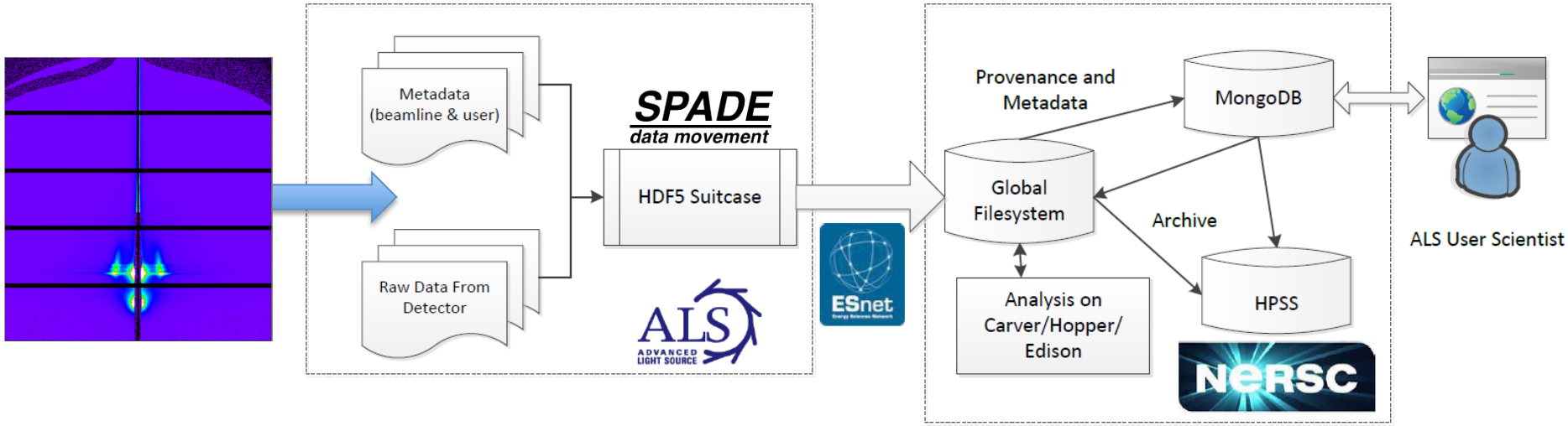

BERKELEY LAB
LAWRENCE BERKELEY NATIONAL LABORATORY
U.S. DEPARTMENT OF ENERGY
Real-time Streaming Analysis for BES User Facilities
Craig E. Tull, PhD
LBNL Computing Research Division
STREAM 2016: Streaming Requirements, Experience, Applications and Middleware Workshop
March 22, 2016 @ Tysons, VA

BES Facilities serve 16,000 users/yr in Materials, Biology, Energy, Medicine, …

- Virtually every area of science and technology are taking advantage of Lightsources, etc.
- The ALS user base is expanding to new areas and includes more 1st timers who cannot afford long investment in learning hardware & software.
- Data volumes are exploding:
  - Lightsources are getting brighter
  - Detectors are getting faster
  - Beamlines are automating
- New mathematical techniques, new architectures, and even new paradigms (eg. Neuromophic, Quantum) are being developed or researched.

SPOT Suite: Integration of ALS, ESnet, and NERSC into a proto-super-facility.



- **Computing Research Div., Advanced Light Source, Material Science Div., ESnet, NERSC**
- Real-time processing needed for: Time-resolved, in-situ experiments & Data Quality Assurance

Daya Bay "Real-Time" Processing

- Use the inverse β-decay reaction:

$\bar{\nu}_e + p \rightarrow e^+ + n$ (prompt signa

~180μs

$\rightarrow + p \rightarrow D + \gamma($

$\rightarrow + Gd \rightarrow Gd^*$

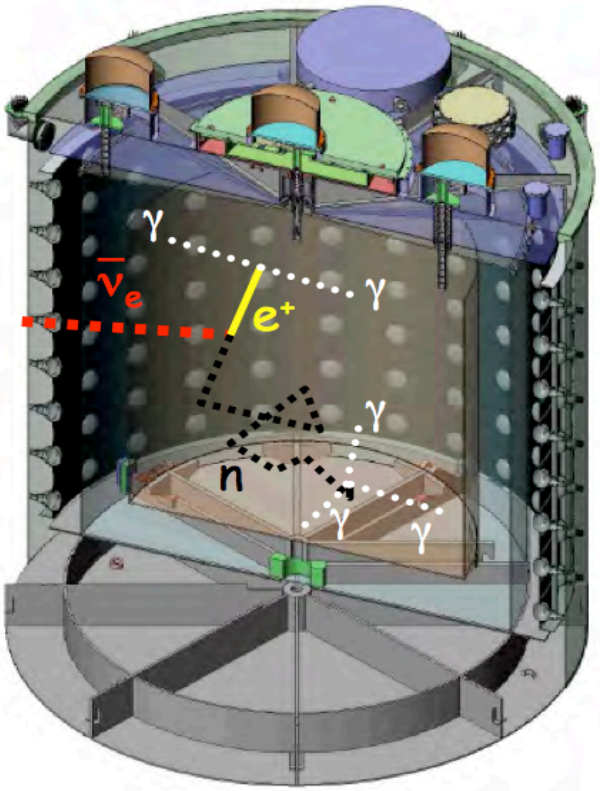
~30μs

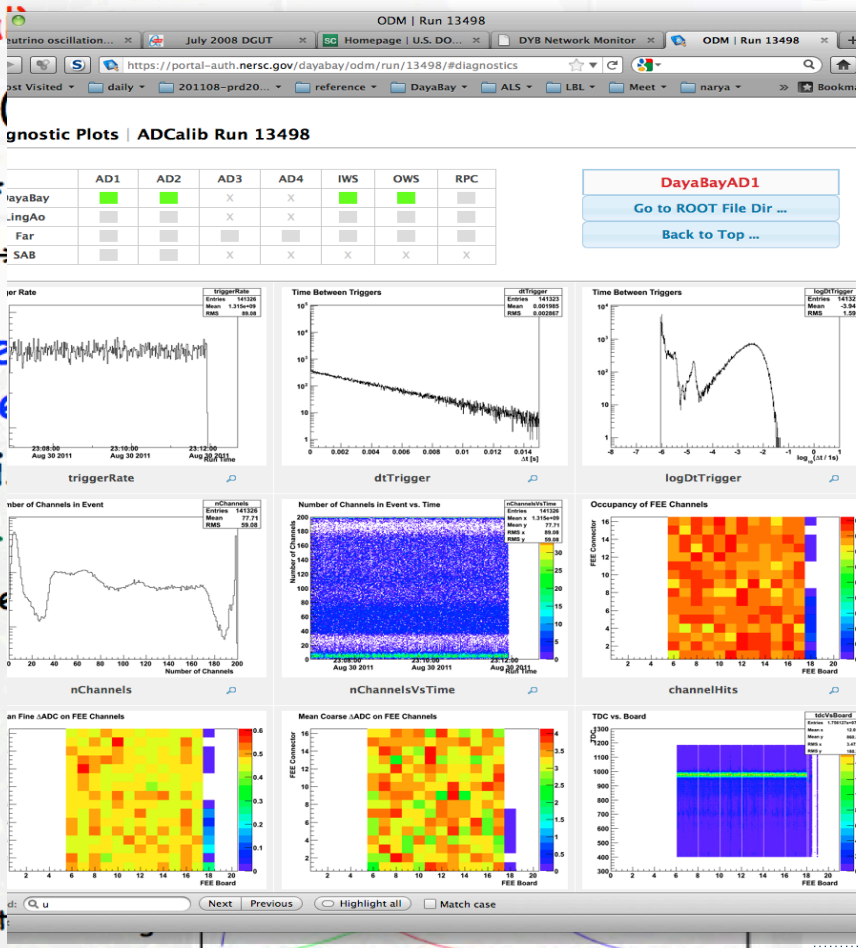
for 0.1% Gd

- Time- and ene
  tool to suppre
- Energy of $\bar{\nu}_e$ i

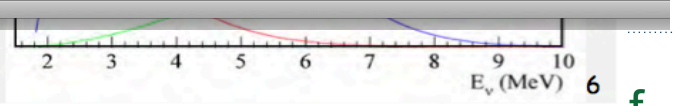
$E_{\bar{\nu}} \approx T_{e+} + T_n +$

10-40 ke



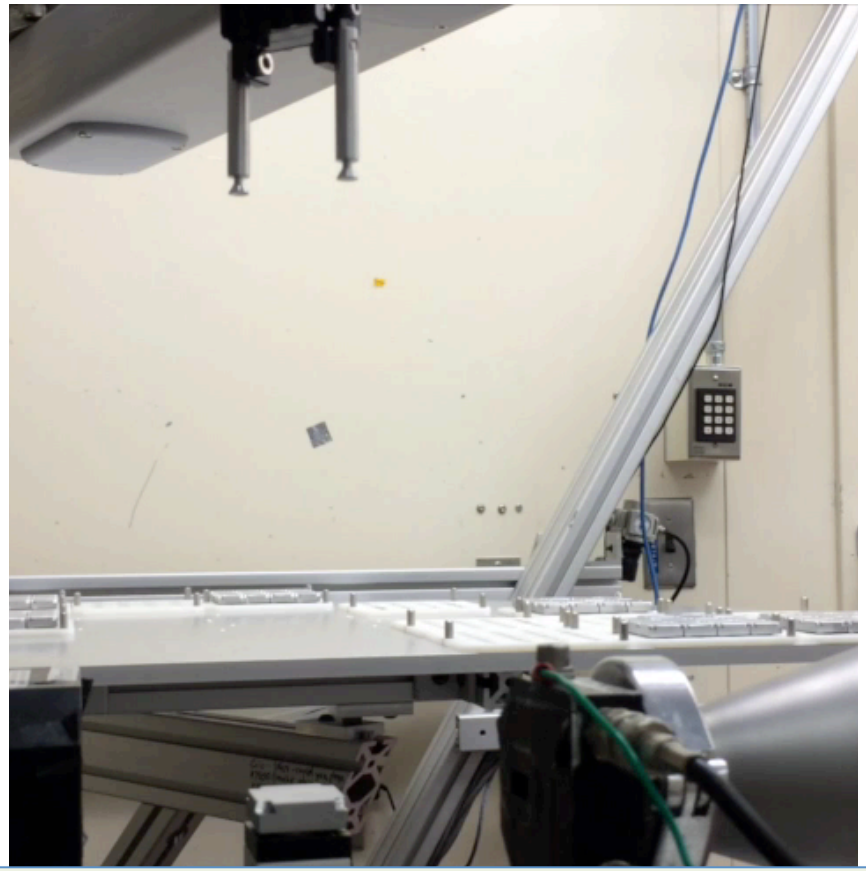


From Bemporad, Grat



Office of Science

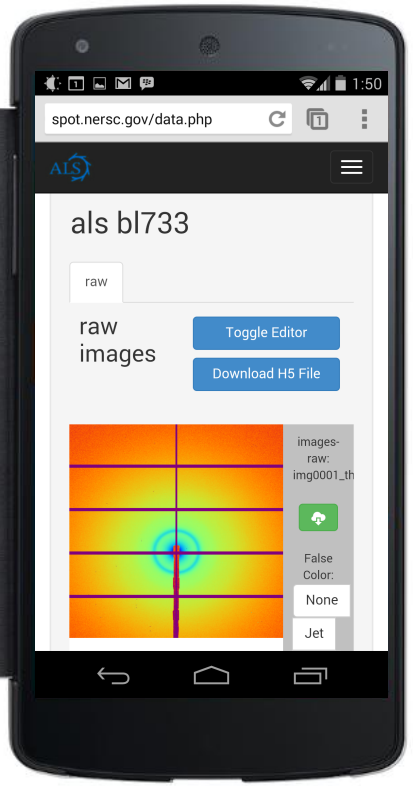
Remote experiments now a reality.



25mar2014:

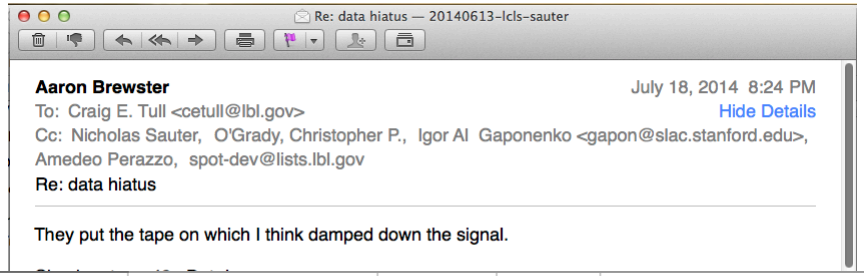
UK scientists conduct remote experiment using new BL 7.3.3 robot and SPOT.

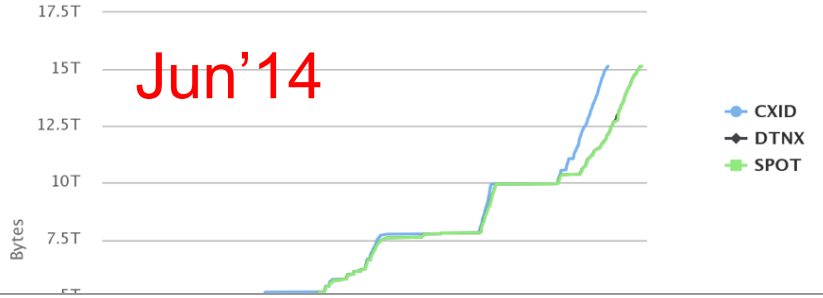
Able to assess experimental data on train to Zurich via mobile interface.



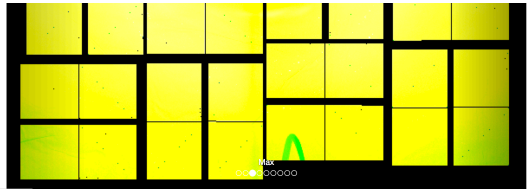
From: Alessandro Sepe as2237@cam.ac.uk -- **Actually, I did not feel any difference between a standard beamtime and this NERSC remotely accessed beamtime, which is quite an extraordinary result.**

“SPOT was like an extra pair of hands working in the background.” – N.Sauter





Date	Experiment	Files	Size	
Jun-14	cxid9114	4536	76	Dry Run
Jul-14	cxid9114	4544	79	Simulated Realtime
Jul-14	xppe0314	1555	24	Photosystem II - XFEL pump probe nano-crystallagraphy
May-15	cxid9114	560	9	Photosystem II - XFEL pump probe nano-crystallagraphy
May-15	cxih8015	4411	76	Photosystem II - XFEL pump probe nano-crystallagraphy
Jun-15	xppi6115	862	3	Photosystem II - XFEL pump probe nano-crystallagraphy
Nov-15	amok5415	931	26	Fluctuation X-Ray Scattering (XFS) of biological nanostructures



Real-time access to ASCR HPC changes the way scientists imagine the facility.

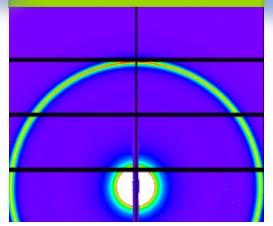
- "I've been having more users bring up the idea of running experiments with a 'digital twin'. Take an initial data set, send to HPC, create a 3d model of their sampleas input to simulation, which they start right away and run as they run experiments at the beamline. Matching up and comparing the results of the simulation with the results of the experiment."
- 1. Simulating flow and reactions underground at the pore scale: Jonathan Ajo-Franklin (ESD, LBNL) David Trebotich (CRD, LBNL): http://ascr-discovery.science.doe.gov/2014/09/pore-samples/
- 2. Simulating material failure in realistic conditions Rob Ritchie (MSD, LBNL), Michael Czabaj (UofU) http://newscenter.lbl.gov/2012/12/10/space-age-ceramics-get-their-toughest-test/
- 3. Simulating heat shield ablation Nagi Mansour, NASA http://www.nas.nasa.gov/publications/articles/feature_TPS_panerai.html

GISAXS Super Facility Demo Data Flow

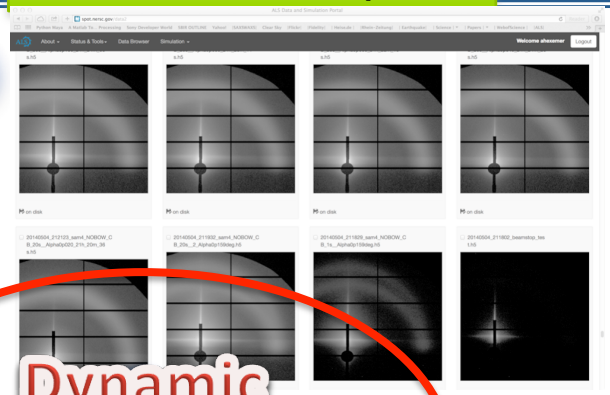
Transfer to NERSC

Static, Production Spade

On-the-fly calibration, processing



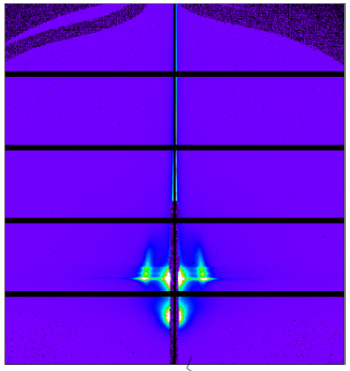
Real-time access via web portal



Combining:
GIXSGUI, dpdak + …

Dynamic Spade + Globus Online

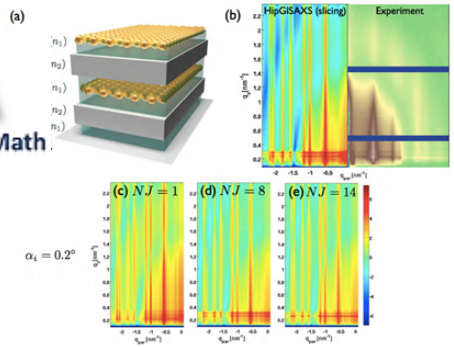
Data collection



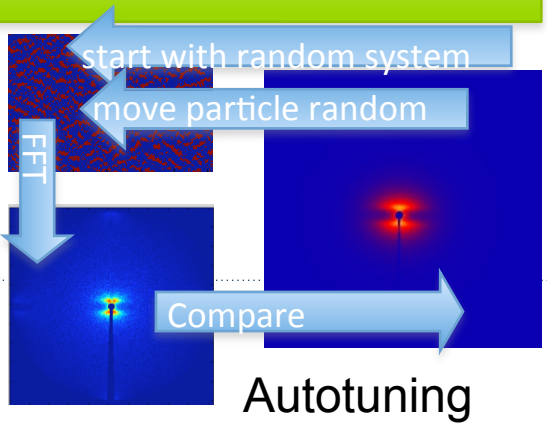
Analysis and mo[illegible]ercomputers:

HipGISAXS simulation

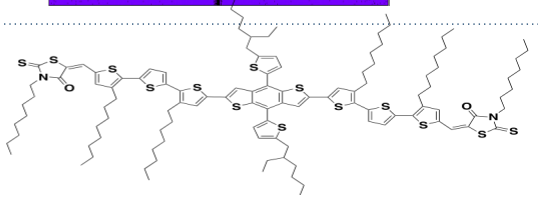
HipRMC fitting



CAMERA Applied Math



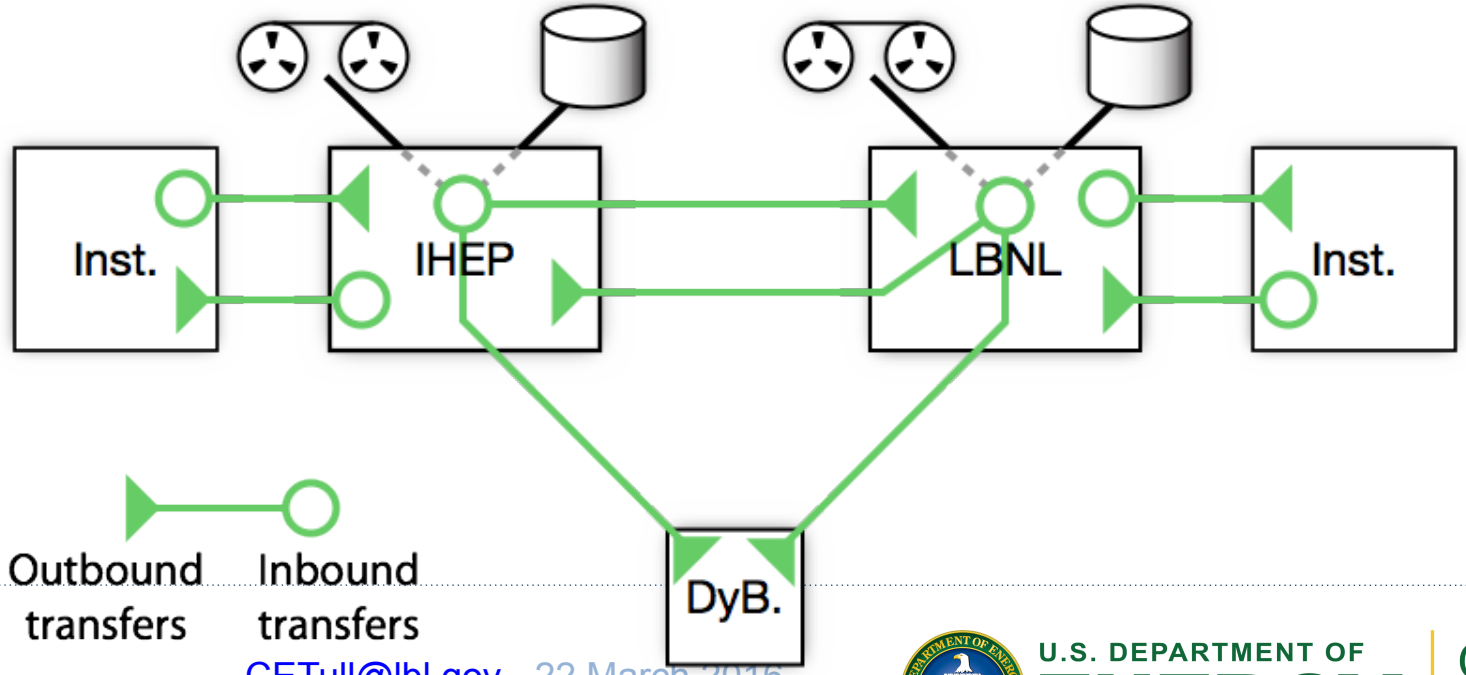
Autotuning



OLCF Oak Ridge Leadership Computing Facility

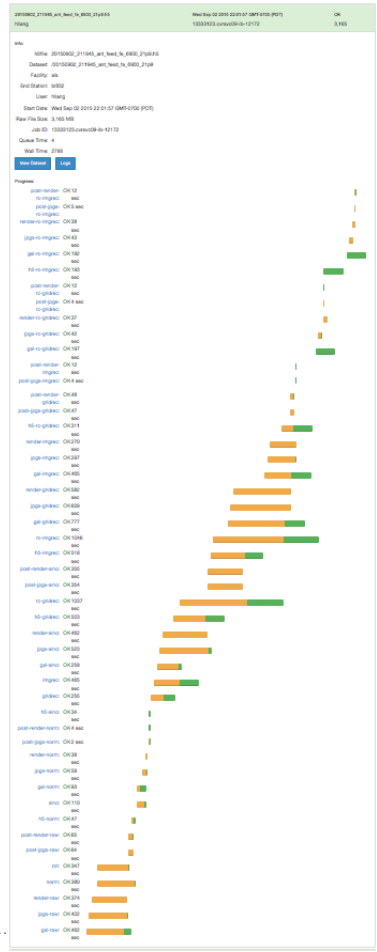
SPADE used for production orchestration of network data movement

- SPADE developed in IceCube, used in Daya Bay & ALS
- Underlying protocols: scp, bbcp, gridftp, Globus Online, RDMA?
- Highly Configurable: push, pull, relay, local
- Integrated warehouse, catalog, monitoring; Highly instrumented



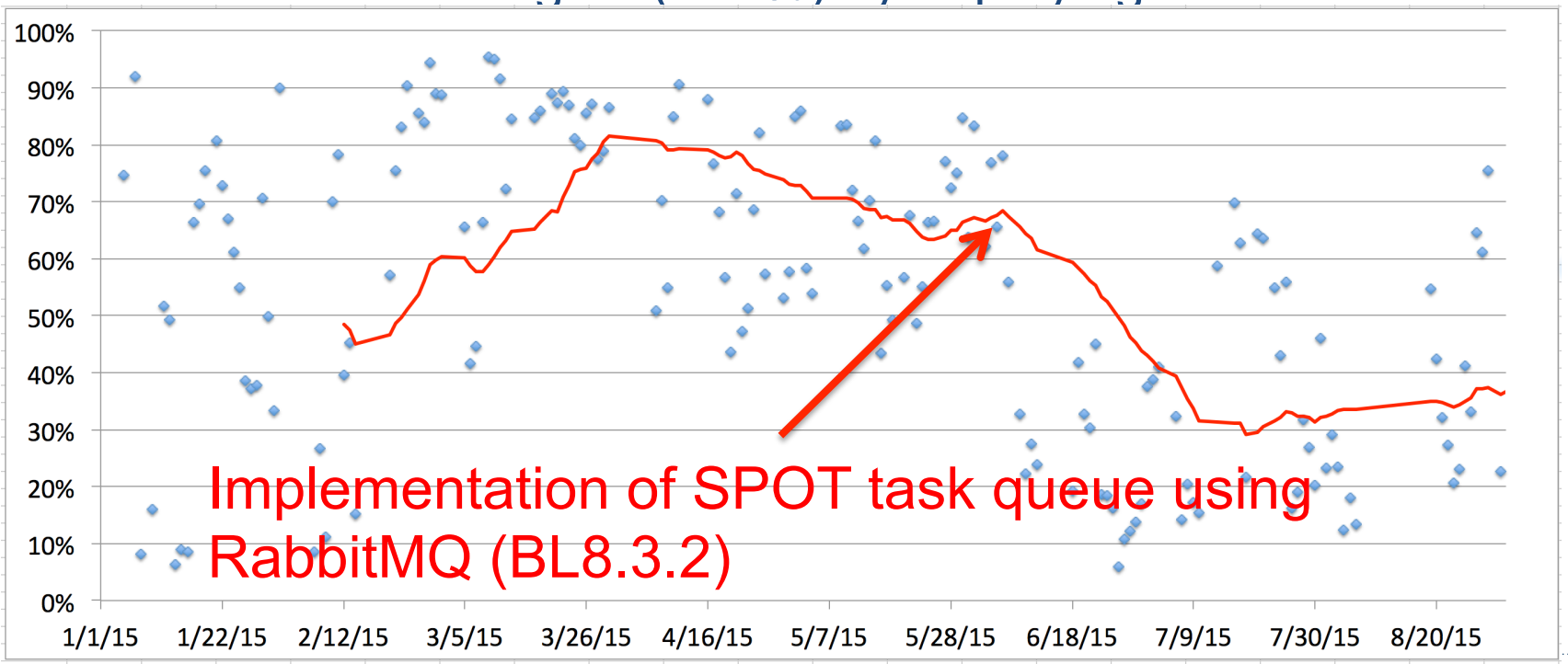
X-SWAP: Time-sensitive processing on a Queue-based facility (NERSC)

- Tomography workflow on NERSC = DAG with 48 graph nodes
- NERSC batch queue wait time penalty was significant.
- Implemented RabbitMQ worker node model (summer 2015)
  - Queue penalty dropped by 50% or more
  - Can be optimized by deploying more workers
  - Provides additional robustness for machine failures (1500 jobs automatically resumed after 1-day NERSC outage)
  - Adopted this same technique to Daya Bay



X-SWAP: Instrumented NERSC workflow provides lever for optimizing throughput.

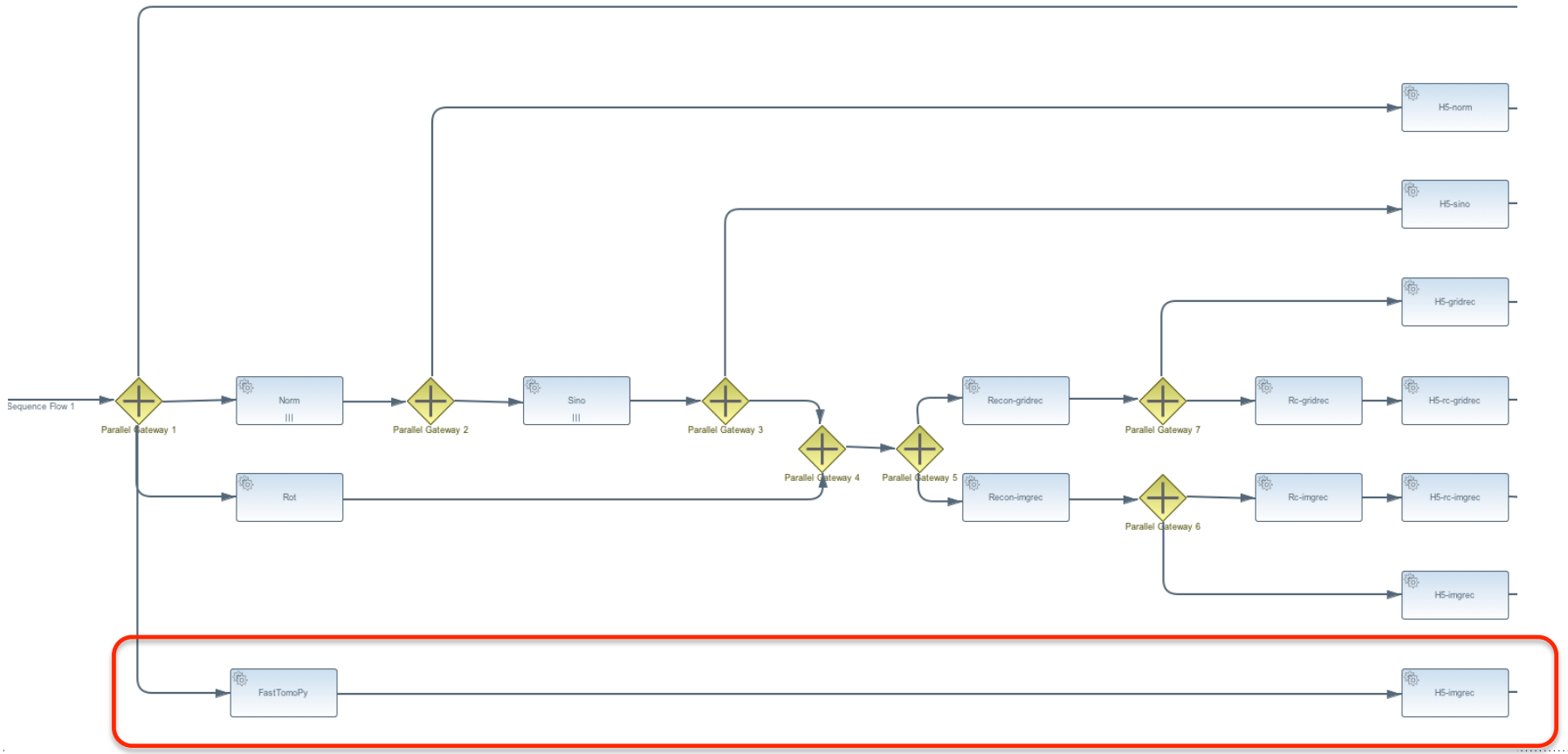
- ALS beamline 8.3.2 (Tomography) queue wait time dropped from 60-70% to 30% of total turn-around time for jobs.
- We can see we will gain (<20%) by deploying more workers.



Experiments' and Facilities' realtime streaming requirements vary.

- Overnight (eg. telescopes, day shift experiments)
  - Plan campaign for next shift/day
- Hourly (eg. stable, long-term HEP experiments)
  - Detect problems; Maintain steady-state data taking
- Minutes (eg. time-resolved, in-situ experiments)
  - Follow experiment evolution; Verify data quality
- "Instantaneous" - like a "software" microscope
- BES Experiments are "new" every day
- Understanding, instrumenting, and modeling the scientific workflow are powerful tools in assessing trade-offs between speed and quality of streaming data analysis.

In a complex workflow, not all paths are of equal value for streaming feedback.

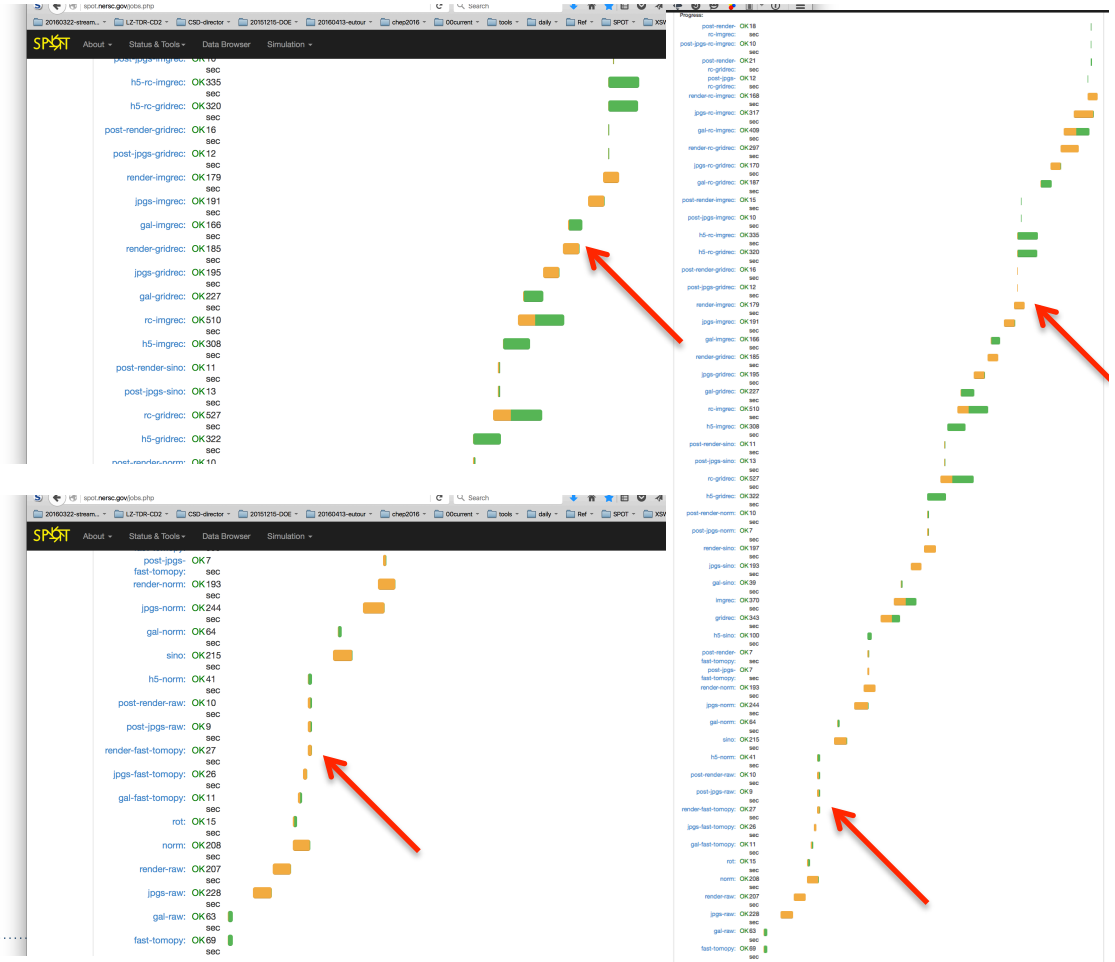


CETull@lbl.gov - 22 March 2016

U.S. DEPARTMENT OF ENERGY | Office of Science

X-SWAP: Instrumenting and modeling to minimize workflow branch latency.

- SPOT Tomographic processing is a DAG of 54 graph nodes.
- Fast feedback on a small subset of data is sufficient for QA.
- Introduce a new DAG branch (Fast TomoPy)
- First feedback reduced from ~16 minutes to ~2
- Trade-off quality & completeness.



CETull@lbl.gov - 22 March 2016

U.S. DEPARTMENT OF ENERGY | Office of Science

Summary

- Real-time processing important for QA, in-situ time-resolved experiments, and for experimental steering.
- The meaning of “real-time” varies with scientific goals.
- Optimizing overall throughput important. But, analysis of workflows yield opportunities to trade off fast user feedback with quality/completeness of results.
- Pairing real-time simulations with real-time analysis increasingly needed to maximize scientific insight.
- X-SWAP: Complex, distributed workflows need instrumentation and modeling to understand and optimize.
- DEDUCE: Need to inject decision-making into data workflows.