PROCEEDINGS OF SCIENCE

Front-end electronics for the B-layer replacement or SLHC in ATLAS

Roberto Beccherle for the ATLAS Pixel Collaboration

INFN - Genova

Via Dodecaneso 33, 16136, Genova, Italy

E-mail: Roberto.Beccherle@ge.infn.it

After the successful installation of the present ATLAS Pixel Detector in 2007 and its commissioning in 2008, the ATLAS Pixel Collaboration is now targeting the needed enhancements foreseen for the Phase-I upgrade of the LHC. This paper will describe the ongoing work for the new "Insertable B-Layer" (IBL) from the point of view of the overall architecture, the electronics and chosen design implementations. We are currently finalizing the global architecture and are working on the implementation of a full scale FE chip, to be implemented in IBM 0.13 μm technology and submitted via CERN in 2009. Work is based on a prototype submission done in 2008. The new IBL will be located at 3.7 cm from the beam pipe, posing new requirements both on total dose and Single Event Upset (SEU) robustness. Foreseen hit rate, at which the chip has to operate, imposes a completely new readout architecture. In addition to that the new detector layer has to operate within the present Pixel DAQ, powering scheme and data link implementation, as only few changes on these systems will be implemented for the upgrade.

The paper will describe chosen implementations and design choices focusing on the new digital readout architecture and interface. The last section will also shortly describe possible future enhancements to the communication protocol, in order to cope with clock data alignment issues, higher bandwidth requirements and allow having multiple communication channels active at the same time.

17th International Workshop on Vertex detectors

Utö Island, Sweden

July 28th - 1st August, 200

PoS(VERTEX 2008)045

1. Need for an upgrade

The ATLAS detector 4 will start taking data in summer 2009. The Pixel detector 4 is the innermost part of the inner tracker and is formed by a barrel with three layers and two end-caps with three disks each. The detector is based on 1744 identical modules. Each module is hosting 16 identical Front End (FE) chips, each bump bonded to a single silicon sensor segmented in 46.080, $50 \times 400\ \mu m^2$ pixels and a digital Module Controller Chip (MCC) that provides configuration and readout capabilities. The electronics, and the overall architecture, has been designed to cope with the design luminosity of $10^{34}\ cm^{-2}s^{-1}$. The Large Hadron Collider (LHC) has planned an upgrade of the machine in two distinct phases during the first 10 years of operation. Phase-I upgrade is planned by 2014, and should allow reaching a luminosity of $3x10^{34}\ cm^{-2}s^{-1}$ in the following years. A second and more radical upgrade of the machine is planned between 2016 and 2018. This second upgrade is identified as Phase-II upgrade and often referred as "Super-LHC" (SLHC).

1.1 Phase-I and the Insertable B-Layer (IBL)

The innermost pixel layer, called B-Layer, will have degraded performance before a new inner detector will replace the existing one for SLHC. At that time the integrated luminosity will be $3 \times 10^{34} cm^{-2}s^{-1}$. Since the replacement of the existing B-Layer has shown to be impossible during LHC shutdowns, the ATLAS collaboration has decided to make a new detector to be inserted, at a radius of 3.7 cm, together with a smaller radius beam-pipe, inside the existing pixel detector.

This choice poses many new constraints to the FE electronics and the sensors. Both will have to sustain a higher radiation dose (200 MRad, that in Silicon correspond to a fluency of $4 \times 10^{15}\ n_{eq}/cm^2$) and there will be increased readout data rates.

Being closer to the beam also imposes an increase in the live area of the modules in order to guarantee good spatial coverage of the modules forming the staves of the new barrel.

All these constraints together with a very tight schedule impose R&D and prototyping by end of 2010. In order to be installed by 2014 the IBL has to be ready in 2013 and therefore construction will be in 2011-2013.

1.2 Super LHC

Work is also ongoing for Phase-II, in which the whole inner detector will be replaced, where the Pixel detector could consist of a four layer barrel covering radii ranging from 3.7 cm

to about 25 cm. Of course a completely new electronics and architecture, will have to be developed for SLHC. As simulation studies show that data rates of the outermost layers are comparable to the ones foreseen for the IBL we decided to develop an electronics that could be used, maybe only with minor modifications, for both the IBL and the outer layers of the detector foreseen for SLHC.

This allows defining the development needed for the IBL as an intermediate step towards SLHC, thus allowing the same community to be focused on both projects.

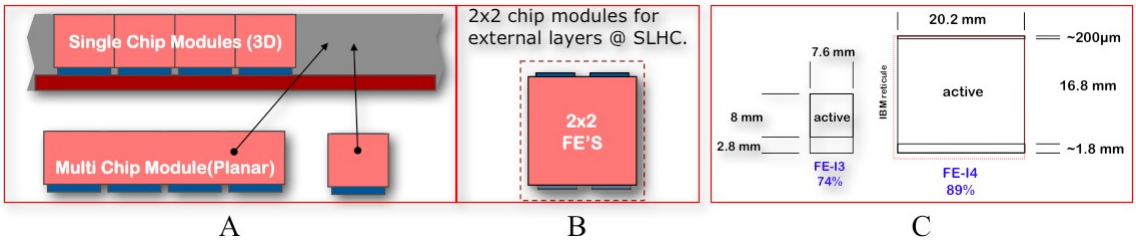

Figure 1: **A**) Possible module choices based on 3D or planar sensors. **B**) External layers at SLHC will have four FE chips bump–bonded to a single sensor. **C**) FE-I4 chip size and active area.

2. New Pixel Front End chip

FE-I3 4, the current Pixel Front-End chip, was designed using IBM 0.25 µm technology using enclosed layout techniques to provide radiation hardness. Each pixel detects the hit, determines the corresponding charge information, using the Time over Threshold (ToT), and its timing inside the pixel cell. Due to space limitations hit information is not stored inside the cells but transferred to the chip periphery, using a common bus shared between two columns of pixels, as soon as possible in order to avoid inefficiencies in the pixels. All transferred data is stored in End of Column (EoC) buffers until a LV1 trigger arrives and eventually validates the stored information. Such column drain architecture has an obvious limiting factor due to the congestion of the bus transferring all information from the Pixels, as soon as one increases the luminosity and therefore the data rate.

Detailed simulations, using a full C^{++} model of the present FE-I3 chip, together with physics simulations of the current detector, have been performed to correctly understand current architecture limitations and in order to help defining the new readout architecture and to optimize its components.

Technology improvements

Technology improvements both in the sensor and in the electronics area allow now designing for a radically different system, which should address most of the issues arising from the increased luminosity and radiation hardness.

New sensor technologies allow building edgeless detectors, thus allowing a stave design based on non-shingled modules. This will end up with staves that are much simpler from the mechanical point of view. A sketch of possible module implementations can be seen in Figure 1-A. Figure 1-B shows the basic concept of a module for the outer layers of SLHC in which we would bump-bond four FE chips to a single sensor.

The new IBM 0.13 μm CMOS technology allows building bigger chips, smaller pixels and higher output speeds. This technology is intrinsically rad-hard up to 100 MRad, without the need of enclosed transistors. This allows the realization of much denser designs even with smaller pixels. The new Front End chip (FE-I4) will have 80 columns and 336 rows of 50 μm x 250 μm large pixels. Total chip size will be 19 mm x 20 mm with an active area of ~89% (see Figure 1-C).

FE-I4 will provide an output data rate of 160 Mb/s per chip, thus allowing connecting the FE chips directly to the optical readout chips. Present pixel detector modules, hosting 16 FE chips and a single Module Controller Chip (MCC) 4, communicate with the ROD using two 80 Mb/s links per module. This allows eliminating the need of an MCC, thus further simplifying the design of the system. The removal of the MCC, however, implies that all communication, configuration, trigger and timing handling issues, together with event building, currently performed by the MCC, will have to be moved inside the new FE chip. This added complexity, together with the digital circuitry to be implemented in the new digital region architecture, imposes a new approach to the design of the digital part of the chip. We therefore chose to use a design flow, for the digital part, that is quite different from what has been done in FE-I3. The circuit to be implemented will be described using a high level behavioural description and will be simulated using a fast digital simulator. All tools used for implementing the design in silicon, simulation, synthesis, timing verification, signal integrity checks and automatic place and route, come from a digital design flow. In addition to this, all critical components, like global signal distribution, clock trees, power distribution will be simulated using full analog simulations with full 3D parasitic extraction. This approach should ensure a reliable implementation of all the desired functionalities.

The following table shows the most important specifications of the new FE chip.

Pixel size	50 x 250	μm^2
DC leakage current tolerance	100	nA
Pixel array size	80 x 336	Col x Row
Pixel input capacitance range	300 – 500	fF
In-time threshold with 20 ns gate (400 pF)	4000	e^-
Hit – Trigger association resolution	25	ns
Double Hit discrimination (single pixel)	400	ns
Single channel ENC sigma (400 fF)	300	e^-
Tuned threshold dispersion (max)	100	e^-
ToT charge resolution	4	bits
Operation voltage range	1.2 ÷ 1.5	V
Analog supply current (@400 fF)	10	µA/Pixel
Digital supply current (@100kHz)	10	µA/Pixel
Average Hit rate	200	MHz/cm^2
Maximum Trigger rate	200	kHz
Maximum Trigger latency	3.2	µs
Single chip data output rate	160	MHz
Radiation tolerance (specs met at this dose)	200	MRad

Table 1: Main specifications of the FE-I4 chip.

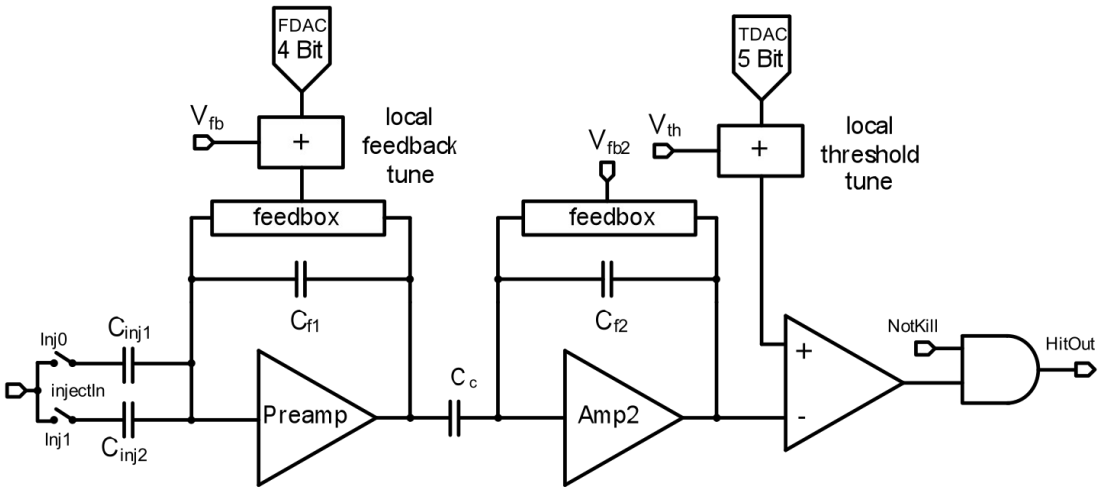

Figure 2: Block diagram of the Analog cell. C_{f1}= 17pF and C_c/C_{f2} = 5.8. Each pixel implements a 4-bit local feedback tuning capability and a 5-bit local threshold tune.

The new Analog Front End

The analog part of the pixel cell of the new pixel Front End chip (FE-I4) will occupy an area of approximately 50 μm x 140 μm out of the total 50 μm x 250 μm. Figure 2 shows a block diagram of the new design. As can be seen the front end circuit is made with a two stages amplifier with a charge preamplifier AC-coupled to a closed loop 2nd stage. The preamplifier uses a triple-well NMOS transistor as the input device and is a high gain regulated cascode preamplifier with an NMOS feedback transistor for constant current discharge of the feedback capacitor.

The leakage current compensation is based on a differential amplifier. Each pixel has a 4–bit local feedback tuning capability built in. The second stage amplifier is a folded regulated cascode that uses a PMOS input. The preamplifier feedback is done with two FETs, CF is 17 fF, and there is an active, differential, leakage compensation circuit. The 2^{nd} amplifier is DC-coupled to a classical two-stage comparator with a global threshold setting and a local per pixel threshold tuning capability based on 5–bit DAC. More details on the design and measurements of the analog FE circuit can be found in 4.

The new digital architecture

The new Pixel FE-I4 will be based on a new architecture, based on the consideration that present limitations arise from the double column data drain model. Both in FE-I3 and FE-I4 the pixels are placed abutting the digital part of the Pixel in order to form a column pair, as can be seen in Figure 3-B. This allows sharing some digital circuitry, a common bus for data transfer and communication from/to the End of Column (EoC). Figure 3-A shows the actual FE-I3 readout architecture, where all hits detected and time stamped in the double column are transferred to the EoC region as soon as possible and subsequently stored in EoC buffers. Here incoming triggers select relevant data for readout, while absence of triggers clears remaining buffers. In this way more than 99% of the data that gets transferred to the chip periphery will be erased inside the EoC buffers, as it does not match a LV1 trigger.

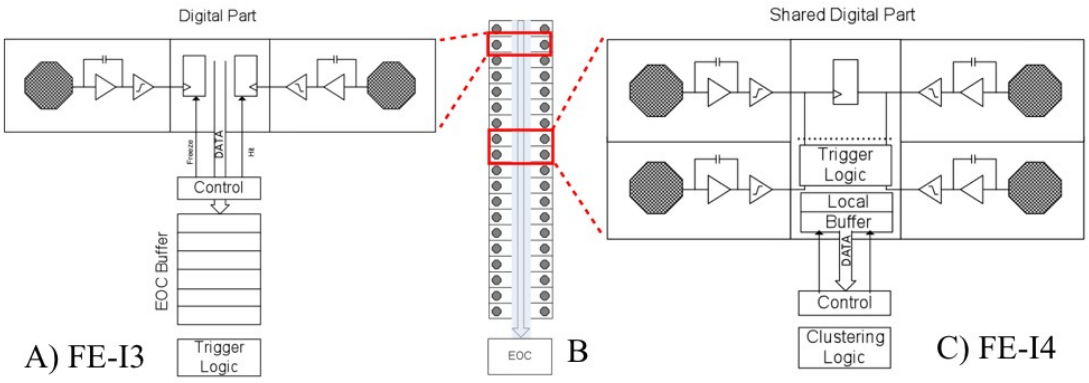

Figure 3*: **A**) Limited digital storage implies that data is transferred to the EoC buffers as soon as possible. **B**) Pixels are grouped in Double Columns. **C**) All information is kept (as long as possible) inside the Pixel region and transferred only if validated by a LV1 trigger.*

Inefficiencies due to buffer overflows and congestion in the readout bus inside each double column are the main limitations of this architecture and become significant already at data rates foreseen for the IBL. The basic idea to overcome this limitation is to store all information of a hit pixel, in digital format, locally inside the pixel and only read out the information if an incoming Trigger validates the stored data. This will end up in a big reduction of the bus activity in the double column region. Such a scheme requires significant storage capabilities inside the Pixels and this is possible due to the much more compact design rules of 130 µm technology. Storing all information directly in the pixels is still not feasible as one buffer per pixel would surely not be enough, and two would probably not fit in the smaller pixel size. In order to organize local storage more efficiently we decided to group the digital part of the pixel cell in so called regions. The digital logic of two pixels belonging to a double column is therefore grouped together. If we group N adjacent pixels of a double column we form a so called "2 x N region" that shares all the digital logic of the 2 x N pixels belonging to the same region. Figure 3-C shows a 2 x 2 digital region. This allows optimizing the region for an optimum number of local storage elements and minimizing the digital logic needed. There has to be a compromise between the number of pixels grouped together which leads to a lower number of buffers, but a higher complexity in the routing and combinational logic needed to deal with the common region.

Another source of inefficiency of the present electronics was coming from the time-walk. A hit that crosses the threshold of the digital comparator is identified by the time stamp of the Leading Edge (LE) and the ToT, calculated as the number of bunch crossings the signal stays above threshold. Signals just above the threshold will be time-stamped later than big signals and the difference in time is the time-walk of the comparator. Since the time stamp is discrete it might happen that a signal gets time stamped in the wrong bunch crossing.

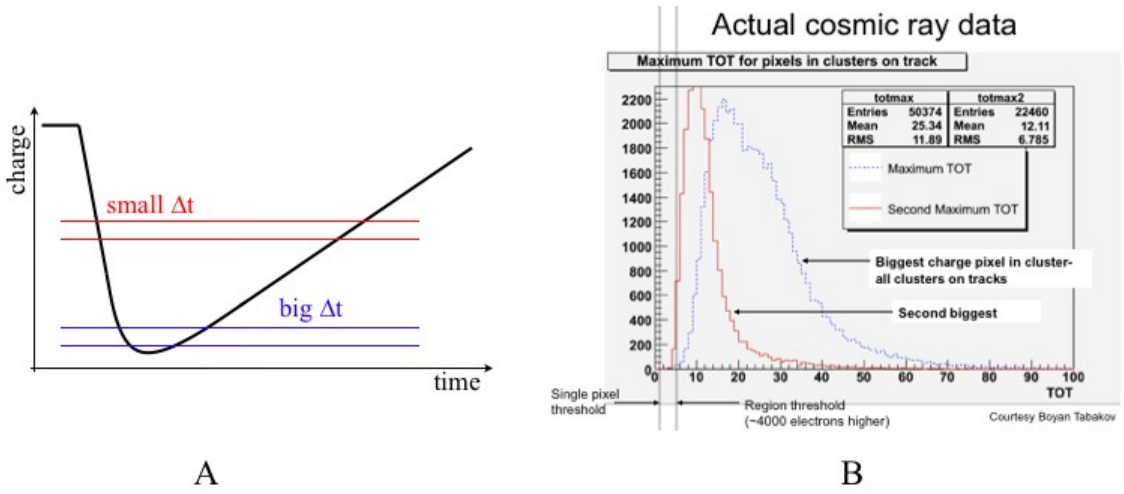

Figure 4: **A**) Small fluctuation in charge produces a big time walk for pulses just over threshold and a small one for signals well over threshold. **B**) Single hits produce signals with a big ToT, while second biggest pulses generate values with small ToT.

This can be seen in Figure 4-A. Since the time stamping discrete if the time-walk exceeds 25ns the LE gets associated to the wrong bunch crossing and the Hit is lost.

The introduction of pixel regions can also be used to drastically reduce the time-walk associated to a single hit, taking into account the fact that real physics hits are usually clustered and that in a single cluster there is always a hit that is big compared to the threshold, as can be seen in Figure 4-B. Using this knowledge we can store the LE information of a whole cluster only once, using the biggest pulse detected inside a single region. The ToT information of the remaining hits of the same region is still stored in order to provide charge information, but is not timing critical. This simple trick allows effectively reducing time-walk and using ToT of small pulses just for tracking information, in local clusters of pixels. Of course it might happen that the big pulse is in one region and the small one in the top or bottom one. Therefore each region also implements a 4-bit deep memory, one for each neighbour pixel, as these pixels, not

belonging to the region itself, might have a hit associated with a big pulse belonging to the cluster. In such a case a small ToT value in a neighbouring region would not be discarded.

Simulations

The new architecture has been fully evaluated using a simulation tool written in C^{++} developed for this purpose. Hit and data rates information was coming from a full physics based simulation of the current layout of the new IBL, that corresponds to three times LHC, while SLHC conditions are approximately ten times LHC in terms of data rates. Simulations are extended up to expected SLHC hit rates in order to evaluate design margins, even if FE-I4 is intended to work for IBL and the outer layers of SLHC. For the innermost layers of SLHC a different solution will have to be implemented. Different options, based on different region sizes, number of buffers per region and different ToT processing algorithms were studied in detail. The final choice will be based on total efficiency and on constraints coming from the total number of buffers that one can actually fit inside a particular region.

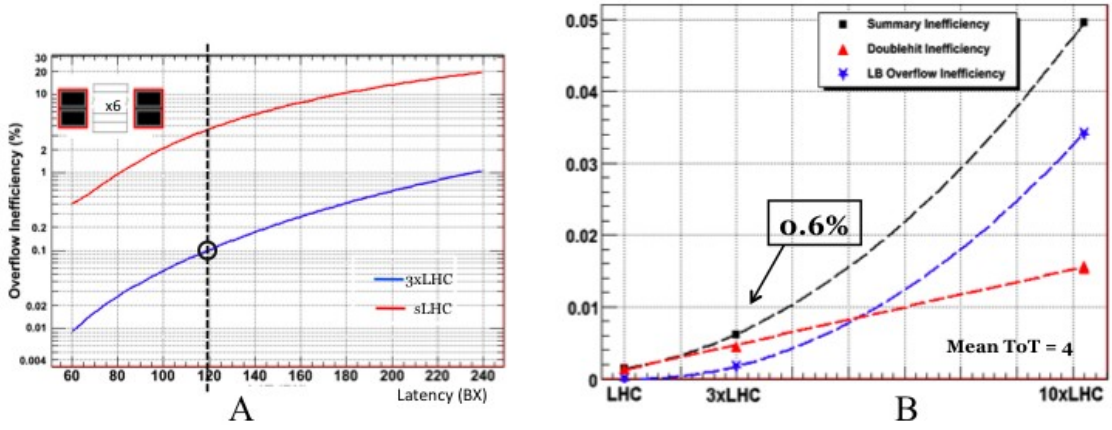

Figure 5: **A**) Buffer overflow inefficiency for a 2 x 2 region with six buffers. **B**) Total (Double Hit and Local buffer overflow) inefficiency in function of luminosity.

Figure 5-A shows the results of buffer overflow inefficiency both for the IBL and for SLHC. A parametric Verilog description of the complete digital region has been developed and automatic place and route tools were used in order to be able to quickly evaluate the maximum number of storage cells that could possibly to fit inside a particular implementation of the digital region. For the time being a target region with size of 2 x 2 pixels has been chosen as the reference design for the FE-I4 chip as it provides good efficiency (see Figure 5-B) while allowing enough room in the region for some small last minute design modifications. Work on

this topic is ongoing and there might very likely be changes in the final implementation of the FE-I4. More details on the simulation work can be found in 4.

Input/Output Data Format

The new chip, used for the IBL, will have to communicate with existing hardware and has to provide hit data to the actual ROD, being part of the actual Pixel DAQ. This poses some constraints to the possible data format extensions allowed for the new chip. These constraints are different in case of downlink or uplink communication. As a consequence the communication protocol between the ROD and the FE chip is essentially the same 40 Mb/s serial protocol used for the present detector. Implemented changes to the present protocol reflect the fact that now we talk directly to the FE chip and not via the Module Controller Chip (MCC) as in the present system. The FE chip for the IBL, being both closer to the beam and more than four times bigger than the present FE chip, will have to handle an increased data rate. The incoming 40 MHz system clock will be internally multiplied by a factor of two and outgoing data will be sampled on both edges of the clock in order to allow a data rate of 160 Mb/s. The ATLAS Pixel system has a data push architecture which means that the system is not able to control data flow (i.e. there is no Busy signal) and the system has always to respond with some data to a request coming from a Trigger command. This implies that even if there are no Hits an empty event has to be produced and that in case of Trigger overflow the system has to be able to signal the number of skipped triggers.

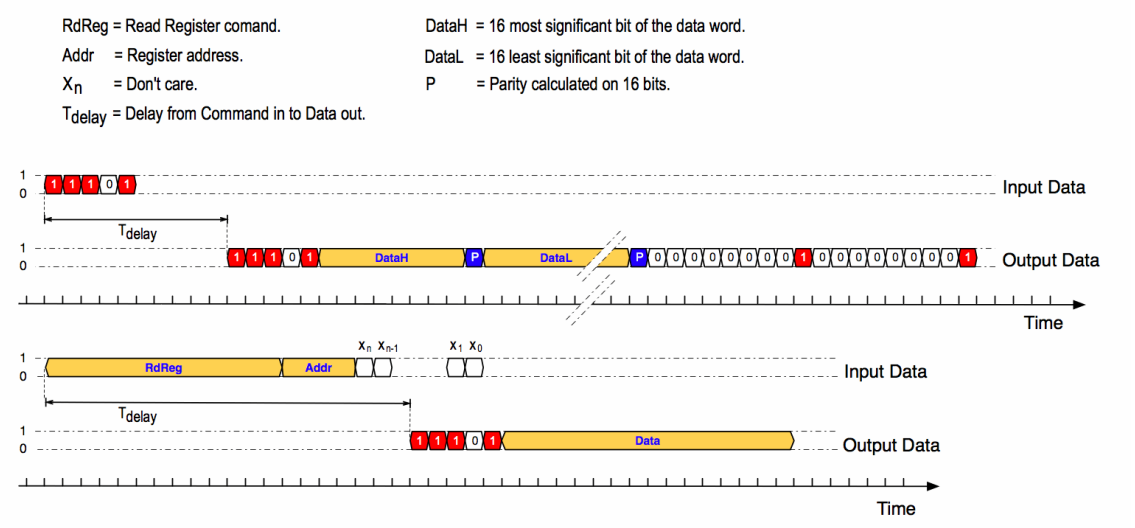

Figure 6: Response to a Trigger and to a RdRegister command. Note that input data is at 40 Mb/s while output data is provided at 160 Mb/s, encoding data on both edges of the clock.

Communication from the FE chip to the Read Out Driver (ROD) allows more flexibility in defining a new protocol as long as we keep the 160 Mb/s data rate. In addition to that there might be significant differences in the data protocol between the IBL version of the FE-I4 and the version to be used for SLHC as a change in specification only involves the redesign of the digital bottom of chip circuitry and not a global redesign of the analog and digital pixel region. The following description applies to the proposed protocol for the IBL, if not otherwise stated.

The FE provides three different types of Output Data: "Configuration Data" in response to a Configuration command, "Detector Data" in response to a Trigger command and "No Data" in all other cases. When we send "Configuration Data" we know in advance the length of the data field, and therefore we need only to provide a Header before the actual data. During event readout ("Detector Data") the length of the data is unknown and therefore one has to foresee a Header, Sync and Trailer format, in order to allow the ROD to know when a data stream, associated to a particular incoming Trigger command, is completed. In case we do not have to transmit any data we still need to transmit a "No Data" command in order to avoid long streams of zeroes or ones. The proposed protocol, as can be seen from Figure 6, uses a 5-bit Header, followed by N Data words and is ended by a Trailer word in case of event readout.

The Data words have a fixed width of 32 bits and in order to keep the Trailer as short as possible (it has to be a pattern that never appears in the actual data stream) two synchronization (Sync) bits are added for each data word. Sync bits are calculated using the even–parity bits of the first and last 16 bits of the data word. The Trailer is constructed using the odd–parity bits in order to make it a unique pattern.

Exact format of the event generated by the FE has still to be defined, but it will provide events by grouping all hits belonging to the same LV1 number. The first event that has been triggered is the first to be sent out. Every event is entirely transmitted before the next event is considered for transmission: no event interleaving is allowed. LV1Id and BCID information is sent out together with each event. Data are sorted and grouped by FE double columns. A token mechanism is used inside the FE chip to collect data coming from all double columns. For the time being each Pixel region (formed by four distinct pixels) has to transmit 20 bit of information (four 4-bit ToT charge information and 4-bit information to know if the neighbouring pixels have some information related to the same event) and is uniquely identified by an 8-bit row and 6-bit column address. Event length is not known until the whole event is transmitted. This makes the event builder simpler, but forbids the usage of forward word counters in the event frame. Event data will be eventually padded with zeroes in order to fit in the correct number of data words.

In addition to this the FE-I4 chip will provide the ability to encode data using an 8b/10b algorithm that is DC balanced, allows some error detection capability and allows clock recovery from the data but has a transmission overhead of 25%. DC balancing will be a user selectable option.

3. Possible protocol extensions for SLHC

Two possible extensions to the present protocol for SLHC are presented in order to overcome possible limitations of the present architecture.

Trigger and Configuration Data

One of the biggest limitations in our present system architecture is the fact that the chip has two distinct, and mutually exclusive, operating modes, RunMode and ConfigMode. Due to the fact that the chip is accessed via a single 40 Mb/s serial protocol and that the Trigger command identifies uniquely the bunch crossing to which it refers (it has a fixed latency) commands are split in Trigger and Fast commands (processed during RunMode) and Configuration commands that can be used only during chip configuration. Usually the chip starts in Configuration mode, it gets configured and tuned and then it is put in RunMode via a dedicated command. Once in this mode the Command Decoder processes only Trigger and Fast (Reset and Sync) commands. If a Configuration command is detected by the system the chip enters ConfigMode and before being able to process Trigger commands again, it has to be rearmed.
This mechanism does not allow issuing monitoring commands during RunMode. Monitoring commands would be very useful to check internal test structures that inform the counting room of anomalies, warnings and/or SEU detected by the chip.

Another useful feature that could be used would be integrating some DCS functionality in the FE electronics, via an IP block, or in the module hosting it using an external component. In such a case one would like to be able to communicate with these structures and monitor them. This should of course be done during data taking and should allow embedding monitoring data inside the output data stream.

One possible solution to overcome this limitation would be to send data to the detector using a higher than 40 MHz clock frequency and splitting the downlink in two logically distinct channels dedicated to different type of communications. One channel would be reserved for Trigger information while the other one could be used for configuration and monitoring with the chip even while the FE is in RunMode.

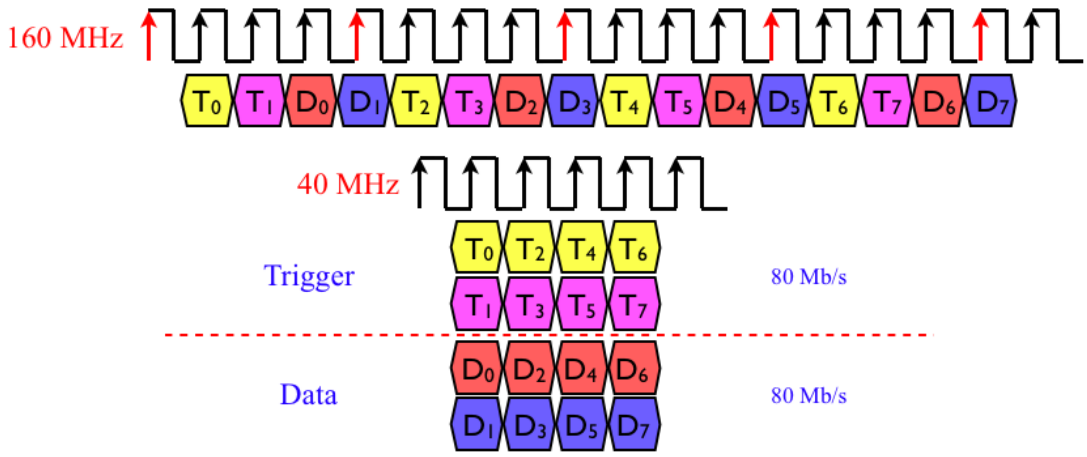

Figure 7: One serial data stream encoded with a 160 MHz clock can be split in two distinct channels of 80Mb/s allowing the concurrent transmission of two different data streams.

One possible implementation of this approach with a 160 MHz clock used on the downlink is shown in Figure 7. In the depicted case the communication channel has been split in two logically distinct 80 Mb/s channels carrying different type of information. Another possible implementation of this idea could be to split the link in more than two channels in order to allow one single physical link to communicate with multiple chips. This is possible as the Trigger channel (carrying only Trigger, Sync and Reset commands) would be common to all chips, while the configuration channels could be different.

Data Strobe encoding

As soon as we start sending to the detector a clock frequency that is a multiple of 40 MHz, we will have to face two different problems. Data must be correctly associated with the rising edge of the original 40 MHz clock (machine clock) and one has to correctly sample the data words with the incoming clock eventually re-phasing the data line. Extracting the 40 MHz clock and aligning data with the right edge of the clock can be performed using some special data alignment commands that can be added to the protocol and that have to be sent once in a while to the chip.

Making sure that data words are correctly aligned with the clock phase, can become a difficult task to perform especially if the cable lengths of clock and data are not the same and in case of noisy environment. Usually this is done with the use of a PLL/DLL circuitry inside a FE chip.

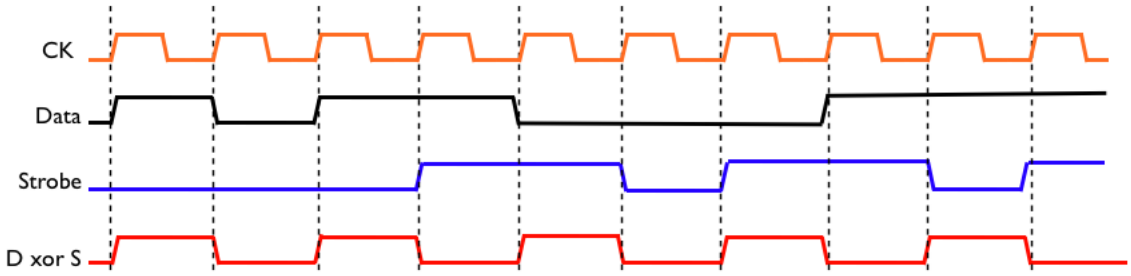

Figure 8: Data/Strobe encoding. Strobe line is obtained from Data and Clock ensuring that every clock cycle only one of the two changes its value, but never both. Data and Strobe signals are sent on the link instead of Data and Clock. Clock is reconstructed from these two lines using an XOR.

One could use the Data/Strobe encoding (as shown in Figure 8) to solve the same problem, by encoding Data and Clock so that either Data or Strobe changes its logical value in one clock cycle, but never both. The link will then carry both Data and Strobe, instead of Data and Clock. This approach, also used in space communications (SpaceWire) and consumer applications (FireWire 400), allows for easy clock recovery with a good jitter tolerance and without the use of PLL/DLL circuitry. Clock information is reconstructed, XORing the two signal line values.

4. Conclusions

Work for the Front-end electronics for the Insertable B-layer is ongoing. The limited amount of time available for R&D and the constraints imposed by the existing hardware and software drive some of the requirements of the new electronics. This, together with constraints coming from the increased radiation levels and particle fluencies, determine the need to develop a new architecture of the front-end readout chip that overcomes the limitations that the present FE-I3 chip would have. The paper describes the main blocks of the new FE-I4 chip with focus on the new chip architecture. There is also a brief description on possible extensions of the data protocol to cope with needs that might come from SLHC requirements.

References

[1] The ATLAS Collaboration and G. Aad et al, 2008 *The ATLAS Experiment at the CERN Large Hadron Collider*, Journal of Instrumentation, number **08**, vol. 3, S08003.

[2] G. Aad et al, 2008 *ATLAS pixel detector electronics and sensors*, Journal of Instrumentation, number **07**, vol. 3, P07007.

[3] I. Peric et al, 2006 *The FEI3 readout chip for the ATLAS pixel detector*, Nucl. Instrum. Meth. A **565** 178.

[4] R Beccherle et al 2002, *MCC: the Module Controller Chip for the ATLAS Pixel Detector*, Nucl. Instrum. Meth. A **492** 117.

[5] D. Gnani et al, *FEI4_P1 a prototype pixel readout chip for the ATLAS B-layer upgrade*, in proceedings of *Pixel 2008* [http://indico.fnal.gov/conferenceDisplay.py?confId=1588].

[6] D. Arutinov et al, *Digital Architecture and Interface of the New ATLAS Pixel Front-End IC for Upgraded LHC Luminosity*, Nuclear Science Symposium Conference Record, 2008. NSS '08. IEEE, 1923 [http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4774765&isnumber=4774073].