TUPPC015

ON-LINE AND OFF-LINE DATA ANALYSIS SYSTEM FOR SACLA EXPERIMENTS

T. SUGIMOTO, Y. FURUKAWA, K. OKADA, R. TANAKA, M. YAMAGA, Y. JOTI, T. KAMESHIMA (JASRI/SPring-8)
and T. ABE (RIKEN SPring-8 Center)

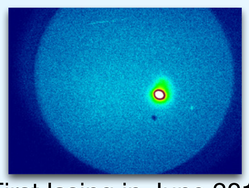
Abstract

User experiments at the SPring-8 angstrom compact free electron laser (SACLA) facility were recently commenced, in March 2012. Typical data rate is up to 5 Gbps using 10 multipart charge-coupled device sensors. To analyze such a large experimental data, we develop data-analysis system for the SACLA experiments. We present an overview of the analysis system and its future.

Overview of SACLA

Background of our Development

SACLA (SPring-8 angstrom compact free electron laser) is X-ray free electron laser (XFEL) facility, located at west region of Japan. First self-amplified spontaneous emission lasing was achieved in June 2011, and user experiments have been performed from March 2012. By utilizing the characteristics of XFEL (high brilliance, coherence, ultra-short pulse), many sciences are carried out.

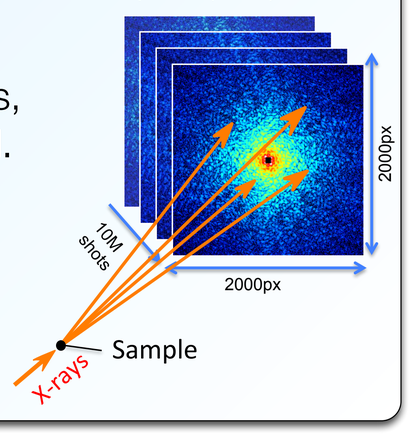


First lasing in June 2011

Characteristic Specification
- Facility length: 700 m
- Electron energy: 8 GeV
- Emittance: 0.77 - 1.1 pimm.mrad
- Repetition rate: 10, 20, 30, 60 Hz
- X-ray wavelength: 0.6 Å (shortest) (15 keV)
- X-ray peak power: 500 uJ/pulse (max.)

One of the sciences at SACLA is structural biology. By analyzing the structure of biomaterials, pharmaceutical development can be accelerated.

To analyze a certain protein using XFEL, 10 million shots of X-ray diffraction images (2K x 2K x 16bit) are required. To accumulate and to analyze such large data, we developed dedicate DAQ system and analysis system.

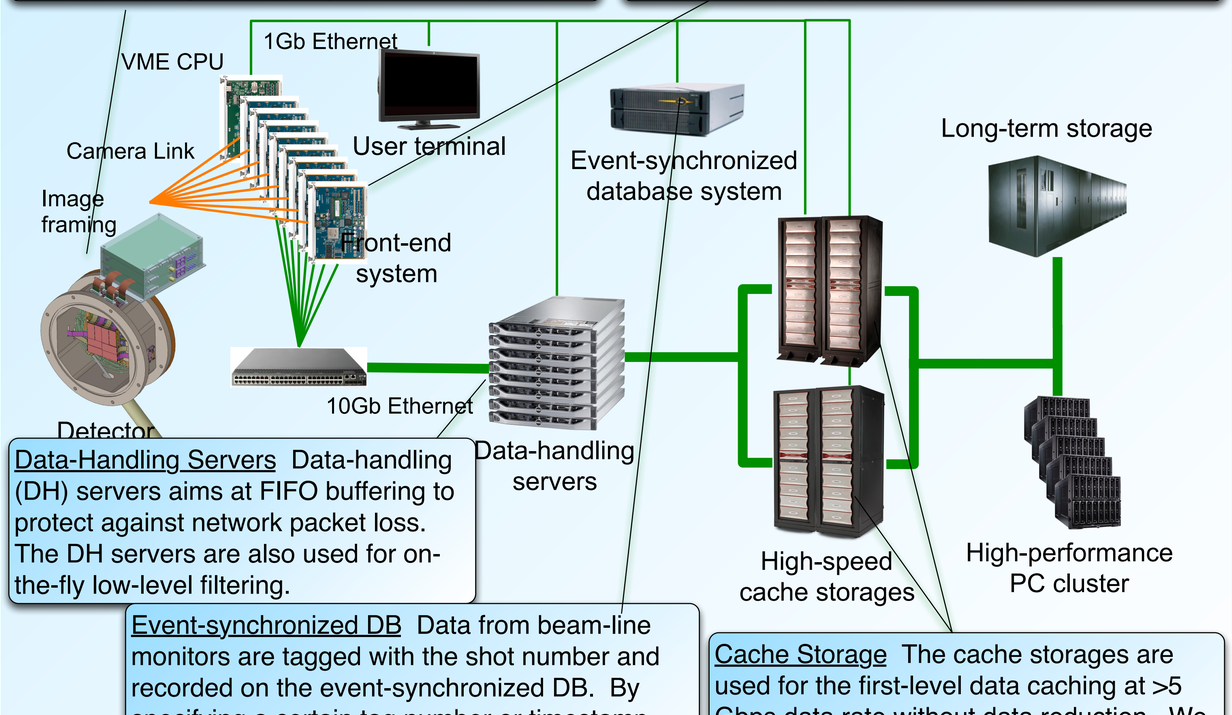


Overview of Data Acquition System

We developed a dedicate DAQ system for SACLA experiments, which is required to satisfy 5 Gbps data rate.

Detector We developed dedicate X-ray 2D image sensor: Multi-Port Charge Coupled Device (MPCCD) sensor. Single MPCCD sensor is 512px × 1024px × 16bit. We use 8+2 sensors at 60 Hz repetition rate. Therefore, maximum data rate to be supported by the DAQ system is 5 Gbps.

Front-End System Camera Link is a standard interface of 2D cameras. The front-end (FE) systems receive and transfer image from Camera Link to Ethernet. We have two type of FE systems; one is PC-based, and another is VME-based. We choose suitable systems for each camera.



Data-Handling Servers Data-handling (DH) servers aims at FIFO buffering to protect against network packet loss. The DH servers are also used for on-the-fly low-level filtering.

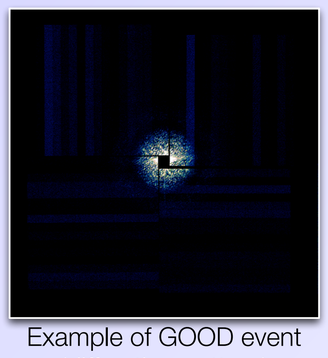
Event-synchronized DB Data from beam-line monitors are tagged with the shot number and recorded on the event-synchronized DB. By specifying a certain tag number or timestamp, we can acquire beam monitor data from the DB. The DB is also used for low-level filtering.

Cache Storage The cache storages are used for the first-level data caching at >5 Gbps data rate without data reduction. We have two cache storages (200TB and 250TB), which are used in rotation.

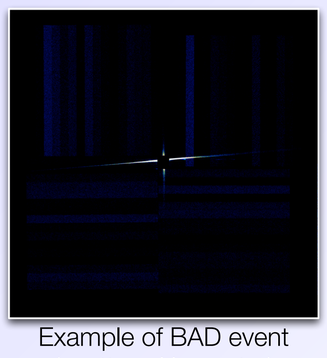
Data Analysis System

Low-Level Filtering

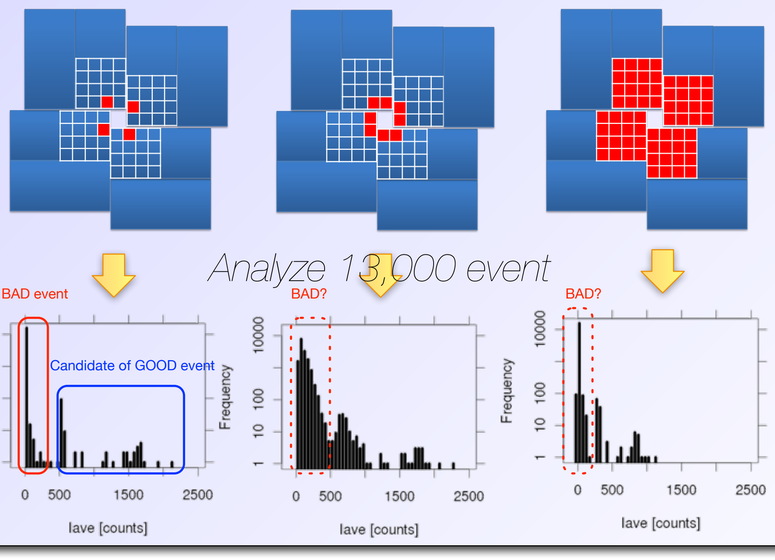
We developed on-the-fly low-level filtering (LLF) system. First LLF application is Region of Interest (ROI) pattern matching. The DH server acquire photon number of each grid on the sensors at 60 Hz repetition rate. These photon numbers are recorded on the event-synchronized DB. In the downstream analysis, we can easily choose candidate events by applying threshold of photon numbers.



Example of GOOD event (diffraction pattern)



Example of BAD event (scattered by water)

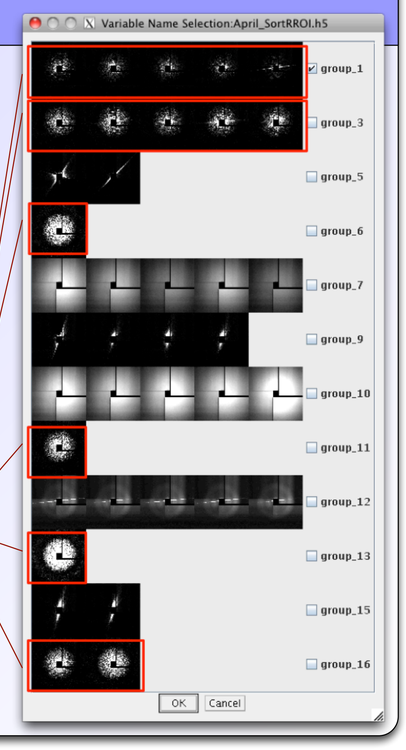


Photon numbers of each grid are recorded on the event-synchronized DB. By applying ROI region (red rectangle), histograms of photon number in the ROIs are extracted.
The left figures show example of ROI patterns. Peak around 0 means BAD event (no diffraction in the ROI).

High-Level Filtering

We develop high-level filtering (HLF) system to classify diffraction images. To reconstruct the real-space 3D image of biomaterial, some million shot of diffraction images are necessary to improve S/N ratio using overlaying technique. Before overlaying, diffraction images must be classified by orientation. For the image classification, we use ImageJ software.

Output of HLF system. Inputs are 38 candidate images filtered by the LLF. Each row shows similar diffraction images classified by ImageJ software. The red rectangle shows GOOD event finally determined by human user.



By utilizing LLF and HLF, it is easy to retrieve GOOD events from several million images.

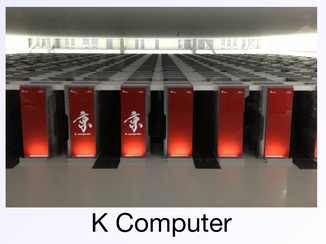
Future Plan

3D image reconstruction requires very high-performance computers. For the reconstruction, we plan to use on-site HPC and off-site K Computer selectively. We are now developing large-bandwidth (> 5Gbps) data transfer system from SACLA to the K Computer.

HPC System

Long-term Storage

data transfer (100 km)



K Computer