APS / CELS Town Hall

A common vision for the future

The Advanced Photon Source (APS) at Argonne National Laboratory (ANL) is a U.S. Department of Energy (DOE) Office of Science-Basic Energy Sciences (BES) scientific user facility. The core mission of the APS is to serve the scientific community by providing experiment facilities utilizing x-rays (beamlines) to allow users to address the most important basic and applied research challenges facing our nation.

Demands for increased computing at the APS are driven by new scientific opportunities, which are enabled by new measurement techniques, technological advances in detectors, multi-modal data utilization, and advances in data analysis algorithms. The priority for the APS is to further improve its world-class programs that benefit most from high-energy, high-brightness, and coherent x-rays. The revolutionized high-energy synchrotron facility of the APS-U era will increase brightness and coherence, leading to further increases in data rates and experiment complexity, creating further demands for advanced scientific computation. Estimates indicate that within the next decade the APS will generate hundreds of PBs of raw data per year and require access to over 50 PFLOPS of computing power for on-demand real-time computing.

Argonne is well poised to employ advanced computing to maintain a world-leading position in the synchrotron light source community. The APS has a world-class photon science program with a large and diverse user base, and the computing divisions within the Computing, Environment, and Life Sciences directorate (CELS) are home to world-leading supercomputing infrastructure and computational expertise. This colocation provides an unprecedented opportunity for collaboration in exciting and innovative areas and to explore how advanced computing and APS-U can together create the leading synchrotron light source instrument worldwide and enable discoveries that would otherwise not be possible.

The APS and the computing divisions within CELS are hosting a series of town hall meetings. The overarching goal is to **develop a common vision of the big challenges and opportunities associated with computing in the APS-U era, the capabilities needed to address those challenges and opportunities, and how the APS and CELS can work together to provide those capabilities**. The primary outcome is a vision and roadmap detailing work that must be undertaken over the next decade as well as near-term steps required to get started.

The APS / CELS Town Hall meetings will be held remotely. There will be a plenary session followed by six parallel breakout working groups focusing on: 1) new algorithms, math, and AI/ML, 2) scalable software tools, 3) workflow and orchestration, 4) the APS-CELS computing architecture, 5) sustainable and discoverable data repositories, and 6) networking. The output of each breakout working group should be a 5-page document (minimum) and accompanying presentation slide deck describing the vision and roadmap over the next decade for its respective topic along with near-term steps required to get started. Each breakout working group will meet in at least two sessions (more if needed). Prior to each breakout, short documents describing current capabilities will be solicited for brief presentations at the beginning of each breakout to provide participants with a baseline and foster discussion. Co-chairs for each breakout working group will lead discussions and are ultimately responsible for generating and reporting the group's output. The final report is due mid-February 2021.

APS / CELS Town Hall
Breakout Groups

New Algorithms, Math, and AI/ML

Co-chairs: Mathew Cherukara (APS), Stefan Wild (MCS)

Describe a vision and roadmap for the development of new methods to enable advanced processing, analysis, and interpretation of data from modeling/theory, simulations, and experiments and observations that drive scientific discovery in the APS-U era. Ongoing development of novel AI-enabled and computational imaging methods that leverage the increased coherent flux are further expected to increase computational demand for ML model training and x-ray data inversion. The roadmap should address topics as needed, including:

- R&D of advanced mathematics and algorithms for the inversion, abstraction, interpretation, analysis, and reduction of data
- R&D of AI/ML for experiment steering, rare event detection, and data reduction both at the edge and on HPC environments
- Methods to couple modeling/theory and simulations with experiment data to design new experiments and generate new knowledge
- Novel methods for multi-modal analysis
- Focus on high-priority APS techniques, including coherence driven and high-energy techniques: Ptychography, Bragg CDI, XPCS, serial crystallography, etc.

The output should be a 5-page document (minimum) and accompanying presentation slide deck.

Scalable Software Tools

Co-chairs: Arthur Glowacki (APS), Tekin Bicer (DSL/APS)

Describe a vision and roadmap for scalable software tools for APS-U era reduction, processing, analysis, interpretation, and visualization that operate across a multi-tiered computing landscape. The roadmap should address topics as needed, including:

- Development and management of scalable common software tools for inversion, abstraction, reduction, processing, analysis, interpretation, and visualization
- Operation across the computing continuum: edge, local, laboratory/campus, and high-end computing facilities
- Rapid development of software with HPC capabilities
- Role of domain specific languages (DSLs)
- Focus on high-priority light source techniques, including Ptychography, Bragg CDI, XPCS, serial crystallography, etc.
- Identify specific software packages that should be developed and/or scaled
- Consider current and next-generation computing platforms anticipated to be available at Argonne

The output should be a 5-page document (minimum) and accompanying presentation slide deck.

Workflow and Orchestration

Co-chairs: Faisal Khan (APS), Ryan Chard (DSL)

Describe a vision for coupling APS beamline instruments at facilities with edge, local, Laboratory, and large-scale computing resources to run data reduction, processing, analysis, interpretation, and visualization tasks. The roadmap should address topics as needed, including:

- Run data processing and analysis routines seamlessly across the continuum of computing resources
- Facile configuration and use by facility staff and users
- Enable both real-time and post-processing, e.g. via science portals
- Track workflow provenance and re-run workflows

The output should be a 5-page document (minimum) and accompanying presentation slide deck.

APS Computing Architecture

Co-chairs: Sinisa Veseli (APS), Ray Bair (CPS)

Describe a vision for the computing architecture needed in the APS-U era. How many and what types of computing resources are needed, and where should those resources be located? For example, should each APS instrument rely on entirely local computing, or should all data processing be performed on Aurora? The roadmap should address topics as needed, including:

- Schedule resources and launch processing jobs in a similar fashion among resources at the edge (FPGA, TPU, etc.), local, laboratory/campus, and high-end computing centers, e.g. via a common API
- On-demand resource utilization for real-time processing, e.g. via pre-emption and checkpointing, virtualization, the use of backfill, and new mechanisms
- Allow the allocation of resources based on higher-level availability
- Enable streaming data directly from instruments in addition to conventional file-based mechanisms

The output should be a 5-page document (minimum) and accompanying presentation slide deck.

Sustainable and Discoverable Data Repositories

Co-chairs: John Hammonds (APS), Ben Blaiszik (DSL)

Describe a vision to provide sufficient and sustainable storage, cataloging, searching, and publication services for knowledge dissemination and future scientific discovery in the APS-U era. The roadmap should address topics as needed, including:

- Deploy sufficient and sustainable data storage for data and results generated at the APS
- Easy-to-use catalogs and search capabilities
- Publication and DOI services for knowledge dissemination
- Standard ontologies for data and metadata capture
- FAIR and AI-ready data to enable future scientific discovery

The output should be a 5-page document (minimum) and accompanying presentation slide deck.

Networking

Co-chairs: David Leibfritz (APS), Raj Kettimuthu (DSL)

Describe a vision for the future of networking at the APS and between the APS, ALCF, LCRC, and other computing resources. The roadmap should address topics as needed, including:

- Describe APS networking needs over the next decade
- Instrument, facility, and laboratory/campus networking brought to a common baseline of capability and operability, e.g. bandwidth improvements, conduits, etc.
- Utilization of networking DMZ models across the facilities for transparent connections between instruments, computing, and storage within the fabric
- Realize an APS computing network, e.g. a virtual network encompassing multiple facilities or beamlines of similar types, together with computing resources required so they all seem to be within a single network.
- Innovative data sources and technologies, e.g., new instruments, 5G, 6G

The output should be a 5-page document (minimum) and accompanying presentation slide deck.