Data Analytics for Free Electron Lasers

February 27th 2018
Simons Institute – Applications in the Natural Sciences and Physical Systems Workshop

Amedeo Perazzo
LCLS Controls & Data Systems Division Director

U.S. DEPARTMENT OF ENERGY
Office of Science
SLAC NATIONAL ACCELERATOR LABORATORY

Outline

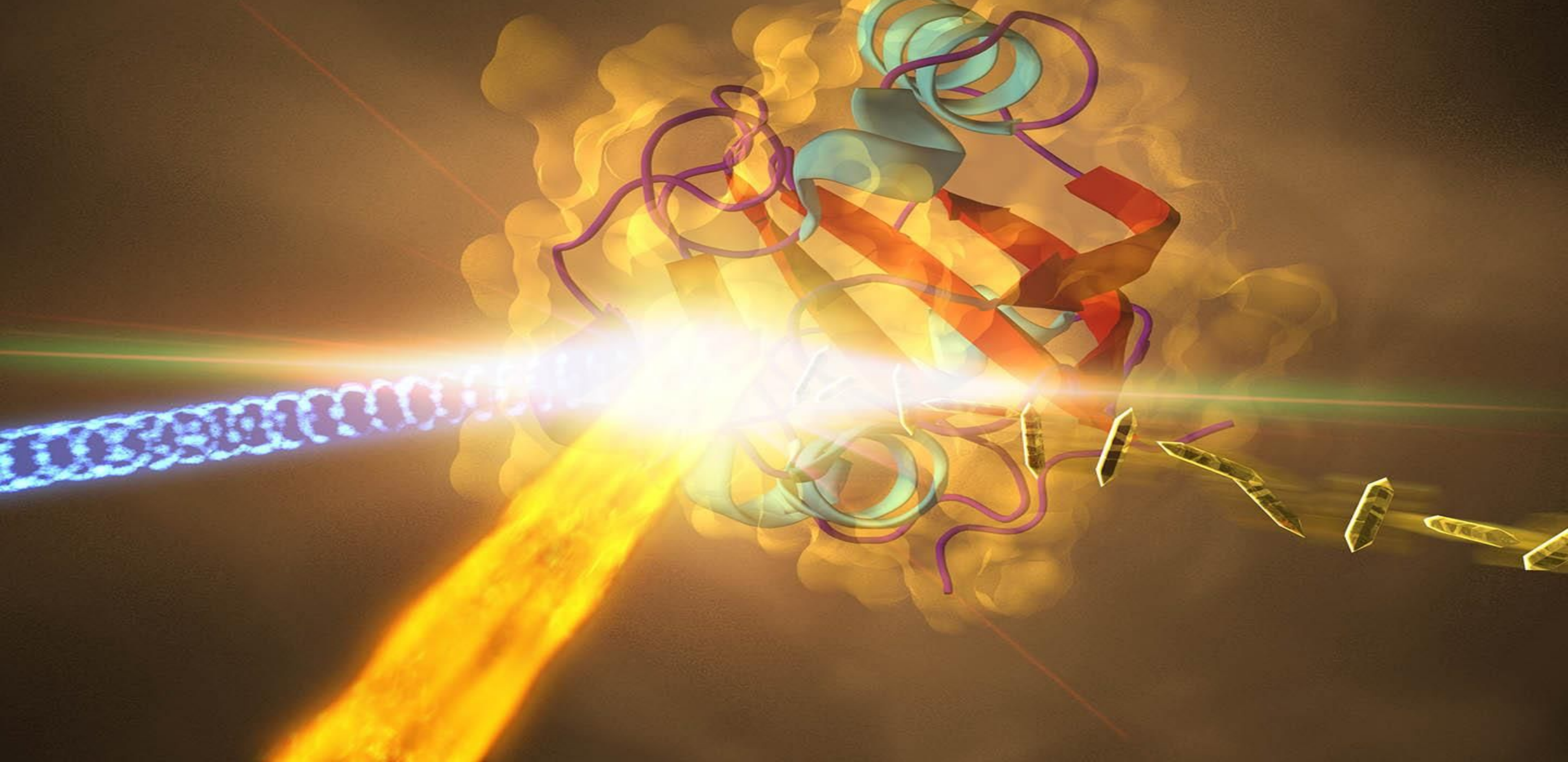
Linac Coherent Light Source (LCLS) instruments and science case

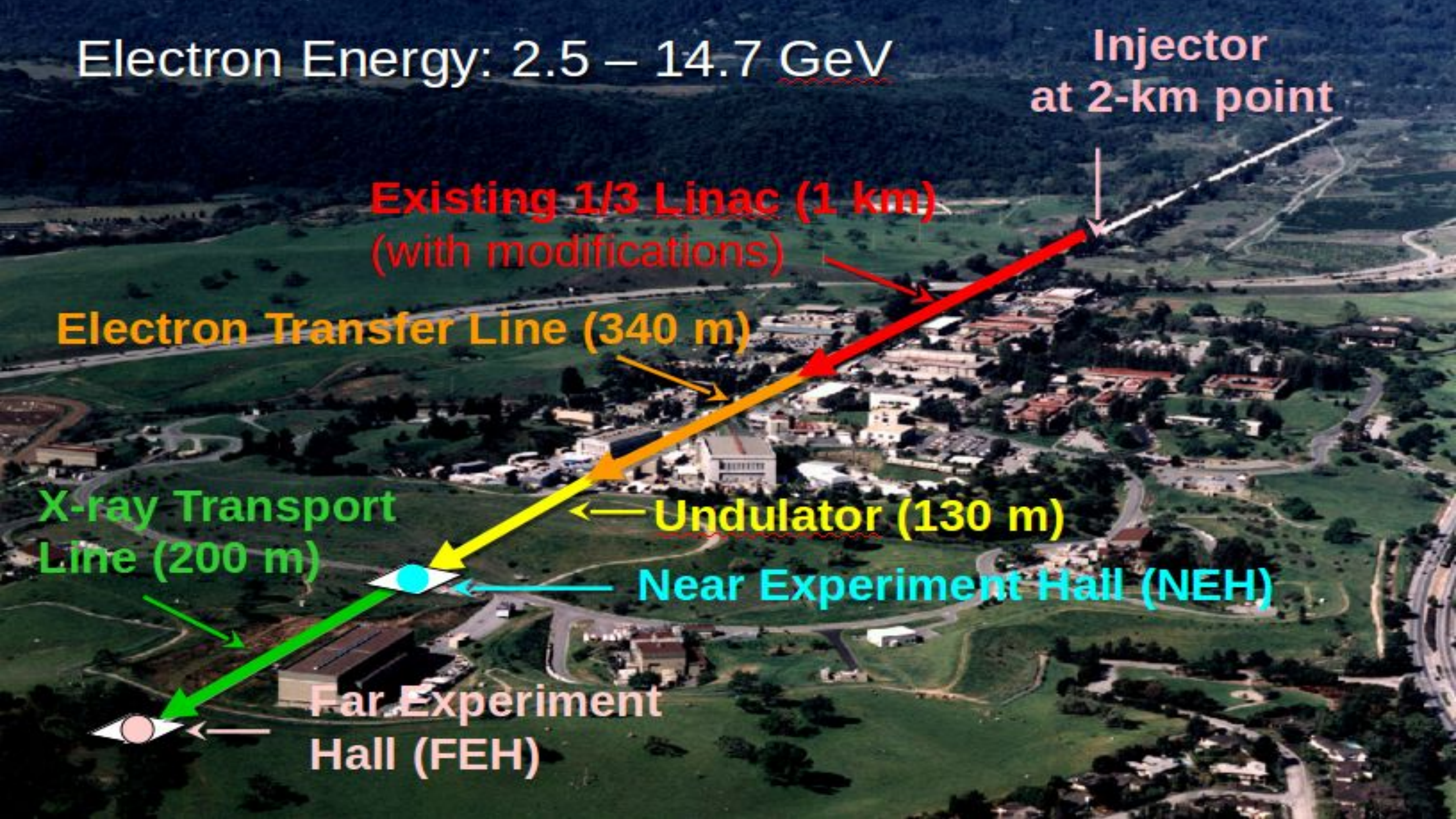
Data systems architecture

On-the-fly data reduction

Quasi real-time data analysis

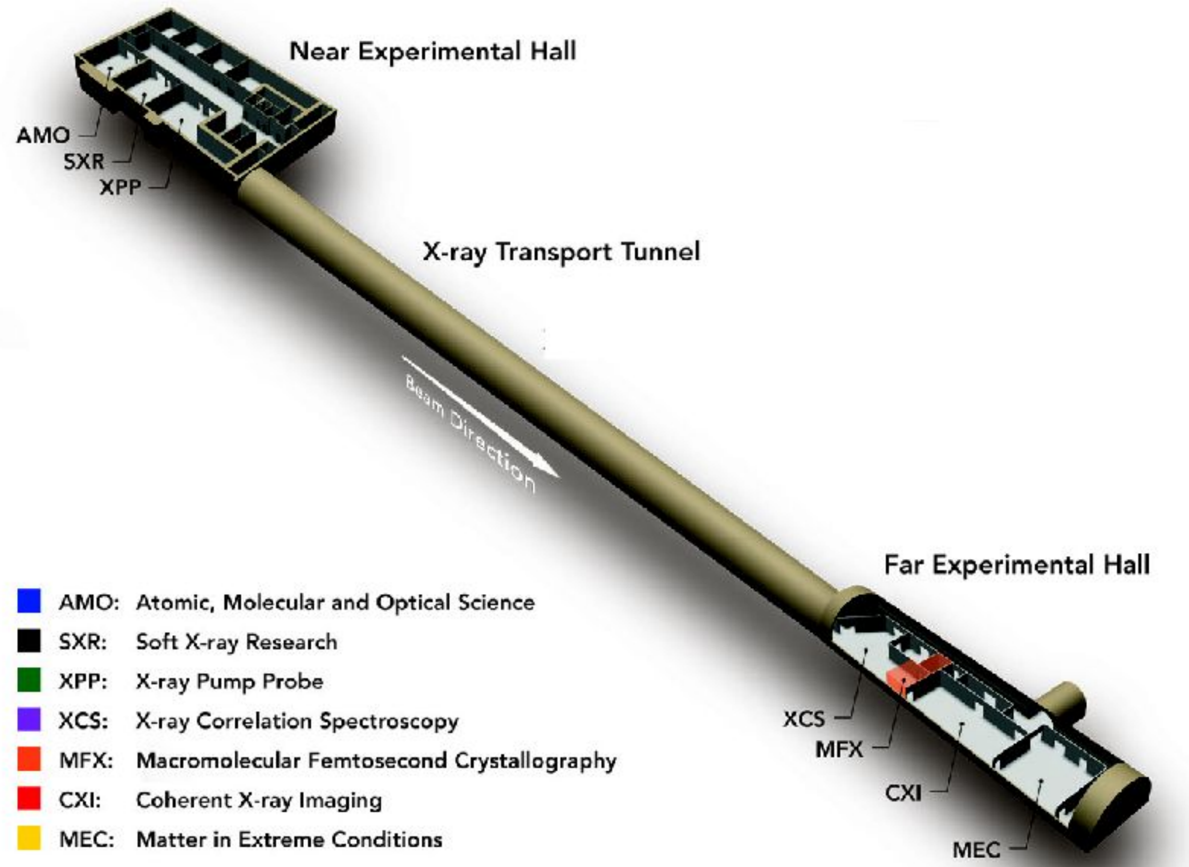
LCLS Science Case





Electron Energy: 2.5 – 14.7 GeV
Injector
at 2-km point
Existing 1/3 Linac (1 km)
(with modifications)
Electron Transfer Line (340 m)
X-ray Transport
Line (200 m)
Undulator (130 m)
Near Experiment Hall (NEH)
Far Experiment
Hall (FEH)

LCLS Instruments



LCLS has already had a significant impact on many areas of science, including:

- ➔ Resolving the structures of macromolecular protein complexes that were previously inaccessible
- ➔ Capturing bond formation in the elusive transition-state of a chemical reaction
- ➔ Revealing the behavior of atoms and molecules in the presence of strong fields
- ➔ Probing extreme states of matter

Data Analytics for high repetition rate Free Electron Lasers

SLAC

FEL data challenge:

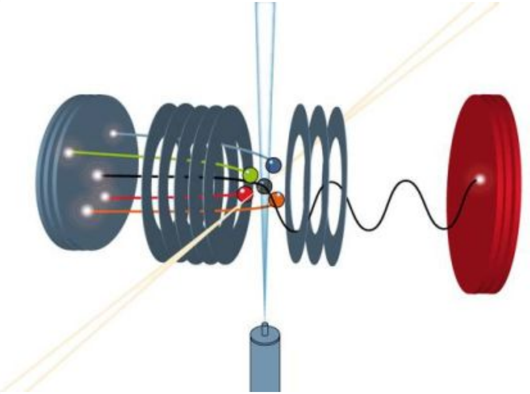
- **Ultrafast X-ray pulses** from LCLS are used like flashes from a high-speed strobe light, producing stop-action movies of atoms and molecules
- Both **data processing** and **scientific interpretation** demand intensive computational analysis

LCLS-II will increase **data throughput by three orders of magnitude** by 2025, creating an exceptional scientific computing challenge

LCLS-II represents SLAC's largest data challenge

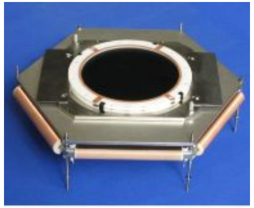
LCLS-II Data Analysis Pipelines: Dynamic Reaction Microscope Example

Experiment



- A gas jet is used to introduce individual molecules into an interaction point where optical and LCLS pulses are overlapped
- Twin position sensitive electron/ion time of flight detectors are used to detect coincident events, elucidating correlated excited state dynamics

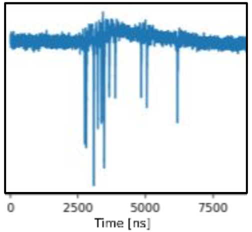
Charged particle detector



- **20 digitizers (1.25 GHz)**
- **50,000 samples per pulse**
- **100 kHz pulse rate**

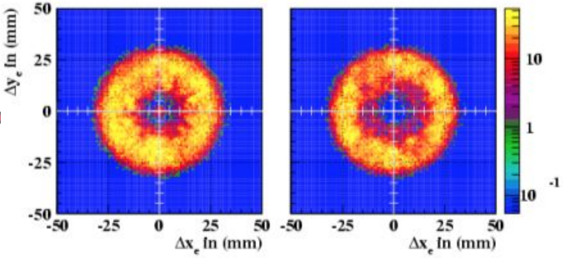
200 GB/s →

X-ray diffraction image



5 GB/s →

Intensity map from multiple pulses



Data Reduction
- Sparse data
- Remove"zeros"
- >50x reduction

< 1 TFlops

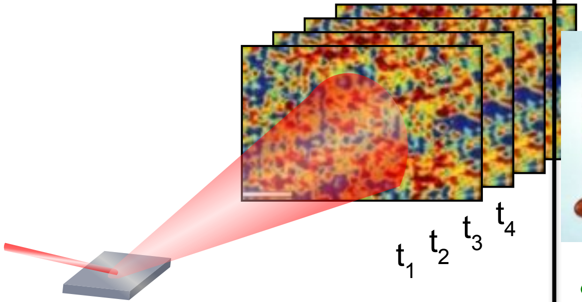
Data Analysis
- Reconstruct 3-dimensional momentum vector as a function of pump-probe time delay

< 1 TFlops

LCLS-II Data Analysis Pipelines: Photon Correlation Spectroscopy Example

SLAC

Experiment Description



- A continuous time series of coherent scattering ("speckle") patterns are collected from disordered system at different momentum transfer
- A time correlation analysis is performed to reveal the dynamic behavior of the system

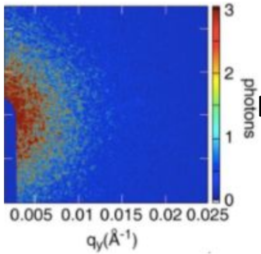
Low-noise area detector

- 5 kHz in 2024 (0.5 MP)
- 40 kHz in 2026 (0.5 MP)

5 GB/s

40 GB/s

Coherent scattering image



Data Reduction

- Sparse data
- Remove"zeros"
- >2x reduction

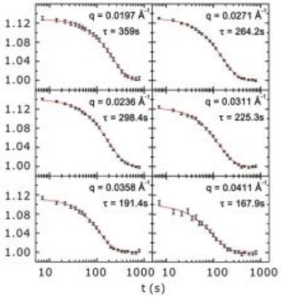
4 TFlops

34 TFlops

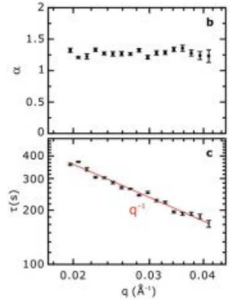
2.5 GB/s

20 GB/s

Intensity map from multiple pulses



Interpretation of system structure / dynamics



Data Analysis

- Droplet algorithm
- Momentum resolved time correlation analysis

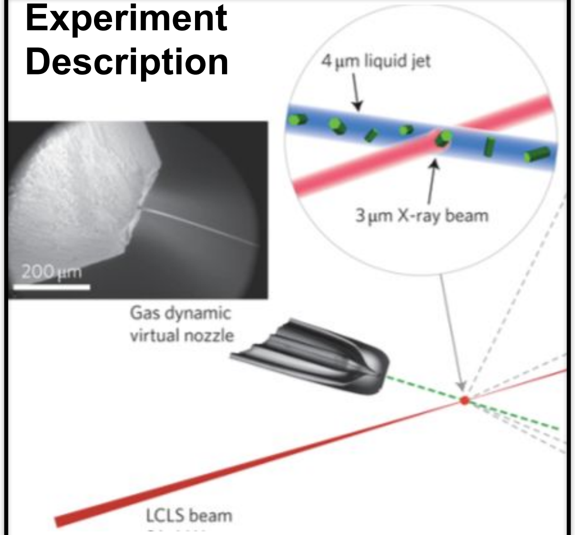
34 TFlops

270 TFlops

LCLS-II Data Analysis Pipelines: Nanocrystallography Example

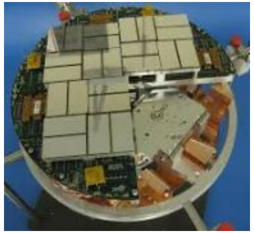
SLAC

Experiment Description



- Individual nanocrystals are injected into the focused LCLS pulses
- Diffraction patterns are collected on a pulse-by-pulse basis
- Crystal concentration dictates “hit” rate

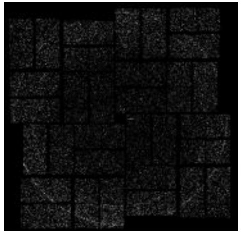
Multi-megapixel detector



- 8 kHz in 2024 (4 MP)
- 40 kHz in 2027 (16 MP)

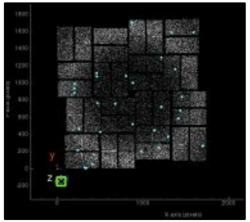
60 GB/s → 1 TB/s

X-ray diffraction image

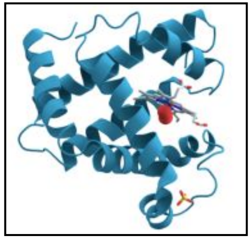


6 GB/s → 100 GB/s

Intensity map from multiple pulses



Interpretation of system structure / dynamics



Data Reduction

- Remove”no hits”
- >10x reduction

3 TFlops
16 TFlops

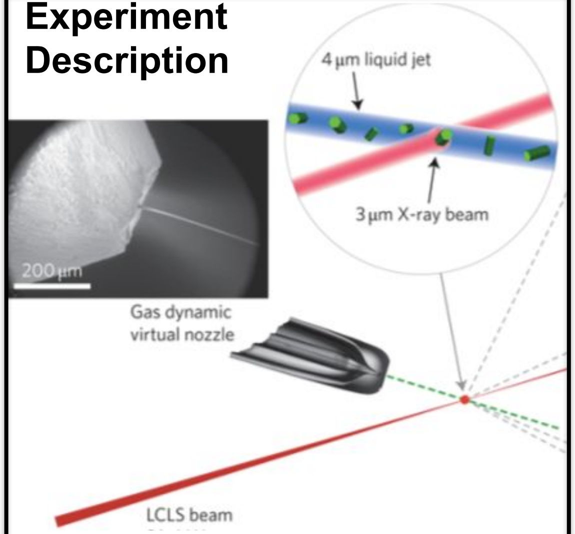
Data Analysis

- Bragg peak finding
- Index / orient patterns
- Average
- 3D intensity map
- Reconstruction

4 PFlops
20 PFlops

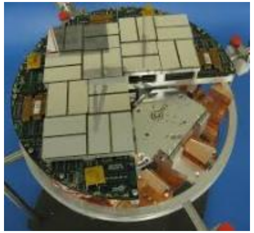
LCLS-II Data Analysis Pipelines: Single Particle Imaging Example

Experiment Description



- Individual particles are injected into the focused LCLS pulses
- Scattering patterns are collected on a pulse-by-pulse basis
- Particle concentration dictates "hit" rate

Multi-megapixel detector



- 8 kHz in 2024 (4 MP)
- 40 kHz in 2027 (16 MP)

60 GB/s
1 TB/s

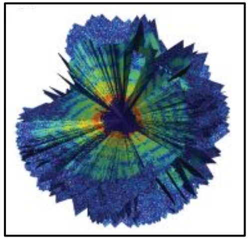
Coherent scattering image

Data Reduction
- Remove"no hits"
- >10x reduction

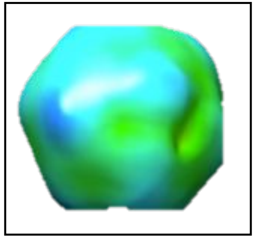
3 TFlops
16 TFlops

6 GB/s
100 GB/s

Intensity map from multiple pulses



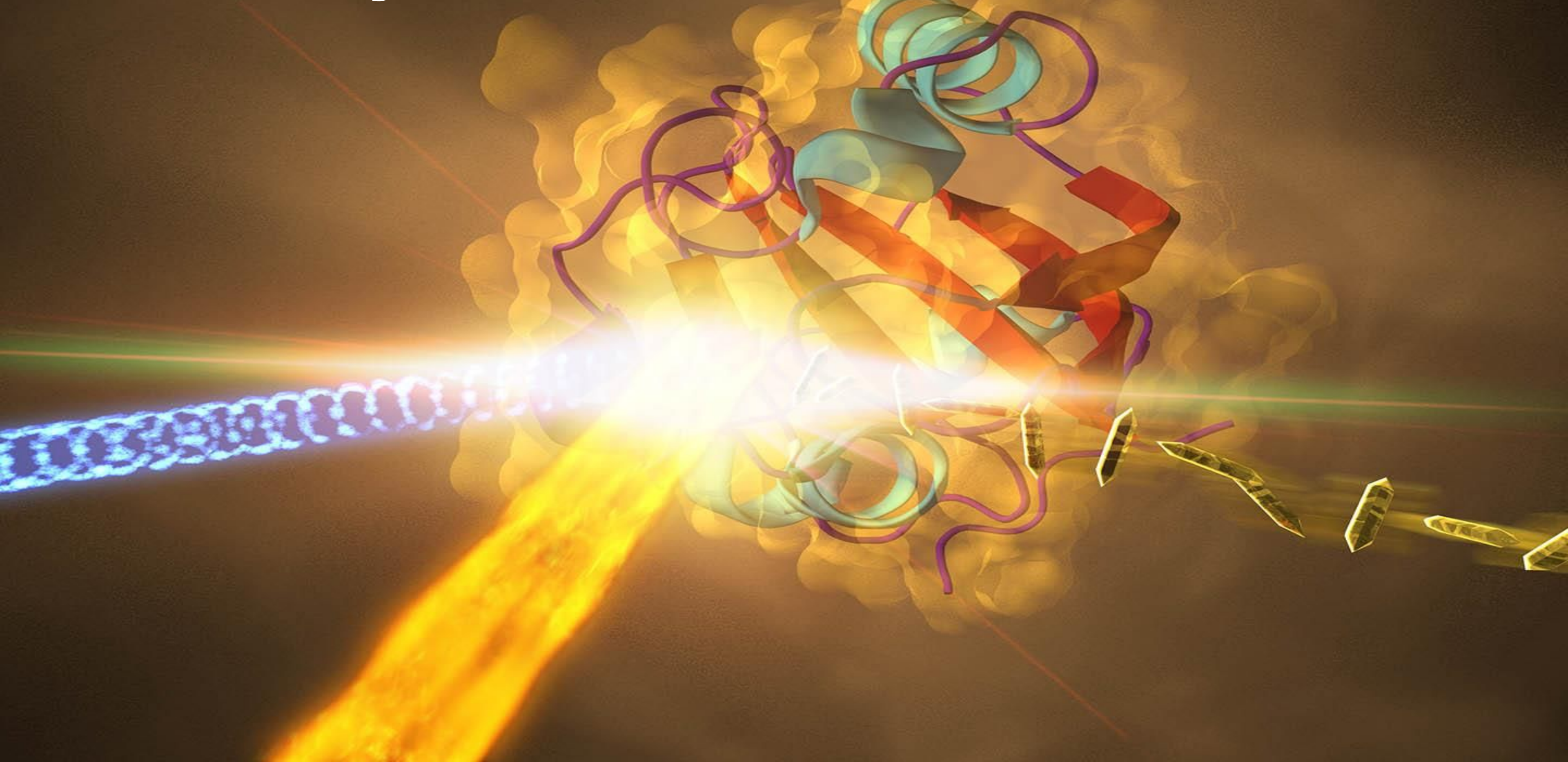
Interpretation of system structure / dynamics



Data Analysis
- Orient patterns
- Average
- 3D intensity map
- Reconstruction

270 PFlops
1340 PFlops

Data Systems Architecture



Computing Requirements for Data Analysis: a *Day in the Life of a User* Perspective

SLAC

- During **data taking**:
  - Must be able to get real time (~1 s) **feedback** about the **quality of data taking**, e.g.
    - Are we getting all the required detector contributions for each event?
    - Is the hit rate for the pulse-sample interaction high enough?
  - Must be able to get **feedback** about the **quality of the acquired data** with a latency lower (~1 min) than the typical lifetime of a measurement (~10 min) in order to optimize the experimental setup for the next measurement, e.g.
    - Are we collecting enough statistics? Is the S/N ratio as expected?
    - Is the resolution of the reconstructed electron density what we expected?
- During **off shifts**: must be able to run **multiple passes** (> 10) of the full analysis on the data acquired during the previous shift to optimize analysis parameters and, possibly, code in preparation for the next shift
- During **4 months** after the experiment: must be able analyze the raw and intermediate data on **fast access storage** in preparation for publication
- **After 4 months**: if needed, must be able to **restore** the archived data to test new ideas, new code or new parameters

The Challenging Characteristics of LCLS Computing

SLAC

1. **Fast feedback** is essential (seconds / minute timescale) to reduce the time to complete the experiment, improve data quality, and increase the success rate
2. **24/7 availability**
3. **Short burst** jobs, needing very short startup time
4. **Storage** represents significant fraction of the overall system
5. **Throughput** between storage and processing is critical
6. Speed and flexibility of the **development cycle** is critical - *wide variety of experiments, with rapid turnaround, and the need to modify data analysis during experiments*

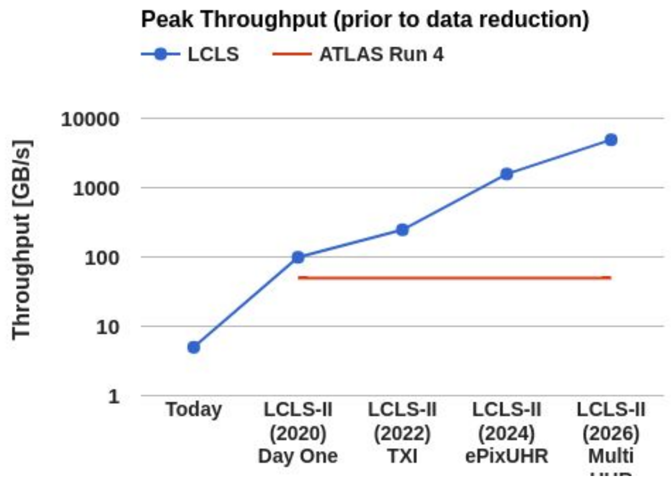
Example data rate for LCLS-II (early science)

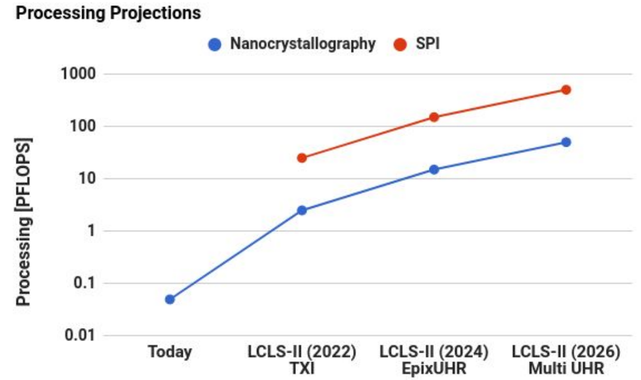
- 1 x 4 Mpixel detector @ 5 kHz = **40 GB/s**
- 100K points fast digitizers @ 100kHz = **20 GB/s**
- Distributed diagnostics 1-10 GB/s range

Example LCLS-II and LCLS-II-HE (mature facility)

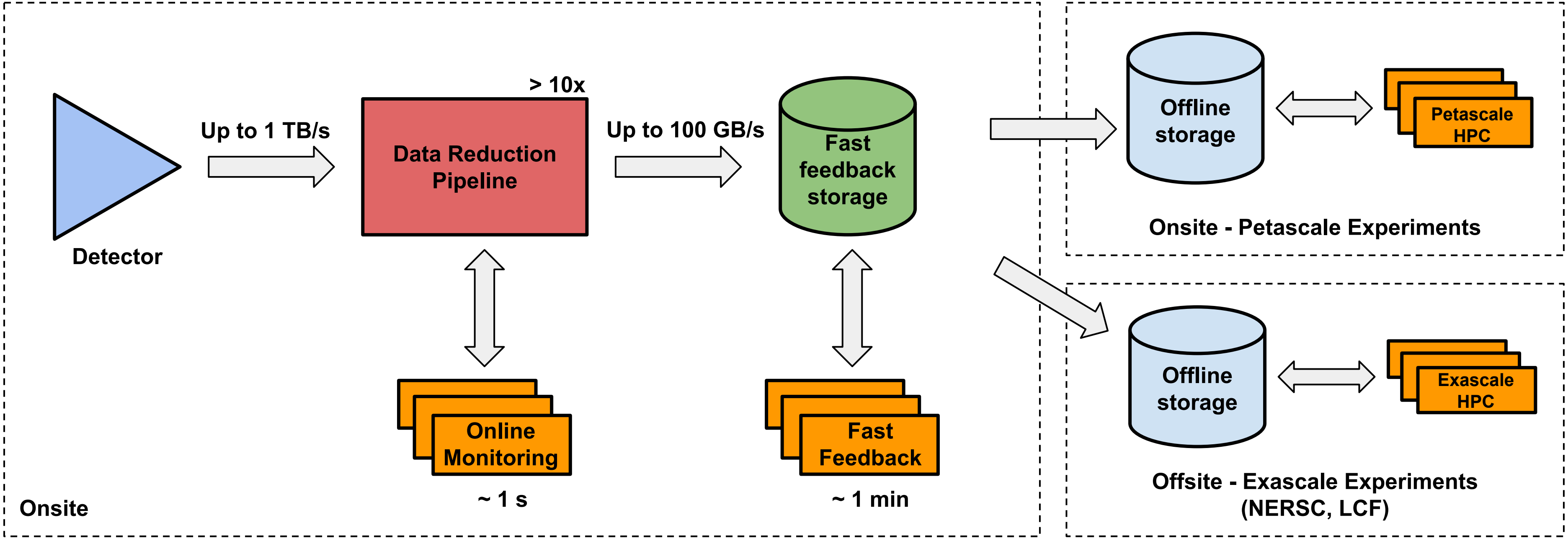
- 2 planes x 4 Mpixel ePixUHR @ 100 kHz = **1.6 TB/s**

Sophisticated algorithms under development within ExaFEL (e.g., M-TIP for single particle imaging) will require exascale machines





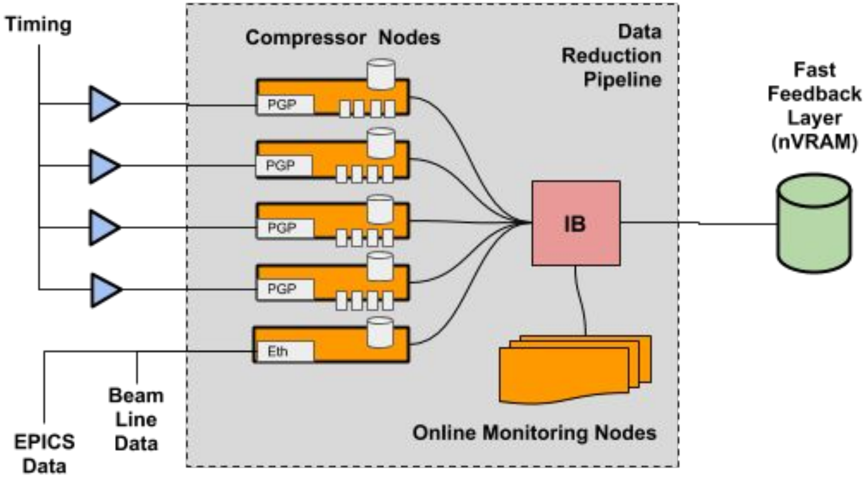
LCLS-II Data Flow



Data reduction mitigates storage, networking, and processing requirements

Data Reduction Pipeline

- Besides cost, there are **significant risks** by not adopting on-the-fly data reduction
  - Inability to move the data to HEC, system complexity (robustness, intermittent failures)
- Developing toolbox of techniques (**compression, feature extraction, vetoing**) to run on a **Data Reduction Pipeline**
- Significant **R&D effort**, both engineering (throughput, heterogeneous architectures) and scientific (real time analysis)

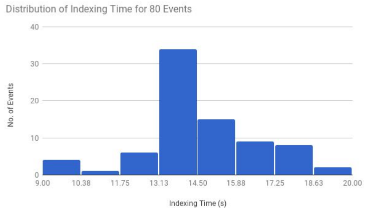


Without on-the-fly data reduction we would face unsustainable hardware costs by 2026

Process for determining future projections

Includes:

1. **Detector rates** for each instrument
2. **Distribution of experiments** across instruments (as function of time, ie as more instruments are commissioned)
3. Typical **uptimes** (by instruments)
4. **Data reduction** capabilities based on the experimental techniques
5. Algorithm **processing times** for each experimental technique



Time per event
x
Data rates

↓

Number of cores required

↓

FLOPS (~invariant)

Undulator	Instrument	Endstation	Technique	Detector	Detector Size	Detector Rate (Hz)	Data Rate (aggregate) (GB/s)	Ultilization Factor (0-1)	Data Reduction Type (1st Cut)	DR Factor (1st cut)	Data Reduction Type (Optimistic))	DR Factor (Optimistic)	FY20 Q1	FY20 Q2	FY20 Q3	FY20 Q4	FY21 Q1	FY21 Q2	FY21 Q3	F
SXU	NEH 1.1	DREAM	COLTRIMS	Digitizer	800000	100000	160.0	0.75	Zero suppression	0.020	Peak Finding	0.0020		1.00	1.00	0.50	0.25	0.25	0.25	[illegible]
SXU	NEH 1.1	DREAM	Time of Flight	Digitizer	1000000	100000	200.0	0.75	Zero suppression	0.020	Peak Finding	0.0020				0.13	0.13	0.13	0.06	[illegible]
SXU	NEH 1.1	LAMP	Time of Flight	Digitizer	1000000	100000	200.0	0.75	Zero suppression	0.020	Peak Finding	0.0020				0.13	0.13	0.13	0.06	[illegible]
SXU	NEH 1.1	LAMP	Imaging	SXR Imag. + Digi.	4000000	10000	82.0	0.45	Veto	0.100	N.A.	0.1000							0.13	[illegible]
SXU	NEH 2.2	LJE	XAS / XES	TES	1000	100000	20.0	0.60	Zero suppression	0.100	Binning	0.0000								[illegible]
SXU	NEH 2.2	LJE	XAS / XES	TES	10000	100000	200.0	0.60	Zero suppression	0.100	Binning	0.0000								
SXU	NEH 2.2	LJE	XAS / XES	RIXS-ccd	4096	1000	0.0	0.60	N.A.	1.000	Accumulating	0.0010				0.25	0.50	0.25	0.25	[illegible]
SXU	NEH 2.2	RIXS	IXS / RIXS	RIXS-ccd	4096	1000	0.0	0.60	N.A.	1.000	Accumulating	0.0010						0.13	0.13	[illegible]
SXU	NEH 2.2	RIXS	XRD / RXRD	SXR Imaging	1000000	10000	20.0	0.60	ROI	0.100	Accumulating	0.0001						0.06	0.06	[illegible]
SXU	NEH 2.2	RIXS	XPCS	SXR Imaging	1000000	10000	20.0	0.60	Compression	0.500		0.1000						0.06	0.06	[illegible]
SXU	NEH 1.2	---	X-ray/X-ray	SXR Imaging	1000000	10000	20.0	0.30	ROI	0.100	Binning	0.0001								
SXU	NEH 1.2	---	Imaging	epix100-HR + Digi.	4000000	5000	42.0	0.45	Veto	0.100	N.A.	0.1000								
SXU	NEH 1.2	---	XAS / XES	RIXS-ccd	4096	1000	0.0	0.60	N.A.	1.000	Accumulating	0.0010								

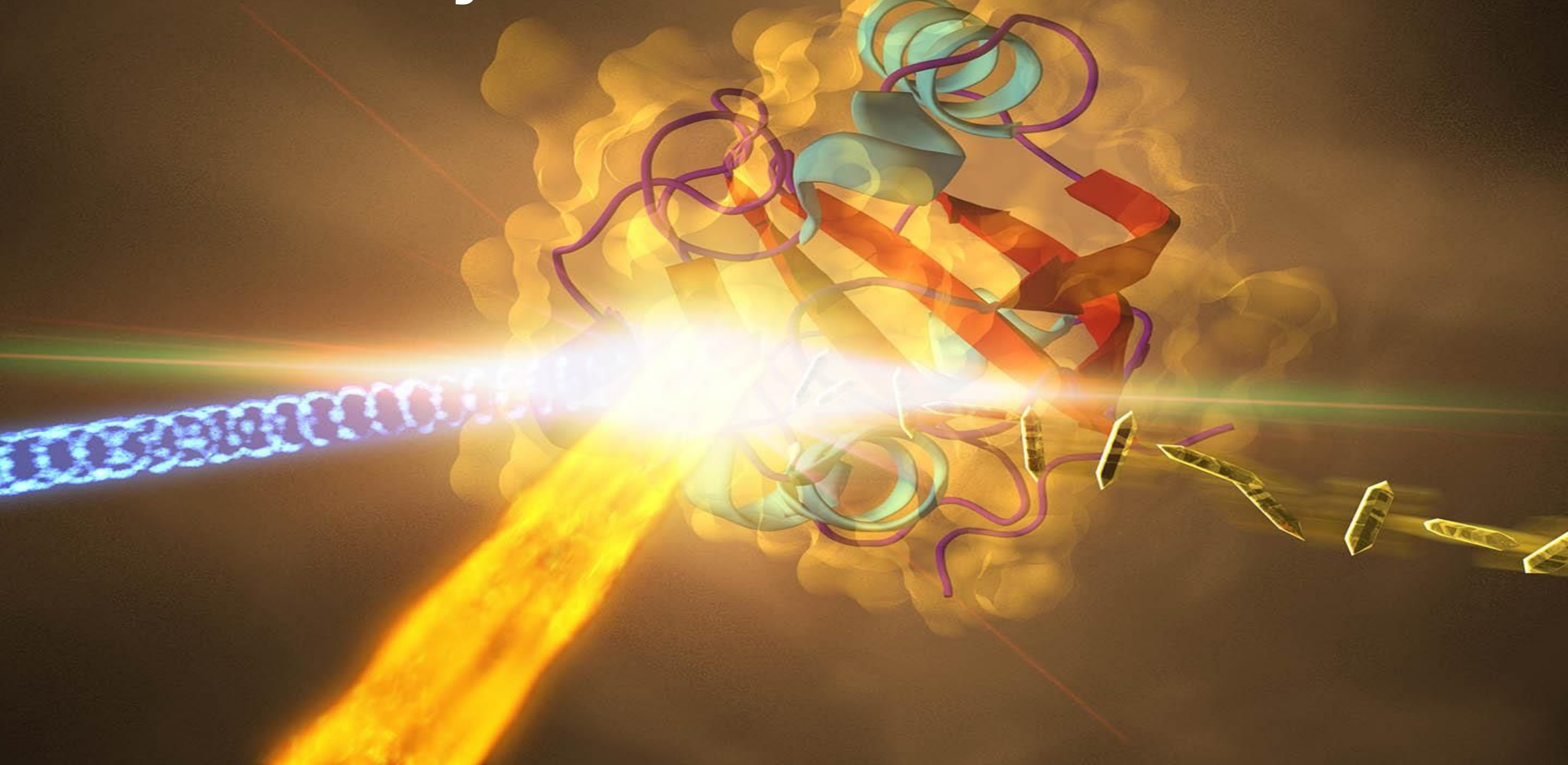
Make full use of national capabilities

LCLS-II will require access to **High End Computing Facilities (NERSC and LCF)** for highest demand experiments (exascale)

Photon Science Speedway

Stream science data files on-the-fly from the LCLS beamlines to the NERSC supercomputers via ESnet

On-the-fly Data Reduction



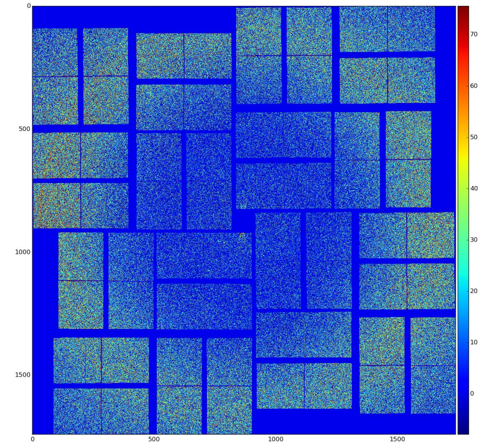
Data Reduction Techniques

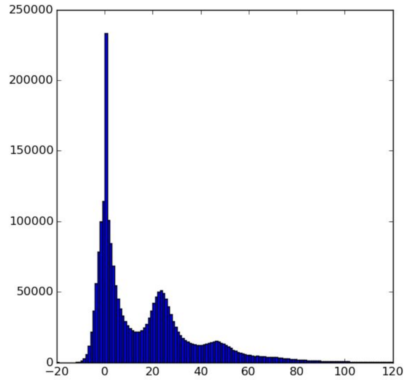
Develop a toolbox of different techniques:

- Compression, lossless and lossy
- Feature extraction
  - sparsification (peak-finders, ROI)
  - calculation (ROI integration, angular integration, binning)
  - data transformation (into space where data are sparse, e.g. wavelet JPEG-style compression)
- Veto (software trigger)
- Multi event reduction
  - MPEG style correlation over events

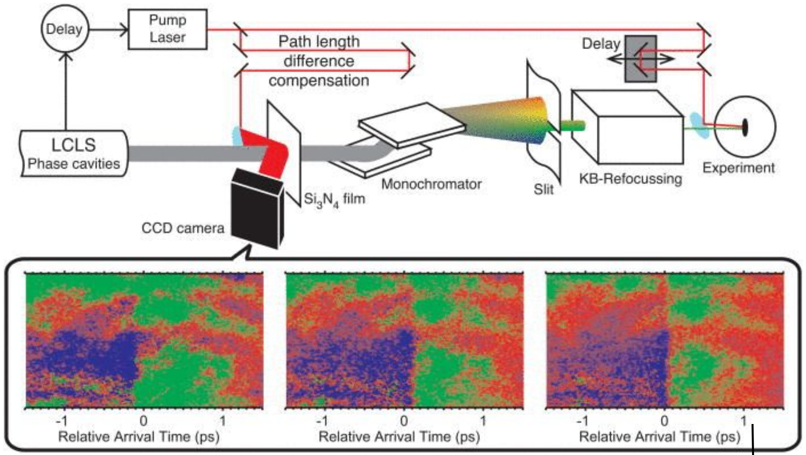
Three examples of these techniques described in this section:

- Time tool processing with FPGA
- Lossy compression for nano crystallography
- Beam center determination for feature extraction





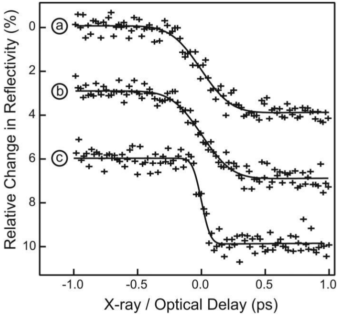
Data reduction examples: FPGA Processing of Time Tool

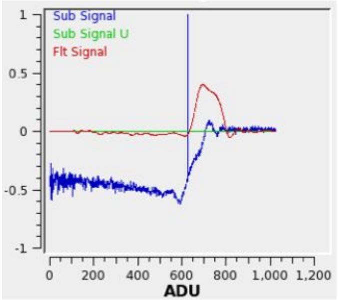


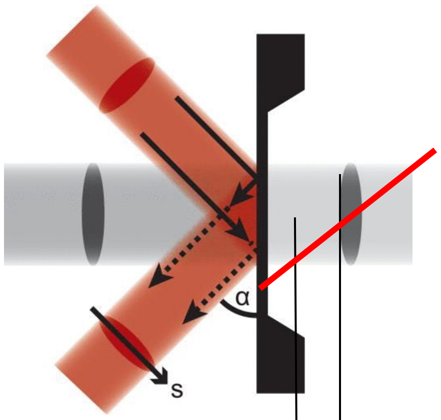
M. Beye, W.F. Schlotter, et. al. Appl. Phys. Lett. 100, 121108

2d projection

FPGA ideal for analysis which are (relatively) stable over time







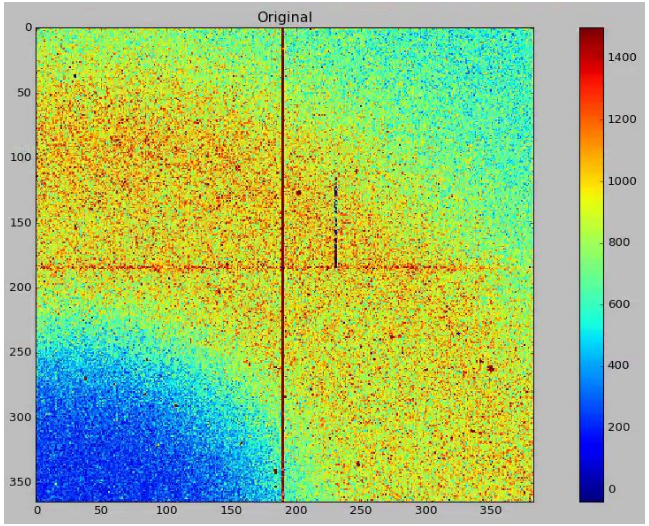
femtosecond delay

this side of laser pulse hits SiN first

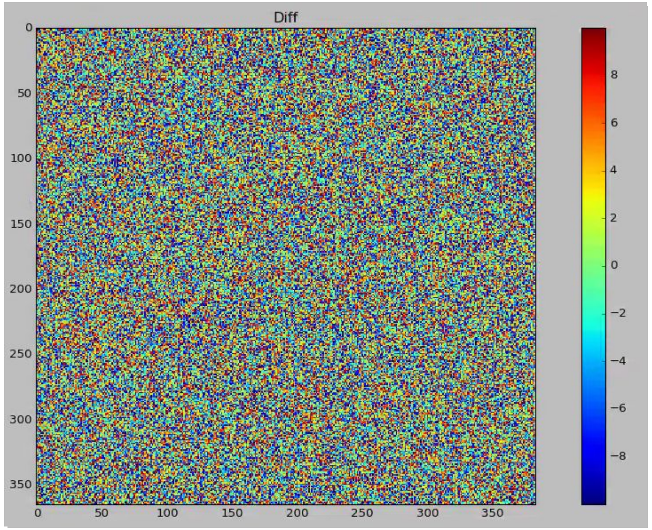
Data reduction examples: Crystallography SZ Compression

- Create simulated SZ compression by adding uniform random noise
- How does noise affect reconstruction?

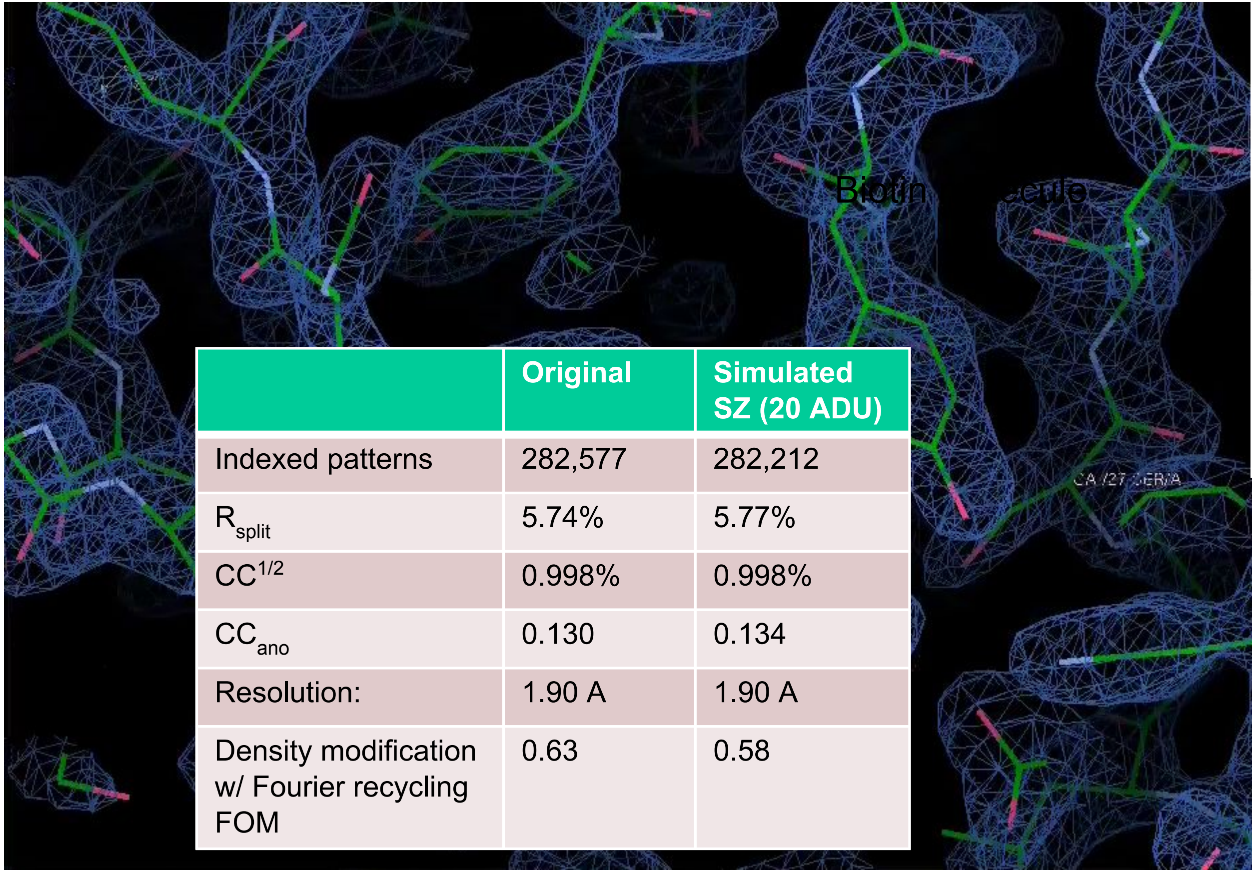
Original



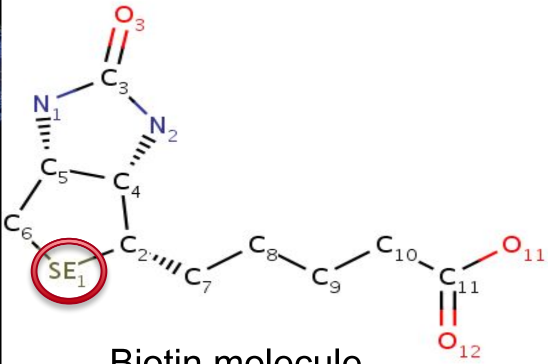
Uniform random noise: +/-10 ADU



Data reduction examples:
SZ Compression Effect on Nanocrystallography



	Original	Simulated SZ (20 ADU)
Indexed patterns	282,577	282,212
R_{split}	5.74%	5.77%
$CC^{1/2}$	0.998%	0.998%
CC_{ano}	0.130	0.134
Resolution:	1.90 A	1.90 A
Density modification w/ Fourier recycling FOM	0.63	0.58

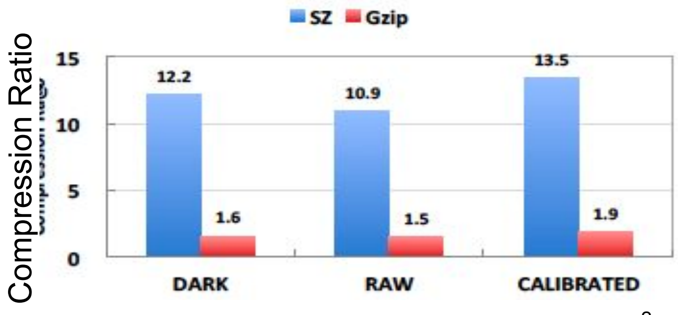


Biotin molecule

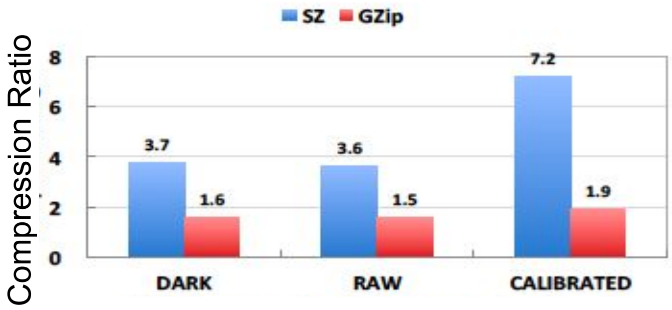
Data reduction examples:
SZ Compression results (work in progress)

Lossy compression with guaranteed error bounds (relative or absolute)
Value-range-based relative error bound = absolute error bound / value range

Value-range-based relative error bound = 10^{-1}



Value-range-based relative error bound = 10^{-2}

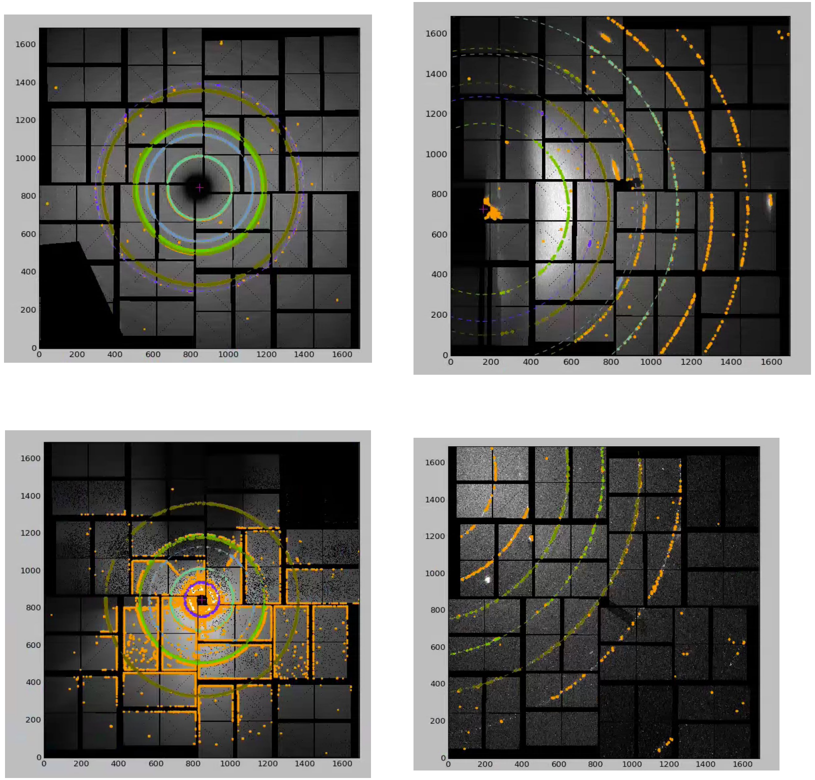


Value-range-based relative error bound = 10^{-3}

Added Noise Level (absolute)	Compression Factor
20ADU	2.5x
100ADU	4x
500ADU	10x

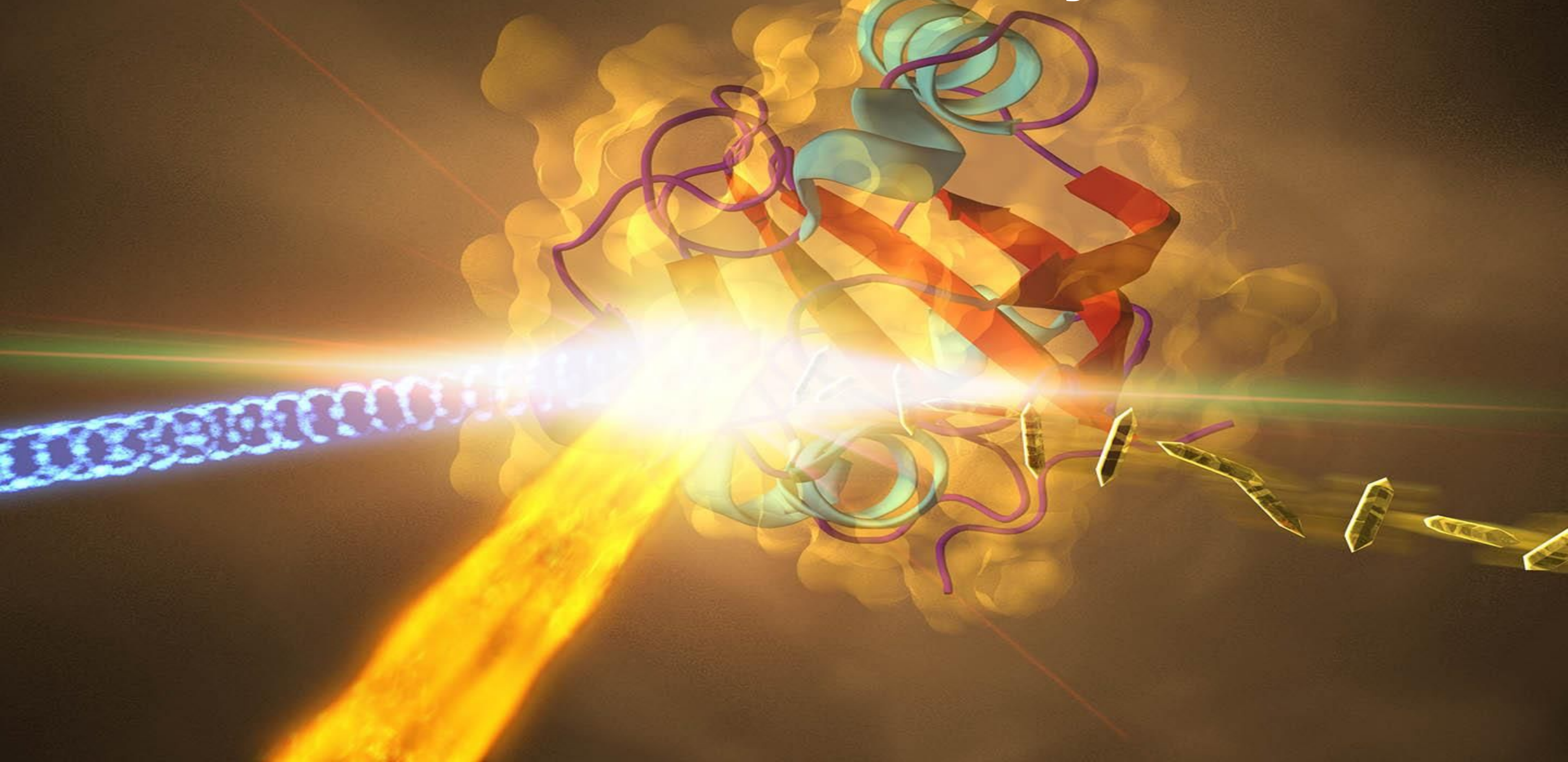
Data reduction examples: Beam Center Finding

- Angular Integration is a common data-reduction technique (e.g. WAXS/SAXS) but need to know beam-center location
- Developed parameter-free beam-center finding from rings using Hough transform, plus “canny” and “ransac” algorithms
- Works in many tricky cases (see pics)
- Two known conditions where this approach fails:
  - very faint/wide rings
  - data with an obvious beamstop that is not masked out



Picture credit: Silke Nelson (SLAC)

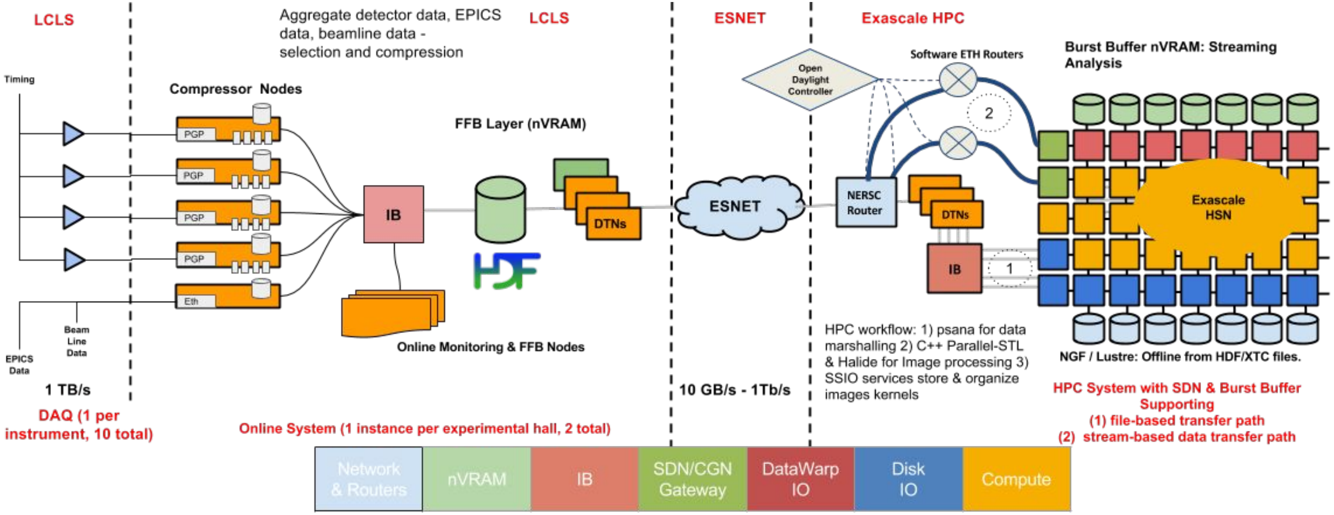
Quasi Real-time Data Analysis



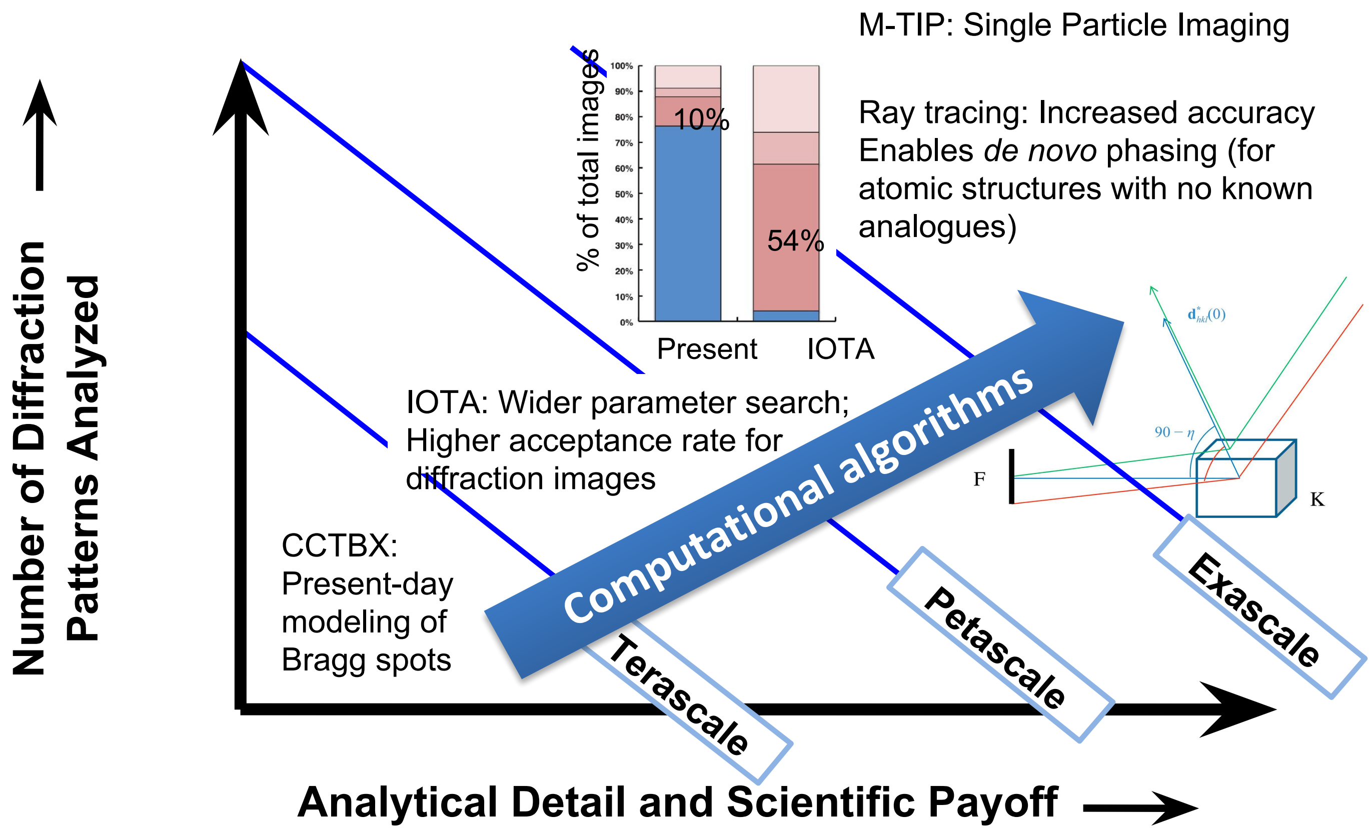
ExaFEL:
Data Analytics at the Exascale for Free Electron Lasers

Application Project within Exascale Computing Project (ECP)

High data throughput experiments	LCLS data analysis framework	Infrastructure
Algorithmic improvements and ray tracing - Example test-cases of Serial Femtosecond Crystallography, and Single Particle Imaging	Porting LCLS code to supercomputer architecture, allow scaling from hundreds of cores (now) to hundred of thousands of cores	Data flow from SLAC to NERSC over ESnet



From Terascale to Exascale

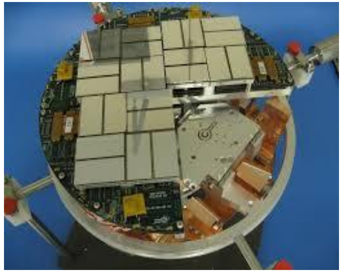


Exascale vastly expands the experimental repertoire and computational toolkit

Picture credit: Kroon-Batenburg et al (2015) Acta Cryst D71:1799

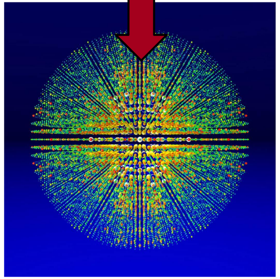
Scaling the nanocrystallography pipeline

- Avoidance of radiation damage and emphasis on physiological conditions requires a transition to fast (fs) X-ray light sources & large (10^6 image) datasets
- Real time data analysis within minutes provides results that feed back into experimental decisions, improving the use of scarce sample and beam time
- *Terabyte diffraction image datasets collected at SLAC / LCLS are transferred to NERSC over ESnet & analyzed on Cori / KNL*

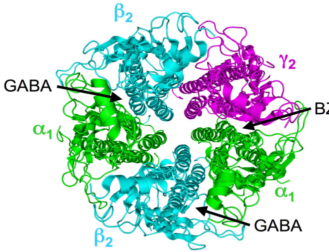


Megapixel detector

X-Ray Diffraction Image
"diffraction-before-destruction"



Intensity map (multiple pulses)

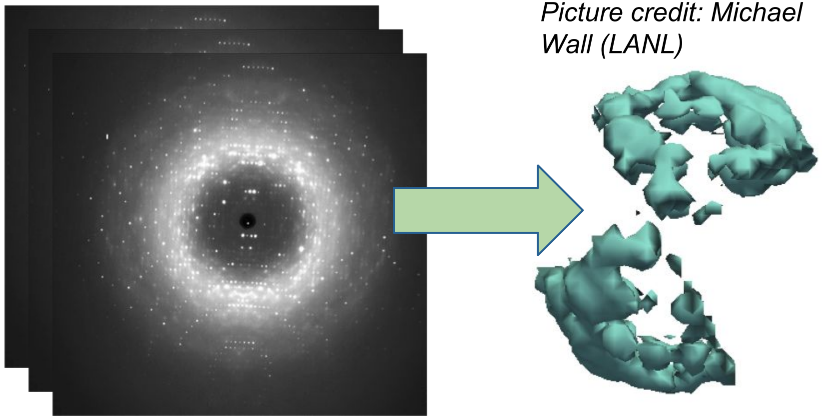


Electron density (3D) of the macromolecule

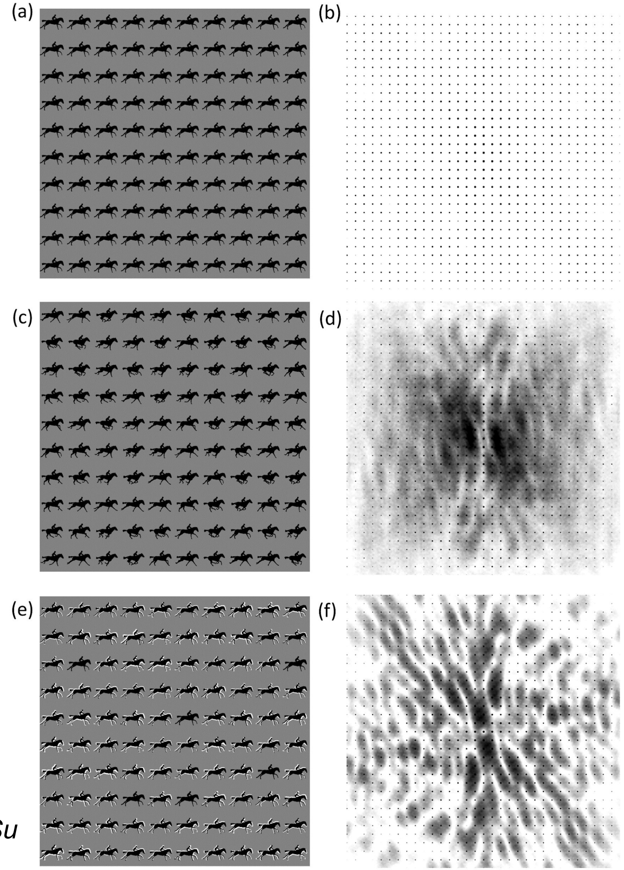
Nick Sauter, LBNL

Analyzing disorder: diffuse scattering

Diffuse scattering is the scattering that arises from any departure of the material structure from that of a perfectly regular lattice



Picture credit: Michael Wall (LANL)

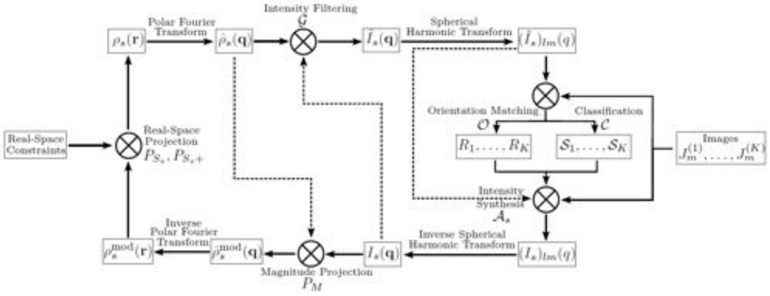


1. *Processing diffraction images to provide input data for scaling and for data integration*
2. *Indexing diffraction images to map pixels to reciprocal space coordinates*
3. *Integration of individual diffraction images to populate thin Ewald sphere slices in reciprocal space*
4. *Merging of the slices into a complete 3D dataset*

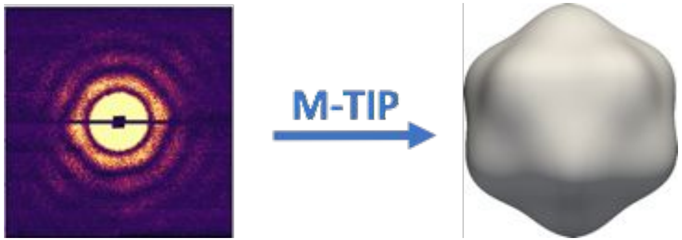
Picture credit: Zhen Su (SLAC)

A new algorithm for single particle imaging: M-TIP

- MTIP (**Multi-Tiered Iterative Phasing**) is an algorithmic framework that **simultaneously determines conformational states, orientations, intensity, and phase** from single particle diffraction images
  - The aim is to reconstruct a 3D structure of a single particle
  - Using diffraction-before-destruction, we measure thousands of diffraction patterns of randomly oriented single particles - we can NOT measure: a) the orientations of the individual particles and b) phases of the diffraction patterns
  - MTIP is an iterative algorithm that deduces these two unknowns given some constraints
- Algorithm published in PNAS [1]
- Modular approach allows modifications to model systematic issues in data and/or incorporate additional constraints



The MTIP framework is a modular, extendable approach to iterative image reconstruction



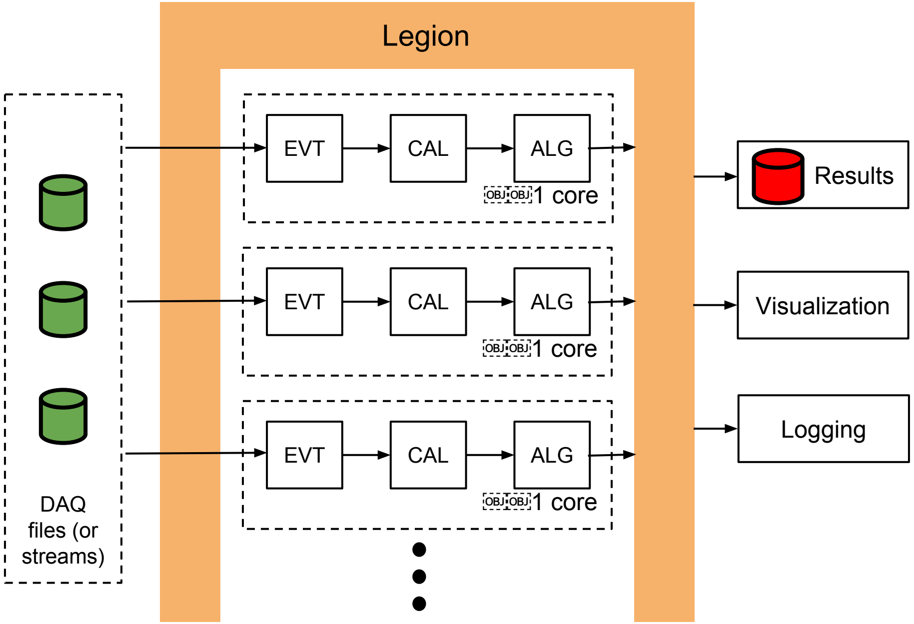
12nm reconstruction of an RDV virus from experimental LCLS data under icosahedral symmetry constraints

[1] Donatelli JJ, Sethian JA, and Zwart PH (2017) Reconstruction from limited single-particle diffraction data via simultaneous determination of state, orientation, intensity and phase. PNAS 114(28): 7222-7227.

From hundreds of cores to hundreds of thousands: scaling the LCLS analysis framework

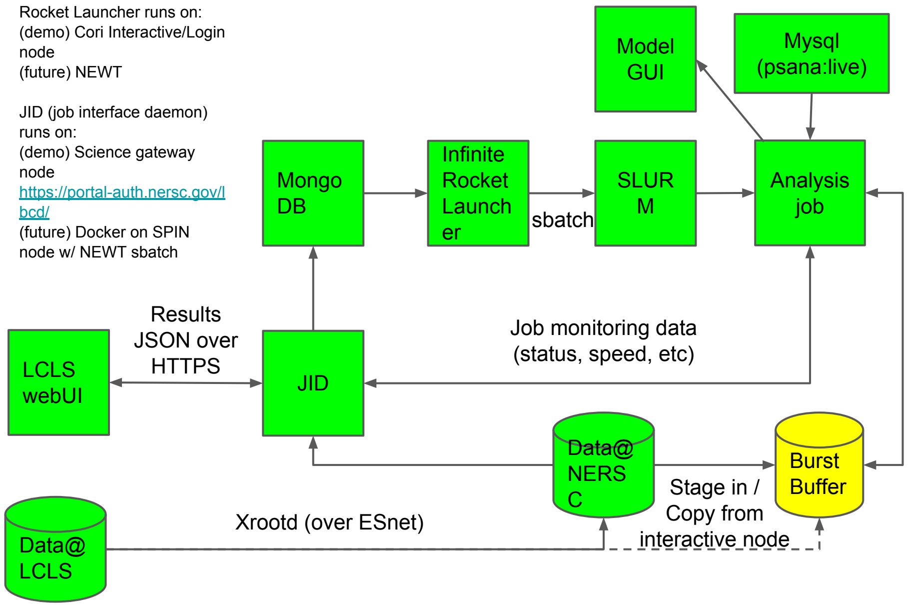
SLAC

- **Complexity of the LCLS analysis framework (psana) is hidden from the users**: parallelization, common algorithms, detector corrections, file formats, visualization
- Reads science data from file or stream, distributes (typically) one event per core, performs detector calibrations and invokes science specific algorithms
- Allows for **real-time analysis** in an identical fashion as offline analysis
- Adopt Legion in psana (currently based on MPI) for LCLS-II to provide
  a. Overlap I/O and compute
  b. Portable performance on new architectures



Towards Automation: end-to-end Workflow

- Workflow manages combination of data streams, hardware system components and applications to derive in quasi real-time the electron density from the diffraction images acquired at the beamline
- Stream the data from the LCLS online cache (NVRAM) to the SLAC data transfer nodes
- Stream the data over an SDN path from the SLAC DTNs to the NERSC DTNs (actual DTNs subset of the supercomputer nodes)
- Write the data to the burst buffers layer (NVRAM)
- Distribute the data from the burst buffers to the local memory on the HPC nodes
- Orchestrate the reduction, merging, phasing and visualization parts of the SFX analysis



ExaFEL FY17 Year End Demo (video)