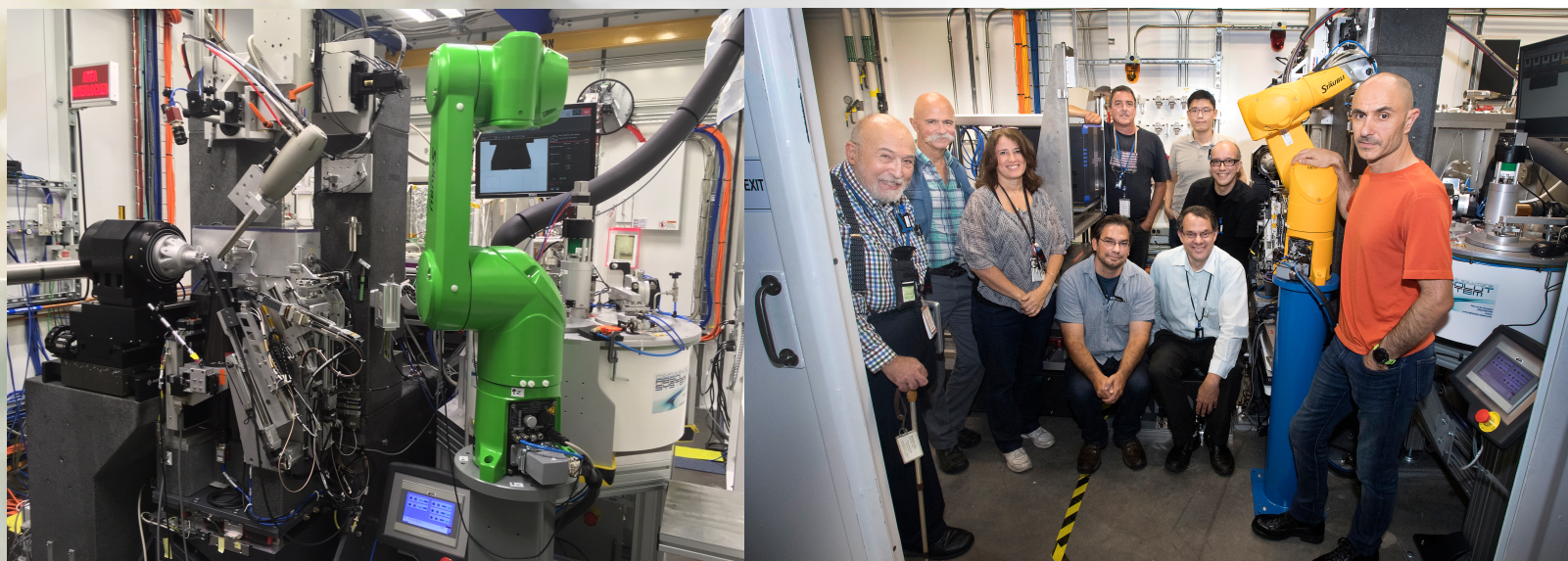
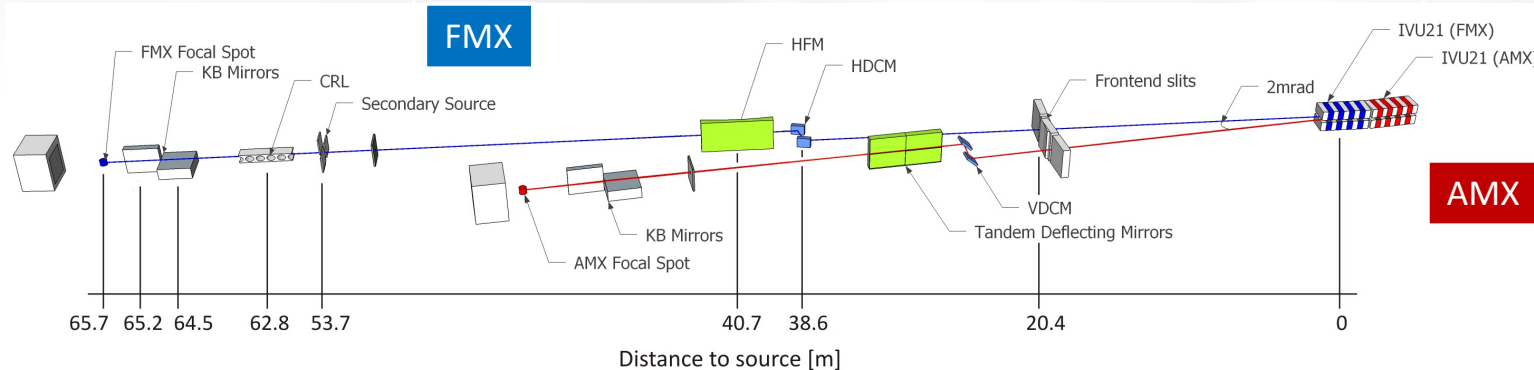


High Data Rate MX (HDRMX) at NSLS-II



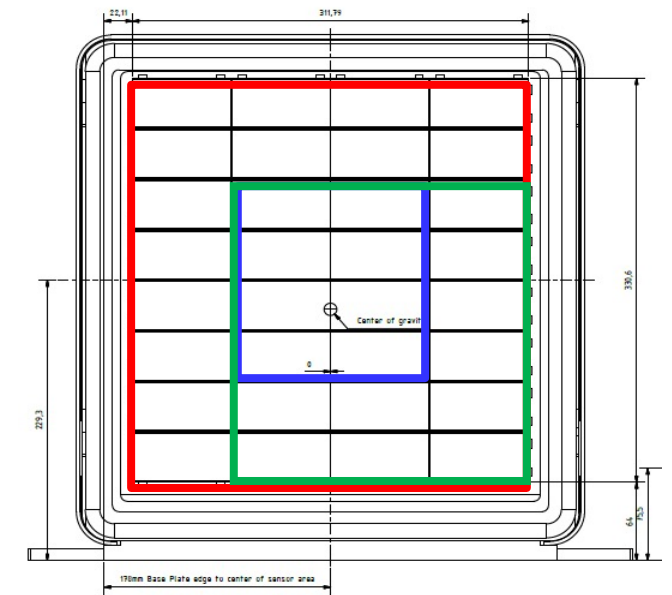
Dale Kreidler
March 17, 2021
Upton, New York

Beamline Characteristics



Specifications

	AMX	FMX
Wavelength range	0.7 – 2.5 Å	0.4 – 2.5 Å
Flux at focus at 12.7 keV	$>4 \times 10^{12}$ ph/s	3.5×10^{12} ph/s
Focal spot min (H×V)	$7 \times 5 \mu\text{m}^2$	$1.5 \times 1 \mu\text{m}^2$
Detector	Eiger 9M (200Hz)	Eiger 16M (100Hz)



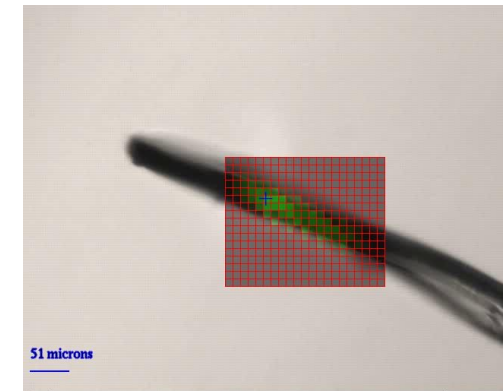
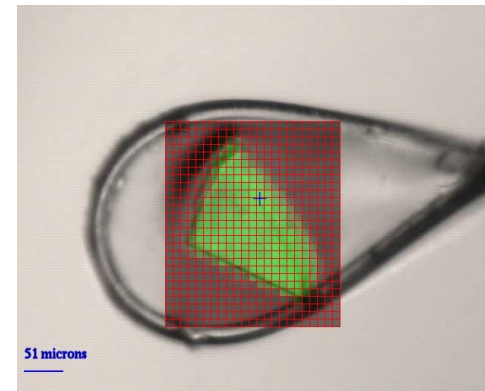
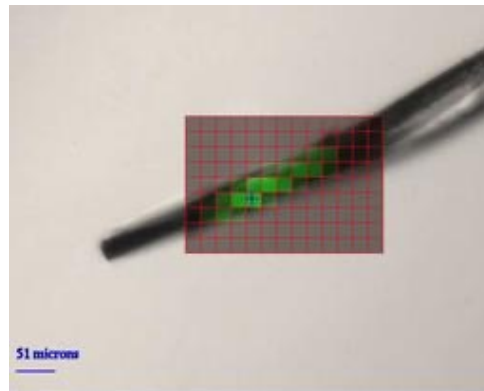
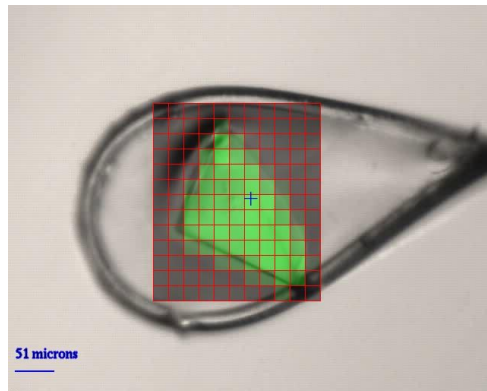
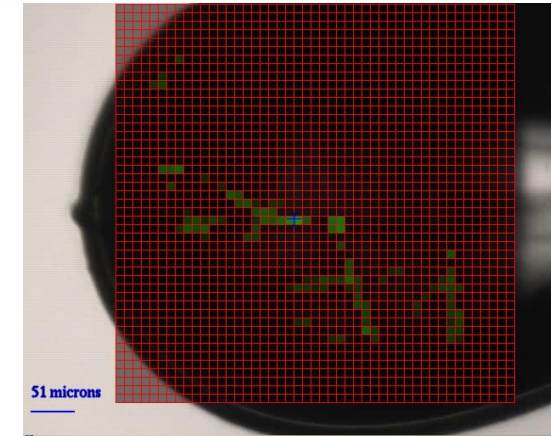
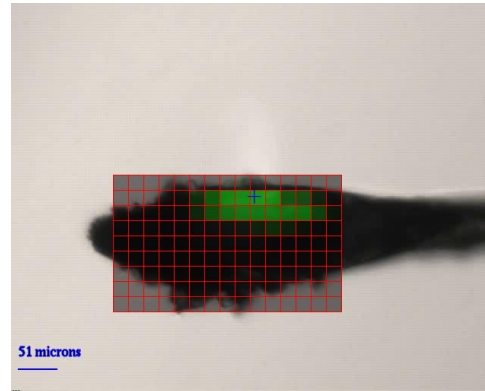
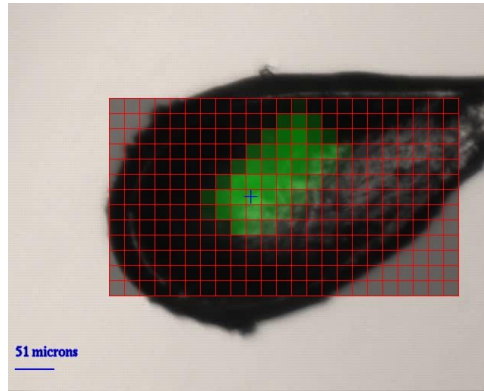
2-30 TB / day / beamline

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

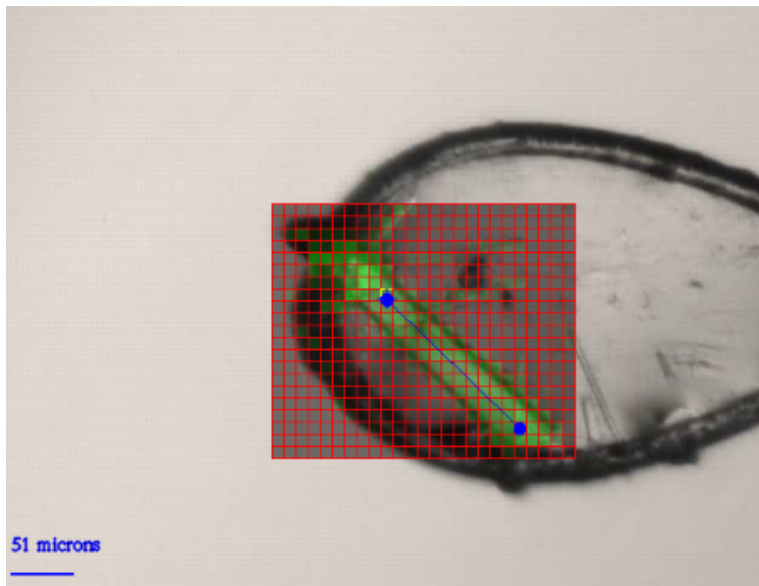
Fuchs, M. R. et al. "NSLS-II biomedical beamlines for micro-crystallography, FMX, and for highly automated crystallography, AMX: New opportunities for advanced data collection *AIP Conf. Proc. SRI2015*, **2016**, 1741, 030006

Raster-based centering

Raster every sample



Vector (helical) data collections



AMX Vector

Parameters

Start X	5068.688 um	5588.981 um
Start Y	-402.200 um	200.100 um
Start Z	1637.500 um	1520.000 um
Start Omega	0.00 deg	-0.00 deg
End X	5068.688 um	5588.981 um
End Y	-402.200 um	200.100 um
End Z	1637.510 um	1520.000 um
End Omega	360.00 deg	360.00 deg
Exposure	400.000000 n	400.000000 n
# Samples	9.000000	9.000000

Expose ☐ Hold ☐ State ● Idle

Go Proceed **ABORT**

More Parameters

Buffer Time 3.000000 ms 3.000000 ms

Status

Data Acq Duration 3600.000000 Shutter Opening Time 3.666675 ms
 Maximum Time to Speed 115.199997 n utter Change in Position 0.000000 ct

	Stack X	Pin Y	Pin Z	Omega
Too Fast	●	●	●	●
Acceleration	5.000 ct/ms2	4000.000 ct/n	4000.000 ct/n	50.000 ct/ms2
Data Acq Distance	0 ct	0 ct	0 ct	20736000 ct
Desired Speed	0.000000 ct/n	0.000000 ct/n	0.000000 ct/n	5760.000000
Time to Speed	0.000000 ms	0.000000 ms	0.000000 ms	115.199997 n
Direction	1.000000	1.000000	1.000000	1.000000
Speed Up Distance	0.000000 ct	0.000000 ct	0.000000 ct	331775.99256
Buffer Distance	0.000000 ct	0.000000 ct	0.000000 ct	17280.000000
Shut Open Distance	0.000000 ct	0.000000 ct	0.000000 ct	21120.050511
Shut Lag Distance	0.000000 ct	0.000000 ct	0.000000 ct	11520.000000
Backup Distance	0.000000 ct	0.000000 ct	0.000000 ct	381696.04305

Motors

Omega	170.673943 deg	170.673943 deg	< 2.000000 deg > STOP More
Pin Y	329.240 um	329.200 um	< 50.000 um > STOP More
Pin Z	1535.138 um	1535.100 um	< 100.000 um > STOP More
Stack X	5555.43 um	5555.43 um	< 350.00 um > STOP More
Shutter Rot	15.00 deg	15.00 deg	< 5.00 deg > STOP More
Open Position	-10.000 deg	-10.000 deg	
Closed Position	15.000 deg	15.000 deg	

Shutter

ZEBRA

PC AND OR GATE DIV PULSE ENC SYS

Diagram illustrating the ZEBRA pulse sequence and gate timing. The diagram shows a sequence of pulses (Pulse, Gate, Step) and gates (Gate, Step) with labels for Width, Delay, and Capture position here. A legend indicates: Disarm - Soft (Hardware Num Gates reached).

Setup

Capture: ☐ Enc1 ☐ Enc2 ☐ Enc3 ☒ Enc4 ☐ Sys1 ☐ Sys2 ☐ Div1 ☐ Div2 ☐ Div3 ☐ Div4

Posn Trig Enc4 Enc4 Time Units ms ms

Posn Dir Positive Positive

Arm

Trig Source External External 1 IN1_TTL Disarm Arm Status ●

Gate

Trig Source Position Position

Gate Start	Gate Width	Gate Step	Num Gates	Gate Status
0.0000 um	20.0000 um	40.0000 um	9	●

Pulse

Trig Source Time Time

Pulse Start	Pulse Width	Pulse Step	Capt Delay	Max Pulses	Pulse Status
0.0000 ms	9.9800 ms	10.0000 ms	5.0000 ms	69	●

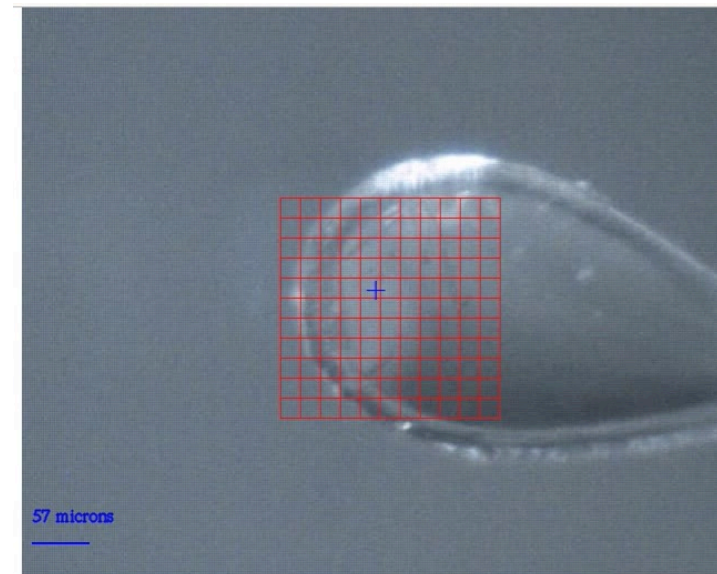
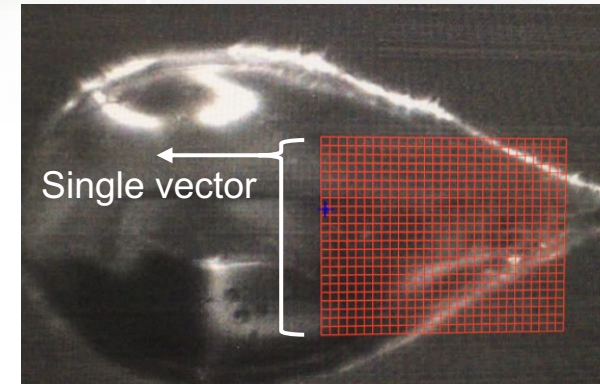
Captured Data

Display Captured Data... Data Download in Progress ●

Comms: Connected Soft In: ☐ INP1 ☐ INP2 ☐ INP3 ☐ INP4 Block State: Reset Exit

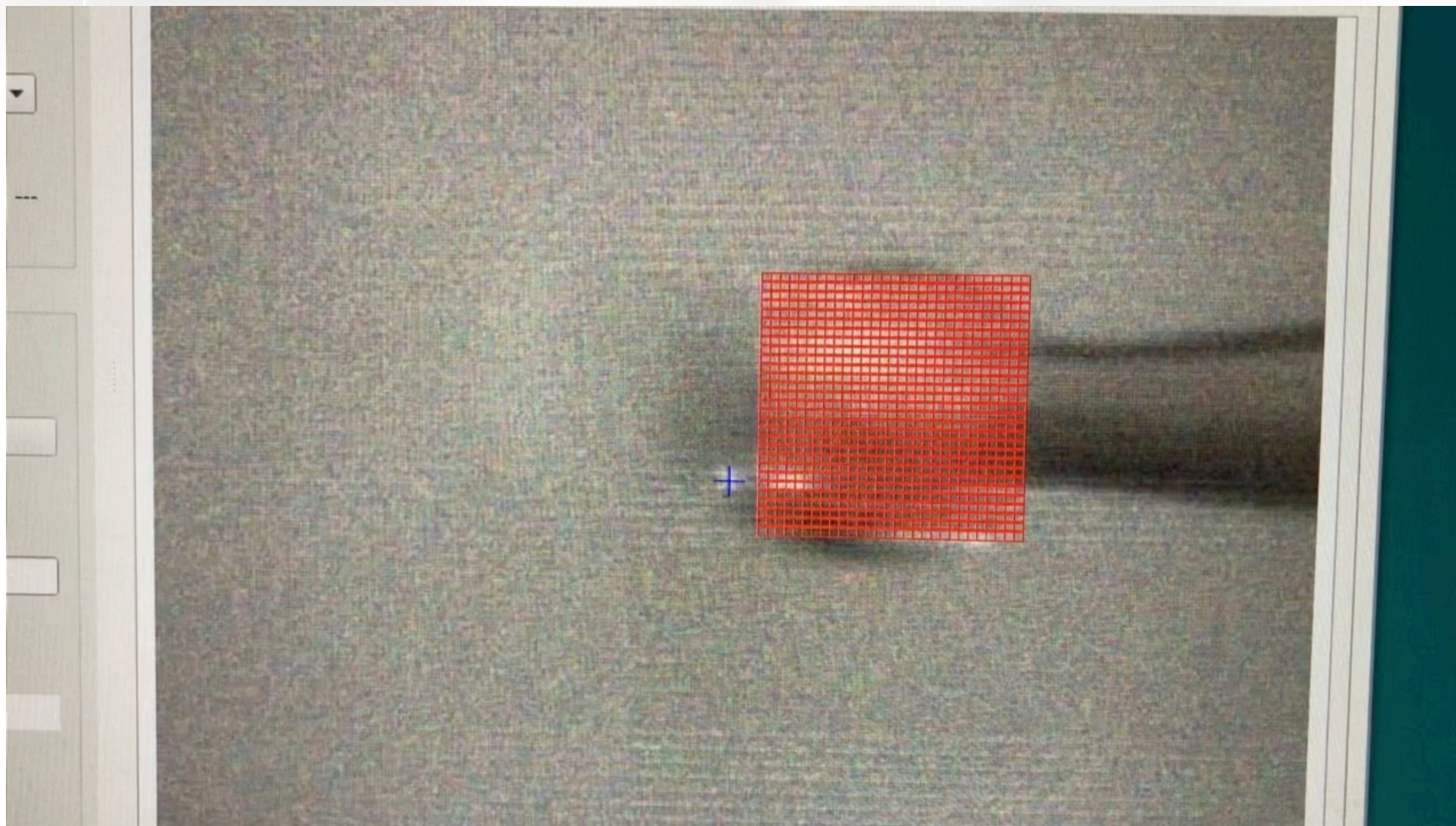
Vector-based rastering and data collection

- Every collection is a 'vector'
- Raster centering instead of optical centering
- Vector collection by row
 - data.h5 file written for each row
 - Python LSDC server generates dozor input file based on metadata from database request object
 - Dozor processing triggered on remote node (per row)
 - Decompression with dextris-neggia
 - Also can reprocess with `dials.find_spots`
 - Only rectangular rasters (single detector arm)

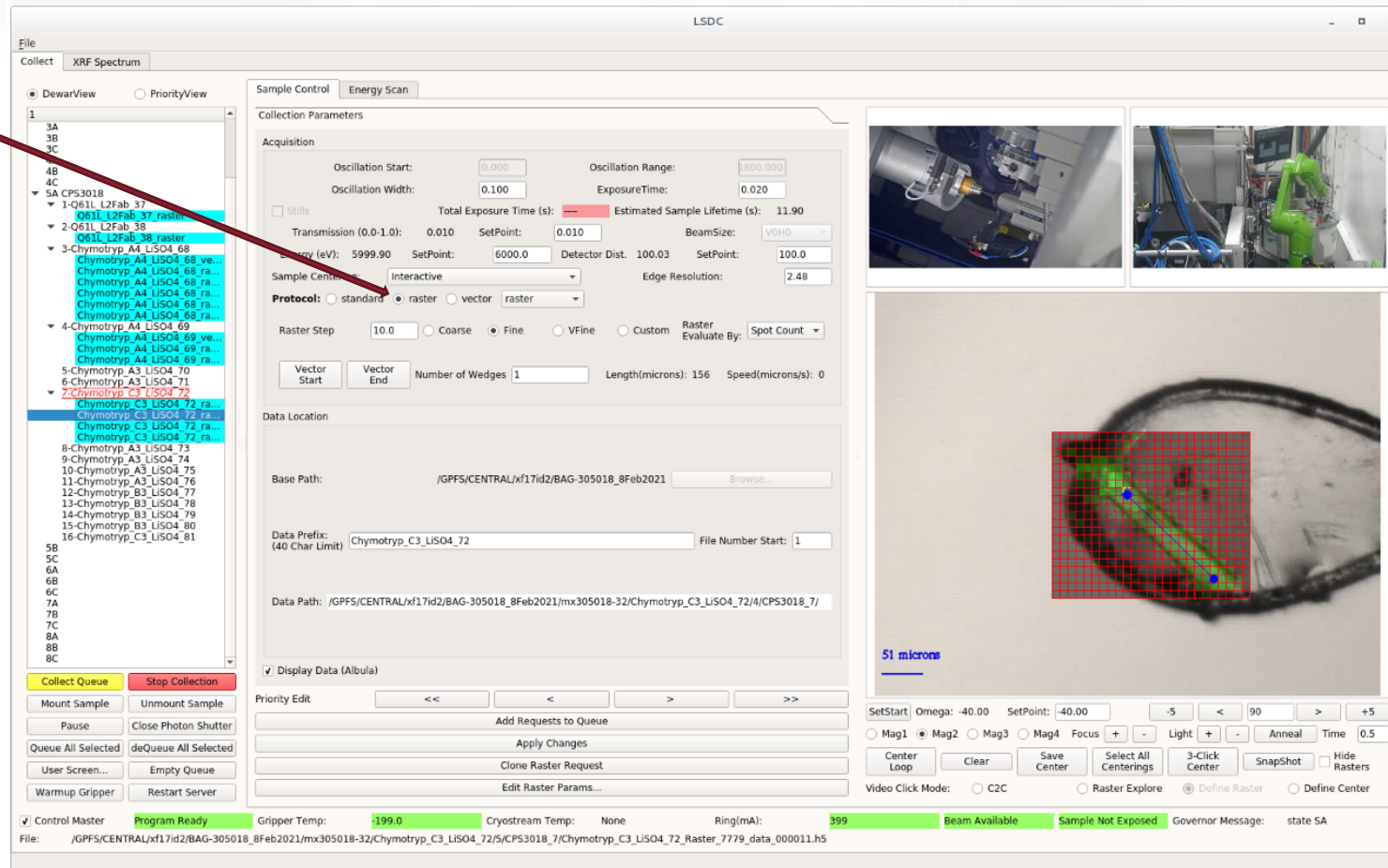


```
PIN_ALIGN: DEBUG enabled
PIN_ALIGN_ROI_WIDTH: 150
PIN_ALIGN_ROI_HEIGHT: 400
PIN_ALIGN_ROI_WIDTH_OFFSET: 410
PIN_ALIGN: Y motor axis is up
PIN_ALIGN: Z motor axis is down
argument 0: /GPFS/CENTRAL/xf17id1/skinnerProjectsBackup/pinAlign_fm.xsh
argument 1: MKP1268_9_4_8_0_PA_0_001.jpg
argument 2: MKP1268_9_4_8_0_PA_90_001.jpg
Task MOUNT isRunning = false
Running Task PARK true
PIN_ALIGN: DEBUG enabled
PIN_ALIGN_ROI_WIDTH: 150
PIN_ALIGN_ROI_HEIGHT: 400
PIN_ALIGN_ROI_WIDTH_OFFSET: 410
PIN_ALIGN: Y motor axis is up
PIN_ALIGN: Z motor axis is down
PIN_ALIGN: DEBUG enabled
PIN_ALIGN_ROI_WIDTH: 150
PIN_ALIGN_ROI_HEIGHT: 400
PIN_ALIGN_ROI_WIDTH_OFFSET: 410
PIN_ALIGN: Y motor axis is up
PIN_ALIGN: Z motor axis is down
Task PARK isRunning = false
Running Task UNLATCHROBGOV true
Task UNLATCHROBGOV isRunning = false
```

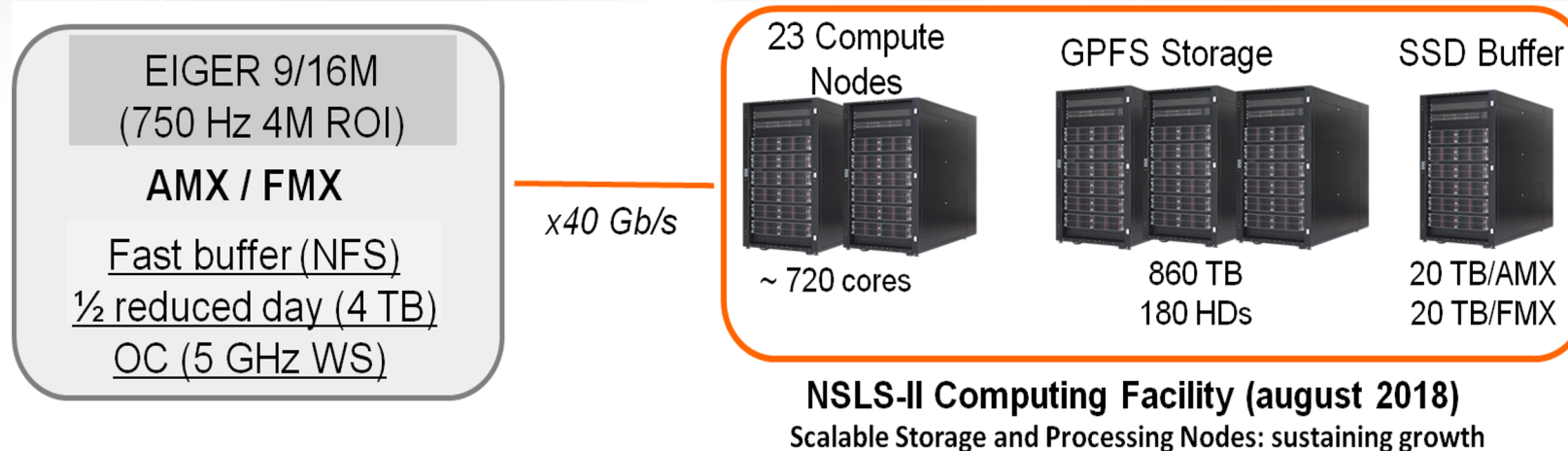
50 Hz; usage typically 50-200 Hz
500 cell raster, ~2.5-10 sec total exposure



Simple raster queueing interface



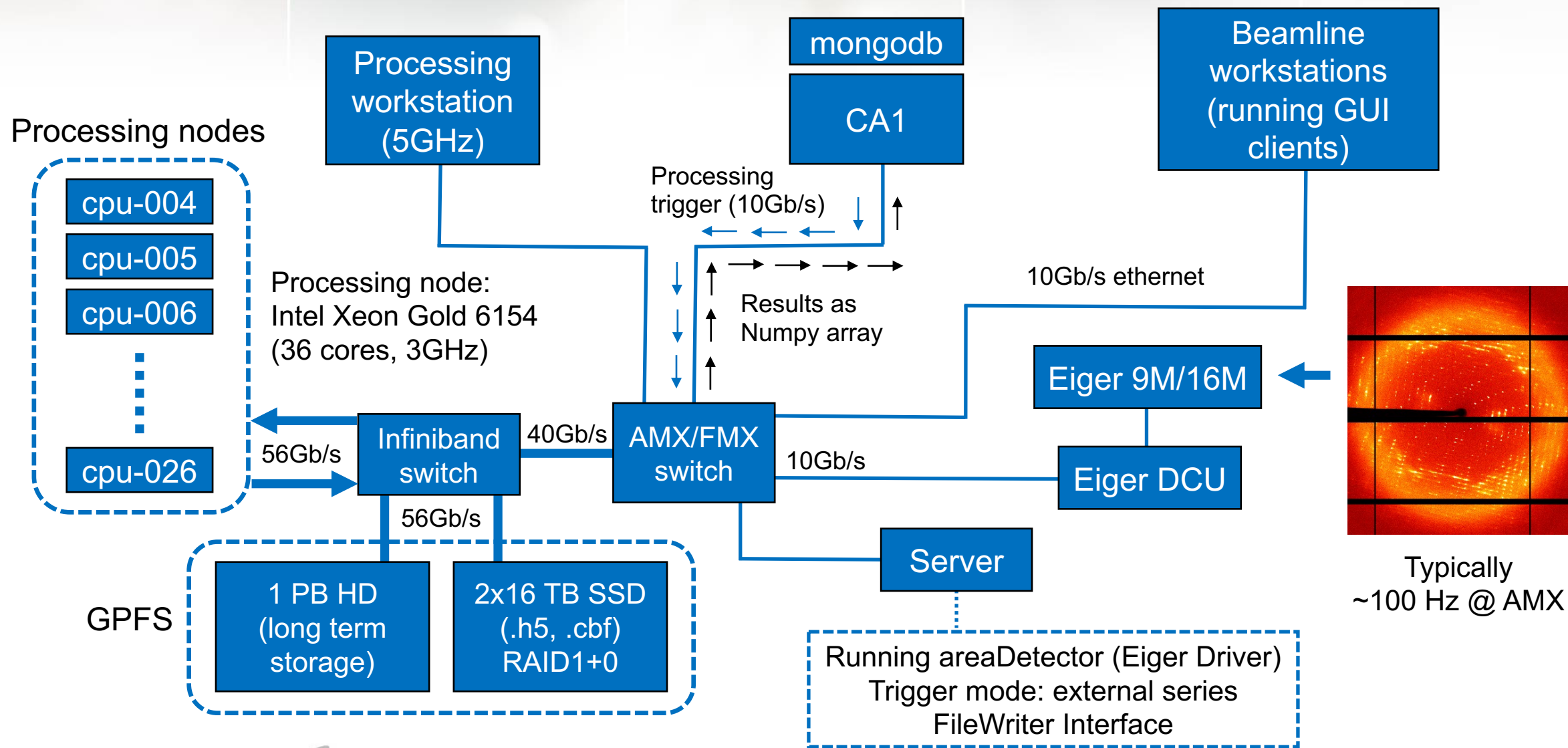
AMX/FMX Infrastructure



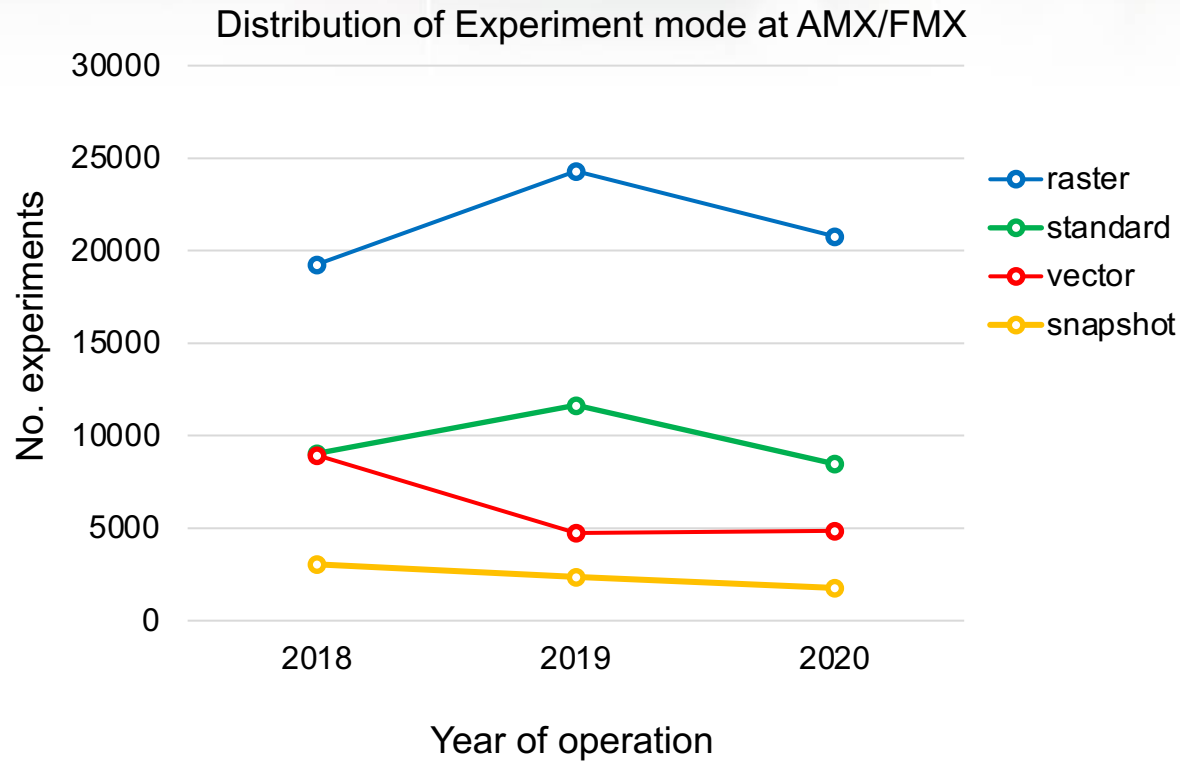
- @ beamlines (2x): 3 WS, 1 data backup WS, 1 OC WS (single processes) and CAs SRVs ... (all on 10 Gbs and on GPFS)
- @ NSLS-II central facility: 23 nodes: ~ 720 cores (56 Gb/s IB to GPFS; 10 Gbs/ node to node)
- @ NSLS-II central facility: 2 x 16 TB SSD fast buffer (NSDs) GPFS

Data transfer to users with globus (<https://www.globus.org>), no hard drives

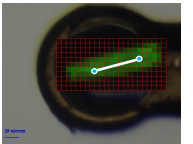
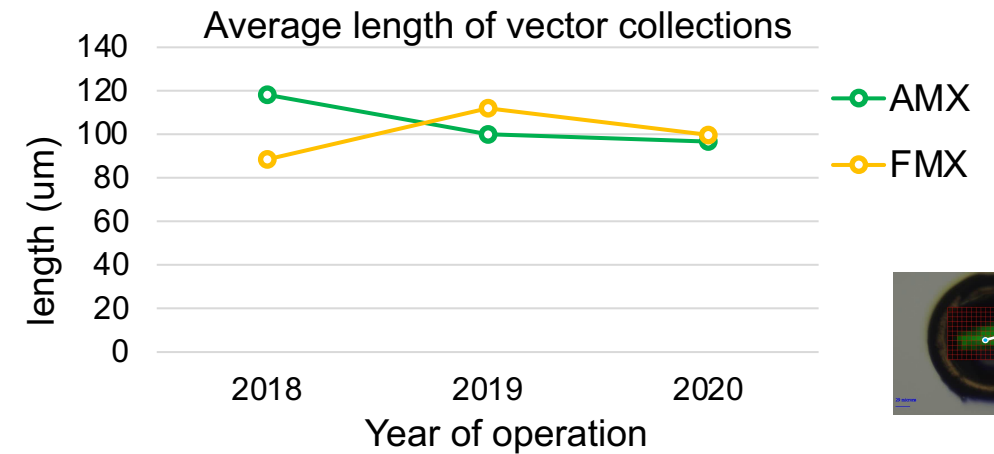
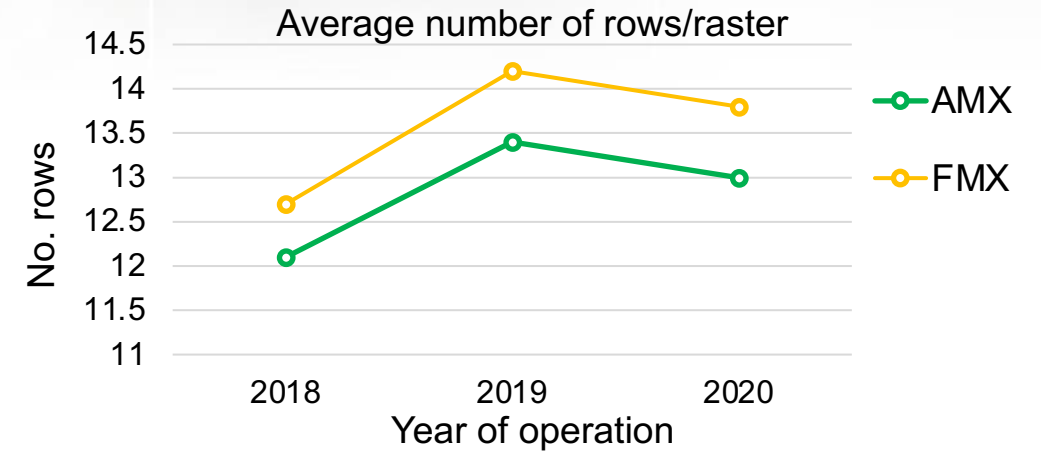
Network architecture for raster processing



Database statistics of AMX/FMX collections

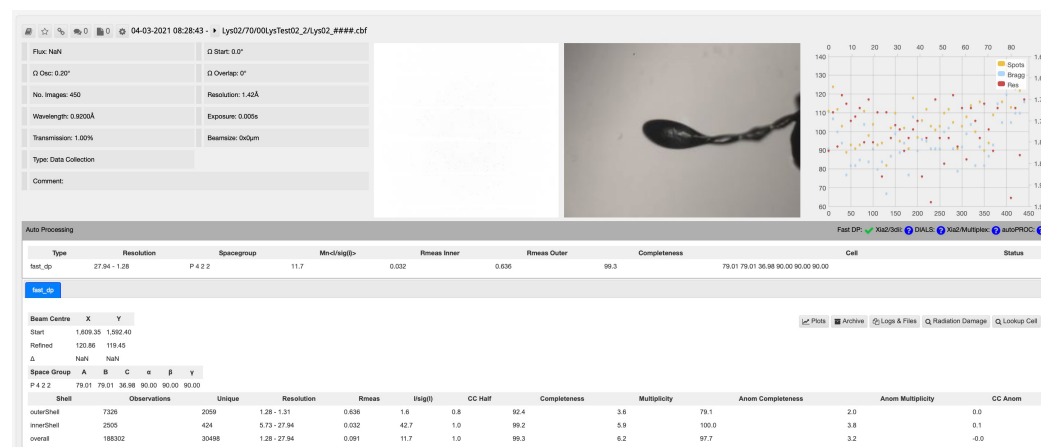
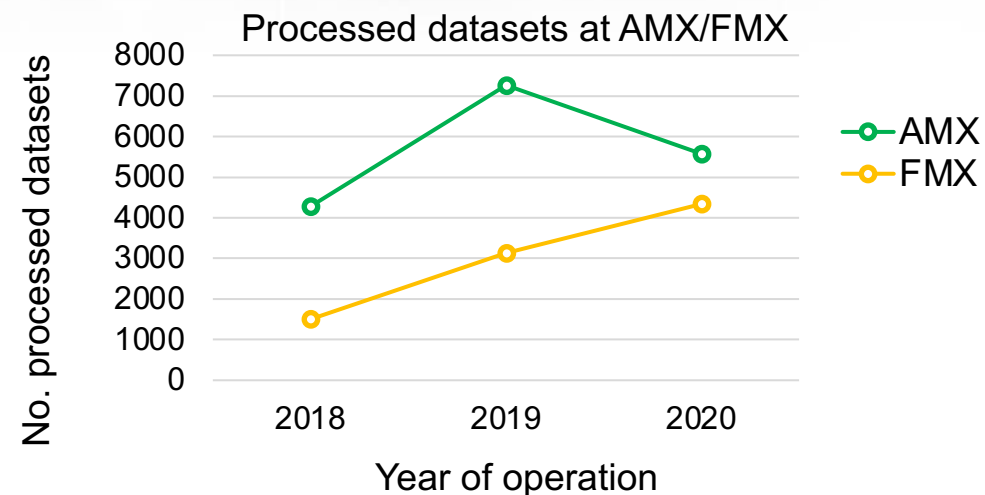


- ~270k vector collections in 2020
- ~2 rasters/dataset
- 10 sec raster time improvement → ~55 h of annual beamtime



Pipelines: In-line processing with fast_dp, dimple

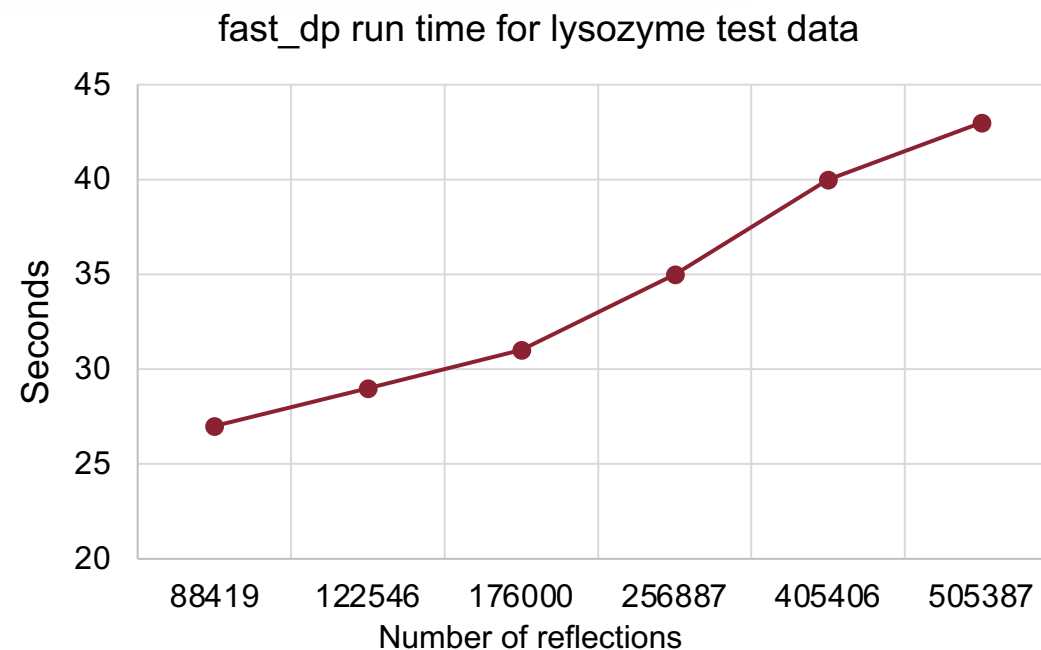
- Autoprocessing of all standard, vector datasets triggered after collection (XDS/fast_dp)
- Typical dataset ~10 seconds (180 deg, 0.2 deg, 100Hz)
- In automated mode, ~1 unipuck/hour
- Well diffracting samples (ligand binding or dynamics) — ~4 min/dataset including sample exchange
- 2 dedicated 36 core nodes for spotfinding, indexing, integration
- pointless/aimless run on overclocked workstation



ISPyB

Data processing timing

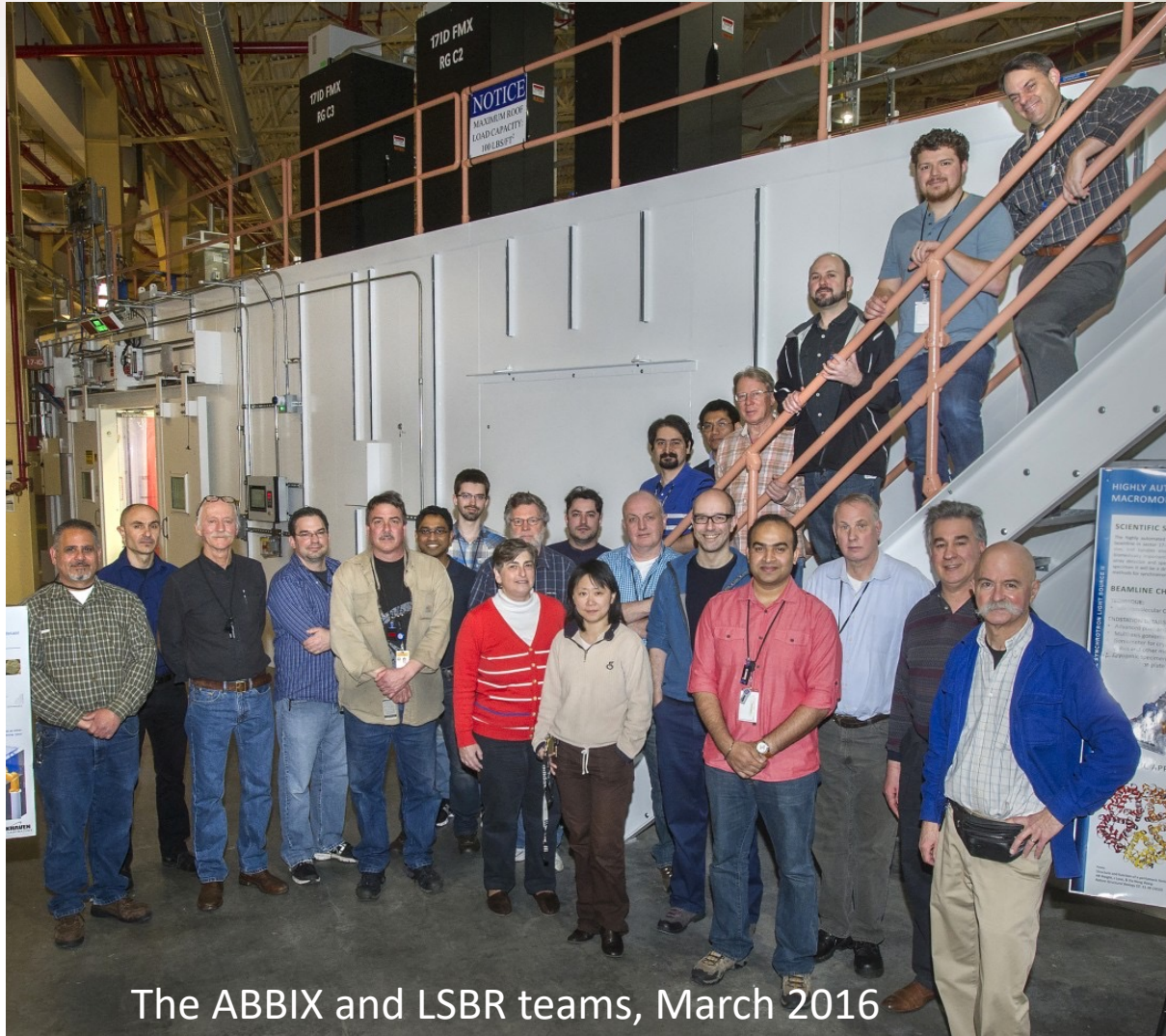
- From lysozyme start up data
- 0.2 deg, 150 deg (750 frames)
- Eiger 9M
- 1 Angstrom resolution
- 8 nodes (XDS):
 - Intel Xeon Gold 6154
 - 36 cores, 3 GHz
- AMX workstation (pointless/aimless):
 - Intel Core i7-8700K
 - 6 cores
 - overclocked @ 5 GHz
- Dimple (tetragonal lysozyme; 12-36 seconds)
- If achieving 1000 samples/day need, processing must keep pace (1 dataset/min)



Considerations for the future

- How to keep pace with throughput without compromising quality?
- Wealth of new information, but avoid analysis paralysis
- Automation vs. manual collection
- New types of experiments enabled by HDRMX: clustering, dynamics, time-resolved, etc. How do we use these new tools?
- Precision of sample stages/goniometers, stable optics required by faster detectors

Acknowledgments



The ABBIX and LSBR teams, March 2016

- John Lara
- Jean Jakoncic
- Dieter Schneider
- Stu Myers
- Tom Langdon
- *Shirish Chodankar*
- Bruno Seiva Martins
- Vivian Stojanoff
- John Skinner
- *Hugo Slepicka*
- Kun Qian
- Sean McSweeney
- Martin Fuchs
- Babak Andi
- Dale Kreidler
- *Lin Yang*
- Neil Whalen
- Dileep Bhogadi
- Rick Jackimowicz
- Lonny Berman
- Edwin Lazo
- Bob Sweet
- Alexei Soares
- Wuxian Shi
- Yuan Gao
- Jim Magill
- Nicolas Guichard
- *Johnathan DiFabio*
- *Oksana Ivashkevych*
- Herbert Bernstein
- Yusuke Yamada

Software:

Dozor -- Gleb Bourenkov, Alexander Popov
fast_dp developers/maintainers
Graeme Winter, Markus Gerstel, Richard Gildea,
Nicholas Devenish, and others

This work is supported by the US National Institutes of Health and the US Department of Energy