

FE-I4 ATLAS Pixel Chip Design

Marlon Barbero¹, David Arutinov, Tomasz Hemperek, Michael Karagounis, Andre Kruth, Norbert Wermes

University of Bonn

Nussallee 12, D-53115 Bonn, Germany

E-mail: barbero@physik.uni-bonn.de

Roberto Beccherle, Giovanni Darbo

INFN Genova

Via Dodecaseno 33, IT-16146 Genova, Italy

Sourabh Dube, David Elledge, Maurice Garcia-Sciveres, Dario Gnani, Abderrezak Mekkaoui

Lawrence Berkeley National Laboratory

1 cyclotron road, Berkeley, CA 94720, United States of America

Denis Fougeron, Mohsine Menouni

CPPM Aix-Marseille Université

CNRS IN2P3, Marseille, France

Vladimir Gromov, Ruud Kluit, Jan David Schipper

NIKHEF

Science Park 105, 1098 XG Amsterdam, The Netherlands

FE-I4 is the new ATLAS pixel chip developed for use in upgraded luminosity environments, in the framework of the Insertable B-Layer (IBL) project but also for the outer pixel layers of Super-LHC. It is designed in a 130 nm CMOS process and is based on an array of 80 by 336 pixels, each $50 \times 250 \mu\text{m}^2$ for an overall size of about $19 \times 20 \text{mm}^2$. Each pixel consists of analog and synthesized digital sections. The analog pixel section is designed for low power consumption and compatibility to several sensor candidates. The digital architecture is based on a 4 pixel unit called region, which allows for a power-efficient, low recording inefficiency design, and provides a solution to record hits timewalk-free. A mixture of techniques is used for yield enhancement. The chip periphery contains a control block, a command decoder and global memory, powering blocks, a data reformatting unit, an asynchronous storage FIFO, an 8b10b coder and a clock multiplier unit, which allows data transmission up to 160 Mb/s for the IBL.

VERTEX 2009 (18th workshop) – VERTEX 2009

Veluwe, the Netherlands

September 13-18, 2009

¹ Speaker

1. The Insertable B-Layer Project and Outer Layers of Super-LHC

CERN's Large Hadron Collider (LHC) has restarted operation at the end 2009. According to the available schedules and to the progresses that will be achieved, the LHC will ramp up both in center-of-mass energy and in luminosity to reach the 14 TeV proton-proton benchmark, and the full design luminosity of 10^{34} particle.cm⁻².s⁻¹. ATLAS [1] is one of the two multipurpose experiments located on the LHC ring. The pixel detector is the innermost element of ATLAS [2]. It provides excellent single point resolution and is an essential ingredient for both precise tracking and the determination of displaced vertices resulting from the production of long-lived particles such as b-quarks. As such, it is a crucial detector for many analyses involving b-tagging. The pixel detector surrounds the beam-pipe in a three cylinder arrangement, with three extra disks at each end to enhance its hermiticity.

The innermost pixel layer is located at a radius of about 5 cm and suffers from high radiation exposure. It is designed in radiation-hard technologies with the requirement to sustain 50 Mrad total irradiation dose. An accurate prediction of the LHC ramp-up is difficult to make, but it is believed that despite its radiation tolerance, the innermost layer may show degradation due to sensor damage in the timescale of a few years (around 2014). An inserted new pixel layer inside the current pixel detector is the most favored option to recover good physics performance even with a radiation-damaged highly inefficient pixel layer at 5 cm. From an engineering point of view, this is the most favorable option too. Hence the project of inserting a smaller radius b-layer at about 3.7 cm together with a smaller radius beam-pipe has started, and is now called the Insertable B-Layer project (IBL).

After the year 2018, the LHC will undergo a major upgrade in order to reach higher luminosity. This second phase, called Super-LHC, will need upgrades to major parts of ATLAS. In particular the inner tracking detector needs to be re-designed to cope with a 10 fold increase in number of hits and radiation. The new tracker will be all-Silicon, with inner pixel layers, followed by short Silicon Strip detectors (2.4 cm long) and then long Silicon Strip detectors (9.7 cm long) at larger radius. It will cover radii from 3.7 cm to close to 1 m. A definitive layout is yet (end of 2009) to be agreed on. Currently, discussions concerning the number of layers of the various technologies, the exact position of layer radii and end-caps, taking into account track reconstruction and b-tagging capabilities, cost, powering and material estimates and time schedule are ongoing. It is believed that the pixel detector will consist of two parts, a central insertable double-layer at radii below 10 cm, and an external double- (or triple-) layer system consisting of fixed layers covering radii from ~15 cm to ~25 cm. The area of the external pixel system being much larger than the area of the inner pixel system, cost constraints and time schedule are more demanding for the outer layers.

FE-I4 is developed to address the needs of an IBL at 3.7 cm around the year 2014 and also fits the needs of the pixel outer layers at Super-LHC. Section 2 gives a general introduction to FE-I4 specifications. Section 3 focuses on the analogue pixel, and section 4 focuses on the digital pixel. Section 5 discusses many peripheral blocks and system issues. In section 6 we extrapolate to future developments.

2. Introduction to FE-I4

The motivations for the redesign of the current pixel Front-End FE-I3 come from several aspects, related to system issues and physics performances of the pixel detector. With a smaller innermost layer radius for the IBL project and an increased luminosity, the hit rate increases to levels which the current Front-End architecture is not capable of handling. In particular, it was shown [3] that the current FE-I3 column-drain architecture scales badly with high hit rates and increased FE area, leading to unacceptable inefficiencies for the IBL (see Figure 1).

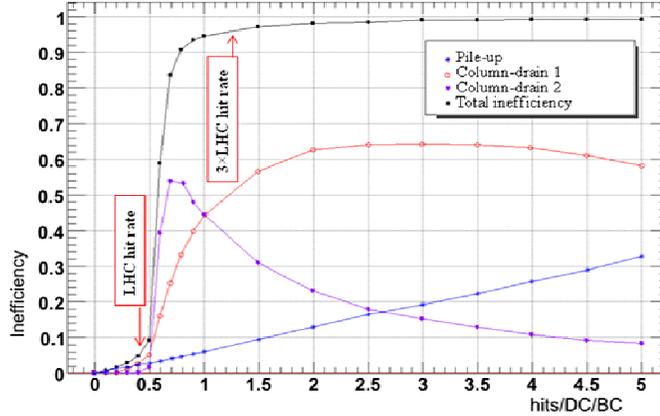


Figure 1: Inefficiencies for a FE-I4 using a FE-I3-like column-drain architecture in a 3.7 cm radius layer, given as a function of the number of hits per Double-Column and Bunch-Crossing. Pile-up inefficiency comes from lost hits due to an already busy analog pixel chain, whereas column-drain 1 (resp. 2) inefficiency comes from lost hits due to busy digital pixel (resp. lost hits due to wrong timestamp association in periphery). See [3] for more details.

FE-I4 stores hits locally to avoid a column-drain based transfer. The FE-I4 pixel size is also reduced, from $50 \times 400 \mu\text{m}^2$ to $50 \times 250 \mu\text{m}^2$ which reduces the pixel cross-section and enhances the single point resolution in z direction. FE-I4 is built up from an array of 80 by 336 pixels, each pixel being subdivided into analog and digital section. The total FE-I4 active size is 20 mm (z direction) by 16.8 mm (ϕ direction), with about 2 mm more foreseen for periphery, leading to an active area of close to 90% of the total. The FE is now a standalone unit avoiding the extra steering of a Module Controller Chip for communication and data output. Communication and output blocks are included in the periphery of the FE (see Section 5).

Going to a bigger FE size is beneficial with respect to active over total area ratio as well as for the building up of modules and staves. This leads to more integrated stave and barrel concepts, and as a consequence reduces the amount of material needed per detector layer. Together with R&D on the thinning down of the chip and the exploitation of thinned sensors, lighter mechanics, new cooling system and the possibility to reduce the cabling scheme (powering), this leads to stave concepts with much reduced material, from about 2.5% x/X_0 for the current pixel layers to about 1.5% x/X_0 for the future staves. Such a reduction of material has a drastic effect on physics performance, e.g. on b-tagging efficiency vs. light quark rejection factor. One of the main advantages of having a big FE is also the cost reduction. Despite progresses from the industry to decrease the cost of bump-bonding, the main driver of the detector costs still is the flip-chip process. This cost scales proportionally to the number of chips

to manipulate: the bigger the FE, the smaller the cost of flip-chipping per unit of detector area. Reducing this cost becomes a determinant factor for the large area of detector foreseen for the outer layers of Super-LHC.

FE-I4 is designed in a 130 nm CMOS process, in an 8 metal option with 2 thick aluminum top layers for enhanced power routing. Particular care has been taken to separate analog and digital power nets. With the thinning down of the gate oxide, the 130 nm CMOS process shows an increased radiation tolerance with respect to previous larger feature size processes [4]. Using rather minimal guidelines (avoiding minimal size transistors and systematically using guard rings for analog and sensitive digital circuitry), a radiation-hardness of more than 200 Mrad is achievable. In particular the use of enclosed layout transistors is generally not required.

3. The Analog Pixel Section

The analog pixel section fits $\sim 50 \times 150 \mu\text{m}^2$, $3/5^{\text{th}}$ of the total pixel size. It is implemented as a 2-stage architecture, optimized for low power, low noise and fast rise time, followed by a discriminator. The first stage is a regulated cascode pre-amplifier, with a triple-well NMOS input. It contains an active slow differential pair, tying the pre-amplifier input to its output, and used to compensate sensor radiation-related leakage current. The DC leakage current tolerance is above 100nA. The second stage is AC coupled to the pre-amplifier, and is implemented as a PMOS input folded cascode. AC-coupling the second stage to the first brings mainly two benefits: This decouples the second stage from leakage current related DC potential shift, and gives an additional gain factor of about 6 (ratio of coupling capacitance to feedback capacitance of the second stage). As a consequence, the feedback capacitance of the first stage can be increased with positive consequences on charge collection efficiency, signal rise time and power consumption, without degrading the signal pulse amplitude at the discriminator input, as was underlined in [5]. The analog pixel is configured with about 20 global settings (e.g. bias currents, feedback currents, discriminator threshold...) and 13 configuration bits for local adjustment (threshold local tuning, pre-amplifier feedback local tuning, charge injection circuitry...).

The analog pixel has already been prototyped in 2008 in a 61 by 14 analog pixel array called FE-I4proto1. This prototype has shown very good tolerance to radiation, with noise increasing by less than 20 % when receiving a dose of 200 Mrad. Pre-irradiation noise was measured to be of order 65 electrons for unloaded channels (resp. of order 100 electrons for channels loaded with a 400fF diode at input to mimic detector capacitance), matching the simulation results well. It must be underlined that these results rely on an internal calibration method (no sensor attached) and thus can be subject to rather large uncertainties. Bonding a sensor to a FE prototype will give in the future more accurate noise estimates. The prototype has shown excellent pre-irradiation un-tuned threshold dispersion of order 160 electrons, increasing only to around 190 electrons after a 200 Mrad dose.

4. The Digital Pixel Region and the Double-Column

To avoid sources of inefficiency related to a column-drain-based architecture “à la FE-I3”, FE-I4 is based on a local storage of pixel hits in buffers located at pixel level, taking advantage

of the small feature size of the CMOS 130 nm process. After initial architectural studies (explained in detail in [3]), a choice was done to use a 4-pixel structure as the base unit for hit recording and storage inside the Double-Column. Besides avoiding the inefficient transfer of un-triggered hits through the Double-Column, this architecture brings further benefits which can be summarized as follows: The choice of a 2 by 2 pixel region leads to an efficient hit recording with hit losses below 0.6% at hit rates corresponding to 3 times LHC full luminosity. As 4 pixels are tied together from the point of view of their digital logic, digital processing can be shared by the 4 pixels together, which leads to area reduction and power savings. In simulation, for typical IBL hit occupancy, the power consumed by a digital pixel is of order $7 \mu\text{W}$ at 1.2V. Finally, as pixels recording a small number of electrons are most of the time located in the vicinity of pixels recording rather large signals (clustered nature of real physics hit in our experiment), small hits can be recovered without being time-stamped, which gives a handle on time-walk. The 4-pixel region is sketched in Figure 2.

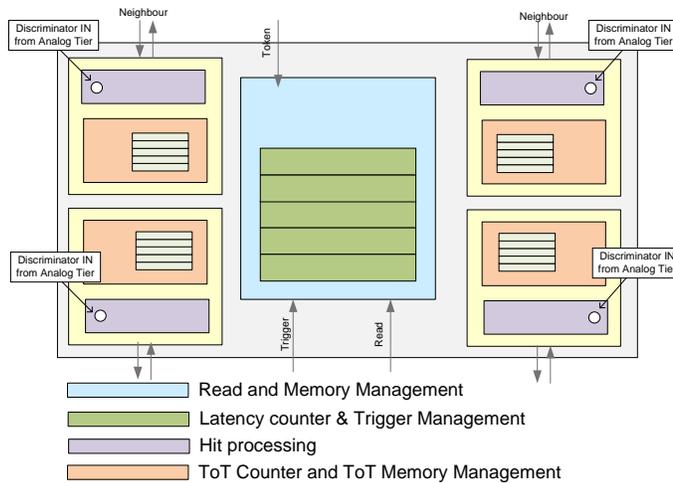


Figure 2: The 4-pixel regional digital logic.

The four pixels of the region form a 2 by 2 logic block inside a Double-Column which is fed by the 40MHz clock (LHC bunch-crossing). Latency counters and trigger management units, as well as read and memory management units are shared between four adjacent pixels. The 8-bit latency counters count down a programmable latency. The pixels still retain individual Time over Threshold (ToT) 4-bit counters, as well as individual hit processing circuitry. Any discriminator that fires in the corresponding four analogue pixels starts the common latency counter, effectively time-stamping a particular event. It is to be noted that even if several pixels are hit in the same bunch-crossing, a single latency counter is allocated. This has the important consequences of reducing digital activity, reducing digital power and improving the efficiency of the architecture. Furthermore, it is possible to distinguish in the digital logic small hits from big hits, by the time the corresponding pixel comparators stay above threshold. The logic allows smaller hits to be associated with bigger hits in their immediate vicinity, either in the same region, or in adjacent regions (so-called “neighbour logic” mechanism). This provides a way to avoid recording small hits with time-walk. There is one hit processing unit per pixel, where the hit leading-edge (the discriminator output rising time) is flagged and the ToT is recorded, and

where the distinction between small and big hits is made (3 programmable modes available for big / small hit digital threshold). An extra bit is stored in the ToT buffers corresponding to the hit information from the off-region neighbour pixel along the Double-Column, effectively crossing the region boundary for small hit recording.

The digital pixel region is entirely designed using automated synthesizing tools, and the automatically generated gates have their bulk node tied to the substrate. To avoid extensive transient coupling and parasitic charge injection to the analog pixel section, the entire digital 4-pixel region is placed in a deep NWELL implant, effectively isolating the digital section from the analog one.

The complete Double-Column consists of 168 regions, with an approximate $200 \times 16800 \mu\text{m}^2$ digital core in the middle surrounded on both sides by the analogue pixel sections ($\sim 150 \times 16800 \mu\text{m}^2$). The bump-bond pads, inputs to the pre-amplifiers from the sensor, are $\sim 12 \mu\text{m}$ hexagonal openings located in the analogue section fitting $50 \mu\text{m}$ pitch constraints both inside the column and with the adjacent neighbour analogue column. Level 1 Trigger information is sent from the periphery to the 4-pixel regions and triggered hits are then read out. The readout is based on a dual token passing scheme (Double-Column / End of Column tokens), made triple redundant with majority voting for yield enhancement. Inside the Double-Column, the data and the thermal-encoded region addresses are propagated down Hamming coded until reaching the End of Column logic and the input to the data storage FIFO (see section 5.1). A pixel configuration shift register runs in each Double-Column for tuning of each analogue pixel locally. For yield enhancement, it is made redundant as well. The End of Column logic is kept very simple and serves only as a dedicated interface between each of the 40 Double-Columns and the digital control block with its FIFO.

5. FE-I4 Periphery

5.1 Data Formatting, Digital Control Block and Storage FIFO:

When region data and address are tagged for readout, they are Hamming coded and sent down Double-Column data busses, then forwarded to the input of the FIFO. There the data are decoded and corrected if need be, before being formatted. After reformatting, data are again Hamming coded for enhanced Single Event Upset (SEU) protection. A sketch of the data transfer until storage to the FIFO is provided in Figure 3.

Reformatting of the pixel region data is done for two purposes. First, this allows reducing the data output bandwidth by removing redundant information related to neighbor logic bits, as well as avoiding to transfer data for pixels having recorded no ToT value. Second, reformatting allows fitting a byte-based format adapted to further processing steps. Raw region data is made of 20 bits, corresponding to four 4-bit ToT values and 4 neighbor logic bits. Based on a GEANT description of the new pixel layer at 3.7cm radius and at a luminosity corresponding to 3 times LHC full design luminosity, event topology and hit cluster shapes were studied as a function of pseudo-rapidity and the choice of a reformatting algorithm. Results of these studies are shown in Table 1.

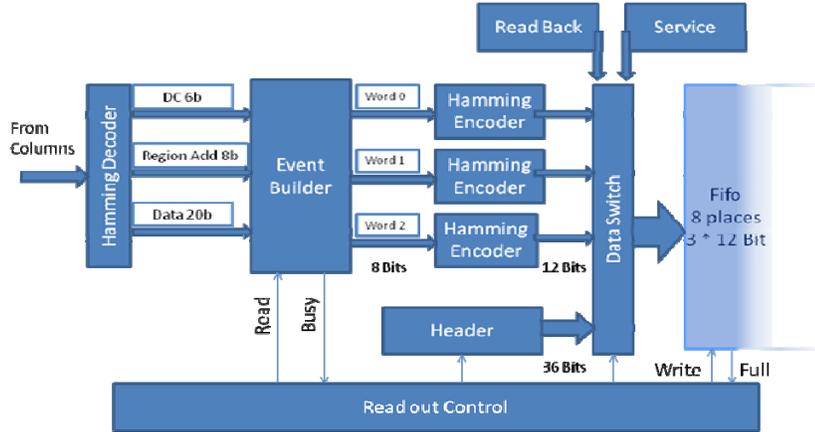


Figure 3: Sketch of the data transfer from Double-Column down to the storage FIFO.

Hit data organization in record	bandwidth reduction	bits in record
Single Pixel Hit Transfer	0%	20
2-Hit Transfer (fixed in region)	4%	23
2-Hit Transfer (across regions)	13%	24
4-Hits (in Double-Column)	5%	31
4-Hits (across Double-Columns)	15%	32

Table 1: Bandwidth reduction with respect to single pixel hit transfer and size of record as a function of record organization for a central module in the IBL.

As a consequence, the reformatting algorithm chosen consists of sending 2 pixels hits adjacent in ϕ together in the same data record, which not only reduces the data bandwidth out of the Front-End and fits a byte-based format, but also is in practice particularly simple to implement. The data records are hence stored in the FIFO in the form of 3 times 8-bit records. The digital control block takes care of providing the logic to perform the transfer of the pixel hits from the 4-pixel regions to the FIFO. It also handles the recording of other types of data such as data header (header for transmission of pixel data), service messages (e.g. error messages), address records and value records (for read back of global or local registers), or the value of the empty record word. Table 2 shows the 6 types of record word. All record words are 24 bits long. Data header, address record, value record and service record can start an event transmission and as such start with 11101 flag (similar to what is done in the current ATLAS pixel module [6]). The data stored in the FIFO are then transmitted out in slices of 8 bits by a mechanism driven from the Data Output Block.

5.2 Clock Multiplier unit, Data Output Block and 8b10b Coder

Simulation has shown [7] that to fit IBL needs, data has to be transmitted out of FE-I4 at a bandwidth of 160 Mb.s⁻¹. The LHC bunch-crossing clock and hence the clock which reaches FE-I4 is 40MHz. As the IBL is inserted inside the present ATLAS pixel detector, it needs to fit constraints of an already built-up system. In particular, sending a higher frequency clock to the

FE is thought impractical as this would require modifications of off detector elements, modifications of opto-components, and a new synchronization protocol in the FE between an incoming 80 MHz clock and the LHC beam crossing clock frequency. It was decided to instead send the 40MHz LHC clock to the FE, and perform clock multiplication in the FE. Off detector, the higher frequency clock needs to be reconstructed from the data received from the IBL. To ease this reconstruction, FE-I4 codes the data using 8b10b protocol [8] which gives an output data stream with favorable engineering properties.

24-bit Record Word	Value	Description
Data Header	11101 001 SR[1] LIT[7] bcID[8]	001 identifies Data Header. SR: flags if service word is attached. LIT: trigger ID. bcID: bunch crossing ID
Data Record	Col[7]Row[1] Row[8] ToTtop[4] ToTbot[4]	7-bit column address, 9-bit row address, 2 adjacent pixel ToT data transferred
Address Record	11101 010 Type[1] Add[7] Add[8]	010 identifies Address Record. Type: Global Register / Shift Register. 15-bit address: Global Register ID / Shift Register position
Value Record	11101 100 Value[16]	100 identifies Value Record. Value: Either contained in Global Register or in Shift Register
Service Record	11101 111 Message[16]	111 identifies Service Record. 16-bit Message information follows
Empty Record	$3 \times \text{ERvalue}[8]$	Programmable 8-bit word which is sent when 8b10b coding is turned off.

Table 2: Description of Record Words.

The FE-I4 Phase Locked Loop (PLL) is based on a phase frequency detector controlling the charging or discharging of a charge pump which feeds a differential voltage controlled ring oscillator running at 640MHz. The high frequency clock is then divided down in a succession of differential divide by two toggle Flip-Flops, and the resulting 40 MHz clock is then fed back to the phase frequency detector where it is compared to the incoming 40 MHz reference clock. The 160 MHz divided down clock is used in the Data Output Block for single edge data stream out at 160 Mb/s. Details concerning the FE-I4 Clock Generation block are given in [9]. It should also be noted that two 6 to 1 configurable multiplexers (MUX) are implemented after the PLL, and allow selecting the clocks. One MUX output clock is then used for the data stream-out in the Data Output Block, the other MUX output clock is used in a data concatenation unit together with a 4 to 1 multiplexer. These blocks allow implementation of a star-based 4-chip module with single channel data transmission from the module unit, at speeds up to 320 MHz for sLHC outer layer prototyping. The clocks fed to the MUXs are the 320, 160, 80 and 40 MHz feedback clock as well as the 40 MHz reference clock, and an auxiliary clock useful for test purposes.

By default, 8b10b coding of the data is performed, but the Data Output Block provides the option of turning 8b10b coding off. It also takes care in a state machine of the complete 8b10b framing using comma words for Start of Frame, End of Frame, and Empty Records with

beneficial properties for off-detector resynchronization in case of loss of synchronization. The output block also contains the needed 8 and 10 bit serializers and the clock divider.

5.3 Command Decoder, Registers and Front-End Configuration

The command decoder is the block handling the decoding of configuration data, both global and local, generating resets for the rest of the logic, and decoding the Level 1 Trigger incoming requests. The FE-I4 command decoder presents many similarities to the command decoder developed for the Module Controller Chip of the present ATLAS pixel FE-I3-based module [6]. Commands are classified in three classes: trigger, fast and slow. The trigger command is the shortest (fastest) command to be decoded and is based on the 11101 5-bit field. As such it fits the ATLAS trigger requirements (need to allow for a minimal delay between two triggers of 5 clock cycles) and is single-bit flip safe (single bit flips are flagged, error code issued but still the trigger is propagated with the correct timing). Fast commands are 9 bits long, and are commands related to performing resets and calibration. Slow commands are mainly used for configuration of the chip global registers or of the pixel local registers, writing and reading back, and putting the chip in or out of run mode.

As global configuration data bank, the chip uses a set of 16-bit deep registers based on custom made SEU-hard latches [10]. The 13 bits of pixel local configuration are programmed through Double-Column-based shift registers with 13 strobe signals. Complete configuration of the FE using the 40 MHz standard clock takes about 10 ms. Special care is taken to ensure the hardness of the command decoder to SEUs: The complete command decoder is triplicated with majority voting and correction provided each clock cycle. By construction, the command decoder state machine returns to idle state very quickly without need for reset. For test purposes in this first full scale FE-I4, the command decoder can be bypassed and global and local registers can be written with a shift register arrangement. Note that the command decoder as well as the digital control block and the data output block are fully scan-able.

6. Conclusion and Future Developments

FE-I4 is the next ATLAS pixel Front-End chip developed in a CMOS 130 nm technology. With a new analog pixel tuned for low power operation and major changes brought to the digital section of the pixel and to the periphery, FE-I4 is well adapted to the innermost layer of the pixel detector corresponding to the first ATLAS pixel upgrade (IBL project), as well as to the outermost pixel layers at Super-LHC. Submission of a first full scale FE-I4 chip is foreseen for beginning of 2010, and will allow both prototyping for the IBL and for 4-chip module-based Super-LHC outer layers.

As for the ATLAS pixel inner layers for Super-LHC, where even higher hit occupancy is expected and for which radiation tolerance is even more of a concern, two main options will be followed. The first one is CMOS technology scaling down and working in a smaller feature size, though the price to pay will be increasing difficulties for analog design, and a possible increase of SEU cross-section. The second is the so-called 3D integration approach, where two tiers are connected using Through Silicon Vias and direct tier to tier bonding technology. One

tier would then shelter the analog array while the other would shelter the corresponding digital one. In both cases, these technologies could achieve the reduction of the pixel size needed to decrease pixel cross-section (and resulting sources of inefficiencies), and might also call for a reorganization of the digital section if not new concepts for the analog pixel.

References

- [1] G. Aad *et al.*, *The ATLAS experiment at the Large Hadron Collider*, *JINST* **3**, S08003 (2008)
- [2] G. Aad *et al.*, *ATLAS pixel detector electronics and sensor*, *JINST* **3**, P07007 (2008)
- [3] D. Arutinov *et al.*, *Digital Architecture and Interface of the new ATLAS Pixel Front-End IC for Upgraded Luminosity*, *IEEE Trans. Nucl. Sci.* **56**, **2** (2009)
- [4] F. Faccio and G. Cervelli, *Radiation-induced edge effects in deep submicron CMOS transistors*, *IEEE Trans. Nucl. Sci.* **52**, **6** (2005)
- [5] M. Karagounis *et al.*, *Development of the ATLAS FE-I4 pixel readout IC for b-layer upgrade and Super-LHC*, “Proceedings of the Topical Workshop on Electronics for Particle Physics 2008”, in Naxos 2008, Electronics for particle physics, CERN-2008-008: 70-75 (2008)
- [6] R. Beccherle *et al.*, *MCC : The Module Controller Chip for the ATLAS Pixel Detector*, *Nucl. Instrum. Meth. A* **492**, 117-133 (2002)
- [7] A. Grillo *et al.*, *I/O Choices for the ATLAS IBL*, ATLAS upgrade document
- [8] A. Widmer and P. Franaszek, *A DC-Balanced, Partitioned-Block, 8B/10B Transmission Code*, *IBM J. Res. Develop.*, Vol. 27, **5** (1983)
- [9] A. Kruth *et al.*, *Charge Pump Clock Generation PLL for the Data Output Block of the Upgraded ATLAS Pixel Front-End in 130 nm CMOS*, “Proceedings of the Topical Workshop on Electronics for Particle Physics 2009”, in Paris 2009, Electronics for particle physics, CERN-2009-006: 548-552 (2009)
- [10] M. Menouni *et al.*, *Design and measurements of SEU tolerant latches*, “Proceedings of the Topical Workshop on Electronics for Particle Physics 2008”, in Naxos 2008, Electronics for particle physics, CERN-2008-008: 402-405 (2008)