The FE-I4 Pixel Readout Integrated Circuit

M. Garcia-Sciveres ^{a,*}, D. Arutinov^c, M. Barbero^c, R. Beccherle^f, S. Dube^a, D. Elledge^a, J. Fleury^b, D. Fougeron^d, F. Gensolen^d, D. Gnani^a, V.

Gromov^e, T. Hemperek^c, M. Karagounis^c, R. Kluit^e, A. Kruth^c, A.

Mekkaoui^a, M. Menouni^d, J-D. Schipper^e

^aLawrence Berkeley National Lab., Berkeley, CA, USA b Laboratoire de l'Accelerateur Lineaire, Orsay, France c University of Bonn, Bonn, Germany ^dCentre de Physique des Particules de Marseille, Marseille, France ^eNationaal Instituut voor Subatomaire Fysica, Amsterdam, The Netherlands f Istituto Nazionale di Fisica Nucleare Sezione di Genova, Genova, Italy

Abstract

A new pixel readout integrated circuit denominated FE-I4 is being designed to meet the requirements of ATLAS experiment upgrades. It will be the largest readout IC produced to date for particle physics applications, filling the maximum allowed reticle area. This will significantly reduce the cost of future hybrid pixel detectors. In addition, FE-I4 will have smaller pixels and higher rate capability than the present generation of LHC pixel detectors. Design features are described along with simulation and test results, including low power and high rate readout architecture, mixed signal design strategy, and yield hardening.

Key words: Pixel detector, ATLAS upgrades, High luminosity, 130 nm PACS: 29.40.Wk, 29.40.Gx, 29.85.Ca

1. Introduction

The FE-I4 integrated circuit contains readout circuitry for 26 880 hybrid pixels arranged in 80 columns on $250 \mu m$ pitch by 336 rows on $50 \mu m$ pitch. It is designed in a 130 nm feature size bulk CMOS process. Sensors will be DC

Preprint submitted to Nuclear Instruments and Methods A October 17, 2009

[∗]Corresponding author

Email address: mgs@lbl.gov (M. Garcia-Sciveres)

coupled to FE-I4 with negative charge collection. Each FE-I4 pixel contains an independent, free running amplification stage with adjustable shaping, followed by a discriminator with independently adjustable threshold. The chip keeps track of the firing time of each discriminator as well as the time over threshold (TOT) with 4-bit resolution, in counts of an externally supplied clock, nominally 40 MHz (the collision rate). Information from all discriminator firings is kept in the chip for a latency interval, programmable up to 256 cycles of the external clock. The design targets an average discriminator firing rate of 50 kHz per pixel (400 MHz/cm^2) . Within the latency interval, information can be marked for retrieval by supplying a trigger, and is otherwise automatically erased. Triggers do not interfere with storage of new data. The data that have been marked for readout are output serially over a current-balanced pair (similar to LVDS). The primary output mode is 8b/10b encoded at 160 Mb/s rate. The FE-I4 is controlled by a serial LVDS input synchronized by the external clock. No further I/O connections are required for regular operation, but several others are supported for testing.

The FE-I4 requirements were driven by two planned upgrades of the ATLAS experiment [1] at the Large Hadron Collier (LHC). The insertable B-Layer (IBL) is the medium term replacement of the beam pipe with one having an integrated pixel detector layer in the central 60 cm section. In the long term, a complete replacement of the ATLAS tracker is expected, and the new tracker must cover a larger area with pixels in order to handle increased rate and radiation from LHC luminosity upgrades. A significant cost reduction of hybrid assemblies relative to present state of the art is required for such large area coverage to be viable. The large format of FE-I4 is a significant step in this direction, since the cost of flip chip bump bonding (which is a cost driver) scales roughly with the number of chips processed, rather than with area processed. The FE-I4 is as large as allowed by the process reticle size. This in turn presents design challenges for power, bias, and signal distribution, management of power transients, and integration of a design with well over 100 M transistors. Table 1 compares the FE-I4 dimensions with presently existing hybrid pixel readout chips.

2. Architecture

The FE-I4 pixel array geometry is derived from that of the ATLAS pixel detector FE-I3 readout chip [6], wherein pixels are arranged in column pairs and power is segregated on vertical bands alternating between analog and

digital. However, the readout architecture of FE-I4 is very different [7]. Instead of moving all hits in each column pair to a global shared memory structure for later trigger processing, the FE-I4 column pairs are further divided into 2 pixel by 2 pixel regions. Each region contains 4 identical analog front ends, ending in a discriminator, and one shared memory and logic block (digital region). The digital region can store up to 5 "events". For each event, a dedicated counter keeps track of the time elapsed since the event took place with resolution given by the external clock period (nominally 25 ns). This requires distributing the clock to all regions with a maximum skew of 2 ns. When an external trigger arrives, it is also distributed to all the regions simultaneously within 2 ns. The trigger selects any event for which the time counter matches the programmed trigger latency value. When the counter exceeds the latency in the absence of a trigger, the event is erased to make room for more. The events selected by a trigger remain in the region until it is their turn to be sent off chip via the serial LVDS output. All regions can be read out if they are all selected. A "stop" readout mode is also provided where the time counters are stopped and all events stored in every region can be read (in this special mode, no acquisition of new data can take place during readout).

The individual discriminator outputs are synchronized with the clock as they feed into the digital region, and all region operations are synchronous. Each synchronized discriminator output is further processed by applying a digital cut on the time over threshold (TOT). Hits smaller than a certain TOT (programmable between 1 and 3 clock periods) are classified as "small hits" and those larger as "large hits". The next available time counter in a given region starts whenever there is at least one large hit. Small hits do not

Chip name	Width	Length	Active	Num.
	$\rm (cm)$	$\rm (cm)$	Area cm^2)	Pixels
$FE-I4$	2.0	1.9	3.36	26880
Medipix $[2]$	1.4	1.7	1.94	65536
ALICE1LHCb ^[3]	1.4	1.6	1.74	8192
PSI46 [4] $(CMS exp.)$	0.8	1.0	0.63	4 1 6 0
$FE-I3$ [5] (ATLAS exp.)	0.7	1.1	0.58	2880

Table 1: Dimensions of existing hybrid pixel readout chips compared to FE-I4

start a counter. Once a counter starts, the TOT value of all 4 pixels is always recorded (which is 0 for a comparator that did not fire). This automatically stores small hits in the same time bin as the large hit(s) belonging to the same charge deposition cluster. Thus small hits, for which timing information is degraded due to time-walk, are associated to the correct time slice by physical proximity to large hits, rather than by their own discriminator firing time. This technique exploits prior knowledge about the physical detection process, which can be qualitatively summarized as: large hits are in time, small hits are next to large hits, and isolated small hits are noise. Thanks to this added digital processing, the time-walk requirements on the analog circuitry are relaxed, allowing lower current operation and therefore worse analog timewalk performance than may be acceptable in the absence of hit association by proximity as well as time. As charge clusters will often straddle a region boundary, in addition to storing all 4 pixels within the region, 4 neighbor bits are stored to flag the presence of small hits in the pixels adjacent to the region within the double column. Clusters that straddle a double column boundary may lose small hits, but this does not result in loss of valuable position information, because the resolution along the $250 \,\mu m$ dimension of the pixel is already relatively poor by design. A time association window of two clock cycles is used to capture small hits both within the region and in the 4 neighbor pixels. Note that by setting the programmable digital TOT cut to its minimum value, one can recover traditional hit discrimination for comparative studies. In this case all comparator firings would be treated as equal and recorded in the time bin they occurred (which, due to timewalk, tends to be the wrong bin for hits just above the analog discriminator threshold).

The performance of this architecture has been evaluated using simulated charge clusters from $250 \mu m$ thick silicon sensors 3.7 cm away from protonproton collision events at 14 TeV center of mass energy. Beam collisions were simulated at a rate of 40 MHz and with a varying number of interactions per collision spanning the range expected for LHC luminosity upgrades. A parametrized model for the analog front end response was coupled with behavioral simulations of a full double column and digital readout chain. Fig. 1 shows the loss of hits observed in these simulations vs. the number of interactions per beam collision. The total fraction of lost hits is shown as well as the component due to simple pile-up within a single pixel (where a particle hits a pixel that is still integrating charge from a previous collision). The pile-up component is a function of the pixel size and amplifier shaping time,

while all other losses are due to the readout architecture. The leading cause of hit loss for IBL operation will be pixel pile-up, but eventually loses in the digital region due to fully occupied memories dominates. The architecture scales to higher rate capability by increasing the number of events that can be stored in a region, which could be done by using denser circuitry, by storing less information per event, or by grouping more pixels per region. Fig. 1 also shows the simulated hit loss for the present ATLAS pixel detector chip (FE-I3) under these same conditions. The simple pile-up component is larger for FE-I3 because the pixels are larger and the 8 bit TOT digitization requires slower pulse shaping. Nevertheless, the FE-I3 rate capability has a strict limit imposed by the readout architecture, wherein all hits (before any trigger) must be copied to a global memory at the end of the column pair. The issue here is not memory, but available bandwidth to copy hits along a column pair. Fig. 1 shows that the FE-I3 chip could not be used for the IBL upgrade. Scaling this architecture to higher rates and/or longer columns is problematic, because bandwidth increases require more power and generate more noise, and because more memory in the periphery means less of the chip footprint can be active (the FE-I4 is more than twice as long as FE-I3).

3. Analog Design

A 2 stage amplifier configuration is used for the analog front end. The $1st$ stage (preamp) is a straight cascode integrator with an NMOS as input device operated in weak inversion. The preamp output is AC-coupled to a 2^{nd} stage folded cascode with PMOS input amplifier. The main motivation of this 2 stage system is to provide enough gain in front of the discriminator while allowing some independence in the optimization of the preamp feedback capacitor. Shaping of the input pulse is achieved by continuous reset of the preamp with a constant current feedback, which is optimized for a return to baseline of 400 ns after a -4 fC input pulse. The DC leakage current of the sensor is supplied by a leakage compensation current source in parallel with the input. The 2^{nd} stage provide only voltage gain and no shaping. The gain is approximately 60 mV/fC (6 V/V) for the first (second) stage, but the charge measurement linear range is greater than suggested by the combined $350 \,\mathrm{mV/fC}$ total gain and the 1.4 V nominal supply voltage, thanks to the TOT digitization method used. Saturation of the 2^{nd} stage does not affect the TOT measurement (recall all shaping is done by the first stage). More details about the analog front end can be found in [8].

The front end has been optimized to remain functional down to very low bias current. The nominal front end current for FE-I4, assuming a 400 fF sensor load, is $10 \mu A$ per pixel. For comparison, the nominal front end current for the present ATLAS pixel detector, scaled to the FE-I4 pixel size, would be $17 \mu\text{A}$. At the nominal FE-I4 current the analog time-walk performance is worse than in FE-I3, and in fact a current close to $17 \mu A$ per pixel would be required to match the present detector analog time-walk. However, as explained in the previous section, the FE-I4 region architecture compensates for the degraded analog timing performance with digital processing. For the ATLAS detector application we define a time-walk figure of merit ("overdrive") as the charge above threshold at which the delay at the discriminator output is 20 ns greater than for a maximum charge pulse. The FE-I4 overdrive at nominal current is $2000 e^-$, while the present ATLAS detector operates with $1500 e^-$ overdrive. This nominal FE-I4 overdrive can be seen in Fig. 2, which is the measured time delay at the discriminator output for a prototype of the FE-I4 front end circuit, after irradiation to 200 MRad (Si).

The threshold dispersion caused by device mismatch is kept low by the AC coupling between the first and second stages. Threshold dispersion can only arise from the 2^{nd} stage and the discriminator itself. In a prototype chip with 854 pixels an RMS dispersion of $300 \pm 100 e^-$ equivalent input charge has been measured after 200 MRad (Si) irradiation. The uncertainty is due to the absolute charge scale calibration, which is challenging for a test circuit with no sensor. The intrinsically low dispersion is further corrected with an independent threshold adjustment in each pixel. A 5-bit DAC with adjustable LSB size is used in order to ensure enough range to correct all pixels in the large FE-I4 chip. The simulated noise with 400 fF load and 100 nA DC leakage current is under $200 e^-$. Existing prototypes without a proper sensor load and leakage current provide only approximate noise information. Noise measurements in prototypes with an approximate load and no leakage current show $120 \pm 50 e^{-}$ after 200 MRad (Si), where again the uncertainty is due to the absolute charge scale.

A characteristic of great interest to detector design is the minimum stable threshold that can be set. Unfortunately this is dominated by coherent effects in the whole system of chip, sensor, off chip power and ground distribution, and power supply bypassing, and reliable simulations are not available. The FE-I4 design attempts to minimize coherent effects by isolating digital activity from the substrate and minimizing digital transients. In the region architecture the digital current increase caused by a new hit is small, because hit data are not transmitted outside the region. The isolation of digital activity from the substrate relies on a deep implant process option. The region logic is synthesized using a commercial standard cell library that itself has no mixed signal provisions, but the entire circuit is placed over a deep implant that isolates it from the rest of the substrate. This isolation method is used for all digital circuits in FE-I4.

4. System Features

The power distribution within FE-I4 relies heavily on low resistivity metal options in the CMOS process used. The possibility to distribute power with negligible voltage drops in spite of the chip large size enables simple bias distribution where current sources in every pixel are controlled by common gate voltages generated by global bias mirrors. The nominal operating voltage for analog circuits is 1.4 V in order to support a wide operating margin, while for digital circuits it is 1.2 V to minimize digital current consumption. Nevertheless all circuits have been designed to meet specifications in the full range 1.2 V to 1.5 V. While internally there is a strict distinction between analog and digital power, the FE-I4 is designed to operate from a single external supply. The chip contains two linear regulators each with shunt and series capability [9] as well as a charge pump divide-by-two voltage converter capable to supply the full chip current at 90% efficiency. These devices can be used in any combination to deliver the internal voltages from a single external supply. The DC-DC converter uses transistors with the maximum gate oxide thickness available in this process, in order to safely operate at 3.3V input. These transistors are expected to degrade with radiation dose more than the thin oxide core transistors. However, they are only used as digital switches for which threshold shifts are relatively unimportant. A small scale prototype of the DC-DC converter was irradiated to 200MRad(Si) and a 30% increase in the switch on-resistance was measured. In a device operating at 90% efficiency before irradiation, this implies a drop to 87% efficiency after irradiation.

The FE-I4 has significant digital functionality both within the pixels (digital region) and in the periphery. A high level description followed by synthesis was therefore used for all digital logic. The configuration memory, however, is produced with custom analog design methods because it must be robust against single event upsets (SEU). Single latches are used to store bits within each pixel (such as the 5-bit threshold adjustment), while triple redundant latches are used in the periphery. All these custom designed latches are powered from the analog voltage net, because the higher voltage increases the SEU threshold, while the latches store only static values and produce no transients during data acquisition. More details can be found in [10]. SEU tolerance is generally not an issue for switched latches within the digital logic, because the time each data bit spends in FE-I4 (of order microseconds) is much less than the worst case mean time between upsets for typical standard cell library latches (of order minutes). Thus the potential data loss due to SEU is negligible compared to the operational data loss caused by pile-up and region memory overflow (Fig. 1). The more pressing challenges for the digital logic are yield and testability. The very large size of FE-I4 would translate into an extremely low fabrication yield if a single defect anywhere in the chip were cause for rejection. However, a pixel array is inherently defect tolerant because 100% of the pixels are not required to function for the application of building a pixel detector. A fraction of dead pixels in the 10[−]⁴ - 10[−]³ range is acceptable, since this is far less than other unavoidable data losses, such as pixel pile-up. Most of the area of FE-I4 can therefore tolerate several single defects other than power shorts. Using defect densities calculated from yields observed in other device fabrications in this process, we estimate a 40% yield for zero faults in circuits extending to more than a single pixel. In order to further improve the expected yield, digital circuits with distributed functionality have been made fault tolerant. The data and address busses in the digital columns are Hamming coded, the read token in the columns is triple redundant, the data management in the chip periphery is Hamming coded, and the configuration register used to serially program all pixels in each column pair is redundant and selectable in a non-volatile memory available in this process, that can be burned at wafer probing (there was not enough space for triple redundancy in this case). With these features the yield for fully functional FE-I4 chips is expected to be between 50% and 70%.

5. Conclusion

The FE-I4 will be the largest hybrid pixel readout chip to date, which will translate into a significant cost reduction for hybrid pixel detectors. Despite its large size a relatively high fabrication yield is expected due to the inherent fault tolerance of pixel arrays and the error tolerant design of extended digital circuits. Beyond the large format, the FE-I4 includes advances in

radiation tolerance, low power, and high rate capability. These arise from a combination of favorable features in the CMOS process used and a new readout architecture. All digital circuits are synthesized from high level descriptions using a commercial standard cell library, and they are isolated form the substrate using a deep well structure provided by the process. The basic analog performance has been validated with small prototypes irradiated to 200Mrad(Si). The first wafer scale submission is planned at the end of 2009. The new "region" architecture scales to higher rates by increasing the amount of memory distributed in the array. Higher rate capability than FE-I4 will be needed for the inner most layers of long term LHC luminosity upgrades. The options to achieve such scaling, while at the same time reducing the pixel size, are smaller feature size processes and 3D integration of two 130nm feature tiers. The FE-I4 pixel array naturally lends itself to the 2 tier approach because the analog pixel front end and the digital region area per pixel are of similar size. Reduction of pixel size must accompany higher rate scaling in order to control the pixel pile-up component of the total hit loss.

References

- [1] G. Aad, et al.,"The ATLAS experiment at the Large Hadron Collider," JINST 3, S08003 (2008).
- [2] X. Llopart and M. Campbell, "First test measurements of a 64k pixel readout chip working in single photon counting mode," Nucl. Instrum. Meth. A 509, 157 (2003).
- [3] A. Kluge [ALICE Collaboration],"The ALICE silicon pixel detector front-end and read-out electronics," Nucl. Instrum. Meth. A 560, 67 (2006).
- [4] H. C. Kastli et al., "Design and performance of the CMS pixel detector readout chip," Nucl. Instrum. Meth. A 565, 188 (2006
- [5] I. Peric et al., "The FE-I3 readout chip for the ATLAS pixel detector," Nucl. Instrum. Meth. A 565 (2006)
- [6] G. Aad et al.,"ATLAS pixel detector electronics and sensors," JINST 3, P07007 (2008).
- [7] D. Arutinov et al.,"Digital Architecture and Interface of the new ATLAS Pixel Front-End IC for Upgraded LHC Lu- minosity," IEEE Trans. Nucl. Sci. **56, No.2**, (2009).
- [8] M. Karagounis et al., "Development of the ATLAS FE-I4 pixel readout IC for b-layer upgrade and Super-LHC," Published in "Naxos 2008, Electronics for particle physics" CERN-2008-008: 70-75, (2008).
- [9] M. Karagounis et al., "An Integrated Shunt-LDO Regulator for Serial Powered Systems", in "Proceedings of the 35th European Solid-State Circuits Conference" (2009).
- [10] M. Menouni et al., "Design and measurements of SEU tolerant latches," Published in "Naxos 2008, Electronics for particle physics" CERN-2008-008: 402-405, (2008).

Figure 1: Simulated data losses for FE-I4 and FE-I3 chips placed at 3.7 cm from protonproton beam collisions at 14 TeV. The expected number of interactions per beam crossings during IBL operation is 75.

Figure 2: Measured time delay from charge injection to discriminator output in an FE-I4 front end prototype after irradiation to 200 MRad (Si).