NSLS-II: updates from AMX and FMX

High Data-Rate Macromolecular Crystallography, HDRMX
MAX IV Laboratory, Lund, Sweden, March 15-17 2017
AMX and FMX

2 canted IVU21 undulators -> 2 independent beamlines

AMX: highly Automated MX
vast number of samples
large assemblies

FMX: Frontier MX
most challenging problems
requiring a μ-focus beam
and/or high X-ray energy

AMX beam: 5-100 μm beam
\( \sim 8 \times 10^{12} \text{ ph.s}^{-1} \) (1 Å)
5-18 keV (0.7 - 2.5 Å)
EIGER 9M (≤238 Hz)

FMX beam: 1-20 μm beam
\( \sim 6 \times 10^{12} \text{ ph.s}^{-1} \) (1 Å)
5-30 keV (0.4 - 2.5 Å)
EIGER 16M (≤133 Hz)

@ 500 mA ring current
Automated Sample Mounting (design & implementation)

High Capacity Automated Sample Changer with 24 unipucks capacity (384 bases).

In-House end-station

6-axis Staubli TX60/L robotic arm with NSLS-II gripper.

See Edwin’s poster
AMX / FMX: Current status (275 mA)

AMX: from 5 to 18 keV
Flux ~ 4 \(10^{12}\) ph.s\(^{-1}\)
Beam ~ 8 x 7 um\(^2\)
Manual sample mounting
Hardware trigger (raster / vector)
Typical data collection:
0.1 deg @ 100 Hz
Data collection: from <1s to 15 secs
Default screening:
10 % transmission & 10 ms/frame
Av data set: 1400 frames
500 frames / data.h5
Compression range: 4.2-25
Aver compression: 10

FMX: from 5 to 25 keV
Flux > 2 \(10^{12}\) ph.s\(^{-1}\)
Beam ~ 10 x 6 um\(^2\)
Automated sample mounting
Software trigger (raster / vector)
Typical data collection:
0.1 deg @ 50 Hz
Data collection: from <1s to 15 secs
Default screening:
<10 % transmission & 50 ms/frame
Av data set: 1800 frames
200 frames / data.h5
Compression range: 4.5-32
Aver compression: 11.6

40 % of available beamtime for user operation.
To be increased to 80 % gradually.
Operate 2 shifts / day, at the moment.

FMX: up to 30 samples / hour (robot)
AMX: up to 15 samples / hour (manual)
Computing / Network Infrastructures

- **BNL Institutional Cluster**
  - > 2000 KNL cores; > 1 PB GPFS

- **NSLS-II Computing Facility**
  - **FMX & AMX**
    - 9M@AMX
    - EIGER Detector
    - 16M@FMX
    - Compute Nodes (LSBR)
      - 208 cores
      - 4+4 nodes
    - GPFS Storage (NSLS-II)
      - 860 TB
      - IB
    - GPFS Buffer (LSBR)
      - 20 TB
      - IB
  - Data collection using LSDC.
  - Data visualization using Albula or ADSC.
  - Manual data processing using HKL2000 or other GUIs (ccp4i, phenix, imosflm ...), index/Best.
  - Data transfer onto USB drives.
  - Direct access to compute nodes over terminal.

- **3 x Data Collection Processing Workstations (GPFS / 10 Gb)**

- **FMX**
  - 16M@FMX
  - **EIGER Detector**
  - **Computing / Network Infrastructures**
    - Rapid Data Analysis and Reduction on dedicated compute nodes and fast buffer
    - Pipelines and packages in use: dials spot finders, dozor, XDS, dials, xia2, fast_dp, fast_mr, dimple, fast_ep, fast_ep_weak, fast_ep_NSLS-II, KAMO, and more.
    - Network backup using dedicated node.

- **40 Gb/s**

- **10 then 40 Gb/s**

- **Post processing, Re-processing, Remote processing, Multiple Crystals data reduction and Hierarchical Cluster Analysis.**

- **To world (1 Gb/s (internet) to 100 Gb/s (EsNet))**
**Eiger 9M @ AMX and 16M @ FMX**

<table>
<thead>
<tr>
<th>Model</th>
<th>Modules</th>
<th>Pixels</th>
<th>Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eiger 9M</td>
<td>18</td>
<td>10 M</td>
<td>238</td>
</tr>
<tr>
<td>Eiger 16M</td>
<td>32</td>
<td>18 M</td>
<td>133</td>
</tr>
<tr>
<td>Eiger 4M ROI</td>
<td>8</td>
<td>4.5 M</td>
<td>750</td>
</tr>
</tbody>
</table>

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Size</td>
<td>233x245</td>
<td>311x328</td>
<td>mm²</td>
</tr>
<tr>
<td>Frequency</td>
<td>238 (750)</td>
<td>133 (750)</td>
<td>Hz</td>
</tr>
<tr>
<td>Resolution</td>
<td>10.2</td>
<td>18.1</td>
<td>M pixel</td>
</tr>
<tr>
<td>Storage</td>
<td>&lt; 5</td>
<td>&lt; 8</td>
<td>M byte (actual)</td>
</tr>
<tr>
<td>Throughput</td>
<td>200-1700</td>
<td>200-1700</td>
<td>MB/s</td>
</tr>
</tbody>
</table>

2-30 TB / day / beamline

Challenges: data storage, transfer, processing and backup. Requires advanced software and computing cluster.
**Rastering**

Free-form selected by user
Automated loop detection (X-rec & InHouse) followed by coarse rastering (face ON), then fine rastering and completed by line scan (at 90 degrees).
For now: each row / or column is a collected as a vector (individual data set).
Files are converted to cbf and processed with dials spot finder client / server.

>> arms for each row or column (~3.5 secs on AMX ~ 10 secs on FMX) : 1 arm for all (in the work)
>> HW trigger: on omega position and rely on clock on Zebra Box (Quantum Detector).
Will trigger at each omega position (or pin xy positions for stills raster)

**Processing at ~ 15 FPS on one node (RAM or GPFS)**
Mixed results when ice rings are present.

>> tuning dials (Nick/Graeme slides), using dozor (Gleb slides).
>>>>> optimizing codes for faster results (ROI, HDF5 read, KNL, OMP ......)
**Data Reduction**

**Fast_dp: on 1 dedicated node (more to be used soon) @ 2-10 FPS.**
data set of 900 frames (lyso test). (all data on GPFS)

<table>
<thead>
<tr>
<th>9M</th>
<th>Total time (s)</th>
<th>FPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>FDP-005</td>
<td>90 (56 RAM)</td>
<td>10 (16)</td>
</tr>
<tr>
<td>FDP-002</td>
<td>86 (79 RAM)</td>
<td>10 (11)</td>
</tr>
<tr>
<td>FDP-010</td>
<td>83 (50 RAM)</td>
<td>11 (18)</td>
</tr>
<tr>
<td>XDS-005</td>
<td>51 (44 RAM)</td>
<td>18 (20)</td>
</tr>
<tr>
<td>XDS-002</td>
<td>61 (55 RAM)</td>
<td>15 (16)</td>
</tr>
<tr>
<td>XDS-010</td>
<td>72 (43 RAM)</td>
<td>12 (21)</td>
</tr>
<tr>
<td>FDP-004-005</td>
<td>TBRD</td>
<td>TBRD</td>
</tr>
<tr>
<td>FDP-002-004-005</td>
<td>TBRD</td>
<td>TBRD</td>
</tr>
</tbody>
</table>

XDS: jobs = all, 72 procs, 16 jobs

CPUS 004, 005: 36 cores@2.3 GHz (10 Gb)
CPUs 009, 010: 44 cores@2.2 GHz (40 Gb)
CPUs 002, 003, 007, 008: 12 cores@3.4 GHz (10 Gb)
Pipelines

DIMPLE (modified)

Start Dimples
Optional input:
1. Sigma cutoff
2. Blob number

Required input:
1. Reduced data
2. Model

In this example:
Blob number = 3
Sigma cutoff = 0.8

Output:
1. Pictures. 2. .wsh and .pol

FAST_EP_WEAK

Fast_ep_weak is a modified version of Fast_ep, which is originally developed at Diamond Light Source. Fast_ep_weak is designed to find a structure solution from weak anomalous signals by finer grid searches with spacegroup, noise and resolution cut off.

KAMO

Multiple small wedge data collection at AMX

Thermolysin Zn-SAD at AMX at 11 keV

Energy: 8
Flux: \(5.3 \times 10^{10}\) ph/sec (85% attenuation)
Total oscillation: 5 deg/spot
Osc. width: 0.2 deg/frame
Exp. time: 0.04 sec/frame (P6M)

A. Soares, Y. Yamada.
G.W., H.B.
LSDC User Interface (under continuous development)

- Layout derived from MxCuBE
- Automatic loop centering, shape recognition (XREC) + In-House code.
- User-defined raster areas & helical scanning data collection (NSLS-II: bluesky; CSS)


Conclusions & Acknowledgements

GPFS works well with 2 writes but not so well with x concurrent reads. Currently implementing the SSD buffer to be tested in the upcoming weeks. Will update nodes network as needed (40 or IB)

With beams 10 microns or less, we see more head scratching.

Data collection from multiple micro-crystals more and more used.

Not all users are prepared to handle the amount of data and reprocessing data: database a must and general cluster with remote access would help.

Learning as we go and implement new features (automation, remote, Multiple crystals data collections, in-situ ...)

Good timing for next meeting > 6 months from now!

Using KNLs: spot finder, dozor, xds? We are testing on Intel cluster: ACA workshop with Intel

ABBIX, LSBR, NSLS-II controls staff.
N. Sauter, G. Winter, S. Popov, G. Bourenkov.
W. Shi, H. Bernstein, Y. Yamada#.