Advanced Real-time Evaluation and Data Quality Monitoring Model Integration with FPGAs for Tokamak High-performance Soft X-ray Diagnostic System

Andrzej Wojenski, Krzysztof T. Pozniak, Didier Mazon and Maryna Chernyshova

Abstract— Based on the publications regarding new or recent measurement systems for the tokamak plasma experiments, it can be found that the monitoring and quality validation of input signals for the computation stage is done in different, often simple, ways. In the paper is described the unique approach to implement the novel evaluation and data quality monitoring (EDQM) model for use in various measurement systems. The adaptation of the model is made for the GEM-based soft X-ray measurement system FPGA-based. The EDQM elements has been connected to the base firmware using PCI-E DMA real-time data streaming with minimal modification. As additional storage, on-board DDR3 memory has been used. Description of implemented elements is provided, along with designed data processing tools and advanced simulation environment based on Questa software.

Keywords— data quality monitoring, system modeling, FPGA, Verilog/VHDL, HDL, GEM detector, SXR plasma diagnostics, modular measurement system, data evaluation

I. INTRODUCTION

THE importance of the data quality diagnostic techniques, especially for the plasma physic, is often an simplified or neglected topic. The Author of the paper analysed several measurement systems working at the various tokamaks [1]–[7]. Many papers do not provide sufficient discussion about the validation of the input data, on the base of the output products are computed. Some of the systems provide information about the input data validation, gathered from the sensors. However, there is no common methodology along various systems. It should be also noted that the implementation of advanced DQM techniques and methods can increase the overall cost of the systems in both terms: development time and hardware costs. In order to design proper system architecture, a system under design should be at the very first stage analysed in scope of the input signals, their validation and final outputs. The methodology is often omitted, focusing mostly on data processing. Many of modern, advanced measurement systems works in real-time mode. This means, that any kind of the additional diagnostic needs to be:

• Suitable for work in real-time mode,
• Optimally designed.

The paper focus on the soft X-ray measurement system (GT-SXR), designed for work in real time mode with the feedback loop to control system of the tokamak. Upon results from the GT-SXR unit, the tokamaks’ actuators section can undertake necessary actions to keep the plasma. It very important that the decisions are based on valid data. The system is based on the GEM detector with 128 channel strip board (presented in Fig. 1).

![Schematic of the 128-channel GT-SXR measurement system](image-url)

Fig. 1. Schematic of the 128-channel GT-SXR measurement system [8]
Due to the proximity of the tokamak vessel, the analog stage of the system is designed to be rad-hard. The preprocessing stage is done in the FPGAs, where the raw signal is selected and streamed through PCI-E to the high-performance server unit. The system post-processes the data in two modes: offline and online mode. The energy and topology histograms with time resolution < 1ms are produced (10ms for real-time mode). More details regarding system construction can be found in [8]–[14]. Since the data will be read by tokamaks’ control system, it is especially important to provide high quality output products, based on correct input signals. For that purpose, in the following sections the Authors describes the adaptation of the real-time evaluation and data quality model for the implementation in the FPGAs of the GT-SXR measurement system.

II. REAL-TIME DATA QUALITY MONITORING MODEL OVERVIEW

Since the data quality monitoring discussion regarding the tokamak plasma experiments cannot be found in many publications, Authors of the paper proposed the model approach for input data evaluation and validation from the detector analog signal stage for the high-performance measurement systems based on the FPGAs. The model is divided into two stages:

- FPGA raw signal selection and preprocessing stage,
- Embedded high-performance real-time data processing algorithms.

The model is based on critical analysis of the most common elements used in FPGA and PC data processing firmware. The paper focuses on the FPGA stage of the system. The model of efficient data quality monitoring indicates the need of the proper input signals analysis and dynamic filtering of trigger. The signals can be marked as incorrect based on the Signal Classification (SC) units. The blocks consist description of the already known malformed signal shapes that can be expected at any time from the detector. In case if the signal is marked by SC, further FPGA stages can either allow to process it or discard it. Since probably many of signal malformation isn’t know yet (they aren’t many publications regarding GEM signal output related to hard environmental conditions), the model offers the implementation of data recording unit – technical sub-diagnostic, based on DDR3 memory, for further offline analysis. In order to have better knowledge on the topological aspects of signal malformation, there is also proposed a statistic sub-diagnostic unit, providing high time resolution histograms of local occurrences of SC flags or trigger events. Main components of the FPGA evaluation and data quality monitoring (EDQM) sub-systems are:

- Signal classification (SC) units,
- Technical sub-diagnostic,
- Statistical sub-diagnostic.

The complete EDQM model is presented in Fig. 2.

It is worth to mention, that the implementation of the model involves very small modifications of the base firmware layout, therefore it is easy to adapt in many projects, also after development of a data processing path.

The complete description of the model can be found in [15] (other EDQM analysis can be found also in [16] and [17]).

III. FPGA FIRMWARE EDQM MODEL ADAPTATION

The base firmware implemented in the GT-SXR FPGAs allows fast raw data acquisition in global trigger mode from all 128 GEM channels at once. The transfer is done through the PCI-Express Gen2 interface. Other important components in the
whole data selection and preprocessing path are:

- ADC interfaces based on SERDES transmission,
- Signal data adjustment (trigger shifting etc.),
- Fast triggering units,
- Fast online offset computation units,
- Global timestamp unit,
- Advanced external trigger connections between multiple FPGA board (currently working with 2 boards) MLVDS-based for:
  - Start of an acquisition,
  - Registering event for all FPGAs,
- Data aggregator unit,
- FSM for flow control.

In order to work with the PCI-E in DMA mode it was also necessary to adapt the common data bus for the system. The computation blocks are based on AXI4-Stream standard with additional buffering units. The separate PCI-E channel is used for slow control of the system based on the Wishbone standard. The complete schematic block is presented in Fig. 3. More details about the FPGA firmware can be find also in [18], [19].

Fig. 3. Basic FPGA firmware implantation for signal selection and preprocessing based on DMA PCI-E interface [15]

In the system are connected 2 FPGA units, in order to provide readout from 128 GEM channels. To provide consistent timestamp scale, both units must start the data acquisition at very same time. The current version of FPGA works in global triggering mode. Therefore, there is also a need to provide mechanism of advanced external trigger synchronization. This is necessary for registering at the same time (within difference up to 3 clock cycles) data events related to the photon pulses among all boards (detector snapshot registration). The external trigger information is based on MLVDS links, connected to FPGAs. In order to provide the best timing-fit between board, special algorithms has been developed for synchronization purposes. There are two synchronization stages:

1. Training of FPGA IDELAYE2 components to detect the center of the data window for MLVDS triggers. Time resolution is several ps. The system is based on pseudo-random patterns, based on the LSFR mechanism.
2. Flip-flops delay compensation in the system structure. Based on signal feedback pulse, the software algorithms detect the delay (in term of clock cycles) of the pulse among different FPGA boards. Based on gathered information, the algorithms then tune all other signals to match the slowest (most delayed) FPGA backplane.

The details about the advanced trigger synchronization, which is also an important part of the EDQM, can be found in [20].

An important benefit from EDQM model is ease of integration with the main FPGA data flow. In this case (as described in [15]), there was only necessary to add delay blocks on the main data path in order to implement advanced data quality analysis using DQM techniques. Additionally, simple AND gate has been implemented and combined with global trigger. This approach allows fast filtering or marking the data based on EDQM blocks decisions. In the FPGA model adaptation, all of the main EDQM blocks has been implemented, that is:

- Signal classification units SC with 4 signal models,
- Technical sub-diagnostic,
- Statistical sub-diagnostic.

Regarding the SC units, 4 signal models has been defined consider as a malformed signals or complex ones, needed to be processed in special way:

- Overflow (OVF) – the signal is out of ADC scale, therefore the signal information is lost and cannot be retrieved. This type of event needs to be rejected, since it contains incomplete information.
- Underflow (UDF) – the signal is at the zero level of ADC. This can be related to signals interferences or input channel malfunctions. This type of event needs to be rejected, since it contains incomplete information.
- Multi event (MULTI_EV) – overlapping (pile-up effect) pulses, in close time resolution. Related to intense photon flux. Needs to be computed by special CPU algorithms. In order to speed-up the CPU software, data can be marked, in order to lower the CPU usage.
- Long event (LONG_EV) – during intense photon fluxes, or multiple pile-up, the signal can be longer than one data frame (40 samples, corresponding to 500 ns). For the moment, there are no algorithms handling
In the Fig. 4 are presented the selected and implemented signal classifiers in the FPGA firmware for EDQM analysis.

The signals of the presented types were selected as for the first EDQM implementation approach, due to extensive experiments carried out at the CELIA laboratory with the first generation of the GEM measurement system. With intense photon flux, based on femtosecond laser pulses, those types of signals were very common (especially the 2 last ones), resulting in fact that it was impossible to register any kind of proper histogram.

However, at that stage, there was no possibility to either perform advanced EDQM analysis or even gather and time relate large quantities of raw data. More details about the experiments and their results can be found in [21]–[24].

The implemented technical sub-diagnostic allows to perform snapshots of the GEM detector, based on EDQM flags, that are in the field of interest. The raw signal is then recorded to the onboard DDR3 memory, for offline data analysis. The memory can store up to 100 000 events. The DDR3 data stream is completely independent from the fast PCI-E DMA main data stream.

Another sub-diagnostic implemented is the statistical sub-diagnostic. This is a very novel approach regarding the plasma measurements in term of high data quality. The sub-diagnostic is counting local EDQM flags activations, together with counting the local triggers that occurred. The analysis is performed on every channel. Additionality, the sub-diagnostic is designed for further use in a real-time mode, therefore it is possible to produce statistical histograms with 10 ms resolution or even less. Since there are no publications related this kind of analysis, the sub-diagnostic model will also be verified in scope of usefulness in term of the output histogram products.

This is the first implementation of such advanced and novel EDQM diagnostic for the GT-SXR measurement system. The complete integration of the described features is presented in Fig. 5. As stated before, the integration is based on very few elements, preserving the structure of the base firmware.
Since the whole concept is based on the reconfigurable signal classifiers SC, with additional use of 2 sub-diagnostic, it is possible to design a complete FPGA firmware with 2 main EDQM functions:

- Input data quality evaluation and filtering based on reconfigurable signal classification models
- Analysis of detector and system behavior in scope of continuous development of newly discovered signal classes based on registered sub-diagnostic data

The firmware with adapted EDQM model provides following data output types:

- PCI-E stream in DMA mode, working in real-time. The signals in the stream are either filtered out (only correct ones are passed further) or marked with appropriate EDQM flags
- DDR3 data storage with read-out using PCI-E DMA interface in offline mode (after measurement). Allows to download to measurement PC data containing EDQM marked events with high number of statistics and periodic statistical histograms

In previous generations of the SXR measurement system [25]–[28], this kind of data analysis and streaming were not available due to hardware construction and available FPGA resources.

IV. ADDITIONAL EDQM DATA PROCESSING AND ANALYSIS TOOLS

As mentioned at the beginning, the EDQM analysis of plasma measurements, especially in term of GEM detector analog signals, are rather poorly explored. The implemented firmware allows to perform standard measurements with additional EDQM discrimination and possibility of offline analysis of new events in form of raw data recordings with high statistics. In order to analyze the EDQM output of the system, several additional tools has been developed. The data processing engine for energy and topology spectra computation is done in Matlab software. The system can output raw data in form of binary files – one is the main real-time PCI-E DMA stream, second is an offline DDR3 readout.

The first set of tools, allows to filter the marked PCI-E data by various flags EDQM in offline mode and compute the output histograms. This allows to check:

- The statistics of occurrence of different flags during the event registration (this also depends of EDQM mode – mark but do not reject)
- The influence on spectra by rejecting different EDQM flags from the output products analysis

This allows to provide more precise image on how rejecting the certain type of data can influence the output algorithms products. The tools also perform EDQM conversion of measurement files from binary to MAT data format.

The second, important mode allows to analyze the output from the 2 sub-diagnostics.

The DQMSignalViewer GUI tools has been developed on the base of the base version of SignalViewer for visualization of the measurement data in scope of EDQM. The tool allows to visualize the analog signal registered from GEM, recorded in the DDR3 memory, based on active EDQM flags (with selection in the software). The event is selected upon user configuration regarding: channel range, event number (or range) and EDQM classification. The events analyzed in offline mode can reach up to 100 000 data packets. The use of this mode is especially important in term of discovery of new type of EDQM events and their influence on the neighbor channels. What is important, the signal shape can be directly visualized, and new SC model can be easily developed and included in new firmware revision with updated SC components.

Second mode allows to view histograms describing local statistics. The data can be viewed in high time resolution (10ms or less), specifying the events local activity on each channel regarding EDQM and triggering aspects. The data is visualized as a 3D plot with time relation. Currently it is possible to view local channels behavior with the following flags: active global triggers in the system, active local triggers with no EDQM flags, active local triggers with active EDQM flags, activity of following EDQM flags on each channel: OVF, UDF, MULTI_EV, LONG_EV. This mode is very important in term of GEM detector behavior during operation either in laboratory or on the tokamak site.

The fast, hard to detect EDQM events can be easily visualized and analyzed in scope of malfunction of the system or discovering a new type of events classes – also including new type of events – topology events. In addition, the raw data mode (technical sub-diagnostic) allows to observe the specified signals form along with their influence on the other measurement channels.

The screenshot from the DQMSignalViewer is presented in Fig. 6.

![Fig. 6. Example view of the additional EDQM data analysis software DQMSignalViewer in Matlab](image)

Other useful mode is a comparison of the histograms. Due to availability of the conversion tools BIN/MAT it is possible to produce the spectra from: base dataflow and EDQM mode switched on, with the same dataset. The software outputs 2 separate spectra plots (without and with EDQM), one plot when both spectra overlap and a plot with indicated percent in term of difference in computation of energy and spectra with EDQM enabled and disabled (both output products). In term of analysis of EDQM usefulness in certain aspects (e.g. for the systems with GEM detectors) this can provide additional output about which flags are especially influencing the output spectra.

V. ADVANCED EDQM FPGA VERIFICATION STAGE

As presented in Fig. 5, the complete FPGA firmware including EDQM is very expanded. Therefore, in order to verify the correct integrity of the EDQM and base data path, as
well as correctness of implemented SC models, it is necessary to perform complex FPGA tests. The verification of implemented firmware was done upon designing an extended and complex testbench module working with Questa simulation software. The testbench include the whole system environment, as implemented in real hardware, especially:

- FPGA backplanes with whole set of algorithms and input ADC channels (ADC interface replaced by simple model), including:
  - PCI-E DMA main data path (optimized),
  - EDQM data path with DDR3 memory (with further optimizations),
- MLVDS connections between FPGA boards with external triggers training algorithms,
- PCI-E switch component for distribution of algorithm start signal,
- The ADC unit simulating the detector part with ADC readout data path,
- Set of start-up commands to initiate the system based on configuration registers.

ADC modules implemented in testbench allows to work with the BIN files from the real measurement system and output the real analog signals on all systems’ input channels in testbench project under Questa simulation software. The use of the real data allows to especially verify the correctness of the implementation of the whole design with EDQM module and signal classifiers SC.

The testbench simulates both the base PCI-E DMA real-time data stream and separate offline DDR3 data from technical and statistical sub-diagnostics. The designed testbench was highly optimized, in order to process as many data as possible in a reasonable time, resulting in simulations based on high statistics. In current version, around 10 000 events are processed within 1 hour of simulation.

The simulation is based on the input data in form of raw signals like from the GEM detector. For that reason, the ADC-simulation block has been designed. It allows to source the BIN files from the real experiments done by the system with base FPGA firmware. Then, in the simulation the component produces the same signals and checks the behavior of the EDQM and base system implementation in FPGA. The components can also work with custom data in term of model signals.

The output from the simulation are the binary files, one representing PCI-E stream, second one representing DDR3 memory readout with EDQM technical and statistical data. The format is the same as from the real working system. Data can be directly sourced by Matlab software tools (described in section IV) to produce output energy and topology spectra. The basic data flow of the simulation engine is presented in Fig. 7.

The described functionality is unique for the verification of the measurement behavior when working with real sources of radiation. The significal advantage is that the simulations are based on high events statistics, more then 10 000 events.

Since it is possible to run the histogram algorithms based on the output of the simulation using real data, it is especially valuable in scope of testing implementation of the new signal classification models. Also, the influence on the output spectra can be verified, before implementing and updating the firmware in the real measurement system.

Fig. 7. Data flow in advanced EDQM FPGA testbench application using Questa simulator

VI. SUMMARY

The model of the EDQM technique in scope of FPGA units has been adapted and implemented for the GT-SXR measurement system. The implementation is designed to work with 2 FPGA boards including 128 measurement channels. For the beginning, 4 signal classifiers has been implemented: overflow, underflow, multievent, long-event. The EDQM model adaptation is complete, including the PCI-E stream monitoring, technical sub-diagnostic and statistical sub-diagnostic. For the sub-diagnostics, the DDR3 available on the FPGA boards has been used, resulting in space to store up to 100 000 raw data events / statistical histograms. Set of Matlab tools has been designed and adapted. One of the most important are converters from binary to Matlab files, for connection with Questa simulation software. Also, the algorithms in Matlab has been adapted to work with EDQM data format. The new DQMSignalViewer has been designed (based on previous version SignalViewer), providing functionality to easily analyze the output of the EDQM algorithms in scope of new type of events and correctness of the behavior of the system. The complete set of the verification stage has been also designed. The testbench in Questa simulation software replicates the whole system hardware with necessary links (e.g. between FPGAs) and components. The algorithms has been optimized to allow the simulations of large number of events. The simulation engine produces the binary data in the same format as real system with additional data from the DDR3 memory as separate file. Therefore, the simulations are directly linked with spectra computation. This mode is unique and can be used e.g. for verification of the correctness of the implemented signal classification model and their influence on the output products from the system.

As for the next steps, it is planned to perform EDQM FPGA firmware verification using the real physical data registered by the base firmware of GT-SXR measurement system. This stage will allow to check the usefulness of the EDQM technique, the possibility of registration the new types of signal malfunction or detection of already known ones, with their influence on the output products.
REFERENCES


