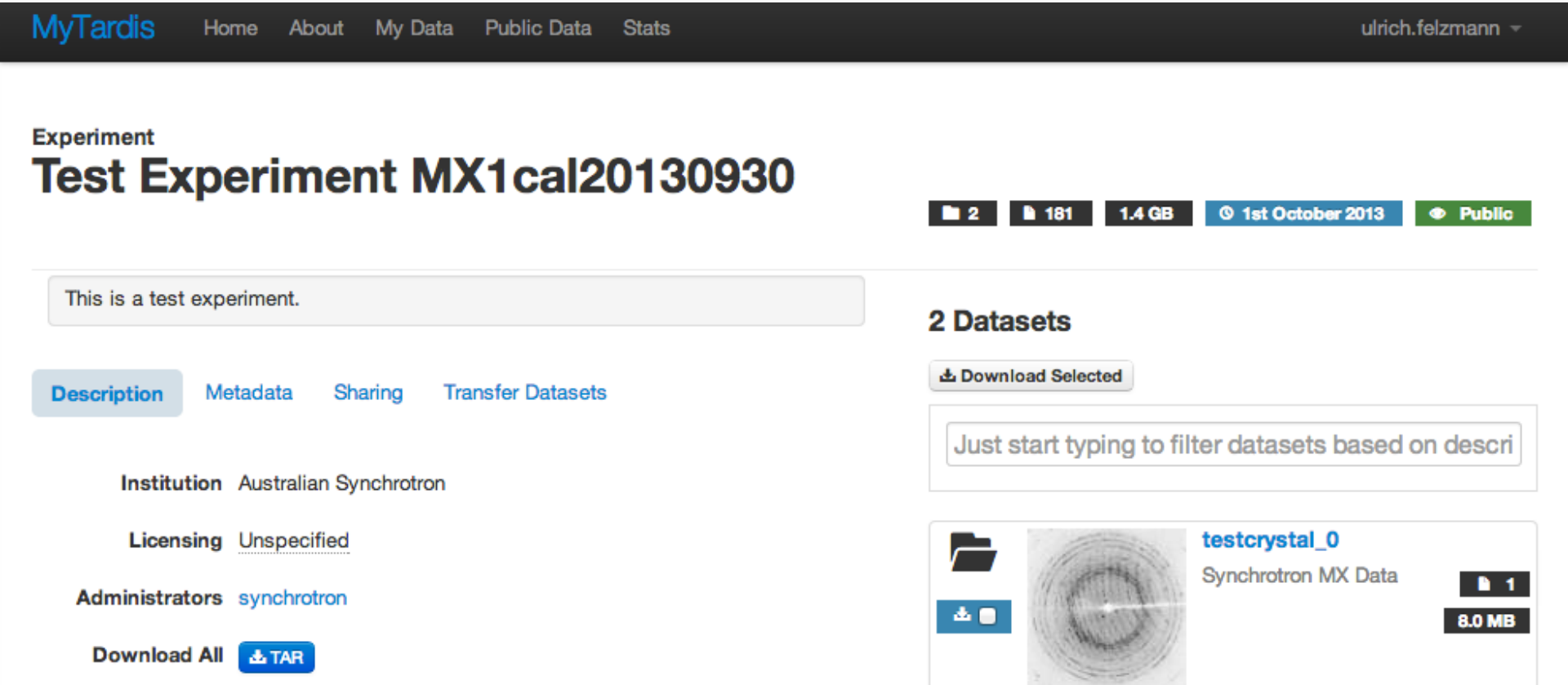




- Deployment of a Synchrotron-hosted Windows-VM (VirtualBox) with IR software stack
- Linked to experimental data-store
- “Single user mode” only, booking required
- 68 booked sessions in 2013

DATA & METADATA MANAGEMENT

- Data- & metadata is managed centrally for 4 beamlines (IR, MX1, MX2, SAXS) at the Australian Synchrotron using the MyTardis software (https://github.com/mytardis/mytardis)
- Accessible for discovery, citation, reuse and collaboration via a web-interface



SUMMARY & CONCLUSION

- Portfolio of tools has been created for the users of the Australian Synchrotron
- Simplify the otherwise laborious task of post-experiment processing large data sets
- Project has shown very good uptake
- Everything released on GitHub https://github.com/AustralianSynchrotron
- Interested in collaborations? Please contact me: ulrich.felzmann@synchrotron.org.au
- Thank you and hope to see you in Melbourne 2015!