Mitigated FPGA Design of Multi-Gigabit Transceivers for Application in High Radiation Environments of High Energy Physics Experiments

M. Brusati, H. Chen

Submitted to the Journal of Measurement

October 01, 2017

Physics Department
Brookhaven National Laboratory

U.S. Department of Energy
USDOE Office of Science (SC),
Basic Energy Sciences (BES) (SC-22)
DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or any third party’s use or the results of such use of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof or its contractors or subcontractors. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.
Mitigated FPGA Design of Multi-Gigabit Transceivers for Application in High Radiation Environments of High Energy Physics Experiments

M. Brusati\textsuperscript{a}, A. Camplani\textsuperscript{a}, M. Cannon\textsuperscript{c}, H. Chen\textsuperscript{d}, M. Citterio\textsuperscript{b}, M. Lazzaroni\textsuperscript{a,b,*,}\textsuperscript{*}, H. Takah\textsuperscript{d}, M. Wirthlin\textsuperscript{e}

\textsuperscript{a}Dipartimento di Fisica, Università degli Studi di Milano, via G. Celoria, 16, 20133 Milano, Italy
\textsuperscript{b}INFN Milano, via G. Celoria, 16, 20133 Milano, Italy
\textsuperscript{c}NSF Center for High Performance Reconfigurable Computing, Department of Electrical and Computer Engineering, Brigham Young University, Provo, Utah, U.S.A.
\textsuperscript{d}Brookhaven National Laboratory, Upton, New York, U.S.A.

Abstract

SRAM-based Field Programmable Gate Array (FPGA) logic devices are very attractive in applications where high data throughput is needed, such as the latest generation of High Energy Physics (HEP) experiments. FPGAs have been rarely used in such experiments because of their sensitivity to radiation. The present paper proposes a mitigation approach applied to commercial FPGA devices to meet the reliability requirements for the front-end electronics of the Liquid Argon (LAr) electromagnetic calorimeter of the ATLAS experiment, located at CERN. Particular attention will be devoted to define a proper mitigation scheme of the multi-gigabit transceivers embedded in the FPGA, which is a critical part of the LAr data acquisition chain. A demonstrator board is being developed to validate the proposed methodology. Mitigation techniques such as Triple Modular Redundancy (TMR) and scrubbing will be used to increase the robustness of the design and to maximize the fault tolerance from Single-Event Upsets (SEUs).

Keywords: Measurement, Instrumentation, Reliability, High Energy Physics, FPGA, Scrubbing

*Corresponding author
Email address: massimo.lazzaroni@unimi.it (M. Lazzaroni)

Preprint submitted to Elsevier January 28, 2017
1. Introduction

In high radiation environment commercial FPGAs are not a common choice, since their internal structure is particularly sensitive to the effects due to the passage of charged and uncharged particles. The change of state in a memory cell due to a particle strike is called Single-Event Upset (SEU) and is the main error source that has to be considered in a reliability analysis on FPGAs. This happens because FPGA architecture is composed of a large amount of memory cells in a very small circuit area, making memory protection a primary concern in these particular applications. Mitigation techniques have been widely investigated for this purpose in hostile environments to ensure the correct operation of FPGAs [1, 2].

Even though the sensitivity of FPGAs and memories (such as SRAM or DRAM) are very similar, they have very different failure modes. Typical memory cells store data and any upset will only corrupt that data. In FPGAs, however, the largest component of memory is used as configuration memory, which define the operation of the configurable logic blocks, routing resources, input/output blocks and other programmable resources. An upset in a configuration memory cell may thus change the operation of the user defined circuit.

The failure modes of FPGAs are also different from custom circuit technology – such as Application-Specific Integrated Circuits (ASICs) – where the design has a fixed functionality and does not need a configuration memory. Since the custom circuit is hard-wired, only latches need radiation protection techniques and the routing and logic are usually considered insensitive to soft errors [3]. FPGAs require SEU mitigation to ensure adequate reliability for many different resources (latches, logic, routing and I/O), since they are subject to alteration from SEUs.

The benefits of this technology are nonetheless very intriguing: the high throughput and parallel computing features are important elements in a high demanding application like the data processing of the ATLAS calorimeter.
The reprogrammability of the board is also an interesting feature as the circuit can be changed by simply loading a new bitfile: it allows firmware upgrades without the need to physically replace the electronics and also provides easy access for error correction, where errors can either be bugs in the original source code or faults due to radiation.

The present work is part of an ongoing collaboration that is based on previous studies on the HEP environment, FPGA reliability, resistance to radiation and mitigation techniques [4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15].

Serializers embedded in the latest generation of FPGAs present an interesting solution for high speed data transmission applications, but can operate unreliably in high radiation environments. In this paper we present an approach that allows the use of commercial off-the-shelf (COTS) FPGAs in a hostile environment (like the ATLAS LAr calorimeter front-end electronics), verified by creating a mitigated design of a multi-gigabit transceiver that will be installed on a validation board in the experiment.

The paper is structured as follows: the next section will describe the radiation environment expected for the LAr calorimeter front-end electronics, followed in Section 3 by a description of the damage caused by radiation to electronics and programmable logics in particular. In Section 4 the mitigation techniques proposed for this project will be introduced and in Section 5 the project layout will be presented. A brief discussion of the results is contained in Section 6 and finally in Section 7 the conclusions are drawn. A final Appendix is given as reference for radiation damage terminology and single event effects classification.
2. High Energy Environment

The LAr calorimeter of the ATLAS experiment is located at the Large Hadron Collider (LHC) in Geneva.\textsuperscript{1} LAr is a part of ATLAS experiment as can be seen in Fig. 1. ATLAS is a very complex apparatus and it is designed to precisely detect very high energetic events. It has observed a particle as massive as the Higgs boson, i.e. near 126 GeV [21].

This section is an overview of the experiment and the electronic readout structure, which is very complicated as well, since it has to process the detector’s output data at very high speeds. Particular attention devoted to the LAr electromagnetic calorimeter and the front-end electronics that are the settings of this project.

The ATLAS detector consists of three main sub-systems: the inner tracking system, the calorimeter, and the muon spectrometer.

Beams of particles from the LHC collide at the centre of the ATLAS de-

\textsuperscript{1}This section is mostly based on ATLAS collaboration articles, where more details can be found on every aspect of this detector’s development and operation [16, 17, 18, 19, 20]
Figure 2: Picture of a section of the ATLAS experiment: the different detectors are meant to measure different quantities of different particles [22].

tector making collision debris in the form of new particles, which fly out from the collision point in all directions. The different detectors placed around the collision point are designed to record the paths, momentum, and energy of the particles, allowing them to be individually identified. Figure 2 shows the different particles generated from a collision and their energy deposition in the various detectors that compose the ATLAS experiment. Each detector has its own electronic chain associated so that different information on the particle can be reconstructed (momentum, direction, energy, ...). The ATLAS experiment structure is explained in more detail in paragraph 2.1.

The present paper deals with the electronics for the barrel LAr electromagnetic calorimeter. The front-end electronics are placed near the LAr Electromagnetic Barrel (EC) to perform the first signal processing (such as preamplification,
Figure 3: In this picture the main structure of the ATLAS experiment is shown. We can see the inner detector, the calorimeters and the muon detector [16].

analog-to-digital conversion and serialization for transmission via optical fiber), and the crate is exposed to a high radiation level and magnetic fields that cause frequent failures in non-protected devices.

The magnetic field that affects ATLAS front-end electronics goes up to 0.1 T, and some care in the choice of the components must be taken to avoid any unwanted effects. Radiation exposure in the experiment is very consistent as well, and is composed of ionizing and non-ionizing particles, constituting a very hostile environment for electronics and in particular for programmable logic devices.

Some basic definitions for dealing with radiation damage are recalled in Appendix A, while a more extensive description of the effects of radiation on electronic devices is given in Section 3.

2.1. The ATLAS Experiment Structure

ATLAS is a general purpose detector meant for measuring collision products and their effects generated in the Large Hadron Collider (LHC) at the CERN
facility in Geneva.

Inside the LHC, bunches of up to $10^{11}$ protons (p) will collide 40 million times per second to provide 14 TeV proton-proton collisions at a design luminosity of $10^{34} \text{cm}^{-2}\text{s}^{-1}$.

The LHC will also collide heavy ions (A), in particular lead nuclei, at 5.5 TeV per nucleon pair, at a design luminosity of $10^{27} \text{cm}^{-2}\text{s}^{-1}$.

The ATLAS experiment is made of three main sub-systems: the inner tracking system, the calorimeter, and the muon spectrometer. It has a cylindrical symmetry and nearly $4\pi$ coverage in solid angle [16], which means that all possible directions of collision products are covered by detectors. As shown in figure 3, the closest part to the collision point of the experiment is an inner detector whose main objective is to track the particles that cross it. It is composed by the pixel detector, the semiconductor tracker and the transition radiation tracker. This inner detectors are surrounded by a 2 T magnetic field generated by a superconducting solenoid. The calorimeter system is composed of sampling detectors and is housed in one barrel and two endcap cryostats. The barrel is made of a LAr Electromagnetic Calorimeter (EC), surrounded by a hadronic calorimeter made of steel and scintillating tile (TileCal). In the endcap region of the detector all of the calorimetry uses LAr as the active material.

The EC is made of alternating layers of accordion-shaped lead absorbers and electrodes as shown in figure 4. Liquid argon is placed between these layers as active medium. Interactions in the absorbers transform the incident energy into a “shower” of particles that are detected by the sensing elements. The outermost level of the ATLAS experiment is composed of muon chambers, that can efficiently track the passage of a muon and measure its momentum.

Particles can be identified with complicated reconstruction algorithms that take into account the relativistic behavior of collision product. In figure 2 and table 1 various particles inside the different detectors are shown: electrons and

---

2Luminosity is the ratio of the number of events detected in a certain time to the interaction cross-section. It is then a measure of how many particles pass through a given area in a given time.
Table 1: Particle’s behavior in ATLAS detector.

<table>
<thead>
<tr>
<th>Particle</th>
<th>Detection and Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma$</td>
<td>no track in ID, energy release in ECAL (not in HCAL)</td>
</tr>
<tr>
<td>$e^\pm$</td>
<td>track in ID, energy release in ECAL (not in HCAL)</td>
</tr>
<tr>
<td>$p, \bar{p}$</td>
<td>track in ID, energy release in ECAL, HCAL</td>
</tr>
<tr>
<td>$\mu^\pm$</td>
<td>track in ID and MS, little energy in ECAL, HCAL</td>
</tr>
<tr>
<td>$\pi^\pm$</td>
<td>track in ID, energy release in ECAL, HCAL</td>
</tr>
<tr>
<td>$K^\pm$</td>
<td>track in ID, energy release in ECAL, HCAL</td>
</tr>
<tr>
<td>$n, \bar{n}$</td>
<td>no track in ID, energy release in ECAL, HCAL</td>
</tr>
<tr>
<td>$K^0_L$</td>
<td>no track in ID, energy release in ECAL, HCAL</td>
</tr>
<tr>
<td>$\pi^0$</td>
<td>two overlapping energy releases in ECAL</td>
</tr>
<tr>
<td>$K^0_S$</td>
<td>two $\pi^\pm$ tracks, displaced vertex</td>
</tr>
<tr>
<td>$\tau^\pm$</td>
<td>decays in lepton (tracks as $\ell, \mu$), or hadrons (jets in ECAL, HCAL)</td>
</tr>
</tbody>
</table>

ID: Inner Detector; ECAL: Electromagnetic CALorimeter; HCAL: Hadronic CALorimeter; MS: Muon Spectrometer.

Positrons ($e^\pm$) are charged and very light particles, so they leave an ionization track in the Inner Detector (ID) and release all their energy in a shower in the Electromagnetic CALorimeter (ECAL); photons are neutral, so they do not leave an ionization track but only release their energy in the ECAL; pions, protons, antiprotons and charged kaons ($\pi^\pm, p, \bar{p}, K^\pm$) are charged particles that interact with hadrons, so they leave a track in ID and also release their energy with hadronic showers in the Hadron CALorimeter (HCAL); neutrons, antineutrons and kaon K-Long ($n, \bar{n}, K_L^0$) are neutral and hadronic-interacting particles, so they only release their energy in the HCAL; muons ($\mu^\pm$) are charged particles that leave an ionization track in the ID and in the outer muon chambers (MS is for Muon Spectrometer), without interacting in the calorimeters.

The other particles listed in the table can be identified as well but their tracks are more difficult to recognize and specific algorithms have been implemented to this purpose.

To identify collision products one has to combine all data from the different
detectors and see which ones have recorded a signal. Additional information can be inferred too: total energy deposited can be reconstructed because it is proportional to the induced current on the electrodes, and momentum is extracted from curvature of tracks – caused by magnetic fields that generate a Lorentz force on charged particles.

![Diagram of detector layers](image)

Figure 4: On the right we can see the segmentation of the detector; the three layers have different granularities to allow a better reconstruction of the incoming energy. On the left the accordion structure of the barrel is shown. The top figure is a view of a small sector of the barrel calorimeter in a plane transverse to the LHC beams [18].

The experiment is under continuous development to measure more precisely the parameters of interest in order to effectively search for new physics phenomena and validate models for physics beyond the Standard Model. The upgrades of detectors and electronic are allowed only during interruptions of the particle beam; to access the cavern it is mandatory to wait for the radiation level to meet the required safety levels. Upgrades can be scheduled during long shutdown periods of the collider as well as in the Year-End Technical Stops (YETS) between December and February every year. An Extended YETS is expected this year from December 2016 to April 2017. Figure 5 shows an estimate for the timeline of LHC development towards High-Luminosity LHC, which will be active after 2022.
2.2. The Readout Electronics

The electronic readout of the ATLAS LAr calorimeters is divided into a Front-End (FE) system of boards mounted in custom crates directly on the cryostat (as shown in figure 6 and 7), and a Back-End (BE) system located in an off-detector underground counting house (called USA15) separated from the detector by approximately 70 m of optical fibers and other cables.

It is important to notice that FPGAs are currently used in the back-end electronics because USA15 is far enough from the detector to consider negligible the effects of radiation. This is why programmable electronics are currently used in this area and considered reliable. The front-end is instead made of ASIC chips that passed very strict radiation testing.

The architecture of the LAr readout system is shown in figure 8: the electronic signal coming from the detector is passed on to the front-end crate, where different boards are hosted. Given the stringent noise requirements, the LAr
FE electronic boards are placed in crates mounted directly on the calorimeter cryostat feedthroughs, circularly disposed around the calorimeter, as shown in figure 7.

The Front-End Boards (FEBs) are designed to amplify, shape, sample, pipeline, and digitize the calorimeter signals. The raw data from LAr are subject to some stages of analog processing: preamplifiers amplify the signal and split the output in three overlapping linear gain scales which are each subject to fast analog shaping. This is done to reduce the dynamic range requirements of the sampling and digitization stages. The shaped signals are sampled at the LHC bunch crossing frequency of 40 MHz by switched-capacitor array (SCA) analog pipeline chips which store the signals in analog form during the L1 trigger latency. If the event is accepted by the trigger, the samples are read from the SCA and
digitized through 12-bit ADCs.

The digitized data are formatted, multiplexed, and then transmitted optically out of the detector to the Readout Driver (ROD) in the back-end via a single 1.6 Gbps optical output link per FEB. Given the fine segmentation of the ATLAS LAr calorimeters, each FEB has to read 128 channels.

The FEB also implements the first two stages in the summing tree producing analog sums for the L1 trigger. The shaper chips sum their four input channels, then the plug-in Layer Sum Boards (LSB) mounted on the FEB performs another sum. The LSB outputs are sent off of the FEB to trigger boards mounted in the same crate, named Tower Builder board, which processes data for the L1 trigger in the back-end.

For configuration and operation, the FEBs require a number of external control signals: The 40 MHz LHC clock, as well as the L1 trigger signal and a few other signals synchronous with the clock, are delivered via the ATLAS Trigger and Timing Control (TTC) system. The FEB is configured and monitored via a dedicated “Serial Protocol for ATLAS Calorimeters” (SPAC) serial control
Figure 8: Block diagram of the architecture of the overall LAr readout electronics. The lower box depicts the cold electrical circuit in the calorimeter cryostat. The central box illustrates the functionality of the front-end boards located in the on-detector front-end electronics crates mounted on the cryostat feedthroughs. The upper layer shows the off-detector back-end electronics components (ROD boards and TTC modules) mounted in their readout crates, together with the LAr front-end tower builder electronics and the interfaces to the L1 trigger system with its central trigger processor (CTP) [16].
link, operating at 10 MHz.

Another very important feature to be considered is the cooling system: no fan can be used here because the magnetic field would interfere with their functioning, so the crate is provided with cooling plates besides every FEB. The coolant circulator is designed to provide water below atmospheric pressure at around 18°C.

Each water-cooled FE board has cooling plates mounted on both sides. The cooling plates are attached to the FE boards leaving a distance of 6 mm from the board surface to the cooling plate, establishing the maximum height allowed for electronic components on the board.

![Trigger Towers](image)

Figure 9: The energy depositions of an electron which carries energy of 70 GeV is illustrated for the Trigger Tower readout, which sums the energy deposition across the longitudinal layers of the calorimeters in an area of Δη × Δφ = 0.1 × 0.1 [23].

The amount of events is large and cannot be all stored and analyzed, so the ATLAS experiment has developed a trigger system that selects only the potentially useful data for storage, discarding other events and reducing the data rate. The selection of events is performed in three steps: L1 trigger is a hardware-based system that uses partial information from the calorimeters and the muon system to reduce the trigger rate to a maximum of 100 kHz. If the event passes L1 trigger, data is sent out for further analysis. The higher levels
of trigger are performed in the back-end and are software-based. The final goal is to reduce the data rate from 40 MHz to 200 Hz.

A new version (denoted as demonstrator) of the trigger board for the LAr calorimeter is under demonstration for the Phase-I upgrade. The purpose of these experimental tests is the study of the behavior of the electronics for the purposes of the design of a new card to be used during the Phase-I.

As can be seen in figure 4, the ATLAS LAr EM calorimeter is highly segmented. The calorimeter cells in each sampling layer have different size by construction. Now the L1 decision for the calorimeter is based on “trigger tower” signals which exploits only partially the granularity of the detector, as shown in figure 9. The trigger signal is created through several stages of on-detector analog electronics (see figure 9): in the shaper ASICS, in the Layer Sum Boards on the Front-End Boards and in the Trigger Tower Boards installed on detector. Typically, a total of 64 read out channels are summed to form an EM trigger tower.

In the Phase-I upgrade of the LAr trigger system a finer granularity is going to be used, based on a “Super-Cells” system. This scheme is shown in figure 10. With this configuration the transverse granularity is increased, since the
resolution/granularity in \( \eta \)-axis \( \Delta \eta \) is decreased from 0.1 to 0.025 in the front and middle layers, as shown in the picture. The readout electronics must match this increased data rate and in the future upgrade it will perform the super-cell sum instead of the tower builder.

Figure 11: Schematic diagram of the architecture of the readout electronics after the Phase-I Upgrade. New components are indicated by red-outlined blocks [20].

The schematic of the LAr calorimeter readout with the upgraded trigger is shown in 11. The old scheme is still used during the demonstration so that backward compatibility is ensured and will also be kept during Phase-I but at some point decommissioned. The new boards are indicated by the red outlines and arrows on the diagram. The analog signal sums for the trigger are modified such that the sums for Super-Cells will be formed first, then further summed into trigger towers for the current L1 trigger function. The new components include new layer sum boards on the existing front-end boards to produce the Super-Cell signals for each layer and new baseplanes in front-end crates for
analog signal transfer.

New LAr Trigger Digitizer Boards (LTDB) receive and digitize analog Super-Cell signals and then transmit them to the back-end. There the LAr Digital Processing System (LDPS) extracts transverse energy values for each Super-Cell at every bunch crossing and transmits this information to the L1 calorimeter trigger system.

With the upgraded design, each LTDB will process up to 320 Super-Cell signals. The signals will be digitized at 40 million samples per second by 12-bit ADCs on the LTDB after one additional stage of shaping. A total of 124 LTDBs will be needed, delivering 25 Tbps digital information to the LDPS. The LDPS is in the back-end and houses FPGAs for dealing with the huge amount of data transmitted by the LTDB. The LDPS is designed to receive the digitized data at a total rate of 25 Tbps, perform digital signal processing in real-time and transmit the processed data to the L1 calorimeter trigger at a total rate of 41 Tbps.

As can be seen in figure 11, the LTDB board must serialize the digitized data from ADCs with a MUX/serializer block. In the board now working as demonstrator this was realized using FPGAs instead of ASICs, because in this way the prototyping can be done more easily. The configuration that is used is not mitigated, so even if there is no severe malfunctioning detected yet, there is no insurance that the electronics is not experiencing errors.

There are currently two LTDB demonstrator boards in one crate of the calorimeter: one was developed at BNL and the other at LAL (Laboratoire de l’accélérateur linéaire - Saclay), shown in figure 12.

More details on the LTDB board are provided in Section 5.1

The LTDB demonstrator board developed at BNL has been the starting point for the project described in this manuscript, since the board already hosts FPGAs and is fully configured to fit the FE crate specifications. The group has been providing all the material and support that we needed to modify the board for our purposes. This contribution has been very time saving because
it allowed us to reuse a project that had already been adopted in the ATLAS experiment.

3. Radiation Effects on Electronics

The radiation exposure of electronics in the ATLAS experiment is very consistent and is composed by both ionizing and non-ionizing particles, constituting a very hostile environment for programmable logics.

The expected background for the LAr calorimeter electronics for Phase II runs are shown in Table 2, while Fig. 13 shows a simulation of the background due to a mixed field of hadrons, electrons and photons. Phase II background will be $\sim 10$ times higher than present.

This section deals with the damage that electronics, and in particular FP-GAs, experience when immersed in a high radiation environment. Circuit damage caused by radiation due to Single-Event Effects (SEE) will be considered. When a single particle strikes the area of a circuit, the system can change state or alter its performance [24, 25, 26]. In this case we call the semiconductor's ra-
Table 2: Expected background for the LAr electronics for Phase II runs.

<table>
<thead>
<tr>
<th>Ionizing Dose</th>
<th>Simulation (one year)</th>
<th>Safety Factor</th>
<th>Test Target (10 years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 MeV eq. Neutron</td>
<td>6.0 · 10^{11} cm^{-2}</td>
<td>2</td>
<td>1.2 · 10^{13} cm^{-2}</td>
</tr>
<tr>
<td>Hadrons (&gt;20 MeV)</td>
<td>8.5 · 10^{10} cm^{-2}</td>
<td>2</td>
<td>2 · 10^{12} cm^{-2}</td>
</tr>
</tbody>
</table>

* 1 LHC year = 10^7 s, σ_{pp} = 80 mb.

Luminosity = 5 · 10^{34} cm^{-2}s^{-1};

Radiation response a Single-Event Effect (SEE). SEE include many different types of errors that are explained in Appendix A.

A further classification divides errors into soft and hard ones. Soft errors are nondestructive functional errors that do not cause a permanent damage to the material. Soft errors are a subset of SEE, and include Single-Event Upset (SEU), Multiple-Bit Upset (MBU), Single-Event Functional Interrupt (SEFI), Single-Event Transient (SET) and Single-Event Latch-up (SEL). These errors are usually associated with erroneous output signals from a latch or memory cell that can be corrected. On the other hand hard errors are irreversible: the device is permanently damaged, data is lost and original operation is no longer recovered even after power reset and re-initialization. Hard errors can be caused by Single-Event Gate Rupture (SEGR) and Single-Event Burnout (SEB).

The radiation that strikes a circuit node will not affect the proper operation if the charge disturbance stays within noise margins. Otherwise, the induced charge can be interpreted as the opposite logic value, generating a malfunctioning in the circuit. If the glitch generated by the passage of radiation strikes a node in a memory cell, it will store the wrong bit until data is written again [27]. Figure 14 shows an SRAM cell (Static Random Access Memory) that ex-
Figure 13: Simulated spectrum of particles for the ATLAS LAr calorimeter electronics.

...experiences an upset: the word line (WL) is low, so the cell is holding the stored data with the two latched inverters. Two NMOS transistors are used to decouple the bitlines BL and BL#. If the particle strike changes the logic level of node 'A', then the inverter will propagate the wrong value to node 'B', which, in turn, drives 'A' with the wrong value. The latch structure itself is causing the storage of a flipped value, and once it happens recovery can be achieved only by rewriting the state via bitlines.

The next paragraph 3.1 explores the architecture of SRAM-based FPGAs explaining why they are so sensitive to radiation. In this project we used a Xilinx device, more precisely a Kintex-7 FPGA (xc7k325tffg900). In paragraph 3.2 is a resume of all the radiation testing done on the evaluation board of this device.
3.1. Radiation Effects on FPGAs

Protection against SEE is important in every radiation tolerant device, but FPGAs are particularly sensitive when exposed to this kind of environment. This is intrinsic in the architecture of Field Programmable Gate Arrays as will be explained in this section. The 7 series family of commercial Xilinx FPGAs – based on 28 nm feature size and on Silicon-On-Insulator (SOI) technologies – are more tolerant to radiation, which is why they have been considered to be used in front-end electronics.

The technologies used for the storage of the configuration bits generate three main categories for FPGAs: anti-fuse, flash-based and SRAM-based FPGAs.

Anti-fuse FPGAs use one-time programmable fuses to permanently set the state of each FPGA configuration bit. It is nonvolatile and configuration memory occupy a small area, but this kind of device cannot be re-programmed since configuration data is fixed once the fuses have been configured.

Flash-based FPGAs use flash memory cells to set the state of the FPGA configuration memory. Flash memory cells use an electrically isolated floating gate to store the state of the memory cell. They are nonvolatile and reprogrammable a limited number of times, due to the degradation of oxide layer during program-erase cycles (most commercially available flash products are guaranteed to withstand around 100,000 program/erase cycles). These memory cells are generally immune to SEUs but are more sensitive to total ionizing
Figure 15: Architecture of a general FPGA.

dose than SRAM-based FPGAs, so they can be used under radiation only for a limited amount of time.

SRAM-based FPGAs are the most widely used type of configurable devices because of their infinite re-programmability and the use of standard CMOS process technology. Memory cells like the one of figure 14 store the configuration bitfile of the device. No additional circuitry beyond standard transistors is required and the latest CMOS technology available can be used. They can therefore benefit from the increased integration, the higher speeds and the lower dynamic power consumption of new processes with smaller minimum geometries [28]. SRAM cells have nonetheless some drawbacks, such as their size and volatility. The first one is mainly due to the number of transistors required for building a memory cell (at least six), while the second implies the use of an external device that stores the configuration data when the FPGA is powered down. Dedicated circuitry is then needed to initialize all the SRAM bits at
Table 3: CLB resources of some Kintex 7 devices.

<table>
<thead>
<tr>
<th>Device</th>
<th>Logic Cells</th>
<th>Slices</th>
<th>6-input LUTs</th>
<th>Distributed RAM (Kb)</th>
<th>Shift Register (Kb)</th>
<th>Flip-Flops</th>
</tr>
</thead>
<tbody>
<tr>
<td>XC7K10T</td>
<td>65,000</td>
<td>10,250</td>
<td>41,000</td>
<td>838</td>
<td>419</td>
<td>82,000</td>
</tr>
<tr>
<td>XC7K325T</td>
<td>325,000</td>
<td>50,050</td>
<td>265,800</td>
<td>4,000</td>
<td>2,000</td>
<td>477,000</td>
</tr>
<tr>
<td>XC7K490T</td>
<td>477,700</td>
<td>74,020</td>
<td>289,000</td>
<td>6,788</td>
<td>3,284</td>
<td>597,000</td>
</tr>
</tbody>
</table>

power up by reading the external memory.

In this project commercial SRAM-based FPGAs were used in order to prove that with proper mitigation even commercial FPGAs can be made resistant to radiation. The following considerations are all referred to SRAM-based FPGAs.

In figure 15 is shown a general architecture of an FPGA: it is composed of a regular array of Configurable Logic Blocks (CLB) that contain the user defined algorithm, connected to each other and to the input/output resources through switch matrices. In this way the routing between used resources is as general as possible, and the specific user design is built by a proper programming of the switches. It is to be noted that routing is transparent to the user, which means that usually the interconnection path is automatically handled by the design software.

Table 3 shows the CLB resources of the device that was used during this project. Each 7 series FPGA slice contains four LUTs and eight flip-flops, and each CLB is formed by a pair of slices [29].

There are two very important uses for SRAM cells: they are used to control the switch matrix (as a select line for multiplexers) and to store the data in the lookup-tables (LUTs) that are typically used in the CLBs.

It is clear then that the information needed to build a user-defined circuit is all contained in these SRAM cells, composing the configuration memory. A large amount of static memory is needed to store the configuration of the FPGA, because it carries the information that defines the function of the programmable logic units, routing switches, internal memory, and other special-purpose structures.

This architecture is very dangerous in a high radiation environment, because if a particle strikes a transistor it can change its state resulting, for example,
in opening a switch that was originally meant to be closed. This can lead to functional change in a logic module, mis-connected or misrouted signals, eventually resulting in a system failure as shown in figure 16.

If, for example, the design is a simple logic function like an AND gate, the VDHL code for it would be the following:

```vhdl
Listing 1: AND gate code

entity and_gate is
Port ( A : in STD_LOGIC;
      B : in STD_LOGIC;
      OUTPUT : out STD_LOGIC
    );
end and_gate;

architecture Behavioral of and_gate is
begin
  OUTPUT <= A and B;
end Behavioral;
```

The software will synthesize this code into a schematic like the one in figure 18. Notice that the FPGA does not use a CMOS AND gate, but implements
Figure 17: On the left is the initial AND gate with the corresponding truth table contained in the LUT. If a SEU occurs and changes the value of the truth table — as shown on the right — the logic function changes completely, becoming an XNOR gate.

the function with a Look-Up Table (LUT). In this way if the user wants a different logic function the software only updates the truth table within the LUT, without having to change the CMOS configuration into the new desired gate. We see that this is a very useful feature for FPGAs, because they can build any function with the same internal configuration, but is very hazardous when used in a high radiation environment. If an upset occurs in the memory that decides the LUT that is in use, it will change the logic function and result in an error, as can be seen in figure 17.

Figure 18: Schematic generated by the synthesizer software from the code given in 1. The software automatically assigns buffers for the input and output pins and uses a Look-Up Table (LUT) to implement the desired function.
There are different components within an FPGA which are subject to SEU and need to be protected with particular techniques. The most important architectural components that are affected by radiation are configuration memory, block memory and user flip-flops.

- Configuration memory: a large amount of memory cells are used to define the operation of the CLBs, routing resources, I/O blocks and other programmable FPGA resources. Upsets in these static memory cells may change the operation of the circuit and actually alter the user design.

- Block memory: this user memory is devoted to many different functions such as data storage, buffering or FIFO. Upsets in the BRAM can be corrected using error correction codes to prevent wrong data to be propagated in the system.

- User flip-flops: these components are used to implement sequential logics such as state machines, counters and registers. Upsets in flip-flops may alter the state and output of the circuit.

The effect of radiation on each of these internal resources depends on the specific device used and on the implemented design.

It must be noted that not all configuration bits are important in the same way in the design; some of them will directly affect the behavior of the circuit when upset while others will not. We will call the first type of configuration bits “sensitive”, while the latter one will be called “insensitive”. Sensitivity depends on the specific design, as it is related to the location and number of configuration bits used for the design. There are a variety of ways to check if a bit is sensitive to SEUs, but the most common are radiation testing and fault injection. The latter is by far cheaper and faster, and has been shown to be a good solution. In [31] the fault injection tool predicted approximately 97% of the output errors found during radiation testing.

We are interested in measuring the failure occurrences of the system, by means of a diagnostic procedure that allows a further distinction into different
modes of disruption (failure modes). We can indeed further classify errors into persistent and nonpersistent ones, as was done in [9].

The distinction is based on the duration of the failure: if the system restores its correct functionality after some time, the error is classified as nonpersistent and the upset bit is stored as a nonpersistent one. If the SEU causes a permanent error in the design performance which can be restored only through a global system reset, the faulty bit is treated as persistent instead.

3.1.1. Nonpersistent Errors

Nonpersistent errors are only temporary errors which are fixed after the scrubbing procedure. This technique rewrites the correct configuration memory on the device, as will be discussed in the next section. We can see the behavior of such an error in Fig. 19: the Device Under Test (DUT) calculates the square number of the input, but when an SEU occurs its output becomes incorrect. The third sequence of the figure shows the arithmetic difference between the expected result and the DUT output. As long as the faulty bit produces an incorrect output the difference is not zero, but after configuration scrubbing occurs correct operation mode of the upset circuit is restored and the sequence is null again. This example shows that nonpersistent behavior appears in circuits where the output depends on the present value of the inputs, so that the correct functioning is restored when the configuration memory is repaired by scrubbing.

3.1.2. Persistent Errors

If alteration of a bit value produces an error that will indefinitely propagate in the circuit even after scrubbing has restored the original value, this bit is classified as persistent. A correction of this type of error can be made only through a system reset.

These errors occur because the scrubbing procedure restores the circuit structure, but the temporary failure may have generated a faulty state which will not be corrected until a system reset. Persistent errors are thus caused by upsets in the configuration memory corresponding to feedback loops and internal state
storage, which receive a faulty state and maintain it until a reset occurs.

An example of a circuit vulnerable to persistent errors is a binary counter: the next value is computed from the current one, so the output depends on the previous state in the counting sequence. If a SEU damages the counter and the output does not change (as shown in Fig. 20), this output will be wrong even after the scrubbing procedure. Correction of the faulty bit assures the restoration of correct operation but the counter will start from the wrong state, losing the correct sequence.

Figure 19: Example of a nonpersistent error in a squarer.

Figure 20: Example of persistent error in a binary counter.
It is clear that the electronics used in the experiment must be protected in order to operate without performance and reliability degradation. In applications such as the ATLAS FE electronics reliability is an important issue, because the boards can be physically accessed only during shutdown time, and remote maintenance must be provided in order to avoid complete data loss.

Direct radiation testing is an important validation step for the chosen device and implemented design, in order to understand the effective sensitivity of the system.

3.2. Radiation Tests on Xilinx Kintex-7

Some radiation testing prior to this project have been conducted on FPGAs and their embedded transceivers. In the last four years the Milan group of Istituto Nazionale di Fisica Nucleare (INFN), Brookhaven National Laboratories (BNL) and Brigham Young University (BYU) collaborated in various experiments to evaluate the radiation tolerance of FPGAs to the mixed radiation field of HEP experiments [11, 32, 33]. The tested FPGA is a Xilinx Kintex-7 (part number XC7K325T-FFG900), that has also been used for the demonstrator LTDB board.

As previously said, the 7 series FPGAs are one of the latest generation of commercial FPGAs available on the market, based on 28 nm feature size and SOI technology. The Kintex-7 family is optimized for best price-performance, providing a high signal processing capability and low power consumption.

To test the memory resistance to radiation and the transceivers reliability, different radiation sources were used on various designs. The benefits of mitigation techniques such as scrubbing and Triple Modular Redundancy (TMR) were tested as well – these techniques are thoroughly discussed in chapter 4.

The Kintex7 device was tested in the following radiation environments:

- **Hadrons and neutrons**: this test took place in 2012 at H4IRRAD, a facility developed for testing the electronic equipment for LHC [34]. The H4IRRAD mixed-field is very similar to the LAr radiation environment,
but presented a large uncertainty on fluence and ionization dose. At this time BRAM and the configuration memory only were tested.

- **Neutrons**: testing with neutrons have been performed at LANSCE (Los Alamos Neutron Science Center) in 2012 and 2013 [35], and at TSL (The Svedberg Laboratory in Uppsala) [36] in 2013. The maximum energy of the wide-spectrum neutron beam is 800 MeV at LANSCE and up to 200 MeV at TSL. The first test at Los Alamos measured the sensitive cross section of configuration and BRAM memory to a wide-spectrum neutron beam, and a periodical readback was implemented to check for upsets. The campaign at TSL used the ANITA wide-spectrum neutron facility and validated the results from LANSCE, as well as verify a preliminary dynamic configuration memory scrubber. The last test performed at LANSCE verified the use of a JTAG scrubber and studied the Multi-Gigabit Transceiver (MGT) sensitivity. The improvement on radiation resistance through TMR techniques was investigated as well.

- **Protons**: the TSL PAULA facility [37] was used to test the sensitive cross section of the CRAM and BRAM with high-energy protons (180 MeV) in 2013, while in 2014 the use of the GTX – the embedded transceivers of Xilinx 7-series FPGA family – was extensively investigated [11]. In the latest test an external scrubbing was used and data on the type of communication errors was collected.

Results of the radiation campaigns are displayed in table 4. The cross section measurement for Block RAM and Configuration RAM are computed with the following equation:

$$\sigma = \frac{N_{\text{events}}}{\phi \cdot N_{\text{elements}}}$$

(1)

where $N_{\text{events}}$ is the number of upsets in the architectural elements under examination (CRAM or BRAM), $\phi$ is the particle fluence of the test and $N_{\text{elements}}$ is the number of architectural elements within the device under test. For the Kintex7 325T, $N_{\text{CRAM}} = 67,930,000$ and $N_{\text{BRAM}} = 16,404,480$. 
Table 4: Cross section comparison [33].

<table>
<thead>
<tr>
<th></th>
<th>CRAM [cm²/bit]</th>
<th>BRAM [cm²/bit]</th>
<th>Fluence [particle/cm²]</th>
</tr>
</thead>
<tbody>
<tr>
<td>H4IRRAD₁ (Hadron)</td>
<td>1.50 \cdot 10^{-14}</td>
<td>1.40 \cdot 10^{-14}</td>
<td>1.8 \cdot 10^{9}</td>
</tr>
<tr>
<td>LANSCE² (Neutron)</td>
<td>6.89 \cdot 10^{-15}</td>
<td>6.15 \cdot 10^{-15}</td>
<td>5.7 \cdot 10^{10}</td>
</tr>
<tr>
<td>TSL³ (Neutron)</td>
<td>6.55 \cdot 10^{-15}</td>
<td>-</td>
<td>&gt; 5.7 \cdot 10^{10}</td>
</tr>
<tr>
<td>TSL³ (Proton)</td>
<td>8.29 \cdot 10^{-15}</td>
<td>8.19 \cdot 10^{-15}</td>
<td>1.13 \cdot 10^{13}</td>
</tr>
</tbody>
</table>

₁ H4 normalization has an error of 50%.
² LANSCE normalization has an error of 10%.
³ TSL normalization has an error of 10 – 15%.
⁴ during proton irradiation more than 300 krad of ionizing dose was deposited on the DUT.

The measures taken at LANSCE are also consistent with the Device Reliability Report from Xilinx, that published test of various devices in the same laboratory: their neutron cross-section per bit for the 7 series Kintex CRAM is 6.99 \cdot 10^{-13} \text{ cm}^2/\text{bit} and for the BRAM is 6.32 \cdot 10^{-13} \text{ cm}^2/\text{bit} [38].

No functional interrupts (SEFI) were observed during any of these tests. A moderate increase in core current was observed while testing with protons, but it did not exceed the specifications. The current draw was significantly reduced when scrubbing techniques were applied but it could not be restored to the initial value. This suggests that radiation at least partially affected the configuration memory causing an activation of unused portions of the circuit and producing an overall increase of power consumption.

During the neutron preliminary test on the use of MGTs only two GTX lanes at 5 Gbps were used and 6 links failures were observed. In the proton testing a more refined design was used for the GTX reliability evaluation: 13 bi-directional GTX lanes at 3.125 Gbps were continuously monitored. The measured error cross section for the GTX transceiver with configuration scrubbing but no TMR is 7.27 \cdot 10^{-19} \text{ errors/lane}/\mu\text{CM}^2, with a lane upset rate of 5.06 \cdot 10^{-8} \text{ upsets/lane/s} or once every 229 days. The total number of lanes used in the calorimeter is 2 \cdot 10^4, meaning that in a full-scale calorimeter system the transceivers would experience an upset every 229/20000 = \approx 11 minutes.

Moreover, the conclusions drawn for a specific device of 7 series Kintex in
[10] can be here directly recalled:

- in this paper authors measure that a specific tested device (the Kintex 7K325) have "1 CRAM upset every 150 seconds and 1 BRAM upset every 670 seconds" (2.5 minutes and about 11.2 minutes respectively).

- authors also suggest that "active SEU mitigation methods will be required to incorporate Kintex7 FPGAs within the ATLAS LAr detector".

Finally, the lane upset rate is higher than the acceptable failure rate in ATLAS, but with the application of mitigation methods as TMR, not applied during the last test, the MGT sensitivity and thus the error rate can be significantly reduced.

In conclusion, it is clear that protection against the effects of radiation on FPGAs is essential to ensure an acceptable upset rate for the LAr FE electronics.

4. Mitigation Techniques

Since FPGAs can operate unreliably in a high radiation environment, mitigation techniques are needed to correct or mask the faults caused by SEUs. In this section the mitigation techniques used to improve the FPGAs robustness will be presented.

The mitigated circuitry itself is subject to SEUs, so it has to be carefully designed to minimize the potential sensitivity and avoid additional reliability losses.

4.1. TMR: Triple Modular Redundancy

TMR is a static hardware redundancy scheme for masking single faults in a digital circuit. The masking is performed by triplicating logic resources and inserting a majority voter, as in Fig. 21. In this way if a SEU changes the state of a module, the majority voter will see two equal results and a flawed one, and will choose to transmit only the good result, assuming that only one module
failed. If a particle strike were to cause two errors or if errors are allowed to accumulate, two domains could fail and TMR would choose an incorrect output. To reduce the probability of fault accumulation, a scrubbing technique is usually implemented to prevent errors from accumulating and breaking TMR.

![Triple Modular Redundancy scheme](image)

The TMR technique has a very high area cost, since it requires to triple the whole design and needs an additional circuitry for the majority voter implementation. To ensure protection against SEUs, the mitigated design will occupy an area that goes from a minimum of three up to six times the original design area [3]. There are also issues related to timing performance and power consumption of the design, since the additional resources needed for mitigation may worsen the design characteristics. In particular the voter circuitry is made of combinational logic which increases the critical path length and negatively affects timing [39].

Previous studies have investigated alternative mitigation techniques to TMR [3]. However, they have shown that common ASIC mitigation techniques are not viable for FPGA design, making TMR the best available choice for reliability and resource requirements.

TMR is therefore one of the best techniques for improving a FPGA’s reliability under radiation: it has been widely used and has shown significant enhancements on different designs. Some tools for automatic TMR implementation are available, such as the BLTMR, developed at BYU and LANL that
will be used in the present work [2].

4.2. Configuration Scrubbing

Configuration scrubbing is a mitigation technique that provides error correction by repeatedly scanning and cleaning configuration upsets. It is usually composed of an external radiation hardened circuit that has access to the FPGA configuration memory and compares it against a golden copy. If a discrepancy is found, the fault is corrected and scanning resumes. Configuration scrubbing is necessary in circuits that use mitigation such as TMR, because it prevents the buildup of configuration faults. Without scrubbing TMR would become ineffective, since faults will accumulate and eventually overcome the voter output.

There are a variety of scrubbing modes that can be utilized for a system. Three of the most common scrubbing methods are blind, readback and hybrid. Blind scrubbing continuously writes the golden bitstream to the FPGA, even if there are no errors in the system. This method is the quickest to correct the errors. However, often the operation is not really necessary and high amount of bandwidth is required. Readback scrubbing reads the configuration memory, compares it to the golden bitstream and then corrects any errors. While not very useful in a deployed system, this method provides useful data about upset rates and failure modes. The final method, hybrid scrubbing, utilizes the on chip scrubber to fix single bit errors (using the ECC codes) and only uses external circuitry and memory when a MBU occurs. This method is slower than blind scrubbing, but utilizes less system bandwidth.

5. Project Setup

To validate that mitigation techniques can make FPGAs reliable for use in the front-end electronics of LAr calorimeter, a custom developed board has been designed and is under production. The board hosts four Xilinx 7-Series Kintex 325T FPGAs, whose behavior under radiation has been extensively studied, as we have seen in 3.2.
Table 5: Data from [40].

<table>
<thead>
<tr>
<th>Kintex XC7K325T Important Features</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>User FFs</td>
<td>407,600</td>
</tr>
<tr>
<td>Logic cells</td>
<td>326,080</td>
</tr>
<tr>
<td>Slices</td>
<td>50,950</td>
</tr>
<tr>
<td>DSP Slices</td>
<td>840</td>
</tr>
<tr>
<td>Block RAM Memory (36 Kb)</td>
<td>445 (total 16,690 Kb)</td>
</tr>
<tr>
<td>Max Distributed RAM (Mb)</td>
<td>4</td>
</tr>
<tr>
<td>Internal memory</td>
<td>16.4 Mb</td>
</tr>
<tr>
<td>Transceivers</td>
<td>16</td>
</tr>
</tbody>
</table>

The ATLAS experiment is very demanding in terms of throughput, since the bunch crossing of the LHC occur every 25 ns. The minimum data throughput of the board using optical links must be 100 Gb/s.

The device is provided with 16 GTX transceivers that can handle 12.5 Gb/s each, providing the needed performance. Other notable features of the used device are shown in Table 5. For a complete overview refer to the Xilinx documentation [40, 41].

The FPGAs will transmit synthetic data that mimic the real calorimeter raw output and will allow corrections from the outside through optical links. The board is built to support many different kinds of scrubbing. All of the scrubbing methods previously explained will be utilized to determine the most effective solution. The different methods can also be implemented in a variety of ways, because access to the configuration memory can be provided in a few ways. The most notable ones are JTAG, SelectMAP and ICAP: the first one is a standard serial protocol, the second one is a parallel high-bandwidth configuration interface while the third one is an IP core that is used for internally self-check the configuration memory. The easiest method is to use JTAG for remotely uploading a new bitfile or correct a corrupted one. This configuration mode is always available and the board is designed to provide both a local and
a remote JTAG path. Since this is a serial interface it is slower than parallel ones, but it is safer since the bitfile to load is located externally to the high radiation environment. A new solution is to implement a “Paired-Self Scrubbing” method, where two FPGAs with the same bitfile will scrub each other using the fast SelectMAP interface. Each scrubbing technique will be studied in detail before deciding upon the final method. In addition to scrubbing, each design will be mitigated with TMR using the latest tools available. Using these tools, we will be able to control the granularity of the domains and measure the trade-offs in the real system as discussed in the following. In fact, the finer the granularity, the greater the reliability improvement, but the resource usage and power consumption greatly increase. We can also investigate the use of partial-TMR (pTMR) to protect the most sensitive and critical parts of the circuit to obtain the greatest reliability improvement while limiting the resource usage and power consumption.

5.1. Modified LTDB Demonstrator

ATLAS electronics must be designed to interface correctly with the existing crate, in particular with the power lines and cooling structure. For this purpose our board is based on the LTDB Demonstrator project, since it already meets the requirements of the crate. This board hosts four FPGAs but does not implement any mitigation, and when errors occur it is only rebooted. The purpose of this board has already been discussed in Section 2 (see page 17).

The reason why the BNL LTDB uses FPGA is that it is easier to prototype (through remote configuration) and it is a small step towards the introduction of programmable logics inside the experiment. Moreover, FPGA devices are ready to be used. The expected life of the demonstrator is of approximately two years and it does not interfere with the data collection and analysis for the ATLAS research purposes, so even if the board cannot be considered reliable for the long term, it serves as a demonstrator and has been accepted. After the Phase I upgrade the LTDBs that will be installed in the FEC will be composed of ASICs that have been thoroughly tested under radiation.
Figure 22: Picture of the LTDB with comments to explain the function of different blocks: analog data from the calorimeter comes from the lower part of the board, then the signals are shaped in the analog mezzanine stage and digitized with the ADCs. The digital data are organized by FPGAs and transmitted to the Back End via optical fibers connected to the PPOD. QSFP block is also connected to optical links and is responsible for remote configuration and slow control of the board.
The different blocks that compose the LTDB are highlighted in 22: from the lower side of the board the three big connectors provide to the on-board circuits the raw data from the calorimeter. The analog mezzanine boards presample and shape the analog data, then the ADCs convert the analog signal to a digital one. Data are then passed on to the FPGAs that encode and transmit them to the Back End through optical links (named PPOD links). Slow control logics and remote JTAG reconfiguration are achieved through another optical link, indicated with QSFP in the image. The most notable characteristics and design choice common to the LTDB board and the modified one will be now briefly presented.

5.1.1. Dimensions

The board is designed to fit the LAr FE Crate: it is 48.90±0.04 cm x 40.93±0.02 cm. The board thickness must be below 2.5 mm and with components mounted cannot exceed the 6 mm height limit for the cooling plates installation.

5.1.2. Clocks

The clock distribution has been very carefully designed to make sure that the different parts of the board are all perfectly synchronized. An external board transmits commands to configure the clock chips on the LTDB using the QSFP optical links.

Figure 23 shows the clock distribution tree that is currently used on the LTDB demonstrator board. The different devices used are now briefly discussed.

- **CDCE62005 (CDCE1 and CDCE2):** this device is a clock generator and jitter cleaner that is programmed from FPGA2 to provide 120 MHz and 160 MHz clocks as output. It is not used currently but is configured with the initial link reset.

- **LMK03200 (LMK1 and LMK2):** this is a 0-delay clock conditioner that is programmed as clock generator using its integrated voltage controlled oscillator (VCO). On the LTDB it is used to generate the 40 MHz
Figure 23: The clock distribution tree of LTDB. The gray links are not currently used in the board. All resources are redundant so that if an error occurs in the chain it is easy to switch to a completely symmetric clock distribution.
clock for the ADCs and the 160 MHz clock for the GTX transceivers of the FPGAs.

- **CDCLVD1208 (CDCLVD1 and CDCLVD2):** this chip is a 2:8 low additive jitter LVDS buffer. The IN_SEL bit that comes from the FPGA determines which one of the two clock inputs will be distributed very precisely to the 5 used outputs.

- **CDC1310 (CDC1 to CDC6):** this device is a low-jitter buffer as the previous one. Its 10 outputs are all single ended (LVCMOS standard) and are used for distributing the 40 MHz clock to the 40 ADC chips. It is very important for all the ADCs to be perfectly synchronized, and to sample at the exact same frequency, so that data from the calorimeter can be effectively compared. In the modified board only one clock for each of these components will be used, since only 4 ADCs will be left on the board.

### 5.1.3. Optical Connections

The front-end crate allows communication with the back-end only through optical fibers, so both data from the calorimeter and control signals must be passed to the LTDB through optical links.

There are two types of optical connectors that are used: the data links and the slow control links. A diagram representing the optical links structure of the LTDB demonstrator is depicted in figure 24.

In order to use existing transceivers for the optical data link it is necessary to cut the motherboard and mount the component on a small mezzanine card, to respect the 6 mm height requirement of FEB boards. On the board there are then four mezzanine boards that connect to an Avago PPOD transmitter (AFBR-810BEPZ) via an FCI connector. Each of these devices can be used to transmit up to 12 links, but only 10 are needed.

The slow control optical links requires fewer connections, so a QSFP+ connector is used (Avago AFBR-79EIDZ); inside this transceiver there are four
receivers and four transmitters. The dimensions of this connector require a cut in the motherboard and the construction of a small mezzanine as for the PPOD.

The four links are used for the slow control of the board through the master FPGA and for the reconfiguration of the FPGAs using a remote JTAG chain.

5.1.4. Power Distribution

The power distribution to all the FE boards is mounted along one side of the crate, while other signal lines are routed on the backplane. In this way the noise of parasitic coupling is minimized.

Power lines are brought to the crate through a power bussbar, and the power distribution of each board is achieved by means of a power comb that provides
a pass-through connection, as can be seen in figure 25 and 26. In figure 6 on page 11 the complete setup of a FEC is shown, with the PowerBus on the side of the crate in view.

After the board is placed into the crate, the comb is inserted mating each pin with a connector on the board and providing the physical contact between the board and power supplies. The contacts are rated for 10 A, and have a nominal resistance of 2.5 mΩ [19].

The supplied voltages are +11V (not used on the board), +7V, +6V, +6V (analog), +4V, -4V, -7V and ground reference. These voltages are distributed to some DC/DC regulators that distribute the correct voltage to the different components on the board, as shown in figure 27.

5.1.5. Cooling Plates

Power dissipation of the board is another issue of great interest: due to the presence of a magnetic field, the use of fans to cool off the electronics is not feasible. Water cooling plates are then the chosen solution for the FE boards in the crate [19]. To avoid leaks, water is kept ≈ 600 mbar under pressure. In this way leaks will result in air entering the cooling system rather than water flowing out. The coolant circulator is designed to provide water at a temperature of almost 18°C.
Figure 27: Diagram of the power distribution.
The FEBs have cooling plates on both sides, attached to the board by means of aluminum stand-offs that establish the distance of 6 mm from surface of the board to cooling plates that sets the constraint on components height.

The gap between cooling plates and active components on the FEB is compensated by heat transfer plates made of aluminum sheets, cut in size and drilled to match the component on the board. The sheets are hydroformed on a mold derived by three-dimensional models of each side of the FEB.

In figure 28 a picture of a cooling plate and a heat transfer plate for a FEB is shown.

Silicone-based thermal pads impregnated with aluminum oxide are used above highly dissipating components on the board. The material has a thermal conductivity of $\approx 1\text{W/m} \cdot \text{K}$ (300 times lower than aluminum) but it is an excellent electrical insulator, which is the reason why it is used to couple active components to the heat transfer plates.

5.1.6. ADCs

There are 40 ADCs on the LTDB that are used to digitize the data from 320 channels and send them to the FPGAs. Since the modified board transmits only known synthetic data, in this project only 4 ADCs have been left. Their inputs will be driven by reference voltages and simple oscillator circuits.
Figure 29: LVDS timing diagram (per ADC channel): the input is sampled at ADCLK, digitized and serialized at 12x ADCLK with a latency of 6.5 clock cycles.

The analog to digital conversion is made by means of a commercial ADC that has been tested under radiation. The ADS5272 is an 8-channel ADC that can provide a sample rate of up to 65 MSPS with 12 bits LVDS output. The resolution of the ADC is calculated as 
\[ R = \frac{V_{FS}}{2^N} \] where \( V_{FS} \) is the full scale voltage range of the converter and \( N \) is the number of bits of the conversion. In our case the resolution is:

\[ R = \frac{2.033V}{2^{12}} = 4.96 \times 10^{-4} V = 496 \mu V \] \hspace{1cm} (2)

All eight channels of the ADS5272 operate from a single clock referred to as ADCLK. The sampling clock is generated internally from this clock reference using a phase lock loop (PLL). The PLL multiplies the clock by a factor of 12, and a 6x and 1x clocks are also output in LVDS format to help synchronize with the data.

The ADC uses a pipelined architecture, that results in a latency of 6.5 clock cycles, as can be seen in figure 29.

The 40 ADCs have some internal registers that are accessed by SPI; using the slow control logic link to FPGA2 one can reset the output, use a defined or
custom pattern, select the MSB or LSB mode or power down some channels.

5.1.7. JTAG

Reconfiguration of the FPGAs with new source code can be achieved with the use of JTAG. There are two ways for accessing the JTAG configuration path on the board: a “local” one, with a physical connector on the board, and a “remote” one that uses two optical links to transmit the bitfile to the FPGAs.

The “local” configuration setup uses a standard JTAG cable that connects to a 14 position header. This component is connected to a series of jumpers (J61-J72) that allow multiple device configuration and choice between the remote or local path. This configuration mode can be used for debugging or during tests: once the board is placed in the FEC its connection to the back-end is only possible with optical links, so a “remote” configuration setup is needed. There are two bidirectional optical links in the QSFP module that are designed to distribute the JTAG commands to the board: the links are passed to two transceivers (TLK1 and TLK2) that serialize and deserialize data at 2.5 Gbps.

The decoded signals are connected to the reserved configuration pins of FPGAs via proper connection of jumpers, as in figure 30.

In figure 31 the block diagram of remote JTAG configuration is shown. The transceiver used for data preparation is the TLK2501, that is optimized for operation at a serial data rate of 2.5 Gbps and has two parallel 8-bit/10-bit decode circuits to convert the 20-bit deserialized word into a 16 bit one. The
Figure 31: Remote JTAG configuration blocks on LTDB board: from the QSFP module the JTAG signals are deserialized and decoded by the TLK2501 devices and then connected to the FPGAs configuration pins - jumpers must be set up correctly as explained in figure 30.
comma detect circuitry can recognize a special character that define the word boundary and synchronize the data stream accordingly. The reference clocks are two oscillators with a frequency of 125 MHz. The internal PLL multiplies this frequency by 20 to match the serialized data rate coming from the back-end.

5.1.8. FPGAs

The LTDB hosts four Xilinx Kintex-7 FPGAs (XC7K325T-FFG900), that do not implement any mitigation in the code. The FPGAs are mainly used as transceivers for the digitized data that come from the 40 ADCs. Figure 32 shows a block diagram of the FPGAs used on the LTDB. The device that can receive commands from the outside is the master FPGA, and the others are slaves. The LTDB is designed so that both FPGA 2 and 3 can be the master, since the board is symmetrical. In the final configuration FPGA 2 has been chosen, following the block diagram of figure 32.

Signals and clocks coming from the ADCs are organized and prepared for transmission via the GTX connected to the PPOD transmitters. Signals coming from the QSFP are interpreted and the corresponding command is issued in the
right way: a finite state machine (FSM) takes care of reading the slow control data and update the addressed register, so for example a reset for the GTXs can be called from remote.

The four FPGAs are connected to each other, so that slow control can also address different FPGAs: the FSM in FPGA2 recognizes that the remote signal is meant for another FPGA and transmits the information to the correct component, that in turn has a state machine that manage to interpret the information correctly.

5.1.9. Back-end Structure

The experimental setup that will be used is shown in figure 33. The back-end system will be composed of two evaluation boards and a remote PC: data is read out by a Zynq evaluation board that checks if the transmission is valid. Slow control links and remote JTAG configuration are entrusted to a Kintex-7 evaluation board. A remote PC will be connected to both these boards to retrieve information on the system. In this way the boards can send information
on the line connection status and messages if an error has been detected. A live check on the board’s behavior under radiation can also be ensured.

6. The final board

In this work we have described the challenges of building a demonstrator board for the ATLAS LAr front-end electronics. The final layout of the board is visible in figure 34. The board is composed of 16 layers and is $48.90 \, \text{cm} \times 40.93 \, \text{cm}$, and will use all 16 GTX links either for communication with the outside or with the other FPGAs of the board, so that statistic on a large number of GTX can be collected. With mitigation techniques reliability of the system is improved and error propagation is limited to the board itself, since fault correction algorithms mask the faulty bit and correct the error before corrupted information spreads in the external system.

7. Conclusion

In the current work we have presented a new approach to radiation resistant front-end electronics for HEP experiments: a programmable-logic based transceiver design that meets the high performance requirements of the ATLAS experiment. In the paper we have discussed the harsh radiation environment of the ATLAS LAr calorimeter, describing in detail the front-end electronic system.

Radiation damage caused to electronic systems and in particular to programmable logic devices have been presented, with a particular attention to Xilinx 7-series FPGA architecture.

To justify the interest to pursue the development of a custom board that hosts mitigated FPGAs, the measurement taken during previous radiation campaigns have been presented. These tests provided an experimental evaluation of Kintex-7 evaluation boards in order to prove that, with the latest technologies, FPGAs are a suitable option for front-end radiation levels.
Mitigation techniques are mandatory, though, if reliability requirement of the ATLAS LAr FEC must be met. For this reason TMR and scrubbing designs have been presented in the paper, since they are our chosen strategy to ensure the correct FPGA operation under radiation.

A custom board based on the LTDB project is currently under production: its main characteristics have been presented and the challenges of software and hardware development have been explained. With the configuration described reliability is improved, since external failure rate is decreased by applying the presented SEU mitigation techniques.

A radiation campaign at the Los Alamos National Laboratory is planned to test the behavior of the board under a neutron spectrum that is very similar to
the one expected in the LAr front-end. After radiation data has been collected and analyzed, the proposal for inserting the board in a free slot of the FEC before April 2017 will be feasible. After the insertion in the crate we will be able to try more mitigation strategies and collect a lot of data for FPGA behavior under radiation.

Acknowledgements

The authors would like to thank Istituto Nazionale di Fisica Nucleare (INFN), Brookhaven National Laboratory (BNL), Università degli Studi di Milano (UNIMI) and Brigham Young University (BYU) for supporting the research described in this paper.

Appendix A. Terminology

The following definitions are used in the paper for radiation damage and error in semiconductor devices \(^{42, 43}\).

Dose is a measure of energy deposited per unit mass of medium and the unit of dose is rad or Gray. The SI unit for absorbed radiation is Gray (Gy) defined as the absorption of one Joule of radiation energy per one kilogram of matter.

Dose rate is the amount of ionizing radiation which an object would receive per unit of time.

Total dose is the total accumulated amount of absorbed ionizing radiation specified at a particular dose rate exposure at +25°C.

Fluence is the number of particles crossing a surface per unit area.

Total Ionizing Dose (TID) is the amount of radiation dose that a device can tolerate before failing to meet published parameter specifications. This is a long-term effect of radiation on an electronic component.

Single-Event Upset (SEU) is a soft error that occurs when a single particle strikes the sensitive volume of a memory cell, generating a charge which
causes a change in the logic state of the cell. That node will remain in the upset state until new data is written into the memory element.

**Single-Event Transient (SET)** is a temporary voltage variation at a node in an integrated circuit caused by a single energetic particle strike. It is a glitch that propagates through the circuit, and if it results in a proper change of state it becomes an SEU.

**Single-Event Functional Interrupt (SEFI)** is a detectable functional failure, often associated with an upset in a control bit or register and can be corrected with a device reconfiguration or a reset. Examples of SEFIs include device-level reset, lockup, initiation of power-on reset, initiation of unique operating mode (brown out, sleep mode, etc.), and device shutdown. Fortunately, the circuit area devoted to these structures is very small and the probability of causing such SEFIs is very low.

**Single-Event Latch-up (SEL)** is an anomalous high-current state caused by the passage of a single particle in the device. If the current is above device specification, it must be immediately cleared with a power cycle to avoid permanent damage to the device. An example is in a CMOS device, where the PNPN structure can be seen as a PNP and an NPN BJT, stacked next to each other. SEL occurs when the parasitic current energy is able to switch on one of the BJTs.

**Single-Event Burnout (SEB)** is a hard error, caused by a high-current state that (unlike SEL) results in device damage.

**Single-Event Gate Rupture (SEGR)** occurs when the gate oxide is damaged and a new current path is created. It is a hard error because it cannot be corrected.

**Multiple-Bit Upset (MBU)** takes place when two or more error bits occur in the same word, and cannot be corrected by a simple single-bit Error Correcting Code (ECC).
References


j.measure.2015.04.002.


55


URL http://stacks.iop.org/1748-0221/3/i=08/a=S08003

URL http://cds.cern.ch/record/331061?ln=it


1748-0221/3/09/P09003.
URL http://cds.cern.ch/record/1133362

URL https://cds.cern.ch/record/2063426


URL https://cds.cern.ch/record/1602230


URL http://stacks.iop.org/0022-3727/38/i=13/a=015


