3 FoCal-E pixel layers

3.1 Overview

As described in Sec. 2, FoCal-E is made up of 22 FoCal-E modules, with 11 modules being stacked on each side of the beam pipe. Each FoCal-E module is segmented longitudinally into 20 module-layers, with two pixel layers and the remainder pad layers. The pixel layers are located at positions 5 and 10 in the 20-layer module stack, as illustrated in Fig. 2. Each pixel module-layer consists of six multi-chip strings (see Sec. 3.3), a simple *Printed Circuit Board* (PCB) called *Transition Card* (TC) providing the interface with the *Readout Unit*s (RUs) (see Sec. 3.5), and supporting mechanical structure.

Figure 16 illustrates the arrangement of the stacked module-layers, which constitutes one complete pixel layer. The beam pipe passes through the rectangular opening in the center. The innermost section of the detector, where the highest data throughput is expected, is instrumented with ALPIDE pixel sensors in the *Inner Barrel* (IB) mode (with 1.2 Gbps), while for the outer sections all sensors have *Outer Barrel* (OB) (master or slave) mode (with 400 Mbps).

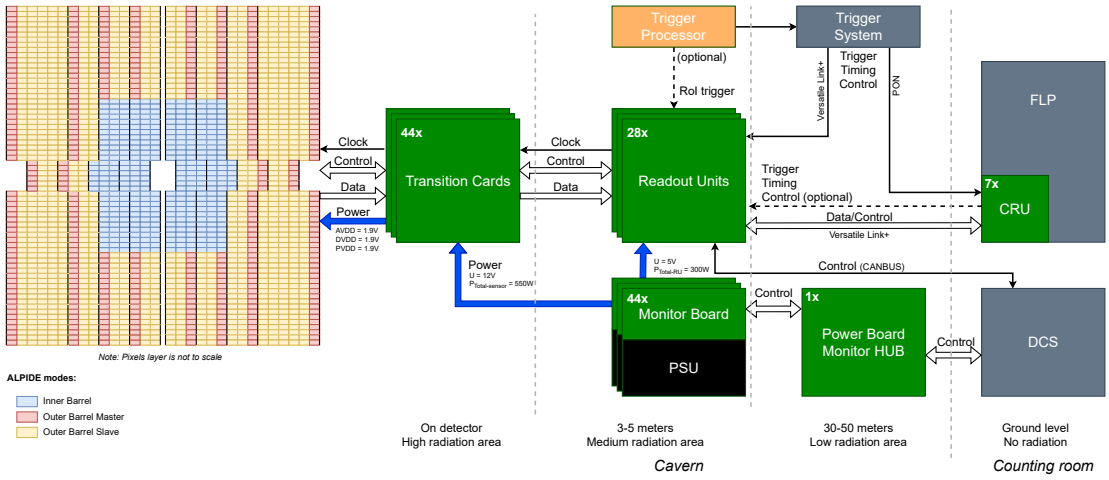


Fig. 16: Sketch of one complete pixel layer, including readout and power scheme integrated in the ALICE O2 system. The different modes of ALPIDEs sensor employed are IB mode (light blue); OB master mode (light red), and OB slave mode (yellow). Components specific to the pixel layer will be discussed in the section. Optional repeaters between TCs and RUs are not shown.

Figure 16 also shows the readout chain. The readout is the same as that for the ITS2 detector [20], modified as follows (explained further below):

- The ALPIDEs are bonded in multi-chip strings with the *Single-point Tape Automated Bonding* (SpTAB) technique instead of wire bonding as in ITS2 [5];
- The TC is connected to the strings of ALPIDEs. There is one TC per pixel module-layer, defined by six neighbouring strings;
- Active repeater boards can be inserted between the RUs and the TCs. This can be needed in the unlikely situation where the readout units must placed further away from the detector, due to space limitations in the default position;
- The RU is an updated version of the ITS2 RU, e.g. the (outdated) *GigaBit Transceiver* (GBTx) *Application-Specific Integrated Circuit*s (ASICs) are replaced with the *low power GigaBit Transceiver* (lpGBT);

[5] A row of ALPIDEs bonded together is called a string in FoCal and a *Hybrid Integrated Circuit* (HIC) in ITS2.

– The RU supports a copper cable input, which can be used for a local trigger in the laboratory or in test beam. In addition, it enables the pixel readout to be triggered directly based on hit information from the pad layers (Sec. 7.3);

– The *Common Readout Unit* (CRU) firmware is updated with a lpGBT *Field-Programmable Gate Arrays* (FPGA) module to support the *Versatile Trans-Receiver PLUS* (VTRx+), located on the RUs;

– The power distribution system is inherited from the *proton Computed Tomography* (pCT) project [12, 21]. It is a homogeneous system with one small power supply and one monitoring board per pixel module-layer. The regulators are situated on the TC and are controlled by the monitoring board.

3.2 Pixel sensor — ALPIDE

ALPIDE is a MAPS which was developed for the ITS upgrade and has been in use since the start of Run 3 [20]. The ALPIDE layout is shown in Fig. 17. It is implemented in 180 nm TowerJazz CMOS technology, containing a matrix of 1024 × 512 pixels in the x×y plane. Total area of the sensor is 30 mm × 15 mm, of which 29.94 mm × 13.76 mm is the pixel area, with the remainder at the lower edge containing digital circuitry for readout and control. The size of an individual pixel is 29.24 μm × 26.88 μm, containing charge collection diode, amplification circuit, and discriminator. The thickness of FoCal ALPIDE sensors will be 100 μm.

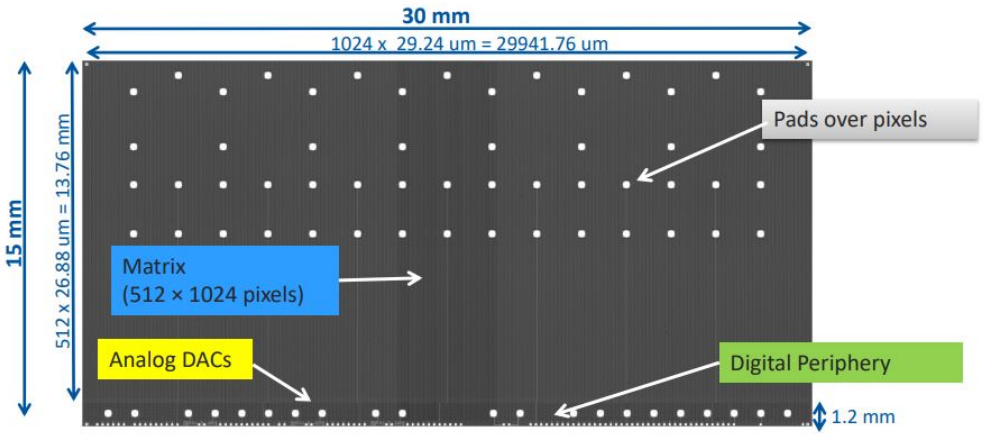


Fig. 17: Layout of a single ALPIDE showing the position of its components.

ALPIDE operates in three different modes: IB, OB Master and OB Slave, which differ in data interface and speed. The IB mode uses the serial data port, with a rate of 1200 Mbps. However, for transmission the data are 8b/10b–encoded, with maximum data throughput of **960** Mbps. The OB Master mode uses the same serial data port, with reduced data rate of 400 Mbps. Its data are likewise 8b/10b–encoded, with maximum data throughput of **320** Mbps. These two modes transmit the data from the detector using differential signaling. The OB Slave mode utilizes a bi–directional parallel data port with single-ended signaling, with maximum data throughput of **320** Mbps. A maximum number of six slaves can be connected to one master. The chips operated in this mode share this data link with each other and with the corresponding chip in OB Master mode. There is only one design of the ALPIDE chip, and the chip mode is selected by choice of bonding pads and address [22].

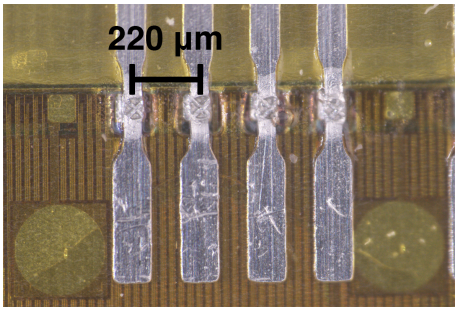
The ALPIDE chip can be operated in two different data acquisition modes: a *continuous* mode in which the ALPIDE is configured to internally generate continuous trigger frames at a frequency set by the user, or a *triggered* mode in which the ALPIDE expects a trigger signal from an external source. One important difference between these two modes is the handling of the multi-event buffers, in particular

494 in situations where the trigger or particle hit rate exceeds the readout data rate of the chip. In triggered
495 mode, the ALPIDE preserves the readout of every event recorded. However, received triggers for a given
496 ALPIDE are skipped by a BUSY-state handling mechanism when there is no free space in the event
497 buffer left. In continuous mode, the ALPIDE ensures always one free event buffer in order to record the
498 next trigger frame. In order to do so, it might abort an event read out, leading to data loss in this specific
499 event. For FoCal, operation in triggered mode is foreseen, as described in detail in Sec. 3.4.3, with the
500 external source to be configured to send a *p*eriodic trigger signal.

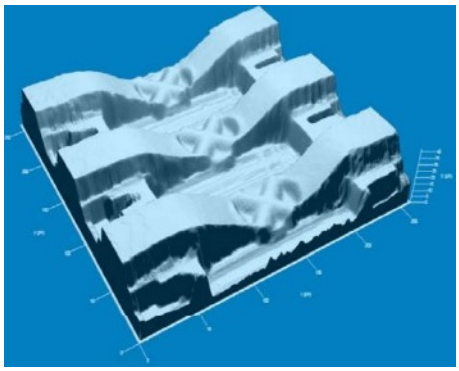
501 3.3 String design

502 *3.3.1 Technological approach for the components*

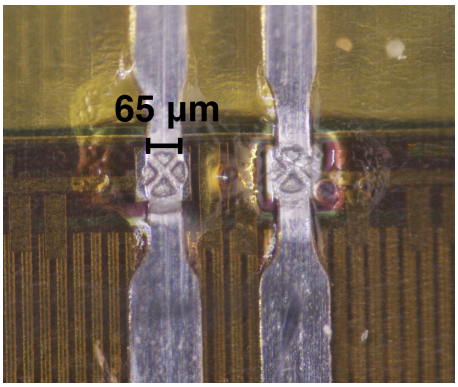
503 Reliable materials, technologies, and approaches are required for realization of the interconnection of
504 components in pixel layers. The materials for single- and multi-layered flexible micro-cables and boards
505 are adhesive-free aluminium-polyimide dielectrics with aluminium conductive layer thickness of 15, 30,
506 and 100 μm and polyimide dielectric thickness from 10 to 20 μm. For the interconnection technique,
507 ultrasonic welding of aluminium ribbon leads (referred to as SpTAB) is used. Such approaches, ma-
508 terials, and techniques have already been used in R&D and production stages by ALICE ITS [23–25],
509 CBM [26, 27], Mu3e [28], and PANDA [29]. They have also been used for prototyping of the *Digital*
510 *Tracking Calorimeter* (DTC) by the pCT project [12, 21] and for the EPICAL-2 digital calorimeter pro-
511 totype [5], achieving good results and reliable operation in both cases.



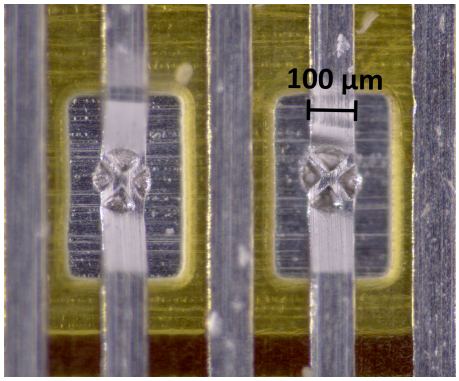
(a) Bond joint to the smaller pads on the periphery of the ALPIDE and the pitch between traces.



(b) Microscopic detail of SpTAB joints [30].



(c) Chip-to-flex SpTAB.



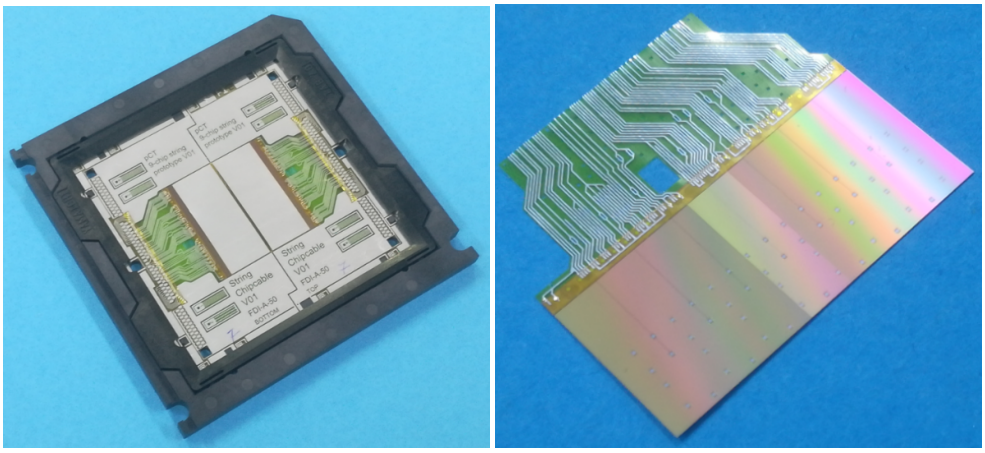
(d) Flex-to-flex SpTAB.

Fig. 18: Details of the SpTAB technique.

3.3.2 SpTAB-technique as interconnection technique

Optimal interconnection techniques are crucial for reliable operation. Several techniques were explored during development of the ITS2ALPIDE chips and detector modules: laser soldering by small alloy balls; ultrasonic wire welding (bonding) by thin aluminium wires; and ultrasonic welding (bonding) of aluminium ribbon leads of flexible layers (SpTAB technique (see Fig. 18). Each such approach has its advantages and disadvantages. The SpTAB technique was chosen for the pixel layers due to the following features:

- using adhesive-less aluminium-polyimide dielectrics with aluminium conductive layer creates a non-corrosive, uniform, reliable and mechanically stable mono-metallic aluminum-aluminum joint, which connects aluminium contact pads on the sensors and the aluminium ribbon leads of aluminium-polyimide interconnection elements.
- no Ni/Au layer on contact pads of the sensor (contrast laser soldering);
- absence of thermal influence on the chip (contrast laser soldering);
- no requirement of extremely high precision at the manufacturing of boards (contrast laser soldering));
- flexible board does not need to be precisely glued to the sensors because the ribbon leads are connected to the sensors during welding (contrast wire bonding);
- half as many weld joints compared to wire bonding because the ribbon leads are directly connected to the sensors during welding;
- all bond joints are encapsulated by glue after welding, which avoids damage of bond joints (contrast aluminium thin wire connections during assembling, where joints are not protected by glue);
- relatively simple repair of bond joints (contrast laser soldering);
- opportunity to use standard automated equipment e.g. wedge–wire bonding machines (with minor tuning) similar to wire bonding.



(a) Framed cables with chips. (b) Cut-off work area with chip.

Fig. 19: Prototype of chip-cable with ALPIDEs in Yamaichi plastic frame (left) and cut-off work area (right).

3.3.3 Chip-cable for ALPIDE mounting

A key point for building complex and high–cost detector modules and their components is high assembly yield. Module failure may arise from MAPS sensor defects that were not detected during testing bare chips. However, in the case of direct mounting MAPS to multilayered flex (e.g. in the case of laser

soldering or wire bonding techniques), it is difficult to substitute bad MAPS after assembly. Special thin micro-cables ("chip-cable") are therefore used in the pixel layers (see Fig. 19), to minimize the usage of bad chips.

These cables connect to MAPS by SpTAB, with the assembly installed in a plastic frame for further testing. After successful testing, the bond joining the chip-cable to the chip is encapsulated in glue, with final testing carried out after glue polymerization. This approach minimizes the use of bad chips for further assembly of the multi-chip strings.

A plastic frame and test socket from the Yamaichi Company (Japan) is used for testing chips on chip-cables. After final testing, the assembled chips are separated from the frame (see Fig. 19). This approach enables testing of both bare chip-cables during their production (to eliminate shorts or breaks of aluminium traces), and assembled MAPS chips with chip-cables during pixel layer assembly (functional tests of chips).

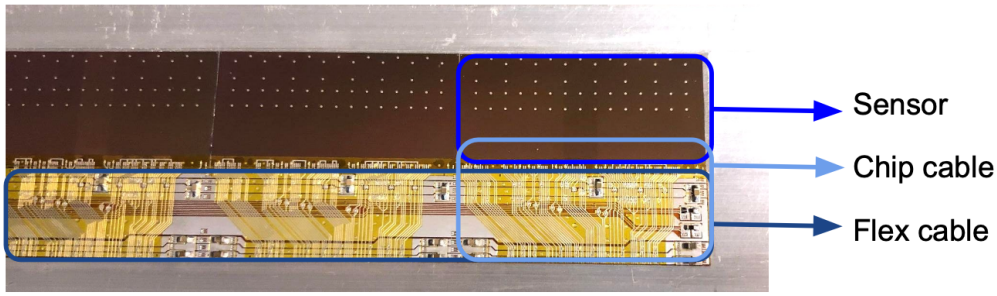


Fig. 20: Detail of the string structure and its various components.

3.3.4 Components of the string

Figure 20 shows the components of a string: a sensor bonded to a chip-cable and the chip-cable bonded to a flex cable. The flex cable is manufactured using the same techniques as the chip-cable, but is longer and has an additional layer to support the traces and power for all sensors. In addition it has small passive components, such as capacitors and termination resistors. The flex cable is glued parallel to the sensors.

3.3.5 Multi-chip string design

The pixel sensors are bonded into multi-chip strings. There are three versions of such strings: IB/OB 15-chip string, OB 15-chip string, and IB/OB 12-chip string. The strings are segmented into groups, as described in Sec. 3.5.1. Figure 21 shows the design and segmentation for all three versions, with features as follows:

- **IB/OB 15-chip string** with six IB (positions 15 to 10), two OB master (positions 7 and 1), and seven OB slave (positions 9, 8, and 6 to 2) chips segmented into four groups;
- **OB 15-chip string** with four OB master (positions 13, 10, 7, and 1) and eleven OB slave (positions 15, 14, 12, 11, 9, 8, and 6 to 2) chips segmented into four groups;
- **IB/OB 12-chip string** with six IB (positions 12 to 7), two OB master (positions 4 and 1), and four OB slave (positions 6, 5, and 3, 2) chips segmented into four groups.

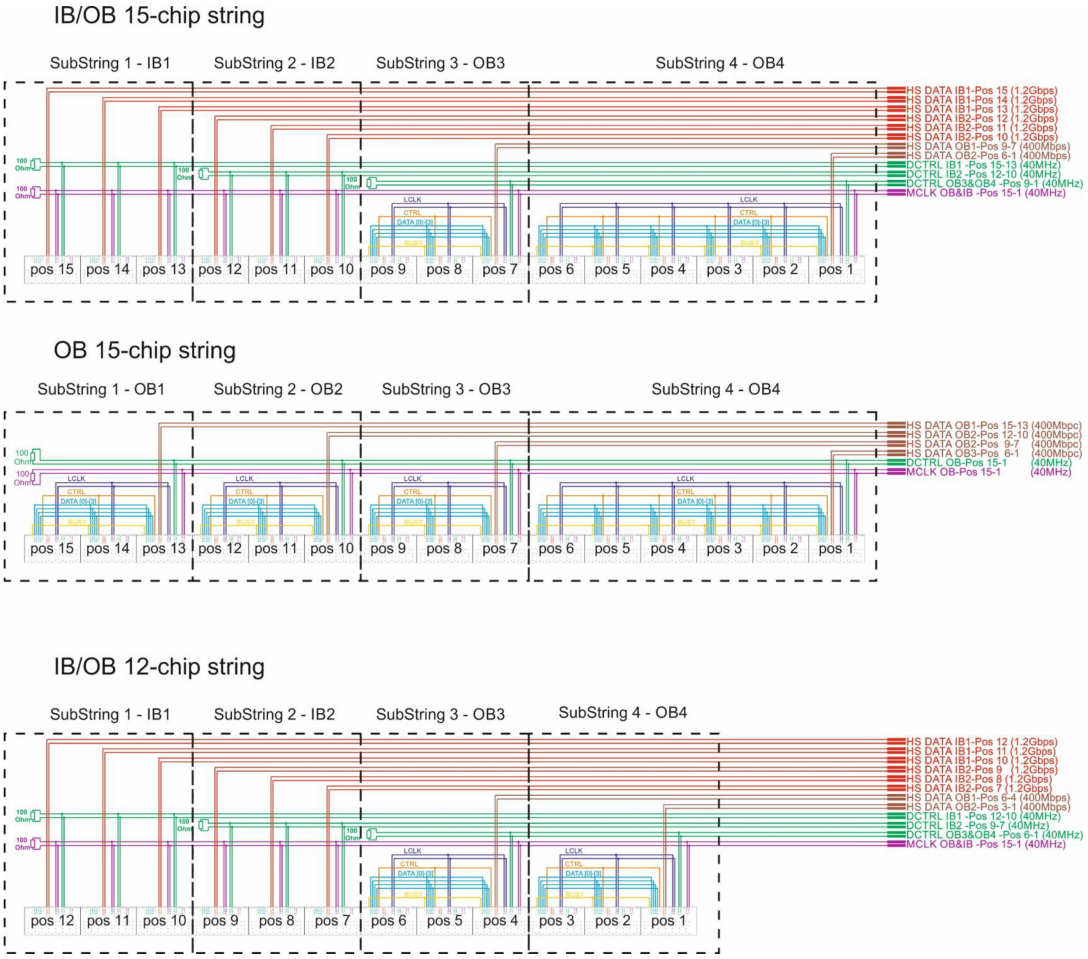


Fig. 21: Multi-chip string design and segmentation for the IB/OB 15-chip string (top panel), the OB 15-chip string (middle panel), and the IB/OB 12-chip string (bottom panel).

3.4 Simulation of readout rates and occupancy

3.4.1 Comprehensive overview and summary

The ALPIDE was designed for tracking, with significant lower occupancies than expected for calorimetry. In particular, when operating at higher interaction rates, this can lead to problems with the amount of data that needs to be read out, since locally more hits are recorded than can be buffered. Hence, a detailed simulation of the occupancy, detector response and readout rates of the pixel layers has been conducted in software. The simulation is performed by using minimum bias physics events generated in the FoCal physics simulation framework, a representation of the pixel string layout and layer geometry in C++, and an implementation of the digital part of the ALPIDE in "SystemC". In the following we present the main results of the simulation, while details are provided in the Sects. 3.4.2–3.4.7.

As a default scenario, pp interaction rates of 1 MHz were assumed. The readout of the ALPIDEs was implemented in a periodically-triggered mode with trigger frame lengths of $10\,\mu$s and a pulse length of $5\,\mu$s in the pixel front-end. The particle hit rate in the pixel layers is dominated by electromagnetic showers from π^0 decays and reaches close to $3.0\,\mathrm{MHz/cm^2}$ at the innermost radii (Fig. 22). With this readout scheme the average expected pile-up in the pixel layers is estimated to be 14 (within the tolerable range of up to about 20 pileup events). The average pixel cluster size (RMS) per particle hit is set to 4 (2), in agreement with results from test beams.

With these parameters, we obtain an average occupancy (Fig. 34) and data rates (Fig. 35) in the inner most ALPIDEs below 2.5% and below 370 Mbps, respectively from the simulation. We simulate maximum instantaneous data rates below 800 Mbps for the IB chips, well below the full link bandwidth of 960 Mbps, normalized to a reference time window of 10 μs. For the outer mode links, the instantaneous data rate reaches occasionally the full link bandwidth of 320 Mbps, but remains on average well below the link saturation (see Fig. 36). The overall data rate of both pixel layers amounts to roughly 75 Gbps.

We observe (Fig. 37) that for the default scenario the readout performance of the pixel layers would be significantly affected by so-called BUSY violations of the ALPIDEs in the inner regions. BUSY violations are an effective dead-time of the ALPIDE chip at high occupancy and high trigger rates, which originates from the matrix readout speed of the pixel matrix. In the pixel layers the BUSY violations originate from high-energetic electromagnetic showers which produce a high regional occupancy in the ALPIDEs. If on analysis level pixel layer frames without any BUSY violation were required, the pixel layers would provide an efficiency of only 20%, in terms of dead-time (Fig. 38).

We therefore consider strategies in order to mitigate the inefficiency originating from the BUSY violations. The key parameters investigated are the reduction of occupancy and trigger rate. The trigger rate can be reduced using a frame length of 20 μs, which however increases the pileup by roughly a factor 2 compared to a frame length of 10 μs. The pixel occupancy can be reduced by applying a back-bias voltage to the epitaxial layer of the ALPIDE, leading to smaller cluster sizes of presumably size ≈ 3. With these two changes of configuration, the efficiency of the pixel layers can be increased to about 85%, in terms of dead-time (Fig. 38).

For further mitigation, we simulate digital masking of pixels with a grid pattern of e.g. every fourth row and column in order to artificially reduce the occupancy by about 50% (which was checked not to significantly reduce the performance of the clusterizer). With all the configurations introduced above, pixel layer efficiencies above 75% are reached for a trigger frame a length of 10 μs, and above 95% for the 20 μs (Fig. 40). For the case of p–Pb collisions no large differences compared to the pp case are observed, and the proposed configurations mitigate the rate of BUSY violations in the same manner as for the pp case (Fig. 41).

3.4.2 Particle hit rates

To estimate the occupancy of the pixel layers, and to understand properties of background events, a total of 50,000 minimum-bias pp events at $\sqrt{s} = 14$ TeV were simulated with the PYTHIA event generator and processed with the FoCal simulation framework. The particle hit rates in the pixel layers 5 and 10 are shown in Fig. 22 for an average interaction rate of 1 MHz in pp collisions. The rates for the two layers are of the same order of magnitude, with layer 5 exhibiting ≈ 10 % higher hit rates. At the innermost radii the particle hit rate reaches nearly 3.0 MHz/cm^2. The hit rate drops quickly with increasing radius, to values below 10 kHz/cm^2 at the outermost positions.

The particle occupancy in layers 5 and 10 is dominated by electromagnetic showers: 2/3 of the particle hits are associated to photons from π^0 meson decays, as shown in Fig. 23. Other leading contributions to the pixel layer hit rates stem from decays of η ($\approx 6-7\%$) and ω ($\approx 2-3\%$) mesons. While the decay of π^0 to two gamma particles has a probability close to 100%, the dominant decay modes of η and ω include both neutral (π^0) and charged ($\pi^\pm$) pions, as well as γ; similarly for other hadronic decays. In layer 5 (10), 75% (70%) of the pixel hits originate from primary γ showers. A fraction of 14% (17%) originates from primary charged pions, $\pi^\pm$, which represent the second major source for pixel hits. Other primary particles, which contribute at most a few percent to the pixel hit rates, are K^0_S, K^0_L, $K^\pm$, p, n, $\overline{p}$, and $\overline{n}$. Primary electrons and positrons contribute only at the few per-mille level to hits in the pixel layers.

Figure 24 shows the energy spectrum of primary γ and $\pi^\pm$which generate hits in the pixel layers. The

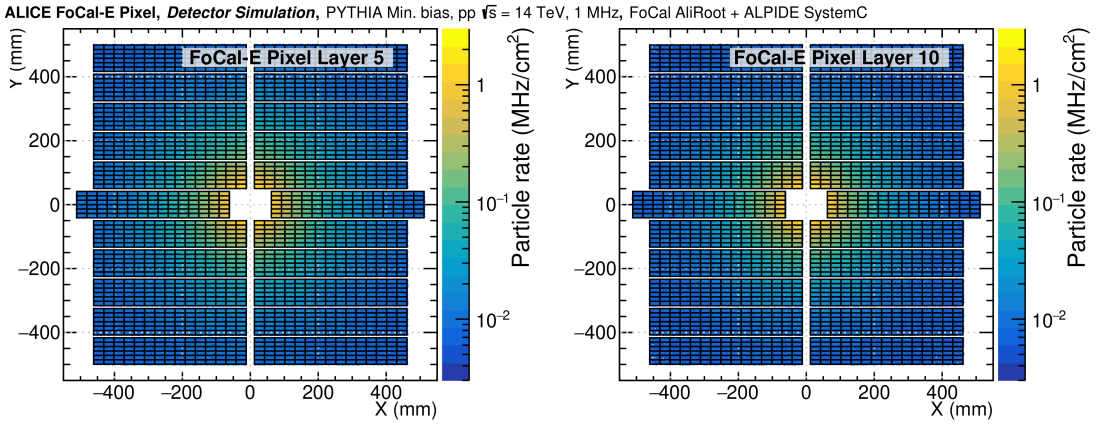


Fig. 22: Estimated particle hit rate in pixel layer 5 (left panel) and layer 10 (right panel) for an interaction rate of 1 MHz in pp collisions.

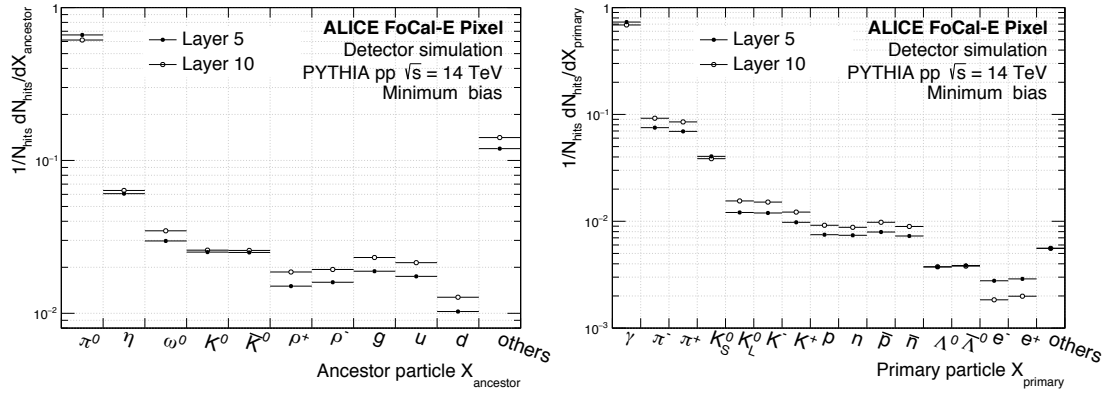


Fig. 23: Fraction of particle hits in layer 5 and 10, differentiated by the species of the ancestor particle precursor to the primary particle initiating the shower (left panel) and the associated primary particle (right panel).

occupancy is dominated by the products of low energy primaries. For electromagnetic showers from γ, around 90 % of the hits originate from electromagnetic showers with an energy below 25 GeV, and more than 40 % originate from showers with an energy below 5 GeV. The spectrum of charged pions is qualitatively similar to that of γ, extending about 100 GeVhigh in energy. Over 90 % of the pixel hits from primary charge pions originate from $\pi^{\pm}$ with an energy below 50 GeV, and around 40 % from energies below 10 GeV.

The cumulative contributions of all hadronic and electromagnetic primaries to the number of particle hits per event in the pixel layers is shown in Fig. 25. The number of pixel hits from hadrons follows an exponential distribution, while for the electromagnetic part high energy contributions appear at the end of the spectrum. On average 400 particles per interaction event traverse layer 5, and 330 particles traverse layer 10. About 75 % of the hits originate from electromagnetic showers.

3.4.3 Pileup and busy violations in the pixel layers

The default scenario for the operation of the pixel layers is the ALPIDE's triggered mode with a periodic trigger. In this mode the ALPIDEs receive an external trigger from the *Local Trigger Unit* (LTU) at a frequency f_{trg}, creating frames of constant time interval. The default frequency is $f_{\mathrm{trg}} = 100$ kHz, or equivalently $\Delta t_{\mathrm{frame}} = 1/f_{\mathrm{trg}} = 10$ μs. Figure 26 illustrates the periodically–triggered mode scheme. At the reception of a trigger, the strobe window of the ALPIDE is opened for the accumulation of pixel

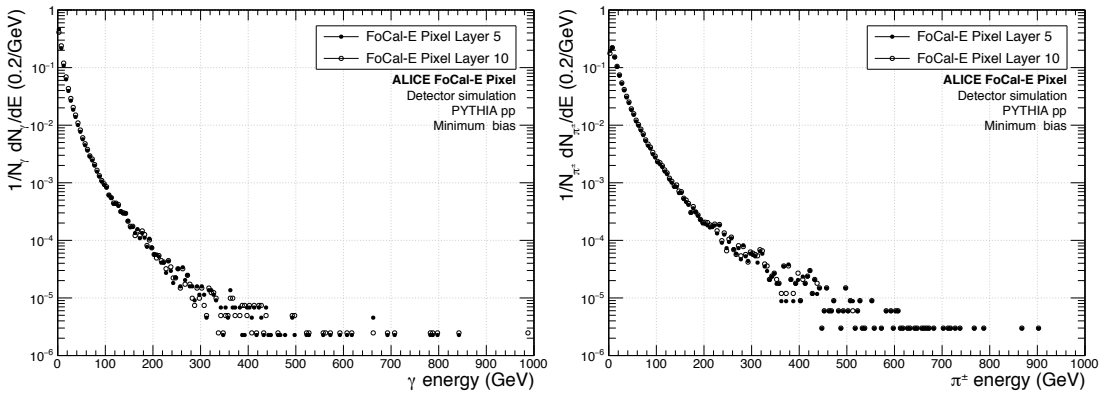


Fig. 24: Energy spectrum of the primary particles (left: γ, right: $\pi^{\pm}$), which particle hits in layer 5 and 10 are associated with.

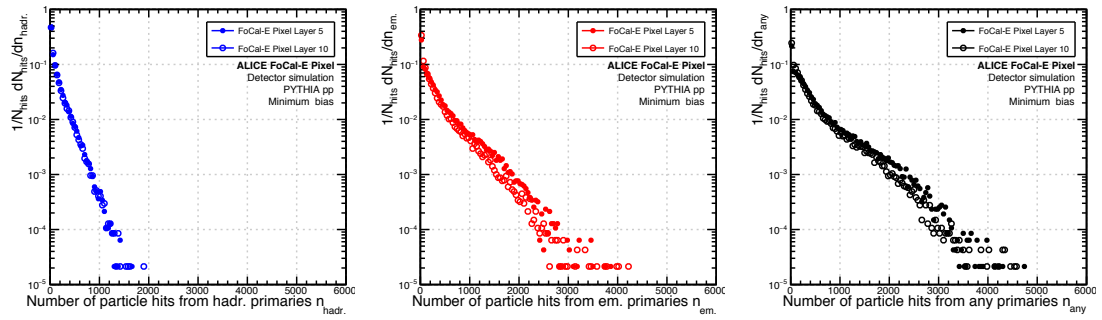


Fig. 25: Simulated number of particle hits per event in layer 5 and 10 for hadronic (left) and electromagnetic (middle) primaries, as well as the cumulative number of hits per event (right).

hits during the time Δt_{frame}. The time constant of the ALPIDE pixel front end is $t_{\text{pulse}} \approx 5\ \mu$s, meaning that once a pixel is activated by a hit of an ionizing particle it remains active (i. e. over the detection threshold) for this duration. When the pixel is active during the opening of the strobe window, the pixel hit is latched into the *Multi-Event Buffer* (MEB) and remains stored there until the matrix readout of the event has started.

If an event happens shortly before the trigger signal is issued, the pixel hits associated with that event will be latched twice, in the time frame the event happened in and in the next time frame, since the pixel front-end will be still over threshold. Thus, time frames can be contaminated with pileup events from the previous time frame. In order to avoid ALPIDE strobe extensions, the strobe window is closed for a short period ($\Delta t_{\text{strobe, inactive}} \lesssim 1\,\mu$s) before the arrival of the next trigger, during which no pixel hits are latched. Because of the relatively long pixel front-end time constant t_{pulse}, pixel hits which happen during $\Delta t_{\text{strobe, inactive}}$will still be over threshold at the start of the next time frame, and thus are not lost. It is even theoretically possible to operate the system with inactive strobe times of $\Delta t_{\text{strobe, inactive}} \approx 2\ \mu$s, which may be beneficial in terms of pile-up, occupancy, and readout speed.

Taking into account t_{pulse}, the mean number of pile-up events per readout frame is given by

$$N_{\text{pile-up}} \approx (\Delta t_{\text{pulse}} + \Delta t_{\text{frame}} - \Delta t_{\text{strobe, inactive}}) \cdot \frac{\mathrm{d}N_{\text{events}}}{\mathrm{d}t}\,. \tag{2}$$

For example, for a frame length Δt_{frame}=10 μs, a strobe inactive time $\Delta t_{\text{strobe, inactive}}$=1 μs, and an interaction rate $\mathrm{d}N_{\text{events}}/\mathrm{d}t = 1$ MHz, the estimated $N_{\text{pile-up}}$ amounts to 14.

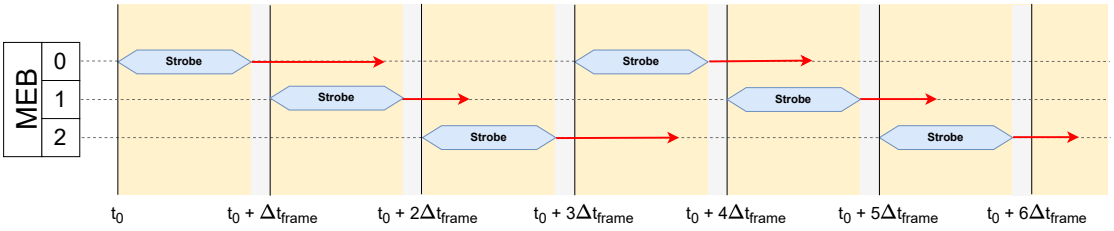


Fig. 26: Schematic illustration of the periodically–triggered readout mode, with a frame length of Δt_{frame}. The red arrows illustrate the time needed for completely reading out the pixel matrix.

The addressing and readout scheme of the ALPIDE pixel matrix is organized in regions. Along its column direction, the pixel matrix is divided into 32 regions. Each region contains 16 pixel double columns and 512 pixel rows, summing to 16,384 pixels per region. When the strobe window is closed, the readout of the 32 regions starts in parallel. For reading out one pixel, two clock cycles are needed, i. e. 50 ns. The duration Δt_{matrix} of the complete matrix readout is therefore determined by the number of pixel hits $N_{\text{region, max}}$ in the region with the maximum occupancy: $\Delta t_{\text{matrix}} = N_{\text{region, max}} \cdot 50\,\text{ns}$. For a maximum occupancy per region of more than 200 pixel hits ($N_{\text{region, max}} > 200$), the matrix readout takes $\Delta t_{\text{matrix}} > 10\,\mu\text{s}$, and reaches the scale of the foreseen frame lengths Δt_{frame}.

If the matrix readout takes longer than the frame length Δt_{frame}, the MEB is filled faster than it is emptied. The ALPIDE chip reports a BUSY flag when, during an open strobe window, the pixel hits are being latched to the last free position in the MEB, indicating that the trigger rate is faster than the readout rate, and hence the chip is busy. This situation is illustrated in Fig. 27.

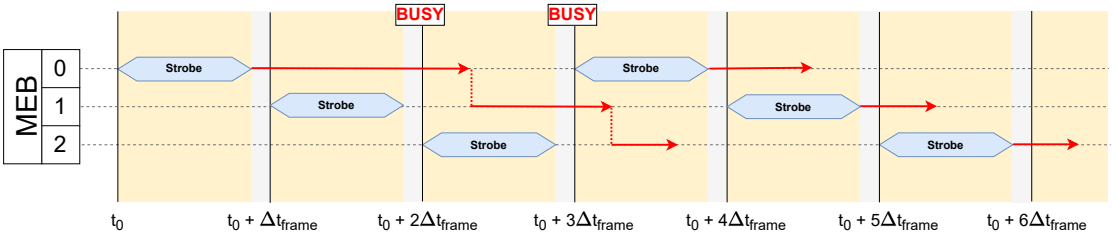


Fig. 27: Schematic illustration of the periodically triggered readout mode with a frame length of Δt_{frame} with the occurrence of BUSY flags if the current strobe is being latched into the last free position of the MEB. The red arrows illustrate the time needed for completely reading out the pixel matrix.

When, on reception of a trigger, a strobe window should be opened but there is no free slot in the MEB left to latch the pixel hits, the chip skips this trigger. This typically happens when the matrix readout of the event from three time frames earlier has not yet been finished. The pixel hits in this chip are entirely lost for the newly requested time frame. The ALPIDE chip reports this behavior with a BUSY violation flag. The event itself that provoked the busy violation is read out completely and is not corrupted. Once a slot in the MEB is freed, the chip is again ready to receive triggers and take data.

This BUSY violation principle is illustrated in Fig. 28. When a high–occupancy event is latched in two consecutive time frames, the pixel hits have to be read out twice. This may lead to a frame in BUSY violation already at lower occupancy, but the frame occurs one frame later than for the previous case (Fig. 29).

The problem of BUSY violations is related only to the regional matrix readout speed, and is independent of the data link speed. Thus it is technically identical for inner and outer mode chips. Event signatures which provoke busy violations are characterized by a high regional particle density, such as from en-

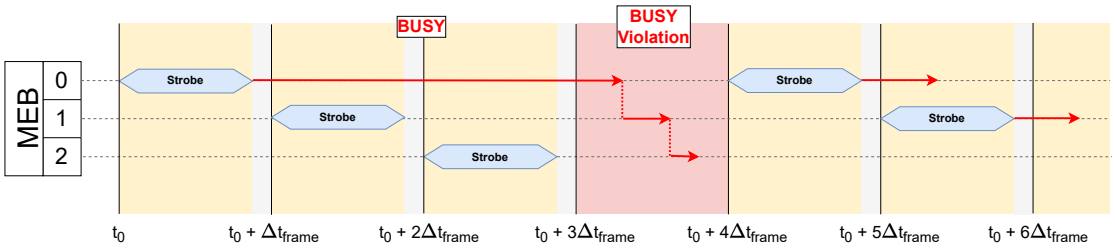


Fig. 28: Schematic illustration of the periodically triggered readout mode with a frame length of Δt_{frame} with the occurrence of BUSY violations, if the current strobe cannot be latched into a free position of the MEB. In this case the occupancy event producing the BUSY violation is latched once in the first frame. The time frame lost because of a BUSY violation is colored light red. The red arrows illustrate the time needed for completely reading out the pixel matrix.

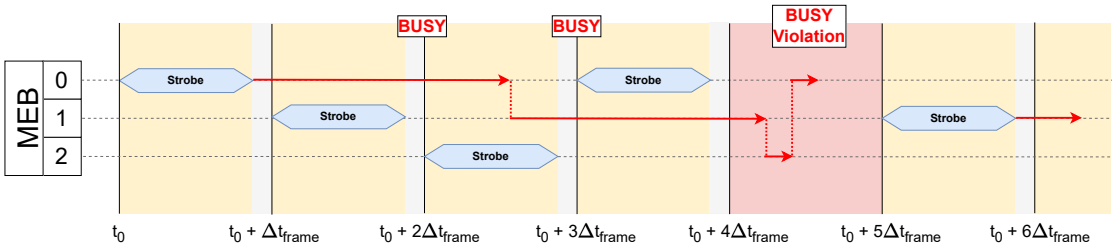


Fig. 29: Schematic illustration of the periodically triggered readout mode with a frame length of Δt_{frame} with the occurrence of BUSY violations if the current strobe cannot be latched into a free position of the MEB. In this case the occupancy event producing the BUSY violation is latched two times, in the first and in the second frame. The time frame lost because of a BUSY violation is colored light red. The red arrows illustrate the time needed for completely reading out the pixel matrix.

ergetic electromagnetic showers. These occur predominantly at high rapidities, and handling of BUSY violations is crucial at the innermost radii.

3.4.4 Measurements of BUSY violations

The mechanism of BUSY violations in the ALPIDE was validated in a series of measurements at SPS H6 in October 2021. Two prototype FoCal-E Pixel layers with ALPIDE chips in inner mode were tested in a prototype tungsten stack at the equivalent depth corresponding to layer 5 and layer 10. An 80 GeV beam with an estimated electron fraction of $\approx 1\,\%$ was used. The ALPIDE chips were operated in periodically triggered mode with a trigger frequency of $f_{\text{trg}} = 100\,\text{kHz}$ ($\Delta t_{\text{frame}} = 10\,\mu\text{s}$).

Figure 30 shows the number of events with BUSY flags or BUSY violations with respect to both the number of hits in the matrix slice that had the highest occupancy in the original event, and the theoretically calculated readout time of the matrix region. The mechanism was found to work as expected from the theoretical design description.

The events are classified in two categories: events which produce one BUSY flag (Fig. 30, left), and events which produce two BUSY flags (Fig. 30, right). The first category with one BUSY flag is directly related to Fig. 28. Events with a phase of arrival which is latched only in one readout frame generate a BUSY flag when the number of pixel hits in the maximum occupied matrix region exceeds 200, which corresponds to a matrix readout time of $\Delta t_{\text{matrix}} > 200 \cdot 50\,\text{ns}$, i. e. $> \Delta t_{\text{frame}}$ (Fig. 30, left). After events with > 400 hits per maximum occupied matrix region, the matrix readout time exceeds $\Delta t_{\text{matrix readout}} > 400 \cdot 50\,\text{ns}$, i. e. $> 2\Delta t_{\text{frame}}$, and a BUSY violation occurs. When the particle arrives in a phase such that the pixel front-end is over threshold in two consecutive readout frames, the pixels associated with

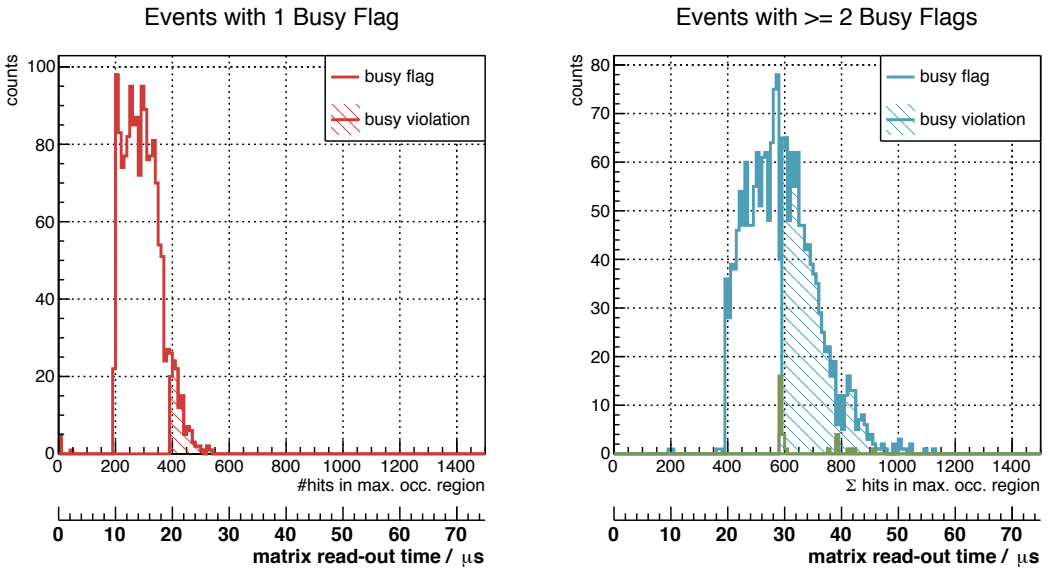


Fig. 30: Number of events with BUSY flags or BUSY violations with respect to the number of hits in the matrix slice that had the highest occupancy in the original event, and its theoretically calculated readout time. The events are classified into two categories: events which produce one BUSY flag + BUSY violation(s) (left panel), and events which produce two BUSY flags + BUSY violation(s) (right panel). The green histogram in the right panel refers to other, less likely BUSY flag and/or BUSY violation sequences not described in detail in this document.

this event are read out once in each frame, as depicted in Fig. 29. A BUSY violation occurs when the sum of the hits to be read out in the two frames exceeds 600, i. e. $\Delta t_{\text{matrix}} > 30\,\mu\text{s} = 600 \cdot 50\,\text{ns}$ (Fig. 30, right).

3.4.5 SystemC simulation framework

For the simulation of the pixel layer occupancy, trigger, readout, and busy rates, a C++ software framework is used that was originally developed for ITS2 [20] and was extended for the needs of the FoCal design [31]. The core of the framework simulates the response of the digital part of the detector front-end (i.e. the 264 strings with 15 ALPIDEs each), utilizing the SystemC [32] package. An overview of the modules and the workflow of the framework is shown in Fig. 31.

The modelled strings are contained in a string interface class which provides the interface to input and output of the model. Simulated minimum bias events, which were generated with the PYTHIA event generator (or any other event generator) and processed with FoCal simulation framework, serve as the input for the pixel hits in the model. An event generator class assigns a timestamp to the event following a randomized interaction probability model. The particle hits in the pixel layers are transformed to pixel hits in the chip internal coordinate system, and a pixel cluster generation model is applied for the formation of a realistic occupancy. A trigger generator class is used to trigger the readout of the system. One readout unit class per string interface, organizes the statistical analysis of the data. Data link parser classes analyse the payload on the data links, and readout unit statistic classes dump data to output files on disk.

After the digitization of the single particle hits in the detector plane, a cluster generator model is applied in order to produce realistic pixel cluster sizes for the simulation of the ALPIDE occupancy. By default, the pixel cluster size is simulated by a Gaussian probability distribution with mean cluster size of $n_{\text{cl. size}} = 4$ and a width of $\sigma_{\text{cl. size}} = 1.4$. More realistic cluster generation models, which are motivated and will be validated by measurements from test beam results, are being worked on.

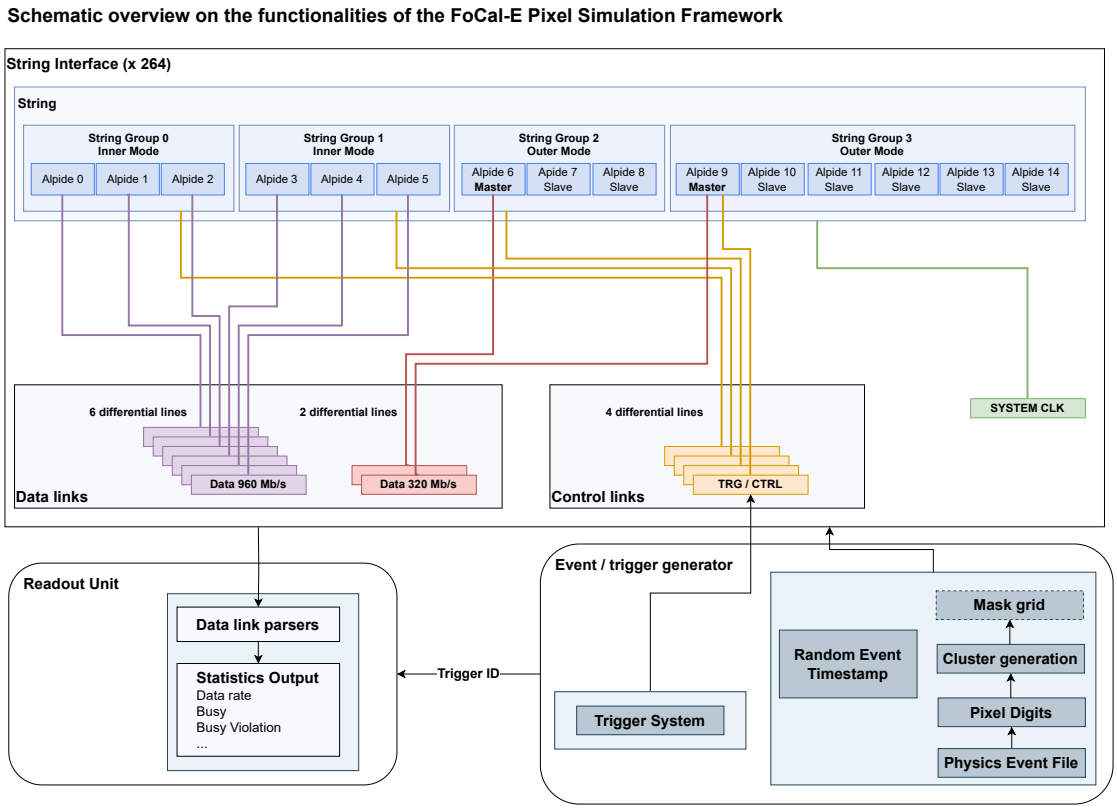


Fig. 31: Schematic overview of the SystemC simulation framework for the FoCal-E pixel layers. The model is fed by a trigger/event generator, and read out by a class which handles the readout statistics. The core of the framework is the digital simulation in SystemC of 264 pixel strings with 15 ALPIDEs each (here shown for an inner mode string).

Random event timestamps with a granularity of a bunch crossing (25 ns) are created from an exponential probability distribution function with a given mean interaction frequency (default 1 MHz). The generated pixel layer event is filled into a pixel hit container which holds the corresponding pixels active for the active time of the ALPIDE pixel front-end (typically 5 μs).

The model can be run in two different trigger modes: in a periodically triggered mode with a constant frequency (typically $f_{\mathrm{trg}} = 100\,\mathrm{kHz}$), or in a triggered mode where the trigger is derived from an external signal. This provides the possibility to implement e. g. a potential trigger from the FIT detector, or a FoCal self-triggered scheme where the trigger would be derived from the FoCal-E pad signal.

The model of the pixel strings describes the geometric arrangement of the strings, the data, trigger, and control links, as well as the functionality of the digital part of the ALPIDE. The ALPIDE string model is implemented providing all essential specifications of the final system. However, small differences exist in order to keep the framework as simple as possible.

Figure 32 shows the geometric arrangement of the inner and outer mode ALPIDE chips. Within a radius of $\eta \gtrsim 4.0$ the detector is equipped with IB chips, and outside of this radius with OB chips.

Figure 33 shows the implementation of the inner and outer string model into the framework. The inner string consists of six inner mode chips and nine outer mode chips, and the outer mode string contains only outer mode chips. Inner mode chips are readout with one data link per chip while the readout of the outer mode chips happens through the corresponding master chip. Since 8b/10b encoding is not implemented

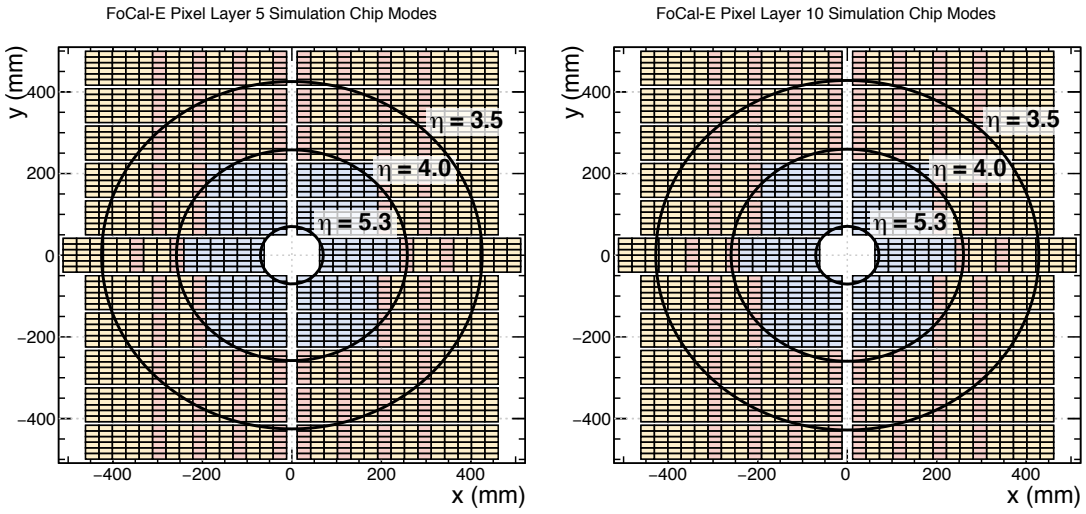


Fig. 32: Schematic view of the arrangement of inner mode chips (light blue), outer mode master chips (light red), and outer mode slave chips (yellow). The arrangement of the chip modes is identical for Layer 5 and 10.

in the framework, the inner mode (outer mode) data links are implemented with a bandwidth of 960 Mbps (320 Mbps) whereas the real link speed with 8b/10b encoding will be 1.2 Gbps (400 Mbps). Both types of strings are divided into four chip groups, three groups of three chips in the inner regions, and one group of six chips in the outer regions. Each group is served by an individual control and trigger link, making it possible to simulate regional trigger schemes with a granularity of a chip group. A "ReadoutUnit" class manages the parsing of the data links and produces result files which for each link contain the information of the ALPIDE and link response.

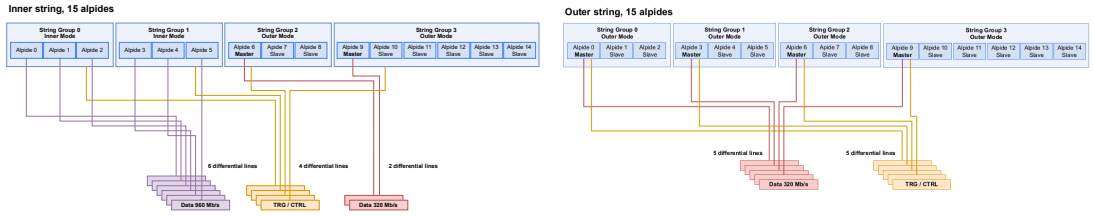


Fig. 33: Schematic view of the inner (left) and outer (right) string implementation in the pixel SystemC simulation framework.

If not mentioned differently, results from the simulations are presented for the following default parameters: Gaussian pixel cluster size distribution $n_{\text{cl. size}} = 4$ with $\sigma_{\text{cl. size}} = 1.4$, a time frame length $\Delta t_{\text{frame}} = 10\,\mu$s, a strobe inactive length $\Delta t_{\text{strobe, inactive}} = 0.1\,\mu$s, and for pp events at an interaction rate of $\mathrm{d}N_{\text{events}}/\mathrm{d}t = 1$ MHz.

3.4.6 Occupancy and data rates

The simulated pixel hit occupancy in the ALPIDEs per time frame is shown in Fig. 34. In the innermost chips an average occupancy of 2.2 % is expected for layer 5 (2.0 % for layer 10), which corresponds to roughly 10,000 active pixels per time frame. If the pixel hits are produced from electromagnetic or hadronic showers, the pixel hits will not be uniformly distributed across the pixel matrix, but will be — unlike for a tracking detector — concentrated in the shower cores where a high local hit density occurs. The chip occupancy drops quickly with increasing radius and is below 0.1 % for radii $\gtrsim 25$ cm.

The resulting average data rates for layer 5 and 10 are depicted in Fig. 35. For the innermost chips average data rates close to 400 Mbps are expected per link. As for the occupancy, the data rate per

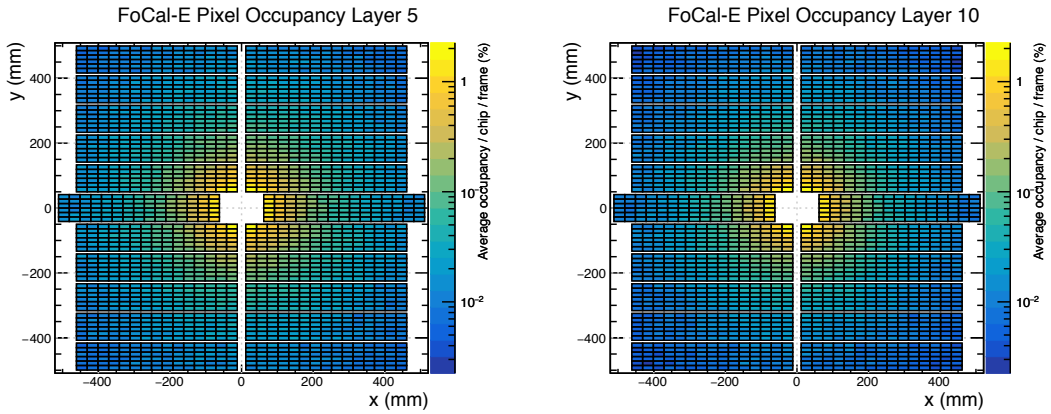


Fig. 34: Average occupancy for layer 5 (left) and 10 (right) for simulated pp events at $\sqrt{s} = 14$ TeV with 1 MHz interaction rate, $\Delta t_{\text{frame}} = 10\,\mu$s, and $\Delta t_{\text{strobe, inactive}} = 0.1\,\mu$s.

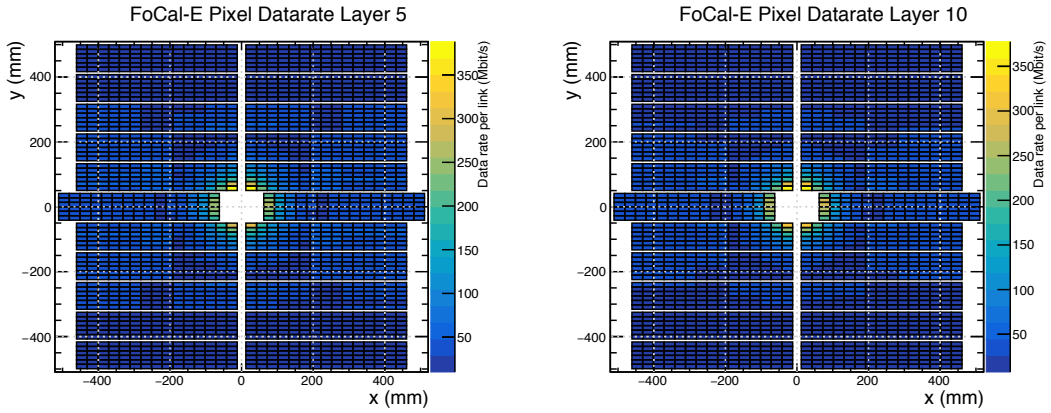


Fig. 35: Average data rates in layer 5 (left) and 10 (right) for same simulation settings as in Fig. 34. The average data rates are far below the actual speed links which are 960 Mbps (inner mode) and 320 Mbps (outer mode). While in the outer mode groups only the master chip drives readout, also the slave chips are filled with the color that corresponds to the data rate of the master chip.

link gets lower with higher radii. In the transition between inner and outer mode groups, the data rate increases a little bit since the master of an outer mode bus has to drive the readout of multiple chips. The simulated data rates are well below the link speed specification (960 Mbps IB and 320 Mbps OB).

Figure 36 shows the simulated data rates per link for the two pixel layers. The inner links will be operated with good margin (instantaneous data rate $< 10\,\%$ smaller than the maximum rate). The outer mode links are simulated to run occasionally on full bandwidth of 320 Mbps after having been hit by an high occupancy event. However the average link speed stays at good margin to the full bandwidth of the link. The total estimated data rate amounts to 35 ± 11 Gbps for layer 5 and 30 ± 10 Gbps for layer 10.

3.4.7 Busy violations and frame efficiency

From high–occupancy events, single chips drop readout timeframes because of BUSY violations (Sec. 3.4.3). This produces an effective inefficiency of the chips. From the simulation framework the fraction of frames in BUSY violation per chip is extracted. It can likewise be understood as the effective inefficiency or dead time of the chip. The fraction of frames in BUSY violation is shown in Fig. 37 for layer 5 and 10. For the default simulation settings is simulated to amount to an average of 5 to 7 % for

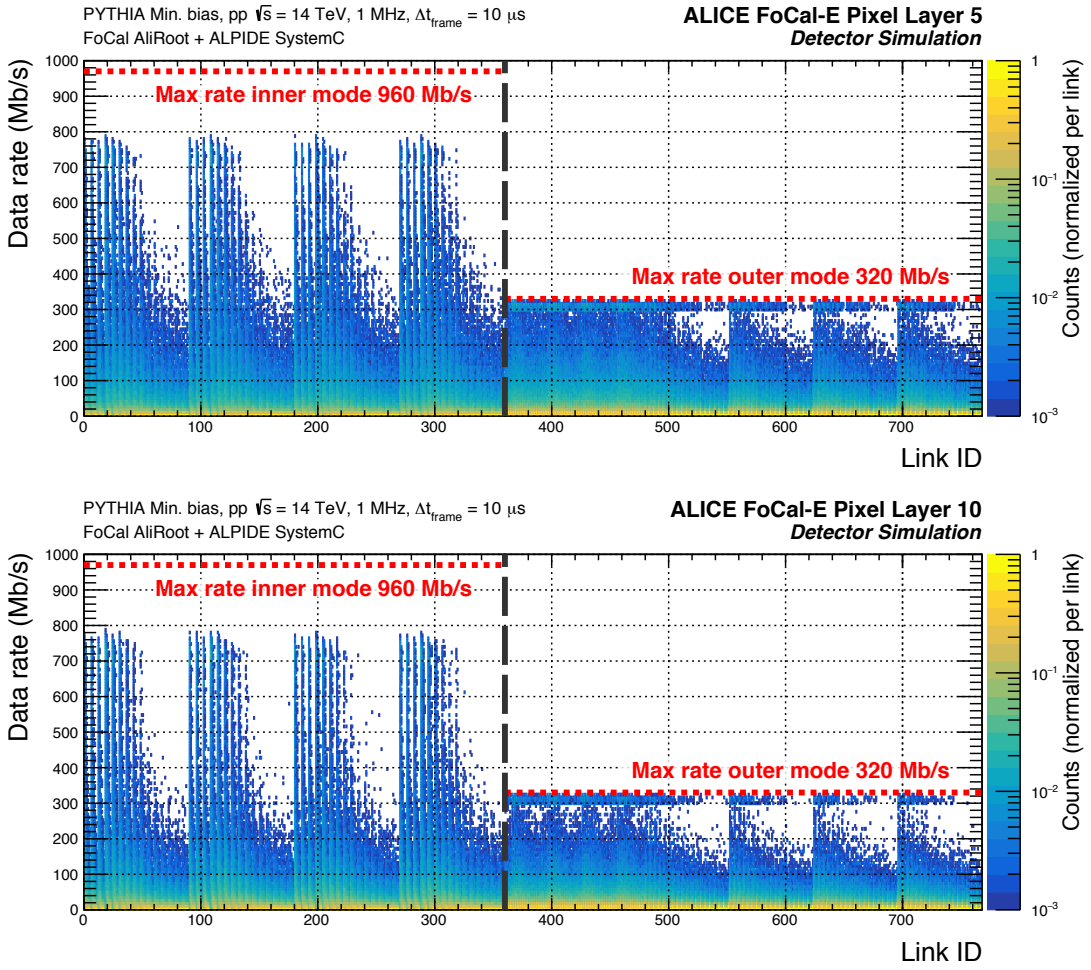


Fig. 36: Instantaneous data rates in layer 5 (top panel) and 10 (bottom panel) versus link ID for same simulation settings as in Fig. 34. The distribution left of the thick black line are links connected to ALPIDEs in IB mode, and the distributions on the right are links connected to master ALPIDEs in OB mode.

the inner most chips, and has values below 1 % outside a radius of 10 cm.

The dead time frames of single chips occur randomly and would lead to random local dead times in the detector. Thus, the signatures of the detector response for the pixel layers would show different and incomplete characteristics for every event frame, making the event reconstruction and handling very difficult at analysis level. It is therefore beneficial to produce complete frames, i. e. frames without readout loss. The fraction of frames with no BUSY violation in both layers is extracted to approx. 20 % for the default simulation settings.

In order to mitigate the effect of frame loss by BUSY violation, various options are being studied. By back biasing the ALPIDEs, the average pixel cluster size $n_{\text{cl. size}}$ can be reduced presumably from 4 to 3, thus reducing the overall occupancy in the pixel by roughly a factor 3/4. Increasing the time frame length provides more time for the pixel matrix regions to be read out, and reduces the number of events which are being latched twice since less triggers are sent. Extending the time frame length Δt_{frame} leads to an increment of pileup. A length of Δt_{frame}=20 μs is considered as an upper limit for the time frame because of limitations in the deconvolution of pileup events. A higher Δt_{frame} produces also higher chip occupancy. However, a significant increase of the pixel matrix occupancy in a specific region by another

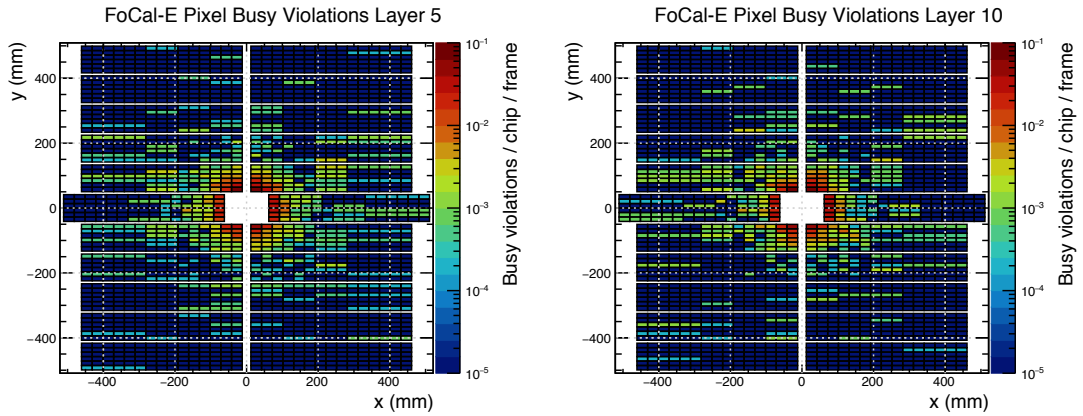


Fig. 37: Fraction of frames in BUSY violation per chip for the default simulation settings in pp-collisions.

.

shower core, which is the driving quantity for BUSY violations, remains unlikely. The reduction of the probability for latching pixel hit events in two time frames and an increased time for the pixel matrix region readout can also be provided by increasing the strobe inactive length time $\Delta t_{\text{strobe, inactive}}$ up to a reasonable high value (up to $\approx 2\,\mu$s).

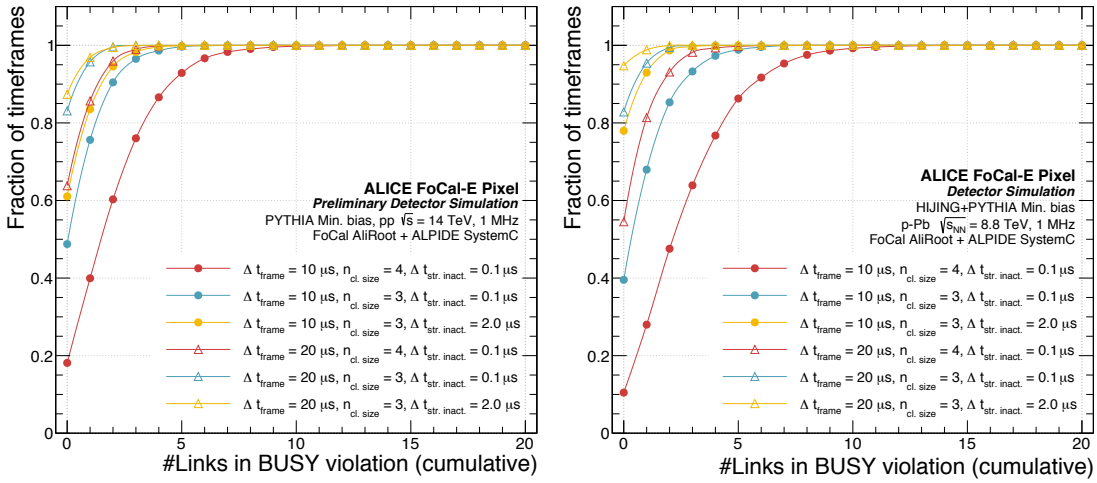


Fig. 38: Fraction of frames with the cumulative number of links in BUSY violation, i. e. #Links. The cumulative curves show the probability to encounter a time frame with a maximum of #Links in BUSY violations. The left figure shows the simulation for running in pp collisions, and the right figure for running in p-Pb collisions.

Figure 38 (left) shows the fraction of frames with the cumulative number of links in BUSY violation for various settings Δt_{frame}, $n_{\text{cl. size}}$, and $\Delta t_{\text{strobe, inactive}}$in pp-collisions. For the default settings ($\Delta t_{\text{frame}} = 10\,\mu\text{s}, n_{\text{cl. size}} = 4, \Delta t_{\text{strobe, inactive}} = 0.1\,\mu\text{s}$), 18 % of the frames are not affected, by any BUSY violation, and 40 % of the frames contain only at maximum one BUSY violation in either pixel layer. The fraction of frames which are not affected by BUSY violations is increased significantly by operating with a smaller mean cluster size ($n_{\text{cl. size}}$= 3) and an $\Delta t_{\text{strobe, inactive}}$of up to $2\,\mu$s. With these settings the fraction of frames without any BUSY violation is increased up to 61 %, and the fraction with maximum one BUSY violation to 83 %. With a longer timeframe window length Δt_{frame}, the efficiency can be increased significantly. For the default $n_{\text{cl. size}}$ and $\Delta t_{\text{strobe, inactive}}$ 64 % of all timeframes are not affected by any BUSY violation, and 86 % contain at most one BUSY violation. With $n_{\text{cl. size}} = 3$ and $\Delta t_{\text{strobe, inactive}} = 2\,\mu$s, a fraction of frames without any BUSY violation of 87 % can be reached, and 97 % of the frames

Table 5: Calculated pile-up and simulated global frame efficiency for various detector operation parameters with the pp simulation dataset.

Varied simulation parameters				Result	
Δt_{frame}	$n_{\text{cl. size}}$	$\Delta t_{\text{strobe, inactive}}$	$d_{\text{grid mask}}$	Pile-up	Global frame efficiency
10 μs	4	0.1 μs	0	14	18.1 %
10 μs	3	0.1 μs	0	14	48.8 %
10 μs	3	2.0 μs	0	13	61.0 %
20 μs	4	0.1 μs	0	24	63.8 %
20 μs	3	0.1 μs	0	24	83.1 %
20 μs	3	2.0 μs	0	23	87.4 %
10 μs	4	0.1 μs	4	14	46.4 %
10 μs	3	0.1 μs	4	14	74.0 %
10 μs	4	0.1 μs	3	14	51.7 %
10 μs	3	0.1 μs	3	14	77.6 %
20 μs	4	0.1 μs	4	24	86.6 %
20 μs	3	0.1 μs	4	24	94.9 %
20 μs	4	0.1 μs	3	24	87.8 %
20 μs	3	0.1 μs	3	24	96.2 %

Table 6: Calculated pile-up and simulated global frame efficiency for various detector operation parameters with the p-Pb simulation dataset.

Varied simulation parameters				Result	
Δt_{frame}	$n_{\text{cl. size}}$	$\Delta t_{\text{strobe, inactive}}$	$d_{\text{grid mask}}$	Pile-up	Global frame efficiency
10 μs	4	0.1 μs	0	14	10.5 %
10 μs	3	0.1 μs	0	14	39.5 %
10 μs	3	2.0 μs	0	13	78.0 %
20 μs	4	0.1 μs	0	24	54.5 %
20 μs	3	0.1 μs	0	24	82.8 %
20 μs	3	2.0 μs	0	23	94.7 %
10 μs	4	0.1 μs	4	14	65.7 %
10 μs	3	0.1 μs	4	14	86.2 %
10 μs	4	0.1 μs	3	14	71.0 %
10 μs	3	0.1 μs	3	14	87.5 %
20 μs	4	0.1 μs	4	24	93.4 %
20 μs	3	0.1 μs	4	24	98.0 %
20 μs	4	0.1 μs	3	24	94.2 %
20 μs	3	0.1 μs	3	24	98.2 %

contain at maximum one link in BUSY violation. The values described above are summarized in Tab. 5. From the variation of the timeframe window Δt_{frame} and the inactive strobe length $\Delta t_{\text{strobe, inactive}}$, the pile-up is estimated to reach values of 24 for $\Delta t_{\text{frame}} = 20\,\mu$s which is also summarized in Tab. 5.

The performance in terms of BUSY violations is simulated to be slightly worse for running in p–Pb collisions, and for the standard configuration in pp-collisions a frame efficiency of only 10.5 % is simulated. However, with the same strategy as the one used for pp, a similar or even better performance can be reached, compared to the one in pp-collisions. The time frame efficiency values corresponding to p–Pb for same parameters as for pp-collsions are summarized in Tab. 6 and shown in Figure 38 (right). Increasing the time frame length $\Delta t_{\text{frame}}$to 20 μs, a smaller pixel cluster size of 3, and a longer inactive strobe length $\Delta t_{\text{strobe, inactive}}$lead to higher values of the frame efficiency. In particular, a frame efficiency higher than 94 % is simulated for Δt_{frame}= 20 μs, $n_{\text{cl. size}}$= 3, and $\Delta t_{\text{strobe, inactive}}$= 2 μs.

Most of the incomplete frames in the pixel layers originate from BUSY violations in the innermost chips at radii $< 10\,\text{cm}$ because of the high occupancy in this very forward detector region. It is therefore considered to artificially reduce the occupancy in those chips by using only a subset of the available pixels. This can be achieved by digitally masking a subset of the pixel matrix with a periodic pattern, e. g. as shown in Fig. 39.

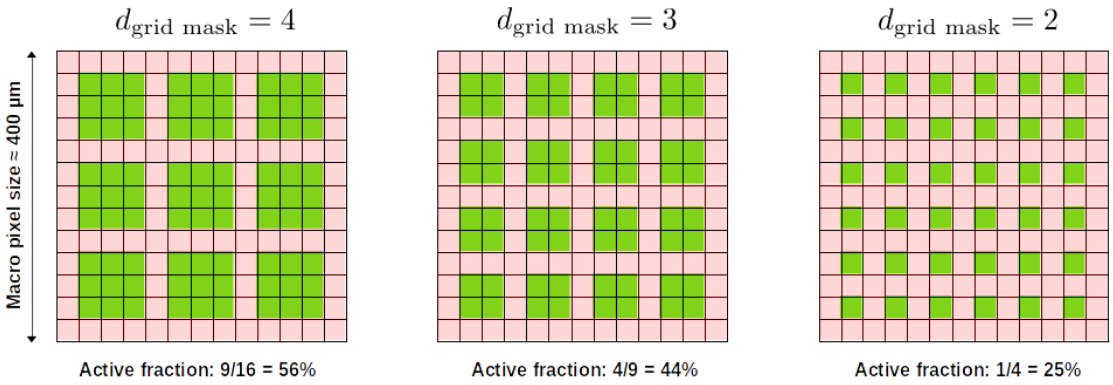


Fig. 39: Simulated grid masks for the inner most chips at radii $< 10\,\text{cm}$. The parameter $d_{\text{grid mask}}$ represents the size of the elementary cell of the applied grid.

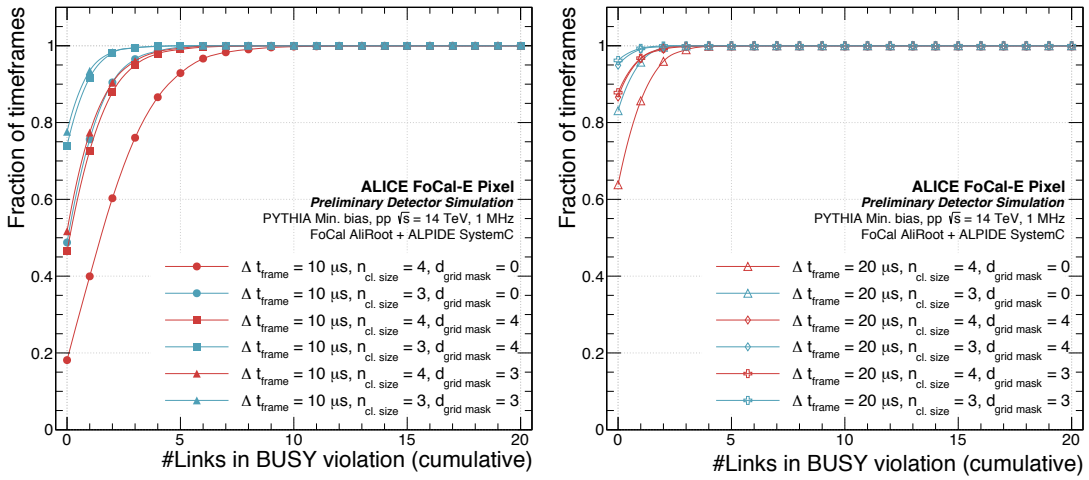


Fig. 40: Fraction of frames with the cumulative number of links in BUSY violation for pp collisions, i. e. #Links with grid mask of $d_{\text{grid mask}} = 4$ and $d_{\text{grid mask}} = 3$, with two different timeframe lengths $\Delta t_{\text{frame}} = 10\,\mu\text{s}$ (left) and $\Delta t_{\text{frame}} = 20\,\mu\text{s}$ (right). The cumulative curves show the probability to encounter a timeframe with a maximum of #Links in BUSY violations.

The parameter $d_{\text{grid mask}}$ denotes the size of the repetitive pattern. In this masking pattern scheme, the active pixels used for data taking can be reduced to 56 %, 44 % and 25 % for the cases where the $d_{\text{grid mask}}$ is 4, 3 or 2, respectively. With such masking patterns applied, some information about the micro structure of the shower would be lost. However, at the analysis level, e. g. for π^0 separation, macro pixel cells with a size of several 100 μm will be utilized which means that the absolute position of individual fired pixels can be neglected. Ongoing analyses are studying the impact of the reduced number of fired pixels on the physics performance. Since high occupancy events happen mostly in the inner region, the grid masks are applied in the simulation to the very inner ALPIDE chips only at radii $\lesssim 10\,\text{cm}$. With this technique, the frame efficiency can be further improved. Figure 40 (left) shows the simulated frame efficiencies for pp-collisions with the grid masks applied for $\Delta t_{\text{frame}} = 10$ μs (left) and $\Delta t_{\text{frame}} = 20$ μs (right). For the case

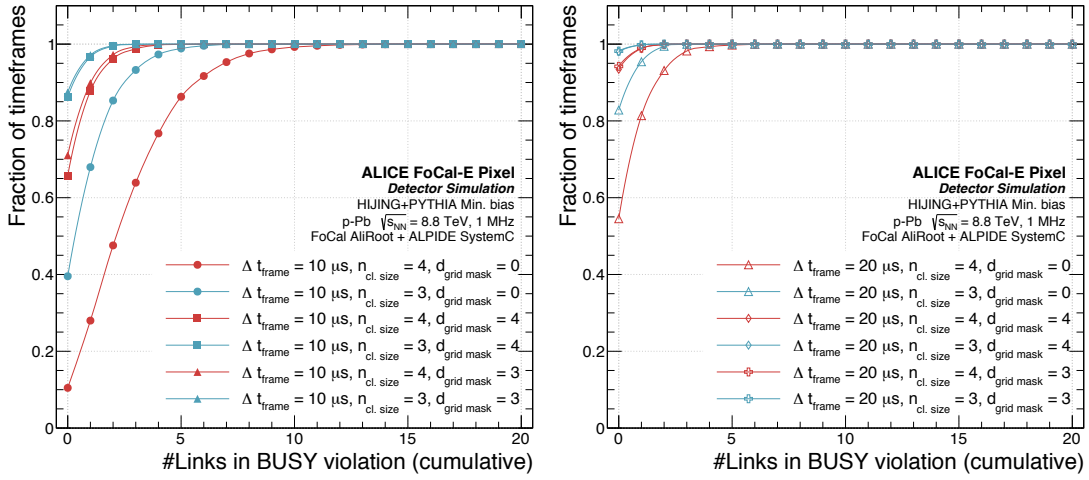


Fig. 41: Fraction of frames with the cumulative number of links in BUSY violation for p–Pb collisions, i. e. #Links with grid mask of $d_{\text{grid mask}} = 4$ and $d_{\text{grid mask}} = 3$, with two different timeframe lengths $\Delta t_{\text{frame}} = 10\,\mu$s (left) and $\Delta t_{\text{frame}} = 20\,\mu$s (right). The cumulative curves show the probability to encounter a timeframe with a maximum of #Links in BUSY violations.

of $\Delta t_{\text{frame}} = 10\ \mu$s and reduced pixel cluster sizes, frame efficiencies of about 75% can be reached. With $\Delta t_{\text{frame}} = 20\ \mu$s, frame efficiencies in the order of 95% are achieved. For p–Pb-collisions the maximum simulated frame efficiencies with $n_{\text{cl. size}}$= 3 and the grid masks $d_{\text{grid mask}}$ of 3 or 4 applied are 86 - 87 %, as shown Figure 41 (left). With the longer time frame, Δt_{frame}= 20 μs, frame efficiencies in the order of or higher than 95 % are simulated with the grid mask applied, as shown Figure 41 (right).

3.5 Readout architecture

Figure 42 presents the pixel layer readout architecture in more detail than Fig. 16. The left panel shows half of a pixel plane, which consists of 6 OB pixel layers and 5 IB/OB pixel layers. The pixel layers connect to TCs which work as intermediate boards for providing power to the ALPIDEs and connection (data, clock, control) between ALPIDEs and the off-detector electronics. If the length of this connection needs to be extended over 8 m, active repeater boards can be inserted. One RU controls and reads out one IB/OB pixel layer or 3 OB pixel layers.

In the right panel of Fig. 42, the full off-detector readout chain representing the readout of all pixel layers is shown. RUs are controlled and read out by CRUs upgraded with the lpGBT interface and the VTRx+. Trigger, timing, and control information are supplied from the LTU directly as implemented by ITS2. Alternatively, it can be supplied from the CRU as is the ALICE baseline mode for continuous data-taking mode.

3.5.1 Pixel module-layers

The design of pixel strings has been introduced in Sec. 3.3.5. As mentioned before, each pixel module-layer contains six multi-chip strings. This section focuses on the number of input and output lines for the readout of both pixel layers. The number of lines and their specification is summarized in Tab. 7. The total number of *Low Voltage Differential Signaling* (LVDS) lines is 72 for an IB/OB layer and 36 for an OB chip string. Note that there is no difference in the number of lines between the 15-chip and 12-chip pixel strings.

The structure of the IB and OB pixel strings is shown in Fig. 43(a). The 6 innermost chips of an IB string are of IB type ALPIDEs, implying a bandwidth of 1.2 Gbps for the offloading datalink. The next 9 chips are in OB mode with 2 of them set up as masters. The bandwidth of the OB master datalink is 400 Mbps, and it supports 3 and 6 OB type chips, respectively. This division is made since it is more likely to have

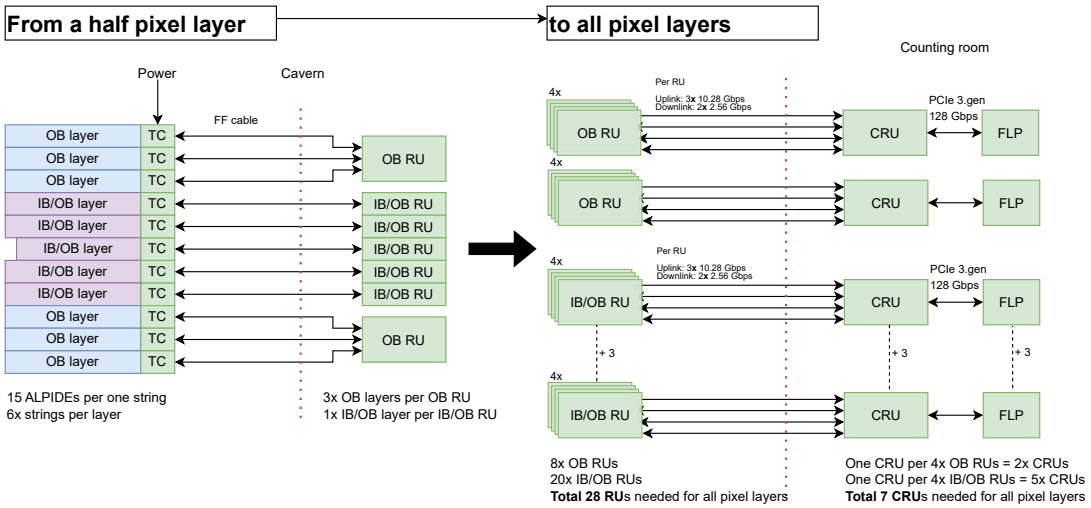
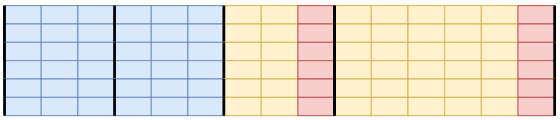
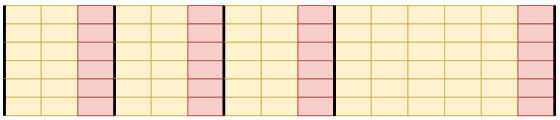


Fig. 42: Left: Readout architecture for half of a pixel layer. Right: The full off-detector readout chain.



(a) IB/OB pixel module-layer



(b) OB pixel module-layer

Fig. 43: The two flavours of pixel strings. Color coding of ALPIDEs: Light blue for IB, yellow for OB slave, light red for OB master.

more hits on the innermost part than the outermost part, implying that the data volume supported by the 2 OB Masters is similar. For more about ALPIDE and its modes see Sec. 3.2.

Figure 43(b) shows the structure of the OB pixel string. This is only equipped with OB type ALPIDEs, and an OB string consists of 4 groups, with the 3 innermost configured as 1 master and 2 slave chips, while the outermost group is configured as 1 master with 5 slave chips. This gives a conformity in the layout that matches the IB string. The OB pixel string are used the farthest from the beam pipe in the vertical direction. For the detailed design see Fig. 21.

On the ALPIDEs, a special short word supplied over the control line provides a trigger to the chip. To enable the possibility of 3 trigger regions on a pixel string, 3 control lines are used per IB/OB pixel string. These are set up in groups of 3 + 3 + 9 ALPIDEs from the inner to the outer part. For a full IB/OB pixel module-layer, this enables the possibility of having a maximum of 18 separate trigger regions since the module-layer is comprised of 6 individual pixel strings with 3 regions each. Each trigger region can then be individually triggered based on the registered amplitudes in the pad layers. For the OB pixel strings, all ALPIDEs on one string share a common control line, i.e. one string is one trigger region. This gives a maximum of six individual trigger regions per OB pixel module-layer in those areas with low expected hit density.

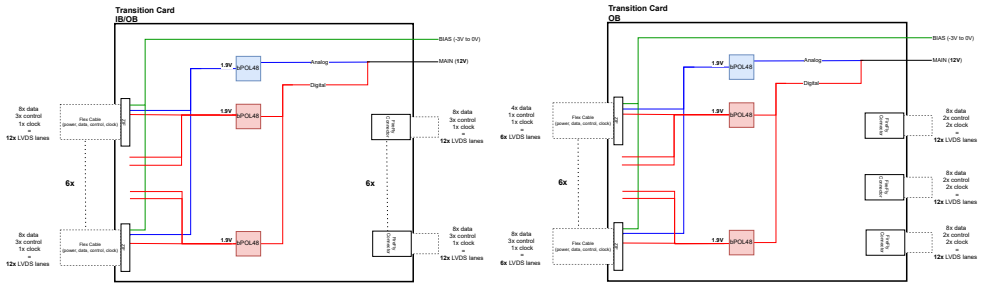
A summary of the different lines on the different pixel layers is given in Tab. 7.

Table 7: Number of lines per IB/OB and OB pixel layers.

Mode	Description	Link	Port	Number
IB/OB pixel layer	LVDS, 100 Ω, high-speed data IB	1200 Mbps	Input	36
	LVDS, 100 Ω, high-speed data OB	400 Mbps	Input	12
	LVDS, 100 Ω, control	40 MHz	Bidirectional	18
	LVDS, 100 Ω, clock	40 MHz	Output	6
				72
OB pixel layer	LVDS, 100 Ω, high-speed data OB	400 Mbps	Input	24
	LVDS, 100 Ω, control	40 MHz	Bidirectional	6
	LVDS, 100 Ω, clock	40 MHz	Output	6
				36

3.5.2 Transition card

The TC is an intermediate PCB that connects one pixel module-layer to the off-detector electronics and it provides power to the layer. The data, clock, and control lines from the pixel sensors are directly routed between the *Zero Insertion Force* (ZIF) connectors on the detector side and *FireFly* (FF) (copper cable assembly solution from Samtec) on the off-detector (readout) side. The bPOL48 converts voltage down from 12 V down to 1.9 V for digital and analog domains separately. There will be one bPOL48 for the analog domain for all strings and two for the digital domain (one per side). The details of the bPOL48 are described in Sec. 4.4.4.



(a) IB/OB configuration (b) OB configuration

Fig. 44: Diagrams of the TC.

However, the number of LVDS lanes is different for the two different pixel layers: IB/OB and OB. It is in fact double the amount of lanes for IB/OB layer. For this, there must be 6 FF connectors on the TC (see Fig. 44(a)) and only 3 for the other configuration (see Fig. 44(b)). This implies that there will be two designs for the TC, but the differences are only related to routing of the signals from the zif-connectors to the FF connectors. Since the designs are so similar, the risk and effort attached to this is regarded to be very low. Also, this should not affect the number of FF cable assemblies needed. In addition, several versions of this PCB have already been produced and are being successfully used in test benches and test beams.

The total number of FF copper cable assemblies needed is **192**. The total number of power cable assemblies is **44** assuming that one assembly contains a wire for main power, ground, bias, and a sense wire.

3.5.3 Repeater board

The preferred position of the RUs is relatively close to the FoCAL detector. This means that is is not foreseen that the FF cables are longer that what is currently in use for the ITS2.

However, in the unlikely case that a longer distance is needed, an active repeater board is considered as an option between the transition boards and the readout units. In the first prototype, showed in Fig. 45,

the datalinks (up to 1.2 Gbps) are repeated via active LVDS buffers (Texas Instruments DS25BR440), while the clock and the control links are forwarded passively.

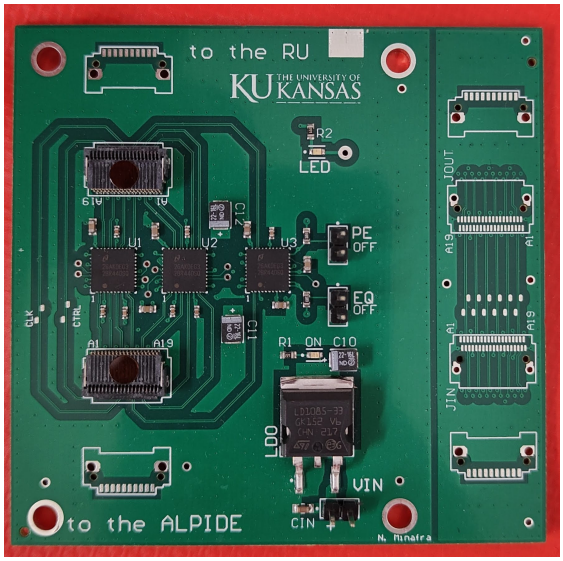


Fig. 45: Prototype of a repeater board currently under test.

Clock and control lines are not repeated since, working at 40 MHz, they can safely be transmitted to larger distances. It was demonstrated that those lines can be transmitted up to 18 m without active repeaters. More tests are needed for the high-speed lines. Furthermore, the LVDS buffers have to be qualified to work in the radiation environment, however, this particular *Integrated Circuit* (IC) was selected based on radiation tests performed on slower models of the same family. Production of these cards will cause delay in the project as well as a slight increase in cost, see table Tab. 38.

3.5.4 Readout unit

A sketch of the RU for the pixel layers in FoCal is shown in Fig. 46. It is designed based on the *Inner Tracking System for Run 3* (ITS2) RU. A few adaptations are needed to match the layout of the FoCal detector layout and requirements, and to replace components that are now out of production.

One RU will read out and control either one IB/OB pixel module-layer consisting of six strings, or 3 OB pixel module-layers with 3 times 6 strings. Reusing the ITS2 solution, the IB ALPIDE datalinks will be connected to dedicated transceiver pins on the FPGA, while the OB master ALPIDE datalinks are connected to LVDS IOs on the FPGA. The FPGA needs then to support 36 transceiver links and 72 LVDS signals for the front-end communication. Additionally, 18 clock signals (40 MHz) and 54 bidirectional control links are envisaged.

To facilitate the two variants of the pixel layers, it is envisaged to design two flavours of a sensor IO mezzanine board. One supports the IB/OB configuration, and the other supports the 3 OB configuration. All signals will be routed to the mezzanine connector, while only the signals required for the given configuration are routed to the firefly connectors on the mezzanine. This is similar to the ITS2 solution to support both ITS and MFT flavours of the ITS2 RU.

The communication to the CRUs [33] in the counting room is handled by lpGBTs and VTRx+. Three lpGBT chips in the 10.28 Gbps *Forward Error Correction* (FEC)12 configuration is sufficient to match the bandwidth from the detector-side of the RU based on the simulations.

The reception of clock and trigger info from the trigger, timing and control system can either be directly from an LTU (like ITS2) or via the CRU using the lpGBT and the VTRx+. Either solution does not have

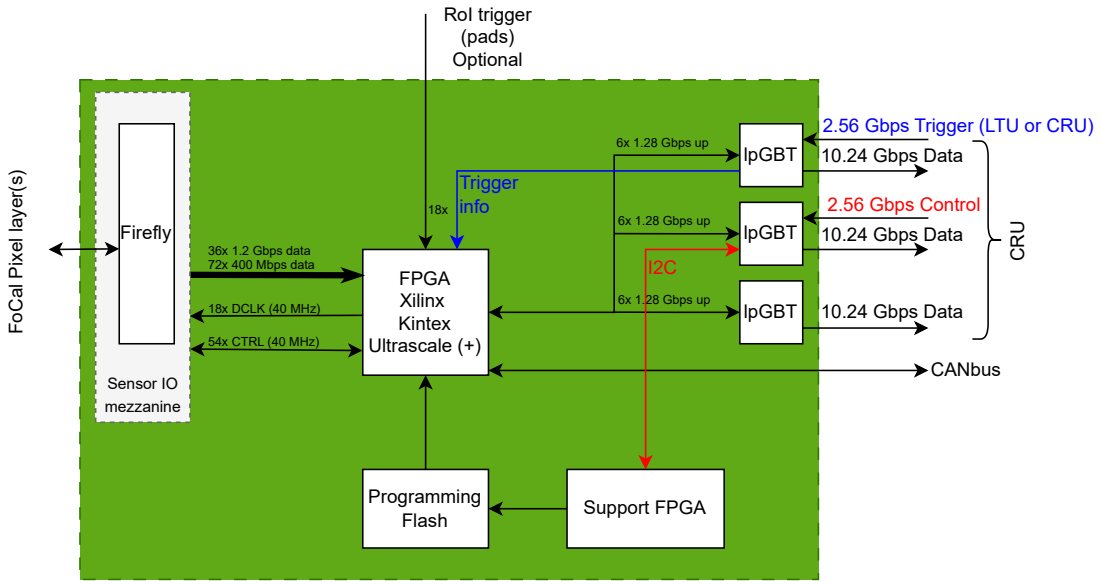


Fig. 46: Simplified sketch of the Focal pixel layer RU.

any impact on the hardware and the decision can be made at a later stage. As an option to support the local trigger from the pad layers, up to 18 LVDS signals, one per potential trigger region for an IB/OB pixel layer, can be provided from a nearby trigger board (see Sec. 7.3).

The suggested FPGA is a Xilinx Kintex Ultrascale KU085. This is the bigger brother of the KU060, used on the ITS2 RU. A test build was done for the pCT project in Bergen using 108 readout lanes, and this used about 40% of the logic resources available in the FPGA. For comparison, the ITS2 firmware has 37 readout lanes. This indicates that this FPGA has sufficient amount of internal resources, LVDS IOs and transceivers available to support the design described above. The programming and scrubbing solution of the *Static Random-Access Memory* (SRAM)-based FPGA is directly copied from ITS2. However, a different flash device must be sourced due to availability reasons. Also copied from ITS2 is the *Controller Area Network* (CAN) bus link towards the *Detector Control System* (DCS), acting as a backup communication path. The availability of the other components on the ITS2 RU was verified.

Table 8: Pixel power estimate.

	Inner Barrel	Outer Barrel Master	Outer Barrel Slave	Comment
Number of ALPIDEs				
15-chip IB/OB string	6	2	7	96x strings
15-chip OB string	-	4	11	144x strings
12-chip IB/OB string	6	2	4	24x strings
In total	720	816	2352	3888 ALPIDEs in total
Power consumption of ALPIDEs				
Per ALPIDE	0.2 W	0.2 W	0.1 W	
15-chip IB/OB string			2.3 W	$\approx$ 220 W for all strings
15-chip OB string			1.9 W	$\approx$ 270 W for all strings
12-chip IB/OB string			2.0 W	$\approx$ 48 W for all strings
				$\approx$ 550 **W in total**

3.5.5 Power estimate and distribution

Table 8 summarizes that the total power consumption of all ALPIDEs in FoCal will be 550 W. As baseline, the power distribution will be done following the pCT project. The same powering solution as for the rest of FoCal will be used, namely the bPOL48 (discussed later in Sec. 4.4.4).

3.5.6 Summary - pixel layers in numbers

- **3888** ALPIDEs - 720 in IB mode, 816 in OB Master mode and 2352 in OB Slave mode;
- ALPIDEs will consume **550 W**;
- **264** multi-chip strings - 96 of 15-chip IB/OB strings, 144 of 15-chip OB strings and 24 of 12-chip IB/OB strings;
- **44** pixel module-layers - 16 of 15-chip IB/OB layers, 24 of 15-chip OB layers and 4 of 12-chip IB/OB layers;
- **648** data links (LVDS) - 1.2 Gbps;
- **792** data links (LVDS) - 400 Mbps;
- **504** control links (LVDS) - 40 MHz;
- **264** clock links (LVDS) - 40 MHz;
- **28** RUs where 8 have the OB sensor IO mezzanine board and the other are IB/OB;
- **7** CRUs where 2 are dedicated to OB RUs and 5 are dedicated to IB/OB RUs.

3.6 Assembly

The separate components of a string design and the bonding technique have been described in detail in Sec. 3.3. This section describes how these components are assembled creating a pixel module-layer.

3.6.1 Single-chip assembly

Figure 47 shows a custom-made jig called Single ALPIDE Bond Tool (SABT) developed for high-precision alignment of a single ALPIDE with the chip-cable during bonding. The tool has a special vacuum mechanism to keep components in place without movement. After the bonding, the chip with chip-cable is placed in a Yamaichi frame, which together with the Yamaichi socket is a type of chip-testing system. These systems are made by a company called Yamaichi, thereby the name. The Yamaichi frame is the black frame with ALPIDEs and chip-cables in the center in Fig. 48. The Yamaichi frame connects all relevant pins on the ALPIDE and the chip-cable to a FPGA via a bed-of-needles design in the Yamaichi socket. This tool has been successfully used for the production of string prototypes.

3.6.2 Production tests

Furthermore, a testing framework has been developed for functional tests between individual assembly steps. After the single-chip assembly, the Yamaichi frame is placed into Yamaichi socket on the *Production Test Box* (PTB) (see Fig. 48).

Using the PTB allows the interface to all utilized bonds of the ALPIDE for performing functional tests such as register test, analog and digital scan, and others. It also ensures that the bonding has been executed successfully. In addition, it allows to perform classification tests of the ALPIDE. The latest version of the PTB even includes a FFs connector for connecting a multi-chip string.

3.6.3 Multi-chip assembly

The final step of the string assembly is to align and glue the flex cables first and then the chips (already bonded to the chip-cable and tested) next to each other on a carrier. After that, the chip-cables are bonded to the long flex cable to create one multi-chip string. So far the alignment has been done by making alignment lines directly on the carrier and then placing small weights on the chips. Figure 49

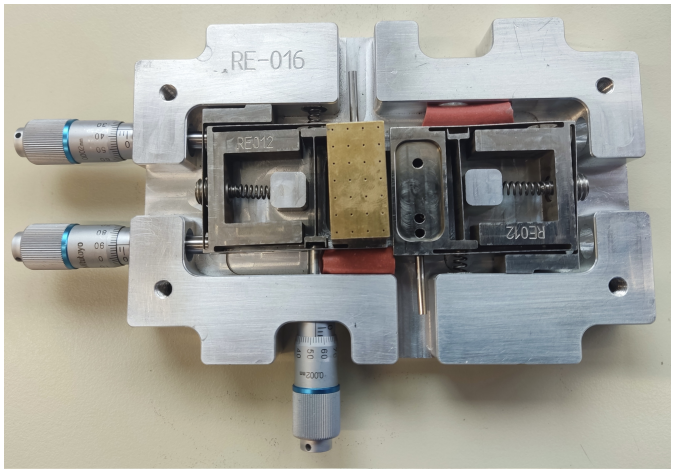


Fig. 47: The SABT developed for alignment during bonding and for mounting the chip into a Yamaichi frame.

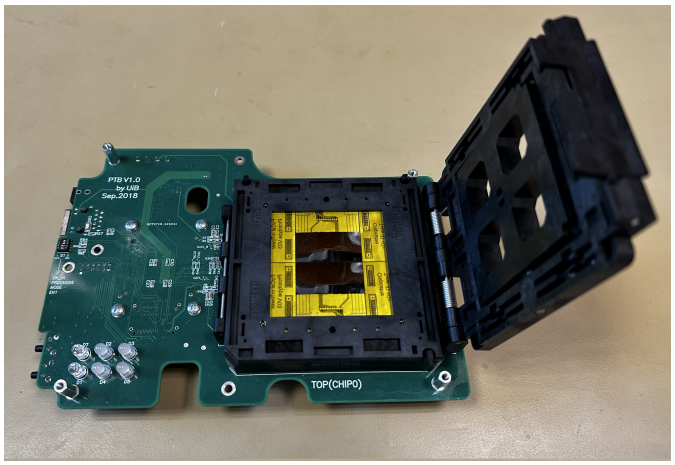


Fig. 48: The PTB developed for functional tests of both chip-cables and strings during production.

presents a similar procedure but with a custom-made jig called MCAT for faster production. It is also proposed to connect a TC to test the whole string at this stage.

3.6.4 Pixel module-layer assembly

For assembling one full pixel module-layer, three 15-chip strings are glued onto an aluminum carrier next to each other in the previous multi-chip assembly. Furthermore, two carriers are assembled together back-to-back with various spacers to create a baseline for mechanical support. The ALPIDE chips form one uniform active area (82.22 mm $\times$ 451.40 mm) with the 2 mm gap of the aluminum carriers. Such configuration will create one pixel layer and this is true for both IB/OB and OB strings. This means that one pixel layer consists of six 15-chip strings corresponding to 90 ALPIDE chips.

Figure 50 shows the structure of the final pixel layer. The layer is 8.5 mm in thickness and it also includes a TC. The TC is a simple PCB that connects data, clock, control, and power between the detector and the off-detector electronics, see Sec. 3.5.2.

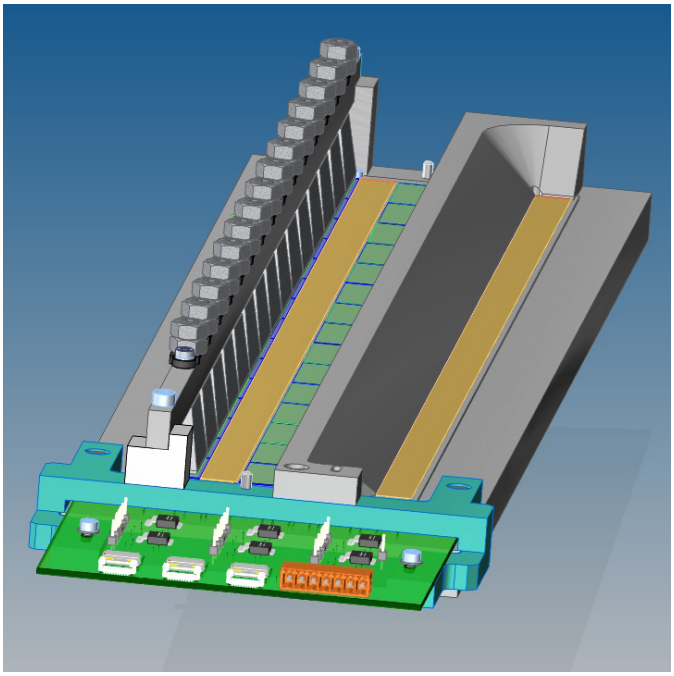


Fig. 49: The MCAT designed for alignment of ALPIDEs already bonded to a chip-cable for further bonding to the flex cable and gluing to the carrier.

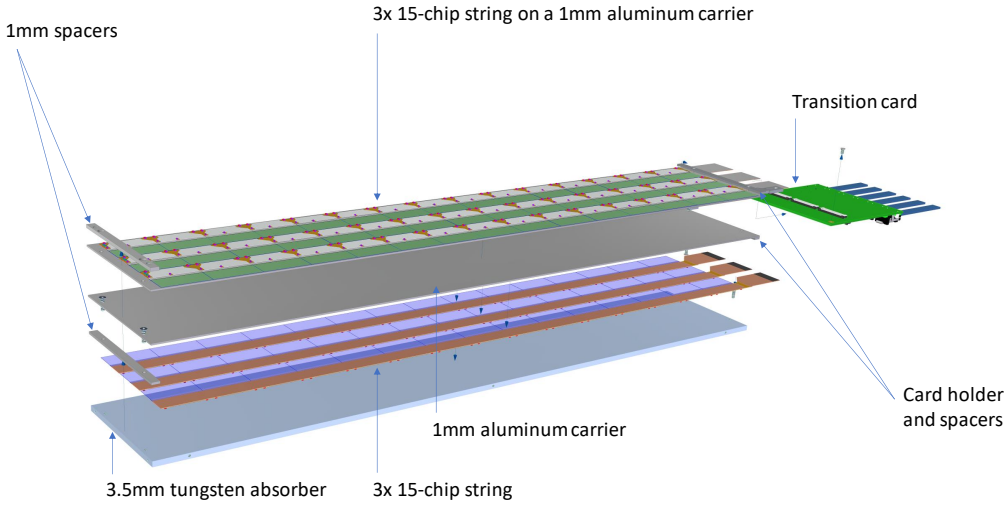


Fig. 50: Schematics of a pixel module-layer; in addition also the TC card is shown.