ULB UNIVERSITÉ LIBRE DE BRUXELLES

PhD Thesis

Development of the DAQ System of Triple-GEM Detectors for the CMS Muon Spectrometer Upgrade at LHC

Thomas Lenzi

CERN-THESIS-2016-215
19/12/2016

Supervisor Gilles De Lentdecker

December 2016

Université Libre de Bruxelles

Département de Physique
IIHE - Interuniversity Institute for High Energies

PhD Thesis

Development of the DAQ System of Triple-GEM Detectors for the CMS Muon Spectrometer Upgrade at LHC

Thomas Lenzi

Jury members

Supervisor Gilles De Lentdecker
Université Libre de Bruxelles

President Laurent Favart
Université Libre de Bruxelles

Secretary Pascal Vanlaer
Université Libre de Bruxelles

Reviewers Paul Aspell
European Organization for Nuclear Research

Kael Hanson
University of Wisconsin–Madison

Nick Van Remortel
Universiteit Antwerpen

Frédéric Robert
Université Libre de Bruxelles

Yifan Yang
Université Libre de Bruxelles

December 2016

Thomas Lenzi

Development of the DAQ System of Triple-GEM Detectors for the CMS Muon Spectrometer Upgrade at LHC

PhD Thesis, December 2016
Reviewers: Laurent Favart, Pascal Vanlaer, Paul Aspell, Kael Hanson, Nick Van Remortel, Frédéric Robert, and Yifan Yang
Supervisor: Gilles De Lentdecker

Université Libre de Bruxelles
IIHE - Interuniversity Institute for High Energies
Département de Physique
Boulevard du Triomphe, 2
1050 - Brussels

À vous.

Abstract

The Gas Electron Multiplier (GEM) upgrade project aims at improving the performance of the muon spectrometer of the Compact Muon Solenoid (CMS) experiment which will suffer from the increase in luminosity of the Large Hadron Collider (LHC). After a long technical stop in 2019-2020, the LHC will restart and run at a luminosity of 2×10^{34} cm^{-2} s^{-1}, twice its nominal value. This will in turn increase the rate of particles to which detectors in CMS will be exposed and affect their performance. The muon spectrometer in particular will suffer from a degraded detection efficiency due to the lack of redundancy in its most forward region. To solve this issue, the GEM collaboration proposes to instrument the first muon station with Triple-GEM detectors, a technology which has proven to be resistant to high fluxes of particles. Within the GEM collaboration, the Data Acquisition (DAQ) subgroup is in charge of the development of the electronics and software of the DAQ system of the detectors.

This thesis presents the work of the author as lead developer of the firmware for the front-end and back-end electronics. These developments have been performed from the ground up and designed to transfer data from the analog front-end to the off-detector electronics while offering extensive control and monitoring capabilities. The developed DAQ chain has been tested extensively during two test beam campaigns which provided results on both the stability of the system and the efficiency of the detectors. The results of these campaigns are described in the present thesis. Further characterization of the electronics is also described along with the procedures and tools built to qualify the components for their installation in CMS. Finally, the results of the irradiation tests performed on the on-detector electronics are presented. These allowed us to understand the effects of radiation on the electronic boards used for the GEM upgrade project and the impact that it has on the design of the firmware for CMS.

Additionally, the work of the author on a new architecture for DAQ systems is described. The latter uses modified network topologies and novel web technologies to increase the available bandwidth on the network and yield an event-driven system.

Résumé

Cette thèse de doctorat s'inscrit dans le projet de mise à niveau du spectromètre à muons du Compact Muon Solénoid (CMS) auprès du Large Hadron Collider (LHC). Après un arrêt pour maintenance prévu en 2019-2020, le LHC reprendra son programme de recherche à une luminosité de 2×10^{34} cm^{-2} s^{-1}, soit deux fois sa valeur nominale. Ceci aura pour conséquence d'accroître le taux de particules auquel seront soumis les détecteurs de CMS et d'entraver l'efficacité de détection de ces derniers. Le spectromètre à muons de CMS sera tout particulièrement touché à cause du manque de redondance dans sa partie avant. Afin de palier à ce problème, il est proposé d'installer des détecteurs Gas Electron Multipliers (GEM) dans la première station à muons. La technologie GEM répond aux besoins de CMS en offrant une excellente efficacité de détection à de hauts flux de particules. Au sein de la collaboration GEM, le sous-groupe en charge du système d'acquisition de données (DAQ) doit développer l'électronique et les logiciels de gestion de la chaîne de lecture des détecteurs.

Nous présentons ici le travail de l'auteur réalisé en tant que principal développeur du firmware pour l'électronique du système DAQ. Ces développements visent à créer une architecture qui achemine les données depuis l'électronique de lecture du détecteur jusqu'aux systèmes situés dans la zone de comptage, tout en offrant la possibilité de contrôler et surveiller l'ensemble des composants du DAQ. Le système mis en place a été largement testé durant deux campagnes de tests en faisceaux qui ont fourni des informations concernant la stabilité du DAQ ainsi que des mesures de l'efficacité des détecteurs. Nous décrivons également l'ensemble des travaux réalisés afin de caractériser les composants électroniques avant leur installation dans CMS ainsi que les résultats des tests d'irradiation effectués sur l'électronique du détecteur. Ces derniers permettent de mieux comprendre les conséquences des radiations sur le système et de développer des méthodes de mitigation.

De plus, le travail de l'auteur sur la création d'une nouvelle architecture de système DAQ est décrit. Cette dernière combine les avancées récentes en terme de technologies web à une topologie de réseaux non-classique afin d'accroître la bande passante disponible et de créer un système réactif.

Acknowledgments

I would like to start by thanking my supervisor Dr. Gilles De Lentdecker for his guidance throughout my PhD and the valuable advice he has given me during these years. I greatly appreciate the freedom he gave me to investigate topics outside of my research scope. I am also thankfull to the members of my jury - Dr. Laurent Favart, Dr. Pascal Vanlaer, Dr. Paul Aspell, Dr. Kael Hanson, Dr. Nick Van Remortel, Dr. Frédéric Robert, and Dr. Yifan Yand - whose insightfull comments have helped to improve this dissertation.

My time at the IIHE would not have been the same without my colleagues and friends. A special thank you to Alexandre, Audrey, Erik, Florian, Nadir, Thierry, and many others for their friendship.

Finally, I would like to thank my friends and family for it is the moments we share that are most precious to me.

Contents

Abstract vii

Résumé ix

Acknowledgments xi

Table of Contents xiii

Introduction 1

I An Introduction to Particle Physics 5

1 The Standard Model 7
- 1.1 Gauge Invariance and Gauge Bosons . . . 8
- 1.2 Electroweak Theory . . . 9
- 1.3 Electroweak Symmetry Breaking . . . 10
- 1.4 Quantum Chromodynamics . . . 12
- 1.5 Beyond the Standard Model . . . 13

2 The Large Hadron Collider 15
- 2.1 The Injection Complex . . . 16
- 2.2 LHC Technical Description . . . 17
- 2.3 Performance Goals . . . 18
- 2.4 Operations and Schedule . . . 19
- 2.5 Scientific Motivations for the LHC Upgrade . . . 20

3 The Compact Muon Solenoid Experiment 23
- 3.1 Overview . . . 23
- 3.2 The Inner Tracking System . . . 25
- 3.3 The Electromagnetic Calorimeter . . . 27
- 3.4 The Hadron Calorimeter . . . 28
- 3.5 The Superconducting Magnet . . . 30
- 3.6 The Muon System . . . 31
  - 3.6.1 Drift Tubes . . . 32
  - 3.6.2 Cathode Strip Chambers . . . 33

3.6.3 Resistive Plate Chambers . . . 33
3.7 The Trigger System . . . 34
3.8 The Data Acquisition System . . . 36
3.9 The Upgrades of CMS . . . 37

II The CMS GEM Upgrade Project 39

4 Gas Electron Multiplier Detectors 41
4.1 Motivations for the GE1/1 Upgrade . . . 41
4.2 Technology Overview of Triple-GEM detectors . . . 44
4.3 Technical Design of GE1/1 chambers for CMS . . . 46
4.4 Detector Design Evolution . . . 48
4.5 GE1/1 Prototyping Results . . . 50
4.5.1 Rate Capability . . . 50
4.5.2 Detection Efficiency . . . 51
4.5.3 Time Resolution . . . 51
4.5.4 Spatial Resolution . . . 52
4.5.5 Response Uniformity . . . 52
4.6 GEM Upgrade Schedule . . . 52
4.7 Developments Beyond GE1/1 . . . 53

5 The GEM Data Acquisition System 55
5.1 The GE1/1 Data Acquisition System . . . 55
5.1.1 The VFAT3 ASIC . . . 56
5.1.2 The GEM Electronic Board (GEB) . . . 58
5.1.3 The OptoHybrid . . . 59
5.1.4 The GBT and Versatile Link . . . 59
5.1.5 The microTCA Standard . . . 61
5.1.6 The CTP7 Advanced Mezzanine Card . . . 63
5.1.7 The AMC13 Advanced Mezzanine Card . . . 64
5.1.8 The IPBus Protocol . . . 65
5.2 The First Prototype of the Data Acquisition System . . . 66
5.2.1 The VFAT2 ASIC . . . 67
5.2.2 The GEM Electronic Board v1 . . . 68
5.2.3 The OptoHybrid v1 . . . 69
5.2.4 The GLIB Advanced Mezzanine Card . . . 69
5.3 The GE1/1 Data Acquisition System for the Test Beam . . . 70
5.3.1 The GEM Electronic Board v2 . . . 70
5.3.2 The OptoHybrid v2a . . . 71
5.4 The GE1/1 Data Acquisition System for the CMS Slice Test . . . 72
5.4.1 The OptoHybrid v2b . . . 72
5.5 Evolution of the DAQ System . . . 72

6 A Data Acquisition System for the Test Beam **75**
6.1 Architecture of the OptoHybrid Firmware . . . 75
6.1.1 Internal Communication Through a Wishbone-Like Bus . . . 76
6.1.2 Encoding Fast Commands . . . 78
6.1.3 Formatting Trigger Data . . . 80
6.1.4 Acquiring Tracking Data . . . 80
6.1.5 Control and Monitoring . . . 82
6.1.6 Calibrating the Systems . . . 83
6.1.7 Optical Communication with the Off-Detector Electronics . . . 85
6.1.8 System Summary . . . 87
6.2 Architecture of the GLIB Firmware . . . 88
6.2.1 The Tracking Data Readout . . . 88
6.2.2 The IPBus Controller and Slaves . . . 88
6.3 The Control and Monitoring Application . . . 89
6.3.1 Architecture of the Application . . . 89
6.3.2 Controlling the Systems . . . 90
6.3.3 Calibrating the VFAT2s . . . 94
6.3.4 Fast and Slow Control of the VFAT2s . . . 94
6.3.5 Data Quality Monitoring . . . 94
6.3.6 Hit Rate Monitoring . . . 97
6.4 The Test Beam Setup . . . 97
6.4.1 The GEM Detectors Setup . . . 98
6.4.2 The Beam Area Setup . . . 100
6.5 Data Analysis . . . 102
6.5.1 VFAT2 Threshold Scans . . . 102
6.5.2 VFAT2 Latency Scans . . . 104
6.5.3 Evolution of the Efficiency with the High-Voltage . . . 105
6.5.4 Evolution of the Efficiency with the VFAT2 Threshold . . . 106
6.5.5 Evolution of the Efficiency with the Trigger Rate . . . 108
6.5.6 Cluster Multiplicity and Size . . . 109
6.5.7 Beam Profile . . . 111
6.6 Towards the Slice Test . . . 112
6.7 Towards the Final DAQ System . . . 113
6.8 Conclusion . . . 114

7 Qualification of the Electronics **117**
7.1 Qualification of the VFAT2s . . . 117
7.1.1 Identification of Faulty Channels . . . 117
7.1.2 Analog Front-End Currents and Voltages . . . 118
7.1.3 Analog Front-End Calibration . . . 119
7.1.4 Analog Front-End Equalization . . . 121
7.1.5 Noise Measurement . . . 122

7.2 Qualification of the GEB . . . 123
7.2.1 The GEB Testing Board . . . 123
7.2.2 Results of the Qualification of the GEB . . . 124
7.3 Qualification of the DAQ System . . . 125
7.4 Conclusion . . . 127

8 Irradiation Tests **129**
8.1 Architecture of an FPGA . . . 129
8.1.1 Configurable Logic Blocks (CLB) . . . 129
8.1.2 Digital Signal Processing (DSP) . . . 131
8.1.3 The Switching Matrix . . . 132
8.1.4 Block RAM (BRAM) . . . 132
8.1.5 The Configuration Memory . . . 133
8.2 Effects of Radiation on FPGAs . . . 133
8.2.1 Energy Losses in Matter . . . 133
8.2.2 Effects of Radiation on Transistors . . . 134
8.2.3 Single Event Transients (SET) . . . 134
8.2.4 Single Event Upsets (SEU) . . . 135
8.2.5 Total Ionizing Dose (TID) . . . 135
8.3 SEU Mitigation Techniques . . . 135
8.3.1 Configuration Memory Scrubbing . . . 135
8.3.2 Logic Triplication . . . 136
8.3.3 Forward Error Correction . . . 136
8.3.4 One-Hot Finite State Machines . . . 136
8.4 Firmware Design for the Irradiated FPGA . . . 137
8.4.1 Configurable Logic Block . . . 137
8.4.2 Block RAM . . . 139
8.4.3 Digital Signal Processing . . . 139
8.4.4 Soft Error Mitigation . . . 139
8.5 Firmware Design for the Control FPGA . . . 140
8.5.1 Communication Protocol . . . 140
8.5.2 ChipScope Core . . . 140
8.6 Irradiation Setup . . . 142
8.7 Data Analysis . . . 142
8.7.1 Evolution of the SEU Cross Section with the Energy . . . 144
8.7.2 Evolution of the SEU Cross Section with the TID . . . 145
8.7.3 Evolution of the SEU Cross Section with the In-Beam Position 146
8.7.4 Interaction with the CLBs and DSPs . . . 147
8.7.5 Effectiveness of Triplication as Error Mitigation Technique . . 148
8.8 Consequences for the GEM Upgrade Project . . . 148
8.8.1 Impact on the GE1/1 Upgrade Project . . . 149
8.8.2 Impact on the Firmware of the OptoHybrid . . . 150

8.8.3 Impact on the GE2/1 and ME0 Upgrade Projects 151
8.9 Conclusion 152

9 Summary **153**

III Data Acquisition System Design **155**

10 A Classical Approach to Data Acquisition System Design **157**
10.1 The Experimental Setup 157
10.2 The Infrastructure of the System 158
10.2.1 The Web Interface 159
10.2.2 The Central Computing Node 162
10.2.3 The High Voltage and Gas Controllers 162
10.3 A First Prototype using the Xilinx SP601 Development Board 162
10.3.1 The VFAT2 FPGA Mezzanine Card 162
10.3.2 The Trigger and Tracking Paths 163
10.3.3 The Microblaze Soft Core Processor 164
10.3.4 IPBus over the Universal Asynchronous Receiver/Transmitter Protocol 165
10.4 Upgrade of the System using the GLIB 166
10.5 Final System using the First Prototype of the OptoHybrid 167
10.6 Conclusion 168

11 Developing Data Acquisition Systems using new WEB Technologies **171**
11.1 The Experimental Setup and System Architecture 171
11.2 Web Server, WebSockets, and TCPSockets 171
11.2.1 The HTTP Standard 172
11.2.2 The WebSockets Protocol 174
11.2.3 Handling Client Requests 174
11.2.4 The TCPSockets Implementation 175
11.3 The Microblaze Processor 175
11.3.1 Lightweight TCP/IP stack 176
11.3.2 Memory File System 178
11.3.3 Mailboxes for Dual Processor Architecture 178
11.4 Innovation of the Architecture 178
11.5 Response Time of Three DAQ Systems 180
11.5.1 Slow Control Request 180
11.5.2 Data Readout 182
11.6 Potential Application for the GE1/1 Upgrade Project 183
11.7 Conclusion 185

12 Summary **187**

Conclusion **189**

List of Figures **191**

List of Tables **201**

List of Acronyms **203**

Bibliography **209**

Introduction

The Large Hadron Collider (LHC), operated since 2008 by the European Organization for Nuclear Research (CERN), is state-of-the-art in the field of particle accelerators and colliders. Designed to provide proton-proton collisions at an energy of 14 TeV in the center of mass reference frame and a luminosity of 10^{34} cm^{-2} s^{-1}, it enables scientists to study particle physics at scales never reached before. The data collected by the experiments recording the collisions of the LHC, among which the Compact Muon Solenoid (CMS), is used to test and refine the current models used to describe the fundamental particles and their interactions. To improve the performance of the LHC and increase the recorded event sample, an upgrade of the machine is foreseen for 2019-2020, during the so-called Long Shutdown 2 (LS2). After this upgrade, the LHC will restart with a luminosity of 2×10^{34} cm^{-2} s^{-1}, twice its nominal value. This will in turn impact the detectors as the rate of proton-proton interactions occurring during each bunch crossing will increase.

The increase in luminosity of the LHC will greatly affect the forward region of the muon spectrometer of CMS where particle rates can reach several kHz cm^{-2}. While most of the muon spectrometer is equipped with two technologies of detectors, either Drift Tubes (DTs) or Cathode Strip Chambers (CSCs) combined with Resistive Plate Chambers (RPCs), the forward region only instruments CSCs. Due to concerns regarding the rate capability of RPCs, the space foreseen to equip these detectors was left vacant in the most forward region, closest to the beam, thus relying on CSCs to perform measurements. If left as is, CMS will experience a loss of efficiency in the triggering and reconstruction of muons after LS2. In this scenario, the level-1 trigger threshold on the transverse momentum of single muons would have to be raised to $p_T \approx 30$ GeV c^{-1} instead of the current 14 GeV c^{-1} which would be a considerable drawback in terms of physics analysis, decreasing the kinematic acceptance by more than 35% for channels such as $H \to \tau_h \tau_\mu$.

To tackle this issue, new detectors relying on the Gas Electron Multiplier (GEM) technology are under development. Studies led by the CMS GEM collaboration

have proven that GEM detectors maintain an efficiency above 98% at particle fluxes as high as 100 MHz cm^{-2}, well above the expected hit rate of 5 kHz cm^{-2} within CMS. Additionally, they provide a spatial resolution of the order of 150 μm and a time resolution better than 5 ns with a gas mixture of $Ar/CO_2/CF_4$. Through simulations, it was shown that the instrumentation of a layer of GEM detectors in the forward region of the muon spectrometer, coupled with the CSCs, would improve the triggering and reconstruction efficiency of CMS. These results led to the approval of the installation of a full ring of detectors in both endcaps of CMS during LS2. In preparation of the LS2 installation, a small scale test with the near final electronics, called the slice test, will take place starting in 2017.

Within the GEM collaboration, the Data Acquisition (DAQ) subgroup is in charge of the development of the electronics and software of the DAQ system. The readout and digitization of the GEM detectors is performed by the VFAT3 analog front-end chip coupled to the OptoHybrid board. The latter concentrates the data from 24 VFAT3s placed on the detector and forwards it, over optical links, to the off-detector electronics composed of the CTP7 board. The latter runs within a Micro Telecommunications Computing Architecture (microTCA) crate, standard adopted throughout CMS, and is the gateway to the central DAQ system of CMS.

The DAQ system of the GEM detectors has been tested several times during test beam campaigns organized at CERN. Using near final electronic components, the system was used to perform measurements on the detectors and test the reliability of the readout chain. Additionally, characterization and commissioning procedures have been developed to qualify the components for their installation in CMS. These procedures are of importance to understand the response of the analog front-end to the passage of particles in the detector. In parallel, irradiation tests have been performed on the OptoHybrid in order to quantify its sensitivity to upsets created by the high background fluxes to which it will be exposed in CMS. These various tests have helped to build a robust DAQ system that will be used during the slice test and in the system installed during LS2.

The first part of this PhD thesis provides the reader with a background in theoretical and experimental particle physics. Chapter 1 describes the Standard Model used to describe the fundamental particles and their interactions. Focus is then given, in Chapter 2, to the technical description and scientific motivations behind the LHC. Chapter 3 finalizes the introduction by describing CMS and its subdetectors.

The second part is dedicated to the GEM upgrade project and the thorough description of its DAQ system and the various tests performed. Chapter 4 introduces the reader to the GEM technology and the performance of the detectors. The DAQ system and its evolution are reviewed in Chapter 5 in which the electronic compo-

nents and setups developed over time are detailed. Chapter 6 presents the firmware and software developments performed by the author for the test beam campaigns organized at CERN in November 2014 and 2015. Additionally, the data recorded during these tests is analyzed and presented, yielding various results on the detection efficiency of the detectors. Finally, Chapters 7 and 8 describe the work performed on the qualification and calibration, and the irradiation tests of the on-detector electronics used in the GEM project.

The last part of this work is dedicated to the design of two different architectures of DAQ systems. Chapter 10 describes the DAQ system developed for a small prototype of GEM detectors. This system relies on well-known methods used to build DAQ architectures. These are challenged by a novel topology, developed by the author and reviewed in Chapter 11, which changes the interactions within a DAQ system.

Part I

An Introduction to Particle Physics

Over the last century, with the advent of quantum physics and special relativity, the field of particle physics has made tremendous progress in the understanding of the elements and the forces that compose our universe. While many questions remain open for future generations to answer, scientists have continuously been improving the models developed to describe the building blocks of nature and their interactions. Benefiting from the technological advancements led by the advent of the transistor and modern computing, physicists have been given the tools required to create more powerful and complex experiments. These advancements allow scientists to confront theoretical predictions and experimental observations while probing deeper the secrets of the universe.

In 2013, the Nobel Prize in Physics was awarded to François Englert and Peter W. Higgs "*for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider*". For the thousands of physicists who helped make this discovery possible, the news was received as the recognition of a search spanning over 40 years. It started in 1964 when François Englert together with Robert Brout [1] and, independently, Peter W. Higgs [2] proposed a mechanism to explain possible mass differences among the gauge bosons of the Standard Model. As a consequence they postulated the existence of a new elementary particle. They thereby solved the ever present question of the origin of mass in the theory as it was formulated, allowing the field of particle physics to leap forward. However, their hypothesis remained unproven until July 4th 2012, day on which two international collaborations of more than 3000 scientists announced that, after two years of experimentation and data analysis, they had discovered the so called Higgs boson [3], the missing particle which would validate the famed theory.

The discovery of the Higgs boson not only solved the problem of the origin of mass of elementary particles but also strengthened the foundations of the Standard Model, the prominent theory in particle physics also considered to be one of the greatest

achievements in the field. The Standard Model is a mathematical construction that classifies the elementary particles and the interactions they undergo. It was developed in the 1970s driven by the idea of grand unification: the description of all forces of nature by a single theory. It combined the electroweak and quantum chromodynamics theories, respectively describing the electromagnetic and weak, and strong interactions. The Standard Model quickly gained credibility with the discoveries of neutral weak currents by the Gargamelle experiment and the charm quark by the Stanford Linear Accelerator Center and the Brookhaven National Laboratory in 1974. Over the years, many other theoretical predictions have proven to be correct through experimental results and discoveries. These successes were made possible thanks to the construction of new experiments that allow physicists to reach new frontiers in the search for answers.

This introduction to particle physics aims to provide the reader with a background on both the theoretical and experimental aspects of the field. Chapter 1 covers the Standard Model and the mathematical tools used to describe the physics supporting it. After a review of the current state of the theory and the description of nature it yields, the limitations of the model are addressed along with the solutions that are currently being investigated. Focus is then given to the experimental side of particle physics in Chapters 2 and 3 which respectively describe the Large Hadron Collider at CERN, the European Organization for Nuclear Research, and the Compact Muon Solenoid experiment. The Large Hadron Collider is state-of-the-art in particle collider engineering which, through the collisions it provides, enables the Compact Muon Solenoid detector to probe the mysteries of nature.

The Standard Model

1

The Standard Model provides a classification and description of the subatomic particles that compose our universe and their interactions. It describes both the particles of matter called fermions, which are subdivided among quarks and leptons, and the force carriers which are bosons. Figure 1.1 illustrates the particle content of the Standard Model and the ways in which they are classified among types and generations. Quarks and leptons, and their antiparticles (antiquarks and antileptons), are further divided into three generations with identical properties but increasing mass, from which the first constitutes the everyday matter, whereas the two others are unstable and quickly decay towards lighter generations. The interactions between particles are carried out through the exchange of bosons. Gluons are the carriers of the strong force and interact solely with themselves and quarks as described by quantum chromodynamics (QCD). The photon, responsible for electromagnetism, couples to all charged particles while the $W^{\pm}$ and Z bosons couple to all particles except photons and gluons, as described by the electroweak theory (EWK). Finally, the Higgs boson as previously stated is the manifestation of the Higgs mechanism at the origin of mass.

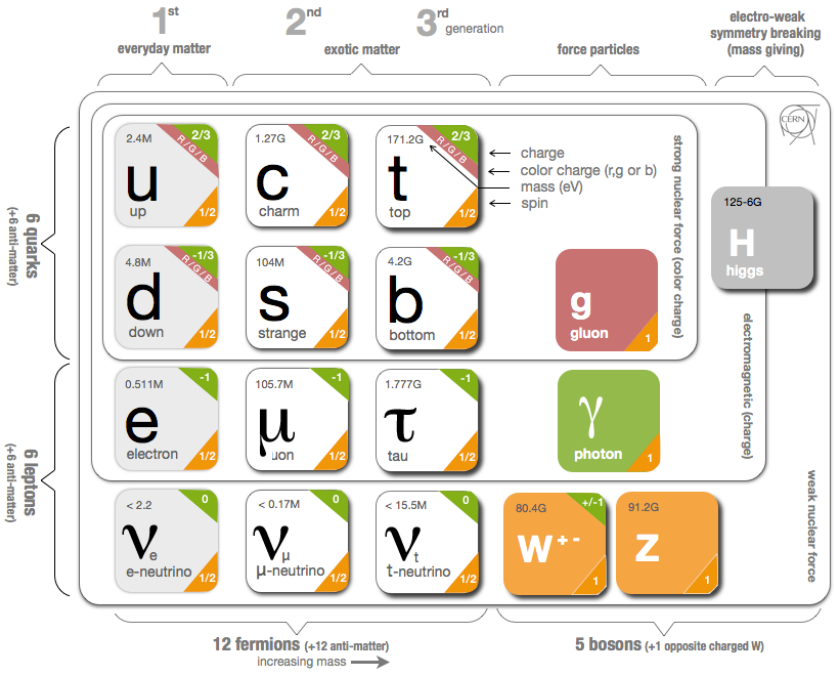

Figure 1.1 – Overview of the elementary particles as described by the Standard Model. Everyday matter is composed of the first generation of quarks and leptons to which two heavier generations are added. Gauge bosons, or force particles are the carriers of the interactions between other particles. The Higgs boson is the manifestation of the mechanism that gives mass to particles [4].

In addition to classifying particles, the Standard Model also provides a strong mathematical model built on top of quantum field theory and gauge invariance that yields an accurate representation of the fundamental processes of the universe. The gauge group of the theory is $SU(3)_C \otimes SU(2)_L \otimes U(1)_Y$, where $SU(3)_C$ and $SU(2)_L \otimes U(1)_Y$ are respectively the symmetry groups of the QCD and EWK sectors. In this paradigm, quarks and leptons are represented as fermionic fields within the Lagrangian of the model while bosons arise from the invariance of the Lagrangian under local gauge transformations.

1.1 Gauge Invariance and Gauge Bosons

In the Standard Model, bosons are mathematical and physical objects that arise from the requirement of gauge invariance of the Lagrangian. This constraint translates the fact that the physics of the system should not change under an arbitrary phase change of the coordinate system. A local transformation of a fermionic field ψ under a gauge group G can be written as

$$\psi \to \psi' = \exp\left(-\frac{i}{2} g \alpha^k(x) T^k\right) \psi, \tag{1.1}$$

where g is the gauge coupling constant, $\alpha^k(x)$ are the local rotation parameters, and T^k are the generators of the representation of G. To preserve the invariance of the Lagrangian of a free massless field ψ,

$$\mathcal{L} = i\bar{\psi}\gamma^\mu \partial_\mu \psi, \tag{1.2}$$

the partial derivate ∂_μ must be rewritten as a covariant derivate to absorb the change in coordinates

$$D_\mu = \partial_\mu - \frac{ig}{2} W_\mu^k T^k, \tag{1.3}$$

where W_μ^k are the emerging gauge fields associated to G. The gauge invariant Lagrangian thus becomes

$$\mathcal{L} = i\bar{\psi}\gamma^\mu D_\mu \psi - \frac{1}{4} W_{\mu\nu}^i W^{i\mu\nu}, \tag{1.4}$$

where $W_{\mu\nu}^i$ are the field strength tensors of the W_μ^i gauge fields given by

$$W_{\mu\nu}^i = \partial_\mu W_\nu^i - \partial_\nu W_\mu^i - g\epsilon^{ijk} W_\nu^j W_\mu^k, \tag{1.5}$$

with ϵ^{ijk} the totally antisymmetric tensor. From the necessity of the invariance of the theory under local transformations, a new massless gauge boson is associated to each generator of the adjoint representation of the symmetry group G. Moreover, interaction terms between the fermionic field ψ and the gauge fields W_μ^i as well as

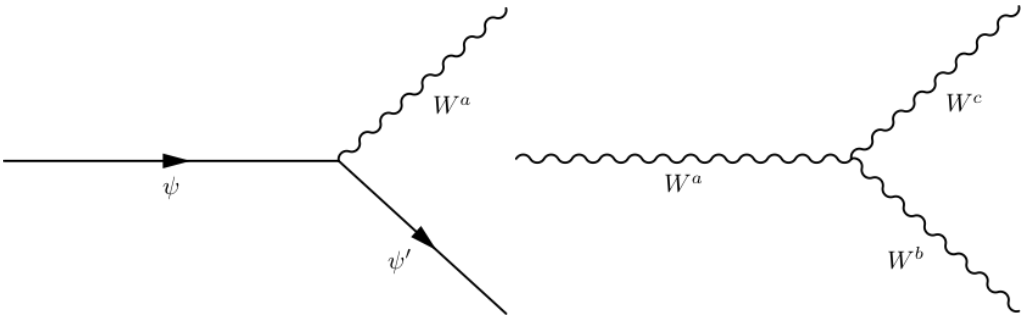

Figure 1.2 – Feynman diagrams representing the interaction terms of the Lagrangian. Left: interaction between a fermion and a boson. Right: self-interaction of gauge bosons.

self-interacting gauge field terms have emerged. These are represented in Figure 1.2 which provides a schematic view of the processes at play using Feynman diagrams.

1.2 Electroweak Theory

The gauge group of the EWK sector of the Standard Model is $SU(2)_L \otimes U(1)_Y$ for which a transformation of a fermionic field ψ can be written as

$$\psi \to \psi' = \exp\left(-\frac{i}{2} g_W \Lambda^k(x) T^k\right) \exp\left(-\frac{i}{2} g'_W \alpha(x) Y\right) \psi, \tag{1.6}$$

where g_W, $\Lambda^k(x)$, and T^k are respectively the gauge coupling, local rotation parameters, and representation of the $SU(2)_L$ weak isospin algebra and g'_W, $\alpha(x)$, and Y are their $U(1)_Y$ hypercharge algebra counterparts. The corresponding covariant derivate is

$$D_\mu = \partial_\mu - \frac{i g_W}{2} W^k_\mu T^k - \frac{i g'_W}{2} B_\mu Y, \tag{1.7}$$

where W^k_μ and B_μ are the emerging gauge fields associated to $SU(2)_L$ and $U(1)_Y$.

Each fermion can couple differently to the gauge fields according to the representation of the $SU(2)_L$ group they are part of. To determine the coupling constants it is important to separate each fermionic field into its left-handed and right-handed components such that

$$\psi = \psi_L + \psi_R \equiv \gamma^L \psi + \gamma^R \psi, \tag{1.8}$$

with

$$\gamma^{L/R} = \frac{1}{2}\left(1 \mp \gamma^5\right) \text{ and} \tag{1.9}$$

$$\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3, \tag{1.10}$$

where γ^μ are the Dirac matrices. It has been shown experimentally that only left-handed particles interact with the $SU(2)_L$ group, meaning that right-handed particles are part of the trivial representation of the group while left-handed particles

are part of the fundamental representation whose generators are given by Pauli's matrices. Particles are thus grouped as follows:

- leptons: e_R, μ_R, τ_R, $\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L$, $\begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L$, and $\begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L$;
- quarks: u_R, d_R, c_R, s_R, t_R, b_R, $\begin{pmatrix} u \\ d \end{pmatrix}_L$, $\begin{pmatrix} c \\ s \end{pmatrix}_L$, and $\begin{pmatrix} t \\ b \end{pmatrix}_L$.

Note that the right-handed neutrinos are not present as they do not interact with other particles and are thus sterile.

To each particle, a weak isospin T_3 can be associated, corresponding to the Casimir operator of $SU(2)_L$. Right-handed particles have $T_3 = 0$, positive quarks and neutrinos have $T_3 = \frac{1}{2}$, and negative quarks and charged leptons have $T_3 = -\frac{1}{2}$. The same can be done for the $U(1)_Y$ group for which an hypercharge Y is defined. Through breaking of the symmetry of the EWK group, as explained in the next section, a new relation can be obtained:

$$Q = T_3 + Y, \tag{1.11}$$

where Q is the electric charge of the particle.

1.3 Electroweak Symmetry Breaking

The physical fields of the photon A_μ and the weak interaction bosons $W^\pm_\mu$ and Z_μ are linear combinations of the B_μ and W^i_μ gauge fields. This is a consequence of the Higgs mechanism that gives mass to the $W^\pm$ and Z bosons but leaves the photon γ massless, thus breaking the gauge symmetry of EWK under $SU(2)_L \otimes U(1)_Y$ into a $U(1)_{em}$ group.

To account for the mass of particles, R. Brout, F. Englert, and P. W. Higgs postulated the existence of a scalar doublet in $SU(2)_L$ with hypercharge $Y = \frac{1}{2}$

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{1.12}$$

for which the Lagrangian can be written as

$$\mathcal{L} = (D_\mu \phi)^\dagger (D^\mu \phi) - V(\phi^\dagger \phi) \tag{1.13}$$

with D_μ the covariant derivate of the $SU(2)_L \otimes U(1)_Y$ gauge group, and V the potential of the so-called Higgs field. The chosen form of V,

$$V(\phi^\dagger\phi) = \lambda\left(\phi^\dagger\phi\right)^2 - \mu^2\phi^\dagger\phi, \tag{1.14}$$

is where the symmetry breaking occurs as the ground state of the system, $V(\phi^\dagger\phi) = 0$, is reached when

$$|\phi| = \sqrt{\frac{\mu^2}{2\lambda}} = \frac{v}{\sqrt{2}}. \tag{1.15}$$

The potential remains invariant under a transformation of three out of the four components of the Higgs field while a transformation along the last component breaks the $SU(2)_L$ symmetry. The spontaneous breakdown of the symmetry thus accounts for the creation of three Nambu-Goldstone bosons which give their masses to the three weak interaction bosons, and one new field H, the Higgs field. It can be shown that the masses of these particles are directly proportional to v:

$$m_H = \sqrt{2\lambda}v \tag{1.16}$$

$$m_Z = \frac{1}{2}\sqrt{g_W^2 + g_W^2}, \tag{1.17}$$

$$m_W = \frac{1}{2}vg. \tag{1.18}$$

Giving masses to fermionic fields occurs through coupling to the ϕ doublet which respects the gauge invariance of $SU(2)_L \otimes U(1)_Y$

$$\mathcal{L} = -\alpha\bar{\psi}_L\phi\psi_R + h.c. \tag{1.19}$$

Requiring that the photon field A_μ be massless, a linear combination of B_μ and W_μ^i must be done as follows

$$A_\mu = W_\mu^3 \sin\theta_W + B_\mu \cos\theta_W, \tag{1.20}$$

$$Z_\mu = W_\mu^3 \cos\theta_W - B_\mu \sin\theta_W, \tag{1.21}$$

$$W_\mu^\pm = \frac{1}{\sqrt{2}}\left(W_\mu^1 \mp iW_\mu^2\right) \text{ and} \tag{1.22}$$

with

$$\tan\theta_W = \frac{g'_W}{g_W}, \tag{1.23}$$

where θ_W is the weak mixing or Weinberg angle. From this transformation, it is straightforward to compute the coupling strengths of particles to the photon field and more particularly the coupling of the positron to the photon

$$e = g_w \sin\theta_W, \tag{1.24}$$

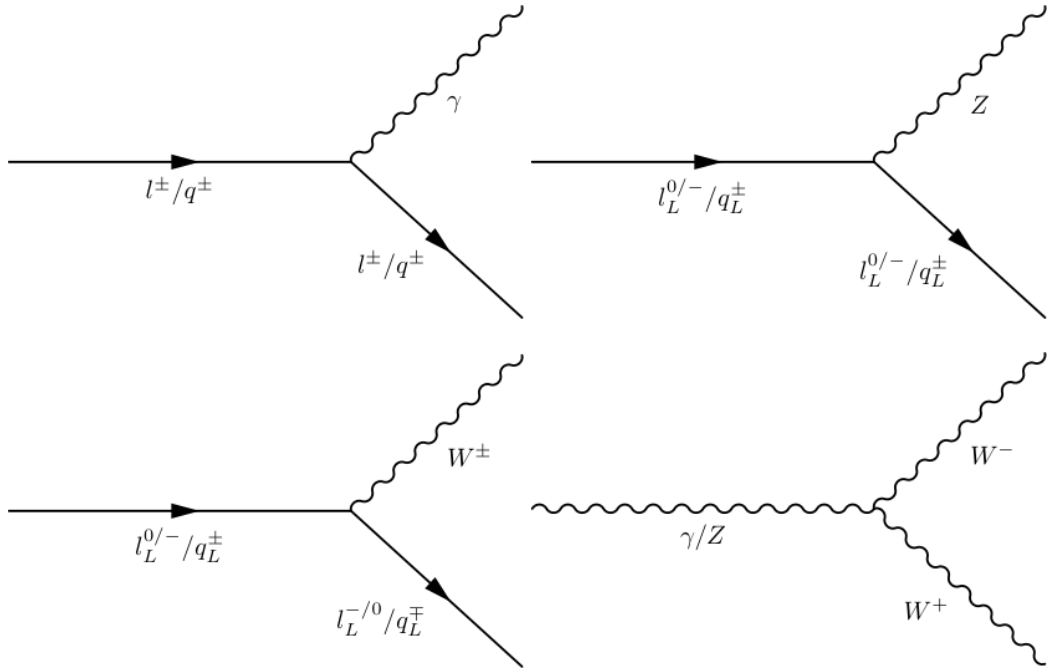

Figure 1.3 – Feynman diagrams representing the interaction of the γ, Z boson, and $W^{\pm}$ bosons to fermions. Top-left: coupling of γ to charged particles. Top-right: diagonal coupling of the Z boson to $SU(2)_L$ doublets. Bottom-left: non-diagonal coupling of the $W^{\pm}$ bosons to $SU(2)_L$ doublets. Bottom-right: self-coupling of the gauge bosons.

which is the charge of the particle. As expected, the coupling to the A_μ field is proportional to the electric charge and thus null for neutral particles. Z_μ on the other hand has a non-zero coupling constant to all particles belonging to the doublet representation of $SU(2)_L$ and is diagonal in matrix representation of the group. This means that it couples each particle to itself, while W_μ is non-diagonal thus providing a coupling between particles of the same doublet. Figure 1.3 summarizes these interactions by showing the coupling of the photon γ, Z boson, and $W^{\pm}$ bosons to fermions in, respectively, the top-left, top-right, and bottom-left diagrams and the coupling of the γ and Z to the $W^{\pm}$ in the bottom-right diagram.

1.4 Quantum Chromodynamics

QCD uses the $SU(3)_C$ gauge group to describe the behavior of quarks and gluons, and introduces the notion of color charge. After the observation of particles composed of three identical quarks in the same spin configuration, a new quantum number had to be introduced in order to obey Pauli's exclusion principle. Through their interaction with the $SU(3)_C$ group, quarks carry a color, arbitrarily labeled red, blue, or green, that is exchanged by the interaction with gluons which carry a color and anti-color combination. However, as given quark configurations were not observed experimentally, a constraint was added and forced hadrons, groups of quarks, to be colorless: either composed of a quark and an antiquark of opposite color for mesons, or composed of three quarks of different colors for baryons. Re-

cently, pentaquarks composed of four quarks and one antiquark have been observed by the Large Hadron Collider Beauty experiment [5].

Quarks are placed in the fundamental representation of the $SU(3)_C$ group which are triplets of the same quark flavor with three different colors, while leptons are singlets as they do not interact under QCD. The covariant derivate becomes

$$D_\mu = \partial_\mu - \frac{ig_S}{2} G^a_\mu \lambda^a \tag{1.25}$$

where g_S is the strong coupling constant, G^a are the eight gluon gauge fields, and λ^a are the Gell-Mann matrices. Gluons are massless as they do not interact with the Higgs fields through the EWK symmetry breaking.

The observation of flavor-changing charged current interactions, interactions mediated through a $W^\pm$ that couple quarks from different $SU(2)_L$ doublets, led N. Cabibbo, M. Kobayashi, and T. Maskawa to introduce the concept of mixing matrix for the quarks. They postulated that the eigenstates of quarks that interact with the EWK gauge bosons, those present in the doublets, are not the eigenstates of mass, thus the physical particles. By introducing a CKM matrix which for each quark redefines a linear combination of the other quarks, they were able to account for the violation of quark flavor in the EWK sector.

1.5 Beyond the Standard Model

While the Standard Model provides a very accurate description of the QCD and EWK theories, it bares some shortcomings in order to be considered the theory of everything.

The success of the unification of the weak and electromagnetic forces into a single EWK theory has motivated physicists to unify QCD and EWK into a single force at high energies, described by a new gauge invariance group. This implies a symmetry breaking mechanism similar to the Higgs mechanism at low energies and thus the existence of new particles of high mass undetected as of today.

QCD and EWK theories, although not yet unified, have successfully been combined in the Standard Model to account for strong, weak, and electromagnetic forces. However, including Einstein's theory of general relativity, which describes gravitational forces, still remains a challenge. Some theories propose to add a new boson to the model, the graviton, to account for gravity but fail to describe all phenomena observed experimentally. To this day, general relativity remains the prominent theory of gravitation, validated by the recent discovery of gravitational waves by the LIGO experiment in February 2016 [6].

Although general relativity provides predictions that have been tested experimentally with great precision, it fails to account for effects such as the excessive rotation speed of galaxies or gravitational lensing. Instead of rewriting Einstein's theory, the existence of dark matter and energy has been postulated which would interact gravitationally but not through any other force. From computation, only 5% of our universe would be composed of Standard Model particles while the remaining 95% would be made of yet unknown matter and energy. Physicists also postulate the existence of a hidden sector which new exotic particles interact only rarely with ordinary matter and at higher energy scales.

Finally, the Standard Model as it is written fails to give mass to the neutrinos while experimental results have proven that neutrinos, although very light, carry a mass. Moreover, neutrino oscillations between generations have been observed by Super-Kamiokande in 1998 [7], meaning that the Standard Model neutrinos, those placed in the $SU(2)_L$ doublets are not the mass eigenstates of the particles.

These examples represent only a few of the questions that remain open. To answer some of them, physicists need to reach higher energies in order to explore new sectors of physics unaccessible under normal conditions. To reach that goal, powerful machines such as particle colliders have been built over the last century. The latest achievement in this field is the construction of the Large Hadron Collider, which has already yielded great success with the discovery of the Higgs boson.

The Large Hadron Collider

2

The Large Hadron Collider (LHC) [8] is a hadron accelerator and collider installed in a 26.7 km long tunnel beneath the Franco-Swiss border near Geneva, at a depth varying from 170 m below the Jura mountains to 45 m below the Leman lake. The tunnel was built by the European Organization for Nuclear Research (CERN) between 1984 and 1989 to host the former Large Electron Positron collider (LEP). As represented in Figure 2.1, which provides a schematic illustration of the LHC, it is composed of eight arcs, eight straight sections (in between the arcs), and two transfer tunnels which connect the LHC to CERN's main injection complex. From the eight possible collision points located in the straight sections of the tunnel, only four are in use and instrumented with a total of seven particle detectors: ALICE [9], ATLAS [10], CMS [11], LHCb [12], TOTEM [13], LHCf [14], and MoEDAL [15].

The construction of the LHC, for which the concept dates back to 1984, was approved by the CERN Council in December 1994 and started in 1998 with the excavation of the caverns that would hold the experimentation sites. In 2003, the first section of the accelerator was assembled inside the tunnel marking the start of the installation

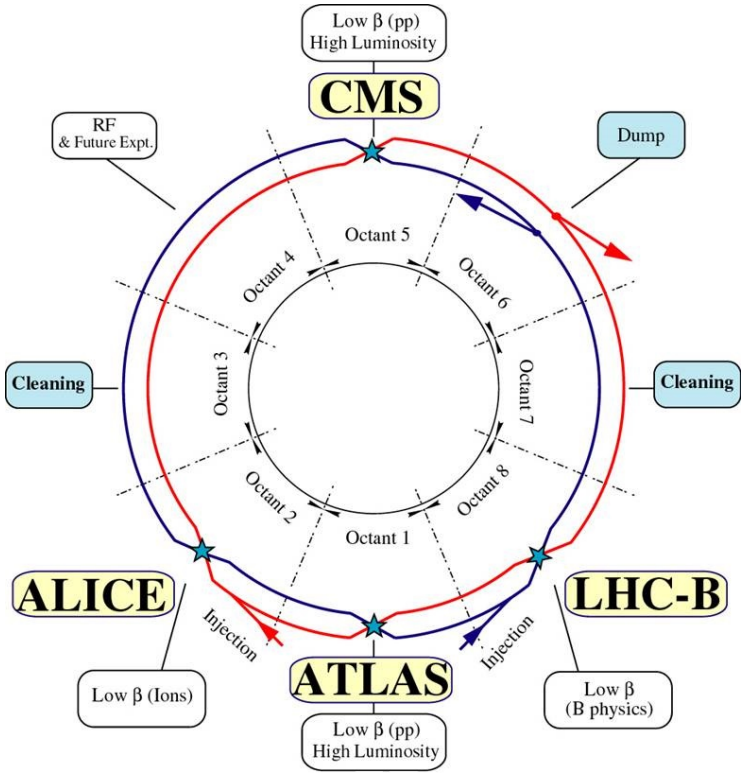

Figure 2.1 – Schematic representation of the Large Hadron Collider and the four main experiments (ALICE, ATLAS, CMS, and LHCb) that analyze the produced collisions [8].

phase of the LHC and its detectors, which would span until 2008. On September 10th 2008, the first beam of protons was circulated inside the LHC only to reveal a major technical issue with the superconducting magnets which halted the machine for nearly a year. In November 2009, reparation works were finished just in time for the LHC to produce its first collisions at 2.36 TeV before a Year-End Technical Stop (YETS), a yearly technical stop during which maintenance is performed on the machine. In February 2010, the LHC restarted with a physics program with collisions at 7 TeV that lasted until February 2013. At that time, the LHC stopped for two years during a Long Shutdown (LS) in order to perform upgrades to prepare it to run at 14 TeV. On June 3rd 2015, the LHC restarted and set a new record with an energy in the center of mass reference frame of 13 TeV.

2.1 The Injection Complex

The particles that enter the LHC are first created and accelerated by the injection complex of CERN which network of accelerators is depicted in Figure 2.2. Protons are extracted from gaseous hydrogen by means of a duoplasmatron, a device that uses a heated filament cathode in conjunction with electric fields to produce electrons that will ionize and break the H_2 gas. The resulting protons have an energy of 100 keV and are injected in the Linac2 linear accelerator.

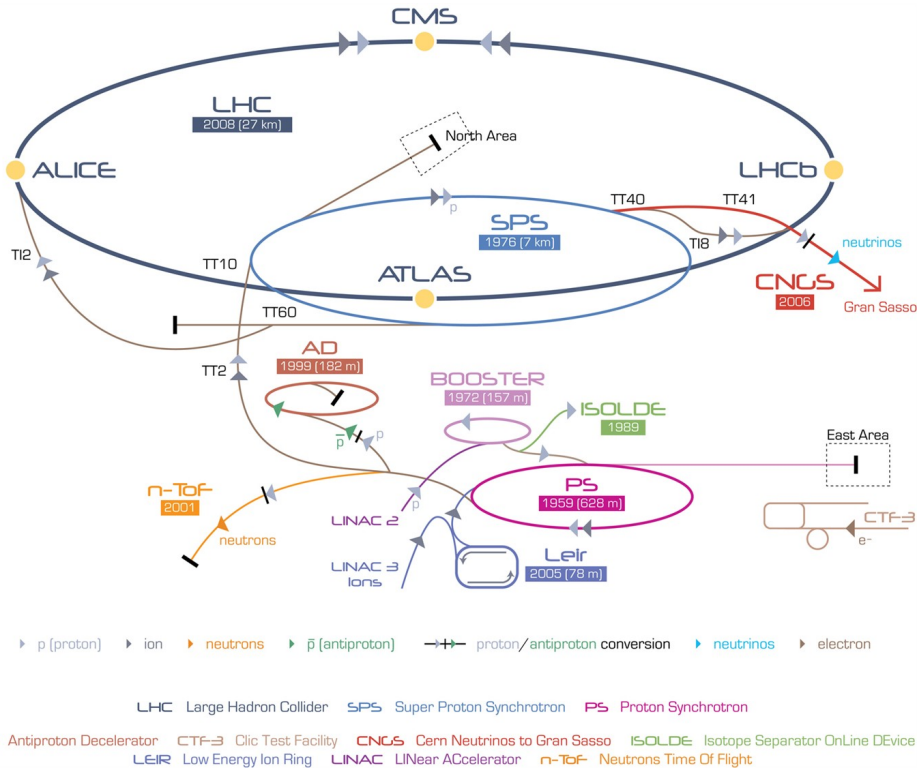

Figure 2.2 – Schematic diagram of the injection chain of the LHC composed of multiple smaller accelerators [16].

Linac2 uses radio frequency cavities that produce an electromagnetic field inside the accelerator to transfer energy to the charged particles. The oscillations of the field allow the formation of bunches of particles by regrouping them on the front of the radio wave. The succession of cavities that form Linac2 boosts the incoming protons to an energy of 50 MeV before they enter the Booster.

Following Linac2, three consecutive circular synchrotrons, namely the Proton Synchrotron Booster (Booster), the Proton Synchrotron (PS), and the Super Proton Synchrotron (SPS), accelerate the beams to energies of respectively 1.4 GeV, 25 GeV, and 450 GeV. Each accelerator is composed of electromagnets that bend the trajectory of the beam and boost it further. From the SPS, the beam is injected in the LHC through two transfer tunnels into two rings where particles rotate in opposite directions to be further accelerated until they reach the desired energy. The SPS also provides beam to the North Area where various experiments take place and future technologies can be tested during test beam campaigns, frequently organized by CERN.

2.2 LHC Technical Description

Once inside the LHC, particles travel through two beam pipes which are subject to a vacuum of 10^{-10} Torr, representing around 3 000 000 molecules per cm^3, which greatly reduces the collisions with the air and allows for greater beam longevity. The beam pipes do not form a perfect circle as they are composed of eight arcs and eight straight sections called insertions.

Each arc contains 154 powerful dipole magnets, represented in Figure 2.3, which produce a magnetic field higher than 8 T to bend the beams. The fields are generated by passing high currents through a coil of NbTi superconductor that surrounds the beam pipes cooled by liquid Helium bellow 2 K. The dipoles use a two-in-one design with a common cooling and housing system for the magnets of the two beam pipes which allows to reduce cost and space, and generate a magnetic flux circulating in opposite direction for the two rings. Due to the low temperature at which the magnets operate, the minimum energy deposition left by particles in the coils needed to trigger a quench, a sudden loss of superconductivity, is also decreased. A tight control of the beam structure must thus be enforced. Therefore, 49 quadrupole magnets are installed per section in order to focus the beams and reduce horizontal and vertical spread of the particles.

The eight straight sections of the LHC have different use cases and designs: four of them are dedicated to the experiments and provide them with collisions, one contains the radio frequency cavities to accelerate the beams, two are used to clean the beam, and one to dump the beam. These functions are fulfilled by using various

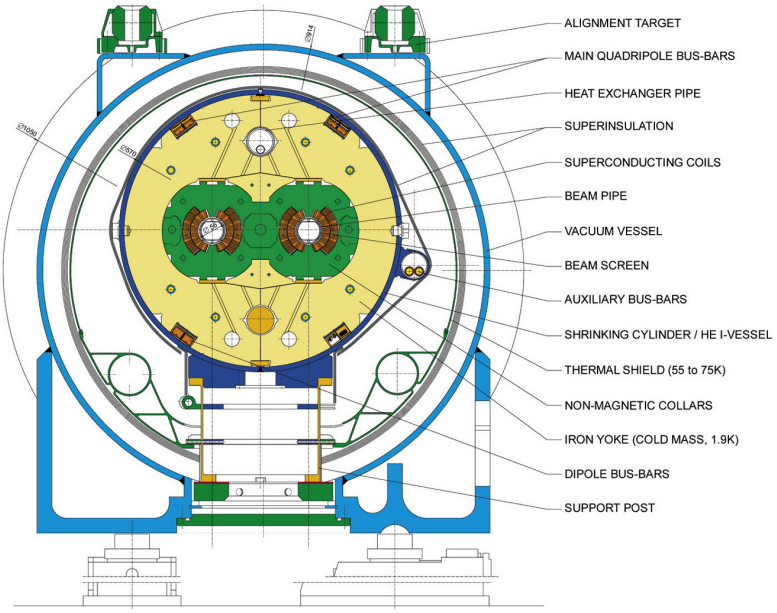

Figure 2.3 – Cross-section of a cryodipole of the LHC representing the two beam pipes equipped with superconducting NbTi coils surrounded by iron cold mass at 1.9 K [8].

magnet designs to bend or deflect the beams. From the four collision sites, two also host the beam injection pipes from the SPS into the LHC. The beam cleaning insertions are used to remove particles with out-of-band momentum by scattering them on collimators. When the integrity of the beam is compromised, fast switching magnets activate at the dump insertion in order to quickly extract the particles from the LHC and redirect them to the dump site: an eight meter long graphite composite block that absorbs the energy of the beam.

2.3 Performance Goals

The beams inside the LHC rings are composed of 2 808 bunches each made of approximately 110 billion protons. They are separated by 25 ns which is the time between two consecutive bunch crossing (BX), yielding a collision frequency of 40 MHz. During a collision, only a small fraction of the protons interact depending on various parameters such as the crossing angle of the beams, collimation of the bunches, etc. To quantify this, the notion of luminosity has to be introduced and is directly related to the frequency of apparition of events in a given interaction process

$$f_{process} = \mathcal{L}\sigma_{process} \ , \tag{2.1}$$

where $f_{process}$ is the number of expected events per second, and $\sigma_{process}$ is the interaction cross-section of the process. For a circular collider, the instantaneous luminosity is defined as

$$\mathcal{L} = f\frac{n_1 n_2}{4\pi\sigma_x\sigma_y}, \tag{2.2}$$

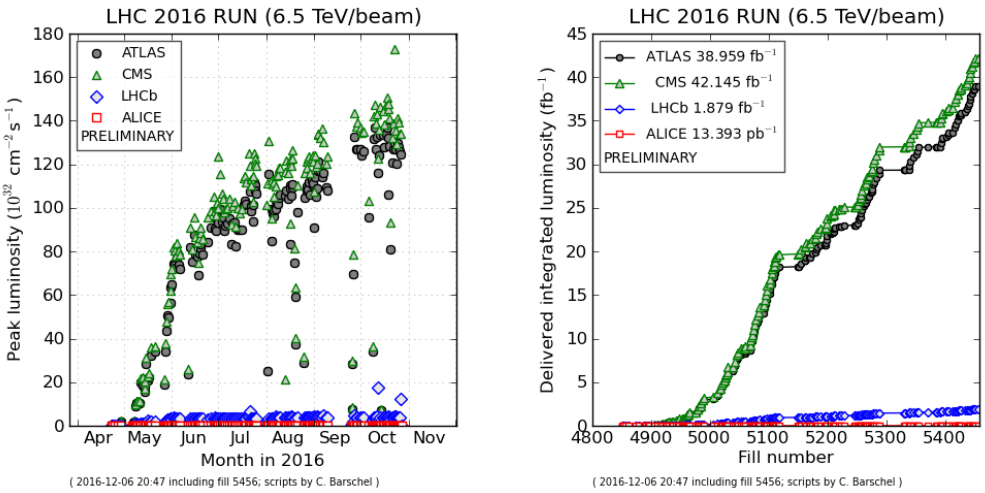

Figure 2.4 – Monthly evolution of the peak delivered luminosity (left) for the four main experiments of the LHC and the corresponding integrated luminosity (right) [17].

where n_i is the number of protons or ions per bunch, f is the frequency of the collisions, and σ is the RMS of the transverse size of the beam in the x and y directions. These parameters change during the operation of the machine as the number of protons per bunch decreases and the bunches spread out.

The instantaneous luminosity can be accumulated over a given period of time in order to obtain the integrated luminosity

$$L = \int \mathcal{L}\, dt, \tag{2.3}$$

which results in the number of events one can expect for a given interaction process

$$N_{process} = L\sigma_{process}. \tag{2.4}$$

Figure 2.4 shows the peak and integrated luminosities delivered by the LHC to its four main experiments during 2016. It can be seen in the picture on the left that the highest instantaneous luminosity reached during that period is on the order of 14×10^{33} cm^{-2} s^{-1}.

2.4 Operations and Schedule

Over the past years, the LHC has been building up in energy but also in luminosity in order to reach its nominal values of 14 TeV and 10^{34} cm^{-2} s^{-1}. Figure 2.5 depicts the schedule of the LHC for the coming years throughout 2037 and the corresponding luminosities: peak on the left axis and integrated on the right axis. It can be seen that several LS are planned in order to upgrade the LHC and allow the experiments to perform maintenance and improvements. A noticeable increase in peak luminosity

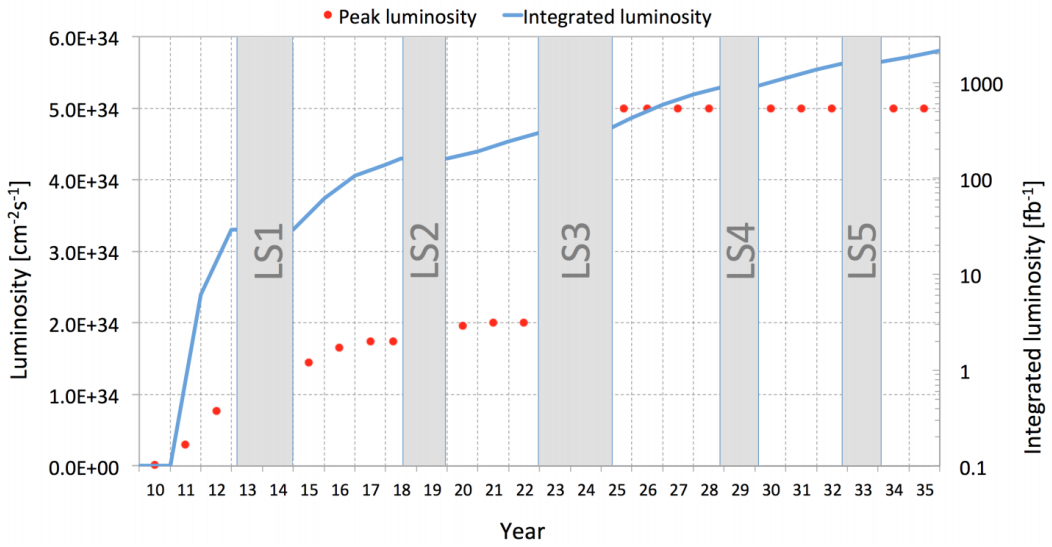

Figure 2.5 – Projected instantaneous and integrated luminosity of the LHC throughout 2037 [4].

will occur after LS3 and mark the beginning of the so-called Phase II operation of the LHC. Such an increase in luminosity will affect the experiments due to the higher number of proton-proton collisions during each BX which will in turn translate to a higher flux of particles in the detectors. Preparing the LHC and the detectors to function at such a high luminosity is a challenge to which physicists and engineers are already providing answers.

2.5 Scientific Motivations for the LHC Upgrade

The high energy and luminosity at which the LHC operates are the key factors of its success. The new energy levels it unlocks enable scientists to explore hidden sectors of physics as new particles can be created through processes that were previously impossible to generate. Moreover, due to the rarity of these events that lie beyond the Standard Model, it is of importance to accumulate extensive statistics and thus run the machine at high luminosity. Not only physics beyond the Standard Model is being explored by the LHC. Measuring the free parameters of the theory and comparing them to theoretical and experimentally obtained values is also an important task which helps validate the models. Running at high luminosity allows the observation of rare processes such as the B_S^0 meson decay, $\bar{b}s \rightarrow \mu + \bar{\mu}$, which was first observed at the LHC [18].

The discovery of the Higgs boson is one of the main goals of the LHC which has already been fulfilled. However, its characterization still requires a great amount of analysis in order to, for example, quantify the coupling constants to fermions or weak bosons, which today have a precision of about 20%. These fine measurements make use of the study of rare events which will greatly benefit from the upgrade

of the LHC towards higher luminosity, to reach a percent-level precision for most coupling measurements. Moreover, it is not excluded that other scalar bosons exist with lower production rates as predicted by some theories. In this case, the upgrade of the LHC becomes a necessity to detect the new exotic physics at play.

The very high integrated luminosity that will be collected during the LHC Phase II will open the doors to searches for new heavy gauge bosons, extra dimensions, lepto-quarks, dark matter candidates, etc. Their respective models predicted an extremely low interaction cross section and thus require extensive data to be detected. Any discovery in this field would be ground breaking and would open a new chapter in the history of particle physics.

The measurements and observations described above are done by the detectors which are installed on the LHC using the collisions it provides. From the four large experiments the LHC hosts, two of them, namely ATLAS and CMS, are considered to be generalist as they are built to detect a wide range of interactions. ALICE and LHCb on the other hand are respectively optimized to study Pb-Pb interactions at lower energy which result in more particles in the final state, and b quark physics to understand the CP violation. The three smaller experiments, namely TOTEM, LHCf, and MoEDAL, use the forward particles detected far away from the interaction point to respectively analyze elastic collisions, simulate cosmic rays, and search for magnetic monopoles.

3 The Compact Muon Solenoid Experiment

The Compact Muon Solenoid (CMS) [11] is a multi-purpose particle detector recording the collisions provided by the LHC. It was, along with ATLAS, the first experiment officially approved for the LHC by the CERN research board in January 1997 after a long evaluation process of the letter of intent published four years earlier by the collaboration. The construction of the detector started in 2005, after the excavation works of the cavern finished, and spanned until 2008. Since then, CMS has been proficiently taking and analyzing data, and announced on July 4th 2012 the discovery of the Higgs boson, one of the most significant results of the collaboration.

3.1 Overview

At nominal energy and luminosity, the LHC produces around 20 proton-proton interactions per BX which results in more than 1 000 particles in the final state. In order to distinguish physical processes with great precision, CMS has to ensure accurate identification and reconstruction of the particles. To this end, the detector was built around four requirements, as stated in [11]:

- Good muon identification and momentum resolution over a wide range of momenta and angles, good dimuon mass resolution ($\approx$ 1% at 100 GeV[1]), and the ability to determine unambiguously the charge of muons with momentum $<$ 1 TeV;
- Good charged-particle momentum resolution and reconstruction efficiency in the inner tracker. Efficient triggering and offline tagging of τ's and b-jets, requiring pixel detectors close to the interaction region;
- Good electromagnetic energy resolution, good diphoton and dielectron mass resolution ($\approx$ 1% at 100 GeV), wide geometric coverage, π^0 rejection, and efficient photon and lepton isolation at high luminosities;
- Good missing-transverse-energy and dijet-mass resolution, requiring hadron calorimeters with a large hermetic geometric coverage and with fine lateral segmentation.

These requirements are met through the subdivision of CMS in various detection systems each specialized in the reconstruction of a given type of particles. The overall

[1]The widely used convention that sets $c = \hbar = 1$ is used through out this work in order to express energies and masses in electronvolts. A factor of c^{-1} and c^{-2} should respectively be applied to these numbers in order to obtain the values in the international unit system.

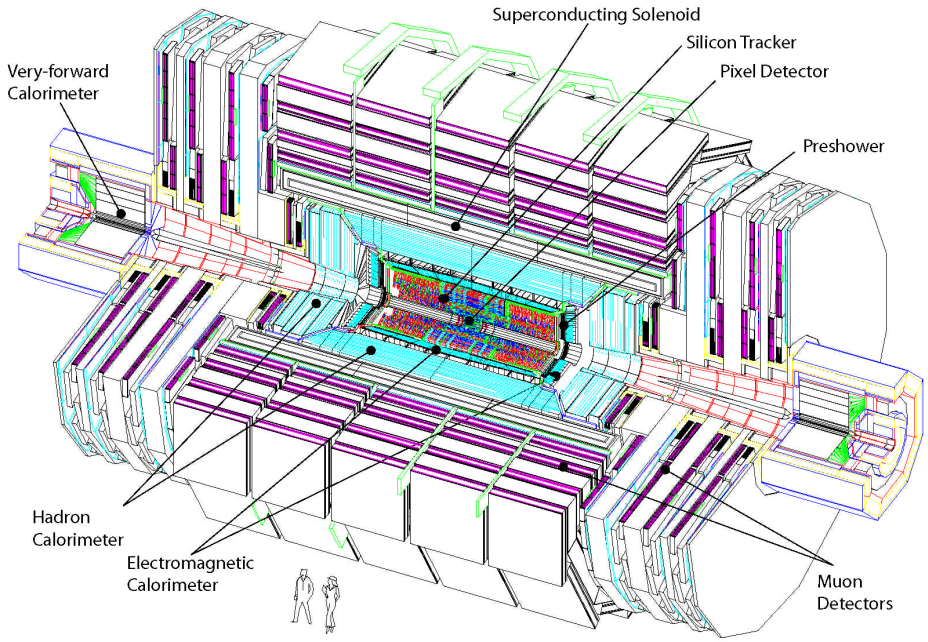

Figure 3.1 – The Compact Muon Solenoid detector and its subsystems installed at LHC [11].

layout of CMS, shown in Figure 3.1, is divided into the barrel and the two endcaps, regions where the detectors are respectively placed parallel and perpendicular to the beam pipe. At the center of the detector, within a radius of 1.25 m around the beam pipe, lies the inner tracking system. Composed of layers of silicon pixels and silicon strips detectors, it is designed to provide around 10 positions on the trajectory of any charged particle with a spatial resolution as high as 10 μm. Surrounding the tracking system are the electromagnetic and hadron calorimeters which respectively measure the energy of electrons and photons, and hadrons through the annihilation of the latter within the detectors. These systems are placed inside a superconducting solenoid magnet which produces a strong 3.8 T field that bends charged particles and allows for precise momentum measurements. The obtained resolution on the transverse momentum is so that $\frac{\Delta p_T}{p_T} \approx 1\%$ for 100 GeV muons in the barrel. Outside of the magnet, three different technologies of muon detectors are placed on large iron yokes.

The coordinate system used in CMS has its origin at the nominal interaction point of the beams, the y-axis pointing upwards, the x-axis pointing toward the center of the LHC, and the z-axis directed along the beam direction. The polar coordinates (r, ϕ) are defined in the x-y transverse plane and the widely used pseudorapidity η is taken to be

$$\eta = -\ln\left(\tan\left(\frac{\theta}{2}\right)\right), \tag{3.1}$$

where θ is the azimuthal angle between the z-axis and the transverse plane.

3.2 The Inner Tracking System

The inner tracking system of CMS provides a spatial resolution on the order of 50 μm for the passage of charged particles. The proximity of the detectors to the beam pipe yields a high particle density of up to 10^8 particles per cm^2 at a radius of 4 cm. This justifies the need for a high granularity and fast response time to unambiguously assign particles to the correct collision. This is achieved through a density of up to 4 000 channels per cm^2 and a time resolution of 6 ns. However, these features come with an elevated power consumption of a few microwatts per channel which in turn requires a cooling infrastructure to maintain the tracker at a temperature of -20 ^{o}C. In order to minimize the material budget of the tracker, a compromise was found to use two technologies: silicon pixels close to the beam pipe to provide a high granularity, and silicon strips further away to minimize the material budget.

The layout of the detectors is shown in Figure 3.2. The silicon pixels (PIXEL) are installed on three cylindrical layers in the barrel and two disks in each endcap. These detectors cover an area between radii of 4.4 cm and 10.2 cm and a pseudorapidity $|\eta|$ < 2.5. The silicon strips (TIB, TID, TOB, and TEC) occupy a radial region between 20 cm and 116 cm and are divided in different sectors with a total of 10 layers in the barrel and 12 disks in each endcap. The segmentation into sectors corresponds to the various strip geometries and arrangement.

The barrel pixel tracker contains 768 modules for a total of 48 million pixels. The disks of the endcaps are divided into 24 segments each composed of 7 modules which yield a total number of 672 modules with 18 million pixels for both endcaps. Each pixel is 150 μm $\times$ 100 μm in size with a thickness of 260 μm to 300 μm resulting in a single point hit resolution of approximately 10 μm in the transverse coordinate and 20 μm to 40 μm in the longitudinal coordinate.

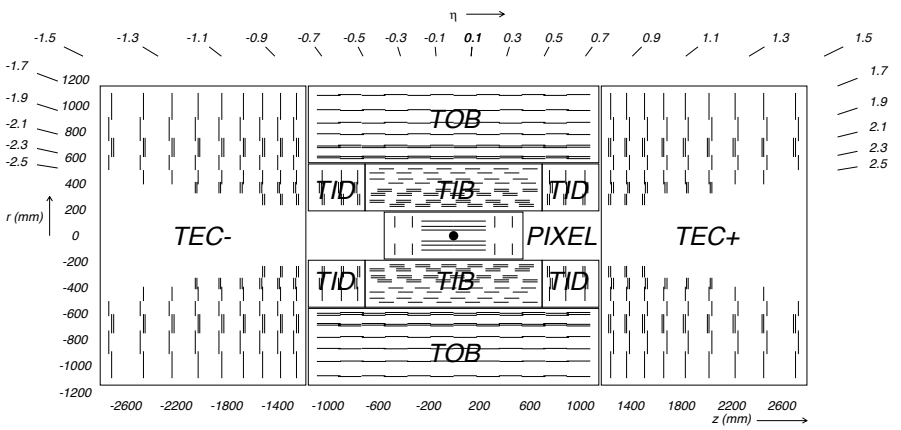

Figure 3.2 – Disposition of the silicon pixels (PIXEL) and strips (TIB, TID, TOB, and TEC) modules inside the tracker of CMS [11].

The silicon strip tracker is divided into three subsystems. The Tracker Inner Barrel and Disks (TIB and TID) are composed of 4 layers in the barrel and 3 disks in each endcap, covering a region up to a radius of 55 cm. The sensors are 320 μm thick and run parallel to the beam pipe in the barrel and radially in the disks. The strip pitch is of 80 μm in the two first layers of the TIB and 120 μm in the two next ones, resulting in a spatial resolution of respectively 23 μm and 35 μm. In the TID, the pitch varies between 100 μm and 141 μm. Surrounding the TIB and TID is the Tracker Outer Barrel (TOB) composed of 6 barrel layers of 500 μm thick sensors. It provides a resolution of 53 μm for the first four layers and 35 μm for the two last ones. Closing the TOB on both sides are the Tracker EndCaps (TEC) which each holds 9 disks. Additionally, the modules of the two inner layers of the TIB and TOB, the two inner rings of the TID and TEC, and the fifth ring of the TEC are equipped with a second strip detector mounted back-to-back with a stereo angle of 100 mrad and provides a measurement of the second coordinate.

The detection efficiency above 99% and spatial resolution below 53 μm of the tracker allows it to achieve a track reconstruction efficiency above 99% with a transverse momentum resolution better than 3% for 100 GeV muons. Figure 3.3 depicts the tracking efficiency measured with a tag-and-probe technique, for muons from Z decays, as a function of the muon η (left) and the number of reconstructed primary vertices in the event (right) for data (black dots) and simulation (blue bands).

The resulting material budget of the tracker is shown in Figure 3.4 in which it is expressed in terms of radiation length X_0 on the left and hadronic interaction length

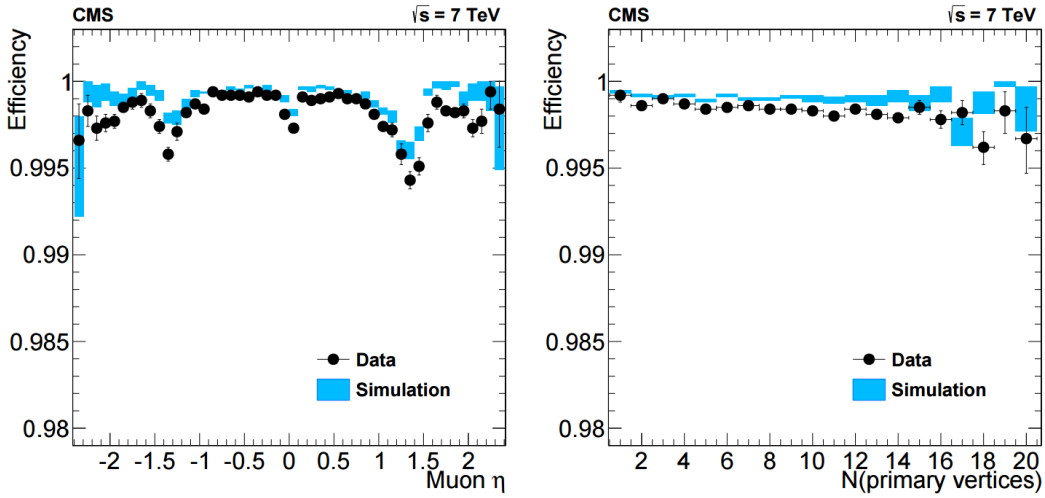

Figure 3.3 – Tracking efficiency measured with a tag-and-probe technique, for muons from Z decays, as a function of the muon η (left) and the number of reconstructed primary vertices in the event (right) for data (black dots) and simulation (blue bands) [19].

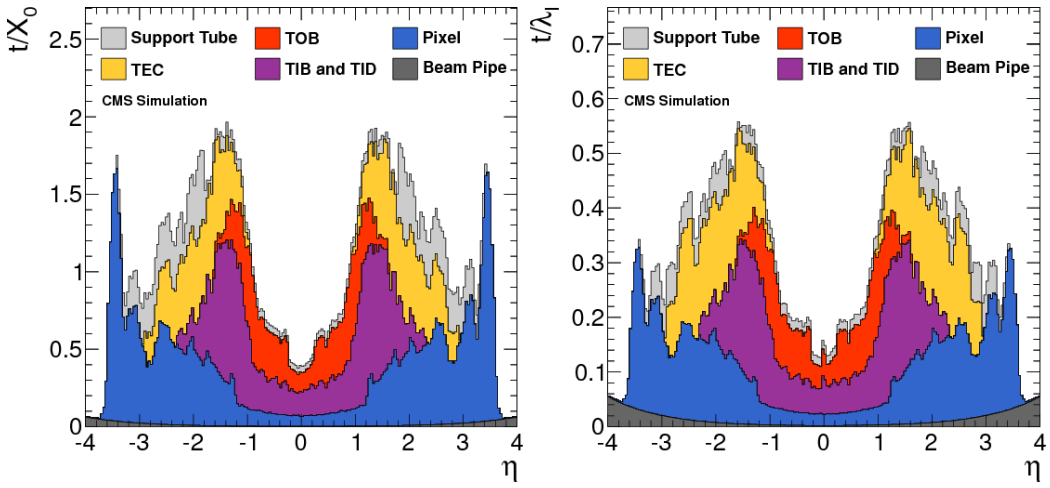

Figure 3.4 – Material budget of the tracker expressed in terms of radiation length X_0 on the left and hadronic interaction length λ_I on the right [19].

λ_I on the right. From the quantity of material in the tracker, 70% of the photons materialize into an e^+/e^- pair and 5% of 5 GeV pions interact within its volume.

3.3 The Electromagnetic Calorimeter

The Electromagnetic Calorimeter (ECAL) of CMS has been designed to provide an energy resolution for electrons and photons of the order of 1% for 100 GeV. It is composed of lead tungsten ($PbWO_4$) crystals that act as homogeneous calorimeters: they both initiate the particle cascades and provide the proportional response to the energy deposition. The barrel counts 61 200 crystals and is closed by 7 324 crystals in each endcap. Additionally, a silicon preshower detector is installed in front of the crystals in the endcaps to improve the detection of collimated particles.

The structure of the ECAL is shown in Figure 3.5. In the barrel, the ECAL covers a pseudorapidity range $|\eta| < 1.479$ and is divided in 360 sections in ϕ and 170 in η. The cross section of the crystals varies between 22 mm $\times$ 22 mm to 26 mm $\times$ 26 mm for a length of 230 mm corresponding to 25.8 radiation lengths. An alveolar structure maintains crystals together in submodules which are grouped into modules of 400 to 500 crystals further assembled in supermodules containing 1700 crystals in total. In the endcaps, modules cover a pseudorapidity of $1.479 < |\eta| < 3.0$ with each crystal having a front cross-section of 28.62 mm $\times$ 28.62 mm and a length of 220 mm. These are grouped in units of 5 $\times$ 5 crystals called supercrystals further inserted in one of the two *Dees* that compose an endcap and maintains the ECAL in place. The preshower uses lead radiators to initiate the avalanche that will be detected by silicon strips to measure the deposited energy and the avalanche profile. Each sensitive module has an active area of 61 mm $\times$ 61 mm divided among 32 strips. The preshower is installed in the $1.653 < |\eta| < 2.6$ pseudorapidity region.

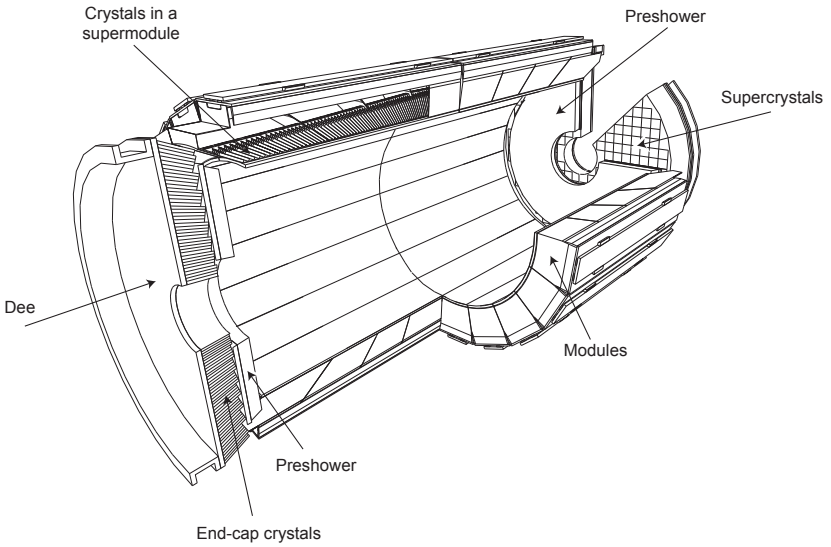

Figure 3.5 – Layout of the ECAL showing the disposition of the crystals, modules, supermodules, and preshower [11].

The energy resolution of the ECAL for particles with energy bellow 500 GeV is parametrized by the following equation:

$$\left(\frac{\sigma}{E[GeV]}\right)^2 = \left(\frac{S}{\sqrt{E[GeV]}}\right)^2 + \left(\frac{N}{E[GeV]}\right)^2 + C^2 \tag{3.2}$$

where S is the stochastic term, N the noise term, and C the constant term. The main contributions to the stochastic term are related to fluctuations in the lateral shower containment, photostatistics contributions, and fluctuations in the energy deposition inside the preshower radiators compared to what is measured by the silicon detector. The noise term includes noise from the electronics, the digitization, and pile-up. Finally, the constant term accounts for non-uniformities in the crystals, calibration errors, and leakage of energy from the back of the crystals. From test beams, the values of S, N, and E, were found to be so that:

$$\left(\frac{\sigma}{E[GeV]}\right)^2 = \left(\frac{2.8\%}{\sqrt{E[GeV]}}\right)^2 + \left(\frac{0.12}{E[GeV]}\right)^2 + (0.30\%)^2 \tag{3.3}$$

3.4 The Hadron Calorimeter

The Hadron Calorimeter (HCAL) provides a measurement of the energy of the hadrons. It is divided into a barrel region (HB), two endcaps regions (HE), and an outer calorimeter (HO) placed right outside the magnet as represented in Figure 3.6. The HB sits between the ECAL and the magnet at radii of 1.77 m and 2.95 m respectively, which limits the amount of material and thus absorption lengths that can be used, justifying the need for the HO to catch the tail of the cascades. An additional forward calorimeter (HF) is placed in the very forward region of CMS to

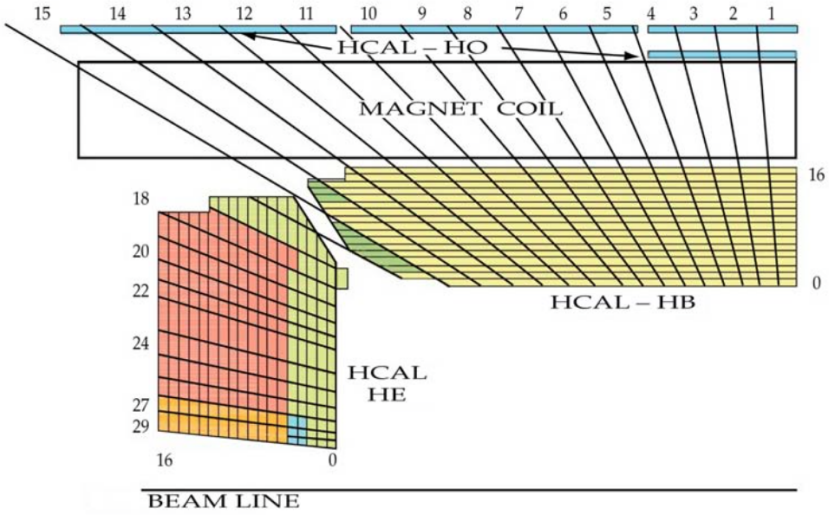

Figure 3.6 – Longitudinal view of the HCAL showing the disposition of the barrel, endcaps, and outer calorimeters [11].

analyze particles that are produced or scattered at $\eta > 3$.

The HB uses a sampling calorimeter design that consists of a succession of metal sheets positioned parallel to the beam that act as absorber, and plastic scintillators to measure the energy deposition. The absorber is made of a 40-mm-thick front steel panel, followed by eight 50.5-mm-thick brass sheets, six 56.5-mm-thick brass panels, and a 75-mm-thick steel back plate for a total minimum of 5.82 interaction lengths. The scintillators are designed out of 3.7-mm-thick Kuraray SCSN81 plastic and placed after each layer of metal. An additional layer of 9-mm-thick Bicron BC408 is placed in front of the first steel panel in order to sample hadronic showers that might have formed in the ECAL. The scintillator is divided into 16 η and 72 ϕ sectors resulting in a segmentation of $\Delta\eta \times \Delta\phi = 0.087 \times 0.087$. The HE uses 79-mm-thick brass plates spaced by 9 mm gaps in which the plastic Bicron BC408 scintillators are inserted. The segmentation of the HE results in a granularity similar to the HB for $|\eta| < 1.6$ and of $\Delta\eta \times \Delta\phi = 0.17 \times 0.17$ for $|\eta| \leq 1.6$. In the $|\eta| < 1.3$ region, hadron showers are not entirely contained in the HB and HE thus justifying the installation of the HO outside of the magnet to catch the trail of the cascades. This has a significant impact on high energy cascades and the measurement of the missing energy of an event, i.e. the energy carried by non-interacting particles. The HO has a minimum thickness of 1.4 interaction lengths and is segmented in 5 rings further divided into segments. Each segment has a granularity similar to the HB of 0.087×0.087 in $\Delta\eta \times \Delta\phi$. Limited by the muon system, the HO has been allocated 40 mm of space, from which 16 mm are used for the scintillators while remaining space is filled with support material.

The energy resolution of the HCAL can be expressed in the same manner as the ECAL, breaking down the result into a stochastic term and a constant term. The obtained resolutions [20] are

$$\left(\frac{\sigma}{E[GeV]}\right)^2 = \left(\frac{120\%}{\sqrt{E[GeV]}}\right)^2 + (9.5\%)^2 \tag{3.4}$$

for the HB and HE combined, and

$$\left(\frac{\sigma}{E[GeV]}\right)^2 = \left(\frac{280\%}{\sqrt{E[GeV]}}\right)^2 + (11\%)^2 \tag{3.5}$$

for the HO.

3.5 The Superconducting Magnet

The superconducting magnet of CMS measures 6 m of diameter for 12.5 m in length and has been designed to generate a 4 T magnetic field. The return of the magnetic field lines is done through the steel yokes that support the muon chambers of CMS as represented in Figure 3.7, in which the amplitude of the magnetic field has been simulated. The magnet is composed of four layers of wound NbTi conductor which offers very low resistance to the nominal 19.14 kA current that is required to generate the field. To reach the superconducting state of the magnet, the latter needs to be cooled down to 1.8 K using liquid Helium. The magnetic field thereby created bends charged particles according to their transverse momentum in the x-y plane

$$R = \frac{p_T}{0.3B}, \tag{3.6}$$

where R is the curvature radius of the trajectory in the transverse plane, p_T is the transverse momentum of the particle expressed in GeV, and B is the intensity of the magnetic field in T.

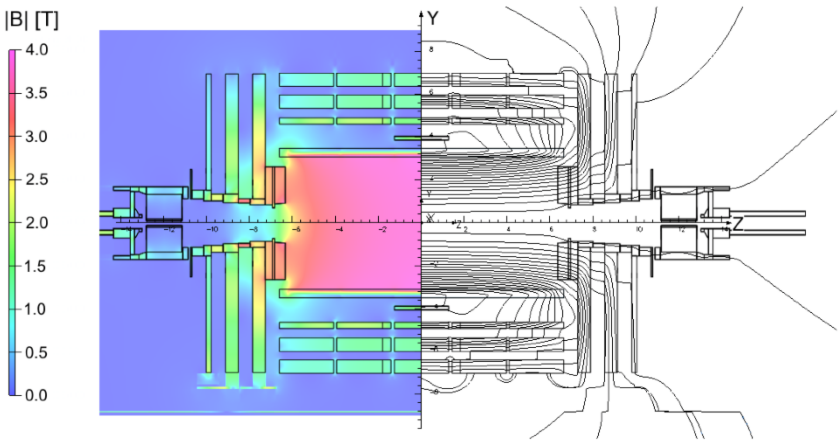

Figure 3.7 – Representation of the amplitude of the magnetic field and the field lines predicted on a longitudinal section of the CMS detector using simulations [21].

3.6 The Muon System

Muons are a recognizable signature from the high background noise produced by the large number of proton-proton interactions. Moreover, they are present in many final states that characterize interesting processes from the Standard Model and beyond. Therefore, providing identification capabilities better than 98% is essential. CMS is equipped with three different muon detectors that together form the muon spectrometer: Drift Tubes (DTs), Cathode Strip Chambers (CSCs), and Resistive Plate Chambers (RPCs). DTs and CSCs are respectively placed in the barrel and the endcaps while RPCs are installed in both regions, adding redundancy to the system by providing additional measurement points. Figure 3.8 represents a quadrant of CMS highlighting the placement of the various technologies inside the detector.

In the barrel, four cylindrical layers of DTs, numbered MB1 to MB4, cover a pseudorapidity range of $|\eta| < 1.2$. The first three layers each contain twelve chambers: eight to measure the r-ϕ bending angle and four to measure the z direction. The fourth station solely tracks the r-ϕ parameter using eight chambers.

In the endcaps, the CSCs cover a region of $0.9 < |\eta| < 2.4$ and are spread over four disks where detectors are placed perpendicular to the beam pipe. Each disk is further divided into two or three rings numbered MEx/y, where x is the disk number

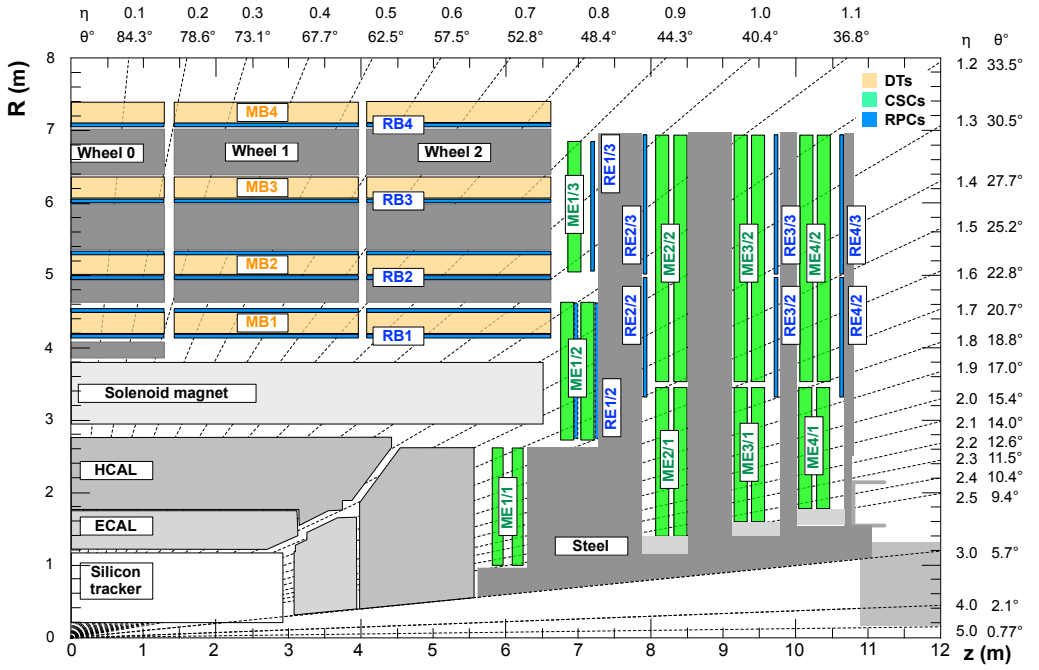

Figure 3.8 – Layout of the muon spectrometer of CMS depicting the three gaseous detectors in use: Drift Tubes (DTs) in orange labeled MBx, Cathode Strip Chambers (CSCs) in green labeled MEx/y, and Resistive Plate Chambers (RPCs) in blue labeled RBx and REx/y [11].

and y is the ring number. CSCs measure both the r-ϕ and the η coordinates using respectively cathode strips that run radially and anode wires that run perpendicular to the strips. Each module is composed of 6 layers of CSCs which provide precise background noise rejection and tracking.

Additionally, RPCs are installed along with the DTs and CSCs label RBx and REx/y respectively. In the barrel, two stations of RPCs are installed in RB1 and RB2 while only one station is instrumented in RB3 and RB4. In the endcaps, one module of RPCs equips each of the stations except for the first ring of each disk which has been left empty. This is due to the high rate of particles in the region close to the beam pipe, up to 50 kHz cm^{-2}, which is not supported by the RPCs.

3.6.1 Drift Tubes

The barrel muon spectrometer consists in 4 cylindrical layers of DT modules: the inner three contain 60 chambers and the outer has 70 for a total of 172 000 sensitive wires. The smallest unit composing the DTs is the drift cell: a 2.4 m long by 42 mm wide tube in which a wire is stretched over the whole length as depicted in Figure 3.9. Filled with a gas mixture of Ar/CO_2, they provide a measurement of the point of impact of the particle in the direction perpendicular to the wire using the ionization of the gas amplified through the applied high voltage. The small flux of particles in the barrel region of CMS, of the order of 2 Hz cm^{-2}, enables the use of these long chambers which limits the number of active channels and thus cost while still providing a spatial resolution in the 80-120 μm range. Multiple drift cells are combined into layers and four layers are further staggered by half a cell width in order to eliminate dead space and form a super layer (SL). SLs vary in size between 1.9 m in MB1 and 4.1 m in MB4. Each chamber contains two SLs that measure the r-ϕ bending angle and one SL that provides a measurement in the z direction. MB4, the fourth station, however does not include the SLs in the z direction. The spatial resolution of a station is on the order of 100 μm on the position and 1 mrad on the direction of the particle, and of 5 ns for the time resolution.

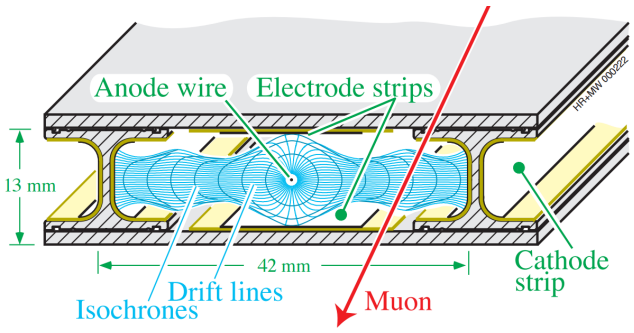

Figure 3.9 – Sketch of the drift cells highlighting the electric field lines converging towards the anode wire [11].

3.6.2 Cathode Strip Chambers

The CSCs are multiwire proportional chambers made of 6 anode wire planes interleaved among 7 cathode strip panels. The largest chambers measure 3.4 m $\times$ 1.5 m and are installed in the ME2/2, ME3/2, and ME4/2 stations while smaller chambers are mounted in the other stations limited by the steel yoke layout. Taking advantage of the multiple detection layers of the chambers, CSCs can provide a detection efficiency better than 99%, 75 μm resolution in the r-ϕ coordinate for ME1/1 and ME1/2 and about 150 μm elsewhere, and a BX assignment efficiency of 99%. The time resolution of the CSCs is on the order of 5 ns.

The structure of the CSCs is shown in Figure 3.10. Seven 16.2-mm-thick trapezoidal panels made of 12.7-mm-thick polycarbonate core with two 1.6 mm FR4 skins glued on each side are assembled in order to form one chamber. FR4 is a material used to create printed circuit board. On six of the planes, 36 μm thick copper strips are milled on one side to form the cathode planes. Gap bars are glued to the planes in order to form six gaps of 9.5 mm once the planes are stacked together. About 1000 anode wires are wound around three of the panels in order to form the anode planes. The 50-μm-thick gold-plated tungsten wires are separated by 3.2 mm and elevated by 4.75 mm with respect to the plane in order to sit in the middle of the gap.

3.6.3 Resistive Plate Chambers

RPCs are used for their excellent time resolution of less than 3 ns which allows them to unambiguously assign events to their corresponding BX while also providing

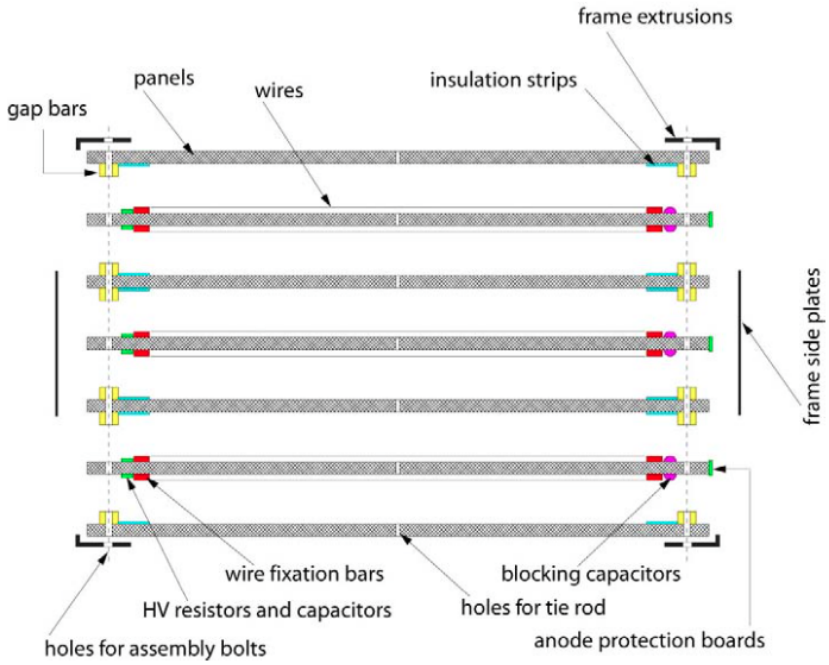

Figure 3.10 – Schematic view of a CSC detailing the different layers that compose it [11].

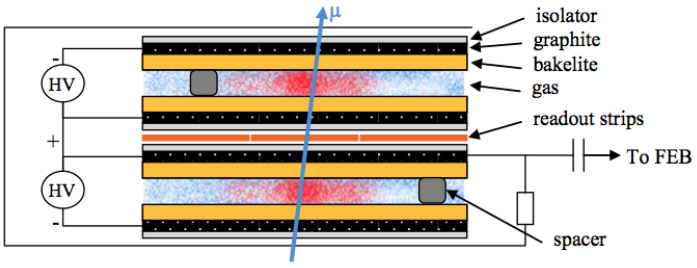

Figure 3.11 – Diagram of the RPCs double gap layout [11].

a spatial resolution of the order of the centimeter. Each module consists of two gas filled gaps, with a common read-out strip layer in the middle, as depicted in Figure 3.11. On each side of the gaps, an insulating backelite panel is covered with a conducting graphite layer on which a high voltage is applied. This layout allows to operate each gap at a lower high voltage and reduce the readout time required, thus improving the time resolution.

In the barrel, RB1 and RB2 each include two RPCs placed on both sides of the DTs. RB3 and RB4 only use one RPC mounted in front of the DTs. In total, 480 chambers are installed, each with a length varying between 2.455 m long and a width between 1.5 m and 2.08 m. In the endcaps, RPCs cover a pseudorapidity range of $|\eta| < 1.6$ divided into two rings. Chambers are trapezoidal and cover 20^o in ϕ for the inner ring and 10^o for the outer ring. This setup leaves an empty region between $1.6 < |\eta| < 2.45$ where CSCs are the only detectors providing measurements.

3.7 The Trigger System

At a luminosity of 10^{34} cm^{-2} s^{-1}, approximately 20 proton-proton interactions resulting from the collisions of the LHC occur every 25 ns, yielding more than 1 000 particles in the final state. Each collision generates more than one MB of data which makes it impossible to process and store as is. Therefore, two trigger stages have been implemented to reduce the 40 MHz rate of collisions to 100 kHz at first and down to 1.5 kHz afterwards. The Level-1 Trigger (L1 Trigger) consists of custom electronics designed to quickly discriminate between the events. In less than 3.2 μs the system has to take the decision to forward the event to the next stage or drop it and lose the data. Accepted events are passed to the High Level Trigger (HLT) which runs on a farm of commercial computers. Unlike the L1 Trigger, which due to the small allocated computation time has restricted access to hit information from the muon detectors and calorimeters, the HLT can use the full information from the detectors to reconstruct events.

At the L1 Trigger stage, only the information from the HCAL, ECAL, and muon chambers are used as depicted in Figure 3.12. The energy depositions collected

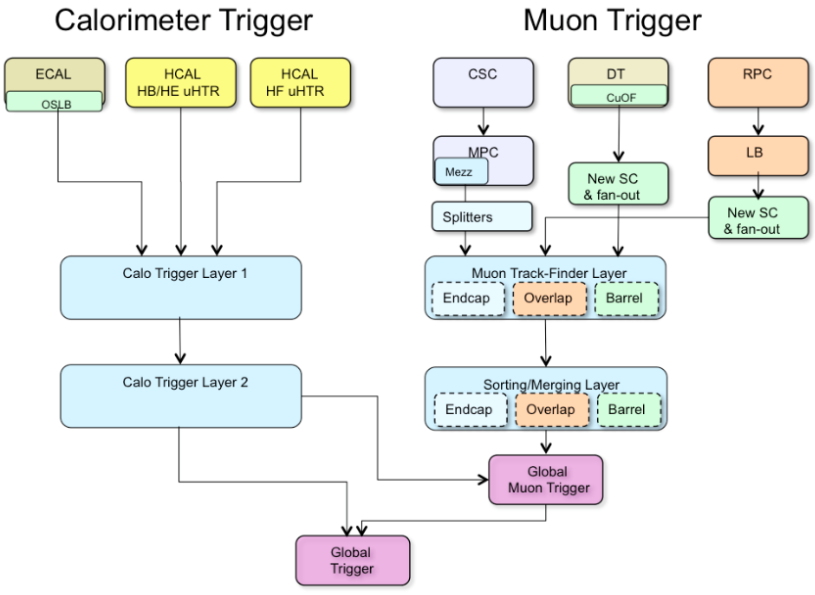

Figure 3.12 – Architecture of the CMS trigger system for both the calorimeter trigger and the muon trigger [22].

by the HCAL and ECAL are sent to the Calo Trigger Layer 1 which performs data formatting and concentration, and forwards events to the Calo Trigger Layer 2. The latter finds the twelve highest transverse energy jet, tau, and electron/photon candidates and computes global energy sums.

The muon trigger is divided into three sections: the barrel, the endcaps, and the overlap region where tracks cross both CSCs and DTs. Hits from the different chambers are combined in the Muon Track-Finder layer which uses track extrapolation and pattern recognition algorithms to create tracks. The reconstructed tracks are sent to the sorting/merging layer which selects the best candidates according to quality parameters. The best tracks are sent to the Global Muon Trigger (GMT) which uses information from the calorimeters to compute isolation values for each muon. This combined information is sent to the Global Trigger (GT) along with the data of the calorimeter.

The GT implements a menu of triggers that perform selections on the reconstructed objects such as electrons, photons, muons, jets, or missing energy and make the decision to accept or reject the events. If an event is accepted, a Level-1 Accept (L1A) is issued and sent to the Trigger Control and Distribution System (TCDS) [22]. The latter forwards the L1A to all the subsystems using the Timing, Trigger and Control (TTC) interface according to the readiness of the subdetectors and the data acquisition (DAQ) system. Upon reception of a L1A, the front-end electronics provides the full granularity data to the read-out DAQ which includes the HLT.

3.8 The Data Acquisition System

The DAQ system of CMS [23], shown in Figure 3.13, reads out the full granularity data of events from more than 650 sources in the whole detector. It is able to handle a data flow of 100 GB per second and redirect data to the HLT which reduces the rate of events by a factor of 1 000 and writes the accepted events to disk.

For every collision, the front-end electronic modules of each subsystem store the full granularity data in buffers and forward reduced information to the trigger system, based on which the latter will issue a L1A or not through the TCDS system. Upon reception of an L1A, the Front-end drivers (FEDs), specific to each subsystem, aggregate and pack the data from multiple modules and push the information uplink to the central DAQ of CMS. The data is received by either the new FrontEnd Readout Optical Link (FEROL) in fragments of 8 kB through 6 Gbps or 10 Gbps optical links or by the legacy FrontEnd Readout Link (FRL) in fragments of 4 kB through 400 MBps electric links. These boards are Peripheral Component Interconnect Express (PCIx) cards that are hosted in computers in the underground service cavern. The FEROLs relay their data and the data from the FRL within the same computer to the surface control building using a simplified version of the Transmission Control Protocol (TCP) over a 10 Gbps Ethernet link. A total of 576 10 Gbps Ethernet links

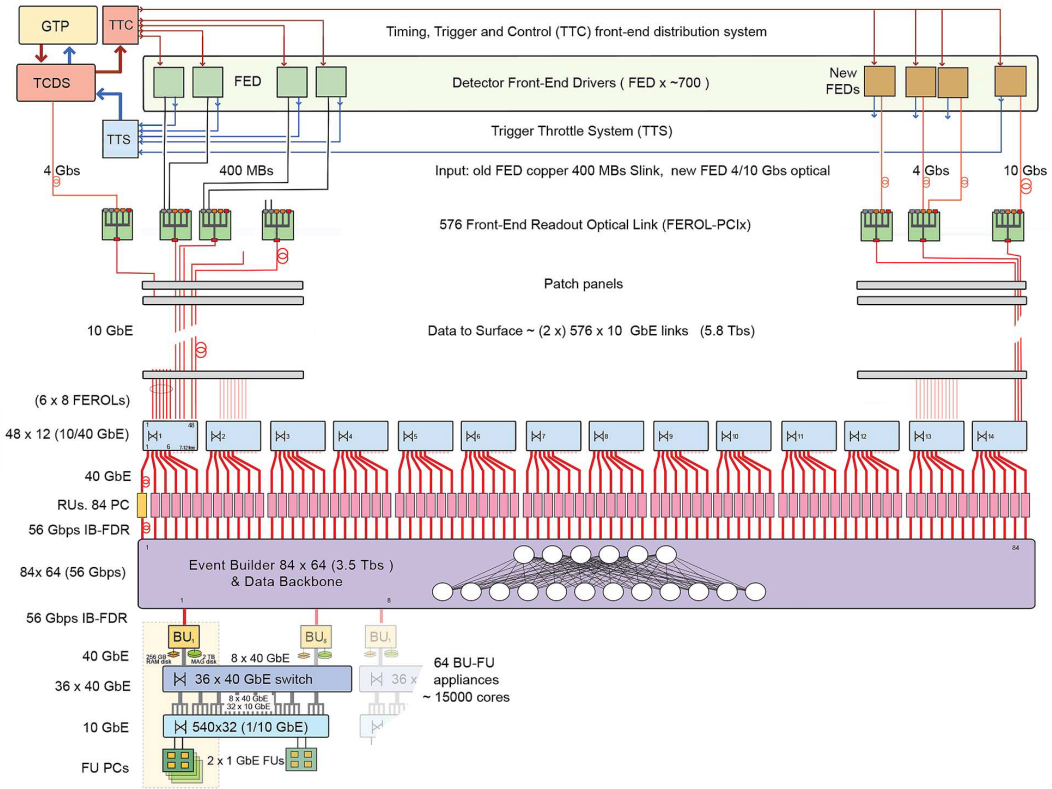

Figure 3.13 – Architecture of the DAQ system of CMS [23].

are used to transmit the data from the cavern of CMS to the surface.

On the surface, a second step of data concentration is performed by 12 Ethernet switches which each merge the data from 48 FEROLs onto 6 40 Gbps links. These are then connected to a total of 84 Readout Units (RUs). Each RU holds data from a reduced number of detectors from multiple events. In order to provide the full information of a single event to the 64 Builder Units (BU), a complex switching network is installed between the RUs and BUs. The switch is able to handle data transfers up to 6 Tbps from which only 3.5 Tbps are used, allowing for future extensions of the DAQ system. Once the full information of an event is collected by a BU, the latter stores the data to be reconstructed by the Filter Unit (FU) on disk. Each BU can store up to 256 GB corresponding to two minutes of data taking. The RUs, BUs, and FUs all use commercially available computers that run custom software.

The FUs connect to the BUs by mounting an image of their hard-drive through a 1/10/40 Gbps Ethernet network. They are in charge of the event reconstruction by running the HLT algorithms programmed in the CMS Software (CMSSW). The results of the HLT are written back on the BUs and then packaged with the original data. The data from the accepted events is then sent from the BUs to the CERN computing center for storage.

3.9 The Upgrades of CMS

Following the maintenance schedule of the LHC, CMS has and will perform upgrades of its detectors to replace aging components and make use of new technologies. Due to the ambient radiation inside the cavern, the aging process of the detectors and their electronics is accelerated. Replacing these elements is thus vital in order to preserve excellent performance. Furthermore, with the increase in luminosity of the LHC, the flux of particles inside CMS will rise, yielding more difficult conditions to detect, identify, and reconstruct events. To solve this challenge, new technologies are being investigated by various collaborations within CMS.

Part II

The CMS GEM Upgrade Project

In the coming years, several upgrades of the LHC and its injection chain are planned with the aim of increasing the performance of the machine. The next upgrade, LS2, is currently set to take place in 2019 and will increase the instantaneous luminosity of the LHC to 2×10^{34} cm^{-2} s^{-1} for the 2020-2022 run. In 2023, the LHC Phase 1 will end with a total integrated luminosity of around 300 fb^{-1} and LS3 will start, allowing the preparation of the systems for the so called high-luminosity LHC or LHC Phase II. The LHC will restart mid-2025 with an instantaneous luminosity of 5 $\times 10^{34}$ cm^{-2} s^{-1}.

The muon spectrometer of CMS must be able to cope with the program of the LHC. It was originally designed to be an hermetic and redundant system relying on the DTs, CSCs, and RPCs to provide efficient muon detection, identification, and triggering. The DTs are installed in the barrel and cover a region of $|\eta| < 1.2$, and the CSCs are located in the endcaps between $1.0 < |\eta| < 2.4$. Additionally, the RPCs provide redundancy in both the barrel and the endcaps but were not instrumented in the $|\eta| > 1.6$ region due to concerns about their tolerance to high fluxes of particles.

Aside from the lack of redundancy, studies have shown that the triggering efficiency of the muon system of CMS will be drastically affected by the increase in luminosity and reduce the physics performance by limiting the parameter phase-space that can be studied.

To tackle these challenges, the CMS GEM collaboration [24] will install during LS2 a set of muon detectors that use the Gas Electron Multiplier (GEM) technology in the $1.6 < |\eta| < 2.2$ region left vacant by the RPCs. The objective is to improve the trigger efficiency in combination with the CSCs. To this end, a new DAQ system has to be designed based on the micro telecommunications computing architecture standard. Small scale systems have been studied during test beam campaigns and five GEM superchambers equipped with the full DAQ system will be installed in CMS during the YETS-2016 to perform a so-called slice test.

This part focuses on the CMS GEM upgrade project and the developments of the author with respect to the design and characterization of the DAQ system. Chapter 4 provides an overview of the GEM technology and the status of the project. It touches on the evolution of the chamber design and the different generations of GEM detectors that have been developed, and covers the performance of the detectors. The architecture of the DAQ system is explained, in details, in Chapter 5 describing each component of the system and their integration in the global CMS DAQ system. Chapter 6 focuses on the test beam campaigns that took place in November 2014 and 2015 and helped to qualify the system providing results on the chamber performance. Following the test beams, modifications have been made to the electronics in order to prepare for the slice test and the final system which are further detailed. Finally, Chapters 7 and 8 describe the work performed on the qualification and calibration, and the irradiation tests of the on-detector electronics used in the GEM project.

All firmware and software developments, analysis, and results exposed in Chapters 6, 7, and 8, have been performed solely by the author unless otherwise specified.

Gas Electron Multiplier Detectors 4

In 1968, G. Charpak revolutionized the field of tracking detectors by introducing the Multiwire Proportional Chamber (MWPC). MWPCs were more reliable than the existing single-wire gas counters providing a higher sub-mm resolution and rate capabilities up to fluxes of several MHz mm^{-2}. Developments of manufacturing techniques and high-density electronics led to a new generation of detectors fulfilling the needs of high-energy physics experimentation. However, their use in high-luminosity experiments revealed some intrinsic weaknesses: the mechanical design of the chamber limits the spacial resolution; long ion drift times affect the efficiency of the detector at high fluxes; solid deposits agglomerate on the wires due to aging, and induce sparks in the chamber.

Twenty years later, Micro-Strip Gas Counters (MSGCs) [25] were developed using photolithographic processes to engrave anode and cathode strips on an insulating support. Both the position resolution and the rate capability were increased by several orders of magnitude. Unfortunately, the detectors were prone to the effect of aging and discharges, irremediably damaging the counters.

The success of the MSGCs led to the development of Micro-Pattern Gas Detectors (MPGDs) less susceptible to discharges and offering comparable performances. They use micro pattern structures to amplify the signal and reduce the probability of sparks. Two technologies have proven to be operational: the MICRO-MEsh GASeous structure (MICROMEGAS) [26] and the Gas Electron Multiplier (GEM) [27].

Already in use in the COMPASS and TOTEM experiments at CERN, GEM detectors have proven to meet the requirements of high-luminosity environments. In 2009, a dedicated R&D program was launched to study the feasibility of the installation of GEMs in the muon spectrometer of CMS to instrument the positions left vacant by RPCs. Six years later, the so called GE1/1 [24] muon detector upgrade project was approved by CMS to be installed during LS2.

4.1 Motivations for the GE1/1 Upgrade

During LS2, that will take place in 2019-2020, the CMS GEM collaboration will install an additional set of muon detectors in CMS in the $1.6 < |\eta| < 2.2$ region. The so-called GE1/1 detectors will be installed near the ME1/1 CSC station as shown in Figure 4.1 which highlights the location of the chambers within the muon spectrometer in a longitudinal view of a quadrant of CMS.

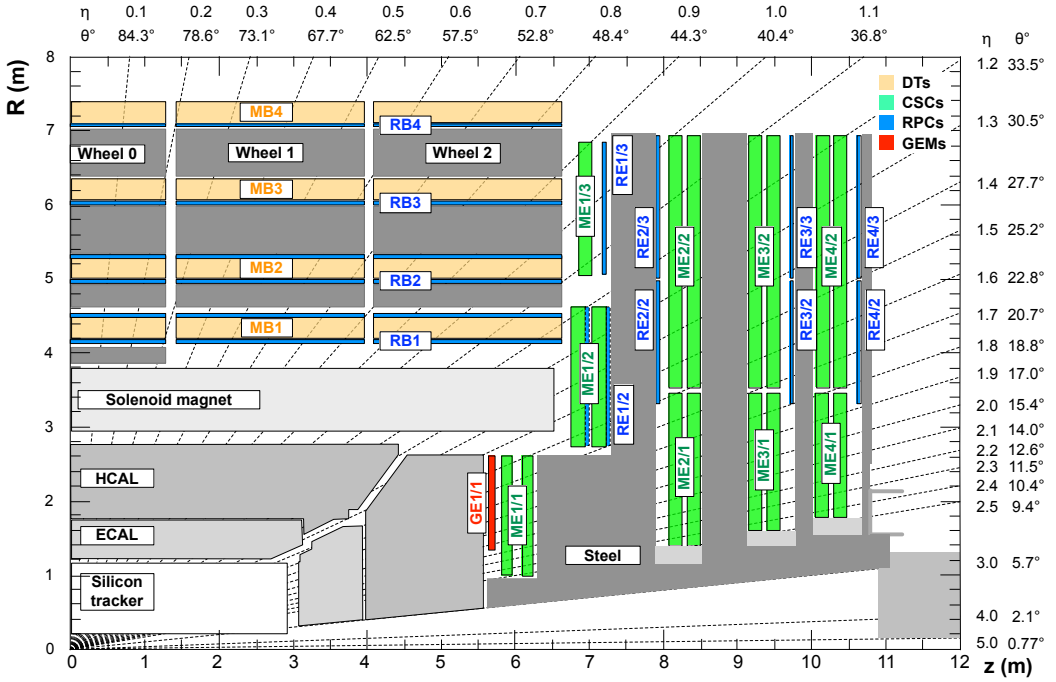

Figure 4.1 – Longitudinal view of a quadrant of CMS highlighting the location of the GE1/1 detectors colored in red within the muon spectrometer. DTs are represented in yellow, CSCs in green, and RPCs in blue [24].

The main motivation for the installation of GE1/1 is the significant impact it has on the triggering system of CMS. Left as is, the performance of the current system would degrade in the coming years with the increase in instantaneous luminosity due to the rate limitations of the ME1/1 station exposed to an intense flux of particles. Muon triggering will in turn suffer from a degradation in the resolution on the transverse momentum, p_T. As muon triggers are limited in bandwidth to a fraction of the total allocated bandwidth of the trigger, typically 5 kHz for the single muon trigger, they must set a cut on the minimum p_T of the muons (around 20 GeV). In the scenario where the muon spectrometer remains unmodified, the threshold would have to be raised to $p_T \approx 30$ GeV which would be a considerable drawback in terms of physics analysis, decreasing the acceptance by more than 35% for given channels.

In order to prevent the degradation of the trigger system, the GEM collaboration investigated the possibility to use the bending angle of tracks between GE1/1 and ME1/1 to estimate the p_T of muons, which trajectory is curved by the magnetic field. Figure 4.2 shows on the left the lever arm given by the 20 cm to 46 cm distance between the two detectors. The variation in distance corresponds to "close" and "far" chambers which are staggered in order to allow for some overlap and avoid dead space. From this, the picture on the right displays the discrimination power between 5 GeV/c and 20 GeV/c muons.

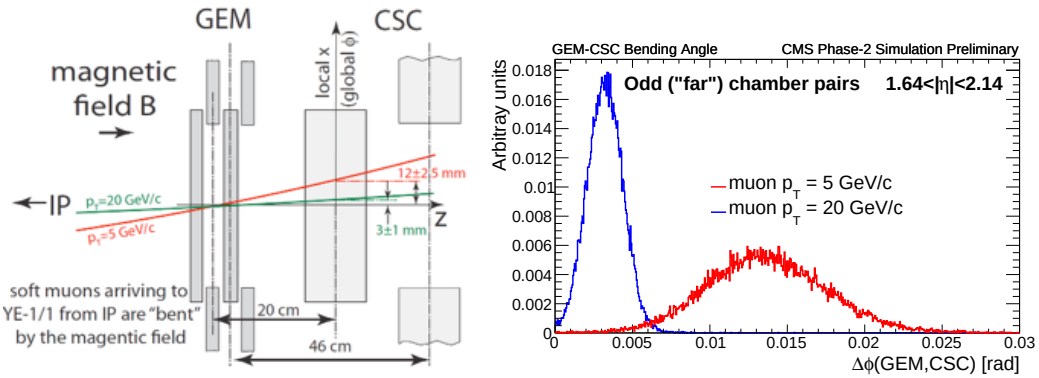

Figure 4.2 – Left: sketch of the measurement of the bending angle with a pair of CSC and GEM chambers, illustrating the difference between lower and higher momentum muons. Right: $\Delta\phi = \phi_{GE1/1} - \phi_{ME1/1}$ distribution for 5 GeV/c and 20 GeV/c p_T muons demonstrating the discrimination power of the combined GEM-CSC system [24].

Using these results, a common effort from the GEM and CSC collaborations ensued to integrate the GEM measurements in the CSC trigger system to provide an additional layer of information. By the means of simulations, it has been shown that the efficiency of the combined system increases, thus yielding a lower rate of fake triggers. These results are shown in Figure 4.3. In the left plot, which shows the efficiency as a function of the pseudorapidity of the particles, it can be seen that the efficiency of the GEM and CSC combined trigger (blue) is superior to the CSC only results (red), recovering the losses occurring at higher luminosity. In the right plot which shows the trigger rate as a function of the cut applied on the transverse momentum of the particle, it is depicted how the trigger rate of the combined system (blue) diminishes by one order of magnitude compared to the current system (red). For a given trigger rate, a smaller muon p_T threshold can be applied due to the lower rate of misreconstructed particles, which improves the performance of the detectors.

Preserving a low p_T threshold is essential for physics analysis to yield a reliable reconstruction efficiency for soft muons, that is muons with low momentum. These particles are important for a wide range of processes from beyond the Standard Model searches to measurements in the Higgs sector. Simulation studies have shown that in the case of $H \rightarrow \tau^+\tau^-$ where one τ decays into a muon and the other hadronically, a decrease of 5 GeV in the p_T threshold results in an increase of 35% in the acceptance of the channel. In addition, the impact is not limited to the single muon trigger but affects all other triggers involving muons analysis, for example, $\mu + jet$ or $e/\gamma + \mu$ signals.

To further improve the trigger efficiency of CMS, it is planned to integrate a new track trigger, a trigger system for the tracker, which would benefit from the high resolution of the silicon subdetector. Foreseen to be installed during LS3, the system

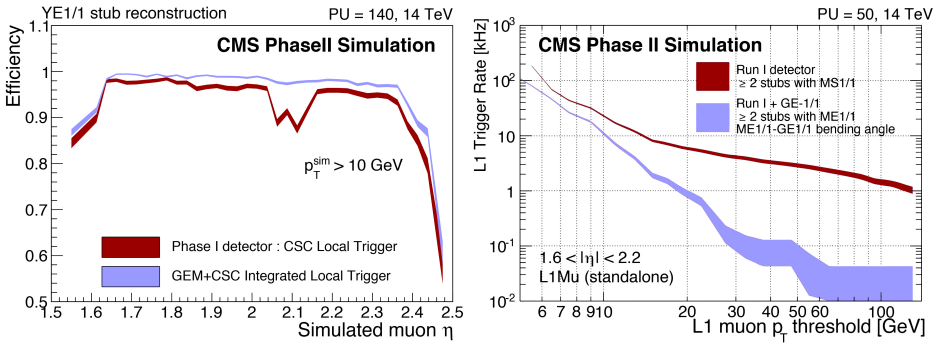

Figure 4.3 – Left: reconstruction efficiency of the integrated GEM-CSC trigger as a function of the simulated muon $|\eta|$, compared to the same for the Phase-I CSC-only algorithm. Right: level 1 muon trigger rates before and after the GE1/1 upgrade at a luminosity of 2×10^{34} cm^{-2} s^{-1}. MS1/1 denotes the first endcap muon station level 1 trigger in both cases, i.e. with CSC-only or with the combination CSC and GEM trigger information [24].

yields excellent results in the identification, reconstruction, and triggering of muons. However, due to the high occupancy of hits and the limited amount of computational time allocated to the trigger, given approximations must be done, degrading the results for certain physics channels. One of the assumptions made by the algorithms is that every particle originates from the interaction point. This has a significant impact on new physics processes involving hidden sectors of low interacting particles such as the decay of a Higgs into two neutralinos n_1, which in turn create a dark photon γ_d

$$h \rightarrow 2n_1 \rightarrow 2n_d + 2\gamma_d. \tag{4.1}$$

The dark photons in turn decay into four muons which originate from a displaced vertex. In this channel, 30% of the muons cross the region where the GEM detectors will be installed. Figure 4.4 illustrates the limitations of the track trigger (red) to reconstruct muons as a function of the distance between the vertex and the interaction point. Efficiency quickly drops to reach 0% around 50 cm. The muon standalone trigger (blue) on the other hand is capable of reconstructing the track with high efficiency underlining the necessity of preserving excellent trigger performances in the muon spectrometer.

4.2 Technology Overview of Triple-GEM detectors

A GEM foil is a 50-μm-thick polymer foil covered with 5-μm-thin copper sheets on both sides, and chemically perforated by a high density of microscopic holes. The polymer used is either Kapton or Apical, both with a dielectric constant of 3.5. The holes are truncated double cones with an outer diameter of 70 μm and a inner diameter of 50 μm and are spaced by 140 μm on an hexagonal grid. The structure of the GEM foil as shown on the left in Figure 4.5, is obtained using photolitography techniques that require precise alignment of the top and bottom masks.

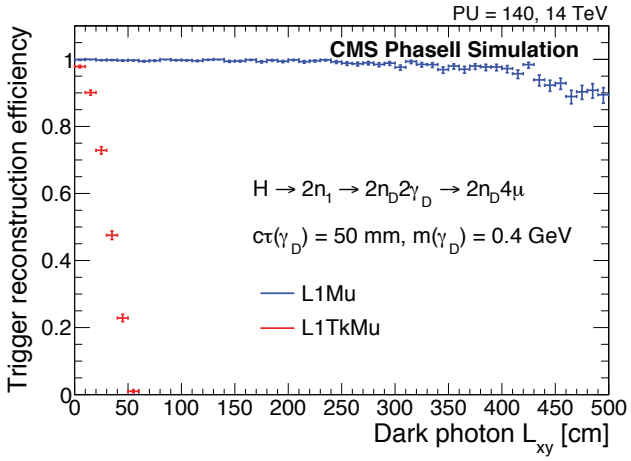

Figure 4.4 – The probability of reconstructing at least one muon candidate produced in the decay of a light long-lived particle decaying to a pair of muons $\gamma_d \to \mu\mu$ as a function of L_{xy}, the distance between the γ_d decay vertex to the beamline in the transverse plane. Standalone muon trigger L1Mu performance is compared to that of L1TrkMu, a trigger based on matching muon and track trigger candidates with the CMS Phase-II detector simulation [24].

The diagram on the right in Figure 4.5 represents the electric field lines that appear when a high voltage difference is applied between the two layers of copper, typically on the order of 300 V, resulting in field densities inside the holes reaching approximately 80 kV cm^{-1}. The structure of the electric field is used to amplify the signal of particles passing through the detector. Electrons resulting from the ionization of the gas by charged particles are directed towards the holes of the foil. When reaching high kinetic energy inside the holes, they themselves ionize the medium, producing secondary avalanches of electrons. Each hole acts as a proportional counter with gains up to 800 at voltages between 500 and 550 V.

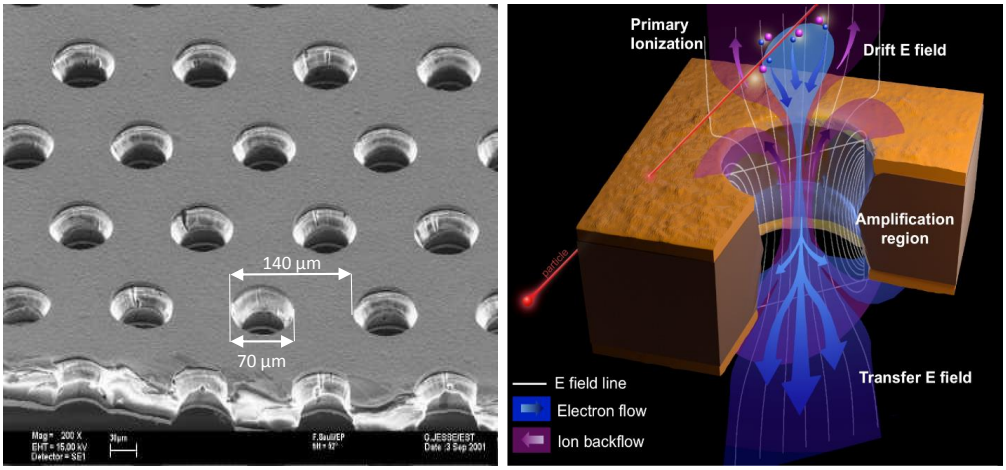

Figure 4.5 – Left: scanning electron microscope picture of a GEM foil. Right: schematic view of the electric field lines (white), electron flow (blue), and ion flow (purple) through a bi-conical GEM hole [24].

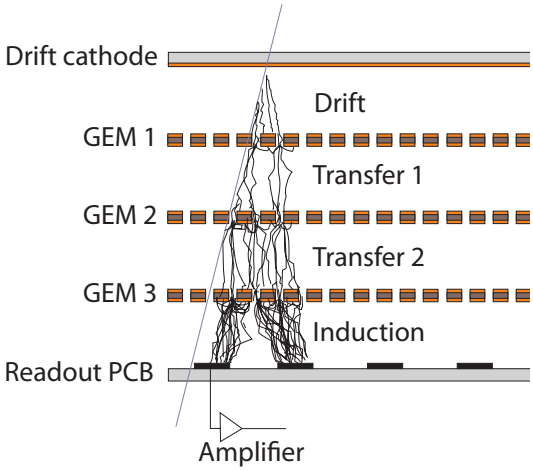

Figure 4.6 – Principle of operation of a generic triple-GEM chamber and definition of drift, transfer, and signal induction gap regions within the detector [24].

To achieve high gains in the detectors, two solutions can be implemented: increase the high voltage on the foils, or use multiple foils. The first option gives rise to discharges when operating at too high electric fields thus damaging the detector. Therefore, the choice to use three GEM foils has been made, hence the name of Triple-GEM detectors. The foils are arranged as shown in Figure 4.6 which depicts the principle of operation of a triple-GEM chamber. A drift cathode is placed on the top of the chamber which creates an electric field between itself and the top copper layer of the first GEM foil in the so-called drift gap measuring 3 mm. Electrons originating from the ionization of the gas mixture (either 70% Argon + 30% CO_2 or 45% Argon + 15% CO_2 + 40% CF_4) by a charged particle drift towards the holes of the foil, while ions are collected by the cathode. Following, the electron shower is amplified by GEM 1, GEM 2, and GEM 3, respectively transferring the signal to the transfer 1, transfer 2, and induction gaps measuring 1 mm, 2 mm, and 1 mm respectively. The size of each gap is of importance in the detector performance and has been optimized over time to increase the speed of the signal and the efficiency. The current configuration is called 3/1/2/1 mm in reference to the gap size. In the induction gap, the drift of the electrons towards the anodes induces a signal on the latter which can be detected by electronics. The choice to use either Ar/CO_2 or $Ar/CO_2/CF_4$ has yet to be made. Although the use of CF_4 improves given parameters of the detector exposed bellow, CMS has limited its use due to its toxicity for the environment.

4.3 Technical Design of GE1/1 chambers for CMS

The GE1/1 chambers are arranged in pairs to form superchambers, each covering 10^o in ϕ for a total of 72 chambers per endcap. Figure 4.7 displays the geometry of the superchambers. Due to mechanical constraints, the trapezoidal chambers are declined in two versions: a short one and a long one. The short chambers cover

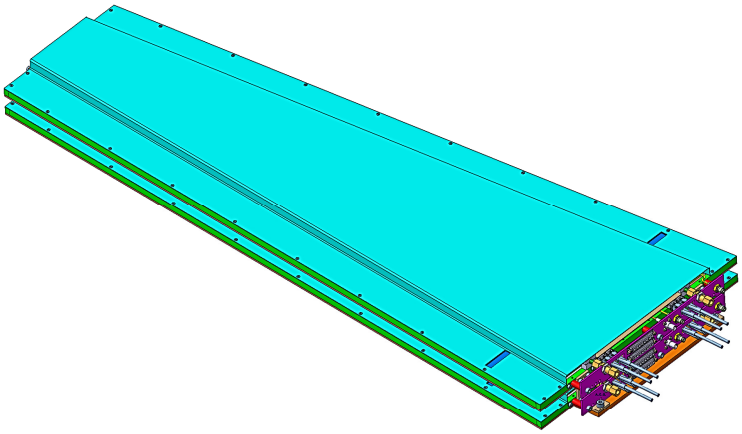

Figure 4.7 – A pair of GEM chambers form a superchamber [24].

$1.61 < |\eta| < 2.18$ and have a large base of 445 mm, a small base of 220 mm, and a height of 990 mm. The long chambers cover $1.55 < |\eta| < 2.18$ and have a large base of 510 mm, a small base of 279 mm, and a height of 1283 mm.

The structure of a single chamber is shown in Figure 4.8 which provides an exploded view of the detector. The drift cathode is placed on the bottom of the detector on which a frame is attached. The three GEM foils are stretched out using screws embedded in the frame along with the seals making the chamber air tight. The anode strips are engraved on the readout board that closes the active region of the detector. One chamber is divided into three ϕ sectors and eight η sectors with 128 strips running radially in each. This results in a mean coverage per strip of 450 μrad and a pitch

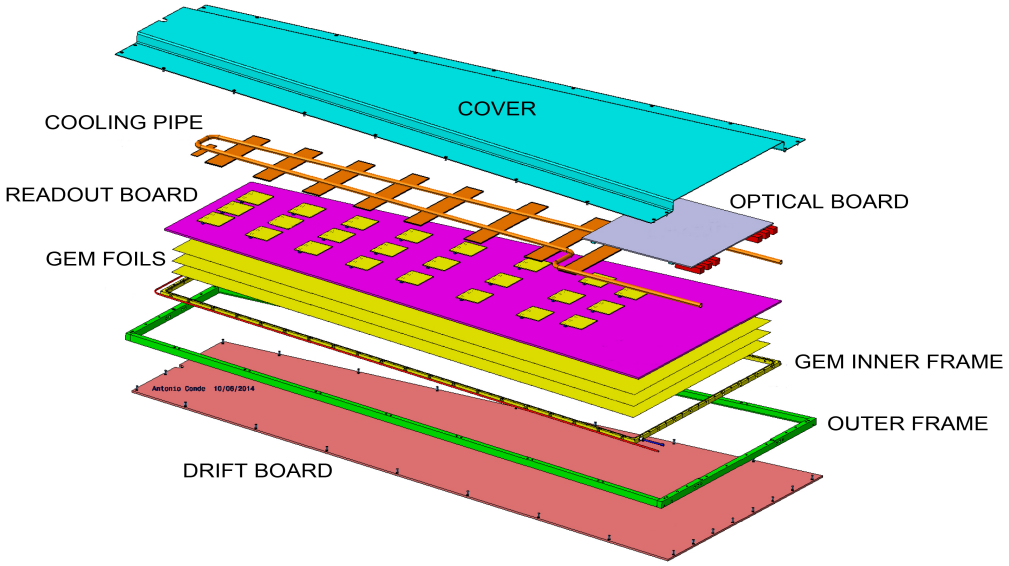

Figure 4.8 – Exploded view of the mechanical design of a Triple-GEM chamber [24].

varying from approximately 500 μm to 1.3 mm. To each group of 128-strips a front-end electronics chip called the VFAT3 is attached. This chip digitizes the signals from the strips returning a single hit/not-hit bit. The VFAT3s are all connected to the GEM electronics board which acts as router to forward the data towards the OptoHybrid, a concentrator board located on the large side of the GEM which is the interface to the off-detector electronics. The VFAT3s and the OptoHybrid are cooled down using cooling pipes attached on the top of the boards. Finally, a protective cover closes and protects the detector. When two chambers are assembled, a patch panel is added in order to facilitate the insertion and removal of the low and high voltage cables, the water for the cooling, the gas, and the optical fibers connected to the DAQ system. Note that a single source of high voltage is needed to power a chamber thanks to the use of a ceramic voltage divider which spreads the voltage over the various layers.

Once assembled, 36 superchambers will be installed in each endcap of CMS as displayed in Figure 4.9 showing the first station of the endcap of CMS. The full ring will be placed in the so-called nose of the endcap, between the HCAL and the ME1/1 CSC stations.

4.4 Detector Design Evolution

The design and fabrication method of the GE1/1 chambers has evolved over the past five years of R&D. Figure 4.10 shows the five generations of detectors that have been developed. GE1/1-I was the first 1-m-long prototype. It was equipped with only eight readout sectors, components were glued together, and GEM foils were separated using spacers. The number of sectors was increased to 24 in GE1/1-II with the full granularity of the current design (384 strips per 10^o). The gap configuration was changed from 3/2/2/2 mm to 3/1/2/1 mm to increase the signal speed. From GE1/1-III and onwards, GEM foils were stretched mechanically using an outer frame glued on the drift board. This design also introduced the use of a miniature ceramic high-voltage divider to apply voltage on the foils. However, the assembly of the chamber induced mechanical tension on the readout board, deforming the detector and resulting in a non-uniform response. This problem was resolved in GE1/1-IV by pre-bending the boards in the opposite direction in order to compensate before being bolted to the frames. This generation was the first assembled without glue, reducing production time to a few hours. However, the pre-beding technique did not yield viable results. Therefore, GE1/1-V uses pull-out pieces to add tension to the foils from the outer frame, on which the drift and readout board are bolted. The outer frame serves as a wall for the chamber and seals it using O-rings. Finally, the version of the design of the chambers that will be installed in CMS is GE1/1-VI which has optimized dimension to maximize the geometrical acceptance.

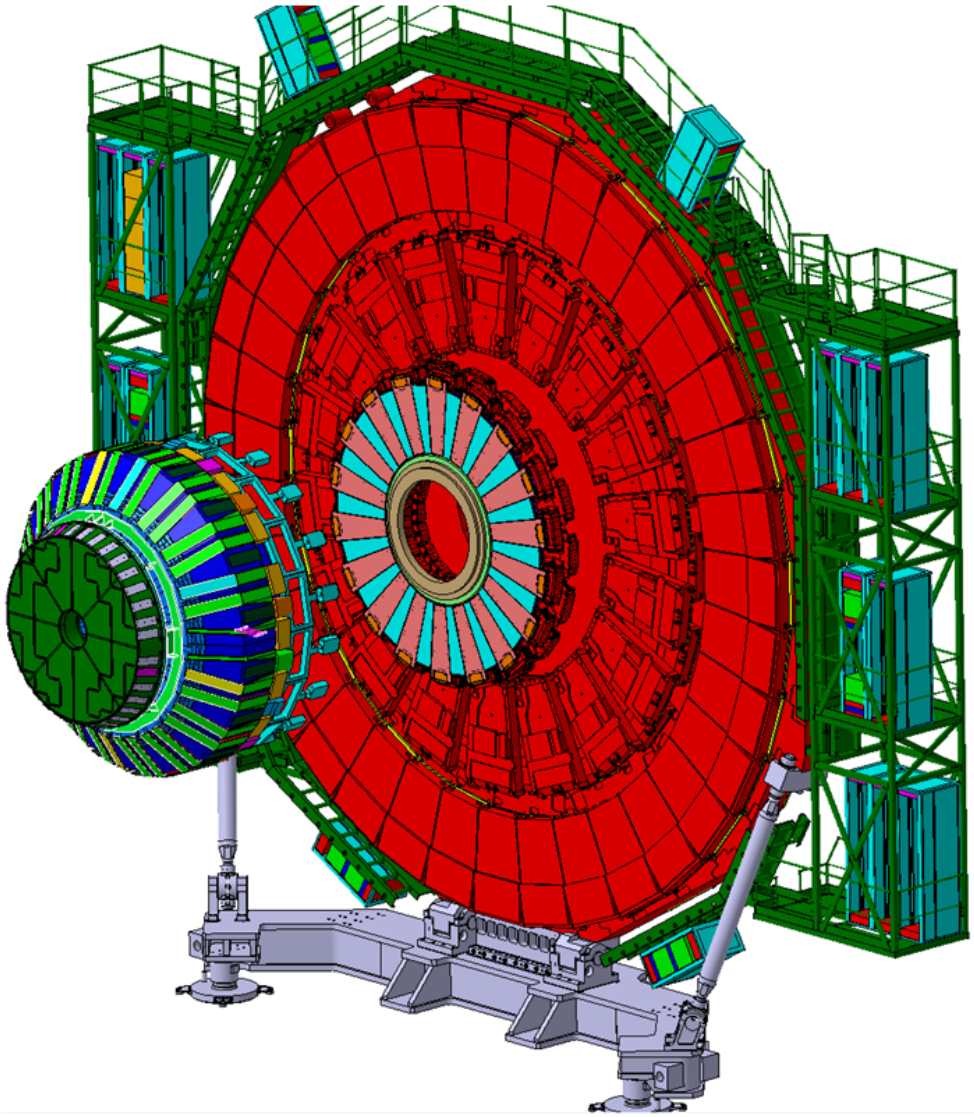

Figure 4.9 – First CMS muon endcap station where the inner ring is equipped with 18 long (red) and 18 short (blue) triple GEM superchambers [24].

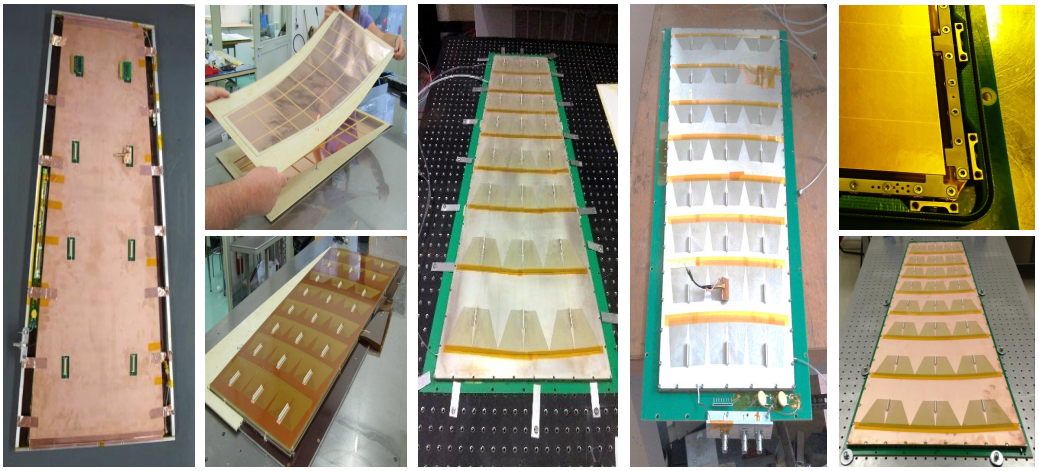

Figure 4.10 – Five generations of GE1/1 prototype chambers constructed and tested by the GEM collaboration in 2010-2014. The split figures for GE1/1-II and GE1/1-V demonstrate the evolution from construction using spacer frames to purely mechanical stretching of GEM foils without any spacers [24].

4.5 GE1/1 Prototyping Results

In order to provide the desired trigger and physics performance, the GE1/1 chambers must meet the following list of requirements:

- Maximum geometric acceptance within the given CMS envelope;
- Rate capability of 10 kHz cm^{-2} or better to cope with the HL-LHC hit rate;
- Single-chamber efficiency of 97% or better for detecting minimum ionizing particles resulting in 99.9% efficiency for a superchamber when signals are combined with a logical OR;
- Timing resolution of 10 ns or better for a single chamber, improving when combining the superchamber measurements, to allow matching with the CSC hits;
- Angular resolution of 300 μrad or better on $\phi_{GE1/1}$ to discriminate high-p$_T$ from low-p$_T$ muons when combining the measurements with ME1/1;
- Gain uniformity of 15% or better across a chamber and between chambers to ensure that no geometrical or trigger biases are present.

4.5.1 Rate Capability

Rate capability has been measured by monitoring the gain of a chamber when subject to a 22 keV Ag X-ray source and a high-intensity 8 keV Cu source. Figure 4.11 depicts

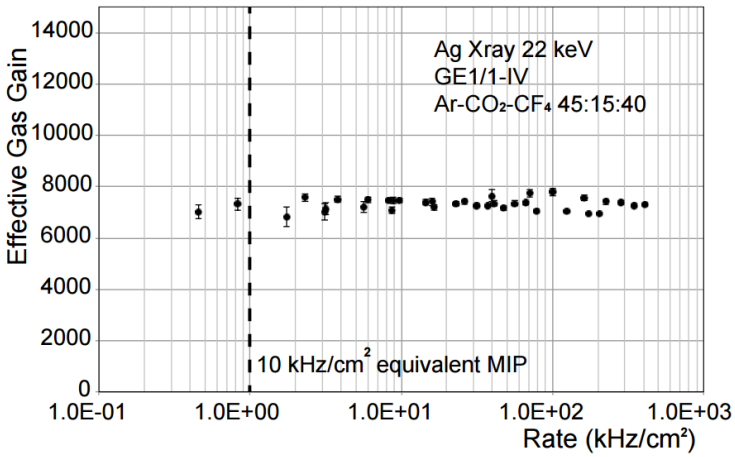

Figure 4.11 – Effective gas gain as a function of the rate of incident particles in a GE1/1-IV detector irradiated with a 22 keV X-ray source [24].

the effective gas gain as a function of the rate of incident particles in a GE1/1-IV detector irradiated with a 22 keV X-ray source. With $Ar/CO_2/CF_4$, gain uniformity is maintained up to 100 MHz cm^{-2}. This confirms that the GE1/1 chambers are able to sustain rates up to 10 000 times higher than what they will experience in the $1.6 < |\eta| < 2.2$ region of CMS.

4.5.2 Detection Efficiency

The detection efficiency is measured using a multi-layer tracking system placed in front of the detector in a beam of particles. Tracks are reconstructed in the former and extrapolated to the latter to match with reconstructed hits. When using the analog readout electronics, an offline cut is applied on the charge of the strips to cancel noise. In the digital readout system, an online threshold is applied. The image on the left in Figure 4.12 plots the evolution of the efficiency as a function of the high voltage applied to the drift electrode for a GE1/1-III chamber operated with Ar/CO_2 at 70% and 30% respectively and for different offline cuts on the collected charge. Both results indicate an efficiency plateau at 97%.

4.5.3 Time Resolution

The timing measurements are done by fitting the spread of the distribution plotting the time difference between a trigger coming from scintillators and the detection of the particle by the GEM detector. The plot on the right in Figure **??** shows the time resolution of a small Triple-GEM detector as a function of the electric field in the drift gap. Time resolution studies have been performed using $Ar/CO_2/CF_4$ as well as Ar/CO_2 gas mixtures. At the highest voltages, the former yields a resolution of 4 ns due to the fast electron transport speed in CF_4 while the latter provides 8 ns [28].

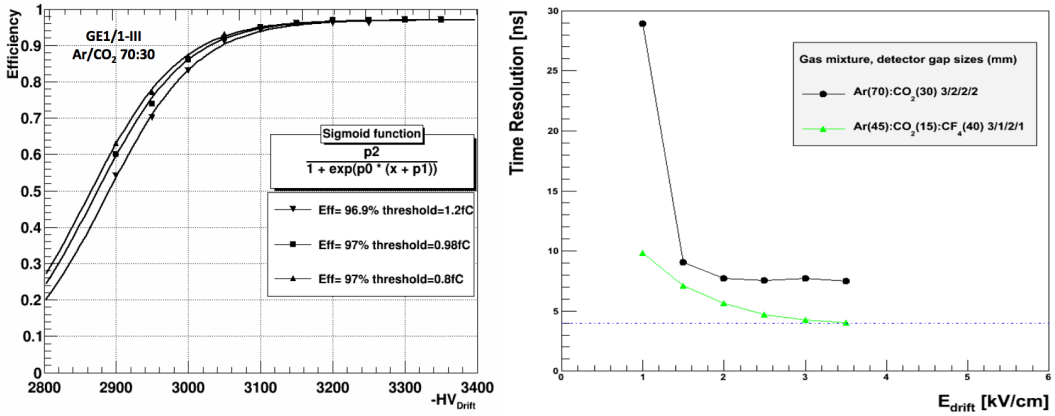

Figure 4.12 – Left: evolution of the efficiency as a function of the high voltage applied to the drift electrode for a GE1/1-III chamber operated with Ar/CO_2 at 70% and 30% respectively and for different offline cuts on the collected charge [24]. Right: time resolution of a small Triple-GEM detector as a function of the electric field in the drift gap [28].

4.5.4 Spatial Resolution

Using the above mentioned tracking system, the angular resolution of the GE1/1 chambers has also been measured during test beams. The resolution is taken to be the difference between the projected track position and the measured position in the detector. A resolution of 137 ± 1 μrad has been obtained setting an upper limit on the angular resolution which is close to the theoretical value computed for a binary readout of

$$\frac{\text{angular strip pitch}}{\sqrt{12}} = \frac{455\ \mu\text{rad}}{\sqrt{12}} = 131\ \mu\text{rad}. \tag{4.2}$$

Similar results were obtained using analog readout systems.

4.5.5 Response Uniformity

The response uniformity of the GEMs is measured for each detector as part of the quality control process. An X-ray generator is used to scan the entire chamber and the peak amplitude of the pulse charge distribution is taken as measurement of the uniformity. Figure 4.13 represents the distribution of the pulse charge measured over different strips on the left, and the evolution of the mean of the distribution of the pulse charge over multiple strips on the right. These tests show that the uniformity of the detector does not vary more than 15%.

4.6 GEM Upgrade Schedule

As previously stated, the objective of the GEM collaboration is to install a full ring of GE1/1 detectors in the first muon station of the endcaps during LS2 to maintain good trigger performances. To this end, the collaboration launched a large R&D program to study the GEM technology and to develop a new DAQ system for the detectors later described in details in Chapter 5. These developments have been and will be tested during test beam campaigns, as well as in CMS to characterize the

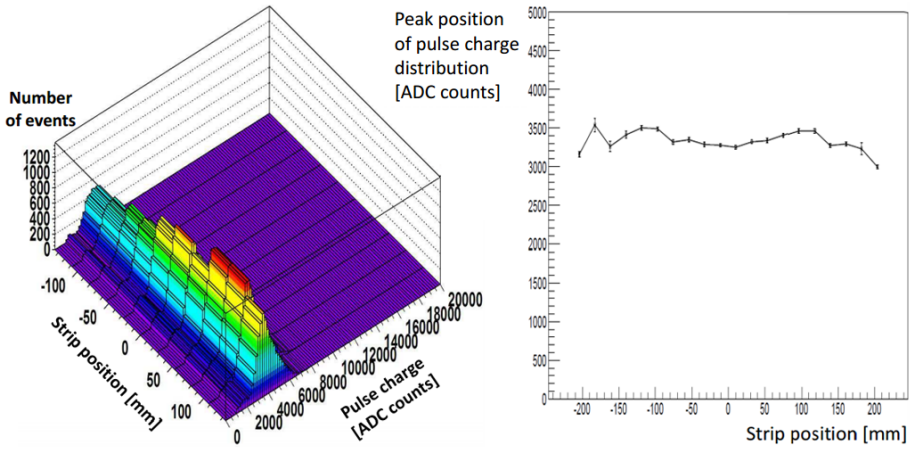

Figure 4.13 – Left: distribution of the pulse charge measured over different strips. Right: evolution of the mean of the distribution of the pulse charge over multiple strips [24].

chambers and prepare the installation in CMS.

Various test beam campaigns have been organized at CERN using the SPS proton beam to produce muons and pions to measure the performance of the detectors and test the developments done on the new DAQ system. The results of the latest test beam are discussed in Chapter 6. Besides test beams, mechanical tests have also been operated to prepare the installation of the chambers in the muon spectrometer. At the beginning of LS1, three replicas of superchambers have been installed in CMS to measure the available space and plan for cable space.

At the end of 2016, during the YETS, five superchambers equipped with a full DAQ system will be installed in CMS and connected to the central DAQ of CMS. This small scale test is called the slice test and will demonstrate the integration of the GEMs with the CSCs before the full installation during LS2. As developments on the DAQ systems are still ongoing, the detectors will be equipped with slightly different components in the readout chain.

Finally, before installation in CMS, each component needs to be qualified and characterized, as well as tested against the effects of radiation. To this end, automated procedures have been developed and are described in Chapters 7 and 8.

4.7 Developments Beyond GE1/1

Next to GE1/1, two other projects are under study by the CMS GEM collaboration: GE2/1 which would instrument the second inner muon station with GEMs, and ME0 which would extend the coverage of the muon spectrometer up to $\eta = 3.5$. Figure 4.14 shows a longitudinal view of a quadrant of CMS highlighting the location of the GE1/1, GE2/1, and ME0 detectors colored in red within the muon spectrometer.

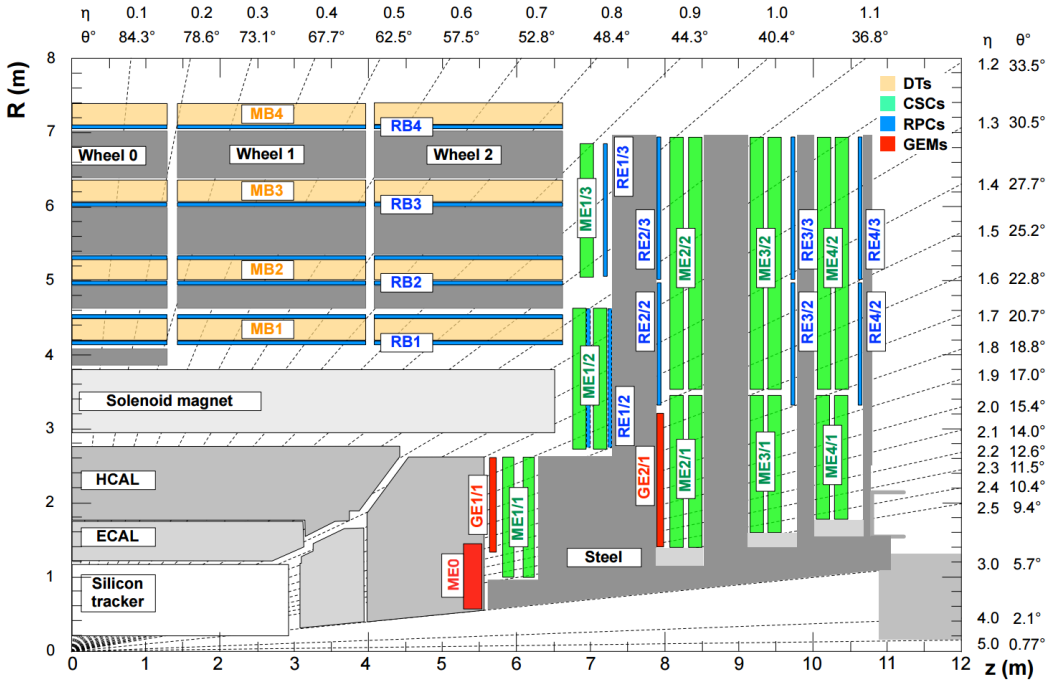

Figure 4.14 – Longitudinal view of a quadrant of CMS highlighting the location of the GE1/1, GE2/1, and ME0 detectors colored in red within the muon spectrometer. DTs are represented in yellow, CSCs in green, and RPCs in blue [28].

These projects are still under development and proposed to be installed during LS3, before the start of the LHC Phase II.

The GE2/1 upgrade proposes to install 36 superchambers of Triple-GEM detectors next to the CSCs ME2/1 stations in the $1.60 < |\eta| < 2.46$ region. Each detector would cover 20^o in ϕ and measure 1.84 m in length, 1.18 m in width, and 81 mm in height. Due to the large size of the detectors, they would be segmented in eight η and six ϕ sectors.

The ME0 project aims at installing a 6-layered Triple-GEM detector behind the future shortened HCAL, which upgrade will take place during LS3. This would extend the coverage of the muon spectrometer in the $2.0 < |\eta| < 3.5$ region and provide efficient muon identification with low background. The detectors will be subject to high rates of particles, of the order of the MHz cm^{-2}, and be exposed to high doses of radiation. The latter will impact the design of the readout electronics and the components that will be selected.

5 The GEM Data Acquisition System

The installation of GEM detectors in CMS and the integration with the CSCs requires the development of a new DAQ system for the GE1/1 project. The understanding of the structure of the GEM readout chain as well as the prototyping version that have been developed is of importance in the scope of this thesis. To this end, these systems are presented in details in the following sections.

5.1 The GE1/1 Data Acquisition System

The architecture of the GE1/1 DAQ system is represented in Figure 5.1. It is divided into two sectors: the on-detector electronics on the left in charge of the management of the detector, and the off-detector electronics on the right responsible for the data handling and the connection to the central DAQ. The two sectors are separated by a few dozen meters and connected through optical fibers.

The on-detector electronics focuses on the control and readout of the VFAT3 Application Specific Integrated Circuit (ASIC) which connects to the strips of the chamber and digitizes the signals. The GEM Electronic Board (GEB), on which the VFAT3s are plugged in, then routes the data to the OptoHybrid which acts as concentrator board and communication relay for the 24 VFAT3s. The communication with the off-detector system is performed through the GigaBit Transceiver (GBT) chipset and

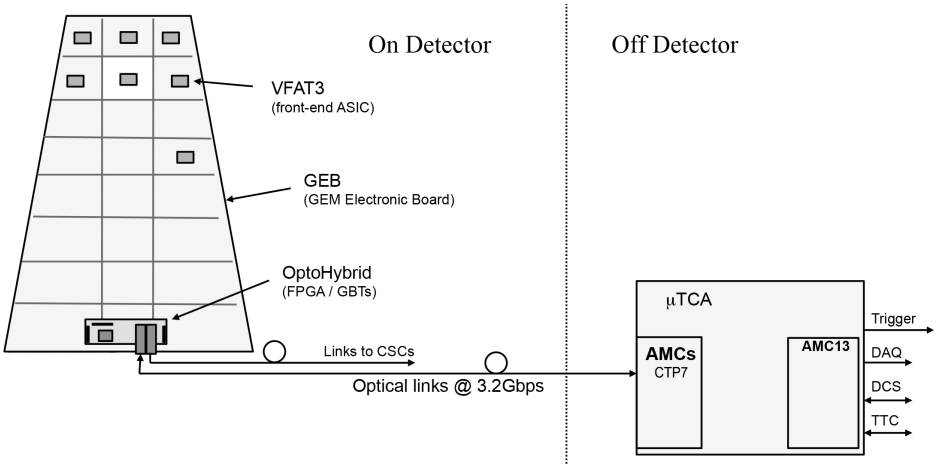

Figure 5.1 – Architecture of the GE1/1 DAQ system divided into two sectors: the on-detector electronics composed of the VFAT3 ASICs, GEB, and OptoHybrid, and the off-detector electronics with CTP7 and AMC13 mezzanines located in a microTCA crate [24].

the Versatile Link installed on the OptoHybrid. Both projects are led by CERN and provide radiation hard tools for LHC experiments.

On the off-detector side, the Micro Telecommunications Computing Architecture (microTCA, MTCA, or μTCA) [29] crate standard is used to power and monitor the Advanced Mezzanine Cards (AMCs) which provide the resources to communicate with the on-detector electronics. Links from 12 OptoHybrids are concentrated on one CTP7 AMC which formats the data and transfers it to the CMS AMC13 mezzanine. The AMC13 is the link between the microTCA crate and the central DAQ of CMS providing the clocking, trigger, and control over the system. The control of the DAQ chain is performed through software using the IPBus protocol over Ethernet.

5.1.1 The VFAT3 ASIC

The VFAT3 ASIC is a binary front-end chip optimized for gaseous detectors which function is to digitize the analog signals coming from the detector and provide fast trigger and tracking data. The trigger data is sent at the LHC clock frequency over a fixed latency path and then used in the algorithms of the L1 trigger to accept or reject events. The tracking data holds the full granularity information of the events that have been accepted and follows a variable latency path. The logic diagram of the chip is shown in Figure 5.2. It is made of an analog front-end which amplifies, shapes, and digitizes the signal from the GEM strips, and of a digital back-end for slow control, fast control, and data readout.

The analog front-end is further optimized for the readout of GEMs in particular. It is composed of 128 channels which amplify and shape the analog signals from the strips with programmable shaping times to allow for various integration lengths of the signal. According to the gas mixture, signal charge from the GEM can last for approximately 60 ns. Increasing the shaping time to fully integrate the charge will result in a higher signal to noise ratio. Simulations [28] performed on the

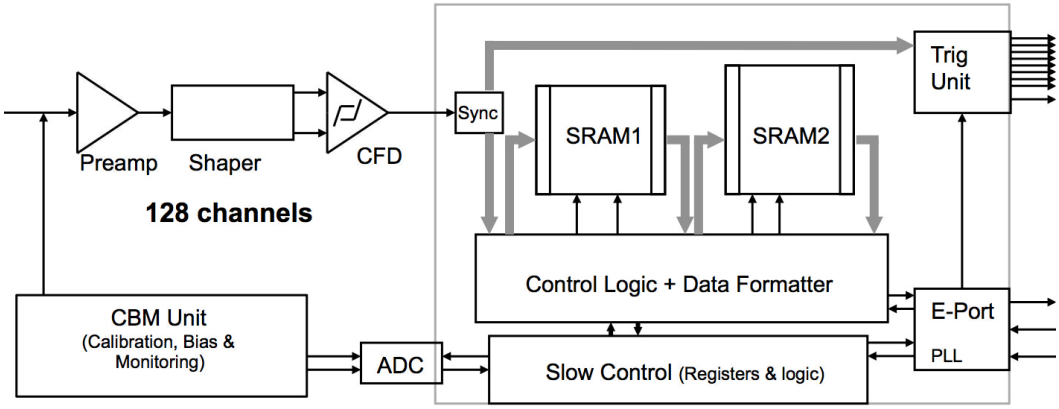

Figure 5.2 – Schematic diagram of the VFAT3 showing the analog front-end on the left and the digital control and readout system on the right [24].

analog front-end show that a time resolution of 7 ns can be achieved by using a Constant Fraction Discriminator (CFD) with a shaping time of 50 ns. After shaping, the amplitude of the analog signal is compared against a programmable threshold by a comparator to yield a binary output flag for each channel.

The fixed latency path is used to provide a fast hit information to the trigger system of CMS at a frequency equal to the LHC clock of 40 MHz. The trigger unit inside the VFAT3 formats the output of the 128 comparators and transmits it over 8 differential pairs at a rate of 64 bits per BX. It allows to encode a logical-OR of two adjacent strips effectively dividing the number of bits to send by a factor two. Ensuring a fixed latency on this path is crucial to maintain predictability in the trigger system and identify the correct BX.

The variable latency path is activated upon reception of an L1A to transmit the full granularity information on an event that has been accepted by the trigger system. The VFAT3 holds two Static Random-Access Memories (SRAMs) which are used to store events. The SRAM1 is a circular buffer filled at a frequency equal to the LHC clock with the output of the 128 comparators. When the VFAT3 receives an L1A, it transfers a given event from the SRAM1 to the SRAM2 and adds the BX Counter (BC) and the Event Counter (EC), which respectively counts the number of BX elapsed and the number of L1A received, to the event data. To determine which event needs to be transferred, the chip uses a programmable parameter called the latency that informs the system of the delay between the digitization of an event and the arrival of an L1A corresponding to the same event. It is a measurement of the response time of the trigger system to a given event, which is a fixed delay. Events stored in the SRAM2 are then formatted and sent over a single differential pair out of the VFAT3. The formatting of the data can be selected, using a programmable register, to be either lossless or zero suppressed.

The slow control module handles the configuration of the internal registers of the VFAT3. These registers define quantities such as the threshold of the analog front-end, the latency, the readout data format, etc. Coupled with the Calibration, Bias & Monitoring (CBM) unit, it allows to perform calibration routines on the chip.

The fast control defines all the time sensitive commands that the VFAT3 receives. These are the Event Counter 0 (EC0), Bunch Counter 0 (BC0), Calibration Pulse (CalPulse), Resynchronisation (Resync), and L1A. The EC0 and BC0 commands respectively reset the EC and BC; the CalPulse is used to send a calibration pulse on given strips in order to calibrate the analog front-end; the ReSync command resets the internal pointers; and the L1A informs the VFAT3 that an event needs to be transferred to the DAQ.

5.1.2 The GEM Electronic Board (GEB)

The limited space in which the GEM detectors will be installed constrains the design of the system by making it impossible to run flat cables for the 24 VFAT3s. Therefore, the GEB, a Printed Circuit Board (PCB) of the same dimension as the GEM detector, was designed to route the signals of the front-end chips to the OptoHybrid located on the large side of the detector, furthest away from the beam pipe. A picture of a prototype of the GEB is shown on the right in Figure 5.3. The functions of the GEBs are to host the VFAT3 ASICs connected to the 24 sectors of the GEM readout board, route their signals to the OptoHybrid, distribute power to the chips, and provide electric shielding to the detector. The GEB is divided into three columns of eight VFAT3s, further divided into two sectors each, as represented in the left picture in Figure 5.3. The clock, slow control, and fast control are common for each column.

The VFAT3s are mounted on small PCBs attached to the GEB and connected to the GEM readout board. The control and data signals from the chips are routed on the PCB towards four connectors located on the large side of the detector which connect to the OptoHybrid. The power for the VFAT3s and the OptoHybrid is distributed by the GEB using radiation hard DC/DC converters developed by CERN, which convert the incoming low voltage to the required voltages. In addition, the bottom layer of the GEB is grounded proving shielding to the detector from external ElectroMagnetic Interference (EMI).

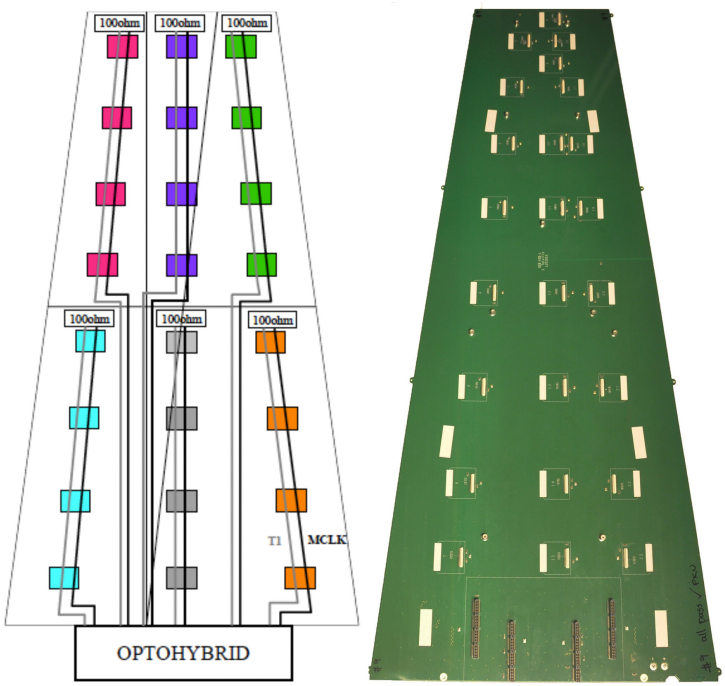

Figure 5.3 – Pictures of a prototype of the GEB, a PCB of the same dimension as the GEM detector designed to route the signals of the front-end chips to the OptoHybrid [24].

5.1.3 The OptoHybrid

The OptoHybrid is the interface between the VFAT3 ASICs and the off-detector system. It is a 10 cm $\times$ 20 cm board mounted on the GEB equipped with a Field Programmable Gate Array (FPGA) and Integrated Circuits (ICs) dedicated to the operation of the optical links. The FPGA solely handles the trigger data by applying zero suppression algorithms, formatting the data, and sending it to the CSC and the GEM trigger system separately over two optical links. The other functions of the VFAT3s are directly connected to the three GBT chipsets installed on the OptoHybrid. Each GBT chipset is capable of handling one column of eight VFAT3s.

5.1.4 The GBT and Versatile Link

The GBT [30] and Versatile Link [31] are projects developed by CERN providing radiation hard tools for optical communication to the LHC experiments. The GBT is an optical data link technology providing bidirectional 4.8 Gbps (Gigabits per second) serial communication designed to connect the on-detector and off-detector electronics. It is produced in two versions: the GBTX which is a radiation hard ASIC, and the GBT-FPGA which is a core that can be implemented inside an FPGA. Complementary to the GBT which provides the data structure and communication protocol between components, the Versatile Link project develops radiation hard optical transceivers which connect to the GBT and generates the signals on the optical fibers. The integration of the GBTX, Versatile Link, and GBT-FPGA projects is shown in Figure 5.4.

The GBT defines a data format shown in Figure 5.5. The stream of data is divided into frames of 120 bits transmitted at 40 MHz, resulting in a data rate of 4.8 Gbps from which 3.36 Gbps are accessible by the user. The frame starts with a four bits header (H) which defines the frame type and ends with 32 bits of Forward Error Correction (FEC). The error correction is done by encoding the data using a Reed-Solomon code that can correct bit flips due to radiation. The 84 remaining bits are user accessible. The first two bits are used for internal slow control (IC) and the two following bits for external slow control (EC), leaving 80 bits of data (D)

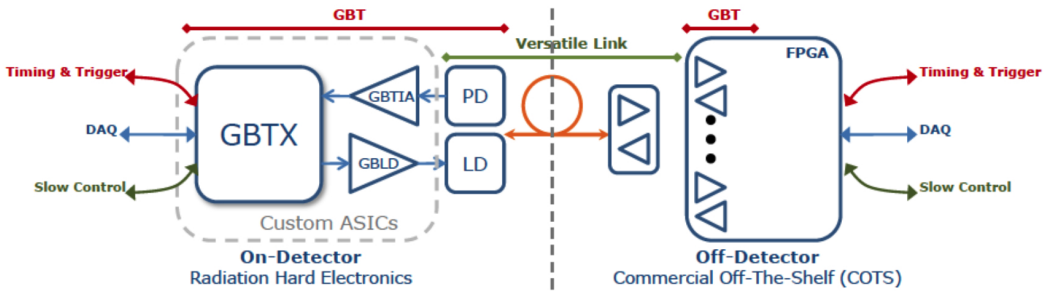

Figure 5.4 – Architecture of the radiation hard optical link setup integrating the GBT and Versatile Link projects [30].

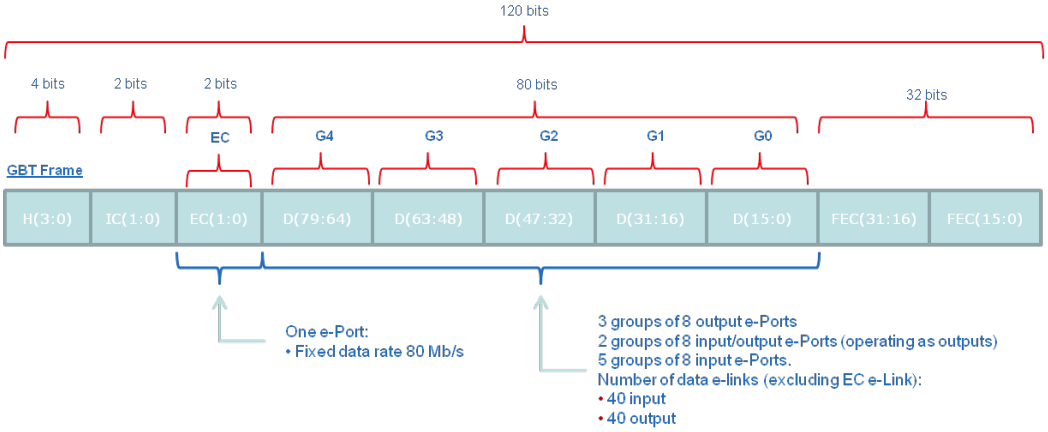

Figure 5.5 – Data format of a GBT frame [30].

accessible to the user for DAQ, TTC, and slow control purposes.

The header is used to synchronize the link by requiring the detection of a series of valid header fields. It is therefore protected by the FEC in order to ensure efficient frame-locking. From the slow control bits, both running at 80 Mb/s, the IC is reserved for the control and monitoring of the GBTX while the EC on the other hand can be freely assigned by the user. The data field is used to transfer DAQ, TTC, and slow control requests to the various subsystems connected to the GBT and can implement any user defined protocol. The FEC uses two Reed-Salomon codes interleaved with 4-bit symbols each capable of correcting double errors. This means that the total frame can recover up to 16 corrupted bits. Additionally, the GBT frame is scrambled in order to maintain DC balance on the transmission lines.

On the FPGA side, the communication protocol is seen by the user as a 84-bit-wide vector that is sampled at a frequency of 40 MHz and then formatted by the GBT-FPGA core. The GBTX ASIC however uses 42 Electric serial Links (E-Links) to handle the data. The structure of the GBTX is shown in Figure 5.6. Each E-Link is composed of three differential pairs: one to transmit data, one to receive data, and one to carry the clock synchronous to the data. From the 42 E-links, one is dedicated to the IC and one to the EC, each transferring data at 80 Mb/s by serializing the two IC or EC bits of the GBT frame on a single link. The use of the remaining 40 E-Links is defined according to the speed at which they operate. The 80-bit-long user data can be configured to either be serialized by a factor 2, 4, or 8 resulting in the use of 40, 20, or 10 E-Links respectively running at 80 MHz, 160 MHz, or 320 MHz to accommodate the data rate. For the GEM project, a serialization factor of 8 is used, meaning that from the 40 E-Links, only 10 are in use, each of them transferring 8 bits per BX.

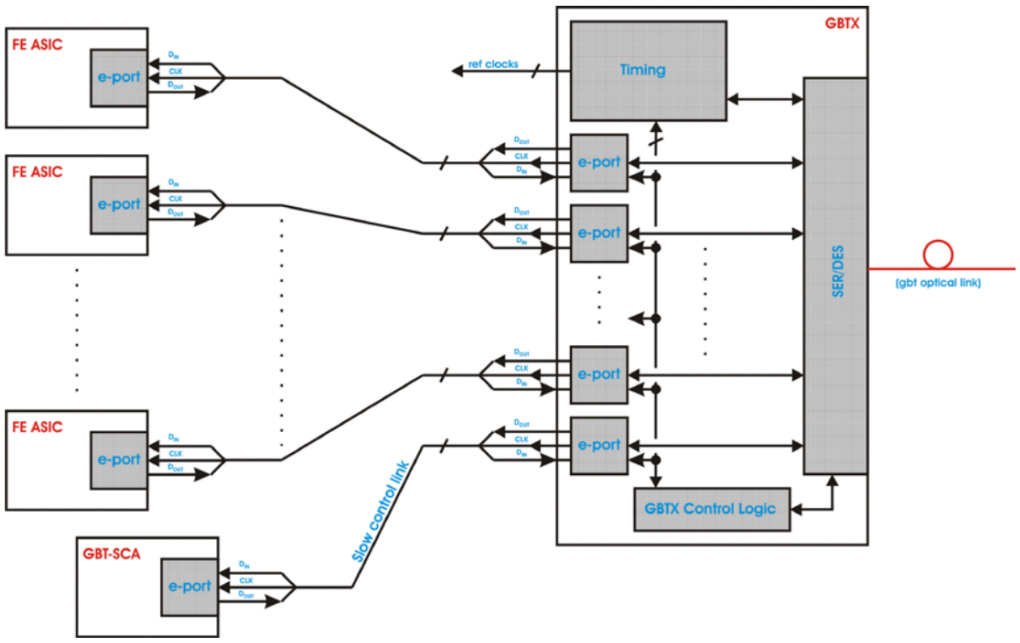

Figure 5.6 – Interaction between the GBTX ASIC and external front-end drivers through E-links [30].

To convert the GBTX signals into optical communication, the Versatile Link project provides two laser diode modules: a transceiver module called VTRx and a dual transmitter module called VTTx. The modules have been designed to withstand high magnetic field and radiation doses as encountered in the LHC experiments. The maximum bit rate is of 4.8 Gbps with a bit error rate inferior to 10^{-12}. Figure 5.7 shows a picture of the VTTx (left) and VTRx (right) optical transceivers of the Versatile Link project.

5.1.5 The microTCA Standard

The off-detector electronics is implemented using the microTCA crate standard to manage the back-end system. The microTCA standard was developed for the telecommunication industry as a successor for the Advanced Telecommunications Computing Architecture (ATCA) technology. It uses a smaller form factor for the AMCs declined into a single width (73.8 mm $\times$ 181.5 mm) or double width (148.8 mm $\times$ 181.5 mm) version. The key feature of the crate is the topology of the

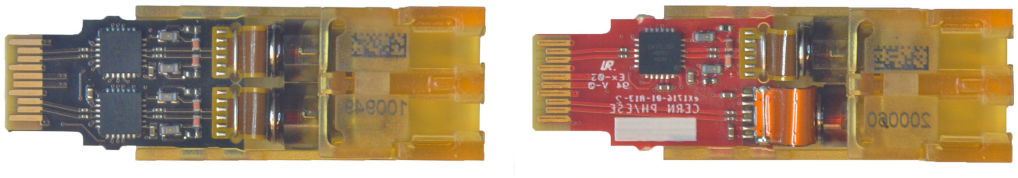

Figure 5.7 – Picture of the VTTx (left) and VTRx (right) optical transceivers of the Versatile Link project.

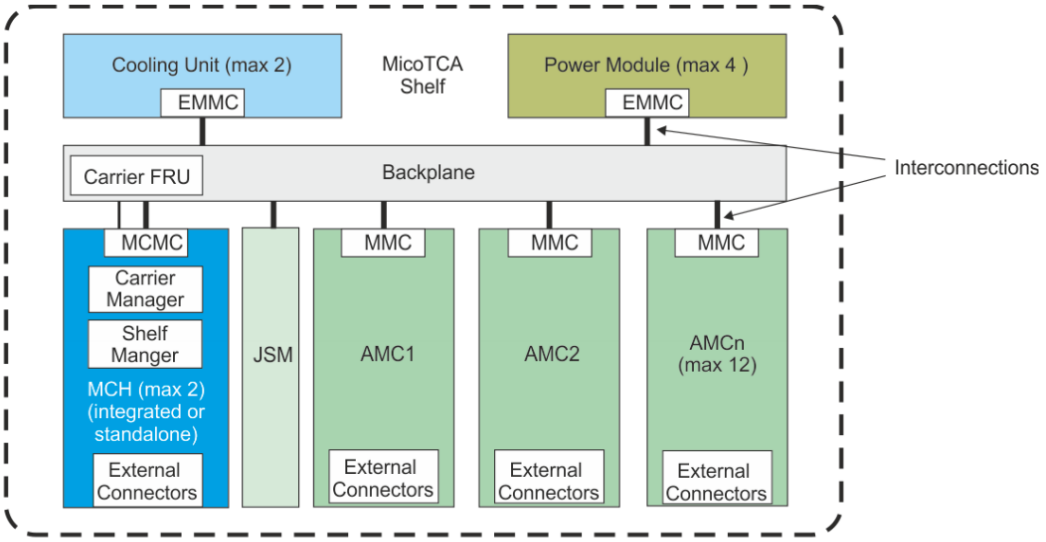

Figure 5.8 – Schematic representation of the architecture of a microTCA crate [32].

backplane connecting the AMCs. The latter is not defined in the specifications, but often implements a dual-star network which provides redundancy and thus avoids critical failures of the system upon malfunction of a component.

The structure of the crate is shown in Figure 5.8. Management and control of the subsystems of the crate are performed by the MicroTCA Carrier Hub (MCH), which in failure critical systems can be accompanied by a second MCH for redundancy. These two MCHs are the central points of the network in dual-star topology. They communicate with the Cooling Units (CUs), Power Modules (PMs), and AMCs through the Intelligent Platform Management Interface (IPMI). The CUs and PMs are equipped with an Enhanced Module Management Controller (EMMC) and the AMCs with a Module Management Controller (MMC) which are used to communicate with the MicroTCA Carrier Management Controller (MCMC) of the MHCs. These controllers implement the IPMI protocol and provide system management functions to the crate. Upon power-on of the crate or hot-swap of an AMC, the MCHs provide minimal power to the EMMCs and MMCs through dedicated lines in order to start the initial transaction sequence. During this exchange, the power requirements of the modules are defined and a boot-up sequence takes place. Once the transaction is closed, power is provided to the rest of the module and normal operation begins. A power-off sequence takes place when the crate is shutdown or an AMC is removed.

The backbone of the microTCA crate is its backplane. It provides point to point connections between the AMCs and with the MCH. The interconnects between elements are called fabrics and are what defines the topology of the crate. On the AMCs, the connections to the backplane are called ports. In the dual-star topology, each fabric is duplicated and connected to a separate MCH. As the backplane

topology is not described in the specifications of the standard, a common topology is described.

- Fabric A is used for Gigabit Ethernet communication between the AMCs and MCH1 on port 0 and the MCH2 on port 1.
- Fabric B is allocated for technologies such as Seriel-ATA through port 2 and 3 for the MCH1 and MCH2 respectively.
- Fabric C is only used in single MCH mode and uses port 3 to connect the AMCs to MCH1.
- Fabric D to G are fat pipes which use multiple protocol agnostic links to connect the AMCs to the MCH1 on ports 4 to 7 and to the MCH2 on ports 8 to 11.
- Some backplanes implement point to point communication between ACMs on ports 12 to 20 following a custom interconnect.

Next to signal buses, the microTCA standard defines three clocks buses (CLKx) that can originate from the MCHs or the AMCs.

5.1.6 The CTP7 Advanced Mezzanine Card

The AMC board used as back-end electronics is the CTP7 [33] developed by the University of Wisconsin and shown in Figure 5.9. It is equipped with a large Virtex-7 FPGA (XC7VX690T) providing extensive computational power along with a Zynq FPGA (XC7Z045) able to run a Linux operating system for monitoring. The optical

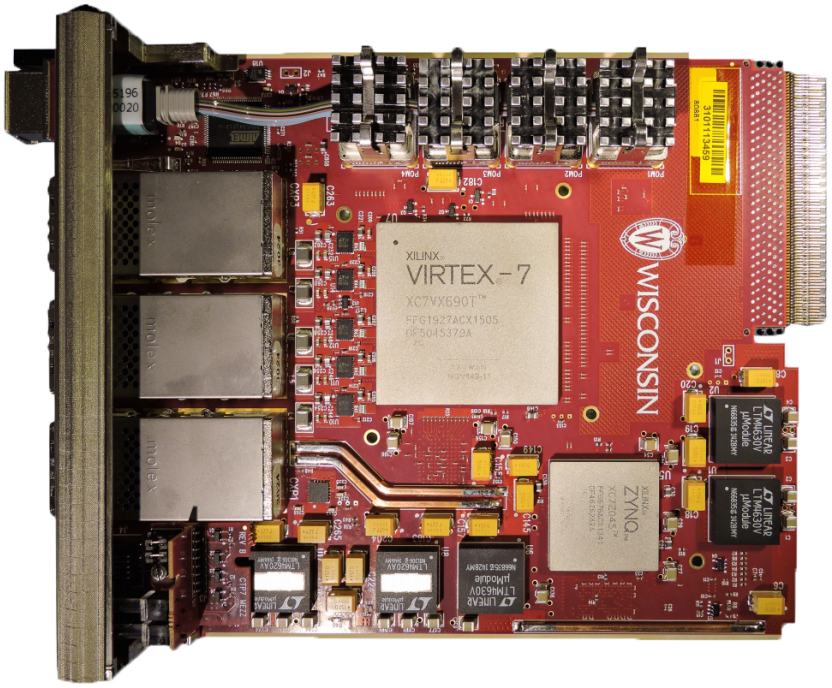

Figure 5.9 – Photograph of the CTP7 AMC developed by the University of Wisconsin used as back-end electronics [33].

interface with the on-detector electronics is done through 36 transmitters and 48 receivers. Using one CTP7, six superchambers can be readout thus requiring a total of six CTP7s per endcap for the GE1/1 project.

The function of the CTP7 is triple: handle slow control requests for the subsystems, interpret TTC signals, and readout the trigger and tracking data from the Opto-Hybrids. The CTP7 is the interface between the control and monitoring software and the detector electronics. It uses the fabric A lines of the MCH1 to connect to a Gigabit Ethernet network and handle TCP requests. It interprets the requests and forwards them to the appropriate subsystem: either itself or over the optical link to the OptoHybrids and VFAT3s. The CTP7 also receives the TTC commands on the fabric B from MCH2 and must forward them to the subsystems along with the LHC clock received on the CLK1 line. Finally, when the VFAT3s or the OptoHybrids push data upwards to the CTP7, the latter must format the data before sending it on the backplane to a dedicated DAQ AMC.

5.1.7 The AMC13 Advanced Mezzanine Card

The module which delivers the TTC signals and the LHC clock to the CTP7s and retrieves the trigger and tracking data from the backplane is the AMC13 [34]. It is a board specifically designed for CMS which replaces the MCH2 in the microTCA crate. Its functional diagram is shown in Figure 5.10. It is equipped with two FPGAs: a powerful Virtex-6 for data handling, and a smaller Spartan-6 for control and monitoring. The TTC signal is received through an optical link from which the

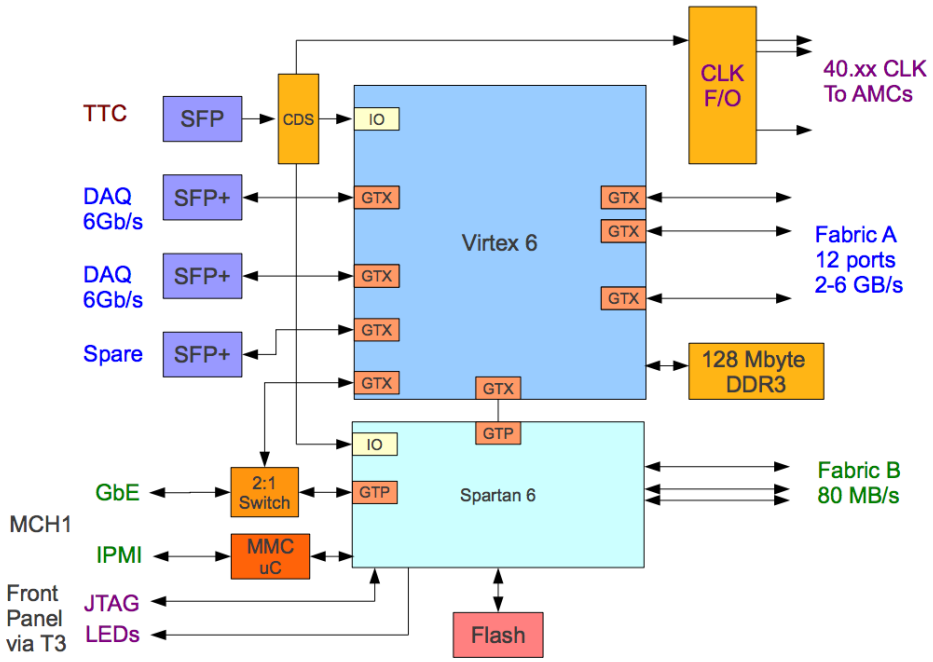

Figure 5.10 – Functional diagram of the AMC13 highlighting the external connections to the TTC system, backplane, and MCH of the microTCA crate [34].

LHC clock is extracted along with the TTC commands. The clock is forwarded on the CLK1 line and the TTC commands are passed to the Spartan-6 FPGA, which formats them on the fabric B lines at 80 MB/s. Being in the MCH2 slot, the AMC13 has direct access to all AMCs, which it uses to retrieve tracking and trigger data over the fabric A ports running at 5 Gbps. For development purposes, the AMC13 integrates internal TTC commands and LHC clock generators.

The Virtex-6 FPGA of the AMC13 is connected to FEROLs of the central DAQ of CMS (see Section 3.8) through the two SFP+ optical links running at 6 Gbps which implement the S-Link Express protocol defined by CMS. An additional SFP+ can be used as a spare if the bandwidth is required. Next to the data, the AMC13 also receives the TTC signals from TCDS and sends TTS signals to the TCDS system to inform it about the status of the readout. The TTS codes can be: hardware failure, imminent overflow, sync lost, busy, ready, or error. Using these, the TCDS system can regulate the L1A rate in order for the system to be able to cope.

5.1.8 The IPBus Protocol

To communicate between the control and monitoring applications, and the subsystems, a new protocol called IPBus has been developed for the CMS subdetectors. It uses the Internet Protocol (IP) and can be sent over User Data Protocol (UDP) or TCP. IPBus transactions are simple read/write operations using a A32/D32 implementation (Address of 32 bits, Data of 32 bits). A sequence of transactions is called a packet which starts with a packet header of 32 bits which format is shown in Figure 5.11.

A transaction is always performed in two stages: a request from the applications is sent to which the system provides a response. The header of the request specifies the request type and request size. Additional 32 bits words are added in function of the type of operation to perform. For a read operation, the base address of the read must be specified. The response will contain N 32 bits words corresponding to the data at the base address and the incremental N addresses where N is the size of the request defined in the header. For a write operation, the request will contain the base address followed by N 32 bits words that should be written at the base address and the N following addresses where N is the size of the request defined in the header. The response will be a single header containing a status code. Variation of the read/write operations exist such as the FIFO-read and FIFO-write, which repetitively read and write words at or to the same address.

On the electronics side, the IPBus protocol is handled by a master which decodes the requests and forwards them to slaves, each responding to a given address range. Requests involving multiple reads and writes are decomposed by the master in single operations. Slaves receive the request type, address, and data as input parameters

Packet Header

	31 – 24	23 – 16	15 – 8	7 – 0	
	Protocol version	Rsvd.	Packet ID (16 bits)	Byte-order qualifier	Packet Type
Word 0	0x2	0	0x0 – 0xffff	0xf	0x0 – 0x2

Read Request

	31 – 24	23 – 16	15 – 8	7 – 0	
Word 0	Version = 2	Transaction ID	Words = READ_SIZE	Type ID = 0	InfoCode = 0xf
Word 1	BASE_ADDRESS				

Read Response

	31 – 24	23 – 16	15 – 8	7 – 0	
Word 0	Version = 2	Transaction ID	Words = READ_SIZE	Type ID = 0	InfoCode = 0
Word 1	Data read from BASE_ADDRESS				
Word 2	Data read from BASE_ADDRESS + 1				
...	...				
Word n	Data read from BASE_ADDRESS + (READ_SIZE - 1)				

Write Request

	31 – 24	23 – 16	15 – 8	7 – 0	
Word 0	Version = 2	Transaction ID	Words = WRITE_SIZE	Type ID = 1	InfoCode = 0xf
Word 1	BASE_ADDRESS				
Word 2	Data for BASE_ADDRESS				
Word 3	Data For BASE_ADDRESS + 1				
...	...				
Word n	Data for BASE_ADDRESS + (WRITE_SIZE -1)				

Write Response

	31 – 24	23 – 16	15 – 8	7 – 0	
Word 0	Version = 2	Transaction ID	Words = WRITE_SIZE	Type ID = 1	InfoCode = 0

Figure 5.11 – Data format of the data packets of IPBus for read and write transaction types [34].

and can then treat it as they like. In return, they provide an acknowledgement and optionally data in the case of a read operation. Even though the software treats the address space of 32 bits as a set of read/write registers of 32 bits (A32/D32), the physical implementation on the electronics can be different. Requests can be connected to registers or be used to activate functions on the PCBs. The implementation is up to the developers.

5.2 The First Prototype of the Data Acquisition System

Due to the ongoing developments on the VFAT3 ASIC and the GBT chipset, and to test the feasibility of the manufacturing of the GEB, a first prototype of the GE1/1 DAQ system was developed in 2014 with slightly different components. The differences with the final system are shown in Figure 5.12 where elements in red are not yet implemented and those in blue are prototypes or variations of the final electronics.

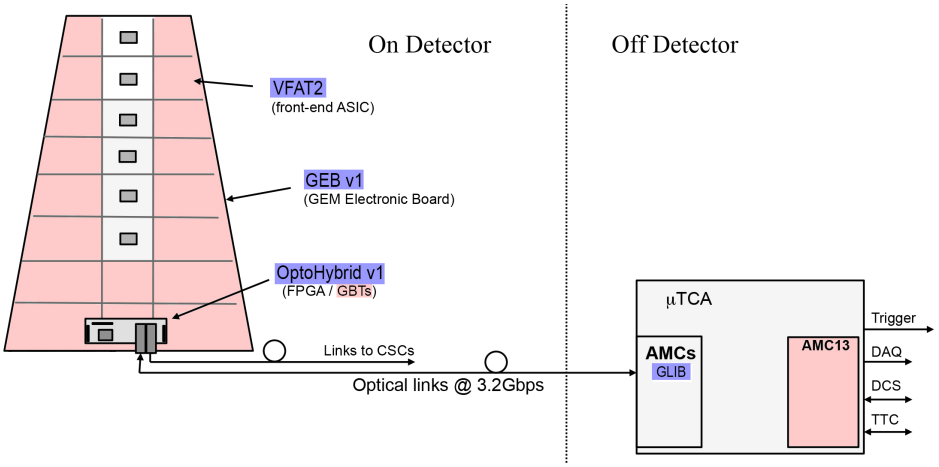

Figure 5.12 – Architecture of the first prototype of the GE1/1 DAQ system with the VFAT2 ASIC, first version of the GEB, first version of the OptoHybrid, and GLIB AMC.

The system was composed of 6 VFAT2 front-end chips, predecessor of the VFAT3, a first prototype of the GEB with only 6 out of the 24 positions active, and a first version of the OptoHybrid optically connected to the GLIB AMC serving as readout board. The usability of the system was tested during a test beam campaign in 2014-2015 which acted as proof of concept for the DAQ system architecture.

5.2.1 The VFAT2 ASIC

Similarly to the VFAT3 ASIC, of which the architecture has been presented above, the VFAT2 is a binary front-end chip used to digitize signals generated by gaseous detectors. A schematic representation of the chip is depicted in Figure 5.13. The 128 analog signals are preamplified, shaped, and discriminated against a settable threshold in the analog front-end which is asynchronous from the LHC clock. Synchronicity is achieved through monostables connected to each channel which allow to generate

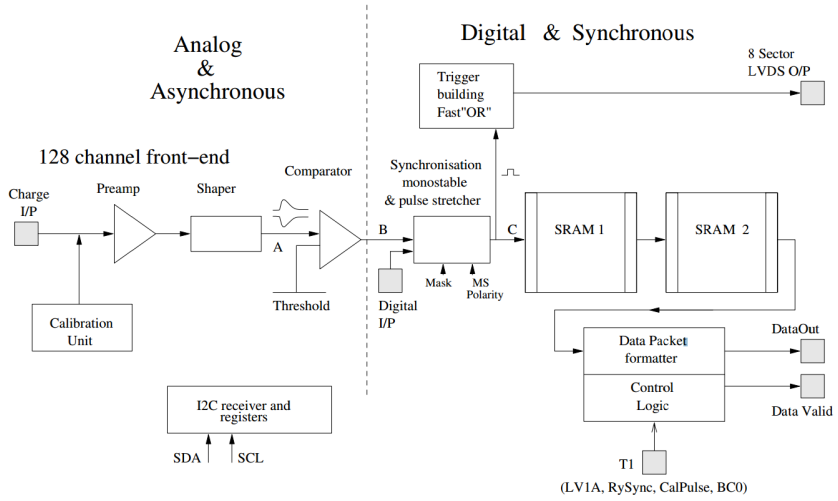

Figure 5.13 – Schematic diagram of the VFAT2 showing the analog front-end on the left and the digital control and readout system on the right [35].

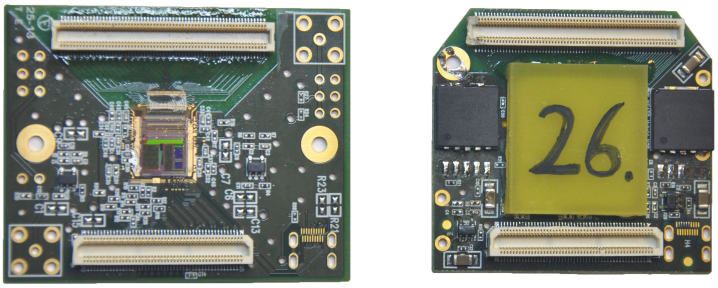

Figure 5.14 – Photograph of the VFAT2 Hybrids on which the VFAT2 ASICs are mounted to be interfaced with the GEB.

a binary pulse each time the threshold is reached in a given BX. Moreover, the duration of the output signal of the monostable can be stretched to cover multiple clock cycles. This function is used in the tracking data stored in the SRAM1 as well as the eight trigger bits, called SBits, formatted by the Trigger unit. Upon reception of a L1A, data is transferred from the SRAM1 to the SRAM2 and serially transmitted over the DataOut and DataValid lines. The DataOut line caries the 192 bits that form an event while the DataValid line signals the validity of the data by going to a logic '1' during transmission. The four fast commands (L1A, Resync, CalPulse, and BC0), called the T1 in the VFAT2 context, are received over a single differential pair and are encoded on three bits, meaning that no two consecutive L1As can be received within three BXs. Finally, the configuration of the ASIC is done using the Inter-Integrated Circuit (I2C) protocol.

The main differences between VFAT2 and VFAT3 are the communication protocols and the reduced bandwidth allocated to trigger data over the fixed latency path. While VFAT3 can be directly connected to the GBT chipset, the VFAT2 has to be controlled and readout by the FPGA located on the OptoHybrid. Furthermore, VFAT3 has a trigger granularity of 64 bits per BX while VFAT2 is limited to 8 bits per BX. Although the configuration of the bits can be chosen through programmable registers, it is common to divide the 128 strips in 8 groups of 16 neighboring strips and transmit one bit for each group.

As the VFAT2 ASIC does not come in a package like standard ICs, it is bound to a small PCB called the VFAT2 Hybrid depicted in Figure 5.14. The Hybrids are equipped with two connectors: one for the 128 analog input channels and one for the digital signals and the power. They are respectively connected to the GEM readout plane on one side, and the GEB on the other side.

5.2.2 The GEM Electronic Board v1

Due to the manufacturing challenge that is the production of a 1-m-long multilayer PCB, a first version of the GEB has been produced with minimal features. The first

prototype was equipped with only six functioning positions in the middle row closest to the small side of the board. The positions shared a common clock and trigger line and were divided into two slow control sectors. As fewer connections were needed between the GEB and the OptoHybrid, a smaller connector was also used.

5.2.3 The OptoHybrid v1

Mounted on the GEB v1 is the OptoHybrid v1, a preliminary prototype designed to demonstrate the concept of an FPGA-based concentrator board to make the communication with the back-end electronics. The OptoHybrid v1, shown in Figure 5.15, is equipped with a Xilinx Spartan-6 FPGA that concentrates the data from the six VFAT2 Hybrids and transmits it over four optical links to the off-detector electronics. It is further equipped with a MSP430 EZ430-RF2500 [36] microcontroller which is used as Analog-to-Digital Converter (ADC) to digitize the analog outputs of the calibration units of the VFAT2s, and some General Purpose Input/Output pins (GPIOs) for debugging.

5.2.4 The GLIB Advanced Mezzanine Card

Designed by CERN, the Gigabit Link Interface Board (GLIB) [37] has been developed as an evaluation platform for the GBT FPGA project (see Section 5.1.4). As shown in Figure 5.16, it is a general purpose microTCA AMC equipped with a Xilinx Virtex-6 FPGA, four SFP optical cages, and two extension headers that can accommodate custom electronics. The GLIB can be used within the microTCA crate in so called crate mode, or as standalone board in bench-top mode. To this end, it is equipped with an Ethernet connector to replace the fabric A line coming from the MCHs while

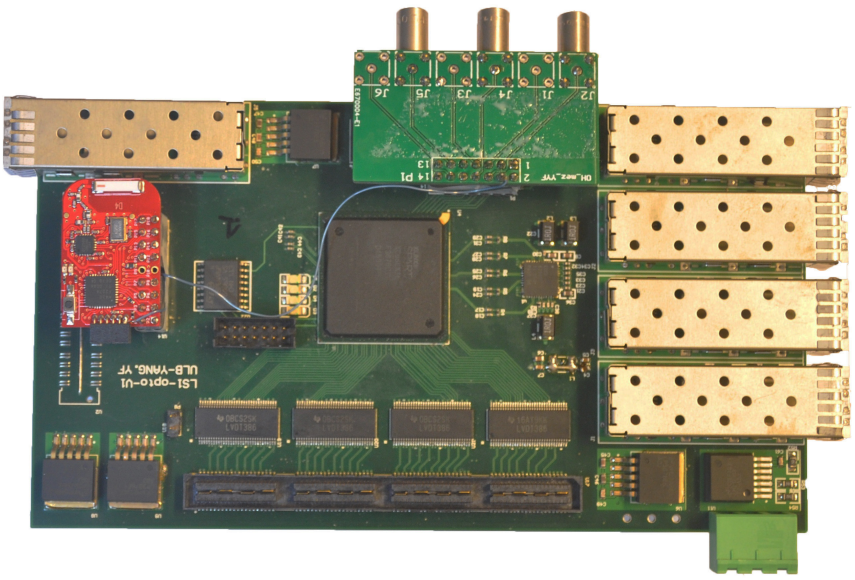

Figure 5.15 – Photograph of the OptoHybrid V1 equipped with a Xilinx Spartan-6 FPGA that concentrates the data from the six VFAT2 Hybrids and transmits it over four optical links to the off-detector electronics.

Figure 5.16 – Picture of the GLIB AMC equipped with a Xilinx Virtex-6 FPGA, four SFP optical cages, and two extension headers that can accommodate custom electronics and are used as back-end electronics [37].

still providing access using IPBus.

In this prototype version, the board has been used as concentrator for the data pushed by the OptoHybrid which is stored locally and then pulled by the monitoring application over IPBus. The slow control requests were also handled directly by the GLIB over IPBus and forwarded to the OptoHybrid.

5.3 The GE1/1 Data Acquisition System for the Test Beam

Incrementing on the first prototype developed in 2014, a second version of the system has been developed in 2015, and tested during test beams in November of the same year of which results are presented in Chapter 6. This system, shown in Figure 5.17, has many similarities with the first prototype, only increasing the number of supported VFAT2 Hybrids to 24 which, however, required a complete redesign of the GEB and the OptoHybrid. The back-end system is still handled by the GLIB board while controlled and read out through IPBus.

5.3.1 The GEM Electronic Board v2

The GEB v2 handles 24 VFAT2 Hybrids divided into three columns for the clock and fast control distribution, and six sectors for the slow control I2C protocol. The

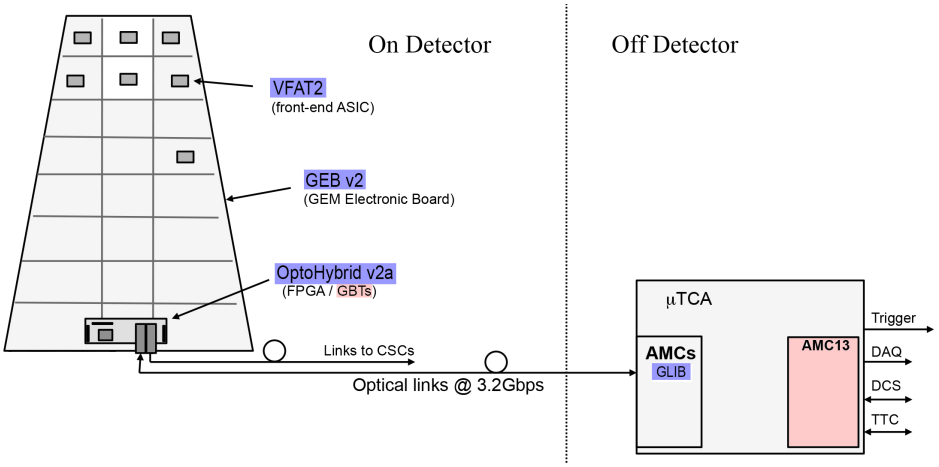

Figure 5.17 – Architecture of the GE1/1 DAQ system used during the test beam with the VFAT2 ASIC, second version of the GEB, second version of the OptoHybrid, and GLIB AMC.

connection to the OptoHybrid is done through four connector that transmit the trigger, tracking, and command signals from the VFAT2s to the FPGA.

5.3.2 The OptoHybrid v2a

The second version of the OptoHybrid is still a prototyping board which offers various configuration schemes to select the best operating mode. Represented in Figure 5.18, the OptoHybridv2a uses a Xilinx Virtex-6 FPGA similar to the GLIB. It can either use 12 SFP optical cages or a QSFP+ transceiver which offers four

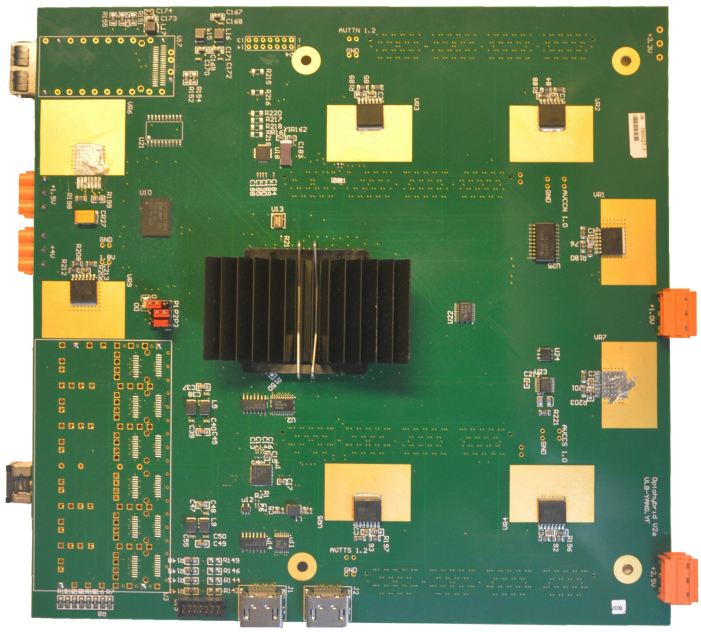

Figure 5.18 – Photograph of the OptoHybrid V2a equipped with a Xilinx Virtex-6 FPGA that concentrates the data from the 24 VFAT2 Hybrids and transmits it to the off-detector electronics.

optical connections in a single module. The LHC clock can be recovered by the Texas Instrument CDCE62005, a programmable clock generator, or by the radiation hard QPLL module developed by CERN specifically for this purpose.

5.4 The GE1/1 Data Acquisition System for the CMS Slice Test

The last prototype of the architecture, before developments for the final system start, is the DAQ system for the slice test, to be installed in January 2017. The architecture of the system is depicted in Figure 5.19. Although this small scale test will be performed using the VFAT2 ASIC as the VFAT3 is still in development, other components are more similar to the final version. The GBT chipset, for which production has started, is used. Readout is performed using the CTP7 and AMC13 connected to the central DAQ of CMS.

5.4.1 The OptoHybrid v2b

The second iteration on the OptoHybridv2 has a U shape to accommodate the constraints of the detector geometry, cooling system, and front-panel as shown in Figure 5.20. It is equipped with four optical links using the SFP cages and the VTRx and VTTx modules. Three of the links are controlled by the FPGA itself while the last one is connected to a GBT chipset in turn controlled by the FPGA.

5.5 Evolution of the DAQ System

The evolution of the DAQ system follows the readiness of the various components foreseen to be used in the final system. Table 5.1 provides a summary of the various prototypes developed and the systems that will be installed during the slice test and in the final setup. Changes are incremental allowing one to acquire expertise in the

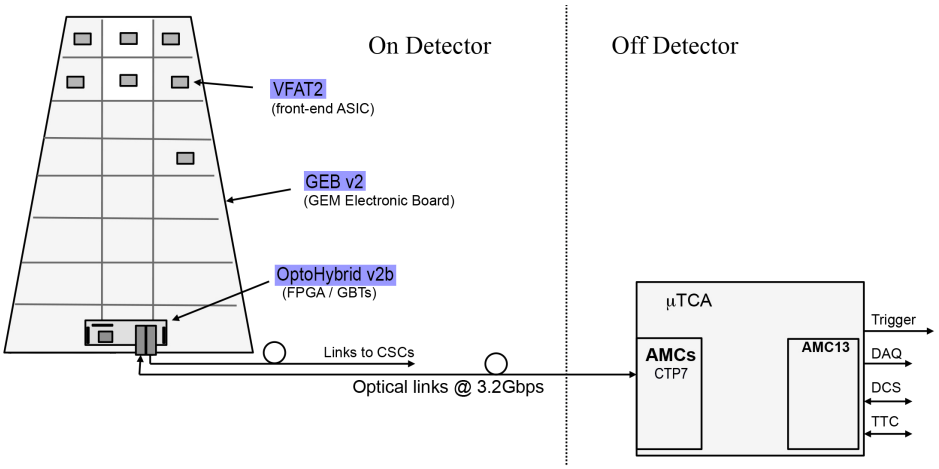

Figure 5.19 – Architecture of the GE1/1 DAQ system foreseen for the slice test with the VFAT2 ASIC, second version of the GEB, second version of the OptoHybrid, CTP7 AMC, and AMC13.

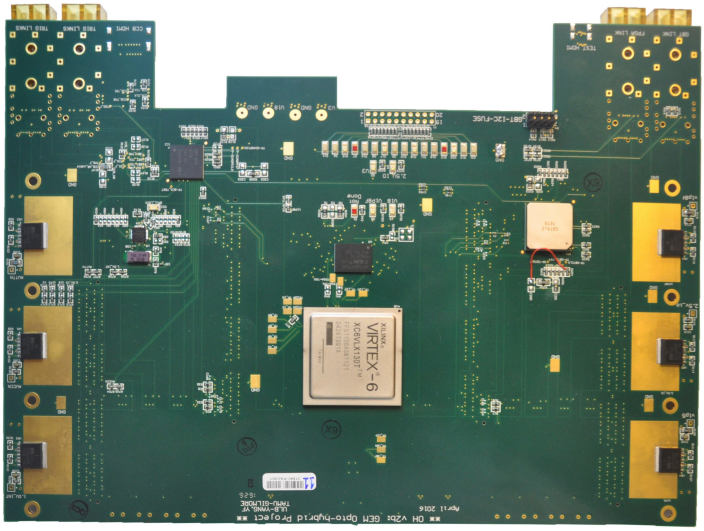

Figure 5.20 – Photograph of the OptoHybrid V2b that provides a testing platform for the GBT chipset.

control of the various technologies interacting in the architecture. They also provide feedback on components that can be updated in future versions or more deeply integrated in the system, such as the GBT chipset or the Versatile Link. To validate the design of the prototypes, test beam campaigns were organized during which the DAQ system was instrumented on detectors and exposed to particles beams.

First prototype	Test beam	Slice test	Final system
6x VFAT2	24x VFAT2	24x VFAT2	24x VFAT3
GEB v1	GEB v2	GEB v2	GEB v3
OptoHybrid v1	OptoHybrid v2a	OptoHybrid v2b	OptoHybrid v3
-	-	1x GBT	3x GBT
GLIB	GLIB	CTP7	CTP7
-	-	AMC13	AMC13

Table 5.1 – Summary of the evolution of the components of the DAQ system of GE1/1.

6 A Data Acquisition System for the Test Beam

During the ongoing process of the development of the DAQ system, the electronics was tested in two test beams organized in Fall 2014 and November 2015 respectively. The first test beam, run with the first prototype of the DAQ, aimed at proving the feasibility of the architecture of the system, not focusing on data taking from the provided pion and muon beam but rather on usability. As the results were encouraging, the second version of the electronics was developed involving a complete redesign of the hardware, firmware, and software. Therefore, we will mainly focus on the second test beam, of which the electronics were described in the previous chapter, during which abundant data was recorded showing both excellent results for the detectors and the DAQ system.

In this chapter, we present the firmware and software developments done for the DAQ system for the test beams, followed by the analysis of the recorded data. First, we will present the firmware architecture of the OptoHybrid and the GLIB to better understand the global layout of the system and the features that have been implemented to control the components. Then we will move on to the back-end applications developed to control and monitor the DAQ system and read out data. Finally, after a presentation of the layout of the test beam setup and its characteristics, we will review the results obtained after analysis of the recorded data.

6.1 Architecture of the OptoHybrid Firmware

The most important function of the OptoHybrid is to transfer data between the VFAT2s and the GLIB. Downwards, from the off-detector to the on-detector electronics, it transfers slow control requests and fast commands while upwards it sends the trigger and tracking data back to the back-end system. Next to the handling of the basic VFAT2 functionalities, it must also handle the optical links and a couple of programmable registers which control the system. Furthermore, procedures that were previously implemented in software requiring extensive computational time have been moved to firmware in order to speed up the system.

Although rather complex, the modules implemented in the firmware of the OptoHybrid can be regrouped under six categories: the fast commands, the trigger data, the tracking data, the slow control, the calibration routines, and the optical links. Even though these will be presented separately, they are tightly interconnected using a wishbone-like [38] architecture for intercommunication. A summary of the

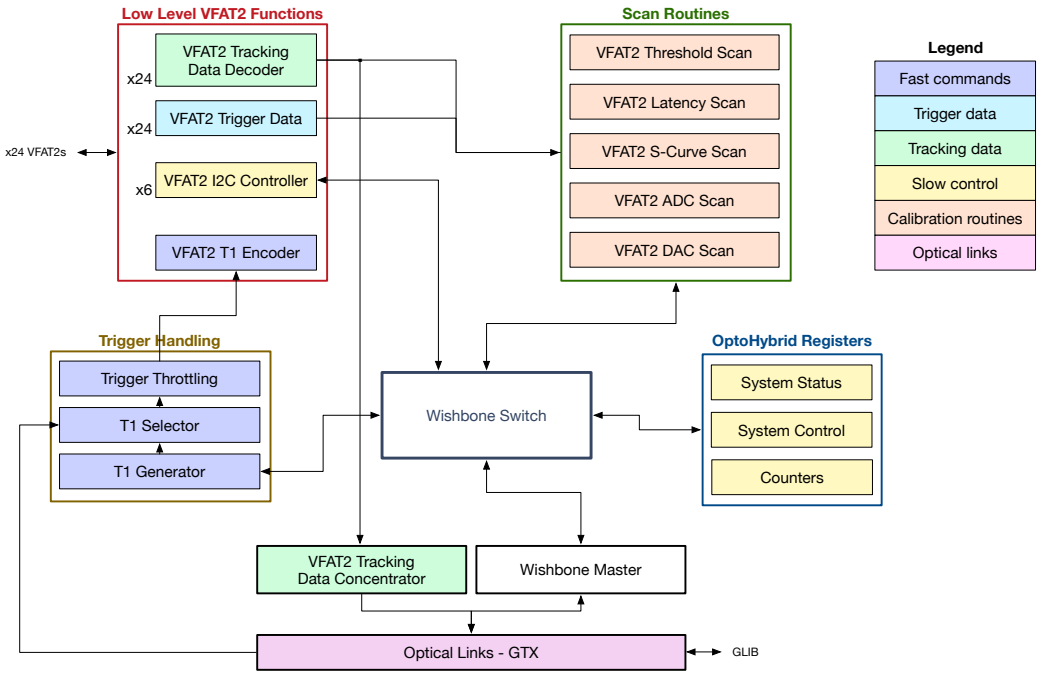

Figure 6.1 – Schematic diagram of the modules implemented in the firmware of the OptoHybrid regrouped by functionality.

system is provided in Figure 6.1 through a schematic representation of the modules implemented in the firmware of the OptoHybrid and the categories they fall in.

6.1.1 Internal Communication Through a Wishbone-Like Bus

Wishbone is an open-source communication protocol used in many projects to enable data transfer between ICs. A light version of the protocol has been implemented in the OptoHybrid to link the various modules of the system. The latter are divided in two categories: masters which initiate requests, and slaves which provide responses. The link between the two is created through a switching hub which redirects the requests to the appropriate slaves by using a 32-bit-long address space mapped to the various modules. Some components implement both a master and a slave module in order to be able to receive requests and propagate them to various other modules. Figure 6.2 is a schematic illustration of the architecture of the system without Wishbone on the left and with Wishbone on the right. In the architecture of the first version of the firmware, all requests had to originate from the software routines which were the only masters of the system. This meant that the firmware modules could not communicate among themselves. In the second version, firmware routines can be developed which act as masters in the system and have direct access to all resources of the FPGA in a unified way. The right part of the figure shows how the software sends a request to the firmware routine in red, and how the firmware routine then handles the communication with the other modules. All the requests transit through the switch and are redirected to the appropriate slaves.

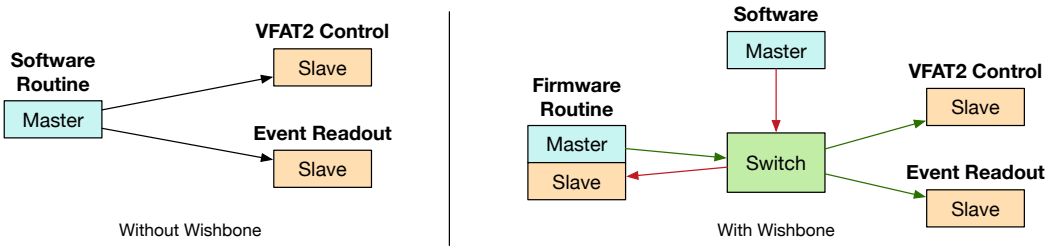

Figure 6.2 – Schematic illustration of the architecture of the system without Wishbone on the left and with Wishbone on the right.

To highlight the benefits of the implementation of the Wishbone-like bus on the architecture, Figure 6.3 presents the flow of operations for a custom calibration routine performed on the system without Wishbone on the left and with Wishbone on the right. The IPBus requests that the system has to perform are highlighted in blue, the I2C requests in orange, and the computations in green. In the system on the left, the software has to perform an IPBus request for every single action it needs to perform. Each request is sent over the network and has thus a large latency, higher than 100 ms. For complex routines, this can add up to several minutes and thus block the system. In the architecture of the second version of the firmware, the OptoHybrid itself handles the operations with the VFAT2. A single IPBus request is sent from the software to start the scan and all other operations are done in parallel by the firmware. The analysis of the data is done directly upon reception of an event and the results are stored in memory. Meanwhile, other operations can be performed by the software on the system. At the end of the scan, the software can collect the results and display them directly.

A request from a master is composed of four signals: a flag signaling the presence of a request, a write-enable bit to indicate the nature of the request (read or write), a

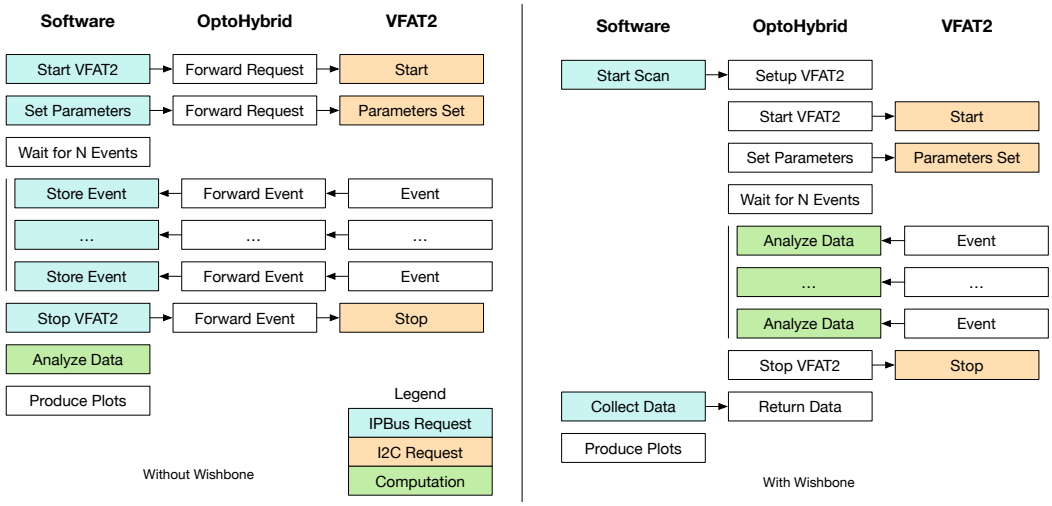

Figure 6.3 – Flow of operations for a custom calibration routine performed on the system without Wishbone on the left and with Wishbone on the right.

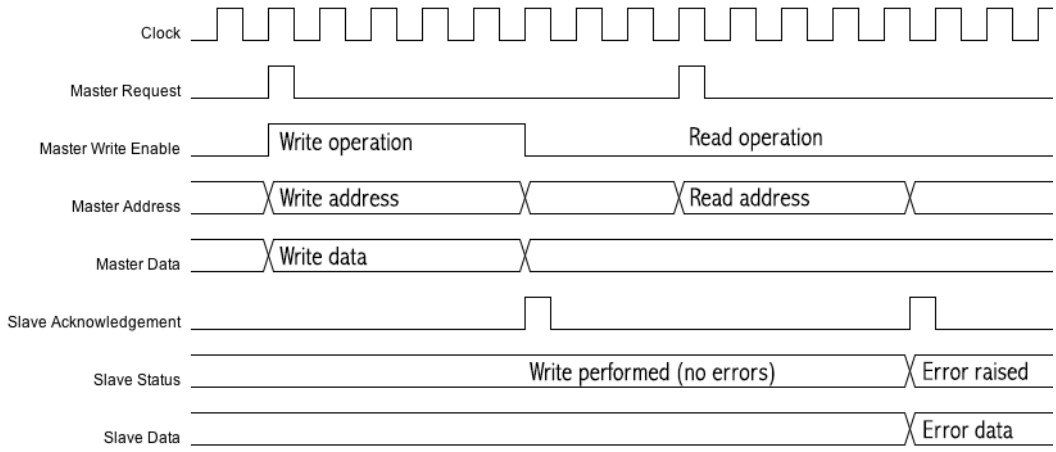

Figure 6.4 – Chronograph of two wishbone communications between a master and a slave for a successful write operation and a read operation resulting in an error.

32-bit-long address to which the request will be redirected, and an optional 32-bit-long data field in case the request is a write operation. The response from a slave consists of: an acknowledgment signal, a 4-bit-long error status in case the operation failed, and an optional 32-bit-long data field holding the response to read requests. Figure 6.4 represents a chronograph of two wishbone communications between a master and a slave for a successful write operation and a read operation resulting in an error. A programmable timeout has also been implemented to avoid blocking operation.

Communication done through the wishbone-like protocol is used mainly for slow control or non-time-critical operations as the latency of a transactions is not fixed. Indeed, the switching hub implements a waiting list functionality, which allows two masters to address the same slave at the same time by storing one of the requests in memory and awaiting for the slave to finish the other transaction.

6.1.2 Encoding Fast Commands

Fast commands such as L1As, Resyncs, BC0s, etc can originate from various sources. In normal data taking runs, the AMC13 receives the TTC signals common to CMS, forwards them to the microTCA AMC which in turn transmits them to the OptoHybrid. When data taking is stopped and calibration runs are performed, the routines implemented in the OptoHybrid are run and require fast commands. To this end, the TTC signals can also be generated locally using the T1 generator module. This entity can generate L1As, Resyncs, BC0s, and CalPulses in three different ways: send a single command a given number of times at fixed interval, send a CalPulse followed by a L1A with a fixed delay, or send a programmable pattern involving all the commands. Figure 6.5 illustrates those functionalities through a chronograph of two events generated by the T1 generator and encoded on the T1 lines for the

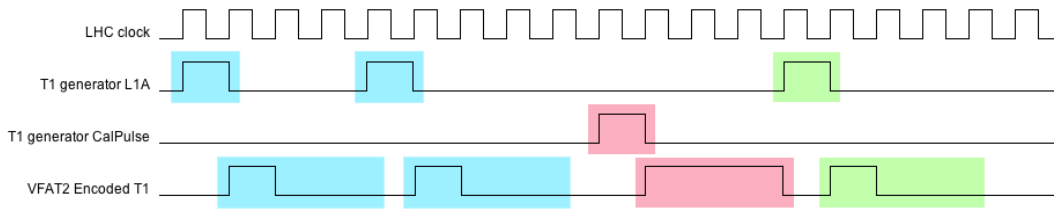

Figure 6.5 – Chronograph of two events generated by the T1 generator and encoded on the T1 lines for the VFAT2s: first two L1As (blue) are sent with a fixed interval, then a CalPulse (red) is generated followed by a L1A (green) after a defined delay. The T1 commands are encoded on three bits for the VFAT2s ("100" for L1As and "111" for CalPulses).

VFAT2s: first two L1As (blue) are sent with a fixed interval, then a CalPulse (red) is generated followed by a L1A (green) after a defined delay. The T1 commands are then encoded on three bits for the VFAT2s (100 for L1As and 111 for CalPulses).

The two remaining sources of TTC commands, and more precisely L1As, are either a loopback from a given VFAT2 trigger bits directly to all the VFAT2s (self-triggering) or the signals coming from an external component through a debugging header on the OptoHybrid. The former can be used to trigger on noise or when no other source of triggers is available; the latter has been used during the test beams in conjunction with scintillators. The switching between the various sources is done by setting a register through slow control operations. Once the source has been selected, the commands are forwarded to the VFAT2s and encoded on the T1 signals. Each command is composed of three bits clocked at 40 MHz. An additional feature has been added to the L1A line to throttle the trigger signals when working in high rate environments. Through registers, the operator can select to send only a fraction of the received commands not to overload the VFAT2 buffers and allow for correct readout of the chip. Figure 6.6 summarizes the encoding of fast commands through a diagram of the switching and throttling process for T1 commands highlighting the various sources of T1 commands (blue) and the involved modules (green) before the signal is sent to the VFAT2 chips (orange).

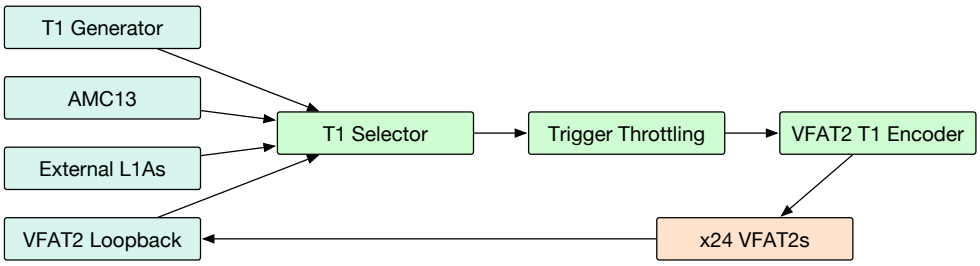

Figure 6.6 – Diagram of the switching and throttling process for T1 commands highlighting the various sources of T1 commands (blue) and the involved modules (green) before the signal is sent to the VFAT2 chips (orange).

6.1.3 Formatting Trigger Data

Each of the 24 VFAT2s transmits eight trigger bits per BX. These are regrouped using a logic OR in the firmware of the OptoHybrid thus yielding 24 bits. The trigger information is used for calibration routines and forwarded over debugging headers to an external electronics crate. As the number of external connections is limited, only six trigger bits can be sent, thus requiring the need for a selection mechanism controlled by programmable registers.

6.1.4 Acquiring Tracking Data

Upon reception of a L1A, the VFAT2 transfers the selected event from its SRAM1 to SRAM2 and then to an encoder which serializes the data. Each event is 192 bits long followed by two idle bits and clocked at 40 MHz. This means that it takes 194 BXs to read out one event and that the VFAT2 cannot handle trigger rates higher than $\approx$ 200 kHz without experiencing an overflow of the buffers. The data is pushed out of the VFAT2 automatically and formatted according to the pattern shown in Table 6.1. The top left bits are the Most Significant Bits (MSBs) which are pushed out first and the bottom right bits or the Least Significant Bits (LSBs) which come in last. The first four bits of the first three 16-bit-long words are constant values. These are completed by the BC which increments at each clock cycle, the EC which counts the number of received L1As, four flags which hold information on the status of the buffers, and a chip ID unique to each VFAT2. Following are the 128 bits reflecting the hit information on each channel. Finally, the VFAT2 uses a Cyclic Redundancy Check (CRC) on 16 bits to detect errors. The CRC uses all other 176 bits to encode its data but does not provide a way to correct errors.

VFAT2 tracking data

15		0
1010	BC[11:0]	
1100	EC[7:0]	Flags[3:0]
1110	Chip ID[11:0]	
Channel data x8		
CRC		

Table 6.1 – Format of the tracking data packets sent by the VFAT2s.

Next to the data line, each VFAT2 provides a DataValid line which is pulled high when the bits exiting the VFAT2 are valid. However, due to the limited number of pins connecting the GEB and the OptoHybrid, this signal has been left unconnected for all but six VFAT2s. Therefore, data is constantly shifted in a 194 bits (192 bits of data and 2 idle bits) serial-to-parallel converter and analyzed. When the fixed pattern of 12 bits is seen, a flag is raised signaling the presence of potential

data. The data packet is split up in its various elements and the CRC is recomputed and compared against the received CRC. Two additional flags respectively hold the results of the comparison of both CRCs indicating if the data is valid or not, and a logic OR of all 128 channel to indicate if the packet contains a hit or not. Figure 6.7 illustrates this process through a chronograph of the serial decoding of an incoming data packet (blue) which is analyzed and placed in parallel busses (yellow) with the CRC check bit and the hit present bit. In theory, only packets with valid CRCs could be transmitted to the back-end electronics. However, sending all packets offers the possibility to identify recurring errors in the CRCs computation and the possibility to correct them offline. To this end, a programmable register is used to select whether packets with a corrupted CRC should be sent or not.

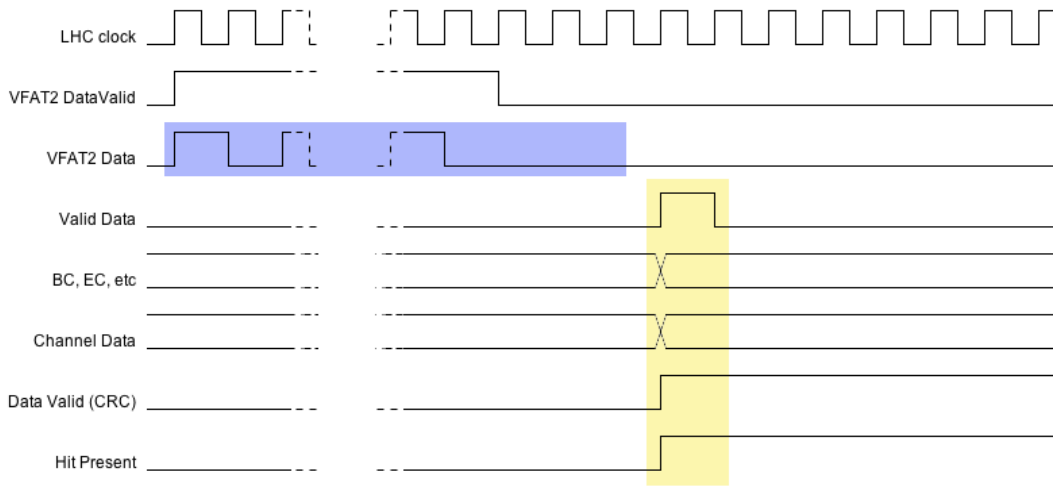

Figure 6.7 – Chronograph of the serial decoding of an incoming data packet (blue) which is analyzed and placed in parallel busses (yellow) with the CRC check bit and the hit present bit.

In parallel to the decoding of the data packets, the OptoHybrid maintains its own BC and EC and stores the value of the counters in a buffer each time a L1A is received. When the 24 decoding modules, one for each VFAT2, report they have detected data, a concentrator module aggregates the information from all available VFAT2s. Every packet received within a window of ten BX is assigned to the same event to which the value of the corresponding BC and EC is appended. The assembled event is stored in a large buffer to be later sent over the optical links to the GLIB. Additionnaly, the data decoder can perform zero suppression on the packets, suppressing them when no strips have been hit in the event.

To prevent positions either not equipped with VFAT2 Hybrids or that are noisy from generating fake data, a 24-bit-long register allows the masking of individual positions

to ignore any packets it generates. Next to this, each decoding module is equipped with two counters respectively counting the number of valid and invalid packets.

6.1.5 Control and Monitoring

The OptoHybrid is used to control itself through wishbone and the VFAT2s through I2C [39]. The slow control of the OptoHybrid mainly consists in selecting TTC command sources, setting the trigger throttling, reading out counters, etc. These operations are performed to control the data flow and the functioning of the DAQ system itself. Commands sent to the VFAT2s on the other hand have direct impact on the physics of the data taking through the bias of the analog front-end readout.

The I2C protocol uses two signals to connect one master to multiple slaves: one clock (SCK) and one data (SDA) line. The master controls the clock signal which defines the reference clock of the exchanges. The data line on the other hand is driven by both the master and the slaves according to a defined data format. The official I2C data frame is shown in Figure 6.8. A communication is always started by the master by driving SDA low when SCK is high. The master then sends seven address bits (yellow) which are used to identify the slave it wants to talk to, followed by a read/write bit (R/W, blue). If a slave was configured to respond to the address, it will drive SDA low to signal an acknowledgement (ACK, green). Afterwards, eight bits (red) are sent from the master to the slave in case of a write operation followed by a slave acknowledgment (ACK, green), or the opposite in case of a read operation. The frame ends when the master drives SDA high while SCK is low.

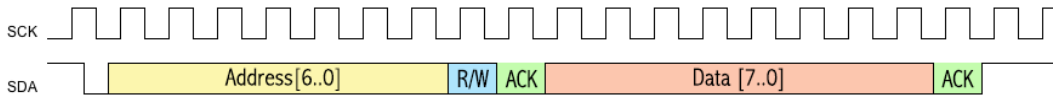

Figure 6.8 – Chronograph of an I2C communication.

To communicate with the VFAT2s, six I2C controllers are implemented in the firmware of the OptoHybrid, one per sector on the GEB. Each controller can access four VFAT2s which are identified using three resistors that can be installed on the GEB. VFAT2 uses a modified I2C protocol with the same frame format as the official one, but a different addressing scheme. For the VFAT2, the first three bits of the address are used to select the chip that is addressed by the master while the four remaining bits are used to select the register that needs to be accessed. This addressing scheme allows the OptoHybrid to access up to 16 registers on the VFAT2s. Figure 6.9 represents a chronograph of the modified I2C communication for the VFAT2.

However, each VFAT2 holds 16 primary registers and 136 extended registers. The latter are accesses using two primary registers: one set to point to an extended

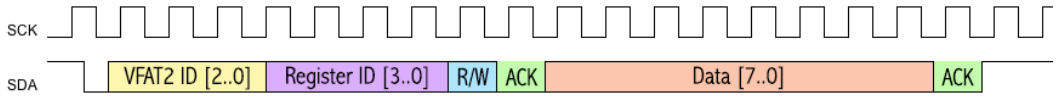

Figure 6.9 – Chronograph of an I2C communication performed with the VFAT2.

register (Extended Pointer), and one to read/write the data in said register (Extended Data). Thus, in order to perform a transaction on a primary register, only one operation is necessary, while two are required to access extended registers. A first operation will modify the content of the Extended Pointer register to point to the desired extended register, and a second operation will read/write the Extended Data register to affect the extended register. Figure 6.10 illustrates how the extended registers are addressed through two primary registers.

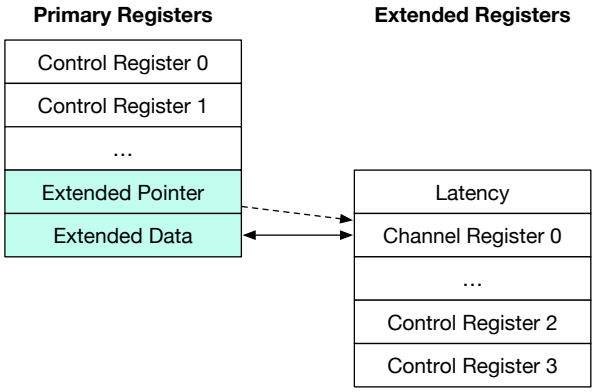

Figure 6.10 – Addressing of the extended registers through two primary registers.

To make this process transparent to the user and flatten out the address space of the VFAT2s, the OptoHybrid maps the primary registers from addresses 0 to 15 and the extended registers from addresses 16 to 151. When performing a transaction with the extended registers, the I2C controllers automatically run the double addressing scheme.

Next to the basic I2C controllers, an extended I2C controller was developed to abstract individual addressing and allow request broadcasting. This module forwards I2C requests to all selected VFAT2s at once to configure the entire system in parallel. A programmable register makes it possible to leave out given VFAT2s from the broadcast while a buffer stores the result of the operations.

6.1.6 Calibrating the Systems

The calibration routines have been ported from software to firmware in order to increase their speed of operation and reduce the number of requests that need to be performed. The routines are as follows: a threshold scan which measures the noise on the channels as a function of the threshold of the VFAT2s; a latency scan

which allows to select the correct BX when receiving L1As; a s-curve scan which characterizes the response of the channels as a function of the collected charge and threshold; an analog-to-digital converter (ADC) scan which reads out the information from the ADC embedded in the FPGA; and a digital-to-analog converter (DAC) scan which is used to convert information stored in the VFAT2 registers to voltages and currents. Two versions of the scans exist: one which performs the scans on a single VFAT2, and one which is able to run on multiple VFAT2s in parallel using a mask to define which ones it should target. The latter improves the speed by a factor of 24 compared to the former when performing systematic scans on the entire detector.

The threshold scan is used to scan each VFAT2 for noise. For each threshold value set on the VFAT2, the percentage of events displaying a hit is recorded and taken as the percentage of noise. A graph, as depicted in Figure 6.11, can be obtained showing the noise decrease as the threshold increases yielding a point at which the system can be operated with minimal noise. The threshold scan can be operated at the chip level using trigger information or the 128 channels as a whole, or on individual channels using tracking data. For the latter, the T1 generator is used as trigger source to generate data as these runs are being performed when the beam is off. No relation between a L1A and physical event is needed to study the noise on the system.

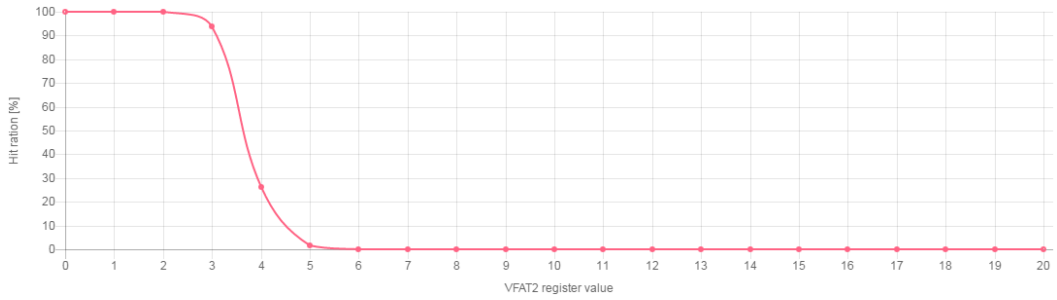

Figure 6.11 – Illustration of a threshold scan representing the hit percentage as a function of the threshold of the VFAT2.

The latency scan allows to determine the best latency value to be set in a VFAT2. The latency is the time difference, in number of BX, between the time of arrival of a L1A and the time at which the related event was stored in the VFAT2 buffer. This module is operated when the particle beam is on and triggers are generated by an external source such as a Photomultiplier (PM) placed in front of the detector. For each value of the latency, the OptoHybrid counts the ratio of events with hits over the total number of events. Figure 6.12 is an illustration of an ideal latency scan representing the hit percentage as a function of the latency of the VFAT2. For a noiseless VFAT2 with 100% detection efficiency, the ratio would be 0% outside the correct latency window and 100% inside. The size of the window can be adjusted by changing the length of the output of the monostables in the VFAT2s.

Figure 6.12 – Illustration of an ideal latency scan representing the hit percentage as a function of the latency of the VFAT2.

The s-curve scan is part of the calibration routines performed for the qualification of the VFAT2s. It yields the response of the VFAT2 channels to an injected charge pulse according to the threshold. It is used in conjunction with the T1 generator which sends a CalPulse followed by a L1A at fixed interval, thus with known latency. The use and results obtained for this module are further detailed in Chapter 7.

The ADC scan reads out the ADC embedded inside the fabric of the FPGA of the OptoHybrid. The Xilinx Virtex-6 FPGAs contain a system monitor [40] which offers the possibility to convert voltages to ADC counts with a resolution of 10 bits over 1 V, resulting in a precision of 977 μV. The inputs of the ADC are connected to voltage lines of the OptoHybrid and to analog signals coming from the VFAT2s.

The DAC scan is used to characterize the analog front-end of the VFAT2s. When a digital value is written in given registers of the VFAT2, it is converted to a voltage or current in the amplification and shaping stages of the electronics. These can optionally be driven out of the VFAT2 for monitoring and be connected to the ADC of the OptoHybrid. The DAC scan provides a routine which associates every value written in the registers of the VFAT2s and the corresponding ADC counts, and thus voltage or current.

6.1.7 Optical Communication with the Off-Detector Electronics

The connection between the OptoHybrid and the off-detector electronics is the optical links from which two are used: one for the trigger data and fast commands, and one for the tracking data and slow control. Each optical fiber is connected to an SFP+ transceiver which contains the Light Emitting Diode (LED) that drives the link. The SFP+ module is in turn connected to a Gigabit Transceiver X (GTX) block inside the FPGA, which is a dedicated component that can handle high-speed serial communications.

The GTX module is in charge of the serialization/deserialization of high-speed data up to 6.6 Gbps. A schematic diagram of the functions it provides is shown in Figure 6.13. The GTX handles both the transmission and reception of data. From the FPGA logic, it receives a parallel vector of bits containing the raw data. The raw data

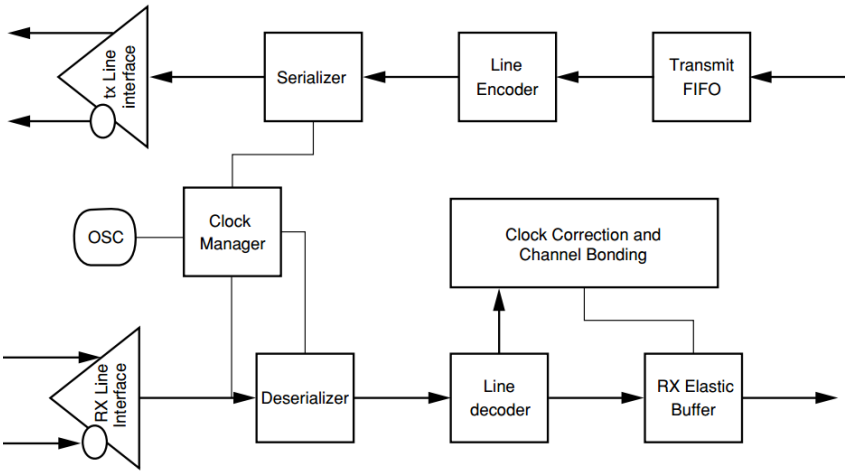

Figure 6.13 – Functionnal diagram of the Gigabit Transceiver X module used inside the Xilinx Virtex-6 FPGAs for high-speed serial communication [41].

enters the Transmit FIFO which acts as a buffer before the Line Encoding module. The latter uses an eight-bit/ten-bit (8b/10b) encoding scheming which is a standard in high-speed communications. 8b/10b encodes every eight bit word on ten bits according to a predefined table. It has the particularity that whatever combination of words might be sent, no more than five successive '0's or '1's will arise in the serial stream. This ensures line balancing: the fact that the electric lines between the GTX and the SFP+ module will remain at an average voltage and not drift to a logic '1' or '0' over time, thus requiring a longer fall or rise time of the signals. 8b/10b also offers a set of K-characters which are unique patterns of bit that can never arise randomly in the stream. Those characters are used to detect the beginning of a frame and allow the receiver to align correctly. Finally, after the encoding of the data frame is done, it is serialized and transmitted to the SFP+ module over differential pairs.

The receiver essentially follows the opposite path as the transmitter except for the Clock Management block. The 8b/10b encoding offers one more advantage which is the possibility to recover a clock from the data stream itself, technique called Clock Data Recovery (CDR). In high-speed signal design, small variations in the clock frequency between the transmitter and the receiver can lead to discrepancy in the read out data. To avoid this, the receiver can tune the frequency of the clock of its deserializer using the transitions of the bits in the incoming stream to align its phase to the data. The GTX block further provides the obtained recovered clock to the FPGA to use in the logic. Correct alignment of the data is checked using the K-characters in the stream.

The fixed latency link is used to transmit trigger data from the OptoHybrid to the GLIB and receive the fast commands at a speed of 3.2 Gbps. The data format used to communicate between the on-detector and off-detector electronics is detailed in

Table 6.2. Of the 64 bits that the downlink transmits every BX, 16 are used to align the serial bit stream on the receiving side through a header and four others are used to encode the four fast commands that the VFAT2 understands. The other 44 bits are not used and left to constant zeros. The uplink uses a 16 bit header to align the stream to which it appends 24 bits, one bit per VFAT2 corresponding to the logic-OR of its trigger bits. These patterns repeat at 40 MHz as data changes every BX.

Fixed latency link

GLIB to OptoHybrid frame

15	0
Header	
TTC[3:0]	Unused[11:0]
Unused	
Unused	

OptoHybrid to GLIB frame

15	0
Header	
Trigger data[15:0]	
Trigger Data[7:0]	Unused[7:0]
Unused	

Table 6.2 – Format of the data packets used to communicate between the GLIB and OptoHybrid on the fixed latency link.

The variable latency link is used to transmit and receive slow control commands and to transmit the tracking data of one VFAT2 at the time to the GLIB. As only one VFAT2 was exposed to the beam at any given time during the test beam, reconstructing and sending the full event from the 24 VFAT2s was not a necessity for the DAQ system. The format used to transmit the data is detailed in Table 6.3. In both directions, a given frame format is constantly sent using two bits in the header to indicate the presence of a valid slow control request and valid VFAT2 tracking data in case of the uplink. Both send a 16 bit header to align the link followed by the frame data. As can be seen in the table on the left, the OptoHybrid receives requests from the optical link which are then converted to wishbone-like requests by the logic. The optical link module is one of the wishbone masters that controls the system and whose responses are forwarded back to the GLIB.

6.1.8 System Summary

The aim of the firmware of the OptoHybrid is to provide a communication medium between the VFAT2s and the off-detector electronics and to speed up the calibration routines by implementing them closer to the detector. The developments that have been listed in this chapter are targeted at the OptoHybrid v2a and the configuration of the system ran during the test beams. Modifications have to be brought for the final system although the foundations of the firmware have been solidly laid out.

Variable latency link

GLIB to OptoHybrid frame

15	0
Header	
Request address (MSB)	
Request address (LSB)	
Request data (MSB)	
Request data (LSB)	
CRC	

OptoHybrid to GLIB frame

15	0
Header	
VFAT2 data 14x 16 bits	
Response data (MSB)	
Response data (LSB)	
CRC	

Table 6.3 – Format of the data packets used to communicate between the GLIB and OptoHybrid on the variable latency link.

6.2 Architecture of the GLIB Firmware

The GLIB development group at CERN provides a system firmware which controls the on-board components, handles the communication with the MCH, and implements the logic to decode IPBus requests described in Section 5.1.8. This code is similar to the operating system in which the users run their applications by adding custom firmware. The latter is project specific and up to the user to develop. In this application, the firmware developed for the GLIB is in charge of the buffering of the tracking data until it is read out by the application, and the forwarding of the IPBus requests over the optical link to the OptoHybrid. In addition, the GLIB holds a couple of counters which monitor the status of the optical links and the number of requests performed. However, the firmware of the GLIB has been designed to be able to handle two OptoHybrids in parallel, seamlessly allowing the user to switch from one to another at any time.

6.2.1 The Tracking Data Readout

The GLIB acts as buffer for the tracking data it reads out from the OptoHybrids and decodes from the optical links. It stores the VFAT2 events until the control and monitoring application reads them out through IPBus. In case the buffers reach their full occupancy, new events are discarded and flags are raised. The buffers can hold up to 16 383 VFAT2 events per connected OptoHybrid. Additionally, the system has been tested and proven to work without overflows up until L1A rates of 200 kHz, the maximum rate the VFAT2s can handle.

6.2.2 The IPBus Controller and Slaves

The IPBus decoding procedure is similar to the wishbone system implemented on the OptoHybrid. According to the address contained in the request, the GLIB selects the slave it needs to address. Each request is made of a flag signaling the presence of a request, a read/write bit reporting the type of operation, a 32-bit-long

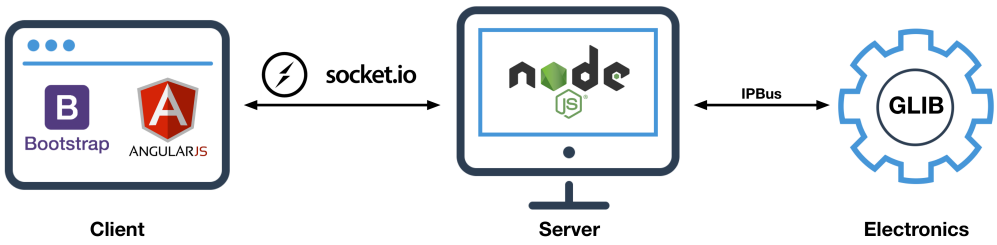

Figure 6.14 – Functional diagram of the architecture of the application used to control and monitor the GEM detectors.

address field, and an optional 32-bit-long data field populated with data in case of a write operation. The IPBus response is composed of an acknowledgment and error flag to signal the validity of a response, and an optional 32-bit-long data field containing the read data. By design, each IPBus request is decomposed in single read or write operations. Therefore, even if the software performs a read request of multiple addresses at once, IPBus slaves will still see this as multiple sequential read operations on single addresses. This in turn propagates to the OptoHybrid wishbone requests which originate from IPBus, meaning that no more than one operation coming from the GLIB is present at any time in the system. As a majority of the requests have to be forwarded to the OptoHybrid, an entire address space of the GLIB is mapped to the former. These are handled by an IPBus slave connected to the GTX blocks of the GLIB. This module encodes the request according to the data format on the left in Table 6.3 and awaits for a response before signaling the IPBus controller.

6.3 The Control and Monitoring Application

The first version of the control and monitoring application that communicated with the system was based on a series of Python scripts running PyChips, a Python implementation of IPBus provided with the GLIB source code. Each script was developed to run a specific task and needed to be run over and over again. This was feasible for the first test beam during which only the system was tested for small amounts of time, but not for the second one that was meant to collect data. Therefore, a web application was developed based on NodeJS [42] as server technology, SocketIO [43] for the client/server exchange, AngularJS [44] as front-end framework, and Bootstrap [45] as design framework. The architecture of the application is depicted in Figure 6.14 which shows how the various technologies interact.

6.3.1 Architecture of the Application

NodeJS is a JavaScript (JS) runtime that has the ability to run JS code outside the context of a web browser. It extends JavaScript from its original use in web content to a wide range of applications by enabling it to manipulate files, data streams, etc. The most common use for NodeJS is the implementation of a web server that has

the ability to be much more reactive than other infrastructures. Where classic server structures display a blocking behavior, meaning they only handle one request at a time, blocking the queue of incoming demands, NodeJS is event driven and thus non-blocking. For example, if a client requests a file, NodeJS will transfer the request to the operating system and continue to respond to other client as long as the file is not served. Interactions are thus more reactive and allow for easy implementation of a real-time control and monitoring system. Alongside the NodeJS server that serves the content of the web application, a version of IPBus in JS was also implemented. The latter allows the server to execute IPBus requests on the GLIB and readout the response.

To link the client side controlled by the user and the server side which sends requests to the GLIB, the SocketIO project is used. It is an abstraction layer on the fairly new WebSocket technology which adds sockets to web browsers and allows for fast communication with servers by bypassing the HTTP protocol. When the client changes settings in its web browser, SocketIO transfers the request to NodeJS which in turn creates an IPBus packet. When the GLIB responded to the request, the server transfers the response back to the client which can handle the data as it likes.

To handle data in a dynamic and convenient way, the web pages use Bootstrap and AngularJS. Bootstrap is a design framework which provides simple and readable themes to layout the pages. The content of the page on the other hand is managed by AngularJS which is a front-end JS framework. It helps render dynamic content in function of the data received from the server. Actions can be triggered or disabled according to the state of the system, data can be updated live, etc. It renders the user experience more fluid and enjoyable by providing a usable set of actions and an efficient navigation through the control pages.

6.3.2 Controlling the Systems

The control over the GLIB, OptoHybrids, and VFAT2s is done through four pages. The home page of the application, which is shown in Figure 6.15, provides an overview of the status of each system. The left panel displays the firmware version of the GLIB and OptoHybrid and the occupancy of the GLIB readout buffers; the middle panel gives the status of the various modules of the OptoHybrid and its temperature; finally, the right panel lists the VFAT2s that are turned on in green, turned off in red, and missing in gray and the rate of data they transmit. By clicking on any of the elements, the user is redirected to the corresponding system control page.

The control page of the GLIB is shown in Figure 6.16. The information it contains is mainly related to counters which allow to see, for example, the number of requests sent to the OptoHybrid and the number of responds received. It also counts the number of errors detected on the optical link providing a quantitative value of the

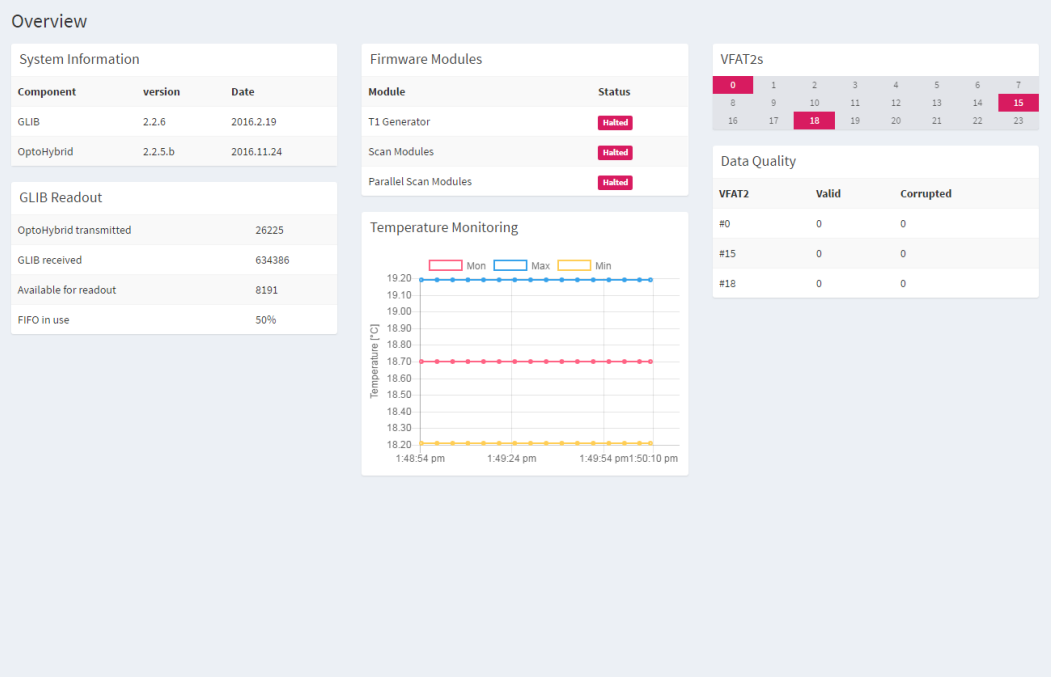

Figure 6.15 – Screenshots of the home page of the web application used to control and monitor the components of the DAQ system.

health of the link.

The OptoHybrid page, shown in Figure 6.17, provides more information and control options to the user. The top-left part of the page is used to modify system register controlling the trigger source, clock source, trigger information sent over the debugging header, etc. This is where the user can change the type of ongoing run by going from a data taking run using external triggers to a calibration run using internally generated triggers. The panel directly on the right provides status information on the various clocking resources in the FPGA. By default, the OptoHybrid uses the on-board oscillator to power the system and the optical links. As communication is established with the GLIB, the GTXs provide the recovered clock from the link which is the system clock. When the latter is stable, the OptoHybrid can decide to switch from its on-board clock to the system clock in order to be synchronous with the whole DAQ system. This operation can be prohibited, forced, or allowed by the user, which can also chose to use an external clock provided through the debugging header. These clocking options are useful for the test beam setup where the clock can be provided externally to synchronize with other electronics equipment. Finally, the bottom part of the page display counter values for the wishbone request, fast commands, link errors, etc. Note that when a discrepancy is observed between two correlated counters, a warning flag is raised. Additionally, the interface allows the user to monitor the evolution of the temperature and voltages of the OptoHybrid. Figure 6.18 is a screenshots of the page which enables this feature.

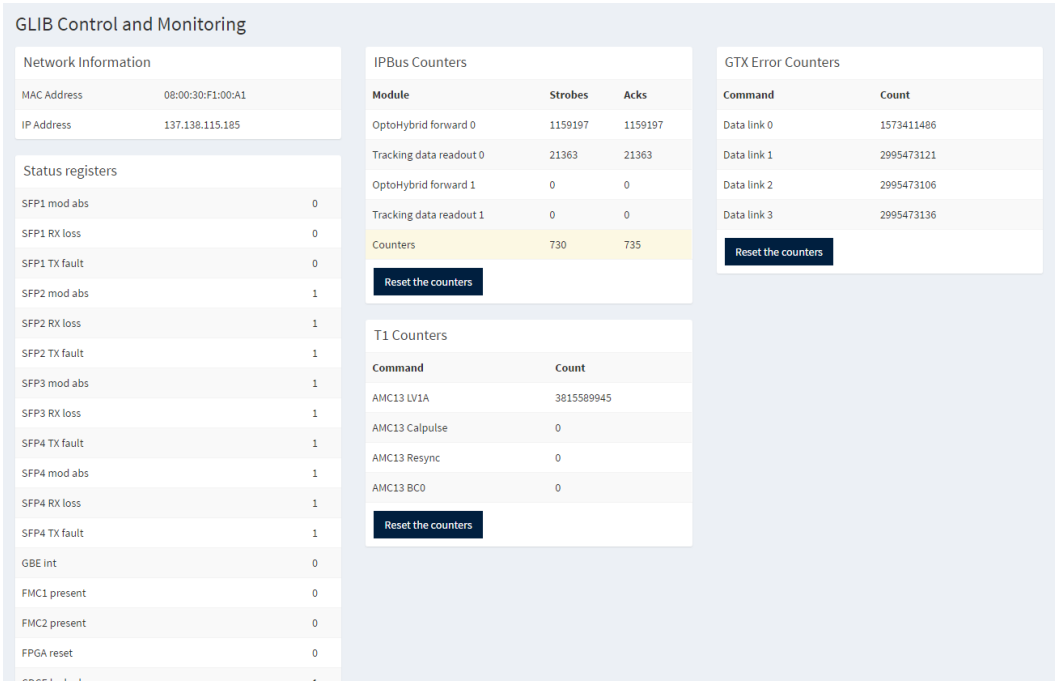

Figure 6.16 – Screenshots of the page used to control and monitor the GLIB.

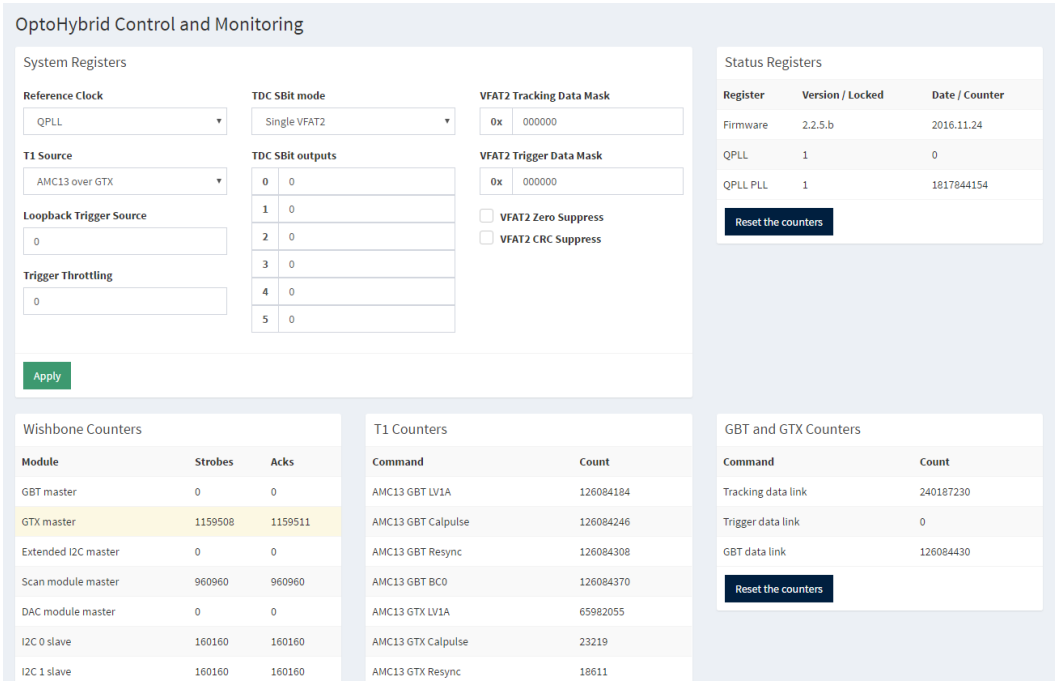

Figure 6.17 – Screenshots of the page used to control and monitor the OptoHybrid.

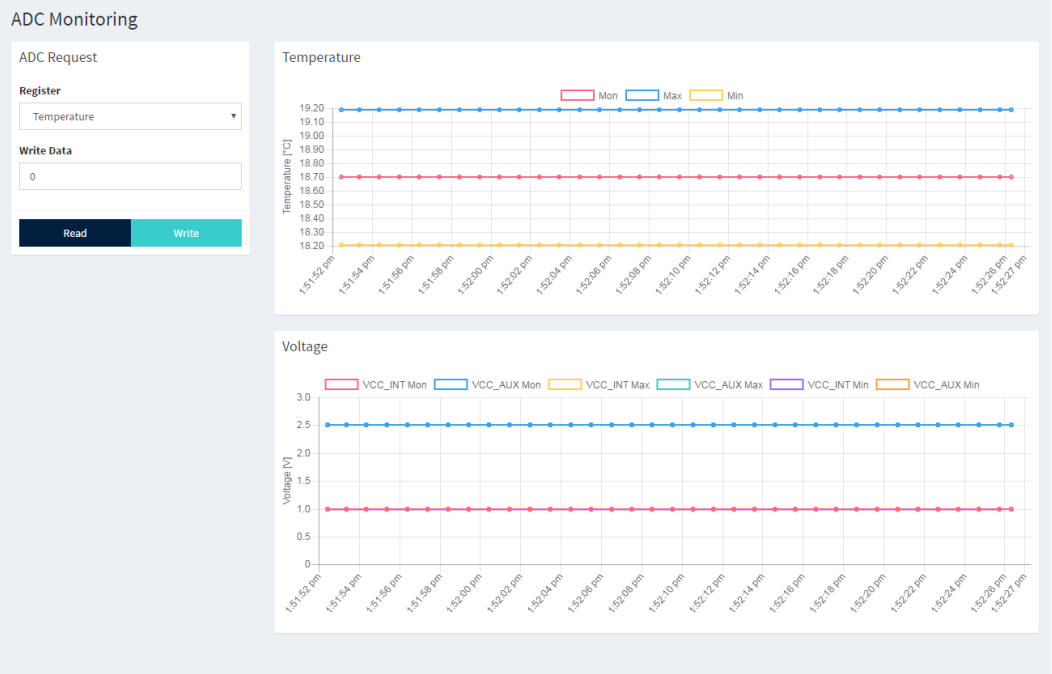

Figure 6.18 – Screenshots of the page used to monitor the temperature and voltages of the OptoHybrid.

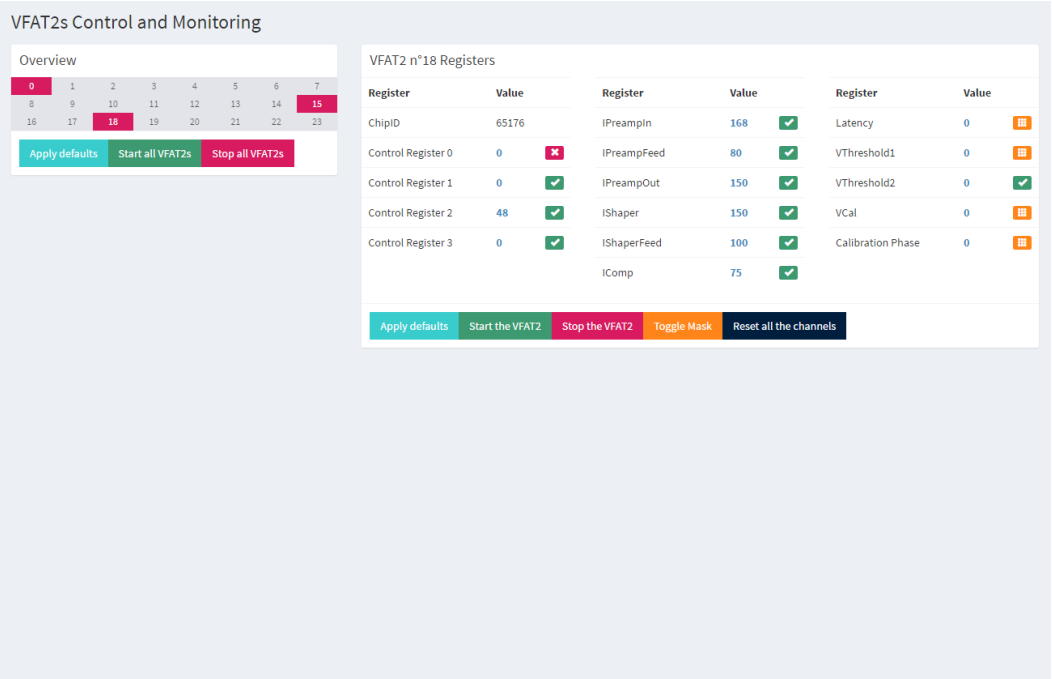

Figure 6.19 – Screenshots of the page used to control and monitor the VFAT2.

Finally, the control page of the VFAT2s, shown in Figure 6.19, gives access to the most significant registers of the front-end chips. It provides a summary of the on/off/absent status of each VFAT2 on the GEB and, by clicking on an individual slot, the value of each register as shown in the panel on the right. Default settings for the registers are saved in memory and compared to the current values, raising warnings when they differ and allowing the user to correct them. The user can also control all the VFAT2s at once by applying the default settings or turning them on or off at the same time.

6.3.3 Calibrating the VFAT2s

The calibration of the VFAT2s is done through a dedicated page which regroups the various operations that can be performed on the chip. Figure 6.20 is a screenshot of the application running a threshold scan on one of the VFAT2s. The left panel allows the user to select all the required parameters such as the type of scan, the range of the scan, the number of events per point, etc. Once the parameters are entered, the user starts the scan and waits for it to finish in firmware. When done, a graph displays the results in the right panel. The user can then dynamically change VFAT2 settings by clicking on a point of the graph to, for example, select the most appropriate threshold or latency value. Furthermore, the user can save the data points to disk for offline analysis.

6.3.4 Fast and Slow Control of the VFAT2s

The fast control or T1 generator is controlled by a dedicated page, displayed on the left of Figure 6.21, which allows to select the type of commands to generate, the pattern to send, and the delay between two pulses. The slow control of the VFAT2s handled by the I2C modules and the extended I2C controller is managed by another page shown on the right. The slow control commands can either target a single VFAT2 as shown in the left panel or an entire set of chips as can be seen in the right panel. In case of a single command, the response is shown inline. For multiple operations, a table of results is shown indicating the value and the corresponding VFAT2. To help the user, the accessible registers are named rather than listed using their address.

6.3.5 Data Quality Monitoring

DQM is used online to check the consistency of the recorded data. This is the last component of the application shown in Figure 6.22. This page monitors the status of the readout buffers of the GLIB and allows the user to acquire data. Once acquisition has started, plots begin to fill with the various parameters extracted from the VFAT2 data format. The user can visualize the number of data packets, BC, EC, flags, and chip ID distribution, the occupancy of the channels, etc, effectively monitoring the validity of the data and if required change the parameters of the run. Although this application has the ability to readout data and store it to disk, this feature was not used to this intend during the test beam. A program developed by the software

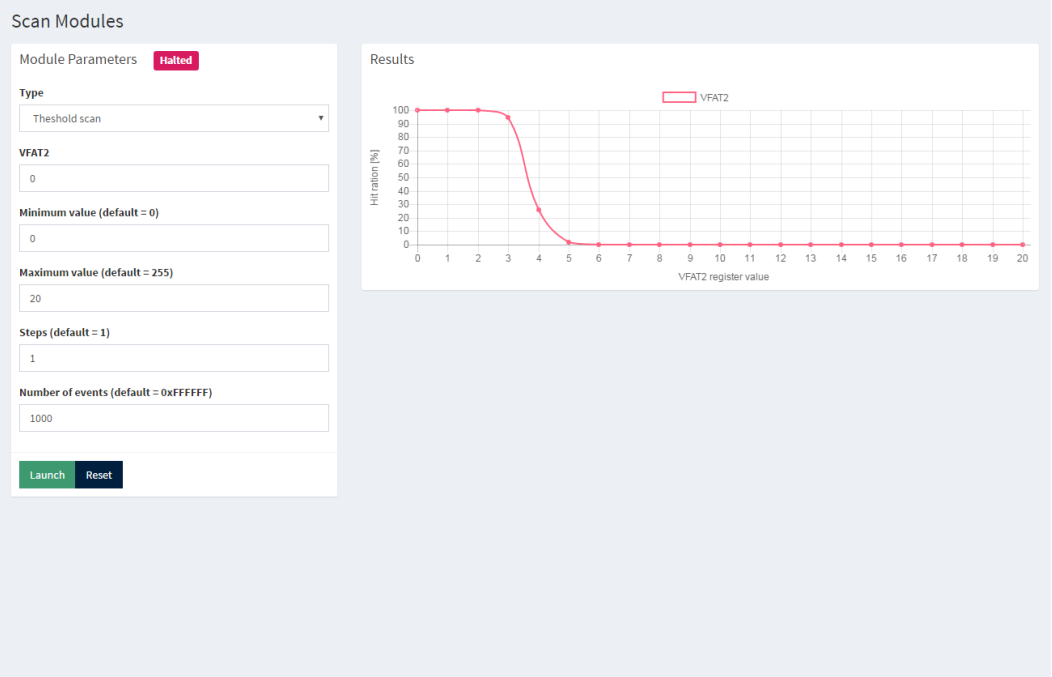

Figure 6.20 – Screenshot of the page of the web application used to perform calibration scans of the VFAT2s.

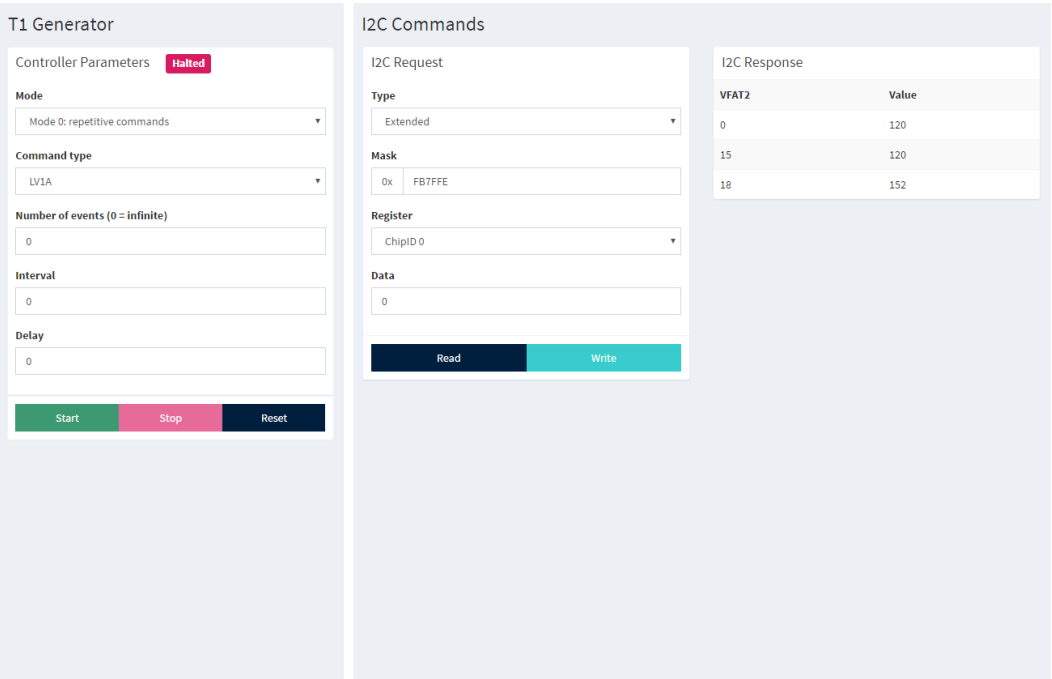

Figure 6.21 – Screenshots of the pages of the web application used to send fast commands (left) and slow control (right) requests to the VFAT2s.

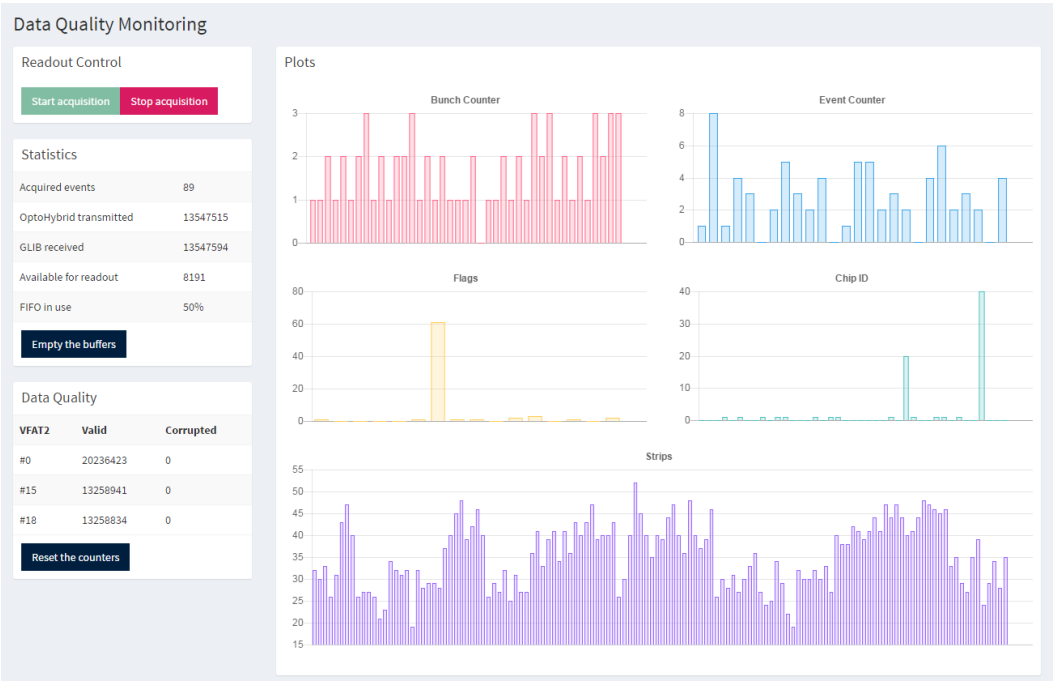

Figure 6.22 – Screenshot of the page of the web application used to do online data quality monitoring.

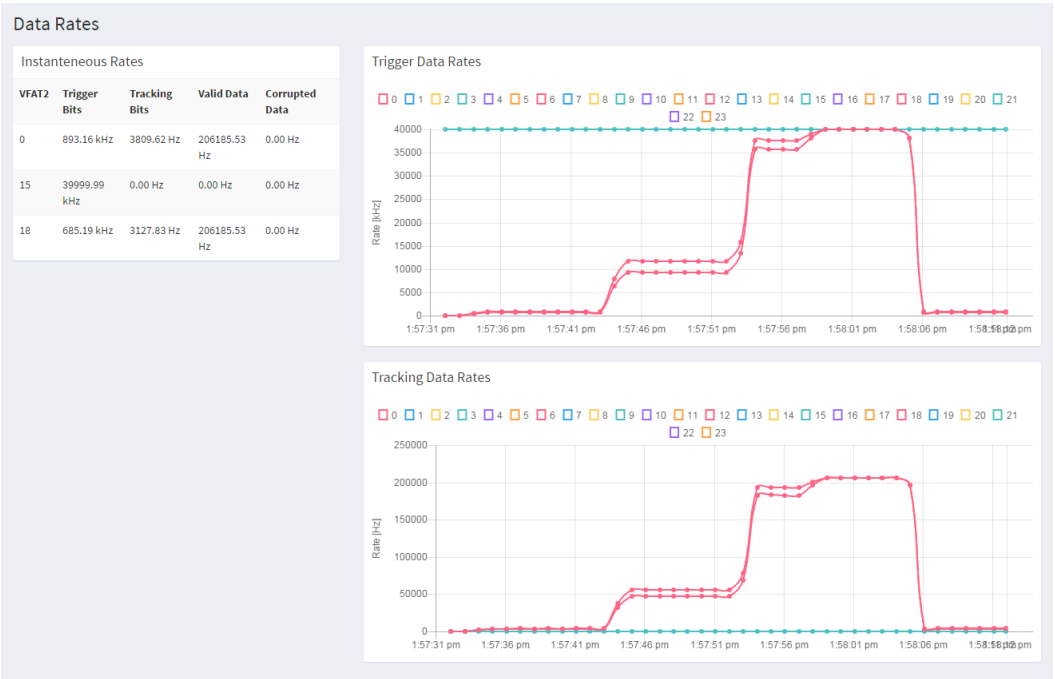

Figure 6.23 – Screenshot of the page of the web application used to monitor the hit rates.

development team of the GEM collaboration was used to acquire raw data from the GLIB and store it on disk in text files for later analysis.

6.3.6 Hit Rate Monitoring

Next to DQM functionalities, the firmware also monitors the hit rate in the trigger and in tracking data forwarded by the VFAT2s. Figure 6.23 is a screenshot of the page of the web application used to monitor these rates. For every connected VFAT2, the application logs the rate of hits in the trigger and tracking data, and the rate of tracking data packets that are received with a valid and corrupted CRC.

6.4 The Test Beam Setup

The test beam campaigns took place at CERN in the North Area of Prevessin, one of the secondary sites. The beam provided to the experiments originates from the SPS proton bunches which are extracted from the accelerator at 400 GeV and interact with a 30-cm-thick Beryllium target to produce the secondary beams mainly composed of electrons and hadrons (protons, pions, and kaons). A selection on the type of particles and their momentum is performed using so-called wobbling magnets that bend the beams and redirect them to the experimental setups. In the transfer tunnels, collimators can be adjusted to increase or reduce the flux of particles and an optional secondary target can be placed in the beam path to create a tertiary beam. Figure 6.24 shows the monitoring page of the SPS indicating the presence of spills of particles in yellow. Spills last between 4.8 s and 9.6 s and repeat every 14 s to 48 s depending on the SPS operation.

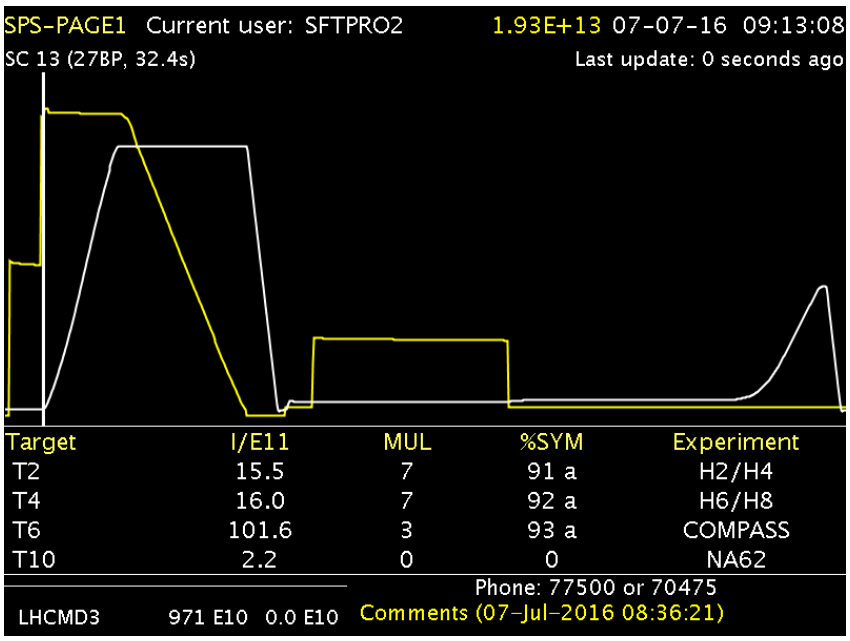

Figure 6.24 – Monitoring page of the SPS indicating the presence of spills of particles (yellow).

The two types of particles used during the test beam were 150 GeV pions and muons. Pions are obtained from the secondary beam by inserting a thin absorber which removes the electronic component of the beam. They can reach intensities between 10^4 and 10^7 particles per spill and produce a very narrow beam of around 5 cm in diameter. From their decay, a muon beam can be obtained by using moderate intensity secondary beams and closing the collimators. The muon beam reaching the experimental setup is wider, with a diameter around 14 cm, and with an intensity of 1% of the incident proton beam.

6.4.1 The GEM Detectors Setup

During the November 2015 test beam campaign, two GE1/1-V detectors have been used to collect data. The first detector in the beam will be referred to as GEM0 and the second as GEM1. Each detector was equipped with 12 VFAT2 Hybrids mounted in the central four rows. The remaining slots, too far away from the beam spot, were covered with grounding equipment to reduce noise. Each GEM was mounted with a GEB v2 and a OptoHybrid v2a. A picture of the setup is shown in Figure 6.25. Both OptoHybrids were connected to the same GLIB placed in a microTCA crate.

The GEM detectors were connected to a gas system providing a mixture of $Ar/CO_2/CF_4$ at 45%, 15%, and 40% respectively. The high-voltage was delivered through a ceramic high-voltage divider for one chamber and by separate channels connected directly to the GEM foils and planes for the other. The high-voltage applied is expressed in terms of current flowing through the divider which, by summing up the individual resistors that it is made of, yields a value in Volts. Table 6.4 enumerates

Current (μA)	Divider (V)	Drift plane (V)	GEM1 top (V)	GEM1 bottom (V)	GEM2 top (V)	GEM2 bottom (V)	GEM3 top (V)	GEM3 bottom (V)
700	3921	3290	2503	2107	1801	1416	805	437
710	3977	3337	2538	2137	1827	1437	816	443
720	4033	3384	2574	2167	1853	1457	828	450
730	4089	3431	2610	2198	1879	1477	839	456
740	4145	3478	2646	2228	1904	1497	851	462
750	4201	3525	2682	2258	1930	1518	862	468
760	4257	3572	2717	2288	1956	1538	874	475
770	4313	3619	2753	2318	1981	1558	885	481
780	4369	3666	2789	2348	2007	1578	897	487
790	4425	3713	2825	2378	2033	1598	908	493
800	4481	3760	2860	2408	2059	1619	820	500
810	4537	3807	2896	2438	2084	1639	931	506

Table 6.4 – Values of the voltages applied to the detectors according to the current flowing through the high voltage resistive divider.

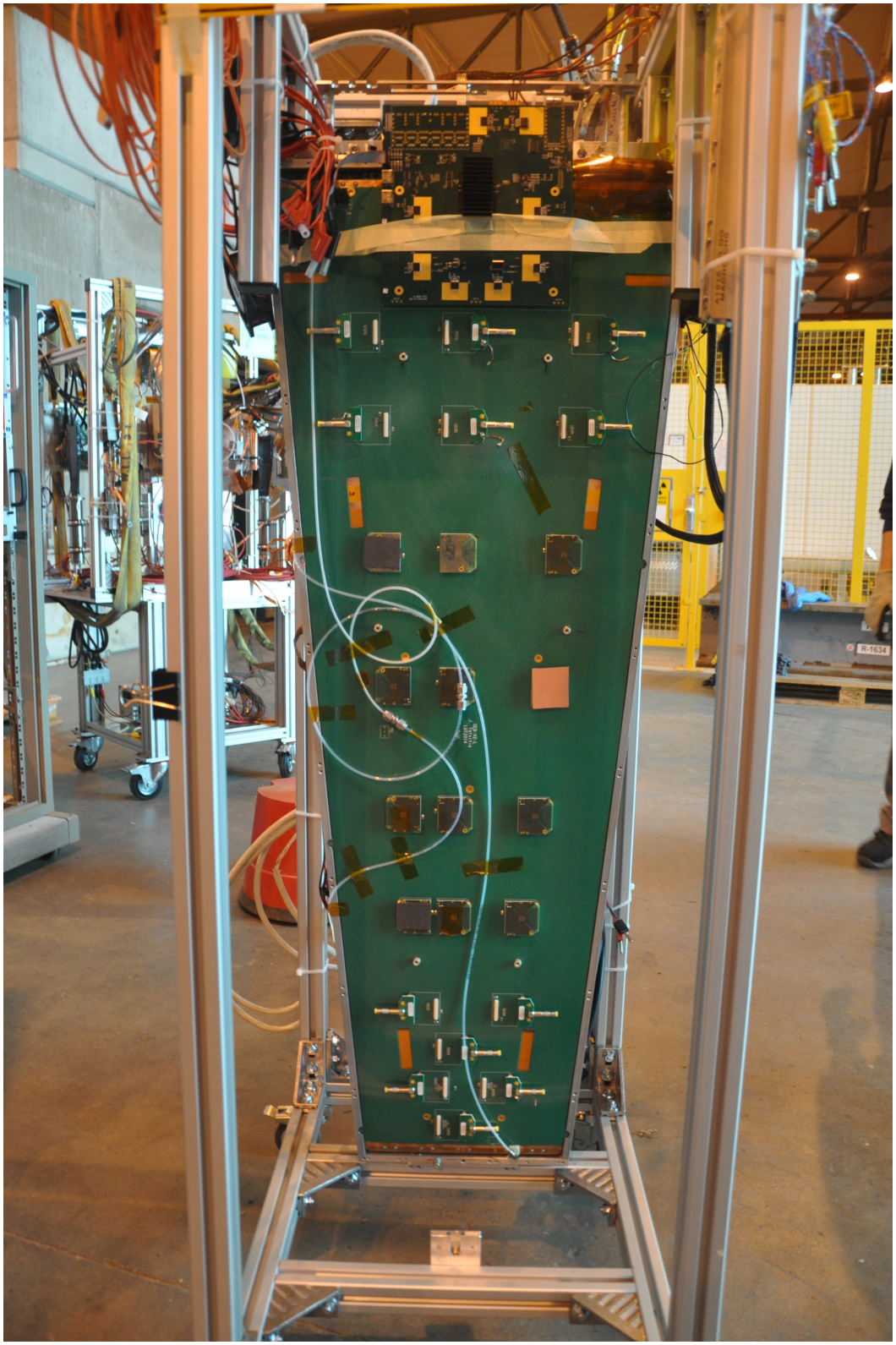

Figure 6.25 – Photograph of the GEM detectors used during the test beam and equipped with the full electronics system.

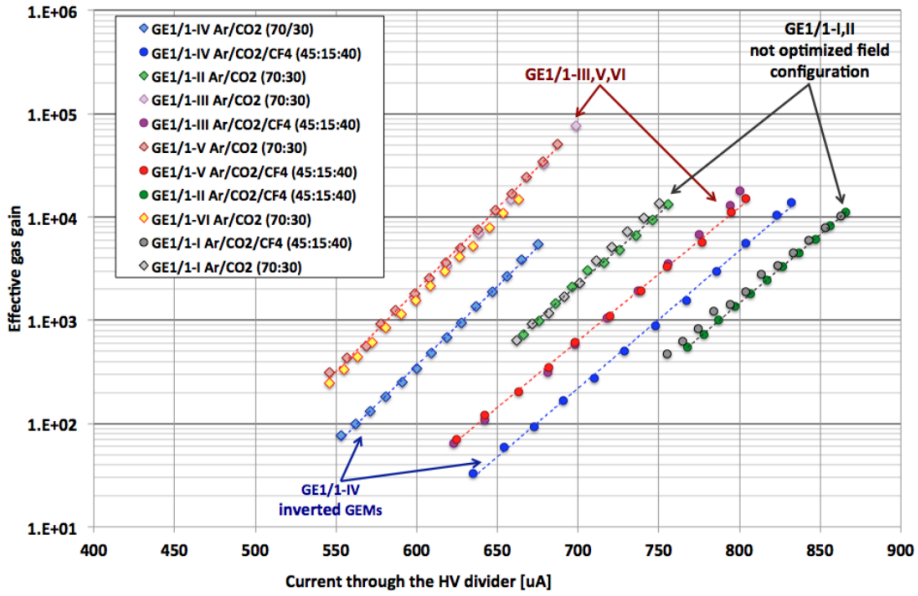

Figure 6.26 – Measured gas gains as a function of the current through the high voltage divider for the different generations of GEM detectors and for Ar/CO_2 70:30 and Ar/CO_2/CF_4 45:15:40 [46].

the different values of the voltages according to the current flowing through the divider. The difference between the divider and drift plane values is due to an input filter placed before the latter which smooths out fluctuations in the voltage. The low voltages for the OptoHybrids (1.8 V and 4 V) and for the GEBs (2.5 V) were delivered by two power boards mounted on top of the detector stand, along with the necessary tools to reprogram the FPGA in case of failure and the debugging headers.

The high voltage applied on the drift plane can be converted to an effective gas gain for the Triple-GEM detector. Figure 6.26 shows the measured gas gains as a function of the current through the high voltage divider for the different generations of GEM detectors and for Ar/CO_2 70:30 and Ar/CO_2/CF_4 45:15:40.

6.4.2 The Beam Area Setup

The GE1/1 detectors were inserted in the beam path and adjusted so that only one VFAT2 in the middle column would be exposed to the beam spot. Figure 6.27 shows the setup on the left placed in front of the beam transfer tunnel on the right side. Four scintillators, three placed in front of the detector and visible in the picture, and one placed in the back of the detector were used to provide triggers to the system. The output of the four scintillators was converted to Nuclear Instrumentation Module (NIM) logic and their coincidence used as the asynchronous trigger source of the system.

NIM uses current driven logic levels which produce -0.8 V over 50 Ω for a logic '1' and 0 V for a logic '0'. As the OptoHybrids use TTL voltage driven levels, namely 2.5

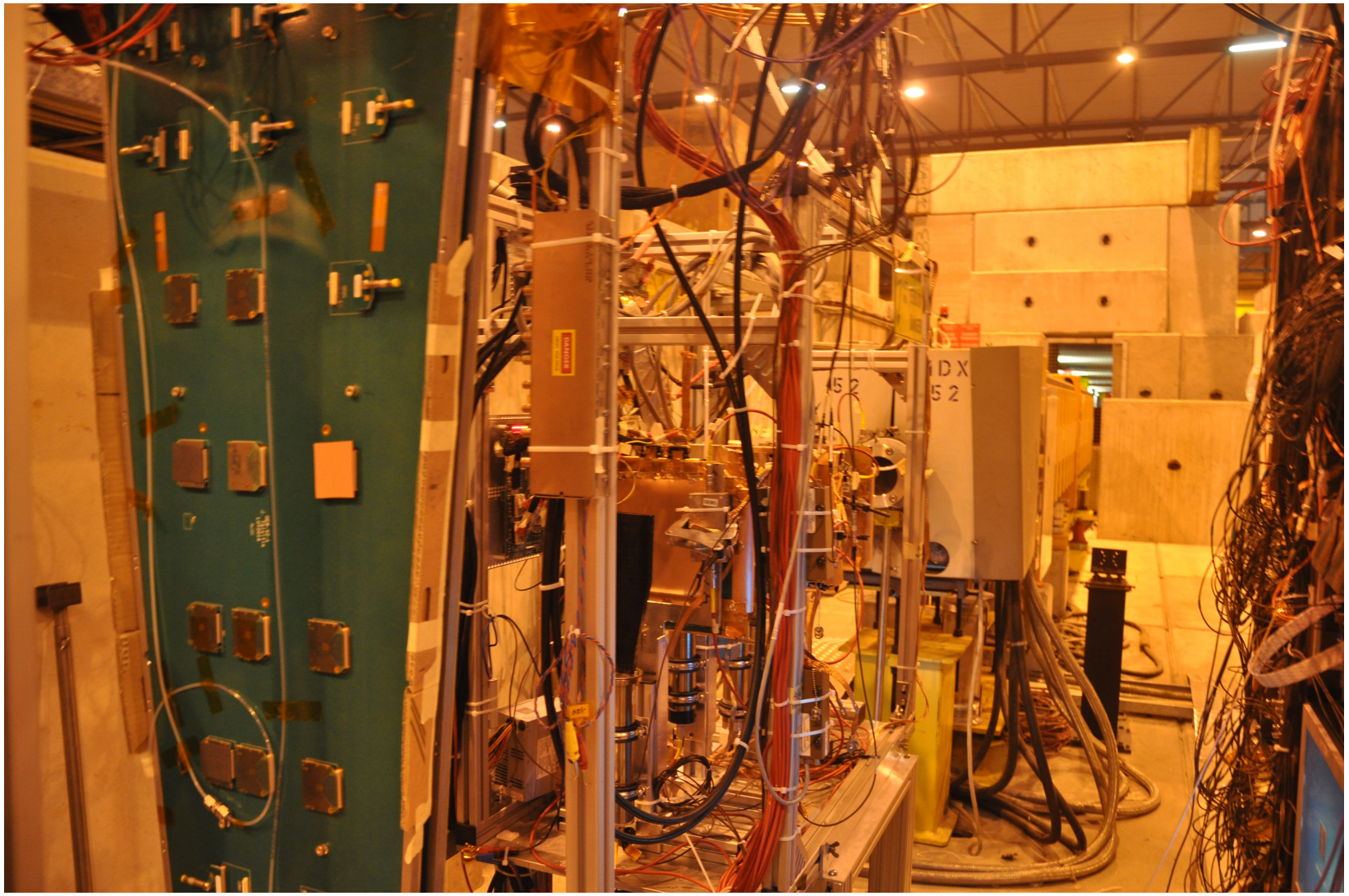

Figure 6.27 – Photograph of the two detectors on the left placed in front of the beam transfer tunnel on the right side.

V for a logic '1' and 0 V for a logic '0', a converter had to be used. It was observed that the output levels of the NIM to TTL module were 0 V - 1.25 V instead of 0 V - 2.5 V. This is close to the switching voltage of 1.2 V and might affect the number of triggers seen by the OptoHybrids, as logic '1's might be discarded. It was computed using the offline data to count the number of missing events that this situation occurred in less than 1/10 000 events and could be solved by the offline analysis. Using other sources of triggers, it was ensured the source of the problem did not originate from the DAQ system itself, which was proven not to be the case. The problem came from the NIM to TTL converter and not the electronics of the GEM detectors.

6.5 Data Analysis

The data collected during the test beam was analyzed to study the performance of the detector according to various parameters. Data was taken at multiple high-voltage values, VFAT2 threshold settings, and trigger rates. An analysis framework has been developed in order to combine the readout data from the two GE1/1 chambers. Alignment of the events was non-trivial and based on the difference between two consecutive BC values of the events. Indeed, due to the problem exposed here above with the NIM to TTL converter the number of events recorded by the two detectors might not be equal. Furthermore, the fact that the two VFAT2s are not guaranteed to have the same value of the BC counter at the same time prevented the framework from using the BC directly. Therefore, the difference between the BCs of two consecutive events in the data stream was computed and matched between the two datasets. When a discrepancy arose, one of the datasets had to be realigned. Once realignment was performed, analysis tools had access to both chambers for each event. This issue will not occur in CMS as the triggers will be distributed by the AMC13 and as all the detectors will be synchronized. It will thus be possible to reset all the VFAT2s at the same time, ensuring that the BC is the same.

6.5.1 VFAT2 Threshold Scans

The first parameter to study is the noise on the VFAT2s induced by the system and by the effect of the channels. To this end, threshold scans have been produced using the web application. These scans are ran when the beam is not present. They use the trigger bits of the VFAT2s which do not need a L1A source to be produced. These are clocked at 40 MHz and for each threshold value the ratio of events containing hits against the total amount of collected events is defined to be the noise. Figure 6.28 plots the noise level as a function of the VFAT2 threshold in terms of VFAT2 units, the decimal value written in the eight bits register afterwards converted to a voltage level, for GEM0 on the left and GEM1 on the right. Although the noise level quickly diminishes, it never reaches the expected value of 0 even at very high threshold. Typically, previous test beams have shown that a value of 15 eliminates the noise. Furthermore, it can be seen that noise level is slightly higher in GEM1, which reaches 20% at a threshold value of 10 against 15% for GEM0. From the

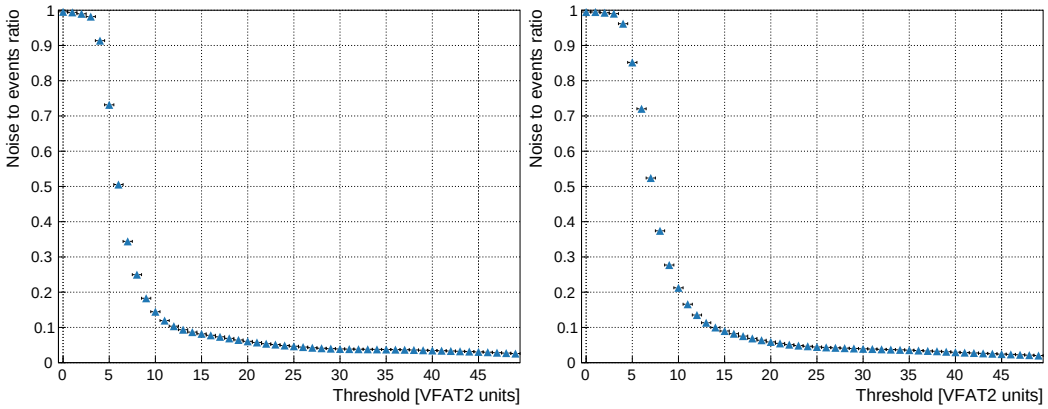

Figure 6.28 – Plots of the noise level as a function of the VFAT2 threshold in terms of VFAT2 units, the decimal value written in the 8 bit register afterwards converted to a voltage level, for GEM0 on the left and GEM1 on the right.

threshold scans, it was decided to run both chambers at a VFAT2 threshold level of 25 for data taking runs. This value is considered to be the default threshold except if otherwise specified.

After the end of the test beam, the noise issue has been investigated thoroughly and two sources have been isolated: the I2C clock used for the slow communication and the LEMO cables used for the trigger input. In the version of the firmware used during the test beam, the I2C clock ran continuously which induced noise on the trigger bits of the VFAT2s at a frequency of 100 kHz, the same as the clock itself. This problem has been solved by activating the I2C clock solely when a slow control operation takes place, limiting the number of corrupted trigger bits. The second issue originated from current loops formed between the low power cables of the OptoHybrids and the LEMO cable connecting the system to the NIM crate. The ground of the latter was connected to the OptoHybrid ground through the shielding of the LEMO cables. Once the two ground systems were isolated, noise

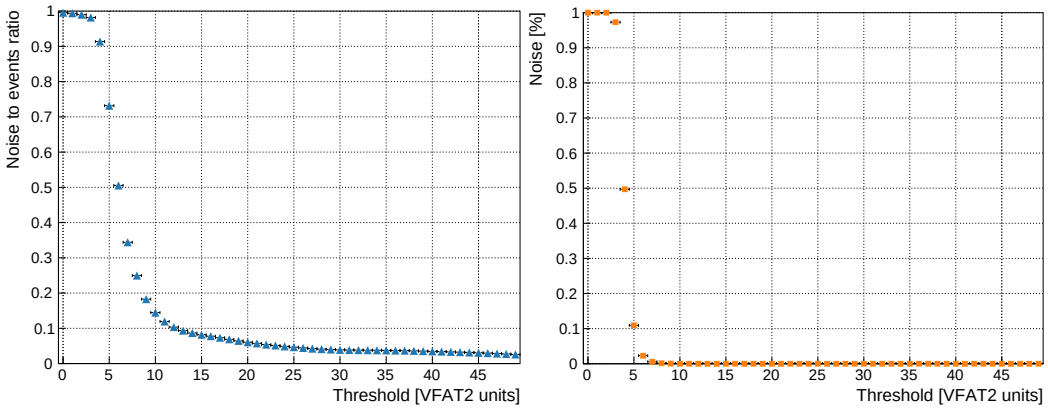

Figure 6.29 – Plots of the noise level as a function of the VFAT2 threshold in terms of VFAT2 units with results obtained during the test beam on the left and with the improved noise cancelation on the right.

level comparable to previous observations were obtained. Figure 6.29 plots the noise level as a function of the VFAT2 threshold in terms of VFAT2 units with results obtained during the test beam on the left and with the improved noise cancelation on the right. The reduction in noise in turn allowed to function at a lower threshold and thus increase the efficiency of the detector, reaching the required 98% mark.

It was also observed that GEM1 experienced a leakage current through the GEM foils of about 10 μA. This had as consequence to reduce the voltage on the GEM foils and thus decrease the gain and efficiency of the detector. The detector also suffered from water condensation inside the gas pipes which caused some problems during the first days of run. These issues can explain the difference in noise observed between the two chambers.

6.5.2 VFAT2 Latency Scans

Once the threshold on the VFAT2s has been set and the beam turned on, a latency scan must be done in order to capture the events corresponding to the triggers coming from the scintillators. Again, the web application is used to run the scan on the VFAT2s. Similarly to the threshold scan, the latency scan counts the ratio of events with a hit against the total number of events. However, it uses the tracking data provided by the VFAT2s and not the trigger bits. Figure 6.30 plots the ratio of hit events as a function of the VFAT2 latency in terms of BXs for a monostable pulse length of four clock cycles for GEM0 on the left and GEM1 on the right with muons. When the latency is too low or too high, the VFAT2 misses the hit and a baseline value of 0.05 is observed corresponding to noise. Once the latency enters the correct window, it rises to a value close to 0.98. These plots illustrate the usefulness of the settable monostable pulse length. Indeed, the two points at approximately 0.5 and 0.6 indicate that hits are falling in two adjacent latency values, each collecting one half of the events. This is due to the asynchronous trigger coming from the PMs. In case of a non-adjustable setting, the maximum efficiency of the system would be reduced by half. From these results, a VFAT2 latency value of 11 has been used for data taking runs. This value is considered to be the default latency except if otherwise specified.

Latency scans are also used to measure the noise and the efficiency of the detector. When looking at out-of-time events, the only contribution to the hits is noise. On the other hand, when considering the window in which the latency is well adjusted, both the noise and the particle hits play a role in the measurement. Taking this value to be the measured efficiency, it corresponds to the ratio between the number of events detected by the GEM detector and those detected by the scintillator system. However, the measured efficiency is the superposition of two Poisson processes: the noise and the real detector efficiency. In order to obtain the computed or real efficiency, simulations were developed to compute upper and lower limits on the real

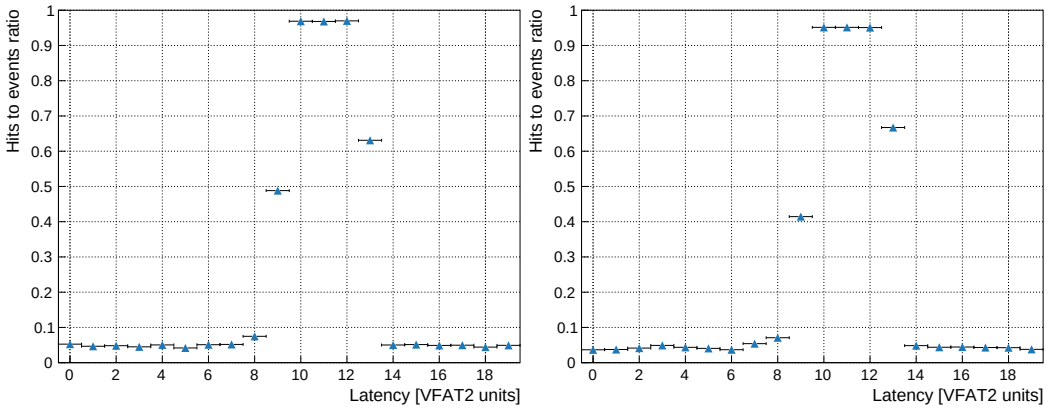

Figure 6.30 – Plots of the ratio of hit events as a function of the VFAT2 latency in terms of BXs for a monostable pulse length of four clock cycles for GEM0 on the left and GEM1 on the right with muons.

efficiency according to the noise levels. This was achieved by running a thousand pseudo-experiments for each possible efficiency value and comparing the result to the measured value. With a noise level of 100%, the real efficiency could be anything between 0% and 100%. However, with a noise level of 1%, the computed efficiency will be close to the measured efficiency. For example, for a measured efficiency of 95% and noise levels of 50% and 20%, the resulting computed efficiency is within a range of 89% - 91% and 93% - 94% respectively. When the noise is lower, the computed efficiency is close to the measured value as the probability of noise faking a hit is less.

6.5.3 Evolution of the Efficiency with the High-Voltage

Using the latency scan to compute noise and efficiency values, the evolution of these parameters has been studied according to the high-voltage applied on the detector. Figure 6.31 plots the measured noise (green), measured efficiency (blue), and computed efficiency (orange) as a function of the high-voltage expressed in microamperes for GEM0 on the left and GEM1 on the right with muons. Both results show an increase in efficiency with the high-voltage, reaching the start of a plateau around 97%. This is caused by the higher gain at which the GEM foils start to operate which in turn amplifies the signals and allows it to be detected by the VFAT2 ASIC. It can be noted that a shift exists between the results for GEM0 and GEM1. Values for the latter are shifted to the left by approximately 10 μA. While GEM0 reaches a plateau at 800 μA, GEM1 requires 810 μA to reach an efficiency around 97%. This is due to the leakage current observed during installation which decreases the voltage applied to the GEM foils and in turn performance. It can also be noted that the noise is constant with the high-voltage and remains at 4%. From this, it was decided to operate both chambers at 800 μA as sparks were sometimes observed at 810 μA. This value is considered to be the default high-voltage except if otherwise specified.

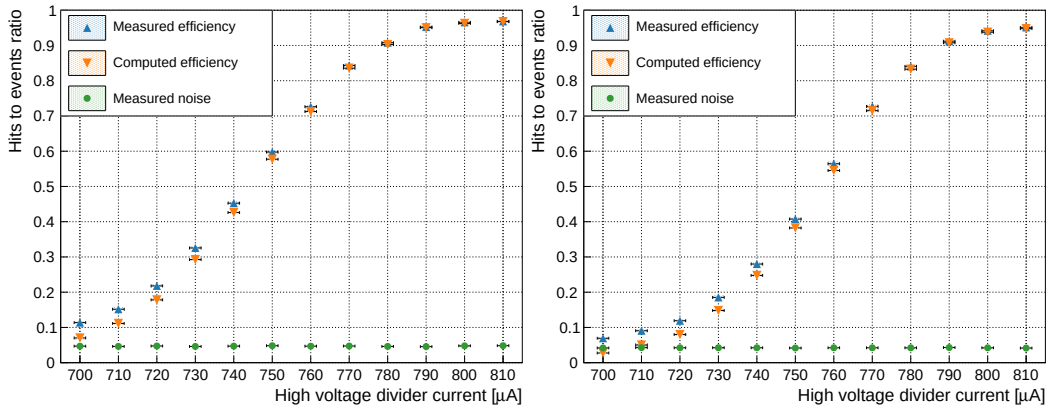

Figure 6.31 – Plots of the measured noise (green), measured efficiency (blue), and computed efficiency (orange) as a function of the high-voltage expressed in μA for GEM0 on the left and GEM1 on the right with muons.

The results obtained with GEM0 can be compared against the scans performed during previous test beams [47, 48] which used the same gas mixture. These scans show the efficiency in terms of gain, which is then expressed in terms of current through the high voltage divider. In the 700 μA to 810 μA range, gains of 100 to 8 000 are obtained. Between these values, the efficiency quickly rises to 97% to reach the very beginning of the plateau. The same behavior is observed in the results exposed here-above. To convert the high voltage from microamperes to a gain, Figure 6.26 is used. The latter yields a range for the gain of 500 to 10 000 in which the efficiency reaches its maximum at the highest values, right at the beginning of the plateau region. A third test beam [49] provides measurements of the efficiency as a function of the high voltage for lower thresholds of 10, 15, and 20. While they reach the efficiency plateau at lower currents, it sits at a value of 95%, compared to 97% for the results showed here-above.

6.5.4 Evolution of the Efficiency with the VFAT2 Threshold

Although it had been decided to run the system with a VFAT2 threshold of 25, a study of the evolution of the noise and efficiency according to the threshold was performed. It aimed at quantifying the amount of noise and signal that was cut when increasing its value allowing to select the configuration yielding the best signal over noise ratio. This procedure uses tracking data and the full granularity information to analyze events. Figure 6.32 plots the measured noise (green), measured efficiency (blue), and computed efficiency (orange) as a function of the VFAT2 threshold in terms of VFAT2 units for GEM0 on the left and GEM1 on the right with muons. Additional points in purple have been extracted from the tracking data. Using the two GEM detectors, it is possible to measure the efficiency of one using the other as reference. For each event, the following cuts were applied on the GEM of reference in order to reduce the dataset:

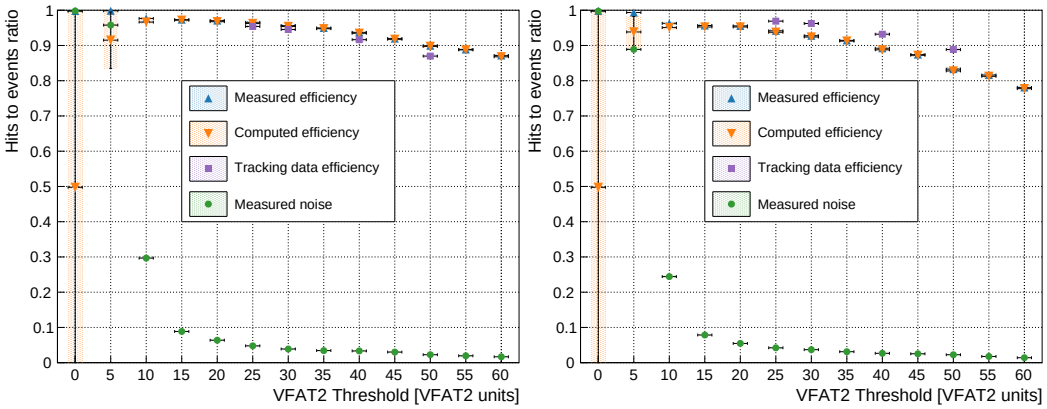

Figure 6.32 – Plots of the measured noise (green), measured efficiency (blue), computed efficiency (orange), and tracking data based efficiency (purple) as a function of the VFAT2 threshold in terms of VFAT2 units for GEM0 on the left and GEM1 on the right with muons.

- only one cluster, group of adjacent channels that have been hit, is present in the event;
- the sole cluster must contain exactly two active channels;
- the center of the cluster must be at least ten channels away from the sides of the chip.

The first condition eliminates events where a hit is accompanied by noise, creating an ambiguity on the cluster associated with the particles. The second condition helps to remove noise related events as they have a cluster size distribution peaked at one channel. Therefore, combining these two criteria creates a stringent condition for the selection of signals generated by particles. The final condition ensures that the hit in the second GEM detector will be located inside the VFAT2 thus eliminating acceptance problems. Using the reduced dataset, the algorithm computed the efficiency by looking for hits in the second GEM detector in a window of ten channels around the center of the cluster of the GEM of reference, which corresponds to a window of 40 mm.

It can be noted that at very low thresholds the computed efficiency displays a large error due to the high noise. As the latter decreases with the increase in the threshold, the algorithm computing the real efficiency is able to tighten its limits ultimately providing a computed efficiency close to the measured value. From a threshold value of 15, efficiency starts to slowly drop from 98% to 90% at 50. This shows how the threshold cuts the signal while only slightly affecting noise which remains at 4% at thresholds above 25. From these results, the choice to operate the system at a threshold of 25 is validated. Although displaying a slightly lower efficiency, still around 97%, than results at a threshold of 15, the lower noise is non-negligible and reaches a local plateau. The deviation between the efficiency obtained using the

latency scans and the tracking data for GEM1 comes from the higher noise to which it is subject. Apart from the requirement that the hits in GEM1 should be within a given window around the hit in GEM0, no other cut is performed on the data in order to not bias the results. Therefore, the higher noise in GEM1 will artificially increase the efficiency when faking a hit.

6.5.5 Evolution of the Efficiency with the Trigger Rate

Similarly to the measurements and computations made for the previous study, the effect of trigger rate on the efficiency was measured. Focus is given on the latter as the noise levels are not influenced by the increase in rate and remain at 4%. Figure 6.33 plots the measured efficiency (blue), computed efficiency (orange), and tracking data based efficiency (purple) as a function of the trigger rates for GEM0 on the left and GEM1 on the right with pions. The trigger rate was adjusted using collimaters placed in front of the beam setup. As noted in the results shown in the previous section, the deviation between the efficiency obtained using the latency scans and the tracking data for GEM1 comes from the higher noise to which it was subject.

A small decrease in efficiency of 1% for GEM0 and of roughly 2% for GEM1 from the initial value of 97% is observed when increasing the trigger rate from 20 kHz to 120 kHz. This is due to the charge-up effect that occurs inside the detector [50]. The electrons and ions created during the avalanche process inside the GEM foils can be absorbed by the Kapton insulator and in turn modify the transparency of the GEM foils. The transparency is defined as the ratio of primary electrons that initiate an avalanche and thus contribute to the formation of the signal, and the total number of primary electrons. Figure 6.34 illustrates the drift path (without diffusion) of electrons (orange) starting 90 μm above the GEM, in an electrostatic configuration (green) without charges on the Kapton surface on the left and with charges on the Kapton surface on the right. As can be seen in the right plot, a significant number

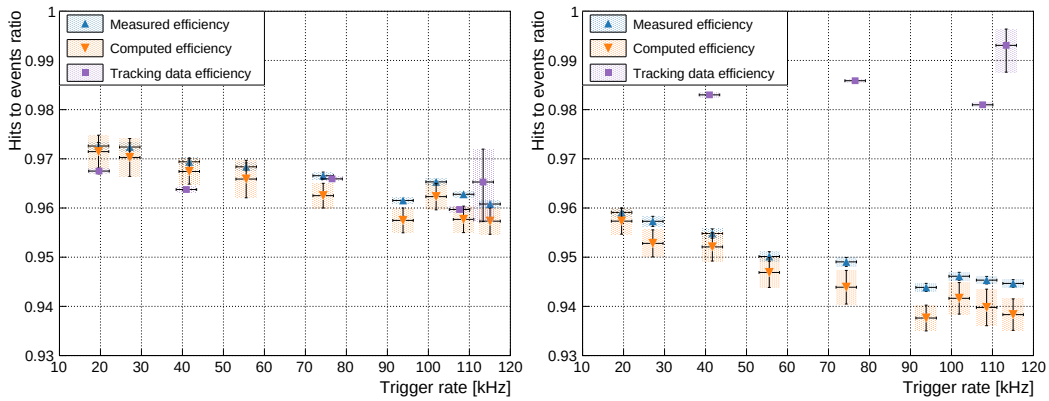

Figure 6.33 – Plots of the measured efficiency (blue), computed efficiency (orange), and tracking data based efficiency (purple) as a function of the trigger rates for GEM0 on the left and GEM1 on the right with pions.

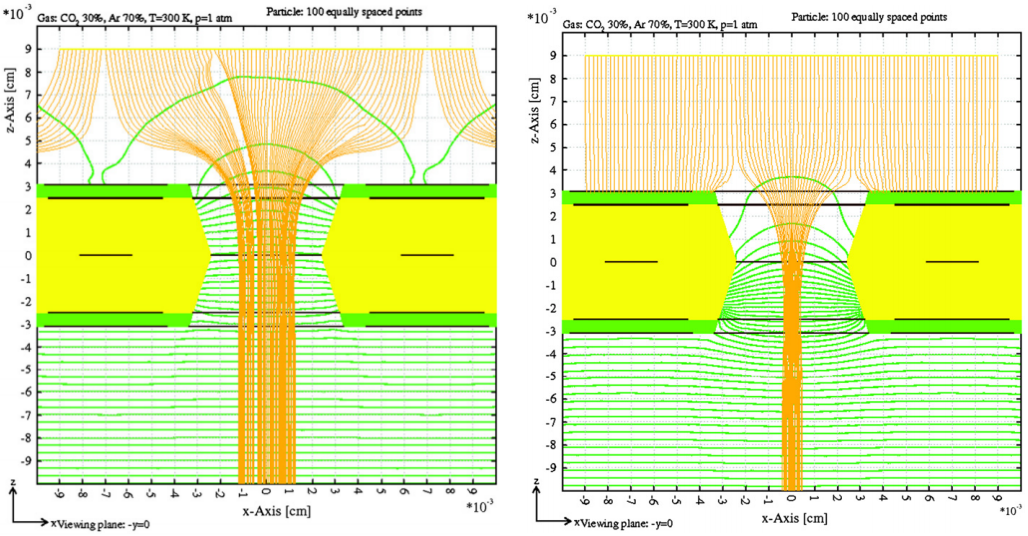

Figure 6.34 – Drift path (without diffusion) of electrons (orange) starting 90 μm above the GEM, in an electrostatic configuration (green) without charges on the Kapton surface on the left and with charges on the Kapton surface on the right [50].

of electrons are captured by the coper layer on top of the Kapton when the latter has accumulated charge, thus reducing the transparancy of the foil. This in turn lowers the amplitude of the signal formed on the readout strips [51] and increases the probability that the signal is bellow threshold, meaning the hit is not recorded by the VFAT2. When the counting rate in the chamber is small, the accumulated charge can dissipate, effectively eliminating this effect.

To corroborate the statement that the loss in efficiency originates from the detector and not the DAQ system, the latter has been tested against trigger rates of 200 kHz, the maximum rate supported by the VFAT2s. By providing a know number of triggers to the system and counting the number of data packets read out, it was shown that no data losses occur in the readout chain. These results support the claim that the GEM chamber and DAQ system can sustain trigger rate of CMS, which is of 100 kHz for the LHC Phase I, while keeping an efficiency above 96%. Finally, it is important to emphasize that the measurements of the efficiency against the threshold show that when lowering the working threshold of the VFAT2s to 15 instead of 25, the efficiency increases by 1% to 2%. This means a final efficiency above 97% could be achieve.

6.5.6 Cluster Multiplicity and Size

Two studies that can be performed offline using the recorded data are the evolution of the cluster multiplicity, the number of clusters, and the cluster size, the number of channels per cluster, with the VFAT2 threshold. Figures 6.35 and 6.36 respectively plot the cluster multiplicity and cluster size for muons (green) and pions (blue) as a function of the VFAT2 threshold in terms of VFAT2 units for GEM0 on the left and GEM1 on the right. The points and error bars respectively correspond to the mean

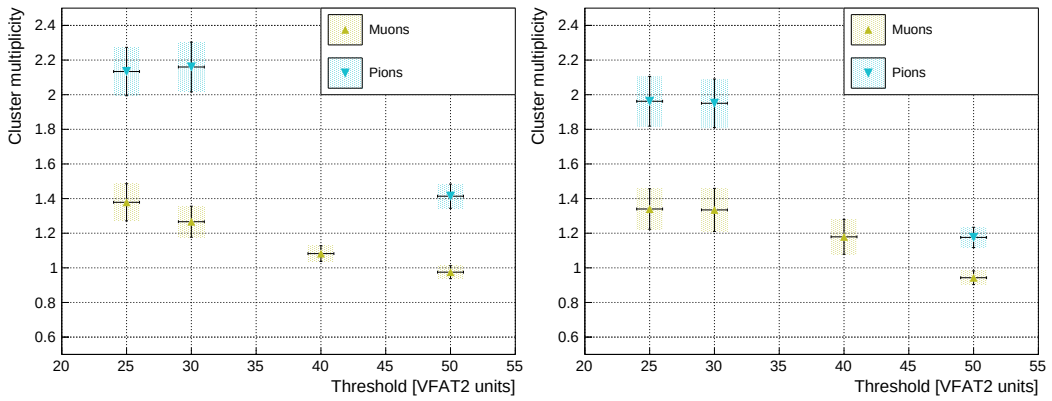

Figure 6.35 – Plots of the cluster multiplicity for muons (green) and pions (blue) as a function of the VFAT2 threshold in terms of VFAT2 units for GEM0 on the left and GEM1 on the right.

and RMS of the distributions.

The cluster multiplicity reflects the number of clusters present in the event, some of which are due to noise, others to signal. Its evolution is tightly linked to the evolution of the efficiency as a function of the VFAT2 threshold. It was observed that for threshold values of 25 and 30, the noise and efficiency were only slightly decreasing and remained at 4% and 97% respectively. At a higher threshold of 50 VFAT2 units however, the efficiency dropped to 90%. This explains that the points at 25 and 30 for both pions and muons in GEM0 and GEM1 have similar values ($\pm$ 0.1). When the threshold is further increased to 50, the effects start to appear, cutting the signal and removing clusters.

The impact of the efficiency according to the threshold is however not directly visible when considering the cluster size. It is clear that the cluster size decreases as the threshold increases, meaning that less and less channels are being fired. At low threshold, clusters contain roughly 1.5 channels and an average of 1.6 clusters are present per event. When increasing the threshold, the average number of channels fired decreases slightly which does not immediately affect the cluster multiplicity: clusters are reduced in size but not in numbers. The counting of cluster is not directly affected by the cluster size when the latter is greater than one. This means that two clusters of size two are equivalent to two clusters of size one. Even if by increasing the threshold, some channels are being deactivated, other channels in the cluster might still remain active thus leaving the cluster statistic unchanged.

Two test beam campaigns [48, 49] studied the distribution of the cluster size and provide a reference point against which the results here-above can be compared. These tests respectively yielded an average cluster size of 1.79 ± 1.16 for a mixed muon/pion beam at a threshold of 20 VFAT2 units, and 1.75 (no errors given) for a muon beam at a threshold of 12 VFAT2 units. The first measurement did not

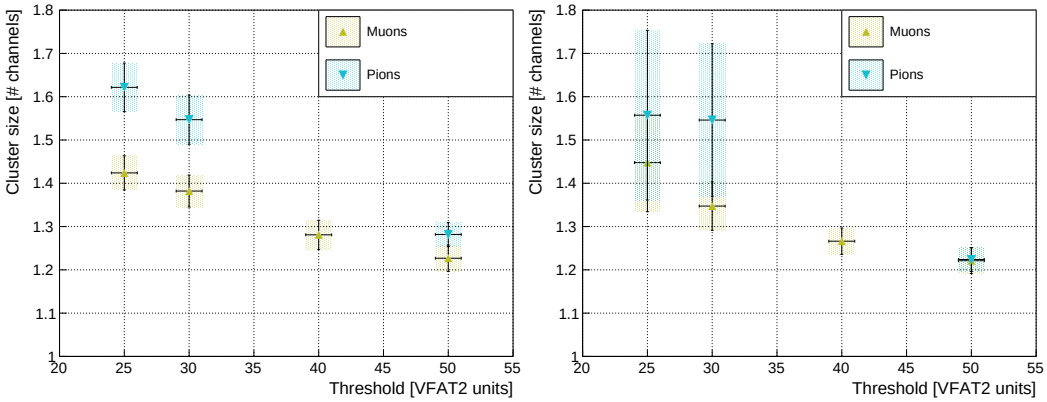

Figure 6.36 – Plots of the cluster size for muons (green) and pions (blue) as a function of the VFAT2 threshold in terms of VFAT2 units for GEM0 on the left and GEM1 on the right.

distinguish data taken during muon and pion runs and is thus an average of the two. Performing the same operation on the results shown above and assuming an equal number of events for both types of particles, an average value of 1.57 ± 1.1 is obtained. In order to compare the data against the second set of results, a quadratic extrapolation of the data points for the muons shown in the left plot of Figure 6.36 was done towards 12 VFAT2 units. The obtained cluster size is of 1.77 ± 0.4. In both cases, the cluster sizes obtained during the previous test beams and the results shown above are well within the error bars and show excellent agreement. Additional measurements have been performed in previous studies [52] at Fermilab using hadron beams. These yielded a cluster size greater than for muon beams at a value of 2.43 ± 1.04, which is also visible in the results above.

These results confirm the different behaviors between muons and pions: the former with a cluster size around 1.42 ± 0.5, and the latter around 1.62 ± 0.6. This originates from the higher interaction rate of the pion beam with the material placed in front of the GEM detectors, which causes the creation of secondary particles. As the beam is collimated and hits around 40 channels on the VFAT2, the clusters resulting from different particles overlap and artificially increase the cluster size. When clusters do not overlap, it is the cluster multiplicity which is increased.

6.5.7 Beam Profile

Finally, to illustrate the structure of the beam, one can create a beam profile obtained by superimposing the collected data and counting the number of hits for each channels. Figure 6.37 plots the beam profile for muons (green) and pions (blue) for GEM0 on the left and GEM1 on the right. The pion beam is well defined and centered on the VFAT2 as it is guided and collimated by the magnets of the transfer tunnels from the SPS to the experimental area. The muon beam on the other side displays a large spread due to the decay process of the pions and the collimators installed to stop them. The sharp cut-off on the left side of the profile is due to the

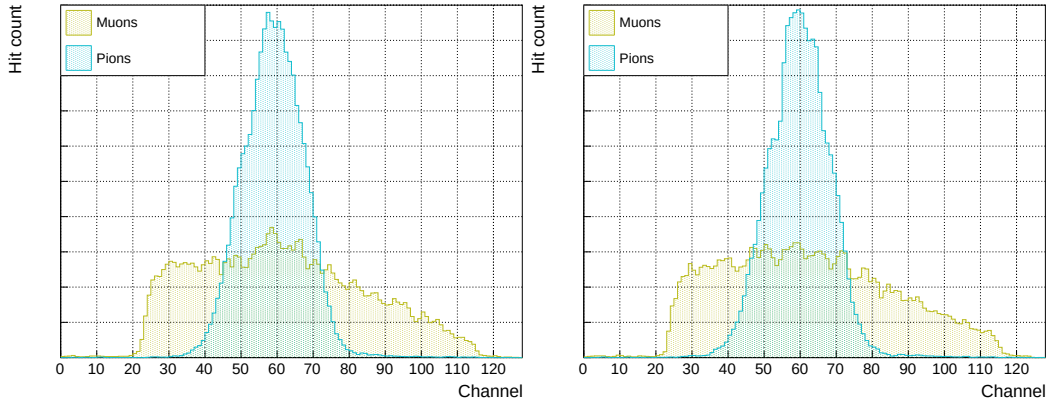

Figure 6.37 – Plots of the beam profile for muons (green) and pions (blue) for GEM0 on the left and GEM1 on the right.

geometrical acceptance of the PM placed in the back of the detector. These plots show that the chips do not have any dead channels.

6.6 Towards the Slice Test

Using the feedback obtained during the test beam, a OptoHybrid v2b was designed which incorporated the GBT chipsets (see Section 5.1.4). In addition to porting the code to the new board, firmware developments were required in order to make use of the latter and move from the GTX and SFP+ optical links to the GBT chipset coupled with the Versatile Link. The resources available in the FPGA and the flexibility of the Wishbone bus allowed for the two technologies to be implemented and control the system at the same time. This increases redundancy and offers the possibility to switch between the GTX and GBT in case of failure.

The GBT chipset is connected to the OptoHybrid through two E-Links for the receiver and four E-Links for the transmitter, all running at 320 MHz. This configuration allows to receive 16 bits and transmit 32 bits per BX. The asymmetry is due to the requirement of a higher bandwidth for the link to the off-detector electronics which carries the slow control and tracking data.

On the FPGA, the data is deserialized and aligned using a communication protocol ran during a synchronization phase between the OptoHybrid and the CTP7. As the 40 MHz clock on the former is generated from the 320 MHz clock, the phase relation with the clock of the latter is not known which causes data to be misaligned. At start up, the CTP7 sends a constant pattern of 0x76BC which the OptoHybrid tries to find in the data. If the pattern is not found, a bitslip is operated, shifting the received data by one unit. The operation is repeated until the link is aligned. At that moment, the OptoHybrid, which was transmitting a stream of zeros, switches to a constant pattern of 0x76BC76BC. When the CTP7 is able to identify this pattern, it returns a single 0x8943 word informing the OptoHybrid that data is aligned. Otherwise, the

CTP7 to OptoHybrid frame

15		0
TTC	Header[3:0]	Address[31:24]
TTC	Address[23:12]	
TTC	Address[11:0]	
TTC	0000	Data[31:24]
TTC	Request Data[23:12]	
TTC	Request Data[11:0]	
TTC	0xABC	

OptoHybrid to CTP7 frame

31	0
Header[3:0]	0x000BCBC
VFAT2 data x7	
Response data	

Table 6.5 – Data format of the packets sent between the off-detector and on-detector electronics over the GBT link.

latter performs a bitslip on the transmitted data.

Once the alignment procedure is done, the two systems can communicate through the data format exposed in Table 6.5. Although similar to the data format used over the GTX links, two fundamental changes have been operated: due to the fact that data is clocked at 40 MHz in the fabric of the FPGA, the TTC commands need to be sent every BX; the bandwidth has been changed modifying the width of the data packets.

Running along the GTX core, the GBT module is defined as a Wishbone master meaning it can operate the same requests as the former. Furthermore, tracking data is duplicated and sent over the two links ensuring redundancy and protection against failure as the CTP7 can dynamically change between the two cores.

6.7 Towards the Final DAQ System

The VFAT2 was designed to handle trigger rates up to 100 kHz and transmit data at a maximum rate of 0.32 Gbps for the trigger data and 0.02 Gbps for the tracking data. These data rates can easily be accommodated by the current system, however, they will be drastically increased for the final system using the VFAT3 ASIC. Furthermore, it is foreseen to multiply the L1A rate by a factor 10 from the current 100 kHz to 1 MHz. It is thus crucial to ensure that the GBT and other components will be able to cope with this increase.

The VFAT3 will output 64 trigger bits per BX which will be collected by the OptoHybrid. Many of these are empty as less than one particle hits the chamber per BX on average. A zero suppressing algorithm is thus implemented on the OptoHybrid in order to only encode bits which are active. Each fired trigger bit requires 11

bits to describe: 5 to address the VFAT3 and 6 to address the strip that is hit. As a maximum of 80 bits per BX can be transmitted over the single 3.2 Gbps data link to the off-detector electronics or the CSCs, the probability that the number of hits exceeds 7 and thus overflows the link can be computed. This results in a probability of less than 6×10^{-5}. To reduce the probability of overflowing the link, double transmitter modules can be used to double the available bandwidth.

The tracking data can be zero suppressed by the VFAT3 itself resulting in data packets of minimum 48 bits compared to the 246 bits when using the lossless data format. Simulations were developed in order to compute the mean size of the data packets while taking into account the rate of particles to which GE1/1 will be exposed. This results in a data rate of 0.17 Gbps and 0.48 Gbps for the zero suppressed and lossless algorithms respectively. Both have a probability of less than 10^{-7} to overflow the three GBT data links. When running at L1A rates of 1 MHz, a factor of 10 is applied to the resulting data rates, which still yield a probability of overflow less than 10^{-7}, limit of the simulation.

6.8 Conclusion

The November 2015 test beam campaign showed that the DAQ system composed of the VFAT2 ASIC, GEB v2, OptoHybrid v2a, and GLIB was able to run flawlessly during the entirety of the campaigns. The latter was able to handle two GE1/1 detectors with ease while providing quality data which has been used to compute the noise and efficiency on the system as a function of various parameters as well as the cluster multiplicity and size.

The flexible firmware architecture we developed for the OptoHybrid, which provides the user with numerous monitoring and control options, ran for the duration of the test beam and carried out all required tasks: handle the slow control with the VFAT2s, read out tracking data, and monitor the system. The wishbone-like communication protocol implemented in the FPGA allows future developments to be easily integrated and modules to communicate with one another. The firmware of the GLIB and the communication through the optical links have ran smoothly showing no errors on the transmitted data. The readout rate was high enough to prevent the buffer from overflowing and the data taking has shown to be well integrated with the software. The latter, namely a web application that we designed, has been used extensively and is still in use for gain and uniformity studies at CERN. It has proven to be a complete solution to control and monitor the DAQ system allowing users to run all the necessary scans.

Finally, we have analyzed the data recorded during the test beam. Threshold scans have shown an elevated noise level around 4%, which has since then been under-

stood and solved, forcing us to work at higher threshold values and thus reducing the efficiency by 1%-2% to 97%. Nonetheless, using the latency scans and the tracking data, we performed various studies of the noise and efficiency levels against high-voltage, threshold, and trigger rates. An efficiency of 97% has been obtained when running at nominal values, and trigger rate capability has been tested up until 120 kHz with a resulting efficiency of 96%. This trigger rate translates to a hit rate of 40 kHz cm^{-2}, 10 times higher than the rate expected in CMS. Studies on the cluster size have yielded an average of 1.42 ± 0.5 for muons and 1.62 ± 0.6 for pions. These results are in excellent agreement with measurements performed during two test beams campaigns.

From the analysis we have performed and the results obtained for the efficiency of the system, we conclude that the GEM detectors equipped with the DAQ system we have designed meet the requirements of the project regarding the efficiency and the rate capability of the chambers.

Qualification of the Electronics 7

To prepare the DAQ system for the slice test, the electronic components need to be tested and qualified. For the VFAT2s, qualification procedures have to be developed to characterize the analog front-end of each chip, record their response to calibration pulses, and optimize the parameters to achieve uniformity over the entire detector. For the GEB, qualification means ensuring that no shorts are present and that the integrity of the signals is kept over the full length of the board. Finally, the system as a whole needs to be tested against communication errors or readout problems.

In this chapter, we present the algorithms we have developed to test and qualify the various systems which rely on the firmware and software development made for the test beam. We detail the calibration procedure of the analog front-end of the VFAT2 which has never been done before. Next, we describe the GEB testing PCB which we have created in order to test the communication through the GEB. Finally, we review the script used to test the system as a whole which is employed at CERN and in various research laboratories to set up DAQ systems for future developments.

7.1 Qualification of the VFAT2s

The qualification process of the VFAT2s has been automated using Python scripts relying on IPBus to run the various procedures using the dedicated firmware modules of the OptoHybrid. The aim of this tools is to reject faulty chips and record the parameters of the analog front-end which slightly vary from component to component.

7.1.1 Identification of Faulty Channels

The first step of the procedure is to run a threshold scan for each channel to reveal those that are not responding or that are noisy. To do so, the tools use the threshold

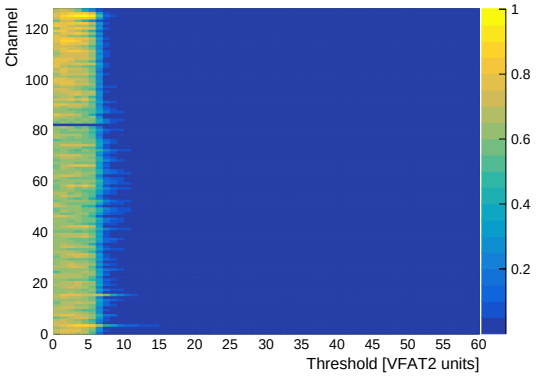

Figure 7.1 – Two-dimensional graph plotting the evolution of the noise level for each channel as a function of the VFAT2 threshold in VFAT2 units.

scan module of the firmware that relies on tracking data to access the full granularity information. Figure 7.1 is a two-dimensional graph plotting the evolution of the noise level for each channel as a function of the VFAT2 threshold in VFAT2 units. As can be seen, one of the channels appears to be damaged as it does not display any change during the whole scan and remains empty. From this result, the VFAT2 would be disqualified and not used for data taking. If performing these scans on instrumented VFAT2s, the results can be used to mask faulty channels and regain acceptable operational conditions.

7.1.2 Analog Front-End Currents and Voltages

Following the threshold scan, the procedure reads out the currents and voltages that are produced by the VFAT2 to bias its analog front-end. When writing values in the eight bits registers defining the threshold and other parameters, a conversion to an analog value is performed inside the chip, which will vary from component to component. It is therefore essential to record the results of each chip. In order to access the analog signals inside the VFAT2, two analog signals are sent out: one for the voltage reference and one for the current reference. These lines are routed on the GEB and transmitted to the OptoHybrid which uses an ADC to digitize the information. To fit the window of operation of the ADC, which has a range of 0 V to 1 V, the voltage signal is sent through a voltage divider and the current signal is measured over a resistor. Figure 7.2 plots the ADC counts, current values on the left, and voltage values on the right of each settable register of the VFAT2 as a function of the eight bits value written in the register.

All but one of the current based registers are related to the amplification and shaping of the analog signal. They power the front-end and are used to define the length of the tail of the signal shape, the rise time of the distribution, etc. The remaining register, IComp, defines the amount of current delivered to the comparator which in turn influences the response time of the latter. These parameters are set to default

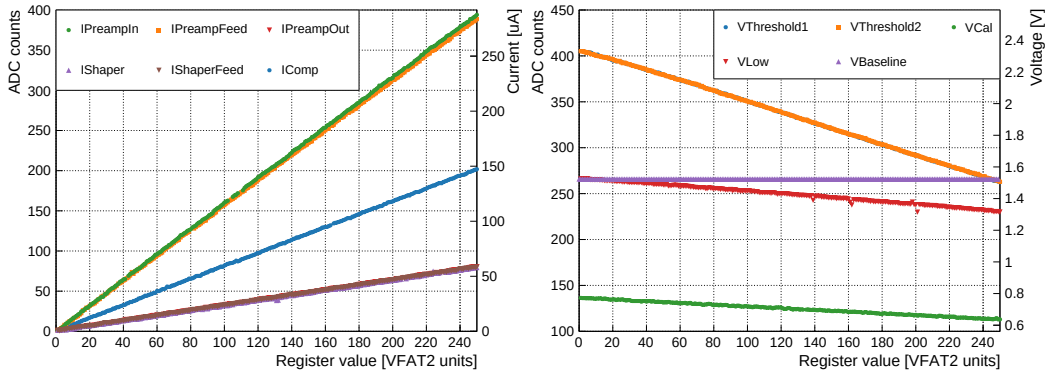

Figure 7.2 – Plots of the ADC counts, current values on the left, and voltage values on the right of each settable register of the VFAT2 as a function of the 8-bit value written in the register.

values provided by the reference manual of the VFAT2.

The voltage based registers are used for the comparator and the calibration modules. The VThreshold1 and VThreshold2 signals are those defining the threshold on the strips. By design, VFAT2 can handle positive and negative pulse and provides one threshold register for each polarity. During operations, one of the registers is set to zero and the other one is used to adjust the threshold, meaning that a single parameter is used. The VBaseline, VCal, and VLow registers are used to define the amplitude of the calibration pulse sent to the channels. Figure 7.3 shows a diagram of the injection and calibration circuit of the VFAT2. The amplitude of the calibration pulse is equal to the difference between VBaseline which has a fixed value of 1.514 $\pm$ 0.009 V and VLow which is programmable between 1.529$\pm$ 0.002 V and 1.314 $\pm$ 0.002 V. The output pulse is sent to one channel over a capacitor of 100 fF resulting in a pulse amplitude in term of charge of

$$Q = 100 \text{ fF} \times (\text{VBaseline} - \text{VLow}) \,. \tag{7.1}$$

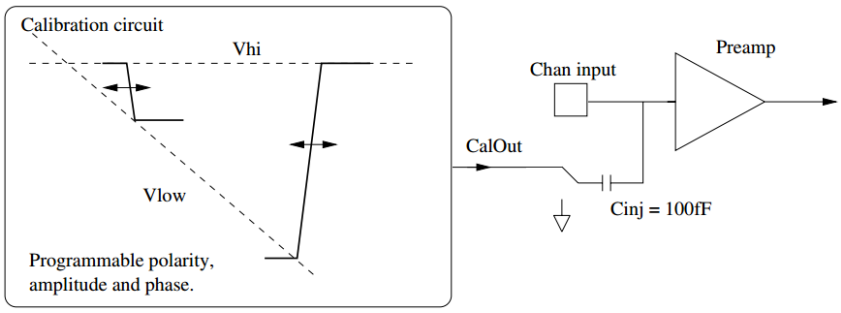

Figure 7.3 – Diagram of the injection and calibration circuit of the VFAT2 highlighting the path of the calibration pulse towards a given channel [53].

By inserting the values of VBaseline and VLow in Equation 7.1, a range of injected charge can be computed and qualitatively compared to the one given in the VFAT2 specifications [53]. The latter states a range of -2 fC to 18.5 fC with a slope of 0.08 fC per VFAT2 unit, while the former results in a range of -1.5 $\pm$ 0.9 fC to 19.2 $\pm$ 0.9 fC with a slope of 0.08 $\pm$ 0.01 fC per VFAT2 unit. Both measurements agree well within the error margins and help to validate the methodology used during the calibration routines.

7.1.3 Analog Front-End Calibration

Using the results from the ADC, the script can determine what amount of charge is injected during the calibration phase of the VFAT2 upon reception of a CalPulse fast command. With this, it is possible to produce s-curves which for a given threshold indicate at which amplitude of the calibration pulse a signal becomes visible. Figure 7.4 plots the hit-to-event ratio as a function of the calibration pulse height in VFAT2 units for a threshold of 25. The calibration pulse height refers to the value written

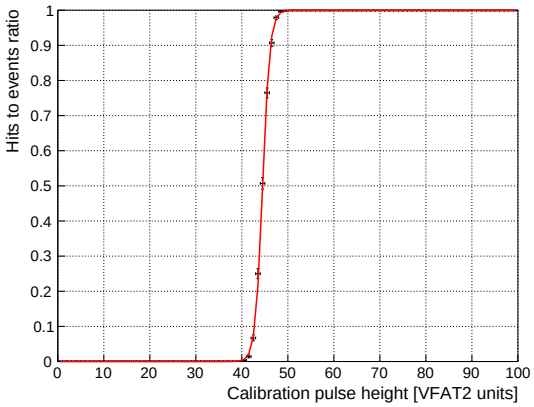

Figure 7.4 – Plot of the hit-to-event ratio as a function of the calibration pulse height in VFAT2 units for a threshold of 25.

in the VLow register which linearly affects the charge deposited on the channels. The graph shows that below a value of 40 no signal is detected due to the threshold level. The turn-on of the curve is defined as the value for which 50% of hit-to-event ratio is reached and is in this case equal to 44.

The operation is then repeated for various threshold values and the turn-on value is extracted automatically using a logistic function to fit the curve. The equation of the function is as follows

$$f(x) = \frac{1}{1 + e^{-\sigma(x - x_0)}} \tag{7.2}$$

where x_0 is the value of the midpoint or turn-on, and σ is the steepness of the slope. Figure 7.5 is a collection of s-curve scans for various thresholds on the left which can be translated to a plot of the calibration pulse height of the turn-on in VFAT2 units and current as a function of the threshold in VFAT2 units on the right.

Using these results, the threshold value applied on the strips can be converted to a charge, which is of importance to compute the equivalent noise charge of the system.

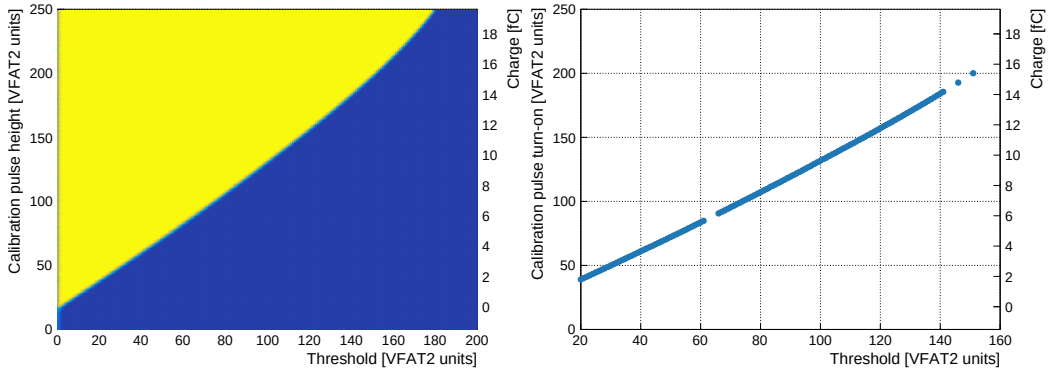

Figure 7.5 – Left: collection of s-curve scans for various thresholds with corresponding current values. Right: plot of the calibration pulse height of the turn-on in VFAT2 units and current as a function of the threshold in VFAT2 units.

This is the amplitude of the noise in terms of electrons, which can easily be compared to the signal induced by the particles which is often expressed in femtocoulombs. The results presented here above are in agreement with those obtained during the initial qualification of the VFAT2 [53] both displaying a qualitative charge range of -2 fC to 18 fC.

7.1.4 Analog Front-End Equalization

The results obtained in the previous procedures were values averaged on all the strips. However, channels display a dispersion around the central value which induces a bias as the effective threshold is not constant across the chip. To eliminate this effect, it is possible to equalize the response of the front-end by using channel by channel programmable registers. Each of them is equipped with a programmable register which slightly adjusts the threshold value. The equalization procedure used to align the front-ends starts by applying a given value to the common threshold register and sets all channel registers to zero. It then performs s-curve scans for each channel. The operation is repeated but with the channel registers set to their maximum value. Figure 7.6 plots the s-curve scans for each channel with the minimum value of the channel registers on the left and the maximum value on the right. The two plots obtained this way are very similar with only the turn-on value changing.

To align the channels, the average value of turn-on for both the minimal and maximal value of the channel registers is taken and set to be the aim of the equalization routine. For each channel, the script then sets its individual register value to the mean value and performs an s-curve scan. According to the turn-on value, the register is incremented or decremented and the procedure is repeated until the channel is aligned. The results of the alignment are shown in Figure 7.7 which plots the s-curve scan for each channel with the aligned values for the register on the left and the dispersion of the turn-on value for the non-aligned and aligned configurations.

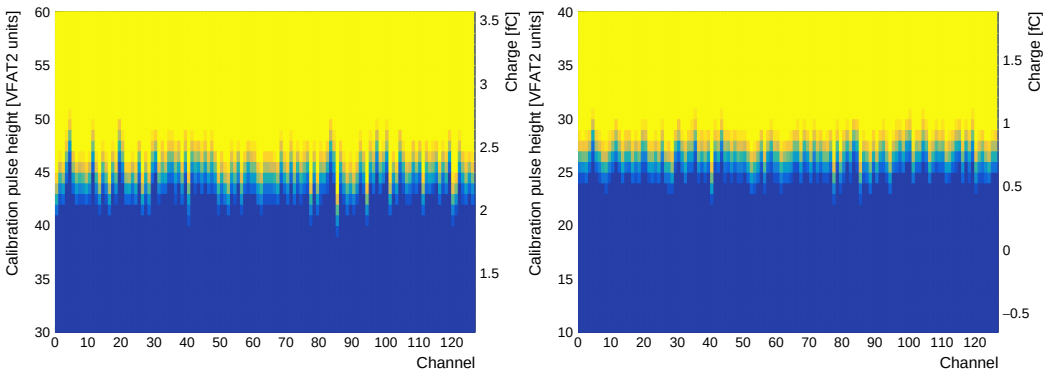

Figure 7.6 – Plots of the s-curve scans for each channel with the minimum value of the channel registers on the left and the maximum value on the right.

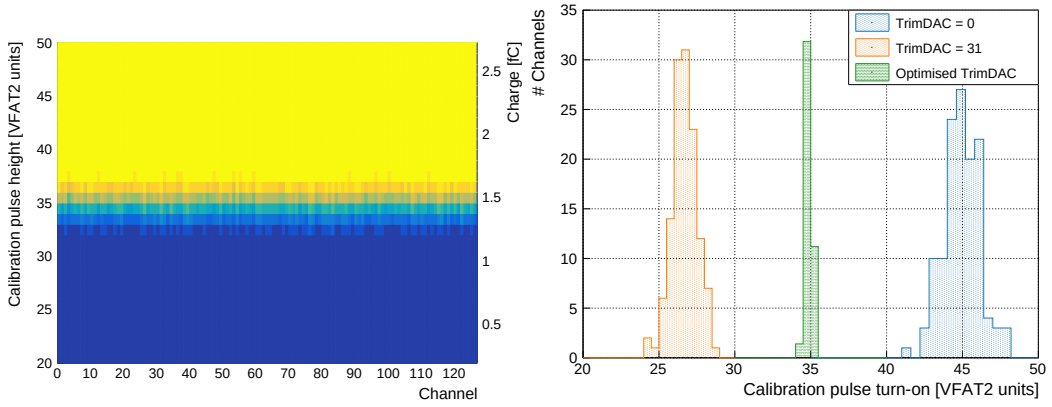

Figure 7.7 – Left: plot of the s-curve scan for each channel with the aligned values for the channel register. Right: dispersion of the turn-on value for the non-aligned (orange and blue) and aligned configurations (green, scaled by a factor of 0.35).

At the end of the script, the channels of the VFAT2 have all been tested against noise issues and have been adjusted so that they all display the same response to calibration pulses. The dispersion of the channels around the central value, defined as the RMS of the distribution, is improved from an initial value of 1.22 to a final value of 0.19 after alignment, corresponding to 0.10 fC and 0.02 fC respectively. This is the first time the full range of possibilities offered by the VFAT2 is used in a DAQ system and demonstrates that these features are within working specification.

7.1.5 Noise Measurement

A more quantitative comparison can be performed against [35] with respect to the electronic noise on the VFAT2. The latter will induce a smearing effect on the charge injected on the channel, either through noise on the voltage used to generate the pulse or on the electronics in the preamplification step. In a perfect system, the curve in Figure 7.4 would be a step function with values at 0 before the threshold is reached, and at 1 above. The noise is thus responsible for the steepness of the slope and defined as the sigma of the curve. After performing a fit of a logistic function, which is given in Equation 7.2, the sigma is found to be 1.27 ± 0.02 VFAT2 Units. This value can be converted to femtocoulombs using the previous results and yields 0.10 ± 0.01 fC, which in turn can be expressed in terms of electrons to obtain

$$\sigma_e = 624 \pm 62\ e^- \ . \tag{7.3}$$

This can be compared to the value listed in the specification of the VFAT2 which is of $589 \pm 84\ e^-$ when no detector is connected to the electronics. Both results are thus compatible within the error bars.

The second measurement that can be performed is to study the variation of the noise from channel to channel. Figure 7.8 plots the Equivalent Noise Charge (ENC), or

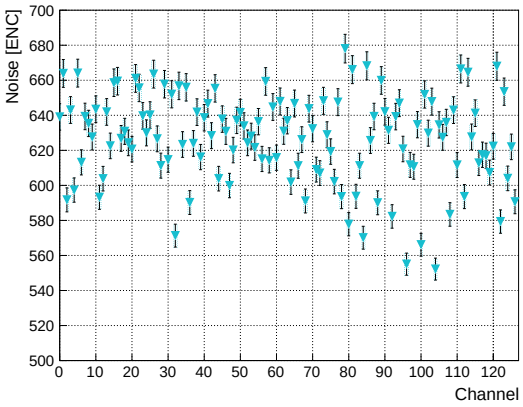

Figure 7.8 – Plot of the Equivalent Noise Charge (ENC) for each channel.

noise in terms of electrons, for each channel. The distribution is centered around 622 e^- with a spread of 31 e^-. Although the previously exposed procedure is able to align the mean of the s-curves, it does not affect their slope and thus noise. This parameter cannot be changed and is intrinsic to the VFAT2.

7.2 Qualification of the GEB

In order to test the GEB, a small PCB equipped with the same connector as the VFAT2 Hybrids was developed. Using signals originating from the OptoHybrid, the board can check the integrity of the data and use on-board LEDs to indicate the results of the tests for the current position.

7.2.1 The GEB Testing Board

The PCB is a four layer board equipped with an FPGA and a microcontroller unit (MCU) to analyze the signals. Figure 7.9 provides 3D models and a PCB layout of the board. Next to the processing units, the board also features an ADC, a DAC, a flash memory, a USB-to-SPI converter to link the universal serial bus (USB) port to the FPGA through serial peripheral interface bus (SPI), and powering options.

Powering of the board is done either through the GEB connector which provides 2.5 V to the system or through the USB port. In case USB is used, the GEB power source is automatically disabled and 3.3 V is given to the system to increase processing speed. The 2.5 V or 3.3 V are further used to generate the 1.2 V required for the internal logic of the FPGA.

The processing units of the board, namely the Xilinx Spartan-6 FPGA (XC6SLX9-2TQG144C) and the ATMEL MCU (ATmega164PA), are what controls the various components. The FPGA is the main element of the board and is connected to all other chips. It can either control them directly or act as a router for the signals originating from the MCU. The communication between the two processing units is done via four serial buses: two universal asynchronous receiver/transmitter (UART), one SPI, and one bus composed of four GPIOs. Next to the links to the FPGA, the

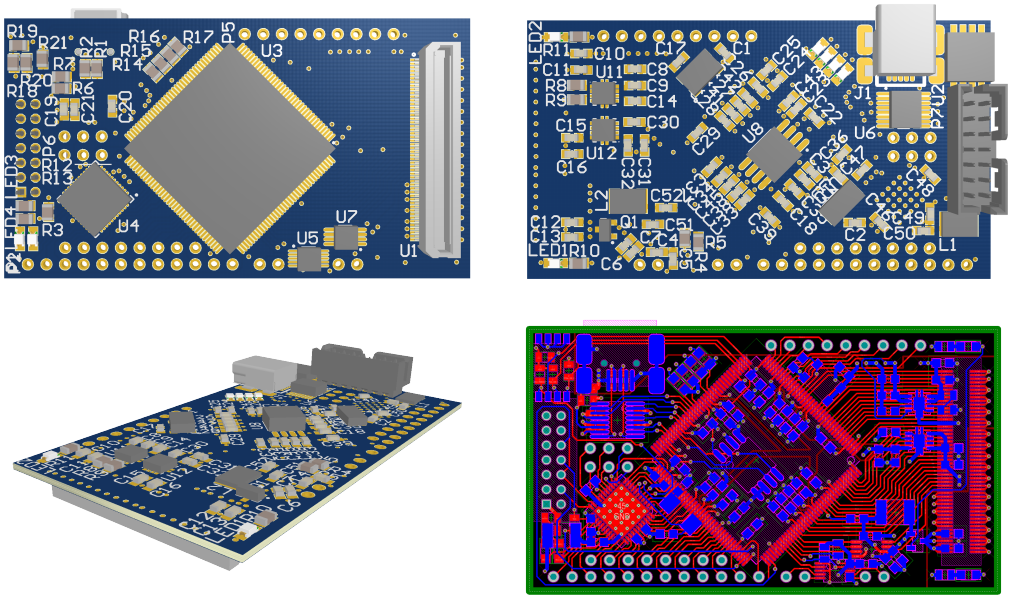

Figure 7.9 – 3D models and PCB layout of the GEB testing board.

remaining 16 IOs of the MCU are broken out on header connectors for debugging purposes. The same is done for 22 IOs of the FPGA which are connected to headers.

The GEB connector holds 30 digital signals, 2 analog signals, and power and ground lines. The digital signals are all connected to the FPGA while the two analog signals are connected to the DAC.

The USB port is used to communicate with a computer using an FTDI chip (FT220XS) which converts the USB protocol to SPI.

The analog part of the board is controlled by the ADC (ADS1015) and the DAC (DAC8532). The two outputs of the DAC are connected to the GEB connector to drive the two analog signals digitized by the OptoHybrid. The ADC signals on the other hand are purely used for debugging purposes and connected to headers.

Programming the system is done through the Joint Test Action Group (JTAG) protocol which connects to the FPGA and the MCU. JTAG allows multiple devices to be connected in series and programs them both by shifting the configuration from one to another. The configuration scheme of the system can be modified to only affect the FPGA by removing a 0 Ω resistor which bypasses the MCU.

7.2.2 Results of the Qualification of the GEB

To test the GEB positions, the OptoHybrid is used to generate a pseudorandom binary sequence (PRBS) and transmit it over the data lines along with a reference clock. PRBS is often used to detect errors on communication lines including optical fibers. It relies on a sequence of bits to compute the next bit in the stream. A vector of bits is loaded in shift registers and is moved by one position every clock cycle. The overflowing position in transmitted on the line while the entering position is

computed by summing the data stored in given positions. PRBS-7, for example, uses 7-bit-long vectors and computes the new element with

$$x_0 = x_6 \oplus x_5 \wedge x_1 \tag{7.4}$$

which yields a repetition period of pattern of 127 clock cycles.

For each differential pair or single ended signal, a different seed is used for the PRBS-7 generator. By receiving the clock along with the data, the GEB tester board can decode the bit streams and ensure that the encoding is correct. In case one of the lines is faulty, the LEDs connected to the FPGA are used to display an error code. The analog lines are tested by generating a given voltage using the DAC. The OptoHybrid is then used to readout the value on the line and check the results are matching.

This procedure was used to test one GEB which was found to be healthy. To validate the functioning of the algorithm, the response to an error was simulated by injecting an error from the OptoHybrid. The latter was successfully detected by the testing board.

7.3 Qualification of the DAQ System

The final qualification procedure developed aims at testing the DAQ system in its entirety. Figure 7.10 provides a screen shot of the output of the script. The latter first attempts to establish communication with the GLIB/CTP7 and the OptoHybrid in section A and B. It then performs a series of read/write operations on the registers ensuring that every request receives a response in section C and D. Once the tests on the GLIB/CTP7 and OptoHybrid are done, the script detects all present VFAT2s in section E through I2C requests. Using these, random read/write operations are made on the registers of the VFAT2s in section F to test the reliability of the communication. Afterwards, in section G, chips are turned on one by one and tracking data is read out. The script verifies that the counters add up and that the data is not corrupted. Then, it performs the same action with all VFAT2s active at the same time in section H. The last test performed using the VFAT2s, in section I, is the maximum readout rate that can be achieved by the system. Finally, section J looks at the error rate on the optical links to ensure no errors appear on the communication lines.

This test is used to check the system configuration and ensure that every component is configured correctly. It can detect errors at any level of the system, from the VFAT2s to the GLIB/CTP7.

```
A. Testing the GLIB's presence
   Trying to read the GLIB board ID... If this test fails, the script
   will stop.
   >   Passed...

B. Testing the OH's presence
   Trying to set the OptoHybrid registers... If this test fails, the
   script will stop.
   >   Passed...

C. Testing the GLIB registers
   Performing single and FIFO reads on the GLIB counters and ensuring
   they increment.
   >   Passed...

D. Testing the OH registers
   Performing single and FIFO reads on the OptoHybrid counters and
   ensuring they increment.
   >   Passed...

E. Detecting the VFAT2s over I2C
   Detecting VFAT2s on the GEM by reading out their chip ID.
   Detected 18 VFAT2s: [1, ..., 23]

F. Testing the I2C communication with the VFAT2s
   Performing random read/write operation on each connect VFAT2.
   >   Passed... #1
     [...]
   >   Passed... #23

G. Reading out tracking data
   Sending triggers and testing if the Event Counter adds up.
   >   Passed... #1
     [...]
   >   Passed... #23

H. Reading out tracking data
   Turning on all VFAT2s and looking that all the Event Counters add up.
   >   Passed...

I. Testing the tracking data readout rate
   Sending triggers at a given rate and looking at the maximum readout
   rate that can be achieved.
   Maximum readout rate 200000 Hz

J. Testing the optical link error rate
   GLIB tracking link error rate is of          0 Hz
   GLIB trigger link error rate is of           0 Hz
   OptoHybrid tracking link error rate is of    0 Hz
   OptoHybrid trigger link error rate is of     0 Hz

K. Results
   A.    >   Passed...
     [...]
   J.    >   Passed...
```

Figure 7.10 – Output of the qualification procedure developed to test the DAQ system in its entirety.

7.4 Conclusion

For the first time, a method was developed to use the calibration capabilities of the VFAT2 to their full extent by taking advantage of the channel by channel optimization. The procedure we designed to qualify the VFAT2s is used to detect faulty units and perform qualification tests which entirely characterize the analog front-end of each VFAT2. The first part of the results can be compared to the specifications of the VFAT2 and yields the range of charge that can be injected on the channels and the noise on the VFAT2. The former results in a range of -1.5 $\pm$ 0.9 fC to 19.2 $\pm$ 0.9 fC with a slope of 0.08 $\pm$ 0.01 fC per VFAT2 unit, while the latter was found to be of 624 $\pm$ 62 e^-. This is in turn used to provide information on the effective threshold applied on the channels in terms of electrons and thus induced charge. The second part of the script aligns the response of each channel around a central value and reduces the threshold disparity from an initial dispersion value of 1.22 to a value 0.19 after alignment in VFAT2 units, corresponding to 0.10 fC and 0.02 fC respectively.

Qualification procedures were also created for the GEB, for which we designed a GEB testing board to test each position and detect broken lines. The board relies on an FPGA coupled with an MCU to communicate with the OptoHybrid and test the integrity of the transmitted signals. The results of each test are displayed using on-board LEDs and inform the user of the quality of the signals and thus the GEB.

Finally, we tested the system as a whole using a script which targets specific components of the architectures. Random read/write operations to the GLIB/CTP7, OptoHybrid, and VFAT2s are performed, tracking data is read out, and stress tests are done to push the system to the limit. A report is provided to the user about each individual point in order to qualify the system or not.

These tools are used for the preparation of the slice test to select appropriate components and install testing facilities at CERN and in other associated research laboratories.

Irradiation Tests

8

The OptoHybrid will be located in a region of CMS exposed to fluxes of particles up to 20 kHz cm^{-2}, some of which might interact with the FPGA and cause errors in the logic. Those could in turn influence the functioning of the system and degrade its performance. To understand and potentially solve this issue, irradiation tests have been performed to measure the interaction cross section of the particles with the various components of the FPGA. Two OptoHybrids v2a programmed with a dedicated firmware designed to detect errors have been placed in a high intensity proton beam of fluxes up to 2×10^8 particles cm^{-2} s^{-1}. The two boards were each controlled by an additional OptoHybrid placed outside the test area which recorded the events and statistics.

In this chapter, we provide the reader with an overview of the internal architecture of an FPGA to better understand the potential sources of errors and the effects of radiation. We then describe the firmware of the irradiated and control FPGAs which has been used during the test. The setup and beam parameters of the irradiation test are reviewed before presenting the results that were obtained after analysis.

8.1 Architecture of an FPGA

To optimize the occupancy of the resources of the FPGA and develop code that uses the full potential of the device, a deep comprehension of the intrinsic architecture of the chip is required. Although families of FPGAs differ in size and complexity, their building blocks remain the same and are described hereafter.

8.1.1 Configurable Logic Blocks (CLB)

Configurable Logic Blocks (CLBs) [54] are the building blocks of the FPGA used to implement sequential and combinatorial logic. They are composed of two important objects: Look-Up Tables (LUTs) and registers. LUTs are components which outputs are a function of the inputs as defined in a programmable table. They implement a truth table for every possible combination of the inputs which defines the value of the outputs. The response of a LUT to a change in the inputs is almost instantaneous. LUTs in the Xilinx Virtex-6 FPGAs can either implement functions with six inputs and a single output or functions with five inputs and two outputs. The outputs of the LUTs can, if so required by the design, be connected to registers which sample the signals at the rising edge of a given clock. Registers are used for their sample-and-hold functionality which makes designs synchronous to clocks. With these two components, CLBs can implement complex functions and describe intricate systems.

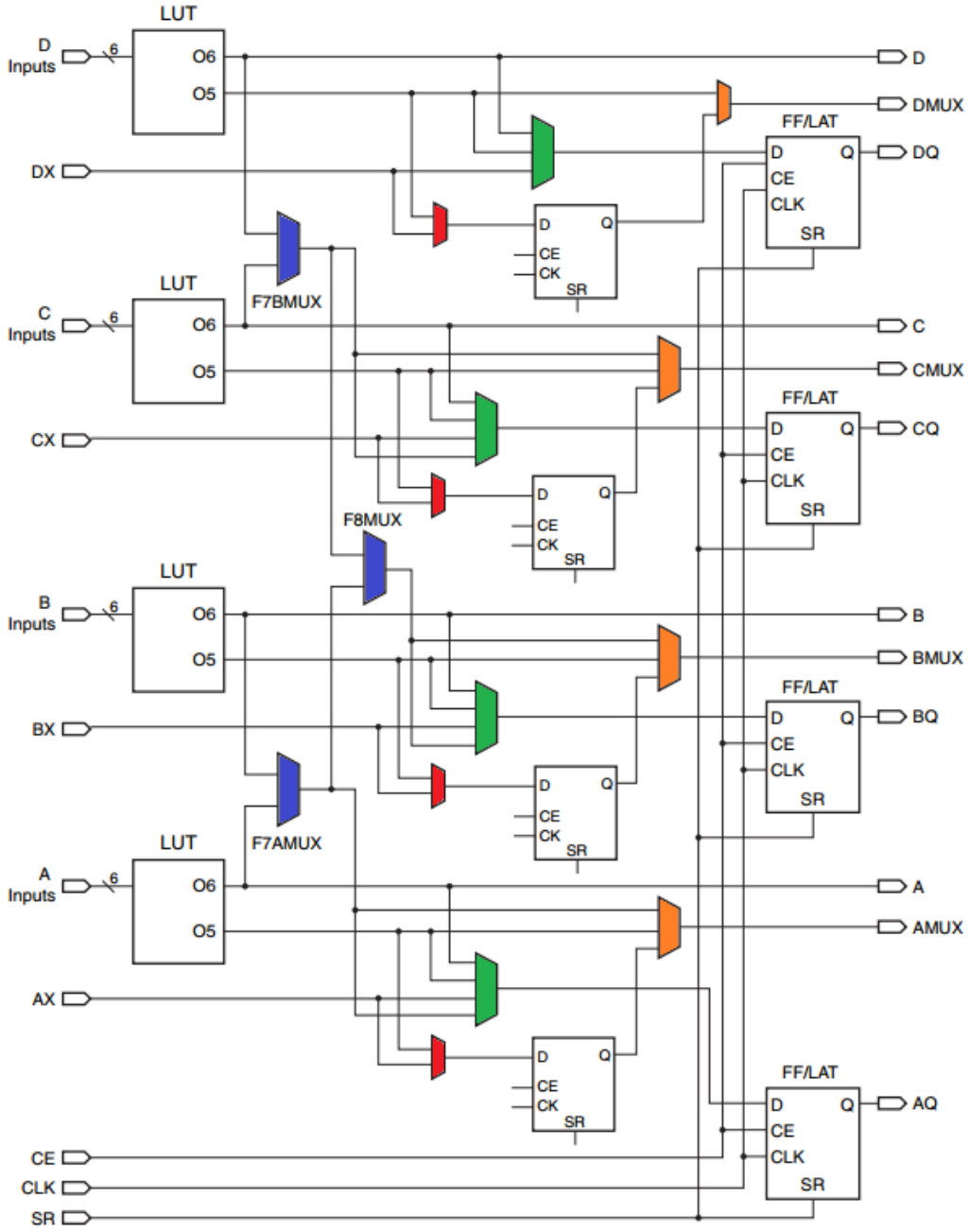

Figure 8.1 – Simplified view of a slice composed of four LUTs on the left and eight registers on the right [54].

Each CLB is composed of two slices each made of four LUTs and eight registers which layout is shown in Figure 8.1. Each LUT is connected to an input bus of six signals (A, B, C, and D input buses) and to an unbuffered output bus of a single bit (O6 to A, B, C, and D). Four additional signals enter the slice (AX, BX, CX, and DX) and are connected to multiplexers (red) which allow to buffer either the former or the second output of the LUTs (O5). The output of the same buffer is then connected to a second multiplexer (orange) which allows to select either the former or an unbuffered output of the LUTs (AMUX, BMUX, CMUX, and DMUX). Finally, a third multiplexer (green) connects a range of signals to the registers on the right which are then connected to four outputs (AQ, BQ, CQ, and DQ). Three additional multiplexers (blue) offer the possibility to mix the signals from the four LUTs to generate a wider range of logical operations. The flexibility offered by this architecture is what enables FPGAs to implement complex designs. The Xilinx Virtex-6 FPGA used in the OptoHybrid v2a (XC6VLX130T) holds 10 000 CLBs with a total of 80 000 LUTs and 160 000 registers.

8.1.2 Digital Signal Processing (DSP)

Digital Signal Processing units (DSPs) [55] are modules which perform mathematical operations using dedicated hardware elements. They are used to quickly solve problems without relying on CLBs which can consume large amount of resources to perform an equivalent task. Figure 8.2 shows a diagram of the DSP48E1 slices present in the Xilinx Virtex-6 family. They include a 30-bits adder, a 25-bits by 18-bits multiplier, and a programmable module which can implement either a subtraction, an addition, or a logic operation. Using the various multiplexers and configuration bits, the user can define the data paths that are followed and thus create DSP modules which meet the requirements of the design.

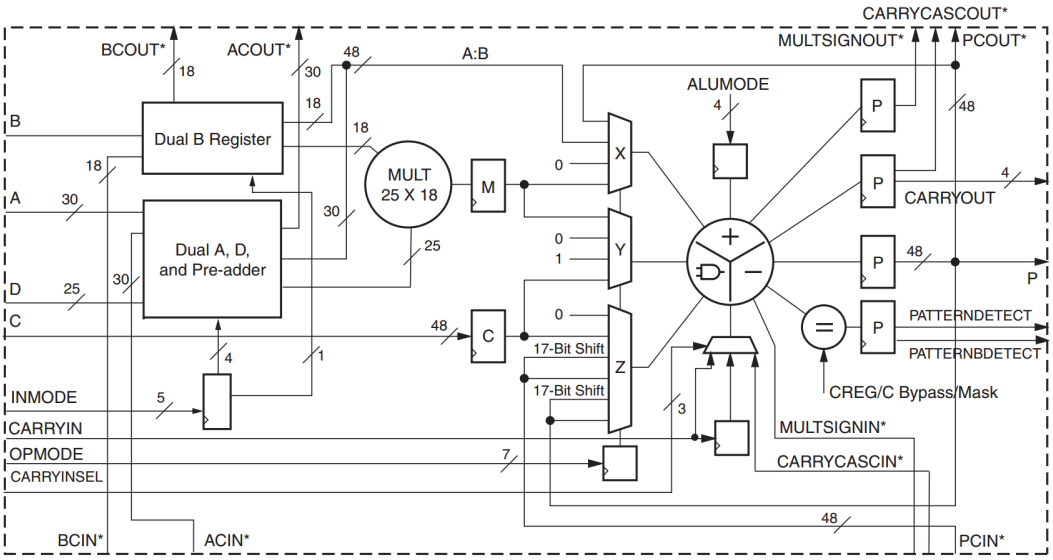

Figure 8.2 – Diagram of the DSP48E1 slices present in the Xilinx Virtex-6 family [55].

8.1.3 The Switching Matrix

The inputs and outputs of the slices of the CLBs as well as of the other components of the FPGA are interconnected through the switching matrix. It is a vast network of wires and switches which are used to route signals between elements. The open or closed state of each switch is programmable and defines the routes signals follow inside the logic. Figure 8.3 shows an element of the matrix connected to two slices (blue) along with the signals that are routed in the network (cyan). Complex designs can span a large area of the FPGA due to the high usage of logic and thus be difficult to place and route by the compiler. A technique called floorplanning is used to reduce compilation time and improve the design by constraining parts of the code to given areas in the FPGA. This reduces propagation delays and logic usage which in turn yield a more efficient design.

8.1.4 Block RAM (BRAM)

Besides programmable logic, the FPGA also includes dedicated storage elements. Block RAMs (BRAMs) [56] can store up to 36 kb of data which can be configured in different ways: 32K x 1 bit, 16K x 2 bits, etc. They can also be used as First In First Out (FIFO) modules which are similar to data queues.

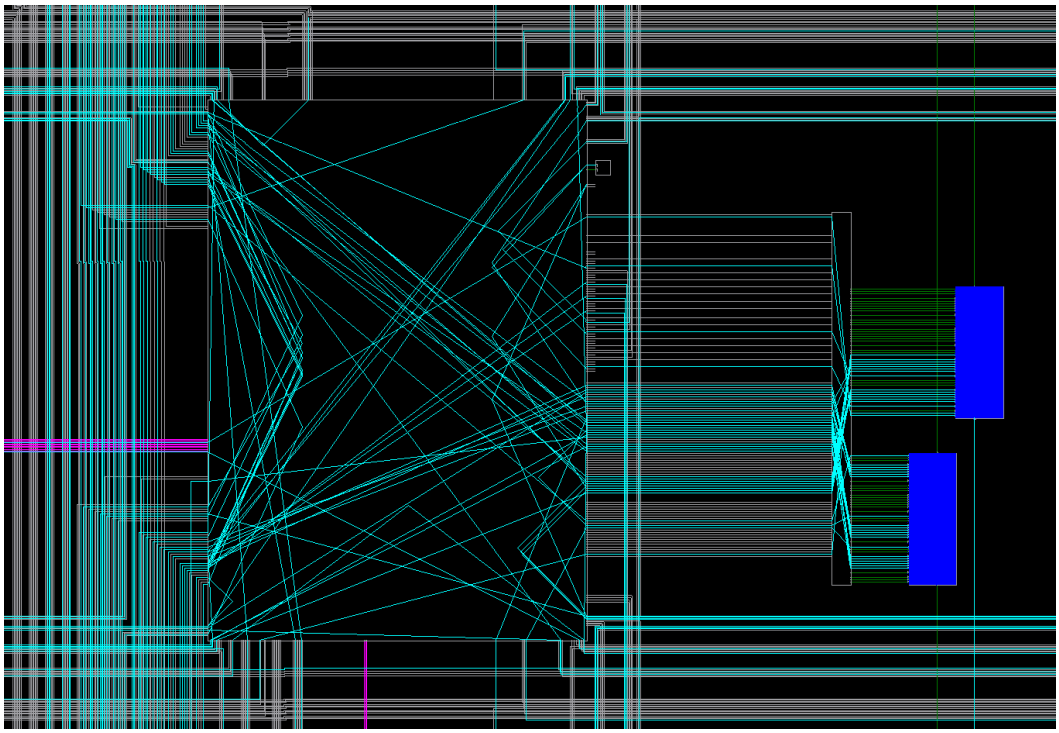

Figure 8.3 – Schematic of two slices (blue) connected to an element of the switching matrix which routes the signals (cyan) between components over the multitude of existing paths (gray).

8.1.5 The Configuration Memory

The configuration memory holds the configuration of the entire FPGA and defines its behavior. It sets the truth table of the LUTs, parametrizes the DSP, creates the connection between elements, etc. It is what implements the design in the FPGA. In most FPGAs, the configuration memory is volatile and will lose its content upon power down or reset. To reload it, the FPGA tries to read it out of an attached memory device or remains in a blank or corrupted state if it fails to do so.

8.2 Effects of Radiation on FPGAs

To understand how radiation affects an FPGA, it is import to know how the interaction between the particles and the silicon occurs and what the energy transfer mechanisms are. Then, the consequences of these energy depositions on the transistors of the chip can be analyzed along with the effects on the FPGA.

8.2.1 Energy Losses in Matter

Radiation affects the functioning of FPGAs through its interactions with the silicon substrate that composes the chip. These interactions take the form of direct or indirect ionization depending on the nature of the incident particle. Charged particles will directly ionize the medium through their Coulomb interactions with the electrons, causing them to escape with a fraction of the energy of the particle. Neutral hadrons on the other hand interact through indirect ionization which results from the scatterings with the medium which transfer energy to the electrons and nuclei. Both processes result in a trail of ions and electrons in the path of the incident particle, which in the context of silicon devices are called holes and electrons.

These energy losses are stochastic by nature and thus described by probability distributions around a mean value. They further depend upon the energy of the incident particle as different physical processes will have different behaviors for their cross sections. However, a value called the Linear Energy Transfer (LET) is used and defined as the amount of energy lost by a particle per unit of distance in the direct vicinity of the track

$$LET = -\frac{dE}{dX}. \tag{8.1}$$

The latter greatly depends upon the medium and is thus normalized by its density and expressed in terms of MeV cm^2 g^{-1}. As reference, 62 MeV protons have a LET of 8.39 $\times$ 10^{-3} MeV cm^2 mg^{-1} while 33 MeV protons have a LET of 1.37 $\times$ 10^{-2} MeV cm^2 mg^{-1}. The LET thus increases as the energy decreases due to the stronger interactions with the medium which will contain the energy deposition to a small cylinder around the track. The maximum LET is achieved when the particle comes to rest in the medium. High energy protons on the other hand will create δ-rays inside the medium which will in turn create an ionization trail which does not contribute

to the LET. The total deposited charge is thus higher and will have more impact on the devices.

8.2.2 Effects of Radiation on Transistors

The holes and electrons created along the track of the particle will either recombine if they are located in the bulk of the substrate, or drift towards the doped regions of the transistor if located near the p-n junction, as depicted in Figure 8.4. The latter gives rise to the creation of a hole and electron current in the transistor and a resulting current spike at the gates of the component. This can propagate in the circuit and affect its status.

Not all holes and electrons are collected by the gates and over time charge builds up in the oxide region of the transistors and screens or enhances the electric field of the gate. This results in a shift of the voltage threshold of the transistor. Additionally, this also affects the mobility of the charge carriers in the region of the transistor where the conductive channel forms between the two doped regions. This can lead to current leakage and an increase in the voltage shift.

Additionally, the crystal lattice of the transistor can be disrupted by incoming particles through their interaction with atoms of silicon. These defects in the lattice will modify the conduction bands of the silicon and thus modify the conductive properties of the medium. New conduction bands can be created at lower energy levels which can, for example, be reached by thermal electrons. These effects will result in an increase of the leakage current or a decrease of the current gain.

8.2.3 Single Event Transients (SET)

When a current spike occurs in a transistor due to the passage of a particle, it can be propagated inside the combinatorial logic of the device. These events are called Single Event Transients (SETs). They typically have a fast rising time on the order of 100 ps, and last for less than 1 ns. When coupled to sequential logic, such signals can easily be recovered from if their duration is small in comparison to the frequency of the clock used to run the circuit. Unless the error is produced exactly at the

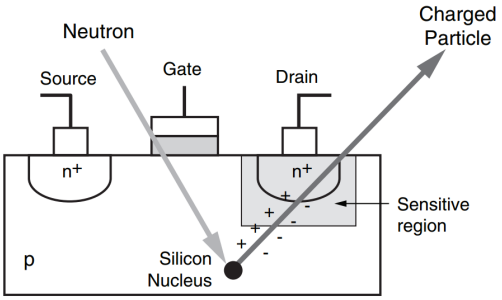

Figure 8.4 – Charge deposition by a charged particle inside a transistor affecting the state of the circuit [57].

sampling time of the registers and thus sampled, it is not stored and thus mitigated. These errors are critical in purely combinatorial logic but can be easily mitigated in sequential logic. They will not be further discusses in the below sections.

8.2.4 Single Event Upsets (SEU)

If the transistor affected by the passage of a particle is located within a memory cell, the event is called a Single Event Upset (SEU). These events can trigger a change of state of the cell and result in a bit flip. Memory cells are composed of a group of transistors running in a stable state. They will retain their value until a change occurs or power is shut down. When an SEU occurs, the current spike causes the memory cell to react and place itself in another stable state, effectively operating a bit flip.

8.2.5 Total Ionizing Dose (TID)

With time, the accumulation of charge by the transistor and the defects in the crystal lattice of the silicon will reduces the efficiency of the device. The threshold voltage of the transistors is shifted such that they no longer fulfill their function and their behavior becomes unpredictable. This leads to a breakdown of the logic gates and thus the higher level functions they are part of. Sectors of the FPGA or the device itself might become unusable. These effects might dissipate over time if the device is removed from the irradiated area.

8.3 SEU Mitigation Techniques

A multitude of mitigation or best practice techniques are used when designing radiation tolerant firmware. Depending on the requirements of the project and the available resources, a combination of one or more of these can be used to correct errors in the various components of the FPGA.

8.3.1 Configuration Memory Scrubbing

To eliminate errors occurring in the configuration memory of the FPGA, the latter can implement a scrubbing mechanism. To this end, the FPGA needs to access its own configuration memory during run time. This is possible through the Soft Error Mitigation (SEM) core, a component which allows the firmware to analyze the data in memory and, using a two bit complement, perform a FEC on the bits. SEM uses two hard components that are embedded in the silicon of the FPGA: the Internal Configuration Access Port (ICAP) which allows the FPGA to read out its configuration memory, and the Error Checking & Correction Frame (ECC_FRAME) which performs the FEC on the data. By handling the data transfer between these two components, SEM is able to detect and correct errors. Although very effective to correct errors on the fly, this technique is time consuming due to the large amount of data to readout and can take up to several milliseconds to identify and repair an error. Furthermore, the FEC will fail to recover the data in case two bits are flipped within the same

word, raising a flag that tells the system a hard reset is needed. While the error remains uncorrected, the components of the FPGA configured through the corrupted bits will encounter errors.

8.3.2 Logic Triplication

A common technique used to mitigate errors in CLBs and DSPs is to triplicate the logic and couple its outputs to a majority voter. This allows to correct the data while SEM solves the corrupted bits in the configuration memory. First, the sequential logic is triplicated, meaning the inputs are sent to three identical modules which output is returned synchronously. The latter are then forwarded to a majority voter which performs a bit-by-bit vote and returns the most probable response. This technique can recover data if only one of the modules is affected by an SEU. In case two modules are corrupted and flip the same bit, data will be corrupted as well. Additionally, a flag can be raised by the majority voter if all three inputs are not identical, meaning an error occurred. A logic diagram of a triplicated module is shown in Figure 8.5. Next to the simple triplication method exposed here-above and represented in black, a more complex triplication can be set in place and is represented in gray. In addition to the triplication of the modules holding the logic, the voter is also triplicated along with its outputs. In that scheme, the three outputs are passed to the next module on three different inputs. This allows to mitigate errors occurring in the voter itself.

8.3.3 Forward Error Correction

SEUs occurring in BRAMs can be corrected using a two complements FEC which can recover single bit errors and detect double bit errors. This option is available on given BRAM components in the FPGA and is transparent to the user. The data returned by the BRAMs is corrected on the fly and a flag is raised if an error is detected.

8.3.4 One-Hot Finite State Machines

A common structure in firmware design is the Finite State Machine (FSM) which is used to describe the process flow of a system through a well-defined number of state

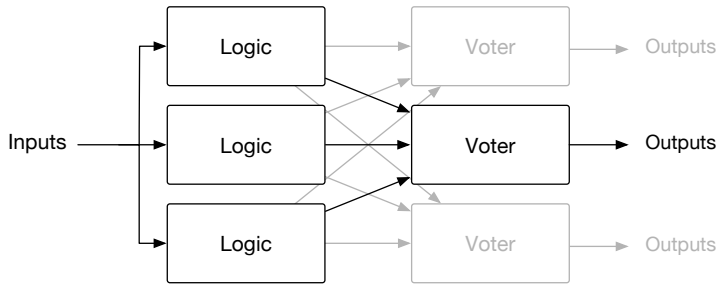

Figure 8.5 – Logic diagram of a triplicated module coupled to a single voter (black) or to three distinct voters (gray).

in which it can be. Each state defines the status of the output signals of the system and the transition between states follows a pre-defined flow. FSMs allow for a clear and well-structured operation of the system.

The defining element of an FSM is the state it is in. Once implemented in an FPGA, the latter is encoded on a vector of bits which informs the logic of the actions to take. The default way to encode states is to assign them with sequential integer values and then encode the latter in binary (000, 001, 010, 011, 100, etc). This limits the number of bits on which the state is coded but makes the FSM vulnerable to SEUs by allowing transitions between states upon bit flips (001 could become 011). To avoid unwanted transitions between states, it is possible to encode the state of the FSM using the one-hot encoding. If the FSM has n states, a vector of n bits will be created. At any given time, only one of these bits will be set to a logic '1' in order to represent the state of the FSM. If an SEU affects the FSM, it will be detected through an invalid state. The register holding the state of the FSM can be further triplicated in order to increase redundancy and allow the FSM to recover.

8.4 Firmware Design for the Irradiated FPGA

The firmware of the irradiated FPGA is designed to use a maximum of the available resources and transmit any error detected during run time. Specific firmware has been developed to test the CLBs, BRAMs, DSPs, and SEM. To optimize the design, floorplanning is used to divide the FPGA in ten regions running identical code. This prevents the compiler from placing elements in different sectors of the device and thus increase routing resources. Figure 8.6 depicts how the design occupies the FPGA: the image on the left represents the FPGA sectors labeled XaYb composed of CLBs in dark blue, BRAMs in red, and DSPs in green; the image in the middle highlights the resources used by the design; and the image on the right shows how the firmware developed for each function occupies the FPGA with CLBs in blue, BRAMs in yellow, DSPs in red, SEM in purple, and the communication protocol in green. The sampling rate of the various error detection tools is of 40 MHz. Figure 8.7 provides a logic diagram of the firmware designed to detect errors in the various components of the irradiated FPGA.

8.4.1 Configurable Logic Block

To test the CLBs, an algorithm performing logic operations and bit swapping on a 32-bit word is triplicated and connected to a majority voter to form a level-1 module. Three level-1 modules are connected together to another majority voter to form a level-2 module. The level-1 modules are used to detect errors in the CLBs and the level-2 modules to signal when the triplication technique fails to recover errors. The 32-bit word fed to the algorithm is generated by shifting a word every clock cycle. The same data is provided to all modules meaning that it is not susceptible to SEUs.

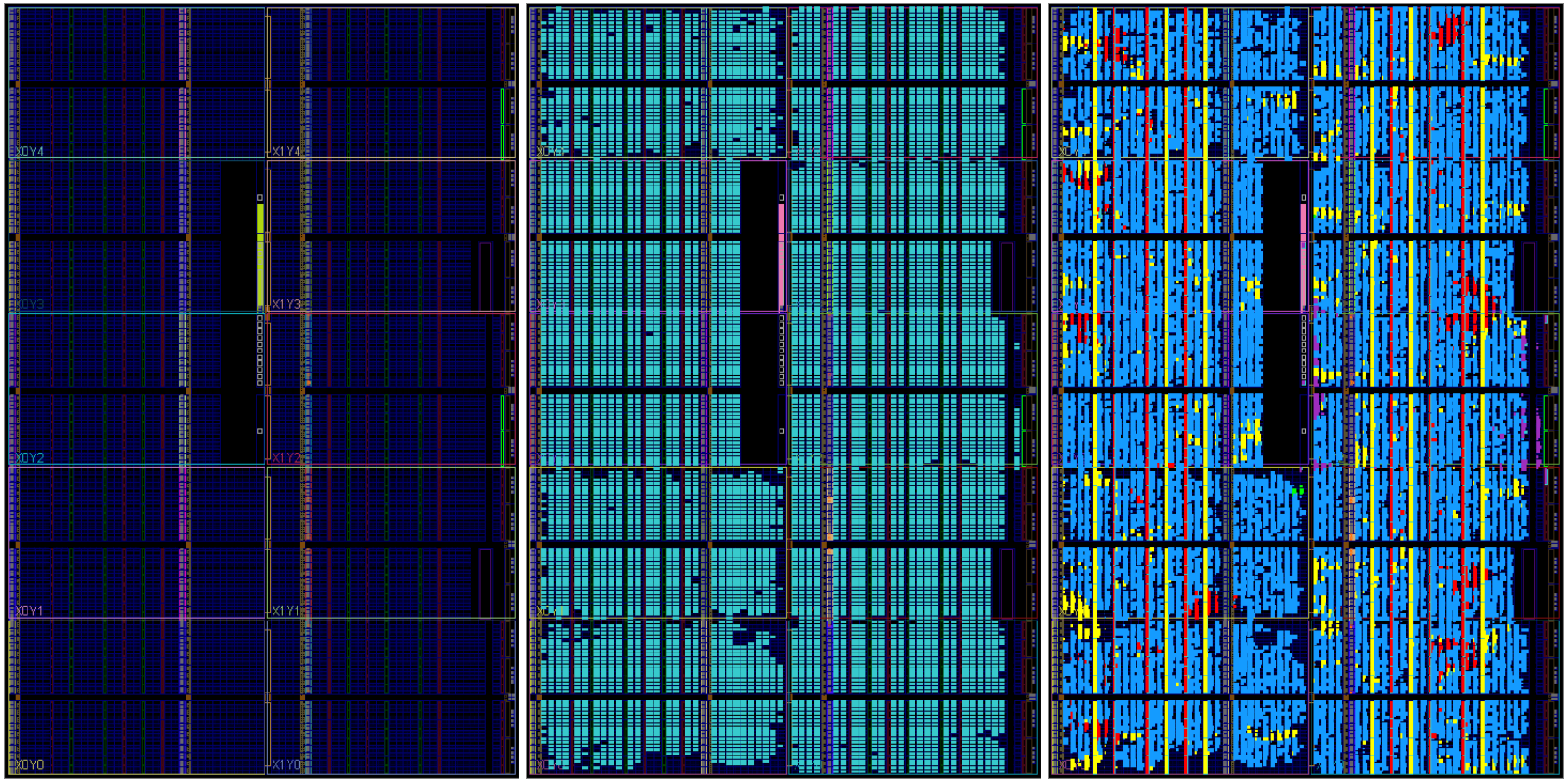

Figure 8.6 – Schematic view of the occupancy of the FPGA. Left: sectors of the FPGA labeled XaYb composed of CLBs in dark blue, BRAMs in red, and DSPs in green. Middle: highlight of the resources used by the design. Right: occupancy of the code developed to test the CLBs in blue, BRAMs in yellow, DSPs in red, SEM in purple, and the communication protocol in green.

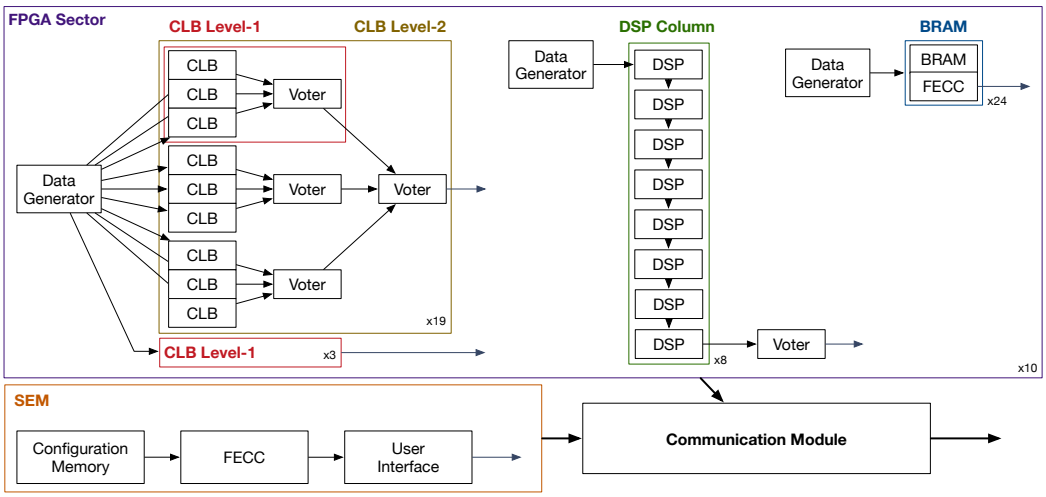

Figure 8.7 – Logic diagram of the firmware designed to detect errors in the irradiated FPGA.

Each of the 10 sectors of the FPGA contains 19 level-2 modules and 3 additional level-1 modules for a total occupancy of the FPGA of 80%. Within each sector the flags of the level-1 and level-2 modules are gathered and sent to the communication module which forwards information to the control FPGA located outside the irradiation area of the tests.

8.4.2 Block RAM

In Xilinx Virtex-6 FPGAs, some of the BRAM components are equipped with error detection and correction features. These can detect double bit flips and correct for single bit flips. During the tests, 90% of the available BRAMs, namely 240 components, are used and constantly written to and read from in order to detect SEUs happening inside the memory cells.

8.4.3 Digital Signal Processing

Each sector of the FPGA is composed of 48 DSPs grouped into 6 columns which share dedicated connections. Each column implements the same series of mathematical operations and is fed with the same 32-bit word generated within the sector. The results of the six columns are then compared to detect single, double, or triple failures according to the number of DSPs which returned a corrupted value.

8.4.4 Soft Error Mitigation

The SEM core included in the Xilinx Virtex-6 FPGA, when activated, continuously scans the configuration memory to detect corrupted bits. The core outputs various signals to report its status: a heartbeat, a detection flag which indicates an error has been identified, a correction flag which indicates an error has been corrected, and a critical flag which indicates the detected error cannot be recovered. In case the critical flag is set, a hard reset is needed to trigger a total reconfiguration of the FPGA from the external memory containing the design files.

8.5 Firmware Design for the Control FPGA

The signals collected by the irradiated FPGA are sent to a control FPGA located outside of the test area. This FPGA implements a set of counters that can be accessed by a computer through a dedicated core in the FPGA called ChipScope. The latter allows to monitor signals and send pulses inside the FPGA. It is used to monitor the communication protocol and read out the counters.

8.5.1 Communication Protocol

To connect the two boards, a High-Definition Multimedia Interface (HDMI) cable is used containing a total of eight wires. The data is always going from the irradiation zone to the control room with four of the wires used to send CLB, BRAM, and DSP errors and four used by the SEM core.

The four wires used to transmit errors from the various components implement a simple protocol and encode data on two 4-bit words clocked at 40 MHz. The first word is composed of all '1' to inform the control FPGA that a communication is happening, and the second word encodes the type of error: CLB level-1, CLB level-2, BRAM single, BRAM double, DSP single, DSP double, or DSP triple. The control FPGA samples the signals at a frequency of 160 MHz in order to perform oversampling and avoid error when decoding the data due to the difference in frequency between the clocks of the two boards.

The remaining four wires connected to the SEM core carry the heartbeat, detection, correction, and critical signals.

8.5.2 ChipScope Core

The Xilinx Virtex-6 FPGA implements a dedicated core called ChipScope that allows signals to be read out and set from a computer. In firmware, two versions of the core exist: one called Virtual Input/Output (VIO) which connects to signals and allows to modify or view them in real-time, and one called Integrated Logic Analyzer (ILA) which reads out the signals and provide a view of their evolution over time. In software, control windows can be created to handle those signals as depicted in Figure 8.8 in which the VIO control window is located on the left where signals in blue are those that are read out, and signals sin green are those that can be modified, and the ILA monitoring windows in showed on the right.

The VIO window provides an overview of the error counters and allows to reset them individually. It also controls a timer which is used to count the elapsed time and a "reconfiguration needed" signal which indicates that the irradiated FPGA requires a hard reset. The ILA monitor displays the incoming bits and the resulting information. The first bus labeled "HDMI" holds the four bits used to encode the error types which can be seen to transition from 0x0 to 0xF and 0x4. The 0xF signals the beginning of

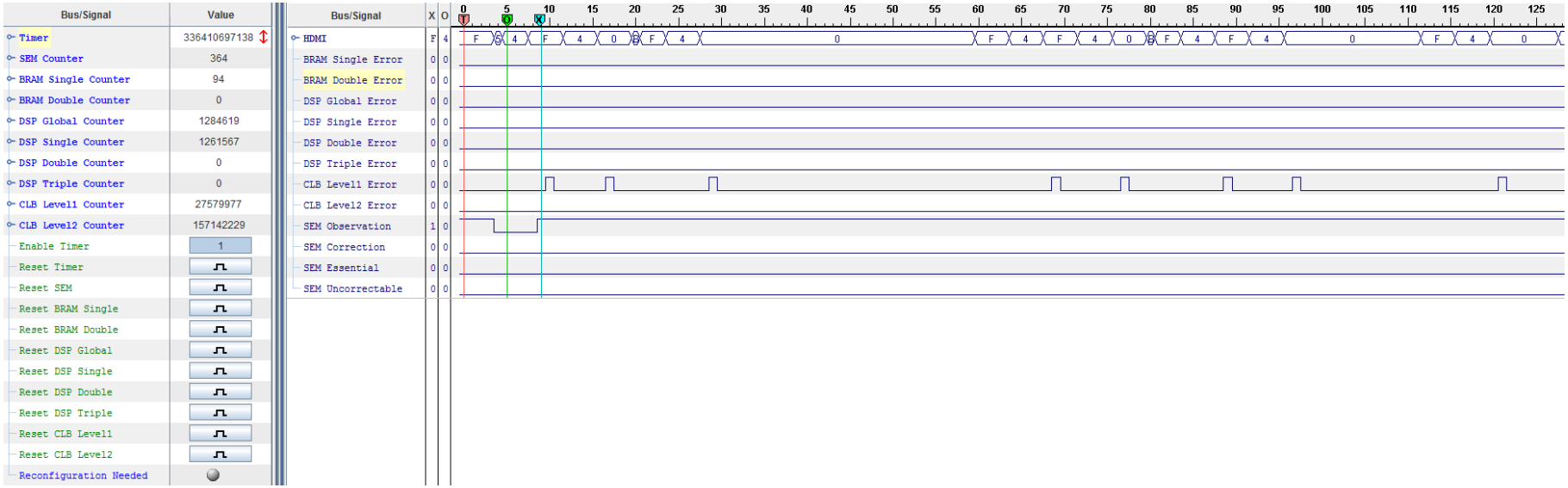

Figure 8.8 – View of the interface of ChipScope in which the VIO control window is located on the left where signals in blue are those that are read out, and signals in green are those that can be modified, and the ILA monitoring windows in shown on the right.

a transmission and the 0x4 indicates a CLB level-1 error which in turn raises one of the error flags. The last four buses are connected to the signals of the SEM core.

8.6 Irradiation Setup

The irradiation tests were performed at the CYCLONE110 of the Cyclotron Resource Centre at the Université Catholique de Louvain [58] using a proton beam with energies between 14.4 MeV and 62 MeV. The machine can deliver fluxes up to 2×10^8 particles cm^{-2} s^{-1} with a homogeneity of 10% within an 80 mm of diameter beam spot. The top picture in Figure 8.9 is a photograph of the beam line delivering the protons in the irradiation area in front of which degraders can be placed in order to change the energy of the particles. The FPGA was the only component on the OptoHybrid to be directly exposed to the beam.

Two OptoHybrids v2a were installed in front of the beam to mimic the setup of a superchamber in CMS. The bottom picture in Figure 8.9 is a photograph of the two boards being aligned in front of the beam using lasers. Next to the low voltage cables required to power the board, two HDMI cables are attached to each OptoHybrid v2a: one used to communicate with the control boards, and one used to send reset signals to the FPGA.

In the control room, two additional OptoHybrids v2a were used to collect data through the HDMI cable and count the errors. These values were read out using ChipScope on the two control boards separately from a computer.

8.7 Data Analysis

The parameters of the beam and disposition of the setup allowed for the study of the dependence of the SEU cross section with the beam energy, TID, and placement of the boards. The number of SEUs was recorded for various energies and fluxes of particles over the course of 24 hours. For every case, the cross section was computed from the counters as being

$$\sigma = \frac{N}{\phi t}, \tag{8.2}$$

where N is the number of recorded SEUs, t is the duration of the test, and ϕ is the flux of particles.

Over the duration of the test, the TID of the FPGAs increased, reaching up to 84 krad. The rad is a unit used to measure the dose absorbed by the object and is a function of the fluence, which is the flux integrated over time, and of the LET. In comparison, the TID accumulated by the GE1/1 detectors in CMS at the end of Phase II is estimated to be of the order of 10 krad. To accelerate the aging process of the FPGAs and accumulate statistics, both boards were exposed to fluxes up to 10 000

Figure 8.9 – Top: photograph of the beam line delivering the protons in the irradiation area in front of which degraders can be placed in order to change the energy of the particles. Bottom: photograph of the two OptoHybrids being aligned in front of the beam using lasers.

times higher than those that they will have to cope within the environment of CMS.

Finally, from the error counting in the CLBs and the two triplication levels, it is possible to extract the effectiveness of the latter method as mitigation technique of the SEUs.

8.7.1 Evolution of the SEU Cross Section with the Energy

The cross section with the configuration memory and the BRAMs were computed from the data collected using the results of the SEM core and the FEC of the BRAMs. Figure 8.10 displays the interaction cross section of particles as a function of the energy of the beam for the configuration memory resulting in recoverable (green) and critical (blue) errors on the left, and with the BRAMs resulting in single (blue) and double (orange) bit flips on the right.

Under 20 MeV, the number of errors is almost null with a cross section less than 1.14×10^{-8} cm^2, in both the configuration memory and BRAMs, with zero errors detected at the lowest energy of 14.4 MeV. This behavior has been observed and explained in previous irradiation tests [59, 60] and is due to the very limited range of the particles. At 10 MeV, protons have a LET of 3.48×10^{-2} MeV cm^2 mg^{-1}, and can only penetrate the first 10 μm in silicon due to the high energy losses they experience. They are thus not able to interact with the core of the FPGA but only with its packaging. At higher energies, the particles start to penetrate the silicon and recoverable errors are detected while critical errors are still rare events with a cross section of 1.88×10^{-8} cm^2 for protons of 40.8 MeV in the configuration memory. The latter are still rare as the total deposited energy is low and only reaches a limited range in the fabric of the FPGA. This generate single bit flips which result in recoverable errors. At even higher energies, above 40 MeV, the number of SEUs in the configuration memory keeps increasing to reach a cross section of 3.08 ± 0.21

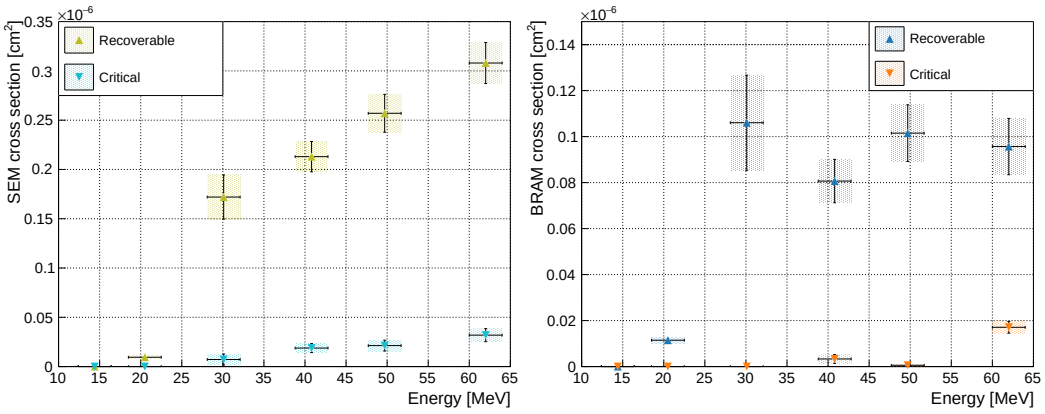

Figure 8.10 – Interaction cross section of particles as a function of the energy of the beam for the configuration memory resulting in recoverable (green) and critical (blue) errors on the left, and with the BRAMs resulting in recoverable (blue) and critical (orange) errors on the right.

$\times 10^{-7}$ cm^2 at 62 MeV while the error count in the BRAM seems to hit a plateau at a cross section of $1.02 \pm 0.12 \times 10^{-7}$ cm^2.

These results can be compared with tests performed for the electronics of the CSCs [59] for the same FPGA, which yielded values of $3.7 \pm 0.50 \times 10^{-8}$ cm^2 and $5.7 \pm 0.60 \times 10^{-8}$ cm^2 for interaction cross sections with the CLBs and BRAMs respectively at 55 MeV. The discrepancy in the values obtained for the configuration memory and CLBs is expected as the measured quantities differ. The configuration memory encapsulates the CLB and other resources present in the FPGA. It is thus expected that the interaction cross section of the former be larger than the latter. Exact numbers relative to the percentage of memory that is used to describe the CLBs are not disclosed by Xilinx, only an estimate of 6.4% is given [61]. This scaling factor can be applied to the cross section of the configuration memory at 62 MeV to find an approximative cross section for the CLBs of $1.98\ (\pm 0.14) \times 10^{-8}$ cm^2. For both the CLBs and the BRAMs, the obtained cross sections do not match within the error bars but however are at the same order of magnitude, only varying by a maximum of 50%.

Finally, a comparison with the reference values provided by Xilinx [62] can be done. The latter uses the irradiation facility of the Los Alamos Neutron Science Center to produce a neutron beam with an energy spectrum similar to the one measured in the atmosphere. The resulting neutrons range in energy from 0.1 MeV to 1 GeV with an unknown flux. A strict comparison cannot be performed, however, a validity check on the order of magnitude of the results is possible. The reference document states a cross section of $1.26 \pm 0.23 \times 10^{-14}$ cm^2 per bit of configuration memory. The FPGA used in the design of the OptoHybrid holds 43 720 728 bits resulting in a cross section for the entire FPGA of $5.51 \pm 0.99 \times 10^{-7}$ cm^2. This is compatible with the value obtained of $3.08 \pm 0.21 \times 10^{-7}$ cm^2 at energy of 62 MeV. The same can be done for the BRAM, for which a cross section of $1.14 \pm 0.21 \times 10^{-14}$ cm^2 per bit of memory is announced. This yields a cross section of $9.84 \pm 0.18 \times 10^{-8}$ cm^2 for the entire FPGA. This can be put in parallel with the value of $1.02 \pm 0.12 \times 10^{-7}$ cm^2 obtained in this study. The values for the configuration memory differ by less than 32% while the values of the BRAM are within the error margins. This validates the experimental process and results presented here-above.

8.7.2 Evolution of the SEU Cross Section with the TID

The OptoHybrids were initially exposed to a total of 33 krad over the course of various tests performed during the first 22h. The remaining two hours were used to irradiate the FPGAs with high fluxes and then perform a data taking run to measure a potential increase in interaction cross section: first, three runs that each provided 10 krad, followed by a run at 20 krad. Figure 8.11 displays the interaction cross section of particles as a function of the TID at an energy of 49.7 MeV for the configuration memory resulting in recoverable (green) and critical (blue) errors on the left, and

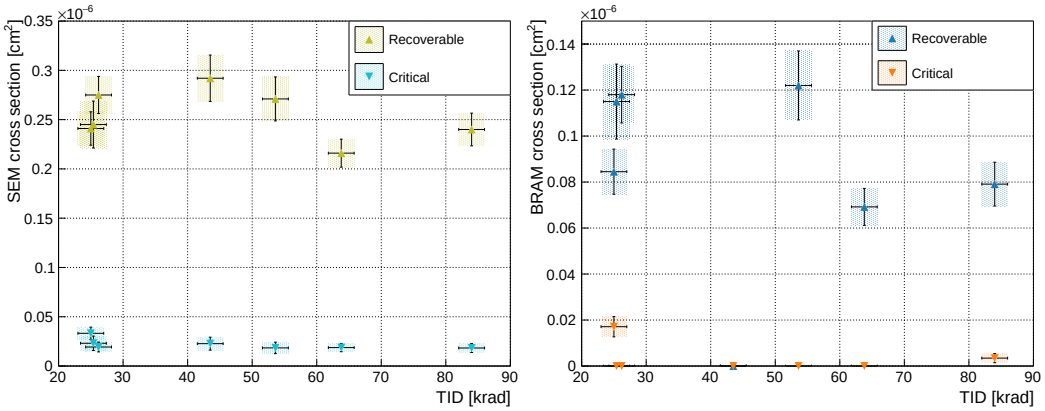

Figure 8.11 – Interaction cross section of particles as a function of the TID at an energy of 49.7 MeV for the configuration memory resulting in recoverable (green) and critical (blue) errors on the left, and with the BRAMs resulting in recoverable (blue) and critical (orange) errors on the right.

with the BRAMs resulting in recoverable (blue) and critical (orange) errors on the right.

From these results, no trend indicating an increase of the interaction cross section with the TID has been observed. This means that the FPGA remains operational at radiation doses that exceed those expected at CMS. Similar conclusions have been made from previous irradiation tests [59] were the TID reached 30 krad.

8.7.3 Evolution of the SEU Cross Section with the In-Beam Position

A comparison between the two boards was operated regarding the evolution of the interaction cross section with the energy. A different behavior is expected from the two FPGAs as the first one shields the second one at low energies. As particles pass through the former, they lose energy and might reach the threshold at which no SEU is produced in the latter. Figure 8.12 displays the interaction cross section of particles as a function of the energy of the beam and the position of the FPGA for the configuration memory resulting in recoverable (green and red) and critical (blue and purple) errors on the left, and with the BRAMs resulting in recoverable (blue and green) and critical (orange and red) errors on the right.

It can be seen that the interaction cross section with the configuration memory of the second FPGA is lower than the first one due to the shielding provided by the latter. An incident beam with less than 40 MeV will not impact the second OptoHybrid due to the low energy remaining in the beam after the first interaction. The effect becomes visible only at energies higher than 40 MeV where the slope of the curve quickly rises. The interaction cross section with the BRAM displayed a plateau in the first FPGA, which is also seen in the second one, with the only difference being in

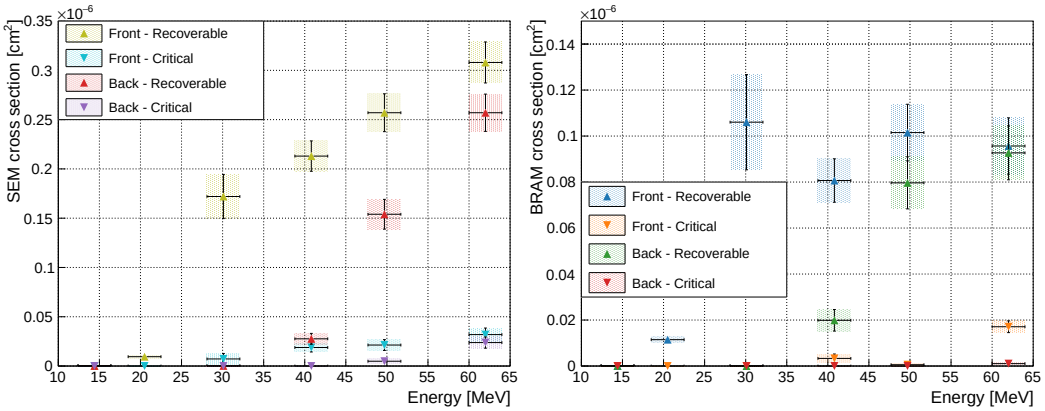

Figure 8.12 – Interaction cross section of particles as a function of the energy of the beam and the position of the FPGA for the configuration memory resulting in recoverable (green and red) and critical (blue and purple) errors on the left, and with the BRAMs resulting in recoverable (blue and green) and critical (orange and red) errors on the right.

the energy at which it appears.

To confirm the fact that no errors are expected in the second FPGA at energies below 40 MeV, the energy losses of the particles through the first FPGA can be computed using the LET. At 40 MeV, the LET is of 1.17 $\times$ 10^{-2} MeV cm^2 g^{-1}. From the specifications of the Xilinx Virtex-6 FPGAs, it can be found that the chip is made of a 1-mm-thick silicon layer, with density of 2.3290 g cm^{-3}, protected by a 1-mm-thick copper lid, with a density of 8.96 g cm^{-3}. The PCB of the OptoHybrid adds an additional 1.8 mm of FR-4 material with a density of 1.850 g cm^{-3}. Using these number, a total energy loss can be computed

$$\Delta E = LET \times \rho \times h, \tag{8.3}$$

where ρ is the density of the material and h its thickness. The combined energy loss is of 17.71 MeV for a particle of 40 MeV. This means it will have an energy of 22.29 MeV when reaching the second FPGA and thus be at the energy threshold to produce an SEU.

8.7.4 Interaction with the CLBs and DSPs

The errors occurring in the CLBs and DSPs are due to bit flips in the configuration memory which change the structure of the design. This means that these SEUs will be detected by both SEM and the triplicated logic. By computing the ratio of errors seen in the CLBs and DSPs, and by making the approximation that all errors produced in the configuration memory are reflected in either of these components, an upper limit can be given on the interaction cross section. From all the events recorded by the triplicated logic, 6.61% affected the DSPs and 93.39% the CLBs. By applying those ratios to the cross section of recoverable errors in SEM, this yields an

interaction cross section of 2.04×10^{-8} cm^2 and 2.88×10^{-7} cm^2 respectively for protons at 62 MeV.

8.7.5 Effectiveness of Triplication as Error Mitigation Technique

The effectiveness of this method was tested during irradiation at various particle fluxes. For this, two levels of triplication were compared: the level-1 which triplicates a basic algorithm, and the level-2 which triplicates the level-1 modules resulting in a double level of triplication. A failure of the level-1 only indicates that an error was detected in one of the module but corrected by the method. A failure detected at the level-2 is a sign that the triplication failed to correct an error in the system. Using these signals, a failure and success rate can be computed for this technique. Figure 8.13 shows the success (blue) and failure (orange) percentage of the triplication method according to the flux of particles when an SEU appears.

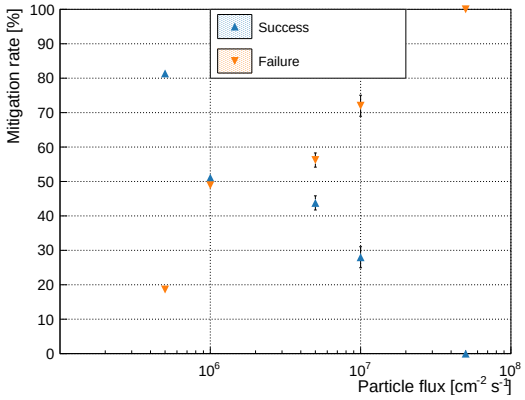

Figure 8.13 – Success (blue) and failure (orange) percentage of the triplication method according to the flux of particles when an SEU appears.

As can be observed, the failure rate increases with the particle flux due to the accumulation of errors in the design. The characteristic time needed to correct an error in the configuration memory is of a few milliseconds which at high fluxes is longer than the time between two SEUs. If two SEUs affect the same area of the logic, the triplication method will fail. At fluxes of 5×10^7 cm^{-2} s^{-1}, the SEU rate is so high that the triplication method cannot recover data at all. However, by extrapolating the data points, it can be seen that this technique will prove efficient and mitigate the SEUs at fluxes below 10^5 cm^{-2} s^{-1}.

8.8 Consequences for the GEM Upgrade Project

From the numbers obtained during the irradiation tests, it is possible to estimate the rate of SEUs that will be observed by the OptoHybrid in GE1/1 during run time in CMS. The latter will then define the requirements of the firmware in terms of implementation of mitigation techniques. The impact on the other GEM upgrade

projects, namely GE2/1 and ME0, can also be estimated and guide the decision that will be made regarding the design of their electronics.

8.8.1 Impact on the GE1/1 Upgrade Project

The dominant background in CMS is composed of photons and neutrons. Figure 8.14 shows the particle fluxes for the GE1/1 station as a function of the pseudorapidity for the LHC Phase II on the left, and the energy spectrum of the background particles in GE1/1 on the right.

The fluxes shown on the left are integrated over the whole energy spectrum and reach a maximum around 100 kHz cm^{-2} for photons and neutrons at the highest η value of 2.1. The OptoHybrid will be located in the lowest η region where fluxes reach 20 kHz cm^{-2}. These values have to be corrected according to the energy of the incident particles and the type of interactions they undergo with the FPGA. Photons do not interact with the silicon of the FPGA as the thickness of the chip ($<$ 1 mm) is four orders of magnitude smaller than their mean free path in silicon (12 cm) [64]. Experimental results [65] showed that electrons have an interaction cross section with silicon devices five orders of magnitudes lower than protons. These results coupled with the flux of electrons which is around 200 Hz cm^{-2} leads to a negligible contribution of the electrons with less than one event per year. Finally, neutrons can be described with the same values as those obtained for protons as they lead to similar rates of SEUs [60].

The 20 kHz cm^{-2} flux is thus reduced as the contribution from the particles below 20 MeV is ignored as it has been shown that they do not interact with the FPGA. Therefore, in order to set an upper limit on the number of SEUs, the interaction cross section is integrated with the flux between 20 MeV and 1 GeV. The cross section for particles with energy higher than 62 MeV is linearly extrapolated from the results shown in Figure 8.10. Table 8.1 lists the types of SEUs encountered along with their

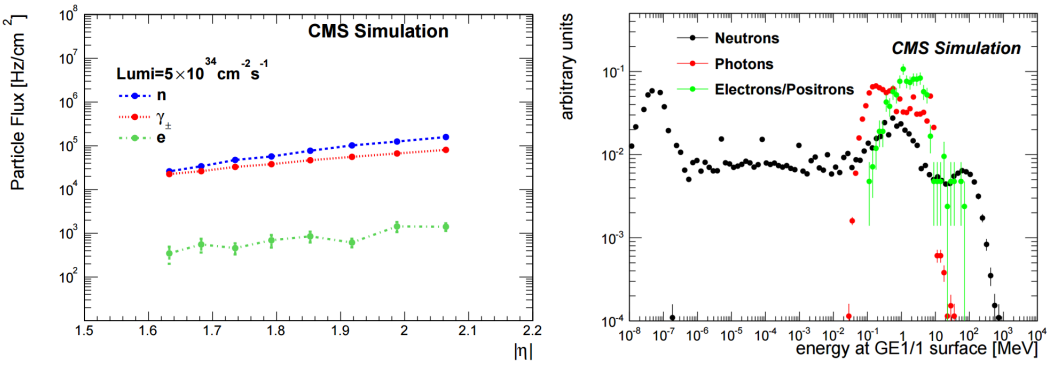

Figure 8.14 – Left: particle fluxes for the GE1/1 station as a function of the pseudorapidity for the LHC Phase II [63].Right: energy spectrum of the background particles in GE1/1 [63].

Event type	cross section	SEUs per day LHC Phase I	SEUs per day LHC Phase II
SEM - recoverable	3.08×10^{-7} cm^2	13.3	33.2
SEM - critical	3.19×10^{-8} cm^2	1.36	3.44
BRAM - recoverable	1.02×10^{-7} cm^2	4.38	11
BRAM - critical	1.71×10^{-8} cm^2	0.54	1.33
CLB	2.88×10^{-7} cm^2	12.42	31
DSP	2.04×10^{-8} cm^2	0.88	2.2

Table 8.1 – Types of SEUs encountered along with their respective interaction cross section and the resulting daily rate of errors in CMS per FPGA assuming that the LHC runs at nominal values during 24h.

respective interaction cross section and the resulting rate of errors in CMS per FPGA for the LHC Phase I ($\mathcal{L} = 2 \times 10^{34}$ cm^{-2} s^{-1}) and LHC Phase II ($\mathcal{L} = 5 \times 10^{34}$ cm^{-2} s^{-1}) conditions. The LHC Phase II conditions correspond to a fluxes 2.5 times higher than for LHC Phase I.

The recoverable errors do not pose a problem for the design as the triplication method can be used to successfully mitigate all errors in the CLBs and DSPs caused by the configuration memory, and ECC can correct those in the BRAM. Indeed, the maximum flux of 2 kHz cm^{-2} is inferior to the 100 kHz cm^{-2} the triplication can handle. The critical errors will require a partial reconfiguration of the FPGA, a process in which the defective bits are recovered from the external memory device and reloaded in real-time. Finally, the double bit flips in the BRAM can be mitigated by triplicating the storage of data in memory, which is allowed due to the large amount of available BRAMs.

8.8.2 Impact on the Firmware of the OptoHybrid

The presented results are upper limits on the cross section of SEUs for an FPGA which resources are fully used. The firmware of the OptoHybrid v2 used for the DAQ system of the slice test occupies 36% of the CLBs and 3% of the BRAMs. This means that the cross section of SEUs that will effectively impact the functioning of the design will be reduced by a factor of 3 for the CLBs and 33 for the BRAMs.

Although the total number of SEUs that will affect the logic per day during the slice test (LHC Phase I) will be on the order of 13, some modules are critical and need to be triplicated in order to mitigate errors. These modules are the tracking data readout and the trigger data formatting, as well as some registers which control the system. The remaining of the system is used for slow control operations and can be affected by SEUs without having an impact on the data taking and thus compromise

the run. The system can afford to wait a few milliseconds for SEM to correct the errors and normal operations to resume.

8.8.3 Impact on the GE2/1 and ME0 Upgrade Projects

The electronics that will be used for GE2/1 will be based on the one developed for GE1/1 due to the similarities between both detector designs. The latter differ only in size and segmentation and thus is number of VFAT3s required to be read out by the electronics. As GE2/1 will be placed in an environment less challenging than the GE1/1 in terms of particle rates, the values obtained for the latter can be used as upper limits for the former. The same electronics will thus be able to withstand the SEU rate by using triplication and the SEM core as described above.

ME0 on the other hand will be subject to a background rate of up to 1 MHz cm^{-2}, exceeding the limits of what the triplication module can handle in a Xilinx Virtex-6 FPGA. Therefore, new solutions need to be investigated to cope with the radiation and ensure the electronics is radiation tolerant. Furthermore, the TID of ME0 will reach 100 krad over the LHC Phase 2 period. Possible solutions include the upgrade to the Xilinx 7-Series (Artix-7, Kintex-7, etc), Flash-based FPGAs, or antifuse FPGAs. The Xilinx 7-Series uses a new 28 nm process for the manufacturing of their FPGAs which renders them less vulnerable to SEUs [62]. Flash-based FPGAs are also less susceptible to SEUs due to the persistent nature of their configuration memory. However they provide less resources and computational power than SRAM-based FPGAs (Xilinx, Altera, etc). Finally, antifuse FPGAs are immune to SEUs but are non-reconfigurable. Although future R&D is required to tackle this complex challenge, the GEM collaboration can benefit from work performed by other experiments at CERN.

The Xilinx Artix-7 is currently being investigated by the LHCb collaboration for the upgrade of their photodetectors. After exposure to 500 krad, corresponding to 2.5 times the maximum TID to which their electronics would be exposed, their analysis did not rule out the Artix-7 as possible candidate. Although an increase in current consumption was observed, the latter disappeared after one day of annealing. Similar studies are done by the CERN engineering department for the Microsemi Smartfusion2 FPGA. The latter is a Flash-based SEU-tolerant FPGA. By using the one-hot FSM and triplication techniques, results have shown that the FPGA is able to withstand a TID around 800 krad before failing. The ATLAs collaboration is also studying the SEU rates in the Xilinx Kintex-7 FPGA for the upgrade of their Liquid Argon calorimeter. The results show a mean time between SEUs of 150 s in the configuration memory and of 670 s in the BRAMs. From these results, they conclude that using triplication and scrubbing the Kintex-7 will be suitable for their system.

8.9 Conclusion

To characterize the behavior of the on-detector electronics used for the CMS GEM project when subject to radiation, two OptoHybrid v2a boards were exposed to proton beams during irradiation tests performed at the cyclotron of Louvain-La-Neuve. Custom firmware was developed for the FPGAs in order to compute the interaction cross section with the various components inside the chip and to study the effectiveness of mitigation techniques.

After analysis of the data, we obtained an interaction cross section of $3.08 \pm 0.21 \times 10^{-7}$ cm^2 and $1.02 \pm 0.12 \times 10^{-7}$ cm^2 for the configuration memory of the FPGA and the BRAM respectively, the two main sources of errors. These values are obtained for protons of 62 MeV, which behavior is similar to neutrons, the prominent source of radiation that affect the FPGA. Within the neutron spectrum, particles below 20 MeV do not interact with the silicon of the chip while rates of particles above 100 MeV are negligible in CMS. When transposed to the environment of CMS, these values yield a rate of 33.2 errors a day and 11 errors a day per FPGA respectively for the LHC Phase II. These errors can be mitigated using triplication techniques which, through extrapolations from results at higher rates, have shown to be efficient at low particle fluxes below 5×10^4 particles cm^{-2} s^{-1}.

Over the course of the tests, the FPGAs were exposed to a total of 84 krad, corresponding to a dose 8 times higher than what will be collected during the totality of the LHC Phase II. We did not observe any increase in the interaction cross section for the components of the FPGAs which continued to function correctly until the end of the tests.

From our studies, we conclude that, although subject to SEUs, the FPGAs used in the OptoHybrid design are suitable for the environment of CMS and will survive for the entirety of the Phase II run. The mitigation technique that we tested, namely triplication, is a suitable method to mitigate any error in the data accompanied by the use of the SEM core to correct the upsets that modify the configuration of the FPGA.

9 Summary

The Triple-GEM upgrade project aims at improving the performance of the muon spectrometer of CMS which will suffer from the increase in luminosity of the LHC in the coming years. The installation of the GE1/1 stations has been approved by the collaboration and will occur during LS2. To prepare for the integration in CMS, a small scale test will take place during the YETS 2016 involving the instrumentation of five superchambers in one endcap of the detector.

In the CMS GEM collaboration, the GEM DAQ group focuses on the design of the DAQ system of the detectors. Within this group, we were in charge of the firmware design of the on-detector and off-detector electronics, namely the OptoHybrid and GLIB boards. We developed both a flexible firmware architecture which provides the user with numerous monitoring and control options, and a web application which allows the user to control the system dynamically. After a successful integration of the components within a stable system, they were tested during two test beam campaigns in November 2014 and November 2015. These demonstrated that the designed DAQ system is able to handle a superchamber of GE1/1 detectors with ease. Furthermore, they enabled the recording of data that yielded various measurements of the efficiency and other significant parameters of the detector. During the analysis of the data, we performed various studies of the noise and efficiency levels against high-voltage, threshold, and trigger rates. We obtained a single chamber efficiency of 97% and tested the rate capability up until a trigger rate of 120 kHz with a resulting efficiency of 96%. The analysis we have performed shows that the GEM detectors equipped with the DAQ system we have designed meet the requirements of the project regarding the efficiency of the chambers and the trigger rate capability.

After the DAQ system had been tested extensively, we created an automated method designed to qualify the electronic components to be used in CMS. For this, we developed a data analysis tool which, for the first time, makes use of the full extend of the capabilities of the VFAT2 to perform the bias and analysis of the analog front-end. The procedure is used to detect faulty units, and characterize and align the analog response of the chip channel by channel. After analysis, we could quantify the noise of the electronics on the channels to 624 ± 62 e^- and determine that the dispersion of the channels was reduced from 0.10 fC and 0.02 fC. Furthermore, we designed a GEB testing board to test each position on the GEB and detect broken lines. The board relies on an FPGA coupled with an MCU to communicate with the OptoHybrid and test the integrity of the transmitted signals. Finally, the system as a whole is tested using a script which targets specific components of the architectures. Random

read/write operations to the GLIB, OptoHybrid, and VFAT2s are performed, tracking data is read out, and stress tests are done to push the system to the limit. The tools we created are used for the preparation of the integration test to select appropriate components and install testing facilities at CERN and in other associated research laboratories.

Finally, to characterize the behavior of the on-detector electronics when subject to radiation, we exposed two OptoHybrid v2a boards to proton beams during irradiation tests. For this, we developed custom firmware for the FPGAs in order to compute the interaction cross section with the various components inside the chip and to study the effectiveness of mitigation techniques. After analysis of the data, we obtained an interaction cross section of 3.08×10^{-7} cm^2 and 1.02×10^{-7} cm^2 for the configuration memory of the FPGA and the BRAM respectively, the two main sources of errors. These values are obtained for protons of 62 MeV, and when transposed to the environment of CMS yield a rate of 33.2 errors and 11 errors a day per FPGA respectively. These errors can be mitigated using triplication techniques which we have proven to be efficient at particle fluxes below 5×10^4 particles cm^{-2} s^{-1}. Over the course of the tests, the FPGAs were exposed to a total of 84 krad, corresponding to a dose 8 times higher than what will be collected during the totality of the LHC Phase II. We did not observe any increase in the interaction cross section for the components of the FPGAs which continued to function correctly until the end of the tests.

All the developments and studies performed on the DAQ system of the Triple-GEM upgrade project have allowed for the design and characterization of a system well suited for installation in CMS from the ground up. The system has been integrated and tested extensively under various conditions and proven to be stable and maintain efficiency in a harsh environments. Along with the firmware developments, we also created software which allows users to monitor and control the system to perform a wide range of tests on the detectors. Finally, the various data analysis have yielded results on the parameters of the detectors as well as the response of the electronics to signals and radiation. These studies are critical to ensure good operation during data taking in CMS.

Part III

Data Acquisition System Design

The backbone of every DAQ system is the integration of its components into a single coherent architecture. The latter must provide a seamless user experience abstracting every aspect of the design and be optimized for efficient data readout and transfer. While companies provide all-in-one systems that are well integrated and include intuitive user interfaces, they are rarely suitable for custom applications that need to interface with components from different vendors. Custom interfaces need to be developed to control the system and transfer data between the subsystems.

The architecture of a DAQ system is mainly defined by the way data is distributed between the components: it is either pulled or pushed. Data is pulled from one component to another when a transfer solely occurs upon request from the receiver, and is pushed when the decision to forward data is taken by the transmitter. As each component either pushes or pulls data, a bottleneck appears in the system due to limitations in the bandwidth, storage capacity, etc. This in turn can cause data losses or system failures if not handle properly. Classical approaches to DAQ system design make use of a central node which handles the system and coordinates the communication between the components. However, new technologies have enabled the development of new topologies which blur the lines, integrating components which fulfill multiple roles.

In this part, two DAQ systems are presented: one following the classical approach which uses a central node, and the other using recent developments in web technologies to create a modern and dynamic system. Both systems are described in details in order to highlight the innovative approach taken during the design of the second architecture.

All firmware and software developments exposed in Chapters 10 and 11, have been performed solely by the author unless otherwise specified.

A Classical Approach to Data Acquisition System Design

10

In the early development stages of the DAQ system for the GE1/1 detector, before the first version of the OptoHybrid was designed, a small prototyping setup was developed to readout a 10 cm $\times$ 10 cm Triple-GEM detector using VFAT2 Hybrids. With time, the setup was improved to use the GLIB and later on the OptoHybrid. Next to the data readout chain, the system also controls the high voltage and gas sources of the detector from a web interface.

In this chapter, we detail the evolution of the DAQ system developed to read out a small Triple-GEM prototype. We describe the technologies that have been used and the developments that were performed in order to integrate the components in the system.

10.1 The Experimental Setup

The experimental setup consists of a 10 cm $\times$ 10 cm Triple-GEM detector placed between two scintillators which coincidence, processed using a NIM crate, is used to produce triggers upon crossing of a cosmic muon. The GEM prototype is equipped with a two-dimensional readout board with 2 times 128 strips in each direction. Figure 10.1 is a photograph of the detector showing the four readout connector, the

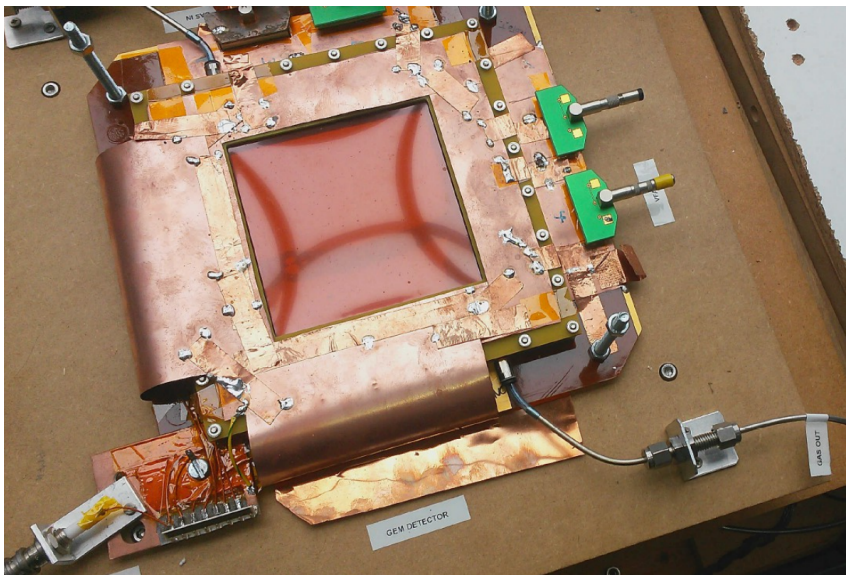

Figure 10.1 – Photograph of the 10 cm $\times$ 10 cm Triple-GEM prototype along with the high voltage resistive divider in the bottom-left corner, the gas supply in the bottom-right corner, and the four readout sectors on the right and top of the image.

high voltage resistive divider in the bottom-left corner, and the gas supply in the bottom-right corner.

The high voltage provided to the GEM and the photomultipliers attached to the scintillators is originating from a CAEN V6533N VME module. VME [66] is a crate standard that defines a high speed communication protocol and an infrastructure for the interconnect of the boards. A CAEN V1718 VME-USB bridge board is used to control the communication of the VME crate from a computer. This modules allows the latter to talk to any board in the crate by initiating requests. The gas flow is regulated by four HORIBA STEC SEC-N112MGRW mass flow controllers: three for the Ar, CO_2, and CF_4 gas bottles and one to monitor the mixture. These are controlled through a serial link communication by a computer. Figure 10.2 shows photographs of the VME high-voltage module on the left and the mass flow meters on the right.

10.2 The Infrastructure of the System

To control and monitor the GEM detector, high voltage, and gas systems, a custom solution was developed around a web interface. The latter allows the user to change parameters of the runs, visualize the evolution of the settings, access logs, etc. All this information is stored in a database located on the central computing node which also hosts the web server. The database is the communication interface between all elements. It stores requests, responses, statuses, etc. From the database, a second computer located near the experimental setup extracts the parameters that need to be set on the high voltage and gas system, and handles the communication protocol

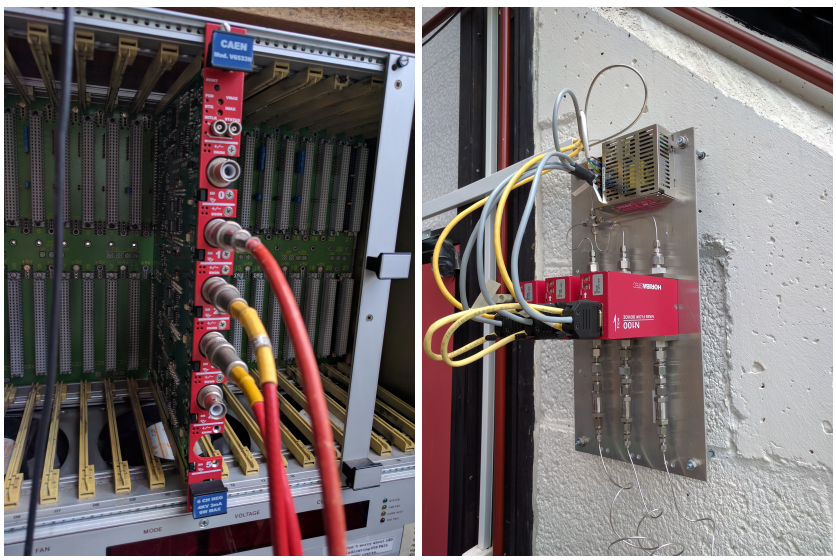

Figure 10.2 – Photographs of the CAEN V6533N VME high voltage module inserted in a VME crate on the left, and of the three mass flow meters connected to the gas bottles on the right.

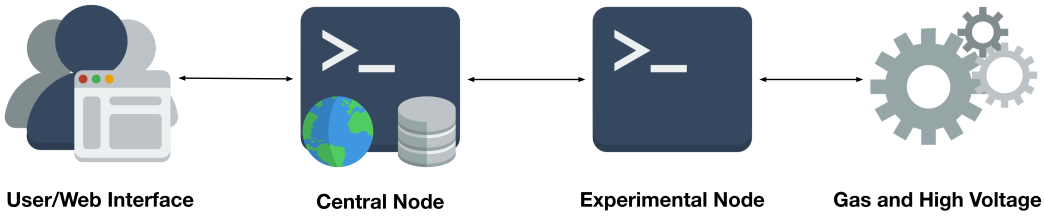

Figure 10.3 – Infrastructure of the system composed of the central node which hosts the database and web server, the experimental node which is located near the setup and controls the high voltage and gas, and the web interface which allows the user to interact with the systems.

with these two components. This infrastructure, shown in Figure 10.3, remained unchanged while the data readout of the GEMs evolved with time.

10.2.1 The Web Interface

Controlling and monitoring the elements of the system is done through a web interface which makes heavy use of JavaScript. The web server runs on NodeJS, which allows JavaScript to be executed on the back-end, and communicates with the client through Socket.IO, an implementation of the WebSocket technology which provides real-time messaging. The rendering of the page is done using Bootstrap for the design and AngularJS for the data handling. Using these technologies, the web interface provides a responsive and user friendly application.

The application is divided into three areas: the run control, high voltage monitoring, and gas monitoring. The run control is where parameters of the system are set when starting a new run. The system is constantly in run status even when the voltages are off or the gas disconnected. In order to differentiate data taking runs from stand-by runs, a logbook can be accessed to retrieve the parameters of the system. Figure 10.4 shows screenshots of the web pages controlling the experimental setup with the home page on the top, and the run page on the bottom. Upon the start of a new run, the user must specify the voltages he wants to apply to the detector and the flows of gas that will be provided.

To monitor the parameters during a run, the user can use the high voltage and gas pages to view an extended list of details. These list the current status of the registers and monitoring values which can be plotted. Figure 10.5 is a list of screenshots from the web application with the high voltage page on the top, and gas page on the bottom. Each page defines alarms that will modify the layout of the interface if error occurs such as incorrect values stored in the system, over heating, etc. The high voltage page allows to monitor the delivered voltages and currents as well as the temperature of the high voltage modules. The gas page displays the evolution of

Overview

High Voltage

Current values delivered by the HV module

Channel	Description	Voltage (V)	Current (I)	Power
0	GEM HV	1.6	0	Off
1	unused	1.7	0	Off
2	unused	1.3	0	Off
3	PM3 (below GEM)	0	0	Off
4	PM4 (below GEM)	3.8	0.95	Off
5	PM5 (above GEM)	0	0	Off

Runs

Information about the current run

Run
574
Start time
Jul 11, 2016 at 12:18:21
Started by
Alexandre Leonard
Comment
Stop gas flow

Gas System

Current values delivered by the Gas module

Channel	Description	Flow (ml/min)
1	Ar	0
2	CO2	0
3	CF4	0
4	Mixture	12

Runs

High Voltage

Channel	Description	Power	Voltage (V)	Maximum Current (uA)	Maximum Voltage (V)	Ramp Up (V/s)	Ramp Down (V/s)	Trip Time (s)
0	GEM HV	Off	0	1000	3800	100	100	1
1	unused	Off	3300	269	4000	20	250	1
2	unused	Off	900	1000	4000	100	100	1
3	PM3 (below GEM)	Off	850	1000	4000	100	100	1
4	PM4 (below GEM)	Off	850	1000	4000	100	100	1
5	PM5 (above GEM)	Off	1000	1000	4000	100	100	1

Gas system

Channel	Description	Flow (ml/min)
1	Ar	0
2	CO2	0
3	CF4	0

Figure 10.4 – Screenshots of the web pages controlling the experimental setup with the home page on the top, and the run page on the bottom.

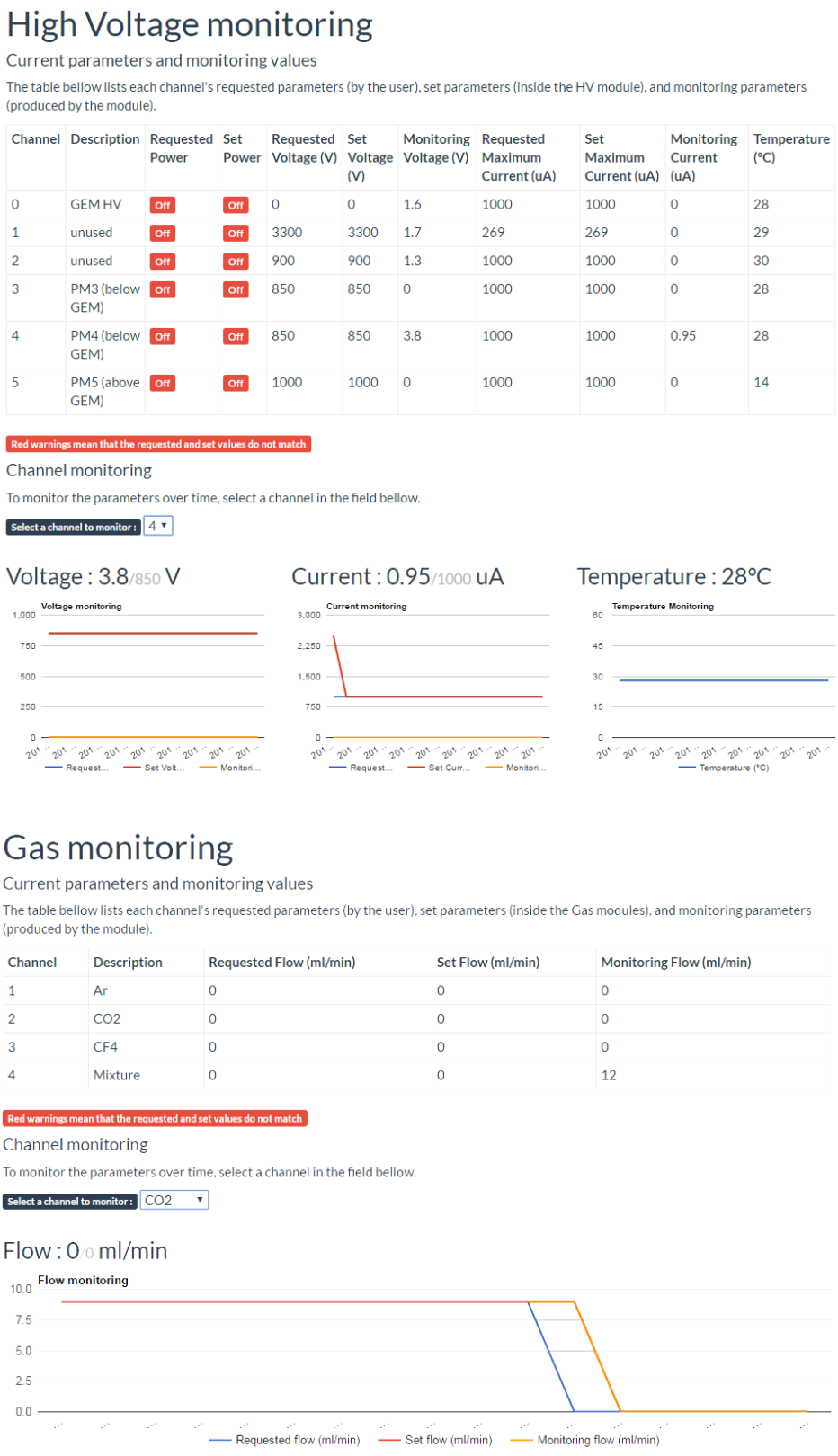

Figure 10.5 – Screenshots of the web pages monitoring the experimental setup with the high voltage page on the top, and gas page on the bottom.

the flows for the individual gases and the mixture over time.

Every action taken or value read is respectively stored or retrieved from the database. The web interface is the link between the user and the latter.

10.2.2 The Central Computing Node

The central computing node is located in a server rack far from the experimental setup. It holds the web server and the MySQL database. Its main role is to log every action and store the monitoring data. It is the bottleneck of the system which limits communication throughput as it has to perform various operations on the database for each request, which are relatively slow. As the central computing node delivers the web application to the client, it is also the main interface between the latter and the database it hosts, and thus the various systems in the readout chain.

10.2.3 The High Voltage and Gas Controllers

The communication between the database and the high voltage and gas systems is done through a computer placed in a rack near the experimental setup. Custom software was developed by the ULB DAQ team to communicate with the VME crate and the mass flow controllers. CAEN provides an integrated library to perform read/write operations on the crate which is used to apply the parameters stored in the database. Communication with the flow meters is done in the same application, which is thus in charge of loading the parameters in the registers and logging the monitoring values.

10.3 A First Prototype using the Xilinx SP601 Development Board

The first prototype of the DAQ system was developed using a Xilinx SP601 development board which is equipped with a Xilinx Spartan-6 FPGA (XC6SLX16-2CSG324) similar to the one used on the first version of the OptoHybrid, 1 Gb of DDR2 SDRAM, an Ethernet PHY, a USB-to-UART bridge, and a VITA 57.1 FPGA Mezzanine Card (FMC) extension header. The SP601 was linked to a VFAT2 Hybrid mounted on the Triple-GEM detector through a specially designed board which plugged onto the FMC connector. On the other side, the SP601 was connected to the computer present in the server rack close to the detector. Figure 10.6 provides a schematic overview of the system and its components.

10.3.1 The VFAT2 FPGA Mezzanine Card

The VFAT2 FMC, shown in Figure 10.7, is a small board which makes the connection between the SP601 and two VFAT2s, provides power to the boards, and is able to receive or send triggers through two LEMO connectors. It connects to the FMC header of the SP601 and forwards the signals of the VFAT2s to the FPGA.

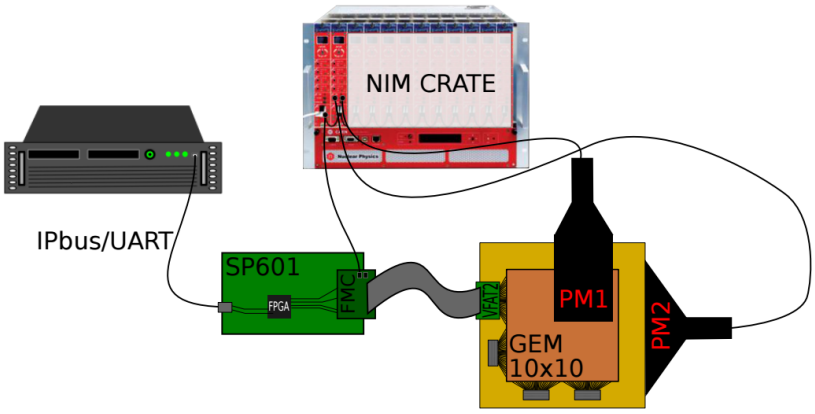

Figure 10.6 – Schematic view of the first prototype of the DAQ system of the 10 cm × 10 cm Triple-GEM prototype. The VFAT2 Hybrid is connected through an FMC module to the SP601 development board, in turn connected to a server through IPBus over UART. A NIM crate is used to perform the coincidence between the photomultipliers and send a trigger to the VFAT2 Hybrid through the FMC [67].

10.3.2 The Trigger and Tracking Paths

The trigger, defined as the coincidence of the top and bottom photomultipliers, is injected in the system through the LEMO connectors of the FMC board and redirected to the FPGA. It is then encoded and forwarded to the VFAT2s as a T1 signal. The trigger data received from the VFAT2s is put on one bit using a logic-OR and transmitted over the second LEMO connector of the FMC board and used to measure

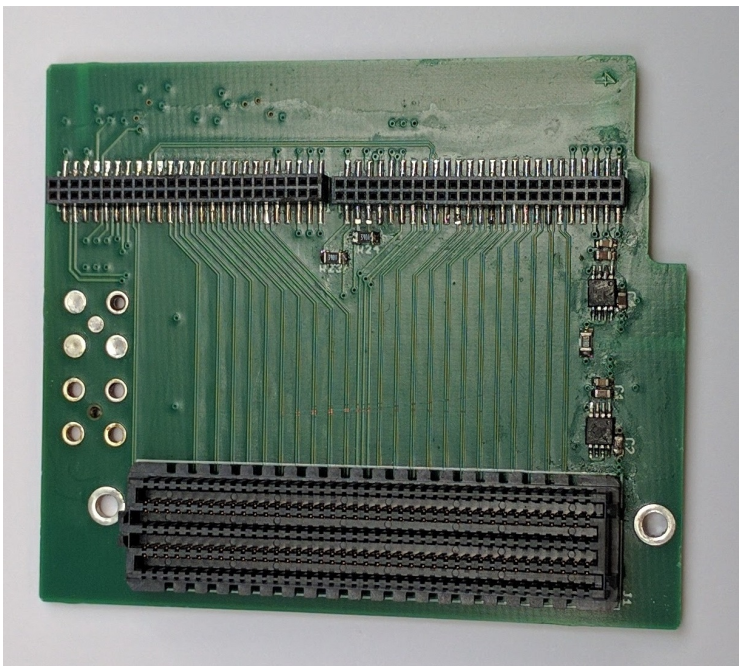

Figure 10.7 – Photograph of the VFAT2 FMC module which connects the SP601 development board to two VFAT2 Hybrids, provides power to the latter, and is able to receive or send trigger through two LEMO connectors.

timing delays.

The tracking data of the VFAT2s is decoded and stored in a local buffer to be later read out by the Microblaze soft core processor implemented on the FPGA.

10.3.3 The Microblaze Soft Core Processor

The Microblaze [68] is an embedded 32-bit processor developed by Xilinx which can be implemented on the fabric of the FPGA. As it is not hard-coded in the silicon of the FPGA, it is said to be a soft core. Figure 10.8 is a functional diagram of the architecture of the processor. The instruction buffer holds the instructions to be executed by the processor. They are decoded by the instruction decoder which forwards the operation to the registers or various logic units: the Arithmetic Logic Unit (ALU) which performs integer operations, the Floating Point Unit (FPU) which performs floating point operations, etc. The result of the operation is either stored in the registers or forwarded on the data bus. Finally, new instructions are fetched on the instruction bus to be stored in the instruction buffer according to the program counter value which points to the next address to execute. The data and instruction buses use a 32-bit addressing scheme to retrieve data from the elements they are connected to, and communicate with the components that attach to the processor.

To extend the computing power offered by the Microblaze processor, the data bus and instruction buses can be connected to other entities that are also implemented on the fabric of the FPGA. For example, a given address range of the instruction bus can be mapped to a memory controller which will interface with the onboard SDRAM to

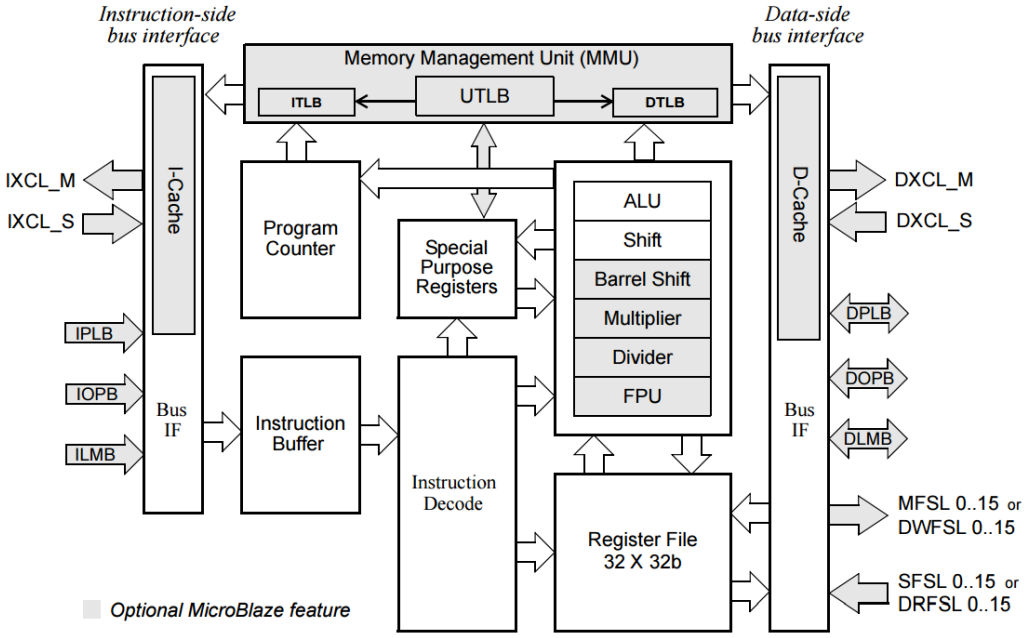

Figure 10.8 – Functional diagram of the architecture of the Microblaze soft core processor developed by Xilinx [68].

store the instructions of the program to run. The data bus on the other hand can be connected to Ethernet controllers, UART modules, etc to extended the IO capability of the processor. The way the instruction and data bus are mapped generates an address map file which describes the components attached to the processor.

To control the processor and its components, a toolkit is provided by Xilinx to compile C or C++ code for the Microblaze. Each component also comes with a driver which is used in conjunction with the address map file to access the dedicated functions. Additional modules can be developed and connected to the Microblaze interface, which has been done to control the buffer holding the tracking data of the VFAT2s. The module is composed of VHDL code to control the buffer on the fabric of the FPGA and of a driver written in C to provide high-level functions to access the data.

The final design of the Microblaze processor is composed of an external memory controller to store the program, an I2C controller to communicate with the slow control of the VFAT2, a dedicated module created to read out the tracking data, and a UART module to communicate with the computer located near the experimental setup.

10.3.4 IPBus over the Universal Asynchronous Receiver/Transmitter Protocol

The connection between the computer and the SP601 is done through UART over USB. The conversion between protocols is done through either drivers on the computer or a dedicated chip on the development board. The output of the chip is further connected to the Microblaze which receives interrupts whenever data is transmitted.

UART is a full-duplex two-wires protocol: RX which receives data, and TX which transmits data. The communicating parties are set on equal ground (no master or slave) and do not exchange a clock, which means they have to agree on a sampling frequency before hand. The latter is the limiting factor as uncorrelated clocks will have slightly different frequencies and thus induce errors when sampling at high speeds. Therefore, UART is limited to speeds of 1 kHz which means data is transmitted at 100 bytes per second. When the lines are idle, they are pulled up in order to indicate the presence of the other party. To initiate a communication, a start bit corresponding to a logic '0' is sent. Then follows the 8 bits of data with the LSB transmitted first. Finally, a stop bit equal to a logic '1' is sent, ensuring that during each communication a transition from '0' to '1' ensues. Figure 10.9 shows the

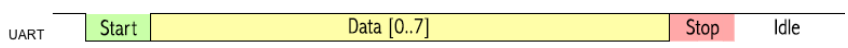

Figure 10.9 – Illustration of the sequence of bits in the UART communication protocol.

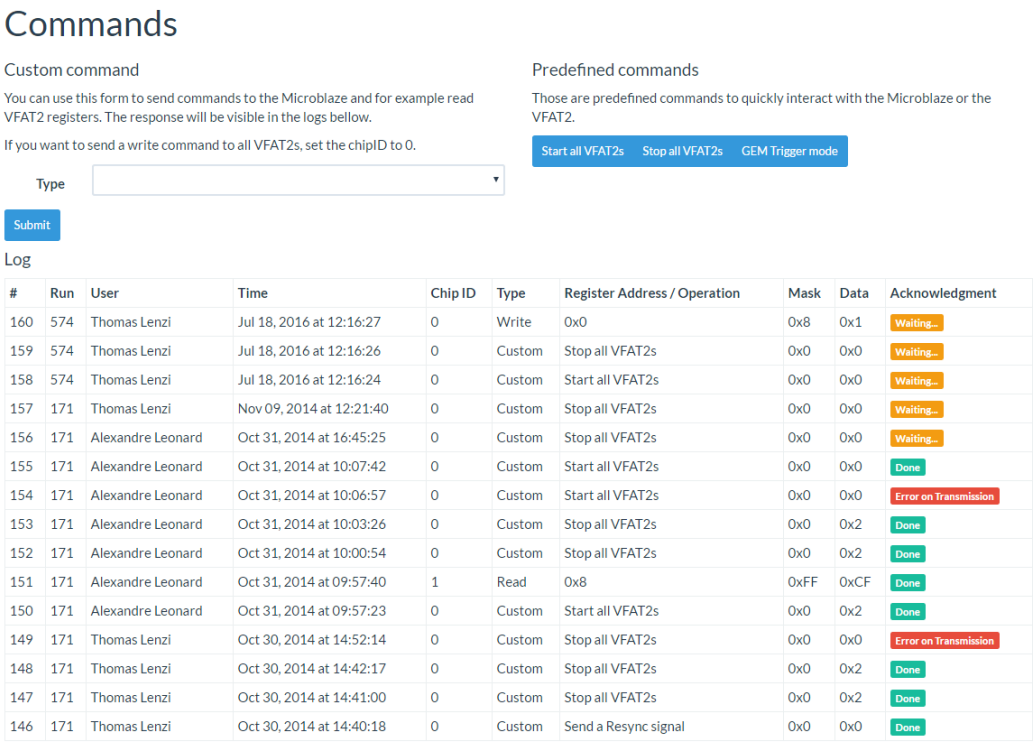

Figure 10.10 – Screenshot of the web interface used by the user to control the VFAT2 Hybrid through the Microblaze.

communication protocol with the relevant bit order.

To control the systems, it was chosen to use an implementation of IPBus over UART to be as close as possible to later developments. To this end, dedicated software has been designed for both the computer and the Microblaze. The software is flexible enough to either use a UART or an Ethernet interface, abstracting the transportation layer from the data. In the case of the UART implementation that was used, the 32-bit structure of the IPBus frame is decomposed in 8-bit packets. Each instruction is then recomposed by the Microblaze and forwarded to the corresponding module. The response is then sent back to the computer.

The software in charge of the communication with the SP601 that is running on the computer is constantly polling the database for new requests to transmit. It also pulls data out of the tracking data buffers that is then stored in the same database. The user can then visualize the events from the web interface or issue commands by polling the server. Figure 10.10 is a screenshot of the interface that allows the user to send commands and read the response from the system.

10.4 Upgrade of the System using the GLIB

When the GLIB became available, developments focused on the back-end electronics and software moved from an IPBus UART to an IPBus Ethernet implementation.

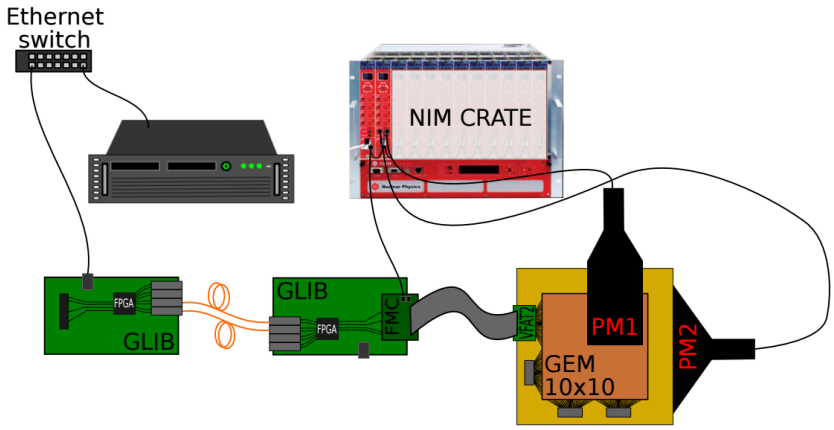

Figure 10.11 – Schematic view of the second prototype of the DAQ system of the 10 cm × 10 cm Triple-GEM prototype. The VFAT2 Hybrid is connected through an FMC module to a first GLIB, in turn connected to a second GLIB through optical fibers. A NIM crate is used to perform the coincidence between the photomultipliers and send a trigger to the VFAT2 Hybrid through the FMC [67].

As the GLIB is also equipped with an FMC connector, it was decided to use it as front-end module coupled with the VFAT2 FMC already used in the first prototype. The choice to separate the system instead of implementing the full design on a single GLIB was made to develop knowledge with respect to the use of optical links. Figure 10.11 displays the configuration of the system using two GLIBs as readout boards.

IPBus requests issued by the computer installed in the rack are sent over a switch to the first GLIB. The latter forwards the data to the second GLIB which further sends the commands to the VFAT2 over the FMC connector. The second GLIB also receives a trigger bit over the LEMO connector of the FMC board which is transmitted to the VFAT2. The tracking data is read out and pushed along with any slow control response, to the first GLIB. The main difference in the communication between elements is the increase in speed: Ethernet and the optical fibers have transfer rates of 1 Gbps and 3.2 Gbps respectively. This reduces the bottleneck in the system. The tracking data is stored in buffers in the first GLIB and read out at regular interval by the computer which saves it in database. In this version, the Microblaze is replaced by custom firmware developed from scratch to be more efficient and reduce delays.

10.5 Final System using the First Prototype of the OptoHybrid

The final system was developed when the first version of the GEB and OptoHybrid became available. The former is equipped with special connector which allow to plug in flexible cables to connect to remote VFAT2s Hybrids. This was used to test the GEB and ensure that signals were transfered correctly from the VFAT2 connected

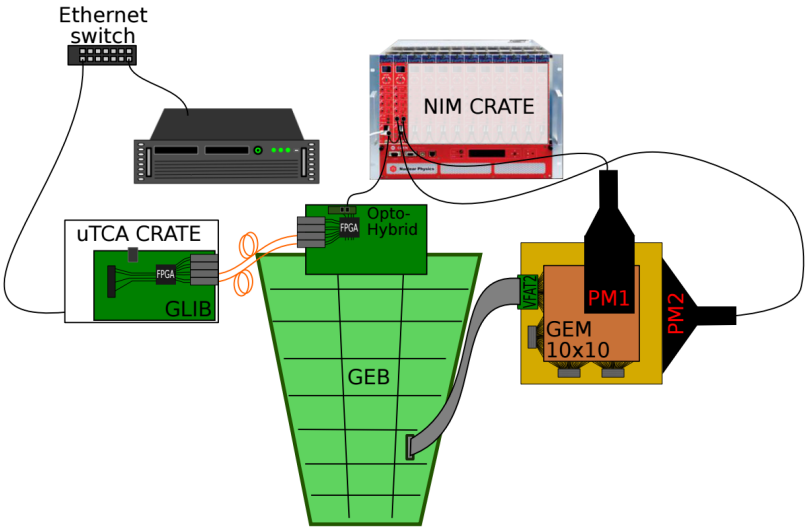

Figure 10.12 – Schematic view of the third prototype of the DAQ system of the 10 cm $\times$ 10 cm Triple-GEM prototype. The VFAT2 Hybrid is connected through the GEB to the OptoHybrid, in turn connected to the GLIB through optical fibers. A NIM crate is used to perform the coincidence between the photomultipliers and send a trigger to the VFAT2 Hybrid through the OptoHybrid [67].

to the small GEM prototype to the OptoHybrid. The trigger is delivered directly to the latter through a debugging connector which could also receive a clock in order to synchronize with the GLIB. The design of this system laid the foundations of the architecture used during the two test beams and allowed for testing with detectors. Figure 10.12 represents the system using the GLIB, OptoHybrid, and GEB to which a VFAT2 Hybrid is attached via a cable. This version can handle up to six VFAT2s unlike the previous ones which could only communicate with two VFAT2s through the VFAT2 FMC board.

10.6 Conclusion

The evolution of the readout system of the 10 cm $\times$ 10 cm GEM prototype followed the readiness of the hardware components. We first started with a Xilinx development board on which we implemented a Microblaze soft core processor as main control unit. Running custom software, the core interacted with a central computing node to propagate user requests to the front-end electronics. Additionally, a small mezzanine board was created in order to connect to the VFAT2 Hybrids. Although very functional, this system lacked in transfer speed as it used UART to transfer data at a maximum of 1 kHz.

The following generations relied on the GLIB at first, which was then backed up by the GEB and OptoHybrid. Dedicated firmware was design for all components in order to increase the speed of the system up to 1 GHz, the maximum speed of the Ethernet connection between the central node and the system. Furthermore, these

setups allowed us to gain experience using the newly created hardware, later on used during test beam campaigns.

All systems described in this chapter follow a classical approach to DAQ system design: the implementation of a star topology using a central node to control the traffic on the network. These systems allow for easy maintenance as all requests transit through a given path. However, this might cause bottlenecks in the architecture and generate system failures.

11 Developing Data Acquisition Systems using new WEB Technologies

In the past years, web technologies have evolved to the point that fully fledged applications can be developed. The main advantage of these applications is that they can run on any platform and do not require installation processes by the client.

In this chapter, we describe the architecture of a system making use of a Microblaze processor to run a web server and enable event driven communication between the front-end electronics and the control and monitoring application.

11.1 The Experimental Setup and System Architecture

The system is composed of the Xilinx SP601 development board to which the VFAT2 FMC is attached and connected to a VFAT2 Hybrid. Figure 11.1 is a photograph of the board showing the various components: the FPGA in the middle, the SDRAM on its right, the Ethernet connector and MAC controller in the bottom-left corner, and the FMC connector on the left. The SP601 communicates with any computer on the network through an Ethernet connection. Additionally, one of the two USB ports is used as debugging module and implements a USB-to-UART interface to stream information from the Microblaze to the exterior.

On the FPGA of the SP601, a dual-core Microblaze processor is implemented: one core handles the Ethernet connection and the other the communication with the VFAT2. The former runs a web server on top of the TCP/IP protocol in order to deliver the content of the web application used by the client to control and monitor the system ; the latter receives commands from the client and transfers them to the VFAT2. The whole system is contained in one board which can connect to various clients at the same time or to an external storage unit.

11.2 Web Server, WebSockets, and TCPSockets

A Microblaze processor is embedded on the FPGA of the SP601 and acts as web server to deliver web applications to clients which connect to the IP address of the board. Once the web application has been downloaded by the client over Hypertext Transfer Protocol (HTTP), a two-way communication is established over

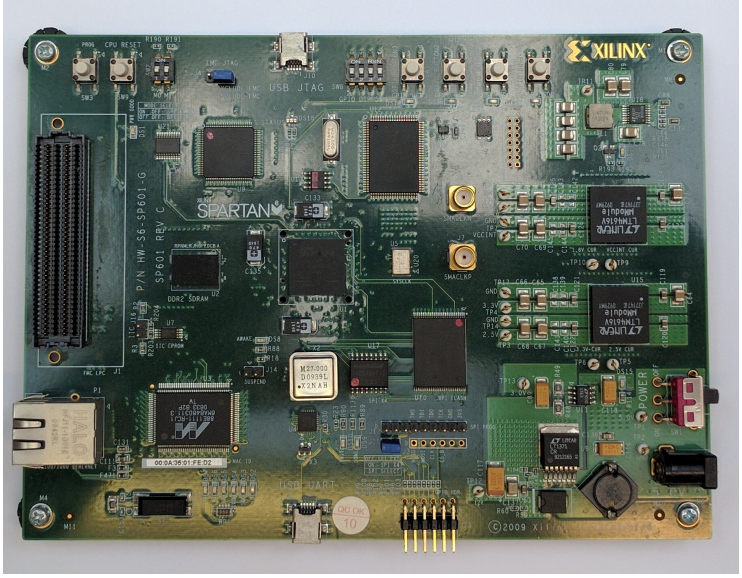

Figure 11.1 – Photograph of the Xilinx SP601 development board showing the various components: the Spartan-6 FPGA in the middle, the SDRAM on its right, the Ethernet connector and MAC controller in the bottom-left corner, and the FMC connector on the left.

WebSockets to transmit requests between parties. Figure 11.2 displays the exchange of information between the client and the server.

11.2.1 The HTTP Standard

When an Internet browser pulls a web page from a server, it sends an HTTP request containing the type of request and the path of the page. The type of request defines what the action of the client is. It can be GET, PUT, DELETE, etc which will respectively ask for a specific page, transfer content, delete content, etc. The path on the other hand can be a reference to a file to download or to a script to handle the new data. The request is encoded in ASCII characters and is thus human-readable. Figure 11.3 displays an example of an HTTP request and response in which the client requests a specific page and the server returns the content of that page.

After analyzing the HTTP request, the server performs the desired operation and returns a status code, information headers, and optionally data. The status codes are defined in the HTTP standard: 200 means the operation was successful, 404 means that the request file is missing, 500 means the server encountered an error, etc. The additional headers contain information regarding the type of content that is returned (text, image, etc), the encoding of the data, the data size, etc. Finally, the response ends with the raw data.

To each request, a single response is provided by the server. Moreover, the server does not initiate requests, meaning the communication has to be started by the

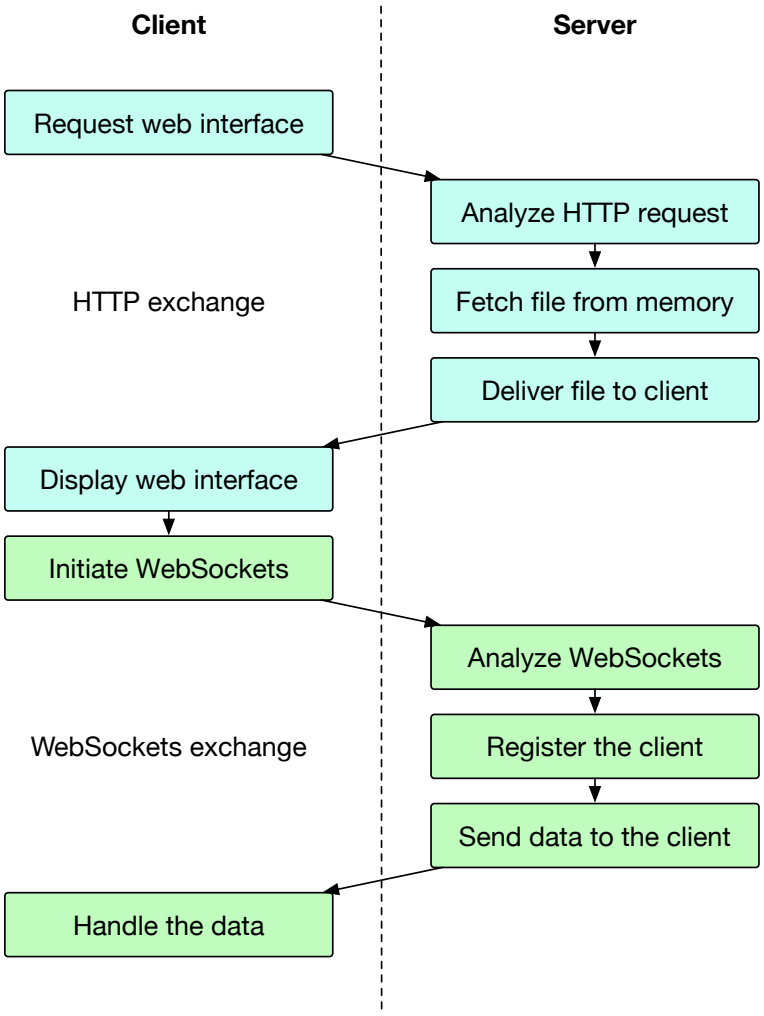

Figure 11.2 – Exchange of information between the client and the server.

HTTP request

```
GET /index.html HTTP/1.1
Host: www.web-daq.com
```

HTTP response

```
HTTP/1.1 200 OK
Date: Thu, 28 July 2016 10:36:02 GMT
Content-Type: text/plain; charset=UTF-8
Content-Encoding: UTF-8
Content-Length: 29
This is the web application !
```

Figure 11.3 – Transcript of an HTTP request and response in which the client requests a specific page and the server returns the content of that page.

client; the server can not push data to the client, but only provides data when probed (pulled).

11.2.2 The WebSockets Protocol

To allow the server to push data and enable event driven communication, the WebSockets standard was developed. WebSockets are sockets that can be used inside the Internet browser to communicate with a server following a given data format. When a connection is established, it remains open on both sides reducing the overhead on each data packet and decreasing the latency.

Upon connection from a client to the server, a dedicated HTTP request is sent informing the latter that the WebSockets protocol is being used. Figure 11.4 shows the request and response that are sent during the handshaking procedure between the client and the server. The request contains an upgrade header which informs the server of the change of communication protocol, along with a random key. The response from the server informs the client that the switch was made and returns a hash which prevents multiple connections from a same client. The server stores the information of the client in a list of active users until the latter disconnects or the connection times out. After this handshaking procedure, data can be sent between parties without restriction.

11.2.3 Handling Client Requests

The server on the Microblaze processor handles both HTTP and WebSocket requests, differentiating them using the HTTP header as selection criteria. It then forwards them to two dedicated functions according to the type of request.

For HTTP requests, the headers are extracted along with the type and path. Only GET requests need to be handled in this application, corresponding to the download of resources: web page, image, etc. In case the requested path is valid, the file is retrieved from memory and sent back to the user. Otherwise, an error 404 is generated to inform the user that the resource is not available. In normal oper-

Handshake request

```
GET /websocket HTTP/1.1
Host: www.web-daq.com
Upgrade: websocket
Connection: Upgrade
Sec-WebSocket-Key:
x3JJHMbDL1EzLkh9GBhXDw==
Sec-WebSocket-Protocol: chat
Sec-WebSocket-Version: 13
Origin: www.web-daq.com
```

Handshake response

```
HTTP/1.1 101 Switching Protocols
Upgrade: websocket
Connection: Upgrade
Sec-WebSocket-Accept:
HSmrc0sMlYUkAGmm5OPpG2HaGWk=
Sec-WebSocket-Protocol: chat
```

Figure 11.4 – Transcript of the request and response that are sent during the handshaking procedure of the WebSockets protocol between the client and the server.

ation mode, the client connects once to the server over HTTP to download the application and then switches over to WebSocket communication. The switching is done through JavaScript code present in the application that is executed by the client.

Once the application is downloaded, the connection with the server over WebSockets is established automatically. The server registers the client and event driven communication between parties can begin.

11.2.4 The TCPSockets Implementation

Although WebSockets provide an implementation of sockets in the web browser, they also define a format for the data being transmitted. This limits the field of application to the connection between a WebSockets client and server. A more generic approach uses the TCPSockets, a raw implementation of sockets in the web browser which is still in experimental phase. This technology enables the connection between any systems that support TCP sockets, including databases. A direct communication between the client and the database can thus be established in order to retrieve stored data, effectively bypassing the need for an intermediary server.

11.3 The Microblaze Processor

Upon startup of the Microblaze processor, the operating system is loaded into the SDRAM memory and executed. To implement the server and handle data, various components require initialization. Figure 11.5 shows the boot-up sequence of the system highlighting the configuration of the various subsystems. First, the caches are enabled allowing for an increased interaction speed between the processor and memory. Then, the interrupt controller is initialized and interrupts from the multiple modules are registered. One of them is the timer which triggers at given interval in order to execute control routines periodically. Finally, three complex modules are initialized: a Lightweight TCP/IP stack (lwIP) which handles the networking protocol,

Memory File System (MFS) which creates a virtual file system, and Mailboxes which handle the communication between processors.

11.3.1 Lightweight TCP/IP stack

lwIP is a light weighted implementation of the TCP/IP stack designed specifically for embedded systems with limited resources. The notion of stack is a reference to the Open Systems Interconnection (OSI) model which characterizes the communication functions of a system by subdividing it in layers, each performing a given task. The model contains seven layers: physical, data link, network, transport, session, presentation, and application. lwIP handles the data link, network, transport, and session layers, providing the user with an API to implement the presentation, and application layers.

The physical layer defines the electrical specifications of the communication medium and the way bits are encoded on the signal line. In this application, an Ethernet connection is used as interface between the SP601 development board and the network.

The data layer formats the raw data stream into packets and enables communication between two directly connected nodes on the network. It communicates directly with the hardware that composes the physical layer to transfer bits through a given network interface. It controls the latter and handles the flow of data. In the TCP/IP stack, the data layer is handled by the Media Access Control (MAC) and setup during the initialization of the network interface.

The network layer provides arbitration over a set of systems connected to a complex network where nodes are not directly connected to each other. It provides an address to each node in order for data to be transfered or routed. The prominent protocol used by networking systems is the Internet Protocol which provides a unique IP address to each node.

The transport and session layers are built on top of the network layer to add functions like error recovery, flow control, timeout, etc. A lightweight protocol implemented at this level is UDP which provides connectionless communication between nodes. They communicate by sending simple data packets without prior communication or acknowledgment system. This yields in a fast yet unreliable protocol. TCP on the other hand implements a full handshaking procedure which allows for features like congestion and flow control or data recovery. This however results in a heavyweight structure that can limit the range of target systems.

The presentation layer handles the encoding of the data which might vary between the transport and application layer. It is responsible for the encryption of the stream.

The application layer is the layer with which the software interacts. In this application, the software handles either HTTP or WebSockets packets.

```
------------------------------------------------
--                                            --
--            Microblaze Web Server           --
--                                            --
--    Developer:                              --
--    Thomas Lenzi - thomas.lenzi@ulb.ac.be   --
------------------------------------------------
Initializing the system...
> Enabling the caches...                      [OK]
> Initializing the interrupt controller...    [OK]
> Initializing the timer...                   [OK]
> Initializing MFS...                         [OK]
Setting up the network interface...
> Board IP: 192.168.1.10
> Netmask : 255.255.255.0
> Gateway : 192.168.1.1
> Initializing lwIP...
> Adding network interface...                 [OK]
> Starting network interface...               [OK]
Setting up the TCP protocol...
> Creating the TCP PCB...                     [OK]
> Binding to port...                          [OK]
> Starting the TCP server...                  [OK]

>> Request: HTTP                             [200]
>> Request: HTTP                             [404]
>> Request: HTTP                             [500]
>> Request: HTTP                             [Unknown]

>> Request: WebSockets Handshake             [Hi]
>> Request: WebSocket Data                   [Text]
>> Request: WebSocket Data                   [Binary]
>> Request: WebSocket Data                   [Unknown]
>> Request: WebSocket Disconnect             [Bye]
```

Figure 11.5 – Boot-up sequence of the operating system of the MicroBlaze highlighting the configuration of the various subsystems.

Using lwIP, a TCP server is started on the Microblaze and is set up to listen to any incoming connection on port 80, the default port used by Internet browsers. Upon reception of data, lwIP transfers the packet from its internal functions (transport and session layers) to user functions (presentation and application layers) which handle the HTTP and WebSockets requests.

11.3.2 Memory File System

When the client requests the download of a file from the Microblaze, MFS is used to retrieve the content from the SDRAM. MFS allows the operating system to maintain a virtual file system in the dynamic memory. This is required as the SP601 development board is not equipped with storage memory to host a fully fledge file system. Part of the SDRAM is allocated to MFS which initializes it using a precompiled image loaded upon startup. The image contains a tree of directories and documents which can be directly accessed from the software. MFS provides a range of functions to manipulate these files as if they where stored on a hard-drive.

11.3.3 Mailboxes for Dual Processor Architecture

To enable communication between the two processors which run different software, a mailbox system is implemented. Mailboxes are buffers which can be filled by either processors and are used to transmit small amount of data. Whenever a buffer is filled, an interrupt is sent to the processor in order to perform a read operation. The mailboxes are used to send instructions from the processor running the server to the one communicating with the VFAT2 and transferring back the response.

From the web application, the client can send requests to the server through WebSockets, which are transfered between processors and forwarded to the VFAT2. The returned data then follows the opposite path back to the client. Data can be broadcasted to multiple clients in order to update the information on each instance of the application at once.

11.4 Innovation of the Architecture

The architecture used in most systems such as the one described in Chapter 10 is shown in Figure 11.6. The whole system relies on the presence of the server which acts as the central node connecting the clients, the database, and the electronics. All transactions are required to transit through it, slowing down the performance of the whole infrastructure when scaling up. Furthermore, the client needs to constantly pull data from the server and the database to keep informed of the changes happening in the system or those submitted by other clients.

The system described in this chapter moves away from a star network in which the server is the central piece and tries to form a mesh network where every component can directly access all other nodes. By taking advantage of the increase in com-

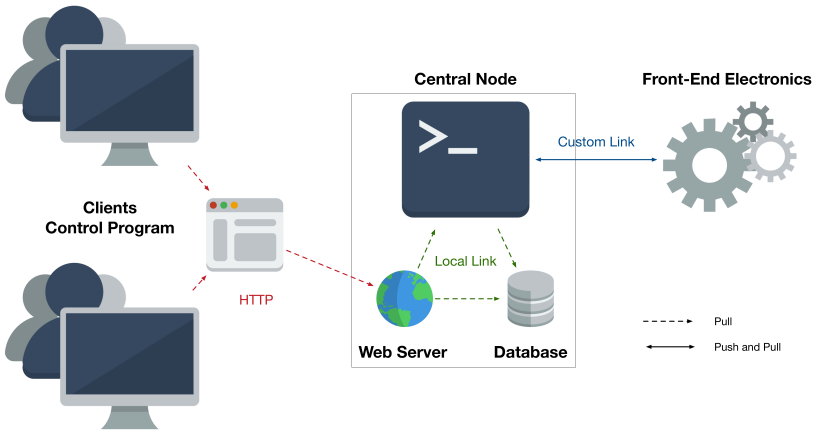

Figure 11.6 – Diagram of the architecture used in most DAQ systems in which the whole system relies on the presence of the server which acts as the central node.

putational power of microelectronics, it is possible to merge the functionalities of the server and the electronics in a single embedded system. Furthermore, developments in web technologies have enabled easy and reliable two-way communication between the client and the server and between the client and the database. The WebSockets technology provides response times up to ten times faster than classical solutions used to pull the server. The obtained system, which topology is shown in Figure 11.7, thus allows for faster inter-node communication, which releases the traffic on given data paths, and allows for more bandwidth and activity. Furthermore, the distribution of the application is centralized by the server which transfer the most up-to-date version of the web pages to every client, ensuring compatibility across time and platforms.

Through the two-way mesh network topology, each node can inform the systems of changes, effectively bypassing the need for a pull system which requests updated

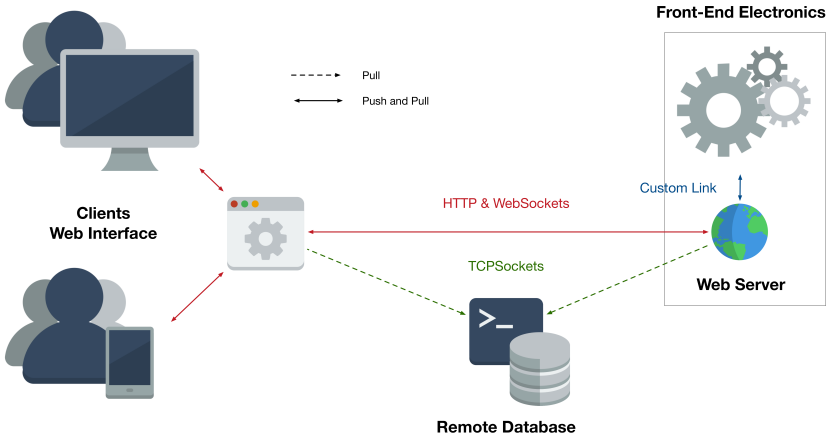

Figure 11.7 – Diagram of the new DAQ system which forms a mesh network where every component can directly access all other nodes.

data at constant interval and generates unwanted traffic on the network. The system becomes truly event driven with minimum delays. The bandwidth allocated to each client on the network for a given number of client N goes from O(1/N) to O(1). As updates are pushed by the server, only one data packet is sent to all clients at once. Each client is not required to handle the pull system, each using up bandwidth.

11.5 Response Time of Three DAQ Systems

To quantify the improvements made by the architecture described above, a comparison of the response time of three DAQ systems is performed, each with a different layout. First is the setup described in Section 10.3 which relies on a database to interface between the user control and the front-end ; second is the setup using IPBus to communicate with two GLIB boards, as detailed in Section 10.4 ; third is the setup described above which uses the SP601 with an embedded Microblaze. The comparison is done for two types of communications: a slow control request sent from the user control to the front-end electronics, and the readout of data coming from the front-end electronics towards the user.

11.5.1 Slow Control Request

In the first scenario, the user wants to send a slow control command to the front-end to change the content of a register. This process is illustrated in Figure 11.8 which depicts the nodes of the DAQ systems and the interactions they have in order to transmit the request from the user control to the front-end electronics. Blue arrows indicate that data is pushed to the next node, and red arrows that the data is pulled from the node. The first setup uses WebSockets between the web interface and NodeJS (A) which then writes the request into the database (B). The same database is pulled by a server (C) which transfers the data to the SP601 over UART (D) and then to the VFAT2 over I2C (E). By considering that every request is 32 bits long, except for the communication with the VFAT2 which is always 8 bits long, the total propagation time for a request in this system is of

$$\begin{aligned} t_{tot}[\text{ms}] &= t_A + t_B + t_C + t_D + t_E \\ &= 2.58 + 3342.8 + (2500 + 35.72) + 40 + 0.2 = 5921.3 \text{ ms}. \end{aligned} \tag{11.1}$$

The two major contributions to the response time are the communication between the NodeJS server and the database (B), and the fact that the server needs to pull data from the database (C). The latter is done every 5 s and thus results in a mean delay of 2.5 s before a request is seen as new, adding to the 35.72 ms needed to query the database. Finally, the UART communication (D) is the third largest contribution as it is one of the slowest protocols used to exchange data. It runs at 1 kHz with an overhead of 2 bits for every 8 bits it transfers, thus requiring a total of 40 bits to transmit the 32 bits of data. This yields 40 ms to transfer a request.

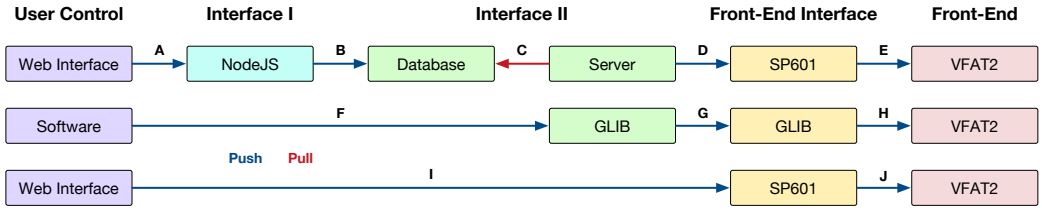

Figure 11.8 – Nodes of the DAQ systems and the interactions they have in order to transmit the request from the user control to the front-end electronics. Blue arrows indicate that data is pushed to the next node, and red arrows that the data is pulled from the node.

The same computation can be done for the second setup in which a Python script communicates with a first GLIB over IPBus (F). The request is then transferred to a second GLIB over optical links (G) and then to the VFAT2 over I2C (H). This setup yields a response time of

$$\begin{aligned} t_{tot}[\text{ms}] &= t_F + t_G + t_H \\ &= 0.44 + 10^{-6} + 0.2 = 0.64 \text{ ms}. \end{aligned} \tag{11.2}$$

As noted, the propagation time of the requests over the optical link is negligible due to the high transfer speed of 3.2 Gbps and the main contribution comes from the IPBus protocol. As for the I2C request (E) to the VFAT2, it is done with a reference clock of 100 kHz and requires the exchange of 20 bits to transfer the request, resulting in a propagation time of 0.2 ms.

Finally, the last system, in which a WebSocket request is sent from the web interface to the SP601 (I) and then transferred to the VFAT2 over I2C (J), totals a propagation time of

$$\begin{aligned} t_{tot}[\text{ms}] &= t_I + t_J \\ &= 2.58 + 0.2 = 2.78 \text{ ms}. \end{aligned} \tag{11.3}$$

This results in a response time over 2 000 times faster than the first setup while being around 4 times slower than the setup using IPBus. The latter result is due to the difference in execution environments in which the softwares are running. The Python scripts that perform the IPBus requests are executed in a terminal with almost no overhead. They have a high priority of execution within the operating system. The web interface on the other hand is executed inside a web browser which adds a layer on top of the system. Additionally, the IPBus protocol relies on UDP which is lightweight and fast in opposition to TCP used by WebSockets.

These measurements were performed for a single request and do not scale linearly when transferring multiple operations at the same time. Table 11.1 lists the response times of the three DAQ systems for 1, 100, 500, and 1 000 operations. When

Number of Operations	Database	GLIB	SP601
1	5921.3 ms	0.44 ms	2.78 ms
100	10 min	42.25 ms	22.91 ms
500	49 min	203.73 ms	103.34 ms
1 000	99 min	225 ms	125 ms

Table 11.1 – Response times of three DAQ systems for 1, 100, 500, and 1 000 operations.

bundling more than 100 requests together, the overhead generated by the IPBus protocol affects the performance of the transmission which reaches response times of 225 ms. This is not the case for the WebSockets which maintain a steady performance around 3 ms of response time for the communication with the server, thus yielding better results.

11.5.2 Data Readout

In the second scenario, a data packet of 224 bits is pushed from the front-end electronics to the user. This process is illustrated in Figure 11.9 which depicts the nodes of the DAQ systems and the interactions they have in order to transmit the data from the front-end electronics to the user control. Blue arrows indicate that data is pushed to the next node, and red arrows that the data is pulled from the node. In the first setup, the SP601 reads out the data from the VFAT2 at 40 MHz (E) and transfers it to the server over UART (D) which then stores it in the database (C). The latter is read out by NodeJS (B) when pulled by the web interface (A). The total propagation time for the data to reach the user is of

$$\begin{aligned} t_{tot}[\text{ms}] &= t_A + t_B + t_C + t_D + t_E \\ &= (2500 + 2.58) + 3342.8 + 35.72 + 280 + 5.6 \times 10^{-3} = 6161.1 \text{ ms}. \end{aligned} \tag{11.4}$$

The delay caused by the pulling of the database has been moved from the server to the web interface. Every five seconds, the latter sends a request to NodeJS which queries the database for new data packets. When new data is present, NodeJS transfers the information back to the web interface.

The second setup relies on the same pulling mechanism from the Python script to the GLIB which stores new events in a large buffer. The propagation time is thus of

$$\begin{aligned} t_{tot}[\text{ms}] &= t_F + t_G + t_H \\ &= (2500 + 0.44) + 10^{-6} + 5.6 \times 10^{-3} = 2500.44 \text{ ms}. \end{aligned} \tag{11.5}$$

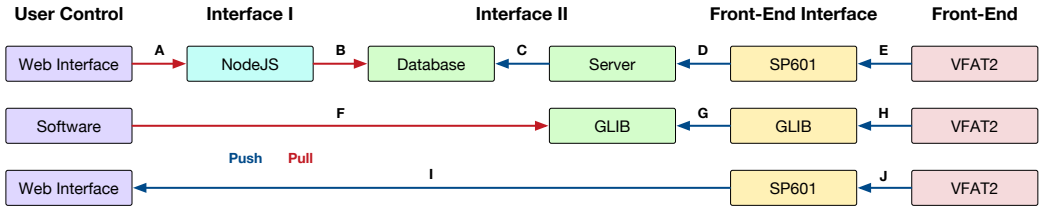

Figure 11.9 – Nodes of the DAQ systems and the interactions they have in order to transmit the data from the front-end electronics to the user control. Blue arrows indicate that data is pushed to the next node, and red arrows that the data is pulled from the node.

The fact that the communication between the nodes is not bidirectional is what adds delay in the system and reduces its efficiency.

In the last setup, the SP601 is able to push data to the web interface directly, which yields a response time of

$$\begin{aligned} t_{tot}[\text{ms}] &= t_I + t_J \\ &= 2.58 + 5.6 \times 10^{-3} = 2.58 \text{ ms}. \end{aligned} \tag{11.6}$$

By allowing the SP601 to inform the web interface directly upon reception of new data from the VFAT2, the response time of the system is reduced by a factor of 2 000 and 1 000 compared to the first and second setup respectively. A responsiveness of 2.58 ms is achieved which allows the system to be truly event driven.

11.6 Potential Application for the GE1/1 Upgrade Project

The web application developed by the GEM software group to control and monitor the DAQ system of the CMS GEM upgrade project (see Chapter 5) relies on Asynchronous JavaScript and XML (AJAX) requests. AJAX is used in JavaScript to create HTTP requests and acquire data from a server without having to refresh the page. It is in every point similar to normal HTTP requests, the only difference being that the returned data can directly be handled by the script of the application. AJAX is the predecessor of WebSockets in that it allows for the client to dynamically (without changing the content of the page) update data by pulling the server. However, it doesn't allow for bidirectional communication and thus relies on pull mechanisms to keep the data up to date. In the same circumstances as the tests performed in the previous section, a single AJAX request takes 4.1 ms to perform, almost twice the time required by WebSockets. After the server is asked to perform an operation through AJAX, it sends an IPBus request to the CTP7 which is the interface between the software and the electronics. The request is then forwarded to the components

through firmware, in a similar fashion as described in Chapter 6.

The control and monitoring application could benefit from the architecture described above by using the microprocessor available on the CTP7. As mentioned in Section 5.1.6, the CTP7 embarks a Zynq FPGA which is composed of programmable logic and two hard processors. The latter can run full-fledged operating systems on top of which one can program a web server. The client could thus directly connect and communicate with software running on the CTP7, as it was done using the SP601.

However, the question of scalability must be raised as a total of 12 CTP7s will be used to control the two endcaps of GE1/1 detectors. The architecture presented before where each board delivers the web application and provides the WebSockets communication becomes unpractical when multiple components have to be monitored at the same time. The user would need to open 12 instances of the same interface and operate the requests on each one individually. The solution comes from the flexibility of the WebSockets which allow to interact with multiple sources within the same web application. The client would connect to a server which sole functionality is to deliver the interface and scripts of the application. The communication with the CTP7s can then be established directly between the client and the electronics through WebSockets. This system centralizes the maintenance of the interface to a single server while still benefiting from the advantages provided by the architecture described above. This is illustrated in Figure 11.10 which is a schematic of the hypothetical architecture of the control and monitoring system of the DAQ system of the GE1/1 system using the techniques described in this chapter.

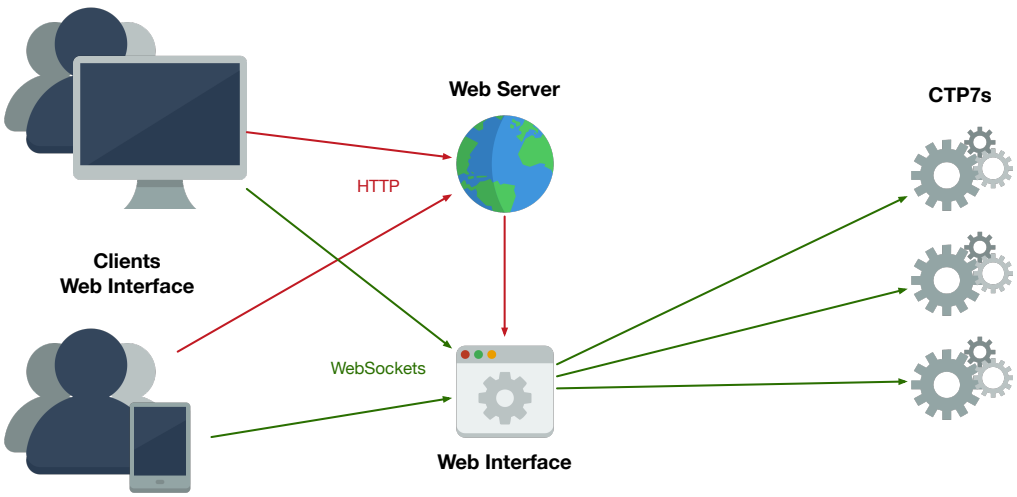

Figure 11.10 – Hypothetical architecture of the control and monitoring system of the DAQ system of the GE1/1 system using the techniques described in this chapter.

11.7 Conclusion

Using recent developments in web technologies, we were able to create an innovative data acquisition system architecture which provides higher flexibility and better utilizes the resources of the nodes in the network. Compared to classic designs which use pull/push techniques to transfer data, the new implementation offers event driven interactions, making use of the bandwidth only when necessary. The response time of the system was thereby improved by a factor of up to 1 000 compared to classic architectures and offered a reactivity of 2.58 ms compared to 2500.44 ms for the other tested systems.

Besides the increase in speed and optimization, the use of a web-based application offers greater portability on a variety of platforms without the need to maintain code for various systems.

This system is a proof-of-concept for a new type of data acquisition systems which relies on event based data analysis and control. It has the advantage to offer great flexibility and reduce the complexity of the overall system, while deeply combining the electronics readout and the server providing data.

Summary

12

Along with the advancements in microelectronics and the rise of new technologies, the architecture of DAQ systems will evolve to become more efficient and data driven. This is illustrated through the work we performed on the design of two DAQ systems: one controlling a small Triple-GEM prototype which followed the classical rules of DAQ design, and one that made use of novel technologies to create an innovative architecture.

The two systems implement radically different topologies to interconnect their nodes and handle the flow of data. The readout system of the 10 cm $\times$ 10 cm GEM prototype, which evolved over time according to the readiness of the hardware components, uses a star network to control the various subsystems. In this configuration, data is pulled from node to node and systematically transits through the central computing node which becomes the bottleneck of the system. The DAQ system relies on a updating scheme with a high frequency of pull requests in order to remain up to date. This in turn increases the traffic on the network and impacts the speed of the communications while diminishing the available bandwidth.

The second system implements a two-way mesh network based on novel web technologies which allows for event driven functionalities with response times as low as 2.58 ms. Helped by the increase in computational power of microelectronics, the system merges the functions of the central computing node and front-end electronics in a single device which runs a web server. The latter delivers the control and monitoring application to the client and pushes the data. Furthermore, a direct connection between the server or client and the database can be established to store information.

The architecture created to develop this proof-of-concept has shown to bring improvements on various fronts of data and system management. It provides a truly event-driven architecture which renders obsolete the need of a constant pull mechanism to ensure that all systems are up to date. This in turn reduces the traffic on the network with response times up to 1 000 times faster in comparison with other DAQ systems. The architecture also limits the number of subsystems in the DAQ system and thus maintenance by merging the functionalities of various components together. Finally, relying on a web application to control the system ensures a wider compatibility and lower maintenance cost than system dependent applications.

Conclusions

The work exposed in this PhD thesis details the contribution of the author to the development, testing, and characterization of the DAQ system of the Triple-GEM upgrade project for the muon spectrometer of CMS. This upgrade aims at improving the triggering and reconstruction performances of CMS after the increase in luminosity of the LHC foreseen for 2018-2019. Through the installation of a ring of GEM detectors in both endcaps of CMS, the CMS GEM collaboration proposes to increase the redundancy of the muon spectrometer and combine the results of the GEMs and the CSCs. To this end, a novel DAQ system was designed and tested.

In the early stages of the project, our developments focused on the design of a robust yet flexible firmware architecture for the OptoHybrid in charge of the communication between the analog front-end and the off-detector electronics. When tested during test beam campaigns at CERN and coupled to the VFAT2 front-end and GLIB AMC, we built a dedicated software suite in order to control and monitor the systems. Using the data recorded by the detector and the analysis we developed, we were able to obtain measurements of the efficiency of the detector according to various parameters. We obtained a single chamber efficiency of 97% and tested the rate capability up until a trigger rate of 120 kHz with a resulting efficiency of 96%. The analysis we have performed show that the GEM detectors equipped with the DAQ system we have designed meet the requirements of the project regarding the efficiency and the rate capability of the chambers.

Our developments then shifted towards the qualification of the electronic systems for their installation in CMS. To automate the characterization of the components, we developed a data analysis tool which makes use of the full extend of the VFAT2 capabilities to perform the bias and analysis of the analog front-end. The procedure is used to detect faulty units, and characterize and align the analog response of the chip channel by channel. We quantified the noise of the electronics on the channels to 624 ± 62 e^- and determine that the dispersion of the channels was reduced from 0.10 fC to 0.02 fC. To test the GEB, we created a small PCB board that can be used to test each position to detect faults in the manufacturing or soldering. The board relies on an FPGA coupled with an MCU to communicate with the OptoHybrid and test the integrity of the transmitted signals. Finally, the system as a whole was tested

using a Python script which targets specific components of the architectures. The tools and procedures we created are still in use for the preparation of the integration test to select appropriate components and install testing facilities at CERN and in other associated research laboratories.

Lastly, we studied the impact of radiation on the FPGA of the OptoHybrid by exposing it to proton beams. To this end, we developed custom firmware in order to compute the interaction cross section with the various components inside the chip and to study the effectiveness of mitigation techniques. After the tests were performed and the data analyzed by our analysis software, we obtained an interaction cross section of 3.08×10^{-7} cm^2 and 1.02×10^{-7} cm^2 for the configuration memory of the FPGA and the BRAM respectively, the two main sources of errors. These values were obtained for protons of 62 MeV which display the same behavior as neutrons, prominent background in CMS. From these results, we calculated the rate of errors that the FPGAs will encounter in the GE1/1 station. This yielded a rate of 33.2 errors and 11 errors a day per FPGA for the configuration memory and the BRAM respectively. Although elevated, these errors can be mitigated using the techniques we studied. Indeed, these showed to be efficient at particle fluxes below 5×10^4 particles cm^{-2} s^{-1}. Finally, a total dose of 84 krad was accumulated over the course of the tests, corresponding to a dose 8 times higher than what will be collected during the totality of the LHC Phase II. We did not observe any increase in the interaction cross section for the components of the FPGAs which continued to function correctly until the end of the tests.

In parallel to our work on the GEM upgrade project, we also investigated novel DAQ systems that make use of new technologies to create an innovative architecture. To this end, we implemented two different DAQ systems that use different topologies to interconnect their nodes and handle the flow of data. The first one, the readout system of the 10 cm $\times$ 10 cm GEM prototype, used a star network to control the various subsystems. In this configuration, data is pulled from node to node and systematically transits through the central computing node which becomes the bottleneck of the system. The DAQ system relied on a updating scheme with a high frequency of pull requests which increased the traffic on the network and impacted the speed of the communications. The second system implemented a two-way mesh network based on novel web technologies which allows for event driven functionalities. A direct connection between the server or client and the database was established to store information thus reducing the bandwidth used on the network and the strain on the central computing node. This architecture showed to bring improvements on various fronts of data and system management as it provided a truly event-driven architecture with a response time as low as 2.58 ms, 1 000 times faster than other DAQ systems.

List of Figures

1.1 Overview of the elementary particles as described by the Standard Model. Everyday matter is composed of the first generation of quarks and leptons to which two heavier generations are added. Gauge bosons, or force particles are the carriers of the interactions between other particles. The Higgs boson is the manifestation of the mechanism that gives mass to particles [4]. 7

1.2 Feynman diagrams representing the interaction terms of the Lagrangian. Left: interaction between a fermion and a boson. Right: self-interaction of gauge bosons. 9

1.3 Feynman diagrams representing the interaction of the γ, Z boson, and $W^{\pm}$ bosons to fermions. Top-left: coupling of γ to charged particles. Top-right: diagonal coupling of the Z boson to $SU(2)_L$ doublets. Bottom-left: non-diagonal coupling of the $W^{\pm}$ bosons to $SU(2)_L$ doublets. Bottom-right: self-coupling of the gauge bosons. 12

2.1 Schematic representation of the Large Hadron Collider and the four main experiments (ALICE, ATLAS, CMS, and LHCb) that analyze the produced collisions [8]. 15

2.2 Schematic diagram of the injection chain of the LHC composed of multiple smaller accelerators [16]. 16

2.3 Cross-section of a cryodipole of the LHC representing the two beam pipes equipped with superconducting NbTi coils surrounded by iron cold mass at 1.9 K [8]. 18

2.4 Monthly evolution of the peak delivered luminosity (left) for the four main experiments of the LHC and the corresponding integrated luminosity (right) [17]. 19

2.5 Projected instantaneous and integrated luminosity of the LHC throughout 2037 [4]. 20

3.1 The Compact Muon Solenoid detector and its subsystems installed at LHC [11]. 24

3.2 Disposition of the silicon pixels (PIXEL) and strips (TIB, TID, TOB, and TEC) modules inside the tracker of CMS [11]. 25

3.3 Tracking efficiency measured with a tag-and-probe technique, for muons from Z decays, as a function of the muon η (left) and the number of reconstructed primary vertices in the event (right) for data (black dots) and simulation (blue bands) [19]. 26
3.4 Material budget of the tracker expressed in terms of radiation length X_0 on the left and hadronic interaction length λ_I on the right [19]. . . 27
3.5 Layout of the ECAL showing the disposition of the crystals, modules, supermodules, and preshower [11]. 28
3.6 Longitudinal view of the HCAL showing the disposition of the barrel, endcaps, and outer calorimeters [11]. 29
3.7 Representation of the amplitude of the magnetic field and the field lines predicted on a longitudinal section of the CMS detector using simulations [21]. 30
3.8 Layout of the muon spectrometer of CMS depicting the three gaseous detectors in use: Drift Tubes (DTs) in orange labeled MBx, Cathode Strip Chambers (CSCs) in green labeled MEx/y, and Resistive Plate Chambers (RPCs) in blue labeled RBx and REx/y [11]. 31
3.9 Sketch of the drift cells highlighting the electric field lines converging towards the anode wire [11]. 32
3.10 Schematic view of a CSC detailing the different layers that compose it [11]. 33
3.11 Diagram of the RPCs double gap layout [11]. 34
3.12 Architecture of the CMS trigger system for both the calorimeter trigger and the muon trigger [22]. 35
3.13 Architecture of the DAQ system of CMS [23]. 36

4.1 Longitudinal view of a quadrant of CMS highlighting the location of the GE1/1 detectors colored in red within the muon spectrometer. DTs are represented in yellow, CSCs in green, and RPCs in blue [24]. 42
4.2 Left: sketch of the measurement of the bending angle with a pair of CSC and GEM chambers, illustrating the difference between lower and higher momentum muons. Right: $\Delta\phi = \phi_{GE1/1} - \phi_{ME1/1}$ distribution for 5 GeV/c and 20 GeV/c p_T muons demonstrating the discrimination power of the combined GEM-CSC system [24]. 43
4.3 Left: reconstruction efficiency of the integrated GEM-CSC trigger as a function of the simulated muon $|\eta|$, compared to the same for the Phase-I CSC-only algorithm. Right: level 1 muon trigger rates before and after the GE1/1 upgrade at a luminosity of 2×10^{34} cm^{-2} s^{-1}. MS1/1 denotes the first endcap muon station level 1 trigger in both cases, i.e. with CSC-only or with the combination CSC and GEM trigger information [24]. 44

4.4 The probability of reconstructing at least one muon candidate produced in the decay of a light long-lived particle decaying to a pair of muons $\gamma_d \to \mu\mu$ as a function of L_{xy}, the distance between the γ_d decay vertex to the beamline in the transverse plane. Standalone muon trigger L1Mu performance is compared to that of L1TrkMu, a trigger based on matching muon and track trigger candidates with the CMS Phase-II detector simulation [24]. 45
4.5 Left: scanning electron microscope picture of a GEM foil. Right: schematic view of the electric field lines (white), electron flow (blue), and ion flow (purple) through a bi-conical GEM hole [24]. 45
4.6 Principle of operation of a generic triple-GEM chamber and definition of drift, transfer, and signal induction gap regions within the detector [24]. 46
4.7 A pair of GEM chambers form a superchamber [24]. 47
4.8 Exploded view of the mechanical design of a Triple-GEM chamber [24]. 47
4.9 First CMS muon endcap station where the inner ring is equipped with 18 long (red) and 18 short (blue) triple GEM superchambers [24]. . . 49
4.10 Five generations of GE1/1 prototype chambers constructed and tested by the GEM collaboration in 2010-2014. The split figures for GE1/1-II and GE1/1-V demonstrate the evolution from construction using spacer frames to purely mechanical stretching of GEM foils without any spacers [24]. 50
4.11 Effective gas gain as a function of the rate of incident particles in a GE1/1-IV detector irradiated with a 22 keV X-ray source [24]. 51
4.12 Left: evolution of the efficiency as a function of the high voltage applied to the drift electrode for a GE1/1-III chamber operated with Ar/CO_2 at 70% and 30% respectively and for different offline cuts on the collected charge [24]. Right: time resolution of a small Triple-GEM detector as a function of the electric field in the drift gap [28]. 52
4.13 Left: distribution of the pulse charge measured over different strips. Right: evolution of the mean of the distribution of the pulse charge over multiple strips [24]. 53
4.14 Longitudinal view of a quadrant of CMS highlighting the location of the GE1/1, GE2/1, and ME0 detectors colored in red within the muon spectrometer. DTs are represented in yellow, CSCs in green, and RPCs in blue [28]. 54

5.1 Architecture of the GE1/1 DAQ system divided into two sectors: the on-detector electronics composed of the VFAT3 ASICs, GEB, and OptoHybrid, and the off-detector electronics with CTP7 and AMC13 mezzanines located in a microTCA crate [24]. 55

5.2 Schematic diagram of the VFAT3 showing the analog front-end on the left and the digital control and readout system on the right [24]. . . . 56
5.3 Pictures of a prototype of the GEB, a PCB of the same dimension as the GEM detector designed to route the signals of the front-end chips to the OptoHybrid [24]. . . . 58
5.4 Architecture of the radiation hard optical link setup integrating the GBT and Versatile Link projects [30]. . . . 59
5.5 Data format of a GBT frame [30]. . . . 60
5.6 Interaction between the GBTX ASIC and external front-end drivers through E-links [30]. . . . 61
5.7 Picture of the VTTx (left) and VTRx (right) optical transceivers of the Versatile Link project. . . . 61
5.8 Schematic representation of the architecture of a microTCA crate [32]. 62
5.9 Photograph of the CTP7 AMC developed by the University of Wisconsin used as back-end electronics [33]. . . . 63
5.10 Functional diagram of the AMC13 highlighting the external connections to the TTC system, backplane, and MCH of the microTCA crate [34]. . 64
5.11 Data format of the data packets of IPBus for read and write transaction types [34]. . . . 66
5.12 Architecture of the first prototype of the GE1/1 DAQ system with the VFAT2 ASIC, first version of the GEB, first version of the OptoHybrid, and GLIB AMC. . . . 67
5.13 Schematic diagram of the VFAT2 showing the analog front-end on the left and the digital control and readout system on the right [35]. . . . 67
5.14 Photograph of the VFAT2 Hybrids on which the VFAT2 ASICs are mounted to be interfaced with the GEB. . . . 68
5.15 Photograph of the OptoHybrid V1 equipped with a Xilinx Spartan-6 FPGA that concentrates the data from the six VFAT2 Hybrids and transmits it over four optical links to the off-detector electronics. . . . 69
5.16 Picture of the GLIB AMC equipped with a Xilinx Virtex-6 FPGA, four SFP optical cages, and two extension headers that can accommodate custom electronics and are used as back-end electronics [37]. . . . 70
5.17 Architecture of the GE1/1 DAQ system used during the test beam with the VFAT2 ASIC, second version of the GEB, second version of the OptoHybrid, and GLIB AMC. . . . 71
5.18 Photograph of the OptoHybrid V2a equipped with a Xilinx Virtex-6 FPGA that concentrates the data from the 24 VFAT2 Hybrids and transmits it to the off-detector electronics. . . . 71
5.19 Architecture of the GE1/1 DAQ system foreseen for the slice test with the VFAT2 ASIC, second version of the GEB, second version of the OptoHybrid, CTP7 AMC, and AMC13. . . . 72

5.20 Photograph of the OptoHybrid V2b that provides a testing platform for the GBT chipset. 73

6.1 Schematic diagram of the modules implemented in the firmware of the OptoHybrid regrouped by functionality. 76
6.2 Schematic illustration of the architecture of the system without Wishbone on the left and with Wishbone on the right. 77
6.3 Flow of operations for a custom calibration routine performed on the system without Wishbone on the left and with Wishbone on the right. . 77
6.4 Chronograph of two wishbone communications between a master and a slave for a successful write operation and a read operation resulting in an error. 78
6.5 Chronograph of two events generated by the T1 generator and encoded on the T1 lines for the VFAT2s: first two L1As (blue) are sent with a fixed interval, then a CalPulse (red) is generated followed by a L1A (green) after a defined delay. The T1 commands are encoded on three bits for the VFAT2s ("100" for L1As and "111" for CalPulses). 79
6.6 Diagram of the switching and throttling process for T1 commands highlighting the various sources of T1 commands (blue) and the involved modules (green) before the signal is sent to the VFAT2 chips (orange). 79
6.7 Chronograph of the serial decoding of an incoming data packet (blue) which is analyzed and placed in parallel busses (yellow) with the CRC check bit and the hit present bit. 81
6.8 Chronograph of an I2C communication. 82
6.9 Chronograph of an I2C communication performed with the VFAT2. . . 83
6.10 Addressing of the extended registers through two primary registers. . . 83
6.11 Illustration of a threshold scan representing the hit percentage as a function of the threshold of the VFAT2. 84
6.12 Illustration of an ideal latency scan representing the hit percentage as a function of the latency of the VFAT2. 85
6.13 Functionnal diagram of the Gigabit Transceiver X module used inside the Xilinx Virtex-6 FPGAs for high-speed serial communication [41]. . 86
6.14 Functional diagram of the architecture of the application used to control and monitor the GEM detectors. 89
6.15 Screenshots of the home page of the web application used to control and monitor the components of the DAQ system. 91
6.16 Screenshots of the page used to control and monitor the GLIB. 92
6.17 Screenshots of the page used to control and monitor the OptoHybrid. . 92
6.18 Screenshots of the page used to monitor the temperature and voltages of the OptoHybrid. 93
6.19 Screenshots of the page used to control and monitor the VFAT2. 93

6.20 Screenshot of the page of the web application used to perform calibration scans of the VFAT2s. 95
6.21 Screenshots of the pages of the web application used to send fast commands (left) and slow control (right) requests to the VFAT2s. . . . 95
6.22 Screenshot of the page of the web application used to do online data quality monitoring. 96
6.23 Screenshot of the page of the web application used to monitor the hit rates. 96
6.24 Monitoring page of the SPS indicating the presence of spills of particles (yellow). 97
6.25 Photograph of the GEM detectors used during the test beam and equipped with the full electronics system. 99
6.26 Measured gas gains as a function of the current through the high voltage divider for the different generations of GEM detectors and for Ar/CO_2 70:30 and Ar/CO_2/CF_4 45:15:40 [46]. 100
6.27 Photograph of the two detectors on the left placed in front of the beam transfer tunnel on the right side. 101
6.28 Plots of the noise level as a function of the VFAT2 threshold in terms of VFAT2 units, the decimal value written in the 8 bit register afterwards converted to a voltage level, for GEM0 on the left and GEM1 on the right.103
6.29 Plots of the noise level as a function of the VFAT2 threshold in terms of VFAT2 units with results obtained during the test beam on the left and with the improved noise cancelation on the right. 103
6.30 Plots of the ratio of hit events as a function of the VFAT2 latency in terms of BXs for a monostable pulse length of four clock cycles for GEM0 on the left and GEM1 on the right with muons. 105
6.31 Plots of the measured noise (green), measured efficiency (blue), and computed efficiency (orange) as a function of the high-voltage expressed in μA for GEM0 on the left and GEM1 on the right with muons. 106
6.32 Plots of the measured noise (green), measured efficiency (blue), computed efficiency (orange), and tracking data based efficiency (purple) as a function of the VFAT2 threshold in terms of VFAT2 units for GEM0 on the left and GEM1 on the right with muons. 107
6.33 Plots of the measured efficiency (blue), computed efficiency (orange), and tracking data based efficiency (purple) as a function of the trigger rates for GEM0 on the left and GEM1 on the right with pions. 108
6.34 Drift path (without diffusion) of electrons (orange) starting 90 μm above the GEM, in an electrostatic configuration (green) without charges on the Kapton surface on the left and with charges on the Kapton surface on the right [50]. 109

6.35 Plots of the cluster multiplicity for muons (green) and pions (blue) as a function of the VFAT2 threshold in terms of VFAT2 units for GEM0 on the left and GEM1 on the right. 110
6.36 Plots of the cluster size for muons (green) and pions (blue) as a function of the VFAT2 threshold in terms of VFAT2 units for GEM0 on the left and GEM1 on the right. 111
6.37 Plots of the beam profile for muons (green) and pions (blue) for GEM0 on the left and GEM1 on the right. 112

7.1 Two-dimensional graph plotting the evolution of the noise level for each channel as a function of the VFAT2 threshold in VFAT2 units. 117
7.2 Plots of the ADC counts, current values on the left, and voltage values on the right of each settable register of the VFAT2 as a function of the 8-bit value written in the register. 118
7.3 Diagram of the injection and calibration circuit of the VFAT2 highlighting the path of the calibration pulse towards a given channel [53]. . . 119
7.4 Plot of the hit-to-event ratio as a function of the calibration pulse height in VFAT2 units for a threshold of 25. 120
7.5 Left: collection of s-curve scans for various thresholds with corresponding current values. Right: plot of the calibration pulse height of the turn-on in VFAT2 units and current as a function of the threshold in VFAT2 units. 120
7.6 Plots of the s-curve scans for each channel with the minimum value of the channel registers on the left and the maximum value on the right. . 121
7.7 Left: plot of the s-curve scan for each channel with the aligned values for the channel register. Right: dispersion of the turn-on value for the non-aligned (orange and blue) and aligned configurations (green, scaled by a factor of 0.35). 122
7.8 Plot of the Equivalent Noise Charge (ENC) for each channel. 123
7.9 3D models and PCB layout of the GEB testing board. 124
7.10 Output of the qualification procedure developed to test the DAQ system in its entirety. 126

8.1 Simplified view of a slice composed of four LUTs on the left and eight registers on the right [54]. 130
8.2 Diagram of the DSP48E1 slices present in the Xilinx Virtex-6 family [55].131
8.3 Schematic of two slices (blue) connected to an element of the switching matrix which routes the signals (cyan) between components over the multitude of existing paths (gray). 132
8.4 Charge deposition by a charged particle inside a transistor affecting the state of the circuit [57]. 134

8.5 Logic diagram of a triplicated module coupled to a single voter (black) or to three distinct voters (gray). 136
8.6 Schematic view of the occupancy of the FPGA. Left: sectors of the FPGA labeled XaYb composed of CLBs in dark blue, BRAMs in red, and DSPs in green. Middle: highlight of the resources used by the design. Right: occupancy of the code developed to test the CLBs in blue, BRAMs in yellow, DSPs in red, SEM in purple, and the communication protocol in green. 138
8.7 Logic diagram of the firmware designed to detect errors in the irradiated FPGA. 139
8.8 View of the interface of ChipScope in which the VIO control window is located on the left where signals in blue are those that are read out, and signals in green are those that can be modified, and the ILA monitoring windows in shown on the right. 141
8.9 Top: photograph of the beam line delivering the protons in the irradiation area in front of which degraders can be placed in order to change the energy of the particles. Bottom: photograph of the two OptoHybrids being aligned in front of the beam using lasers. 143
8.10 Interaction cross section of particles as a function of the energy of the beam for the configuration memory resulting in recoverable (green) and critical (blue) errors on the left, and with the BRAMs resulting in recoverable (blue) and critical (orange) errors on the right. 144
8.11 Interaction cross section of particles as a function of the TID at an energy of 49.7 MeV for the configuration memory resulting in recoverable (green) and critical (blue) errors on the left, and with the BRAMs resulting in recoverable (blue) and critical (orange) errors on the right. 146
8.12 Interaction cross section of particles as a function of the energy of the beam and the position of the FPGA for the configuration memory resulting in recoverable (green and red) and critical (blue and purple) errors on the left, and with the BRAMs resulting in recoverable (blue and green) and critical (orange and red) errors on the right. 147
8.13 Success (blue) and failure (orange) percentage of the triplication method according to the flux of particles when an SEU appears. 148
8.14 Left: particle fluxes for the GE1/1 station as a function of the pseudorapidity for the LHC Phase II [63].Right: energy spectrum of the background particles in GE1/1 [63]. 149

10.1 Photograph of the 10 cm $\times$ 10 cm Triple-GEM prototype along with the high voltage resistive divider in the bottom-left corner, the gas supply in the bottom-right corner, and the four readout sectors on the right and top of the image. 157

10.2 Photographs of the CAEN V6533N VME high voltage module inserted in a VME crate on the left, and of the three mass flow meters connected to the gas bottles on the right. 158
10.3 Infrastructure of the system composed of the central node which hosts the database and web server, the experimental node which is located near the setup and controls the high voltage and gas, and the web interface which allows the user to interact with the systems. 159
10.4 Screenshots of the web pages controlling the experimental setup with the home page on the top, and the run page on the bottom. 160
10.5 Screenshots of the web pages monitoring the experimental setup with the high voltage page on the top, and gas page on the bottom. 161
10.6 Schematic view of the first prototype of the DAQ system of the 10 cm $\times$ 10 cm Triple-GEM prototype. The VFAT2 Hybrid is connected through an FMC module to the SP601 development board, in turn connected to a server through IPBus over UART. A NIM crate is used to perform the coincidence between the photomultipliers and send a trigger to the VFAT2 Hybrid through the FMC [67]. 163
10.7 Photograph of the VFAT2 FMC module which connects the SP601 development board to two VFAT2 Hybrids, provides power to the latter, and is able to receive or send trigger through two LEMO connectors. 163
10.8 Functional diagram of the architecture of the Microblaze soft core processor developed by Xilinx [68]. 164
10.9 Illustration of the sequence of bits in the UART communication protocol.165
10.10 Screenshot of the web interface used by the user to control the VFAT2 Hybrid through the Microblaze. 166
10.11 Schematic view of the second prototype of the DAQ system of the 10 cm $\times$ 10 cm Triple-GEM prototype. The VFAT2 Hybrid is connected through an FMC module to a first GLIB, in turn connected to a second GLIB through optical fibers. A NIM crate is used to perform the coincidence between the photomultipliers and send a trigger to the VFAT2 Hybrid through the FMC [67]. 167
10.12 Schematic view of the third prototype of the DAQ system of the 10 cm $\times$ 10 cm Triple-GEM prototype. The VFAT2 Hybrid is connected through the GEB to the OptoHybrid, in turn connected to the GLIB through optical fibers. A NIM crate is used to perform the coincidence between the photomultipliers and send a trigger to the VFAT2 Hybrid through the OptoHybrid [67]. 168

11.1 Photograph of the Xilinx SP601 development board showing the various components: the Spartan-6 FPGA in the middle, the SDRAM on its right, the Ethernet connector and MAC controller in the bottom-left corner, and the FMC connector on the left. 172

11.2 Exchange of information between the client and the server. 173
11.3 Transcript of an HTTP request and response in which the client requests a specific page and the server returns the content of that page. 174
11.4 Transcript of the request and response that are sent during the hand-shaking procedure of the WebSockets protocol between the client and the server. 175
11.5 Boot-up sequence of the operating system of the MicroBlaze highlighting the configuration of the various subsystems. 177
11.6 Diagram of the architecture used in most DAQ systems in which the whole system relies on the presence of the server which acts as the central node. 179
11.7 Diagram of the new DAQ system which forms a mesh network where every component can directly access all other nodes. 179
11.8 Nodes of the DAQ systems and the interactions they have in order to transmit the request from the user control to the front-end electronics. Blue arrows indicate that data is pushed to the next node, and red arrows that the data is pulled from the node. 181
11.9 Nodes of the DAQ systems and the interactions they have in order to transmit the data from the front-end electronics to the user control. Blue arrows indicate that data is pushed to the next node, and red arrows that the data is pulled from the node. 183
11.10 Hypothetical architecture of the control and monitoring system of the DAQ system of the GE1/1 system using the techniques described in this chapter. 184

List of Tables

5.1 Summary of the evolution of the components of the DAQ system of GE1/1. 73

6.1 Format of the tracking data packets sent by the VFAT2s. 80
6.2 Format of the data packets used to communicate between the GLIB and OptoHybrid on the fixed latency link. 87
6.3 Format of the data packets used to communicate between the GLIB and OptoHybrid on the variable latency link. 88
6.4 Values of the voltages applied to the detectors according to the current flowing through the high voltage resistive divider. 98
6.5 Data format of the packets sent between the off-detector and on-detector electronics over the GBT link. 113

8.1 Types of SEUs encountered along with their respective interaction cross section and the resulting daily rate of errors in CMS per FPGA assuming that the LHC runs at nominal values during 24h. 150

11.1 Response times of three DAQ systems for 1, 100, 500, and 1 000 operations. 182

List of Acronyms

μTCA ..... Micro Telecommunications Computing Architecture

ADC ..... Analog-to-Digital Converter

AJAX ..... Asynchronous JavaScript and XML

ALU ..... Arithmetic Logic Unit

AMC ..... Advanced Mezzanine Card

ASIC ..... Application Specific Integrated Circuit

ATCA ..... Advanced Telecommunications Computing Architecture

BC ..... BX Counter

BC0 ..... Bunch Counter 0

Booster ..... Proton Synchrotron Booster

BRAM ..... Block RAM

BU ..... Builder Unit

BX ..... Bunch Crossing

CalPulse ..... Calibration Pulse

CBM ..... Calibration, Bias & Monitoring

CDR ..... Clock Data Recovery

CERN ..... European Organisation for Nuclear Research

CFD ..... Constant Fraction Discriminator

CLB ..... Configurable Logic Block

CMS ..... Compact Muon Solenoid

CMSSW ..... CMS Software

CRC ..... Cyclic Redundancy Check

CSC ..... Cathode Strip Chamber

CU ..... Cooling Unit

DAC ..... Digital-to-Analog Converter

DAQ ..... Data Acquisition

DQM ..... Data Quality Monitoring

DSP ..... Digital Signal Processing

DT ..... Drift Tube

EC ..... Event Counter

EC0 ..... Event Counter 0

ECAL ..... Electromagnetic Calorimeter

ECC_FRAME ..... Error Checking & Correction Frame

E-Link ..... Electric serial Link

EMI ..... ElectroMagnetic Interference

EMMC ..... Enhanced Module Management Controller

ENC ..... Equivalent Noise Charge

EWK ..... Electroweak (theory)

FEC ..... Forward Error Correction

FED ..... Front-End Driver

FEROL ..... FrontEnd Readout Optical Link

FMC ..... FPGA Mezzanine Card

FPGA ..... Field Programmable Gate Array

FPU ..... Floating Point Unit

FRL ..... FrontEnd Readout Link

FSM ..... Finite State Machine

FU ..... Filter Unit

GBT ..... GigaBit Transceiver

GEB ..... GEM Electronics Board

GEM ..... Gas Electron Multiplier

GLIB ..... Gigabit Link Interface Board

GMT ..... Global Muon Trigger

GT ..... Global Trigger

GTX ..... Gigabit Transceiver X

HB ..... Barrel Hadron Calorimeter

HCAL ..... Hadron Calorimeter

HDMI ..... High-Definition Multimedia Interface

HE ..... Endcaps Hadron Calorimeter

HF ..... Forward Hadron Calorimeter

HLT ..... High Level Trigger

HO ..... Outer Hadron Calorimeter

HTTP ..... Hypertext Transfer Protocol

I2C ..... Inter-Integrated Circuit

IC ..... Integrated Circuit

ICAP ..... Internal Configuration Access Port

ILA ..... Integrated Logic Analyzer

IP ..... Internet Protocol

IPMI ..... Intelligent Platform Management Interface

JS ..... JavaScript

JTAG ..... Joint Test Action Group

L1 Trigger ..... Level-1 Trigger

L1A ..... Level-1 Accept

LED ..... Light Emitting Diode

LEP ..... Large Electron Positron (collider)

LHC ..... Large Hadron Collider

LS ..... Long Shutdown

LSB ..... Least Significant Bit

LUT ..... Look-Up Table

lwIP ..... Lightweight TCP/IP stack

MAC ..... Media Access Control

MCH ..... MicroTCA Carrier Hub

MCMC ..... MicroTCA Carrier Management Controller

MCU ..... MicroController Unit

MFS ..... Memory File System

MICROMEGAS ..... Micro-mesh gaseous structure

microTCA ..... Micro Telecommunications Computing Architecture

MMC ..... Module Management Controller

MPGD ..... Micro-Pattern Gas Detectors

MSB ..... Most Significant Bit

MSGC ..... Micro-Strip Gas Counters

MTCA ..... Micro Telecommunications Computing Architecture

MWPC ..... Multiwire Proportional Chamber

NIM ..... Nuclear Instrumentation Module

OSI ..... Open Systems Interconnection

PCB ..... Printed Circuit Board

PCIx ..... Peripheral Component Interconnect Express

PIXELS ..... Silicon pixels

PM ..... Power Module

PM ..... Photomultiplier

PRBS ..... PseudoRandom Binary Sequence

PS ..... Proton Synchrotron

QCD ..... Quantum chromodynamics

RAM ..... Random Access Memory

Resync ..... Resynchronisation

RPC ..... Resistive Plate Chamber

RU ..... Readout Unit

SEM ..... Soft Error Mitigation

SET ..... Single Event Transient

SEU ..... Single Event Upset

SL ..... Super Layer (Drift Tubes)

SPI ..... Serial Peripheral Interface bus

SPS ..... Super Proton Synchrotron

SRAM ..... Static Random-Access Memory

TCDS ..... Trigger Control and Distribution System

TCP ..... Transmission Control Protocol

TEC ..... Tracker EndCap (silicon strips)

TIB ..... Tracker Inner Barrel (silicon strips)

TID ..... Tracker Inner Disk (silicon strips)

TID ..... Total Ionizing Dose

TOB ..... Tracker Outer Barrel (silicon strips)

TTC ..... Timing, Trigger and Control

UART ..... Universal Asynchronous Receiver/Transmitter

UDP ..... User Data Protocol

USB ..... Universal Serial Bus

VIO ..... Virtual Input/Output

YETS ..... Year End Technical Stop

Bibliography

[1] F. Englert et al. „Broken Symmetry and the Mass of Gauge Vector Mesons“. In: *Phys. Rev. Lett.* 13 (9 1964), pp. 321–323. URL: http://link.aps.org/doi/10.1103/PhysRevLett.13.321 (cit. on p. 5).

[2] Peter W. Higgs. „Broken Symmetries and the Masses of Gauge Bosons“. In: *Phys. Rev. Lett.* 13 (16 1964), pp. 508–509. URL: http://link.aps.org/doi/10.1103/PhysRevLett.13.508 (cit. on p. 5).

[3] G. Aad et al. „Combined Measurement of the Higgs Boson Mass in pp Collisions at $\sqrt{s} = 7$ and 8 TeV with the ATLAS and CMS Experiments“. In: *Phys. Rev. Lett.* 114 (19 2015), p. 191803. URL: http://link.aps.org/doi/10.1103/PhysRevLett.114.191803 (cit. on p. 5).

[4] *CERN | Accelerating science*. 2012. URL: http://home.cern/ (cit. on pp. 7, 20).

[5] Roel Aaij et al. „Observation of $J/\psi p$ Resonances Consistent with Pentaquark States in $\Lambda_b^0 \to J/\psi K^- p$ Decays“. In: *Phys. Rev. Lett.* 115 (2015), p. 072001. arXiv: 1507.03414 [hep-ex] (cit. on p. 13).

[6] B. P. Abbott et al. „Observation of Gravitational Waves from a Binary Black Hole Merger“. In: *Phys. Rev. Lett.* 116 (6 2016), p. 061102. URL: http://link.aps.org/doi/10.1103/PhysRevLett.116.061102 (cit. on p. 13).

[7] Y. Fukuda et al. „Evidence for oscillation of atmospheric neutrinos“. In: *Phys. Rev. Lett.* 81 (1998), pp. 1562–1567. arXiv: hep-ex/9807003 [hep-ex] (cit. on p. 14).

[8] Lyndon Evans et al. „LHC Machine“. In: *Journal of Instrumentation* 3 (2008), S08001 (cit. on pp. 15, 18).

[9] The ALICE Collaboration. „The ALICE experiment at the CERN LHC“. In: *Journal of Instrumentation* 3.08 (2008), S08002. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08002 (cit. on p. 15).

[10] The ATLAS Collaboration. „The ATLAS Experiment at the CERN Large Hadron Collider“. In: *Journal of Instrumentation* 3.08 (2008), S08003. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08003 (cit. on p. 15).

[11] The CMS Collaboration. „The CMS experiment at the CERN LHC“. In: *Journal of Instrumentation* 3.08 (2008), S08004. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08004 (cit. on pp. 15, 23–25, 28, 29, 31–34).

[12] The LHCb Collaboration. „The LHCb Detector at the LHC“. In: *Journal of Instrumentation* 3.08 (2008), S08005. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08005 (cit. on p. 15).

[13] The TOTEM Collaboration. „The TOTEM Experiment at the CERN Large Hadron Collider“. In: *Journal of Instrumentation* 3.08 (2008), S08007. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08007 (cit. on p. 15).

[14] The LHCf Collaboration. „The LHCf detector at the CERN Large Hadron Collider“. In: *Journal of Instrumentation* 3.08 (2008), S08006. URL: http://stacks.iop.org/1748-0221/3/i=08/a=S08006 (cit. on p. 15).

[15] B. Acharya et al. „The Physics Programme Of The MoEDAL Experiment At The LHC“. In: *Int. J. Mod. Phys.* A29 (2014), p. 1430050. arXiv: 1405.7662 [hep-ph] (cit. on p. 15).

[16] *TE-EPC-LPC in LHC*. 2012. URL: http://te-epc-lpc.web.cern.ch/te-epc-lpc/machines/lhc/general.stm (cit. on p. 16).

[17] *Luminosity plots pp*. 2015. URL: http://lpc.web.cern.ch/lpc/lumiplots_2015_pp.htm (cit. on p. 19).

[18] CMS Collaboration et al. „Observation of the rare $B_s^0 \to \mu^+\mu^-$ decay from the combined analysis of CMS and LHCb data“. In: *Nature* 522.7554 (June 2015), pp. 68–72. URL: http://dx.doi.org/10.1038/nature14474 (cit. on p. 20).

[19] Serguei Chatrchyan et al. „Description and performance of track and primary-vertex reconstruction with the CMS tracker“. In: *JINST* 9.10 (2014), P10009. arXiv: 1405.6569 [physics.ins-det] (cit. on pp. 26, 27).

[20] G Baiatian et al. *Energy Response and Longitudinal Shower Profiles Measured in CMS HCAL and Comparison With Geant4*. Tech. rep. CMS-NOTE-2006-143. Geneva: CERN, 2007. URL: http://cds.cern.ch/record/1049929 (cit. on p. 30).

[21] S Chatrchyan et al. „Precise Mapping of the Magnetic Field in the CMS Barrel Yoke using Cosmic Rays“. In: *Journal of Instrumentation* 5 (2010), T03021. arXiv: 0910.5530 [physics.ins-det] (cit. on p. 30).

[22] A Tapper et al. *CMS Technical Design Report for the Level-1 Trigger Upgrade*. Tech. rep. CERN-LHCC-2013-011. CMS-TDR-12. Additional contacts: Jeffrey Spalding, Fermilab, Jeffrey.Spalding@cern.ch Didier Contardo, Universite Claude Bernard-Lyon I, didier.claude.contardo@cern.ch. 2013. URL: http://cds.cern.ch/record/1556311 (cit. on p. 35).

[23] Tomasz Bawej et al. „The New CMS DAQ System for Run-2 of the LHC“. In: *IEEE Trans. Nucl. Sci.* 62.3 (2015). [,7097437(2014)], pp. 1099–1103 (cit. on p. 36).

[24] The CMS Collaboration. *CMS Technical Design Report for the Muon Endcap GEM Upgrade*. Tech. rep. CERN-LHCC-2015-012. CMS-TDR-013. Geneva: CERN, 2015. URL: https://cds.cern.ch/record/2021453 (cit. on pp. 39, 41–47, 49–53, 55, 56, 58).

[25] L Alunni et al. „Performance of MSGC manufactured on electronic and ionic conductivity substrata in various operational conditions“. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 348.CERN-PPE-93-179 (1993), 344–350. 14 p. URL: http://cds.cern.ch/record/254631 (cit. on p. 41).

[26] Ioanis Giomataris et al. „Micromegas: a high-granularity position-sensitive gaseous detector for high particle-flux environments“. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 376.DAPNIA-SED-95-04. 1 (1995), 29–35. 20 p. URL: http://cds.cern.ch/record/299159 (cit. on p. 41).

[27] F. Sauli. „GEM: A new concept for electron amplification in gas detectors“. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 386.2 (1997), pp. 531 –534. URL: http://www.sciencedirect.com/science/article/pii/S0168900296011722 (cit. on p. 41).

[28] Thierry Maerschalk. „Study of Triple-GEM detectors for the CMSmuon spectrometer upgrade at LHC“. PhD thesis. Brussels U., 2016. URL: http://iihe.ac.be/publications.php (cit. on pp. 51, 52, 54, 56).

[29] *MicroTCA® Overview*. 2015. URL: https://www.picmg.org/openstandards/microtca/ (cit. on p. 56).

[30] P Moreira et al. „The GBT Project“. In: (2009). URL: http://cds.cern.ch/record/1235836 (cit. on pp. 59–61).

[31] C Soos et al. „The Versatile transceiver: Towards production readiness“. In: *Journal of Instrumentation* 8 (2013), p. C03004. URL: http://cds.cern.ch/record/1609037 (cit. on p. 59).

[32] *MicroTCA Overview, A Brief Introduction to Micro Telecommunications Computing Architecture Concepts*. 2016. URL: http://www.vadatech.com/media/article_MicroTCA_Overview.pdf (cit. on p. 62).

[33] A. Svetek et al. „The Calorimeter Trigger Processor Card: the next generation of high speed algorithmic data processing at CMS“. In: *Journal of Instrumentation* 11.02 (2016), p. C02011. URL: http://stacks.iop.org/1748-0221/11/i=02/a=C02011 (cit. on p. 63).

[34] *AMC13 Information Page*. 2016. URL: http://bucms.bu.edu/twiki/bin/view/BUCMSPublic/HcalDTC (cit. on pp. 64, 66).

[35] P Aspell et al. „VFAT2: A front-end system on chip providing fast trigger information, digitized data storage and formatting for the charge sensitive readout of multi-channel silicon and gas particle detectors“. In: (2007), 5 p. URL: https://cds.cern.ch/record/1069906 (cit. on pp. 67, 122).

[36] *MSP430 Wireless Development Tool*. 2016. URL: http://www.ti.com/tool/ez430-rf2500 (cit. on p. 69).

[37] P Vichoudis et al. „The Gigabit Link Interface Board (GLIB), a flexible system for the evaluation and use of GBT-based optical links“. In: *Journal of Instrumentation* 5 (2010), p. C11007. URL: http://cds.cern.ch/record/1359270 (cit. on pp. 69, 70).

[38] *SoC Interconnection: Wishbone*. 2016. URL: http://opencores.org/opencores,wishbone (cit. on p. 75).

[39] *I2C Info – I2C Bus, Interface and Protocol*. 2016. URL: http://i2c.info/ (cit. on p. 82).

[40] Xilinx. *Virtex-6 FPGA System Monitor*. UG370 v1.2. 2014. URL: http://www.xilinx.com/ (cit. on p. 85).

[41] Abhijit Athavale et al. *High-Speed Serial I/O Made Simple*. 1st ed. Connectivity Solutions, Apr. 2005 (cit. on p. 86).

[42] *NodeJS - Event-driven I/O server-side JavaScript environment*. 2016. URL: https://nodejs.org/en/ (cit. on p. 89).

[43] *Socket.IO - enabling real-time bidirectional event-based communication*. 2016. URL: http://socket.io/ (cit. on p. 89).

[44] *AngularJS - Superheroic JavaScript MVW Framework*. 2016. URL: https://angularjs.org/ (cit. on p. 89).

[45] *Bootstrap - Designed for everyone, everywhere*. 2016. URL: http://getbootstrap.com/ (cit. on p. 89).

[46] Jeremie Alexandre Merlin. „Study of long-term sustained operation of gaseous detectors for the high rate environment in CMS“. Presented 25 Apr 2016. PhD thesis. Strasbourg U., 2016. URL: https://cds.cern.ch/record/2155685 (cit. on p. 100).

[47] D Abbaneo et al. „GEM based detector for future upgrade of the CMS forward muon system“. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 718 (2013), 383–386. 4 p. URL: http://cds.cern.ch/record/1709907 (cit. on p. 106).

[48] D Abbaneo et al. *Test beam results of the GE1/1 prototype for a future upgrade of the CMS high-η muon system*. Tech. rep. arXiv:1111.4883. RDSI-Note-2011-013. 2011. URL: http://cds.cern.ch/record/1401079 (cit. on pp. 106, 110).

[49] D Abbaneo et al. *Beam Test Results for New Full-scale GEM Prototypes for a Future Upgrade of the CMS High-eta Muon System*. Tech. rep. arXiv:1211.3939. Comments: 5 pages, 9 figures, submitted to Proc. 2012 IEEE Nucl. Sci. Symposium, Anaheim, CA. 2012. URL: http://cds.cern.ch/record/1494965 (cit. on pp. 106, 110).

[50] M. Alfonsi et al. „Simulation of the dielectric charging-up effect in a {GEM} detector“. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 671 (2012), pp. 6 –9. URL: http://www.sciencedirect.com/science/article/pii/S0168900211022789 (cit. on pp. 108, 109).

[51] V Peskov et al. „Development of novel designs of spark-protected micropattern gaseous detectors with resistive electrodes“. In: *Journal of Instrumentation* 7.01 (2012), p. C01005. URL: http://stacks.iop.org/1748-0221/7/i=01/a=C01005 (cit. on p. 109).

[52] D Abbaneo et al. „Performance of a Large-Area GEM Detector Prototype for the Upgrade of the CMS Muon Endcap System“. In: *PoS* EPS-HEP2013.arXiv:1412.0228 (2014). Comments: 8 pages, 32 figures, submitted to Proc. 2014 IEEE Nucl. Sci. Symposium, Seattle, WA, 123. 8 p. URL: http://cds.cern.ch/record/1973272 (cit. on p. 111).

[53] P Aspell et al. „VFAT2: A front-end "system on chip" providing fast trigger information and digitized data storage for the charge sensitive readout of multi-channel silicon and gas particle detectors“. In: (2009). URL: http://cds.cern.ch/record/1267947 (cit. on pp. 119, 121).

[54] Xilinx. *Virtex-6 FPGA Configurable Logic Block*. UG364 v1.2. 2012. URL: http://www.xilinx.com/ (cit. on pp. 129, 130).

[55] Xilinx. *Virtex-6 FPGA DSP48E1 Slice*. UG369 v1.3. 2011. URL: http://www.xilinx.com/ (cit. on p. 131).

[56] Xilinx. *Virtex-6 FPGA Memory Resources*. UG363 v1.8. 2014. URL: http://www.xilinx.com/ (cit. on p. 132).

[57] Xilinx. *Considerations Surrounding Single Event Effects in FPGAs, ASICS, and Processors*. WP402 v1.0.1. 2012. URL: http://www.xilinx.com/ (cit. on p. 134).

[58] *CENTRE DE RESSOURCES DU CYCLOTRON*. 2016. URL: http://www.cyc.ucl.ac.be/ (cit. on p. 142).

[59] B. Bylsma et al. „Radiation testing of electronics for the CMS endcap muon system“. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 698 (2013), pp. 242 –248. URL: http://www.sciencedirect.com/science/article/pii/S0168900212010455 (cit. on pp. 144–146).

[60] M Huhtinen et al. „Computational method to estimate Single Event Upset rates in an accelerator environment“. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 450.1 (2000), pp. 155 –172. URL: http://www.sciencedirect.com/science/article/pii/S0168900200001558 (cit. on pp. 144, 149).

[61] Xilinx. *Soft Error Mitigation Controller v4.1*. PG036. 2015. URL: http://www.xilinx.com/ (cit. on p. 145).

[62] Xilinx. *Device Reliability Report*. UG116 v10.4. 2016. URL: http://www.xilinx.com/ (cit. on pp. 145, 151).

[63] Florian Zenoni. „Study of Triple-GEM detectors for the CMSmuon spectrometer upgrade at LHC and study of the forward-backward charge asymmetry for the search of extra neutral gauge bosons“. PhD thesis. Brussels U., 2016. URL: http://iihe.ac.be/publications.php (cit. on p. 149).

[64] K. A. Olive et al. „Review of Particle Physics“. In: *Chin. Phys.* C38 (2014), p. 090001 (cit. on p. 149).

[65] Samaras A et al. „Experimental Characterization and Simulation of Electron-induced SEU in 45nm CMOS Technology“. In: CNES / ESA Days, 2015 (cit. on p. 149).

[66] VMEbus International Trade Association. *American National Standard for VME64*. ANSI/VITA 1-1994. 1994. URL: http://www.vita.com/ (cit. on p. 158).

[67] Alexandre Léonard. „Measurement of Z boson production in association with jets at the LHC and study of a DAQ system for the Triple-GEM detector in view of the CMS upgrade“. Presented 10 Jun 2015. PhD thesis. Brussels U., 2015. URL: http://cds.cern.ch/record/2065693 (cit. on pp. 163, 167, 168).

[68] Xilinx. *Microblaze Processor Reference Guide*. UG081 v10.3. 2009. URL: http://www.xilinx.com/ (cit. on p. 164).