National Synchrotron Light Source II

DAMA Update

Stuart Campbell
Group Leader, Data Acquisition, Management and Analysis

28th March 2019

U.S. DEPARTMENT OF ENERGY | Office of Science
BROOKHAVEN NATIONAL LABORATORY

Outline

- NSLS-II Computing and Data Infrastructure Plans
- Collaborations with CSI / lightsources
- 5 Light Source Data Working Group
- BES User Facilities Data Management and Analysis Resource Needs (BES/ASCR Data Call)

Computing and Data Infrastructure Plans

What do we want to provide users ?

- Access to data
  - Via a remote session (SSH, Remote Desktop, Jupyter, Webservice) wherein the data doesn't leave BNL
  - Bulk transfer to home institution as files in appropriate format(s)
  - Local mirror at home institution

Data Infrastructure (current situation)

- Separate GPFS per beamline (except for AMX/FMX)
- NSLS-II Central GPFS
  - Intention is to sync beamlines → NSLS-II Central
  - Sync is not operational at the moment
- Experiencing performance and reliability issues

- No direct access from outside BNL
  - Users need a BNL (VPN) + NSLS-II Accounts

Data Infrastructure (Future)

- Beamline cache storage → Sync to Central Disks
- Users always ‘process’ from central disk to avoid write contention

- Question – Where should the ‘central’ disk reside ?
  - NSLS-II
  - BNL Scientific Data and Computing Centre (SDCC)
  - Commercial Cloud (Amazon, Google, Microsoft)

- Given resources and space constraints, SDCC looks the best option

Accessing Data

Moving data to SDCC as test, end goal to allow users to send data home

Now

- Testing Globus to move data to SDCC
- Proof of concept already successful
- 100 TB Test Allocation
- 500 MB/sec NSLS-II to SDCC test

6 months (start) - 18 months (finish)

- Assuming all tests are positive / policy decisions made
- Start deploying in production to beamlines

1+ years

- Public facing Databroker API
- 'Push' to cloud storage

Storage Costs

- 1 PB Tape Storage
  - CSI: $29,696 / year
  - Amazon Glacier: $50,331 / year

- 1 PB Fast Disk
  - CSI: $116,736 / year
  - Amazon S3: $301,990 / year

- Transfer Cost of 1 PB to AWS: $41,943 (using Amazon Datasync)

What do we want to provide users ?

- Access to data
  - Via a remote session (SSH, Remote Desktop, Jupyter, Webservice) wherein the data doesn't leave BNL
  - Bulk transfer to home institution as files in appropriate format(s)
  - Local mirror at home institution

- Data Analysis Processing Resources
  - Requires sufficient hardware - Memory, CPU, GPU
  - Provide software environments

Remote Access (Jupyter)

Allow users to analyze data remotely in a familiar interface

- SDCC are standing up a *jupyterhub* ‘service’

- Plan to test this service
  - with selected staff from CSX and CHX

- What computational resources should we use ?
  - SDCC
  - New Servers
  - Move unused servers from NSLS-II → SDCC

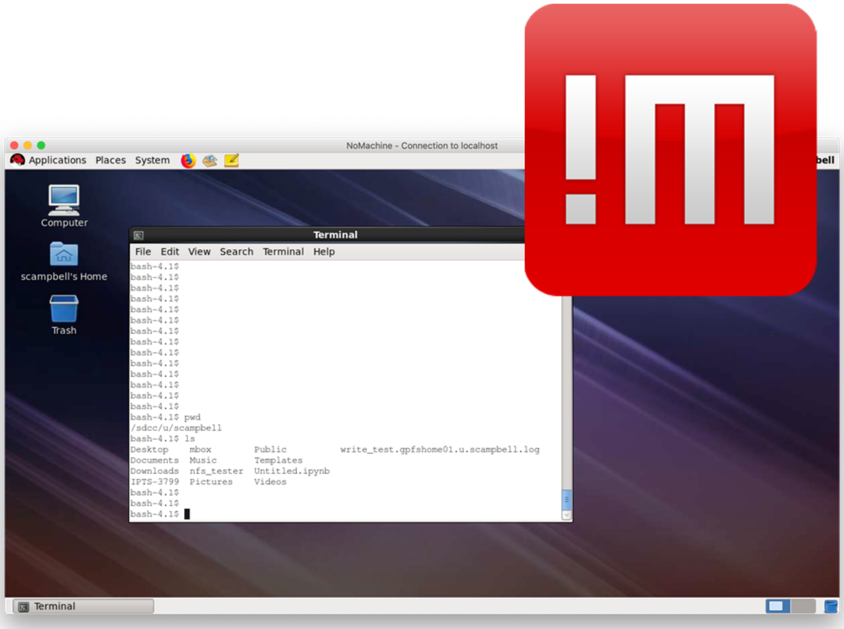
- Requires a single SDCC Account (No BNL or NSLS-II Accounts required)
- Works outside BNL Firewall (using OTP 2-factor)

Remote Analysis Access (Desktop)

Allow users to analyze data remotely using desktop applications

- Testing SDCC NX Remote Desktop Service
- Evaluate and standardize software deployment methods and technologies

- Question – Is it reasonable for users to only have home directories at SDCC and not NSLS-II controls ?



Building relationships to SDCC

- Hosting of two servers in their existing RedHat Enterprise Virtualization environment.
- Will be used to host test services for pilot system
  - Databroker server
  - MongoDB
- Used as development resources for DAMA (+others) effort
  - Build systems
  - Prototyping services
  - Reduces current use of AWS

RED HAT ENTERPRISE VIRTUALIZATION

Computational Resources

If we leverage SDCC expertise, it won't be free.

- What computing resources will we use ?
  - Existing SDCC capacity
  - Move unused machines over to SDCC
  - Buy new machines
- What will the cost from SDCC be for operating NSLS-II hardware?
- Need to bear in mind the hidden costs of doing things ourselves.

Using NERSC

- Cannot rely on NERSC for all computing infrastructure
  - On demand access
  - Operational cycles do not coincide

- Superfacility project
  - Deploy large scale computing and storage resources
  - Provide reusable building blocks for experimental scientists to build pipelines
  - Provide scalable infrastructure to launch services
  - Provide expertise on how to optimize pipelines

- Provide users access to large scale compute to support their experiments

Collaborations

Examples of Collaborations with CSI

Overall 12+ projects

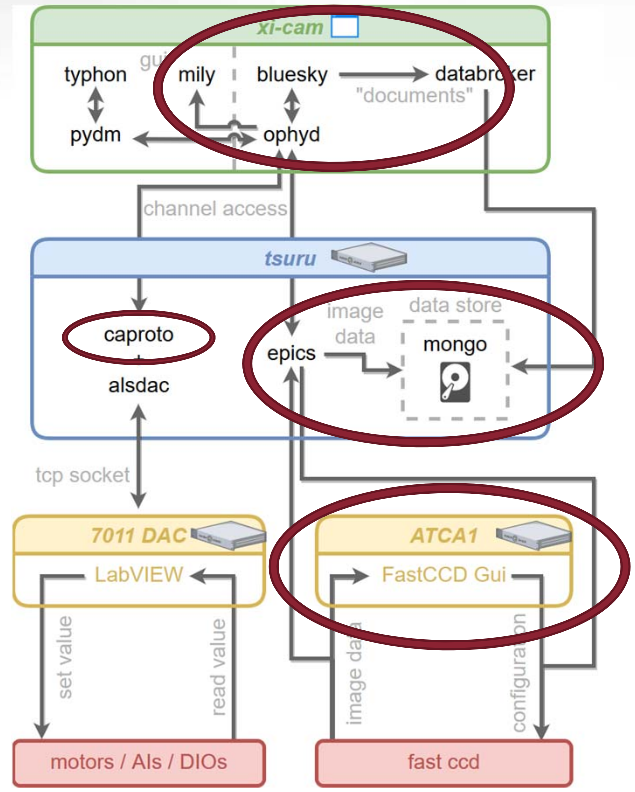
- Machine Learning assisted Materials Discovery
- Visualization toward White-box Machine Learning for Multi-modal Image Analysis
- HXN Pytchography GUI
- High-Performance Toolkit for X-ray Diffraction Forward Simulations and Data Analysis at HXN
- Provenance-enabled sample measurements for multi-modal analysis with ISS and XPD
- DiffPy-CMI – a software toolbox for structure analysis by Complex Modeling method
- Dynamic visualization and visual analytics for scientific data
- Applications of Visualization
- Machine-learning for Autonomous X-ray Scattering
- EPICS network design
- SDCC workshop on supporting photon science
- Data Management using Rucio

Collaboration between Lightsources

- Building a community around software tools
  - ALS (+ CAMERA), APS, LCLS and NSLS-II
- Desire to adopt common solutions
  - e.g. Single XPCS package

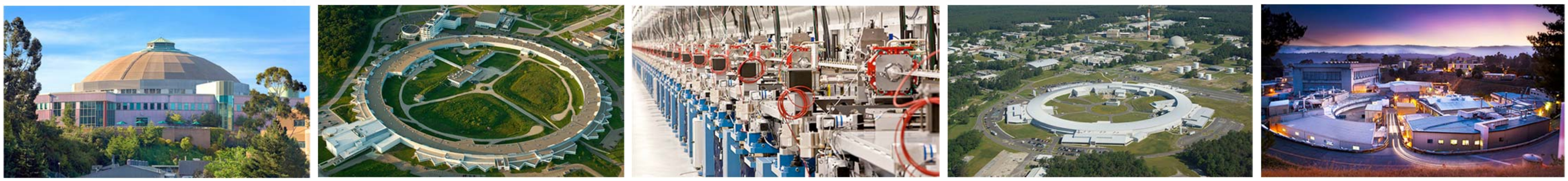
- The COSMIC beamline at the ALS for Soft X-ray XPCS has adopted the NSLS-II DAMA software for next generation experimental control and data management.

- Leveraging efforts at other facilities rather than re-inventing the wheel



Data Working Group

The 5 Light Sources Data Working Group



- Serves as a resource for the 5 US DOE BES light source (ALS, APS, LCLS, NSLS-II, SSRL) directors
- Provides information and recommendations on working together in the areas of data and computing
- Agreed to adopt a common data policy template between the 5 facilities by October 2018; data will be kept for at least 1 year

Members

Nicholas Schwarz (*Chair*)
Group Leader, Scientific Software Engineering & Data Management
Advanced Photon Source (APS)

Amedeo Perazzo
Division Director, Controls and Data Systems
Linac Coherent Light Source (LCLS)

Stuart Campbell
Group Leader, Data Acquisition, Management, and Analysis
National Synchrotron Light Source II (NSLS-II)

Alex Hexemer
Program Lead, Computing
Advanced Light Source (ALS)

BES User Facilities Data Management and Analysis Resource Needs (BES/ASCR Data Call)

Increase in Advanced Computing Needs

Multiple order-of-magnitude increase in demand for computing resources over next decade

By 2028, the BES Light Sources will

- Generate in the **exabyte (EB) of data** range per year
- Require **tens to 1,000 PFLOPS of on-demand computing** resources per year
- Utilize **billions of core hours** per year

Unified solutions across the facilities are required in order to leverage efficiencies of scale, and to provide facility users with the ability to easily and transparently manipulate data across the complex

The facilities do not currently have the operations resources to address these needs

High-Priority Shared Needs

Mission requires computing advances in four main areas

1. **Data management and workflow tools** that integrate beamline instruments with computing and storage resources, for use during experiments, as well as facile user access for post-experiment analysis
2. **Real-time data analysis capabilities** to significantly reduce the data volumes and provide feedback during experiments to improve data quality and to drive the direction of ongoing measurements; the application of advanced machine learning algorithms and the integration of simulations and model-based approaches will allow automated steering of data collection
3. **On-demand utilization of super-computing environments** to enable quasi real-time data processing
4. **Data storage and archival** resources to house the continually increasing amounts of valuable scientific data produced by the BES Light Sources

Data Institute Bridging BES and ASCR

Distributed Initiative Across the Five Light Sources

- Goal is to develop
  - Data management and workflow tools / streaming from the beamline to computing
  - Adaptive and scalable automated experiment steering capabilities
  - Development of novel machine learning, artificial intelligence, and other automated segmentation and registration techniques and approaches
- Build on and institutionalize the progress made by ASCR-BES initiatives, such as the ExaFEL grant and CAMERA
- **Complementary to CAMERA**, which focuses on applied math and algorithms, this would deliver production level, customized code for workflow, data analysis, and data management for the over 200 instruments within the complex

Summary

- Continue and expand collaborations with other light sources, CAMERA and NERSC
- Continue to build a relationship and trust with SDCC (and CSI)
- Use SDCC for storage and scientific compute

Thank you

Questions?

Acronyms

- ASCR – DOE's Office of Advanced Scientific Computing Research
- BES – DOE's Office of Basic Energy Science
- CAMERA – Center for Advanced Mathematics for Energy Research Applications
- CSI – Computational Science Initiative
- GPFS – IBM's General Parallel File System (now called Spectrum Scale)
- SDCC – Scientific Data and Computing Center

Additional Slides

BES Light Source Data Generation Estimates

Year	Facility				
	ALS	APS	LCLS/LCLS-II	NSLS-II	SSRL
2021	3 PB	7 PB	30 PB	42 PB	15 PB
2028	31 PB	243 PB	300 PB	85 PB	15 PB

Estimated data generation rates per year at the BES Light Sources. At the ALS and APS, data generation will stop during 2025 and 2023, respectively, due to installations of new storage rings. Aggregate data generation across the BES Light Sources will approach the exabyte (EB) range per year by 2028. The LCLS volumes assume that the data will be reduced on-the-fly, in some cases by several orders of magnitude, through reduction techniques that will depend on the experiment and that will range from vetoing to features extraction to lossless and lossy compression.
The differences in data generation rates across the facilities depend on the number, rate and resolution of the detectors at each instrument which in turn depend on factors like the brightness of the source and the actual requirements of the experimental technique specific to that instrument.

Data Storage Costs

There is currently no single DOE complex tape-based storage system with the capacity for the sustained storage of anticipated data volumes from the BES light sources.

The location of storage systems for the facilities could be local (either at the facility or host laboratory), combined into shared locations (single or east/west locations), or stored on the commercial cloud.

The cost to store the aggregate total of data from the BES light sources (mentioned previously) will continue to increase.

The estimated incremental cost* per year to store one year's worth of data from all of the BES light sources will be:

- By 2021 - DOE complex: $1.6M/yr to $5.5M/yr | Commercial cloud: $3.9M/yr to $13.5M/yr
- By 2028 - DOE complex: $9M/yr to $30M/yr | Commercial cloud: $22M/yr to $75M/yr

** In 2018 dollars, unless disruptive solutions are found to the evolutionary path this assumes.*

Light Source On-Demand Computing Estimates

Year	Facility				
	ALS	APS	LCLS/LCLS-II	NSLS-II	SSRL
2021	0.1 PFLOPS	4 PFLOPS	1 - 100 PFLOPS	2.5 PFLOPS	< 1 PFLOPS
2028	30 PFLOPS	50 PFLOPS	1 - 1,000 PFLOPS	45 PFLOPS	< 1 PFLOPS

Estimated PFLOPS of on-demand computing resources required by each of the BES Light Sources by 2021 and 2028. Compute jobs requiring < 10 PFLOPS are common and best run on local resources; compute jobs requiring > 10-20 PFLOPS are best suited to run at an high-end computing facility.

For comparison:

NERSC CORI is 30 PFLOPS

NERSC Perlmutter (2020) will be 90+ PFLOPS