APS Scientific Computing Strategy

2024

Table of Contents

1 Overview

1.1 Introduction

The Advanced Photon Source (APS), located at Argonne National Laboratory, is a synchrotron light source funded by the U.S Department of Energy (DOE), Office of Science-Basic Energy Sciences (BES) to produce high-energy, high-brightness x-ray beams. The APS has been operating since 1996 and has become of the largest scattering user facility in the world, averaging 5,500 unique users and producing more than 2,000 scientific publications every year. Scientists from every state in the US and international users utilize the 68 beamlines at the APS to conduct cutting-edge basic and applied research in the fields of materials science; biological and life science; physics; chemistry; environmental, geophysical, and planetary science; and innovative x-ray instrumentation.

To maintain the APS at the forefront of X-ray science, BES and Argonne launched the APS-U project, a $815M investment that will deliver a comprehensive upgrade of the electron storage ring and X-ray beamlines. After more than a decade of planning and preparation, the storage ring has been replaced by a new Hybrid Multi-Bend Achromat during the period April 2023 - April 2024. After 3 months of machine commissioning, the new storage ring has delivered a current of 125mA and a world leading electron emittance of 41pm.rad, almost 2 orders of magnitude smaller than the previous machine. Return to user operation is underway, and the next few months will see a progressive commissioning of the 9 new and 15 enhanced beamlines offering transformative analytical capabilities for X-ray scattering spectroscopy, and imaging. The transformation of the APS will be expanded further in the next 5 years by upgrading an additional 8 beamlines and investing in additional computing infrastructure, through the project EXCEL@APS (Extend X-ray Capabilities with Extreme Light).

More than ever before in the history of the APS, advanced computational approaches and technologies are essential in fully unlocking the scientific potential of the new facility. The upgraded source opens the door for new measurement techniques and increase in throughput, which, coupled to technological advances in detectors, new multi-modal data, and advances in data analysis algorithms, including artificial intelligence and machine learning (AI/ML), will open a new era of synchrotron light source enabled research. In particular, the high-brightness, and increase in coherent x-ray flux at the new APS is leading to significant increases in data rates and experiment complexity that can only be addressed with advanced computing.

The overarching goal of the APS is to harness the power of computing, and the value contained in data to elevate the overall scientific impact and boost the operational efficiencies of the APS beamlines and accelerator complex. This will enable the APS and its users to help unlock new scientific opportunities and accelerate science through innovative data science, including AI/ML, scientific software, data management, and computing technologies. In order achieve this, the APS must remain at the cutting-edge of computational and computer science-, data-, and AI-driven discovery.

The APS has developed a comprehensive strategy in order to realize this goal, the APS Scientific Computing Strategy. This strategy identifies five high-level strategic goals, along with high-level approaches to achieve these goals, for computing at the APS (see Table 1-1). The strategy then identifies activities and plans in the areas of AI/ML, data reduction and analysis software, computing infrastructure, data management, workflows, and science portals, controls, data acquisition, and detector integration, network architecture and infrastructure, and collaborations required over the coming years.

Given the prominence of computing and data science presently, and in the foreseeable future, the APS is establishing a Computing Advisory Committee. This committee will provide specialized advice, expertise, and guidance to the Associate Laboratory Director for Photon Sciences (ALD-PSC) / APS Director on strategic development, review of practices and organization matters, and on the development of future policies where applicable. The Committee's advice and guidance will help ensure that computational methodologies and technologies are leveraged by, and integrated into, the research and operations activities of the APS.

Table 1-1 High-level APS strategic computing goals and approaches.

Computing Strategic Goals	Computing Approaches / Strategy
1. **Enable High-Speed Discovery** - *Harness the unprecedented power of computing and data science expertise at Argonne to reduce the time to science.*	1. Deliver state-of-the-art network performance, and security and resiliency (Section1.2). 2. Utilize Argonne Leadership Computing Facility (ALCF), and other supercomputing facilities, including the future High-Performance Data Facility (HPDF), for large-scale computing needs (Section 1.5, Section 1.8). 3. Deploy HPC-enabled software and workflows using Globus for near real-time data processing (Section 1.6, Section 1.4). 4. Apply AI/ML inference to accelerate data processing (Section 1.7).
2. **Unlock Exceptionally Challenging Experiments** - *Apply the confluence of Argonne's strengths in AI/ML and autonomous experiments, large-scale simulations, and x-ray experiments to unlock otherwise unreachable complex scientific solutions.*	1. Deploy advanced controls infrastructure and tools in order to realize experiment feedback, for example, for rare event detection (Section 1.3). 2. Combine large-scale simulations with experiment data during beamtime to drive experiments (Section 1.3, Section 1.5). 3. Utilize edge computing for fast feedback at beamlines, and develop and apply algorithms for advanced computational techniques (Section 1.5, Section 1.6). 4. Leverage AI/ML and digital twin capabilities to design and steer autonomous experiments (Section 1.7).
3. **Leverage Data for New Science** - *Take advantage of the opportunities created by the explosion of data at the APS coupled with AI to advance new science that is only achievable using combined knowledge beyond individual experiments.*	1. Leverage opportunities from the DOE Frontiers in Artificial Intelligence for Science, Security, and Technology (FASST) initiative (Section 1.7, Section 1.8). 2. Utilize a facility-wide data management and workflow system and the Globus ecosystem to collect and organize data consistently (Section 1.4). 3. Develop plan for FAIR data, open data, and metadata (Section 1.4). 4. Apply persistent identifiers to APS output (e.g. awards and data) (Section 1.4).
4. **Empower Users to Realize the Full Potential of the APS** - *Provide APS users with the cutting-edge tools needed to produce world leading science.*	1. Deploy advanced beamline controls and experiment control systems to support real-time feedback and AI applications (Section 1.3). 2. Deliver data management, workflows, and portals for facile management and manipulation of data generated at the APS (Section 1.2, Section 1.4). 3. Focus on data processing software aligned to beamline and user needs, including multi-modal data utilization (Section 1.6). 4. Place AI/ML advances into operational use at instruments (Section 1.7).
5. **Expand Collaborations** - *Leverage partnerships and initiatives to amplify the impact of Argonne and APS resources.*	1. Plan for an integrated facility approach to leveraging computing and data resources as part of the DOE SC ASCR Integrated Research Infrastructure (IRI) program and HPDF project (Section 1.5, Section 1.8). 2. Develop an engagement roadmap to integrate light source-enabled research into the FASST initiative (Section 1.7, Section 1.8). 3. Leverage the 6-way collaboration among the DOE BES light and neutron sources to develop shared solution and approaches to common data and computing challenges (Section 1.8).

The APS has a decentralized computing organization with functional groups located in two divisions, the X-ray Science Division (XSD) and the APS Engineering Support (AES) division. The XSD Beamline Controls (BC) group is responsible for beamline data acquisition, through control and operations systems and software. The XSD Computational X-ray Science (CXS) group is mainly responsible for the development of theory, mathematical models, algorithms, and software for interpreting x-ray measurements. The XSD Scientific Software Engineering & Data Management (SDM) group is responsible for software engineering for data analysis applications and data management tools, enabling high-performance computing (HPC). The management and support of information technology resources are in the AES division Information Technology (IT) and Information Solutions (IS) groups.

The outline of the rest of the document is as follows. The remaining sections provide a breakdown of strategies and tactics proposed in each functional areas in support of the 5 overarching strategic goals. This includes plans for networking architecture & infrastructure (see Section 1.2), controls, data acquisition, and detector integration (see Section 1.3); data management, workflows, and science portals (see Section 1.4); computing infrastructure (see Section 1.5); data reduction and analysis (see Section 1.6); AI/ML (see Section 1.7); and effort funding and collaborations (see Section 1.8). Specific needs and plans for the APS-U feature beamlines are documented in Section 2.

In addition to this strategy, each support group maintains its own detailed documents and plans describing goals for the current and next fiscal year:

1.2 Network Architecture and Infrastructure

A state-of-the-art high-performance network that is secure and resilient is a key underlying capability required to meet the facility's strategic computing goals. Such a network is required to *enable high-speed discovery* (Goal 1) so the APS and its users can collect data at the high-rates now possible with the new source and cutting-edge detectors, quickly transfer data to large-scale computing resources, both on campus and geographically distributed, and retrieve results quickly. This network is also required to *allow users to effectively access, manage, and manipulate data and experiments* (Goal 4) from outside of the APS.

Current State

As data rates and volumes continue to grow (see Section 1.4), greater demands will be placed on the APS network. This is especially true for the new and upgraded beamlines part of the APS-U project, and future upgrades a part of EXCEL@APS and subsequent projects. The APS has had a long-term commitment to continuously update and enhance its network architecture and infrastructure to adequately serve beamline and user needs, and to follow cybersecurity requirements required by DOE. The APS Network Team has been working over the past years to develop a network architecture and infrastructure plan and to implement that plan to support the future needs of the facility. Figure 1-1 depicts the new APS network architecture and infrastructure plan.

The center of the beamline APS network consists of a pair of core switches (HPE Aruba 6410) located in the APS data center. These Tier 1 switches provide all routing to beamline subnets and to other parts of the APS, Argonne, and the Internet via ESnet (https://www.es.net/), the interlaboratory network and internet service provider for the DOE laboratories. The core switches provide multiple 10/40/100 Gbps line rate ports. These core switches are connected via 4 x 40 Gbps uplinks to the APS Tier 2 firewall, which in turn connects to the Argonne Tier 1 firewall with 4 x 100 Gbps uplinks. These Tier 1 core switches have been augmented with distribution switches (HPE Aruba 9300) that provide 32 x 100/200/400 Gbps ports for beamline high speed uplink and storage connectivity. The upgraded Tier 1 core switches were installed in the spring of 2021. The additional distribution core switches were installed in 2023 and 2024. The Tier 1 Argonne firewall connects to the Internet via ESnet using 2 x 400 Gbps uplinks. The same core switches connect to the storage systems for the APS Data Management System (see 1.4), sector data storage systems, the dservs that host beamline control system configurations and software, and the APS accelerator network.

Each sector at the APS has Tier 2 network switch(es) that connect beamline devices and connect the beamline to the core Tier 1 switches. The Tier 2 switch(es) connect to beamline computers, control system EPICS IOCs, detectors and data acquisition servers, wireless access points, cameras, and controls hardware. Each Tier 2 beamline network switch will provide line rate 10/100/1000 Mbps ports for many of devices at the beamline, as well as high speed line rate 10/40/100 Gbps ports for data acquisition as needed.

In this plan, a Tier 3 managed switch with 48 x 10/100/1000 Mbps ports is deployed at each experiment hutch for controls hardware stations to provide additional ports, a dynamic cabling environment, and to isolate beamline controls hardware traffic.

An additional 96 pairs of single mode fiber have been installed from APS data center to each of the Laboratory Office Module (LOM) network closets; 768 pairs in total. This additional fiber infrastructure will provide sufficient network bandwidth from the beamlines to the data center for the next decade.

A science DMZ has been established to directly connect the APS beamline network core to the ALCF. Currently four pair of single mode fibers connect the beamline core to the ALCF network core at 4 x 100Gbps. Static routing has been configured at both facilities to provide this direct path without passing through any security devices such as firewalls or IDS systems that would inspect this traffic and present network delays.

Tactics related to the APS network architecture and infrastructure required to realize high-level goals:

1. *Complete implementation of network switch and cabling infrastructure for new beamlines to support performance requirements.*
2. *Transition APS beamlines to Supervisory Control and Data Acquisition (SCADA) architecture for improved security and resiliency.*
3. *Implement a terabit per second (Tbps) network between the APS and the ALCF to support performance requirements and utilization of ALCF for high-speed data processing.*
4. *Create plans to upgrade additional beamline network infrastructure, as required by EXCEL@APS.*

Planning has been completed for 7 of the 9 APS-U feature beamline networks. Installation of these networks in underway. The APS aims to complete planning for the remaining beamlines by the end of this calendar year.

The APS is adopting a SCADA architecture for the beamline control system network. Controls and data analysis network traffic will be separated and isolated from outside networks for maximum performance and security. The APS is converting beamlines to this model in a phased plan that is currently underway. The APS plans to complete this transition by the end of calendar year 2026.

The APS plans to upgrade the network to the ALCF to a 1.6 Tbps network consisting of 4 x 400Gbps uplinks. This network will support the performance required to utilize the ALCF for real-time and on-demand workloads. This will be completed by the end of calendar year 2025.

The APS will follow this same infrastructure and architecture plan to provide a consistent and state-of-the-art network as additional beamlines are upgraded. Note that EXCEL@APS is currently in the early stage (mission need approval). Detailed planning will be created post CD-1 approval.

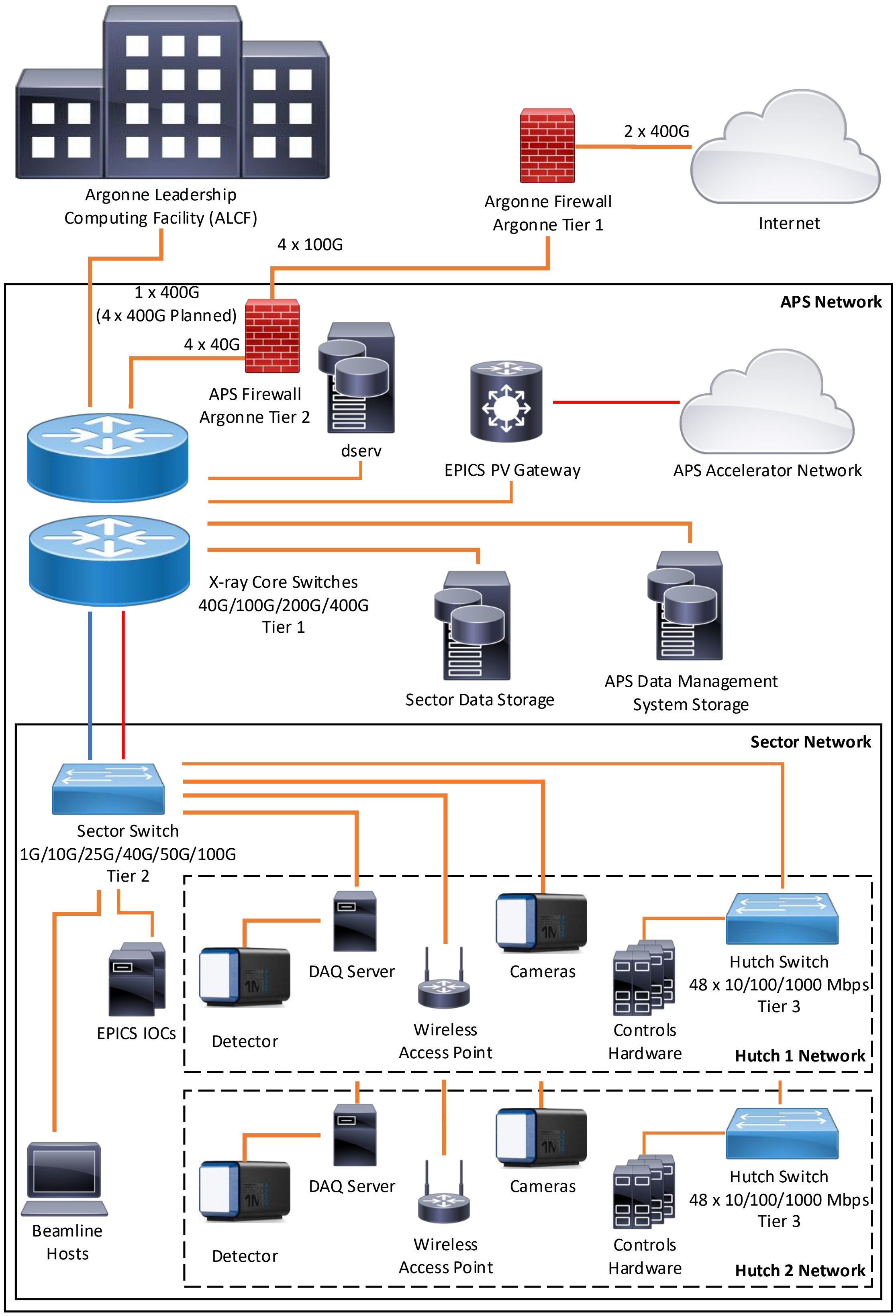


Figure 1-1 APS network architecture and infrastructure.

1.3 Controls, Data Acquisition, and Detector Integration

The deployment of advanced controls infrastructure and tools is key to enabling real-time experiment feedback required to *unlock exceptionally challenging experiments* (Goal 2). This is particularly valuable for rare event detection, where timely adjustments can significantly enhance the outcome of experiments. Additionally, advanced control systems are required to enable the integration of large-scale simulations with experimental data during beamtime to enable data-driven decision-making, optimizing experimental conditions in real time. Advanced beamline controls and experiment control systems will enhance the user experience better *empowering users to realize the full potential of the APS* (Goal 4).

Current State

Before APS-U, XSD operated a small number of scanning microscopes, the majority of which provided only a single data acquisition channel (e.g., fluorescence data acquisition). With APS-U, many more rapidly scanning microscopy instruments will be deployed, all requiring the rapid acquisition of positioner readback signals, rapid triggering and coordination of multiple sophisticated detectors, and high-bandwidth feedback to positioners. Figure 1-2 illustrates the complexity of implementing one scanning microscope system for APS-U.

Also, before APS-U, XSD beamline controls relied heavily on VME-based infrastructure for experiment control. Each XSD-operated beamline typically has several VME crates, each with an input-output-controller (IOC) running a proprietary real-time operating system (VxWorks) and housing several VME cards that provide a variety of basic beamline functions such as motor step and direction signals as well as lower-bandwidth beamline input-output (IO) such as analog, digital, relays, and thermocouple readback. VME has proven reliable over the decades, which is why there is still a large installed base in XSD. Nevertheless, various components of our VME installations are at end-of-life, and VME itself is gradually ebbing away. In addition, the pricing and licensing model for VxWorks has worsened in recent years, and there is uncertainty (and, therefore, risk) about how this will evolve.

These considerations and others instigated a re-examination of the portfolio of supported hardware, software, and firmware to meet the needs of the APS-U instruments and the APS's strategic computing goals. Proposed solutions were vetted with the beamline community at, for example, the 2019 APS Advanced Controls Workshop, and regular updates were made in various other forums for XSD. Briefly, the solutions adopted and vetted with instrument stakeholders are the following:

a) ACSMotionControl motor controllers and drives are the new standard APS-U motion solution. This motion system uses EtherCAT technology to coordinate motors across multiple devices and includes built-in machine safety functions.
b) A network-based motion system (OMS MXA) is available when high-density/low-duty cycle motors need support.
c) Networked industrial IO devices from LabJack and Measurement Computing are to be used for lower-bandwidth beamline input-output (IO) needs (analog, digital, relays, thermocouples) readback.
d) The BC-designed softGlueZynq system provides FPGA performance for fast triggering/timing, and streaming data acquisition capabilities needed for APS-U instruments. softGlueZynq allows beamline users to construct FPGA circuits, and layout interconnections using EPICS process variables (PVs).
e) A commercial FPGA DAQ appliance (ACQ2106) from D-TACQ Solutions Ltd. will deploy softGlueZynq solutions customized for each APS-U instrument's requirements. The appliance provides FMC slots for FPGA I/O cards, which will be selected to match each instrument's needs.

The beamline controls group (BC) has used the lead-up to APS-U to develop support for this new infrastructure and test it in various environments. In addition to meeting the APS's strategic computing goals, all solutions deployed to the beamlines must also provide maintainability and adaptability, which are hallmarks of Beamline Controls support within XSD.

Tactics related to APS beamline controls required to realize high-level goals:

1. *Transition to a Fieldbus-based approach for device control.*
2. *Transition to EPICS v7 in order to better facilitate data streaming and data integration at APS beamlines.*
3. *Deploy Bluesky widely at the APS as an experiment control system.*

APS-U has jump-started work on deploying Tactics 1, 2, and 3. For instance, the 8-ID beamline (XPCS beamline) has been built and begun successful commissioning, employing Tactics 1, 2, and 3. Despite the jump-start provided by APS-U, XSD will continue to have a large installed (and slowly withering) base of VME. We suggest an XSD or PSC project to accelerate the transition to Fieldbus-based infrastructure. Another VME-related gap is support for high-fidelity beamline trigger signals synchronized with the storage ring (e.g., for laser pump and x-ray probe experiments). The VME hardware and software currently used to support this is EOL and no longer supported except in a maintenance mode. We propose an ASD-led project with XSD to develop new FPGA-based tools for this capability.

Tactic 2 is to further deploy EPICS v7 IOCs to the beamlines. EPICS is a robust, large-scale distributed control system. Its widespread and longstanding use at similar facilities worldwide enables collaboration with instrument scientists and controls communities within the APS, other DOE facilities, and abroad. A key motivator for the switch to EPICS v7 is the pvAccess protocol, which enables the manipulation and transport of structured data over the network. pvAccess might be considered an extremely high-performance data broker so it will form the bridge between instruments and the high-performance data stores and data reduction tools.

One example of the power and utility of pvAccess, is the need to support the imminent deployment of new high data-rate detectors that can generate thousands of frames per second using tens of Gbps of bandwidth. The newly developed C++ and Python-based pvaPy framework (https://github.com/epics-base/pvaPy/blob/master/documentation/streamingFramework.md) provides a streaming data framework for EPICS v7 that can combine multiple data sources and process data at thousands of frames per second (see Figure 1-3). The pvaPy framework will be the basis for fast data handling for XSD beamlines with first deployments at APS-U feature beamlines.

As an example of the planned deployment of pvaPy, we consider the APS-U Era Beamline Data Pipeline (BDP) Project. The BDP is a cross-XSD effort to identify, demonstrate, and deploy a portable data pipeline solution that addresses APS-U Era beamline data needs. This project will create a template for APS beamlines and support groups to follow when deploying new instruments or data pipelines addressing detector integration, data movement, network infrastructure, storage systems, multi-tiered computing, and cyber security. The template has been validated in a laboratory setting and is in the midst of its first deployment to APS-U Feature Beamlines and APS-U Enhancement Projects.

Tactic 3 is to further deploy the Bluesky controls framework to XSD beamlines. Bluesky is an organized set of Python-based libraries that enable experimental science at, e.g., beamlines within XSD. Based on Python, it provides structured access to the entire scientific Python ecosystem and enables interdisciplinary teams due to the common language. Figure 1-4 illustrates this concept. Python has also become the de facto tool for AI/ML applications using libraries like PyTorch and TensorFlow. Bluesky also handles multi-dimensional and asynchronous event-based sources and facilitates the creation of a common metadata catalog. For example, through the BDP, we have demonstrated integration into automated workflows that leverage the APS Data Management system, capture metadata, and initiate data reduction processes on Argonne's leadership computing facilities, such as Polaris.

One particularly exciting possibility and driver for Bluesky deployment is the first-class support for adaptivity and informed decision-making in experimental workflows that Bluesky enables. Bluesky provides a flexible API called Bluesky Adaptive that supports various adaptive algorithms, from simple rule-based adjustments to artificial intelligence and machine learning models. Three examples nearing first deployments are automated beamline alignment via integration of a digital twin, optimal scanning of XANES spectra facilitated via Bayesian optimization, and automatic monitoring and gain changes on pre-amplifiers to allow automated data acquisition. In the first example, illustrated in Figure 1-5, our Optics group has created comprehensive digital twins [using Oasys (https://github.com/oasys-kit)] of selected APS beamlines to test automated alignment and optimize desired beam properties. Using in-line wavefront sensors developed by this same group and deployed on request at XSD beamlines, the system will be capable of objectives that are far more complicated than, e.g., minimum focal spot size. A particular wavefront at the sample position is an example of a more complicated objective that could be achieved assuming appropriate beamline optics.

A second example nearing deployment is optimized XANES energy scans via Bayesian optimization that again leverages Bluesky Adaptive. In this example, a coarse energy scan is first performed. Then successive cycles of Bayesian optimization tell Bluesky, i.e., the beamline where to measure next to complete the scan with appropriate resolution and quality. The interconnection between Bluesky, i.e., beamline 25-ID, and the scan optimization algorithm is illustrated in Figure 1-6. In the third example, Bluesky and Bluesky Adaptive will monitor and automatically set the gain for pre-amplifiers used to collect spectroscopy data at beamline 25-ID. The capabilities enabled by examples 2 and 3, and, to a degree, example 1, means that 25-ID can collect spectroscopy data automatically for extended periods (number of samples) and of high quality. We eagerly await further deployments like this for, e.g., adaptive scanning in microscopy, as areas of interest are identified and fed back to the control system for more comprehensive scanning.

Bluesky deployment has been slower than anticipated for various technical and cultural reasons. We have under-estimated the staff development needs to transition effectively from procedural programming to historically used at the APS in spec to an object-oriented approach best suited to Bluesky. Focused training and upskilling using experienced Python trainers will be put in place to boost the Python literacy of beamline staff, and in strategic areas in XSD. We will conduct frequent assessments to determine the effectiveness of this campaign. In addition, the amount of hands-on support we can provide to beamlines interested in switching is limited for staffing reasons. There are several approaches we are proposing to address this challenge. In the early phases of deployment, we will prioritize and centralize our support to a limited number of beamlines to reach critical mass targeting a small set of objectives. These beamlines are chosen based on the need for innovative features offered or easily developed in Bluesky and, when possible, in areas where we can maximize staff participation. In the medium term, hiring additional dedicated support staff will be necessary to ensure this effort grows and succeeds. Hires in this area are in line with the APS's objective to grow in the area of data science and AI/ML, as part of the 5-year staffing plan. Lastly, we recognize that change management is not straightforward. We plan to actively communicate the benefits of this technology shift by regularly featuring, through relevant forums, the innovative capabilities of Bluesky, such as Bluesky Adaptive that cannot be accomplished via spec, for instance, and engage with stakeholders in support of this initiative

As it stands, the APS plans to rollout Bluesky to a select number of beamlines (1-ID, 2-ID/19-ID via queueserver and GUI), 4-ID, 8-ID, and 25-ID and USAXS (12-ID) over the next two years with a re-evaluation to follow.

In parallel to these efforts, Bluesky-related information caches such as its GitHub repository, online documentation, and Jupyter notebooks have enabled training efforts on and early exploration of a large-language model (LLM) called CALMS that will facilitate the use of Bluesky. The concept is illustrated in Figure 1-7. CALMS should make it relatively straightforward for inexperienced users to create robust Bluesky plans and perform tasks like moving a diffractometer to a selected peak position in reciprocal space based on the state of a beamline and a sample's material properties retrieved from the Materials Database.

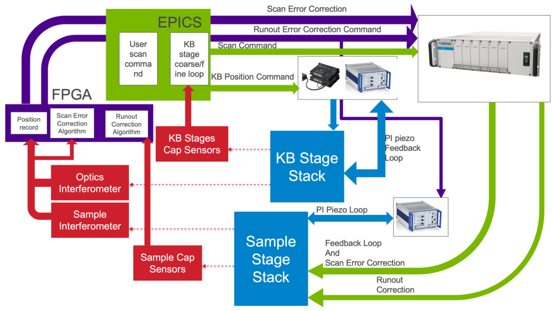


Figure 1-2 Schematic of fast scanning and interferometry for the APS-U InSitu Nanoprobe (ISN) Feature beamline.

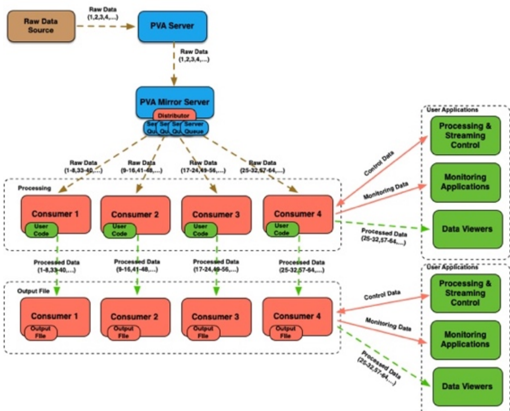


Figure 1-3 pvaPy streaming framework depicting multiple data processing consumers.

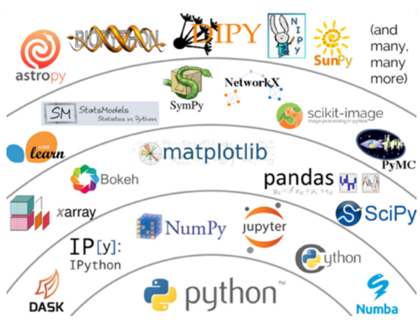


Figure 1-4 A portion of the scientific Python ecosystem that Bluesky can leverage. Credit: Jake vanderPlas, "The Unexpected Effectiveness of Python in Science", PyCon 2017.

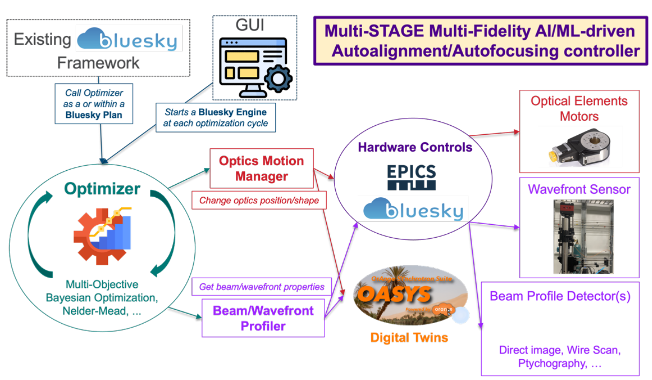


*Figure 1-5 Schematic of Oasys digital twin integrated with a beamline via Bluesky and its Bluesky Adaptive framework. First tests are imminent at Sectors 8-ID and 25-ID. See L. Rebuffi, S. Kandel et al., Opt. Express **31**, 39514 (2023). DOI: 10.1364/OE.505289.*

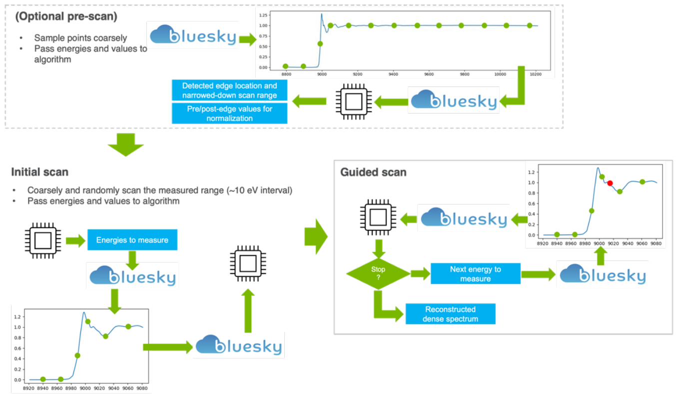


Figure 1-6 Optimized XANES scan enabled via Bayesian optimization and Bluesky Adaptive. Contact: Ming Du.

Figure 1-7 CALMS: an LLM-enabled scientific companion for, among many other things, facilitating the use of Bluesky at beamlines. Contacts: Eric Codrea and Mathew Cherukara. Figure courtesy of E. Codrea and F. Rodolokis.

1.4 Data Management, Workflows, and Science Portals

Data management, efficient workflows, and advanced science portals are essential for *enabling high-speed discovery* (Goal 1) at the APS. By deploying HPC-enabled software and utilizing the Globus Compute platform, near real-time data processing using large-scale computing systems, including leadership class systems becomes possible, accelerating research outcomes. The ability to *leverage massive data for new science* (Goal 3) hinges on effective data management. Utilizing a facility-wide data management and workflow system, researchers can consistently collect and organize data and metadata. The creation of a plan for adherence to FAIR data principles and open data is one of the prerequisites to the reuse of data for new scientific exploration beyond the intention of the original experiment. Through these efforts, *users will be empowered to fully tap the potential of the APS* (Goal 4), with seamless data management, workflows, and portals facilitating data analysis and innovation.

Current State

The need for data management, workflow, and distribution tools, and data storage resources continues to grow. Prior to the replacement of the storage ring and the new and upgraded instruments as part of the APS-U project, the APS X-ray Science Division beamlines collected on the order of 5 PB of raw data per year. Over the next decade, it is estimated that the data storage needs of the APS are anticipated to increase by at least two orders of magnitude to 100s of PBs of raw data per year (see Figure 1-8), reaching a peak when the facility and new and upgraded instruments are fully commissioned and the User community is back in full force.

The APS Data Management System consists of software and hardware tools that automate the transfer of data between acquisition devices, computing resources, and data storage systems. Ownership and access permissions are granted to the users signed-up to perform a particular experiment. A metadata catalog allows beamline staff to populate experiment conditions and information for access via a web portal. Users can download data at their home institutions or transfer data to computing centers using Globus Transfer (globus.org) or SFTP. At present, approximately 60 APS beamlines (XSD and non-XSD), i.e. the vast majority of beamlines, take advantage of this system.

Medium-term data storage is available within the APS; longer-term storage systems are provided by the ALCF (see Figure 1-9). Currently, the APS provides approximately 10 PB of central disk storage for medium-term data retention, and several Data Transfer Nodes (DTNs) for reliable, high-speed data movement internally and externally. The ALCF currently provides approximately 10 PB of tape storage (easily expandable to meet future APS needs) for longer-term data retention. The ALCF has recently deployed a 100 PB community file system (Eagle) and a 100 PB project file system (Grand) along with additional tape storage that is available for APS use.

The APS continues working with Argonne's Data Science and Learning Division (DSL) and the Globus team to develop a computational data fabric for end-to-end data lifecycle management. This fabric, Globus Compute, connects and automates many stages of the data lifecycle from acquisition to processing to publication. Science web portals will allow APS users to view and download their data and reprocess their data on ALCF and other large-scale computing resources using Globus tools. The Materials Data Facility (MDF) and the DOE Office of Scientific and Technical Information will serve as a DOI generating service for APS datasets.

The APS and Globus team have prototyped this computational fabric with the ALCF for many APS techniques (see Figure 1-10), including XPCS, ptychography (see Figure 1-11), High-Energy Diffraction Microscopy (HEDM), Laue microdiffraction, serial crystallography, and Bragg Coherent Diffraction Imaging (BCDI).

Tactics related to data management, workflows, and science portals required to realize high-level goals:

1. *Deploy the next-generation storage system for operations use by beamlines.*
2. *Harden existing and develop new data pipelines as beamlines return to operation.*
3. *Develop a plan for the use of metadata, FAIR data, DOIs and persistent identifiers, and open data to expand and accelerate the generation of scientific knowledge and impact of the facility.*

The APS is in the procurement process for a new high-performance data buffer for APS beamlines. The system will be installed in phases over four years, resulting in a system with approximately 50 PB of storage and hundreds of GB/s of aggregate throughput available with GPFS. The system is a hybrid NVMe SSD/spinning disk system, where NVMe SSDs provide a high-speed landing space for data and spinning disks provide storage capacity. Due to the need for high-speed storage, the APS has accelerated its plans for implementing this system. The first two phases of this storage will be in operations use during the first quarter of calendar year 2025 and provide approximately 24 PB of usable storage with 300 GB/s of aggregate throughput. The storage capacity and throughput will increase as the final two phases are installed by the end of calendar year 2027. As more beamlines come online, the APS will reevaluate its needs for performance (potentially as high as 1 TB/s) vs. capacity, and can modify system specifications for the final two phases to accommodate. This system will enable the direct connection of detectors and analysis computers which will better enable near real-time views into data, data analysis, and experiment feedback and autonomous control, as well as reduce the turnaround time to training AI/ML models and quickly incorporating newly acquired data into models. The APS plans to stabilize its investment in storage annually so it keeps up to date with emerging technologies, needs, and sponsor requirements.

As more beamlines return to operation, the APS plans to refine existing workflows and deploy new workflows using the APS Data Management System and the Globus Compute infrastructure.

In order to fully take advantage of the opportunities presented by the massive volumes of data generated at the APS, the APS will develop a plan for activities to collect and better use metadata and follow FAIR principles. The APS now creates DOIs for beamtime awards, and work is underway to create DOIs for beamlines and instruments. The APS will continue to explore DOIs for datasets, both to keep up with sponsor requirements for persistent identifiers, and to best leverage DOIs to advance scientific discovery through the reuse of data. These tools will help to serve as the basis for enabling searchable data catalogs and adopting FAIR, and possibly open, data practices. To help chart a path for these activities, the APS will take direction from the upcoming series of DOE SC Roundtables scheduled for the fall of 2024, and the APS will work closely with the 6-way light and neutron source collaboration (see 1.8 for more details).

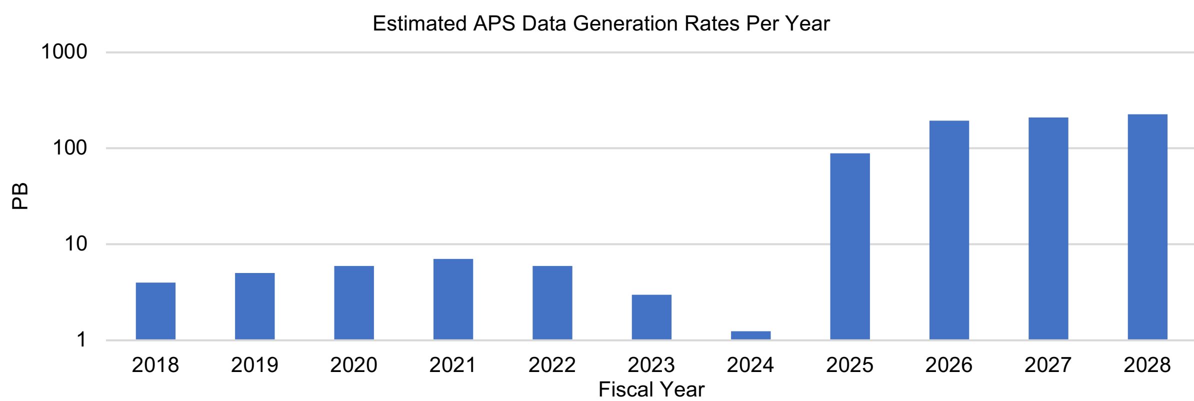


Figure 1-8 Anticipated aggregate APS X-ray Science Division data generation per year. Data generation during FY23 and FY24 is estimated to be lower due to the storage ring replacement period followed by the storage ring and beamline commissioning periods.

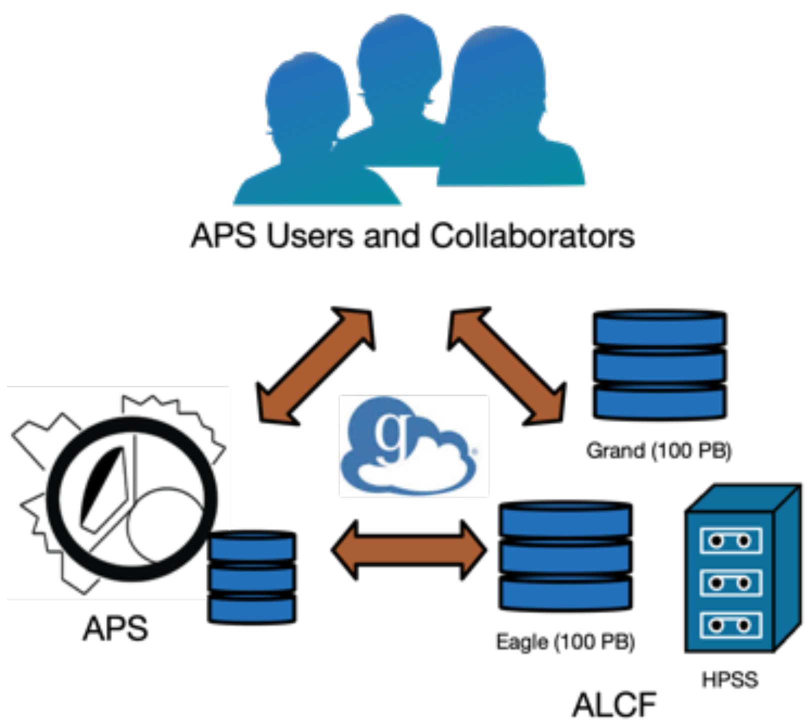


Figure 1-9 Storage available for APS beamlines. A multi-PB data storage system located at the APS serves medium-term needs. The ALCF provides multiple systems for long-term storage. Capacity will be expanded as needed to meet sponsor requirements.

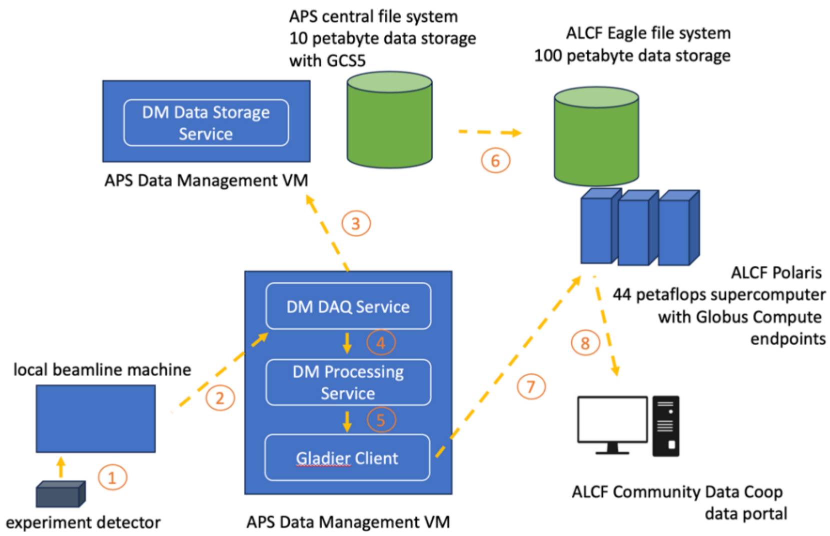


Figure 1-10 Standard workflow for file-based processing. 1) A detector generates an image from an x-ray experiment and writes to a local machine at the beamline. 2) The APS Data Management System monitors the local file system and uploads data files to central storage as they appear. 3) The APS Data Management System organizes data into directories for each beamline and experiment and assigns user access permissions. Users can then access data from Globus. 4) As data is uploaded, the APS Data Management System launches a processing job and performs any local preprocessing tasks. 5) The processing job launches a Globus Compute client. 6) The client initiates a Globus transfer to ALCF data storage. 7) The client submits a job as a service account to the on-demand queue for Polaris using the Globus Compute endpoint. 8) The processed results are published to a data portal created with the Globus Portal Framework as part of the ALCF Community Data Coop.

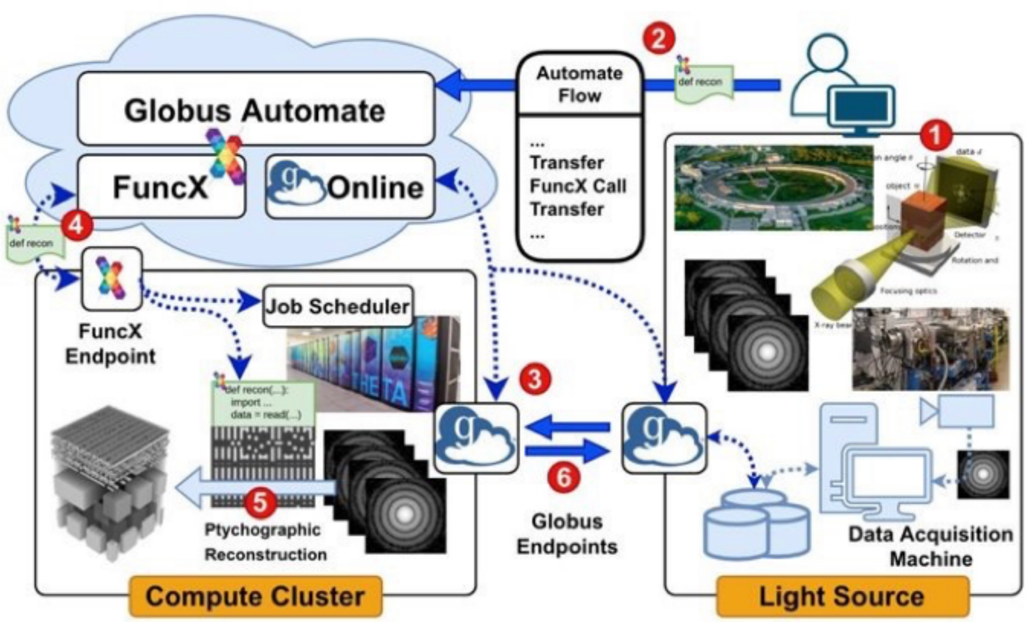


Figure 1-11 Automation used to perform on-demand analysis of ptychography data using computing resources at the ALCF. 1) Diffraction patterns are collected at the APS. 2) A beamline scientist submits a ptychography workflow definition file to the Globus Automate service. 3) The Globus Automate service begins executing the workflow and transfers the data from the beamline to the ALCF compute resource. 4) Remote function calls are triggered that 5) run the ptychography reconstruction code. 6) Results are transferred back to the beamline.

References

[1] Veseli, S., Schwarz, N. & Schmitz, C. (2018) APS Data Management System, J. Synchrotron Rad. 25, 1574-1580, DOI: 10.1107/S1600577518010056.

[2] Schwarz, N., Veseli, S., Jarosz, D. (2019) Data Management at the Advanced Photon Source, Synchrotron Radiation News 32:3, 13-18, DOI: 10.1080/08940886.2019.1608120.

[3] Park, JS., Horn, C., Ramanathan, P., Kenesei, P., Veseli, S. (2019) Data management and processing workflow for the Materials Physics and Engineering group beamlines at the Advanced Photon Source, J. Synchrotron Rad. 26, 373-381, DOI: 10.1107/S1600577519000584.

[4] Zhang, Q., Dufrense, E.M., Nakaye, Y., Jemian, P.R., Sakumura, T., Sakuma, Y., Ferrara, J.D., Maj, P., Hassan, A., Bahadur, D., Ramakrishnan, S., Khan, F., Veseli, S., Sandy, A.R., Schwarz, N., Narayanan, S. (2021) 20 μs-resolved high-throughput X-ray photon correlation spectroscopy on a 500k pixel detector enabled by data-management workflow, J. Synchrotron Rad. 28, 259-265, DOI: 10.1107/S1600577520014319.

[5] Vescovi, R., Chard, R., Saint, N., Blaiszik, B., Pruyne, J., Bicer, T., Lavens, A., Liu, Z., Papka, M. E., Narayanan, S., Schwarz, N., Chard, K., Foster, I. (2022) Linking Scientific Instruments and HPC: Patterns, Technologies, Experiences, DOI: 10.2139/ssrn.4141629.

[6] Bicer, T., Yu, X., Ching, D., Chard, R., Cherukara, M., Nicolae, B., Kettimuthu, R., Foster, I. (2021) High-Performance Ptychographic Reconstruction with Federated Facilities. In Nichols, J., Maccabe, A., Nutaro, J., Pophale, S., Devineni, P., Ahearn, T., Verastegui, B. (eds) Driving Scientific and Engineering Discoveries Through the Integration of Experiment, Big Data, and Modeling and Simulation. SMC 2021. Communications in Computer and Information Science, vol 1512. Springer, Cham.

[7] Parraga, H., Hammonds, J., Henke, S., Veseli, S., Allcock, W., Côté, B., Chard, R., Narayanan, S., Schwarz, N. (2023) Empowering Scientific Discovery Through Computing at the Advanced Photon Source. In Proceedings of the SC ’23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis, SC-W ’23, 2126–2132, DOI: 10.1145/3624062.3624612.

[8] Prince, M., Gürsoy, D., Sheyfer, D., Chard, R., Côté, B., Parraga, H., Frosik, B., Tischler, J., Schwarz, N. (2023) Demonstrating Cross-Facility Data Processing at Scale with Laue Microdiffraction. In Proceedings of the SC ’23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis, SC-W ’23, 2133–2139, DOI: 10.1145/3624062.3624613.

1.5 Computing Infrastructure

To *enable high-speed discovery* (Goal 1), the APS must leverage large-scale supercomputing resources, such as the ALCF, and the future HPDF [1], to address its computational demands. In order to *unlock exceptionally challenging experiments* (Goal 2), the APS must apply Argonne’s strengths in computing for large-scale simulations and x-ray data during beamtime to enhance the experimental process. Edge computing devices will provide rapid feedback at beamlines, ensuring that experiments can adapt dynamically and efficiently.

Current State

Demands for increased data processing at the APS are driven by new experimental capabilities enabled by the brightness increase and advances in detector data rates. Moreover, the increased reliance on multi-modal experiments to answer new scientific questions requires more complex and sophisticated data processing algorithms. Increases in computing power are needed by advanced algorithms for existing techniques that, for example, provide higher-fidelity results, and to train AI/ML models. The need for real-time analysis and feedback to make crucial experiment decisions and enable autonomous experiment steering also requires more computing cycles than have been traditionally utilized.

As with data storage, the computing resources required by the APS are anticipated to grow dramatically from typically TFLOP/s of computing resources. to tens of PFLOP/s of on-demand computing resources for the most advanced beamlines. There is wide variability in the computational requirements among techniques and processing approaches with those instruments and techniques that benefit most from high-energy, high-brightness, and coherent x-rays driving most requirements [2].

The APS adopts a graded approach to resource utilization. Small-scale resources, such as multi-core processors and GPUs, local to beamlines will be used when sufficient. For moderate computational needs, the APS maintains many powerful GPU-equipped workstations and a computing cluster in the APS computer room, and ANL maintains computing resources as a part of the Laboratory Computing Resource Center (LCRC). For the most demanding computational problems, large-scale computing facilities must be used, including the ALCF, the National Energy Research Scientific Computing (NERSC) Center, and the Oak Ridge Leadership Computing Facility (OLCF). To mitigate challenges surrounding processing and storing such large, anticipated data volumes, the APS is exploring the utilization of edge computing resources coupled closely to detectors and instruments, to run AI/ML data reduction algorithms. Figure 1-12 shows the ALCF systems deployed over the past decades. See Figure 1-13 for a list of computing resources available at Argonne.

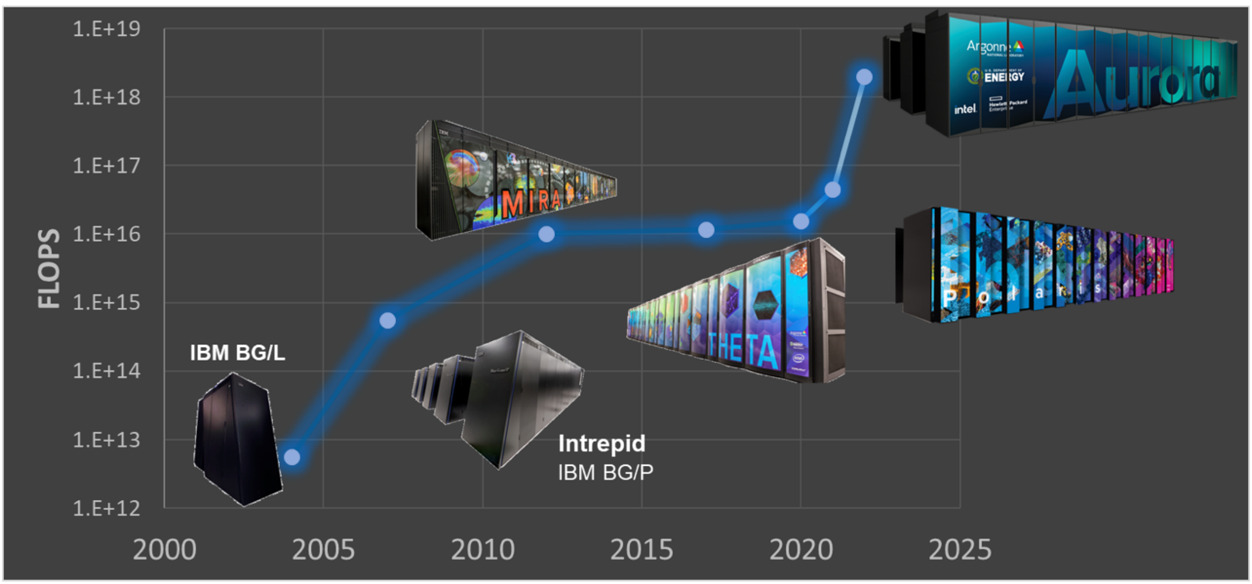


Figure 1-12 Timeline showing ALCF systems deployed over the past decades.

Integrated ALCF Supercomputing Resources at the APS: A New Era of APS Computing

The new storage ring and new and enhanced instruments, provide data rates that cannot reasonably be supported only by local APS resources. The colocation of the APS and world-leading supercomputing infrastructure at the ALCF on Argonne's campus provides an unprecedented opportunity for collaboration. The APS and ALCF have partnered to deliver a new model of computing, tightly coupling APS experiment instruments with ALCF supercomputers, to accelerate scientific discovery.

The ALCF has deployed a new computing system, Polaris, in 2022. Polaris is a combination commodity CPU/GPU system with performance of approximately 44 PFLOP/s. Up to 4 PFLOP/s of computing is prioritized to explore on-demand use of high-end computing resources by experimental and observational facilities, including the APS.

On-demand scheduling queues have been deployed for immediate access for APS jobs. Gateway nodes on this system will provide the ability for the APS to stream data directly to Polaris from detectors, avoiding local file I/O. The APS is working with Argonne's Data Science and Learning Division and the Globus team to develop a computational data fabric for end-to-end data lifecycle management, Globus Compute (see 1.4 above). A combined team of APS and ALCF scientists and engineers are developing end-to-end workflow pipelines for many techniques and beamlines that will connect APS instruments to this new resource [3].

The APS has been involved in many activities aimed at using centralized and large-scale computing resources. Notable activities include:

- A team comprising of staff at the APS, ALCF, and DSL have successfully demonstrated the first use of the new ALCF Polaris supercomputer for near real-time processing of XPCS data.
- ALCF researchers demonstrated processing XPCS and serial crystallography (SX) jobs in an on-demand fashion [4].
- At SC'19, ALCF researchers demonstrated large-scale real-time reconstruction and visualization of tomography data. This demonstration won the first annual SCinet Technology Challenge (TC) [5].
- Argonne scientists have demonstrated tomographic reconstructions of a fixed adult mouse brain specimen on the Summit Supercomputer [6, 7]. This work won the SC'20 Best Paper Award.
- A team of staff at the APS, ALCF, and DSL have demonstrated utilizing over 50 nodes of the ALCF Polaris supercomputer for on-demand near real-time processing of Laue microdiffraction data [8]. This work was demonstrated at SC'23.

Edge computing offers the ability to process data quickly on or near detectors and experiment instrumentation without the need to first transfer all data to high-end computing resources. This is particularly promising for handling large data when coupled with machine-learning methods. Using only a subset of data, machine-learning models may be trained on supercomputers. The trained model is then run using edge computing devices to process newly acquired data, providing fast feedback for experiment steering. See Figure 1-14. For example, APS and Argonne researchers have developed deep neural networks that perform ptychography reconstructions hundreds of times faster than the conventional iterative approaches and require up to 5 times less data [9], and 200 times faster than the conventional pseudo-Voigt profiling to locate Bragg peak positions [10].

Tactics related to computing infrastructure required to realize high-level goals:

1. *Work closely with the ALCF to determine a suitable and sustainable access/allocation model for routine utilization of ALCF systems by the APS.*
2. *Develop plans to utilize the resources of the future HPDF facility, and use the APS position as an IRI Pathfinder project to leverage developments from the IRI program.*
3. *Take advantage of edge computing devices to complement computational tasks performed on large-scale computing resources.*

This model coupling the APS and the ALCF to more seamlessly utilize large computing resources to enable the data processing needed at the APS, as it is refined using Polaris, will be deployed on more computing resources at the ALCF and at Argonne. These capabilities will be deployed for many other APS techniques and beamlines for data processing during beam time and for post-processing by APS Users after allocated experiment time is over. The Aurora exascale supercomputer is designed to support numerical simulation, data analysis, and deep learning applications. To this end, it is architected with a mix of Intel CPUs and GPUs to deliver sustained performance of greater than one exaflop/s full-precision floating point operations per second, and substantially higher compute rates at reduced precision including AI applications. The APS will utilize this new class of supercomputer to couple the results of simulations and modeling with experiment data and train ML models in real-time. The APS will work closely with the ALCF to determine appropriate access mechanisms to ALCF systems and an appropriate allocation model for utilization by APS beamlines.

The HPDF [1] will be a new DOE computing facility. Currently, in the early stages of planning, it is envisioned as a state-of-the-art resource for data science and research. The HPDF holds great potential to be a key enabler for the APS and its user community. The HPDF could provide seamless access to interoperable, scalable, and resilient computing resources: smaller-scale resources located at light source and neutron source Spokes, and larger-scale resources located at computing-facility Spokes and at the Hub(s), for real-time data processing, post experiment refinement, and simulation. Additionally, the HPDF could provide data catalogs and storage connected to computing resources for data management and retention and post experiment data processing for the User community, and especially aid AI/ML and digital twins efforts. The APS, and its light source and neutron source partners, continues to keep abreast of developments as HPDF planning progresses.

The new DOE SC IRI program (https://www.osti.gov/biblio/1984466) is being brought into formation. This new effort aims to provide researchers with seamless interoperability of DOE's unique data, user facilities, and computing resources. IRI is intended to be the infrastructure support, software, interface standards, and policies – that layers on top of existing DOE facilities making complex data-intensive workflows seamless and fast for research teams. The APS continues to remain involved in IRI activities, and most recently has been selected to participate as an IRI Pathfinder project.

R&D for the application of edge devices and methods will accelerate with effort and funding from the recent projects, X-ray & Neutron Scientific Center for Optimization, Prediction, & Experimentation (XSCOPE), and Intelligent Learning for Light Source and Neutron Source User Measurements Including Navigation and Experiment Steering (ILLUMINE) projects.

References

[1] The High-Performance Data Facility Project, https://www.hpdf.science.

[2] Schwarz N., Campbell S., Hexemer A., Mehta A., Thayer J. (2020) Enabling Scientific Discovery at Next-Generation Light Sources with Advanced AI and HPC, Driving Scientific and Engineering Discoveries Through the Convergence of HPC, Big Data and AI. SMC 2020. https://doi.org/10.1007/978-3-030-63393-6_10.

[3] Parraga, H., Hammonds, J., Henke, S., Veseli, S., Allcock, W., Côté, B., Chard, R., Narayanan, S., Schwarz, N. (2023) Empowering Scientific Discovery Through Computing at the Advanced Photon Source. In Proceedings of the SC '23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis, SC-W '23, 2126–2132, DOI: 10.1145/3624062.3624612.

[4] Salim, M., Uram, T., Childers, J. T., Vishwanath, V., Papka, M., "Balsam: Near Real-Time Experimental Data Analysis on Supercomputers," 2019 IEEE/ACM 1st Annual Workshop on Large-scale Experiment-in-the-Loop Computing (XLOOP), 2019, pp. 26-31, DOI: 10.1109/XLOOP49562.2019.00010.

[5] Liu, Z., Bicer, T., Kettimuthu, R., Foster, I., "Deep learning accelerated light source experiments," 2019 IEEE/ACM Third Workshop on Deep Learning on Supercomputers (DLS), 2019, pp. 20-28, DOI: 10.1109/DLS49591.2019.00004.

[6] Hidayetoglu, M., Bicer, T., de Gonzalo, S. G., Ren, B., Gursoy, D., Kettimuthu, R., Foster, I. T., Hwu, W. W., "MemXCT: Memory-Centric X-ray CT Reconstruction with Massive Parallelization", in Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis (SC'19), Denver, Colorado, Nov. 2019 DOI: 10.1145/3295500.3356220.

[7] Hidayetoglu, M., Bicer, T., de Gonzalo, S. G., Ren, B., Andrade, V., Gursoy, D., Kettimuthu, R., Foster, I. T., ßHwu, W. W., "Petascale XCT: 3D Image Reconstruction with Hierarchical Communications on Multi-GPU Nodes", International Conference for High Performance Computing, Networking, Storage and Analysis (SC'20), Nov. 2020, pp. 1-13. DOI: 10.1109/SC41405.2020.00041.

[8] Prince, M., Gürsoy, D., Sheyfer, D., Chard, R., Côté, B., Parraga, H., Frosik, B., Tischler, J., Schwarz, N. (2023) Demonstrating Cross-Racility Data Processing at Scale with Laue Microdiffraction. In Proceedings of the SC '23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis, SC-W '23, 2133–2139, DOI: 10.1145/3624062.3624613.

[9] Cherukara, M. J., Zhou, T., Nashed, Y., Enfedaque, P., Hexemer, A., Harder, R. J., Holt, M. V., "AI-enabled high-resolution scanning coherent diffraction imaging." Applied Physics Letters 117, no. 4 (2020): 044103.

[10] Liu, Z., Sharma, H., Park, J. S., Kenesei, P., Miceli, A., Almer, J., Kettimuthu, R., Foster, I. T. (2022) BraggNN: fast X-ray Bragg Peak Analysis Using Deep Learning, IUCrJ 9, 104-113.

Argonne Leadership Computing Facility (ALCF)

Polaris
~44 PFLOP/s

Aurora
> 1 EXAFLOP/s

~4 PFLOP/s of Polaris is prioritized for prototype on-demand use by experimental and observational facilities; when Aurora is in User operations, all of Polaris will be prioritized for on-demand use

Synergy
Planning is underway for the next generation on-demand system prioritized for experimental and observational facilities

Next Generation Supercomputer
Planning is underway for the next generation leadership class supercomputer

Argonne Laboratory Computing Resource Center (LCRC)

Improv
~2.51 PFLOP/s
825 nodes with 2 AMD EPYC CPUs each

Bebop
~1.75 PFLOP/s
672 nodes with 36 Intel Broadwell cores each

Swing
~925 TFLOP/s
48 NVIDIA A100s | 768 AMD EPYC cores

Advanced Photon Source (APS)

APS general purpose distributed-memory compute cluster
- ~20 TFLOP/s CPU cores

~50 High-Performance Computing Workstations
- Califone – 8 H100s
- Ecto – 8 RTX A6000s
- Refiner – 4 A100s
- Many others…

Edge Computing Devices
- 1 x NVIDIA Jetson AGX Orin
- 2 x NVIDIA BlueField-3 DPUs

ALCF AI Platforms

GroqRack
188 TFLOP/s (FP16) and 750 TFLOP/s (INT8)
72 accelerators

Cerebras (CS-2)
5780 TFLOP/s (FP16)
580,000 processor cores

SambaNova Dataflow
660 TFLOP/s (BF16)
64 accelerators

Graphcore Bow Pod64
250 TFLOP/s (FP16)
64 accelerators

Habana Gaudi-1
150 TFLOP/s (FP16)
16 accelerators

Figure 1-13 Computing resources and respective specifications and performance available for use by the APS at Argonne.

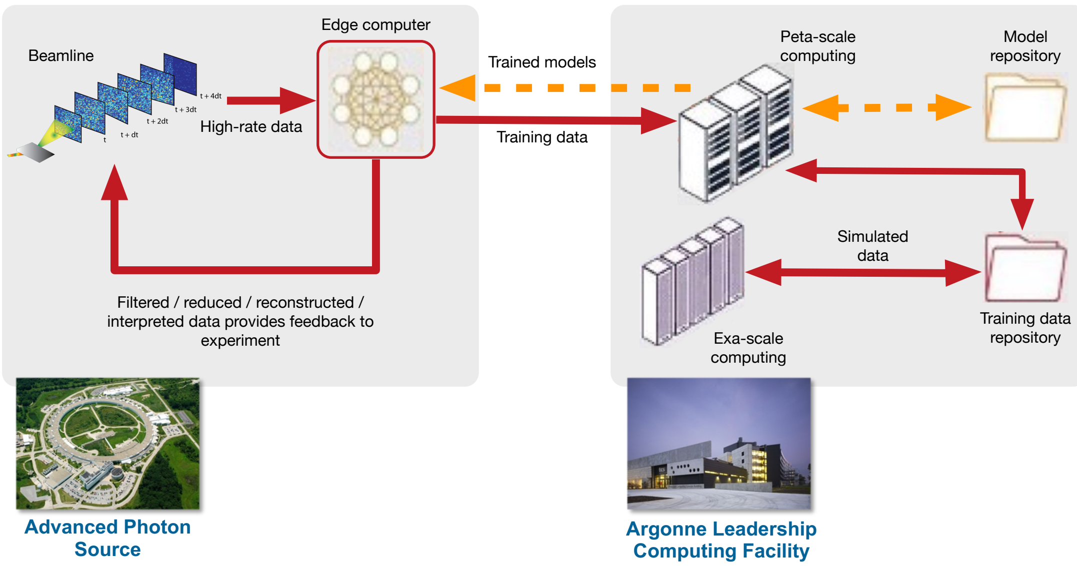


Figure 1-14 Edge computing architecture using machine-learning models trained on supercomputers.

1.6 Data Reduction and Analysis Software

Data reduction and analysis plays a crucial role in accomplishing APS strategic computing goals. To *enable high-speed discovery* (Goal 1), the deployment of HPC-enabled software and workflows using Globus is essential for facilitating near real-time data processing. This approach accelerates the ability to extract meaningful insights from experiment data. For *unlocking exceptionally challenging experiments* (Goal 2), the application of advanced computational algorithms enables and enhances the ability to tackle difficult scientific problems. To *empower users to realize the full potential of the APS* (Goal 4), data processing software is being developed and refined to align closely with the specific needs of beamlines and users. This includes the ability to integrate multi-modal data, ensuring that users can fully leverage the diverse types of data generated by the new and upgraded beamlines at the APS.

Current State

The APS is focusing data analysis algorithm and software development in the areas needed to answer novel scientific inquiries enabled by the renewed APS. These areas are techniques driven by coherence, imaging, and high-energy, as well as multi-modal techniques. Algorithms and software are being developed to analyze and reconstruct massive data volumes, bridge across length and time scales, combined and understand data from multiple modalities, identify and classify features and patterns, and provide feedback to experiments dynamically using real-time reduction.

Coherence, imaging, high-energy, and multi-modal techniques are already the most computationally intensive techniques performed at the APS. Throughput demands are expected to grow by as much as multiple orders of magnitude due to improved detectors and the upgraded source. Data reduction and analysis will rely heavily on the use of high-performance computing (HPC), utilizing appropriate technologies such as multi-threading, General Purpose Graphical Processing Units (GPUs), edge devices, and distributed computing environments to obtain results with near real-time completion, so that results enable user-driven or even automated steering of experiments.

Most software is developed as open source and is made available with user community code contributions encouraged. A graded approach according to impact and priority is applied to development. Packaging and active support either as distributable applications or as Software-as-a-Service (SaaS) is provided for software systems that have been deemed to be most important for the success of APS and its users.

New efforts are underway to address the development of new algorithms and HPC software for multi-modal analysis, including

- Fluorescence tomography
- Fluorescence ptychography
- Magnetic ptychography
- Tomography diffraction
- Bragg CDI and ptychography
- New approaches to Laue diffraction reconstructions

Key software developments that have been made in this area over the past years that align with priority needs are shown in Figure 1-15. Detailed descriptions and plans for algorithm and software development for each APS-U feature beamline, including funding sources and collaborative efforts, may be found in each APS-U feature beamline's area in section 2. A summary of required core software capabilities for these beamlines, software packages that are intended to address these needs, and their status may be found in Table 1-2. Beamlines not directly a part of the APS-U project still benefit from the reuse of tools developed for these priority applications.

Table 1-2 Summary of software capabilities and software packages and their status for the APS-U feature beamlines.

Beamline	Capability	Software	Status
ATOMIC	Conventional Phase Retrieval Reconstructions for Bragg CDI	cohere	Complete
	Faster Conventional Phase Retrieval Reconstructions for Bragg CDI	cohere	In Progress
	High-Resolution Conventional Phase Retrieval for Bragg CDI	cohere	In Progress
	AI / Automatic Differentiation (AD) Methods	cohere	In Progress
CHEX	Conventional Phase Retrieval Reconstructions for Bragg CDI	cohere	Complete
	Ptychography in Bragg geometry	TBD	To Do
	Rod Analysis (COBRA)	TBD	To Do
	Dark Field X-ray Microscopy (DFXM)	R&D Prototyping	In Progress
	Multi-Tau & Two-Time XPCS Correlations	XPCS-Eigen, XPCS-Boost	Complete
	Surface XPCS	XPCS-Eigen, XPCS-Boost	In Progress
	Higher-Order XPCS Correlations	R&D Prototyping	In Progress
CSSI	Thin-film Structure Indexing	GI-SAXS GUI	Complete
	Coherent Surface Scattering Imaging	R&D Prototyping	In Progress
	Surface XPCS	XPCS-Eigen, XPCS-Boost	In Progress
HEXM	Near-field Diffraction	MIDAS	Complete
	Far-field Diffraction	MIDAS	Complete
	Diffraction (Scattering) Tomography	MIDAS	In Progress
	Imaging Tomography	TomoPy, MIDAS	Complete
ISN	XRF Elemental Fitting	XRF-Maps	Complete
	XRF Self-absorption Correction	XRF-Maps	In Progress
	Conventional Ptychography Reconstructions	tike, ptychodus	Complete
	Improved Ptychography Quality	tike, ptychodus	In Progress
Polar	Hard Resonant Magnetic Scattering	XMCD tools, PyMCA	Complete
	Hard Resonant X-ray Ptychography	R&D Prototyping	In Progress
	Bragg CDI Magnetic Contrast	R&D Prototyping	In Progress
	Tomographic CDI	TBD	To Do
PtychoProbe	XRF Elemental Fitting	XRF-Maps	Complete
	XRF Self-absorption Correction	XRF-Maps	In Progress
	Conventional Ptychography Reconstructions	tike, ptychodus	Complete
	Improved Ptychography Quality	tike, ptychodus	In Progress
XPCS	Multi-Tau & Two-Time XPCS Correlations	XPCS-Eigen, XPCS-Boost	Complete
	Higher-Order XPCS Correlations	R&D Prototyping	In Progress
	Spatial-Temporal XPCS Cross-Correlations	R&D Prototyping	In Progress
3DMN	Wire Scan Laue Reconstructions	LaueGo	Complete
	Mask Scan Laue Reconstructions	R&D Prototyping	In Progress

Applications will continue to be developed for improved performance and new algorithms. A more complete list of software produced at the APS can be found at https://www.aps.anl.gov/Science/Scientific-Software and https://github.com/AdvancedPhotonSource.

Tactics related to data reduction and analysis software required to realize high-level goals:

1. *Support existing and develop new HPC-enabled software for priority areas of coherence, high-brightness, and high-energy driven techniques.*
2. *Develop real-time streaming tools for feedback during experiments.*
3. *Explore and deploy software and algorithms for use on edge devices.*
4. *Develop algorithms and methods for advanced computational techniques and multi-modal data fusion and utilization.*
5. *Apply robust optimization and uncertainty quantification.*

Figure 1-15 Key software packages under development at the APS.

1.7 Artificial Intelligence / Machine Learning (AI/ML)

Artificial Intelligence / Machine Learning (AI/ML) advances are critical to achieving high-level goals at the APS. AI/ML inference must be applied to accelerate data processing, allowing for quicker analysis and interpretation of complex datasets to *enable high-speed discovery* (Goal 1). In order to *unlock exceptionally challenging experiments* (Goal 2), AI/ML, combined with digital twin technologies, will enable the design and steering of autonomous experiments, providing real-time insights and adjustments during experiments. AI/ML plays a pivotal role in analyzing and understanding the vast and diverse datasets generated at the APS, enabling the extraction of new scientific insights, *leverage data for new science* (Goal 3). To *empower users to realize the full potential of the APS* (Goal 4), advancements in AI/ML must be integrated into operational workflows at instruments, ensuring that users can benefit from cutting-edge technology to enhance their research capabilities.

Current State

Advanced data processing and analysis methods will be crucial to keep pace with the expected data rates and volumes, enabling real-time experiment steering and more efficient scientific discovery. Historical trends indicate that data rates and computing demands at x-ray sources have consistently outpaced available computing resources. Figure 1-16-A illustrates this by plotting the brightness at facilities over the past 60 years against Moore's Law; it clearly shows that the increase in brightness – and consequently, the associated data and computing requirements – have exceeded transistor scaling. Similarly, Figure 1-16-B compares the brightness at various generations of facilities with the capabilities of the fastest supercomputers at those times, further demonstrating that the scaling of x-ray facility has surpassed the growth in available computational power.

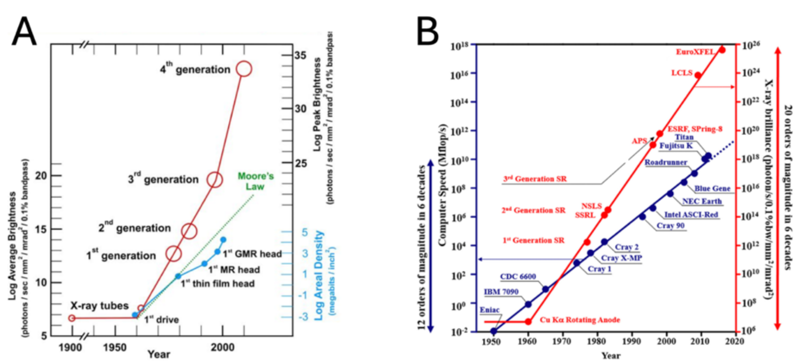


Figure 1-16 Data and compute needs are outpacing traditional computer scaling. Brightness of facilities over the last 60 years is plotted against Moore's law A) and the evolution of the fastest supercomputer in the world B).

The development of new x-ray characterization techniques has historically depended on the simultaneous invention of algorithms and mathematical models. These computational tools are vital for analyzing and interpreting the data each new technique generates. For instance, several synchrotron imaging techniques, such as ptychography and tomography, owe their feasibility to specific computational imaging methods. While numerical algorithms have historically enabled groundbreaking science at synchrotron sources, the next-generation light sources pose significant computational challenges that many current algorithms may struggle to meet due to the sheer volume of data expected in the future.

In response to these challenges, Artificial Intelligence and Machine Learning (AI/ML) techniques are emerging as powerful tools. They not only accelerate x-ray data analysis but also enhance the robustness and expand the potential applications of these methods. The overarching motivation of leveraging AI/ML at the APS is to unlock new scientific capability from existing instruments while improving and enhancing the users' experience and scientific productivity.

Efforts in this space can be divided into four main thrusts:

- **AI4Analysis:** Real-time analysis and visualization that is >100X faster and (sometimes) more accurate analysis than conventional methods. Enables real-time analysis of Gb/s data streams
- **AI4Steering:** AI-guided steering of accelerator, instrument and experiments allowing autonomous accelerator and instrument tuning along with autonomous tracking and targeted acquisition of data
- **AI4Knowledge:** Extracting scientific insight from very large, complex multi-modal datasets
- **LLMs as Scientific Co-Pilots:** Context-aware large language models (LLMs) that help users navigate different aspects of experimentation at the APS; experiment planning, guidance on using complex instruments and analysis software, and directly performing basic instrument operations

The APS has realized early successes applying AI in various aspects of accelerator and beamline operations. Some examples of how AI has been used to speed up data analysis, automate operations, improve scientific knowledge extraction, and assist users in experiment planning and execution follow.

Work continues to develop and apply new AI/ML methods. The APS has and continues to invest heavily in AI/ML develops through the LDRD program (see the list of LDRDs in Section 1.8). With recent funding from the DOE for Artificial Intelligence and Machine Learning at DOE Scientific User Facilities (see Section 1.8), the APS is collaborating on AI/ML tools for spectroscopy data analysis, a digital twin for in silico time-resolved experiments, high-energy diffraction microscopy data reduction, accelerator tuning and optimization, and sharing and cataloging ML models and data. See https://www.anl.gov/ai for a full list of AI/ML developments.

AI4Analysis - High Performance Computing (HPC) and AI@Edge Enables Real-time Imaging

A representative example for the sheer volume of data that needs to be processed in real-time is ptychography. For example, conducting a single raster scan over a 1 mm by 1 mm area with a step size of 100 nm can generate 200 terabytes of raw data when using a moderately sized detector with one million pixels and a 16-bit dynamic range. This results in a data transfer rate of 16 gigabits per second and necessitates approximately petaflops of computational power to carry out phase retrieval.

Using ptychography as a representative technique, we demonstrated an AI-enabled workflow that can provide analysis on streaming data up to 8 KHz, facilitating real-time coherent imaging. The approach leverages High-Performance Computing (HPC) resources for online neural network training, alongside a cost-effective, compact embedded GPU system running a trained neural network positioned at the beamline for near-instantaneous phase retrieval. Figure 1-17-A shows a schematic of the workflow. Through Globus Compute, data from the instrument is moved to a HPC resource like Polaris as soon as it is collected. HPC is used for conventional analysis and distributed model training. The trained model, in this case PtychoNN, learns to solve the inverse problem of phase retrieval for ptychography. This trained model is pushed to the edge computing device to provide live inference from the raw data stream from the detector which is displayed on the user screen. Figure 1-17-B, C shows examples of the raw detector image, the prediction from the neural network and the cumulative image obtained by stitching the predictions together. Online or continual training, i.e. the process of constantly updating the model during the course of the experiment ensures that the model is trained.

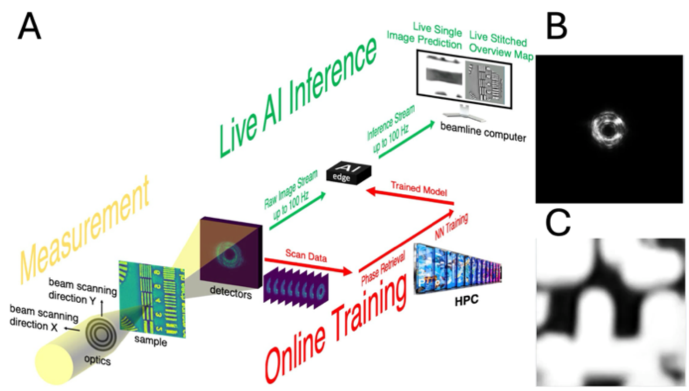


Figure 1-17 HPC + AI@Edge enables real-time, streaming ptychography. A) Schematic of workflow moving data from instrument to Polaris for labeled data generation and training. The trained model is pushed to the edge for live inference. B) Detector image that is input to the trained network. C) Single-shot prediction from the network.

AI4Analysis - AI-accelerated Spectroscopy

ML is used in x-ray emission and absorption spectroscopy for data processing and information extraction. The Argonne X-ray Emission Analysis Package (AXEAP) has been developed to process and analyze X-ray emission spectroscopy (XES) data collected with a two-dimensional (2D) position sensitive detector. AXEAP is designed to convert a 2D XES image into a one-dimensional XES spectrum and perform quantitative analysis in real time using unsupervised machine learning. K-means clustering is used to automatically determine regions of interest from 2D XES images. AXEAP can determine regions of interest (ROIs) in the 2D image data space at a rate similar to data collection, allowing real time comparisons during data collection, reducing the amount of data stored from gigabyte-sized image files to kilobyte-sized text files. With a user-friendly interface, AXEAP includes data processing for non-resonant and resonant XES images from multiple edges and elements.

The determination of charge, spin, and chemical bonding information from extracted XES as well as x-ray absorption near edge structure (XANES) also benefit significantly from the use of ML. One important aspect is the featurization of 1D spectroscopy data for the supervised learning of chemical and bonding information. Because of the difficulty involved in obtaining experimental datasets with labeled ground truths, computational simulations of XES and XANES data are relied upon. Featurization approaches such as peak extraction, principal component analysis, variational autoencoders, and continuous wavelet transforms are used in conjunction with supervised ML methods such as multi-layer perceptron, random forest, and gradient boost regression models, to determine the most effective approach for extracting coordination numbers and other structural descriptors from raw spectra.

AI4Steering - Autonomous Control of Complex Optical Systems

To exploit the opportunities offered by the novel fourth-generation synchrotron radiation facilities like the upgraded APS, highly focused x-ray beams with minimum wavefront distortion, high stability, and variable focal sizes are required. Optical elements are being designed with challenging manufacturing requirements. However, optical aberrations, mechanical vibration and drift, and heat loading can all deform the beam wavefront and degrade the quality of the detectable signal. A possible solution to compensate for these wavefront aberrations is to use adaptive optics (AOs) combined with a real-time wavefront sensor and intelligent automatic control system. Several AO applications have been demonstrated, such as a prototype zoom mirror system and an active cooling mirror, using the traditional feedback control system based on the linear response model. However, time and history-dependent behaviors of AOs cannot be well represented by linear models, especially when considering timesteps on the order of seconds. Therefore, our research has been directed toward developing control systems based on AI, capable of rapidly achieving and keeping optimal and desired wavefront characteristics. In collaboration with the Advanced Light Source (ALS), the APS has successfully demonstrated ML control of a piezo-bimorph mirror, including a feed-forward neural network to predict the mirror shape and an optimization model to drive to the desired surface shape. The system was initially developed and tested using a visible-light interferometer capable of recording the mirror surface profile. Recently, we demonstrated the system's efficacy at an APS beamline by training the ML system with the absolute phase and the local radius of curvature of the wavefront measured with a wavefront sensor with different voltages applied to piezo electrodes. The system can systematically create wavefronts with desired shapes, such as a spherical wavefront focusing on the desired location. Fast and automatic alignment of optics is beneficial for experiments and can maintain the conditions on which the ML data is collected to ensure the model repeatability. The APS is currently developing AI-driven control systems to optimize beam properties, such as the focal spot position and size, by acting on positioning motors, e.g. the pitch angle of a mirror. To efficiently study different approaches, we have developed ultra-realistic digital twins using the OASYS simulation libraries to represent real components such as bendable mirrors. The APS envisions that the combination of the AI-driven auto-alignment system and the ML control of AOs will become a new beamline standard for next-generation light sources, where exceptional beam stability, repeatability, and total control of an aberration-free wavefront are required.

AI4Steering - AI-guided Scanning Microscopy

The APS has implemented the Fast Autonomous Scanning Toolkit (FAST) that uses the SLADS-Net NN, route optimization, and efficient and modular hardware controls to make on-the-fly sampling and scan path choices for a synchrotron-based scanning microscopy experiment. The NN used here, a fully connected network containing 5 hidden layers with 50 nodes each, is trained using only a generic natural image with no experiment-specific tuning. The FAST framework has a low computational cost that is negligible compared to the acquisition time, even when used within a low-power edge-computing device placed at the synchrotron beamline. This fast and sample-agnostic autonomous experimentation technique is well-suited for application in a synchrotron beamline that is versatile by nature.

A schematic of the FAST workflow is shown in Figure 1-18. The FAST workflow is initialized by measuring a quasi-random selection of 1% of the sample area, then transferring these measurements to an edge device, a NVIDIA Jetson Xavier AGX, connected to the beamline computer. The edge device uses the inverse distance-weighted interpolation to reconstruct the full image from the sparse measurement set. The SLADS-Net NN then uses the measurement information to identify a batch of 50 unmeasured points that would most improve the reconstructed image. The coordinates of these points are supplied to a route optimization algorithm to generate the shortest part for the scan motors to visit all these points. This path is supplied to the scan motors, two piezoelectric linear translation motors in step mode, through an EPICS interface, which restarts the data acquisition phase. The data acquisition and analysis steps are repeated until a pre-specified stopping criterion is achieved.

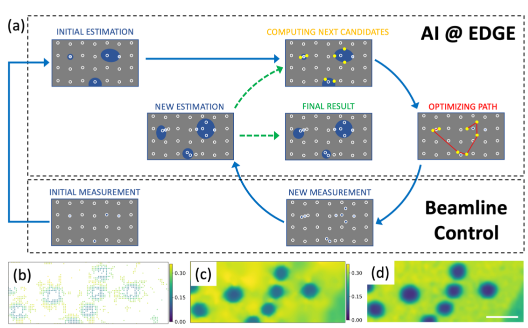


Figure 1-18 a) Schematic of the FAST workflow. A series of quasi-randomly selected points were picked for the initial measurements. The result was transferred to the edge device which proceeded to generate an initial sample estimate. The rest of the experiment was conducted in a sequential and iterative manner: The edge AI determined the candidate points for the next measurements and calculated an optimal path. The path was scanned by the beamline control which sent the new results back to the edge AI for a new estimation. This iterative process continued until a predefined completion criterion is met. b) The points scanned at 20% of coverage. c) Estimation of the sample by edge AI at 20% coverage. d) Ground truth obtained after raster scanning at 100% coverage. The scale bar is 2 um.

AI4Knowledge - AI-NERD for XPCS Analysis

Understanding and interpreting the dynamics of functional materials in situ represents a significant challenge in the fields of physics and materials science, primarily due to the complexities involved in experimentally probing materials across various length and time scales. X-ray photon correlation spectroscopy (XPCS) is particularly well-suited for examining material dynamics across a broad range of time scales. However, the spatial and temporal heterogeneity observed in material behavior can complicate the interpretation of experimental XPCS data. The APS has developed an unsupervised deep learning (DL) framework designed for the automated classification of relaxation dynamics from experimental data, eliminating the need for prior physical knowledge of the system. We showed how this method can expedite the exploration of massive datasets to pinpoint snapshots of interest. Furthermore, we employed this approach to directly link microscopic dynamics with macroscopic properties in a model system. Notably, this DL framework is both material and process agnostic, representing a significant advancement towards autonomous materials discovery.

LLMs as Scientific Co-Pilots - Context-Aware Language Model for Science (CALMS)

The ever-increasing instrument complexity at light source instruments poses greater challenges for domain scientists in designing experiments that effectively utilize and operate these sophisticated instruments. Large language models (LLMs) have the capability to perform intricate information retrieval, assist in knowledge-intensive tasks across various applications, and provide guidance on tool usage. We reported initial experiments with a Context-Aware Language Model for Science (CALMS) aimed at aiding scientists with instrument operations and complex experimentation. CALMS can access and retrieve pertinent information from facility documentation, thereby answering questions about scientific capabilities and other operational procedures. Moreover, with its ability to interface with software tools and experimental hardware, CALMS can facilitate conversational operation of scientific instruments. Figure 1-19-A shows an overview of CALMS. CALMS consists of a LLM with access to document stores to answer user queries and access to control systems to perform instrument operation. Figure 1-19-B shows examples of CALMS answering questions on experimental planning, user operations and directly driving an instrument.

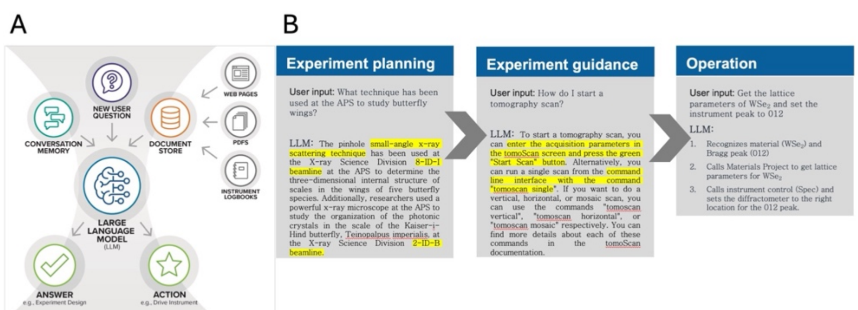


Figure 1-19 A) Overview of CALMS: CALMS uses a large language model in conjunction with conversational memory, document stores, and experimental tools to answer user queries or take action to drive an instrument. B) Examples of CALMS interacting with a user in the experiment planning stage, guiding an user through data acquisition and directly calling instrument controls with user input.

In addition to these advances, the APS is engaged in many other impactful AI/ML projects, including:

- *HPC & AI@Edge Enables Real-Time Peak Fitting:* The APS has developed BraggNN, a deep-learning based method that can determine peak positions much more rapidly (200 times faster) and more accurately than conventional pseudo-Voigt peak fitting.
- *Accelerator Tuning and Fault Mitigation:* The APS is developing ML methods to efficiently achieve and maintain optimal accelerator performance through reinforcement learning (RL) and Bayesian optimization (BO).

- *Adaptive XANES Experimentation Through Physics-informed AI:* The APS developed an adaptive sampling algorithm, which incorporates an acquisition function designed with prior structural knowledge about XANES spectra.
- *Learning Material Models from Diffraction Data:* ML is used to scale up small models of ab initio level calculations to larger molecular dynamics box sizes, capturing the structural, dynamic, and physical properties of a material over a wide range of temperatures and pressures. These ML models provide new physical insights into the temperature dependent coordination environment of the disordered state as well as densities, self-diffusion constants, and ionic conductivities.
- *Data-Driven Discovery of Dynamics from Coherent Scattering*: The APS developed a data-driven framework which employs neural differential equations to parameterize unknown real-space dynamics and a computational scattering forward model to relate real-space predictions to reciprocal-space observations. This framework was shown to recover dynamics of several computational model systems, including domain synchronization, particle clustering, and source fluctuation, without solving the phase reconstruction problem for the entire time series of diffraction patterns.
- *AutoPhaseNN - Unsupervised Physics-aware Deep Learning of 3D Nanoscale Bragg Coherent Diffraction Imaging:* A neural network that uses known physics and learns to invert 3D Bragg CDI data completely unsupervised, i.e., without ever being shown sample images during training, that is more than 10x faster than traditional methods while also being more accurate.

Tactics related to AI/ML required to realize high-level goals:

1. *Align with goals and activities related to the DOE Frontiers in Artificial Intelligence for Science, Security, and Technology (FASST) initiative, especially in the areas of AI-ready data and AI applications.*
2. *Develop physics-aware AI methods and exploration of LLMs for experiments; continue development of AI-enabled high-resolution coherent imaging methods including sparse-sampled ptychography, coherent surface scattering and Bragg ptychography with dynamical effects; and continue exploration of unsupervised deep learning (DL) techniques for phase retrieval that will accelerate and improve quality of reconstructions without human input.*
3. *Create a plan to bring AI/ML advances into operational use at beamlines.*
4. *Develop digital twins of beamlines and sample environments and integrate them into model training workflows.*

1.8 Effort, Funding, and Collaborations

Fruitful collaborations support all the APS strategic computing goals. By leveraging the 6-way collaboration among the DOE BES light and neutron sources, institutions can develop unified solutions and approaches to address common data and computing challenges, enhancing efficiency and innovation. Creating an engagement roadmap to integrate light source-enabled research into the DOE's Frontiers in Artificial Intelligence for Science, Security, and Technology (FASST) initiative ensures that cutting-edge AI advancements are applied to important scientific frontiers and that data generated at the light sources is leveraged to the fullest. An integrated facility approach to computing and data resources, as part of the DOE SC ASCR IRI program and HPDF project, will open the door to cohesive utilization of an increased number of advanced resources. These collaborative efforts will enable the APS to amplify its impact and effectiveness.

Current State

Effort

In addition to the efforts described in Section 1.1, instrument and accelerator scientists from across the APS also provide effort towards this strategy in a multitude of ways, including setting goals, acting as liaisons for the user community, developing and applying algorithms as a part of technique and instrumentation development, assisting in the analysis and interpretation of data, and writing prototype and operational software tools.

Funding

The APS-U project provides funding for networking infrastructure within the APS-U feature beamlines. Controls systems for the APS-U feature beamlines are also supported by the APS-U project. The APS-U project may provide funding for certain local computing resources at APS-U feature beamlines but the majority of resources and effort are outside of APS-U project scope.

One way Argonne National Laboratory supports computational efforts at the APS is via Laboratory Directed Research & Development Funding (LDRD) funding. Beginning in FY11, the *Tao of Fusion* LDRD helped seed the *TomoPy* application and the APS Data Management System; likewise, the FY13 *Next Generation Data Exploration: Intelligence in Data Analysis, Visualization and Mining* LDRD was aimed at multi-modal analysis. Other previously funded LDRDs include *Visualization and Mining, Modeling, Analysis, and Ultrafast Imaging (MAUI)*, *Multimodal Imaging of Materials for Energy Storage (MIMES)*, *Enabling Nanometer-scale X-ray Fluorescence Tomography*, and *Coherent Surface Scattering Imaging*.

Additional funded LDRDs of direct benefit to the APS in the computing space include:

- FY17 *Integrated Imaging*
- FY17 *A Universal Data Analytics Platform for Science*
- FY17 *COHED: Coherence for High-Energy Diffraction*
- FY17 *Developing Advanced Coherent Surface Scattering Reconstruction Method Incorporating Dynamical Scattering Theory*
- FY17 *Enabling Multidimensional X-ray Nano-Tomography*
- FY17 *The Perfect Thermodynamics of Imperfect Materials*
- FY18 *A.I. C.D.I.: Atomistically Informed Coherent Diffraction Imaging*
- FY18 *Integrated Approach to Unravel Four Dimensional Spatiotemporal Correlation in Highly Transient Phenomena: Ultrafast X-ray Imaging and High-Performance Computing*
- FY18 *Novel Capabilities for Ultra-fast and Ultra-low-dose 3D Scanning Hard X-ray Microscopy*
- FY19 *Enabling Automatic Learning of Atmospheric Particles through APS-U*
- FY19 *Finding Critical Processes of Deformation in Structural Materials with Artificial Intelligence*
- FY19 *Learning and Differentiating: Using Artificial Intelligence to Image Beyond the X-ray Depth of Focus Limit*
- FY19 *Machine Learning Enabled Advanced X-ray Spectroscopy in the APS-U Era*
- FY20 *Machine Learning Methods for Spectral Data from X-ray Transition Edge Sensor Arrays*
- FY20 *Tomographic Data Analysis Accelerated by Deep Learning*
- FY20 *Self-supervised deep learning for x-ray imaging without reference data*
- FY20 *Coded Apertures for Depth Resolved Diffraction*
- FY20 *Intelligent Ptychography Scan via Diffraction-Based Machine Learning*
- FY20 *AI-steer: AI-driven online steering of light source experiments*
- FY20 *AI patterns for executable end-to-end biological programming experiments*
- FY20 Innovate *High-Energy X-ray Diffraction and Machine Learning Driven Molecular Dynamics Simulation Study of Molten Chloride Salts*
- FY21 *AutoPtycho: Autonomous, Sparse-sampled Ptychographic Imaging*
- FY21 *Scalable DL-based 3D X-ray nanoscale imaging enabled by AI accelerators*
- FY21 ALCF Expedition *Scalable DL-based 3D X-ray Nanoscale Imaging Enabled by AI Accelerators*
- FY21 *High Pressure Material Characterization in 3-Dimensions Using X-Ray Diffraction Contrast Computed Tomography*
- FY22 *Intelligent Analysis of Scattering and Spectroscopic Signatures of Quantum Materials*
- FY22 *Development of 3D dichroic ptychography at the APS*
- FY22 *High Energy X-Ray Imaging for Non-Destructive and Rapid Nuclear Forensics*
- FY22 *Intermittent Dynamics in Hard and Soft Materials enabled by APS-U*

- FY22 ALCF Expedition *Deep Learning Accelerated X-ray Data Analysis for Experiment Steering*
- FY22 ALCF Expedition *Machine earning at the edge for real-time analysis in X-ray ptychography enabled by hardware AI accelerators*
- FY22 ALCF Expedition *AI accelerator for scalable DL-based 3D X-ray nanoscale imaging*
- FY22 ALCF Expedition *Exploring Groq as a Real-time AI Inference Accelerator for Scientific Instruments*
- FY22 ALCF Expedition *Scalability Study of AI-based Surrogate for Ptychographic Image Reconstruction on Graphcore*
- FY23 *AI-Driven, Real-Time Optics Control System to Achieve Aberration-Free Coherent Wavefronts at 4th Generation Synchrotron Radiation Beamlines*
- FY23 *Auto Parameter Calibration for X-ray Fluorescence Spectrum Fitting Using Machine Leaning*
- FY23 *AI/ML Accelerated High-Performance Image Analysis using Supercomputers*
- FY23 *Three-Dimensional Multiscale Diffraction Imaging of Nano-Scale Defect Kinetics During Corrosion and Mechanical Deformation*
- FY23 *Developing Point-Focus High-Energy Diffraction Microscopy to Reveal Battery Material Degradation*
- FY23 *The Development of Combining Ptychography and Small-Angle X-ray Scattering for Nanomaterial Characterization*
- FY24 *Simulation-Guided High-Energy X-ray Diffraction Experiment Framework for Understanding Irradiation-Induced Slip Anisotropy in Structural Metals*
- FY24 *Pushing the Experimental Envelope: High-Resolution and Ultra-High-Speed X-ray Imaging via Spatiotemporal Image Fusion*
- FY24 *A Pipeline for Autonomous Comprehensive Reduction and Analysis of Sequential 2-Dimensional Diffraction and Scattering Data*
- FY24 LDRD Innovate (Seed) *Crystal Graph-based Generative Model for X-ray Crystallography*

The APS has received funding and personnel support from the ALCF Data Sciences Program (ADSP):

- *Large-Scale Computing and Visualization on the Connectomes of the Brain*
- *Developing High-Fidelity Dynamic and Ultrafast X-ray Imaging Tools for APS-Upgrade*
- *X-ray Microscopy of Extended 3D Objects: Scaling Towards the Future, and Dynamic Compressed Sensing for Real-Time Tomographic Reconstruction*
- *Dynamic Compressed Sensing for Real-Time Tomographic Reconstruction*

The NERSC Exascale Science Applications Program supported the APS on the *Optimization of data-intensive tomography workflows at light sources* project.

The APS receives funding for AI/ML efforts in part from collaborative awards from the DOE for Artificial Intelligence and Machine Learning at DOE Scientific User Facilities**,** Lab 20-2261, and subsequent renewal awards:

- *A Collaborative Machine Learning Platform for Scientific Discovery*, Principal Investigator (PI) - Alex Hexemer (Advanced Light Source, Lawrence Berkeley National Laboratory [LBNL]), Subramanian Sankaranarayanan (CNM-Argonne), Nicholas Schwarz (APS-Argonne)
- *A Digital Twin for In Silico Time-resolved Experiments*, PI - Subramanian Sankaranarayanan (CNM-Argonne), Maria Chan (CNM-Argonne), Mathew Cherukara (APS-Argonne), Pierre Darancet (CNM-Argonne), Ross Harder (APS-Argonne), Haidan Wen (APS-Argonne), Jianguo Wen (CNM-Argonne)
- *Actionable Information from Sensor to Data Center*, PI - Jana Thayer (Linac Coherent Light Source, SLAC National Accelerator Laboratory), Ian Foster, Zhengchun Liu (DSL-Argonne), Peter Kenesei, Antonino Miceli, Nicholas Schwarz (APS-Argonne)
- *Machine Learning for Autonomous Control of Accelerators*, PI - Daniel Ratner (SLAC National Accelerator Laboratory), Michael Borland (APS-Argonne)

- *Integrated Platform for Multimodal Data Capture, Exploration and Discovery Driven by AI Tools*, PI - Eli Stavitski (National Synchrotron Light Source II, Brookhaven National Laboratory) Chengjun Sun, Steve Heald, Nicholas Schwarz (APS-Argonne) Maria Chan (CNM-Argonne)

The APS receives funding from DOE for Advanced Scientific Computing Research for DOE User Facilities, Lab 23-3030:

- *X-ray & Neutron Scientific Center for Optimization, Prediction, & Experimentation (XSCOPE)*, PI - Sven Leyffer (MCS-Argonne), Ian Foster (DSL-Argonne), Nicholas Schwarz (APS-Argonne), et al.
- *ILLUMINE - Intelligent Learning for Light Source and Neutron Source User Measurements Including Navigation and Experiment Steering*, PI - Jana Thayer (LCLS), Stuart Campbell (NSLS-II), Alexander Hexemer (ALS), Nicholas Schwarz (APS-Argonne), Jonathan Taylor (SNS/HFIR), Vivek Thampy (SSRL)

Additionally, the APS receives funding from other DOE awards:

- *Randomized algorithms for optimal data acquisition in Bayesian inverse Problems*, PI - Youseff Marzouk (MIT), Zichao Wendy Di (MCS/XSD-Argonne)
- *Privacy-Preserving Federated Learning on Multimodal Data*, PI - Kibaek Kim (ANL/MCS), Zichao Wendy Di (MCS/XSD-Argonne)

The *Expand X-ray Capabilities with Extreme Light at APS (EXCEL@APS)* DOE MIE project proposes approximately $6.7M for advanced computing infrastructure at the APS. This project is currently in the DOE CD0 to CD1 phase.

Collaborations

Collaborations play a key role in the computing strategy for the APS. The APS actively collaborates with other facilities and organizations, and members of the APS User community to develop data analysis algorithms and software. As examples, most Argonne-funded LDRDs in this area involve collaborators from Argonne's Mathematics and Computer Science Division, Computational Science Division, or Data Science and Learning Division. Select APS User groups have contributed greatly to analysis algorithms and software.

The Center for Advanced Mathematics for Energy Research Applications (CAMERA) at Lawrence Berkeley National Laboratory aids in the development of software, modeling, and mathematics. For example, CAMERA helped develop GISAXS algorithms and tools, and the *SHARP* ptychographic reconstruction package. Most recently CAMERA has been involved in the development of the *XPCS-Eigen* correlation application for XPCS and in the application of the Multi-Tiered Iterative Phasing (M-TIP) algorithm for the reconstruction of Coherent Surface Scattering Imaging (CSSI) data. APS staff and researchers participate regularly in annual workshops for tomography, ptychography, and XPCS organized by CAMERA.

Innovative APS applications, improved Globus-based data management and transfer capabilities, and the Globus Compute platform has benefited, and continues to benefit, from ASCR support to Argonne research projects, such as RAMSES: Robust Analytical Models for Science at Extreme Scales and Braid: Data Flow Automation for Scalable and FAIR Science. Effort for CDI ptychography was initially funded by ASCR and then via Intelligence Advanced Research Projects Activity (IARPA) and Northwestern University. Early efforts for the *MIDAS* software for High-Energy Diffraction Microscopy (HEDM) data processing were funded by APS industrial partners. The APS and the National Synchrotron Light Source II (NSLS-II) at Brookhaven National Laboratory (BNL) have developed a comprehensive computing collaboration plan so as to best utilize our scare resources, especially related to expanding Bluesky use at the APS. Work on support for multi- and distributed-GPU N-dimensional complex FFTs is supported by NVIDIA and the ALCF.

The APS has been involved in the NOBUGS conference community and maintains active participation in the series of hack-a-thons organized by the Experimental Facilities Computing (ExFaC) Working Group.

Researchers at the APS and Argonne's DSL Division co-organize the annual Workshop on Extreme-Scale Experiment-in-the-Loop Computing (XLOOP) at SC, The International Conference on High Performance

Computing, Networking, Storage and Analysis. This workshop focuses on the intersection of large-scale experimental science from user facilities, such as the APS, with high-performance computing. A peer review process led by the workshop's program committee selects manuscripts for presentation. Accepted manuscripts are published by the IEEE or ACM. The program committee selects the recipient of the best paper award, and the workshop attendees selects the recipient of the best presentation award. The novel work presented during at this workshop will help the APS develop solutions critical to handling massive amounts of data generated during the APS-U era.

APS and Argonne scientists also co-chair several conferences and/or serve on program committees, including the Parallel and Distributed Algorithms for Data Science track at the IEEE International Parallel and Distributed Processing Symposium (IPDPS'23), IEEE Conference on Image Processing (ICIP), IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Computational Imaging (COIMG), Machine Learning for Scientific Imaging (MLSI), High Performance Computing for Imaging (HPCI'22) at Electronic Imaging and the Denver X-ray Conference. These conferences cover wide range of large-scale imaging and data science problems.

Scientists also organize AI/ML workshops and symposia at the annual Materials Science and Technology (MS&T) meeting, the International Materials Research Congress (IMRC) meeting, the Denver X-ray Conference (DXC), The Minerals, Metals and Materials (TMS) annual meeting, the American Crystallographic Association, and the International Union of Crystallography.

The APS and the computing divisions within Argonne's Computing, Environment, and Life Sciences (CELS) directorate hosted a series of town hall meetings in December 2020. Over 150 attendees participated from across Argonne. The goal is to develop a common vision for the future of APS computing within Argonne. Breakout sessions focused on new algorithm, math, and AI/ML, scalable software tools, workflow and orchestration, computing architecture, sustainable and discoverable data repositories, and networking.

The X-ray Science Division has organized the APS Scientific Computation Seminar Series since 2015. This seminar series focuses on scientific computation for APS experiments. The series focuses on advanced software and computing infrastructure for analysis, reduction, reconstruction, and simulation. It provides an opportunity to learn about state-of-the-art computational techniques and tools and how they are being applied to science at the APS.

The directors of the 5 BES funded light sources chartered a Data and Computing Working Group (also called the Light Source Data & Computing Steering Committee) in 2017. The role of the committee is to develop and maintain, with input from the directors, a strategic plan in computing and data. This is defined to include data acquisition, analysis, visualization and management, and the associated hardware and software infrastructure. The committee also advises and assists the directors in the coordination and execution of work in this area, consistent with that strategic plan, and is responsible for reporting and responding to charges, achieving consensus on paths forward, coordinating proposal submissions, and tracking funded activities. In 2024, this group expanded to include collaboration with the Spallation Neutron Source (SNS) and High Flux Isotope Reactor (HFIR) neutron source from Oak Ridge National Laboratory (ORNL).

The Light Source Data & Computing Steering Committee has developed a common vision for computing across the light sources, the Distributed Infrastructure for Scientific Computing for User Science (DISCUS), and a decade long roadmap to achieve the vision. This vision proposes a transformative computational fabric that covers the full lifecycle of data generated at the BES Light Sources to accelerate discovery and insight. See Figure 1-20.

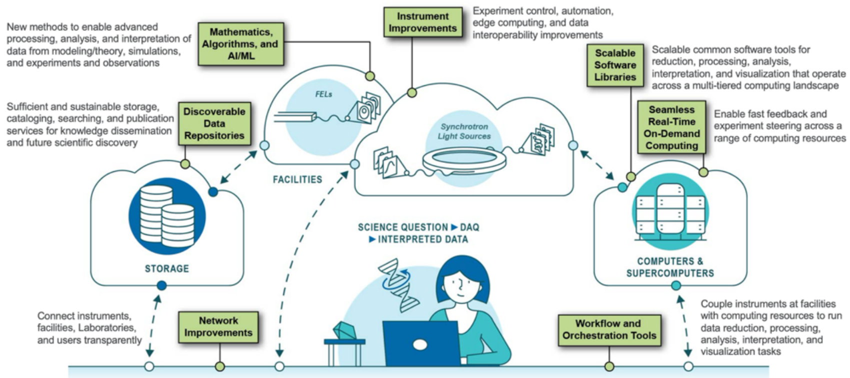


Figure 1-20 The Distributed Infrastructure for Scientific Computing for User Science (DISCUS) vision for computing at the light sources.

In 2019, the directors of the five BES funded light sources and the directors of the 4 ASCR computing and networking facilities charted the BES Light Source and ASCR Computing Facilities Directors' Data Working Group tasked with identifying how the ASCR facilities can help meet the needs of the BES facilities regarding data and computing. Membership is from the US DOE light sources and the US DOE supercomputing and networking facilities, and observers from the US neutron sources and Nano Science Research Centers (NSRCs) (ALS, APS, LCLS, MF, NSLS-II, SNS, SSRL, ALCF, ESnet, NERSC, and OLCF). The working group has formulated a plan for the desired data management architecture across the facilities, identified gaps in current planning, suggested a balance of responsibilities among the facilities, suggested next steps, and has undertaken pilot activities to utilize ASCR computing and networking facilities for processing and storing light source data. See Figure 1-21.

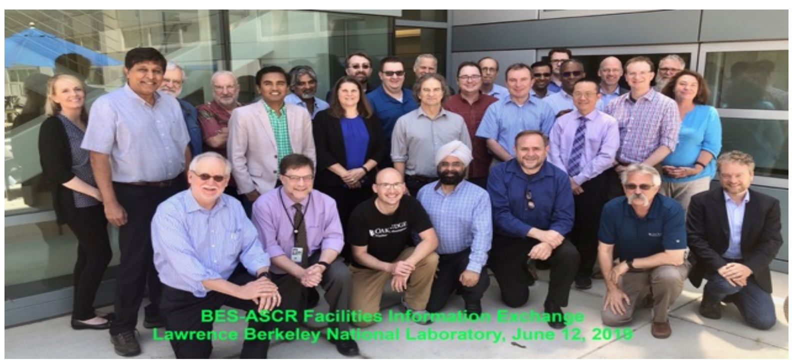


Figure 1-21 The BES-ASCR Facilities Information Exchange held at Lawrence Berkeley National Laboratory on June 12, 2019 established a working group across the BES light sources and the ASCR facilities.

The BES Data Solutions Task Force Pilot Project is a 2-year pilot project to develop common software for data acquisition, management, and analysis across the five BES light sources (ALS, APS, LCLS, NSLS-II, SSRL). The project aims at creating a synergistic approach to software where the five light sources work as a team to deliver common solutions across the facilities. This is being achieved by leveraging tools and expertise from all the BES light sources and integrating complementary components, including Bluesky from NSLS-II, *Xi-Cam* from CAMERA and ALS, and *XPCS-Eigen* and *TomoPy*, high-performance data processing software, from the APS. The project is focusing on X-ray Photon Correlation Spectroscopy (XPCS), ptychography, and tomography beamlines across the facilities. At the APS, Bluesky and *Xi-Cam* were successfully deployed at the 8-ID XPCS beamline.

Beginning in 2021, members of the APS and the other five BES light sources have participated in the DOE SC IRI Architecture Blueprint Activity (IRI ABA). This activity aims to produce the reference conceptual foundations to inform a coordinated whole-of-SC strategy for an integrative research ecosystem. The overarching motivation is to achieve a more seamlessly composable, interoperable, and extensible ecosystem of SC experimental and observational user facilities with SC advanced computing, data, and networking infrastructure. This ecosystem approach is critical to accomplishing many envisaged SC, DOE, and national R&D priorities such as AI for Science, Advanced Computing Ecosystem, National AI Research Resource, Future of Advanced Computing Ecosystem, Earthshots, and other initiatives and priorities. Participants worked to gather insight from across SC programs and design cross-cutting blueprints addressing SC needs.

The APS is part of a $10 million collaborative effort led by SLAC National Accelerator Laboratory, along with other DOE national labs: BNL, LBNL, and ORNL. The project is called Intelligent Learning for Light Source and Neutron Source User Measurements Including Navigation and Experiment Steering (ILLUMINE). It will focus on the testing, delivery and productive use of advanced computing methods and tools across DOE's x-ray and neutron sources.

Most recently, APS scientists have begun collaborating with Diamond Light Source on software for tomography reconstructions and AI/ML applications.

Tactics related to enhancing collaborations required to realize high-level goals:

1. *Leverage the 6-way collaboration among the DOE BES light and neutron sources to develop shared solution and approaches to common data and computing challenges.*
2. *Develop an engagement roadmap to integrate light source-enabled research into the DOE Frontiers in Artificial Intelligence for Science, Security, and Technology (FASST) initiative.*
3. *Plan for an integrated facility approach to leveraging computing and data resources as part of the DOE SC ASCR IRI program and HPDF project.*

Most recently, the DOE Frontiers in Artificial Intelligence for Science, Security, and Technology (FASST) initiative is beginning to take shape. This initiative leverages DOE's enabling infrastructure to deliver key assets for the national interest, including national security, attracting and building a talented workforce, harnessing AI for scientific discovery, addressing energy challenges, and developing technical expertise needed for AI governance. FASST is an exciting opportunity to harness the unique capabilities of the APS, the other light and neutron sources, the User community, and BES researchers with AI to unlock new science. The explosion of data coupled with AI creates opportunities to harness the value of data beyond individual experiments. Coupling AI/ML and experiment data with modeling/theory/simulation will enable new and more complex experiments. The APS be both contributors to and beneficiaries of advanced AI/ML efforts as both large data providers and as consumers of AI/ML output. To this end, the APS intends to align with goals and activities related to the FASST initiative.

The new DOE IRI program is being brought into formation. This new effort aims to provide researchers with seamless interoperability of DOE's unique data, user facilities, and computing resources. IRI is intended to be the infrastructure support, software, interface standards, and policies – that layers on top of existing DOE facilities making complex data-intensive workflows simple and fast for research teams. The APS continues to remain involved in IRI activities, and most recently has been selected to participate as an IRI Pathfinder project.

The HPDF will be a new DOE computing facility. Currently, in the early stages of planning, it is envisioned as a state-of-the-art resource for data science and research. The HPDF holds great potential to be a key enabler for the APS and its user community. The HPDF could provide seamless access to interoperable, scalable, and resilient computing resources: smaller-scale resources located at light source and neutron source Spokes, and larger-scale resources located at computing-facility Spokes and at the Hub(s), for real-time data processing, post experiment refinement, and simulation. Additionally, the HPDF could provide data catalogs and storage connected to computing resources for data management and retention and post experiment data processing for the User community, and especially aid AI/ML and digital twins efforts. The APS, and its light source and neutron source partners, continues to keep abreast of developments as HPDF planning progresses.

2 APS-U Feature Beamlines

Table 2-1 Summary of APS-U feature beamlines.

Feature Beamline	Synopsis
ATOMIC	Uses the enhanced coherence of the APS-U x-ray beam for high-resolution studies of the structural, chemical, and physical properties exhibited by advanced functional materials by acquiring atomistic structural information across many length scales in full three-dimensional detail.
Coherent High-Energy X-rays (CHEX)	Use coherent x-ray techniques to advance the frontier for in situ, real-time studies of advanced materials synthesis and chemical transformations in natural operating environments, employing condensed-matter physics and environmental science.
Coherent Surface-Scattering Imaging (CSSI)	Combines a surface X-ray probe using novel coherent scattering methods with state-of-the-art X-ray optics and detectors to study a range of materials surface and interface phenomena.
High-Energy X-ray Microscope (HEXM)	Investigates structure and evolution within bulk materials, often in extreme environments, with the established high-energy X-ray scattering techniques and novel coherence-based techniques enabled by APS-U.
In Situ Nanoprobe (ISN)	An x-ray nanoprobe designed to have a relatively large optical working distance enabling investigation of complex functional materials and materials systems such as catalysts, batteries, photovoltaic systems, and nanoscale Earth and environmental samples, during synthesis, operation, and under actual environmental conditions.
Polarization Modulation Spectroscopy (Polar)	Generates photon beams with highly tunable and modulated polarization states for imaging electronic and magnetic inhomogeneity in quantum materials with ~ 50 nm resolution as well as discovery of novel electronic states of matter at extreme pressure conditions ($P < 7$ Mbar).
PtychoProbe	Realizes the highest possible spatial-resolution X-ray microscopy both for structural and chemical information, with the goals of focusing an X-ray beam to a 5-nanometer spot and ultra-fast scanning of the beam across the sample being studied.
X-ray Photon Correlation Spectroscopy	Advances studies in physics and materials science and engineering including dynamic heterogeneity, structural dynamics in super-cooled liquids, and fluctuations associated with competing mesoscale interactions in emergent materials.
3D Micro and Nano (3DMN) Diffraction	Addresses a wide range of problems in materials science, physics, and geoscience by providing small, intense X-ray spots (between 50 and 200 nanometers) to investigate spatial variations and correlations of strain and structure that define a wide range of scientifically and technologically important materials.

2.1 ATOMIC APS-U Feature Beamline

2.1.1 Summary

The ATOMIC APS-U feature beamline will be dedicated to coherent x-ray diffraction imaging experiments for a diverse scientific community; experiments will exploit the brilliance of the upgraded source to study fundamental materials structures.

In the APS-U era, the ATOMIC APS-U feature beamline will perform Bragg CDI acquisitions in two modes: fast and high-resolution. Table 2-2 shows estimated data generation rates at the ATOMIC APS-U feature beamline, and current data rates at the 34-ID-C instrument, for comparison. The ATOMIC APS-U feature beamline is anticipated to collect approximately 250 to 300 TB of raw data per year, in comparison to approximately 0.65 TB of data collected today at the 34-ID-C Bragg CDI instrument. This represents a nearly 400x increase in data. These data generation estimates form the basis for networking infrastructure, controls, data management, and data processing planning.

Table 2-2 Data generation rates today at the 34-ID-C Bragg CDI instrument (for comparison) and estimated data generation rates at the ATOMIC APS-U feature beamline.

	Technique	Detector	Frame Size (MB)	Written Frame Rate (Hz)	Written Data Rate (MB/s)*	Daily Utilization (%)**	Data Set Size (MB)+	Data Per Day (GB)**	Annual Utilization (%)++	Data Per Year (TB) ‡
Today	Bragg CDI	ASI Si Timepix, ASI GaAs Timepix+++	0.25	0.2	0.05	80	60	3.38	90	0.65
APS-U Era++++	Fast Bragg CDI	TBD	1.00	20.0	20.00	80	500	1,350.00	80	220
	High-Resolution CDI	TBD	61.22	0.2	12.24	80	> 6,000	826.47	20	35

* The collection rate is high, but the frames are combined and written at a lower rate.
** Based on 1,440 minutes in one day.
\+ The data set sizes are approximate and representative of typical experiments, as this value varies.
++ Based on 210 days of beam time per fiscal year.
+++ The number of pixels for ASI Si Timepix is 65,536 and ASI GaAs Timepix is 262,144, however the frame size is typically cropped.
++++ The APS-U project has descoped certain parts of the ATOMIC Feature beamline, including detector purchases. Although the detectors listed in the table may not be purchased as a part of the APS-U project, this table represents the desired long-term potential capabilities intended for this beamline.

2.1.2 Network Architecture and Infrastructure

Network infrastructure needs will be addressed by the facility plan described in 1.2.

2.1.3 Controls, Data Acquisition, and Detector Integration

Controls, data acquisition, and detector integration needs will be addressed by the facility plan described in 1.3. Detailed plans and schedules are described in respective APS-U beamline controls documents, Equipment Specification Documents (ESDs), and Instrument Specification Documents (ISDs).

2.1.4 Data Management, Workflows, and Science Portals

The APS-U ATOMIC feature beamline will leverage the data management, workflow, and science portal efforts described in 1.4. The APS Data Management System, the facility-wide software and hardware system for managing data and workflows, will integrate data collection and processing workflows, manage user permissions based on experiment groups, and enable user access to data. Globus tools will provide a science portal for viewing and accessing data and automating the re-processing of data after the allotted experiment time has ended. For the ATOMIC APS-U feature beamline, workflows will provide a pipeline to automatically run tools to remove artifacts from data, reconstruct Bragg CDI data set, and view results.

2.1.5 Computing Infrastructure

Computing infrastructure needs will be addressed by the facility plan described in 1.5. Some local computing resources will be provided for on-the-fly data processing and experiment steering. Computing capacity for larger data processing tasks and for post-experiment processing and analysis will be provided by computing centers, including the ALCF and Argonne's Laboratory Computing Resource Center (LCRC). The APS Data Management System and Globus tools will be used to seamlessly integrate these resources.

2.1.6 Data Reduction and Analysis Software

The preliminary step is finding diffraction peaks from a crystal sample. This is accomplished by using the micro-diffraction technique and analyzing the captured data with the *LaueGo* software package. The data collected during this phase is only used to find the coordinates to guide the sample stage during data acquisition and is not retained.

The APS develops and supports the *cohere* software package for Bragg CDI data. The software is available as an open-source package (https://github.com/AdvancedPhotonSource/cohere). The package performs routine data correction, formatting, reconstruction, and visualization for Bragg CDI data. *cohere* currently implements conventional phase retrieval algorithms. It is written in Python and uses the ArrayFire package for data processing on CPUs and GPUs. Work is underway to add additional backend library choices so that ArrayFire can be replaced with CuPy or NumPy for easier distribution and deployment at computing centers.

The current feature set and performance of *cohere* is adequate for most of today's needs. However, the estimated approximate 400-fold increase in overall data that will be generated at the ATOMIC APS-U feature beamline, and the increase in size of individual datasets necessitates improvements and advances in software and algorithms.

The APS is currently developing higher-performance implementations of conventional phase retrieval algorithms and exploring novel AI/ML methods that may replace computationally complex phase retrieval methods. Table 2-3summarizes Bragg CDI data reduction needs, approaches, and status for the ATOMIC APS-U feature beamline.

AutoPhaseNN - Unsupervised Physics-aware Deep Learning of 3D Nanoscale Bragg Coherent Diffraction Imaging: A neural network that uses known physics and learns to invert 3D Bragg CDI data completely unsupervised, i.e., without ever being shown sample images during training, that is more than 10x faster than traditional methods, enabling real-time analysis, while also being more accurate. AutoPhaseNN has been integrated in to *cohere*, providing users the option of starting phase retrieval from an initial guess provided by AutoPhaseNN (see Figure 2-1).

The APS is optimizing implementations of conventional phase retrieval algorithms in *cohere* for better performance. In order to process data quickly in the APS-U era using conventional phase retrieval approaches, the APS is developing distributed-memory CPU and multi-GPU implementations of presently utilized algorithm. N-dimensional complex FFTs are at the core of conventional phase retrieval (and many ML) algorithms. The anticipated size of high-resolution Bragg CDI datasets and intermediate results will be too large to fit in the memory of a single GPU. The APS is working with the ALCF and a team at NVIDIA to realize better, optimized multi- and distributed-GPU support for N-dimensional complex FFTs.

Argonne researchers are exploring the use of AI and Automatic Differentiation (AD) as a high-performance alternative to conventional phase retrieval algorithms for Bragg CDI. This new workflow leverages a library of pre-computed, large-scale Molecular Dynamics (MD) simulations to provide on-the-fly, best guess structure to measured diffraction data through a trained deep convolutional neural network. Predictions are displayed in real-time at the instrument and are also used as the initial guess for iterative refinement through AD. This two-step approach will enable real-time feedback to an experiment and provide the highest possible fidelity in image reconstruction. A recently funded ALCF Expedition LDRD is focusing on using AI accelerators, such as SambaNova for Bragg CDI calculations.

Table 2-3 Summary of Bragg CDI data reduction needs, approaches, and status for the ATOMIC APS-U feature beamline.

Capability	Algorithm / Software Requirement	Status
Conventional Phase Retrieval Reconstructions	CPU and GPU software for Bragg CDI reconstructions	Done – APS Operations
Faster Conventional Phase Retrieval Reconstructions	Scalable distributed-memory CPU and GPU implementation of conventional phase retrieval algorithms	In Progress – APS Operations
High-Resolution Conventional Phase Retrieval Reconstructions	Support for multi- and distributed-GPU N-dimensional complex FFTs	In Progress – APS Operations working with the ALCF and NVIDIA
AI / Automatic Differentiation (AD) Methods	A deep learning (DL) approach to structure and strain prediction from raw X-ray diffraction data without the use of phase retrieval algorithms	Demonstrated at low resolution – LDRD
	A CNN training set generator and a trained CNN for the study of metals; this can grow to other advanced materials without changes to the underlying workflow	Demonstrated – LDRD
	Physics based image generation workflow installed at the CDI instrument to analyze coherent diffraction data in real-time	In Progress – LDRD

	Network optimization and combining deep learning with automatic differentiation to enable highest possible image reconstruction accuracy	Demonstrated – LDRD
	Scale to TB dataset sizes	To do – APS Operations

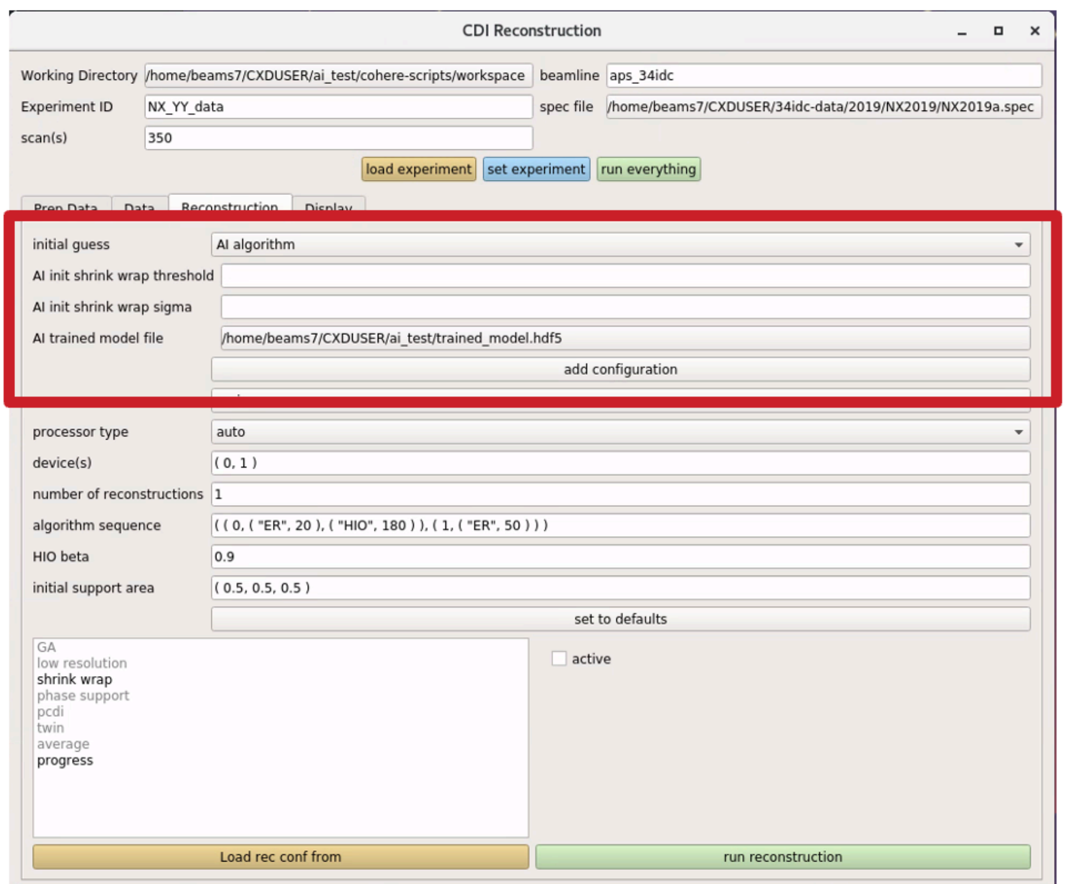
2.1.7 Effort, Funding, and Collaborations

Network infrastructure to the edge of the ATOMIC APS-U feature beamline will be provided by APS-U funding. All other network infrastructure is funded by APS Operations funding as described in 1.2. Facility-wide data management, workflows, and science portal effort is funded by APS Operations (approximately 2.5 FTE per year) and the Globus Services team as described in 1.4. Some local computing resources dedicated to the ATOMIC APS-U feature beamline for on-the-fly processing and experiment steering will be provided from APS-U funding. Large-scale computing resources are provided by APS Operations funding, ALCF funding, and Argonne Laboratory Computing Resource Center (LCRC) funding as described in 1.5. The APS dedicates approximately 1 FTE per year for Bragg CDI software development from APS Operations funding.

Work on support for multi- and distributed-GPU N-dimensional complex FFTs is supported by NVIDIA and the ALCF.

The following LDRD funding was awarded to support these efforts:

- *A.I. C.D.I.: Atomistically Informed Coherent Diffraction Imaging* (FY18)
- *Finding Critical Processes of Deformation in Structural Materials with Artificial Intelligence* (FY19)
- *Scalable DL-based 3D X-ray Nanoscale Imaging Enabled by AI Accelerators* (ALCF Expedition LDRD - FY21)



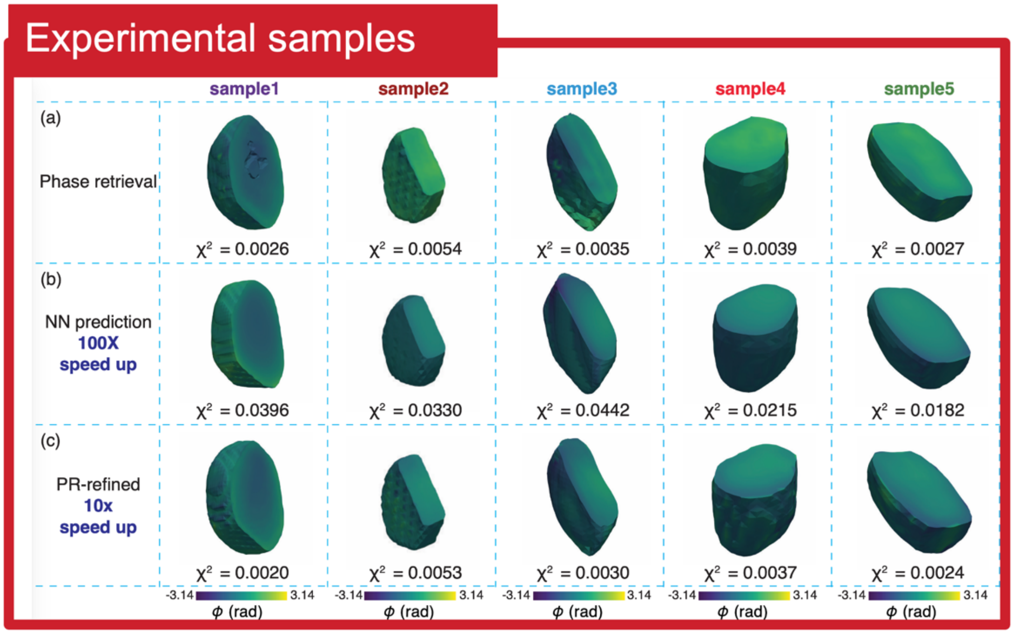


Figure 2-1 AutoPhaseNN within cohere - AutoPhaseNN provides an initial guess for iterative phase retrieval run through cohere. Results are obtained faster and are more accurate than conventional iterative phase retrieval methods.

2.2 Coherent High-Energy X-rays (CHEX) APS-U Feature Beamline

2.2.1 Summary

The Coherent High Energy X-rays (CHEX) APS-U feature sector will advance the frontier for in situ, real time studies of materials synthesis and chemical transformations in natural operating environments, using the unprecedented coherence of the high energy X-ray beams that will be provided by the upgrade. State-of-the-art experimental techniques that will be used, include, but are not limited to, Bragg coherent diffraction imaging (BCDI), Bragg ptychography, coherent Bragg rod analysis (COBRA), dark field X-ray microscopy (DFXM), and X-ray photon correlation spectroscopy (XPCS). These approaches will be used to provide transformative insight into materials structure, heterogeneity and disorder, chemical and long-range interactions, atomic-level dynamics, and structural, chemical, and morphological evolution under challenging environmental conditions and a wide range of time frames.

When fully built out, the CHEX sector will consist of four branch lines, with each line having two experimental stations. Two canted undulators will be used to operate one of the branches at tunable energies from 5-60 keV, while the other three branch lines will operate at fixed selectable energies of 15, 25, or 35 keV (D/E and F hutches) or 45, 75, or 105 keV (G hutch). The multiplexed nature of the design will allow up to four separate experiments to be performed simultaneously (one per branch line). The in-situ studies that will be performed at these stations are often data-intensive, due to desired frequent periodic monitoring of processes over long times, e.g., millisecond resolution over seconds, or second resolution over thousands of seconds. The current plans are to primarily use pixel array detectors with small (55-75 μm) pixel sizes, e.g., Lambda 750K or Eiger 1M. Under these conditions, a preliminary conservative estimate is that on the order of 1.5 TB of data can be generated per 24 hours of operation. As faster detectors with smaller pixel sizes become available, this data generation rate will increase correspondingly.

2.2.2 Network Architecture and Infrastructure

Network infrastructure needs will be addressed by the facility plan described in 1.2

2.2.3 Controls, Data Acquisition, and Detector Integration

Controls, data acquisition, and detector integration needs will be addressed by the facility plan described in 1.3. Detailed plans and schedules are described in respective APS-U beamline controls documents, Equipment Specification Documents (ESDs), and Instrument Specification Documents (ISDs).

2.2.4 Data Management, Workflows, and Science Portals

The CHEX APS-U feature beamline will leverage the data management, workflow, and science portal efforts described in 1.4. The APS Data Management System, the facility-wide software and hardware system for managing data and workflows, will integrate data collection and processing workflows, manage user permissions based on experiment groups, and enable user access to data. Globus tools will provide a science portal for viewing and accessing data and automating the re-processing of data after the allotted experiment time has ended. For the CHEX APS-U feature beamline, workflows will provide a pipeline to automatically run data processing software for preliminary analysis of coherent imaging and XPCS data, with a goal of providing near-real-time feedback useful in planning and modifying experiments.

2.2.5 Computing Infrastructure

Computing infrastructure needs will be addressed by the facility plan described in 1.5. Some local computing resources will be provided for on-the-fly data processing and experiment steering. Computing capacity for larger data processing tasks and for post-experiment processing and analysis will be provided by computing centers, including the ALCF and Argonne's Laboratory Computing Resource Center (LCRC). The APS Data Management System and Globus tools will be used to seamlessly integrate these resources.

2.2.6 Data Reduction and Analysis Software

The experimental techniques that will be used at CHEX, include, but are not limited to, Bragg coherent diffraction imaging (BCDI), Bragg ptychography, coherent Bragg rod analysis (COBRA), dark field X-ray microscopy (DFXM), and X-ray photon correlation spectroscopy (XPCS). Since it is anticipated that there will be at least an order-of-magnitude increase in data rates and volumes at the APS over the next decade, combined with continual rapid developments in these coherent imaging and spectroscopy, there is a great need for there to be a concomitant emphasis on the development of advanced data analysis approaches that will enable realization of the full potential of the CHEX beamlines to elucidate materials behavior. APS is currently devoting significant effort to developing data analysis packages relevant to experimental techniques of interest to the CHEX sector. For example:

- The *cohere* software package is being developed to provide tools for reconstruction of images from data obtained using Bragg Coherent Diffraction Imaging techniques
- The *MIDAS* software package is being developed to enable users to non-destructively image the microstructure of crystalline materials in 3D
- Development of the *tike* toolbox is enabling tomographic reconstruction of 3D objects from ptychography data

Implementations of the above software packages will be of great value to CHEX users. It is anticipated, however, that many planned experiments at CHEX will have specific, unique data analysis requirements that will require significant refinements/extensions of the above-mentioned packages and/or developments of new analysis approaches. Experiments with these unique needs will form the foundation of many of the initial science campaigns at CHEX and will serve as a backbone for future science at CHEX. For example, several planned early experiments at CHEX will have specific data analysis needs that due to unique experimental requirements at CHEX that require additional algorithm and software development, including:

- High-energy BCDI capabilities for imaging experiments via phase retrieval with high-energy coherent focused beams, possibly implemented in *cohere* or *MIDAS*

- High-q pixel mapping and two-time and higher-order correlation function incorporation into XPCS analysis packages requires significant new development, especially for experiments planned at CHEX that will target very weakly scattering surface-sensitive regions of q-space
- Support, in *tike*, for example, for ptychography in the Bragg geometry, and optimization for studies of the epitaxial films of interest to thin film synthesis programs that are planned at CHEX
- Software for COBRA and DFXM analyses are required as both of these techniques are anticipated to be employed in key roles in planned in-situ synthesis and materials processing experiments at CHEX.

2.2.7 Effort, Funding, and Collaborations

Network infrastructure to the edge of the CHEX APS-U feature beamline will be provided by APS-U funding. All other network infrastructure is funded by APS Operations funding as described in 1.2. Facility-wide data management, workflows, and science portal effort is funded by APS Operations (approximately 2.5 FTE per year) and the Globus Services team as described in 1.4. Some local computing resources dedicated to the CHEX APS-U feature beamline for on-the-fly processing and experiment steering will be provided from APS-U funding. Large-scale computing resources are provided by APS Operations funding, ALCF funding, and Argonne Laboratory Computing Resource Center (LCRC) funding as described in 1.5.

To accelerate progress in the development of the advanced data analysis tools to support state-of-the-art science at CHEX, a joint scientific staff appointment between the APS and MSD has been proposed, requesting 50% funding by APS Operations funds. Through this collaboration, a co-located staff member will both adopt the computational data science approaches of currently supported APS software packages to meet the specific needs of a broad range of CHEX users and be involved in the planning and execution of pioneering coherent imaging and XPCS materials science experiments at CHEX that will be carried out by the MSD Synchrotron Studies of Materials Group. Creation of this joint appointment will ensure a healthy data analysis ecosystem at CHEX and will simultaneously improve the breadth and impact of the suite of ongoing software package design by addressing and eliminating current key blind spots in software packages used for analysis of BCDI, XPCS, Bragg geometry ptychography, COBRA, and DFXM data.

2.3 Coherent Surface-Scattering Imaging (CSSI) APS-U Feature Beamline

2.3.1 Summary

The Coherent Surface-Scattering Imaging (CSSI) APS-U feature beamline will take advantage of the MBA lattice's dramatically improved x-ray beam coherence for probing and understanding mesoscopic structures and dynamics at surfaces and interfaces.

In the APS-U era, the CSSI APS-U feature beamline will employ two primary operation modes: Coherent Surface Scattering Imaging (CSSI) and Grazing-Incidence X-ray Scattering (GIXS). The latter includes Grazing-Incidence Wide-Angle X-ray Scattering (GIWAXS), GIWAXS with XPCS, Grazing-Incidence Small-Angle X-ray Scattering (GISAXS), and GISAXS with XPCS. In addition, to characterize fast kinetics across a broad range of length scales, a fast data acquisition mode will be provided where both GIWAXS and GISAXS detectors are operated at high frame rates for short periods. Table 2-4 shows estimated data generation rates at the CSSI APS-U feature beamline. The CSSI APS-U feature beamline is anticipated to collect approximately 17 PB raw data per year. These data generation estimates form the basis for networking infrastructure, controls, data management, and data processing planning.

Table 2-4 Estimated data generation rates at the CSSI APS-U feature beamline.

	Technique	Detector	Frame Size (MB)	Frames Per Dataset	Compression Factor*	Compressed Dataset Size (GB)	Daily Utilization (%)**	Data Per Day (TB)**	Annual Utilization (%)***	Data Per Year (PB)***

APS-U Era	GIWAXS+	Eiger 9M	33.750	40	1	1.3	13	3.5	38	0.27
	GIWAXS-XPCS+	Eiger 9M	33.750	120,000	10	197.8	25	69.5		5.42
	GISAXS+	Eiger 16M	61.035	40	1	2.4	13	6.3		0.49
	GISAXS-XPCS+	Eiger 16M (4M XPCS mode)	15.26	135,000	10	201.2	25	70.7		5.51
	Fast GIXS ++	Eiger 9M	33.750	30,000	1	988.8	14	38.9	8	0.64
		Eiger 16M	61.035	30,000	1	1788.1	14	70.4		1.16
	CSSI+++	Eiger 16M	61.035	60,000	10	357.6	80	40.2	46	3.80

* A compression factor of 1.0 represents no compression.

** Based on 1,440 minutes in one day. Routine GIXS: 2 min data out of 20 min/sample = 10%. Spin GIXS: 5 min data out of 40 min/sample = 12/5%.

*** Based on 210 days of beam time per fiscal year.

+ GIXS: Static simultaneous GIWAXS and GISAXS for 1 dataset, followed by 1 dataset simultaneous GIWAXS/GISAXS surface XPCS, assuming 5 min alignment.

++ Fast GIXS: Simultaneous GIWAXS/GISAXS for 5 minutes, assuming 30 minutes required to set up sample.

+++ CSSI: 20% of each day required for alignment and sample motions during scan.

2.3.2 Network Architecture and Infrastructure

Network infrastructure needs will be addressed by the facility plan described in 1.2

2.3.3 Controls, Data Acquisition, and Detector Integration

Controls, data acquisition, and detector integration needs will be addressed by the facility plan described in 1.3. Detailed plans and schedules are described in respective APS-U beamline controls documents, Equipment Specification Documents (ESDs), and Instrument Specification Documents (ISDs).

2.3.4 Data Management, Workflows, and Science Portals

The APS-U CSSI feature beamline will leverage the data management, workflow, and science portal efforts described in 1.4. The APS Data Management System, the facility-wide software and hardware system for managing data and workflows, will integrate data collection and processing workflows, manage user permissions based on experiment groups, and enable user access to data. Globus tools will provide a science portal for viewing and accessing data and automating the re-processing of data after the allotted experiment time has ended.

All operation modes of the APS-U CSSI feature beamline will generate data at high rates. The APS Data Management System will coordinate data transfer, data backup, preprocessing, and analysis, and provide visualization of analysis results to users.

2.3.5 Computing Infrastructure

Computing infrastructure needs will be addressed by the facility plan described in 1.5. Some local computing resources may be provided for on-the-fly data processing and experiment steering. The anticipated high data rate and large data volume generated by the CSSI beamline makes processing data likely beyond the capability of local workstations. Computing capacity for these data processing tasks and for post-experiment processing and analysis will be provided by computing centers, including the ALCF and Argonne's Laboratory Computing Resource Center (LCRC). To develop performant codes suitable for ALCF, a local workstation with four GPUs has been commissioned as a test bed. The APS Data Management System and Globus tools will be used to integrate these resources.

2.3.6 Data Reduction and Analysis Software

Coherent Surface-Scattering Imaging (CSSI) Data Processing

CSSI is a coherent imaging technique for creating quantitative 3D high resolution images of surface and interface structures. The method combines ideas from scanning coherent diffractive imaging (ptychography) and computed laminography to create very large diffraction datasets. The process here is done in two steps: the datasets for each laminographic rotation angle must fed into ptychographic reconstruction algorithms to create various 2D projection images of the overall 3D sample structure. These 2D projection images will then be combined in a computed laminography algorithm to synthesize the final 3D sample image.

Conventional transmission-based ptychography and laminography algorithms for data inversion must be modified to account for CSSI's geometry and the multiple-scattering effects that are significant at low incidence angles. One particular laminography challenge here is due to the very shallow incident angle encountered in CSSI, the “missing cone” problem can be quite severe. Novel variations of iterative constrained laminography must be developed using multiple GPUs due to the large 3D sample arrays used in CSSI (which are too large for a single GPU to contain). Additionally, we can also vary the incident angle and perform further laminography data collection, but this obviously further greatly increases the data volume that must be computationally processed.

For the ptychography step, one must fundamentally change the forward model to account for non-kinematical scattering. Recent investigations have determined that a multislice Fresnel propagation approach is the most accurate model here. One challenge encountered with attempting to simply extend existing ptychography inversion algorithms like rPIE is that the number of slices in the multislice approach cannot be too large otherwise multiplicative error propagation becomes too dominant. We can somewhat “ignore” the multiple-scattering effects with careful selection of appropriate regions of the diffraction measurements, but this ultimately amounts to only reconstructing lower resolution 2D projection images; to achieve the best possible spatial resolution other ptychographic multislice methods are necessary.

Another option here is to move beyond this “two-step” ptychography-laminography process and combine both into a single unified “one-step” 3D ptychographic phase retrieval problem. In this way, the Fresnel multislice model (which inherently represents a complex valued probing wavefield interacting with and propagating through a 3D sample volume) will be used in a 3D ptychography algorithm where the laminographic rotation is treated as a type of “rotational” diversity which is used as a constraint alongside the usual 2D translational diversity. The mathematics of the numerical nonlinear optimization problem have already been established and demonstrated on small dimensionality problems to establish proof of principle; the work that must be done here is to scale this algorithm up for realistic CSSI problem sizes for computation on multiple GPUs.

Additionally, a Multi-Tiered Iterative Phasing (M-TIP) approach to decompose the larger problem into smaller solvable parts is being developed in collaboration with CAMERA.

GIXS and GIXS-XPCS Data Processing

Today, GIXS data analysis is performed with the APS developed and supported MATLAB package, GIXSGUI, and the CAMERA developed Python-based package, Xi-CAM, for long sequences of time-resolved measurements. The integration of GIXSGUI with the APS Data Management System to automate data reduction and analysis at the 8-ID-E beamline is underway. High-performance algorithms for near real-time GIXS data processing will be implemented.

At CSSI, the data production rate will be many orders-of-magnitude higher. Multiple-detector collection modes will be routine, adding further complexity for data processing. These challenges necessitate large volume and multiple-dimension real-time data visualization. The APS Data Management System will accommodate increased data volumes. Work will be undertaken to replace the current MATLAB-based tools with a new higher-performance and scalable Python-based toolkit. Thin-film structure peak indexing capabilities will also be improved.

For XPCS data reduction and analysis, CSSI will leverage the resources and tools available and being developed for the XPCS beamline.

Table 2-5 Summary CSSI APS-U feature beamline data reduction and processing capabilities and needs.

Capability	Algorithm / Software Requirement	Status

Data Visualization and Preprocessing	Single image	Done – *GIXSGUI* – APS Operations
	Multiple images	Done – *Xi-Cam* – CAMERA
	Support for scattering vector *q*	To do – APS Operations
	New Python-based software package	To do – APS Operations
	Near real-time processing	To do – APS Operations
Thin-film Structure Indexing	Basic implementation of space groups and indexing	Done – APS Operations developed *GIXSGUI,* CAMERA developed *Xi-Cam,* and SSRL developed a thin-film structural indexing package *SIIRkit*.
	New scalable Python-based software package that integrates surface scattering (Distorted Wave Born Approximation)	To do – APS Operations
Coherent Surface Scattering Imaging (CSSI)	Image reconstruction algorithms	Done – APS Operations & DOE Early Career Award developed an algorithm to reconstruct CSSI ptychography data
		In Progress – CAMERA is developing an M-TIP based CSSI reconstruction algorithm
	Scalable CPU and GPU software	In Progress – APS Operations & DOE Early Career Award is developing a multi-GPU CSSI ptychography reconstruction software package
Surface XPCS	XPCS correlation algorithms and software	In Progress – APS Operations work is underway as part of effort for the XPCS APS-U feature beamline

2.3.7 Effort, Funding, and Collaborations

Network infrastructure to the edge of the CSSI APS-U feature beamline will be provided by APS-U funding. All other network infrastructure is funded by APS Operations funding as described in 1.2. Facility-wide data management, workflows, and science portal effort is funded by APS Operations (approximately 2.5 FTE per year) and the Globus Services team as described in 1.4. Some local computing resources dedicated to the CSSI APS-U feature beamline for on-the-fly processing and experiment steering will be provided from APS-U funding. Large-scale computing resources are provided by APS Operations funding, ALCF funding, and Argonne Laboratory Computing Resource Center (LCRC) funding as described in 1.5. The APS dedicates approximately 1.5 FTE per year for CSSI related algorithm and software development from APS Operations funding.

CAMERA provides effort in support of CSSI algorithm development.

The following awards support these efforts:

- *Unraveling Mesoscale Spatial-temporal Correlations in Materials Using Coherent X-ray Probes* (FY15 LDRD)
- *Developing Advanced Coherent Surface Scattering Reconstruction Method Incorporating Dynamical Scattering Theory* (FY17 LDRD)
- *Development of Coherent Surface Scattering Imaging with Nanometer Resolution for Revealing 3D Mesoscaled Structures* (DOE Early Career Award)

2.4 High-Energy X-ray Microscope (HEXM) APS-U Feature Beamline

2.4.1 Summary

The High-Energy X-ray Microscope (HEXM) APS-U feature beamline is designed to investigate structure and evolution within bulk materials, often in extreme environments, with established high-energy x-ray scattering techniques and novel coherence-based techniques enabled by the APS-U.

Table 2-6 shows estimated data generation rates at the HEXM APS-U feature beamline. The HEXM instrument will perform near- and far-field, diffraction tomography, and imaging tomography measurements. The HEXM APS-U feature beamline is anticipated to collect approximately 20 PB of raw data per year and 5 PB of compressed raw data per year in comparison to approximately 4 PB of raw data and approximately 1 PB of compressed raw data collected today at the 1-ID High-Energy Diffraction Microscopy (HEDM) instrument. Both uncompressed and compressed data sizes are given because uncompressed data is often required for data processing. This represents an approximately 4x increase in data. These data generation estimates form the basis for networking infrastructure, controls, data management, and data processing planning.

Table 2-6 Data generation rates today at the 1-ID High-Energy Diffraction Microscopy (HEDM) instrument (for comparison) and estimated data generation rates at the HEXM APS-U feature beamline.

	Technique	Detector	Frame Size (MB)	Average Frame Rate (Hz)	Peak Data Rate (MB/s)	Raw Data set Size (GB)	Compression Factor *	Daily Utilization (%)**	Raw Data Per Day (TB)**	Compressed Raw Data Per Day (TB)**	Annual Utilization (%) ***	Raw Data Per Year (TB) ***	Compressed Raw Data Per Year (TB) ***
Today	Near-Field	Qimaging Retiga 4000DC (4MP, CCD, 12-bit)	8	3.3	26	11	4	54	1	0.25	25	62	15
	Far-Field	Varex 4343CT (8MP, 14-16-bit)	16	15	237	22	4	63	12	3	25	647	162
	Far-Field	GE RT41 (4MP, 14-bit)	8	7	56	11	2	63	3	1.5	25	153	76
	Far-Field	Hydra 4x GE RT41 (16MP, 14-bit)	32	7	224	45	2	63	12	6	15	366	183
	Far-Field	Pilatus 2M CdTe (20-bit)	7	250	1,771	10	4	63	92	23	15	2,896	724
	Diffraction Tomography	GE RT41 (4MP, 14-bit)	8	2	16	281	2	72	1	0.5	5	10	5
	Diffraction Tomography	Hydra 4x GE RT41 (16MP, 14-bit)	32	2	64	1,125	2	72	4	2	5	40	20
	Diffraction Tomography	GE RT41 (4MP, 14-bit)	8	7	56	2,250	2	72	3	1.66	5	35	17
	Diffraction Tomography	Hydra 4x GE RT41 (16MP, 14-bit)	32	7	224	9,000	2	72	13	6.5	5	140	70
	Imaging Tomography	PointGrey CMOS (2.3MP, 12-bit)	4	5	22	15	1	54	1	1	20	41	41
APS-U Era	Near-Field	FLIR Oryx (5MP, 12-bit)	10	40	383	13	4	54	17	4	25	894	223
	Far-Field	Dectris Pilatus 6M (20-bit)	18	125	2,226	25	4	63	116	29	40	9,706	2,426
	Diffraction Tomography	Dectris Pilatus 2M CdTe (20-bit)	7	50	354	249	4	72	21	5	10	441	110
	Diffraction Tomography	Dectris Eiger 16M CdTe (12-bit)	52	50	2,595	1,825	4	72	154	39	10	3,233	808
	Diffraction Tomography	Sydor SMM-PAD CdTe (22-bit)	0.75	50	38	26.37	4	72	2.22	0.56	10	47	12
	Diffraction Tomography	Dectris Pilatus 2M CdTe (20-bit)	7	250	1,771	1,993	4	72	105	26	5	1,103	276
	Diffraction Tomography	Dectris Eiger 16M CdTe (12-bit)	52	133	6,902	14,596	4	72	410	102	5	4,300	1,075
	Diffraction Tomography	Sydor SMM-PAD CdTe (22-bit)	0.75	1,000	750	211	4	72	45	11	5	467	117
	Imaging Tomography	FLIR Oryx 5MP (12-bit)	10	10	96	34	1	54	4	4	15	134	134
	Fast Imaging Tomography	FLIR Oryx 5MP (12-bit)	10	100	956	34	1	12	10	10	5	99	99

* A compression factor of 1.0 represents no compression.
** Based on 1,440 minutes in one day. Utilization reflects the overhead associated with detector duty cycles and motion, as well as related setup time for alignment, calibration, sample changes and sample alignment, in situ environment modification, etc.
*** Based on 210 days of beam time per fiscal year.

2.4.2 Network Architecture and Infrastructure

Network infrastructure needs will be addressed by the facility plan described in 1.2.

2.4.3 Controls, Data Acquisition, and Detector Integration

Controls, data acquisition, and detector integration needs will be addressed by the facility plan described in 1.3. Detailed plans and schedules are described in respective APS-U beamline controls documents, Equipment Specification Documents (ESDs), and Instrument Specification Documents (ISDs).

2.4.4 Data Management, Workflows, and Science Portals

The HEXM APS-U feature beamline will leverage the data management, workflow, and science portal efforts described in 1.4. The APS Data Management System, the facility-wide software and hardware system for managing data and workflows, will integrate data collection and processing workflows, manage user permissions based on experiment groups, and enable user access to data. Globus tools will provide a science portal for viewing and accessing data and automating the re-processing of data after the allotted experiment time has ended. For the HEXM APS-U feature beamline, workflows will provide a pipeline to automatically run data processing software for near- and far-field, diffraction tomography, and imaging tomography data reconstructions, and to view results.

2.4.5 Computing Infrastructure

Computing infrastructure needs will be addressed by the facility plan described in 1.5. Some local computing resources will be provided for on-the-fly data processing and experiment steering. Computing capacity for larger data processing tasks and for post-experiment processing and analysis will be provided by computing centers, including the ALCF and Argonne's Laboratory Computing Resource Center (LCRC). The APS Data Management System and Globus tools will be used to seamlessly integrate these resources.

2.4.6 Data Reduction and Analysis Software

The HEXM APS-U feature beamline requires data processing algorithms and software for near- and far-field, diffraction tomography, and imaging tomography measurements. Descriptions of current efforts and plans in each of these areas follow. Table 2-7 summarizes capabilities for each of these modes, respectively.

Near- and Far-Field Diffraction Microscopy

The APS develops and supports the *Microstructural Imaging using Diffraction Analysis Software (MIDAS)* software package for near- and far-field diffraction microscopy data processing. The software is available as an open-source package (https://github.com/marinerhemant/MIDAS). *MIDAS* is written in C and uses Python for scripting. Distributed-memory parallelization is achieved using the SWIFT parallel execution framework. Time-critical parts of the code have been ported to run on GPUs with CUDA. *MIDAS* has been demonstrated to scale to tens-of-thousands of cores on supercomputers at the ALCF and at the National Energy Research Scientific Computing Center (NERSC). An average size data set today is typically processed within a few minutes. The APS will continue to develop *MIDAS* by enabling the processing of data taken with 3D scans, implementing intensity fitting, closely integrating with tomographic reconstruction algorithms and software, and scaling and optimizing performance to support APS-U Era data rates and sizes.

In addition to APS developed software, *IceNine* supports processing near-field diffraction data, and *Fable* and *HEXRD* support far-field data processing.

Diffraction Tomography

APS staff in the Materials Physics & Engineering group have developed prototype software in MATLAB to reconstruct diffraction tomography data. This prototype MATLAB software serves as a proof-of-principle for algorithm quality. The APS will develop production ready, higher-performance software for diffraction tomography reconstructions, for APS-U Era data. Algorithmic work will continue to integrate more advanced algorithms uses for imaging tomography, such as Algebraic Reconstruction Technique (ART) based reconstruction methods.

As part of the *High Pressure Material Characterization in 3-Dimensions Using X-ray Diffraction Contrast Computed Tomography* LDRD, modules have been added in MIDAS for rapid and automated analysis of diffraction tomography data. These include rapid correction and transformation of diffraction data, extraction of peak properties and tomographic inversion.

Imaging Tomography

The APS develops and supports the *TomoPy* tomographic reconstruction library. *TomoPy* is available as an open-source library (https://github.com/tomopy/tomopy). *TomoPy* is primarily written in Python and has integrated MPI-based and GPU-based routines for performance. Reconstruction algorithms in TomoPy have been scaled to run on supercomputers at the ALCF, the National Energy Research Scientific Computing Center (NERSC), and the Oak Ridge Leadership Computing Facility. In the APS-U Era, close integration of tomography reconstruction algorithms with *MIDAS* will improve performance and add convenience for users.

AI/ML Developments for the HEXM APS-U Feature Beamline

Researchers at the APS and from Argonne's DSL division have developed a deep neural network called BraggNN. This method enables extraction of precise Bragg peak locations from far-field High-Energy Diffraction Microscopy (HEDM) data. The model runs more than 200 times faster than the conventional pseudo-Voigt profiling to locate Bragg peak position (see Figure 2-2).

The APS is researching Point Focused High-Energy Diffraction Microscopy (PF-HEDM) as a technique that pushes the limits of HEDM techniques to smaller grains to obtain sub-granular information. Preliminary algorithms developed using tomography-like reconstructions are promising. Researchers are developing inversion tools using AI/ML to improve the quality of reconstructions and obtain higher-quality answers. These developments may provide an alternative to conventional diffraction tomography methods. The *Developing Point-Focus High-Energy Diffraction Microscopy to Reveal Battery Material Degradation* LDRD has been funded in FY23 for development of the PF-HEDM technique and leverage the beam capabilities of the APS-U era.

Table 2-7 Summary of near- and far field-diffraction, diffraction tomography, and imaging tomography data processing needs and status for the HEXM APS-U feature beamline.

Science Driver	Capability	Algorithm / Software Requirement	Status
Smaller grains Greater dispersity Higher deformation	Near-Field Diffraction	Scalable distributed-memory CPU and GPU implementation	Done – APS Operations – *MIDAS*
		Intensity fitting	To do – APS Operations
Smaller grains Greater dispersity Higher deformation	Far-Field Diffraction	Scalable distributed-memory CPU and GPU implementation	Done – APS Operations – *MIDAS*
		Multi-panel support	Done – APS Operations – *MIDAS*
		3D scanning support	To do – APS Operations
Nano-grains Amorphous materials	Diffraction (Scattering) Tomography	Prototype implementation	Done – APS Operations - MATLAB
		Scalable distributed-memory CPU and GPU implementation	Done – LDRD – *MIDAS*
		Integrate more advanced tomographic reconstruction algorithms, e.g., ART	To do – If needed – APS Operations
Faster processes (sub-second)	Imaging Tomography	Scalable distributed-memory CPU and GPU implementation	Done – APS Operations – *TomoPy*
		Integration with *MIDAS*	To do – APS Operations

2.4.7 Effort, Funding, and Collaborations

Network infrastructure to the edge of the HEXM APS-U feature beamline will be provided by APS-U funding. All other network infrastructure is funded by APS Operations funding as described in 1.2. Facility-wide data management, workflows, and science portal effort is funded by APS Operations (approximately 2.5 FTE per year) and the Globus Services team as described in 1.4. Some local computing resources dedicated to the HEXM APS-U feature beamline for on-the-fly processing and experiment steering will be provided from APS-U funding. Large-scale computing resources are provided by APS Operations funding, ALCF funding, and Argonne Laboratory Computing Resource Center (LCRC) funding as described in 1.5. The APS dedicates approximately 1 FTE per year for HEDM software development and approximately 1 FTE per year for *TomoPy* development from APS Operations funding.

In addition to APS Operations funding, the APS benefits from long-term collaborations with the Air Force Research Laboratory (AFRL), Carnegie Mellon University (CMU), in particular an NSF-MRI grant to CMU supported the development of a new APS High-Throughput High-Energy Diffraction Microscopy (HEDM) beamline at 6-ID-D, and past and future industrial partnerships with GE and Pratt & Whitney.

AI/ML BraggNN work is funded by *Information from Sensor to Data Center* (PI: Jana Thayer, SLAC National Accelerator Laboratory, LAB 20-2261).

The following LDRD funding was awarded to support these efforts:

- *Finding Critical Processes of Deformation in Structural Materials with Artificial Intelligence* (FY19)
- *AI-steer: AI-driven Online Steering of Light Source Experiments* (FY20)
- *High Pressure Material Characterization in 3-Dimensions Using X-ray Diffraction Contrast Computed Tomography (FY21)*
- *Developing Point-Focus High-Energy Diffraction Microscopy to Reveal Battery Material Degradation (FY23)*

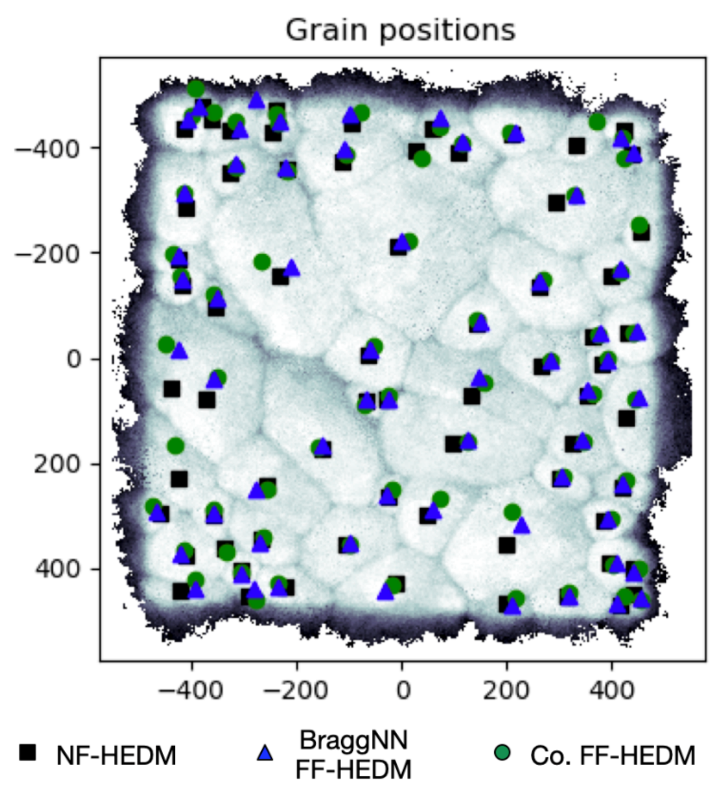


Figure 2-2 BraggNN: Grain center positions in microns determined by three methods, with a full high-resolution grain map from Near-Field HEDM superimposed in background. The Near-Field HEDM results provide the highest accuracy against which the grain-averaged Far-Field HEDM results can be compared. On average BraggNN provided slightly smaller position error than the conventional method.

2.5 In Situ Nanoprobe (ISN) APS-U Feature Beamline

2.5.1 Summary

The In Situ Nanoprobe (ISN) APS-U feature beamline is designed to study advanced materials during fabrication and operation. Its large working distance enables broad in situ environments, including heating, cooling, flow of process gases and fluids, and application of electric fields. The ISN beamline takes advantage of the upgraded source and is ideally suited for applications requiring diffraction-limited focusing. The ISN instrument will be a scanning nanoprobe, with x-ray fluorescence (XRF) detection and ptychography as major contrast modes. A secondary area detector will collect diffracted x-rays and provide some capability to identify local crystalline states.

Table 2-8 shows estimated data generation rates at the ISN APS-U feature beamline, and current data rates at the 2-ID-D ptychography, 2-ID-E XRF, and Bio Nano-Probe (BNP) XRF instruments for comparison. The ISN APS-U feature beamline is anticipated to collect approximately 10 PB of raw data per year and 1 PB of compressed raw data per year, in comparison to approximately 730 TB of raw data and approximately 73 TB of compressed raw data collected today across the 2-ID-D ptychography, 2-ID-E XRF, and Bio Nano-Probe (BNP) XRF instruments. Both uncompressed and compressed data sizes are given because uncompressed data is often required for data processing. This represents a nearly 15x increase in data. These data generation estimates form the basis for networking infrastructure, controls, data management, and data processing planning.

Table 2-8 Data generation rates today at the 2-ID-D ptychography and diffraction, 2-ID-E XRF, and Bio Nano-Probe (BNP) XRF instruments (for comparison) and estimated data generation rates at the ISN APS-U feature beamline.

	Technique	**Detector**	**Frame Size (MB)**	**Average Frame Rate (Hz)**	**Data Rate (MB/s)**	**Raw Data set Size (GB)**	**Compression Factor***	**Daily Utilization (%)****	**Raw Data Per Day (GB)****	**Compressed Raw Data Per Day (GB)****	**Annual Utilization (%)*****	**Raw Data Per Year (TB)*****	**Compressed Raw Data Per Year (TB)*****
Today	2-ID-D Ptychography	Dectris Eiger 500K	1.010	100	100	40	10	80	6,912	691	50	726	72.6
	2-ID-E XRF	Vortex ME4	0.008	20	0.15	1.91	5	100	13.18	2.64	80	2.16	0.43
	BNP XRF	Vortex ME4	0.008	20	0.15	0.69	15	100	13.18	0.88	80	2.16	0.14
APS-U Era	ISN XRF	2 X Vortex ME7	0.1484375	1,000	144.95	579.83	5	80	1,956	116	80	329	65
	ISN Ptychography	Dectris Eiger 1M	2.092	3,000	5,000	204,322	10	35	185,362	18,536	20	7,603	760
	ISN Diffraction	Dectris Eiger 1M	2.092	3,000	5,000	40.86	10	10	52,960	5,296	20	3,528	326

* A compression factor of 1.0 represents no compression.
** Based on 1,440 minutes in one day.
*** Based on 210 days of beam time per fiscal year.

2.5.2 Network Architecture and Infrastructure

Network infrastructure needs will be addressed by the facility plan described in 1.2.

2.5.3 Controls, Data Acquisition, and Detector Integration

Controls, data acquisition, and detector integration needs will be addressed by the facility plan described in 1.3. Detailed plans and schedules are described in respective APS-U beamline controls documents, Equipment Specification Documents (ESDs), and Instrument Specification Documents (ISDs).

2.5.4 Data Management, Workflows, and Science Portals

The ISN APS-U feature beamline will leverage the data management, workflow, and science portal efforts described in 1.4. The APS Data Management System, the facility-wide software and hardware system for managing data and workflows, will integrate data collection and processing workflows, manage user permissions based on experiment groups, and enable user access to data. Globus tools will provide a science portal for viewing and accessing data and automating the re-processing of data after the allotted experiment time has ended. For the ISN APS-U feature beamline, workflows will provide a pipeline to automatically run tools to remove artifacts from data, reconstruct the XRF, Ptychography, and Diffraction data set, and view results.

2.5.5 Computing Infrastructure

Computing infrastructure needs will be addressed by the facility plan described in 1.5. Some local computing resources will be provided for on-the-fly data processing and experiment steering. Computing capacity for larger data processing tasks and for post-experiment processing and analysis will be provided by computing centers, including the ALCF and Argonne's Laboratory Computing Resource Center (LCRC). The APS Data Management System and Globus tools will be used to seamlessly integrate these resources.

2.5.6 Data Reduction and Analysis Software

The ISN APS-U feature beamline requires three modes of data processing: elemental fitting for XRF microscopy data, ptychography data reconstruction, and space-mapping for diffraction data. Descriptions of current efforts and plans in each of these areas follow. Table 2-9, Table 2-10, and Table 2-11 summarize capabilities for each of these three modes, respectively. Many of the data processing requirements for the ISN APS-U feature beamline are like those of the PtychoProbe APS-U feature beamline described in 2.7.

Elemental Fitting for XRF Microscopy Data

The APS develops and supports the *XRF-Maps and uProbeX* software packages for XRF microscopy data processing and visualization (see Figure 2-3). This software is available as open-source packages (https://github.com/AdvancedPhotonSource/XRF-Maps and https://github.com/AdvancedPhotonSource/uProbeX). The *XRF-Maps* package performs elemental map fitting and the *uProbeX* application is a GUI for visualizing *XRF-Maps* results. *XRF-Maps* and *uProbeX* are both written in C++. *XRF-Maps* supports multi-core data processing in a shared-memory CPU environment and has a Python wrapper which allows all the functionality to be called from a Python environment.

APS-U enhancements will allow for larger scan areas resulting in larger datasets. These larger datasets may not be able to fit in system memory. To accommodate this *XRF-Maps* implements a streaming architecture that allows processing a dataset spectra by spectra without having to load the entire dataset. Only a limited number of spectra are loaded based on memory limits, processed, and saved to an HDF5 file until the whole dataset is processed. As data sizes increase, it may be become necessary to develop GPU-based and distributed-memory CPU- and GPU-based elemental fitting software.

The higher intensity x-ray beam generated by the APS-U storage ring necessitates the use of self-absorption correction when generating elemental maps. APS researchers and instrument staff are working on developing new self-absorption correction algorithms in collaboration with staff at the National Synchrotron Light Source II (NSLS-II) at Brookhaven National Laboratory (BNL). These algorithms are being implemented and tested in the *XRF-Maps* software.

Table 2-9 Summary of XRF microscopy elemental mapping data processing needs and status for the ISN APS-U feature beamline.

Capability	Algorithm / Software Requirement	Status
XRF Elemental Map Fitting	Algorithms for elemental map fitting	Done
	Multi-core shared-memory CPU implementation	Done – APS Operations

	Streaming data processing / operate on out-of-core data	Done – APS Operations
	Distributed-memory CPU and GPU implementation	To do – If required
XRF Self-Absorption Correction	Self-absorption correction algorithm development	In Progress – APS Operations and collaborations with NSLS-II
	Self-absorption correction implementation in *XRF-Maps*	In Progress – APS Operations

Ptychography Reconstruction

Ptychography has emerged as a powerful technique at synchrotron light sources. It will play a central role in answering many emerging scientific questions that the upgraded APS will help solve. Advanced ptychographic reconstruction algorithms and software are critical to take advantage of this new and innovative technique.

Multiple ptychographic reconstruction algorithms are required to achieve reasonable reconstruction quality to best analyze ptychography data collected for different domains and of varying sample characteristics. The APS has implemented the extended Ptychographic Iterative Engine (ePIE), regularized Ptychographic Iterative Engine (rPIE), conjugate gradient, Difference Map (DM), and iterative Least-SQuares solver for generalized Maximum-Likelihood (LSQ-ML) methods. Algorithms to help improve reconstruction quality, such as position and probe variation correction, and affine position regularization, are being developed and implemented.

Due to the computationally complex nature of ptychographic reconstruction algorithms and due to the anticipated increase in data rates and sizes in the APS-U Era, distributed high-performance implementations of ptychography reconstruction software are required. The APS with collaborators in Argonne's Mathematics & Computer Science (MCS) division developed *PtychoLib*, a distributed-memory GPU implementation of the extended Ptychographic Iterative Engine (ePIE) in 2014 and integrated Difference Map (DM) algorithms in 2018. *PtychoLib* was written in C++ and uses MPI and CUDA. The software was shown to scale on up to 256 GPUs on the ALCF's Cooley GPU cluster. This software has been supported and extended since then and has been the main tool used for high-performance ptychography reconstructions at APS beamlines. *PtychoLib* has been the main tool used for high-performance ptychography reconstructions at APS beamlines for the past decade. *PtychoPy* (https://github.com/kyuepublic/ptychopy) was developed as a Python wrapper and GUI for *PtychoLib*. Since then, the APS has consolidated ptychography development into the *tike* (https://github.com/AdvancedPhotonSource/tike) toolkit in order to make installing and developing new ptychography features and algorithms easier. This toolkit is written in Python and uses CuPy as the underlying GPU framework. All the reconstruction features of *PtychoLib* have been reimplemented in the tike toolkit including MPI and thread-based parallelism. *Ptychodus*, is a new pyQT-based GUI/workflow manager for ptychography reconstruction workflows has also been created in order to provide live reconstruction visualization and analysis.

Table 2-10 Summary of ptychography reconstruction needs and status for the ISN APS-U feature beamline.

Capability	Algorithm / Software Requirement	Status
Conventional Reconstruction	GPU implementation of extended Ptychographic Iterative Engine (ePIE) method	Done – APS Operations
	GPU implementation of Difference Map (DM) method	Done – APS Operations
	GPU implementation of the iterative Least-SQuares solver for generalized Maximum-Likelihood (LSQ-ML) method	Done – APS Operations
Improved Reconstruction Quality	Position correction	In Progress – Implemented in *tike* and currently being tested – APS Operations
	Probe variation correction	In Progress – Implemented in *tike* and currently being tested – APS Operations
	Multi-probe retrieval	In Progress – Implemented in *tike* and currently being tested – APS Operations

	Mini-batches	In Progress – Implemented and currently being tested – APS Operations
	Multi-wavelength	In Progress – APS Operations
	Arbitrary fly-scan	To do – APS Operations
	Multi-slice ptychography	To do – APS Operations
	Integration with CNN denoising and priors (regularization)	In Progress – APS Operations
	Affine position regularization	In Progress – APS Operations
High-Performance Implementations	Scalable distributed-memory GPU implementation of extended Ptychographic Iterative Engine (ePIE) method	Done – APS Operations and ASCR funding
	Scalable distributed-memory GPU implementation of Difference Map (DM) method	Done – APS Operations and ASCR funding
	Scalable distributed-memory GPU implementation of iterative least-squares solver for generalized maximum-likelihood (LSQ-ML) method	Done – APS Operations
	Ptychographic reconstruction using AI/ML	In Progress – APS Operations & LDRD

APS-U Era data rates are expected to be so large that traditional algorithms may not be able to keep up with acquired data. These data rates are so large, and the scientific problems that APS-U Era capabilities can enable are so great, that porting and scaling current models and algorithmic approaches may not realize the full promise of next-generation light sources. Using AI techniques, APS researchers have developed an approach to improve the performance of ptychographic reconstructions. A deep neural network model is trained to predict and reconstruct ptychographic x-ray data. This approach, PytchoNN, can then perform reconstructions up to 300 times faster than conventional iterative approaches and uses up to 5 times less data, speeding up both data acquisition and data reconstruction (see Figure 2-4).

Space-Mapping for Diffraction Data

The APS develops and supports the *RSMap3D* tool for diffraction space-mapping (see Figure 2-5). This software is available as an open-source package (https://github.com/AdvancedPhotonSource/rsMap3D). The tool allows users to examine the volume of collected data and select portions on which to apply transformations that convert detector pixel locations from diffractometer geometry to reciprocal-space units, and then map pixel data onto a 3D reciprocal-space grid. This application uses diffractometer angles, the energy of the scan and sample to detector distances to calculate either q-vector component values or HKL values. These values are calculated for each detector pixel and scan position. The calculated q/HKL value and pixel intensity is then binned in a 3D grid based on the selected q/HKL values. The core routines utilize OpenMP to parallelize operations across multiple cores on a shared-memory CPU. Data too big to fit entirely into memory at one time are processed in smaller chunks and reassembled to form the final output volume, allowing users to process arbitrarily large input datasets. It will be straightforward to extend this application to operate in a distributed-memory CPU environment if needed. These parallel and out-of-core computational techniques will be critical to handle larger data rates expected in the APS-U Era.

Table 2-11 Summary of diffraction space-mapping data processing needs and status for the ISN APS-U feature beamline.

Capability	Algorithm / Software Requirement	Status
RSM for Diffraction Data	Multi-core shared-memory CPU implementation	Done – APS Operations in collaboration with DESY
	Operate on out-of-core data	Done – APS Operations
	Distributed-memory CPU and GPU implementation	To do – If required

Spectroscopy Tools

An additional need for ISN is a spectroscopy tool. The relevant data acquisition mode is 2D or (possibly 3D) spatial scans with a full spectrum XANES spectrum at each pixel. Software should display spectra at each pixel. Related is a tool that allows PCA analysis of the data within an interactive feature that that allows display of the local spectra for each major component. For example, PCA would identify uniform/crystalline areas of a multicrystalline sample and grain boundaries, and extraction of the spectra for the relevant principal components would enable direct visualization of the spectral differences between uniform/crystalline areas and grain boundaries.

2.5.7 Effort, Funding, and Collaborations

Network infrastructure to the edge of the ISN APS-U feature beamline will be provided by APS-U funding. All other network infrastructure is funded by APS Operations funding as described in 1.2. Facility-wide data management, workflows, and science portal effort is funded by APS Operations (approximately 2.5 FTE per year) and the Globus Services team as described in 1.4. Some local computing resources dedicated to the ISN APS-U feature beamline for on-the-fly processing and experiment steering will be provided from APS-U funding. Large-scale computing resources are provided by APS Operations funding, ALCF funding, and Argonne Laboratory Computing Resource Center (LCRC) funding as described in 1.5. The APS dedicates approximately 1 FTE per year for XRF elemental mapping and diffraction space-mapping software development from APS Operations funding.

The following LDRD funding was awarded to support these efforts:

- *Novel Capabilities for Ultra-fast and Ultra-low-dose 3D Scanning Hard X-ray Microscopy* (FY18)
- *Enabling Automatic Learning of Atmospheric Particles through APS-U* (FY19)
- *Learning and Differentiating: Using Artificial Intelligence to Image Beyond the X-ray Depth of Focus Limit* (FY19)
- *Intelligent Ptychography Scan via Diffraction-Based Machine Learning* (FY20)
- *AutoPtycho: Autonomous, Sparse-sampled Ptychographic Imaging* (FY21)
- *AI Accelerator for 3D X-ray Phase Retrieval with Automatic Differentiation* (FY21)

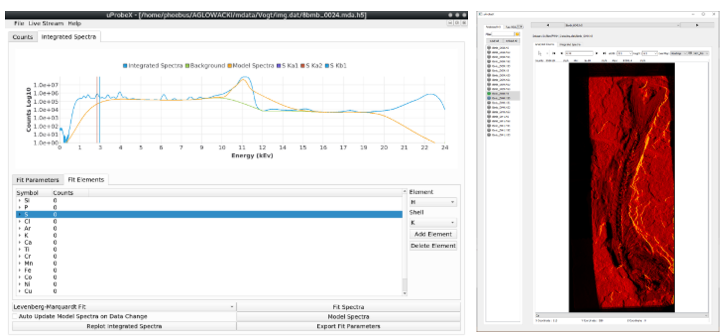


Figure 2-3 Left: uProbeX displaying integrated spectra from a dataset in blue, background subtraction in green, modeled spectra in orange, and elemental lines for element S. Right: uProbeX displaying Calcium quantities of an analyzed fish fossil. Elemental maps are generated with XRF-Maps.

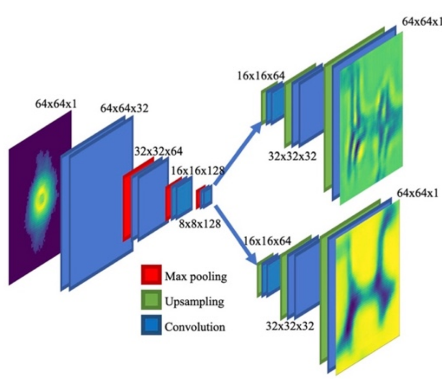


Figure 2-4 Architecture of PtychoNN, a deep convolutional neural network that can predict real-space amplitude and phase from input diffraction data alone.

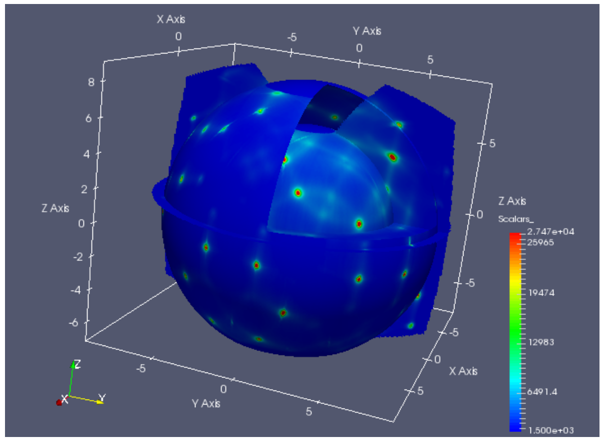


Figure 2-5 Visualization of a dataset from a pixel from a Silicon analyzer crystal processed with RSMap3D. This data shows a combination of diffraction (bright spots) and thermal diffuse scattering (broad lines connecting diffraction spots). This data was taken on the High Energy Resolution Inelastic X-Ray Spectrometer (HERIX) at APS Sector 30.

2.6 Polarization Modulation Spectroscopy (Polar) APS-U Feature Beamline

2.6.1 Summary

The Polarization Modulation Spectroscopy (Polar) APS-U feature beamline will use the polarization dependence of resonant absorption and scattering to study emergent quantum states in novel electronic and magnetic materials. Emphasis is placed on tuning/controlling competing ground states and electronic inhomogeneity with a combination of extreme high-pressures, low temperature, and high magnetic fields. Brilliant beams with tunable circular-and linear-polarization states will allow reaching extreme pressures as well as mapping electronic inhomogeneity in both real and reciprocal space.

In the APS-U era, the Polar APS-U feature beamline will continue to support techniques relying on x-ray polarization control but will augment its capabilities by leveraging the enhanced brilliance and coherence of APS-U beams, coupled with extreme sample environments. A future installation of a pair of polarizing undulators that leverage use of round insertion device vacuum chambers made possible in APS-U will provide exquisite polarization control (circular, elliptical, arbitrary linear) and extend the energy range of polarization modulated spectroscopies to high energy resonances up to 27 keV. Dichroic techniques currently supported include X-ray Magnetic Circular Dichroism (XMCD), X-ray linear dichroism (XLD), and Resonant Magnetic Scattering/Reflectivity. New techniques in Polar that require software development, and entail significant increases in data volumes, are: (1) Dichroic Ptychography, including fly scanning/interferometry and tomographic modes, for imaging of electronic/magnetic domains in reciprocal space with ~ 10 nm resolution; (2) Scanning dichroic x-ray absorption imaging, including fly scanning/interferometry and tomographic modes, for imaging of electronic/magnetic domains in real space with ~ 200 nm resolution. New sample environments such as a 3-axis 9-1-1 T magnet also enable expansion of dichroic technique modalities to include X-ray Magnetic Linear Dichroism (XMLD) and Dichroic emission spectroscopy (RXES-MCD). The new dichroic imaging techniques are implemented in combination with extreme conditions of ultra-high pressure, low temperature, and high magnetic field. The data volume estimates for these techniques form the basis for networking infrastructure, controls, data management, and data processing planning.

Table 2-12 shows the estimated data generation rates at the Polar APS-U feature beamline. The Polar APS-U feature beamline is anticipated to collect approximately 84 TB of compressed raw data per year. These data generation estimates form the basis for networking infrastructure, controls, data management, and data processing planning.

Table 2-12 Estimated data generation rates at the Polar APS-U feature beamline.

	Technique	Detector	Frame Size (MB)	Data Rate (MB/s)	Raw Data set Size (MB)	Daily Utilization (%)*	Annual Utilization (%) **	Raw Data Per Year (GB) **
Today	4-ID-D XAS/XMCD/XLD	Si Drift Multi-Element, photodiodes	2.73	0.0014	547	90	15	3
	4-ID-D Dichroic Resonant Scattering/Reflectivity	Scintilator (NaI) / Avalanche Photodiode	0.38	7.6E-7	0.3815	90	10	0.0012
APS-U Era	4-ID-G Dichroic Ptychography imaging 2D/3D modes (fly scans, interferometry)	Dectris EIGER2 X 1M	4.16	208	21,000	90	20	76,000
	4-ID-G Dichroic Resonant Scattering imaging (200 nm)	Dectris EIGER 2 X 1M	4.16	41.6	4,000	90	10	7,500
	4-ID-G Dichroic Absorption Tomography (200 nm)	Photodiodes	0.38	7.6E-7	0.3815	90	10	0.0012
	4-ID-H Dichroic Absorption XAS/ XMCD/XMLD (mapping 300 nm, high-pressure 7 Mbar)	Si Drift 7-Element Photodiodes	1.302	0.33	1172	70	25	51
	4-ID-H Dichroic X-ray emission (RXES-MCD)	Lambda 250k	1	1	20	90	10	1

* Based on 1,440 minutes in one day.

** Based on 210 days of beam time per fiscal year.

2.6.2 Network Architecture and Infrastructure

Network infrastructure needs will be addressed by the facility plan described in 1.2.

2.6.3 Controls, Data Acquisition, and Detector Integration

Controls, data acquisition, and detector integration needs will be addressed by the facility plan described in 1.3. Detailed plans and schedules are described in respective APS-U beamline controls documents, Equipment Specification Documents (ESDs), and Instrument Specification Documents (ISDs). Dichroic Ptychography in 2D and 3D (tomographic) modes will make use of interferometry to inform on actual sample and beam position during fly scans, information to be used for image alignment before reconstructions. Integration of interferometry with beamline controls will be done with FPGA-based *softGlueZynq*.

2.6.4 Data Management, Workflows, and Science Portals

The APS-U Polar feature beamline will leverage the data management, workflow, and science portal efforts described in 1.4. The APS Data Management System, the facility-wide software and hardware system for managing data and workflows, will integrate data collection and processing workflows, manage user permissions based on experiment groups, and enable user access to data. Globus tools will provide a science portal for viewing and accessing data and automating the re-processing of data after the allotted experiment time has ended. For the Polar APS-U feature beamline, workflows will provide a pipeline to automatically run analysis and reconstruction tools.

2.6.5 Computing Infrastructure

Computing infrastructure needs will be addressed by the facility plan described in 1.5. Some local computing resources will be provided for on-the-fly data processing and experiment steering. Under the scope of LDRD project 2022-0008, *Development of 3D Dichroic Ptychography at APS*, a 2 x A100 80GB GPU server was purchased in FY22 to provide a dedicated platform for reconstruction of ptychography data at 4-ID-D. Computing capacity for larger data processing tasks and for post-experiment processing and analysis will be provided by computing centers, including the ALCF and Argonne’s Laboratory Computing Resource Center (LCRC). The APS Data Management System and Globus tools will be used to seamlessly integrate these resources.

2.6.6 Data Reduction and Analysis Software

The processing and analysis of dichroic data currently collected at 4-ID-D leverages a variety of in-house scripting tools and open-source software, such as *GenX* code for Dichroic Resonant Reflectivity data (https://aglavic.github.io/genx), and *polartools* for python-based processing of dichroic and resonant diffraction data (https://github.com/APS-4ID-POLAR/polartools).

Reconstruction algorithms for processing Dichroic Ptychography data, including tomographic mode, are currently being developed with funding from LDRD 2022-0008, *Development of 3D Dichroic Ptychography at APS*. The APS will leverage the *tike* ptychography toolkit as a framework for implementing algorithms in this area. If required, the APS will develop scalable distributed-memory CPU and GPU implementations for processing the large volumes of Dichroic Ptychography data generated by the Polar APS-U feature beamline, especially for fly-scan implementations.

Reconstruction algorithms for scanning Dichroic Tomography are not yet defined. The APS is currently performing preliminary R&D in this area under LDRD 2022-0008. The APS will leverage the *TomoPy* toolkit as a framework for implementing algorithms in this area. If required, the APS will develop scalable distributed-memory CPU and GPU implementations for processing the large volumes of Tomographic CDI data generated by the Polar APS-U feature beamline.

Table 2-13 Summary of data reduction needs, approaches, and status for the Polar APS-U feature beamline.

Capability	Algorithm / Software Requirement	Status
Dichroic Resonant Absorption and Scattering	Algorithm development	Done – APS Operations
	Single CPU software implementation	Done – APS Operations – *polartools* Open Source - *GenX*, etc
	Scalable distributed-memory CPU and GPU implementation	In progress – LDRD, APS operations
Dichroic Ptychography	Algorithm development	In progress – LDRD
	Single CPU software implementation	In progress – LDRD
	Scalable distributed-memory CPU and GPU implementation	In progress – LDRD, APS operations
Dichroic Tomography	Algorithm development	In progress – APS Operations
	Single CPU software implementation	In progress – APS Operations – Leverage *TomoPy* tools for tomographic reconstruction
	Scalable distributed-memory CPU and GPU implementation	In progress – LDRD, APS operations

2.6.7 Effort, Funding, and Collaborations

Network infrastructure to the edge of the Polar APS-U feature beamline will be provided by APS-U funding. All other network infrastructure is funded by APS Operations funding as described in 1.2. Facility-wide data management, workflows, and science portal effort is funded by APS Operations (approximately 2.5 FTE per year) and the Globus Services team as described in 1.4. Some local computing resources dedicated to the Polar APS-U feature beamline for on-the-fly processing and experiment steering will be provided from APS-U funding. Large-scale computing resources are provided by APS Operations funding, ALCF funding, and Argonne Laboratory Computing Resource Center (LCRC) funding as described in 1.5. LDRD 2022-0008, *Development of 3D Dichroic Ptychography at APS* is providing effort to develop reconstruction algorithms for dichroic ptychography data (2D and 3D), and resources to augment local computing infrastructure.

The following LDRD funding was awarded to support these efforts:

- *Development of 3D Dichroic Ptychography at APS* (FY22)

2.7 PtychoProbe APS-U Feature Beamline

2.7.1 Summary

The PtychoProbe APS-U feature beamline (Ptychography + Nanoprobe) is designed to realize the highest possible spatial resolution X-ray microscopy both for structural and chemical information. The unprecedented brightness of the APS MBA lattice will be exploited to produce a nm-beam of focused hard X-rays to achieve the highest possible sensitivity to trace elements, and ptychography will be used to further improve the spatial resolution for structural components to its ultimate limit. The proposed beamline will enable high resolution two- and three-dimensional imaging of thick objects and bridge the resolution gap between contemporary X-ray and electron microscopy. Extending X-ray microscopy into the nanoscale is crucial for understanding complex hierarchical systems on length scales from atomic up to meso and macroscales, and time scales down to the microsecond level, and is applicable to scientific questions ranging from biology to earth and environmental materials science, to electrochemistry, catalysis and corrosion, and beyond.

Table 2-14 shows estimated data generation rates at the PtychoProbe APS-U feature beamline. The PtychoProbe APS-U feature beamline is anticipated to collect approximately 48.8 PB of raw data per year and 4.8 PB of compressed raw data per year, in comparison to approximately 730 TB of raw data and approximately 73 TB of compressed raw data collected today across the 2-ID-D ptychography, 2-ID-E XRF, and Bio Nano-Probe (BNP) XRF instruments. Both uncompressed and compressed data sizes are given because uncompressed data is often required for data processing. This represents an approximately 100x increase in data. These data generation estimates form the basis for networking infrastructure, controls, data management, and data processing planning.

Table 2-14 Data generation rates today at the 2-ID-D ptychography, 2-ID-E XRF, and Bio Nano-Probe (BNP) XRF instruments (for comparison) and estimated data generation rates at the PtychoProbe APS-U feature beamline.

	Technique	Detector	Frame Size (MB)	Average Frame Rate (Hz)	Data Rate (MB/s)	Raw Data set Size (GB)	Compression Factor *	Daily Utilization (%)**	Raw Data Per Day (GB)**	Compressed Raw Data Per Day (GB)**	Annual Utilization (%) ***	Raw Data Per Year (TB) ***	Compressed Raw Data Per Year (TB) ***
Today	2-ID-D Ptychography	Dectris Eiger 500K	1.010	100	100	40	10	80	6,912	691	50	726	72.6
	2-ID-E XRF	Vortex ME4	0.008	20	0.15	1.91	5	100	13.18	2.64	80	2.16	0.43
	BNP XRF	Vortex ME4	0.008	20	0.15	0.69	15	100	13.18	0.88	80	2.16	0.14
APS-U Era	PtychoProbe XRF	Vortex ME7	0.054	1,000	53	214	5	80	631	42.7	80	106	21
	PtychoProbe Ptychography – Slow	Dectris Eiger 2XE 1.5M	3	2,000	6,000	120	10	80	414,720	41,472	40	34,836	3,483.6
	PtychoProbe Ptychography – Fast	TBD – Small fast detector (200x200)	0.08	30,000	2,400	3.2	10	80	165,888	16,589	40	13,934	1,393.4

* A compression factor of 1.0 represents no compression.
** Based on 1,440 minutes in one day.
*** Based on 210 days of beam time per fiscal year.

2.7.2 Network Architecture and Infrastructure

Network infrastructure needs will be addressed by the facility plan described in 1.2.

2.7.3 Controls, Data Acquisition, and Detector Integration

Controls, data acquisition, and detector integration needs will be addressed by the facility plan described in 1.3. Detailed plans and schedules are described in respective APS-U beamline controls documents, Equipment Specification Documents (ESDs), and Instrument Specification Documents (ISDs).

2.7.4 Data Management, Workflows, and Science Portals

The PtychoProbe APS-U feature beamline will leverage the data management, workflow, and science portal efforts described in 1.4. The APS Data Management System, the facility-wide software and hardware system for managing data and workflows, will integrate data collection and processing workflows, manage user permissions based on experiment groups, and enable user access to data. Globus tools will provide a science portal for viewing and accessing data and automating the re-processing of data after the allotted experiment time has ended. For the PtychoProbe APS-U feature beamline, workflows will provide a pipeline to automatically run tools to remove artifacts from data, reconstruct the XRF and Ptychography data set, and view results.

2.7.5 Computing Infrastructure

Computing infrastructure needs will be addressed by the facility plan described in 1.5. Some local computing resources will be provided for on-the-fly data processing and experiment steering. Computing capacity for larger data processing tasks and for post-experiment processing and analysis will be provided by computing centers, including the ALCF and Argonne's Laboratory Computing Resource Center (LCRC). The APS Data Management System and Globus tools will be used to seamlessly integrate these resources.

2.7.6 Data Reduction and Analysis Software

The PtychoProbe APS-U feature beamline requires two modes of data processing: elemental fitting for XRF microscopy data and ptychography data reconstruction. Descriptions of current efforts and plans in each of these areas follow. Table 2-15 and Table 2-16 summarize capabilities for each of these two modes, respectively. Many of the data processing requirements for the PtychoProbe APS-U feature beamline are like those of the ISN APS-U feature beamline described in 2.5.

Elemental Fitting for XRF Microscopy Data

The APS develops and supports the *XRF-Maps and uProbeX* software packages for XRF microscopy data processing and visualization (see Figure 2-6). This software is available as open-source packages (https://github.com/AdvancedPhotonSource/XRF-Maps and https://github.com/AdvancedPhotonSource/uProbeX). The *XRF-Maps* package performs elemental map fitting and the *uProbeX* application is a GUI for visualizing *XRF-Maps* results. *XRF-Maps* and *uProbeX* are both written in C++. *XRF-Maps* supports multi-core data processing in a shared-memory CPU environment and has a Python wrapper which allows all the functionality to be called from a Python environment.

APS-U enhancements will allow for larger scan areas and/or finer pixel sizes resulting in larger datasets. These larger datasets may not be able to fit in system memory. To accommodate this *XRF-Maps* implements a streaming architecture that allows processing a dataset spectra by spectra without having to load the entire dataset. Only a limited number of spectra are loaded based on memory limits, processed, and saved to an HDF5 file until the whole dataset is processed. As data sizes increase, it may be become necessary to develop GPU-based and distributed-memory CPU- and GPU-based elemental fitting software.

The higher intensity x-ray beam generated by the APS-U storage ring necessitates the utilization of self-absorption correction when generating elemental maps. APS researchers and instrument staff are working on developing new self-absorption correction algorithms in collaboration with staff at the National Synchrotron Light Source II (NSLS-II) at Brookhaven National Laboratory (BNL). These algorithms are being implemented and tested in the *XRF-Maps* software.

Table 2-15 Summary of XRF microscopy elemental mapping data processing needs and status for the ISN APS-U feature beamline.

Capability	Algorithm / Software Requirement	Status
XRF Elemental Map Fitting	Algorithms for elemental map fitting	Done
	Multi-core shared-memory CPU implementation	Done – APS Operations

	Streaming data processing / operate on out-of-core data	Done – APS Operations
	Distributed-memory CPU and GPU implementation	To do – If required
XRF Self-Absorption Correction	Self-absorption correction algorithm development	In Progress – APS Operations and collaborations with NSLS-II
	Self-absorption correction implementation in *XRF-Maps*	In Progress – APS Operations

Ptychography Reconstruction

Ptychography has emerged as a powerful technique at synchrotron light sources. It will play a central role in answering many emerging scientific questions that the upgraded APS will help solve. Advanced ptychographic reconstruction algorithms and software are critical to take advantage of this new and innovative technique.

Multiple ptychographic reconstruction algorithms are required to achieve reasonable reconstruction quality to best analyze ptychography data collected for different domains and of varying sample characteristics. The APS has implemented the extended Ptychographic Iterative Engine (ePIE), regularized Ptychographic Iterative Engine (rPIE), conjugate gradient, Difference Map (DM), and iterative Least-SQuares solver for generalized Maximum-Likelihood (LSQ-ML) methods. Algorithms to help improve reconstruction quality, such as position and probe variation correction, and affine position regularization, are being developed and implemented.

Due to the computationally complex nature of ptychographic reconstruction algorithms and due to the anticipated increase in data rates and sizes in the APS-U Era, distributed high-performance implementations of ptychography reconstruction software are required. The APS with collaborators in Argonne's Mathematics & Computer Science (MCS) division developed *PtychoLib*, a distributed-memory GPU implementation of the extended Ptychographic Iterative Engine (ePIE) in 2014 and integrated Difference Map (DM) algorithms in 2018. *PtychoLib* was written in C++ and uses MPI and CUDA. The software was shown to scale on up to 256 GPUs on the ALCF's Cooley GPU cluster. This software has been supported and extended since then and has been the main tool used for high-performance ptychography reconstructions at APS beamlines. *PtychoLib* has been the main tool used for high-performance ptychography reconstructions at APS beamlines for the past decade. *PtychoPy* (https://github.com/kyuepublic/ptychopy) was developed as a Python wrapper and GUI for *PtychoLib*. Since then, the APS has consolidated ptychography development into the *tike* (https://github.com/AdvancedPhotonSource/tike) toolkit in order to make installing and developing new ptychography features and algorithms easier. This toolkit is written in Python and uses CuPy as the underlying GPU framework. All the reconstruction features of *PtychoLib* have been reimplemented in the tike toolkit including MPI and thread-based parallelism. *Ptychodus*, is a new pyQT-based GUI/workflow manager for ptychography reconstruction workflows has also been created in order to provide live reconstruction visualization and analysis.

Table 2-16 Summary of ptychography reconstruction needs and status for the PtychoProbe APS-U feature beamline.

Capability	Algorithm / Software Requirement	Status
Conventional Reconstruction	GPU implementation of extended Ptychographic Iterative Engine (ePIE) method	Done – APS Operations
	GPU implementation of Difference Map (DM) method	Done – APS Operations
	GPU implementation of the iterative Least-SQuares solver for generalized Maximum-Likelihood (LSQ-ML) method	Done – APS Operations
Improved Reconstruction Quality	Position correction	In Progress – Implemented in *tike* and currently being tested – APS Operations
	Probe variation correction	In Progress – Implemented in *tike* and currently being tested – APS Operations
	Multi-probe retrieval	In Progress – Implemented in *tike* and currently being tested – APS Operations

	Mini-batches	In Progress – Implemented and currently being tested – APS Operations
	Multi-wavelength	In Progress – APS Operations
	Arbitrary fly-scan	To do – APS Operations
	Multi-slice ptychography	To do – APS Operations
	Integration with CNN denoising and priors (regularization)	In Progress – APS Operations
	Affine position regularization	In Progress – APS Operations
High-Performance Implementations	Scalable distributed-memory GPU implementation of extended Ptychographic Iterative Engine (ePIE) method	Done – APS Operations and ASCR funding
	Scalable distributed-memory GPU implementation of Difference Map (DM) method	Done – APS Operations and ASCR funding
	Scalable distributed-memory GPU implementation of iterative least-squares solver for generalized maximum-likelihood (LSQ-ML) method	Done – APS Operations
	Ptychographic reconstruction using AI/ML	In Progress – APS Operations & LDRD

APS-U Era data rates are expected to be so large that traditional algorithms may not be able to keep up with acquired data. These data rates are so large, and the scientific problems that APS-U Era capabilities can enable are so great, that porting and scaling current models and algorithmic approaches may not realize the full promise of next-generation light sources. Using AI techniques, APS researchers have developed an approach to improve the performance of ptychographic reconstructions. A deep neural network model is trained to predict and reconstruct ptychographic x-ray data. This approach, PtychoNN, can then perform reconstructions up to 300 times faster than conventional iterative approaches and uses up to 5 times less data, speeding up both data acquisition and data reconstruction (see Figure 2-7).

2.7.7 Effort, Funding, and Collaborations

Network infrastructure to the edge of the PtychoProbe APS-U feature beamline will be provided by APS-U funding. All other network infrastructure is funded by APS Operations funding as described in 1.2. Facility-wide data management, workflows, and science portal effort is funded by APS Operations (approximately 2.5 FTE per year) and the Globus Services team as described in 1.4. Some local computing resources dedicated to the PtychoProbe APS-U feature beamline for on-the-fly processing and experiment steering will be provided from APS-U funding. Large-scale computing resources are provided by APS Operations funding, ALCF funding, and Argonne Laboratory Computing Resource Center (LCRC) funding as described in 1.5. The APS dedicates approximately 1 FTE per year for XRF elemental mapping and diffraction space-mapping software development from APS Operations funding.

The following LDRD funding was awarded to support these efforts:

- *Novel Capabilities for Ultra-fast and Ultra-low-dose 3D Scanning Hard X-ray Microscopy* (FY18)
- *Enabling Automatic Learning of Atmospheric Particles through APS-U* (FY19)
- *Learning and Differentiating: Using Artificial Intelligence to Image Beyond the X-ray Depth of Focus Limit* (FY19)
- *Intelligent Ptychography Scan via Diffraction-Based Machine Learning* (FY20)
- *AutoPtycho: Autonomous, Sparse-sampled Ptychographic Imaging* (FY21)
- *AI Accelerator for 3D X-ray Phase Retrieval with Automatic Differentiation* (FY21)

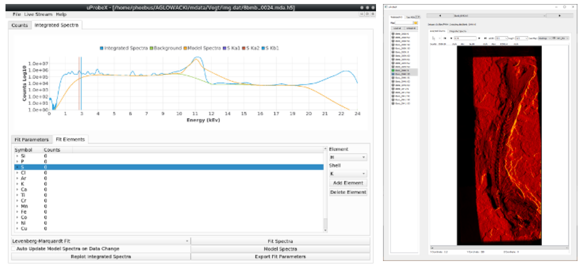


Figure 2-6 Left: uProbeX displaying integrated spectra from a dataset in blue, background subtraction in green, modeled spectra in orange, and elemental lines for element S. Right: uProbeX displaying Calcium quantities of an analyzed fish fossil. Elemental maps are generated with XRF-Maps.

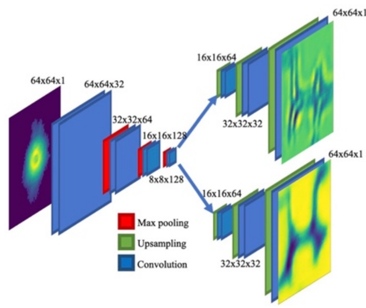


Figure 2-7 Architecture of PtychoNN, a deep convolutional neural network that can predict real-space amplitude and phase from input diffraction data alone.

2.8 X-ray Photon Correlation Spectroscopy (XPCS) APS-U Feature Beamline

2.8.1 Summary

The X-ray Photon Correlation Spectroscopy (XPCS) APS-U feature beamline will be dedicated to time-resolved coherent x-ray scattering experiments for a diverse scientific community; experiments will exploit the brilliance of the upgraded source to study fundamental materials structures. Since the signal to noise for XPCS scales as the square of the brilliance which will increase 500x in the APS-U era, it will be possible to measure faster dynamics and weaker scattering systems. The small- and wide-angle instruments will probe dynamics in soft and hard matter respectively.

In the APS-U era, the XPCS APS-U feature beamline will operate in modes that vary between collecting time series of area detector frames at very high frame rates (up to 100 kHz) and at moderate frame rates (a few kHz to Hz). The XPCS APS-U feature beamline is anticipated to collect up to approximately 20 PB of raw data per year, in comparison to approximately 0.1 PB of data collected today at the 8-ID beamline. These data generation estimates form the basis for networking infrastructure, controls, data management, and data processing planning. Table 2-17summarizes the data rates and total data accumulation for the anticipated experimental configuration at the XPCS APS-U feature beamline.

Table 2-17 Estimated data generation rates at the XPCS APS-U feature beamline.

	Technique	Detector	Frame Size (MB)	Detector Rate (Hz)	Data Rate (GB/s) *	Raw Dataset Size (GB)	Daily Utilization (%)**	Data Per Day (TB)**	Annual Utilization (%)***	Data Per Year (TB)***
Today	XPCS	LAMBDA 750K	1.5	500	0.74	14.84	25	1.57	25	82
	XPCS – Fast	UHSS 500K	0.5	56,000	27.34	48.83	5	2.31	5	20

APS-U Era	XPCS – Fast	Eiger 4M	4.0	4,000	15.63	390.63	20	65.92	20	2,765
	XPCS – Fast	UHSS 3M	3.0	56,000	164.06	292.97	15	207.64	20	8,725
	XPCS – Average	Eiger 4M	16.0	1,000	15.63	468.75	15	65.92	20	2,765
	XPCS – Average	UHSS 3M	6.0	8,500	49.80	351.56	20	168.09	20	7,055

* Raw uncompressed data rate.
** Based on 1,440 minutes in one day.
*** Based on 210 days of beam time per fiscal year.

2.8.2 Network Architecture and Infrastructure

Network infrastructure needs will be addressed by the facility plan described in 1.2.

2.8.3 Controls, Data Acquisition, and Detector Integration

Controls, data acquisition, and detector integration needs will be addressed by the facility plan described in 1.3. Detailed plans and schedules are described in respective APS-U beamline controls documents, Equipment Specification Documents (ESDs), and Instrument Specification Documents (ISDs).

2.8.4 Data Management, Workflows, and Science Portals

The APS-U XPCS feature beamline will leverage the data management, workflow, and science portal efforts described in 1.4. The APS Data Management System, the facility-wide software and hardware system for managing data and workflows, will integrate data collection and processing workflows, manage user permissions based on experiment groups, and enable user access to data. Globus tools will provide a science portal for viewing and accessing data and automating the re-processing of data after the allotted experiment time has ended. For the XPCS APS-U feature beamline, workflows will provide a pipeline to automatically transfer data to computing resources for processing, launch processing jobs, and save results for visualization. A streaming data pipeline will be developed so that the data is processed in near real-time.

2.8.5 Computing Infrastructure

Computing infrastructure needs will be addressed by the facility plan described in 1.5. Some local computing resources may be provided for on-the-fly data processing and experiment steering. Computing capacity for these data processing tasks and for post-experiment processing and analysis will be provided by computing centers, including the ALCF and Argonne's Laboratory Computing Resource Center (LCRC). The APS Data Management System and Globus tools will be used to seamlessly integrate these resources.

2.8.6 Data Reduction and Analysis Software

The main technique used for analyzing XPCS data involves auto-correlating time-resolved signals. Table 2-18 shows algorithm requirements based on science drivers, along with current algorithm development status. The Multi-Tau and Two-Time correlation algorithms are most used when processing data that studies equilibrium and non-equilibrium dynamics, respectively. These algorithms are already well developed and in common use. Higher-order time correlations are required to study spatial and temporal heterogeneity, intermittent dynamics, and avalanches. The study of speckle metrology, nanoscale flow, and velocimetry require the use of spatial-temporal cross-correlations. Development of these latter two classes of algorithms is underway at the APS and with APS Users and collaborators, including collaborators at CAMERA.

Table 2-18 Summary of algorithm requirements for the XPCS APS-U feature beamline.

Science Driver	Algorithm Requirement	Status
Equilibrium Dynamics	Multi-Tau Correlation	Done
Non-Equilibrium Dynamics	Two-Time Correlation	Done
Spatial and temporal heterogeneity, intermittent dynamics, avalanches	Higher-Order Time Correlations	In Progress – APS Operations, APS User group collaborations, and CAMERA
Speckle metrology, nanoscale flow, and velocimetry	Spatial-Temporal Cross-Correlations	In Progress – APS Operations, APS User group collaborations, and CAMERA

The APS develops and maintains the high-performance *boost-corr* auto-correlation software package for processing XPCS data. This tool utilizes multiple CPU cores and a single GPU in a shared-memory environment to quickly produce auto-correlations of XPCS data using the Multi-Tau and Two-Time algorithms. The *pyXpcsViewer* tool (https://github.com/AdvancedPhotonSource/pyXpcsViewer) helps users visualize and analyze correlation results generated from *boost-corr* (see Figure 2-8). A GPU implementation of the Multi-Tau algorithm has been developed that shows significant performance improvements over the current CPU implementation.

The current feature sets and performance today's software is adequate for today's needs. However, the estimated increase in overall data that will be generated at the XPCS APS-U feature beamline necessitates improvements and advances in software. The APS is planning to develop implementations of higher-order time correlation and spatial-temporal cross-correlation algorithms and develop higher-performance distributed-memory CPU and GPU software applications. Table 2-19 summarizes XPCS software capabilities and current development statuses for the XPCS APS-U feature beamline.

Table 2-19 Summary XPCS APS-U feature beamline data reduction and processing software capabilities and needs.

Capability	Software Requirement	Status
Multi-Tau Correlation	Shared-memory CPU implementation	Done – APS Operations
	Distributed-memory CPU implementation	To do – APS Operations
	Single GPU implementation	Done – APS Operations
	Multiple GPU implementation	To do – APS Operations
Two-Time Correlation	Shared-memory CPU implementation	Done – APS Operations
	Distributed-memory CPU implementation	To do – APS Operations
	Single GPU implementation	Done – APS Operations
	Multiple GPU implementation	To do – APS Operations
Higher-Order Time Correlations	CPU implementation	To do – APS Operations – pending algorithm development
	GPU implementation	To do – APS Operations – pending algorithm development
Spatial-Temporal Cross-Correlations	CPU implementation	To do – APS Operations – pending algorithm development
	GPU implementation	To do – APS Operations – pending algorithm development

Physics-Informed Machine Learning from Speckle Patterns

While fitting measured correlation functions to approximate models is often used to extract physical insights from raw speckle patterns, there are potential benefits to learning physics directly from the data. To this end, we outline and test a proof of concept for recovering physical equations from measured speckle patterns based on neural Ordinary Differential Equations (ODEs). In contrast to a traditional neural ODE workflow, the real-space dynamics of the system probed by coherent scattering are considered inaccessible apart from the initial condition. Instead, a neural network model of the ODE is trained by minimizing a loss function of the predicted and true sequence of speckle patterns. Using the trained model, we can then not only accurately predict future dynamics but also extract the model's dependence on the system variables to recover quantitative information about the governing equations. The extension of this framework to more complex systems and more realistic simulations is ongoing and will seek answers to additional questions, particularly the suitable balance between model flexibility and interpretability.

Automated Classification of Experimental Data

Recent work by APS scientists has applied unsupervised machine learning to automate the processing of XPCS Two-Time correlations. While algorithms exist for calculating these correlations from scattering data, methods for interpretation and quantification of Two-time correlations are still needed for studying non-equilibrium dynamics. A deep neural network was trained to recognize and reproduce spatial patterns in two-time correlations and encode these features into a low-dimensional representation. After using this neural network to encode entire datasets, clustering was performed to classify experimental data without requiring any input or physical knowledge form the user. In a similar manner, this automated method can be used to suggest similar two-time correlations based on a user-specified feature of interest, drastically reducing the analysis time required to comb through large XPCS datasets and identify interesting results. Future work aims to again use unsupervised machine learning to detect anomalous events and changes in the dynamic behavior of evolving systems to enable on-line data analysis at the beamline.

2.8.7 Effort, Funding, and Collaborations

Network infrastructure to the edge of the XPCS APS-U feature beamline will be provided by APS-U funding. All other network infrastructure is funded by APS Operations funding as described in 1.2. Facility-wide data management, workflows, and science portal effort is funded by APS Operations (approximately 2.5 FTE per year) and the Globus Services team as described in 1.4. Some local computing resources dedicated to the XPCS APS-U feature beamline for on-the-fly processing and experiment steering will be provided from APS-U funding. Large-scale computing resources are provided by APS Operations funding, ALCF funding, and Argonne Laboratory Computing Resource Center (LCRC) funding as described in 1.5. The APS dedicates approximately 1.0 FTE per year for XPCS related algorithm and software development from APS Operations funding.

CAMERA provides effort in support of XPCS algorithm development. An Argonne MGM Fellow provides effort related to physics-informed machine learning.

The following LDRD funding was awarded to support this effort:

- *Intermittent Dynamics in Hard and Soft Materials enabled by APS-U* (FY22)

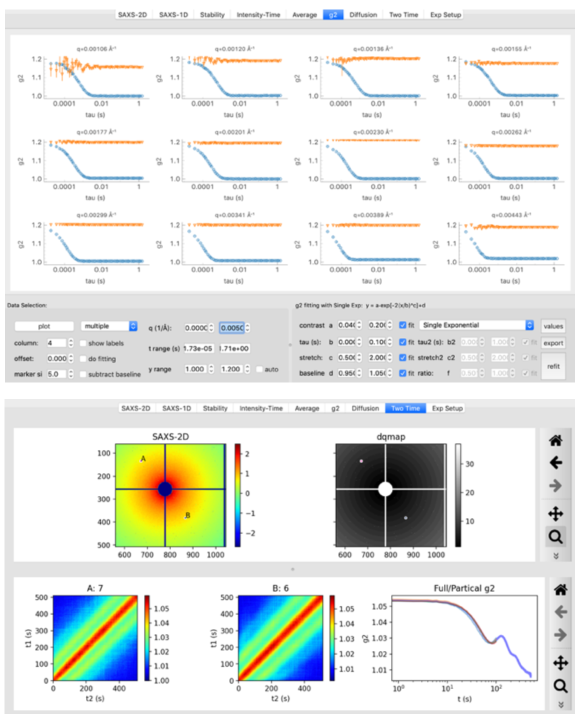


Figure 2-8 Top: g2 plots of the multi-tau correlation results. Users have many options to visualize data. Bottom: Visualization of a two-time correlation of a rubber sample. Users can select the region of interest by clicking the mouse.

2.9 3D Micro and Nano (3DMN) Diffraction APS-U Feature Beamline

2.9.1 Summary

The 3D Micro and Nano Diffraction (3DMN) APS-U feature beamline is designed to address a wide range of spatially inhomogeneous materials problems at the mesoscopic scale. These problems range over many areas of science where previous x-ray diffraction techniques are insufficient due to the short length scale of the inhomogeneities in the materials. 3DMN proposes to overcome current difficulties by using the bright MBA source to provide small intense x-ray spots (50-200nm) to investigate the important spatial variations of strain and structure that define this wide range of scientifically and technologically important materials.

In the APS-U era, the 3DMN feature beamline will perform Laue depth reconstruction diffraction scans. 3DMN will be able to operate in a mode like the current wire or knife-edge scan mode. This should allow analysis to work with some adjustments for data volume. 3DMN's updated detectors will lead to an increase in the size of acquired data from 6 megapixels (from 3 detectors) to 10 megapixels per collection point. To optimize use of updated beam parameters, it will be necessary to further decrease the size of the steps in a scan thus increasing the data volume. A new more efficient data collection mode is proposed that uses scans with a mask in place of a wire. Instead of blocking off one row with a wire in the scattered beam, a mask is used which passes only the previously blanked row. This new method allows for processing a larger data volume at each point by holding the number of scanned points close to the current wire scan. This will keep data volumes per dataset lower but requires implementing a new algorithm that is currently being developed. The new mode will allow more datasets of equal quality to be collected in the same amount of time as compared to the current wire scan method.

Table 2-20 shows the estimated data generation rates at the 3DMN APS-U feature beamline, and current data rates at the 34-ID-E instrument, for comparison. The 3DMN APS-U feature beamline is anticipated to collect approximately 2.8 PB of raw compressed data per year, in comparison to approximately 400 TB of data collected today at the 34-ID-E instrument. This represents slightly less than a one-order-of-magnitude increase in data. Note that in the APS-U era, it is anticipated that the 3DMN instrument will likely use the mask scan mode instead of the wire scan mode, assuming sufficient algorithmic developments. Although the overall amount of data generated by the two modes is the same in Table 2-20, the mask scan mode estimates represent a much larger amount of individual sample scans, and thus more final data. These data generation estimates form the basis for networking infrastructure, controls, data management, and data processing planning.

Table 2-20 Data generation rates today at the 34-ID-E Laue diffraction instrument (for comparison) and estimated data generation rates at the 3DMN APS-U feature beamline.

	Technique	Detector	Frame Size (MB)	Detector Rate (Hz)	Data Rate (GB/s)	Frames Per Dataset	Raw Dataset Size (GB)	Compression Factor*	Compressed Dataset Size (GB)	Dataset Collection Time (min)	Daily Utilization (%)**	Data Per Day (TB)**	Annual Utilization (%)***	Data Per Year (TB)***
Today	Wire Scan	PE XRD 1620 AN	8	4	0.03	500	3.91	1	3.91	2.08	50	1.32	45	125
		PE XRD 1620 AN + 2 x PE XRD 0820 AN	12	4	0.05	500	5.86	1	5.86	2.08	50	1.98	5	21
		PE XRD 1620 AN	8	8	0.06	500	3.91	1	3.91	1.04	50	2.64	45	249
		PE XRD 1620 AN + 2 x PE XRD 0820 AN	12	8	0.09	500	5.86	1	5.86	1.04	50	3.96	5	42
		Pilatus 6M	24	100	2.34	1000	23.44	3.5	6.7	0.17	50	28.25	20	1,187

APS-U Era[++]	Wire Scan – Fast[+]	Pilatus 6M + 2 x 2 MP Detectors	40	100	3.91	1000	39.06	3.5	11.16	0.17	50	47.08	5	494
	Wire Scan – Average[+]	Pilatus 6M	24	25	0.59	1000	23.44	3.5	6.7	0.67	50	7.06	70	1,038
		Pilatus 6M + 2 x 2 MP Detectors	40	25	0.98	1000	39.06	3.5	11.16	0.67	50	11.77	5	124
	Mask Scan – Fast[+]	Pilatus 6M	24	100	2.34	200	4.69	3.5	1.34	0.03	50	28.25	20	1,187
		Pilatus 6M + 2 x 2 MP Detectors	40	100	3.91	200	7.81	3.5	2.23	0.03	50	47.08	5	494
	Mask Scan – Average[+]	Pilatus 6M	24	25	0.59	200	4.69	3.5	1.34	0.13	50	7.06	70	1,038
		Pilatus 6M + 2 x 2 MP Detectors	40	25	0.98	200	7.81	3.5	2.23	0.13	50	11.77	5	124

[*] A compression factor of 1.0 represents no compression.
[**] Based on 1,440 minutes in one day.
[***] Based on 210 days of beam time per fiscal year.
[+] It is anticipated that the 3DMN instrument will likely use the mask scan mode instead of the wire scan mode, assuming sufficient algorithmic developments. Although the overall amount of data generated by the two modes is the same, the mask scan mode estimates represent a much larger amount of individual sample scans.
[++] The APS-U project has descoped certain parts of the 3DMN Feature beamline, including detector purchases. Although the detectors listed in the table may not be purchased as a part of the APS-U project, this table represents the desired long-term potential capabilities intended for this beamline.

2.9.2 Network Architecture and Infrastructure

Network infrastructure needs will be addressed by the facility plan described in 1.2.

2.9.3 Controls, Data Acquisition, and Detector Integration

Controls, data acquisition, and detector integration needs will be addressed by the facility plan described in 1.3. Detailed plans and schedules are described in respective APS-U beamline controls documents, Equipment Specification Documents (ESDs), and Instrument Specification Documents (ISDs).

2.9.4 Data Management, Workflows, and Science Portals

The APS-U 3DMN feature beamline will use the data management, workflow, and science portal efforts described in 1.4. The APS Data Management System, the facility-wide software and hardware system for managing data and workflows, will integrate data collection and processing workflows, manage user permissions based on experiment groups, and enable user access to data. Globus tools will provide a science portal for viewing and accessing data and automating the re-processing of data after the allotted experiment time has ended. For the 3DMN APS-U feature beamline, workflows will provide a pipeline to automatically run the wire or mask scan Laue depth reconstruction processing software and view results.

2.9.5 Computing Infrastructure

Computing infrastructure needs will be addressed by the facility plan described in 1.5. Some local computing resources will be provided for on-the-fly data processing and experiment steering. Computing capacity for larger data processing tasks and for post-experiment processing and analysis will be provided by computing centers, including the ALCF and Argonne's Laboratory Computing Resource Center (LCRC). The APS Data Management System and Globus tools will be used to seamlessly integrate these resources.

2.9.6 Data Reduction and Analysis Software

Processing Laue micro- and nano-diffraction microscopy data generally consists of three main steps in the following order: depth reconstruction, peak searching and indexing, and q-space histogram generation. The depth reconstruction process generates new images corresponding to the scattering observed from a single depth. Peak searching and indexing finds all the peaks and indexes them to get the crystal orientation of a Laue pattern. With energy scans, a 1D or 3D histogram of intensity in q-space may also be generated.

The APS develops and maintains the *LaueGo* software for Laue depth reconstructions of wire scan mode data. The software is available as an open-source package (https://github.com/34IDE/LaueGo). It performs peak searching and indexing, depth reconstruction, and q-space histogram generation for wire scan data. Versions are available in both Igor and C. A CUDA GPU implementation is available to improve performance on GPU equipped workstations.

The current feature set and performance of *LaueGo* and the corresponding GPU implementation is adequate for today's needs. However, the estimated increase in overall data that will be generated at the 3DMN APS-U feature beamline necessitates improvements and advances in software and algorithms.

In order to improve data collection time in the APS-U era, the APS is developing coded-aperture scans that may replace the current wire-scans for obtaining depth reconstructions. Along with higher-performance implementations of the contemporary wire scan mode data, high-performance implementations of coded-aperture reconstruction methods are being developed. Table 2-21 summarizes Laue depth reconstruction data reduction needs, approaches, and status for the 3DMN APS-U feature beamline.

Table 2-21 Summary of Laue depth reconstruction data reduction needs, approaches, and status for the 3DMN APS-U feature beamline.

Capability	Algorithm / Software Requirement	Status
Reconstruct Laue microscopy wire scan data	Algorithms for Laue microscopy wire scan reconstruction	Done – APS Operations
	CPU and GPU software for Laue microscopy wire scan reconstructions	Done – APS Operations
	Parallel distributed-memory CPU and GPU software for APS-U era wire scan data	To do – APS Operations
Reconstruct Laue microscopy mask scan data	Algorithms for Laue microscopy coded-aperture reconstruction	In Progress – Past LDRD, APS Operations
	CPU and GPU software for Laue microscopy wire scan reconstructions	In Progress – APS Operations
	Parallel distributed-memory CPU and GPU software for APS-U era mask scan data	In Progress – APS Operations

2.9.7 Effort, Funding, and Collaborations

Network infrastructure to the edge of the 3DMN APS-U feature beamline will be provided by APS-U funding. All other network infrastructure is funded by APS Operations funding as described in 1.2. Facility-wide data management, workflows, and science portal effort is funded by APS Operations (approximately 2.5 FTE per year) and the Globus Services team as described in 1.4. Some local computing resources dedicated to the 3DMN APS-U feature beamline for on-the-fly processing and experiment steering will be provided from APS-U funding. Large-scale computing resources are provided by APS Operations funding, ALCF funding, and Argonne Laboratory Computing Resource Center (LCRC) funding as described in 1.5. A postdoc is dedicated to development of the new mask scan algorithm. The APS will dedicate appropriate software development effort from APS Operations funding.

The following LDRD funding was awarded to support this effort:

- *Coded Apertures for Depth Resolved Diffraction* (FY20) www.anl.gov

ARGONNE NATIONAL LABORATORY

- U.S. Department of Energy research facility
- Operated by the University of Chicago
- Midwest’s largest federally funded R&D facility
- Located in Lemont, IL, about 25 miles (40 km) southwest of Chicago, IL (USA)
- Conducts basic and applied research. in dozens of fields
- Unique suite of leading-edge and rare scientific user facilities