Abstract

The purpose of this paper is to present a new architecture design for a complex electronic system to be used in medical (imaging based) diagnosis. This work focuses a subsystem of the Data Acquisition Electronics (DAE) of a PEM (Positron Emission Mammography) system, namely, the Data Acquisition Read Out Controller (DAQ ROC). A preliminary version has been designed. A main objective of the work reported in this paper has been to increase system performance. To do so, a new strategy for memory management that leads to significant memory efficiency improvement and new structures for insuring synchronism in the communication busses have been implemented. The system is implemented in FPGA (Field Programmable Gate Array) technology. The robustness of the solutions has been ascertained with hardware validation. Results of test and validation on FPGA, boards and buses are presented.

1 Introduction

New product development in a tight time-to-market context poses huge challenges. In this context, power management, quality, costs and design productivity are key factors of success. High quality design and test (D&T) flows become mandatory and efficient methodologies and tools are required.

System design is usually bounded by a large set of constraints. In some cases, stringent timing requirements demand that significant parts of system functionality need to be implemented in hardware. In other cases (as in the present system), physical signals need to be sensed and converted into electrical signals, prior to digital processing. Synchronism and communication constraints are also problems that frequently traverse the entire system design process.

The work presented in this contribution is a subsystem of the electronic data acquisition and processing system of a diagnosis system for medical purposes, based on PET (Positron Emission Tomography) technology [1]-[3]. The Clear-PEM [1] characteristics that mainly constraint the electronic requirements are high data volumes and high data rates. These characteristics strongly influence system working frequency and communication structures.

The desired functionality is the identification of meaningful data out of the huge amount of data that come from sensor arrays. However, other constraints to system design come from the fact that we are dealing with an imaging system for medical diagnosis. This fact demands that results be obtained at the highest speed. It demands also that the correctness of the implemented functionality during system lifetime is guaranteed. This latter requirement demands the presence of robust Functional Built-In Self Test (FBIST).

The system described in this contribution is being developed within a research context. As a consequence, system requirements changes along the time, as well as, the reconstruction algorithms and calibration procedures. It is easy to understand that reconfigurable technology should be used in system implementation, since it allows low-cost system design adaptation. To use FPGA technology is, therefore, a natural choice. An additionally advantage of using FPGA technology is the need to implement BIST structures. Unallocated FPGA resources can be assigned (without additional hardware cost) to implement self-test functionality [4] - [6]. Test planning can also include chip, board and system level test [7], thus dramatically decreasing the costs of the test process and allowing cost-effective lifetime test.
The design methodology has been described in [8]. It has been adapted from the software domain [9], where this kind of modelling is used as a “thinking tool” to characterize the problem under analysis, as completely as possible, prior to initiate system design.

The paper is organized as follows. In section 2, a brief description of the PEM system is presented. In section 3, the DAE functionality is briefly described. Section 4 presents the overall DAE architecture. In section 5, the FPGA physical implementation is described. Validation results are presented in section 6. Finally, section 7 summarizes the main conclusions and future work.

2 PEM System

In order to understand the design methodology, as well as the design options, a brief description of the PEM underlying physics is presented in this section. The system is constituted by a scanner constituted by two planes of scintillating crystals that react with γ-ray photons emitted by the human body. When this occurs, crystals emit light that is afterwards converted into electric signals by Avalanche Photo Diodes (APD). Human cells emit γ rays when a radioactive substance is injected into the human blood stream. When this occurs, 2 photons are emitted in opposite directions over a linear trajectory. Cancerous cells radiate much more, due to their high metabolic rate (in comparison with normal cells) thus contrasting with a background radiation. Emission sources are detected by the intersection of trajectories [2]. Image reconstruction uses data originated in the PEM scanner, and identifies the presence of the radiation source (the cancerous cells).

Potential meaningful data correspond to the simultaneous detection of γ-ray photons in both crystal plans. This is the basic principle of the Clear-PEM scanner, a high-resolution PEM system, capable of detecting tumours with diameters down to 2 mm [1].

Crystal arrays in the Clear-PEM scanner are organized in modules and sub-modules in a hierarchical way [1]. Data are provided by 12288 readout channels (there are two channels per crystal, one for the top and one for the bottom planes). These channels are organized in 296 identical detector modules distributed by two crystal planes.

It is obvious that a complex data processing and data storage are required. The electronic Data Acquisition System (DAS) is constituted by two main systems, (1) the Front End (FE) electronics, an analogue/mixed signal system, responsible for capturing γ-ray photons emitted by the human body and for the first processing steps, the analogue/digital conversion and communication with the off-detector, (2) the Data Acquisition Electronics (DAE) system, the digital system responsible for digital signal processing and for sending relevant processed data to the computer where image reconstruction takes place. As mentioned before, the focus of this work is the DAE system, specifically one of its key modules, the Data Acquisition Read-Out Controller (DAQ ROC).

The DAE architecture maps crystal organization. Data reaches the DAE in parallel. Thus, it should, has much as possible, be processed in parallel. On the other hand, data are transmitted from the FE to the DAE by a large number of cables that may introduce diverse delays. This has lead to the development of stringent techniques to guarantee system synchronism.

3 DAE Functionality

Specific requirements to the DAE are as follows. The system should support a data acquisition rate of 1 million events per second, under a total single photon background rate of 10 MHz [10]. An event or hit (photoelectric event or Compton - according to the associated energy [1]) is defined as the interaction of a γ-ray with a crystal. These specifications have been defined using Monte Carlo simulation and mathematical models of the system [11][12].

Data to be analyzed and processed correspond to the energy originated in the different crystals as a consequence of these collisions. Relevant data is associated with relevant events. A relevant event is characterized by the simultaneous occurrence of hits in both crystal planes (229 bit/crystal).

Since huge amount of data require processing and physical memory resources are finite, an efficient memory management strategy must be implemented.

On the other hand, data is transmitted from the FE to the DAE by a large number of cables that introduce diverse delays. Data arrives with the corresponding clocks. This means that the clock phase may be diverse for the different channels. To guarantee that a detected coincidence is effectively a coincidence, it is mandatory to restore system synchronism.

4 DAE Architecture

In Figure 1, the top level diagram of the DAE architecture is depicted. A comprehensive description of the DAE system can be found in [13]. For the sake of clarity, a brief description of the constituting elements is provided here.

As can be observed, the DAE is constituted by 4 DAQ (Data Acquisition) boards and 1 TGR/DCC (Trigger/Data Concentrator) Board, communicating among them through buses. DAQ boards are mainly responsible for the acquisition and digital processing of data coming from the FE electronics. Each DAQ board has 2 (four million gates/FPGA) FPGAs.
The DAQ boards carry out the first data filtering to identify probably useful data, out of all the data that comes from the FE. Data is classified as probably useful when it may correspond to the emission of a pair of photons from a human cell. The criterion for this classification is the detection of coincidences, that is, the simultaneous interactions with crystal pairs within a given discrete time interval, to which an identifying Time Tag is associated. This Data is, afterwards classified either as useful and stored, or as noise, in which case it is discarded.

For the reasons presented before, main functionality is implemented using reconfigurable devices, namely FPGA technology. Each DAQ board houses 2 DAQ FPGA, each containing a read out controller - the DAQ ROC module. In [13] - [14] detailed descriptions of several design and test aspects are reported.

TGR/DCC board is responsible for the identification of coincidences. When a coincidence is detected, TGR/DCC generates a trigger signal that notifies the DAQ boards of the situation, picks up the corresponding data and concentrates it according to a given protocol, to be sent to the external PC for image reconstruction. This communication is based on a commercial Bus.

### 4.1 DAQ ROC Architecture

The preliminary architecture of the DAQ ROC is the one represented in Figure 2, hereby referred as v1. As far as memory management is concerned, different options are possible. A first option (see Figure 2) is as follows: memory write and read operations (the entire data package) are carried out in each single clock cycle. Each data package has 229 bit: (2 (APD) x10 (samples) x10 (bit) +17 (TimeTag Delta) +10 (Channel Id) +2 (Hit Type)). Therefore, memory width must at least 229 bit. This forces the partition of the available memory banks into 4 blocks, due to the available physical memory, although, the number of inputs is 8. With this approach, a complex multiplexing strategy for addressing the memories is required and has been implemented.

However, such large memories can not work at the specified 100 MHz frequency in the target FPGA device (Xilinx xc2v4000-4bf957 [15]). This forces the use of 2 clock frequencies, namely, 50 MHz (DAQ ROC module) and 100 MHz (DAQ and Filter), and the use of a frequency adaptation module.

This architecture has been implemented. However, in the DAE validation phase, problems have been detected in association with the DAQ ROC (v1) module, namely, large dead times associated with data transmission and reception and also with the arbiters, and deadlocks. These, associated with a relatively low efficiency of memory usage, contribute to performance degradation.

The above mentioned problems associated with the first architecture led to the need of developing the new architecture (v2) shown in Figure 3.

Main goals behind the development of the new architecture are to improve memory management efficiency and to solve the synchronism related problems in order to improve data integrity and system performance. Subsidiary goals are those indicated in Figure 4.

1. To use a single clock
2. To avoid input multiplexing
3. To improve processing time
4. To reduce dead times
5. To prevent deadlocks

As can be observed in Figure 3, the new architecture is composed by 5 main functional blocks, namely, GBus Controller, Memory and FIFOs.

All these blocks work in parallel, resulting in a reduction of transmission time to around 65 clock cycles.

Transmission time has been further reduced on another 6 clock cycles by refining the communication protocol. The refinement comes mainly from the reduction of redundant data. Robustness is guaranteed in the new protocol by the implemented error detection scheme, namely, by using voting schemes in critical transmissions and parity otherwise.

### 4.2 Synchronism Issues

In this system, a master clock controls the behaviour of the entire system. However, system functionality is distributed over different subsystems, each one behaving as a clock domain. The subsystems are separated by non negligible physical distance (see Figure 1). It is easy to understand that synchronisation problems may occur related with the communication channels.
In fact, communication problems occur on one side between DAQ and FE and on the other side between the different FPGAs.

In the first case, a control signal is used to restore synchronism between FE and DAQs, thus ensuring that data arriving from the FE is meaningful. This control signal is used to carry out pipeline dynamic self-adjustment, responding to delay variations.

In the second case, GBUS (see Figure 3) handles the communication process among FPGAs. Working clock frequency is the same for the various FPGA. However, each FPGA is, itself, a clock domain, that is, it is locally synchronous. GBus guarantees the communication among these clock domains.

Communication among multi-clock domains requires the resynchronisation of data associated with the different clocks. To achieve this purpose, several approaches may be followed [17]-[20], corresponding to synchronous or asynchronous buses. GBus tries to unite the best characteristics of the synchronous (speed) and the asynchronous (robustness) solutions.

GBus is a proprietary asynchronous Multi-Source, Multi-Drop, Multi-Master BUS, that is, ‘everybody’ accessing the BUS can write, can read, and can assume the GBus control (initiate the communication process).

Details of the implementation of the GBus structure can be found in [21].

4.3 Memory Management

The Memory is responsible for data storage and access. In architecture v1, 4 memory banks were accessed 2 in a time in every other clock cycle. In the new architecture, all 8 memory banks can be accessed in every clock cycle.

![New DAQ ROC Memory Structure](image)

Figure 5- New DAQ ROC Memory Structure

The memory blocks and FIFO, in Figure 3, include the respective controllers. As an example, Figure 5 shows the structure of the memory module, highlighting the 8 memory write controllers (one for each memory bank) and the single memory read controller.

Memory partition has been revised to comply with goals 1 and 2 in Figure 4. There is no advantage in writing or reading the entire package in a single clock, since data arrives in 10 (conditioned by the FE) clock cycles and is read in 4 clock cycles (conditioned by the GBus protocol).

Therefore, it has been decided to store each package in 4 clock cycles. In Figure 6 (a), the new memory map is depicted. The memory addresses are: (TimeTag) TTmsb 00 through TTmsb 11. The first position contains the type of hit (event or Compton), the TimeTag, the Delta, the (crystal) ID and sample data. This format is compatible with the GBus protocol (see Figure 6 (b)), thus avoiding later processing.

![GBus protocol](image)

Figure 6- (a) memory format; (b) GBus protocol

In architecture v2, 4 memory banks of 512x229 bit (1st architecture) have been substituted by 8 smaller memory banks of 512x62 bit. This avoids the use of complex multiplexing structures (Selector1 in Figure 2). However, this has been achieved with reduction of the actual physical memory (lower in architecture v2). With this new partition, it is possible to work at full clock speed (100 MHz). A memory bank for each data source guarantees that no data package is lost.

In Figure 3, FIFO structures are used to store the request signals, TGR FIFO, to store the memory address where single event data is stored, Sng FIFO, and to store output data (to GBus) Out FIFO.

The three FIFO operate in push mode. In this mode, the read controller converts the regular pull operation (Empty/Read) that characterizes the conventional FIFO behaviour into a push operation (Ready/Ack). By doing so, FIFO can use the dead time that corresponds to the processing, by the next processing block, of the previous data that has been read to make available the next data.

The pull protocol implements a data-on-demand approach. In such an approach, the availability of data requires that data is first requested (Read instruction), and then delivered to the data consumer. Hence, in a pull operation, data is available one clock after the read command is issued by the data consumer. In a push operation, there is no Read command, since Data is already available. Instead, a Ready signal is issued to the data consumer. After reading, the data consumer issues an Acknowledge (Ack) signal. By substituting the Pull protocol by the Push protocol, a cumulative saving of one clock per reading operation can be obtained, in most situations, in the overall data processing time.

---

1 Time Tag is the clock cycle associated to an event, and is used to identify the occurrence of a coincidence.

2 Delta is the clock phase associated with the maximum value of the energy. It is defined as the difference between the time instant associated with the energy peak and time instant associated with the highest sampled value.
5 DAE Physical Implementation

In Figure 7, the proprietary DAQ board is shown. As can be observed, the DAQ functionality has been implemented with 8 Xilinx™ xc2v4000-4bf957 FPGA (2 FPGA per DAQ Board), i.e., 4 million equivalent gates and 957 pins FPGA. The TGR/DCC functionality is implemented with one Xilinx™ xc2v3000-4bg728, 3 million equivalent gates and 728 pins FPGA [15]. This is the same platform that has been used for the implementation of the previous architecture.

The communication between the FE and DAE subsystems is carried out through LVDS (Low Voltage Differential Signalling) channel links (Figure 1). LVDS channel links de-serializer convert the high speed, serialised, long distance communication lines into the standard LVTTL (Low Voltage TTL (Transistor Transistor Logic)) electrical signals at the input/output of DAQ board’s logic components.

Transceivers are used as FPGAs gateways to the internal BUSes. Transceivers serve also as buffers for these BUSes.

The EPROM devices on each board store the FPGA’s configuration bit stream.

Figure 7 – DAQ Board Architecture

Proprietary buses have been implemented to establish communication among boards. This has been necessary to meet system requirements. Two main reasons forced this solution.

First, the need to transmit two different types of data packages, namely, one with 4 Bytes (event identifier) and another type up to 81 Bytes (event data) of information. The second reason is related with the required bandwidth (around 250MB/s).

6 Validation Results

The electrical test of the DAE constituting boards has been carried out prior to the FPGAs functional test. Some errors have been detected during the test of the FPGAs, mainly opens and bridges in the transceivers pin-outs. These errors have not been detected in the electrical test of the boards. In this case, Built-in functional test of the FPGAs and ChipScope™ made possible the detection of such errors.

The test of the PEM System Data Acquisition Electronics has been carried out on a system prototype. The strategy is as follows. First, the functionality of the FPGAs has been validated in silicon (DAE system and board test). Data in Figure 8 and Figure 9 are real experimental data, obtained in the validation process. Validation conditions are as follows. Input data (input events) are provided by a previously developed system that emulates the Front End (FE) functionality. Output data has been obtained with the DAE [13].

During this validation, some dynamic errors have occurred with the first version of the DAQ ROC architecture that were not foreseen by simulation. An example of such errors is illustrated in Figure 8. As it can be observed, GBus event rate for architectures v1 is linear until 400 Kevent/s but beyond this value, GBus transmission ceases due to pileup of out-of-valid data at GBus communication which is in turn caused by consecutive transmission delays built-ups. As mentioned before, the system was specified to work with 1M event/s.

Figure 8 – GBus event rate (architecture v1) (experimental result)

The lack of transmission rate and the fact that the entire system enters in a deadlock if input rate is over 400K event/s make v1 architecture unusable for final implementation.

In Figure 9, the GBus event rate for architectures v1 and v2 are represented.

Figure 9- DAQ ROC v1 and v2 GBus event rate (experimental result)

For v2 architecture, the GBus keeps working for the entire scope of the validation inputs. It is worthwhile to explain the shape of the curve obtained for v2 architecture. In fact, GBus is linear until almost 700
Kevent/s. After this value, starts to flare until reaching a maximum transmission rate of 800 Kevent/s. After that, it starts dropping. This is caused by the inefficiency of the arbiter. In the linear region, this inefficiency is absorbed by the fact that the required bandwidth is below the maximum.

This transmission rate has been considered acceptable for the current implementation, but expansion of the system will be limited (ex. larger crystal sensor array).

It is expected that with a more efficient arbiter, the linearity of the GBus is verified until higher values and remains almost constant after that, which will also lead to a wider possibility for expansion.

7 Conclusions and Future Work

In this paper, a new strategy for memory management and to deal with synchronism problems has been presented. The design methodology has been applied to the Data Acquisition Read-Out Controller (DAQ ROC) of the Data Acquisition Electronics (DAE) sub-system of the PEM (Positron Emission Mammography) system for medical image-based diagnosis.

By increasing the use of parallelism, processing times has been reduced and dead times have been reduced, thus improving system performance.

The novel architecture takes full advantage of FPGA technology. With this v2 architecture, effective memory capacity has been significantly increased, while the use of physical memory has been significantly reduced. Significant efficiency on memory usage and significant improvements in module performance have been achieved.

The new architecture has been validated in hardware, using a test setup structure developed to the effect, which includes a Front-End emulator. Functional Built-In Self Test solutions are now under development, in order to allow at-speed testing.

The use of FPGA technology to implement the major data processing functionality has been fully exploited for debug and test purposes.

Additional improvements of the DAQ ROC sub-system are expected with the review of the GBus arbiter, which is currently under development. Results will be reported in the future.

References


