

Exploring FPGA hardening solutions at the detector level for the future high luminosity phase of the CMS experiment at the LHC

Jérôme Lemaire

Master thesis submitted under the supervision of
Prof. dr. ir. Mister Frédéric Robert

The co-supervision of
Prof. dr. ir. Mister Gilles De Lentdecker

In order to be awarded the Master's Degree in
Electronics and Information Technology
Engineering

Academic year
2015-2016

Abstract

The objective of the GE1/1 project which is under development since 2010, is to install triple-GEM detectors in the endcaps area of Compact Muons Solenoid (CMS) experiment, one of the four main experiments of the Large Hadron Collider (LHC). This master thesis is a contribution to the development of the Data Acquisition electronics (DAQ) of the project, more precisely on the necessity of the radiation hardening of the FPGA on the Opto-Hybrid board (OH). First of all, an overview of the interactions between incident particles and electronic devices and the categorisation of the different possible induced misbehaviours is given. Moreover, the simulations on the environment inside CMS are presented and are followed by a description of different existing mitigation techniques for electronic devices found in the literature. The second part of this thesis is about the development of a firmware in order to test the different resources (CLB, BRAM and configuration memory) of the FPGA with and without mitigation techniques. In order to test the CLB, the Triple Modular Redundancy (TMR) mitigation technique was chosen and applied to the zero suppression module. The latter is a critical part of the firmware which will be implemented on the FPGA at the end. Concerning the BRAM, the IP "block memory generator" has a feature called Error Correction Capability (ECC) that allows to detect and correct upsets by the use of Hamming code. Finally, the configuration memory use the Soft Error Mitigation (SEM) IP in order to detect and correct upset in it.

The firmware was tested at the UCL cyclotron Cyclone. The FPGA was irradiated by a beam of protons with a flux several order higher than those predicted inside CMS. This allowed to observe a significant amount of upsets in it but also to reach a Total Ionizing Dose (TID) of 84 krad comparing to the 10 krad predicted by simulations as the accumulated dose at the end of operations of the LHC.

The analysis revealed the FPGA is still working well after a TID of 84 krad. The FPGA is then adapted to working properly until the end of operations of the LHC. The ECC for the BRAM was able to correct every single upsets and detect double upsets. Concerning SEM, upsets in configuration memory were detected and around 90 % were corrected within a few milliseconds. This reveals the necessity of the mitigation of logic implemented on the FPGA. Unfortunately, TMR could not be tested due to a malfunction of the firmware. Thus, more tests will be necessary in order to determine the efficiency of this mitigation technique.

Keywords: FPGA, GE1/1 project, CMS, LHC, Triple-GEM, DAQ.

Jérôme Lemaire - Electronics and Information Technology Engineering - 2015-2016
Exploring FPGA hardening solutions at the detector level for the future high luminosity phase
of the CMS experiment at the LHC

Résumé

L'objectif du projet GE1/1, en développement depuis 2010, est d'installer des détecteurs triple-GEM dans la zone des bouchons au sein de l'expérience Compact Muons Solenoid (CMS), l'une des quatre principales expériences du Large Hadron Collider (LHC). Ce mémoire est une contribution au développement de l'électronique d'acquisition du projet et plus particulièrement la nécessité du durcissement aux radiations du FPGA présent sur la carte opto-hybrid (OH). Premièrement, une vue d'ensemble des interactions entre les particules incidentes et l'électronique et la catégorisation des mauvais fonctionnements induits est donnée. Les résultats des simulations de l'environnement l'intérieur de CMS effectuées est également présenté suivi par une description de différentes techniques de mitigation possible trouvées dans la littérature. La seconde partie de ce mémoire se penche sur le développement d'un firmware afin de tester les différentes ressources du FPGA (CLB, BRAM et mémoire de configuration) avec et sans technique de mitigation. Afin de tester les CLB, la technique de triplification a été choisie et appliquée au module de suppression de zéro. Ce dernier est une partie critique du firmware qui sera implémenté sur le FPGA à la fin du projet. Concernant la BRAM, l'IP "block memory generator" possède une fonction appelée "Error Correction Capability" (ECC) qui permet de détecter et corriger des erreurs grâce à l'utilisation du code de Hamming. Finalement, la mémoire de configuration utilise l'IP "Soft Error Mitigation" (SEM) afin d'y détecter des erreurs.

Le firmware a été testé au cyclotron de l'UCL. Le FPGA a été irradié à l'aide d'un faisceau de protons dont le flux est bien plus élevé que ceux prédit à l'intérieur de CMS. Ceci a permis d'observer un nombre significatif d'erreurs mais aussi d'atteindre une dose totale déposée d'environ 84 krad. Ce qui est plus que les 10 krad atteint en fin de vie du LHC prédits par les simulations.

L'analyse des résultats a révélé que le FPGA fonctionne toujours correctement après une dose déposée de 84 krad. Le FPGA est donc adapté pour une utilisation jusqu'à la fin des opérations du LHC. Au niveau de la BRAM, l'ECC a corrigé toutes les erreurs simples et détecté les erreurs double. Concernant SEM, les erreurs dans la mémoire de configuration ont été détectées et environ 90 % ont été corrigées en quelques microsecondes. Ceci révèle la nécessité de la mitigation de la logique implémentée dans le FPGA. Malheureusement, la triplification n'a pas pu être testée à cause d'un mauvais fonctionnement du firmware. D'autres tests seront donc nécessaires afin de déterminer l'efficacité de la triplification.

Mots-clés FPGA, projet GE1/1, CMS, LHC, Triple-GEM, électronique d'acquisition

Jérôme Lemaire - Electronics and Information Technology Engineering - 2015-2016
Exploring FPGA hardening solutions at the detector level for the future high luminosity phase
of the CMS experiment at the LHC

Samenvatting

Het doel van het GE1/1 project dat in ontwikkeling is sinds 2010, is om triple-GEM detectoren in de eindkappen van de Compact Muons Solenoid (CMS) experiment te installeren, n van de vier belangrijkste experimenten van de Large Hadron Collider (LHC) . Deze masterthesis is een bijdrage tot de ontwikkeling van de data-acquisitie elektronica voor dit project, meer bepaald het stralingsbestendig maken van de FPGA op het Opto-Hybride (OH) bord. Eerst wordt een overzicht gegeven van de interacties tussen inkomende deeltjes en elektronische apparatuur, samen met de indeling van de verschillende mogelijk genduceerde perturbaties. Verder, worden de simulaties van binnen CMS milieu gepresenteerd, gevolgd door een beschrijving van de verschillende mitigatie technieken voor bestaande elektronische apparaten volgens de huidige literatuur. In het tweede deel van deze masterthesis, werd de firmware ontwikkeld om de verschillende bronnen (CLB en configuratie geheugen BRAM) van de FPGA te testen met en zonder mitigatie technieken. Om de CLB testen, werd de Triple Modular Redundancy (TMR) compensatietechniek gekozen en toegepast op de nul-onderdrukkingsmodule. Deze is een cruciaal onderdeel van de firmware dat finaal op de FPGA zal worden gecomplementeerd. Met betrekking tot de BRAM, de IP "block memory generator" heeft een functie met de naam "Error Correction Capability" (ECC) dat toelaat om fouten op te sporen en te corrigeren via de Hamming code. Tot slot, gebruikt het configuratie geheugen de Soft Error Mitigation (SEM) IP om verstoringen op te sporen en te corrigeren. De firmware werd getest aan de UCL cyclotron Cyclone. Hiervoor werd de FPGA bestraald met een bundel protonen met een hogere flux dan de voorspelde flux binnen de CMS. Een aanzienlijke hoeveelheid verstoringen kon worden gedetecteerd in de FPGA , waarbij een totale ioniserende dosis (TID) van 84 krad werd gemeten. Deze geaccumuleerde dosis is veel hoger dan de voorspelde 10 krad voor alle werkzaamheden van de LHC . Uit de analyse bleek de FPGA is nog steeds goed werkte na blootstelling aan een TID van 84 krad. De FPGA werd dan aangepast om te werken totaan het einde van de operaties van de LHC. De ECC van de BRAM kon elke verstoring corrigeren en dubbele verstoringen detecteren. Met betrekking tot SEM, werden de verstoringen in de configuratie geheugen gedetecteerd waarvan ongeveer 90% binnen een paar milliseconden werden gecorrigeerd. Deze bevindingen onthullen de noodzaak van de mitigatie van de logica uitgevoerd op FPGA. Helaas kon de TMR niet getest worden als gevolg van een storing van de firmware. Daarom zijn meer testen nodig om de efficiëntie van deze mitigatietechniek bepalen.

Sleutelwoorden: FPGA, GE1/1 project, CMS, LHC, Triple-GEM, data-acquisitie elektronica

Jérôme Lemaire - Electronics and Information Technology Engineering - 2015-2016
Exploring FPGA hardening solutions at the detector level for the future high luminosity phase
of the CMS experiment at the LHC

Acknowledgments

I am grateful to my supervisors Frédéric Robert and Gilles De Lentdecker for their good advice, their guidance and their patience all along this year.

I would like to thank Thomas Lenzi who shared his experience and helped me in the design of the firmware.

I would also like to thank people from IIHE for the good meals time and their direct or indirect involvement in the testbeam at UCL.

Last but not least, I would like to thank Vincent Van Eeghem for the abstract translation in Dutch and also Nathalie Fiévet and my godfather who helped me in the re-reading of this thesis.

Contents

Introduction	1
1 The LHC and CMS experiment	3
1.1 The Large Hadron Collider (LHC)	3
1.1.1 The LHC Characteristics	4
1.1.2 The LHC Prevision Calendar	5
1.2 The Compact Muon Solenoid (CMS)	5
1.2.1 Particle Detection Principles	6
1.2.2 The CMS Sub-Detectors	6
1.2.3 The CMS Trigger System	7
2 The GE1/1 project	9
2.1 Overview of the Different Technologies Inside the Muon Tracker	9
2.2 Presentation of the GE1/1 Project	10
2.2.1 Listing of the Constraints	10
2.3 Description of the Detector Components	11
2.3.1 The Triple-GEM Detector	11
2.3.2 Data Acquisition Electronics	14
3 The radiation effects	19
3.1 The Physical Effects on Matter	19
3.1.1 Energy Loss Mechanism	20
3.2 The Functional Effects on Devices	22
3.2.1 The Single Event Effects	22
3.2.2 Total Ionizing Dose	25
3.2.3 Conclusion	25
3.3 The Cross-Section	25
3.4 The CMS Environment	26
4 The Mitigation Techniques	29
4.1 The Fault Detection Methods	29
4.2 The Triple Modular Redundancy in Details	31
4.2.1 The Implementation of TMR	34
4.3 The Reconfiguration of the FPGA	34
4.4 The Fault Repair Methods	35

5	Test beam description	37
5.1	FPGA Irradiation Tests Principle	37
5.2	Description of the Test Bench	37
5.3	Description of the Testing Firmware	40
5.3.1	The Irradiated FPGA	41
5.3.2	The Reference FPGA	44
6	The test results	47
6.1	Firmware performances	47
6.2	Data Analysis	48
6.2.1	The Communication Analysis	48
6.2.2	The Upset Counters Analysis	48
6.2.3	Upset Rate in CMS	49
6.3	The Total Ionizing Dose	50
6.4	For the Future	50
	Conclusion	51
	List of Figures	53
	List of Tables	55
	Bibliography	57

Introduction

The European Organization for Nuclear Research (CERN) is one of the biggest research center in the world and where the Large Hadron Collider (LHC), the biggest and most powerful particle accelerator in the world is established. In order to produce new particles, two beams collide in four points of the LHC. This is where the main experiments are. One of them is the Compact Muons Solenoid (CMS), composed of many sub-detectors detecting the momentum and the energy of different types of particles.

During the lifetime of the LHC, the luminosity is gradually increased producing more data to analyse. However, in 2023, the LHC will enter in phase 2 and has the objective to increasing the luminosity by a factor 2.5 resulting in a significant increase of the amount of collected data. Thus, an upgrade of the different sub-detectors has to be done. The GE1/1 project is part of the upgrade of CMS. It has started in 2010 and is planned to be installed during the Long Shutdown 2 (LS2) in 2018. This project is developing a gaseous detector using the Triple-GEM technology and the data acquisition electronics for the forward regions of CMS.

This master thesis has been realised with The Inter-University Institute of High Energy (IIHE), a ULB-VUB collaboration, working with other laboratories on the GE1/1 project. The main contribution of this thesis is about the radiation hardening of the FPGA on the opto-hybrid board composing the data acquisition electronics. Simulations revealed the background noise is especially high in the region of GE1/1 project and can have a serious impact on the detectors and the electronics. Since the FPGA is not designed to be used in a radioactive environment, studies have to be performed in order to determined the capability of the FPGA to work in such environment.

This master thesis begins with a description of the LHC and CMS (Chapter 1), followed by the description of the Triple-GEM detector principle and its data acquisition electronics (Chapter2). Secondly an overview of the physical principles of the interactions between particles and matter and the resulting malfunctions are described (Chapter 3) and followed by a description of different existing mitigation techniques (Chapter 4). Then, the description of the developed firmware is presented (Chapter 5). Finally, the testbeams performed at UCL cyclotron are described (Chapter 6) and results are presented (Chapter 7).

Chapter 1

The LHC and CMS experiment

1.1 The Large Hadron Collider (LHC)

The Large Hadron Collider (LHC) is the largest and most powerful particle accelerator in the world and started up on 10th September 2008 and should be operating until 2035. It is a part of the accelerator complex at CERN with other smaller accelerators visible on figure 1.1. The LHC

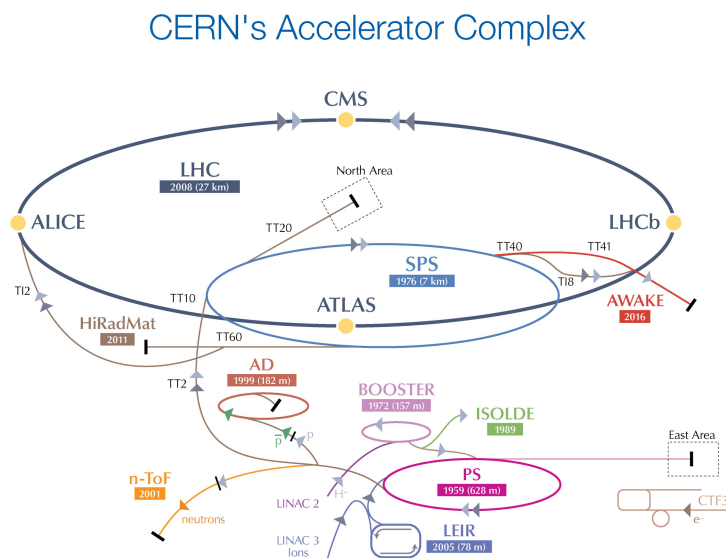


Figure 1.1: CERN's accelerator complex schematic with all the different accelerators and connection between it[1].

appears as a ring of 27 kilometres diameter. Inside of it, two beams of hadrons travel at a speed close to the speed of light in opposite directions and collide in four specific places as depicted on figure 1.2. Accordingly, this is where the different experiments are : Alice, Atlas, LHC-B and CMS.

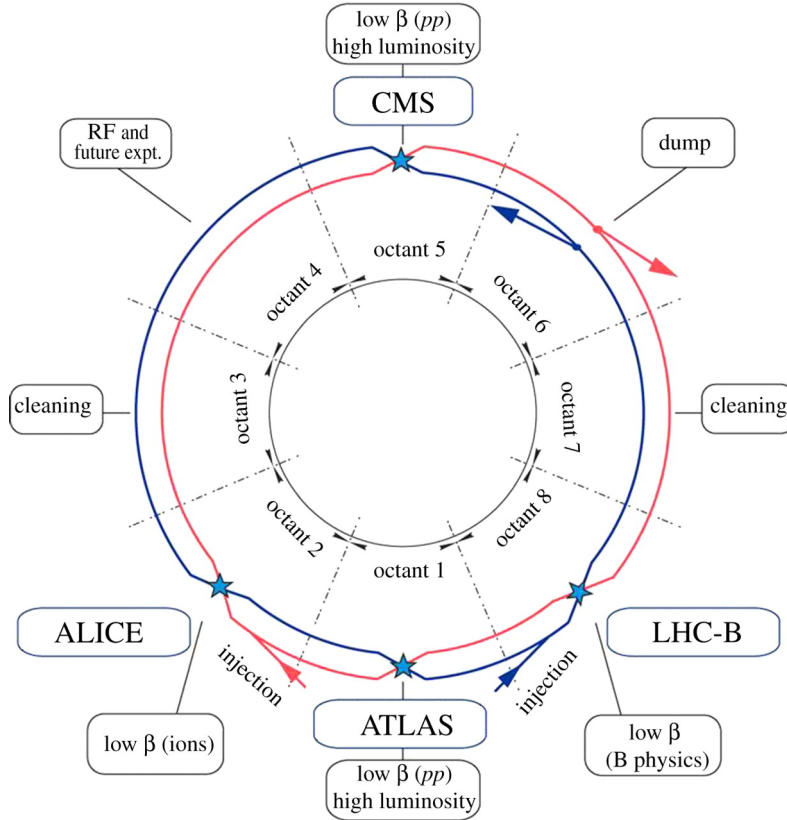


Figure 1.2: LHC schematic with the experiment and beams representation[2]

1.1.1 The LHC Characteristics

Two characteristics of the LHC are relevant for the experiments: the luminosity and the energy of collision. These two are briefly explained below.

Luminosity The luminosity is used to characterise the number of collisions inside the LHC and only depends on beam characteristics. In the beam, protons are not uniformly distributed but packed in bunches. There are around 2000 bunches of 10^{11} protons in each beam performing collisions at a frequency of 40MHz . At each crossing only a small portion of protons will collide. The number of collisions per second N_c can be defined by equation 1.1 [3]

$$N_c = \mathcal{L}\sigma \quad (1.1)$$

where \mathcal{L} is the *instantaneous luminosity* and σ is the cross section. Typical values for these parameters in CMS are $\mathcal{L} = 10^{34}\text{cm}^{-2}\text{s}^{-1}$ and $\sigma = 10^{-25}\text{cm}^2$. Then, N_c is around 25 at each beam crossing and decreases slowly as there are less and less protons in the beam. Moreover, there is also the *integrated luminosity* which is the instantaneous luminosity integrated over the duration of operation. The units are the inverse femtobarn ($\text{fb}^{-1} = 10^{39}\text{cm}^2$) and it can be used to express the amount of produced data [4].

Energy of collision The energy of a collision is expressed in eV . In the case of LHC, the nominal energy is 14TeV and it is running at 13TeV for now for proton-proton collisions.

1.1.2 The LHC Prevision Calendar

The activity of the LHC during its lifetime is depicted in figure 1.3. It is a succession of periods of data acquisition and periods of *long shutdown* (LS). Repairs, upgrades and additions of new features are performed during these long shutdowns. The prediction of the evolution of peak luminosity and integrated luminosity are also presented. This calendar can be divided in two phases, the phase 1 from 2008 to 2023 where the peak luminosity increases gradually as the energy of collision and the phase 2 from 2022 to 2035. This second phase will start with a very long shutdown of more than two years. The LHC will enter in a phase of high luminosity (HL-LHC) and will increase the instantaneous luminosity by a factor of 2.5 and will stay constant during the whole phase. During the redaction of this master thesis, the LHC is in the second period of phase 1 and the integration of GE1/1 project inside CMS is planned during LS2.

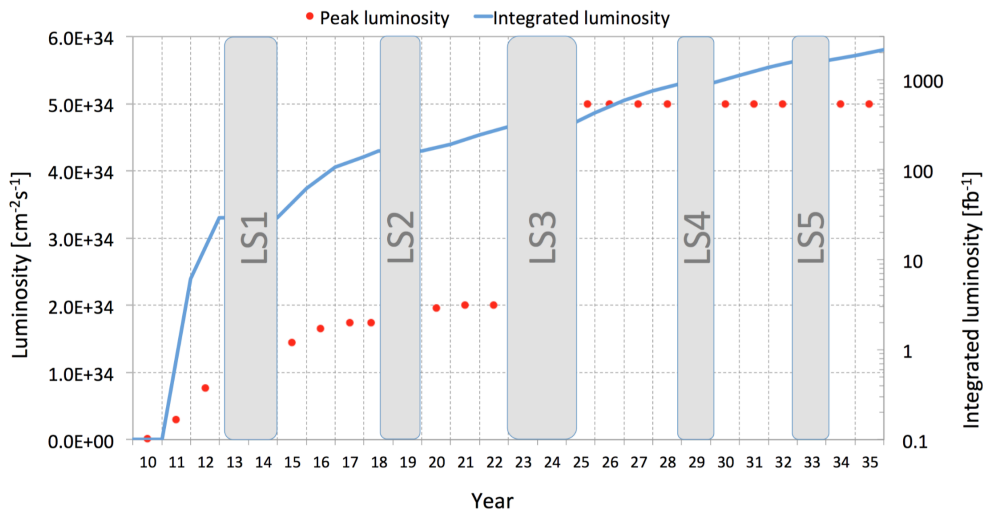


Figure 1.3: Long term LHC calendar with the integrated and peak luminosity previsions

1.2 The Compact Muon Solenoid (CMS)

CMS experiment is one of the four main detectors around the LHC and is used to observe a wide range of particles. It has a cylinder shape of 28.7m long and 15m diameter for a weight of 14,000 tons [5]. Two areas are defined: the *barrel* forming the cylinder and the *endcaps* closing the two extremities. CMS is composed of several layers of components clearly visible on figure 1.4. Each of these layers of sub-detectors is designed for a specific type of particles. When a collision occurs, it produces around 1000 other particles in all directions and will have different trajectories before being absorbed. Then each particle has its own signature.

In the next sub-sections, the different detection principles will be briefly explained followed by the disposition of the different layers inside CMS and finally, the system of trigger of CMS to filter events will be introduced. More details can be found in the master thesis of Hugo Dewitte [6] and Baptiste Herregods [4].

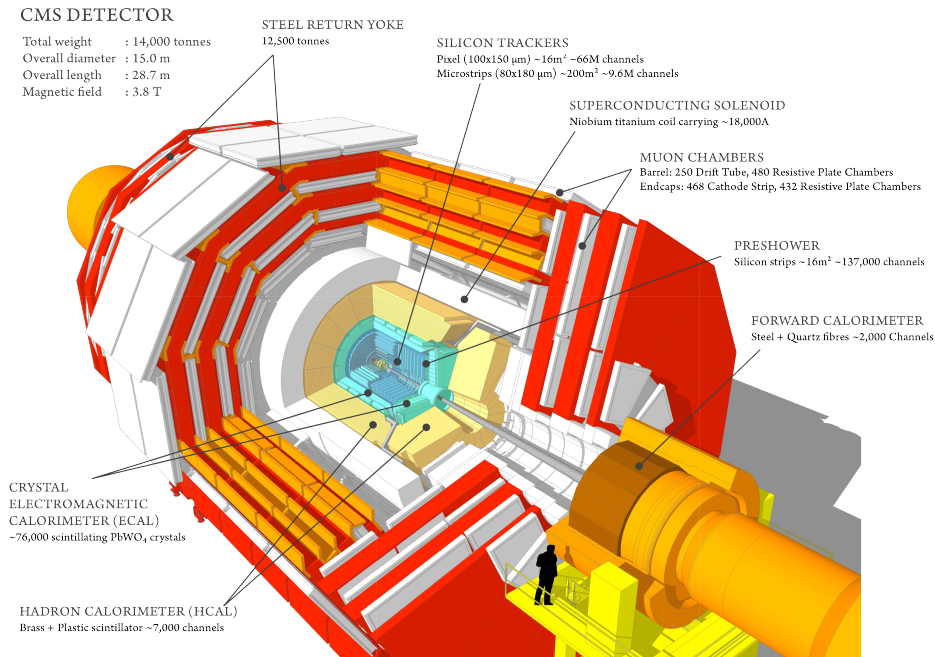


Figure 1.4: CMS experiment 3D representation with its components [5]

1.2.1 Particle Detection Principles

Two parameters of the particles need to be measured: their energy and their momentum. Hence the sub-detectors can be grouped in two categories. The first ones measure the particles energy by absorption and are called *calorimeters*. These are thick enough in a medium that interact with particles in order to stop the particles and absorb all their energy. The second category of sub-detectors measure the particles trajectory and are called *trackers*. The momentum can be recovered by studying the curvature of the trajectory through a known magnetic field produced by the superconducting solenoid visible in white on figure 1.4. These trackers are as thin as possible in order to minimise the interaction with the particles.

1.2.2 The CMS Sub-Detectors

The figure 1.5 is a representation of each layer in the barrel of CMS. Only one sector is represented here. In total, several identical sectors complete the revolution and cover the whole barrel. Furthermore, it is the same disposition for the endcaps where the GE1/1 detector will be placed as it will be explained in more detailed in next chapter.

Right next to the collision point are the silicon trackers. These trackers are used to determine the collision point and the momentum of charged particles. The next layer is the electromagnetic calorimeter (ECAL) that detects electrons and photons. The medium used for ECAL is lead tungstate crystals, a highly transparent material producing light when electrons or photons pass through it. Photodetectors are placed on the back in order to detect the produced light and convert it in electrical signals. A second kind of calorimeter is placed after the ECAL. The hadron calorimeter (HCAL) that are designed in order to detect hadrons. It is composed of

alternating absorbing and scintillating layers. Optic fibers collect the generated light and send it to other photodetectors. All these layers are inside the supraconductive solenoid. A giant magnet of 12000 tons generating a magnetic field of 4 Tesla that bends the trajectory of particles. Such high magnetic field is necessary since trajectory bending is proportional to the energy which is very high in this case. Consequently, all the inner layers are made in non magnetic materials in order to not disturb the magnetic field. In addition some return yokes are used to guide the magnetic field and also to filter the background noise produced by hadronic shower generated by the HCAL. After the solenoid, only muons and neutrinos remain due to their large penetration range. Hence, the muons tracker is placed here. This tracker detects muons, computes their momentum and also participates to the creation of the trigger signal that will be explained in the next section. Due to the distance from the collision point, the surface is around $25000m^2$. Gas detectors are then more suitable as their cost is smaller than semiconductor detectors. The GE1/1 project is part of the muons tracker and is planned to be used in the upgrade of the muon tracker in the endcaps.

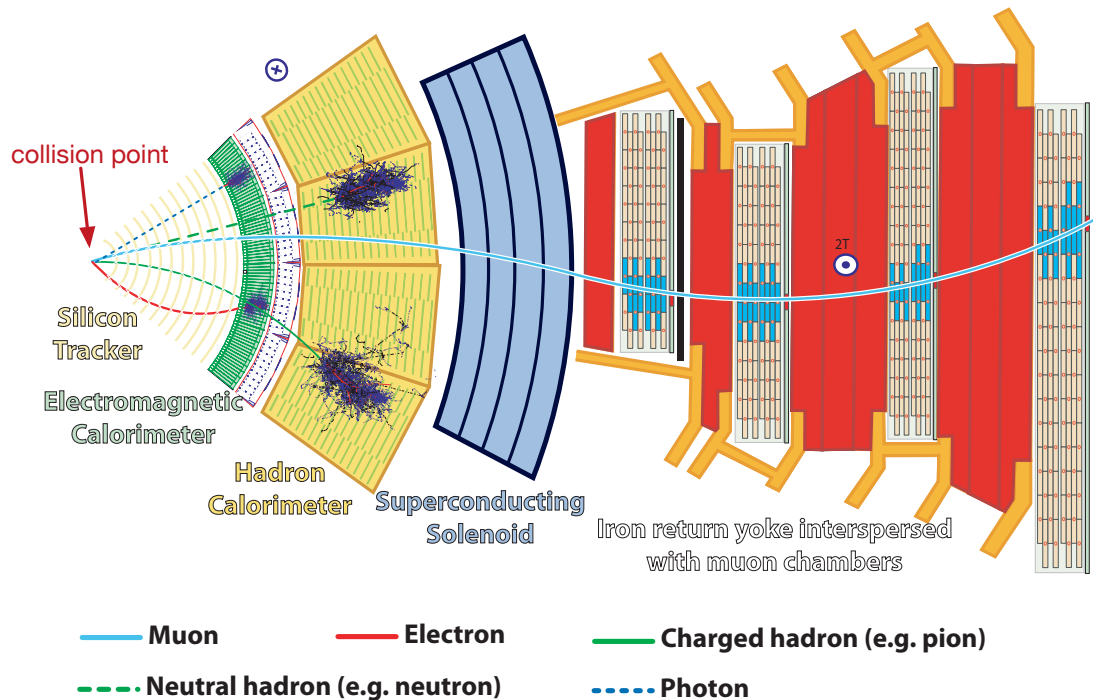


Figure 1.5: Representation of the different components of one sector inside CMS and typical particle trajectories [5]

1.2.3 The CMS Trigger System

The amount of data produced every second by around 10^9 collisions inside CMS does not allow to collect and process every events in real time. Consequently, a system of trigger is needed to filter pertinent events from the others. At LHC, the scientists are looking for theoretical predicted particles with specific characteristic such as charge, energy or mass. Such parameters are used to

define the trigger and select only pertinent events. The CMS trigger system is depicted on figure 1.6. Briefly, the trigger system is composed of two levels of selection, the Level-1 trigger (LV1) and the High Level Trigger (HLT) . The LV1 continuously receives data at 40MHz , stores it in a FIFO memory and has $3.2\mu\text{s}$ to perform basic algorithms in order to make a decision quickly. This process is repeated every 25ns to avoid congestion. Such time constraints do not allow to analyse events in detail. If the event is judged pertinent, data go to HLT. If not, it will be erased by the next event. Such filtering allows to decrease the event rate from 40MHz to 100kHz . The HLT is able to perform more powerful algorithms that take into account the totality of data of the event since the time constraint is around 1s . At the output, the event rate is around 100Hz .

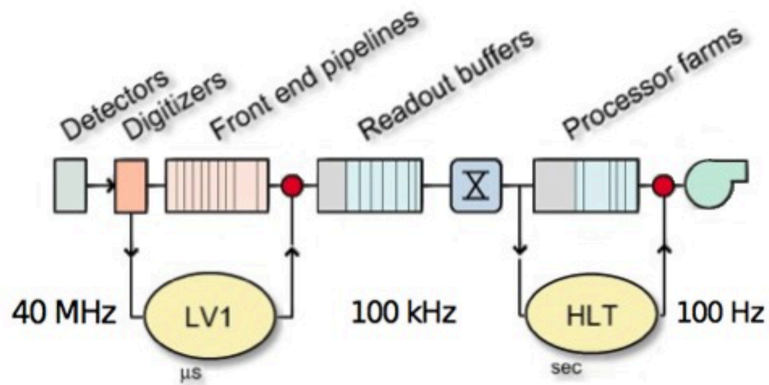


Figure 1.6: Schematic of the trigger system principle [7]

Chapter 2

The GE1/1 project

The GE1/1 project is developing a new gaseous detector based on triple GEM technology. This detector will be added at the muon tracker in the CMS endcaps during the second long shutdown in 2018. In this chapter, each component of the detector and its data acquisition electronics will be presented in details.

2.1 Overview of the Different Technologies Inside the Muon Tracker

Different technologies are present inside the muon tracker : the Drift Tubes (DT), present in the barrel region, the Cathode Strip Chambers (CSC), present in the endcaps region and the Resistive Plate Chambers (RPC), present in both. The disposition of all the sub-detectors can

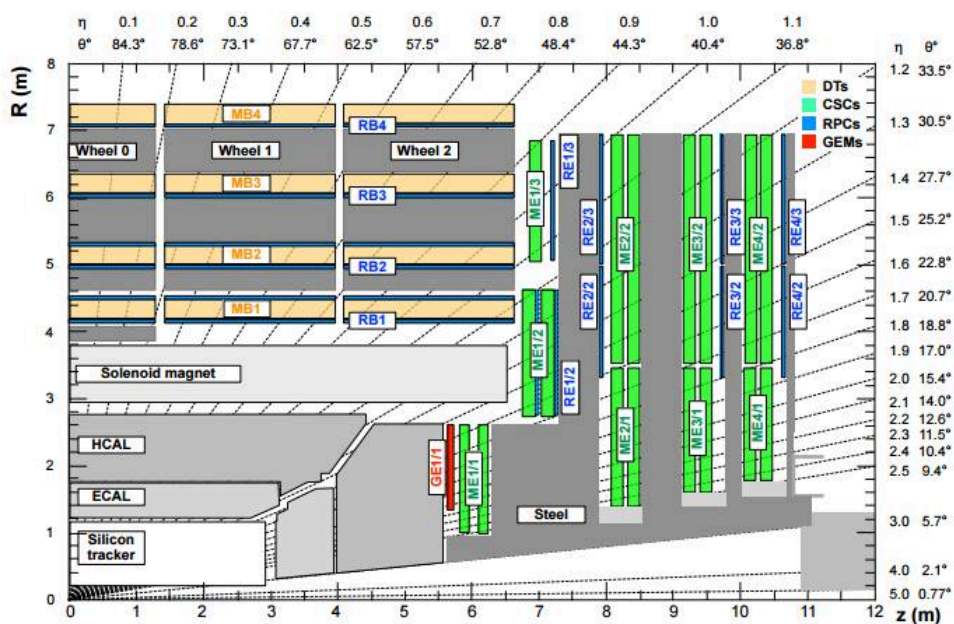


Figure 2.1: Longitudinal section of CMS with the different technologies of detector inside the muon tracker including the GE1/1 project in red [8]

be observed on figure 2.1 which is a transverse section of CMS where GE1/1 detectors are in red, right after the HCAL. In order to define clearly the position of one sub-detector, a coordinate system and a name convention are used. The coordinate can be expressed in a cartesian system $(R; z)$ where the collision point is in $(0; 0)$. Another way is to express the position related to the pseudo-rapidity η that can be expressed by :

$$\eta = -\ln(\tan(\theta/2)) \quad (2.1)$$

Where θ is the polar angle. Each sub-detector has a specific name composed of two letters and two numbers. The first letter represents the technology (G for the GEM), the second letter indicates the region where the sub-detector is (E for endcap). About the numbers, the first one is for the distance in z . The second number is for the distance in R . With such a name it is then possible to define precisely the location inside CMS.

2.2 Presentation of the GE1/1 Project

When the LHC will enter in the high luminosity phase, the number of events will be multiplied by 5 and will generate a significant increase of the produced particles and noise. An upgrade of the sub-detectors is then necessary. The GE1/1 project will use the location of RPC detectors since the new requirements can not be fulfilled by this technology. The Triple-GEM technology should be more adapted than RPC. Moreover, three other projects have already successfully implemented Triple-GEM detectors. Even if these detectors are smaller, their performances are very encouraging to fulfil the constraints presented in the next subsection [6].

2.2.1 Listing of the Constraints

The GE1/1 project has to meet several constraints. The main ones are listed below.[6, 8]

Geometry As this project will take the location of RPC detectors, it has to fit the empty area that is 10cm thick. Furthermore, the whole setup has to completely cover the zone without blind spot.

Detection rate The particle rate at the level of GE1/1 detector is predicted to reach $5kHz/cm^2$ during the HL-LHC phase. A safety factor of 2 is applied. Thus, the detector should be able to handle a particle rate of $10kHz/cm^2$

Efficiency The efficiency has to reach 97% or better. With the use of a double chamber, this efficiency should increase to 99.9%

Angular resolution The angular resolution should be better than $300\mu rad$ in the azimuthal direction (ϕ)

Time resolution The time resolution should be at least of $10ns$

Radiation hardening The total radiation level is expected to be $100mC/cm^2$ during the complete lifespan of the detector. A safety factor of 2 is applied. Thus, the detector should resist to a total dose of at least $200mC/cm^2$

In the next section, the Triple-GEM detector and its different components will be explained.

2.3 Description of the Detector Components

One single detector has a trapezoidal shape and covers 10° in ϕ . So as to satisfy the geometry constraint, 36 detectors are needed. The figure 2.2 is a representation of one endcap region with the 36 detectors depicted in blue and pink. As already mentioned in the efficiency constraint, two chambers are placed in the endcap to improve the efficiency. Moreover this is how the level 1 trigger is achieved by performing a logic OR between two independent detectors. Since, there are two endcaps, 144 detectors are needed in total, covering an area of around $25m^2$.

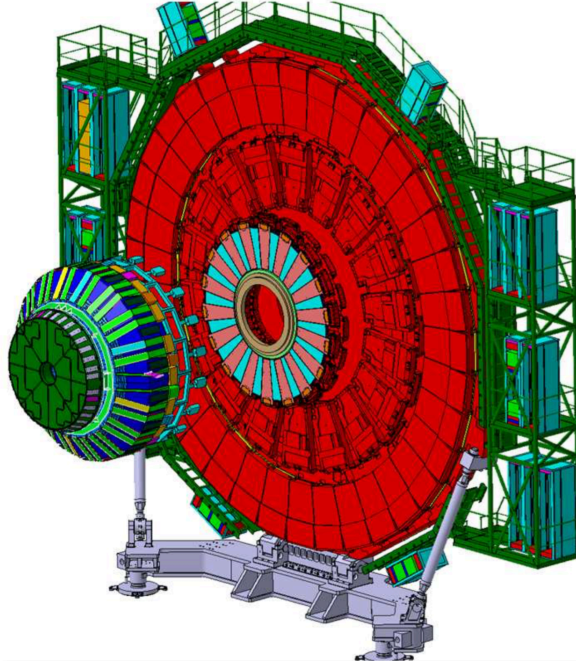


Figure 2.2: 3D representation of one of the endcaps zone with the GE1/1 detector in blue and pink [8].

Figure 2.3 depicts the different components composing the detector. In the next sub-sections, a description of each one will be made starting the GEM foils that actually detect particles until the end of the acquisition electronic.

2.3.1 The Triple-GEM Detector

The Triple-GEM detector is a gaseous detector where the gas is in between an anode and a cathode with a difference of potential. When a particle goes through the gas, it creates electron-ion pairs. The electron from this pair drift due to the electric field applied in the direction of the anode which is the readout PCB. However, an amplification is needed in order to have a readable signal. This is realized by creating an avalanche by the means of a very high electric field ($> 10kV/cm$). Indeed, the electrons are accelerated and acquire enough energy to ionise the gas and create new electron-ion pairs that can create other pairs and so on. In the detector, the avalanche phenomenon is performed by three GEM foils hence the name of *Triple-GEM*.

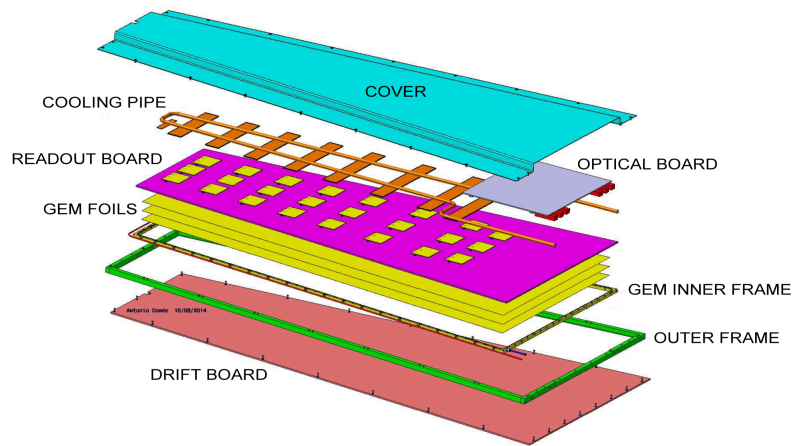


Figure 2.3: 3D representation of each component composing the detector [8].

The GEM Foils

Basically, a GEM foil is a sheet of dielectric (kapton) $50\mu m$ thick in between two copper sheets $5\mu m$ thick and pierced by $70\mu m$ holes with a pitch of $140\mu m$ as it can be seen on figure 2.4. A differential voltage is applied on the copper sheets inducing an electric field around $60kV/cm$

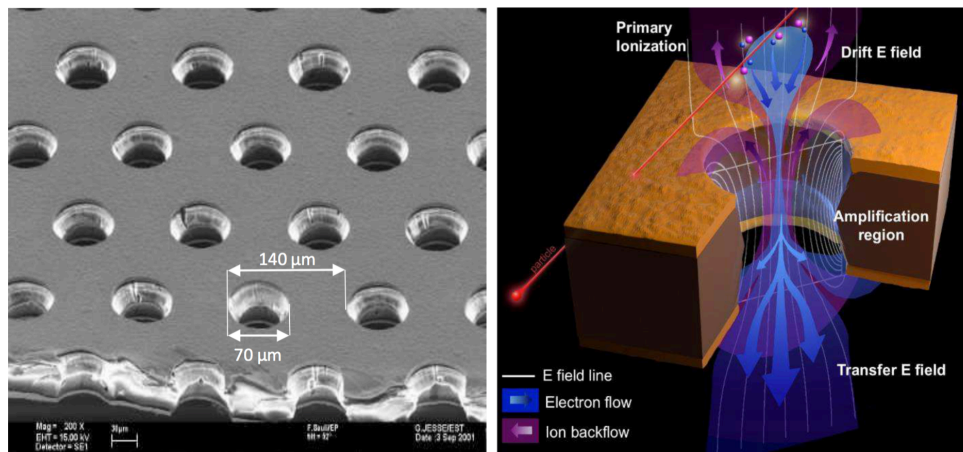


Figure 2.4: Left: Zoomed in view of a GEM foil. Right: Schematic view with the electric field line in white, the electron flow in blue and the ion flow in purple through one hole in the foil [8].

or higher. Typical values for the different foils are depicted on the right side of figure 2.5. The electric field lines are focus on the hole center, attracts the electrons and accelerates them inducing an avalanche visible on the left side of figure 2.5. Typical gain factor is around 20 for each foil leading to a final gain around 8000. After the amplification all the electrons continue their drift until the readout board.

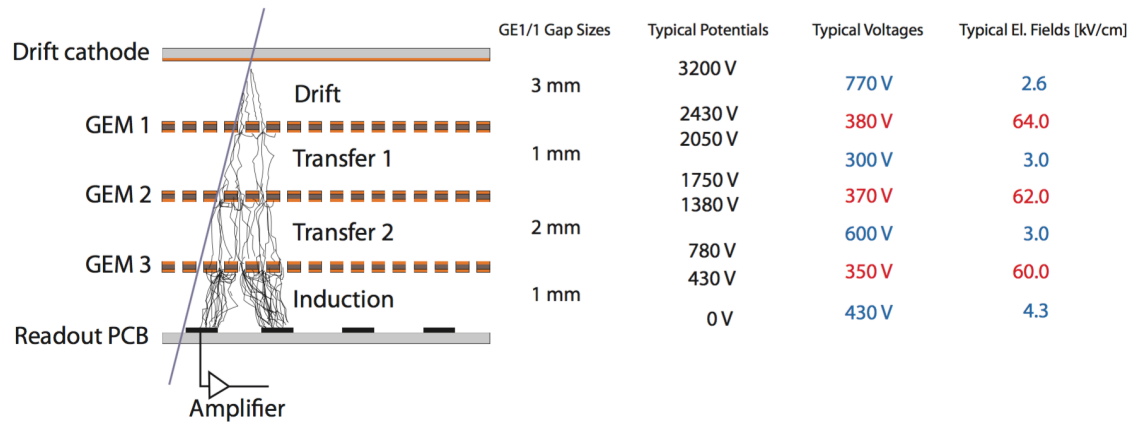


Figure 2.5: Left: Details of a GEM foil principle. Right: Typical size and electrical parameters [8].

The Readout Board

The readout board is covered of gold-plated copper strips on one of its sides that actually capture the electrons created by the GEM foils. The readout board is divided in 8 sectors along its length and has 384 strips along its width and are visible on figure 2.6. About the dimensions of the strips, it is $230\mu\text{rad}$ width with pitch of $463\mu\text{rad}$. Such dimensions and some algorithm to recover the position of a particle allows to meet the angular resolution constraint ($300\mu\text{rad}$). On the other side of the readout board, the 384 strips are gathered in groups of 128 and plugged to a male connector leading to 24 sections (3 groups per sector and 8 sectors). These 24 sections are plugged to 24 VFAT chips thanks to the GEM electronic board.

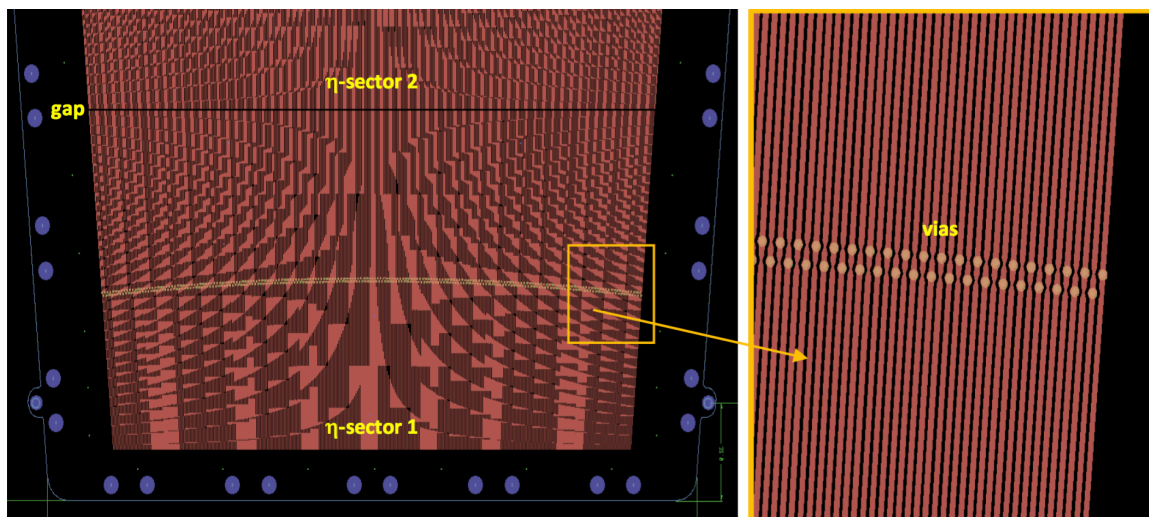


Figure 2.6: Picture of the strips of the readout board. "Hyperbolic geometric pattern is an artifact of the display on a screen" [8]. Strips are visible in the zoomed-in picture on the right.

The GEM Electronic Board (GEB)

The GEB supplies the data acquisition electronics and transmit generated data by the VFATs to the Opto-hybrid board. It also isolates the detector from the noise generated by the electronics. A 3D representation is depicted on figure 2.7 and a photo of a portion of the GEB with 2 VFATs on figure 2.8. As for the readout board, the GEB is divided in 24 sectors. For each sector, a hole let pass the male connector from the readout board and another male connector is present to send data to the Opto-hybrid board. The VFAT chip is connected to these two connectors.

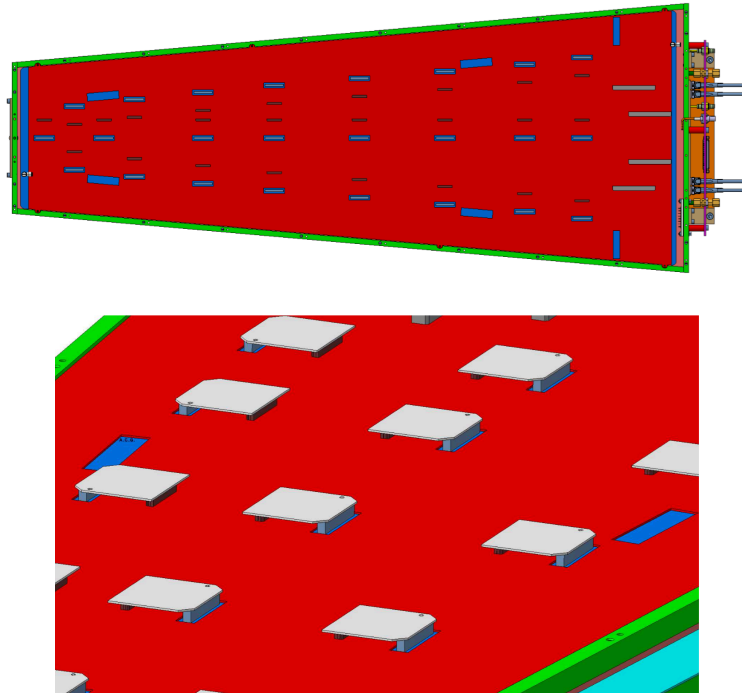


Figure 2.7: Top: global view. GEB(red) and the readout board (blue). Bottom: zoom in on VFAT chips[8].

2.3.2 Data Acquisition Electronics

Until now, particles passing through the detector has been transformed into an electrical signal. From now on, signals have to be measured, processed and transmitted. This is done by the data acquisition electronics (DAQ) which contains several components described below.

The VFAT Chip

The first DAQ component is the VFAT chip. Actually, the version 2 is used. However at the end, it will be the version 3 which is able to handle more higher data throughput. The inputs are the 128 strips from one sector of the readout board. As it can be seen of figure 2.9, inputs are first amplified, filtered and compared. The comparator compares the incoming signal with a defined threshold and produces a logic 1 or 0. This value is stored in the monostable (included in the "sync" block) during one clock cycle ($40MHz$) after being overwritten by the new result of the

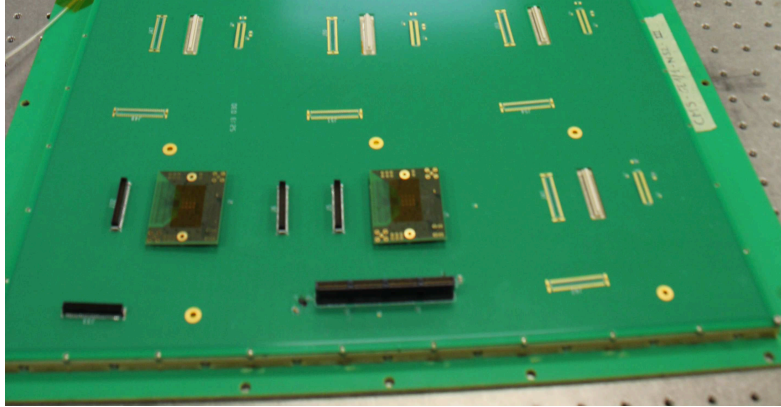


Figure 2.8: Picture of the actual GEB with some sector without connected VFAT that allows to see the male connectors [8].

comparator. All the operations are made in parallel for the 128 inputs. After the monostable, the path split in two. The first one is the trigger path that quickly detect a hit and send information at small granularity to the opto-hybrid board. The second path is the tracking path that send information at full granularity to the opto-hybrid board. Concerning the trigger path, the 128 bits are grouped in order to form 1, 2, 4 or 8 sectors. A fast OR is performed on each sector and produces what is called a *sbit*. Accordingly, a maximum of 8 sbits are generated and are sent to the opto-hybrid board at $320MHz$ in order to avoid congestion. Finally, for the tracking path, all the 128 bits are stored in the SRAM1 waiting for the CMS LV1 trigger received via the T1 command. If the event is judged relevant, corresponding data in SRAM 1 are sent to the SRAM2 waiting to be sent to the opto-hybrid board.

To retrieve the right data in the SRAM1 when LV1 trigger signal is received is not trivial. More informations can be found in the master thesis of Hugo Dewitte [6] which focused on the calibration procedure of the VFATs.

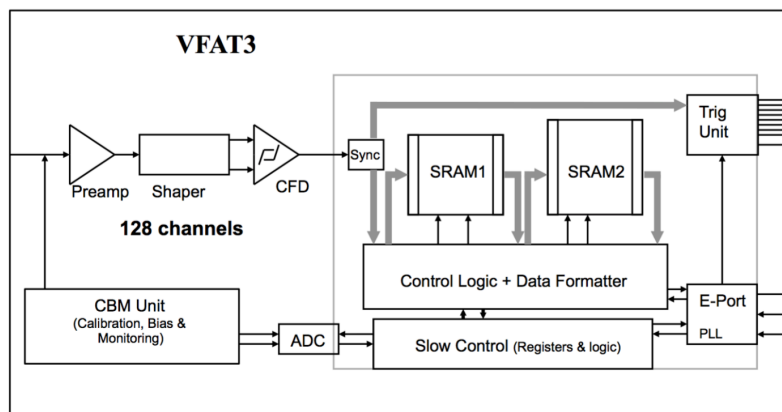


Figure 2.9: Block diagram of the VFAT chip version 3 [8].

The Opto-Hybrid Board

The opto-hybrid board makes the link between the VFATs and the off-detector electronics. It is located on the large base of the GEB and can be seen on figure 2.10. The role has to compress and synchronise the data from the 24 VFATs and send it on the outside via optical link. Moreover, it also transmit all the command signals to the VFATs. The different notable components on the opto-hybrid board are: an FPGA, four gigabit transceivers (GBT) and one 8b/10b module. Concerning the resistance to irradiations, the GBT with versatile links that are developed by CERN, can resist to a total deposited dose of 200Mrad which is much more than the expected dose obtained by simulations (around 1krad in this region of the detector).



Figure 2.10: Picture of the GEB, the 24 VFATs and the opto-hybrid in the bottom [6].

The FPGA

The FPGA is a programmable integrated circuit. It is a good compromise between performances and costs since the amount of chips needed is really limited (around 150). Unlike the GBT, the chosen FPGA for this project (Virtex 6) is not radiation tolerant by nature. Some studies have to be made in order to be sure the behaviour is correct in a harsh environment. Since the FPGA can be programmed a great variety of solutions can be considered in order to insure the correct behaviour when FPGA is operating. Since it is actually the main topic of this work, this will

be explained in details later in a next chapter. Another concern is to know if the FPGA will be able to work properly during 20 years since it will accumulate a dose around $1krad$. According to Xilinx, the company producing these FPGA, Virtex 6 should be capable to resist at a dose around $100krad$. Nevertheless, some test has been made within the scope of this work and will be explained later in another chapter.

The Gigabit Link Interface Board (GLIB)

The GLIB is the board that make the link between the off-detector electronics outside CMS (PC and other signals from CMS and the LHC) and the opto-hybrid. Moreover, it also makes a comparison of between trigger signals from the two chambers in order to filter the noise. The GLIB is located in CMS in a μTCA crate [9], a technology used in telecommunication that is highly redundant and hence, suitable for use in harsh environment.

Chapter 3

The radiation effects

In the previous chapters, the detector and its data acquisition electronics have been detailed. Subsequently, it is necessary to pay attention to its environment. Indeed, the energy and luminosity regime of HL-HLC leads to an extreme radiation environment [10]. As a consequence, in addition to the muons present in the muon tracker, a whole variety of particles are present such as neutrons, photons, electrons, etc. These particles are referred as background particles. Such background can degrade detector performances or even make it inoperable. Moreover, the data acquisition electronics is also concerned. A particular attention has to be made about the FPGA on the opto-hybrid board which is not intrinsically.

In this chapter, a description of the different physical phenomenons between particles and media will be made. Nowadays the electronics devices are tremendously complex, trying to describe the global behaviour with these phenomenons is not doable. Therefore, another approach at a higher level of abstraction has to be used. This will be explained on the second section of this chapter where the different misbehaviours are classified. Then, the cross-section will be introduced in order to have an insight of the electronic devices behaviour when parameters are tuned. Finally, a rigorous description of the GE1/1 project environment will be made.

3.1 The Physical Effects on Matter

An electronic device in a radioactive environment will suffer of several effects. An incident particle called *primary particle* penetrates in the material and makes collisions with its nuclei. It results in an exchange of energy between the incident particle and the material and a production of secondary particles if the energy is high enough [11]. Actually the real path of one particle into the material is scattered that can nevertheless be approximated as linear as shown in figure 3.1.

This linear path is the length of a cylinder where radius describe certain energy Δ . If the secondary particles have a greater energy than Δ , it will go out this cylinder. Otherwise it remains in it. In order to characterise a material in a radioactive element, several concepts are introduced. First of all, the *stopping power* (sp) of one incident particle in a material. This is the energy loss per unit distance in the material and the expression is :

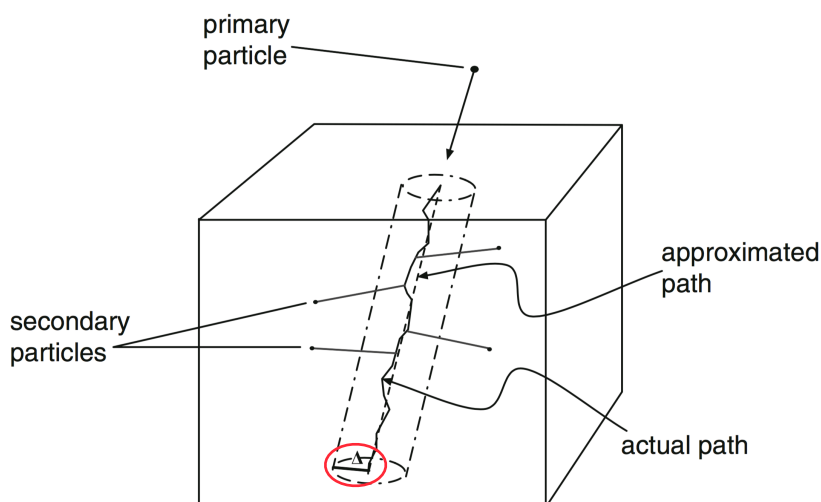


Figure 3.1: Penetration path of a particle in the material [11]

$$sp = \frac{-dE}{ds} \quad (3.1)$$

Moreover, it is quite straightforward to see this parameter is dependent of the material and particle. Another important parameter is the *Linear Energy Transfer* (LET) which is the transfer of energy from the incident particle to the material in the vicinity of the track by means of secondary particles emission. The expression is the following:

$$LET_{\Delta} = \frac{-dE_{\Delta}}{ds} \quad (3.2)$$

The LET only takes into account secondary particles with an energy smaller than Δ and then included in the cylinder in figure 3.1. If Δ tends to infinity, all the secondary particles are taking into account and the LET tends to the stopping power. In practice, LET can usually be approximated to the stopping power and plays an important role in error rate computation. LET is usually expressed in $MeVcm^2/mg$.

3.1.1 Energy Loss Mechanism

As explained in the previous section, the interaction is an exchange of energy. However, there are two major mechanisms that lead to an exchange of energy: the *ionisation* and the *bremstrahlung*.

Ionisation is the release of a charge in the material. It could be either *direct* or *indirect*. Direct ionisation means the incident particle is charged like heavy ions, electrons, positrons and alpha particles. When the incident particle goes through the material, it can take or give electrons to the atoms and provoking a movement of charge along the track particle. On the other hand, indirect ionisation is led by uncharged particles like neutrons and photons. Bremsstrahlung is a production of an electromagnetic radiation due to the acceleration of a charged particle when deflected by another charged particle. This phenomenon is mainly due to electrons with

ions. In case of high energy bremsstrahlung with pair production, an avalanche phenomenon could happen until their energies are not sufficient enough. These two mechanisms can produce different effect explained in the next sub-sections.

Funneling effect

The *funneling effect* is the result of the mechanisms previously discussed. As illustrated in figure 3.2. When an ionising particle penetrates into a semi-conductor device through the depletion layer, it creates a track of ionisation composed of electrons and holes. These charges create a distortion of the equipotential surfaces of the electric field in the depletion layer in a funnel shape and can extend into the substrate. It results in two opposite currents composed of electrons and holes respectively. If the collected charge is sufficient and sufficiently close to one of the nodes of the device, it may be collected leading to in a malfunction of the global circuit. This effect is temporary as the charge accumulation disappears when it is collected.

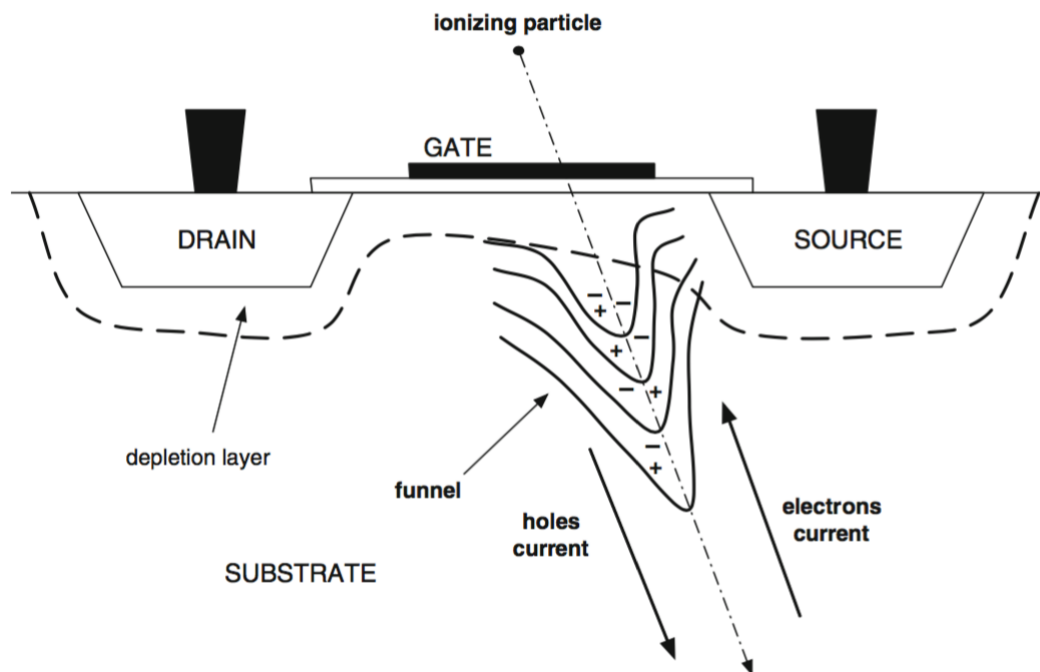


Figure 3.2: Illustration of the funneling effect in a CMOS transistor [11]

Displacement effect

When an incident particle goes through the material, it could damage the crystal lattice by changing the arrangement of atoms. It leads to an increase of the number of recombination locations and then a worsening of the analog properties of the junctions in the device. These damages are permanent but can be reduced by heating the device if there are not too important.

Charge Accumulation Effect

This effect is due to funneling and displacement effect combined. An accumulation can appear in the lattice defect and in the insulation oxydes. It results in a gradual degradation of the performances over time. Like displacements, the device can be healed by heating the device.

3.2 The Functional Effects on Devices

As already mentioned in the introduction of this chapter, studying the behaviour of an electronic device in a radioactive environment considering all the physical phenomena is too complex and time consuming since actual chips contain millions or even billions of transistors. Consequently, another approach has to be taken at a higher level of abstraction. Rather focusing on every particles and their interactions with matter, it is simpler to observe the internal misbehaviours induced by these particles. These misbehaviours are called *faults* and are classified in two categories, the Single event effects (SEE) and the Total Ionizing Dose (TID). Furthermore, a fault can propagate through the device and produce an *error* if it reaches an output. The misbehaviour is then external. Both of them can be mitigated but not at the same level and not with the same facility. Since faults are internals, the designer can apply some mitigation techniques on the logic. However, it is much more difficult to correct errors since it is a consequence of one or more faults and has to be done on the outside [11]. Some mitigation techniques for the faults are presented in the next chapter. On the other hands, no discussion will be made about errors.

In the next sub-sections, each type of fault are described with also a brief insight of when, in the development, these faults has to be a concern.

3.2.1 The Single Event Effects

The Single Event Effects (SEE) is the category that includes the effects due to one particule hitting the device. Furthermore, all the SEEs can also be divided into two subcategories regarding their effects: the soft errors and the hard errors. The soft are temporary and do not damage the device itself. The second ones are permanent and can produce damages.

Single Event Transient

A Single Event Transient (SET) occurs when a funnelling effect takes place. In case this event occurs close enough to one transistor node, the total charge is collected by the electric field and transmitted to this node. Whenever the charge is sufficient, a spurious variation (*glitch*) of the voltage can be seen at the output. This is a typical SET. The shape is described on figure 3.3. Several parameters describe the SET like t_{rise} and t_{fall} which describe the time needed to go from 10% to 90% of the maximal value V_{abs} and inversely. Since a SET is a double transition of the value, two kinds of SET can be defined. On one hand, the positive SET which is the double transition $\{0 \rightarrow 1 \rightarrow 0\}$ and on the other hands the negative SET which is the opposite i.e. $\{1 \rightarrow 0 \rightarrow 1\}$. Furthermore, t_{rise} and t_{fall} are dependent of the SET sign and could have an impact on the mitigation techniques. SETs are transient faults that last a very short time (around several picoseconds to several nanoseconds). Nevertheless the risk is it could be sampled by a memory

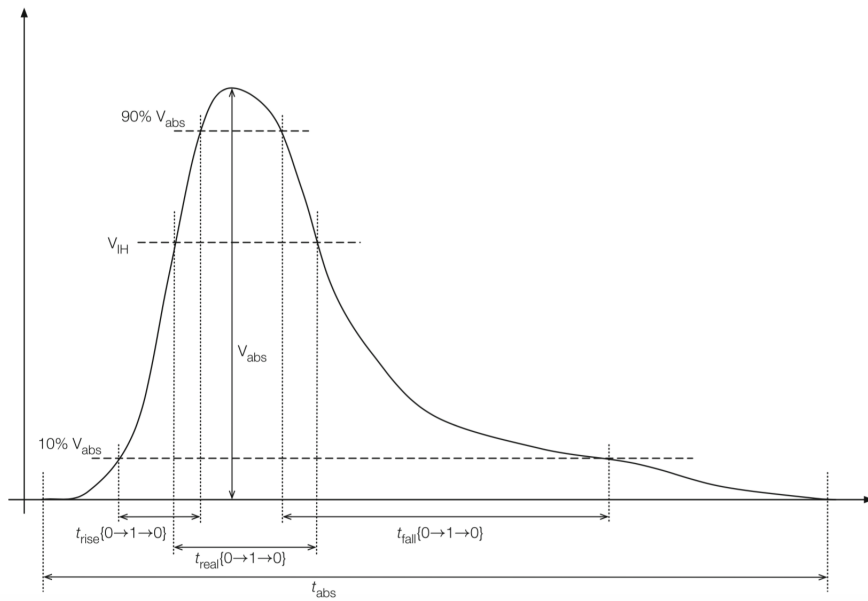


Figure 3.3: SET $\{ 0 \rightarrow 1 \rightarrow 0 \}$ shape and characteristic values [11].

element and by propagating into the device could lead to an error. This is particularly critical for clock and reset signals.

Single Event Upset

An event is considered as a Single Event Upset (SEU) when the value of a memory element is changed by an incident particle (also called *upset* or *bit-flip*). An SEU is a consequence of an SET sampled by a memory element. Due to the positive feedback loop present in a memory cell, the value is modified until a writing operation overwrite the value or a reset of the device is performed. As an example, the arising of an SEU in an SRAM cell in figure 3.4 which is the memory type embedded in Virtex FPGA is explained. The initial value stored is a logic 1 as depicted on the left side of the figure 3.4. If a SET is generated at the input of the inverter M2-M1, it could be propagated until the gate of M3 and force the transistor to open if the amplitude is high enough. As a result, the value at the input of inverter M3-M4 changes. Once the first inverter is flipped, due to feedback the second inverter is flipped too and the cell is in a new stable state which is the opposite of the initial one as depicted in the right side of the figure 3.4.

Multiple Cell Upset

An Multiple Cell Upset (MCU) is the same effect than SEU except that it upsets more than one memory element. It depends on the energy of the particule and the incident angle but also on the size of a cell and the distance between them. This is why latest devices are more sensitive to MCU than older devices. Although MCU are less frequent than SEU, it is less convenient to get rid of it since multiple errors could corrupt the same ressource. For instance, a word in a RAM. A simple way to correct data is to use one parity bit that can detect single upset except that it can not detect a double upset. In this case, to have a good mitigation against MCU, more

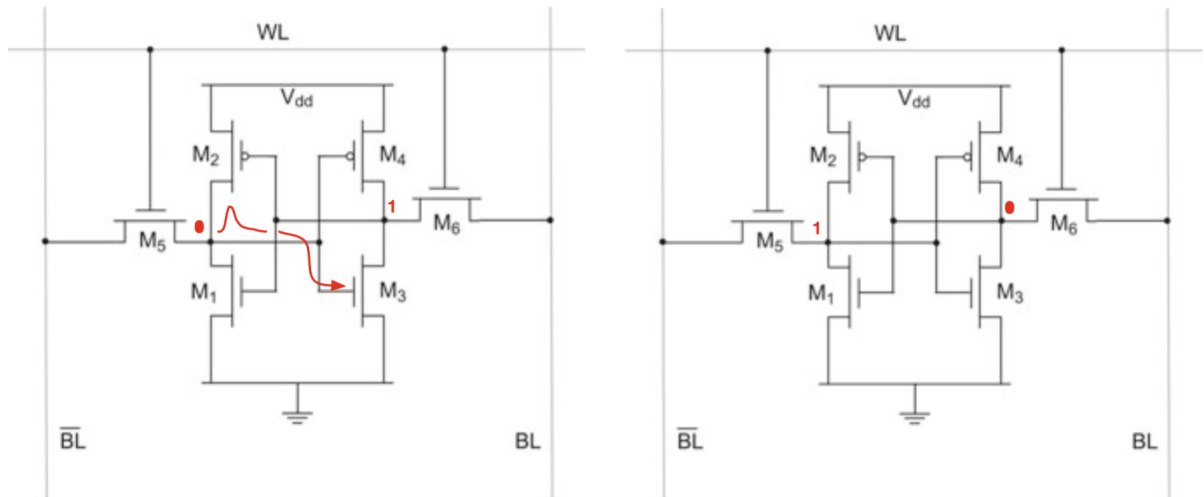


Figure 3.4: Bit-flip in a SRAM cell [11].

complex techniques have to be implemented. Moreover, it is not always possible to deal with MCU, i.e. it is making MCU a real concern.

Single Event Functional Interrupt

A Single Event Functional Interrupt (SEFI) is an error that affects an element of the device and results in a dysfunction of the whole device. These errors are impossible to delete unless by a global reset of the device or even a power cycle. A typical example is the corruption of a counter leading a state machine in a wrong state.

Single Event Latch-Up

All the previous events are the soft errors and can be resolved by the user in the case of FPGA. The last two events are the hard errors on which users can not act on it, it is then a concern for the manufacturers. Nevertheless, a small description is made for the completeness. The first hard error is the Single Event Latch-up (SEL) which is a permanent error leading to an increase of the current in the device due to positive feedback until the destruction of the device itself if it is not stopped fast enough. The only way to stop a SEL is to shutdown the device. Then it is important to try to avoid this effect by using a proper technology.

Single Event Gate Rupture

The last single effect can occur in non-volatile memories. It is not very relevant in our case because Virtex6 FPGA uses SRAM memory which is volatile. Single Event Gate Rupture (SEGR) is a destructive effect as its name suggests. When a heavy particle strikes the gate of the transistor while a high electric field is present. It may lead to the rupture of the gate. The reason is the track created by the particle is highly conductive and leads to the discharge of the capacitance between the gate and the bulk. This phenomenon can lead to excessive heating and then to the destruction of the device.

3.2.2 Total Ionizing Dose

The Total Ionizing Dose is an accumulation of charges inside the device due to radiations. It is then an effect depending on the time exposure, the flux of the particules and their LET. Consequently, it results in a global worsening of the device performances. The transistors slow down, the power consumption increases due to a higher leakage current, the sensibility of logic gate can varies asymmetrically, the threshold voltage in flash memory changes and may leads to a loss of reprogramming possibility and so on. Moreover, sensitivity to SEE can also increase and could lead to the presence of *stuck bits* which is a bit stored in a memory element with no possibility to be restored. Finally, the TID can be annealed by heating the device in order to restore the crystal.

3.2.3 Conclusion

Electronics devices in radioactive environment suffer from different effects. In the case of FPGA, some event are manufacturers concerns like SEL and SEGR whereas SET,SEU,MCU and SEFI are users concerns by modifying the original logic implemented on it in order to mitigate it against SEE. In this work, every SEE are called *upsets* without distinction. It is worth to notice that all the SEE are punctual events and can be erased in the case of soft errors. On the contrary, TID is a gradual effect and represent the lifetime of the electronic device. Nevertheless, some techniques in order to improve this lifetime are possible and will be explained in the next chapter.

3.3 The Cross-Section

The cross-section σ is a parameter that quantify the amount of upsets induced by a certain flux of particles at a certain energy and can be expressed as :

$$\sigma = \frac{SEU}{F} \quad (3.3)$$

Where SEU is the number of upset observed and F is the fluence, the flux integrated over time.

A typical curve is depicted on figure 3.5. This curve is not correlated with the FPGA used

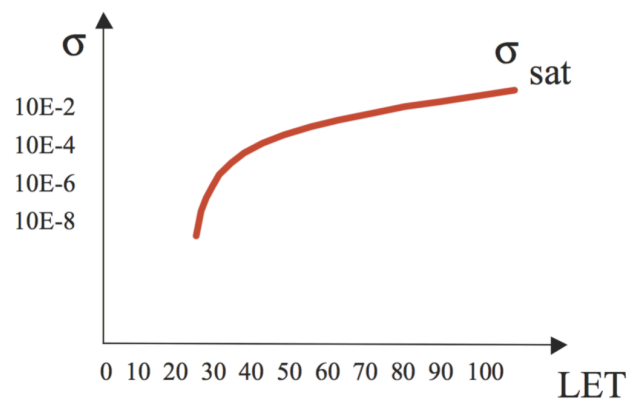


Figure 3.5: Example of a cross-section curve in function of LET [12].

in GE1/1 project but is used as an example in order to know the behaviour of the cross-section. Two comments could be made. On the one hand, there is a certain threshold LET_{th} below which the cross-section drop. Consequently, the number of upsets can be considered as zero. On the other hands, a saturation can be observed in the high LET values. The saturation value is dependent on the device but remains far below 1 as results shown in [13].

Since each ressource inside FPGA is different, one cross-section is assigned to each one of them and are unknown. Consequently, some tests have to be made in order to determine them. These tests were done in July and results are presented in a following chapter.

3.4 The CMS Environment

In order to perform some tests on the FPGA, it is important to know the environment inside CMS and more particularly at the muon chambers. As its name suggests, the sub-detectors in there are designed to detect muons. However, muons are drowned into parasitic particles such as neutrons, photons, electrons, etc. All these unwanted particles form the *background* of the CMS cavern. Unfortunately, the location of GE1/1 project highly suffers from these backgrounds where the neutrons are the main contribution with a range of energy is between the thermal region and a few GeV as it can be seen on figure 3.6. The origin of these neutrons are the

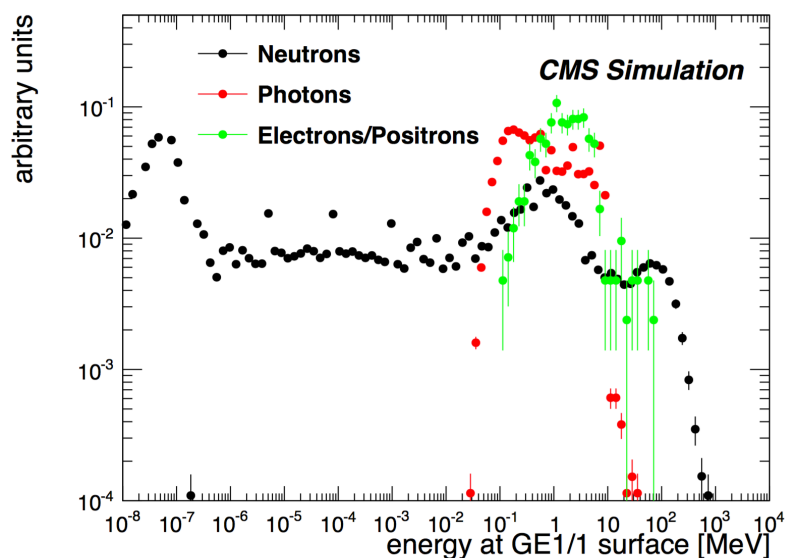


Figure 3.6: Energy spectrum of the incident particles at the GE1/1 location. Results obtained by means of FLUKA simulations [8].

interactions of hadrons produced during collisions of the beams with the beam pipe, the beam collimator, the shielding and the HCAL. Furthermore, neutrons interact with matter producing secondary particles which is a part of background as well. Consequently, some simulations has been made with the FLUKA simulations [14] which takes into account the CMS geometry and an instantaneous luminosity of $5 \times 10^{34} cm^{-2} s^{-1}$ which is the planned value during phase 2 of LHC. The available results are a plot of the flux of the different particles at the location of

GE1/1 project and also a map of the deposited dose. Both results are visible on figure 3.7 and 3.8 respectively. Furthermore, the opto-hybrid board is located in R around 240 and 220 cm. Then, a flux around $5 \times 10^4 \text{cm}^{-2} \text{s}^{-1}$ and a deposited dose around 10Gy (1krad) are expected. These value will serve as reference for the tests of the FPGA detailed in chapter 4.4.

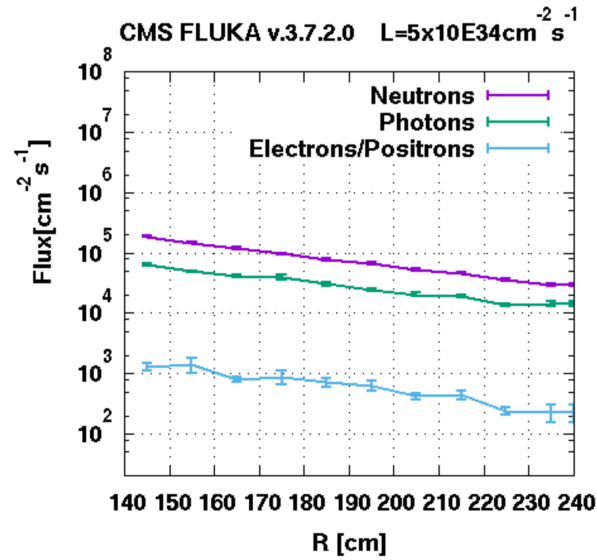


Figure 3.7: Flux estimations of neutrons, electrons/positrons and photons at the location of GE1/1 project. Results obtained by means of FLUKA simulations [14].

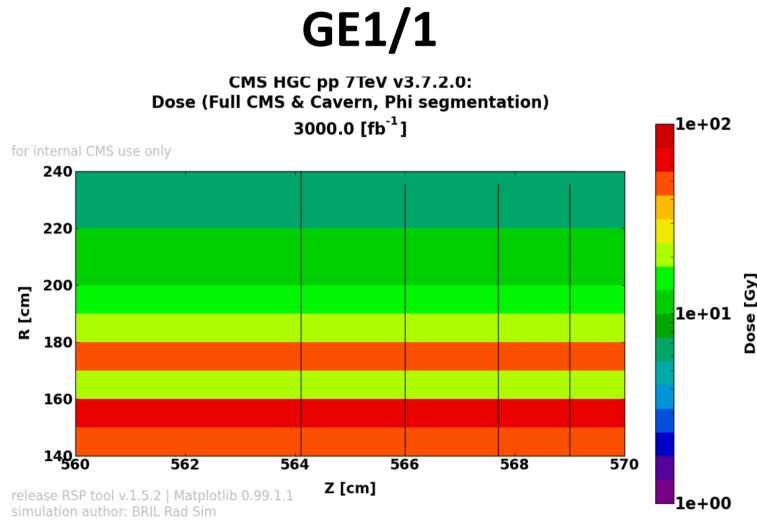


Figure 3.8: Absorbed dose estimation for an integrated luminosity of 3000fb^{-1} at the location of GE1/1 project. Results obtained by means of FLUKA simulations [14].

Chapter 4

The Mitigation Techniques

In this chapter an overview of the different possibilities to mitigate the logic of the FPGA will be made. Three categories can be defined [15]. The first one is the detection fault methods. As the name suggests, these methods only detect and inform that a fault is present in the logic and a reconfiguration is needed. The second category contains the reconfiguration methods which are used when an accumulation or uncorrectable faults are present inside the FPGA. The last category contains the repairs fault methods which are used when logic inside FPGA are unavailable due to TID. These researches were made at the beginning of this thesis. At this moment, it was not known if the FPGA will need such methods or not. After the tests performed at the end, it reveals the FPGA is perfectly capable to handle much larger TID than those predicted inside CMS. Therefore, the repairs methods will only be briefly presented for the completeness. After that, the chosen methods will be explained in more details. Others tools can detect and correct upsets in other resources of the FPGA like the BRAM and the configuration memory. These will be detailed in the next chapter with the description of the firmware developed for the tests performed at the end of this thesis.

4.1 The Fault Detection Methods

The fault detection methods are techniques used to detect faults inside the logic. Consequently, it does not performs any correction of the configuration memory which is made by the reconfiguration methods. As it can be seen on figure 4.1, a whole variety of methods exists. Each of these methods have their own benefits and drawbacks. In general, speed of detection and high granularity is at the cost of resource usage and performances. However, the classification presented here describes a family of methods and not a cooking recipe to its implementation. Indeed, depending on the logic that has to be mitigate and its specificities, the designer could implement it in a slightly different way than for another logic or even perform a mix of different techniques.

Concurrent Error Detection (CED) This family of methods verifies if the results of the logic are correct during normal operation [16]. It can be intrusive or non intrusive which means the original logic is modified or not. An example of generic non intrusive CED schematic is depicted in figure 4.2 where the CED circuit is plugged to the original one and listens

Method	Speed of detection	Resource overhead	Performance overhead	Granularity	Coverage
Modular Redundancy	Fast – As soon as fault is manifest	Very Large – Triplicate plus voting logic	Very Small – Latency of voting logic	Coarse – Limited to size of module	Good - All manifest errors are detected.
Concurrent error detection	Fast – As soon as fault is manifest	Medium – trade-off with coverage	Small – Additional latency of CRC logic	Medium – trade-off with resource	Medium – Not practical for all types of functionality.
Off-line BIST	Slow – only when off-line	Very small	Small – Slight start-up delay	Fine – Possible to detect the exact error	Very Good – All faults including dormant.
Roving (Segmented Interconnect)	Medium – order 1 seconds	Medium – Empty test block plus test controller	Large – Clock must be stopped to swap blocks. Critical paths may lengthen	Fine – Possible to detect the exact error	Very Good – Multiple Manifest and latent faults are detected.

Figure 4.1: Overview of the different detection methods [15].

the inputs and outputs. As the output can be predicted with the knowledge of the input, a comparator is used to compare actual and predicted output. Consequently a flag informs if an error occurred or not. Moreover, CED is an online fault detection method, the original circuit is working without interruption.

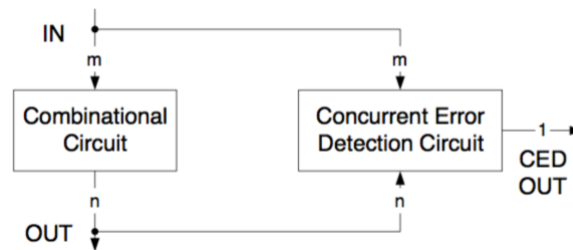


Figure 4.2: Schematic of non intrusive CED [16]

Off-line fault detection / Built-In Self-Test (BIST) The BIST methods are a family of methods capable to detect errors without any external device such as CED methods [17]. On the contrary, such methods perform detection off-line which means the FPGA must be in a test mode and is then not operational. According to [17], duration of test mode is around 1 second. Since the FPGA in GE1/1 project receives input at 40 MHz from the VFATs, such methods are not possible.

Roving The roving methods try to improve the BIST methods by reducing the duration of the test mode. In order to do that, the FPGA is divided into regions and one region at a time will be in test mode. Moreover, to avoid a too long interruption, the region which is in test mode is copied into another empty region. Therefore, the FPGA can continue to work and only encounters small interruptions around $25\mu s$ [15]. In conclusion, as for BIST methods, Roving is not well suited for this project.

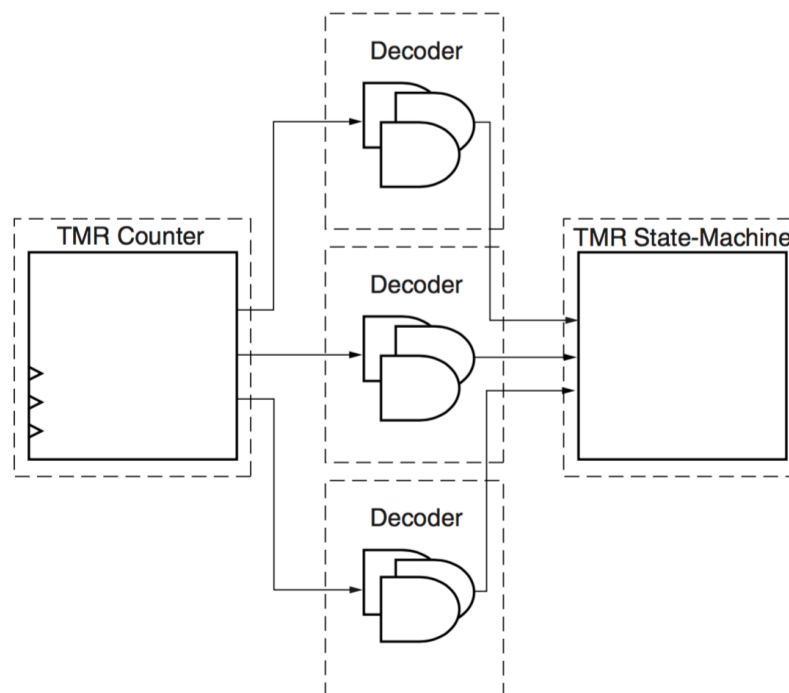
Triple Modular Redundancy (TMR) The principle of TMR is quite simple. The different elements composing the logic can be triplicated and connected to a majority voter which is able to detect one fault and correct it. This is a major difference comparing to the other methods that only detects faults. Moreover, TMR is also the fastest, generally one clock

cycle is required to perform the vote. On the other hand, the resource overhead is very large. Indeed, the final logic is more than 3 times bigger than the original. Fortunately, a lot of resources are still available in the FPGA for this project. For all these reasons, TMR is the chosen technique to mitigate logic. In the next section, more details on this methods are explained in order to provide a better understanding.

4.2 The Triple Modular Redundancy in Details

The TMR method is a mitigation technique acting against SEU and SET. For the ASICs, only the flip-flops have to be triplicated because SEU only affects memory elements and path logic is not reconfigurable. Regarding the FPGA, almost all the logic is reconfigurable. Then a complete triplication of the design is needed in order to have an efficient mitigation. In the ideal case, no single point of failure exists anymore. In the reality, this could be a very complex task. Depending on the logic type, TMR is not implemented the same way. Basically, it can be classified in four categories [18]: throughput logic, state machine logic, I/O logic and special features. The whole design is generally a succession of these categories.

Throughput logic also called combinatorial logic. The signals go through the logic without performing a loop. In other words, the output never depends on the previous inputs. Thus, it is very simple to apply TMR on it. Logic has just to be triplicated as shown in figure 4.3. If a corrupted bit appears, there is no reason the next one will be corrupted too.



XAPP197_11_030801

Figure 4.3: Schematic of triplicated throughput logic between two state machines in a logic module [18]

State machine logic also called sequential logic. Here, a feedback is present, i.e. outputs

depends on the previous inputs. A simple example is a counter as shown in figure 4.4. State machine is much more complex to triplicate since if the output is corrupted, the next ones will be corrupted too. One possibility is to add a unique majority voter after the replication as depicted in figure 4.5. Consequently, if one state machine is corrupted, the final output is still correct. However, this state machine will stay corrupted since there is no correction logic on the feedback. Consequently, if another state machine becomes corrupted, the output of the majority voter will be permanently corrupted until a global reset of the FPGA is performed. Another way to perform TMR in order to avoid such kind of problems is shown in figure 4.6. Here, the feedback is located after one majority voter replication and single point failure is completely avoided. As a result, such logic will not fail unless two simultaneous upsets arise.

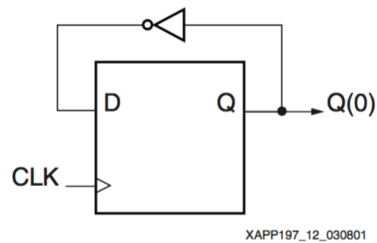


Figure 4.4: Schematic of an unmitigated bit counter with the feedback visible [18].

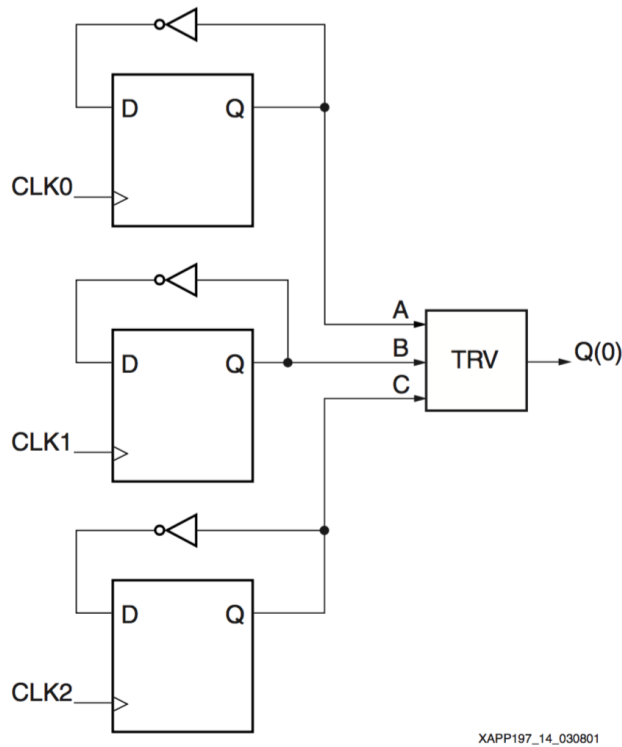


Figure 4.5: Schematic of a triplicated bit counter with a unique majority voter [18].

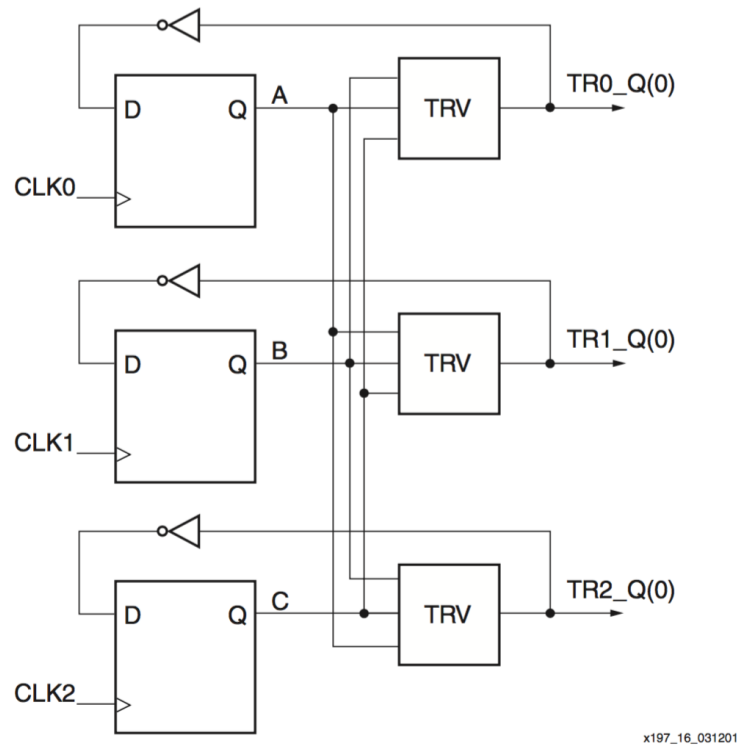


Figure 4.6: Schematic of a triple voted Triplicated counter [18].

I/O logic Inputs and outputs of the FPGA can also be triplicated. About the inputs, nothing special has to be done inside the FPGA but the signal must be triplicated before and connected to the pins. Regarding the outputs, such circuit is depicted in figure 4.7. In order to check a last time the signals before leaving the FPGA, a minority voter is added to each output and corrects a potential error avoiding single point of failure.

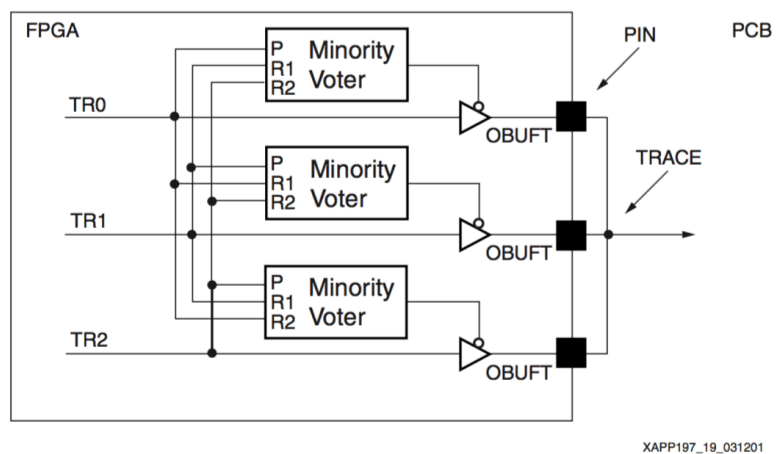


Figure 4.7: Schematic of a triplicated output with minority voters [18].

Special features Some specific features like BRAM, distributed RAM, shift registers, etc. can also be mitigated. Since these methods are specific for each one of them, it will not be explained here.

4.2.1 The Implementation of TMR

Once the TMR method has been described, the implementation has to be done on some logic. As already explained in chapter 2, VFATs send two types of data to the FPGA: the trigger data and the tracking data. Each one could be triplicated but since trigger data integrity is essential, this is the chosen logic to test TMR. In short, the 24 VFAT send 8 sbits to the FPGA at 40 MHz representing a bandwidth around 7Gbps which is a concern regarding the available bandwidth of the Gigabit Transceiver (GTX) at the output of the FPGA. Since most of the sbits will be equal to zero (simulations predict only few detected particles per clock cycle), a data compression can be performed. Rather than transmit every sbits, those set to one (with a maximum of 5) will be encoded on 8 bits. 5 bits for the VFAT ID and 3 bits for the sbit ID. Accordingly, data are reduced from 192 bits to 40. This piece of the firmware is called *zero suppression*.

In the framework of this thesis, the zero suppression module has been written in VHDL. The basic principle is the following, the five first bits set to "1" are selected and then encoded. In order to perform it, a counter with a variable (and not a signal) is used. The result was an indivisible module which is throughput logic. Thus, triplication includes 3 independent replications. Since the the opto-hybrid is already designed, the output pins of the FPGA can not be triplicated. Accordingly a majority voter is used at the end of the triplication. In Addition, the voter do not compare the whole outputs but performs a vote bit to bit and in parallels in order to increase the robustness. Indeed, in this case, two simultaneous upsets on two different bits will not lead to a corrupted output.

4.3 The Reconfiguration of the FPGA

In addition to the fault detection methods, reconfiguration of the FPGA has to be performed. Indeed, some upsets are uncorrectable and their accumulation could lead to a complete failure of the logic. The simplest way is to reconfigure the whole FPGA losing every data stored in RAM. A second possibility is the *scrubbing* [19] that performs a reconfiguration of the logic and keep data stored in RAM. A last possibility is to use the partial reconfiguration capability [20] reducing the time requirement for the reconfiguration.

In order to choose one method, it is important to know the time requirement imposed to the FPGA. Regularly, a global reset of all sub-detectors is performed. Moreover, the FPGA is reconfigured by the use of a flash memory placed on the opto-hybrid board which allows reconfiguration in a few microseconds. As a result, a simple reconfiguration of the whole FPGA should be sufficient. Nevertheless, a computation of the upset rate has to be performed in order to settle it.

4.4 The Fault Repair Methods

Finally, the fault repair methods are introduced. These methods have to be considered when resources in FPGA are unavailable because of TID and have to be avoided. Reconfigure the FPGA with the same bitstream will not change anything as the problem is not an SEU in the configuration memory. At the end of this master thesis, it reveals the FPGA can handle much higher TID than the dose predicted inside CMS until the end of operations. Then, these methods are not necessary but for completeness, they will be briefly introduced.

An overview of the different fault repairs methods with their benefits and drawback is presented on figure 4.8. All of these are based on the reconfiguration capability of the FPGA and unused resources. Basically, the FPGA is divided in clusters. Some methods like cluster reconfiguration reconfigure differently a cluster when a fault is detected. In addition, pebble shifting defines spare clusters that will be used when another cluster is faulty. This can be performed at different granularities depending on the amount of unused resources. The other methods are not explained here but are similar to the previous one with some variance in the way to split the FPGA logic and the spare resources use.

Method	Fault Pattern Tolerance	Resource Overhead	Performance Overhead / Degradation	Complexity of Repair
Hardware Level	Poor – Limited number and distribution tolerated.	Medium	Low overhead and little degradation	Low – Transparent to configuration
Multiple Configurations	Poor – Limited number and distribution tolerated. Interconnect tolerance causes complexity	Low – Uses naturally spare resources, but requires ROM for configurations	Low – Each configuration can be fully optimised	Medium – Selection and loading of configurations
Chain Shifting	Poor – Limited number and distribution tolerated. Poor for interconnect	Medium – A set of interconnect must be reserved	Low – Alternative routing is pre-determined	Low – Alternative routing already reserved
Pebble Shifting	Medium – Relies on nearby spare PLBs	Low – Uses naturally spare resources	Medium, rerouting causes uncertainty	High – Re-routing necessary
Cluster Reconfig.	Poor – Reliant on spare resource in cluster. Poor tolerance in interconnect	Low – Uses naturally spare resources	Low – Changes only local interconnect, slight uncertainty	Medium – Analysis of logic, no re-routing
Cluster Reconfig.+ Pebble Shifting	Good – Flexible solutions possible.	Low – Uses naturally spare resources	Low – Usually a fast alternative will be found, medium uncertainty	High – Analysis of logic and rerouting
Evolutionary	Good – Implementation is completely flexible	Large – Configuration grading and storage	Variable – Solution is arrived through random mutations.	Massive – May take a long time to repair

Figure 4.8: Overview of the different repair methods [15].

Chapter 5

Test beam description

During this master thesis, a firmware was developed in order to test the FPGA in real conditions. In this chapter, an explanation about the test principle is made and followed by a rigorous description of the test bench used during the actual tests. Finally, the developed firmware is detailed

5.1 FPGA Irradiation Tests Principle

The on-detector FPGA of the GE1/1 detectors will have to operate for almost 20 years in a harsh environment mainly composed of neutrons which flux will eventually reach around $50kHz/cm^2$. It is therefore mandatory to test the FPGA radiation robustness and the efficiency of the mitigation techniques in order to be sure the GE1/1 detectors inside the CMS experiment will work correctly. The most accurate and realistic test would be to put the device directly inside CMS. However this is not straightforward. First of all the access to the CMS detector is very limited while the LHC is in operation. Secondly one can not wait for several years before making the choice of technology. Obviously the tests have to be performed in another facility allowing to speed-up the irradiation process. This technique is called *accelerated radiation testing* [11]. The principle is to use a particle accelerator to produce a flux of particles several orders higher than in the real conditions inside CMS with an appropriate energy and put the FPGA through the beam particle to produce upsets in it. This testing method is the chosen one for the FPGA of GE1/1 and test were performed with *CYCLONE110*, the cyclotron of Louvain-la-neuve [21]

5.2 Description of the Test Bench

The opto-hybrid board in the yellow box *A* in figure 5.1 and 5.2 with the irradiated FPGA indicated by the purple arrow A_{bis} are put into the irradiation chamber. These devices are duplicated in order to have more data to analyse. However, since one is in front of the other, the penetration range of the beam, which is proportional to the energy, has to be large enough in order to produce upsets on the second FPGA too. During the test it was observed that at an energy below $40MeV$ no more upsets are observed on the second FPGA. In the back in the blue box *C* in figure 5.1 is the cyclotron extremity where the beam comes out. A shield is placed in

front of the opto-hybrid boards and is visible in orange box on figure 5.2. The shield is a metallic plate with a hole of $80mm$ diameter at the center. The beam homogeneity after the shield is $\pm 10\%$ [21]. Since the FPGA dimensions is a square of $35mm$ side, one can consider the flux constant all over the device. The cyclotron produce a beam at $65MeV$, in order to lower the energy, some degraders are used. These are the white circle in red box D in figure 5.3. After the degraders is the ionization chamber used to compute the beam intensity. With a combination of those, the energy at the end can be fixed in a range between $10MeV$ and $62MeV$ [21]. About the particles flux, it can be tuned from $50part/s.cm^2$ up to $2 \times 10^8 part/s.cm^2$ [21] which is around 3 order higher than the expected maximum flux inside CMS [14]. Then a total deposited dose of $10krad$ can be reach in a few minutes and a significant amount of upset can be observed in order to have relevant statistics to compute the different cross-sections. It is worth to notice that here, the particles are protons and not neutrons like in CMS but since the upset susceptibility of these two particles at energies higher than $20 MeV$ are equivalent [22] the use of protons for testing is adequate.

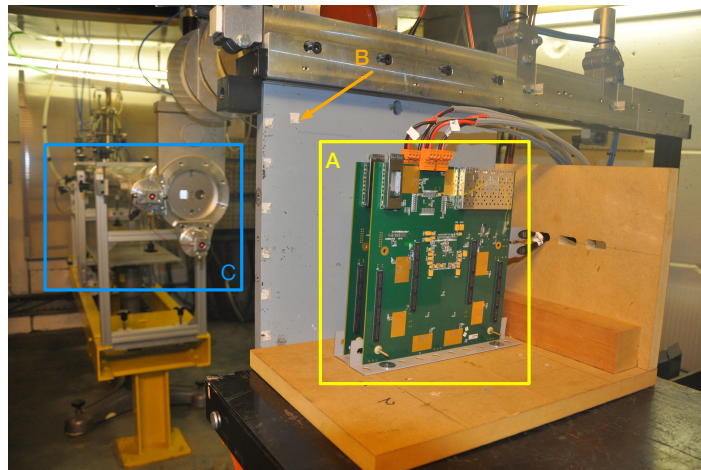


Figure 5.1: picture of the device under test inside the irradiation chamber

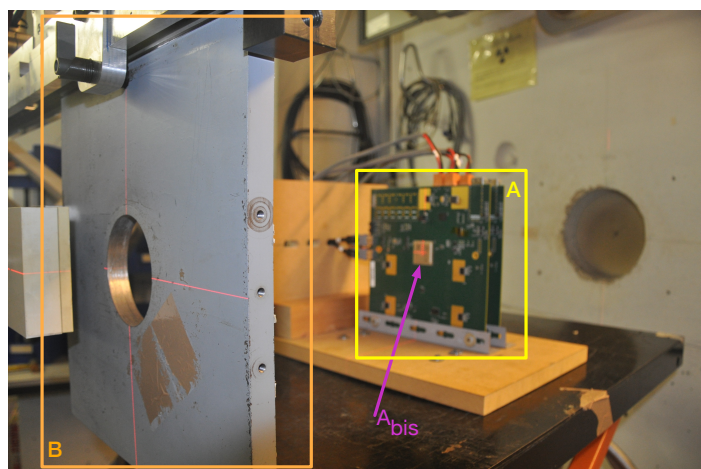


Figure 5.2: Another viewpoint of the device under test with the shield more visible

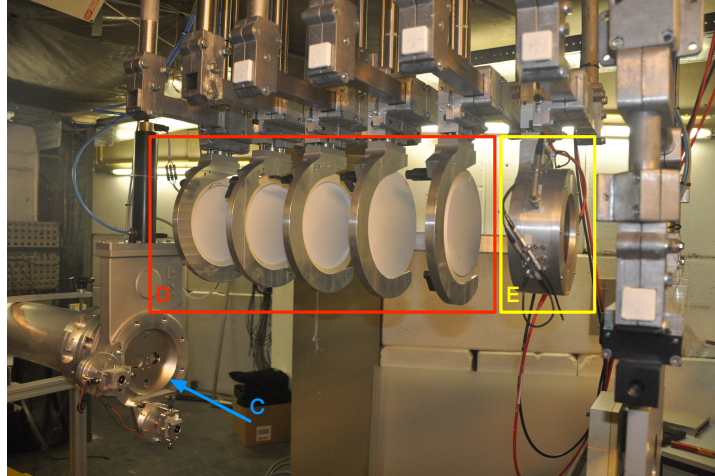


Figure 5.3: Focus on the source of the beam, the degraders and the ionization chamber

Outside the irradiation chamber are the two opto-hybrid with the reference FPGA that collects data from irradiated FPGA and are indicated by the blue box 1 in figure 5.4. Each

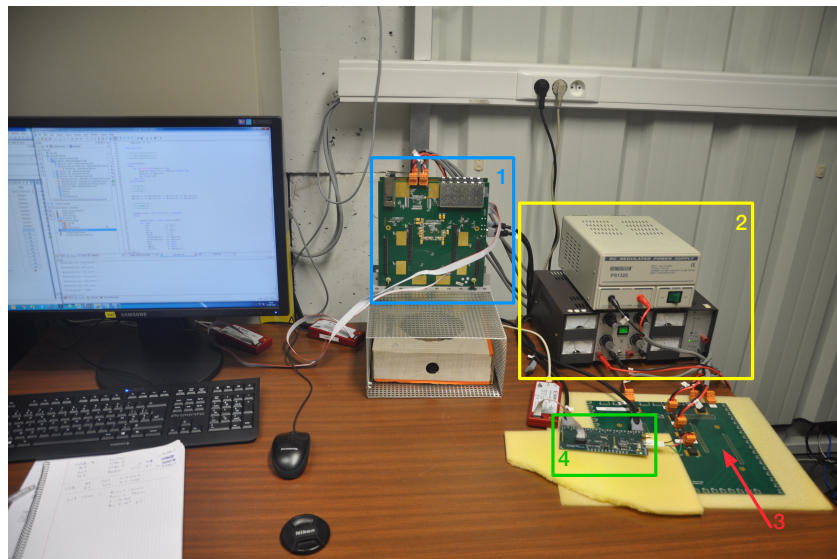


Figure 5.4: Test bench with the power supplies, the reference FPGAs and the computer

reference FPGA is linked via an HDMI cable to one irradiated FPGA. Power supplies are marked by the yellow box 2, it provides power to the power board which supply the opto-hybrid board and then the FPGA. The power board for the reference FPGA is indicated by the red 3 while the power board for irradiated FPGA is inside the irradiation chamber next to the irradiated FPGA. However, this is non visible on figure 5.1 and 5.2. In the green box 4 is the *mezzanine board*. It allow to reconfigure the irradiated FPGA by pushing a button manually. In addition, to have a better understanding of the whole setup, one can see the schematic on figure 5.5 where the opto-hybrid boards are not duplicated in order to keep it clear.

Some components are not visible on the photos but are represented on the schematic. First,

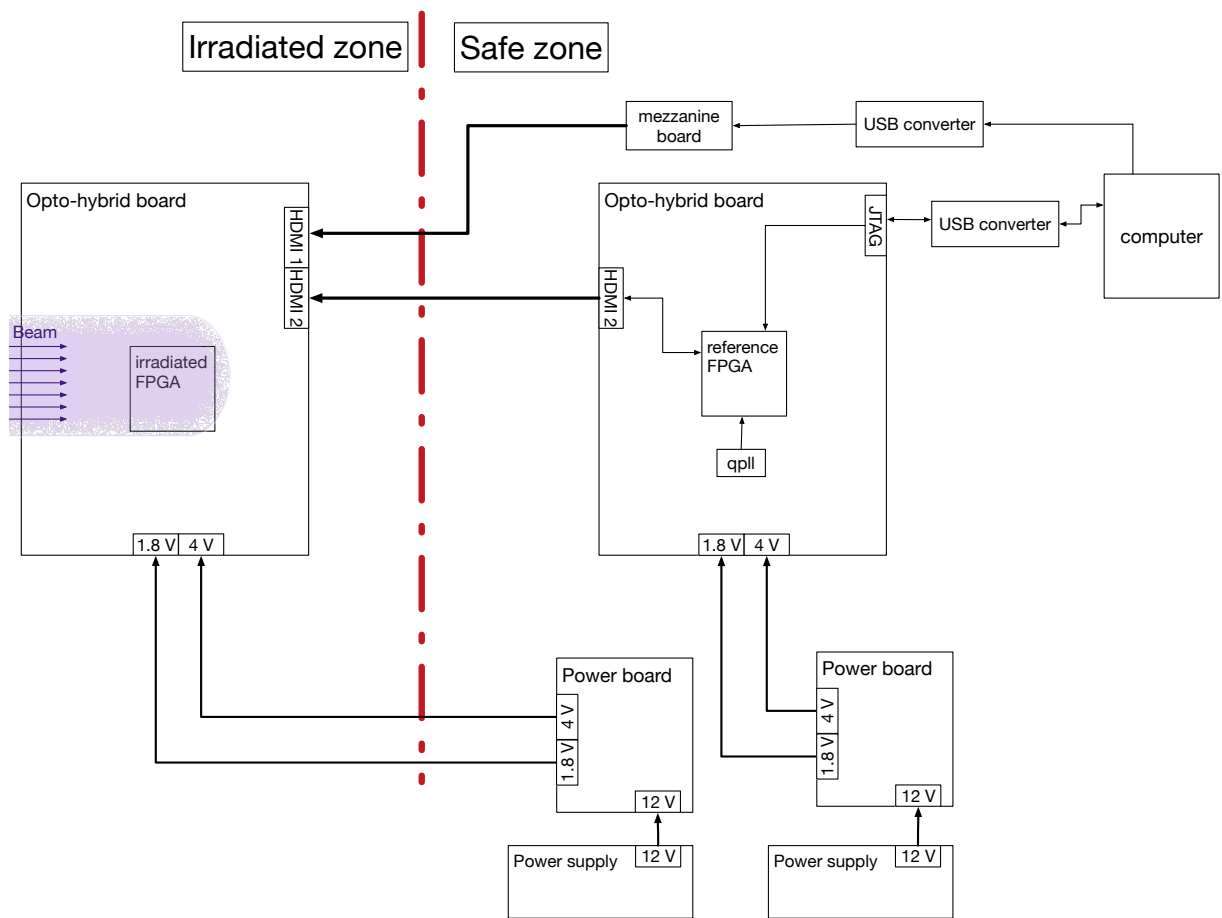


Figure 5.5: Testbench schematic

the flash memory on the opto-hybrid board in the irradiated zone. This memory stock the bitstream for the irradiated FPGA reconfiguration and allows to perform in a few millisecond instead of several seconds. Secondly, the qpll which provide a clock at $40MHz$ for the reference FPGA. This clock is also sent to the irradiated FPGA via the HDMI cable. Finally, the USB converter that is used to communicate between the computer and the FPGA. The computer is used to save and check data collected by the reference FPGA.

5.3 Description of the Testing Firmware

To actually compute the different cross-section, the firmware on the irradiated FPGA should use each tested resources as much as possible: CLB, BRAM and configuration memory in this case. This is called the *tested logic*. On the other hand, there is the *control logic* which is used to extract data from tested logic and send it to the reference FPGA. This logic has to remain as small as possible in order to avoid upsets as much as possible. The different part of the firmware and their modules are explained below.

5.3.1 The Irradiated FPGA

Configuration memory For the configuration memory, the IP core *soft error mitigation* (SEM) is used. This IP core is able to detect and correct a single error by the use of a parity bit and detect only double error which are located in configuration memory. A bunch of output flags inform what the SEM is currently doing SEM scan continuously the configuration memory, flag *observation* is set to one. When an upset rose, there is a certain delay before the detection. This delay depends on the working SEM frequency and is around a few millisecond. When the upset is detected, SEM enter in correction mode and set the corresponding flag. This takes a few milliseconds too. At the end, if the error is correctable, SEM return in observation mode and the FPGA can continue to run properly. If not, it set the flag uncorrectable and a reconfiguration of the FPGA is needed. Other flags exist but are not used in this firmware. Then it is not explained here. In this firmware, the instantiation example available in the IP file was taken. Since this module is autonomous, all the flags are simply linked to the transmitter module described below.

BRAM To instantiate the BRAM the IP core *block memory generator* is used. In order to maximise the amount of upset in it, around 98% of the available BRAM are used. The words are 64 bits long and the address is 17 bits long which lead to 131072 addresses. This IP core offers the possibility to add a *Hamming error correction capability* (ECC) which can detect and correct single error and only detect double error. It advises the user in case an error is detected during a reading operation by setting to one the corresponding flags *sbiterr* and *dbiterr* as it can be seen on figure 5.6. All the "A" ports are for writing

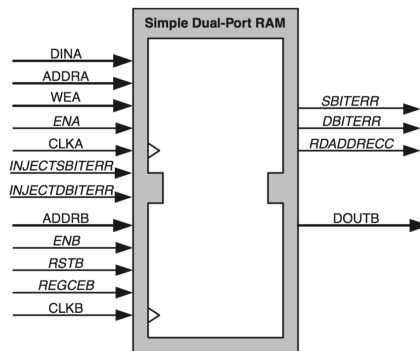


Figure 5.6: Simple dual port BRAM schematic [23]

operations and "B" ports for reading operations. There are two last input ports which are *injectsbiterr* and *injectdbiterr* that can be used during debugging phase in order to inject single or double error inside the BRAM at a chosen address by the use of *addr a*. These two ports can be removed by deselecting them in the configuration IP core window. It is also worth to notice that the ECC module correct data for single error at the output and not the data inside the BRAM. Hence in this firmware, a small module rewrite data at the corresponding address in BRAM when errors flags are set. It is possible in this case to correct double error since data inside BRAM is "0101..." for even addresses and the complementary one for odd addresses. The choice of such pattern was made in order to have an equal amount of zeroes and ones to try to know if there is a difference between

upset susceptibility between them. The constrain with the use of ECC is that one can not use an initialisation file in order to load the BRAM. Then a module was written the generate the data and was implemented on the reference FPGA to be preserve it from upset.

Finally, you can see the protocol for the BRAM on the figure 5.7. The FPGA is (re)configured, the BRAM content is received from the reference FPGA. Once it is loaded, reading operation can start and will go on until reconfiguration is performed by the user. If an error is detected, the address content is send to the serialiser in order to send it to the reference FPGA. At the same time, data are corrected inside the BRAM. After that, the address is incremented and a new address is readed. Since there is no critical error that can occur in this firmware part, the FPGA will never be reconfigured because of it.

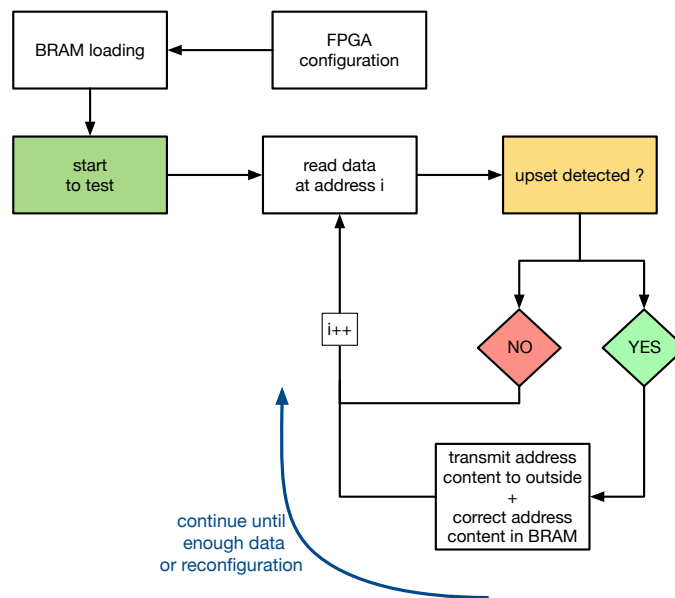


Figure 5.7: Protocol for the BRAM use during test beam

CLB For the CLB, the zero suppression module is used. As it can be seen on figure 5.8 this module is triplicated a first time in order to mitigate it. With the first level of replication one can have access to the output of each replicated module and the output and the flag of the majority voter (MV). This flag is set to one when there is a mismatch between the inputs of MV. The level two of replication triplicates a second time in order to have additional data to extract. Indeed, with only the first level of replication, it is quite tricky to know if the MV output is correct or not since two identical errors in two inputs lead to a wrong output and a flag that remains to zero. With the second replication, the flag from level 2 MV directly informs if the level 1 MV has worked well or not. In the test firmware, it was planned to replicate the whole thing N times in order to fill the FPGA as much as possible. Unfortunately, the synthesis tool XST has some limitations in the optimisation process maybe due to congestion. One replication of the design as in figure 5.8 takes around 10% of the LUT available but two replications use 115%. After some efforts to

resolve this issue we were unable to reduce the amount of used resources and only one replication was used during the test. It means that when an error in configuration memory occurs, there is only 10% of chance that there is an impact on the logic. This is not optimal but still doable. Another parameter that encounter limitation is the working frequency. Since the logic is quite complex, the critical path, hence the maximum frequency of the design is reduced. At the end, a maximal frequency of $20MHz$ has been reached instead of $40MHz$ which is the frequency of the inputs coming from VFATs.

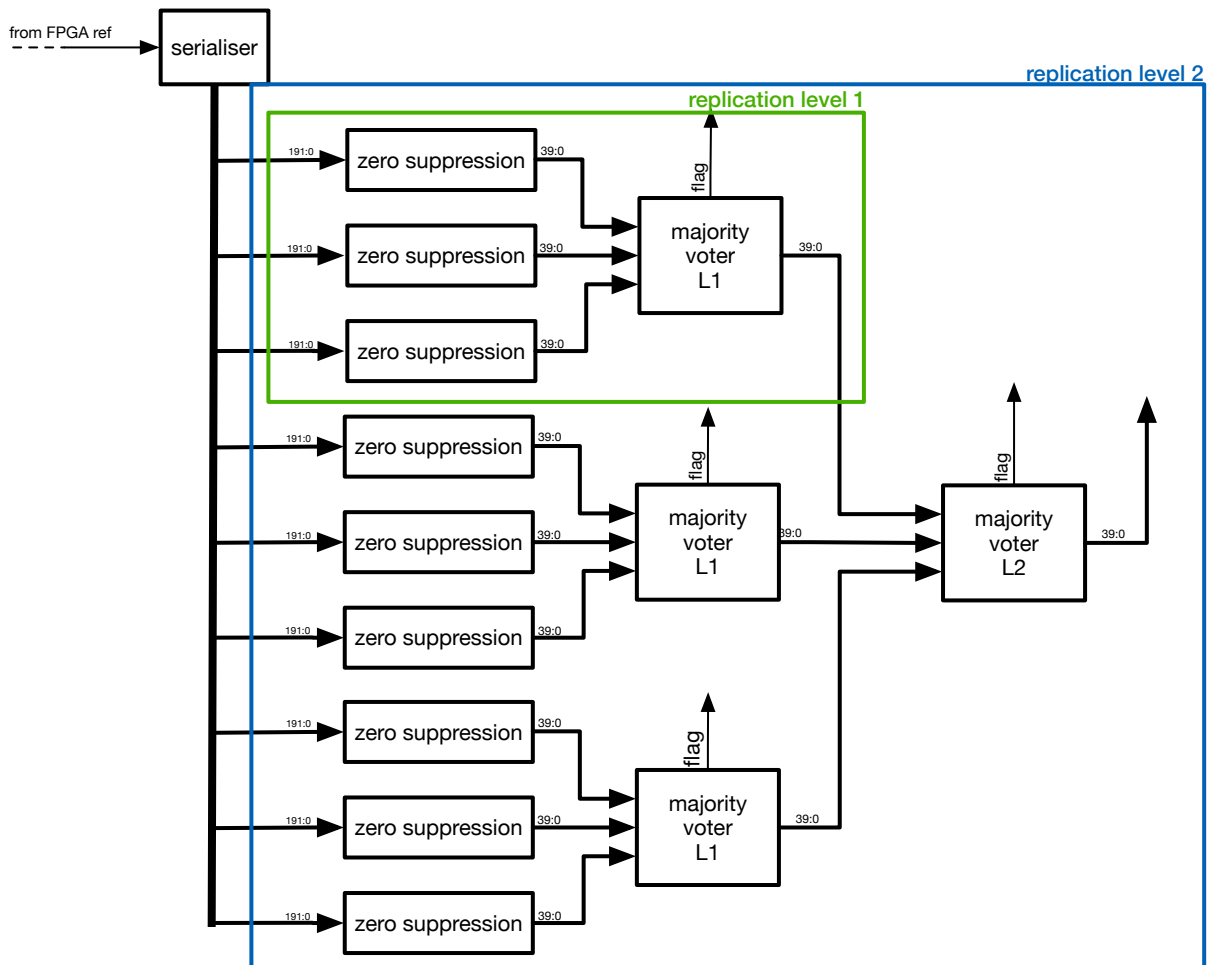


Figure 5.8: Design implemented on the FPGA for the test beam

The transmitter Every other modules described before send data and flags to the transmission module. This one, is the biggest module of the control logic but since it only use a buffer to store incoming data and the VHDL code is a succession of *case* statements that XST optimise very well, the used resources are acceptable. The principle of the transmission is quite simple. When nothing has to be sent, the transmitter always send "1". When one flag among all the available one is set to one, all data at this time is stored in a buffer. The transmitter begins with a header, continue with three replication of the data where the format can be seen of figure 5.9. These data contain the BRAM address where the

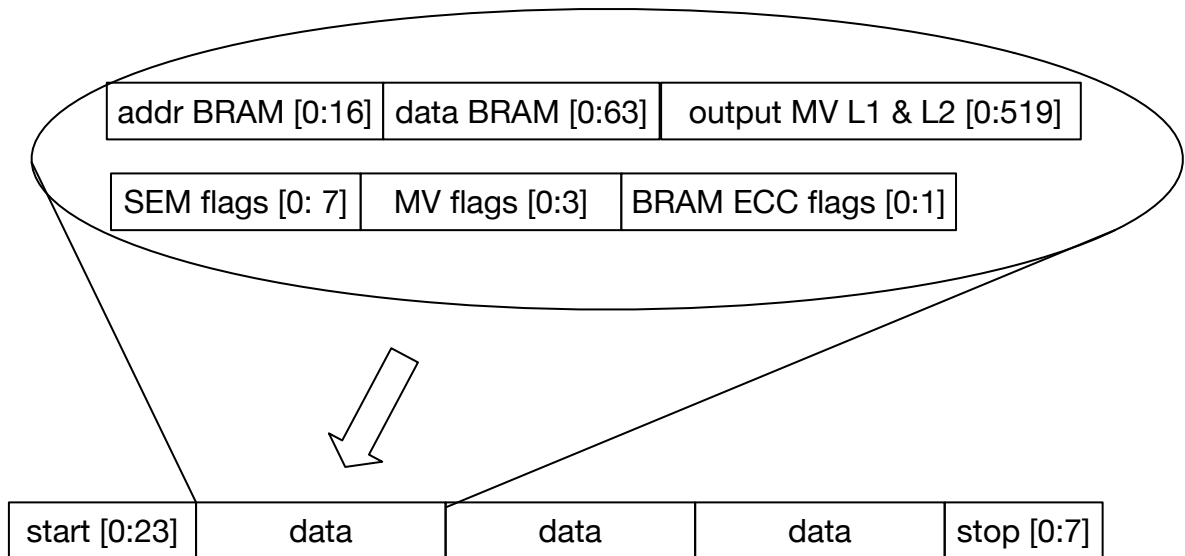


Figure 5.9: Data sent during transmission

upset is detected and the corresponding data. If the upset do not comes from BRAM, the address and data have no meaning. Next to BRAM data is the output of the three level 1 majority voters and the level 2 majority voter which are the presumed corrected data. Finally come all the different flags from SEM, majority voters and BRAM. These are the data informing an upset arose. At the receiver, a majority voter will correct an eventual error. The transmitter ends with some stop bits to inform receiver all data are sent. An important remark is the transmitter will ignore flags during the transmission. Since, all the flags are pulse signals, informations will be lost. In order to avoid this case, the beam flux has to be low enough in order to produce an upset rate low enough. Since the upset rate regarding the flux is an unknown value, it has to be tuned during the real tests at the cyclotron.

5.3.2 The Reference FPGA

The reference FPGA is the link between the computer and the irradiated FPGA. It receives raw data from the irradiated FPGA, make some basic process and send everything to the computer via Chipscope module. The adjective "reference" is not well chose. In fact, during meeting, the original idea was to run two FPGA in parallel and compare the results : one through the proton beams, the irradiated FPGA and the other one by the side. Hence the name "reference FPGA" which was kept.

The receiver This is a simple deserialiser waiting to receive start bits from the irradiated FPGA. If there is an error during start bits, a counter is incremented and if there are too much consecutive errors, the flag is set to one as the irradiated FPGA is considered down.

Upset counter It is a simple module with several counters. One for each transmitted flags. This allows us to compute the cross-section of the different ressources.

Critical error detector There are two possibility of critical error. First one is when the "uncorrectable" flags is set to one. The other is when the receiver detect too much consecutive errors. When one of these inputs is equal to one, it set the signal *reconfiguration needed* to one to inform the user to perform a manual reconfiguration via the mezzanine board.

Virtual input output (VIO) This is a Chipscope module. It allows to observe internal signals and generate inputs. The outputs can be displayed in some different ways like a LED for single bit or a number in hexadecimal, unsigned, etc. for busses. The inputs can also be customised in order to produce pulse, pulse train, switch, etc.

Integrated logic analyser (ILA) Like VIO, it is a Chipscope module but here, it displays N samples (up to 2^{16}) and have a complex triggering system in order to capture signals as we want.

Chapter 6

The test results

The test bench described in the previous chapter was tested during two hours at the UCL cyclotron. At the end, communications between the two FPGA were saved in a text file on the computer. The data analysis was performed with Matlab.

In this chapter, the performances and limitations of the firmware will be explained. After that, a discussion about the communication capture will be made and will be followed by the results that can be extracted from it. Finally, a proposition about the possibilities for this work in the future will be given.

6.1 Firmware performances

Before analysing the data saved during the testbeam, the performances and the limitations of the firmware have to be described. The firmware has been written to extract a maximum of data in order to be able to analyse what happened when an upset occurred inside the FPGA. The goal was to reach a flux as high as possible that can be handled by the communication between the two FPGA as maximise the amount of collected data. Concerning the communications, each one is 1877 bits long and takes around $93\mu s$ which is acceptable to have a large amount of data. Unfortunately, the Chipscope procedure to save the data is very slow, around 2 or 3 second to save all the signals into a text file. The consequence is an inability to save upsets arising too fast. However, the upset counters remain operational and then the possibility to count correctly. Despite this limitation, Chipscope was kept and the flux of particles during the testbeam was decreased. During the tests, a flux of $5 \times 10^5 cm^{-2}s^{-1}$ was used which lead to an upset rate around one upset every 5 seconds. Since upsets are random, some of them will still occur with an interval less than 2 or 3 seconds. Hence, a few communications will be probably missed. Such flux is around one order higher than expected flux inside CMS during phase 2. The energy of particles was 62 MeV

6.2 Data Analysis

6.2.1 The Communication Analysis

At the end of the tests, 183 communications were saved with 183 potentials upsets on the different FPGA resources. However 15 of these transmissions were corrupted. It is not possible to determine what happened with the available data. Considering the uncorrupted transmissions, no difference is observed on the three replications of transmitted data which is a good news since FPGA outputs can not be replicated. Among all the upsets detected, 59 upsets come from a single bit upset inside BRAM and 103 upsets come from the configuration. Every BRAM single bit upset were corrected. Such results were expected since Error Correction Capability (ECC) is able to correct single error [23]. Concerning the upsets in the configuration memory, no modification in the triplicated logic output was observed. Moreover, an "SEM correction" signal could leads to a "SEM uncorrectable" signal if the SEM IP was unable to correct the detected upset. No "SEM uncorrectable" signal was captured by chipscope but some were detected by the counters. This will be detailed in the next section.

6.2.2 The Upset Counters Analysis

The upset counters are adapted to the transmission rate since both are working at the same working frequency of the irradiated FPGA logic (20MHz). Thus the upsets missed by Chipscope were captured by the counters. The amount of upset for the different flags and their mean rate are listed in table 6.1. The flags sbiterr and dbiterr correspond to single and double upset respectively.

flags	#upsets	mean upset rate [s^{-1}]	cross-section [cm^2]
sbiterr	79	5.7×10^{-2}	6.64×10^{-8}
dbiterr	12	8.66×10^{-3}	1.01×10^{-8}
SEM correction	221	1.59×10^{-1}	1.86×10^{-7}
SEM uncorrectable	22	1.59×10^{-2}	1.85×10^{-8}
MVL1(0)	0	-	-
MVL1(1)	0	-	-
MVL1(2)	0	-	-
MVL2	0	-	-

Table 6.1: Amount of upsets detected by the different counters for each flag and their corresponding cross-section with a flux of protons of $5 \times 10^5 p/s.cm^2$ at 62 MeV

Every single upsets were corrected by ECC. As expected, double upsets are less frequent than simple one but it is not negligible (15%). Since double upsets can be detected but not corrected, this is a concern regarding data integrity. Concerning the memory configuration, 221 upsets were detected (SEM correction) and 22 of those can not be corrected (SEM uncorrectable) by the SEM IP which represent around 19% of the upsets in configuration memory. It is important to note that SEM have a latence of correction around several microseconds depending on the frequency working. Thus, during this time, the logic involved where the upset occurs is corrupted. This is why TMR is implemented on logic. Unfortunately, probably due to a malfunction of the firmware, no data is available for the triplicated logic. During the elaboration of the firmware, an upset simulator module was plugged in the replicated design in order to test the flags of majority

voters. At the end, this module was removed in order to not interfere with the real upsets. At this point, it was not possible to test it anymore. Furthermore, other upgrades were made in the transmitter module that collects all the flags. Accordingly, the bug is probably in the trigger of the transmitter leading to an inability to observe and count upset in triplicated logics.

6.2.3 Upset Rate in CMS

The flux of particles used during the tests was $5 \times 10^5 \text{cm}^{-2}\text{s}^{-1}$ with proton at 62MeV . In comparison, the expected flux at GE1/1 location is around $5 \times 10^4 \text{cm}^{-2}\text{s}^{-1}$ for the whole energy spectrum of neutrons. Then, a study of the number of upset induced by neutrons regarding their energy has to be performed. Therefore a second set of measurements has been performed at the UCL cyclotron during which the upset rate in the configuration memory has been recorded as a function of the proton energy, as shown in figure 6.1. It can be observed that the upset

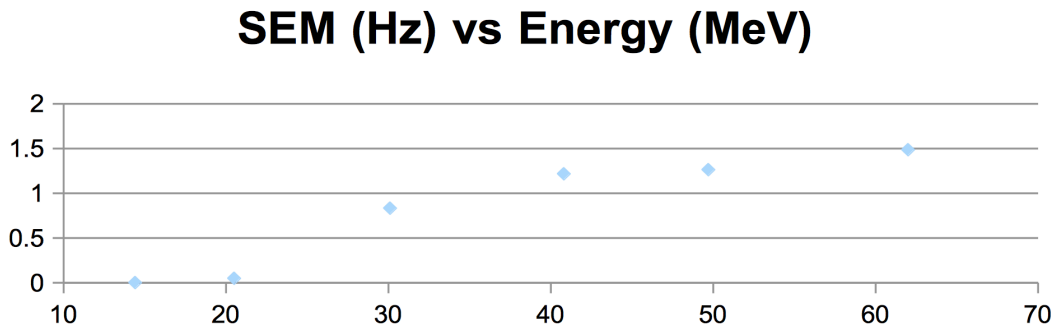


Figure 6.1: Upset rate in configuration memory in function of the protons energy.

rate decrease with the energy and is almost equal to zero at 14.4 MeV. On the other hand, the increasing seems to saturate at high energy. This is confirming the observations made in similar tests for another project of CMS sub-detectors [24] [22]

Then one can only considers neutrons with an energy higher than 14.4 MeV. The impact of the flux on the upset rate was also computed and are shown in figure 6.2.

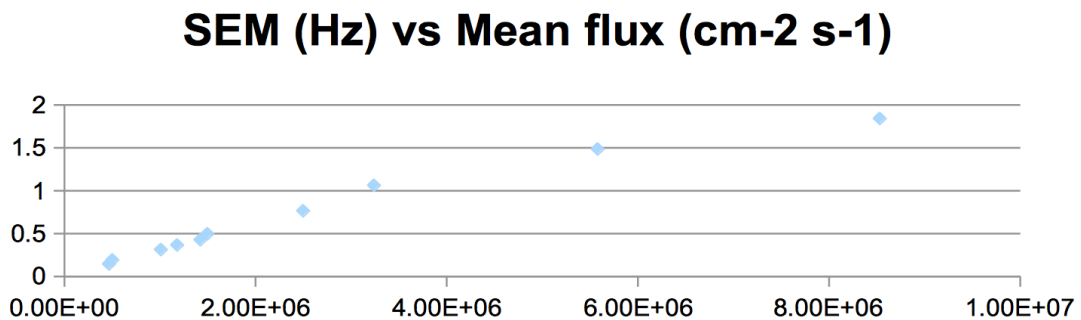


Figure 6.2: Upset rate in configuration memory in function of the protons flux.

It appears to have a linear dependency. Then a simple scale can be used. In figure 3.6 which represents the energy spectrum of neutrons, only a small fraction has an energy greater than 14.4 MeV. Considering this portion equal to 5% of the total flux the actual flux to be considered is equal to $2.5 \times 10^3 \text{cm}^{-2}\text{s}^{-1}$. Moreover, considering the upset rate is constant over the energy range, the upset rate inside CMS would be around 200 times smaller than during testbeam at UCL. Table 6.2 is an overview comparing the amount of upsets per second during the testbeam and those expected inside CMS. As a result, it can be observed that double upsets in BRAM

flags	during tests [s]	inside CMS [s]
sbiterr	17	3506
dbiterr	115	23083
SEM correction	6	1253
SEM uncorrectable	63	12590

Table 6.2: Mean time between 2 upsets for each FPGA resources during testbeam at UCL and the expected mean time inside CMS.

and uncorrectable upsets in configuration memory would be very uncommon (around every 6.5 hours and 3.5 hours respectively). Consequently, it will not be a problem. Since these upsets are uncorrectable but detected by mitigation techniques, a reconfiguration will be requested when such upset occurs. Regarding single bit upset in BRAM and correctable upset in configuration memory, mean time between 2 upsets is more frequent (around every 58 minutes and 21 minutes respectively). Correction by means of mitigation techniques is then desirable. As already seen in previous section, ECC fulfil the requirement. Moreover, concerning the TMR, even if no data is available, one can expect some convincing improvements.

6.3 The Total Ionizing Dose

As already explained, the TID predictions for the FPGA is around 1 krad [14]. On the other hand, the FPGA was exposed to a total dose around 84 krad during the testbeams. Nevertheless, the FPGA is still working correctly. As a consequence, considerations about fault repairs methods are not necessary.

6.4 For the Future

Since no data is available for the triplicated logic, more tests are necessary in the future. It is also interesting to note that SEM has an upset injection feature. This could allow to simulate the effects of the upsets on the implemented logic without the use of a cyclotron.

Conclusion

Within the framework of the GE1/1 project, an overview of the possibilities to mitigate the FPGA on the opto-hybrid board has been detailed. Consequently, a firmware to test the intrinsic FPGA behaviour and with mitigation techniques in a radioactive environment has been developed. The FPGA was also submitted to an important dose of radiations in order to determine if it will resist until the end of operations of the LHC.

Since the environment of GE1/1 project is extreme and the FPGA is not intrinsically mitigated against radiations, some tests had to be performed.

Among all the different techniques the TMR appears as the best choice. The detection is instantaneous and it is the only method that actually correct corrupted signals. TMR had to be tested since its efficiency can greatly varies from one implemented logic to another. Regarding the BRAM, the BRAM generator IP core provides mitigation by the use of ECC. Consequently, it had to be tested in order to know if BRAM is robust enough with it. Same tests had to be done for the SEM IP core which provides detection and correction of the configuration memory.

All these techniques were tested at the UCL cyclotron. It appears that ECC is able to correct every single upsets and detect double upsets which represent around 15% of the upsets in BRAM. Equivalently, SEM is able to correct single upsets and detects the uncorrectable ones which represent around 9% of upsets in configuration memory. Concerning the TMR, no data has been collected due to a malfunction of the firmware. Finally, the FPGA was exposed to a TID around 84 krad which is much more than the expected dose at the end of lifetime of the LHC and reveals to still work properly.

In conclusion, at the end of this thesis, we have a better knowledge of the FPGA behaviour and its limits.

List of Figures

1.1	CERN's accelerator complex schematic with all the different accelerators and connection between it[1].	3
1.2	LHC schematic with the experiment and beams representation[2]	4
1.3	Long term LHC calendar with the integrated and peak luminosity previsions . . .	5
1.4	CMS experiment 3D representation with its components [5]	6
1.5	Representation of the different components of one sector inside CMS and typical particle trajectories [5]	7
1.6	Schematic of the trigger system principle [7]	8
2.1	Longitudinal section of CMS with the different technologies of detector inside the muon tracker including the GE1/1 project in red [8]	9
2.2	3D representation of one of the endcaps zone with the GE1/1 detector in blue and pink [8].	11
2.3	3D representation of each component composing the detector [8].	12
2.4	Left: Zoomed in view of a GEM foil. Right: Schematic view with the electric field line in white, the electron flow in blue and the ion flow in purple through one hole in the foil [8].	12
2.5	Left: Details of a GEM foil principle. Right: Typical size and electrical parameters [8].	13
2.6	Picture of the strips of the readout board. "Hyperbolic geometric pattern is an artifact of the display on a screen" [8]. Strips are visible in the zoomed-in picture on the right.	13
2.7	Top: global view. GEB(red) and the readout board (blue). Bottom: zoom in on VFAT chips[8].	14
2.8	Picture of the actual GEB with some sector without connected VFAT that allows to see the male connectors [8].	15
2.9	Block diagram of the VFAT chip version 3 [8].	15
2.10	Picture of the GEB, the 24 VFATs and the opto-hybrid in the bottom [6].	16
3.1	Penetration path of a particle in the material [11]	20
3.2	Illustration of the funneling effect in a CMOS transistor [11]	21
3.3	SET $\{ 0 \rightarrow 1 \rightarrow 0 \}$ shape and characteristic values [11].	23
3.4	Bit-flip in a SRAM cell [11].	24
3.5	Example of a cross-section curve in function of LET [12].	25

3.6	Energy spectrum of the incident particles at the GE1/1 location. Results obtained by means of FLUKA simulations [8].	26
3.7	Flux estimations of neutrons, electrons/positrons and photons at the location of GE1/1 project. Results obtained by means of FLUKA simulations [14].	27
3.8	Absorbed dose estimation for an integrated luminosity of $3000fb^{-1}$ at the location of GE1/1 project. Results obtained by means of FLUKA simulations [14].	27
4.1	Overview of the different detection methods [15].	30
4.2	Schematic of non intrusive CED [16]	30
4.3	Schematic of triplicated throughput logic between two state machines in a logic module [18]	31
4.4	Schematic of an unmitigated bit counter with the feedback visible [18].	32
4.5	Schematic of a triplicated bit counter with a unique majority voter [18].	32
4.6	Schematic of a triple voted Triplicated counter [18].	33
4.7	Schematic of a triplicated output with minority voters [18].	33
4.8	Overview of the different repair methods [15].	35
5.1	picture of the device under test inside the irradiation chamber	38
5.2	Another viewpoint of the device under test with the shield more visible	38
5.3	Focus on the source of the beam, the degraders and the ionization chamber	39
5.4	Test bench with the power supplies, the reference FPGAs and the computer	39
5.5	Testbench schematic	40
5.6	Simple dual port BRAM schematic [23]	41
5.7	Protocol for the BRAM use during test beam	42
5.8	Design implemented on the FPGA for the test beam	43
5.9	Data sent during transmission	44
6.1	Upset rate in configuration memory in function of the protons energy.	49
6.2	Upset rate in configuration memory in function of the protons flux.	49

List of Tables

6.1	Amount of upsets detected by the different counters for each flag and their corresponding cross-section with a flux of protons of $5 \times 10^5 p/s.cm^2$ at 62 MeV	48
6.2	Mean time between 2 upsets for each FPGA resources during testbeam at UCL and the expected mean time inside CMS.	50

Bibliography

- [1] *CERN document server CERN's accelerator complex pictures*. <http://cds.cern.ch/record/1621583/files/CERN%20s-accelerator-complex2013.jpg?version=1>. (Visited on 07/25/2016).
- [2] *Royal society website The Large Hadron Collider article*. <http://rsta.royalsocietypublishing.org/content/370/1961/831>. (Visited on 07/25/2016).
- [3] Lyndon Evans and Philip Bryan. "LHC machine". In: *Journal of Instrumentation* (2008).
- [4] Baptiste Herregods. "Contribution à l'étude de détecteurs équipés de GEM pour les "upgrades" du spectromètre à muons de CMS". MA thesis. Université Libre de Bruxelles, 2015.
- [5] URL: <http://cms.web.cern.ch/news/what-cms> (visited on 07/27/2016).
- [6] Hugo Dewitte. "Implementation of an FPGA-based calibration procedure at the detector level for the future high-luminosity phase of the CMS". MA thesis. Université Libre de Bruxelles, 2016.
- [7] Erik Verhagen. "Development of the new trigger and data acquisition system for the CMS forward-muon spectrometer upgrade". PhD thesis. Université Libre de Bruxelles, 2015.
- [8] *CMS Technical design report for the muon endcap GEM upgrade*. 2014.
- [9] VadaTech. "A Brief Introduction to Micro Telecommunications Computing Architecture Concepts". In: *MicroTCA overview* (2014).
- [10] Florian Zenoni. "Study of Triple-GEM detectors for the CMS muon spectrometer upgrade at LHC and study of the forward-backward charge asymmetry for the search of extra neutral gauge bosons". PhD thesis. Université Libre de Bruxelles, 2016.
- [11] Niccolo Battezzati, Luca Sterpone, and Massimo Violante. *Reconfigurable Field Programmable Gate Arrays for Mission-Critical Applications*. Springer, 2011.
- [12] Fernanda Lima Kastensmidt, Luigi Carro, and Ricardo Reis. *Fault-Tolerance Techniques for SRAM-based FPGAs*. 2006.

- [13] M. Huhtinen and F. Faccio. “Computational method to estimate Single Event Upset rates in an accelerator environment”. In: *Nuclear instruments and methods in physics research* (1999).
- [14] Alfredo Castaneda. *Muon Phase-2 background simulation*. 2016.
- [15] Edward Stott, Pete Sedcole, and Peter Y. K; Cheung. “FAULT TOLERANT METHODS FOR RELIABILITY IN FPGAs”. In: *IEEE Xplore* (2008).
- [16] Sobeeh Almkhaizim, Petros Drineas, and Yiorgos Makris. “Concurrent Error Detection for Combinational and Sequential Logic via Output Compaction”. In: *IEEE* (2004).
- [17] Charles Stroud et al. “Built-In Self-Test of Logic Blocks in FPGAs”. In: *IEEE* (1996).
- [18] Carl Carmichael. “XAPP197 : Triple Module Redundancy Design Techniques for Virtex FPGAs”. In: *Xilinx application note* (2006).
- [19] “XAPP138 Virtex FPGA Series Configuration and Readback”. In: *Xilinx application note* (2005).
- [20] Carl Carmichael. “Correcting Single-Event Upsets Through Virtex Partial Configuration”. In: *Xilinx application note* (2000).
- [21] Centre de ressources du cyclotron. *Parameters of the proton beam facility*. <http://www.cyc.ucl.ac.be/LIF/LIF.php>. (Visited on 07/20/2016).
- [22] B. Bylsma et al. “Radiation testing of electronics for the CMS endcap muon system”. In: *Elsevier Nuclear Instruments and Methods in Physics Research A* 698 (2013), pp. 242–248. DOI: [url{http://dx.doi.org/10.1016/j.nima.2012.09.017}](http://dx.doi.org/10.1016/j.nima.2012.09.017).
- [23] Xilinx. “LogiCORE IP Block Memory Generator v7.3 Product Guide”. In: (2012).
- [24] J. Gilmore et al. “Very forward muon trigger and data acquisition electronics for CMS: design and radiation testing”. In: *IOP PUBLISHING FOR SISSA MEDICAL AB* (2013).