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, Sayo, Hyogo 679-5198, Japan
T. Abe, RIKEN SPring-8 Center, Sayo, Hyogo 679-5148, Japan

Abstract

User experiments at the SPring-8 angstrom compact free electron laser (SACLA) facility commenced in March 2012. Typical data rates are up to 5 Gbps using 10 multi-port charge-coupled device sensors. To analyze such a large volume of experimental data, we developed a data analysis system for SACLA experiments. In this paper, we present an overview of the analysis system and its future.

OVERVIEW OF SACLA

The SPring-8 angstrom compact free electron laser (SACLA) facility is an X-ray free electron laser (XFEL) facility located at the SPring-8 site in Japan. This facility is characterized by its compact design; it is about 700 m long and includes the accelerator, undulator, and experimental buildings. To achieve a compact XFEL design, we developed the relevant technical systems such the C-band accelerating cavity, in-vacuum undulator, timing, optics, detectors, and data-acquisition system. A significant advantage of the SACLA-SPring-8 experimental facility is that two X-rays, one from SACLA and one from SPring-8, can be used simultaneously. The first self-amplified spontaneous emission (SASE) lasing of SACLA was achieved in June 2011, and X-ray laser beams have been delivered to users since March 2012 [1]. XFEL wavelengths in the subangstrom region (0.6 Å) have been achieved. During the first year of SACLA, many experiments in the areas of atomic, molecular, and optical physics, ultrafast and material sciences, and structural biology have been carried out.

To support such a wide range of experiments, we developed a dedicated data-acquisition (DAQ) system and a data-analysis system. The DAQ system consists of a number of components: detectors, a front-end system, data-handling servers, a cache storage system, and an event-synchronized database (DB). The analysis system consists of long-term archive storage and a high-performance computing (HPC) system. We also use an external supercomputer for precise analysis. We chose standard interfaces to connect components with each other. To match the needs of future experiments, the DAQ performance will be upgraded by replacing its system components. Figure 1 shows an overview of the SACLA DAQ and analysis system. In the next two sections, we give a short overview of the DAQ system and the SACLA data analysis system.

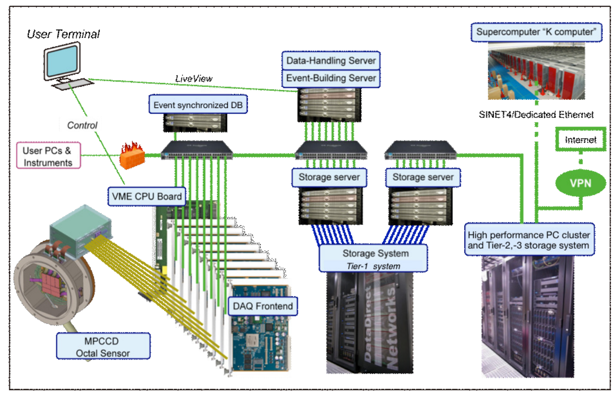

Figure 1: Overview of the DAQ system. The current DAQ system supports up to 10 multi-port charge-coupled device sensors.

OVERVIEW OF THE DATA-ACQUISITION SYSTEM

Supported Detectors

The current DAQ system must support several cameras, as various experiments require different types of sensors. To do this, we adopted the standard digital interface Camera Link [2] and developed a front-end (FE) system for Camera Link cameras. By using Camera Link-supported cameras, various experiments can share the components downstream from the FE system.

One of the supported cameras is a multi-port charge-coupled device (MPCCD) sensor. The MPCCD sensor is a two-dimensional X-ray detector developed for SACLA experiments [3]. Each pixel of the MPCCD is $50 \times 50\ \mu m^2$ with a 16-bit data depth. A single MPCCD sensor module has 512×1024 net pixels. Inclusive of calibration data, 512×1032 gross pixels per shot are acquired by the sensor at a repetition rate of 60 Hz. MPCCD sensors can be used in one of the following typical sensor configurations: 1 (single), 2 (dual), or 8 (octal). We decided to use either the single or the octal + dual combined configurations that have data rates of about 500 Mbps and 5 Gbps, respectively. We developed downstream components for the DAQ system to satisfy such high data rate requirements.

We also support commercially available cameras such as IMPERX [4] and OPAL [5]. By satisfying the data-rate requirements of the MPCCD sensors, these commercial cameras are also supported by the DAQ system.

ISBN 978-3-95450-139-7

Front-End System

Because the physical range of Camera Link is short (several meters), the front-end system receives and transfers data from Camera Link to Ethernet. We developed two different front-end systems: a PC-based system and a VME-based system [6]. The PC-based system is a one-rack-unit Linux PC equipped with a commercial Camera Link grabber board. The advantage of using the PC-based system is its ability to support more than 100 commercial cameras, including full configuration Camera Link cameras. The VME-based system has a dedicated field programmable gate array (FPGA) processor to forward image data from Camera Link to Gigabit Ethernet (GbE). The advantage of the VME-based system is its efficient hardware processing; we implemented on-the-fly lossless image compression using the FPGA processor. The I/O interfaces are implemented with a processor PCI mezzanine card that will be upgraded in the future. For practical applications, we choose the more suitable system for each detector; the PC-based and VME-based systems are used for the OPAL-2000 camera and MPCCD detector arrays, respectively.

Data-Handling Server

Data-handling servers are inserted between the front-end and storage systems and each data-handling server corresponds to an individual sensor. The data-handling servers are for in-line buffering and processing. We implemented a pipeline fast-in-fast-out (FIFO) buffer system on the data-handling server that keeps the DAQ system stable and protects against network packet loss.

The data-handling servers are used not only for the FIFO buffer but also for low-level analysis; we use an x86-based PC server as new low-level filters can be applied as a plug-in. An overview of the current low-level filtering system is described in the next section.

Storage System

Experimental data are accumulated in cache storage. The cache storage systems are for first-level caching without data reduction. We calculated the capacity requirement for the cache storage system to be about 200–300 terabytes (TB) per week, which was estimated from the data rate of the MPCCD sensors (10 sensors), each with a 50% duty factor. We have two cache storage systems (System-A and System-B). The storage capacities of System-A and System-B are 200 and 250 TB, respectively. We intend to use the two systems in rotation; for example, while one system is used for data readout to the analysis system, the other system is used for data acquisition without performance degradation. We chose a single-namespace file system (StorNext and GPFS) for the cache systems because the numbers of sensors can vary by experiment. The minimum guaranteed throughput of the cache storage system is 70 MBps with 10 simultaneous streams and satisfies the required data rate of the MPCCD sensors (60 MBps, $8 + 2$ sensors). It should also be noted that the average throughput of the cache storage system is 500 MBps, which satisfies the estimated data rate of the silicon-on-insulator photon imaging array sensor (SOPHIAS) [7] currently under development.

Event-Synchronized Database

The XFEL is a discrete beam operating at a 60 Hz repetition rate. To distinguish the event data corresponding to each X-ray shot, we developed an event tag system and an event-synchronized DB system [8]. The timing signal is distributed from the XFEL master trigger at 60 Hz and is used by the front-end system both for acquiring the trigger and for incrementing the shot-number. Data acquired from the beam-line monitors are tagged with the shot number and recorded on the event-synchronized DB. An experimenter can acquire specific beam monitor data from the event-synchronized DB by specifying a certain tag number or timestamp.

OVERVIEW OF THE DATA-ANALYSIS SYSTEM

The analysis procedure must be faster than the DAQ throughput of 5 Gbps. To achieve the required throughput, we developed a dedicated data-analysis system for SACLA experiments. Current analysis systems are classified into two types: “on-line” and “off-line” analysis. The on-line analysis is for fast event marking, while the off-line analysis performs a detailed analysis of the marked event. In this section, we give an overview of the analysis system and these procedures.

Computing Resources

High-Performance Computing System Because the images obtained from X-ray diffraction experiments are reciprocal lattice spaces, a sophisticated calculation is necessary to extract real-space images [9]. To perform such calculations, we installed an HPC system in the SACLA computer room. (Figure 2) The HPC system consists of an 80-node cluster system of 12.7 TFLOPS CPUs, and a single-node symmetric multiprocessing (SMP) system at 289 GFLOPS with 1 TB memory. For the working area, 170 TB shared storage is interconnected by an InfiniBand QDR network. The cluster system is used for parallel processes such as image filtering and 2D Fourier transform calculations. The SMP system is used for large-scaled calculations such as fine-meshed 3D Fourier transforms. The HPC system is used not only for off-line analysis but also for on-line instant visualization.

Long-Term Storage System The capacities of the cache storage systems (200 TB and 250 TB) are not large enough to archive experimental data. Hence, to archive experimental data for more than one year, we installed long-term storage in the SACLA computer room. (See Figure 3.) The long-term storage is a hybrid storage system that consists of a 1 PB disk array and automated tape-library system. The current tape-library system has a capacity of 6 PB

ISBN 978-3-95450-139-7

Figure 2: High-performance computing system installed in the SACLA computer room. The HPC system consists of an 80-node cluster system and a single-node SMP system.

Figure 3: Long-term storage system installed in the SACLA computer room consisting of a 1 PB disk array and 6 PB tape library.

and is equipped with 1700×4 TB cartridges. The tape-library has 7500 tape-cartridge slots so that the capacity is upgradable up to 28 PB.

On-Line Analysis

Low-Level Filtering As described in the previous section, the DAQ system accumulates all the 2D image data in cache storage without data reduction. To easily search for a particular event, we developed an on-the-fly event-marking system. This on-line analysis system records certain physical quantities of the candidate shot on the event-synchronized DB. By thresholding the physical quantities, candidate tag numbers are found in the event-synchronized DB and candidate shot images can be selectively read out from the cache storage. Event marking lowers the readout throughput such that it is below that of accumulation.

The on-the-fly marking system is implemented on the

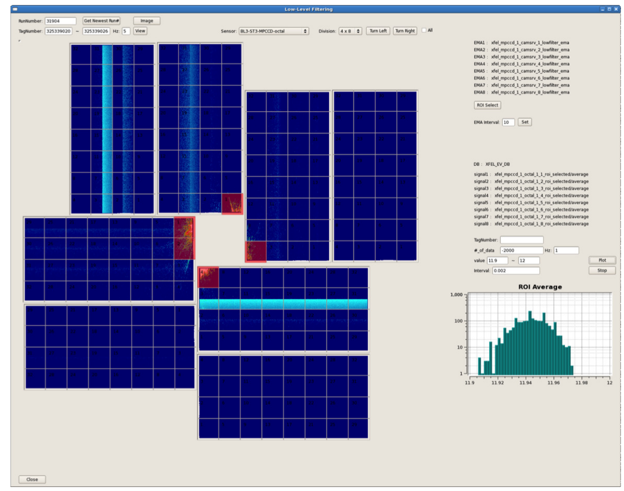

Figure 4: GUI image of the low-level filtering system. The region to the left shows MPCCD image data. A histogram of the ROI average photon number is shown in the lower right.

data-handling servers. Since the marking period per shot must be less than 16.6 ms (= 60 Hz) only the primitive physical quantities can be calculated. The marking system is a so-called "low-level" filtering system.

The first low-level filtering implemented is a grid-based region-of-interest (ROI) statistical analysis. Figure 4 shows a graphical user interface (GUI) image of the filtering system. The region to the left shows MPCCD image data and a histogram of the ROI average photon number is shown on the lower right. In the current filtering system, each region is split into a 4×8 grid. The average photon number for each grid is calculated by the data-handling server for each shot and is recorded on an event-synchronized DB. By choosing a region of interest using the GUI, primitive image-pattern selection can be performed without the need to read the entire image from the storage system.

Instant Visualization Real-space images from on-line reduced analyses help to determine the feasibility of carrying out the experiment. By examining these images, we can determine whether the experimental conditions need to be changed. To perform such instant visualization, we use the HPC system in an interactive mode.

Off-Line Analysis

High-Level Filtering To reconstruct the real-space 3D image of a biomaterial, several million shots of reciprocal-lattice-space images are required. Since the number of photons detected per shot is not very accurate, thousands of shots must be overlaid to improve the signal-to-noise ratio. Before overlaying, images are classified by orientation using ImageJ software [10] on the HPC system.

Figure 5 shows an example of this high-level filtering. This classification was automatically performed on 13,000 shots. Similar images, as classified by ImageJ, are arranged

ISBN 978-3-95450-139-7

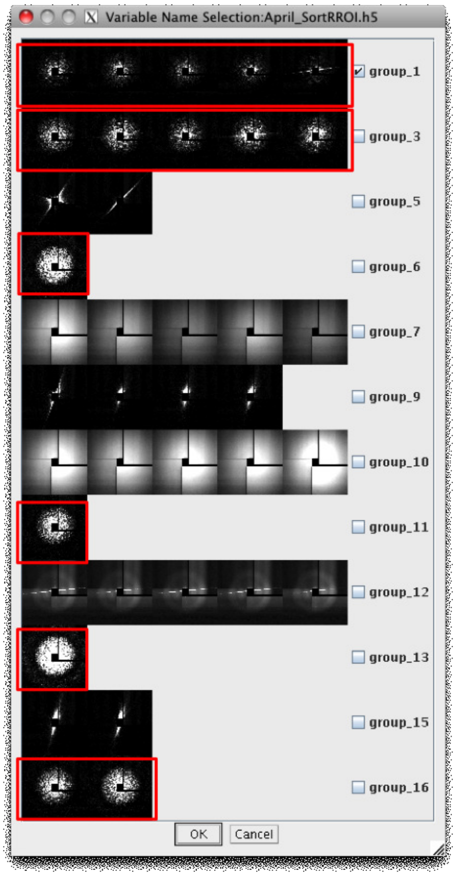

Figure 5: Example output of the high-level filtering. Images that have been classified as similar are shown in rows. The high-level filtering system helps to determine good events in the accumulated data.

in rows. The 15 shots enclosed by red rectangles represent good events as determined by a human user. By utilizing low-level and high-level filtering, it is easy to retrieve good events from several million 2D images.

3D Image Reconstruction Using a Supercomputer To extract real-space 3D images of large biomaterial requires two weeks using the current HPC system. Therefore, we plan to use an external supercomputer, the K computer [11] located at Kobe, about 100 km from SACLA. The K computer has a 10 petaFLOPS performance, and is configured for large-scaled calculations such as 3D fast Fourier transformations. By utilizing the K computer, we expect that an analysis of a biomaterial could be completed within 20 minutes. Therefore, we have begun to develop an off-line analysis system for the joint use of the SACLA and K computers.

One of the items in development is data sharing between the SACLA and K computers for which network bandwidth is important. Data transfer trials between SACLA and the K computer have been performed since 2011 and we have achieved 6.4 Gbps bandwidth, faster than the current DAQ data rate. We have also developed data transfer GUIs for experimenters that will be in operation in October 2013.

UPGRADE PLANS

Computing Resources Upgrade

We plan to install a K computer compatible supercomputer at the SPring-8 site. The planned machine will have 90 TFLOPS, which is 1/100th the performance of the K computer. By placing a supercomputer on site, we will be able to: (1) avoid the bandwidth loss due to data transfer over WAN, (2) determine queue priority by ourselves, and (3) selectively choose between the K computer and the on-site supercomputer, depending on the scale of the analysis. The new supercomputer will be installed in March 2014.

REFERENCES

[1] T. Ishikawa et al., “A compact X-ray free-electron laser emitting in the sub-angstrom region,” Nature Photonics **6**, 540 (2012).

[2] Camera Link®, Automated Imaging Association, `http://www.visiononline.org/`.

[3] T. Kameshima et al., “Development status of X-ray 2D detectors for SPring-8 XFEL,” IEEE Nuclear Science Symposium, Knoxville, TN, USA, October 2010.

[4] IMPERX Camera Link Cameras, `http://www.imperx.com/`.

[5] Adimec OPAL-2000 Camera, `http://www.adimec.com/`.

[6] A. Kiyomichi et al., “Development of image data acquisition system for 2D detector at SACLA,” Proceedings of ICALEPCS2011, Grenoble, France, 947 (2011).

[7] T. Hatsui et al., “Development and deployment status of X-ray 2D detector for SACLA,” International Workshop on Semiconductor Pixel Detectors for Particles and Imaging (PIXEL2012).

[8] M. Yamaga et al., “Event-synchronized data acquisition system of 5 giga-bps data rate for user experiment at the XFEL facility, SACLA,” Proceedings of ICALEPCS2011, Grenoble, France, 581 (2011).

[9] K. J. Gaffney and H. N. Chapman, “Imaging atomic structure and dynamics with ultrafast X-ray scattering,” Science **316**, 1444 (2007).

[10] “ImageJ”, Image Processing and Analysis in Java, National Institutes of Health, USA. `http://rsbweb.nih.gov/ij/`

[11] K computer, RIKEN Advanced Institute for Computational Science, `http://www.aics.riken.jp/`; The K computer is the fourth fastest supercomputer in the world as of June 2013, `http://www.top500.org/lists/2013/06/`.

ISBN 978-3-95450-139-7