

# BEAMLINE TO HPC: CODED APERTURE LAUE MICRODIFFRACTION

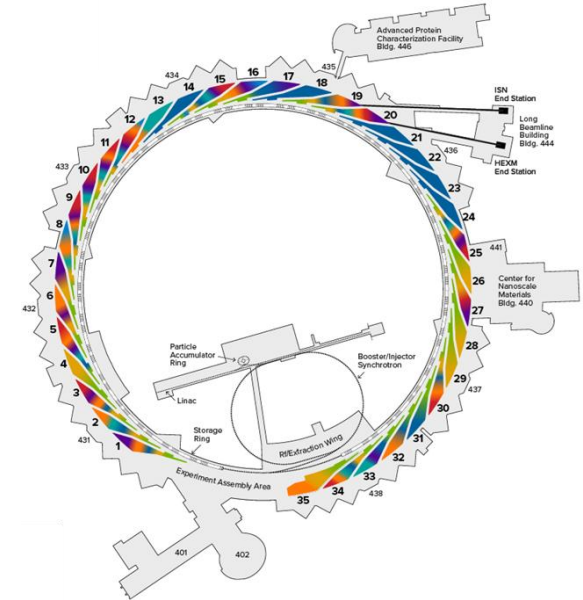
**Michael Prince**, Doğa Gürsoy, Dina Sheyfer,  
Ryan Chard, Benoit Côté, Hannah Parraga,  
Barbara Frosik, Jon Tischler, Nicholas Schwarz

# A NOTE ON THIS PRESENTATION

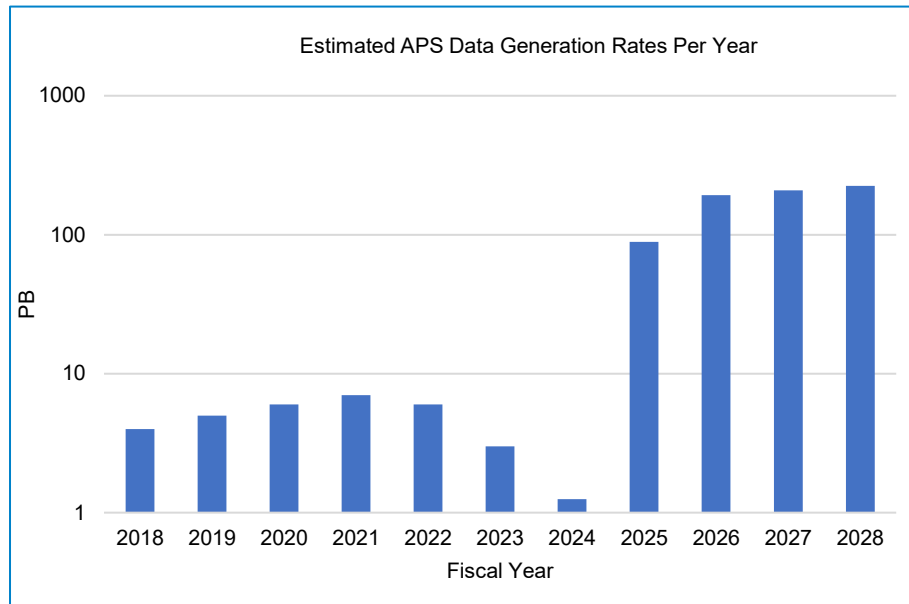
- Most work done in late 2022-2023
  - Pre-AI
- Beamline has been down since mid-2023
  - Early 2D experiments ~2 weeks ago
  - HPC-scale experiments starting in winter

# ADVANCED PHOTON SOURCE

- 72 beamlines
- 5,500 users per year
- ~2000 publications per year
- 20% of AlphaFold's training data was characterized at the APS



# HANDLING THE APS UPGRADE

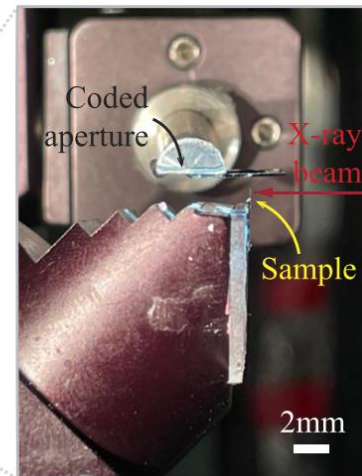
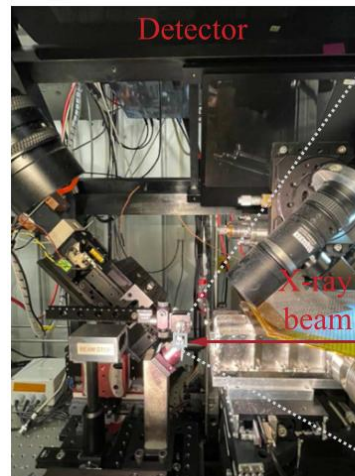
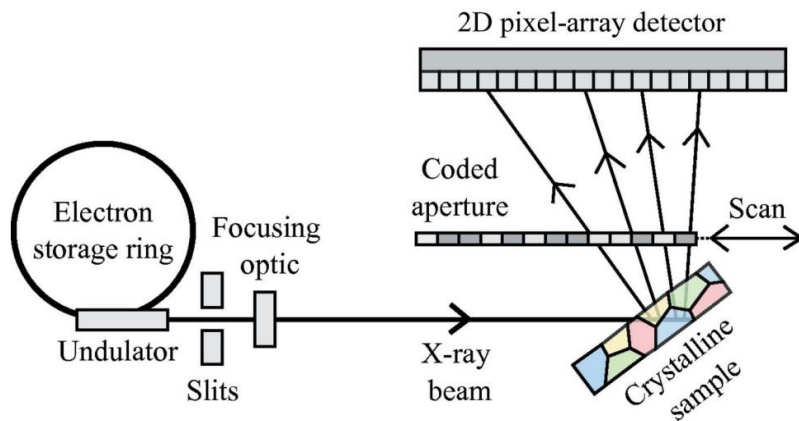


# BEAMTIME IS EVERYTHING

- Users are assigned shifts to conduct experiments
- **Any** time not putting atoms in front of beams feels wasted
- Culture of “fire and forget”

| Feb |  | Mar |  | Apr |  | May |  |
|-----|--|-----|--|-----|--|-----|--|
| 1   |  | 1   |  | 1   |  | 1   |  |
| 2   |  | 2   |  | 2   |  | 2   |  |
| 3   |  | 3   |  | 3   |  | 3   |  |
| 4   |  | 4   |  | 4   |  | 4   |  |
| 5   |  | 5   |  | 5   |  | 5   |  |
| 6   |  | 6   |  | 6   |  | 6   |  |
| 7   |  | 7   |  | 7   |  | 7   |  |
| 8   |  | 8   |  | 8   |  | 8   |  |
| 9   |  | 9   |  | 9   |  | 9   |  |
| 10  |  | 10  |  | 10  |  | 10  |  |
| 11  |  | 11  |  | 11  |  | 11  |  |
| 12  |  | 12  |  | 12  |  | 12  |  |
| 13  |  | 13  |  | 13  |  | 13  |  |
| 14  |  | 14  |  | 14  |  | 14  |  |
| 15  |  | 15  |  | 15  |  | 15  |  |

# CODED APERTURE LAUE MICRODIFFRACTION



**MatSci applied to microelectronics, batteries, geology, etc**  
**Up to a 5x reduction in beamtime vs wire scans**

Gürsoy, Doğa, Dina Sheyfer, Michael Wojcik, Wenjun Liu, and Jonathan Z. Tischler. "Depth-resolved Laue microdiffraction with coded apertures." *Journal of Applied Crystallography* 55, no. 5 (2022).

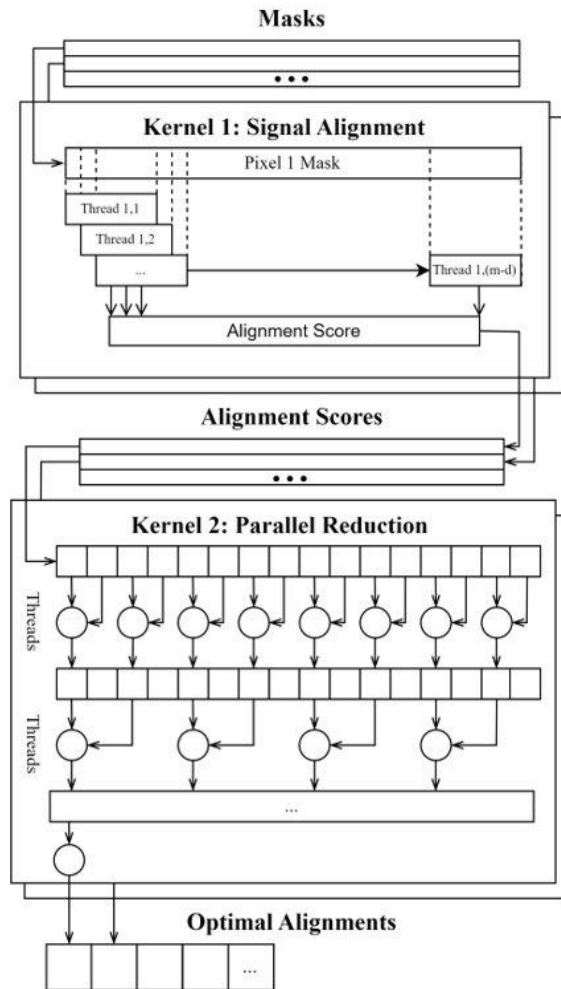
Gürsoy, Doğa, et al. "Digital autofocusing of a coded-aperture Laue diffraction microscope." *Review of Scientific Instruments* 94.1 (2023): 013702.

# TRADEOFF: COMPUTATIONAL COMPLEXITY

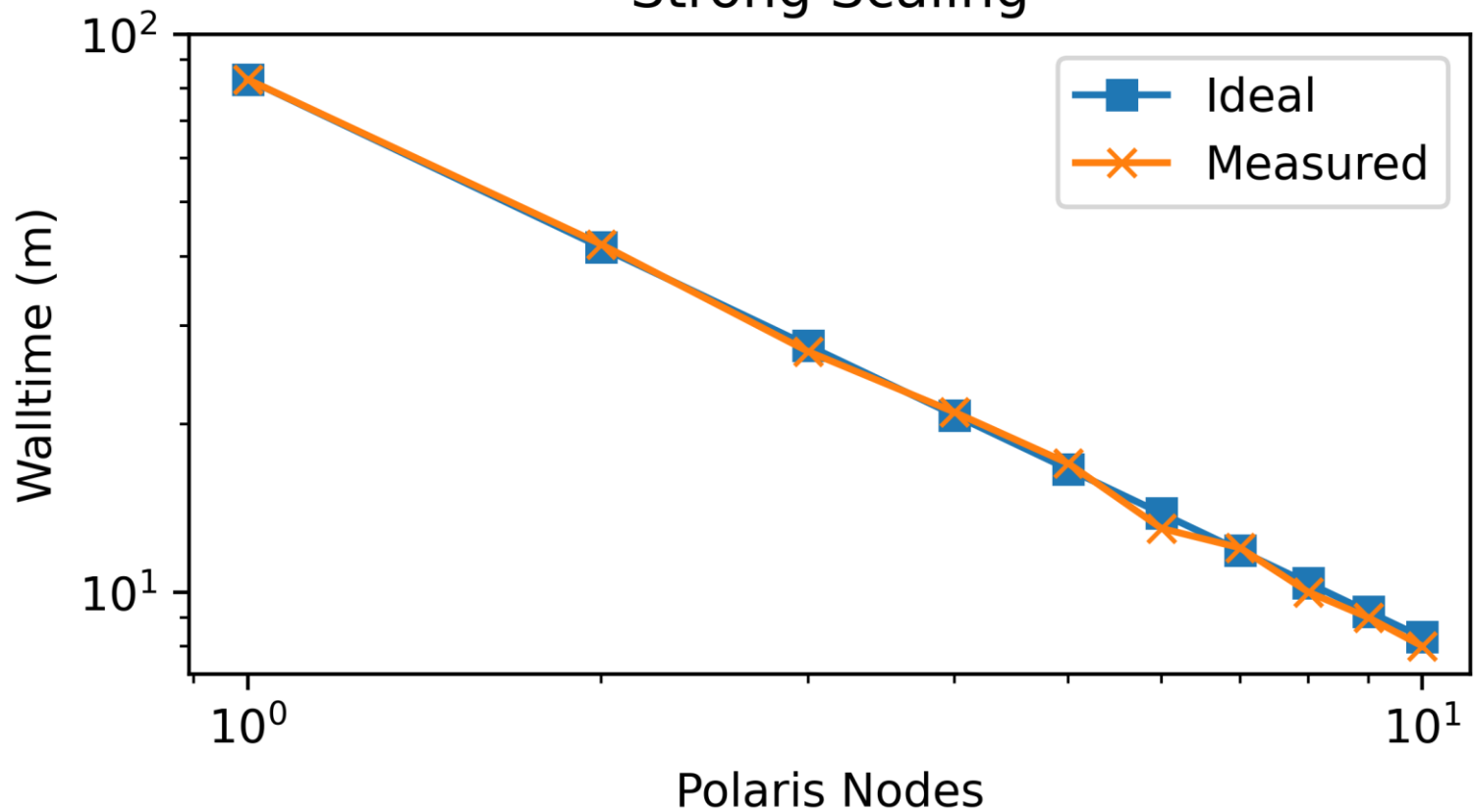
- ~450 points per scan
  - ~1.5 min per point
  - ~**11 hours per scan**
- Est. Processing (CPU Implementation, Oct. 2022)
  - Single core: 15 days/point (**18.75 years/scan**)
  - 1 Polaris Node: 20 hours/point (**375 days/scan**)
- **Unique Position:** reason to use HPC pre-APS-U
  - Allocated 55 Polaris nodes & preemptable queue
  - Experiments scheduled in 6 months
  - Can we keep up with the detector?

# OPTIMIZATION

- GPU: Custom CUDA kernels
    - Heavy memory optimization
    - 800x segment speedup
    - (When GPT-o1 came out, this is the first thing I tested on it)
  - CPU: Jax JIT compilation
    - Cache & memory optimizations
    - 4x segment speedup
- **14-15x speedup in practice**
- ~~20 hours/point\* (375 days/scan)~~
  - **1.4 hours/point (26 days/scan)**



## Strong Scaling



# POINT-LEVEL PARALLELISM

- Increases throughput
  - ~~26~~ days/scan
  - **12 hours/scan @ 55 nodes**
- Mostly there for optimization.

```
-----  
440154.polaris-p* laue_aps_realti* mprince 25:30:56 R demand  
440155.polaris-p* laue_aps_realti* mprince 19:05:28 R demand  
440156.polaris-p* laue_aps_realti* mprince 11:18:02 R demand  
440157.polaris-p* laue_aps_realti* mprince 08:57:42 R demand  
440158.polaris-p* laue_aps_realti* mprince 02:00:51 R demand  
440159.polaris-p* laue_aps_realti* mprince 0 Q demand  
440160.polaris-p* laue_aps_realti* mprince 0 Q demand  
440161.polaris-p* laue_aps_realti* mprince 0 Q demand  
440162.polaris-p* laue_aps_realti* mprince 0 Q demand  
440163.polaris-p* laue_aps_realti* mprince 0 Q demand  
440164.polaris-p* laue_aps_realti* mprince 0 Q demand
```

# DATA TRANSFER SYSTEM REQUIREMENTS

- Autonomously detect and process new data from the APS
- Transfer data to and from the facilities
- Create and monitor jobs on Polaris



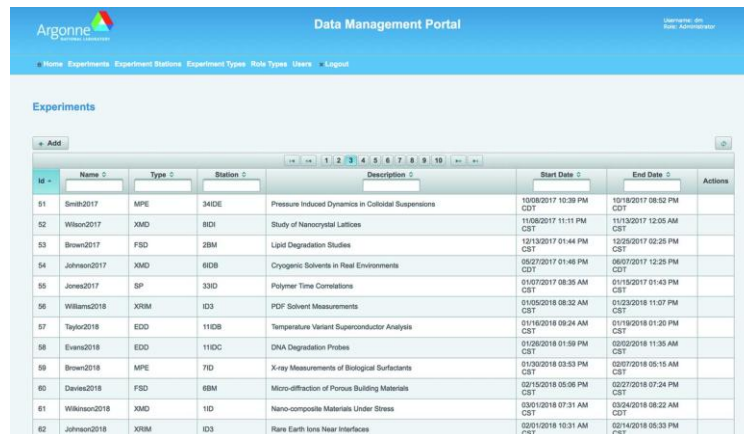
# DATA TRANSFER SYSTEM TOOLS

- Data Management System

- User Isolation
- Internal Data Processing

- Globus

- Grid FTP
- Remote SSO
- Remote Queue Access

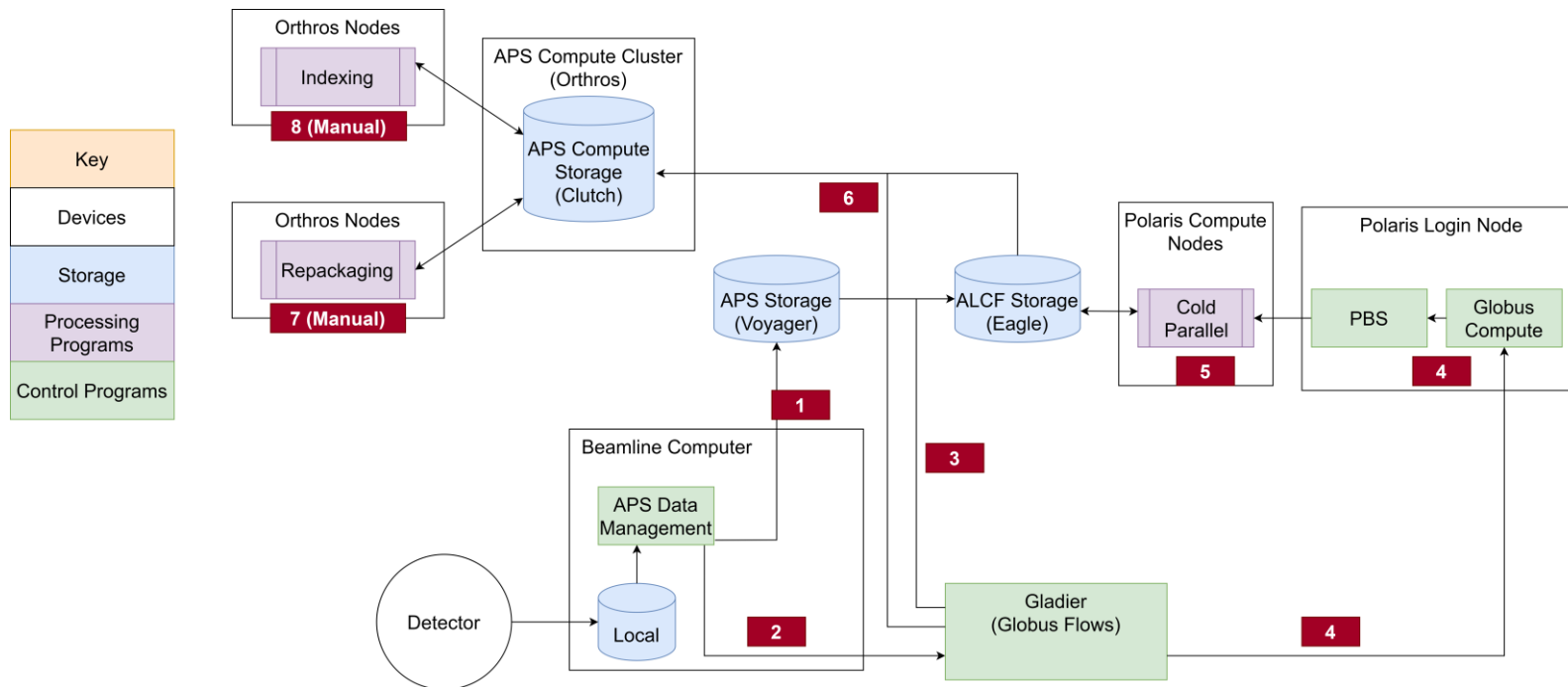


The screenshot shows the Argonne Data Management Portal. At the top, there is a navigation bar with the Argonne logo and the text "Data Management Portal". Below the navigation bar, there is a section titled "Experiments" which contains a table of experimental data. The table has columns for ID, Name, Type, Station, Description, Start Date, End Date, and Actions. The data is organized into 10 rows, each representing a different experiment.

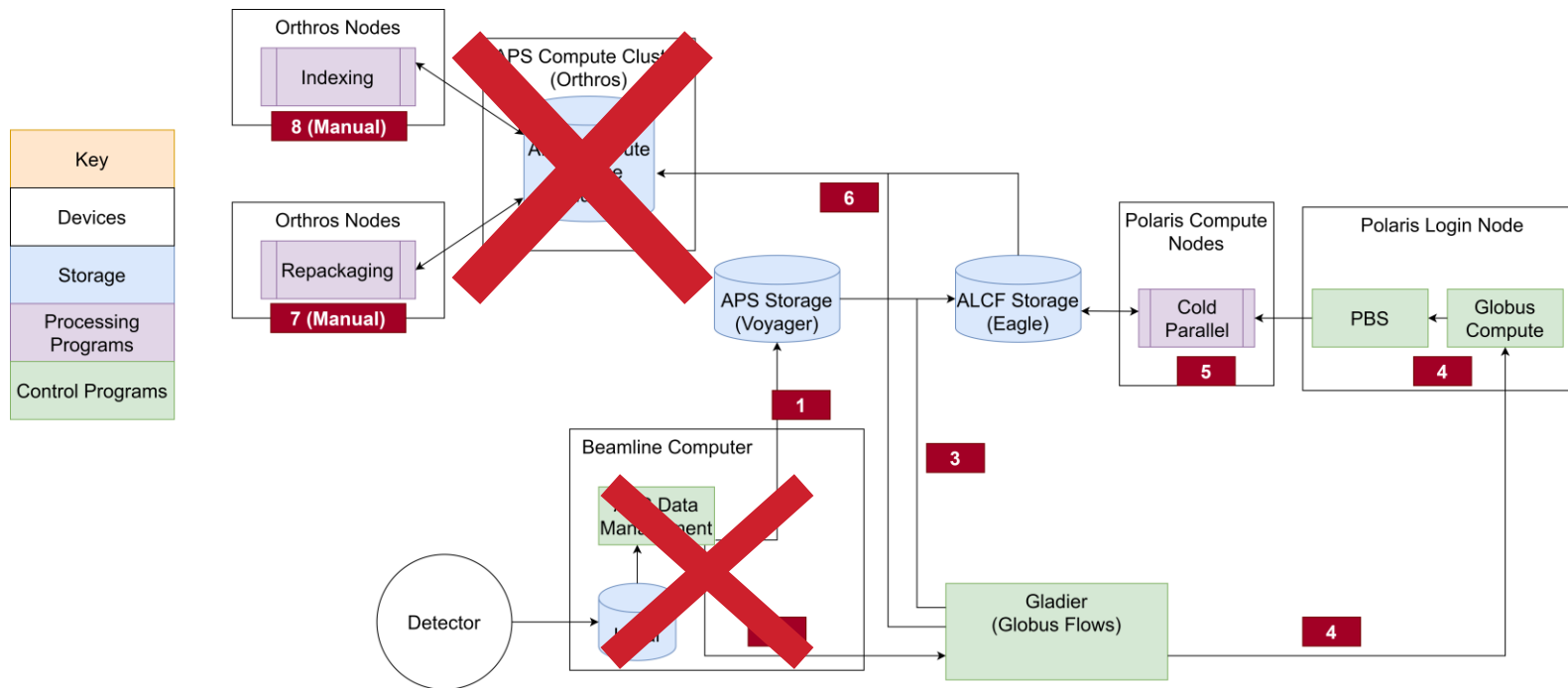
| ID | Name         | Type | Station | Description                                        | Start Date              | End Date                | Actions |
|----|--------------|------|---------|----------------------------------------------------|-------------------------|-------------------------|---------|
| 51 | Smith2017    | MPE  | 34IDE   | Pressure Induced Dynamics in Colloidal Suspensions | 10/18/2017 10:39 PM CST | 10/18/2017 08:52 PM CST |         |
| 52 | Wilson2017   | XMD  | 8DI     | Study of Nanocrystal Lattices                      | 11/08/2017 11:11 PM CST | 11/13/2017 12:05 AM CST |         |
| 53 | Brown2017    | FSD  | 2BM     | Lipid Degradation Studies                          | 12/15/2017 01:44 PM CST | 12/29/2017 02:25 PM CST |         |
| 54 | Johnson2017  | XMD  | 8IDB    | Cryogenic Solvents in Real Environments            | 05/27/2017 01:48 PM CST | 06/07/2017 12:25 PM CST |         |
| 55 | James2017    | SP   | 33D     | Polymer Time Correlations                          | 01/07/2017 08:35 AM CST | 01/15/2017 01:43 PM CST |         |
| 56 | Williams2018 | XRM  | ID3     | PDF Solvent Measurements                           | 01/05/2018 08:32 AM CST | 01/23/2018 11:07 PM CST |         |
| 57 | Taylor2018   | EDD  | 11IDB   | Temperature Variant Superconductor Analysis        | 01/16/2018 09:24 AM CST | 01/16/2018 01:20 PM CST |         |
| 58 | Evens2018    | EDD  | 11IDC   | DNA Degradation Probes                             | 01/26/2018 01:59 PM CST | 02/02/2018 11:35 AM CST |         |
| 59 | Brown2018    | MPE  | 7ID     | X-ray Measurements of Biological Surfactants       | 01/30/2018 03:53 PM CST | 02/07/2018 05:15 AM CST |         |
| 60 | Davies2018   | FSD  | 6BM     | Micro-diffraction of Porous Building Materials     | 02/15/2018 05:06 PM CST | 02/27/2018 07:24 PM CST |         |
| 61 | Wilson2018   | XMD  | 1ID     | Nano-composite Materials Under Stress              | 03/01/2018 07:31 AM CST | 03/24/2018 08:22 AM CST |         |
| 62 | Johnson2018  | XRM  | ID3     | Rare Earth Ions Near Interfaces                    | 02/01/2018 10:31 AM CST | 02/14/2018 05:33 PM CST |         |



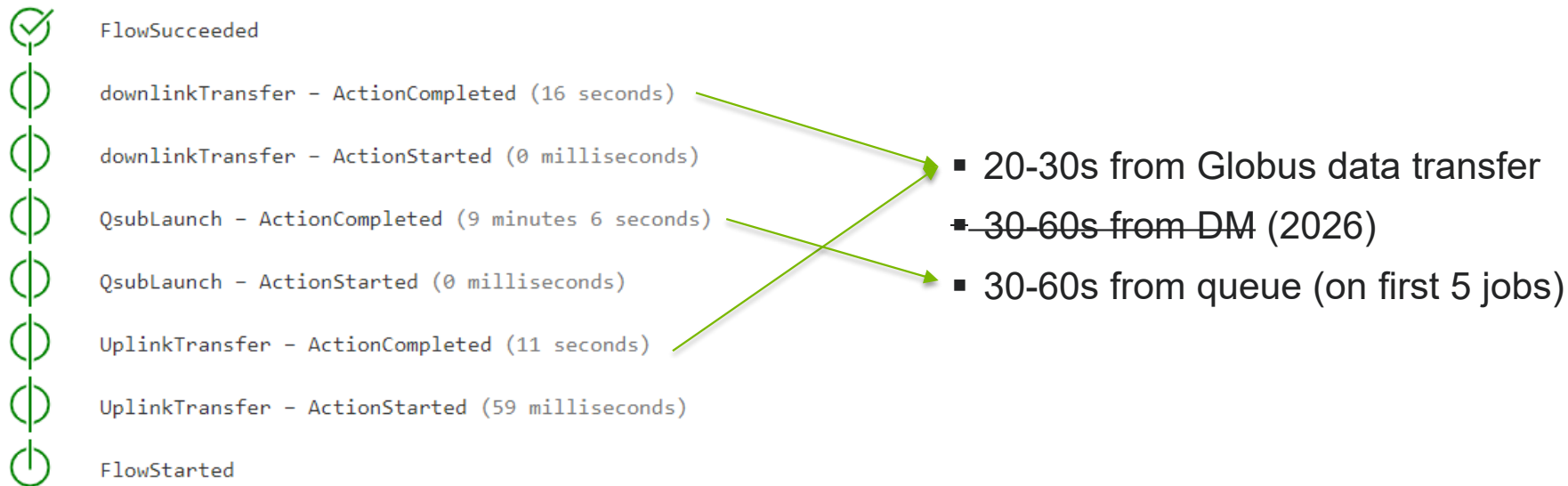
# DATA TRANSFER SYSTEM



# DATA TRANSFER SYSTEM 2026

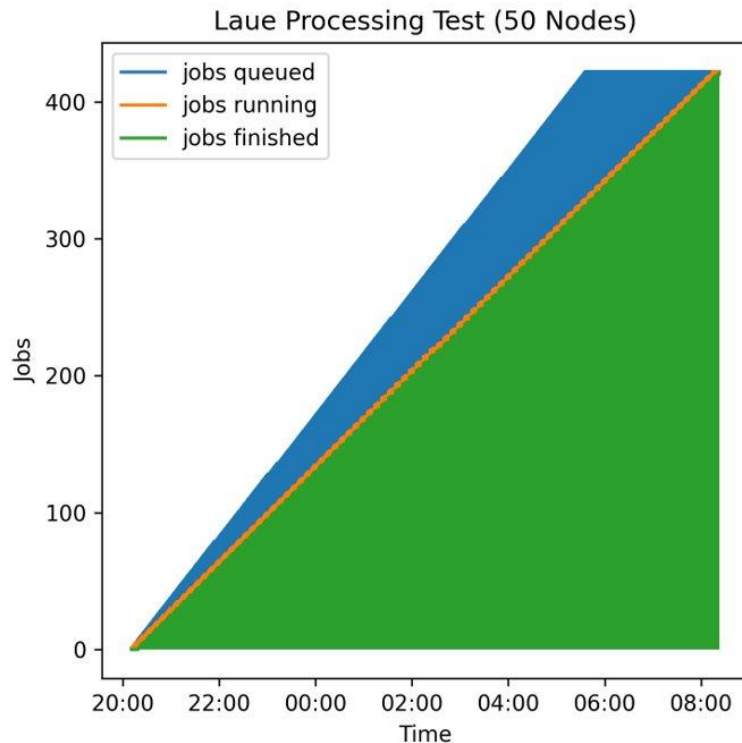


# DATA TRANSFER SYSTEM OVERHEAD

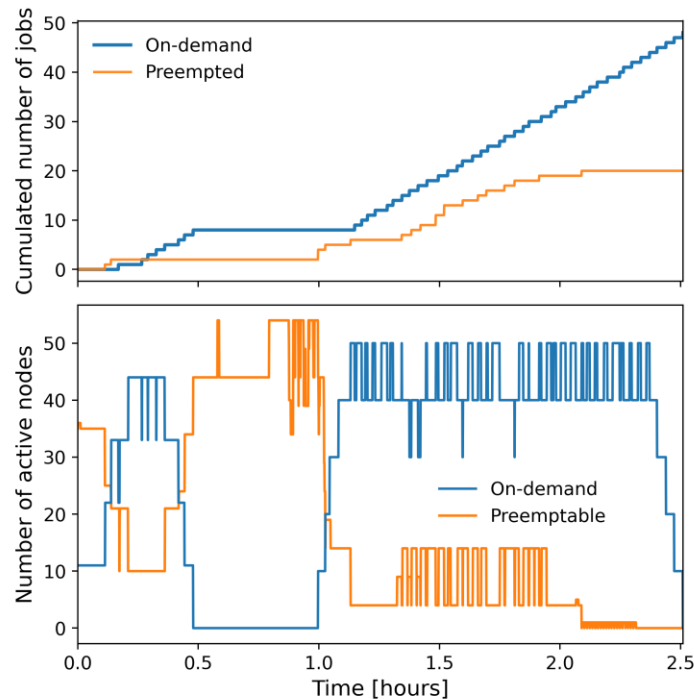
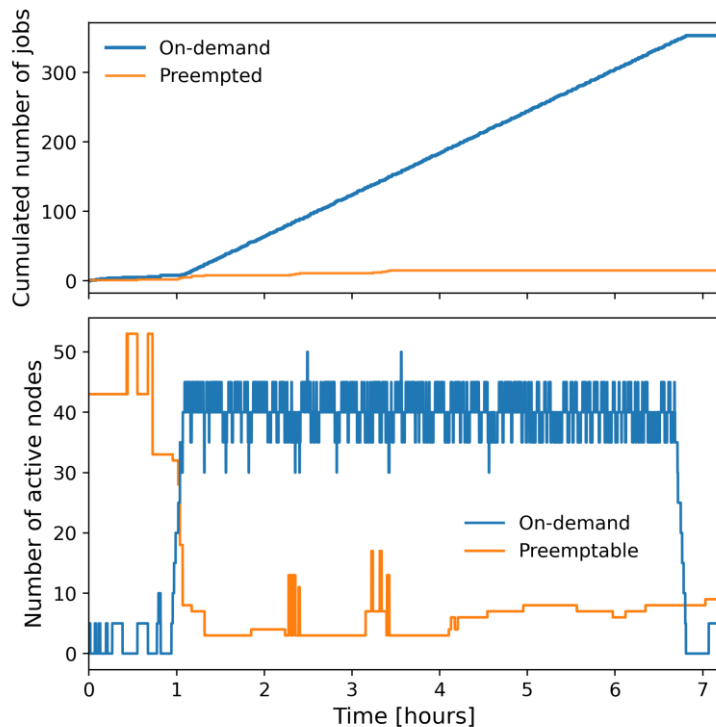


# FEB 23, 2023 FIRST FULL-SCALE RUN

- Data Rate
  - 1 point every 72 s
  - 428 points (9 hrs)
- Data Volume:
  - Input: 673 GB
  - Output:
    - Test 1: 2.6 TB
    - Test 2: 9.4 TB
- Success Rate:
  - 844/856 (98.5%)
  - (Failed file copies on eagle)



# PREEMPTION CHARACTERISTICS

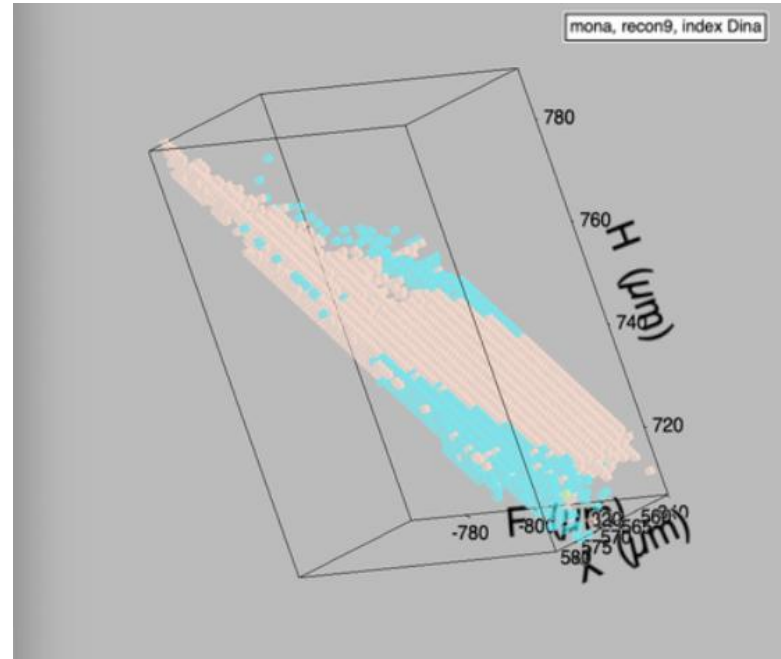
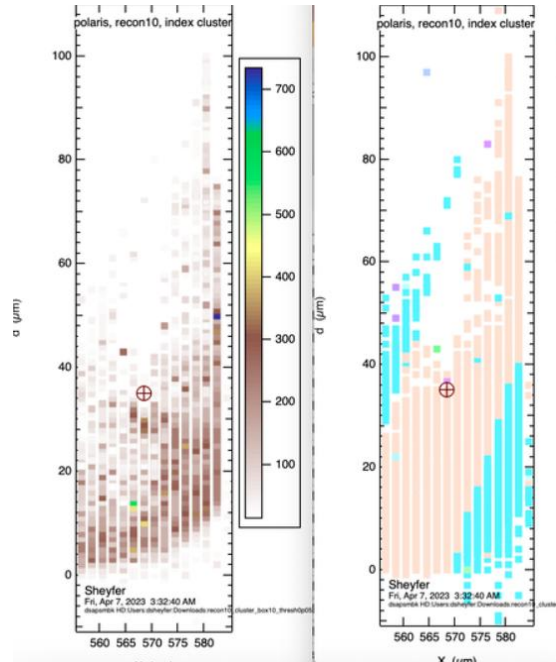
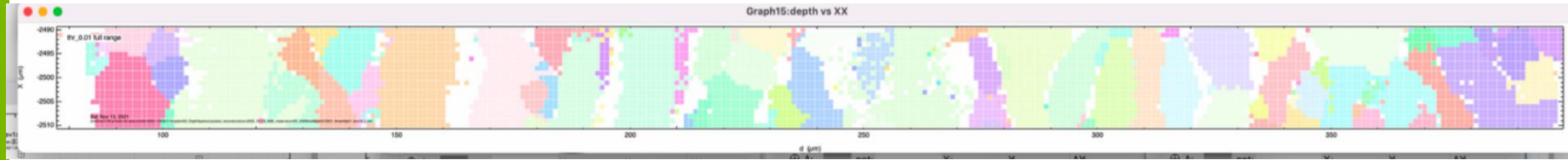


# ALL TEST RUNS

| <b>Dataset</b> | <b>Points</b> | <b>Nodes</b> | <b>Node H.</b> | <b>NH/Job (SD)</b> |
|----------------|---------------|--------------|----------------|--------------------|
| AL (2/9)       | 40            | 11           | 54.11          | 1.71 (0.22)        |
| AL (2/10)      | 40            | 10           | 60.80          | 1.52 (0.01)        |
| Ni2(1) (2/25)  | 422           | 10           | 598.68         | 1.42 (0.01)        |
| Ni2(1) (2/28)  | 422           | 10           | 600.79         | 1.42 (0.72)        |
| Ni2(2)* (4/1)  | 345           | 5            | 228.52         | 0.66 (0.04)        |

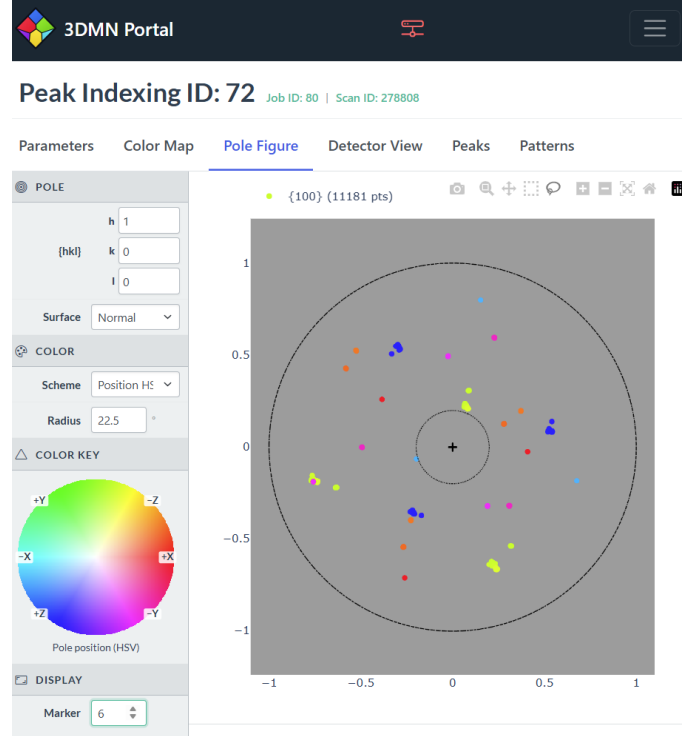
\*Processing conducted on a later software version

# EARLY SCIENTIFIC RESULTS



# LOOKING FORWARD

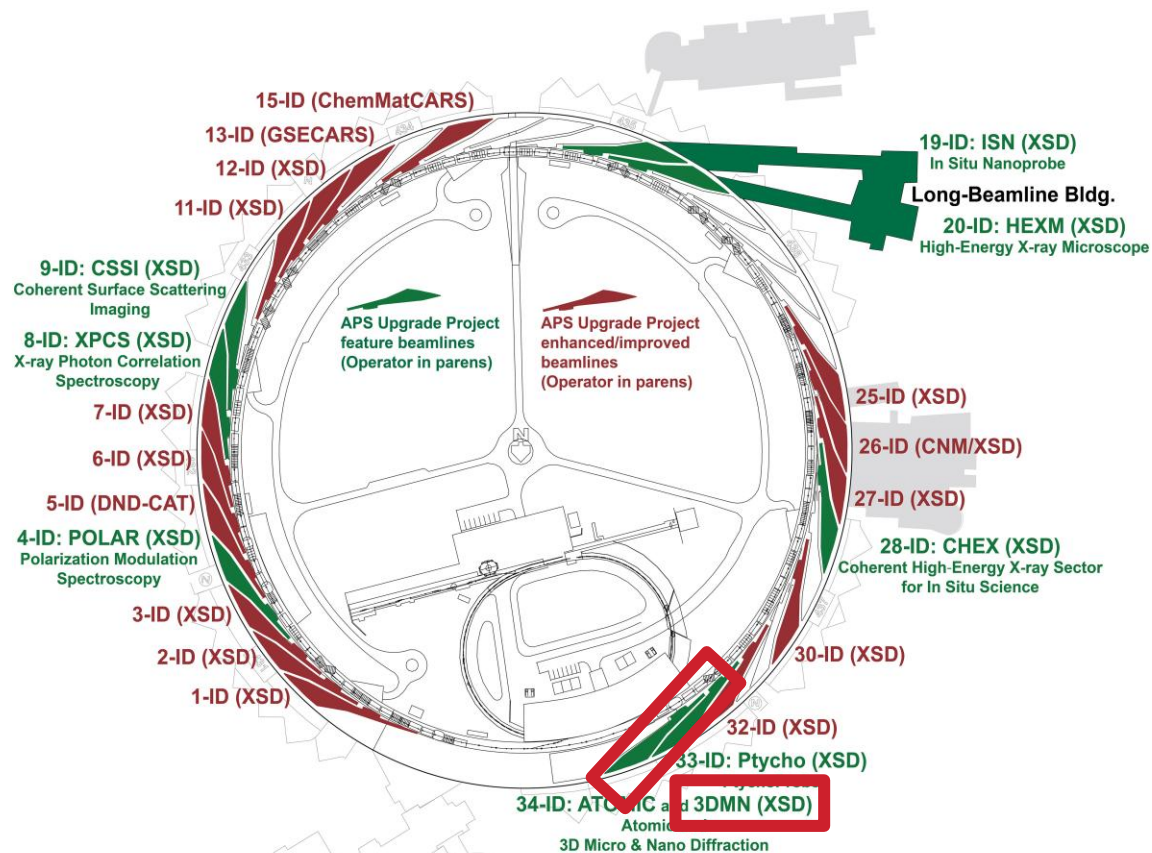
- Web Portals & Viz
- LLM Agents
  - User Support
  - Automated Experimentation
- 34-ID-E is planning to get a new detector:
  - 10x collection speed
  - 10x more processing
  - Excellent target for autoresearch/LLMs



# ZOOMING OUT

- 3DMN Uses a lot more software:
  - **Coded Aperture** <- Today's presentation
  - CA Calibration
  - Wire Reconstructions (HPC possible with Eiger)
  - Indexing
  - Laue Matching (Streaming/instant views for 2D only)
  - Assorted Analysis and Viz

# ZOOMING OUT... AGAIN



# CONCLUSION & TAKEAWAYS

- Multifaceted, full stack problem: UI/UX -> data management + transfer -> HPC -> GPU kernels.
- More beamlines are likely to follow, and we need to learn to work in parallel
- As detector technology improves, this problem compounds

# ACKNOWLEDGEMENT

This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science user facility at Argonne National Laboratory and is based on research supported by the U.S. DOE Office of Science-Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. This research used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02 06CH11357.

# ACKNOWLEDGEMENT

The U.S. Government retains for itself, and others acting on its behalf, a paid-up nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan. <http://energy.gov/downloads/doe-public-access-plan>.

# QUESTIONS?



Argonne National Laboratory is a  
U.S. Department of Energy laboratory  
managed by UChicago Argonne, LLC.

