Two Day Workshop on Bio-XFEL Data Analysis

BioXFEL
A National Science Foundation Science and Technology Center

August 21 - 22, 2014

at the

Lawrence Berkeley National Lab, Berkeley, California

BERKELEY LAB
Lawrence Berkeley National Laboratory

Application Deadline June 13, 2014

We are pleased to announce the first Workshop on Bio-XFEL Data Analysis, organized as a collaborative effort between the NSF BioXFEL Science and Technology Center, and the Physical Biosciences Division of Lawrence Berkeley National Laboratory.

Serial crystallography is a rapidly growing field with a correspondingly rapidly growing user base. One of the main bottlenecks in serial crystallography is the ability to rapidly filter and analyze large datasets to arrive at accurate structure factors for structure solution and refinement. This workshop is designed as an introduction to serial crystallography analysis tools, and discussion of how to recognize and address data processing challenges and assess the data quality compared to 'normal' synchrotron crystallography.

Day 1 will consist of lecture sessions covering an introduction to serial (femtosecond) crystallography (SFX), software suites available for SFX data analysis and key issues in data processing and assessment of SFX data quality.

Day 2 will involve hands-on computer tutorials and live demonstrations covering the software presented on day 1. These computer lab sessions are limited to 30 participants. The tutorials will be run by the software developers with a team of expert users. LCLS will be providing computing facilities.

Some travels funds available for students. See the website below for an application.

Speakers and Instructors

Paul Adams, *LBNL*
Anton Barty, *CFEL, DESY*
Wolfgang Brehm, *Uni Konstanz*
Aaron Brewster, *LBNL*
Johan Hattne, *HHMI*
James Holton, *LBNL*
Karol Nass, *Max Planck Institute for Medical Research*
Nick Sauter, *LBNL*
Thomas White, *CFEL, DESY*
Nadia Zatsepin, *ASU*
Oliver Zeldin, *Stanford*

NSF · CFEL SCIENCE · ASU Arizona State University · DESY · SLAC National Accelerator Laboratory · LCLS

For more information: www.bioxfel.org/events/details/9

BioXFEL
A National Science Foundation Science and Technology Center

Thursday, August 21, 2014 (8:30 am - 7:00 pm)

Speaker	Title	Description
Nadia Zatsepin	Introduction to workshop and to serial crystallography	
John Spence	How XFEL MX differs from MX and why we do it	
Paul Adams	BioXFEL in the greater context of crystallography	
Anton Barty	Cheetah: Data Reduction and Pre-Processing	Photons, detectors and data streams: how to get at the data and what it records. Data reduction and pre-processing: how it works and how to optimize it.
Karol Nass	CASS: Rapid Data Quality Feedback and Online Monitoring using CASS	Description of the layout of CASS software suite and its functionality on the LCLS online and offline data streams. Real time monitoring of the SFX experiment: hit rate, resolution, pixel saturation, 2D powder pattern. Rapid offline processing of the raw LCLS data: crystal hit finding, extraction of images into HDF5 or cbf formats, pulse characterisitcs, fluorescence spectra, other data specific to an experiment.
Tom White	Indexing and Integration using CrystFEL	CrystFEL is a suite of programs that deal with viewing, indexing, integrating and merging serial crystallography data, evaluating its quality and simulating patterns. At the core of CrystFEL is an automated, high throughput processing pipeline which indexes and integrates each diffraction pattern in a serial crystallography data set. Merging the results yields diffraction data which can be imported into standard crystallographic processing packages for further analysis.

Thursday 21st Auguts	Speaker	Title	Summary
8:30 AM	REGISTRATION		
9:00 AM	Nadia Zatsepin	Introduction to workshop and to serial crystallography	
9:15 AM	John Spence	*How XFEL MX differs from MX and why we do it*	
9:30 AM	Paul Adams	*BioXFEL in the greater context of crystallography*	
9:45 AM	Anton Barty	*Cheetah: Data Reduction and Pre-Processing*	*Photons, detectors and data streams: how to get at the data and what it records. Data reduction and pre-processing: how it works and how to optimize it.*
10:30 AM	Karol Nass	*CASS: Rapid Data Quality Feedback and Online Monitoring using CASS*	*Description of the layout of CASS software suite and its functionality on the LCLS online and offline data streams. Real time monitoring of the SFX experiment: hit rate, resolution, pixel saturation, 2D powder pattern. Rapid offline processing of the raw LCLS data: crystal hit finding, extraction of images into HDF5 or cbf formats, pulse characterisitcs, fluorescence spectra, other data specific to an experiment.*
11:15 AM	COFFEE BREAK		
11:30 AM	Tom White	*Indexing and Integration using CrystFEL*	*CrystFEL is a suite of programs that deal with viewing, indexing, integrating and merging serial crystallography data, evaluating its quality and simulating patterns. At the core of CrystFEL is an automated, high throughput processing pipeline which indexes and integrates each diffraction pattern in a serial crystallography data set. Merging the results yields diffraction data which can be imported into standard crystallographic processing packages for further analysis.*
12:30 PM	LUNCH		
1:30 PM	Nick Sauter	*How good are my data?*	*A design goal of the data reduction package cctbx.xfel is to discover and apply the numerous systematic corrections necessary to correctly model XFEL diffraction images. Numerous challenges must be addressed to interpret serial femtosecond diffraction, including an inaccurate knowledge of the detector's micron-scale metrology, diffraction patterns with multiple lattices and different resolution cutoffs, and the difficulties in modeling crystal orientation from still shots. The ultimate endpoint is to be able to obtain the most accurate structure factors possible by scaling and merging a limited number of images.*
2:10 PM	Aaron Brewster	*Practical considerations during processing of serial crystallographic XFEL data.*	*Fundamental differences exist between processing rotational data collected at a synchrotron source and data collected during a serial crystallographic XFEL experiment. Practical techniques will be described to enable the user to get the most out of their data using the cctbx.xfel package, from properly tuning parameters used during spot finding to common pitfalls encountered while indexing and merging. Specific attention will be paid to real use cases from users in the past, and how problems with their data were addressed.*
2:55 PM	Johan Hattne	*From cctbx to cctbx.xfel and beyond*	*Algorithms to process X-ray diffraction images into accurate atomic models of macromolecules have undergone steady refinement over the past decades. As methods by which crystallographic diffraction patters are obtained are further developed, these techniques continue to provide a solid base from which new processing programs can be bootstrapped. The Computational Crystallography Toolbox (cctbx) collects a wealth of such routines for data analysis and provided the functionality to get cctbx.xfel running within a few months. The presentation will highlight past developments in cctbx.xfel and touch on further generalizations towards processing diffraction data collected from microcrystals using electron microscopes.*
3:35 PM	COFFEE BREAK		
3:50 PM	Oliver Zeldin	*Handling sample heterogeneity in serial cry*	*Serial crystallography data is made up of images obtained from a population of slightly in-homogenous crystals. We will discuss clustering and filtering approaches to (1) characterizing sample heterogeneity in unmerged data, and (2) identifying optimal sub-sets of images for merging and post-refinement.*
4:20 PM	James Holton	*Winning the Battle of Signal vs Noise*	*I will review the major sources of error that contribute to macromolecular crystallography (MX) data with a particular focus on MX data from X-ray Free Electron Lasers. The absolute limits will be defined, and the potential of new and future technologies for overcoming them will be discussed.*
5:20 PM	Wolfgang Brehm	*Clustering and detwinning of SFX data*	*Clustering approaches and variance analysis are explained using the example of the indexing ambiguity in serial crystallography. Recently published algorithms for detwinning and advancements thereof will be shown. Variance analysis can lead to a more accurate result with less data by correcting the main dependent errors. Detecting anisotropy and enhancing the analysis of time dependant data are future perspectives.*
6:05 PM	PANEL discussion		
6:50 PM	Nadia Zatsepin	*Closing remarks*	
7:00 PM	DINNER		
8:30 PM			

SFX data reduction and preprocessing

research papers

Journal of
Applied Crystallography
ISSN 1600-5767

Received 14 November 2013
Accepted 4 April 2014

Cheetah: software for high-throughput reduction and analysis of serial femtosecond X-ray diffraction data

Anton Barty,[a*] Richard A. Kirian,[a] Filipe R. N. C. Maia,[b,c] Max Hantke,[b] Chun Hong Yoon,[a,d] Thomas A. White[a] and Henry Chapman[a,e]

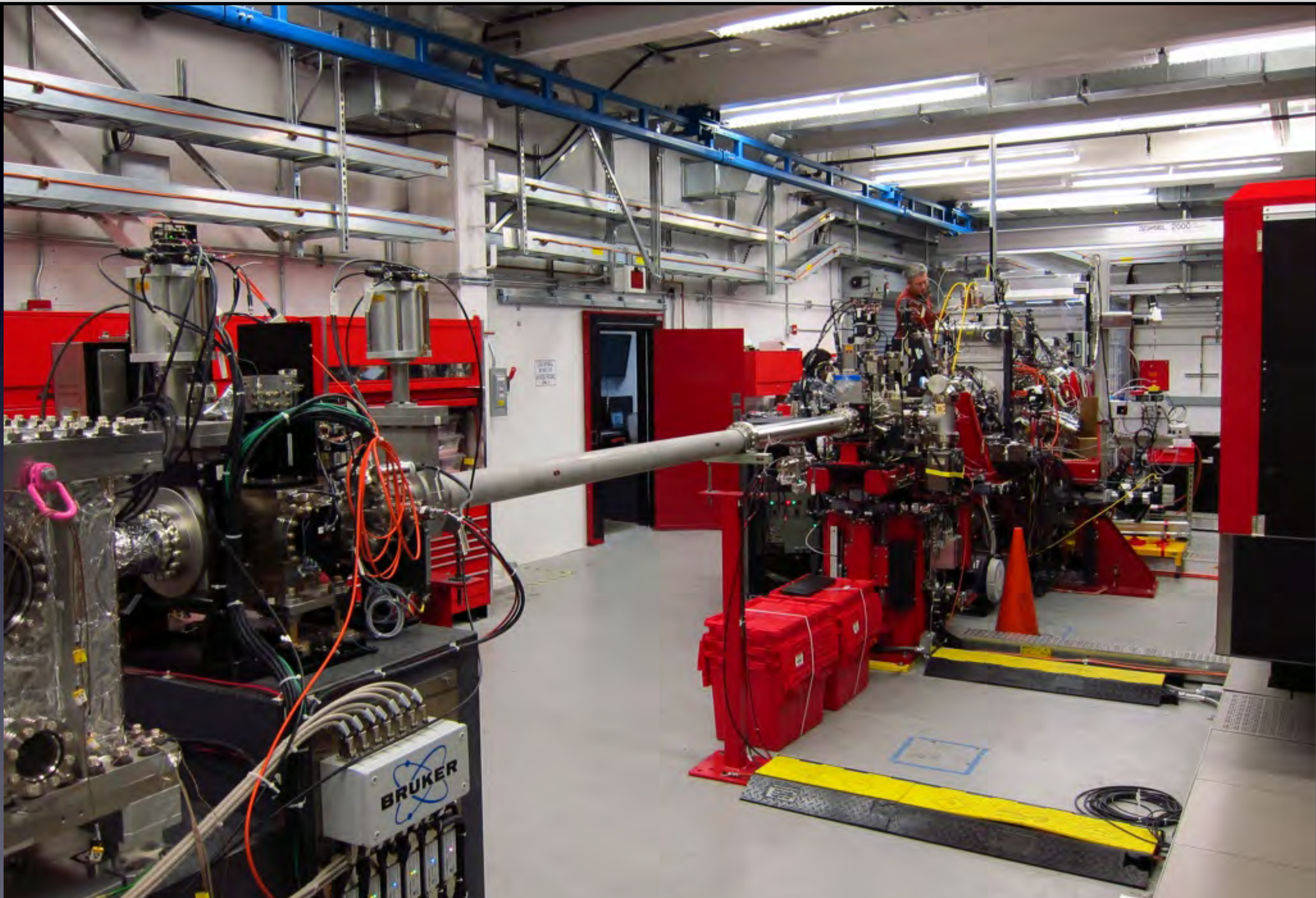
1118 doi:10.1107/S1600576714007626 *J. Appl. Cryst.* (2014). **47**, 1118–1131

Not necessarily sexy, but a necessary evil

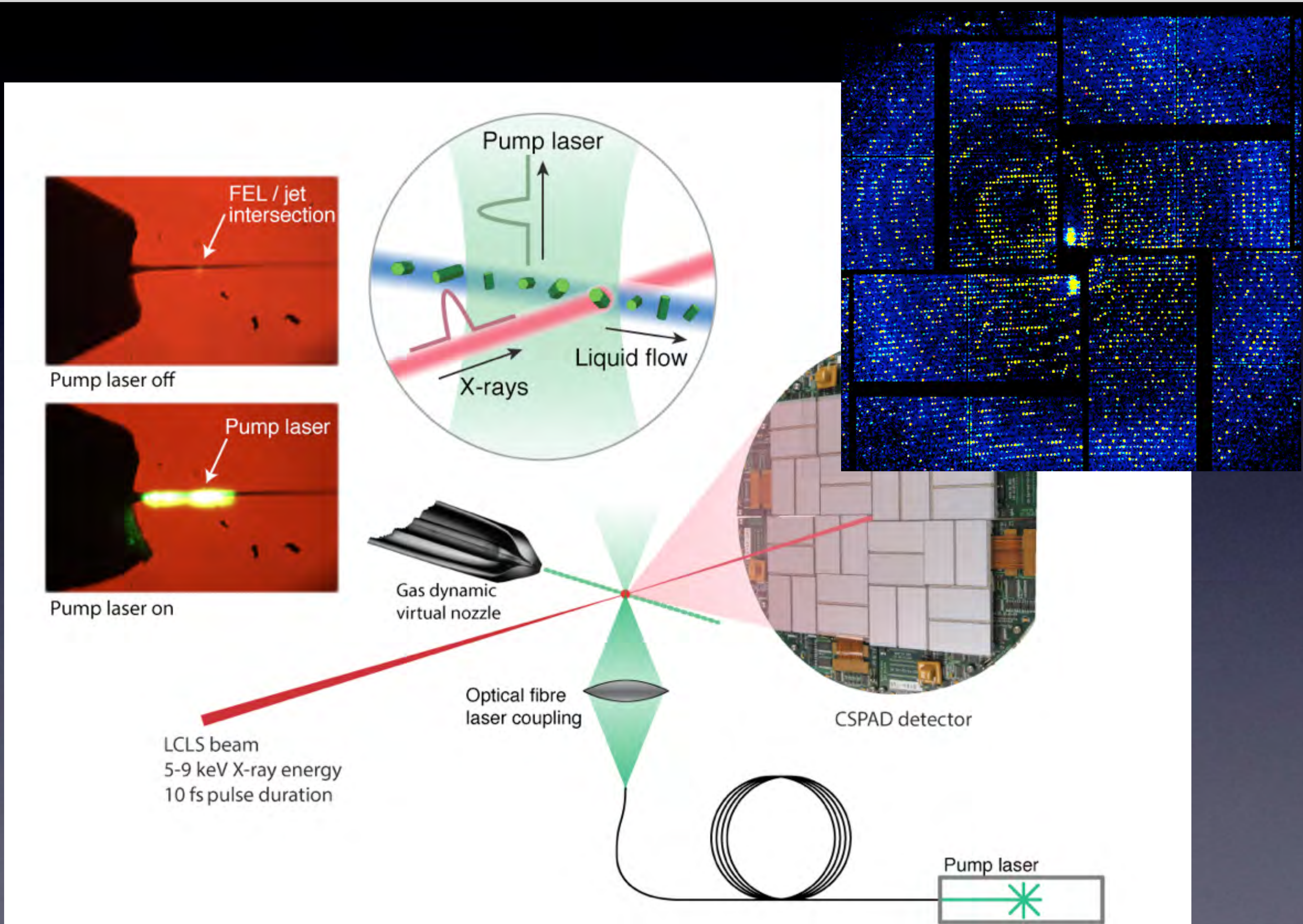
CFEL
SCIENCE

SLAC
NATIONAL ACCELERATOR LABORATORY

The CXI instrument at LCLS delivers more than 10^{12} photons into a 1 μm or 0.1 μm focal spot at 4-9 keV X-ray energy

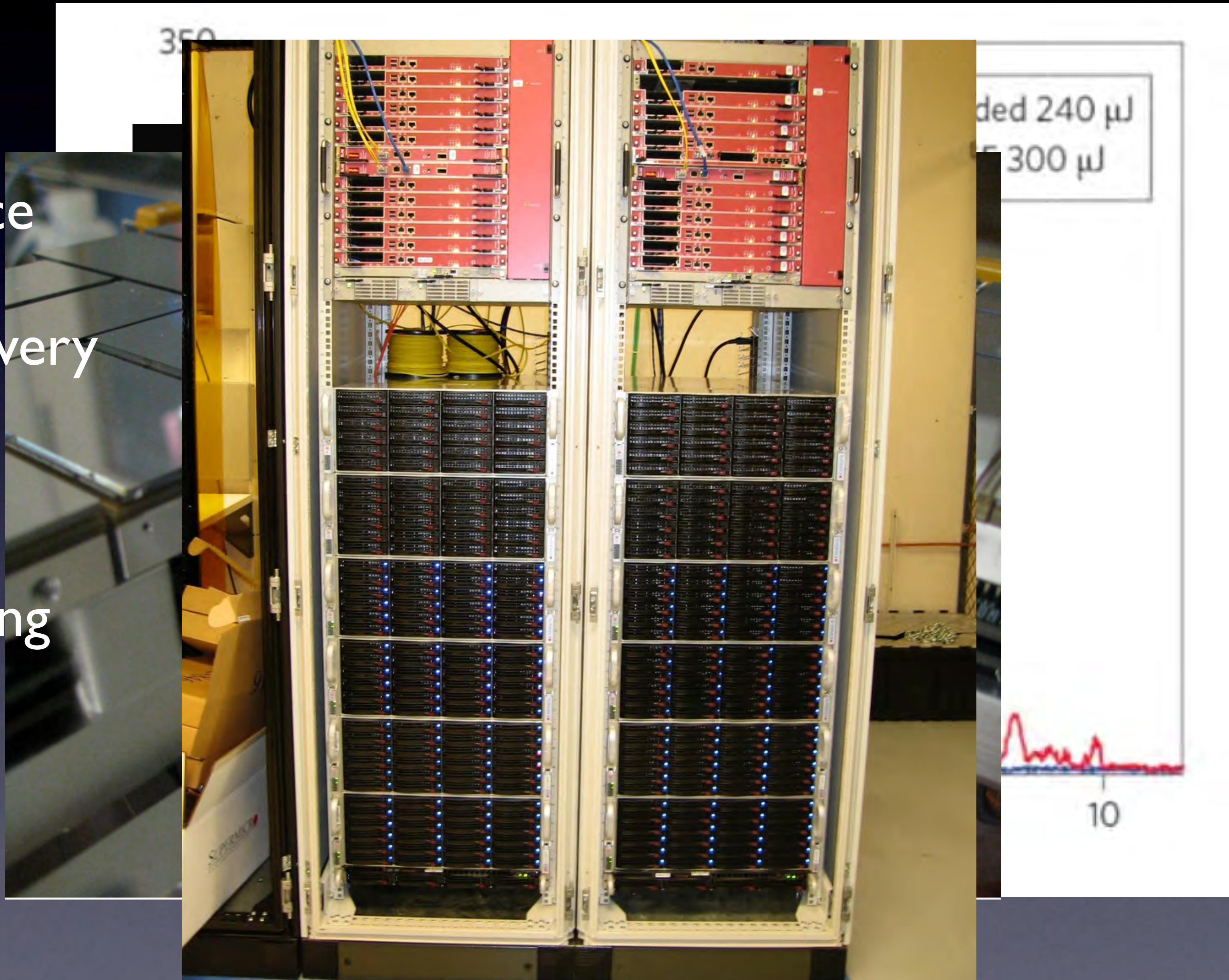


A liquid jet continuously delivers fresh sample to the X-ray focus



SFX experiments at LCLS pose some unique challenges

1. X-ray source
2. Sample delivery
3. Detector
4. Data handling



CFEL SCIENCE

You will soon drown in data

120 frames per second

432,000 frames per hour

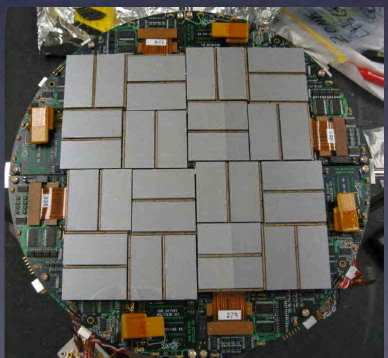
Where is my data ?

WTF is an XTC ?

How do I take 100 TB home ?

How do I read an XTC anyway ?

What now ?



Server 1 → Online analysis → Fast Feeback buffer → Mass storage (XTC format)

Server 2 ...

Server 3 ...

Server 4 ...

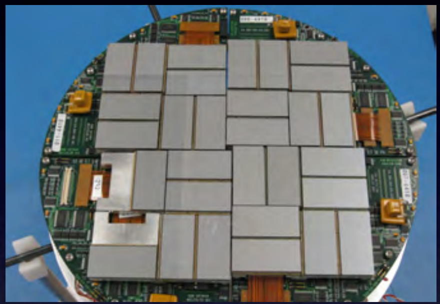
Server 5 ...

```
[psexport01:barty]cxi/cxi52212/xtc> ls -alh *0155*
-r--r-----+ 1 psdatmgr ps-data 92G Oct 23 11:06 e182-r0155-s02-c00.xtc
-r--r-----+ 1 psdatmgr ps-data 20G Oct 23 11:22 e182-r0155-s02-c01.xtc
-r--r-----+ 1 psdatmgr ps-data 92G Oct 23 11:06 e182-r0155-s03-c00.xtc
-r--r-----+ 1 psdatmgr ps-data 20G Oct 23 11:21 e182-r0155-s03-c01.xtc
-r--r-----+ 1 psdatmgr ps-data 93G Oct 23 11:06 e182-r0155-s04-c00.xtc
-r--r-----+ 1 psdatmgr ps-data 20G Oct 23 11:22 e182-r0155-s04-c01.xtc
-r--r-----+ 1 psdatmgr ps-data 93G Oct 23 11:06 e182-r0155-s05-c00.xtc
-r--r-----+ 1 psdatmgr ps-data 20G Oct 23 11:21 e182-r0155-s05-c01.xtc
-r--r-----+ 1 psdatmgr ps-data 92G Oct 23 11:06 e182-r0155-s06-c00.xtc
-r--r-----+ 1 psdatmgr ps-data 20G Oct 23 11:21 e182-r0155-s06-c01.xtc
```

CFEL SCIENCE

Data processing is an exercise in massive data reduction

Detector



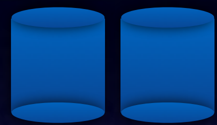
120 Hz
4 MB/event
1.8 TB/hr

Storage

> 4,000,000 frames

Store all data
(no corrections)

Offline reduction

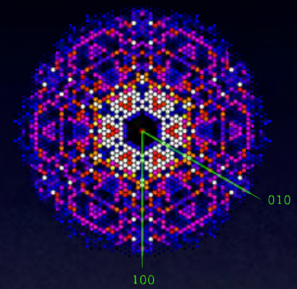


~ 400,000 hits

~ 10,000 hits

Retain only 'hits'
(detector corrected)

Analysis



~ 300,000 indexed

~ 8,000 indexed

Index and
integrate

Automated high volume image processing is essential
(eg: background correction, weeding useful data from useless data)

CFEL SCIENCE

Step 1: Organise your data
A shared spreadsheet is ideal for remembering what is in each run

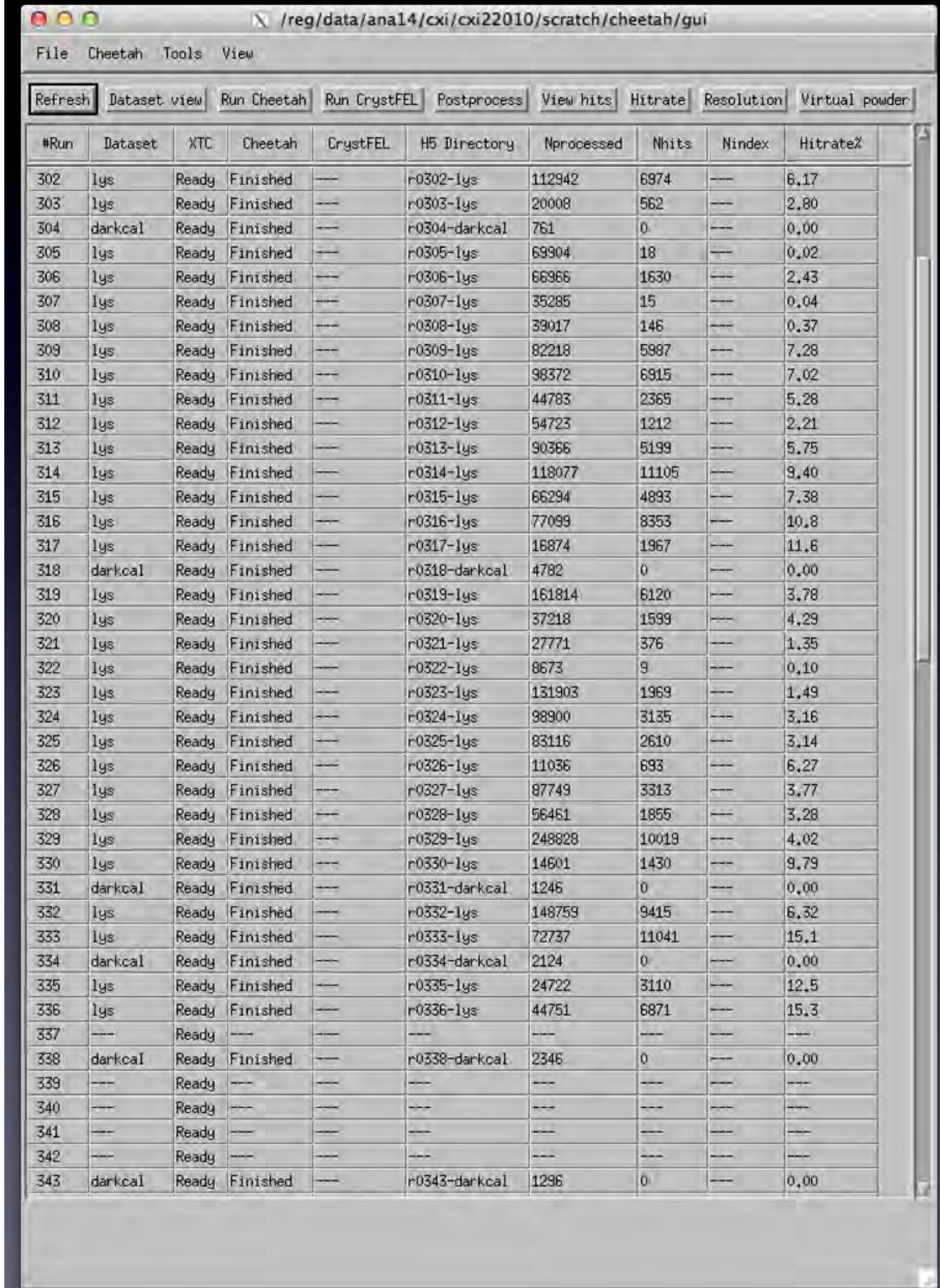
CXI L669 Chapman June 2013 MAD beamtime

Electronic logbook: https://pswww.slac.stanford.edu/apps/portal/index.php?exper_id=251 — Total number of p

	Run #	Date	Time	Samp Numb	Sample	Sample owner	Comments	Filter/Transmission	Post Sample Transmission	Analysis results
89	85	27 Jun 2013	02:58:23	4	KI 0.05 M		Calibration for filter checks, catcher z= 3.5, x= 1.1775, y= -0.742 T=100% 25uL/min 200psi			
90	86	27 Jun 2013	03:00:57	4	KI 0.05 M		Calibration without filters, same catcher pos.			
91	87	27 Jun 2013	06:58:12		Dark Run			Filter = post attenuation		
92	88	27 Jun 2013	07:00:24	5	~~Cat B (conc)~~	Francesco	1.5 x Concentrated sample 3, no hits	No Filter T=7%		water by mistake
93	89	27 Jun 2013	07:02:17	5	~~Cat B (conc)~~		jet not stable, no hits 10uL/min 1050psi	No Filter T=7%		water by mistake
94	90	27 Jun 2013	07:08:49	5	~~Cat B (conc)~~		chasing crystals	No Filter T=14%		water by mistake
95	91	27 Jun 2013	07:15:56	5	~~Cat B (conc)~~		z=3, out of focus to try to increase hit chance	No Filter T=52%		water by mistake
96	92	27 Jun 2013	07:32:43	5	~~Cat B (conc)~~		inline filter 2um to 10um PEEK, 8uL/min 1000psi	No Filter T=14%		water by mistake
97	93	27 Jun 2013	07:54:21	6	~~Lysozyme 1A~~	ilme	removed Ti/In filters, no hits apparently just running water, checking connections	No Filter T=7%		water by mistake
98	94	27 Jun 2013	08:13:47		Dark Run					
99	95	27 Jun 2013	08:21:27	5	Cat B (conc)	Francesco	10uL/min 900psi 10 um PEEK inline filter	No Filter T=14%		
100	96	27 Jun 2013	08:37:37	5	Cat B (conc)		10uL/min 1060psi 10um inline filter No Ti/In filter	No Filter T=14%		
101	97	27 Jun 2013	08:47:16	5	Cat B (conc)		15uL/min 1500psi 10um inline filter No Ti/In filter	No Filter T=14%		
102	98	27 Jun 2013	09:00:34		Dark Run					
103							End of 2nd shift			
104	99	27 Jun 2013	23:33:30		Dark Run					
105	100	27 Jun 2013	23:41:28	7	Cat B (pooled)	Francesco	10uL/min 2um PEEK	No Filter T=3.7%		
106	101	27 Jun 2013	23:52:58	7	Cat B (pooled)	Francesco	10 -> 15uL/min 2um PEEK	No Filter T=3.7%		
107	102	28 Jun 2013	00:01:53	7	Cat B (pooled)	Francesco	15uL/min 2um PEEK	No Filter T=3.7%		
108	103	28 Jun 2013	00:10:51	7	Cat B (pooled)	Francesco	15uL/min 2um PEEK	No Filter T=3.7%		
109	104	28 Jun 2013	00:19:44	7	Cat B (pooled)	Francesco	15uL/min 2um PEEK 300psi	No Filter T=14%		really short run
110	105	28 Jun 2013	00:19:48	7	Cat B (pooled)	Francesco	15uL/min 2um PEEK	No Filter T=14%		
111	106	28 Jun 2013	00:29:54	7	Cat B (pooled)	Francesco	15uL/min 2um PEEK	No Filter T=14%		
112	107	28 Jun 2013	00:31:12	7	Cat B (pooled)	Francesco	15uL/min 2um PEEK	No Filter T=27%		
113	108	28 Jun 2013	00:42:38	7	Cat B (pooled)	Francesco	15uL/min 2um PEEK	No Filter T=27%		

Run Log | Nozzles | Sample Log | Fluores | Sample List | Travel Dates | notes | inventory (non shipped)

Cheetah is used for data reduction, rapid data evaluation, and translation



#Run	Dataset	XTC	Cheetah	CrystFEL	H5 Directory	Nprocessed	Nhits	Nindex	Hitrate%
302	lys	Ready	Finished	---	r0302-lys	112942	6974	---	6.17
303	lys	Ready	Finished	---	r0303-lys	20008	562	---	2.80
304	darkcal	Ready	Finished	---	r0304-darkcal	761	0	---	0.00
305	lys	Ready	Finished	---	r0305-lys	69904	18	---	0.02
306	lys	Ready	Finished	---	r0306-lys	66966	1630	---	2.43
307	lys	Ready	Finished	---	r0307-lys	35285	15	---	0.04
308	lys	Ready	Finished	---	r0308-lys	39017	146	---	0.37
309	lys	Ready	Finished	---	r0309-lys	82218	5987	---	7.28
310	lys	Ready	Finished	---	r0310-lys	98372	6915	---	7.02
311	lys	Ready	Finished	---	r0311-lys	44783	2365	---	5.28
312	lys	Ready	Finished	---	r0312-lys	54723	1212	---	2.21
313	lys	Ready	Finished	---	r0313-lys	90366	5199	---	5.75
314	lys	Ready	Finished	---	r0314-lys	118077	11105	---	9.40
315	lys	Ready	Finished	---	r0315-lys	66294	4893	---	7.38
316	lys	Ready	Finished	---	r0316-lys	77099	8353	---	10.8
317	lys	Ready	Finished	---	r0317-lys	16874	1967	---	11.6
318	darkcal	Ready	Finished	---	r0318-darkcal	4782	0	---	0.00
319	lys	Ready	Finished	---	r0319-lys	161814	6120	---	3.78
320	lys	Ready	Finished	---	r0320-lys	37218	1599	---	4.29
321	lys	Ready	Finished	---	r0321-lys	27771	376	---	1.35
322	lys	Ready	Finished	---	r0322-lys	8673	9	---	0.10
323	lys	Ready	Finished	---	r0323-lys	131903	1969	---	1.49
324	lys	Ready	Finished	---	r0324-lys	98900	3135	---	3.16
325	lys	Ready	Finished	---	r0325-lys	83116	2610	---	3.14
326	lys	Ready	Finished	---	r0326-lys	11036	693	---	6.27
327	lys	Ready	Finished	---	r0327-lys	87749	3313	---	3.77
328	lys	Ready	Finished	---	r0328-lys	56461	1855	---	3.28
329	lys	Ready	Finished	---	r0329-lys	248828	10019	---	4.02
330	lys	Ready	Finished	---	r0330-lys	14601	1430	---	9.79
331	darkcal	Ready	Finished	---	r0331-darkcal	1246	0	---	0.00
332	lys	Ready	Finished	---	r0332-lys	148759	9415	---	6.32
333	lys	Ready	Finished	---	r0333-lys	72737	11041	---	15.1
334	darkcal	Ready	Finished	---	r0334-darkcal	2124	0	---	0.00
335	lys	Ready	Finished	---	r0335-lys	24722	3110	---	12.5
336	lys	Ready	Finished	---	r0336-lys	44751	6871	---	15.3
337	---	Ready	---	---	---	---	---	---	---
338	darkcal	Ready	Finished	---	r0338-darkcal	2346	0	---	0.00
339	---	Ready	---	---	---	---	---	---	---
340	---	Ready	---	---	---	---	---	---	---
341	---	Ready	---	---	---	---	---	---	---
342	---	Ready	---	---	---	---	---	---	---
343	darkcal	Ready	Finished	---	r0343-darkcal	1296	0	---	0.00

1. 'Control panel' interface to data and LCLS analysis cluster

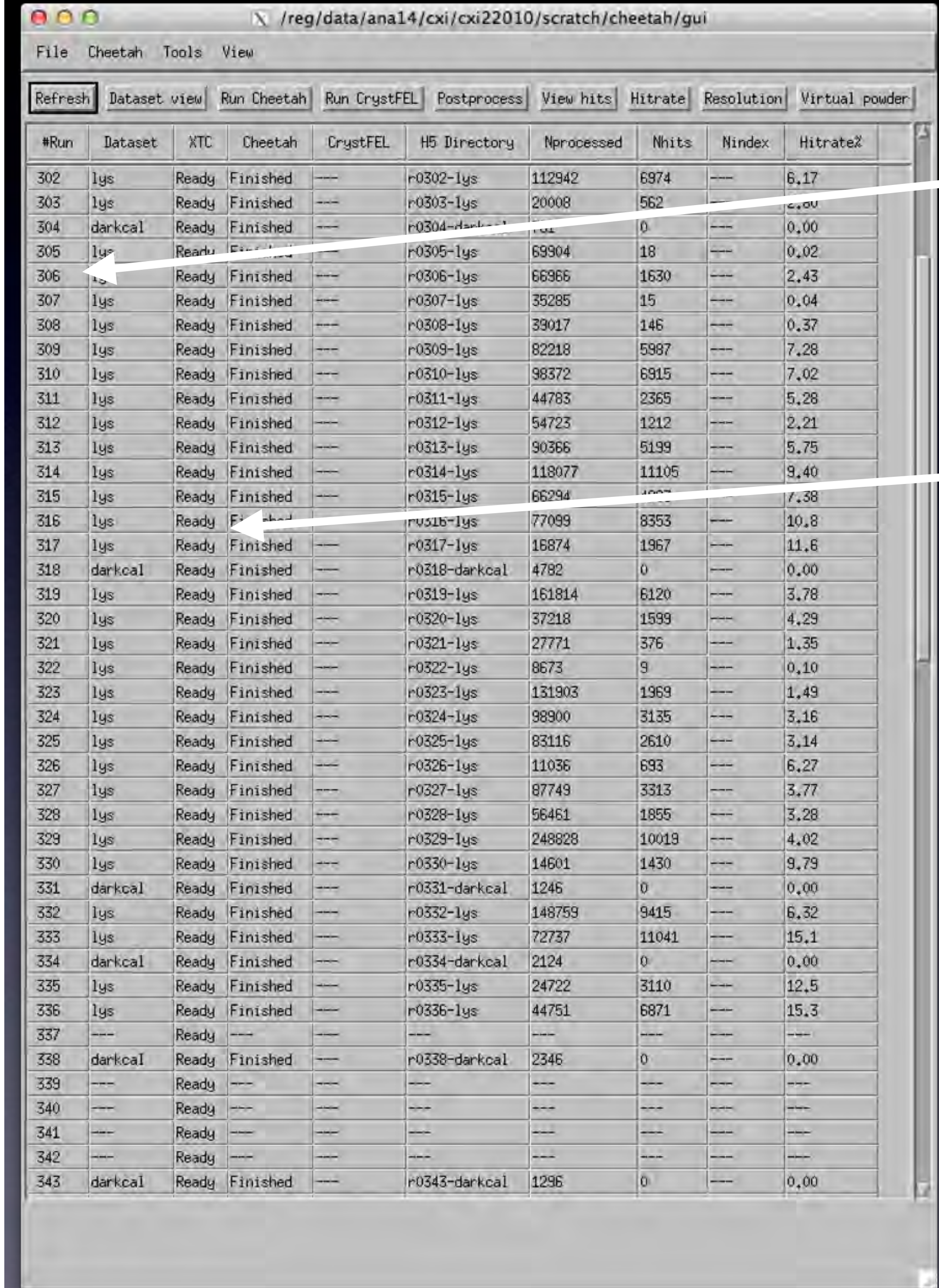
2. Rapid feedback
   Hit rate, resolution, diffraction quality
   Quickly viewing images

3. Data reduction
   Keeps only useful events crystals (ie: frames with crystal diffraction)

4. Data translation
   XTC data is converted to a facility independent format (HDF5)

5. Data organisation
   Summarises what is in each run; easy to group data by sample; summarises statistics

Cheetah functionality: XTC monitor

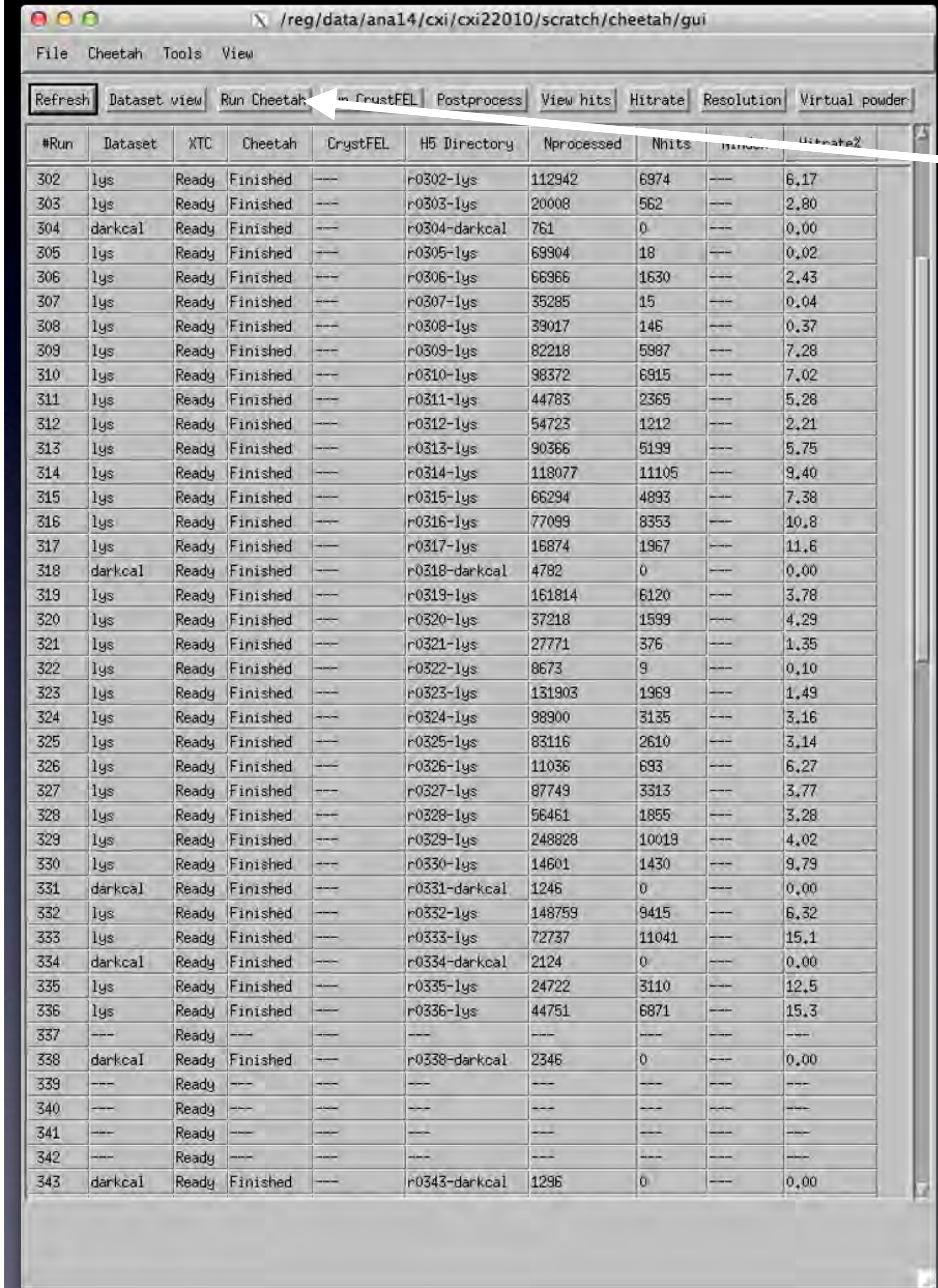


#Run	Dataset	XTC	Cheetah	CrystFEL	H5 Directory	Nprocessed	Nhits	Nindex	Hitrate%
302	lys	Ready	Finished	---	r0302-lys	112942	6974	---	6.17
303	lys	Ready	Finished	---	r0303-lys	20008	562	[illegible]	[illegible]
304	darkcal	Ready	Finished	---	[illegible]	[illegible]	0	---	0.00
305	[illegible]	[illegible]	[illegible]	[illegible]	r0305-lys	69904	18	---	0.02
306	[illegible]	Ready	Finished	---	r0306-lys	66966	1630	---	2.43
307	lys	Ready	Finished	---	r0307-lys	35285	15	---	0.04
308	lys	Ready	Finished	---	r0308-lys	39017	146	---	0.37
309	lys	Ready	Finished	---	r0309-lys	82218	5987	---	7.28
310	lys	Ready	Finished	---	r0310-lys	98372	6915	---	7.02
311	lys	Ready	Finished	---	r0311-lys	44783	2365	---	5.28
312	lys	Ready	Finished	---	r0312-lys	54723	1212	---	2.21
313	lys	Ready	Finished	---	r0313-lys	90366	5199	---	5.75
314	lys	Ready	Finished	---	r0314-lys	118077	11105	---	9.40
315	lys	Ready	Finished	---	r0315-lys	66294	[illegible]	[illegible]	[illegible]
316	lys	Ready	[illegible]	[illegible]	[illegible]	77099	8353	---	10.8
317	lys	Ready	Finished	---	r0317-lys	16874	1967	---	11.6
318	darkcal	Ready	Finished	---	r0318-darkcal	4782	0	---	0.00
319	lys	Ready	Finished	---	r0319-lys	161814	6120	---	3.78
320	lys	Ready	Finished	---	r0320-lys	37218	1599	---	4.29
321	lys	Ready	Finished	---	r0321-lys	27771	376	---	1.35
322	lys	Ready	Finished	---	r0322-lys	8673	9	---	0.10
323	lys	Ready	Finished	---	r0323-lys	131903	1969	---	1.49
324	lys	Ready	Finished	---	r0324-lys	98900	3135	---	3.16
325	lys	Ready	Finished	---	r0325-lys	83116	2610	---	3.14
326	lys	Ready	Finished	---	r0326-lys	11036	693	---	6.27
327	lys	Ready	Finished	---	r0327-lys	87749	3313	---	3.77
328	lys	Ready	Finished	---	r0328-lys	56461	1855	---	3.28
329	lys	Ready	Finished	---	r0329-lys	248828	10019	---	4.02
330	lys	Ready	Finished	---	r0330-lys	14601	1430	---	9.79
331	darkcal	Ready	Finished	---	r0331-darkcal	1246	0	---	0.00
332	lys	Ready	Finished	---	r0332-lys	148759	9415	---	6.32
333	lys	Ready	Finished	---	r0333-lys	72737	11041	---	15.1
334	darkcal	Ready	Finished	---	r0334-darkcal	2124	0	---	0.00
335	lys	Ready	Finished	---	r0335-lys	24722	3110	---	12.5
336	lys	Ready	Finished	---	r0336-lys	44751	6871	---	15.3
337	---	Ready	---	---	---	---	---	---	---
338	darkcal	Ready	Finished	---	r0338-darkcal	2346	0	---	0.00
339	---	Ready	---	---	---	---	---	---	---
340	---	Ready	---	---	---	---	---	---	---
341	---	Ready	---	---	---	---	---	---	---
342	---	Ready	---	---	---	---	---	---	---
343	darkcal	Ready	Finished	---	r0343-darkcal	1296	0	---	0.00

Newly collected data (new runs) appear automatically ready to process

Status of data collection

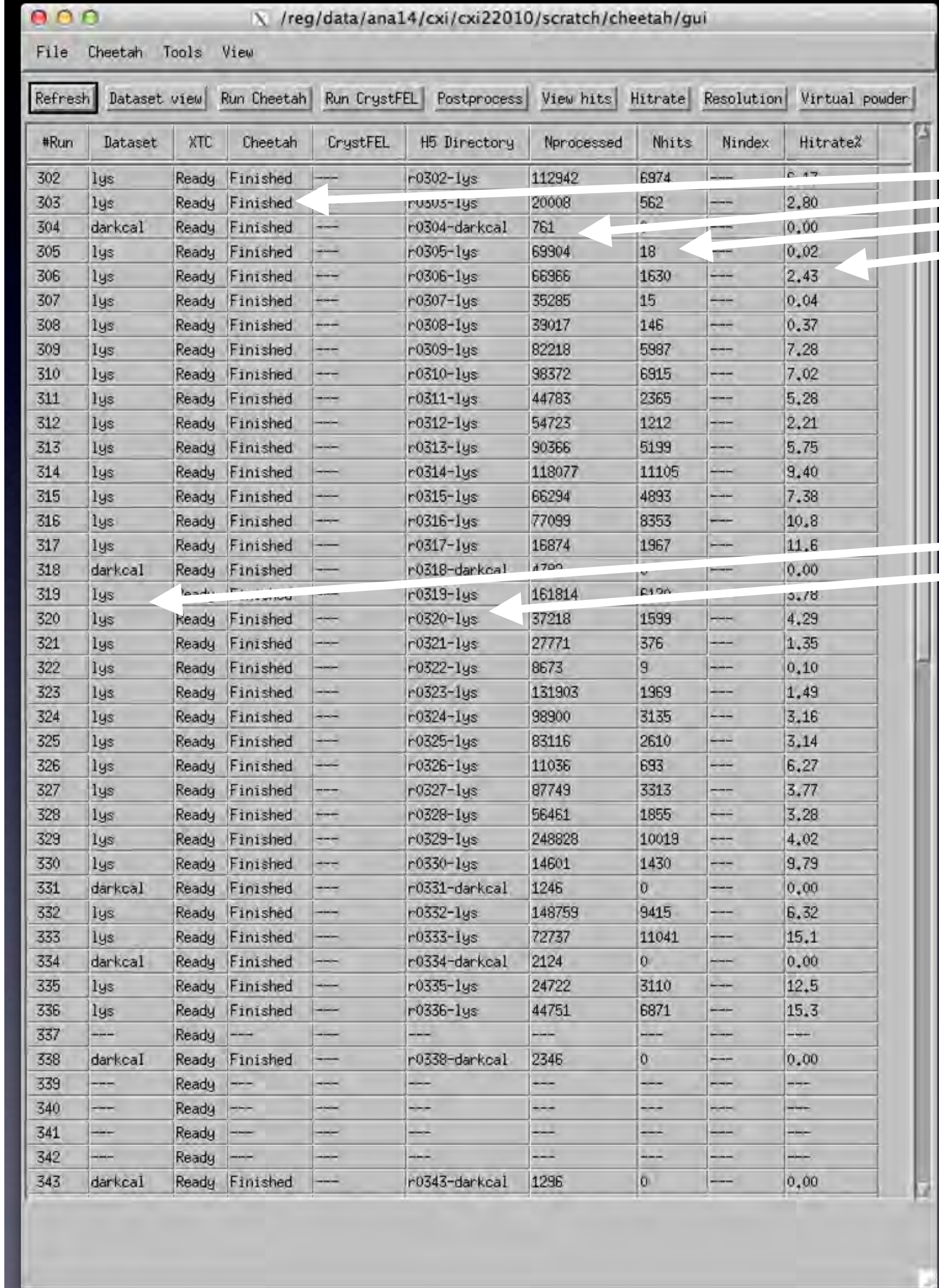
Cheetah functionality: Processing control



#Run	Dataset	XTC	Cheetah	CrystFEL	H5 Directory	Nprocessed	Nhits	[illegible]	Hitrate%
302	lys	Ready	Finished	---	r0302-lys	112942	6974	---	6.17
303	lys	Ready	Finished	---	r0303-lys	20008	562	---	2.80
304	darkcal	Ready	Finished	---	r0304-darkcal	761	0	---	0.00
305	lys	Ready	Finished	---	r0305-lys	69904	18	---	0.02
306	lys	Ready	Finished	---	r0306-lys	66966	1630	---	2.43
307	lys	Ready	Finished	---	r0307-lys	35285	15	---	0.04
308	lys	Ready	Finished	---	r0308-lys	39017	146	---	0.37
309	lys	Ready	Finished	---	r0309-lys	82218	5987	---	7.28
310	lys	Ready	Finished	---	r0310-lys	98372	6915	---	7.02
311	lys	Ready	Finished	---	r0311-lys	44783	2365	---	5.28
312	lys	Ready	Finished	---	r0312-lys	54723	1212	---	2.21
313	lys	Ready	Finished	---	r0313-lys	90366	5199	---	5.75
314	lys	Ready	Finished	---	r0314-lys	118077	11105	---	9.40
315	lys	Ready	Finished	---	r0315-lys	66294	4893	---	7.38
316	lys	Ready	Finished	---	r0316-lys	77099	8353	---	10.8
317	lys	Ready	Finished	---	r0317-lys	16874	1967	---	11.6
318	darkcal	Ready	Finished	---	r0318-darkcal	4782	0	---	0.00
319	lys	Ready	Finished	---	r0319-lys	161814	6120	---	3.78
320	lys	Ready	Finished	---	r0320-lys	37218	1599	---	4.29
321	lys	Ready	Finished	---	r0321-lys	27771	376	---	1.35
322	lys	Ready	Finished	---	r0322-lys	8673	9	---	0.10
323	lys	Ready	Finished	---	r0323-lys	131903	1969	---	1.49
324	lys	Ready	Finished	---	r0324-lys	98900	3135	---	3.16
325	lys	Ready	Finished	---	r0325-lys	83116	2610	---	3.14
326	lys	Ready	Finished	---	r0326-lys	11036	693	---	6.27
327	lys	Ready	Finished	---	r0327-lys	87749	3313	---	3.77
328	lys	Ready	Finished	---	r0328-lys	56461	1855	---	3.28
329	lys	Ready	Finished	---	r0329-lys	248828	10019	---	4.02
330	lys	Ready	Finished	---	r0330-lys	14601	1430	---	9.79
331	darkcal	Ready	Finished	---	r0331-darkcal	1246	0	---	0.00
332	lys	Ready	Finished	---	r0332-lys	148759	9415	---	6.32
333	lys	Ready	Finished	---	r0333-lys	72737	11041	---	15.1
334	darkcal	Ready	Finished	---	r0334-darkcal	2124	0	---	0.00
335	lys	Ready	Finished	---	r0335-lys	24722	3110	---	12.5
336	lys	Ready	Finished	---	r0336-lys	44751	6871	---	15.3
337	---	Ready	---	---	---	---	---	---	---
338	darkcal	Ready	Finished	---	r0338-darkcal	2346	0	---	0.00
339	---	Ready	---	---	---	---	---	---	---
340	---	Ready	---	---	---	---	---	---	---
341	---	Ready	---	---	---	---	---	---	---
342	---	Ready	---	---	---	---	---	---	---
343	darkcal	Ready	Finished	---	r0343-darkcal	1296	0	---	0.00

One-click to start the processing of data sets

Cheetah functionality: Processing status monitor

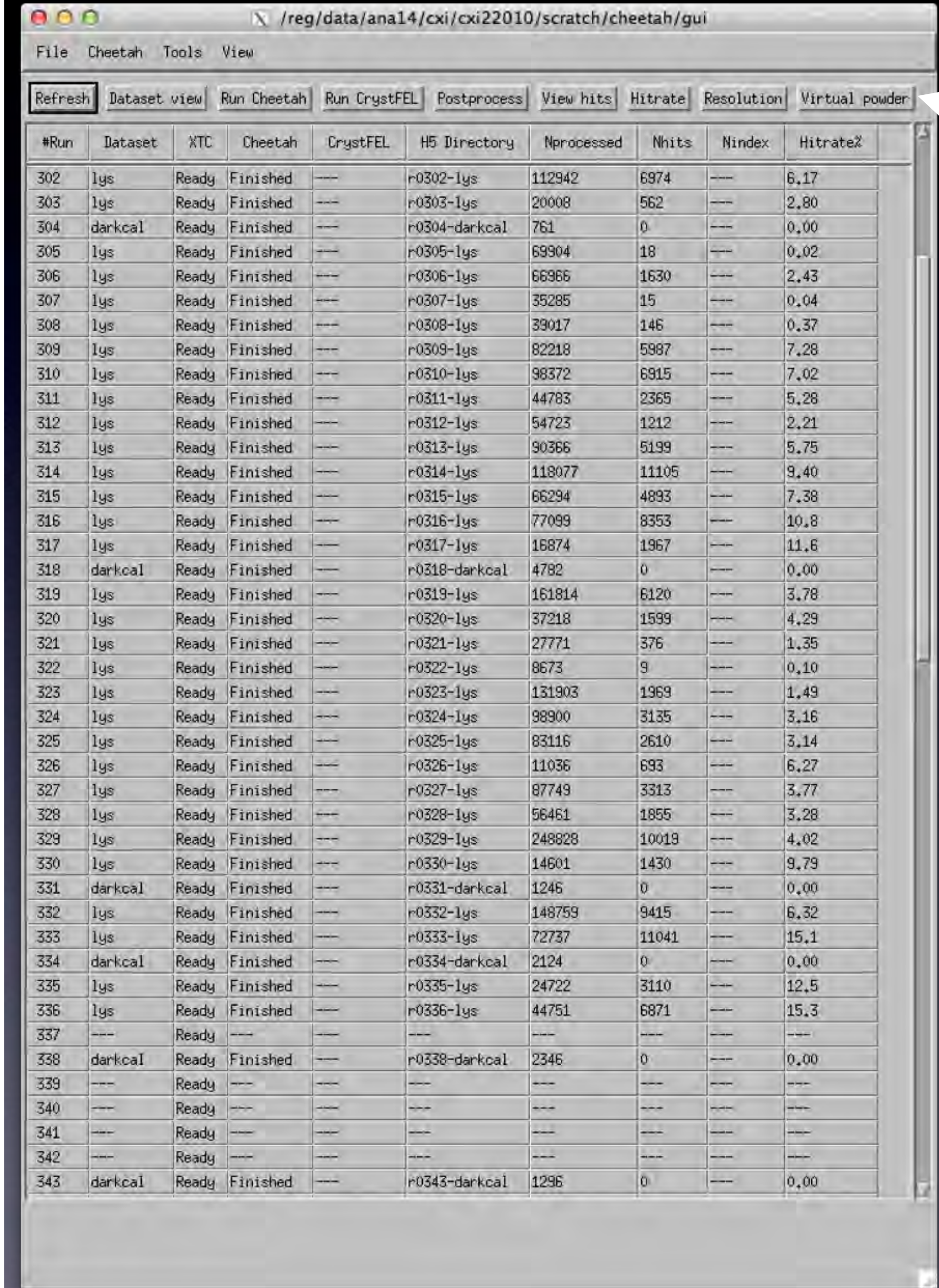


#Run	Dataset	XTC	Cheetah	CrystFEL	H5 Directory	Nprocessed	Nhits	Nindex	Hitrate%
302	lys	Ready	Finished	---	r0302-lys	112942	6974	---	[illegible]
303	lys	Ready	Finished	[illegible]	[illegible]	20008	562	---	2.80
304	darkcal	Ready	Finished	---	r0304-darkcal	761	[illegible]	[illegible]	0.00
305	lys	Ready	Finished	---	r0305-lys	69904	18	[illegible]	0.02
306	lys	Ready	Finished	---	r0306-lys	66966	1630	---	2.43
307	lys	Ready	Finished	---	r0307-lys	35285	15	---	0.04
308	lys	Ready	Finished	---	r0308-lys	39017	146	---	0.37
309	lys	Ready	Finished	---	r0309-lys	82218	5987	---	7.28
310	lys	Ready	Finished	---	r0310-lys	98372	6915	---	7.02
311	lys	Ready	Finished	---	r0311-lys	44783	2365	---	5.28
312	lys	Ready	Finished	---	r0312-lys	54723	1212	---	2.21
313	lys	Ready	Finished	---	r0313-lys	90366	5199	---	5.75
314	lys	Ready	Finished	---	r0314-lys	118077	11105	---	9.40
315	lys	Ready	Finished	---	r0315-lys	66294	4893	---	7.38
316	lys	Ready	Finished	---	r0316-lys	77099	8353	---	10.8
317	lys	Ready	Finished	---	r0317-lys	16874	1967	---	11.6
318	darkcal	Ready	Finished	---	r0318-darkcal	[illegible]	[illegible]	---	0.00
319	lys	[illegible]	[illegible]	[illegible]	r0319-lys	161814	[illegible]	[illegible]	[illegible]
320	lys	Ready	Finished	---	r0320-lys	37218	1599	---	4.29
321	lys	Ready	Finished	---	r0321-lys	27771	376	---	1.35
322	lys	Ready	Finished	---	r0322-lys	8673	9	---	0.10
323	lys	Ready	Finished	---	r0323-lys	131903	1969	---	1.49
324	lys	Ready	Finished	---	r0324-lys	98900	3135	---	3.16
325	lys	Ready	Finished	---	r0325-lys	83116	2610	---	3.14
326	lys	Ready	Finished	---	r0326-lys	11036	693	---	6.27
327	lys	Ready	Finished	---	r0327-lys	87749	3313	---	3.77
328	lys	Ready	Finished	---	r0328-lys	56461	1855	---	3.28
329	lys	Ready	Finished	---	r0329-lys	248828	10019	---	4.02
330	lys	Ready	Finished	---	r0330-lys	14601	1430	---	9.79
331	darkcal	Ready	Finished	---	r0331-darkcal	1246	0	---	0.00
332	lys	Ready	Finished	---	r0332-lys	148759	9415	---	6.32
333	lys	Ready	Finished	---	r0333-lys	72737	11041	---	15.1
334	darkcal	Ready	Finished	---	r0334-darkcal	2124	0	---	0.00
335	lys	Ready	Finished	---	r0335-lys	24722	3110	---	12.5
336	lys	Ready	Finished	---	r0336-lys	44751	6871	---	15.3
337	---	Ready	---	---	---	---	---	---	---
338	darkcal	Ready	Finished	---	r0338-darkcal	2346	0	---	0.00
339	---	Ready	---	---	---	---	---	---	---
340	---	Ready	---	---	---	---	---	---	---
341	---	Ready	---	---	---	---	---	---	---
342	---	Ready	---	---	---	---	---	---	---
343	darkcal	Ready	Finished	---	r0343-darkcal	1296	0	---	0.00

Status of processing is continually updated

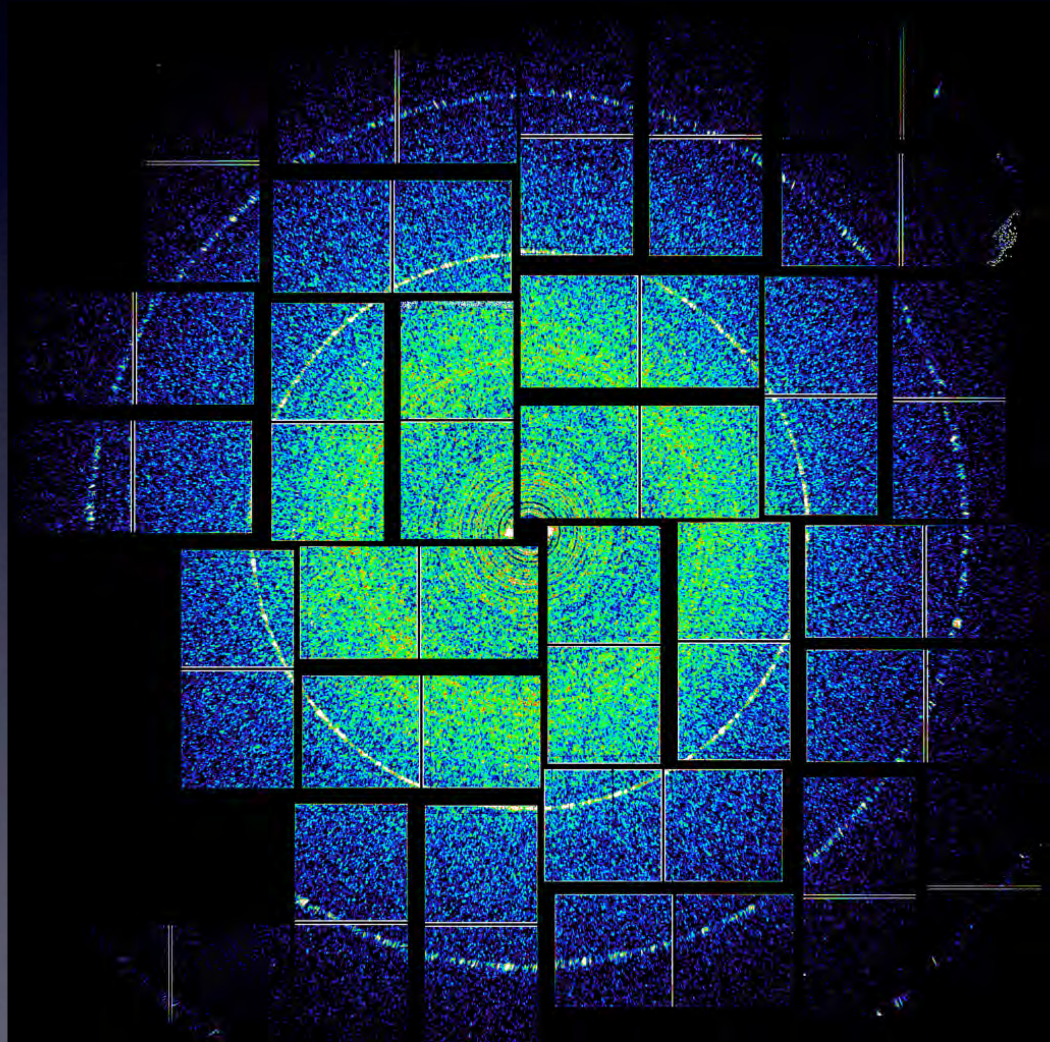
Contents of each run and associated data directory

Cheetah functionality: Run summaries

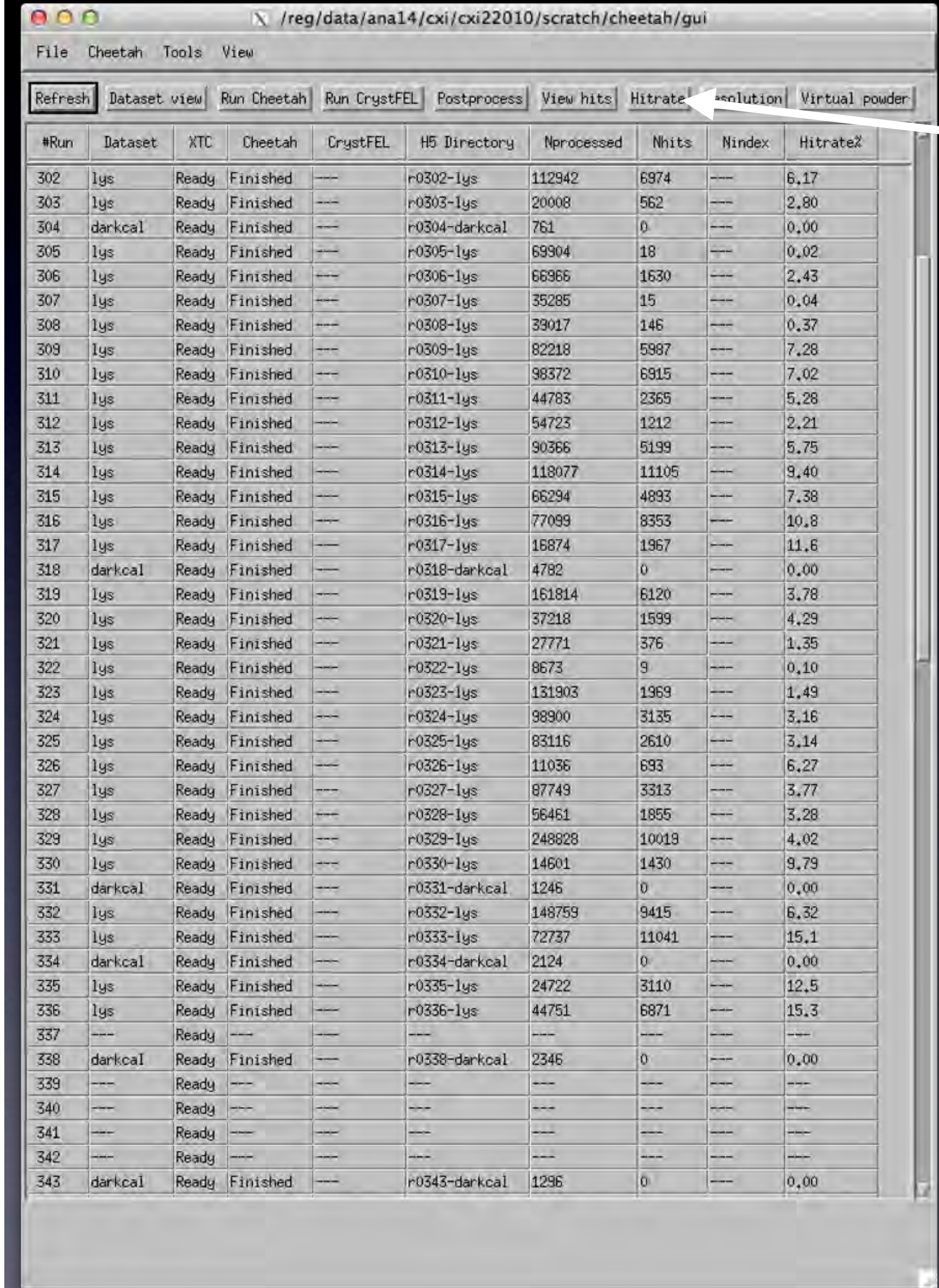


#Run	Dataset	XTC	Cheetah	CrystFEL	H5 Directory	Nprocessed	Nhits	Nindex	Hitrate%
302	lys	Ready	Finished	---	r0302-lys	112942	6974	---	6.17
303	lys	Ready	Finished	---	r0303-lys	20008	562	---	2.80
304	darkcal	Ready	Finished	---	r0304-darkcal	761	0	---	0.00
305	lys	Ready	Finished	---	r0305-lys	69904	18	---	0.02
306	lys	Ready	Finished	---	r0306-lys	66966	1630	---	2.43
307	lys	Ready	Finished	---	r0307-lys	35285	15	---	0.04
308	lys	Ready	Finished	---	r0308-lys	39017	146	---	0.37
309	lys	Ready	Finished	---	r0309-lys	82218	5987	---	7.28
310	lys	Ready	Finished	---	r0310-lys	98372	6915	---	7.02
311	lys	Ready	Finished	---	r0311-lys	44783	2365	---	5.28
312	lys	Ready	Finished	---	r0312-lys	54723	1212	---	2.21
313	lys	Ready	Finished	---	r0313-lys	90366	5199	---	5.75
314	lys	Ready	Finished	---	r0314-lys	118077	11105	---	9.40
315	lys	Ready	Finished	---	r0315-lys	66294	4893	---	7.38
316	lys	Ready	Finished	---	r0316-lys	77099	8353	---	10.8
317	lys	Ready	Finished	---	r0317-lys	16874	1967	---	11.6
318	darkcal	Ready	Finished	---	r0318-darkcal	4782	0	---	0.00
319	lys	Ready	Finished	---	r0319-lys	161814	6120	---	3.78
320	lys	Ready	Finished	---	r0320-lys	37218	1599	---	4.29
321	lys	Ready	Finished	---	r0321-lys	27771	376	---	1.35
322	lys	Ready	Finished	---	r0322-lys	8673	9	---	0.10
323	lys	Ready	Finished	---	r0323-lys	131903	1969	---	1.49
324	lys	Ready	Finished	---	r0324-lys	98900	3135	---	3.16
325	lys	Ready	Finished	---	r0325-lys	83116	2610	---	3.14
326	lys	Ready	Finished	---	r0326-lys	11036	693	---	6.27
327	lys	Ready	Finished	---	r0327-lys	87749	3313	---	3.77
328	lys	Ready	Finished	---	r0328-lys	56461	1855	---	3.28
329	lys	Ready	Finished	---	r0329-lys	248828	10019	---	4.02
330	lys	Ready	Finished	---	r0330-lys	14601	1430	---	9.79
331	darkcal	Ready	Finished	---	r0331-darkcal	1246	0	---	0.00
332	lys	Ready	Finished	---	r0332-lys	148759	9415	---	6.32
333	lys	Ready	Finished	---	r0333-lys	72737	11041	---	15.1
334	darkcal	Ready	Finished	---	r0334-darkcal	2124	0	---	0.00
335	lys	Ready	Finished	---	r0335-lys	24722	3110	---	12.5
336	lys	Ready	Finished	---	r0336-lys	44751	6871	---	15.3
337	---	Ready	---	---	---	---	---	---	---
338	darkcal	Ready	Finished	---	r0338-darkcal	2346	0	---	0.00
339	---	Ready	---	---	---	---	---	---	---
340	---	Ready	---	---	---	---	---	---	---
341	---	Ready	---	---	---	---	---	---	---
342	---	Ready	---	---	---	---	---	---	---
343	darkcal	Ready	Finished	---	r0343-darkcal	1296	0	---	0.00

Virtual powder pattern

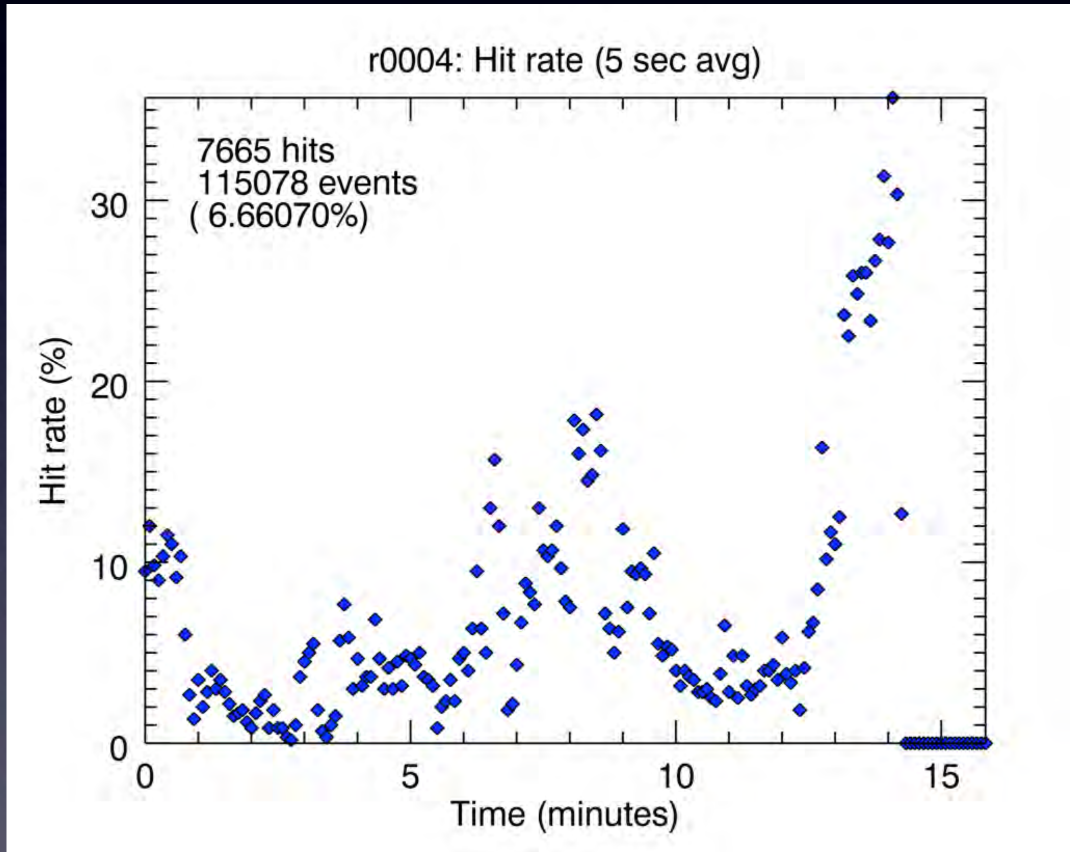


Cheetah functionality: Hit rates

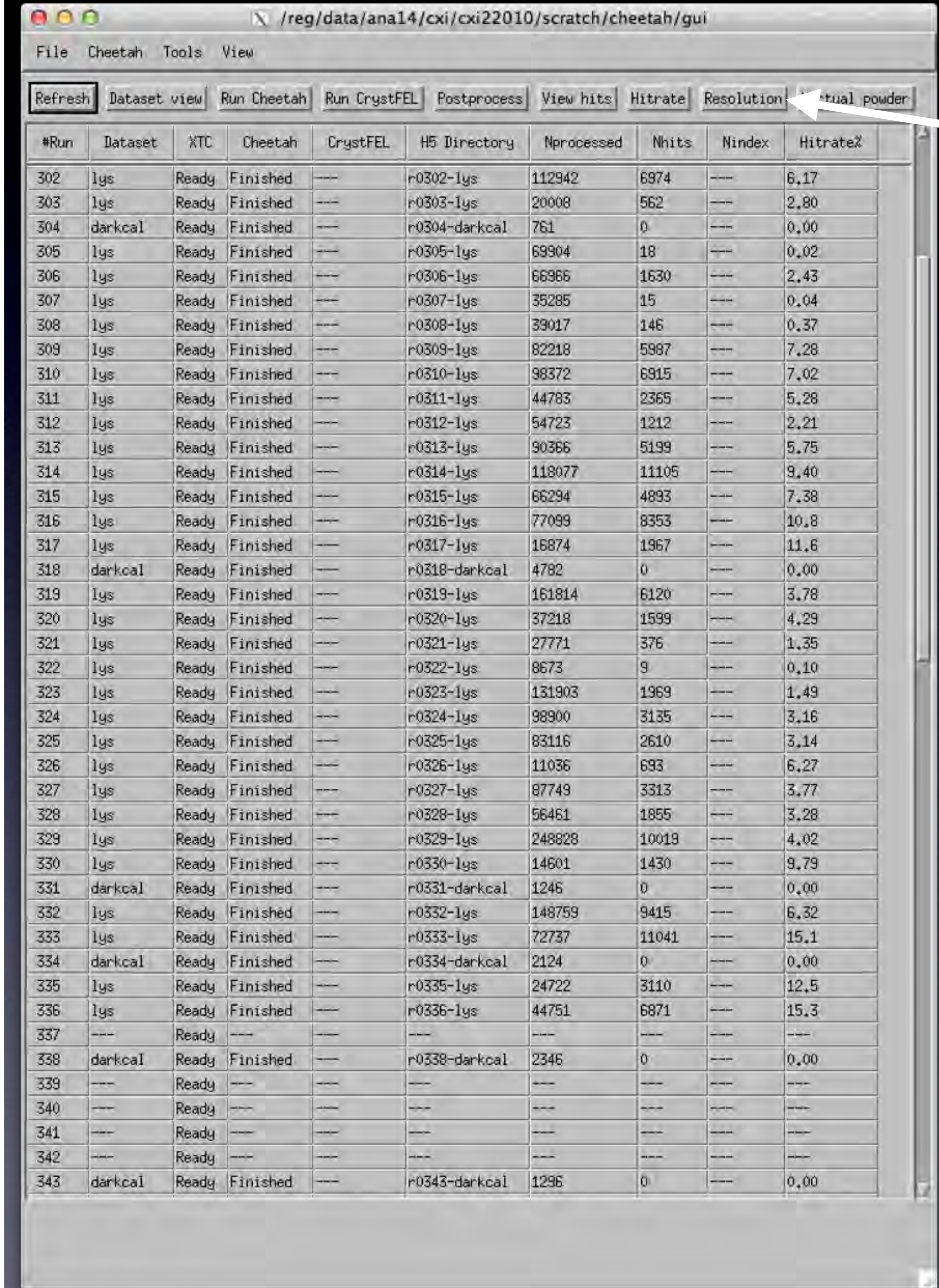


#Run	Dataset	XTC	Cheetah	CrystFEL	H5 Directory	Nprocessed	Nhits	Nindex	Hitrate%
302	lys	Ready	Finished	---	r0302-lys	112942	6974	---	6.17
303	lys	Ready	Finished	---	r0303-lys	20008	562	---	2.80
304	darkcal	Ready	Finished	---	r0304-darkcal	761	0	---	0.00
305	lys	Ready	Finished	---	r0305-lys	69904	18	---	0.02
306	lys	Ready	Finished	---	r0306-lys	66966	1630	---	2.43
307	lys	Ready	Finished	---	r0307-lys	35285	15	---	0.04
308	lys	Ready	Finished	---	r0308-lys	39017	146	---	0.37
309	lys	Ready	Finished	---	r0309-lys	82218	5987	---	7.28
310	lys	Ready	Finished	---	r0310-lys	98372	6915	---	7.02
311	lys	Ready	Finished	---	r0311-lys	44783	2365	---	5.28
312	lys	Ready	Finished	---	r0312-lys	54723	1212	---	2.21
313	lys	Ready	Finished	---	r0313-lys	90366	5199	---	5.75
314	lys	Ready	Finished	---	r0314-lys	118077	11105	---	9.40
315	lys	Ready	Finished	---	r0315-lys	66294	4893	---	7.38
316	lys	Ready	Finished	---	r0316-lys	77099	8353	---	10.8
317	lys	Ready	Finished	---	r0317-lys	16874	1967	---	11.6
318	darkcal	Ready	Finished	---	r0318-darkcal	4782	0	---	0.00
319	lys	Ready	Finished	---	r0319-lys	161814	6120	---	3.78
320	lys	Ready	Finished	---	r0320-lys	37218	1599	---	4.29
321	lys	Ready	Finished	---	r0321-lys	27771	376	---	1.35
322	lys	Ready	Finished	---	r0322-lys	8673	9	---	0.10
323	lys	Ready	Finished	---	r0323-lys	131903	1969	---	1.49
324	lys	Ready	Finished	---	r0324-lys	98900	3135	---	3.16
325	lys	Ready	Finished	---	r0325-lys	83116	2610	---	3.14
326	lys	Ready	Finished	---	r0326-lys	11036	693	---	6.27
327	lys	Ready	Finished	---	r0327-lys	87749	3313	---	3.77
328	lys	Ready	Finished	---	r0328-lys	56461	1855	---	3.28
329	lys	Ready	Finished	---	r0329-lys	248828	10019	---	4.02
330	lys	Ready	Finished	---	r0330-lys	14601	1430	---	9.79
331	darkcal	Ready	Finished	---	r0331-darkcal	1246	0	---	0.00
332	lys	Ready	Finished	---	r0332-lys	148759	9415	---	6.32
333	lys	Ready	Finished	---	r0333-lys	72737	11041	---	15.1
334	darkcal	Ready	Finished	---	r0334-darkcal	2124	0	---	0.00
335	lys	Ready	Finished	---	r0335-lys	24722	3110	---	12.5
336	lys	Ready	Finished	---	r0336-lys	44751	6871	---	15.3
337	---	Ready	---	---	---	---	---	---	---
338	darkcal	Ready	Finished	---	r0338-darkcal	2346	0	---	0.00
339	---	Ready	---	---	---	---	---	---	---
340	---	Ready	---	---	---	---	---	---	---
341	---	Ready	---	---	---	---	---	---	---
342	---	Ready	---	---	---	---	---	---	---
343	darkcal	Ready	Finished	---	r0343-darkcal	1296	0	---	0.00

Hit rate

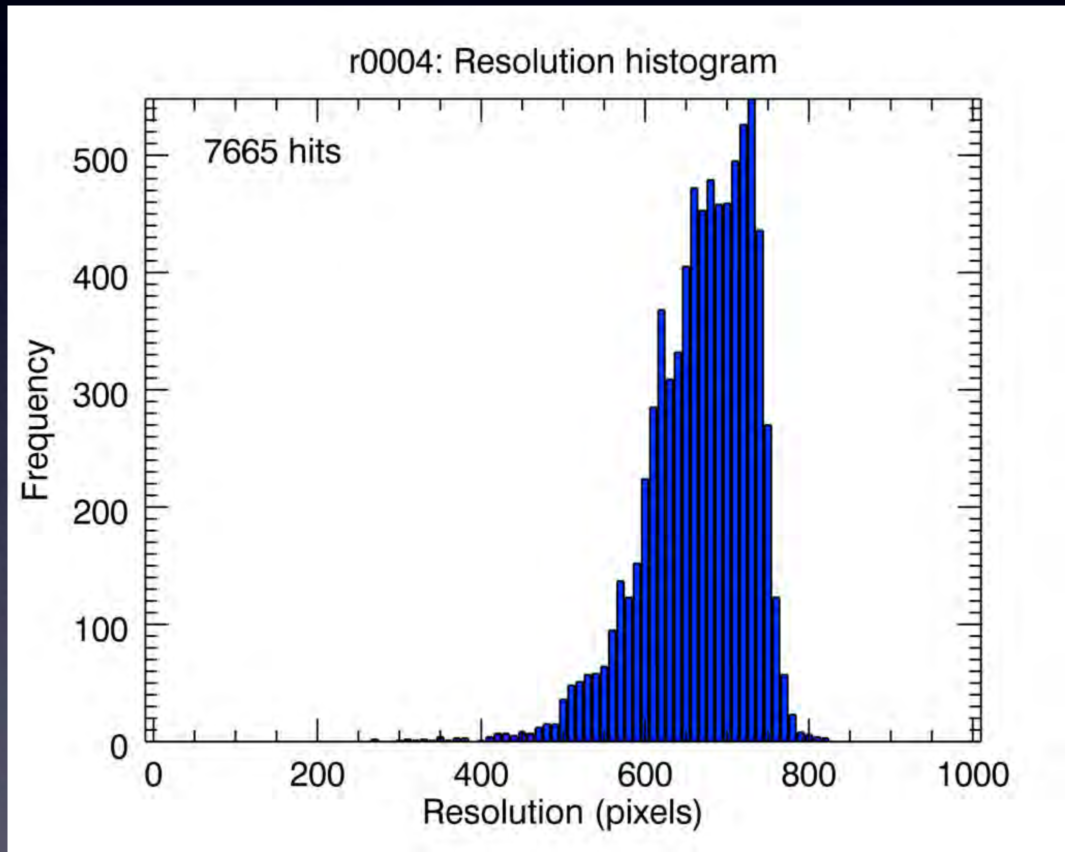


Cheetah functionality: Resolution



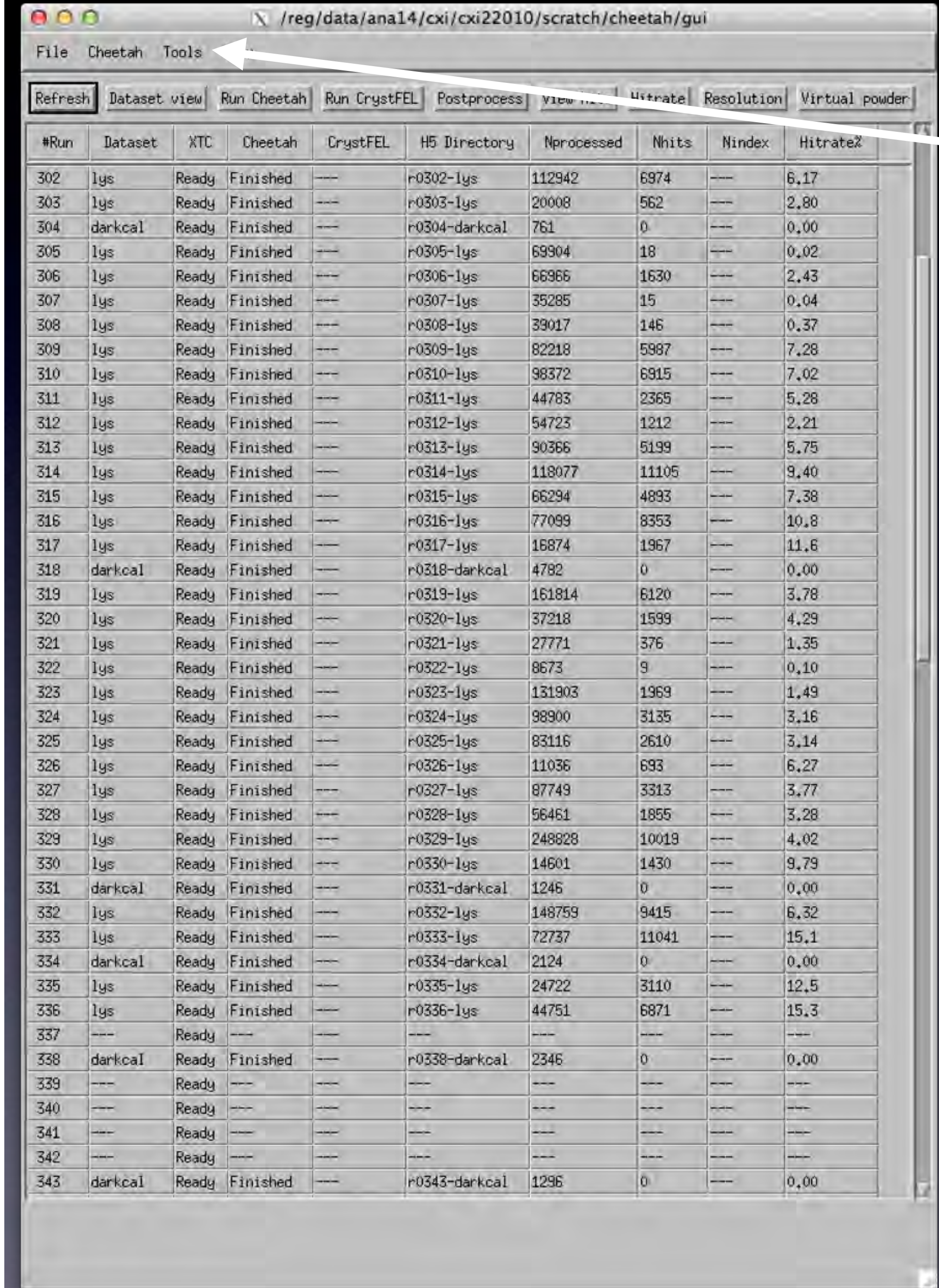
#Run	Dataset	XTC	Cheetah	CrystFEL	H5 Directory	Nprocessed	Nhits	Nindex	Hitrate%
302	lys	Ready	Finished	---	r0302-lys	112942	6974	---	6.17
303	lys	Ready	Finished	---	r0303-lys	20008	562	---	2.80
304	darkcal	Ready	Finished	---	r0304-darkcal	761	0	---	0.00
305	lys	Ready	Finished	---	r0305-lys	69904	18	---	0.02
306	lys	Ready	Finished	---	r0306-lys	66966	1630	---	2.43
307	lys	Ready	Finished	---	r0307-lys	35285	15	---	0.04
308	lys	Ready	Finished	---	r0308-lys	39017	146	---	0.37
309	lys	Ready	Finished	---	r0309-lys	82218	5987	---	7.28
310	lys	Ready	Finished	---	r0310-lys	98372	6915	---	7.02
311	lys	Ready	Finished	---	r0311-lys	44783	2365	---	5.28
312	lys	Ready	Finished	---	r0312-lys	54723	1212	---	2.21
313	lys	Ready	Finished	---	r0313-lys	90366	5199	---	5.75
314	lys	Ready	Finished	---	r0314-lys	118077	11105	---	9.40
315	lys	Ready	Finished	---	r0315-lys	66294	4893	---	7.38
316	lys	Ready	Finished	---	r0316-lys	77099	8353	---	10.8
317	lys	Ready	Finished	---	r0317-lys	16874	1967	---	11.6
318	darkcal	Ready	Finished	---	r0318-darkcal	4782	0	---	0.00
319	lys	Ready	Finished	---	r0319-lys	161814	6120	---	3.78
320	lys	Ready	Finished	---	r0320-lys	37218	1599	---	4.29
321	lys	Ready	Finished	---	r0321-lys	27771	376	---	1.35
322	lys	Ready	Finished	---	r0322-lys	8673	9	---	0.10
323	lys	Ready	Finished	---	r0323-lys	131903	1969	---	1.49
324	lys	Ready	Finished	---	r0324-lys	98900	3135	---	3.16
325	lys	Ready	Finished	---	r0325-lys	83116	2610	---	3.14
326	lys	Ready	Finished	---	r0326-lys	11036	693	---	6.27
327	lys	Ready	Finished	---	r0327-lys	87749	3313	---	3.77
328	lys	Ready	Finished	---	r0328-lys	56461	1855	---	3.28
329	lys	Ready	Finished	---	r0329-lys	248828	10019	---	4.02
330	lys	Ready	Finished	---	r0330-lys	14601	1430	---	9.79
331	darkcal	Ready	Finished	---	r0331-darkcal	1246	0	---	0.00
332	lys	Ready	Finished	---	r0332-lys	148759	9415	---	6.32
333	lys	Ready	Finished	---	r0333-lys	72737	11041	---	15.1
334	darkcal	Ready	Finished	---	r0334-darkcal	2124	0	---	0.00
335	lys	Ready	Finished	---	r0335-lys	24722	3110	---	12.5
336	lys	Ready	Finished	---	r0336-lys	44751	6871	---	15.3
337	---	Ready	---	---	---	---	---	---	---
338	darkcal	Ready	Finished	---	r0338-darkcal	2346	0	---	0.00
339	---	Ready	---	---	---	---	---	---	---
340	---	Ready	---	---	---	---	---	---	---
341	---	Ready	---	---	---	---	---	---	---
342	---	Ready	---	---	---	---	---	---	---
343	darkcal	Ready	Finished	---	r0343-darkcal	1296	0	---	0.00

Resolution



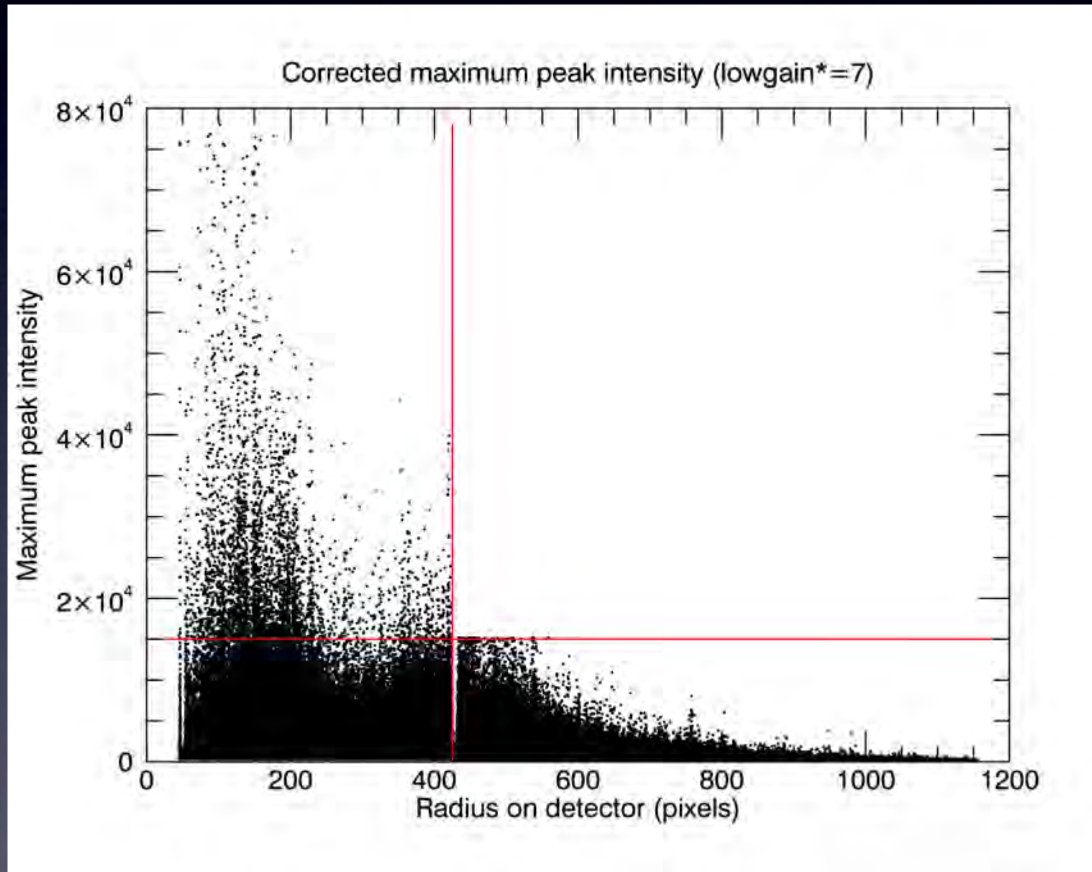
Circle containing 80% of found peaks

Cheetah functionality: Detector saturation check



#Run	Dataset	XTC	Cheetah	CrystFEL	H5 Directory	Nprocessed	Nhits	Nindex	Hitrate%
302	lys	Ready	Finished	---	r0302-lys	112942	6974	---	6.17
303	lys	Ready	Finished	---	r0303-lys	20008	562	---	2.80
304	darkcal	Ready	Finished	---	r0304-darkcal	761	0	---	0.00
305	lys	Ready	Finished	---	r0305-lys	69904	18	---	0.02
306	lys	Ready	Finished	---	r0306-lys	66966	1630	---	2.43
307	lys	Ready	Finished	---	r0307-lys	35285	15	---	0.04
308	lys	Ready	Finished	---	r0308-lys	39017	146	---	0.37
309	lys	Ready	Finished	---	r0309-lys	82218	5987	---	7.28
310	lys	Ready	Finished	---	r0310-lys	98372	6915	---	7.02
311	lys	Ready	Finished	---	r0311-lys	44783	2365	---	5.28
312	lys	Ready	Finished	---	r0312-lys	54723	1212	---	2.21
313	lys	Ready	Finished	---	r0313-lys	90366	5199	---	5.75
314	lys	Ready	Finished	---	r0314-lys	118077	11105	---	9.40
315	lys	Ready	Finished	---	r0315-lys	66294	4893	---	7.38
316	lys	Ready	Finished	---	r0316-lys	77099	8353	---	10.8
317	lys	Ready	Finished	---	r0317-lys	16874	1967	---	11.6
318	darkcal	Ready	Finished	---	r0318-darkcal	4782	0	---	0.00
319	lys	Ready	Finished	---	r0319-lys	161814	6120	---	3.78
320	lys	Ready	Finished	---	r0320-lys	37218	1599	---	4.29
321	lys	Ready	Finished	---	r0321-lys	27771	376	---	1.35
322	lys	Ready	Finished	---	r0322-lys	8673	9	---	0.10
323	lys	Ready	Finished	---	r0323-lys	131903	1969	---	1.49
324	lys	Ready	Finished	---	r0324-lys	98900	3135	---	3.16
325	lys	Ready	Finished	---	r0325-lys	83116	2610	---	3.14
326	lys	Ready	Finished	---	r0326-lys	11036	693	---	6.27
327	lys	Ready	Finished	---	r0327-lys	87749	3313	---	3.77
328	lys	Ready	Finished	---	r0328-lys	56461	1855	---	3.28
329	lys	Ready	Finished	---	r0329-lys	248828	10019	---	4.02
330	lys	Ready	Finished	---	r0330-lys	14601	1430	---	9.79
331	darkcal	Ready	Finished	---	r0331-darkcal	1246	0	---	0.00
332	lys	Ready	Finished	---	r0332-lys	148759	9415	---	6.32
333	lys	Ready	Finished	---	r0333-lys	72737	11041	---	15.1
334	darkcal	Ready	Finished	---	r0334-darkcal	2124	0	---	0.00
335	lys	Ready	Finished	---	r0335-lys	24722	3110	---	12.5
336	lys	Ready	Finished	---	r0336-lys	44751	6871	---	15.3
337	---	Ready	---	---	---	---	---	---	---
338	darkcal	Ready	Finished	---	r0338-darkcal	2346	0	---	0.00
339	---	Ready	---	---	---	---	---	---	---
340	---	Ready	---	---	---	---	---	---	---
341	---	Ready	---	---	---	---	---	---	---
342	---	Ready	---	---	---	---	---	---	---
343	darkcal	Ready	Finished	---	r0343-darkcal	1296	0	---	0.00

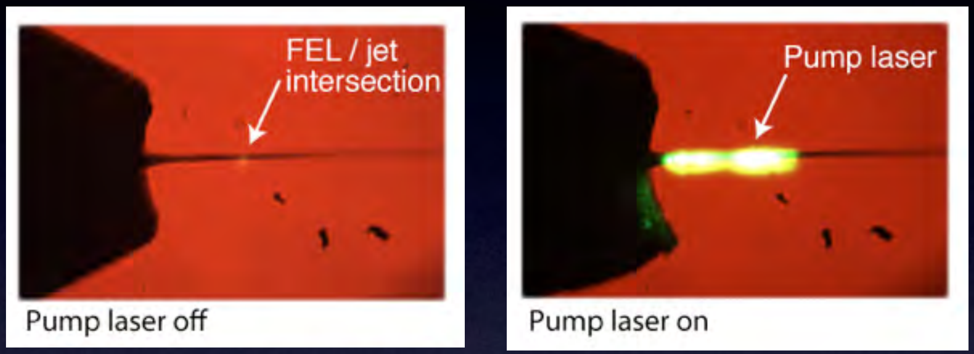
Saturation check



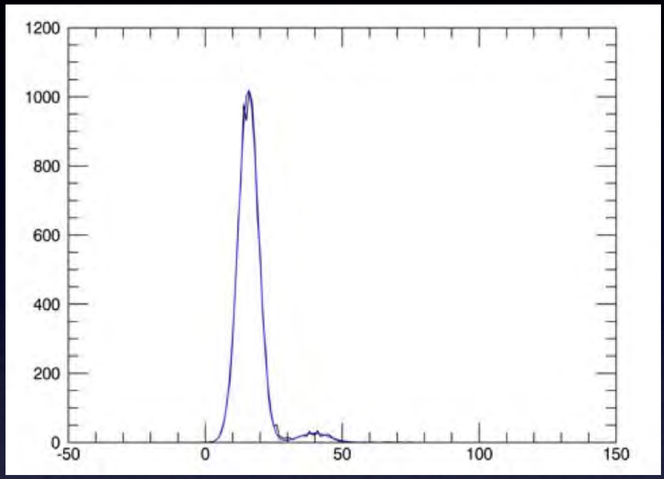
+ support for dual gain mode

Other useful outputs

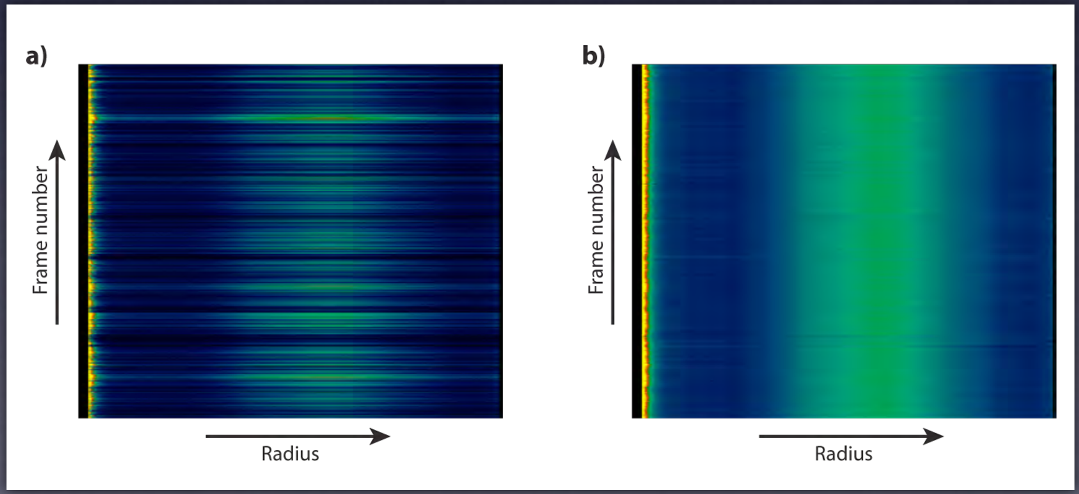
Pump/probe sorting



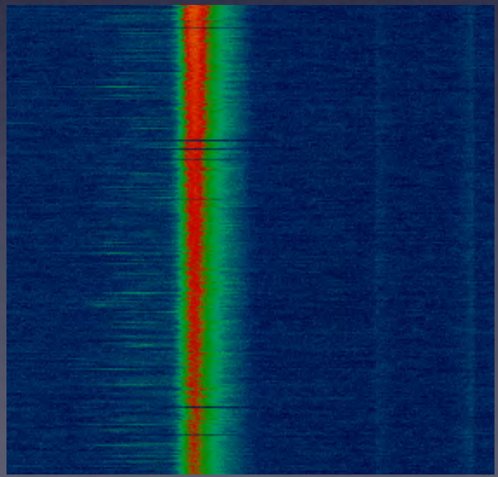
Full detector histogram



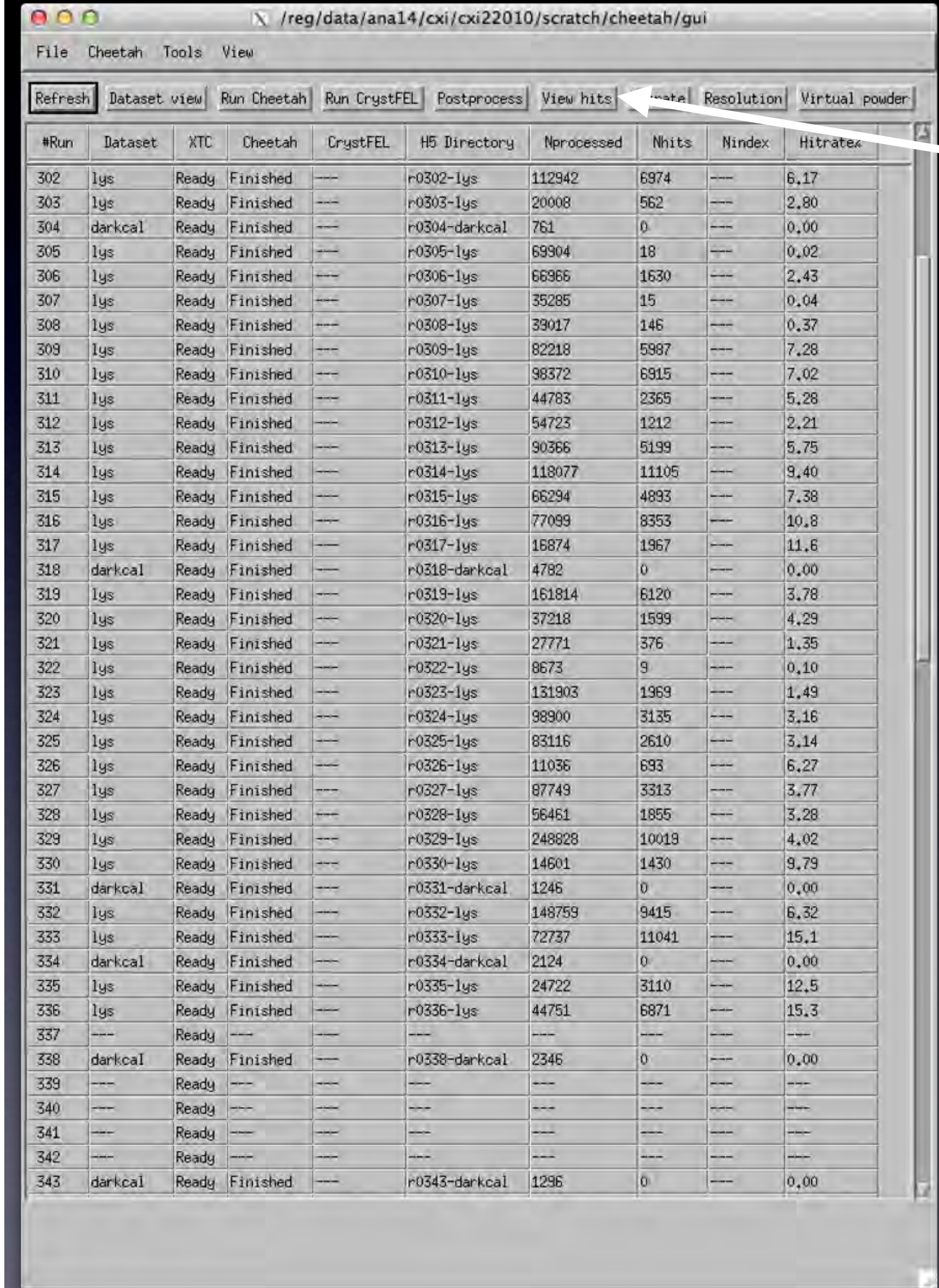
Radial stacks



Spectral stack

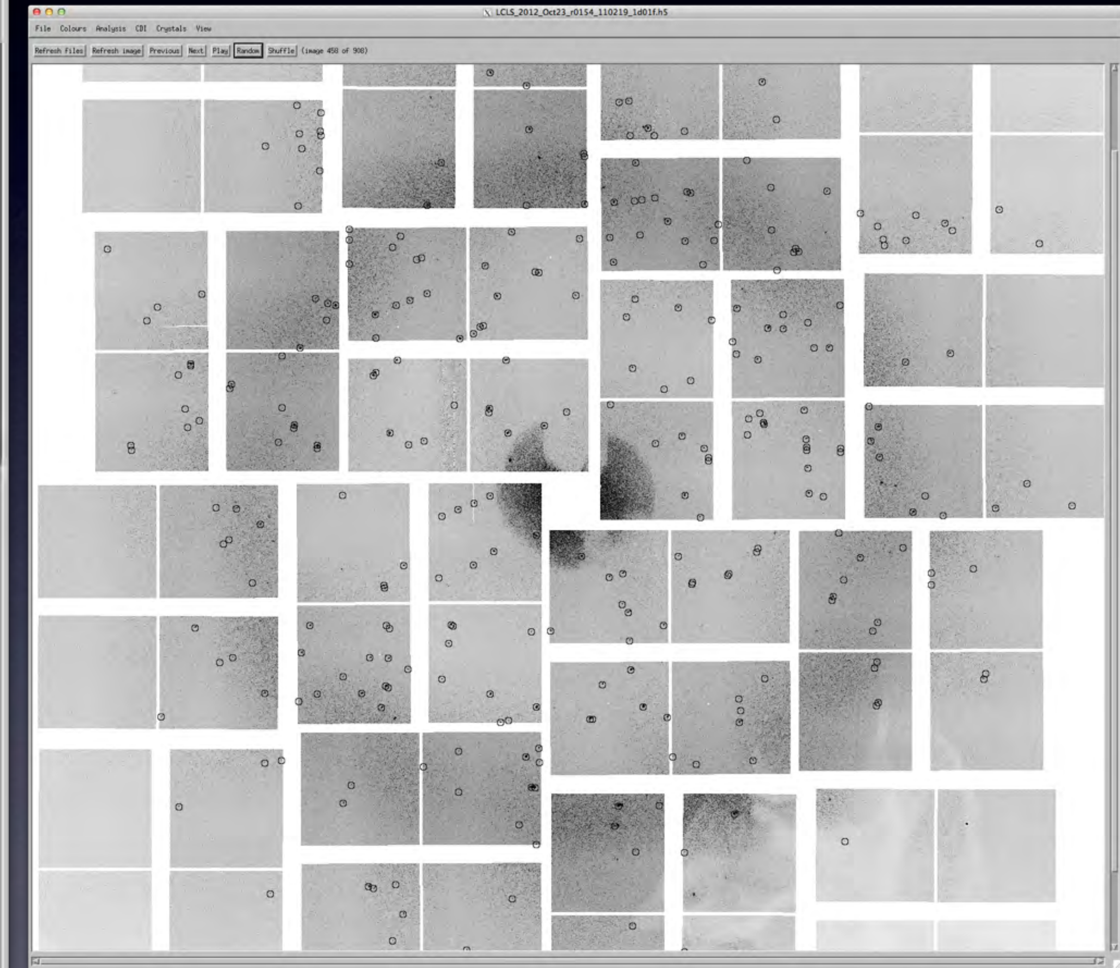


Cheetah functionality: Data inspection

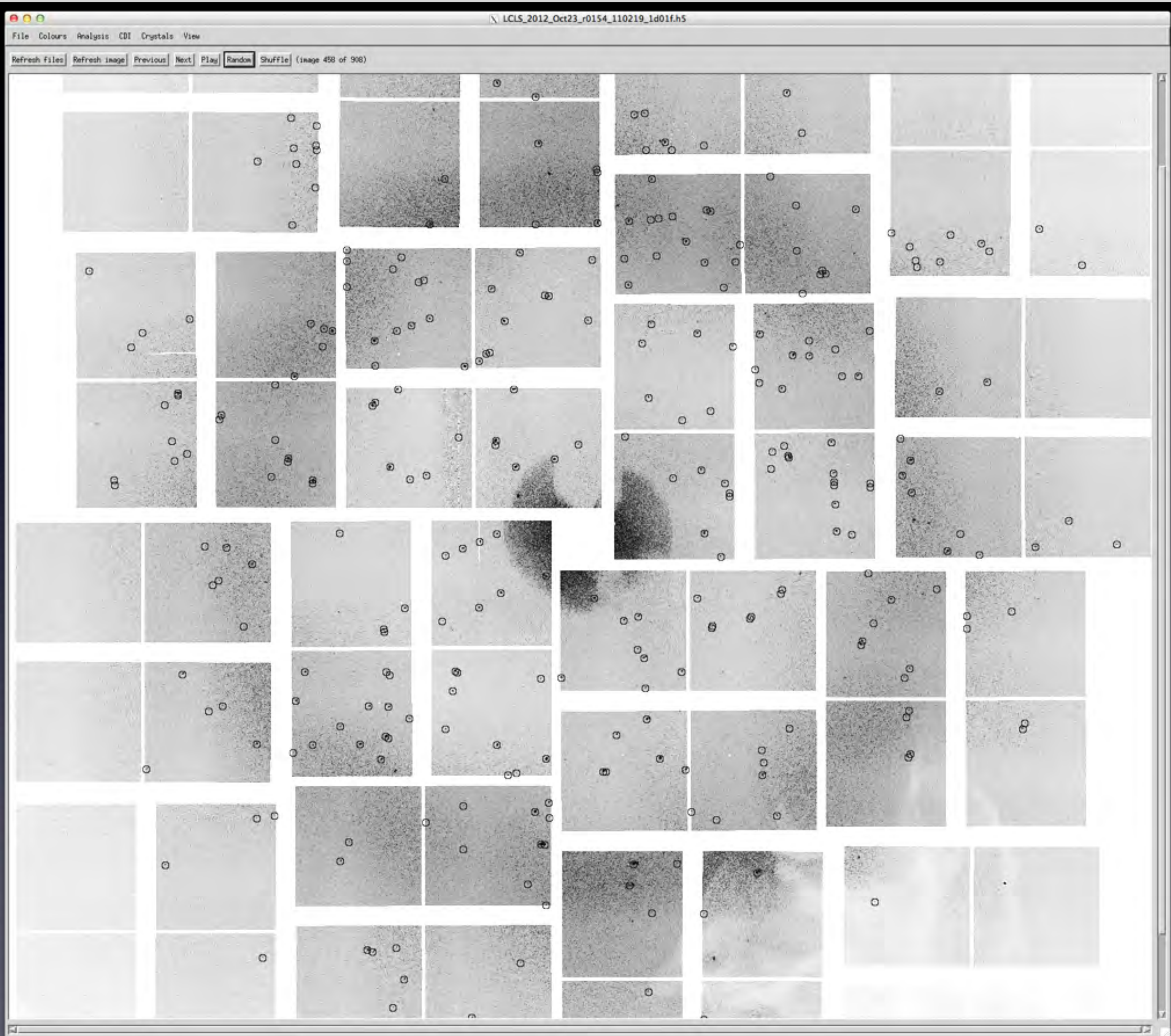


#Run	Dataset	XTC	Cheetah	CrystFEL	H5 Directory	Nprocessed	Nhits	Nindex	Hitrate%
302	lys	Ready	Finished	---	r0302-lys	112942	6974	---	6.17
303	lys	Ready	Finished	---	r0303-lys	20008	562	---	2.80
304	darkcal	Ready	Finished	---	r0304-darkcal	761	0	---	0.00
305	lys	Ready	Finished	---	r0305-lys	69904	18	---	0.02
306	lys	Ready	Finished	---	r0306-lys	66966	1630	---	2.43
307	lys	Ready	Finished	---	r0307-lys	35285	15	---	0.04
308	lys	Ready	Finished	---	r0308-lys	39017	146	---	0.37
309	lys	Ready	Finished	---	r0309-lys	82218	5987	---	7.28
310	lys	Ready	Finished	---	r0310-lys	98372	6915	---	7.02
311	lys	Ready	Finished	---	r0311-lys	44783	2365	---	5.28
312	lys	Ready	Finished	---	r0312-lys	54723	1212	---	2.21
313	lys	Ready	Finished	---	r0313-lys	90366	5199	---	5.75
314	lys	Ready	Finished	---	r0314-lys	118077	11105	---	9.40
315	lys	Ready	Finished	---	r0315-lys	66294	4893	---	7.38
316	lys	Ready	Finished	---	r0316-lys	77099	8353	---	10.8
317	lys	Ready	Finished	---	r0317-lys	16874	1967	---	11.6
318	darkcal	Ready	Finished	---	r0318-darkcal	4782	0	---	0.00
319	lys	Ready	Finished	---	r0319-lys	161814	6120	---	3.78
320	lys	Ready	Finished	---	r0320-lys	37218	1599	---	4.29
321	lys	Ready	Finished	---	r0321-lys	27771	376	---	1.35
322	lys	Ready	Finished	---	r0322-lys	8673	9	---	0.10
323	lys	Ready	Finished	---	r0323-lys	131903	1969	---	1.49
324	lys	Ready	Finished	---	r0324-lys	98900	3135	---	3.16
325	lys	Ready	Finished	---	r0325-lys	83116	2610	---	3.14
326	lys	Ready	Finished	---	r0326-lys	11036	693	---	6.27
327	lys	Ready	Finished	---	r0327-lys	87749	3313	---	3.77
328	lys	Ready	Finished	---	r0328-lys	56461	1855	---	3.28
329	lys	Ready	Finished	---	r0329-lys	248828	10019	---	4.02
330	lys	Ready	Finished	---	r0330-lys	14601	1430	---	9.79
331	darkcal	Ready	Finished	---	r0331-darkcal	1246	0	---	0.00
332	lys	Ready	Finished	---	r0332-lys	148759	9415	---	6.32
333	lys	Ready	Finished	---	r0333-lys	72737	11041	---	15.1
334	darkcal	Ready	Finished	---	r0334-darkcal	2124	0	---	0.00
335	lys	Ready	Finished	---	r0335-lys	24722	3110	---	12.5
336	lys	Ready	Finished	---	r0336-lys	44751	6871	---	15.3
337	---	Ready	---	---	---	---	---	---	---
338	darkcal	Ready	Finished	---	r0338-darkcal	2346	0	---	0.00
339	---	Ready	---	---	---	---	---	---	---
340	---	Ready	---	---	---	---	---	---	---
341	---	Ready	---	---	---	---	---	---	---
342	---	Ready	---	---	---	---	---	---	---
343	darkcal	Ready	Finished	---	r0343-darkcal	1296	0	---	0.00

Individual diffraction patterns

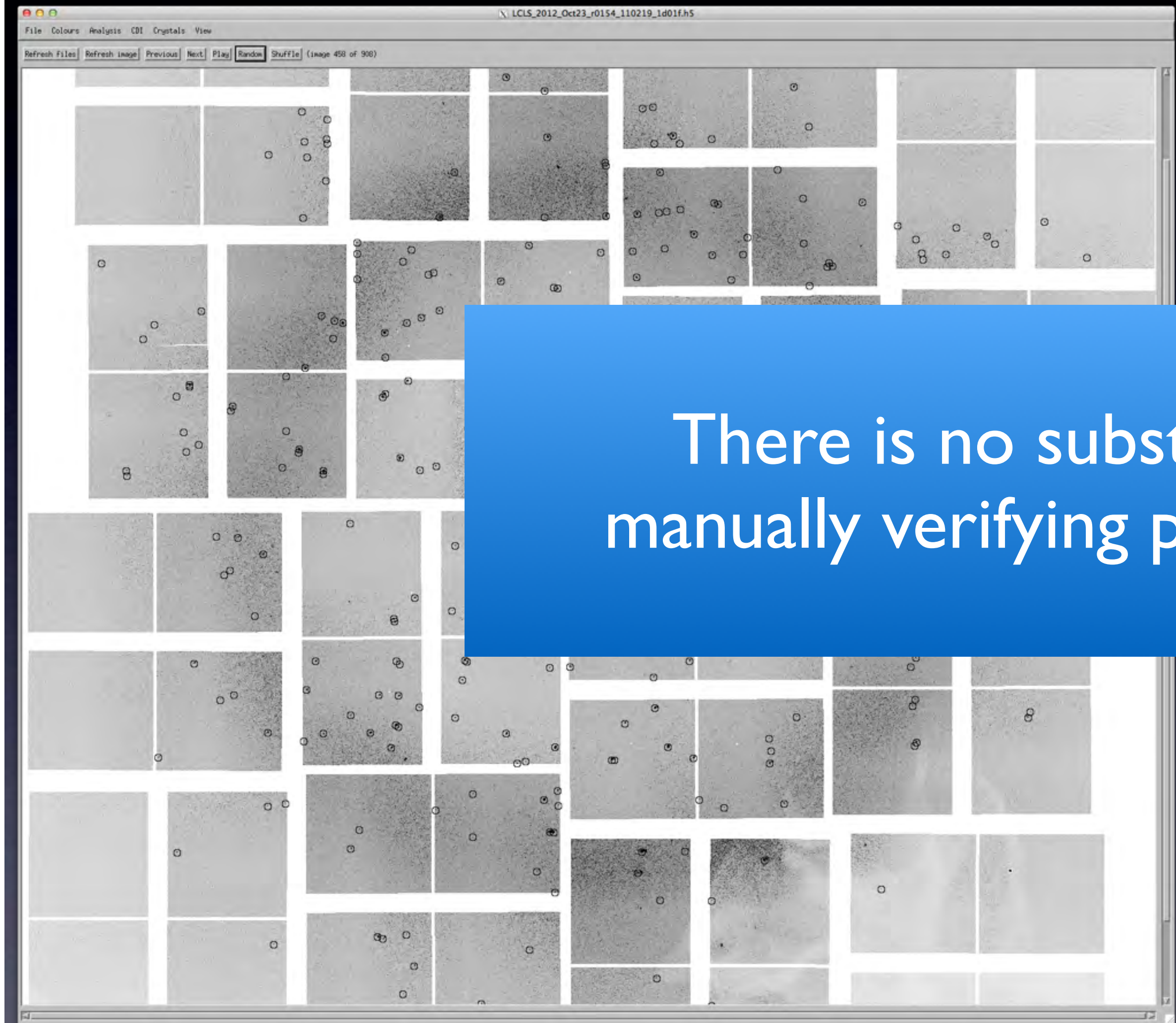


Cheetah functionality: Data viewer for checking peak finding



CFEL SCIENCE

It is all boils down to quick but accurate (enough) peak finding



"Hitfinder 8"

1. Calculate radial SNR and offset

2. thresh(r) = offset + s*sigma
   nADC

   an npix connected
   esh(r)

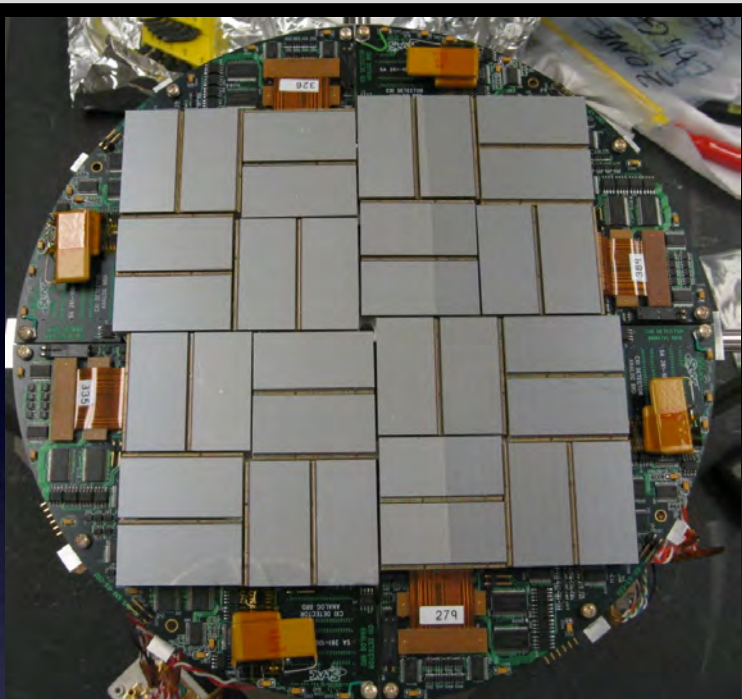
4. Require more than n peaks

There is no substitute for manually verifying peak finding

Critical inputs:
- SNR
- npix
- minADC
- rmin, rmax

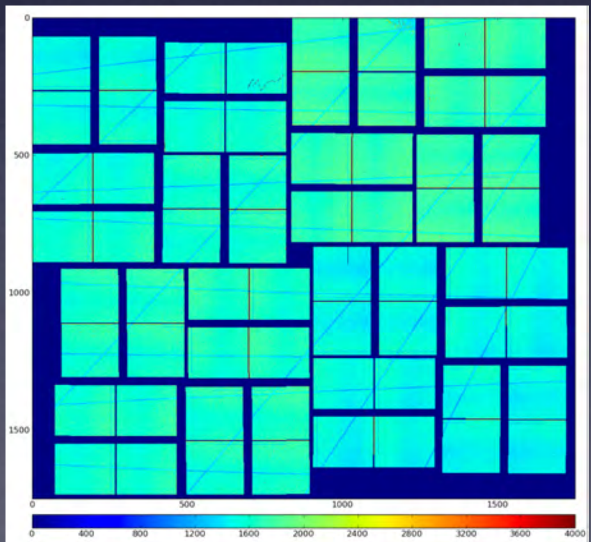
CFEL SCIENCE

The full cspad detector consists of many tiles of smaller detectors

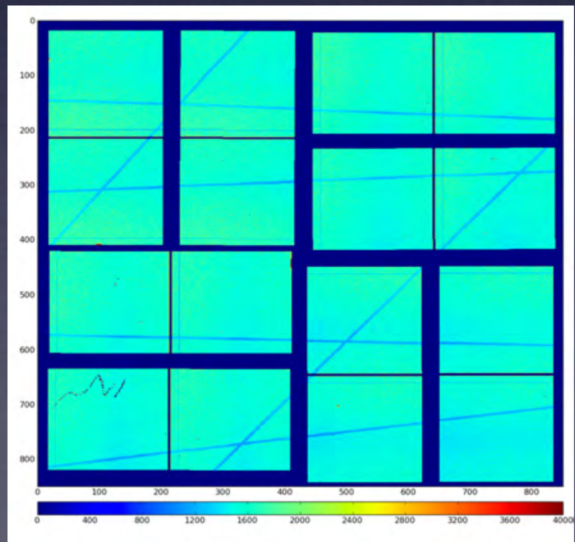


- Each ASIC is a separate detector
- Each quadrant can move independently

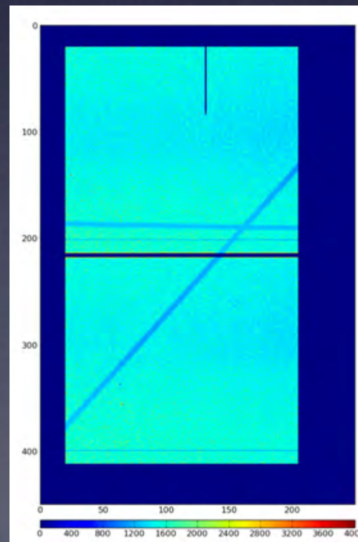
Detector geometry is very important
(and not a trivial problem)



Full detector

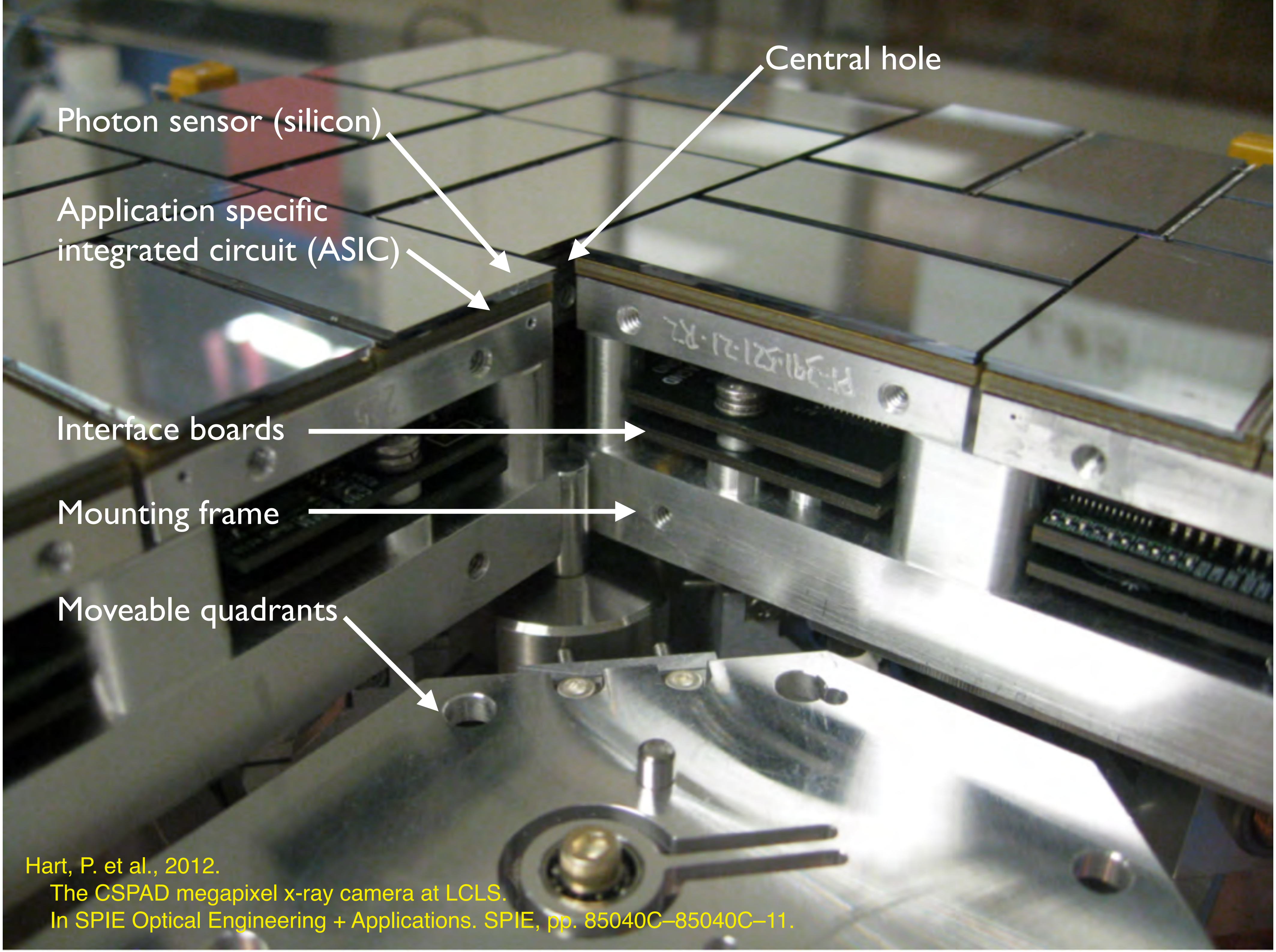


Quadrant



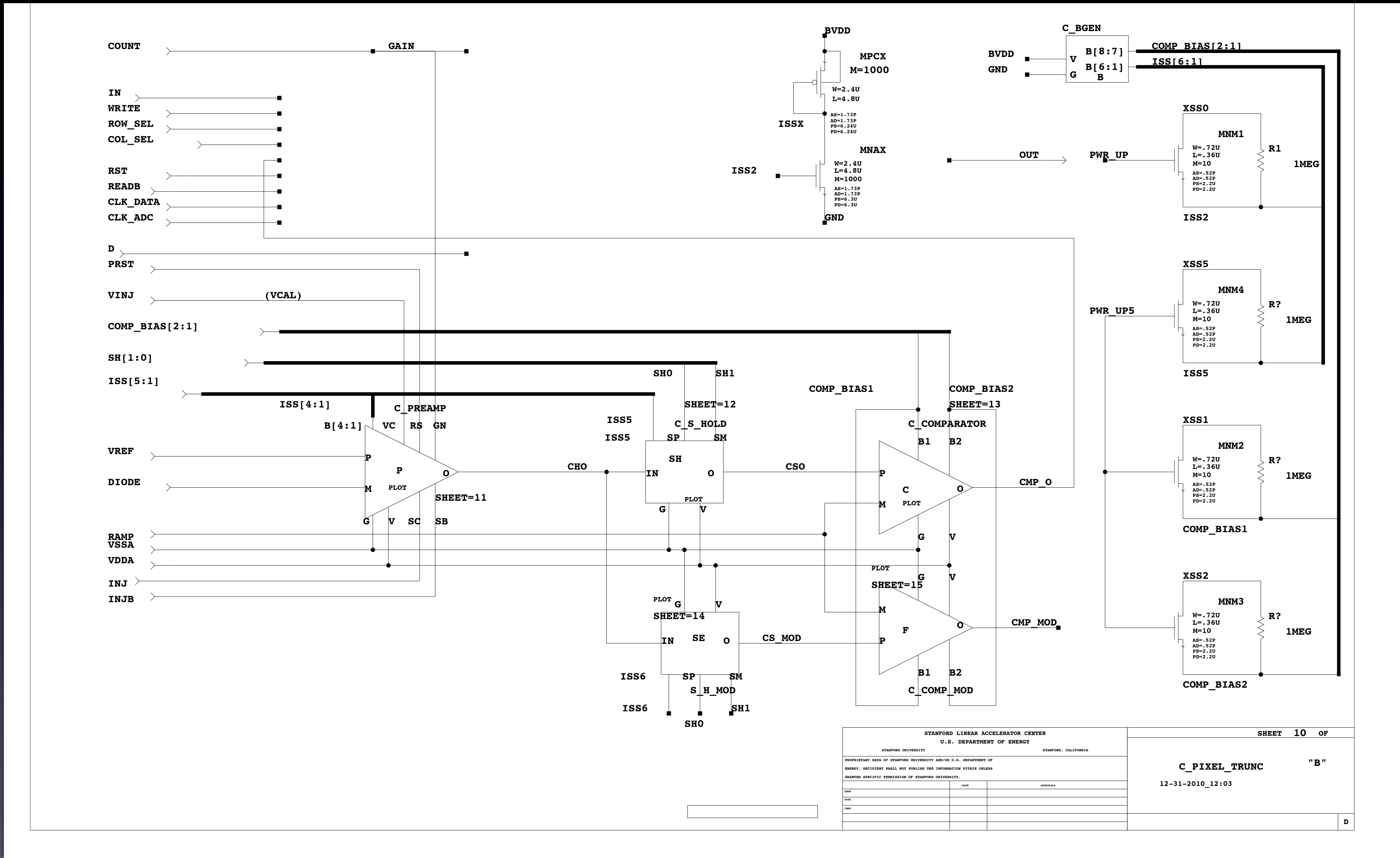
2x1

CFEL SCIENCE

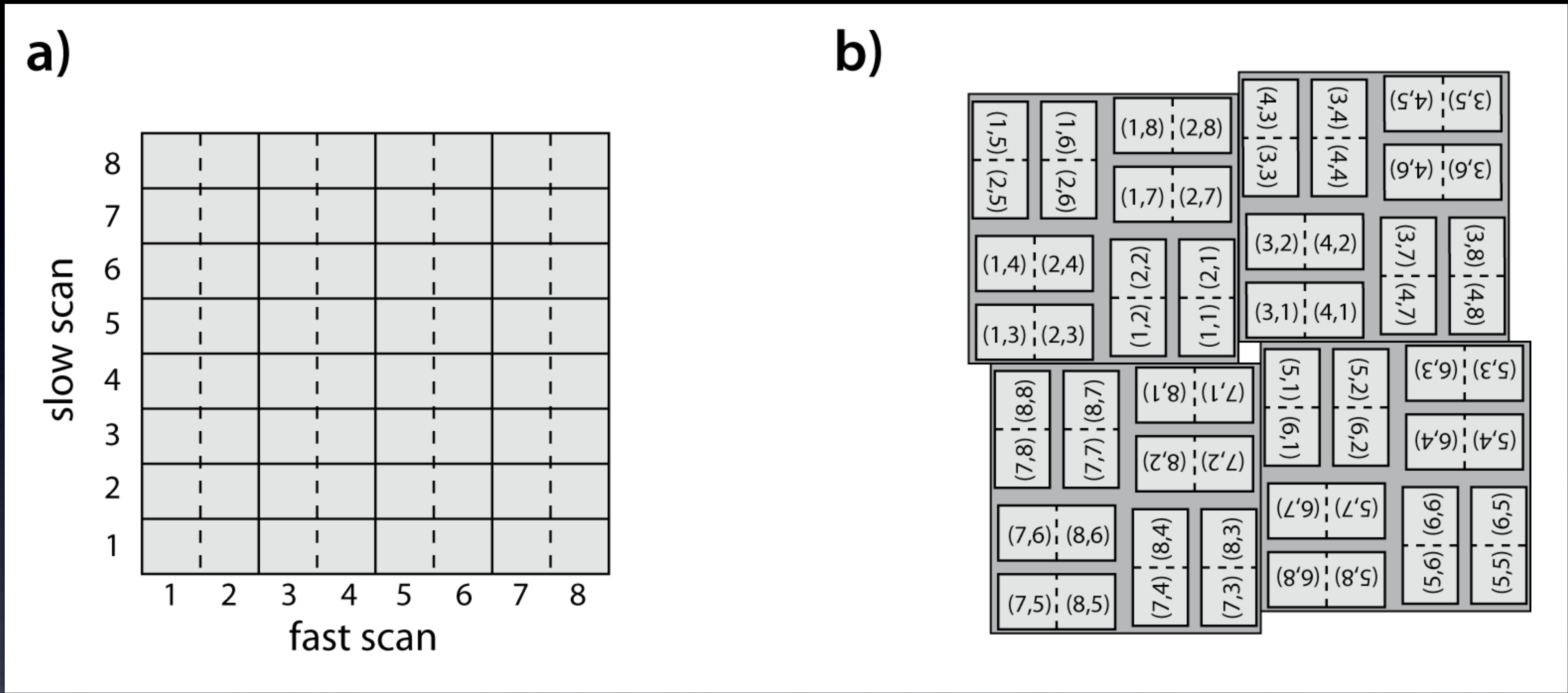


Hart, P. et al., 2012.
The CSPAD megapixel x-ray camera at LCLS.
In SPIE Optical Engineering + Applications. SPIE, pp. 85040C–85040C–11.

The insides of a cspad detector



There is no need to assemble one geometrically correct image; each module is a collection of pixels placed somewhere in space



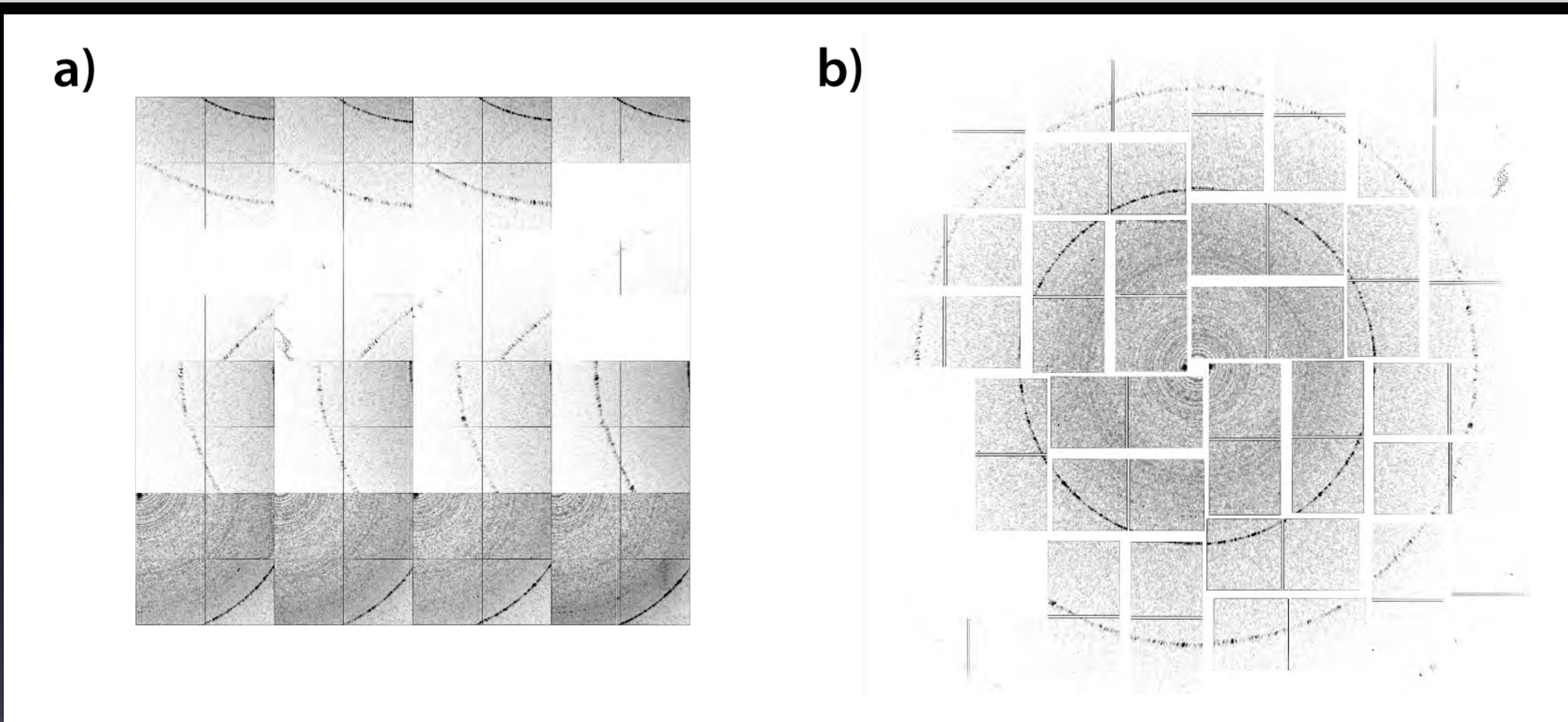
Data layout

Physical layout

Geometry is specified in a pixel map:
HDF5 file with (x,y,z) coordinate of each pixel in experiment space
as seen when looking downstream (ie: looking at image projected onto front of detector)

CFEL SCIENCE

There is no need to assemble one geometrically correct image; each module is a collection of pixels placed somewhere in space

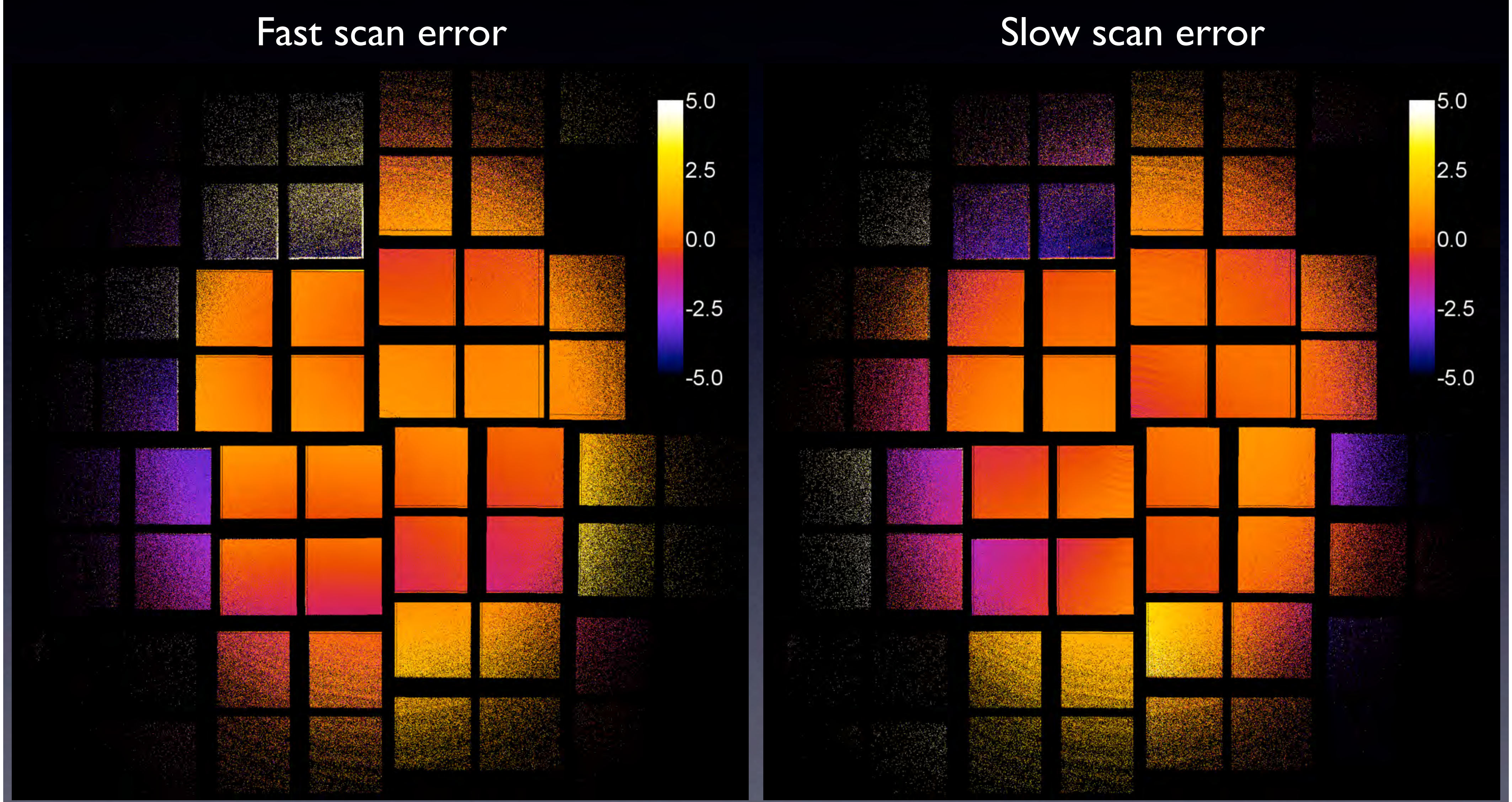


Data layout in data file

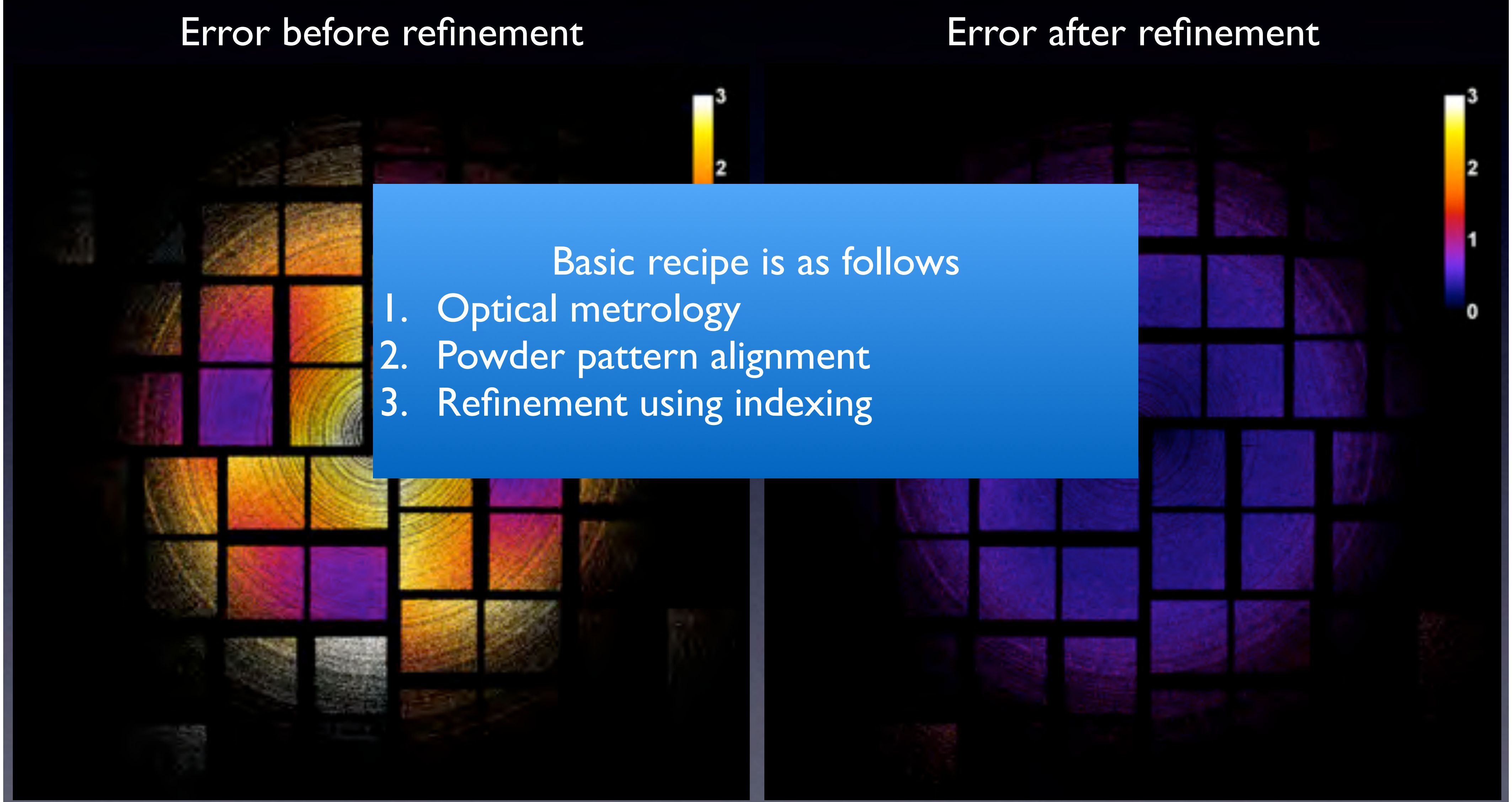
Physical layout

Geometry is specified in a pixel map:
HDF5 file with (x,y,z) coordinate of each pixel in experiment space
as seen when looking downstream (ie: looking at image projected onto front of detector)

Detector geometry is accurately determined by comparing observed and predicted peak locations



Detector geometry is accurately determined by comparing observed and predicted peak locations

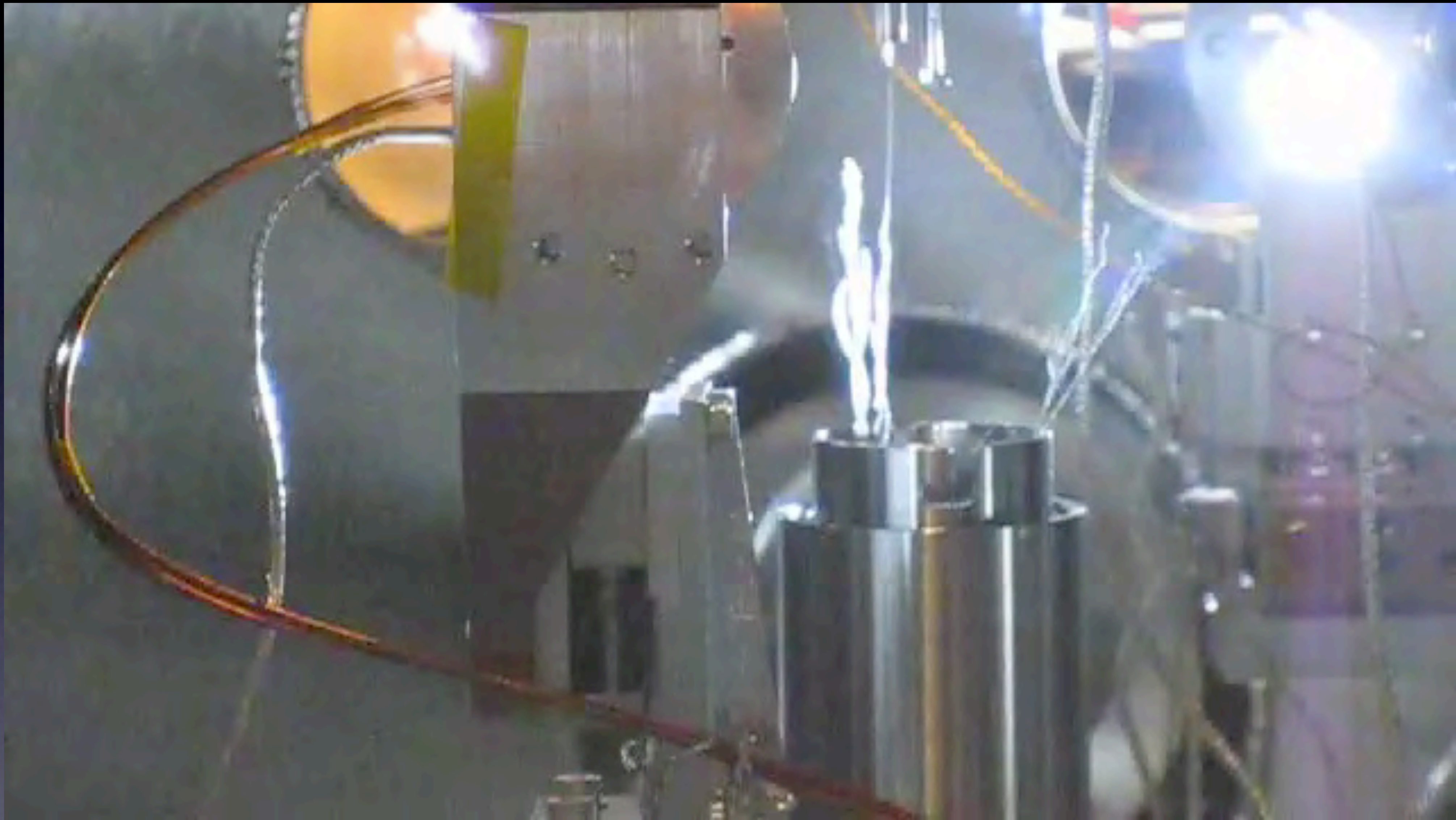


Basic recipe is as follows

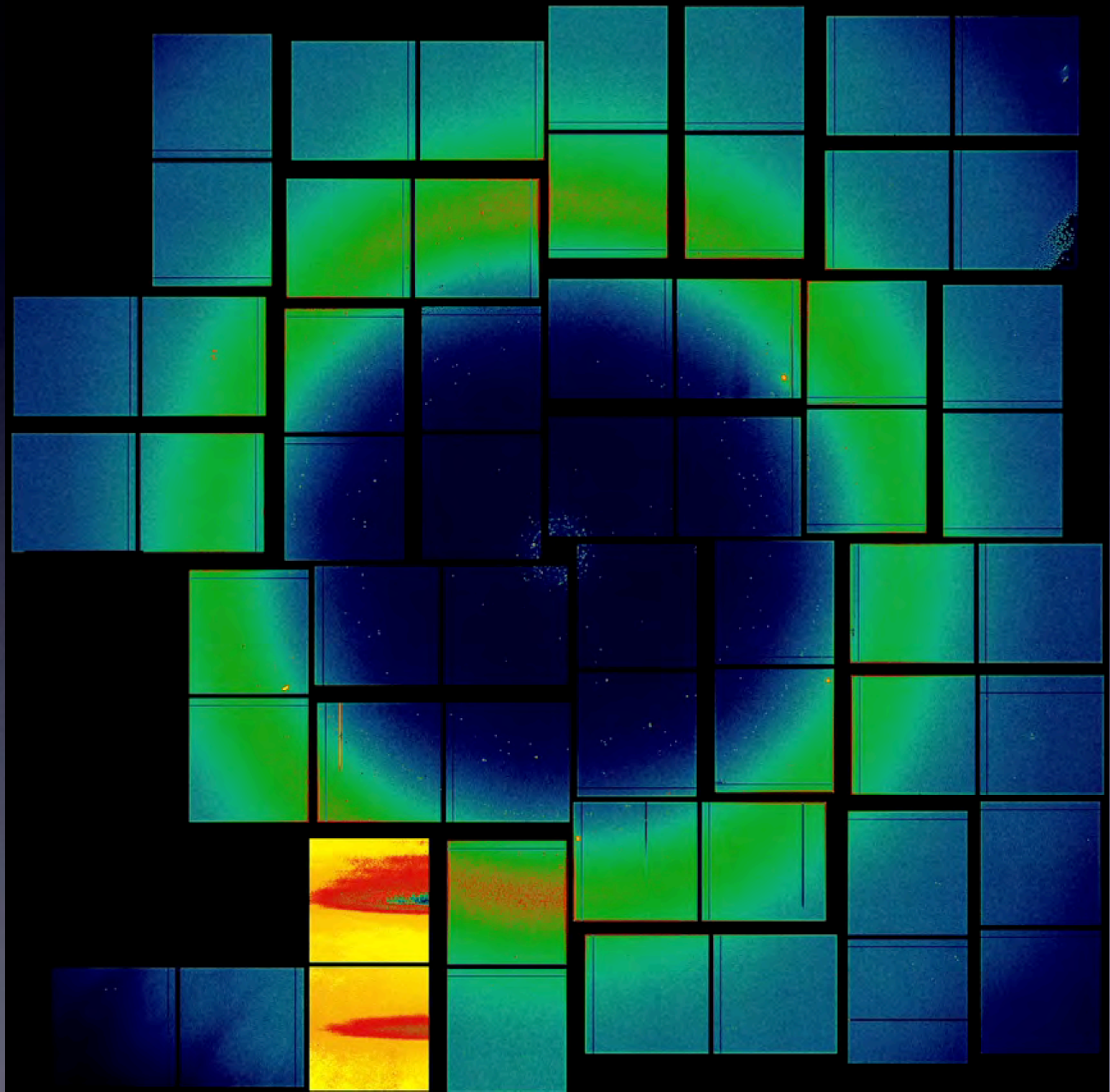
1. Optical metrology
2. Powder pattern alignment
3. Refinement using indexing

CFEL SCIENCE

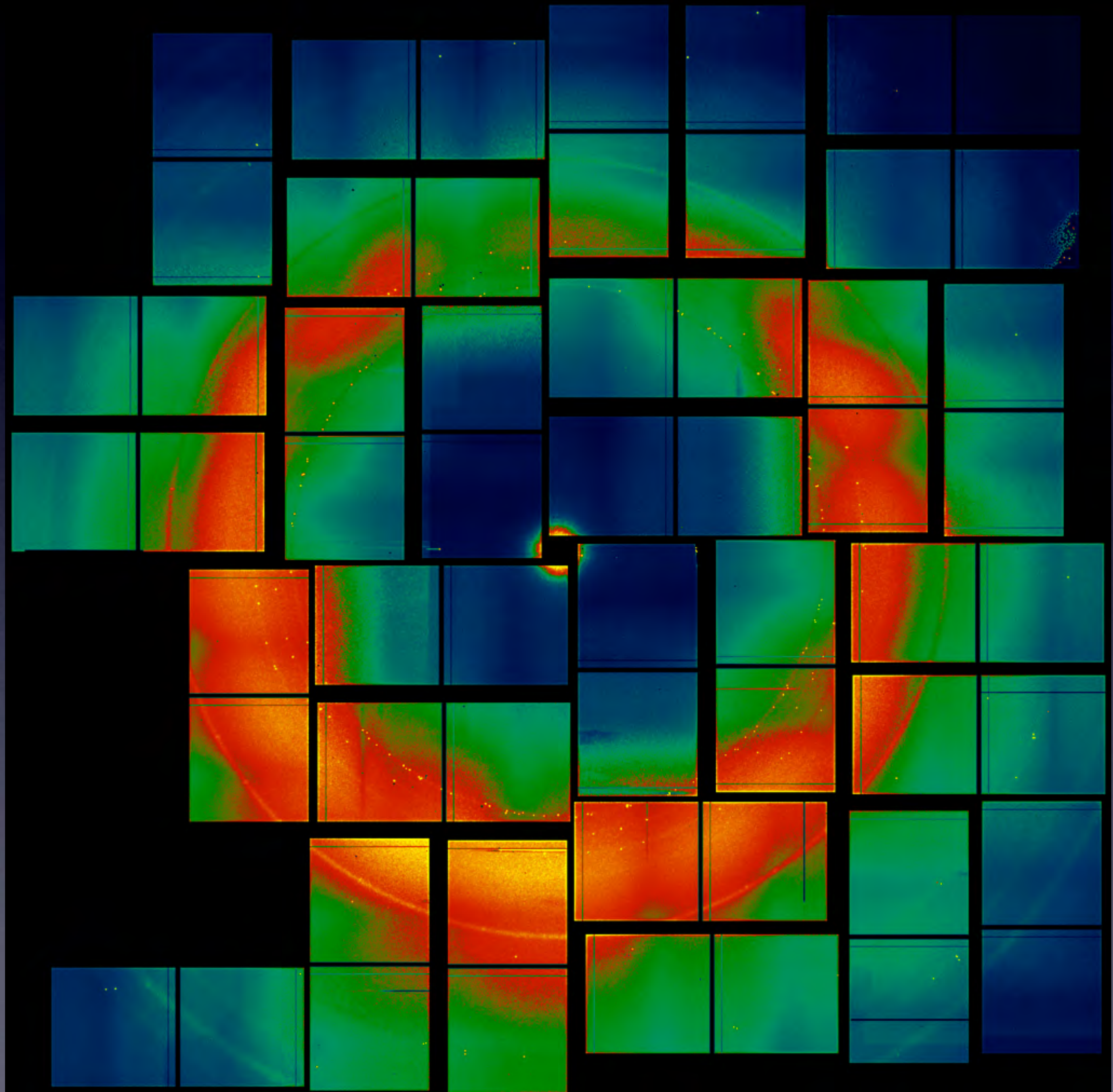
Detectors are not disposable, but unfortunately they are easily damaged during the course of an experiment



CFEL SCIENCE



Sum of all frames is dominated by water ring background



LCLS pulses: 2,292,468

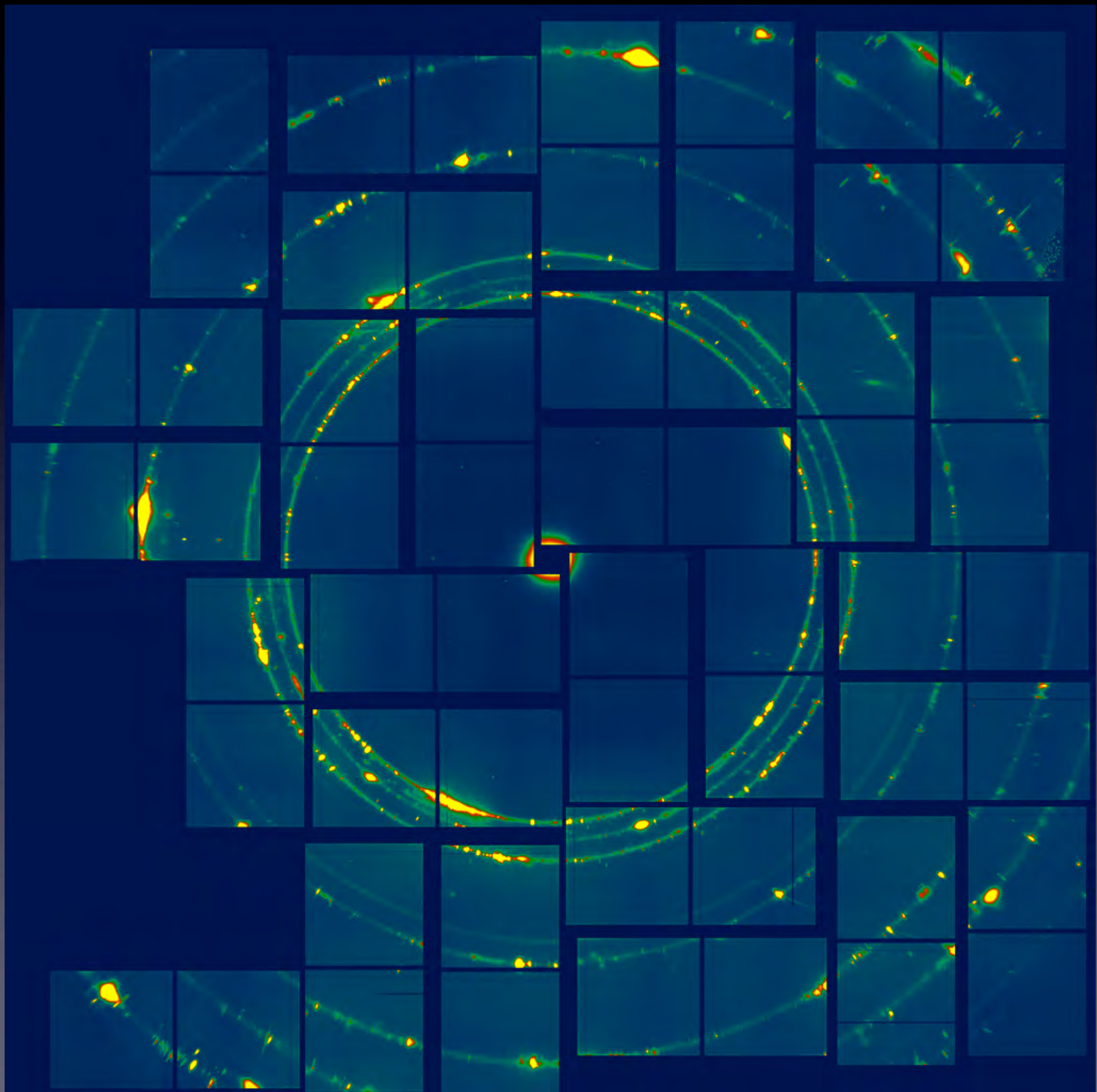
Acquisition time: 5 hr 18 min

Photon energy: 9.4 keV

Up to 5.2×10^8 ADU/pixel

CFEL SCIENCE

Ice gives rise to strong diffraction peaks on the detector



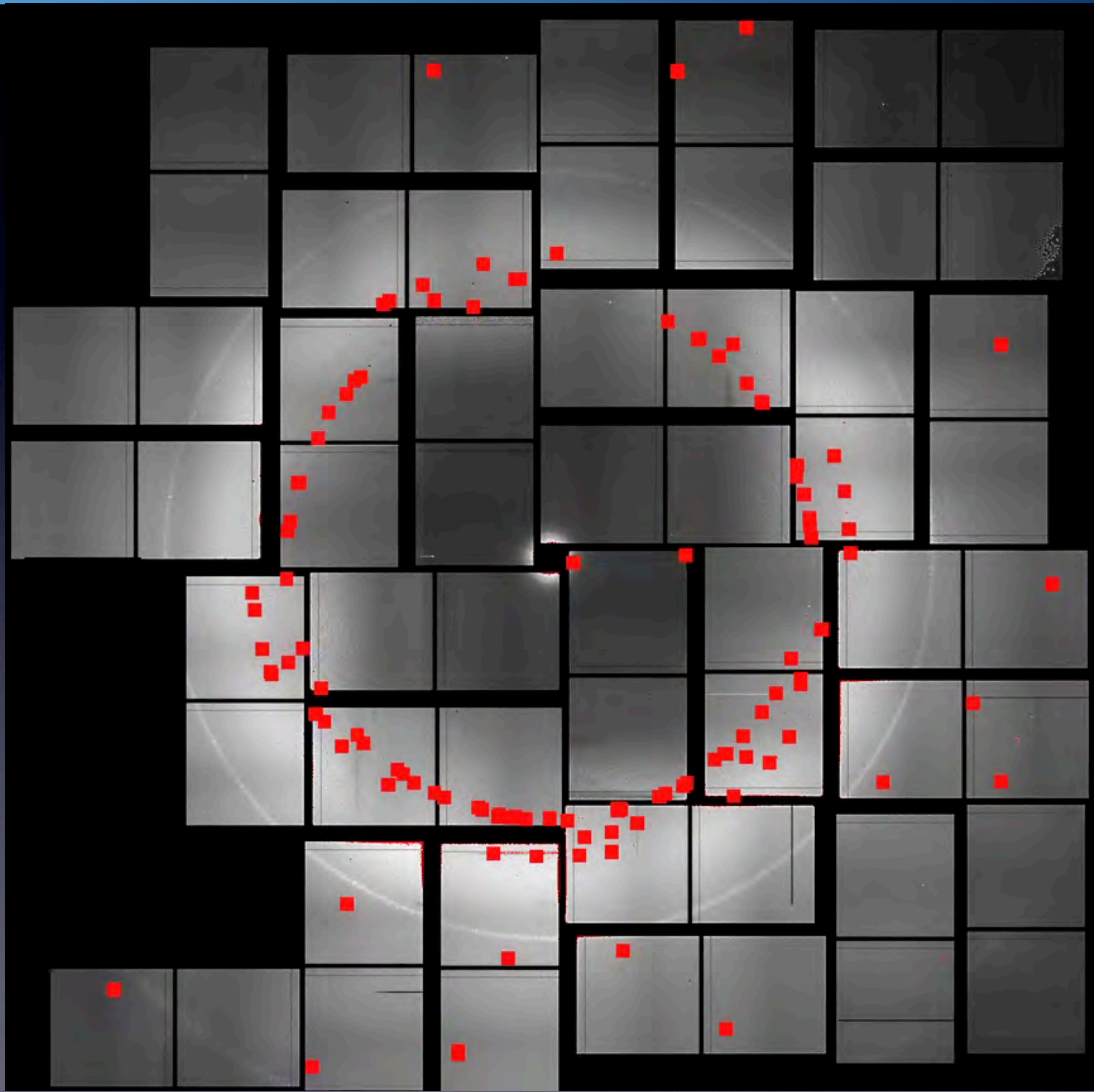
LCLS pulses: 4,293

Acquisition time: 35 seconds

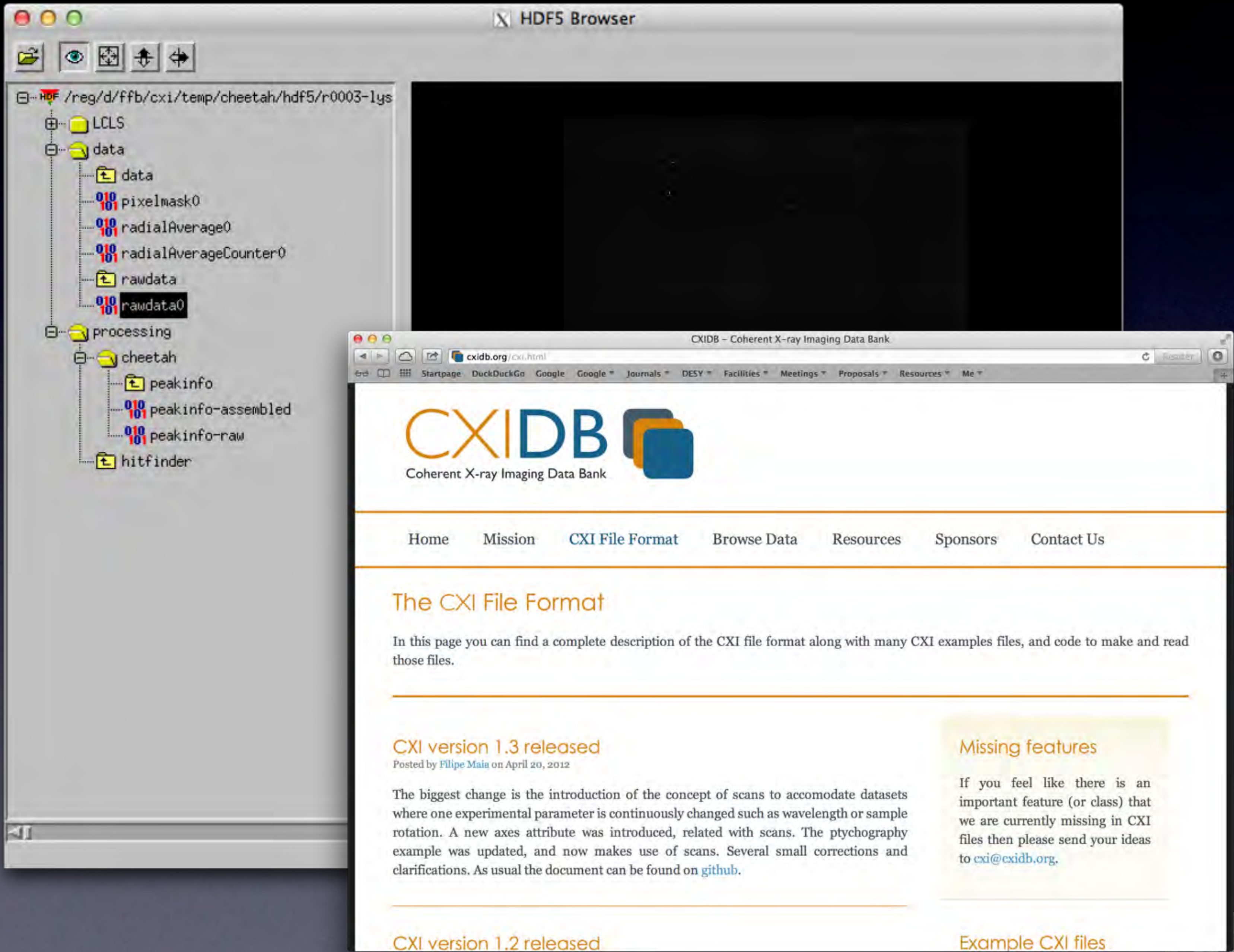
Photon energy: 9.4 keV

Up to 4.1×10^7 ADU/pixel

Dead pixels are identified by Cheetah as they accumulate during the course of the experiment



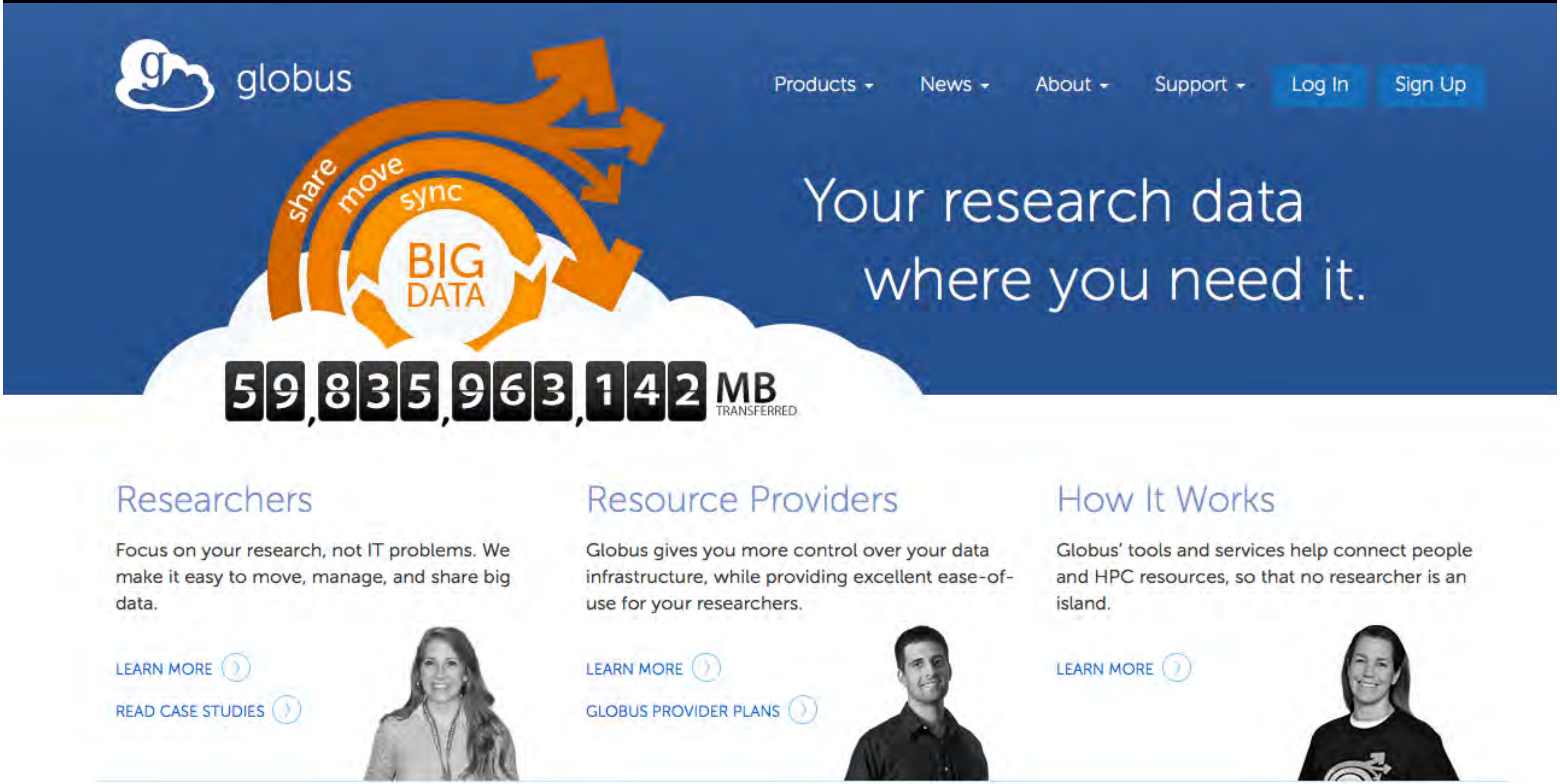
Reduced data is output in facility independent HDF5 format



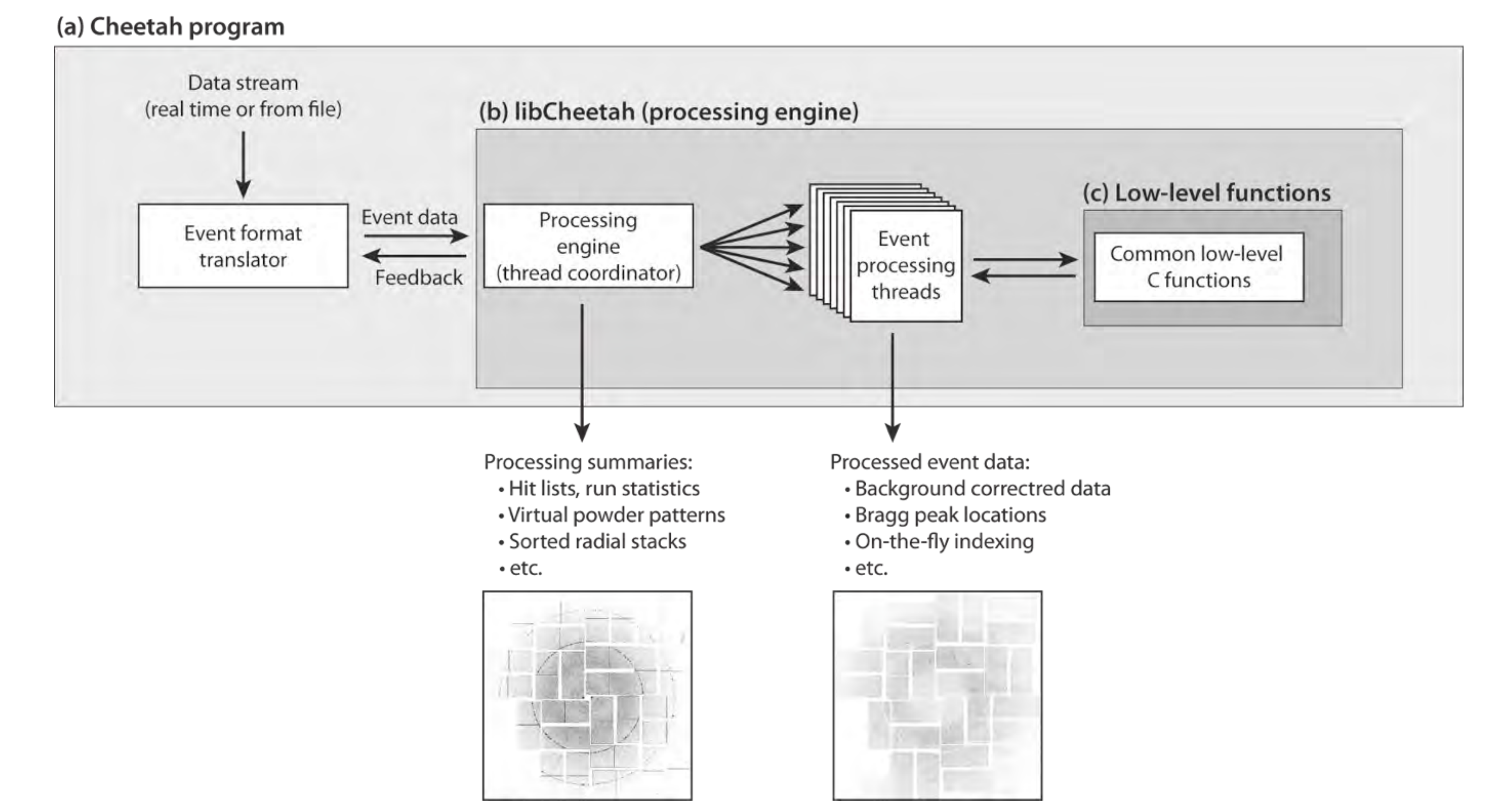
FEL SCIENCE

Globus online is an excellent tool for transferring terabyte datasets to your home institution

https://www.globus.org



Cheetah can easily be deployed at other facilities



Has been used at:
SACLA (.h5)
ESRF (.edf)
Petra III (.cbf)

- Output format is always the same
- CrystFEL does not have to change, and is free from facility dependencies
- The only change is in the file reader

Cheetah is modular by design and open for code reuse

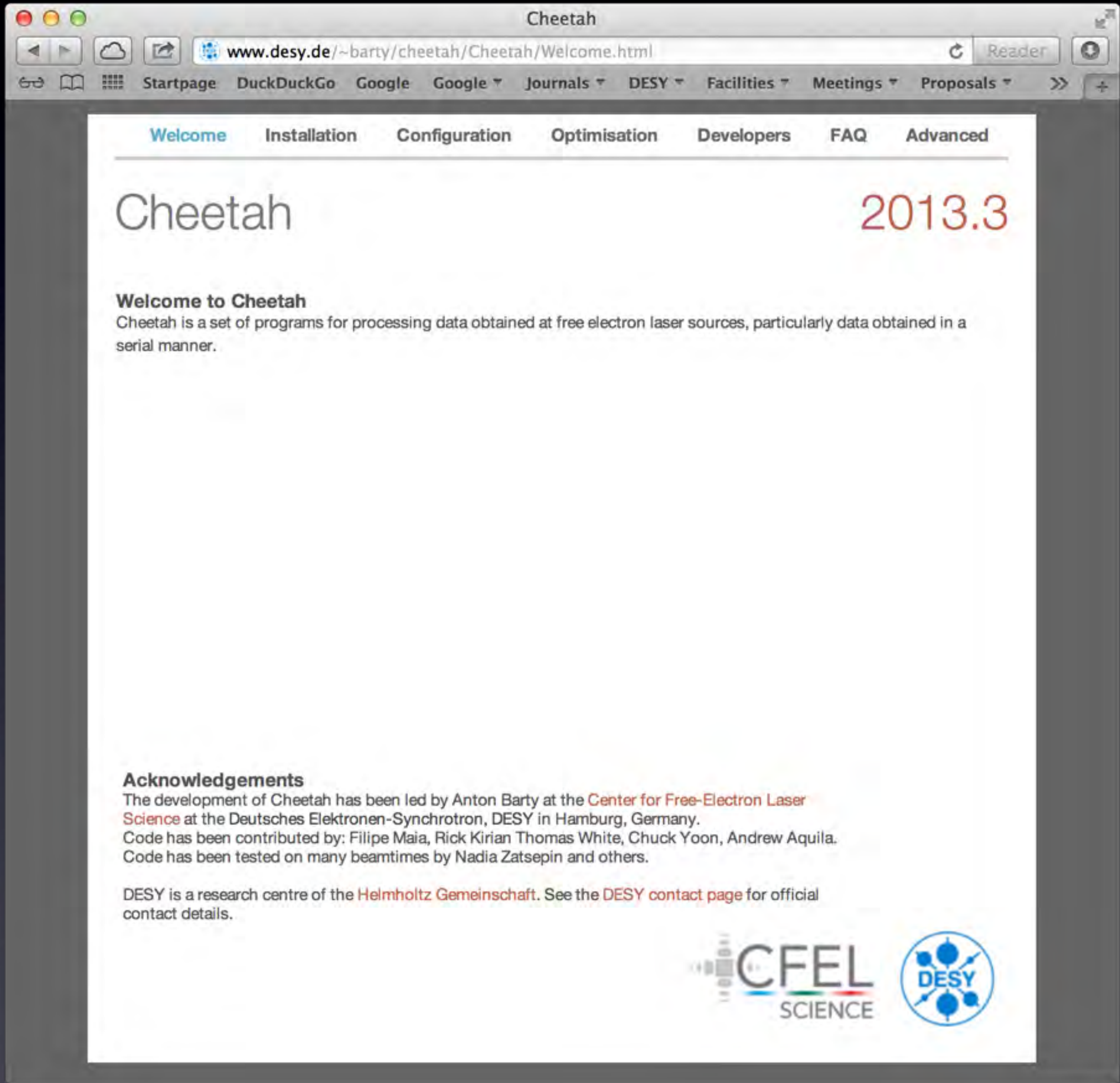
```
/*
 *  Function prototypes
 */
void *worker(void *);

// detectorCorrection.cpp
void subtractDarkcal(cEventData*, cGlobal*);
void applyGainCorrection(cEventData*, cGlobal*);
void applyBadPixelMask(cEventData*, cGlobal*);
void cspadModuleSubtract(cEventData*, cGlobal*);
void cspadModuleSubtract2(cEventData*, cGlobal*);
void cspadModuleSubtract(cEventData*, cGlobal*, int);
void cspadSubtractUnbondedPixels(cEventData*, cGlobal*);
void cspadSubtractBehindWires(cEventData*, cGlobal*);
void calculateHotPixelMask(cGlobal*);
void identifyHotPixels(cEventData*, cGlobal*);
void applyHotPixelMask(cEventData*, cGlobal*);
void calculateHaloPixelMask(cGlobal*);
void updateHaloBuffer(cEventData*, cGlobal*,int);

void subtractDarkcal(float*, float*, long);
void applyGainCorrection(float*, float*, long);
void applyBadPixelMask(float*, uint16_t*, long);
void cspadModuleSubtract(float*, uint16_t*, float, long, long, long, long);
void cspadSubtractUnbondedPixels(float*, uint16_t*, long, long, long, long);
void cspadSubtractBehindWires(float*, uint16_t*, float, long, long, long, long);
long calculateHotPixelMask(uint16_t*, int16_t*, long, long, long);
long calculateHaloPixelMask(uint16_t*, float*, float, long, long);
```

```
// assemble2DImage.cpp
void assemble2Dimage(cEventData*, cGlobal*);
void assemble2Dmask(cEventData*, cGlobal*);
void assemble2Dimage(int16_t*, float*, float*, float*, long, long, long, int);
void assemble2Dmask(uint16_t*, uint16_t*, float*, float*, long, long, long, int);
void downsample(cEventData*, cGlobal*);
void downsampleImage(int16_t*, int16_t*, long, long, long, long);
void downsampleMask(uint16_t*, uint16_t*, long, long, long, long);
```

CFEL SCIENCE

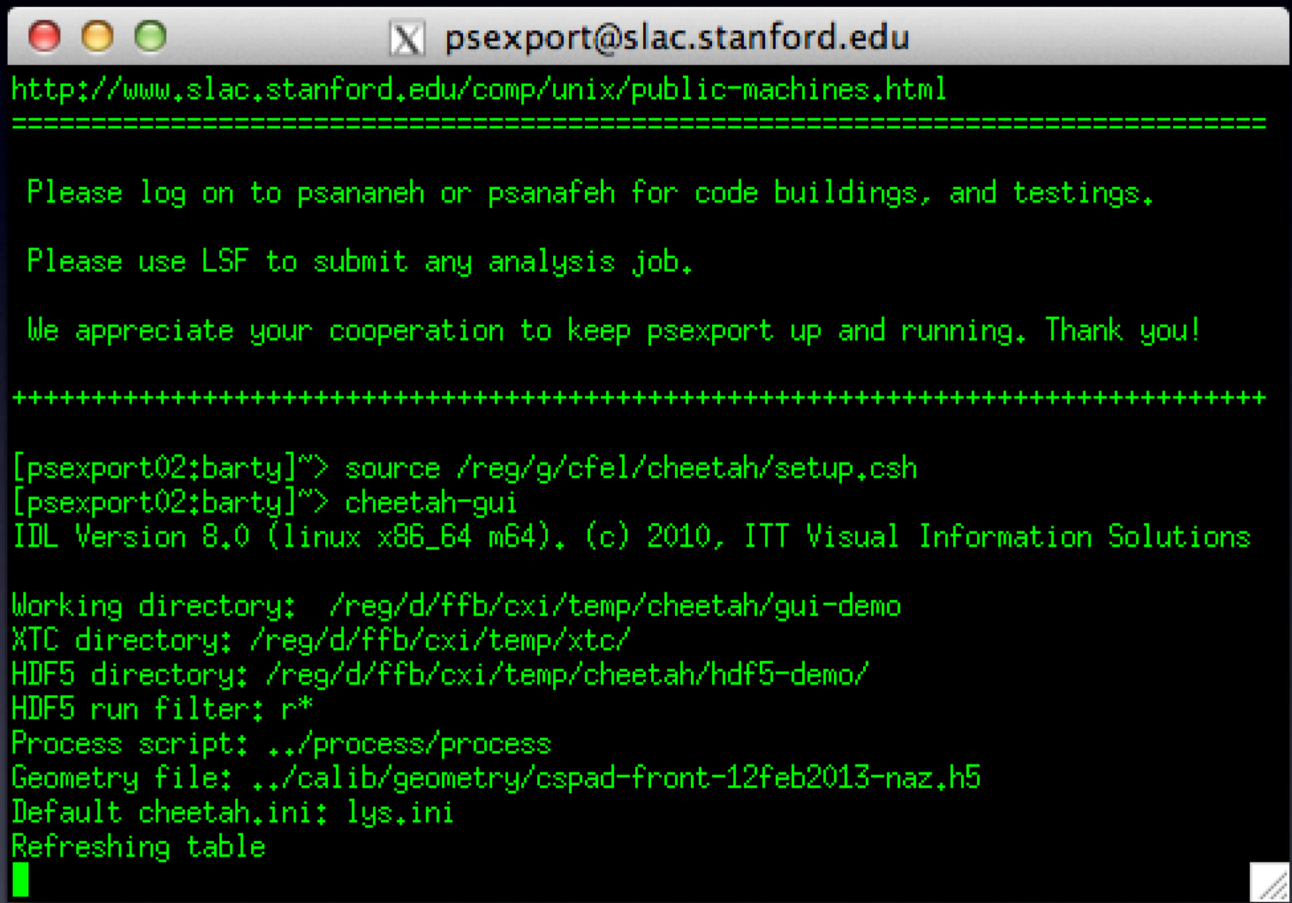
Resources can be found on the web

http://www.desy.de/~barty/cheetah/



Cheetah is pre-installed at SLAC

/reg/g/cfel/cheetah



> source /reg/g/cfel/cheetah/setup.csh
> cheetah-gui

CFEL SCIENCE

The end result is an accurate set of reflection intensities for structure determination

CFEL SCIENCE

A very good summary of cctbx.xfel vs Cheetah + CrystFEL

Protein crystal structure obtained at 2.9 Å resolution from injecting bacterial cells into an X-ray free-electron laser beam

Michael R. Sawaya[a,b,1], Duilio Cascio[a,b,1], Mari Gingery[a,b,1], Jose Rodriguez[a,b], Lukasz Goldschmidt[a,b], Jacques-Philippe Colletier[c,d,e], Marc M. Messerschmidt[f,2], Sébastien Boutet[f], Jason E. Koglin[f], Garth J. Williams[f], Aaron S. Brewster[g], Karol Nass[h], Johan Hattne[g], Sabine Botha[h], R. Bruce Doak[h,i], Robert L. Shoeman[h], Daniel P. DePonte[f], Hyun-Woo Park[j,3], Brian A. Federici[j,k], Nicholas K. Sauter[g], Ilme Schlichting[h], and David S. Eisenberg[a,b,l,4]

[a]UCLA–DOE Institute for Genomics and Proteomics, [b]Department of Biological Chemistry, and [l]Howard Hughes Medical Institute, University of California, Los Angeles, CA 90095-1570; [c]Université Grenoble Alpes, [d]Centre National de la Recherche Scientifique, and [e]Commissariat à l'Energie Atomique, Institut de Biologie Structurale, F-38044 Grenoble, France; [f]Linac Coherent Light Source, SLAC National Accelerator Laboratory, Menlo Park, CA 94025; [g]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720; [h]Max Planck Institute for Medical Research, 69120 Heidelberg, Germany; [i]Department of Physics, Arizona State University, Tempe, AZ 85287; and [j]Department of Entomology and [k]Graduate Program in Cell, Molecular and Developmental Biology, University of California, Riverside, CA 92521

Contributed by David S. Eisenberg, July 23, 2014 (sent for review April 22, 2014)

It has long been known that toxins produced by *Bacillus thurin-*

(3, 9). The crystals for this study were *not* grown in artificial crys-

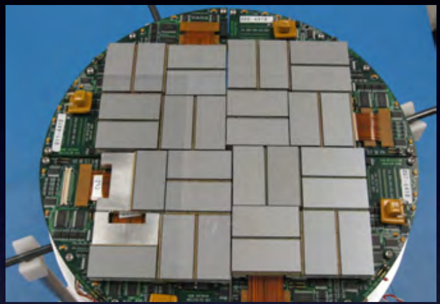
has published results on nine different macromolecular systems since its inception in 2009 (Table 1). One system in particular, cathepsin B, marks an advancement toward in vivo crystallography

[4]To whom correspondence should be addressed. Email: david@mbi.ucla.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1413456111/-/DCSupplemental.

Hit finding may not be as sexy as indexing, but it is very practical

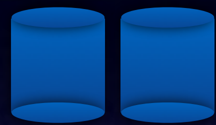
Detector



Storage

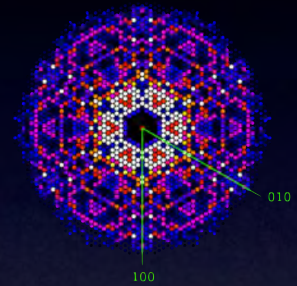
> 4,000,000 frames

Offline reduction



~ 400,000 hits

Analysis



~ 300,000 indexed

Why do it:

1. Quickly evaluate diffraction quality (including seeing your data)
2. Select and retain only the useful events (hit finding)
3. Take home only the useful data for analysis using CrystFEL (or cctbx)

CFEL SCIENCE

So long and thanks for listening

CFEL SCIENCE