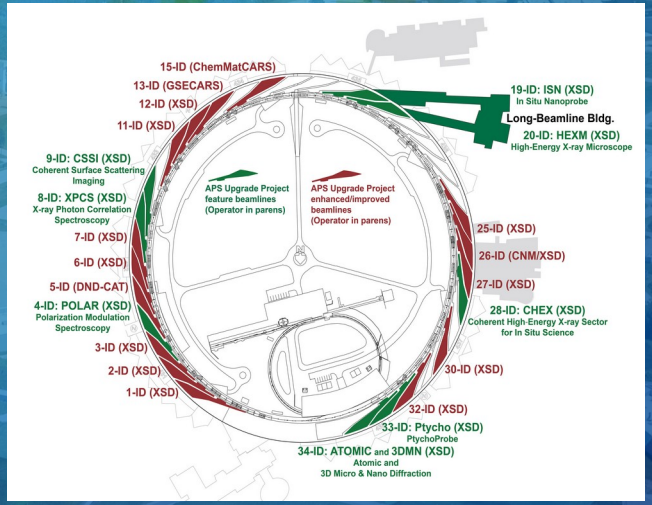
BEAMLINE DATA PIPELINE DESIGN EXPERIENCES UNDER APS UPGRADE

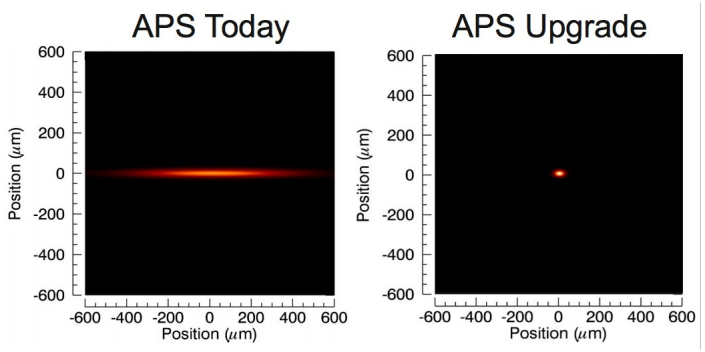


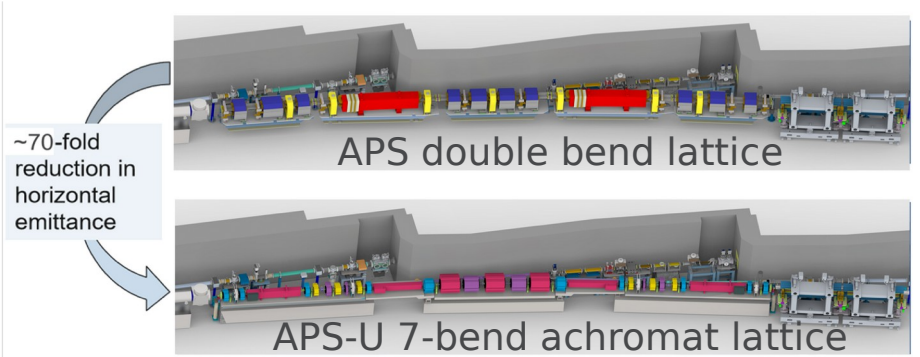
KEENAN LANG
(BEAMLINE CONTROLS)

APS-U

APS Light Source Improvements

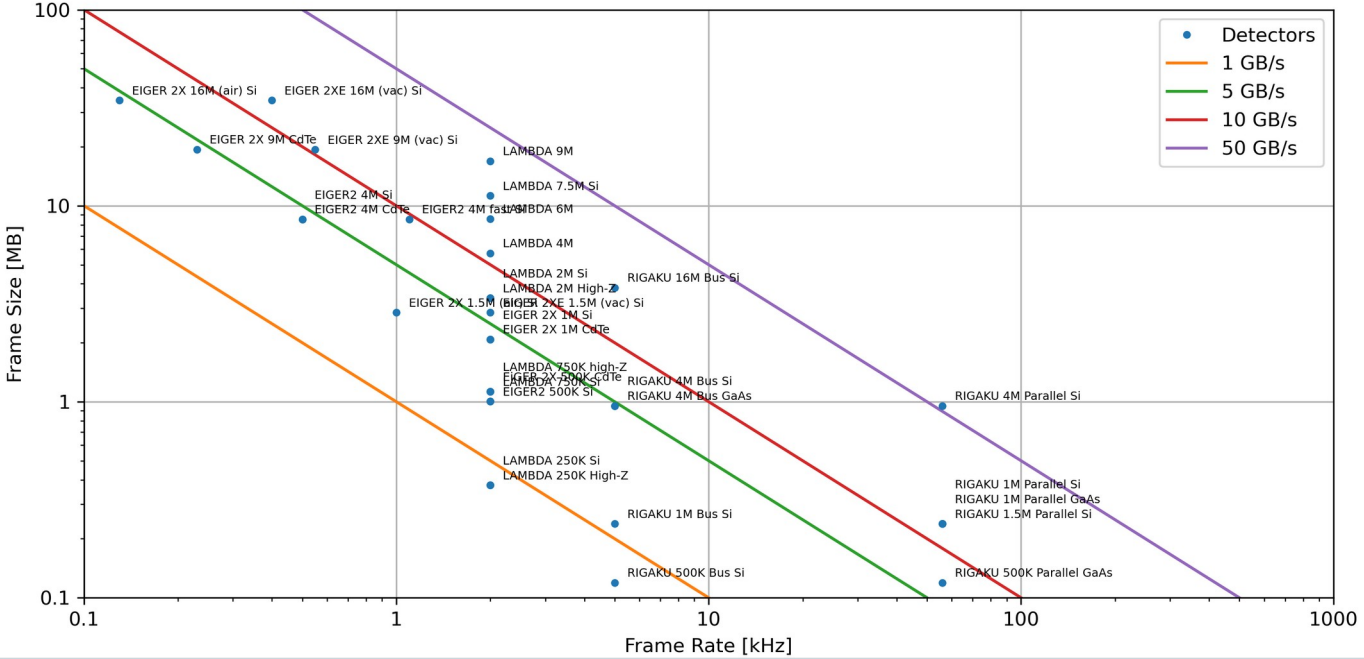
- X-ray detector capabilities are constantly improving: bigger frames, higher frame rates => more raw data
- APS Upgrade: Higher brightness => more x-rays can be focused onto a smaller area => more raw data in greater detail and less time
- Typical APS-U detectors support ~MB frames at ~kHz rates => GB/s data rates





Data Deluge: Next Generation Detectors Have Higher Spatial Resolution and More Photons Enable Faster Detector Operation

CANDIDATE APS-U DETECTOR DATA RATES



New High Data Rate Detectors Necessitate Software and Hardware Infrastructure Capable of Bridging Detectors, and Compute and Storage Resources

SCALE OF THE CHALLENGE

Multiple Order-of-Magnitude Increase in Demand For Computing Resources Over Next Decade

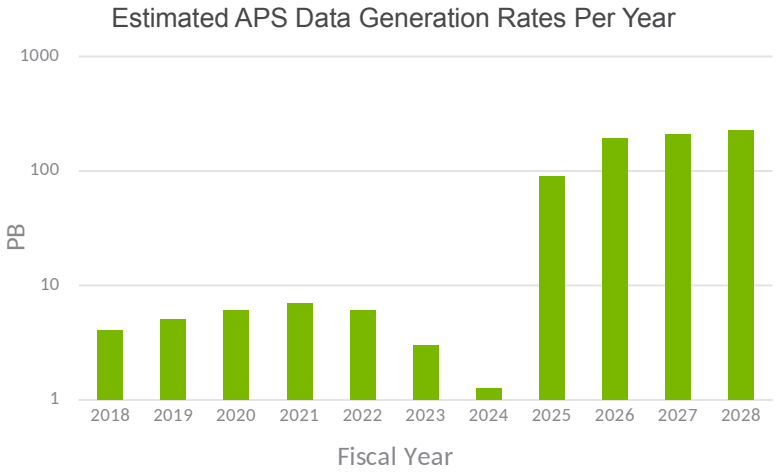
APS-U Era

- ~68 beamlines
- 9 feature beamlines and many enhanced beamlines

Over the next decade the APS will

- Generate 100s of petabytes (PBs) of raw data per year
- Require 10s of PFLOP/s of on-demand computing power for first pass data processing and reduction

For more information see the APS Scientific Computing Strategy document



Log Scale: Estimated data generation volumes per year at the APS. Data generation will drop temporarily due to the installation of new storage ring and beamline upgrades.

GUIDING PHILOSOPHY

- Solution will prefer existing APS and synchrotron light source community tools and technologies, for example
  - EPICS (including areaDetector and synApps)
  - Bluesky
  - APS Data Management System
  - Globus

- Entirely new developments will only be undertaken when necessary and no suitable existing solutions are found.

- We will deliver a working solution that is deployable at a facility-wide scale that makes best use of existing resources and support mechanisms.

REBUILDING THE BEAMLINE

BEAMLINE HARDWARE

- Moving away from VME for anything new

- Focusing on networked based devices
  - Motors: ACS MP4U
  - Comms: Moxa Serial Terminals
  - I/O
    - Labjack
    - Measurement Computing
    - TetrAMM
    - SoftGlueZynq

BEAMLINE SOFTWARE (IOCS)

- All IOCs using or upgraded to synApps 6-3 built against base-7.0.8
  - https://github.com/EPICS-synApps/assemble_synApps

- EPICS v7 / PVA
  - For AreaDetectors, still primarily using CA

- Control scripts allowing remote start/stop/restart features
- Automatic iocConsole logs
- iocshLoad-enabled scripts for ease of loading support
  - `iocshLoad("ADGeniCam.iocsh", "PREFIX=8idaSoft:, INSTANCE=flag1, CAM_IP=x.x.x.x, MODEL=AVT_Alvium_G1-510m, XSIZE=2464, YSIZE=2064, TYPE=UInt8")`

BEAMLINE SOFTWARE (UI)

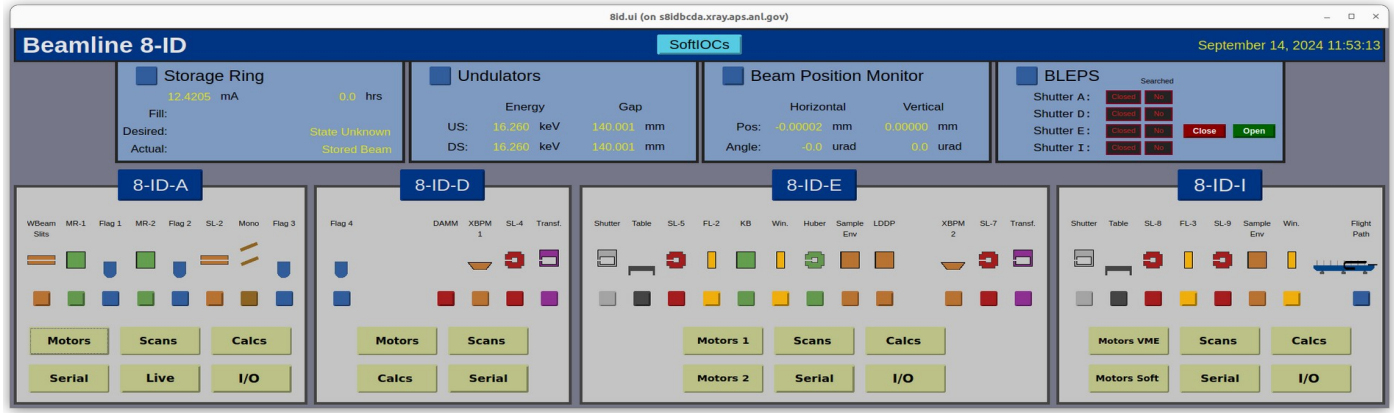
- Using CaQtDM (PSI)
  - will likely be moving to CSS-Phoebus in the future
- **Takeaway:** Use generation software for screens
  - Phoebusgen (https://github.com/als-epics/phoebusgen)
  - GESTALT (https://github.com/BCDA-APS/gestalt)

```
Hutch_Links: !VAnchor:HCenter:Grid
    margins: 0x0x0x15
    padding: 15
    aspect-ratio: 1.5
    repeat-over: "buttons"

    children:
        - !RelatedDisplay
            background: $BBC187
            geometry: 114x45
            font: -Liberation Sans Mono -Bold -9
            text: "{title:^14s}"
            links: "{links}"

Hutch_Optics: !HRepeat
    geometry: 15x20 x 0x0
    repeat-over: "optics"
    padding: 9
    children: [ !Apply:OPS_Slice ]
```

+

```
- name: "Flag 4"
  image: "/net/s8iddserv/xorApps/epics/ui/screens/img/flag.gif"
  links:
    - {label: "Flag 4 Motor", file: "motorx", macros: "P=8iddSoft:CR8-D1:,M=m1"}
    - {label: "AD Screen", file: "ADVimba", macros: "P=8iddSoft:,R=cam1:,C=FLIR_BFS_PGE_31S4C"}
    - {label: "AD Collect", file: "ADCollect", macros: "P=8iddSoft:,R=cam1:"}

- used: False

- used: False

- name: "DAMM"
  link-color: $A51D1E
  links:
    - {label: "Base", file: "motor2x1", macros: "P=8iddSoft:CR8-D1:,M1=m2,M2=m3"}

- name: "XBPM"
  name2: "1"
  image: "/net/s8iddserv/xorApps/epics/ui/screens/img/BD.gif"
  link-color: $B87333
```

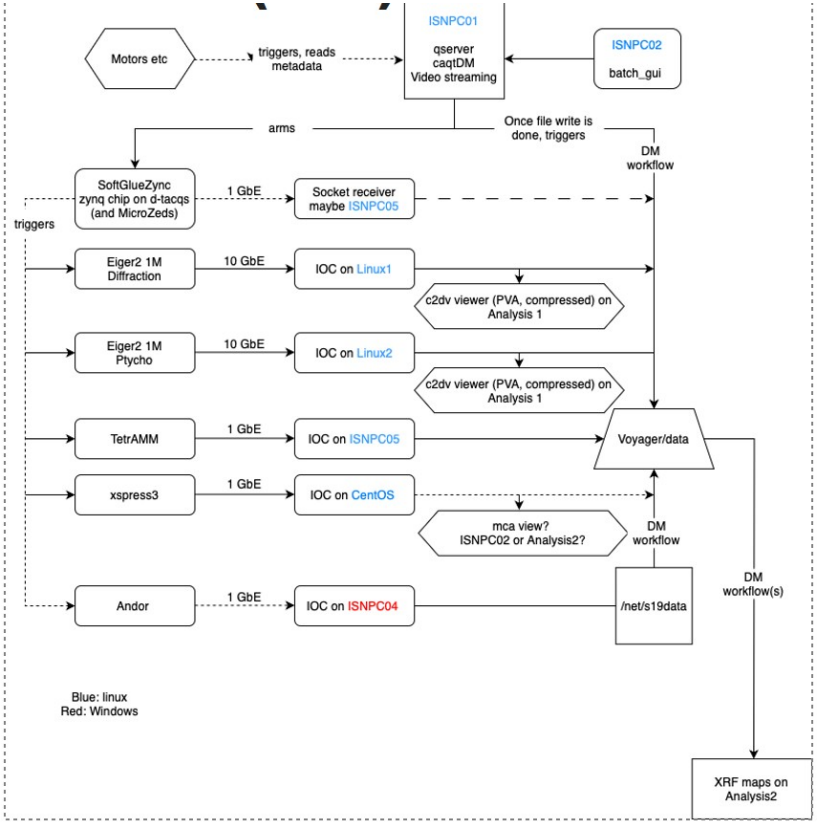
=



BEAMLINE DATA PIPELINE

DATA ACQUISITION

- Example Detectors:
  - Rigaku 3m
    - 3 megapixels, 2 bit @ 56,000hz
    - ~300 Gbps
  - Lambda 2m
    - 2 megapixels, 16 bit @ 2,000hz
    - ~50 Gbps
- Fast detector computers receive GPFS links
  - Write to shared file system
  - Accessible by Computing Facility
- Standardized on HDF5 file format
  - Flexible metadata structure
  - Allows us to connect multiple pieces of data together using references



DATA MANAGEMENT AND ANALYSIS

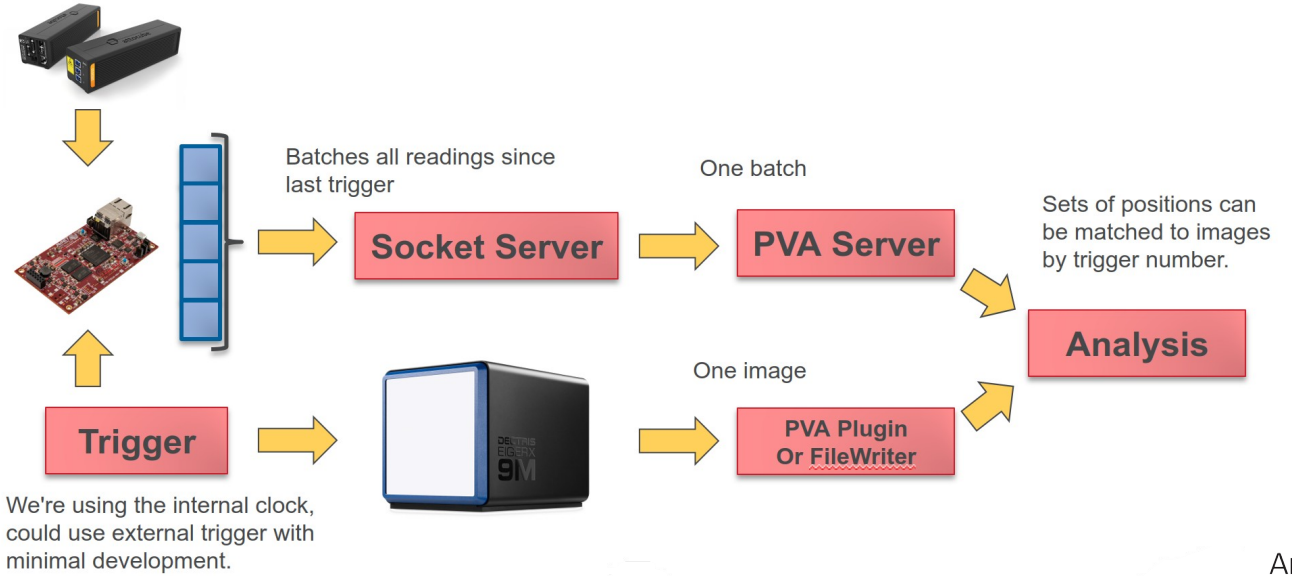
- Use Bluesky with Bluesky Queue Server to control experiments

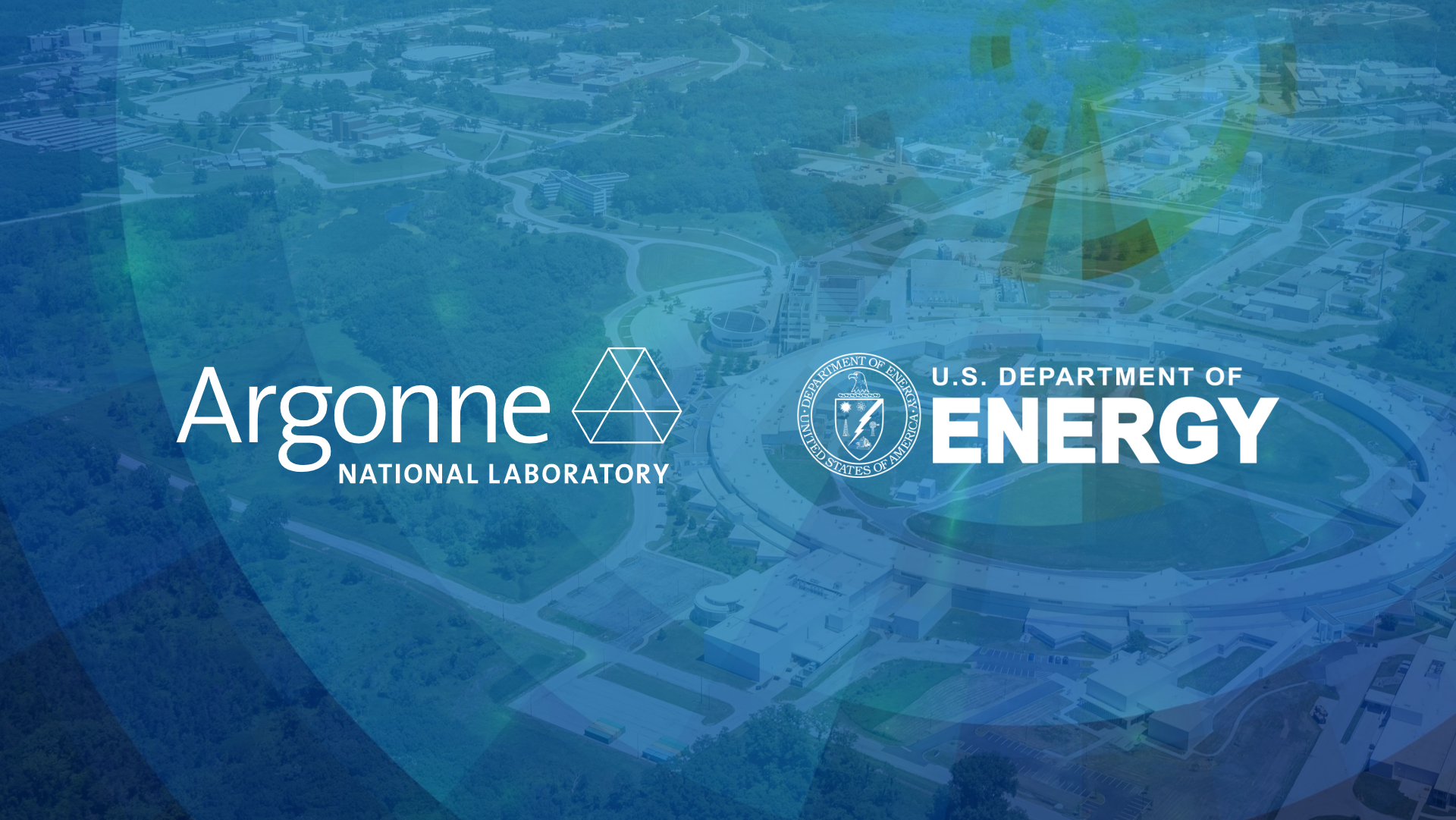
- Integrates with APS Data Manager Workflow API
  - https://www.aps.anl.gov/Science/Scientific-Software/DataManagement
  - JSON-based
  - Commands and execution flow that are run on remote computer
  - Queues time on polaris
  - Can be triggered manually or watch files/folders

```
{
    'name': 'tomocupy-polaris',
    'owner': 'bdp',
    'description': 'Tomocupy Workflow on Polaris',
    'stages': {
        '01-TOMOCUPY' : {
            'command' : 'source activate dm-gladier && python /clhome/BDP/DM/DEV/gladier/tomocupyGladierClient.py --f
            'outputVariableRegexList' : [
                'Flow Action ID: (?P<flowActionID>.*)',
                'URL: (?P<url>.*)',
                'Status: (?P<gladierStatus>.*)'
            ]
        },
        '02-MONITOR' : {
            'command': 'source activate dm-gladier && python /clhome/BDP/DM/DEV/gladier/checkStatus.py --flowID $flow
            'repeatPeriod': 5,
            'repeatUntil': '"$gladierStatus" == "SUCCEEDED" or "$gladierStatus" == "FAILED"',
            'maxRepeats': 99999,
            'outputVariableRegexList' : [
                'Status: (?P<gladierStatus>.*)'
            ]
        },
        '03-DONE'    : {
            'command': '/bin/echo "Job done"'
        },
    }
}
```

IN DEVELOPMENT

- Updating more analysis programs to support PVA as input
  - Using pvaPy as integration library (https://github.com/epics-base/pvaPy)
  - Live data analysis
  - Including streaming to and from our Computing Facility
- Interferometry Position Streaming and Association




Argonne
NATIONAL LABORATORY
U.S. DEPARTMENT OF
ENERGY