

## Data collection strategy and Data reduction in JBlueICE-EPICS

Sudhir Babu Pothineni, Nagarajan Venugopalan, Craig Ogata, Mark Hilgart  
GM/CA@APS  
TWG Meeting, Aug 16<sup>th</sup> 2012

# Software Highlights of GM/CA@APS

- JBluice-EPICS beamline control software
- Automated screening
- Automated diffraction/fluorescence rastering for invisible crystals and to select best diffracting areas of a visible crystal
- Auto centering on visible crystals
- Automated data collection strategy
- Single site & multiple site data collection (Vector data collection)
- Automated data reduction

The screenshot displays the JBluice-EPICS Beamline BM Version 2017.2 Build 4828 software interface. The window title is "JBluice-EPICS: Beamline BM Version 2017.2 Build 4828". The menu bar includes "File", "Network", "Options", "Tools", and "Help". The main interface is divided into several sections:

- Control Buttons:** "Start" and "Cancel".
- Attenuation:** Value 34.98.
- Omega:** Value 359.989.
- 2-Theta:** Value -0.000.
- Energy:** Value 12.6580.
- Width:** Value 0.045.
- Height:** Value 0.026.
- Distance:** Value 935.000.
- Beamstop:** Value 30.000.
- Automation Panel:** "Loop Centering", "Crystal Centering", "Stop", "Enter Hutch", "Optimize Beam".
- Sample Selection:** "Sample: Low Res", "Sample: High Res", "View Hutch".
- LowRes Zoom:** Value 8.00.
- Rotate +/-n:** Value 1.00 deg.
- Resolution Predictor:** A circular plot with a red circle and a white square, showing values 5.86 and 8.18.
- Status Bar:** APS Current: 102.1, Shutter Permit: Enabled, A Shutter: N/A, Endstation Shutter: Closed, Endstation Secure: No, State: idle, ETA: ---, EMERGENCY STOP, Mono: 12.658 keV, IZero: 0.05 V, Control: Passive, Shutter: Closed.

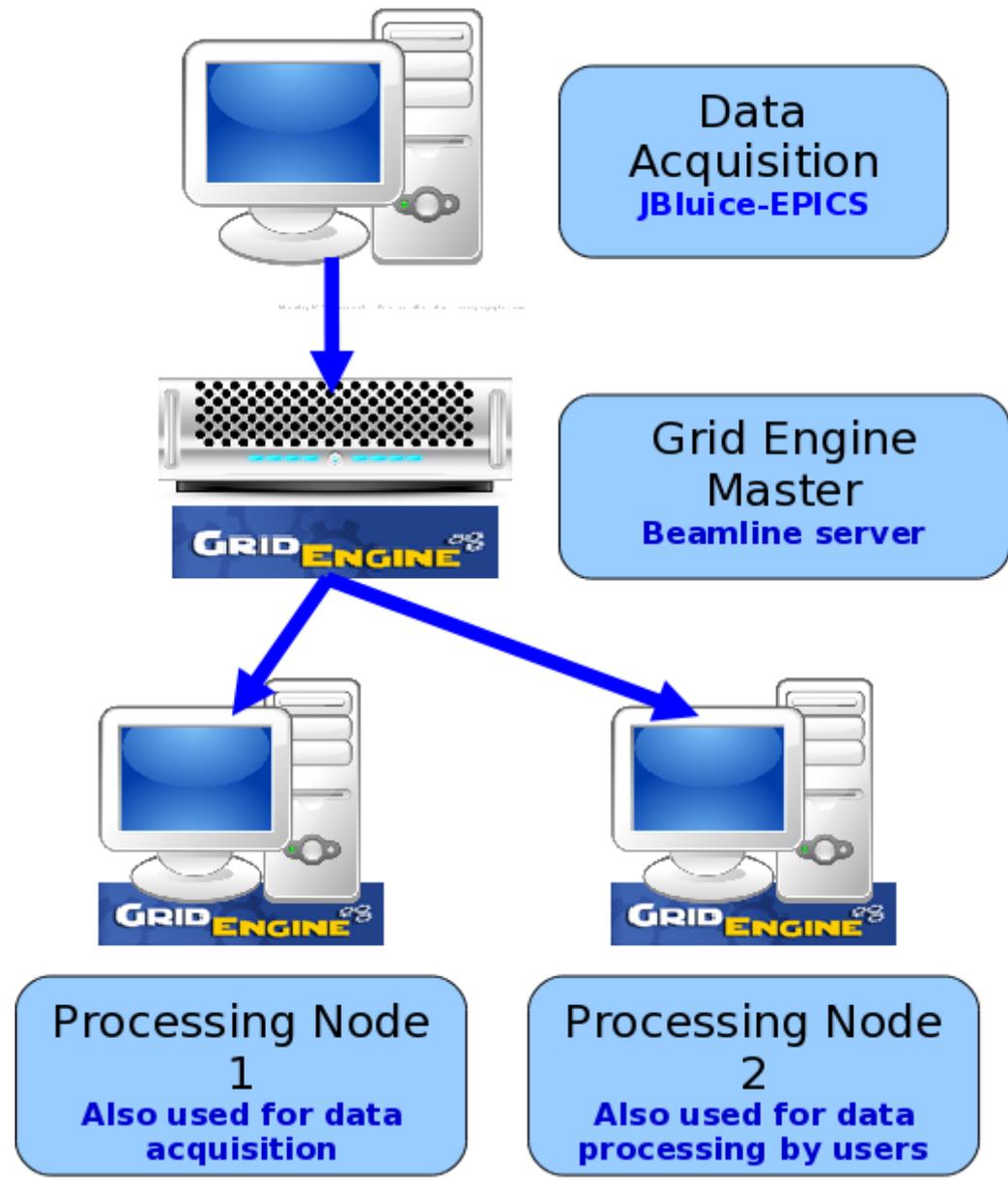
<http://www.gmca.anl.gov/jbluice-epics/>



# Data Processing in JBluice-EPICS

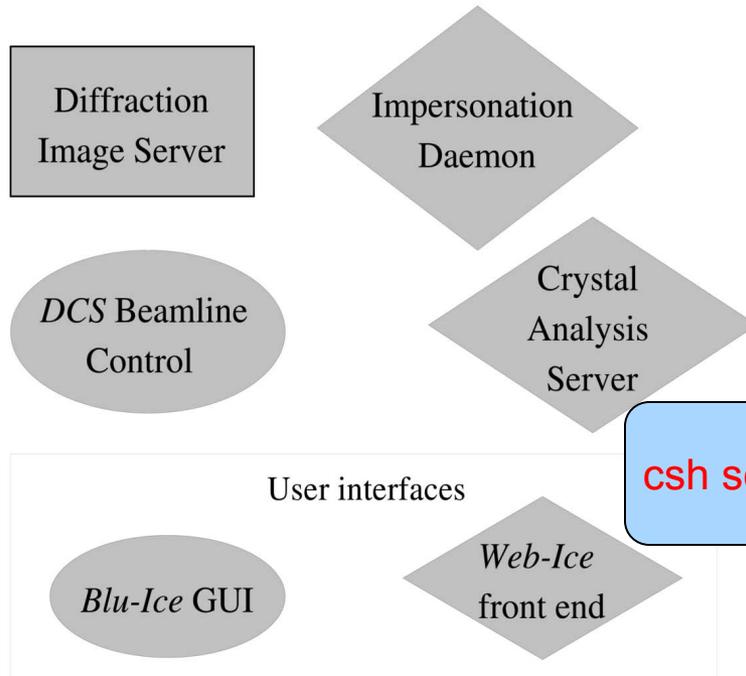
- Efficient use of available resources.
- JBluice uses Grid Engine for running strategy and data reduction jobs.
- Grid Engine is an open source batch queuing system/job scheduler.
- JBluice submits and monitors the Grid Engine jobs using DRMAA java binding and parses the result log files for quality parameters.
- Scalable to dedicated multiple node cluster.

- SSRL Webice based data collection strategy.
- XDS based data reduction

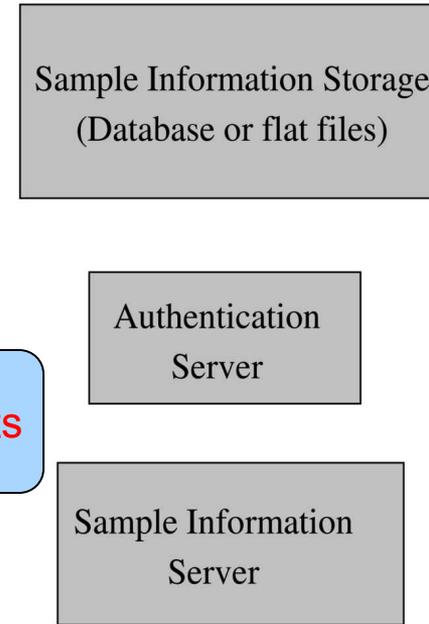


# SSRL WebIce

Scalable resources (site deploys as many instances as needed)



Single-instance resources  
(only one deployed per site)



csh scripts

*LABELIT* (autoindexing )  
*MOSFLM* (integration)  
*BEST* (strategy)

**Web-Ice: integrated data collection and analysis for macromolecular crystallography.** A. González, P. Moorhead, S.E. McPhillips, J. Song, K. Sharp, J.R. Taylor, P.D. Adams, N.K. Sauter and S.M. Soltis. *J. Appl. Cryst.* 41, 176-184 (2008).



## Porting Webice to JBluice

Webice	JBluice-EPICS
Impersonation daemon, crystal-analysis server	Grid Engine
Sample information server	MySQL
Image server	Image server
JSP web interface	Java swt application interface
SSRL csh scripts for strategy calculations	Modified SSRL csh scripts for strategy calculations

- Maintenance of tomcat web server, authentication server, ssl certificates etc.. are avoided. Parallelization added.



# Strategy initiated from screening tab

The screenshot shows the JBIulce-EPICS Beamline ID-B software interface, Version 2012.2 Build 4832. The interface is divided into several sections:

- Navigation:** File, Network, Options, Tools, Help.
- Tabbed Interface:** Hutch, Sample, **Screening**, Raster, Scan, Collect, Analysis, Users, Log.
- File Management:** /mnt/software/common/ALS\_template.xls, Import spreadsheet, Edit, Export spreadsheet.
- Table:** A table with columns: Select, Port, CrystalID, Directory, Comment, Resolution, Score. Rows 1-7 are checked.
- Tasklist/History:** Task Progress: A1 MountNextXtal, A1 Paused4Center, A1 Rotate0, A1 Collect0, A1 Rotate90, A1 Collect90, A1 Paused4Inspection.
- Puck Type:** ALS,  Puck A, Puck B, Puck C, Puck D, Puck E, Puck F. Mounted: none.
- Collection Parameters:** Prefix: F15\_0, Path: /mnt/share2/sliu, Current position: Gonio = 180.901, Detector = 900.000, Attenuation = 3.002. Settings: Distance(mm): 250.000, Delta(deg): 1.0, Expose(sec): 2.0, Attenuation(factors): 20.00. Wash/Rinse by Robot: 3, wash, Go.
- Actions:**  Mount Next Crystal,  Auto Centering (Loop/Crystal),  Pause4Center,  Jpeg,  Collect Image (0, 1),  Collect Image (45, 1),  Collect Image (90, 1),  Pause4Inspection. Start, Resume, Show color code info.
- Robot:** Ready, Pause, Mount, Unmount, Gripper Warmup. Robot Status: Current: safetyMat cleared, -1: pause, -2: SMR: ready. Command: .
- Status Bar:** [13:06:22] NOTE: puck type ALS. APS Current: -0.0, Shutter Permit: Disabled, A Shutter: Closed, Endstation Shutter: Closed, Endstation Secure: No, State: Idle, ETA: ---, EMERGENCY STOP, Mono: 12.000 keV, IZero: 0.02 V, Control: Active, Shutter: Closed.

- Strategy calculation initiated automatically after collecting 2 images in screening tab.

# Strategy initiated from collect tab

The screenshot shows the JBLuce-EPICS Beamline ID-D software interface. The main window is titled "JBLuce-EPICS: Beamline ID-D Version 2012.2 Build 4803 \*Simulation Level 1\*". The interface is divided into several tabs: Hutch, Sample, Screening, Raster, Scan, Collect, Analysis, Users, and Log. The "Collect" tab is active, and the "Strategy" sub-tab is selected. The "Run 0 (inactive)" panel on the right contains various configuration fields, including Prefix, Dir, Distance, Atten., Site, Beam size, Delta, Time, Frame, and Gonio. The "Strategy" checkbox is checked, and an orange arrow points to it. The "Current position" section shows Gonio = 14.803, Detector = 900.001, Attenuation = 15.055, and Beamstop = 35.000. The "Resolution Predictor" section shows a circular plot with values 4.19 and 5.87. The "Run sequence" table is also visible.

Filename	Angle	Ener
test_0.0001	14.80	12.8
test_0.0003	104.80	12.8

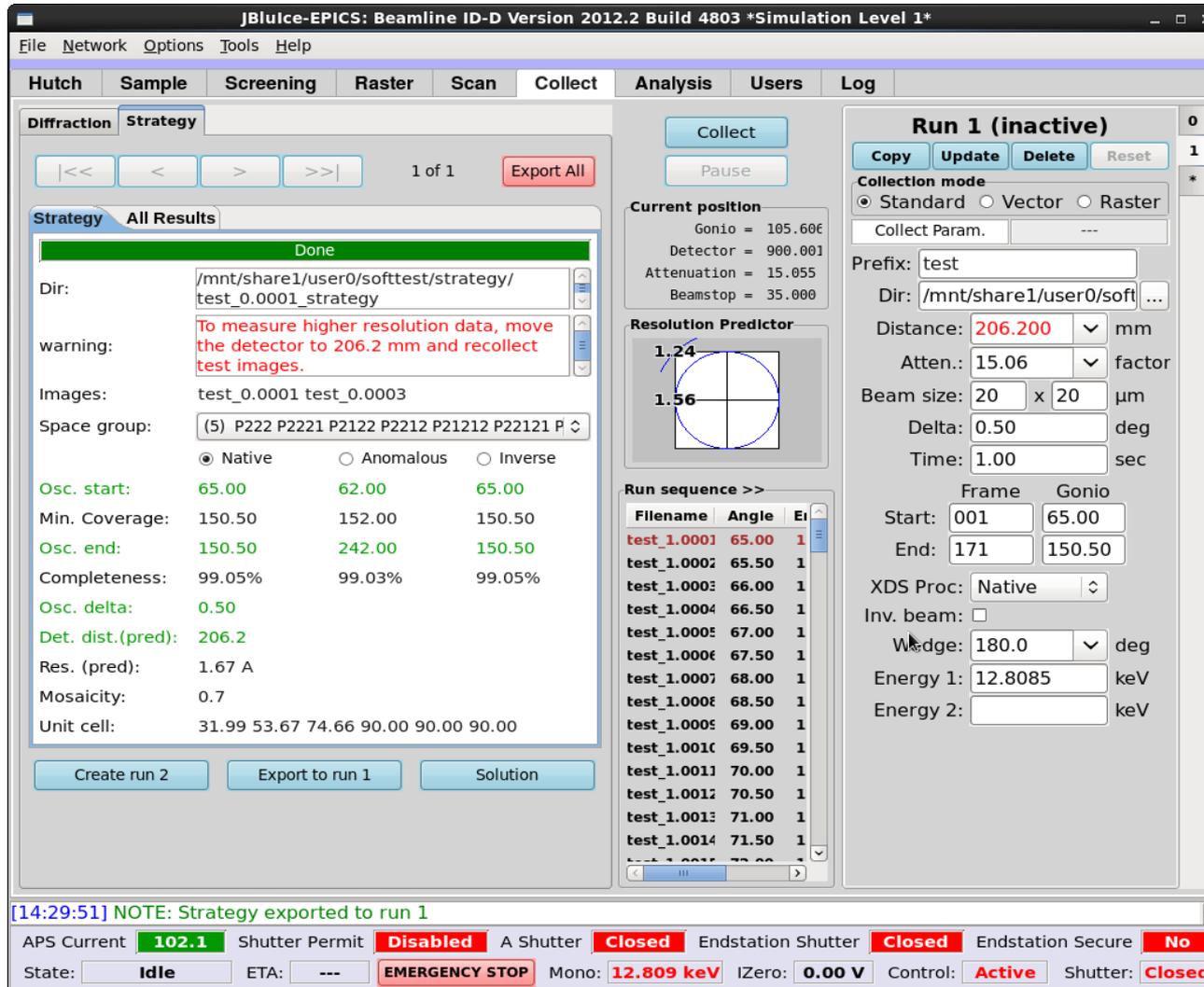
[14:27:44] NOTE: [collimatorCtrl.pl]

APS Current: 102.5 Shutter Permit: Disabled A Shutter: Closed Endstation Shutter: Closed Endstation Secure: No  
State: Idle ETA: --- EMERGENCY STOP Mono: 12.809 keV IZero: 0.00 V Control: Active Shutter: Closed

- Strategy can be initiated from collect tab '0' run for manually mounted samples



# Strategy parameters



**JBluice-EPICS: Beamline ID-D Version 2012.2 Build 4803 \*Simulation Level 1\***

File Network Options Tools Help

Hutch Sample Screening Raster Scan Collect Analysis Users Log

Diffraction Strategy

1 of 1 Export All

Strategy All Results

Done

Dir: /mnt/share1/user0/softtest/strategy/test\_0.0001\_strategy

warning: To measure higher resolution data, move the detector to 206.2 mm and recollect test images.

Images: test\_0.0001 test\_0.0003

Space group: (5) P222 P2221 P2122 P2212 P21212 P22121 P

Native  Anomalous  Inverse

Osc. start:	65.00	62.00	65.00
Min. Coverage:	150.50	152.00	150.50
Osc. end:	150.50	242.00	150.50
Completeness:	99.05%	99.03%	99.05%
Osc. delta:	0.50		
Det. dist.(pred):	206.2		
Res. (pred):	1.67 Å		
Mosaicity:	0.7		
Unit cell:	31.99 53.67 74.66	90.00 90.00	90.00

Create run 2 Export to run 1 Solution

Collect Pause

Current position  
Gonio = 105.600  
Detector = 900.001  
Attenuation = 15.055  
Beamstop = 35.000

Resolution Predictor  
1.24  
1.56

Run sequence >>

Filename	Angle	El
test_1.0001	65.00	1
test_1.0002	65.50	1
test_1.0003	66.00	1
test_1.0004	66.50	1
test_1.0005	67.00	1
test_1.0006	67.50	1
test_1.0007	68.00	1
test_1.0008	68.50	1
test_1.0009	69.00	1
test_1.0010	69.50	1
test_1.0011	70.00	1
test_1.0012	70.50	1
test_1.0013	71.00	1
test_1.0014	71.50	1

Run 1 (inactive)

Copy Update Delete Reset

Collection mode  
 Standard  Vector  Raster

Collect Param. ---

Prefix: test

Dir: /mnt/share1/user0/soft ...

Distance: 206.200 mm

Atten.: 15.06 factor

Beam size: 20 x 20 µm

Delta: 0.50 deg

Time: 1.00 sec

Frame	Gonio
Start: 001	65.00
End: 171	150.50

XDS Proc: Native

Inv. beam:

Wedge: 180.0 deg

Energy 1: 12.8085 keV

Energy 2: keV

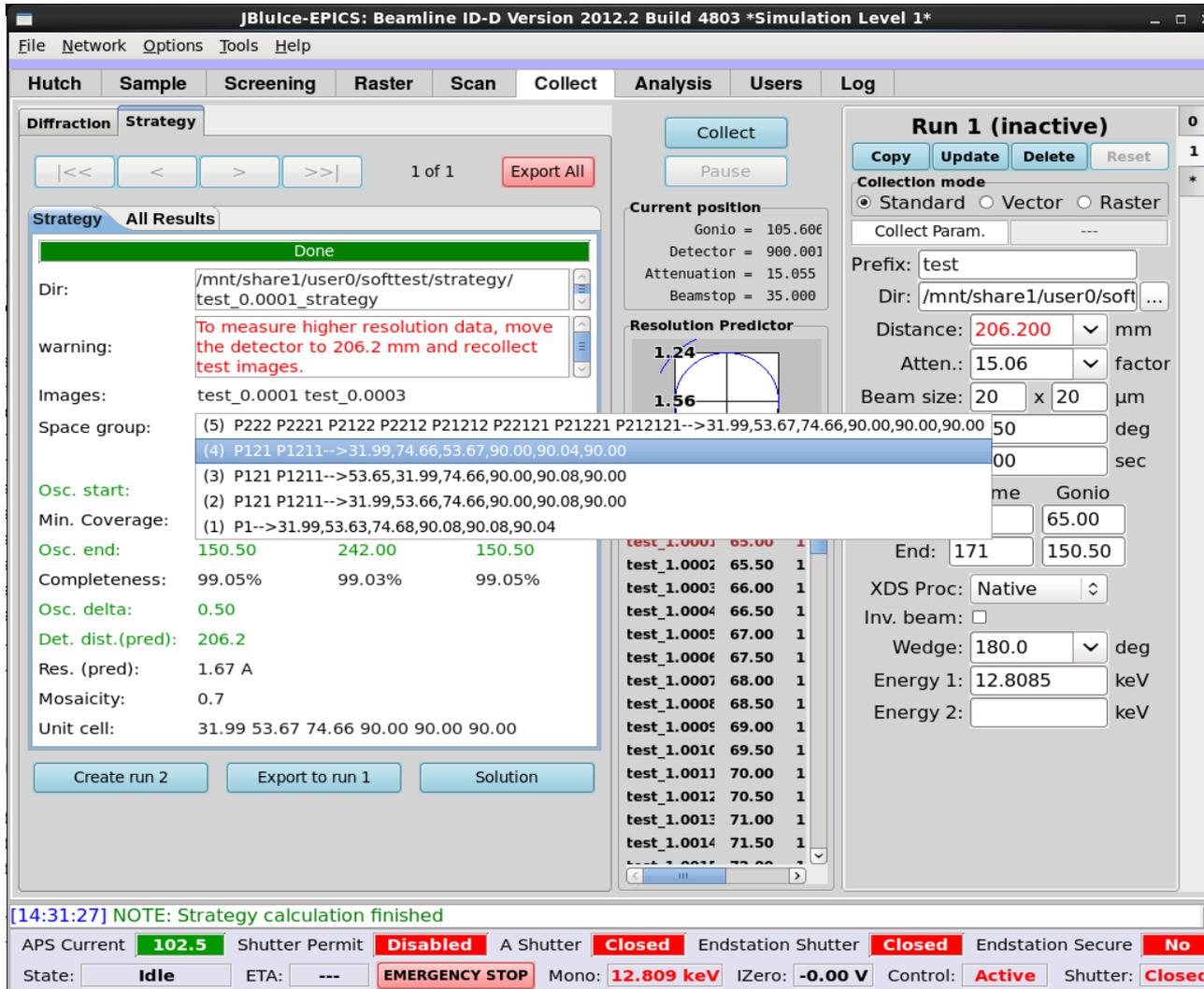
[14:29:51] NOTE: Strategy exported to run 1

APS Current: 102.1 Shutter Permit: Disabled A Shutter: Closed Endstation Shutter: Closed Endstation Secure: No

State: Idle ETA: --- EMERGENCY STOP Mono: 12.809 keV IZero: 0.00 V Control: Active Shutter: Closed

- After strategy calculation parameters are displayed on strategy sub tab

# Selecting lower symmetry spacegroup



**JBluIce-EPICS: Beamline ID-D Version 2012.2 Build 4803 \*Simulation Level 1\***

File Network Options Tools Help

Hutch Sample Screening Raster Scan Collect Analysis Users Log

Diffraction Strategy

Collect Pause

Run 1 (inactive)

Copy Update Delete Reset

Collection mode  
 Standard  Vector  Raster

Collect Param. ---

Prefix: test

Dir: /mnt/share1/user0/soft...

Distance: 206.200 mm

Atten.: 15.06 factor

Beam size: 20 x 20  $\mu$ m

50 deg

00 sec

me Gonio

End: 171 150.50

XDS Proc: Native

Inv. beam:

Wedge: 180.0 deg

Energy 1: 12.8085 keV

Energy 2: keV

Current position  
 Gonio = 105.60E  
 Detector = 900.001  
 Attenuation = 15.055  
 Beamstop = 35.000

Resolution Predictor

1.24  
 1.56

Strategy All Results

Done

Dir: /mnt/share1/user0/softtest/strategy/test\_0.0001\_strategy

warning: To measure higher resolution data, move the detector to 206.2 mm and recollect test images.

Images: test\_0.0001 test\_0.0003

Space group: (5) P222 P2221 P2122 P2212 P21212 P22121 P21221 P212121-->31.99,53.67,74.66,90.00,90.00,90.00  
 (4) P121 P1211-->31.99,74.66,53.67,90.00,90.04,90.00  
 (3) P121 P1211-->53.65,31.99,74.66,90.00,90.08,90.00  
 (2) P121 P1211-->31.99,53.66,74.66,90.00,90.08,90.00  
 (1) P1-->31.99,53.63,74.68,90.08,90.08,90.04

Osc. start:

Min. Coverage:

Osc. end: 150.50 242.00 150.50

Completeness: 99.05% 99.03% 99.05%

Osc. delta: 0.50

Det. dist.(pred): 206.2

Res. (pred): 1.67 A

Mosaicity: 0.7

Unit cell: 31.99 53.67 74.66 90.00 90.00 90.00

Create run 2 Export to run 1 Solution

test\_1.0001 65.00 1  
 test\_1.0002 65.50 1  
 test\_1.0003 66.00 1  
 test\_1.0004 66.50 1  
 test\_1.0005 67.00 1  
 test\_1.0006 67.50 1  
 test\_1.0007 68.00 1  
 test\_1.0008 68.50 1  
 test\_1.0009 69.00 1  
 test\_1.0010 69.50 1  
 test\_1.0011 70.00 1  
 test\_1.0012 70.50 1  
 test\_1.0013 71.00 1  
 test\_1.0014 71.50 1

[14:31:27] NOTE: Strategy calculation finished

APS Current 102.5 Shutter Permit Disabled A Shutter Closed Endstation Shutter Closed Endstation Secure No

State: Idle ETA: --- EMERGENCY STOP Mono: 12.809 keV IZero: -0.00 V Control: Active Shutter: Closed

- Lower symmetry spacegroup can be selectable for to re-run strategy



## Data Collection Strategy

- List all possible space groups from Labelit solutions.
- Solutions from same crystal symmetry are filtered.
- By default only space groups from highest symmetry are processed initially.
  - They are processed in parallel so it is faster.
- User can choose lower symmetry solution, JBluice will process and display strategy.
- User has an option to choose strategy from BEST/Mosflm.
- Displays Anomalous information, There are two options.
  - Anomalous continuous where osc. range calculated is doubled.
  - Anomalous Inverse where it will use inverse beam with 'Native' osc. range.
- Improved error/warning handling.
- All the strategy results can be saved in Excel file.
- Processing time about 30 sec (if there is index solution).



## Data Reduction

- There is XDSProc option (None, Native, Anomalous) on each run on collect tab, if checked it will initiate XDS processing through Grid Engine after a dataset is collected.
- JBluice will write XDS input file XDS.INP.
- Runs XDS, POINTLESS, SCALA and TRUNCATE
- Any dataset with more than 10 images collected is automatically reduced.
- XDS runs second time with DEFPIX, INTEGRATE, CORRECT incase of failure with INSUFFICIENT PERCENTAGE of spots.
- XDS can process Anomalous data for SAD & MAD data collection.

**Run 1 (inactive)**

Copy Update Delete Reset

Collection mode  
 Standard  Vector  Raster

Collect Param. ---

Prefix: test

Dir: /mnt/share1/user0/soft ...

Distance: 206.200 mm

Atten.: 15.06 factor

Beam size: 20 x 20  $\mu$ m

Delta: 0.50 deg

Time: 1.00 sec

	Frame	Gonio
Start:	001	65.00
End:	171	150.50

XDS Proc: Native

Inv. beam:

Wedge: 180.0 deg

Energy 1: 12.8085 keV

Energy 2: keV



# Analysis Tab

- Display data quality parameters and plots automatically on Analysis tab.
- Direct access to XDS, SCALA, TRUNCAE logs with the click of a button.
- The scaled mtz file, and other input and intermediate files stores in a subdirectory of data.
- User can change XDS.INP and reprocess the data.

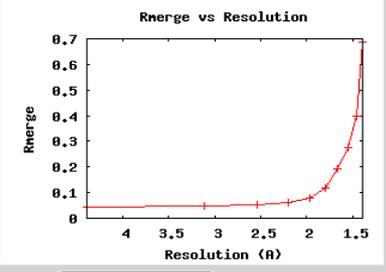
JBluice-EPICS: Beamline ID-D Version 2012.2 Build 4803

File Network Options Tools Help

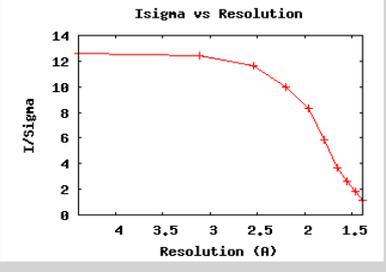
Hutch Sample Screening Raster Scan Collect Analysis Users Log

1 of 1 Export All

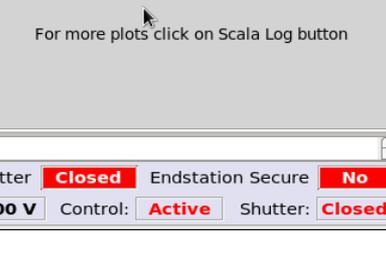
Rmerge Vs Resolution



Energy vs Resolution



I/Sigma Vs Resolution



For more plots click on Scala Log button

**Data Quality** All Results

Done

Sample Name: beth1\_2

Dir: /mnt/staffhome/spothineni/testxsderrors//beth1\_2\_xds

Images: /mnt/staffhome/spothineni/testxsderrors//beth1\_2.####

warning: DONE

Unit Cell: 49.61 55.26 84.24 90.00 90.00 90.00

Space Group: P 2 21 21

	Overall	InnerShell	OuterShell
Low Res.:	46.21	46.21	1.46
High Res.:	1.39	4.39	1.39
Rmerge:	0.085	0.043	0.685
Completeness:	97.4	94.7	98.1
Anom.Completeness:	93.2	91.7	94.4
Multiplicity:	4.1	3.9	4.1
Anom.Multiplicity:	2.2	2.3	2.2
I/Sigma:	9.7	24.8	1.8

XDS Log Scala Log Truncate Log

APS Current **102.4** Shutter Permit **Disabled** A Shutter **Closed** Endstation Shutter **Closed** Endstation Secure **No**

State: **Idle** ETA: --- **EMERGENCY STOP** Mono: **12.809 keV** IZero: **0.00 V** Control: **Active** Shutter: **Closed**



## Processing time & specs

Dataset :

360 MAR ccd images

JOB= ALL

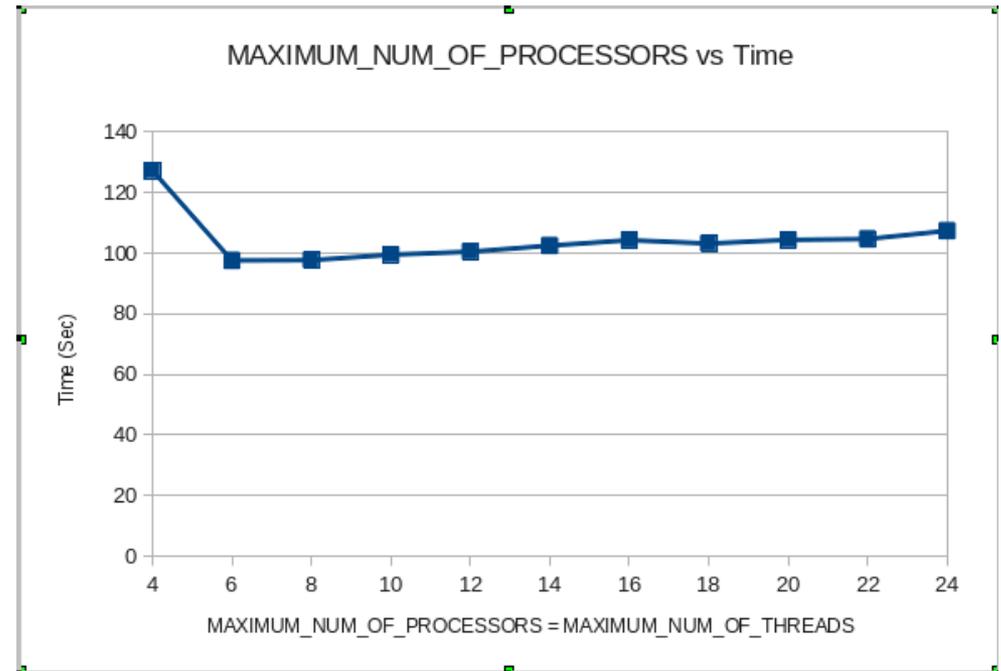
SPOT\_RANGE= 1 180

MAXIMUM\_NUM\_PROCESSORS= 6

MAXIMUM\_NUM\_THREADS= 6

DELTAPHI= 5

Processing time = 97sec.



Since manual processing overlapping with auto processing on same workstations using  
MAXIMUM\_NUM\_PROCESSORS= 4  
MAXIMUM\_NUM\_THREADS= 4

Processing time = 127sec.

Workstation : Intel Xeon with 2 x 6 cores, hyper-threading enabled, total 24 cores.

- limitation : 32 bit.

OMP\_STACKSIZE set according to number of cores.

- Ex. 64m for 24 core machine.



## Future plans

- Installing dedicated machines for data processing.
- Adopting fast\_dp from Diamond Light Source.
  - Error correction and reprocessing
- Running data processing in parallel to data collection.
- Processing vector data with overlaps and inverse beam.
- Processing data collected with raster sites.
- Molecular replacement for structure solution.

## Thanks

JBluice-EPICS

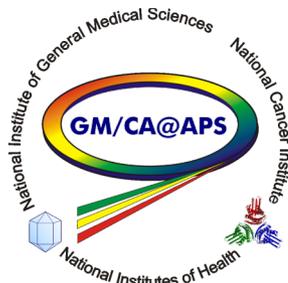
Mark Hilgart  
Sudhir Babu Pothineni

GM/CA Management

Janet Smith  
Robert Fischetti

Controls

Sergey Stepanov  
Oleg Makarov



Crystallography

Craig Ogata  
Nagarajan Venugopalan  
Ruslan sanishvili (Nukri)  
Michael Becker

LABELIT

Nicholas Sauter

