

## Chapter 7

# Trigger

The LHCb experiment plans to operate at an average luminosity of  $2 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$ , much lower than the maximum design luminosity of the LHC, reducing the radiation damage to the detectors and electronics. Furthermore, the number of interactions per bunch crossing is dominated by single interactions, which facilitates the triggering and reconstruction by assuring low channel occupancy. Due to the LHC bunch structure and low luminosity, the crossing frequency with interactions visible<sup>1</sup> by the spectrometer is about 10 MHz, which has to be reduced by the trigger to about 2 kHz, at which rate the events are written to storage for further offline analysis. This reduction is achieved in two trigger levels [214] as shown in figure 7.1: Level-0 (L0) and the High Level Trigger (HLT). The L0 trigger is implemented using custom made electronics, operating synchronously with the 40 MHz bunch crossing frequency, while the HLT is executed asynchronously on a processor farm, using commercially available equipment. At a luminosity of  $2 \times 10^{32} \text{ cm}^{-2} \text{ s}^{-1}$  the bunch crossings with visible pp interactions are expected to contain a rate of about 100 kHz of  $b\bar{b}$ -pairs. However, only about 15% of these events will include at least one B meson with all its decay products contained in the spectrometer acceptance. Furthermore the branching ratios of interesting B meson decays used to study for instance CP violation are typically less than  $10^{-3}$ . The offline analysis uses event selections based on the masses of the B mesons, their lifetimes and other stringent cuts to enhance the signal over background. For the best overall performance the trigger was therefore optimised to achieve the highest efficiency for the events selected in the offline analysis, while rejecting uninteresting background events as strongly as possible.

The purpose of the L0 trigger is to reduce the LHC beam crossing rate of 40 MHz to the rate of 1 MHz with which the entire detector can be read out. Due to their large mass, B mesons decays often produce particles with large transverse momentum ( $p_T$ ) and energy ( $E_T$ ) respectively. The Level-0 trigger attempts to reconstruct:

- the highest  $E_T$  hadron, electron and photon clusters in the calorimeters,
- the two highest  $p_T$  muons in the muon chambers.

In addition, a pile-up system in the VELO estimates the number of primary pp interactions in each bunch crossing. The calorimeters calculate the total observed energy and an estimate for the num-

<sup>1</sup> An interaction is defined to be visible if it produces at least two charged particles with sufficient hits in the VELO and T1–T3 to allow them to be reconstructible.

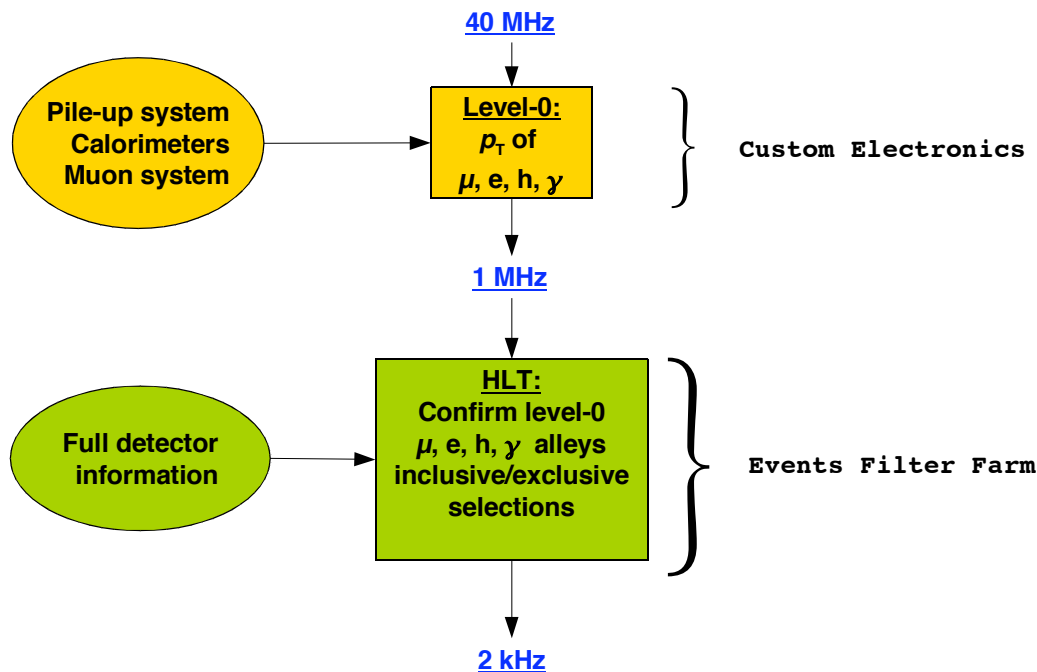


Figure 7.1: Scheme of the LHCb trigger.

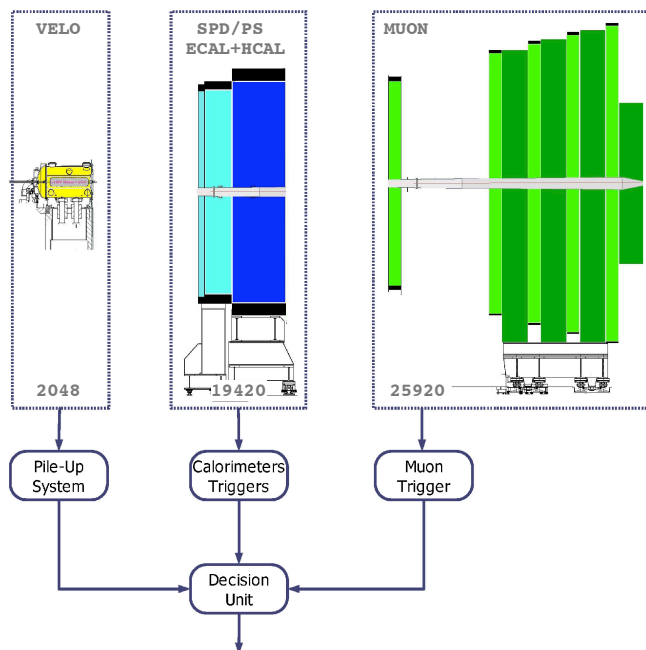
ber of tracks, based on the number of hits in the SPD. With the help of these global quantities events may be rejected, which would otherwise be triggered due to large combinatorics, and would occupy a disproportionate fraction of the data-flow bandwidth or available processing power in the HLT.

A Level-0 Decision Unit (DU) collects all the information and derives the final Level-0 trigger decision for each bunch crossing. It allows for overlapping of several trigger conditions and for prescaling.

The L0 trigger system is fully synchronous with the 40 MHz bunch crossing signal of the LHC. The latencies are fixed and depend neither on the occupancy nor on the bunch crossing history. All Level-0 electronics is implemented in fully custom-designed boards which make use of parallelism and pipelining to do the necessary calculations with sufficient speed.

In order to be able to reduce the event rate from 1 MHz down to 2 kHz, the HLT makes use of the full event data. The generic HLT algorithms refine candidates found by the Level-0 trigger and divide them into independent *alleys* (see section 7.2). The alleys to be followed are selected from the Level-0 decision. The alley selections are based on the principle of confirming a previous trigger stage by requiring the candidate tracks to be reconstructed in the VELO and/or the T-stations. Requiring candidate tracks with a combination of high  $p_T$  and/or large impact parameter reduces the rate to about 30 kHz. At this rate interesting final states are selected using inclusive and exclusive criteria.

Generally speaking, selection cuts are relaxed compared to the offline analysis, in order to be able to study the sensitivity of the selections and to profit from refinements due to improved calibration constants. A large fraction of the output bandwidth is devoted to calibration and monitoring. In order to monitor trigger efficiencies and systematic uncertainties both trigger levels can be emulated fully on stored data.



**Figure 7.2:** Overview of the Level-0 trigger. Every 25 ns the pile-up system receives 2048 channels from the pile-up detector, the Level-0 calorimeters 19420 channels from the scintillating pad detector, preshower, electromagnetic and hadronic calorimeters while the Level-0 muon handles 25920 logical channels from the muon detector.

## 7.1 Level 0 trigger

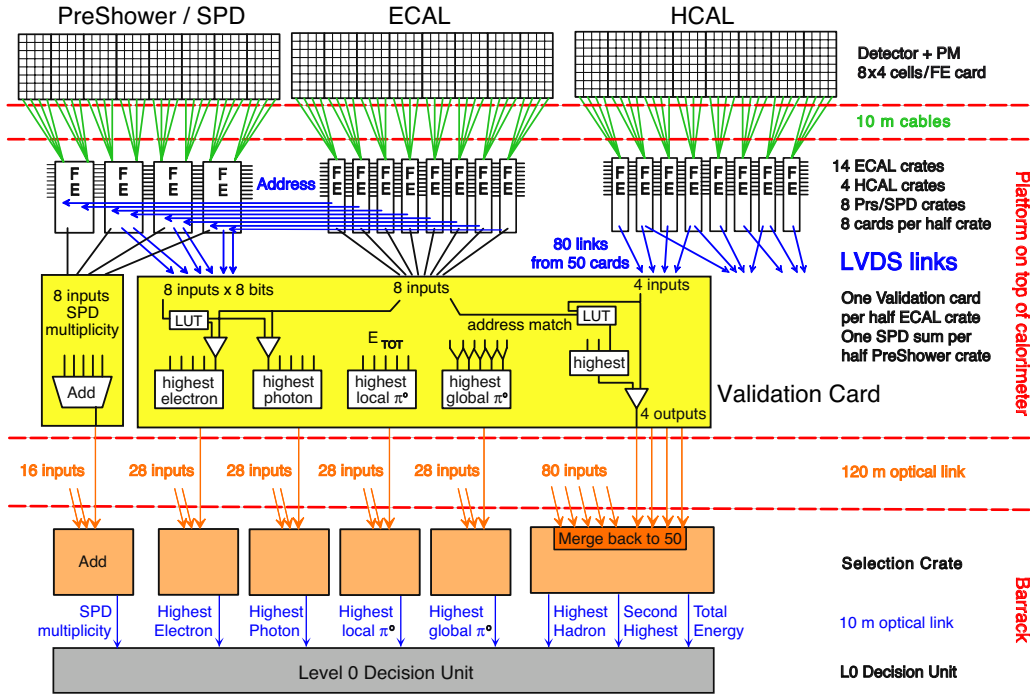
### 7.1.1 Overview

As shown in figure 7.2, the Level-0 trigger is subdivided into three components: the pile-up system, the Level-0 calorimeter trigger and the Level-0 muon trigger. Each component is connected to one detector and to the Level-0 DU which collects all information calculated by the trigger systems to evaluate the final decision.

The pile-up system aims at distinguishing between crossings with single and multiple visible interactions. It uses four silicon sensors of the same type as those used in the VELO to measure the radial position of tracks. The pile-up system provides the position of the primary vertices candidates along the beam-line and a measure of the total backward charged track multiplicity.

The Calorimeter Trigger system looks for high  $E_T$  particles: electrons,  $\gamma$ 's,  $\pi^0$ 's or hadrons. It forms clusters by adding the  $E_T$  of  $2 \times 2$  cells and selecting the clusters with the largest  $E_T$ . Clusters are identified as electron,  $\gamma$  or hadron based on the information from the SPD, PS, ECAL and HCAL Calorimeter. The  $E_T$  of all HCAL cells is summed to reject crossings without visible interactions and to reject triggers on muon from the halo. The total number of SPD cells with a hit are counted to provide a measure of the charged track multiplicity in the crossing.

The muon chambers allow stand-alone muon reconstruction with a  $p_T$  resolution of  $\sim 20\%$ . Track finding is performed by processing elements which combine the strip and pad data from the five muon stations to form towers pointing towards the interaction region. The Level-0 muon trigger selects the two muons with the highest  $p_T$  for each quadrant of the muon detector.



**Figure 7.3:** Overview of the Level-0 calorimeter trigger architecture.

The Level-0 DU collects all information from Level-0 components to form the Level-0 trigger. It is able to perform simple logic to combine all signatures into one decision per crossing. This decision is passed to the Readout Supervisor (see section 8.3) which transmits it to the front-end electronics.

The latency of Level-0, i.e. the time elapsed between a pp interaction and the arrival of the Level-0 trigger decision at the front-end electronics, is fixed to  $4 \mu\text{s}$ . This time which includes the time-of-flight of the particles, cable delays and all delays in the front-end electronics, leaves  $2 \mu\text{s}$  for the processing of the data in the Level-0 trigger to derive a decision.

### 7.1.2 Architecture

#### Calorimeter trigger

A zone of 2 by 2 cells is used, since it is large enough to contain most of the energy, and small enough to avoid overlap of various particles. Ultimately, only the particle with the highest  $E_T$  enters into the trigger decision. Therefore, to minimize the number of candidates to be processed, only the highest  $E_T$  candidate is kept at this stage.

These candidates are provided by a three step selection system as shown in figure 7.3:

- a first selection of high  $E_T$  deposits is performed on the Front-End card, which is the same for ECAL and HCAL. Each card handles 32 cells, and the highest  $E_T$  sum over the 32 sums of  $2 \times 2$  cells is selected. To compute these 32 sums, access to cells in other cards is an important issue.

- the Validation Card merges the ECAL with the PS and SPD information prepared by the preshower front-end card. It identifies the type of electromagnetic candidate, electron,  $\gamma$  or  $\pi^0$ . Only the highest  $E_T$  candidate per type is selected and sent to the next stage. The same card also adds the energy deposited in ECAL to the corresponding hadron candidates. Similar cards in the PreShower crates compute the SPD multiplicity.
- the Selection Crate selects the candidate with the highest  $E_T$  for each type, and also produces a measure of the total  $E_T$  in HCAL and the total SPD multiplicity.

The first two steps are performed on the calorimeter platform, at a location where the radiation dose is expected to be below 50 Gy over the whole lifetime of the experiment, and where single event upsets are expected to occur. Each component has been tested for radiation tolerance and robustness against single event upsets. Anti-fuse FPGAs are used, as well as *triple-voting* techniques.

The trigger interface is housed in one anti-fuse FPGA from ACTEL for ECAL/HCAL front-end cards and in one flash EEPROM based FPGA for PS/SPD front-end boards. There is a large data flow between these components at a frequency of 40 MHz, through a dedicated backplane, where interconnections are realized by point-to-point links running a multiplexed LVDS signals at 280 MHz. The same backplane is used for PreShower, ECAL and HCAL crates.

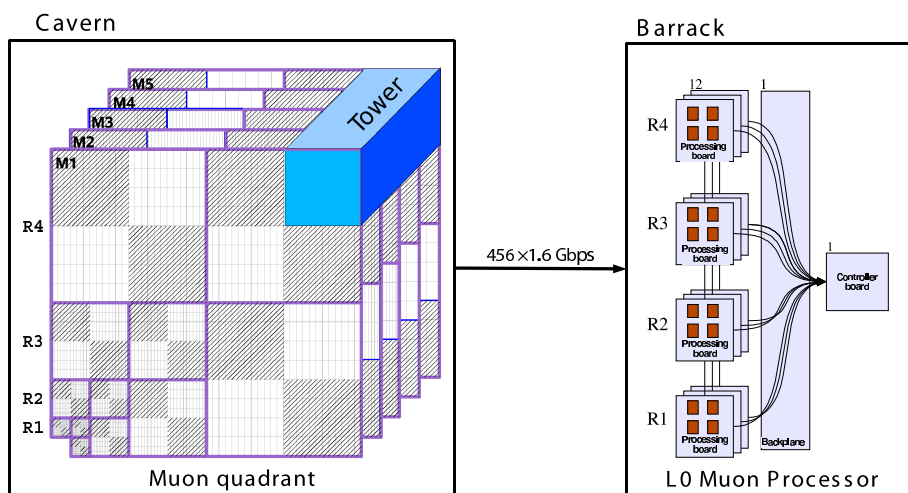
The validation card is a 9U board with 16 layers. Clusters, PS and SPD hit maps arrive through the backplane via 20 LVDS links running at 280 MHz. The cluster identification is performed by two ProAsic FPGAs from ACTEL. Electron,  $\gamma$ , hadron and  $\pi^0$  candidates are transmitted to the selection crate via an 8-channel optical mezzanine which serializes data at 1.6 Gbps and drives a ribbon of 12 fibres. The control of the validation and calorimeter front-end cards are performed by a SPECS interface.

The selection crate is located in the counting house in a radiation free environment. It is a modular system containing eight 16-layer 9U VME selection boards. The design of the selection boards is unique and adapted to perform both the electromagnetic and the hadron clusters selection. The electromagnetic cluster selection is performed on one board for each cluster type (electron,  $\gamma$ ,  $\pi^0$ ) while the hadron selection requires three boards. The results of the two first boards are transmitted to the third one where the final selection is performed. Finally, one board is used to sum the SPD multiplicity. Inputs arrive via 28 optical links grouped into three ribbons of 12 fibres. High-speed serial signals are deserialized by 28 TLK2501 chips.<sup>2</sup> The selection of the highest  $E_T$  candidate of each type is performed by six FPGAs from the Xilinx Virtex II family. The selected candidates are sent to the Level-0 DU via a mezzanine with 1-channel high speed optical link. Inputs and outputs of the Selection Boards are sent to the data acquisition system via two high speed optical links connected to the TELL1 board. The Selection Boards are controlled by a credit card PC.

The types and total numbers of boards for the Level-0 Calorimeters Trigger are summarized in table 7.1.

**Table 7.1:** Boards of the Level-0 calorimeters trigger.

Boards	Number
ECAL/HCAL front-end	246
PS/SPD front-end	100
8-channels optical mezzanine	80
1-channels optical mezzanine	40
Validation card	28
SPD Control board	16
Selection Board	8

**Figure 7.4:** Overview of the Level-0 muon architecture.

### Muon trigger

An overview of the Level-0 muon architecture is given in figure 7.4 and a detailed description in [215]. Each quadrant of the muon detector is connected to a Level-0 muon processor via 456 optical links grouped in 38 ribbons containing 12 optical fibres each. An optical fibre transmits serialized data at 1.6 Gbps over a distance of approximately 100 meters. The 4 Level-0 muon processors are located in the counting house, a place immune to radiation effects.

A L0 muon processor looks for the two muon tracks with the largest and second largest  $p_T$ . The track finding is performed on the logical pads. It searches for hits defining a straight line through the five muon stations and pointing towards the interaction point. The position of a track in the first two stations allows the determination of its  $p_T$ . The final algorithm is very close to the one reported in the Technical Proposal [1] and in the Muon Technical Design Report [164].

Seeds of the track finding algorithm are hits in M3. For each logical pad hit in M3, an extrapolated position is set in M2, M4 and M5 along a straight line passing through the hit and the interaction point. Hits are looked for in these stations in search windows termed Field Of Interest (FOI) which are approximately centred on the extrapolated positions. FOIs are opened along the  $x$ -

<sup>2</sup>from Texas Instrument, USA.

axis for all stations and along the  $y$ -axis only for stations M4 and M5. The size of the FOI depends on the station being considered, the level of background and the minimum-bias retention allowed. When at least one hit is found inside the FOI for each station M2, M4 and M5, a muon track is flagged and the pad hit in M2 closest to the extrapolation from M3 is selected for a subsequent use.

The track position in station M1 is determined by making a straight-line extrapolation from M3 and M2, and identifying, in the M1 FOI, the pad hit closest to the extrapolation point.

Since the logical layout is projective, there is a one-to-one mapping from pads in M3 to pads in M2, M4 and M5. There is also a one-to-one mapping from pairs of pads in M2 and M3 to pads in M1. This allows the track finding algorithm to be implemented using only logical operations.

To simplify the processing and to hide the complex layout of the stations, the muon detector is subdivided into 192 towers (48 per quadrant) pointing towards the interaction point. All towers have the same layout with 288 logical pads<sup>3</sup> each. Therefore, the same algorithm can be executed in each tower. Each tower is connected to a processing element, the basic component of the Level-0 Muon processor.

To collect data coming from a tower spread over five stations and to send them to the processing element, a patch panel close to the muon processor is used.

Processing elements have to exchange a large number<sup>4</sup> of logical channels with each other to avoid inefficiencies on borders of towers. The topology of the data exchange depends strongly on the location of the tower.

A processing element runs 96 tracking algorithms in parallel, one per M3 seed, on logical channels from a tower. It is implemented in a FPGA named Processing Unit (PU). A processing board contains four PUs and an additional FPGA to select the two muons with the highest transverse momentum within the board. A Level-0 Muon processor consists of a crate housing 12 Processing Boards, a custom backplane and a controller board. The custom backplane is mandatory to exchange logical channels between PUs. The controller board collects candidates found by the 12 Processing Boards and selects the two with the highest  $p_T$ . It also distributes signals coming from the TTC.

The Level-0 Muon implementation relies on the massive use of multigigabit serial links deserialized inside FPGAs. Processors are interfaced to the outside world via optical links while processing elements are interconnected with high speed copper serial links.

The Processing Board contains five FPGAs from the Altera Stratix GX family and 92 high speed serial links with serializers and deserializers embedded in FPGAs. The board sends data to the data acquisition system via two high speed optical links. The processing board is remotely controlled via Ethernet by a credit card PC running Linux. The size of the printed circuit is  $366.7 \times 220$  mm and is composed of 18 layers and a total of 1512 components. The power consumption is less than 60 W.

The Controller Board contains two FPGAs from the Stratix GX family. The board shares many common functionalities with the Processing Board: the same credit card PC, the same mechanism to send information to the data acquisition system. The printed circuit measues

<sup>3</sup> 48 pads from M1,  $2 \times 96$  pads from M2 and M3,  $2 \times 24$  pads from M4 and M5.

<sup>4</sup> A processing element handles 288 logical pads. It sends a maximum of 224 and receives a maximum of 214 logical channels from neighbouring elements.

**Table 7.2:** Boards of the Level-0 muon trigger.

Boards	Number
Processing Board	48
Controller Board	4
Backplane	4

366.7 × 220 mm and is composed of 14 layers with 948 mounted components. The power consumption is less than 50 W.

The backplane contains 15 slots: 12 for the Processing Boards, one for the Controller Board and two for test. It distributes power supplies, signals coming from the TTC, and assures the connectivity between the processing elements via 288 single-ended links (40 MHz) and 110 differential high speed serial links (1.6 Gbps). The size of the 18-layer printed circuit board is 395,4 × 426,72 mm.

The types and total numbers of boards for the Level-0 Muon Trigger are summarized in table 7.2.

### Pile-Up system

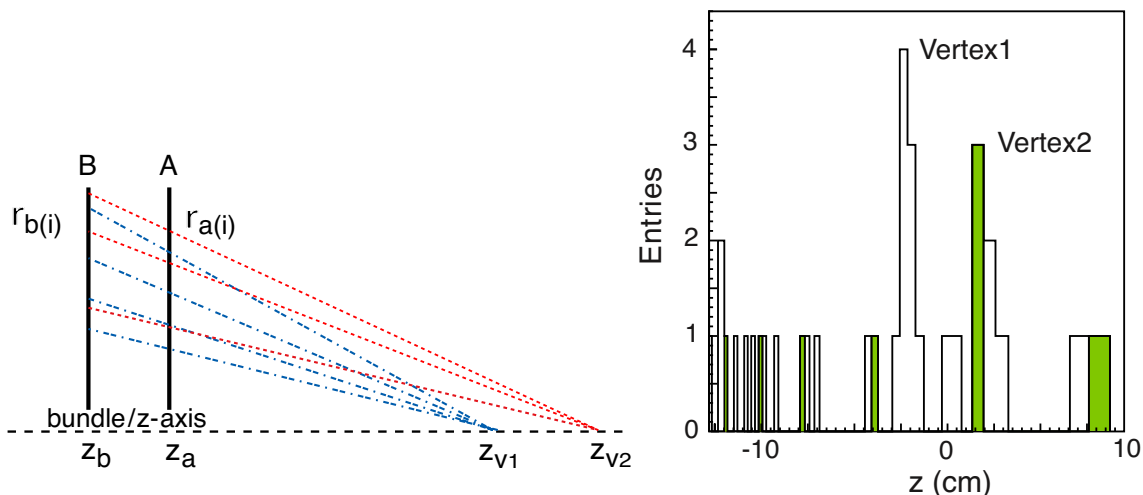
The pile-up system consists of two planes (*A* and *B*) perpendicular to the beam-line and located upstream of the VELO (see figure 5.1). Each 300 μm thick silicon plane consists of two overlapping VELO R-sensors which have strips at constant radii, and each strip covers 45°. In both planes the radii of track hits,  $r_a$  and  $r_b$ , are measured. The hits belonging to tracks from the same origin have the simple relation  $k = r_b/r_a$ , giving:

$$z_v = \frac{kz_a - z_b}{k - 1} \quad (7.1)$$

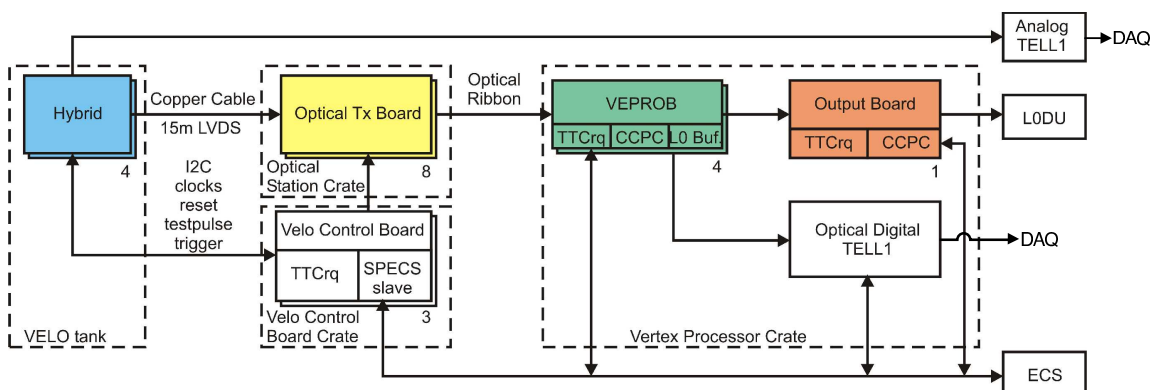
where  $z_b$ ,  $z_a$  are the detector positions and  $z_v$  is the position of the track origin on the beam axis, *i.e.* the vertex. The equation is exact for tracks originating from the beam-line. All hits in the same octant of both planes are combined according to equation 7.1 and the resulting values of  $z_v$  are entered into an appropriately binned histogram, in which a peak search is performed, as shown in figure 7.5. The resolution of  $z_v$  is limited to around 3 mm by multiple scattering and the hit resolution of the radial measurements. All hits contributing to the highest peak in this histogram are masked, after which a second peak is searched for. The height of this second peak is a measure of the number of tracks coming from a second vertex. A cut is applied on this number to detect multiple interactions. If multiple interactions are found, the crossing is vetoed.

The architecture of the pile-up system is shown in figure 7.6. It uses the signals of the integrated comparators of the Beetle chips located on the four hybrids. The outputs of neighbouring comparators are OR-ed in groups of four, resulting in 256 LVDS links running at 80 Mbit/s per hybrid, which send the Level-0 signals to eight Optical Transmission Boards. Two Optical Transmission Boards cover one quadrant. They time align and multiplex input hit maps to four Vertex Processing Boards. Hit maps of one bunch crossing are sent to one of the four Vertex Processing Board (VEPROB) in four consecutive clock cycles, while hit maps of the following bunch crossing





**Figure 7.5:** The basic principle of detecting vertices in an event. The hits of plane A and B are combined in a coincidence matrix. All combinations are projected onto a  $z_v$ -histogram. The peaks indicated correspond to the two interaction vertices in this particular MonteCarlo event. After the first vertex finding iteration, the hits corresponding to the two highest bins are masked, resulting in the hatched histogram.



**Figure 7.6:** Overview of the Level-0 pile-up Architecture.

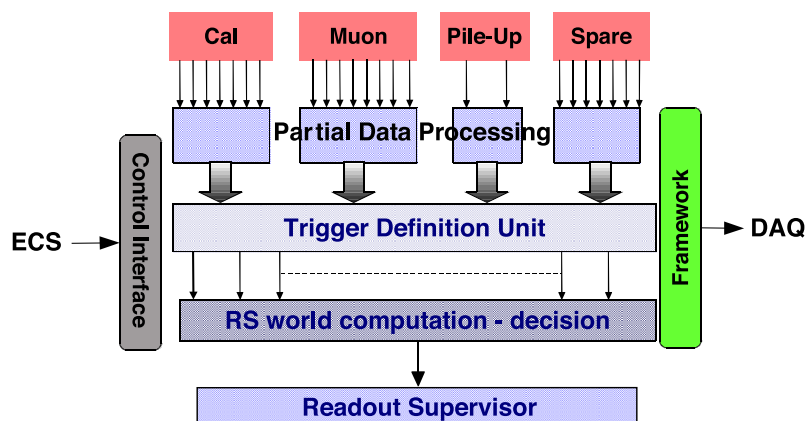
are sent to the second VEPROB in four consecutive clock cycles. Bunch-crossings are distributed over the four Vertex Processing Boards in a round-robin fashion. The Optical transmission board is a 9U board controlled by a SPEC interface via the VELO control board.

Vertex Processing boards are 9U boards located in the radiation-free electronics barracks. They are connected to the Optical Transmission boards via 24 high speed optical links. The vertex processing board is the key component of the pile-up system. It houses a large FPGA from Xilinx Virtex II family which runs the pile-up algorithm. A board handles one of four events and sends its trigger decision to the output board via a high speed copper link (1.6 Gbps). The VEPROB is controlled by a credit card PC and sends the inputs and outputs of the vertex finding algorithm to the DAQ system via two high speed optical links.

The output board is a simple 9U board multiplexing the inputs from the vertex processing

**Table 7.3:** Boards of the pile-up system.

Boards	Number
Hybrids	4
Optical Transmission Board	8
Vertex Processing Board	4
Output Board	1

**Figure 7.7:** Level-0 DU architecture.

board and sends the number of primary pp interactions for each bunch crossing to the Level-0 DU. In addition, the output boards make histograms of trigger decisions made by the pile-up system. These histograms are accessible via the ECS interface.

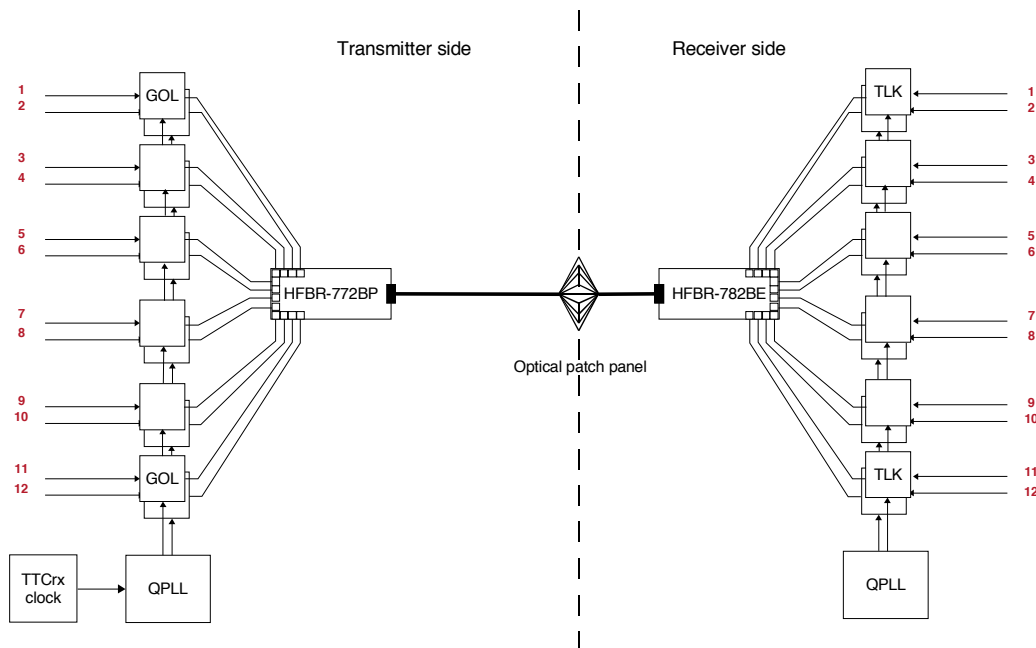
The types total numbers of boards for the pile-up system are summarized in table 7.3.

### Decision Unit

The Level-0 DU receives information from the calorimeter, muon and pile-up sub-triggers at 40 MHz, which arrive at different fixed times. The computation of the decision can start with a sub-set of information coming from a Level-0 sub-trigger, after which the sub-trigger information is time aligned. An algorithm is executed to determine the trigger decision. The decision is sent to the Readout Supervisor, which makes the ultimate decision about whether to accept an event or not. The Readout Supervisor is able to generate and time-in all types of self-triggers (random triggers, calibration, etc.) and to control the trigger rate by taking into account the status of the different components in order to prevent buffer overflows and to enable/disable the triggers at appropriate times during resets.

The architecture of the Level-0 DU is shown in figure 7.7. For each data source, a Partial Data Processing system performs a specific part of the algorithm and the synchronisation between the various data sources. Then a trigger definition unit combines the information from the above systems to form a set of trigger conditions based on multi-source information.

The trigger conditions are logically OR-ed to obtain the Level-0 decision after they have been individually downscaled if necessary.



**Figure 7.8:** Overview of ribbon optical link.

The Level-0 DU is based on the TELL1 board with optical cards replaced by a single mezzanine in which Level-0 DU hardware is implemented. Inputs are received on two ribbons of 12 high speed optical links. Serial signals are deserialized by a 24 TLK2501 chip.<sup>5</sup> and sent to two large FPGAs from the Stratix Family. Electron,  $\gamma$ ,  $\pi^0$ , hadron and muon candidates as well as intermediate and final decisions are sent to the DAQ via the TELL1 mother boards. This information can be used later on by the HLT to confirm the Level-0 candidates using more refined algorithms.

### 7.1.3 Technology

The implementation of the Level-0 trigger relies on the massive use of large FPGAs, high speed serial links and common techniques which simplify debugging and commissioning.

#### High speed links

The transport of information from the front-end electronics to Level-0 trigger boards located in the barrack is based on three concepts:

- serialization of the detector data;
- use of optical links as transport media;
- use of high density devices.

High speed serial transmission reduces the number of signal lines required to transmit data from one point to another. It also offers a high level of integration with many advantages: high reliability

<sup>5</sup>from Texas Instrument, USA.

for data transfer over 100 meters; complete electrical isolation avoids ground loops and common mode problems. In addition, the integration of several high speed optical links in a single device increases data rate while keeping a manageable component count and a reasonable cost.

Ribbon optical links integrate twelve optical transmitters (fibres, receivers) in one module. The important benefit of ribbon optical links is based on low-cost array integration of electronic and opto-electronic components. It also results a low power consumption and a high level of integration.

An overview of the ribbon optical link developed for the Level-0 trigger is shown in figure 7.8. The emitter stage relies on twelve serializer chips connected to one optical transmitter. The serializer is the GOL, a radiation hard chip designed by the CERN microelectronic group, which every 25 ns, transforms a 32-bit word into a serial signal with a frequency of 1.6 GHz using a 8B/10B encoding. High frequency signals are converted into optical signals by the 12-channel optical transmitter from Agilent HFBR-772BE. The module is designed to operate multimode fibres at a nominal wavelength of 850 nm.

Initially the LHC clock distribution was not intended to be used for optical data transmission and hence, does not fulfill the severe jitter constraints required by high speed serializers. The GOL requires a maximum jitter of 100 ps peak to peak to operate correctly whereas the LHC clock jitter is as large as 400 or 500 ps. To reduce the jitter, a radiation hard chip, the QPLL, designed by the CERN microelectronics group is used. It filters out the jitter up to an acceptable value with the help of a reference quartz crystal associated to a phase locked loop.

The emitter side is close to the detector in a place where the radiation dose is below 50 Gy over 10 years where single event upsets (SEU) are expected to occur. The GOL and QPLL chips are radiation hard chips immune to SEU. However, the optical transceiver is a commercial component designed to work in an environment free of radiation. An irradiation campaign took place at the Paul Scherrer Institute in December 2003. The component was found to work within its specifications up to a total dose of 150 Gy. The cross-section for single event upsets is equal to  $(4.1 \pm 0.1) \times 10^{-10} \text{ cm}^2$  per single optical link. The expected SEU rate is 1 every 220 minutes for the Level-0 muon trigger. When this happens, a single optical link emitter is not synchronized with its receiver anymore. All emitter/receiver pairs are resynchronized automatically at the end of each LHC cycle. Therefore, the link will not transmit data during a maximum of one LHC cycle or 89  $\mu\text{s}$ . The corresponding inefficiency is negligible.

The physical media between the front-end electronic boards and the Level-0 trigger board consist of ribbons of twelve fibres with MPO connectors on both sides ( $\sim 10 \text{ m.}$ ), MPO-MPO patch panels, long cables containing eight ribbons with MPO connectors ( $\sim 80 \text{ m.}$ ), fanout panels (MPO-MPO or MPO-SC), short ribbons of twelve fibres ( $\sim 3 \text{ m}$ ) with MPO connector on one side and a MPO or 12 SC connectors on the other side.

The receiving side is the mirror of the emitting side. Optical signals are converted into 1.6 Gbps serial electrical signals by the 12-channel optical receiver HFBR-782BE. The twelve high-frequency signals are deserialized into 16-bit words at 80 MHz by twelve TLK2501 chips. The receiving side is located in the counting room. Therefore standard components can be used. In the muon processing board, where the density of input signal is high, TLK2501 chips are replaced by serializers and deserializers embedded in the Stratix GX FPGA.

The routing of the differential high speed traces between serializer/deserializer and the optical transceiver requires considerable care since the geometry of the tracks must be totally controlled to guarantee good impedance matching and to minimize electromagnetic emissions to the environment as well as sensitivity to electromagnetic perturbations from the environment.

The performance of the optical link has been measured with several setups in different ways. The bit error ratio measured with Lecroy SDA11000 Serial Data Analyser is below  $10^{-16}$  for a single fibre of 100 m long.

### Field Programmable Gate Arrays

Three FPGA technologies are used in the Level-0 trigger. They are characterized by the way they are configured:

- Anti-fuse based FPGAs (ACTEL AX family), that can be programmed only once;
- Flash-EEPROM based FPGAs (ACTEL pro-ASIC family), that can be erased and reprogrammed;
- RAM based FPGAs (Altera Acex, Flex, Stratix and Stratix GX families or Xilinx Virtex family) that can be reprogrammed an unlimited number of times.

Anti-fuse and flash FPGAs are used in the front-end boards close to the detector and are therefore exposed to significant radiation doses. These components have been tested in heavy ion beams and have shown very low sensitivity to single event upsets and single event latch-up. Special mechanisms such as triple-voting or horizontal and vertical parity are implemented to increase the protection of registers containing critical data. Dose effects begin to appear in Flash based FPGAs for doses an order of magnitude above the total dose received during 10 years by the trigger front-end electronics.

RAM-based FPGAs are known to be very sensitive to single event upsets. For this reason their use is restricted to boards located in the barracks which is a radiation free area.

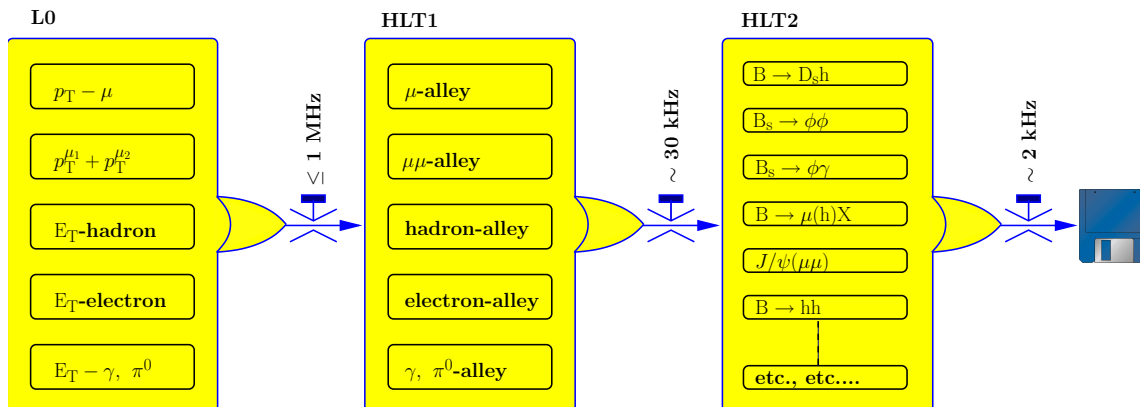
All the FPGAs used in the trigger provide for good visibility of internal node behavior during the debug phase by providing embedded logic analyzer features (Silicon Explorer for ACTEL, SignalTap for the largest components of the Altera family and Chipscope for the Xilinx family).

### Debugging and monitoring tools

Each Level-0 trigger board includes either a credit card PC or a SPECS component interfaced to the embedded FPGAs by a custom 16-bit bus. By this means the operation of any of the FPGAs is controlled and error detection mechanisms, such as error counters, spy and snooping mechanisms are implemented.

To test a complete sub-trigger in stand-alone mode, a data injection buffer to substitute input data is implemented. Results of the processing can be read back via the credit card PC at the output of dedicated SPY memories

The level-0 trigger is a very complex system. Any malfunctions can therefore be difficult to understand and interpret. At each stage the input and results of the processing are logged. In



**Figure 7.9:** Flow-diagram of the different trigger sequences.

addition, a software emulator was developed which reproduces the behaviour of the hardware at the bit level. By comparing results computed by the hardware with those of the emulator run on the same input data, any faulty components can quickly be located.

## 7.2 High Level Trigger

The High Level Trigger (HLT) consists of a C++ application which runs on every CPU of the Event Filter Farm (EFF). The EFF contains up to 2000 computing nodes and is described in section 8. Each HLT application has access to all data in one event, and thus, in principle, could execute the off-line selection algorithms. However, given the 1 MHz output rate of the Level-0 trigger and CPU power limitations, the HLT aims to reject the bulk of the uninteresting events by using only part of the full event data. In this section, the algorithm flow is described which, according to MonteCarlo simulation studies, is thought to give the optimal performance within the allowed time budget. However, it should be kept in mind that since the HLT is fully implemented in software, it is very flexible and will evolve with the knowledge of the first real data and the physics priorities of the experiment. In addition the HLT is subject to developments and adjustments following the evolution of the event reconstruction and selection software.

A schematic of the overall trigger flow is shown in figure 7.9. Level-0 triggers on having at least one cluster in the HCAL with  $E_T^{\text{hadron}} > 3.5 \text{ GeV}$ , or the ECAL with  $E_T^{e, \gamma, \pi^0} > 2.5 \text{ GeV}$ , or a muon candidate in the muon chambers with  $p_T^\mu > 1.2 \text{ GeV}$ , or  $p_T^{\mu_1} + p_T^{\mu_2} > 1. \text{ GeV}$ , where  $\mu_1$  and  $\mu_2$  are the two muons with the largest  $p_T$ . The above thresholds are typical for running at a luminosity of  $2 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ , but depend on luminosity and the relative bandwidth division between the different Level-0 triggers. All Level-0 calorimeter clusters and muon tracks above threshold are passed to the HLT as part of the Level-0 trigger information as described in section 7.1.2, and will be referred to as Level-0 objects henceforward.

The HLT is subdivided in two stages, HLT1 and HLT2. The purpose of HLT1 is to reconstruct particles in the VELO and T-stations corresponding to the Level-0 objects, or in the case of Level-0  $\gamma$  and  $\pi^0$  candidates to confirm the absence of a charged particle which could be associated to these objects. This is called Level-0 confirmation, and the details of how this is achieved within the

CPU time budget is explained below. HLT1 should reduce the rate to a sufficiently low level to allow for full pattern recognition on the remaining events, which corresponds to a rate of about 30 kHz. At this rate HLT2 performs a combination of inclusive trigger algorithms where the B decay is reconstructed only partially, and exclusive trigger algorithms which aim to fully reconstruct B-hadron final states.

### 7.3 HLT1

HLT1 starts with so-called *alleys*, where each alley addresses one of the trigger types of the Level-0 trigger. About  $\sim 15\%$  of the Level-0 events are selected by multiple triggers, and will consequently pass by more than one alley. To confirm the Level-0 objects each alley makes use of the following algorithms:

**L0→T:** The Level-0 objects are assumed to originate from the interaction region, which defines the whole trajectory of the candidate in the spectrometer. So-called T-seeds are reconstructed in the T-stations, decoding only the hits in a window around the trajectory, or in case of the calorimeter clusters the two trajectories corresponding to the two charge hypothesis. The seeds are required to match the Level-0 object in both space and momentum.

**L0→VELO:** VELO-seeds are reconstructed in two stages. First the information from the R-sensors are used to reconstruct 2D-tracks. The  $\chi^2$  is calculated for the matching of a 2D track with the Level-0 object, and only candidates with a sufficiently low  $\chi^2$  are used to reconstruct a VELO-seed using the  $\phi$ -sensor information. These VELO-seeds in turn are required to match the Level-0 object with a sufficiently small  $\chi^2$ . In addition the 2D-tracks are used to reconstruct the primary vertexes in the event [216].

**VELO→T:** The VELO-seeds above define a trajectory in the T-stations, around which a T-seed is reconstructed completely analogue to the L0→T algorithm described above.

**T→VELO:** this algorithm finds the VELO-seeds which match a T-seed, using an algorithm analogue to the L0→VELO algorithm, but now starting from a T-seed, rather than a Level-0 object.

Each HLT1 alley uses a sequence of the above algorithms to reduce the rate. An algorithm common to all alleys is used for computing the primary vertex with the 2D tracks reconstructed in the VELO. While the alleys are operating independently, care has been taken to avoid having to reconstruct the same track or primary vertex twice to avoid wasting precious CPU power.

While the bandwidth division between the alleys has not been defined, the performance of the alleys will be illustrated with two typical alleys, the muon and hadron alleys running at a luminosity of  $2 \times 10^{32} \text{ cm}^{-2}\text{s}^{-1}$ .

The HLT1  $\mu$ -alley input rate will be  $\sim 230$  kHz, and contain 1.2  $L0^\mu$  objects per event.  $L0^\mu \rightarrow T$  reduces the rate to 120 kHz, while the number of candidates increases to 1.8 T-seeds per event. T→VELO reduces the rate to 80 kHz. Requiring the remaining candidates to have an impact parameter to any primary vertex larger than 0.1 mm reduces the rate to 10 kHz. The HLT1 hadron-alley input rate will be  $\sim 600$  kHz, and contain 1.3  $L0^{\text{hadron}}$  objects per event.  $L0^{\text{hadron}} \rightarrow \text{VELO}$ , requiring a 0.1 mm impact parameter of the VELO-seeds to any primary vertex reduces the rate to 300 kHz which contain 2.2 VELO-seeds per event. VELO→T reduces this rate to 30 kHz with 1.2 candidates per event. Since this rate is still too large for the HLT2 stage, a further reduction is obtained by requiring a VELO-track with a distance of closest approach to the confirmed Level-0

track of less than 0.2 mm, and a  $p_T$  of at least 1 GeV. This reduces the rate to 11 kHz with 3.2 candidate secondary vertices per event. The other HLT1 alleys employ similar strategies.

## 7.4 HLT2

The combined output rate of events accepted by the HLT1 alleys is sufficiently low to allow an off-line track reconstruction as described in section 10.1. The HLT-tracks differ from the off-line in not having been fitted with a Kalman filter to obtain a full covariance matrix since this is too CPU intensive. Prior to the final selection, a set of tracks is selected with very loose cuts on their momentum and impact parameter. These tracks are used to form composite particles, such as  $K^* \rightarrow K^+\pi^-$ ,  $\phi \rightarrow K^+K^-$ ,  $D^0 \rightarrow hh$ ,  $D_s \rightarrow K^+K^-\pi^-$  and  $J/\psi \rightarrow \mu^+\mu^-$ , which are subsequently used for all selections to avoid duplication in the creation of final states.

The HLT2 stage uses therefore cuts either on invariant mass, or on pointing of the B momentum towards the primary vertex. The resulting inclusive and exclusive selections aim to reduce the rate to about 2 kHz, the rate at which the data is written to storage for further analysis. The exclusive triggers are sensitive to tracking performance, while the inclusive triggers select partial B decays to  $\phi X$ ,  $J/\psi X$ ,  $D^* X$ ,  $\mu^\pm X$ ,  $\mu^\pm h X$  and  $\mu^+\mu^- X$  and therefore are less dependent on the on-line reconstruction. However, the exclusive selection of these channels produces a smaller rate, thus allowing for a more relaxed set of cuts. The final trigger is the logical OR of the inclusive and exclusive selections.

## 7.5 HLT monitoring

Each HLT1 alley and HLT2 selection produces summary information which is written to storage for the accepted events. This summary contains the information of all tracks and vertexes which triggered the event. It is foreseen to reserve a significant fraction of the output bandwidth for triggers on semi-leptonic B-decays, hence a sample in which the trigger did not bias the decay of the accompanying B-hadron. The summary information is used to check if an event would have triggered, even if the B decay of interest would not have participated in the trigger. It therefore allows to study the trigger performance. The summary information also guarantees that during the analysis the trigger source of an individual event is known.

To assure that during off-line analysis the trigger conditions are known, the combination of trigger algorithms with their selection parameters will be assigned a unique key, the Trigger Configuration Key (TCK). All trigger configurations with their associated TCK are pre-loaded in the EFF before a fill. To change from one trigger configuration to another one, for example to follow the decaying luminosity in a fill, a new TCK must be selected. This TCK is attached by the Time and Fast Control system (TFC, see section 8.3) to each event, and it steers the configuration of the algorithms on the EFF and allows full traceability of the used configuration.