ULB Faculté des Sciences

Development of the data acquisition system, commissioning, and operations of the CMS experiment GE1/1 Triple-GEM detector

Thesis submitted by Laurent PÉTRÉ
in fulfilment of the requirements of the PhD Degree in Physics (“Doctorat en Physique”)
Academic year 2024-2025

Supervisor: Prof. Gilles DE LENTDECKER

Co-supervisor: Prof. Frédéric ROBERT

Thesis jury:

Prof. Laurent FAVART (Université libre de Bruxelles, Chair)
Prof. Ioana MARIS (Université libre de Bruxelles, Secretary)
Prof. Gilles DE LENTDECKER (Université libre de Bruxelles, Supervisor)
Prof. Frédéric ROBERT (Université libre de Bruxelles, Co-supervisor)
Dr. Laurent MIRABITO (CNRS - Lyon)
Dr. Pieter EVERAERTS (University of Wisconsin)
Dr. Isabelle DE BRUYN (Université libre de Bruxelles)

fnrs LA LIBERTÉ DE CHERCHER

Abstract

In the coming years, the CERN LHC (Large Hadron Collider) will undergo a series of upgrades to increase its instantaneous luminosity up to $10 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, targeting a total integrated luminosity of $4000\,\mathrm{fb}^{-1}$ after ten years of operation. The CMS (*Compact Muon Solenoid*) experiment, like the other LHC detectors, will face increasingly harsh conditions due to the higher particle flux and the larger radiation doses, particularly in its most forward regions. To meet these challenges, CMS is undergoing several detector upgrades, including the installation in the endcaps of three new muon stations based on the Triple-GEM technology: ME0, GE1/1, and GE2/1.

This thesis focuses on the GE1/1 station, installed during the Long Shutdown 2 (LS2) between July 2019 and September 2020, and operated since the start of Run-3. It first describes the evolution of the data acquisition (DAQ) system from a small-scale project - used for quality control and demonstrator setups - into a fully featured solution suitable for large-scale operations. This evolution required an almost complete redesign and rewrite of the software stack, resulting in a system with robust control and monitoring capabilities, that have been validated during the commissioning and operations in LS2 and Run-3. Special emphasis was placed on scalability and compatibility with the future ME0 and GE2/1 stations.

This thesis then summarizes key aspects of the electronics quality control, including the optimization strategies implemented which contributed to the successful delivery and completion of the project.

Further, this thesis discusses the commissioning challenges. Notably, how major communication instabilities with front-end readout electronics were mitigated, and how the electronics noise levels were reduced to below 1 fC. Finally, the most recent operational performance results are presented. Through careful monitoring, the readout channel damages due to high-voltage discharges were limited to less than 0.23% after 4 years of operation. The detector timing alignment, together with additional configuration refinements, led to an average measured efficiency of 94%, approaching the 97% design target.

Résumé

Dans les années à venir, le LHC (*Large Hadron Collider*) du CERN fera l'objet de mises à niveau afin d'augmenter sa luminosité instantanée jusqu'à $10\times10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, avec pour objectif d'atteindre une luminosité intégrée totale de $4000\,\mathrm{fb}^{-1}$ après dix ans d'exploitation. L'expérience CMS (*Compact Muon Solenoid*), comme les autres détecteurs du LHC, sera confrontée à des conditions de plus en plus difficiles, en raison d'un flux de particules plus important et de doses de radiation accrues, en particulier dans ses régions avants. À ces fins, CMS modernise ses détecteurs, incluant l'installation dans les bouchons de trois nouvelles stations à muons basées sur la technologie Triple-GEM : ME0, GE1/1 et GE2/1.

Cette thèse se concentre sur la station GE1/1, installée durant le *Long Shutdown 2* (LS2) entre juillet 2019 et septembre 2020, et en fonctionnement depuis le début du Run-3. Elle décrit l'évolution du système d'acquisition de données (DAQ), passé d'un projet à petite échelle - utilisé pour le contrôle de qualité et les prototypes - à une solution complète adaptée aux opérations à grande échelle. Cette évolution a nécessité une refonte presque intégrale de la pile logicielle, aboutissant à un système de contrôle et de surveillance robuste, validé lors de la mise en service et des opérations durant le LS2 et le Run-3. Une attention particulière a été accordée à son évolutivité et à sa compatibilité avec futures stations ME0 et GE2/1.

Cette thèse résume les principaux résultats du contrôle de qualité de l'électronique, incluant les stratégies d'optimisation qui ont permis de mener à bien le projet dans les délais.

Cette thèse aborde ensuite les défis rencontrés lors de la mise en service, en particulier la manière dont des instabilités majeures de communication avec l'électronique de lecture ont été traitées, et comment les niveaux de bruit de l'électronique ont été réduits à moins de 1 fC. Enfin, les résultats opérationnels les plus récents sont présentés. Grâce à une surveillance rigoureuse, les dommages causés aux canaux de lecture par les décharges haute tension ont été limités à moins de 0,23% après quatre ans de fonctionnement. L'alignement en temps du détecteur, combiné à des optimisations de configuration, a permis d'obtenir une efficacité moyenne de détection de 94%, proche de l'objectif de conception de 97%.

Acknowledgments

Contents

Abstract iii

Acknowledgments v

Contents vii

Introduction 1

1 The CMS experiment at the LHC 3
1.1 The Large Hadron Collider . . . 3
1.1.1 Main characteristics . . . 5
1.1.2 The LHC schedule . . . 6
1.2 The CMS experiment . . . 7
1.2.1 The tracker . . . 8
1.2.2 The calorimeters . . . 10
1.2.3 The solenoid magnet . . . 11
1.2.4 The muon spectrometer . . . 12
1.2.5 The trigger and data acquisition systems . . . 17
1.2.6 The Phase-II upgrades . . . 21

2 The CMS GEM project 25
2.1 Gaseous detectors working principles . . . 25
2.1.1 Energy losses in medium . . . 25
2.1.2 Gaseous detectors . . . 26
2.1.3 GEM detectors . . . 30
2.1.4 Factors affecting the time resolution . . . 33
2.2 The GE1/1 project . . . 35
2.2.1 Requirements on GE1/1 . . . 36
2.2.2 The GE1/1 detector . . . 37
2.2.3 The front-end readout electronics . . . 39
2.2.4 The back-end readout electronics . . . 46
2.2.5 The detector services . . . 50
2.3 The GEM Phase-II upgrades, ME0 & GE2/1 . . . 52
2.3.1 GE2/1 highlights . . . 53

2.3.2 ME0 highlights . . . 55
2.3.3 DAQ highlights . . . 57
2.4 Summary . . . 59

3 The CMS GEM data acquisition system 61
3.1 Common CMS software frameworks . . . 62
3.1.1 xDAQ . . . 62
3.1.2 RCMS . . . 62
3.2 The legacy software . . . 63
3.3 Software architecture . . . 65
3.3.1 The back-end services . . . 67
3.3.2 Control applications . . . 68
3.3.3 System configuration mechanisms . . . 72
3.3.4 The Function Manager . . . 74
3.4 Local readout . . . 74
3.5 Calibration suite . . . 76
3.5.1 GBT phase scan . . . 77
3.5.2 DAC scan . . . 79
3.5.3 S-bit rate scan . . . 80
3.5.4 S-curves . . . 82
3.5.5 Threshold scan . . . 87
3.5.6 Latency scan . . . 87
3.5.7 Scans scalability . . . 89
3.6 Monitoring suite . . . 90
3.6.1 Online monitoring system . . . 90
3.6.2 DCS interface . . . 90
3.6.3 Archiving . . . 93
3.6.4 Contextual information . . . 94
3.7 Automatic actions . . . 95
3.7.1 Automatic masking . . . 95
3.7.2 Automatic recovery . . . 96
3.8 Summary . . . 98

4 Quality control 101
4.1 Overview . . . 101
4.2 GEB quality control . . . 104
4.3 OptoHybrid quality control . . . 107
4.3.1 Temperature sensors . . . 107
4.4 Combined GEB & OptoHybrid quality control . . . 112
4.5 Database storage . . . 114
4.6 Outcome & lessons learned . . . 116

5 Commissioning & Operations 119
5.1 Commissioning . . . 119

135 5.1.1 Trolley test . . . 122
136 5.1.2 Mapping validation . . . 122
137 5.2 Communication instabilities . . . 124
138 5.2.1 GBTx fuses corruption . . . 124
139 5.2.2 VTRx outgassing . . . 125
140 5.2.3 Powering weakness . . . 126
141 5.2.4 Saturation of the VFAT bandwidth . . . 128
142 5.2.5 High-voltage discharges . . . 129
143 5.2.6 Random communication failures . . . 129
144 5.3 VFAT damages due to HV discharges . . . 129
145 5.3.1 Channel damages . . . 130
146 5.3.2 Channel damages on the GE2/1 demonstrator . . . 131
147 5.3.3 DAC circuits . . . 132
148 5.4 Electronics noise . . . 133
149 5.5 Latency scans . . . 138
150 5.6 High multiplicity events . . . 140
151 5.7 Efficiency measurements . . . 147
152 5.7.1 Dataset . . . 147
153 5.7.2 Workflow . . . 147
154 5.7.3 Analysis methodology . . . 148
155 5.7.4 Results . . . 150
156 5.7.5 Prospects . . . 154
157 5.8 Summary . . . 156

158 **Conclusions** **159**

159 **List of Acronyms** **161**

160 **List of Figures** **169**

161 **List of Tables** **177**

162 **Bibliography** **179**

Introduction

The Large Hadron Collider (LHC), built at CERN between 2000 and 2008 and operated since then, is the world's largest and most powerful particle accelerator. Initially designed to deliver proton-proton collisions at a rate of 40 MHz, with a center-of-mass energy of 14 TeV, and an instantaneous luminosity of $1 \times 10^{34}\,\mathrm{cm^{-2}s^{-1}}$, the LHC has undergone - and will continue to undergo - major upgrades. Between 2019 and 2020, during the Long Shutdown 2 (LS2), the machine's instantaneous luminosity was increased to $2 \times 10^{34}\,\mathrm{cm^{-2}s^{-1}}$, twice its design value. Between 2026 and 2028, the Long Shutdown 3 (LS3) will prepare the LHC for a new operation phase, called High-Luminosity LHC (HL-LHC), targeting an instantaneous luminosity of $10 \times 10^{34}\,\mathrm{cm^{-2}s^{-1}}$.

While these upgrades will significantly enhance the LHC's discovery and precision measurement potential, they also represent major challenges for the detectors due to the higher particle fluxes and the radiation doses. The Compact Muon Solenoid (CMS) experiment, like the other LHC detectors, was not originally designed for such intense operating conditions. Therefore, it must itself undergo a series of upgrades, particularly in the most forward regions of the muon spectrometer at pseudo-rapidities $|\eta|$ above 1.6.

One of the main challenges in this region consists of maintaining the efficiency and the low background rate of the Level-1 Trigger (L1T). Failing to do so would require increasing the threshold on the transverse momentum of single muons, compromising the experiment's physics reach. The GEM Collaboration proposed the GE1/1 project, designed to complement the existing Cathode Strip Chambers (CSC) and maintain the performance of the muon subsystem. The GE1/1 station consists of 144 trapezoidal gaseous detector chambers based on the Triple-GEM technology, installed in the space left vacant by the innermost layers of Resistive Plate Chamber (RPC). Extensive R&D proved that the Triple-GEM micro-pattern gaseous detectors are suited for the HL-LHC environment in terms of time resolution ($\sim$10 ns), spatial resolution ($\sim$300 µm), detection efficiency (>97 %), and rate capability (up to O(MHz/cm^2)). The GE1/1 station will later be completed by the GE2/1 station - placed adjacent to CSC chambers as well -, and by ME0, which extends the coverage of the muon sub-system to $\eta = 2.8$.

Before their installation in CMS during LS2, all 144 GE1/1 chambers underwent a rigorous sequence of quality control steps, scrutinizing all operational parameters and validating the performance. Following their installation, a thorough commissioning phase ensured their proper functioning in the final environment. Meanwhile, the data acquisition (DAQ) system was developed and continuously refined, incorporating feedback

from the early operations to optimize both the DAQ and and detector performances, while mitigating emerging issues.

The work presented in this thesis took place within the context of the production and commissioning of the GE1/1 detectors in view of their operation during the LHC Run-3. Nevertheless, one of my priorities has always been to ensure that most of the work and established procedures could be extended naturally to the GEM Phase-II upgrades following LS3. Initially, I actively participated in the Quality Control (QC) of the GE1/1 on-detector electronics at ULB, gaining in-depth knowledge of the hardware and the limitations of the DAQ software in use at that time. I then pursued my work taking a leading role in the commissioning and operations of the GE1/1 detectors after their installation within CMS. The main challenges addressed were the handling of severe communication instabilities with the front-end readout electronics, the reduction of the electronics noise to acceptable levels, and the minimization of the damages due to high-voltage discharges. During this time, GE1/1 recorded its first-ever muons, marking a significant milestone for the GEM collaboration. This also resulted in the observation of the high-multiplicity event, also called "flower events", which I characterized and provided satisfactory mitigation for the level of the Level-1 Trigger. Overall, these efforts yielded the first efficiency measurements, reaching 94% after optimization. In parallel, I also became the lead developer of the GEM DAQ system, evolving it from a small-scale project designed for QC setups to a large-scale project suited for CMS operations. The synergy between those two roles contributed to the project's success, allowing for rapid issue resolution and system improvement.

The content of this thesis is structured as follows:

- Chapter 1 describes the LHC and the CMS experiment as of Run-3, along with the planned HL-LHC upgrades.
- Chapter 2 focuses on the CMS GEM project. It introduces the principles of gaseous detectors with an emphasis on the GEM technology. The GE1/1 detector and its readout electronics are then described in detail, followed by an overview of the GEM Phase-II upgrades: ME0 and GE2/1.
- Chapter 3 dives into the development of the GEM DAQ, focusing on its evolution from the so-called *legacy software* to a robust and scalable solution meeting the current and future operational needs.
- Chapter 4 presents the quality control procedures applied to the readout electronics and their results.
- Chapter 5 details the commissioning challenges and summarizes the most recent operational performances.

Original content can be found in Chapters 3 to 5.

Chapter 1

The Compact Muon Solenoid experiment at the Large Hadron Collider

1.1 The Large Hadron Collider

The LHC [36] is part of the CERN accelerator complex and is the world's largest and most powerful particle accelerator to this date. Its construction spans from 2000 to 2008 in the tunnel previously used by the Large Electron-Positron (LEP) collider and lies between 45 m and 170 m under the France–Switzerland border, near Geneva.

The LHC is built around a 26.7 km ring of two parallel pipes where beams of particles, confined by superconducting dipole magnets, circulate in opposite directions within an ultrahigh vacuum. In four locations, called Interaction Points (IP), the beams are focused and crossed by superconducting quadrupole and octupole magnets in order to collide at a nominal frequency of 40.079 MHz. While most of the LHC program consists of proton-proton collisions, the beams can also be composed of heavy ion bunches allowing, for example, lead-lead collisions or lead-proton collisions.

The particles are progressively accelerated and split into bunches through the CERN accelerator complex before being provided to the LHC (see Figure 1.1).

In the case of proton beams, the sequence is the following. A source provides negative hydrogen ions (H^-) to a 86 m linear accelerator, the LINAC4, designed to increase their energy to 160 MeV [109]. Between their extraction from the LINAC and their injection into the Proton Synchrotron Booster (PSB), the ions are stripped from their electrons, resulting in bare protons. The PSB is where the protons are accelerated up to 1.4 GeV before their injection into the Proton Synchrotron (PS). The PS then accelerates the bunches further up to an energy of 28 GeV. This is also where they acquire their final 25 ns spacing that will be maintained up to the LHC. Once extracted from the PS, the bunches are accelerated one last time before injection to the LHC via the Super Proton

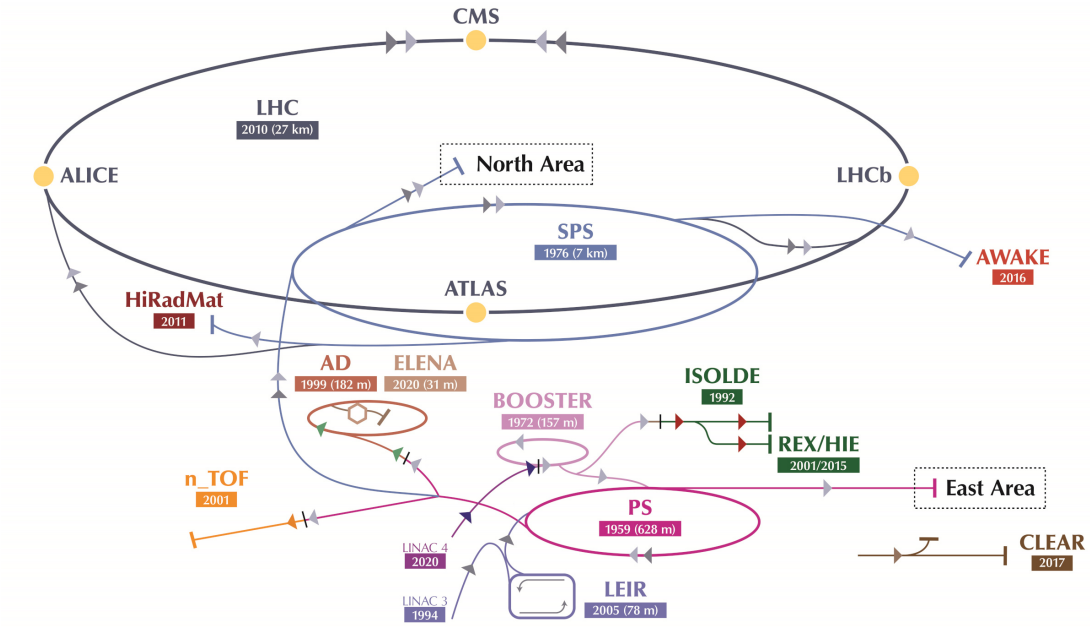


Figure 1.1: Overview of the CERN accelerator complex as of January 2022. The LHC and its four main experiments are drawn in dark blue. The LHC proton accelerating chain in composed (in order) by the LINAC4 (purple), the BOOSTER (pink), the PS (magenta) and the SPS (lighter blue) [80].

Synchrotron (SPS), which speeds up the protons up to 450 GeV. The bunches are finally transferred to the LHC where the final acceleration takes place into radiofrequency cavities. The protons have reached their nominal energy of 7 TeV. They will revolve around the LHC with a frequency of 11.2455 kHz until the beam dump.

Four experiments are located in caverns at the four interaction points. ATLAS (A Toroidal LHC ApparatuS) [17] and CMS (Compact Muon Solenoid) [47] are general-purpose detectors located, respectively, at the IP1 (Interaction Point 1) and the IP5 of the LHC. Their main scientific goals consist in the precise measurement of the standard model as well as searches for new physics. One of their major discoveries was the Brout-Englert-Higgs boson in 2012. [40, 16]. Despite similar goals and principles of operation, they feature different designs and technologies and fulfill the need for confirmation of physics results.

ALICE (A Large Ion Collider Experiment) [9], located at IP2, is tailored to the analysis of heavy ion collisions. Such collisions permit the study of the Quantum ChromoDynamics (QCD) and the improvement of the understanding of quark-gluon plasma properties. LHCb (Large Hadron Collider beauty experiment) [77], located at IP8, is specialized in the precise study of bottom quark physics. It aims at better understanding possible violations of charge and parity (CP) conservation in fundamental interactions and the matter-antimatter asymmetry in the universe.

Smaller experiments are installed in the same caverns as or in caverns adjacent to the four bigger experiments. LHCf (Large Hadron Collider forward) [78] measures the neutral particles produced in the direction of the beams. It provides calibration data for models used in the study of ultra-high-energy cosmic rays. TOTEM (TOTal Elastic and diffractive cross section Measurement) [105] is designed to measure the total cross-section of the proton with detectors located in the LHC tunnel at a distance of 200 m from the CMS IP. Since 2018 and the LHC Run-3, it has now evolved into a standard CMS sub-detector [7]. MoEDAL (Monopole and Exotics Detector At LHC) [88] searches for new hypothetical particles such as magnetic monopoles or massive stable charged particles. MilliQan [20] is designed to search for milli-charged particles produced in the proton-proton collisions at the LHC. FASER (ForwArd Search ExpeRiment) [6] is designed to search for new, yet undiscovered, light and weakly-interacting particles and study the interactions of high-energy neutrinos.

1.1.1 Main characteristics

In any accelerator, the two most important parameters for the physics are the center-of-mass energy and the luminosity:

- The center-of-mass energy ($\sqrt{s}$) is the total amount of energy available in a collision. It is function of the energy E_i and the momentum $\vec{p_i}$ of each of the two beams. As the LHC is a symmetric collider, the momentums are opposite ($\vec{p_1} = -\vec{p_2}$) and therefore the total energy is the sum of the energy of the two beams.

$$\sqrt{s} = \sqrt{(E_1 + E_2)^2 - (\vec{p_1} + \vec{p_2})^2} = E_1 + E_2$$

- The instantaneous luminosity is a function of n_i, the number of particles by bunch, σ_x and σ_y, the root mean squared transverse beam sizes, and f, the frequency of collision [27]:

$$\mathcal{L}_{inst} = f\frac{n_1 n_2}{4\pi\sigma_x\sigma_y}$$

  By integrating this instantaneous luminosity over a time period, one gets the total luminosity $\mathcal{L}$:

$$\mathcal{L} = \int_{t1}^{t2} \mathcal{L}_{inst}\mathrm{d}t$$

  It is generally expressed in barn^{-1} [1]. By definition that quantity is proportional to the expected number of events of a physical process of cross-section σ_{evt}:

$$N_{evt} = \mathcal{L}\sigma_{evt}$$

[1] $1\,\text{barn} = 1 \times 10^{-24}\,\text{cm}^2$

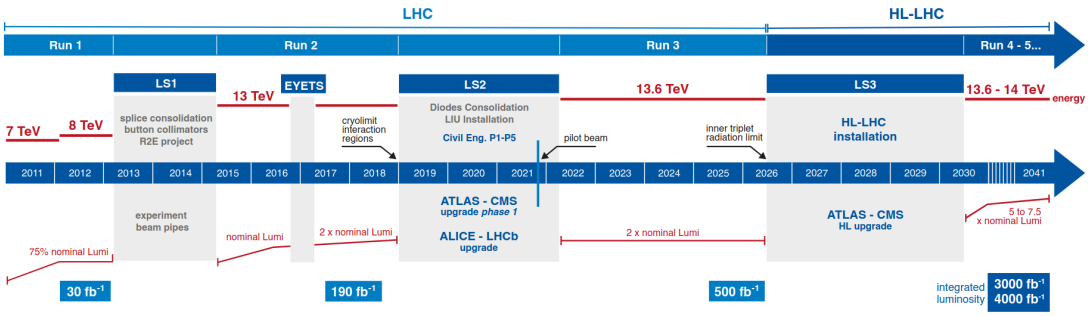


Figure 1.2: The planned LHC long-term calendar. It will lead to the commissioning of the HL-LHC, High Luminosity-LHC around 2028. [35]

The LHC nominal center-of-mass energy is 14 TeV while its nominal instantaneous luminosity is $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

1.1.2 The LHC schedule

Year after year the LHC is upgraded to improve its performance and reliability while replacing possibly aging parts. To fulfill its ambitious physics program, both the center-of-mass energy and luminosity were and will be increased to help in the discovery of new and rare physics processes.

A summary of the past and future upgrades of the LHC can be visualized in Figure 1.2. The LHC first operated with a center-of-mass energy of 7 TeV and was progressively upgraded to reach 13.6 TeV during the Run-3 period. At the same time, the instantaneous luminosity was also improved to reach $2.06 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, more than twice the design value of $1 \times 10^{33}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ [90]. At the end of Run-2, in 2018, CMS had recorded an integrated luminosity of $150.78\,\mathrm{fb}^{-1}$ at an energy of 13 TeV [102].

The next big leap is what is called the High Luminosity-LHC (HL-LHC) or LHC Phase II [13]. While its center-of-mass energy will be limited to 14 TeV, the luminosity will peak at $10 \times 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

In order to prepare for such a jump in luminosity, the Long Shutdown (LS2) took place between 2019 and 2020. CMS used this opportunity to complete the first phase of its upgrade program.

The LHC and detector upgrades will be finalized during the Long Shutdown (LS3) scheduled between 2026 and 2028. At the end of the HL-LHC operation, the total integrated luminosity should reach $3000\,\mathrm{fb}^{-1}$ to $4000\,\mathrm{fb}^{-1}$.

1.2 The CMS experiment

CMS [54, 47, 39] is one of the four big LHC experiments and one of its two general-purpose detectors. It is shaped like a cylinder with a length of 21 m and a diameter of 15 m for a total weight of 14 000 tons. The detector is designed nearly hermetic around the interaction point to trigger and identify efficiently electrons, hadrons, muons, and photons.

CMS follows a somewhat traditional detector architecture, both in its barrel section and in its endcaps. Figure 1.3 represents one octant of a slice of the CMS barrel and allows to follow the path of a particle produced near the collision point, through the multiple layers of sub-detectors [57]. The first layer consists of an inner tracking system, followed by electromagnetic and hadronic calorimeters, and finally a muon spectrometer. One of its iconic features is the compact solenoid capable of generating a strong magnetic of 3.8 T in its inner part.

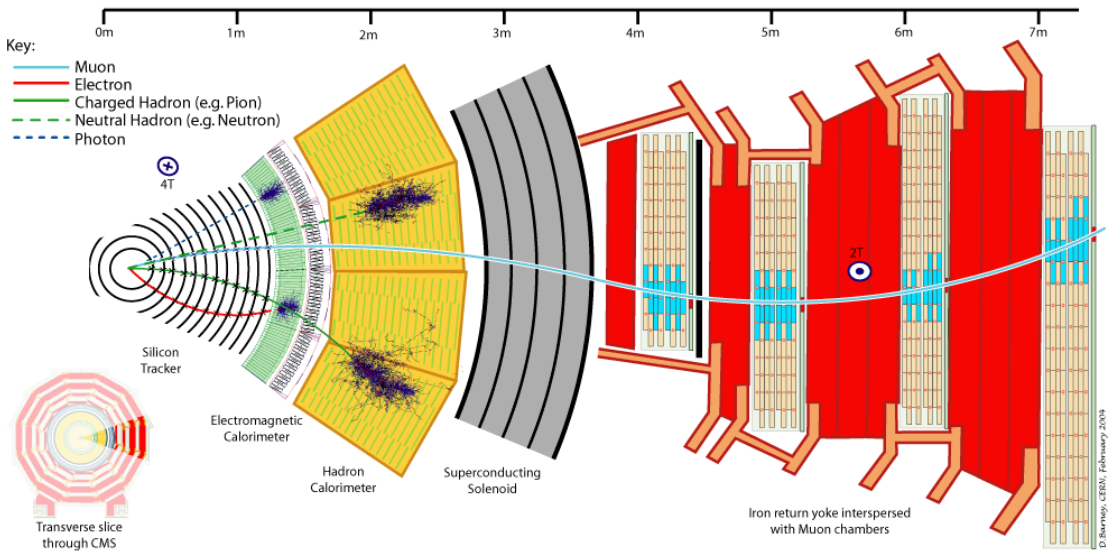


Figure 1.3: Transverse view of one octant of the barrel of CMS. The lines represent the path of different particles through the different sub-detectors : the tracker in black, the ECAL in green, the HCAL in yellow and the muon system in orange. The solenoid is colored in grey. [74]

The CMS system of coordinates is centered on the collision point with the Y-axis pointing vertically, the X-axis pointing to the center of the LHC ring, and the Z-axis pointing in the direction of the beams in a right-handed way. Due to the geometry of CMS, a spherical system of coordinates is the most appropriate: R is defined as the distance from the beamline ($R = \sqrt{x^2 + y^2}$), ϕ as the azimuthal angle in the XY plane,

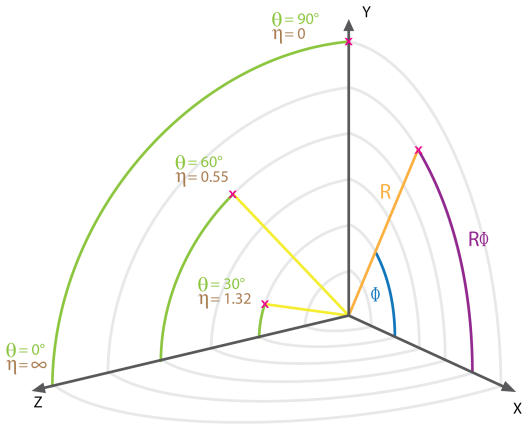


Figure 1.4: CMS system of coordinates. The beamline is along the Z axis. In green and brown, one can see the comparison between the polar angle θ and the pseudo-rapidity η. [76]

and θ as the polar angle with respect to the Z-axis. However, these coordinates do not present any Lorentz invariance. In particular, the particles issued from collisions are distributed very unevenly over the θ coordinate. Therefore it is usually replaced by the pseudo-rapidity [112]:

$$\eta = -\ln \frac{\theta}{2}$$

This quantity has the nice property of being a good approximation of the rapidity for highly relativistic particles:

$$\eta \approx y = \frac{1}{2} \ln \left(\frac{p + p_z}{p - p_z} \right)$$

Differences of pseudo-rapidity is therefore Lorentz invariant and the particle production is equally distributed in the coordinate.

The two systems of coordinates are illustrated in Figure 1.4.

1.2.1 The tracker

The inner tracker (shown in concentric black lines in Figure 1.3) is the first detector encountered by the particles produced in a collision. Due to the high number of particles ($\mathcal{O}(1000)$) produced from overlapping proton-proton collisions during each Bunch CROSSing (BX), a detector with high-granularity as well as sufficient time resolution is required. Those two requirements allow the reliable identification of the trajectories and their association with the correct bunch crossing. The momentum of the particles is extracted from the bending of the tracks and the production point is estimated from backward interpolation.

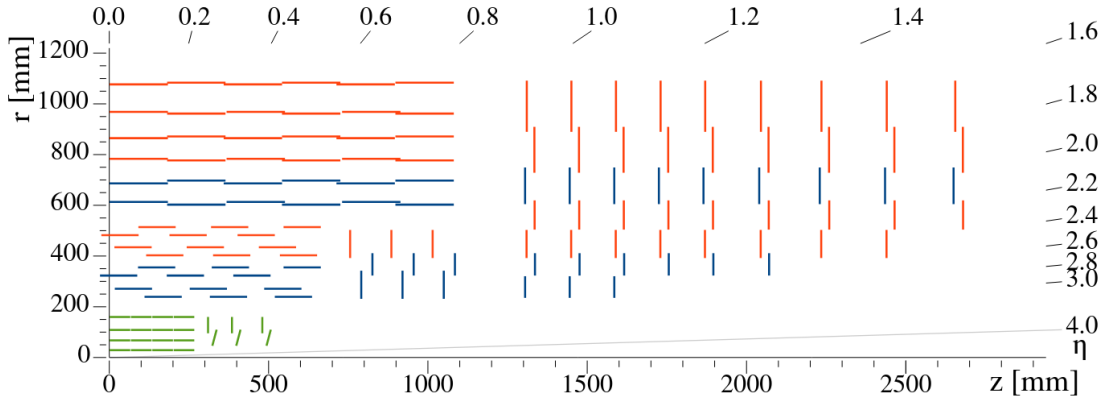


Figure 1.5: Schematic view of one quadrant in the RZ plane of the CMS tracker. The single-sided (double-sided) strip modules are colored in red (blue). The pixel detector is shown in green. [46]

The CMS tracker occupies a cylindrical volume of 5.4 m of length by 2.4 m of radius at the heart of CMS and is composed of two subsystems: the silicon pixel and the Silicon Strip Tracker (SST) both fully based on silicon sensors. A schematic view of the different layers of the CMS tracker is available in Figure 1.5. It achieves performances of 20 to 25 µm in primary vertex reconstruction [50].

Silicon pixel detector The innermost part of the CMS tracker is the silicon pixel detector. It provides high-resolution three-dimensional measurement points close to the LHC interaction point allowing the precise spatial reconstruction of primary and secondary vertices.

Since its replacement in 2017 with the so-called Phase-I upgrade, four detection layers are available in any direction up to a pseudo-rapidity $|\eta| < 3$. Due to the intense radiation, the first layer of the barrel was already replaced during LS2 in order to maintain the tracking performance throughout Run-3. It consists of four barrel layers (BPix) located at radii of 29, 68, 109, and 160 mm from the beamline, and three disks (FPix) located at 291, 396, and 516 mm from the center of the detector. A total of 1756 modules each comprised of 160×416 pixels of a size of $100\,\mu\text{m} \times 150\,\mu\text{m}$ are used, totalling 124 million readout channels.

Silicon strip tracker In addition to the pixel detectors, the inner tracking system is completed by the silicon strip tracker covering pseudo-rapidity in the range $|\eta| < 2.5$. It is composed of ten layers in the barrel region: four in the Tracker Inner Barrel (TIB) and six in the Tracker Outer Barrel (TOB). In the endcap regions, the TIB is completed with three layer Tracker Inner Disks (TID) made out of three rings each, whereas the TOB is completed with nine layers of Tracker EndCaps (TEC) made out of up to 7 rings. The full tracker comprises 15182 modules with a pitch size ranging from 80 µm to 180 µm for a length of 10 cm. Some of the modules, called stereo modules and colored in blue in

Figure 1.5, are used to provide a coarse measurement of the coordinate orthogonal to the strips. This is achieved by mounting two detectors back-to-back with an angle of 100 mrad. The SST amounts for 9.3 million readout channels.

1.2.2 The calorimeters

As mentioned earlier, the CMS calorimeter system is composed of an (internal) Electromagnetic Calorimeter (ECAL) and an (external) Hadronic Calorimeter (HCAL). The former measures the energy of photons and electrons whereas the latter measures the energy of the hadrons, consisting mainly of neutrons, protons, and pions at this distance from the interaction point.

Electromagnetic calorimeter The CMS ECAL is made out of 75 848 $PbWO_4$ crystals which have a section of 22 mm × 22 mm for a length of 23 cm in the barrel and a section of 26.8 mm × 26.8 mm for a length of 23 cm in the endcaps. The barrel part is readout through Avalanche Photodiodes (APD) whereas the endcap part is readout through Vacuum PhotoTriodes (VPT).

The barrel ECAL covers the range $|\eta| < 1.479$ and has an energy resolution of [104]:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{2.7\%}{\sqrt{E}}\right)^2 + \left(\frac{210}{E}\right)^2 + 0.55\%$$

For its part, the endcap ECAL covers the range $1.479 < |\eta| < 3$ and has an energy resolution of [104]:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{5.7\%}{\sqrt{E}}\right)^2 + \left(\frac{245}{E}\right)^2 + 0.55\%$$

Located in front of the barrel and endcap crystals, the preshower detector has been installed to distinguish high-energy photons from close pairs of low-energy photons originating from the decay of neutral pions. It consists of two lead absorbers followed by silicon sensors.

Hadronic calorimeter The CMS HCAL is mainly divided into a barrel region (HB, Hadron Barel) and two endcap regions (HE, Hardon Endcap), and provides coverage up to $|\eta| < 3$. The limited amount of space in the barrel region, and thus stopping power, between the ECAL and the CMS solenoid however calls for an outer calorimeter (HO, Hadron Outer) which allows to catch the tail of the hadronic showers. Additionally, the forward calorimeter (HF, Hadron Forward) is installed further down the beam pipe at 11.2 m from the interaction point and increases the coverage up to $|\eta| < 5.2$. This configuration is depicted in Figure 1.6.

The HCAL is designed as a sampling calorimeter where the HB and HE absorbers are made out of brass - except for the inner and outer layers - alternating with layers of scintillating plastic. Originally readout by Hybrid PhotoDetectors (HPD), the readout system has been upgraded to Silicon PhotoMultipliers (SiPM) as part of the Phase-I

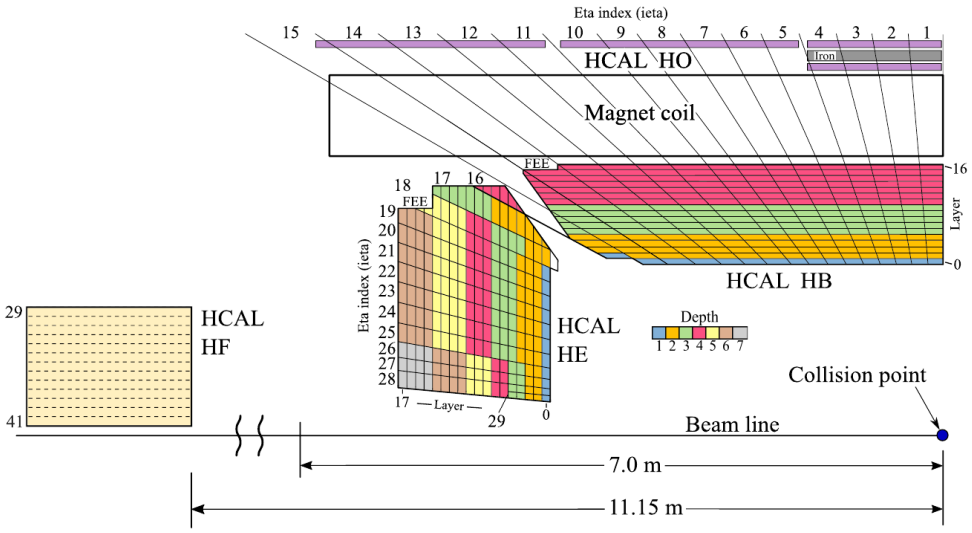


Figure 1.6: Longitudinal view of the HCAL showing its segmentation and regions for Run 3. [101]

upgrades. The HF region is based on the sampling calorimeter technique as well with absorbers only made out of steel and a readout based on PhotoMultiplier Tubes (PMT).

The HB and HE reach an energy resolution of [18]:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{120\%}{\sqrt{E}}\right)^2 + 9.5\%$$

whereas the HF reaches an energy resolution of [5]:

$$\left(\frac{\sigma}{E}\right)^2 = \left(\frac{280\%}{\sqrt{E}}\right)^2 + 11\%$$

1.2.3 The solenoid magnet

The superconducting solenoid magnet (colored in grey in Figure 1.3) forms the centerpiece of CMS and produces the magnetic field essential to the identification of the charged particles. It bends the trajectories of the particles, allowing for the measurement of their momentum and the determination of their charge.

The CMS magnet measures 12.5 m in length for an inner diameter of 6 m and a weight of 220 t [69]. The 1.5 m thick iron yoke adds another 10 000 t to the overall magnet system. Thanks to the extremely low resistance of its NbTi supraconducting material, it is able to produce a magnetic field of 3.8 T in its internal part - while the external magnetic field drops to 2 T - when powered at 18.16 kA.

1.2.4 The muon spectrometer

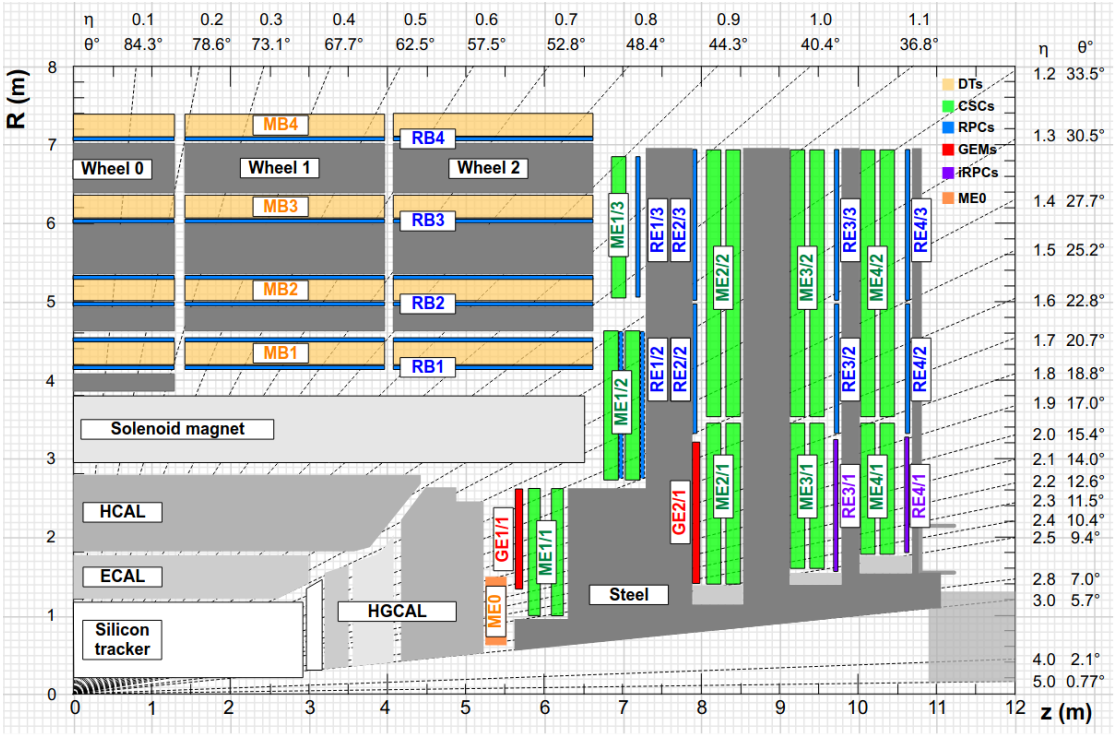


Figure 1.7: Location of the different detectors used in the muon system. The collision point is in the bottom left corner. The DTs (in light orange) are only present in the barrel. The CSCs (in green) are located only in the endcaps. The RPCs (in blue) are located both in the barrel and the endcaps. The Phase-II detectors are GE1/1 and GE2/1 (in red), ME0 (in orange), and the iRPC (in purple). [45]

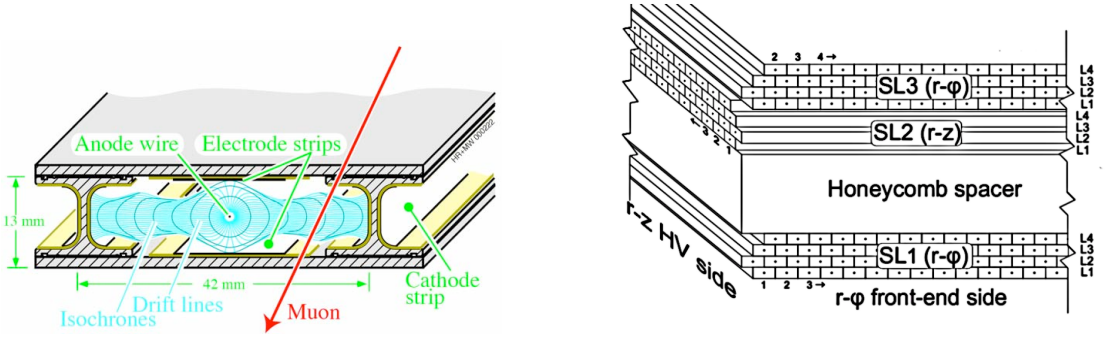
As suggested by the experiment's name, the muon detectors play a very important role in CMS and are a cornerstone of the project since its design phase. Located in the outermost part of CMS (in orange in Figure 1.3), this subsystem can be compared to a giant tracker dedicated to the muon detection. It aims at providing a robust and precise muon identification, momentum measurement, and triggering capabilities.

As the muons have a high penetrating power, locating the muon stations in the outside part of the experiment is well suited. The sub-detectors in the inner part of the experiment, the solenoid, and the return yoke create a shielding that filters particles other than muons in order to produce a clean signal; the rate of background particles is lower, so is the radiation; the space constraints are less; and the accessibility increased. The main drawback of placing the muon detectors behind the return yokes is the degradation of the track positioning. Indeed, the muons undergo multiple scattering and placing the detectors behind a larger amount of material will increase the deflections.

The gaseous detector characteristics make them ideal candidates due to their reliability, robustness, and relatively low cost. Until Run-3, the CMS experiment was composed of three technologies of gaseous detector as depicted in Figure 1.7 for a coverage area of around 25 000 m^2. The Drift Tubes (DT) are present only in the barrel while the Cathode Strip Chambers (CSC) are located solely in the endcaps. The Resistive Plate Chambers (RPC) are present both in the barrel and the endcap. Since Run-3, the GE1/1 station based on the Gas Electron Multiplier (GEM) technology has been added as part of the Phase-II upgrades. All CMS muon chambers are slotted between the return yokes made of steel.

The DTs and the CSCs provide a good spatial resolution ($\sim$ 100 µm) in order to accurately measure the momentum of the muon. For their part, the RPCs have a lower spatial resolution ($\sim$ 1 mm), but they are fast and have a very good time resolution ($\sim$ 1 ns). As the muons subsystem is an essential part of the trigger by the clean signature muons provide, the good time resolution is a significant advantage. Moreover, the usage of two independent technologies provides redundancy to inefficiencies or false positive in the trigger [8].

Drift tube



(a) Cross-section of one tube cell with a representation of the electric field [47].

(b) Schematic view of the elementary tube cells within a single chamber [100].

Figure 1.8: Illustrations of a CMS DT chamber.

The DT basic element are drift tube cells of rectangular sections 13 mm × 42 mm for a length of 2.4 m as represented in Figure 1.8a The anode is composed of a single 50 µm gold-plated stainless steel wire stretched in the chamber length while the cathode is composed of two strips on two sides of the chamber. Additionally, the electric field is shaped with electrode strips running in length on the top and bottom part of the cell. The cathode and electrode strips are respectively set at a voltage of −1200 V and −1800 V whereas the anode wires are operated at a voltage optimized individually for each chamber and ranging from 3000 V to 3600 V. The tubes are filled with a gas mixture of 85%Ar and 15%CO2 which offers ideal properties for those conditions and provides a drift velocity around 55 µm ns^{-1}.

In the CMS barrel, four layers of tube cells are stacked with an offset of half a cell to form what is called a superlayer. Three (or two) superlayers are then assembled to form one DT chamber. In the three superlayers version, one of them is orthogonal to the two others to measure the Z coordinate. A schematic of this organization is shown in Figure 1.8b.

The long deadtime associated with the size of the tubes limits the use of the DTs to low particle rate regions such as in the CMS barrel ($\sim 10\,\mathrm{Hz/cm}$), covering a pseudo-rapidity range $|\eta| < 1.2$. Single chamber reconstructed segments reach a spatial resolution of 100 µm for a time resolution of 5 ns.

Cathode strips chamber

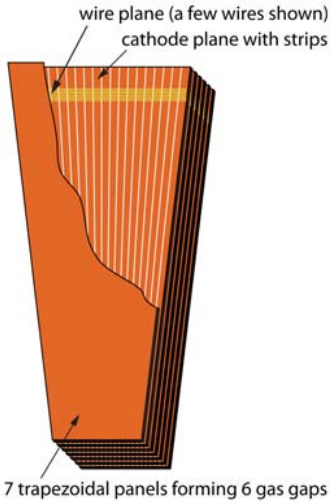


Figure 1.9: Layout of a CSC chamber made of 7 trapezoidal panels, forming 6 gas gaps. A few anode wires as well as cathode strips are revealed by the top panel cut-out. The former, the wires, are stretched in the azimuthal direction, providing measurement in the radial direction. The latter, the strips, run radially and provide a measurement in the azimuthal direction [47].

The CSC are multiwire proportional chambers optimized for the endcaps of CMS. They are composed of 6 layers of anode wires and 7 layers of finely segmented cathode strips (see Figure 1.9). The nominal gas mixture is 40%Ar + 50%CO2 + 10%CF4 has ideal properties in terms of quenching power and aging, allowing for a gas gain of 7×10^4 with an electric field of up to 3.9 kV/cm.

One layer of anodes contains about 1000 wires spaced by 3.2 mm. They provide a coarse measurement in the radial direction. Each layer of cathode is made of about 80 strips orthogonal to the anode wire. They provide a precise measurement in the muon trajectory bending plane via the repartition of the charges induced on the cathode strips. This dual readout system enables the extraction of a 2D measurement in each of the gaseous volumes, and the multiple layers provide 3D segment reconstruction.

The CSC are used in the endcaps where the particle rate can be as high as $\sim 1\,\mathrm{kHz/cm^2}$, covering a pseudo-rapidity range $0.9 < |\eta| < 2.4$. The reached spatial resolution can be as good as 75 µm, depending on the position, with a segment time resolution better than 4 ns.

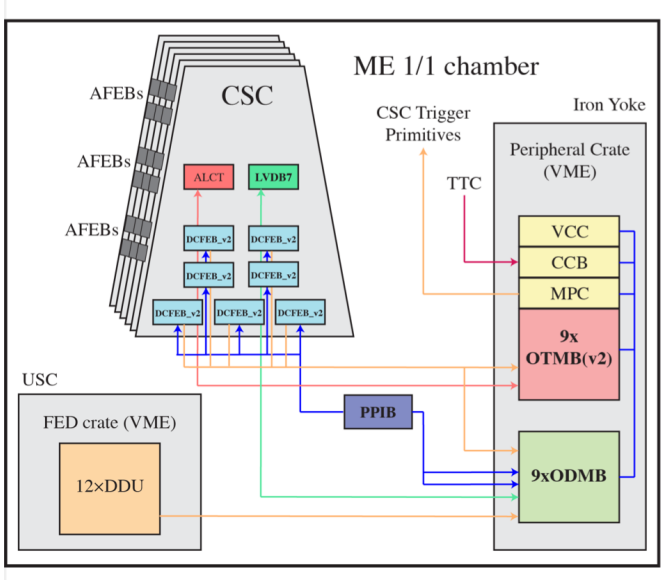


Figure 1.10: Details of the electronics present on the ME1/1 chambers [66].

The CMS CSC electronics [52, 66] is admittedly more complex than the one of the CMS GEM detectors (described further in Section 2.2.3). Figure 1.10 shows a schematic of the electronics for the ME1/1 station, companion of GE1/1, which can be divided into three groups: the on-chamber electronics, the off-chamber electronics in the Peripherical Crates (PC) located in the CMS Underground Experimental Cavern (UXC), and the FED (Front-End Driver) system located in the CMS Underground Service Cavern (USC).

The Anode Front-End Board (AFEB) amplifies and digitizes the anode signals from 16 wire groups, and sends the resulting data to the Anode Local Charged Track (ALCT) baseboard. The ALCT mezzanine present on each baseboard hosts a Field Programmable Gate Array (FPGA) that quickly finds particle tracks in the wire groups which are used as anode triggering information. The cathode strip signals are amplified and digitized the by Cathode Front-End Board (CFEB). Additionally, the Low Voltage Distribution Board (LVDB) and Motherboard (LVMB) respectively distribute the low-voltage power to the other on-chamber boards and provides monitoring features.

Moving on to the off-chamber electronics, the ALCT and CFEB of each chamber are connected to a pair of Data MotherBoard (DMB) and Trigger MotherBoard (TMB). The DMB is responsible for the data acquisition (DAQ) and collects data from the ALCT, CFEB, and TMB. The TMB is responsible for combining the triggering information from the ALCT and CFEB to build Local Charge Track (LCT) candidates that will be

further propagated to the Level-1 trigger. Those two cards are hosted in VME crates, each responsible for 9 CSC chambers. Additionally, each crate contains one Clock and Control Board (CCB) - which forwards the CMS clock and fast control signals to the CSC system -, one VME Crate Controller (VCC) - which allows slow-control communication with the on-chamber and off-chamber electronics for control and monitoring purposes -, and one Muon Port Card (MPC) - which collects the LCT from the TMB and forwards them to the Level-1 trigger.

The FED system for its part comprises Detector-Dependent Unit (DDU) boards hosted in VME crates and is responsible for collecting the data from all DMB and sending it to the central CMS DAQ system.

The particle rate increase in HL-LHC is expected to lead to large buffer overflows and thus significant event losses and readout inefficiencies. In anticipation of LS3, the on-chamber electronics of the innermost rings have been upgraded during LS2. The ALCT mezzanines have received larger FPGA with more embedded RAM for deeper buffers. The CFEB have been updated to their digital counterpart, the DCFEB and xDCFEB, which use flash ADC and digital pipelines rather than Switch Capacitor Arrays (SCA) to buffer the data before readout. The off-chamber electronics also received a set of updates in order to handle the new optical links to and from the on-chamber electronics: those are the ODMB and OTMB, the Optical counterparts of the DMB and TMB.

Significant improvements to the triggering algorithm during LS2 are the CCLUT (Comparator Code LookUp Table) and HMT (High-Multiplicity Trigger). The former improves the parameters estimation of the LCT stub, essentially acting as a precomputed fit, whereas the latter allows the detection of hadronic showers in the CMS muon subsystems, which are a signature of potential exotic long-lived particles (LLPs).

The upgrade program will be completed during LS3 with the replacement of the ODMB with newer versions, ODMB5 and ODMB7, as well as the removal of the MPC, and the replacement of the FED system.

Resistive plate chamber

The RPCs are gaseous parallel-plate detectors having an excellent time resolution. Figure 1.11 represents the section of a 2 gaps RPC detector. It consists of two chambers made of parallel electrodes coated with graphite and separated by only 2 mm. While this technology is usually operated in *streamer* mode, the CMS RPC are kept in *avalanche* mode by applying a voltage of 9.6 kV between its electrodes. In order to sustain the avalanches, the readout electrodes, located in between the two gaps, must be protected by a resistive layer made of bakelite, and the chamber filled with a mixture of 96.2% C2H2F4, 3.5% iC4H10, and 0.3% SF6.

The very good time resolution of the RPCs is achieved by producing faster avalanches than the previous two detectors in smaller gaps. As a result, the time resolution is as good as 1 ns at the expense of a spatial resolution of 1 mm. These specific characteristics make them ideal to unambiguously associate an event to the right bunch crossing and participate in the trigger. The main drawback is their inability to operate in high particle rate, which limits their pseudo-rapidity coverage to $|\eta| < 1.9$.

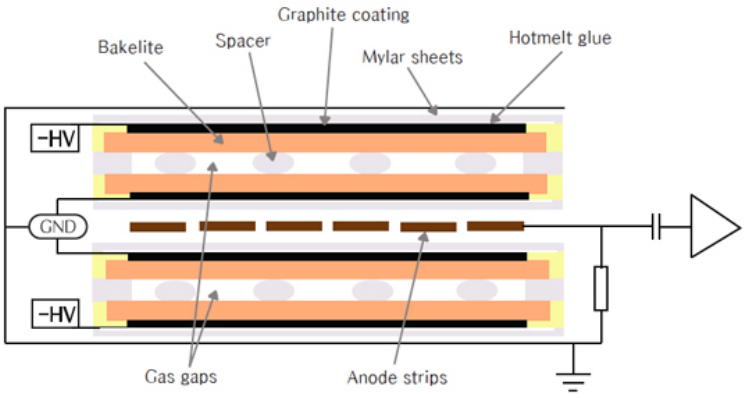


Figure 1.11: Sectional view of a 2 gaps RPC detector. The readout strips (in brown) are sandwiched by the two gas volumes. The bakelite plates which protect the electrodes are colored in orange. [93]

Gas electron multiplier

The GEM detector is the most recent addition to CMS via its GE1/1 station which has been installed during the LHC LS2. As the main topic of this dissertation, the detector and its electronics will be covered in details in Chapter 2.

1.2.5 The trigger and data acquisition systems

At the CMS interaction point, a collision between bunches occurs at most every 25 ns. As the event data generated for each collision has an average size of 1MB, processing and storing every event is unpractical but also useless as interesting physical processes are not present in every bunch crossing (BX). Currently, the data storage capabilities "only" allow for a maximum recorded event rate of a few kHz.

The online trigger system is responsible for selecting the events of interest before their storage, achieving the much-needed rate reduction. The CMS experiment has chosen a two-levels trigger approach as represented in Figure 1.12. The Level-1 Trigger (L1T), implemented in custom electronics, reduces the rate to a maximum of approximately 120 kHz with a latency of 4 µs based on coarse signals from the calorimeters and the muons detectors. The High-Level Trigger (HLT), based on a farm of commercial computers, processes fully assembled events with algorithms similar to those used in the offline reconstruction in a few hundred milliseconds.

The Level-1 trigger system

The Level-1 trigger is designed to reduce the event rate from 40 MHz to about 120 kHz. To achieve this goal, it receives the trigger data, a coarse-grained dataset, from the calorimeters and the muon subsystem. During this decision process, the complete dataset

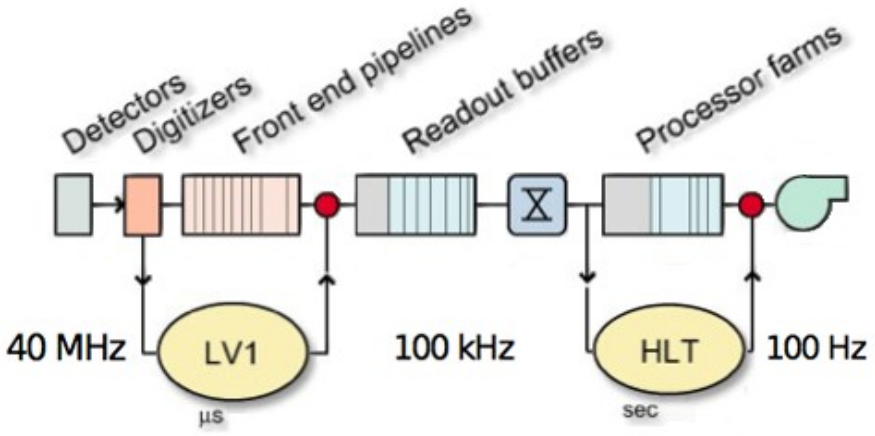


Figure 1.12: Overview of the CMS trigger system. The yellow bubbles represent the two levels of trigger processing. They successively select the events to a rate that can be stored.

remains stored in pipeline memories on the frontend electronics. That is why the L1T must issue a decision every 25 ns with a fixed latency of 4 µs.

The Run-2 upgraded L1T is fully implemented in custom-developed hardware, using Xilinx Virtex-7 Field Programmable Gates Arrays (FPGAs), Application Specific Integrated Circuits (ASICs), and LookUp Tables (LUTs) integrated in Advanced Mezzanine Cards (AMCs), building blocks of the µTCA architecture [96]. The data transfers between modules profit from multi-Gb/s serial optical links.

Figure 1.12 presents the architecture of the CMS L1T used from the beginning of Run-2. Two main paths can be identified: the calorimeter trigger which processes data coming from the ECAL and HCAL and generates for electrons, photons, and jets; and the muon trigger which processes data from the muon subsystems and generates muon candidates.

The calorimeter trigger consists of two layers. The first layer processes the trigger primitives (TPs) coming from the ECAL and HCAL with a granularity of $\Delta\eta \times \Delta\phi$ of 0.87×0.87 in the barrel and 0.17×0.17 in the endcaps. It applies LUT-based energy corrections to the deposits, sorts them, and finally forwards them to the second layer. The second layer reconstructs, calibrates, and sorts physics objects such as electrons and photons, tau leptons, and jets. As the calorimeter trigger uses a time-multiplexed design, the full event information is available on each board, alleviating the regional boundaries and providing the full information to compute energy sums. However, as the trigger chain downstream does not use the same concept, an additional demultiplexer board (DeMux) is required to format the events to be sent to the µGlobal Trigger (µGT) processors.

The muon trigger system consists of 3 independent Muon Track Finders (MTFs)

depending on the region of the detector covered. The Barrel Muon Track Finder (BMTF) covers the pseudo-rapidity region $|\eta| < 0.83$ and takes as inputs the DTs and RPCs. The Overlap Muon Track Finders (OMTFs) cover the pseudo-rapidity regions $0.83 < |\eta| < 1.2$ and take as inputs the CSC, DTs, and RPCs. The Endcap Muon Track Finders (EMTFs) cover the pseudo-rapidity regions $1.2 < |\eta| < 2.4$ and take as inputs the CSCs and RPCs. Additionally, in the barrel, the DT and RPC TPs are first pre-processed by the TwinMux boards to form the so-called "super-primitives", which benefit from the DTs higher spatial resolution and the RPCs better timing.

Each MTFs identifies muon candidates in their respective regions and assigns parameters (e.g. p_T) to the tracks. In the BTMF, the track assembler unit is responsible for identifying muon candidates by extrapolation of the super-primitives via LUTs. The track parameters assignment is done via LUTs as well. The OTMB identifies and assigns parameters to the muon candidates via fine-grained pattern matching. The EMTF finds muon candidates via a pattern-matching logic and the parameters are assigned via LUTs trained on a Boosted Decision Tree (BDT).

At the end of the muon trigger path, the µGlobal Muon Trigger (µGMT) receives up to 108 muon candidates from the three muon track finders, sorts them, removes the duplicates, and forwards the best 8 candidates to the µGT.

Finally, the µGlobal Trigger (µGT) combines the muon candidates and calorimeter objects, and executes every algorithm present in the menu to issue the final trigger decision.

The event builder

As the events are selected for readout by the Level-1 trigger, the data is extracted from the front-end buffers and shipped to the central CMS DAQ where the full event building occurs. Figure 1.14 represents an overview of the Run-3 CMS DAQ system.

The Event Builder (EvB) is split into two stages. The first stage, called *FED Builder*, aggregates in RU/BU (Readout Unit/Builder Unit) computing nodes event fragments from around 760 detector back-end boards, called FED (Front-End Driver) using an Ethernet switched network. Each set of FEDs is assigned one specific RU/BU node. The fragments are transmitted from the FED to the underground FEROL (Front-End Readout Optical Link), using custom protocols in the SLink family, where they are translated to standard TCP/IP. The second stage, called *Core Event Builder*, assembles all fragments for a given event to a specific destination RU/BU node, assigned on a per event basis, using a 100 Gbps switched network. In the Run-3 DAQ system, up to 62 nodes participate in the process, each able to handle around 2 kHz of fully-built events which are then stored in a 200 GB large RAM buffer. In total, the event builder must handle a throughput of about 100 GB/s.

In parallel to the so-called Global runs in which all sub-detectors participate and take physics data, the Run-3 DAQ system allows the so-called MiniDAQ runs. Those runs are traditionally used by one or a reduced number of sub-detectors for calibration, commissioning, or test activities, and have been proven extremely precious during operations.

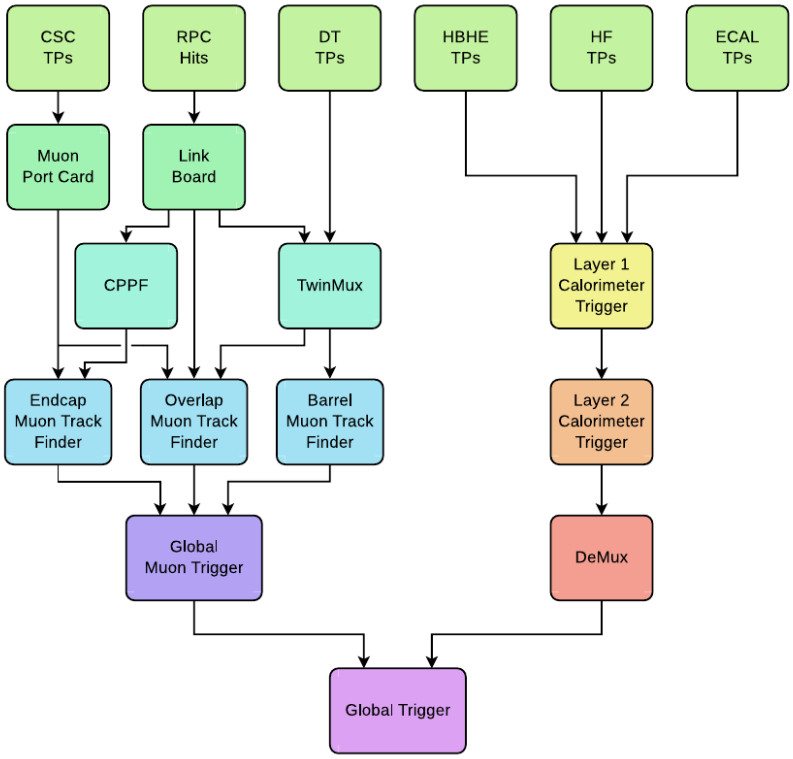


Figure 1.13: Block diagram of the CMS Level-1 trigger used during Run-2 and Run-3 [96].

The high-level trigger

Running on a cluster of 200 nodes, called Filter Units (FUs), each equipped with two AMD EPYC "Milan" 7763 CPUs, two NVIDIA T4 GPUs, and 256 GB of memory, the HLT consists in the second stage of events filtering. It analyzes and selects events interesting for future physics analyses at a rate of approximately 120 kHz, accepting events at a rate of a few kHz. As mentioned, the reconstruction and filtering are performed in software, implemented within the standard CMS SoftWare framework (CMSSW) [21]. The usage of the same framework, code base, and algorithms for the online and offline reconstruction reduces the development, testing, and maintenance needs. Support for offloading to accelerators, such as GPUs, has also been added at the start of Run-3, significantly reducing the average processing time from 690 ms per event to 385 ms and acting as real-world demonstrator for Phase-II upgrades.

Since Run-2, the HLT uses a File-based Filter Farm (FFF) approach to fully decouple the input data flow, processing algorithms, and output data flow [10]. The RU/BU ram buffers in which the fully-built events are written are mounted on the FU nodes via the NFS protocol. On each of the FU nodes, a Python-written application orchestrates the HLT operations. Whenever new raw data files appear in the ram buffers, signaling the

start of a new run, the HLT daemon starts the CMSSW applications and manages their lifetime (e.g. monitor their status and restart them in case of unexpected failure).

The HLT daemon also combines all selected events belonging to a single data stream and writes them back to the RU/BU nodes. From there on, the Storage and Transfer System (STS) will progressively merge together the streams coming from all RU/BU nodes and store them on a 1.2 PB shared Lustre filesystem. At this stage, the data can be transferred out of the CMS local cluster to the Tier-0 located at CERN for repacking, complete reconstruction, archiving, and distribution in worldwide data centers for physics analyses.

The timing and control distribution system

From the start of Run-2, the Timing and Control Distribution System (TCDS) [68] distributes the LHC reference clock along with fast control information synchronous with the LHC beam, such as the Level-1 Accept (L1A) signal, to all sub-detector back-end and/or front-end boards. In the other direction, the TCDS receives and merges the readiness status from all sub-detectors. Both of these interactions are required to keep the data-taking synchronized across the numerous sub-detectors.

The TCDS stems from an evolution of the Run-1 TTC & TTS (Timing, Trigger & Control, & Trigger Throttling System) systems to provide more independent detector partitions, support the new µTCA systems, simplify the operations by using a more unfired approach, and bring new functionalities (e.g. add synchronization markers for luminosity measurements). As part of the Run-2 trigger and DAQ upgrades, the TCDS is implemented in the µTCA architecture, maintaining perfect hardware compatibility with the legacy systems.

1.2.6 The Phase-II upgrades

In order to withstand the massive increase in luminosity foreseen by the HL-LHC, CMS, similarly to the other experiments at the LHC, must be upgraded. Whereas a pile-up (number of proton-proton collisions in a given bunch crossing) of 65 is routinely reached during Run-3, the most promising scenarios for the HL-LHC foresee a pile-up of 200. Not only it provides more data for the physics analyses, but it also significantly increases the particle rates, thus the irradiation levels in the detectors and their electronics, and the readout data rates.

A program of detector upgrades, spanning from LS2 to LS3, has been specifically designed to address those challenges. All current detectors will require changes: some will be fully replaced to enhance their granularity, recover the damages due to irradiation accumulated so far, or enhance their functionalities, whereas others will only receive electronics upgrades to cope with the new requirements. Some brand new sub-detectors will also be added to the CMS detector aiming to fully exploit the LHC physics potential. The CMS Phase-II upgrades are briefly discussed below:

- The Level-1 discrimination power will be enhanced by allowing more time for computation, bumping the latency from 4 µs to 12.5 µs, and by using more and higher

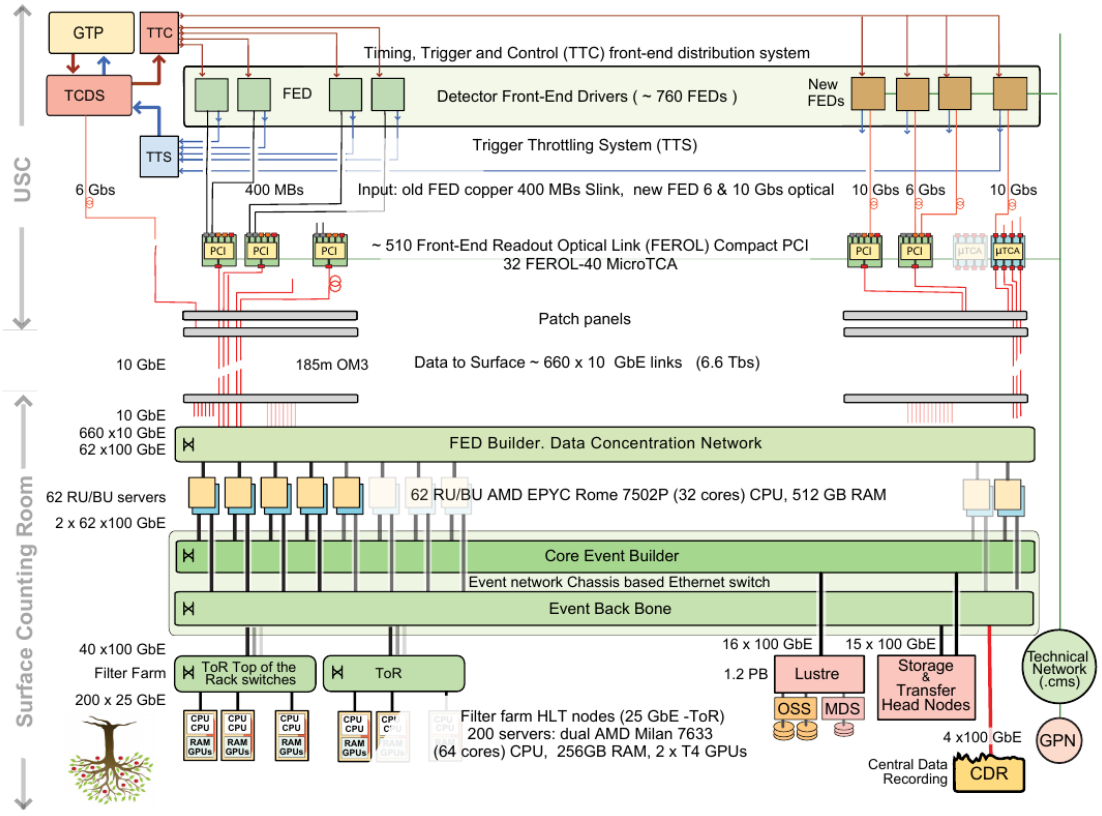


Figure 1.14: Diagram of the Run-3 DAQ system. [39]

granularity triggering information, particularly from the tracker subsystem [44]. Following the success of the heterogeneous computing model in Run-3, the HLT will fully embrace this new acceleration technique [42]. Advances in electronics and computing will also allow to bump the Level-1 trigger rate from 100 kHz to 750 kHz, and the HLT output rate from 1 kHz to 7.5 kHz.

- The MIP Timing Detector (MTD) is a new detector planned for CMS Phase-II [38]. It aims at bringing the new ability to measure the production time of particles with a precision better than 50 ps. Such a capability is required to disentangle the 200 nearly simultaneous pile-up interactions. The barrel layer is based on crystals readout via SiPMs and the endcap layer is composed of Low Gain Avalanche Diodes (LGAD).

- To cope with the growing particle rate, the increased radiation levels, and the new requirements from the Level-1 trigger, the inner tracker must be completely replaced [46]. Its new fully silicon-based design features a granularity multiplied by a factor ~ 4 with respect to the Run-3 tracker, a much lower material budget - reducing the multiple scattering -, and a coverage up to $|\eta| < 3.8$. In addition, it will be able to provide track information directly to the Level-1 trigger.

- The barrel ECAL will be equipped with new front-end electronics boards featuring trigger-less readout at 40 MHz to meet the trigger and rate requirements [41]. The electronics upgrade also aims at fully exploiting the crystal and APD precise timing for pile-up mitigation thanks to the usage of Analog-to-Digital Converter (ADC) sampling the signals at 160 MS/s. On the detector side, the same crystals will be used but cooled at a lower temperature to reduce the damage from irradiation [41].

- The barrel HCAL active material and new electronics installed in LS2 will not require any change and will continue to be used in Phase-II. However, the µTCA-based back-end electronics requires to be upgraded in order to sustain the planned 750 kHz Level-1 trigger rate. The back-end board hardware and its development are mutualized with the barrel ECAL ones.

- The endcap calorimeters however will be totally replaced by a new calorimeter design integrating both the ECAL and HCAL functions. This new calorimeter, called High Granularity Calorimeter (HGCAL) [43], aims at providing a granularity of roughly 1 cm^2 with time resolution ranging from 20 ps to 150 ps. Such capabilities allow a full 4D reconstruction of the electromagnetic and hadronic showers for enhanced pile-up rejection.

- Whereas the actual muon chambers are able to cope with the increased particle rate and radiation, their electronics must be upgraded. The DT will see both its front-end and back-end electronics replaced whereas only the RPC back-end electronics requires an upgrade [45]. The situation of the CSC, which stages its upgrades between LS2 and LS3, has been discussed in Section 1.2.4.

- The GE1/1 station will be joined by the GE2/1 and ME0 stations, extending the muon sub-detector coverage to $|\eta| < 2.8$. Sections 2.2 and 2.3 cover these detectors in more detail.

- In parallel, the RPC subsystem will be complemented with an improved version of RPC (iRPC) [45] in the region $1.8 < |\eta| < 2.4$. These new RPC detectors feature thinner electrodes and narrowed gas gaps, both of 1.4 mm, and are operated at a lower high voltage with respect to the currently installed RPCs. Effectively reducing the gain, the aging process is slowed down, at the expense of requiring higher gain in the electronics and better signal-to-noise ratio.

As one can notice, the back-end trigger and readout electronics of all sub-detectors require a full replacement. Taking advantage of this opportunity, CMS has decided to transition to a uniform platform based on the Advanced Telecommunications Computing Architecture (ATCA). Its implementation in the GEM case is described in Section 2.3.

Chapter 2

The CMS GEM project

As mentioned in the previous chapter, the GEM technology has been chosen to complete the muon spectrometer of CMS in order to face the challenges of the LHC Phase-II and maintain its performance.

This chapter is twofold. First, it introduces the working principle of gaseous detectors, with an emphasis on the GEM technology. Second, it describes the CMS GEM project and its motivations, focusing on the GE1/1 station and its electronics, while succinctly depicting the GEM Phase-II upgrades.

2.1 Gaseous detectors working principles

Gaseous detectors are widely used in High Energy Physics (HEP) thanks to their multiple advantages, the main ones being:

- they are relatively inexpensive – compared to silicon detectors for example – which makes them suitable to cover large areas;
- their sensitive volume, gases, have low densities which reduce the effect of multiple scattering as well as the energy losses when a particle crosses them.

In this section, we review the generic working principles of gaseous detectors focusing on the parameters most relevant for the intrinsic efficiency and time resolution, that is the ionization of the gas mixture and drift velocity [65, 25]. A detailed description the Triple-GEM detectors follows [95]. This section finally ends with a discussion on the intrinsic time resolution of a Triple-GEM gaseous detector.

2.1.1 Energy losses in medium

As particles travel through matter, they undergo multiple interactions with its constituents. In the case of charged particles these interactions can be: inelastic Coulomb scattering, bremsstrahlung, Cerenkov effect,... During these interactions, energy is lost and particles are scattered. This is the starting point of particle detection systems.

For the particles we are interested in, the muons, the dominant energy loss in a material results from electromagnetic interactions. A small fraction of the incoming particle energy is transferred to the electrons present in the material for each scattering. The Bethe-Bloch formula best describes the mean loss of energy of a particle in a medium [27]:

$$-\langle\frac{dE}{dx}\rangle = Kz^2\frac{Z}{A}\frac{1}{\beta^2}\left[\frac{1}{2}\ln\frac{2m_ec^2\beta^2\gamma^2T_{max}}{I^2} - \beta^2 - \frac{\delta(\beta\gamma)}{2}\right]$$

$$K = 4\pi N_A r_e^2 m_e c^2$$

where N_A is the Avogadro's number, r_e the classical electron radius, m_e the electron mass at rest, z the charge of the incident particle, Z and A the atomic number and the atomic mass of the medium and T_{max}, the maximum energy transfer in a single collision[1]. I is defined as the mean excitation energy of the medium while $\delta(\beta\gamma)$ is the density effect correction. The equation remains valid for $0.1 \lesssim \beta\gamma \lesssim 1000$.

Figure 2.1 shows the energy loss normalized by material density for muons going through copper. The Bethe-Bloch equation is drawn in red. At lower energies the particle velocity becomes comparable to atomic electron velocity while at higher energies the radiative loss becomes the dominant process, both increasing the total energy loss compared to the bare Bethe-Block equation.

The minimal energy loss is situated around $\beta\gamma \sim 3 - 4$. In HEP experiments, the particles of interest are typically at or above that threshold and are called Minimum Ionizing Particles (MIP).

It is important to emphasize that this formula only gives the mean total energy loss. Indeed, the interaction process is an ionization of the medium by the traversing particles. For each ionization, a localized group of electrons is deposited in the medium. This stochastic energy loss process is strongly asymmetric and is described by a Landau distribution. Only in the case of thick materials, the energy loss would become Gaussian due to the large number of interactions.

2.1.2 Gaseous detectors

Total ionization

The gaseous detectors take advantage of the ionization of the gas and the production of electron-ion pairs by the incoming ionizing particles. The electrons ejected during this process are called primary electrons, n_p. Some of these electrons may themselves possess enough kinetic energy to ionize new atoms and free new electrons, called secondary electrons, n_s. The group of electrons gathered around a primary ionization forms a cluster.

[1] $T_{max} = \frac{2m_ec^2\beta^2\gamma^2}{1+2\gamma_e/M+(m_e/M)^2}$ where M is the mass of the incident particle.

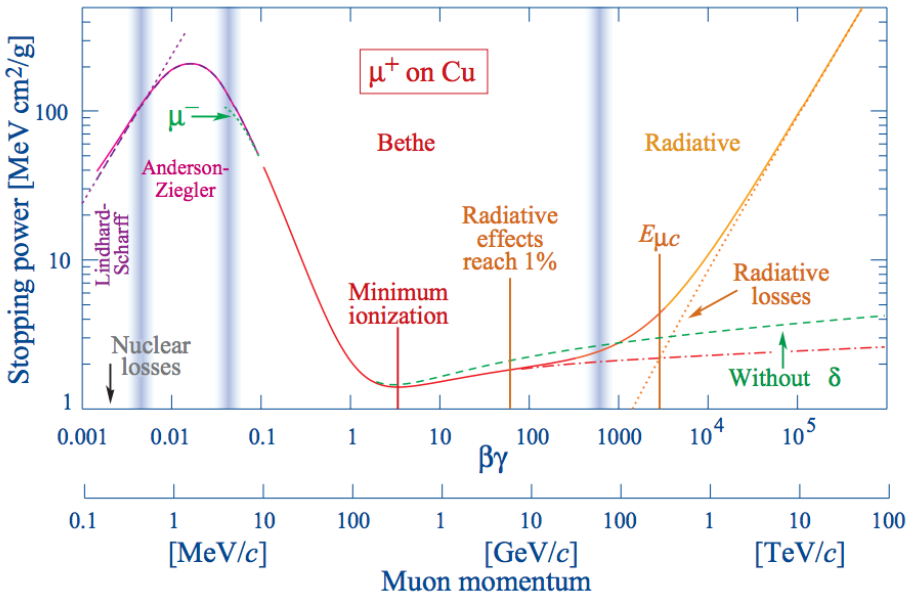


Figure 2.1: The energy loss for a positively charged muon in copper. The Bethe-Bloch formula is represented in red. [27]

The sum of the primary and secondary electrons forms the total ionization, $n_{tot} = n_p + n_s$. For a given path length, L, it can be computed by dividing the mean energy loss of the particle by the mean ionization energy of the medium, W_I:

$$\langle n_{tot} \rangle = \frac{\langle \frac{dE}{dx} \rangle L}{W_I}$$

The differential energy loss, the average primary and total number of electrons, and the average ionization potential is determined by the gas or gas mixture. They are listed in Table 2.1 for the gases typically used in GEM detectors.

Gas	$\frac{dE}{dx}\|_{min}$ [keV/cm]	W_I [eV]	n_p [cm^{-1}]	n_{tot} [cm^{-1}]
Ar	2.53	26	25	97
CO_2	3.35	34	35	100
CF_4	6.38	52	63	120
CH_4	1.61	30	28	54

Table 2.1: Properties of gases typically used in the Triple-GEM detectors at normal temperature and pressure (NTP: 20 °C, 1 atm) for a MIP. [27]

The intrinsic efficiency of a gaseous detector is directly related to the number of primary electrons. Indeed, the probability that a particle produces k primary ionizations while traversing a medium is dictated by a Poisson law with n_p as mean:

$$P(k) = \frac{n_p^k}{k!} e^{-n_p} \tag{2.1}$$

The intrinsic inefficiency is the probability that a particle does not produce any primary ionization while traversing the medium:

$$\text{inefficiency} = P(0) = e^{-n_p} \tag{2.2}$$

The required efficiency constrains the choice of the gas. As the noble gases usually have a lower ionization energy, they are the primary choice.

Signal amplification

The typical total ionization is about a few hundreds of e^- (for example 97 e^- are produced on average in 1 cm of argon for a MIP) while an electronic amplifier typically has an electronics noise of around a thousand e^-. If the signal was only amplified electronically, it would be completely masked by the noise. So, gaseous detectors amplify the signal by using an avalanche process.

A strong electric field (several tens of kV/cm) is produced in the gaseous detector chambers to accelerate the electrons. Once they have gained enough kinetic energy, they ionize the medium in turn. This process is repeated for each newly produced electron, creating an exponential growth of the avalanche.

Figure 2.2 details the formation of an avalanche around an anode wire. In Figures (a), (b), and (c) we can see the electrons moving toward the anode while the ions are drifting to the cathode. During the drift, electrons gain enough energy to free new electrons and develop the avalanche. This process stops when the internal electric field compensates the external field.

Figure (d) shows the avalanche totally surrounding the anode wire. In Figure (e) the electrons have been collected and the ions are slowly moving toward the cathode. In this geometry, the ions generate most of the measured signal.

The gain of the avalanche, also called the gas gain, is defined as:

$$A = \frac{N}{N_0}$$

where N is the final number of electrons while N_0 is the total number of electrons at the start of the avalanche.

N depends of the distance travelled by the electrons, D, and the first Townsend coefficient, α. The variation of the number of electrons along a path is thus given by:

$$dN = N\alpha dx$$

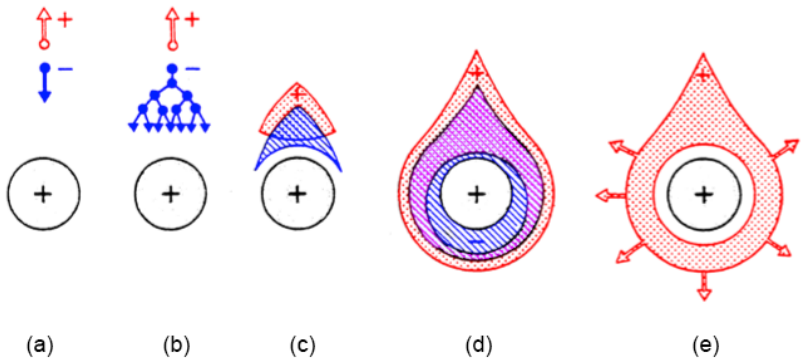


Figure 2.2: Formation of an avalanche around an anode wire. [94]

Its integration along a path gives the following relation:

$$N = N_0 \exp(\alpha D)$$

The first Townsend coefficient is an empiric parameter which depends on the applied electric field as well as the gas temperature and pressure. It is usually represented as $\alpha(E, p, T)$. While it has to be measured or simulated for each gas mixture, certain tendencies appear. In particular, noble gases tend to have a larger Townsend coefficient than more complex gases, since they have fewer degrees of freedom that can absorb energy.

However, noble gases can emit photons by electron-ion recombination with an energy high enough to ionize another atom via photo-ionization. That phenomenon could spoil the measurements by creating a secondary avalanche far from the main one, which would then be misidentified as another ionizing particle. It could also lead to operational concerns if the photon is able to ionize the electrode surface, leading to sustained avalanches. In order to avoid such unsought effects, it is possible to add a quenching gas, for example CO2, to the gas mixture. Those molecules hold plenty of vibration and rotation states which are able to absorb the photons emitted during the de-excitation of the other gases.

Drift velocity

As the electron-ion pairs are created and in the absence of an electric field, they diffuse in the gas and recombine. This is the reason why an electric field has to be applied to collect and focus the electrons toward the amplification area, in the direction of the anode. As they move through the gas, they are subject to both the drift caused by the electric field and the collision with the atoms and molecules of the gas mixture.

The drift velocity, v_{drift} of electrons in gases stems from the equation of motion of free charges in a gas, in the presence of an electric $\vec{E}$ and a magnetic field $\vec{B}$:

$$m\frac{d\vec{v}}{dt} = e\left(\vec{E} + \vec{v} \times \vec{B}\right) + \vec{Q}(t)$$

Since we are interested in the average velocity, we consider time intervals larger than the characteristic time between collisions of the electrons with the gas ($\Delta t \gg \tau$). We also focus on the typical situation for drift gas where the gaseous detectors have constant fields. The collisions with the gas atoms and molecules, represented by $\vec{Q}(t)$ in the above equation, can be thought as a friction force, proportional and opposite to the velocity. This yields to

$$\vec{v}_{drift} = \mu\vec{E} + \omega\tau(\vec{v}_{drift} \times \vec{B})$$

whose solution is

$$v_{drift} = \frac{\mu}{1+\omega^2\tau^2}\left(\vec{E} + \frac{\vec{E}\times\vec{B}}{B}\omega\tau + \frac{(\vec{E}.\vec{B})\cdot\vec{B}}{B^2}\omega^2\tau^2\right)$$

where $\mu = \frac{e\tau}{m}$ is the particle mobility, $\omega = \frac{eB}{m}$ is the cyclotron frequency, and τ is the average time between collisions with the medium [65].

One can notice that the velocity is proportional to the electric field and the mobility. Henceforth the lighter the drifting particle is, the faster it is, as expected. The drift speed of the electrons and positive ions plays an important role in the time resolution of the detector as we will see in Section 2.1.4.

2.1.3 GEM detectors

Over time, multiple technologies of gaseous detectors have been developed to improve their characteristics and meet specific requirements. In particular,

- increase the spatial resolution thanks to a higher readout granularity;
- increase the time resolution by having faster avalanches and drift velocities;
- improve the counting rate by collecting the positives ions faster and avoid disturbing the electric fields;
- increase the gain to improve the efficiency;
- reduce the risk of discharges.

As presented in section 1.2.4, CMS currently uses four different technologies of gaseous detectors: the drift tubes, the cathode strip chambers, the resistive plate chambers, and the Triple-GEM. The latter ones, introduced in 1997 by Fabio Sauli at CERN, are part of the larger Micro-Pattern Gaseous Detectors (MPGD) family, a novel technology based on microscopic structures to amplify the electrons.

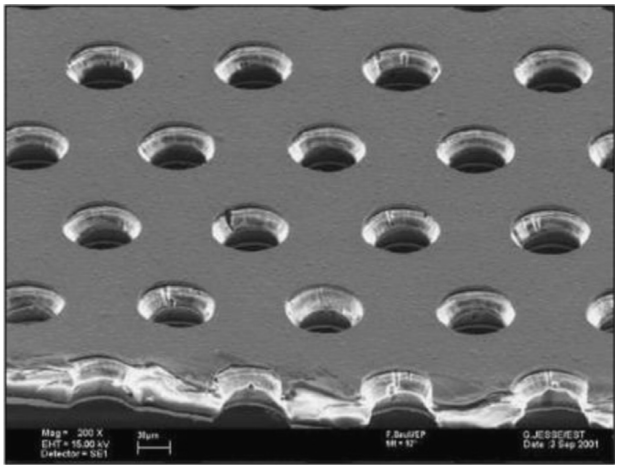


Figure 2.3: Electron microscope picture of a standard GEM foil. The holes have a typical inner diameter of 50 µm, outer diameter of 70 µm, with a pitch of 140 µm. [95]

One standard GEM foil is made of a polyimide foil, thick of 50 µm, and coated with a 5 µm layer of copper on each side. The foil is then etched through photolithography to obtain small regularly spaced holes. Typical hole dimensions are 50 µm for the inner diameter and 70 µm for the outer diameter, with a pitch of 140 µm. An electron microscope picture of a GEM foil is presented in Figure 2.3.

By applying a potential difference between the two sides of the GEM foil, one creates an electric field whose intensity is maximal within the holes as shown in Figure 2.4. For foils of small size ($\sim 10\,\text{cm} \times 10\,\text{cm}$), manufactured with good quality, the field within the holes can reach hundreds of kV/cm representing gas gains of 10^3. The avalanche process takes place only within the foils. This remarkable feature of GEM detectors provides significant advantages.

The first advantage is the decoupling of the amplification zone from the readout electrode, which avoids damage to the electronics in case of discharge.

The second advantage stems from the charge collection and signal formation speeds. One can see the electrons' drift path in blue and the ions' in red in Figure 2.4. A large fraction of the ions is quickly collected by the ion trap of the top of the GEM foil, which avoids ions backflow in the gas volume. Hence a reduction of the detector dead-time and an increased rate capability. Additionally, as the signal is formed by electrons, it is much faster than if it was formed by ions. This enables a better time resolution as we will see in the next section.

The third advantage, and probably most interesting one, is the possibility of sharing the total gas gain between multiple amplification stages. The high gain required for experiments would imply very high voltages across the GEM foil. The problem thus raised is the risk of discharges which would damage or destroy the detector and its

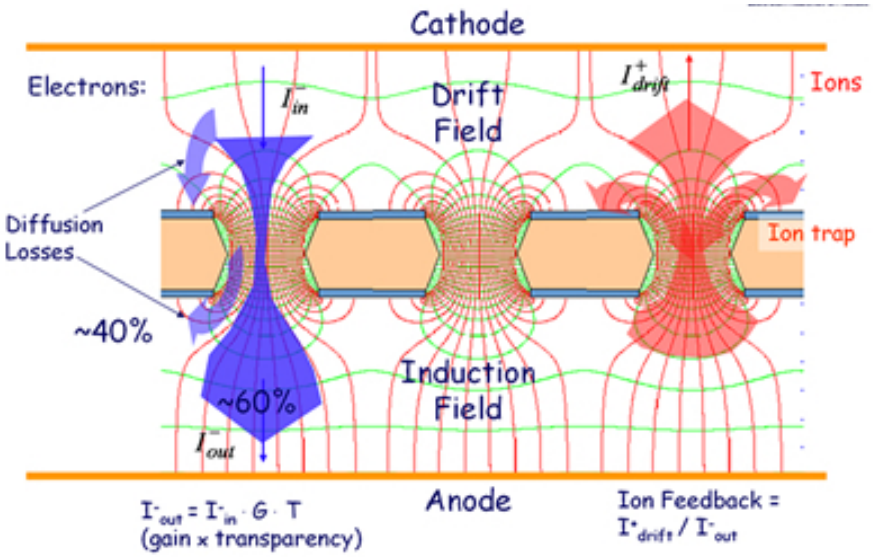


Figure 2.4: Schematic of a GEM foil encloses between a cathode and an anode. The electric field lines are colored in red. The drift paths of the electrons and the ions are represented in blue and in red respectively. [110]

electronics. The GEM technology provides a solution to this problem common within MPGD family by enabling multiple foils to be stacked. It is typical to stack 3 foils as shown in Figure 2.5, hence the Triple-GEM naming.

The effective gas gain, defined as the ratio between the detected charge and the primary ionization charge, is not solely determined by the voltage difference across the foils; it is also influenced by the voltage differences applied across the individual gaps. For the electrons to be amplified, they need to be focused in the holes, the amplification region, and not get lost on the copper on the top of the foil. The ratio between the number of electrons produced above the holes and collected in the holes is called *collection efficiency*. Similarly, the electrons must be confined well enough in the center of the holes to escape with enough energy to avoid being collected on the copper at the bottom of the foil. The ratio between the number of electrons extracted from the holes and produced inside the holes is named *extraction efficiency*. The combination of the collection and extraction efficiency forms the *foil transparency*. Typical examples of how the effective gas gain and transparency are affected by the high voltage parameters are depicted in Figure 2.6.

In the case of the CMS GEM project, R&D has established that the ideal geometry consists of 3 stacked GEM foils. As written in Figure 2.5, the terminology used for the different gaps is, in order: drift, transfer 1, transfer 2, and induction. The primary and secondary ionization takes place in the 3 mm thick drift gap and is amplified by the first GEM foil. The transfer gaps of 2 and 1 mm drive the electrons from one foil to the

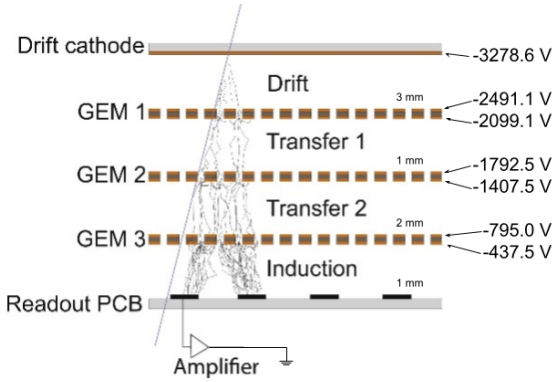


Figure 2.5: Disposition of the three foils in a triple-GEM detector. The voltages shown correspond to an equivalent voltage divider current of 700 µA. Based on [95].

next one. The 1 mm induction gap finally guides the electrons to the readout anode, participating in the signal shaping.

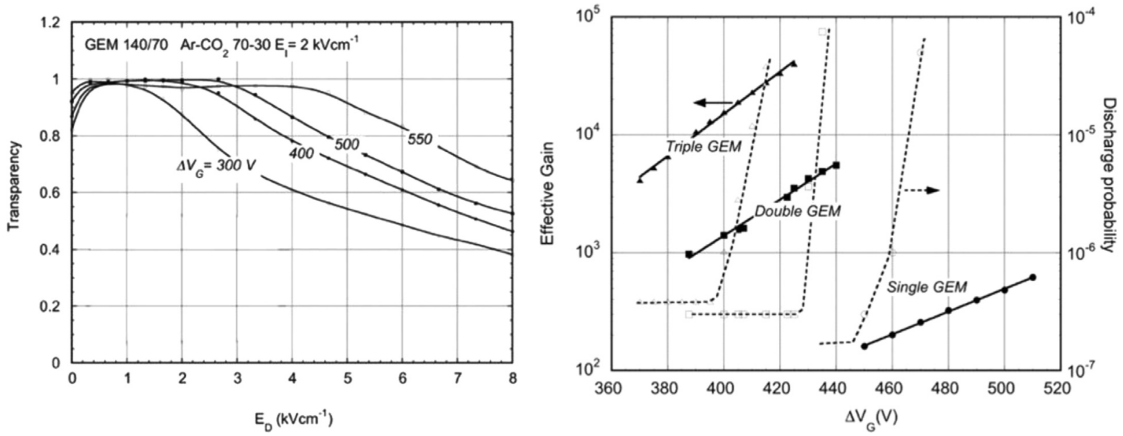
The operation of a single triple-GEM chamber requires seven voltages – one for each of its electrodes (two per GEM foil, plus one for the drift electrode). While these voltages are typically derived from a single source using a ceramic high-voltage divider – which imposes fixed ratios between the electrode voltages – this approach proved challenging in the CMS operating conditions. Indeed, it limits the ability to monitor the behavior of individual electrodes in details, partially masks the discharge transients, and prevents any adjustments in response to issues or optimization needs. For these reasons, the final CMS GEM system uses a stacked multi-channel high-voltage power supply, allowing independent control and monitoring of the voltages applied to each electrode. The voltage ratios remain based ceramic HV divider value, but are calculated in software. Therefore, the *equivalent divider current* terminology remains widely used throughout the project and in this document.

Similarly, the gas mixture judged the most appropriate during the R&D phase consists of 70% of Argon and 30% of CO_2. It is safe since not flammable; does not contribute to the greenhouse effect; and possesses excellent properties against aging for operation in harsh environments.

Together, they yield the best overall trade-off between efficiency, spatial resolution, time resolution, and discharge reduction.

2.1.4 Factors affecting the time resolution

As anticipated in the previous sections, the intrinsic time resolution of a gaseous detector depends on the gas mixture used and on the geometry of the detector. In the GEM case, the signal is formed by the successive electron avalanches produced by the multiple clusters created in the drift gap. Supposing that the cluster closest to the first



(a) (b)

Figure 2.6: (a) GEM foil transparency as a function of the gap field for different GEM foil voltages. (b) Effective gas gain and discharge probability as a function of the GEM foil voltage for different multi-GEM detector configurations. For both figures, ΔV_G is the voltage difference applied across the foil, and E_D is the electric field magnitude of the drift gap. Taken from [95].

electrode can create a strong enough signal, its drift time will determine the minimal time uncertainty.

Considering that the production of electron clusters follows a Poisson statistics, the space distribution of the i-th cluster created at a distance x from the first GEM foil can be expressed as:

$$p_i^n(x) = \frac{x^{i-1}}{(i-1)!} n^i e^{-nx}$$

where n is the number of primary electrons per unit of length.

As we are interested in the position of the first cluster and more specifically in its drift time $t_{drift} = x v_{drift}^{-1}$:

$$p^n(t_{drift}) = n v_{drift} \exp(-n v_{drift} t_{drift})$$

Therefore, the intrinsic time resolution goes as $(n v_{drift})^{-1}$ [8]. It is immediate from this relation that the number of primary ionizations that occurs in the drift region and the drift velocity play a critical role. A fast gas mixture releasing a large number of primary ionizations as well as a large electric field should be used in order to maximize the time resolution of the detector.

This quick derivation, however, only represents a lower limit by assuming that the first cluster is always amplified by the detector and measured by the electronics. If that is not the case, the second, third, etc. cluster will then produce the signal and degrade the detector time resolution. As discussed in the previous section, the GEM foil transparency is the factor that limits the collection of the electrons, thus the detector

time resolution. We deduce from Figure 2.6a, that the drift electric field then needs to be set as compromise between a reduced drift time and an high transparency.

A complete study of the GEM timing, including simulations and the impact of the parameters previously mentioned, can be found in [82].

2.2 The GE1/1 project

As explained in Section 1.2.6, the CMS muon subsystem needs to be enhanced to cope with the increased rate of particles produced at the HL-LHC. In CMS, the dominant background sources are neutrons, produced by the interactions of primary collision hadrons with the material of the beam pipe and the very forward structures, and photons, which originate from the neutron interactions with the material of the detector. In the GE1/1 region, the fluxes, integrated over the whole energy spectrum, are expected to reach peak values around $150\,\mathrm{kHz/cm^2}$ for neutrons and $76\,\mathrm{kHz/cm^2}$ for photons at the highest pseudo-rapidity ($\eta \sim 2.1$). Convoluted with the GEM detector sensitivity for each particle type, this results in a total hit rate expected to reach around $1.2\,\mathrm{kHz/cm^2}$.

The aim of the muon subsystem upgrade is thus to maintain the Level-1 muon trigger efficiency while improving the muon track reconstruction in high pile-up conditions as well as increasing the redundancy in the most forward – and challenging – region.

As briefly mentioned in Section 1.2.4, the GE1/1 project is part of the CMS Phase-II muon upgrades. Its new detectors have been installed during the LHC LS2 in the area that was left vacant for the innermost layer of RPC detectors in the muon endcap. If the installation of the RPCs was first staged for financial reasons at the start of the LHC program, it then became clear that standard RPCs could not sustain the particle rate ($> 1\,\mathrm{kHz/cm^2}$) and dose expected in this region during HL-LHC.

The primary goal of GE1/1 is to improve the muon transverse momentum measurement (p_T) within the combined GEM-CSC muon station at the L1T level. Thanks to a larger lever arm within the first muon station, a better measurement of the bending angle is achieved, hence a better estimation of the particle transverse momentum, and thus a reduction of the fake rate. The fakes triggers originate from two main effects:

1. The misestimation of the muon transverse momentum, where low-p_T muons are incorrectly promoted to higher p_T above the L1T threshold. This effect how on which the larger arm has the most impact.

2. The combinatorial background, where unrelated hits are mistakenly combined into fake muons tracks due to the high hits occupancy originating from background particles. This is the effect on which the additional measurement points have the most effect.

The combined anticipated trigger rate improvement is presented in Figure 2.7.

The secondary goal of GE1/1 is to provide additional redundancy in its region of the muon subsystem if any of the current detectors suffer degradation or failure due to the higher particle, larger integrated dose, or aging.

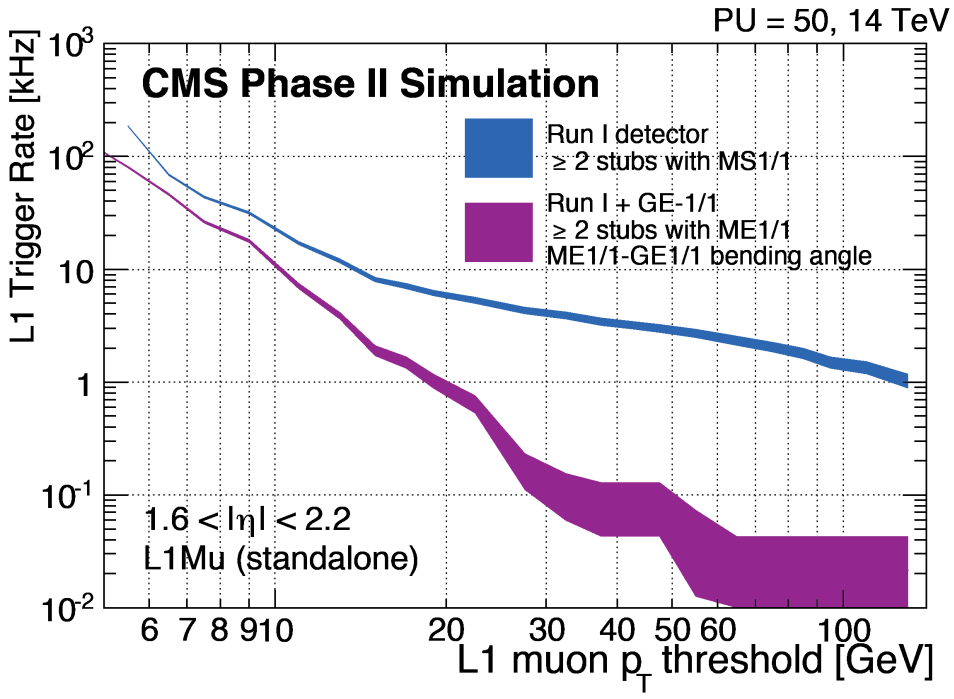


Figure 2.7: Level-1 muon trigger rates as a function of the L1T muon candidate track p_T threshold before (blue) and after (purple) the GE1/1 upgrade at a luminosity of $2 \times 10^{34}\,\mathrm{cm}^{-1}\,\mathrm{s}^{-1}$, a for constant efficiency of 94%. [1]

2.2.1 Requirements on GE1/1

In order to meet its physics objectives, a detector in this region of CMS should fulfill multiple requirements. Those have initially been defined in the original GE1/1 Technical Design Report (TDR) and have later been refined based on production experience [1, 62].

Geometric acceptance The complete detector should not let any blind spots in the region it covers, taking into account the limit envelope with only 10 cm of thickness, space left by the RPCs.

Rate capability During the HL-LHC era, the maximum hit rate is expected to reach $5\,\mathrm{kHz/cm^2}$. Considering a safety factor of 2, the required rate capability must be $10\,\mathrm{kHz/cm^2}$ or better.

Efficiency An efficiency of 95% or more must be reached for each single detection layer. The two detection layers thus provide a minimum efficiency of 99.7% per chamber.

Angular resolution The angular resolution should be better than 300 µrad in ϕ to reliably discriminate between low-p_T and high-p_T muons.

Timing resolution A timing resolution better than 10 ns must be reached to provide reliable Level-1 trigger primitives.

Gain uniformity Defined as the ratio between standard deviation and the mean of the distribution of the gain measured across all readout strips, the relative gain uniformity characterizes must be 37% or better within a chamber.

Radiation tolerance The detector is expected to receive an integrated charge of 100 mC/cm^2 after 20 years of operation in the HL-LHC. Taking into account a safety factor of 2, the detector must not suffer from any gain loss or other losses in response after 200 mC/cm^2 of integrated charge.

The COMPASS (Common Muon and Proton Apparatus for Structure and Spectroscopy) [49], TOTEM and LHCb experiments already developed and used medium-size (30x30 cm^2) Triple-GEM detectors with the required performances. These encouraging results showed that the technology could be used in the CMS muon subsystem. However, it was not clear at that time if building even larger Triple-GEM detectors (1 m in length) was possible while meeting all these constraints. It took more than 5 years of R&D for the CMS GEM project to show that it was feasible. [48, 72, 82]. This effort led to the installation of 5 *slice test* super-chambers in early 2017, followed by the beginning of the mass production in October 2017.

2.2.2 The GE1/1 detector

Single layer chambers are the elementary unit of the GE1/1 detector. Their trapezoidal geometry with an angular opening of 10.15°, allowing overlap between neighbors, provides maximal coverage. Due to the structure of the CMS nose, however, two alternating types of chambers are required: the *short* type covers the region $1.61 < |\eta| < 2.18$ whereas the *long* type covers the region $1.55 < |\eta| < 2.18$. This results in chambers with the following dimensions: the small side of the trapeze is always 28.5 cm wide while the larger side is 48.4 cm (51.0 cm) wide and the length is 113.5 cm (128.3 cm) for the short (long) type.

The chambers are assembled into a double-layer structure, the so-called *super-chambers*, which are installed in a ring fashion into each of CMS noses. This organization is depicted in the 2D drawing in Figure 2.8. In total, 144 single chambers, 72 shorts and 72 longs, are required to complete the GE1/1 project.

To refer to individual chambers, the following naming scheme is adopted: **GEe1/1/nn Lyl**. In this notation, *e* indicates the CMS endcap, `-` or `+`, corresponding to the sign of the CMS Z coordinate; *nn* denotes the slot number, following the CMS phi azimuthal coordinate; and *l* specifies the layer number, `1` or `2`, with layer 1 being the closest to the interaction point. Chambers installed in odd-numbered slots are short, while those in even-numbered slots are long. The first and second `1`'s in the scheme respectively indicates the CMS endcap disk and ring.

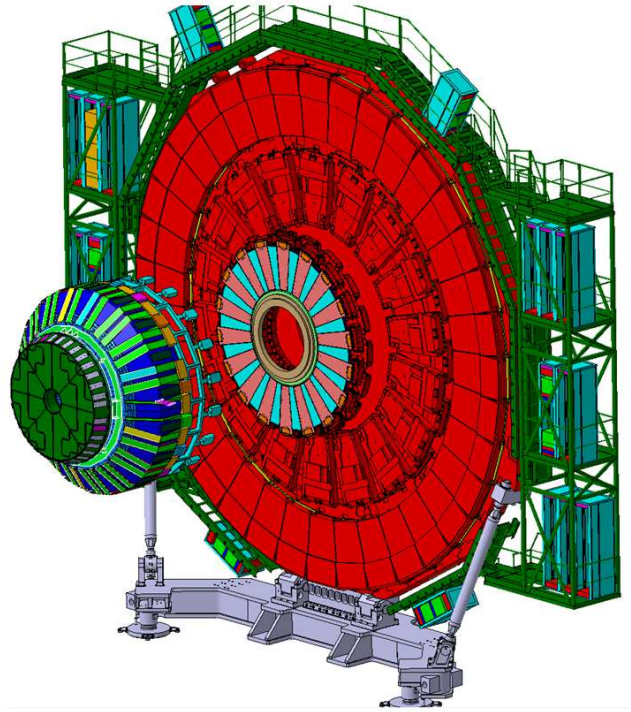


Figure 2.8: The CMS endcap first disk and nose. The GE1/1 super-chambers are highlighted in light red and cyan. [1]

Figure 2.9 presents an exploded view of one GE1/1 single chamber. The Triple-GEM structure can be recognized at the bottom: the drift board in light red, the GEM foils in cyan, and the readout board in blue. The readout electronics comprises, from bottom to top, the GEM Electronic Board (GEB) in red, the VFAT3 chips in green, and the OptoHybrid (OH) in light brown. Those elements are detailed in the next section. The detector is completed by the cooling plate, in light gray, providing adequate cooling to the readout electronics, and the chimney, in light blue, acting both as a mechanical protection layer and a Faraday cage.

The readout board (in blue) is divided into 24 sectors: 3 along ϕ and 8 along η. Each sector is further divided into 128 strips along its width for a total of 3,072 strips per chamber and a grand total of 442,368 readout channels for the GE1/1 detector. The chosen division ensures the required angular resolution of 300 µrad.

A specificity of the large-size GEM detectors is their susceptibility to discharges. Such events can damage the GEM foils themselves - when occurring inside the GEM holes - or the readout electronics - when *propagating* from one foil to another. To mitigate the risks, the top of each foil is segmented into slices perpendicularly to the $|\eta|$ direction while the bottom is made of a single plane. The segments are designed such that their area is about 100 cm^3, a compromise between the effectiveness of the solution and the

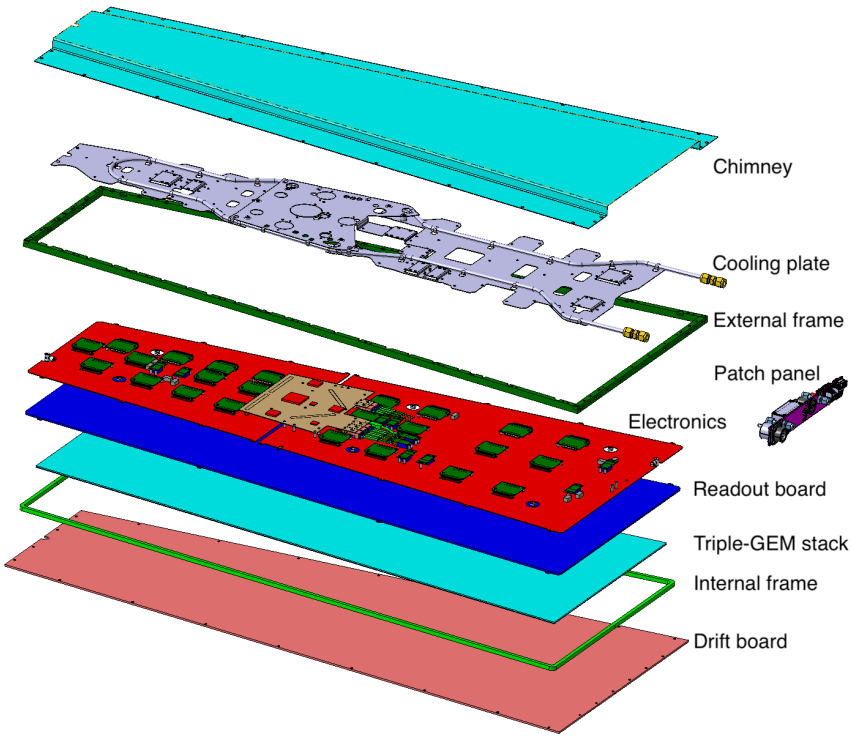


Figure 2.9: Exploded view of a GE1/1 chamber. [1]

added dead zone. Most importantly, the segments are each powered through a dedicated 10 MΩ protection resistor. This provides three benefits:

1. The amount of energy stored in each segment and the current flowing through a spark are reduced.

2. During a discharge, the current flowing through the protection resistor reduces the voltage applied to the foil, quenching the discharge.

3. In case a short circuit is generated across a foil, the protection resistance limits the current drawn and effectively limits the dead area to the damaged segment.

This powering scheme is presented in Figure 2.10. The HV filter reduces noise injected in the readout electronics by the HV power system.

2.2.3 The front-end readout electronics

The GE1/1 on-chamber, or front-end, readout electronics is mounted directly on top of the chamber and thus constitutes the first stage of signal processing. Due to its location within the experimental cavern, it is continuously exposed to radiation. As such, its

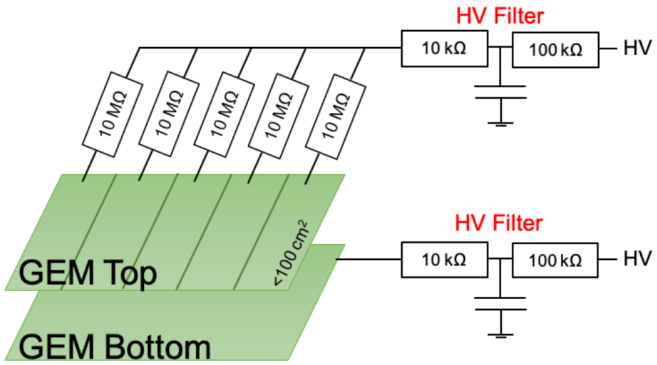


Figure 2.10: High voltage scheme of one single segmented GE1/1 foil.

components must be either intrinsically radiation resistant (RadHard) or protected by appropriate mitigation strategies.

An overview of the GE1/1 readout electronics is shown in Figure 2.11. Each GE1/1 chamber is equipped with 24 VFAT3 custom readout ASICs, 1 OptoHybrid concentrator board, and 1 GEM Electronics Board (GEB). The VFAT3 features 128 input channels with binary readout, each composed of a charge-sensitive preamplifier, a shaping network, and a constant-fraction discriminator [15]. The OptoHybrid board serves as the communication hub between the counting room and the VFAT3s via optical links. The GEB, a 2-piece PCB sharing the dimensions of the GE1/1 detector, serves as the signals and power carrier between the OptoHybrid board and the VFAT3s.

VFAT3

The VFAT3 ASIC is the third generation of the VFAT (Very Forward ATLAS and TOTEM) chip family [14, 108]. It constitutes a complete redesign to satisfy the specific needs of the CMS GEM project. Figure 2.12 shows a picture of a VFAT3 ASIC mounted on an external PCB and ready for integration into a GE1/1 detector. In the case of GE1/1, it is colloquially referred to as *hybrid*.

The block diagram of the VFAT3 is provided in Figure 2.13. The design can be divided into two domains: the analog side, responsible for signal amplification and digitization; and the digital side, responsible for communication, control, monitoring, and data buffering.

On the analog side, each of the 128 channels includes a preamplifier, a shaper, and a Constant Fraction Discriminator (CFD). The charge-sensitive preamplifier can be configured in *low*, *medium*, and *high* gain modes using programmable feedback resistors and capacitors, yielding amplification factors of 1, 2, or 6 mV/fC. The shaper peaking time is also configurable, with effective values of 15, 25, 36, or 45 ns, adjusted through

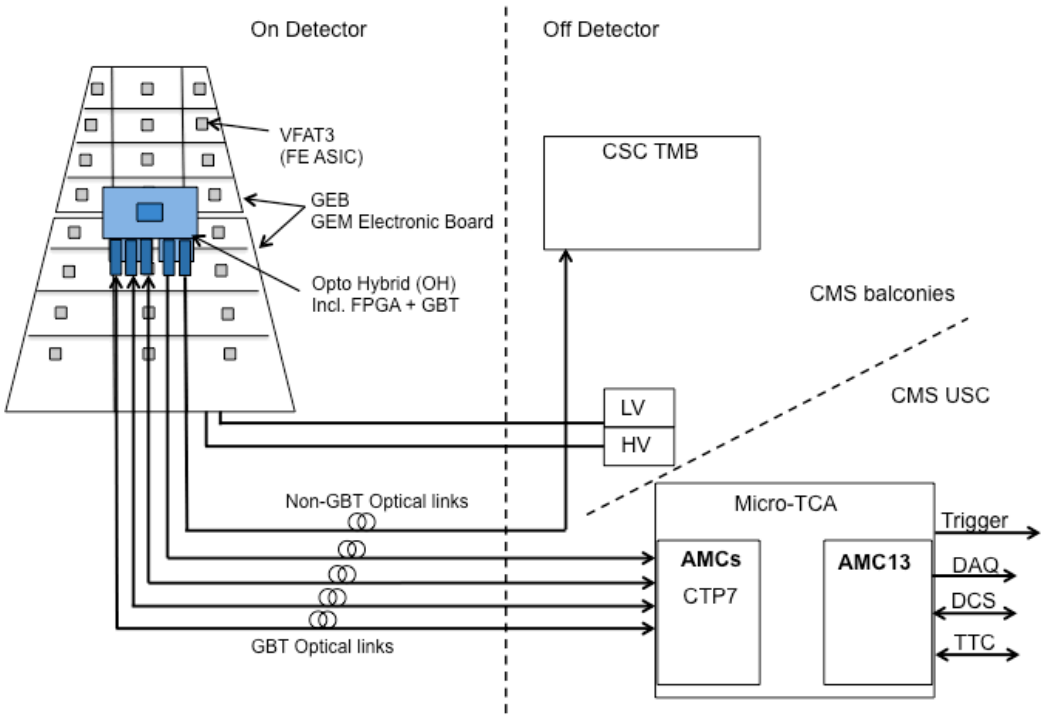


Figure 2.11: Overview of the GE1/1 readout electronic. [75]

programmable capacitors. [14]. A differential copy of the shaped signal is fed to CFD, which digitizes it.

The VFAT CFD block comprises two comparators [3]. The *arming comparator* (ARM) determines whether the signal amplitude exceeds the threshold. However, signals of identical shape but varying amplitudes will cross the threshold at different times. This phenomenon, known as time-walk, degrades the final time resolution of the detector. The *zero-crossing comparator* (ZCC) avoids this effect by accurately timestamping the time of arrival of the signal.

The resulting binary signal is then synchronized with the LHC clock sent in two paths: the trigger path and the tracking data path.

The trigger path features a low and fixed latency transmission optimized for the Level-1 Trigger. The VFAT3 trigger primitives, called *S-bits*, are a coarser resolution version of the full channel data. They consist of the logical OR of two adjacent channels, yielding 64 bits of data per bunch crossing. Their transmission is achieved through 8 differential pairs running at 320 Mbps using Time-Division Multiplexing (TDM).

Figure 2.12: Picture of a GE1/1 VFAT3 hybrid.

The tracking data path transmits the full-resolution hit data for triggered events with a variable latency. The digitized signals from all 128 are continuously stored in a circular buffer (SRAM1). Its depth of 1024 entries, corresponding to 12.8 µs at 40 MHz, is compatible with the CMS Phase-II DAQ requirements. Upon trigger, the relevant data is fetched in the SRAM1 and copied to the SRAM2 buffer along with additional metadata: Bunch Crossing Number (BCN) and the Event Counter (EC). The BCN is increased for each LHC clock cycle, while the EC is a counter increased for every L1A. The content of the SRAM2 is then progressively flushed through the VFAT3 Comm-Port.

The VFAT3 Comm-Port is the main communication channel with the VFAT [81]. It enables the LHC clock recovery; features a bi-directional slow-control channel to the VFAT; accepts fast control commands to reset or trigger (L1A) the chip; and streams the tracking data to external systems.

The chip also includes a Calibration, Bias, and Monitoring (CBM) unit, which houses a set of Digital-to-Analog Converters (DACs) and Analog-to-Digital Converters (ADCs) to, respectively, bias the analog circuitry and monitor the applied currents and voltages. Moreover, the CBM embeds a calibration module capable of sending voltage or current calibration pulses of varying amplitude for channel characterization and health monitoring.

Channel damages During the GEM slice test in 2017-2019, unexpected front-end channel losses were experienced as shown in Figure 2.14. This issue was eventually traced to discharges propagating to the anode plane and to the VFAT input channels. At the time, the GEM technology was considered free from propagating discharges,

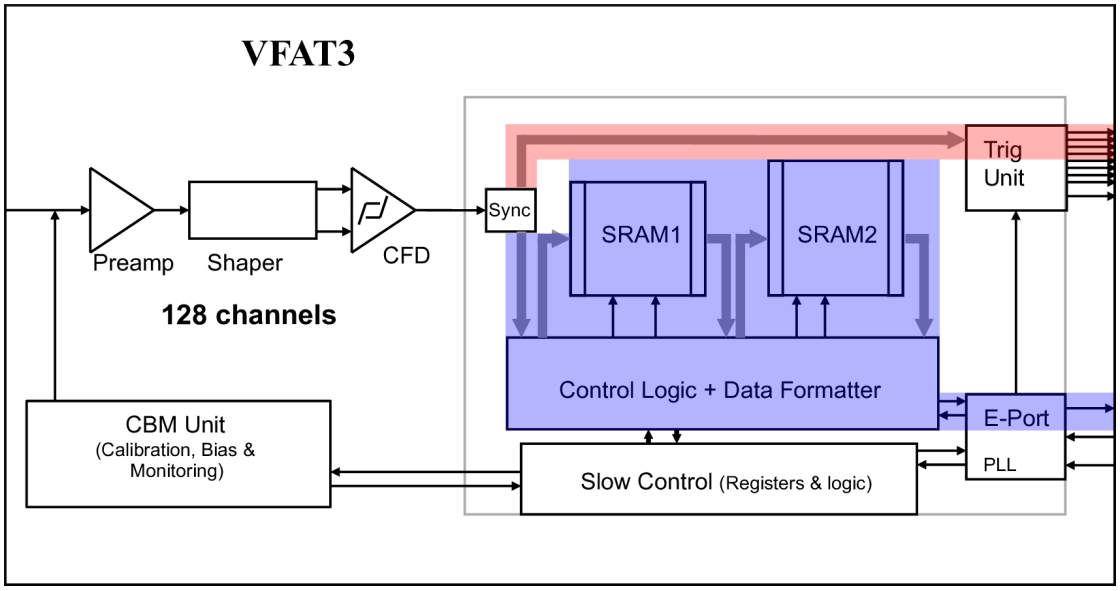


Figure 2.13: Block diagram of the VFAT3. The trigger path is highlighted is red and the tracking data path is in blue. [14]

although discharges internal to one GEM foil were well-known. As such, no protection mechanism was implemented on the ASIC inputs.

The channel loss rate, with estimations for the GE1/1 system, is given in Equation 2.3.

$$\begin{aligned} \text{channel loss rate} = \text{background rate}(&\sim 3.15 \times 10^{3}\,\text{VFAT}^{-1}) \\ &\times \text{discharge prob.}(\sim 1.24 \times 10^{-9}) \\ &\times \text{propagation prob.}(\sim 50\%) \\ &\times \text{damage prob.}(\sim 3\%) \end{aligned} \tag{2.3}$$

Two mitigation strategies have been implemented in GE1/1. The input channel protection has been enhanced by the utilization of 470 Ohm protection resistors in series (on the VFAT hybrid PCB) to lower the damage probability; and the high-voltage filter resistance has been increased to reduce the discharge propagation probability [97]. Considering the advanced stage of construction of the project, no further mitigation strategies could be implemented.

OptoHybrid

The OptoHybrid (OH) board, installed at the center of the GE1/1 detector, is a custom processing board designed to interface the VFATs with the back-end electronics located in the counting room. Figure 2.15 shows an OH picture with its different components labeled.

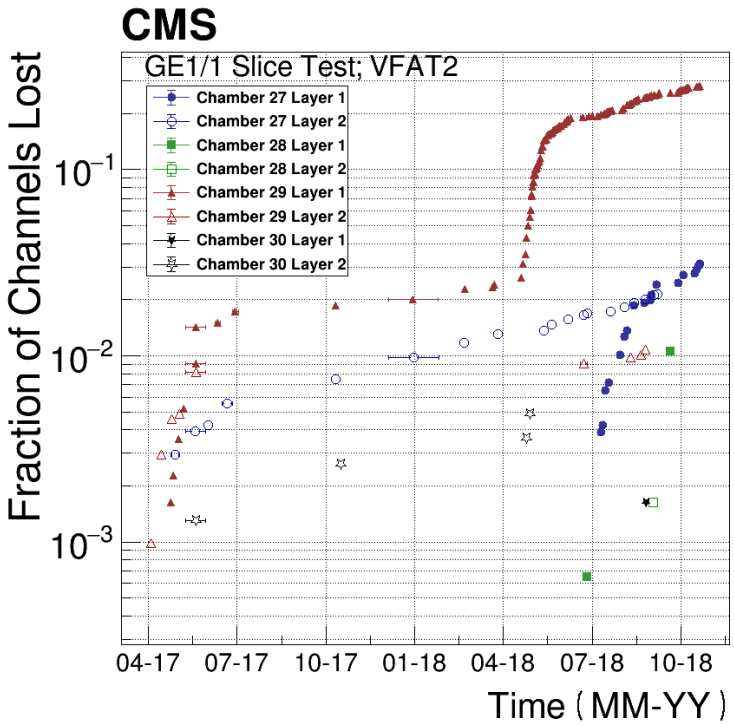


Figure 2.14: Cumulative channel losses for the GE1/1 slice test chambers before the implementation of any mitigation measure against discharges [89, 37].

The OptoHybrid tasks are twofold. First, it provides a direct communication path between the back-end electronics and the 24 VFATs via 3 GBTx ASIC and their associated VTRx (Versatile Transmitter-Receiver) transceivers [84, 106]. These components, developed as part of the CERN's GigaBit Transceiver (GBT) project, aim to provide reliable multi-gigabit (4.8 Gbps) data transmission through optical fibers in high-radiation environments. Second, an embedded Xilinx Virtex-6 FPGA processes and transmits the trigger information coming from the VFATs to the CMS Level-1 trigger via 2 VTTx (Versatile Transmitter-Transmitter) optical transceivers. The VTTx is a variant of the VTRx that contains only two transmitters. Additionally, a Slow-Control ASIC (SCA) provides comprehensive monitoring in a radiation-hardened chip [34].

The GBTx-VFAT connections are unbalanced: the first GBTx connects to 6 VFATs, while the other two connect to 9 VFATs each. This configuration is due to the first GBTx also handling the slow-control communication with the FPGA, as well as its configuration via the PROM-less method. Indeed, as no commercially available Programmable Read-Only Memories (PROM) withstanding the dose of radiation expected during the GE1/1 lifetime were found, the FPGA firmware is loaded remotely from the back-end board.

The OptoHybrid FPGA firmware plays an essential role in the integration of GE1/1

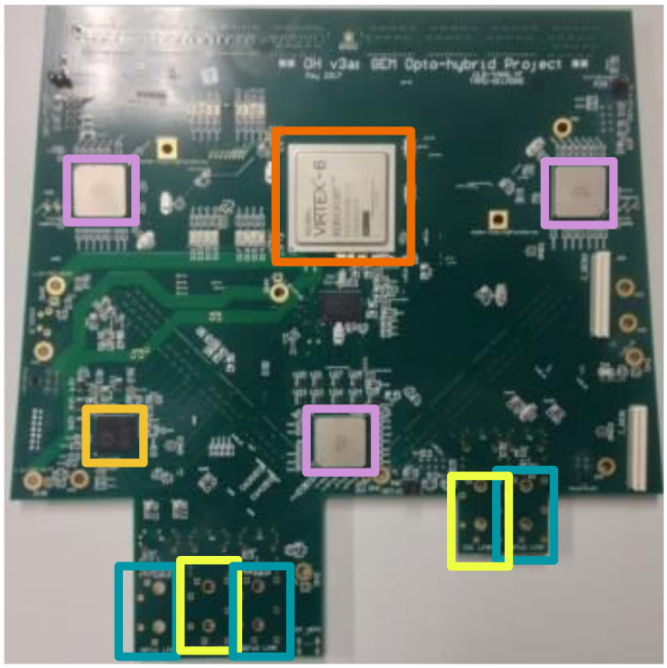


Figure 2.15: Picture of a GE1/1 OptoHybrid. The Virtex-6 FPGA is framed in orange; the 3 GBTx ASICs in pink; and the SCA in light orange. The VTRx and VTTx are situated underneath the PCB at the locations designated by the teal and yellow rectangles, respectively.

into the Level-1 Trigger. Sending the raw trigger data generated by the VFAT without any processing is infeasible as it would require a bandwidth of 61.44 Gbps per detector. At its core, the firmware implements a low-latency, fully pipelined compression algorithm. It receives and deserializes the 8 TDM streams coming from each VFAT; applies corrections to ensure accurate data mapping, both in space and in time; and identifies up to 8 S-bit clusters per bunch crossing among the 1536 bits sent by the VFATs in 3.25 BX. Only the cluster positions and sizes are encoded, achieving the required bandwidth reduction. The trigger primitives are then formatted and sent to the OTMB over two 8b10b-encoded 3.2 Gbps optical links and to the EMTF over two 8b10b-encoded 4 Gbps optical links transiting via the GEM back-end. Additionally, the firmware provides extensive control and monitoring features to best configure the system.

FEAST

While using a single lower voltage power input per detector simplifies the cabling and installation, the on-chamber electronics requires multiple voltages to operate. Therefore, voltage conversion must be performed on the detector itself.

The solution adopted for the GE1/1 project is the FEASTMP_CLP module, built around the FEAST2 ASIC developed at CERN [56]. The FEASTs are synchronous step-down buck DC-DC converters designed to operate in high radiation and magnetic field

environments. Accepting an input voltage comprised between 5 and 12 V, each FEAST is capable of delivering up to 4 A and 10 W.

Each GE1/1 detector uses 9 FEASTs: 4 for the VFATs and 5 for the OptoHybrid board. Their purpose is described in Table 2.2.

Table 2.2: Quantity, type, and purpose of the GE1/1 FEASTs DC-DC converters.

Quantity	Voltage	Usage
1	1 V	FPGA core
1	1 V	FPGA multi-gigabit transceivers
1	1.2 V	FPGA multi-gigabit transceivers
1	1.55 V	GBTx, SCA, VTRx
1	2.58 V	VTRx
4	$\sim$ 1.3 V	VFATs, 4 power domains with 6 VFATs each

GEM electronics board

The GEM Electronics Board (GEB) is a passive, eight-layer Printed Circuit Board (PCB) mounted on top of the readout board, outside of the chamber's gaseous volume. It is designed to interconnect the various components of the readout electronics. The GEB serves three primary functions:

- distributing power to both the VFATs and OptoHybrid;
- connecting the VFATs to the OptoHybrid with sufficient signal integrity;
- shielding the detector from the high-frequency noise generated by the front-end electronics.

Therefore, the PCB must employ high-density layout techniques, routing approximately 500 differential signal pairs that run at 320 Mbps, in addition to the necessary power and ground lines. It must also withstand significant mechanical stress, as it hosts a large number of rigid connectors. Due to these requirements and the manufacturing limitations, the final revision consists of two separate PCB sections, as illustrated in Figure 2.16. The 4 central black SAMTEC connectors interface with the OptoHybrid, while the 24 white 100-pin Panasonic connectors, distributed across the boards, are used to connect to the VFATs. Additionally, the 9 black 12-pin connectors are used to connect to the DC-DC converters.

2.2.4 The back-end readout electronics

Represented on the right side of Figure 2.11, the GE1/1 back-end electronics is located in the CMS service cavern. It provides the interfaces between the on-chamber electronics

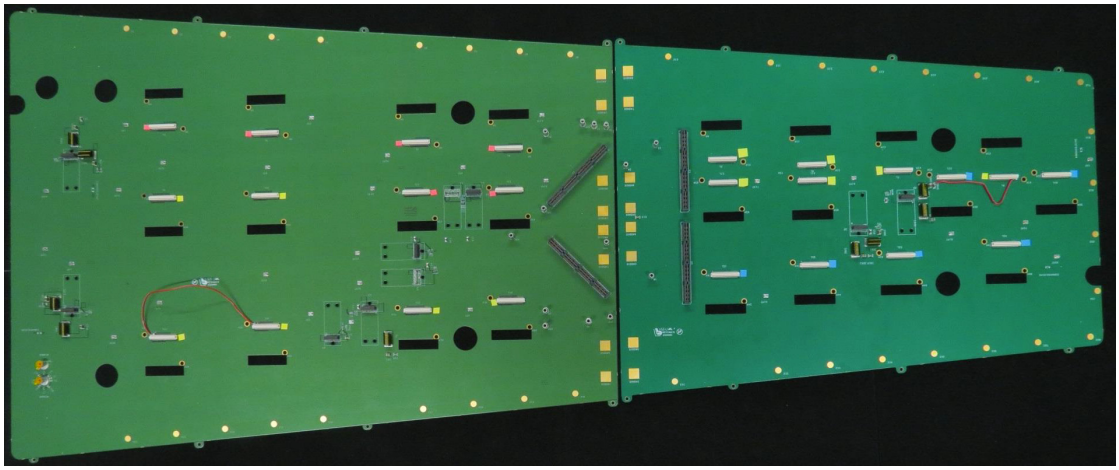


Figure 2.16: Picture of the two parts of the final GEB revision (v3) [75].

and the external control, monitoring, trigger, and readout systems. Like the other Phase-I CMS upgrades, the back-end electronics is based on the µTCA standard.

Originally introduced in 2006 for the telecommunication industry, the µTCA is a standard has since been widely adopted by the high-energy community. It provides a unified framework for developing high-available, modular, and scalable systems. Among extensive aspects, the standard defines the communication backplane, the management interface, and the mechanical form factors. The MCH (MicroTCA Carrier Hub) provides low-level hardware management capabilities through the IPMI (Intelligent Platform Management Interface) as well as Ethernet switching functionality. Custom functionalities are implemented in specialized Advanced Mezzanine Cards (AMCs).

Figure 2.17 shows a picture of the GE1/1 back-end electronics. The top (bottom) µTCA crate, or shelf, serves the negative (positive) GE1/1 endcap. Each crate hosts 6 CTP7 cards and 1 AMC13 module, which together constitute a GEM FED.

CTP7

The Calorimeter Trigger Processor 7 board (CTP7), originally designed for the CMS calorimeter Level-1 Trigger, is used as the GE1/1 back-end board thanks to its high flexibility [98]. The board features a fairly standard design centered around a Xilinx Virtx-7 FPGA. A total of 48 TX and 67 RX optical links has been certified to operate at 10 Gbps: three hot-swappable Avago CXP modules each provide 12 TX and 12 RX links, while a set of Avago MiniPODs installed on the board PCB provides the additional 12 TX and 31 RX links. Additional Multi-Gigabit Transceivers (MGTs) are connected to the backplane to enable direct communication with other boards in the same µTCA crate.

The distinctive feature of the CTP7 is the integrated Xilinx Zynq-700 SoC (System-on-Chip), which runs an embedded Linux Operating System (OS) on a dual-core ARM

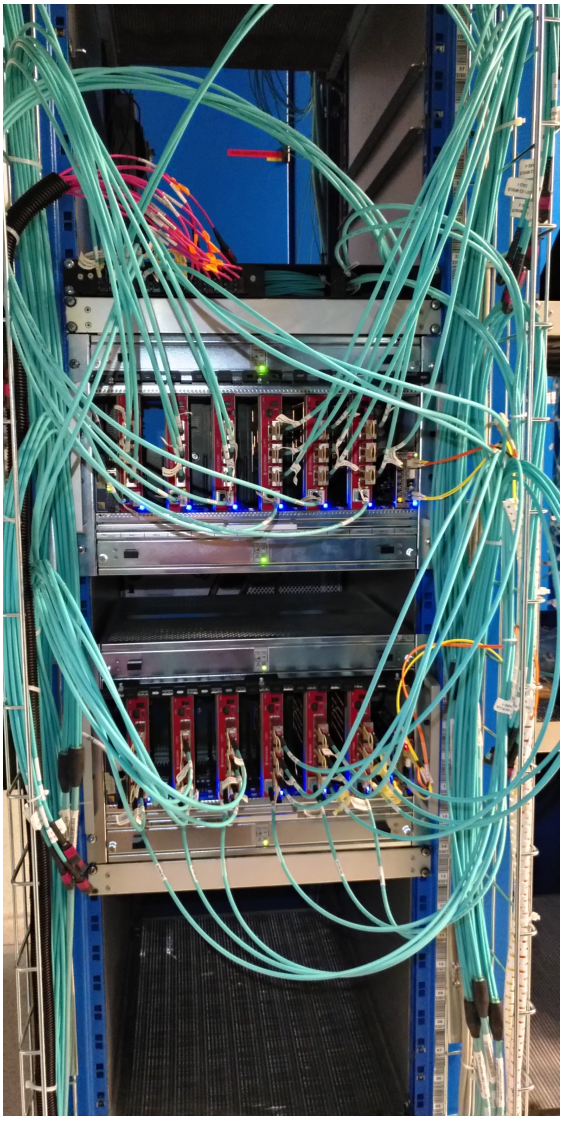


Figure 2.17: Picture of the GE1/1 back-end electronics in the CMS service cavern. The top (bottom) µTCA crate serves the negative (positive) GE1/1 endcap. Visible in each crate are 6 CTP7, 1 AMC13, and 1 MCH.

Cortex-A9 CPU. This SoC also includes a small FPGA – compared to the Virtex-7 – used to implement ancillary functions, primarily the communication with the Virtex-7 main FPGA. The CTP7 was the first (and still is the only) AMC board in CMS to integrate a CPU.

The following desirable features were developed on the Zynq-7000 system:

1. Remote programming and debugging of the Virtex-7 FPGA.
2. Integrated eye-scan link engine for optical (and copper) link quality monitoring.

3. Fast register access library for both the Zynq-7000 and the Virtex-7 ($< 10\,\mu s$).
4. Remote Procedure Call (RPC) service to offload remote processing tasks.
5. Board monitoring features (e.g. temperatures, optical power levels).

The onboard CPU has proven to be a crucial asset for the DAQ system control and monitoring in the GEM online software, as discussed in Chapter 3. Features 3, 4, and 5 have been reimplemented for better integration into the software stack, while features 1 and 2 are used as-is.

The Virtex-7 firmware implements the core processing functionalities of the board. It supports the following features:

- Forwarding of the TTC clock and commands to the on-chamber electronics
- Provision of slow-control interface to the GBTx, SCA, OptoHybrid FPGA, and VFAT3, and Virtex-7 itself
- Concentration of the trigger primitives coming from multiple GE1/1 detector into fewer, higher-speed links to the EMTF
- Readout and event building for all connected detectors

The available resources in each CTP7 (FPGA capacity, optical link quantity, and CPU processing power) allow one board to support up to 12 GE1/1 OptoHybrids, and thus GE1/1 detectors. In total, 12 CTP7s, distributed across two µTCA crates, are required to drive the entire GE1/1 station.

AMC13

The AMC13, originally designed for the needs of the CMS HCAL Phase-I upgrade, serves as the interface between the sub-detector DAQ hardware and the central DAQ (cDAQ) systems [67]. To fulfill this role, the AMC13 performs the following functions:

- It receives the TTC signal from the TCDS and broadcasts it, along with the recovered LHC clock, to all AMCs;
- It receives the TTS signals from all AMCs, aggregates them, and transmits the summary status to the TCDS;
- It receives the DAQ streams at 5 Gbps from all AMCs, builds events, and ships them to the cDAQ FEROLs via 10 Gbps S-Link Express links.

These functionalities are enabled by inserting the AMC13 within the second, redundant MCH slot of the µTCA crate, thereby providing direct communication with all AMC slots.

In addition to its role in CMS, the AMC13 offers extensive features to run, trigger, and debug the system in standalone, local mode.

Figure 2.19 shows a production AMC13 board.

Figure 2.18: Picture of the CTP7 board.

2.2.5 The detector services

To operate the GE1/1 detector, a set of essential services must be provided. This section briefly describes the low-voltage (LV) and high-voltage (LV) systems, along with their associated control systems, which play an important role in the rest of this dissertation.

The **low-voltage** power distribution follows a two-stage conversion chain based on the CAEN EASY3000 system, specifically designed to operate in environments with magnetic fields or radiation concerns [32]. First, CAEN A3486 MAO modules convert the 3-phase 220V AC main power available in the CMS experimental cavern into 48V DC [31]. Next, the 48V DC voltage is regulated down to 8 V by CAEN A3016HP boards housed in EASY3000 crates [30].

In total, the LV system comprises 4 dedicated racks, each containing 2 MAO units, 2 EASY3000 crates, and 6 LV boards. Each low-voltage channel is mapped one-to-one to a single detector, enabling fully independent operation and monitoring.

System control is assured by CAEN A1676A Branch Controllers located in CAEN SY4527 mainframes installed in the CMS service cavern, isolated from the radiations [29, 33]. Each branch controller can manage up to 6 EASY3000 crates and their associated MAOs via a single 50-pin flat cable that embeds 6 independent CAN (Controller Area Network) buses.

The **high-voltage** power is provided by CAEN A1515TG floating HV boards designed for the operation of Triple-GEM detectors [28]. Each board supports two fully

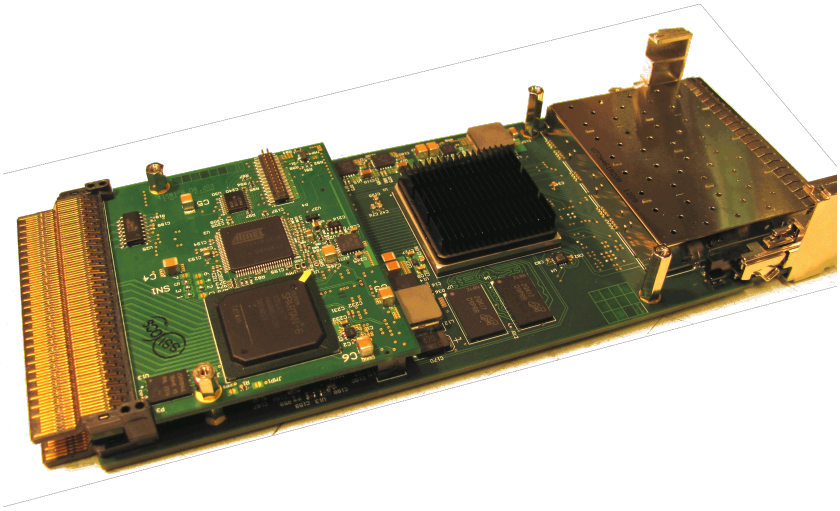


Figure 2.19: Picture of the AMC13 board.

independent HV complex channels, each capable of powering all electrodes of a detector.

Installed in the CMS services cavern, a total of 36 boards are hosted in 4 SY4527 mainframes, each serving one quadrant of the GE1/1 sub-system. Due to cost reasons, each HV complex channel is typically shared between 2 detectors with similar working points. This sharing is implemented using Y-splitting cables connecting the HV board to HV patch panels, from where detectors are connected with point-to-point connections.

If needed – e.g. to maximize the detection efficiency or improve stability – additional HV boards can be added in the mainframes, and the mapping between detectors and HV boards adjusted accordingly.

The **DCS (Detector Control System)** is conceptually similar to the DAQ online software: where the online software controls and monitors the data-acquisition process, the DCS is responsible for the same tasks regarding the detector services and associated hardware components.

Implemented using the solution adopted by all LHC experiments, the WinCC OA SCADA (Supervisory Control and Data Acquisition) framework [55], the DCS interfaces directly with the hardware components (such as the LV and HV boards) and with external systems (such as the GCS (Gas Control System) and the DAQ).

Operational parameters, such as the HV currents, are continuously monitored and archived in a conditions database for long-term tracking. Upon detection of abnormal conditions or events, the DCS can emit alarms, notify operators, or trigger automatic safety actions.

As for the DAQ system, the DCS can be operated in either local or central mode. In local mode, the GEM users and experts have full control over the system. In central mode, only commands originating from the central CMS DCS or from the shift crew are allowed. This centralized mode of operation enables all CMS sub-systems to be

operated by a reduced number of trained shifters. Regardless of the operating mode, safety actions are handled automatically.

The **DSS (Detector Safety System)** is designed to preserve the integrity of the detectors in case of critical failures or hazardous conditions. Due to high-reliability requirements, the DSS is implemented exclusively using industrial-grade PLCs (Programmable Logic Controllers), avoiding the risk of software-induced failures. For example, the electronics racks are forcefully switched off if smoke is detected. Or, in the GEM sub-system, the high-voltage racks are "killed" if a fast CMS magnet discharge occurs. However, from an operational standpoint, it would arguably be beneficial to replace this action with an interlock of the high-voltage channels. This would offer the same level of safety (as implemented fully in hardware) but without losing control and monitoring over the system for post-mortem analysis and recovery.

2.3 The GEM Phase-II upgrades, ME0 & GE2/1

As the GE1/1 station is built, installed, and operated, the GEM collaboration prepares the development, production, installation, and operation of two new GEM Phase-II stations: GE2/1 and ME0. These new stations, like GE1/1, aim to maintain and improve the tracking and trigger capabilities in the forward region of the CMS muon spectrometer, which will face increasingly challenging conditions during the LHC Phase-II era.

As of early 2025, the ME0 station has entered its final stage of production. Through extensive R&D and test beam campaigns, the designs of all components have been validated. Production is now underway, and complete detectors are under assembly before a rigorous set of quality control procedures. Within 2 years, the ME0 production must be completed to be ready to meet the timeline of its installation in the CMS nose. Located closest to the beamline and interaction point, ME0 will face the harshest conditions experienced by any CMS GEM detectors and thus provide a specifically optimized design.

The GE2/1 station production, by contrast, is currently on hold until an indeterminate date. Initially scheduled for installation in the 2023 and 2024 YETS, the project suffered significant delays due to defects in the manufacturing process of the readout PCBs. Given the narrow time window for the ME0 installation, the project's resources have been redirected accordingly. Nevertheless, the first five GE2/1 production detectors have already been installed in CMS and are operational. As with GE1/1, each GE2/1 detector complements a companion CSC detector.

Moreover, the back-end electronics must be entirely replaced to comply with new CMS requirements and design principles.

This section summarizes the highlights of the GEM Phase-II upgrades.

2.3.1 GE2/1 highlights

The GE2/1 station, installed in front of its companion ME2/1 station, covers a pseudorapidity range of $1.62 < |\eta| < 2.43$. Each of the two GE2/1 endcaps consists of 18 super-chambers, each covering approximately 20° in phi. A super-chamber comprises two independent, staggered layers – named FRONT and BACK chambers – each containing different module types to mitigate dead zones caused by geometry acceptance effects. Each layer is made out of four individual trapezoidal detector modules, labeled either M1-M4 or M5-M8 depending on the layer. This configuration for a BACK chamber is shown in Figure 2.20. In total, the GE2/1 station comprises 288 modules.

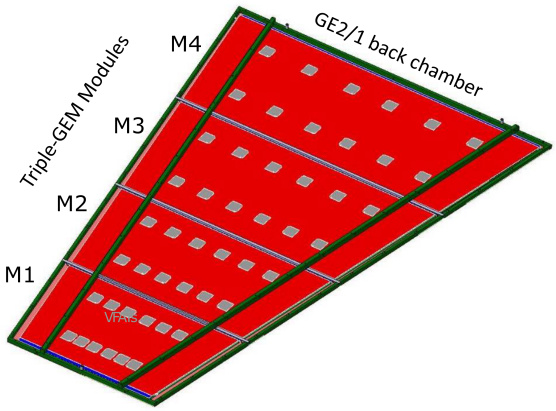


Figure 2.20: Layout of a GE2/1 BACK detector and its division in 4 modules, M1 to M4 [45].

The GE2/1 on-chamber electronics is largely similar to the GE1/1 one. Each module is equipped with 1 OptoHybrid, 12 VFAT3s, 1 GEB, and 5 FEASTs, adapted to the half angular coverage relative to a GE1/1 chamber. The most notable electronics differences are:

- The OptoHybrid FPGA forwards trigger primitives to the GEM back-end board via eLinks rather than optical links thanks to unused bandwidth in the GBT links.
- The analog and digital VFAT power domains are decoupled by 2 independent FEASTs, reducing the electronics noise levels.
- The VFAT3 hybrid board has evolved into a so-called *Plugin card* featuring a flexible PCB between the readout board connection and the main VFAT PCB. This reduces the mechanical stress created by rigid parts and results in much better assembly reliability.

Additionally, the VFAT3 had been equipped with a new input protection mechanism [97]. Allowed by the complete redesign of the front-end PCB, this protection scheme implements an AC decoupling capacitor and a drain resistor. The decoupling

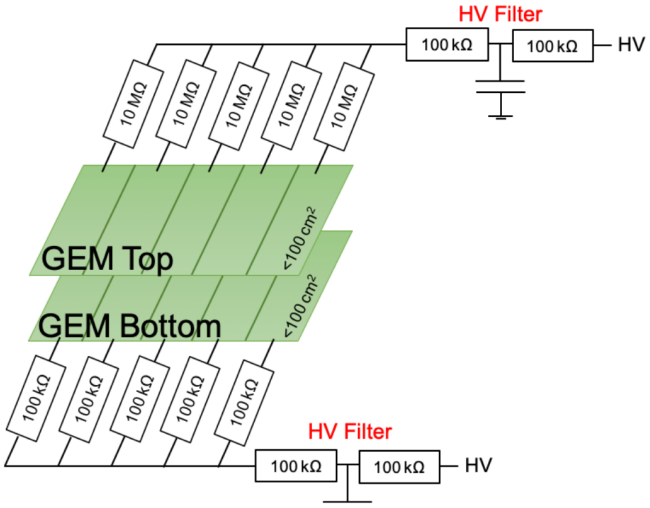


Figure 2.21: High voltage scheme of one double segmented GE2/1 foil.

capacitor reduces the damage probability as it prevents the ASIC input channel from absorbing the full energy of the discharge. The drain resistor reduces the discharge propagation probably to the anode by quenching the discharges precursor current (through the increase of the readout strip potential and the suppression of the induction field).

However, as it will be demonstrated in Section 5.3, and despite promising results on small-scale Triple-GEM detectors, this protection scheme led to severe channel losses. Therefore, the definition protection mechanism chosen across all three GEM stations remains the use of protection resistors.

On the detector side, the key differences are:

- The addition of central "pillars" between the drift and readout PCB, passing through the GEM foils. They prevent mechanical deformations and improve the response uniformity over the detector.

- The implementation of a double foil segmentation, as shown in Figure 2.21 (in contrast to the GE1/1 single-segmentation). This limits the energy stored in the foil capacitance, reducing both the probability and the energy of propagating discharges. To limit the cross-talk signals (see Section 5.5), the bottom foil remains single-segmented.

Finally, a single BACK-layer GE2/1 demonstrator was installed in CMS in November 2022. It served as the final testing and validation platform of the GE2/1 station for multiple installation and operational aspects. Fully integrated into the DAQ and DCS systems, it enabled the evaluation of the updated discharge protection scheme (both foil design and VFAT3 protection circuit), noise levels (for grounding validation), and detection efficiency.

Additional production GE2/1 chambers have been installed in subsequent years, bringing the total number of operational GE2/1 chambers in CMS to six (including the demonstrator).

2.3.2 ME0 highlights

The ME0 station, covering a pseudorapidity range of $2.03 < |\eta| < 2.8$, is the GEM station located closest to the beamline and to the interaction point, where background rates can reach up to $150\,\mathrm{kHz\,cm^{-2}}$. Each of the two ME0 endcaps consists of 18 stacks, each covering approximately 20° in phi. A stack comprises 6 layers, called modules. To avoid dead areas, the modules are flipped between adjacent stacks, resulting in a staggered installation. In total, the ME0 station comprises 216 modules. This organization is depicted in Figure 2.22

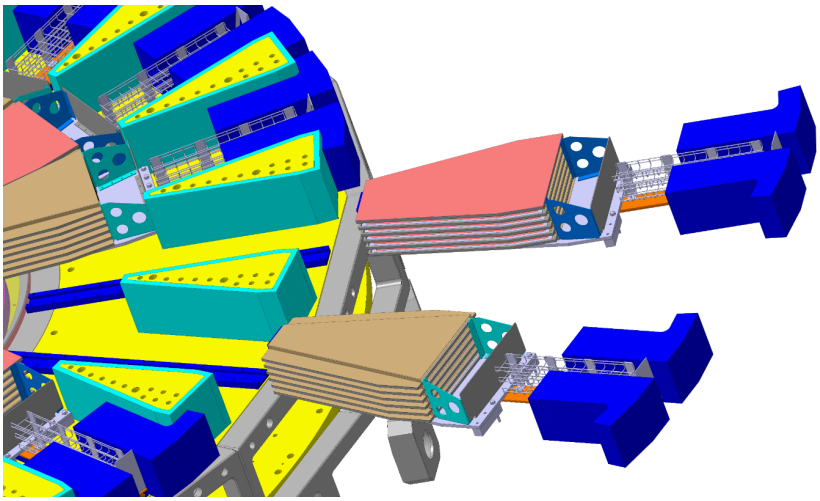
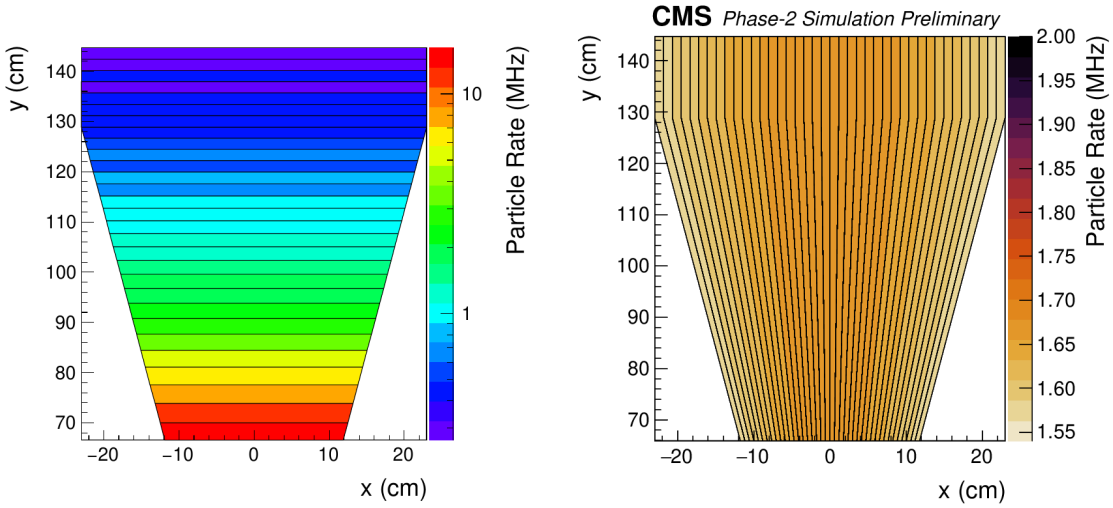


Figure 2.22: Drawing of design, installation, and placement of ME0 stacks in the CMS endcap. Note the alternating orientation of modules in the stacks [45].

Despite some similarities with GE1/1, the ME0 on-chamber electronics introduces significant differences. Each chamber is equipped with 1 GEB and 24 VFAT3s, serving the same roles as in GE1/1. However, DC-DC conversion is handled by bPOLs, devices similar to the FEASTs but designed to withstand higher radiation doses.

Due to the high radiation levels, the OptoHybrid cannot include a processing FPGA; as a result, no S-bit compression is performed on-chamber. Communication between the VFAT3s and the back-end is assured by 8 LpGBT (Low-power GBT) ASICs and their 4 associated VTRx+ (Versatile Transmitter-Receiver Plus) transceivers, which are mounted on 4 independent ASIAGO boards [85, 107]. This distributed system is conventionally still referred to as the OptoHybrid. The LpGBT is the next-generation version of the GBTx, developed within the same CERN project. It features a 2.56 Gbps down-



(a) Horizontal segmentation

(b) Vertical segmentation

Figure 2.23: Simulation of the background particle rates per GEM foil sector in ME0 segmentation [12].

link and a 10.24 Gbps uplink. The VTRx+ is the evolution of the VTRx, tailored for use with the LpGBT, and provides 1 RX and 4 TX links.

Trigger primitives in ME0 are also significantly more advanced. Because there is no companion CSC chamber, and thanks to the presence of six detection layers, online segments can be reconstructed directly in the back-end firmware and sent to the Level-1 Trigger.

On the chamber side, the most significant change concerns the GEM foil segmentation. In the highly non-uniform background rate region of ME0, the horizontal segmentation used in the GE1/1 and GE2/1 is not suitable. High currents flowing through the segment protection resistors would lead to gain drops of up to 40%. While the resistance values could theoretically be adjusted to match the rate profile, any inaccuracies in simulation could render the ME0 performances suboptimal. To address this, the foil segmentation for ME0 has been redesigned. It features a vertical (azimuthal) segmentation scheme, thereby eliminating all dependence on the pseudorapidity for the total rate per HV segment. Figure 2.23 compares the total particle rate per segment for the two configurations.

2.3.3 DAQ highlights

As foreseen in Section 1.2.6, the back-end trigger and readout electronics must be upgraded to support the latency and bandwidth requirements imposed by the CMS Phase-II DAQ system. For coherence and standardization, CMS mandates that all sub-detector back-end hardware follow the ATCA standard, an evolution of the µTCA standard[2]. Due to limitations in optical link density and thermal dissipation, the µTCA system was no longer suitable.

The GEM back-end board for Phase-II will be the **X2O**, a modular platform composed of:

- a power module – integrating all control and management features via a Kria SoC;
- a processing module – embedding a powerful Xilinx Virtex+ VU13P FPGA;
- an optical module – hosting up to 30 QSFP28 transceivers.

The final board revision is shown in Figure 2.24. This architecture closely follows the one of the CTP7, including an onboard Linux OS for control and management purposes. Therefore, minimal firmware and online software changes will be necessary to ensure adequate support.

Each ATCA crate will also be equipped with (at least) one **DAQ & Timing Hub** (DTH) board, shown in Figure 2.25 [42]. Serving as the interface between the sub-detector hardware and the central DAQ systems, the DTH essentially combines the functions of the AMC13, FEROL, and TCDS interface into a single unit.

The TCDS stream is received via an optical transceiver on the DTH front panel and then broadcast, along with the recovered LHC clock, to all sub-detector boards through the ATCA backplane. In the reverse direction, the TTS signals from all sub-detector boards are collected, summarized, and forwarded to the TCDS.

On the readout side, the DTH can receive data from up to 24 FEDs through its front panel optical links by using the custom 25 Gbps S-Link Rocket protocol. Internally, the DTH assembles event fragments for each orbit, individually for each FED. The resulting data stream is transmitted to the CMS event builder via one TCP/IP stream per FED on one of the 5 100 Gbps Ethernet ports.

[2]While the µTCA standard is technically derived from ATCA and was originally designed for simplification purposes, ATCA can still be considered an evolution in terms of features and capabilities.

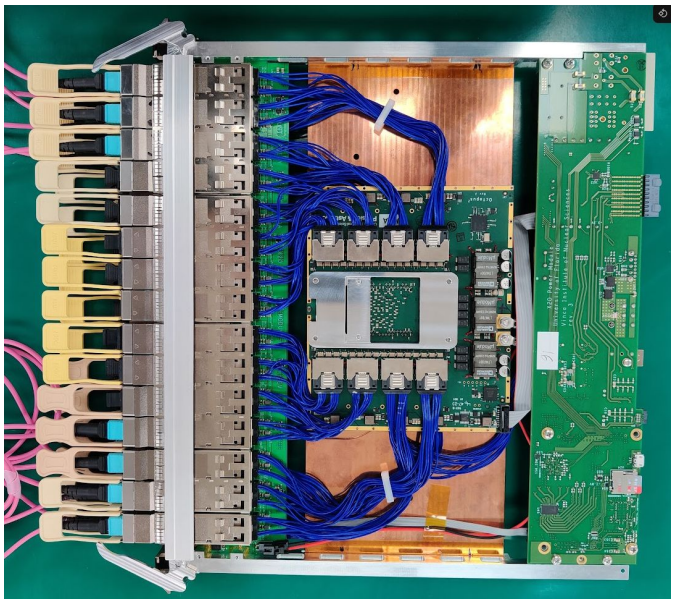


Figure 2.24: Picture of the X2O leaf board.

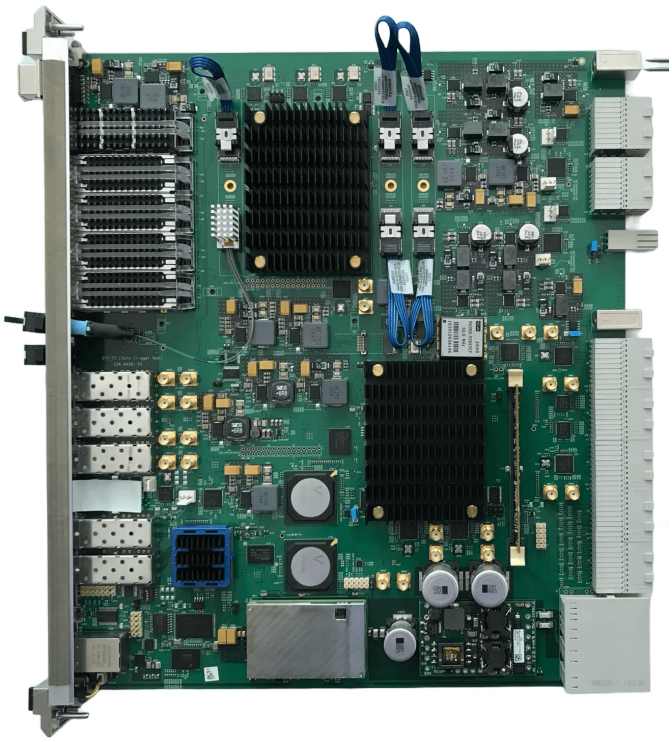


Figure 2.25: Picture of the DTH-400 ATCA hub board.

2.4 Summary

The GE1/1 project represents the first use of the Triple-GEM detector technology within the CMS muon spectrometer. Selected for its suitable rate capability, radiation tolerance, efficiency, timing, and angular resolution, it complements the previously installed CSC chambers in the most forward regions of the endcaps. Its primary goal is to enhance the trigger and reconstruction performance in anticipation of the challenging conditions expected during the HL-LHC era.

The GE1/1 station consists of 144 trapezoidal chambers – 72 short and 72 long – ensuring complete coverage without blind spots in its region. Each chamber is subdivided into 24 eta-phi regions each read out by a custom VFAT3 ASIC featuring 128 binary input channels. Based on the lessons learned from the GE1/1 slice test, the VFAT3 hybrid PCB was refined to include a resistor-based protection circuit, along with optimization of the chamber high-voltage filter, to reduce the input channel losses due to HV discharges. The communication between the counting room and the on-chamber electronics is handled by the OptoHybrid, a custom communication hub built around 3 GBTx ASICs and one Xilinx Virtex-6 FPGA. The GBTx chips manage the optical data transmissions, while the FPGA performs essential compression of the VFAT3 S-bits trigger primitives. The on-chamber electronics is complemented by the FEAST, which provides DC-DC power conversion, and the GEB, which provides connectivity between all components.

The back-end electronics, located in the counting room, is based on the µTCA telecommunication standard, providing a high-reliability platform for HEP DAQ systems. The core of the GEM back-end system is the CTP7 AMC, a versatile multi-purpose board that provides slow control, clock and trigger distribution, event building, and trigger concentration to all GE1/1 chambers. Its onboard CPU makes it particularly well-suited to the online software discussed later in this dissertation. The communication with the central CMS DAQ systems is handled by an AMC13 board.

The GE1/1 station is only the first of several GEM stations to be installed in CMS. The ME0 station, scheduled for installation during LS3, will consist of 36 six-layer stacks (18 per endcap), located closer to the beamline and interaction point than any other CMS muon detector, covering a pseudorapidity range of $2.03 < |\eta| < 2.8$, with particle rates reaching up to $150\,\mathrm{kHz\,cm^{-2}}$.

The GE2/1 station, whose production and installation have been postponed until after LS3, will complement the CSC system. Positioned further away from the interaction point than GE1/1, it will consist of 72 chambers, each composed of four independent modules.

In addition, the increased rates, readout latency, and throughput expected by the CMS Phase-II operations require a completely new back-end electronics design. Based on the ATCA standard, each crate will house one DTH-400 board – replacing the role of the ACM13 – and eight or nine X2O boards, developed in part by the GEM project – serving a role similar to the CTP7.

Chapter 3

The CMS GEM data acquisition system

A data acquisition system includes all the hardware and software frameworks necessary to collect, digitize, and process signals from sensors through permanent storage. The previous chapter described the hardware aspects, detailing the readout electronics tailored for the GEM detector technology and the challenges and constraints of the CMS environment. This chapter turns to the online software (sometimes abbreviated OS), which is required to operate the detector once it is built.

The online software serves two primary functions: control and monitoring of the data-taking.

Control First, the online software manages the configuration of the electronics and orchestrates all processes required to achieve optimal and coordinated data-taking operations. This includes executing calibration routines to fine-tune the detector's stability and performance.

Monitoring Second, the online software provides real-time access to the status and performance of the system. This enables the prompt detection of issues to ensure optimal data quality data-taking. In some cases, corrective actions can be triggered automatically in response to detected problems.

The chapter begins with an overview of the common software frameworks used across the CMS DAQ systems. It then presents the current GEM online software, explaining its design choices in light of the limitations of the so-called legacy software, inherited from the GE1/1 slice test. The discussion then covers the calibration aspects, essential for efficient data-taking. Finally, the chapter covers the monitoring tools and automatic actions that are implemented.

3.1 Common CMS software frameworks

To simplify and standardize the development and operation of DAQ systems across all sub-detectors, CMS has developed and maintains two dedicated software frameworks. xDAQ (Cross Platform Data Acquisition) is used for controlling the readout and trigger electronics, as well as for developing the CMS event builder [23]. RCMS (Run Control and Monitoring System) manages the systems at a higher level, providing a hierarchical view of the run status and centralized controls [26].

As with the other CMS sub-systems, the GEM online software is developed around those two frameworks. However, its design is strictly bound to them: industry-standard alternatives are adopted when suitable, and features of the frameworks conflicting with the GEM design goals are avoided. Notably, the monitoring suite described in Section 3.6 avoids xDAQ altogether due to its complexity and limited documentation.

3.1.1 xDAQ

The xDAQ framework, implemented in C++, is a fully-featured platform designed to meet the requirements of the CMS distributed trigger and DAQ systems. The framework provides a comprehensive set of libraries and tools covering all aspects of the online software: multi-threading, message serialization, database access, structured logging, application monitoring, Finite State Machines (FSMs),...

At the framework core is the xDAQ executive, a process that dynamically loads user applications at runtime from shared libraries. Each xDAQ executive loads a *context* defined in an XML (eXtensible Markup Language) configuration file. This file describes the applications, their parameters, and their interconnections.

In addition to the local context, the configuration file can also describe remote contexts, enabling seamless and efficient communication between a distributed set of applications using peer-transport mechanisms. Among the typical communication mechanisms are SOAP (Simple Object Access Protocol), which enables Remote Procedure Calls (RPCs) over HTTP, and I2O (Input/Output), which enables binary messages exchange over TCP/IP.

Another key feature of xDAQ, particularly notable at the time of its original in the early 2000s, is the built-in support for web-based user interfaces. Through a Common Gateway Interface (CGI), the xDAQ executive redirects HTTP requests to application-defined handlers, allowing the developers to implement custom web interfaces for configuration, control, and monitoring purposes. Thanks to this system, the operators and developers are able to interact with the DAQ software from anywhere using a standard web browser, thus facilitating remote operations.

3.1.2 RCMS

The RCMS framework, implemented in Java, is used to implement the CMS top-level control system. It organizes the CMS DAQ into a hierarchical control tree composed of so-called Function Managers (FMs). Each sub-system that participates in the global

data-taking is required to implement a dedicated FM that follows a standardized FSM. The resulting FMs can be operated independently – for standalone runs – or steered by higher-level FMs during global runs.

The development and deployment of the Function Managers are supported by several services provided by the framework:

- A security service to ensure authenticated user access and prevent unsafe detector operations.
- A resource service that gives access to the description of the control tree and the configuration of the distributed processes.
- A logging service used to collect, store, and retrieve both unstructured logs and structured metadata.
- A job controller to manage the lifecycle of xDAQ processes in a system distributed across dozens of machines.

Like xDAQ, RCMS provides a web-based interface, allowing remote operations via a standard web browser. However, unlike xDAQ, RCMS intentionally exposes a limited number of configuration parameters to the user. This design choice ensures that the entire CMS detector can be operated efficiently by a reduced number of trained shifters.

3.2 The legacy software

The so-called *legacy software*, inherited from the GE1/1 slice test and used throughout the GE1/1 quality control phase, represents the first version of the GEM online software fully capable of controlling and monitoring the final GE1/1 electronics across a broad range of scenarios: from small-scale operations within CMS – during the slice test with 5 super-chambers installed –, to the final GE1/1 quality control test stand with cosmic muons (QC8) at CERN, as well as various remote test stands. Only the minimal set of tools required for integration within CMS data-taking was implemented using the xDAQ framework. Most other functionalities relied on a large collection of organically developed Python scripts, designed to support exclusively the quality control procedures.

Despite its success in supporting the GE1/1 production and its wide range of features, the legacy system exhibited a number of architectural and operational shortcomings that hindered its viability for large-scale deployments.

Fragmented codebase The source code was split across a large number of Git repositories, making the system difficult to understand, build, deploy, and maintain. Adding new features or fixing a bug often required changes in two, three, or even four, different repositories, significantly complicating the development and slowing progress. Additionally, this fragmentation led to a steep learning curve, preventing streamlined operations and integration of new contributors, always relying on the

same few experts. At its worst, only a single person[1] was able to build the entire software stack from scratch.

Scalability limitations The legacy software could not scale to the operations of the full GE1/1 station in CMS. For example, a complete trimming procedure during the QC8 validation step – where up to 30 detectors were hosted – necessitated an entire night of exclusive data-taking and analysis. This duration was incompatible with the tight CMS operational constraints. Even when configured with a limited number of detectors, the initialization phase alone required 3 to 4 minutes, an unacceptable delay from the CMS operations' point of view. Two factors mainly contributed to this limitation. First, the operations were not parallelized, even where it would have been trivially achievable. Second, intrinsically slow routine were used throughout the code path, even where not required and easily replaceable.

Code path incoherences Since not all features were implemented using the xDAQ framework, nor any other unified framework, the integration of the various components with each other and with CMS presented a major challenge. Moreover, maintaining synchronization between competing features or redundant hardware access paths required continuous and cumbersome oversight to avoid conflicts or incoherences.

Inadequate error handling The legacy software lacked robust error detection and handling mechanisms to respond to communication failures with the front-end electronics. Some errors could occur silently, resulting in consequences only discovered much later – such as empty scans, unexpected analysis software crashes, or incorrect front-end configuration. In other cases, communication instabilities with the front-end would cause online software crashes without apparent reason reported to the users, preventing smooth operations in CMS, blocking the data-taking, and requiring continuous expert intervention for error diagnosis and resolution.

Obsolete code artifacts Due to its long lifetime, the codebase became cluttered with remnant code paths meant to handle earlier versions of the electronics, leading to a significant maintenance overhead and, more importantly, reduced code clarity.

Given these issues, the legacy software was deemed unsuitable for future use. The amount of resources required to improve it would have overpassed the amount of resources required for a new implementation. It was therefore set in maintenance-only mode – in order to keep supporting the ongoing operations – while designing a new and modern software stack capable of meeting the future needs of the GEM project.

[1] The author of this dissertation.

3.3 Software architecture

Given the aforementioned shortcomings, a new development workflow was established and a new software architecture was designed. This redevelopment effort was guided by the following key objectives.

Ease of development Considering the evolving nature of the collaboration – with contributors joining and leaving the project regularly –, the development workflow had to be tailored to ensure a seamless onboarding process. In particular, it should be straightforward for any new developer – or advanced user – to build the entire software stack from scratch and use it across various test environments.

Unified codebase and portability Considering the multiple flavors to support – a combination of GEM stations and backend boards –, efforts had to be made to centralize and commonalize the development as much as possible. Platform-specific dependencies must be minimized, and modularity must be a guiding principle.

Robustness against communication errors As detailed further in Section 5.2, major instabilities affect the communication between the front-end and the back-end electronics. Therefore, the software must be inherently tolerant against such communication errors to ensure optimal operational continuity, even under suboptimal conditions. Robust error handling and recovery mechanisms must be implemented throughout the software stack.

Phase-II readiness A special emphasis must be placed on ensuring compatibility with the demands of Phase-II operations. This includes scalability needs, as well as sufficient flexibility to handle evolving frameworks, interfaces, and deployment mechanisms.

These objectives were addressed through two main approaches. First, the number of repositories was drastically reduced and limited to two: one for the core online software used during data-taking (`cmsgemos`), and one for the analysis routines used in the offline calibration processing (`cmsgemos-analysis`). This structure, loosely following the mono-repository development model, was driven mainly by the use of different programming languages: C++ for the core system, and Python for the analysis part. Merging the two repositories constitutes a possible further simplification and remains under consideration. The reduction of repositories immediately simplified the development and maintenance by providing a unified view of the entire system, eliminating dependency-related errors, and improving the overall system integration.

Secondly, all variations between GEM stations and back-end boards were abstracted as low in the DAQ stack as possible. Ideally in the firmware, if feasible; otherwise, in the software components running on the back-end board; and, only as a last resort, in the higher-level control software. This abstraction strategy promotes modularity and reusability, at the cost of a slightly higher complexity during the initial implementation of new features.

The software architecture implemented as a result of these design approaches is illustrated in Figure 3.1. Its various constituents are described thoroughly in the following sections.

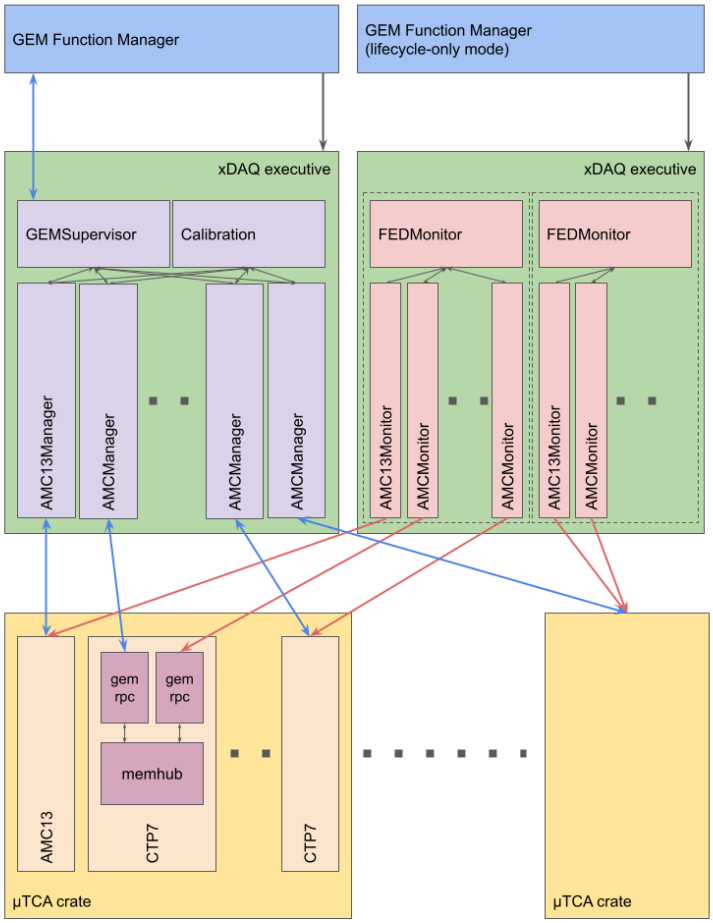


Figure 3.1: Functional block diagram of the CMS GEM online software architecture. Hardware components are shown in yellow, with embedded software services highlighted in purple. xDAQ-based processes are represented in green, whereas RCMS-based processes are marked in dark blue. Control-related components are shown in light blue, while monitoring functionalities are colored in red.

3.3.1 The back-end services

The GEM online software architecture conceives the back-end board processing system as an accelerator layer, offloading frequent and low-level hardware interactions from the high-level control applications. To this end, the back-end hosts two services: `memhub`, a memory-access layer that provides low-level access to the electronics registers, and `gemrpc`, an RPC server that implements stateless high-level routines.

Registers access

Systematically parsing the XML address table provided in each firmware release on the CTP7 embedded Zynq CPU proved to be too slow and resource-intensive to be usable for operations at a large scale[2]. Therefore, it was decided to store the register map in an LMDB (Lightning Memory-Mapped Database) database [99], after parsing the address table only once upon firmware updates. The LMDB is a key-value, memory-mapped database chosen for its high performance, low overhead, and zero-copy access pattern. This allows the register map to be loaded within milliseconds and individual registers to be looked up within microseconds.

The `memhub` library provides register access by directly poking the memory-mapped addresses, once their names are resolved via LMDB. In addition, the `memhub-server` application centralizes the memory access: it serves the correct memory segment to clients and provides locking primitives shared among all clients. The latter is essential for ensuring multi-processing safety and preventing race conditions during non-atomic operations. For example, interfacing with the GTBx requires reading and writing to multiple backend firmware registers. Without proper locking, simultaneous access from multiple clients could be interleaved, causing erroneous transactions or software crashes. Importantly, the server is only contacted during the initialization phase to retrieve the memory handle and locking primitives, so that the actual register access performance remains unaffected.

Remote procedure calls

Alongside the `memhub` library and service, the `gemhardware` package is designed to run on the backend board. It provides a set of fully stateless routines to accelerate management operations (e.g. configuration, control, and monitoring) leveraging the low-latency, direct access to hardware while reducing the load on the control applications.

As mentioned previously, a key component of this package is the `gemrpc` service, which enables control applications to transparently delegate processing tasks via RPC methods. Despite its essential role, the RPC interface initially inherited from the legacy software was neither robust nor maintainable. Its limitations can be summarized as follows:

- The requests and responses had to be manually built.

[2]Parsing the full-scale GE1/1 address table in the legacy Python-based software was leading to systematic out-of-memory kills, artificially reducing the number of OptoHybrids supported per CTP7.

- Missing keys – typically caused by software updates – could silently crash the server or the client.
- Exceptions thrown in the RPC methods could not be caught, leading to the termination of the server.

To address these limitations, the RPC system was completely overhauled with the following goals:

- Build automatically the requests and decode the responses based on the C++ function signatures.
- Check the argument types at compile time.
- Forward exceptions from the server to the calling site, even if remote.
- Minimize the overhead at runtime.

The solution, heavily leveraging C++ templates, avoids the need for external preprocessing tools and streamlines the integration of new developers. A key design principle is that the RPC methods are declared only once and shared between both the client and the server. This guarantees consistency between the call site and the implementation. This also allows the function to be called both locally (from other RPC methods) or remotely (from the client code) without code duplication.

The RPC method arguments are serialized to a binary format using the `cereal` library [64], a fast and lightweight header-only C++11 serialization library. It supports all C++ standard library data structures and can be extended to handle custom data types. And, although the semantic meaning of the arguments cannot be checked (for obvious reasons), the system ensures syntactic and type correctness. Similarly, the return values are also serialized via `cereal`, and exceptions thrown by the RPC methods are caught and re-thrown at the calling site instead of silently returning invalid data.

As a user, though, this is only an implementation detail: an RPC method behaves exactly like a regular C++ function call (with a specific call syntax), but runs on a remote node. Listing 3.1 provides a simple, self-explanatory example of an RPC method declaration, implementation, and call. Notably, one can witness the absence of code duplication and, on line 39, the transparent offloading of the task to the backend board processing system.

3.3.2 Control applications

This section provides a concise overview of the core control applications deployed on the control machine: the `GEMSupervisor`, the `AMCManager`, and the `AMC13Manager`. These applications form the essential part of the high-level control infrastructure, coordinating the configuration, the state transitions, and the running of the detector. Applications related to the calibration and the monitoring of the detector are presented further in the dedicated Sections 3.5 and 3.6, respectively.

```
//// memory.h -- RPC Method declaration

namespace Memory
{
    struct Read : public RPC::Method
    {
        std::vector<std::uint32_t> operator()(std::uint32_t address,
                                              std::uint32_t count) const;
    };
}

////  memory.cpp -- RPC Method implementation

#include "memory.h" // Common header!

std::vector<std::uint32_t> Memory::Read::operator()(std::uint32_t address,
                                                    std::uint32_t count) const
{
    std::vector<std::uint32_t> result(count);
    if (0 == memhub_read_block_raw(address, count, &result.front())) {
        return result;
    } else {
        // Exceptions are forwarded to the client
        throw std::runtime_error(
            std::string("Read error: ") + memhub_get_last_error());
    }
}

//// client_test.cpp -- Remote RPC Method call

#include "memory.h" // Common header!

int main(int argc, char **argv)
{
    try {
        RPC::Connection connection;
        connection.connect("my-ctp7");

        auto mem = connection.call<Memory::Read>(0x6640000c, 1);

        std::cout << "CTP7 Virtex-7 firmware release : "
                  << std::hex << word.at(0) << std::endl;
    } catch (const std::exception &e) {
        std::cout << e.what() << std::endl;
    }

    return 0;
}
```

Listing 3.1: Example of an RPC Method declaration, implementation, and remote call.

The GEMSupervisor

The `GEMSupervisor` serves as the top-level controller of the GEM software stack. It acts as the primary interface between the GEM online software and any control system – either through direct user interaction in standalone deployments, or through indirect control via higher-level external systems such as the GEM Function Manager.

The `GEMSupervisor` first responsibility is to relay commands to the supervised applications. It orchestrates their state transitions of all applications under its control, ensuring proper sequencing. Additionally, it collects and aggregates the status and error reports from all supervised applications, providing a coherent and summarized system state to the user.

The supervised applications fall into to categories: GEM and non-GEM. Supervising non-GEM applications is useful in two scenarios. First, on test stands, applications provided by the cDAQ group – typically controlled by their own Function Manager – can instead be controlled in a unified and simplied way. This is commonly used to spawn a small-scale copy of the CMS event builder for development purposes. Second, in the production system at CMS, the TCDS control applications run as a service. Each sub-system must control the services managing its own partitions; for GEM, this role is fulfilled by the `GEMSupervisor`.

The AMCManager

Named for historical reasons after the µTCA AMC board, the `AMCManager` is responsible for all management and control tasks related back-end electronics board and associated firmware. One instance of this application is instantiated per CTP7.

Compared to the legacy software architecture, the scope of the `AMCManager` has been significantly expanded. Not only does it manage the back-end itself, it also interfaces with the front-end electronics connected to it. While it might appear conceptually cleaner to separate these responsibilities using an `OptoHybridManager` application, doing so would complicate the overall implementation.

For example, consider the synchronization of the readout and trigger masks – located in the back-end firmware – with the state of the front-end electronics. A two-application design would require frequent back-and-forth communication to maintain consistency applications and electronics components, including at least one request for every state update, and likely more to capture the global state of the system. By contrast, a one-application approach avoids this overhead. It removes the needs for constant communication and coordination, which is particularly beneficial given the highly dynamic nature of the system and the fact that all communication with the front-end must pass through the back-end board.

Instead of separating concern across applications, the `AMCManager` maintains a clear internal structure: the front-end and back-end features are encapsulated in distinct functions, while sharing a unified global stated.

The AMC13Manager

Finally, the `AMC13Manager` is responsible for controlling the AMC13 board. Unlike the `AMCManager`, its scope is relatively limited, as the firmware AMC13 implementation is largely automatized and does not require calibration features. The `AMC13Manager` only configures a small set of hardware registers to ensure proper forwarding of the TCDS clocks and commands to the CTP7, and to enable the appropriate slots in the event building.

The finite state machine

As anticipated in the previous subsections, the behavior of the control applications is driven by a Finite State Machine (FSM). The FSM ensures that all operations follow a well-defined sequence of transitions with well-defined hardware states, preventing unforeseen actions. Each application contains one independent FSM instance.

Figure 3.2 represents the FSM implemented in the GEM online system. For clarity, only stable states are shown. During each transition, the FSM passes through an intermediate state, signaling that the command has been received and is being processed. Only once the operation is complete does the FSM move into the next stable state.

Upon creation, the GEM FSM begins in the **Initial** state. At this stage, the application is instantiated with the default configuration parameters defined in the xDAQ XML configuration file.

The *Initialize* command transitions the FSM to the **Initializing** intermediate state. During this phase, the application self-configures based on additional configuration files (see Section 3.3.3). It establishes communication with external components – such as databases, or other GEM and non-GEM applications – to ensure full functionality. The application is no longer in standalone mode. Nevertheless, no communication with the readout hardware is yet performed, as the final list of hardware components will only be provided later. Once the initialization phase is complete, the FSM enters the **Halted** state.

The *Configure* command moves the FSM to the **Configuring** intermediate state. The configuration sequence is arguably the most important. During this step, the entire electronics chain is reset and reconfigured. This includes reloading the back-end FPGAs firmware and reconfiguring all front-end ASICs – GBTx, SCA, VFAT – as well as FPGAs. All configuration parameters are subsequently written to the corresponding hardware registers. At this point, valid trigger primitives are already sent to the other Level-1 Trigger boards. Only the final steps of enabling the event building and data readout upon trigger remain to be done. At the end of the configuration sequence, the FSM enters the **Configured** state: the hardware is ready for data acquisition, depending on the type of run requested.

While in the **Configured** state, the system can be re-configured at any time, restarting the full configuration cycle, which can help clear any potential hardware issues.

The *Start* command transitions the FSM to the **Starting** intermediate state. This sequence turns out to be relatively lightweight, as the electronics chain is already fully configured and sending trigger data. Its main task is to enable the event building and ensure that all components are synchronized. During this transition, the final run number is assigned and can be used for tagging logs and metadata. Additionally, the automatic masking and recovery mechanisms (described further in Section 3.7) are enabled, ensuring reliable data-taking. Upon completion, the FSM reaches the **Running** state, where the full data acquisition system is active and ready to accept triggers.

The **Paused** state is reachable from the **Running** with the *Pause* command (and the intermediate Pausing state) and exited through the *Resume* command (and the intermediate **Resuming** state). Although included in the standard GEM FSM for consistency, the `Paused` state is only used by the `AMC13Manager` and has no operational aspect in other applications.

Finally the *Halt* command, via the **Halting** intermediate state, returns the system to a state identical to **Halted** following the initialization. Similarly, the *Stop* command, through the **Stopping** intermediate state, puts back the system to a state compatible with **Configured**, but without reconfiguring the electronics.

From any state, the FSM can transition to the **Error** state, spuriously or not, via the *Fail* transition. This transition should be reserved for critical, unexpected errors that the software is unable to handle – such as invalid configuration parameters – and recoverable issues that are part of normal operations, – such as communication instabilities with the front-end electronics. The **Error** state is terminal: no further commands are accepted, and recovery is only possible by destroying and restarting the process from scratch.

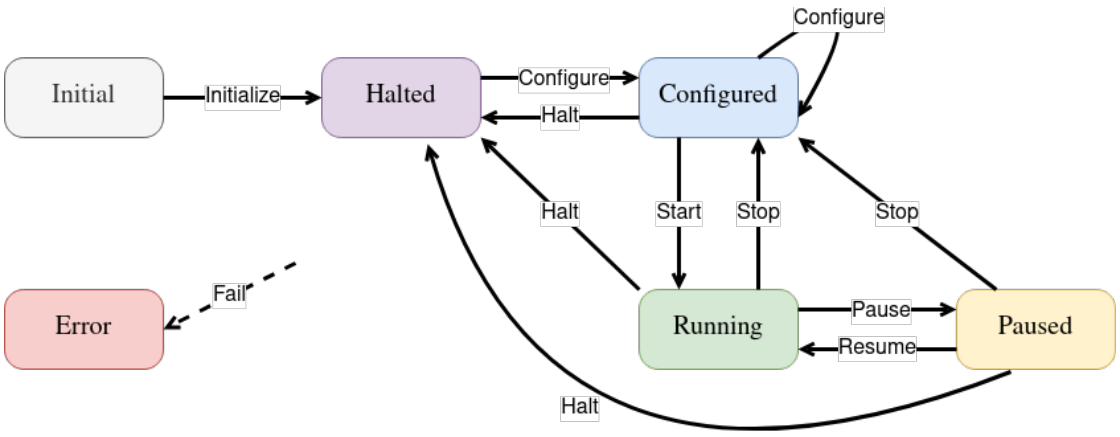


Figure 3.2: Finite State Machine used by the GEM online software applications. For clarity, only the stable states are represented.

3.3.3 System configuration mechanisms

To operate the system, all its components must be properly configured: the online software must be set up; the hardware must be described to sufficient details; and the

hardware parameters must be set to their optimal values. And, while unification is always sought, the technologies involved and their respective purposes differ significantly. Therefore, the three independent configuration mechanisms described below are implemented.

xDAQ executive

As GEM applications are inherently xDAQ applications, they must be configured using its standard mechanism: the executive is only able to instantiate applications described in XML configuration files. However, unlike other online software projects that fully embrace the xDAQ framework, its usage within the GEM online software is reduced to the minimum. Only the most essential configuration parameters required for each application to self-configure are included in the XML files.

This minimalist approach is motivated by maintainability. On large-scale systems, XML configuration files can quickly become unmanageable, as they describe a large number of applications. Moreover, no standard generation tool exists, so these files are usually generated by hand – minimizing their complexity then significantly reduces the operational overhead and the potential for human error.

Layout tree

A key component of the configuration infrastructure is the *layout tree*. Stored in YAML (YAML Ain't Markup Language) format [113], it provides a complete description of the hardware structure that the online software instance is expected to control: crates, back-end boards, GBT links, OptoHybrid FPGA, VFAT, and more.

Once parsed and loaded into C++ structures, the layout tree becomes a foundational building block for the rest of the GEM online software. Whereas the xDAQ configuration focuses on minimalism and the detector configuration strictly on register values, the layout tree provides a global view of the entire detector system. It gives each application the possibility to self-configure dynamically, based on its location in the system hierarchy.

For example, the list of detectors expected to be included in the data-taking is defined exclusively in the layout tree. This information is then used to adjust the web interface to accurately represent the system, or to interface with external systems which know only about existing detectors.

Detector configuration

Finally, detector configuration parameters – such as VFAT thresholds or bias values – must be stored and retrieved consistently. The current storage mechanism relies on a set of key-value text files, one per configurable electronics element. While fully functionally sufficient, it presents a major limitation in traceability: tracking configuration changes over time is only possible through careful bookkeeping by the operator.

To partially address this limitation, a JSON-based snapshot of the system's most relevant parameters has been implemented. For each run, this dump allows to unambiguously recover quantities such as the VFAT3 readout latency or threshold.

This is however not the complete solution. A new, more robust configuration system, based on a PostgreSQL relational database, is currently under development. This system aims to support efficient retrieval of any past or present configuration, support long-term archival, and provide full tracking of the configuration changes over time. Nonetheless, the text-based configuration mechanism will remain for small-scale setups, testing environments, and development purposes, where simplicity is preferable.

3.3.4 The Function Manager

While one of the design goals of the GEM online software is to ensure broad compatibility with multiple usage scenarios, its primary role remains to support the data-taking activities within CMS. This is where the Function Manager comes into play. Implemented using the RCMS framework, the GEM FM interfaces with the central CMS DAQ system, enabling shifters and experts to control the lifecycle of the GEM applications, issue commands, and monitor the status of the GEM sub-system.

To meet these divergent objectives, the GEM Function Manager was conceived as a lightweight shim between the central DAQ and the GEM environments. All core functionalities are implemented directly within the `GEMSupervisor` application. The FM only spawns the xDAQ executives through the `Job Controller`, forwards central commands to the supervisor, and monitors its status, which it then reports to the central DAQ interface.

The GEM Function Manager is not designed to take calibration procedures or special-purpose runs. However, as the primary entry point for the users working within the CMS context, one additional feature was introduced: the ability to operate in *lifecycle-only* mode. In this mode, the FM solely manages the lifecycle of the GEM applications without forwarding commands or status between the central and GEM environments. This option greatly simplifies the deployment and operations in standalone local scenarios.

Figure 3.3 shows the screenshot of the GEM sub-system fully integrated in a MiniDAQ-3 run via its Function Manager.

3.4 Local readout

As it will be made clear in this and the following sections, the local readout is a key component of the GEM DAQ system. It provides a high-bandwidth, Ethernet-based data path that operates independently of the central DAQ system described in Section 1.2.5. Despite some apparent similarities, the local readout does not aim at replicating the central DAQ functionalities. Instead, it aims at complementing the central DAQ system with features it does not provide. Notably, the local readout performs no event building: the event fragments either are stored in a raw form for future analysis or immediately processed for online analysis. The rationales for implementing this parallel readout system are the following:

Calibration scans The backend firmware includes a ”DAQ monitor” block, which allows a coarse inspection of the DAQ data from a single OptoHybrid. However,

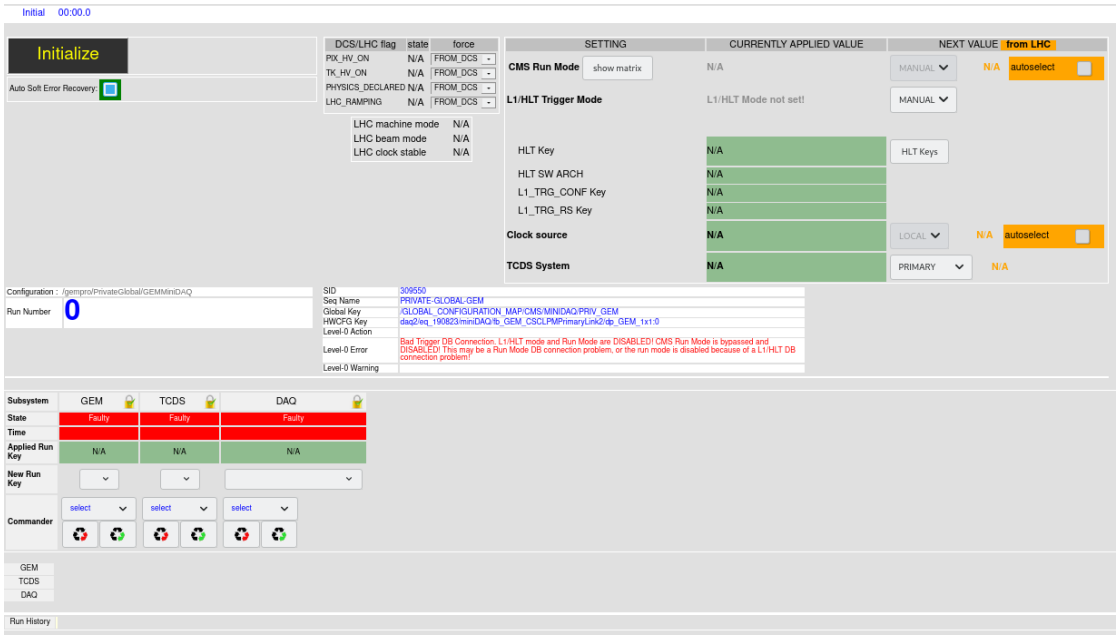


Figure 3.3: Web interface of the Level-0 Function Manager running in a MiniDAQ-3 instance, and using the GEM Level-1 function to interface with the GEM sub-system.

this approach does not scale to the level required for CMS operations. To lift this limitation, this local readout was designed to allow for the full DAQ data stream to be dumped to storage. Its flexibility to implement data formats specifically optimized for calibration purposes, along with the high available bandwidth, further enhanced this capability. Moreover, the local readout enables the possibility of performing real-time calibration analysis, a possibility currently not exploited.

Local operations The MiniDAQ-3 system presented in Section 1.2.5 is ideal to mimic the global CMS operation conditions and investigate any integration issues. However, it is not suitable for all use cases. Its usage is occasionally disrupted by cDAQ testing; its dataflow incurs a latency of days or hours - as data must be packed at Tier-0 - which prevents prompt analysis; and it is unavailable at test stands. Supporting fully all use cases with MiniDAQ-3 would require significant duplication of the efforts. On the contrary, the local readout allows for seamless development and usage in all conditions.

Pre-HLT data analysis The central DAQ only records data that passes the HLT selections. While this is essential for physics analysis, this interferes with the processes of understanding the detector and the diagnosing anomalies. For example, analyzing unfiltered data chunks around corrupted events is not possible. Since the local readout runs in parallel with the cDAQ, it can apply an entirely different filtering logic, enabling the capture of such data. The ability to process

pre-HLT data also supports running a local Data Quality Monitoring (DQM) fed with larger samples than those typically available through the central online DQM. This has proven to be a valuable asset for detector health monitoring and problem debugging in the CSC system.

Trigger data readout Currently, no trigger primitives are stored in the GEM DAQ stream, limiting their accessibility for analysis. While this does not pose a fundamental problem for the Level-1 Trigger operations, it significantly affects the debuggability of the system. Indeed, the trigger primitives must be captured via other paths - in the CSC OTMB or the L1T EMTF DAQ streams -, which prevents standalone investigations. The local readout path, through a dedicated data format, provides the ability to record the trigger data directly.

Figure 3.4 illustrates the architecture eventually developed for the local readout system. The backend boards transmit raw Ethernet frames containing the event data over 10 Gbps links. The frames sent by a given backend board are directed to a single MAC address manually configured via firmware registers to match the address monitored by the Network Interface Card (NIC) on the receiving computer. To minimize the processing overhead, Jumbo frames of up to 9000 bytes are used. Events exceeding this size are split across multiple frames – but one frame never contains data from more than one event. In small-scale setups, the Ethernet switch can be omitted, and a direct connection between the backend board and the computer NIC can be established. In the CMS production system, 40 Gbps links are used between the switch and the readout computers to avoid any bottleneck.

The readout application running on the receiving computer is developed using the Data Plane Development Kit (DPDK) framework [79]. This framework provides a set of high-performance libraries designed to accelerate the packet processing on traditional CPUs by bypassing the kernel and accessing the NIC directly from user-space. This enables the system to capture raw Ethernet frames at high rates and throughput with minimal overhead.

The readout application also implements on-the-fly compression, which reduces the size of the output files. This is particularly useful when the processing rate is limited by the storage speed. On the other hand, if the performance is limited by the CPU, the compression can be disabled.

After optimization of the buffer sizes and assignment of the processing threads to specific CPU cores (CPU pinning), the system has proven to be able to saturate a 10 Gbps link, which is sufficient for calibration scans, quality control, and test beam campaigns.

3.5 Calibration suite

The calibration suite contains all test and calibration routines required to ensure optimal operation of the detector. Its implementation is distributed across the software stack

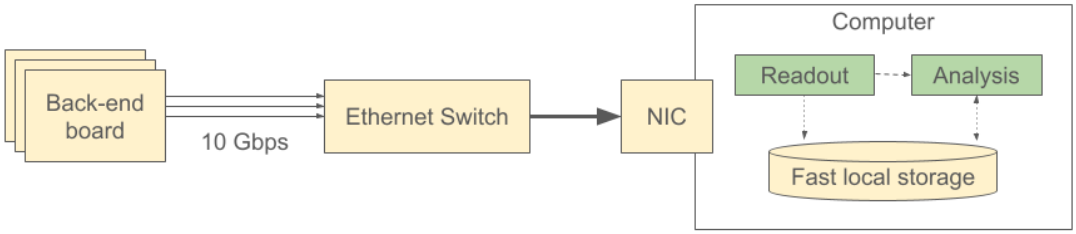


Figure 3.4: Flow of the local readout data. The switch component is optional and only used in the largest scale systems. The hardware components are colored in yellow whereas the software components are colored in green.

described previously in Section 3.3 and Figure 3.1. The scan routines themselves are implemented in the `gemhardware` package as RPC methods. Those methods are invoked and their results are saved to persistent storage by the `AMCManager` application. An additional xDAQ application – creatively named `Calibration` – orchestrates and controls the scans. Figure 3.5 presents the calibration suite web interface used to configure and initiate the scans – in this example, an S-curve scan.

One of the key factors driving the new software stack is scalability. The calibration suite is no exception: all scans are implemented to support execution on the full-scale system, including in the presence of major communication instabilities. Contributions that break this requirement are not accepted. Consequently, the current implementation is unable to orchestrate scans on a subset of the detector. However, since such functionality can easily be justified, work is planned to add per-OptoHybrid scans in future updates.

The available scans are described in the subsections below. The *GBT phase scan* establishes communication with the front-end electronics; the *DAC scan* tunes the VFATs' analog front-end bias settings; the *S-bit rate scan* determines the appropriate VFAT discriminator thresholds; the *S-curve scan* characterizes the electronics noise; the *threshold scan* diagnoses connectivity problems; and the *latency scan* aligns the detector sampling point in time relative to the trigger. The first three scans are qualified as *non-tracking data* scans, as they do not use the VFAT3 tracking data, but instead extract the data through the slow-control path and save the output as text files. In contrast, the last three scans are qualified as *tracking data* scans, as they utilize the VFAT3 tracking data recorded via the local readout application and saved in the standard GEM DAQ data format.

Additional scans implemented due to specific operational needs will be described further in Chapter 5.

3.5.1 GBT phase scan

The downlink eLink from the GBTx ASIC and the VFAT3 transmits both clock and data signals with similar delays, making it straightforward for the VFAT receiver to sample the data line at the correct moment. However, uplink eLink from the VFAT

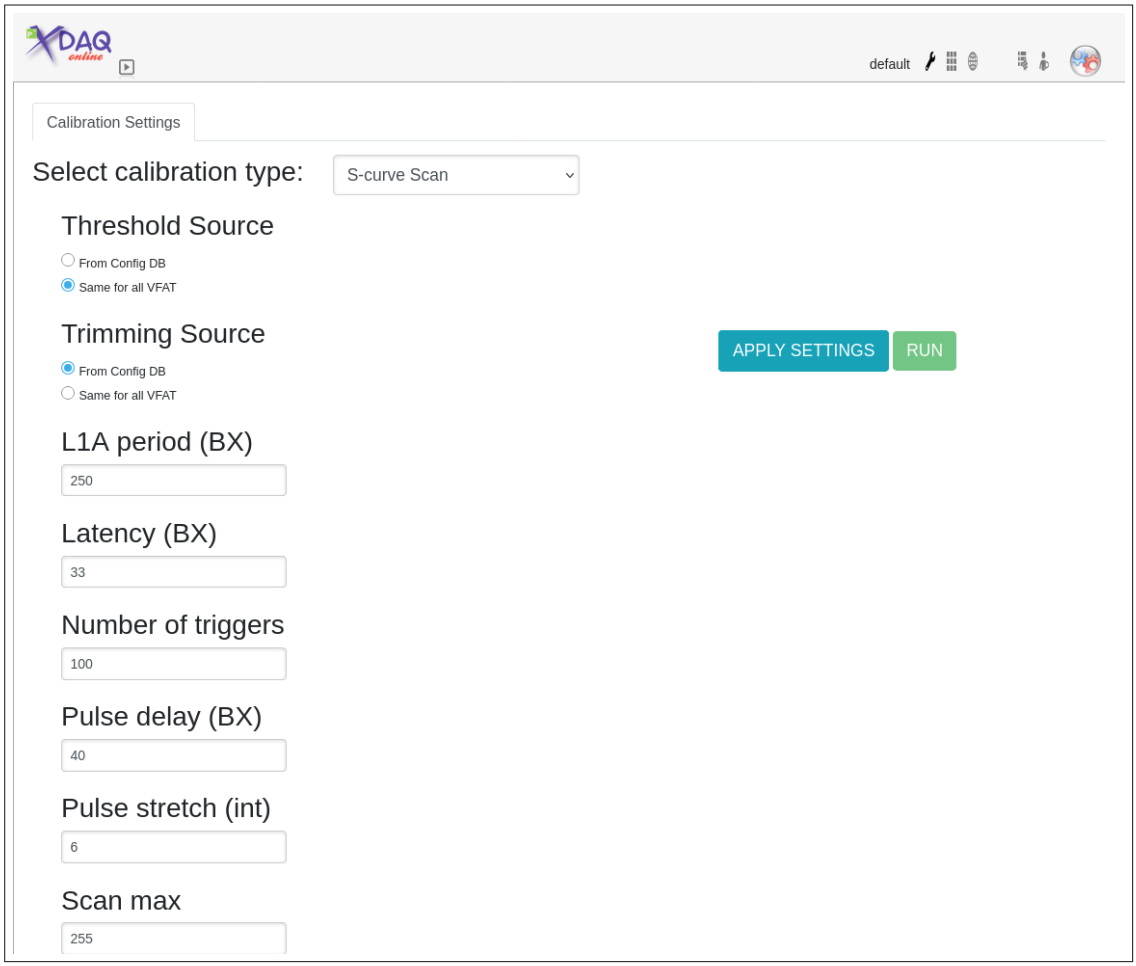


Figure 3.5: Screenshot of the calibration suite web interface for taking a S-curve.

to the GBTx does not transmit an independent clock signal. As a result, the phase relationship between the incoming data and the GBTx sampling is unknown. To resolve this issue, the GBTx is equipped with a dedicated phase-aligner circuit for each eLink. This circuit allows the selection of the optimal sampling point for the incoming data, with a resolution of one-eight of the 320 MHz sampling clock (~ 391 ps).

The *GBT phase scan* is performed to determine the optimal phase setting for each VFAT. For each possible phase value, a large number of slow-control transactions are issued to test registers. If any transaction fails, the corresponding phase is flagged as bad. Similarly, a sequence of L1A commands is sent to the VFATs, and the resulting DAQ packets are analyzed by the backend firmware. If too few or corrupted packets are received, the phase is likewise marked as bad. Only phases for which no communication errors have been detected are considered good. The optimal phase for each VFAT is then selected as the central one within the window of good phases – essentially, the middle

of the opening of this coarse eye diagram. In the case where all phases are marked as good, the selection process falls back to default values derived from previously tested detectors of the same type.

Figure 3.6 shows the aggregation of 22 such phase scans performed on long GE1/1 detectors For each VFAT and GBT pair, the number of times the given phase was considered valid is displayed. These results were used to construct the lookup table default values.

Beyond calibration, the GBT phase scan is also a useful debugging tool. A complete absence of communication or a valid window narrower than expected can point to various hardware faults, such as a broken FEAST or an incorrectly plugged VFAT.

Regarding the scalability, few optimizations are possible as the scan is dominated by the slow-control transactions duration.

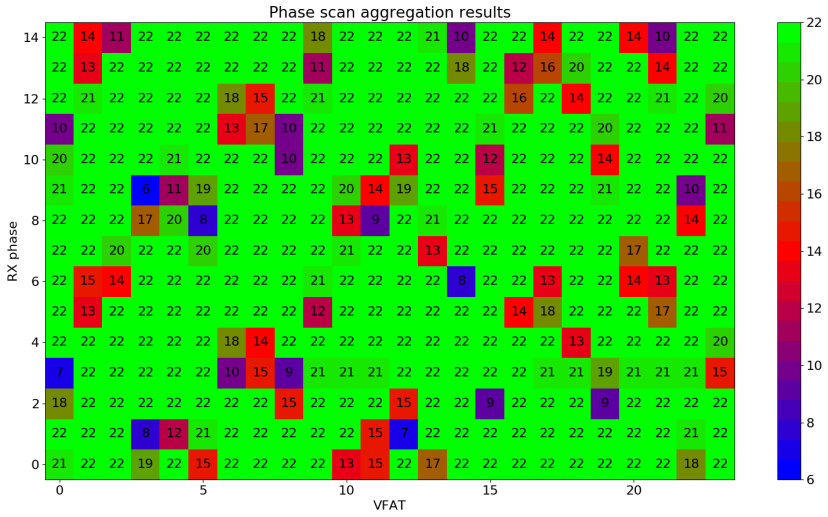


Figure 3.6: Aggregation of the GBT phase scan results from 22 long GEBs during the electronics quality control. The number corresponds to the number of GEBs for which the phase was measured as valid for a given VFAT slot. Large windows of ”22” show that the bad phases mostly depend on the electronics design.

3.5.2 DAC scan

In order to achieve consistent and nominal performance, the various circuits of the VFAT3 ASIC must be biased according to their specifications. Dedicated 6-bit or 8-bit DACs are used to provide and fine-tune the required voltages and currents. Integrated monitoring ADCSs are available either the current – via a 10 kOhm resistor – or the voltage generated by the respective DAC circuits.

For each DAC, the *DAC scan* routine sweeps its full range and records the corresponding ADC readings. After converting these ADC readings back into physical units representing the signals generated by DAC, the optimal DAC value is determined as the one whose reading is the closest to the specifications.

In addition, two sanity checks are performed. First, the computed ideal DAC value is checked to be within the allowable range. If it exceeds the maximum value, it is clipped and a warning is emitted. Second, a fifth-degree polynomial fit is attempted on the recorded ADC response. If the resulting reduced χ^2 is above a predefined threshold, a warning is also emitted. The thresholds were determined empirically to flag, with high probability, all VFATs that require a human check, while accepting a reasonable false-positive rate.

Figure 3.7 shows a typical result of a DAC scan for a single VFAT.

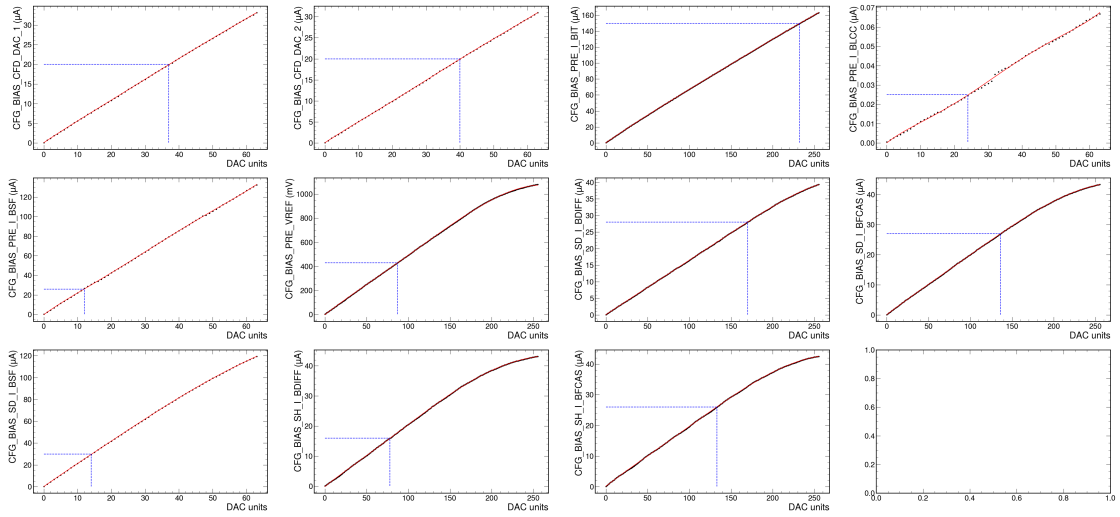


Figure 3.7: Typical DAC scan results. The generated current or voltage is shown a function of the DAC set point for each of the VFAT3 DAC circuits.

3.5.3 S-bit rate scan

As described in Section 2.2.3, the VFAT3 features a binary readout. In other words, the digitized signal will be either recorded (logic '1') or discarded (logic '0'). The decision threshold – i.e. the level at which the signal is considered significant – is set by the arming comparator threshold (*THR_ARM_DAC*). This threshold is controlled by an 8-bit DAC circuit shared between all 128 channels.

Unlike the DAC circuits presented in the previous section, which are configured according to VFAT3 specification values, the comparator threshold must be chosen based on the acceptable hit rate due to the electronics noise. The threshold is set such that 0 Hz of noise is accepted per VFAT, essentially rejecting any hits coming from the electronics noise.

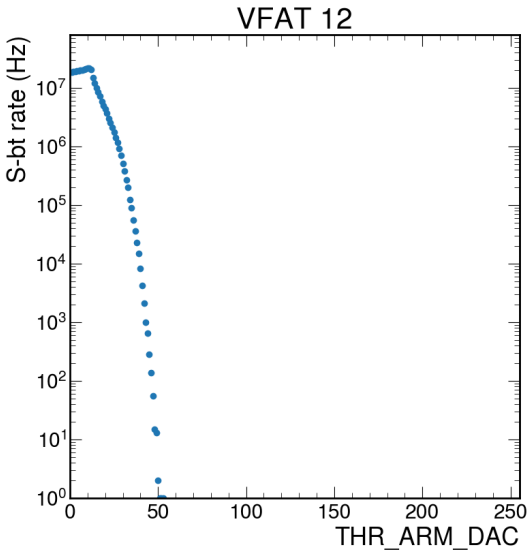


Figure 3.8: Typical example of an S-bit rate scan in the absence of detector signal for a single VFAT. The X-axis represents the *THR_ARM_DAC* value, scanned from its minimum (0) to its maximum (255); the Y-axis represents the measured S-bit rate, in Hertz.

Rather than relying on simulations or ideal scenarios, the relationship between comparator threshold and the noise is directly measured. This is achieved via the *S-bit rate scan*, which counts during a configurable number of bunch crossings how many times any of the S-bits from a VFAT fires, as a function of the comparator threshold. The S-bits are particularly advantageous in this context hit information can be recorded for every BX, significantly improving the statistical precision of the measurement.

The operating threshold is chosen as the lowest DAC value that ensures a noise rate below the predefined acceptable value. Figure 3.8 shows a typical result from one such scan in the absence of detector amplification. While the chamber high-voltage can be switched on for cross-check purposes, it must remain at least partially off during scans used to establish thresholds to avoid contamination from dark counts or signals from genuine muons.

The scan duration is limited only by the time spent accumulating S-bit counts at each threshold value – typically 1 second per point. Moreover, it has been shown that the intrinsic noise of the VFAT3 can change when toggling its run mode status [62]. To reduce power consumption, the VFAT3 can be placed in *sleep* mode, which powers off of the analog part of the readout channels. Conversely, the *run* mode powers on the analog circuitry and correctly biases the readout channels. The transition between those two modes – occurring at each configuration – results in this noise instability. The only effective mitigation to this date is to average multiple scans (by default 6), leading to a total acquisition time of about 30 minutes for a reliable determination of the thresholds.

3.5.4 S-curves

While the threshold-setting method based on the recorded electronics noise described in the previous section is effective, it does not provide information on the noise magnitude in physical units, nor allow to characterize the detector noise levels. Due to the binary nature of the VFAT3 readout, the input charge pedestal cannot be determined directly. Instead, it can be obtained only by injecting calibration charges into each readout channel, one at a time, in the so-called *S-curve scan*.

The S-curve scan is performed as follows: at constant comparator threshold (*THR_ARM_DAC*), each input channel is injected with a sequence of delta-like voltage calibration pulses increasing charge, defined by the *CAL_DAC* DAC. For each charge level, the ratio of hits to the number of injections is recorded. The resulting curve resembles the shape of an "S" as show on Figure 3.9, hence the scan name.

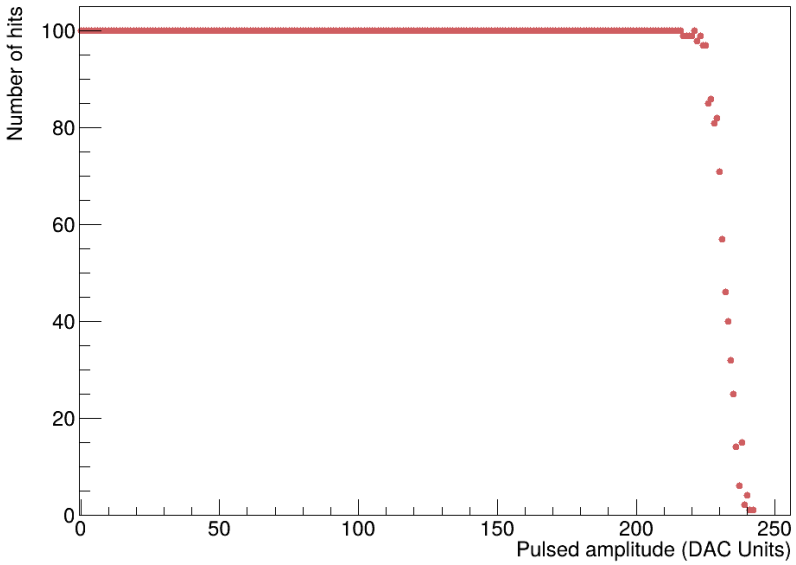


Figure 3.9: Example of an S-curve scan for a single channel on one VFAT. The plot shows the number of hits as a function of the injected calibration pulse charge, at a fixed VFAT threshold. For each charge values, 100 pulses are injected. The decreasing trend is due to the fact that smaller (larger) DAC values correspond to higher (lower) charges.

In an ideal, noise-free system, one would expect a step function: no hits are recorded when the injected charge is below the discriminator threshold, and hits are recorded with a 100% efficiency once the charge exceeds the threshold. However, in reality, the electronics noise smears this step transition. Assuming the noise amplitude follows a Gaussian distribution, the response curve becomes the convolution of a Heaviside function and a normal distribution:

$$f(q) = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{q-\mu}{\sqrt{2}\sigma}\right)$$

where μ represents the effective threshold – i.e. the 50% efficiency point – and σ is the Equivalent Noise Charge (ENC). If the injected charge is converted from DAC units to physical units (fC), the ENC can be used to compute the Signal-to-Noise Ratio (SNR) of the electronics. Similarly, the effective threshold is available in physical units and can be used to get a coarse idea of the expected efficiency.

Figure 3.10 shows the raw result of an S-curve scan for all 3072 channels of a GE1/1 detector, grouped by VFAT, as a color map. The color code represents the hits ratio; the X-axis corresponds to the channel number and the Y-axis to the injected charge. The threshold can be inferred from to the position of the transition region between white (0% hit) and yellow (100% hit) colors; the noise is related to the width of this region. An experienced eye can therefore quickly estimate the noise, threshold, and threshold variation across channel based on the raw data.

Figure 3.11 shows boxplots of the extracted threshold and ENC values for each VFAT. In the boxplots, the boxes span from the 25th to the 75th percentile; the whiskers extend to 1.5 times the InterQuartile Range (IQR); and the individual circles represent the outliers. Within each box, the solid line indicates the median whereas the dashed line represents the mean.

In addition to the overall VFAT characterization, the S-curve scan is an essential diagnostic tool to identify problematic channels. Disconnected or noisy channels exhibit an anomalously low or high ENC. Dead channels, which do not respond to calibration pulses at all, appear as white vertical lines in the color map summary plot.

While the data readout can be efficiently handled via the local readout mechanism and does not require specific optimization attention, the scan procedure does. Indeed, it is particularly intensive in slow-control transactions, which could hinder the scan duration and thus scalability. To mitigate this, two optimization strategies were implemented:

1. The results from the register lookups in LMDB are cached in local C++ structure, avoiding repeating queries that are fast but still too slow when executed in large quantities.

2. The scan was structured to minimize the number of register accesses by scanning the CAL_DAC first for each channel.

This resulted in a drastic performance gain compared to the legacy software. The S-curve scan over the entire GE1/1 detector, now completes in under 4 minutes, it used to take approximately 3 hours with the legacy software!

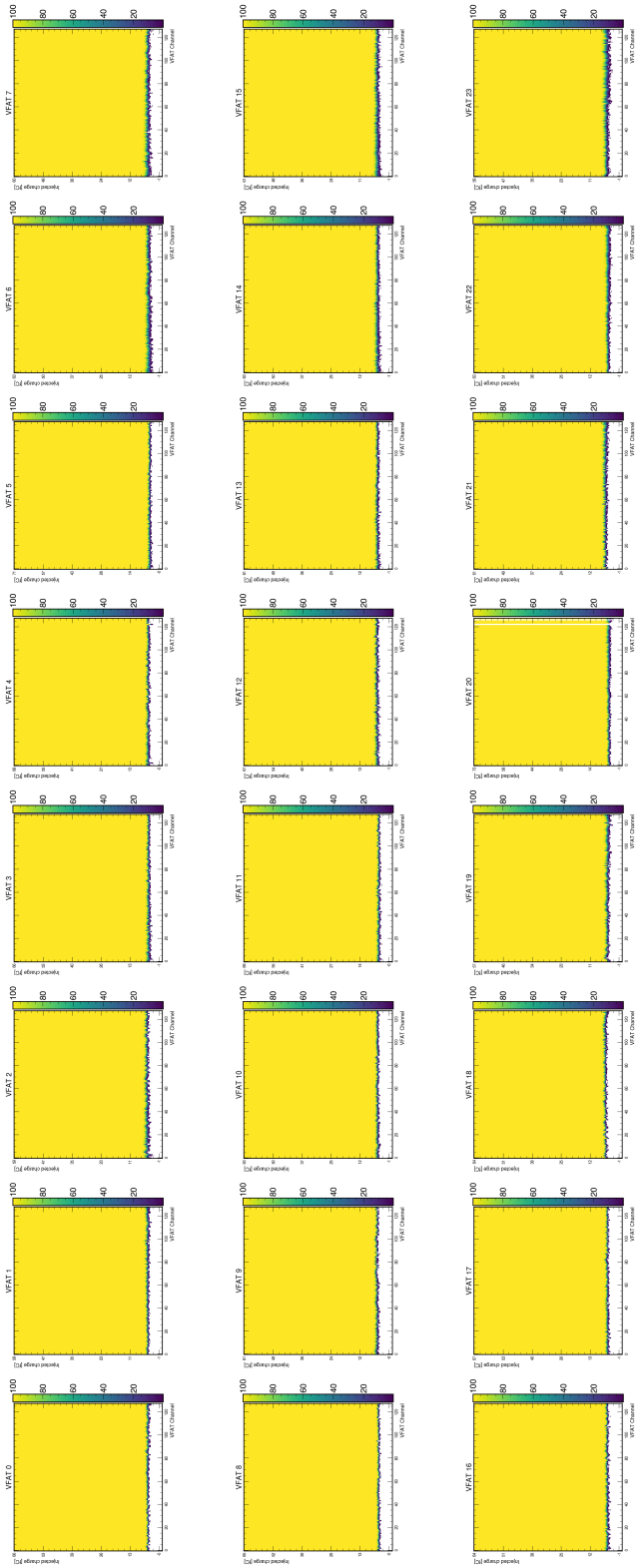
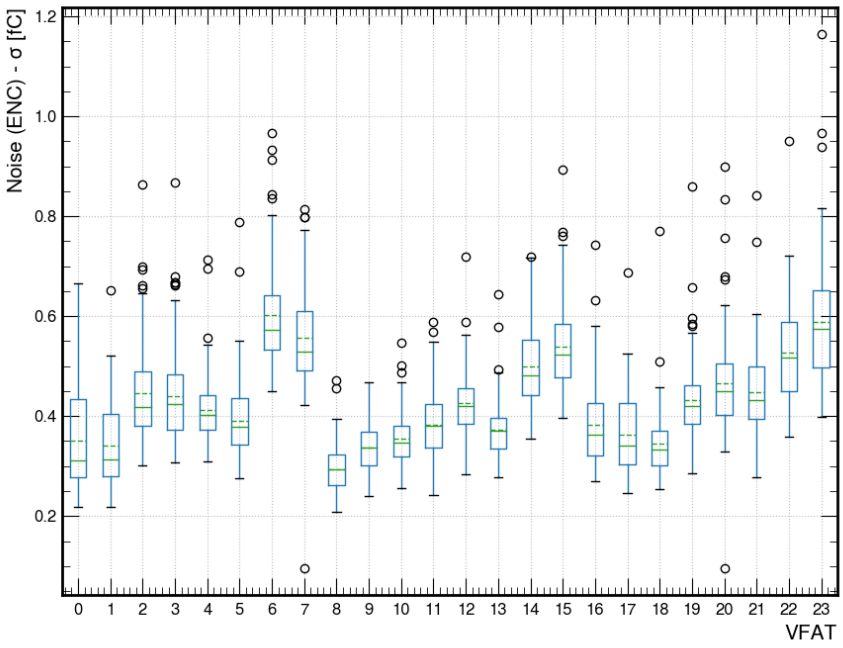
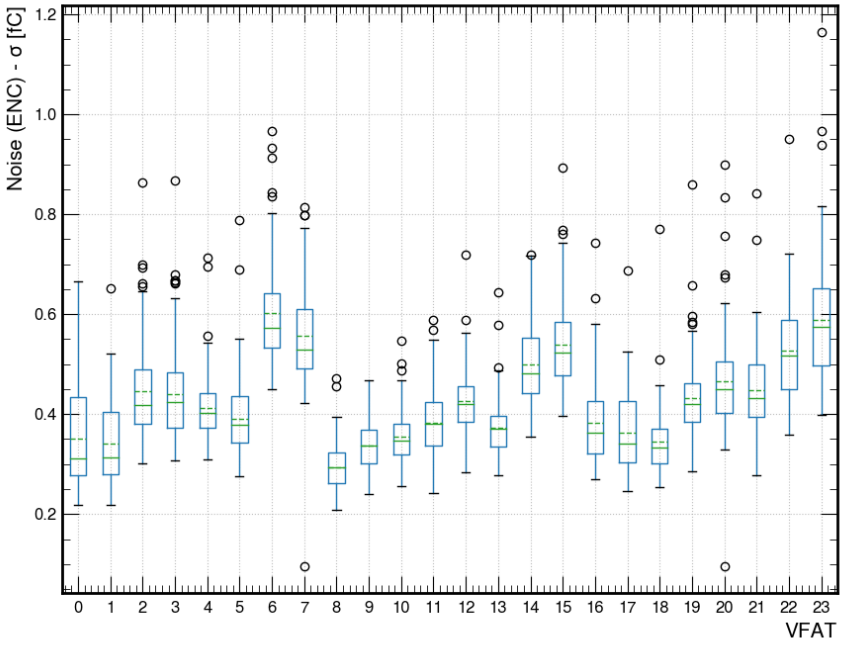


Figure 3.10: Typical S-curve scan summary plots. It shows the raw result of an S-curve scan for all 3072 channels of a GE1/1 detector, grouped by VFAT, as a color map. The color code represents the hits ratio; the X-axis corresponds to the channel number and the Y-axis to the injected charge. The threshold can be inferred from to the position of the transition region between white (0% hit) and yellow (100% hit) colors, while the noise is related to the width of this region.



(a) Threshold



(b) Noise (ENC)

Figure 3.11: Boxplots representing the extracted threshold (a) and ENC (b) values for all 3072 channels of the GE1/1 detector, grouped by VFAT. The boxes span from the 25th to the 75th percentile; the whiskers extend to 1.5 times the interquartile range; and the individual circles represent the outliers. Within each box, the solid line indicates the median whereas the dashed line represents the mean.

Trimming

Because of variability in the VFAT3 ASIC manufacturing process, the response to a signal can slightly differ from one channel to another. To compensate for these imperfections and achieve a more uniform response within a single VFAT3 chip, the discriminator threshold can be adjusted – or trimmed – independently for each channel in a limited range. This is made possible by dedicated circuits present in the VFAT3.

More specifically, each channel includes a 7-bit DAC, referred to as *TRIM_DAC*, which controls a local threshold offset. By modifying the DAC value, the threshold voltage applied to the channel discriminator can be individually tuned.

This tuning procedure, called *trimming*, can be designed to meet two exclusive objectives:

1. Equalize the effective threshold, μ, across all channels.

2. Equalize the "turn-on point" across all channels. The turn-on point represents the input charge at which hits start to be recorded. It is defined as $X = \mu - N\sigma$, where N is an empirically chosen constant, typically 4.

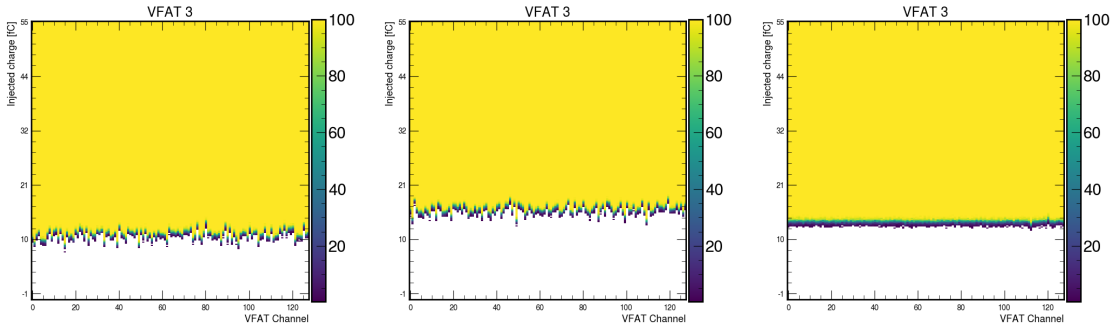
It has been shown in detail in [62] that the second procedure yields the lowest thresholds. This is because the first strategy is limited by the noisiest channels, which drive the global threshold higher than necessary for most channels. On the contrary, aligning the turn-on point equalizes the response to noise from all channels, yielding the lowest thresholds set for all channels. As it leads to the highest detection efficiency, this second option has been adopted as the default trimming strategy for regular operations.

It has also been demonstrated optimal trimming parameters evolve only marginally with the global. Therefore, trimming does not need to be re-performed after each new set of thresholds derived via the S-bit rate scan.

The trimming procedure originally implemented in the legacy software was performed using an iterative approach. An initial S-curve scan was taken with all *TRIM_DAC* registers set to 0. Based on scan results and assuming an ideal DAC behavior, the *TRIM_DAC* values were computed to achieve the chosen equalization objective. A second S-curve scan was then taken to evaluate the effect of the adjustments. This process was repeated until convergence.

It is needless to detail why this procedure could not be applied to the CMS operations, particularly given the long S-curve scan duration. The new software stack implements a more efficient approach. A set of S-curve scans is taken for several fixed *TRIM_DAC* values, typically -63, -32, 0, +32, and +63. This enables a direct calibration of the response of each individual DAC. Once this calibration is established, the optimal trimming parameters can be determined in a single step.

Overall, the procedure is significantly simplified and does not require more effort than performing 5 S-curve scans. Figs. 3.12a and 3.12b represent the S-curves for one VFAT with the *TRIM_DAC* values set to -63 and +63, respectively, before trimming. Figure 3.12c represents the post-trimming S-curves for the same VFAT, showing significant improvement in the uniformity of the response across channels.



(a) (b) (c)

Figure 3.12: Example of S-curves for a single VFAT before and after the trimming procedure. In Figures (a) and (b) show, all channels are configured with the extremum *TRIM_DAC* values, respectively -63 and +63. Figure (c) shows the results after trimming. The response across channels is significantly more uniform that before the trimming procedure.

3.5.5 Threshold scan

The *threshold scan* is conceptually similar to the S-bit rate scan: the arming comparator threshold is swept and the resulting hit rate is recorded. However, unlike the S-bit rate scan, which uses coarse granularity trigger data, the threshold scan relies on the data from the VFAT tracking path, providing access to the full channel-level granularity data. To extract data from the VFAT and check for the presence of noise hits, a periodic trigger signal is used. As with the S-curves, the local readout is used to acquire data at high rate, bypassing the slow "DAQ monitor" firmware module.

Figure 3.13 shows a typical output histogram. The X-axis corresponds to the VFAT channels, the Y-axis to the threshold values, and the color code represents the number of hits recorded for each channel-threshold pair.

This scan has proven particularly useful for detecting VFAT channels disconnected from the readout board due to incorrect assembly. A reduced hit rate on adjacent channels on one side of the VFAT is a clear indication of such defects.

3.5.6 Latency scan

The latency of the VFAT3 is defined as the delay, measured in bunch crossings, between the reception of an external trigger signal (L1A) and the memory location where the corresponding tracking data is stored. In CMS, the Level-1 Trigger determines the events of interest. In standalone systems, this selection is usually performed by scintillating tiles.

The total delay between the passage of the particle through the detector and the arrival of the corresponding trigger signal is influenced by several factors, including the particle time-of-flight, the signal processing delays, the trigger generation delay, and the readout optical fiber propagation delays. As these factors depend on the exact

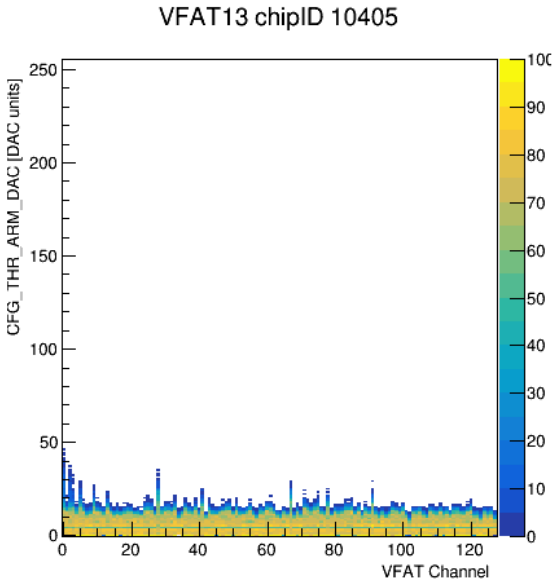


Figure 3.13: Threshold scan during QC7 test (with cooling plate) of the GE1/1-X-S-CERN-0012 chamber.

system, the exact latency is a priori unknown and must be determined empirically for each VFAT. This value is then configured in the VFAT3 through its CFG_LATENCY register, which specifies the position to read from in the VFAT3 SRAM1 circular buffer, where the binary hit data is stored.

In a *latency scan*, the latency parameter is swept across a range of values while the system continuously records data. Crucially, the scan must be completely transparent to ongoing data-taking – pausing CMS triggers is neither allowed nor possible[3]. To this end, the online software flags the data as bad during the transitions between latency points. Moreover, the scan runs cyclically and continues until the end of the run.

Data is readout and stored permanently through the local readout in the standard tracking data format, bypassing any possibly sub-optimal HLT selection. For each latency setting, the number of hits detected in the chambers is analyzed. The optimal latency value corresponds to the peak in the hit distribution, indicating the maximum signal detection efficiency.

Figure 3.14 shows a typical latency scan result for a single VFAT during a test beam campaign. Ideally, the hit distribution would be a rectangular signal, with a width corresponding to the VFAT-configured pulse stretch and a height matching the number of triggers sent per point. In practice, though, the peak is smeared due to detector time resolution, and the overall efficiency is reduced due to detector inefficiencies and limited geometrical acceptance. Section 5.5 discusses latency scan results with real proton-proton collision data in CMS, along with the associated challenges and additional detector effects.

[3]If everything is technically possible in software, some tasks are strongly recommended against.

Finally, fine timing adjustments below the BX resolution are also possible. They can be achieved in two complementary ways:

- By shifting the GBTx clock in steps of ~ 208 ps, which adjusts the detector signal sampling point of all VFATs linked to that GBTx.
- By bit-shifting the VFAT data stream itself, effectively adjusting its 40 MHz sampling clock with a resolution of 3.125 ns.

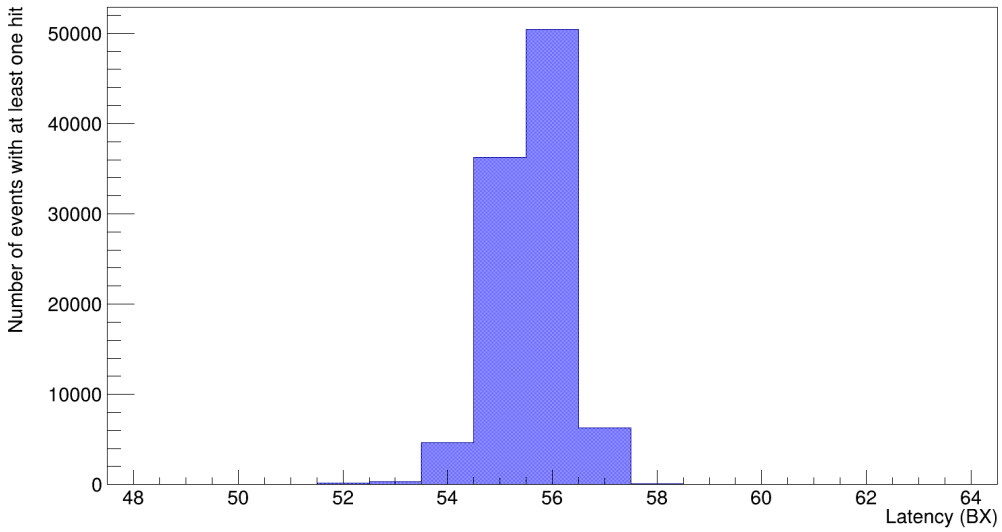


Figure 3.14: Typical output plot of a latency scan for a single VFAT. This specific plot was taken during a test beam campaign in Spring 2022 with a pulse length of 1 BX. The X-axis represent the scanned latency values, whereas the Y-axis represents the number of time the specific VFAT recorded at least one hit on any of its channels.

3.5.7 Scans scalability

One of the key motivations behind the redesign of the GEM online software was the need to scale seamlessly to the full GE1/1 station during Run-3, as well as the larger ME0 and GE2/1 stations anticipated for the GEM Phase-II upgrade. With exception, this requirement guided the reimplementation of the calibration scan routines.

Table 3.1 summarizes the performance improvements obtained on a CTP7-based setup. The measurements are reported per `AMCManager`, as the operations of multiple instances are fully independent and parallelized, leading to perfect scalability. At the time of writing, no further optimizations are under consideration: the current performance is sufficient to ensure smooth and effective operations and additional gains would

come at the cost of lower code clarity and higher complexity. While the algorithm complexity is expected to remain consistent with other back-end boards, the absolute performance could improve when deployed on newer and more powerful back-end boards.

Figure 3.15 shows how the duration of the S-curve scan evolves with the number of VFATs included in the data-taking. This specific scan is highlighted as it has undergone the most extensive optimization effort.

3.6 Monitoring suite

At this stage, all desired control features of the DAQ system have been implemented. This section addresses the other side of the online software: the monitoring. The monitoring suite ensures the operational health of the system, both from the hardware perspective (e.g. temperatures, voltages) and from the data-taking perspective (e.g. link instabilities, VFAT synchronization issues). The following subsections cover these aspects.

3.6.1 Online monitoring system

The online monitoring system is based on the continuous polling of relevant metrics by dedicated xDAQ applications: one `AMCMonitor` instance per CTP7 board, and one `AMC13Monitor` instance per AMC13. These applications are designed to run at all times, even between data-taking runs. Any metric relevant to the user can be included: trigger rates, voltages, temperatures, DAQ status,... Over time, the set of metrics was refined to match the evolving operational needs.

The metrics are organized in groups according to the ”location” in the electronics (backend vs. frontend), functional purpose (e.g. links, DAQ, trigger), and configurability needs. If a group update fails, the refresh of all metrics from that group is skipped. Each group can be independently configured in terms of refreshing period and activation status.

A subset of the metrics is displayed in the `FEDMonitor` web interface, which aggregates data from all monitoring applications associated with the same FED. Figure 3.16 shows part of a typical monitoring page during a MiniDAQ-3 run. This interface is currently implemented using raw JavaScript, but a migration to the `Vue.js` framework is planned to facilitate the maintenance and enhance the integration across the system.

3.6.2 DCS interface

One critical aspect of the monitoring suite is related to the detector safety. The GEM DAQ is required to continuously report the OptoHybrid FPGA core temperature - the highest measurable temperature available on a chamber - to the GEM DCS. If this temperature exceeds a configurable limit (defined per chamber), or if it is not reported for a predefined duration (defaulted to 20 minutes), the DCS will automatically power off the low-voltage as a safety precaution.

Table 3.1: Scalability of the calibration or health scans taken in the CMS GEM project. The numbers are always accounted per `AMCManager` as the operations between managers are fully independent and parallelized. The calibration pulses and cluster mask scans are described further in Sections 5.3 and 5.6.

Scan	Scalability
Latency	N/A
DAC	$\mathcal{O}(0.7\,\mathrm{s/VFAT})$ per 6 bits DAC
	$\mathcal{O}(2.8\,\mathrm{s/VFAT})$ per 8 bits DAC
Calibration pulse*	$\mathcal{O}(0.9\,\mathrm{s/OH})$
Cluster mask*	$\mathcal{O}(30s)$
Sbit-rate	$\mathcal{O}(260s)$
S-curves	$\mathcal{O}(0.8\,\mathrm{s/VFAT})$
Threshold	$\mathcal{O}(0.5\,\mathrm{s/VFAT})$
GBT phase	$\mathcal{O}(30\,\mathrm{s/VFAT})$

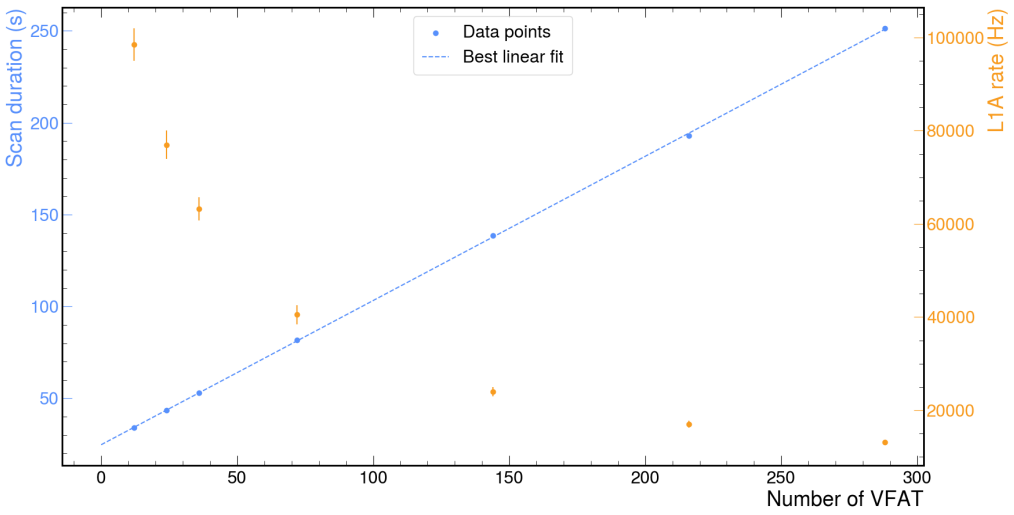
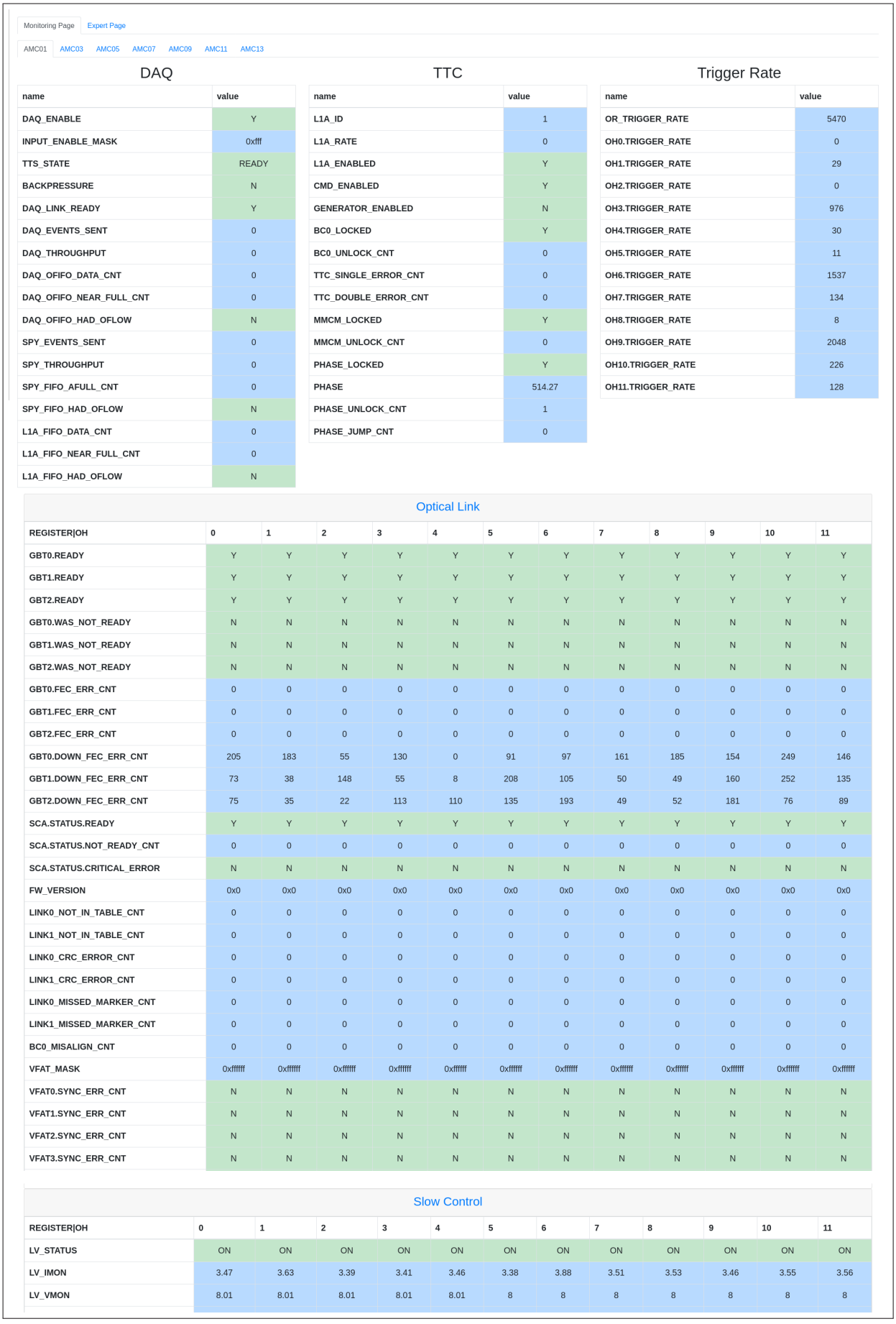


Figure 3.15: Evolution of the duration (blue) and trigger rate (orange) of the S-cruve scan, as function of the number of VFATs included in the data-taking. The scan lasts a constant 24 s, plus 0.8 s for each additional VFAT.



Monitoring Page | Expert Page

AMC01 | AMC03 | AMC05 | AMC07 | AMC09 | AMC11 | AMC13

DAQ

name	value
DAQ_ENABLE	Y
INPUT_ENABLE_MASK	0xfff
TTS_STATE	READY
BACKPRESSURE	N
DAQ_LINK_READY	Y
DAQ_EVENTS_SENT	0
DAQ_THROUGHPUT	0
DAQ_OFIFO_DATA_CNT	0
DAQ_OFIFO_NEAR_FULL_CNT	0
DAQ_OFIFO_HAD_OFLOW	N
SPY_EVENTS_SENT	0
SPY_THROUGHPUT	0
SPY_FIFO_AFULL_CNT	0
SPY_FIFO_HAD_OFLOW	N
L1A_FIFO_DATA_CNT	0
L1A_FIFO_NEAR_FULL_CNT	0
L1A_FIFO_HAD_OFLOW	N

TTC

name	value
L1A_ID	1
L1A_RATE	0
L1A_ENABLED	Y
CMD_ENABLED	Y
GENERATOR_ENABLED	N
BC0_LOCKED	Y
BC0_UNLOCK_CNT	0
TTC_SINGLE_ERROR_CNT	0
TTC_DOUBLE_ERROR_CNT	0
MMCM_LOCKED	Y
MMCM_UNLOCK_CNT	0
PHASE_LOCKED	Y
PHASE	514.27
PHASE_UNLOCK_CNT	1
PHASE_JUMP_CNT	0

Trigger Rate

name	value
OR_TRIGGER_RATE	5470
OH0.TRIGGER_RATE	0
OH1.TRIGGER_RATE	29
OH2.TRIGGER_RATE	0
OH3.TRIGGER_RATE	976
OH4.TRIGGER_RATE	30
OH5.TRIGGER_RATE	11
OH6.TRIGGER_RATE	1537
OH7.TRIGGER_RATE	134
OH8.TRIGGER_RATE	8
OH9.TRIGGER_RATE	2048
OH10.TRIGGER_RATE	226
OH11.TRIGGER_RATE	128

Optical Link

REGISTER|OH	0	1	2	3	4	5	6	7	8	9	10	11
GBT0.READY	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
GBT1.READY	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
GBT2.READY	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
GBT0.WAS_NOT_READY	N	N	N	N	N	N	N	N	N	N	N	N
GBT1.WAS_NOT_READY	N	N	N	N	N	N	N	N	N	N	N	N
GBT2.WAS_NOT_READY	N	N	N	N	N	N	N	N	N	N	N	N
GBT0.FEC_ERR_CNT	0	0	0	0	0	0	0	0	0	0	0	0
GBT1.FEC_ERR_CNT	0	0	0	0	0	0	0	0	0	0	0	0
GBT2.FEC_ERR_CNT	0	0	0	0	0	0	0	0	0	0	0	0
GBT0.DOWN_FEC_ERR_CNT	205	183	55	130	0	91	97	161	185	154	249	146
GBT1.DOWN_FEC_ERR_CNT	73	38	148	55	8	208	105	50	49	160	252	135
GBT2.DOWN_FEC_ERR_CNT	75	35	22	113	110	135	193	49	52	181	76	89
SCA.STATUS.READY	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
SCA.STATUS.NOT_READY_CNT	0	0	0	0	0	0	0	0	0	0	0	0
SCA.STATUS.CRITICAL_ERROR	N	N	N	N	N	N	N	N	N	N	N	N
FW_VERSION	0x0	0x0	0x0	0x0	0x0	0x0	0x0	0x0	0x0	0x0	0x0	0x0
LINK0_NOT_IN_TABLE_CNT	0	0	0	0	0	0	0	0	0	0	0	0
LINK1_NOT_IN_TABLE_CNT	0	0	0	0	0	0	0	0	0	0	0	0
LINK0_CRC_ERROR_CNT	0	0	0	0	0	0	0	0	0	0	0	0
LINK1_CRC_ERROR_CNT	0	0	0	0	0	0	0	0	0	0	0	0
LINK0_MISSED_MARKER_CNT	0	0	0	0	0	0	0	0	0	0	0	0
LINK1_MISSED_MARKER_CNT	0	0	0	0	0	0	0	0	0	0	0	0
BC0_MISALIGN_CNT	0	0	0	0	0	0	0	0	0	0	0	0
VFAT_MASK	0xffffff	0xffffff	0xffffff	0xffffff	0xffffff	0xffffff	0xffffff	0xffffff	0xffffff	0xffffff	0xffffff	0xffffff
VFAT0.SYNC_ERR_CNT	N	N	N	N	N	N	N	N	N	N	N	N
VFAT1.SYNC_ERR_CNT	N	N	N	N	N	N	N	N	N	N	N	N
VFAT2.SYNC_ERR_CNT	N	N	N	N	N	N	N	N	N	N	N	N
VFAT3.SYNC_ERR_CNT	N	N	N	N	N	N	N	N	N	N	N	N

Slow Control

REGISTER|OH	0	1	2	3	4	5	6	7	8	9	10	11
LV_STATUS	ON	ON	ON	ON	ON	ON	ON	ON	ON	ON	ON	ON
LV_IMON	3.47	3.63	3.39	3.41	3.46	3.38	3.88	3.51	3.53	3.46	3.55	3.56
LV_VMON	8.01	8.01	8.01	8.01	8.01	8	8	8	8	8	8	8

Figure 3.16: Monitoring suite web interface for one `AMCMonitor` during a MiniDAQ-3 run.

The implemented solution is based on the Distributed Information Management (DIM) framework, developed and thoroughly used at CERN [60]. Implemented following a publish–subscribe pattern, it provides a communication platform between loosely coupled systems for efficient data exchange in distributed systems. Any client can read the value of a service and subscribe to it to receive updates.

The `AMCMonitor` application was then extended to expose one DIM service per OptoHybrid, publishing all available temperature sensor values. This fully satisfies the monitoring requirements imposed on the GEM system. In the other direction, relevant values - namely the low- and high-voltage statuses - are also exposed via DIM by the DCS and integrated into the online monitoring suite. This provides a unified overview of all parameters relevant to the data-taking in a single interface.

3.6.3 Archiving

Rather than developing a custom solution from scratch, the metrics archiving was left to Commercial-Off-The-Shelf (COTS) applications.

InfluxDB is used to store the metrics. It is a high-performance Time Series DataBase (TSDB) [71]. Compared to standard SQL databases, it is particularly well suited to the ingestion, storage, and retrieval of time series data thanks to optimized protocols, indexing, and storage and compression engines. It was chosen over alternative tools because of its native support for *push-based* monitoring. In contrast, most of the other COTS monitoring systems periodically *pull* the data for well-known and configured endpoints. Therefore, fire-and-forget systems can more trivially be implemented with InfluxDB, not requiring any new configuration of the database server for new data sources.

Grafana is used for visualization. It is an open-source web platform that enables the real-time visualization of data originating from a wide variety of sources, among which InfluxDB [63]. Users can easily create dynamic and interactive dashboards, providing custom views of time series data for efficient monitoring and analysis. Additionally, alerting mechanisms can be configured to notify the users in case of abnormal situations.

The current archiving workflow is presented in Figure 3.17. It uses a custom Python logger, which regularly fetches the metrics exposed by the `AMCMonitor` applications. After reformatting to meet the InfluxDB standards, the metrics are pushed and archived into the TSDB. The visualization is handled by Grafana, where dashboards and panels are created following the operation team's needs.

Despite the advanced compression algorithms implemented in InfluxDB, the most efficient solution remains to send no useless data. Therefore, de-duplication is implemented by the logger: values unchanged between two polling iterations are not pushed. However, to preserve readability in the visualizations, a minimum update frequency of one point every 10 minutes is guaranteed.

A future improvement consists of shifting toward a fully push-based model, fully exploiting the potential of InfluxDB. Rather than regularly polling the `AMCMonitor` applications via an external program, it envisioned for the `AMCMonitor` to send the metrics directly to InfluxDB.

During data-taking at full rate, an average of 6GB worth of monitoring data is collected and saved to disk each day. During idle periods, this amount drops to less than 1GB/day. To manage long-term storage, the retention period for all raw metrics has been limited to 1 year, using InfluxDB *shard* data organization. On the contrary, summary metrics or metrics of particular interest are kept forever.

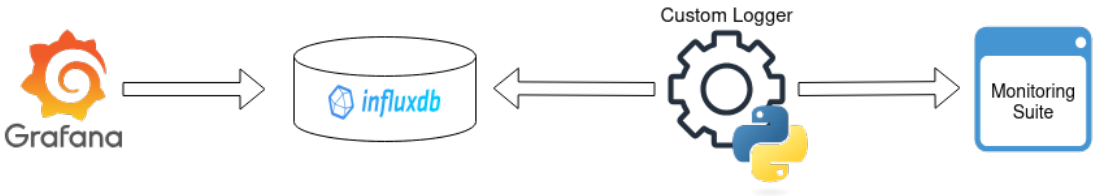


Figure 3.17: Architecture of the metrics archiving.

3.6.4 Contextual information

While DAQ-related metrics are essential for diagnosing data-taking problems, contextual information greatly helps this process. To that end, dedicated *bridges* were developed to enrich Grafana visualizations with additional information.

The first bridge interfaces with the CMS DCS condition database, providing access to all parameters monitored and archived by the GEM DCS, such as the gas, low-voltage, high-voltage, and temperature metrics. Figure 3.18 illustrates the architecture implemented. Grafana issues queries formatted in JSON, which the bridge application translates into a set of SQL queries directed to the OracleSQL database containing the DCS conditions. The results are then converted back to JSON and returned to Grafana. Because the Oracle database contains a very large amount of data, shared by all CMS systems, queries over long time intervals can be slow (tens of seconds). Therefore, a cache can optionally be used to drastically speed up the performance of frequent requests. This cache, based on the Redis in-memory database [91], provides access to the results of frequent queries within milliseconds. All bridge applications are deployed in the general-purpose OpenShift instance provided by the CERN IT, for simplified maintenance and improved fault tolerance.

A second bridge interfaces with the CMS OMS aggregator API [111, 11]. Following an architecture similar to the DCS bridge (excluding the Redis cache), it gives access to metadata such as the list of runs and fills. It is heavily extensible would new parameters of interest appear.

Thanks to these developments, the Grafana monitoring interface has evolved from a simple DAQ status visualization tool into a full GEM operations cockpit.

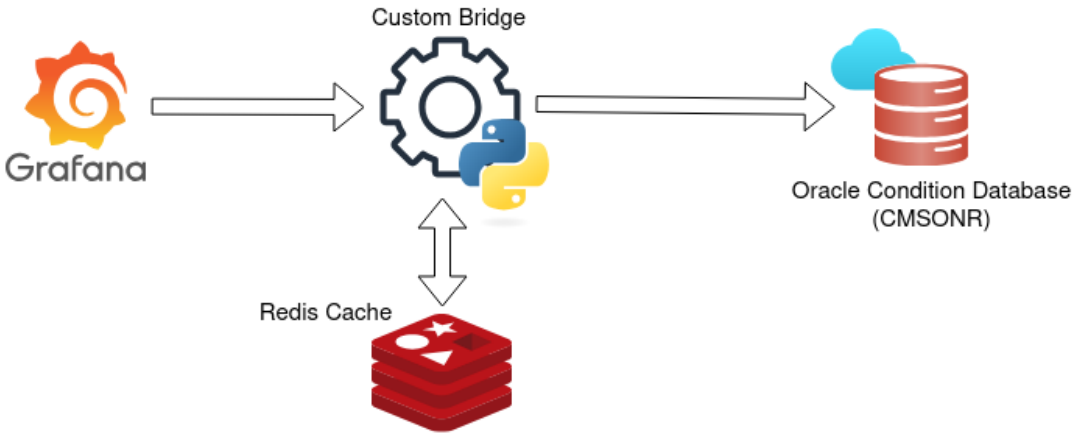


Figure 3.18: Architecture of the CMS DCS to Grafana bridge. The Redis cache provides large query speed up for frequent queries.

3.7 Automatic actions

The GEM data acquisition system is a dynamic system, like any other CMS sub-system. Due to various reasons elaborated in Section 5.2, communication with certain parts of the detector can occasionally become unstable or even impossible. More generally, any abnormal condition must be detected, and the appropriate corrective action taken to guarantee the quality of the data.

The monitoring system described in the previous section is conceived to present the current and past status of the system, but it cannot trigger preventive or corrective actions. This is where *Automasker* comes into play – a system embedded into the `AMCManager` and responsible for taking automatic actions based on the detector's status.

3.7.1 Automatic masking

The primary function of the *Automasker* is to dynamically mask faulty parts of the hardware. This process complements the static masking configured by the operator and targets components that are identified as problematic during the operations. Its design is driven by two main principles:

- **Disable unreliable components**, whose configuration or status cannot be reliably asserted to prevent arbitrary data from being sent downstream, either in the Level-1 Trigger or in the DAQ.
- **Avoid any impact on a working part of the detector**, either in GEM or in CMS; thus ensuring the continuity and integrity of the data-taking.

The first round of dynamic masking is performed during the configuration sequence. Any component that cannot be communicated with or that may be not properly configured is excluded from the upcoming data-taking, until the next re-configuration.

The second round of dynamic masking is performed when the `AMCManager` is in the *Running* state. The status of the front-end links is periodically polled and a masking action is triggered in two situations:

1. If a GBT link loses lock, and even if it re-establishes almost immediately, the linked VFAT are almost certainly out-of-sync with the rest of the data-taking or misconfigured.

2. If a VFAT eLink is no longer seen as synchronized, it will not report any reliable data.

No masking action is taken for components that are only out-of-sync with respect to the rest of the data-taking, as they do not pose a threat to the overall data-taking. Moreover, periodic TTC ReSyncs are likely to recover the situation.

As a general rule of the *Automasker*, whenever a component is masked, all downstream components are masked as well. In GE1/1, losing GBT #0 results in masking the whole chamber, including the trigger links. Conversely, losing GBT #1 or #2 masks "only" the 9 VFATs associated with each link.

Finally, any masked component is saved in the GEM DAQ data stream on a per-event basis for further analysis. As outlined in Section 5.7, this feature plays a critical role in ensuring reproducible efficiency measurements.

Overall, the masking mechanism implementation is very robust, yielding to no luminosity lost due to the issues within the GEM sub-system over the last years!

3.7.2 Automatic recovery

While the *Automasker* reliably masks faulty parts of the detector, a corresponding recovery mechanism is required to minimize the impact on the overall GEM detector performance.

A notable example is the following: during collision runs at nominal luminosity mid-2024, groups of 12 chambers - each connected to the same CTP7 - were suddenly powered off by the GEM DCS in nearly every run. This behavior was quickly traced to the automatic safety mechanism described in Section 3.6, which powers off the low-voltage after 20 minutes without receiving updates on the FPGA core temperature.

Post-mortem investigations showed that updates from the monitoring group including the core temperature were abruptly failing, triggering the safety response. However, the very narrow time window during the issue could be observed live made the detailed diagnostic difficult. Eventually, clear evidence of abnormal OptoHybrid firmware behavior was observed: inconsistent register values, wrong address decoding, access errors, etc.

Curiously, this issue was never observed in previous years, even under similar luminosity conditions. It was eventually discovered that a now-fixed bug in the GEM online monitoring system would result in publishing stale values to the DIM service used for the DCS communication.

Very quickly the root cause of the erratic behavior was identified as a Single Event Upset (SEU) occurring within the OptoHybrid FPGA.

An SEU consists of the change of state of a memory cell caused by ionizing radiation passing through the substrate of a silicon device. Unlike Single Event Transients (SETs), which are temporary, the change caused by an SEU is permanent until the affected memory cell is reconfigured or power-cycled. Although SEUs do not physically damage the device, the resulting bit-flip can affect its behavior. In the case of an FPGA, any change in configuration memory will result in unpredictable effects on its firmware.

At that time, no proper SEU correction, monitoring, or counting was available in the GE1/1 OptoHybrid FPGA firmware. Xilinx FPGAs such as those used in the GEM project support SEU mitigation techniques. One consists of triplicating the most critical logic, assuming that no two replicas would be impacted simultaneously; another one consists of scrubbing the configuration memory in the background, and correcting found errors, if any. Due to resource constraints, the latter option was selected. The embedded scrubbing logic can correct single-bit errors and detect (but not correct) double-bit errors within each configuration frame.

This example highlighted the need for a recovery mechanism, especially for the most radiation-sensitive components. Even if overall CMS data-taking integrity remained intact, the frequent loss of entire chamber groups was unacceptable.

CMS provides several central recovery mechanisms:

- TTC ReSync sequences
- TTC HardReset sequences
- RCMS SoftErrorRecovery commands

The first two are issued via the TCDS system (see Section 1.2.5), while the last one is handled through RCMS. All these can triggered either automatically or manually. The GEM sub-system opportunistically reacts only to the ReSync sequences, which propagated to all components, but never request any by itself.

The main drawback of those mechanisms is the creation of CMS-central deadtime or downtime for, respectively, the TCDS and RCMS sequences. For this reason, a more subtle, localized recovery mechanism was implemented within GEM. It avoids global event loss by recovering only the parts of the detector affected by an uncorrectable SEU. More specifically, when an uncorrectable SEU is detected by the FPGA scrubbing routine, the *Automasker* takes the following actions:

1. Masks the affected OptoHybrid from the Level-1 Trigger.
2. Reprograms and reconfigures its FPGA from scratch, clearing the error.
3. Re-enable it in the L1T.

The entire recovery sequence takes less than one second and is completely transparent to the ongoing data-taking.

This mechanism has only been introduced recently, and its behavior and performance are still under careful evaluation. As an example, after initial debugging, a total of 314 recoveries were successfully completed in a single 10h:12m:16s long run, during which $752.78\,\mathrm{pb}^{-1}$ of integrated luminosity was delivered.

Looking back, unexplained corruption of the GE1/1 trigger primitives received at the OTMB and EMTF had been reported. Typically, one or a few strips appeared to fire continuously – a condition that VFAT3 should not produce. These issues were originally suspected to be linked to communication instabilities, but no clear correlation could be established between unstable chambers and those sending corrupted data. However, since the introduction of the automatic recovery sequence upon non-correctable SEUs, no such errors have been observed. This suggests that the observed corruption may have been caused by SEUs.

Looking forward, the goal is to extend this recovery approach past the OptoHybrid FPGA SEUs to address any recoverable failure mode affecting the GEM front-end during data-taking.

3.8 Summary

The GE1/1 slice test left the GEM project with a functional DAQ system, adequate for small-scale setups and quality control, but not suitable for large-scale systems such as the full GE1/1 station, particularly on its online software component. This motivated a redevelopment, not only to support GE1/1 but also to address the future needs of the GEM project.

The software architecture presented here is based on a few core principles: reducing the number of repositories to simplify the development; unifying the codebase by abstracting hardware differences as low as possible in the software stack; ensuring robustness, particularly against communication errors; and guaranteeing readiness for both Phase-II GEM and non-GEM systems through design flexibility.

The back-end services, running on the onboard CPU of the back-end board, serve as an accelerator system. Through an original and robust C++11-based approach, the RPC services offload the higher-level software from frequent and low-level hardware interactions, fully leveraging the specifies of the GEM hardware.

The control and monitoring applications, loosely based on the xDAQ and RCMS frameworks, provide web-based tools to the end-users. The control applications drive the data-taking through well-defined status implemented via an FSM. The calibration suite offers all scans necessary to optimize the data-taking performance or assess the health of the detector in a robust and scalable fashion. The monitoring suite provides all metrics required to understand the detector status, both in real-time and post-mortem. It is complemented by contextual information coming from the DCS database and CMS OMS API available in the interface. The local readout features a high-throughput readout path independent from the central CMS DAQ, for usage during calibrations, troubleshooting, and on test stands.

Finally, automatic actions ensure that only the functional and properly configured parts of the detector are included in the data-taking. This both protects CMS from potentially disruptive behaviors and GEM from impacting working parts of the detector.

Over the past years, the online software has proven extremely successful during GE1/1 operations in CMS – with no luminosity loss –, as well as Phase-II test beam campaigns and quality control procedures. Nevertheless, continuous improvements are always pursued, not only to enhance the detector performance and user experience but also to ensure its scalability and maintainability across a broad range of hardware throughout the system's expected decade-long lifetime.

Chapter 4

Quality control

Before installation in CMS, each GE1/1 detector and each of its components must pass a thorough quality control (QC) in order to assert their characteristics, performances, and longevity. This is achieved in a series of progressively more stringent steps, applied to larger and larger parts of the final detector throughout the assembly process.

In this chapter, we first present the overall structure of the GE1/1 QC. We then focus on the tests performed at ULB regarding the GEB and the OptoHybrid board validation. Finally, we conclude with the lessons learned during the GE1/1 electronics QC at ULB, including suggestions of enhancements for the future CMS GEM stations.

4.1 Overview

As the detector parts production and detector assembly are distributed amongst multiple institutes across the world, so is the QC process. Figure 4.1 depicts the overall quality control flow for the GE1/1 project. A few logical groups can be identified: the tests related solely to the GEM chamber, the tests related exclusively to its electronics, and finally the integration and validation tests after the full detector assembly.

The QC1 to QC5 steps focus on the GEM chamber itself, without integration with its final readout electronics. A comprehensive description of all tests, their motivation, and typical results are available in [4] and [24], which we summarize below.

In parallel, the front-end electronics go through their own QC processes. The GEB and OptoHybrid board testing performed at ULB is detailed in further sections. The main goal is to ensure a smooth assembly without worrying about connectivity or mechanical issues when the GEB and OH are mounted on the chambers at CERN. Regarding the VFAT3 quality control and characterization procedures at CERN, a comprehensive discussion of the procedure and the results can be found here [2].

Finally, the QC6 to QC8 steps are performed centrally at CERN by the GEM production group with the help of shifters [62]. QC6 is used as an acceptance for the chambers delivered to CERN. QC7 and QC8 are run on GEM chambers with their full electronics installed and constitute the final qualification round.

Although very linear on paper, the QC process can work "backward" at times. In the event that a chamber or one of its components would be found problematic at a later stage, it would be sent back for investigations, potential fixes, and revalidation at an earlier QC step. Typical examples include VFAT damages occurring during QC8, or, as we will see later, communications problems with the GEB or OH when performing QC7.

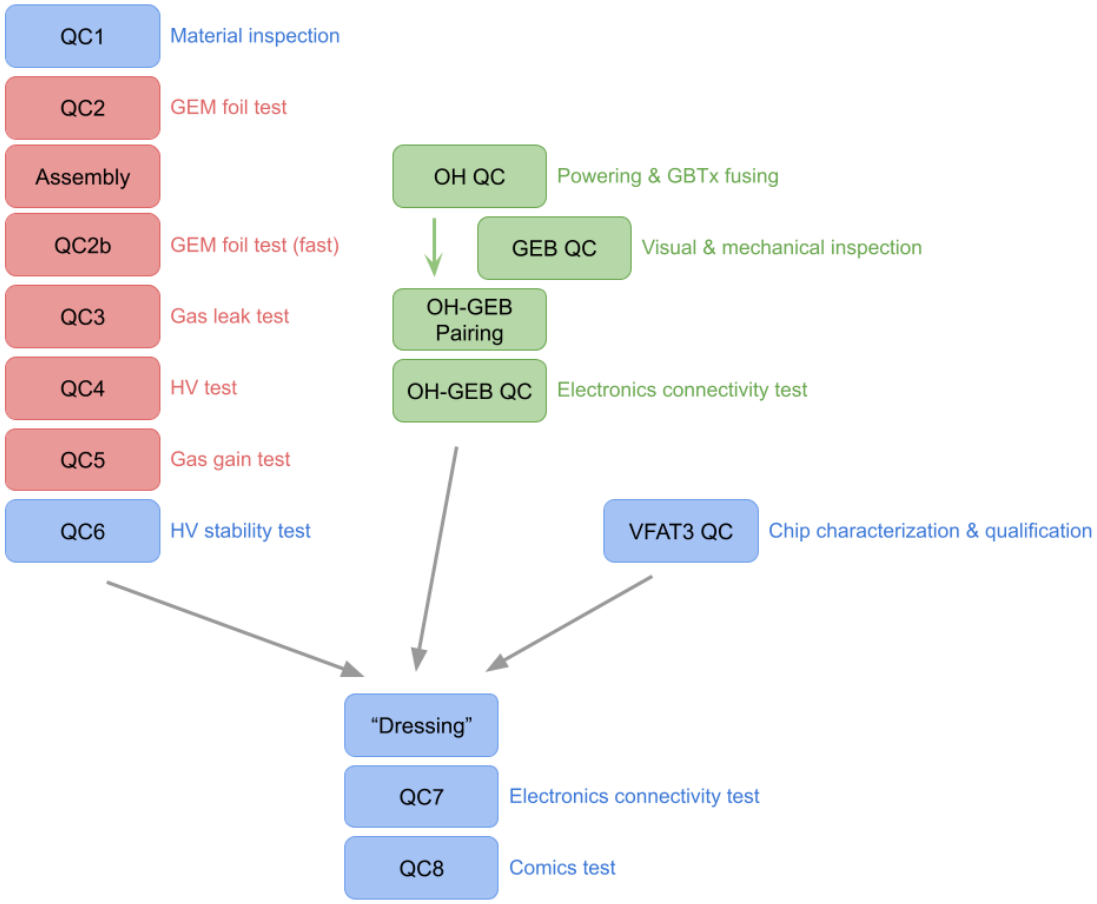


Figure 4.1: GE1/1 quality control flow summary. The steps in blue are performed at CERN, in green at ULB, and in red at the various detector production sites.

QC1 The quality control starts with the inspection of all components that will compose the main structure of the chamber: readout and drift boards, internal and external frames, O-rings,... An emphasis is put on verifying compliance with the specifications, particularly regarding the sizes, and the planarity of the PCB.

QC2a The GEM foils, as the amplification core of the chamber, are tested not only optically, but also electrically. A sane foil, free of defects, dust, or chemical contamination must present a resistance above 60 GΩ in an environment with relative humidity below 50%. This is ensured before starting the assembly procedure and then multiple times during the rest of the assembly via a Giga-Ohm meter applying a voltage of 550 V across the foils.

QC2b This step consists of the long-term counterpart of QC2a and aims at detecting minor issues that could affect the long-term stability of the chamber. This is achieved by monitoring the behavior of the HV current when applying 600 V across the foil in a pure nitrogen atmosphere. Each foil is required to leak less than 1 nA and experience less than 3 discharges per 5 hours.

QC3 Once the chamber is fully assembled, it is critical to ensure the gas tightness of its volume to avoid any leak or contamination by external pollutants, which would deteriorate the amplification process, degrade the detector's performance, and fasten its aging. The maximum accepted gas leak rate was set to 1% of the total incoming flow for an over-pressure of 25 mbar. While measuring such a low leak rate is experimentally very challenging, the internal pressure loss can instead be measured when both the gas inlet and outlet are closed. The time constant required to accept a GE1/1 chamber is then 3.04 h.

QC4 This QC step aims at measuring the current-voltage curve, also known as IV curve, as well as quantifying the spurious, or intrinsic noise, signal rate. Such signals are not caused by ionizing particles, but arise from coronal discharges, which are normal during the operation of MPGD detectors at high gain. In order to distinguish the effect of ionizing particles from spurious signals, the detectors are flushed in pure CO_2. During this step, the current drawn by the detector and the spurious signal rate are measured as a function of the applied voltage. A GE1/1 chamber is validated if the rate of spurious signal does not exceed 100 Hz in normal operating conditions, which is negligible with respect to the background expected in CMS.

QC5a It was determined that an effective gas gain of $\left(2 \times 10^4\right) \pm 37\%$ would allow to reach both an efficiency above 97% and a time resolution better than 8 ns [4]. The aim of QC5a is thus to measure the chamber gain by comparing the output current induced on the readout electrode after amplification with the primary current induced in the drift gap by the ionizing particles. A chamber is considered validated if the nominal gain can be reached for HV values far enough from the breakdown voltage that would trigger potentially damaging discharges.

QC5b While QC5a focuses on measuring the effective gas gain as a function of the high voltage, the QC5b step measures the uniformity of this parameter over the chamber. Combining the effects coming from the variation of the GEM holes geometry – affecting the foil gain – and the electric field uniformity – affecting the foil transparency –, the selection limit was set to 37%. It must also be noted that, due to the limitations in the signal amplitude accepted by the APV15 ASIC [58] used in the QC5 readout system, the detector must be operated at very low gains of the order of 500-600.

QC6 Once a chamber is fully assembled, it gets shipped to CERN. During transportation, it could suffer damages, but also mechanical stress and vibrations that would release dust on the foils. In order to ensure the high-voltage stability of the detector and clean it from any impurities, QC6 has been designed to perform the so-called *HV training*. Each electrode is progressively brought to the desired setting in pure CO_2 - to prevent any signal amplification - with the aim of vaporizing any impurity. Only if the chamber remains stable at the chosen working points, it is deemed validated. The procedure will be repeated multiple times throughout the detector's lifetime, including regularly during operations with the final gas mixture.

QC7 Once a GE1/1 chamber is fully validated at the end of QC6, it goes through the "dressing" procedure: the front-end electronics is installed on the bare chamber, followed by the cooling plate, and finally the chimney. A series of scans, described later, is performed to ensure proper connection of all parts, reliable communication between all components, and an acceptable level of noise.

QC8 constitutes the final qualification step the detectors must undergo before installation. The five-row, three-column stand shown in Figure 4.2 is a small-scale replica of the final CMS system able to host a maximum of 15 GE1/1 super-chambers. The same low-voltage, high-voltage, and DAQ systems are set up, along with similar gas and cooling services. The efficiency of each chamber is carefully measured with cosmic muon data: only those detectors with an efficiency above 95% are qualified to be installed into CMS. The detectors which do not pass the requirement are removed from the stand for debugging and repair. They will go through QC8 a second time.

4.2 GEB quality control

Despite the visual, electrical, and mechanical quality controls performed at production by the manufacturer and before shipment by the responsible institute, a significant number of problems fell through, delaying the OptoHybrid quality control at ULB, and preventing a smooth detector assembly and QC7 at CERN. And despite improvements batch after batch during production following our feedback, it was decided to impose an additional GEB-only quality control at ULB, fixing the simplest issues in order to

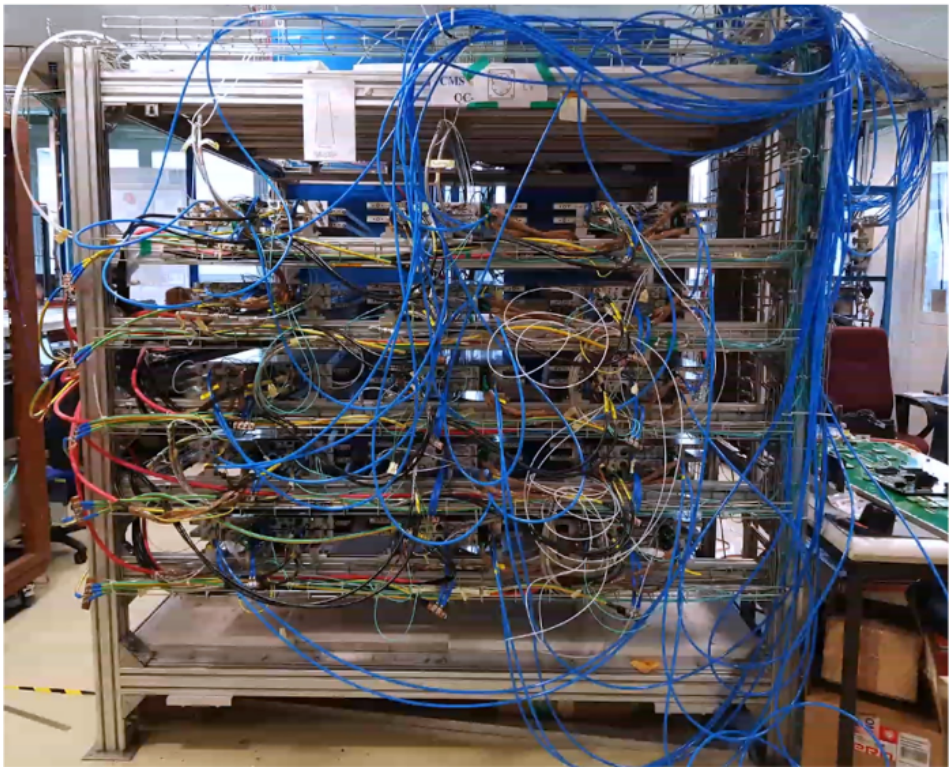


Figure 4.2: Picture of the GEM hodoscope located in the laboratory at CERN and used for the final QC step, QC8.

ensure a sufficient yield. Figure 4.3 displays the typical issues found after delivery that the new routine aimed at detecting, as well as some of the solutions put in place.

The GEB-only quality control at ULB at the end of production consisted of the following steps:

1. Visual inspection
2. Consolidation of the standoffs soldering
3. Verification of the standoffs soldering by screwing every screw up to a defined torque
4. Cleaning of each connector and the board with flux remover, isopropyl alcohol, and pressurized air
5. Checking the power delivery
6. Taking the PCB bending measurements

This order evolved with experience and was specifically chosen to include the slowest steps after the ones with high failure or damage rate. For example, the standoff soldering

can damage some connectors, or the cleaning procedure is extremely time-consuming. In case of doubt at any step, the GEB was put aside in order to increase the throughput. Would we have run out of parts, the stash of early discarded components remained available.

The rationale behind each of the steps is as follows. The visual inspection enables a quick assessment of the packaging, flatness, cleanliness, and soldering quality. It was decided that the height at the periphery of the GEB had to remain below 5 mm to ensure a smooth detector assembly. Any GEB bent past this limit was immediately discarded (Figure 4.3a). The precise board bending measurements themselves were however only taken very last due to their tediousness.

Some transportation damages such as broken fixation eyelets were for the largest part ignored during the visual inspection since they did not pose a problem during assembly: once the chimney is screwed onto the detector, the GEB PCB is kept in place.

The last point of interest during the visual inspection was the detection of any suspicious soldering or rework. Experience has shown that reworked connectors often lead to communication or VFAT biasing issues during the combined GEB & OH testing.

The standoffs constitute a weakness of the GE1/1 electronics design. In addition to mechanically coupling multiple parts (narrow GEB, wide GEB, OH), they are used to feed low-voltage power from the GEB to the OptoHybrid. If any of those standoffs breaks, the functionality of the detector is compromised and the affected GEB must be replaced. To avoid such faith, each standoff was tested with a torque-limiting screwdriver. The constraints on the torque have been continuously increased to reach at the end of the QC, 0.5 Nm, 0.35 Nm, and 0.35 Nm for the OH, VFAT, and FEAST screws, respectively. Nevertheless, very few standoffs were initially able to sustain such constraints: consolidation of their soldering was deemed required. It is however to be noted that no validation could be made on the force required to pull or tilt a standoff – those actions remain the most damaging during assembly of the electronics and cooling on the chamber.

Most of the GEB arrived extremely dirty at ULB with flux residues, stains, and even **food** in the connectors. If left in that state, communication troubles were almost assured. This alone justifies the implementation of the thorough cleaning procedure before connectivity testing.

Finally, before installing any VFAT3 or OptoHybrid, we ensured the continuity of the power lines. This test served two purposes. For one, it helped ensure that any human error in the installation of the FEAST was caught in time before damaging the downstream electronics. For two, it enabled us to catch missing components in the powering circuits (Figure 4.3c).

It took time to find the right QC protocol and tune it to smooth the QC7 step. While almost all GEB were at ULB beginning of 2019; by the end of May 2019, only 35 narrow + wide pairs were qualified.

The GEB remains the component that took the longest time to be qualified due to the large amount of interventions, in particular regarding the mechanical re-enforcement and testing, and the cleaning procedure. The QC of a good GEB, without debugging, took about 2.5 h with ~ 1.5 h dedicated to the standoffs.

If a relatively low number ($\mathcal{O}(15)$) of parts has been discarded at CERN during QC7, many more have been at ULB ($\mathcal{O}(60)$), which shows the effectiveness of the procedure put in place. It must also be noted that some GEB got irremediably damaged by too light or too tight packaging, even before starting any QC procedure at ULB.

4.3 OptoHybrid quality control

The quality control of the OptoHybrid in standalone mode was quite limited. While it was originally thought to simply power on the board via an external power supply, fuse the GBTx, ensure the back-board to GBT communication, and ship to CERN, the need for enhanced GEB testing changed the plans.

In addition to the QC procedure itself, the main tasks consisted of finalizing the production with the addition of wires to enable GBTx fusing and implementing some modifications where required. The second task originated from two small assembly changes decided in June 2019:

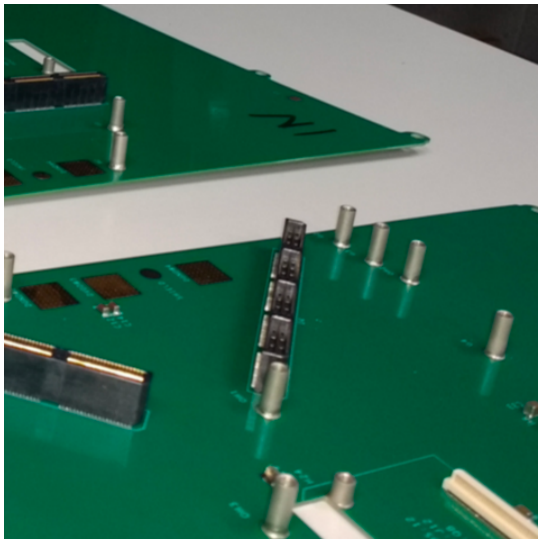
- The PCB gold ring that was in contact with the tin-plated brass standoffs had to be covered with tin to avoid the risk of galvanic oxidation (Figure 4.4). That was required for long-term operations since those standoffs carry critical power. This was later enhanced on GE2/1 where the standoffs are only used for mechanical support.
- The SCA-based temperature monitoring system had to be adjusted to improve the accuracy of the measurements. This change is covered in depth in the next section.

As those changes were implemented late in the production, some boards had to be refurbished in the late stages of the detector assembly, leading to delays. Therefore, the least critical changes were not systematically implemented on all OH.

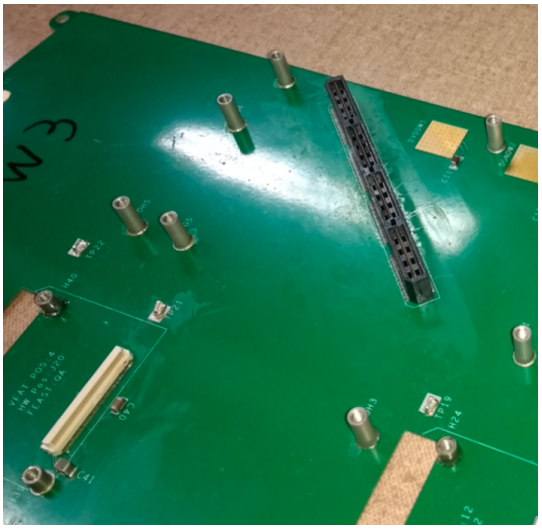
Very few ($\mathcal{O}(5)$) boards were found problematic during the initial tests at ULB. However, several ($\mathcal{O}(5)$) SAMTEC connectors got damaged during assembly at QC7. In most of the cases, repairs were attempted either locally at ULB, or by the replacement of the problematic part. The testing and repair operations only stopped once the total number of fully functional boards delivered at CERN was enough to guarantee the full GE1/1 production. Table 4.1 summarizes the different repairs attempted and their eventual success. Note that the unsuccessful VTRx replacement corresponds to the failed GBTx replacement.

4.3.1 Temperature sensors

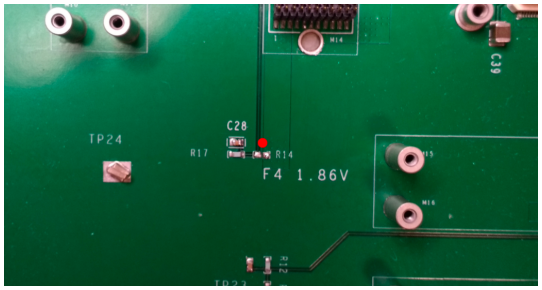
As the temperatures were measured for the first time via the PT100 sensor on the original version of the OptoHybrid board, a large bias compared to the expected temperature



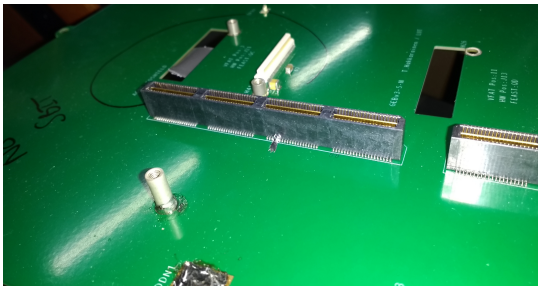
(a) Bended GEB PCB.



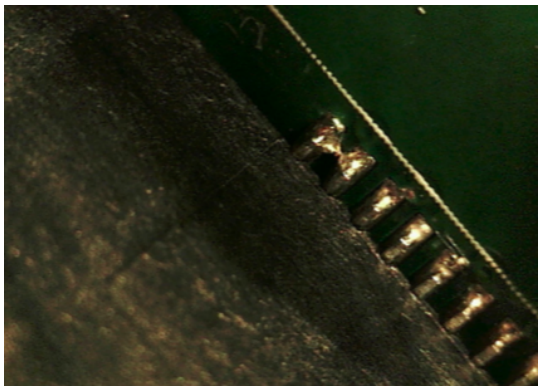
(b) Very dirty board.



(c) Missing resistor and standoff.



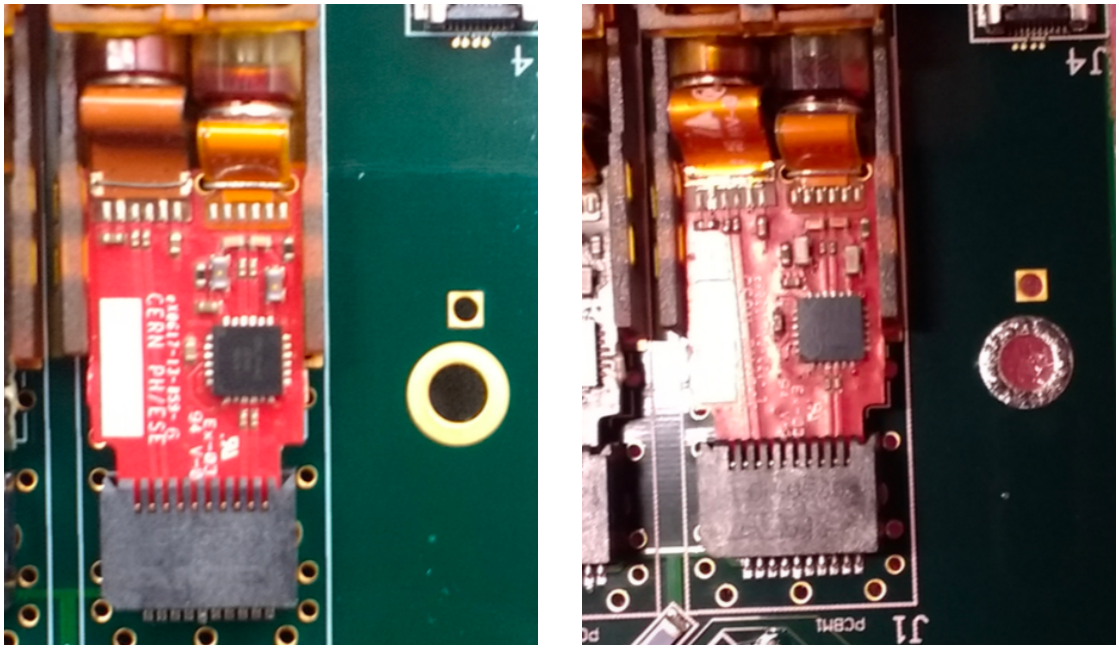
(d) Badly soldered standoffs and SAMTEC connector.



(e) Shorted SAMTEC connector pins (before fix).

(f) Shorted SAMTEC connector pins (after fix).

Figure 4.3: Miscellaneous issues discovered during the GEB QC (a, b, c, d, e) and possible fixes (f).



(a) (b)

Figure 4.4: The OptoHybrid standoff pads before (a) and after (b) applying tin.

Table 4.1: Summary of the repairs attempted on the OptoHybrid boards.

Intervention	Successes	Failures
VTRx replacement	5	1
GBTx replacement	3	1
SAMTEC connector replacement	4	1

was observed on some of the boards. Figure 4.5 compares temperature measurements of a sensor located on the OptoHybrid PCB just below the programmed Virtex-6 FPGA.

This phenomenon was traced back to three complementary factors:

1. The SCA ADC current source is uncalibrated and only provides *nominally* 100 µA;
2. The PT100 sensor does not match appropriately the characteristics of the SCA ADC;
3. The resistance of the routing traces on the PCB is large with respect to the PT100 resistance.

The first factor accounts for a bias of up 30 °C in the most extreme cases. Figure 4.6 shows the distribution of the SCA ADC current for all GE1/1 chambers installed in

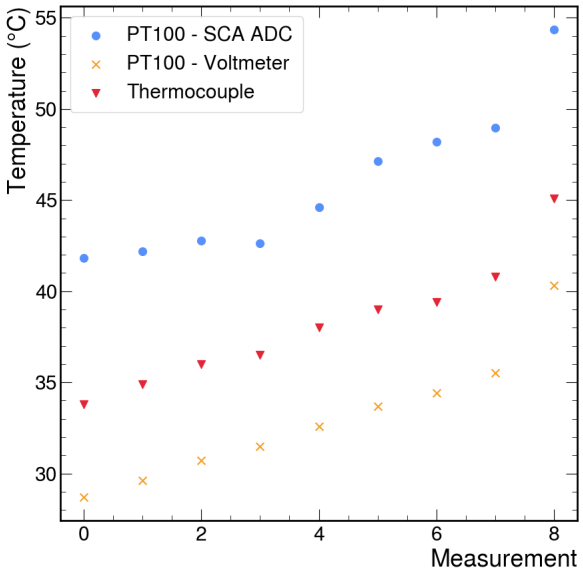


Figure 4.5: FPGA external temperature measured via different probes. In blue, via the on-board PT100 readout by the SCA ADC (averaging 250 measurements); in red, via a thermocouple located on top of the PT100; in yellow, via the on-board PT100 measured via a voltmeter (but using the SCA current source).

CMS. The significant difference from the nominal value calls for a calibration of the current source.

The second factor accounts for a bias of up to 12 °C. If the SCA 12-bits ADC specifications guarantee an Integrated Non-Linearity (INL) better or equal to 2 LSB units, the PT100 resistance variation with temperature remains too large.

Finally, the third factor accounts for a bias of up to 3 °C. Even if the smallest contribution of the three, this is also the simplest to compensate for without any hardware change. The resistance of the routing traces has been measured via a 4-wire method[1] on 4 blank OptoHybrid PCB (Table 4.2).

[1]The 4-wire method is a resistance measurement technique that uses one pair of conductors to carry the reference current and another pair of conductors to measure the voltage drop across the resistive element. It provides more accurate measurements that simpler methods as no current flows through the voltage sensing pair.

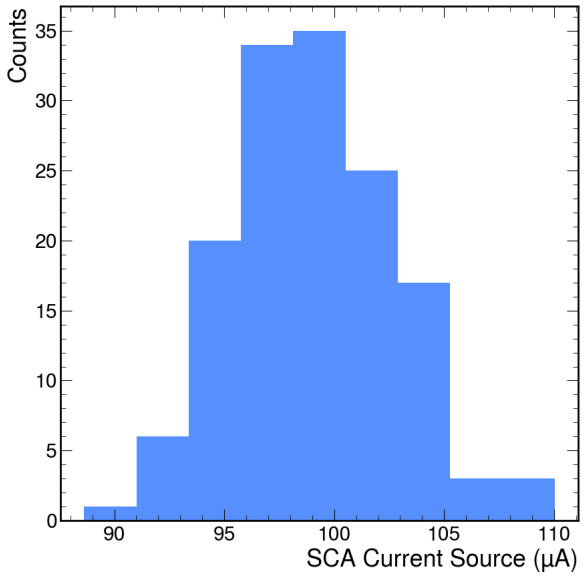


Figure 4.6: Distribution of the current generated by the SCA ADC current generator for the GE1/1 OptoHybrid installed in CMS.

Table 4.2: Resistance of the traces on the OptoHybrid board PCB for the different sensors connected to the SCA ADC.

	Current source (Ω)	GEM VTTx (Ω)	GBT0 (Ω)	CSC VTTx (Ω)	FPGA (Ω)
PCB #1	0.495	0.389	0.661	1.171	0.550
PCB #2	0.420	0.446	0.740	1.282	0.564
PCB #3	0.357	0.323	0.557	1.052	0.395
PCB #4	0.416	0.357	0.621	1.103	0.432
Average	0.42 ± 0.06	0.38 ± 0.05	0.65 ± 0.08	1.2 ± 0.1	0.48 ± 0.08

As a result, three solutions have been put in place to improve the measurement accuracy:

1. The 1.8 V voltage measurement input of the ADC, unused in the final OptoHybrid board design, has been converted to readout the SCA ADC current source with the help of a high precision $1\,\mathrm{k}\Omega \pm 0.05\%$ resistor;

2. The PT100 has been replaced by a PT1000, whose variations better fit the characteristics of the SCA ADC;

3. The resistance of the routing traces on the PCB is taken into account in the temperature computations.

With the full solution implemented, the accuracy of the PT1000 temperature measurement improves to ±4 °C, taking into account the ADC INL for both the current source calibration and the PT1000 resistance measurement.

However, due to production constraints, not all the OptoHybrid boards are modified. Three variants exist, from the least to the most accurate:

1. Original version with the PT100 and no current measurement, which can lead to temperature measurement errors exceeding ±30 °C;
2. Partially patched version with the PT100, but a precision resistor for current measurement, reducing the error to ±12 °C;
3. Entirely patched version with the PT1000 and the precision resistor, achieving errors within ±4 °C.

It is expected that all OptoHybrid boards will be refurbished during the LHC LS3 in order to implement variant 3, finally providing accurate temperature readings on all boards.

4.4 Combined GEB & OptoHybrid quality control

While the GEB-only and OH-only steps were initially thought to be sufficient, it was quickly realized that some issues could only be detected when testing the system as a whole: communication chain instabilities, long-term instabilities, impossibility to properly bias the VFAT circuits,... This led to the creation of a new quality control step at ULB where each OptoHybrid is tested with at least one full GEB.

After assembling the OptoHybrid, VFAT, and FEAST on the GEB, the QC7 test routine was run. It goes through the following sequence:

1. Configure the GBTxs through the GBT optical links
2. Program the OptoHybrid FPGA
3. Perform a GBT phase scan and set the correct phases for all VFATs
4. Perform a VFAT DAC scan and set the correct DAC values
5. Configure the chamber for data-taking, including all VFAT registers
6. Take & analyze an S-bit rate scan
7. Take & analyze a set of S-curves
8. Take & analyze a threshold scan

In essence, the test routine is identical to QC7, but without requirements on the detector behavior, only on its electronics. More specifically, the last 3 steps (6-8) are used exclusively to test the communication reliability. Differently, in a standard QC7, the results are scrutinized to also ensure a good connection to the readout strips, acceptable noise levels, and that the VFAT has not been damaged in any way during the assembly procedure. Additionally, the QC@ULB implemented a few complementary steps:

1. Perform a fast S-bit mapping test
2. Perform 10 power cycles of each board
3. Monitor the GBTx and trigger link error counters overnight (more than 10 hours)

All the aforementioned tests are communication intensive and thus properly exercise the system in order to ensure a reliable communication between all components. However, some tests also served additional purposes. The OptoHybrid FPGA programming ensures that the board can run a regular firmware and that the GEB is able to provide enough current.

The DAC scans for their part test the ability to properly bias the VFAT analog front-end (Figure 4.7). Indeed, when none of the nominal currents can be reached for a given VFAT, this is a strong indication of a powering issue. Upon inspection, the VFAT Panasonic connector was systematically found improperly soldered, leading to the GEB to be eventually rejected. Over time, the GEB visual inspection got more and more stringent, reducing the chance to encounter such an error.

Whereas the QC7 S-bit rate scan can identify misconnections, the S-bit mapping test is much stricter: each S-bit sent by the VFAT must be correctly received by the OH FPGA. Additionally, it provides a way to quickly report the anomalous differential pairs which can then be investigated by the operator.

Once all tests passed, the final routine consisted in monitoring the links stability overnight. Errors on the GBTx links and trigger links were monitored for more than 10 hours. Only systems reporting no error were considered validated. This corresponds to a Bit Error Rate (BER) of less than 10^{-13} with a confidence level above 99.99%. The overnight test was accompanied by the repetition of 10 power cycles which aimed at ensuring a strong GBTx lock.

Overall, in the absence of connectivity issues, the GEB and OptoHybrid combined testing takes around 30 min. The overnight test is obviously to be added on top, but does require very little human intervention.

While proceeding with the testing, it was quickly realized that the bottleneck at ULB was this last step of combined testing. In particular, the whole testing chain had to be put on pause if any problem was discovered.

In order to limit the delays, the setup was scaled up to enable the testing of up to six GEB + OH in parallel. Figure 4.8 shows the entire testing system in its most optimized

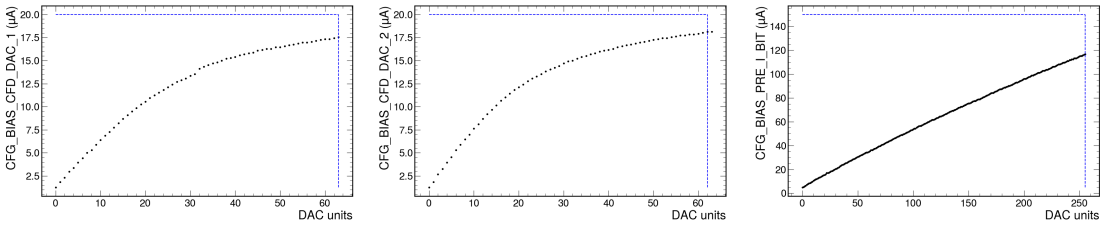


Figure 4.7: Failed VFAT3 DAC scans for which the nominal current values cannot be reached. This is the typical symptom of a chip powering issue.

configuration. Each of the six tables hosts the electronics of a full GE1/1 detector. The DAQ is based on a scaled down version of the final μTCA system and shared between all devices under tests. The LV is based on two laboratory bench power supplies, controlled via RS485. With this setup, most of the tests can run simultaneously, or at least without human intervention. It really allowed to systematically run the automated power cycle and the overnight stability testing.

In order to save debugging time, the procedure also envisioned to always replace only one of the two components, either the GEB or the OH.

4.5 Database storage

As the quality control progressed and chamber were assembled at CERN, recording and sharing the results became critical. Not only for curiosity but to keep track of the different parts and their status: the OptoHybrid, GEB, FPGA, GBTx serial numbers, the status of the PT100 & resistor changes, the GBTx fused configuration, the QC steps run and their outcome,... This exceeds the simple GEM-internal need for traceability as parts tracking is essential to meet the CERN radioprotection requirements.

Multiple databases, based on Oracle technologies, linked to the life of the CMS experiment, and targeting the storage of non-event data, are suitable for this task [53].

The **CMSONR** database cluster is located in the CMS online cluster and can operate at any time, independently from connectivity outages to the outside world. The two best know databases are *OMDS* (Online Master Database System) and *ORCON* (Offline Reconstruction Condition DB - ONline subset). The former stores the detector configuration data and records the condition data produced by all sub-detectors; the latter stores the subset of condition data required for physics reconstruction and detector performance studies in the online realm, typically the HLT and Online DQM.

The **CMSR** database cluster is located in the CERN main data center in Meyrin. Among other databases, it contains *ORCOFF* (Offline Reconstruction Condition DB - OFFline subset), which stores a copy of the *ORCON* database as well as the historical and present information required for any physics reconstruction.

Both clusters are also replicated within the CERN computing infrastructure for backup, redundancy, and accessibility.

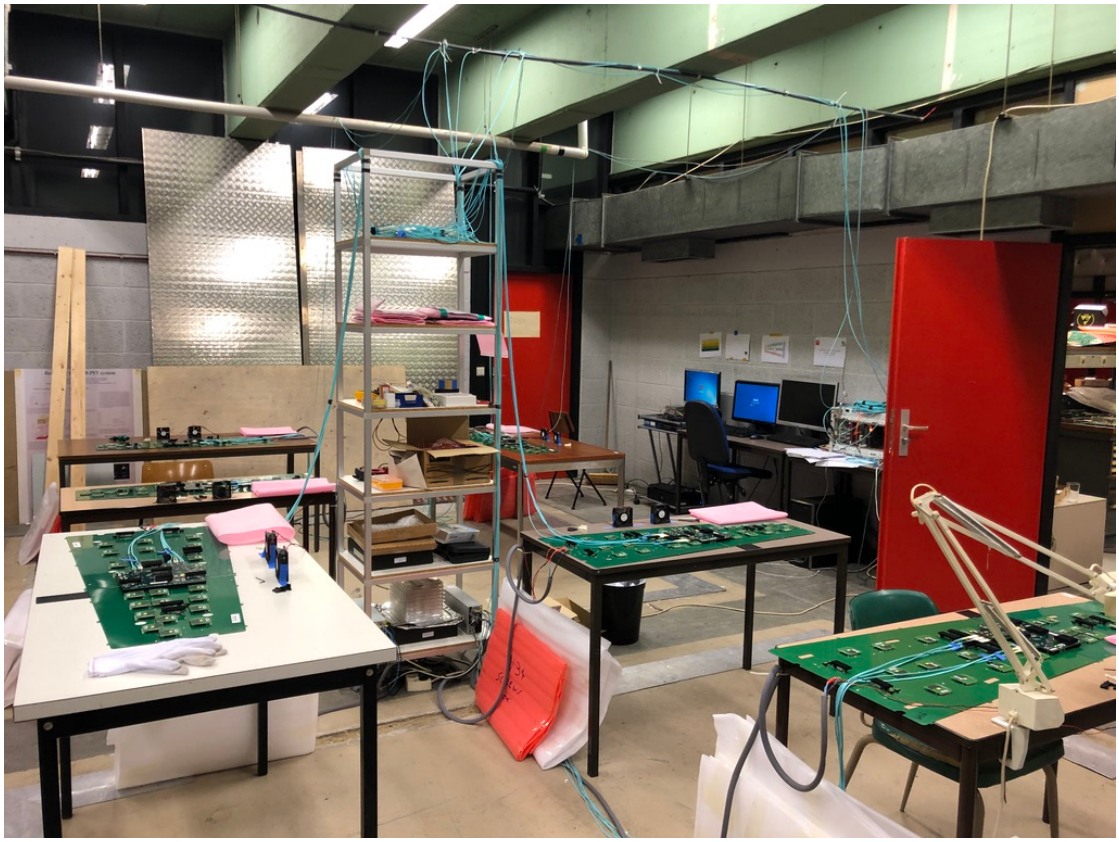


Figure 4.8: Picture of the 6 test stands for the combined GEB & OH quality control. The µTCA-based DAQ system is visible on the right.

The GEM group chose the *ODMS* to archive its construction database[2]. It was also decided to use the *DBLoader* framework [70]. This framework provides standard and extension SQL schemas serving the construction and condition purposes.

Rather than connecting directly to the SQL database to insert new items, an XML file interface is provided. Files representing the parts and their QC can be prepared in advance as the tests go and later push to the database, typically as they are shipped. In the other direction, data can be retrieved via SQL views, abstracting the database internals, the CMS OMS web interface, and the GEM custom web interface. At this stage, the latter one is the most developed and typically used (Figure 4.9).

[2]Considering that the information is unused for online purposes, there is currently a push to move toward *CMSR* which provides easier accessibility outside of the CMS cluster.

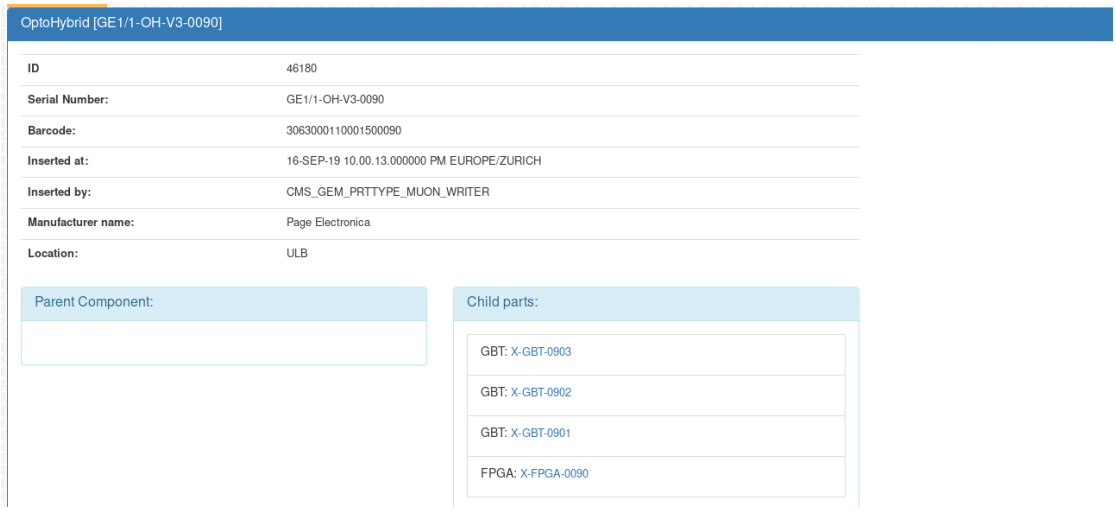


Figure 4.9: Screenshot of the web interface used to display the GEM construction database content. The detailed description of the OptoHybrid board *GE1/1-OH-V3-0090* is shown.

4.6 Outcome & lessons learned

If the quality control performed at ULB did not lead to any delay in the GE1/1 production and smoothened the detector assembly experience at CERN, it remained a very time-consuming process and was always in lean manufacturing.

A significant part of the problem stems from the absence of full detector integration and testing with all designers and experts before the beginning of production. Unexperienced personnel had to discover all the features and shortcomings of a system produced by others. Another part of the problem lies in the late availability of the final revision of the electronics components and different priorities for the CMS GEM project at the time. Indeed, the first batch of production components was ”tortured” during the so-called sustained operation studies focusing on the VFAT3 channel losses and protection circuits optimization. This eventually led to the VFAT3 hybrids being available but not in their final revision due to insufficient input protection, the QC8 hodoscope not being commissioned, and the firmware and software required for the QC process not being developed and exercised in advance.

It took more than 6 months to develop and then tune the final quality control steps, both the electronics QC at ULB and the QC7 & QC8 at CERN, based on mutual feedbacks. The full and stable production speed was only reached in Q2 of 2019 while it originally started in Q4 of 2018. For reference, the first GE1/1 chamber installation took place in July 2019.

Retrospectively, the main weakness of the qualification procedure was that it remained only a functional testing and not a systematic point-by-point testing. This will have to be improved for the future GEM stations. Each component should meet its re-

quirements, must properly implement the interfaces, and must be thoroughly validated. Most importantly, the system as a whole must be tested together to ensure that all high-level requirements are met. It is critical to test every single feature, even if implemented only as a fallback option.

The main weakness of the GE1/1 electronics remains its mechanics. The two pairs of SAMTEC connectors with complex geometry create important difficulties to mount the OptoHybrid on the two GEB parts in case of imperfect alignment. This is further accentuated by the rigid interconnections of the readout board with each of the 24 VFAT3, and each of the VFAT3 with the GEB. Together, the VFAT3 Panasonic connectors and the OH SAMTEC connectors can create and suffer from significant mechanical stress. Retrospectively, using a single GEB with more a conventional connectors geometry would have been better despite the challenges to produce such a large and thin PCB.

The GEB-OH standoffs also remain a critical part of the design, particularly since they are used to carry the power to the OptoHybrid board. Even after reinforcement of the soldering, they can break during chamber assembly, quality control, or even after installation in CMS, requiring chamber extraction for repairs. This issue is further emphasized by a heavy, rigid, and, sometimes, unfitting cooling system.

The project has learned the lessons for the GE2/1 and ME0 designs: the GEB PCB is made of a single piece and no power is carried through the standoffs, limiting their purpose to be solely mechanical. Significantly more flexibility is also brought to the design via the usage of BGA-packaged VFAT3 ASIC and the conception of a host PCB which includes a flex section. Additionally, the cooling system is minimized to reduce its weight and uses copper braids to increase its flexibility.

Finally, Table 4.3 and Table 4.4 present the final statistics on the status of the OptoHybrid and GEB after completing the GE1/1 detector production. The numbers are shared to the best of our knowledge. Once the parts were delivered to the assembly laboratory at CERN, the lack of proper tracking and precise error reporting prevented giving accurate counts, exact yields, and rejection causes. It must also be noted that some damages likely occurred at CERN during the detector assembly.

Table 4.3: Summary of the OptoHybrid board status after quality control and chamber assembly. If multiple problems apply, only the most significant one is accounted for.

Status	OH
Fully operational	162
Communication issue(s) with VFAT	6
FPGA not working with some firmware builds	1
FPGA not receiving all clocks	1
GBTx not detected by the I2C dongle	1
New and untested	4
Total	175

Table 4.4: Summary of the GEB status after quality control and chamber assembly. If multiple problems apply, only the most significant one is accounted for.

Status	Short Narrow	Short Wide	Long Narrow	Long Wide
Fully operational	77	80	79	86
Broken OH standoff	3	6	1	1
"Noisy"	-	1	-	-
Communication issues(s) with VFAT	6	4	4	1
Impossible to bias some VFAT	-	2	-	-
Shorts in the VFAT connector	-	1	-	1
Tilted VFAT connector	2	4	-	3
Shorts in the FEAST connector	-	1	-	-
Broken FEAST connector	1	-	-	-
Broken SAMTEC connection	-	1	1	1
Damaged PCB during transportation	1	-	-	-
PCB too bended	-	-	3	-
Connector soldering issues	-	-	2	4
Unknown problem (at CERN)	10	7	4	3
Unknown problem (at ULB)	5	2	0	1
Partially validated	0	0	3	3
New/untested	11	6	13	4
Total	116	115	110	110

Chapter 5

Commissioning & Operations

September 2020 marked the completion of the GE1/1 detectors installation, concluding more than a decade of research, development, and construction.

The first two super-chambers were installed in July 2019 in the negative endcap of the CMS experiment. This served as real scale rehearsal of the entire installation sequence, ranging from the transportation between the GEM QC laboratory on the CERN Prévessin site and the CMS experimental cavern in Cessy, to the insertion in the CMS nose, connection of the services, and the detector checkout. The remaining 34 GE1/1 super-chambers for the negative endcap were installed shortly thereafter, between September and October 2019. Finally, the installation of the positive endcap super-chambers took place between July and September 2020, following a particularly challenging production phase defined by the constraints of the COVID-19 pandemic.

This chapter details the GE1/1 detector commissioning and early operations of the GE1/1 detectors, with a focus on the electronics and DAQ-related issues encountered, as well as the first performance results obtained.

5.1 Commissioning

While the installation of the GE1/1 detectors into the CMS experiment represented a major milestone for the GEM project, it merely marked the beginning of the commissioning and operation phases. The initial phase, the commissioning, encompasses all activities required to prepare the detectors for data-taking.

The commissioning overall objective can be summarized as follows: to become familiar with the detectors in their new and final environment, and to ensure they operate optimally.

The commissioning begins with the verification of the detector services. During installation, each detector is connected to its readout optical fibers, cooling pipes, low-voltage and high-voltage cables, as well as grounding. The *services commissioning* phase ensures that none of the connections were damaged before installation and that all are

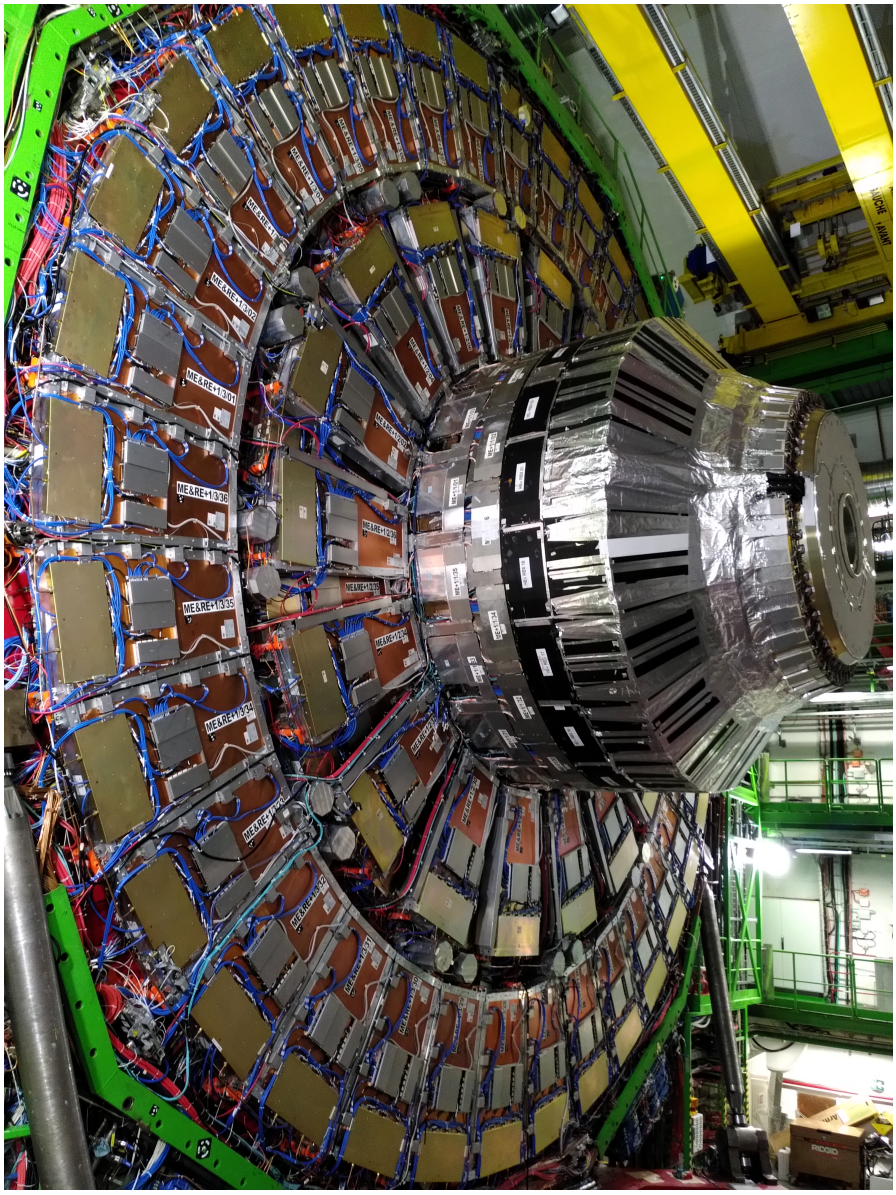


Figure 5.1: Picture of the CMS YE+1 disk and nose, including the newly installed GE+1/1 chambers located under aluminium cover plates, and ready for closure.

reliably and securely attached. This step is an essential prerequisite before starting the detector operations.

The commissioning then enters into the *local commissioning* phase, in which the GE1/1 subsystem is tested and characterized in standalone mode. This marks the very first time the detectors are operated within the CMS environment, requiring scrutiny of

both the chambers and their associated electronics components.

The testing procedure largely follows the QC6 to QC8 sequence, with the necessary adjustments to accommodate CMS-specific constraints. The HV training is first performed in pure CO_2, as in QC6, and then in the Ar/CO_2 mixture, as in QC8. The electronics undergo a similar procedure to that of QC7.

After this initial testing, the local commissioning focuses on fine-tuning the detector configuration. Following a procedure similar to QC8, the front-end settings are optimized, with particular attention given to the VFAT noise and thresholds, which directly impact the detection efficiency.

At this stage of the commissioning, a limited number of detectors can still be extracted and replaced with spares if necessary. Consequently, any non-conformities must be identified and evaluated to determine whether they will affect the detector's performances during operations.

The final stage of commissioning, the *global commissioning*, includes all activities required to fully integrate GE1/1 with the rest of the CMS subsystems. Integration into the central DSS and DCS grants GE1/1 the rights to operate in unattended mode, relieving the GEM operator from continuous monitoring tasks.

Inclusion in the global DAQ subsystem allows the GEM data to be read out in the CMS global data stream, alongside the other sub-detectors - a requirement for participating in the muon reconstruction. Outside of data-taking periods, this effort is supported by regular Mid-Week Global Runs (MWGRs), when all CMS subsystems operate together to identify and resolve any remaining operability or performance issues. These run periods also allowed for the initial spatial and time alignments of GE1/1 with the rest of the experiment using cosmic ray muons.

Subsequent efforts focused on adding GE1/1 to the Level-1 Trigger. The optical links between the GE1/1 OptoHybrid and the CSC OTMB were troubleshooted and validated. Similarly, the trigger links between the GEM back-end boards and the EMTF were commissioned. These steps enabled the recording of the GE1/1 trigger primitives in the DAQ systems of both the OTMB and EMTF, allowing their debugging and validation.

In parallel, the front-end electronics, back-end system, and chamber configurations were refined, in order to improve the detector reliability and performance, and ultimately ensure optimal data-taking with CMS during Run-3.

Although the sequence outlined above may appear linear to the reader, the reality is vastly different. The individual tasks and activities previously mentioned take place in an interleaved and sometimes concomitant manner, driven by the immediate needs as well as the system availability in the overall CMS schedule.

Finally, the author would argue that the commissioning phase is still ongoing, as some optimizations remain to be done, and as certain detector behaviors remain to be understood. Nonetheless, both the GE1/1 station and GE2/1 demonstrator are now fully integrated into CMS and actively participate in the standard data-taking.

5.1.1 Trolley test

During the early commissioning of the first endcap, several electronics issues were identified across multiple chambers. Most of these problems were traced back to broken trigger lines, resulting from a partial disconnection of the Panasonic connectors that interface the VFATs with the GEBs. The commonly admitted hypothesis is that these disconnections occurred during the transportation of the chambers to the CMS cavern, likely due to mechanical vibrations.

With this new information, a dedicated test was developed to prevent similar problems during the installation of the second endcap. This test was designed to be performed on the experimental cavern floor, after transportation, and just before the installation of the chambers in CMS. Due to the strict time constraints and the challenging environment, the resulting test was constrained to complete within 20 minutes, avoiding any risk of the chambers overheating as no cooling was available in such conditions. Required services were also limited to a single low-voltage bench power supply and a single trunk fiber connection to the existing backend in the service cavern. This test, resembling a fast version of QC7 with particular emphasis on verifying the VFAT-to-GEB connection, became known as the *trolley test*, named after the trolley used to host the chambers during the testing.

By the end of the second endcap installation, the trolley test had identified 5 chambers with connectivity issues. These were returned to the GEM QC laboratory, repaired, and then installed in the experiment.

A detailed description of the test procedure and its results can be found at [92].

5.1.2 Mapping validation

One of the essential steps before beginning in-depth commissioning is the validation of the services mapping, as it is unfortunately too easy to misconnect cables when working in the dense CMS nose region. Two separate checks were performed: the **DAQ-to-LV** mapping and the **DAQ-to-HV** mapping.

The *DAQ vs. low-voltage* mapping consists of powering up a single detector and recording which readout link becomes active. Any mismatch with the expectations indicates a mistake in the optical fiber connections or the low-voltage cabling chain, which must then be thoroughly checked. This validation can be performed immediately after the service connections are complete.

The *DAQ vs high-voltage* mapping , on the other hand, cannot be validated until an advanced stage of high-voltage training, which, in CMS, is a lengthy procedure spanning over multiple months. Indeed, it relies on the observation of ”muon activity” in the detector, originating either from cosmic ray muons or from environmental radioactivity. Figure 5.2 shows the typical behavior of the S-bit rate as a function of the front-end threshold, both with (orange) and without (blue) HV applied. When no HV is applied, the absence of gas amplification causes the rate to drop sharply. Once the HV is applied to all electrodes, the rate reaches a plateau

due to the gas amplification. As with the low-voltage mapping check, any discrepancy requires a careful inspection of the HV hardware chain. Considering the HV distribution architecture described in Section 2.2.5, an ambiguity remains: the two chambers connected to the same HV channel could be inadvertently swapped.

This method assumes that at least one service is properly connected: a complete swap of both the DAQ, LV, and HV chains would remain undetected. Nevertheless, it is the only viable validation method available during the early commissioning phase. Further validation of the DAQ fiber mapping can be performed later using muon track reconstruction. This is, however, not applicable during the early phases of commissioning. At the time of writing, this last check has been completed and has confirmed the accuracy of the full system mapping.

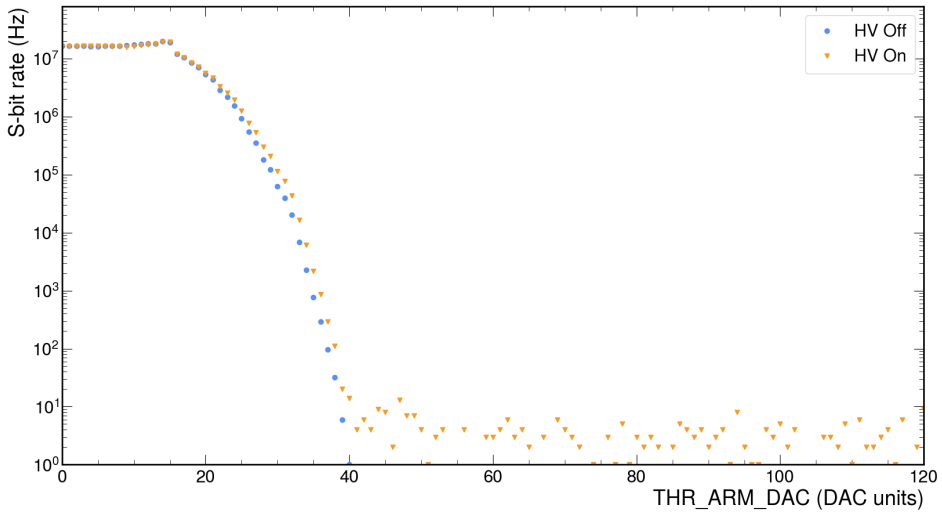


Figure 5.2: S-bit rate as function of the comparator threshold with the high-voltage off (blue) and on (orange), and the VFAT pre-amplifier configured in medium gain. Taken in March 2022 on GE+1/1/07 Ly2 VFAT #14. Each point represents 1 second of accumulated data, recorded sequentially from low to high threshold. The large rate fluctuations observed at high thresholds with the HV on (orange) are attributed to the stochastic nature of the signals, originating from either cosmic muon showers or from high-voltage events such as coronal discharges.

Going one step further, and provided that the electronics is tracked reliably, one can verify which chamber was physically installed in each CMS nose sector. In addition to revealing electronics tracking inconsistencies, this process has proven that multiple chambers were installed in sectors different from those that were previously assumed. While

such mismatches do not pose any major operational concerns, traceability is essential due to the radioprotection constraints enlightened in Section 4.5.

In summary, the mapping validation process allowed the identification and correction of service mismatches, as well as the resolution of component tracking inaccuracies.

5.2 Communication instabilities

As anticipated in Chapter 3, GE1/1 suffers from a wide range of unexpected communication failures with the front-end electronics (GBTx, VFAT, OptoHybrid FPGA). Observed very soon after installation in CMS, those instabilities affect single or multiple slow-control transactions, the fast-controls, the tracking data readout stream, or the trigger primitives, both during calibration scans and regular data-taking. In-depth investigations led to the discovery of a combination of multiple effects that are described in the sections below.

The communication instabilities are part of the main operational reasons that participated to the redesign of the firmware and online software. The end goal was very clear: gracefully handling any kind of communication instabilities with the smallest possible impact on the detector performance, without compromising the discovery of other, potentially fixable, problems, and without negative impact on the overall CMS data-taking. As it will be shown below, the errors are handled by statically or dynamically masking components, or attempting retries depending on the specific context. Additionally, the exact status of the system has been included in the GEM DAQ data format to provide bookkeeping with a per-event granularity. The final VFAT-level granularity handling was eventually reached in March 2022.

The two main lessons learned from these investigations are the following. First, *always* validate the data sent by the front-end, and react *accordingly* in case of data corruption. "Accordingly" is of the utmost importance since the appropriate reaction depends on the exact failure mode. Second, start the commissioning of the full system as soon as possible with enough time for long-term stability tests, regardless of the quality control put in place.

5.2.1 GBTx fuses corruption

This first issue is caused by a now well-known defect in the GBTx chip that may fetch a corrupted configuration at power-up or when its watchdog is triggered. As the GBTx reads its initial configuration from fuses, sampling errors can silently occur in a few percent of the cases, increasing with the number of chips on the same power line. Consequently, if critical configuration bits are corrupted, the GBTx is unable to lock on the downlink stream and provide a valid uplink stream.

The observable symptom consists of random GBTx links not initializing properly after powering the front-end or configuring the back-end board (as the GBTx watchdog is triggered). As this is a defect in the ASIC design, a higher-level solution has to be implemented. And, as the issue affects chips randomly, the solution must be dynamic.

The GEM workaround takes advantage of the command channel of the DIM protocol already used in the monitoring system (Section 3.6): each GE1/1 chamber with problematic links is automatically power cycled while the GEM subsystem is being configured. Since another link from the same chamber could fail after the power cycle, it is limited to one per configuration. On average, about 3 GE1/1 chambers are power cycled for this problem at each configuration of the GEM subsystem since the deployment in 2021.

5.2.2 VTRx outgassing

Another major source of communication failures, which amounts to $\mathcal{O}(10\%)$ of the VFAT excluded in the readout, is linked to the VTRx transceivers. Early during the GE1/1 commissioning, it was observed that communication with the GBTx could be lost at random moments without a clear sign on the origin of the failure besides that the longer the LV was powered on, the more important the issue was. It also observed that rather than completely losing communication with the GBTx, groups of VFAT connected to a given GBTx could misbehave. Simultaneously, the CMS HCAL group investigated and reported a similar problem to the CERN electronics group, which triggered its own investigation campaign due to the large usage of VTRx throughout CERN projects [86].

It was discovered that the root cause of the issue lies in the out-gassing of the glues present in the VTRx ROSA (Receiver Optical Sub-Assembly) that then condenses on the cooler optical fiber ferrule. The deposit thus leads to a loss of the received optical power and, more importantly, refraction of the optical signal, causing the observed communication instabilities. As the deposits move or evaporate, the situation can quickly evolve from non-working to recovered and vice-versa, explaining the correlation with the powering duration.

The solution initially proposed by the CERN electronics group was to use VTRx baked at 85 °C for 500 hours. This however proved not to be sufficient. If the epoxy glue could be cured with that method, the UV-glue could not. The lack of complete curing also allows the mixing of the two adhesives and unwanted reactions between them, leading to a larger variety of volatile compounds.

The definitive solution, confirmed by the CERN electronics group, comes from the CMS HCAL group which showed that if the temperature gradient between the fiber and VTRx ROSA remains below 10 °C, no out-gassing takes place [51]. The fiber ferrule remains clean, whether or not the VTRx has been baked. At this point, however, GE1/1 was fully installed and no last-minute modification to the cooling system could take place for Run-3. Thus, the issue had to be fully mitigated within the DAQ system. This is achieved by the aforementioned developments: in case of instability related to the GBTx communication, the link and all its downstream components are dynamically masked.

It is suspected that the problem was not observed earlier during the QC steps because the optical power was higher thanks to the shorter fibers and reduced number of connections; the system was never operated for a long enough duration; and the closed environment in CMS promotes temperature increases.

Despite the implementation of an appropriate workaround and considering the scale of the issue, the two most problematic GE1/1 super-chambers were extracted from CMS for inspection and refurbishment. Their OptoHybrid boards were then modified to enable the VTRx Receiver Signal Strength Indicator (RSSI) readings (Figure 5.3), obtaining the measurements shown in Figure 5.4a for a single GE1/1 chamber. The RSSI provides a proxy to the optical power received by the VTRx, itself a proxy for the presence of contaminants on the optical path. A decreasing and jumpy behavior is typical of the out-gassing issue.

As baking the VTRx was initially considered to alleviate the out-gassing process, baked transceivers were installed for long-term monitoring in the CMS environment. Those chambers are now refurbished with baked VTRx and now carefully monitored. Figure 5.4 shows the RSSI behavior before and after the refurbishment.

As the previous paragraphs have shown, this is however not the end of the story: only proper cooling of the VTRx can completely avoid the out-gassing process. Therefore we proposed to design a cooling solution adapted to its GE1/1 electronics. Due to the tight space and the absence of any cooling on the bottom part of the OptoHybrid, the design was particularly challenging. The final prototype shown in Figure 5.5 provides the best compromise between effectiveness, ease of assembly, and robustness found so far. The latter two points are critical for an expected mass refurbishment scheduled during LS3 because of the VTRx flex-PCB fragility. Since the beginning of 2024, this final prototype has been installed on two GE1/1 super-chambers, serving as the final demonstrator of the solution.

The GE2/1 design already foresees proper RSSI readings, baked VTRx, and VTRx cooling, eluding most of the associated risks. All three enhancements installed in CMS (GE1/1 baked, GE1/1 cooled, GE2/1 baked and cooled) remain under scrutiny to spot any problem and anticipate the GE1/1 refurbishment during the LS3 as well as the end of the GE2/1 production.

5.2.3 Powering weakness

Over time, groups of VFAT not linked to a given GBTx started to misbehave: low efficiency, communication errors, impossibility to properly bias the front-end, large number of channels reported as dead,... It took time to identify those groups as matching the VFAT FEAST power domains. Based on the failure mode and pattern, there is a strong evidence that the FEAST is unable to provide enough voltage or current to the VFAT. In the absence of a direct method to measure the current or voltage of the VFAT power rails, confirmation however remains impossible until the extraction of the chamber from CMS. It is highly suspected that the electronics mechanics is pulling the FEAST away from its connector as it was already observed during assembly, QC8, or installation. In this specific case, the issue can be subtle and leads to poor physics performance, so it has been decided to mask statically the affected VFAT. To improve monitoring and diagnostic in the future, the addition of voltage readout capabilities for the FEASTs powering

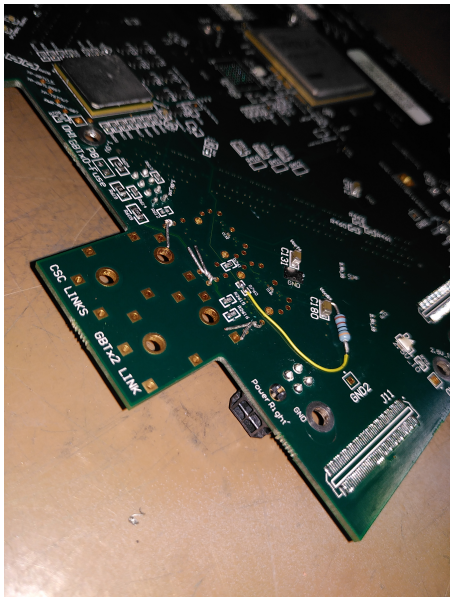
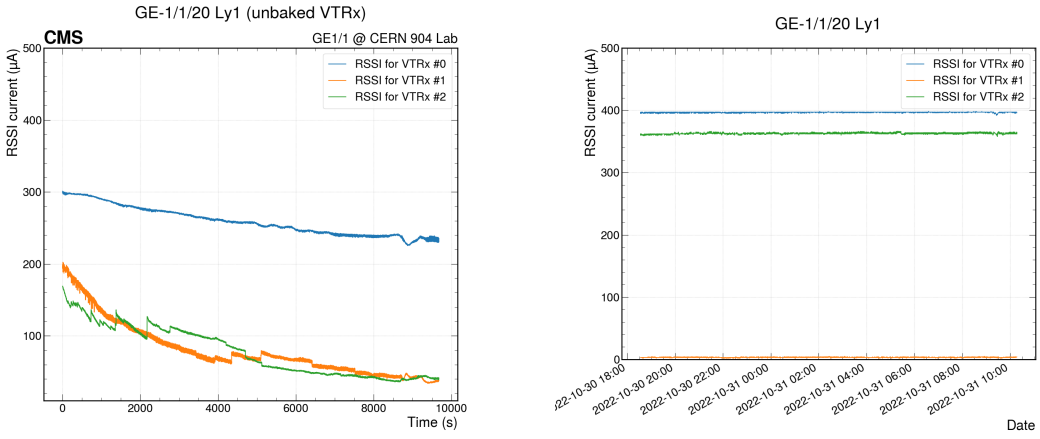
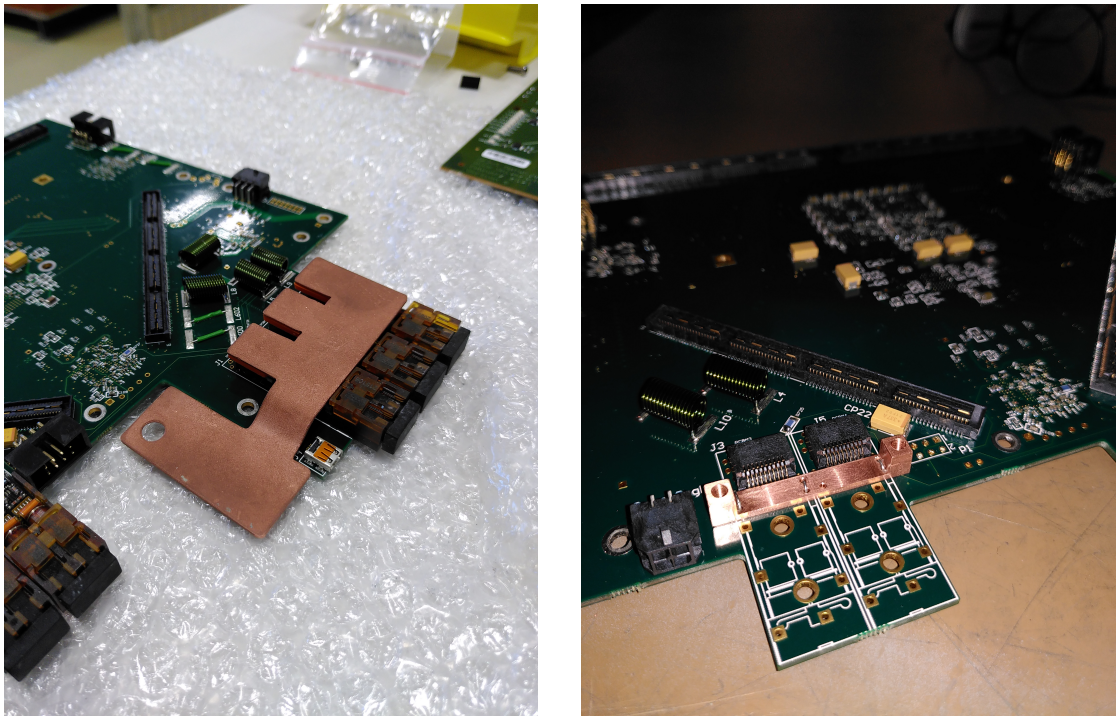


Figure 5.3: Modification (fix) of the RSSI readout circuit on one of the GE1/1 OptoHybrid boards.



(a) Original VTRx (b) Baked VTRx

Figure 5.4: Measurement of the RSSI of the three VTRx installed on the chamber GE-1/1/20 Ly1. In the original situation, the rapid drop of the indicators, as well as their fluctuations, are characteristic of VTRx optical transceivers suffering from glue out-gassing. After VTRx baking, the indicators remain very stable on extended periods of time. The almost null value for VTRx #1 is attributed to a failure in the readout circuit. The differences in behavior are explained by a chip-to-chip variation of the VTRx receiver optical sub-assembly block, in particular its glue, and the different transmitters and fibers.



(a) Cooling plate for the VTRx, with the heat extraction tongue visible on the bottom left.

(b) Mechanical support located below the VTRx PCB tongue.

Figure 5.5: Prototype of the GE1/1 VTRx cooling system.

the VFATs is under consideration for all GEBs and OptoHybrid boards as part of the GE1/1 refurbishment planned in LS3.

5.2.4 Saturation of the VFAT bandwidth

As described in Section 2.2.3, the VFAT3 communication port is multiplexed between clock, fast control, tracking data readout, and slow-control transactions. It is quite natural for the slow-control transactions to receive the lowest priority in order to ensure critical information is brought to and retrieved from the chip as quickly as possible. However, it comes with the drawback that some slow-control transactions can take an extended amount of time to complete.

The back-end firmware implements a hard timeout of 40 µs on any slow-control transaction to avoid deadlocks or large bus contention and to increase the system responsiveness. In case that the VFAT3 communication port is intensively occupied sending tracking data, the slow-control transactions can timeout even in the absence of link instability.

Such a situation causes a specific problem to the latency scan that requires online changes of the latency parameter on all VFAT in the system. Other scans do not suffer from the same concern. So, rather than masking the VFAT for any failed transaction, it

was decided to attempt each slow-control transaction up to 10 times during the latency scan.

5.2.5 High-voltage discharges

It was observed early on during QC8 that the HV discharges could cause the electronics to lose communication. As the detector started to be operated with a higher particle background, the HV discharge rate increased. The few events recorded on GE1/1 during QC8 became more during operations in CMS. This behavior was eventually traced back to the HV discharges being picked up by the readout board strips and creating voltage fluctuations in the LV power rails, which in turn reset the VFAT and GBTx. On the ME0 detectors where the behavior was studied, the situation is such that there is an almost 100% probability of resetting the VFAT in the vicinity of the discharge. Since the VFAT is fully reset, only a full reconfiguration and resynchronization process is able to help. This procedure is being worked on as part of the automatic masking and recovery described in Section 3.7.2. Until then, the VFAT need to be masked.

5.2.6 Random communication failures

Finally, the last kind of communication failure encompasses all random errors that can be observed with the front-end chips. In some cases, the failure was identified as linked to faulty connectors. Indeed, the large size of the GEB PCB (∼60 cm) and the rigidity of the interfaces induce significant mechanical stress on the GEM readout board-to-VFAT, VFAT-to-GEB, and GEB-to-OptoHybrid connectors in case of slight misalignment as already described in Chapter 4. Experience has shown that the VFAT-to-GEB connector is the most susceptible to loosening because of the vibrations induced during installation in CMS.

Since the root cause of those issues is extremely difficult to establish online, the only realistic option is to mask the corresponding component in case of error. Failing to do so would result in potentially recurring errors, which would impact much more strongly the operations, the data-taking, and the data quality.

5.3 VFAT damages due to HV discharges

As presented in Section 2.2.3, the VFAT3 chips are sensitive to channel losses due to propagating GEM discharges. And despite the mitigation measures implemented on VFAT3 hybrid and the GE1/1 chamber HV filter, it remains essential to monitor the system health during commissioning and operations. Firstly, it enables the adjustment of the detector operations would the number of damaged channels increase too quickly. Secondly, it helps ensure the effectiveness of the mitigation measures put in place in view of the GE2/1 and ME0 station development.

5.3.1 Channel damages

A readout channel can be classified as active or inactive. The active channels defined are those fully functional and included in the readout. The previous section showed that a significant fraction of the inactive channels in GE1/1 are due to communication instabilities. Channels can however get damaged during the high-voltage operations and stop recording hits while still being included in the readout. We distinguish two modes of failure:

Disconnected channels Channels for which the wire bond between the ASIC and pin has been destroyed. It still responds to calibration pulses but does not record signals from the readout strips anymore. The electronics noise extracted from S-curves is much lower that for the other channels within the same VFAT.

Dead channels Channels for which the analog channel in the ASIC is damaged. Such channels do not react to calibration pulses anymore and do not see signals from the detector.

The damaged channels can be identified using four different methods:

The per-channel efficiency method is the most exhaustive one and should catch any kind of problem. It however requires precise track reconstruction and large statistics, which make it inapplicable in the absence of collisions at nominal luminosity.

The low threshold runs method aims at detecting dead channels by taking data with random triggers and very low *THR_ARM_DAC* threshold. The analysis is currently being refined by the GEM Detector Performance Group (DPG). It is also relatively slow to take (~ 30 minutes) and thus not taken more than a few times a week, outside of collision periods.

The S-curves scan detects both disconnected and dead channels. While technically not complicated, it was one of the last scans to be optimized for large-scale execution. Indeed, it heavily relies on the local readout system, which was developed late in the DAQ timeline. Therefore, this method was not available during the detector commissioning phase. Furthermore, its analysis workflow is slightly more cumbersome than that of the next method.

The calibration pulse scan method was then developed to provide a swift procedure that can be used daily. This scan effectively emulates a one-point S-curve with low thresholds and maximal injected charge. In this case, the local readout is not required, taking advantage of firmware counters and alleviating the need to use the full analysis workflow. While this method scales poorly for complete S-curves, it was deemed fast enough for a one-point measurement, achieving a scan of the full system in less than 30 seconds. The analysis is also trivial: a channel is considered dead if it does not record a single hit during the scan. Figure 5.6 shows a summary wheel plot for all the chambers in the GE-1/1 Layer 1 disk. Additionally, the results are pushed to the monitoring database described in Section 3.6 to provide

trend plots online (e.g. Figure 5.7). If extremely fast, this technique is however insensitive to disconnected channels.

In August 2024, less than 0.23% of the channels are reported dead by the latter method, not including damaged VFAT3 biasing circuits.

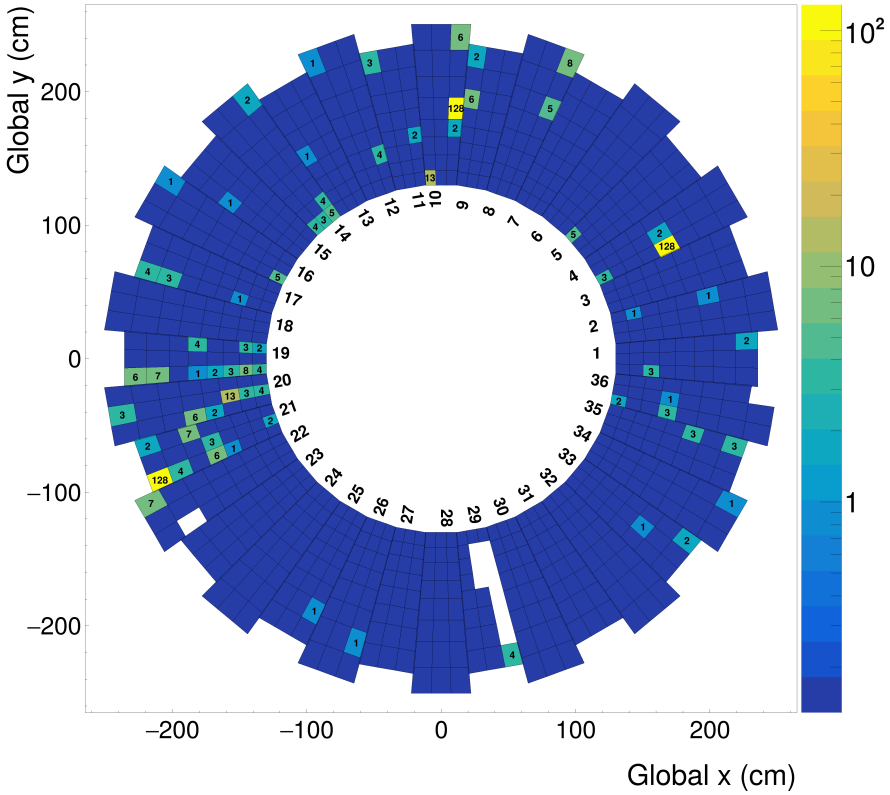


Figure 5.6: Summary of the number of dead channels per VFAT of the GE-1/1 Layer 1 based on the calibration pulses scan. VFAT with communication errors at the calibration scan time have been excluded and do not appear in the plot. VFAT with all 128 input channels reported as dead are improperly configured or biased.

5.3.2 Channel damages on the GE2/1 demonstrator

As one goal of the periodic health checks is to ensure the effectiveness of the mitigation measures, careful monitoring of the GE2/1 demonstrator makes particular sense. Figure 5.7 shows the evolution of the number of dead channels on the M3 module during its first year of operations. The channel losses are striking, particularly for VFAT #2 which suffered more than 60% of dead channels.

This behavior triggered in-depth studies in the laboratory which confirmed the ineffectiveness of the AC-coupled protection circuit (described in Section 2.3.1) on large-scale GEM detectors. The AC-coupled protection circuit had been designed to reduce the damage probability by using a decoupling capacitor (to prevent the ASIC input channel from absorbing the full energy of the discharge) and to reduce the discharge propagation probability to the anode by using a the drain resistor (to quench the discharges precursor current). However, the large capacitance of the full-size GEM foils, and the corresponding large stored energy, negate the effectiveness of this protection scheme. Eventually, the protection circuit was reverted to the regular one with resistors in series.

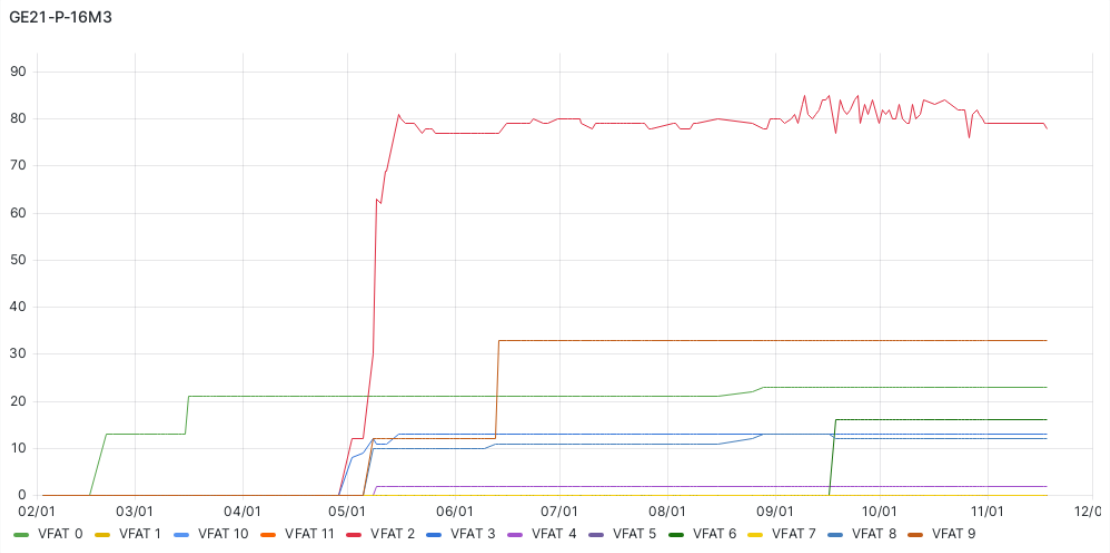


Figure 5.7: Evolution of the number of dead channels for each VFAT3 on GE2/1 demonstrator M3 module (GE+2/1/16 Ly2 C) during its first year of operation.

5.3.3 DAC circuits

It has been reported during QC8 that GEM discharges can destroy (part of) the calibration module of the VFATs or create shorts internal to the VFAT3 ASIC. The former is usually seen as empty S-curves due to the absence of calibration pulses, while the latter increases the chamber LV current consumption and often affects the VFAT operations.

The same effect is also seen in the GE1/1 system in CMS: some detectors start drawing more low-voltage current after a high-voltage discharge. For example, the discharge shown in Figure 5.8 coincides with an increase in the LV current of 0.35 A. In the absence of other effects, this does not pose any problem to the operations as long as the total current drawn of the chamber remains below the 6 A that the power supply can provide.

However, another behavior potentially linked to those HV discharges is increasingly worrisome: it is suspected that they may damage the VFAT biasing circuits, leading to poor performance of the affected chip.

For example, the discharge reported above is likely linked to the so-called "broken DAC circuit" shown in Figure 5.9. The patterns usually suggest some broken bits in either the biasing DAC or monitoring ADC circuit. It is however quite difficult to assert the correlation between the two events since the DAC scans are taken only on-demand at irregular intervals. In this specific case, the last clean scan was taken on the 30th of July 2024, whereas the HV event occurred on the 10th of August, and the verification scan was taken on the 12th of August.

Deeper studies are ongoing to find evidence of whether or not the damage is actually due to discharges. At the time of writing, 29 VFAT3 have confirmed damaged biasing circuits (out of 3456).

5.4 Electronics noise

Immediately after installation in CMS during the early commissioning in 2020, the electronics noise levels showed a significant increase with respect to the quality control tests performed in the laboratory. From an equivalent noise charge below 1 fC, the noise rose to more than 5 fC in the worst cases. Such increases in ENC would have required to push up the VFAT thresholds and thus reduce the detection efficiency.

Figure 5.10b presents the average ENC for each chamber in the GE1/1 negative endcap just after installation. These results were later confirmed in the GE1/1 positive endcap (Figure 5.10a).

An extensive investigation campaign was thus carried on in spring 2021 over 6 months, taking and analyzing over 3000 S-curves, in order to reduce the noise to acceptable levels (<1 fC). At first, the potential impact of a wide variety of hypotheses were studied:

- Environmental factors:
  - Daily fluctuations
  - Proximity of the CSC ME1/1 chambers
  - Proximity of the HCAL endcap detector
- High-voltage system
- Low-voltage system
- Interaction between GE1/1 chambers
- Interaction between GE1/1 super-chambers

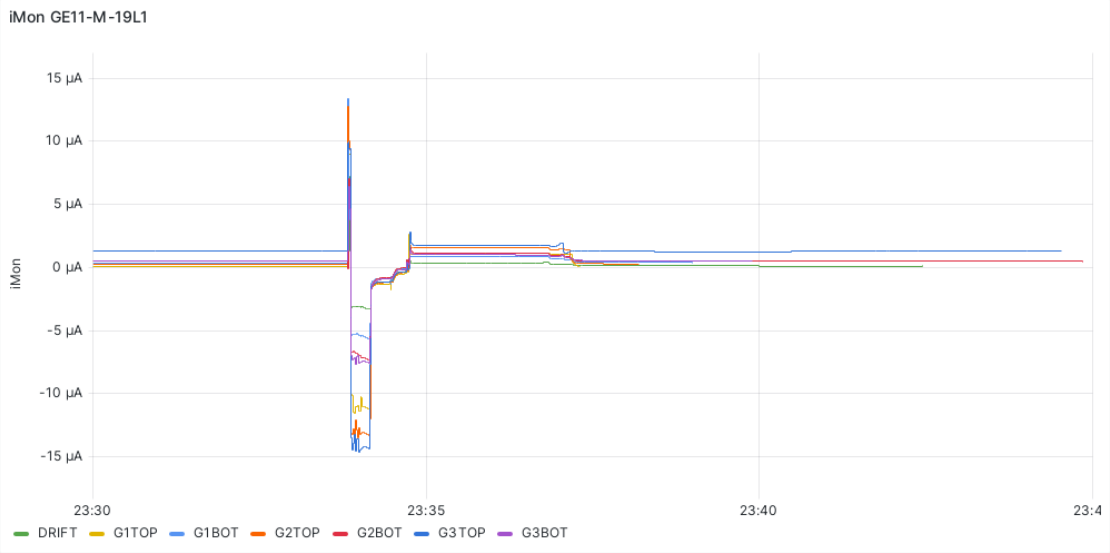


Figure 5.8: High-voltage discharge recorded in the 10th of August 2024 on chamber GE-1/1/19 Ly1.

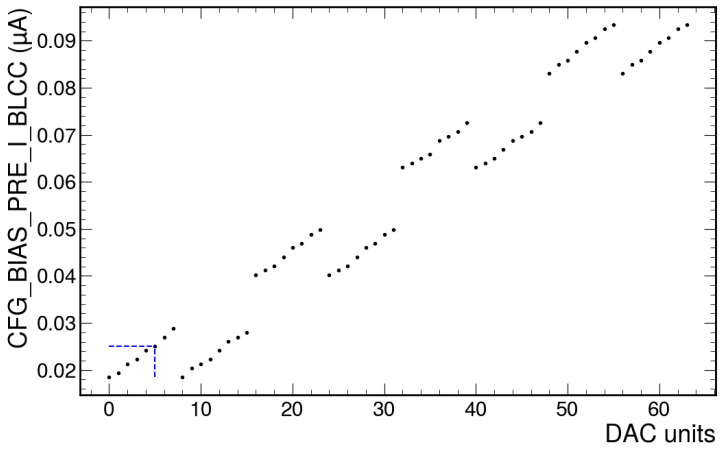
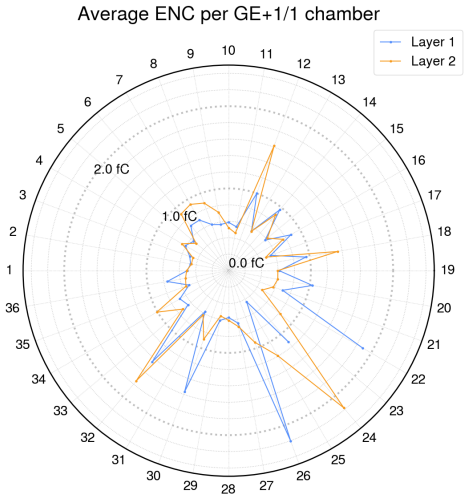
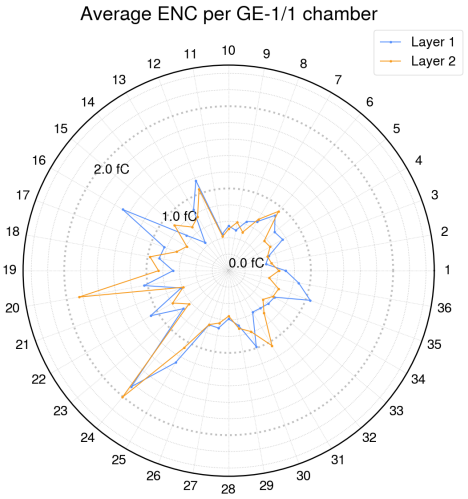


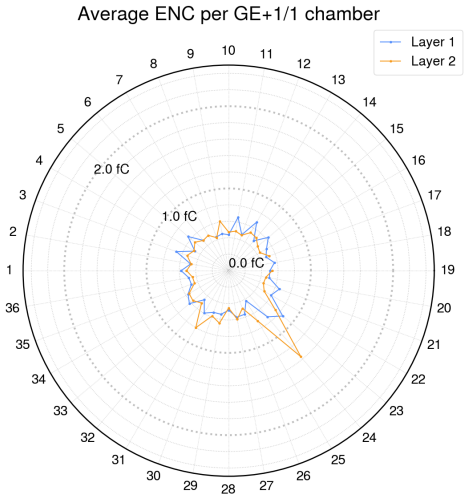
Figure 5.9: Analysis results of the DAC *BIAS_PRE_I_BLCC* scan for VFAT #0 on chamber GE-1/1/19 Ly1 taken on the 12th of August 2024. The pattern suggests broken bits in the biasing DAC circuit.



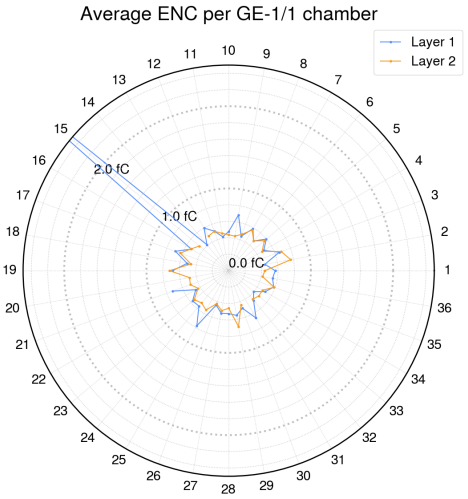
(a) Before mitigations - Positive endcap



(b) Before mitigations - Negative endcap



(c) After mitigations - Positive endcap

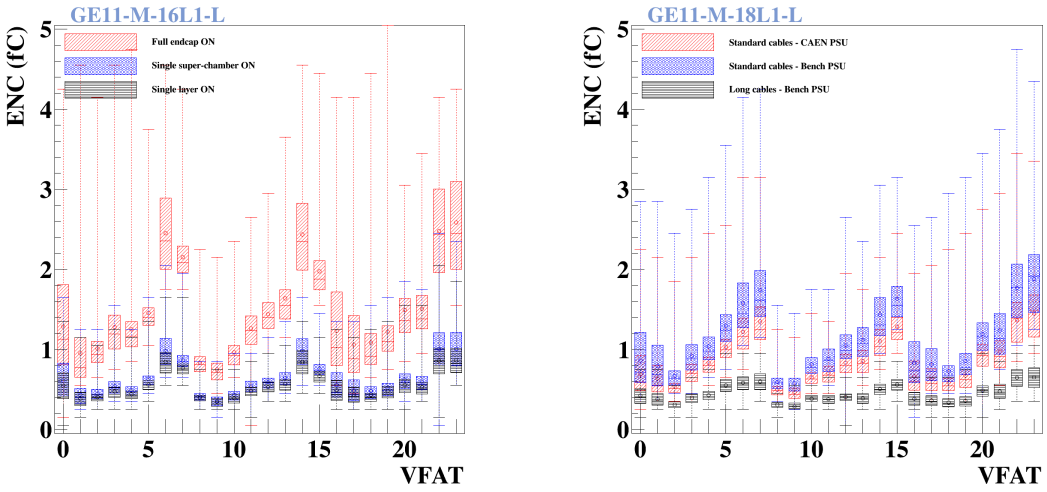


(d) After mitigations - Negative endcap

Figure 5.10: Average ENC, extracted from the S-curve width, for all GE1/1 chamber before and after the implementation of the mitigation strategies. The wheels are oriented as if they were looked from the interaction point. Chambers that could be scanned show no data.

Amongst those factors, no impact was found except for the low-voltage system. At this stage, it was very clear that the high noise levels were due only to a group effect in the lower-voltage system. As Figure 5.11a shows, the noisy behavior was observed only when a significant number of GE1/1 chambers were powered on. The larger the number, the higher the noise.

The symmetry between the positive and negative endcaps (Figures 5.10b and 5.10a) gave a strong hint that the problem could be due to the services routing. Further studies performed on the low-voltage cables showed that the standard cables routed on the CMS disk and nose were the cause of the electronics noise. Figure 5.11b compares the standard cables with a long ($\mathcal{O}(20\,\mathrm{m})$) flying cable hanging to the CMS cavern floor. The long flying cable does not show any noise increase with respect to the QC laboratory even when powering the full detector.

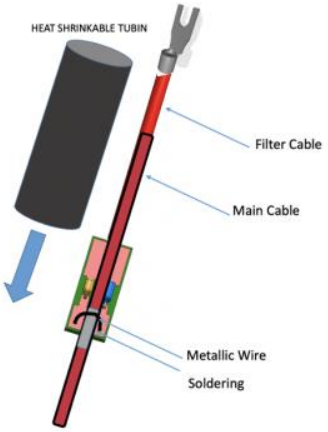


(a) Group effect on chamber GE-1/1/16 Ly1. (b) Cables effect on chamber GE-1/1/18 Ly1.

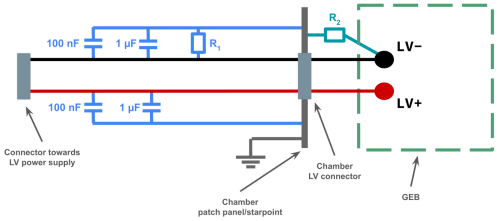
Figure 5.11: Equivalent noise charge comparison candle plots highlighting in (a) the presence of a group effect when the detector low-votlage is fully powered on; and in (b) the effect of the low-voltage cables. The front-end electronics becomes noisy when using the standard LV cables and powering the whole detector.

All attempts to characterize the noise better failed. The only established fact is that high-frequency noise is induced via the LV cables. It was then decided to develop and test a low-pass filter to be installed before each GE1/1 chamber. The production version is shown in Figures 5.12. The capacitor values, defined entirely empirically, showed the best results, restoring the QC laboratory noise levels.

(a) Picture of a production part.



(b) 3D schematic of the filter assembly.



(c) Functional representation of the filter.

Figure 5.12: Representations of the GE1/1 low-voltage filter. The filter is installed between the LV power supply cable and the chamber's LV connector to reduce the electronics noise to acceptable levels. The grounding of the LV- line is improved using a $2.5\,\mathrm{mm}^2$ cable (represented by R_1), directly connected to the chamber start point at the patch panel. This constitues a significant improvement over the original internal grounding using a $< 1\,\mathrm{mm}^2$ cable (represented by R_2) The capacitor installed in parallel to R_1 likely have no effect. A low-pass filter is implemented on the LV- line using two capacitors connected to ground via the chamber star point.

Overall, three modifications were implemented. The first one consists of the LV filters already mentioned and remains the dominant improvement factor. The second one, developed with the help of CAEN engineers, consists of the addition of a capacitor inside the power supply to further filter out the common mode noise. The third one consists of the LV cable shielding connection to both the power supply side and the chamber side. While connecting both sides to the ground should create a ground loop and increase the noise levels, that behavior was not observed. It is thought that the limited shield layer conductivity and the length of the cables essentially break the ground loop. This would also explain why the shielding layer appears ineffective in both preventing the noise from escaping the cables and from being induced in the cables.

Figures 5.13 present the effect of these mitigation measures. The noise levels are dramatically reduced on originally noisy chambers, reaching less than 1 fC for all VFAT. On the other hand, the noise levels are not increased in well-behaved chambers, proving that no noise degradation is incurred.

In May 2021, after implementing all modifications, a significant improvement in the noise level was measured for most VFAT chips in the GE1/1 system (>95 %) reaching an ENC below 1 fC. For the noisiest chambers, this improvement enabled a threshold reduction from approximately 12 fC to around 4 fC, corresponding to efficiency gain of more than 5 percentage points.

Figures 5.10d and 5.10c present an overview of the final noise level in the GE1/1 system. It is hypothesized that chambers GE-1/1/15 Ly1 and GE+1/1/24 Ly2 are noisier due to the incorrect installation of the LV filters. For the former, because of the low accessibility of short chambers in the CMS nose. For the latter, some connections may have been damaged as this chamber was frequently used for noise studies.

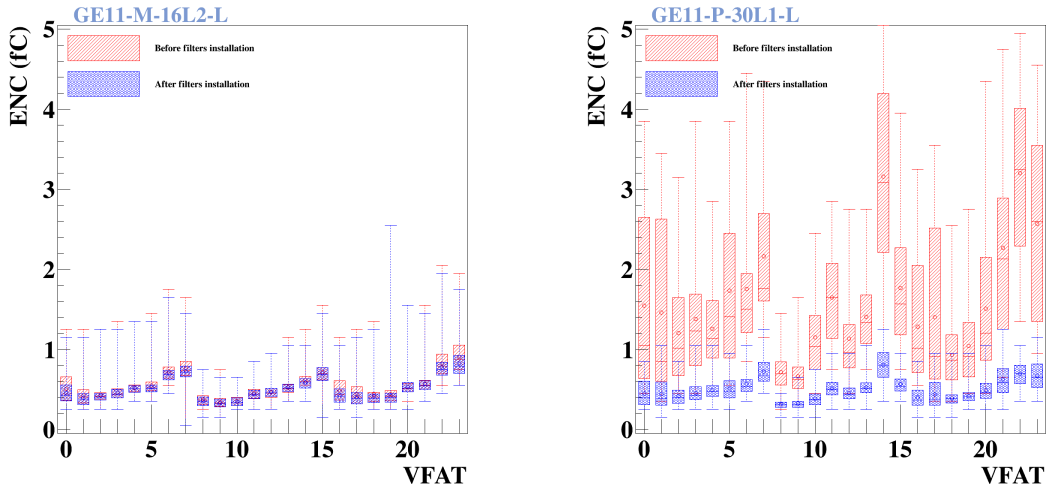
5.5 Latency scans

The very first latency scan was performed in global data-taking with cosmics during the MWGR in October 2020. On this occasion, the full latency range permitted by the Phase-I CMS Level-1 trigger was covered to ensure the success of the scan. It led to the first signs of in-time particles detected in the GE1/1 station, completing a major milestone for the CMS GEM project! Figure 5.14a shows a typical plot of those results with a clear peak around 155 BX.

The latency scan and analysis workflows were subsequently refined to reach an acceptable per-chamber precision in the MWGR of April 2021. To that end, an empiric filtering algorithm was put in place allowing an easier identification of the latency peak:

Removal of the noisiest channels At the time of the scan, no online masking of the noisiest channels was implemented, requiring the implementation of a post-processing workflow.

Removal of the channel with a large number of hits While entirely empiric at the beginning, this filter was later justified by the cross-talk phenomenon (see later) and the high-multiplicity events (Section 5.6).



(a) (b)

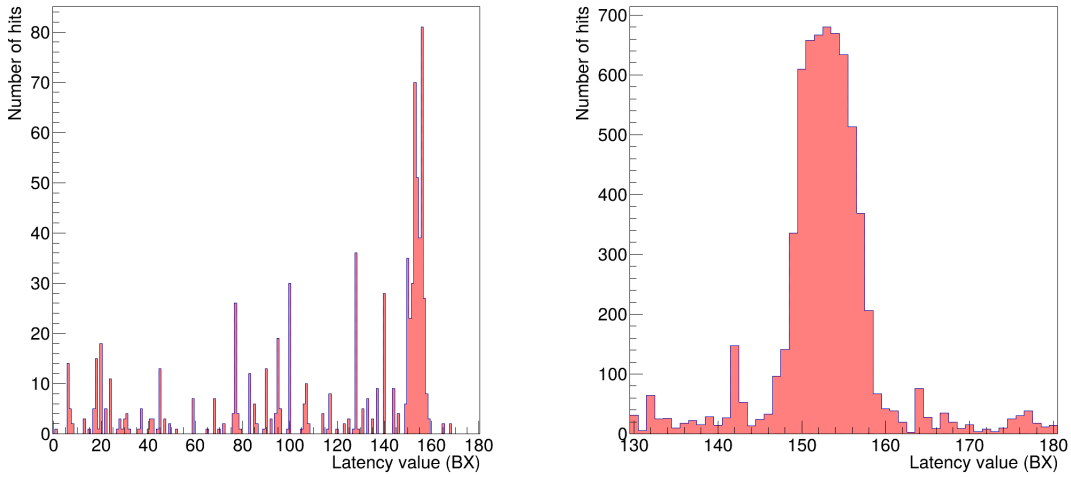
Figure 5.13: Equivalent noise charge before (red) and after (blue) the implementation of all mitigation measures for a quiet chamber (a) and a noisy chamber (b).

The results of a typical narrow latency scan after post-processing are presented in Figure 5.14b. The ideal latencies were determined to be in the range $[150, 155]$ in April 2021.

Over the years, the latency scan procedure was further refined and adjusted to time in GE1/1 during proton-proton collisions. The GEM system typically takes advantage of the LHC van der Meer (VdM) scans [19] to perform its own latency scans. Those fills, aiming at luminosity calibration, provide very clean signals thanks to the large colliding bunches spacing – usually every 525 ns – and the low background. Additionally, the Level-1 trigger menu is well suited to the purpose with almost exclusively zero bias triggers on the colling bunches. This is of particular interest since the latency scan routine uses the local readout and thus completely bypasses the HLT selection algorithms. Enough statistics can be gathered to allow taking data with a pulse length of 1 BX and analyzing the data with a per-VFAT granularity (Figure 5.15).

The careful reader will notice two secondary peaks beside the ideal one at 162 BX. The first one is at a latency of 141 BX. It can be explained by the 525 ns spacing of the fill: this signal coming from the next colliding bunch.

The second one is at a latency of 155 BX in the ”normalized number of hits” green histogram. It corresponds to the so-called cross-talk effect [83]. Cross-talk signals consist of parasitic signals of opposite polarity induced on the electronics front-end channels facing the same HV sector as the one with the real particle hit. In the case of Highly-Ionizing Particles (HIPs), the undershoot of the large amplitude signal can cross the threshold and be recorded as a fake hit. Optimized foil design will limit the efficiency



(a) Broad latency scan for GE-1/1/29 Ly2.

(b) Narrow latency scan for GE-1/1/32 Ly2.

Figure 5.14: Readout latency scan results taken during the GE1/1 station commissioning with cosmics. The left plot is taken from the very first latency scan performed on GE1/1 (run 337973) and shows the first muons ever recorded in the station. The right plot is produced with data from run 341288 providing a narrower latency range, hence higher statistics, and post-processed data. The large width of the peak is due to the VFAT3 pulse stretch set to 7 BX.

loss below 1%, representing a compromise between effectively reducing the discharge propagation probability while keeping the cross-talk under control. This consideration is the main factor behind keeping the GE2/1 and ME0 bottom foils single-segmented – like for GE1/1 – despite their worst discharge propagation properties (see Section 2.3.1). The higher the capacitance of the electrode facing the readout strips, the lower the cross-talk.

5.6 High multiplicity events

The first runs at a high Level-1 trigger rate ($\mathcal{O}(\sim 50\,\text{kHz})$) showed the presence of high-multiplicity events in the GE1/1 data. Those events are characterized by the presence of thousands of large clusters, in opposition to regular events containing less than 100 clusters, each spanning over 2.7 strips on average. In the CMS and GEM jargon, the wording *flower event* is often used, describing accurately the typical pattern those events exhibit on an event display (Figure 5.16). A posteriori, precursor signs could have been seen in the latency scans where events with large multiplicities had to be filtered out or during early EMTF trigger testing where GEM hits were systematically seen synchronously after an L1A.

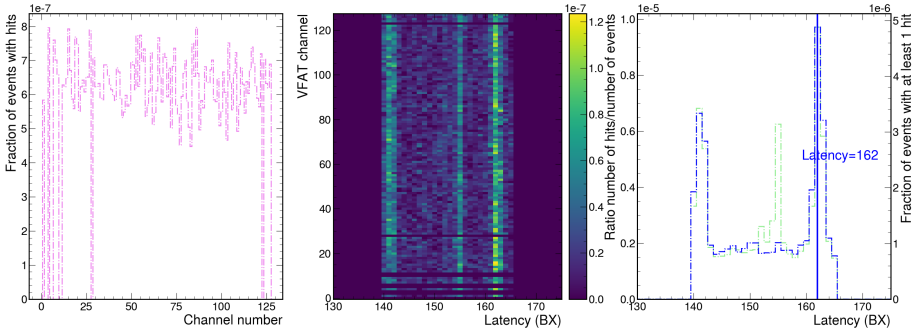


Figure 5.15: Most recent latency scan for VFAT #20 in chamber GE+1/1/23 Ly2. The left histogram represents the per-channel occupancy distribution; the center plot represents the fraction of events where a given channel recorded a hit at a given latency; the right histograms show the fraction of events with at least one hit (blue) and the total number of hits normalized by the number of events (green). The narrow peaks are achieved by configuring the VFAT3 pulse stretch to 0 BX. Data taken during LHC fill 9639, totalling 1.19 pb^{-1} on the 17th of May 2024.

Investigations were immediately triggered on the test system in the laboratory to understand better the phenomenon. As the issue seemed related to the Level-1 trigger rate, the first idea was to perform a latency scan over the whole VFAT3 latency range in the following conditions:

- Keep the HV powered off
- Ensure that calibration pulses are disabled (both the per-VFAT and per-channel calibration circuits)
- Send triggers with a periodic of 1000 BX
- Use a VFAT3 configuration similar to the production system in CMS

In the absence of amplification and calibration pulses, no signal is expected and the latency scan should be flat. However, that is not what was observed. As Figure 5.17 shows, there is a clear signal peak ~ 10 BX after the L1A is received by the VFAT.

Investigations were also carried out on the data from the GE1/1 detectors installed in CMS. It is natural to look at the number of reconstructed hits (RecHits) as a function of the delay between consecutive L1A. This yields to the histogram shown in Figure 5.18. Once again, the time between consecutive L1A should not have impacted the hits multiplicity. Instead, it surges at a delay of around 160 BX.

This value is explained by the chronogram depicted in Figure 5.19. After each L1A, it takes about 10 BX to form a spurious signal within the VFAT3. The signal will however

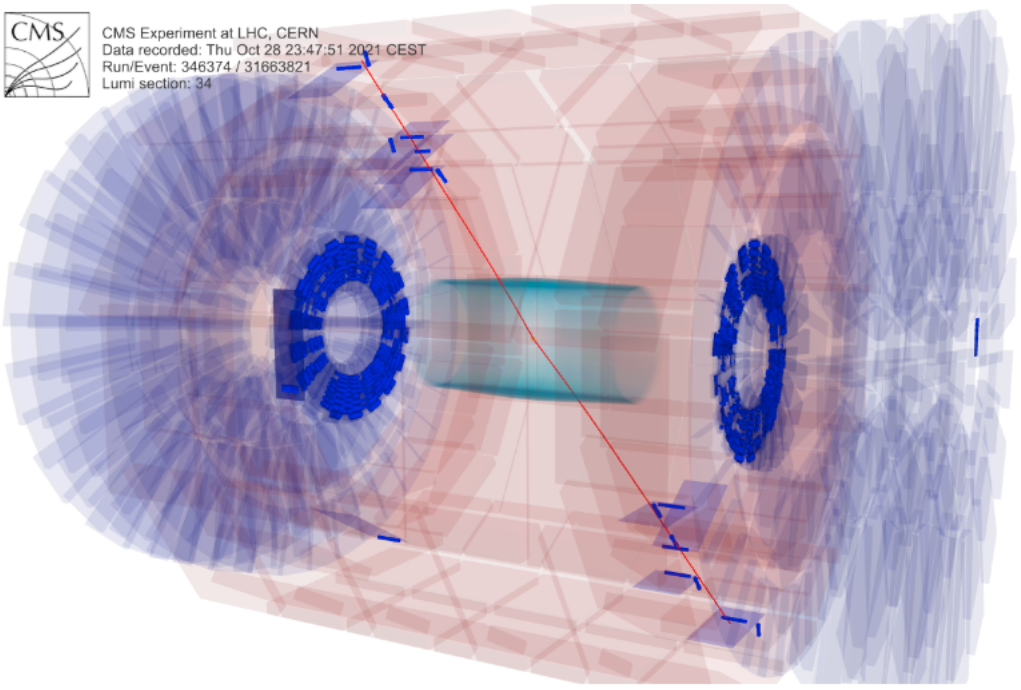


Figure 5.16: CMS event display showing a typical "flower event". A pair of muons, in red, originating from the interaction point cross the CMS barrel. The almost complete GE1/1 station is lit up (in dark blue) in a pattern reminding us of a flower.

be recorded in the tracking data only in the events triggered NBX later, where N is the readout latency configured in the VFAT. All in all, the high-multiplicity events can be observed in the tracking data if and only if the following condition is met:

$$\Delta\text{L1A} = \text{signal formation time} + \text{readout latency} \tag{5.1}$$

Of course, the exact delay depends on the VFAT since each of them can be configured with a different latency.

The same experience can also be reproduced up to the 4th previous L1A. Past this number, the Phase-I CMS trigger rules do not allow any L1A to meet the visibility condition.

The root cause was eventually traced to a design issue within the VFAT3 ASIC design. Whenever the VFAT3 receives an L1A signal, hits data are retrieved from SRAM1 and are written to SRAM2 for buffering before transmission on the tracking path. This operation causes an instantaneous current draw in the digital power domain that induces noise in the analog power domain via the ASIC substrate. The additional synchronous noise is estimated to be around 0.8 fC. Considering that the noisiest channels drive the threshold when untrimmed, only the noisiest channels will react to the additional induced noise. On the contrary, when trimmed, the noise response is similar between

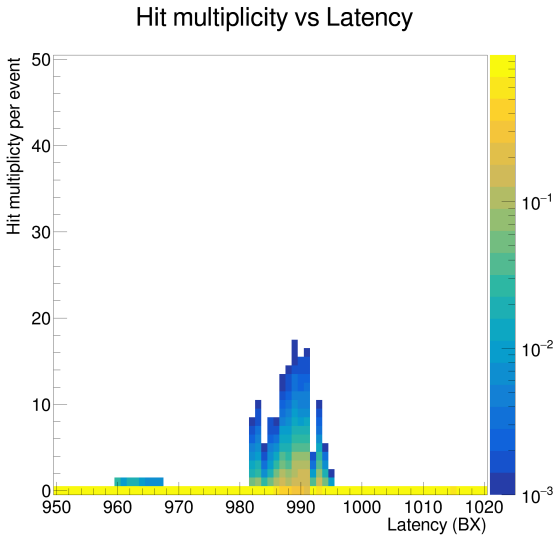


Figure 5.17: Hit multiplicity as function of the readout latency for VFAT #11 with triggers sent periodically with an interval of 1000 BX and without high-voltage or calibration pulses. The peak around latency 990 correspond to the high-multiplicity events.

all channel and thus it is expected that all channels will fire. The time between the L1A signal and the spurious signal is 9 BX = 225 ns, as shown in Figure 5.20, matching the observations done previously on the GE1/1 detector. However, at this stage of the project, no mitigation could be implemented in the ASIC.

Mitigations measures

As this defect causes large number of fake hits both in the tracking and trigger paths, mitigation measures must be put in place. This is particularly true in the trigger path where all bunch crossings are processed and no selection on the events is made – each L1A will necessarily lead to at least one bunch crossing with very high occupancy. Not only it could impact the triggering algorithms, but it also leads to transmission bandwidth overflows.

A few strategies are under investigation and evaluated both in terms of effectiveness and impact on the data quality:

1. Increase the thresholds to cut the L1A induced noise
2. Mask the RecHits in the reconstruction, either timing- or pattern-based
3. Mask the S-bit clusters in the trigger path, either timing- or pattern-based
4. Implement a trigger-less readout

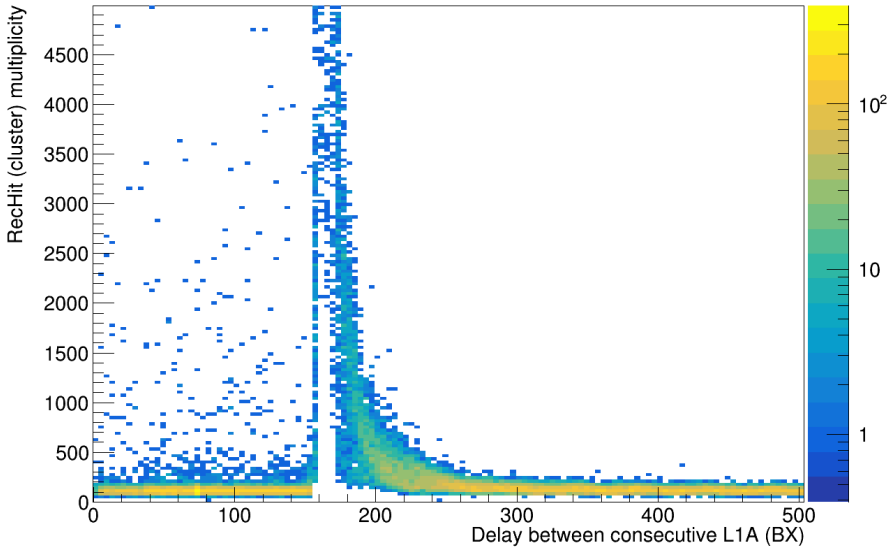


Figure 5.18: This plot shows the number of GE1/1 RecHits as a function of the delay, in LHC bunch crossings, between two consecutive Level-1 Accept. The RecHit multiplicity increase around a delay of 160 BX is attributed to the flower events. Produced with 138702 events from run 346247.

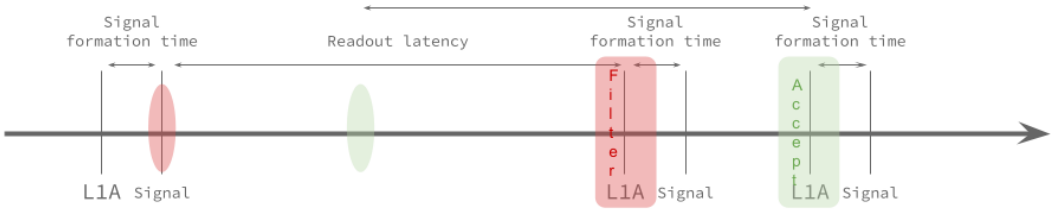


Figure 5.19: Chronogram explaining when high-multiplicity events can be observed in the tracking data path. The data saved by the second L1A will include the synchronous noise induced by the first L1A.

5. Implement regional triggering

The pattern-based approaches take advantage of knowing the full event characteristics to detect, and mask, the events resembling high-multiplicity events. Typically, a cut on the number of GEM hits and/or clusters can be used as a filtering method. The timing-based approaches use only local information and thus can easily be implemented within an FPGA and be applied in real-time with tight constraints. They stem from a deep understanding of the high-multiplicity events source.

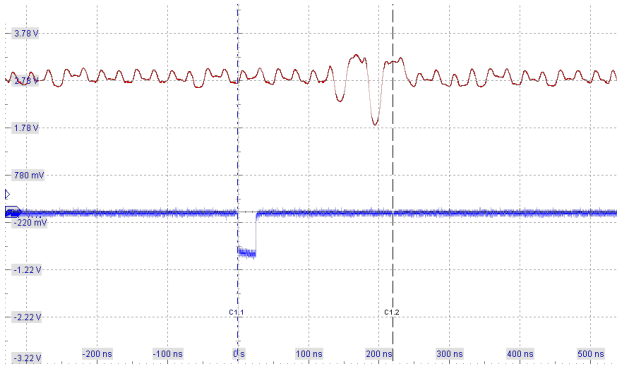


Figure 5.20: Oscilloscope capture showing the timing between the VFAT3 Level-1 Accept signal (blue) and the noise induced at the output of the charge-sense amplifier of a test channel (red). Credits to the GEM Electronics Group.

During the LHC Run 3, the S-bit clusters will be masked in the trigger path based on their timing with respect to the L1A. For the tracking path and the reconstruction, the GEM DPG group is still investigating the best approach. However, the time-based RecHits masking method is already implemented for some analyses (e.g. background rate measurements).

New OptoHybrid firmware and online software features have been developed to implement the trigger path masking. In firmware, the so-called cluster mask, is responsible for canceling all S-bit clusters for any configurable combination of BX (in a window of 32 BX) after it receives the L1A signal. In software, the so-called cluster mask scan, has been implemented. It measures, BX by BX, the impact of the cluster mask on the rate of clusters built by the OptoHybrid firmware. The rates are computed, simultaneously, in the presence and in the absence of masking while generating a 100 kHz periodic L1A signal.

If the BX corresponds to a high-multiplicity event, a drastic reduction of the cluster rate is observed. On the contrary, if the BX does not correspond to a high-multiplicity event, the reduction will be marginal. Typical results of the scan are shown in Figure 5.21. In order to avoid creating too much dead time, 5% of high-multiplicity events remain accepted.

With the current implementation the average dead time introduced per GE1/1 chamber in the trigger path amounts to 5.91 BX per L1A, corresponding to 1.8% at a rate of 120 kHz. Future improvements include the implementation of a per VFAT S-bit masking, aiming at reducing the number of masked BX and reaching the smallest dead time possible.

For the LHC Run-4 and beyond, the trigger-less readout and regional triggers options are being evaluated, as the increased L1A rate of 750 kHz would results in a dead time above 10% (assuming the number of masked BXs cannot be reduced compared

to the current conditions). The former option can be implemented only in the GEx/1 electronics whereas the latter option is best suited for the ME0 station which locally builds L1T stubs.

The trigger-less readout option uses advanced features of the VFAT3 to transmit full granularity trigger data over S-bit lines at twice the nominal data rate – known as DDR (Double Data Rate) mode. If the GEB can be qualified to handle these higher bit rates, the OptoHybrid firmware then needs to be enhanced to provide event-building features. In this case, no L1A signal is sent to the VFATs. And, if no L1As are sent to the VFAT, no high-multiplicity events are generated.

The regional triggers option consists of sending the L1A signal only to those VFATs susceptible to record a muon track. In the ME0 station, this can be achieved by issuing L1As only in the areas where a L1T stub is generated. Nevertheless, this method must be complemented by another as high-multiplicity events will still be produced. In the case where the regions are defined with an eta partition granularity, the L1A rates in the innermost eta partition are expected to drop to 3%, corresponding to an inefficiency of 0.5%, and a further reduction factor of six is anticipated in the outermost partition.

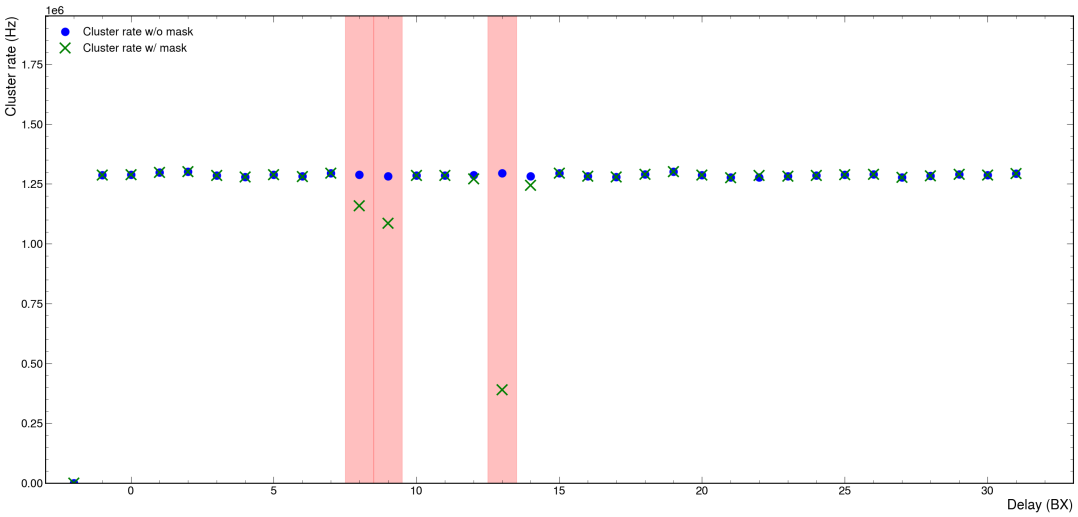


Figure 5.21: Cluster mask scan results for chamber GE-1/1/14 Ly2 taken with an L1A rate of 100 kHz. A delay of −1 BX means "no L1A and not masking applied". A delay of 0 BX mean "no masking applied". The delays colored in red will be masked, allowing up to 5% of high-multiplicity events in the trigger path.

5.7 Efficiency measurements

After completing the initial commissioning steps, as well as the time and spatial alignments, the efficiency of the GE1/1 detectors is carefully studied using CMS global runs. This is achieved by measuring and reconstructing muon tracks in the other CMS subdetectors, and, then, propagating the said tracks to the GE1/1 chamber surfaces. Next, a corresponding GE1/1 hit is searched for in the vicinity of the propagated hit. If such a hit is found, it is called a *matched hit*. The detection efficiency is thus defined as:

$$\text{efficiency} = \frac{\#\text{ matched hits}}{\#\text{ propagated hits}}$$

5.7.1 Dataset

Evaluating the performance of the GEM detector requires the usage of the appropriate dataset. The datasets are organized within the CMS experiment in multiple data tiers, depending on the included content [22]. For our analysis, two data tiers are of particular interest as they provide access to both the detector hits and the muon tracks:

RAW which stores unprocessed information directly from the detector FED as well as the HLT decisions;

RECO which stores completely reconstructed events, including detector hits and clusters, and physic objects and tracks.

Within those two data tiers, the choice of a specific dataset is driven by the presence of a muon content large enough to provide low statistical uncertainty. The method developed most recently and presented here uses the `RPCMonitor` dataset. It consists of the raw FED data coming solely from the muon detectors and muon trigger processing boards, and is seeded by the logical *OR* of the following Level-1 Trigger algorithms: `L1_SingleMu5`, `L1_SingleMu7`, `L1_SingleMu9`, `L1_SingleMu18`, `L1_SingleMu20`, `L1_SingleMu22`, and `L1_SingleMu25`. As their names suggest, they require the presence of a single muon candidate with at least the given p_T (given in GeV/c).

The restricted content of the dataset allows for a low prescale, which was set to 20 in 2023-2024, leading to a rate of about 450 Hz. If necessary, the prescale can also be temporarily removed for dedicated studies requiring higher statistics. This was done in 2024 during the so-called HV scan (see Section 5.7.4).

5.7.2 Workflow

The method described in this section follows a traditional two-step workflow. The first step uses the C++-based CMSSW framework to convert the dataset made available after processing at Tier-0 into a lightweight N-tuple. During this process, muon reconstruction from the raw FED data is performed, and only the most skimmed output is saved for further processing. Little-to-no selection logic is applied at this stage, minimizing the need to reprocess from scratch an old dataset.

The second step runs on top of the N-tuple and is written in Python using columnar analysis for both flexibility and computing efficiency reasons. All data selections and matching criteria described further are implemented at this stage.

In order to simplify the operations and reduce the load on the operator, automated workflows driven by the CMS ECAL automation framework [87] have been implemented. This is particularly beneficial as the `RPCMonitor` dataset remains on disk for only a few days, and requesting data from tape archive would significantly impact the feedback timeline.

The automation framework is based on a set of industry-standard applications deployed in the general-purpose OpenShift instance provided by the CERN IT: Jenkins [73], Grafana [63], and InfluxDB [71]. Jenkins is used to periodically schedule and monitor the execution of all workflows that are part of the automated analysis. InfluxDB serves as a bookkeeping database that records the status of all workflows, tasks, and jobs. Grafana provides a convenient user interface to visualize the detailed status of each workflow, and is fully integrated with the online monitoring infrastructure described in Section 3.6.

Using the batch processing resources available at CERN, the following performances are reached:

- Prompt – The processing delay is less than 3 days, that is less than 24 hours after the files become accessible to the users.
- Reprocessing – In case of bugs or improvements, a full reprocessing is completed in less than 8 hours. Debugging of the analysis can be carried on during the day, and the complete set of results will be available the next morning, leading excellent turnaround capabilities.

5.7.3 Analysis methodology

Muon selection

The muon reconstruction in CMS defines two types of muons. The *standalone muons* are based solely on the information from the muon subsystems outside of the solenoid volume. The *global muons*, typically used for physics analyses, are the combination of a standalone muon with a track from the inner tracker. The `RPCMonitor` dataset content we use here restricts the choice to the standalone muons, as no tracker information is available. This restriction has a negative impact on the p_T resolution as well as purity. To mitigate the impact on the p_T resolution, the standalone muon tracks are further refined with the position of the beamspot in this analysis [59].

To ensure quality, the selection criteria shown in Table 5.1 are applied to the muon tracks.

For this efficiency analysis, we also ensure that the GEM hits are excluded from the muon reconstruction to avoid any bias.

Table 5.1: Muon selection criteria used for the `RPCMonitor`-based efficiency analysis using *standalone muons*.

Cut	Value
Muon transverse momentum p_T	$> 10\,\text{GeV}$
Propagation error in global R	$<= 1\,\text{cm}$
Propagation error in global φ	$<= 0.005\,\text{rad}$
Fiducial cut in local Y	$1.5\,\text{cm}$
Fiducial cut in local φ	$0.0075\,\text{rad}$
Track reduced χ^2	$<= 5$
Number of hits used in the track	$>= 15$
Stations contained in track	Must contain an hit in the adjacent CSC station[1]

Muon propagation

To determine the crossing of a muon track with a GEM detector surface, each standalone muon track is propagated to all GEM surfaces using the *Steppping Helix Propagator* – the default propagation algorithm used in CMS for the regions outside of the tracker volume. This method takes into account the magnetic field geometry as well as the effects of multiple scattering and energy losses in the traversed material [103].

However, this propagation implies uncertainties in the predicted muon position. These are due both to initial uncertainties on the initial track parameters (position and momentum) and to the stochastic nature of the particles-material interactions. As a result, if the propagated hit lies near the edge of a surface, it only has a limited probability of occurrence and cannot be reliably associated with the surface.

To mitigate this effect, the peripheral region of each eta partition is excluded from the efficiency analysis. This exclusion zone – referred to as the *fiducial cut* – ensures that only well-contained tracks are considered. The fiducial cut size is conservatively chosen to account not only for propagation errors, but also for potential detector misalignments, which would result in a similar effect.

Additional selection criteria

On top of the muon selection criteria, two additional selections are optionally applied. Together, they allow measuring the bare detector efficiency, discarding any inefficiency due to the data-taking conditions. They constitute a key element to establish perfect reproducibility between runs.

High-voltage masking It is a normal behavior for the GEM detectors' high-voltage to trip during operations as a consequence of a discharge. During such events, the detector obviously becomes totally inefficient. The high-voltage masking procedure

[1] I.e. ME1/1 for GE1/1 and ME2/1 for GE2/1.

thus removes data from the events for which the monitored voltage deviates from the set voltage by more than 5 µA equivalent divider current. It is applied for each chamber independently with a per-lumisection granularity.

DAQ error masking As the electronics exhibit significant communication instabilities, the active part of the detector varies run-by-run, and even event-by-event within a given run. It is desirable to be able to compare the results from one run to another, so this behavior must be handled appropriately. In the offline analysis performed here, the data from a whole chamber is discarded if any error is recorded for any of its VFAT. This is applied with a per-event granularity in order to maximize the statistics while being robust. Masking the whole chamber was deemed acceptable, even though discarding the data only from the affected VFAT is more conservative from the statistical point of view.

Hit matching

Hit matching between propagated hits and recorded hits, or `RecHits`, is performed by associating each propagated hit with the closest recorded hit within the same eta partition in a recursive fashion. In other words, even if two propagated hits are within the matching distance of a single `RecHit`, only the closest would be associated. The farthest propagated hit would be associated with the second closest `RecHit`, if any.

Once the overall association is complete, the pairs of hits are considered matched if their distance is below 4 cm:

$$R\Delta\varphi <= 4\,\text{cm} \tag{5.2}$$

This value has been chosen as it does not introduce significant bias - by matching unrelated hits - and it can cover for residual spatial misalignment of the chambers.

5.7.4 Results

Figure 5.22 displays the efficiency produced with a per-VFAT granularity for a representative chamber. Similar plots are produced for each chamber and for each run by the automation framework. They are then accessible to the operator for further analysis and performance monitoring.

Working point optimization

While all GE1/1 detectors have been qualified with an efficiency greater than 97%, each detector requires a specific set of parameters to achieve optimal performance [62]. One of the primary goals of the commissioning process is, therefore, to determine the best set of parameters for each detector.

In Spring 2024, a so-called *HV scan* was performed. The term *HV scan* is actually an understatement: it involved a complete study of the GE1/1 efficiency, including variation of the VFAT3 front-end parameters, with a particular focus on the three preamplifier gain settings.

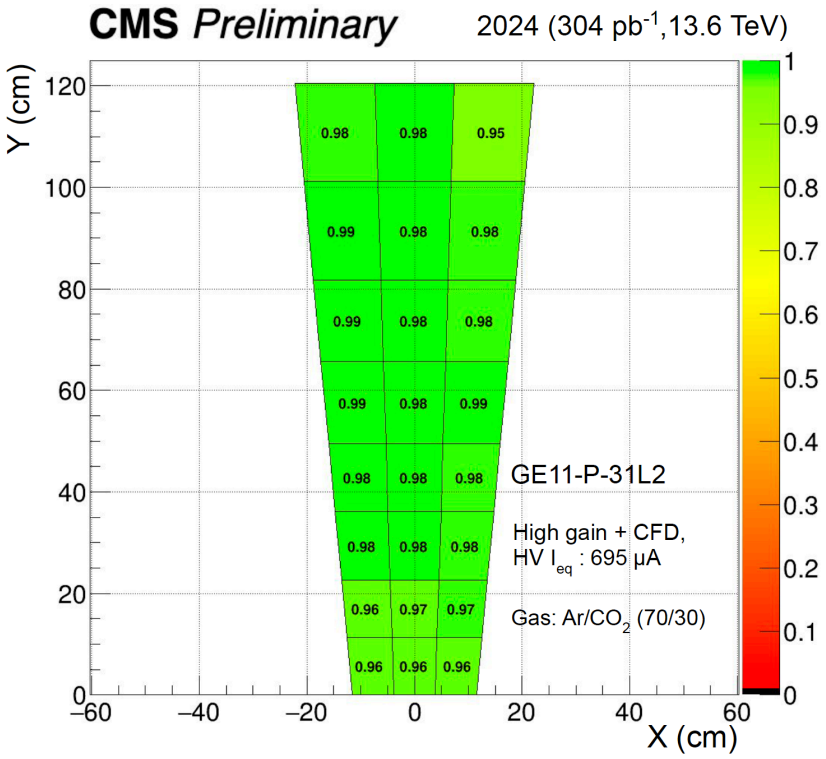


Figure 5.22: Per-VFAT efficiency map for chamber GE+1/1/31 Ly2. [61]

Results for two representative detectors are shown in Figure 5.23. The detector on the left exhibits an ideal behavior, whereas the detector on the right exhibits a lower performance, failing to reach its design efficiency even at the highest operating points.

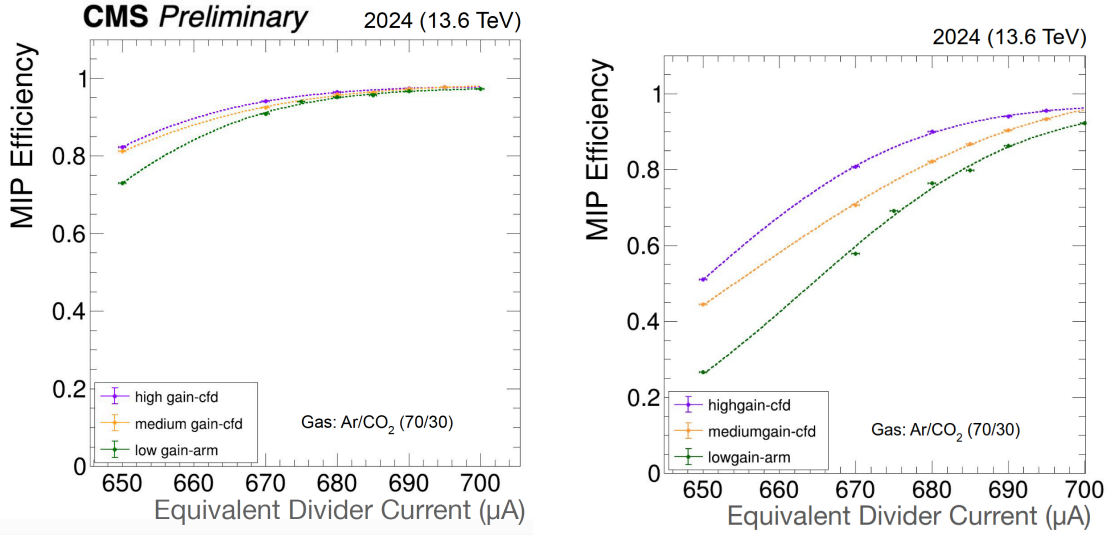
Each curve is fitted with a sigmoid curve:

$$C \times \mathrm{erf}\left(\frac{I_{eq} - I_{50}}{W}\right) + 1 \tag{5.3}$$

where I_{eq} is the operating equivalent divider current, I_{50} is the equivalent divider current at 50% efficiency, W is the width of the curve, and C is a normalization constant. The same procedure is applied individually for each chamber and serves as the base of the working point optimization.

For simplicity reasons, it was decided to configure all VFATs across all chambers in high-gain mode. This choice systematically yields the highest efficiency at a lower HV working point, thereby increasing the chamber stability, without any side effect. The high-voltage working point is then adjusted per-chamber and chosen as the lowest equivalent divider current that ensures full detection efficiency.

As discussed in Section 2.2.5, the HV system has a key limitation: each channel powers two detectors in parallel. Therefore, to fully exploit each detector's potential, a hardware remapping was performed so that two detectors with the same HV working point were paired on a given HV channel.



(a) GE+1/1/31 Ly1 (b) GE+1/1/16 Ly1

Figure 5.23: High-voltage scan results for two representative GE1/1 detectors. A complete scan has been performed for a sub-set of the VFAT3 pre-amplifier gains and discriminator modes: low gain in arming (ARM) mode (green), medium gain in constant fraction discriminator (CFD) mode (orange), and high gain in CFD mode (purple). Detector (a) reaches its maximal efficiency around an equivalent divider current of 695 µA regardless of the VFAT3 operation mode; detector (b) reaches its maximal efficiency at an equivalent divider current of 700 µA only if the VFAT3 is configured in medium or high gain.

The impact of the optimization can be seen in Figure 5.24. The orange histogram represents the non-optimized configuration used during the 2023 data-taking, while the blue histogram represents the optimized configuration obtained by the end of Summer 2024. Factoring out the electronics-related issues and instabilities, the average GE1/1 efficiency after optimization reaches 94%, approaching the 97% design target. Further improvements - such as enhanced chambers pairing, refinements in the readout latency, and VFAT3 channel trimming - are expected to close the gap in 2025.

However, the electronics troubles cannot be addressed before the next long shutdown of the LHC. For instance, chamber GE-1/1/15 Layer 1 exhibits a very low efficiency (close to 25%) due to high readout thresholds required to suppress noise hits.

High-granularity plots

The implemented analysis workflow enables high-granularity measurements of the detector efficiency at no additional cost. This is particularly the case when the prescale of the `RPCMonitor` dataset is reduced to 0.

A typical output plot is shown in Figure 5.25. It important to note, however, that

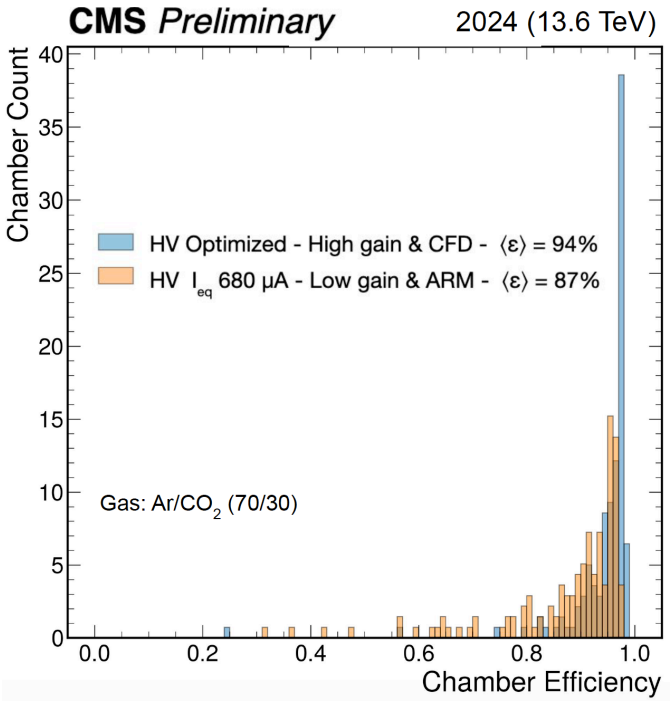


Figure 5.24: Efficiency of the GE1/1 chambers with HV I_{eq} at 680 µA and low gain for the VFAT3 chip settings in a run of 408 pb^{-1} (orange), and after the initial optimization study of the VFAT3 front-end chip configuration and HV settings in a run of 935 pb^{-1} (blue). The efficiency is calculated for chambers properly communicating and operating at nominal HV settings. [61]

the statistical significance of such plots is limited, as each individual bin may contain as few as 10 propagated hits. Nevertheless, these plots remain a precious diagnostic tool to understand various detector features:

Bassin effect The so-called *bassin effect*, present primarily – but not exclusively – on the long (evenly numbered) GE1/1 chambers, is characterized by an efficiency non-uniformity localized around VFAT14, which is located in the central region of the detector. This effect has been thoroughly studied and is attributed to a physical curvature of the chamber drift and readout PCBs [62]. The bending of the PCBs results in three main consequences: (1) a reduced drift field lowers the collection efficiency of the first foil, (2) a reduced induction field lowers the extraction efficiency of the third foil, and (3) the weakened induction field stretches the GEM signal on the readout strips, making it too long to be fully integrated by the VFAT3 analog front-end. Combined, these effects drastically reduce the local efficiency. While increasing the electric field and lowering the front-end thresholds can partially mitigate the effect, a complete correction would require a new detector design.

Unstable GBT link The L-shaped white patterns observed on the chambers in sectors 26 and 34 are characteristic of GBT communication instabilities.

Masked channels To enable the use of the lowest possible front-end thresholds, the noisiest channels must be masked. Since the dominant source of localized noise is capacitive coupling between the strips and the ground plane, the VFAT channels located at the edges of a readout sector tend to be noisier. This results in recurring patterns of masked channels, observable as two narrow lines at constant phi on most detectors – for example, clearly visible in sector 20).

Shorted or disconnected HV sectors As discussed in Section 2.2.2, the GEM foils are segmented for detector protection purposes. In the event of a shorted foil segment, the rest of detector can continue operating normally. However, the affected segment records an efficiency close to zero. This failure mode appears as a straight, narrow band with low efficiency (below 50%) across a specific eta partition – for example visible in sector 18 or 24).

Under-performing VFATs Some isolated VFATs, covering a third of one eta partition, can be seen as significantly less efficiency than the surrounding areas. This effect is predominantly due to broken DAC circuits, as presented in Section 5.3.3. It can be observed, for instance, in sector 36.

It also is possible to detect features non related to the GE1/1 detectors themselves:

Broken CSC DCFEB As one of the muon selection criteria requires the presence of a segment in the ME1/1 CSC chamber corresponding to the GE1/1 one, any malfunctioning of the CSC electronics will result in the absence of propagated hits in the offline reconstruction. Consequently, the efficiency cannot be measured in the overlapping GE1/1 region. This is the case on chamber GE+1/1/12, where it appears as a white bar in the efficiency map.

The diagnostic tool provided by the high-granularity plots help guide th repairs and enhancement plans for GE1/1 during the LHC LS3. While the bassin effect cannot be resolved without replacing the foils, several other interventions are planned: cooling will be added to the VTRx transceivers to prevent GBTx link instabilities; detector grounding will be improved to reduce the number of masked channel – particularly at the edges of the VFATs; attempts will be made to clear the shorted HV sectors – though this is particularly challenging for the second foil, which is inaccessible due to the GE1/1 construction; and under-performing or damaged VFATs will be replaced with spares.

5.7.5 Prospects

At the time of writing, several ongoing efforts aim to further refine the efficiency measurement results. This section highlights the most significant enhancements currently under development:

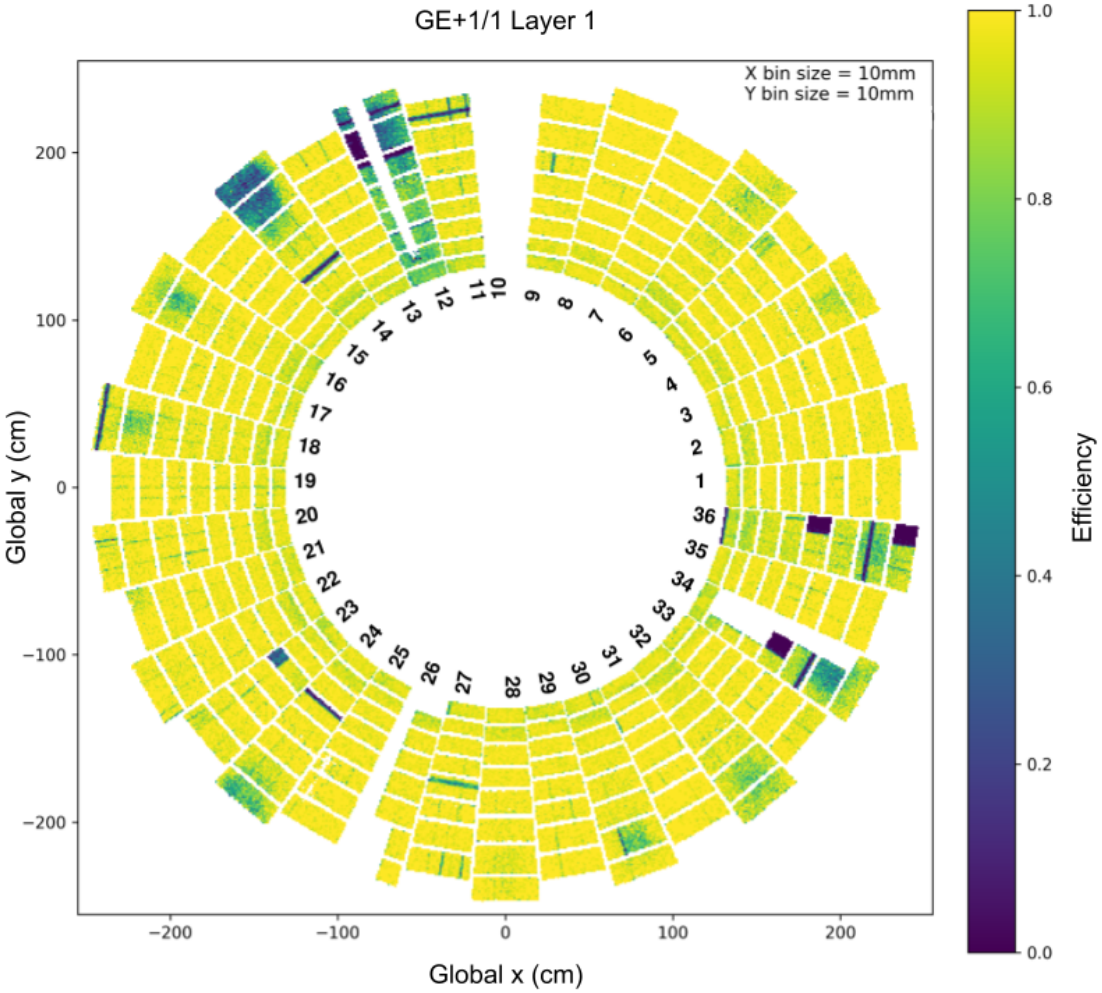


Figure 5.25: High-granularity efficiency measurement of the GE+1/1 Layer 1 disk taken at an equivalent divider current of 690 µA and the VFAT front-end configured in high gain mode (run 381075). Each bin covers an area of 1 cm × 1 cm which contains at least 10 propagated hits.

DAQ status masking for GE2/1 While the full offline analysis supports both the GE1/1 and GE2/1 stations, the DAQ status masking feature has been developed specially for the GE1/1 electronics layout. As a result, it is not directly compatible with the GE2/1 station. Indeed, one GE1/1 chamber contains a single OptoHybrid, whereas one GE2/1 chamber is composed of four independent OptoHybrids. This architectural difference prevents the direct usage of the same masking implementation.

Masked and dead channels So far, the masked and dead channels are not tracked in the conditions database available in CMSSW. Therefore, no specific handling

could be implemented in the reconstruction workflow. Whenever this information becomes available, it will become possible to enhance the `RecHits` building logic – more specifically, by allowing adjacent fired strips to be merged into a single `RecHit` even if they are separated by a missing (dead or masked) strip.

Handling of high-multiplicity events As presented in the previous section, high-multiplicity events contaminate the data recorded by the GE1/1 detector and no specific handling during the reconstruction currently exists. This subset of the data artificially increases the measured efficiency, as matched hits are likely to be found in the vicinity of nearly any propagated track hit. While the optimal approach for CMS-central reconstruction is under evaluation, a straightforward mitigation can be implemented in the offline analysis: any such event can be entirely discarded. Such cleanup of the dataset reflects more accurately the bare chamber efficiency.

5.8 Summary

After the completion of its installation in September 2020, GE1/1 underwent an intensive commissioning phase, followed by sustained operation. However, despite thorough quality control, numerous electronics instabilities were identified and investigated. Among others, the two main issues are: (1) the GBTx fuse corruption, which occurs whenever the system is power-cycled or the chip watchdog is triggered, and (2) the VTRx outgassing, which renders the GBTx links to the back-end unstable. Both issues are inherently dynamic and were thus addressed accordingly: the chamber low-voltage is now power-cycled whenever GBTx links do not come up during configuration (applied on average to 3 GE1/1 detectors), and the parts of the detector affected by the VTRx outgassing are masked (usually accounting for about 10% of the detector).

During commissioning, noise levels that would hinder the detection efficiency were measured. A six-month investigation resulted in the design and installation of ”LV filters”, which significantly improved the detector grounding and thereby reduced the average noise level to below 1 fC.

The first operations at a higher trigger rate (> 70 kHz) revealed the presence of high-multiplicity events. These were quickly and thoroughly analyzed, and are attributed to the VFAT3 ASIC design itself – no corrective action could be implemented at this stage of the project. As the most sensitive to such events, the Level-1 Trigger path received prompt mitigation: such events are masked based on their timing relative to the L1A signal, introducing some deadtime, but preventing further complications.

The first cosmic muons were seen in April 2021, following an initial latency scan that timed in the entire GE1/1 detector. A process remains ongoing in the pursuit of regular optimization. Likewise, the detection efficiency is closely monitored through a set of automation tools. Following the most recent optimization, the efficiency reached 94%, approaching the design target of 97%.

Similarly, the matching efficiency of the GE1/1 trigger primitives is being studied, both offline and online. The offline efficiency measurements assess the generation of

the trigger primitives themselves, whereas the online efficiency measurements evaluate their timing as well as the algorithms implemented in the OTMB and EMTF. Very preliminary results show a drop of efficiency of approximately 3% from the DAQ data-based measurements to the offline matching trigger primitives efficiency. An additional drop of about 8.5% is observed from the offline matching to the online one. Both those effects are currently being thoroughly investigated.

Throughout the commissioning and operations, particular attention is given to the damages caused to the electronics by high-voltage discharges. As shown during the GE1/1 slice test, the most serious risk is the destruction of input readout readout channels. Thanks to careful monitoring and optimized operations, less than 0.23% of the channels are currently flagged as dead.

Like all detectors, particularly young ones, this is however not the end of the story. The detector health and performance of the GE1/1 station – and of a few GE2/1 detectors – are continuously scrutinized and improved wherever possible. New issues are gradually discovered, but promptly troubleshooted and addressed. The lessons learned from GE1/1 are transferred to the upcoming ME0 and GE2/1 stations, to avoid repeating the same mistakes. This includes the mitigation of VTRx outgassing through a cooling system, the protection of the VFAT3 input channels via a dedicated protection circuit, and the reduction of the electronics noise through improved grounding implemented early in the design phase.

Conclusions

In 2020, the GE1/1 station was installed as part of the CMS upgrades for the High-Luminosity LHC. This marked a major milestone for the CMS GEM collaboration, as this was the first-ever large-scale deployment of Triple-GEM detectors within the experiment. Since then, the GE1/1 has been continuously and seamlessly taking data alongside the other CMS sub-detectors.

The GEM project upgrades, including GE1/1, aim to enhance the triggering and reconstruction performances of the muon spectrometer of CMS in response to the progressive increase of luminosity delivered by the LHC in recent years and anticipated in the future. The upcoming ME0 and GE2/1 stations will benefit from the experience acquired with the production, commissioning, and operation of GE1/1. The ME0 station is in active production as of 2025, with an installation scheduled in 2027 during the LHC LS3. It represents the most advanced iteration of the CMS GEM detectors. In contrast, the GE2/1 station is currently on hold due to production issues with the readout PCBs, despite the six chambers already installed and operated in CMS. Its construction and installation are postponed beyond LS3. Nonetheless, its R&D phase provided value insights that are being transposed to both ME0 and GE1/1.

This dissertation describes the author's contributions to the development and operation of the GEM DAQ acquisition system, focusing on its online software component for GE1/1; the quality control of the on-chamber readout electronics; and the initial commissioning and operations of the GE1/1 detector in CMS.

The first contributions were made to the quality control of the on-chamber electronics at ULB before its shipment to CERN. Each of the 144 OptoHybrids and GEB pairs was qualified after production and before installation on detectors. This step proved essential as many defects – mostly originating from mechanical stress – were detected early, saving a considerable amount of time during detector assembly at CERN. Due to the duration of the testing process, the testing workflow was gradually optimized and scaled up, reaching six concurrent testing stations, and allowing the completion of the project on time. Simultaneously, improvements to electronics were implemented to enhance its temperature monitoring. Finally, all the test results were archived into CMS databases for long-term tracking. Many of the lessons and outcomes from this testing campaign have been learned, and have since been applied to the ME0 and GE2/1 stations.

Building on the experience acquired on the GE1/1 DAQ system throughout its quality control, it became evident that a complete redesign of its online software was required.

The legacy software lacked the scalability, error handling, and maintainability for integration and operation of the full GE1/1 station in CMS. The work focused on the creation of a new, robust, and maintainable software stack ready for both current and future GEM stations.

At its core, the hardware abstractions implemented low in the DAQ stack provide support for a wide range of GEM flavors. Over the years, through extensive usage in CMS, quality control of the Phase-II detectors, and test beam campaigns, the online system has proven to be feature-complete and reliable. The calibration suite offers the necessary tools to optimize the detector's performance and assess its health. The monitoring suite provides real-time observations of the detector and its DAQ status, with long-term metric archiving and visualization for post-mortem analyses. The local readout application enables a fast ($\sim 10\,\text{Gbps}$) and independent way to record the detector data, bypassing the central CMS systems when needed.

The GE1/1 detector has since then reached its final home: the noses of the CMS experiment at the LHC Point 5 in Cessy, France. It marked the beginning of a lengthy – but fascinating – commissioning and operational phase. During the commissioning, issues arose and have been fixed or satisfactorily addressed. The noise levels have been reduced from 5 fC to $<1\,\text{fC}$ by the installation of LV filters; the communication issues have all been workaround in firmware and software aiming at including the largest fraction of the detector in the data-taking while maintaining the data acquisition stability; the discharge probability has been mitigated, and the channel damages are carefully monitored showing less than 0.23% of affect channels over 4 years of sustained operation.

During the early global data-taking with CMS, the first-ever muons were observed in the GE1/1 detector following the time alignment performed through latency scans. Some challenges also emerged: the VFAT3 exhibited high-multiplicity events – also poetically called flower events – when operated at a high rate ($> 70\,\text{kHz}$) due to a design flaw. Prompt mitigations in the trigger path were implemented via temporal masking.

Most recently, the detection efficiency of GE1/1 has been measured with high precision and resolution, studied in depth, and optimized to reach 94%, approaching the 97% design target. Further refinements are to, hopefully, narrow the gap. Meanwhile, the detector's health and performance remain scrutinized in terms of stability, channel loss, and efficiency.

Several technical, performance-related, and operational aspects of the GE1/1 station remain under active investigation. The trigger primitives are being characterized in terms of spatial resolution, efficiency, and timing. Integration within the Level-1 Trigger, both via the OTMB path and the EMTF path, is ongoing with, most of the work currently focusing on the former. Finally, a significant refurbishment campaign is being prepared for GE1/1 during the LHC LS3, aiming to address all shortcomings identified during these first years of operations.

List of Acronyms

AC Alternating Current

ADC Analog-to-Digital Converter

AFEB Anode Front-End Board (CSC)

ALCT Anode Local Charged Track (CSC)

ALICE A Large Ion Collider Experiment

AMC Advanced Mezzanine Card

APD Avalanche Photodiodes

API Application Programming Interface

ARM ARMing comparator

ASIC Application-Specific Integrated Circuit

ATCA Advanced Telecommunications Computing Architecture

ATLAS A Toroidal LHC ApparatuS

BCN Bunch Crossing Number

BDT Boosted Decision Tree

BER Bit Error Rate

BGA Ball Grid Array

BTMF Barrel Muon Track Finder

BU Builder Unit

BX Bunch CROSSing

CAN Controller Area Network

CBM Calibration, Bias, and Monitoring (VFAT)

CCB Clock and Control Board (CSC)

CCLUT Comparator Code LookUp Table (CSC)

cDAQ central DAQ

CERN European Organisation for Nuclear Research

CFD Constant Fraction Discriminator

CFEB Cathode Front-End Board (CSC)

CGI Common Gateway Interface

CMS Compact Muon Solenoid

CMSSW CMS SoftWare (framework)

COMPASS Common Muon and Proton Apparatus for Structure and Spectroscopy

COTS Commercial-Off-The-Shelf

CP Charge and Parity

CPU Central Processing Unit

CSC Cathode Strip Chamber

CTP7 Calorimeter Trigger Processor 7

DAC Digital-to-Analog Converter

DAQ Data Acquisition System

DB DataBase

DC Direct Current

DCFEB Digital CFEB (CSC)

DCS Detector Control System

DDR Double Data Rate

DDU Detector-Dependent Unit (CSC)

DIM Distributed Information Management

DMB Data MotherBoard (CSC)

DPDK Data Plane Development Kit (framework)

DPG Detector Performance Group

DQM Data Quality Monitoring

DSS Detector Safety System

DT Drift Tube

DTH DAQ & Timing Hub

EC Event Counter

ECAL Electromagnetic Calorimeter

EMTF Endcap Muon Track Finder

ENC Equivalent Noise Charge

EvB Event Builder

FASER ForwArd Search ExpeRiment

FED Front-End Driver

FEROL Front-End Readout Optical Link

FFF File-based Filter Farm

FM Function Manager

FPGA Field Programmable Gate Array

FSM Finite State Machine

FU Filter Unit

GBT GigaBit Transceiver

GBTx GigaBit Transceiver ASIC

GCS Gas Control System

GEB GEM Electronic Board

GEM Gas Electron Multiplier

HB Hadron Barel (calorimeter)

HCAL Hadronic Calorimeter

HE Hardon Endcap (calorimeter)

HEP High Energy Physics

HF Hadron Forward (calorimeter)

HGCAL High Granularity Calorimeter

HIP Highly-Ionizing Particles

HL-LHC High-Luminosity LHC

HLT High Level Trigger

HMT High-Multiplicity Trigger (CSC)

HO Hadron Outer (calorimeter)

HPD Hybrid PhotoDetectors

HTTP HyperText Transfer Protocol

HV High voltage

I2C Inter-Integrated Circuit

I2O Input/Output

INL Integrated Non-Linearity

IP Interaction Point or Internet Protocol

IPMI Intelligent Platform Management Interface

IQR InterQuartile Range

IT Information Technologies

IV Current-Voltage (curve)

JSON JavaScript Object Notation

L1A Level-1 Accept

L1T Level-1 Trigger

LCT Local Charge Track (CSC)

LEP Large Electron-Positron

LGAD Low Gain Avalanche Diodes

LHC Large Hadron Collider

LINAC LINear ACcelerator

LLP Long-Lived Particle

LMDB Lightning Memory-Mapped Database

LpGBT Low-power GBT ASIC

LSB Least Significant Bit

LUT LookUp Table

LV Low Voltage

LVDB Low Voltage Distribution Board

LVMB Low Voltage Distribution Motherboard

MAC Medium Access Control, address

MCH MicroTCA Carrier Hub

MGT Multi-Gigabit Transceivers

MicroTCA Micro Telecommunications Computing Architecture

MIP Minimum Ionizing Particles

MoEDAL Monopole and Exotics Detector At LHC

MPC Muon Port Card (CSC)

MPGD Micro-Pattern Gaseous Detector

MTD MIP Timing Detector

MWGR Mid-Week Global Run

NFS Network File System

NIC Network Interface Card

ODMB Optical Data MotherBoard (CSC)

OH OptoHybrid

OMDS Online Master Database System

OMS (CMS) Online Monitoring System

OMTF Overlap Muon Track Finder

ORCOFF Offline Reconstruction Condition DB - OFFline subset

ORCON Offline Reconstruction Condition DB - ONline subset

OS Operating System or Online System or Online Software

OTMB Optical Trigger MotherBoard (CSC)

PC Peripherical Crates (CSC) or Personal Computer

PCB Printed Circuit Board

PLC Programmable Logic Controllers

PMT PhotoMultiplier Tubes

PROM Read-Only Memories

PS Proton Synchrotron

PSB Proton Synchrotron Booster

QC Quality Control

QCD Quantum ChromoDynamics

RAM Random Access Memory

RCMS Run Control and Monitoring System

ROSA Receiver Optical Sub-Assembly (VTRx)

RPC Resistive Plate Chamber

RSSI Receiver Signal Strength Indicator (VTRx)

RU Readout Unit

RX Receiver

SCA Switch Capacitor Arrays (CSC) or Slow-Control ASIC (GEM)

SCADA Supervisory Control and Data Acquisition

SET Single Event Transient

SEU Single Event Upset

SiPM Silicon PhotoMultipliers

SNR Signal-to-Noise Ratio

SOAP Simple Object Access Protocol

SoC System-on-Chip

SPS Super Proton Synchrotron

SQL Structured Query Language

SRAM Static Random Access Memory

SST Silicon Strip Tracker

STS Storage and Transfer System

TCDS Timing and Control Distribution System

TCP Transmission Control Protocol

TDM Time-Division Multiplexing

TDR Technical Design Report

TEC Tracker EndCaps

TIB Tracker Inner Barrel

TID Tracker Inner Disk

TMB Trigger MotherBoard

TOB Tracker Outer Barrel

TOTEM TOTal Elastic and diffractive cross section Measurement

TSDB Time Series DataBase

TTC Timing, Trigger & Control

TTS Trigger Throttling System

TX Transmitter

USC Underground Service Cavern

UV UltraViolet

UXC Underground Experimental Cavern

VCC VME Crate Controller (CSC)

VdM Van der Meer (LHC luminosity scan)

VFAT Very Forward ATLAS and TOTEM

VME Versa Module Eurocard

VPT Vacuum PhotoTriodes

VTRx Versatile Transmitter-Receiver

VTRx+ Versatile Transmitter-Receiver Plus

xDAQ Cross Platform Data Acquisition

xDCFEB DCFEB embedding a GBTx ASIC (CSC)

XML eXtensible Markup Language

YAML YAML Ain't Markup Language

YETS Year-End Technical Stop

µGMT µGlobal Muon Trigger

µGT µGlobal Trigger

µTCA Micro Telecommunications Computing Architecture

List of Figures

1.1 Overview of the CERN accelerator complex as of January 2022. The LHC and its four main experiments are drawn in dark blue. The LHC proton accelerating chain in composed (in order) by the LINAC4 (purple), the BOOSTER (pink), the PS (magenta) and the SPS (lighter blue) [80]. 4

1.2 The planned LHC long-term calendar. It will lead to the commissioning of the HL-LHC, High Luminosity-LHC around 2028. [35] 6

1.3 Transverse view of one octant of the barrel of CMS. The lines represent the path of different particles through the different sub-detectors : the tracker in black, the ECAL in green, the HCAL in yellow and the muon system in orange. The solenoid is colored in grey. [74] 7

1.4 CMS system of coordinates. The beamline is along the Z axis. In green and brown, one can see the comparison between the polar angle θ and the pseudo-rapidity η. [76] 8

1.5 Schematic view of one quadrant in the RZ plane of the CMS tracker. The single-sided (double-sided) strip modules are colored in red (blue). The pixel detector is shown in green. [46] 9

1.6 Longitudinal view of the HCAL showing its segmentation and regions for Run 3. [101] 11

1.7 Location of the different detectors used in the muon system. The collision point is in the bottom left corner. The DTs (in light orange) are only present in the barrel. The CSCs (in green) are located only in the endcaps. The RPCs (in blue) are located both in the barrel and the endcaps. The Phase-II detectors are GE1/1 and GE2/1 (in red), ME0 (in orange), and the iRPC (in purple). [45] 12

1.8 Illustrations of a CMS DT chamber. 13

1.9 Layout of a CSC chamber made of 7 trapezoidal panels, forming 6 gas gaps. A few anode wires as well as cathode strips are revealed by the top panel cut-out. The former, the wires, are stretched in the azimuthal direction, providing measurement in the radial direction. The latter, the strips, run radially and provide a measurement in the azimuthal direction [47]. 14

1.10 Details of the electronics present on the ME1/1 chambers [66]. 15

1.11 Sectional view of a 2 gaps RPC detector. The readout strips (in brown) are sandwiched by the two gas volumes. The bakelite plates which protect the electrodes are colored in orange. [93] 17

1.12 Overview of the CMS trigger system. The yellow bubbles represent the two levels of trigger processing. They successively select the events to a rate that can be stored. . . . 18
1.13 Block diagram of the CMS Level-1 trigger used during Run-2 and Run-3 [96]. 20
1.14 Diagram of the Run-3 DAQ system. [39] . . . 22

2.1 The energy loss for a positively charged muon in copper. The Bethe-Bloch formula is represented in red. [27] . . . 27
2.2 Formation of an avalanche around an anode wire. [94] . . . 29
2.3 Electron microscope picture of a standard GEM foil. The holes have a typical inner diameter of 50 µm, outer diameter of 70 µm, with a pitch of 140 µm. [95] 31
2.4 Schematic of a GEM foil encloses between a cathode and an anode. The electric field lines are colored in red. The drift paths of the electrons and the ions are represented in blue and in red respectively. [110] . . . 32
2.5 Disposition of the three foils in a triple-GEM detector. The voltages shown correspond to an equivalent voltage divider current of 700 µA. Based on [95]. 33
2.6 (a) GEM foil transparency as a function of the gap field for different GEM foil voltages. (b) Effective gas gain and discharge probability as a function of the GEM foil voltage for different multi-GEM detector configurations. For both figures, ΔV_G is the voltage difference applied across the foil, and E_D is the electric field magnitude of the drift gap. Taken from [95]. . . . 34
2.7 Level-1 muon trigger rates as a function of the L1T muon candidate track p_T threshold before (blue) and after (purple) the GE1/1 upgrade at a luminosity of $2 \times 10^{34}\,\mathrm{cm}^{-1}\,\mathrm{s}^{-1}$, a for constant efficiency of 94%. [1] . . . 36
2.8 The CMS endcap first disk and nose. The GE1/1 super-chambers are highlighted in light red and cyan. [1] . . . 38
2.9 Exploded view of a GE1/1 chamber. [1] . . . 39
2.10 High voltage scheme of one single segmented GE1/1 foil. . . . 40
2.11 Overview of the GE1/1 readout electronic. [75] . . . 41
2.12 Picture of a GE1/1 VFAT3 hybrid. . . . 42
2.13 Block diagram of the VFAT3. The trigger path is highlighted is red and the tracking data path is in blue. [14] . . . 43
2.14 Cumulative channel losses for the GE1/1 slice test chambers before the implementation of any mitigation measure against discharges [89, 37]. . . . 44
2.15 Picture of a GE1/1 OptoHybrid. The Virtex-6 FPGA is framed in orange; the 3 GBTx ASICs in pink; and the SCA in light orange. The VTRx and VTTx are situated underneath the PCB at the locations designated by the teal and yellow rectangles, respectively. . . . 45
2.16 Picture of the two parts of the final GEB revision (v3) [75]. . . . 47
2.17 Picture of the GE1/1 back-end electronics in the CMS service cavern. The top (bottom) µTCA crate serves the negative (positive) GE1/1 endcap. Visible in each crate are 6 CTP7, 1 AMC13, and 1 MCH. . . . 48
2.18 Picture of the CTP7 board. . . . 50
2.19 Picture of the AMC13 board. . . . 51

2.20 Layout of a GE2/1 BACK detector and its division in 4 modules, M1 to M4 [45]. 53
2.21 High voltage scheme of one double segmented GE2/1 foil. 54
2.22 Drawing of design, installation, and placement of ME0 stacks in the CMS endcap. Note the alternating orientation of modules in the stacks [45]. 55
2.23 Simulation of the background particle rates per GEM foil sector in ME0 segmentation [12]. 56
2.24 Picture of the X2O leaf board. 58
2.25 Picture of the DTH-400 ATCA hub board. 58

3.1 Functional block diagram of the CMS GEM online software architecture. Hardware components are shown in yellow, with embedded software services highlighted in purple. xDAQ-based processes are represented in green, whereas RCMS-based processes are marked in dark blue. Control-related components are shown in light blue, while monitoring functionalities are colored in red. 66
3.2 Finite State Machine used by the GEM online software applications. For clarity, only the stable states are represented. 72
3.3 Web interface of the Level-0 Function Manager running in a MiniDAQ-3 instance, and using the GEM Level-1 function to interface with the GEM sub-system. 75
3.4 Flow of the local readout data. The switch component is optional and only used in the largest scale systems. The hardware components are colored in yellow whereas the software components are colored in green. 77
3.5 Screenshot of the calibration suite web interface for taking a S-curve. 78
3.6 Aggregation of the GBT phase scan results from 22 long GEBs during the electronics quality control. The number corresponds to the number of GEBs for which the phase was measured as valid for a given VFAT slot. Large windows of ”22” show that the bad phases mostly depend on the electronics design. 79
3.7 Typical DAC scan results. The generated current or voltage is shown a function of the DAC set point for each of the VFAT3 DAC circuits. 80
3.8 Typical example of an S-bit rate scan in the absence of detector signal for a single VFAT. The X-axis represents the *THR_ARM_DAC* value, scanned from its minimum (0) to its maximum (255); the Y-axis represents the measured S-bit rate, in Hertz. 81
3.9 Example of an S-curve scan for a single channel on one VFAT. The plot shows the number of hits as a function of the injected calibration pulse charge, at a fixed VFAT threshold. For each charge values, 100 pulses are injected. The decreasing trend is due to the fact that smaller (larger) DAC values correspond to higher (lower) charges. 82

3.10 Typical S-curve scan summary plots. It shows the raw result of an S-curve scan for all 3072 channels of a GE1/1 detector, grouped by VFAT, as a color map. The color code represents the hits ratio; the X-axis corresponds to the channel number and the Y-axis to the injected charge. The threshold can be inferred from to the position of the transition region between white (0% hit) and yellow (100% hit) colors, while the noise is related to the width of this region. 84
3.11 Boxplots representing the extracted threshold (a) and ENC (b) values for all 3072 channels of the GE1/1 detector, grouped by VFAT. The boxes span from the 25th to the 75th percentile; the whiskers extend to 1.5 times the interquartile range; and the individual circles represent the outliers. Within each box, the solid line indicates the median whereas the dashed line represents the mean. 85
3.12 Example of S-curves for a single VFAT before and after the trimming procedure. In Figures (a) and (b) show, all channels are configured with the extremum *TRIM_DAC* values, respectively -63 and +63. Figure (c) shows the results after trimming. The response across channels is significantly more uniform that before the trimming procedure. 87
3.13 Threshold scan during QC7 test (with cooling plate) of the GE1/1-X-S-CERN-0012 chamber. 88
3.14 Typical output plot of a latency scan for a single VFAT. This specific plot was taken during a test beam campaign in Spring 2022 with a pulse length of 1 BX. The X-axis represent the scanned latency values, whereas the Y-axis represents the number of time the specific VFAT recorded at least one hit on any of its channels. 89
3.15 Evolution of the duration (blue) and trigger rate (orange) of the S-cruve scan, as function of the number of VFATs included in the data-taking. The scan lasts a constant 24 s, plus 0.8 s for each additional VFAT. 91
3.16 Monitoring suite web interface for one `AMCMonitor` during a MiniDAQ-3 run. 92
3.17 Architecture of the metrics archiving. 94
3.18 Architecture of the CMS DCS to Grafana bridge. The Redis cache provides large query speed up for frequent queries. 95

4.1 GE1/1 quality control flow summary. The steps in blue are performed at CERN, in green at ULB, and in red at the various detector production sites. 102
4.2 Picture of the GEM hodoscope located in the laboratory at CERN and used for the final QC step, QC8. 105
4.3 Miscellaneous issues discovered during the GEB QC (a, b, c, d, e) and possible fixes (f). 108
4.4 The OptoHybrid standoff pads before (a) and after (b) applying tin. 109
4.5 FPGA external temperature measured via different probes. In blue, via the on-board PT100 readout by the SCA ADC (averaging 250 measurements); in red, via a thermocouple located on top of the PT100; in yellow, via the on-board PT100 measured via a voltmeter (but using the SCA current source). 110

4.6 Distribution of the current generated by the SCA ADC current generator for the GE1/1 OptoHybrid installed in CMS. 111
4.7 Failed VFAT3 DAC scans for which the nominal current values cannot be reached. This is the typical symptom of a chip powering issue. 114
4.8 Picture of the 6 test stands for the combined GEB & OH quality control. The µTCA-based DAQ system is visible on the right. 115
4.9 Screenshot of the web interface used to display the GEM construction database content. The detailed description of the OptoHybrid board *GE1/1-OH-V3-0090* is shown. 116

5.1 Picture of the CMS YE+1 disk and nose, including the newly installed GE+1/1 chambers located under aluminium cover plates, and ready for closure. 120
5.2 S-bit rate as function of the comparator threshold with the high-voltage off (blue) and on (orange), and the VFAT pre-amplifier configured in medium gain. Taken in March 2022 on GE+1/1/07 Ly2 VFAT #14. Each point represents 1 second of accumulated data, recorded sequentially from low to high threshold. The large rate fluctuations observed at high thresholds with the HV on (orange) are attributed to the stochastic nature of the signals, originating from either cosmic muon showers or from high-voltage events such as coronal discharges. 123
5.3 Modification (fix) of the RSSI readout circuit on one of the GE1/1 OptoHybrid boards. 127
5.4 Measurement of the RSSI of the three VTRx installed on the chamber GE-1/1/20 Ly1. In the original situation, the rapid drop of the indicators, as well as their fluctuations, are characteristic of VTRx optical transceivers suffering from glue out-gassing. After VTRx baking, the indicators remain very stable on extended periods of time. The almost null value for VTRx #1 is attributed to a failure in the readout circuit. The differences in behavior are explained by a chip-to-chip variation of the VTRx receiver optical sub-assembly block, in particular its glue, and the different transmitters and fibers. 127
5.5 Prototype of the GE1/1 VTRx cooling system. 128
5.6 Summary of the number of dead channels per VFAT of the GE-1/1 Layer 1 based on the calibration pulses scan. VFAT with communication errors at the calibration scan time have been excluded and do not appear in the plot. VFAT with all 128 input channels reported as dead are improperly configured or biased. 131
5.7 Evolution of the number of dead channels for each VFAT3 on GE2/1 demonstrator M3 module (GE+2/1/16 Ly2 C) during its first year of operation. . . 132
5.8 High-voltage discharge recorded in the 10th of August 2024 on chamber GE-1/1/19 Ly1. 134
5.9 Analysis results of the DAC *BIAS_PRE_I_BLCC* scan for VFAT #0 on chamber GE-1/1/19 Ly1 taken on the 12th of August 2024. The pattern suggests broken bits in the biasing DAC circuit. 134

5.10 Average ENC, extracted from the S-curve width, for all GE1/1 chamber before and after the implementation of the mitigation strategies. The wheels are oriented as if they were looked from the interaction point. Chambers that could be scanned show no data. 135

5.11 Equivalent noise charge comparison candle plots highlighting in (a) the presence of a group effect when the detector low-votlage is fully powered on; and in (b) the effect of the low-voltage cables. The front-end electronics becomes noisy when using the standard LV cables and powering the whole detector. . 136

5.12 Representations of the GE1/1 low-voltage filter. The filter is installed between the LV power supply cable and the chamber's LV connector to reduce the electronics noise to acceptable levels. The grounding of the LV- line is improved using a $2.5\,\mathrm{mm}^2$ cable (represented by R_1), directly connected to the chamber start point at the patch panel. This constitues a significant improvement over the original internal grounding using a $< 1\,\mathrm{mm}^2$ cable (represented by R_2) The capacitor installed in parallel to R_1 likely have no effect. A low-pass filter is implemented on the LV- line using two capacitors connected to ground via the chamber star point. 137

5.13 Equivalent noise charge before (red) and after (blue) the implementation of all mitigation measures for a quiet chamber (a) and a noisy chamber (b). . . 139

5.14 Readout latency scan results taken during the GE1/1 station commissioning with cosmics. The left plot is taken from the very first latency scan performed on GE1/1 (run 337973) and shows the first muons ever recorded in the station. The right plot is produced with data from run 341288 providing a narrower latency range, hence higher statistics, and post-processed data. The large width of the peak is due to the VFAT3 pulse stretch set to 7 BX. 140

5.15 Most recent latency scan for VFAT #20 in chamber GE+1/1/23 Ly2. The left histogram represents the per-channel occupancy distribution; the center plot represents the fraction of events where a given channel recorded a hit at a given latency; the right histograms show the fraction of events with at least one hit (blue) and the total number of hits normalized by the number of events (green). The narrow peaks are achieved by configuring the VFAT3 pulse stretch to 0 BX. Data taken during LHC fill 9639, totalling $1.19\,\mathrm{pb}^{-1}$ on the 17th of May 2024. 141

5.16 CMS event display showing a typical "flower event". A pair of muons, in red, originating from the interaction point cross the CMS barrel. The almost complete GE1/1 station is lit up (in dark blue) in a pattern reminding us of a flower. 142

5.17 Hit multiplicity as function of the readout latency for VFAT #11 with triggers sent periodically with an interval of 1000 BX and without high-voltage or calibration pulses. The peak around latency 990 correspond to the high-multiplicity events. 143

5.18 This plot shows the number of GE1/1 RecHits as a function of the delay, in LHC bunch crossings, between two consecutive Level-1 Accept. The RecHit multiplicity increase around a delay of 160 BX is attributed to the flower events. Produced with 138702 events from run 346247. 144
5.19 Chronogram explaining when high-multiplicity events can be observed in the tracking data path. The data saved by the second L1A will include the synchronous noise induced by the first L1A. 144
5.20 Oscilloscope capture showing the timing between the VFAT3 Level-1 Accept signal (blue) and the noise induced at the output of the charge-sense amplifier of a test channel (red). Credits to the GEM Electronics Group. 145
5.21 Cluster mask scan results for chamber GE-1/1/14 Ly2 taken with an L1A rate of 100 kHz. A delay of −1 BX means "no L1A and not masking applied". A delay of 0 BX mean "no masking applied". The delays colored in red will be masked, allowing up to 5% of high-multiplicity events in the trigger path. 146
5.22 Per-VFAT efficiency map for chamber GE+1/1/31 Ly2. [61] 151
5.23 High-voltage scan results for two representative GE1/1 detectors. A complete scan has been performed for a sub-set of the VFAT3 pre-amplifier gains and discriminator modes: low gain in arming (ARM) mode (green), medium gain in constant fraction discriminator (CFD) mode (orange), and high gain in CFD mode (purple). Detector (a) reaches its maximal efficiency around an equivalent divider current of 695 µA regardless of the VFAT3 operation mode; detector (b) reaches its maximal efficiency at an equivalent divider current of 700 µA only if the VFAT3 is configured in medium or high gain. 152
5.24 Efficiency of the GE1/1 chambers with HV I_{eq} at 680 µA and low gain for the VFAT3 chip settings in a run of 408 pb^{-1} (orange), and after the initial optimization study of the VFAT3 front-end chip configuration and HV settings in a run of 935 pb^{-1} (blue). The efficiency is calculated for chambers properly communicating and operating at nominal HV settings. [61] 153
5.25 High-granularity efficiency measurement of the GE+1/1 Layer 1 disk taken at an equivalent divider current of 690 µA and the VFAT front-end configured in high gain mode (run 381075). Each bin covers an area of 1 cm × 1 cm which contains at least 10 propagated hits. 155

List of Tables

2.1 Properties of gases typically used in the Triple-GEM detectors at normal temperature and pressure (NTP: 20 °C, 1 atm) for a MIP. [27] 27
2.2 Quantity, type, and purpose of the GE1/1 FEASTs DC-DC converters. . . . 46

3.1 Scalability of the calibration or health scans taken in the CMS GEM project. The numbers are always accounted per `AMCManager` as the operations between managers are fully independent and parallelized. The calibration pulses and cluster mask scans are described further in Sections 5.3 and 5.6. 91

4.1 Summary of the repairs attempted on the OptoHybrid boards. 109
4.2 Resistance of the traces on the OptoHybrid board PCB for the different sensors connected to the SCA ADC. 111
4.3 Summary of the OptoHybrid board status after quality control and chamber assembly. If multiple problems apply, only the most significant one is accounted for. 118
4.4 Summary of the GEB status after quality control and chamber assembly. If multiple problems apply, only the most significant one is accounted for. . . . 118

5.1 Muon selection criteria used for the `RPCMonitor`-based efficiency analysis using *standalone muons*. 149

Bibliography

[1] A. Colaleo et al. *CMS Technical Design Report for the Muon Endcap GEM Upgrade*. Technical Design Report. CERN-LHCC-2015-012; CMS-TDR-013. June 3, 2015.

[2] Aamir Irshad. "The CMS GEM Detector Front-end Electronics – Characterization and Implementation". PhD thesis. Université Libre de Bruxelles, 2022.

[3] D. Abbaneo et al. "Design of a constant fraction discriminator for the VFAT3 front-end ASIC of the CMS GEM detector". In: *Journal of Instrumentation* 11.1 (Jan. 2016), p. C01023. ISSN: 1748-0221. DOI: 10.1088/1748-0221/11/01/C01023.

[4] M. Abbas et al. "Quality Control of Mass-Produced GEM Detectors for the CMS GE1/1 Muon Upgrade". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 1034 (July 2022), p. 166716. ISSN: 01689002. DOI: 10.1016/j.nima.2022.166716. arXiv: 2203.12037[hep-ex,physics:physics].

[5] S. Abdullin et al. "Design, performance, and calibration of CMS forward calorimeter wedges". In: *The European Physical Journal C* 53.1 (Jan. 1, 2008), pp. 139–166. ISSN: 1434-6052. DOI: 10.1140/epjc/s10052-007-0459-4.

[6] Henso Abreu et al. "The FASER detector". In: *Journal of Instrumentation* 19.5 (May 2024). Publisher: IOP Publishing, P05066. ISSN: 1748-0221. DOI: 10.1088/1748-0221/19/05/P05066.

[7] M. Albrow et al. *CMS-TOTEM Precision Proton Spectrometer*. Technical Design Report. CERN-LHCC-2014-021; TOTEM-TDR-003; CMS-TDR-013. 2014.

[8] Alice Alfonsi, Luigi Benussi, and Emanuele Santovetti. "Time resolution of a Triple-GEM detector for future upgrade of the CMS muon system". Master thesis. Laboratori Nazionali di Frascati, 2017.

[9] The ALICE Collaboration et al. "The ALICE experiment at the CERN LHC". In: *Journal of Instrumentation* 3.8 (2008), S08002. ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08002.

[10] J. M. André et al. *File-Based Data Flow in the CMS Filter Farm*. Issue: 8 Pages: 082033 Publication Title: J. Phys.: Conf. Ser. Volume: 664. 2015. DOI: 10.1088/1742-6596/664/8/082033.

[11] Jean-Marc André et al. *Presentation layer of CMS Online Monitoring System*. 2019. DOI: 10.1051/epjconf/201921401044.

[12] Antonello Pellecchia. "Performance of micro-pattern gaseous detectors at the LHC and future collider experiments". PhD thesis. Bari University, 2023.

[13] G. Apollinari et al., eds. *High-Luminosity Large Hadron Collider (HL-LHC): Preliminary Design Report*. CERN Yellow Reports: Monographs. Geneva: CERN, Dec. 17, 2015. 285 pp. ISBN: 978-92-9083-422-9 978-92-9083-423-6.

[14] P. Aspell. *VFAT3 Basic Specification V1.3*. May 2015. URL: https://indico.cern.ch/event/399814/contributions/950580/attachments/800678/1097303/VFAT3_Basic_Specification_v1.3.pdf.

[15] P. Aspell et al. *VFAT3: A Trigger and Tracking Front-end ASIC for the Binary Readout of Gaseous and Silicon Sensors*. Pages: 8824655. 2019. DOI: 10.1109/NSSMIC.2018.8824655.

[16] The ATLAS Collaboration. "Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC". In: *Physics Letters B* 716.1 (Sept. 2012), pp. 1–29. ISSN: 03702693. DOI: 10.1016/j.physletb.2012.08.020. arXiv: 1207.7214.

[17] The ATLAS Collaboration et al. "The ATLAS Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.8 (2008), S08003. ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08003.

[18] G. Baiatian et al. *Energy Response and Longitudinal Shower Profiles Measured in CMS HCAL and Comparison With Geant4*. 2007.

[19] Vladislav Balagura. "Van der Meer Scan Luminosity Measurement and Beam-Beam Correction". In: *The European Physical Journal C* 81.1 (Jan. 2021), p. 26. ISSN: 1434-6044, 1434-6052. DOI: 10.1140/epjc/s10052-021-08837-y. arXiv: 2012.07752[hep-ex,physics:physics].

[20] Austin Ball et al. *A Letter of Intent to Install a milli-charged Particle Detector at LHC P5*. July 15, 2016. DOI: 10.48550/arXiv.1607.04669. arXiv: 1607.04669[hep-ex,physics:physics].

[21] G. L. Bayatian et al. *CMS Physics: Technical Design Report Volume 1: Detector Performance and Software*. Technical design report. CERN-LHCC-2006-001; CMS-TDR-8-1. 2006.

[22] G. L. Bayatyan et al. *CMS computing: Technical Design Report*. Technical design report. CERN-LHCC-2005-023; CMS-TDR-7. 2005.

[23] L. Berti et al. *Using XDAQ in Application Scenarios of the CMS Experiment*. May 26, 2003. DOI: 10.48550/arXiv.hep-ex/0305076. arXiv: hep-ex/0305076.

[24] Michele Bianco, Brian Dorney, and Jeremie A. Merlin. *GE1/1 Quality Control : instructions*. CMS GEM Internal Note. Sept. 28, 2016.

[25] Walter Blum, Werner Riegler, and Luigi Rolandi. *Particle detection with drift chambers*. 2. ed. Particle acceleration and detection. OCLC: 254908951. Berlin: Springer, 2008. 448 pp. ISBN: 978-3-540-76683-4 978-3-540-76684-1.

[26] V. Brigljevic et al. "Run Control and Monitor System for the CMS Experiment". In: Computing in High Energy and Nuclear Physics. La Jolla CA, June 18, 2003. DOI: 10.48550/arXiv.cs/0306110.

[27] C. Patrignani et al. "Review of Particle Physics". In: *Chin.Phys.* C40 (Oct. 3, 2016), p. 100001. DOI: 10.1088/1674-1137/40/10/100001.

[28] CAEN. *A1515B: 16/14 Channel 1-1.3kV (1 - 3 mA) Individual Floating Channel Dual Range Boards for Quadruple and Triple GEM detectors*. URL: https://www.caen.it/products/a1515b/.

[29] CAEN. *A1676A: EASY3000 Branch Controller*. URL: https://www.caen.it/products/a1676a/.

[30] CAEN. *A3016: 6 Channel 8 V / 16 A / 90 W Power Supply Board*. URL: https://www.caen.it/products/a3016/.

[31] CAEN. *A3486: 3-phase 220/400 Vac. 48 Vdc (2 ch x 2 kW/1 ch x 4 kW) Converter*. URL: https://www.caen.it/products/a3486/.

[32] CAEN. *EASY3000: Crate for EASY3000 Power Supply System for Hostile Area*. URL: https://www.caen.it/products/easy3000/.

[33] CAEN. *SY4527: Universal Multichannel Power Supply System*. URL: https://www.caen.it/products/sy4527/.

[34] A. Caratelli et al. "The GBT-SCA, a radiation tolerant ASIC for detector control and monitoring applications in HEP experiments". In: *Journal of Instrumentation* 10.3 (Mar. 2015), p. C03034. ISSN: 1748-0221. DOI: 10.1088/1748-0221/10/03/C03034.

[35] CERN. *LHC / HL-LHC Plan: January 2025*. Jan. 2025. URL: https://hilumilhc.web.cern.ch/sites/default/files/2025-01/HL-LHC_Plan_January2025.pdf.

[36] CERN : Education, Communications and Outreach Group. *LHC : the guide*. Feb. 2017.

[37] CMS Collaboration. *GE1/1 Slice Test Performance*. CMS Detector Performance Note. CMS-DP-2019-010. 2019.

[38] The CMS Collaboration. *A MIP Timing Detector for the CMS Phase-2 Upgrade*. Technical Design Report. CERN-LHCC-2019-003; CMS-TDR-020. 2019.

[39] The CMS Collaboration. "Development of the CMS detector for the CERN LHC Run 3. Development of the CMS detector for the CERN LHC Run 3". In: *JINST* 19.5 (2024), P05064. DOI: 10.1088/1748-0221/19/05/P05064.

[40] The CMS Collaboration. "Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC". In: *Physics Letters B* 716.1 (Sept. 2012), pp. 30–61. ISSN: 03702693. DOI: 10.1016/j.physletb.2012.08.021. arXiv: 1207.7235.

[41] The CMS Collaboration. *The Phase-2 Upgrade of the CMS Barrel Calorimeters*. Technical Design Report. ISBN: 9789290834588. 2017.

[42] The CMS Collaboration. *The Phase-2 Upgrade of the CMS DAQ Interim Technical Design Report*. Technical Design Report. ISBN: 9789290834670. 2017.

[43] The CMS Collaboration. *The Phase-2 Upgrade of the CMS Endcap Calorimeter*. Technical Design Report. 2017. DOI: 10.17181/CERN.IV8M.1JY2.

[44] The CMS Collaboration. *The Phase-2 Upgrade of the CMS L1 Trigger Interim Technical Design Report*. Technical Design Report. 2017. DOI: 10.17181/CERN.UUWZ.FMIS.

[45] The CMS Collaboration. *The Phase-2 Upgrade of the CMS Muon Detectors*. Technical Design Report. 2017. DOI: 10.17181/CERN.5T9S.VPMI.

[46] The CMS Collaboration. *The Phase-2 Upgrade of the CMS Tracker*. Technical Design Report. 2017. DOI: 10.17181/CERN.QZ28.FLHW.

[47] The CMS Collaboration et al. "The CMS experiment at the CERN LHC". In: *Journal of Instrumentation* 3.8 (2008), S08004. ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08004.

[48] Stefano Colafranceschi. "A study of materials used for muon chambers at the CMS Experiment at the LHC: interaction with gas, new materials and new technologies for detector upgrade". PhD thesis. Rome University, 2012.

[49] COMPASS Collaboration and P. Abbon. *The COMPASS Experiment at CERN*. Mar. 30, 2007. DOI: 10.48550/arXiv.hep-ex/0703049. arXiv: hep-ex/0703049.

[50] The CMS Collaboration. "Description and performance of track and primary-vertex reconstruction with the CMS tracker". In: *Journal of Instrumentation* 9.10 (Oct. 2014), P10009. ISSN: 1748-0221. DOI: 10.1088/1748-0221/9/10/P10009.

[51] G. Cummings and on behalf of the CMS Collaboration. "CMS HCAL VTRx-induced communication loss and mitigation". In: *Journal of Instrumentation* 17.5 (May 2022). Publisher: IOP Publishing, p. C05020. ISSN: 1748-0221. DOI: 10.1088/1748-0221/17/05/C05020.

[52] Isabelle Helena J. De Bruyn. "Electronics upgrade for the CMS CSC muon system at the High Luminosity LHC". In: 27th International Symposium on Nuclear Electronics & Computing. CMS-CR-2019-270. Becici, Budva, Montenegro: CERN, 2019.

[53] M. De Gruttola et al. *First experience in operating the population of the condition database for the CMS experiment. First experience in operating the population of the condition databases for the CMS experiment*. Place: Geneva _eprint: 1001.1676. 2010. DOI: 10.1088/1742-6596/219/4/042046.

[54] *Detector | CMS Experiment*. URL: https://cms.cern/detector (visited on 07/30/2024).

[55] ETM professional control GmbH. *WinCC OA*. URL: https://www.winccoa.com/index.html.

[56] F. Faccio et al. "FEAST2: A Radiation and Magnetic Field Tolerant Point-of-Load Buck DC/DC Converter". In: *2014 IEEE Radiation Effects Data Workshop (REDW)*. 2014 IEEE Radiation Effects Data Workshop (REDW). ISSN: 2154-0535. July 2014, pp. 1–7. DOI: 10.1109/REDW.2014.7004569.

[57] Francesco Fallavollita. "Triple-Gas Electron Multiplier technology for future upgrades of the CMS experiment: construction and certification of the CMS GE1/1 detector and longevity studies". PhD thesis. Pavia University, 2019. 338 pp.

[58] M. J. French et al. "Design and results from the APV25, a deep sub-micron CMOS front-end chip for the CMS tracker". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. 4th Int. Symp. on Development and Application of Semiconductor Tracking Detectors 466.2 (July 1, 2001), pp. 359–365. ISSN: 0168-9002. DOI: 10.1016/S0168-9002(01)00589-7.

[59] G. Abbiendi et al. *Muon Reconstruction in the CMS Detector*. CMS Analysis Note. CMS-AN-2008/097. Aug. 2009.

[60] C Gaspar, P Charpentier, and M Dönszelmann. "DIM, a portable, light weight package for information publishing, data transfer and inter-process communication". In: *Comput. Phys. Commun.* 140.1 (2001), pp. 102–9. DOI: 10.1016/S0010-4655(01)00260-0.

[61] *GEM performance results with 2024 data*. 2024.

[62] Giovanni Mocellin. "Performance of the GE1/1 detectors for the upgrade of the CMS muon forward system". PhD thesis. RWTH Aachen University, 2022.

[63] Grafana Labs. *Grafana*. URL: https://grafana.com/.

[64] W. Grant and Randolph Voorhies. *cereal - A C++11 library for serialization*. 2017. URL: http://uscilab.github.io/cereal/.

[65] Claus Grupen and Boris Shwartz. *Particle detectors*. Second Edition. Cambridge Monographs on Particle Physics, Nuclear Physics and Cosmology, 2008.

[66] Yacine Haddad and on behalf of the CMS Collaboration. "Upgrade of the CSC Muon System for the CMS Detector at the HL-LHC". In: *Proceedings of The European Physical Society Conference on High Energy Physics — PoS(EPS-HEP2021)*. Vol. 398. Conference Name: The European Physical Society Conference on High Energy Physics. SISSA Medialab, May 12, 2022, p. 837. DOI: 10.22323/1.398.0837.

[67] Eric Hazen et al. "The AMC13XG: a new generation clock/timing/DAQ module for CMS MicroTCA". In: *Journal of Instrumentation* 8 (Dec. 1, 2013), pp. C12036–C12036. DOI: 10.1088/1748-0221/8/12/C12036.

[68] Jeroen Hegeman et al. "The CMS Timing and Control Distribution System". In: *2015 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. 2015 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC). Oct. 2015, pp. 1–3. DOI: 10.1109/NSSMIC.2015.7581984.

[69] A Hervé. "The CMS detector magnet". In: *IEEE Trans. Appl. Supercond.* 10.1 (2000), pp. 389–94. DOI: 10.1109/77.828255.

[70] Muhammad Imran et al. "An Overview of the Database Framework for GEM Detector under CMS Experiment at CERN". In: *Proceedings of International Symposium on Grids & Clouds 2019 — PoS(ISGC2019)*. Vol. 351. Conference Name: International Symposium on Grids & Clouds 2019. SISSA Medialab, Nov. 21, 2019, p. 016. DOI: 10.22323/1.351.0016.

[71] InfluxData. *InfluxDB*. URL: https://www.influxdata.com/.

[72] Jeremie Alexandre Merlin. "Study of long-term sustained operation of gaseous detectors for the high rate environment in CMS". PhD thesis. Strasbourg University, 2016.

[73] Kohsuke Kawaguchi and Sun Microsystems, Inc. *Jenkins*. 2004. URL: https://www.jenkins.io/.

[74] Marzena Lapka. *Interactive Slice of the CMS detector*. July 2010. URL: https://cms-docdb.cern.ch/cgi-bin/PublicDocDB/ShowDocument?docid=4172 (visited on 05/10/2018).

[75] G. De Lentdecker. *Lessons learned from GE1/1*. CMS GEM/CSC Forward Muon Upgrade Workshop. Apr. 9, 2018. URL: https://indico.cern.ch/event/712513/contributions/2952971/attachments/1629383/2596461/20180409-Phase2Workshop-v1.pdf (visited on 05/15/2018).

[76] Thomas Lenzi. "Development and Study of Different Muon Track Reconstruction Algorithms for the Level-1 Trigger for the CMS Muon Upgrade with GEM Detectors". In: *arXiv:1306.0858 [hep-ex, physics:physics]* (June 4, 2013).

[77] The LHCb Collaboration et al. "The LHCb Detector at the LHC". In: *Journal of Instrumentation* 3.8 (2008), S08005. ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08005.

[78] The LHCf Collaboration et al. "The LHCf detector at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.8 (2008), S08006. ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08006.

[79] Linux Foundation. *Data Plane Development Kit (DPDK)*. 2015. URL: http://www.dpdk.org/.

[80] Ewa Lopienska. *The CERN accelerator complex, layout in 2022. Complexe des accélérateurs du CERN en janvier 2022*. 2022. URL: https://cds.cern.ch/record/2800984 (visited on 07/30/2024).

[81] M. Dabrowski et al. "The VFAT3-Comm-Port: a complete communication port for front-end ASICs intended for use within the high luminosity radiation environments of the LHC". In: *Journal of Instrumentation* 10.3 (Mar. 16, 2015), pp. C03019–C03019. ISSN: 1748-0221. DOI: 10.1088/1748-0221/10/03/C03019.

[82] Thierry Maerschalk. "Study of Triple-GEM detector for the upgrade of the CMS muon spectrometer at LHC". PhD thesis. Brussels University, 2016.

[83] Jeremie Merlin. *GEM stack optimization for GE21 and ME0 (cross-talk studies, discharge mitigation)*. Sept. 24, 2020. URL: https://indico.cern.ch/event/911950/contributions/3879507/ (visited on 08/19/2024).

[84] P. Moreira et al. "The GBT-SerDes ASIC prototype". In: *JINST* 5 (2010), p. C11022. DOI: 10.1088/1748-0221/5/11/C11022.

[85] Paulo Moreira et al. "lpGBT: Low-Power Radiation-Hard Multipurpose High-Speed Transceiver ASIC for High-Energy Physics Experiments". In: *IEEE Transactions on Nuclear Science* 72.1 (Jan. 2025), pp. 24–37. ISSN: 1558-1578. DOI: 10.1109/TNS.2024.3506753.

[86] Lauri Olantera. *The Curious Case of VTRx Receiver Failures*. Jan. 18, 2022. URL: https://indico.cern.ch/event/1099169/ (visited on 08/18/2024).

[87] Simone Pigazzini. *Automatic data processing for prompt calibration of the CMS ECAL*. 2023.

[88] J. L. Pinfold. "The MoEDAL Experiment at the LHC – a New Light on the Terascale Frontier". In: *Journal of Physics: Conference Series* 631.1 (2015), p. 012014. ISSN: 1742-6596. DOI: 10.1088/1742-6596/631/1/012014.

[89] N. Pozzobon. "The CMS Muon System: performance during the LHC Run-2". In: *Journal of Instrumentation* 14.11 (Nov. 2019), p. C11031. ISSN: 1748-0221. DOI: 10.1088/1748-0221/14/11/C11031.

[90] Corinne Pralavorio. *Record luminosity: well done LHC | CERN*. Dec. 14, 2017. URL: https://home.cern/about/updates/2017/11/record-luminosity-well-done-lhc (visited on 05/11/2018).

[91] Redis Inc. *Redis*. URL: https://redis.io/.

[92] Brendan Regnery. "The BEST Thesis: The Boosted Event Shape Tagger, A Search for Vector-like Quarks, and A Real GEM in CMS. La tesi BEST: il Boosted Event Shape Tagger, una ricerca di Vector-like Quarks e un vero GEM in CMS". PhD thesis. UC Davis, 2023.

[93] S. K. Park et al. "CMS endcap RPC gas gap production for upgrade". In: *Journal of Instrumentation* 7 (Nov. 15, 2012), P11013. DOI: 10.1088/1748-0221/7/11/P11013.

[94] Fabio Sauli. "Principles of operation of multiwire proportional and drift chambers". In: CERN, Geneva, 1975. DOI: 10.5170/CERN-1977-009.

[95] Fabio Sauli. "The gas electron multiplier (GEM): Operating principles and applications". In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*. Special Issue in memory of Glenn F. Knoll 805 (Jan. 1, 2016), pp. 2–24. ISSN: 0168-9002. DOI: 10.1016/j.nima.2015.07.060.

[96] A. M. Sirunyan et al. "Performance of the CMS Level-1 trigger in proton-proton collisions at $\sqrt{s} = 13$ TeV". In: *Journal of Instrumentation* 15.10 (Oct. 2020), P10017. ISSN: 1748-0221. DOI: 10.1088/1748-0221/15/10/P10017.

[97] Elizabeth Starling. "Electronics System of the CMS GE1/1 Muon Upgrade and Lessons Learned From the Slice Test During the 2017-2018 LHC Runs". In: Mar. 9, 2020, p. 118. DOI: 10.22323/1.370.0118.

[98] A Svetek et al. *The Calorimeter Trigger Processor Card: the next generation of high speed algorithmic data processing at CMS*. Issue: 02 Pages: C02011 Publication Title: JINST Volume: 11. 2016. DOI: 10.1088/1748-0221/11/02/C02011.

[99] Sysmas. *LMDB*. URL: https://www.symas.com/lmdb.

[100] The CMS Collaboration. "Calibration of the CMS drift tube chambers and measurement of the drift velocity with cosmic rays". In: *Journal of Instrumentation* 5.3 (Mar. 2010), T03016. ISSN: 1748-0221. DOI: 10.1088/1748-0221/5/03/T03016.

[101] The CMS Collaboration. "Calibration of the CMS hadron calorimeters using proton-proton collision data at $\sqrt{s} =$ 13 TeV". In: *Journal of Instrumentation* 15.5 (May 5, 2020), P05002–P05002. ISSN: 1748-0221. DOI: 10.1088/1748-0221/15/05/P05002. arXiv: 1910.00079[hep-ex,physics:physics].

[102] The CMS Collaboration. *LumiPublicResults < CMSPublic < TWiki*. URL: https://twiki.cern.ch/twiki/bin/view/CMSPublic/LumiPublicResults (visited on 07/30/2024).

[103] The CMS Collaboration. "Performance of the CMS muon detector and muon reconstruction with proton-proton collisions at $\sqrt{s}=$ 13 TeV". In: *Journal of Instrumentation* 13.6 (June 19, 2018), P06015–P06015. ISSN: 1748-0221. DOI: 10.1088/1748-0221/13/06/P06015. arXiv: 1804.04528[physics].

[104] The CMS Collaboration. *The CMS electromagnetic calorimeter project*. Technical design report. CERN-LHCC-97-033; CMS-TDR-4. ISBN: 9290831221; 92908311221. 1997.

[105] The TOTEM Collaboration et al. "The TOTEM Experiment at the CERN Large Hadron Collider". In: *Journal of Instrumentation* 3.8 (2008), S08007. ISSN: 1748-0221. DOI: 10.1088/1748-0221/3/08/S08007.

[106] J Troska et al. "Versatile transceiver and transmitter production status". In: *JINST* 8 (2013), p. C12030. DOI: 10.1088/1748-0221/8/12/C12030.

[107] Jan Troska et al. "The VTRx+, an Optical Link Module for Data Transmission at HL-LHC". In: Mar. 5, 2018, p. 048. DOI: 10.22323/1.313.0048.

[108] *VFAT3 for the CMS GEM Muon upgrades*. Indico. June 12, 2018. URL: https://indico.cern.ch/event/689062/ (visited on 10/20/2024).

[109] M. Vretenar et al. "The LINAC4 project: Overview and status". In: 2009. DOI: 10.5170/CERN-2009-004.36.

[110] *Welcome on triple GEM detector R&D : IMAGEM and GEMINI*. URL: https://web.infn.it/GEMINI/index.php/component/content/article?id=2:scientific-program (visited on 07/12/2018).

[111] Christian Wernet. "Unifying access to data from heterogeneous sources through a RESTful API using an efficient and dynamic SQL-query builder". Bachelor thesis. Hochschule, Eng. Econ., Karlsruhe, 2018.

[112] Cheuk-Yin Wong. *Introduction to High-Energy Heavy-Ion Collisions*. WORLD SCIENTIFIC, Sept. 1994. ISBN: 978-981-02-0263-7 978-981-4277-54-9. DOI: 10.1142/1128.

[113] YAML Language Development Team. *YAML Ain't Markup Language™*. URL: https://yaml.org/.