Alma Mater Studiorum · Università di Bologna

DOTTORATO DI RICERCA IN

FISICA

CICLO XXXI

Settore Concorsuale:
02/A1 - FISICA SPERIMENTALE DELLE INTERAZIONI FONDAMENTALI

Settore Scientifico Disciplinare:
FIS/01 - FISICA SPERIMENTALE

ATLAS Pixel Detector and readout upgrades for the improved LHC performance

Presentata da:
NICO GIANGIACOMI

Coordinatore Dottorato:
Chiar.ma Prof.ssa
Silvia Arcelli

Supervisore:
Chiar.mo Prof.
Alessandro Gabrielli

Abstract

Since the moment it was first started in 2008, the LHC particle accelerator at CERN continued to constantly increase its center-of-mass energy and luminosity. The entire LHC *lifetime* can be divided into several *phases*; in the first period the collider was running at an energy of 7-8 TeV and a luminosity of $\sim 10^{33}\,\mathrm{cm^{-2}s^{-1}}$. After that, the energy was increased to 13 TeV and the luminosity to $1\text{-}2{\cdot}10^{34}\,\mathrm{cm^{-2}s^{-1}}$ (*Phase-0*). During next years, LHC will undergo two more series of upgrades; after the first one it will reach the design energy of 14 TeV and a luminosity of $2\text{-}3{\cdot}10^{34}\,\mathrm{cm^{-2}s^{-1}}$ (*Phase-I*), and in the last phase (*Phase-II*) the luminosity will be increased to $\sim 7 \cdot 10^{34}\,\mathrm{cm^{-2}s^{-1}}$.

To keep up with the augmented detector performance, the LHC detectors where (and will be) upgraded as well. This work will focus on the ATLAS detector - one of the four main experiments of LHC - and in particular on its *Pixel Detector*. The ATLAS Pixel Detector was first upgraded in 2015, with the introduction of a new pixel layer - called IBL - to compensate for the B-layer inefficiencies and dead pixels and to increase the tracking performance for Phase-0 and Phase-I. IBL features smaller pixel size compared to the other layers, and higher granularity. The detector layout, combined with the higher LHC luminosity, led to an increased amount of data to be transmitted and analyzed, constituting a challenge for the read-out system. For this reason the previous readout chain was completed renovated and two new boards, called IBL-ROD and IBL-BOC, were designed to interface IBL. The two cards provide higher bandwidth and feature more recent technologies and high level control capabilities.

Between 2016 and 2018 the collider continued to increase its luminosity, exceeding its design value. As a result, the old readout chain still used for the rest of the Pixel Detector was completely saturated, and it was gradually replaced by the new system (IBL-ROD and IBL-BOC). While the hardware is already in place, the firmware and software utilities of the Pixel Detector readout chain are in constant evolution, in order to be able to provide good quality data even at the harsher environmental conditions of Phase-I LHC.

The second major upgrade involving the ATLAS Pixel Detector will be in 2024-2026, when the Inner Detector will be completely replaced by *ITk*, entirely made of silicon sensors. The new pixel detector will feature even smaller pixels built with 65 nm technology and higher granularity and data rates. To be able to sustain the more difficult conditions, another readout upgrade will be required; the final design has not been decided yet and is still under consideration. Two of the main candidates to implement the final system are the *πLUP project* in Bologna - which produced the πLUP readout board - and the *FELIX collaboration* - which involves several institutes all over the world and produced the readout card FLX-712, that will be used by some ATLAS sub-detectors during Phase-I upgrade.

This work will give an overview on the ATLAS Pixel Detector and will analyze the motivations that led to its upgrades. The current and future DAQ systems will also be discussed, focusing on the technologies adopted, the detector requirements and the results obtained.

Contents

Introduction

Since the late 50s, **particle accelerators** have been used by physicists all over the world to investigate the fundamental structure of the Universe. During the years, the energy and the performance obtained improved a lot. Currently, the largest and most powerful particle accelerator is the **Large Hadron Collider** (LHC) located at CERN in Geneva, which is capable of reaching a center of mass energy of 13 TeV. Independently on their advancement, particle accelerators would be useless without **detectors** capable of measuring the properties of all the particles generated during collisions. Therefore, the evolution of accelerator technologies requires also an evolution of the detectors, that must adapt to the new environment and exploit new strategies to provide the desired results. There are four main experiments at LHC: A Toroidal LHC ApparuS (ATLAS), Compact Muon Solenoid (CMS), A Large Ion Collider Experiment (ALICE) and LHC-beauty (LHCb). Each one exploits different technologies and is specialized to perform a specific task. In the next few years LHC and all the experiments will undergo a series of major upgrades in order to extend searches on physics. The first upgrade will increase the LHC beam energy - that will reach the design value of 14 TeV - and luminosity; it will start in early 2019 and will end in 2021, with the beginning of **Phase-I LHC**. The second upgrade will start in 2024 and will end in 2026 (**Phase-II LHC**) and is meant to ulteriorly increase the collider luminosity.

This thesis is divided in two main parts and will focus on my contributions to the **ATLAS experiment** during the last three years as a PhD student. The first part of this work will discuss the current status and results of the detector (Phase-0 LHC) while the second part will discuss future upgrades (Phase-I/II LHC). ATLAS is a general purpose detector, composed of several sub-detectors which are optimized to detect all kinds of *interesting* particles and cover a large portion of the solid angle. A more detailed overview of the whole ATLAS experiment and its goals will be given in Chapter 1; however, due to the monumental complexity of the subject, many aspects will be left out or treated only superficially.

My work concerned directly the innermost ATLAS sub-detectors: the Pixel Detector, which will be the main subject of Chapter 2. The Pixel Detector was originally composed of three coaxial layers (called B-Layer, Layer-1 and Layer-2) and three end-cap disks. After few years of operation, due to the LHC luminosity increment, the aging of the pixels and the need for more precise measurements, the necessity to

upgrade the detector arose. So, starting from the beginning of 2015, a fourth layer - called **Insertable B-Layer** (IBL) - became part of the Pixel Detector. More details on the reasons that led to the introduction of a fourth pixel layer will be given in Chapter 2. Being more recent and technologically advanced, IBL provides an unprecedented granularity and data-rate, overpassing the limitations of the readout system used by the rest of the Pixel Detector. The readout chain was hence unavoidably upgraded and it was replaced with a new one based on two new boards, called IBL-Back of Crate (**IBL-BOC**) and IBL-ReadOut Driver (**IBL-ROD**). In the following year, the same system was then used to upgrade the entirety of the Pixel Detector readout, to overcome the bandwidth saturation problems caused by the increase of LHC luminosity, pile-up and trigger rate. The readout upgrade time-line is shown in Fig. 1.

During my PhD I worked together with the ATLAS Pixel Detector Data Acquisition group and I collaborated to upgrade the system. In particular, I became the main firmware expert of the ROD board and I designed and developed software, tests and procedures meant to improve the on-line data quality. More details on the ATLAS readout system and in particular on the Pixel Detector data acquisition chain can be found in Chapter 3.

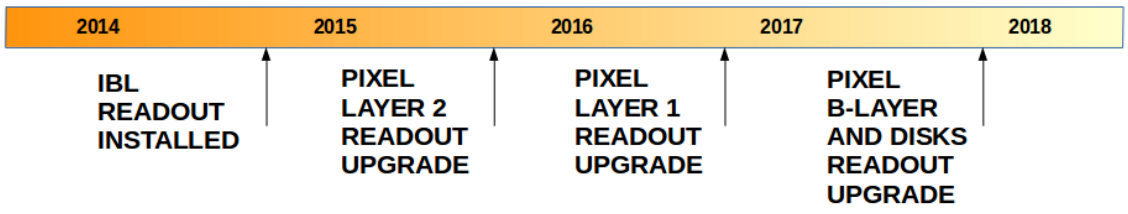


Figure 1: Time-line of the ATLAS Pixel Detector Readout upgrade.

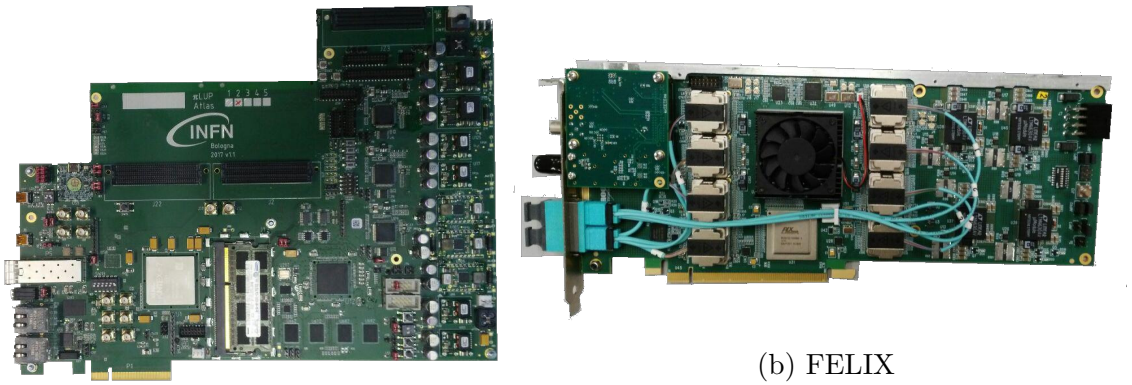
The second part of this thesis will discuss the future LHC plans, focusing again on the pixel detector and its readout chain. To withstand the harsher Phase-II conditions, the ATLAS Inner Detector (which comprehends the Pixel Detector) will be completely substituted with the *Inner TracKer* (ITk), entirely based on silicon technologies. The goals of ITk, its design and the strategies and technologies adopted will be discussed in Chapter 4.

The new Phase-II detector bandwidth, together with the increased pile-up and trigger rate, led to the necessity to redesign completely the readout chain that would otherwise be saturated and incapable of providing good quality data. The final Phase-II data acquisition system has not been finalized yet and several strategies are taken in consideration. During my PhD I worked in collaboration with two groups that realized two electronic boards candidates for the final Phase-II ATLAS readout: the **πLUP Project** and the **FELIX Project**. The πLUP board - shown in Fig. 2 (a) - was developed by University and INFN of Bologna as a natural upgrade of the IBL-BOC and IBL-ROD system. This card exploits recent technology and

is able to provide a bandwidth that would be sufficient to cope with the Phase-II requirements. I joined this project since the very beginning as the main firmware developer of this board, as well as software and integration designer.

The FELIX project involves several institutes, among which Nationaal Instituut voor Kernfysica en Hoge-EnergieFysica (NIKHEF), Brookhaven National Laboratory (BNL) and CERN. The purpose of the collaboration is to provide an electronic board that will revolutionize the ATLAS readout system chain, replacing the role of the ROD. While the Phase-II card has not been designed yet, a first version of the card, called **FLX-712** (Fig. 2 (b)), has been realized and will take part in the data acquisition system for some ATLAS sub-detectors during Phase-I. Between 2017 and 2018 a collaboration between the πLUP and FELIX projects started; during this time I spent six months at BNL where I contributed to the design of the FELIX firmware and I developed a joint setup involving the two boards.

Chapter 5 will present in detail the Phase-II ATLAS readout strategies, the two projects, their collaboration and the results achieved.



(a) πLUP

(b) FELIX

Figure 2: Picture of a πLUP (v1.1) and a FELIX (FLX-712) card.

Part I

LHC Phase-0

Chapter 1

The ATLAS Experiment at CERN

The fundamental goal of all particle physicists is to understand the basic structure of the Universe and uncover its mysteries. After many years of experiments and discoveries, the physics community concluded that the Universe is made of particles that are arranged together according to sets of rules summarized by the *Standard Model*. Particle accelerators provide a way to artificially recreate particles in a laboratory system, so that it is possible to measure their properties and their structure. Currently, the most powerful accelerator in the World is the Large Hadron Collider (LHC) at Conseil Europeen pour la Recherche Nucleaire (CERN). It is placed in an underground tunnel whose circumference is $\sim 27\,\text{km}$ long. Seven experiments are placed along the accelerator ring; this Chapter will present one of those experiments, called A Toroidal Lhc ApparatuS (ATLAS), which is the main subject of this thesis. ATLAS is a very complex system, with more than 3000 people working on it. It is therefore extremely difficult and beyond the purpose of this thesis to provide a complete and detailed treatment of all the aspects of the experiment; this chapter will only give a short overview on the general structure of the detector and its physics program, starting with a brief introduction on the accelerating system.

1.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [1] at CERN is the biggest and most powerful accelerator in the World. It is placed $\sim 100\,\text{m}$ underground under the border between France and Switzerland, in the pre-existing tunnel that between 1989 and 2000 was occupied by the *Large Electron-Proton* collider (LEP). LHC is a proton-proton (p-p) or ion-ion collider, and was built to reach a design center-of-mass energy $\sqrt{s} = 14\,\text{TeV}$ for p-p collisions, although at the current status the maximum energy reached is $13\,\text{TeV}$. The next sections will present a short overview on the LHC structure and some of its main parameters.

1.1.1 LHC structure

The LHC complex accelerates two beams of protons or ions to the designed energy; the protons (ions) are then collided into four points around which the four main experiments (ATLAS, CMS, LHCb and ALICE) are built.

The acceleration process is very complex and requires several consecutive steps; in each stage a smaller accelerator provides more and more energy to the beam before injecting it into the next stage. The whole injection chain is shown in Fig. 1.1. After production, protons enters the linear accelerator *Linac-2*, which increases their energy to 50 MeV. In 2020 Linac-2 will be replaced by a new linear accelerator, called *Linac-4*. The protons are successively injected at the rate of 1 Hz into the *Proton Synchrotron Booster* (PSB) where they reach an energy of 1.4 GeV. In the next stage the *Proton Synchrotron* (PS) further increase the beam energy to 25 GeV and, due to the intrinsic nature of synchrotron accelerators, it divides it in bunches. As the last step before entering LHC, the bunches of protons are then fed to the *Super Proton Synchrotron* (SPS), where they reach an energy of 450 GeV.

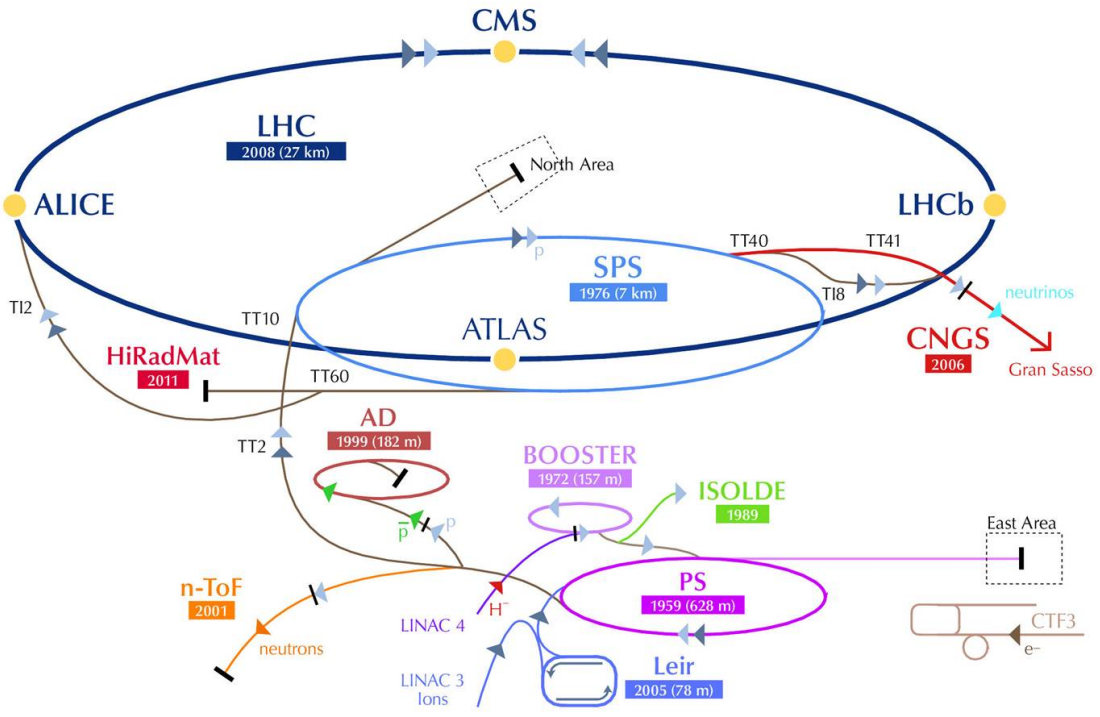


Figure 1.1: The injection chain of the LHC complex, each ring representing a different accelerator. The chain starts with the linear accelerator *Linac-2*, followed by the *Proton Synchrotron Booster* (PSB). Two other synchrotrons, *Proton Synchrotron* (PS) and *Super Proton Synchrotron* (SPS) further increase the beam energy before injecting it directly into LHC.

The acceleration process of ions is slightly different, and begins at the linear

accelerator *Linac-3*, that boosts lead ions that are fully stripped of their electrons ($^{208}Pb^{82+}$). The lead ions are then injected into the *Low Energy Ion Ring* (LEIR), where they are compressed. From the LEIR the ions are sent to the PS, and then they follow the same steps as the protons. LHC accelerates each nucleon up to 2.76 GeV, yielding a total center-of-mass energy of 1.15 PeV.

Since they have the same charge, in LHC the two particle beams travel in two separate counter-rotating rings, differently from particle-antiparticle colliders (such as LEP and Tevatron) where only one ring is needed. The beam curvature is obtained by using a set of dipole magnets, providing a strong magnetic field. The design of LHC was constrained by the fact that it was built into the pre-existing mono-ring LEP tunnel, which could not fit two separate rings of magnets. This problem was overcome by using twin bore magnets, consisting of two sets of coils and beam channels contained within the same mechanical structure and cryostat as shown in Fig. 1.2. The dipole magnetic field required for 14 TeV operations is 8.33 T, which can only be reached through superconducting technology. The 1232 LHC magnets are realized using niobium-titanium wires that are capable of sustaining a operational temperature of 2 K.

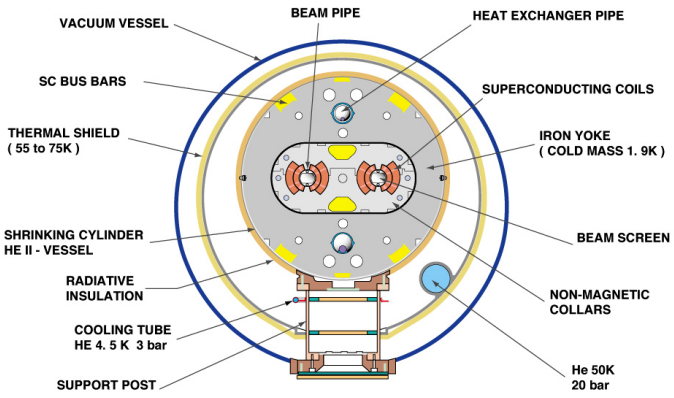


Figure 1.2: Cross section of an LHC dipole. [2]

The decision of using hadrons instead of electron-positron pairs (as was LEP) or proton-antiproton pairs (as was Tevatron) was done to achieve higher center-of-mass energy and higher amount of bunches available in the accelerator. In fact, electron-positron accelerators are strongly limited by energy losses from the synchrotron radiation, an electromagnetic radiation generated by radially accelerated charged particles. This loss can be expressed as

$$\frac{dE}{dt} \propto \frac{E^4}{m^4 R^2}$$

where m is the accelerated particle mass and R is the radius of the accelerator.

This effect is extremely reduced for protons, since their mass is ~ 2000 times higher than the mass of the electrons. In proton-antiproton accelerators, the problem is of a different nature. In fact, it is extremely difficult to produce anti-protons and accumulate them, so the number of bunches and particles inside the collider is very limited.

1.1.2 LHC operational parameters

The qualities of particle accelerators can be described using a series of parameters, such as the maximum energy reached, the length of the accelerator, the number of bunches in the ring, the number of particles in each bunch, the separation between two bunches, and so on. It is possible to express the collider performance using a single quantity, called *Instantaneous Luminosity* $\mathcal{L}$, which is a combination of many of those parameters. The instantaneous luminosity is a function of the number of particles in the accelerator (n_i), the revolution frequency (f) of the bunches and the transversal dimensions og the beam (σ_x and σ_y) and is expressed as:

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

By taking in consideration in a more detailed way other accelerator aspects, the LHC instantaneous luminosity expression can be rewritten as:

$$\mathcal{L} = \frac{N_b^2 \cdot n_b \cdot f \cdot \gamma_r}{4\pi \cdot \epsilon_m \cdot \beta^*} F \tag{1.1}$$

where:

- n_b is the number of bunches inside the ring;
- N_b is the number of particle contained in each bunch;
- γ_r is the relativistic gamma factor of the particles;
- ϵ_m is the normalized transverse beam emittance;
- β^* is the beta function at the collision point;
- F is the geometric luminosity reduction factor, due to the incidence angle of the beams at the collision point.

The importance of the instantaneous luminosity is given by the fact that it correlates the rate of production of a certain event N_e to its cross section σ_e, according to the formula:

$$\frac{dN_e}{dt} = \mathcal{L} \cdot \sigma_e \tag{1.2}$$

The total number of events produced can be obtained by integrating the previous equation in a given time interval:

$$N_e = \int \frac{dN_e}{dt}\, dt = \int \mathcal{L} \cdot \sigma_e \, dt = L \cdot \sigma_e \tag{1.3}$$

where L is the *integrated luminosity*. Equations 1.2 and 1.3 lead to two main con-

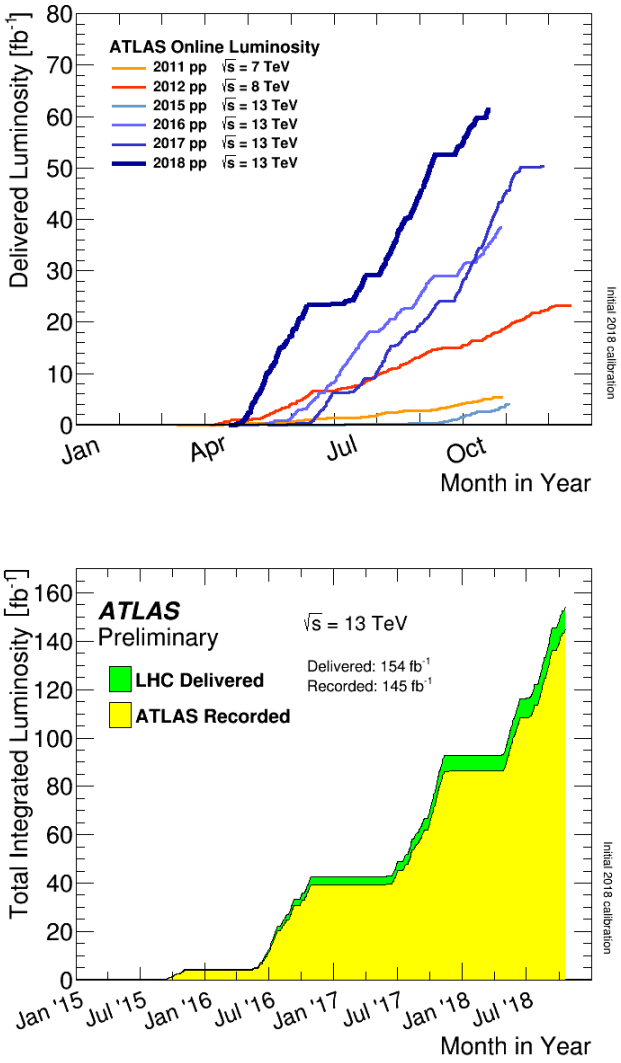


Figure 1.3: (a) Cumulative luminosity versus day delivered to ATLAS during stable beams and for high energy p-p collisions. (b) Cumulative luminosity versus time delivered to ATLAS (green) and recorded by ATLAS (yellow) during stable beams for pp collisions at 13 TeV centre-of-mass energy in LHC Run 2. [ATLAS Luminosity public plots]

clusions; firstly, the cross section of a certain event can be calculated by measuring the number of events produced and the accelerator luminosity. Secondly, in order to produce events with a very low cross sections, it is necessary to have a very high

luminosity. Fig. 1.3 shows as an example the cumulative luminosity delivered by LHC in the ATLAS experiment.

1.1.3 LHC roadmap

LHC was designed to provide a center-of-mass energy $\sqrt{s} = 14\,\text{TeV}$ and an instantaneous luminosity $\mathcal{L} = 1 \cdot 10^{34} cm^{-2} s^{-1}$. However, before reaching the design operational values, LHC passed trough several phases and was (and will be) subjected to several upgrades, meant to gradually improve its performance.

At the beginning of its lifetime, LHC ramped its center-of-mass energy from 700 GeV

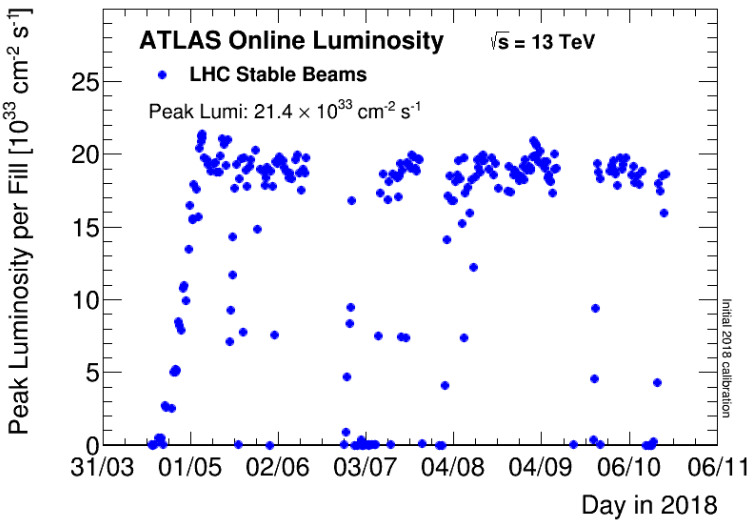


Figure 1.4: The peak instantaneous luminosity delivered to ATLAS during stable beams for pp collisions at 13 TeV centre-of-mass energy is shown for each LHC fill as a function of time in 2018. The luminosity is determined using counting rates measured by the luminosity detectors, and is based on an initial estimate from Van dr Meer beam-separation scans during 2017. [ATLAS Luminosity public plots]

to 7-8 TeV and reached a peak luminosity of $7.7 \cdot 10^{33}\,\text{cm}^{-2}\text{s}^{-1}$, even if at double the design bunch-crossing separation (50 ns instead of 25 ns). This phase, which lasted from 2008 and 2013, is called *Run-I*, during which the most notable achievement was the discovery of the Higgs Boson.

From February 2013 to April 2015, the LHC complex was shut down to consolidate the magnet interconnection (to reach higher energy); this phase is called *Long Shutdown 1* (LS1). LS1 marked the beginning of the so-called *Phase-0* where LHC reached an energy close the design one (13 TeV). The data-taking period of Phase-0 is called *Run-II*, during which the instantaneous luminosity surpassed the design value of $1 \cdot 10^{33}\,\text{cm}^{-2}\text{s}^{-1}$, up to a peak value $\mathcal{L} = 2.14 \cdot 10^{34}\,\text{cm}^{-2}\text{s}^{-1}$ (Fig. 1.4). Run-II

will end in December 2018 with a second upgrade phase - called Long Shutdown-2 (LS2) - starting the LHC *Phase-I*.

LS2 is meant to ulteriorly increase the LHC energy - that will reach the design value of 14 TeV - and luminosity. Among the other improvements, a new linear injector, called *Linac-4*, will substitute its predecessor (Linac-2). After LS2, in 2021, a new data-taking period - called *Run-III* - will begin. LHC is expected to reach a peak luminosity $\mathcal{L} = 3\text{-}4 \cdot 10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ and to deliver a total integrated luminosity of $300\,\mathrm{fb}^{-1}$ during the entirety of Run-III, that will last till 2024.

In 2024 LHC will be shut down again (*Long Shutdown-3* LS3) and will undergo a new series of major upgrades, entering a new phase called *Phase-II* or *High Luminosity LHC* (HL-LHC). During LS3, between 2024 and 2026, new superconducting magnets will be installed and the machine peak luminosity will be increased of several times ($\sim 5/7$ times) respect to the current one. HL-LHC is expected to deliver a total integrated luminosity of $3000\,\mathrm{fb}^{-1}$.

The various LHC phases, runs and shut-down periods are shown in Fig. 1.5.

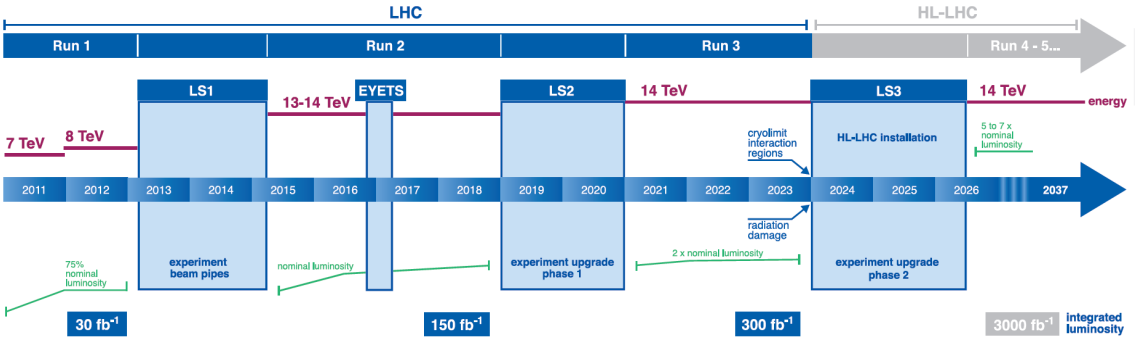


Figure 1.5: The LHC upgrade roadmap. The Run periods (during which data is acquired) are separated by Long Shutdown periods (LS), for hardware maintenance and upgrade. Each period comprising a LS and a Run is called a Phase. [3]

1.2 The ATLAS Experiment

A Toroidal LHC ApparatuS (ATLAS) is one of the four main LHC experiments, together with *Compact Muon Solenoid* (CMS), *A Large Ion Collider Experiment* (ALICE) and *LHC-beauty* (LHC-b). It is operated by an international collaboration comprising more than 3000 scientists from 181 institutions around the world. ATLAS is a general purpose experiment meant to fully exploit all the physics searches enabled by LHC energy and luminosity [4]. This means that it is not focused on studying a particular interaction or process, but instead is optimized to detect all the *interesting* events that are produced with high energy and momentum. By analyzing the properties of the particles produced by each event, it is possible to reconstruct the entire production process; this way ATLAS is able to measure with

high precision Standard Model interactions (QCD, electroweak, flavor physics) and to search for new physics within and outside the Standard Model. The most notable result of the ATLAS experiment so far is the discovery of the Higgs boson announced in 2012, together with the CMS experiment [5] [6].

The detector is $\sim 44\,\text{m}$ long, $\sim 25\,\text{m}$ high and $\sim 25\,\text{m}$ wide, and has a cylindrical geometry, with a central barrel and two end-cap regions at the sides. It is built around one of the LHC collision points - called *Point 1* - and is composed of several concentric cylindrical sub-detectors, as shown in Fig. 1.6.

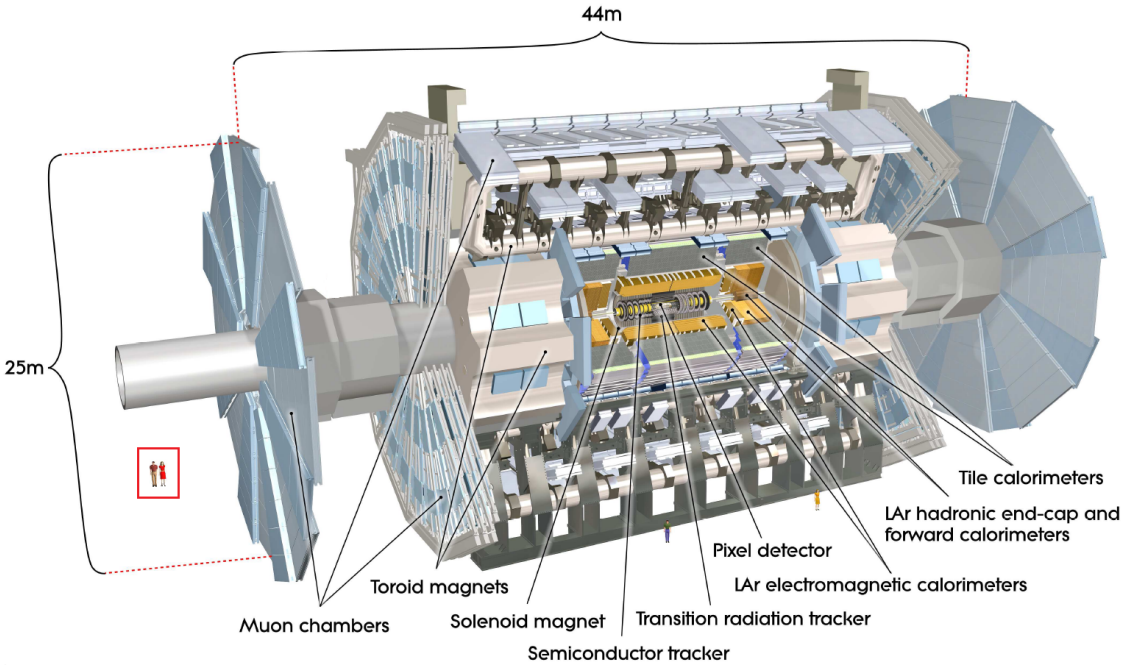


Figure 1.6: View of the ATLAS Detector. The system is built with a cylindrical symmetry, with one central barrel and two end-caps at the sides. It is composed of several cylindrical sub-detectors, expanding concentrically from the interaction point outwards. Each sub-detector can be identified in the picture. It is possible to compare the size of the ATLAS experiment with the humans in the red box.

All the events in ATLAS are described using right-handed spherical coordinates, which origin is set in the *Interaction Point* (IP). The beam direction identifies the z-axis, and the *transverse x-y plane* is defined by the x-axis pointing towards the center of the LHC ring and the y-axis pointing upwards. Since the beam travels along the z-axis, it transverse momentum is null, meaning that also for the collision products there are some observables that are conserved in the transverse plane, such as the transverse momentum p_T or energy E_T.

The same coordinate system can be expressed using polar coordinates: r is the distance from the IP, the azimuthal angle ϕ is the angular distance respect to the x-axis and the polar angle θ is the angular distance respect to the beam axis (z). It is

useful to find quantity that are Lorentz-invariant to describe the particles traveling into ATLAS; one of those quantity is the *rapidity* y, defined as

$$y = \frac{1}{2} ln \left(\frac{E + p_L}{E - p_L} \right)$$

where p_L is the linear momentum of the particle. The angular separation between two particles can be expressed in terms of their rapidity, using the Lorentz-invariant relationship

$$\Delta R = \sqrt{(\Delta y)^2 + (\Delta \phi)^2} \tag{1.4}$$

Another quantity widely used in LHC experiments is the *pseudorapidity* η, defined as

$$\eta = -ln \left[tan \left(\frac{\theta}{2} \right) \right]$$

Pseudorapidity is related to the rapidity according to the equation

$$y = ln \left(\frac{\sqrt{m^2 + p_T^2 cosh^2 \eta} + p_T sinh \eta}{\sqrt{m^2 + p_T^2}} \right)$$

For relativistic boosted particles, where $m << p_T$, pseudorapidity becomes equal to rapidity ($\eta \approx y$), and the angular distance between two particles (Equation 1.5) can be expressed using only angular quantities:

$$\Delta R = \sqrt{(\Delta y)^2 + (\Delta \phi)^2} \approx \sqrt{(\Delta \eta)^2 + (\Delta \phi)^2} \tag{1.5}$$

For this reason the ATLAS coordinate system is often expressed in terms of he pseudorapidity η, instead of the polar angle θ.

1.2.1 ATLAS Detector overview

The purpose of the ATLAS detector is to identify particles traveling through it and measure their momentum and energy, with exception for neutrinos. Its structure is optimized to operate in the LHC environmental conditions, dictated by the fact that proton-proton collisions involve mostly QCD processes, producing a very large amount of high energy jets. Moreover, due to the high luminosity of the beam, multiple proton collisions happen simultaneously; this effect is called *pile-up*. In order to enable the investigation of the properties of the collisions, to reach the physics goals of the ATLAS community, the detector must satisfy several requirements:

- fine granularity: the granularity of the detector must be high enough to face the high pile-up and to distinguish between different collisions;
- geometric acceptance: the detector must cover the largest possible fraction of solid angle;

- fast timing: the time for signal shaping and propagation must be inferior to the bunch crossing separation of 25 ns;
- good energy resolution: the measurement of particle energy must be very precise for a wide range of energies;
- tracking: the detector must be able to identify the track of charged particles with a very precise spatial resolution, in order to identify the primary vertex and eventual secondary vertexes;
- fast and reliable electronics: all the electronics systems interfacing the ATLAS detector must be tolerant to the very high dose of radiation affecting them and must be fast enough to cope with the large ATLAS data sizes;
- efficient trigger system: the trigger system must be able to select data in the most efficient way possible, in order to have optimal performance with the lowest rate possible.

The aforementioned requirements could not be satisfied by a unique monolithic detector; for this reason ATLAS is composed of several sub-detectors, each one with a specific task. As shown in Fig. 1.6, the sub-detectors have a cylindrical geometry and are built concentrically from the collision point outwards.

All the ATLAS sub-detectors can be divided in three main sections: *Tracker*, *Calorimeters* and *Muon Spectrometer*, that work together to measure energy and momentum of the traveling particles. A system of magnets surrounds the detectors with a magnetic field, so that the charged particles will be subjected to the Lorentz force and will curve.

Detector component	**Required resolution**	η **coverage** (measurement)	η **coverage** (Trigger)
Tracking	$\frac{\sigma_{p_T}}{p_T} = 0.05\% p_T \oplus 1\%$	$\|\eta\| < 2.5$	
EM calorimetry	$\frac{\sigma_E}{E} = \frac{10\%}{\sqrt{E}} \oplus 0.7\%$	$\|\eta\| < 3.2$	$\|\eta\| < 2.5$
Hadronic calorimetry (jets)			
barrel and end-caps	$\frac{\sigma_E}{E} = \frac{50\%}{\sqrt{E}} \oplus 3\%$	$\|\eta\| < 3.2$	$\|\eta\| < 3.2$
forward	$\frac{\sigma_E}{E} = \frac{100\%}{\sqrt{E}} \oplus 10\%$	$3.1 < \|\eta\| < 4.9$	$3.1 < \|\eta\| < 4.9$
Muon spectrometer	$\frac{\sigma_{p_T}}{p_T} = 10\%$ *at* $p_T = 1\ TeV$	$\|\eta\| < 2.7$	$\|\eta\| < 2.4$

Table 1.1: Requirements for each of the ATLAS detector subsystem. The minimal resolution required to achieve the expected results is presented for the pseudorapidity region the sub-detector is expected to cover. Pseudorapidity requirements for the trigger system of each sub-detector are also presented.

The innermost detector is the Tracker, also called *Inner Detector* (ID); its goal is to track the charged particles, in order to measure their curvature and hence to reconstruct their momentum. The ID performs not-destructive measures, meaning that the particles passing through it will not be absorbed and the energy loss is

negligible. The Inner Detector is surrounded by a *Central Solenoid* magnet, which provides a 2 T magnetic field.

Moving outwards, the next detectors encountered are the Calorimeters, whose purpose is to completely absorb the particles and measure their energy (destructive measure). Due to the intrinsic differences between electromagnetic and hadronic processes, the technology needed to absorb electromagnetic particles (mostly electrons and photons) is very different from the one needed to absorb hadrons. For this reason, the calorimeters are divided in *Electromagnetic Calorimeter* (EC) and *Hadronic calorimeter* (HC), with the latter placed after the former because hadrons are generally more penetrating and can pass through the EC.

The last ATLAS detector section is the Muon Spectrometer; muons are *Minimum Ionizing Particles* (MIP) and can pass almost unaffected through the rest of the detector. However, they cannot be ignored, since their importance is vital to perform any type of physical analysis. The Muon Spectrometer is in charge of tracking them and measuring their momentum. To bend the muons, another magnetic field surrounding the spectrometer is required. This magnetic field is produced by eight very large air-core superconducting barrel loops and two end-caps air toroidal magnets, all situated within the muon system.

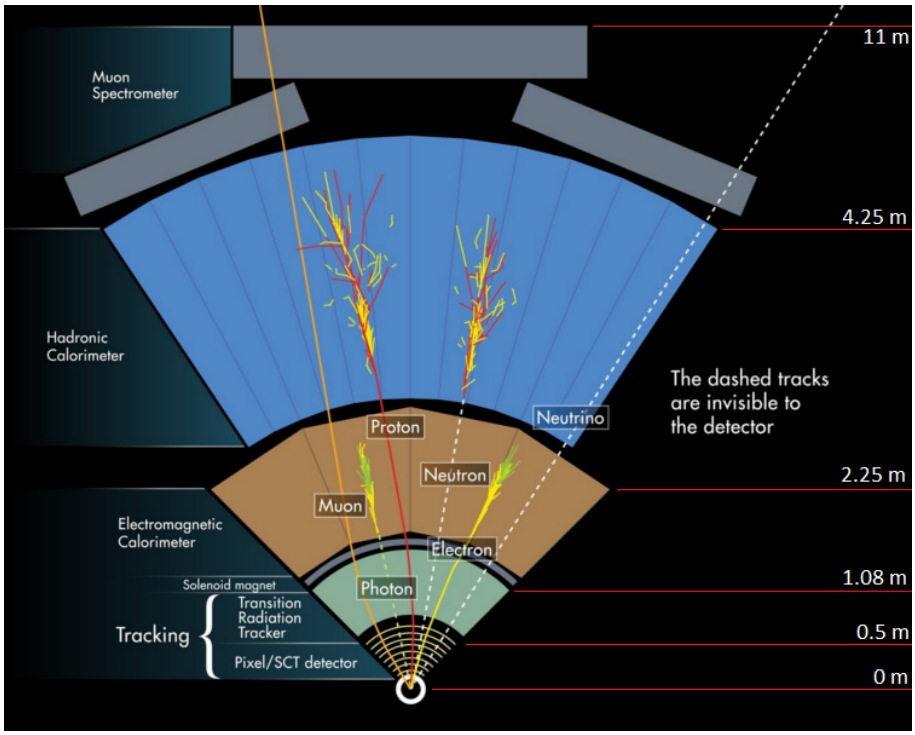


Figure 1.7: Section of the ATLAS detector, highlighting the sub-detectors and their interaction with the particles produced in the collisions. On the bottom, a vertical section of the pipeline in which the particles travel near the speed of light, colliding inside the ATLAS detector.

A section of the overall ATLAS detector, illustrating all the sub-detectors and their geometry, is shown in Fig. 1.7. Table 1.1 summarizes the requirements of each ATLAS subsystem, that will be described in a more detailed way in the next sections.

Inner Detector

The Inner Detector, shown in Fig. 1.8 and Fig. 1.10, is the innermost ATLAS subsystem, in charge of measuring the particles curvature and hence their momentum.

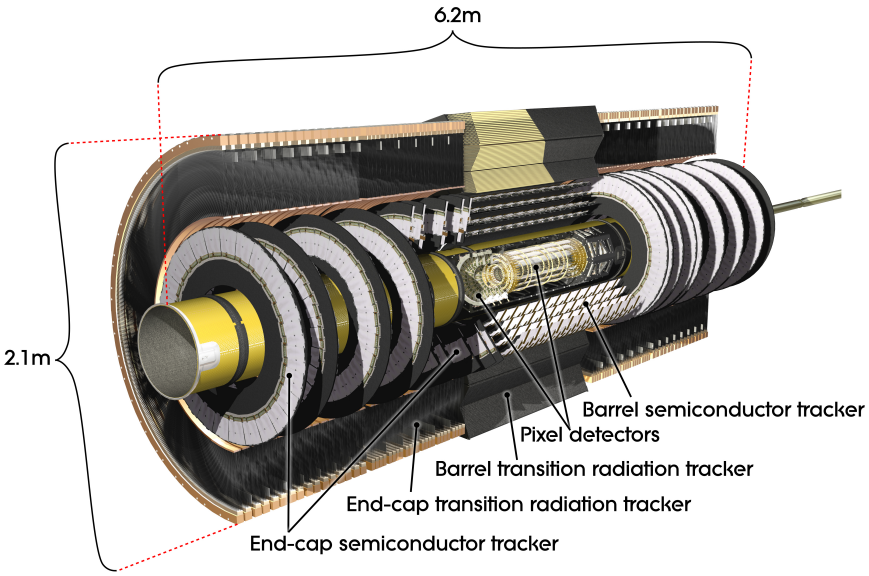


Figure 1.8: The ATLAS Inner Detector (ID). It is composed by Pixel Detector, SemiConductor Tracker (SCT) and Transition Radiation Tracker.

It is divided in three parts:

- ***Pixel Detector***: it is the closest ID sub-detector to the beam pipe and is composed of 3+1 barrel layers and three end-cap disks of pixel matrices, covering a pseudorapidity region $|\eta| < 2.5$. Being the central part of this thesis work, the Pixel Detector description is postponed to the next chapter;

- ***SemiConductor Tracker* (SCT)** [7]: it is composed of four cylindrical layers (numbered from 3 to 6) and two end-caps of silicon strips. Each end-cap consists of 9 disks and covers a pseudorapidity $1.1 < |\eta| < 2.5$, while the barrels cover a pseudorapidity $|\eta| < 1.1 - 1.4$, depending on the layer. The SCT basic structure is called module; each module comprises four silicon detectors, containing 768 readout strips, each of 80 μm pitch. The two sides of a module are glued together with a small (40 mrad) stereo angle to provide positional information along the z-direction. Differently from pixels, offering excellent 2-d

resolution, strips provide a better resolution over one privileged coordinate. To achieve better performance, the strips of two consecutive barrels are aligned over different orthogonal directions; with this configuration, called *u-v orientation*, the SCT reaches a spatial resolution of 17 μm in the $R - \phi$ direction, and of 580 μm in the z direction. A summary of the SCT barrel properties is given in Table 1.2;

Barrel	Radius	Modules	z-alignment
Barrel 3	299 mm	384	parallel
Barrel 4	371 mm	480	perpendicular
Barrel 5	443 mm	576	parallel
Barrel 6	514 mm	672	perpendicular
Total		2112	

Table 1.2: The SCT barrel geometry: the radius is to the centre of a module; the last column gives the u-v orientation. All the barrels have a length of 1492 mm. [7]

- ***Transition Radiation Tracker* (TRT)** [8]: it is the outermost part of the Inner Detector and is composed of a cylindrical barrel and two end-caps. Differently from the Pixel Detector and the SCT - which use semiconductor technology - the TRT is a gaseous detector, composed of carbon-polyamide straw tubes filled with a mixture of Xe (70%), CO_2 (27%) and O_2 (3%). The straws are contained in *modules*, which are the basic component of the detector. Each module is a quadrilateral prism with front and back faces in a plane perpendicular to the local radial ray. There are three types of modules of different sizes, as shown in Fig. 1.9; the main properties or the modules are shown in Tab. 1.3.

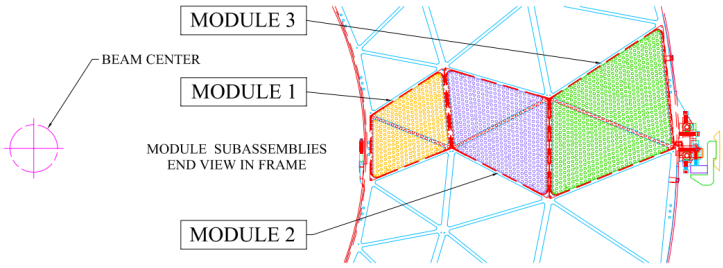


Figure 1.9: The three types of modules are mounted in the Barrel Support System. The orientation with respect to the beam intersection area is shown to scale. The triangular sections on the space frame are radially symmetric. [9]

Particles produced at LHC collisions travel trough the straw tubes at relativistic speed, emitting *transition radiation* photons whose intensity is proportional to the particle relativistic factor $\gamma = E/m$. The choice of the gas mixture was

Module	**Inner Radius (m)**	**# of modules**	$\|\eta\|$	**# Straws**	Mass (kg)
Type 1	0.56 m	19	1.06	329	2.97
Type 2	0.70 m	24	0.89	520	4.21
Type 3	0.86 m	30	0.75	793	6.53
Total for Barrel	73			52544	439

Table 1.3: The SCT barrel geometry: the radius is to the centre of a module; the last column gives the u-v orientation. All the barrels have a length of 1492 mm. [9]

done to maximize the transition radiation photons absorption, to provide good electrical stability and to have a high tolerance to radiation. Information over the particles γ is very useful to discriminate different particles, especially for $e/\pi^{\pm}$ identification. TRT is able to track particle position with an intrinsic resolution of 120 μm at low pile-up, and with a resolution of 200 μm if the detector occupancy (due to high pile-up) is above 80%. While those values are low when compared to the rest of the Inner Detector, the additional informations provided by the TRT together with the particle discrimination capabilities have a great contribution in the track definition and hence in the momentum resolution.

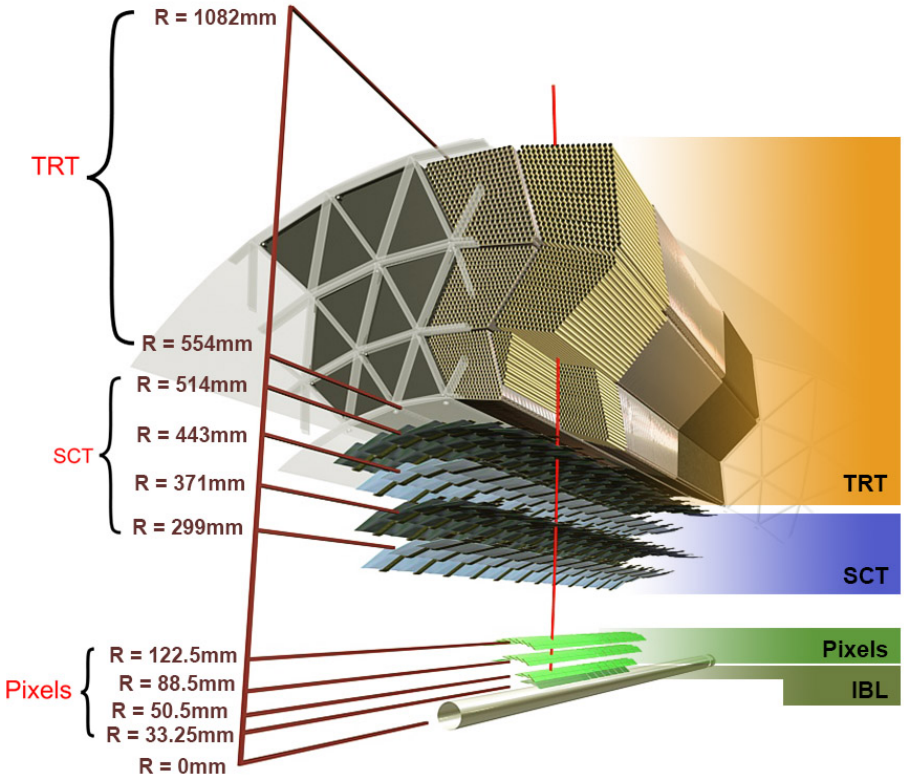


Figure 1.10: Section of the ATLAS Inner Detector (ID). All the sub-detectors and their positioning respect to the beam pipe are shown.

An overall summary of the Inner Detector is illustrated in Table 1.4.

Table 1.4: Summary of the characteristics of each Inner Detector sub-detector.

	Hits/track	Element size	Hit resolution [μm]
Pixel, $\|\eta\| < 2.5$			
4 barrel layers	3	$50 \cdot 400 \mu m^2$	10(R $-$ ϕ), 115 (z)
2 $\times$ 3 end-cap disks			10(R $-$ ϕ), 115 (R)
SCT, $\|\eta\| < 2.5$			
4 barrel layers	8	$50 \mu m$	17(R $-$ ϕ), 580 (z)
2 $\times$ 9 end-cap disks			17(R $-$ ϕ), 580 (R)
TRT, $\|\eta\| < 2.0$			
73 barrel tubes	~ 30	d=4 mm, l=144 cm	130/straw
160 end-cap tubes		d=4 mm, l=37 cm	

Calorimeters

The ATLAS calorimetry system, shown in Fig. 1.11, is in charge of measuring the energies of the particles, while at the same time providing some information on their position and identity. The ATLAS calorimeter is divided into subsystems, because of the different nature of electronic and hadronic interactions and the requirements on resolution and pseudorapidity coverage.

The subsystems can be divided in two main categories:

- ***Electromagnetic Calorimeter* (ECAL)** [10]: it is in charge of absorbing and measuring the energy of electromagnetic particles, i.e. electrons and photons. It is a *sampling* calorimeter, using lead as the absorber and Liquid Argon (LAr) as active area. The Liquid Argon was chosen as active material because it is relatively dense - so no signal amplification is needed - the signal response is linear with the energy, it is stable with time and it is resistant to radiation. The ECAL is composed of an *ElectroMagnetic Barrel* (EMB), covering a pseudorapidity $|\eta| < 1.475$ and an *ElectroMagnetic End-Cap* (EMEC), covering $1.375 < |\eta| < 3.2$. When electrons or photons crosses the ECAL, they produce an *electromagnetic shower*, whose shape is significantly different from the *hadronic shower* produced by hadrons. The Electromagnetic Calorimeter design uses this difference in the shower development to discriminate between photons and neutral pions π^0. In fact, the barrel is divided in three main regions along the z-axis; the first region, called η*-strip layer* is finely granulated to exploit the shower structure and is able to distinguish between photons and pions over a wide energy range ($\sim 5\,\text{GeV}$ - $\sim 5\,\text{TeV}$). The other regions feature high granularity as well, which is of extreme importance in the reconstruction of the missing transverse energy. The length of the calorimeter is constrained by the need of fully containing the electromagnetic showers; since the 99% of

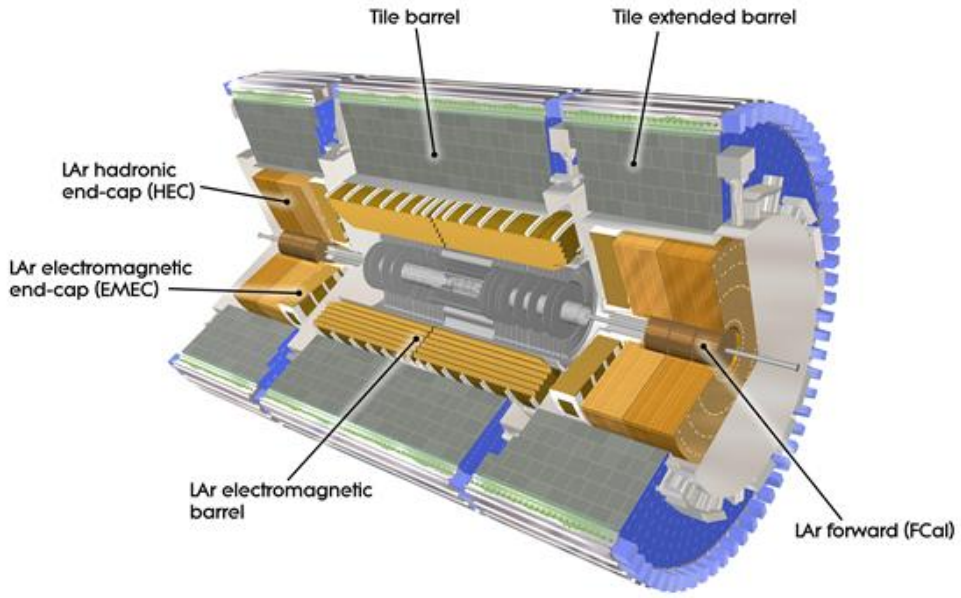


Figure 1.11: View of the ATLAS calorimeter system, for measures of the energies and positions of charged and neutral particles. It consists of a Liquid Argon (LAr) electromagnetic calorimeter and an Hadronic Calorimeter. Interactions in the absorbers transform the energy into a "shower" of particles that are detected by the sensing elements.

the shower is contained in 20 radiation lengths (X_0), the total thickness of the ECAL is 22 X_0 in the barrel region and 24 X_0 in the end-caps. The resolution on energy measurements of the Electromagnetic Calorimeter was measured to be

$$\frac{\sigma(E)}{E} = \frac{9.4\%}{\sqrt{E(GeV)}} \oplus 0.1\%$$

where the first term is the stochastic term and the second is the constant term.

- ***Hadronic Calorimeter*** **(HCAL)** [11]: it is in charge of absorbing and measuring the energy of hadronic particles and jets. It is composed of a *Hadronic Tile Calorimeter* (HTC), - which covers a pseudorapidity interval $|\eta| < 1.7$ - a *Hadronic End-Cap Calorimeter* (HEC) - which covers $1.5 < |\eta| < 3.2$ - and a *Forward Calorimeter* (FCAL) - which covers $3.1 < |\eta| < 4.9$.

  The HTC itself is divided in a Tile barrel (covering a pseudorapidity region $|\eta| < 1.0$) and two smaller tile barrels, called Tile Extended barrels (covering $0.8 < |\eta| < 1.7$). It is a sampling calorimeter (4 radiation lengths for the barrel and 1.8 radiation lengths for the extended barrels), using steel as absorber and scintillator plates as active material; the scintillators are coupled to photomultiplier tubes through wavelength shifting fibers. The energy resolution of the HTC combined with the Electromagnetic Calorimeter to isolated charged

pions is:

$$\frac{\sigma(E)}{E} = \frac{52\%}{\sqrt{E(GeV)}} \oplus 3\%$$

The HEC is also a sampling calorimeter (12 radiation lengths), using copper as absorber and Liquid Argon (as the ECAL) as active material. The energy resolution for charged pions is:

$$\frac{\sigma(E)}{E} = \frac{71\%}{\sqrt{E(GeV)}} \oplus 1.5\%$$

The FCAL is meant to provide very high pseudorapidity coverage for both hadronic and electromagnetic calorimeters. It is a sampling calorimeter (2.6 radiation lengths) using again Liquid Argon as active material, and copper and tungsten as absorber. Its energy resolution is:

$$\frac{\sigma(E)}{E} = \frac{94\%}{\sqrt{E(GeV)}} \oplus 7.5\%$$

Muon Spectrometer

Being *Minimum Ionizing Particles* (MIP), muons can travel through both the calorimeters without being absorbed. The *muon system*, which is the outermost ATLAS sub-detector, is in charge of measuring the momentum of muons in a pseudorapidity range $|\eta| < 2.7$. As shown in Fig. 1.12, the muon system contains both the *Muon Spectrometer* and the toroidal magnets needed to bend the tracks, so that the momentum can be measured. The toroidal magnetic field is provided by a *Barrel Toroid* and two *End-Cap Toroids*, that together can be seen as a unique big magnet; the field is not uniform inside the detector, and varies between 2 and 8 Tesla.

While it is extremely important to measure with high precision and resolution the muon momentum, the informations on the transverse impulse p_T are also valuable when selecting the interesting events to be analyzed. For this reason the Muon Spectrometer is divided in two main functional parts: the *Precision Chambers*, that are very precise but slow, and the *Trigger Chambers*, that are faster and contribute to the ATLAS Level-1 trigger.

The **Precision Chambers** are composed of several sub-detectors:

- ***Monitored Drift Tubes*** **(MDT)** [12]: they are *drift chambers* of two multilayer drift tubes, with 30 mm diameter aluminium walls filled with a gas mixture made of Argon (93%) and CO_2 (7%). The position of the particle is obtained by measuring the drift time in a single tube; MTDs are optimized to produce precise measurements of the z coordinate, and reach a spatial resolution of 80 μm. The Monitored Drift Tubes cover a pseudorapidity region $|\eta| < 2$;

- ***Cathode Strip Chambers*** **(CSC)** [13]: they are multi-wire proportional chambers with segmented cathode readout. The cathodes are segmented in strips; one has the strips parallel to the wires and the other has strips perpendicular to them, to measure with good precision the 2-dimensional coordinates. The chambers are filled with a gas mixture composed of Ar (80%) and CO_2 (20%). There are two types of chambers with different active area, called *Small* and *Large*; they are mounted on a disk structure called *ATLAS Small Wheel* and they cover a pseudorapidity region $1.0 < |\eta| < 2.7$. A summary on the CSC properties is given in Table 1.5.

Table 1.5: Summary of the characteristics of the Cathode Strip Chambers. [13]

Number of chambers	2×16
Number of layers/chamber	4
Layer separation	25 mm
Inclination angle	11.6°
Gas Mixture	Ar/CO_2 (80%/20%)
Anode-cathode distance	2.5 mm
Anode wire pitch	2.5 mm
Anode wire diameter	30 μm
Operating Voltage/Gain	1900 V / 6×10^4
Active area/chamber (Small)	0.50 m^2
Active area/chamber (Large)	0.78 m^2

The Muon Spectrometer **Trigger Chambers** cover a pseudorapidity region $|\eta| < 2.4$ and are composed of:

- ***Thin Gap Chambers*** **(TGC)** [14]: they are very thin multi-wire proportional chambers positioned on the end-cap region of the detector. The chambers are filled with a high quenching gas mixture, made of CO_2 (55%) and C_5H_{12} (45%). Since they must contribute to the Level-1 trigger, the signal generated must be very fast. This is achieved by using a very small anode-cathode spacing (leading to a very short drift time) and by operating in saturation regime. As a drawback, the resolution obtained is less precise; the TGC achieve a spatial resolution of 4 mm in the radial direction and 5 mm in the azimuthal direction. Nevertheless, the TGC spatial measurements are combined together with the results of the Precision Chambers to increase the resolution along the ϕ coordinate;

- ***Resistive Plate Chambers*** **(RPC)** [15]: they are gaseous detectors operating under a very high electric field, typically 4.9 kV/mm. The electrodes are made of a mixture of phenolic resins, which has a volume resistivity ρ_V between 10^9 and 10^{12} Ω·cm. The chambers are filled with a gas mixture made of $C_2H_2F_4$ (94.7%), C_4H_{10} (5%) and SF_6 (0.3%), whose main component is an

electronegative gas, with high enough primary ionization production but low free path for electron capture. Due to the combination of the gas and the high electric field, the RPCs work in *avalanche* regime, reaching a very good time resolution of $\sim 1\,\text{ns}$. The spatial coordinates are measured by strips parallel to the wires (η coordinate) and strips orthogonal to the wires (ϕ coordinate), with a resolution of $\sim 1\,\text{cm}$.

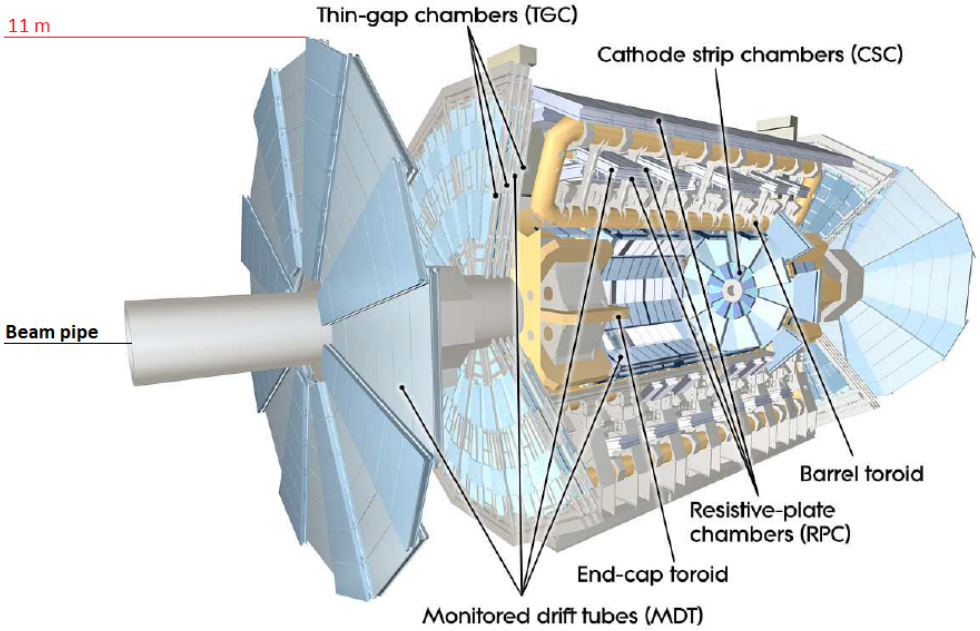


Figure 1.12: View of the ATLAS muon system, containing the *Muon Spectrometer* and the toroidal magnets (*Barrel Toroid* and *End-Cap Toroid*). The Muon Spectrometer can be divided in two: *Precision Chambers* - composed of *Monitored Drift Tubes* (MDT) and *Cathode Strip Chamber* (CSC) - and *Trigger Chambers* - composed of *Thin Gap Chambers* (TGC) and *Resistive Plate Chambers* (RPC).

1.2.2 Future upgrades

During LS2 (2018-2022) and LS3 (2024-2026), LHC will undergo a series of upgrades meant to increase its energy and luminosity, as discussed in Section 1.1.3. In the same years ATLAS will need to renovate some of its systems, in order to be able to sustain the more challenging environmental conditions and to maintain enough precision to carry on its physics program. Before Phase-I the ATLAS Detector will upgrade part of the Muon Spectrometer - with the introduction of the *New Small Wheel* (NSW) that will replace the Small Wheel - and part of the Calorimeters, in particular the LAr Calorimeter. Phase-II LHC will require a major upgrade of all the ATLAS

subsystems. Since it would require a very detailed description of each sub-detector - which is beyond the purpose of this work - the discussion of the upgrades will not be included in this thesis, with the exception of the Phase-II Inner Detector upgrade, that will be discussed in Chapter 4.

Chapter 2

ATLAS Pixel Detector

The ATLAS Pixel Detector is the innermost part of the ATLAS detector and the first stage for the tracking system. It serves different tasks, such as the identification of the primary collision vertex, the reconstruction of the secondary vertex and the ability to separate events coming from simultaneous collisions. This is possible thanks to its high granularity and very good 3-dimensional spatial and timing resolution. Due to its properties, the Pixel Detector exerts a particularly critical role in b-tagging and b-triggering operations. There are many experimental, technical and technological aspects that led to the design of the current ATLAS Pixel Detector, initially composed of three layers and six end-cap disks and subsequently upgraded with a fourth internal layer, called Insertable B-Layer (IBL). This chapter will provide a brief overview of the properties and problematics of a semiconductor solid-state detector and will focus on the techniques and strategies adopted by the ATLAS experiment to reach the required performance.

2.1 Semiconductor detectors overview

The semiconductor diode detectors, or solid state detectors, use the semiconductor materials properties (briefly detailed in Appendix A) to detect the passage of charged particles and measure their properties. The main advantages in adopting these detectors are the compact size of the sensors - resulting in a the possibility to reach a very high granularity - a good energy resolution and the relatively fast timing characteristics. However, the material budget and the cost per volume unit is sensibly higher when compared to gaseous detectors, both for the cost of the material itself and of all the electronic required to readout the elevated number of channels. For those reasons the semiconductor detectors are typically placed as close as possible to the beam pipe, to maximize the benefits of the high granularity and tracking resolution while at the same time reducing the volume needed. A typical semiconductor sensor consists in a p-n junction where an external electric field is applied. When a charged particle passes through such a detector, it loses some energy through ionization processes and the main significant effect is the production

of electron-holes couples along the track of the particle. The role of the electric field is to *reverse bias* the sensor, acting as an extension of the *depletion region* and hence extending the volume over which radiation-produced charge will be collected. In absence of the external electric field, the contact potential of about 1 V that is formed spontaneously across the junction is inadequate to generate enough mobility for the charge carriers, and as a result the performance of the sensor - such as the noise discrimination - would be very low. If the intensity of the reverse bias voltage applied is high enough, the depletion region thickness can extend to the wafer walls, creating a *fully depleted* detector, to be distinguished from the *partially depleted* detectors, where the depletion region does not reach the surface. For practical cases, the applied voltage is sufficient to create a fully depleted detector, while at the same time being relatively far from a potentially catastrophic breakdown. There are several advantages compared to the partially depleted sensors, such as a more uniform electric field inside the material and a lower capacitance (and hence signal rise time). One of the best features of the semiconductor detectors is that the quantity of energy necessary to produce an electron-hole couple, called *ionization energy* ϵ, is of the order of 3 eV, several times lower than the typical amount of energy necessary to create a ion pair in a gaseous detector (approximately $\sim$30 eV for the most common used gas). In a first approximation, the ionization energy is independent from the type of the incident radiation and its energy, although it is experimentally proven that the value ϵ is significantly higher for heavy ions than for electrons or alpha particles, resulting in a lower amount of charge carriers produced along the track. During the constructions of a real semiconductor detector, there are some particularly crucial operational parameters that must be taken in consideration and that will be detailed in the next paragraphs.

Leakage Current In every inverse polarized junction a small current (of the order of a fraction of microampere) is usually observed; this current is called *leakage current* and is originated both in the *bulk* volume of the semiconductor and from surface effects. The direction of the polarization and the voltage applied to the detector solves the purpose of repelling all the majorities charge carriers (electrons or holes) away from the junction, towards the extremity of the depletion region. However, minority carriers are continuously present, due to thermal generation, and their motion will oppose the one of the majority carriers, producing a small leakage current. This effect is highly influenced by the volume and the temperature of the semiconductor and, while for practical use of silicon detectors the value of the current is very small, the effects in germanium detectors are not negligible and particular precautions must be adopted. Surface leakage effects take place at the edges of the junction where high voltage gradients are present in a small distance. There are several factors influencing the surface leakage current, such as humidity, the type of encapsulation used or the amount of contaminations present. The main effects of the leakage current are the overall reduction of energy resolution (since the leakage current will bias the current produced by the passing of a charged particle) and the

screening of the effective potential applied to the semiconductor. It's a good rule to monitor the leakage current during detector operations; during steady operations the current should be stable and any sudden variation could indicate a variation in the state of the system and energy resolution. Also, monitoring the long-term effects of the leakage current provide a good estimate of the amount of radiation damage inside the detector.

Radiation Damage As already stated, thanks to their properties, solid state detectors are often placed very close to the particle beam, where the radiation dose is extremely high. Unfortunately the functionalities of semiconductors are strictly related to the near perfection of the crystalline lattice structure that can be severely damaged by the incoming radiation. While the ionization processes leading to the creation of an electron-hole couple are reversible, non-ionizing energy losses acting on the atoms of the lattice, mostly due to the passage of heavy ions, creates irreversible changes to the structure. The radiation-induced damages affect both the bulk and the surface of the sensor. The main type of bulk radiation damage is the *Frenkel defect*, produced when the interaction with an incoming particle displaces an atom of the semiconductor material from its normal lattice site. The vacancy that is created and the atom now placed in an interstitial position constitute a *trapping site* for the normal charge carriers.

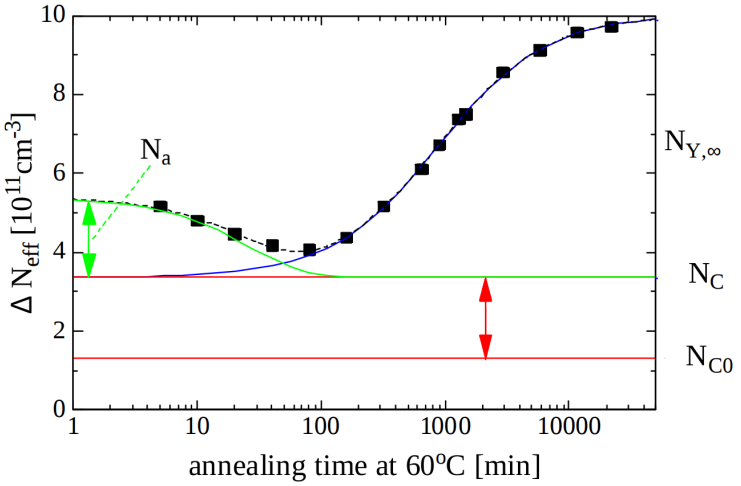


Figure 2.1: Long time dependence of $\Delta N_{eff}(t)$ of a silicon detector irradiated with a fluence of $\Phi_{eq} = 1.4 \cdot 10^{13}\, cm^{-2}$ and storage at a temperature of T = 60°C. N_c is the contribution of stable primary defects; N_a is the contribution of defects disappearing with time due to annealing; N_y is the contribution of secondary defects developing with time due to reverse annealing. [16]

According to the type and the energy of the incident particle, these trapping sites could involve a single atom, creating the so-called *point defects*, or could affect a cluster of atoms along the track. Since the energy loss per distance is greater for

heavy ions, the number of Frenkel defects produced by those particles is very high, and is estimated to be about 100-1000 times greater than that produced by an alpha particle [17]. One interesting effect of this process is the *annealing*, where interstitial atoms already present in the crystal occupy one of the created vacancies and both the defects disappear. After some time, however, the annealing process inverts and secondary defects develop and worsen the radiation damage with time; this phase is called *reverse annealing*. The effects of annealing and reverse annealing influence the effective doping concentration N_{eff}, as is shown in Fig. 2.1. The decrease of N_{eff} is generally beneficial, since the minimum voltage required to fully deplete the semiconductor is directly dependent on the amount of doping concentration. Another effect of radiation damage is the increase of the leakage current. This effect mostly concern the surface of the detector and is closely related to the ionization created within the oxide, causing an increase in leakage current fluctuation and hence a loss of detector resolution.

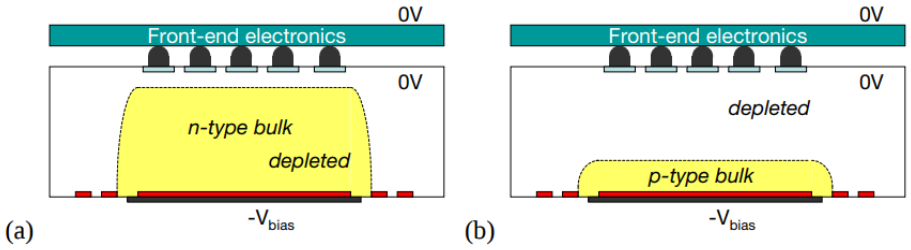


Figure 2.2: Comparison of a n^+-in-n pixel sensor structure before (a) and after (b) type inversion. Before type inversion the electrical field grows from the backside (top) and reaches the pixel implants (full depletion). After type inversion the depletion zone grows from the pixel side and allows operation even if the bulk is not fully depleted. [18]

A third consequence of radiation in solid-state sensors is the *type inversion*, that involves n-type semiconductors. The irradiation produces mainly acceptor like defects and removes donor type defects.

The effective doping concentration N_{eff} decreases until the n-type silicon changes its behavior becoming a p-type silicon. After the inversion point the effective concentration begins increasing again. As shown in Fig. 2.2, the polarity of the bias voltage required for depletion during normal operation is the same before and after type inversion, although eventually the voltage level required for full depletion will exceed the breakdown limit, and the device will have to operate in partial depletion mode.

The evolution of charge densities after the type inversion and the effects of reverse annealing can be greatly moderated by adding oxygen impurities to the silicon, improving the long term performance of the detector. This behavior is showed in detail in Fig. 2.3.

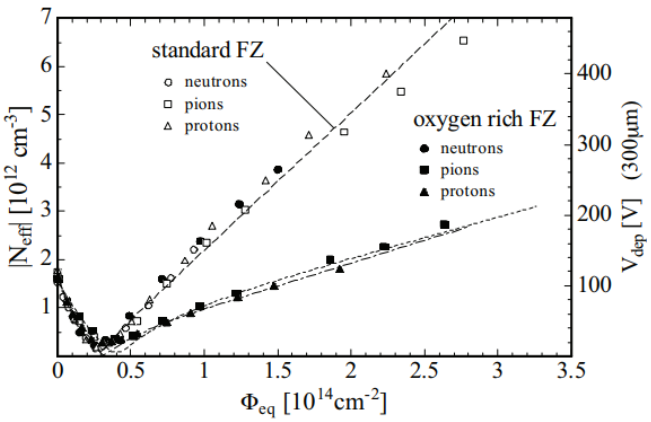


Figure 2.3: Evolution of effective charge densities and full depletion voltage in standard and oxygenated silicon during irradiation with various hadrons. In oxygenated silicon the increase after type inversion induced by charged particles is significantly lower. [18]

Energy Resolution and Noise In every type of detector, one of the main aspects to take in consideration is the energy resolution, which is generally particularly good for semiconductor detectors since the ionization energy is relatively low and the amount of collected charge and statistic is high. The energy resolution is influenced by the noise present in the detector. In semiconductors, some sources of noise are the fluctuations of the leakage current, both the one generated in the bulk and the one generate on the surface, and the noise associated to poor electrical contacts or series resistances. Other contributions to the degradation of the energy resolution are due to the *trapping effects* - where some impurities in the lattice act as a trapping site for the charge carriers - and to the fluctuations of energy loss of incident particles. When detecting particles with a very high dE/dx such as heavy ions, the energy loss that take place before the active volume of the detector can be significant and will affect the overall performance of the measure. The space where the particle looses its energy without being detected is called *dead layer* and its thickness depends on the technology and techniques involved in the realization of the semiconductor. Another mechanism involving mostly heavy ions and contributing to the degradation of energy resolution is the *pulse height defect*, defined as the difference between the true energy of the heavy ion and its measured energy.

Fig. 2.4 shows the relation for different type of incident radiation of the true energy and the measured output of a silicon detector, showing differences up to 3-7 MeV. There are three main mechanisms contributing to pulse height defect; the first one is related to the presence of the dead layer, since heavier particles lose more energy per length unit and the amount of energy undetected inside the dead layer will be higher. The second contribution to the pulse height defect involves the fact that heavy particles tend to lose energy with other mechanisms apart from ionization, such as nuclear recoil, which does not result in the production of an electron-hole pair. Lastly, since the electron-holes pair density produced by heavy

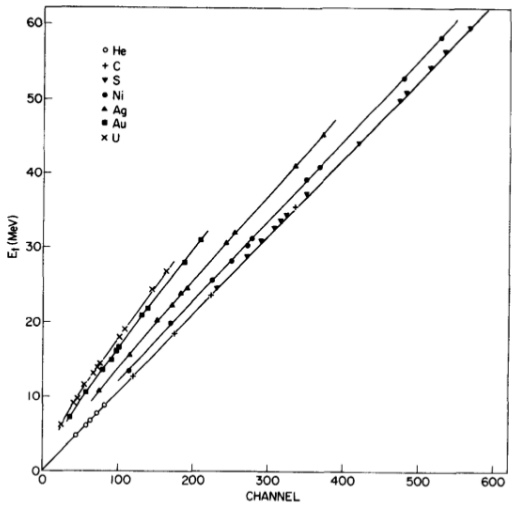


Figure 2.4: The true energy of various ions versus the pulse height channel number from a silicon surface barrier detector. Different particles having the same energy produce a different response in the pixel detector. [17]

ions is generally very high, the recombination probability increases. This last effect can be reduced by augmenting the bias voltage over the sensor.

Pulse Rise Time Solid state detectors are among the fastest radiation detectors available, since under normal operation the pulse-rise time is of the order of $\sim 10\,\text{ns}$ (See Appendix A.1). The detector pulse rise-time has two main contribution: *charge transit time* and the *plasma time*. The charge transit time is the migration time of the electron and holes formed within the depletion region towards the surfaces. It is highly influenced by the intensity of the applied voltage and the width of the depletion region, which corresponds to the dimension of the sensor in case of fully depleted detectors. The plasma time, on the other hand, is observed if the incident radiation is composed by heavy charged particles. In this scenario, the very high amount of electron-hole couples formed in a short distance is sufficient to create a plasma-like structure inside the material, shielding the effects of the external voltage applied and slowing the charge collection. The plasma time can hence be roughly defined as the time required for the plasma charge cloud to disperse to the point where normal collection proceeds. The effects of the plasma time consist in a delay of several nanoseconds in the signal formation. This delay has been measured to be on the order of 1-3 ns for alpha particles and 2-5 ns for heavy ions [17].

Channeling The crystalline structure of semiconductors confers a privileged orientation to the detector, that will respond to incoming radiation according to the incident angle. In fact, particles that travel parallel to the crystal planes usually show a lower rate of energy loss if compared to particles directed in an arbitrary direction. This effect can be minimized by orienting the silicon cut so that the crystalline

structure is perpendicular to the wafer surface.

2.2 The ATLAS Pixel Detector

The performance requirements formulated in the ATLAS Inner Detector Design Report (TDR) [19] brought to the original design of the ATLAS Pixel Detector composed of three concentric barrel layers and three endcap disks per each side, as shown in Fig. 2.5.

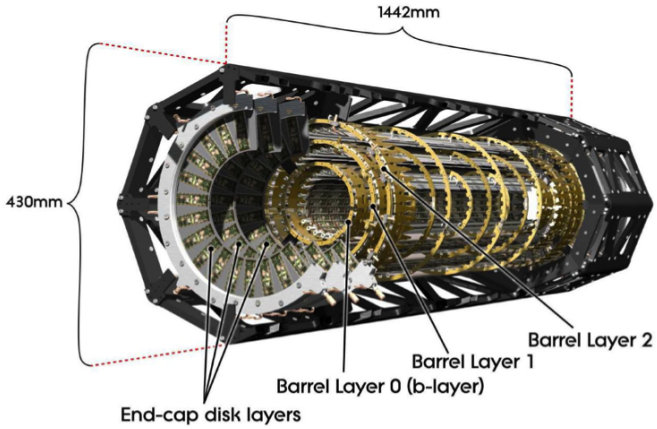


Figure 2.5: A schematic view of the active region of the ATLAS Pixel Detector consisting of barrel and endcap layers. [18]

The general requirements for the pixel detector were the coverage of the maximum possible solid angle (pseudorapidity $|\eta| < 2.5$), a good three dimensional vertexing capability (with a transverse impact parameter resolution $\leq 15\,\mu m$), the minimization of material usage in all the elements of the system - to reduce particle energy loss inside the detector - a good primary ($\sigma_z < 1\,mm$) and secondary vertex reconstruction, an excellent pixel efficiency and a radiation hardness to operate after an estimate lifetime dose of 500 kGy. The choice of a three-point pixel hit layout, the placement of the innermost barrel layer (B-layer) at 5 cm from the beam pipe and the usage of the smallest technologically achievable pixel size at the time fulfilled all the requirements. The silicon pixel sensors, the front-end electronics and the control circuits are arranged in *modules*, which constitute the basic building block of the detector. The nominal pixel size is $50\,\mu m$ in the azimuthal (ϕ) direction and $400\,\mu m$ in the longitudinal (z, along the beam pipe axis, for the barrel sensors) or radial (for disks) direction. Each module contains 46080 pixel electronics channels. Modules are mounted in structures, providing mechanical support and cooling. Those supports are called *staves* in the barrel region and *disk sectors* for the disk layers. All the staves, each including thirteen modules, and the disk sectors have the same layout and are identical for all the layers. Tables 2.1 and 2.2 summarize the overall

properties of the Pixel Detector; as is shown, there are approximately 80 millions active pixel sensors covering an active area of about $1.7\,m^2$.

Table 2.1: Basic parameters for the barrel region of the ATLAS Pixel Detector system

Layer Number	Mean Radius [mm]	Number of Staves	Number of Modules	Number of Channels	Active Area [m^2]
0	50.5	22	286	13178880	0.28
1	88.5	38	494	22763520	0.49
2	122.5	52	676	31150080	0.67
Total		112	1456	67092480	1.45

Table 2.2: Basic parameters of the endcap region of the ATLAS Pixel Detector system

Layer Number	Mean Radius [mm]	Num. of Sectors	Number of Modules	Number of Channels	Active Area [m^2]
0	495	8	48	2211840	0.0475
1	580	8	48	2211840	0.0475
2	650	8	48	2211840	0.0475
Total	(1 endcap)	24	144	6635520	0.14
Total	(2 endcaps)	48	288	13271040	0.28

2.2.1 Sensors

The sensitive part of the pixel detector used to detect charged particles is an array of solid-state ionization chambers, composed by a n-type bulk semiconductor with implantations of high positive p^+ and negative n^+ dose regions on the sides of the wafer.

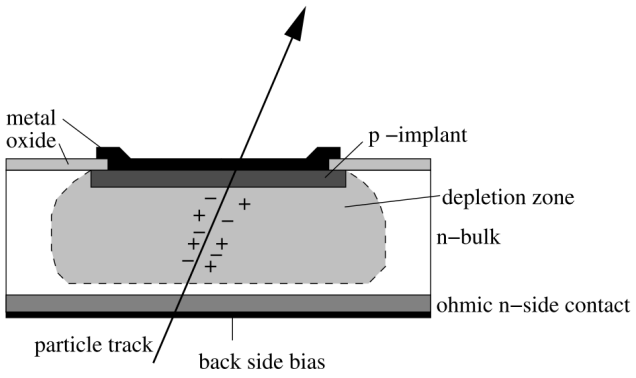


Figure 2.6: Single pixel layout.

Fig. 2.6 shows a schematic view of the pixel. The sensor is reverse polarized to extend the depletion region over the entire bulk volume; before type inversion the

depletion starts at the back (p) side, and after it moves to the front (n) side, as shown in Fig. 2.2.

The necessity of meeting exact geometry constraints, the technology available at the time of production and the necessity of sustaining a massive dose of radiation led to the choice of the material and the design of a $256 \pm 3\,\mu m$ thick n-bulk. Oxygen impurities have been added to the bulk to reduce the increase of the effective concentration of charges N_{eff} due to radiation damage (as shown in Fig. 2.3) and reverse annealing. The effects of the irradiation over N_{eff} and subsequently the minimum voltage required to fully deplete the sensors have been modeled and the results are shown in Fig. 2.7.

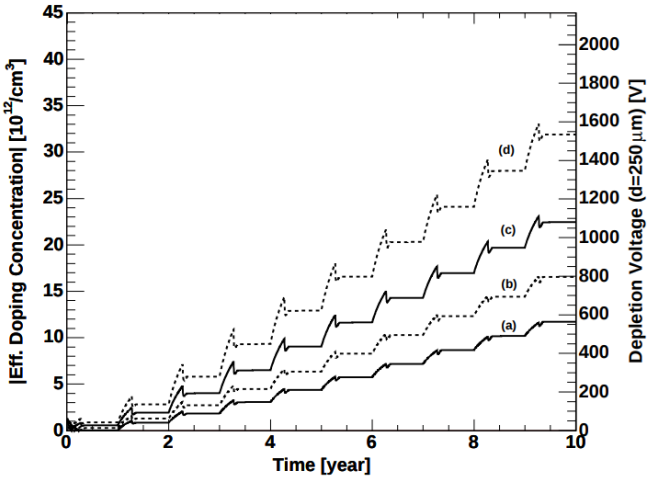


Figure 2.7: Change of the effective doping concentration and the minimum voltage necessary for full depletion of oxygenated sensors affected by irradiation in a standard (solid) and elevated (dashed) radiation scenario.
(a) Layer 1 at 8.85 cm radial distance from interaction point with a standard fluence of $0.9 \cdot 10^{14}\, cm^{-2} year^{-1}$, (b) the same as (a) with a 50% elevated fluence, (c) b-layer at 5.05 cm radial distance from the interaction point with a standard fluence of $2.4 \cdot 10^{14}\, cm^{-2} year^{-1}$, (d) the same as (c) with a 50 % elevated fluence.[18]

Each sensor tile contains 47232 pixels implants, arranged in 144 columns and 328 rows; the 88.9% of those pixels, positioned in 128 columns, have a size of $385.2 \times 30 \mu m^2$, corresponding to a $400 \times 50\,\mu m^2$ pitch and for the remaining 11.1% (in 16 columns), the size is $582.5 \times 30\,\mu m^2$ and the corresponding pith is $600 \times 50\,\mu m^2$. In each column eight pair of pixels are connected to a common read-out, and as a result the independent readout rows are 320 and the readout pixel channels are 46080. This arrangement was chosen to allow 16 front-end chips to interface one sensor tile. All the channels are connected to a common bias grid structure as shown in Fig. 2.8 to ensure isolation between pixels.

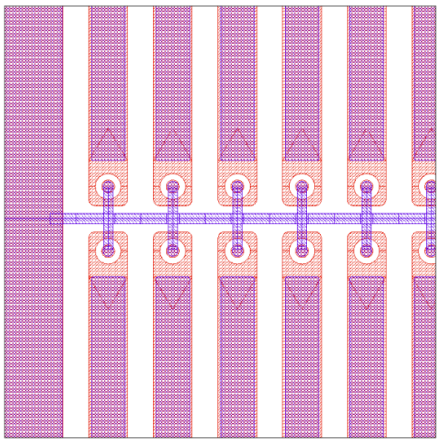


Figure 2.8: Layout detail of the bias grid visible in the production mask for a pixel double row. [18]

2.2.2 Front-end electronics

As already stated in the previous sections, the total number of readout channels for the ATLAS Pixel Detector is over 80 millions. The electronics involved in the readout of such a huge amount of channels must be adequate and carefully chosen. Since at least part of all the electronics is placed in contact with the sensor, hence very close to the beam pipe, particular care must be taken in the choice of radiation-hard materials and design. The full readout system, depicted in Fig 2.9 comprises front-end chips, an optical system for data transmission and off-detector electronics, whose discussion will be postponed to the next chapter.

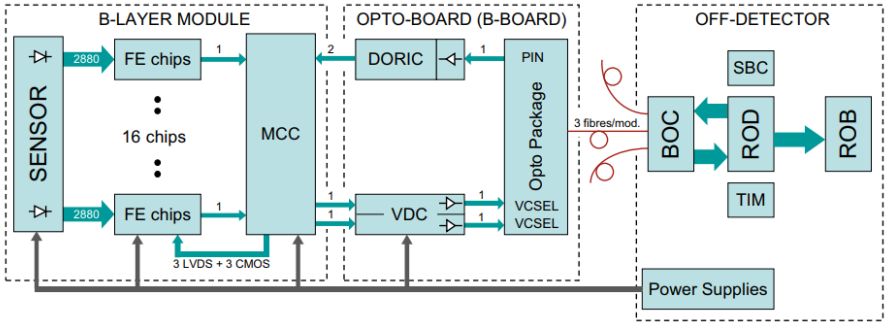


Figure 2.9: Block diagram of the pixel detector system architecture; front-end (left), optics (center) and off detector (right). [18]

Each sensor tile is connected to sixteen FE-I3 front-end chips arranged in two rows of eight chips; the FE-I3s are readout by a Module Control Chip (MCC) through Low Voltage Differential Signaling (LVDS) serial links. MCC data are then converted to optical signals by Opto-Boards and sent to the off-detector electronics.

FE-I3

Studies on the design of a front-end chip with properties adequate to be used by the ATLAS Pixel Detector readout started in 1996 and resulted in 2003 in the realization of the FE-I3 [20] which was realized using a 250 nm CMOS technology and a radiation-tolerant layout. A summary of all the front-end chips design attempts before FE-I3 is shown in Tab. 2.3.

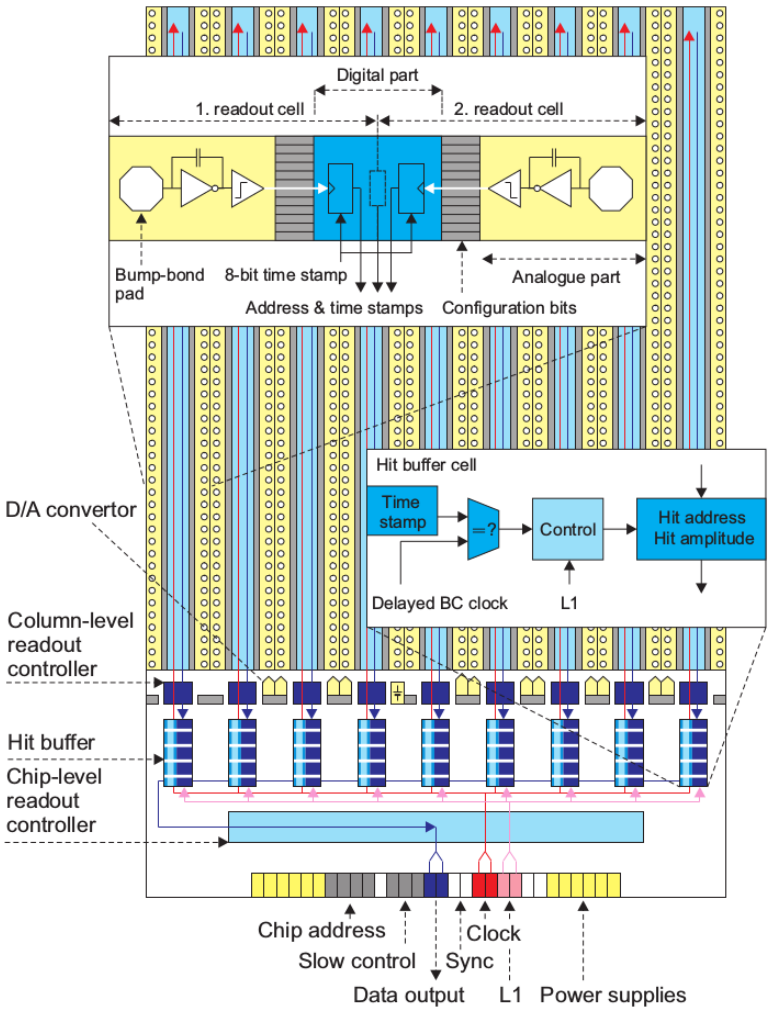


Figure 2.10: Schematic plan of the FE-I3 front-end chip with main functional elements. [18]

The FE-I3 contains 2880 pixels cells arranged in a 18×160 matrix and is composed of an analog circuitry part and a digital part; Fig. 2.10 represents a schematic view of the chip.

Each pixel cell is connected to an analog block where the charge signal is amplified before being digitalized. The charge-sensitive *preamplifier*, shown in Fig. 2.11 has a folded-cascode topology and is optimized to interface systems affected by radiation

Table 2.3: History of the ATLAS pixel front-end chips. [18]

Chip	Year	Cell Size $[\mu m^2]$	Col×Row	Transistors	Technology
Beer&Pastis	1996	50×436	12×63	-	0.8 μm BiCMOS
M27b	1997	50×536	12×64	-	0.8 μm CMOS
Marebo	1997	50×397	12×63	0.1 M	0.8 μm BiCMOS
FE-B	1998	50×400	18×160	0.8 M	0.8 μm CMOS
FE-A/C	1998	50×400	18×160	0.8 M	0.8 μm BiCMOS
FE-D1	1999	50×400	18×160	0.8 M	0.8 μm BiCMOS
FE-D2	2000	50×400	18×160	0.8 M	0.8 μm BiCMOS
FE-I1	2002	50×400	18×160	2.5 M	0.25 μm CMOS
FE-I2/I2.1	2003	50×400	18×160	3.5 M	0.25 μm CMOS
FE-I3	2003	50×400	18×160	3.5 M	0.25 μm CMOS

damage, that can produce leakage currents up to 100 nA. The main features of the preamplifier are a 5 fF capacitor with a current source continuous reset, a 15 ns rise-time and operation at about 8 μA bias.

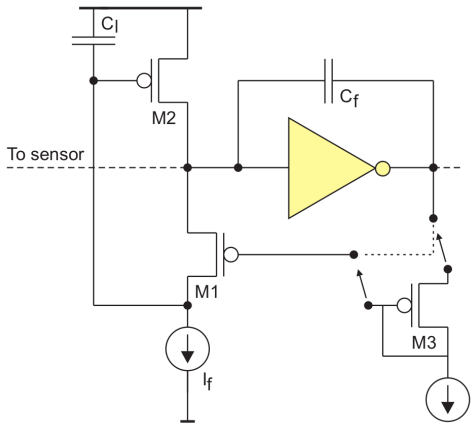


Figure 2.11: Charge preamplifier feedback circuit. M1 acts as a feedback capacitor reset, M2 provides leakage current compensation and M3 acts as a level shifter. [18]

To drain the leakage current and prevent it from influencing the reset circuitry (M2), a compensation circuit has been added (M1) using PMOS devices. Signals with high output amplitudes saturate the current provided by the reset device and, as a consequence, the return to the baseline (hence the pulse width) is nearly linear and directly proportional to the input charge. The signal amplitude can consequently be measured by the *Time over Threshold* (ToT), the width of the preamplifier output. The duration of the ToT is measured by counting cycles of the 40 MHz ATLAS bunch-crossing clock. After amplification, the signal passes through a *discriminator* where is compared to a programmable threshold, provided by a threshold generator integrated in every pixel. The value of the threshold, the threshold current and of

other parameters, such as pixel masking, shutdown or select, are configured in the FE-I3 using a 5 MHz serial protocol and are handled by a *Pixel Cell control logic* shown in Fig. 2.12. There are 231 global configuration bits plus 14 bits for each of the 2880 pixels.

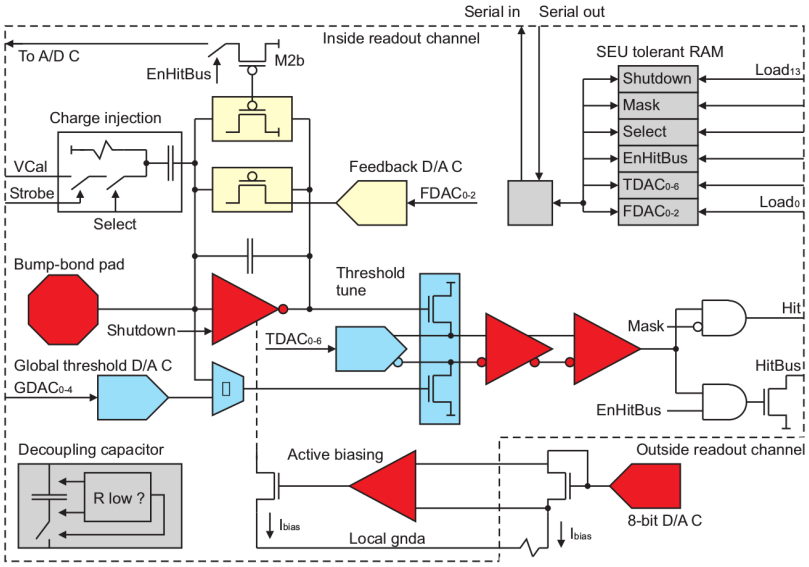


Figure 2.12: Pixel cell control logic diagram [18]

After discrimination, the digital part of the front-end chip assigns a 8-bit Gray-coded time stamp to the *leading edge* (LE) and *trailing edge* (TE) of each hit. When the hit processing is complete (i.e. the TE information is ready) the LE, TE and row number (8 bits) are transferred to the *End of Column* (EOC) buffers by a priority mechanism favoring the top rows. If a trigger arrives at a timing corresponding to the LE time stamp plus a programmable latency, the hit is flagged as belonging to that trigger and data are collected by the *chip-level readout controller*. This unit collects hits from off the EoC buffers, serializes the data and transmits them to the Module Control Chip (MCC). When a L1 trigger arrives, the current bunch-crossing identifier is stored in a First In First Out (FIFO) memory with a depth of 16 locations. The chip can hence keep track of 16 pending triggers and, if another trigger arrives before at least one FIFO location is emptied, this trigger will be *skipped*. This *skipped trigger mechanism* will be rediscussed in the next chapter and plays a key role in the synchronization of the system.

MCC

The *Module Control Chip* (MCC) [21] interfaces sixteen FE-I3 front-end chips and is in charge of the I/O system operation. It performs three main tasks: loading the configuration parameters into the chips, distributing timed signals such as L1 trigger or bunch crossing identifier and reading out the FE chips. The first prototype of this

chip (MCC-D0) was built in 1999 and the final version (MCC-I2), whose diagram is shown if Fig. 2.13, was realized in 2003.

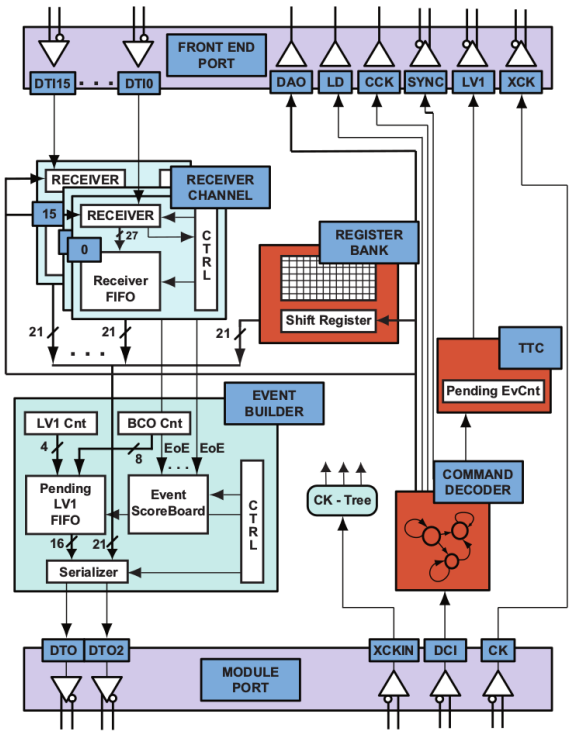


Figure 2.13: Module Control Chip (MCC) diagram. [18]

The off-detector DAQ system communicates to the MCC via a 40 Mb/s serial line to send configuration data and commands such as triggers or resets. After a trigger is received, the MCC proceeds to send it to all the FE-I3, as long as there are less than 16 triggers to be processed. In case of overflow, the trigger will not be propagated to the FE chips and will be lost; the number of lost triggers is stored in a 4-bit register and is propagated to the rest of the DAQ chain to keep synchronization. The MCC is also in charge of building the event by performing two concurrent processes: collection of data of all the 16 FE chips (*Receiver*), and event ordering and construction (*Event Builder*). The event are then transmitted to the off-detector electronics via a serial stream that can be run at 40, 80 (on a single line) or 160 Mbps (on two lines).

Radiation-hardness requirements

The design of the pixel detector front-end electronics, FE-I3 and MCC, was extremely influenced by the necessity of having a high radiation-tolerance, since the chips are positioned very close to the particle beam. The effects of radiation on transistor-based electronic devices can be divided in two big categories: *Total Ionizing Dose* TID effects and *Single Event Effects* (SEE). TID effects are mostly in-

dependent on the rate of the incident radiation and are only due to the cumulative effects of all the absorbed dose. The long-term effects of radiation on CMOS technology are related to the design and structure of MOS transistors, showed in Fig. 2.14.

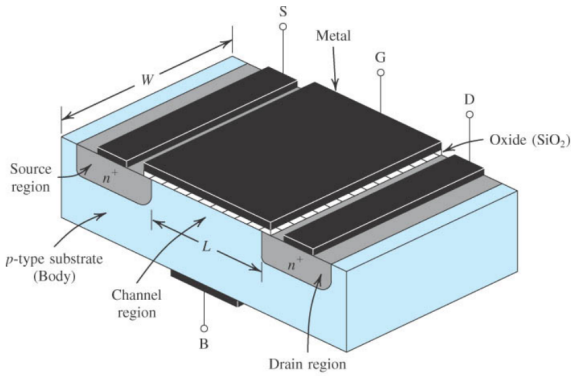


Figure 2.14: MOS transistor diagram. [22]

In particular the most dominant effects are the buildup of trapped charges in the gate oxide (used as an isolator), the increase of number of interface traps and the increase of the number of traps in the oxide bulk. Those effects result in an increase of the transconductance, hence the rise-time of the transistor, in an increase of the leakage current and in a shift of the operating parameters. A careful choice of constructing technologies, as well as an optimization of W/L, has a great influence on the resistance of the device to TID; using a smaller technology is also beneficial since the oxide thickness decreases as well, and with it the effects of charge accumulations and trapping. For this reason FE-I3 and MCC are implemented with 250 nm technology (corresponding to $\sim$ 5 nm oxide thickness), the smallest readout chip technology available at the time. Single Event Effects (SEE), on the other hand, depend on the particle radiation fluence and can be divided in *soft* effects or *hard* effects, where the integrity of the system is compromised. The most typical soft effects are the *Single Event Upset* (SEU) - happening when radiation-induced charge causes a transition in a memory cell and a bit-flip is generated (Fig. 2.15) - and the *Single Event Transient* - happening when the charge collected from an ionization event discharges in the form of a spurious signal traveling through the circuit. While in the case of SEU the 0→1 transitions and the 1→0 transitions are mostly symmetrical, the SET mechanism can cause asymmetries in the bit-flips. FE-I3 and MCC adopts triple-redundancy logic and detection-correction schemes to reduce SEU and SET, although an ulterior improvement can be provided by constantly refreshing the chip configuration, as will be discussed in next chapters.

An example of a *hard* SEE effect is the *Single Event Latchup* (SEL), occurring when particle radiation (in particular heavy-ionizing particles such as heavy ions) creates a latchup, shorting the PMOS and NMOS transistors and generating extremely high currents, potentially damaging the chip. To prevent irreversible

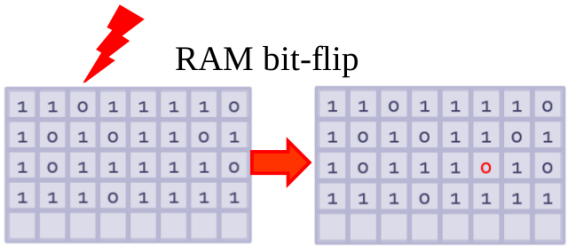


Figure 2.15: Example of Single Event Upset (SEU) in a RAM memory.

damages from SEL effects, the current must be constantly monitored and, in case of sudden changes, the transistor must be switched off.

2.2.3 Modules, staves and geometry

The ATLAS Pixel Detector is composed of 1744 identical modules, covering a total sensitive area of $\sim 1.7\,m^2$. Each module, shown in Fig. 2.16, is composed of a sensor tile, containing 47232 pixels, sixteen FE-I3 front-end chips bump-bonded (Fig 2.16) to the sensor, a *flex-hybrid* circuit - a flexible printed circuit of about $100\,\mu m$ thickness to route signals and power - a MCC chip positioned on the flex-hybrid, and a *pigtail* (only for barrel modules), a flexible foil providing connection to electrical services via a microcable. Each Pixel module has an active surface of $6.08\times1.64\,cm^2$; the active fraction is about 75%, caused by the dead-space between sensors and the fact that a sizable part of the front-end chip is dedicated to End Of Column logic.

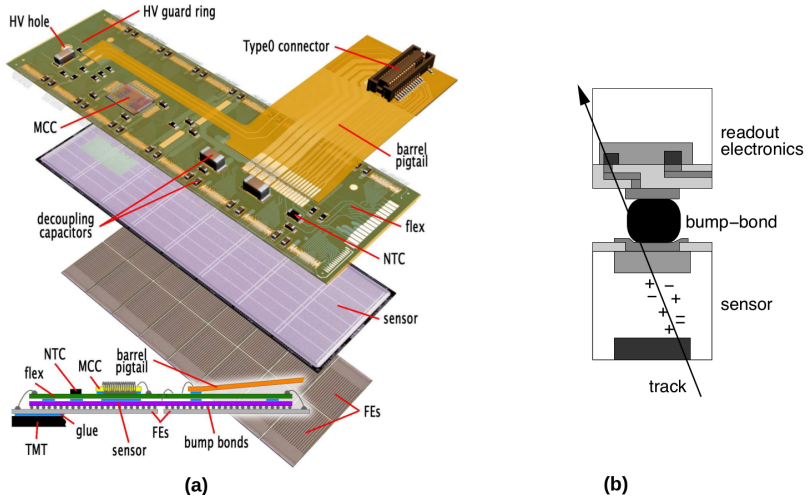


Figure 2.16: (a). The elements of a pixel barrel module. (b). Cross section of a hybrid pixel detector, showing one bump-bonded connection between a sensor and an electronics pixel cell. [18]

To compose the entire detector, thirteen modules are mounted on a supporting structure called *stave*. In total there are 112 identical staves; on each one, the modules are tilted on z axis by 1.1 degrees to face the interaction point as shown in Fig. 2.17 (b). Furthermore, in order of reduce the dead space at the extremity of the staves and to allow overlapping, the staves are tilted by 11-15 degrees on the x-y plane as shown in Fig. 2.17 (a); this solution offers a great coverage of the solid angle and a particle detection efficiency close to 100%.

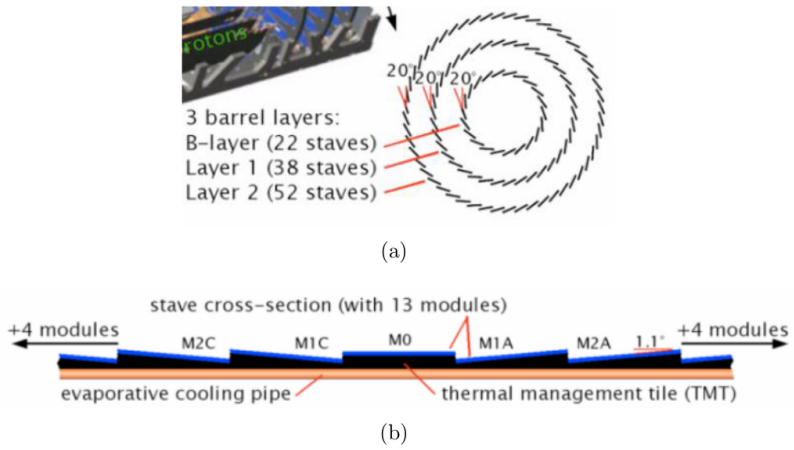


Figure 2.17: ATLAS Pixel Detector staves disposition around the beam pipe (a), and modules layout inside each stave (b).

2.2.4 Optical interface

All the communication between the front-end modules and the off-detector electronics happens via optical connection. The optical system is composed mainly of two boards: the *Opto-Board* on the module side and the *Back Of Crate* BOC board at the off-detector end, as shown in Fig. 2.18. The requirements that led to the design of the system were the need to sustain a high dose of radiation, to implement electrical decoupling and to minimize the material budget. For those reasons all the optical components are not mounted on the detector modules, but at a distance of about 1 meter from them, inside the Pixel Support Tube. The communication happens on individual fibers, one for trigger, clock and configuration (down-link) and one (or two) for front-end data transmission (up-stream). Down-link signals are encoded using *Bi-Phase Mark* (BPM) [23], while up-link signals use a *Non-Return-to-Zero* (NRZ) encoding [24]. In the Opto-Board, consisting in a beryllium-oxide (BeO) printed circuit, the down-link connection is implemented by a PiN diode array, that receives data coming from the BOC, and a *Digital Optical Receiver Integrated Circuit* (DORIC), which amplifies the signal and extracts clock and data from the BPM encoded signals. The DORIC was designed to withstand a total integrated dose of

170 kGy over 10 years of ATLAS operation while maintaining a bit error rate lower than 10^{-11}.

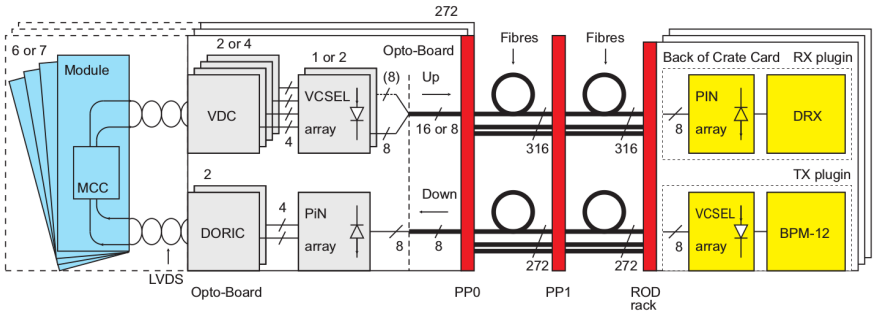


Figure 2.18: Optical link system architecture. [18]

The up-link connection is realized through a *Vertical-Cavity Surface-Emitting Laser* (VCSEL) array, used to optically transmit data with very low currents and fast rise/fall time, and a *VCSEL Driver Chip* (VDC), adapting LVDS signal (from the MCC) to single-ended signals suitable to drive the VCSEL.

2.2.5 Material usage

The purpose of the Pixel Detector, and of the Inner Detector in general, is to track charged particles without stopping them. For this reason it is important to minimize the material budget in the system layout. The *stopping power* of the detector can be quantified as the number of *radiation length* [25] seen by particles crossing at normal incidence. Simulation results are shown in Tab. 2.4, Tab. 2.5 and Fig. 2.19 [26].

2.3 IBL

The original ATLAS Pixel Detector operated since the very beginning of the LHC operations and at the time was the innermost part of the ATLAS Inner Detector. At the end of Run-1, however, the limitations of the tracker started to become evident and the need of a system improvement arose. The main problematics of the original detector layout were:

- the aging of the Pixel Detector was causing efficiencies in tracking identification and reconstruction;
- the Pixel Detector was designed to sustain a total integrated luminosity of 340 fb^{-1} ad a maximum instantaneous luminosity of $1 \cdot 10^{14}\, cm^{-2} s^{-1}$, while the maximum luminosity foreseen during Phase-I operations is higher;
- the need of tracking precision and b-tagging performance improvement.

Table 2.4: Pixel Barrel material. [26]

Item	Material	Size	Radiation Length (%)	
			no overlap	overlap
Pixel Sensors in B-Layer	Silicon	6.24×1.84 $cm^2 \cdot 200\mu m$	0.21	0.27
Pixel Sensors in Layer-1 & Layer-2	Silicon	6.24×1.84 $cm^2 \cdot 200\mu m$	0.27	0.34
Average			0.25	0.32
Interconnection hybrids + cables	Aluminium, Plastic	5.7×1.66 $cm^2 \cdot 200\mu m$	0.18	0.21
Opto-fibres			0.03	0.03
Stave		39.12·2.14 cm^2	0.66	0.88
Support structure	Carbon			0.22
Thermal barrier				0.07
Outboard cylinder	Carbon			0.07
Total average				1.80

Table 2.5: Pixel End-cap material. [26]

Item	Material	Size	Radiation Length (%)	
			no overlap	overlap
Pixel Sensors	Silicon	6.24×1.84 $cm^2 \cdot 200\mu m$	0.27	0.37
Interconnection hybrids + cables	Aluminium, Plastic	5.7×1.66 $cm^2 \cdot 200\mu m$	0.15	0.16
Electronic chips	Silicon	6.06×2.14 $cm^2 \cdot 200\mu m$	0.16	0.25
Support Disks		11.6-20 cm	0.73	0.73
Stave		39.12·2.14 cm^2	0.66	0.88
Support cones + thermal barrier	Carbon			0.11
Total				1.62

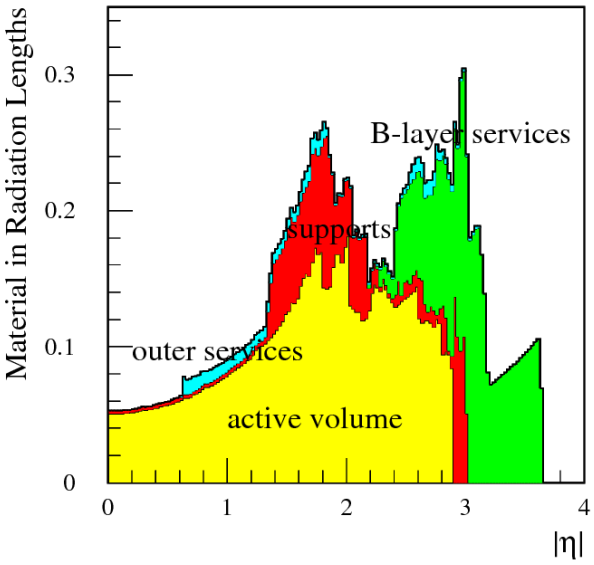


Figure 2.19: Material in Radiation Length in the Pixel system as a function of $|\eta|$. The *active volume* includes: sensors, hybrids, chips and links. The *supports* include mechanical supports and services inside a radius of 25 cm; the *outer services* include services outside this radius as well as the B-layer patch panel. [26]

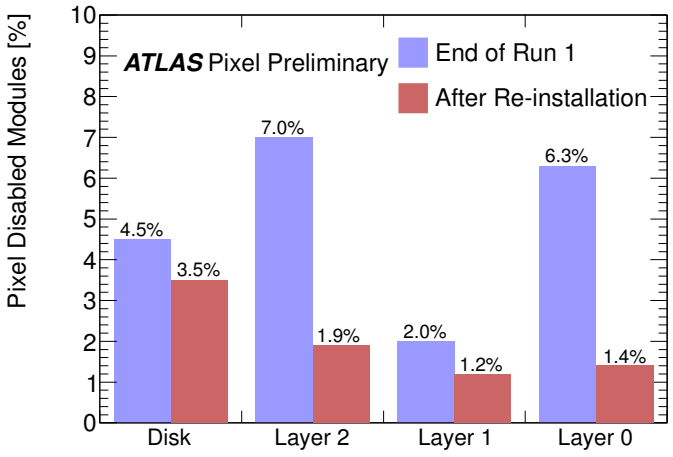


Figure 2.20: Percentage of disabled modules at the end of Run 1 and after the re-insertion of Pixel Detector into the ATLAS Experiment for disk and three layers. [ATL-INDET-SLIDE-2014-388]

The effects of time and radiation inevitably affect the sensors and electronic circuitry, causing aging and failures, which are classified in order of their effect over the system. A *Pixel Failure* is the lightest problem, since a single pixel is affected. It is mainly caused by disconnected bumps, electronic dead channels, and masked noisy channels. A channel is considered noisy if its occupancy is greater than 10^{-5} hits/bunch-crossing. The next problem in order of gravity is a front-end (FE-I3) failure, which affects 2880 pixels, followed by a MCC failure (16 front-end affected) and an opto-board failure (6-7 MCC affected). The status of the ATLAS Pixel Detector at the end of Run-I is shown in Fig. 2.20 and Fig. 2.21.

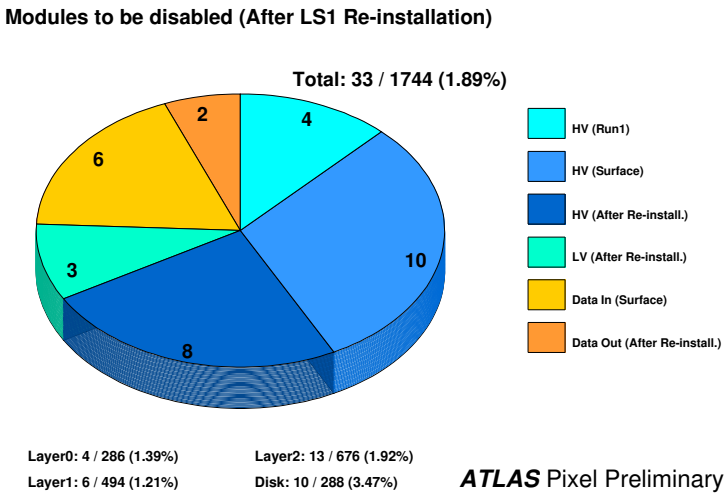


Figure 2.21: Number of modules of the Pixel Detector to be disabled after refurbishment and re-installation in ATLAS (May 2014) during LS1 classified by failure mode (High Voltage *HV*/ Low Voltage *LV*/ module configuration *Data In*/ data readout *Data Out*) and the phase of causing problems (End of Run1/ Surface / After re-installation). Modules having issues but being operable are not included. [ATLAS-COM-CONF-2014-043]

To solve the limitation of the current detector, an ulterior pixel layer, called *Insertable B-Layer* (IBL), was installed during LHC Long Shutdown 1 (LS1, 2013-2015) and is operating since Run-II. IBL is placed at about 33.5 mm from the beam axis, between the B-Layer and the beam pipe, which was substituted with a smaller one (inner radius of 23.5 mm) to fit the new detector. A schematic view of IBL is shown if Fig. 2.22. Due to the proximity to the beam, IBL must sustain a high dose of radiation dose, estimated to be of the order of $5 \times 10^{15}\, n_{eq}/cm^2$ during Phase-I operations. This problem, combined with the need of minimizing the material budget for the system, led to the design of new pixel sensors and a new front-end chip: FE-I4.

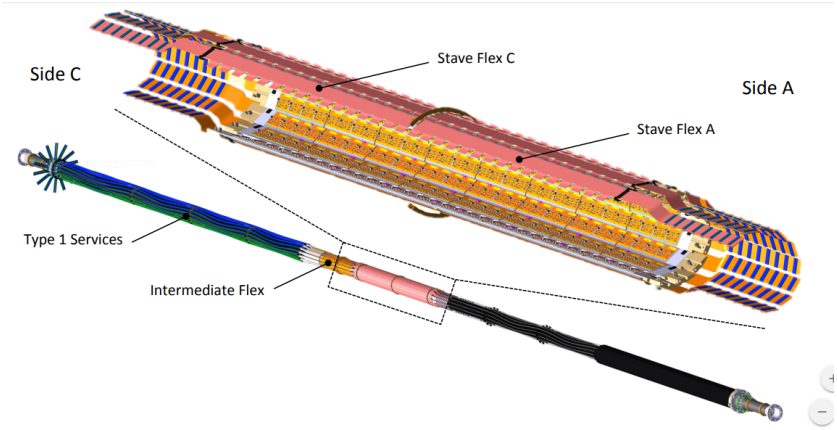


Figure 2.22: Longitudinal view of the IBL detector and its services. The insert shows an enlarged 3-dimensional view of the detector with its modules arranged cylindrically around the beam pipe. [27]

2.3.1 Sensors

For IBL pixel sensors two concurrent technologies were exploited: the well known planar design, already used in the rest of the Pixel Detector, and 3D sensors.

Planar design

Planar sensors have proven their excellent performance during Run-I operations and their technology is well-developed. IBL planar sensors design is similar to the one used for the barrel and disks, shown in Fig. 2.6, but with stricter requirements. In particular, the sensors were designed to resist to a fluence of up to $5\times10^{15}\, n_{eq}/cm^2$ and to reduce the inactive border from $\sim 1\, mm$ to $\sim 450\, \mu m$. To solve the last problem, the pixel length were reduced from $400\, \mu m$ to $250\, \mu m$ and the guard-rings were reduced from 16 to 13 and moved on the p-doped side. Fig. 2.23 shows a graphic scheme of the planar design.

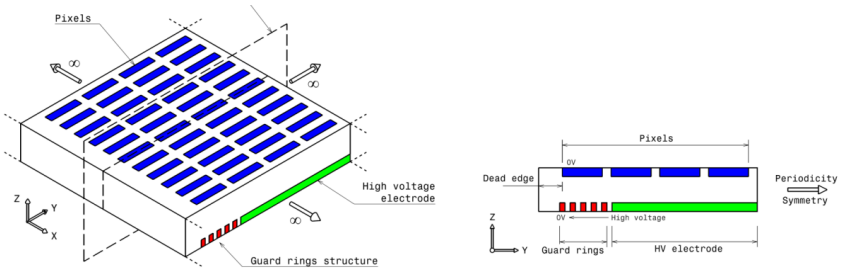


Figure 2.23: Planar Sensor graphic.

3D design

IBL was the first physics detector to implement a 3D pixel design in its structure. The main structural difference between 3D and planar sensors is that the electrodes penetrates the bulk in form of column instead of being implanted on the surface. With this configuration the depletion is parallel to the wafer surface and the pixel configuration is due to the position and distance of the doped columns, that are $\sim 10\,\mu$m wide. Since the column can be positioned at a distance smaller than the pixel thickness, the voltage necessary for full depletion is smaller, as well as the charge-collection distance and consequently the pulse rise-time. With a small depletion voltage, also the power dissipation due to current leakage is smaller, and with it also the cooling requirements. Another consequence of the reduced distance between electrode is that the probability of trapping sites formation in the active area is smaller, hence the sensor provide higher resistance to high radiation. On the other side, the capacitance is higher if compared to the planar sensors, causing higher pixel noise. In conclusion, the doped column are a not sensitive part of the detector and provide inefficiencies. This last problem can be reduced by tilting the pixels; IBL 3D sensors provide a 99.8% efficiency when tilted of $\sim 10°$. Two different 3D designs were fabricated at *Fondazione Bruno Kessler* (FBK) University in Trento and Centro National de Microelectronica (CNM) in Barcelona.

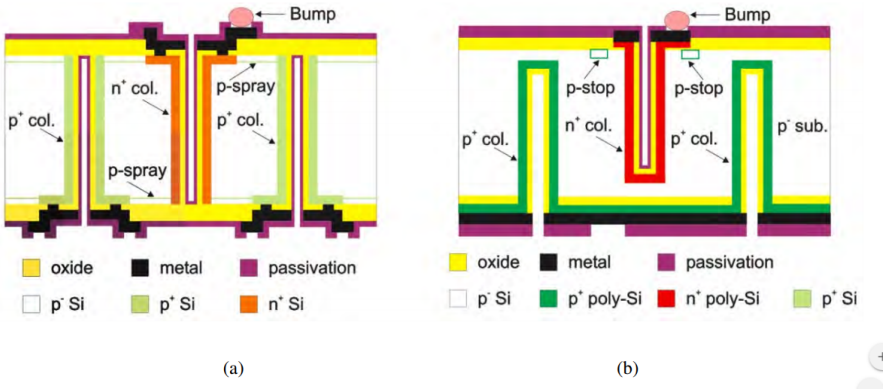


Figure 2.24: Design of the columns of (a) FBK and (b) CNM 3D sensors. [27]

The main difference between the two designs, showed in Fig. 2.24, is that in the FBK sensors the electrodes reach the other side (pass-through) while in the CNM they are stopped $\sim 20\mu m$ before reaching the other side. Tab. 2.6 summarizes the properties of IBL sensors.

2.3.2 Front-end electronics: FE-I4

There are two main reasons that led to the redesign of the FE-I3, the previous generation front-end chip used in the rest of the Pixel Detector. The first reason is that the new chip had to be installed very close to the beam pipe and hence it must

Table 2.6: Summary of the main design specifications for the planar and 3D sensors of the IBL detector. [27]

Parameter	Planar	3D FBK	3D CNM
Tile dimension [μm^2]	41315×18585	20450×18745	20450×18745
Sensor thickness [μm^2]	200	230	230
Sensor resistivity [$k\Omega cm$]	2-5	10-30	10-30
Pixel size (normal) [μm^2]	250×50	250×50	250×50
Pixel size (edge) [μm^2]	500×50	250×50	250×50
Pixel size (middle) [μm^2]	450×50	-	-
Nominal operating bias voltage V (non-irradiated/ $5\times10^{15} n_e q/cm^2$)	-80/-1000	-20/-160	-20/-160
Maximum operational power [mW cm^{-2}] ($5\times10^{15} n_e q/cm^2$)	90	15	15

provide a high tolerance to radiation. Secondly, the FE-I3 column-drain architecture scales badly with high hit rates, causing high amount of inefficiencies for IBL, as shown in Fig. 2.25.

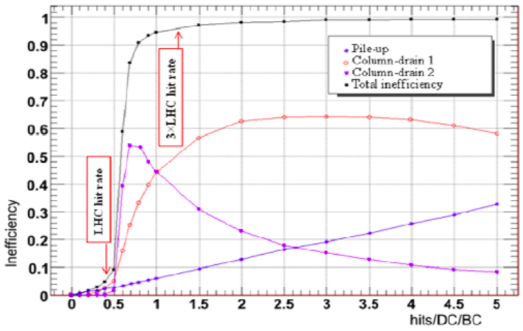


Figure 2.25: Inefficiencies for a FE-I4 using a FE-I3 like column-drain architecture in a 3.7 cm radius layer, given as a function of the number of hits per Double-Column and Bunch-Crossing. Pile-up inefficiency comes from lost hits due to an already busy analog pixel chain, whereas column-drain 1 inefficiency comes from lost hits due to busy digital pixel. [28]

A new front-end chip, called FE-I4, was hence fabricated in 2011 to fully meet the IBL requirements. It is designed using a 130 nm CMOS architecture and contains readout circuitry for 26880 pixels arranged in 80 columns by 336 rows. Its total active size is 20 mm by 16.8 mm with 2 mm occupied by periphery circuitry, leading to a 90% active area. Differently from FE-I3, the FE-I4 is a standalone unit and does not require a Module Controller Chip, since all the input-output communication is implemented in the peripheral part of the chip. It is divided in an analog and digital

part, with particular attention to separate the two power nets. A diagram of the FE-I4 is shown in Fig. 2.26.

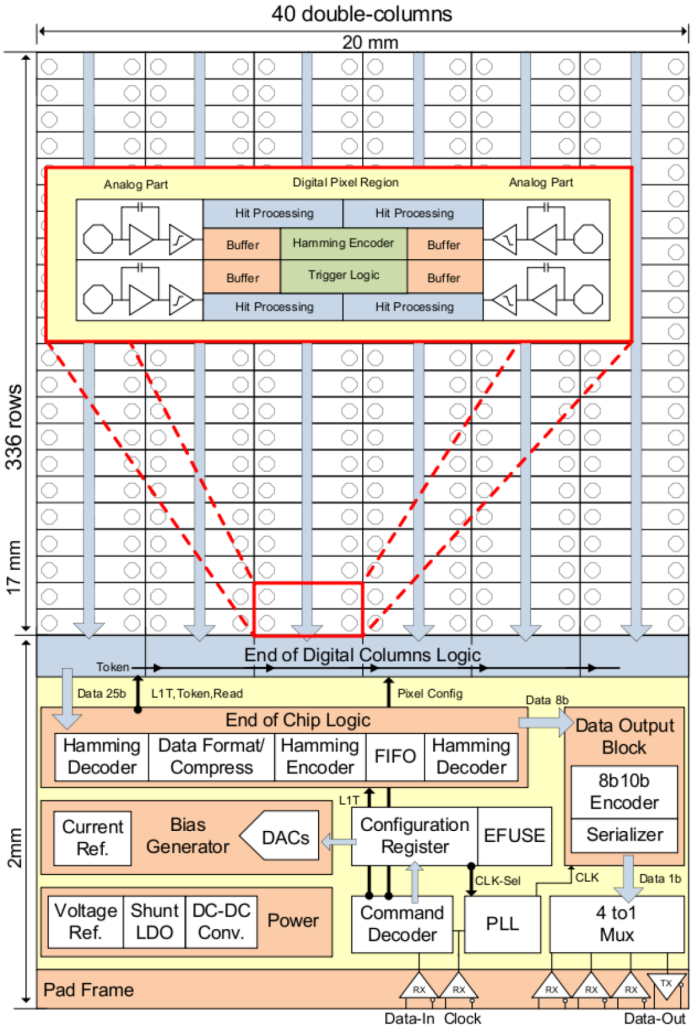


Figure 2.26: FE-I4 chip diagram, not to scale. [29]

Analog Pixel Section

The analog pixel section of the FE-I4, whose diagram is shown in Fig 2.27, is composed of a 2-stage amplifying architecture followed by a discriminator. It has been optimized for low power, low noise and fast rise-time. The first stage is a regulated cascode preamplifier input containing an active slow differential pair, tying the preamplifier input to its output. This configuration compensates sensor radiation-related leakage current, allowing a DC current tolerance of above 100 nA [27]. The second stage is a folded cascode amplifier AC coupled to the preamplifier. The AC coupling is beneficial and has two main advantages: it decouples the second stage

from DC potential shifts caused by leakage current and it gives an additional gain factor of about 6 to the signal. The intensity of the bias current, the feedback currents, the discriminator threshold and many other parameters are configurable via 20 global registers, and there are 13 configuration bits for local adjustment.

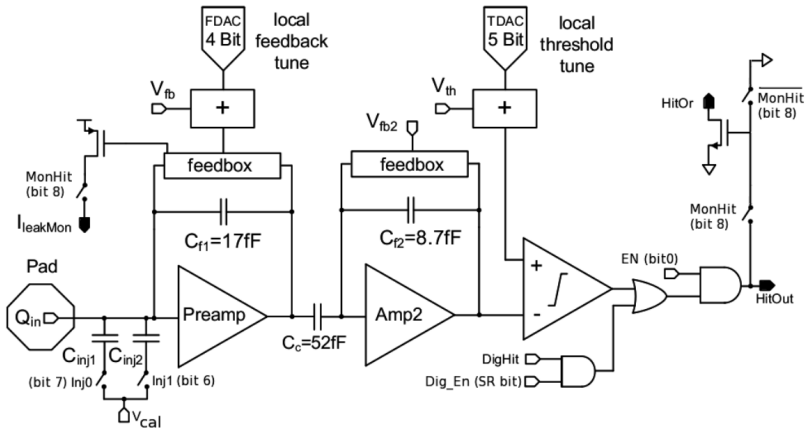


Figure 2.27: FE-I4 analog pixel schematic diagram. Output pins are solid, input pins are open. [29]

The Digital Region

In the previous sections it was discussed that the FE-I3 column-drained architecture would cause a high amount of inefficiencies if applied to IBL. To overcome this issue, the front-end structure was redesigned to store hits in local buffers located at pixel level. The base unit for hit recording and storage inside the Double-Column is the *Pixel Digital Region* (PDR) processing data from 4 discriminators (2 rows by 2 columns) as shown in Fig. 2.28. Since four pixels are tied together in a PDR, most of the digital processing, such as latency counters, trigger management units and read and memory management, is shared by the 4 pixels. When a pixel is *fired*, a common latency counters starts, providing a time-stamp for all the four pixels. However, each pixel has is own hit processing circuitry providing Time Over Threshold (TOT) information and discriminating the hit into small hit/big hit categories. This structure offers several advantages; there are no more time loss due to a triggerless hit propagation time along the column. Also, in case of multiple hits in a four-pixel structure within a bunch crossing, a single latency counter is started, reducing digital activity and improving the efficiency of the architecture. Another consequence is the reduction of time-walk effects, since the clustered nature of the LHC collisions is such that small hits are often located close to big hits and can be recovered without being time-stamped. To ulteriorly take advantage of this behavior, the smaller hits are associated to bigger hits within the same region or in the immediate next region, using a mechanism called *neighbor logic*.

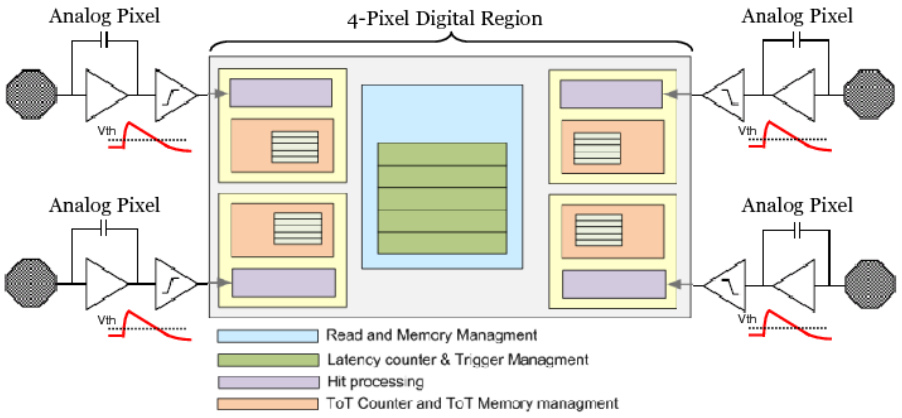


Figure 2.28: The FE-I4 four-pixel regional digital logic. [30]

When triggered hits are available in a Double-Column, data are passed to the periphery of the chip towards a dual token passing scheme (Double-Column / End of Column tokens). The End of Column logic is a very simple interface between each of the 40 Double-Columns and the digital control block with its FIFO. The chip radiation hardness is achieved by using triple redundancy with majority voting and by using configuration memory with hard custom cells latches with a Dual Interlocked Cell (DICE) architecture [31]. Also, inside the Double-Column, the data and the thermal-encoded region addresses are propagated down Hamming coded until reaching the End of Column logic.

Peripheral Logic

The peripheral logic of the FE-I4 front-end chip is in charge of communication and operational tuning of the IC. Particular care has been taken in the organization of the data read back and fast data output serialization. Input communication happens through to LVDS signals: one is the clock (40 MHz, Bunch Crossing frequency of LHC) and the other is dedicated to commands (40 Mb/s). FE-I4 commands comprehend local pixel configuration (13 bits per pixel), global register configuration and trigger commands, mirroring the architecture of the MCC chips. When a trigger is decoded, it is propagated to the End of Column logic and from there to the pixels. In case of coincidence of the trigger with a latency counter inside a 4-PDR, data stored in that PDR ToT buffers are sent to the periphery and associated to the Bunch-Crossing (BC) corresponding to that specific trigger. In the periphery data are then re-formatted and stored in a FIFO to be sent out. The FIFO also includes read back informations from the pixels and global registers and diagnostic information, such as error messages. A final block in the chip, called *Data Output Block* encodes the data in 8b10b protocol and serializes them at 160 Mb/s. The usage of the 8b10b protocol has two main advantages, since it offers protection against bit-flips and it is DC-balanced.

2.3.3 Modules, staves and geometry

The connection of a sensor and a FE-I4 chip, plus a *flex hybrid* - a double-sided, flexible printed circuit that allows connection to external services - constitute a *module*, the basic building block of the IBL detector. According to the sensor used, there are two types of module: planar modules, where a planar sensor is connected to two FE-I4 chips, and 3D modules, where a 3D sensor is connected to a FE-I4. The support structure, holding together 20 modules, electrical services and a cooling pipe il called *stave*. The modules in a single stave are not identical; as shown in Fig. 2.29(a), 12 planar modules are placed at the center of it, while 4 3D modules are at the side, to minimize the inefficiencies of perpendicular incident particles. The entire IBL detector is composed of 14 staves covering the full azimuthal angle and a pseudorapidity $|\eta| < 3$. To minimize the dead space between the staves, they are tilted by 14° so that there is a 1.82° overlap in the azimuthal direction (Fig. 2.29(b)). A final comparison between the Pixel Detector and IBL is given in Table 2.7.

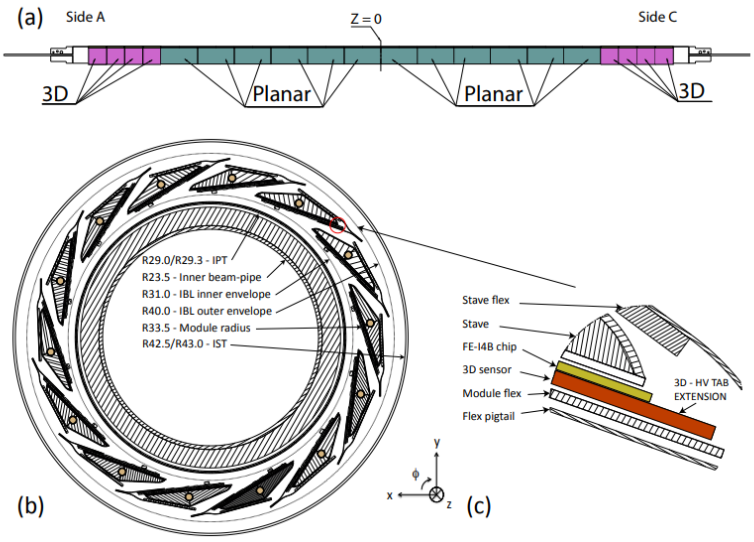


Figure 2.29: IBL detector layout. (a) Longitudinal layout of planar and 3D modules on a stave. (b) An $r-\phi$ section showing the beam pipe, the Inner Positioning Tube (IPT), the staves of the IBL detector and the Inner Support Tube (IST), as viewed from the C-side. (c) An expanded $r-\phi$ view of the corner of a 3D module fixed to the stave. [27]

2.3.4 Problematics of IBL within ATLAS

The production, test and installation of IBL within the ATLAS detector required several years; the assembled modules where firstly tested inside test facilities created in on-surface laboratories. The *Quality Assurance* (QA) procedure consisted in a measurement of the operational parameters (voltage, current, temperature) and in a series of scans (Digital, Analog, ToT scans). Before October 2014 IBL was assembled

Table 2.7: Comparison of the main characteristics of the Pixel and IBL detectors. [27]

Technical Characteristic	Pixel	IBL
Active surface (m^3)	1.73	0.15
Number of channels ($\times 10^6$)	1.73	0.15
Pixel size (μm^2)	50×400	50×250
Pixel array (columns×rows)	160×18	336×80
Front-end chip size (mm^2)	7.6×10.8	20.2×19.0
Active surface fraction (%)	74	89
Analog current (μA/pixel)	26	10
Digital current (μA/pixel)	17	10
Analog voltage (V)	1.6	1.4
Digital voltage (V)	2.0	1.2
Data out transmission (Mbit/s)	40-160	160
Sensor type	planar	planar / 3D
Sensor thickness (μm)	250	200 / 230
Layer thickness ($\%X_0$)	2.8	1.88
Cooling fluid	C_3F_8	CO_2

and installed in the ATLAS cave and the same series of tests was repeated, showing no difference with the on-surface results and proving that the transportation and installation was completely damage-free. However, after the integration with the ATLAS system, some unexpected problems appeared; after some intervention all the problem effects were mitigated and the data quality of IBL was not affected.

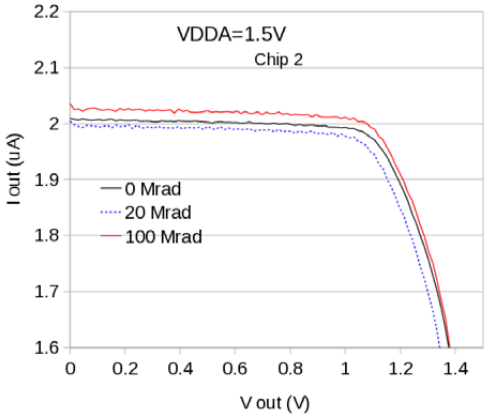
Bond oscillations

The IBL pixel sensors and the FE-I4 front-end chips are interconnected together via bump-bonding. The interconnected sensor/front-end are then wire-bonded to the flex-hybrid module; the bondings were not encapsulated so the wires are free to oscillate when subjected to external forces. Since IBL operates in a 2 T magnetic field, an electric current circulating into the wires would create oscillations, damaging the detector. To avoid antenna effects inside the bondings, which are maximized in case of a constant pulse frequency, a *Fixed Frequency Trigger Veto* (FFTV) was introduced at the DAQ level in the range 3-40 kHz.

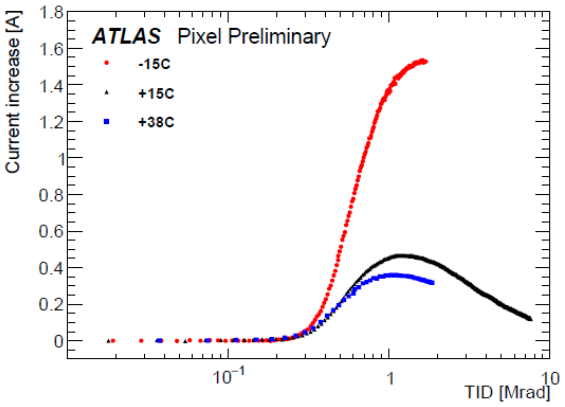
Current consumption

As discussed in Section 2.3.2, the FE-I4 front-end chips used by IBL were designed to be resistant to radiation damage. To certify the radiation tolerance, the chips were subjected to an irradiation test at the PS accelerator at CERN, using a 24 GeV proton beam. The test showed that the TID effects on the front-end operational parameters (current, voltage) were very limited, shifting the current value of $\sim 2\%$ and the single pixel noise of $\sim 10\%$, as shown in Fig. 2.30 (a). After IBL was constructed

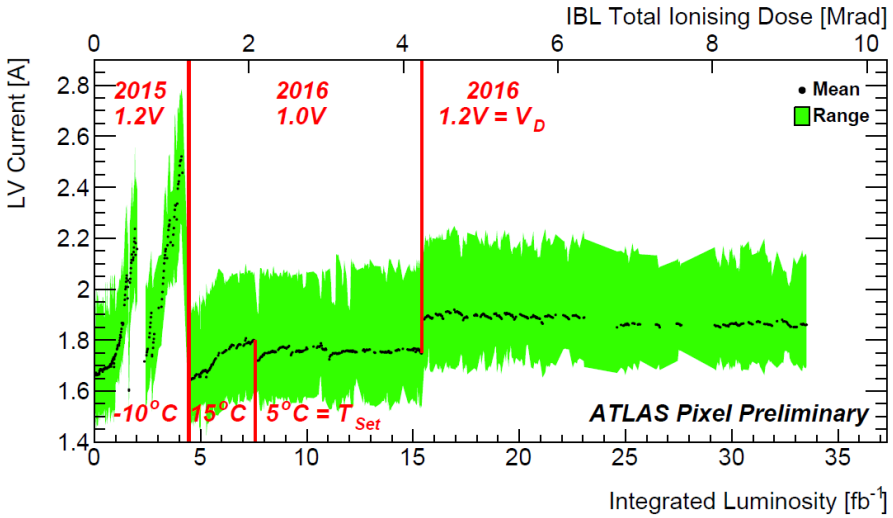
and integrated within the ATLAS experiment, the FE-I4 current consumption during stable beams started to increase dramatically (Fig. 2.30 (b)). This effect was



(a) Irradiation test



(b) Data within ATLAS



(c) Data within ATLAS

Figure 2.30: (a) Mean Low Voltage (LV) current consumption of the FE-I4B front-end chip in function of voltage. The measure was repeated at different Total Ionizing Doses (TID) after irradiation with the CERN-PS 24 GeV proton beam. Results show a 2% current variation and a 10% increase of the single pixel noise. (b) LV current consumption of the FE-I4B front-end chip in function of TID during ATLAS operations. The current consumption increase was measured at different detector temperatures and didn't reflect test result, proving to be significantly higher. (c) LV current in IBL FE-I4 chips during stable beam against integrated luminosity and Total Ionising Dose.

identified to result from N-MOS transistor leakage currents after the build-up of charge at the SiO_2 interface in the 130 nm CMOS process [27]. By running the detector at different temperatures and providing different voltage, the effects of increasing current consumption was drastically reduced, due to the fact that annealing

procedures were introduced.

IBL distortion

During stable beam operations, the IBL data-quality community noticed a distortion of the IBL staves. The distortion, consisting in the twisting of the staves of few μm/°C on the R-ϕ plane, were caused mainly by two effects:

- the stave and the stave flex have a different thermal expansion coefficient;
- the stave flexes are attached asymmetrically, due to mechanical constraints.

This problem was minimized by operating the detector at a very stable temperature (less than 0.2°C temperature stability) and by developing in-time alignment correction procedures.

2.4 Pixel Detector performance and results

This chapter introduced to the design of the ATLAS Pixel Detector, focusing on its most important parameters and the motivations that led to the introduction of IBL as a fourth pixel layer. This section shows some of the results obtained by the Pixel Detector and IBL.

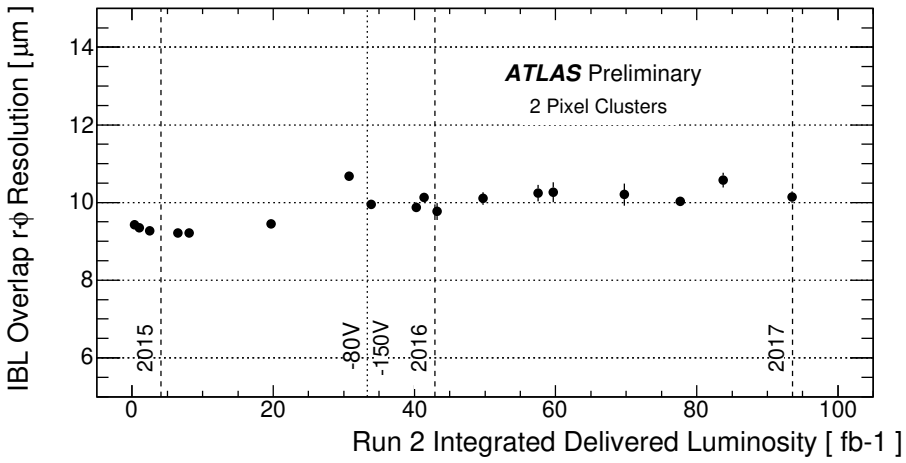


Figure 2.31: Spatial resolution of IBL hits associated to reconstructed particle tracks in 13 TeV collision di-jet events as a function of the integrated luminosity in Run 2. The IBL spatial resolution is determined by the corrected transverse positions of the two reconstructed IBL clusters associated to a charged particle track in the regions where the IBL modules overlap, using the technique described in ATL-INDET-PUB-2016-001. Only clusters with two pixels in the transverse coordinate are considered here. The slight worsening of the spatial resolution observed over the three years is correlated with the reduction of charge collection efficiency as a result of radiation damage. [PIX-2018-002]

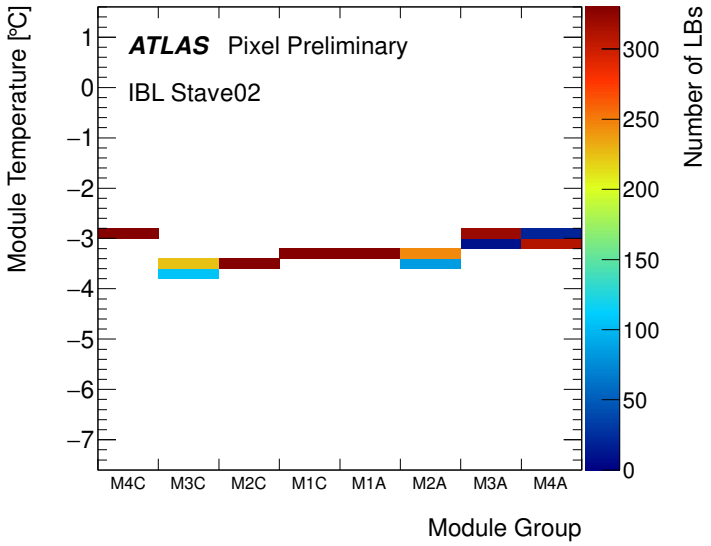


Figure 2.32: Stability of temperature of IBL modules of one run (270953) for a stably operated stave. The colour scale shows the number of luminosity blocks (atomic unit of ATLAS data of the run) which have the temperature record in the corresponding temperature bin. Note that one temperature sensor monitors every four front-end chips (eight sensors in total for a stave). [PIX-2015-005]

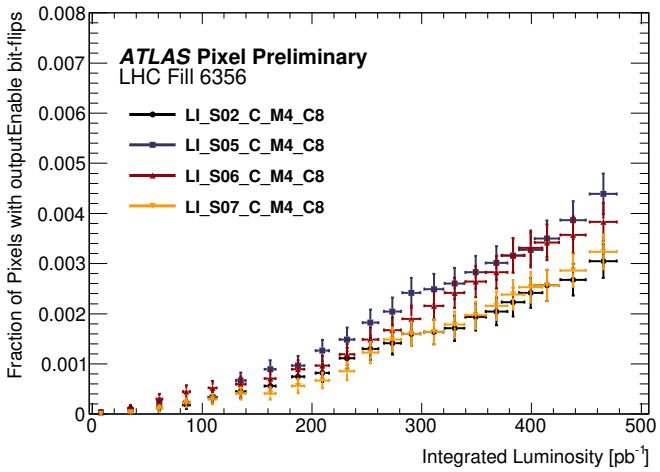


Figure 2.33: Fraction of number of pixels, initially disabled at the beginning of the LHC Fill 6356, that fire when traversed by a charged particle, shown for four different FE-I4 front ends placed on various IBL 3D modules as a function of the integrated luminosity of the fill. The horizontal error bars are obtained by applying 2.4% systematic uncertainty on the integrated luminosity measurement, while the vertical bar is the statistical error. [PIX-2018-007]

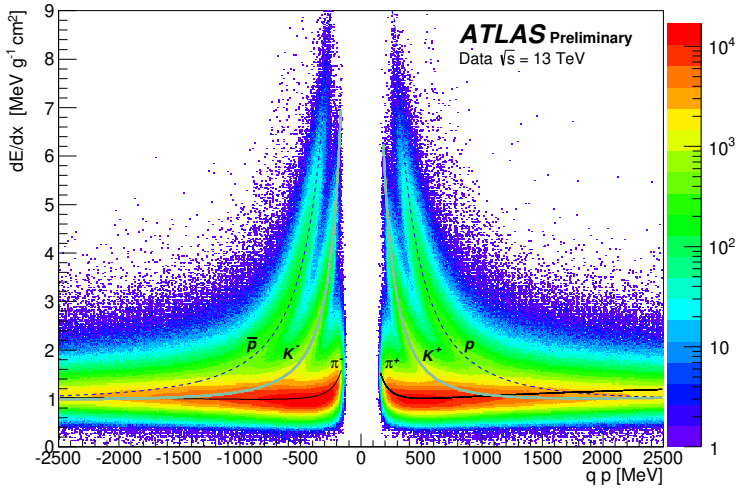


Figure 2.34: Bidimensional distribution of dE/dx and momentum for data. The distributions of the most probable value for the fitted probability density functions of pions, kaons and protons are superimposed for *numberOfUsedHitsdEdx* category equal to 2. [PIX-2015-002]

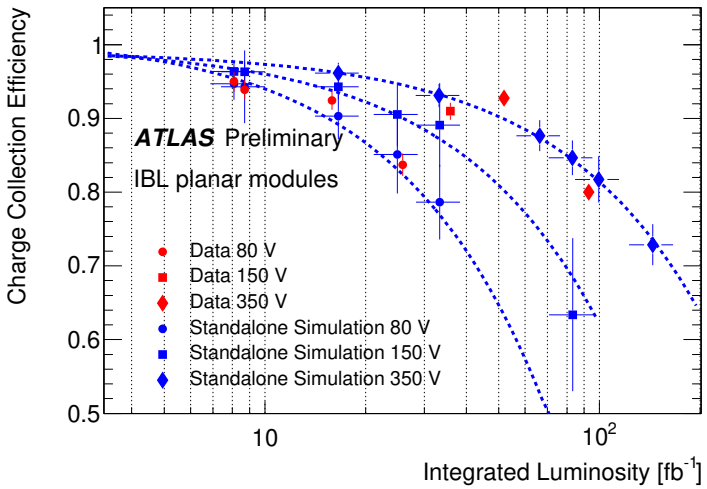


Figure 2.35: Run 2 Data and Standalone Simulation, based on Geant4, of the Charge Collection Efficiency relative to the zero fluence case as a function of the integrated Luminosity for IBL modules with $|\eta| < 0.2$. Simulation based on Geant4 inputs, Electric field maps produced with TCAD simulation based on the Chiochia radiation damage model [32], and accurate description of electrons and holes drift in the silicon detector. Data are corrected to account for ToT drift in the calibration process. [PIX-2017-004]

Chapter 3

ATLAS Pixel Detector Readout system

The ATLAS experiment is a very challenging environment in terms of readout, since there are hundreds of millions of readout channels that must be acquired at a very high frequency. Since the LHC bunch crossing has a period of 25 ns, ideally the ATLAS experiment should be able to retain all the information from all the sub-detectors at a 40 MHz frequency. However, this is impossible because the technology does not offer us yet the possibility to cope with such a bandwidth and the storage resources that would be necessary.

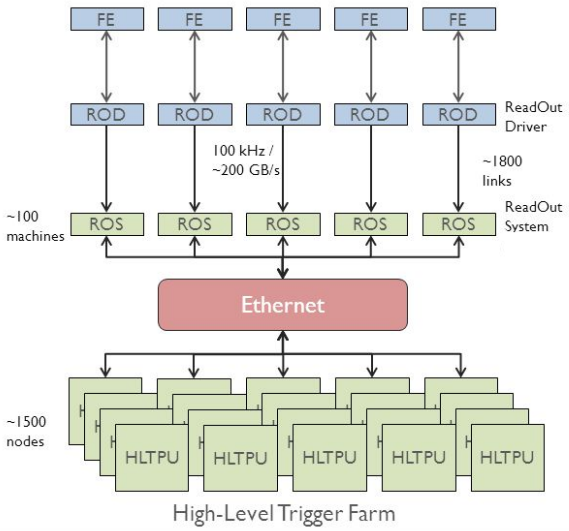


Figure 3.1: Block scheme of the Trigger and DAQ system. Custom electronics is colored in blue, while commercial electronics is green.

For this reason particular care must be put in the design of a readout system capable of carefully selecting the maximum amount of data possible, keeping the system aligned and efficiently merging all the sub-detectors fragments to fully build

the entire events. The system in charge of handling trigger and data propagation is called *Data AcQuisition/High Level Trigger* (DAQ/HLT) or TDAQ. In ATLAS the DAQ rate is reduced from 40 MHz to $\sim$ 500 Hz by a three-level trigger architecture and the data flow is distributed through several nodes.

Fig. 3.1 shows that all the parts composing the TDAQ chain can be divided in two main categories: detector-specific electronics and commercial electronics (mostly PCs). To interface a detector and its front-ends, custom boards called *Read-Out Drivers* (ROD) have been designed; each sub-detector has its own ROD, since the performance and the problematics involved are very specific. The connection point between the custom and the commercial boards used in the ATLAS readout chain is the *Read-Out System* (ROS), after which data are propagated via Ethernet to the HLT farms. My work mostly concerned the development and integration of the ROD board initially built to interface IBL. This board, called *IBL-ROD*, was successively also used to replace the previous readout system - that from this point on will be called *old readout* - of the entire Pixel Detector.

In this chapter a brief overview of the ATLAS readout system architecture will be presented, focusing on the Pixel Detector readout chain, its main components and the reasons that led to its upgrade that happened during 2015-2018.

3.1 ATLAS readout system overview

The ATLAS TDAQ system is in charge of reading-out all the front-end electronics and to mass-storing data into the CERN computing farms. It is implemented by using both custom electronic boards and commercial components, such as computer, Ethernet switches, etc. The two main goals of the DAQ/HTL are data collection an trigger reduction which are greatly interconnected and depend on each other. The trigger system is implemented on three levels, the *Level-1* (L1) - that is implemented in hardware - the *Level-2* (L2) and the *Event Filter* (EF). The L2 and EF together constitute the *High Level Trigger* that is implemented on software running on server computers. The L1 trigger selects events at a maximum design frequency of 100 kHz; after an event has been accepted by the L1 trigger algorithms, it is moved from the front-end electronics to the *Read-Out Driver* (ROD), a detector-custom piece of hardware that packages the events and redirects them to a common Read-Out System (ROS). Up to this point, the transmission is implemented either via electrical or custom optical protocols, such as the S-Link [33].

The ROS receives input data via the *Read-Out Links* (ROLs), that cross the boundary between sub-detector specific readout electronics and common readout system. The ROLs are physically implemented on custom PCI cards residing in the ROS PCs, called *Read Out Buffer IN* (ROBIN). Tab. 3.1 summarizes the total amount of ROS and ROLs used for all the ATLAS sub-detectors.

From the ROS PCs data are accessible via a dedicated network, called *Data-Collection Network* and are distributed to the L2 trigger system and to the *Event Builder* (EB). Only a subset of data ($\sim$ 5% of the L1 accepted rate) is processed

Table 3.1: Numbers of ROLs and ROS PCs per detector TTC partition. [34]

TTC Partition			Number of ROLs	Number of ROS PCs
Inner detector	Pixel	Layer 0	44	4
		Disks	24	2
		Layers 1-2	64	6
	SCT	End-cap A	23	2
		End-cap C	23	2
		Barrel A	22	2
		Barrel C	22	2
	TRT	End-cap A	64	6
		End-cap C	64	6
		Barrel A	32	3
		Barrel C	32	3
Calorimetry	LAr	EM barrel A	224	20
		EM barrel C	224	20
		EM end-cap A	138	12
		EM end-cap C	138	12
		HEC	24	2
		FCal	14	2
	Tile	Barrel A	16	2
		Barrel C	17	2
		Ext. barrel A	16	2
		Ext. barrel C	16	2
Muon Spectrometer	MDT	Barrel A	50	4
		Barrel C	50	4
		End-cap A	52	4
		End-cap C	52	4
	CSC	End-Cap A	8	1
		End-Cap C	8	1
L1	Calorimeter	CP	12	2
		JEP	10	2
		PP	32	3
	Muon RPC	Barrel A	16	2
		Barrel C	16	2
	Muon TGC	End-Cap A	12	1
		End-Cap C	12	1
	MUCTPI		1	1
	CTP		1	1
Forward Detectors	BCM		3	1
	LUCID		1	1
	ALFA		2	1
	ZDC		4	1
		Total	1583	151

by L2; this pre-selection happens by defining *Regions of Interest* (ROI) among all the L1 events. The ROI are regions where the L1 trigger has identified clusters or tracks and are collected by the *Region of Interest Builder* (RoIB) to be used by the L2 system, whose algorithm selects the events reducing the rate to $\sim 5\,kHz$. The events that passed L2 selection are then built by the Event Builder; finally, the last filtering process - the Event Filter implemented off-line on software - further reduces the rate to the storage rate ($\sim 500\,Hz$). Fig. 3.2 illustrates the whole TDAQ system. A complete and detailed description of all the components of the ATLAS TDAQ chain is beyond the purpose of this thesis, and next sections will focus on the custom off-detector electronics used in the Pixel Detector.

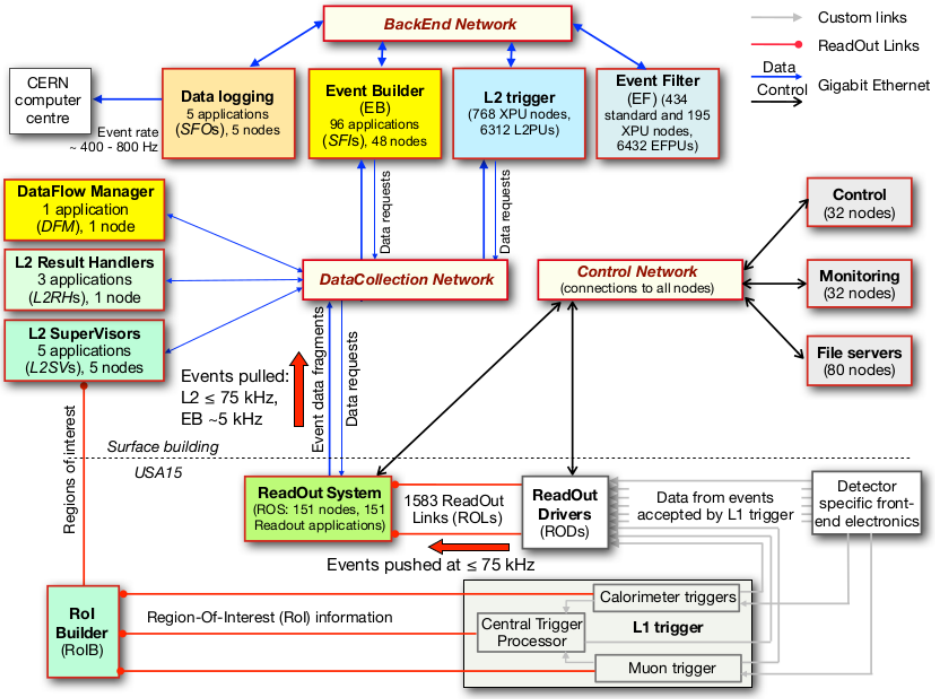


Figure 3.2: Block scheme of the Trigger and DAQ system. XPU nodes are nodes that can be used either for the L2 trigger or for Event Filter processing, L2PUs and EFPUs are applications executing the L2 and EF trigger algorithms respectively. For clarity only a few of the Control Network connections are shown. [34]

3.2 The *old* ATLAS Pixel Detector Readout chain

The Pixel Detector original readout system operated from the beginning of ATLAS Run-I and during Run-II was gradually replaced by an upgraded one; for this reason it will be referred as *old readout*. The motivations that led to the readout upgrade will be discussed later in this chapter. This section will only discuss the custom-made readout components interfacing the pixel detector to the ROS system. The whole

chain can be divided in two main groups: on-detector and off-detector electronics. The on-detector electronics comprises the FE-I3 and MCC front-end chips and the opto-board, already extensively discussed in Chapter 2. The off-detector electronics comprises several components:

- *Silicon ROD* (SiROD), the custom Pixel Detector ROD, a *Versa Module Europa* (VME) board in charge of data processing, histogramming and front-end configuration;
- *Back Of Crate* (BOC) board responsible of input-out communications;
- *Timing and Trigger Control Interface Module* (TIM), which interfaces ATLAS Level-1 Trigger system signals to Pixel sub-detectors;
- *Single-Board Computer* (SBC), a VME-controller computer residing in the VME crate.

The SiROD and the BOC board are placed back-to-back in a VME crate and are connected together by the backplane signals inside the crate. One crate can hold up to 16 SiROD-BOC pairs; 44 pairs (three crates) were necessary for B-Layer, 66 for Layer-1/2 (four crates) and 24 (two crates) for disks. One slot of each crate was occupied by the TIM board - interfacing the L1 trigger and distributing it to all the SiRODs - and another slot was used for the SBC card, which communicates to the SiRODs via the VME bus. During normal data acquisition, the front-end data, after being processed by the SiRODs, are sent to the ROS servers via the BOC S-Links; after this stage they are either discarded or proceed to full event building at the L2 trigger rate. During detector calibration, on the other hand, the data-path is different and the ROS is not involved. A calibration procedure consists in a sequence of injections of a known charge into pixel amplifiers. Each pixel response is then monitored and changes following front-end configuration parameters. The final goal of a calibration scan is to tune the front-end parameters (e.g. digital threshold, fine delay, preamplifier currents, etc.) in order to obtain the best achievable configuration of the detector. To reach maximum precision and to minimize the calibration duration, the front-end data must be extracted at the maximum speed supported by the front-ends and the ROS, which is designed to extract data at L2 trigger rate, is not the best solution. For this reason calibration data are stored by the SiROD in an internal memory buffer and are extracted by the SBC computer via the VME bus, from where they are sent to other servers for further analysis. The VME bus is also used by the DAQ software, running on the SBC, to configure the front-end modules.

3.2.1 Pixel BOC

The Back-of-Crate (BOC) board has two main functions: it executes all the input-output operations between the ROD and the front-end electronics and it is in charge

of distributing the timing to the system. Since the front-end and off-detector electronics are placed $\sim 100\,\text{m}$ apart from each other, all the communication lines are implemented through optical fibers. The connection to the optical fibers, as well as the opto-electrical conversion, is located in two plug-in cards: the *TX plug-in*, for transmission of optical data, and the *RX plug-in* for optical data reception. The TX plug-in has an 8-channel Vertical-Cavity Surface-Emitting Laser (VCSEL) array and a *Bi-Phase Mark* (BPM-12), an ASIC used to adjust the phase between the 40 MHz BOC clock and each detector module bunch-crossing. The RX plug-in has an 8-channel PiN diode array and a *DRX*, an ASIC that amplifies, discriminates and converts the signal from the PiN array to LVDS.

3.2.2 Silicon ROD

The Silicon ROD (or SiROD), shown in Fig. 3.3 is a 9U VME board and is the main building block of the old ATLAS Pixel Detector readout chain. The SiROD handles data coming from the front-ends and handles the data-transfer towards the ROS system or the calibration farm.

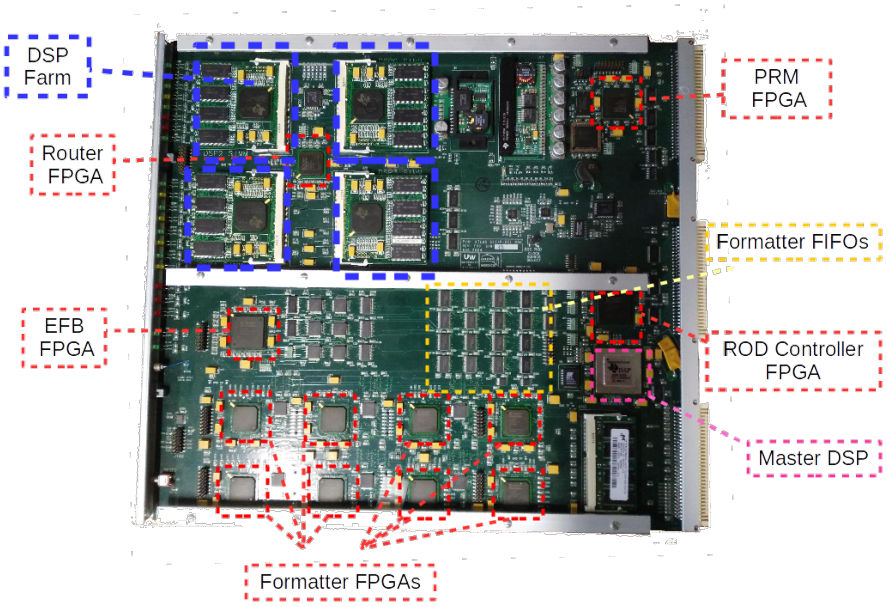


Figure 3.3: Silicon ROD picture with main components highlighted.

The design of the SiROD can be divided in three main sections: *control path*, *data-flow path* and *Digital Signal Processing* (DSP) farm. The control path block performs several actions:

- programs and resets the SiROD board;

- interfaces the SBC via VME, receiving and processing commands;
- receives module configuration via VME and transmits it to the modules;
- controls calibration procedures;
- propagates the triggers from the TIM to the modules.

Physically, the control path section is implemented on a Texas Instrument DSP (TI320C6201), operating at 160 MHz with a 32 MB SDRAM module, and on two Xilinx Spartan-3 [35] Field Programmable Gate Arrays (FPGA), a *Program Reset Manager* (PRM) FPGA and a *ROD Controller FPGA*. The PRM FPGA acts as a VME slave controller, allowing read and write access to all SiROD and BOC registers. It also allows VME access to the flash memory chip that contains the FPGA configuration files, to allow the users to easily upgrade the firmware on the board. The Master DSP receives commands and transmits replies to the VME host and coordinates the configuration, calibration and data-taking modes of the ROD. The ROD Controller FPGA is used in the control path as an interface for the Master DSP to the DSP farm, the BOC, and all of the internal ROD registers in the data flow FPGAs. It also controls all of the required functions on the ROD, including serial transmission of commands to the FE modules, calibration mode trigger generation, and transmission of TIM generated triggers and fast commands.

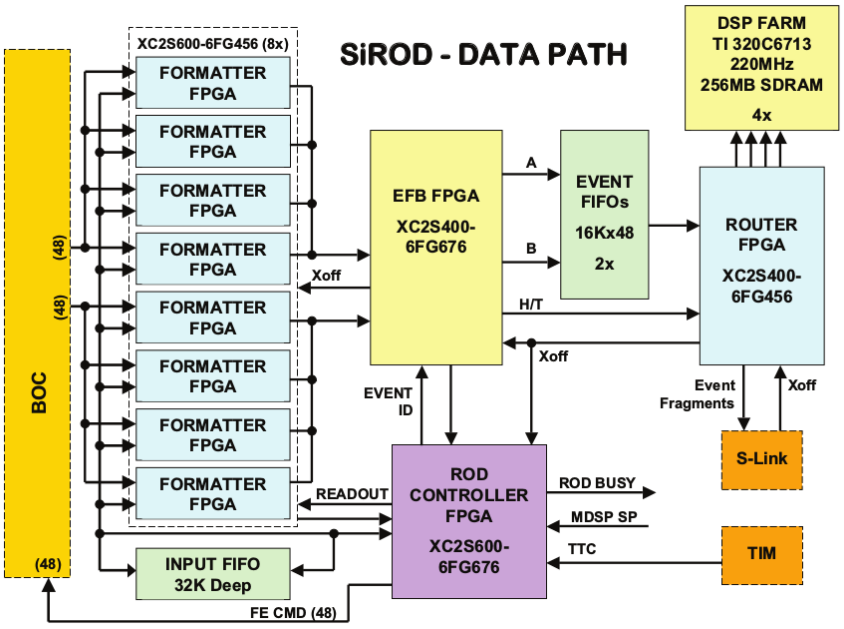


Figure 3.4: Data routing inside the Silicon ROD. All the main nodes and components are shown in the figure. [34]

The data-flow path, shown in Fig. 3.4, is implemented through several Spartan-3 FPGAs. Eight *Formatter* FPGAs receive data from the front-ends and deserialize

them; a *Event Fragment Builder* (EFB) FPGA collects the data from the eight formatters and build a ROD event, adding trigger information received from the TIM and finally a *Router* FPGA directs the output to the S-Link during data-taking or to the DSP farm during calibration.

The DSP farm is composed of four *Slave* DSP processors (TMS320C6713) connected to the Router FPGA. During normal data acquisition the DSP farm performs system monitoring, computing average occupancy to detect noisy pixels or errors; during detector calibration the DSP farm accumulates histograms from the various scans and performs data-compression operations. The histograms are temporally stored into the DSP memories, to be transferred to the SBC in a second time.

3.2.3 TIM

The TTC (Timing, Trigger and Control) Interface Module *TIM* (Figure 3.5) interfaces the ATLAS Level-1 Trigger system signals to the Pixel Read-Out Drivers using the LHC-standard TTC and Busy system. The TIM performs the following tasks:

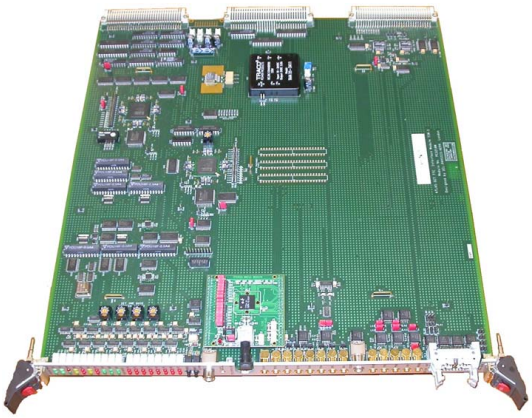


Figure 3.5: Photograph of a Timing, Trigger and Control Interface Module (TIM) card.

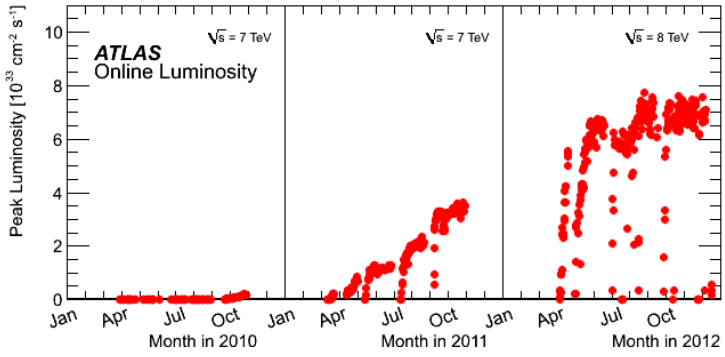
- propagates the TTC clock all over the experiment;
- propagates the triggers all over the experiment;
- keeps *Bunch and Event Counters* synchronized with main TTC;
- propagates the previous informations to the ROD system.

There are two Xilinx Spartan-2 FPGAs [36] on the TIM; one is the *Board Manager* supporting the more generic board functions (VME Interface, local bus control, board reset) and providing status information on the other FPGA. The second one hosts all the TIM specific functions and provides interfaces to front-panel and ROD backplane signals.

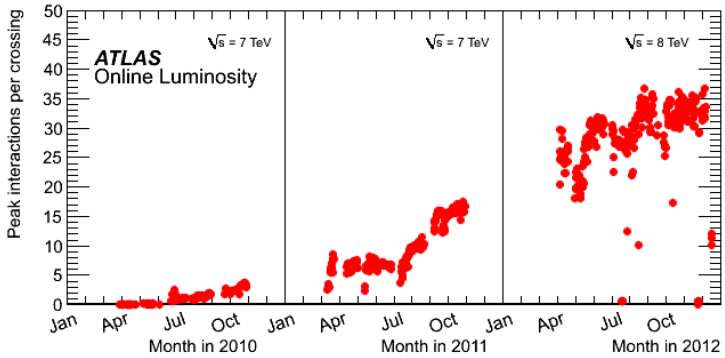
3.3 Effects of luminosity on the Pixel Detector readout

The original ATLAS Pixel Detector readout was designed and optimized to fulfill all the requirements needed for Run-I, when LHC was running with 7 TeV center-mass energy and the accelerator luminosity was relatively low compared to the design one. After Long-Shutdown I, between 2013 and 2015, a new component of the pixel detector - IBL - was introduced providing new readout challenges and different technologies. Also, the beam energy was raised to 13 TeV and the instantaneous luminosity was increased till it passed the design value of $1 \cdot 10^{13} cm^{-2} s^{-1}$. As already stated, the instantaneous luminosity ($\mathcal{L}$) is defined as the amount of events (N) detected in a certain time (t) divided by the interaction cross section (σ) as described in Equation 3.1:

$$\mathcal{L} = \frac{1}{\sigma} \frac{dN}{dt} \tag{3.1}$$



(a)



(b)

Figure 3.6: The peak instantaneous luminosity (a) and maximum mean number of events per beam crossing (b) versus day during the p-p runs of 2010,2011 and 2012. The online luminosity measurement is used for this calculation as for the luminosity plots. Only the maximum value during stable beam periods is shown.

It depends exclusively on the collider parameters, such as the number (n_i) of particles in a bunch, the revolution frequency f of the bunches in the accelerator and the transversal dimensions of the beam σ_x and σ_y. The general formula for instantaneous luminosity is given by:

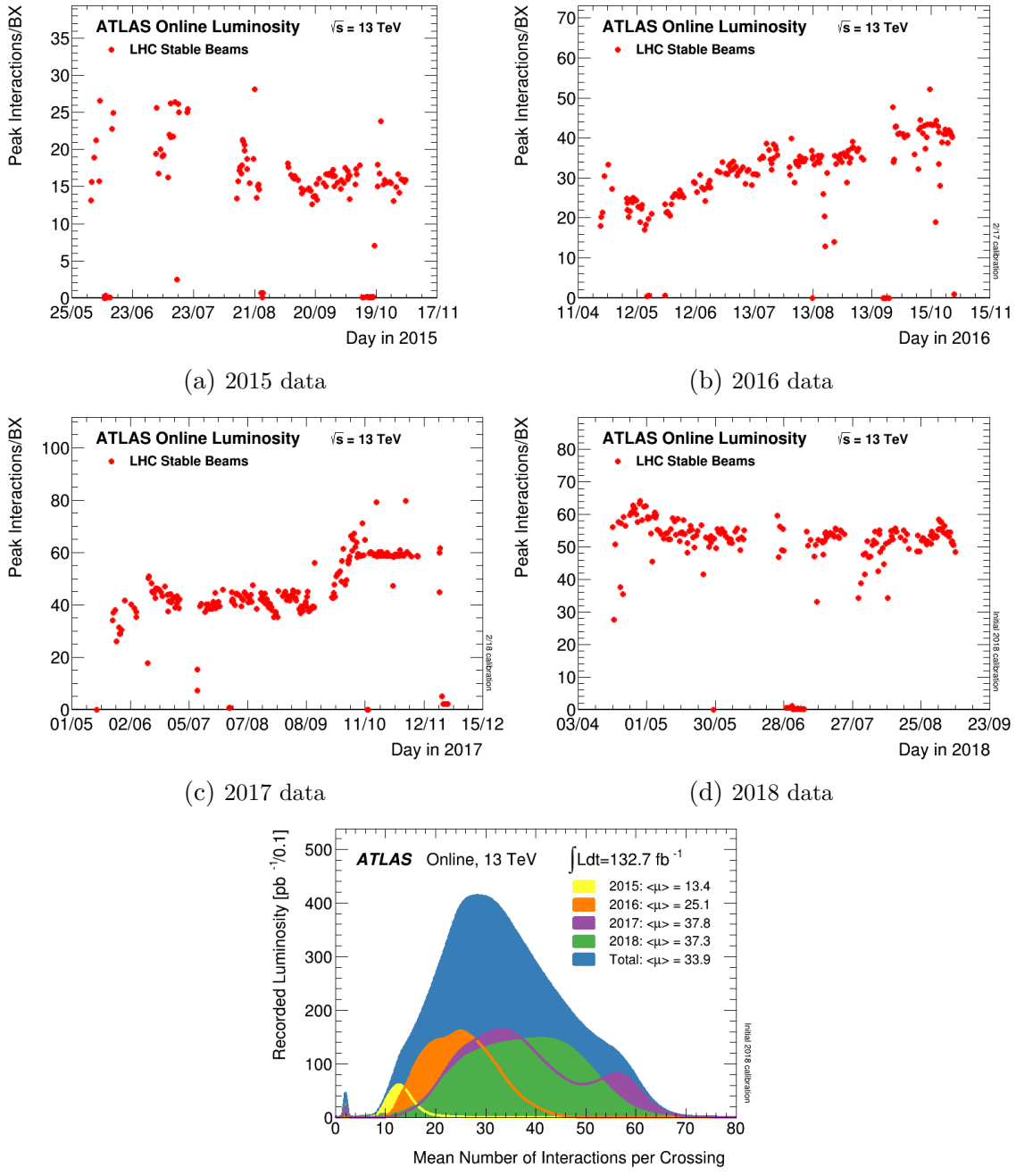
$$\mathcal{L} = f\frac{n_1 n_2}{4\pi\sigma_x\sigma_y}$$

Equation 3.1 shows that the number of collisions per time unit increases with the luminosity; the number of multiple collisions happening within the same bunch crossing window, called *pileup* μ, is strictly correlated to both the luminosity and the interaction cross section, which depends on the beam energy. Fig. 3.6 shows the luminosity trend during Run-I; the correlation between instantaneous luminosity and pileup can also be seen from the plots. The maximum and mean pileup measured during Run-II is shown in Fig. 3.7.

3.3.1 Old readout limitations

The effects of luminosity and pileup increment are particularly visible on the Pixel Detector, since it is the closest detector to the beam pipe. The electronic readout must be able to cope with the increased *data throughput*, the amount of data per time unit to be transmitted. The system must provide enough *bandwidth*, the maximum amount of data that can be transmitted per time unit, to favor the readout of all the events without any data-loss. Losing even only an event would lead to catastrophic consequences, since it would not only affect the immediate performance of the detector, but it could potentially cause the loss of system synchronization. A functioning readout system must hence guarantee that the *link occupancy*, defined as *data throughput/bandwidth* never saturates ($\geq$100%). In the original Pixel Detector readout design, optimized for a Run-I scenario, only part of the maximum bandwidth provided by MCC front-end chips was used by the BOC-SiROD couple. As described in Section 2.2.2, a MCC can run at three different readout speeds: 40 Mbps and 80 Mbps on one line or 160 Mbps on two lines. Only B-Layer, being the closest to the beam and hence the one with more hit density, was readout at the full speed provided by the module (160 Mbps). The Disks and Layer-1 were readout at 80 Mbps and Layer-2, being the least densely populated, was readout at 40 Mbps. Such a configuration was particularly efficient for a Run-I scenario, with a maximum pileup of $\mu \sim 50$, a maximum trigger rate of $\sim$ 75kHz and a bunch crossing separation of 50 ns. After Long Shutdown I, however, the LHC conditions became more critical and the system was not anymore capable of coping with the amount of data without any intervention. Tables 3.2 and 3.3 shows the extrapolated link occupancy for all the Pixel Detector layers at different bunch crossing spacing and pileup, respectively with a mean trigger frequency of 75 kHz and 100 kHz.

According to the two tables, the Pixel Detector layer in the most critical situation is the Layer-2, the one running at the lowest readout speed. In fact, even if the



(a) 2015 data

(b) 2016 data

(c) 2017 data

(d) 2018 data

(e) Mean pileup (all years)

Figure 3.7: The maximum number of inelastic collisions per beam crossing (μ) during stable beams for pp collisions at 13 TeV centre-of-mass energy is shown for each fill in 2015 (a), 2016 (b), 2017 (c) and 2018 (d). (e) Luminosity-weighted distribution of the mean number of interactions per crossing for the 2015-2018 pp collision data at 13 TeV centre-of-mass energy. The preliminary luminosity measurement is used to determine the number of interactions per beam crossing as $\mu = \mathcal{L}_{bunch} \times \sigma_{inel}/f_r$ where L_{bunch} is the per-bunch instantaneous luminosity, σ_{inel} is the inelastic cross-section at 13 TeV, which is taken to be 80 mb, and f_r is the LHC revolution frequency of 11.245 kHz. The number of interactions shown is averaged over all colliding bunch pairs, and only the peak value per fill during stable beams is shown. Data collected by ATLAS through 5 September 2018 (LHC fill 7125) are shown.

Table 3.2: Link occupancy for a L1-trigger rate of 75 kHz. The link occupancy is defined as the ratio between the detector data throughput and the bandwidth of the readout system. Orange values are limit values, meaning the readout system is operating at full-capacity. Red values shows scenarios in which the readout system bandwidth is too low to cope with the data throughput, leading to de-synchronization errors.

	μ	B-Layer	Layer 1	Layer 2	Disks
50 ns	37	39%	34%	52%	30%
25 ns	25	35%	31%	48%	27%
	51	53%	59%	66%	39%
	76	71%	73%	111%	64%

overall link occupancy was still far from reaching saturation, in 2012 (Run-I) the Pixel Detector Layer 2 was showing some problems due to the readout limitation. The overall effect was that data-losses inside the detector were causing the loss of synchronization with the rest of the system, creating the so-called *de-synchronization* errors, as shown in Fig. 3.8.

Table 3.3: Same as Table 3.2 but run at 100 kHz trigger rate.

	μ	B-Layer	Layer 1	Layer 2	Disks
50 ns	37	51%	45%	69%	40%
25 ns	25	47%	42%	65%	37%
	51	71%	67%	88%	52%
	76	95%	97%	148%	75%

3.4 The ATLAS IBL Readout

With the addition of IBL as a fourth layer of the Pixel Detector during Long Shutdown I, the ATLAS Experiment had to face the decision of adapting its current readout system to the new detector or moving to a new system, with new technologies and better performance. Run-I showed that the Pixel Detector readout had limited resources, a bad scalability when moving to higher energy, trigger rate and pileup, and was already showing de-synchronization problems for Layer 2 (Fig. 3.8). The readout system showed several incompatibilities for interfacing IBL: the components and FPGAs (Spartan 3) used were obsolete, the BOC-SiROD pair total bandwidth was limited (1.04 Gbps per pair), the calibration data were passing through VME bus, that could only support limited speed, and the overall control and system recovery capabilities were insufficient. With those reasons in consideration, the design of a new readout was inevitable. Many aspects of the new design were driven by the necessity to have a backward compatibility with the previous system, in order to leave untouched many parts of the chain while replacing only

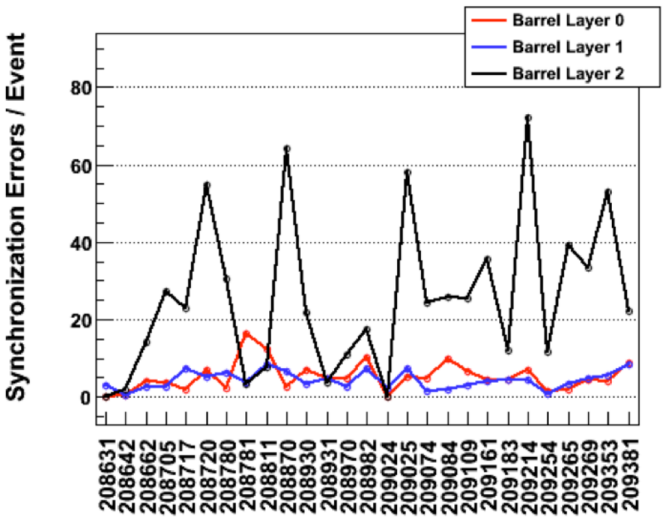


Figure 3.8: Number of desynchronized modules observed during 2012 (Run-I) in pp collisions with a center-of-mass energy of $\sqrt{s} = 8$ TeV measured by the ATLAS Pixel Detector. While B-Layer (red) and Layer-1 (blue) show stable performances during the year, Layer-2 exhibits an high number of de-synchronizations due to readout limitations.

the *obsolete* ones. The final decision was to maintain the same VME-crate system, keeping the same TIM card and SBC as the one discussed in Section 3.2, and to produce new pairs of ROD/BOC, specifically called *IBL-ROD* and *IBL-BOC* [37] [38]. The two cards feature more recent technology devices, such as Xilinx Spartan-6 and Virtex-5 FPGAs, and provide a total bandwidth of 5.12 Gbps divided on four S-Links. While the VME bus was maintained for retro-compatibility, its usage is restricted only to board programming and low-level register monitoring, and all the detector-calibration payload was moved to Gbit-Ethernet ports placed on the IBL-ROD. All communications with the ATLAS TDAQ software are mediated by an embedded processor inside the ROD board, a PowerPC (PPC), which provides the system a high level monitoring and recovery capability. Fig. 3.9 shows the new IBL-ROD and IBL-BOC, their interconnection and the input-output data-path with external devices.

3.4.1 IBL-BOC

The IBL Back-Of-Crate (IBL-BOC) [39], also called simply BOC, was designed as a replacement of the previous Pixel Detector BOC. The board, shown in Fig. 3.10, is responsible of handling all the input/output transmission from and to the detector and of receiving the ATLAS 40 MHz trigger from the TIM board and distributing it to the ROD and the modules. The IBL-BOC operations are handled by three Xilinx Spartan-6 FPGAs on the board, one *BOC Control FPGA* (BCF) and two *BOC Main FPGAs* (BMF). The optical interface to and form the detector is implemented on two equivalent data paths, each one consisting of two commercial SNAP12 receivers

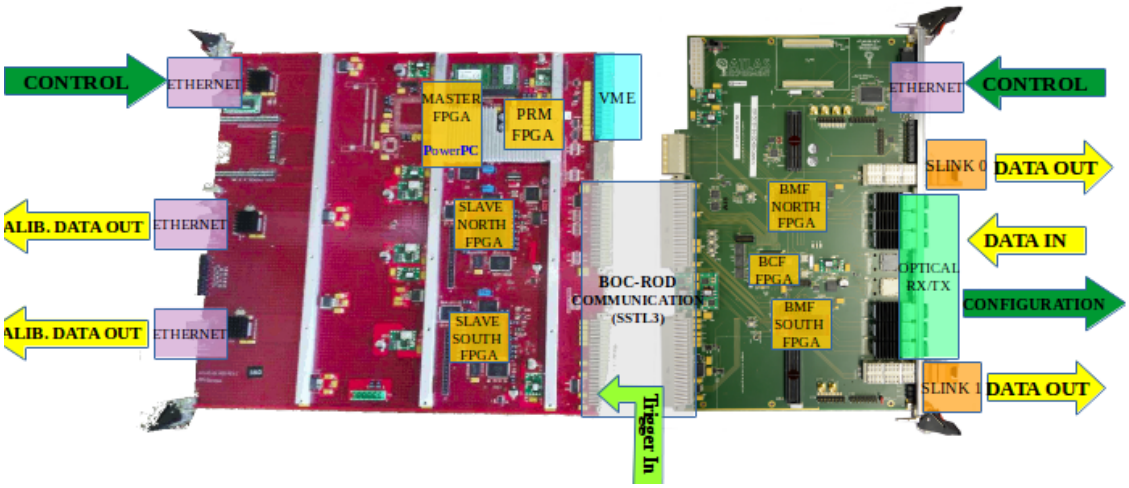


Figure 3.9: Part of the upgraded ATLAS Pixel Detector readout chain. The IBL-ROD board (on the left) features 3 Spartan-6 FPGAs (Program-Reset Manager (PRM), Slave North and Slave South) and one Virtex-5 FPGA (Master). The IBL-BOC board (on the right) features 3 Spartan-6 FPGAs (Board Controller FPGA (BCF), BOC Main Fpga (BMF) North and BMF South). The input-output connections are shown in the figure.

and one SNAP12 transmitter. Two *Quad Small Form-factor Pluggable transceivers* (QSFPs), connected to the Spartan-6 BMF fast serializers, host the S-Link connection towards the ROS system.

BOC Control FPGA

The BOC Control FPGA (BCF) is in charge of board programming and all control operations. All the peripherals on the card are accessible through a *Wishbone* Bus, a open source bus purposely created to allow integrated circuits communications. This bus can be accessed from the ROD, via a *Setup-Bus*, an asynchronous configuration interface between the two cards providing 16 address, 8 data and 3 control lines. Another way to communicate to all the peripherals and the BOC registers is via Ethernet connector implemented on the BOC itself; all the Ethernet access and control operations are driven by a Xilinx *Microblaze* processor, a soft-core processor implemented on the Spartan-6 FPGA.

BOC Main FPGA

The two BOC Main FPGAs (BMFs) host the data-transmission path and the opto-electrical logic. Each BMF interfaces two commercial SNAP12 receivers, one SNAP12 transmitter and a QSFP, hosting the S-Link transmission to the ROS system. The TX path is used to send 40 Mbps commands and triggers to the modules encoded with a 80 MHz clock. In normal detector operation incoming data are *Bi-Phase Mark* (BPM) encoded; the detector timing is adjusted using coarse and fine delay blocks. The coarse delay is implemented using a variable-tap shift register clocked with 160

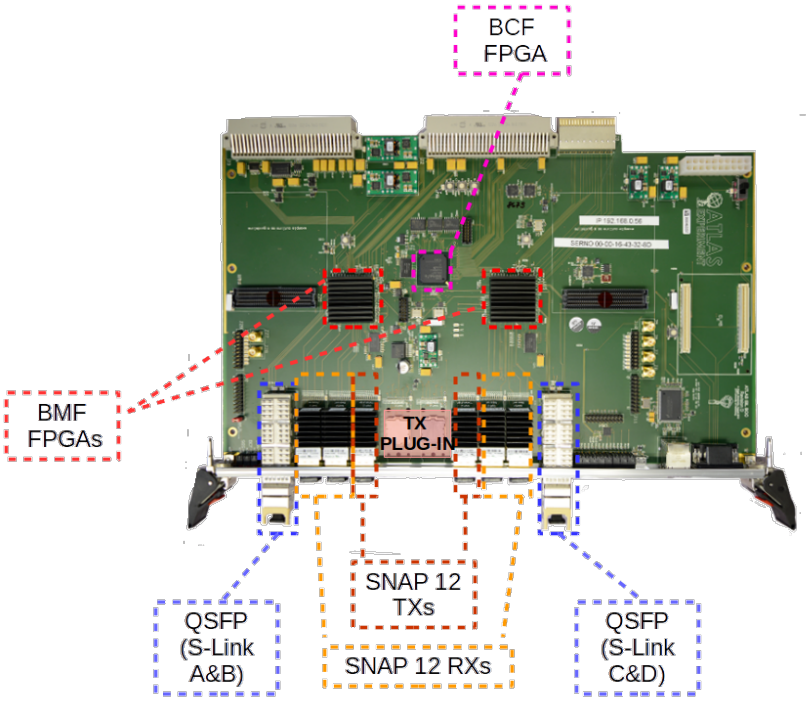


Figure 3.10: Picture of the IBL-BOC. The main components are highlighted.

MHz. The RX path is responsible for the reception and decoding of incoming detector data, that are 4x oversampled by the FPGA *Input SERializers DESerializers* (ISERDES). Data from the detector are 8b/10b decoded, collected and multiplexed to the ROD.

3.4.2 IBL-ROD

The IBL-Read-Out Driver (IBL-ROD) [37] [38] [40] [41] [42], also called simply ROD, is a 14-layer 9U Versa Module Europa (VME) 64x board comprising three fundamental logical blocks: operating blocks, memory and communication interface. Every component is showed in Figure 3.11.

The operating blocks include four FPGAs and a Phase-Locked Loop (PLL):

- a Xilinx Spartan-6 XC6SLX45-FGG484, a Program Reset Manager (PRM) device which routes the Joint Test Action Group (JTAG) connections of the other FPGAs and the EEPROMs in a chain, as depicted in Figure 3.12. Furthermore, the PRM communicates with the Single Board Computer (SBC) via VME bus;

- a Xilinx Virtex-5 XC5VFX70T-FF1136, the Master FPGA hosting an embedded PowerPC (PPC) CPU which controls and communicates with the required external devices (e.g., PC, TIM board);

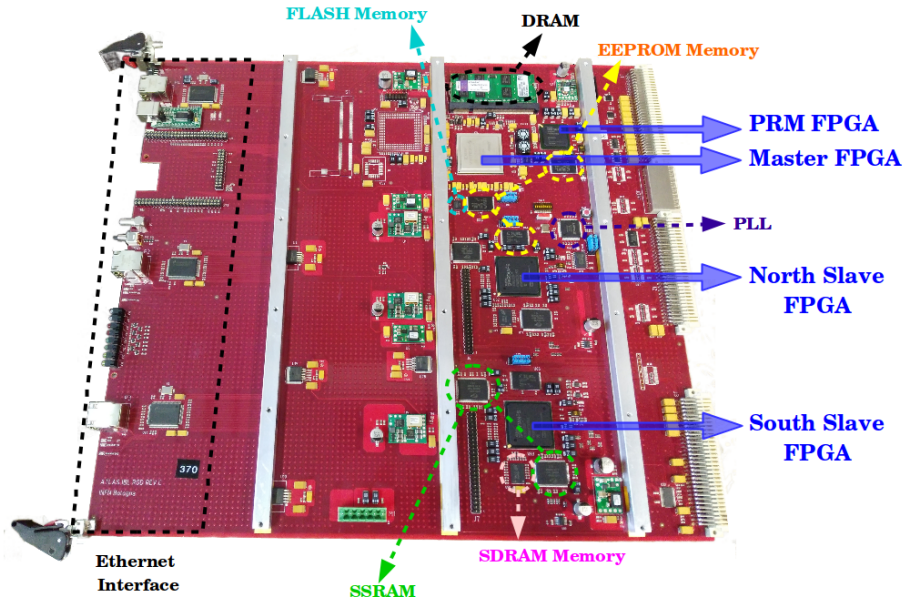


Figure 3.11: ROD board and main components. [41]

- two Xilinx Spartan-6 XC6SLX150-FGG900, the Slave FPGAs hosting a MicroBlaze CPU which performs data gathering, event fragment building and histogramming. MicroBlaze is a soft processor core, differently from the PowerPC which is a hard core;

- a LATTICE ispClock5600, a PLL which synchronizes the clock and selects its source, which can be internal or external coming from BOC.

Different types of memory are present in the ROD; a 16 MByte EEPROM for storing the PRM configuration file, a 32 MByte EEPROM for storing the Master configuration file, a 32 MbByte and a 16 Mbyte EEPROMs for each Slave FPGA and several types of RAM memories, either embedded or external, for the PPC and the MicroBlazes.

The communication between the different blocks inside the ROD is established by using several protocols. The PRM communicate with the VME Bus, with the Master FPGA via Serial Peripheral Interface Bus (SPI Bus) and with the PowerPC via HPI. The latter is a mechanism composed by a Processor Local Bus (PLB) Master peripheral implemented into the ROD Master, directly connected to the PowerPC bus, and a block into the PRM which translates VME access into operations. This mechanism allows the Single Board Computer to access the PPC address space from VME. Communication between Master and Slave operates via a 4 kB Dual-Ported RAM memory inside each Slave FPGA, accessible to the Master via a 40 MHz 16 bit bus named ROD-bus. PowerPC and MicroBlazes use a PLB or a On-Chip Peripheral Bus (OPB) to talk to the peripherals (e.g. ethernet chip, RAMs).

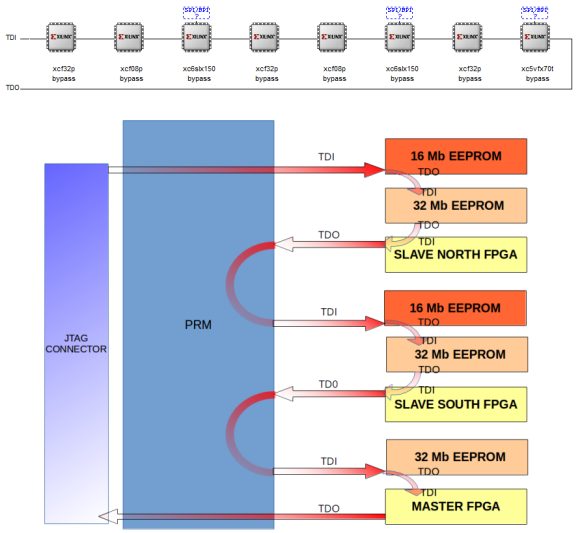


Figure 3.12: ROD board JTAG chain scheme (on top) and snapshot taken from Xilinx Impact Tool (on bottom). On bottom, from left to right: 16 MByte EEPROM (for Slave North), 32 MByte EEPROM (for Slave North), Spartan 6 FPGA (Slave North), 16 MByte EEPROM (for Slave South), 32 MByte EEPROM (for Slave South), Spartan 6 FPGA (Slave South), 32 MByte EEPROM (for Master), Virtex 5 FPGA (Master)

IBL-ROD data-path

The Read-Out Driver was designed to work in two operating modes: calibration and data-taking. The main tasks of the board are:

- to process data coming from BOC, deserializing and arranging them in a format compatible with the data-processing requirements;
- to organize data in histograms (only in calibration mode);
- to send configuration commands to front-end modules.

Calibration mode is used during the interval occurring between two LHC beam fills to adjust the threshold of the pixels, which reduces the noise and compensates the radiation effects. Histograms are created to show the pixel response to an injected charge at FE level. In this mode, the trigger is artificially generated by the PowerPC software, and the charge injected in the pixel is generated by an internal Digital to Analog Converter (DAC) in each FE chip, which provides a programmable voltage.

When operating in data-taking mode, on the other hand, the trigger is the ATLAS Level 1 Trigger, propagated to the ROD through the TIM board, and the charge injected in the pixels is deposited by the passing particles.

Both in calibration and data-taking, the ROD incoming data pass through several logical blocks in order to be processed. These logical blocks, i.e., the Formatters, the Event Fragment Builders (EFBs), the Routers, the Histogrammers and the Merger are opportunely designed by the Firmware of the ROD Slave FPGAs for Pixel Layer

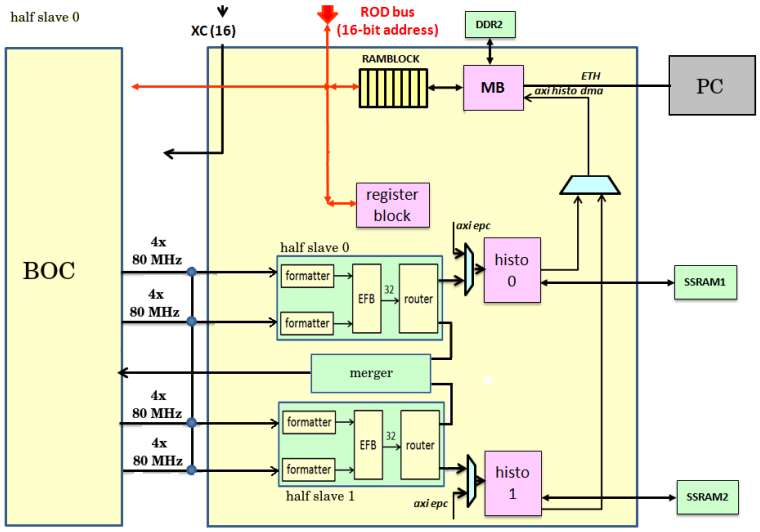


Figure 3.13: Firmware block scheme of the Pixel Layer 2 ROD Slave FPGA. Only one FPGA is represented.

1 and Layer 2. Two formatter, one EFB and one router are grouped in the *half_slave* block, which is duplicated for each FPGA. Figure 3.13 shows the ROD Firmware blocks and their connection.

The Formatter block is the first stage of the data path in the ROD slave. Each formatter manages data coming from four links and performs data deserialization and buffering in one 2048x32 First In First Out (FIFO) per link. Also, the formatter block sends one event per link channel in a fixed order starting from link 0 up to link 3.

The EFB block processes data coming from two Formatters adding header and trailer informations such as trigger and event ID. Another task of the EFB block is to check that the ATLAS Level 1 Trigger ID and the Bunch Counter ID coming from the modules matches the one received from the TIM board. If there is a mismatch between the two IDs (e.g. if some triggers arrive to the module when its internal buffer is full and the triggers are skipped) an error flag is added to the output data stream.

The Router block is in charge of routing data to the Merger in data-taking mode or to the Histogrammer in calibration mode. In data-taking mode the Merger takes data coming from the routers of both the half_slaves and creates a unique data-stream which is sent back to the BOC board. The BOC then sends data to the ATLAS PC Farm via two 1.28 Gbit/s S-Links.

In Calibration mode, viceversa, data are sent to the Histogrammer block, which collects information about the pixels, such as the Time Over Threshold (TOT), and creates a map of pixels calibration responses inside a dedicated memory. These pieces of information are sent to a monitoring PC via Gigabit Ethernet Ports, driven by the MicroBlaze Software.

Comparison between SiROD and IBL-ROD

The IBL-ROD was designed to be an upgrade of the Silicon Read-Out Driver (SiROD), which has been working since the very beginning of the ATLAS experiment. Tab. 3.4 lists the main differences between the old SiROD and the new ROD. The number of integrated circuits (FPGAs, FIFOs and DSPs) was decreased

Table 3.4: Feature comparison between old SiROD and new ROD

	SiROD	**ROD**
Number of FPGAs	12	4
Number of FIFO chips	12	0
Number of DSP chips	5	0
Max bandwidth	Mbits/s	Mbits/s
Layer 1	80	160
Layer 2	40	80
Software Interaction	VME	Ethernet

thanks to the technological improvements of the digital components. For instance, the SiROD Spartan 2 was replaced with a Spartan 6, performing higher number of tasks in a single chip. Moreover, these more recent FPGAs permit to use CPU cores, i.e. MicroBlaze, a highly configurable IP core, and PowerPC, a CPU core embedded in the Virtex-5 IC. As a consequence, DSP chips, which previously performed online histogramming and interacted with the software infrastructure, are no longer required. The primary advantage of this choice consists in speeding up the communication with the software infrastructure by using Ethernet protocol instead of VME, which requires a protocol master (the SBC inside the crate).

The upgraded ROD features lower power consumption and lower active area. However, retro-compatibility was required and the ROD size could not be changed. As a consequence, the lower IC density reduces the noise in the PCB and the possible couplings among analog and digital signals (crosstalk). At the same time, the heat dissipation is favored. Fig. 3.14 depicts the IC density differences between the two Read-Out Drivers.

3.4.3 Readout upgrade for Pixel Barrel and Disks

In Section 3.3.1 the limitations of the old Pixel Detector readout were shown. As it can be seen from Table 3.3, it is clear that, with the increasing performance of the LHC accelerator and the more critical conditions of Run-II, the old readout could not provide solid and stable results. The main limiting factor for the old readout system is the total bandwidth of the Pixel BOC/SiROD pair: 1.28 Gbps. This limitation led to the original choice of interfacing the FE-I3 and MCC front-end electronics at lower speed with respect to the maximum readout speed supported by the modules. While this design was efficient and produced good results during Run-I, in Run-II it became unsustainable, and the need of running the modules at higher

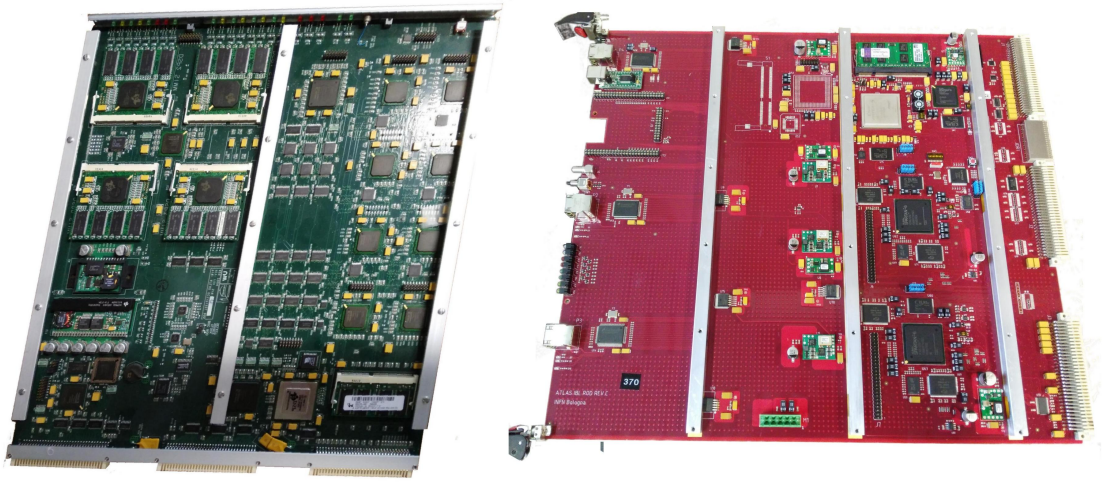


Figure 3.14: Comparison between the IC density of the SiROD (left) and the ROD (right).

readout speed arose. To increase the readout speed of each module, while at the same time maintaining the same amount of modules connected to each off-detector board and hence not augmenting the total amount of ROD-BOC pairs required, the only solution is to replace the ROD-BOC boards with ones with more bandwidth available. The IBL-ROD and IBL-BOC proved to be excellent candidate for this readout upgrade, since they showed optimal results with IBL. The upgrade started in the 2015-2016 LHC Shutdown with Layer-2, the most critical one, whose module readout speed was doubled from 40 Mbps to 80 Mbps. The same year 6 out of 38 Layer-1 boards were upgraded, without changing the readout speed. During 2016-2017 Shutdown the rest of Layer-1 boards were upgraded, and its readout speed was doubled from 80 Mbps to 160 Mbps.

Table 3.5: Module Link Occupancy estimation based on 2016 Run at different pileups (μ) assuming Level 1 trigger rate of 100 kHz and 13 TeV energy.

μ	Link Occupancy				
	B-Layer	Layer-1		Layer-2	
	both readouts (160 Mbps)	old readout (80 Mbps)	new readout (160 Mbps)	old readout (40 Mbps)	new readout (80 Mbps)
40	60%	81%	41%	**119%**	59%
60	81%	**103%**	52%	**159%**	79%
80	**101%**	**125%**	63%	**188%**	98%

Finally, during 2017-2018 LHC Shutdown, Disks and B-Layer readouts were upgraded as well. In the latter case the readout speed of the previous system was already the maximum supported by the MCC chip and could not be ulteriorly increased. However, the upgrade was anyway beneficial, since the number of RODs

required to interface the entire sub-detector was halved and the system was uniformed. Table 3.5 highlights the benefits, in terms of link occupancy, of the readout upgrade for all the barrel layers of the Pixel Detector, based on data extrapolated from 2016 run.

Differences between IBL and Pixel ROD firmware

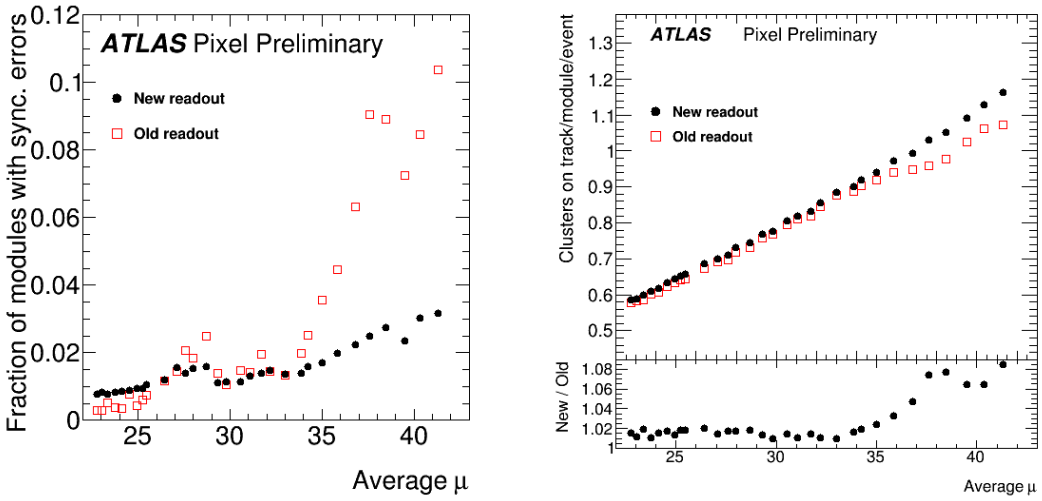
The ROD Firmware for Pixel Layers 1 and 2 is based on the Firmware designed for IBL. However, while the PRM and Master firmwares are unchanged, the Slaves firmware needed few changes, especially at the Formatter and Merger levels. The main differences between IBL and Pixel read-out system lie in the front-end data format and the S-Links blocks. The different FE-I4 and MCC data format led to the need for designing two different decoding mechanisms inside the Formatter. Moreover, the IBL BOC has four S-Links (two for each ROD Slave FPGA), whereas the Pixel BOC has only two. As a consequence, IBL does not require the Merger block whose task is to merge the output of two Router blocks into a unique data-stream.

3.5 Challenges and system development

The ATLAS Pixel Detector DAQ system is in continuous evolution; although the hardware components are already in place, many changes and improvements can be done on the TDAQ software and the firmware running on the boards. This section will present some of the most critical readout aspects for the Pixel Detector and the solutions and strategies adopted to solve, or at least mitigate, the problems. The task of all the developers and experts working for the ATLAS detectors is to provide a solid and reliable system capable of recording the maximum amount possible of good quality data minimizing dead times. There are three main categories of problems that affect the quality of the readout system and constituted a big challenge for the developers. The first one involves the memories (in the ROD or in the ROS) where data coming from the front-ends are stored before further processing. If whose memories are full, the system sends a back-propagating signal to the trigger system to slow-down the trigger rate. This signal is called *Busy signal*, or simply *busy*, and constitutes a major source of dead-time for the ATLAS detector and it is particularly problematic because it doesn't affect only the Pixel Detector but also all the other ATLAS sub-detectors. Another problem for the TDAQ is the *loss of synchronization*, happening when the data propagated to the event reconstruction are associated to the wrong L1 trigger. Finally, the third source of readout inefficiencies is due to the module loss of configuration caused by bit-flips in the module registers. This effect is due to radiation passing into the detector and it is more evident when the beam luminosity is at its peak.

3.5.1 *Busy* dead-time

The *busy* signal is a signal asserted by the readout chain which is back-propagated to the trigger system to slow down or stop completely the trigger rate. It is generated when the memories containing the front-end data to be processed gets full. There are several sources of *busy* in the detector, each one acting in a different memory and with different consequences. This section will investigate in detail some of the busy sources and the actions that were taken in order to mitigate their effects. A very first approach to reduce the dead-time is to increase the size of the memories where data are stored; while sometimes this can be done in a relatively effortless way, in many situation this strategy is not feasible and a workaround must be found. The busy signal is generated principally by the ROD, although under particular conditions, that will be discussed later in this section, it could be generated by the ROS system.



(a) Fraction of modules with desynchronization error

(b) Clusters on track

Figure 3.15: The fraction of modules with synchronization errors and number of clusters on track per module per event as a function of pileup (μ) in the old and new readout at Layer-1 in fill 5446 in 2016. The new readout shows less synchronization errors with respect to the old readout in the same eta coverage, especially in the high pile-up region. This achievement was possible thanks to a mechanism that flushes FIFOs in the readout electronics at each Event Counter Reset (ECR), introduced for the new read-out in 2016. In 2017, this mechanism was also introduced for the old readout

A proper understanding of the ROD data-path, already described in Section 3.4.2, is fundamental to comprehend the busy origins. The *Formatter* block inside the ROD slave FPGA collects data from several front-ends and stores them into *First-In-Fist-Out* (FIFO) memories. When data from all the front-ends are available, the data-stream is moved to the *Event Fragment Builder* (EFB) block for further processing. The size of the Formatter FIFOs is a very important factor in the readout

system; they must be big enough to contain at the very least a full front-end event, but at the same time they must fit within the FPGA resources. Due to the occupancy and trigger rate increase, the FIFO size was moved from the original value of 1 k-Words to 3 k-Words. If for some reason one of the front-ends stops sending data or is slower than the others, the events of the other modules get stuck inside the Formatter FIFOs, that will quickly get full, generating a busy signal. This effect can be significantly mitigated by adding a *Timeout* on the time to be waited for a front-end response. The timeout setting must be tuned and it depends on the internal propagation time of the front-end and on the event size; in fact, especially in the FE-I3 and MCC modules, a larger occupancy corresponds to a larger response time, as an effect of the column-driven architecture. After several attempts, the timeout thresholds were set to $\sim 50\,\mu s$ for IBL and $\sim 500\,\mu s$ for the barrel and disks. Another way to avoid busy signal generation due to data getting trapped into the various FIFOs present in the ROD logic structure, is to regularly send a *reset* signal that empties the content of the FIFOs. A perfect occasion for sending the reset is provided by the ATLAS trigger structure; in fact, every ~ 5 seconds, a *Event Counter Reset* (ECR) signal is propagated to every ATLAS sub-detector, and the trigger is stopped for $\sim 1\,ms$, time more than sufficient for the ROD firmware logic to empty the content of all its memories. A direct measure of the effects of this mechanism was done in 2016 in Layer-1, since at the time the old readout - not implementing this feature - and the new readout were coexisting. Fig. 3.15 shows the results of the *reset at ECR mechanism*, that can be quantified also in terms of amount of modules de-synchronized and track clusters reconstruction abilities.

Another very important source of busy in the Pixel Detector readout is given by the presence of *bugs* in the firmware and software system. While this can seem obvious, experience showed that many problems arose only under very specific circumstances and were generating very different and apparently uncorrelated effects. The most typical final result is the generation of busy signal in the ROS system. In fact, every time a ROD fragment is incomplete or presents an error, e.g. there is a missing word in the event header or trailer, data from a particular module are incomplete, or there is a mismatch on the number of words counted, the ROS PC corresponding to that ROD stores the problematic word in a particular memory that is emptied at a slower frequency with respect to the the *standard* event buffer. If there is a burst of incorrect fragments, the ROS memory gets completely full and the busy signal is generated. Those memories are not cleared by the arrive of the ECR signal, so the consequences could be potentially catastrophic until the point that the entire ROD have to be disabled and then re-enabled in a second moment. In 2016, when the new readout was first installed for Pixel Layer-2 (with different modules compared to the FE-I4s of IBL), some timing propagation issues inside the ROD firmware were causing a high amount of corrupted fragments, and to avoid a 100% dead time caused by the ROS busy, the ATLAS trigger rate was lowered to $\sim 50\,kHz$ until the problem was solved. 2018 runs also showed data-corruption problems influences the ROS performance. In particular, due a a combination of a

bad opto-link tuning and temporization of a ROD FIFO, together with a very high hit occupancy, the total number of words of a full S-Link fragment was reported incorrectly. Fig. 3.16 shows that, after a firmware intervention, the amount of fragment corruption per run dropped drastically to ~ 0.

Figure 3.16: Number of *Event words counter mismatch* errors in ROD-generated fragments. After a ROD firmware intervention the amount of errors dropped to ~ 0 as well as the busy fraction produced by the ROS.

3.5.2 De-synchronization errors

The ATLAS TDAQ system must collect data from all the hundreds of thousands of channels of all the sub-detectors, select them and proceed to build the full event. It is hence extremely important to maintain the system completely synchronized; if the synchronization is lost the DAQ chain will not be able to properly build the event. If only a fraction of the modules is disaligned, the ATLAS reconstruction algorithms will still be able to build the fragments, but with less precision; every ATLAS event is associated to a *weight*, according to its accuracy. For the Pixel Detectors, the system synchronization is verified through the monitoring of two information: the *L1 IDentifier* (L1ID), which is a counter of the L1 triggers sent, and the *Buch Crossing IDentifier* (BCID), that identify the bunch number within the LHC beam. Those two information are monitored independently by the trigger system (TIM) and by the modules (FE-I3/MCC/FE-I4) and are compared by the ROD during event fragment building. If there is a mismatch between the front-end and the TIM counters, the module data are flagged as *de-synchronized* and will be discarded by the ATLAS reconstruction chain. There are several sources of module de-synchronization; the DAQ system must be able to minimize the errors and limit the damages. The main mechanisms generating de-synchronization are *timeout* and *skipped triggers*. As already discussed in Section 3.5.1, the timeout mechanism prevents the system to get stuck in case a front-end stops sending data.

However, in case of a particularly big event size on the module (i.e. several pixel hits), sometimes the timeout protection takes action before the front-end was able to transmits its data. In this scenario the module data are associated to the following event, which corresponds to a different trigger; all the following events will be shifted as well, and the system will be kept in a constant de-synchronized situation. To avoid similar situations where the system is in a permanent erroneous condition, the ATLAS system sends periodically two reset signals: the *Event Counter Reset* (ECR), already introduced in Section 3.5.1 and sent every $\sim 5\,s$, and the *Bunch Crossing Reset* (BCR), that resets the BCID counter after each orbit ($\sim 100\,\mu s$). If all the ROD FIFOs are emptied when the ECR is asserted, the system will be *refreshed* to a status where all the modules are perfectly in line. The impact of this reset mechanism on the system can be seen in Fig. 3.15.

The second main source of de-synchronization is due to *skipped triggers*, a mechanism shortly introduced in Sections 2.2.2 and 2.3.2 together with the basic functioning principles of the FE-I3, MCC and FE-I4. The front-ends can store up to 16 triggers in their internal buffers; if more than 16 triggers arrives before the chip was able to process the first one, the extra triggers are *skipped*, meaning that all the hit information are not recorded and are not propagated by the front-end. The amount of skipped trigger generated depends on the trigger rate, the mean hit occupancy and the internal chip propagation time, which is sensible higher for the FE-I3/MCC than for FE-I4. For this reason the de-synchronization due to skipped triggers is higher for the barrel and disks and is lower for IBL, even if the latter is closer to the beam-pipe and subject to more charge density. As for timeout, if a trigger is skipped and no corrective action is taken, the system will be permanently de-synchronized till the next ECR signal. Many intervention were done in the ROD firmware to provide a solution to this problem.

Skipped triggers re-synchronization

The MCC and FE-I4 front-end chips send informations on the amount of skipped triggers that were generated in the system. The MCC transmits a 4-bits counter of the amount of skipped triggers between two processed triggers, while the FE-I4 sends to the ROD a 11-bits counter of the cumulative amount of skipped triggers. The information are encoded in the front-end data stream together with others *Service Records* and other error messages. The ROD firmware must use those information to re-align the system.

The skipped trigger re-alignment algorithm was added in 2016, after observing that the total amount of de-synchronization for Layer-2 was reaching unbearable levels. The algorithm working principle is very simple; in case of skipped triggers, the module stream is paused and empty fragments, properly flagged as *fake data*, are added to the data-stream. In this scenario the system will be perfectly aligned afterwards. Fig. 3.17 and Fig. 3.18 show the effect of the introduction of the algorithm over the system; after the ROD firmware was changed for Layer-2, the amount

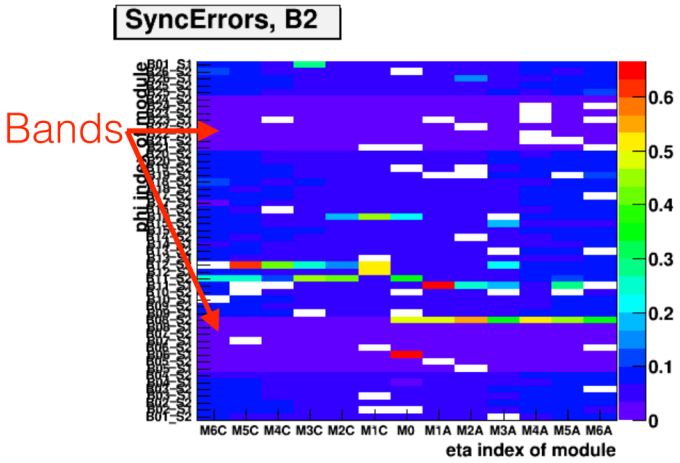


Figure 3.17: Screenshot of the online Layer-2 monitoring de-synchronization plot from a ATLAS Run of June 2016. In this run a new ROD firmware, implementing the *skipped trigger re-synchronization algorithm*, was flashed in only a sub-set of all the Layer-2 RODs, corresponding to the purple bands. The plot shows that the new algorithm greatly reduced the total amount of de-synchronization.

of de-synchronization dropped of a factor ~ 10.

Re-synchronization limits

The implementation of the skipped triggers re-synchronization algorithm showed great results in the barrel and disk layers of the Pixel Detector. However, a further study of the detector behavior, showed that even after the introduction of the realignment mechanism, an important fraction of *untreated* de-synchronized events was still present in the system. If the system is not able to restore the modules alignment in real time, the de-synchronization errors will be propagated till the next ECR signal. After the ECR the system is synchronized again; however, the time spacing between two ECRs (five seconds) is insufficient to provide a stable solution of the problem.

Fig. 3.19 shows that the total de-synchronization presents two main components; one irreducible amount, caused by skipped triggers, and several spikes caused by *untreated* situations, whose magnitude is $\sim 10^3$ times greater than the others. The origins of the residual de-synchronization come from timeout effects, already discussed in the previous section, and the unreliability of the skipped trigger information sent by the modules. In fact, even if the MCC and FE-I4 transmits to the ROD the amount of triggers that were skipped between two consecutive events, this information is not always accurate. Simulations and tests with the real chips proved that the FE-I4 information is completely unreliable and cannot be used at all; for this reason the skipped triggers re-alignment algorithm was never implemented for IBL. Fortunately, the chip itself produces a very low amount of skipped triggers,

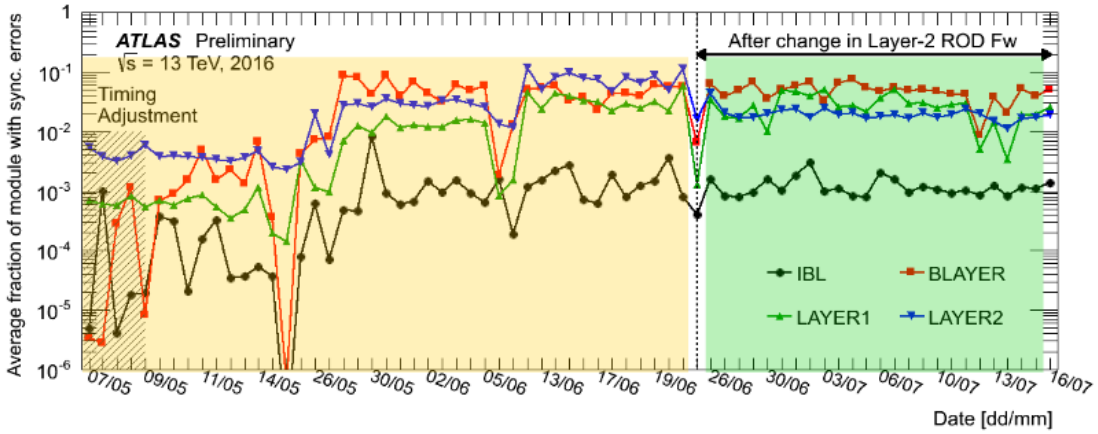


Figure 3.18: Fraction of modules with de-synchronization errors for all the layers of the Pixel Detector after the readout upgrade of Layer-2 (2016). After the introduction of skipped triggers re-synchronization algorithm, it can be observed that the amount of de-synchronization of Layer-2 (new readout) is comparable to or better than Layer-1 or B-Layer (old readout). For IBL the de-synchronization is generally lower, thanks to a combined effect of the new front-end module and the new off-detector readout.

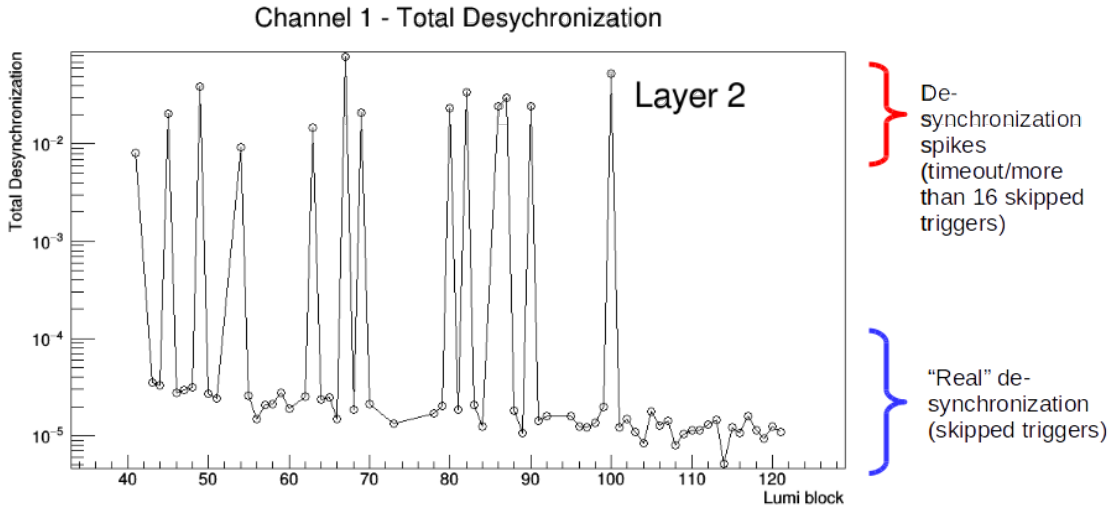


Figure 3.19: Total fraction of ROD de-synchronized events for one Layer-2 channel in a 2016 run. The plot shows two different contributions: the *normal* skipped trigger behavior, corrected by the firmware, and a series of de-synchronization spikes, where the ROD was not able to re-align the system before the arrive of ECR signal. The spikes are induced by the presence of timeout and of more than 16 skipped triggers.

so this problem never significantly affected the readout system. The information is more reliable in the MCC front-end, which also produces a greater amount of skipped triggers (due to the column driven architecture of the front-end chip). However, the MCC transmits the amount of skipped triggers information via a 4-bit counter, meaning that it can count up to 16 skipped triggers. While this is sufficient for situations with a low occupancy and a low acquisition rate, problems may arise when the system is run in more critical conditions. A study on the distribution of the skipped triggers as reported by the MCCs showed that the situation producing more than 16 skipped triggers are present in a small but not negligible fraction. In case of a bad report, the re-alignment algorithm is not able to properly re-synchronize the system, and new strategies must be exploited.

Smart re-synchronization algorithm

The first implementation of the re-synchronization algorithm was based on the skipped triggers information provided by the front-end modules. However, as it was discussed in the previous section, those information are not always reliable, and alternative solutions could be exploited. Since both the TIM and the modules were providing independent L1ID and BCID information, in 2017 a study to use the two set of counters to perform on-line re-alignment began.

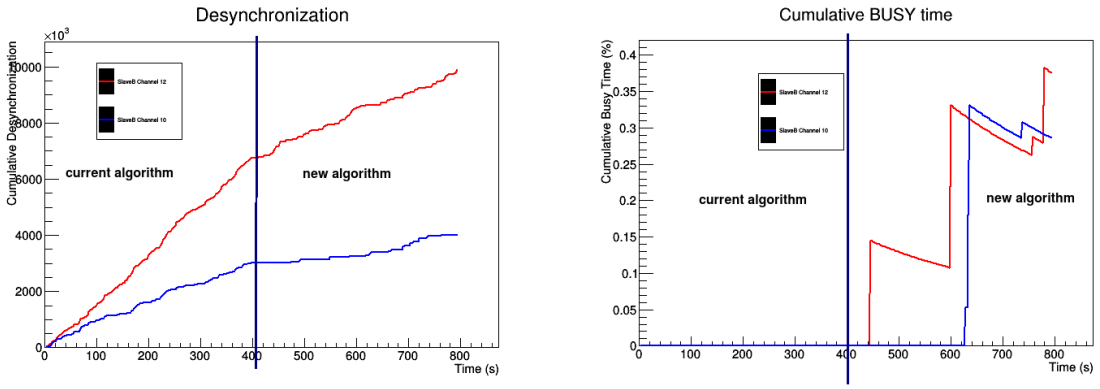


Figure 3.20: Comparison of total amount of de-synchronized events between the old and new re-alignment algorithm. The test was done in a *SR1* test setup at a 40 kHz trigger rate and average occupancy of 20 hits/event. While the new algorithm shows a lower amount of de-synchronization, it also shows a certain percentage of busy dead-time.

The idea between the new algorithm is very simple, if the two L1ID information does not coincide, the module data are stopped and fake empty fragments are inserted until re-alignment, that is obtained only if both the L1ID and BCID counters are in line. Unfortunately, this solution presents few problems:

- it is highly influenced by bit-flips in the data-stream. In fact, if there is a bit-flip on the module L1ID counter information, the ROD will erroneously

identify the module as de-synchronized and insert empty fragments when not necessary;

- the MCC L1ID information is stored in a 4-bits counter, meaning that a maximum de-alignment of 16 triggers can be observed, although previous studies already showed that more than 16 skipped triggers could be generated. Fortunately, re-applying the algorithm until the system is aligned overcomes this issue;

- the algorithm stops the module data stream and add empty triggers until the system alignment is reached. If the re-alignment is not obtained, module data are trapped forever in the ROD FIFOs, that will get quickly full generating a *busy* signal.

The validation of the new algorithm was realized in the ATLAS test setup, called *SR1*, where many sub-detectors are replicated to allow fast development. In the Pixel Detector test system, two modules are particularly noisy and are perfectly suited to reproduce a situation with high amount of de-synchronization and to test new solutions. The final results (Fig. 3.20) showed that the new algorithm was performing very well in terms of reduction of de-synchronized events, but it was also generating some not negligible percentage of *busy* dead time, for the reason aforementioned.

Smart L1 forwarding

So far, all the mechanism developed for reducing the amount of de-synchronized events were focused on correcting the data-stream in case of skipped trigger. However, the increasing occupancy and trigger rate foreseen for the next years will be a cause of stress for the modules, and the DAQ developers started to search a solution to relieve the chips from this stress. In particular, the system should dynamically be able to prevent the generation of skipped triggers without at the same time affecting the ATLAS trigger rate. A study and development of such a mechanism, called *Smart L1 forwarding*, started in 2018 and is still incomplete. The only way to prevent generation of skipped events is to stop sending L1 triggers to module in *critical conditions*. L1 triggers are still generated, since no back-propagation signal is sent to the trigger system, but they are not forwarded to the modules whose buffers are full. Such a mechanism implicates several logical steps:

1. *identification of modules in critical situation.* A skipped event is generated if a module receives more than 16 triggers before it was able to process the first one. To prevent the generation of skipped events, the amount of *pending triggers* of each module must be constantly monitored. The amount of pending triggers is simply calculated as: *number of triggers sent to the module - number of fragments arrived from the module*. If the amount of pending triggers is above a configurable threshold, the module is flagged as *critical*.

2. *L1 triggers are sent only to the non-critical modules.* If a module is flagged as *critical*, the ROD stops sending it the L1 triggers until the amount of pending triggers lowers under the critical threshold. Trigger are sent in the usual way to all the other modules.

3. *Empty event insertion.* The ROD must insert as many empty event as many triggers were not forwarded to each module. This way the total data-stream remains aligned.

4. *Module L1ID compensation.* If a module does not receive a trigger, its internal L1ID counter does not increase and will be disaligned with respect to the TIM L1ID. The ROD must be able to calculate and compensate for this misalignment.

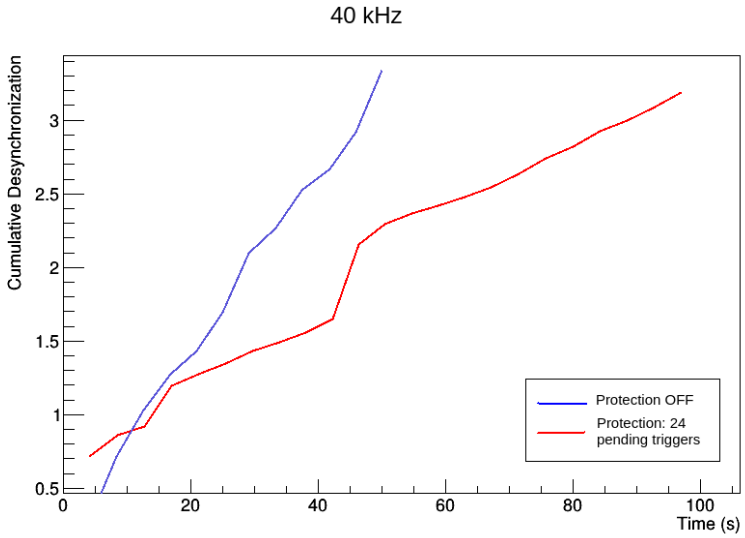


Figure 3.21: Comparison of the effect of the *smart L1 forwarding* protection algorithm over de-synchronization. The test was performed in using one SR1 module at a trigger rate of $\sim$ 40 kHz and an average occupancy of $\sim$ 20 hits/trigger. When the protection is enabled, the total amount of de-synchronized events is significantly lower.

Test results performed in the SR1 test-setup showed that, at the current status, the *smart L1 forwarding* protection mechanism exhibits a higher number of de-synchronized events if the *critical threshold* is set to a low value ($\sim$ 16 pending triggers), but better results if the threshold value is rised to $\geq$ 24 pending triggers (Fig. 3.21). This mechanism is still under testing and development, and will be released for Run-III.

3.5.3 Loss of module configuration

The ATLAS Pixel Detector is located very close to the particle beam and is subjected to a huge amount of radiation dose. Although several precautions were taken

to design a radiation hard electronic system, during a several hours ATLAS run several *Single Event Upset* (SEU) events occur. The SEU was already discussed in Section 2.2.2 and afflicts the DAQ causing bit-flips in data transmission or in the registers of the chip. While data-transmission errors don't afflict the overall behavior of the system in a permanent way, bit-flips in the configuration registers can be very problematic, since the chip can lose its working parameters and behave in a completely unpredicted way.

From the operational point of view, when a module loses its configuration there are two main possible effects on the readout chain:

- the module stops sending data, and its readout channel is constantly in *timeout*, causing a limitation of the trigger rate;
- the module starts sending *noisy* data, filling all the ROD FIFOs and hence producing *busy* signals that pause the trigger.

Once the module chip design is ultimated, it is impossible to reduce the amount of SEU effects on it. However, the DAQ developers can provide a system to quickly and efficiently detect *problematic* modules and try to recover their status before the overall chain is affected. In the next few sections some of the strategies adopted by the Pixel Detector DAQ experts will be introduced and discussed.

Quick Status

Quick Status is a firmware/software system running on the ROD board. It constantly monitors the status of the system, identifies problems in real time and performs actions to *fix* them. The purpose of Quick Status is to automatize the recovery procedures of the front-end modules and to reduce the dead time to ~ 1 ms, which is several orders of magnitude faster than human response time. It uses at the same time the best properties of the FPGA, that can perform real time data check, and CPU software, where complex algorithms can be implemented. The ROD Slaves FPGAs constantly analyze data, storing information and error flags for each module such as *timeout errors percentage*, *busy fraction percentage*, *de-synchronization errors percentage* or *occupancy percentage* in internal registers. The PowerPC embedded in the ROD Master FPGA runs a software that reads the content of those registers every millisecond and decides if the module is not behaving as expected. The decision is done comparing the content of the registers to user-defined threshold that can be tuned to maximally optimize the system behavior. If a module is identified as problematic, the software performs a series of sequential actions; firstly, the front-end is *masked* from the readout, meaning that its data are not propagated to the ROS PC and all the errors coming from its associated FIFO memory are ignored. After the module has been *disabled*, the software proceeds with *reconfiguration*, sending again all the proper register values. As the final step, Quick Status waits till the next ECR signal and re-includes the module in the readout chain. The entire operation lasts $\sim 5/10$ seconds but, since the module has been immediately masked, the rest of the

DAQ chain is almost completely unaffected during this time. If after reconfiguration the module is still in a *problematic* state, Quick Status repeats the operation from scratch; in case of multiple reconfigurations failure, the module is disabled until the end of the run.

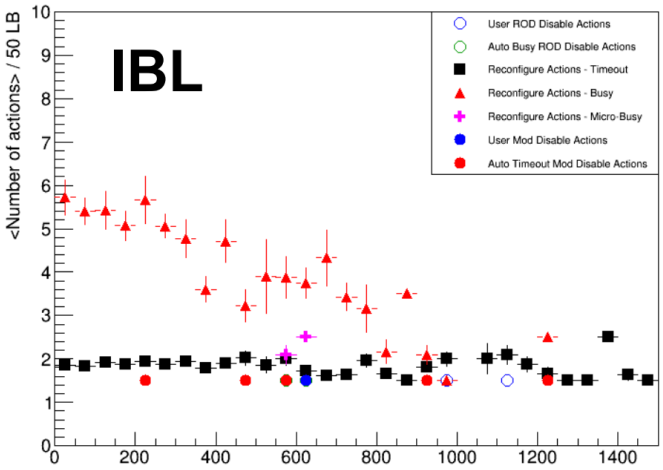


Figure 3.22: Number of Quick Status actions over 50 Lumi-Blocks for IBL in 2018. All the different problems in the ROD causing the actions are highlighted.

The average number of reconfiguration actions operated by Quick Status software is shown in Fig. 3.22, while the results on the system in terms of number of events out of synchronization is shown in Fig. 3.26

Register reconfiguration at ECR

Quick Status acts as a *corrective* system and takes action only after the module configuration has been corrupted. Preventing SEU events is practically impossible, but a lot of improvement could be obtained by continuously refreshing the front-end registers. Unfortunately, both FE-I3/MCC and FE-I4 have a unique line for sending both trigger and configuration commands, so it is impossible to send reconfigurations during data-taking operations. However, the ECR signal from the ATLAS trigger system is sent every ~ 5 seconds and stops the triggers for 1 ms. During this time the module command line is normally unused and configuration registers can be sent without affecting the data-taking at all. In a first implementation, the time immediately after the ECR was used to send only the *Front End Sync* command to the FE-I3; this command acts as an internal front-end counters reset and its usage resulted in a general improvement in terms of reduction of de-synchronization errors (Fig. 3.26). Further studies on the overall system stability showed that the triggerless millisecond after the ECR is sufficient to configure completely both the global and the pixel-local front-end registers. Local registers reconfiguration is needed because it was observed that the increase of noisy and quiet pixels within a run, showed

in Fig. 3.23, is due to corruption on Pixel Latches. Since the timing for reconfiguration is very tight and the software does not have a fixed and easily predictable response time, the mechanism uses a combination of software and firmware logic. The software stores the proper configuration values in a memory of the BOC, which provides to forward the commands to the front-ends in real time. The full mechanism is still under development, and its deployment is foreseen for end of Run-II or the beginning of Run-III.

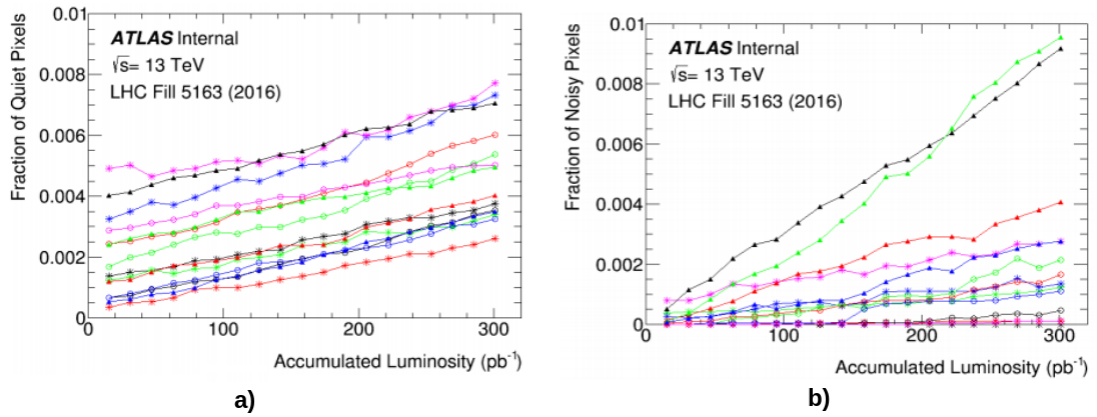
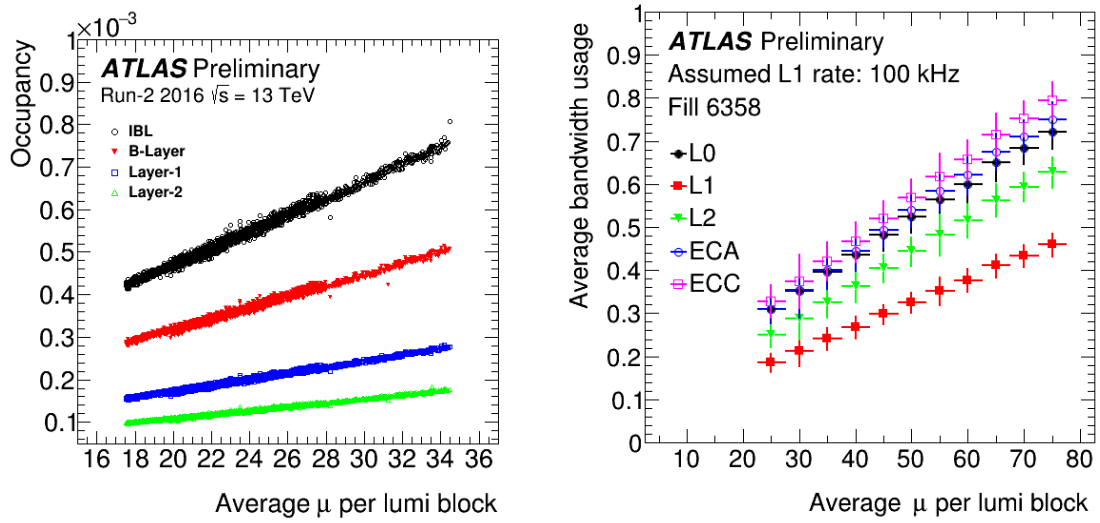


Figure 3.23: Fraction of quiet (a) and noisy (b) pixels over integrated luminosity during LHC Fill 5163 (2016).

3.6 Results

In this chapter the main goals, objectives and architecture of the ATLAS Pixel Detector readout chain were introduced, as well as all the technical difficulties and challenges that led to its upgrade during Run-II. Particular focus was put in the main problems involving the TDAQ and all the solutions and strategies adopted to improve the system stability and performance. The LHC beam and luminosity conditions put the Pixel Detector in a very challenging situation; the total bandwidth was nearly saturated and many radiation-related problems arose. However, the constant effort of all the developers resulted in a reduction of the total readout dead-time and in the improvement of data quality. Overall, the Pixel Detector TDAQ system performed very well during Run-I and Run-II and was perfectly able to meet all the specifics required and to provide ATLAS with all the data needed for its searches on physics. This section will present some of the results obtained during Phase-II, focusing on the system conditions and the improvements obtained by the DAQ upgrade.



(a) PIX-2016-007 (b) PIX-2017-009

Figure 3.24: (a) Number of hits per pixel per event in IBL and Pixel barrel layers in two runs in 2016 over pileup μ. (b) Projection of the average usage of the MCC output bandwidth for the three barrel layers (L0, L1, L2) and the end-caps (ECA, ECC) of the ATLAS pixel detector. The available MCC output bandwidths are 160 Mbit/s for L0 and L1, and 80 Mbit/s for L2 and end-caps, respectively.

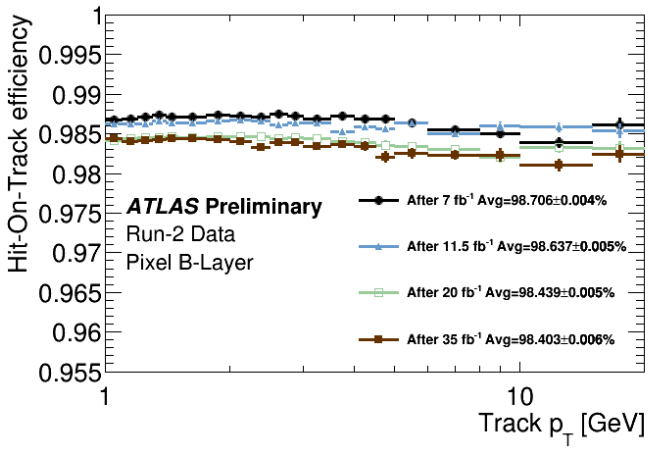


Figure 3.25: Hit-on-track efficiency in the Pixel B-Layer as a function of the track pT measured with the LHC fills 5021 (black dots), 5080 (blue triangles), 5199 (green squares) and 5437 (red squares) that have been collected after a total integrated luminosity of about 7 fb-1, 11.5 fb-1, 20 fb-1 and 35 fb-1, respectively.

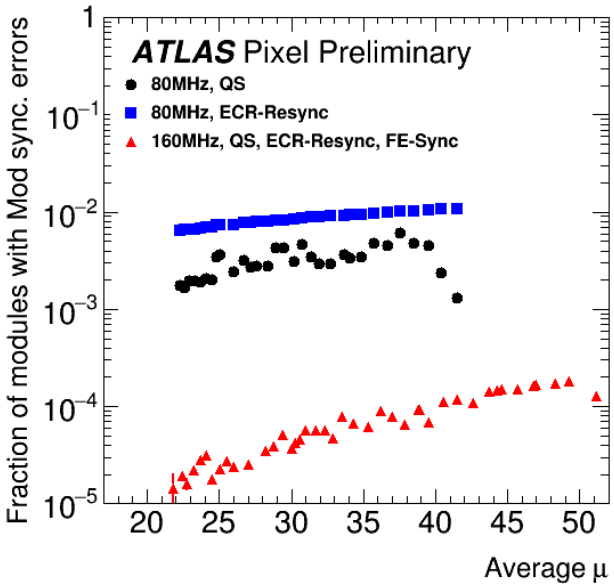


Figure 3.26: The fraction of modules with synchronization errors at the module level as a function of $\mu\times$ *trigger rate* averaged over 20 LBs at Layer-1 in fill 5446 in 2016 and fill 6243 in 2017. μ is the number of interactions per bunch-crossing. This plot is made with physics_ZeroBias. The data points include error bars. There are several methods to improve sync. errors: automatic recovery in software (QS: Quick Status), readout speed from a front-end chip (80 or 160 MHz), mechanism to reset the machine state and clear all the FIFOs in the backend readout electronics at each Event Counter Reset (ECR) (ECR-Resync) and functionality to reset the internal trigger FIFO and clears all memory of the events currently stored in the front-end chip at ECR (FE-Sync). The data were taken with QS and 80 MHz readout speed in the old readout in 2016 (black circle), with 80 MHz readout, ECR-Resync and FE-Sync in the new readout in 2016 (blue rectangle), and with QS, 160 MHz readout, ECR-Resync and FE-Sync in 2017 (red triangle). The sync. error at the module level was reduced mainly by adopting QS, 160 MHz readout speed, FE-Sync. [PIX-2017-007]

Part II

LHC Phase-I/II

Chapter 4

ATLAS Pixel Detector for Phase-II LHC

At the end of Phase-I, foreseen for the end of 2024, LHC will undergo a phase called *Long Shutdown 2* (LS2) that will last for approximately two years. During this period many parts of the accelerator will be upgraded to enhance the instantaneous luminosity to a nominal value of $7.5 \cdot 10^{34}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, several times higher than the Phase-0 and Phase-I luminosity. The period of data-taking after the upgrade is called *Phase-II* or *High Luminosity LHC* (HL-LHC) and will start at the end of 2026. A detailed overview of HL-LHC time-schedule, technical aspects and motivations can be found in [43], [44] and [45]. The final goal of the project is to obtain a total integrated luminosity of $\sim 3000\,\mathrm{fb}^{-1}$ to highly increase the signal statistics necessary to deepen the searches for new physics and precision measurements. As a direct consequence of the luminosity increase, all the LHC detectors will have to face a higher total amount of irradiation and particle multiplicity. During LS2 all the LHC experiments will be upgraded as well, to be able to withstand the harsher environmental conditions while at the same time maintaining an optimal detecting performance. For the ATLAS Experiment the upgrade will mostly involve the whole Inner Detector (ID) - that will be completely replaced by a *Inner Tracker* (ITk) - and the TDAQ system. This chapter will present a short overview of ITk, focusing on its *Third Generation Pixel Detector* and comparing it to the current ATLAS Pixel Detector.

4.1 ITk overview

The ITk final design has not been completed yet and studies on different possible solutions are still ongoing. Being the closest ATLAS sub-detector to the beam-pipe, ITk will have to cope with a very extreme environment. In particular, the innermost pixel layer will be subjected to a total fluence of about $2{\times}10^{16}$ $\mathrm{n}_{eq}/\mathrm{cm}^2$ and the expected pileup at HL-LHC luminosity will reach ~ 200 events/bunch crossing corresponding to an unprecedented data-rate. Notwithstanding the critical conditions

that it will have to face, ITk must meet all the requirements and performance needed to reach the physics goals set by the ATLAS Experiment. The design and optimization of the detector are finalized to fulfill all the Phase-II requirements as outlined by the *ATLAS Phase-II Letter of Intent* [46] and by the *ATLAS Phase-II Upgrade Scoping Document* [47]. In particular, ITk must be able to cover a great portion of the solid angle, to provide robust tracking even in presence of sensor inefficiencies or module failures and it should minimize the material budget and consequently the stopping power of the detector. Another important factor that must be taken in consideration during the design of the tracker is the CPU time needed for reconstruction; in fact, this is one of the cost drivers of the experiment and it should be reduced to the lowest amount possible while at the same time keeping the same precision and detection efficiency.

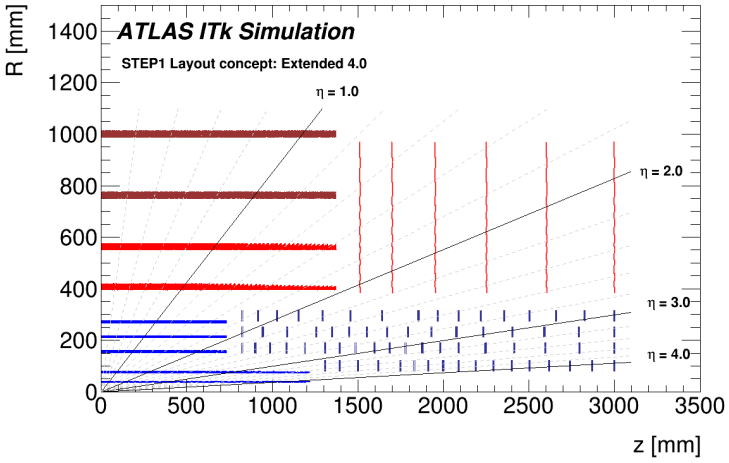
4.1.1 ITk Layout

Differently from the current ATLAS Inner Detector, ITk will be a full-silicon detector, composed by a Pixel Detector surrounded by a Strip Detector. The two sub-detectors will be separated by a *Pixel Support Tube* (PST). The geometry of the detector is still under investigation and its design is driven by the need of having a pseudorapidity coverage up to $|\eta| \leq 4$ and an optimal tracking efficiency with a minimal amount of *fake tracks* identification. After a continuous design improvement, the number of hits needed for efficient track reconstruction was reduced from 14 to 9. The set of cuts that must be applied for track reconstruction depending on pseudorapidity intervals is shown in Table 4.1 and drives an ITk geometry consisting of five layers Pixel Detector, four barrels Strip Detector and six End-Cap disks.

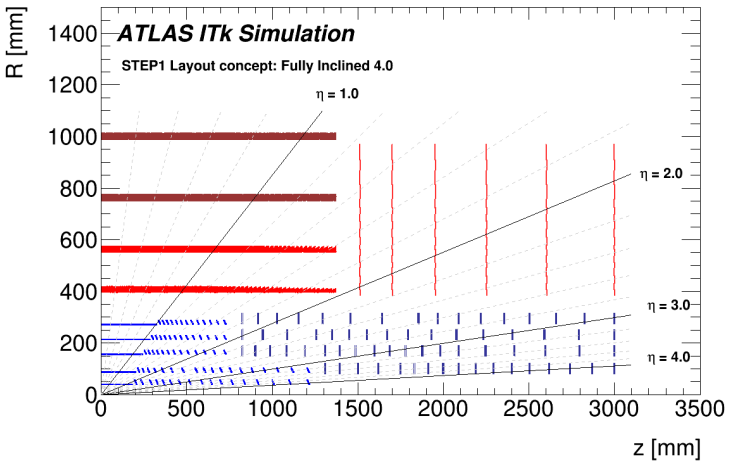
Table 4.1: Set of cuts applied during the track reconstruction depending on the pseudorapidity interval. Holes are counted if track candidates cross active sensors on which no hit was found, double holes are two consecutive active sensors crossed without a hit found. Here d_0 and z_0 are the radial and axial distance defined with respect to the mean position of the beam pipe. [48]

Requirement	Pseudorapidity interval		
	$\|\eta\| < 2.0$	$2.0 < \|\eta\| < 2.6$	$2.6 < \|\eta\| < 4.0$
Pixel + Strip hits	≥ 9	≥ 8	≥ 7
Pixel hits	≥ 1	≥ 1	≥ 1
Holes	< 2	< 2	< 2
Double holes	≤ 1	≤ 1	≤ 1
Pixel holes	< 2	< 2	< 2
Strip holes	< 2	< 2	< 2
p_T [MeV]	> 900	> 400	> 400
d_0	$\leq 2\,\text{mm}$	$\leq 2\,\text{mm}$	$\leq 10\,\text{mm}$
z_0	$\leq 20\,\text{cm}$	$\leq 20\,\text{cm}$	$\leq 20\,\text{cm}$

Thanks to the experience gained with the Phase-0 and Phase-I Pixel Detector, the two innermost layers of ITk will be removable to avoid mechanical problems in case of need of replacement due to radiation damage. To achieve the $|\eta| < 4$ pseudorapidity coverage, two concurrent designs are currently considered: *Extended Layout* and *Inclined Layout*, shown in Fig. 4.1.



(a) Extended Layout

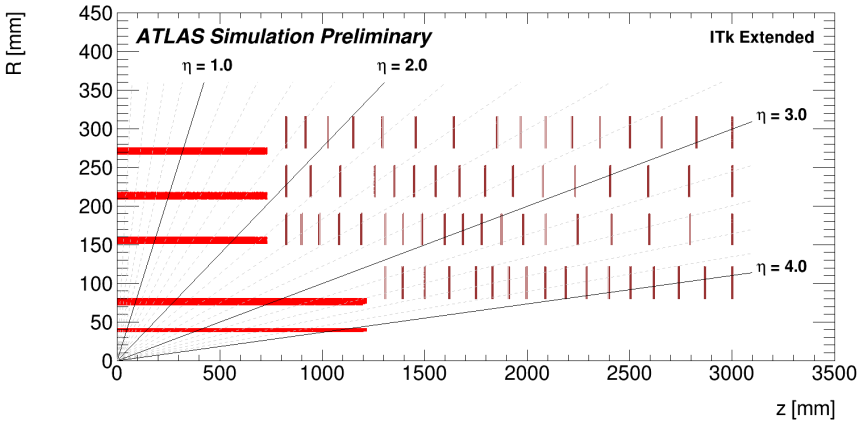


(b) Inclined Layout

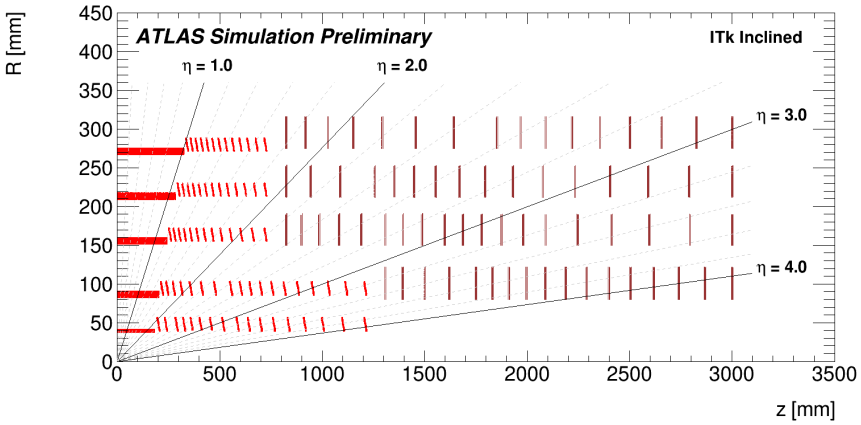
Figure 4.1: Two possible layouts for the ITk. In each picture a quarter of the detector is represented where the strips are shown in red and the pixel detectors in blue. In (a) the extended layout with a pseudo-rapidity coverage of the barrel up to $|\eta| = 4$ is shown. In (b) the fully inclined layout is shown. The pseudorapidity coverage of this layout is also up to $|\eta| = 4$. [ATL-PHYS-PUB-2016-025]

Extended Layout and Inclined Layout comparison

The ITk Extended Layout proposed geometry is similar to the current Pixel Detector layout, composed by *flat* staves with the two innermost pixel layers extended to allow a pseudorapidity coverage of respectively $|\eta| < 3.2$ and $|\eta| < 4$ [49]. The two extended pixel layers together with the rest of the pixel barrel in this layout are shown if Fig. 4.2 (a). This configuration presents two major drawbacks: the occupancy due to fake tracks and pileup is higher and the material budget is large in the forward regions (Fig. 4.3 (a)).



(a) Extended Layout

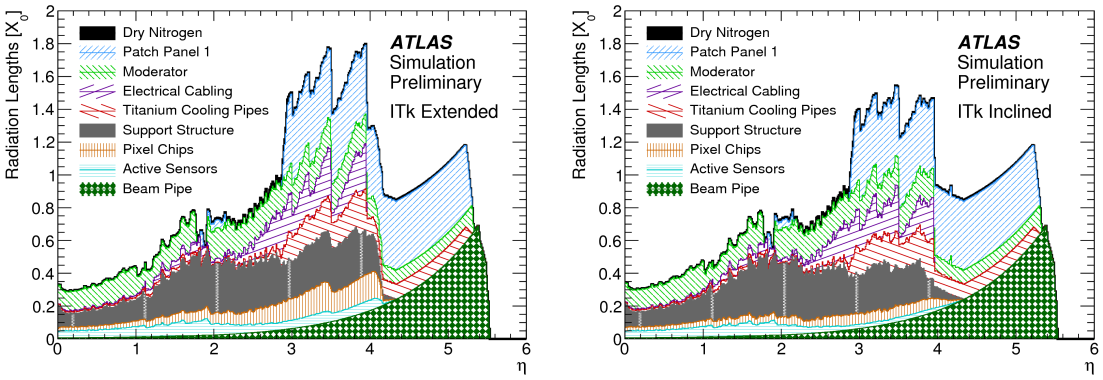


(b) Inclined Layout

Figure 4.2: Diagrams showing simulated energy deposits in active layers for the ITk Extended Layout (a) and Inclined Layout (b) zoomed in on the Pixel barrel. [ATL-PHYS-PUB-2016-025]

The idea behind the Inclined Layout is to have a geometry such that the particles at high pseudorapidity will cross the detector at similar incidental angle with respect to the particles at lower $|\eta|$. This is obtained by tilting the modules sited towards the end of the staves. Such a layout offers some advantages because it reduces the

overall occupancy and the material budget (Fig. 4.3 (b)) and it increases the number of points per particle track close to the interaction point in the forward region. The main drawbacks of this design are that the cooling is particularly challenging due to the larger distance of the cooling pipes from the sensors and the complicated geometry requires more computing time for track reconstruction. A comparison between the material budget and mean expected occupancy at 200 $\langle\mu\rangle$ pileup is given respectively in Fig. 4.3 and Table 4.2. As already discussed in Chapter 3, higher occupancy leads to higher throughput, which represent a challenge for the readout electronics. Therefore, a system featuring lower occupancy is preferable when possible.



(a) Extended Layout (b) Inclined Layout

Figure 4.3: Composition of simulated detector material in radiation lengths, shown as a function of η for a simulated ITk with either an (a) Extended or (b) Inclined Pixel barrel. [ATL-PHYS-PUB-2016-025]

Table 4.2: Average hits per chip per event for $50\times50\,\mu\text{m}^2$ pixels using $t\bar{t}$ events with 200 pileup. Listed are results layer by layer for the flat and inclined barrel regions and the end-caps. [48]

Layer (Ring)	Flat Barrel	Inclined Barrel	End-cap
0	223.0	136.7	80.9
1	26.6	27.8	37.7
2	19.3	20.1	21.0
3	12.9	12.7	13.3
4	9.9	9.1	9.3

4.2 ITk Pixel Detector

The design of the ITk Pixel Detector benefits from both the experience gained during Run-I and Run-II and from the new technological progresses. The composition of

the sensors and the front-end chips is chosen taking in consideration three main factors: overall cost, capability to withstand high dose of radiation and very high detection efficiency. Table 4.3 shows the required radiation tolerance for the sensors. ITk Pixel Detector will be composed of five layers, with the innermost two made replaceable; the exact positioning and geometry of each single layer will depend on the adopted design and on the type of sensors and module used. This section will present a short overview of the technologies and some strategical solutions that are taken in consideration for the final ITk implementation.

Table 4.3: The maximal 1 MeV neutron equivalent fluences and total ionising dose for different parts of the Pixel Detector, for the baseline replacement scenario for the inner section. All valus have been multiplied by a safety factor of 1.5. [48]

Luminosity	Layer	Location	R (cm)	z (cm)	Fluence ($10^{14} n_{eq}/cm^2$)	Dose (MGy)
2000 fb^{-1}	0	flat barrel	3.9	0	131	7.2
		inclined barrel	3.7	25.9	123	9.9
		end-cap	5.1	123.8	68	6.3
2000 fb^{-1}	1	flat barrel	9.9	24.3	27	1.5
		inclined barrel	8.1	110.0	35	2.9
		end-cap	7.9	299.2	38	3.2
4000 fb^{-1}	2-4	flat barrel	16.0	44.6	28	1.6
		inclined barrel	15.6	110.0	30	2.0
		end-cap	15.3	299.2	38	3.5

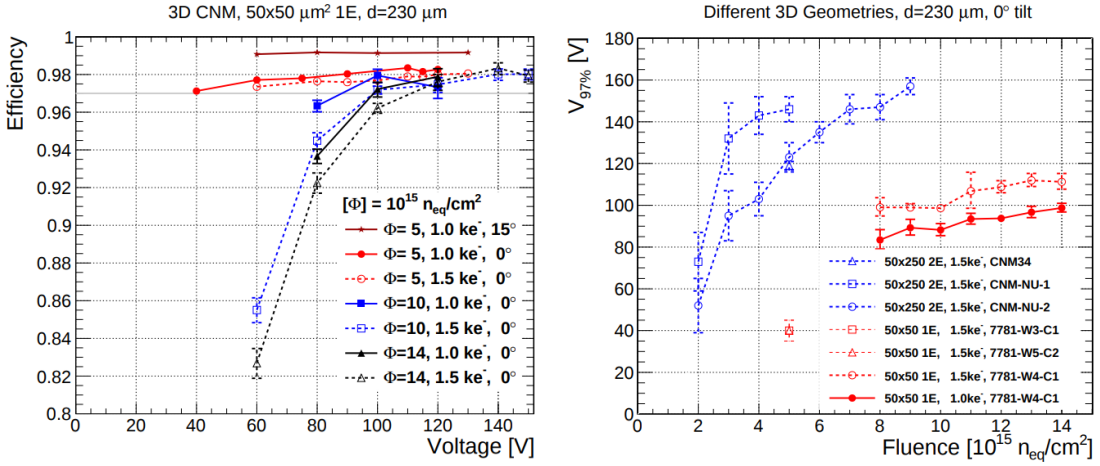
4.2.1 Sensors

The choice of the sensors for ITk is dictated by the radiation hardness requirements; the fluences affecting the detector are described in Tab. 4.3 and are maximized in Layer-0, where the total dose will be 7.2 MGy after 2000 $fb^{-}1$ of integrated luminosity. At the same time, the detector must be able to maintain a random single-pixel hit inefficiency $< 3\%$ and to minimize the leakage current and hence the power consumption. Pixel sizes of $50 \times 50 \mu m^2$ and $125 \times 25 \mu m^2$ are exploited. Beyond the planar and 3D sensors, already used in the current Pixel Detector and IBL, another sensor technology is being investigated: *High Voltage CMOS* (HV-CMOS).

3D sensors

3D sensors are currently used in the sides module of IBL staves. As described in Section 2.3.1, they offer two great advantages: lower power dissipation due to leakage current and higher radiation tolerance. In fact, the depletion region dimensions depends on the distance between the implanted electrodes and it is decoupled from the sensor thickness. The voltage required to "fully deplete" 3D pixels is also relatively

small; Fig. 4.4 (a) shows that less than 130 V are needed to reach a hit efficiency of more than 97% with perpendicular incident particles, even after irradiation up to $9 \times 10^{15}\, n_{eq} cm^{-2}$. ITk sensors, whose size is $50 \times 50\, \mu m^2$, show higher radiation tolerance and lower voltage compared to IBL 3Ds, whose size is $250 \times 50\, \mu m^2$, as illustrated in Fig. 4.4 (b). The main disadvantages of 3D sensors is that they are generally noisier and more expensive, since the process of implanting columnar electrodes is complicated. Thanks to their properties, 3D pixels are the best candidate for the innermost two layers of the ITk pixel detector, which will be the most exposed to radiation and will cover the least surface.



(a) (b)

Figure 4.4: (a) Hit efficiency as a function of the bias voltage of 3D pixel sensors irradiated with protons for different fluences, thresholds and tilts. The solid black line represents the target efficiency of 97%. (b) $V_{97\%}$ as a function of fluence comparison between IBL generation 3D sensors ($250 \times 50\, \mu m^2$ size, 2E) and ITk 3D sensors ($50 \times 50 \mu m^2$ size, 1 E). [48]

Planar sensors

Planar silicon technologies have been used in pixel detectors for many years, which ensued in a mostly consolidated design during this time. The technology chosen for ITk planar pixels is n-in-p, since only one side of the wafer have to be processed, differently from the n-in-n technology used in the current ATLAS Pixel Detector and IBL. The active thickness of the sensors depends on the required radiation hardness and must be optimized to maximize hit efficiency and minimize power dissipation. For the innermost two layers, that will be subjected to the highest radiation dose, the best performance is obtained by using $100\, \mu m$ thickness, while a $150\, \mu m$ thickness is foreseen for the outer layers. Fig. 4.5 compares the planar efficiencies for different pixel thicknesses after a fluence of $1 \times 10^{16}\, n_{eq}/cm^2$.

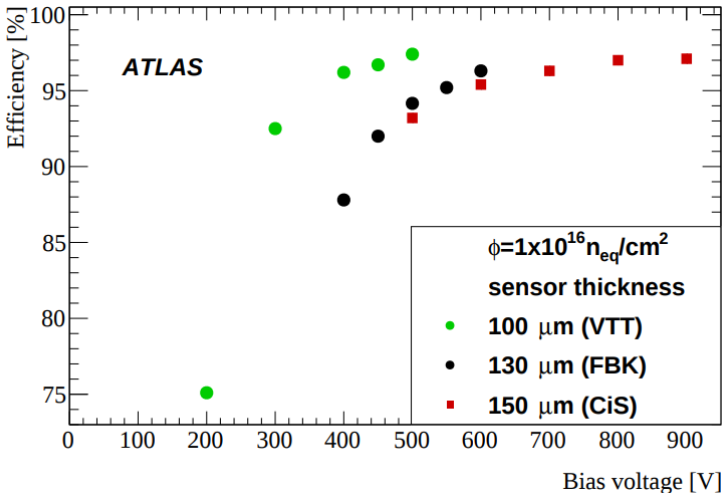


Figure 4.5: Hit efficiency measured at normal incidence in test beams at DESY and CERN SPS with FE-I4 modules assembled with 100, 130 and 150 μm thin planar sensors after a fluence of $1 \times 10^{16}\, n_{eq}/cm^2$. The sensors were designed in different laboratories: the VTT Technical Research Centre of Finland, the Fondazione Bruno Kessel (FBK) in Italy and the CiS laboratory in India.

HV-CMOS sensors

The idea behind High Voltage CMOS (HV-CMOS) sensors is to implement pixel sensors into active CMOS components. The structure of this device, illustrated in Fig. 4.6, consists in a n-well/p-substrate sensor diode, which is depleted with high voltages to facilitate charge collection. The active electronics, typically consisting in a charge sensitive amplifier with continuous reset, is made with transistors placed inside the n-well, which is at the same time used as a charge collection electrode. The main advantage of the HV-CMOS technology is that the cost for production and connection to the front-end chip is sensibly lower; for this reason it is a candidate for the outermost layer of the Pixel Detector, the one covering the largest surface. Recent results proved that the HV-CMOS devices have a good resistance to radiation, since an efficiency of more than 99% has been measured with unirradiated chips and 95% with the chips irradiated to $10^{15}\, n_{eq}/\mathrm{cm}^2$ with neutrons. However, time resolution is still far from meeting the ATLAS requirements, since the average time required to collect 95% of the signal is 100 ns, four times higher than the 25 ns bunch crossing interval.

4.2.2 Front-end electronics: RD53

ITk Pixel Detector requirements are very strict in terms of radiation hardness, amount of data to be handled and trigger rate so that the current IBL front-end chip, FE-I4, is incapable to meet all the restrictions. Therefore a new collaboration - called *RD53 Collaboration* - was born with the purpose of designing a new front-end

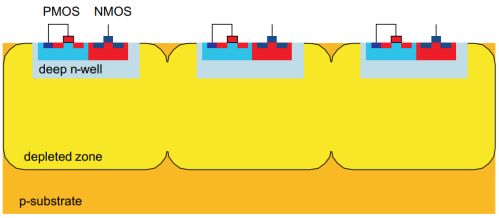


Figure 4.6: HV-CMOS sensor structure. [50]

chip (named RD53) able to withstand the HL-LHC environmental conditions and to meet the ATLAS Detector physics goals. The requirements for the front-end chips are summarized in Table 4.4. One of the major differences between Phase-II and Phase-I ATLAS is that the trigger rate will be raised from $\sim 100\,\text{kHz}$ to $\sim 1\,\text{MHz}$, causing the increase of the detectors throughput. To be able to provide enough bandwidth and to meet all the requirements imposed by Phase-II ATLAS Experiment, RD53 will use 65 nm CMOS technology, smaller than the 250 nm and 130 nm technologies of respectively FE-I3 and FE-I4. The chip dimensions will be approximately $20\,\text{cm} \times 21\,\text{cm}$ hosting 153600 pixels divided in 400 rows and 384 columns. The 65 nm technology is needed because it is more radiation hard and at the same time it allows higher integrability needed to fit all the 500 million transistors present in the RD53. The chip architecture and its layout organization are shown in Fig. 4.7.

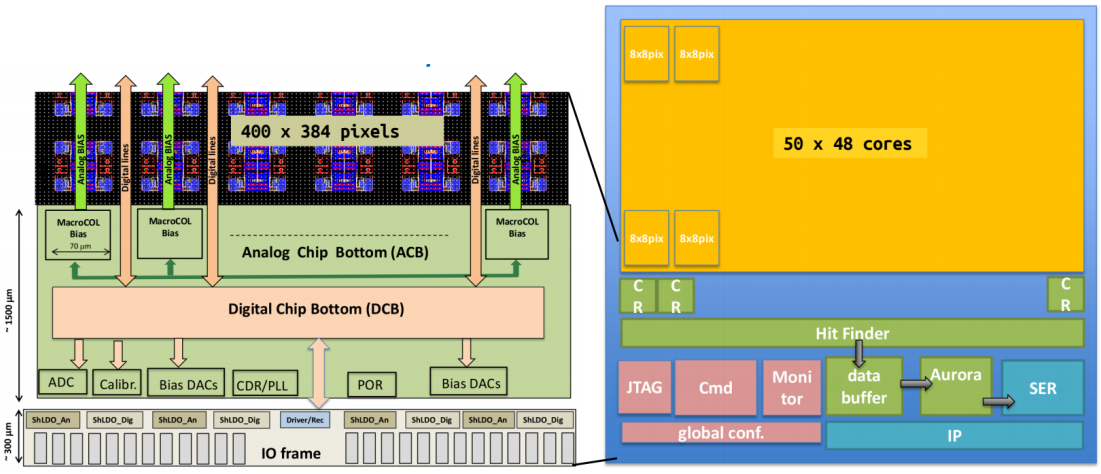


Figure 4.7: Diagram of the production read-out RD53 chip. The right side shows the digital hierarchy while the left side shows the top level organization including analog circuits. [48]

Since the same front-end chip will be used for the innermost layer - defining the most critical conditions - and all the outer layers - generally presenting a more *relaxed* environment, the internal settings of the RD53 are made programmable, so that they can be optimized to reduce power consumption for each single layer.

Table 4.4: Physical, power, environmental and performance requirements for the ITk Pixel Detector Front-End chip. [48]

Requirement	**Value**	**Source**
Matrix size columns × rows	400×384	Layout, IC design constraints
Input pitch	$50\,\mu m \times 50\,\mu m$	Tracking performance, occupancy
Final thickness	$150\ \mu m$	Material, bump bonding, local supports
Power dissipation	$< 0.7\,W/cm^2$	Cooling, local supports, sensor
Current consumption	$< 1.5\,A$/chip	Services
Temperature range	$-40°C$ to $40°C$	Cooling, operation, integration
Total ionizing dose tolerance	$\geq 500\,Mrad$	Inner layer lifetime before replacement
Full chip SEU upset probability	5% per hour	Inner layer operation
Pixel loss due to SEU	$< 1\%$ per run	Hit efficiency, inner layer operation
Trigger rate	1 MHz (4 MHz)	Trigger and DAQ
Trigger latency	$< 35\mu s$	Trigger and DAQ
Trigger Protocol	Tagged trigger	DAQ, operation
Single Pixel Noise (ENC)	$< 100\,e^-$	Threshold, resolution
Threshold dispersion after tuning	$40\ e^-$	Uniformity, efficiency
Threshold variation	$< 10\%/K$	Operation
Noise occupancy per pixel	$< 10^{-6}$	Tracking performance
Hit loss at 75 kHz hit rate	$\leq 1\%$	Efficiency for inner layer
Recovery from saturation	$< 1\,\mu s$	Efficiency, pileup
Charge measurement resolution	$< 600\,e^-$	Tracking performance

Analog circuits

The analog part of the front-end consists of a two stage amplifier and a comparator and is shown if Fig. 4.8. The first stage pre-amplifier is a cascode circuit with NMOS input transistor in weak inversion. It contains a leakage current compensation system that is meant to cancel DC leakage current in the sensor. The shape of the signal, in particular the trailing edge, is defined by a continuous current reset and must be optimized to allow the full signal digitalization and at to keep at the same time in-pixel pileup below 1% hit loss. The reset and feedback current, as well as the threshold voltage for the comparator, are provided by programmable ADC implemented in the digital core of the chip.

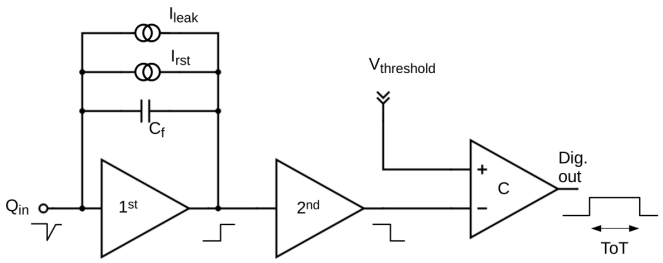


Figure 4.8: Simplified diagram of RD53 analog part consisting in a two stage cascode amplifier and a comparator. [48]

The analog circuitry is implemented in 2×2 analog pixel units (called analog island) surrounded by digital circuitry. This layout is called *analog island in a digital sea*, shown in Fig. 4.9, and was chosen to optimize the routing and placement of the logic area.

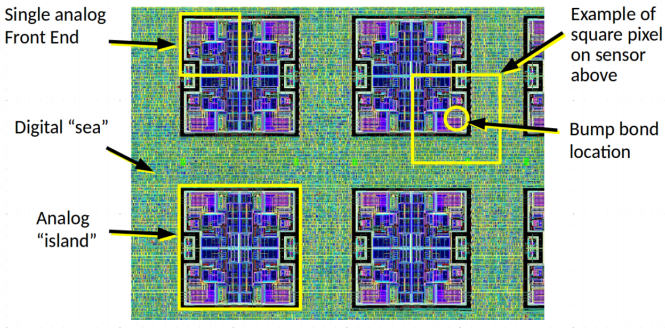


Figure 4.9: Layout detail from RD53A illustrating the concept of islands of analog circuitry (blue) embedded in a *digital sea* of synthesized logic (green). [48]

Digital matrix

The pixel matrix is composed of *digital cores* of 64 pixels (8 rows by 8 columns) in charge of all the digital processing of the pixel output, timing, triggering and readout. They are also responsible of providing configuration bits to the analog islands and of receiving the outputs from each island. The purpose of the digital core architecture is to solve the problems of the column driven architecture featured by FE-I3. Each core is divided into *regions* comprising multiple pixels; the timing and hit informations are stored by region, so the pixels in a region can be seen as a unique *macro-pixel* with an internal structure. Even if all the regions inside a core have the same functionality, their layout is different because all the digital parts are synthesized in one flat layout (digital sea). There are two region architectures, called *Distributed Buffer Architecture* (DBA) and *Central Buffer Architecture* (CBA). The DBA size is four pixel, spanning over one row and four columns and covering an

effective area of 50 μm$\times$200 μm. The ToT information is stored in a four-bit counter inside each pixel (explaining the name *distributed*). The CBA size is sixteen pixels, spanning over four rows and four columns. The ToT information is stored in a common memory space in the region, allowing ToT zero-suppression but requiring a hit map.

Analog and Digital Chip Bottom

The *Chip Bottom* logic, shown in Fig. 4.10, contains all the blocks implementing control and processing functionalities. In particular, it handles the command reception and propagation to the pixel matrix, the clock distribution, the trigger handling and data transmission. The main and most important differences between RD53 and its predecessors logic concern the configuration modality, the output protocol and the available bandwidth.

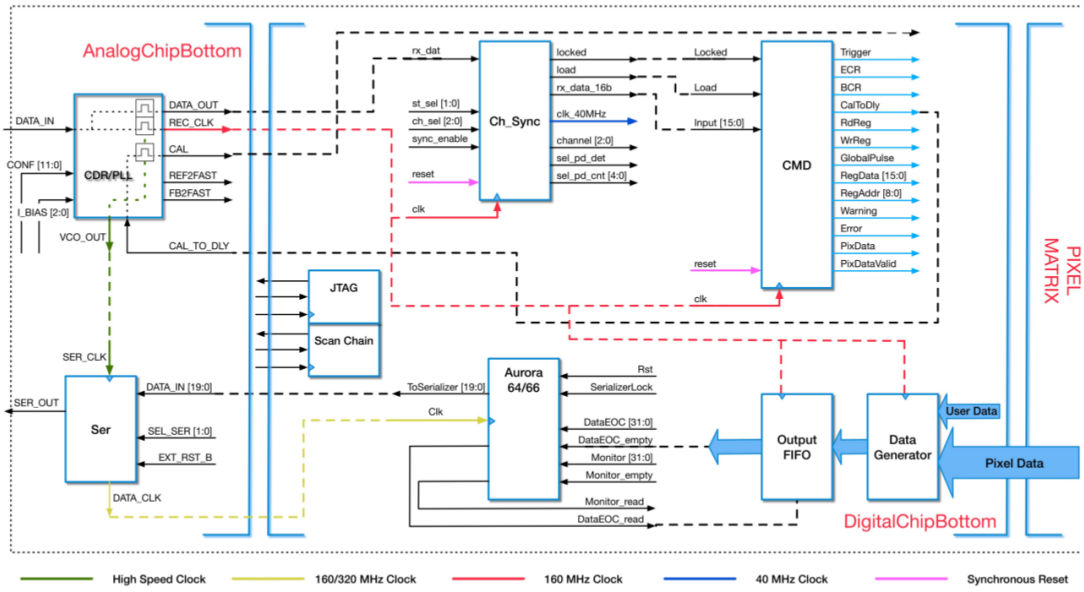


Figure 4.10: Block diagram of digital bottom of chip showing the command input and data output paths. [51]

The configuration protocol uses a custom DC-balanced code at 160 Mbps; the clock is recovered from the serial line by a Phase Locked Loop (PLL) in the chip. Each command consists in a frame of 16 bits which are transmitted serially to the RD53, meaning that 100 ns, or four bunch crossing periods, are required to fully transmit a command frame. The trigger commands must specify which of the 4 bunch crossings were triggered; all the fifteen possible configuration are specified and are coded in a *Trigger Table* shown in Table 4.5. A 5-bits trigger identifier, called *trigger tag*, completes the trigger command. While the FE-I4 required at least five trigger-less bunch crossings between two consecutive triggers, RD53 allows

a continuous triggering with no *dead time*. The real innovation with respect to the previous front-end devices is that the configuration and trigger commands are time-multiplexed in a so-called *trickle configuration*. This way it is possible to continuously send the chip reconfiguration - even during data-taking operations - to reduce Single Event Upset (SEU) effects.

Table 4.5: List of trigger command and symbols used to encode the 15 possible trigger patterns spanning four bunch crossings. Note there is no 0000 pattern as that is the absence of an trigger. [51]

Trigger Pattern	**Encoding**	**Trigger Pattern**	**Encoding**
		T000	0011_1010
000T	0010_1011	T00T	0011_1100
00T0	0010_1101	T0T0	0100_1011
00TT	0010_1110	T0TT	0100_1101
0T00	0011_0011	TT00	0100_1110
0T0T	0011_0101	TT0T	0101_0011
0TT0	0011_0110	TTT0	0101_0101
0TTT	0011_1001	TTTT	0101_0110

The output bandwidth of the RD53 has been increased from 160 Mbps (MCC and FE-I4) to 5.12 Gbps (split on four lanes) to cope with the increased amount of data and trigger rate of HL-LHC. The protocol chosen for data transmission is Aurora 64b/66b (B.2) split over four lanes, which is DC-balanced but more efficient than the 8b/10b used by FE-I4.

4.2.3 Modules

The modules are the basic component of the ITk pixel detector structure. The final design is still under development, and three different module layouts are considered. The layouts are called *singles*, *doubles* and *quads* and can interface respectively one, two or four front-end chips (and hence different sensor area). Tables 4.6, 4.7 and 4.8 summarizes the proposed layout parameters for pixel flat barrel, pixel inclined barrel and pixel end-caps.

4.3 Physics performance

The goal of the ITk detector is to fulfill the physics requirement for Phase-II ATLAS and its layout is optimized to identify charged particles and measure their properties with high efficiency and purity. The new detector must be able to preserve, and possibly exceed, the current ATLAS Inner Detector performance, even if the environmental conditions are more critical. The ITk tracking efficiency, fake rates, track parameter resolution, robustness of tracking and primary vertex reconstruction or

Table 4.6: Main layout parameters for the pixel flat barrel. The number of sensors per row refers to a half row (z> 0 mm) in the central, flat part of the barrel where sensors are placed parallel to the beam line. The number of hits indicates how many hits are expected in the layer for particles originating from z= 15 cm. The total length in z of thhe inner layer and outer barrel sections is 110 cm. [48]

Layer	**Radius [mm]**	**Rows of sensors**	**Sensors per row**	Type	Hits
0	39	16	6	duals	1
1	99	20	6	quads	1
2	160	30	11	quads	1
3	220	40	12	quads	1
4	279	50	13	quads	1

Table 4.7: Main layout parameters for the pixel inclined section. [48]

Layer	**Radius [mm]**	**Rows of sensors**	**Sensors per row**	Type	Hits
0	36	16	6	singles	2-3
1	80	13	6	quads	2-3
2	155	11	11	duals	1
3	215	13	12	duals	1
4	274	13	13	duals	1

Table 4.8: Main layout parameters for the pixel end-caps. The radii refer to the innermost point of sensors on a ring. The number of hits indicates how many hits are expected in the layer for particles originating from z= 15 cm. [48]

EC Layer	**Radius [mm]**	**Rings**	**Sensors per ring**	Type	Hits
0	50	4	16	quads	3
1	78	11	22	quads	3-4
2	152	10	32	quads	2
3	211	8	44	quads	1
4	271	9	52	quads	1

b-tagging capability have been simulated for different configurations using Monte Carlo simulated events; the optimal detector configuration is chosen accordingly to be best results, that will be compared to the current ID Run-2 ones.

Robustness against aging and component failures

ITk must be able to provide reliable results even in case of aging of the detector or component failures. In general, the trackers can present two types of defect:

- inactive modules caused by component failures during operations, usually referred as *known detector inefficiencies*;
- single channels afflicted by radiation damage causing loss of hits, usually referred as *unknown inefficiencies*.

The known inefficiencies are typically easier to treat; according to the ITk approach, when a track candidate crosses an inactive module, the missed measurement is not treated as a *hole*, but as a valid hit. This choice was adopted to avoid the penalization of area in the detectors where several inactive modules are present. The effects of unknown inefficiencies are more difficult to compensate, and can be simulated by introducing some random hit loss in the detector.

4.3.1 Tracking efficiency

The tracking efficiency ϵ is defined as the number of tracks reconstructed from the detector data divided by the real number of tracks and can be calculated using simulated data. The fake rate is defined as the rate of *fake tracks* reconstructed inside the detector, i.e. how many times randomly placed pixels fired by independent particles are reconstructed as a track.

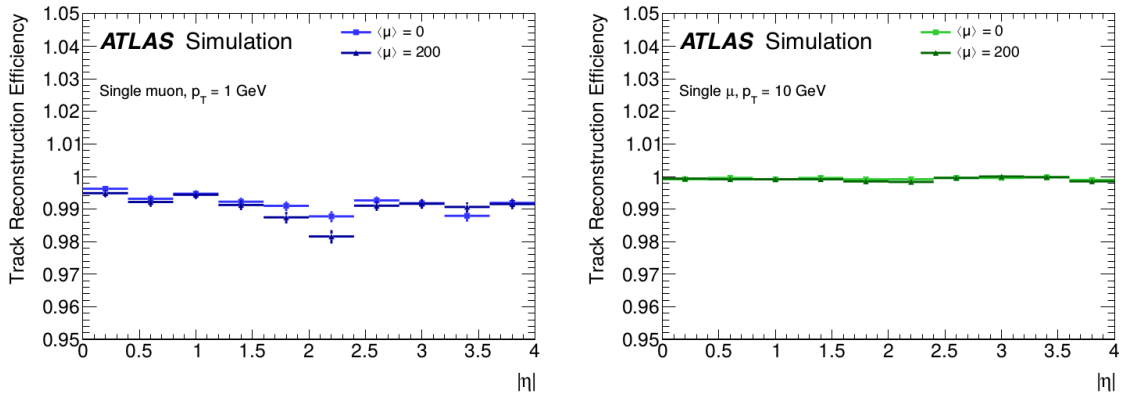
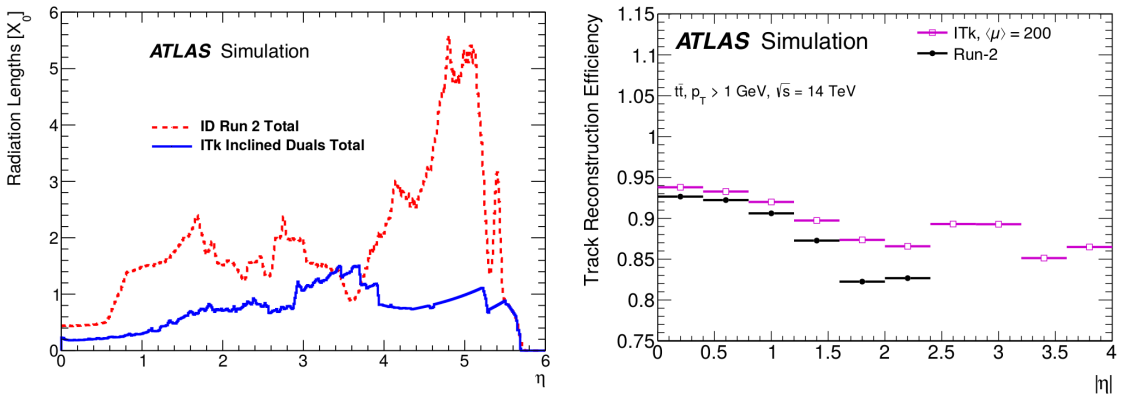


Figure 4.11: Tracking efficiency for muons without pile-up ($\langle\mu\rangle = 0$) and with an average of 200 pile-up events ($\langle\mu\rangle = 200$). Left: for $p_T = 1\,\text{GeV}$ muons. Right: for $p_T = 10\,\text{GeV}$ muons. [48]

The tracking efficiency is strictly influenced by the type of particle crossing the detector and its energy deposition mechanisms. A first validation of the detector geometry can be obtained by measuring (or simulating) muon track efficiency; in fact, muons are not influenced by hadronic interactions and the detection efficiency for muons over the entire pseudorapidity coverage should be close to 100%. Fig. 4.11 shows a simulation of ITk tracking efficiency for single muons (no pile-up) and with an average pile-up of 200 events. The tracking efficiency loss for high pile-up is very small, proving that the detector layout is well designed and provides enough resolution to resolve multiple tracks. The reconstruction efficiency for pions and electrons, unlike muons, does not depends only on the geometry and granularity of the detector. In fact they interact also with the inactive materials (support tubes, colling, electronics) leading to inelastic hadronic interactions or Bremsstrahlung. This effect can be mitigated by reducing the material budget of the detector. Fig. 4.12 shows that, compared to the current ATLAS Inner Detector, the ITk material budget has been reduced, resulting in an overall tracking efficiency increase.



(a) (b)

Figure 4.12: (a) Comparison of radiation length X_0 versus η between the current ATLAS Inner Detector and ITk. (b) Comparison of track reconstruction efficiency for a top-pair sample with an average of 200 pile-up events between the current Run-2 detector and ITk. [48]

4.3.2 Resolution and primary vertex reconstruction

The main goal of ITk is to provide precise measurements of the longitudinal (z_0) and transverse (d_0) impact parameters of tracks, as well as polar (Θ) and azimuthal (ϕ) angles and transverse momentum (p_T) from the tracks curvature. The ITk primary vertex reconstruction, b-tagging and lepton/quark jets reconstruction capabilities depend directly from the measurements resolution. Fig. 4.14 shows a comparison between the simulated track parameters resolution of ITk and the current Inner Detector one. For a pseudorapidity $|\eta| < 3.0$ the estimated intrinsic d_0 resolution is

10 μm, which is slightly worse than the one of the current Pixel Detector because the current first layer of the Pixel Detector (IBL) is placed at a smaller radius than ITk Layer-0. On the other side, the ITk longitudinal resolution z_0 is better, because the sensor pitch along the z axis has been reduced from 250 μm to 50 μm. All the impact parameter informations are used to reconstruct the primary vertexes, after applying some cuts: a minimal p_T of 0.9 GeV is required, each track must have at least three pixel clusters and the tracks must present a transverse resolution $\sigma(d_0) < 0.3$ mm and longitudinal resolution $\sigma(z_0) < 0.5$ mm. Fig. 4.13 shows the number of reconstructed primary vertexes as a function of pile-up for $t\bar{t}$ events with a pile-up $\langle\mu\rangle$ between 30 and 270. The results are then compared to simulated samples using the Run-II Pixel Detector and vertex reconstruction algorithms.

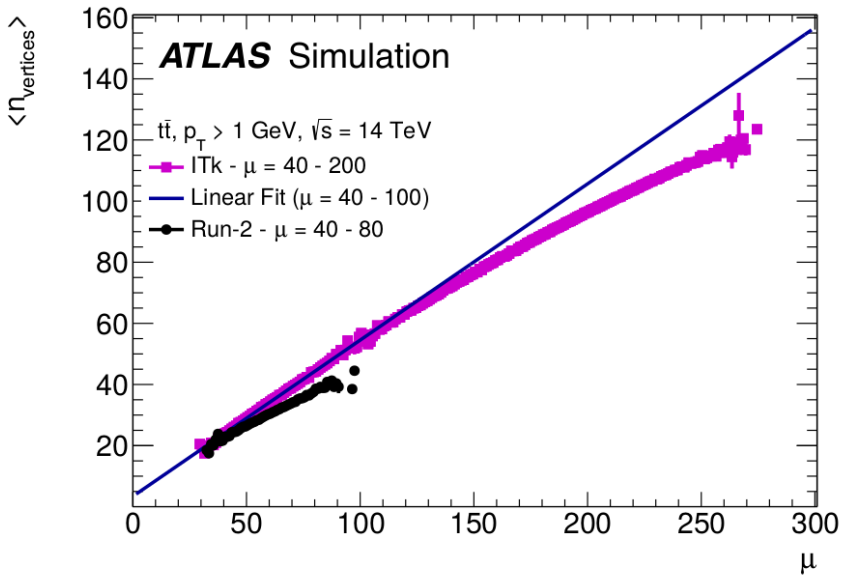


Figure 4.13: The number of reconstructed primary vertexes as a function of pile-up for $t\bar{t}$ events with a pile-up between 30 and 270. Results are shown for a pixel size of $50 \times 50\,\mu m^2$. As an illustration a linear fit to the number of vertexes for pile-up values between 40 and 100 is shown, extrapolated up to 270 pile-up. Results for a Run-II simulation sample using the Run-II primary vertex reconstruction code are shown as well. [48]

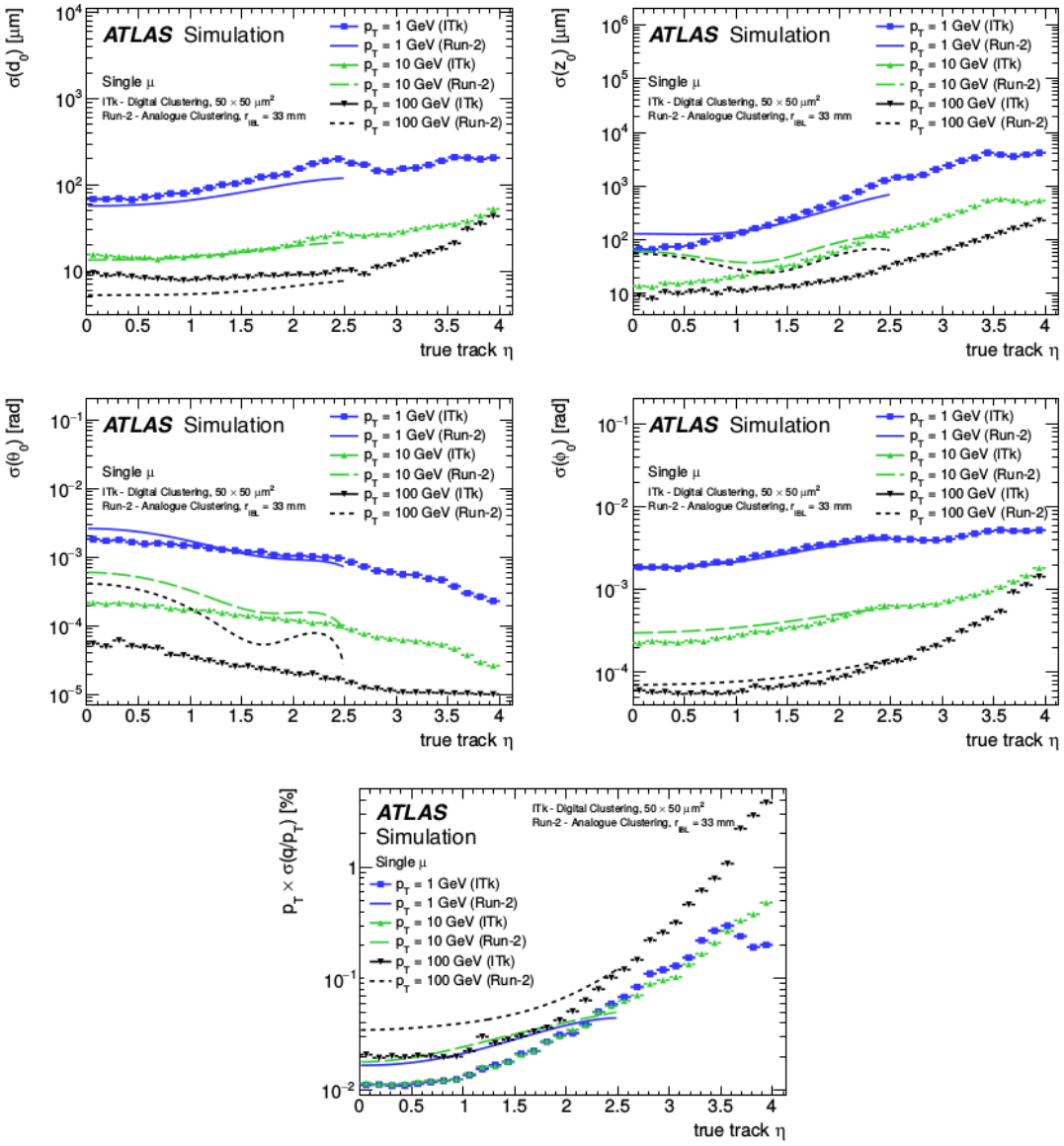


Figure 4.14: Track parameter resolution in d_0 , z_0 , θ, ϕ and p_T as a function of η for an ITk Pixel Detector with $50 \times 50\,\mu m^2$ pixels. Results are shown for single muons of 1, 10 and 100 GeV in p_T . The reconstruction uses digital clustering information. For comparison, the resolutions for the current Run 2 Inner Detector are also shown. [48]

Chapter 5

ATLAS Phase-II DAQ

Last chapter showed that High Luminosity LHC upgrade, that will be completed around 2026, will have a great impact over all the experiment at LHC, most of which will be completely revisited and upgraded. The key factors that will affect all the detectors are two: the increase of instantaneous luminosity - corresponding to an increase of the simultaneous collisions and hence of the amount of total data per time unit - and of the trigger rate, that will be on the order of 1 MHz, ten times higher than the current rate of ∼100 kHz. The combination of those two factors constitutes a major challenge for the electronic readout systems, since it directly effects the total throughput, i.e. the amount of data transmitted per time unit. Therefore, all the readout systems will have to provide a higher total bandwidth, capable of coping with the increased data throughput. The Phase-II ATLAS readout system will present several differences compared to the current one; although its complete layout has not been finalized, the basic structure has been outlined. The main component of the DAQ upgrade will be an electronic card called *Front-End LInk eXchange* (FELIX), that will be developed by the FELIX Collaboration. A first version of this card, called *FLX-712*, will already be used by some ATLAS sub-detectors during Run-III, while the Phase-II version is still under development. The University and INFN of Bologna also produced a readout board, called *PIxel detector high Luminosity UPgrade* (πLUP), that is a general purpose card that can play an important role in the DAQ chain, both in phase of prototyping or validation and in the final setup. This chapter will give a short overview on the general strategy behind the ATLAS Phase-II readout upgrade, introducing to the FELIX board and its results. Finally, the Bologna πLUP card will be presented, focusing on its main features and the qualities that led to the beginning of a collaboration with the FELIX Group.

5.1 Phase-II ATLAS DAQ System Overview

The fundamental structure of the Phase-II ATLAS readout chain, shown in Fig. 5.1, is similar to the current one and includes a two level trigger followed by an Event Filter. However, the technologies involved and the target rate are different and adapted

to the technological evolution of recent years. Level-0 trigger will reduce the trigger rate from 40 MHz to about 1 MHz using information from calorimeters and muon system. It will be completely implemented on hardware on a *Central Trigger Processor* that gathers all useful data and uses algorithms to take the trigger decision. Calorimeter informations are extracted by three boards called *Feature EXtractors* (FEX), whose role is to identify electromagnetic objects (*eFEX*), to search for jets and missing energy (*jFEX*) and to analyze boosted objects (*gFEX*).

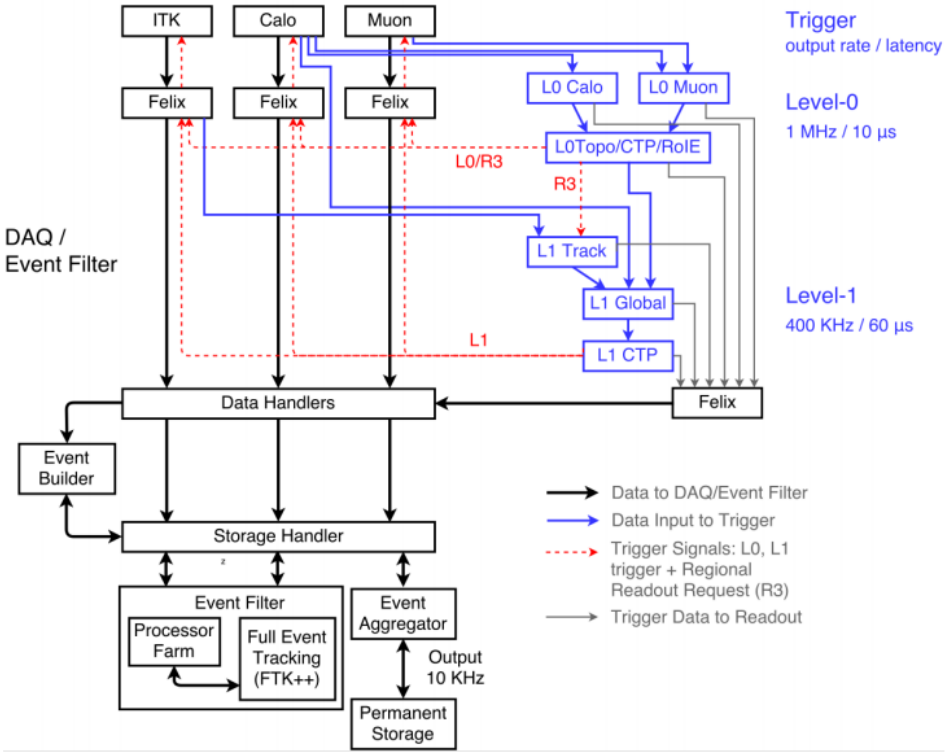
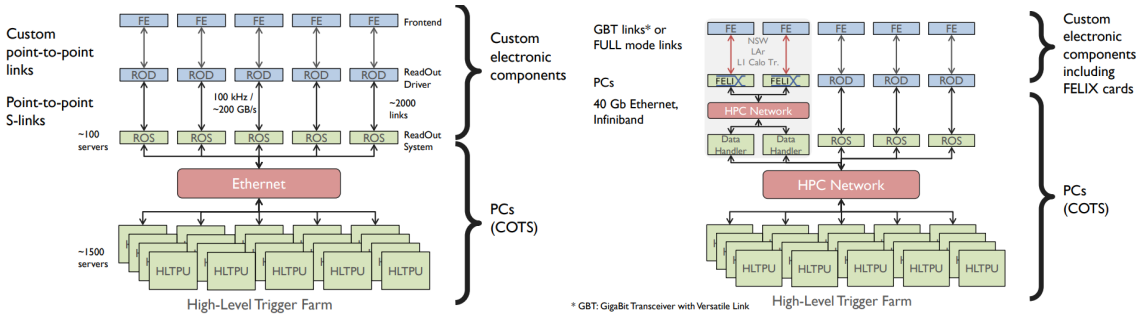


Figure 5.1: Schematic overview of the upgraded Trigger/DAQ system architecture. [52]

The second trigger stage, Level-1 trigger, will also be implemented on hardware and will further reduce the acquisition rate to 400 kHz. The trigger decision will be based on finer-grained calorimeter and tracking information using associative-memory chips for pattern recognition and FPGA-based χ^2 fitting. Finally, the Event Filter will use software algorithms to reach the storage data-rate of 10 kHz. It will likely be a heterogeneous system which can be implemented both in multi-core CPUs and GPUs. The Phase-II ATLAS *Fast TracKer* (FTK++) will provide information contributing to the final decision making.

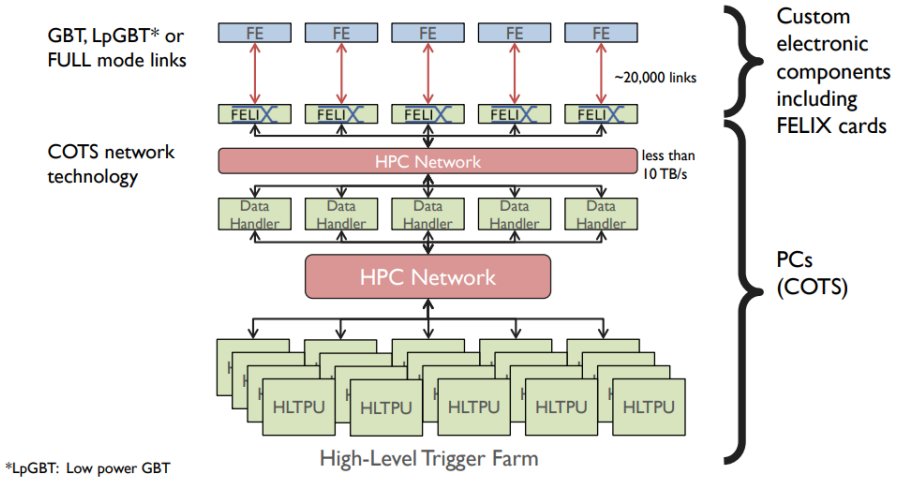
Another main innovations of the new readout system with respect to the current one, is the abolition of sub-detector specific Read-Out Drivers (RODs) in order to move towards a common board that will interface to all the front-ends, called *Front End LInk eXchange* or simply FELIX. This choice will greatly reduce the system complexity, as well as the amount of experts needed for support and bug-fixing of

the boards. A FELIX-based architecture will already be adopted by Run-III ATLAS DAQ chain. In fact three sub-detectors - New Small Wheel (SNW), Liquid Argon (LAr) calorimeter and calorimeter Level 1 trigger - will be upgraded together with their readout system, and their RODs will be substituted by FELIX boards. The evolution of the ATLAS DAQ system is schematically shown in Fig. 5.2.



(a) Current readout system

(b) Phase-I upgrade readout system



(c) Phase-II upgrade readout systemt

Figure 5.2: Evolution of ATLAS readout system. (a) Current DAQ architecture using exclusively custom sub-detectors Read-Out Drivers. (b) Phase-I upgrade hybrid architecture using Read-Out Drivers and FELIX boards. (c) Phase-II upgrade architecture using exclusively FELIX boards.

5.2 The FELIX Project

Chapter 3 showed that the current ATLAS DAQ system can be divided in two general categories: custom made boards/devices and commercially available ones. As shown in Fig. 3.1, the custom electronics involves the early stages of the chain, after which the data-stream is handled by commercial network routers and PCs. The problem of the current DAQ is that each sub-detector has its own Read-Out Driver

(ROD), a custom made board meant to interface a specific front-end to the PCs. Developing and maintaining a *large* number of custom devices proved to be difficult during Phase-0. The Phase-II upgrade will revolutionize the first stages of the ATLAS readout chain by introducing the FELIX board, that is still custom designed but will be used by all the front-ends. It will only act as an interface between the detector and the PCs, and all the front-end specific operations will be handled by the software framework (Software-ROD). The FELIX cards will be adopted already during Run-III (Phase-I), to readout three ATLAS sub-detectors that will be upgraded: the New Small Wheel (SMW), the Liquid Argon (LAr) calorimeter and the calorimeter Level-1 trigger. Run-III DAQ will hence be heterogeneous, with FELIX and RODs coexisting at the same time. This configuration will be a perfect testbench to fully exploit and validate the potentialities of the future readout upgrade. The Phase-I FELIX card is a generation-3 PCIExpress (PCIe) [53] card that can be inserted in a PCIe slot of a server PC. The bus is used for data-transfer from/to the PC, while the connection with the front-end is implemented on optical fibers through GBT (see Appendix B.3) or Full Mode (see Appendix B.4) protocols. In the first stages of the project, the board used for prototyping was an evaluation board from Xilinx: the VC-709. This card, called *Mini-FELIX*, features a 8 lanes gen-3 PCIe connector, four optical links connectors and a Virtex-7 [54], a 7^{th} series Xilinx FPGA. The final Phase-I FELIX design was completed by Brookhaven National Laboratory (BNL), which produced a card called *FLX-712*.

5.2.1 FLX-712

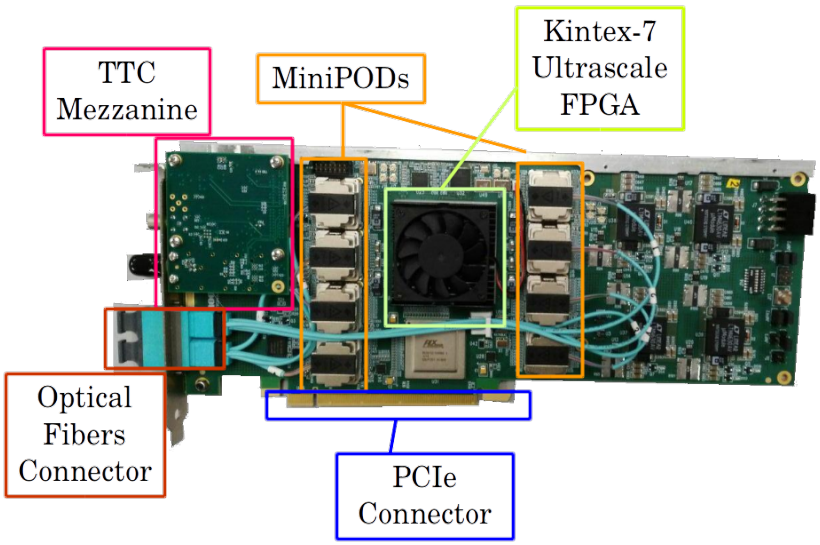


Figure 5.3: FLX-712 board. The main components are highlighted in the picture.

FLX-712 board, shown in Fig. 5.3, was developed by Brookhaven National Laboratory in collaboration with the FELIX Project to be used as the Phase-I FELIX

card. It is a standard height PCIe card featuring a Xilinx Kintex-7 Ultrascale FPGA [55], a 16 lanes gen.3 PCIe connector and two *Multi-fiber Push On* (MP0) connectors capable of connecting to 48 optical fibers each. The Kintex-7 Ultrascale FPGA presents 64 high speed serializers/deserializers called *GTH transceivers* [56], 16 of which are connected to the 16 lanes of the PCIe bus, and the other 48 to the optical fibers transmitters (TX) and receivers (RX). The electrical to optical and vice-versa conversion is done by four TX and four RX *MiniPOD* connectors on the board; each MiniPOD can handle 12 TX or RX fibers. Finally, two patch-cords connect four MiniPODs each (so 48 fibers) to two MPO couplers, placed at the extremity of the board, where the external fibers will be plugged.

The FELIX also includes a Timing and Trigger Control (TTC) mezzanine card plugged on the main board. It serves the scope of recovering, cleaning and propagating the clock, triggers and other signals generated by the TTC system. The FELIX board is provided with an integrated firmware and a software package that allow the user to configure the card, monitor its status and handle the I/O transmission.

5.2.2 FELIX firmware and software package

The FELIX firmware, whose block diagram is shown in Fig. 5.4, is divided in two nearly identical macro-blocks, each one composed of three main parts: *GBT (Full Mode) FPGA Wrapper*, *Central Router* and *Wupper PCIe Engine*. The GBT FPGA Wrapper is a modified version of the GBT-FPGA firmware distributed by CERN [57]; it is in charge of handling the input/output communication implemented using the custom CERN GBT protocol (see Appendix B.3). This block encapsulates the Forward Error Correction (FEC) encoder/decoder, a scrambler/descrambler and a gearbox architecture.

The PCIe Engine interfaces the board with the PCIe bus by using a Xilinx PCIe Gen.3 hard block - functioning as a PCIe End Point - connected to a simple Direct Memory Access (DMA) logic called *Wupper*. Since the Integrated Block for PCI Express supports at most 8-lane operation, a PCIe switch *PEX8732* is used to connect two 8-lane endpoints to the 16-lane PCIe slot. The Wupper engine transfers data between a user-logic FIFO and the host server memory and it is logically divided in two main groups: DMA control - parsing and monitoring DMA descriptors - and DMA write/read - processing data-stream in both directions.

The Central Router block routes all the read-write data-stream inside the FELIX and works as a bridge between the GBT FPGA Wrapper block, the Wupper PCIe Engine block, and the TTC and busy logic coming from the TTC mezzanine.

The interface between the FELIX firmware functionalities and the user is provided by the FELIX Software suite, which can be divided in three levels: low level software, test software and production software. Two device drivers control the access to the FELIX hardware level: *flx driver* - a conventional character driver for PCIe cards - and *cmem rcc driver*. The latter has been developed by the ATLAS TDAQ project and allows the application software to allocate large buffers of con-

tiguous memory. Data transfers between the FELIX card and the memories are handled by the *felixcore application* through a dedicated library called NetIO. The full transmission procedure requires a few steps: firstly, the DMA engine transfers a data stream into a contiguous circular buffer. Secondly, the rcc driver allocates the buffer content in the memory of the host server, where it is inspected for integrity. Finally, a header is attached to the recombined data and the whole stream is published to the network through NetIO.

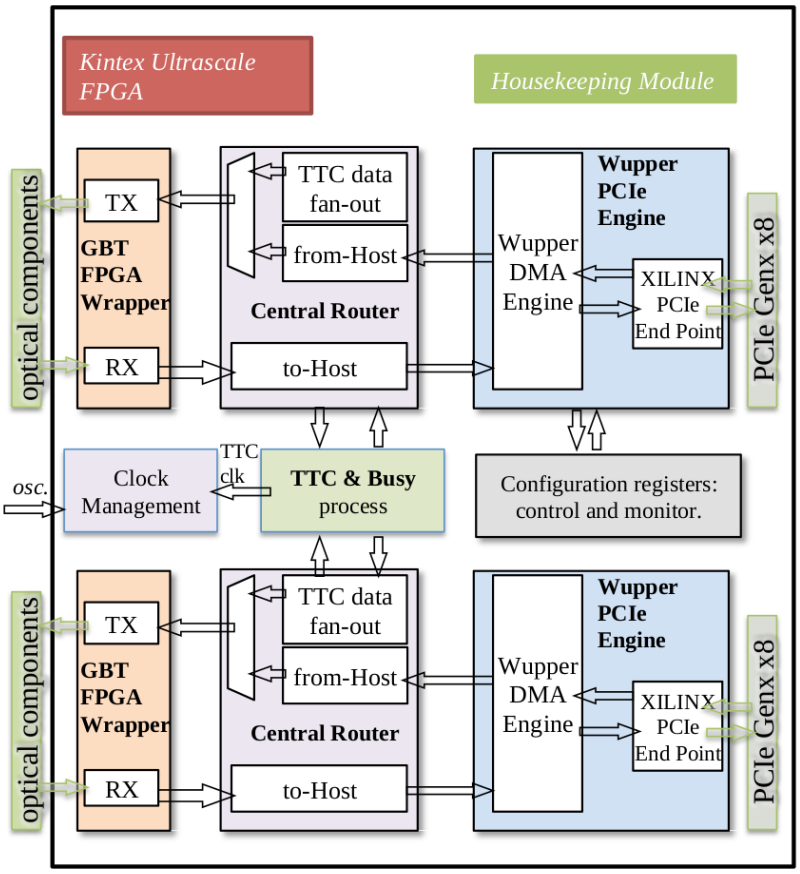


Figure 5.4: Block diagram of the FLX-712 firmware.

5.2.3 Test procedure for FLX-712 procurement

The Phase-I upgrade requires that more than 100 FLX-712 cards are produced and tested before the end of the first half of 2019. Part of this thesis work consisted in the development of a test procedure, meant to be executed by the production contractor, finalized to assure the quality of the boards and their components in the most efficient way possible. The first step of the procedure consists in the configuration of the non-volatile memories of the boards, such as the FLASH PROM, where the FPGA configuration files are stored, or the MINIPOD registers. Secondly, the optical

fibers are connected in loop-back and a custom firmware is loaded on the board to monitor the data transmission quality at the maximum bandwidth used by FELIX, which is 9.6 Gbps.

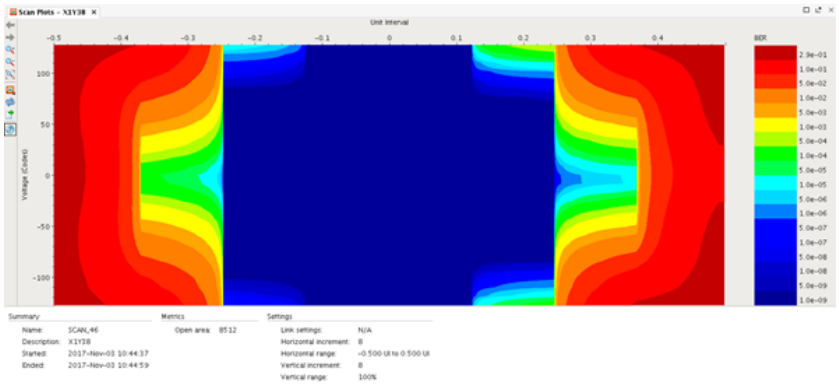


Figure 5.5: Eye-diagram scan from one of the 48 channels of a FLX-712. The *Open Area* is identified by the blue color. [58]

Table 5.1: Average Open Area results for all the FLX-712 channels at 9.8 Gbps speed.

Channel	Open Area	Channel	Open Area	Channel	Open Area
0	8832	16	9280	32	9088
1	9280	17	8576	33	9024
2	8576	18	8320	34	8576
3	9536	19	9088	35	8768
4	8448	20	9600	36	9728
5	9152	21	8640	37	9216
6	8128	22	9728	38	9344
7	8192	23	9216	39	9664
8	8704	24	8448	40	8384
9	9280	25	9216	41	9536
10	9280	26	8896	42	8640
11	9280	27	9216	43	8768
12	8576	28	9216	44	9536
13	9728	29	9216	45	8000
14	9792	30	8576	46	8704
15	8448	31	9152	47	9856

A good estimate on the quality of the link transmission is given by the so called *Eye Diagram* scan. The Eye Diagram scan is obtained by persistently sampling with an oscilloscope a digital signal using the data rate to trigger the horizontal sweep. Since the signal is digital, the theoretical shape obtained would be a rectangle; however, in real situations, effects such as signal rise time or jitter creates imperfections over the quality of the signal and the shape obtained on the oscilloscope resembles

an eye. The area contained inside the eye is called *Open Area*; the bigger the Open Area, the better the signal quality. A typical Eye Diagram Scan for the FLX-712 links, measured using Xilinx tools, is shown in Fig. 5.5, while the Open Area results are listed in Table 5.1.

The final step of the FLX-712 test procedure consists in loading the standard FELIX firmware and testing standard board procedures, such as throughput measurement with data emulated by the host server and loop-back GBT communication test. All the tests have been optimized and can be run on four FLX-712 cards in parallel, which is the maximum number of boards that can be contained in a standard host server provided by CERN.

5.3 The πLUP Project

The πLUP board (or PiLUP), shown in Fig. 5.6, was developed jointly by University and INFN Bologna as a Phase-II candidate for the ATLAS Pixel Detector readout chain.

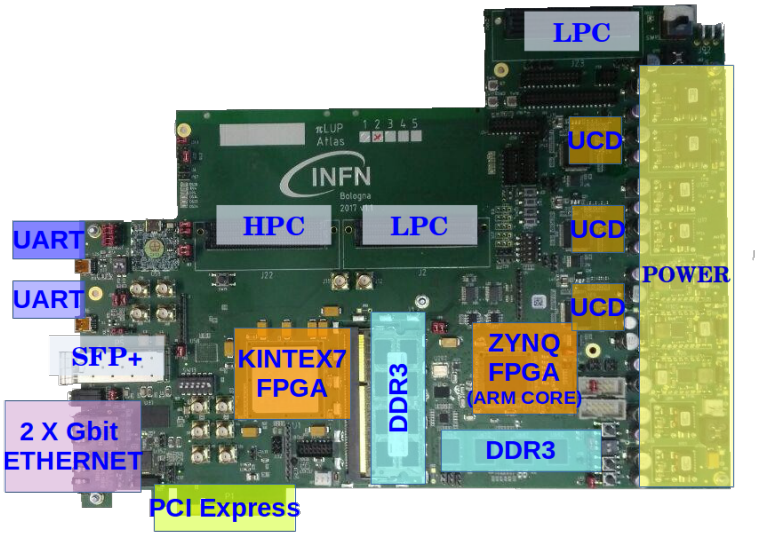


Figure 5.6: Bologna πLUP v1.1 board. FPGAs, I/O ports and other main components are highlighted in the picture. [59]

It was designed as a natural upgrade of the current Pixel Detector DAQ system, mainly composed of two electronic cards: Back Of Crate (BOC) - responsible for handling the control interface to the detector and the data from the detector - and the ReadOut Driver (ROD) - responsible of data processing and packaging - as described in Sec.3.4. The ROD and BOC boards are connected together through a Versa Module Eurocard (VME) crate and provide a total bandwidth of 5.12 Gbps. On the other hand, the πLUP board abandoned the VME connector, moving towards the solution of a 8 lanes *Peripheral Component Interconnect Express* (PCIe) bus. By

exploiting the most recent technologies, it also merges in a single board both the I/O and the data processing capabilities and can provide a total bandwidth of 80 Gbps. Mirroring the ROD structure, the πLUP features two FPGAs in a Master/Slave architecture. Both the FPGAs are from Xilinx (7^{th} generation); the Master FPGA is a Zynq-7 and the Slave is a Kintex-7 [54]. The Zynq-7 includes an embedded dual core ARM processor, which fulfills the same role as the PowerPC processor embedded in the ROD Master FPGA.

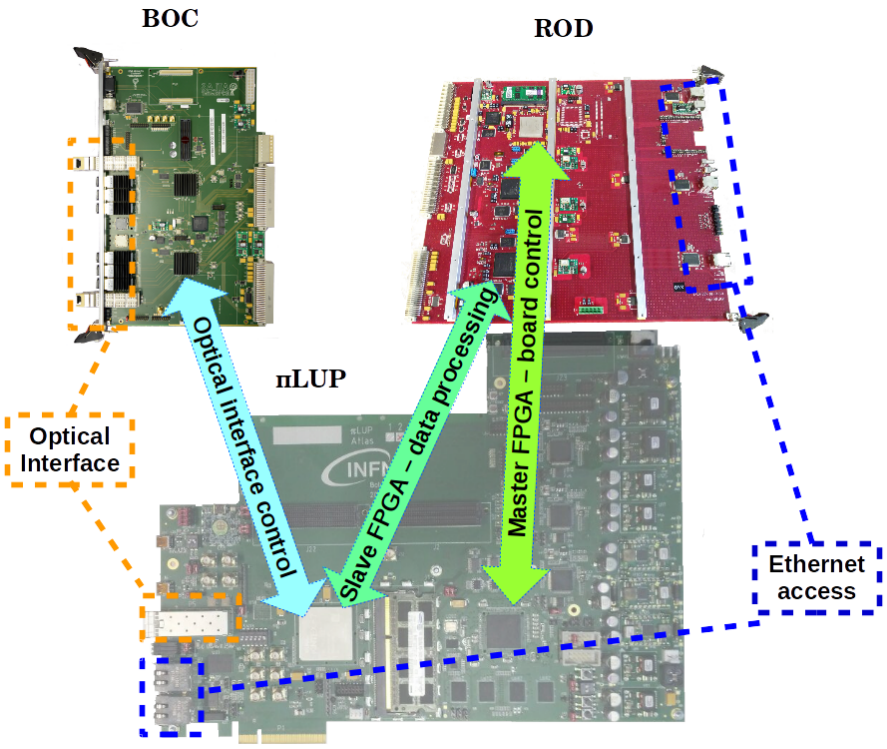


Figure 5.7: πLUP board with the BOC-ROD equivalent interface.

Apart from the PCIe, the πLUP card features a huge variety of I/O connectors, such as two Universal Asynchronous Receiver-Transmitter (UART) ports, one 1 Gbps Ethernet port, one 10 Gbps Ethernet port, one Small Form-factor Pluggable (SFP+) connector and three FPGA Mezzanine Card (FMC) connectors. Having a wide choice of different I/O interfaces attributes a great versatility to the πLUP card, making it perfectly suited to act as a general purpose readout board. In fact, although it was designed to fulfill a specific task, it can be used to interface several types of front-end chips or electronic systems.

Two first prototypes of Bologna πLUPs (version 1.0) were produced in 2016. Most of the I/O connectors and the internal functionalities were successfully tested. However, some small patches were required and the shape of the board had to be revisited to properly fit one of the FMC connectors. Those revisions led to the fabrication of four new boards (version 1.1) in 2018; the two version are shown in Figure 5.8.

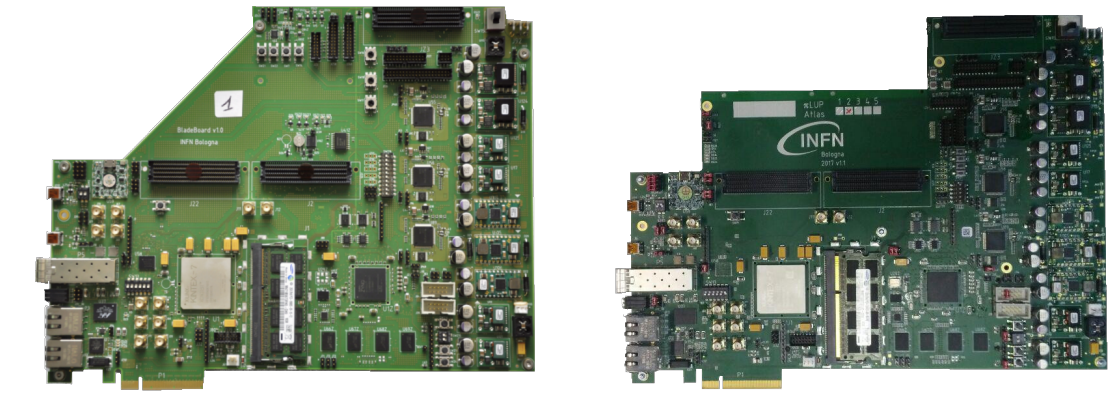


Figure 5.8: Bologna πLUP v1.0 board (left) and πLUP v1.1 (right). The shape was modified to fit all the FMC Mezzanines.

In the following sections a technical overview on the main components of the πLUP board will be presented, as well as its possible applications in different projects and the results obtained.

5.3.1 πLUP board overview

The Bologna πLUP card is a 16 layers PCI Express board capable of interfacing to several different other boards or front-ends and processing data at high speed. Figure 5.6 shows the main components on the board.

The πLUP features two Xilinx 7^{th} series FPGAs arranged in a Master/Slave architecture and connected together by a bus - namely KZbus - composed of 5 single ended and 21 differential lines. A Zynq XC7Z020-1CLG484C - embedding a physical dual-core ARM Cortex-A9 processor - is the Master FPGA and is in charge of controlling the data-flow and status of the Slave FPGA, a Kintex XC7K325T-2FFG900C. The Kintex device handles all the high speed I/O communications through 16 internal physical transceivers (GTx) [56] running at up to 12 Gb/s. The transceivers are connected to different types of physical ports such as the 8 lanes PCI Express, FMCs, SFP+ and 10 Gbit-Ethernet as described in Table 5.2. There are three FMC connectors on the board; one of them, a Low Pin Count (LPC) FMC connector is connected to the Zynq FPGA and the others, a High Pin Count (HPC) FMC connector and another LPC FMC connector are connected to the Kintex 7 FPGA. Those connectors are built accordingly to the VITA Standard 57.1 and can support any FMC mezzanine that respects the same standard.

Clock Distribution

Several clock sources are present in the board. The Zynq-7 FPGA is associated to three main clock sources. A 200 MHz system clock is provided by a SiTime SiT9102,

Table 5.2: GTx transceiver and reference clocks connections in πLUP Kintex-7 FPGA

Bank	REFCLK0	REFCLK1	MGT	I/O Port
115		PCIe REFCLK	0	PCIe lane 7
			1	PCIe lane 6
			2	PCIe lane 5
			3	PCIe lane 4
116	SMA REFCLK	LPC REFCLK	0	PCIe lane 3
			1	PCIe lane 2
			2	PCIe lane 1
			3	PCIe lane 0
117	125 MHz clk source	Si5326	0	SMA
			1	Gb-Ethernet
			2	SFP+
			3	FMC LPC
118	HPC REFCLK 0	HPC REFCLK 1	0	FMC HPC 0
			1	FMC HPC 1
			2	FMC HPC 2
			3	FMC HPC 3

a differential output programmable oscillator providing ±10 ppm frequency stability with sub-piscosecond phase jitter; a configurable user clock is provided by a Silicon Labs Si570, a low jitter oscillator that supports frequencies between 10 and 1400 MHz; the Processing System (PS) clock is provided by a 50 ppm 33.33 MHz oscillator. The Kintex-7 device features another 200 MHz SiT902 system clock and programmable Si570 user clock, as well as other clock inputs required by the GTx transceivers as reference clocks [56]. Some of those reference clocks are embedded on the πLUP itself, while others must be provided from the outside. The two sources provided by the πLUP are a 125 MHz Ethernet reference clock, provided by the combination of a 25 MHz crystal oscillator and a Integrated Device Technology 844021I-01 crystal oscillator interface, and a programmable reference clock, provided by a Silicon Labs Si5326 jitter cleaner. The external reference clock sources must be provided by the PCIe connector, by the LPC and HPC FMC connectors or by the SMA connectors. Table 5.2 shows how the reference clocks are associated to the different GTx transceivers.

5.3.2 Applications for the πLUP board

As already stated, the πLUP board was designed to fulfill a specific task, i.e. the readout upgrade for the next generation ATLAS Pixel Detector, merging in a single board both I/O connections and data processing. Nevertheless, the πLUP features a huge variety of I/O connectors and three FMC connectors, making the board highly versatile and able to interface a wide variety of different other electronic devices

and Front-End chips. The choice of having two FPGAs connected in a Master/Slave mode guarantees enough power to perform high level control operation on the board (Zynq-7 ARM core) and handle I/O communications through several different protocols (Kintex 7) while at the same time maintaining a relatively low price.

The main possible applications for the πLUP board are three:

- readout control system; the πLUP can be used to directly interface a front-end device performing data-processing, data transfer to the PC via PCIe bus, online on-chip histogramming and system control. The maximum bandwidth in this scenario is limited by the PCIe data transfer rate, i.e. 4 GB/s for the 8 lane gen. 2 PCIe bus;

- bridge between two different systems; the πLUP can be used as a bridge to connect two different readout systems that use different protocols or different communication physical layers. The maximum bandwidth in this scenario is highly influenced by the interfaced systems;

- data generator/ front-end emulator; the πLUP can be used to generate/emulate data to be sent to other systems, for example to validate a reconstruction or data-processing algorithm. The maximum bandwidth in this case would be 8 GB/s, which is the maximum speed of the I/O connectors, excluding the PCIe.

A practical example of each application will be presented in the following sections. Regardless of the final application of the πLUP, a common firmware/software structure has been designed to give to the embedded ARM processor full control over the main board functionalities. The kernel running on the ARM processor is a special Linux distribution developed by Xilinx, called *Petalinux*, which allows the user to interface to the board via UART or SSH protocols. Using special registers hard-coded in the Kintex firmware, the software is also able to automatically recognize the πLUP application and proceeds to configure it accordingly.

Software architecture

The two FPGAs mounted on the board are intended to be used in a master- slave configuration, with the Zynq, or more precisely its ARM-based Processing System (PS), controlling any peripheral in the system and acting as a main interface to the user. A diagram of the setup is shown in Figure 5.9. Inside the Zynq, the PS communicates with the FPGA through the AMBA AXI protocol. This channel is extended to the Kintex by the AXI Chip2Chip IP core offered by Xilinx. This core transparently bridges a 32-bit AXI bus to the slave device so that any peripheral present in the Kintex can be addressed from the ARM as if it was directly implemented in the Zynq. The physical interface is quite flexible and can be adapted to a limited pin count; in this case the communication employs 20 differential lines operating at 200 MHz Double Data Rate DDR (9 data bits plus clock for each direction). On startup the Chip2Chip automatically performs a deskewing self-calibration and

then is immediately ready to use. In any configuration the C2C master shows a single AXI slave port and the C2C slave a single AXI master port, so the bridge is not exactly "symmetrical", but this do not entail a limitation in this design. Four interrupt ports for each direction are also present. The C2C channel multiplexer assigns higher priority to those over AXI data. The Zynq PS runs an embedded

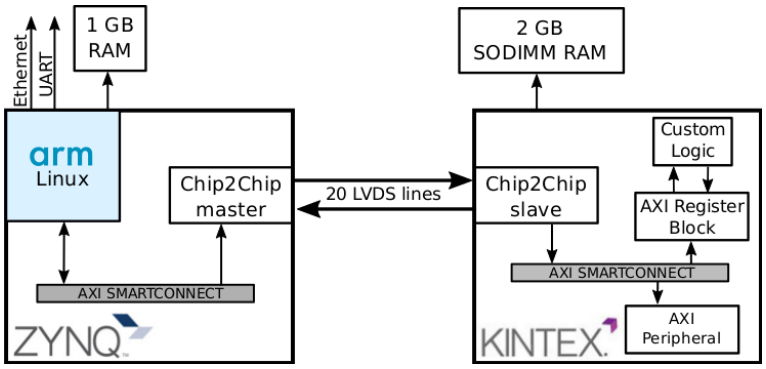


Figure 5.9: Diagram of the πLUP control infrastructure.

Linux distribution generated with the Xilinx Petalinux tools, providing a high level interface to any functionality present in the board (including web services such as an SSH server). During the boot-up, the Linux image can be loaded from the on-board flash chip or downloaded from a remote server with the Trivial File Transfer Protocol (TFTP) protocol. Generally most AXI cores offered by Xilinx also ship a driver often included in the Linux kernel tree. For custom-made cores without an AXI interface, a control interface is offered by an AXI-addressable register block, that is directly accessed from Linux user-space using the generic-User Input Output (UIO) driver. The UIO driver greatly simplifies the development of drivers that does not require a custom kernel module and fits very well with the view of offering a higher level interface to the functionalities implemented in the FPGA. Other off-chip devices, such as the I2C-programmable Si570 clock generator, Si5326 PLL and bus multiplexer, can also be controlled directly from Linux by means of an AXI-based I2C controller. The kernel already includes drivers for the bus multiplexer and the Si570; the former transparently manages the multiplexer and the kernel is simply presented with a number of buses than can be directly accessed. In this application the Si5326 is programmed by a custom user-space software that calculates the required values of its internal registers and writes them with a simple file access to the character devices representing the muxed bus associated to the device.

PCIExpress interface

The workflow to validate and measure the performance of the PCIe Gen 2 bus on the πLUP required the design of a custom firmware implemented on the Kintex 7 FPGA and the development of custom Linux drivers allowing read and write operations from and to the RAM memory on the board, plugged in one of the PCIe

slots of Linux PC. The test design consisted in using the PCIe bus to write and read the 2 GByte DDR3 RAM associated to the Kintex device, measuring BER and speed. The firmware was entirely designed using the Vivado^{TM} IP Integrator as it is shown Figure 5.10. It is composed of a DMA/Bridge subsystem for PCIe (XDMA), a Memory Interface Generator (MIG) and other support logic needed to correctly connect these two blocks.

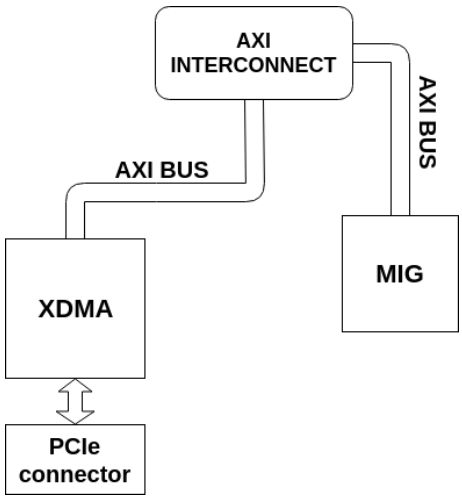


Figure 5.10: Block Diagram of the PCI Express validation firmware, realized with Vivado^{TM} IP Integrator. The main blocks are the DMA/Bridge subsystem for PCIe (XDMA) and the Memory Interface Generator (MIG).

The XDMA is an IP block that implements a high performance, configurable Scatter Gather DMA for use with the PCIe Gen2.1 and Gen3.x that can be configured to be a bridge between the PCI Express and AXI memory spaces. The master side of this block reads and writes requests on the PCIe and its core enables the user to perform direct memory transfers, both Host to Card (H2C), and Card to Host (C2H). The MIG IP core is a controller and physical layer for interfacing 7-series FPGA to DDR3 memory.

The custom drivers required to perform the test were developed for a Linux Ubuntu 16.04 Operative System. The test showed a peak user payload of 3.5 GBps data transfer when using buffers of 2 Mbyte and a BER$\leq 10^{-14}$ corresponding to 24 TByte data transferred without errors.

5.3.3 Front-end readout system

The original and main target of the Bologna πLUP board is to readout the new generation ATLAS Pixel Detector front-end chip: RD53, already introduced in Section 4.2.2. The πLUP acts as an interface between the chip and the user software, that sends configuration commands to the front-end and reads-back the response. The πLUP must hence be able to be connected to both the RD53 and a PC which

runs the main TDAQ software. The connection to the PC is implemented on the PCIExpress bus, which can run up to 4 GB/s, or 32 Gbps. The chip interface, on the other hand, requires an external FMC mezzanine that must be plugged on one of the FMC connectors on the board.

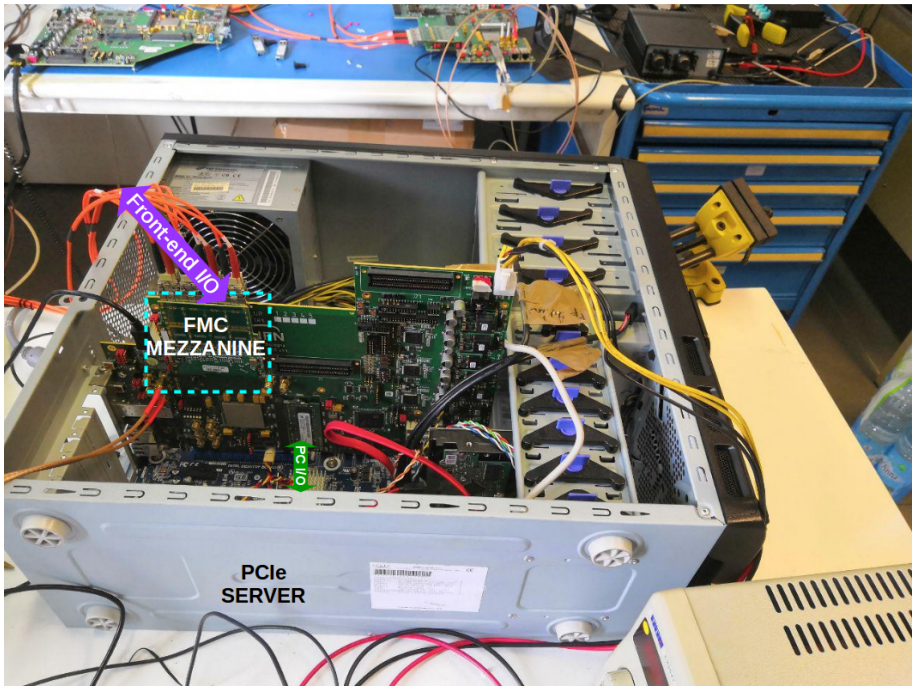


Figure 5.11: Picture of the test setup of the RD53A readout system using πLUP board. The πLUP is inserted in a Gen.2 PCIExpress PC and a FMC mezzanine is connected to interface the front-end chip. In this scenario an emulator is used to reproduce RD53A data and the connection is implemented on optical fibers instead of DP cable.

For the first prototype of RD53, called *RD53A*, the input/output interface runs on a *Display Port* (DP) cable, and hence a DP mezzanine is needed on the πLUP. The advantage of having a FMC connector on the card is that, by using different mezzanines, the πLUP can interface any type of front-end chip. Fig. 5.11 shows the setup realized in a laboratory of University of Bologna; the card is plugged in PCIe slot of a server and an optical mezzanine is used to interface it to a RD53A emulator. The DAQ system firmware implemented on the πLUP board is shown in Fig. 5.12; three main blocks can be distinguished:

- *PCIe interface*, handling the PCIe communication on both directions between the πLUP and the host PC;
- *Aurora decoder*, which decodes the input data from RD53A (Aurora 64b/66b running on four serial lanes at 1.28 Gbps) and deserializes them;
- *Timing and Trigger Control (TTC) Interface*, which receives RD53A configuration and trigger commands from PCIe, encodes them in the front-end format,

serializes and transmits them on a 160 Mbps serial line.

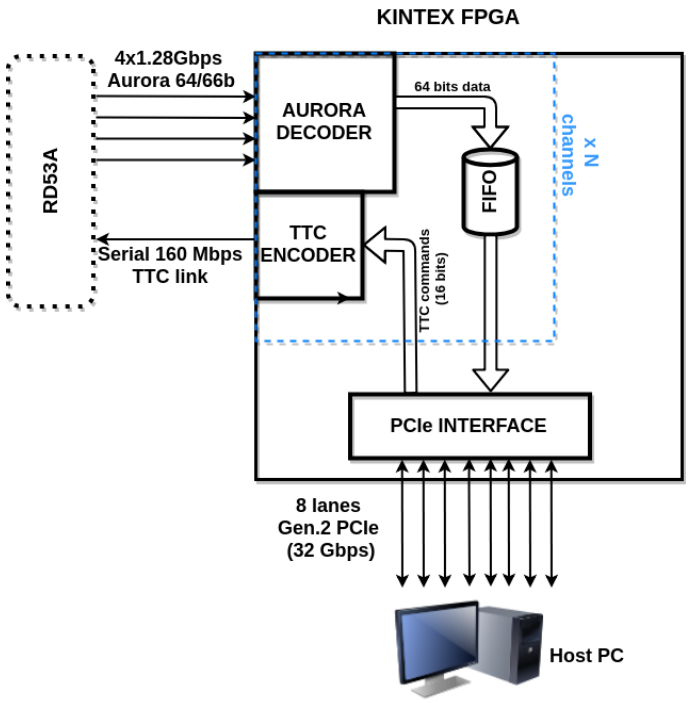


Figure 5.12: Block diagram of RD53A readout firmware running on πLUP. Three main blocks can be identified: Aurora decoder, TTC encoder and PCIe interface.

The number of front-end chips that can be interfaced to the πLUP depends on the FMC mezzanine used and is limited by the PCIe bandwidth, which can fit a maximum of six RD53A chips, whose maximum bandwidth is 5.12 Gbps. Nevertheless the πLUP can interface more than six front-end chips simply by running them at a lower readout speed. While the final ATLAS Phase-II readout system will probably be based on the FELIX board, the πLUP setup is still useful and can be used by laboratories during the chip testing phase. In fact, the system based on πLUP is generally simpler and requires less components with respect to the FELIX one; moreover, the software structure driving the configuration and readout of the chip will be compatible to the ATLAS Phase-II one.

5.3.4 Interface with FELIX

A first proof of the many possibilities of the πLUP board came from a integration test with FELIX boards from the Felix Project. The πLUP was connected through optical fiber to a Mini-Felix card (Xilinx VC709 evaluation board) and a FLX-712 card. The test showed that the two boards were able to establish a communication via both GigaBit Transceiver (GBT, 4.8 Gb/s) and custom Felix Full Mode (9.6 Gb/s) protocols. For both configurations, the πLUP used the Common Phase Locked Loop

(CPLL) of the transceivers to recover the clock from the incoming data stream; the clock was then cleaned using the jitter cleaner Si5326 on the board and propagated to the Quad PLL (QPLL) for the GTx transmitters. This way the acquisition system is totally synchronous and the data transmission frequency and phase is aligned with the incoming one. Figure 5.13 shows the clock distribution of the πLUP board.

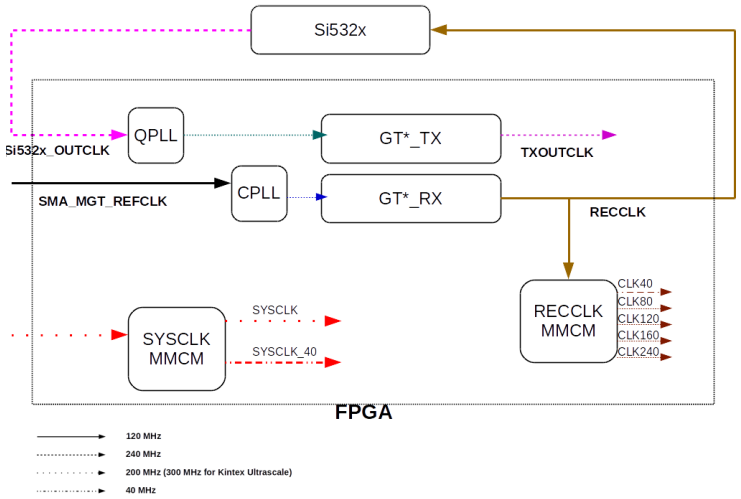


Figure 5.13: Clock distribution of the πLUP GBT/Full Mode protocol firmware.

Using a Faster Technology FM-S14 FMC HPC Mezzanine Card, providing four additional SFP+ connectors, four link connections were simultaneously established between the πLUP and the Felix cards, resulting in a total throughput of 19.2 Gbps in GBT configuration and 38.4 Gbps in Full Mode configuration. Both configurations were tested for about one hour and no errors were found, demonstrating the reliability of the connections.

Protocol converter

The capability of the πLUP of interfacing both the FELIX cards and the RD53A chips was put to practical use in the creation of a readout chain for the ITK front-end chip based on the FELIX. The need of using the πLUP board as an interface system arise from the physical and protocol incompatibilities between the Felix card and the RD53A chip. The first communicates via optical fibers through either 4.6 Gbps GBT or 9.8 Gbps Full Mode protocols, while the latter communicates via Display Port (DP) connectors through 160 Mbps E-link (input) and four lanes 1.28 Gbps Aurora 64/66 protocol (output). The role of the πLUP is hence to act as a bridge between this two systems, handling both the Felix-to-RD53A data-path (downlink) and the RD53A-to-Felix path (uplink). This is done through different firmware blocks, as shown in Figure 5.14. The GBT_FPGA block decodes the GBT-formatted data from Felix containing the configurations commands for the RD53A chip and also synchronizes to the Felix clock, recovering it from the data-stream.

Both configuration commands and clock are then propagated to the TTC Encoder firmware block, which is in charge of converting the commands to a RD53A compatible format and of encapsulating them in a single 160 Mbps serial line, connected to one of the DP connector data lanes.

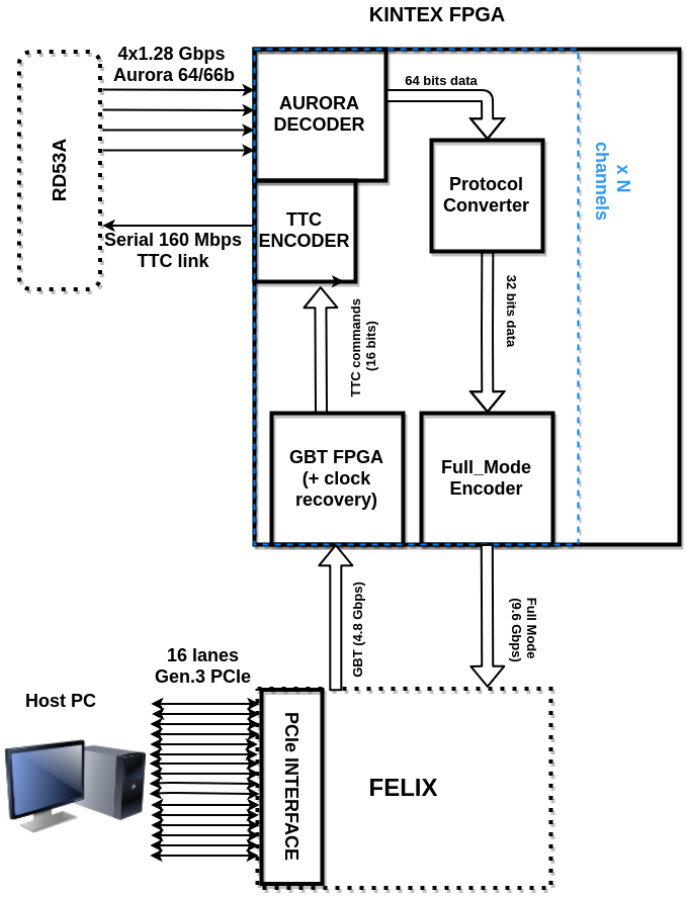


Figure 5.14: Block diagram of πLUP Protocol Converter firmware. The πLUP is used as an interface between the RD53A, communicating with Aurora 64/66b and serial TTC line (electrical) and the FELIX card, communicating with GBT and Full Mode protocols (optical fibers).

Concurrently the πLUP receives and decodes Aurora 64/66 data from the RD53A chip, coming from the other four data lanes of the DP connector. Those four lanes 1.28 Gbps data (resulting in a total throughput of 5.12 Gbps) are then passed to the Protocol Converter firmware block, which merges them in a single Full Mode stream that is transmitted to Felix via optical connection.

Although the πLUP doesn't include a DP connector in its design, the usage of FMC cards can sort through this problem. In particular, two custom FMC mezzanines were developed to be used as an interface to the RD53A: Single Module Adapter (SMA), a HPC FMC mezzanine featuring two DP connectors, and Multiple

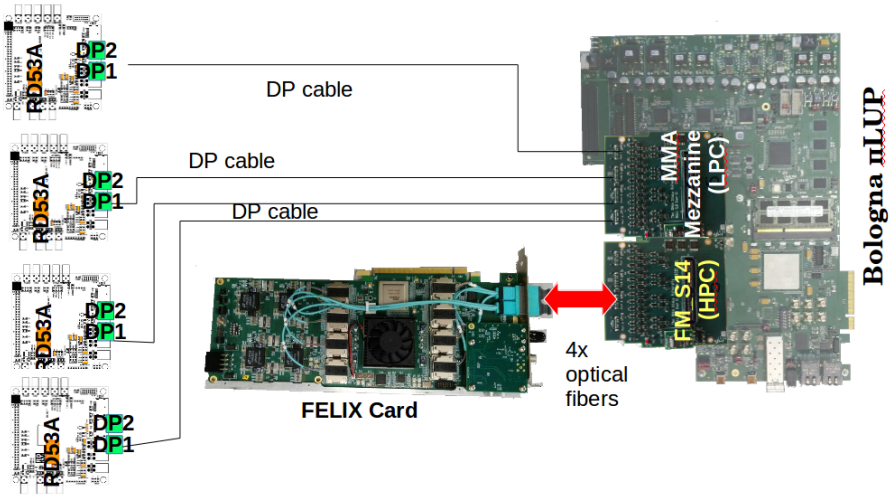


Figure 5.15: Configuration between Felix and four RD53A modules, using the Bologna πLUP as an interface. This configuration requires the usage of the MMA FMC LPC mezzanine (four DP connectors) and the FM-S14 FMC HPC mezzanine (four SFP+ connectors).

Module Adapter (MMA), a LPC FMC mezzanine featuring four mini-DP connectors. The maximum throughput for the πLUP can be obtained by the usage of both the MMA LPC mezzanine and the FM-S14 HPC mezzanine (featuring four SFP+ connectors). Using this configuration, shown in Figure 5.15, a Felix can interface four RD53A chips, resulting in a total throughput of $4 \times 5.12\,\text{Gbps} = 20.48\,\text{Gbps}$.

Data generator

In order to prepare the full readout system even before the RD53A chip was available, the FELIX/πLUP chain needed a device that could emulate the behavior of the front-end, reading configuration commands and sending fake data with the appropriate protocols. This task was fulfilled by the πLUP board itself; the digital part of the RD53A was implemented in the firmware of the Kintex-7 FPGA, that includes at the same time all the Protocol Converter functionalities. Even if the RD53A emulator cannot execute all the functions of the real chip, it can still perform read/write registers operations, receive input commands and generate random data in case of trigger or calibration pulse. Fig. 5.16 shows a RD53A digital scan obtained with the FELIX DAQ system and the πLUP board used as front-end emulator and Protocol Converter. This configuration proved to be more than sufficient to validate the whole readout chain.

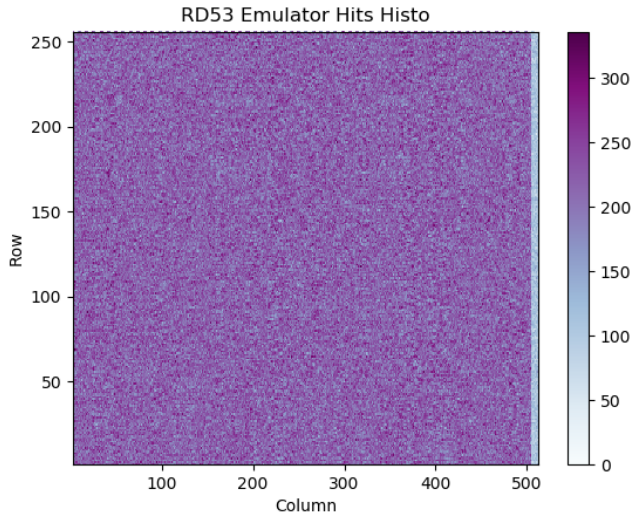


Figure 5.16: RD53A digital scan obtained using the FELIX card and DAQ software and the *pi*LUP ad RD53A Emulator and Protocol Converter. Pixel hits are randomly generated and are not uniform over the sensor.

5.4 πLUP performance

To evaluate and monitor the performance of the GTx transceivers on the πLUP board, the LogiCORE^{TM} IP Integrated Bit Error Ratio Test (IBERT) core for 7 series FPGA [60] was used. This IPcore generates the eye diagrams and calculates the Open Area and Bit Error Rate (BER) for the different I/O interfaces, that were connected in loopback mode. To test the four transceivers in the FMC HPC connector, a Faster Technology FM-S14 Mezzanine Card was used. This Mezzanine Card implements four SFP+ connectors and two IDT ICS8N4Q001 programmable reference clocks.

Table 5.3: Open Area results run at 5/10 Gbps, PRBS 31-bit and BERR 10^{-9}

I/O Connector	Open Area (5 Gbps)	Open Area (10 Gbps)
FMT HPC MGT 0	11952	2784
FMT HPC MGT 1	9008	2272
FMT HPC MGT 2	10480	2512
FMT HPC MGT 3	11280	2480
SFP+	10432	2320
SMA MGT	6704	512

The BER and eye diagram scans were performed at 5 Gbps and at 10 Gbps; the two speeds were chosen to be slightly higher than the design operation mode protocol speeds, i.e. GBT (4.8 Gbps) and Full Mode (9.6 Gbps). The tests were performed

using a PseudoRandom Binary Sequence (PRBS) of 31 bits and requiring a BER $\leq 10^{-9}$. The BER test was then continued until the error rate reached was $\leq 10^{-14}$. Figure 5.17 shows the eye diagram of the tests run at 5 Gbps and Fig. 5.18 shows the results at 10 Gbps; Table 5.3 shows the open area results.

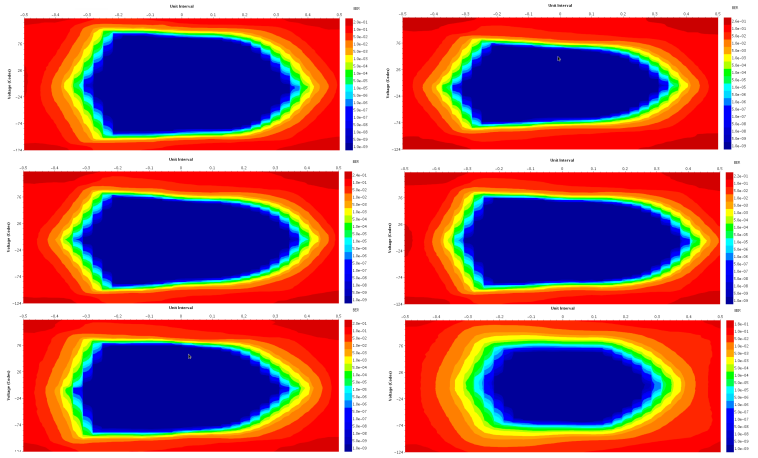


Figure 5.17: Eye Diagram scan results run at 5 Gbps, PRBS 31bit and BER 10^{-9} of: FMT HPC MGT 0 (top left), FMT HPC MGT 1 (top right), FMT HPC MGT 2 (middle left), FMT HPC MGT 3 (middle right), SFP+ (bottom left) and SMA MGT (bottom right).

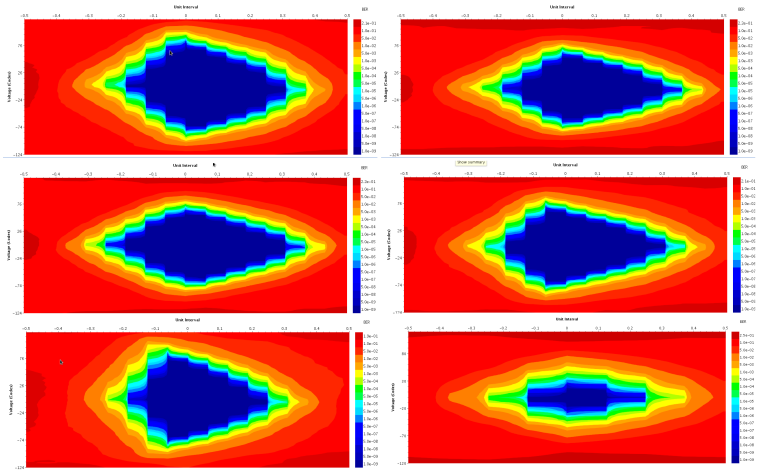
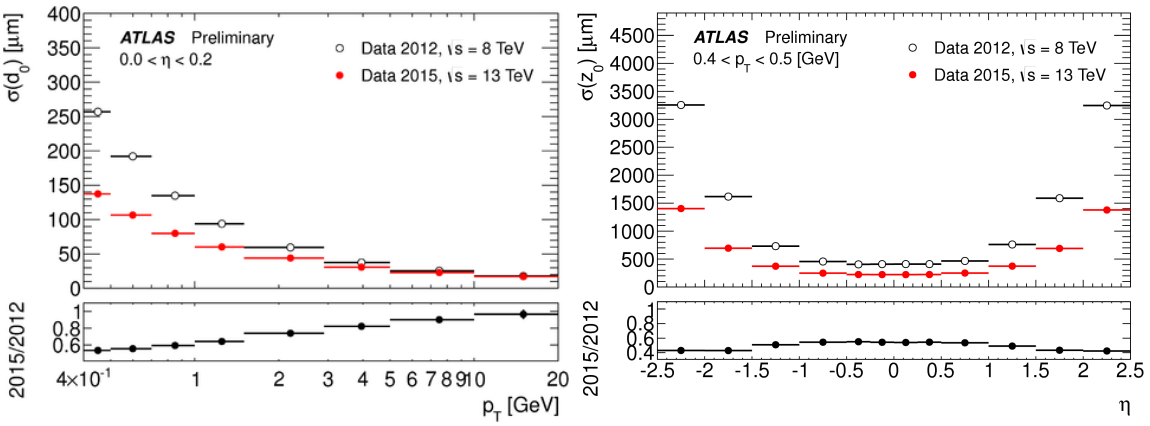


Figure 5.18: Eye Diagram scan results run at 10 Gbps, PRBS 31bit and BER 10^{-9} of: FMT HPC MGT 0 (top left), FMT HPC MGT 1 (top right), FMT HPC MGT 2 (middle left), FMT HPC MGT 3 (middle right), SFP+ (bottom left) and SMA MGT (bottom right).

Conclusions

In this thesis I reported the work I did during the last three years as a PhD student at University of Bologna, in the framework of the ATLAS Experiment at CERN.

My work concerned the innermost ATLAS subsystem, the Pixel Detector and in particular its readout chain. Since the moment it was built, the ATLAS detector underwent a series of gradual upgrades to keep up with the increasing LHC accelerator performance. The first upgrade involving directly the Pixel Detector was the insertion of Insertable B-Layer (IBL), a fourth pixel layer that was added in 2014-2015 between the B-Layer - which was the innermost layer of the Pixel Detector - and the beam pipe. IBL was added to solve the problem of the aging of the B-Layer pixels and to increase the tracking resolution, in order to fulfill the ATLAS physics program requirements. The reasons that led to the Pixel Detector first upgrade and the results achieved were presented in detail in this thesis, in Chapter 2.



(a) (b)

Figure 5.19: Unfolded transverse impact parameter resolution measured from data in 2015 at $\sqrt{s} = 13$ TeV with the Inner Detector including IBL, as a function of p T , for$0.0 < \eta < 0.2$, (a) and η and $\sqrt{}$ for $0.4 < p_T < 0.5$, (b) compared to that measured from data in 2012 at $\sqrt{s}$=8 TeV.

As a result of the upgrade, the overall performance of the ATLAS Inner Detector improved considerably, in terms of spatial resolution, tracking capabilities and vertex reconstruction; Fig. 5.19 shows as an example the improvements on *transverse*

impact parameter resolution.

The introduction of IBL required a modification in the readout system, since the different data format of the front-ends, the higher occupancy and the larger rate could not be handled by the previous DAQ chain. The new readout is based on two VME boards, the IBL ReadOut Driver (IBL-ROD) and the IBL Back Of Crate (IBL-BOC), that together provide enough bandwidth and high level control capabilities. The same system was used in the following year to interface the rest of the Pixel Detector, replacing the previous boards that were incapable to cope with the increasing data throughput. Chapter 3 presented an overview on the readout upgrade, its motivations and its current status. Table 5.4 summarizes the main properties of the new DAQ system, comparing it to the old one.

Table 5.4: Comparison between old and new readout system.

Layer	**Link Readout Speed (Mbps)**		**Bandwidth Saturation at $\mu = 60$ (%)**		**# of RODs installed**	
	old	**new**	**old**	**new**	**old**	**new**
IBL	-	160	-	50	-	15
B-Layer	160	160	81	81	22	44
Layer 1	80	160	103	52	38	38
Layer 2	40	80	159	79	26	26
Disks	40	80	85	63	12	12

During my PhD I collaborated with the *ATLAS Pixel Detector DAQ Group* supervising the readout upgrade of the Pixel Detector, providing support and designing software and firmware tools. As the main ROD firmware developer, I introduced several firmware mechanisms intended to reduce the amount of data de-synchronization and loss of module configuration affecting the quality of the data acquisition. These mechanisms (Smart resynchronization, Smart L1 forwarding, Quickstatus register monitoring), described in Chapter 3, had a great impact on data quality and will allow the system to maintain good performance during Phase-I LHC, where the luminosity will be 2-3 times higher than the current one. Fig. 5.20 shows the results of the DAQ improvements in terms of amount of synchronization errors; it can be noticed that, even if the instantaneous luminosity increases over time, the amount of errors is kept under control.

At the end of Run-III, in 2022, the ATLAS Pixel Detector will be upgraded again, to be able to sustain the HL-LHC environmental conditions during Phase-II. In fact the entire Inner Detector will be replaced with a new one, called Inner Tracker (ITk), entirely based on silicon sensors. ITk will use smaller pixels made using different technologies; the usage of 3D sensors, already exploited by IBL, will be largely expanded especially in the innermost pixel layers, to provide a very high tolerance to radiation. ITk will present an innovative design - even if the technical

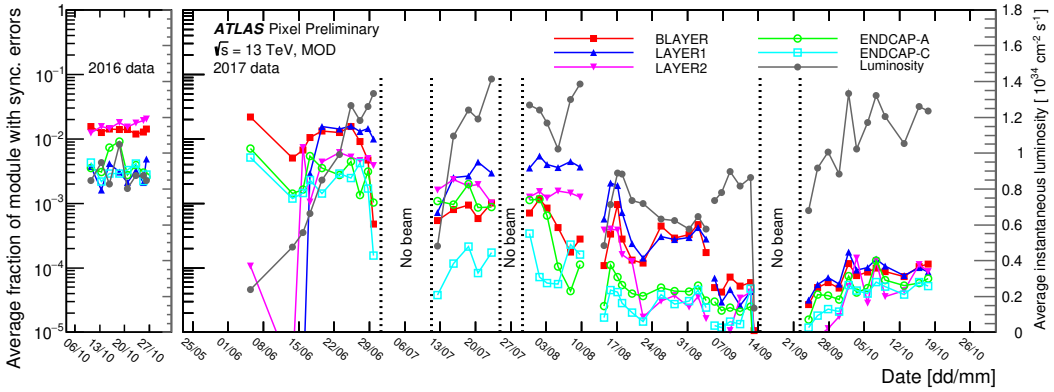


Figure 5.20: The average fraction of Pixel modules with synchronization errors at the module level per event in 2016 and 2017 runs. Each point shows the average fraction in a given run. The synchronization error signals a discrepancy between the level-1 trigger or bunch crossing identifiers recorded in the front-end chips and those stored in the central acquisition system. The synchronization error rate at the module level is decreased in all layers in 2017 mainly due to implementation of automatic recovery action in software for new readout electronics used for Layer-1 and Layer-2 and the mechanism to reset the internal trigger FIFO and clear all memory of the events currently stored in the FE chip at each Event Counter Reset (ECR) for all layers. [PIX-2015-003]

details have not been finalized yet - to maximize the pseudorapidity coverage and minimize the material budget. The new front-end chip that will be used for ITk is called *RD53*, designed by the RD53 Collaboration, involving several institutes around the world. It will be realized using 65 nm technology and will contain more than 500 millions of transistors; in order to reduce effects of Single Event Upset, it will be designed using special techniques, such as triple register redundancy and trickle configuration. The final version of the chip has not been realized yet, but a first prototype, called *RD53A*, was already produced and tested.

The ATLAS conditions during Phase-II, when the pile-up and hence the average occupancy will be very high and the trigger rate will be increased from $\sim$ 100 kHz to $\sim$ 1 MHz, will require another upgrade of the readout chain. The final system has not been decided yet and many proposals are currently under investigation. During my PhD I worked together with the *πLUP Project* and the *FELIX Collaboration* that developed two possible DAQ boards to be used by Phase-II ATLAS. The πLUP board was built by INFN and University of Bologna as an upgrade of the ROD-BOC cards and it was originally designed to readout ITk. Featuring newer technologies, such as 7^{th} generation Xilinx FPGAs and PCIExpress bus, the πLUP provides more bandwidth than the ROD-BOC and is perfectly suited to readout the RD53 front-end chips. Moreover, it features an embedded dual core ARM processor and a variety of different connectors, making it extremely versatile and usable in multiple applications. Table 5.5 shows the differences between the ROD-BOC and

the πLUP boards.

Table 5.5: Comparison between ROD-BOC and *pi*LUP boards.

	ROD-BOC	**πLUP**
Total bandwidth	5.12 Gbps	70 Gbps
Data transmission	Optical	Electrical/optical
Communication Protocol	S-Link	GBT/Full Mode/Aurora
External interface	VME	PCIExpress (Gen.2)
Embedded processor	PowerPC	Dual-core ARM
Number of FPGAs	7	2
FPGAs series	Virtex-5 / Spartan-6	Kintex-7 / Zynq-7
Board support	VME crate	PC case
Number of FE-I4 that can be readout	32	> 400

I personally contributed to the πLUP Project by designing the firmware and the software structure of the board for all its possible applications, that were described in detail in Chapter 5. Also, starting from 2017, I worked in a collaboration between the πLUP Project and the FELIX Project. The first test involving both boards consisted in a bidirectional communication using GBT and Full Mode protocols (Fig. 5.21). The test proved the compatibility between the two systems and marked the starting point for the beginning of the collaboration.

```
->3583 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26487.8 MB, file 0 B; Buffer: 0%, wraps 24
->3584 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26495.2 MB, file 0 B; Buffer: 0%, wraps 24
->3585 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26502.6 MB, file 0 B; Buffer: 0%, wraps 24
->3586 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26510.0 MB, file 0 B; Buffer: 0%, wraps 24
->3587 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26517.4 MB, file 0 B; Buffer: 0%, wraps 24
->3588 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26524.8 MB, file 0 B; Buffer: 0%, wraps 24
->3589 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26532.1 MB, file 0 B; Buffer: 0%, wraps 24
->3590 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26539.5 MB, file 0 B; Buffer: 0%, wraps 24
->3591 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26546.9 MB, file 0 B; Buffer: 0%, wraps 24
->3592 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26554.3 MB, file 0 B; Buffer: 0%, wraps 24
->3593 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26561.7 MB, file 0 B; Buffer: 0%, wraps 24
->3594 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26569.1 MB, file 0 B; Buffer: 0%, wraps 24
->3595 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26576.5 MB, file 0 B; Buffer: 0%, wraps 24
->3596 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26583.9 MB, file 0 B; Buffer: 0%, wraps 24
->3597 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26591.3 MB, file 0 B; Buffer: 0%, wraps 24
->3598 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26598.7 MB, file 0 B; Buffer: 0%, wraps 24
->3599 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26606.1 MB, file 0 B; Buffer: 0%, wraps 24
->3600 sec, Rates: recv   7.4 MB/s, file   0.0 MB/s; Total: recvd 26613.5 MB, file 0 B; Buffer: 0%, wraps 24
**STOP**
-> Data checked: Blocks 25989673, Errors: header=0 trailer=0
Exiting..
[pixeldaq@pcatlasfelixbo01 ~]$
```

Figure 5.21: First communication test between FELIX and πLUP boards via GBT protocol. The test was run over 1 hour and no transmission errors were found.

As a result of the collaboration, I spent six months at Brookhaven National Laboratory (BNL), where I contributed to the development of the FELIX firmware. Also, I designed a test procedure for the FELIX cards FLX-712, that will be pro-

duced in 2018/2019 and will enter the Phase-I readout chain for a portion of the ATLAS sub-detectors.

During the same period I designed a RD53A readout chain using both the FELIX and the πLUP boards: the FELIX acts as the main DAQ system, and the πLUP is used as an interface between the FELIX and the front-end chip. Also, since the RD53A was not available at the time, the πLUP acts also as a chip Emulator, so that the full chain could be validated. The first test setup, created at BNL and replicated at University of Bologna, is shown in Fig. 5.22.

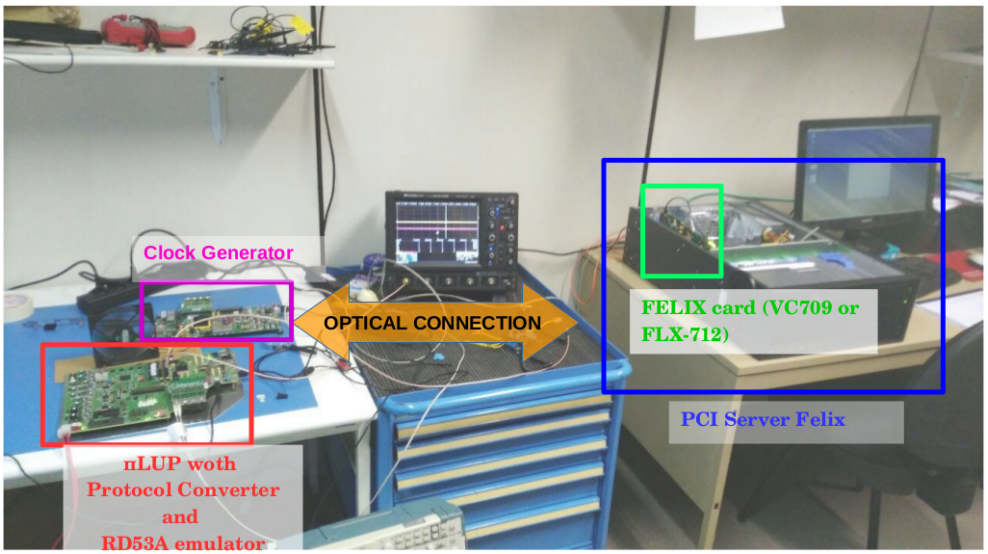


Figure 5.22: Picture of the test setup of the RD53A using the FELIX ad πLUP cards. The πLUP acts as a RD53A emulator and a Protocol Converter at the same time. The transmission between the two boards is implemented on optical fibers.

All the work described in this thesis had a great impact on the ATLAS Experiment performance and data quality. To achieve these results, cutting-edge technologies have been exploited, incorporating the latest advancements in each upgrade. The expertise acquired in these projects proved to be extremely valuable and opened the path towards several collaborations with international institutes.

The constant evolution of the system and its continuous developments will also guarantee stable and durable conditions for the ATLAS Experiment in the upcoming years. As a result, ATLAS will be able to achieve unprecedented performance and to extend its physics program.

Appendices

Appendix A

Semiconductor Properties

The crystalline structure of solid materials results in the formation of well defined energy levels and bands for the electrons of that solid that may be separated by gaps of forbidden energy. The energy bands of most interest are the *valence band*, containing the outer-shell electrons bounded within the crystal, and the *conduction band*, the immediately higher energy band corresponding to the electrons that are free to migrate through the material. According to the value of the gap of forbidden energy between the valence and conduction bands, the solids can be roughly divided in three main categories, as shown in Fig. A.1.

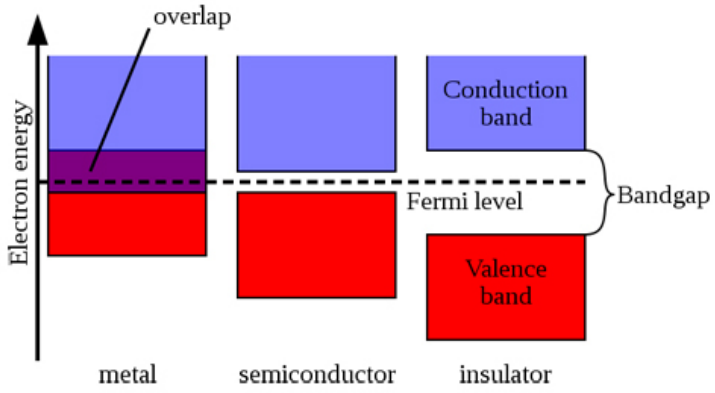


Figure A.1: Band structure for electrons in metals, semiconductors and insulators.

In metals there is an overlap of the two energy bands because the highest occupied energy band is not completely full, and the electrons can easily migrate within the material since they only need a very small increase of energy to enter the conduction band. In semiconductors and insulators, on the other hand, the valence and conduction bands are well separated by a gap of forbidden energies, and the electrons must have enough energy to fill the gap before crossing to the next energy level. The main difference between insulators and semiconductors, conferring completely different properties to the two kinds of crystals, is the value of the energy

gap, that is generally greater than 5 eV for insulators and on the order of 1 eV, hence easier to be overcome, for semiconductors. The main diffused materials that fulfills the semiconductor properties are *silicon* and *germanium*, although the first is largely the most diffused and will be the main subject of the rest of this overview.

A.1 Electron-hole pairs in semiconductors

In a pure semiconductor an electron can pass from the valence to the conduction band if it gains enough thermal energy. The probability of this process is strictly dependent on the temperature and can be express as

$$p(T) = CT^{3/2} \cdot e^{-\frac{E_g}{2KT}} \tag{A.1}$$

where:

- T = absolute temperature (K);
- E_g = gap energy value;
- k = Boltzmann constant
- C = constant characteristic of the material.

When an excited electron occupies the conduction band, a vacancy, or *hole*, in the otherwise full valence band is created; the combination of the two is called *electron-hole pair*. In a pure *intrinsic* semiconductor, the concentration (number per volume unit) n_i of electrons in the conduction band is equal to the concentration p_i of holes in the valence band and corresponds to $\approx 1.5 \cdot 10^{10}\, cm^{-3}$ [17] in silicon. If an external electric field is applied to the material, both the electrons and the holes will start to *drift* in opposite directions (holes move from one position to another if an electron leaves a normal valence site to fill an existing hole), contributing to its electrical conductivity. When the electric field E applied is moderate, the drift velocity ν of electrons and holes is proportional to the field itself, and the proportional constant μ is called *mobility*:

$$\begin{aligned} \nu_e &= \mu_e \cdot E \\ \nu_h &= \mu_h \cdot E \end{aligned} \tag{A.2}$$

Since electrons and holes, having opposite charge, drift in opposite directions, the generate electric current has an additive contribution from both the movements, and can be express in terms of the area A, the electronic charge e and the drift velocity:

$$\begin{aligned} I = I_e + I_h = Ae\left(n_e\nu_e + n_h\nu_h\right) = An_ie\left(\nu_e + \nu_h\right) = \\ An_ieE\left(\mu_e + \mu_h\right) = An_ie\frac{V}{l}\left(\mu_e + \mu_h\right) \end{aligned} \tag{A.3}$$

Defining the *resistivity* ρ as

$$\rho = R \cdot \frac{A}{l}$$

and combining it with the previous equation, a general expression for the resistivity for semiconductor solids can be given as

$$\rho = R \cdot \frac{A}{l} = \frac{V}{I} \cdot \frac{A}{l} = \frac{1}{e n_i (\mu_e + \mu_h)} \tag{A.4}$$

For a pure silicon material at room temperature, the typical value of resistivity is $\sim 230\, k\Omega \cdot$cm.

At higher electric field values, the drift velocity tends to increase more slowly until it reaches a saturation value, usually of the order of $\sim 10^7$ cm/s [17] for both electrons and holes. For practical cases, the saturation velocity is often reached, making the semiconductor detectors one among the fastest time-response detectors available, since the time required to collect the charge over a typical 0.1 cm dimension would be under 10 ns.

A.1.1 Doped Semiconductors

The intrinsic concentration of electrons or holes in a semiconductor can be artificially modified by taking advantage of the atomic composition of the material. Silicon atoms, for example, are tetravalent and in a normal crystalline structure they form a bond with the four nearest silicon atoms. If the material is *doped* with a small fraction ($\sim 10^{-8}$) of pentavalent atoms (such as phosphorus), those atoms will substitute some silicon ones in the crystal, as shown in Fig. A.2 (a). In this scenario, one of the five valence electrons of the pentavalent atom is very weakly bonded with the lattice structure, and very little energy is required to dislodge it. The amount of electrons in

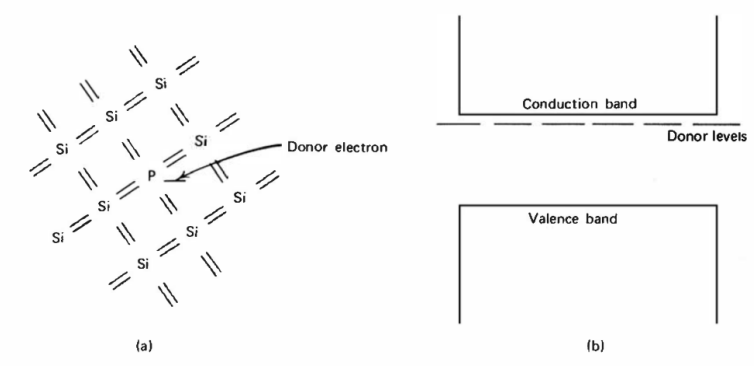


Figure A.2: (a) Representation of a donor impurity (phosphorus) occupying a substitutional site in a silicon crystal. (b) Corresponding donor levels created in the silicon bandgap. [17]

the valence band is hence dominated to the concentration of pentavalent atoms, that

for this reason are called *donors*. In this scenario the concentration of electrons is much higher than the concentration of holes, that are only thermally produced; the semiconductor is called *n-type doped semiconductor*, the electrons are the *majority charge carriers* and the holes are the *minority charge carriers*.

The opposite scenario happens if the silicon semiconductor is doped with trivalent atoms. The atoms will still substitute the silicon in the crystal structure, but it has one fewer valence electron than the surrounding silicon atoms and therefore one covalent bond is left unsaturated. This vacancy represents a hole similar to that left behind when a normal valence electron is excited to the conduction band, meaning that each impurity acts as a hole. In this case the semiconductor is *p-type doped*, the holes are the *majority charge carriers* and the electrons, that are only thermally generated, are the *minority charge carriers*.

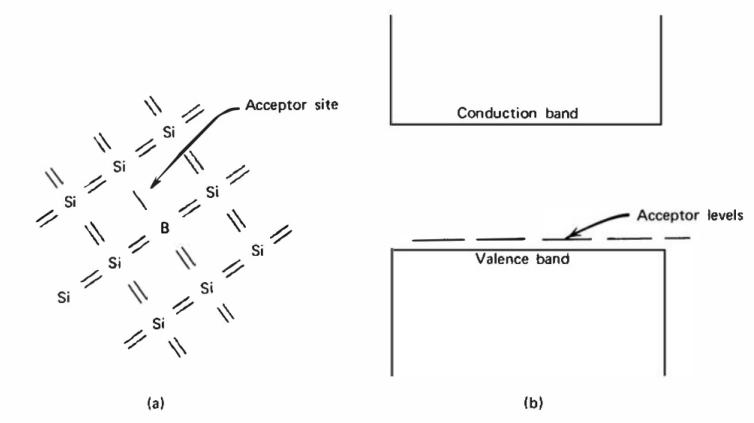


Figure A.3: (a) Representation of an acceptor impurity (boron) occupying a substitutional site in a silicon crystal. (b) Corresponding acceptor levels created in the silicon bandgap. [17]

A.2 P-N junction

If a p-type semiconductor and a n-type semiconductor are joined together they behave in a very different way producing what is generally known as a *PN Junction*. Due to the intrinsic different nature of the majority charge carriers, a very large density gradient exists along the connection line of the two semiconductors and a *contact potential* is formed. The potential intensity can be expressed by the formula;

$$\Delta V = V_T \cdot ln \frac{N_A N_D}{n_i^2} \tag{A.5}$$

where:

- $V_T = q_e \cdot k \cdot T$ is the thermal voltage;

- N_A is the concentration of acceptors in the p-type semiconductor;
- N_D is the concentration of donors in the n-type semiconductor;
- n_i is the thermal electrons/holes concentration.

The free charges will hence start to diffuse on the other side of the junction, attracted by the contact potential, till a situation of equilibrium is reached (shown in Fig. A.4) and certain area inside the junction is emptied of mobile charges, called *depletion region*.

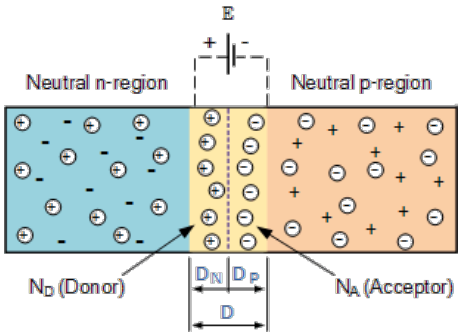


Figure A.4: Charge distribution in a PN junction after reaching equilibrium.

After equilibrium, the majority charge carriers on both sides of the junction encounter a potential barrier when trying to cross the depletion region, while the thermally generated minority carriers will be accelerated, causing a constant but very low current called *leakage current*. The shift of the conduction and valence energy bands in a PN junction is shown in Fig. A.5. If an external inverse potential is applied to the junction, the depletion region size will increase, obstructing the majority charge diffusion to the other side of the contact. This condition is called *inverse polarized diode*, opposed to the *direct polarized diode* where a direct potential is applied to the junction and the diffusion is facilitated.

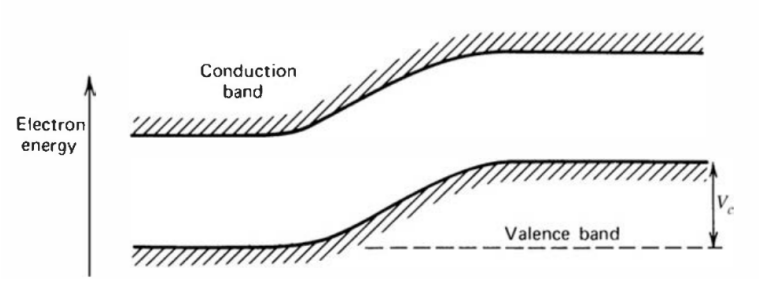


Figure A.5: Electron energy bands across the P-N junction. The curvature is reversed because an increase in electron energy corresponds to a decrease in conventional electric potential defined for a positive charge. [17]

Appendix B

Communication Protocols

Every time a system needs to transmit data to another device, it is necessary to establish a common set of rules for both the systems to avoid data corruption or misinterpretation. Communication protocols are formal descriptions of digital message formats and rules and they cover authentication, error detection/correction, and signaling. The protocols can be application-specific custom made or commercially available. This section will describe some of the protocols encountered in this thesis.

B.1 8b/10b

8b/10b protocol maps 8-bit symbols into 10-bit symbols to achieve DC-balance and bounded disparity. Since two extra-bits are added to the original word, some of the 8-bit symbols (256 possibilities) can be encoded in two different 10-bit words (1024 possibilities). Each 10-bit word is made such that the difference between the numbers of "0" and "1" is always 0 or ± 2; the *parity value* at the beginning of the trasnmission is conventionally set to -1 and at the end of each data transmission it can only assume the values -1 or $+1$. This mechanism is called *Running Disparity* (RD) and establishes a set of rules (shown in Tab. B.1) meant to achieve long-term DC-balance in the data-stream, granting that the number of "0" and "1" transmitted is almost identical.

Table B.1: Rules for running disparity

Previous RD	Disparity of code word	Disparity chosen	Next RD
-1	0	0	-1
-1	± 2	$+2$	$+1$
$+1$	0	0	$+1$
$+1$	± 2	-2	-1

To proceed to the final encoding, the 8 bit word is firstly divided in two sub-words of five and three bits ($HGFEDCBA \rightarrow HGF \quad EDCBA$) that are then encoded

using respectively 5b/6b and 3b/4b encoding as shown in Table B.2 and B.3. Some of the 10-bit symbols does not correspond to any 8-bit word, but are reserved for control messages. The list of control symbols, called *K-words*, is shown in Tab. B.4.

Table B.2: 5b/6b encoding table.

Input		**RD = −1**	**RD = +1**	**Input**		**RD = −1**	**RD = +1**
	EDCBA	abcdei			EDCBA	abcdei	
D.00	00000	100111	011000	D.16	10000	011011	100100
D.01	00001	011101	100010	D.17	10001	100011	
D.02	00010	101101	010010	D.18	10010	010011	
D.03	00011	110001		D.19	10011	110010	
D.04	00100	110101	001010	D.20	10100	001011	
D.05	00101	101001		D.21	10101	101010	
D.06	00110	011001		D.22	10110	011010	
D.07	00111	111000	000111	D.23	10111	111010	000101
D.08	01000	111001	000110	D.24	11000	110011	001100
D.09	01001	100101		D.25	11001	100110	
D.10	01010	010101		D.26	11010	010110	
D.11	01011	110100		D.27	11011	110110	001001
D.12	01100	001101		D.28	11100	001110	
D.13	01101	101100		D.29	11101	101110	010001
D.14	01110	011100		D.30	11110	011110	100001
D.15	01111	010111	101000	D.31	11111	101011	010100

Table B.3: 3b/4b encoding table.

Input		**RD = −1**	**RD = +1**	**Input**		**RD = −1**	**RD = +1**
	HGF	fghj			HGF	fghj	
D.x.0	000	0100	1011	D.x.4	100	0010	1101
D.x.1	001	1001		D.x.5	101	1010	
D.x.2	010	0101		D.x.6	110	0110	
D.x.3	011	0011	1100	D.x.P7	111	0001	1110
				D.x.A7	111	1000	0111

B.2 Aurora 64b/66b

Aurora is a lightweight link-layer protocol developed by Xilinx. It that can be used to move data point-to-point across one *Aurora channel*, that can be composed of one or more high-speed serial lanes. Aurora 64b/66b uses 64b/66b encoding, which is DC-balanced similarly to the 8b/10b encoding but is more efficient, with only 3.125% overhead.

The Aurora protocol specifies the following rules:

- data are packed in *frames* before being transferred to the Aurora channel;

Table B.4: 8b/10b K-words.

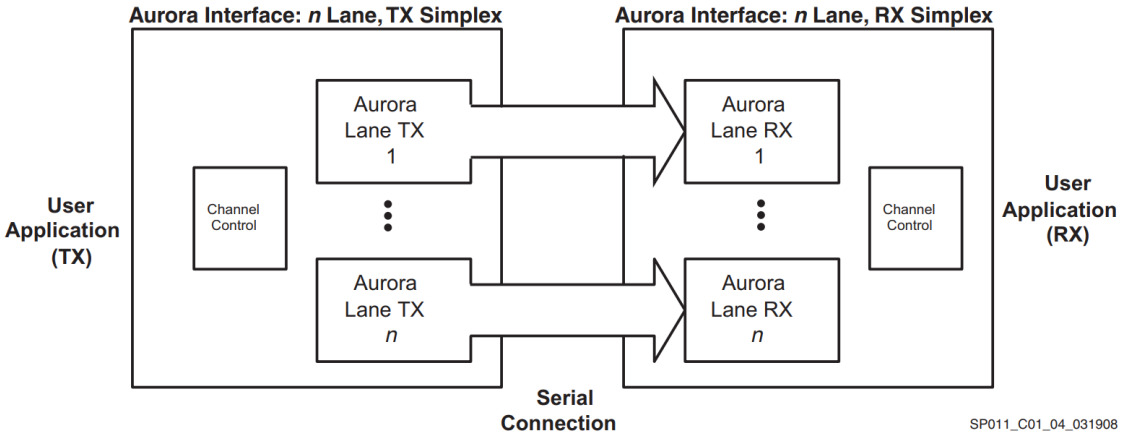
	Input			**RD = −1**	**RD = +1**
	DEC	HEX	HGF EDCBA	abcdei fghj	abcdei fghj
K.28.0	28	1C	000 11100	001111 0100	110000 1011
K.28.1	60	3C	001 11100	001111 1001	110000 0110
K.28.2	92	5C	010 11100	001111 0101	110000 1010
K.28.3	124	7C	011 11100	001111 0011	110000 1100
K.28.4	156	9C	100 11100	001111 0010	110000 1101
K.28.5	188	BC	101 11100	001111 1010	110000 0101
K.28.6	220	DC	110 11100	001111 0110	110000 1001
K.28.7	252	FC	111 11100	001111 1000	110000 0111
K.23.7	247	F7	111 10111	111010 1000	000101 0111
K.27.7	251	FB	111 11011	110110 1000	001001 0111
K.29.7	253	FD	111 11101	101110 1000	010001 0111
K.30.7	254	FE	111 11110	011110 1000	100001 0111

- the Aurora channel is shared by data frames, control information, clock compensation sequences and idles;
- frames can be of any length, and can have any format;
- frames can be interrupted at any time by flow control messages or idles.

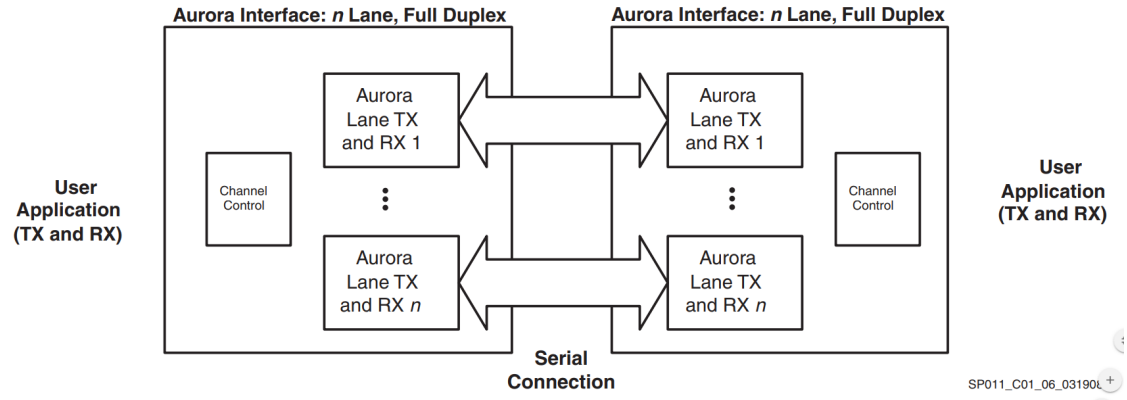
An Aurora communication can be implemented as *Simplex transmission*, meaning that each channel lane is either a receiver (RX) or transmitter (TX), or *Full Duplex transmission*, meaning each lane can transmit and receive data at the same time. All the data are packed in 64-bits codes called *blocks*; there are ten different Aurora block types and, since each lane can send only one block per time, they are given a transmission priority. All possible blocks in order of priority are:

1. *Clock Compensation block*: a special idle block used only in asynchronous transmissions. It is used to prevent data corruption due to small differences between the recovered clock and local reference clock and it must be sent for at least three consecutive blocks every ~ 10000 cycles;
2. *Not Ready Block*: this block is used by full-duplex Aurora transmissions during the initialization procedure of the channel;
3. *Channel Bonding Block*: a special idle block used by multi-lane Aurora transmissions. It is used to correct the eventual delays caused by the difference on the physical lanes lengths;
4. *Native Flow Control Block*: this blocks is sent to request a *Native Flow Control* operation, during which the receiver pauses the transmission of data blocks for a certain time specified by the command itself;

5. *User Flow Control Block*: this block is sent to request a *User Flow Control* operation, which is defined by the user application;
6. *User K-Block*: those blocks are not special control messages that are not decoded by the Aurora interface but are passed directly to the user;
7. *Data block*;
8. *Idle block*.



(a) Aurora multi-lane simplex transmission.



(b) Aurora multi-lane full-duplex transmission.

Figure B.1: Block diagram of an Aurora multi-lane simplex (a) and full-duplex (b) transmission. [61]

B.3 GBT

The GibaBit Transceiver (GBT) protocol was designed by CERN with the purpose of creating a high speed and reliable communication protocol capable of providing

a high amount of radiation tolerance. The transmission is meant to be implemented on a bidirectional point-to-point optical fiber. The GBT is optimized to interface high energy physics experiments, and it is structured to easily handle DAQ, TTC and Slow Commands. A GBT frame consists of a 120-bits word sent at the LHC bunch frequency of 40 MHz, for a total bandwidth of 4.8 Gbps. The 120 bits are divided in the following way: 4 bits are reserved for frame *Header*, 4 bits for *Slow Control*, 80 bits for *Data* and 32 bits for *Forward Error Correction*, as shown in Fig. B.2.

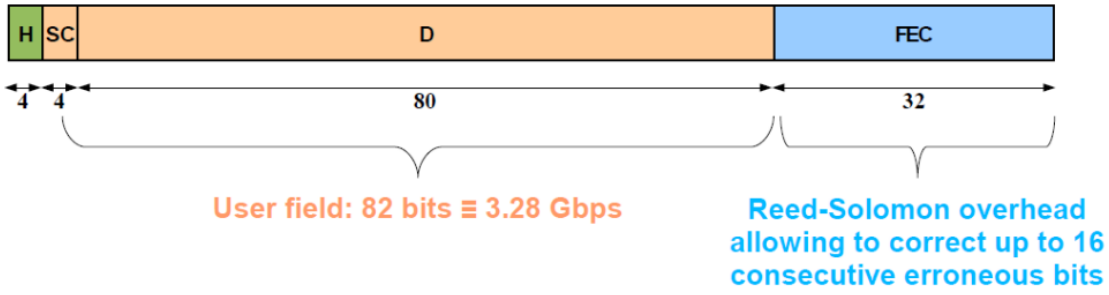


Figure B.2: GBT frame overview. H: header → 4 bits. SC: Slow Control → 4 bits. D: Data → 80 bits. FEC: Forward Error Control → 32 bits. [62]

To provide a high resistance to SEU errors due to the extremely irradiated environment, the GBT reserves part of the frame to error correction, meaning that the real data bandwidth is reduced to 3.36 Gbps, 73% of the total bandwidth. The code chosen for error correction is the *Reed-Solomon* code [63], capable of correcting a sequence of 16 consecutive wrong bits.

B.3.1 GBT FPGA

GBT FPGA is a firmware package meant to implement the GBT transmitter and receiver on FPGAs.

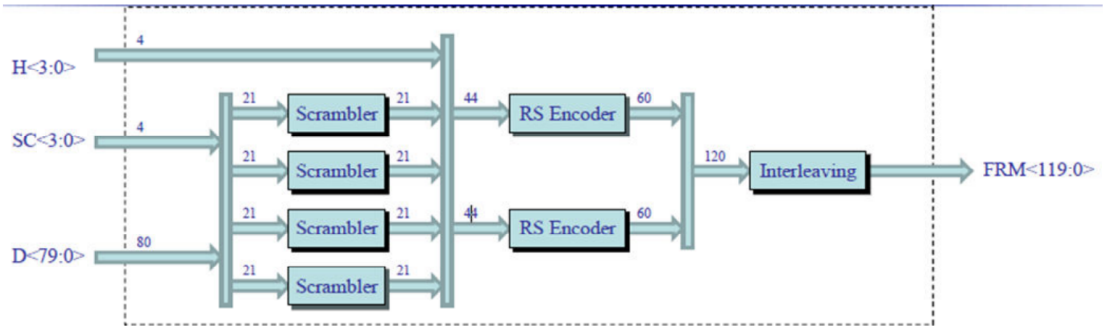


Figure B.3: GBT FPGA encoding scheme [62]

The package was developed by CERN, but a slightly modified version has been

built at Nikhef and BNL and is used by the FELIX and πLUP Projects. A diagram of the encoding scheme is shown in Fig. B.3. The GBT FPGA supports several FPGAs from Xilinx and Altera families and is optimized to occupy the least amount possible of resources.

B.4 Full Mode

The Full Mode protocol is a high speed communication protocol developed by the FELIX collaboration to fully exploit the GTx and GTH transceivers bandwidth. While the GBT speed is 4.8 Gbps with a 3.36 Gbps read data bandwidth, the Full Mode runs at 9.6 Gbps with a maximum payload throughput of 7.68 Gbps. Data are structured in packets, composed of 32-bits words, which are encoded using 8b/10b protocol. Since the Full Mode was born to be used in High Energy Physics applications, it supports the BUSY forwarding from the Front-end to the DAQ system (typically FELIX). Also, it offers flow control capabilities and provides means to detect errors through a 20-bits checksum that is calculated independently by the transmitter and the receiver at the end of each packet. The format of transmitted data is shown in Fig. B.4.

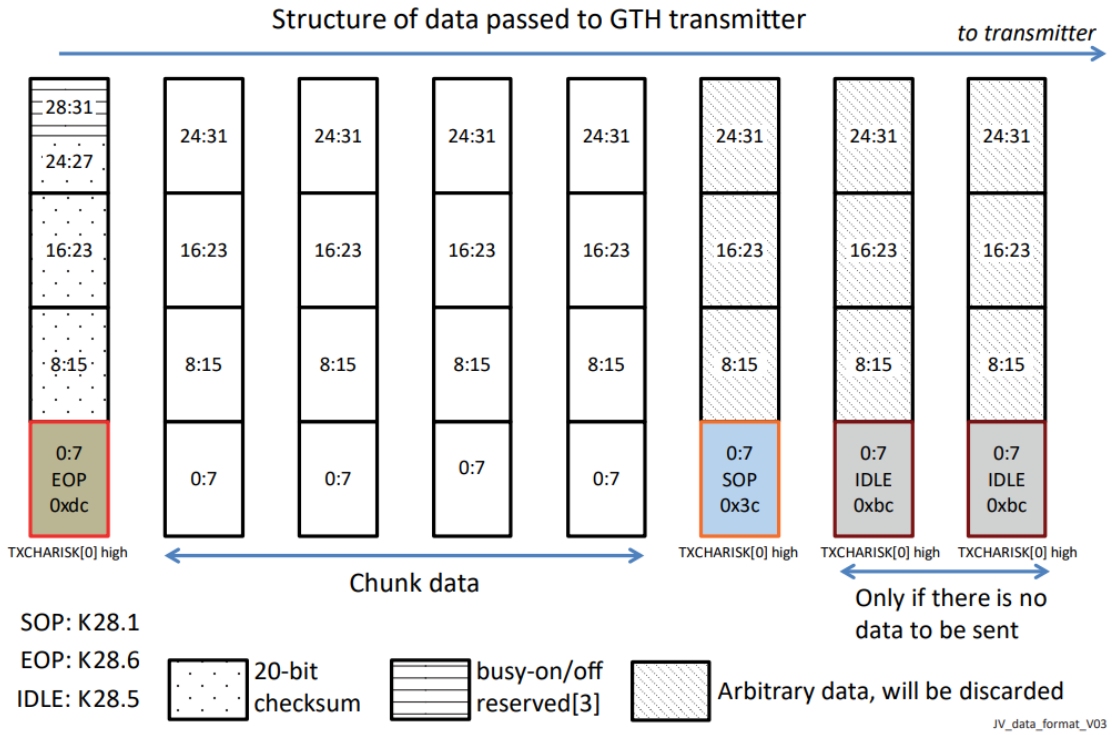


Figure B.4: : The format of the data transmitted between the serializer and deserializer of the Full mode wrapper. [64]

List of Figures

List of Tables

Bibliography

[1] L. Evans and Bryant P. Lhc machine. *JINST*, 3(S08001), 2008.

[2] Jean-Luc Caron. Cross section of LHC dipole.. Dipole LHC: coupe transversale. AC Collection. Legacy of AC. Pictures from 1992 to 2002., May 1998.

[3] P La Rocca and F Riggi. The upgrade programme of the major experiments at the Large Hadron Collider. *Journal of Physics: Conference Series*, 515(1):012012, 2014.

[4] [The ATLAS Collaboration]. The ATLAS Experiment at the CERN Large Hadron Collider. *Journal of Instrumentation*, 3(08):S08003, 2008.

[5] [ATLAS Collaboration]. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716(arXiv:1207.7214. CERN-PH-EP-2012-218):1–29. 29 p, Aug 2012. Comments: 24 pages plus author list (38 pages total), 12 figures, 7 tables, revised author list.

[6] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.

[7] ATLAS Collaboration. Technical Design Report for the ATLAS Inner Tracker Strip Detector. Technical Report CERN-LHCC-2017-005. ATLAS-TDR-025, CERN, Geneva, Apr 2017.

[8] *ATLAS inner detector: Technical Design Report, 1.* Technical Design Report ATLAS. CERN, Geneva, 1997.

[9] The ATLAS TRT collaboration. The ATLAS TRT Barrel Detector. *Journal of Instrumentation*, 3(02):P02014, 2008.

[10] M et al. Aleksa. ATLAS Liquid Argon Calorimeter Phase-I Upgrade Technical Design Report. Technical Report CERN-LHCC-2013-017. ATLAS-TDR-022, Sep 2013. Final version presented to December 2013 LHCC.

[11] *ATLAS tile calorimeter: Technical Design Report.* Technical Design Report ATLAS. CERN, Geneva, 1996.

[12] Michele Livan. Monitored drift tubes in atlas. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 384(1):214 – 218, 1996. BEAUTY '96.

[13] Konstantinos Nikolopoulos et al. Cathode strip chambers in ATLAS : Installation, commissioning and in situ performance. In *Proceedings, 2008 IEEE Nuclear Science Symposium, Medical Imaging Conference and 16th International Workshop on Room-Temperature Semiconductor X-Ray and Gamma-Ray Detectors (NSS/MIC 2008 / RTSD 2008): Dresden, Germany, October 19-25, 2008*, pages 2819–2824, 2008.

[14] Koichi Nagai. Thin gap chambers in atlas. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 384(1):219 – 221, 1996. BEAUTY '96.

[15] Giordano Cattani and the RPC group. The resistive plate chambers of the atlas experiment: performance studies. *Journal of Physics: Conference Series*, 280(1):012001, 2011.

[16] M. Moll G. Lindström and E. Fretwurst. Radiation hardness of silicon detectors - a challenge from high energy physics. *Nucl. Instrum. Meth*, A426:1–15, 1999.

[17] Glenn F. Knoll. *Radiation Detection and Measurement*. John Wiley & Sons, fourth edition, 2010.

[18] [ATLAS Collaboration] G Aad et al. Atlas pixel detector electronics and sensors. *Journal of Instrumentation*, 3(07):P07007, 2008.

[19] *ATLAS inner detector: Technical Design Report, 1*. Technical Design Report ATLAS. CERN, Geneva, 1997.

[20] Kevin Einsweiler. *ATLAS On-detector Electronics Architecture*. CERN, 3 edition, sep 2003.

[21] ATLAS Pixel Collaboration. *MCC-I2.1 Specifications*. CERN, 3 edition, dec 2003.

[22] T. P. Ma and Paul V. Dressendorfer. *Ionizing Radiation Effects in MOS Devices and Circuits*. John Wiley & Sons, first edition, 1989.

[23] F.W. Vaandrager and A.L. de Groot. Analysis of a biphase mark protocol with uppaal and pvs. *In Formal Aspects of Computing Journal*, 18(4):433–458, 2003.

[24] Gorry Fairhurst. https://www.erg.abdn.ac.uk/users/gorry/course/phy-pages/nrz.html, jan 2001.

[25] William R. Leo. *Techniques for Nuclear and Particle Physics Experiments: A How-to Approach*. Springer Science & Business Media, second edition, 2012.

[26] Norbert Wermes and G Hallewel. *ATLAS pixel detector: Technical Design Report*. Technical Design Report ATLAS. CERN, Geneva, 1998.

[27] B. Abbott et al. Production and Integration of the ATLAS Insertable B-Layer. *JINST*, 13(05):T05008, 2018.

[28] [ATLAS Collaboration] Barbero et al. Fe-i4 atlas pixel chip design. 01 2009.

[29] FE-I4 Collaboration. *The FE-I4B Integrated Circuit Guide*. CERN, 2.3 edition, dec 2012.

[30] J Albert et al. Prototype ATLAS IBL Modules using the FE-I4A Front-End Readout Chip. *JINST*, 7:P11010, 2012.

[31] M. Menouni et al. Design and measurements of SEU tolerant latches. In *Electronics for particle physics. Proceedings, Topical Workshop, TWEPP-08, Naxos, Greece, 15-19 September 2008*, pages 402–405, 2008.

[32] V. Chiochia et al. A double junction model of irradiated silicon pixel sensors for lhc. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 568(1):51 – 55, 2006. New Developments in Radiation Detectors.

[33] Erik van der Bij CERN European Laboratory for Particle Physics ECP division. S-link implementation guide, jun 1998.

[34] The ATLAS TDAQ Collaboration. The atlas data acquisition and high level trigger system. *Journal of Instrumentation*, 11(06):P06008, 2016.

[35] Xilinx. *Spartan-3 Generation FPGA User Guide*, ug331 (v1.8) edition, jul 2011.

[36] Xilinx. *Spartan-II FPGA Family Data Sheet*. Xilinx, jun 2013.

[37] Nico Giangiacomi. Readout board upgrade for the Pixel Detectors: reasons, status and results in ATLAS. *PoS*, EPS-HEP2017:790, 2018.

[38] N. Giangiacomi. Readout upgrade for the ATLAS Pixel Detector: Reasons, status and results. *Nuovo Cim.*, C41(1-2-2):76, 2018.

[39] N Schroer. The atlas ibl boc prototype. *Journal of Instrumentation*, 6(12):C12063, 2011.

[40] G Balbi, M Bindi, S P Chen, D Falchieri, T Flick, A Gabrielli, S Hauck, S C Hsu, M Kretz, A Kugel, L Lama, P Morettini, R Travaglini, and M Wensing. "the read-out driver" rod card for the insertable b-layer (ibl) detector of the atlas experiment: commissioning and upgrade studies for the pixel layers 1 and 2. *Journal of Instrumentation*, 9(01):C01044, 2014.

[41] N. Giangiacomi, G. Balbi, A. Damilano, D. Falchieri, A. Gabrielli, L. Lama, and R. Travaglini. New updates on the atlas rod board implementation for pixel layers 1 and 2. *IEEE Transactions on Nuclear Science*, 65(8):2338–2343, Aug 2018.

[42] G. Balbi, D. Falchieri, A. Gabrielli, L. Lama, N. Giangiacomi, and R. Travaglini. Atlas pixel detector rod card from ibl towards layers 2 and 1. *Journal of Instrumentation*, 11(01):C01021, 2016.

[43] ECFA High Luminosity LHC Experiments Workshop: Physics and Technology Challenges. 94th Plenary ECFA meeting. Nov 2013.

[44] ECFA High Luminosity LHC Experiments Workshop: Physics and Technology Developments Summary submitted to ECFA. 96th Plenary ECFA meeting. Jan 2015.

[45] CERN Collabortation. Hl-lhc high luminosity large hadron collider: The hl-lhc project.

[46] Collaboration ATLAS. Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment. Technical Report CERN-LHCC-2012-022. LHCC-I-023, CERN, Geneva, Dec 2012. Draft version for comments.

[47] ATLAS Phase-II Upgrade Scoping Document. Technical Report CERN-LHCC-2015-020. LHCC-G-166, CERN, Geneva, Sep 2015.

[48] ATLAS Collaboration. Technical Design Report for the ATLAS Inner Tracker Pixel Detector. Technical Report CERN-LHCC-2017-021. ATLAS-TDR-030, CERN, Geneva, Sep 2017.

[49] S. Terzo. The phase-ii atlas itk pixel upgrade. *Journal of Instrumentation*, 12(07):C07023, 2017.

[50] I. Perić, R. Eber, F. Ehrler, H. Augustin, N. Berger, S. Dittmeier, C. Graf, L. Huth, A.-K. Perrevoort, R. Phillipp, J. Repenning, D. vom Bruch, D. Wiedner, T. Hirono, M. Benoit, J. Bilbao, B. Ristić, and D. Muenstermann. Overview of hvcmos pixel sensors. *Journal of Instrumentation*, 10(05):C05021, 2015.

[51] Maurice Garcia-Sciveres. The RD53A Integrated Circuit. Technical Report CERN-RD53-PUB-17-001, CERN, Geneva, Oct 2017.

[52] William Keaton Balunas. ATLAS Trigger and Data Acquisition Upgrades for High Luminosity LHC. Technical Report ATL-DAQ-PROC-2016-019, CERN, Geneva, Nov 2016.

[53] PCI-sig. *PCI Express® Base Specification Revision 3.0*, nov 20180.

[54] Xilinx. *7 Series FPGAs Data Sheet: Overview*. DS180 (v2.6), month =.

[55] Xilinx. *UltraScale Architecture and Product Data Sheet: Overview*. DS890 (v3.5), month =.

[56] Xilinx. *7 Series FPGAs GTX/GTH Transceivers User Guide*. UG476 (v1.12.1), aug 2018.

[57] M. Barros Marin, S. Baron, S.S. Feger, P. Leitao, E.S. Lupu, C. Soos, P. Vichoudis, and K. Wyllie. The gbt-fpga core: features and challenges. *Journal of Instrumentation*, 10(03):C03021, 2015.

[58] N. Giangiacomi, K. Chen, and M. Joos. *Test Procedure of FLX-712 card*. CERN, 2018.

[59] Nico Giangiacomi, Alessandro Gabrielli, Fabrizio Alfonsi, Gabriele D'Amen, Gabriele Balbi, Davide Falchieri, Giuliano Pellegrini, and Riccardo Travaglini. General purpose readout board PiLUP: overview and results. *Submitted at TNS*.

[60] Xilinx. *Integrated Bit Error Ratio Tester 7 Series GTX Transceivers v3.0*. PG132, June 2016.

[61] Xilinx. *Aurora 64B/66B Protocol Specification*. SP011 (v1.3), October 2014.

[62] S Baron, J P Cachemiche, F Marin, P Moreira, and C Soos. Implementing the GBT data transmission protocol in FPGAs. 2009.

[63] S. Lin and D. J. Costello. *Error Control Coding*. Prentice Hall, second edition, 2004.

[64] ATLAS. Specifications for the 3 4 felix full mode link. Technical report, ATLAS FELIX Group, 2017.