



eResearch Tools for the Australian Synchrotron User Community

Uli Felzmann Australian Synchrotron

ICALEPCS 2013 – San Francisco

AUSTRALIAN SYNCHROTRON



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 (<u>www.ansto.gov.au</u>) 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

Control Room / Beam Current 200.3 mA	Duty Officer Ex Beam Lifetime 22.8 Hrs	tn 4123 Beam Current x Lifetime 4.56 AHrs	Mode: UserBeam Integrated Current 6,067.6 AHrs	Position X -14.6 um -	
			240	Y 1.7 um Size (FWHM)	
200 -			200	x 337.3 um	
				Y 240.9 um Beamline Status	ID Ga
			160	PSS Master Shutter Enable	
				Infra Red	
			-120	Micro Crystallography	7.4 m
				Macromolecular Crystallography	
				X-ray Fluorescence Microscopy	7.2 m
80 -				Medical Imaging	3.0
				Powder Diffraction	19.2 -
40			-40		10.2 1
		hard		Scattering	8.3 m
				Soft V-ray Spectroscopy	26 E m



Uli Felzmann – ICALEPCS2013 – WECOBA03



ustralian Governmen



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)

No. of experiments last 12 months









HIGH PERFORMANCE COMPUTING

- Multi-modal Australian ScienceS Imaging and Visualisation Environment (<u>www.massive,org.au</u>)
- 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 (<u>http://www.nectar.org.au/</u>)



- 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 <u>www.synchrotron.org.au/index.php/aussyncbeamlines/saxswaxs/software-saxswaxs</u>









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 (<u>http://www.embl-hamburg.de/biosaxs/software.html</u>)
- Providing real-time results during an experiment
- Easing the otherwise laborious task of post-experiment processing large data sets.



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

Australian	Time series for an in-situ ther showing development of an a	mal treatment experiment
<image/> <image/>	Time series for an in-situ ther showing development of an a representation of the second of the seco	mal treatment experiment dditional phase
	StateBan Pit StateBan V1.730 A. Surface	Hin a Her q 20

- 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



Uli Felzmann – ICALEPCS2013 – WECOBA03

Supported bv



POWDER DIFFRACTION VISUALISATION PROCESSING AND REPORTING TOOL (PDViPeR)

- Process data from the Mythen microstrip detector (<u>http://www.synchrotron.org.au/pdviper</u>)
- 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 <u>https://www.enthought.com/products/canopy/</u>
- 285 downloads since our first release in September 2013
- Source code: https://github.com/AustralianSynchrotron/pdviper

C PDViPeR		Plot generator
Open files		© Stacked © 2d surface ⊚ 3d plot (slow) linear ▼ Reset view Save plot Copy to clipboard
Edit datasets	22500 -	Azimuth: -90 0 -70 X label: Angle (2\$\Theta\$) X lower: 0.0 X upper: 82.000404
Help		Elevation: 0 0 30 Y label: Dataset
View Process Backgrnd Od O	20000 -	BZA-scan p1 0002 xye Quality: 1 U 5 1 Name Color Active Markers Marker size Line width ▲ BZA-scan p1 0003 xye Flip order: □ Zlabel: Intensity - N Zlower: 0.0 Z upper: 22630.21875
Positions to process:	17500 -	BZA-scan_p1_0000.xye I I.0 I.0 E BZA-scan_p1_0005 xye BZA-scan_p1_0000.xye BZA-scan_p1_0005 xye BZA-scan_p1_0005 xye BZA-scan_p1_0005 xye
p1+p2 p12+p34 p3+p4		BZA-scan_p1_0001.xye I I I I I BZA-scan_p1_0002.xye I I I I I
Load partners	15000	BZA-scan_p1_0003.xye I <thi< th=""> <thi< th=""> I <thi< th=""></thi<></thi<></thi<>
Align positions:	2 12500	Modify all selected items:
Select peak Align	insit	Color: green
Zero correction: 0.0	<u>ا ا ا</u> 0000 ا	Markers:
Splice: 🖉 Merge: 🕅	7500 -	Line width: 1.0
Normalise: Grid:		
p1	5000 -	
Plot new Plot old and new	2500 -	
Apply		$\frac{1}{20} \frac{30}{30} \frac{40}{40} \frac{50}{50} \frac{60}{70} \frac{2}{80} \frac{2}{2}$
Save		20 40 60 80
		Angle (20)

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 $\mu\text{m})$
- The in-vacuum undulator provides a high-flux density spot (typical size 25 μ m high x 30 μ m 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)

B	u-ice 5.0 for BL_3BM1.	
Hutch \/ Sample \/ Scan \/ Collect \/ Users \/ Log \/ Staff \		
Move Cancel Abort	Chamber 59 counts	3 v deg
Energy 16002.0 v eV Beam Width 199.5 v um x 200.0 v u Optimize Beam	e Beaus	Distance 799.955 mm
	<u> </u>	Resolution Predictor
Select Zoom Level		428A 2422 A 5.38A
01 Apr 2008 13:56:14 (bad line from userLog file: "21 Jan 2008 17:23:56" "no	ot"21 Jan 2008 17:53:44"	<pre>"note" "collecting" "======blctl start collectRums 1====="}</pre>
Detector Ready E	nergy: 16002.013 eV	Abort User: Passive Shutter: closed 01:56:40 PM

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	Туре	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	Туре
C03_1	6863	Datase
C03_0	6863	Indexi
C03_0	6863	Indexi
A16_1	6863	Datase
A16_0	6863	Indexi
A15_1	6863	Datase
A15_0	6863	Indexi
A15_0	6863	Indexi
A14_1	6863	Datase
A14_0	6863	Indexi
A13_0	6863	Indexi
A12_1	6863	Datase
A12_0	6863	Indexi
A12_0	6863	Indexi
A09_1	6863	Datase

C03_1 - dataset					×	
Epn	6863					
Started At	2013-03-27 18:57:41 EST					
Status	Success					Status
Sample	C03_1				s/C03	Success
Directory	/data/frames/6863/kakuda	s/C03			×/C03	Success
No Frames	181					
Last Frame	/data/frames/6863/kakuda	s/C03/C03_1_183.i	mg		\$/C03	Success
Resolution	2.05				s/A16	Success
Space Group	P212121					Success
Unit Cell	[44.31, 52.44, 131.1, 90.0, 90.0, 90.0]					Success
Average Mosaicity	0.22				NA13	ouccess
					s/A15	Success
		Overall	Low	High	s/A15	Success
Low Resolution Limit		48.69	48.69	2.1	s/A14	Success
High Resolution Limit		2.05	8.79	2.05		
Completeness		99.6	99.6	95.6	s/A14	Success
I/Sigma		10.2	29.5	2.6	s/A13	Failed
Rmerge		0.156	0.031	0.581	A12	Success
Rpim(I)		0.067	0.014	0.254	////10	Queenee
Multiplicity		7.1	5.9	7.0	DIATZ	Success
Anomalous Completeness		98.9	100.0	93.7	s/A09	Success



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



Australian





Remotely Accessible Infrared Data Analysis Resource

A COMPANY		S	Austr	alian 🍏	Image: State of the state o	
IL MOME Compute Access (DLAN3 House)		edback Synchro Remotely A Data Analy	SEARCH SEARCH Accessible In sis Resourc	nfrared e	Register South	
Remote analysis session Booking - Contemport Booking - Contemport Solution - Solution -	arted: 25th March ipires: 26th March ing time: 5 hours, 20 V femote session	2013 2013 minutes, 52 seconds Power off the	remote session			☆ 0
Data access	and during your sea	ion .			Contraction of the state o	
Home tokiers Each experiment folder contains a blobe tableted norm. S Contraw, The content of the home blobers is not delated a Section Section there The content of the home blobers and during the way to booking ands. There are currently 8 lies and blobers in the social hilder Booky your remote analysis session	ore the results of your formatically and can th download the results IP file from this websit	analysis into this folder and erefore be accessed at an form a remote analysis. Sin e. However, the content of	d retrieve its contern y time, even if you a mply copy your files. The scratch folder is	t via the Access Storage are not booked for a remote flotders into the Windows deleted as soon as your		
You are booked for the following remote analysis sessions:					04	
Date	Start time	End time	Blatus	Cancel	20 Co	
25th March 2013 Block a remote analysis session	12:00 AM	12:00 AM next day			0.7 0.7 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5	
Help						
The remote analysis visualisation makes use of HTML5 and J	vaScript. Please make 0 or higher Safari 5 or	sure you have JavaScript higher Opera 10.5 or high	t enabled in your bro	owser and are using one of the		1000 500
KORDWING DRUWBERS, IE. 9 OF PROPERT FITTERUX & OF PROPERT CITEDRA			55			
Administration					help dispontant setcour chemistance c	

- Deployment of a Synchrotron-hosted Windows-VM (VirtualBox) with IR software stack
- Linked to experimental data-store

RAIDAR

- "Single user mode" only, booking required
- 68 booked sessions in 2013



Uli Felzmann – ICALEPCS2013 – WECOBA03

Supported



DATA & METADATA MANAGEMENT

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



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
 <u>https://github.com/AustralianSynchrotron</u>
- Interested in collaborations? Please contact me: <u>ulrich.felzmann@synchrotron.org.au</u>
- Thank you and hope to see you in Melbourne 2015!





Uli Felzmann – ICALEPCS2013 – WECOBA03

Supported



