Construction of the fast track reconstruction system (FTK) at the LHC-ATLAS experiment

LHC-ATLAS 実験における 高速飛跡再構成システム（FTK）の構築

December, 2017

Tomoya IIZAWA

飯澤 知弥

Construction of the fast track reconstruction system (FTK) at the LHC-ATLAS experiment

LHC-ATLAS実験における 高速飛跡再構成システム（FTK）の構築

December, 2017

Waseda University
Graduate School of Advanced Science and Engineering
Department of Pure and Applied Physics
Research on Particle Physics Experiment

Tomoya IIZAWA

飯澤 知弥

Abstract

Searches for new physics beyond the Standard Model and precise measurements of the Standard Model are performed at the European Organization for Nuclear Research (CERN) located in Geneva, Switzerland. The Large Hadron Collider (LHC) accelerates protons, and the protons collide at a center-of-mass energy on the TeV scale. The Higgs boson was discovered in 2012 by the ATLAS and CMS experiments, which was a significant milestone for particle physics. To perform further precise measurements and searches at the ATLAS experiment, one of the major challenges is to maintain the phase space at the data-taking stage in high luminosity environments where many simultaneous interactions per bunch crossing occur. The peak number of simultaneous interactions was more than 70 in 2017, which made it difficult to distinguish signals and background.

Proton bunches collide at a rate of 40 MHz in the LHC, and it is impossible to store all the collision information due to the finite size of the data storage. The ATLAS experiment implements a trigger system that performs event selection to reduce the data rate before recording the data to storage. Since events discarded at the trigger stage cannot be recovered anymore, the trigger must be designed to maintain a large phase space for signal events to provide opportunity for offline analyses while rejecting as many background events as possible. One method to improve the ATLAS trigger system is to make better use of track information. Track information is essential for particle identification and vertex reconstruction due to its high resolution and granularity. In the current trigger system, this information is not used efficiently since it takes a large amount of processing time and many CPUs to reconstruct the track information.

Fast TracKer (FTK) is a dedicated hardware system that performs fast track finding in the trigger system. The FTK is installed in the trigger system and reconstructs all tracks that have a transverse momentum greater than 1 GeV in the entire event with an average latency of less than 100 μs. The FTK is driven by more than 8000 ASIC chips and 2000 FPGAs. With the FTK, the trigger system can access the track information much earlier and obtain additional time, which enables to apply more sophisticated algorithms. For example, all the vertices in an entire event can be reconstructed, which enables a more precise correction of the kinematic variables in the high luminosity environment. In addition, it is possible to apply a lower threshold for the transverse momentum by reducing the event rate with better particle identification using track information, such as better b-jet tagging and the identification of hadronically decaying τs. This can retain more signal events and larger phase space for later offline analyses.

This dissertation describes all about the construction of the FTK system including hardware development and production, installation, and commissioning, as well as the performance and physics impact evaluation and the hardware emulation by the simulation. The performance of FTK is evaluated by constructing a dedicated simulation and it is shown that the FTK can achieve high track reconstruction performance in the trigger system in the high luminosity environment. The processing time is also estimated by the simulation, and it is found that the FTK can operate at sufficient speed within the trigger system. Hardware development has also been performed. The boards consisting the FTK were designed, tested, and produced in-house instead of using commercial boards since each board needs to have functions specific to the FTK. The functions were implemented in the firmware and has been developed via numerous tests. Mass production of the FTK input receivers was completed, and the yield rate after the quality control tests was 100%. The boards were integrated, installed, and commissioned to the ATLAS DAQ system. Therefore smooth operation of the FTK system was achieved with the test stand, and the first output of the entire FTK system with real data in the ATLAS DAQ was obtained at the end of 2016.

Acknowledgments

This thesis would not have been completed without great help from many people who were involved in my research.

Initially, I would like to express my special gratitude to my supervisor, Professor Kohei Yorita. He provided me many opportunities and huge amount of advice during my Ph.D research period based on his tremendous experience and deep thought on the research of particle physics. He also taught me what is important to live as a mature man. I believe I was able to learn very much not only as a researcher, but also as a mature man. I hope we can indulged in fruitful discussions and collaborate in the future.

I wish to express my appreciation to Professor Masakazu Washio and Jun Kataoka for their advice to write this thesis. They helped me during the process of Ph.D examination, and their advice filled important pieces of my Ph.D thesis.

Dr. Naoki Kimura is so helpful and it is impossible to express my gratitude to him in words. He was the closest to my age in the staff members of the Kohei Yorita laboratory and communicated with me very frankly. He advised me about research, the way of thinking, how to live in foreign countries as a researcher, and even several daily manners. When I was disappointed, he encouraged me by his own unique method. He was sitting next to me at CERN and indulged in useful discussion with me a lot of times. My research could not have been completed without his help. I hope we can perform some research together in the future.

I deeply appreciate Associate Professor Masashi Tanaka for his advice. His words pierced my heart many times and have strengthened my will. It improved me and gave me clues to think what is really important for physics and to have my own opinions, no matter what others say. I hope that I continue to get more such words and that we can collaborate in the future.

Further, I would like to extend my gratitude to Dr. Koji Ebina. He was from the same high school as myself and communicated friendly, but he was strict when required. He maintained the computing environment of the laboratory which made it possible to perform smooth research. We went to Izakaya (Japanese style bar) together many times and talked on several topics. When I was disappointed, he took me to Izakaya and encouraged me. It was one of the fond memories of my research life in the laboratory.

I am very grateful to Dr. Junji Naganoma. Though we did not perform research together, he gave me advice during meeting and daily communication. We went to Izakaya together which was a very memorable time for me. His attitude that he does what he thinks is important is one of the things that I would like to learn from him.

I deeply appreciate Dr. Yasuyuki Okumura. He was a postdoctoral researcher at the University of Chicago, working for FTK and staying at CERN when I started my Ph.D. I went to his office very frequently to discuss many issues. He always welcomed me and gave me appropriate advice at the field level. I believe those discussions and advice improved my way of logical thinking for research. He also introduced me several theses and textbooks which were all very interesting and easy to understand. I hope we can collaborate for research in the future.

My deep gratitude goes to Dr. Alberto Annovi who was the project leader of the FTK group. When I started working on FTK, I was a beginner and what I said would be wrong at instances, and my English was very poor. However, he understood what I would like to say even better than I, and he provided me many opportunities and helpful advice many times. It encouraged me very much. We went to restaurants

together and talked about several kinds of topics such as research, life, and even politics. This was a great experience and made me confident to perform research in a foreign country.

I appreciate Dr. Guido Volpi. He was responsible for the FTK offline software development when I started working on the FTK offline software. Though I went to his office every single day, he welcomed me, indulged in fruitful discussion and provided me advice. My work for the FTK offline software could not be accomplished without his advice.

I would like to extend my gratitude to Dr. Jahred Adelman. He was responsible for FTK offline software development after Guido when I was working as one of the coordinator of the FTK offline software. He was very logical and led the discussion in the right direction. We were involved in exciting discussions several times and I believe it improved the FTK offline software very much.

I would like to thank all of team members of the FTK offline software development, especially Dr. Daniel Britzger and Dr. Stefan Schmitt. They worked on optimization of the pattern bank which is very important to achieve good performance. When I was a coordinator of the FTK offline software, they discussed with me sincerely even though I was a Ph.D student.

I appreciate Dr. Calliope-Louisa Sotilopourou. We discussed many times about the development of firmware of Input Mezzanine Card, and she gave me useful advice on firmware development. This advice enabled to develop the firmware smoothly.

Dr. Pietro Albicocco developed IMs with Artix-7 FPGA and some part of the firmware together with me. His knowledge on the electric circuit board was of great help for the development. I respect his attitude that he does not sleep until he thinks that enough work has been completed. We also went to bar and restaurant together several times which was an exciting experience for me.

A hearty appreciation has to be extended to Dr. Mykhailo Lysovyi. He was a coordinator of the FTK online software, and he gave me advice and had discussion together on the development of the FTK online software. He was very kind, strict when required, and logical which enabled smooth development of the FTK online software.

Further, my deep appreciation goes to Dr. Lucian Stefan Ancu. He was a run coordinator of FTK group and kept saying what is important for smooth installation. He came to Japan on honeymoon and we went to Izakaya together, which was one of the warming memory for me.

I would like to express my gratefulness to Dr. Melvyn J. Shochet and Dr. Paola Jannetti. Mel was a project leader and Paola was a deputy project leader of the FTK group, and provided me advice when I started working for FTK. Paola hosted me when I was staying at Pisa, and made a comfortable environment to concentrate on research.

I cannot express my gratitude enough to Dr. Takashi Kubota. We discussed on the FTK online software and he gave me advice based on his experience on the development of the online software. He also extended sincere advice for writing my Ph.D thesis when I was concerned about it. His words encouraged me to complete Ph.D thesis.

I was very happy to work with Zihao Jiang and Karol Krizka. They were Ph.D students just like me and we frankly engaged in discussion many times. We went to restaurant together several times and enjoyed our time with a lot of alcohol. It gave me variety on my research life in CERN. I hope we can be involved in research together again some time in the future.

I would like to give my thanks to Mrs. Atsuko Sakamoto who is the secretary of Kohei Yorita Laboratory. She is always cheerful and very much helpful to perform desk work such as reservation of airline tickets and hotels for research trip, and some other tasks that are related to university. I can completely concentrate on research because of her great help.

A big thankfulness has to be extended to Yuki Sakurai and Takashi Mitani. They started their Ph.D course a year earlier than me. We lived together in the same house and shared a lot of time. It was a wonderful experience for me to share fun times as well as hard times with them. I hope we can collaborate and perform some research together in the future.

Further, I want to show appreciate to Toshiaki Kaji. He worked on the development of FTK simulation taking over from me. Though the code of the simulation was complicated, he worked and improved it. The performance of the FTK at the operation is owing to him.

I was very lucky to share time with Ryosuke Fuchi, Midori Fuchi (i.e. Mrs. Green), Kota Kasahara, Mutsuto Hagihara and Tatsumi Nitta. They gave me opportunities to discuss and to have leisure time at CERN. I would not have been able to complete my Ph.D without their existence.

A special thanks goes to members and graduates of Kohei Yorita laboratory, especially to Asuka Okamoto, Toshiharu Ohya, Shinichiro Sugita and Kaoru Fujisaki. We discussed quite a lot about research and also had some leisure time. Even after their leaving, we meet sometimes and talk on what they are doing. It is exciting and relaxed times for me and I am feeling refreshed to perform research.

I would like to give my thanks to Balazs Akos Nemeth and Szoke Zsolt. They shared a house with me together during the early times of my Ph.D course. Though I was anxious when I started life in foreign country, they warmly welcomed me and enjoyed together. It was an exciting experience for me that I joined the home party. I believe that we will be friends, no matter what we are doing and where we are.

Now is the turn to thank my friends, Shuya Ishii and Kei Kato. They have been my friends from my under graduate life. We met on occasions even after the completion of our bachelor course and it was relaxed time for me. Shuya was also a Ph.D student at Waseda University, so it was easy to meet him when I was in Japan. Kei had some tough times and I believe we encouraged each other.

Last but not the least, I would like to express my gratitude to my family, Hideto, Yuriko, Takafumi, Anna, Mi, Maro, and Eri. They have encouraged me always. So far, I have not been able to give you anything, but I hope I can do so in the future.

Contents

1 Introduction

1.1 Standard Model of Particle Physics

The particle physics is one of the academic fields trying to answer what matters are made of and what kind of interaction works among them. Now so-called Standard Model (SM) of particle physics [1–4] has been developed to answer these questions. In the SM, matters consist of six quarks and six leptons, and interaction is mediated by four kinds of gauge bosons. Mass of the particles is mediated by the Higgs boson via Higgs mechanism which breaks electro-weak symmetry [5–8]. Table 1, 2 summarize the particles which constitute matter and mediate interaction in the SM, respectively.

	First Generation	Second Generation	Third Generation
Quark	up	charm	top
	down	strange	bottom
Lepton	electron-neutrino	muon-neutrino	tau-neutrino
	electron	muon	tau

Table 1: Particles which constitute matter in the SM.

Interaction	Gauge Boson
Strong	gluon
Weak	W, Z
Electro-Magnetic	photon

*Mass is mediated by scalar Higgs boson.

Table 2: Particles which mediate interaction in the SM.

The SM describes three kinds of interactions which are strong interaction, weak interaction and electro-magnetic interaction. The strong interaction is described by the Quantum Color Dynamics (QCD) which is based on $SU(3)_c$ gauge invariance, and the Electro-Weak interaction is based on $SU(2)_L \otimes U(1)_Y$ gauge invariance, and then the SM is described by $SU(3)_c \otimes SU(2)_L \otimes U(1)_Y$. The SM is based on the relativistic quantum field theory for which the Lagrangian describes characteristic of the matters and the interactions.

Scalar

Spin-0 relativistic particle is described by excitation of scalar field $\phi(x)$. This follows Klein-Gordon equation expressed as Eq. 1.

$$(\partial^\mu \partial_\mu + m^2)\phi = 0, \tag{1}$$

here it is described with natural system of units, and $\partial^\mu, \partial_\mu$ stand for

$$\partial_\mu \equiv \left(\frac{\partial}{\partial t}, \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right), \tag{2}$$

$$\partial^\mu \equiv \left(\frac{\partial}{\partial t}, -\frac{\partial}{\partial x}, -\frac{\partial}{\partial y}, -\frac{\partial}{\partial z}\right). \tag{3}$$

Eq. 1 is required to satisfy the Euler-Lagrange equation.

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi)} - \frac{\partial \mathcal{L}}{\partial \phi} = 0. \tag{4}$$

Thus the Lagrangian which gives Klein-Gordon equation is as following,

$$\mathcal{L}_S = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2}m^2\phi^2. \tag{5}$$

Fermion

Spin-$\frac{1}{2}$ relativistic particle is described by the Lagrangian which gives Dirac equation described as Eq. 6.

$$i\gamma^\mu \partial_\mu \psi - m\psi = 0. \tag{6}$$

The Lagrangian is expressed as

$$\mathcal{L}_D = i\overline{\psi}\gamma^\mu \partial_\mu \psi - m\overline{\psi}\psi. \tag{7}$$

Vector

Maxwell equation for electro-magnetic field $A^\mu = (\phi, \vec{A})$ is expressed as

$$\Box^2 A^\nu - \partial^\nu(\partial_\mu A^\mu) = \partial_\mu F^{\mu\nu} = j^\nu, \tag{8}$$

$$F^{\mu\nu} \equiv \partial^\mu A^\nu - \partial^\nu A^\mu, \tag{9}$$

where j^μ stands for four-current. Corresponding Lagrangian is

$$\mathcal{L}_{EM} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - j^\mu A_\mu. \tag{10}$$

Considering free field, it becomes

$$\mathcal{L}_{EM} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{11}$$

Interaction

Interaction is introduced by requiring invariance for local gauge transformation,

$$\psi(x) \rightarrow \psi'(x) = e^{iq\chi(x)}\psi(x). \tag{12}$$

Requiring this gauge invariance for U(1) transformation, the Lagrangian of fermion expressed by Eq. 7 is transformed into

$$\mathcal{L} \rightarrow \mathcal{L}' = \mathcal{L} + iq\chi(x)\overline{\psi}\gamma^{\mu}(\partial_{\mu}\chi)\psi. \tag{13}$$

Thus the Lagrangian for free Dirac particle is not invariant for U(1) local gauge transformation. Gauge invariance is recovered by replacing ∂_{μ} with covariant derivative D_{μ} defined as

$$\partial_{\mu} \rightarrow D_{\mu} = \partial_{\mu} + iqA_{\mu}, \tag{14}$$

where A_{μ} is newly introduced field. To remove extra term in Eq. 13, A_{μ} is introduced as

$$A_{\mu} \rightarrow A'_{\mu} = A_{\mu} - \partial_{\mu}\chi. \tag{15}$$

Thus gauge invariant Lagrangian is expressed as

$$\mathcal{L} = i\overline{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\overline{\psi}\psi - q\overline{\psi}\gamma^{\mu}A_{\mu}\psi. \tag{16}$$

This can be interpreted as the Lagrangian of free Dirac particle is not gauge invariant, but it becomes gauge invariant if interaction is taken into account. Including kinetic term of photon, the Lagrangian of Quantum Electro Dynamics (QED) is expressed as

$$\mathcal{L}_{QED} = i\overline{\psi}\gamma^{\mu}\partial_{\mu}\psi - m\overline{\psi}\psi - q\overline{\psi}\gamma^{\mu}A_{\mu}\psi - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{17}$$

To state clearly, the first term represents for the kinetic term of fermion, the second term for potential for mass energy, the third term for interaction between fermion and photon, and the forth term for the kinetic term of photon. The weak interaction and the strong interaction can be introduced by requiring invariance for $SU(2)_L$, $SU(3)_c$ local gauge transformation, respectively.

Higgs Mechanism

In actually, symmetry of the Lagrangian is broken for local gauge transformation due to mass term of particle field. To achieve the symmetry, mass term is generated by a modification of stable vacuum via spontaneous symmetry breaking. The Higgs mechanism introduces a complex scalar doublet to make spontaneous symmetry breaking.

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \tag{18}$$

The Lagrangian for this complex doublet is expressed as

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^\dagger (\partial^\mu \phi) - V(\phi), \tag{19}$$

where the Higgs potential $V(\phi)$ is described as

$$V(\phi) = \mu^2 \phi^\dagger \phi + \lambda(\phi^\dagger \phi)^2, \tag{20}$$

where the first term corresponds to mass and the second term corresponds to self-interaction of the scalar field. λ is positive number since the potential must have the minimum value to form stable vacuum, while μ^2 can take both positive value and negative value. In the case of $\mu^2 < 0$, the minimum value is taken at a set of points of $\phi \neq 0$ expressed as

$$|\phi|^2 = \frac{v^2}{2} = -\frac{\mu^2}{2\lambda}. \tag{21}$$

Here vacuum expectation value is described as

$$\langle 0|\phi|0\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}. \tag{22}$$

The Lagrangian which includes spontaneous symmetry breaking introduces one scalar field which has mass and three massless Goldstone boson. The three massless Goldstone bosons are absorbed to longitudinal mode of $W^\pm$ and Z bosons by choosing appropriate gauge (unitary gauge). With the unitary gauge, the doublet is expressed as

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{23}$$

where h represents the Higgs field to stress it is "physical real field". To make the Lagrangian symmetry for $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ local gauge transformation, introduce interaction with weak gauge boson by replacing ∂_μ with covariant derivative D_μ which is defined as

$$\partial_\mu \rightarrow D_\mu = \partial_\mu + i g_W \vec{T} \cdot \vec{W}_\mu + i g' \frac{Y}{2} B_\mu, \tag{24}$$

where $\vec{T}$ represents three generators of SU(2) group, expressed as $\vec{T} = \frac{1}{2}\sigma$, Y represents hypercharge, and W, B represent field of gauge boson. By choosing the unitary gauge, mass term of the Lagrangian can be calculated as

$$\frac{1}{8} v^2 g_W^2 (W_\mu^{(1)} W^{(1)\mu} + W_\mu^{(2)} W^{(2)\mu}) + \frac{1}{8} v^2 (A_\mu, Z_\mu) \begin{pmatrix} 0 & 0 \\ 0 & g_W^2 + g'^2 \end{pmatrix} \begin{pmatrix} A^\mu \\ Z^\mu \end{pmatrix}, \tag{25}$$

where Z and A represent Z boson and photon respectively, which are generated by mixing of $W_\mu^{(3)}$ and B_μ. Thus mass of bosons are extracted as

$$m_H = 2\lambda v^2, m_A = 0, m_W = \frac{1}{2} g_W v, m_Z = \frac{1}{2} \sqrt{g_W^2 + g'^2}. \tag{26}$$

The Higgs mechanism for $\mathrm{SU(2)_L} \otimes \mathrm{U(1)_Y}$ local gauge transformation also introduces mass of fermions. If spontaneous symmetry breaking occurs and the unitary gauge is chosen, the Lagrangian for coupling between scalar field and fermion field is described as

$$\mathcal{L} = \frac{g_Y}{\sqrt{2}} v (\overline{u}_L u_R + \overline{u}_R u_L) - \frac{g_Y}{\sqrt{2}} h (\overline{u}_L u_R + \overline{u}_R u_L), \tag{27}$$

where g_Y is so called "Yukawa coupling constant" which represents couplings between scalar field and fermion field expected as

$$g_Y = \frac{\sqrt{2} m_f}{v}. \tag{28}$$

By far now the SM reproduces the results of experiments with very high precision. The theory and the experiment have been developed mutually in the field of the particle physics. But there are some incompleteness in the SM such as the existence of the mass of the neutrino, the existence of dark matter, and the gravity is not included in the SM. To construct more accurate and fundamental theory beyond the SM is one of the most important tasks in the field of the particle physics.

1.2 High Energy Frontier Experiment

To obtain further knowledge on the particle physics, precise measurements of the SM and searches for physics beyond the SM are performed with the Large Hadron Collider (LHC) at European Organization of Nuclear Research (CERN) built in Switzerland and France. Protons are accelerated and collides at a center of mass energy of TeV scale. Particles which are generated by interactions of the protons are detected by a combination of the detectors. Many break-through results have been obtained such as discovery of the Higgs boson with a mass of 125 GeV by the ATLAS and the CMS experiments [9, 10]. Precise measurement of the property of the Higgs boson is one of the crucial tasks for the particle physics to obtain any clue for new physics.

At the LHC, the proton bunches collide at a rate of 40 MHz, and it is impossible to store all the collision information due to finite size of data storage. The ATLAS experiment implements trigger system which performs event selection to reduce data rate before recording data to the storage. Figure 1 shows a schematic overview of the current ATLAS trigger system with the Fast TracKer (FTK). The trigger system consists of two stages which are the Level-1 (L1) trigger and the High Level Trigger (HLT). Since events discarded at the trigger stage cannot be used anymore, it is critical to implement efficient selection which keeps as many signal events as possible while reducing background events.

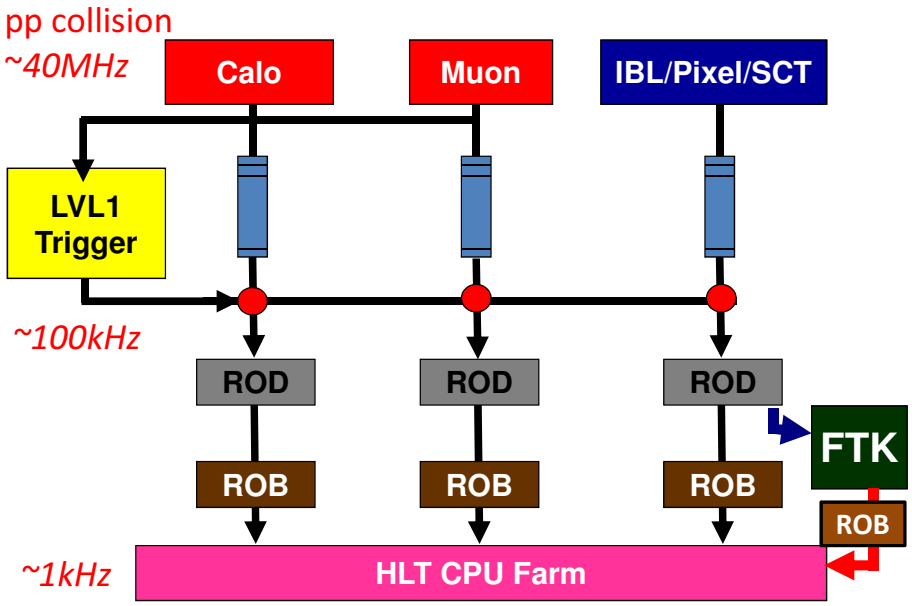

Figure 1: An overview of current ATLAS trigger system with the FTK.

One of the major challenges for the trigger system in high luminosity environment is to cope with the increase of multiple proton-proton interactions per bunch crossing (pile-up) which makes it difficult to distinguish signals from backgrounds, and makes resolution worse. A schematic view of mechanism of the pile-up and an event display of the bunch crossing with 78 pile-ups are shown in Fig. 2. In 2017, the peak number of simultaneous interactions is more than 70. A straight forward way to reduce event rate in high pile-up environment is applying higher threshold on transverse momentum of each object, while it also reduces signal events and phase space for later analysis. For the case of the Higgs boson, its mass is measured with 125 GeV so that the decay products have relatively lower transverse momentum. Thus raising threshold significantly reduces signal events and makes it difficult to perform precise measurement

since it requires a lot of signal statistics. It is favorable to reduce data rate without raising threshold at the trigger to keep phase space for signal events. It is as well the case for the searches for new particles with a mass around 100 GeV.

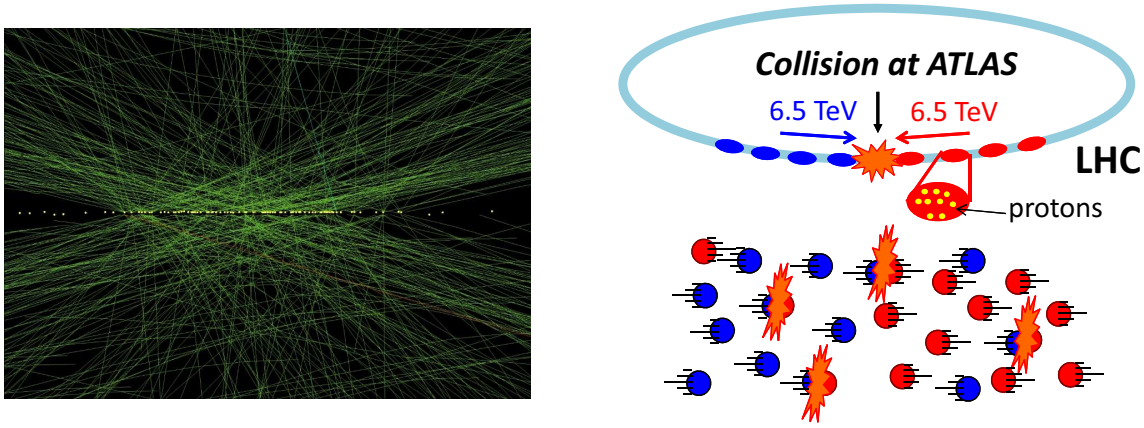

Figure 2: An event display of the bunch crossing with 78 pile-ups [11] (left). A schematic view of the pile-up mechanism (right).

In that situation, efficient use of track information has large potential to improve decision at the trigger significantly. Since the track information has high resolution and granularity and it is measured by the detector placed at the most inner part of the ATLAS detector complex, it provides precise information near collision point. Currently track information is not used at the L1 trigger, and it is used at the HLT but only limited region since it takes much time and a lot of CPUs to calculate track information.

The FTK is a newly installed electric hardware system that reconstructs tracks rapidly for an entire event at the trigger system. It is implemented between the L1 trigger and the HLT, receives hit information and calculates track information for entire detector region for all events which pass the L1 trigger. With the FTK system, the HLT does not need to calculate track information, and then can use time and track information provided by the FTK for more sophisticated algorithms which can reduce data rate without raising threshold.

1.3 Utility of Track Information and Expected Benefit to Install the FTK

Track information is robust for pile-up as well as useful for identification of physics objects. Here ATLAS coordinate system is introduced first, then utility of track information for vertex reconstruction, τ identification, and b-jet tagging are described. In addition, expected benefit at the HLT achieved with the FTK track information is also shown.

ATLAS Coordinate System

The ATLAS coordinate system is defined as right handed Cartesian coordinate system (x, y, z) with its origin at the nominal IP in the center of the detector. z-axis is defined as in along with the beam line, x-axis from the IP to the center of the LHC ring, and y-axis points upwards. A cylindrical coordinate system (r, ϕ, z) and a polar coordinate system (r, θ, ϕ) are defined in the transverse plane as

$$r = \sqrt{x^2 + y^2}, \tag{29}$$

$$\phi = \arctan \frac{y}{x}, \tag{30}$$

$$\theta = \arctan \frac{r}{z}, \tag{31}$$

where the azimuthal angle ϕ is defined with respect to x-axis between -π and π, where $\phi > 0$ represents positive x-axis, and the polar angle θ is defined with respect to z-axis between 0 and π, where $\theta > 0$ represents positive z-axis. The rapidity (y_{rap}) is defined as

$$y_{rap} = \frac{1}{2} \ln \frac{E + p_z}{E - p_z}, \tag{32}$$

where E and p_z are the energy and the momentum along the beam direction, respectively. The pseudo-rapidity (η) which is equal to the rapidity in ultra-relativistic approximation is also often used. It is defined as

$$\eta = -\ln \tan \frac{\theta}{2}. \tag{33}$$

The distance in the pseudo-rapidity and azimuthal angle space (ΔR) is defined as

$$\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2}, \tag{34}$$

where $\Delta\eta$ and $\Delta\phi$ are the distance in the pseudo-rapidity and the azimuthal angle space, respectively.

The transverse and longitudinal impact parameters, d_0 and z_0, are the distance of closest approach of the track to the primary vertex point in the $r - \phi$ plane and in the longitudinal plane, respectively. The track parameters are visualized in Fig. 3.

The transverse momentum and energy are defined in $x - y$ plane as

$$p_T = p \times \sin\theta, \tag{35}$$

$$E_T = E \times \sin\theta. \tag{36}$$

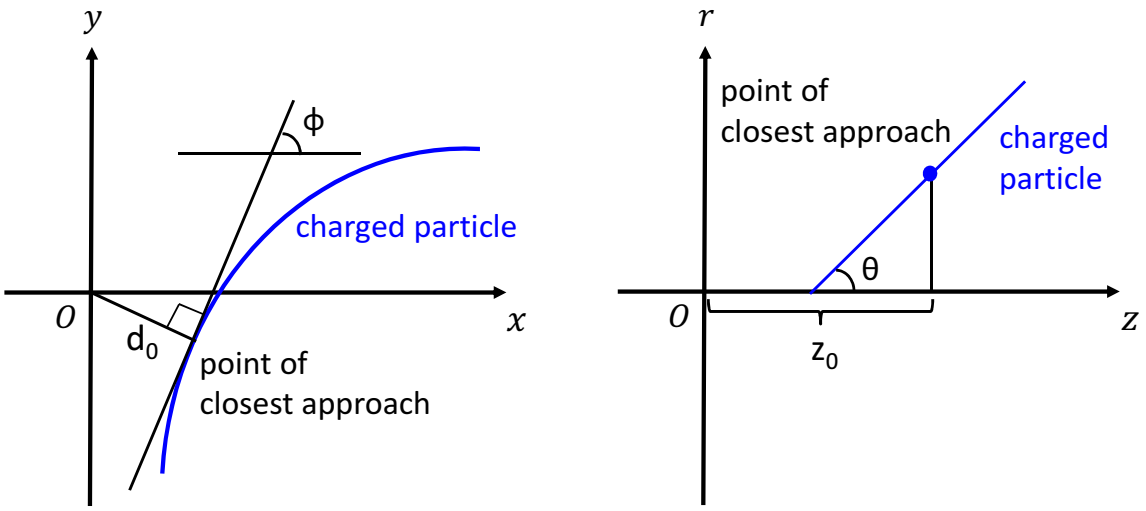

Figure 3: Track parameters in $x-y$ plane (left) and $r-z$ plane (right).

1.3.1 Vertex Reconstruction

Vertex reconstruction [12] is critical at high pile-up environment. There are several pile-up dependent corrections based on the number of vertices. The position of primary vertex is important to separate tracks from signals and backgrounds. For example, Jet Vertex Taggar (JVT) [13] is used at the offline analysis to distinguish jets from signal events and pile-up interactions, which utilizes information whether tracks in jets are coming from primary vertex or pile-up interactions. In addition,

Vertex reconstruction needs information of entire detector region and very much CPU intensive so that it is challenging for trigger to perform event-level vertex reconstruction. The FTK provides all tracks in an event within several tens of μs which enables the HLT to reconstruct all vertices in the event. Primary vertex finding will become easier by searching around the track which has the highest transverse momentum in the event. The FTK also provides hit cluster information which helps to find a seed of vertices.

Vertex reconstruction performance is evaluated by comparing between vertices reconstructed with the offline algorithm using the offline tracks and with the algorithm which can be implemented at the HLT using the FTK tracks. The offline algorithm utilizes Adaptive Vertex Fitter [14] while simpler algorithm of Fast Vertex Fitter [15] would be fit to reconstruct vertices in the time-limited trigger system.

Figure 4 shows performance of vertex reconstruction by Adaptive Vertex Fitter with the offline tracks (offline algorithm) and by Fast Vertex Fitter with the FTK tracks (online algorithm) as well as truth vertex. Similar performance can be obtained between vertices reconstructed by the offline algorithm and the online algorithm. As a consequence, it is validated that 95% of vertices reconstructed by the offline algorithm can be reconstructed by the online algorithms, and processing time of the online algorithm is less than 1 ms which corresponds to the very beginning of the entire HLT processing.

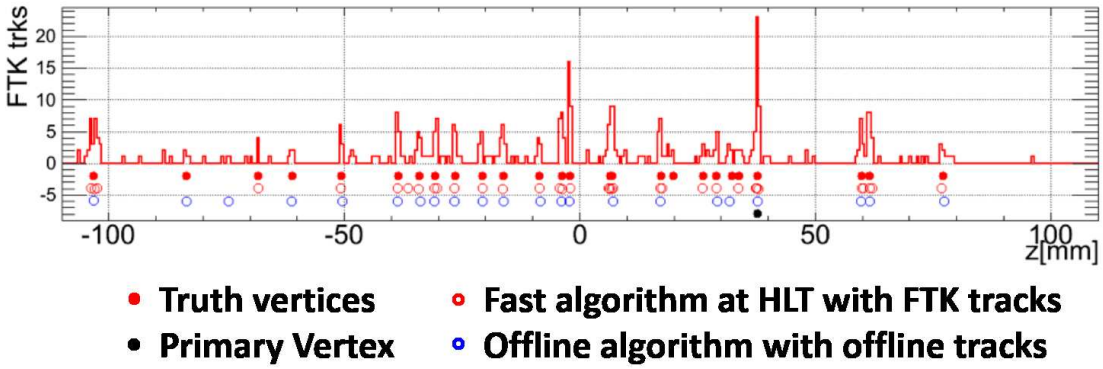

Figure 4: Vertex reconstruction performance for an event.

1.3.2 τ Identification

τ is the heaviest known lepton which plays an important role such as the decay of the Higgs boson. It is also expected to be included in the final state of several physics beyond the SM. τ can decay into other leptons or hadrons since it is the third generation lepton and has mass of 1.78 GeV. Track information is important for identification of hadronically decaying τ lepton (τ_{had}) [16].

One of the challenges to identify τ_{had} is separation with QCD jets since decay signature of τ_{had} is similar to that of QCD jets as shown in Fig. 5, and production cross section of QCD jets is extremely larger than τ_{had}s. To separate τ_{had}s from QCD jets, decay topology is used. τ_{had} typically decays to one or three charged hadrons which are called 1-prong and 3-prong, respectively. Core cone ($0 < \Delta R < 0.2$) and isolation region ($0.2 < \Delta R < 0.4$) are defined around τ_{had} candidate. τ_{had} has one or three charged particles in the core cone and no or very little activity in isolation region, while QCD jet has wider activity. The information of the number of primary vertex and the average number of interactions per bunch crossing are also used for identification of τ_{had}.

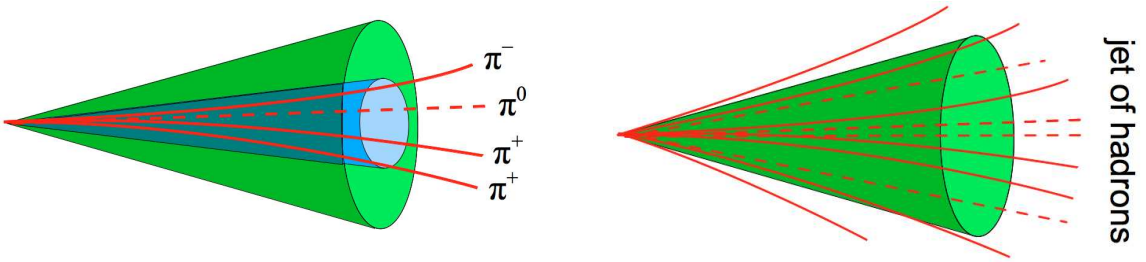

Figure 5: A schematic view of jet originated from hadronically decaying τ lepton (left) and from QCD jets (right).

Figure 6 left (right) shows track multiplicity distribution in the core cone (isolation region). A signal sample of gluon fusion $h \rightarrow \tau_{\text{had}}\tau_{\text{had}}$ is shown in blue and a multijet QCD background sample is shown in red. Both signal and background samples are generated with the average number of interactions is 40 at $\sqrt{s}$ = 14 TeV. At the L1 trigger, τ_{had} candidates are required to be less than 4 GeV of calorimetric isolation energy and above a transverse momentum threshold of 12 GeV. The direction of τ_{had} candidate is defined as the direction of the leading track within ΔR < 0.2 from L1 τ_{had} candidate. Core cone and isolation region are defined around the direction of the leading track. Tracks considered to calculate in multiplicity

is required to have $p_T > 1$ GeV, $|d_0| < 2$ mm, $|z_0| < 150$ mm, $\chi^2/N_{dof} < 4$, and the difference between $|z_0|$ of the track and $|z_0|$ of the leading track should be less than 2 mm. For signal, only L1 τs matching within $\Delta R < 0.1$ to truth τ_{had}s are considered. The distribution of QCD multijet events in the isolation region is broader than the one of signal events.

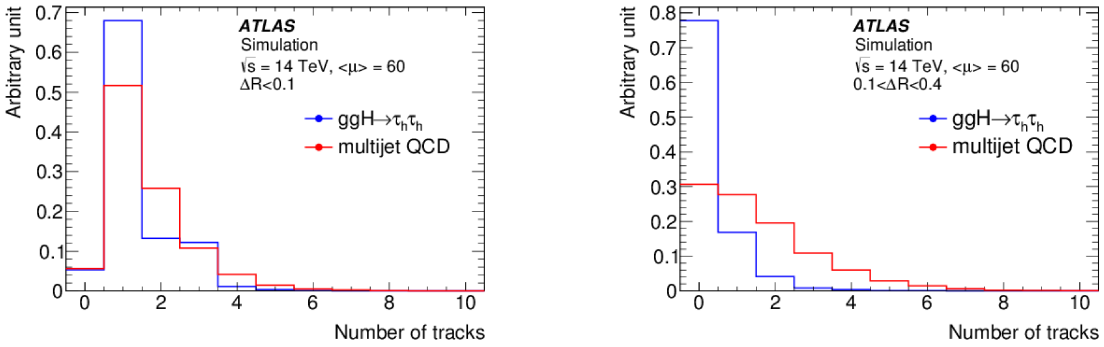

Figure 6: Track multiplicity distribution in the core cone (left) and isolation region (right). A signal sample of gluon fusion $h \rightarrow \tau_{had}\tau_{had}$ is shown in blue and and a multijet QCD background sample is shown in red.

Figure 7 shows τ identification efficiency as a function of offline-τ transverse momentum when applying FTK selection (blue) and calorimeter cluster based selection (red) at the HLT. The efficiency is defined as the fraction of L1 τ matched to a true τ_{had} decay and to an offline τ identified with the offline loose τ identification. FTK selection at the HLT requires $N_{tracks} \leq 3$ in the core cone and $N_{tracks} \leq 2$ in the isolation region, while the calorimeter cluster based selection relies on information of energy deposit in calorimeter such as total energy and the fraction of energy deposition at each calorimeter. With calorimeter cluster based selection, efficiency rises around 20 GeV and becomes flat at around 30 GeV, while FTK selection achieves high efficiency and flat distribution at that region, thus lower threshold can be applied.

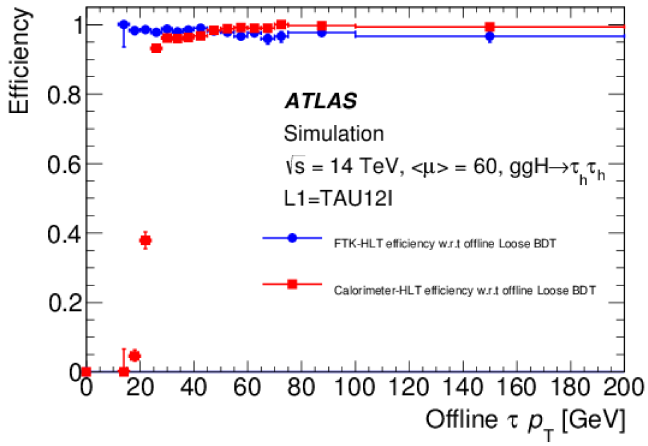

Figure 7: τ identification efficiency as a function of offline-τ p_T when applying FTK selection (blue) and calorimeter cluster-based selection (red) at the HLT.

1.3.3 B-jet Tagging

B-quark is third generation quark and it has the highest branching ratio of the Higgs decay and also expected to be included in the final state of several new physics processes. The identification of jets originated from bottom quarks (b-jet) [17], b-jet tagging, is challenging since production cross section of jets originated from light quarks or gluons (light-jet) is extremely larger. Since B-hadron has relatively long life time ($\sim$ 1.5 ps, $c\tau \sim 450$ μs), the decay products has characteristics to have displaced vertices and large impact parameter values. Characteristics of jets and b-jet is shown in Fig. 8.

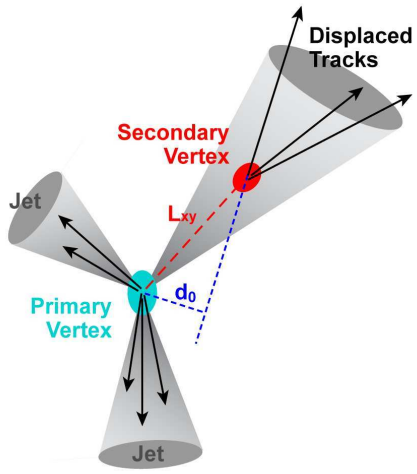

Figure 8: Characteristics of light jets and b-jet [18].

Thus track information such as signed transverse impact parameter and its significance is crucial for the b-jet tagging. The sign of the transverse impact parameter is defined positive (negative) if the point of the closest approach of the track to primary vertex is in front (behind) of the primary vertex with respect to the jet direction. The transverse impact parameter significance is defined as transverse impact parameter divided by its associated uncertainty.

Figure 9 and 10 show transverse impact parameter and its significance, respectively. Tracks are associated to light-flavor (black) and heavy-flavor (red) jets. The solid lines show the distribution for offline tracks, whereas the points show tracks which are quickly re-fitted at the HLT using the FTK tracks as seeds. The distribution is almost same between offline tracks and re-fitted FTK tracks for both transverse impact parameter and its significance, and b-jets tend to have more positive transverse impact parameter and its significance than those of light jets.

Figure 11 shows several working points of b-jet trigger for the HLT rate versus event level signal efficiency of $t\bar{t}H(H \rightarrow b\bar{b})$. Pre-selection is applied for each working point requiring four L1 jets with transverse momentum above 20 GeV. The black points show working points without FTK track information at the HLT, while red points show the option with FTK track information. Track finding can run with looser HLT jet thresholds without adding loads on HLT processors with the FTK. Efficiencies are quoted with respect to inclusive signal and include L1 efficiency. The trigger names specify jet multiplicities, transverse momentum thresholds and b-tagging algorithms. The lower transverse momentum threshold can be applied utilizing FTK track information since the rate is reduced by requiring tighter b-jet tagging.

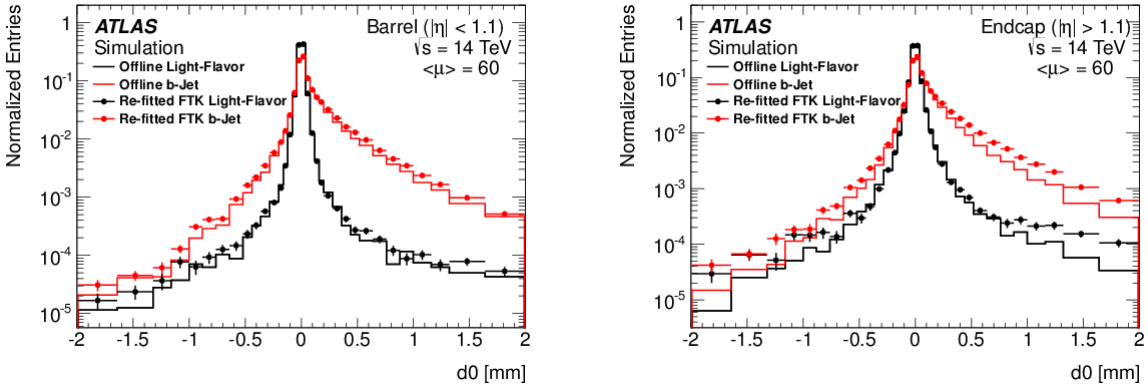

Figure 9: The transverse impact parameter distributions are shown. The left plot is detector barrel region ($|\eta| < 1.1$) and the right plot is detector endcap region ($|\eta| > 1.1$). Tracks associated to light-flavor (black) and heavy-flavor (red) jets are shown. The solid lines stand for offline tracks, whereas the points show re-fitted FTK tracks.

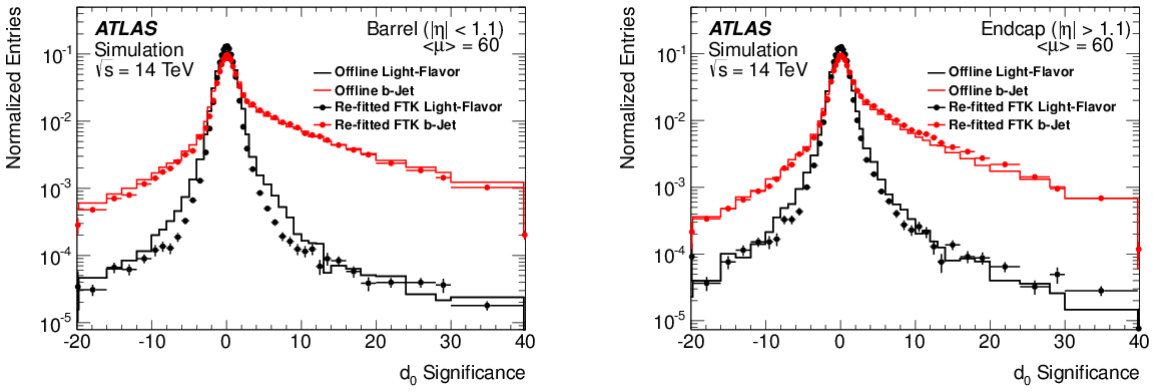

Figure 10: The transverse impact parameter significance distributions are shown. The left plot is detector barrel region ($|\eta| < 1.1$) and the right plot is detector endcap region ($|\eta| > 1.1$). Tracks associated to light-flavor (black) and heavy-flavor (red) jets are shown. The solid lines stand for the offline tracks, whereas the points show re-fitted FTK tracks.

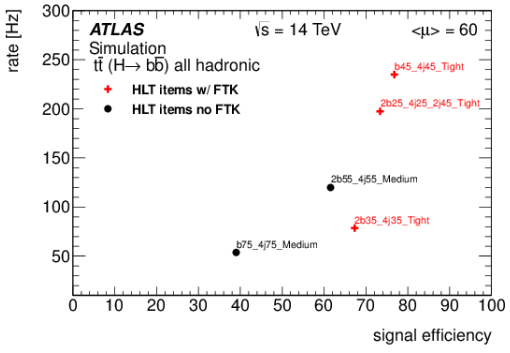

Figure 11: Several working points of b-jet trigger for the HLT rate versus event level signal efficiency of $t\bar{t}H(H \rightarrow b\bar{b})$. The black points represent working points without FTK track information at the HLT, while red points represent the option with FTK track information.

Consequently track information and additional time provided by the FTK enable the HLT to implement more sophisticated algorithms to keep threshold lower, and to keep more signal events and larger phase space for wide range of physics processes at high pile-up environment.

This dissertation describes all about construction of the FTK. Sec. 2 describes the details of the LHC and the ATLAS detector. Sec. 3 explains working principle of the FTK and its hardware. Expected performance of the FTK studied by simulation is shown in Sec. 4. In Sec. 5 the detail of the development and production of the Input Mezzanine Card is described. Sec. 6 describes system integration, installation and commissioning. Sec. 7 is summary and conclusion. Future prospects are described in Sec. 8.

2 LHC and ATLAS detector

In this section, details of the apparatus LHC and the ATLAS detector are described.

2.1 The Large Hadron Collider

The LHC [19] which is the current world's largest and the most powerful proton-proton collider, was built at CERN in Geneva, Switzerland. There are more than 4000 scientists, engineers, technicians participating the experiment. The main ring of the LHC is built in a tunnel with circumference of 26.7 km and at depth varying between 45 and 170 m. The LHC utilizes the ring used at the Large Electron Positron (LEP) experiment [20]. The initial design center-of-mass energy is 14 TeV and the peak instantaneous luminosity is 1.0×10^{34} cm^{-2}s^{-1}.

The LHC has two general purpose experiments with high luminosity of proton-proton collisions, the ATLAS [21] and the CMS [22], as well as two experiments named the LHCb [23] and the ALICE [24]. The LHCb is targeting for B-physics, and the ALICE experiment is utilizing lead-lead ion collision events aiming mainly for quark gluon plasma.

The advantage of proton-proton collisions compared with electron-positron collisions is the ability to achieve higher center-of-mass energy. The energy loss (ΔE) of accelerated particles by synchrotron radiation in the circular accelerator is described as following.

$$\Delta E = \frac{1}{6\pi\epsilon_0} \frac{e^2 c}{(mc^2)^4} \frac{E^4}{\rho^2}, \tag{37}$$

where ϵ_0 is permittivity of vacuum, e the electron charge, c the speed of light, m and E the mass and energy of the moving particle and ρ the bending radius. By this equation, energy loss mainly depends on the mass of the particle. Thus protons can achieve higher center of mass energy than electron and positron. The limitation of acceleration of the LHC is due to bending power of the magnetic field provided by the dipole super-conducting magnet.

2.1.1 The Injector Chain

At first of the accelerating chain, protons are obtained from a duoplasmatron source. It ionizes hydrogen atoms, accelerates the protons by applying electric field, and separates the protons from electrons with kinetic energy of 90 keV. The protons are sent to a radio frequency quadrupole (QRF), where the protons are accelerated further and formed into bunches by the RF field provided by four vanes. Then the protons are forwarded to the LINAC2 (a LINAC to accelerate the protons) and accelerated up to 50 MeV. The protons which pass the LINAC2 reach the Proton Synchrotron Booster (PSB) which is a circular accelerator consisting of four identical rings. The protons are accelerated to 1.4 GeV at the PSB. Then the protons are injected into the Proton Synchrotron (PS). The PS is the circular accelerator which accelerates the protons to 25 GeV, and forms bunch spacing. Next the protons go into the Super Proton Synchrotron (SPS) and are accelerated to 450 GeV and then sent to the LHC main ring.

2.1.2 The LHC Main Ring

In the LHC main ring, two proton beams are going round with opposite direction each other in separate beam pipes and collide at four interaction points (IP), where the ATLAS, CMS, ALICE, and LHCb detector are placed. The main role of the LHC cavities is to keep the proton bunches to achieve high luminosity at the IPs, and also providing radio frequency power to the beams during acceleration.

The luminosity is described by machine parameters as following.

$$L = \frac{N^2 n_b f_r}{4\pi\sigma_x\sigma_y} F = \frac{N^2 n_b f_r \gamma}{4\pi\epsilon_n\beta_*} F, \tag{38}$$

where N is the number of protons in a bunch, n_b the number of bunches stored in the LHC main ring, f_r the revolution frequency and $\sigma_{x,y}$ characterize the transverse beam profiles in the x, y direction, ϵ_n the normalized transverse beam emittance, β^* the beta function at the IP and the geometrical luminosity reduction factor F is expressed as

$$F = \frac{1}{\sqrt{1 + \frac{\theta_c^2 \sigma_z^2}{4\pi\sigma_x\sigma_y}}}, \tag{39}$$

with θ_c the full crossing angle at the IP and the σ_z the RMS of bunch length.

The LHC beam parameters of 2012, 2016 and original design are shown in Tab. 3.

Parameters	Design	2012 (Run1)	2016 (Run2)
Center-of-mass energy [TeV]	14	8	13
Bunch spacing [ns]	25	50	50/25
Number of bunches	2808	1318	2200
Number of protons per bunch	1.15×10^{11}	1.15×10^{11}	1.25×10^{11}
β^* [cm]	50	80	40
Peak Luminosity [$\mathrm{cm}^{-2}\mathrm{s}^{-1}$]	1.0×10^{34}	7.7×10^{33}	1.38×10^{34}

Table 3: The LHC beam parameters of 2012, 2016 and original design.

2.2 The ATLAS Detector

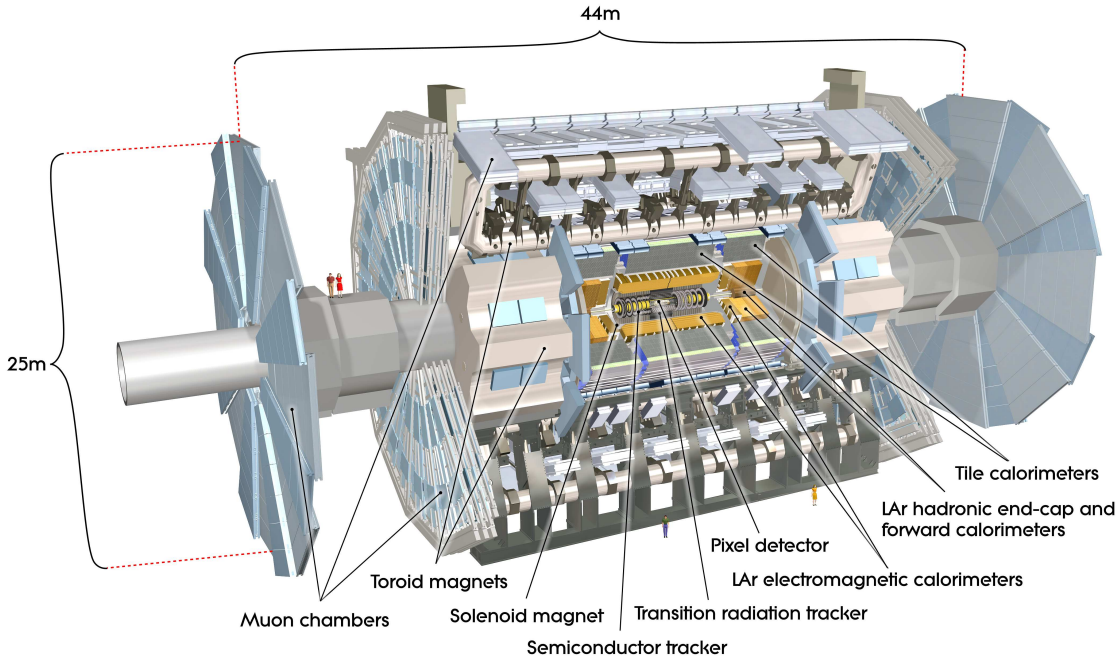

Figure 12: A cutaway view of the ATLAS detector.

The ATLAS detector is the general purpose detector which is placed at one of the IPs of the LHC main ring. A cutaway view of the ATLAS detector is shown in Fig. 12. It is designed as a mirror symmetric manner with respect to the IP. The detector is 25 m in height and 44 m in length, and the total weight is approximately 7000 tons. It aims to search for new particles as well as precise measurement of the SM. It covers almost 4π regions around the IP, and can identify objects such as electrons, photons muons, jets, hadronically decaying τs, and b-jets. In addition, the imbalance of detected transverse energy which is called missing transverse momentum (E_T^{miss}) representing high p_T neutrino in the SM is calculated by summing up all calorimeter transverse energy vectorically. To measure these values precisely, the ATLAS detector consists of an ensemble of sub-detectors.

- Inner detector to measure precise position and momentum of charged particles.
- Electro-magnetic calorimeter which can measure energy with good resolution as well as the position and direction of particles such as electron/gamma.
- Hadronic calorimeter to measure jets and missing transverse energy for wide coverage.
- Muon spectrometer to identify muons and to measure their momentum for a range of 1 GeV up to a few TeV.

The reaction when objects pass in the ATLAS detector is depicted in Fig. 13. Electrons leave tracks in the inner detector and deposit the energy mainly in the electro-magnetic calorimeter. Photons also deposit the energy in the electro-magnetic calorimeter, but do not leave tracks in the inner detector. Neutral hadrons only leave energy in the hadronic calorimeter, while charged hadrons leave tracks in the inner detector and electro-magnetic calorimeter as well. Muons pass through the detectors and leave tracks in the muon

spectrometer. Neutrinos do not leave any signature in the detectors. The details of each sub-detector are described in the following subsections.

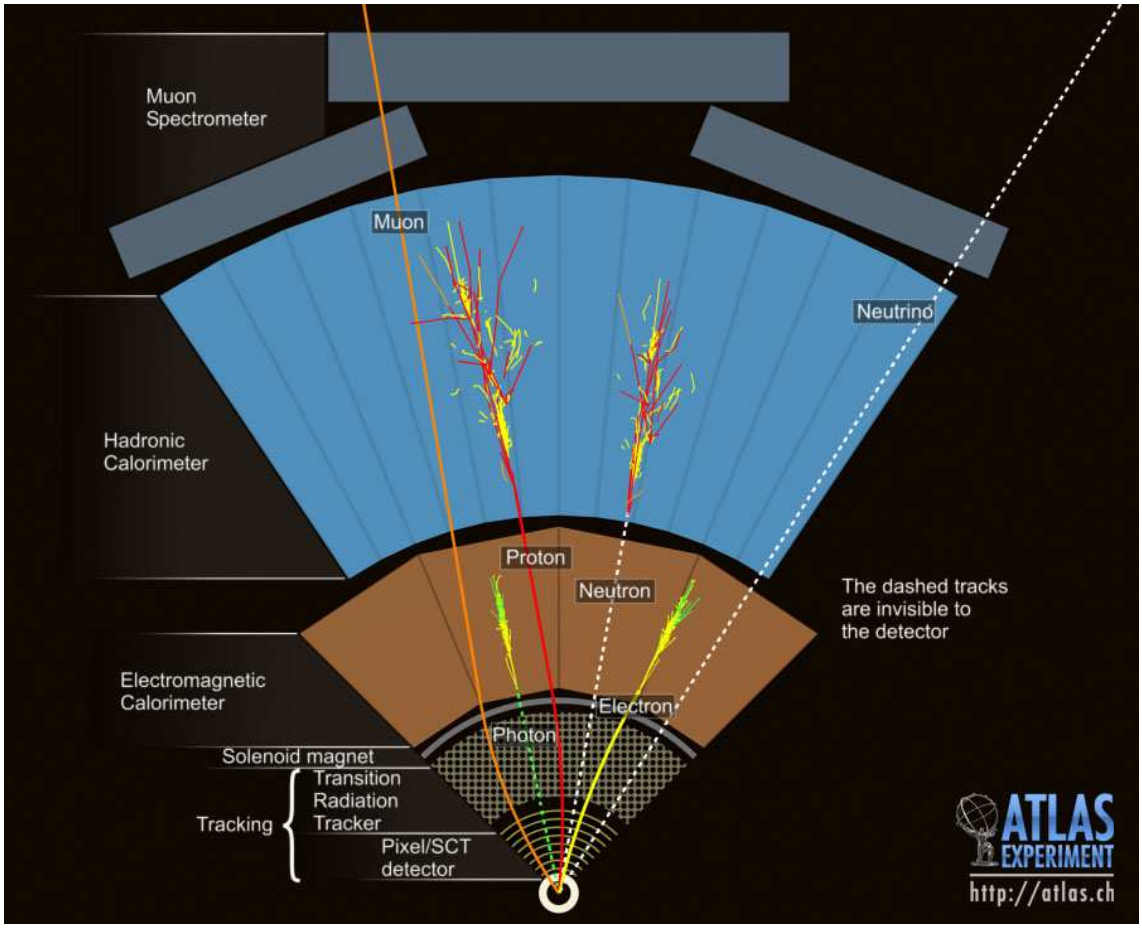

Figure 13: A schematic view of particle detection in the ATLAS detector ensemble.

2.2.1 Magnet System

The ATLAS detector has a unique hybrid system of the four large superconducting magnets, which consists of one solenoid and three toroids. This magnetic system is 22 m in diameter, 26 m in length and its stored energy is 1.6 GJ. A cutaway view of the magnetic system is shown in Fig. 14. Eight barrels and two sets of eight endcaps are depicted as red windings. The solenoid magnet is a red cylinder inside the hadronic calorimeter. They provide magnetic field, bend trajectory of charged particles and make it possible to measure their momentum.

Solenoid Magnet

The solenoid magnet [25] is placed between the inner detector and the calorimeter. It has 2.46 m in inner diameter, 2.56 m in outer diameter, and 5.8 m in length. The solenoid is a layer coil which is made of Al-stabilized NbTi conductor. It provides homogeneous axial magnetic field of 2 T at the nominal

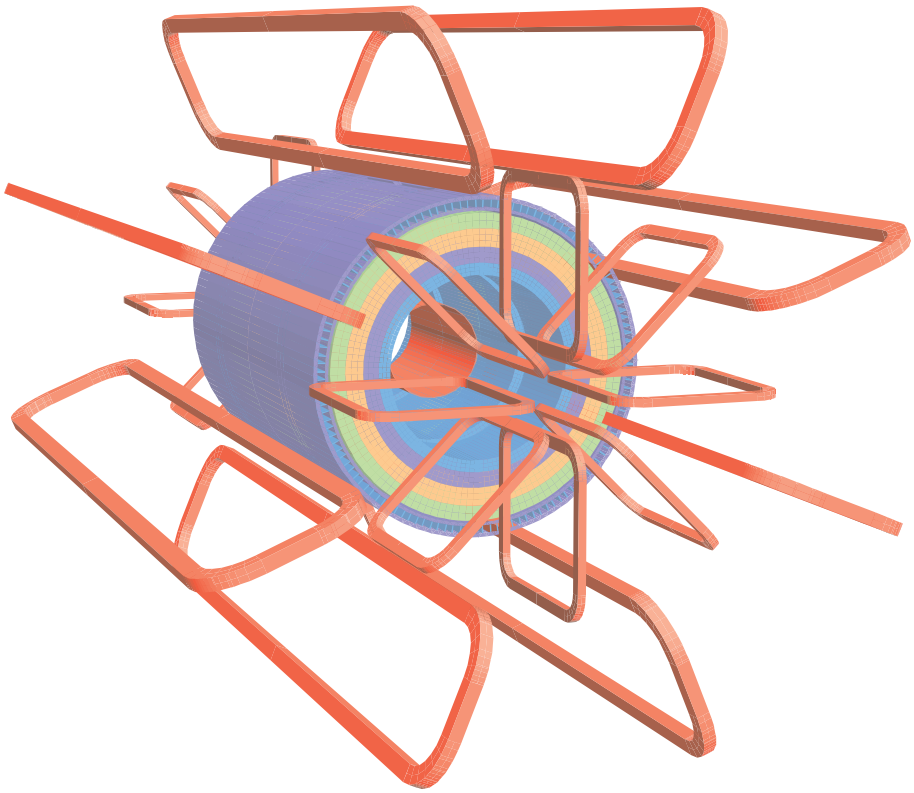

Figure 14: A cutaway view of the ATLAS superconducting magnetic system.

operation current of 7.730 kA and temperature of 4.5 K. The used materials are as low as possible to keep the performance of the calorimeter higher. The assembly of the solenoid is 0.66 X_0 radiation length. The flux of the magnetic field is returned by the steel of hadronic calorimeter.

Toroid Magnets

The air-core toroid magnet [26] consists of one barrel and two endcaps. The barrel toroid magnet is cylindrical volume surrounding calorimeters and in the muon spectrometer. The size of the barrel toroid magnet is 9.4 m in inner diameter, 20.1 m in outer diameter, and 25.3 m in length. It provides magnetic field of 0.5 T at the nominal current, and can provide up to 3.9 T at maximum. Two endcap toroid magnets form magnetic field for the endcap of the muon spectrometer. They have 1.65 m in inner diameter, 10.7 m in outer diameter, and 5.0 m in length. They provide magnetic field of 1 T at the nominal current, and can provide up to 4.1 T at maximum. The conductor and coil-winding technology is based on winding a pure Al-stabilized Nb/Ti/Cu conductor.

2.2.2 Inner Detector

The inner detector [27–29] is placed at the most inner part of the ATLAS detector. A cutaway view is shown in Fig. 15, and a schematic view of the detailed configuration is shown in Fig. 16. It is used to measure precise position and momentum of charged particles, and to reconstruct vertices. It consists

of Insertable B-Layer (IBL), Pixel detector, Semi-Conductor Tracker (SCT), and Transition Radiation Tracker (TRT). The solenoid magnet provides magnetic field of 2 T to measure the momentum of charged particles. The resolution of the transverse momentum for charged particles is described as

$$\frac{\sigma_{p_T}}{p_T} = 0.05 \times p_T(\text{GeV}) \oplus 1\%. \tag{40}$$

The data from the IBL, the Pixel detector and the SCT are used as an input of the FTK. The details of each sub-detector is described in the following subsections.

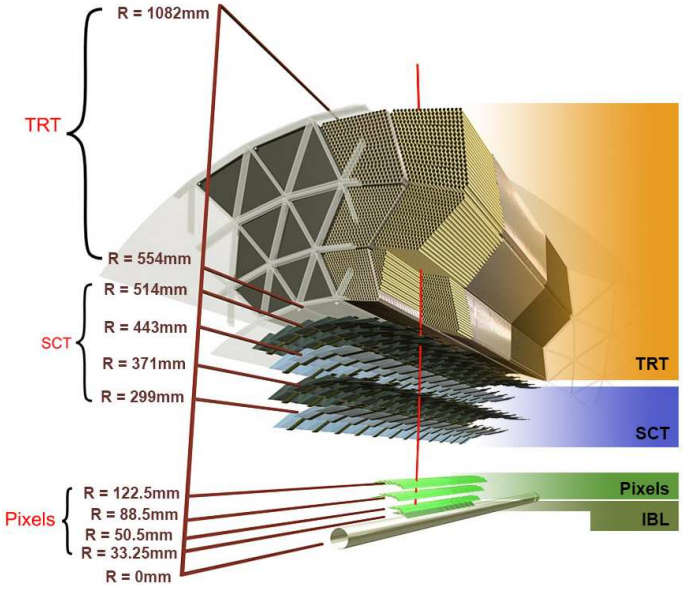

Figure 15: A cutaway view of the ATLAS inner detector [30].

Insertable B-Layer

The Insertable B-Layer (IBL) [31, 32] was installed in the innermost layer of the inner detector in 2015 as one of the ATLAS upgrade program. It is the fourth layer of the Pixel detector covering $|\eta| < 2.9$, and can be repairable to keep tracking performance high when it is damaged by radiation under high instantaneous luminosity. It improves quality of impact parameter which is important for b-tagging and vertex reconstruction.

The IBL consists of 14 staves which support modules. It is very light with less than 0.6% X_0. Each IBL stave is made of carbon fiber material of 64 cm in length tilting 14° around beam pipe. In total 12M pixels are used. It is inserted with beryllium beam pipe of 25 mm radius. To cover entire ϕ region, the modules have 0.18° overlaps which cover inefficiency at the boundary and Lorentz angle in the magnetic field of 2 T provided by the solenoid magnet. The average distance between the staves and the center of the beam pipe is 33.25 mm. Figure 17 shows the IBL structure in the $r - \phi$ plane. The modules should be kept in low temperature (< -20°C) to minimize radiation damage. The chip design utilizes CMOS technology of 130 nm. 26880 pixels are placed, 80 in columns and 336 in rows. The size of the pixel is 50×250 μm^2.

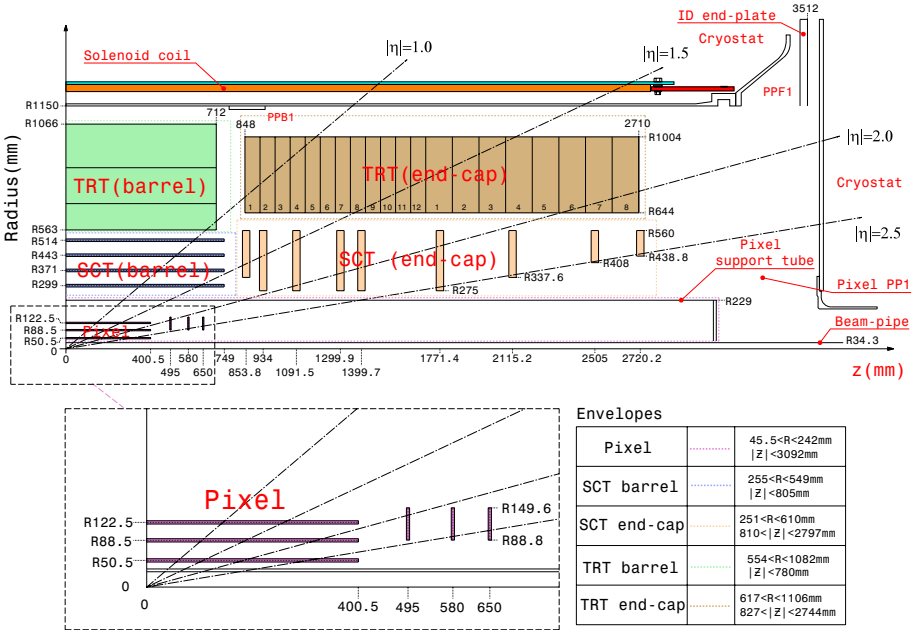

Figure 16: A detailed configuration of the ATLAS inner detectors.

Two types of sensors [33] are utilized, which are planar and 3D. The planar modules cover 75% in η, and the 3D modules cover remaining 25%.

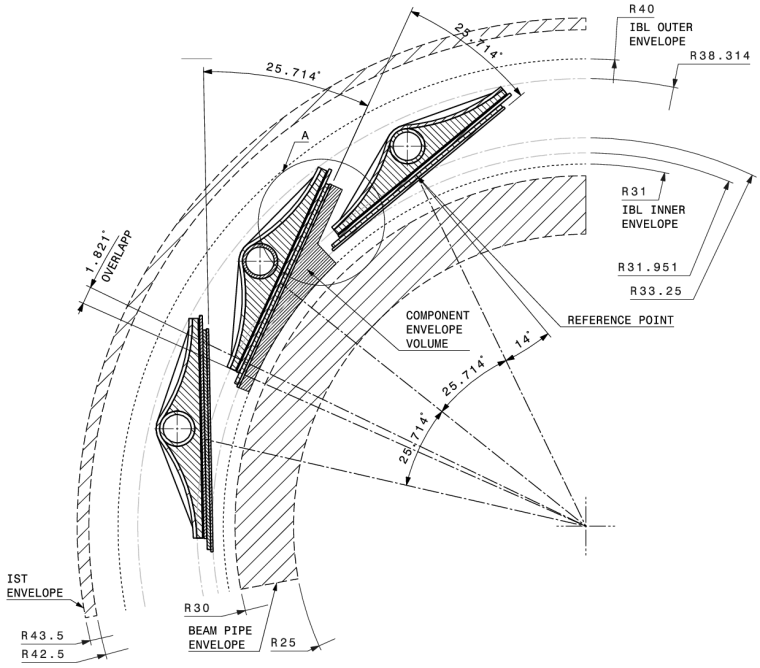

Figure 17: A schematic view of the IBL structure in $r-\phi$ plane.

Pixel Detector

The Pixel detector [34] is placed in the next outer part of the IBL. The most inner part is 51 mm from the LHC beam pipe. It consists of three cylindrical layers in the barrel, and four disks standing vertically to beam line in the endcap, covering $|\eta| < 2.5$. Figure 18 shows a schematic view and a picture of pixel modules. There are 1744 modules in total. The modules are made of semi-conducting silicon sensors. The size of the modules is 50×400 μm^2. Signals are readout by 16 front-end chips with 2880 channels. Some pixels have 600 μm in length and some have two pairs ganged into single readout to cover clearance region. Since the pixels are small, the measurement of the space point can be performed with high resolution. The barrel modules are aligned to measure z and ϕ coordinate, while endcap modules are for z and r coordinates.

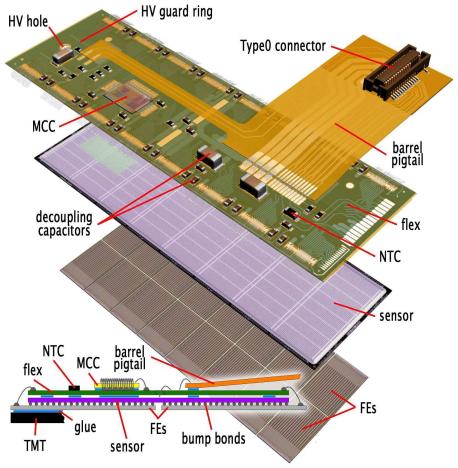

Figure 18: A schematic view of the pixel modules.

Semi-Conductor Tracker

The SCT [35, 36] is placed at the next outer part of the Pixel detector. The SCT covers from 300 mm to 560 mm in radius from the LHC beam axis. It consists of four cylindrical layers in the barrel ($|\eta| < 1.1$), and nine disks in the endcap ($|\eta| < 2.5$). Each layer has axial and stereo layer attached with the 40 mrad stereo angle each other which enables to measure space point. There are 4088 modules in total, where 2112 modules are in the barrel, and 988 modules are in each endcap. A schematic view and a picture of a SCT module are shown in Fig. 19. The resolution is 17 μm in ϕ direction, 580 μm for z direction in the barrel, and 17 μm in ϕ, 580 μm for z direction in the endcap. It consists of double-sided semi-conducting strip sensor with 80 μm pitch. All the sensors have 285 μm thickness made of p-type implanted high-resistive n-type bulk silicon. Each module has 128 mm in length, and 768 active channels in total. They are readout via 6 front-end chips.

The SCT readout system is designed to work with 0.2-0.5% hit occupancy for 6.3 million sample strips under the originally planned LHC luminosity of 1.0×10^{34} cm^{-2}s^{-1} where the number of pile-up interactions is 23. Information of strips is readout by radiation hardness frond-end ABCD chips [37]. Each channel of ABCD chips has preamplifier and shaper. Common discriminator threshold is applied for all 128 channels, and it is typically a charge of 1 fC. 1 fC is determined to ensure low noise occupancy

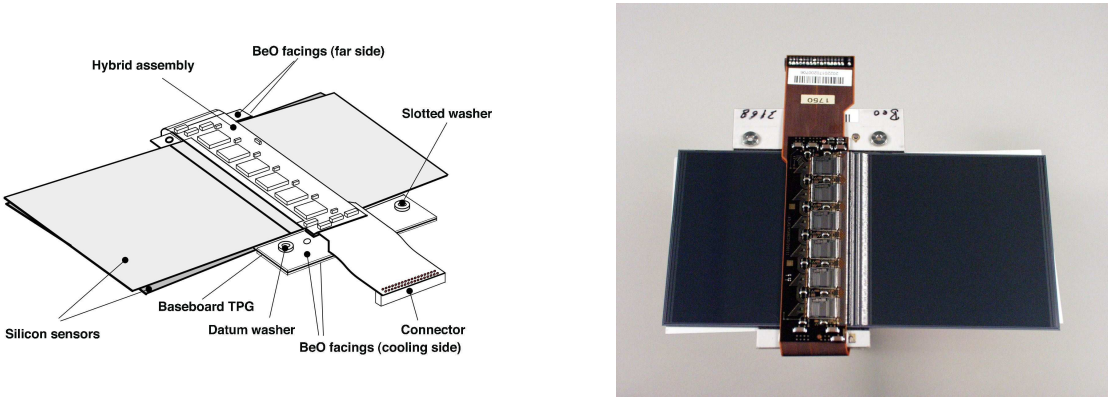

Figure 19: A schematic view and a picture of a SCT module.

($< 5 \times 10^{-4}$) and high hit efficiency ($> 99\%$). To cover variation of the threshold among channels, it is possible to set off-set channel by channel.

Off-detector readout system [38] consists of 90 9U readout-driver boards (ROD) and 90 back-of-crate boards (BOC) implemented on 9U VME. Each ROD processes 48 modules at the maximum, and the the BOC is an optical interface between the ROD and the modules. Figure 20 shows the BOC sends reformatted data to the ATLAS data acquisition (DAQ) chain via optical fiber which is called SLINK. There are two data streams returned from each module which correspond to two sides of the module. The SCT ROD has three modes of output data format which are expanded mode, condensed mode, and super-condensed mode. The difference of the data mode is how the data is compressed to cope with the increase of luminosity and then hit occupancy. The expanded mode sends data of less compressed format, while the super-condensed mode sends data of the most compressed format. The expanded mode was used until middle of 2016, and switched to the super-condensed mode. Redundancy is implemented on both transmitting (TX) side and receiving (RX) side in case of problem on readout links. One side of the module can be readout by the other link in case there is problem for one of them. Though readout by both sides reduces readout bandwidth, it is within the limit of the design.

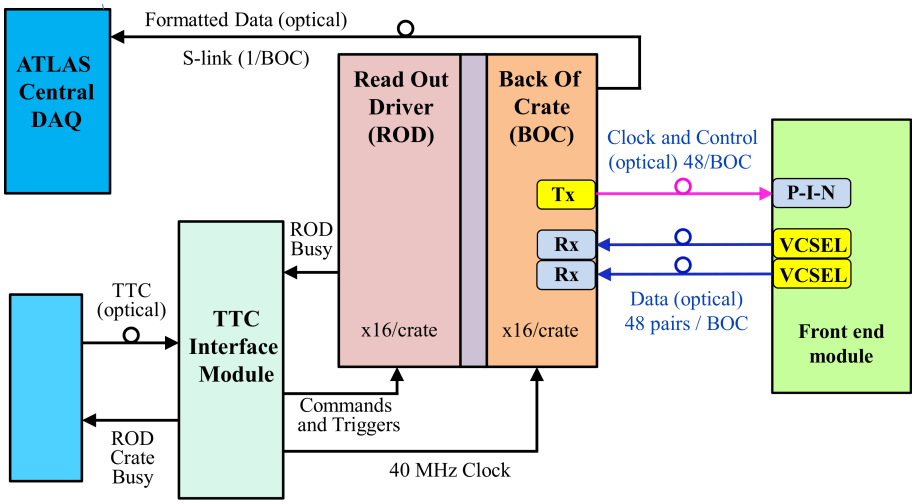

Figure 20: A schematic diagram of the SCT data acquisition.

The SCT keeps powered regardless of the LHC beam status. When the LHC is not stable beam state, the SCT modules are applied reduced high voltage of 50 V to deplete a part of silicon sensor. At the normal data-taking, bias voltage of 150 V is applied to maximize hit efficiency for tracking. The switching of the state from standby of 50 V to nominal 150 V is called "warm start". There are mainly three sources of data-taking inefficiency which are time to switch on after the LHC becomes stable beam, error on a chip and it is not available, and "busy signal" from the SCT DAQ. The inefficiency by the busy signal is dominant but it is less than 1% inefficiency of data-taking, and other sources have little impact.

The SCT operation is basically trouble-free. The number of disabled modules is fluctuated a little through a year, but rarely increase more than 5-10 modules. The online monitoring system is prepared which enables immediate feedback and investigation of the SCT conditions.

The SCT is designed to have low detector occupancy. The mean strip occupancy is expected to be less than 1% at the original LHC design luminosity of 1.0×10^{34} cm^{-2}s^{-1} and beam crossing rate of 40 MHz. Occupancy is defined as the ratio of the number of strips which exceed threshold to the number of total strips. Figure 21 shows the occupancy for four barrels and one endcap. Strips which are defined as noisy are removed from consideration. The plots below 40 interactions per bunch crossing corresponds to $\sqrt{s} = 7$ TeV and above 40 corresponds to $\sqrt{s} = 8$ TeV. Occupancy is expected to be 1.03 times more in $\sqrt{s} = 8$ TeV than in $\sqrt{s} = 7$ TeV. Taking this into account, good linearity can be seen up to 70 interactions per bunch crossing.

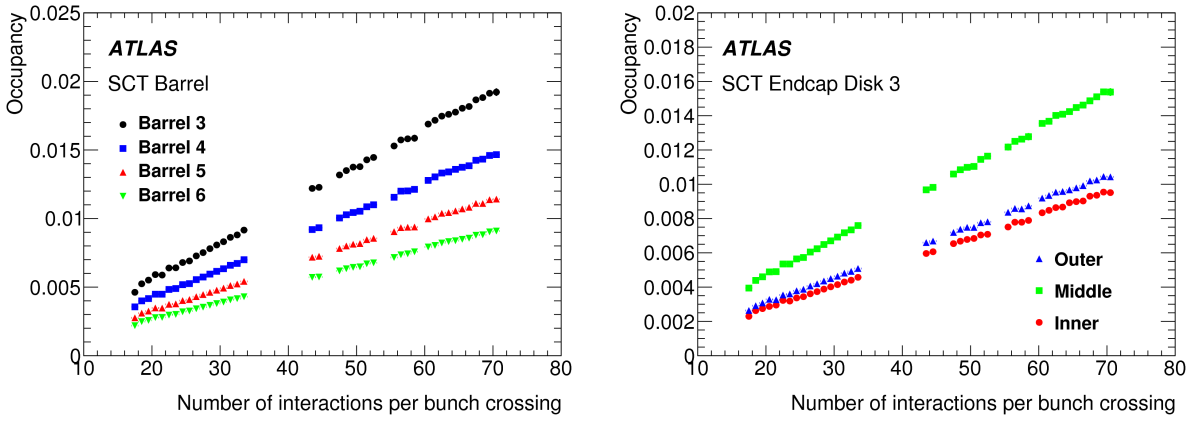

Figure 21: Mean occupancy of each barrel (left) and inner, middle and outer modules of 3 endcap disks (right) as a function of the number of interactions per bunch crossings.

Hit efficiency is probability that a charged particle deposits hits (clusters) when it traverses sensitive region of the detector. Here disabled sensors, chips, modules are not taken into account. The intrinsic efficiency is measured by extrapolating tracks through entire detector and counting hits on the track and strips which have hits and should have hits but not which is called "hole". The efficiency is defined as the ratio of the number of found clusters to the number of holes. It is expressed as

$$\epsilon = \frac{N_{clusters}}{N_{clusters} + N_{holes}}, \tag{41}$$

where $N_{clusters}$ is the number of clusters found and N_{holes} is the number of holes. The hit efficiency for each endcap and disk is shown in Fig. 22. The overall efficiency is $99.74 \pm 0.04\%$. The fluctuation of the

efficiency among disks and endcaps depends on the number of disabled strips which is also shown in the figure. The efficiency and the number of disabled strips have clear correlation.

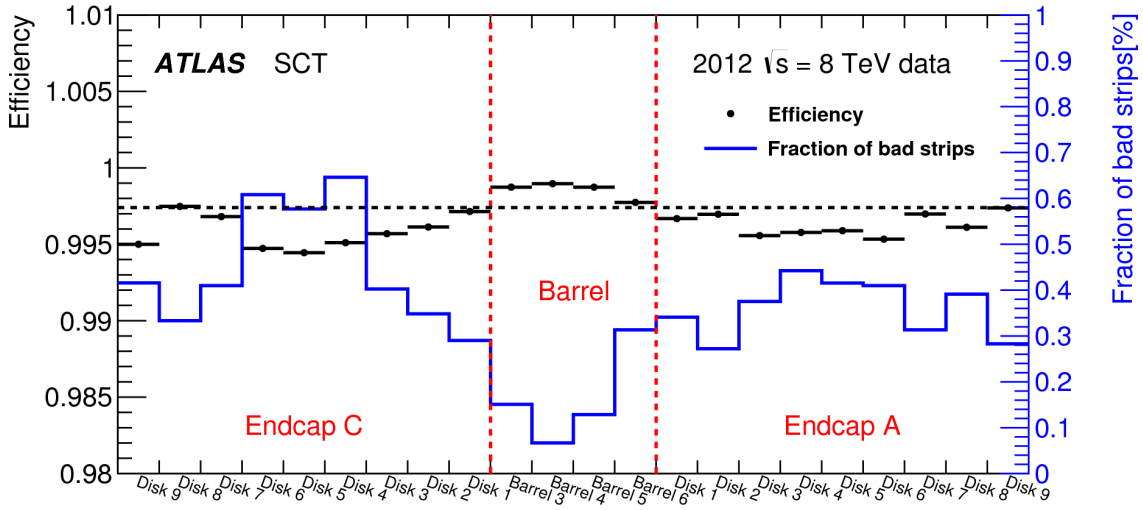

Figure 22: Mean intrinsic efficiency for each layer of the SCT.

Although the SCT is not designed to measure energy loss, dE/dx, and to perform particle identification, it has discrimination power to some extent from information of time bins above threshold and the number of strips in a cluster. A heavily ionizing particle deposits much charge and makes large pulse height which is above threshold in time bins. That particle tends to make a large size of cluster. It can be confirmed by measuring energy loss, dE/dx, defined as following equation.

$$dE/dx = \frac{1}{N}\sum_{i=1} w_i cos\alpha_i, \tag{42}$$

where i runs for all the strips in all the clusters associated to the track, N is the number of clusters associated to the track, w_i is the number of time bins above threshold for the strip. α_i is the angle between the track and silicon sensor, and $cos\alpha_i$ corrects path length of the track within silicon.

Figure 23 left shows energy loss measured in the SCT as a function of momentum multiplied by charge in barrel region for events collected by minimum bias trigger. To increase protons in the sample, particles originated from secondary interactions with detector materials are enhanced by requiring large impact parameter with respect to primary interaction vertex. Protons are seen in the figure which correspond to the band of highly ionizing particles in the low-momentum positive-charge region.

Particle identification is performed by likelihood method. The probability density functions for different particles are determined, in some momentum ranges, by fitting dE/dx with Gaussian for particles identified as protons, kaons, pions by dE/dx of the Pixel detector. Figure 23 right shows proton tagging efficiency as a function of kaon mistag rate. Efficiency is defined as the fraction of the number of protons identified by dE/dx of the SCT to the number of protons identified by dE/dx of the Pixel detector, while kaon mistag rate is defined as the fraction of the number of kaons identified by dE/dx of the SCT to the number of protons identified by dE/dx of the Pixel detector. Thus the SCT has some discrimination power and ability for particle identification.

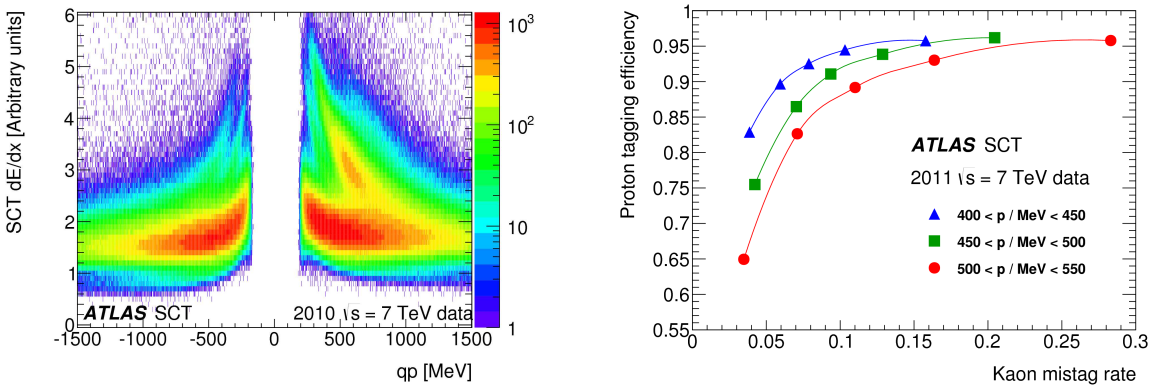

Figure 23: (left) Energy loss measured in the SCT as a function of momentum multiplied by charge. Protons are enhanced in the sample by collecting particles originated from secondary interaction with detector materials. (right) Proton tagging efficiency as a function of kaon mistag rate.

Transition Radiation Tracker

The TRT [39–42] is surrounding the SCT, and consisting of tubes aligned in parallel to the beam axis, covering $|\eta| < 2.0$ and inner radius is 563 mm from the LHC beam pipe. Straw tubes are filled with $Xe/CO_2/O_2$ (70/27/3%) mixture gas and used as cathode, while a gold plated tungsten wires are used as anode. When charged particles pass the TRT, they ionize the gas. Generated electrons drift to the wires, amplified and readout. The cathode is tube made of carbon, aluminum, and kapton. The diameter is 4 mm, and the gain of 2.5×10^4 can be obtained by the typical bias voltage of 1530 V between anode and cathode. Maximum drift time is 48 ns, and the resolution of the drift circle measurement is 130 μm. The left-right ambiguity can be resolved by extrapolating tracks from the Pixel detector and the SCT. The length of the tube is 114 cm and 37 cm in barrel and endcap, respectively. The average number of hits generated by charged particles is 30. The TRT utilizes transition radiation which occurs when charged particles pass two dielectric materials. The intensity of transition radiation is proportional to Lorentz factor γ of the particle. Electrons deposit high energy from transition radiation, while energy deposit by minimal ionizing particles is 250 MeV.

2.2.3 Calorimeter

The ATLAS calorimeter system [43] is equipped at the outer part of the solenoid magnet, covering $|\eta| < 4.9$. A cutaway view of the ATLAS Calorimeter is shown in Fig. 24. It utilizes changing granularity in $\eta \times \phi$ to measure the energy of particles. There are two types of calorimeters, the electro-magnetic calorimeter and the hadronic calorimeter. The energy resolutions for each calorimeter are given as

$$\frac{\sigma_E}{E} = \frac{10}{\sqrt{E(GeV)}} \oplus 0.7\%, \ (|\eta| < 3.2 \text{ for EM measurements}) \tag{43}$$

$$\frac{\sigma_E}{E} = \frac{50}{\sqrt{E(GeV)}} \oplus 3\%, \ (|\eta| < 3.2 \text{ for hadronic measurements}) \tag{44}$$

$$\frac{\sigma_E}{E} = \frac{100}{\sqrt{E(GeV)}} \oplus 10\%. \ (3.2 < |\eta| < 4.9 \text{ for hadronic measurements}) \tag{45}$$

The details of each calorimeter are described in the following subsections.

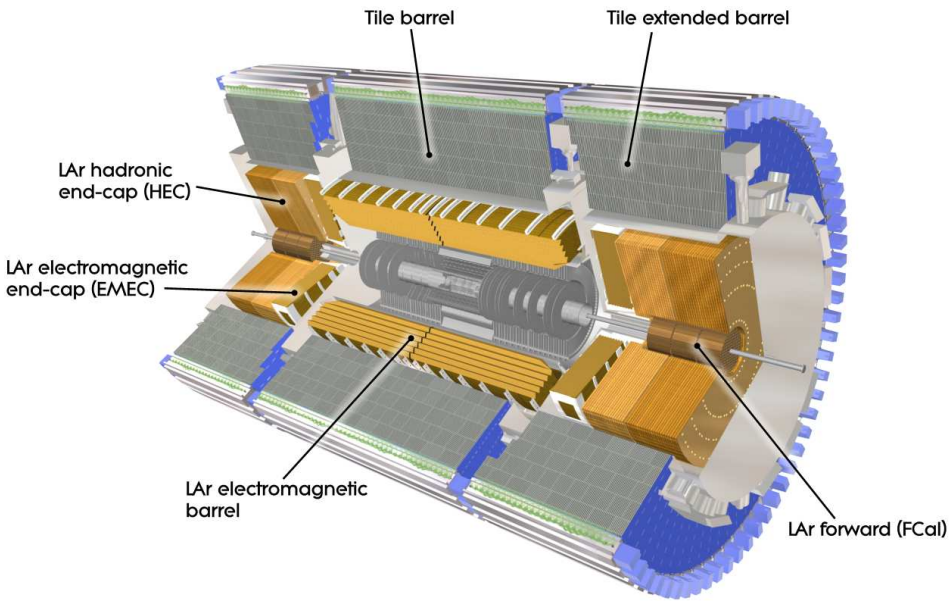

Figure 24: A cutaway view of the ATLAS Calorimeters.

Electro-Magnetic Calorimeter

The electro-magnetic (EM) calorimeter [44] is a sampling calorimeter which consists of lead plates of 1.5 mm in thickness with accordion shape, and there are 2.1 mm gap filled with liquid argon used as active layers. Due to its accordion shape, it can cover entire phi region without azimuthal crack and enables to do fast signal extraction at the rear and the front of the electrodes. The energy of charged particles is measured by detecting the ionization of the charged particles in the liquid argon. The ionization charge is collected by copper clad kapton which is also accordion shape. The total drift time is 450 ns by applying 2000 V to the gap. The thickness of the lead plates is optimized for the performance in energy resolution. The total thickness is exceeding 22 X_0 radiation length in the barrel, and 24 X_0 radiation length in the endcap, which provides good performance of EM shower. The EM calorimeter consists of one cylindrical barrel calorimeter ($|\eta| < 1.475$) with small gap at $|z| = 0$, and two wheel endcap calorimeters ($1.375 < |\eta| < 3.2$). Each calorimeter has three layers with high angular resolution and full shower shape detection. The first layer has $\eta \times \phi = 0.0031 \times 0.0098$, 4.3 X_0 radiation length, which is so called strip layer. The second layer has $\eta \times \phi = 0.025 \times 0.0245$, 16 X_0 radiation length. Most energy by the EM shower are measured in this layer. This layer is also used as a seed to reconstruct electrons and photons. The third layer measures the tail and has $\eta \times \phi = 0.050 \times 0.025$. In addition there is a thin layer in front of the strip layer called pre-sampler, which has 11 mm in thickness and $\eta \times \phi = 0.025 \times 0.1$. It covers $|\eta| < 1.8$ and is used for correction of the energy loss caused by the interaction with materials of the upstream of the calorimeter. A sketch of a barrel module and a photograph of a partly stacked barrel modules are depicted in Fig. 25. There are large materials in $1.37 < |\eta| < 1.52$, where the energy and position resolution are worse.

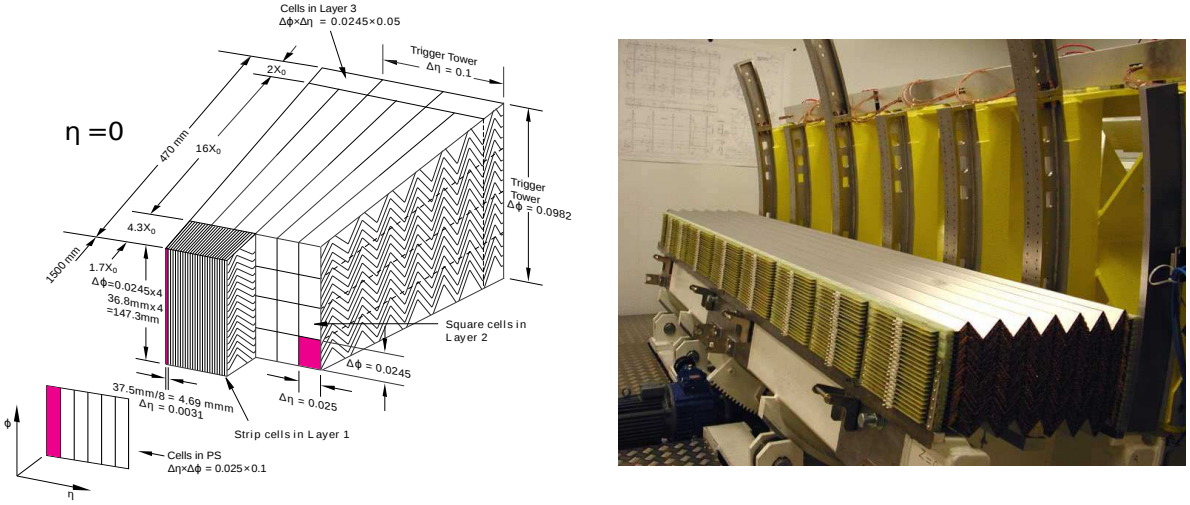

Figure 25: (left) A sketch of a barrel module of the Electro-Magnetic calorimeter. (right) A photo of a partly stacked modules of the barrel Electro-Magnetic Calorimeter.

Hadronic Calorimeter

The hadronic calorimeter is surrounding the EM calorimeter, consists of one barrel scintillator tile calorimeter ($|\eta| < 1.7$) [45] and two Hadronic Endcap Calorimeters (HEC, $1.5 < |\eta| < 3.2$) [46], and the Forward Calorimeter (FCal, $3.1 < |\eta| < 4.9$) [47]. The scintillator tile calorimeter is sampling calorimeter which utilizes steel plates as absorber and plastic scintillator as active material. Its granularity is $\eta \times \phi = 0.1 \times 0.1$. The central barrel covers $|\eta| < 1.0$ and two extended barrels cover $0.8 < |\eta| < 1.7$. Scintillation photons are collected, guided by the wavelength shifter fibers, and readout by photo-multiplier tubes. A schematic view of mechanical assembly and optical readout of the tile calorimeter is shown in Fig. 26.

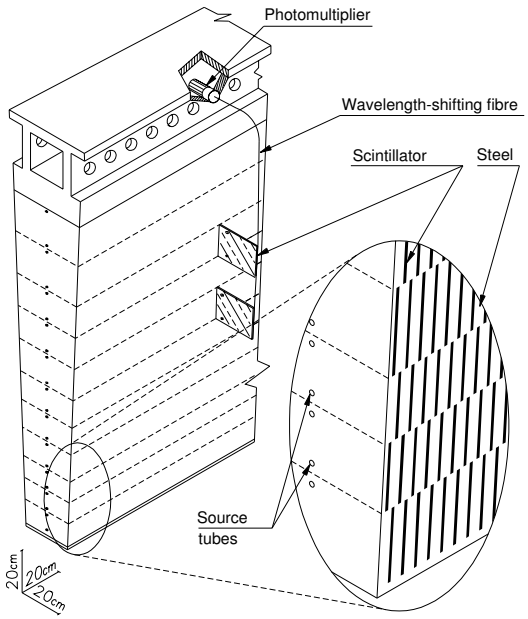

Figure 26: A schematic view of the mechanical assembly and the optical readout of the tile calorimeter.

The endcap calorimeter consists of two wheels and cryostat which is common with the Electro-Magnetic Endcap Calorimeter (EMEC). It utilizes copper as absorber and liquid argon as active media. Each calorimeter has 10 interaction length which is enough to decrease punch through hadrons. A schematic view of a HEC module and a HEC readout structure are shown in Fig. 27.

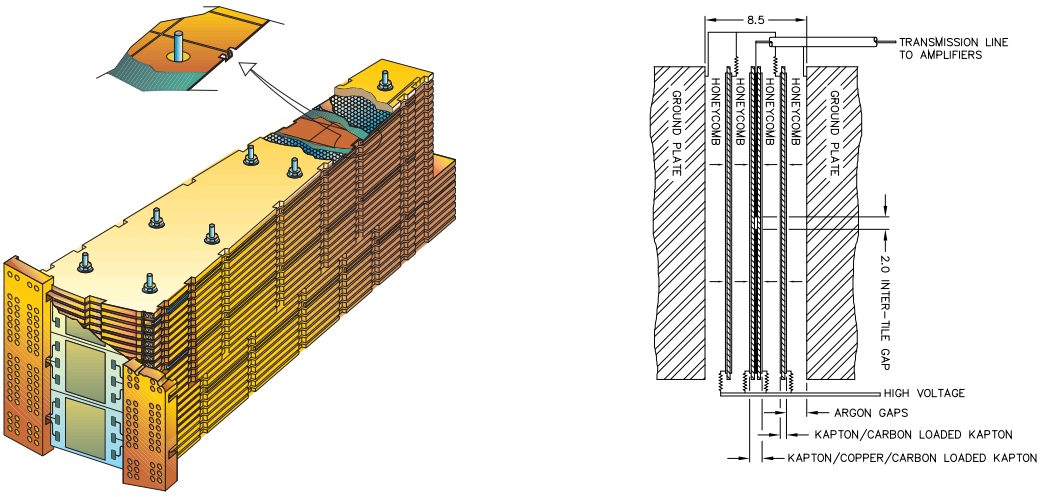

Figure 27: (left) A schematic view of a HEC module with the readout structure and the active-pad. (right) A schematic view of a HEC readout structure. All dimensions are shown in mm.

One of the forward calorimeters measure EM interaction (FCal1), and two measure hadronic interaction (FCal2 and FCal3). As absorber, copper is used for the EM and tungsten is used for the hadronic, both utilize liquid argon as active media. To achieve high radiation durability, there are very small liquid argon gaps. A schematic diagram of the three FCal modules is shown in Fig. 28.

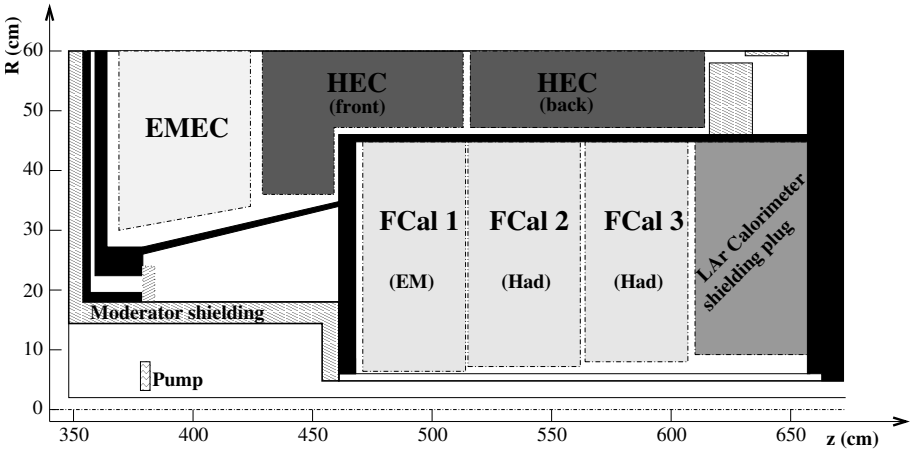

Figure 28: A schematic diagram of the three FCal modules located in the endcap.

2.2.4 Muon Spectrometer

The Muon Spectrometer (MS) [48] is the outer most detector of the ATLAS detector. A cutaway view of the MS is shown in Fig. 29, a schematic $R - Z$ view is shown in Fig. 30, and a schematic $R - \phi$ view is shown in Fig. 31. It covers $|\eta| < 2.7$ and consists of two precision chambers and two trigger chambers. Two precision chambers are called Monitored Drift Tubes (MDT) and Cathode Strip Chamber (CSC), and two trigger chambers are called Thin Gan Chamber (TGC) and Resistive Plate Chamber (RPC). The momentum of muons is measured by the detecting deflection of muon trajectory in the magnetic field provided by air-core toroid magnets. The MS aims to identify muons efficiently and to measure their momentum up to 1 TeV with a resolution better than 10%. The detail of the sub-detectors are described in the following subsections.

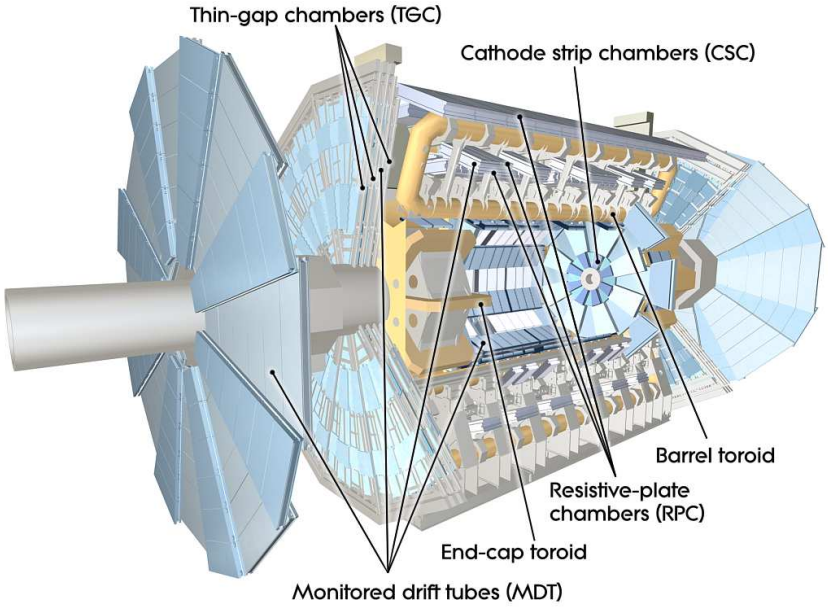

Figure 29: A cutaway view of the ATLAS Muon Spectrometers.

Monitored Drift Tubes

The MDT is a tracking chamber with high resolution. It covers $|\eta| < 2.0$, and z coordinate is measured in the barrel and radial coordinate is measured in both barrel and endcap. The MDT has aluminum drift tubes of 30 mm diameter which are filled with Ar/CO_2 gas with proportion of 93/7% under 3 bar pressure. As an anode, tungsten-rhenium (W-Re) wires of 50 μm diameter is used with 3080 V applied, by which the maximum drift time from the wall to the wires is 700 ns. There are 1150 MDTs, 592 cover barrel and 558 cover endcap. The average resolution of z direction is 80 μm per tubes, 35 μm per chamber. Cross-section of a MDT tube and mechanical structure of a MDT chamber are shown in Fig. 32.

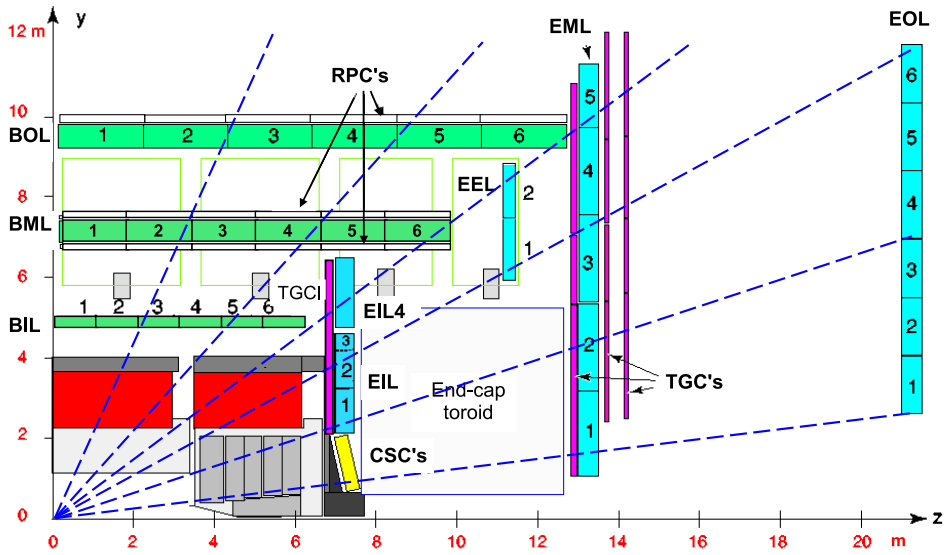

Figure 30: A schematic $R - Z$ view of the ATLAS Muon Spectrometer.

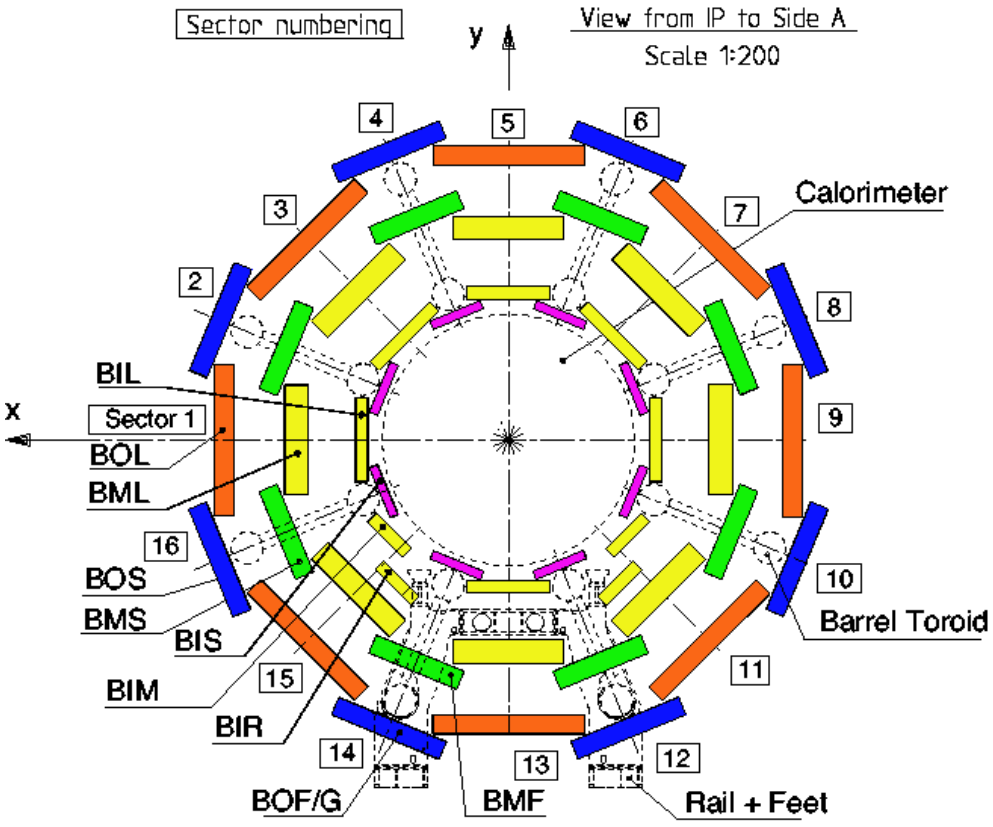

Figure 31: A schematic $R - \phi$ view of the ATLAS Muon Spectrometer along with other sub-detectors.

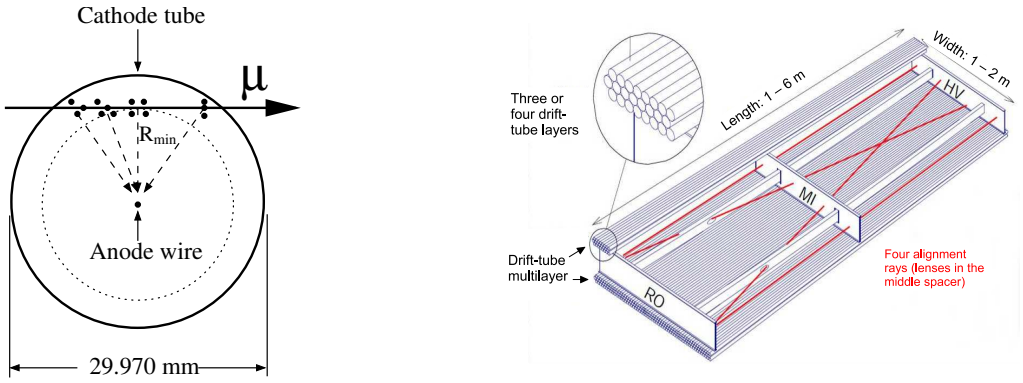

Figure 32: (left) Cross-section of a MDT tube. (right) Mechanical structure of a MDT chamber.

Cathode Strip Chamber

The CSC is also used for precision tracking, covering $2.0 < |\eta| < 2.7$. It consists of two disks, each has eight chambers segmented in ϕ. Each chamber has four wire planes, which are similar configuration as the MDT system. The CSC is multi-wire proportional chamber filled with Ar/CO_2 gas with proportion of 80/20%. The anode wires are gold-plated tungsten with 3% rhenium of 30 μm diameter. The cathode plane is consists of strips in orthogonal directions, which enables to measure both coordinates. The resolution of the chamber is 40 μm in bending plane and about 5 mm in transverse plane. The drift time of electrons is less than 40 ns.

Thin Gap Chamber

The TGC is a trigger chamber covering $|\eta| > 1.05$. It is multi-wire proportional chamber which operates in saturated (limited proportional) mode. The advantages of saturated mode are high single hit efficiency and small dependence of incident angle of muons. It is very strong quenching gas mixture filled with CO_2/n-C_5H_{12} with proportion of 55/45%, which is appropriate to operate in saturated mode. The structure of the TGC is shown in Fig. 33. The characteristics of this detector is small distance between anode and cathode, and strong electric field, which reduces drift component of the ionization clusters and achieves good timing resolution.

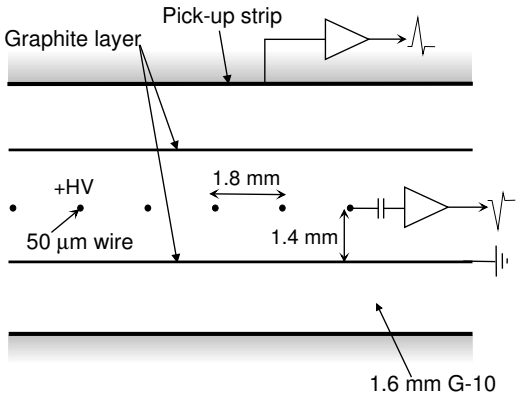

Figure 33: A schematic view of the TGC structure.

Resistive Plate Chamber

The RPC is also used for trigger decision, covering $|\eta| < 1.05$. It is gas parallel electrode plate detector. It consists of two resistive plates made of phenolic-melaminic plastic laminate, which are placed in parallel each other by insulating spaces in 2 mm distance. The used gas is $C_2H_2F_4$/Iso-C_4H_{10}/SF_6 with proportion of 94.7/5.0/0.3%. The position resolution for ϕ is about 10 mm.

2.2.5 Minimum Bias Trigger Scintillator

The Minimum Bias Trigger Scintillator (MBTS) [49] covers $2.1 < |\eta| < 3.8$, which consists of 16 counters per side, 8 of which form an inner ring and the other 8 an outer ring. A schematic view and a photo of the MBTS are shown in Fig. 34. They are placed at $|z| = 3.6$ m between the inner detector and the EM calorimeter. The signals are readout through the electronics of the hadronic calorimeter. The MBTS is used for trigger if it is fired and the voltage exceeds a pre-defined threshold, which is called minimum bias trigger. The minimum bias trigger is designed to select all types of inelastic interactions with minimal bias across full acceptance of the ATLAS detector.

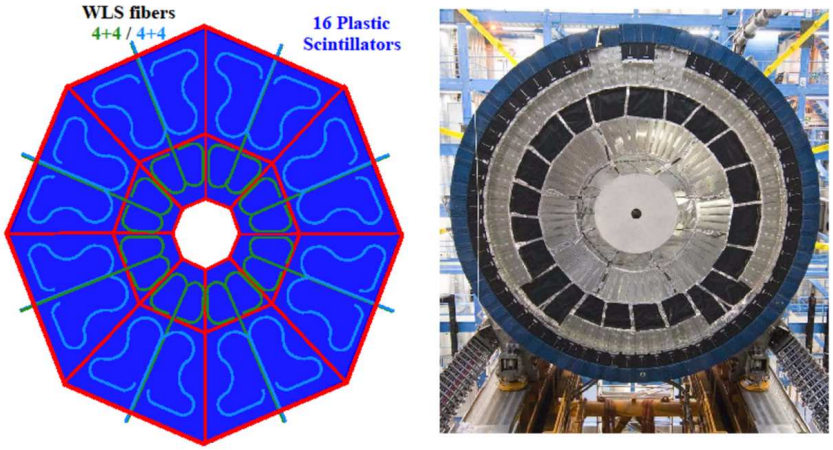

Figure 34: A schematic view and a photo of the Minimum Bias Trigger Scintillator [50].

2.3 ATLAS Trigger System

As briefly described in Sec. 1, the ATLAS experiment implements trigger system [51, 52] that performs event selection at the data taking stage to keep interesting events as well as to reduce event rate. At LHC Run1, the ATLAS experiment implemented three stages of trigger system, which are the Level-1 trigger, the Level-2 trigger and the Event Filter [53]. It consists of two stages which are called the Level-1 trigger (L1) and the High Level Trigger (HLT) at LHC Run2 and will be at LHC Run3. The current overall trigger system including the FTK is shown in Fig. 35.

The hardware-based L1 trigger utilizes information of the calorimeters (L1Calo) and the Muon Spectrometers (L1Muon) with reduced resolution for the fast decision. It searches for the high energy deposit in the calorimeters, high-p_T muon tracks in the MS, and large E_T^{miss} and total transverse energy. The event rate is reduced to 100 kHz after the L1 decision. The software-based HLT makes precise trigger decision

based on Region of Interest (RoI) defined by the detector regions around the objects found by the L1 trigger. The RoI provides the information on coordinates, energy, and type of signatures to reduce the amount of data. The HLT algorithm is performed with the information on the calorimeters, the Muon Spectrometers, as well as the inner detectors with full granularity and precision. The HLT algorithm reduces event rate to 1 kHz. The data which pass the HLT are recorded and used for later analysis.

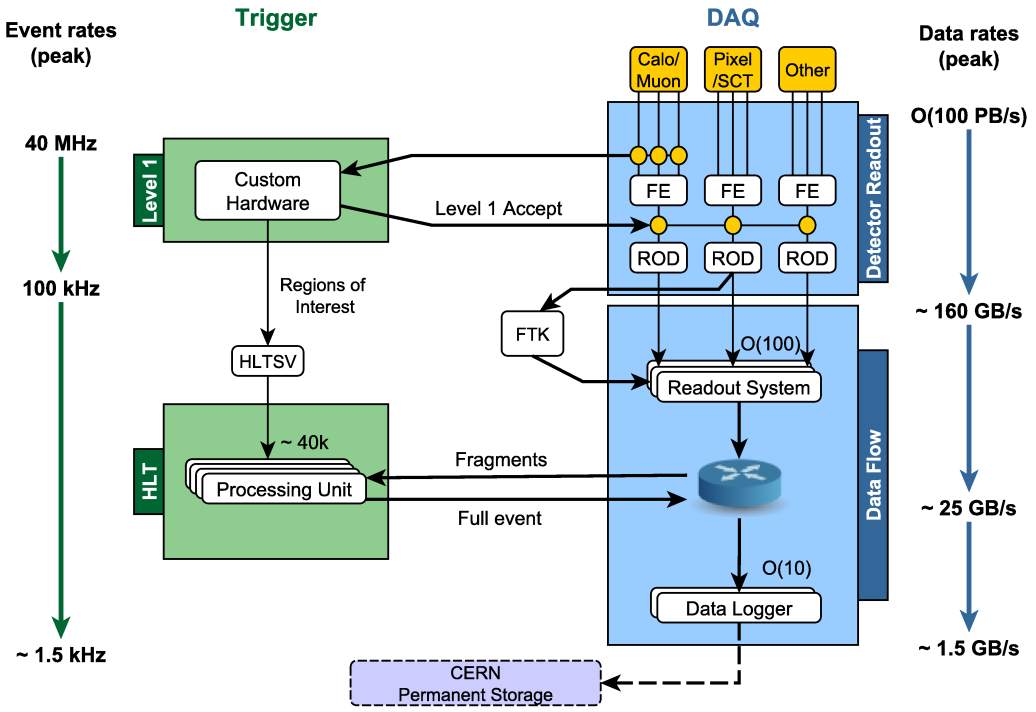

Figure 35: Functional diagram of the ATLAS Trigger and Data Acquisition system at LHC Run2.

The trigger rate with same luminosity and trigger criteria increased at LHC Run2 due to the increase of the center of mass energy from 8 TeV to 13 TeV. To cope with the increase of trigger rate, several improvements were implemented. For example, new FPGA-based topological trigger (L1Topo) [54] was added to the L1 and commissioned in 2016. The L1Topo module enables geometric and kinetic selection for trigger objects forwarded from the L1Calo and the L1Muon. In addition the merge of the Level-2 trigger and the Event Filter at the HLT improved usage of the CPU resource and removed limit of the bandwidth between the two stages.

To meet trigger rate and bandwidth, trigger menu defines the list of the L1 and the HLT consisting of:

Primary Triggers
Used for physics analysis.

Support Triggers
Used for efficiency and performance measurements and monitoring. The rate is low, prescale is set.

Alternative Triggers
Implements alternative reconstruction algorithms to compare performance with primary triggers. They have large overlap with primary triggers.

Backup Triggers
Implements tighter selections and operates with low rate.

Calibration Triggers
Used for calibration of the detector. They are typically high rate, but stores very few events necessary for calibration.

Constituents of the trigger menu and trigger threshold are optimized to maximize physics output while the rate and bandwidth meet the limitation. To constitute optimized trigger menu while satisfying the limitation, prescale factors can be applied which reduces trigger output rate by accepting pre-defined fraction of events. The fraction can be changed during physics run to be optimized for luminosity decrease.

An example of physics trigger rate for the L1 trigger and the HLT as a function of the number of luminosity blocks are shown in Fig. 36. It was taken in July 2016 with a peak luminosity of 1.02×10^{34} cm^{-2}s^{-1} which corresponds to the average number of interaction is 24.2. A luminosity block corresponds to on average 60 s. The rate for individual trigger physics objects are shown. Common structure for all the rates are due to their exponential decay with decreasing luminosity during a LHC fill. The rates periodically increase by the change of prescales.

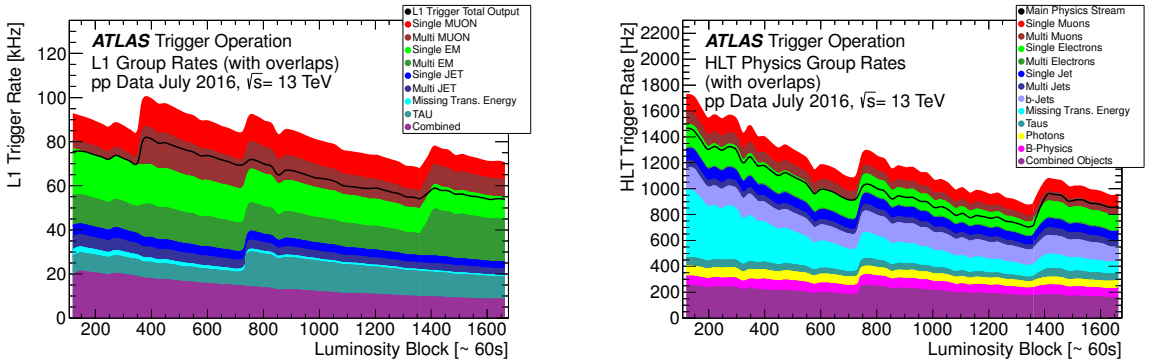

Figure 36: Physics trigger rate for the L1 trigger (left) and the HLT (right) as a function of the number of luminosity blocks [55].

Track Reconstruction at HLT

Track reconstruction at the HLT consists of two stages, which are the fast tracking based on trigger-specific pattern recognition algorithms and the precision tracking based on algorithms used at the offline reconstruction. These algorithms run in only RoIs defined by L1.

Tracking efficiency with respect to offline tracks are measured for electrons and muons. Tracks are required to have two hits in the Pixel detector and six hits in the SCT, and to be within coverage of the inner detector ($|\eta| < 2.5$). Tracks reconstructed at the HLT are considered as match with offline tracks if they satisfy $\Delta R < 0.05$ with the closest offline track. Figure 37 shows the track reconstruction efficiency

with respect to offline tracks as a function of p_T for tracks of $p_T > 20$ GeV of the 24 GeV electron trigger (left) and for tracks of $p_T > 6$ GeV of the 6 GeV muon trigger (right). The overall efficiency for fast and precision tracking exceeds 99%. Small efficiency loss can be seen at lower p_T region for electron due to bremsstrahlung energy loss.

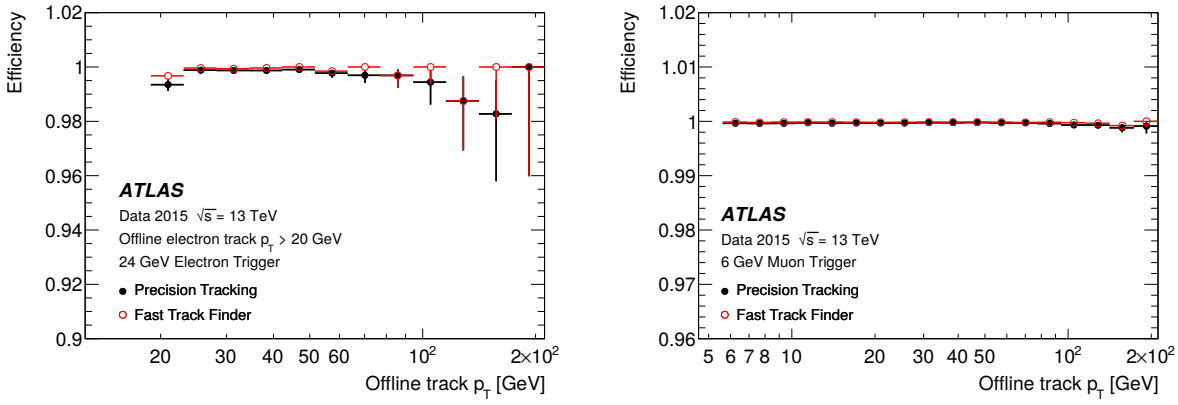

Figure 37: Tracking efficiency at the HLT for tracks of 24 GeV electron trigger (left) and 6 GeV muon trigger (right).

Track reconstruction efficiency with respect to offline tracks are also measured for tracks with $p_T > 1$ GeV within jets. The efficiency for events of jet 55 GeV trigger is shown in Fig. 38. The offline tracks are required to have at least one pixel cluster and at least 4 SCT clusters, with no more than two holes in the silicon detectors and should have at least one hit in the innermost pixel layer. High overall reconstruction efficiency is achieved for tracks within jets.

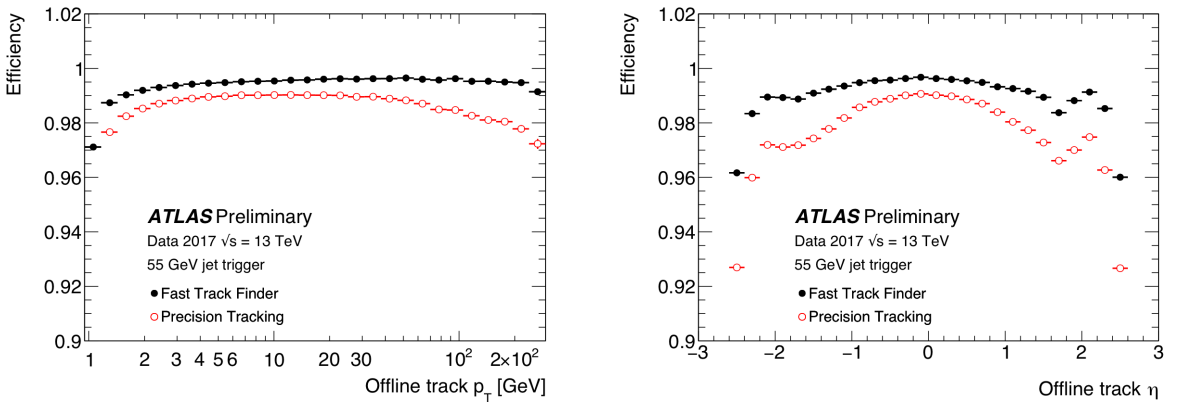

Figure 38: Tracking efficiency at the HLT for tracks within a jet for 55 GeV jet trigger as a function of p_T (left) and η (right) [56].

The processing time of the fast tracking and the precision tracking for the electron trigger per RoI for events which pass 24 GeV electron trigger is shown in Fig. 39. The precision tracking takes less processing time since it utilizes tracks found by fast tracking as a seed. Tracks are reconstructed in order of 10 ms for each RoI.

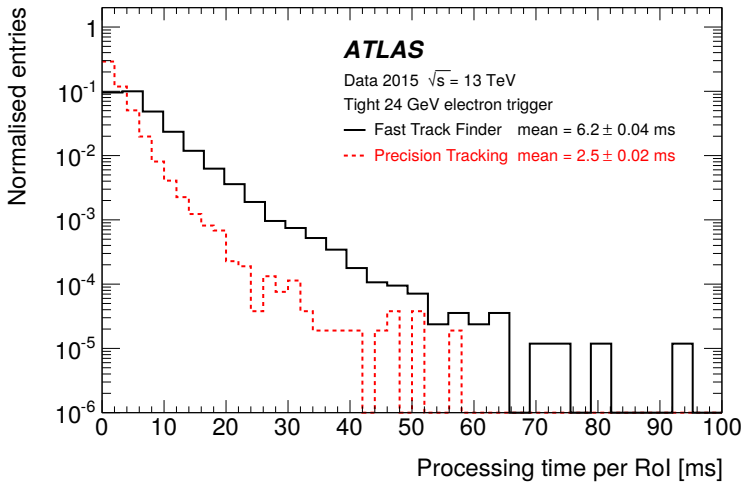

Figure 39: The processing time of the fast tracking and the precision tracking for the electron trigger per RoI for events which pass 24 GeV electron trigger.

3 The Fast TracKer

At LHC Run 2, which started operation in June 2015 at a center of mass energy of 13 TeV, the instantaneous luminosity has already exceeded its design value of 1.0×10^{34} cm^{-2}s^{-1}, while the ATLAS trigger system is designed to work under the design luminosity. Many simultaneous interactions per bunch crossing make detector measure worse, and it is challenging especially for trigger system to separate signal events and background events. In addition, it is important to perform precise object identification such as b-jet and τs at the trigger level. One of the key ingredients to cope with these challenges is "track information" which has high resolution and granularity. As described in Sec. 1.3, track information plays critical role to identify b-jets and hadronically decaying τs as well as vertex reconstruction. In contrast, it takes much processing time and a lot of CPUs to reconstruct track information. Thus current L1 trigger does not use it, while the HLT uses it but only within RoIs and takes much time to reconstruct it.

The FTK [57] is combination of electric circuit boards which are newly installed to ATLAS to make better use of track information at the trigger stage. The FTK is a highly-parallel hardware system that finds and reconstructs all tracks with $p_T > 1$ GeV in the entire detector region ($|\eta| < 2.5$) for every event that passes the L1 trigger at a maximum event rate of 100 kHz. The reconstructed track information and hit information are sent to the ATLAS ReadOut System (ROS) and used at the HLT. The hit information stands for cluster centroid and charge deposition of particles at the IBL, the Pixel detector, and the SCT. The FTK is based on the success of the Silicon Vertex Tracker (SVT) [58, 59] implemented at the Collider Detector at Fermilab (CDF).

After the FTK starts the operation, the HLT can access track information earlier and does not need to reconstruct tracks. This enables to implement more sophisticated time-consuming algorithms such as tighter b-tagging which can make jet energy threshold lower. The HLT can also refit tracks using FTK tracks as a seed and reconstruct better quality tracks. The FTK reconstructs tracks for the entire region of inner detector not only RoIs so that the HLT can reconstruct all the vertices in an event which makes better separation of primary vertex and pile-ups. The target instantaneous luminosity of the FTK operation is 3.0×10^{34} cm^{-2}s^{-1}, which will be reached at LHC Run3 starting from 2021 until 2023.

3.1 Working Principle of the FTK

The FTK is a massively parallel processing hardware system which reconstructs track information. It uses 1 layer of the IBL, 3 layers of the Pixel detector, and 8 layers of the SCT, in total 12 layer information, which consists of about 100M readout channels. The FTK gets the identical copy of the inner detector data which is sent to the ATLAS ROS. The FTK receives hit information at a maximum input event rate of 100 kHz. To deal with such large input rate, the FTK segments detector regions into 64 η-ϕ towers, each has its own processor which runs in parallel. The segmentation is 16 in ϕ and 4 in η. Each tower has overlap region to avoid inefficiency at tower boundaries due to the finite size of beam luminous region in z direction and the finite curvature of charged particles in the magnetic field. The segmentation is depicted in Fig. 40.

The FTK operates with two stages. At first stage, 3 layers of the Pixel detector and 5 layers of the SCT, in total 8 layers are used to find track candidates quickly. At second stage, track fitting is performed with entire 12 layer information to improve resolution of track parameters and to reduce fake tracks. Key algorithms for fast track reconstruction are "pattern matching" and "linear approximation". At the pattern matching, track candidates are found among pre-calculated track patterns, which are implemented

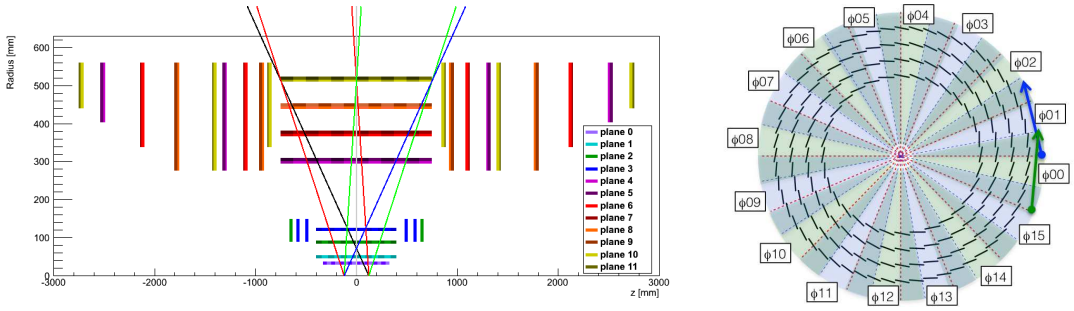

Figure 40: A schematic view of the detector segmentation for the FTK processing. η is segmented into 4 regions and ϕ is into 16 regions. The FTK algorithms run in parallel with these 64 regions.

on ASIC chips designed and optimized for FTK. After track candidates are found, track fitting of linear approximation is performed by pre-calculated fit constants. These processes run in parallel with 64 η-ϕ towers.

3.1.1 Pattern Matching

The pattern matching finds track candidates using information of 3 Pixel detector layers and 5 SCT layers, in total 8 layers. The algorithm is as following.

Before Operation

1. Prepare huge amount of track patterns (> 1 billion patterns) in advance which consist of coarse resolution hit information which is called Super-Strip (SS).

During Operation

2. Transform hit information from inner detector into the SSs.
3. Compare the input SS and the SSs in the prepared patterns, and mark all of the matched SSs in the prepared patterns.
4. Detect patterns which all layers or 7 out of 8 layers of the SSs are marked.
5. Iterate processes 2 to 4 until all the hit information in an event are loaded.

A schematic view of the pattern matching algorithm is shown in Fig. 41. This is a toy detector assuming 4 layers instead of 8 layers. The blue dot stands for strip/pixel, the red dot for strip/pixel on which track passes in the event, the blue rectangle for modules, and the yellow rectangle for SS. Prepared patterns are shown in Fig. 41 right. Each track candidate is often called as "road". If a track passes as a red winding line shown in Fig. 41 left, the SS number of 6, 8, 11, and 13 are fired. They are searched one by one in the large amount of pre-calculated patterns. In this case, all the SSs in Road4 are marked, then the road is detected as track candidate.

When the input SS matches with the SS in the stored pattern, Set-Reset Flip-Flop (SR-FF) becomes high. Since it is just taking logical "AND", the SS comparison with all patterns finishes at one clock after the hit is loaded. Thus pattern matching is completed at the next clock after the final hit in the event is received.

The patterns are stored in ASIC chips developed specifically for FTK, which are called AM chips [60, 61]. The AM chip has been developed significantly for recent years. The initial SVT at the CDF experiment utilized LVSI to store patterns. It works at 30 MHz clock cycle and can store 128 patterns per chip. Now the FTK utilizes the AMChip06 which works at 100 MHz clock cycle and can store 128k patterns per chip. The initial SVT holds 256 LVSI chips thus it can hold 32k patterns in total. In contrast, FTK uses 8192 chips thus it can store more than 1 billion patterns in total. The development of the AM chip is summarized in Tab. 4.

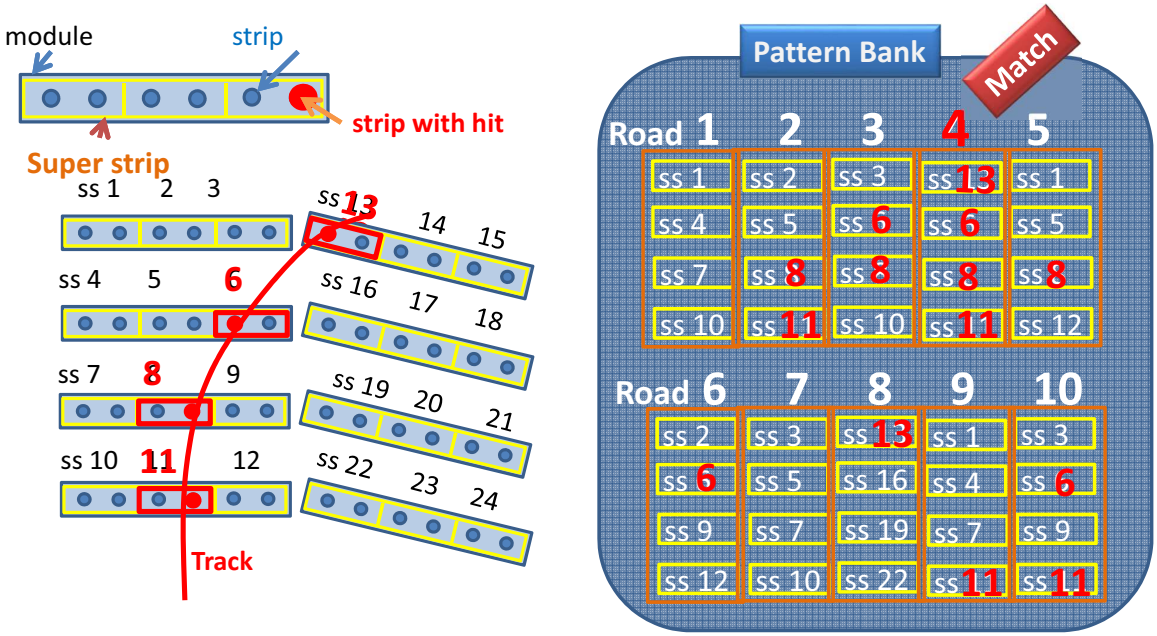

Figure 41: A schematic view of the pattern matching algorithm. This is a toy detector assuming 4 layers though 8 layers are used at the real use case. The blue dots stand for strip/pixel, the red dots for strip/pixel which track passes in the event, the blue rectangle for modules, and the yellow rectangle for SS.

Chip	Technology	#Patterns/chip	#Chips	Clock (MHz)	I/O	Purpose
LVSI	700 nm	128	256	30	parallel bus	CDF SVT (1992)
AMChip03	180 nm	5k	100	50	parallel bus	CDF SVT (2003)
AMChip04	65 nm	8k		100	parallel bus	FTK R&D
AMChip06	65 nm	128k	8192	100	serdes	FTK

Table 4: The history of AM chip development.

3.1.2 Track Fitting

The FTK reconstructs tracks rapidly by performing a linear approximation of local hit positions with full resolution instead of performing helical fit. The term "track fitting" and "track fitter" stand for this linear approximation in this thesis unless specified otherwise. Five track parameters and χ^2 components are calculated by this linear approximation. The linear approximation is expressed by a set of scalar products of hit coordinates and pre-calculated fit constants that take geometry and alignment of the detector into account. The equation for five track parameters p_i $(1 \leq i \leq 5)$ is expressed as following.

$$p_i = \sum_{l=1}^{N} C_{il} x_l + q_i, \tag{46}$$

where C_{il} and q_i are pre-calculated constants and x_l are N hit coordinates. χ^2 is defined as:

$$\chi^2 = \sum_{i=1}^{5} \Big(\sum_{j=1}^{11} S_{ij} x_j + h_i \Big)^2, \tag{47}$$

where S_{ij} and h_i are pre-calculated constants. If hit coordinates are given, track parameters and χ^2 components are calculated immediately just as referring to Look Up Table (LUT). Resolution of the track parameters obtained by this method is almost same as the one by helical fit if the region which each fit constant covers is enough small. In the FTK that region is called "sector" which consists of typically a few cm wide single silicon module in each layer. A set of fit constants are defined for each sector.

At the first stage, track fitting is performed with 8 layers information of the inner detector, and tracks which pass χ^2 cut are passed to the second stage processing. If the track has hits in all layers and it fails χ^2 cut but it is not so large value, the track is refitted a number of times by removing a hit in a layer each time. This is called "recovery fit" and allows the case that a hit is lost due to detector inefficiency and random hit is picked up instead. If more than two tracks pass χ^2 cut in a road and they share more than a pre-defined number of hits, they are considered as duplicated tracks and only the best track is kept based on the χ^2 and the number of hits it has.

Tracks passing the first stage are sent to the second stage where track fitting is performed using all 12 layers by extrapolating 4 layers which are not used at the first stage. Here 2 missing hits are allowed, but both are not in the Pixel detector or the SCT. χ^2 cut is applied and duplicated tracks are removed with the same way as the first stage. Here duplicated tracks are searched not only in a road but in the entire event. The tracks passing the second stage are sent to the ATLAS ROS and used at the HLT.

Optimization of Track Reconstruction Performance: Don't Care Bit Feature

It is one of the most important and complicated tasks for the FTK development to optimize resources for the pattern matching and the track fitting. Since the FTK patterns have some pre-defined size (i.e. SS size), it is possible that a road has silicon hits from multiple tracks and even detector noise. This may produce fake roads which are not derived from a single particle. The number of fake roads is higher when the SS size is wider and hit occupancy is higher. In addition, the number of fits in a road increases when the SS size is wider since the road can have more hits and then hit combinatorics of one hit per layer increases. The number of hit combinatorics becomes higher exponentially when the number of hits in a layer increases.

For these reasons, the quality of the pattern bank is determined by the coverage and the rate of fake roads. The coverage is defined as the fraction of tracks matching with at least one of the patterns in the bank. This only represents the geometric coverage of the bank. In contrast, tracking efficiency of the FTK includes all of the FTK algorithms such as clustering of hit information and efficiency of the track fitting which is

defined as χ^2 cut. If the SS size is fixed, coverage of the bank is higher when the number of patterns in the bank is higher, while the number of fake roads is approximately proportional to the number of patterns. In addition, when instantaneous luminosity increases, the number of fake roads increases rapidly since hit occupancy of detector increases and hit combinatorics increases exponentially. To avoid producing many fake roads, the SS size needs to be narrow. If the SSs of narrower size are implemented with a standard AM chip technology, the number of patterns increases to keep coverage, which makes the FTK system very large and expensive.

One of the way to solve this is to use variable resolution patterns. Each pattern and even each layer in a pattern can have optimal width by applying "Don't-Care" (CD) feature [62] (which is inspired by ternary content-addressable memories or CAMs). This feature enables to keep track reconstruction efficiency while reducing fake rate and the number of patterns in a bank which would be larger by only reducing the overall SS size.

The shape of patterns which have the optimized SS size by the DC-bit is quite different from the one without the DC-bit. When many tracks pass wide range of patterns or layers, applying the DC-bit for those patterns or layers would produce many narrower patterns and the number of fake roads does not decrease so much. Thus these patterns or layers need to be kept wider. In contrast, there are patterns or layers that tracks are traversing just locally in the SS. Those patterns or layers need to be narrower to keep the bank size and the fake rate small. An example of the optimization of pattern size is shown in Fig. 42.

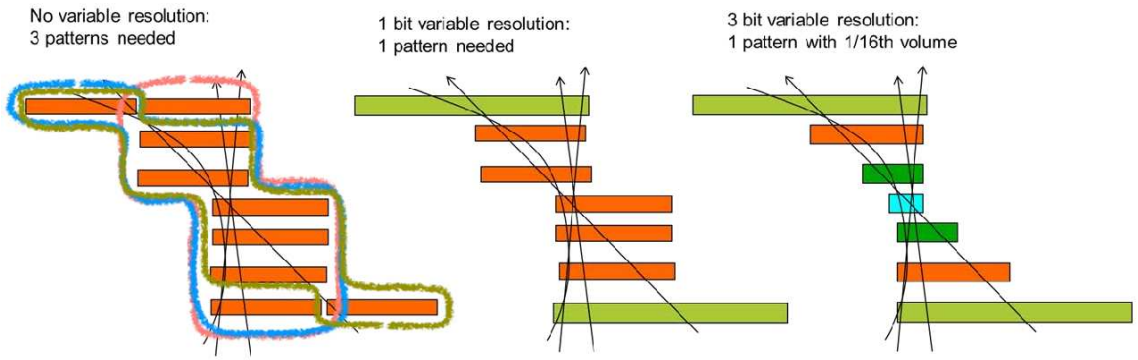

Figure 42: A schematic view of the example of applying the DC-bit.

The optimal pattern width is determined layer by layer based on simulation. For example, when the SS in a specific layer is divided into two parts, in some cases tracks traverse both halves, and in other cases tracks only pass either of the half. For the former case, the DC-bit is applied to the layer for the least significant bit of the pattern word which makes the layer wider by a factor of two during pattern matching process. For the latter case, the layer is kept with high resolution and reduce fake rate generated by uncorrelated hits from multiple tracks or detector noise.

Optimization of Track Reconstruction Performance: Wild Card Algorithm

Some modules of the inner detector are disabled due to the defects of module itself or of its readout links. If a track passes more than two disabled modules, the FTK cannot reconstruct that track since the FTK requires to have hits in at least 7 out of 8 layers. The number of disabled modules is expected to increase in the future since the detector will run in radiation harder environment by the LHC machine upgrade.

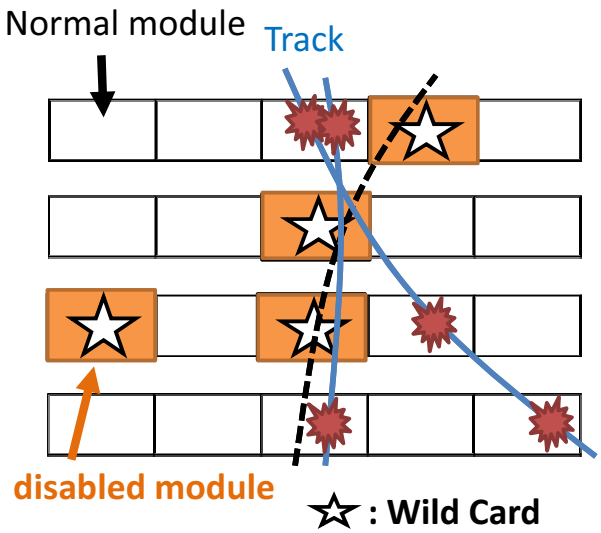

Figure 43: A schematic view of detector layers with disabled modules and the Wild Cards.

To recover this inefficiency, the FTK applies "Wild Card" algorithm which assumes all the SSs in the disabled modules are fired to enable pattern matching for the patterns which have disabled modules in more than two layers. Though the Wild Card algorithm recovers inefficiency caused by the disabled modules, the fake rate and the number of fits will increase, and not certain how it affects to the resolution of reconstructed tracks. Thus it is necessary to validate and optimize the way to apply the Wild Card and take the best compromise of the performance.

3.1.3 Data Flow in the FTK system

First FTK receives hit information from IBL, Pixel detector and SCT, clusters and formats it. The clustered information is distributed into parallel processors based on the detector region. Parallel processors transform cluster information into SSs, and then perform pattern matching and track fitting. The track parameters, cluster information and χ^2 components are forwarded to ATLAS ROS. An overview of the FTK main algorithms are depicted in Fig. 44.

The basic concept of dataflow in the FTK system is simple pipeline handled by control words. 8b/10b encoding is used for serial data stream. Input words in each processing engine are pushed into First-In First-Out (FIFO) buffer. All the valid words with write enable bit applied enter the FIFO. The FIFO buffer is placed without depending on functions in the destination. Source and destination are two different logic functions which are placed in other boards or in some cases other functions in same boards. There is no handshake word by word to maximize processing speed. A HOLD signal is implemented and issued to prevent data loss when destination is busy and the FIFO becomes ALMOST_FULL. To use ALMOST_FULL instead of FULL gives source enough time to respond to the HOLD signal and suspend dataflow. In case the communication between two boards is performed via optical fiber, this mechanism is implemented as XOFF protocol. Since any board or function in the FTK system can have multiple input channels, same event arrives from different sources at different timing. Thus event synchronization needs to be performed among input channels of the function. The FIFOs are placed at the input of functions for this purpose, and they need to be enough deep to cover the fluctuation of arrival time among input channels.

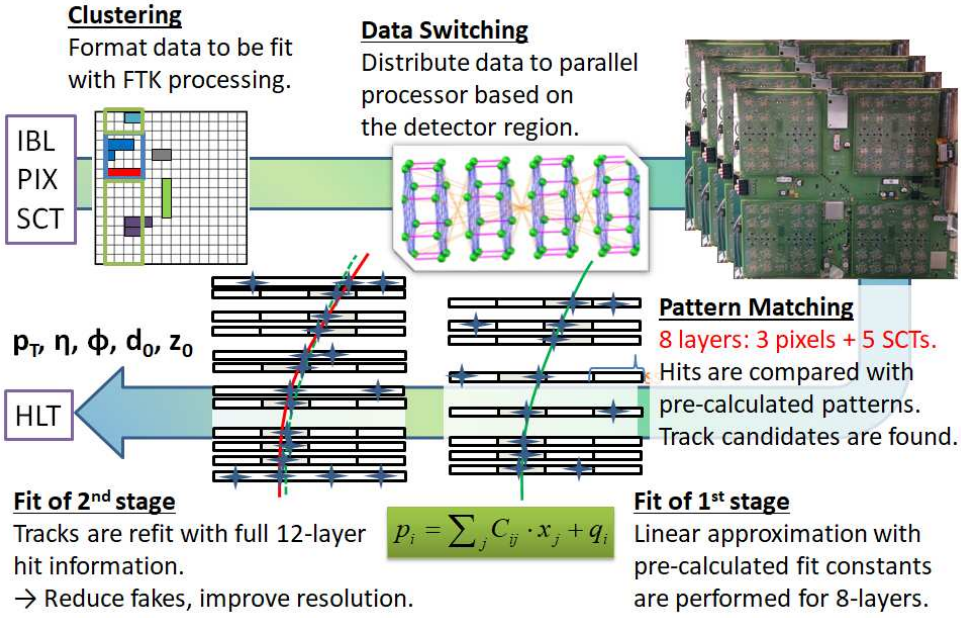

Figure 44: A schematic view of main algorithms of the FTK.

3.2 FTK Hardware Subsystems

The FTK consists of several electronics circuit boards controlled by Field Programmable Gate Arrays (FPGAs). A functional overview of the system is shown in Fig. 45. The data is processed as following in the FTK system.

1. The Dual-Output HOLA or the BOC provides identical copy of hit information from the IBL/Pixel/SCT Read-Out Drivers (RODs) to the FTK Input Mezzanine Card (IM).
2. The IM receives the hit information, performs clustering and sends the cluster information to its mother board, the Data Formatter (DF).
3. The DF shares the clusters among overlap regions, and distributes the clusters to appropriate parallel processors of the AUXiliary Card (AUX) for the first stage processing, and the Second Stage Board (SSB) for the second stage processing.
4. The AUX transforms the clusters into the SSs and forwards the SSs to the AM.
5. The AM performs pattern matching, and returns matched roads to the AUX.
6. The AUX transforms the SSs in a road into original clusters and performs track fitting using 8 layers of the Pixel detector and the SCT. χ^2 component and hit information for tracks passing χ^2 cut are sent to the SSB.
7. The SSB performs track fitting using entire 12 layer information, and sends the tracks which pass χ^2 cut and duplicated track removal to the FTK to the Level-2 Interface Card (FLIC).
8. The FLIC transforms the data of the FTK format into the ATLAS global format, and forwards the track information and hit information to the ATLAS ROS.

The detail of each subsystem is described in the following subsections.

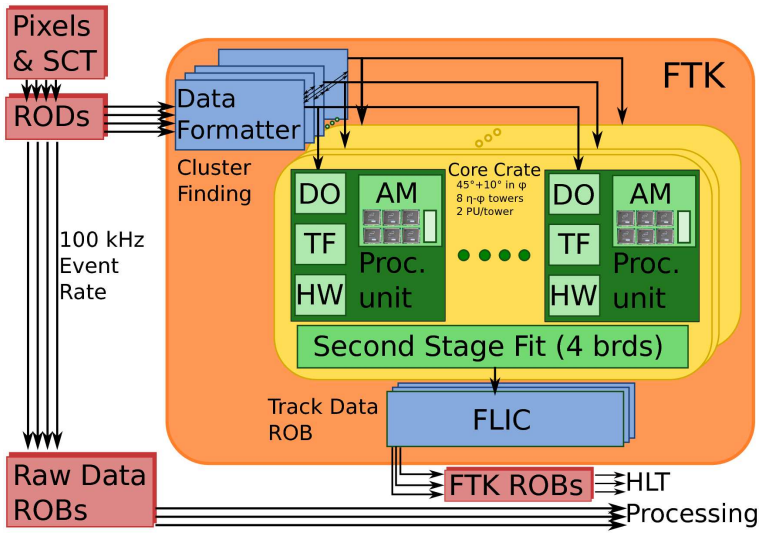

Figure 45: Functional overview of the FTK system.

3.2.1 Dual-Output HOLA

The Pixel detector and the SCT RODs send silicon hit information to the ATLAS ROSs for each L1 trigger accept. HOLA mezzanine cards mounted on the RODs transform parallel data into serial data, and transmit the data via optical fibers which implement the SLINK protocol. To send data to the FTK, the HOLA with two output ports was designed and produced, which is called Dual-Output HOLA (D-HOLA). A photo of the D-HOLA is shown in Fig. 46. 32-bit words from the RODs are entering the HOLAs at a maximum rate of 40 MHz. A Serializer-Deserializer (SERDES) which was originally placed at external part of FPGA is implemented as a pair of function within FPGA. The SERDES serializes 32-bit parallel data and feeds them to a pair of optical fibers.

The channel on the D-HOLA which feeds data to the ATLAS ROS is called DAQ channel, while the other channel which feeds data to the FTK is called FTK channel. The DAQ channel implements full SLINK protocol including link down, link reset and general return signals. In contrast, the FTK channel implements only XOFF protocol to suspend dataflow. The XOFF protocol needs to work only at the normal FTK operation, and it must be under control to install and commission the FTK without interfering ATLAS data taking during LHC physics runs. Thus XOFF enable register is implemented on the FTK channel of the D-HOLA and it is set to 0 by default, and not set to 1 until special control words are coming from the IM. The protocol to start data communication occurs between the DAQ channel on the D-HOLA and the ATLAS ROS. The link comes up even if a fiber is not connected to the FTK channel. Modified SLINK protocol is implemented on the FTK channel which does not require initial handshake. Thus the link on the FTK channel comes up immediately when a fiber is connected and starts sending data.

A prototype D-HOLA was produced in 2011, and a series of thorough tests including bit-error-rate test which requires the error rate of entire board to be $< 1.0 \times 10^{-15}$ were performed to the board. It passes CERN production readiness review, and all the D-HOLAs which are to be used for the FTK were already produced, tested and installed, since the D-HOLAs feed data to the FTK and it is required to be ready to test other boards of the FTK system.

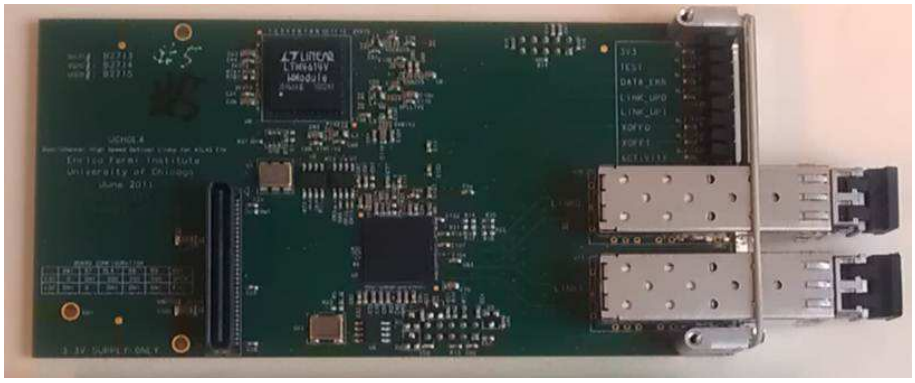

Figure 46: A photo of the dual-output HOLA.

3.2.2 Input Mezzanine

Functions of the IM [63] are to receive the IBL, the Pixel detector and the SCT data from the inner detector RODs, to perform clustering, and then to forward the clusters to the DF. Each IM receives four channels at maximum. Two FPGA are mounted on a board, and each FPGA can process up to two channels that one is from the IBL or the Pixel detector and the other from the SCT. Four data streams are processed and sent to the DF independently. Thus event synchronization are performed on the DF at first in the FTK system.

An IM board is a mezzanine card with the size of 149 mm × 74 mm consisting of 12 layers. It connects to the DF with a High Pin Count FPGA Mezzanine Card (FMC) connector. Two FPGAs are mounted on a board. Four SFP+ connectors receive data from four SLINK protocol optical fibers, and the data are sent to two FPGAs directly. There are two kinds of the IM boards, one implements Spartan-6 FPGA (XC6SLX150T) and the other Artix-7 FPGA (XC7A200T). They are called S6 IM and A7 IM, respectively. The A7 IMs have more FPGA resource than the S6 IMs have, thus the A7 IMs are used to cluster hits from the IBL which has high occupancy, while the S6 IMs are not used to cluster hits from the IBL. Photos of the S6 IM and the A7 IM are shown in Fig. 47. Each FPGA on a board has a 18 Mb external RAM and a 32 Mb flash memory. Power is provided from mother board via FMC connector or from external power connector which is used to power the IM standalone where there is no mother board. JTAG chain for FPGA configuration is accessible from both the FMC connector and an external connector. The SLINK runs 3.1 Gb/s at the maximum.

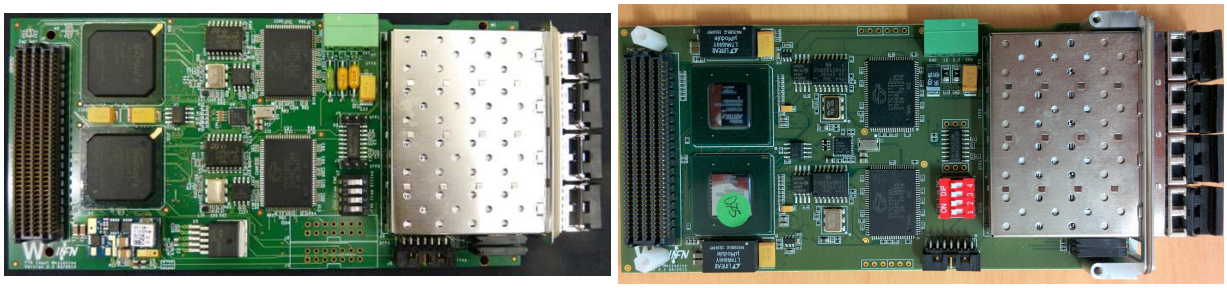

Figure 47: Photos of the IM with Spartan-6 FPGA (left) and Artix-7 FPGA (right).

Clustering Algorithm

One of the main functions of the IM is to cluster hit information from the IBL, the Pixel detector and the SCT. The purpose of the clustering is to reduce data volume, to transform the data to the appropriate format for the FTK processing, and to determine cluster centroid. 2D-clustering algorithm is implemented to cluster hits from the IBL and the Pixel detector, while 1D-clustering is implemented for the SCT.

1D-clustering transforms the data from the SCT RODs to clusters of the FTK format. This algorithm finds micro-strips fired consecutively, and compresses them into 32-bit data which represents the size of the cluster and twice of the position of cluster centroid. The SCT has three modes to format data which are expanded mode, condensed mode and super condensed mode. Appropriate clustering algorithms for each data format are implemented in the IM firmware.

It is challenging to perform 2D-cluster finding without significant delay in latency. A multi-core FPGA-based 2D-clustering algorithm are used to cluster the IBL and the Pixel detector data. The algorithm implements a sliding window technique which utilizes windows of pre-defined size to minimize FPGA resources for cluster identification. A key point of this algorithm is to instantiate clustering engines which work for each sliding window independently in parallel to maximize performance by better use of FPGA resources. In addition to this parallelization, the clustering algorithm is executed as a pipeline so that data preparation, clustering and post clustering processes as determining cluster centroid work in parallel.

The incoming data is transformed from the original ROD format data to the format appropriate for the FTK processing. As a pre-processing step, it is classified such as event header/trailer words, module header/trailer words, and detector hit words, and formatted to be fitted with clustering algorithm. Along with transformation of the data, the incoming data is re-ordered. A pixel module has 16 FE chips in 2 columns so that 8 chips are in each column. They are numbered anti-clock wise, while they are readout from the same side. Thus FE chips on one side are readout in reverse order with respect to the other side. Since the clustering algorithm assumes that the FE chips are readout with increasing FE chip number, they need to be re-ordered to be as such. This is achieved by pushing the FE chips which are readout in reverse order into Last-In First-Out (LIFO) buffer. For the case of the IBL, it has two types of modules, called planar modules and 3D modules. The order to readout FE chips on the planar modules is shown in Fig. 48. There are 2 FE chips on a module, and each FE chip is divided into two parts. First, FE chips on the left half of the right module are readout from left side (1), and then the right half of the right module are readout from the right side (2). Next, the left half of the left module are readout from left side (3), and finally the right half of the left module are readout from the right side (4). In the IM, they are re-ordered by pushing the FE chips of (1) and (3) into FIFO, then (2) and (4) into LIFO. For the case of the 3D module, there are only (1) and (2).

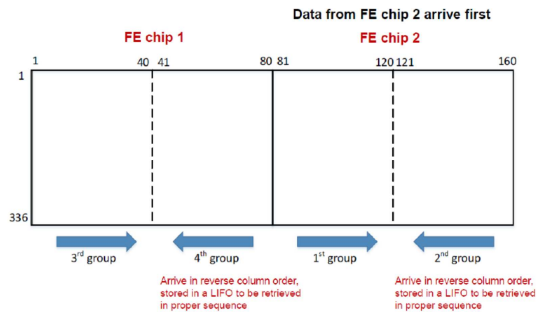

Figure 48: A schematic view of the readout order of FE chips of the IBL planar module.

Cluster finding logic starts when first hit arrives to the logic. It defines cluster window (the size is 21×8 by default) with respect to the first hit position. Then it loads all hits within the window. When all hits within the window are loaded, all neighboring hits sitting next to the first hit are selected in a clock. This is a key part of the algorithm to avoid a lot of hit loops and reduce processing time. The selection of hits keep running until all the hits within the window are selected. Finally cluster centroid is calculated among the selected hits. The cluster finding procedure is depicted in Fig. 49.

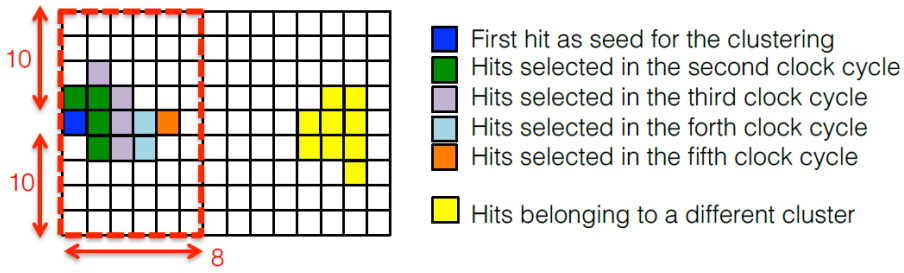

Figure 49: Cluster finding procedure with the sliding window technique and two-dimensional hit selection method.

The post processing step of clustering performs further reduction of data volume and improves precision of cluster centroid. Information on charge which is provided by the detector FE chips as Time-Over-Threshold (ToT) can be taken into account for cluster centroid calculation.

3.2.3 Data Formatter

The FTK consists of parallel engines which work in parallel for 64 towers. The DF [64] receives clusters from the IM, remaps the ATLAS detector hit information into the FTK tower structure, and sends the clusters to the processor units. A photo and schematic view of the DF are shown in Fig. 50.

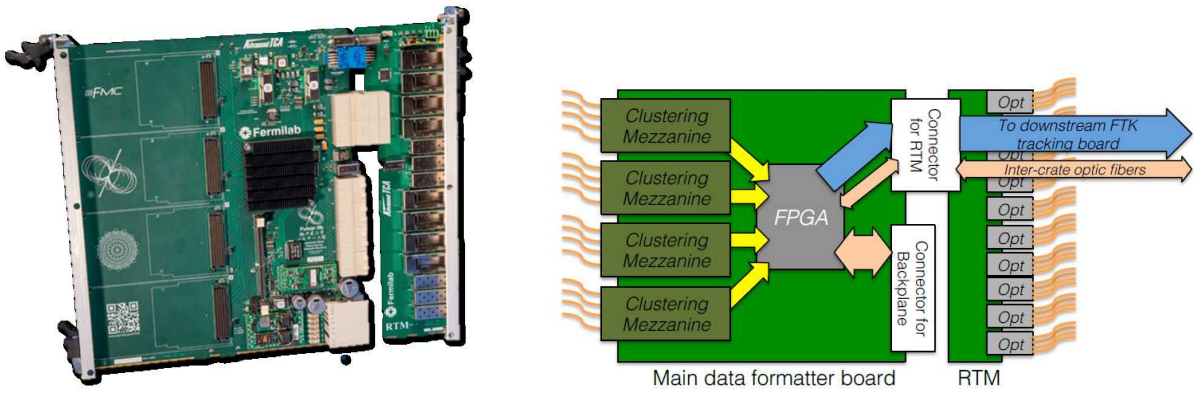

Figure 50: A photo (left) and a functional view (right) of the DF and its RTM.

The input fiber configuration of the FTK system is arranged to minimize data sharing at the DF. Figure 51 left shows data sharing that is needed among the towers in 64×64 matrix after optimization of input fiber configuration. The off diagonal elements represent data sharing between different towers. The red boxes represent 4 crates which are optimized to minimize inter-crate communication.

The ideal hardware platform of data distribution should be enough flexible for future expansion and the change of input fiber and module configuration. To match with these requirements, the function of the DF is implemented on a board which is mounted on a full-mesh Advanced Telecommunications Computing Architecture (ATCA) back plane. (DF board is called "Pulsar-IIb" in general when it is apart from FTK project.) A photo of the ATCA crate is shown in Fig. 51 right. The boards within the same crate can communicate each other over multiple point-to-point links. Data transmission between crates are performed via optical fibers. 32 DF boards are used to process 64-towers, and in total 4 ATCA crates are used so that one ATCA crate holds 8 DF boards. A Rear Transition Module (RTM) performs data sharing with other crates as well as sends data to the downstream boards.

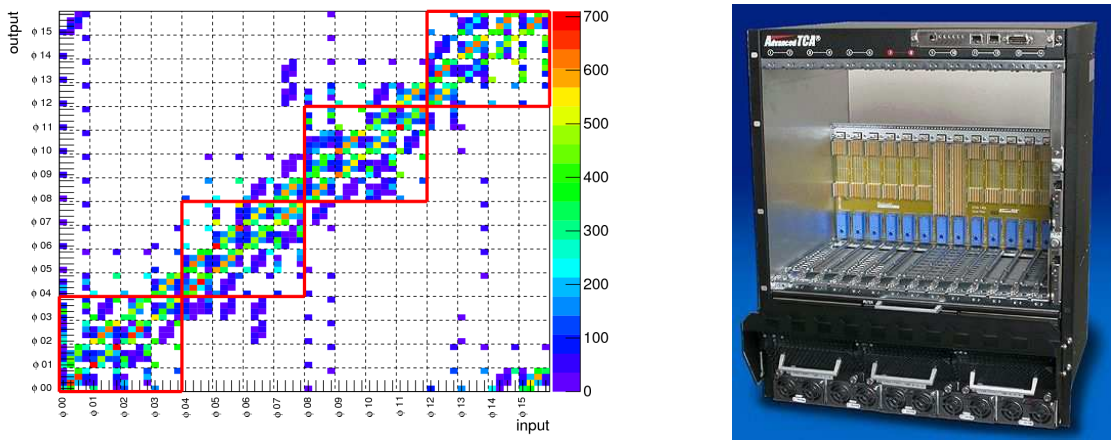

Figure 51: This 64×64 matrix shows data sharing among the FTK $\eta - \phi$ towers in the DF. The red boxes indicate the assignment of towers to 4 ATCA crates to minimize inter-crate data sharing. The color scale indicates the number of clusters shared between towers per event for 8 TeV data with 50 ns bunch separation (left). A photo of the ATCA crate (right).

The DF board connects to the IM through FMC connector, to the ATCA back plane through Zone-2 connector, and to the RTM through Zone-3 connector. Processing engine of the DF is Virtex-7 FPGA (XC7V690T-2). The FPGA has 80 high speed serial transceivers (GTH) with maximum data rate of 10 Gb/s which connect to the ATCA back plane and the RTM transceivers. There are additional circuits on a DF board which are for power distribution, slow control, and clock distribution. A RTM has 8 QSFP+ connectors and 6 SFP+ connectors. Cluster data are sent to the AUX cards for the first stage processing and to the SSB for the second stage processing from the RTM. Along with main data sharing function, an ATCA has user control interfaces. They are implemented on two dedicated slots for hub switching boards with two separate dual-star networks, one is for slow control networks which is called Base Interface, and the other is for high speed data transmission (up to 40 G Ethernet) which is called Fabric Interface. IPbus protocol is implemented [65] for board control and monitoring. UDP packets are transmitted between external computers and the DF boards via hub switching board. In addition, I2C signals to configure and monitor the IM are driven through IPbus interface, which enables to control both IM and DF from external computers.

3.2.4 AUXiliary Card

The AUX [66] receives cluster information of 3 layers of the Pixel detector and 5 layers of the SCT from the DFs, maps the cluster information into the SSs, and performs the first stage track fitting. The AUX

consists of three functions which are Data Organizer (DO), Track Fitter (TF), and Hit Warrior (HW). At first clusters are stored in the DO and mapped to the SSs which are used for pattern matching, and the SSs are sent to the AM Board. When the AM Board finds a road with at least 7 out of 8 SSs matched, road number is sent back to the AUX and all hits within the road are retrieved. Since hits are stored in the DO based on the SS address and each layer consists of a single SS in a road, hits can be retrieved rapidly. Hit information, road number, and sector number are sent to the TF function. The TF performs the first stage track fitting of linear approximation with pre-calculated fit constants which are associated to each sector. Tracks passing χ^2 requirement at the TF are sent to the HW which performs duplicated track removal.

A photo and functional diagram of the AUX are shown in Fig. 52. The AUX card is a back of crate card of 9U VME and holds six Altera Arria V FPGAs. Cluster information from the DFs enters the AUX through two QSFP+ connectors where each fiber transmits clusters of each layer at up to 6.4 Gb/s. The SSs for each cluster are generated within input FPGAs. The SSs are sent to the AM system through VME P3 connector at data rate of 2 Gb/s. The SS number and hit information of full resolution are distributed to 4 processor FPGAs. Each FPGA has functions of DO, TF, and HW in its firmware. Addresses of matched patterns are sent back from the AM system through P3 connector. Processor FPGAs receive them, extract hit information to the roads, and perform track fitting. Duplicated track removal is performed for tracks passing χ^2 requirement, and tracks passing the HW functions are sent to the SSBs via optical fibers fed by SFP+ connectors. There are 10 words information in a track, 8 of them are hit information, 1 is for road number, and the other represents sector number and the HitMap which shows layers that have hits.

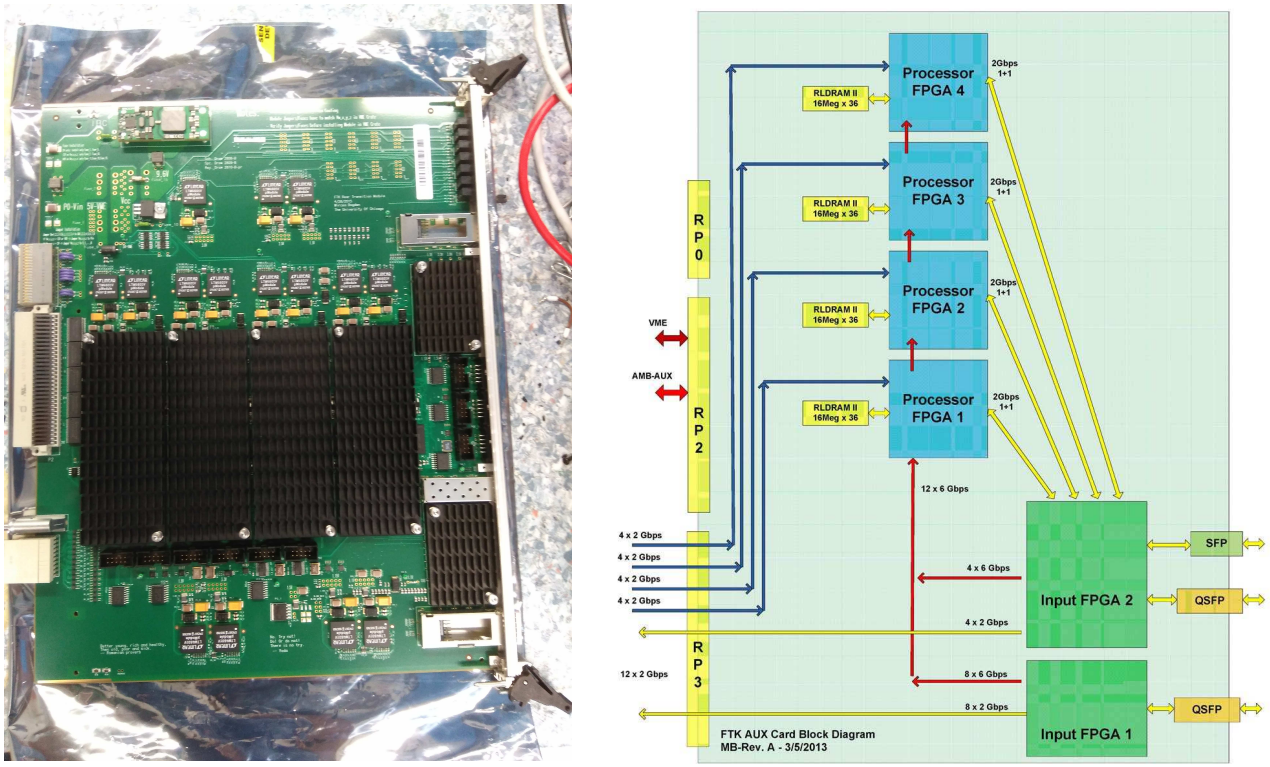

Figure 52: A photo (left) and functional overview (right) of the AUX.

Data Organizer

Input FPGAs look up SS addresses using SSMap memory for each cluster coming from the DFs. The SS addresses are necessary to perform pattern matching at the AM system and also to determine where hit information needs to be stored in the DO. The DO stores cluster information of silicon detector in a database which enables rapid extraction of all hits within a road. The database is filled with cluster information from the DFs at the start of each event. Since hits are stored according to the SS numbers, once it receives road number from the AM system, all hits in each detector layer which are associated to the SSs are retrieved and sent to the TF function immediately. The DO has two working modes which are WRITE MODE and READ MODE. The firmware is duplicated in each processor. When one of the DO is working as the WRITE MODE for event n, the other is working as the READ MODE for event $(n-1)$. In the WRITE MODE, clusters from the DFs are written to the DO, while in the READ MODE, hits are retrieved corresponding to the road number sent back from the AM system. When the WRITE MODE finishes, it switches to the READ MODE of the event. When the READ MODE finishes, it switches to the WRITE MODE of the next event. The WRITE MODE and the READ MODE are two step pipelines for each event.

Track Fitter

Track fitting of linear approximation is performed for roads instead of a helical fit as described in Sec. 3.1.2. At track fitting of first stage, 8 layers and 11 hit coordinates are used in total if there are hits on all layers. (One hit of Pixel detector has two coordinates.) Here it is enough to determine which tracks need to be sent to the second stage processing, only χ^2 components are calculated. To cover efficiency loss caused by detector inefficiency, three ways of track fitting are implemented.

Nominal Fit
: Using each combination of one hit per layer in a road.

Majority Fit
: Allowing missing hit in one layer. χ^2 is calculated for track candidates with hits in 7 out of the 8 layers by assuming hit location in the missing layer that would minimize the overall χ^2.

Recovery Fit
: If χ^2 calculated with hits on all 8 layers is above requirement a little, such track candidates are refitted a number of times with one hit in a layer dropped each time. This allows detector inefficiency and picking up a random hit from other tracks or detector noise. If a refitted track with the best χ^2 satisfies χ^2 requirement, the track is accepted.

All tracks passing χ^2 requirement are sent out of the TF.

Hit Warrior

If more than two tracks are within the same road and share more than a pre-defined number of hits, the tracks are considered as duplicated and a track with the best quality is kept. The quality is characterized by the number of hits in layers of the Pixel detector and the SCT, and χ^2 value. The track which has the most number of hits is kept at first, and if the number of hits is same, the one with the best χ^2 is kept. Tracks passing the HW function are sent to the SSB which performs the second stage processing.

3.2.5 Associative Memory

The AM system performs pattern matching by comparing a large number of pre-calculated track patterns with input SSs which are coarse resolution hit information. The AM system is organized with AM Board (AMB), Local Associative Memory Board (LAMB), and AM chips which are ASIC chips designed and optimized specifically for the FTK. The AMB is 9U VME board and 4 LAMBs are mounted on an AMB. Figure 53 shows a photo of an AMB (left) and AMB layout highlighting a LAMB on left top side in yellow (right). A network of high speed serial links determines bus distribution on an AMB. These buses are connected to the AUX card through high frequency ERNI P3 connector. The data rate for each serial link is 7 Gb/s at the maximum. 8192 chips are mounted in total on 128 AMBs.

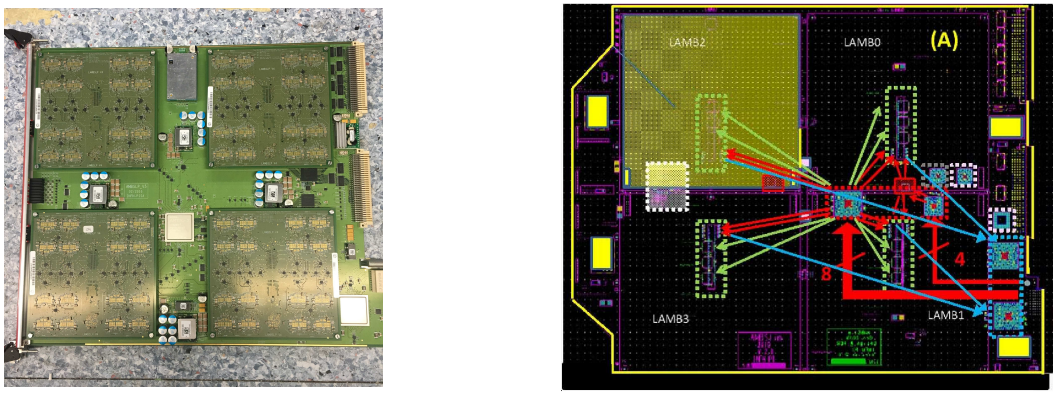

Figure 53: A photo of an AMB with 4 LAMBs mounted (left) and a schematic overview of the data flow in an AMB (right).

The AMB has flexible control logic within FPGA chips. Xilinx Artix-7 FPGAs are used which have Low-Power Gigabit Transceivers (GTP). Road IDs going out from LAMB (4 links/LAMB) are sent to the FPGAs (in the blue boxes in Fig. 53 right) near the P3 connector. The LAMB and the AMB communicate each other through SMD connector which is placed in the center of the LAMB. Each LAMB holds 16 AM chips. Each input chip receives an input road bus from the upstream chip. Roads flow in 8 pipelines, and they are gathered and merged by GLUE chip sitting at the bottom of the LAMBs. GLUE chips have high-speed GTP serial links to send roads to mother board through SMD connector. Figure 54 shows a photo of the LAMB board with 16 AM chips mounted. SSs in each event are organized into 8 buses, each bus corresponds to detector layer. The SSs are sent from the AMB to the LAMBs through SMD connector. They are distributed to the AM chips.

The important factors for the AM chip development are the pattern density and the power consumption. As described in Sec. 3.1.1, the version of the AM chip for the FTK is AMchip06 which implements a full custom cell. A full custom cell has possibility to implement new strategy to reduce power consumption of the chips. Power consumption can be reduced by the reduction of processing speed. Thus the pattern matching is performed with "pre-match" technique which compares the 4 least significant bits of pattern words in each layer in advance. Patterns passing pre-match are continued comparison. The pre-match technique can reduce power consumption up to 80% at maximum.

An AM chip consists of an array of CAM cells as shown in Fig. 55. Columns are called as search bit lines (or just bitlines) which are used to distribute the SS information through vertical buses, while rows are used as write lines (signals which enable write operation flow) and match lines which are used at the SS comparison. Each row in an array corresponds to one pattern. A row consists of sub-blocks of 18

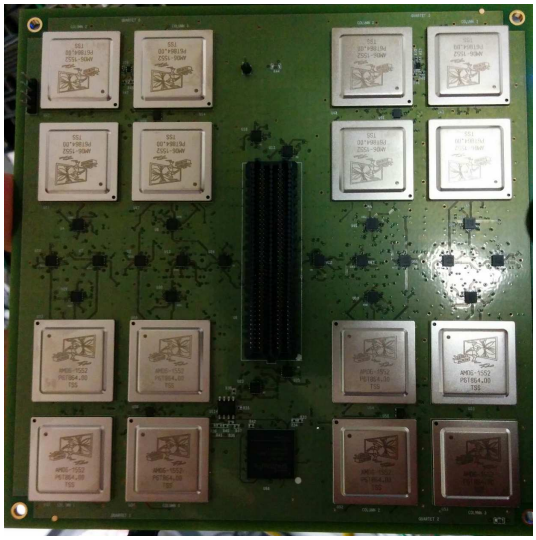

Figure 54: A photo of a LAMB board with 16 AM chips mounted.

CAM cells, called layer block. Each layer block stores hit information of a pre-calculated track pattern in that layer in a unit of a SS. Each cell in 18 CAM cells stores 1 bit, 12 bits of them are used for 12 most significant bits of the word and the other 6 bits are used as 3 pairs of ternary cells storing 0, 1 or X values. The X value represents for DC-bit. In general the least significant bit of each layer corresponds to a specific region of the detector. This region means the width of the matching window of a pattern. When the DC-bit is set to the bit, valid pattern size of the bit becomes twice, since it can match with 2 SSs. The input data flowing in the columns are compared with data stored in the layer blocks. When all 18 bits of layer blocks including the DC-bits match with input data, SR-FF is set as high value. If the number of the SR-FFs which are set as high is 7 or 8, the patterns are readout.

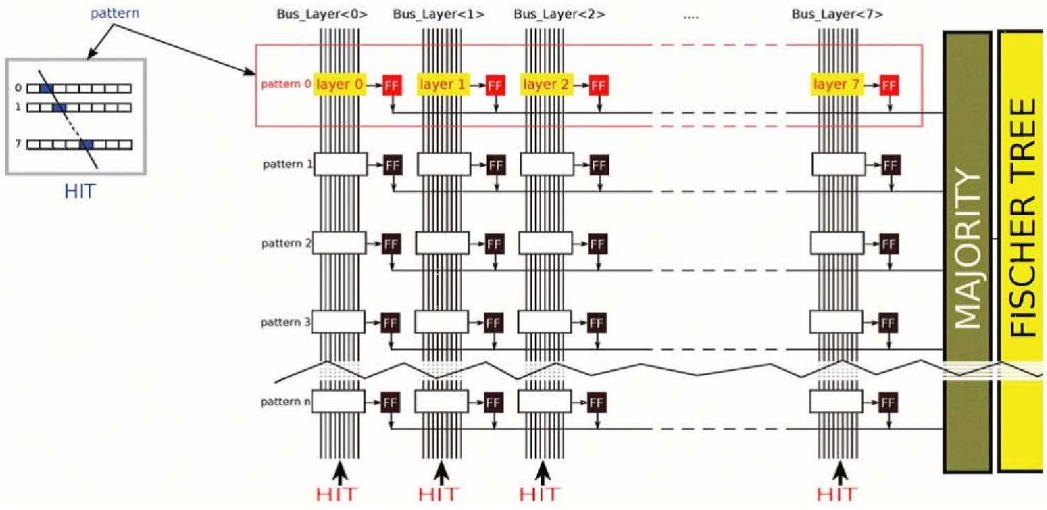

Figure 55: A schematic view of the AM chip array.

3.2.6 Second Stage Board

The SSB receives hit information, tower numbers, sector numbers and road numbers from the AUX card, and hit information in the layers that are not used at first stage processing from the DF. The SSB performs track fitting using hit information of entire 12 layers of inner detector to reduce fake tracks and improve resolution of track parameters.

Figure 56 shows a photo of the SSB. A SSB has Xilinx Kintex-7 FPGAs. Each SSB connects to a RTM through ERNI-204781 connector. The RTM is an optical fiber transceiver interface connecting to other SSBs or downstream FLICs. RTM implements a series of SFP+ modules whose lines connect to P3 area and corresponding connector on the SSB.

Each SSB receives output from 4 AUX cards and hit information which are not used at the first stage processing from the DFs through a RTM for 2 $\eta - \phi$ towers. The SSB needs to share data with other SSBs to remove duplicated tracks and identical tracks in overlap regions. The SSB merges the FTK data in a core crate and sends it to the FLIC through optical fibers fed by the RTM.

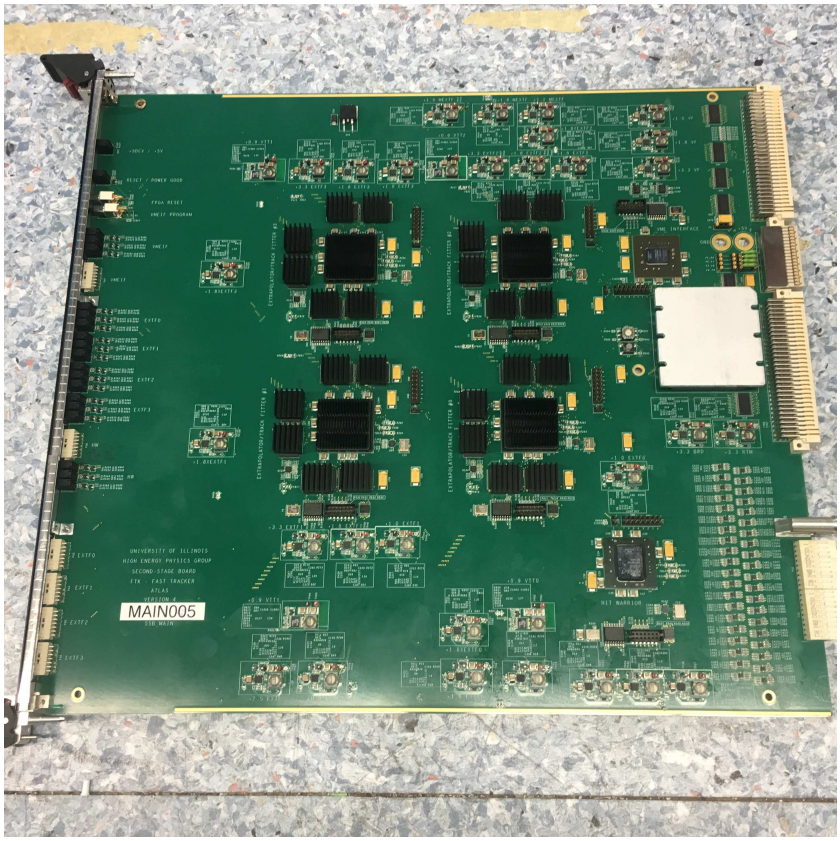

Figure 56: A photo of the SSB.

The main functions of the SSB consist of Extrapolator, Track Fitter, and Hit Warrior which are implemented in the firmware. These functions are summarized in Fig. 57.

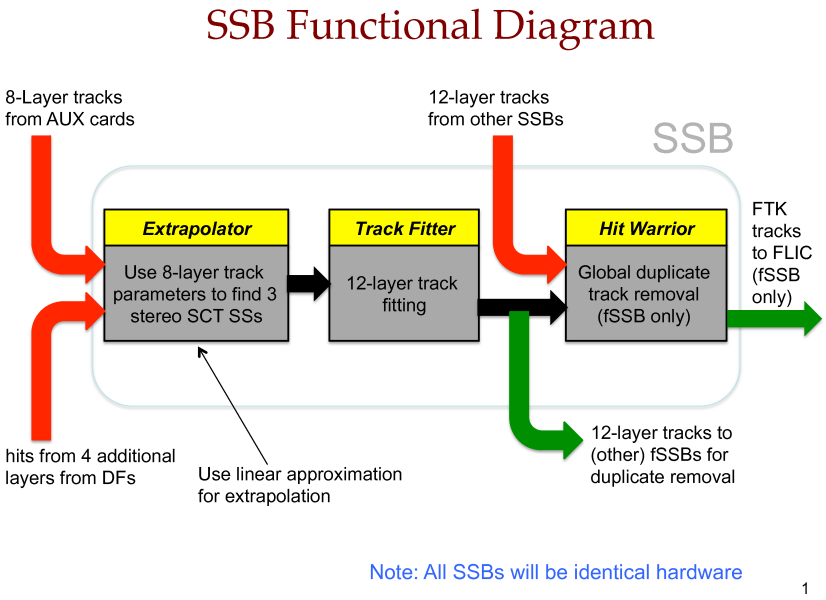

Figure 57: A functional diagram of the SSB.

Extrapolator

The SSB uses hit information of full 12 layers, while it is 8 layers at the first stage. The Extrapolator merges hit lists in the layers not used at the first stage as following steps.

1. Receive hit information from the DF that is not used at the first stage processing.
2. Look up the SS based on the hit information.
3. Make a list of hits in that SS and the 2 adjacent ones.
4. Check if there are hits in more than 10 layers. If track passes this requirement, hit information of 8 layers used at the first stage, sector ID of 12 layers, and the hit lists are sent to the Track Fitter process.

Track Fitter

Linear approximation with full 12 layer information is performed. Track parameters and χ^2 components are calculated to reduce fake tracks and to improve resolution. If a track passes χ^2 cut of pre-defined value, the track is kept for further processing. Three ways of track fitting are implemented which are same as the first stage. To ensure high efficiency, up to 2 missing hits are allowed for the Majority Fit algorithm. The Recovery Fit algorithm runs to recover tracks with a little large χ^2 value by refitting the track a number of times by dropping hits in additional 4 layers in each time. If a track with the best χ^2 passes the threshold, the track is kept. Tracks passing χ^2 requirement are sent to the HW function.

Hit Warrior

The main task of the HW function in the SSB is to remove duplicated tracks in an entire event, while the HW function in the AUX card removes duplicated tracks only in a road. Two tracks are considered as duplicated if they share more than pre-defined number of hits. The number of hits and χ^2 value are used to determine which tracks to be kept. An example of the criteria of duplicated track removal is summarized in Tab. 5.

Track 1	Track 2	Criteria
12 layers	12 layers	Remove track with the higher χ^2
12 layers	11 layers	Remove 11-layer track
11 layers	11 layers	Remove track with the higher χ^2

Table 5: The criteria for duplicated track removal.

SSB Functional Overview

It is the most efficient to concentrate the HW function on the final SSBs which sends data to the FLIC since they need to remove duplicated tracks in an entire event. Thus there are two types of the SSBs which are called primary SSB (pSSB) and final SSB (fSSB). They have identical board, and identical Extrapolator and TF functions but different HW function in the firmware. The pSSBs perform track fitting with 12 layer information, and send tracks to the $+\phi$ neighboring fSSB in the same core crate to send data to the FLIC. They also send tracks to the next neighboring fSSBs for duplicated track removal. The fSSBs receive tracks from the corresponding pSSBs, remove duplicated tracks, and send tracks to the FLIC.

As shown in Fig. 58, each SSB processes 2 $\eta - \phi$ towers (4 AUX cards) with an overlap in η, and the neighboring SSB also has overlap in η for the same ϕ region. A "core crate" in a VME rack consists of 8 $\eta - \phi$ towers, and one tower covers a quarter of rapidity range and 22.5° in ϕ, thus a core crate covers the full rapidity range and 45° in ϕ.

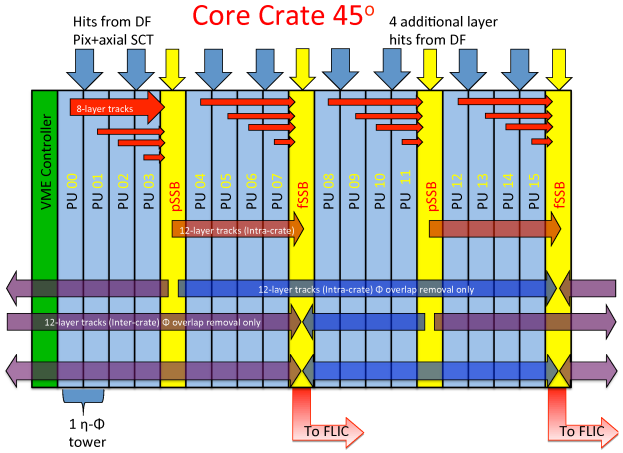

Figure 58: The SSBs in a core crate and data flow in and out of the SSBs.

3.2.7 FTK to Level-2 Interface Crate

After processing in the SSB, the data is sent to the FLIC [67]. The input data of the FLIC includes track parameters and hit information associated to each tracks. Each core crate has 2 data links to the FLIC and covers 22.5° in ϕ. The FLIC receives and combines data from 2 data links of the fSSBs, transforms data from the FTK format to the ATLAS global format, and sends them as a single data stream to the ATLAS ROS where the HLT can access. The FLIC has possibility to have additional functions such as primary vertex finding and determination of LHC beam spot.

Since it is possible that processing time for a specific event is different among core crates, data of the same event can arrive at the FLIC at different timing. Thus event data among all FLICs are synchronized based on the Level-1 ID information along with other data such as event record header sent from the fSSBs. Core crates send event information such as headers and trailers even if there is no track information in the event. This enables the HLT to notice if there is lost event.

A photo of the FLIC board is shown in Fig. 59 left and the design concept is shown in Fig. 59 right. ATCA crate is chosen to operate FLIC. The FLIC is implemented on an individual ATCA crate from the ones used for the DF boards. A 6-slot shelf with full mesh back plane is used. There are two FLIC main boards and two RTMs. A FLIC holds 4 Vertex-6 FPGAs to drive its functions. Each core crate has 2 SFP+ links to send data to front panel of the FLIC main board. The data is processed by the FLIC main board and forwarded to the RTM through Zone-3 connector of ATCA shelf. A RTM has 8 SLINK connectors to send data to the ATLAS ROS.

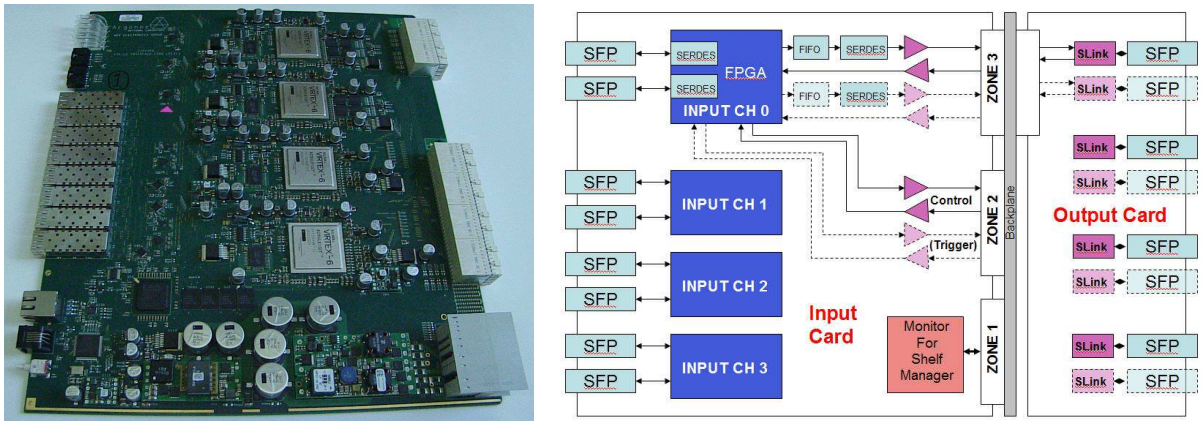

Figure 59: A photo of the FLIC (left) and a functional overview of the FLIC (right).

The ATCA technology enables to have an option to process entire event within the shelf. The required architecture for that function is dual-star back plane configuration which allows all boards in the crate can communicate with 2 dedicated slots, one is necessary for board control and monitoring, and the other is for more complex data processing and triggering. To facilitate board control and readout of monitoring information, processor blade utilizes second two star connections on Zone-2 back plane. This board communicates with external computers over standard Ethernet.

Under the expected LHC Run3 luminosity of 3.0×10^{34} cm^{-2}s^{-1}, there are typically 300 tracks with p_T >1 GeV in average. The data size for each track is expected to be 100 bytes, and additional 40 bytes are used for event header and trailer information. This leads to the average data rate of the FLIC to be 3 GB/s per event and 3 Gb/s from each core crate. The FLIC does not increase data size during its processing,

just add little in the event header and the trailer which is not so large with respect to the entire information. Thus the output data rate of the FLIC is almost same as the input data rate.

The data record sent from core crate to the FLIC is 16-bit word format and each record corresponds to an event. Each input record has multiple track information which includes track parameters of specific tracks in that event and hit information. Track frames sent from core crates include tower number and sector number as well. These data are transformed into the ATLAS detector global module number by the FLIC so that the HLT can use these information easily.

3.3 Schedule of the FTK Operation

Figure 60 shows the schedule of the FTK operation along with the LHC run plan. The FTK will start operation with half number of processor units, 64 AMBs and 64 AUXs, in 2018 at LHC Run2 to accumulate operation experience and to evaluate the performance in the real environment. During LHC shutdown planned from 2019 to 2020, the FTK will be upgraded to full system. And the FTK will start operation with full system at LHC Run3 which will start from 2021. After the end of LHC Run3, next generation system which will operate at High Luminosity LHC (HL-LHC) will be developed.

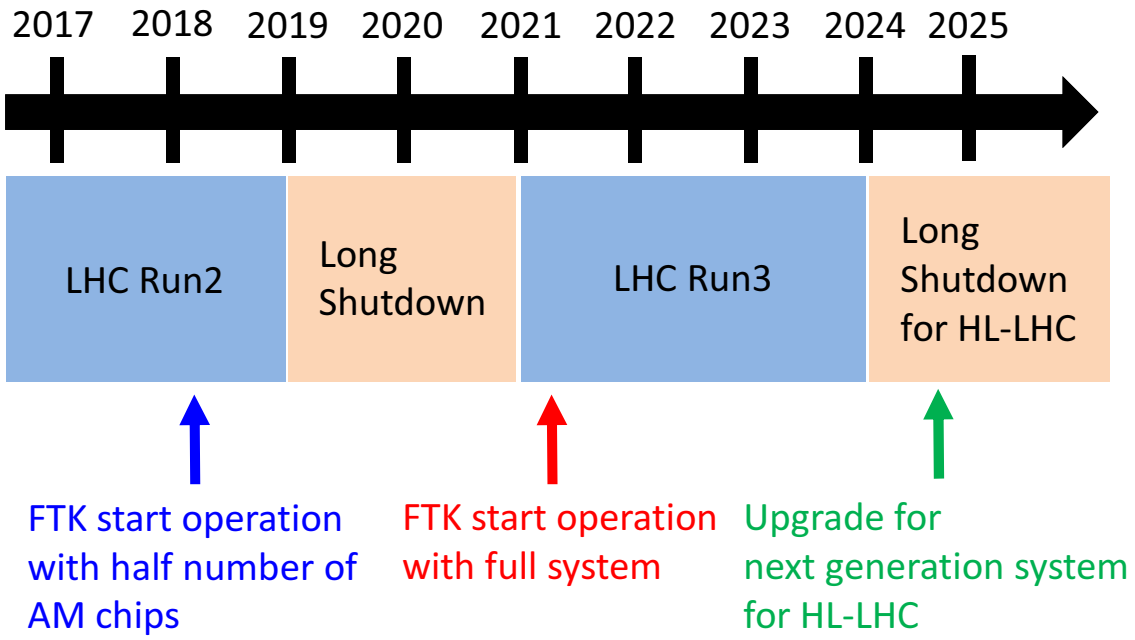

Figure 60: Schedule of the FTK operation along with the LHC run plan.

4 Performance Validation by Simulation

Software simulation of the FTK (FTKSim) [68] has been developed to validate and improve performance of the FTK. Although it is challenging to emulate massively parallel hardware system by software simulation, such simulation is important to optimize design and maximize performance of entire system since it emulates each stage of the FTK logic. It has to be integrated and working under the ATLAS software framework.

The FTKSim covers wide range of fields not only evaluation of the quality of reconstructed tracks but also emulation of the hardware, such as processing time evaluation and bit level validation. The FTKSim covers items shown in the following:

- Validation and optimization of track reconstruction quality.
- Generation of the "configuration files" (e.g. sectors, patterns, fit constants) which are used for hardware operation.
- Sample generation to study physics impact with the FTK track information.
- Validation and improvement of the performance for physics such as particle identification.
- Processing time evaluation.
- Bit level validation of hardware.
- Trigger chain building with the FTK track information.

The simulation plays a key role to achieve high performance and smooth operation of the FTK. The detail is described in the following subsections.

4.1 Sector/constant and Pattern Bank Generation

Fit constants are generated associated with each sector. They are critical for track reconstruction performance of the FTK. There are several processes to generate patterns and fit constants. An overview of the generation process is depicted in Fig. 61.

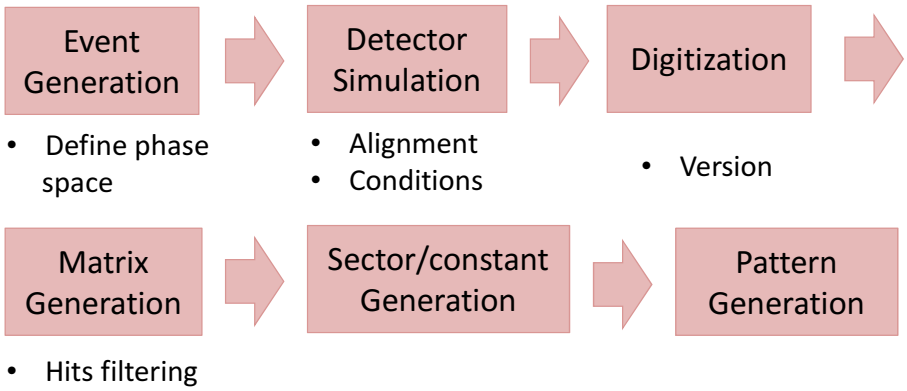

Figure 61: Sector, fit constant and pattern generation process.

4.1.1 Event Generation to Sector Generation

At first, 1.5 billion single muon tracks are generated. The number is estimated to be enough to ensure high track reconstruction efficiency (> 90%). Muons of both negative and positive charge are generated with flat distribution with the following parameter space.

- $-2.2 \leq d_0 \leq 2.2[mm]$
- $-120 \leq z_0 \leq 120[mm]$
- $-0.8 \leq 1/p_T \leq 0.8[GeV^{-1}]$
- $-3.0 \leq \eta \leq 3.0$
- $-\pi \leq \phi \leq \pi$

Training samples are generated by the ATLAS software framework. The samples are passed to the ATLAS detector simulation which is based on Geant4 [69] and then they are digitized. In this process, there are several important parameters such as detector status, condition and geometry which should be set compatible with realistic situation to achieve high track reconstruction performance in real operation.

Hits used for sectors and fit constants generation are filtered to select clean tracks. If there are more than one hit in the same module in an event, the tracks are rejected. Hits from secondary particles or noise are removed. Tracks which do not leave hits in all detector layers are not used.

After processing all tracks, complete list of sectors for each $\eta - \phi$ tower is generated and sorted by the frequency of tracks in each sector in descending order. The frequency of sector is important information and often called as coverage.

4.1.2 Fit Constant Generation

Fit constants for linear approximation are calculated. A set of fit constants are associated to each sector. Linear fit procedure expresses track parameter ($\tilde{p}_i$) as

$$\tilde{p}_i = \sum_{l=1}^{N} C_{il} x_l + q_i, \tag{48}$$

where C_{il} and q_i are constants defined in a sector, x_l are N hit coordinates, and i runs for 5 track parameters ($\mathrm{p_T}, \eta, \phi, \mathrm{d_0}, \mathrm{z_0}$). The coordinate is local cluster centroid position in each module. There are two coordinates for a pixel hit and one for a SCT hit. To evaluate the fit constants, parameters that minimize the difference between the linear parameter ($\tilde{p}_i$) and true parameter (p_i) value is defined.

$$min\big(\big\langle(\tilde{p}_i - p_i)^2\big\rangle\big), \ \forall i = 1, ..., 5, \tag{49}$$

where i runs over 5 track parameters. The solution of Eq. 49 for ith parameter can be expressed as

$$\sum_{m=1}^{N} \left(\langle x_l x_m \rangle - \langle x_l \rangle \langle x_m \rangle \right) C_{mi} - \langle x_m p_i \rangle + \langle x_m \rangle \langle p_i \rangle = 0, \ \forall l = 1, ..., N, \tag{50}$$

$$\sum_{l} C_{il} \langle x_l \rangle + q_i = \langle p_i \rangle, \tag{51}$$

where the averages are over training muons. In Eq. 50, covariance matrix can be defined between the coordinates,

$$V_{lm} = \langle x_l x_m \rangle - \langle x_l \rangle \langle x_m \rangle. \tag{52}$$

The solution of Eq. 50 can be obtained using the inverse of the covariance matrix V^{-1},

$$C_{il} = \sum_{m} V_{lm}^{-1} \langle x_m p_i \rangle - \langle x_m \rangle \langle p_i \rangle. \tag{53}$$

The constants q_i can be obtained by inserting C_{il} from Eq. 53 to Eq. 51. From the covariance matrix, a track quality parameter in the limit of validity of the linear approximation can be calculated as,

$$\chi^2 = \sum_{i,j=0}^{N} \left(x_i - \langle x_i \rangle \right) V_{ij}^{-1} \left(x_j - \langle x_j \rangle \right). \tag{54}$$

Decomposing Eq. 52 into its eigenvalues and eigenvectors produces 5 large eigenvalues that do not contribute to χ^2, while the others are very small and make real contribution. Eq. 54 can be formed into

$$\hat{\chi}^2 = \sum_{i=1}^{N-5} \left(\sum_{j=1}^{N} A_{ij} x_j + k_i \right), \tag{55}$$

where $A_{ij} = u_{ij}/\sqrt{e_i}$ with u_{ij} is the component j of the eigenvector i, and e_i is its eigenvalue, and $k_i = \sum_k A_{ik} \langle x_k \rangle$. The calculations are repeated in each sector. For sectors with low coverage, since the covariance matrix cannot be inverted, the sectors are rejected.

4.1.3 Pattern Generation

Pattern generation by simple training requires a huge amount of single muon samples and then take long time to achieve enough coverage. Instead, the patterns are generated by the method that inverts the way of linear approximation of track fitting, called "Patterns from Constants".

By using randomly distributing 5 track parameters and $N-5$ Gaussian distributed constrains, equations below are calculated,

$$\sum_j C_{lj} x_j + q_l = \hat{p}_l, \ \forall 1, ..., 5, \tag{56}$$

$$\sum_j A_{mj} x_j + k_m = \hat{\chi}_m, \ \forall 1, ..., N-5, \tag{57}$$

where C_{lj}, A_{mj}, q_l and k_m are the same as Eq. 48 and Eq. 55. The coordinates x_j can be extracted and transformed into the SSs. The patterns are sorted as descending order by track frequency as are done for sectors.

Here the patterns are generated with the highest resolution SSs which all patterns cannot be stored in the AM chips. The bank size needs to be reduced while keeping performance as much as possible. It is achieved by applying variable resolution, DC-bit feature described in Sec. 3.1.2. In the real use case, the first step is to cluster the SSs of the highest resolution into the SSs with the size of real use in the AM chips. The size is defined as $\Delta SS_{AM} = \Delta SS_{high-res} \times 2^{N_{DC}}$, where N_{DC} is the number of the DC-bits to be applied. Then if the wider SS has tracks in the highest resolution in both sides, the wider SSs are used, and if only in either sides, the highest resolution SSs are used.

4.1.4 Improvement of Processing Time for Generation

Patterns, fit constants and sectors are necessary to reflect real-time conditions of the LHC beam and the detectors to keep high track reconstruction performance. Thus the generation process must be as fast as possible to cope with the change of these conditions. There has been several optimization to make the generation process faster.

One of the optimizations is better parallelization of sector and fit constant generation. The ATLAS grid system is used to generate them, which utilizes CPU resources distributed all over the world and enables to instantiate a number of parallel jobs among the CPUs. At first, sectors and fit constants were generated in parallel with the FTK $\eta - \phi$ towers. The number of parallelizations was limited by the number of towers in this way. And if even one of the jobs fails in the middle of processing, it should be restarted and other jobs need to wait for the failed job to finish, which makes whole process much longer.

Instead of instantiating parallel jobs based on towers, parallelization is done with the number of events. In this way, the number of parallelization can be optimized easily. In addition, if generation starts with a plenty of extra number of muons to achieve high performance, failure of several jobs does not have significant impact on entire performance and can be neglected. In Eq. 50 and 51, fit constants can be obtained with the average over training muons, thus it can be merged easily into towers by taking the average of constants over all the parallel jobs. Consequently, it took roughly one month to generate

patterns when parallelization was done with towers, while it was improved to be about 10 days with the new way.

4.2 Validation of Track Reconstruction Performance

It is important to evaluate performance of the generated patterns and the fit constants. System-wise track reconstruction performance can be checked only after generation of the patterns and the fit constants and should be validated to have enough performance. In addition, it becomes important feedback for improvement at the next round of generation.

Track reconstruction performance of the FTK is validated with both MC sample and real data. Tracks are required to pass following selections.

- Number of Pixel detector plus SCT hits on a track plus the number of inactive modules crossed by the track must be at least ten.
- There must be no pixel holes on the track.

In addition to the cuts shown above, following phase space cuts are applied to tracks to ensure that tracks are within FTK training phase space.

- $|d_0| < 2[mm]$
- $|z_0| < 110[mm]$
- $p_T > 1[GeV]$

In addition, offline tracks are required to pass following "loose selection" of track quality.

- Tracks are required to be $|\eta| < 2.5$.
- The number of hits in the Pixel detector and the SCT is equal to or more than 7.
- The number of hits shared with multiple tracks are no more than 1.
- The number of hole layer in the Pixel detector and the SCT that the track passes is no more than 2.
- The number of hole layer in the Pixel detector that the track passes is no more than 1.

Track reconstruction efficiency with respect to truth tracks is defined as the ratio of the number of FTK tracks matched with truth tracks to the number of truth tracks. The matching with truth tracks are based on the fraction of the number of hits on the track which are produced by the track. Tracks with at least 50% of hits shared with truth track are considered as match.

Offline tracks are required to pass the same selections of the number of hits and phase space as FTK tracks and loose selection of track quality. The efficiency with respect to offline tracks is defined as the ratio of the number of FTK tracks matched with offline tracks to the number of offline tracks. If there is at least one FTK track around the offline track in ΔR < 0.05, it is considered as matched, where dR is defined as

$$dR = \sqrt{(\eta_{FTK} - \eta_{Off})^2 + (\phi_{FTK} - \phi_{Off})^2}. \tag{58}$$

4.2.1 Track Reconstruction Performance with Full System

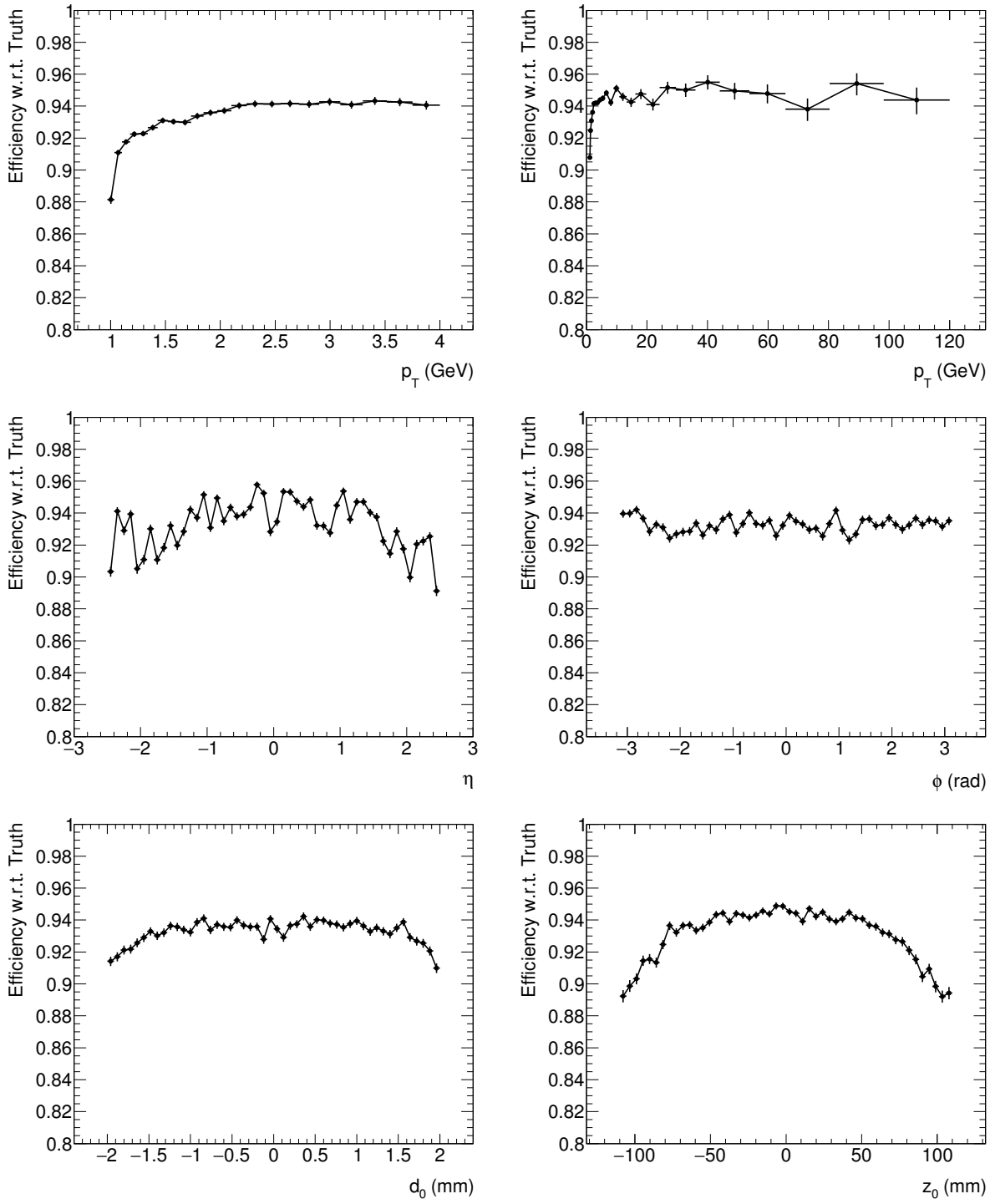

Figure 62: Track reconstruction efficiency with respect to truth tracks for single muon events.

Figure 62 shows the track reconstruction efficiency with respect to truth tracks using single muon MC sample which is generated with the kinematic distribution described in Sec. 4.1.1. The number of patterns

and fit constants corresponds to the FTK full system, where high efficiency (> 90%) can be achieved. The inefficiency mainly comes from geometric coverage of the pattern banks and detector inactive modules. Inefficiency of low p_T tracks around 1 GeV is considered that it is around the boundary of FTK fiducial.

4.2.2 Optimization for Half System

The FTK will run with a half number of patterns compared with full system at early stage of operation as described in Sec. 3.3. The performance with this condition needs to be investigated and validated.

Figure 63 shows the track reconstruction efficiency with respect to truth tracks for single muon events. Here the number of used patterns are half compared with the full system . The black plot shows the one without any optimization so that the efficiency is dropped especially in the transition region ($|\eta| = 1.1$). Since module alignment is complicated in this region, more patterns are necessary to keep high efficiency. The red plot shows the efficiency which the distribution of the patterns is optimized by generating more patterns in the transition region with compromise of fewer patterns in the barrel region. It is found that the efficiency in the transition region is recovered with a little loss of efficiency in the barrel region. Thus overall efficiency can be kept high (90%) with the half number of the Processor Units if the number of pile-up interactions is not so high.

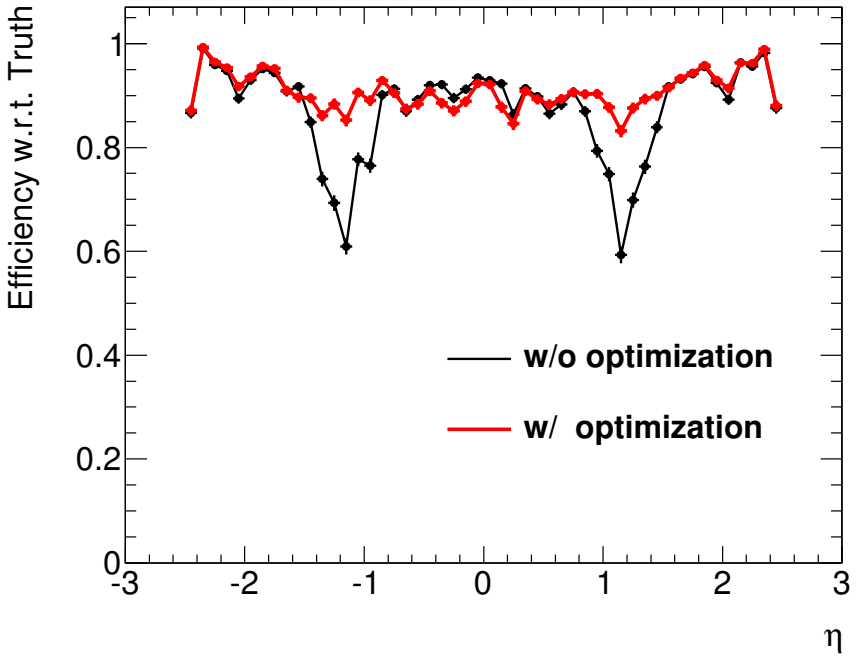

Figure 63: Track reconstruction efficiency with respect to truth tracks for single muon events with (red) and without (black) optimizing the distribution of the patterns. The half number of patterns compared with the full system is assumed.

4.2.3 Performance Evaluation for Several Physics Processes

Track reconstruction performance for physics process dependence is evaluated. Here after, the bank which has the half number of patterns is used unless specified otherwise. Figure 64, 65 show the track reconstruction efficiency for $Z \rightarrow \mu\mu$ events with 40 pile-up interactions, which corresponds to about 1.5×10^{34} cm^{-2}s^{-1}, with respect to truth tracks and offline tracks, respectively. All the tracks which pass the requirements described in Sec. 4.1.1 are considered. The efficiency with respect to truth tracks are less and fluctuated compared with the case of single muon events, which is due to interactions with materials in the detector by pile-up jets and many secondary particles are generated. The FTK tracks can reconstruct more than 85% of the offline tracks since the effect is similar between FTK tracks and offline tracks.

Figure 66, 67 show the track reconstruction efficiency with respect to offline tracks for the MC events which fire minimum bias trigger and $t\bar{t}$ events with 40 pile-up interactions, respectively. Track reconstruction efficiency is high for each process. Thus it is validated that track reconstruction performance of the FTK does not depend on the physics processes.

Resolution with respect to truth tracks is also evaluated. One option of FTK track usage at the HLT is to refit them at the HLT using FTK tracks as a seed to improve resolution, which is called "refitted FTK tracks". The resolution of offline tracks, FTK tracks and refitted FTK tracks with respect to truth tracks as a function of curvature are shown in Fig. 68. The resolution is defined as the width of a Gaussian fit in a range of 3 times the RMS of the distribution. At large absolute values of curvature where the resolution is affected by multiple scattering, the resolution gets worse. Tracks of this quality can be reconstructed by the FTK, and better resolution can be achieved by refitting FTK tracks at the HLT.

4.2.4 Setup Compatibility against Operating Environment

In pattern generation process, it is important to set all configuration parameters such as detector status, conditions, geometry, and digitization version to be compatible with the environment where the FTK works. The FTK can reconstruct tracks with high resolution due to its high granularity at track fitting. In other words, if there are incompatibility of configuration parameters in the generation process, track reconstruction performance of the FTK is affected and in some cases significantly degraded.

Figure 69 shows track reconstruction efficiency with respect to offline tracks where the detector geometry and conditions are quite different between pattern generation and applied data. Significant inefficiency comes from the incompatibility of the detector geometry, the detector conditions and the version of digitization. Although this is an extreme case, it can be convinced that it is critical to keep track of and cope with the change of the operating environment since the parameters and environment of the LHC and the ATLAS such as the detector conditions or the detector geometry change frequently.

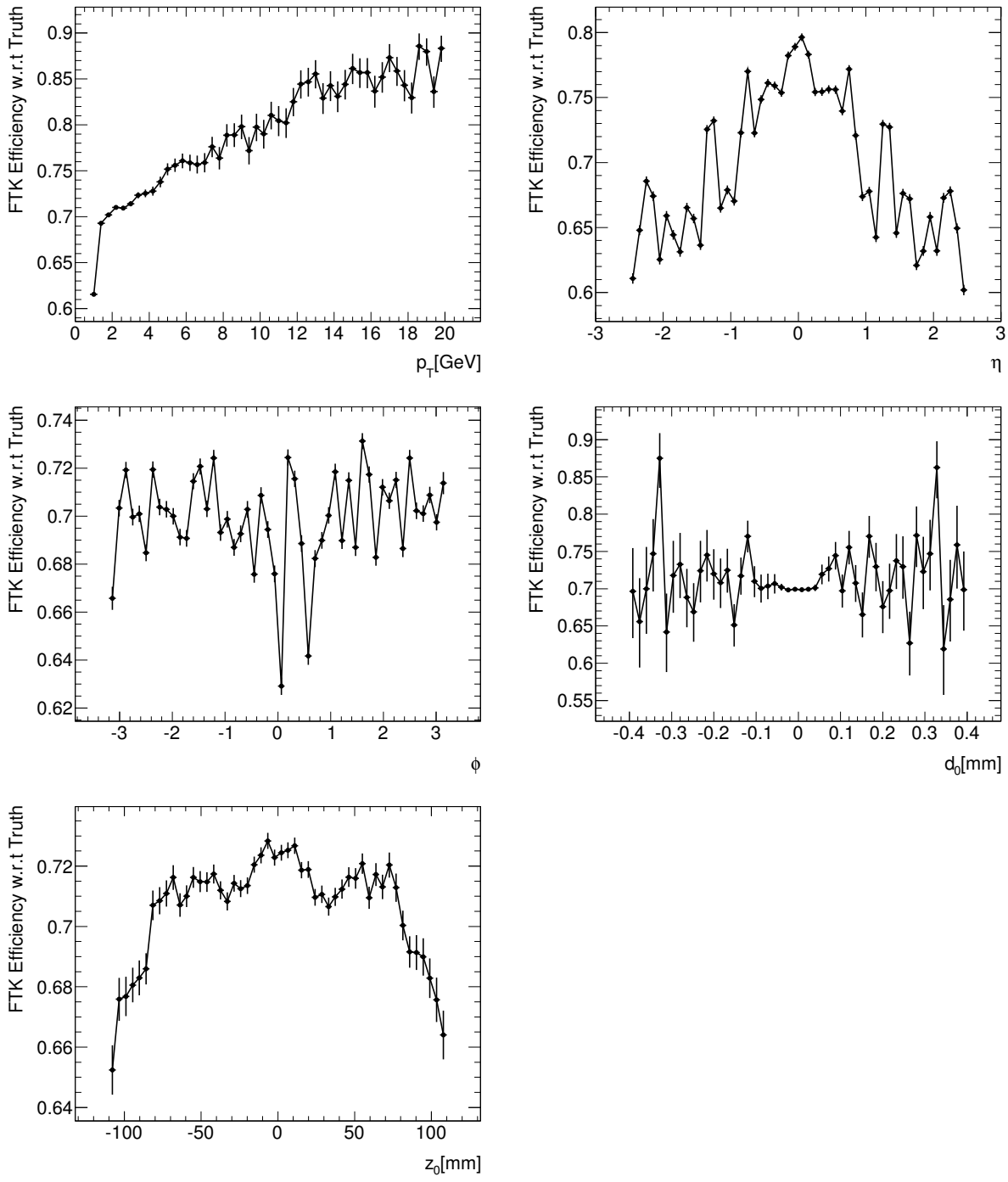

Figure 64: Track reconstruction efficiency with respect to truth tracks for $Z \rightarrow \mu\mu$ events with 40 pile-up interactions.

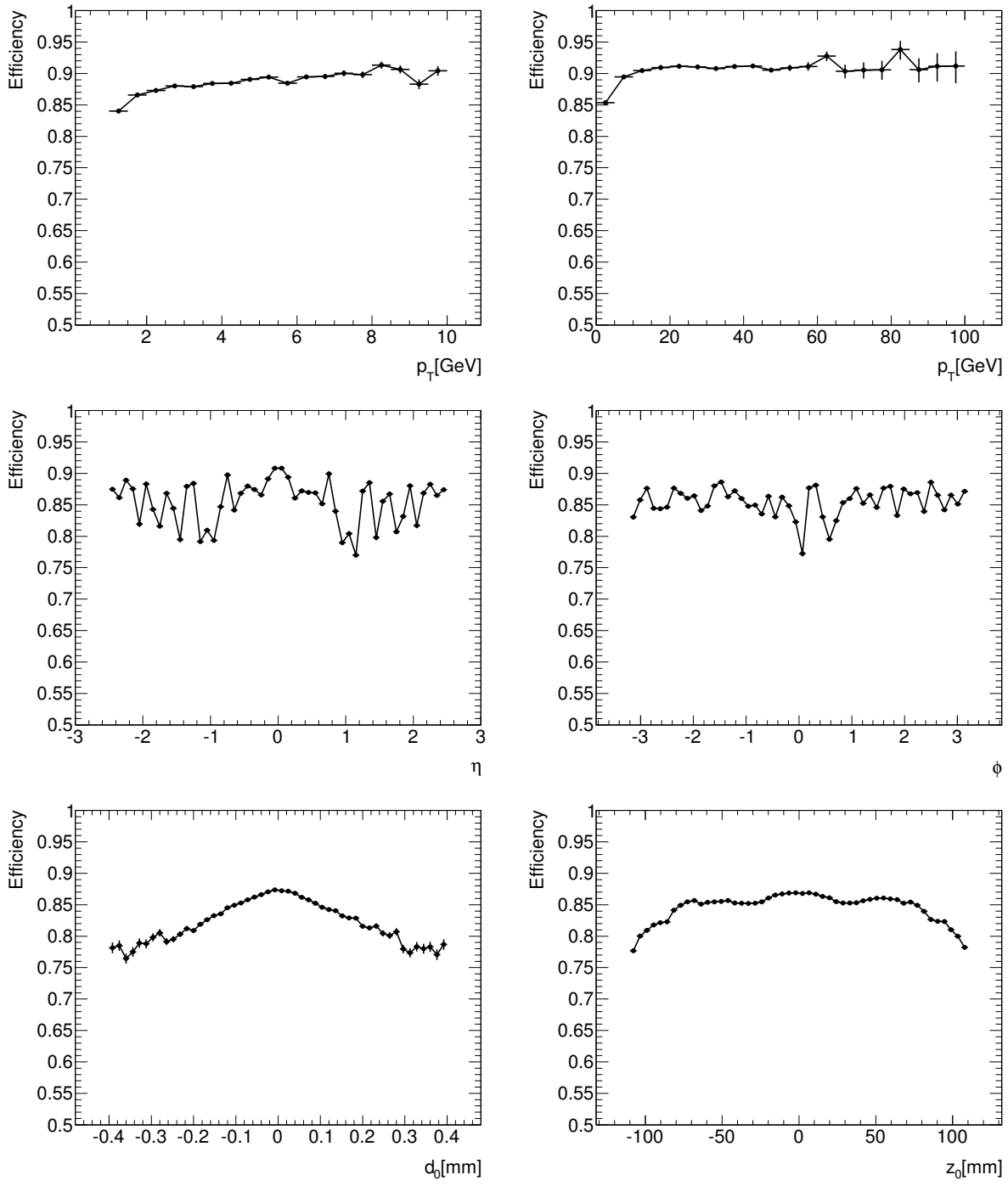

Figure 65: Track reconstruction efficiency with respect to offline tracks for $Z \rightarrow \mu\mu$ events with 40 pile-up interactions.

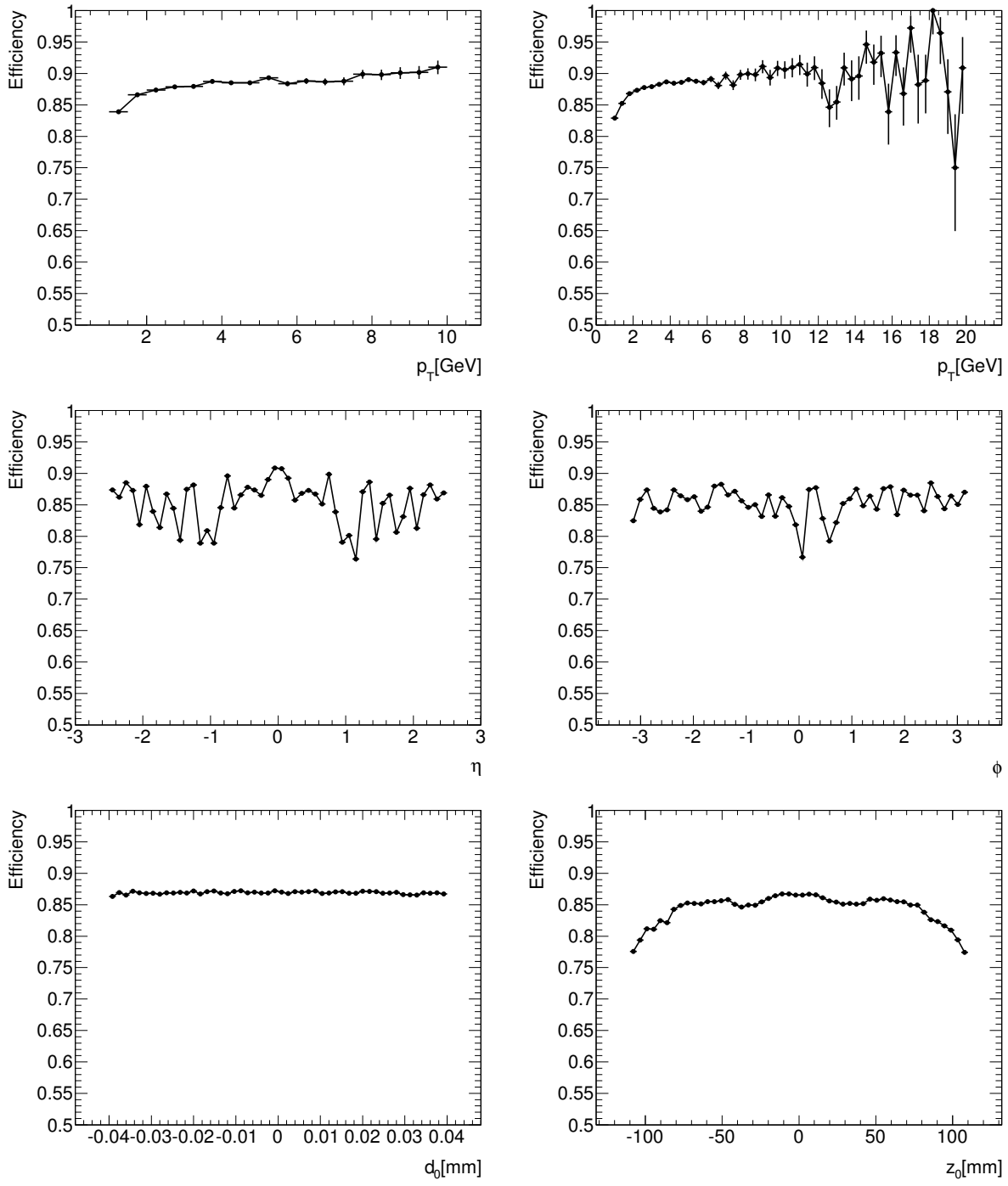

Figure 66: Track reconstruction efficiency with respect to offline tracks for events which fired minimum bias trigger with 40 pile-up interactions.

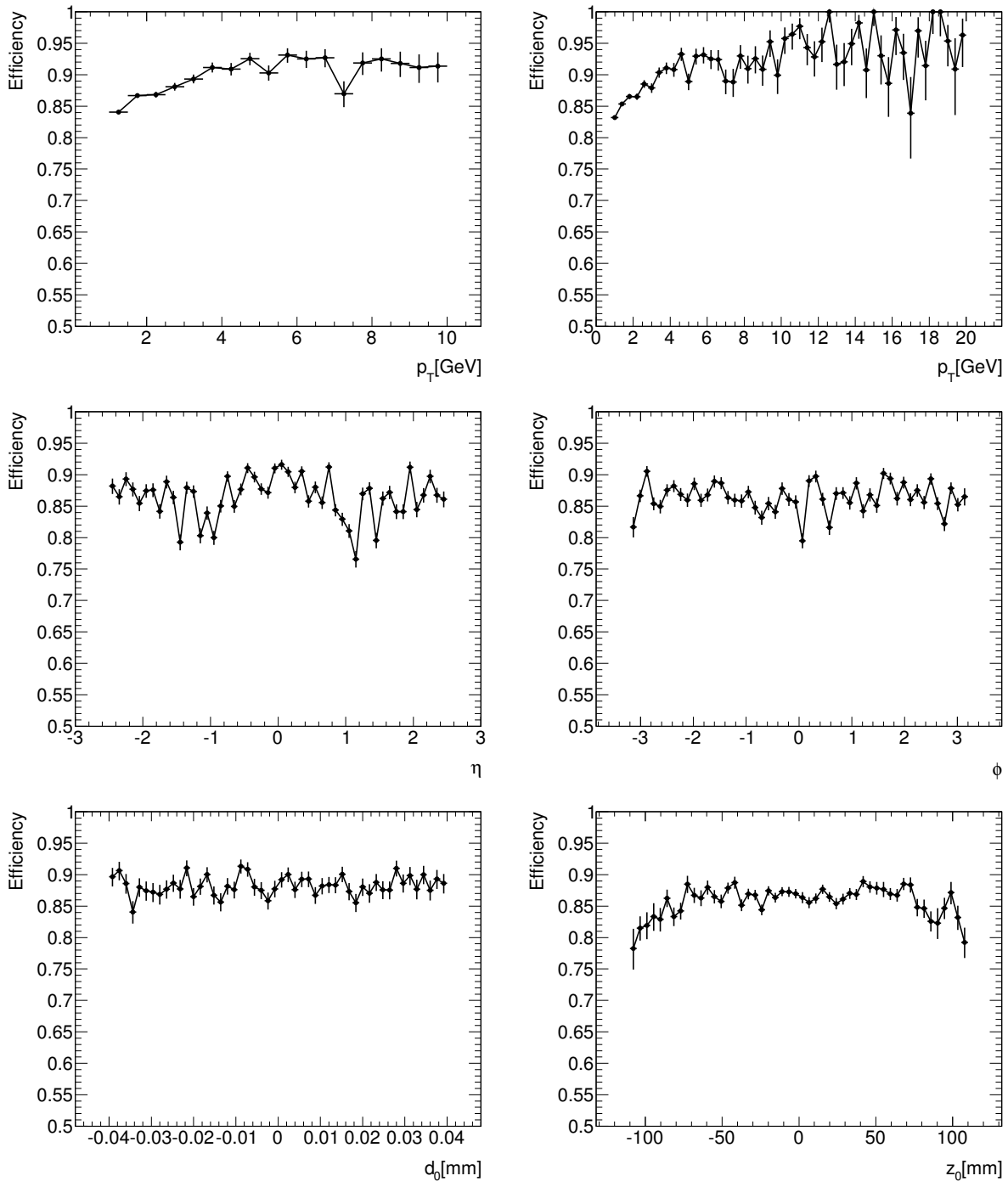

Figure 67: Track reconstruction efficiency with respect to offline tracks for the $t\bar{t}$ event with 40 pile-up interactions.

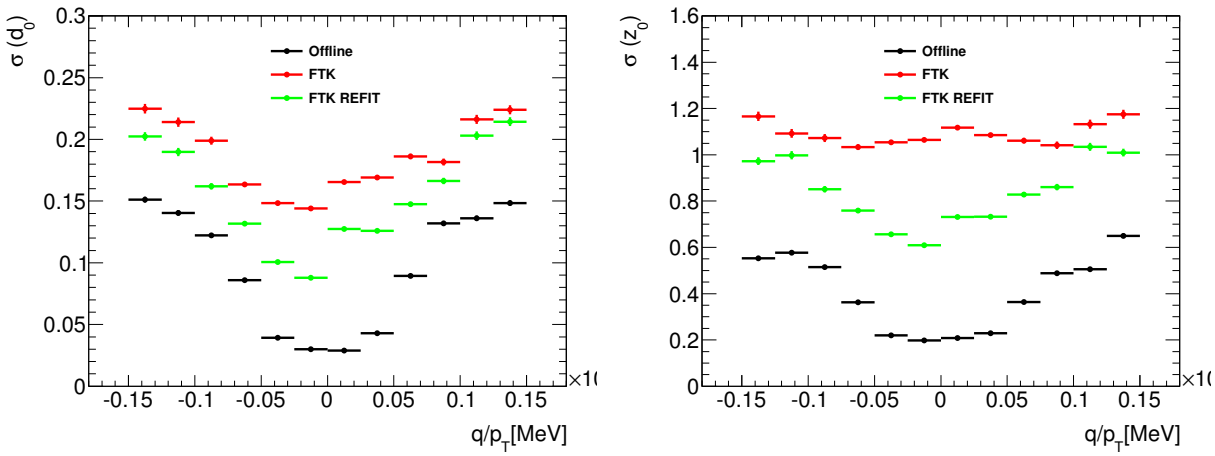

Figure 68: Resolution with respect to truth tracks as a function of curvature. $Z \rightarrow \mu\mu$ events with 40 pile-up interactions are used.

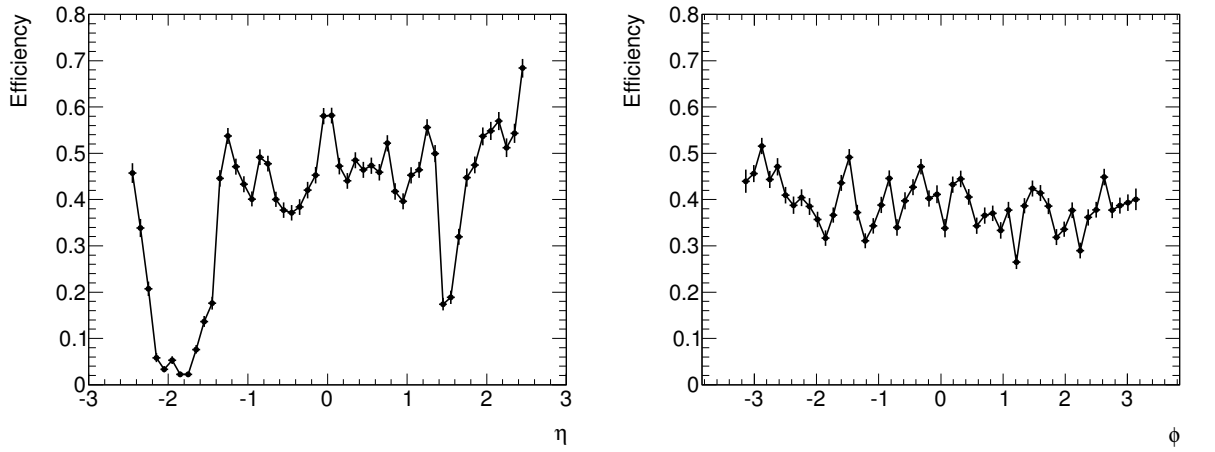

Figure 69: Track reconstruction efficiency with respect to offline tracks for events which fired minimum bias trigger with 40 pile-up interactions. The configuration setup is quite different between pattern generation and applied data.

Optimization for the Change of the LHC Beam Spot

Beam spot of the LHC changes gradually, sometimes drastically after machine upgrade. The change of beam spot affects track reconstruction performance of the FTK since FTK sectors, constants and patterns are trained with samples assuming a specific beam spot.

Figure 70 left shows a schematic view of the ATLAS inner detector layers of $x - y$ plane. If the FTK patterns are trained using the sample of the red beam spot and the beam spot of real operation environment shifts to the blue point, the FTK patterns cannot cover some regions and efficiency gets worse. Figure 70 right shows how much the track reconstruction efficiency is affected by the shift of the beam spot. It is found that the efficiency can be kept when the beam spot shift is less than 0.4 mm, and gets worse when the beam spot moves more than 0.4 mm. It also shows that the efficiency loss comes mainly from the

pattern matching stage. Thus, patterns, sectors and constants need to be replaced when the beam spot moves more than 0.4 mm. Practically, they will be prepared in advance with 0.4 mm grid, and appropriate patterns/sectors/constants can be used at the next physics run since they can be loaded run by run.

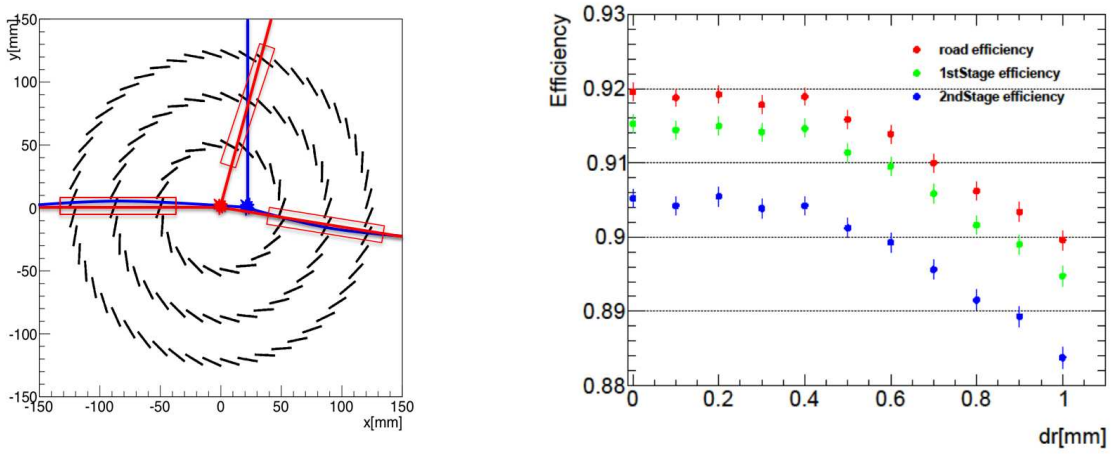

Figure 70: (left) A schematic view of the FTK patterns and tracks from 2 different beam spots. (right) Track reconstruction efficiency as a function of the beam spot shift. The red points show the efficiency of pattern matching, the green ones are the efficiency after the track fitting of first stage, and blue ones are the efficiency after the second stage.

4.2.5 Track Reconstruction Performance for Real Data

Track reconstruction performance should be evaluated and validated using real data. It is important to prepare patterns which have enough performance at the real operation. The efficiency for real data is validated by comparing FTK tracks with offline tracks. The matching criteria is the same as mentioned previously that requires $\Delta R < 0.05$.

Figure 71 shows the track reconstruction efficiency for real data sample applied as an input. The FTK patterns are trained with the configuration compatible with the input data. Here the data has average 40 interactions per bunch crossing, and following requirements are applied at the offline level. The vertex with the highest Σp_T^2 of the associated tracks has at least 3 tracks and $|z| < 150$ mm. Then, either or both of the filters shown below are required.

- Two muons have $p_T > 25$ GeV.
- The invariant di-muon mass must be greater than 55 GeV.

It is validated that high track reconstruction efficiency can be achieved for the real data. The performance is expected to improve with using the full system.

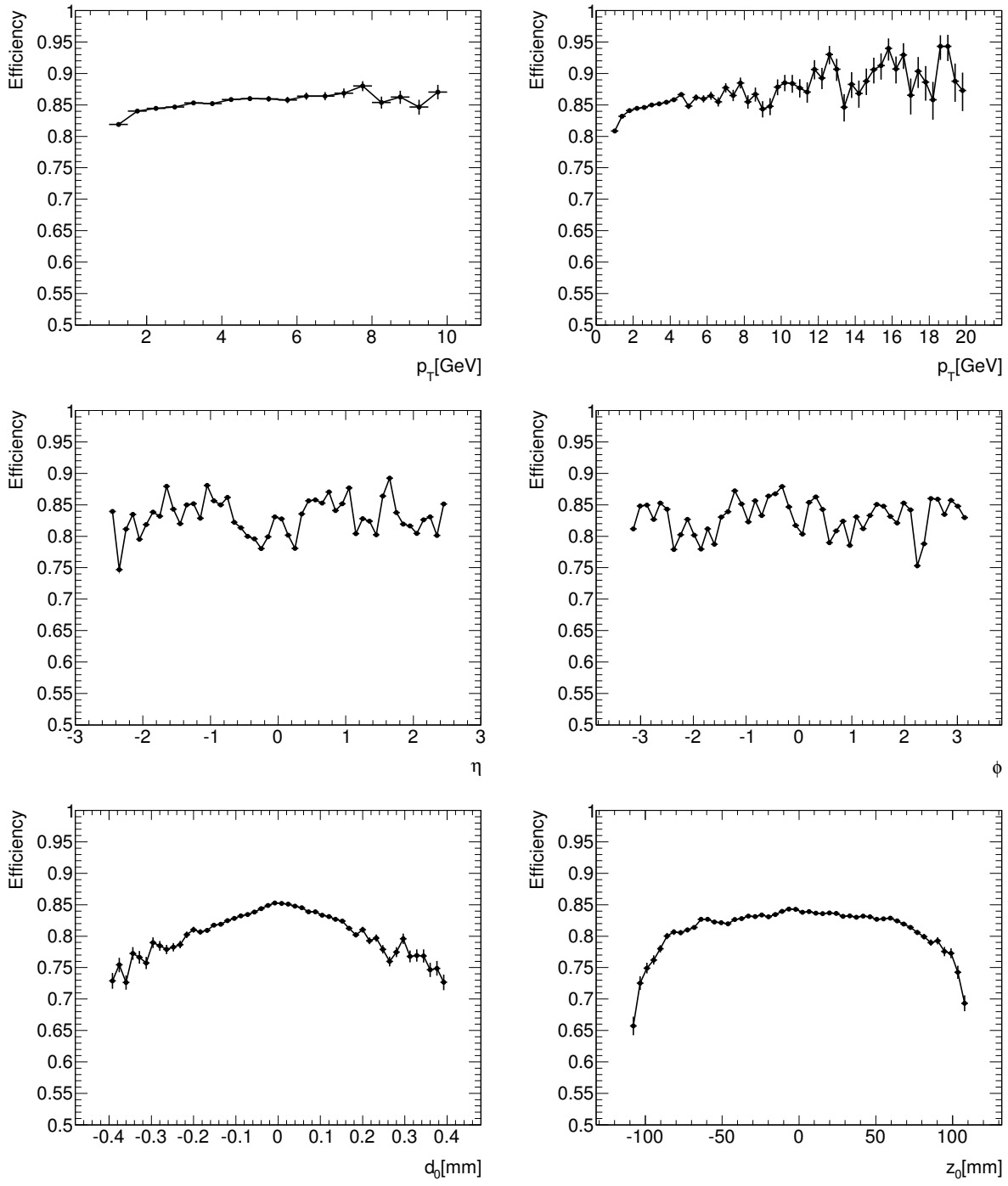

Figure 71: Track reconstruction efficiency with respect to offline tracks for events which passes requirements. Real data is used as an input.

4.2.6 Reconstruction of Physics Process

The performance to reconstruct physics process by FTK tracks is simulated. In advance, FTK track information is integrated into the ATLAS standard sample generation process which is used to generate ATLAS common samples. Here FTK tracks and offline tracks are required to pass the selections of the number of hits and phase space described in Sec. 4.1.1.

The process that a Z boson is produced and decaying into a pair of muons with 40 pile-up interactions is used. To collect $Z \rightarrow \mu\mu$ events and to reject QCD event, simple selections which can be easily implemented in trigger are required for the two leading tracks in an event. Both tracks are required to have $p_T > 22$ GeV, and the two leading tracks have opposite charge. To require the tracks decaying to back-to-back, $\Delta\phi > 2.7$ is applied. With these requirements, the rate of QCD event is significantly reduced. Here requirement for muon identification as well as matching with muon object are not applied.

The invariant mass of the two leading tracks for events which pass the requirements for offline tracks and refitted FTK tracks is shown in Fig. 72. The mass peak of Z boson is reconstructed by refitted FTK track information with this resolution. It is validated that physics process can be reconstructed by FTK track information at the HLT with only simple requirements are applied. If applying further requirements, the quality can be further improved.

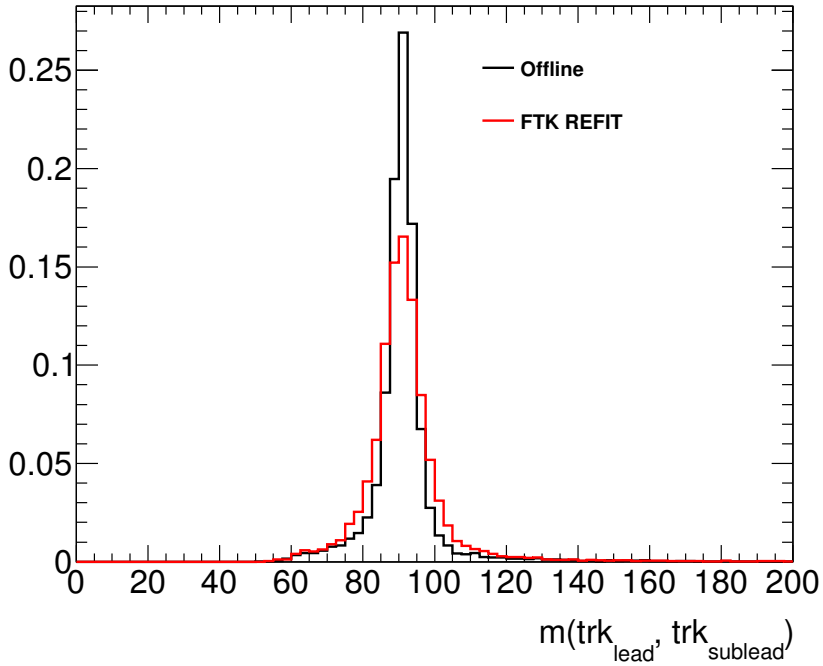

Figure 72: The invariant mass of the two leading tracks for Z $\rightarrow \mu\mu$ candidates.

4.3 FTK Processing Time Evaluation

Since the FTK will operate in the trigger stream, it is required to work fast enough in time for the trigger decision. The processing time of the FTK is one of the key ingredient to operate the FTK in the trigger system. The simulation to evaluate the FTK processing time has been developed to tune the entire system

and ensure that the FTK can handle 100 kHz level-1 trigger rate. The simulation is divided into the functional blocks, DF, DO write mode, AM, DO read mode, TF, HW, and SSB. For each functional block, the time of the first and last words into and out of the block are calculated. The FTK event execution time ends when the last word exits the busiest crate for that event since each core crate operates independently. The processing time for each block is estimated from the architecture and clock cycle. The property of each functional block used for the processing time evaluation is shown in Tab. 6.

Functional Block	Operational Clock	Latency
DF	-	1000 ns
DO write	200 MHz	40 ns
AM	200 MHz	200 ns
DO read	200 MHz	500 ns
TF	1 GHz	300 ns
SSB	1 GHz	300 ns

Table 6: The properties of the functional blocks used for the evaluation of FTK processing time.

Figure 73 shows two examples of the processing time for an event for each functional block as well as entire FTK system. The used event is $t\bar{t}$ with 69 pile-up interactions which correspond to 3.0×10^{34} cm^{-2}s^{-1}. In case of the left figure, processing time is almost same among the parallel engines, while in case of the right figure, one of the core crate takes much more time in TF function than others so that entire system takes long time. Thus event by event validation can be performed.

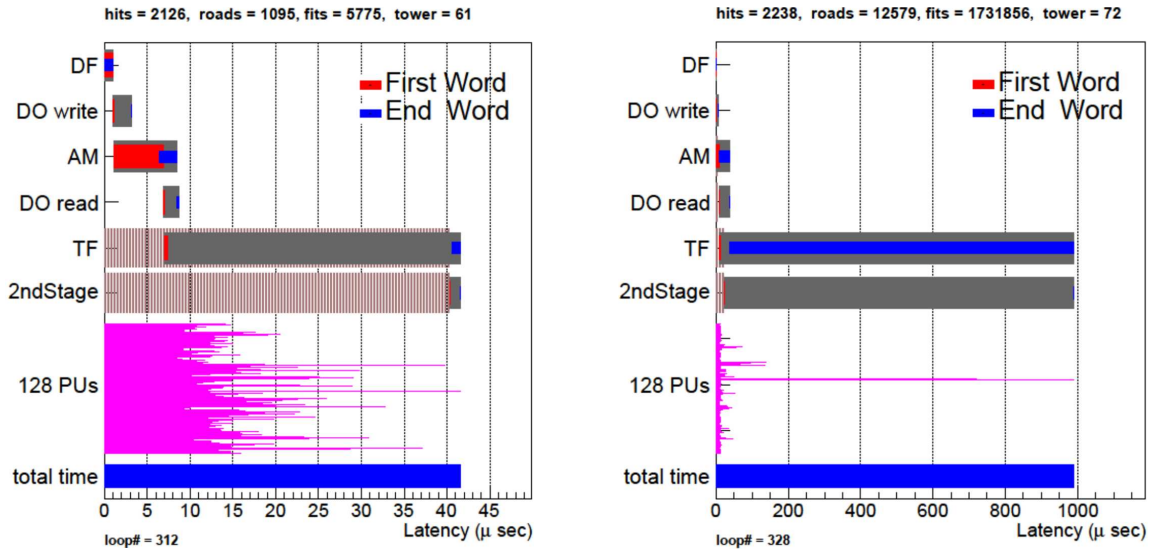

Figure 73: Processing time for an event for each functional block as well as the entire FTK system. The used event is $t\bar{t}$ with 69 pile-up interactions.

To evaluate the processing time for consecutive events, the processing time of consecutive 500 events of the same sample is emulated assuming an input event rate of 100 kHz. Figure 74 shows the total processing time for $t\bar{t}$ events with 69 pile-up interactions of 500 events. In the left figure, though this is extreme case that very heavy event such as $t\bar{t}$ with 69 pile-up interactions comes consecutively, the FTK works in time but a little getting stack. One of the key element related to FTK processing time is the number of fits per road. Since the number of fits per road is determined by the combination of hits in a

road, it increases as a power of the number of hits. Thus the processing time becomes much longer in high luminosity environment. To reduce processing time, it is effective to limit the maximum number of fits per road. Figure 74 left sets the maximum number of fits with 1024 which is currently defined as default, while Fig. 74 right sets it with 16. The processing time is significantly reduced.

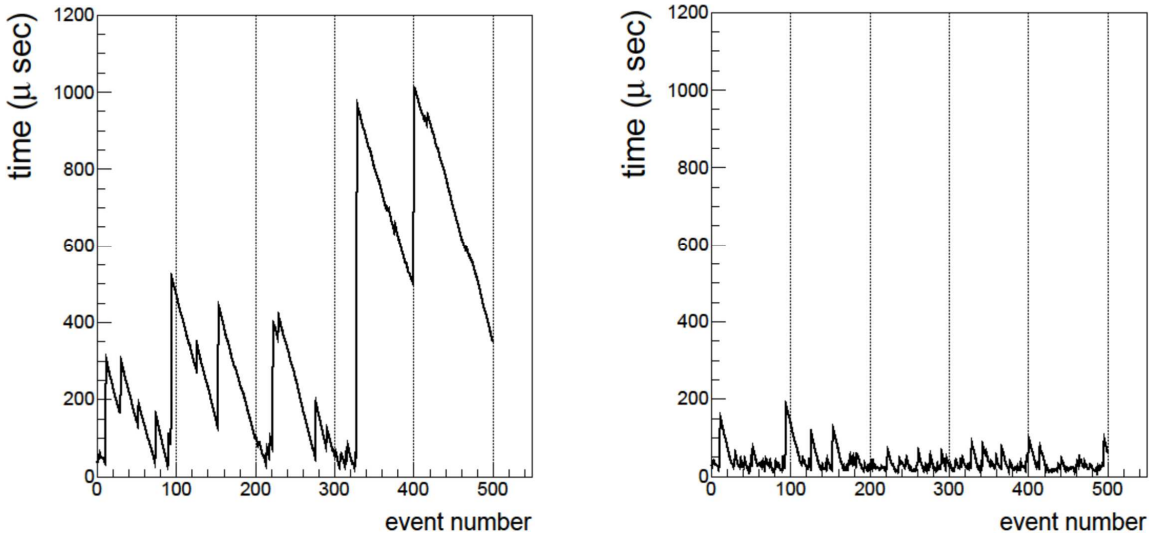

Figure 74: Total processing time for 500 events of $t\bar{t}$ with 69 pile-up interactions. The maximum number of fits per road is 1024 (left) and 16 (right).

The track reconstruction efficiency and resolution with respect to truth tracks for the case of maximum number of fits per road is 1024 and 16 are shown in Fig. 75 and Fig. 76, respectively. As described in Section. 4.2.3, the overall efficiency is not so high. The difference of the efficiency between the case of maximum number of fits per road is 1024 and 16 is less than 1%, and the difference of the resolution is less than 10%. Thus it is validated that the FTK can have options to optimize processing speed.

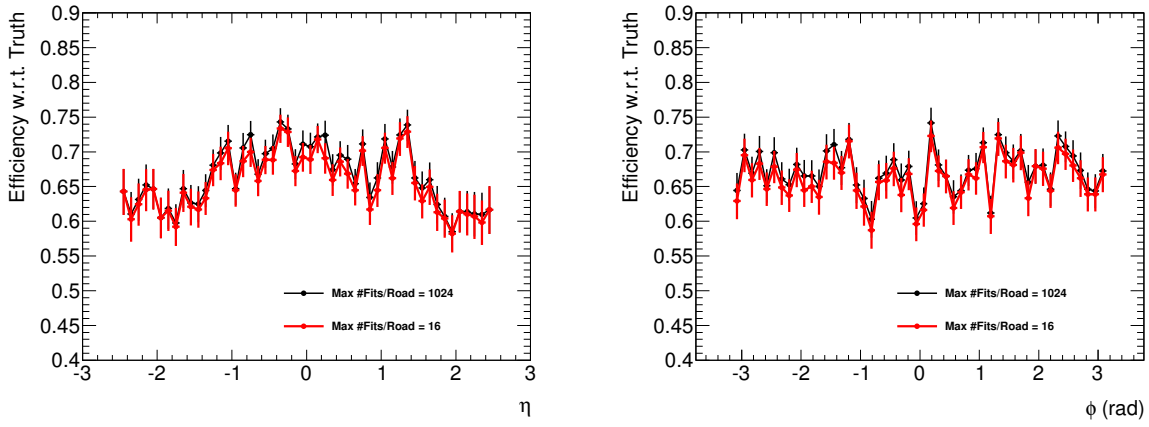

Figure 75: Track reconstruction efficiency with respect to truth tracks are compared between the maximum number of fits per road is 1024 and 16. The used event is $t\bar{t}$ with 69 pile-up interactions.

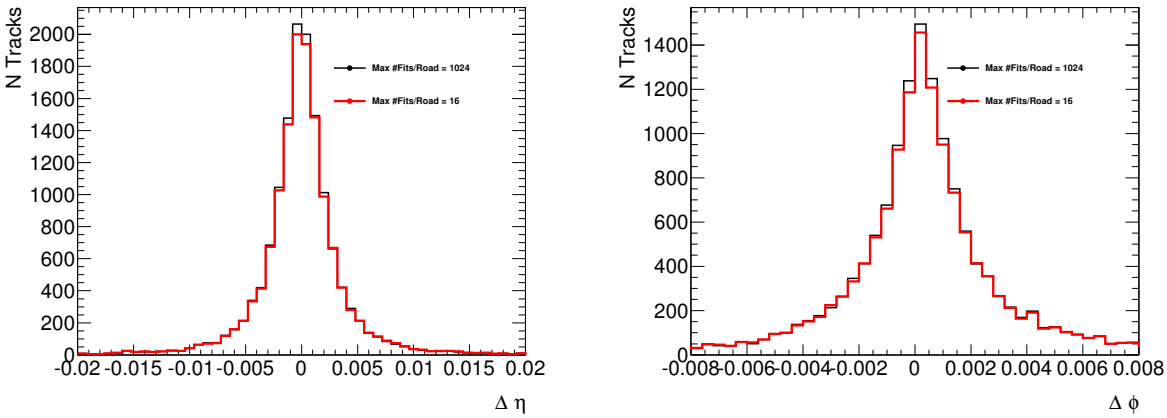

Figure 76: Resolution with respect to truth tracks are compared between the maximum number of fits per road is 1024 and 16. The used event is $t\bar{t}$ with 69 pile-up interactions.

5 Development of the Input Mezzanine Board

The FTK consists of several electric circuit boards. Since each of them is optimized for the FTK, it is necessary to do board design, production, and testing by ourselves instead of using a commercial one. Waseda university is responsible for board design, production, firmware development, installation, commissioning and operation of the IM.

As described in Sec. 3.2.2, the required functions for the IM are to receive hit information from the IBL, the Pixel detector and the SCT, to cluster the hit information, and to forward it to the DF without delay in latency at maximum 100 kHz input rate. The IM was designed and developed to match with this requirement. In this section, production, development and testing about the IM are described.

5.1 Prototype Board Production and Initial Tests

At the very beginning of board production, preliminary concept of the IM board was determined to achieve the requirements. Two Spartan-6 FPGAs were chosen to be implemented on a board to control all processing of the IM. Each FPGA receives two fibers, one from the IBL or the Pixel detector and the other from the SCT. The SFP connector was chosen as an input port of SLINK fibers from the IBL, the Pixel detector and the SCT. For early prototype of the IM boards, the HPC connector was implemented to connect the IM with the test mother board which is called EDRO instead of FMC connector since the prototype DF board was also under production and not ready yet. The EDRO board provides power to the IM and makes possible to do initial tests for the IM. Other components such as JTAG connector, power connector, SRAM, LEDs, and clock were also implemented. The prototype board of the IM was designed referring to the board which had similar components. A photo of the first prototype of the IM is shown in Fig. 77.

Figure 77: A photo of the first prototype board of the IM.

The initial tests were performed to validate and check the function of the board. The test items are:

- Check all the components are mounted.
- Check the components are placed in correct position with correct direction.
- Measurement of the value of resistors to check they are conducted correctly.
- Power on test.
- Download firmware via JTAG.

There is possibility of the deficit of components or it is mounted but with wrong direction. Thus it is important to find those mistakes by eye or measuring value of resistors before power it on since it would result in corruption of the components. When it was validated that the components were placed correctly, the board was powered on and firmware installation was tested to check if it works without problem.

If the board passes all of these initial tests, next step was to implement functions in the firmware. It was necessary to setup test bench to verify the functions of the IM. Boards which emulated the ATLAS inner detector ROD and ROS were used. The used instruments to be setup are:

Single Board Computer (SBC)
The SBC is a board which is inserted in VME, and controls and monitors VME by communicating with external computers via Ethernet cable. The board type number used for the IM test was vp315.

EDRO board
The EDRO board is a mother board which was used at an early stage of the development of the IM when the DF board was not produced yet. It is implemented on VME and holds a Stratix-II FPGA. It can hold two IMs and a transceiver card, HOLA, which are mounted via HPC connector. The EDRO board can imitate a part of functions of the DF board. A photo of the EDRO board with one IM and HOLA mounted is shown in Fig. 78.

QUad SLINK Transceiver (QUEST)
The QUEST is a PCI card which sends data by same protocol as the ATLAS ROD. Four links can be fed simultaneously at maximum. Figure 79 left shows a photo of the QUEST card.

Two Input SLINK for ATLAS Readout (TILAR)
The TILAR is a PCI express card which receives data with same protocol as the ATLAS ROS. Two links can be fed simultaneously at maximum. Figure 79 right shows a photo of the TILAR card.

Figure 78: A photo of the EDRO board with one IM and HOLA mounted. The top most card is HOLA and the next one is an IM. The mezzanine mounted at the center holds Stratix II FPGA.

Initial functions of the IM firmware were tested with using a set of these boards. The setup built at Waseda is shown in Fig. 80. The data is fed from the QUEST card, the IM receives the data and performs clustering, forwards it to the EDRO board through HPC connector. The EDRO sends the data from the

Figure 79: Photos of the QUEST card (left) and the TILAR card (right).

HOLA to the TILAR card without processing the data, and then the data is received and recorded at the TILAR card.

Figure 80: A schematic view (left) and photo (right) of the test bench built at Waseda. An EDRO board which holds two IMs and one HOLA card is in VME rack. The TILAR card is implemented on the computer under the VME rack, and the QUEST card is implemented on the right PC.

Goals to be established for the initial IM firmware were:

- Receive data from the QUEST.
- Output data to the EDRO.
- Validate the output data is as expected.
- Establish stable data flow at high input event rate.

The SLINK [70] is used as one of the standard protocol for board communication through optical fibers in the ATLAS experiment. It can be used as virtual ribbon cable to connect front-end and readout at any stage in a data flow. It also holds return channel to send error signal, flow control signal and user defined return signals. Maximum clock frequency and data flow rate is 40 MHz. The data is transferred in 32-bit. There is a protocol of SLINK reset to perform handshake between transceiver side and receiver side. In

case of implementation for the FTK, the protocol starts by transceiver side to send reset for the receiver side first and make it ready. Then receiver side tells that information to transceiver side via SLINK return channel, and then transceiver side is reset and gets ready. The data communication does not start if both sides are not ready. Figure 81 shows a schematic view of this SLINK reset procedure.

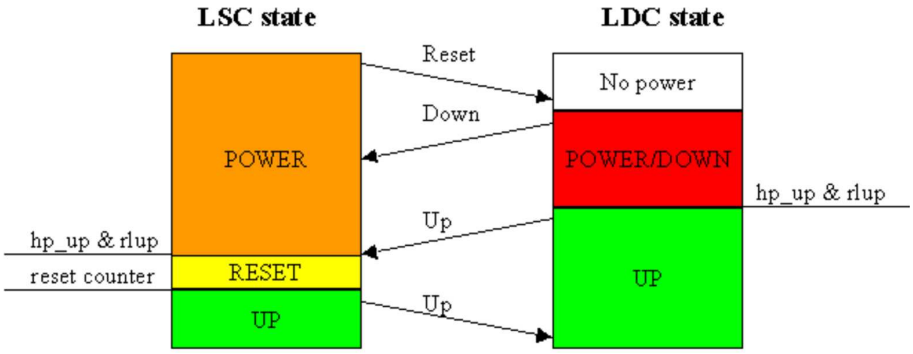

Figure 81: A schematic view of the SLINK reset procedure [71].

Data sent from the QUEST card is encoded with 8b/10b transformation which is same as the ATLAS ROD. It is high-speed serial data transfer which maps 8b words to 10b words, including a clock. It is a safer way to transfer data compared with parallel data transfer because it can avoid parity difference caused by the different length of the lines and achieve easy clock recovery with a loss of bandwidth of 20%. It is required for the firmware of the IM to decode 10b words to 8b words.

The reset protocol of SLINK and 8b/10b decoding were implemented in the firmware of the IM. Firmware reset is sent from an external computer via SBC, which cleans up FIFOs and initializes signals in the firmware. Reset of the QUEST is issued directly from the computer on which the QUEST is mounted. With this setup and functions, the data transfer between the QUEST and the IM was verified and established.

Next, the function of hit clustering was implemented and tested. It is the core part of the IM firmware and required to work stable under any realistic condition. The input data was made by hand referring to the ATLAS ROD users manual [72]. The ISE Simulator (ISim) which is a tool provided by Xilinx to simulate behaviour of signals in the firmware was used as a first check. The behaviour of signals in the firmware are visualized at clock level. An example of waveform simulated by the ISim is shown in Fig. 82. Signals such as a clock which is in the most upper part (clk), and input data (din) and output data (dout) can be seen. By changing setup and input data before running the simulation, behavior of the signals for that configuration can be simulated and verified. Since to perform test with real hardware takes long time as compiling firmware, generating file to be installed to hardware, and to setup hardware, The ISim is very useful tool to reproduce and detect issue quickly.

Issues such as delay of processing, stack of dataflow, and unexpected output were resolved one by one. The delay of processing is resolved by increasing clock speed or improvement of algorithm itself. The stack of dataflow is mainly caused by wrong handling of backpressure signals and then the critical signal for state control of the firmware is lost. It should be reproduced with the ISim at first, and then implement correct handling of backpressure signals. The unexpected output can be fixed by tracking all signals in the ISim and detecting signals which are not as expected.

When it was validated that clustering function worked stable and the output was as expected, behavior of the firmware was tested with real board. Here input data was fed from the QUEST, and the data was sent out from the IM to the EDRO board and finally recorded at the TILAR. To forward data from the IM to

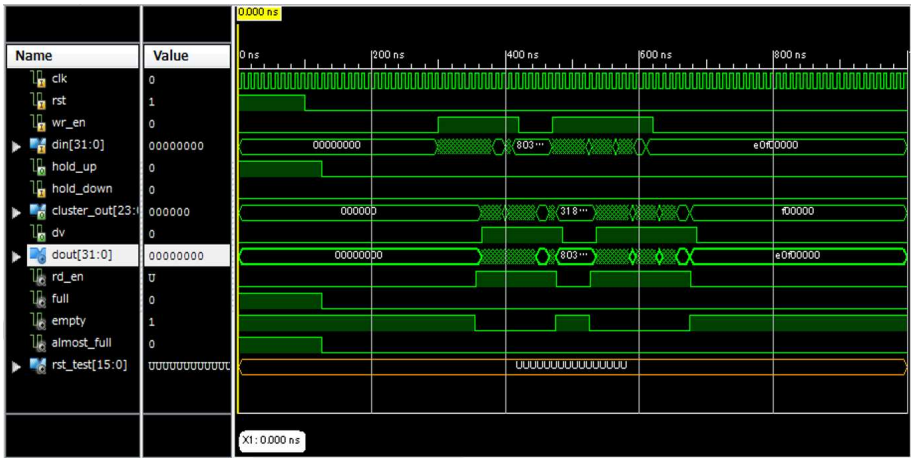

Figure 82: An example of the ISim results.

the EDRO, the lines fed from FPGA to the connector should be connected correctly in the firmware. The connection of those lines is defined in schematics, and it is necessary to match with the EDRO board.

Stability at high input rate was checked next. The QUEST can feed 32-bit data at the rate of 40 MHz which is same as the ATLAS ROD. The IM utilizes several FIFO buffers which have ability to do flow control. If FIFO becomes almost full, it issues ALMOST_FULL signal to the upstream and dataflow is suspended until data in the FIFO is processed and ALMOST_FULL signal is released. It is necessary to take care of handling of ALMOST_FULL signal since mis-handling of the signal could lead to loss of data. The loss of data itself is an issue, and in case the lost data is important to control state of the firmware, dataflow gets stack. The higher input rate means the higher probability to become ALMOST_FULL of FIFO. Iterating tests using the ISim and real hardware, wrong handled ALMOST_FULL signal was detected and resolved. Finally, stable dataflow was achieved with test bench of real hardware at the maximum input rate.

5.2 A Series of Production of the Prototype Boards

There were several points to be modified for the prototype of the IM and some points are found to be optimized through testing. Getting feedbacks from testing experience, a series of prototype boards were produced. Here the summary of modification at each version of the IM is described.

version 1.0 Produced on February 2011. First prototype board of the IM, shown in Fig. 77.

version 2.0 Produced on April 2012. Some minor changes.

version 3.1 Produced on February 2013. Connector to motherboard was changed to FMC connector.

version 3.2 Produced on November 2013. Power supply and SFP connector are changed.

version 3.3 Produced on March 2014. Minor changes to be more useful before mass production.

Photos of the IMs of version 3.1 and 3.2 are shown in Fig. 83.

In addition, the firmware had been modified to be integrated with the DF and also to work against difference of data format among the IBL, the Pixel detector, and the SCT. Integration with the DF is described in

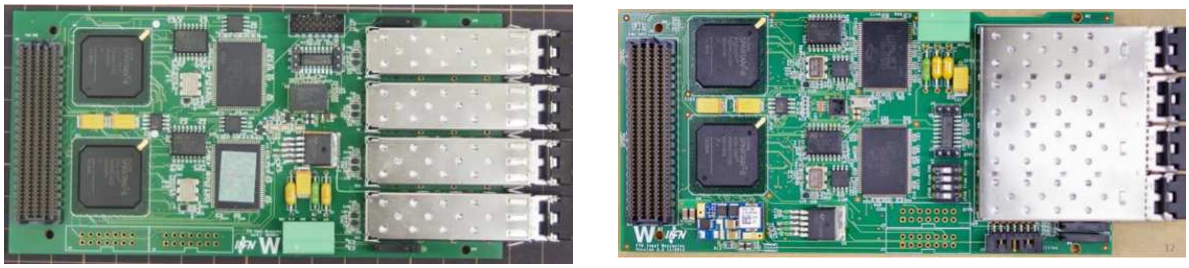

Figure 83: Photos of the IMs of version 3.1 (left) and 3.2 (right).

Sec. 6.2.1. Board design and results of integration test were reviewed by the ATLAS committee and got accepted for mass production.

5.3 Mass Production and Quality Control Test

Final design of components of the IM is shown in Fig. 84 left and structure of 12 layers of IM board is shown in Fig. 84 right. Since the components are densely located, it is important to perform circuit wiring without short-circuit and cross-talks.

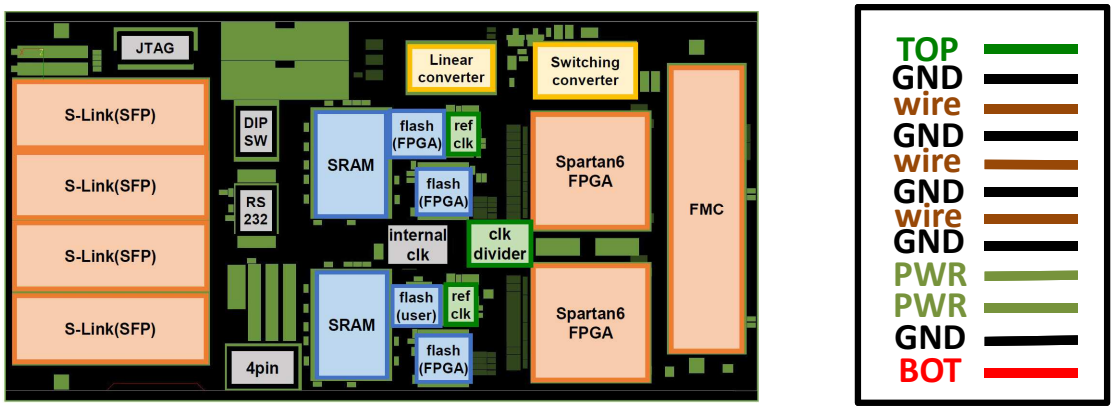

Figure 84: Final components of the IM board are shown in the left and the IM board structure of 12 layer is shown in the right. TOP and BOT stand for top layer and bottom layer, GND for ground layer, wire for the layer which lines is running, and the PWR is the layer used for power supply lines.

252 fibers are fed from the IBL and the Pixel detectors and 128 fibers are fed from the SCT, in total 380 fibers are fed from the IBL, the Pixel detector and the SCT. For the IM, one FPGA has ability to process two fibers, but clustering of hit information from the IBL and the Pixel detector requires so many FPGA resources that one FPGA cannot work if both fibers are fed from the IBL and the Pixel detector. Thus the necessary number of IMs is defined by the number of fibers from the IBL and the Pixel detector. Consequently at least 126 IMs are necessary to process 252 fibers fed from the IBL and the Pixel detector. 80 S6 IMs and 80 A7 IMs, in total 160 IMs were produced including enough spares.

Waseda team produced 80 S6 IMs in February 2015. The components were collected by ourselves to save cost rather than asking company to collect them. At the mass production, initial check was done at factory to detect errors quickly. After transporting to Waseda, more detailed quality control tests were

performed. The tests were done for all the produced 80 S6 IMs. The test items are as following, which are sophisticated through accumulated testing experience and knowledge.

- Same tests as performed for the first prototype board listed in Sec. 5.1.
- Validation of the basic firmware functions.
- Bit Error Rate (BER) measurement.
- High load test by feeding data from all the four channel inputs.
- Stability and durability test.

When problem was found, the reason was investigated and resolved immediately. The BER is defined as error rate of lines in the entire IM board, thus it is invert of the number of lines in an IM times seconds. The requirement of the BER for boards installed to the ATLAS is to be less than 10^{-15}. Duration were verified by keeping operating boards with high load. It was validated that the IM runs without any error for more than 100 days.

Problematic boards were quickly resolved. Consequently, the yield ratio achieves 100% after the quality control tests. 79 boards were shipped to CERN. Same tests were performed after shipping to CERN, and again all the boards passed the tests. A photo of 80 mass produced S6 IMs and some members of laboratory are shown in Fig. 85 left.

The A7 IMs were produced by Frascati team in Italy. They were performed same quality control tests as the S6 IMs. A photo of the mass produced A7 IMs is shown in Fig. 85 right.

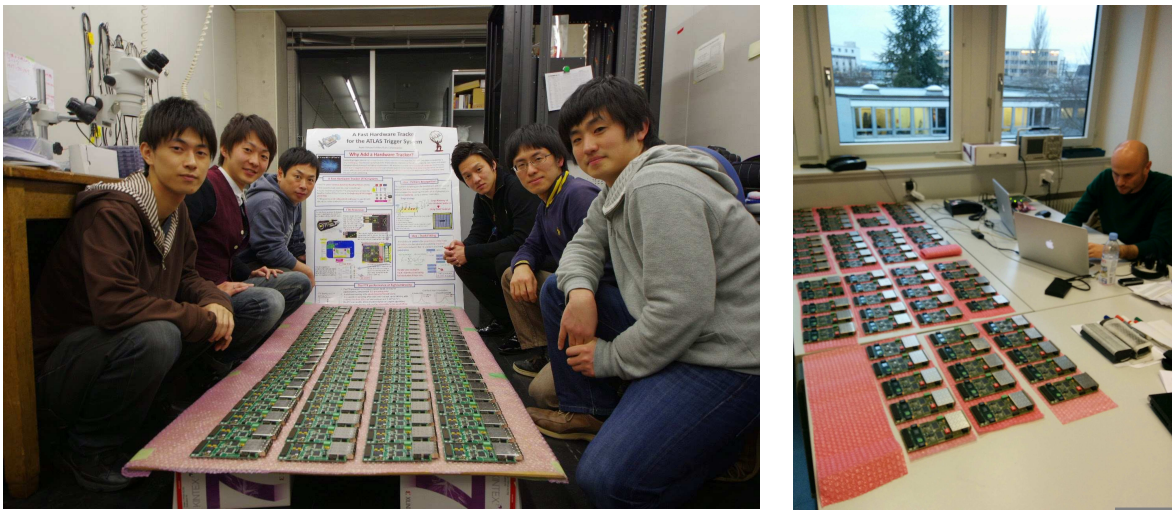

Figure 85: A photo of the mass produced S6 IMs and the members who performed board production, initial tests at the factory, and the quality control tests (left). A photo of the mass produced A7 IMs (right).

6 System Integration, Installation and Commissioning

After production and quality control tests of each board were completed, they were integrated to work as a system, and the entire system was installed into the ATLAS trigger. Performance of the hardware was validated and improved through commissioning and operation. The history of integration, commissioning and operation is summarized in Fig. 86. The first step was integration with the SCT which was performed in 2012, called Vertical Slice Test. Connection between the inner detector and the FTK input interface was established at this time. Next, integration of the entire FTK system were done from 2013 to 2014, called FTK integration. From 2015, installation of the FTK system into the ATLAS trigger was started, and then commissioning has been performed. Detail of a series of the tests, system integration, installation and commissioning are described in this section.

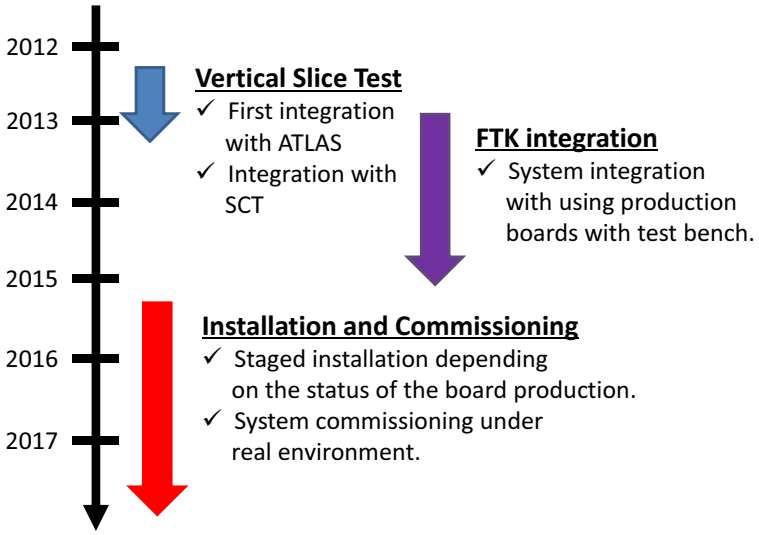

Figure 86: The history of the major FTK tests, installation and commissioning.

6.1 Vertical Slice Test: Communication with SCT

After the prototype of the IM board was produced and basic performance was checked with the test stand at Waseda, integration with the SCT was performed, which is called Vertical Slice Test. It was the first time that a part of the FTK system was integrated with the ATLAS inner detector. Goals of the Vertical Slice Test were to validate and establish communication between the IM and the SCT ROD and to achieve stable dataflow without affecting the ATLAS main data stream (parasitic mode) at the LHC physics run. Another purpose is to build prototype of online software (Run Control) which are used to configure and monitor the FTK and integrated into the ATLAS Run Control.

From the end of 2012 to the beginning of 2013 (i.e. at the end of LHC Run1), the Vertical Slice Test was performed. This test was performed by receiving data from a part of region of the SCT. The detector regions used at the Vertical Slice Test are shown in Fig. 87. It corresponds to 45° in ϕ and half of the barrel in z direction.

The D-HOLA is implemented on the SCT ROD. As described in Sec. 3.2.1, it has two output ports, one sends data to the ATLAS main data stream and the other sends to the FTK. Communication between the

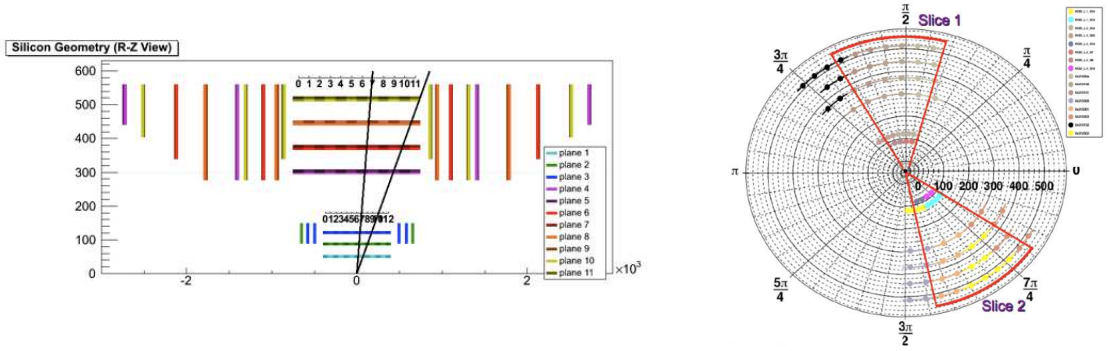

Figure 87: A schematic view of detector regions used at the Vertical Slice Test.

FTK channel of the D-HOLA and the IM must be established. Before start of the Vertical Slice Test, backpressure from the IM to the D-HOLA must be off since the FTK should operate with parasitic mode and should not affect the ATLAS main data stream during the the Vertical Slice Test. The backpressure is propagated through the SLINK fiber and the HOLD protocol is called as XOFF. It was tested and validated in advance at the interval of the LHC physics runs, and confirmed that the FTK did not send XOFF at all. Thus the FTK was accepted to start the Vertical Slice Test.

The FTK channel of the D-HOLA worked with modified SLINK protocol called as slave mode which ignored return signals from the IM except for XOFF, and communication can start without the SLINK reset procedure described in Sec. 5.1. With slave mode, the IM can receive data immediately after fiber is connected, and basically the IM can do anything without affecting the SCT ROD. The firmware of the IM was modified to be fit with the SLINK slave mode and data receiving was established between the D-HOLA and the IM.

The test setup was installed and built in the ATLAS DAQ room which is called USA15. A photo and schematic view of the setup of the Vertical Slice Test are shown in Fig. 88. Two IMs, EDRO, AM, ROD and ROS were used. Here the ROS was not the ATLAS ROS but the one especially prepared for the FTK test. Data was fed from the D-HOLA of the SCT ROD, the IMs received the data, performed clustering and forwarded it to the EDRO through HPC connector. The EDRO sent the data to the AM through VME bus and the AM sent back fake roads to the EDRO. Finally a HOLA card mounted on the EDRO sent the data to the ROS where the data was recorded.

First, clustering function was verified. The goal was to establish stable processing under real environment and to validate the outputs were as expected. Since spy buffer was not implemented on the IM firmware at the Vertical Slice Test, the spy buffer of the EDRO was used to collect input and output data of the IM. The spy buffer data in the EDRO can be retrieved through VME bus to an external computer. The output of the IM was stored in input spy buffer on the EDRO board. In addition, one of the two channels of the IM was operated with "pass through" mode which sent input data directly to the EDRO board without any processing. Thus input data and output data of the IM were both stored in the spy buffer of the EDRO board.

The input data was used as, for example, an an input of simulation. The output of the simulation and the output taken from spy buffer in the EDRO were compared. When they were not compatible, it would be investigated and resolved by developing firmware. One of the reasons of the incompatibility was that the data format was slightly different than as expected (since it was the first time for the FTK to take data from the real SCT ROD). It was resolved by modifying firmware to be able to cope with the format. Other

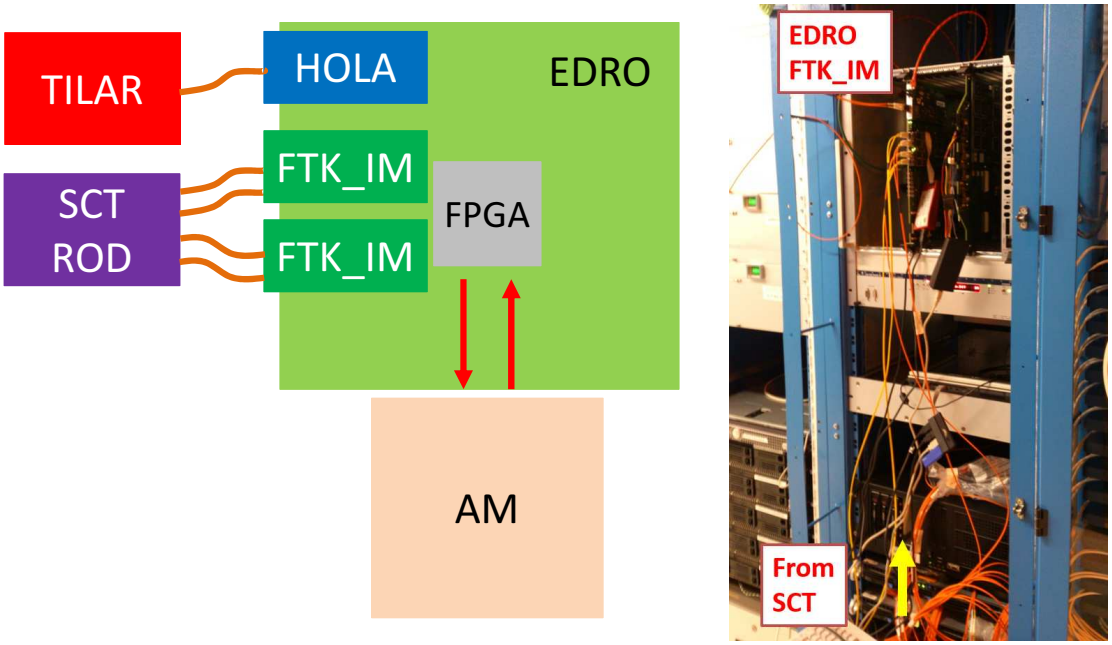

Figure 88: A schematic view and photo of the setup for the Vertical Slice test.

reason was that even if the data format was as expected, some special order of data sequence got firmware into unexpected state which led to wrong output. Iterating the data taking with real hardware, firmware development and validation with simulation, the output of the IM hardware matched with the expected output.

Next stability against high input rate was established. Though dataflow works without problem at low input rate, it is not ensured to be able to operate stable at high input rate since the FIFOs become ALMOST_FULL more frequently and unexpected data sequence comes more frequently that the IM cannot process correctly. It was also resolved by iterating hardware test and simulation.

The online software to monitor and configure the FTK was developed and established as well. The Run Control is a GUI developed to enable non-experts of the FTK to configure and monitor the FTK at real operation. It has the same format as the one of other ATLAS detectors. They are integrated into one system which is called ATLAS Run Control. The Run Control of the FTK needs to be also integrated into the ATLAS Run Control, and operated in the same way as the one of other detectors. A photo of the ATLAS Run Control when a part of the FTK Run Control was integrated into at the first time is shown in Fig. 89. The behavior was tested at the dedicated time slot for the FTK, and it was established to operate with other detectors at low input rate. At high input event rate around 75 kHz, dataflow stopped by Central Trigger Processor (CTP). It monitors all the ATLAS detectors and synchronizes event among them. Dataflow stopped since the FTK was not able to process data at high rate without delay in latency and lost some data including event information. To deal with high rate input was one of the remaining issue of the Vertical Slice Test.

In summary of the Vertical Slice Test, the communication between the IM and the SCT ROD was established, clustering function was validated and developed with real data, and the procedure to validate hardware output was established. A prototype of the FTK Run Control was also developed and tested. Stack of dataflow at high input rate was the remaining issue.

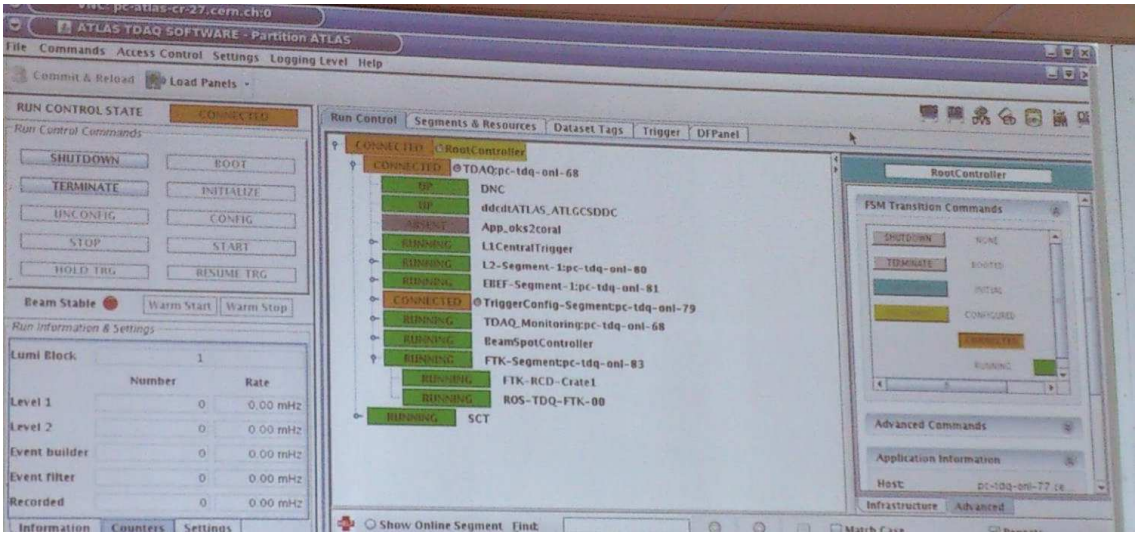

Figure 89: A photo of the ATLAS Run Control with including the FTK first time. The FTK related nodes can be found in the second last part of the list.

6.2 Integration of Full FTK System

When production version boards of each sub-system became ready, the FTK system integration was performed. The boards were integrated when they were produced and ready. The goal was to establish communication with neighbouring board one by one, and finally establish stable dataflow throughout the entire FTK system.

Items to be established to integrate two neighbouring boards are as following.

- Dataflow from input to output in a single board, at least with pass through mode.
- Data receiving from the upstream board.
- Flow control with the upstream.
- Data sending to the downstream board.
- Flow control with the downstream.
- Stable dataflow with full function.

At first, the integration had been performed with test stand built at the CERN DAQ test room which is called Lab4. When stable system was built at the Lab4, they were transferred to the USA15.

6.2.1 Integration of IM and DF

The IM and the DF were integrated at first of the entire FTK system. In case of IM and DF integration, what needs to be validated was not only data transmission but also other functions such as power supply and JTAG lines through FMC connector since the IM is the mezzanine card of the DF. Connection through the FMC connector was tested and validated using test board at first to ensure that the IM and the DF can be integrated without problem, which was important to avoid damaging the boards.

For the IM, it was tested and validated with Xilinx evaluation board named KC705. A photo of the KC705 with an IM mounted are shown in Fig. 90. It has two FMC connectors and Kintex-7 FPGA which can emulate pin connection and a part of functions of the DF. The pin connection between the IM and the KC705 was validated using digital multimeter. It was confirmed that power was supplied from the KC705 to the IM through the FMC connector and the IM can be powered on. The JTAG lines were also validated to be connected correctly and installation of firmware was succeeded to work through the FMC connector. Simple data transmission was tested and validated by sending data through the FMC connector and checking it in the KC705. And all the other lines were checked and validated to be connected and worked as expected. For the DF, similar tests and validation were performed using test mezzanine card.

Figure 90: A photo of the KC705 with an IM mounted .

When connection was validated for each of the IM and DF with the test board, the IM and DF were integrated. Similar tests were performed as with the case of test board and the connection was validated.

Next, data transmission was developed and established. Signals sent from the IM were checked in the DF if they were transmitted as expected. Since 32-bit data are sent from the IM to the DF through parallel Double Data Rate (DDR) lines at 200 MHz clock and the length of each line is not perfectly same in the logic of FPGA, it is possible that timing to readout each bit are slightly different in the DF. Thus it is necessary to set “delay value” for each channel to adjust the timing to readout data in the DF.

Figure 91 shows concept to set delay value. Here it is assumed that the IM sends 4-bit data with all the bits high and the figure shows shape of each bit around clock decision of the DF. Each bit corresponds to line1 to line4. The timing to arrive at the DF is slightly different among lines due to different line length in the logic of FPGA. In case of Fig. 91, if the DF reads out at timing3, wrong value (1010) is readout, while timing5 to timing9 gives correct value (1111). Thus one of the value among timing5 to timing9 is chosen and used during operation. In firmware, it is implemented as IM keeps sending pre-defined data pattern to the DF at idle state which is the state waiting for data coming, and the DF searches for optimal delay value which the expected words can be readout.

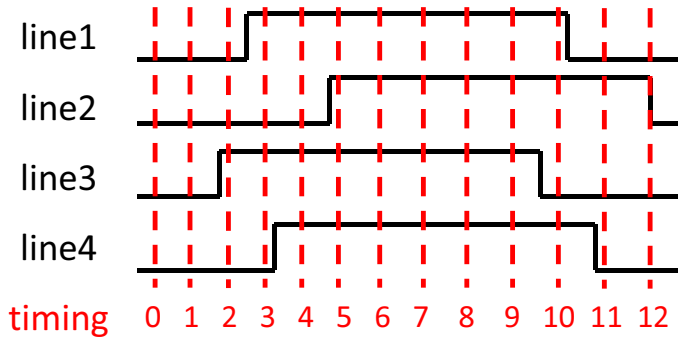

Figure 91: A schematic view of signals in the DF sent from the IM.

Some register signals which can be used for monitoring and configuration were implemented in the IM firmware. They are accessed via Inter-Integrated Circuit (I2C) between the IM and the DF. The I2C consists of two lines, one is called SCL which corresponds to clock line and the other is SDA which corresponds to data line. The I2C consists of master and slave where the DF is master and the IM is slave in case of the IM and the DF.

Figure 92 shows working principle of the I2C. The concept of the I2C implementation is as following. Basically the I2C data packet consists of two blocks of 8-bit words, the former block stands for address of register to be accessed and the latter block for data to be readout from or written to the register of indicated address. The 8-bit signals are transmitted one bit by one bit. The process always starts by master. When both SCL and SDA lines are up, it means idle state. When master gets SCL down while SDA is kept up, it means the start of the I2C communication. When SCL is down, the next bit of SDA is set to 0 or 1, and when SCL becomes up, the value is readout. After the value is readout, SCL gets down again and then the next bit of SDA is set. This process is repeated 8 times thus 8-bit words are transmitted.

In the first block, first seven bits are assigned for address and last one bit stands for read/write bit. If it is up, the data are readout at the latter 8-bit transmission, otherwise data are written to the register. Acknowledgement bit which corresponds to error checker is sent after each 8-bit transmission. The I2C communication ends when SDA becomes up while SCL is up.

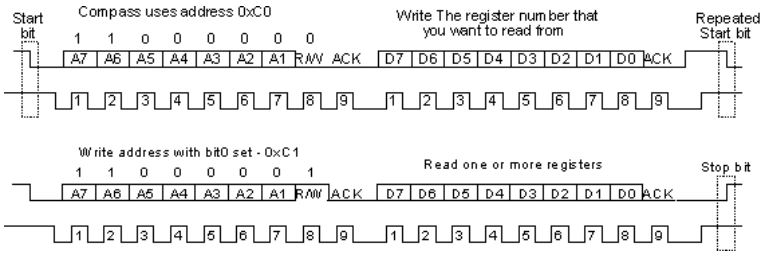

Figure 92: A schematic view of working principle of the I2C. The upper figure shows the case to write data from master to slave, and the lower shows the case to read data from slave to master. In each figure, the upper stands for SDA and the other for SCL [73].

The communication between registers in DF and an external computer are performed by IPbus protocol over Ethernet. The protocol to access registers in IM and DF system is shown in Fig. 93. The IPbus protocol is transactional and a packet-based control protocol to readout and modify registers within FPGA-based hardware devices which can be accessed by IP. For each Read and Write operation, the IPbus client

(typically software) sends a request to device, and then the device sends back a response. Control Hub is used which is a software application that forms a single point of access for IPbus control of each device. It checks the simultaneous access from multiple control applications as well as packet loss, which increases IPbus stability. Thus the monitoring and configuring method was established which enabled smooth hardware testing.

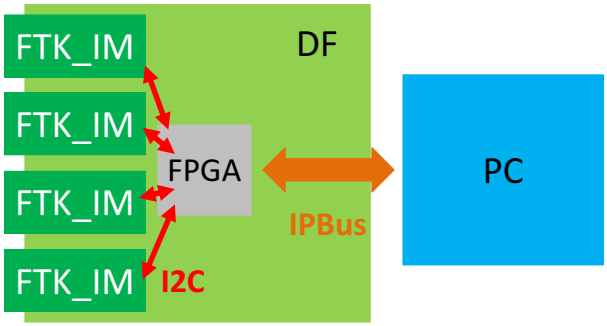

Figure 93: A schematic view where I2C and IPbus work.

After establishing infrastructure to perform hardware test, stability of dataflow through the IM and the DF were checked. A schematic view and a photo of the test setup are shown in Fig. 94. The test stand was built at the Lab4. At this test, a RAM in the IM firmware was used to feed input data which can start anytime when command was sent over the IPbus and the I2C. It is very useful way to validate and develop firmware since it is apart from uncertainty caused by other systems so that it can reproduce results perfectly with real hardware. It was used to feed input data along with or instead of the QUEST.

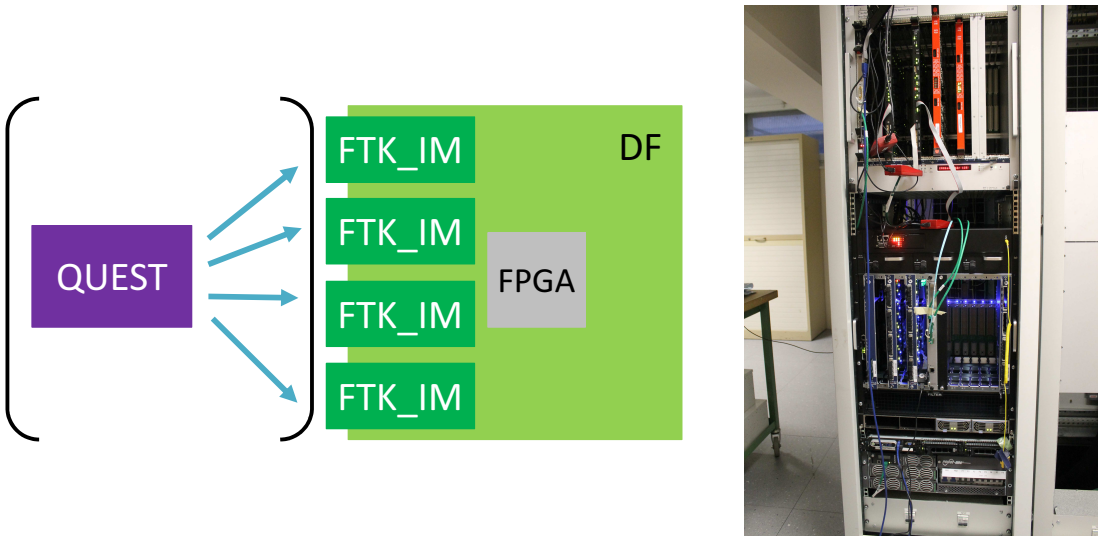

Figure 94: A schematic view and photo of the test stand setup built at the CERN. Boards in the lower rack are the DFs and in the upper rack are the AMBs and the SSBs in the photo. The AMBs have red front panel, while the SSBs have black front panel. The AUX cards are installed behind the AMBs.

One important requirement is that the DF needs to perform event synchronization among all input channels to share hit information. In terms of firmware, it is implemented as the DF sends backpressure to channels which finish processing of the event faster to wait for channels which takes long time for processing of the event. If event processing are finished for all channels, backpressure from the DF is released and the DF

starts processing the next event. Processing time for each channel of the IMs depends on the size of input data so that the regions where information is dense in an event take longer time. The dense region would be almost random event by event so that processing speed would be balanced among channels for long time operation. In terms of the IM, backpressure from the DF makes the FIFO ALMOST_FULL more frequently which should cope with. In case backpressure is not treated correctly in the IM, data is lost or duplicated and dataflow stops. Since the behaviour of backpressure from the DF cannot be emulated thoroughly with only IMs, it is tested using real hardware of the DF and the IM.

To establish stable dataflow, a lot of kinds of dataflow tests have been performed with several kinds of input data. One example of the occurred issues was that since event number information can take all values by definition, thus when it took the format of moduleID, the IM treated it as module and the IM entered wrong state, then dataflow got stack. It was resolved by attaching a bit along with event number information which indicates as such. Other example was that a DF knows all moduleIDs it receives from the IMs to share data among the DFs based on detector region, thus when moduleIDs are duplicated or wrong moduleIDs are sent from the IMs due to mis-handling of backpressure, the DF stops processing. It is resolved by detecting where the mis-handling happens and fixing it properly.

Iterating this kind of tests and firmware development, finally a system of 4 IMs on a DF can achieve stable dataflow with all 16 channels running simultaneously at the input rate of 100 kHz with the test stand.

6.2.2 Integration of AUX, AMB, SSB and FLIC

Down stream boards were also developed and integrated as the IM and the DF. The AUX card is the next downstream board of the DF. It is implemented on back plane of VME crate. It connects to the DF through the SLINK. Dataflow was checked and established with the IM-DF-AUX system. Input data was fed from the IM RAM, it was clustered in the IM and forwarded to the DF, the data was sent to the AUX. At first a single DF board was used and it shared data only within the board, and sent the data to the AUX. The process of development was similar as the case of the IM and the DF integration. The communication was established with one board for each at first, then the number of boards were increased.

Next, the AUX and the AMB were integrated. The AMB is implemented in front of the VME crate. Data was fed to the AUX from the QUEST directly at early stage of integration to reduce uncertainty by including the IM and the DF. After communication was established between the AUX and the AMB, the IM-DF-AUX-AMB system integration was performed.

In parallel, the SSB and the FLIC integration was performed. Since the SSB receives data from both AUX and DF, development of dataflow among the boards are complicated. Thus dataflow between the SSB and the FLIC was established in advance. A photo of the test stand of the FLIC is shown in Fig. 95. The SSB is implemented in the VME crate and the FLIC is implemented on the ATCA. They communicate each other through the SLINK. The SLINK protocol was implemented. Test input data was fed from the SSB RAM. The FLIC and the ROS integration was then performed. Communication is also performed through the SLINK.

Next communication between the DF, the AUX and the SSB was established. Since the SSB merges hit information sent from the DF and the AUX, event synchronization is required to perform global tracking. The communication among them is performed through the SLINK. And finally, the entire FTK system communication was developed and established.

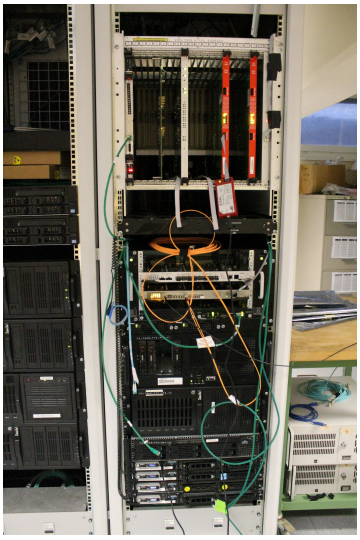

Figure 95: A photo of test setup of the FLIC crate. The FLIC can be seen in the middle crate. The rack is placed at the next of the one of Fig. 94.

6.3 Installation and Commissioning

After board production was done and stable performance was established with test stand, the FTK system was installed to the USA15. First, infrastructure such as racks, VME and ATCA crates were installed, and network configuration was setup. In total, 8 VMEs and 5 ATCAs were installed to hold boards. Figure 96 shows a schematic view of the arrangement of the FTK racks in the USA15. With the FTK full system, 16 AMBs and 4 SSBs are installed in each VME rack, and 8 DFs are in each ATCA crate. One ATCA crate is for the FLIC. Boards are installed when they are produced and ready. Installed boards were commissioned with real data. Photos of the installed boards are shown in Fig. 97, 98 and 99.

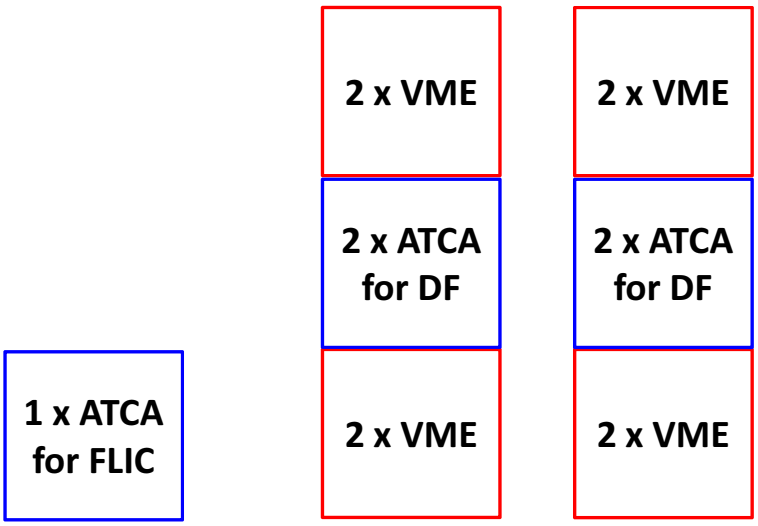

Figure 96: A schematic view of the arrangement of the FTK racks in the USA15 seen from the top.

Figure 97: Photos of the DFs and IMs installed in an ATCA crate (left) and the RTMs on its back plane (right).

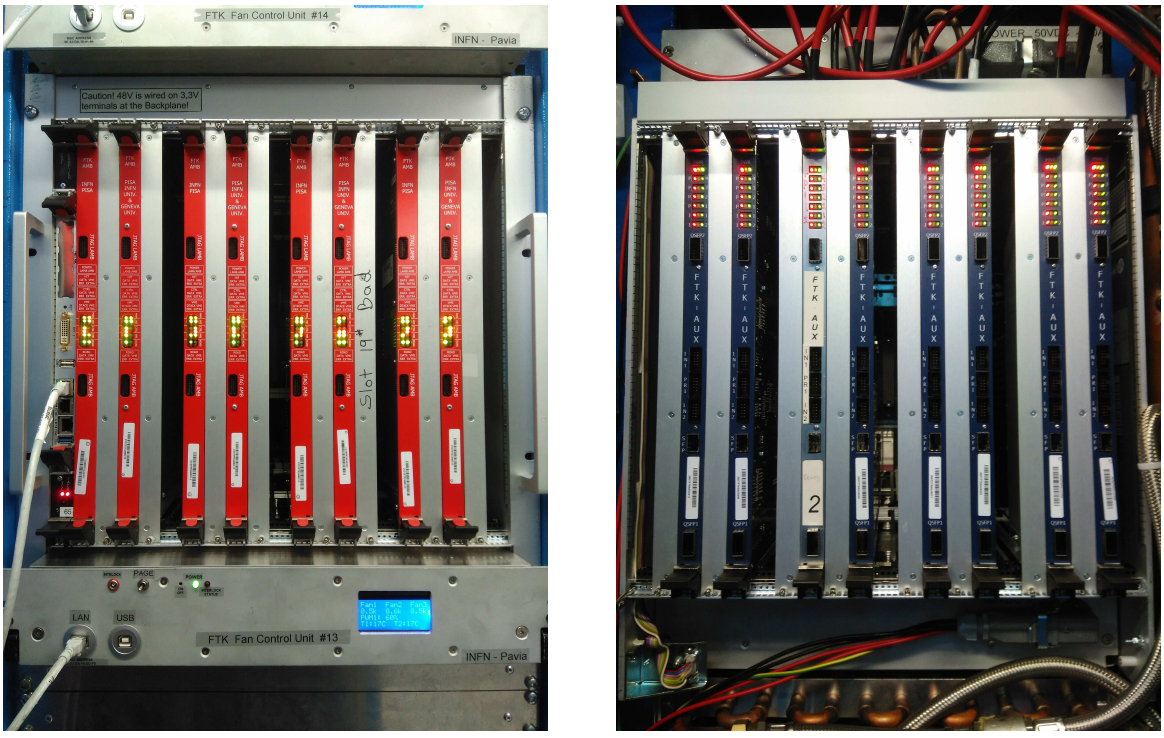

Figure 98: Photos of the AMBs and SSBs installed in a VME rack (left) and the AUX cards on its back plane (right).

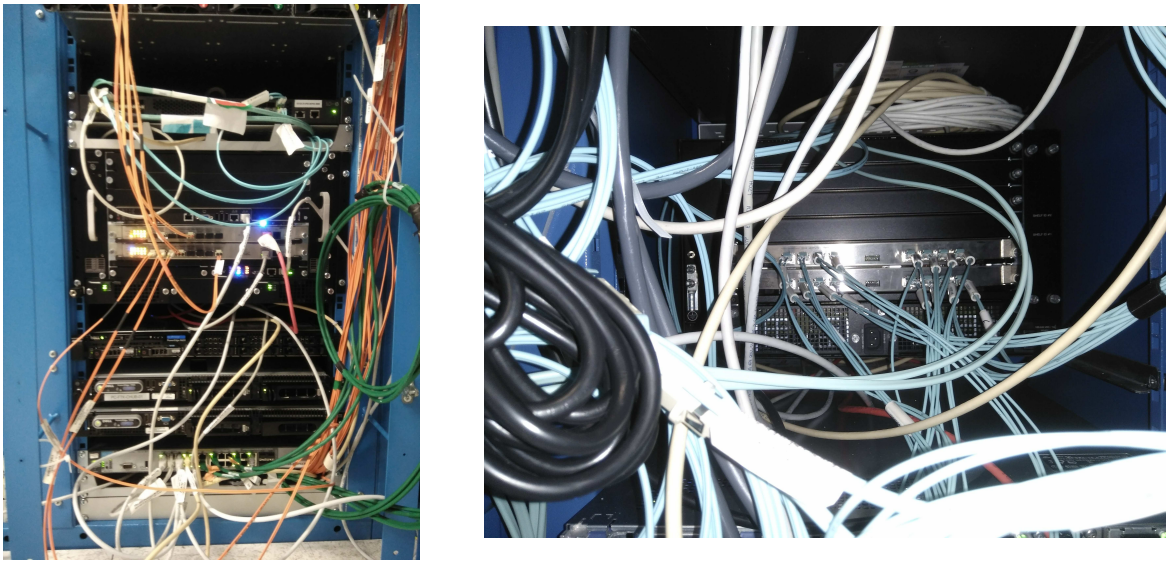

Figure 99: Photos of the FLIC cards installed in one of the VME racks (left) and the RTMs on its back plane (right).

Items to be developed are basically same as the ones at the integration test. At real environment of high input rate, to establish stable dataflow is major challenge. The difficulty is to cope with many kinds of input data patterns as well as to perform event synchronization among all channels to reconstruct track information for an entire event. An event arrives at different timing from each fiber and the data size in an event is also different. These lead to more backpressures to perform event synchronization, make the FIFO ALMOST_FULL more frequently, and require correct handling of them. This can be coped with by implementing efficient processing algorithm to maximize processing speed, increasing the size of buffers in the FTK system as well as implementing correct handling of backpressure. The goal is to establish stable processing for the entire FTK system for any event without any delay and loss of data.

When firmware was much developed and error rate becomes very small as an error per a few hours, it is very difficult to reproduce and resolve the error. Recovery system is implemented to make the FTK keep running at operation. If some FIFOs become ALMOST_FULL and it is not released for more than pre-defined clock cycles, reset is issued automatically inside the FTK and the FTK resumes processing.

Toward operation of the FTK system, the Run Control of the FTK has been also developed and integrated into the ATLAS Run Control.

6.3.1 Installation of IM, and Integration of IM and Inner Detector ROD

The IM and the inner detector ROD were integrated in advance of other boards since it is the upper most part of the entire FTK system. At first, to prevent sending XOFF unintentionally during commissioning, XOFF enable signal was implemented and validated as performed at the Vertical Slice Test. By default the D-HOLA ignores XOFF from the FTK channel when it is powered on. When the IM sends “enable XOFF” signal to the D-HOLA through the SLINK return channel, the D-HOLA gets ready to receive XOFF from the FTK channel. In addition, registers which work as a switch and password are prepared in the IM firmware and necessary to set correct values to send XOFF from the IM to the D-HOLA.

This procedure to enable XOFF was tested and validated with test stand at first. The sub-boards used for the test were the IM, the D-HOLA, and the FILAR card which receives and records data as the ATLAS ROS. The FILAR card is four I/O version of the TILAR. A front panel of the D-HOLA and the FILAR card are shown in Fig. 100. The left card in each photo is the FILAR card and the right one is the D-HOLA. The upper channel of the D-HOLA is main channel which is connected to the FILAR card and the other is the FTK channel which is connected to the IM. The left photo shows the situation that fibers are connected to both main channel and FTK channel, and the link up LED turns on. This is the case when XOFF is not issued from the IM or it is issued but XOFF enable protocol is not applied. The condition of the right photo is that XOFF enable protocol is applied, and also XOFF is issued from the IM. The XOFF LED of the FTK channel turns on (red one). Thus it was validated that XOFF was not issued without XOFF enable procedure, while XOFF can be propagated correctly in case it is applied. This satisfied initial requirement to install the IMs and integrate with the inner detector RODs.

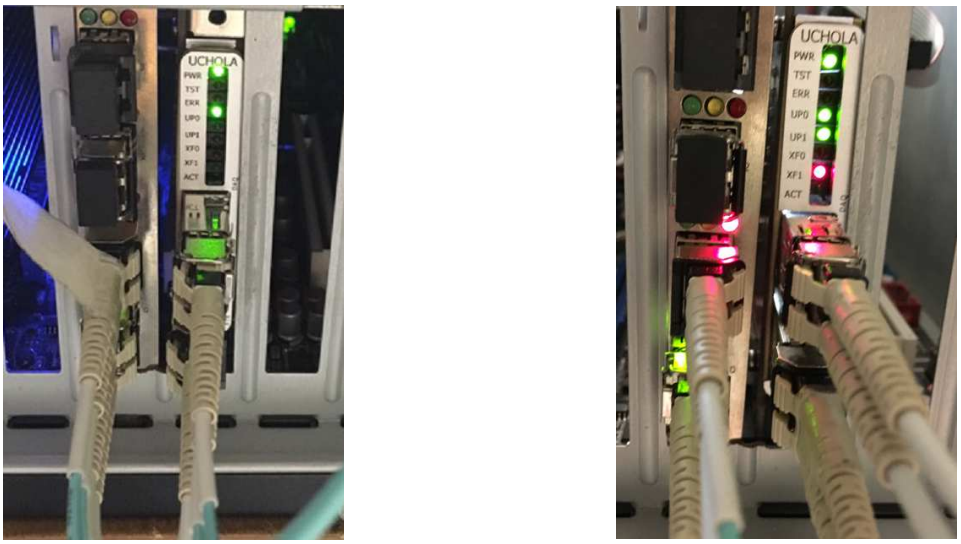

Figure 100: Photos of the front view of the FILAR (left card in each photo) and the D-HOLA (right card in each photo). The fiber connected to upper channel of the D-HOLA goes to the FILAR, and the other goes to the IM. The LEDs stands for POWER ON, Test, Error, Link up for main channel, Link up for the FTK channel, XOFF from main channel, XOFF from the FTK channel, action from the top to bottom.

128 IMs are installed in USA15. 380 fibers are fed from the IBL, the Pixel detector and the SCT in total and connected to the IMs. The configuration of fiber connection is optimized to minimize load for data sharing. The fiber connection was checked that all the IMs and fibers are connected correctly and data transmission was validated.

During shutdown of the LHC from February 2012 to June 2015, the IBL and a part of the Pixel detectors implement the BOC [74] as a readout instead of the D-HOLA. Figure 101 shows a photo of the BOC. Firmware of the BOC was different from that of the D-HOLA. The major difference related to the FTK was that it did treat return signals from the IM so that it was possible for the IMs to affect behavior of the BOC. In addition it had correlation between main DAQ channel and FTK channel. The IM firmware was modified to have more reliability to avoid sending unexpected return signals. For example, the IM switches off SLINK communication in the firmware during reset of the firmware to avoid sending unexpected signals. Firmware of the BOC was also modified not to have correlation between DAQ channel and FTK channel. Thus connection was established also for the BOC.

Figure 101: A photo of the BOC.

6.3.2 Installation and Commissioning of Entire FTK System

The IM and the DF were installed and commissioned at first of the entire FTK system since production boards were ready and validated in advance of other boards. Dataflow in the boards have been tested and developed, and function of clustering and data sharing has been improved by the feedback of tests at real environment. One example of issues was that it was found that with a little frequency, very big events were received from the IBL and the Pixel detector. In that case, most of the hits are along with z direction. Thus those are considered to originate from beam-halo and not from hard scatter events. The algorithm to eliminate those hits has been implemented. Iterating tests and firmware modification, it has been developed to operate at real environment.

The AUX was installed and commissioned after the IM and the DF. First test aimed to establish dataflow throughout the IM-DF-AUX. Data was fed from the SCT, then it was clustered in the IM and forwarded to the DF. The DF sent it to the AUX, and the AUX sent the data to the ATLAS ROS without any processing (pass-through mode). A schematic view of this test is shown in Fig. 102. With this test, the AUX was installed and integrated at the USA15. In addition, the FTK was integrated with the ATLAS ROS and data was recorded for the first time.

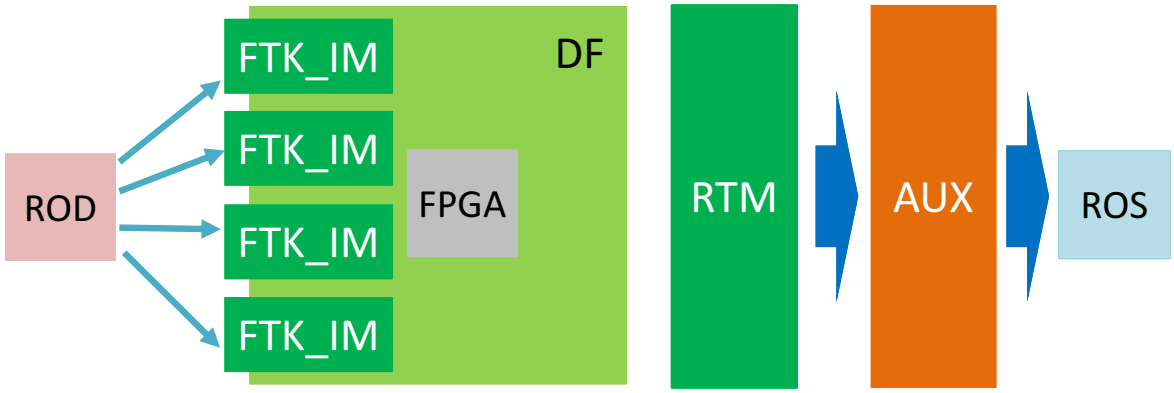

Figure 102: A schematic view of setup of the IM-DF-AUX integration at the USA15.

The downstream boards as the AMB, SSB, and FLIC has been installed and commissioned as similar way as the case of the Lab4. At first, communication was established with simple data, and then increasing loads. Several tests have been performed and firmware has been developed with the feedback of tests. Finally, the FTK succeeded to achieve dataflow throughout the entire FTK system and to send data to the ATLAS ROS. A photo when it was accomplished is shown in Fig. 103.

Figure 103: A photo when the entire FTK processing worked for real data and the data was recorded at the ATLAS ROS.

Toward the start of regular operation, there are several issues to be resolved. One major issue is that dataflow of the entire system is not stable at high input rate. Flow control and event synchronization among many combination of boards needs to be improved. This requires to perform tests, accumulate feedback, and develop firmware to be more stable. Another issue is to increase boards and achieve higher parallelization. In terms of performance, it is necessary to evaluate processing time with real hardware and to validate the FTK can work in time for trigger system. In addition trigger chain using the FTK information should be built, and it needs to be validated at the early stage of real operation.

7 Conclusions

The LHC will operate at higher instantaneous luminosity where there are a greater number of simultaneous interactions per bunch crossing, which makes it difficult to separate signals from the background. One of the major challenges for high energy physics experiments is to maintain the phase space at the data-taking stage in such an environment by fully utilizing the finite resources. The FTK is a newly installed hardware system at the ATLAS experiment that can reconstruct all tracks in an event within the time-limited trigger system. The HLT can access the track information at the beginning of the processing and can have additional time to perform more sophisticated algorithms. For example, all the vertices in an entire event can be reconstructed, which is critical to distinguish hard scatter events and pile-ups. In addition, better particle identification can be implemented at the HLT, which enables to apply a lower threshold for the transverse momentum and to maintain a larger phase space at the trigger stage.

Simulation of the FTK system has been developed not only to evaluate the track reconstruction performance but also to emulate the behavior of the hardware system. It is validated that the FTK can reconstruct tracks with high efficiency and high resolution even in the high luminosity environment. In addition, the simulation shows that the FTK is a sufficiently flexible system against changes in the operation environment, such as the LHC beam spot shift and detector degradation. The FTK processing time was also evaluated, and it is validated that the FTK can process events with a high pile-up condition in time for trigger decision at the HLT. Furthermore, the FTK has an option that limits the maximum number of fits per pattern to optimize the processing speed with little compromise in the tracking performance.

Hardware development has been performed. The FTK consists of several electric boards and each of them is required to have functions optimized for the FTK, therefore the boards need to be designed, produced, and tested in-house. The Waseda team is responsible for the IM boards, and in the end 80 IMs were mass produced. All of the boards passed the quality control tests, which means they were validated to have sufficient quality to be installed into the ATLAS. Other sub-boards of the FTK system have also been produced and are being installed into the ATLAS.

Integration of the entire FTK system and installation of the FTK system into the ATLAS have been performed. It is challenging to achieve a stable dataflow since the flow control needs to be implemented correctly with the combination of multiple electric boards. The system has been established one board at a time, and finally entire system was integrated. The first output of the entire FTK system with real data was sent and recorded at the ATLAS ROS at the end of 2016. To start operation, the remaining issues include the establishment of a stable dataflow with a full system of higher parallelization, and trigger chain building.

The operation of the FTK will start in 2018 with a half number of the processor units. It will be upgraded during the LHC shutdown from 2019 to 2020 and start operation with the full system in 2021. The FTK will run until 2023, and then next-generation tracking systems will be developed.

8 Future Prospect

The LHC plans to run with upgraded luminosity of 5.0×10^{34} cm^{-2}s^{-1}, or in extreme case, 7.5×10^{34} cm^{-2}s^{-1}, from 2026 which is called High Luminosity LHC (HL-LHC) [75]. The average number of interactions per bunch crossing will be 140 in the former case, and it will be 200 in the latter case. An event display for $t\bar{t}$ events with 200 simultaneous interactions is shown in Fig. 104. To take fully advantage of the accelerator upgrade and to cope with higher detector occupancy, detectors and trigger systems will be also upgraded drastically. Theses upgrade will allow more feasibility to search for physics Beyond the Standard Model (BSM) as well as rare decay of the SM. For example, search for new particles with masses above 1 TeV will be more accessible, and Higgs boson decaying to a pair of muons with more than 5σ significance will be reachable.

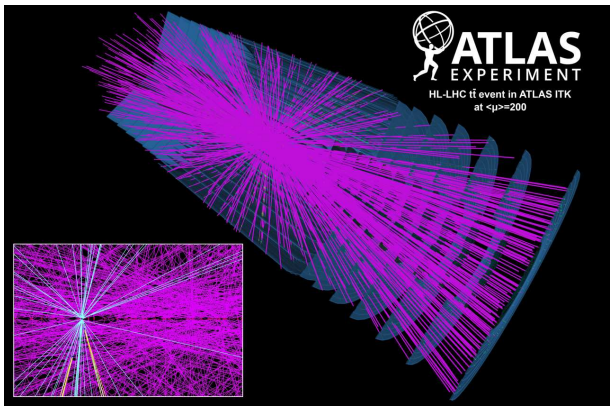

Figure 104: An event display of simulation for $t\bar{t}$ events with 200 simultaneous interactions [76].

Schedule for the upgrade of the LHC is shown in Fig. 105. The upgrade of accelerators and detectors will be performed during two major shutdowns of the LHC. One is called Long Shutdown2 (LS2) which is planned to be taken place in 2019-2020, the other is called Long Shutdown3 (LS3) which is planned to be performed in 2024-2026. The integrated luminosity of 300 fb^{-1} will be accumulated at the end of Run3 (2024), and 3000 fb^{-1} will be recorded by the HL-LHC.

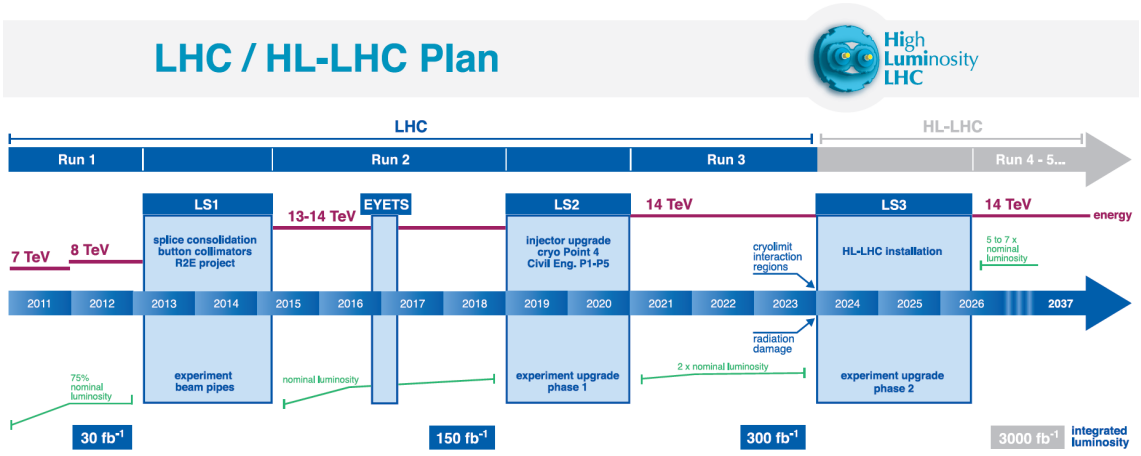

Figure 105: Plan for the High Luminosity LHC upgrade.

About the upgrade of inner detector system, the TRT will not be usable due to high occupancy caused by high luminosity. The Pixel detector and the SCT will be damaged by high radiation dose during long term operation, and both sensors and readout will be deteriorated. Thus all the inner detectors will be replaced with new tracker system which is called ATLAS Inner Tracker (ITk) [77, 78]. The ITk will utilize all silicon-based devices to measure charged particle tracks with high precision. The inner layers will be pixel detectors with the size of 50×50 μm^2 and 25×100 μm^2 are planned, consisting of five barrel layers. The outer region will be covered by 74.5 μm pitch double-sided strip detectors which are attached with 40 mrad stereo angle. They will consist of four barrel layers and six endcap layers per side. The number of channels are designed to be about 600 million for pixel, and 70 million for strip. They are readout at a rate of 1 MHz. A visualization of ITk system in simulation is shown in Fig. 106 left, and one of the options of the planned ITk layout is shown in Fig. 106 right.

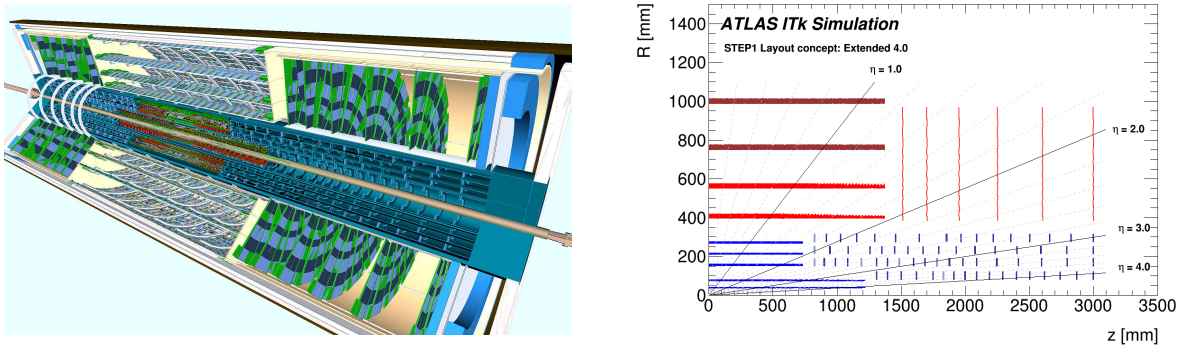

Figure 106: A visualization of the ITk in simulation (left) and one of the options of the planned ITk layout (right). In the layout, the blue ones represent pixel detectors and the red ones represent strip detectors.

To cope with huge amount of data with large input rate provided by upgraded detector systems, trigger system will be also upgraded. Two stages of custom-hardware triggers which are called Level-0 Trigger (L0) and Level-1 Trigger (L1) will be implemented. A schematic view of the planned L0 and L1 trigger system at the HL-LHC is shown in Fig. 107. They allow for data streaming off-detector either after the L0 or at the full 40 MHz bunch crossing rate. The FPGAs will be implemented to process data fast and apply sophisticated trigger algorithms, instead of ASICs which are used currently. The L0 trigger will make trigger decision based on calorimeter and muon detector information, and forms RoIs. The L0 trigger accept rate will be 500 kHz to 1 MHz, and the latency is 6 μs. the L1 trigger will utilize more robust calorimeter information, muon detector information, and also the ITk information within RoIs for decision. The Level-1 Track Trigger (L1Track) [79, 80] will perform track reconstruction for the coverage of the ITk, find tracks with p_T above pre-defined threshold (e.g. 4 GeV or 8 GeV) within RoIs defined by the L0 trigger. The L1 trigger accept rate will be 400 kHz and the latency is 60 μs. The HLT is software-based trigger which will consist of the upgraded Fast TracKer system (FTK++) and the Event Filter (EF). The overview of the data acquisition and the EF system is shown in Fig. 108.

The working principle of track reconstruction by the L1Track and the FTK++ will be similar to that of the FTK, while the goal of the two is slightly different. The L1Track will work for all the RoIs defined by the L0 calorimeter trigger and the muon trigger, and reconstruct tracks within the RoIs based on the ITk information. At first step, the hit information is clustered into “Super-Strips (SS)” and pattern matching is performed by the AM chips based on the SS information. And then FPGA-based linear approximation by pre-defined fit constants are performed to calculate track parameters. To cover full acceptance of the tracker system, about 3.2 billion patterns will be necessary. The number of patterns are more than the case

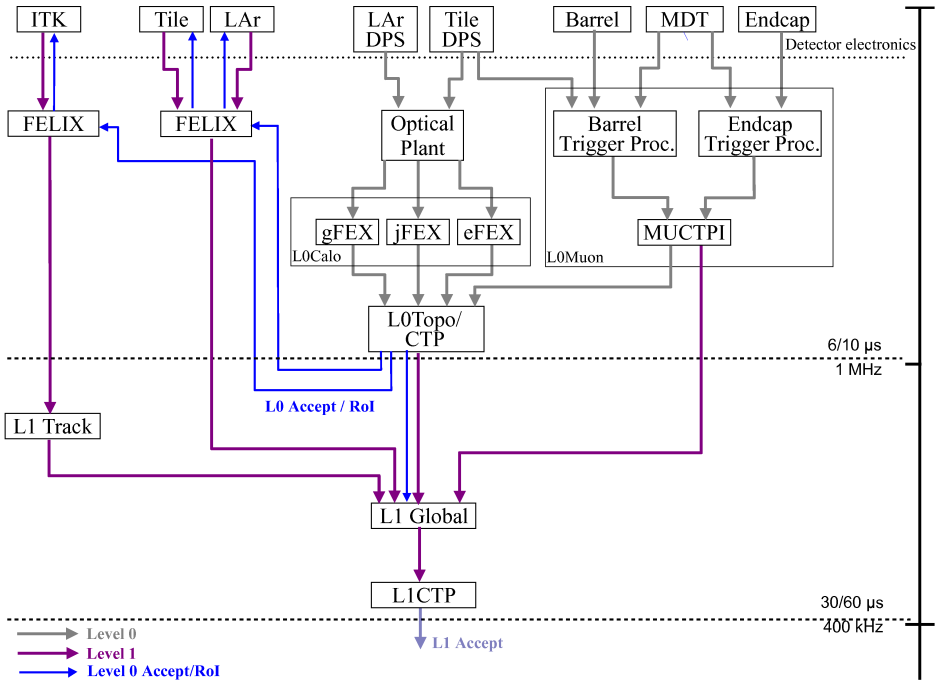

Figure 107: A schematic view of the ATLAS L0 and L1 trigger system planned at the HL-LHC.

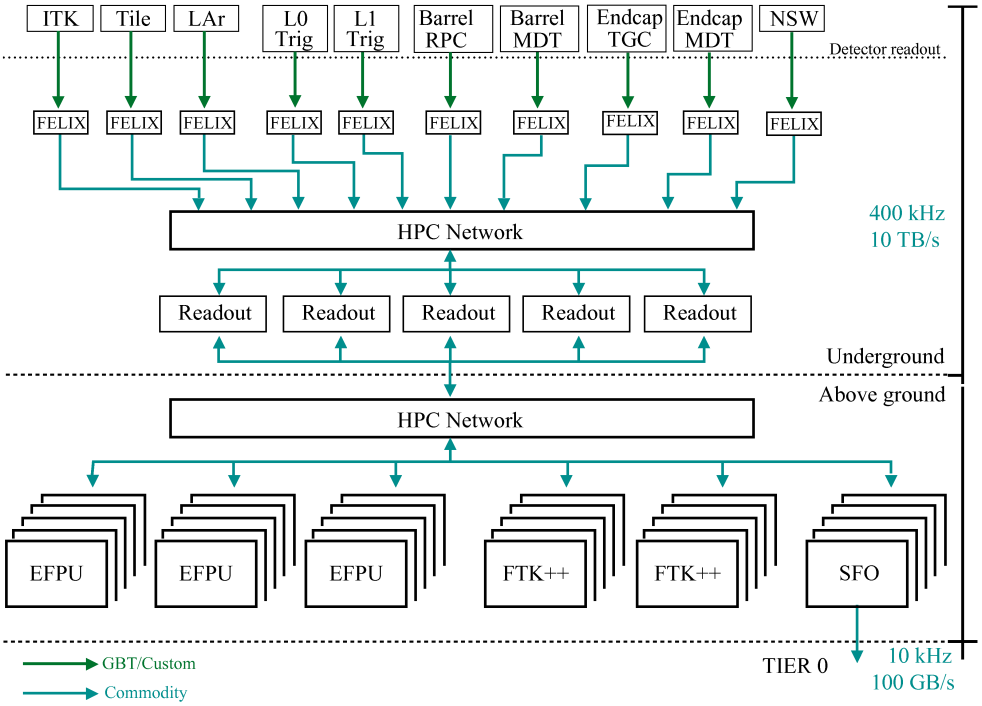

Figure 108: Overview of the Phase-II Data Acquisition and the EF systems showing the flow of data from the detector readout to the EF computing farm.

of the FTK (1 billion patterns) since the number of input channels increases by detector upgrade and also detector occupancy is high due to higher pile-up condition. In contrast, the number is not so large since p_T threshold for a track is higher (e.g. 4 GeV or 8 GeV) than the case of FTK (1 GeV). Full ATCA-based highly modular system is planned which consists of mother board (PulsarIIb) for data sharing and two kinds of mezzanine cards, pattern matching mezzanine and track fitting mezzanine. The other option from the AM chip approach is considered, which is "self-seeded scheme" aiming to reconstruct high p_T tracks (e.g. $>$ 10 GeV). The scheme will reject hits from low-p_T tracks by the size of clusters or "straightness" of tracks by connecting hits for several layers. The overall tracking performance is better for the AM chip approach than that of self-seeded scheme so that current primary option is the AM chip approach. The advantage of self-seeded scheme is that it is not necessary to use the AM chips and full FPGA-based system can be built. The latency required for the L1Track is less than 30 μs since the overall L1 latency is 60 μs thus it should finish at the early stage of the L1 decision.

The purpose to use tracks from the L1Track will be to confirm the L1 trigger decision and reduce background rate by, for example, matching calorimeter or muon object with tracks, and/or by requiring track-based isolation. For this purpose, it is enough to reconstruct high p_T tracks only within RoIs. The track reconstruction efficiency is studied to be above 95% with respect to offline reconstructed tracks and the resolution of 10 mm on the track position along the beam axis. The tracks reconstructed by the L1Track will be used for the input for the trigger selection (L1Global).

The FTK++ will be designed to have similar algorithm to find tracks as the L1Track. Though the L1Track will find only tracks within RoIs, the FTK++ will perform track reconstruction for the entire detector region. It will allow longer latency and lower input rate so that tracks with lower transverse momenta (e.g. 1 GeV) will be covered. The FTK++ system is planned to have 13 billion patterns to reconstruct all the tracks for the entire coverage of the tracker system. To achieve further performance improvement, the hit filtering by cluster size and higher p_T cut is also considered. Track fitting speed is expected to be about four times faster than the FTK with using the newer FPGA. The hardware used for the L1Track and the FTK++ will be identical, the major difference between the two is the required latency. Parameters and requirements for the L1Track and the FTK++ are summarized in Tab. 7.

Parameters	L1Track	FTK++
Latency [μs]	30	not fixed latency
Number of patterns	3.2 billion	13 billion
p_T threshold [GeV]	4	1
Work for	only RoIs	entire detector region

Table 7: Parameters and requirements for the L1Track and the FTK++.

Abbreviations

AM
Associative Memory which is a part of the FTK system, performs pattern matching.

ATCA
Advanced Telecommunications Computing Architecture which is a shelf holding the DFs and the FLICs.

ATLAS
A Toroidal LHC ApparatuS which is the particle detector with an ensemble of sub-detectors located at one of the interaction points of the LHC.

AUX
AUXiliary Card which is a part of the FTK system, maps clusters to SSs, feeds the SSs to the AM, receives roads from the AM, retrieves clusters, performs track fitting and χ^2 cut, and sends the cluster information and χ^2 to the SSB for tracks passing the χ^2 cut.

BOC
Back of Crate which is readout board for the IBL and a part of the Pixel detectors utilized instead of D-HOLA.

CSC
Cathode Strip Chamber which is a precision chamber consisting of the Muon Spectrometer.

DC-bit
Don't Care bit which is assigned to consist variable resolution SS.

DF
Data Formatter which is a part of the FTK system, shares clusters from the IM, and distributes the clusters to the appropriate parallel processor.

D-HOLA
Dual-output HOLA which is mounted on the ROD and has two ports to feed data.

DO
Data Organizer which is a function of the AUX, maps clusters to the SSs, and retrieves the clusters to the SSs.

FCal
Forward Calorimeter which is a part of the hadron calorimeter placed at the forward region.

FIFO
First-In First-Out buffer which can store data of pre-defined size at the maximum and feed them with coming order.

FILAR
Four Input LINK for ATLAS Readout is a PCI express card which receives data with same protocol as the ATLAS ROS. Four links can be fed simultaneously at maximum.

FLIC
: FTK to Level-2 Interface Card which is a part of the FTK system, transforms data of FTK format to ATLAS global format, and sends FTK output to the ATLAS ROS.

FTK
: Fast TracKer which is a combination of electric circuit boards installed between the L1 trigger and the HLT, and reconstructs all the tracks in an event.

HEC
: Hadronic Endcap Calorimeter which is a part of the hadron calorimeter placed at the endcap.

HLT
: High Level Trigger which is the second stage of the current ATLAS trigger system.

HW
: Hit Warrior which is a function of the AUX, removes duplicated tracks.

IBL
: Insertable B-Layer which is a part of the inner detector inserted in the inner most part of the ATLAS detector.

IM
: Input Mezzanine which is a part of the FTK system, receives hit information, clusters the hits, and forwards it to the DF.

S6 IM
: The IM which holds Spartan-6 FPGA.

A7 IM
: The IM which holds Artix-7 FPGA.

ISim
: ISE Simulator which is a tool provided by Xilinx to simulate behaviour of signals in the firmware.

L1
: Level-1 trigger which is the first stage of the current ATLAS trigger system.

LAMB
: Local Associative Memory Board which is mounted on AM board, holds 16 AM chips per board.

LHC
: Large Hadron Collider which accelerates protons and collides them at a center of mass energy of TeV scale.

LIFO
: Last-In First-Out buffer which can store data of pre-defined size at the maximum and feed them with reverse order.

MDT
: Monitored Drift Tube which is a precision chamber consisting of the Muon Spectrometer.

QUEST
: QUad SLINK Transceiver which is a PCI card and sends data by same protocol as the ATLAS ROD. Four links can be fed simultaneously at maximum.

ROD
ReadOut Driver which formats data from the detectors and feeds them to the trigger system.

RoI
Region of Interest which is a region where an interesting object is found by the L1. The HLT processing works around the RoI.

ROS
ReadOut System which receives data from L1 trigger and feeds them to the HLT.

RPC
Resistive Plate Chamber which is a trigger chamber consisting of the Muon Spectrometer.

RTM
Rear Transition Module which is inserted at the back plane and has ports of optical fibers.

SCT
Semi-Conductor Tracker which is a part of the inner detector consisting of silicon strips.

SS
Super Strip which is detector hit information with reduced resolution.

SSB
Second Stage Board which is a part of the FTK system, performs second stage track fitting with full 12 layer information.

TF
Track Fitter which is a function of the AUX, performs track fitting.

TGC
Thin Gap Chamber which is a trigger chamber consisting of the Muon Spectrometer.

TILAR
Two Input LINK for ATLAS Readout is a PCI express card which receives data with same protocol as the ATLAS ROS. Two links can be fed simultaneously at maximum.

TRT
Transition Radiation Tracker which is a part of the inner detector consisting of straw tubes.

References

[1] S. L. Glashow, *Partial-symmetries of weak interactions*, Nucl. Phys. **22** (1961) 579.

[2] S. Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (1967) 1264.

[3] A. Salam, *Weak and electromagnetic interactions*, in Elementary particle theory: relativistic groups and analyticity (1968) 367, Proceedings of the eighth Nobel symposium.

[4] G.'t Hooft and M. Veltman, *Regularization and Renormalization of Gauge Fields*, Nucl. Phys. **B44** (1972) 189.

[5] F. Englert and R. Brout, *Broken symmetry and the mass of gauge vector mesons*, Phys. Rev. Lett. **13** (1964) 321.

[6] P. W. Higgs, *Broken symmetries, massless particles and gauge fields*, Phys. Lett. **12** (1964) 132.

[7] P. W. Higgs, *Broken symmetries and the masses of gauge bosons*, Phys. Rev. Lett. **13** (1964) 508.

[8] P. W. Higgs, *Spontaneous symmetry breakdown without massless bosons*, Phys. Rev. **145** (1966) 1156.

[9] 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. **B716** (2012) 1.

[10] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. **B716** (2012) 30.

[11] http://cms.web.cern.ch/news/reconstructing-multitude-particle-tracks-within-cms.

[12] ATLAS Collaboration, *Vertex Reconstruction Performance of the ATLAS Detector at* $\sqrt{s} = 13$ *TeV*, ATLAS-NOTE, ATL-PHYS-PUB-2015-026 (2015).

[13] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at* $\sqrt{s} = 8$ *TeV using the ATLAS detector*, Eur. Phys. J. **C76** (2016) 581.

[14] W. Waltenberger, R. Fruhwirth and P. Vanlaer, *Adaptive vertex fitting*, JOURNAL OF PHYSICS G: NUCLEAR AND PARTICLE PHYSICS **34** (2007) N34–N356.

[15] D. Emeliyanov, *A Fast Vertex Fitting Algorithm for ATLAS Level 2 Trigger*, Proceedings of Science **058** (2007).

[16] ATLAS Collaboration, *Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment*, Eur. Phys. J. **C76** (2016) 295.

[17] ATLAS Collaboration, *Expected performance of the ATLAS b-tagging algorithms in Run-2*, ATLAS NOTE, ATL-PHYS-PUB-2015-022 (2015).

[18] A. G. Barbaro, F. Margaroli and I. Volobouev, *Precision measurements of the top quark mass from the Tevatron in the pre-LHC era*, Rept. Prog. Phys. **75** (2012) 056201.

[19] L. Evans and P. Bryant, *LHC Machine*, JINST **3** (2008) S08001.

[20] LEP Collaboration, *LEP Design Report: Vol.2. The LEP Main Ring*, Technical Report, CERN (1984).

[21] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[22] CMS Collaboration, *The CMS experiment at the CERN LHC*, JINST **3** (2008) S08004.

[23] LHCb Collaboration, *The LHCb detector at the LHC*, JINST **3** (2008) S08005.

[24] ALICE Collaboration, *The ALICE experiment at the CERN LHC*, JINST **3** (2008) S08002.

[25] A. Yamamoto et al., *Design and development of the ATLAS central solenoid magnet*, IEEE Transactions on Applied Superconductivity **9** (1999) 852.

[26] T-K. Herman, *The superconducting magnet system for the ATLAS detector at CERN*, IEEE Transactions on Applied Superconductivity **9** (1999) 841.

[27] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report 1*, CERN-LHCC-97-016, ATLAS-TDR-4 (1997).

[28] ATLAS Collaboration, *ATLAS inner detector: Technical Design Report 2*, CERN-LHCC-97-017, ATLAS-TDR-5 (1997).

[29] ATLAS Collaboration, *The ATLAS Inner Detector commissioning and calibration*, Eur. Phys. J. **C70** (2010) 787–821.

[30] K. Potamianos, *The upgraded Pixel detector and the commissioning of the Inner Detector tracking of the ATLAS experiment for Run-2 at the Large Hadron Collider*, Proceedings of Science **EPS-HEP2015** (2015) 261.

[31] ATLAS Collaboration, *ATLAS Insertable B-Layer: Technical Design Report*, CERN-LHCC-2010-013, ATLAS-TDR-19 (2010).

[32] A. Miucci, *The ATLAS Insertable B-Layer project*, JINST **9** (2014) C02018.

[33] J. Weingarten, *ATLAS IBL sensor qualification*, JINST **7** (2012) C01039.

[34] ATLAS Pixel Collaboration, *ATLAS pixel detector electronics and sensors*, JINST **3** (2008) P07007.

[35] ATLAS SCT Collaboration, *The barrel modules of the ATLAS semiconductor tracker*, Nucl. Instrum. Meth. **A568** (2006) 642.

[36] ATLAS SCT Collaboration, *The ATLAS semiconductor tracker end-cap module*, Nucl. Instrum. Meth. **A575** (2007) 353.

[37] F. Campabadal et al., *Design and performance of the ABCD3TA ASIC for readout of silicon strip detectors in the ATLAS semiconductor tracker*, Nucl. Instrum. Meth. **A552** (2005) 292.

[38] A. Abdesselam et al., *The Data acquisition and calibration system for the ATLAS semiconductor tracker*, JINST **3** (2008) P01003.

[39] ATLAS TRT Collaboration, *The ATLAS Transition Radiation Tracker (TRT) proportional drift tube: design and performance*, JINST **3** (2008) P02013.

[40] ATLAS TRT Collaboration, *The ATLAS TRT Barrel Detector*, JINST **3** (2008) P02014.

[41] ATLAS TRT Collaboration, *The ATLAS TRT end-cap detectors*, JINST **3** (2008) P10003.

[42] ATLAS Collaboration, *Performance of the ATLAS Transition Radiation Tracker in Run 1 of the LHC: tracker properties*, JINST **12** (2017) P05002.

[43] ATLAS Collaboration, *ATLAS calorimeter performance: Technical Design Report*, CERN-LHCC-96-40, ATLAS-TDR-1 (1996).

[44] ATLAS Collaboration, *ATLAS liquid-argon calorimeter : Technical Design Report*, CERN-LHCC-96-41, ATLAS-TDR-2 (1996).

[45] ATLAS Collaboration, *ATLAS tile calorimeter : Technical Design Report*, CERN-LHCC-96-42, ATLAS-TDR-3 (1996).

[46] D. M. Gingrich et al., *Construction, assembly and testing of the ATLAS hadronic end-cap calorimeter*, JINST **2** (2007) P05005.

[47] A. Artamonov et al., *The ATLAS Forward Calorimeter*, JINST **3** (2008) P02010.

[48] ATLAS Collaboration, *ATLAS muon spectrometer : Technical Design Report*, CERN-LHCC-97-22, ATLAS-TDR-10 (1997).

[49] R. Kwee, *Minimum Bias Triggers at ATLAS, LHC*, 34th International Conference on High Energy Physics (2008).

[50] http://project-atlas-lucid.web.cern.ch/project-atlas-lucid/taskforce/main_10.html.

[51] Y. Nakahama, *The ATLAS Trigger System: Ready for Run-2*, Journal of Physics: Conference Series **664** (2015) 082037.

[52] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. **C77** (2017) 317.

[53] ATLAS Collaboration, *Performance of the ATLAS Trigger System in 2010*, Eur. Phys. J. **C72** (2012) 1849.

[54] ATLAS Collaboration, *Technical Design Report for the Phase-I Upgrade of the ATLAS TDAQ System*, CERN-LHCC-2013-018, ATLAS-TDR-023 (2013).

[55] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TriggerOperationPublicResults.

[56] https://twiki.cern.ch/twiki/bin/view/AtlasPublic/HLTTrackingPublicResults.

[57] ATLAS Collaboration, *Fast TracKer (FTK) Technical Design Report*, CERN-LHCC-2013-007, ATLAS-TDR-21 (2013).

[58] S. Belforte et al., *The CDF trigger Silicon Vertex Tracker (SVT)*, Nuclear Science Symposium and Medical Imaging Conference, 1994., 1994 IEEE Conference Record **2** (1994) 824.

[59] J. Adelman et al., *The Silicon Vertex Trigger upgrade at CDF*, Nucl. Instrum. Meth. **A572** (2007) 361.

[60] A. Annovi et al., *AM06: the Associative Memory chip for the Fast TracKer in the upgraded ATLAS detector*, JINST **12** (2017) C04013.

[61] C. L. Sotiropoulou et al., *The Associative Memory System Infrastructures for the ATLAS Fast Tracker*, IEEE Transactions on Nuclear Science **64** (2017) 1248.

[62] T. Iizawa, *Fast tracker performance using the new "variable resolution associative memory" for ATLAS*, 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC) (2012) 1392.

[63] N. Kimura et al., *A highly parallel FPGA implementation of a 2D-clustering algorithm for the ATLAS Fast TracKer (FTK) processor*, 2014 19th IEEE-NPSS Real Time Conference (2014) 1.

[64] Y. Okumura et al., *ATCA-based ATLAS FTK input interface system*, JINST **10** (2015) C04032.

[65] C. G. Larrea et al., *IPbus: a flexible Ethernet-based control system for xTCA hardware*, JINST **10** (2015) C02019.

[66] V. Cavaliere et al., *Design of a hardware track finder (Fast Tracker) for the ATLAS trigger*, JINST **11** (2016) C02056.

[67] R. Wang et al., *The FTK to Level-2 Interface Card (FLIC) for the ATLAS experiment*, 2015 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC) (2015) 1.

[68] J. Adelman et al., *ATLAS FTK Challenge: Simulation of a Billion-fold Hardware Parallelism*, Procedia Computer Science **66** (2015) 540.

[69] S. Agostinelli et. al., *Geant4 - a simulation toolkit*, Nucl. Instrum. Meth. **A506** (2003) 250.

[70] E. van der Bij et al., *S-LINK, a data link interface specification for the LHC era*, IEEE Trans. Nucl. Sci. **44** (1997) 398.

[71] http://hsi.web.cern.ch/hsi/s-link/devices/hola/hw_spec.pdf.

[72] J. Joseph et al., *ATLAS Silicon ReadOut Driver (ROD) Users Manual*, (2009).

[73] http://www.robot-electronics.co.uk/i2c-tutorial.

[74] M. Wensing, *Testing and firmware development for the ATLAS IBL BOC prototype*, JINST **7** (2012) C12027.

[75] ATLAS Collaboration, *ATLAS Phase-II upgrade Scoping Document*, CERN-LHCC-2015-020, LHCC-G-166 (2013).

[76] http://cerncourier.com/cws/article/cern/63704.

[77] ATLAS collaboration, *Technical Design Report for the ATLAS Inner Tracker Pixel Detector*, CERN-LHCC-2017-021, ATLAS-TDR-30 (2017).

[78] ATLAS Collaboration, *Technical Design Report for the ATLAS Inner Tracker Strip Detector*, CERN-LHCC-2017-005, ATLAS-TDR-25 (2017).

[79] A. Cerri, *L1Track: A fast Level 1 track trigger for the ATLAS high luminosity upgrade*, Nucl. Instrum. Meth. **A824** (2016) 263.

[80] A. De Santo, *Towards a Level-1 Tracking Trigger for the ATLAS Experiment*, Nucl. Part. Phys. Proc. **273-275** (2016) 1059.

Publication List

Selected Publications

1. T. Iizawa et al., the ATLAS Collaboration, “Performance of the ATLAS Trigger System in 2015”, The European Physical Journal C, 77:317, May, 2017.
2. T. Iizawa et al., the ATLAS Collaboration, “Operation and Performance of the ATLAS semiconductor tracker”, JINST9, P08009, 2014.
3. T. Iizawa et al., the ATLAS Collaboration, “The ATLAS FTK System: how to improve the physics potential with a tracking trigger”, Nuclear and Particle Physics Proceedings, Vol. 273-275, pp. 2521-2523, ELSEVIER, April-June. 2016.
4. T. Iizawa et al., the ATLAS Collaboration, “Fast Tracker Performance Using the New Variable Resolution Associative Memory for Atlas”, 2012 IEEE NSS/MIC Conference Record, pp. 1392-1395, IEEE, July. 2013.
5. J. Anderson, T. Iizawa, et al., “FTK: a Fast Track Trigger for ATLAS”, JINST, IOP Publishing, October. 2012.
6. T. Iizawa et al., the ATLAS Collaboration, “ATLAS Fast TracKer Technical Design Report” ATLAS-TDR-021, June. 2013.

Article

1. T. Iizawa N. Kimura, K. Yorita, “Development and construction of the Fast TracKer (FTK) at the ATLAS experiment”, High Energy News, Vol. 35, No. 3, pp. 146-154, Japan Association of High Energy Physicists, October-December. 2016.

Conference and Workshop Presentations

1. “The ATLAS Fast Tracker System”, TWEPP 2017 Topical Workshop on Electronics for Particle Physics, Santa Cruz, California, U.S., September, 2017.
2. “Installation and commissioning of the Fast TracKer (FTK) at the LHC-ATLAS experiment”, The 72nd annual conference of JPS, Osaka, Japan, March 2017.
3. “Input Mezzanine Card for the Fast Tracker at ATLAS”, 2016 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), Strasbourg, France, October. 31, 2016 (Poster).
4. “The performance evaluation of the Fast Tracker at the LHC-ATLAS experiment”, The 71st annual conference of JPS, Miyagi, Japan, March 2016.
5. “Phase-1 upgrade: FTK”, Meeting of the Research of the New Academic Domain, Tera-Scale 2015, Tokyo, Japan, December 2015 (Invited talk).
6. “Fast Tracker(FTK): A Hardware Track Finder for the ATLAS Trigger”, 4th International Conference on Modern Circuits and Systems Technologies (MOCAST), Thessaloniki, Greece, May. 2015 (Poster).
7. “Optimization of the MET trigger for the new particle search at the LHC-ATLAS Run2 experiment”, The 70th annual conference of JPS, Tokyo, Japan, March 2015.
8. “ATLAS FTK: Fast Tracker”, Vertex2014: The 23rd International Conference on Vertex Detectors, Doksy, The Czech Republic, September. 2014.
9. “The ATLAS FTK system: how to improve the physics potential with a tracking trigger”, 37th International Conference on High Energy Physics (ICHEP), Valencia, Spain, July. 2014 (Poster).
10. “The design and development of the input receiver module of the FTK and evaluation of its performance”, The 69th annual conference of JPS, Kanagawa, Japan, March 2014.
11. “The development and construction of the Fast Tracking trigger system (1)”, The 68th annual conference of JPS, Hiroshima, Japan, March 2013.
12. “Fast Tracker Performance Using the New Variable Resolution Associative Memory for Atlas”, 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC), Anaheim, California, U.S., October. 2012 (Poster).
13. “Optimization of the Fast Tracker at the ATLAS experiment”, The annual conference of JPS in autumn 2012, Kyoto, Japan, September 2012.

Seminar

1. “The frontier of the research of the origin of the universe”, Reitaku junior and senior high school, Chiba, Japan, May, 2017.
2. “Construction of the Fast Tracker and its working principle at the ATLAS experiment”, Fundamental Physics Laboratory Seminar at Nagoya University, Aichi, Japan, April, 2016.