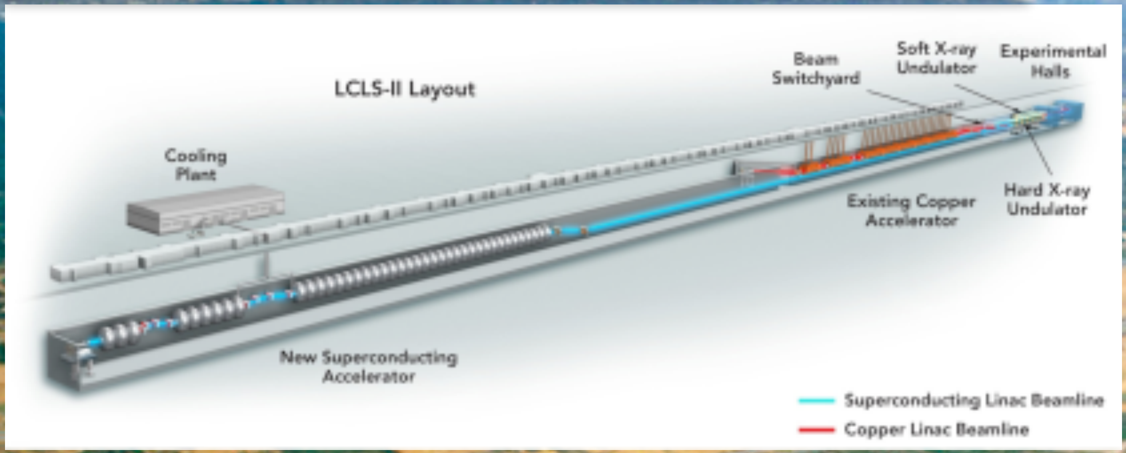
Linac Coherent Light Source (LCLS) Data Transfer Requirements

Les Cottrell,
Ernesto Paiser,
Riccardo Veraldi
SLAC

SC17 talk
Denver Nov 2017

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

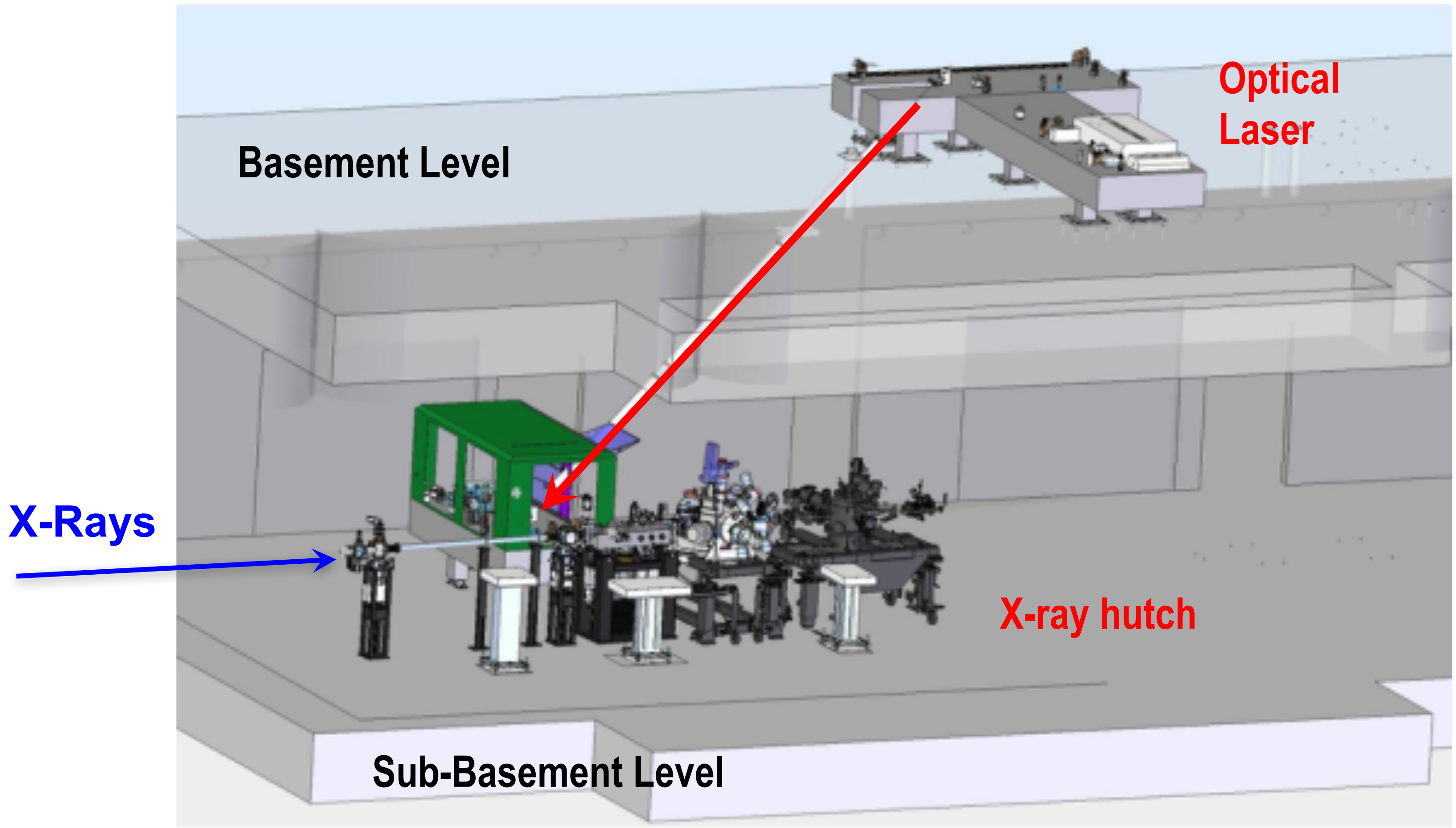
LCLS-II, a major (~ B$) upgrade to LCLS is currently underway. Online in 2020.



Video

Basic instrument layout:
Optical laser pump, and x-ray laser probe

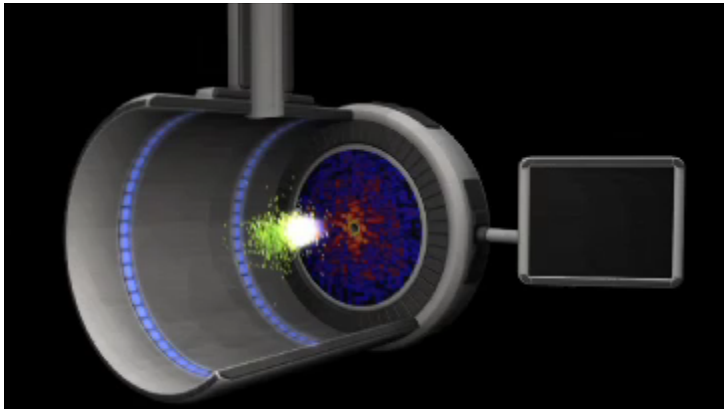
SLAC



Data Analytics for high repetition rate Free Electron Lasers

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

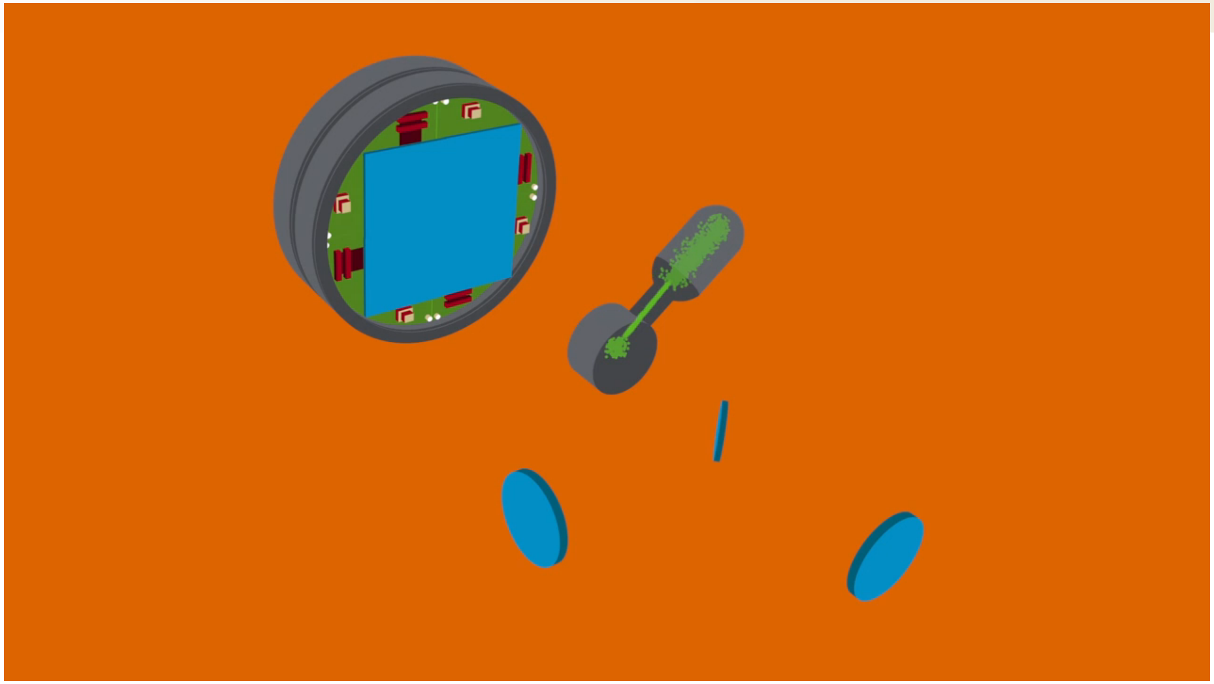


LCLS-II will increase **data throughput by orders of magnitude** by 2026, creating an exceptional scientific computing challenge

LCLS-II represents SLAC's largest data challenge

Example experiment #1
'Molecular Movie' Captures Ultrafast Chemistry in Motion

SLAC



Scientific Achievement

Time-resolved observation of an evolving chemical reaction triggered by light.

Method

LCLS X-ray pulses were delivered at different time intervals, measuring
the structural changes on an X-ray area detector.

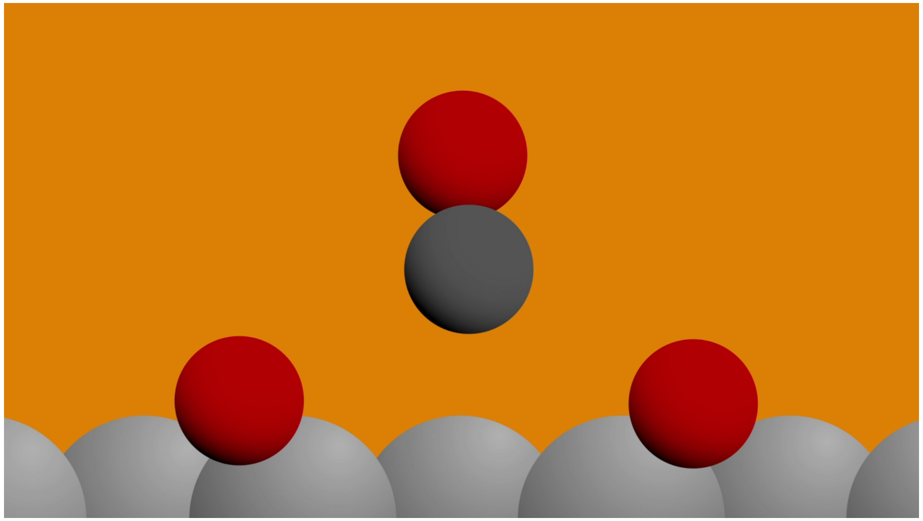
Significance and Impact

Results pave the way for a wide range of X-ray studies examining gas phase chemistry and the structural dynamics associated with the chemical reactions they undergo.

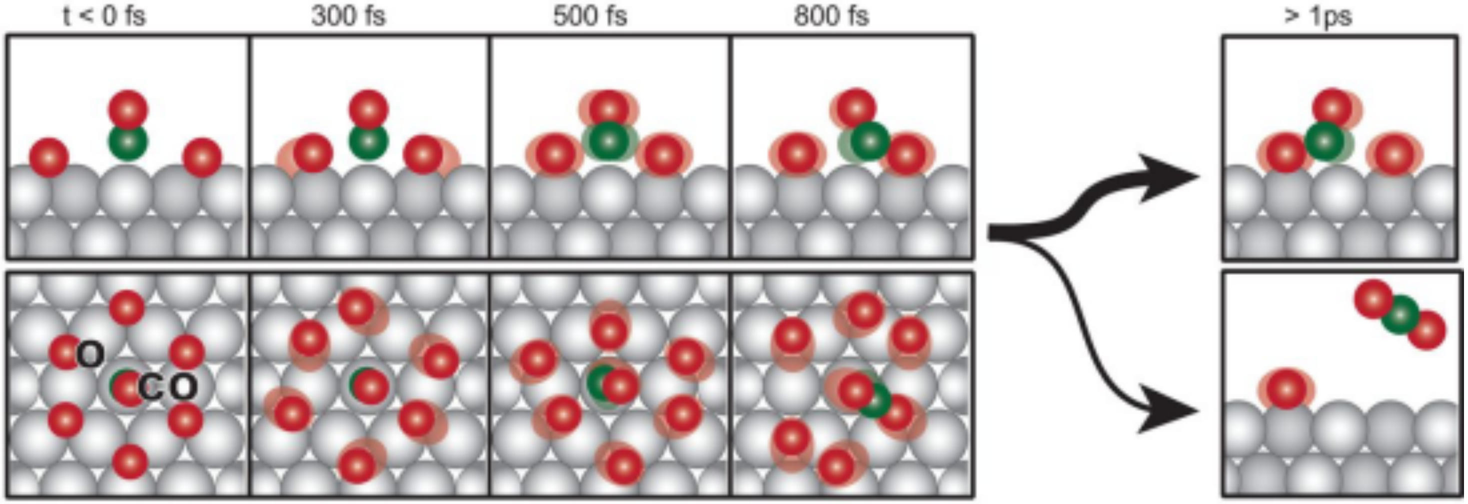
M.P. Minitti, J.M. Budarz, et al., *Phys. Rev. Lett.*, 114, 255501 (2015) (COVER ARTICLE)

Example experiment #2: Catalytic converter – transient dynamics resolved at the atomic scale

- Surface catalysis of CO oxidation to CO_2
- Sub-picosecond transient states, monitored via appearance of new electronic states in the O K-edge x-ray absorption spectrum.



H. Öström et al., **Science** 2015



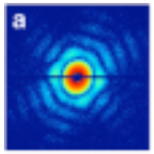
LCLS-II data flow: from data production, to online reduction, real-time analysis, and offline interpretation

SLAC

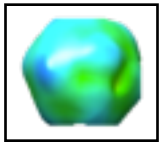
Megapixel detector

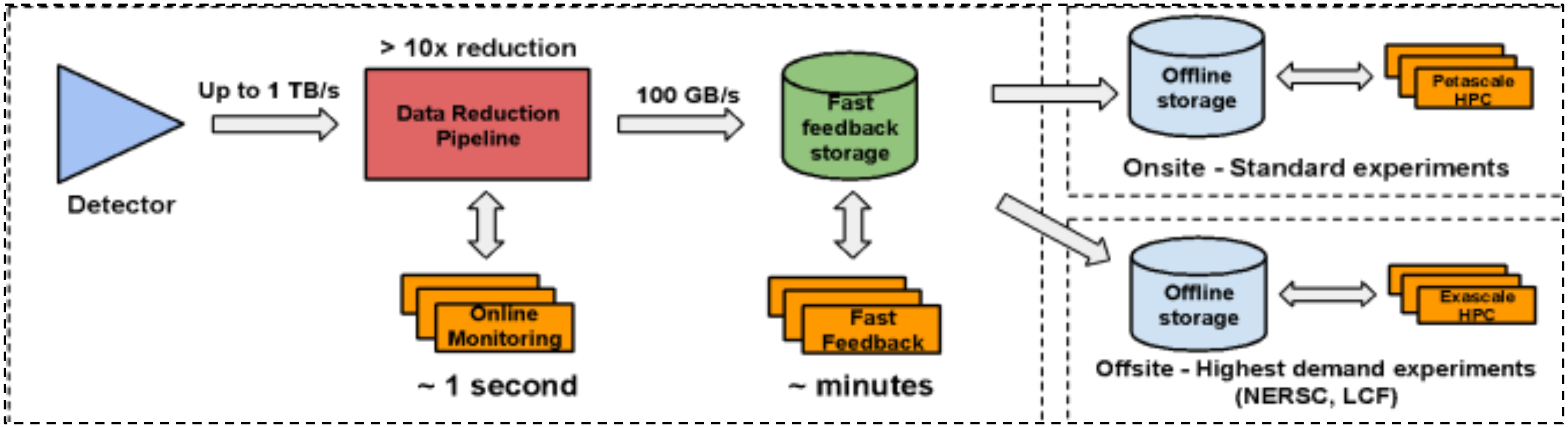
X-ray diffraction image

Intensity map from multiple pulses



Interpretation of system structure / dynamics

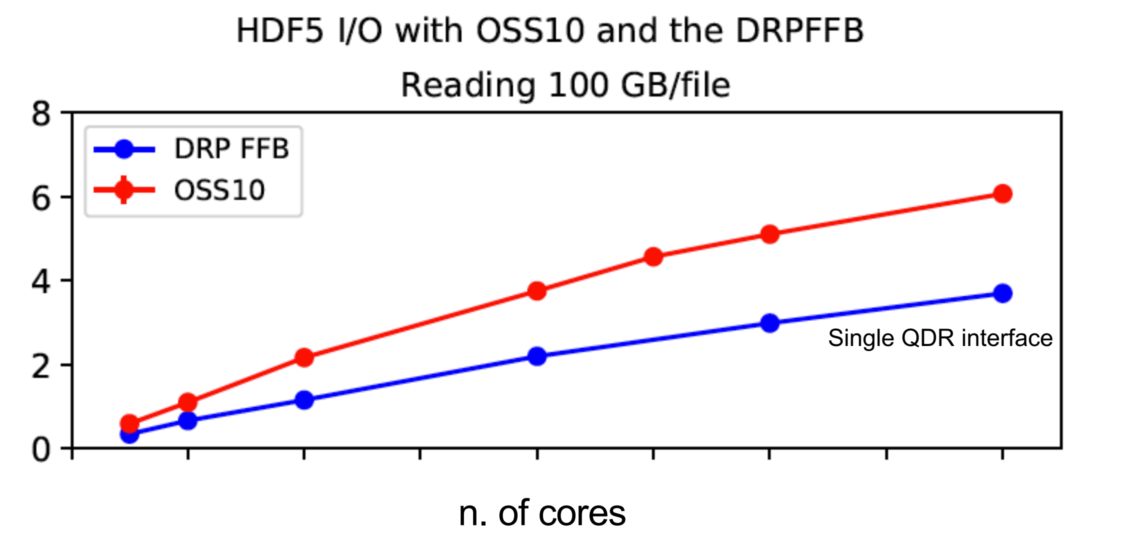




Currently seeking user input on acceptable solutions for data reduction & analysis

Possible storage architecture for Data Reduction Pipeline FFB

- Lustre/ZFS on NVMe SSD devices
- Tested Intel SSD DC P3520
  - ZFS raidz local test vs. Infiniband write



Possible storage architecture for Data Reduction Pipeline FFB (2)

- Lustre/ZFS (raidz) Local obj storage test on 1 OSS

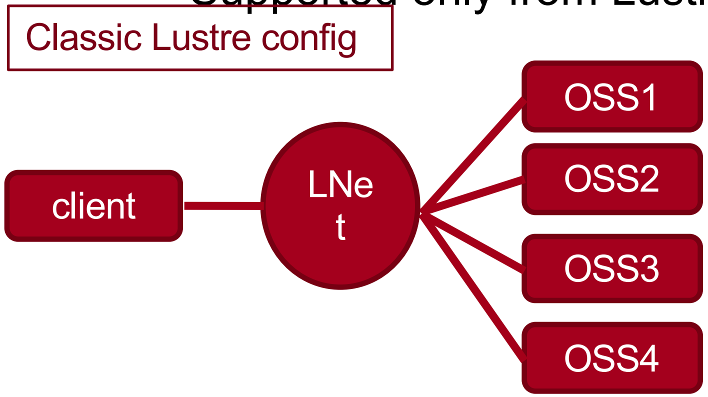
nobjhi=4 thrhi=4 size=10240 case=disk /usr/bin/obdfilter-survey
Sun Nov 5 18:42:57 PST 2017 Obdfilter-survey for case=disk from drp-tst-ffb01
ost 1 sz 10485760K rsz 1024K obj 1 thr 1 write 6367.46
ost 1 sz 10485760K rsz 1024K obj 1 thr 2 write 6839.58
ost 1 sz 10485760K rsz 1024K obj 1 thr 4 write 6436.80
ost 1 sz 10485760K rsz 1024K obj 2 thr 2 write 6689.20
ost 1 sz 10485760K rsz 1024K obj 2 thr 4 write 6979.65
ost 1 sz 10485760K rsz 1024K obj 4 thr 4 write 6737.02

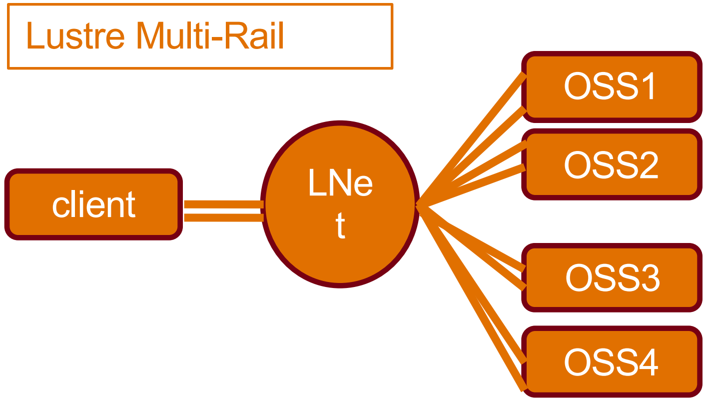
Single QDR interface

n. of cores

Lustre Multi-Rail

- Possibility to improve I/O over InfiniBand between Client and OSS
  - Increasing server bandwidth adding more interfaces to servers and clients
    - Using those interfaces requires a redesign of the LNet network
    - Supported only from Lustre 2.10.*





  - Still under testing

LCLS-II experiments will present challenging computing requirements, in addition to the capacity increase

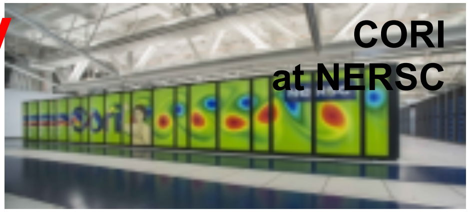
SLAC

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
   Very disruptive for computers that typically host simulations that run for days
4. **Storage** represents significant fraction of the overall system, both in cost and complexity
5. **Throughput** between storage and processing is critical
   Currently most LCLS jobs are I/O limited
6. Speed and flexibility of the **development cycle** is critical
   Wide variety of experiments, with rapid turnaround, and the need to tune data analysis during experiments

These aspects are instrumental also for other SLAC facilities (eg CryoEM, UED, SSRL, FACET-2)

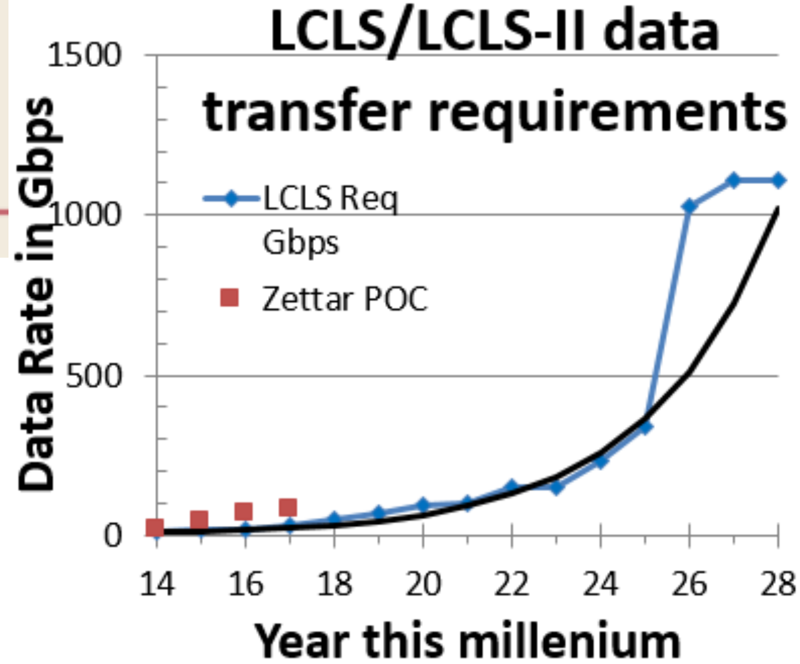
Critical Requirement for Offsite Resources for LCLS Computing

- Several experiments require access to leading edge computers for detailed data analysis. This has its own challenges, in particular:
  - **Need to transfer huge amounts of compressed data from SLAC to supercomputers at other sites**
    - **Providing near real-time results/feedback to the experiment**
    - **Has to be reliable, production quality**



The requirements will need long term agreement between BES and ASCR

Data Transfer requirements

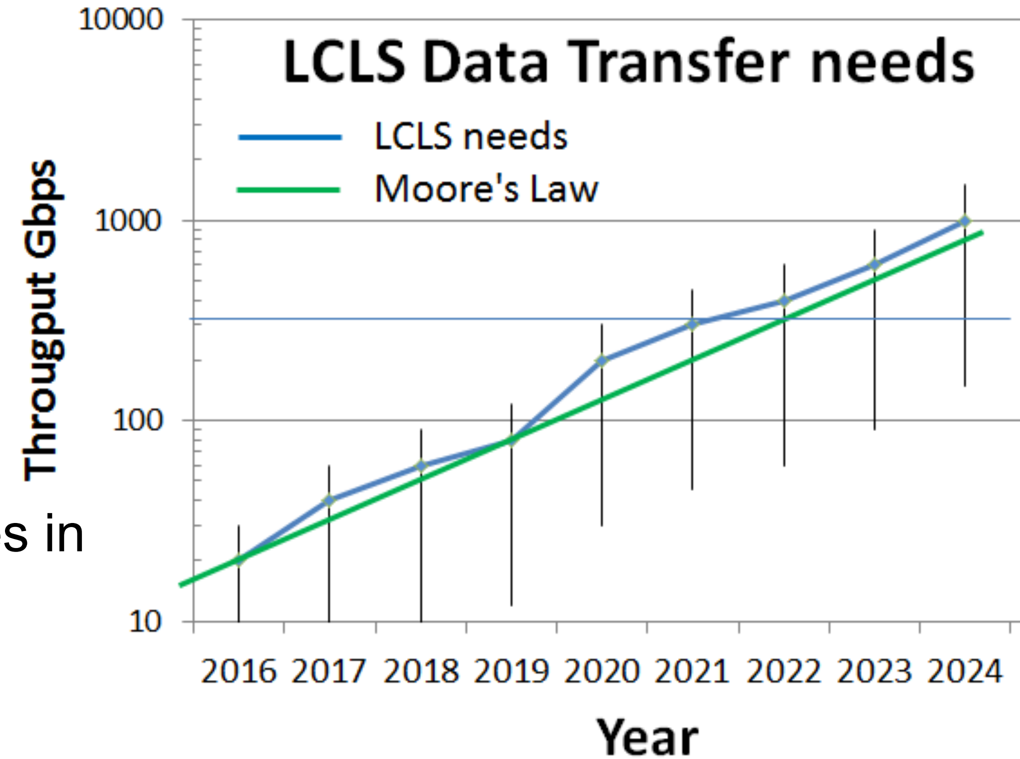


- Today using existing equipment
  - DTNs at NERSC & SLAC
  - With Lustre file systems at ends
  - Today 100Gbps link
  - Using widely used bbcp and xrootd tools
    - Currently ~55Gbps, exploring limitations

- Zettar/zx:
  - Provide HPC data transfer solution (i.e. SW + transfer system reference design):
    - state of the art, efficient, scalable high speed data transfer
  - Over carefully selected demonstration hardware

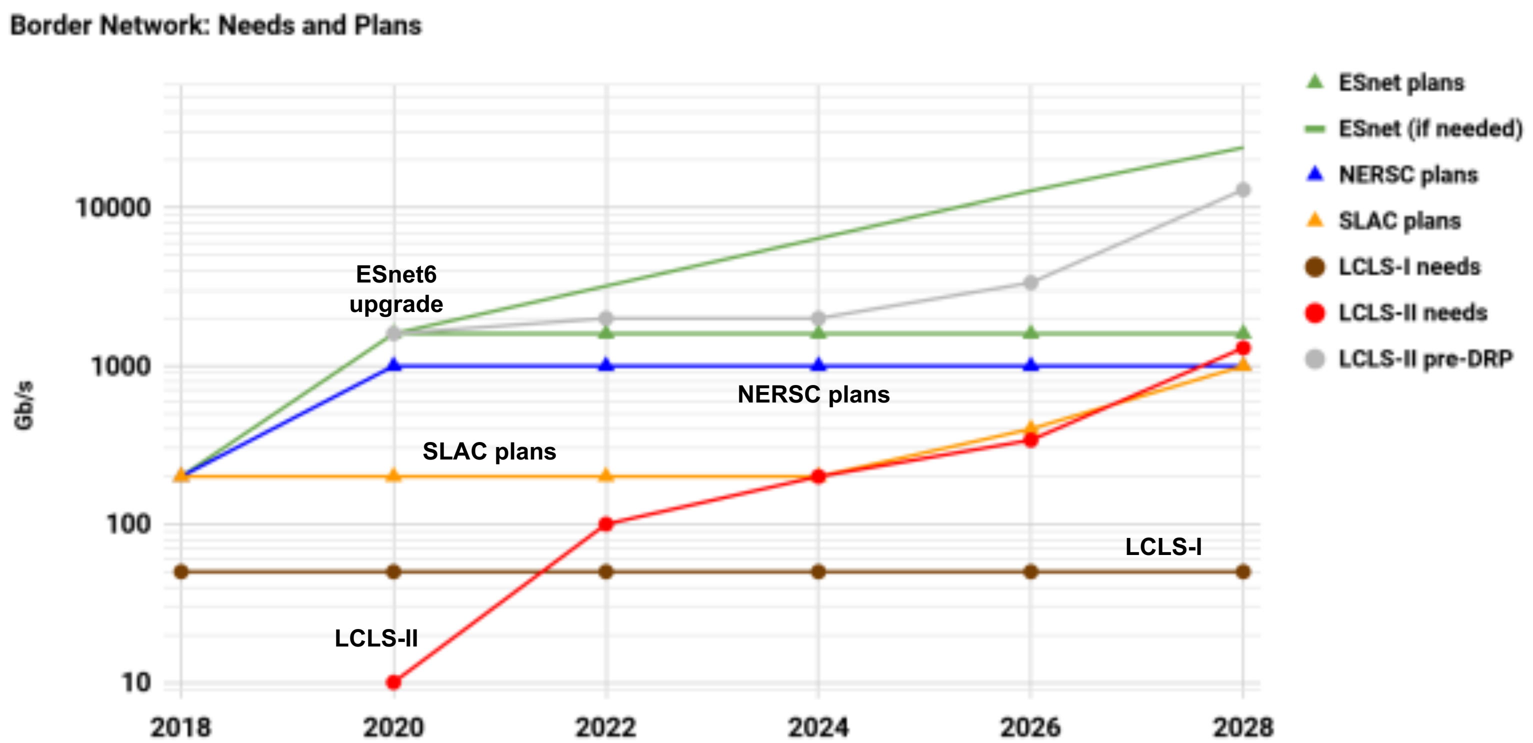
Requirement

- Today 20 Gbps from SLAC to NERSC =>70Gbps 2019
  - Experiments increase efficiency & networking
- 2020 LCLS-II online, data rate 120Hz=>1MHz
  - LCLS-II starts taking data at increased data rate
- 2024 1Tbps:
  - Imaging detectors get faster



- LHC Luminosity increase 10 times in 2020 SLAC ATLAS 35Gbps=>350Gbps

Offsite Data Transfer: Needs and Plans



LCLS-II needs are compatible with SLAC and NERSC plans

More Information

- LCLS SLAC->NERSC 2013
  - http://es.net/science-engagement/case-studies/multi-facility-workflow-case-study/
- LCLS Exascale requirements, Jana Thayer and Amedeo Perazzo
  - https://confluence.slac.stanford.edu/download/attachments/178521813/ExascaleRequirementsLCLSCaseStudy.docx

Questions