ISBN: 978-3-95450-193-9 doi:10.18429/JACoW-ICALEPCS2017-THDPL02

GigaFRoST (GIGABYTE FAST READ-OUT SYSTEM FOR TOMOGRAPHY): CONTROLS AND DAQ SYSTEM DESIGN

T. Celcer†, Paul Scherrer Institut, 5232 Villigen PSI, Switzerland

Abstract

The GigaFRoST (Gigabit Fast Read-out System for Tomography) detector and readout system used at the tomographic microscopy beamline TOMCAT of the Swiss Light Source will be presented. GigaFRoST was built at Paul Scherrer Institute (PSI) and designed to overcome the limitations of existing commercially available high-speed CMOS detectors. It is based on a commercial CMOS fast imaging chip (pco.dimax) with custom-designed readout electronics and control board. The latter is used for detector configuration, coordination of image readout process and system monitoring. The detector can acquire and stream data continuously at 7.7 GB/s to a dedicated backend server, using two data readout boards, each equipped with two FPGAs, and each directly connected with the server via four 10 Gbit/s fiber optics connections. The paper focuses on the implementation of the EPICS control system, data acquisition (DAQ) system, integration of the detector into the beamline infrastructure and implementation of efficient distribution of TTL triggers between the devices involved in the experiments (i.e., GigaFRoST detector, sample rotation stage, arbitrary external devices).

INTRODUCTION

The observation of the full volumetric structural evolution of a sample during a dynamic process with both high temporal and spatial resolution has been a key challenge for the tomographic microscopy beamline TOMCAT [1] at the Swiss Light Source. While spatial resolution has remained at the micrometer level, the temporal resolution has witnessed a considerable and rapid growth in the last years. Today it is shown that 20 tomographic scans may be acquired within one second [2]. Naturally this reflects in the vast amount of generated experimental data (up to several tens of TB/day of raw data). The main limitations under such conditions arise from the characteristics of commercially available high-speed CMOS detectors. The fact that these detectors store data in on-board memory results in (a) the inability of real-time process observation and (b) inability to acquire large number of frames due to the limited available memory. In addition, the consequent transfer of data from the internal memory onto the beamline storage server prevents the efficient use of the detector for experimental needs. In order to overcome these limitations, the GigaFRoST camera system was designed at PSI.

GIGAFROST CAMERA SYSTEM

GigaFRoST (Gigabit Fast Read-out System for Tomography) is a high-speed camera and readout system, built at Paul Scherrer Institute (PSI), and being able to sustain kHz frame rate image acquisition over prolonged periods of time. It uses the commercial CMOS sensor headboard from the pco.dimax detector with fast imaging chip with 2016x2016 pixels, 12 analogue-to-digital converters (ADC), and some auxiliary electronic in order to use its excellent high-speed characteristics (pco.dimax provides sub 20 µm pixel size and a multi-kHz frame rate) [3]. It is also possible to readout a limited Region of Interest (ROI), however due to the design of the imaging sensor (divided in 4 quadrants) and the implementation of the readout process (it stars from the centre of the sensor towards the outer part) the ROI has to be centred.

The problem with the limited on-board storage and lack of image preview is solved with custom-designed readout electronics. Two data readout boards, each equipped with two FPGAs are connected to the sensor headboard (each FPGA is responsible for reading out one quadrant of the imaging sensor (3 ADCs)). Each readout board is then directly connected with the dedicated backend server via four 10 Gbit/s fibre optics connections, which enables the detector to acquire and stream data continuously, without the need for an on-board storage (apart from a minimum amount needed for on-the-fly pixel correction). The maximum data throughput that can be reached is 7.7 GB/s.

The last component is a control board equipped with one FPGA with a built-in PowerPC (PPC) processor (PPC440) and attached RAM. It is responsible for the configuration of the detector system, system monitoring and coordination of the readout process. The control board is equipped with a 1Gbit/s network module used for communication with the camera and with a serial connection which can be used for debugging purposes (direct connection to the console).

Additionally, 6 BNC connectors (2 input and 4 output) using TTL logic, can be used for external image acquisition synchronization. The inputs are used for external enable and trigger signals, while the outputs are used for reporting different camera states (busy, exposure sync-out with optional delay, etc.).

Figure 1 shows a schematic view of the GigaFRoST camera system architecture, while Figs. 2 and 3. show the photos of the GigaFRoST camera (without housing covers).

† tine.celcer@psi.ch

ISBN: 978-3-95450-193-9 doi:10.18429/JACoW-ICALEPCS2017-THDPL02

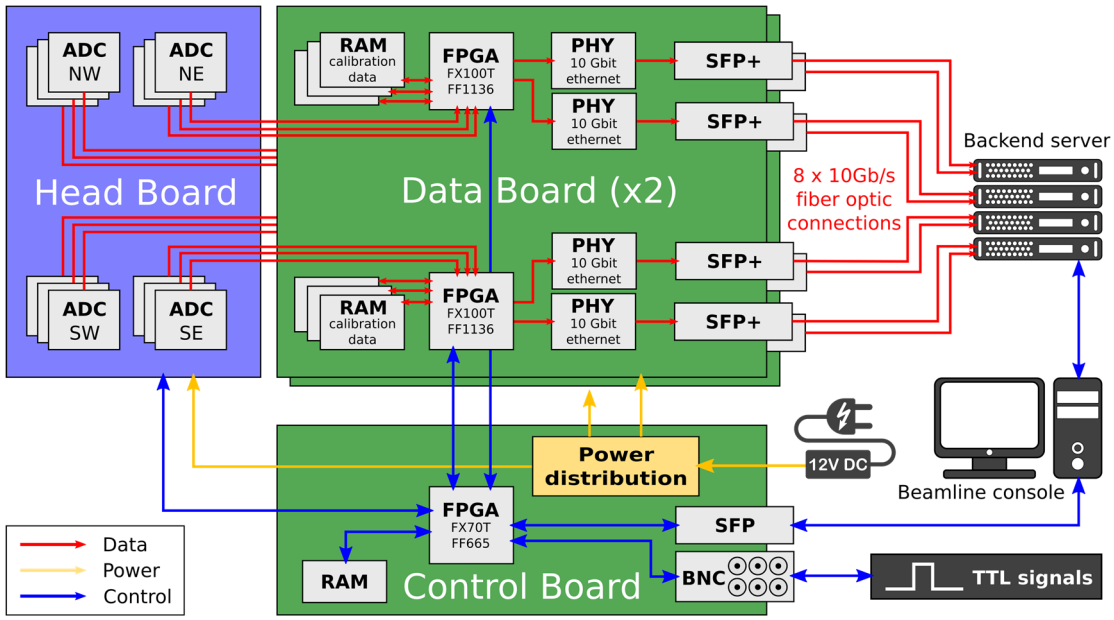

Figure 1: Schematic view of the GigaFRoST camera system architecture.

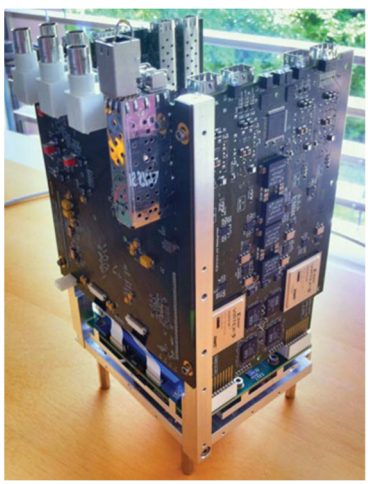

Figure 2: Stripped GigaFRoST camera. Visible are one readout board and the control board with BNC connectors and single Ethernet connection.

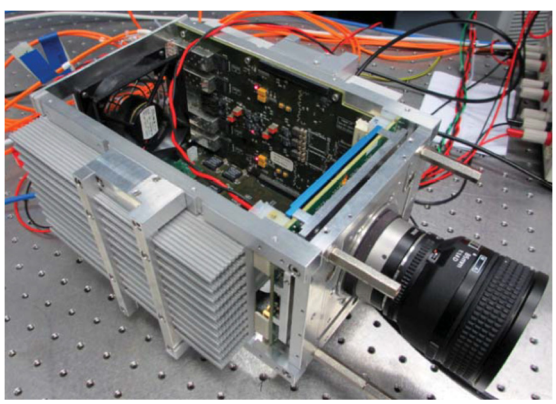

Figure 3: GigaFRoST camera (without housing cover) with the objective, sensor headboard in front, control board at the bottom and one readout board on the side.

CAMERA CONTROL

As already mentioned in the previous section, the GigaFRoST control board is responsible for coordinating the readout system. Its FPGA directly controls the sensor configuration and sensor triggering (configuration of different triggering/operation modes), coordinates the image readout and performs system status monitoring (temperatures and voltages of the control board itself, readout boards and sensor headboard). It also provides a possibility to remotely restart or power cycle either the complete detector, or only individual readout boards and even individual FPGAs.

The GigaFRoST detector is integrated into the control system via an EPICS [4] Input-Output Controller (IOC), running on the control board's PPC which runs an embedded Linux operating system (OS). Thus, all standard communication during experiments takes place via channel access calls (i.e., EPICS process variables (PVs)).

EPICS device driver support was written at PSI to interface the EPICS IOC with all major detector functionalities. Device driver currently provides support for five different EPICS record types (analog input (ai), analog output (ao), binary input (bi), binary output (bo) and waveform (wf) records). Device support consists of the EPICS specific part and a low-level, generic, control system independent part, which facilitates the integration of the Gigafrost detector into other control systems. The in-house, custom defined detector API enables straightforward integration of the detector into the control system as well as an easy upgrade of camera features and their integration into the control system in case of future requests.

The design strategy of the EPICS interface is that each camera parameter is not written to the detector at the moment when its corresponding EPICS PV is processed, but it is rather stored in a local data structure, which is

ISBN: 978-3-95450-193-9 doi:10.18429/JACoW-ICALEPCS2017-THDPL02

sent to the detector in one single call, whenever a 'master' EPICS PV is processed. Similar approach is taken when reading the values from the detector into the EPICS control system. The detector periodically writes the data into the local data structure and sends an interrupt which triggers the processing of corresponding EPICS PVs (readout processing can be triggered also manually, through an EPICS channel). There exists of course a subset of EPICS records which have the direct access to the camera, e.g., restarting/shutting down the camera or its individual readout boards or FPGAs, or setting some low-level camera operation parameters. One important feature is also the provision of the so-called 'in-sync' flag, which defines if the value stored in EPICS PVs are the same as the values on the detector.

Figure 4 shows the expert caQtDM Graphical User Interface (GUI) used to control the GigaFroST detector.

DATA ACQUISITION SYSTEM

In order to efficiently stream and manipulate the vast amount of data generated by GigaFroST detector a high performant data backend system has to be provided. The IT infrastructure consists of a capable backend server responsible for reception and assembling of data sent over eight 10 Gbit/s fiber optics connections (received by a specialized network switch which routes data to the backend server over two 40 Gbit/s ethernet connections). Assembled images are then streamed further to the file storage server, file writer server and the reconstruction cluster.

The data (unidirectional) is streamed between different processes using the ZeroMQ (ZMQ) [5] distributed messaging protocol (one sender can send to many receivers), while the control of individual processes (bi-directional messaging) is passed via Representational State Transfer (REST) interface [6]. The DAQ infrastructure and its topology is shown in Fig. 5.

Backend Server

The data backend server is equipped with two CPUs with 14 cores each. Per CPU, one dedicated 40 Gbit/s network card and 256 GB of RAM are available. Each CPU receives the data from one of the readout boards (one half of the image divided into 4 incoming streams). The data streams are assigned to separate receiver processes running on dedicated cores. Therefore, in the first step each half of the image is reassembled on separate CPUs and then put together and placed in a specialized ring buffer architecture used for high-frequency memory access. Once the frame is assembled successfully, it can be sent further to downstream clients.

The backend server is connected with the rest of the infrastructure (file storage system, file writer server and reconstruction cluster are connected to the same switch) via two 56 Gbit/s Infiniband FDR connections. A separate stream - usually (but not necessarily) with reduced rate - is also sent from the backend server (via a smaller ring buffer) to serve for live image previewing in the experimental control room. However, processing of those frames has a lower priority compared to the main data stream.

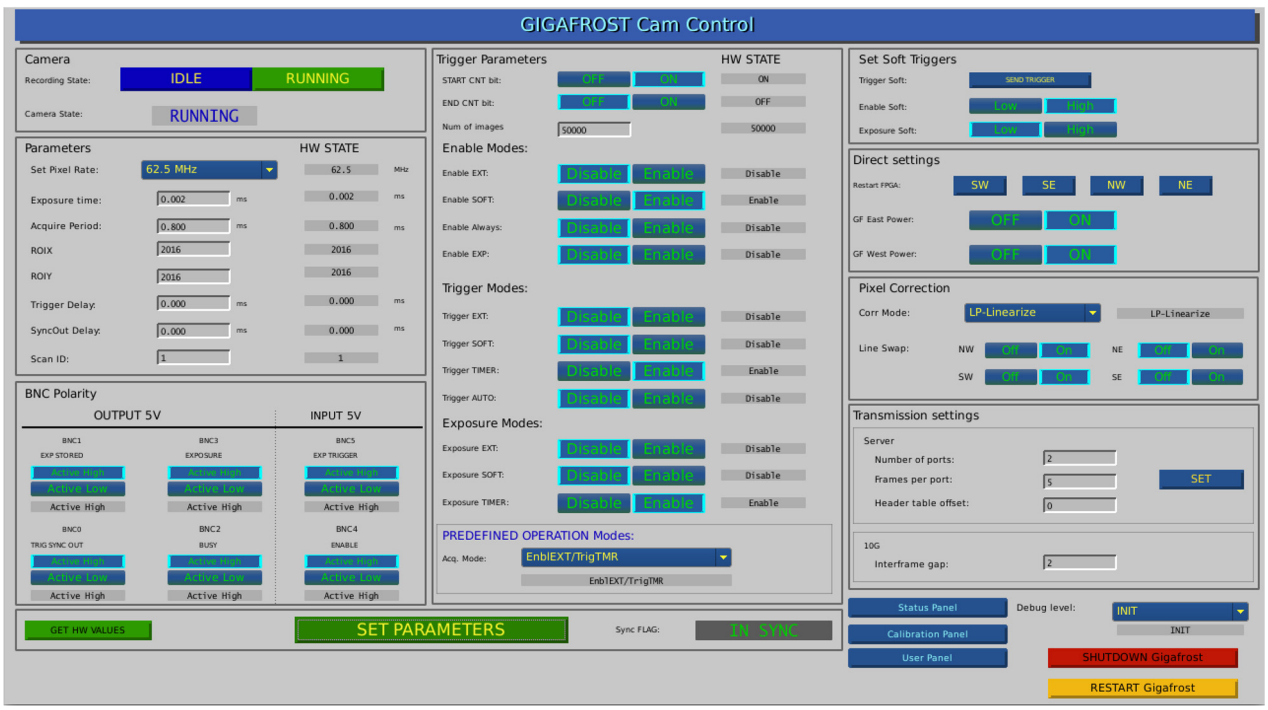

Figure 4: GigaFRoST expert control caQtDM GUI.

ISBN: 978-3-95450-193-9 doi:10.18429/JACoW-ICALEPCS2017-THDPL02

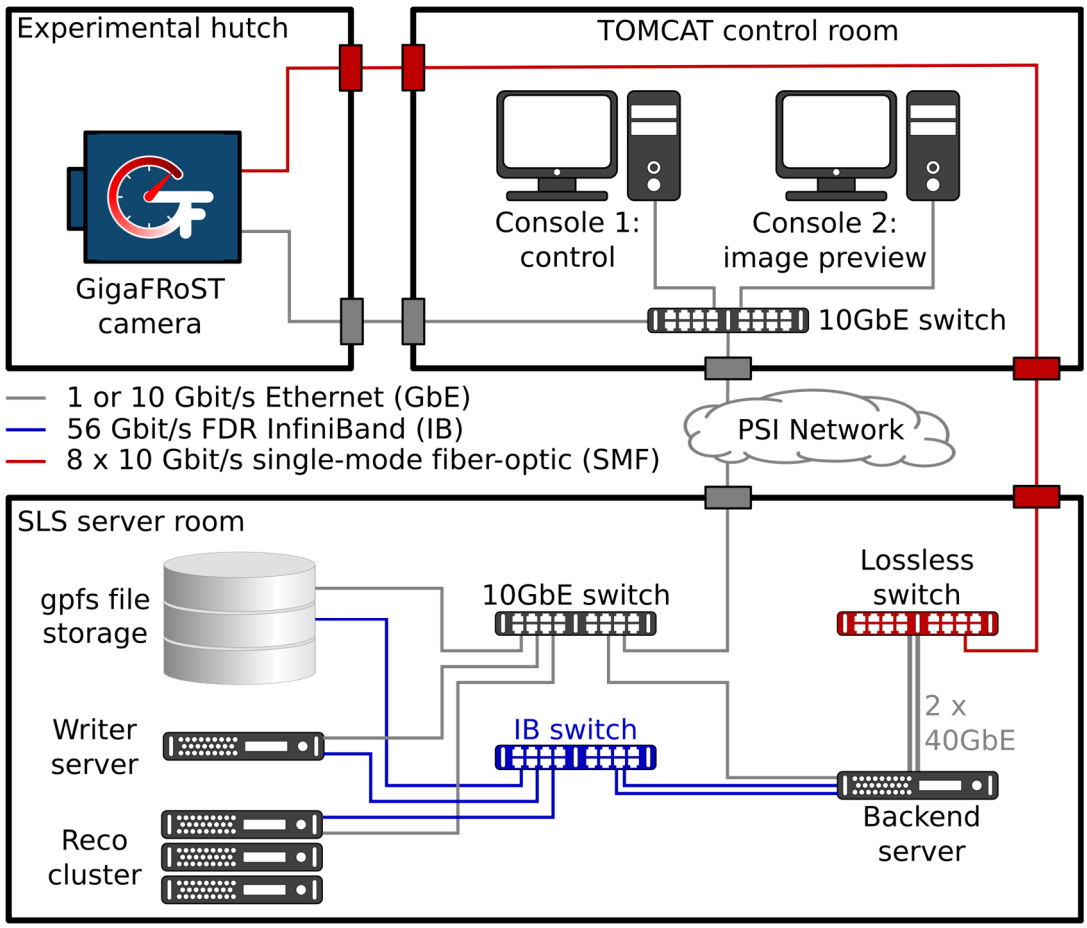

Figure 5: GigaFRoST network infrastructure and its topology.

File Writing and Storage

The writer process can subscribe to the ZMQ stream published by the backend server. Upon receiving the image and header data it packs them into a HDF5 files. Scientific Data Exchange format [7] is used for the file structure of HDF5 files. The central file storage server uses the General Parallel File System (GPFS) [8] and is connected to the beamline network with standard ethernet connections.

Limitations and Expandability

Currently the limiting factor in the entire data chain is the distribution of data from the backend server ring buffer to any downstream consumers. The throughput from the backend server is currently limited to roughly 2.2 GB/s. As a consequence, the ring buffer eventually saturates when using high data rates and starts dropping frames, limiting the amount of recorded data to the size of the internal server memory (currently 512 GB). This issue is currently being investigated and possible bottlenecks (e.g., specifics of a chosen ZMQ stream protocol message structure, internal memory bandwidth, process of reassembling the images) are being identified. Ideally, the writing speed would match the acquisition rate, limiting the amount of continuously taken data to only the size of available disk storage.

Due to the modularized design of system components, future improvements are relatively straightforward. Measures like distribution and parallelization of backend over multiple servers processing only subsets of images, splitting file writing over several processes or insertion of additional modules through subscriptions to the ZMQ stream (e.g., dedicated data compression) could be taken.

SUPPORTING BEAMLINE INFRASTRUCTURE

The GigaFRoST camera image acquisition is synchronized with the experiment setup using 6 BNC connectors with TTL logic. The sample is rotating on the fast air-bearing Aerotech rotation stage rotating at maximum velocity of 3600 deg/s (10 Hz). The rotation stage is integrated into the EPICS control system via a softIOC running on a PC with the Windows operating system with real time extension and communicating with the Aerotech controller over a firewire connection. The Aerotech controller provides a Position Synchronized Output (PSO) signal, that can be used to either enable or externally trigger camera acquisition based on the position of the rotation stage. Camera and/or the rotation stage could be triggered also by any arbitrary external device or signal

ISBN: 978-3-95450-193-9 doi:10.18429/JACoW-ICALEPCS2017-THDPL02

source, taking part in the experiment. The routing of the trigger signals between the camera and external devices is provided via a custom designed, FPGA based 'signal box'. This box was designed at PSI and is equipped with several BNC input and output connectors and the connector for Aerotech input/output signals. Input-output pairs can be configured dynamically using EPICS serial communication stream device driver. If required by the experiment an arbitrary delay can be set for each output as well.

CONCLUSION

A high-speed camera system was developed at PSI, able to continuously stream data at a rate of 7.7 GB/s, process it in real time (assemble the image), send them forward as a ZMQ stream and finally write the data to the final storage as HDF5 files. Additionally, a separate stream serving for a live image preview is provided. The GigaFRoST camera is integrated into the EPICS control system via an IOC running on the camera control board's PPC, where an embedded Linux operating system is running.

ACKNOWLEDGEMENT

I would like to acknowledge the work of people who greatly contributed to the development of the GigaFRoST system. Rajmund Mokso from the TOMCAT beamline for starting and coordinating the project as well as his successor Christian Schlepütz for continuing his work and successfully driving the project forward. Elmar Schmid and Gerd Theidel for their work on the HW design and camera firmware implementation. Heiner Billich for developing the backend server part and last but not least Simon Ebner and Thierry Zamofing for setting up the HDF5 file writing and ZMQ streaming infrastructure.

REFERENCES

[1] Tomcat beamline, SLS, PSI, Switzerland, https://www.psi.ch/sls/tomcat/tomcat
[2] R. Mokso *et al.,* "Advantages of phase retrieval for fast x-ray tomographic microscopy", *J. Phys. D: Appl. Phys.*, vol. 46, no. 49, 494004, 2013.
[3] PCO AG, www.pco.de
[4] EPICS, http://www.aps.anl.gov/epics
[5] ZeroMQ, http://zeromq.org
[6] R.T. Fielding *et al.,* "Principled design of the modern Web architecture", in *Proc. ICSE '00,* Limerick, Ireland, Jun. 2000, pp. 407–416.
[7] F. De Carlo *et al.*, "Scientific data exchange: a schema for HDF5-based storage of raw and analyzed data", *J. Synchrotron Rad.*, vol. 21, no. 6, pp. 1224–1230, 2014.
[8] F. Schmuck *et al.,* "GPFS: A shared-disk file system for large computing clusters", in *Proc. FAST'02,* pp. 231-244, Jan. 2002.