Prepared for submission to JINST

RD53 Pixel Readout Integrated Circuits for ATLAS and CMS HL-LHC Upgrades

CERN-LHCC-RD53 Collaboration

G.Alimonti[11] **A.Andreazza**[11] **F.Arteche**[22] **M.B.Barbero**[1] **P.Barrillon**[1] **R.Beccherle**[16] **G.Bonomelli**[20] **G.M.Bilei**[15] **W.Bialas**[3] **D.Bortoletto**[5] **G.Calderini**[26] **A.Caratelli**[3] **A.Cassese**[9] **J.Christiansen**[3] **E.Conti**[3,15] **F.Crescioli**[26] **M.Daas**[35] **L.Damenti**[9,10] **S.D'Auria**[11] **F.De Canio**[13] **G.De Robertis**[8] **N.Demaria**[17] **J.DeWitt**[31] **Y.Dieter**[35] **A.Dimitrievska**[25] **W.Erdmann**[29] **S.Esposito**[3] **D.Exarchou**[3] **D.Fougeron**[1] **L.Gaioni**[13] **M.Garcia-Sciveres**[25] **D.Gnani**[25] **C.Gozalez Renteria**[36] **M.Grippo**[17,18] **A.Guardino**[36] **M.Hamer**[35] **T.Heim**[25] **T.Hemperek**[35] **F.Hinterkeuser**[35] **S.Huiberts**[34] **L.M.Jara Casas**[3] **J.J.John**[5] **J.Kampkötter**[33] **M.Karagounis**[33] **I.Kazas**[27] **Y.Khwaira**[23] **R.Kluit**[28] **D.Koukola**[3] **A.Krieger**[25] **H.Krüger**[35] **J.Lalic**[3] **M.Lauritzen**[34] **F.Licciulli**[8] **Peilian Liu**[21] **F.Loddo**[8] **E.Lopez Morillo**[6] **A.Lounis**[23] **F.Luongo**[17,18] **M.Manghisoni**[13] **S.Marconi**[3,15] **F.Marquez Lasso**[6] **C.Marzocca**[31] **K.Mauer**[35] **A.Mekkaoui**[4] **Lingxin Meng**[24] **M.Menichelli**[15] **M.Menouni**[1] **M.Minuti**[16] **M.Mironova**[25] **S.Miryala**[2] **M.Missiroli**[29,32] **E.Monteil**[17,18] **K.Moustakas**[35] **F.Muñoz Chavero**[6] **G.Neue**[7] **S.Orfanelli**[3] **A.Paccagnella**[12] **L.Pacher**[17,18] **F.Palla**[16] **F.R.Palomo Pinto**[6] **A.Papadopoulou**[25] **A.Paterno**[17,18] **A.R.Petri**[11] **P.Placidi**[15] **R.Plackett**[5] **A.Pradas**[22] **A.Pulli**[3] **B.Raciti**[19] **L.Ratti**[14] **V.Re**[13] **A.Rehman**[34] **P.Rymaszewski**[35] **P.Sander**[20] **M.C.Solal**[23] **M.Standke**[35] **B.Stugu**[34] **E.Thompson**[25] **G.Traversi**[13] **D.Vogrig**[12] **M.Vogt**[35] **Tianyang Wang**[35] **Hongtao Yang**[37] **J.Zdenko**[7]

[1] *Aix Marseille Université, CNRS/IN2P3, CPPM, Marseille, France*

[2] *Brookhaven National Laboratory, Upton, NY, USA*

[3] *CERN, European Organization for Nuclear Research, Geneva, Switzerland*

[4] *Clevert Systems LLC , West Henrietta, NY, USA*

[5] *Dept. of Physics, Oxford University, Oxford, United Kingdom*

[6] *ETSI, Universidad de Sevilla, Sevilla, Spain*

[7] *Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague, Prague, Czech Republic*

[8] *INFN Sezione di Bari, Bari, Italy*

[9] *INFN Sezione di Firenze, Florence, Italy*

[10] *Università di Firenze, Florence, Italy*

[11] *INFN Sezione di Milano and Università degli Studi di Milano, Milano, Italy*

[12] *INFN Sezione di Padova and Università di Padova, Padova, Italy*

[13] *INFN Sezione di Pavia and Università di Bergamo, Bergamo, Italy*
[14] *INFN Sezione di Pavia and Università di Pavia, Pavia, Italy*
[15] *INFN Sezione di Perugia and Università di Perugia, Perugia, Italy*
[16] *INFN Sezione di Pisa, Pisa, Italy*
[17] *INFN Sezione di Torino, Torino, Italy*
[18] *Università di Torino, Torino, Italy*
[19] *Institut für Experimentalphysik, Universität Hamburg, Hamburg, Germany*
[20] *Institute for Particle Physics, ETH, Zurich, Switzerland*
[21] *Institute of High Energy Physics, Beijing, People's Republic of China*
[22] *Instituto Tecnologico de Aragon, Zaragoza, Spain*
[23] *Laboratoire de Physique des 2 Infinis Iréne Joliot Curie, Orsay, CNRS / Universitè Paris-Saclay, Paris, France*
[24] *Lancaster University, Lancaster, United Kingdom*
[25] *Lawrence Berkeley National Laboratory, Berkeley, CA, USA*
[26] *LPNHE, Sorbonne Université, Université Paris Cité, CNRS, Paris, France*
[27] *National Center for Scientific Research, DEMOKRITOS, Agia Paraskevi, Greece*
[28] *National Institute for Subatomic Physics (NIKHEF), Amsterdam, Netherlands*
[29] *Paul Scherrer Institut, Villigen, Switzerland*
[30] *Politecnico di Bari, Bari, Italy*
[31] *SCIPP, University of California, Santa Cruz, CA, USA*
[32] *Universität Zürich, Zürich, Switzerland*
[33] *University of Applied Sciences and Arts Dortmund, Dortmund, Germany*
[34] *University of Bergen, Bergen, Norway*
[35] *University of Bonn, Bonn, Germany*
[36] *University of California, Berkeley, CA, USA*
[37] *University of Science and Technology of China, Hefei, China*
E-mail: jorgen.christiansen@cern.ch

Abstract: The RD53 collaboration has since 2013 developed new hybrid pixel detector chips with 50×50 um^2 pixels for the HL-LHC upgrades of the ATLAS and CMS experiments at CERN. A common architecture, design and verification framework has been developed to enable final pixel chips of different sizes to be designed, verified and tested to handle extreme hit rates of 3 GHz/cm^2 (up to 12 GHz per chip) together with an increased trigger rate of 1 MHz and efficient readout of up to 5.12 Gbits/s per pixel chip. Tolerance to an extremely hostile radiation environment with 1 Grad over 10 years and induced SEU (Single Event Upset) rates of up to 100 upsets per second per chip have been major challenges to make reliable pixel chips. Three generations of pixel chips, and many specific mixed signal building blocks and radiation test chips, have been submitted and extensively tested to get to final production chips. The large, complex and high rate pixel chips have been developed with a strong emphasis on low power consumption together with a concurrent development and qualification of novel serial powering at chip, module and system level, to minimize detector material budget.

Keywords: VLSI circuits, Radiation-hard electronics, Front-end electronics for detector readout, Particle tracking detectors (Solid-state detectors)

Contents

1 Introduction and requirements

This paper gives an overview of the general requirements, design, architecture and measured performance of the RD53 pixel chips, developed for the ATLAS and CMS High Luminosity Large Hadron Collider (HL-LHC) upgrades. This development has been a major effort by a large number of people (~100) over 10 years. The RD53 collaboration [1], with 24 institutes, was established in 2013 to develop the required hybrid pixel detector readout integrated circuits for the ATLAS [2] and CMS [3] pixel detector upgrades for the HL-LHC. The two experiments have very similar requirements to their pixel detector upgrades and both are using LPGBT (Low Power GigaBit Transceiver) links [4] for control and readout. It was therefore agreed to do such a challenging chip development in common among ATLAS and CMS pixel detector groups with ASIC design and test experience. A common architecture, design and verification framework has been developed to make final production pixel chips, with slightly different chip sizes to enable optimal integration into the two pixel detector systems.

The general layout of the ATLAS and CMS pixel detectors are indicated in figure 1 together with figure 2 showing the anticipated track density for each 40 MHz bunch collision at the HL-LHC.

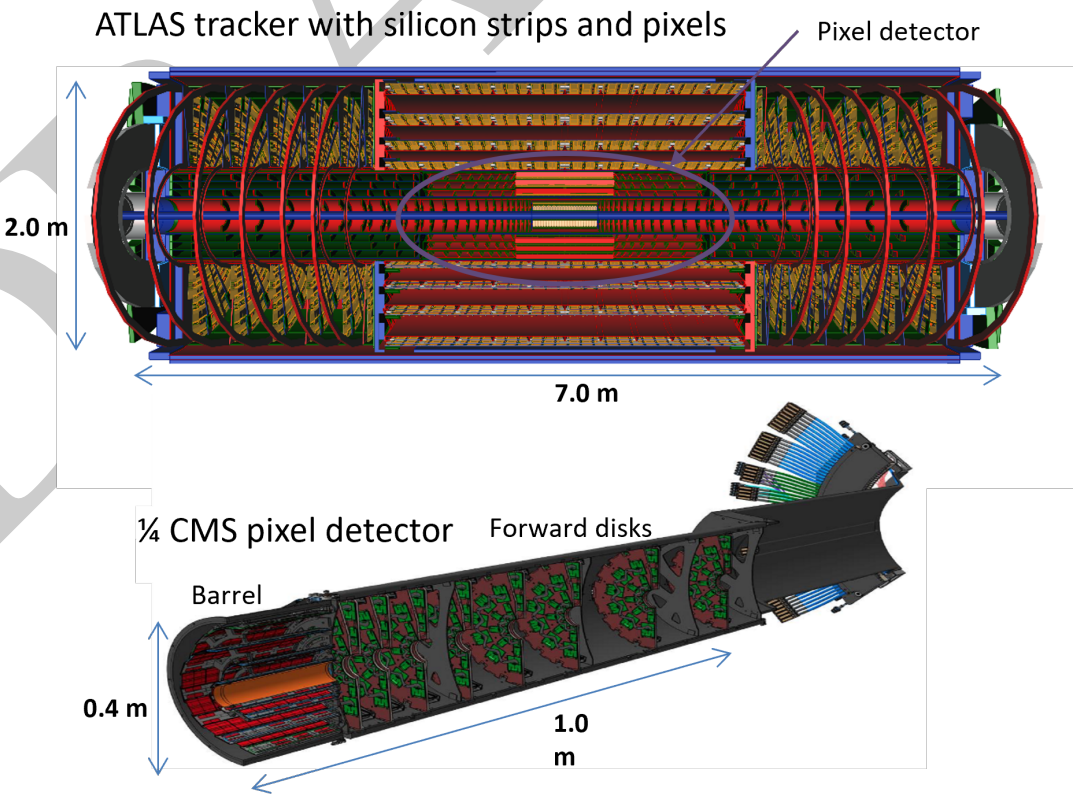


Figure 1. Upper: ATLAS tracker layout with pixel detector at its centre. Lower: One quarter of CMS pixel detector layout. Both pixel detectors are highly compact at the centre of the experiments with critical material budget and difficult access. Inner layers are specifically constructed to enable partial replacement during long shutdowns, in case of significant performance degradation from radiation damage in pixel sensors or pixel chips.

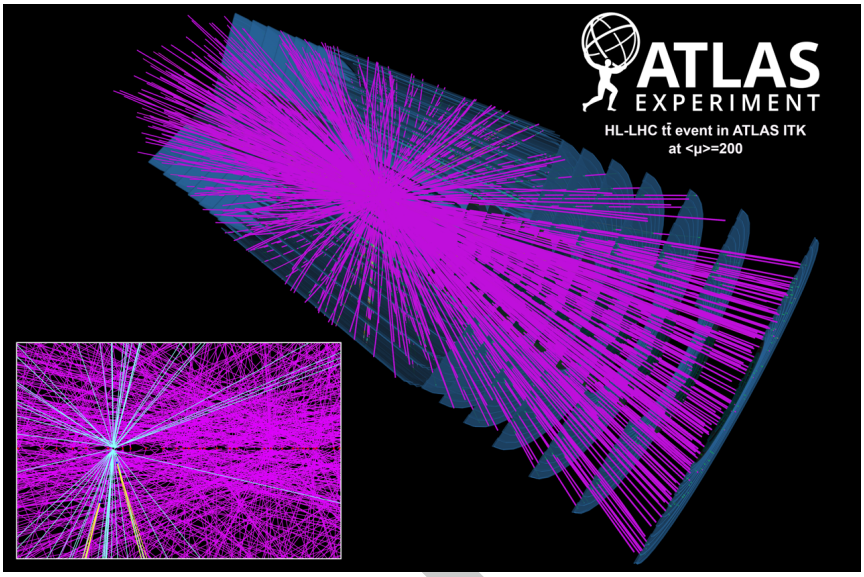


Figure 2. Anticipated track density at HL-LHC in the ATLAS tracker detector with a pileup of 200 inelastic proton-proton collisions per 25 ns bunch crossing with up to 3 GHz/cm^2 pixel hits in innermost pixel layer.

RD53 chips have been developed to meet the stringent rate and radiation requirements for operation at the HL-LHC, projected to begin operation in 2030. The HL-LHC will operate at an instantaneous luminosity of up to 7.5 10^{34} cm^{-2}s^{-1} corresponding to an average pileup of 200 inelastic proton-proton collisions per bunch crossing. This translates into an average pixel hit rate of up to 3 GHz/cm^2 in the innermost pixel layer at the 40 MHz bunch crossing rate. Inner pixel layers will have to work reliably in an extremely hostile radiation environment with up to 1 Grad Total Ionizing Dose (TID) and a Non Ionizing Energy Loss (NIEL) dose of 10^{16} 1 MeV n eq cm^{-2} over 10 years operation. It is assumed that innermost pixel layer(s) will possibly need replacement after 5-10 years, depending on the actual integrated luminosity and pixel sensor and chip performance degradation. The integration of the pixel detectors in the experiments has been made to enable partial replacement of inner pixel layer(s). The extreme radiation levels require the pixel chip design to be made with a strong emphasis on rad-hard design and effective SEE (Single Event Effects) protection. An inner layer pixel chip can be estimated to have up to 100 Hz of SEUs (Single Event Upsets) and SETs (Single Event Transients) and must function reliably despite these upsets in its internal data buffers, state-machines and configuration registers. This unprecedented radiation tolerance requirement is a factor ~10 higher than what has previously been made for High Energy Physics (HEP) applications and a factor ~10,000 higher than normally required for rad-hard space applications.

The general chip requirements are outlined in table 1. The required pixel size of 50×50 μm^2 for forward pixel layers, or 25×100 μm^2 for central barrel layers is ~4 times smaller than previous generation ATLAS/CMS pixel chips. This combined with the increased hit rate (factor ~4) and extended trigger latency (factor ~2) implies that effective trigger latency hit buffering has been increased by a factor of more than 10 compared to current ATLAS [7] and CMS [8] pixel detectors. The increased trigger rate, from 100 kHz to 1 MHz, combined with higher hit rate and smaller pixels implies that effective readout bandwidth is increased by a factor ~100, maintaining a 4 bit

charge measurement per pixel hit. Requirements for pixel sensor capacitance and radiation induced leakage together with appropriate charge detection threshold have been determined from scaling from previous pixel detectors and measurements on pixel sensor prototypes in the two experiments.

Parameter	Value (CMS/ATLAS)
Technology	65 nm CMOS
Max. hit rate	3.0 GHz/cm^2
Trigger rate	750 kHz / 1 MHz
Trigger latency	12.5 µs
Pixel size (chip)	50 x 50 µm^2
Pixel size (sensor)	50 x 50 µm^2 or 25 x 100 µm^2
Pixel array	432 x 336 pixels / 400 x 384 pixels
Chip dimensions	21.6 x 18.6 mm^2 / 20 x 21 mm^2
Detector capacitance	< 100 fF (200fF for edge pixels)
Detector leakage	< 10 nA (20nA for edge pixels)
Min. threshold	1000 e-
Charge measurement	4 bit TOT, max 1% deadtime at 3.0 GHz/cm^2
Radiation tolerance	1 Grad over 10 years at -15°C
SEE tolerance	SEU rate, innermost: ~100Hz/chip
Power	< 1W/cm^2, Serial powering
Readout data rate	1-4 links @ 1.28Gbits/s = max 5.12 Gbits/s
Temperature range	-40°C ÷ 40°C (Nominal operation=-20°C ÷ -10°C)

Table 1. General requirements to RD53 pixel chips for use in ATLAS and CMS pixel detectors at HL-LHC.

At the pixel detector level it is critical to keep the material budget of pixel detector modules and related services as small as possible, so as not to deteriorate significantly tracking performance from particle scattering and conversions in the pixel detector and its related cooling, powering and readout services. The chip power consumption must be kept as small as possible, at similar level as the previous generation pixel chips, despite significantly higher pixel density and complexity with higher hit and readout rates. The use of a scaled CMOS technology is critical to keep an acceptable power consumption, from reduced capacitive loading of on-chip gates and the reduction of power supply voltage ($V_{DD}{}^2$ dependency). An unfortunate side effect of power supply voltage scaling is that for the same power consumption, the required power supply current increases, posing problems making an appropriate low mass power distribution system. The use of switched mode DC-DC power conversion on the pixel chip, or on the pixel module, was investigated, but excluded because of the required radiation tolerance and associated space and material budget of local inductive or capacitive power converters. A novel serial powering scheme has therefore been adopted, with on-chip SLDO (Serial Low DropOut) power regulators, based on initial feasibility demonstrations

with the FEI4 chip [5][26][31]. This particular serial power distribution system has been developed, tested and qualified at the chip, module and system level, while the RD53 pixel chips were actively being developed.

The first years of development in RD53 were focused on radiation tolerance studies of the chosen 65nm CMOS technology and implementing and testing the required radiation hard building blocks: Digital to Analog Converters (DAC), Analog to Digital Converter (ADC), Analog pixel Front-Ends (AFE), Biasing structures, Band-gap reference, Phase Locked Loop (PLL), Input-Outputs (IO), SLDO power regulator and temperature and radiation sensors. An appropriate hit buffering, processing and readout architecture for the high hit and trigger rates was developed and extensively simulated and verified in a flexible simulation and verification framework with detector Monte Carlo hit data.

A first 1/2 sized pixel chip called RD53A, submitted in 2017 on a shared submission, has been used for verification of developed building blocks and general architecture. RD53A has also been instrumental as a test vehicle to test and qualify different pixel sensors [6] and for system studies, covering serial powering, design and testing of pixel modules and testing with LPGBT based readout system with optical links to the off-detector DAQ. A large set of irradiation test campaigns have been made with this chip to get a good understanding of reliable functionality of such a complex chip covering TID (Total Ionizing Dose) effects as a function of temperature, dose rate effects and initial SEE tests. Three different analog front-ends were present in this chip together with two different trigger latency buffering schemes to determine the most appropriate implementation for final chips.

A second generation of RD53 chips, named RD53B-ATLAS & RD53B-CMS [9][10], are complete full sized pixel chips made with the chosen latency buffer architecture and improved building blocks. RD53 developed a flexible parameterized design and verification environment where full custom macros and Register Transfer Level (RTL) code are instantiated according to the specific ATLAS or CMS implementations. The RD53B generation chips were made specifically for each experiment (RD53B-ATLAS, known as ITkPixV1 in ATLAS, and RD53B-CMS, known as CROCv1 in CMS) with their specific AFEs and chip size adapted to specific integration constraints of each experiment. These two chips, submitted in 2020 and 2021, are functionally equivalent with the same control and readout interfaces, with minor specific features related to the analog front-ends and specific features and bugs. The RD53B-ATLAS chip unfortunately had non-functional TOT (Time Over Threshold) charge measurement and could only be used with binary readout (but had working Hit-OR TOT readout). The RD53B generation chips have been instrumental for extended chip testing in RD53 and pixel module and system developments, testing and qualification in the ATLAS and CMS pixel detector groups. The evolution of the RD53 chips is shown in figure 3.

Bug fixes and improvements have been made in the final generation production chips: RD53C-ATLAS and RD53C-CMS [11]. Monitoring functions have been improved and extended. SEU and SET tolerance have been significantly improved based on extensive RD53B ion, proton and laser beam testing and SEU/SET simulations at transistor, gate and RTL level. Testing of serially powered quad chip pixel detector modules, in the ATLAS and CMS pixel detector groups, have

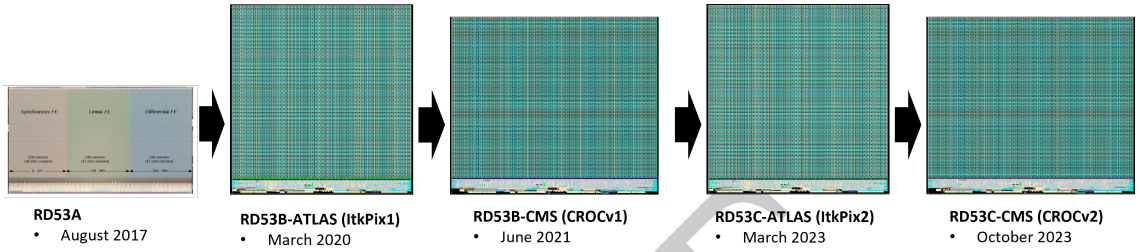


Figure 3. RD53A to RD53C chip generations with chip submission dates.

enabled system issues to be identified and corrected. An extended verification framework was specifically developed for exhaustive functional and SEU/SET verification.

The large scale prototype chips RD53A, RD53B-ATLAS and RD53B-CMS have been produced and extensively tested as reported in this paper. Final production version chips, RD53C-ATLAS and RD53C-CMS, have recently been submitted and are in production for use in the experiment upgrades. The RD53C chips have recently been through extensive chip testing and characterization, with test results as reported in this paper. They are used for pixel module pre-production for large scale system tests. Wafer level production test setups have been developed and qualified for the two experiments.

Test results shown in this paper is in general for the bare pixel chip without a pixel sensor, unless specifically mentioned in the figure caption. Bump bonded pixel chip and pixel sensor assemblies have only recently become available in sufficient quantity and quality to make detailed chip characterization these and measurements of these are shown in Section 12. Extensive pixel module test, characterization and qualification is currently ongoing in the ATLAS and CMS pixel groups with their specifically chosen pixel sensors.

The chip architecture and implementations are outlined together with circuit details of critical blocks to achieve required performance in the hostile radiation environment. Most of the discussions make no distinction between the ATLAS and CMS chips, as these are based on a common architecture with only minor implementation differences. The paper is organized as follows. Section 2 provides a short overview of the planned use of the pixel chips in the ATLAS and CMS pixel detectors. Section 3 gives an overview of the pixel chip architecture. Section 4 describes the analog front-ends, performing pixel hit detection with a 4 bit charge measurement. Section 5 outlines hit data buffering during the trigger latency and following data processing. Section 6 defines the control and readout interfaces. Section 7 covers the on-chip serial power regulator and the generation of biasing and references. Section 8 describes the implemented on-chip monitoring features. Section 9 summarizes radiation tolerance aspects. Section 10 outlines the integration and implementation. Section 11 describes final functional and SEU/SET verification. Section 12 summarizes general test results and wafer probing. Finally section 13 concludes the paper.

2 Pixel detector system

The upgraded pixel detector systems of ATLAS [2] and CMS [3] are made to have the same front-end control, readout and powering interfaces defined for the RD53 pixel chips. Pixel chips are integrated on dual, triple or quad pixel chip modules with a single bump-bonded pixel sensor. Effective hit rates and required readout rates have a strong dependency (r^{-2}) on the radial distance to the interaction point. Inner layer chips require up to 5.12 Gbits/s readout bandwidth, while pixel chips in outer layers only need a factor 20 - 50 lower readout bandwidth (depending on detector layout, number of layers and barrel versus forward). The readout via the LPGBT has therefore been defined to allow a high level of readout link modularity to minimize the number, and material, of required readout links. Each pixel chip can use from 1 up to 4 serial Electrical links (E-links) at 1.28 Gbits/s (LPGBT max E-link speed). RD53 pixel chips can also be used in a primary-secondary configuration, where readout data from 2 or 4 pixel chips are merged into a single 1.28 Gbits/s link as illustrated in figure 4 with pixel module prototypes shown in figure 5. Multiple chips on the same pixel module are controlled with a single 160 Mbits/s control link carrying clock, configuration and real-time control information. In ATLAS, control and readout links between pixel chips and the LPGBT are up to 8 m long [24], using a dedicated cable driver and equalizer GBCR ASIC [23] . In CMS, with an E-link distance limited to 1.5 m, the pixel chips are connected directly to the LPGBT.

Serial powering is used in both pixel detector systems to minimize the material budget for the power distribution. The power supply current on a single power cable pair is used to power up to 16 pixel modules in series as shown in figure 6. Multiple pixel modules are powered in series with a constant current and on-chip SLDO regulators dynamically adjust their chip power impedance, to have constant and well regulated local voltages for the analog and digital parts of the chip. In such a powering scheme it is critical to minimize fluctuations in circuit power consumption and have sufficient current headroom and local decoupling capacitors to enable the local power regulators to absorb such fluctuations. Each chip has separate analog and digital SLDO regulators, connected in parallel to the common input power, to assure the best possible isolation of the sensitive analog front-ends from switching noise induced in the power rails by the digital circuits. The effective load impedance is dynamically regulated with a controlled shunt current, maintaining constant input voltages and currents, independently of the actual current consumed by the analog and digital circuits in the chip.

Such a constant current (and constant voltage) powering system is highly advantageous in systems where low noise is primordial and where long power cables will have significant voltage drops. Assuring constant power supply currents also prevents power cables in strong magnetic fields to have induced dynamic forces with possible resonances. System drawbacks of serial powering are the power dissipated in the on-chip regulators, with necessary current (10-20 %) and voltage (0.2-0.3 V) headrooms, adding up to a total power overhead of 20-40 %. In particular, if the pixel chip goes into a low power state, the on-chip regulators will have to dissipate the full nominal chip power. Both pixel detectors will be cooled with highly efficient high pressure CO_2 cooling systems that will be designed to cope with this. Another system issue with serial powering is the requirement that all control and readout links to/from the pixel chips must have AC coupling, with appropriate DC balanced link encoding. Special care must be taken for system grounding as pixel module

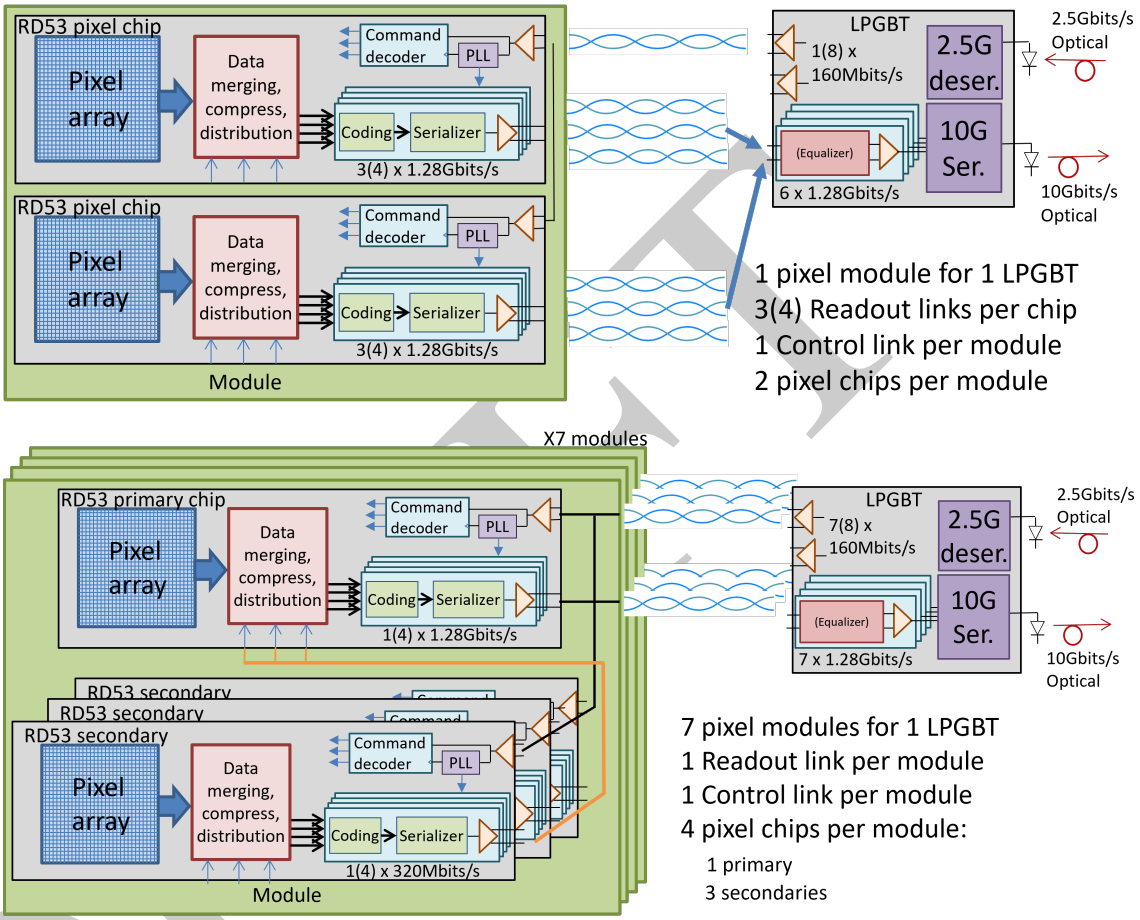


Figure 4. Pixel module control and readout. Upper: Dual chip inner module with multiple (CMS:3 or ATLAS: 4) readout E-links per pixel chip. Lower: Quad chip outer module with data merging to single readout link.

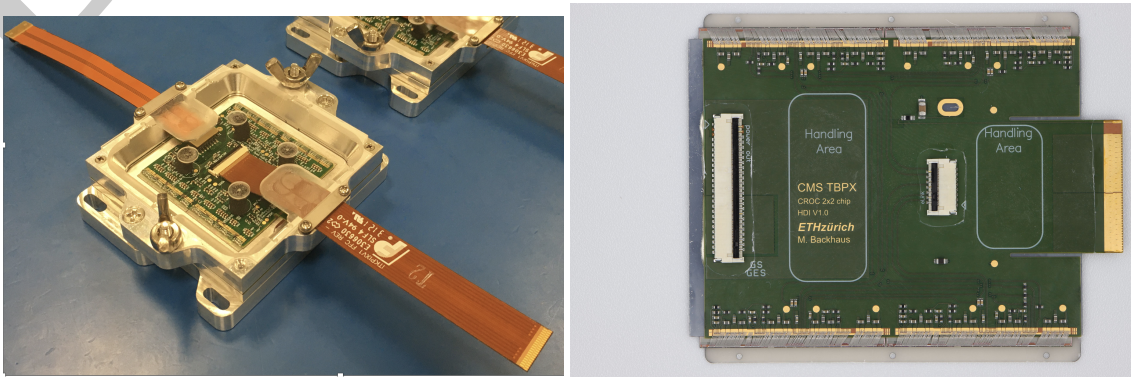


Figure 5. Pixel modules with RD53B chips and bump-bonded sensor (behind Kapton pixel hybrid with wire-bonds visible along top and bottom sides). Left: ATLAS quad module in its protection box. Right: CMS quad module.

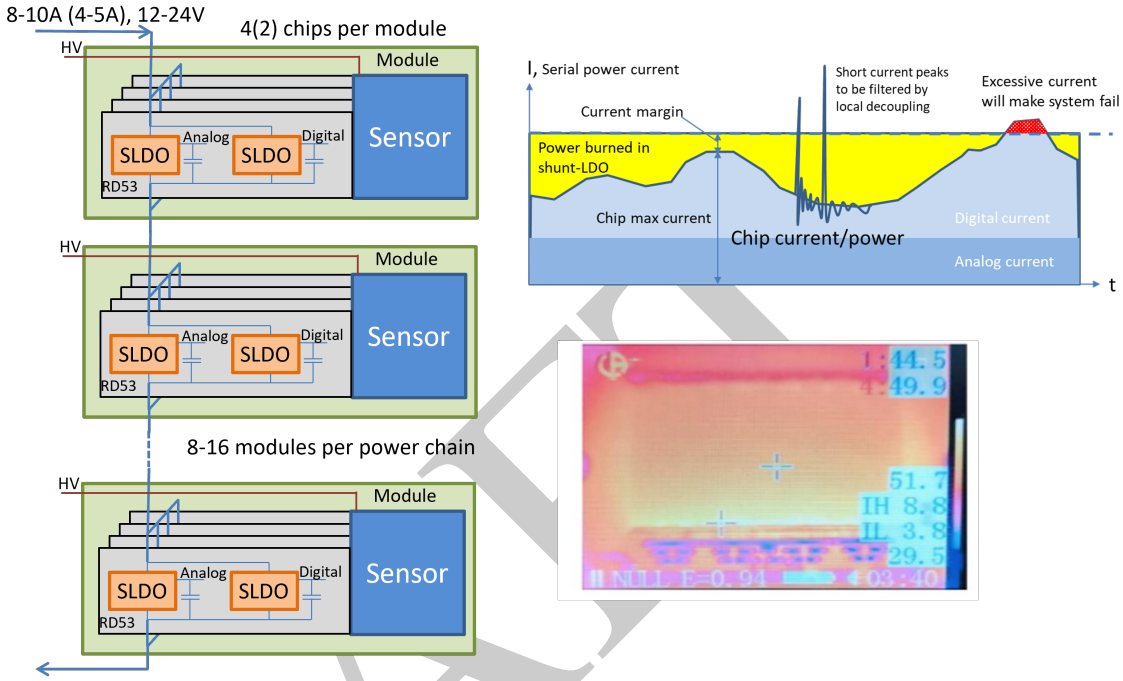


Figure 6. Serial powering of pixel modules with multiple chips in parallel per module. Issues with varying load current (upper right) and SLDO hot spot (lower right)

grounds can not be connected to system ground, which implies the need for good galvanic isolation between pixel chips and local CO_2 cooling. HV biasing of the pixel sensors (Planar: 100-1000 V, 3D: 20-100 V) will have voltage differences between modules along the serial powering chain of up to 12-24 V, depending on the number of modules in the serial power chain.

The novel serial powering scheme developed for the RD53 chips also enables multiple chips on a pixel module (up to 4) to be powered in parallel. This enables all pixel chips on the same module, and bump-bonded to a common pixel sensor, to be at the same potential. It also improves significantly the overall reliability of serial powering, as individual chips can be allowed to fail with a power-open, without affecting other chips and modules on the serial power chain (e.g. 16 quad pixel modules with a total of 64 pixel chips in single serial power chain). To assure appropriate power by-passing, in case of a failing chip on a module, the shunt capability of the on-chip regulators can handle up to 200 % of nominal chip current. As an additional protection capability, chips have a built-in 2 V over-voltage clamp to assure that local voltage fluctuations do not cause damage during power cycling or during local power anomalies. Extensive chip, module and system tests of serial powering have been made to qualify such a novel powering scheme as shown in figure 7 with appropriate cooling infrastructure.

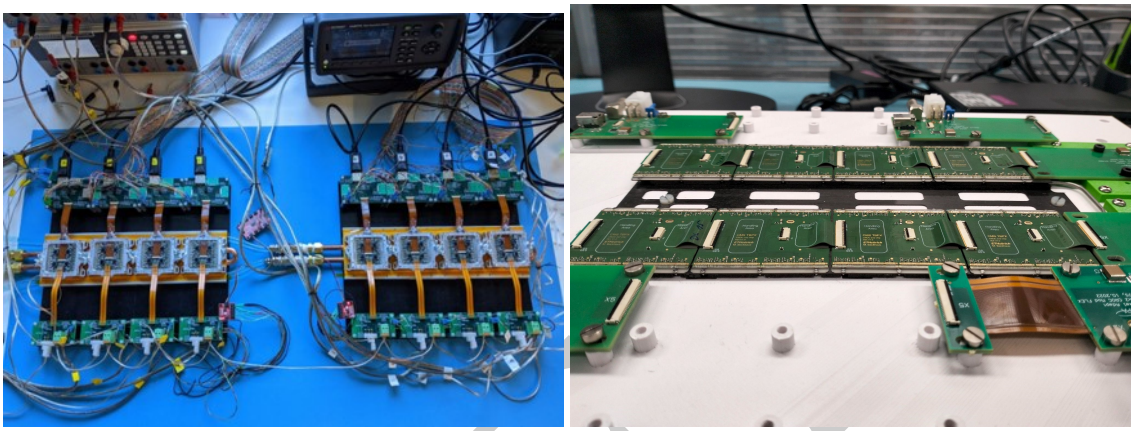


Figure 7. System tests with multiple quad RD53B chip pixel modules with serial powering and cooling. Left: ATLAS quad chip pixel module testing with serial power being connected to pixel modules with Kapton flexes to a power distribution flex [32][33]. Right: CMS quad chip pixel module testing with serial power, and power return, passing from one module to the next for the barrel part of the CMS pixel detector.

3 Chip architecture

RD53 pixel chips capture pixel hits across the 2×2 cm^2 pixel array with appropriate timing to enable triggering of hits in individual bunch crossings for readout. Charge deposited in the pixel sensor, bump bonded to the pixel chip, is amplified and shaped so it can be sampled precisely in the correct bunch crossing, with associated 4 bit charge information. Sampled and zero-suppressed hit information is stored during the trigger latency in latency buffers distributed across the pixel array in local pixel regions consisting of 4 pixel cells. The transfer of triggered and extracted pixel hit data from the pixel regions is organized into 8×8 pixel cores (consisting of 2×8 pixel regions) via columns of pixel core buses to the Digital Chip Bottom (DCB). Triggered hit data are then assembled and processed before being queued in derandomizer FIFOs for event readout over serial links in their original trigger order.

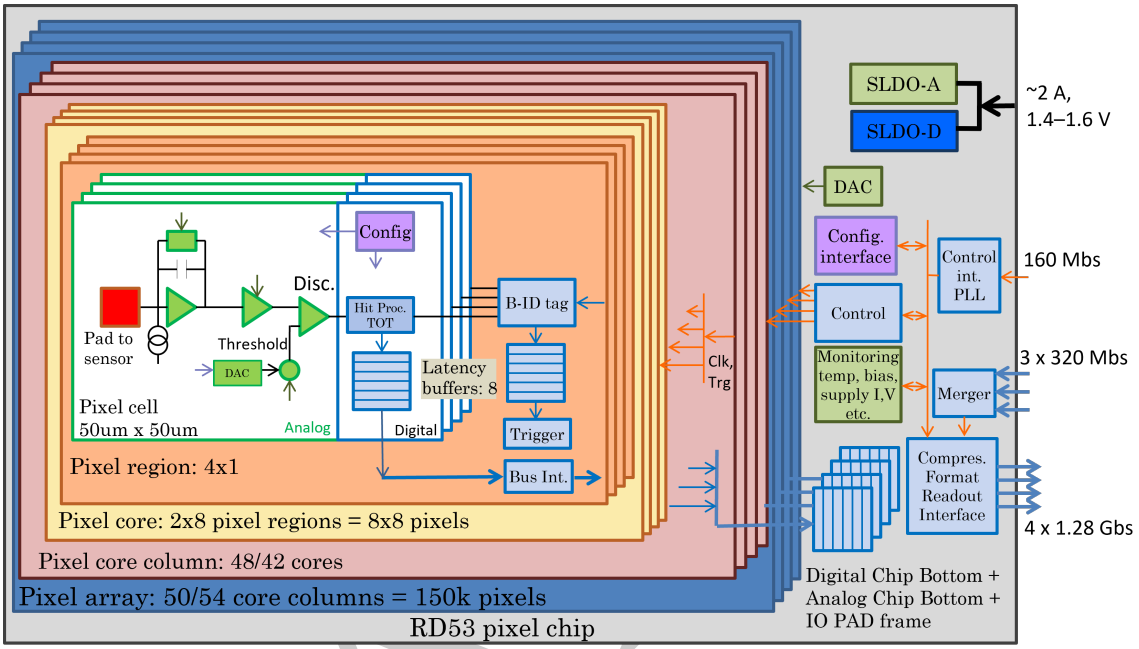


Figure 8. RD53 Data flow architecture.

The RD53 architecture is organized in a hierarchical fashion of which some parts are related to the logical data flow, as shown in figure 8, while others are related to the implementation floorplan of the chip as shown in figure 9 :

- **Pixel**: Experiment specific AFE with threshold adjust, charge injection and configuration.
- **Pixel region**: 4×1 pixel digitization with TOT and 8 latency buffer locations.
- **Pixel core**: 2×8 pixel regions, 8×8 pixels.
- **Pixel core column**: 48 cores in ATLAS chip, 42 cores in CMS chip.
- **Pixel core row**: 50 cores in ATLAS chip, 54 cores in CMS chip.
- **Pixel array**: ATLAS: 8×8×50×48 = 153.600 pixels, CMS: 8×8×54×42 = 145.152 pixels.
- **Analog pixel island**: 2×2=4 pixel AFEs.
- **Analog biasing column**: 2 pixels wide columns for AFE biasing with drivers.
- **Digital Chip Bottom (DCB)**: Digital logic outside the pixel array.
- **Analog Chip Bottom (ACB)**: Peripheral analog circuitry: Bandgaps, Biasing DACs, Monitoring ADC, PLL.
- **Pad frame**: IO with wire-bond pads, ESD protection, distributed SLDO power regulators and power-on reset.

It is important to notice that the effective organization of pixels into pixel regions, with shared latency buffer, depends on the organization and routing on the pixel sensor. For square 50×50 μm^2 pixels the sensor array and the chip have the same array structure and the pixel regions are seen as 4×1 pixels covering 200×50 μm^2. Such an elongated pixel region shape is advantageous for latency buffer sharing at the end of the detector barrel, where hit clusters are elongated because of particle

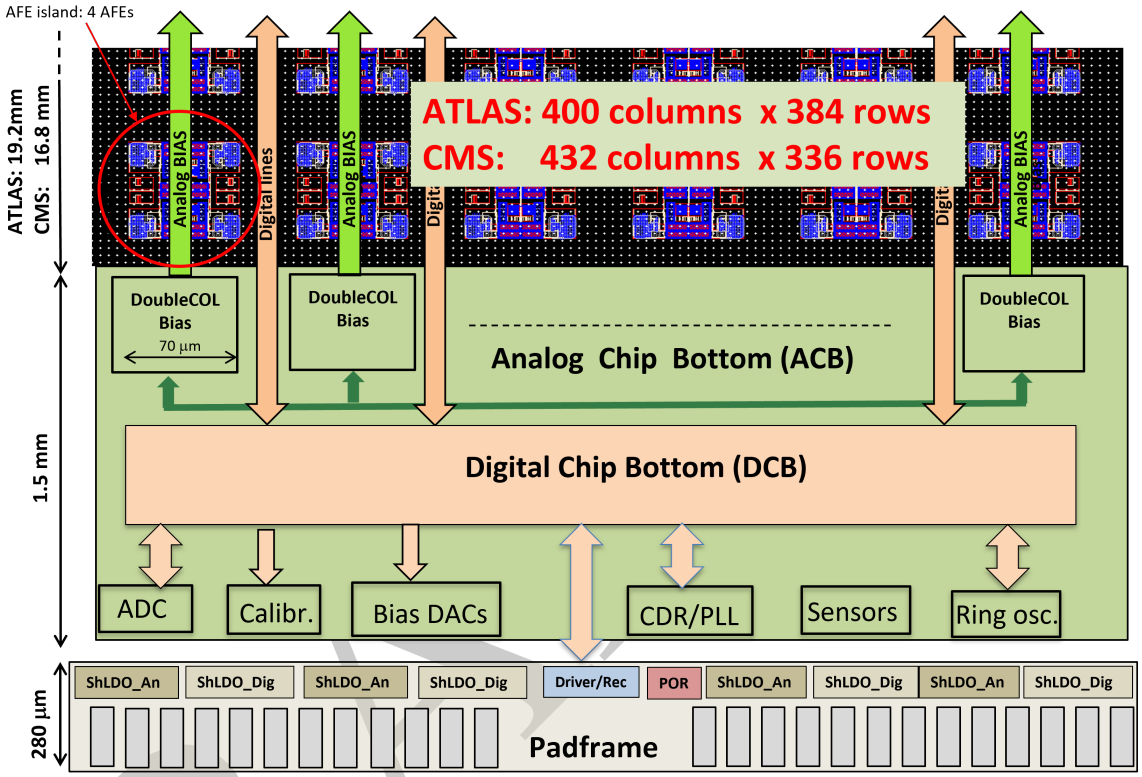


Figure 9. Physical implementation floorplan with analog pixel islands surrounded by digital pixel logic and digital/analog chip bottom.

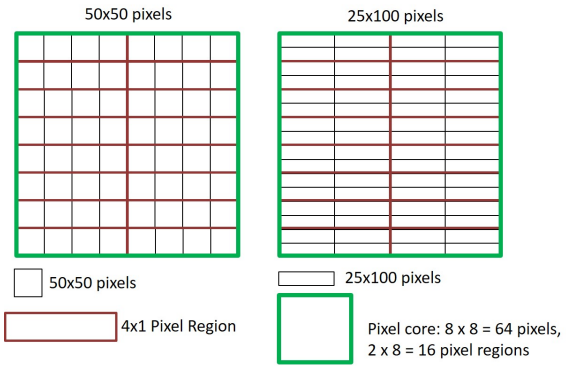


Figure 10. Pixel region organization with 50×50 μm^2 and 25×100 μm^2 pixels.

angle and active sensor thickness. When having elongated pixels of 25×100 μm^2 on the sensor, the pixel region is effectively seen as a 2×2 pixel region covering the same 200×50 μm^2 area as indicated in figure 10.

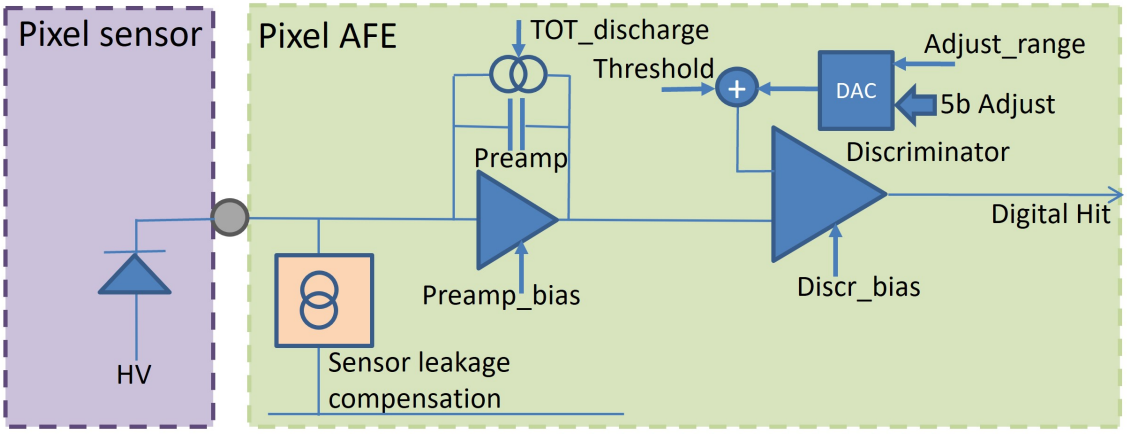


Figure 11. Block diagram of generic pixel Analog Front-End (AFE).

4 Analog front-ends and hit digitization

The Analog Front-Ends (AFE) are grouped in 4 pixels and implemented as small analog islands in a sea of digital pixel array logic, as illustrated in figure 9. The grouping of 4 AFEs into analog islands, where the AFE layout has been mirrored and flipped to fit together, can cause minor systematic mismatch differences, that after appropriate threshold tuning become negligible. The analog islands are isolated from the surrounding noisy pixel logic using a deep N-Well (triple well) available in the chosen 65 nm technology. AFEs also use a separate analog power supply (together with Analog Chip Bottom) from a dedicated analog SLDO regulator.

Charge induced by traversing particles in the pixel sensor is transferred to the pixel chip via fine pitch bumps as shown in figure 11. The AFE in each pixel integrates the collected charge in a pre-amplifier stage, with charge integrating feedback, followed by appropriate signal shaping and buffering. The discharge of collected integrated charge is done with a configurable discharge current (in some implementations referred to as Krummenacher current [14]), resulting in an analog pulse width proportional to charge. The shaped signal, with fast rising edge and slowly decreasing falling edge, as illustrated in figure 12, is transformed into a 1 bit digital hit signal by a discriminator with programmable threshold.

The AFEs work asynchronous to the 40 MHz bunch crossing clock and hits are synchronized to the 40 MHz chip clock at the entry to the digital pixel hit processing in the pixel regions. The effective threshold per pixel is determined by a global threshold bias together with a 5 bit threshold adjust per pixel, to compensate for threshold dispersion among pixels. The range of the threshold adjustment is defined by a global programmable bias to allow its dynamic range and resolution to be optimized for the observed pixel array threshold dispersion.

The simple small area and low power hit detection with TOT charge measurement has some side effects as indicated in figure 12. The prolongation of the analog hit signal, to get an appropriate TOT charge resolution, gives an analog dead-time hit loss of the order of: (Average TOT time) $\times$ (Hit rate per pixel) = (from 50 ns to 200 ns) $\times$ 75 kHz/pixel = 0.4-1.5 % for the highest hit rate

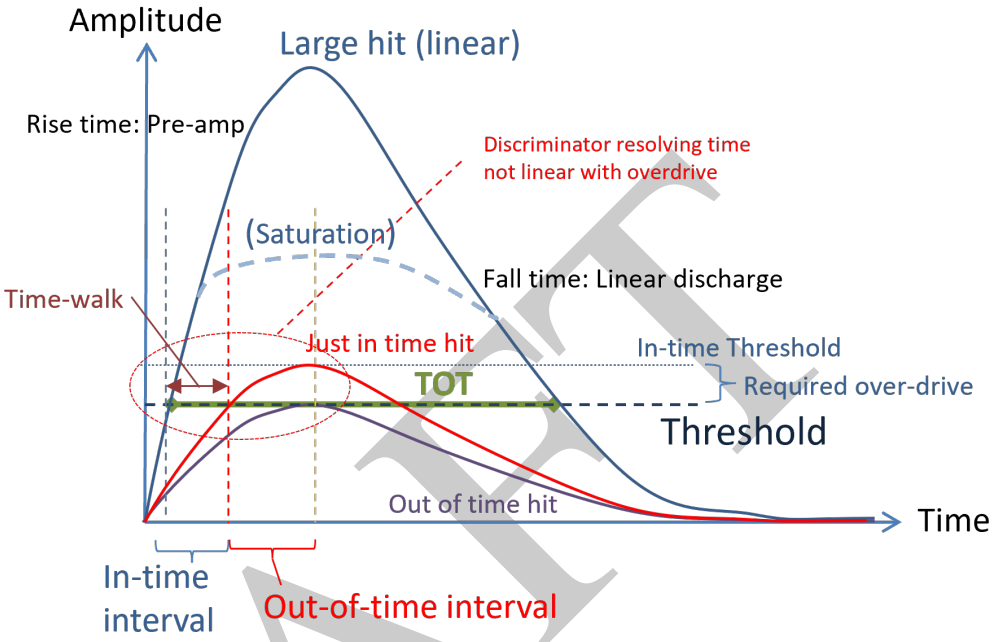


Figure 12. Charge measurement with TOT and indication of related time-walk, dead-time and amplitude saturation effects.

in inner layers. The maximum TOT dead-time loss in inner layers have been specified to be 1 %, enforcing a short average TOT time below 133 ns. In outer low rate layers a longer TOT time can be used to get better charge resolution.

It can be mentioned that if the charge integrated analog signal has amplitude saturation effects, the TOT charge measurement can be corrected off-line, when used at low thresholds as indicated in figure 12.

Small hit signals, just above threshold (e.g. at edge of pixel cluster), will get time walk from the combined effect of analog signal shape and the discriminator reacting slower when having small signal over-drive. For a time-walk below 20 ns (25 ns clock period minus 5 ns sampling margin for system jitter and time alignment to collisions) the hit will be detected in the appropriate bunch crossing for triggered readout. For time walk larger than this, the hit can be seen as a low TOT/charge "noise" hit in the following bunch crossing and will therefore not be read out for the corresponding trigger (unless forcing double triggering).

Each experiment has chosen their specific low power AFE implementation based on their specific emphasis on particular performance characteristics. The differential AFE used in the ATLAS chip has particular emphasis on low noise and small time-walk. The linear AFE used in the CMS chip has particular emphasis on linearity (~ 5%) and capability to work with short analog/TOT dead-time in inner high rate pixel layers [16]. The two AFEs have similar effective area and interfaces and are integrated into the RD53 design framework with a few design configuration parameters.

Multiple biasing levels for the AFEs are defined by global configuration registers connected to

biasing DACs driving the analog pixel array via column drivers as shown in figure 13:

- **Pre-amplifier bias**: Determines effective speed of AFE charge integration, but also affects effective gain, noise and dispersion. Major contributor to AFE power consumption.
- **TOT discharge current**: Determines discharge rate of integrated charge, thereby defining TOT resolution and related analog dead-time.
- **Discriminator bias**: Determines effective speed (and time-walk) of discriminator.
- **Global threshold**: Global chip threshold onto which local pixel threshold adjustment is applied.
- **Threshold adjust range**: Determines range of local threshold adjust DACs.

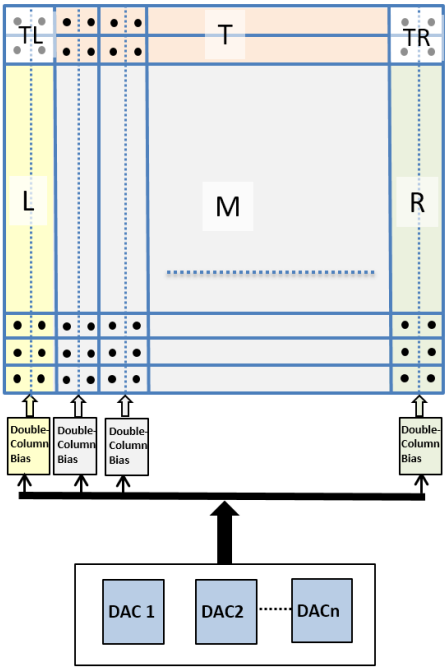


Figure 13. AFE Bias distribution with different Pre-amp biasing of different pixel regions: Main (M), Left side (L), Top Left side (TL) , Top (T), Right side (R), Top Right (TR).

Biasing of the AFEs has significant effects on their behavior and is a delicate optimization to be done according to the allowed power consumption, hit rate, noise, time walk and pixel sensor characteristics. It will also be needed to take into account accumulated radiation effects in both pixel sensor and the pixel chip itself. Edge and corner pixels have separate pre-amplifier biasing as they are typically 2-4 times larger. This is organized in 6 groups: Main, Left side, Right side, Top, Top Left corner, Top Right corner, as indicated in figure 13, to enable flexible adaptation to different pixel sensor and module configurations with enlarge pixels in boundary regions between pixel chips.

The leading edge of the discriminator hit signal is a measure of particle Time Of Arrival (TOA), with associated time walk, and Time Over Threshold (TOT) is proportional to collected charge. The leading edge is synchronized to the 40 MHz sampling clock and pulse width is measured with rising

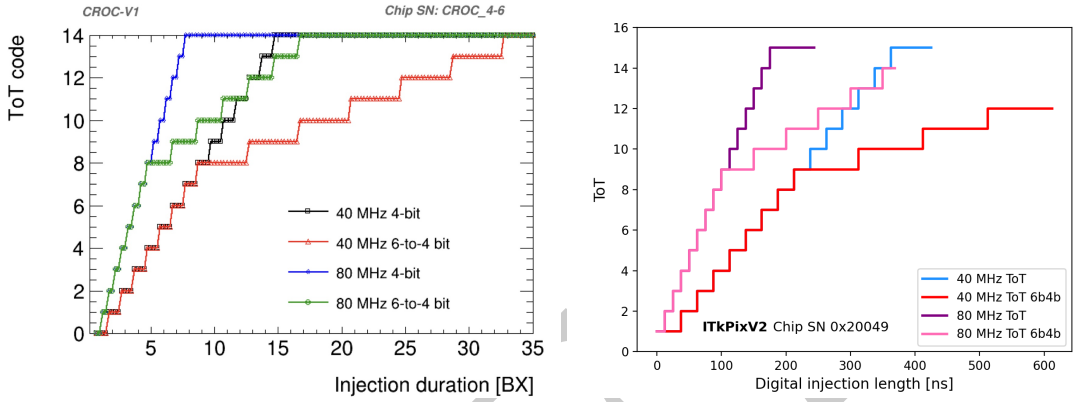


Figure 14. Measured Linear and Dual slope TOT with 40 MHz and 80 MHz sampling using digital pulse width injection. Left: CMS chip with pulse width in BX units (25 ns) and TOT=15 representing no hit (so not shown). Right: ATLAS chip with pulse width injection in ns and TOT=0 representing no hit. Does not including analog non-linearity of the AFE (described later).

or both edges of the 25 ns sampling clock (40 and 80 MHz TOT sampling) for a 6 bit TOT count. The 6 bit TOT count can be mapped directly into 4 bit TOT, ignoring 2 MSB (Most Significant Bits) bits with saturation, or can be mapped into 4 bits with a dual slope mapping. With dual slope mapping the full TOT resolution is maintained in the first half of the 4 bit dynamic range, whereas extended dynamic range is obtained in the second half as illustrated in figure 14. Dual slope mapping assures good position interpolation at the edge of pixel clusters, where the collected charge is normally small, and high dynamic range for dE/dx measurements of pixel clusters that can be used to identify highly ionizing particles and contribute to general particle identification [15].

Capture of the discriminated hit signal can be performed synchronously or asynchronously. For synchronous/simple sampling, short hits not present at the rising edge of the sampling clock will be ignored. For asynchronous/latched sampling, a discriminator hit pulse is kept high until it has been captured by the first coming 40 MHz sampling clock edge. Async sampling is guaranteed to capture short hit pulses, but will have slightly higher sensitivity to noise hits as indicated in figure 15. One should also be aware that short/small hits will typically also be affected by time-walk. In RD53B chips the hit sampling mode was configurable. For the final RD53C chips, ATLAS has chosen the async sampling mode, given the low noise and low time-walk AFE and issues fitting both modes in available area. For the RD53C-CMS chip it has been possible to maintain both sampling modes.

The 40 MHz sampling clock for the hit detection in the individual pixels is carefully distributed across the pixel array with typical (maximum) time skew across the whole array below 1 ns (2 ns) to assure that hits are captured in the correct bunch crossing for triggering and readout. This gives short digital power surges over the pixel array at the rising edge of the clock that can affect the analog front-ends and hit digitization. Local decoupling capacitors are distributed across the array to minimize this effect and the AFEs have been decoupled as much a possible from the digital with separate power domains and the use of separate triple wells for analog and digital. A threshold variation over the 40 MHz clock period has been observed in the full RD53 chips as indicated

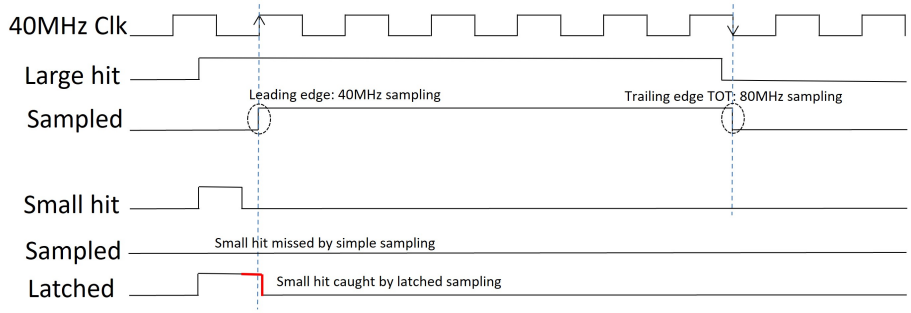


Figure 15. Hit sampling with synch/asynch sampling. Indication of optional 80 MHz TOT sampling.

in figure 16 and figure 17. This can be seen to be related to on-chip power distribution with four distributed sub-instances for each SLDO. The magnitude of this effect is dependent on the sampling mode used (sync or async) and on the configured TOT discharge rate (fast/slow). The Async sampling mode, assured to capture incoming hits in the full clock cycle is intrinsically less sensitive to the arrival time of hits. At a 1000 e threshold the async sampling will have a 50-100 e threshold variation effect. This has been measured to be proportional to the number of pixel core columns actively being clocked and will therefore also in practice depend on power decoupling and wire-bonding (resistance and inductance) on a pixel module. In sync sampling mode the effective hit capture threshold has a relatively large dependency on the relative phase to the sampling clock and TOT discharge rate as shown in figure 17, because short hits can be missed depending on its relative timing/phase to the sampling clock. The measured threshold variation across the pixel array, for the async sampling mode, is caused by static and dynamic voltage drops in the on-chip power distribution network within the pixel array and four distributed instances of the SLDOs. It has not been possible to improve this further for a one-side powered chip, needed for quad chip pixel modules with no dead detection zones between chips. It should be noted that direct particles from the bunch collisions of the LHC arrive within a relatively narrow time window (~1ns) and will therefore not be significantly affected by this. It is for each experiment to determine how they want to time align their pixel detector to the bunch collisions and how to obtain appropriate absolute threshold and charge measurement calibration.

Discriminated hit signals from individual pixels, with local enables, are also connected to a configurable hit-OR column network to measure Time Of Arrival (TOA) and TOT with 640 MHz precision TDCs (Time to Digital Converter) in the Digital Chip Bottom (DCB). The hit-OR signals are also used for a flexible self-trigger function that can be used for detailed chip and sensor characterization in test beams and with radioactive sources. An example of using the precision TOA is shown in figure 18 to measure calibration injection and hit-OR skew along pixel columns in the RD53B-ATLAS chip (improved in RD53C chips to have reduced skew).

An analog injection circuit is implemented in each pixel, with an equivalent circuit as shown in figure 19, to perform precise threshold tuning and calibration. The calibration injection circuit uses two distributed DC voltages (Vcal_Med and Vcal_Hi), from two on-chip 12 bits DACs, followed by in-pixel switches to generate charge injections via a pixel injection capacitor (Cinj). Having

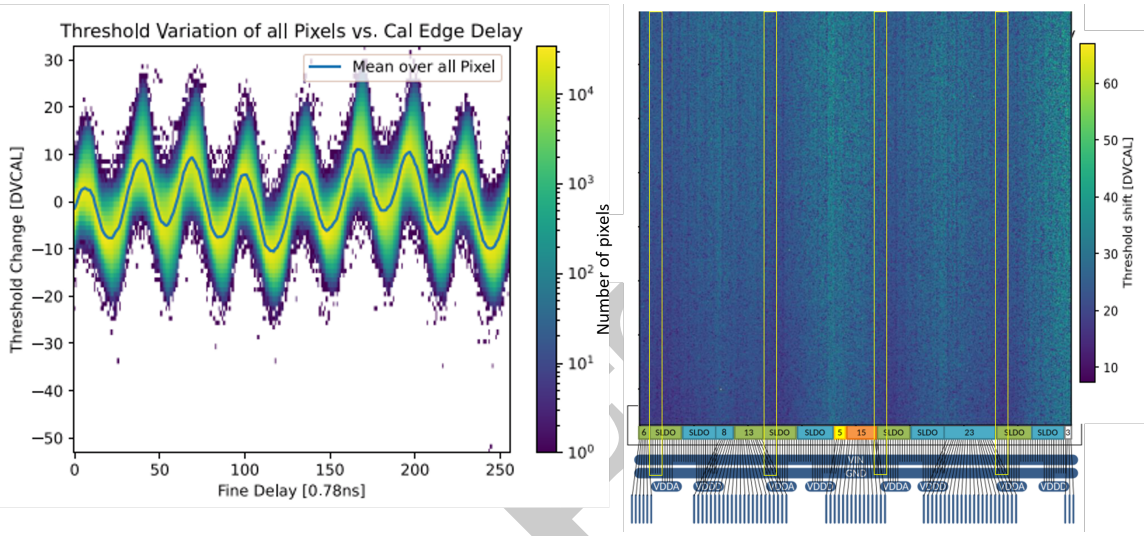


Figure 16. Left: RD53B-ATLAS Threshold variation (async sampling) across clock cycle from power supply perturbations [56]. Right: Threshold variation amplitude across pixel array overlayed on RD53B layout with indication of distributed SLDOs in four groups. DVCAL is digital threshold configuration in steps of 5 e.

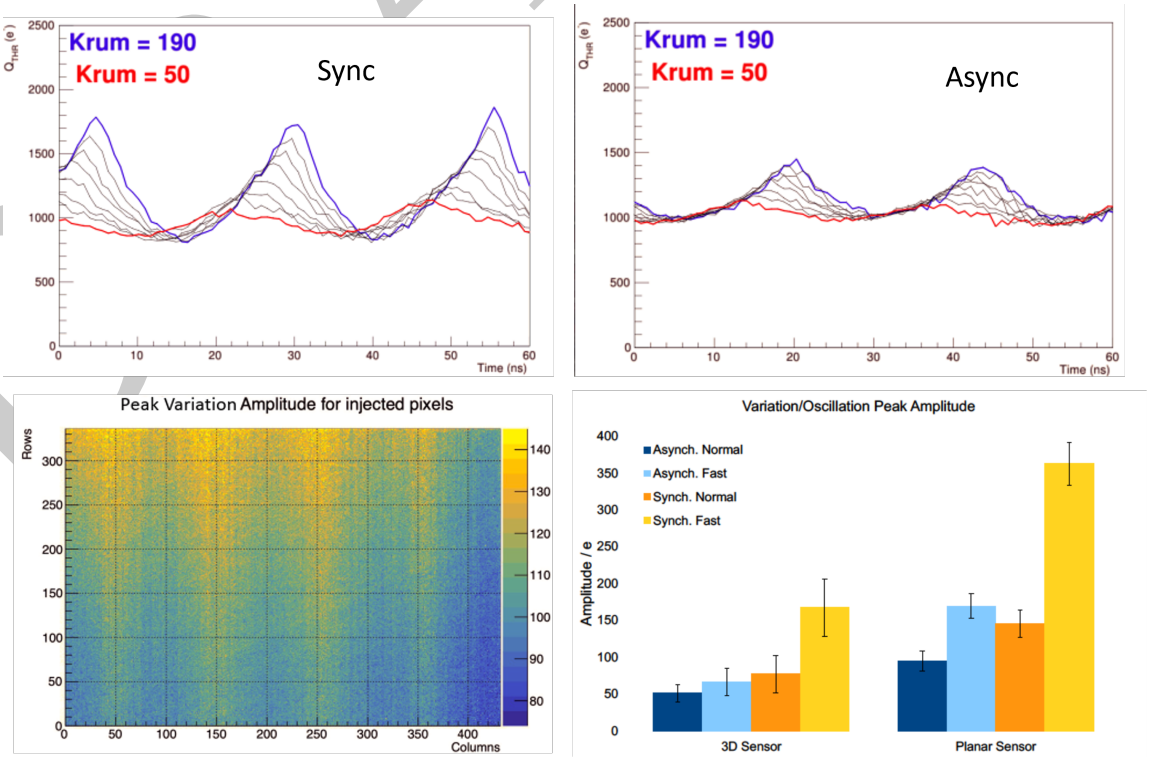


Figure 17. RD53B-CMS Threshold variation across clock cycle [17]. Upper left: Sync sampling mode for different TOT discharge rates from Fast (Krum=190) to Slow (Krum=50). Upper right: Async sampling mode. Lower left: 2D variation across pixel array in fast mode with 10 e color step. Lower right: Estimated max threshold variation effect for fast (inner) and normal (outer layers) discharge rate for 3D and planar pixel sensors.

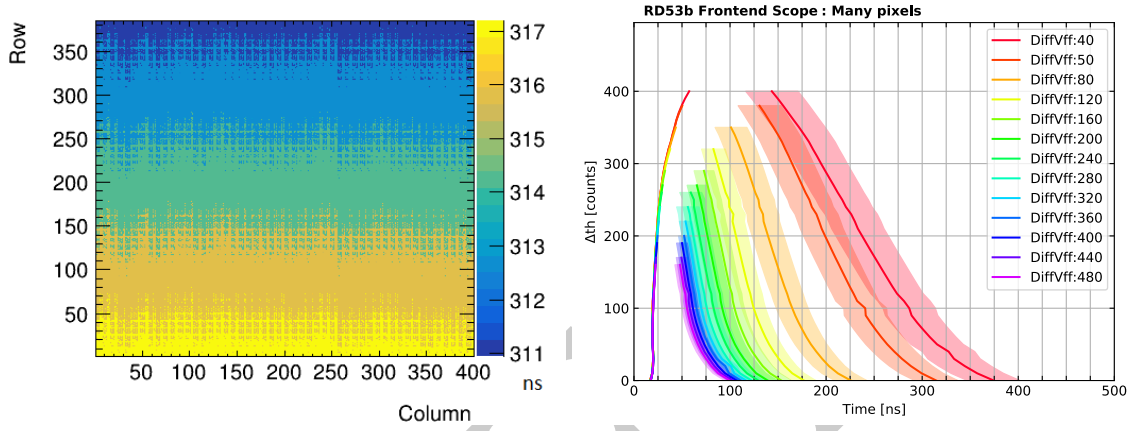


Figure 18. RD53B-ATLAS precision TOA and TOT. Left: Digital pixel injection across pixel array with measured TOA using the high resolution TDC via the hit-OR network. Right: Analog pulse shape reconstructed from measured TOA and TOT with high resolution TDC, over different charge injections. DiffVff sets charge integration discharge current of the Differential AFE and shaded area indicates spread among pixels in the pixel array.

two charge injection voltages enables precise differential charge injections (Vcal_Hi - Vcal_Med), independent of ground voltage drops across the pixel array, as well as making two consecutive injections (Vcal_Hi - Vcal_Med followed by Vcal_Med - ground) into the same pixel. The timing of the injection is controlled by a digital pulse generator with programmable injection time (0.78 ns resolution) and time between two consecutive injections, using the two calibration voltages plus local ground. The same pulse generator can be used for direct digital hit injections.

Two 12 bit voltage DACs located in the ACB generate the charge injection voltages, driven with dedicated voltage drivers to the in-pixel charge injection circuits in each pixel. The DAC characteristics are shown in figure 19. Calibration injection has an effective resolution of ~5 e when being used concurrently on a limited number (~100) of pixels. Used for massive concurrent injections in a large number of pixels, the effective precision is deteriorated by dynamic capacitive loading from the voltage switches in the pixels. Injection capacitance spread among chips on the same wafer has been seen to be 1 - 1.5 % with a 5-10 % difference between wafers.

A dedicated charge injection capacitor calibration circuit, shown in figure 20, is available in the ACB to make a precise injection capacitor measurement per chip, during wafer probing to enable calibrated charge injections in the final systems.

Shown AFE test results are in general for bare chips without bump-bonded pixel sensors. Bump bonded assemblies with sensors have only recently become available in sufficient quantity and quality to make detailed AFE characterization with these. No significant changes in pixel chip performance have been seen when tested with a bump bonded pixel sensor, except an AFE noise increase of 10-30 e as can be expected when having increased input capacitance [19].

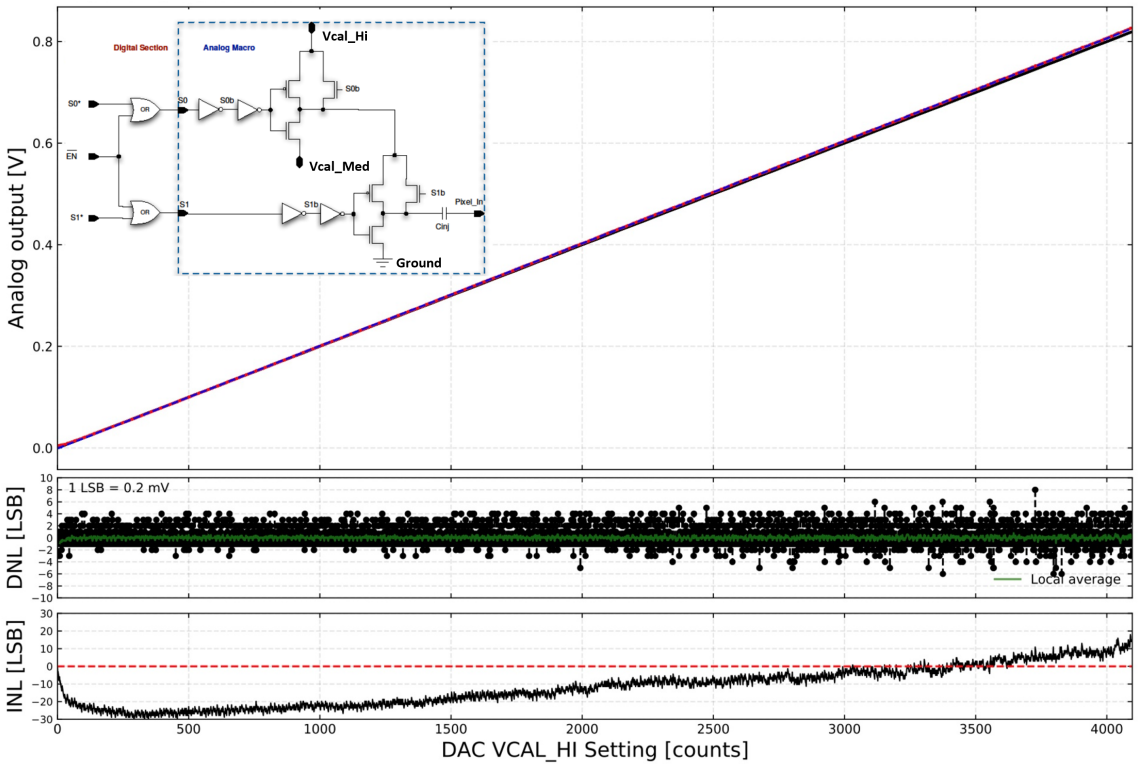


Figure 19. Calibration injection voltage as function of DAC setting with a linear fit and extracted DNL and INL. Pixel charge injection circuit shown as an insert with its two injection voltages (Vcal_Med, Vcal_Hi) and local ground, enabling two consecutive charge injections to be made.

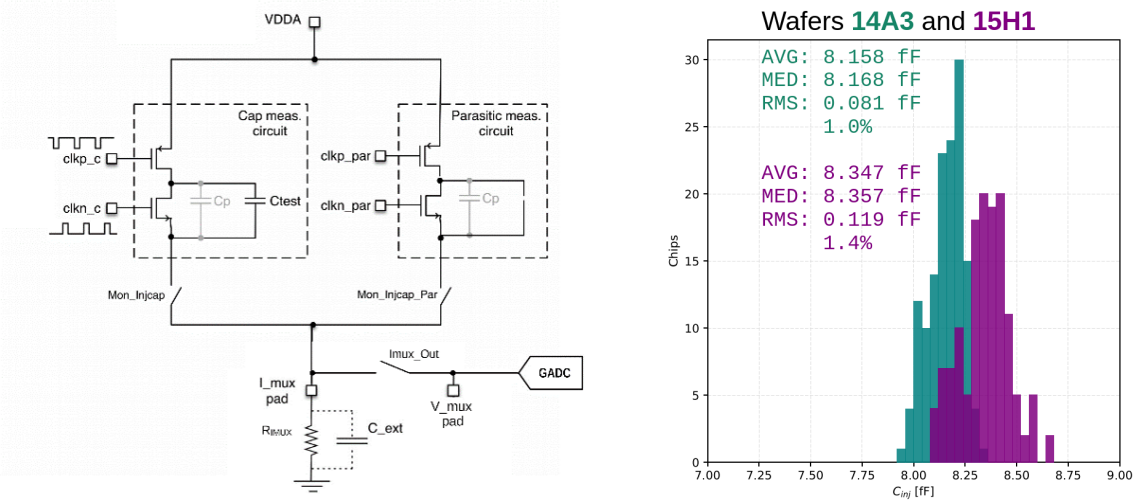


Figure 20. Left: Pixel Injection capacitor (Ctest) measurement circuit measuring the average current, with on-chip ADC (GADC) or external pin (V_Mux_pad), when charging and discharging injection capacitor at a constant injection rate. An equivalent branch, without injection capacitor, enables to measure parasitic capacitance (Cp) of the circuit. Right: Measured injection capacitor dispersion over two wafers from the same production lot.

4.1 CMS Linear front-end

The schematic of the linear analog front-end [18][19] adopted in the RD53B/C-CMS chip is shown in figure 21 with a Charge Sensitive Amplifier (CSA) with Krummenacher feedback [14] complying with the expected radiation induced detector leakage and providing a linear discharge of the feedback capacitor C_F. The choice of a single amplification stage is dictated by power consumption and area constraints with a charge sensitivity, set by C_F, of around 26 mV/ke$^-$. The signal from the CSA is fed to a low power comparator with a 5 bit, current-mode binary weighted DAC for local threshold tuning. The front-end has been optimized for a linear response for an input charge up to 30 ke and features an overall current consumption of 5 μA.

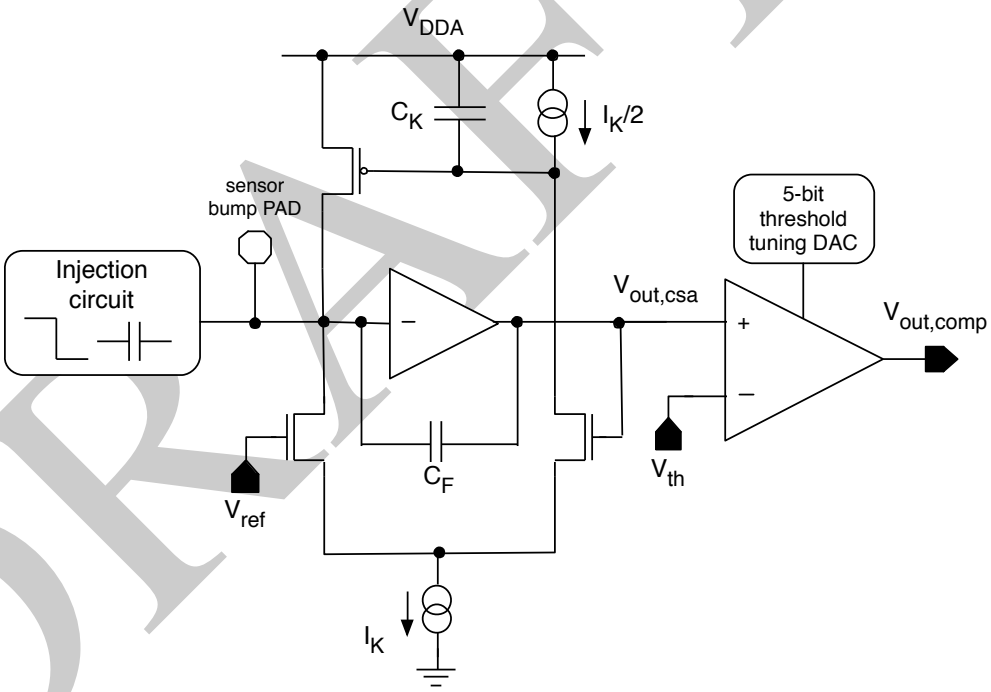


Figure 21. Schematic of CMS linear analog front-end.

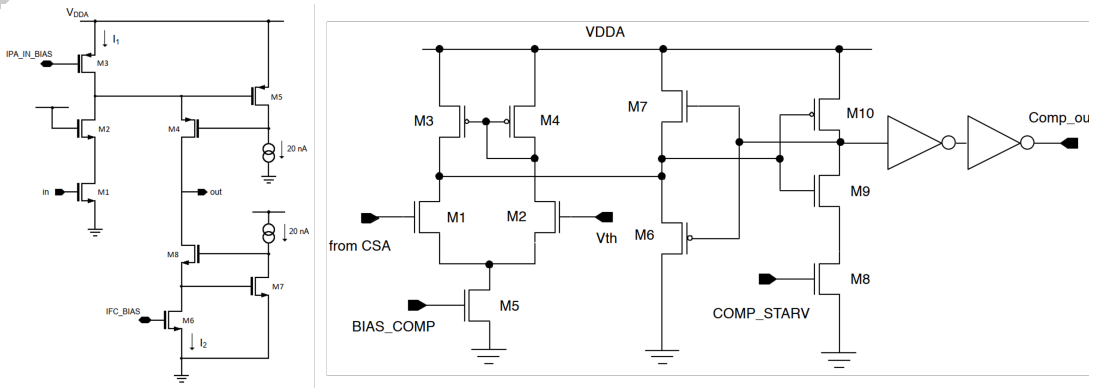


Figure 22. Transistor level implementation of linear AFE pre-amplifier (left) and comparator (right).

The charge sensitive amplifier, shown in figure 22 left, is based on a folded cascode input stage with two local feedback networks, composed of the M4-M5 and M7-M8 pairs, boosting the signal resistance at the output node. A 3 μA biasing current in the input branch and 200 nA in the cascode branch are responsible for most of the power consumption with a simulated open-loop DC gain of 76 dB with -3 dB cutoff frequency at 140 kHz with an effective closed loop peaking time of 22 ns. Noise is dominated by the input device and the PMOS transistor in the feedback. The comparator shown in figure 22 right, has a transconductance stage (M1-M5) followed by a Trans-Impedance Amplifier (TIA) (M6-M10) for fast switching, with an optimized feedback network (M6 and M7) for acceptable time-walk. Two inverters are used at the output to assure fast signal transitions to the digital pixel sampling logic. The layout and measured analog pulse shape are shown in figure 23.

A reconstructed AFE pulse shape from a combined scan of injection time and threshold with TOT is shown in figure 24 together with threshold linearity as function of global threshold setting. Figure 25 shows the time-walk measured as function of injected charge together with time-walk dependency on chip temperature with sync and async sampling and for fast (inner layers) and slow (outer layers) TOT discharge times. TOT linearity, with saturation, is shown in figure 26 together with its dependency on sampling clock phase for different charge injections and TOT spread across the pixel array. Figure 27 shows untuned (before local threshold trimming) threshold dispersion over the full pixel array together with a 2D map of appropriate pixel trimming to obtain tuned pixel threshold dispersion as shown in figure 28, when using the optimal trimming DAC range to cover the full dispersion range with the best possible resolution. Tuned pixel threshold dispersion at 1000 e before and after 1 Grad irradiation is shown in figure 29 with only a small degradation of threshold dispersion (after re-tuning at 1 Grad). Finally pixel noise distribution is shown in figure 30 at room temperature and cold with mean noise as function of temperature for fast and slow TOT discharge. No noticeable change of noise has been observed with irradiation up to 1 Grad.

No differences have been seen for the linear AFE in the RD53B-CMS and RD53C-CMS chips. In short it can be summarized that the linear AFE with a planar (or 3D) bump-bonded pixel sensor works fully satisfactory for the CMS pixel detector upgrade at a 1000 e threshold with ~50 e dispersion, mean noise below ~70 e (80-100 e with pixel sensor), time walk below 17 ns with a linear TOT charge measurement and radiation tolerance up to 1 Grad. At the time of writing, extended testing of the RD53C-CMS chip is ongoing in the CMS pixel detector project with different sensor types on pre-production modules in test beams and after irradiation.

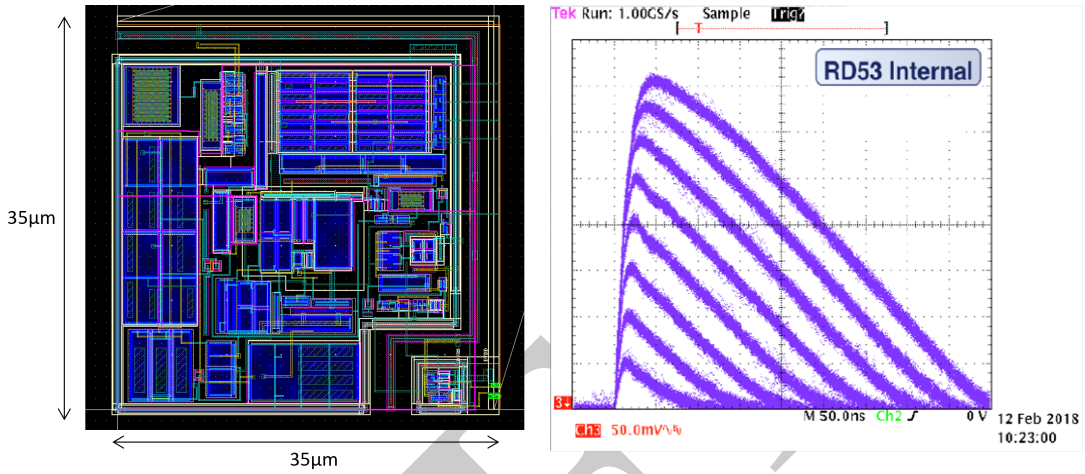


Figure 23. Left: Linear AFE layout. Right: typical analog waveform before comparator, with 1-10 ke charge injections, measured on analog output of the RD53A chip.

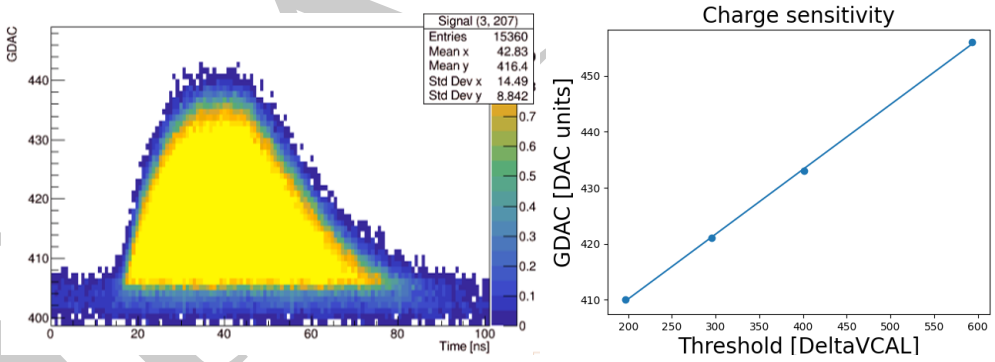


Figure 24. RD53B-CMS Linear AFE pulse shape and threshold linearity. Left: Reconstructed pulse shape from threshold and time scan combined with precision TDC TOT information, in fast mode with short TOT charge encoding. Right: Threshold as function of global threshold setting. GDAC: global threshold setting, DeltaVCAL: Injection voltage DAC setting (5 e per LSB).

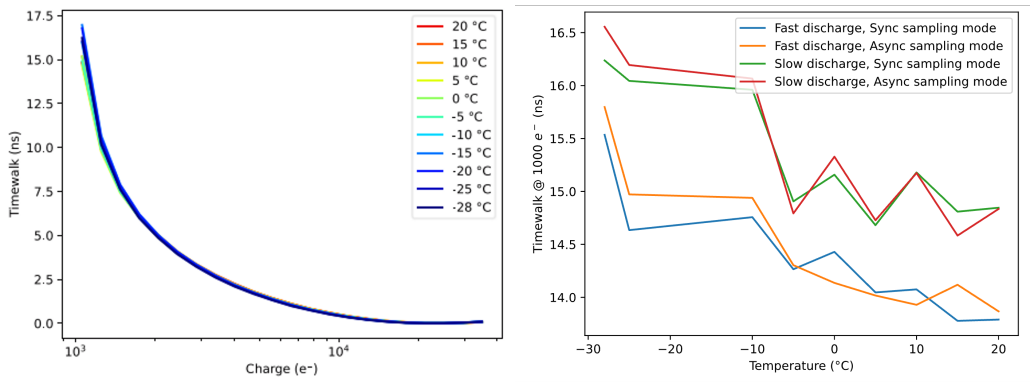


Figure 25. RD53B-CMS Linear AFE time-walk as function of injected charge (Left), at slow discharge and sync mode, and at different temperatures (right), at 1000 e threshold for different combinations of sync/async mode and fast/slow discharge rate.

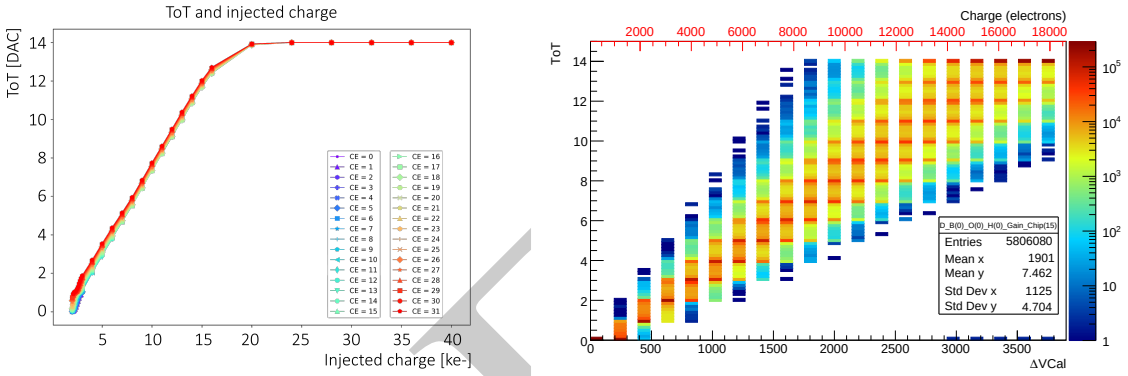


Figure 26. RD53C-CMS Linear AFE TOT linearity and spread. Left: Average TOT value as function of injected charge with linear encoding up to 15 ke, of interest for hit position interpolation between pixel hits in pixel cluster, for different relative clock phases (CE unit = 25 ns/32 = 0.78 ns). Right: Measured TOT linearity and spread across pixel array.

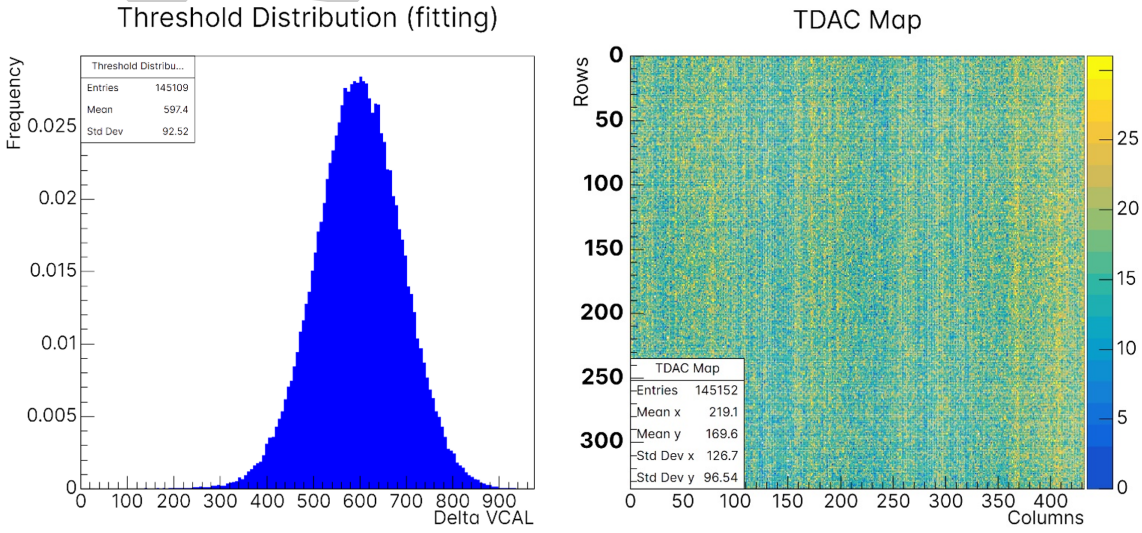


Figure 27. RD53B-CMS Linear AFE untuned threshold dispersion together with 2D trim DAC values to get uniform threshold (effectively shows untuned threshold map). Delta VCAL = 5 e. A column structure is clearly visible, coming from columns of pixel islands with their biasing drivers.

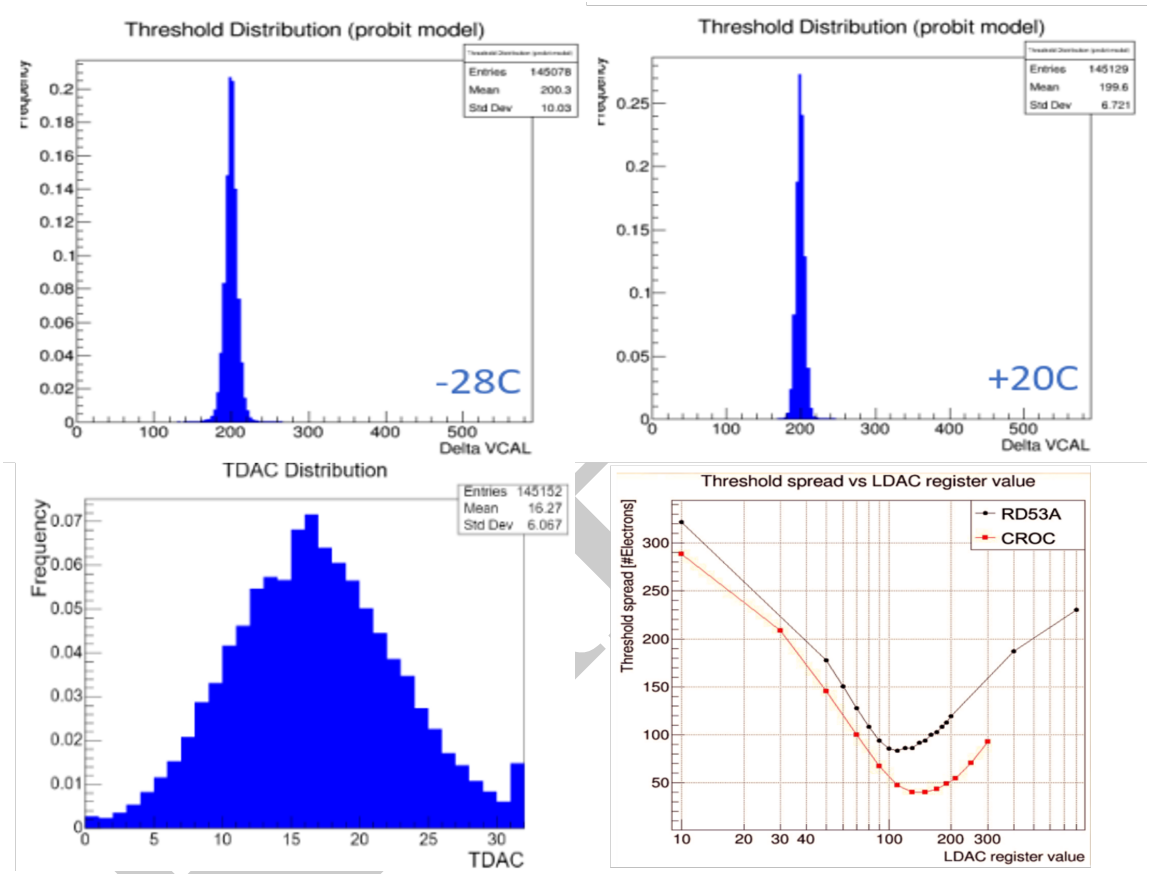


Figure 28. RD53B-CMS Linear AFE tuned threshold dispersion at 1000 e threshold at cold and room temperature together with trim-DAC tuning (TDAC) distribution and effective threshold spread as function as used trim-DAC range (LDAC). Delta VCAL = 5 e. CROC = RD53B-CMS.

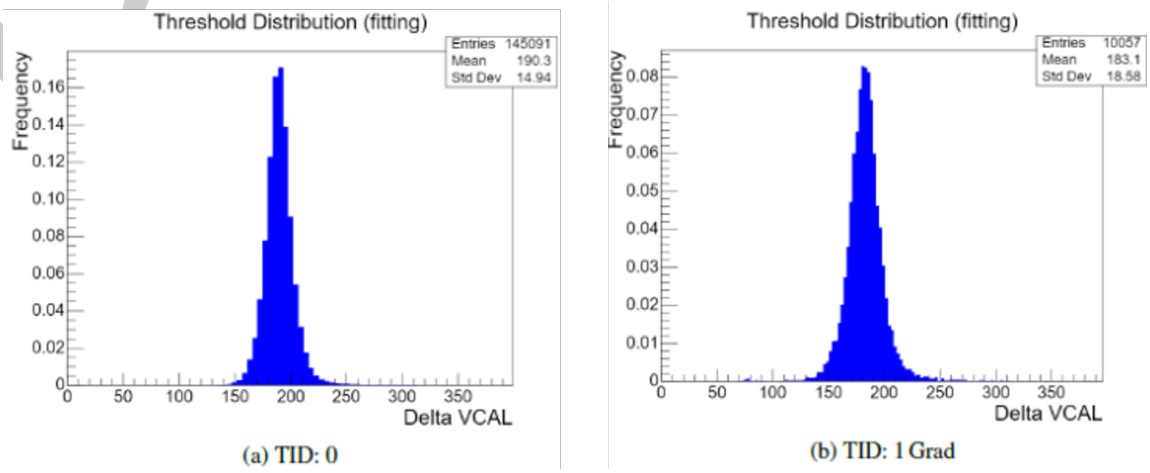


Figure 29. RD53B-CMS Linear AFE tuned threshold dispersion at 1000 e before and after irradiation to 1 Grad. Delta VCAL = 5 e.

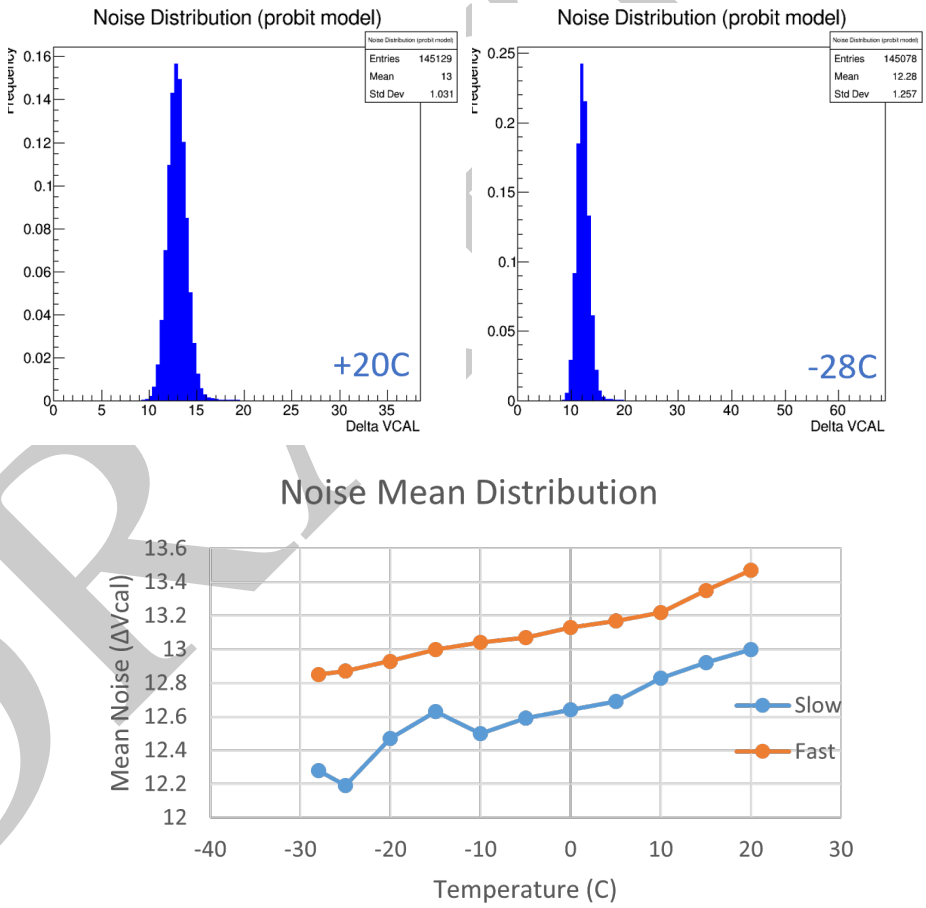


Figure 30. RD53B-CMS Linear AFE noise (without sensor) as function of temperature and for fast and slow TOT discharge. Delta VCAL = 5 e.

4.2 ATLAS Differential front-end

The differential AFE , as shown in figure 31, consists of 4 main stages: Pre-amplifier, Leakage Current Compensation (LCC), Pre-comparator, and Comparator. It is made from a single ended pre-amplifier followed by a pseudo differential comparator circuit.

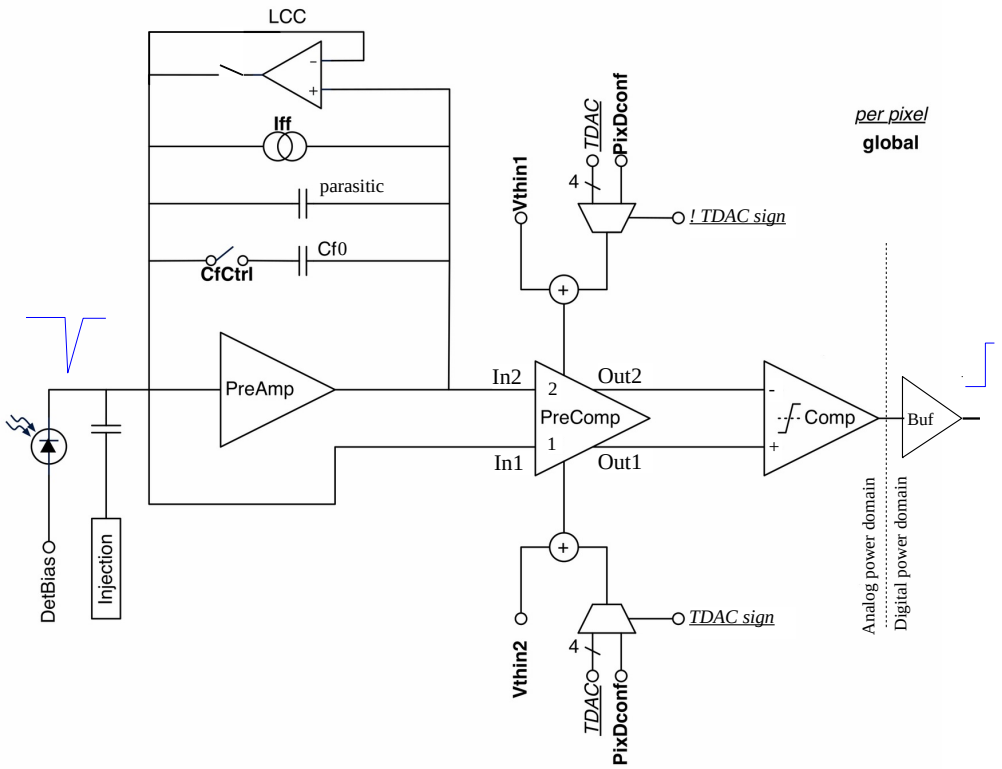


Figure 31. General schematic of ATLAS differential analog front-end. It can be noticed that the differential comparator has two 4 bit trim DACs, driven by the same 4 bit TADC values. The TDAC_sign bit determines which of the two differential branches is used for threshold trimming (effectively the 5th threshold adjust bit).

The topology of the pre-amplifier Charge Sensitive Amplifier (CSA) is based on a regulated cascode, with programmable constant current feedback (Iff) for the TOT charge measurement as shown in figure 32. The gate length and inversion region of the input device, as well as the implemented feedback circuit, gives the lowest possible noise at low power supply current. Due to its simplicity (small area), the pre-amp exhibits gain compression and non-linear discharge for large signals. This non-linear behavior can be considered advantageous for a pixel detector, giving good resolution at low input charge and compressed dynamic range for large signals, in a similar fashion as the optional digital 6 to 4 bit dual slope TOT encoding. The pre-amp gain is maximized using only the intrinsic input device and parasitic routing capacitance for charge integration (parasitic). A reduced gain programmable option is available, adding a feedback capacitor in parallel (Cf0).

Leakage Current Compensation (LCC) can be enabled with an optional auxiliary feedback path with a tunable low-pass pole to drain detector leakage. LCC is required for a pixel sensor leakage current above 2 nA. It eliminates DC operation point shifts of the pre-amp, that will otherwise reduce

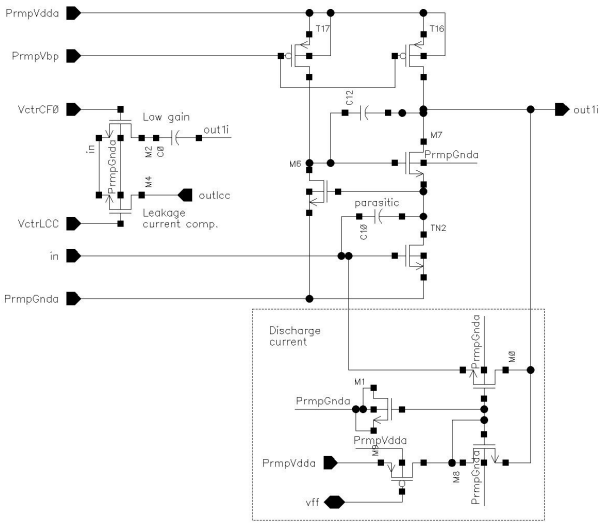


Figure 32. Transistor level schematic of single-ended analog pre-amp using parasitic capacitance for charge integration and its discharge circuit. Optional Leakage Current Compensation (LCC) and low gain feedback capacitance shown on left side.

effective dynamic range of the pre-comparator and pixel threshold in the presence of detector leakage. It also reduces leakage current induced noise by having a reduced feedback bandwidth.

The pre-comparator implements two essential features of the AFE. First, it utilizes the DC working point of the pre-amplifier input in concert with the pre-amplifier output (nominally at the same DC level) to form a differential thresholding circuit. Second, it includes a differential trim DAC to compensate for pixel-to-pixel threshold variation. Both the global and trimmed thresholds are set by source-followers, in each branch of the differential pre-comparator. This improves power supply noise rejection for internal and external power supply noise. A differential to single-ended comparator with a two stage open-loop class-A amplifier generates the discriminated hit signal, followed by an inverter for output buffering. Layout and measured analog pulse shape are shown in figure 33.

Measured threshold dependency on threshold settings is shown in figure 34 together with its temperature dependency. Figure 35 shows the time-walk of the differential AFE and figure 36 shows TOT linearity and dispersion together with TOT spread for a constant charge injection. There is no in-pixel tuning for pixel to pixel TOT dispersion, as this can be corrected offline if required. Un-tuned threshold dispersion is shown in figure 37, before and after irradiation. Tuned threshold dispersion is shown in figure 38 pre-irradiation and after 1 Grad. An indication of threshold de-tuning with radiation is shown in figure 39 showing only noticeable de-tuning during the first few Mrad of irradiation (typical of several CMOS technologies). Noise variation is shown in figure 40 with no significant noise increase after 1 Grad. Finally figure 41 shows the number and fraction of noisy pixels as function of tuned threshold, and can be seen to be very low for thresholds above 500 e.

In short it can be summarized that the differential AFE with a bump-bonded planar (or 3D) pixel sensor works fully satisfactory for the ATLAS pixel detector upgrade at a 1000 e threshold with ~50 e dispersion, noise of ~55 e (65-85 e with pixel sensor), time walk as low as ~15 ns, a compressed TOT charge measurement, and radiation hardness up to 1 Grad. At the time of writing, extended testing and qualification of the RD53C-ATLAS chip is ongoing in the ATLAS pixel detector project with final pixel sensors on pre-production modules in test beams and after irradiation.

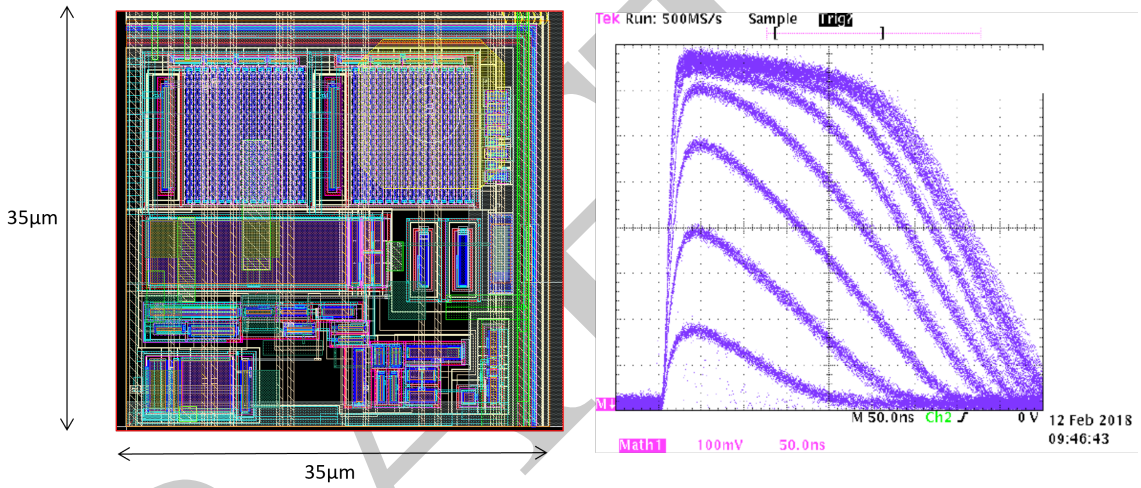


Figure 33. Left: Differential AFE layout. Right: typical analog waveform before comparator, with 1-10 ke charge injections, measured on analog output of the RD53A chip without pixel sensor. Saturation of the analog signal giving TOT compression at large charge injections can be noticed.

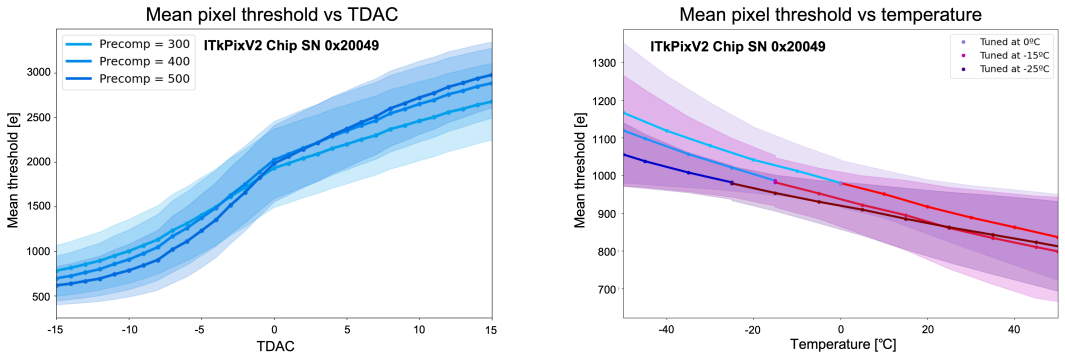


Figure 34. Left: RD53C-ATLAS Differential AFE threshold as function of threshold trimming (TDAC) for different Precomp biasing settings for the discriminator (affects power consumption, offsets, dispersion and time-walk. Default setting = 400). Right: Threshold temperature dependency for threshold tuned to 1000 e at different temperatures (0, -15, -25 °C) with nominal Precomp=400.

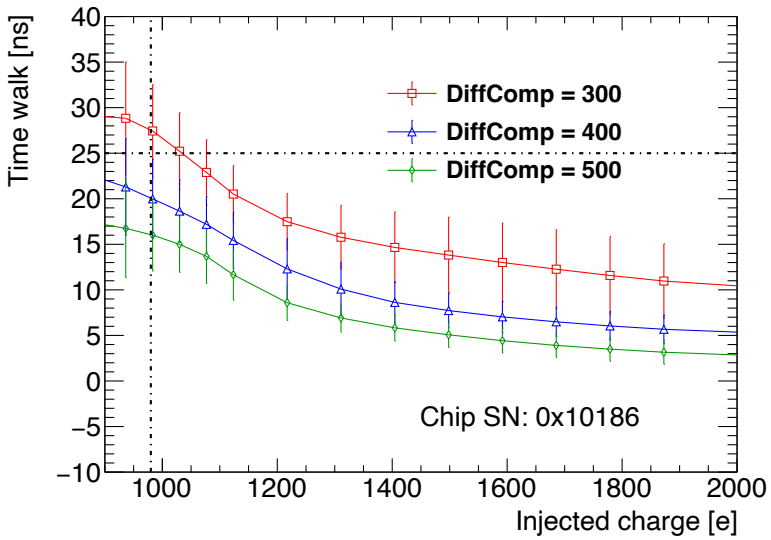


Figure 35. RD53C-ATLAS Differential AFE time-walk for different pre-comp biasing settings (DiffComp). To be noticed that for large charges the curves end up with different time offsets so effective time-walk change is difference between time for 1000 e injection and time for 2000 e injection (~15 ns for DiffComp=500 and ~20 ns for DiffComp=300).

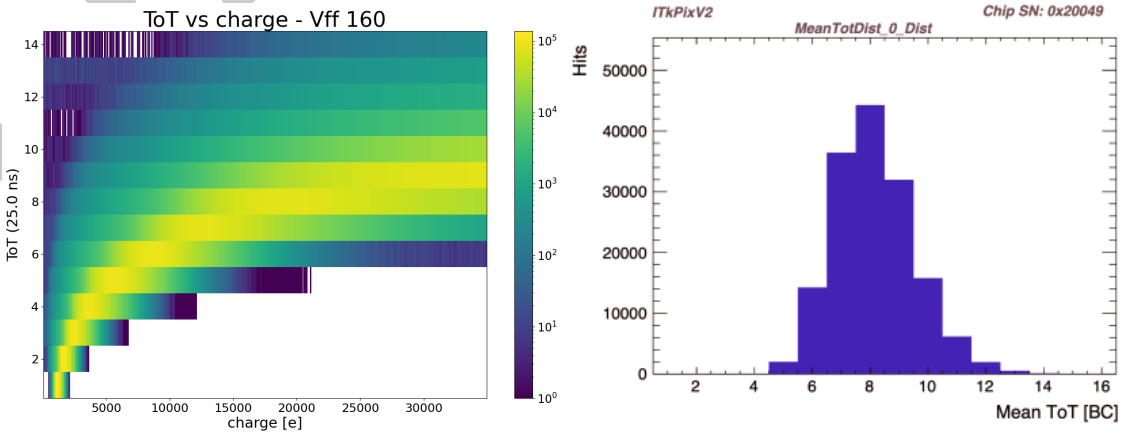


Figure 36. RD53C-ATLAS Differential AFE TOT linearity and spread across pixel array at 1000 e threshold. Left: Measured TOT as function as injected charge. Right: TOT spread across pixels for a constant 8 ke charge injection. TOT discharge does not have in-pixel tuning. Pixel TOT charge measurement non-linearity and variation between pixels can be compensated for off-line as needed.

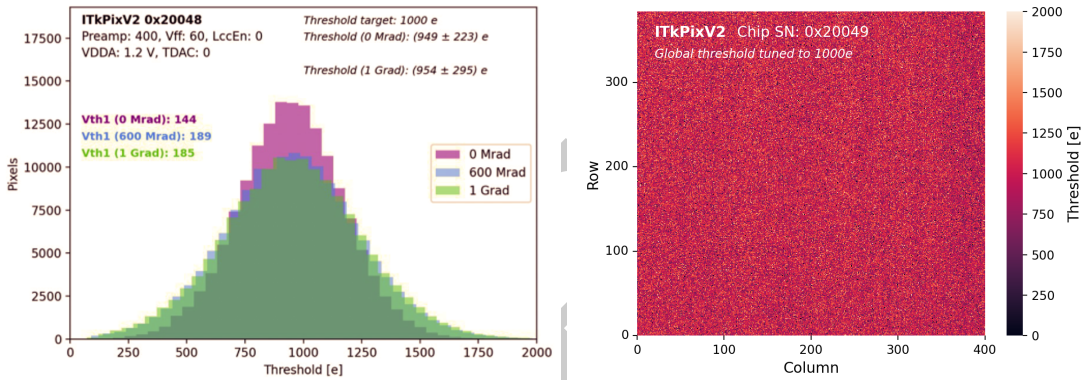


Figure 37. RD53C-ATLAS Differential AFE untuned threshold dispersion. Left: Before and after irradiation. Right: Un-irradiated dispersion across pixel array at 1000 e threshold.

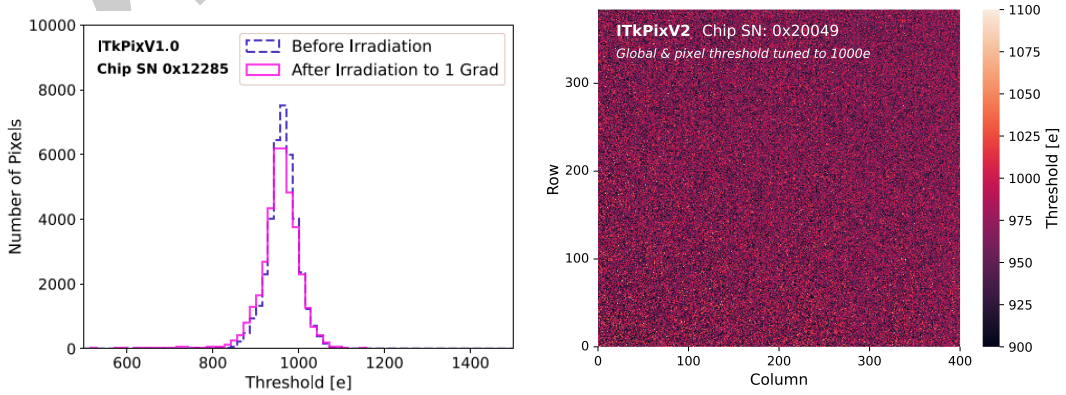


Figure 38. Differential AFE tuned threshold dispersion. Left: RD53B-ATLAS Dispersion before and after 1 Grad irradiation. Right: RD53C-ATLAS Un-irradiated dispersion across pixel array at 1000 e threshold.

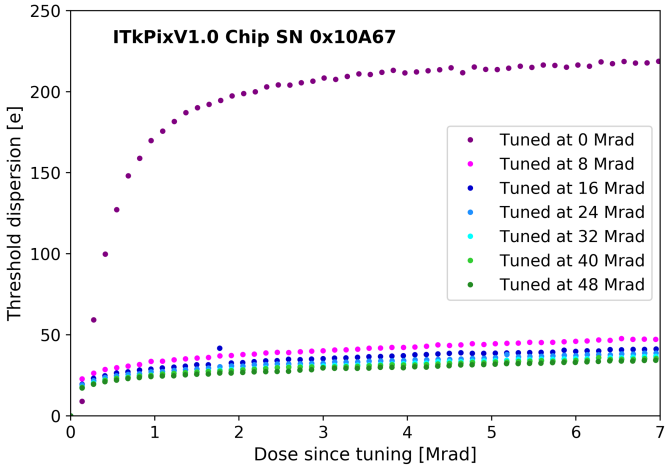


Figure 39. RD53B-ATLAS Differential AFE threshold de-tuning at different irradiation levels. The large transistors used in the AFE mainly have (minor) parameter changes during the first 0-2 Mrad irradiation (this has also been seen for other CMOS technologies). Thereafter radiation effects saturate and remain constant. Relatively frequent threshold tuning will therefore be required during initial running of the pixel detectors at high luminosity.

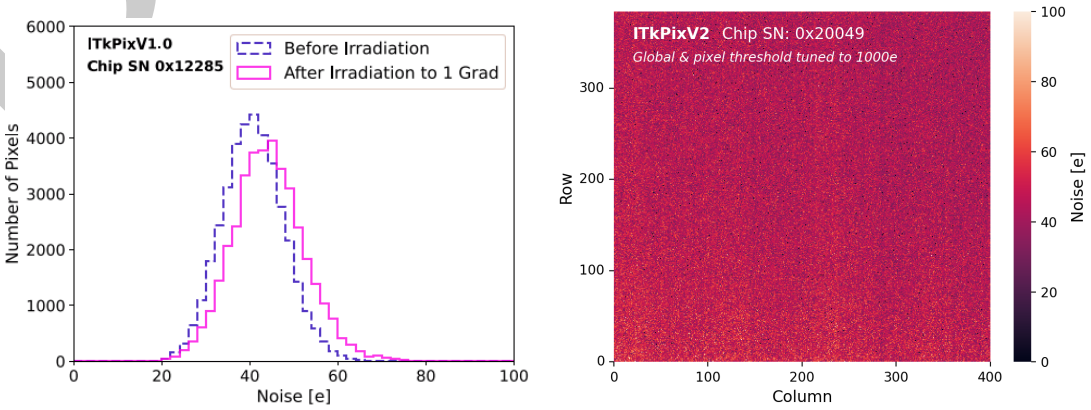
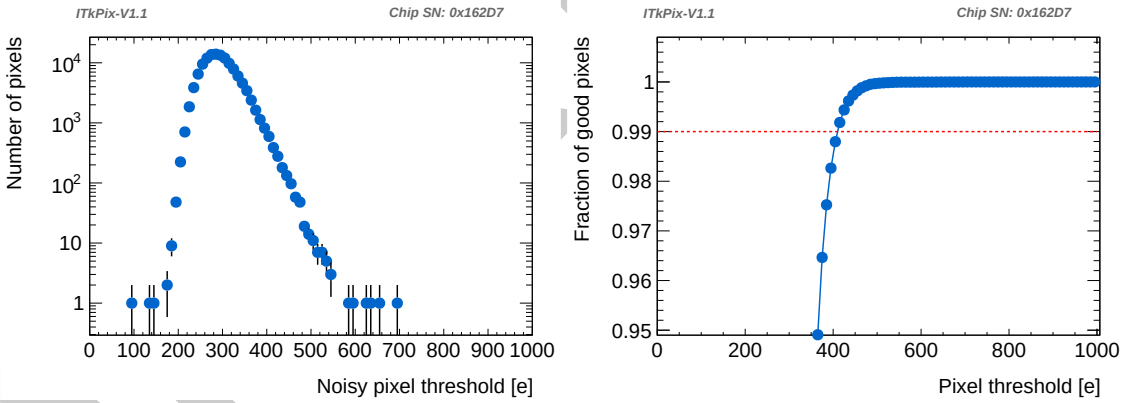


Figure 40. RD53C-ATLAS Differential AFE noise variation at 1000 e threshold. Left: Before and after 1 Grad Irradiation. Right: Variation across pixel array after 1 Grad



(a) (b)

Figure 41. RD53B-ATLAS Differential AFE noisy pixels as function of threshold. Left: Number of noisy pixels (relative noise occupancy greater than 10^{-6}). Right: Fraction of not-noisy pixels. The differential AFE can be seen to have excellent noise performance with thresholds as low as 500 e.

5 Data buffering and triggering

Alternative hit buffering and triggering architectures have been evaluated to choose a final implementation fulfilling trigger latency buffering requirements, with the lowest possible hit loss and acceptable power consumption. Fitting this in available area in the pixel array is a critical design constraint. Sharing of hit buffering between 4 neighbor pixels was quickly identified to be critical to profit from locally clustered hits from a single particle (typically from 1 - 4 pixel hits per cluster). Initial studies found a pixel region of 2×2 pixels to be ideal for the high hit rate in the middle of the inner barrel layer. Further studies, with detailed Monte Carlo hit data from different parts of the detectors, with both 50×50 μm^2 and 25×100 μm^2 sized pixels, determined that a pixel region of 4×1 pixels is a better overall optimization for the two pixel detector layouts. Pixel hits are clustered from traversing particles depending on multiple factors: location of traversing particle, particle angle, sensor thickness, magnetic field, and also radiation damage in the pixel sensor. Two alternative buffering architectures were implemented in the RD53A prototype [12]. The "zero-suppressed FIFO" architecture uses two levels of shared FIFOs to minimize required number of storage bits, at the cost of increased logic complexity. The "distributed latency counter" architecture minimizes logic complexity, at the cost of an increased use of memory bits. Both schemes were found fully functional in simulations and in the RD53A chip. The final choice of using the distributed latency counter architecture was based on effective hit losses, and minimizing logic and layout complexity to assure best possible SEU/SET tolerance.

Sampled pixel hit signals are processed and buffered in small local pixel regions consisting of 4 pixels. When one, or multiple, pixels in a pixel region have a hit, a 4 bit TOT register per pixel stores the measured TOT. The four TOT values in the pixel region are stored in a local latency buffer location together with a 9 bit Bunch ID time-stamp from a central 40 MHz Bunch-ID counter, as indicated in figure 42. A TOT register value of 1111 bin indicates that no pixel hit has been detected. Writing to a 4 pixel buffer location is completed when all 4 TOT counting measurements are finalized. The pixel region hit capture and buffering is non-blocking so a new hit arriving in following clock cycles, on a pixel not part of the first cluster, is captured in the next free buffer location. Each pixel region has 8 local latency buffer locations. Hit losses from this, at the highest hit rates of 3 GHz/cm^2, have been modeled and simulated with Monte Carlo hit data and shown to be well below 1 % [12], as shown in figure 43. Significant design efforts have been invested to fit the required latency buffering in the highly constrained pixel area, using a custom made compact multi-bit latch and highly optimized logic. Effective hit losses has been measured with X-ray irradiations of a pixel module, as shown in figure 44 and scaled to anticipated HL-LHC hit rates (compensated for different cluster size between X-rays and particles in the HL-LHC environment).

When a latency buffer location is in active use, the stored Bunch ID is continuously compared to a global latency counter with a relative offset, defining effective trigger latency. When they match and an active trigger is generated, the buffer location is flagged as triggered, or the buffer location is released. Bunch ID information is then replaced with a trigger event ID to handle the readout of multiple pending triggered events with hit data.

Digital logic in the pixel array uses optimized clock gating to obtain significant power savings. Pixel

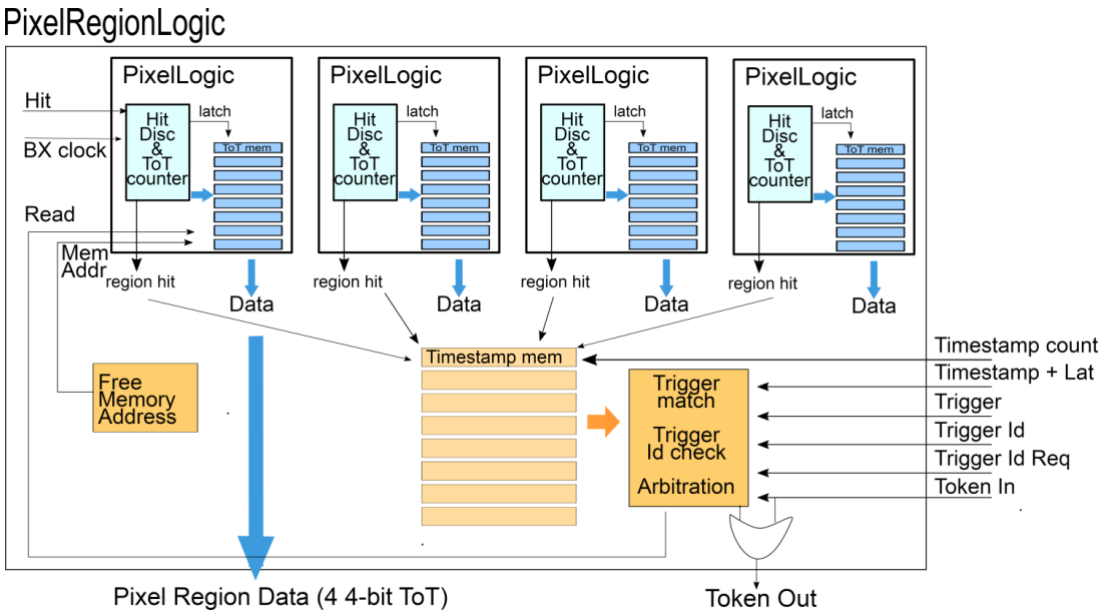


Figure 42. Pixel region logic for distributed latency counter buffering with TOT and BX-ID time tags (Timestamp count) [12]. Hit detection is made per pixel, with storage of associated hit TOTs (blue). Hit time stamps are stored in common with associated buffer management logic, handling triggering and token based readout from the pixel array (yellow).

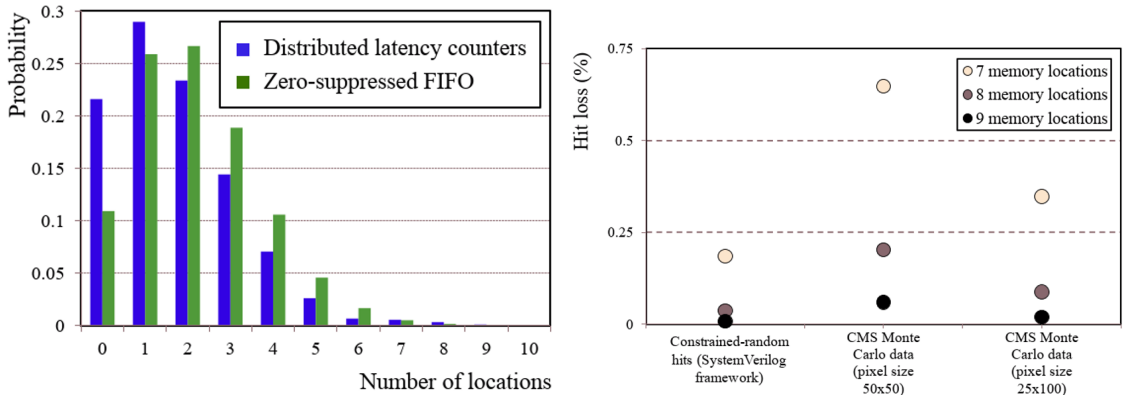


Figure 43. Left: Pixel region latency buffer occupancy probability for two alternative architectures evaluated [12]. Selected architecture is distributed latency counter buffers for 4×1 pixel regions as it has lowest hit loss for high hit rates and has simplest and most compact implementation. Right: Hit loss probability for distributed latency counter buffers with detector Monte Carlo hits at 3 GHz/cm^2 for 7, 8 and 9 buffer locations. 8 Buffer locations are used in final chip implementations as it fits available area and have acceptable hit loss in the highest rate regions (below 0.25 %)

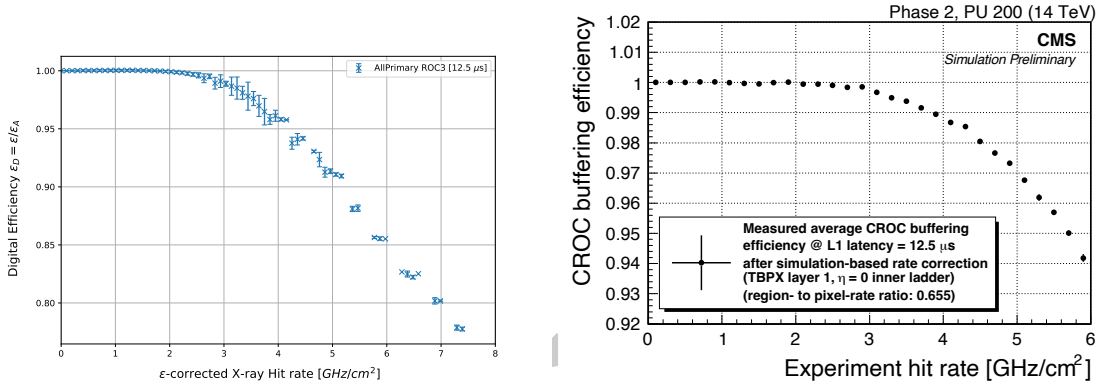


Figure 44. RD53B-CMS Measured X-ray hit loss probability in a non irradiated chip and planar sensor from latency buffering as function of hit rate and corrected for analog dead time (left) and shown for estimated equivalent HL-LHC hit rate (right), with particle cluster size of 1.53 hits/cluster.

region hit capture logic only has active local clocking during the capture window of a hit (effective time window depends on TOT length), making the instantaneous power consumption dependent on hit rates. This requires carefully optimization of local power decoupling capacitors on-chip and on pixel modules, to work reliably with serial powering.

Readout of triggered hit data from the local pixel region latency buffers is controlled by a core column readout controller at the end of each core column bus. Pixel cores, consisting of 2×8 pixel regions (8×8 pixels), share a core column readout bus, with its associated controller in the DCB. Readout from the pixel array is initiated by the core column controller signaling the event ID and asserting a readout token. Pixel regions having triggered hit data await the arrival of the readout token and then assert their hit data on the readout bus together with its pixel region address and passes the token. When the token finally returns to the pixel core column controller, all event data in the core column for this event ID has been collected. Pixel core columns have independent readout controllers that can be in the process of reading out different events. This improves the effective readout rate from the array when having multiple pending triggered events. A central trigger table keeps track of events awaiting readout from the pixel array.

A pixel core column bus is covering a large number of pixel regions. This limits the effective readout speed on this long bus and makes it significantly affected by radiation degradation of its bus drivers and handshake logic. The effective readout time is two clock cycles per pixel region with hit data, that with radiation degradation can get as long as 3(4) clock cycles (configurable). It has been confirmed in simulations that such a reduced pixel array readout speed is compatible with required hit and trigger rates. In practice it has not yet been seen necessary to use this extended readout period for highly irradiated chips.

It is possible by configuration to constrain the maximum number of pixel regions to read out from each core column per event, to prevent possible readout congestion from events with excessive number of hits. It is also possible to constrain the maximum time available to readout all pixel core columns, thereby effectively constraining the maximum number of hits per event.

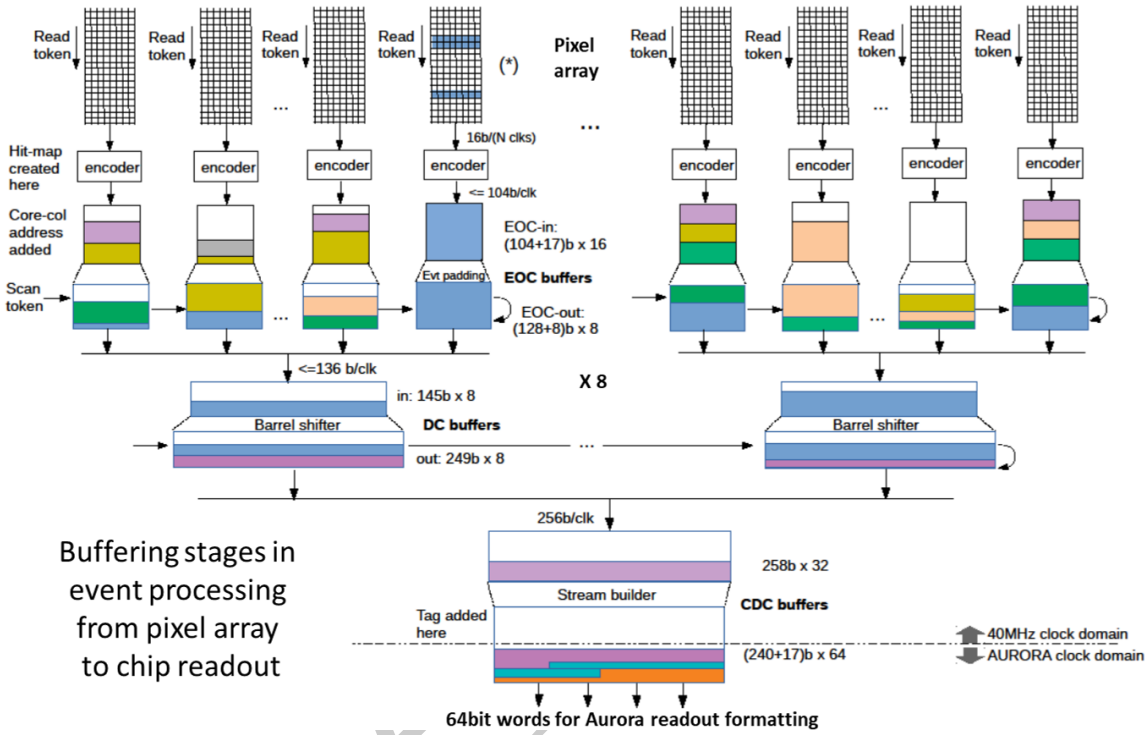


Figure 45. Outline of processing and buffering of event data in multiple stages from pixel core columns, processed by End Of Column (EOC) logic, to final readout link Aurora formatting via Clock Domain Crossing (CDC) buffer. Intermediate FIFO's are used for data buffering to enable the different stages to work concurrently to sustain required bandwidth. Barrel-shifters are used for effective data merging and re-packaging of zero-suppressed event data between processing stages. Colors shown in data buffers represent event data, belonging to same triggered event, in different stages of processing.

An extended two level trigger mode for potential future trigger upgrades has been implemented. In this mode, L0 triggered hits remain in the pixel region latency buffers for a configurable time-out period (max 25.6 μs). During this time-out period (L1 trigger latency), events can be flagged for readout (L1 accept), or by default be rejected (L1 reject).

Event data accepted for readout will go through multiple levels of processing, event building, buffering and formatting, as shown in figure 45, before being ready for final readout via the serial readout links. Total event data buffering before final readout is of the order of 25 kBytes in the DCB. Significant hit data buffering also takes place in the pixel array, from a trigger is received until having been read out from the pixel array. This buffering assures efficient hit data de-randomization that enables good readout bandwidth utilization.

6 Control and readout

The control and readout interfaces of the RD53 chips are highly constrained from their specific use in an inner high rate and low mass detector with LPGBT based optical links to DAQ and control systems. An efficient variable length hit data encoding format, called binary tree encoding, has been developed to minimize readout bandwidth. The use of 1, 2, 3 or 4 readout links per pixel

chip, and the option of merging data from 2 or 4 chips into one link, enables the number of required readout cables to be optimized and minimized for different system configurations. Control and readout links use DC balanced encoding, for the AC coupled links required in a serially powered detector system.

A 160 Mbits/s DC balanced differential control link, with transmission error detection, has been specifically developed to address up to 15 chips (e.g. chip specific configuration) and with broadcast capability (e.g. common configuration). It has an embedded 40 MHz reference clock with sub-ns timing control, to appropriately align pixel hit sampling with bunch collisions. Real time commands at 25 ns level have priority over control, configuration and monitoring commands. The control link has sufficient bandwidth to perform continuous scrubbing of pixel configuration, in case needed in the hostile radiation environment (see chapter 9).

A radiation hard 1.28 GHz PLL (Phase Locked Loop) has been developed for appropriate Clock and Data Recovery (CDR) from the control link and generate the clock for the serial readout links. Initial prototypes have been extensively tested and gradually improved to get lower jitter with sufficient TID and SEU/SET tolerance[21][22]. The PLL locks to the 160 Mbits/s control stream and generates required on-chip clocks. The 40 MHz hit sampling clock is generated with a frame alignment circuit, based on regular sync symbols. It can be phase shifted in steps of 0.78 ns to perform precise time alignment to particles from the HL-LHC collisions. The PLL is separately powered to allow additional external filtering of the analog chip power in case needed. The classical PLL architecture, with frequency and phase detectors, is shown in figure 46 with measured jitter and eye diagram of a readout link shown in figure 47.

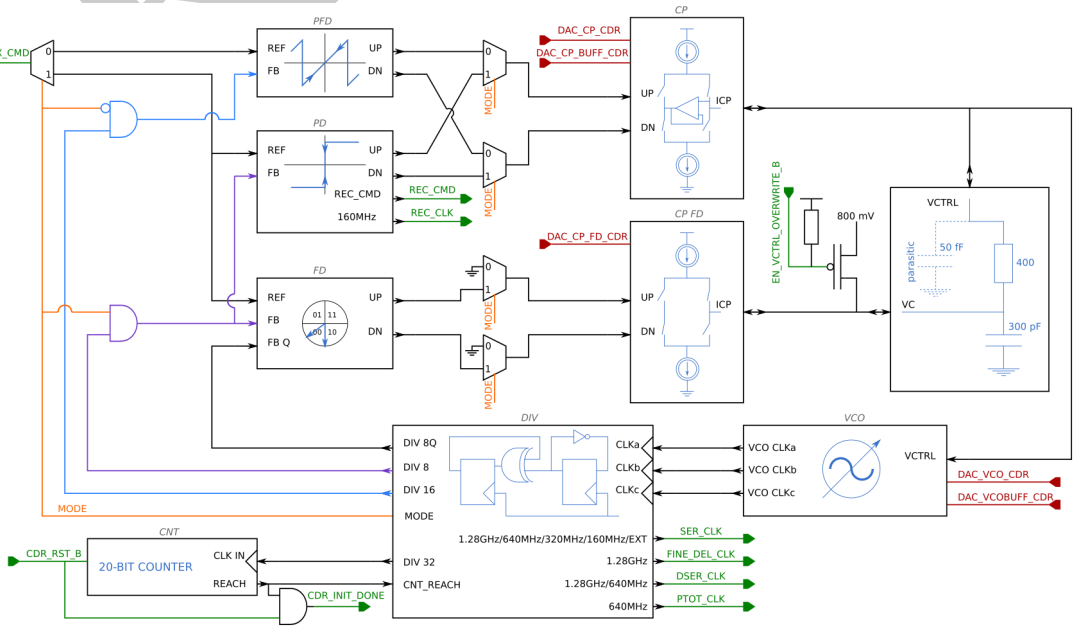


Figure 46. PLL generating high frequency clocks used in the chip. PLL control loop with combined Phase Detector (PD) and Phase - Frequency Detector (PFD) controlling a Voltage Controlled Oscillator (VCO) via analog Charge Pumps (CP) and loop filter. Frequency multiplication from the 160 Mbits/s control link to the 1.28 GHz serializer clock is obtained with SEU protected counters (CNT and DIV).

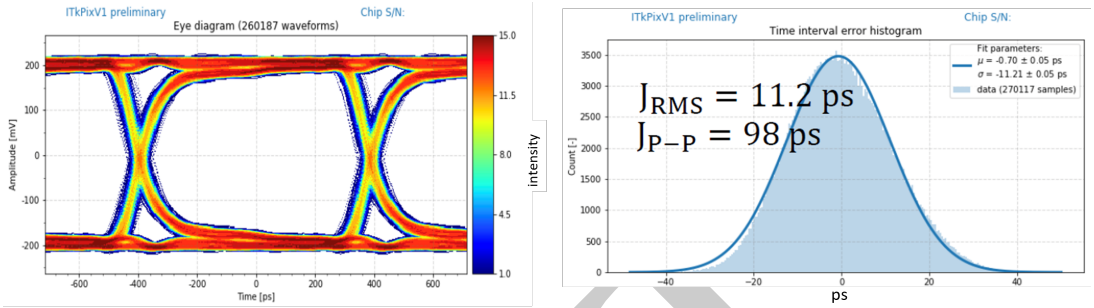


Figure 47. Measured eye diagram and PLL jitter on 1.28 Gbits/s serial readout.

Particular emphasis has been put on efficient and reliable startup, resetting and configuration of the chips for use in a serially powered detector system in a hostile radiation environment. At power-on startup (see also SLDO startup in section 7) the chip will initially use a default hardwired configuration. Only when all critical re-configuration data has been downloaded, will these settings be activated with a dedicated enable code. Full chip data path and buffers can be cleared quickly or specific parts of the chip can be reset by specific commands. In a worst case scenario where control link synchronization is lost, and it does not self recover as it should normally be the case, a dedicated link reset can be applied that initializes all chip configuration and starts a full chip re-synchronization (as done at power up). This is done by running the control link at a low frequency (invalid link bit rate, but still compatible with AC coupling), that is detected by the chip to be out of normal working range. This removes the need of using power cycling to recover chip operation, which is highly undesirable in a large serial powered system with high voltage sensor biasing.

Up to 4 readout links of 1.28 Gbits/s (or 640 Mbits/s or 320 Mbits/s) are available per chip for readout and monitoring. A subset of the Aurora encoding [20] is used as it supports all the required features: DC balanced 64B/66B encoding, framing with minimum overhead, multi lane support, data and service type frames. Aurora formatting is well documented and well supported for FPGAs in test and DAQ systems, with general event and service data formatting as indicated in figure 48. When used in final ATLAS/CMS pixel detectors with LPGBT optical links, two levels of link encoding (Aurora 64B/66B plus LPGBT FEC) will be present, to be decoded by the DAQ system FPGAs. It should be noted that Aurora formatting does not have Forward Error Correction (FEC). Single bit transmission errors (or SEUs in pixel chip serializer) can therefore occasionally cause the corruption of event fragments. The use of FEC was considered but it was found to have too large bandwidth overhead, especially in combination with the LPGBT that has extensive error correction. The RD53C-CMS chip has the option to add a CRC (Cyclic Redundancy Check) at the end of each event.

Raw zero-suppressed hit data from the pixel regions consist of a pixel region address followed by 4 bit TOT information from the 4 pixels in the region, with TOT=1111 bin indicating no hit. This is already a relatively efficient data format for clustered hits, compared to individual pixel hit addresses with TOT (18 bit pixel address + 4 bit TOT = 22 bit per hit). Hit cluster size and shape from a single particle depends on many factors: particle location among pixels, particle angle, sensor thickness,

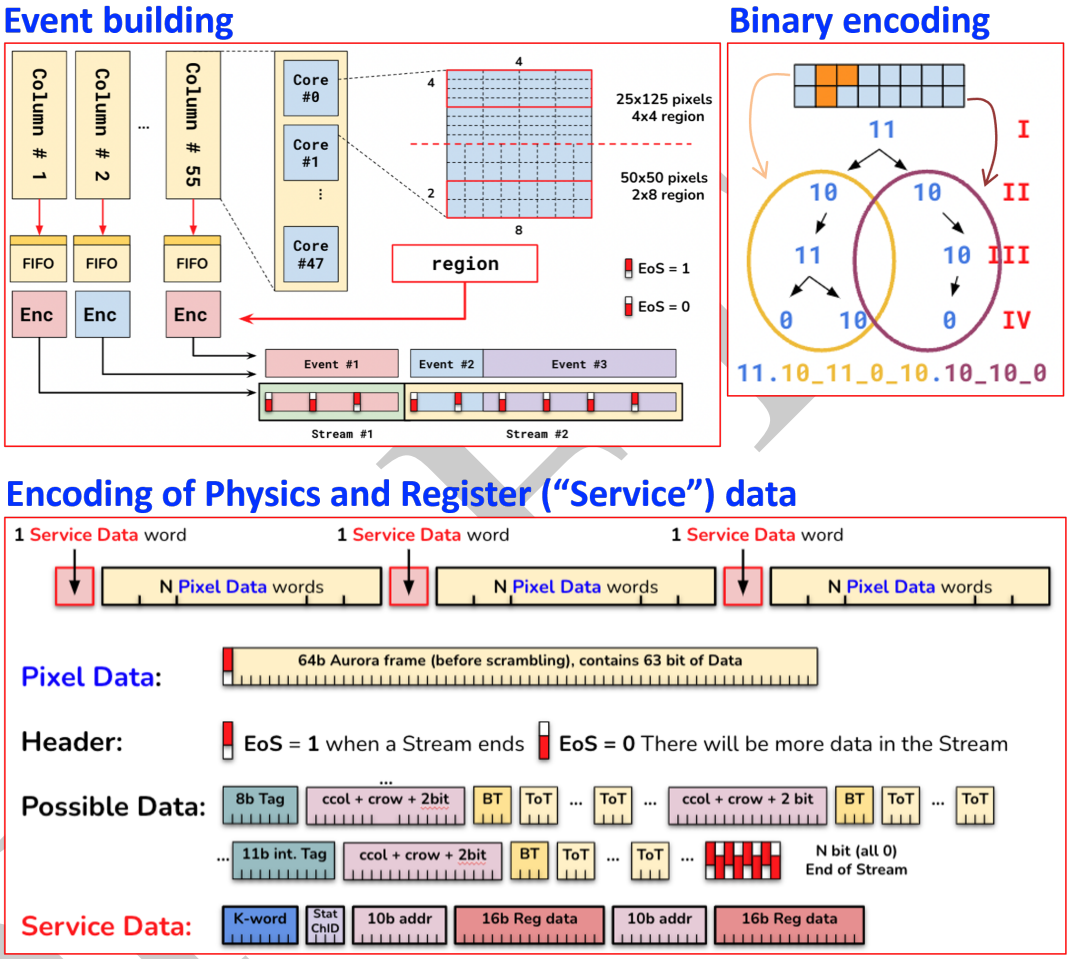


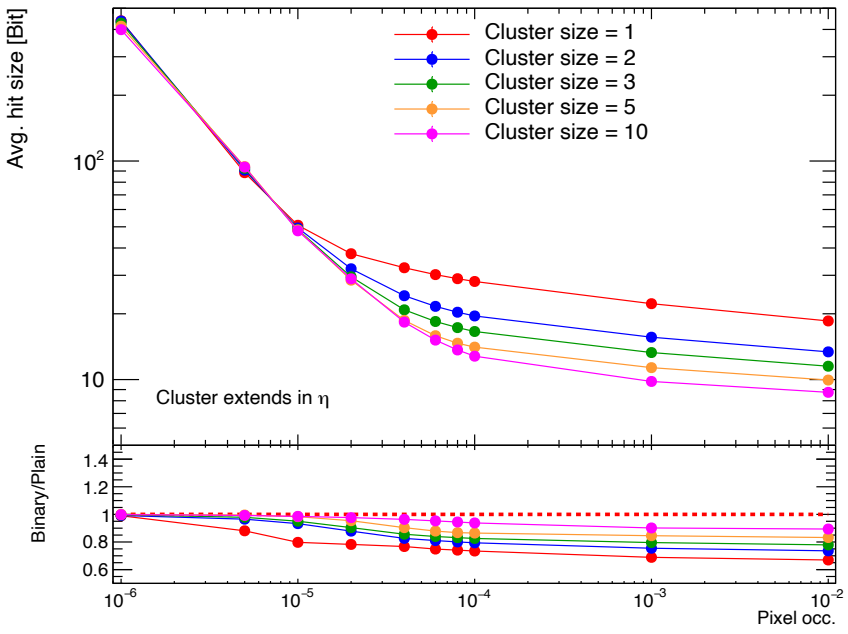
Figure 48. Outline of event building, binary hit encoding and Readout event encoding [11]. Event building is performed on-chip between event fragments from pixel array columns. Binary tree hit encoding taking advantage of multiple pixel hits in local clusters. Encoding of physics data (pixel hits) and register data (monitoring or register reads) being indicated in 66 bit frame header. Hit data consists of binary tree encoded pixel addresses, followed by corresponding TOT information that can optionally be omitted. Event data from single events can be contained in well separated single-event streams (indicated with EOS=1), with required 66 bit frame bit padding at the end, or alternatively as a multi-event stream with reduced bit padding overhead. Service frames are sent a regular intervals (configurable) with requested monitoring and configuration read-back data.

magnetic field and radiation damage in the pixel sensor. An optimized binary tree hit encoding scheme can further reduce readout bandwidth by 10-20 %. It is also possible to suppress TOT charge readout information, having only binary hit information, giving a data reduction of ~30 %.

In binary three encoding, the 16 bit hit map from 4 pixel regions covering 8×2 pixels, is encoded to produce a compressed hit map representation, with fewer than 16 bits per hit on average for clustered hit data. The algorithm divides the hit map, containing one or multiple hits, in half (e.g. upper and lower half as shown in upper right corner of figure 48) and labels each half as containing hits (1) or not (0). This is applied recursively to every non-empty hit pattern until only 2 bit hit

patterns are left. A bit code substitution is then applied to compress this. In any step, the two-bit code 00, for the two halves without hits, is zero-suppressed. Code 01 is represented by a single bit set to 0, while 10 and 11 are kept as a 2 bit code. This results in a compressed 16-pixel hit map with between 5 bits (single hit) and 30 bits (all 16 pixels hit). The encoded hit map is preceded by a pixel hit map address (location in full pixel array) and followed by pixel hit TOT(s). This encoding is in the chip implemented with simple and fast logic based on a small look-up table. A complex Huffman encoding was in simulations seen to obtain 10-20 % better compression on HL-LHC data, but requires complex on-chip processing. The effective number of bits per hit with binary tree encoding has with Monte Carlo hit data been seen to be in the range of 10-15 bits/hit, as shown in figure 49 depending on cluster size. This can be compared to the raw zero-suppressed hit data format with 14-28 bits/hit.

When having a small number of pixel hits per chip, as is the case for outer pixel layers, a relatively large overhead is used for event header information and required 66 bit frame padding at the end. An optional multi-event stream formatting, with multiple events sharing a single transmission stream, can reduce this overhead. It must though be kept in mind that a single bit transmission error (or SEU, SET) can then corrupt multiple events. It is therefore encouraged to use single stream event formatting when ever possible, as this enables correct data decoding to be reestablished at the start of each event.



(a)

Figure 49. Simulated number of bits per hit with binary tree encoding and 4 bit TOT as function of pixel hit occupancy and cluster size. Lower: Relative data bandwidth gain from binary tree encoding.

Data merging between chips is available for low rate outer pixel layers, to merge readout data from 2 or 4 chips on a pixel module into a single readout link, thereby significantly reducing the number

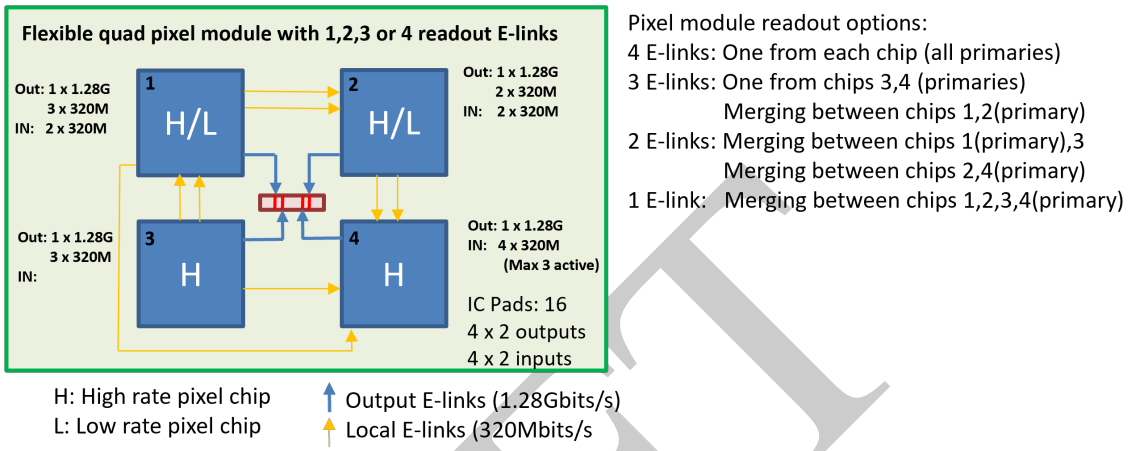


Figure 50. Data merging between primary and secondary chips on a flexible quad module that can be configured to have 4, 3, 2 or 1 active readout links. Up to four 1.28 Gbits/s readout links shown in blue on module connector in center. Local 320 Mbits/s data merging links between chips shown in yellow.

of required readout links/cables. A primary chip is driving a single 1.28 Gbits/s readout link and 1 or 3 secondary chips drive 320 Mbits/s serial data on a local single or dual lane link to the primary chip. Chip to chip data merging requires the chips to be driven by the same control link and the interface is based on oversampling at 640 MHz in the primary chip. Used in a flexible manner, a quad pixel chip module can be configured to have 4, 3, 2 or only 1 readout links as shown in figure 50. Data merging uses simple frame by frame time multiplexing with a 2 bit source ID, so the DAQ system must handler 2 or 4 independent event streams on the readout link.

Event readout latency (time interval from receiving trigger message to completion of corresponding event readout) will have a complex dependency on statistical fluctuations in hits and triggers and the available readout bandwidth as indicated in Monte Carlo simulations shown in figure 51. The readout latency can become particularly long if the readout bandwidth is highly utilized (e.g. above 90 %). The chip will eventually be forced to drop events, if on-chip data buffers run full (flagged in chip monitoring).

Readout links are driven by differential CML (Current Mode Logic) drivers with configurable drive current and pre-emphasis. The driver has a 100 Ohm differential output impedance, assuring the best possible matching to low mass 100 Ohm differential electrical cables (twisted pair, flex micro strip-lines, twinax), with the absorption of possible transmission reflections.

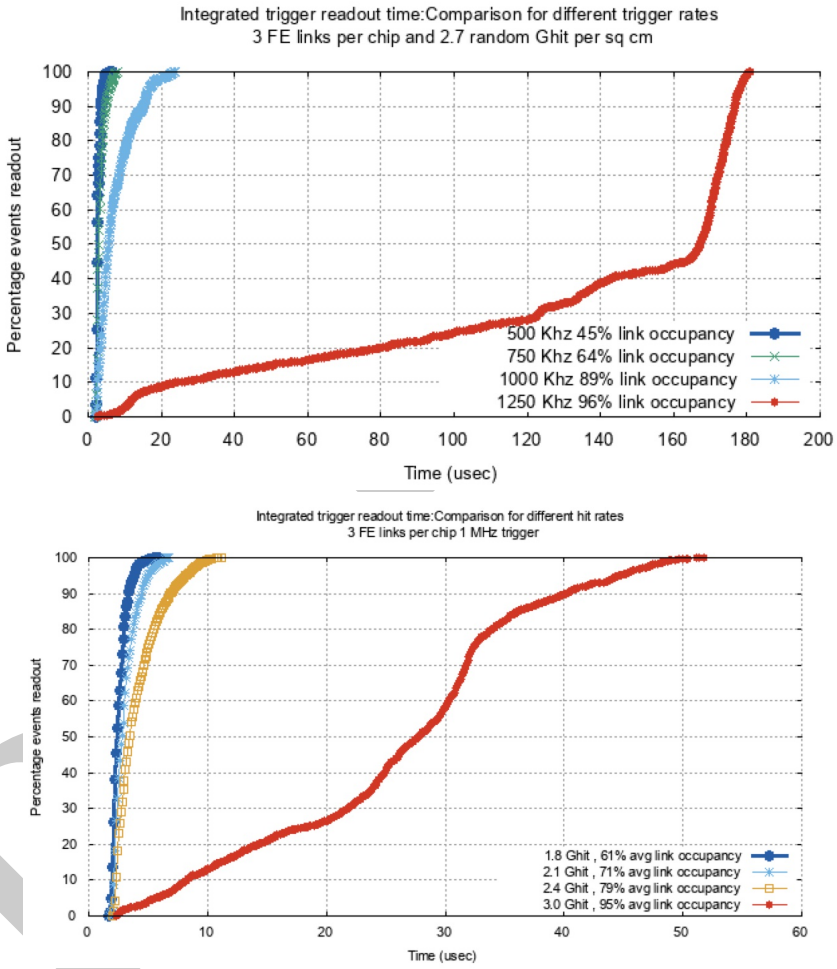


Figure 51. Simulated readout latency for pixel chip with 3 readout E-links. Upper: at 2.7 GHz/cm^2 hit rate, for different trigger rates. Lower: at 1 MHz trigger rate, for different hit rates. It can be noticed that the readout latency gets excessively long above an average readout link utilization of 90-95 % . Above a readout link utilization of 95% there is a significant risk of loosing event fragments.

7 Power and references

RD53 chips have on-chip SLDO power regulators for serial powering of pixel modules, with chips on the same module connected in parallel [30]. These SLDOs can also be (hardware) configured as classical LDOs for parallel powering with on-chip voltage regulation or be by-passed for direct powering. Separate SLDOs are implemented for analog and digital power domains to assure the best possible noise isolation between the digital and analog parts of the chip.

In SLDO mode, being powered by a constant input current, the chip I-V characteristics is as indicated in figure 52. Above a given minimum operation current, the SLDO actively regulates a constant output voltage, independent of the load current drawn, with a well-defined input impedance. This is in practice obtained with a regulated shunt current at the output of a LDO. A simplified schematic of the SLDO is shown in figure 53 with the large pass-device and shunt power MOSFETs marked

in red. A classical shunt power regulator [25] normally has a "flat" I-V curve with very low input impedance, and in practice it can not be used in parallel with controlled current sharing. The RD53 SLDO has configurable current sharing between multiple parallel chips, by having a small and well controlled input impedance with a configurable offset voltage V_0. Both impedance and V_0 are configurable by external resistors. Using programmable configuration registers for this has been considered potentially hazardous in case of misconfiguration or SEUs. One-time programmable E-fuses have been seen to become unreliable when exposed to more than 100 Mrad. The offset voltage (V_0) can optionally be shared among multiple chips on the same module, with effective current sharing then given by the impedance ratio among chips connected in parallel. Appropriate current sharing between digital and analog regulators, on the same chip, is done in the same fashion. The SLDOs have distributed power output stages along the chip edge, as shown in figure 9, to distribute generated heat and assure best possible power uniformity across the chip.

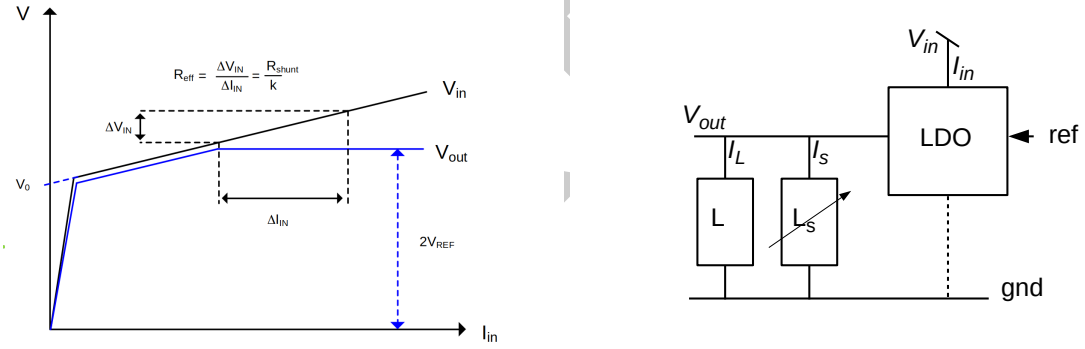


Figure 52. Left: Ideal SLDO I-V characteristics, without load, where operational range is from point where output voltage V_{OUT} is stabilized. Right: SLDO made from special LDO with regulated shunt current (L_s) in parallel to load (L).

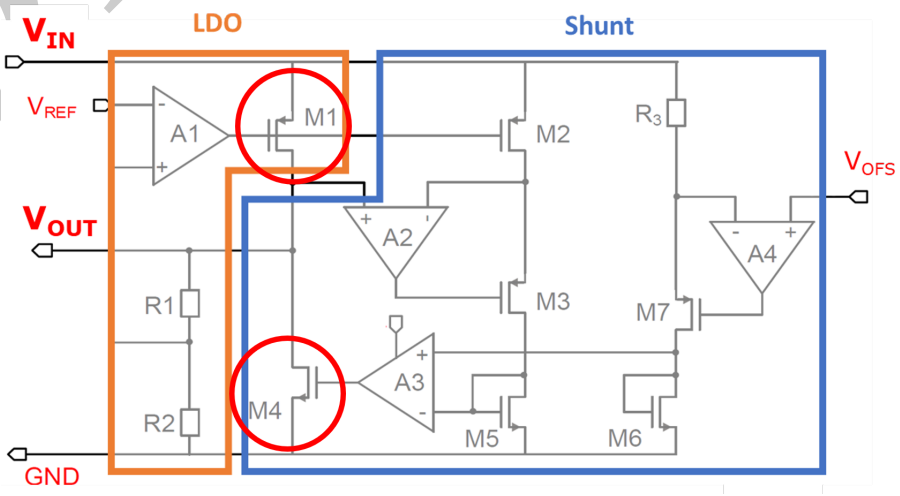


Figure 53. Simplified SLDO implementation with power FETs M1 (LDO pass-device) and M4 (shunt) indicated in red. V_{OFS} determines effective SLDO offset voltage V_O.

The RD53 SLDO regulator has gradually been improved in different chip generations with additional

and improved features for reliable start-up independent of system configuration and current ramping rate, overload protection, over-voltage protection, and improved current monitoring [27].

A dedicated low current pre-regulator has been introduced in latest chip generations to assure reliable and stable over voltage protection together with power start up and generate stable biasing levels for analog blocks in the chip, as shown in figure 54. The pre-regulator (PreReg), using a fixed over-voltage tolerant bandgap reference, supplies power for a high precision and configurable core bandgap (CoreBGR) [28] to derive all analog voltage and current biases via an external high precision resistor (R_IREF). 4 wire-bond pads (IREF_TRIM pads) are available to make process and chip specific fine adjustment if needed, based on wafer probing characterization. This assures precise and stable biasing, independent of temperature and input voltage, as shown in figure 55 with less than 1 % variation in the nominal operation temperature range of -20 - -10 °C. Bandgap voltages have consistently been seen to be 40-50 mv different than what have been predicted by detailed circuit simulations (490 mv versus 450 mv). This has been compensated for by the external bias current resistor and voltage adjustment DACs.

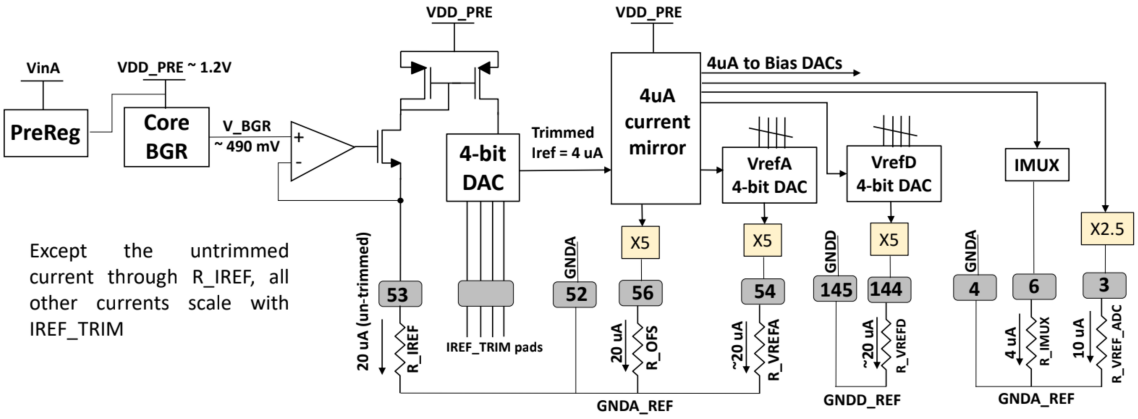


Figure 54. Pre-regulator (PreReg) and high precision core bandgap (CoreBGR) used to derive all analog chip voltages and current biases (AFE, PLL, ADC, Drivers, Receivers). The bandgap voltage reference is converted to a reference bias current via the external R_IREF resistor and 4 bit tuning DAC, set by bond-wiring (IREF_TRIM pads) on pixel module. On-chip SLDO generated analog and digital supply voltages can be tuned with 4 bit configuration DACs (VrefA and VrefD) with adjustment range set by external resistors (R_VREFA, R_VREFD).

Voltage and current references, driven by the core bandgap, have been seen to have a near linear drift of 10 % in final chips when being exposed to 1 Grad TID, as shown in figure 56. Initial chip versions had up to 20 % TID drifts. It is therefore anticipated that SLDO output voltages and AFE biasing will have to be adjusted yearly. The SLDO regulated analog and digital power supply voltages can be adjusted by configuration in a limited safe range of 1.1 - 1.4 V. The voltage reference drift to the monitoring ADC also requires appropriate off-line compensation (see section 8). The pixel detector groups are currently investigating how best to deal with this issue at system level.

Guaranteed and safe startup of serial powering has been a major challenge that has required several incremental SLDO improvements and specific features over the different chip generations [29]. A

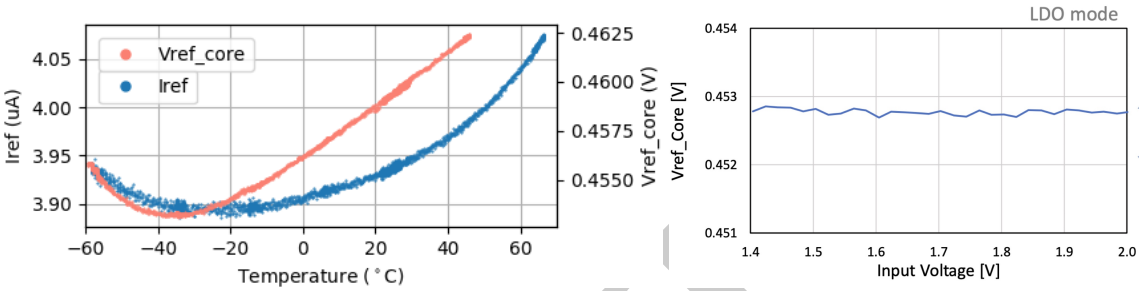


Figure 55. Reference current stability as function of temperature (left) and precision bandgap voltage as function of input voltage (right).

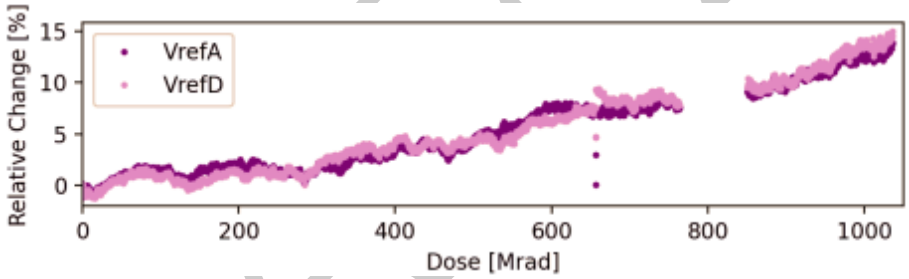


Figure 56. RD53B-ATLAS Biasing reference dependency on irradiation.

large set of variable system and chip parameters affects this: Number of chips in parallel, number of modules in series, chip differences, temperature, radiation effects, loading differences, dynamic load variations, local power decoupling, cable inductance and startup current ramping. Startup of analog and digital SLDOs have been verified in a multitude of different system configurations at different operation conditions. As an example, correct startup at 20 C° and at -50 C° are shown in figure 57. The SLDO output voltages are seen to be well stabilized at 1.25 V with a well controlled SLDO input impedance, when injected current above the consumed on-chip current (1.0 - 1.5 A depending on chip configuration). At low input current the SLDO can not actively regulate the on-chip supply voltages. -40 C° is considered the absolute coldest startup temperature, when the detector cooling is running and the pixel chips are off and not dissipating power.

The SLDO has a special low power mode for detector connectivity verification without active CO_2 cooling during installation. In normal mode each SLDO will typically be configured for an operation current of 0.8-1.0 A. In low power mode, the SLDO is forced to work with an increased offset, Vo, giving a sufficiently high input voltage at an input current of 0.1 A. The low power mode is enforced at power-up with a dedicated ~100 kHz AC control signal, as all control signals in a serially powered system must use local AC coupling.

Active over-voltage protection has been included to protect the system, and pixel chips, against potentially damaging dynamic and static over voltages occurring during power cycling and during potential power anomalies. This is controlled by the over-voltage tolerant pre-regulator and works as an effective voltage clamp. In case of open chip failures, as shown in figure 58, the voltage clamp will constrain the module voltage to 2 V when two out of four chips do not carry any current (or for

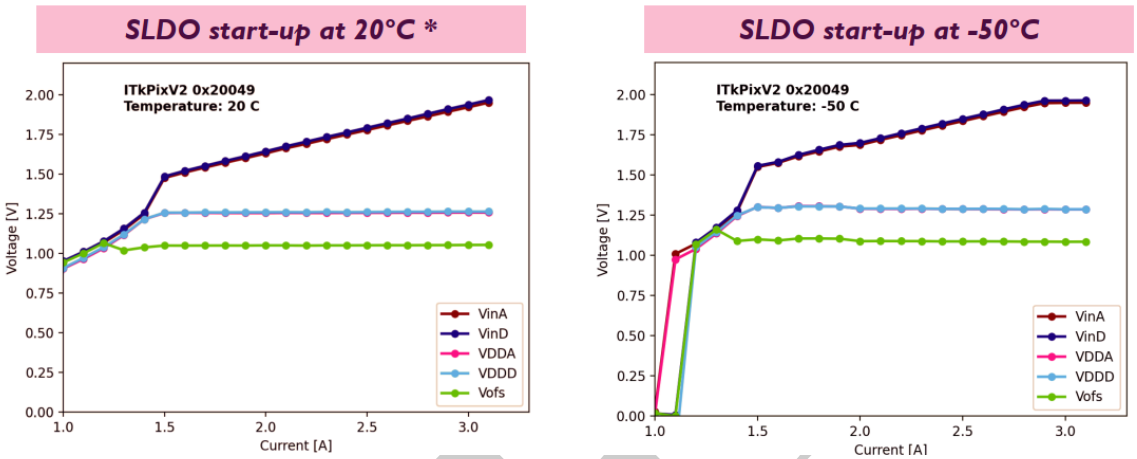


Figure 57. Analog and Digital SLDO startup at 20 C° and -50 C°. Vofs is an internal voltage reference setting the voltage offset of the input impedance. Input current is the sum of the analog and digital SLDO input currents. Analog and digital SLDO input and output voltages are overlaying, so hard to distinguish.

781 a dual chip module with one chip with open power failure) .

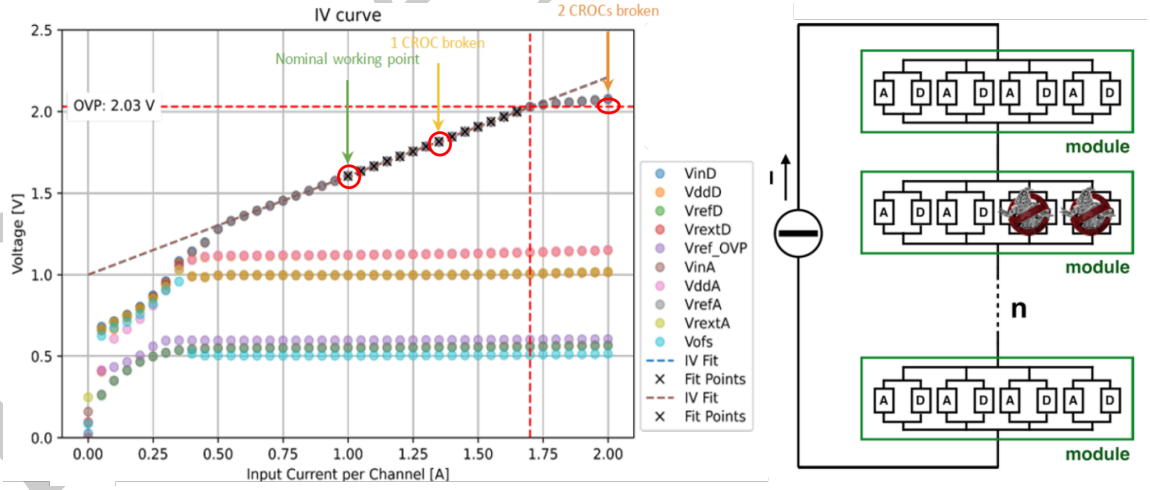


Figure 58. Serial power chain with 0, 1 or 2 failing chips with a power-open on a quad chip module. Each chip shown with their analog (A) and digital (D) SLDO regulators. SLDO transfer characteristics (input current - input voltage) shown for a single working chip during power ramp up. The SLDO current of a working chip (Input current per channel) is normally 1.0 A when all four chips on the module works correctly. With a failing chip on the module, the input current of the working chip increases by 33 % , where Vin is still seen to be on the linear SLDO regulation curve. When two chips on the module fails, the current per working chip increases by 100 % and the over-voltage protection can be seen to become activated, limiting the module input voltage to 2.0 V. Regulated supply voltages for the working chip can be seen to maintained at constant voltages, which in this particular case has been configured to be 1.0 V for digital and 1.2 V for analog. Internal reference voltages, shown in lower part of the plot at 0.5 and 0.6 V are also maintained constant in the different failure scenarios.

782 In a serial powering system, with local parallel connected loads, a single chip with excessive current

consumption can provoke their common input voltage to become reduced or collapse. This only affects parallel connected chips, and not other pixel modules in the serial power chain, as long as the serial power supply current is maintained. An optional overload protection can be enabled to protect the system against a single chip consuming excessive load current.

It must be noted that power opens are the critical failure mode, as this can prevent the serial power supply current to flow. Grounding faults between local chip grounds and system ground will also required a whole serial power chain to be turned off. Local power shorts, or excessive load current, in a single chip on a pixel module will only prevent a single module to be appropriately powered, as long as the serial current is conducted to the other modules in the serial power chain.

The typical power supply current required by a pixel chip for an inner high rate pixel layer is 0.8 A for analog and 0.8 A for digital, giving a total active chip power of ~2.0 W. For outer pixel layers with lower hit rates this can potentially be reduced by 0.1-0.3 A (lower biasing for AFEs and lower digital hit activity). For a serial powered system a current headroom of 10 - 20 % must be added together with a SLDO voltage drop of minimum 0.2 V as operation margins. These add up to a total power dissipation of ~3 W per chip. The choice of serial power current headroom is a delicate balance between cooling, margins for production variations between chips (checked during wafer probing), radiation induced power consumption changes (measured to be very small) and anticipated hit and readout rates in the different pixel layers.

The SLDO shunt current capability is designed to be up to 200 % of normal operation current to assure, in case of a single chip power failure on a module, that remaining chip(s) can correctly pass the full serial power chain current. A chip with increased shunt current will in this case have significantly increased power dissipation that must be appropriately cooled. The SLDO hot spot, shown in figure 6 is on the edge of the pixel modules, typically located very close to CO_2 cooling pipes.

Extensive serial powering tests made of pixel modules with RD53B and RD53C chips have demonstrated fully satisfactory functionality for parallel connected chips on serial connected modules [35][33][34]. The general chip performance figures (e.g. thresholds, noise, monitoring) are not affected noticeably by their position in a serial power chain. Figure 59 shows the correct function and current sharing among four chips on a quad pixel module. Radiation has not been seen to have any noticeable effects on the SLDOs, as all its circuits are made with large transistors. Further system studies are ongoing also covering HV sensor biasing with local filters.

On chip power distribution is critical for such a large mixed signal chip. Making quad (2×2) pixel modules requires having all (power) wire-bonds on one side of the chip with the SLDOs distributed along this chip edge. The very tightly packaged pixel array made it impossible to implement distributed SLDO regulators across the pixel array. On-chip power distribution uses all metal layers available and in particular an ultra thick copper layer and a thick aluminium top redistribution layer. Voltage drops have been evaluated with specific power verification tools. This can be compared with measured voltage drops in the analog ground as shown in figure 60. The regular pattern seen is related to the location of 4 distributed instances of the SLDO power stages at the chip periphery as shown in figure 9.

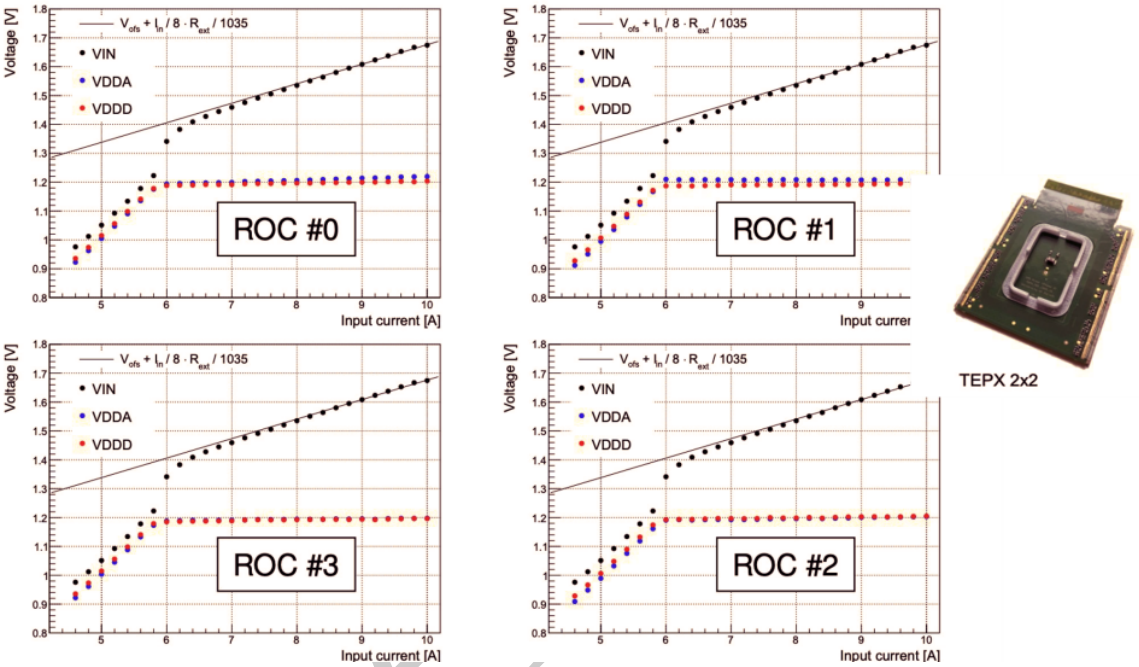


Figure 59. SLDO I-V characteristics of 4 RD53B-CMS chips on a quad pixel module [34]. The shared VIN is the same for the 4 chips that all have their local regulated analog and digital supply voltages well stabilized. The small output voltage variation with input current is known to have been caused by a resistive ground voltage drop in the test setup.

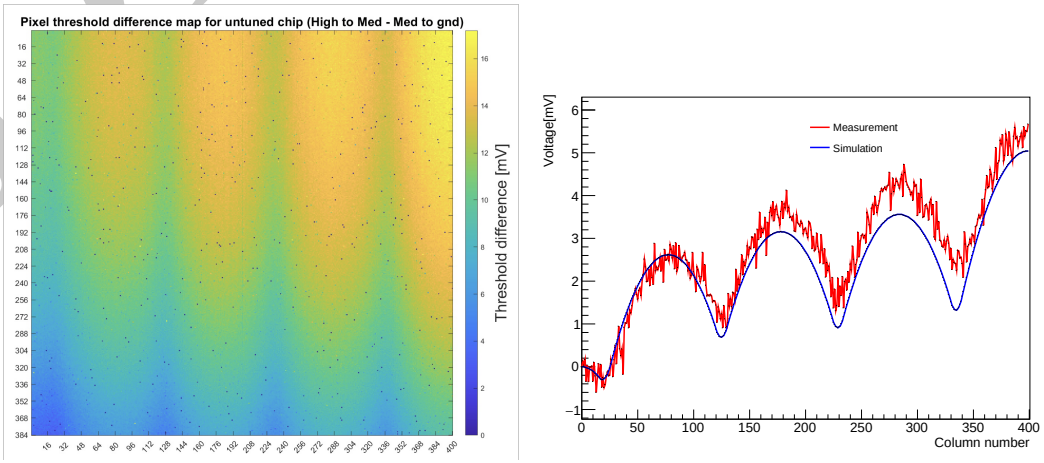


Figure 60. Analog ground variation across pixel array (left) and across pixel columns (right). Measured indirectly by showing effective threshold difference between differential charge injection (Vhigh - Vmed) and single-ended charge injection (Vmed - gnd), being directly sensitive to analog ground variations.

8 Monitoring

RD53 chips have extensive on-chip monitoring capabilities [36] covering: On-chip temperature sensors, Pixel module temperature with external NTC (Negative Temperature Coefficient) thermistor, SLDO input and output voltages and currents, internal references and biasing levels plus radiation effects monitoring. Analog monitoring is performed with an ADC conversion request command on the control link followed by an ADC read request, with monitoring data being read out in dedicated Aurora service frames on the readout links as indicated in figure 48.

Analog monitoring is made with a multiplexed 12 bit switched capacitor ADC [37][38] with layout shown in figure 61 and with measured resolution and linearity as shown in figure 62. Absolute ADC calibration per chip is performed during wafer probing, with ADC calibration parameters for each chip being stored in a central data base.

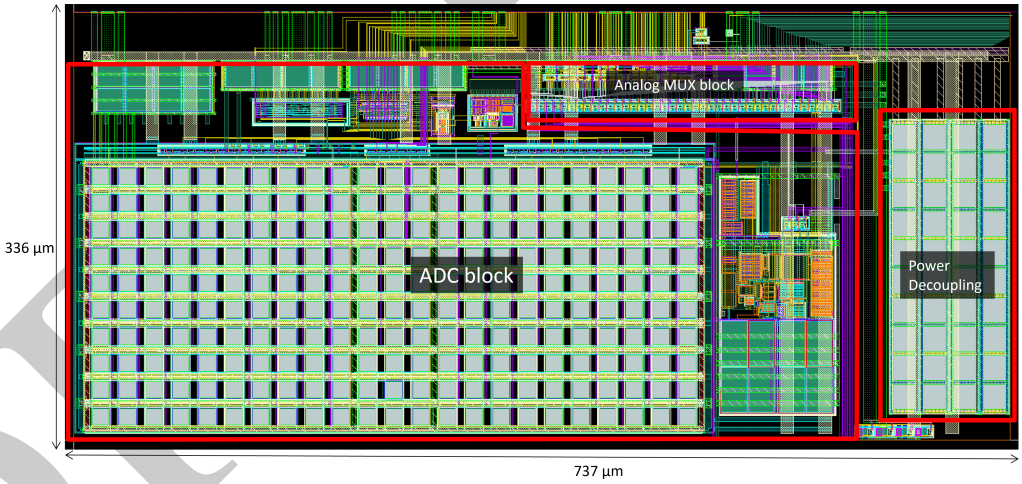


Figure 61. 12 bit switched capacitor monitoring ADC layout with Analog monitoring multiplexer.

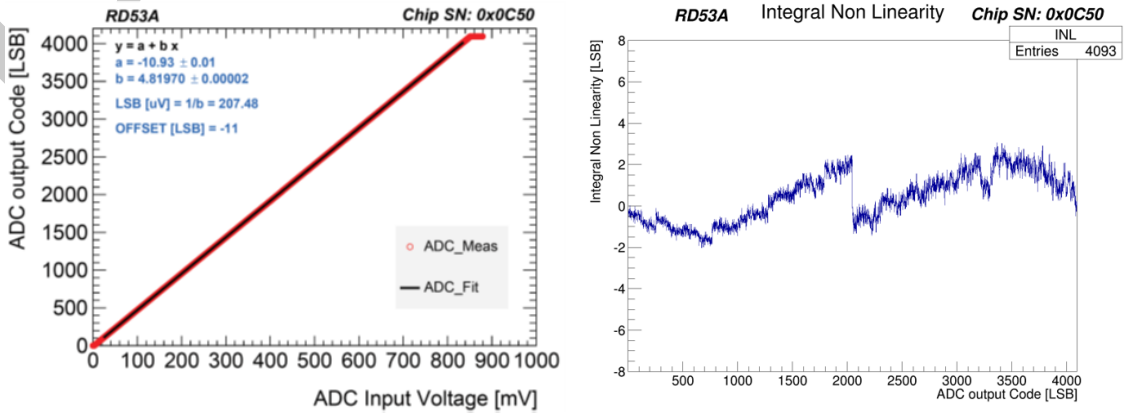


Figure 62. 12 bit ADC linearity and INL.

The switched capacitor ADC core has been shown to have excellent radiation tolerance. However, radiation test campaigns have shown a significant drift of about 5% for 500 Mrad, 10% for 1 Grad,

for monitored voltages. This originates from a TID drift of the ADC reference, which impacts directly voltage measurements as shown in figure 63. A method to correct for the TID drift of the ADC reference has been developed, based on specific properties of the used temperature sensors as described below.

Three temperature sensors in the chip bottom, close to the SLDO power regulators, are based on large-area NMOS transistors biased in sub-threshold region. Temperature measurement accuracy has been greatly improved by making multiple measurements at different currents (configurable) in the NMOS sensor. It has been possible to demonstrate an effective temperature linearity and resolution of ~1 °C as shown in figure 64 with good TID tolerance. This measurement does not depend on the ADC reference voltage and allows to contain TID effect drifts to values lower than 2 °C at 500 Mrad [36]. This allows to deduct the real voltage across the temperature sensors and use the apparent voltage measured by the ADC to correct other voltage measurements accordingly to get an effective TID drift of 1-2% for voltage monitoring. Two resistive temperature sensors with lower resolution are available to measure temperature gradient across the pixel array. These resistive temperature sensors are very narrow to fit on top of the pixel array.

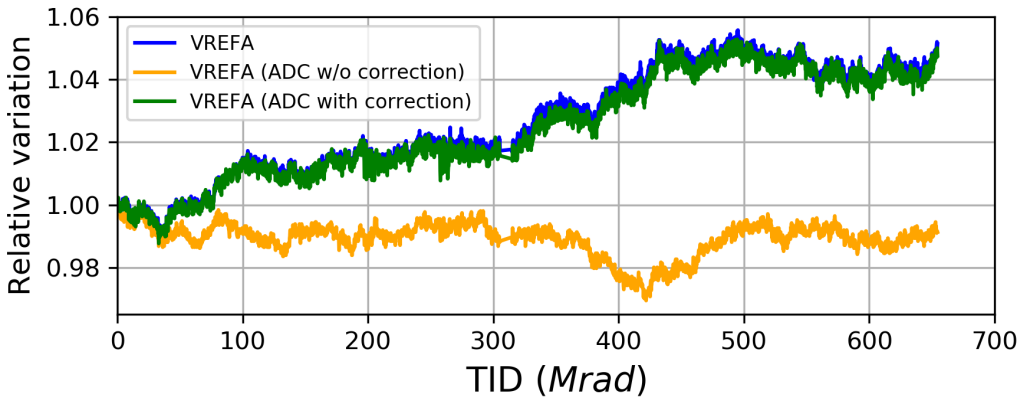


Figure 63. RD53B Monitoring dependency on radiation induced ADC reference drift. When using the ADC to measure radiation drifts in the biasing reference, VREFA, the apparent TID drift seems small (yellow), as both references have similar TID drifts. If VREFA is measured with an external ADC (blue & green) then its real 5 % TID drift at 500 Mrad becomes visible.

Radiation effects monitoring of digital logic is made with a set of digital ring oscillators with different gate types and transistor sizes and for analog transistors by direct analog measurements on a few reference MOS transistors. The ring oscillator frequency is measured using the 40 MHz chip/system clock as reference, from which the effective gate delay can be calculated. Typical gate delays of the implemented ring oscillators can be seen in figure 66 as function of TID.

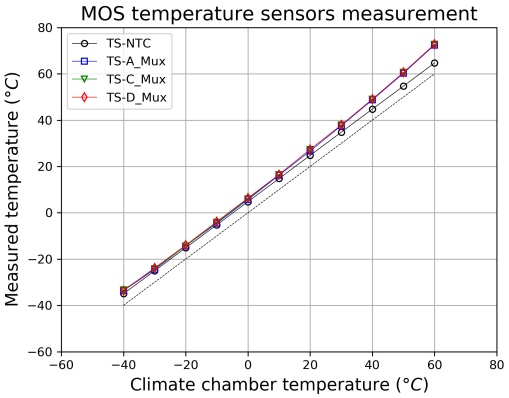


Figure 64. Calibration of 3 (A,C,D) on-chip MOS based temperature sensors in climatic chamber. The NTC thermistor is a reference sensor on the test board. The measured temperature difference between the climatic chamber and the NTC sensor is caused by the pixel chip power dissipation.

9 Radiation tolerance

Achieving the required radiation tolerance of the RD53 pixel chips of 1 Grad for 10 years operation in inner pixel layers at HL-LHC has been a major challenge. The used 65 nm technology has been seen to have excellent radiation tolerance up to dose levels of ~100 Mrad [39]. To reach an effective TID tolerance of up to 1 Grad the RD53 collaboration has invested significant efforts on radiation effects studies with dedicated radiation test chips. Initially it was thought impossible to implement such a complex mixed signal chip with 1 Grad radiation tolerance, as radiation tests showed large transistor and circuit degradation at radiation levels above 100-200 Mrad. With systematic radiation tests of different transistor types and sizes under different irradiation conditions (voltage, temperature, dose rate) indications were found on how to reach 500 Mrad, and potentially higher, radiation tolerance with specific design constraints and under particular operation conditions. Dedicated analog and digital circuit test chips were made to confirm that this seemed viable [41][43]. Finally, it has been confirmed with full sized pixel chips (RD53B/C generations) that 1 Grad is feasible, when using specific design precautions and operation conditions as outlined below.

Wide transistors have excellent radiation tolerance without significant leakage (other similar technologies have been seen to have transistor leakage issues). This has enabled appropriately designed analog circuits to demonstrate excellent radiation tolerance. In the RD53B chips, issues were seen with increased mismatch in critical current mirrors for biasing different parts of the chip. The origin of this was traced to be a x-ray shielding effect from thick top copper routing layers above critical transistors, as was seen to be significantly smaller with proton irradiation. In final RD53C chips it is assured to have the same thick copper routing above critical current biasing transistors.

Narrow transistors show large radiation degradation, with additional detrimental annealing when operated at elevated temperature, especially under specific biasing conditions, as indicated in figure 65. Narrow 65 nm gate length transistors are critical in high density logic, particularly needed in the pixel array logic. Transistors in digital logic are only under worst case biasing conditions during very short signal transitions. Not using the highest density digital library (with name

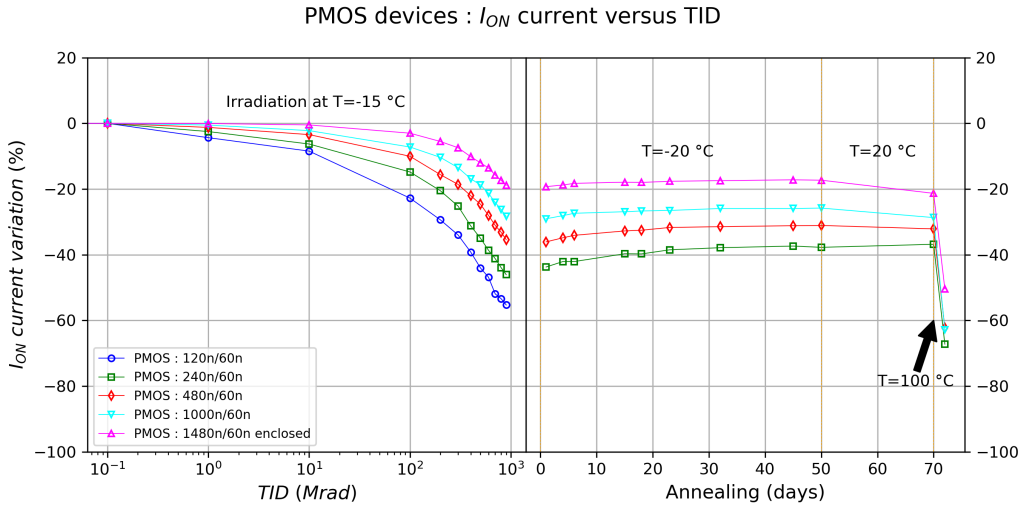


Figure 65. PMOS Ion drive capability radiation degradation of short channel (60 nm digital) transistors of different width under worst case biasing conditions at -15 C° (left) and related detrimental annealing (right) at different temperatures [41].

extension Drive0) but instead the second highest density library (Drive1), with wider transistors, gives significantly improved radiation tolerance. Finally, it was determined that if used cold (below -10 C°) and never (less than a few days) powered at room temperature after high TID exposure, the observed detrimental annealing can be kept under control. The basic mechanism behind this behavior has now been understood [39] and confirmed. It is caused by radiation induced trapped charges in gate spacers that at elevated temperature and under specific biasing conditions drift into active gate regions. Initially the pixel chip operation temperature was estimated to be -20 C°, but has with more detailed thermal modeling of the pixel detectors been seen to be up to -10 C° in certain locations. No significant difference in radiation tolerance has been seen between -20 C° and -10 C° operation temperature.

With these design and operation constraints (below -10 C°), digital logic will after 1 Grad still have a ~50 % speed degradation, when irradiated at high dose rates (1 Grad in 1 week). When irradiated at low dose rates the speed degradation was found to be significantly more. Low dose rate effects have been characterized in dedicated long term X-ray, cobalt source and Kr85 source irradiations [40] with an effective speed reduction of a factor 2-3 for the used digital library, as shown in figure 66. This is taken into account in the RD53 design flow using a specific extreme timing corner case. Initially a dedicated radiation corner for the used digital libraries was developed in RD53. It was then realized that using a very low voltage corner provided by the foundry (0.9 V, worst case process and -40 C°) in practice results in similar timing and have been used for timing closure of final RD53C designs.

Full scale RD53B chips have in low dose rate irradiation tests [49] indicated a projected radiation tolerance to the Grad level as shown in figure 67. Integrated on-chip analog and digital radiation effects monitoring (see section 8) allows this to be monitored during operation and can be used to predict if an inner pixel layer needs to be replaced. The current prediction is that final RD53C

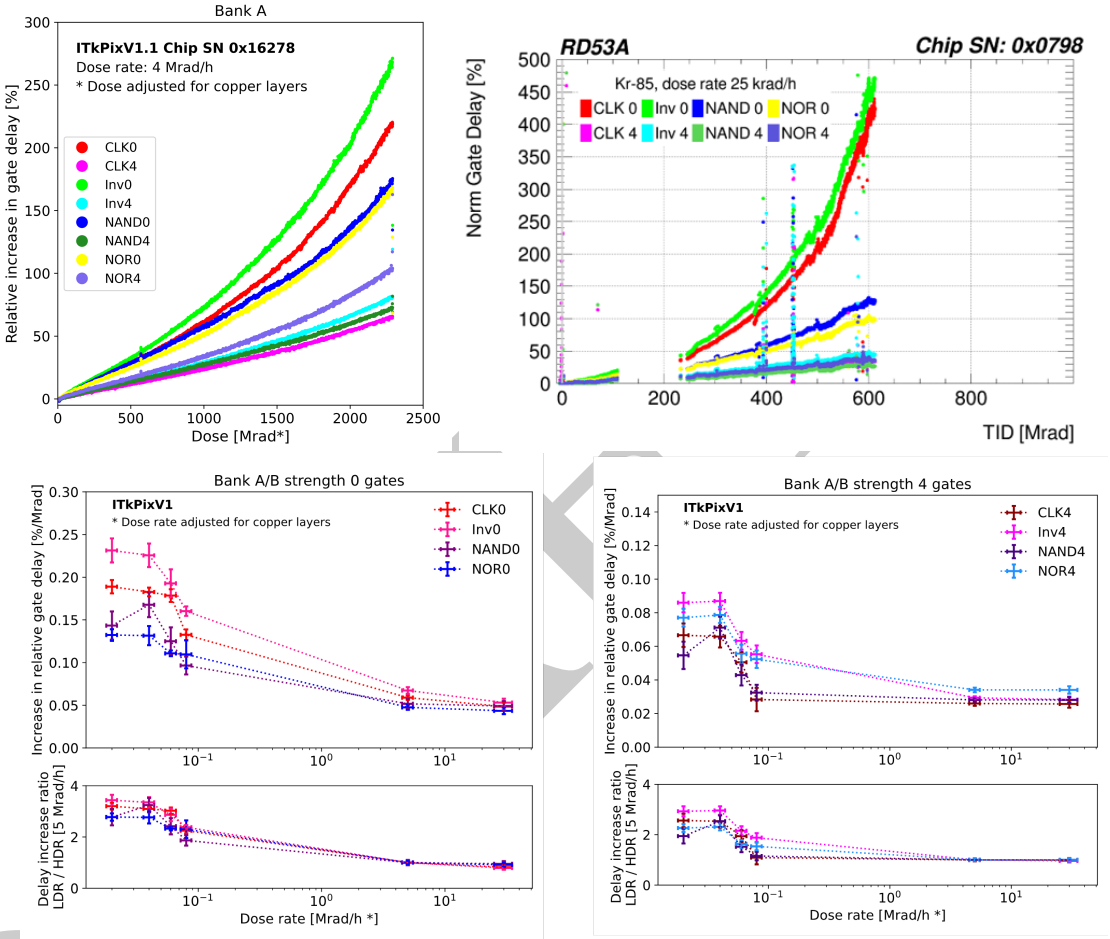


Figure 66. Gate delay degradation of Drive0 and Drive4 gates at high dose rate (upper left) and low dose rate (upper right). Relative delay degradation between Low dose rate and High dose rate at different dose rates. Drive4 uses same transistor width as Drive1, but with multiple parallel output transistors for higher drive capability. They therefore have similar relative radiation degradation.

chips will be capable of taking 1-1.5 Grad over 10 years of operation in an appropriately cooled and operated pixel detector. Pixel assemblies have been exposed to 10^{16} hadrons/cm^2 and have as expected not been seen to be affected by NIEL (Non Ionizing Energy Loss), as it is generally the case for CMOS processes. Final production chips will be produced in the same fab as the prototypes, as there have been indications of differences between fabs of the same technology node. Radiation tolerance tests of production batch samples will be necessary to assure that required radiation tolerance is maintained during the wafer production period. Production batches with indications of reduced radiation tolerance can if needed be used for outer pixel layers (~100 Mrad and majority of chips needed). It is planned to make further irradiation tests of final pixel chips to extremely high levels (multiple Grad) until until they show signs of failing because of TID.

Tolerance to SEU and SET effects is the other critical requirement for a chip with complex digital logic. A RD53 chip is estimated to have ~100 SEU upsets per second in inner pixel layers, based on the measured SEU cross-section of used memory elements. SETs can be assumed to be of the

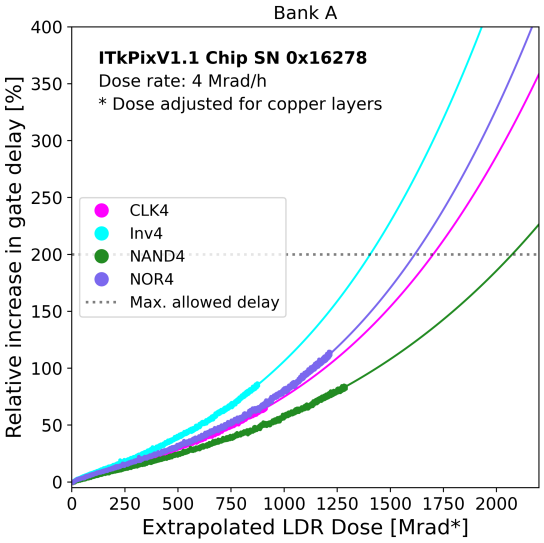


Figure 67. Projected low dose rate limit of RD53 chip based on low dose rate irradiation of Drive4 (and Drive1) gates. Gate delays have been measured with low dose rate radiation characterization of ring oscillator test circuits. A maximum tolerable gate delay increase of 200 % have here been defined as the limit, as this is the effective timing margin obtained with the used gate timing models.

same order of magnitude. This makes it a major challenge to assure sufficiently reliable operation of thousands of chips. Systematic use of well known general TMR (Triple Modular Redundancy) schemes, and related specific tools [44], can resolve this, but at an excessive cost in terms of area and power overhead (factor 3).

In particular in the large and dense pixel array logic it is not feasible to fit TMR protection. Critical pixel configuration bits have triplicated latches, without auto-correction feedback. Continuous pixel re-configuration can be done at a rate of up to 10 times per second with the available control link bandwidth. Protection from SEUs and SETs with TMR in remaining pixel array logic can not fit in the highly constrained area. This has with simulations been estimated to cause fake or or lost hits below 0.01 % of the actual hit rate.

The Digital Chip Bottom (DCB) contains critical chip functions that can not be allowed to be upset by SEU/SET, as the chip may then get into a dead-locked state, lose system synchronization or get mis-configured. Critical functions (global configuration, Trigger table, state machines, buffer pointers and critical event information) have full TMR. However, hit data in data buffers and processing pipelines are not protected. This strategy has resulted in ~25 % of registers in the DCB to have TMR protection. With such a partial protection scheme it is critical not to overlook critical memory elements that require protection. This is a delicate task requiring careful verification with SEU simulations (see section 11). It can also be mentioned that RD53 chips specifically use event tags, included in trigger commands, to prevent event de-synchronization to occur because of SEUs in local chip event ID counters.

TMR of selected registers can be done in different fashions that must be carefully chosen based on

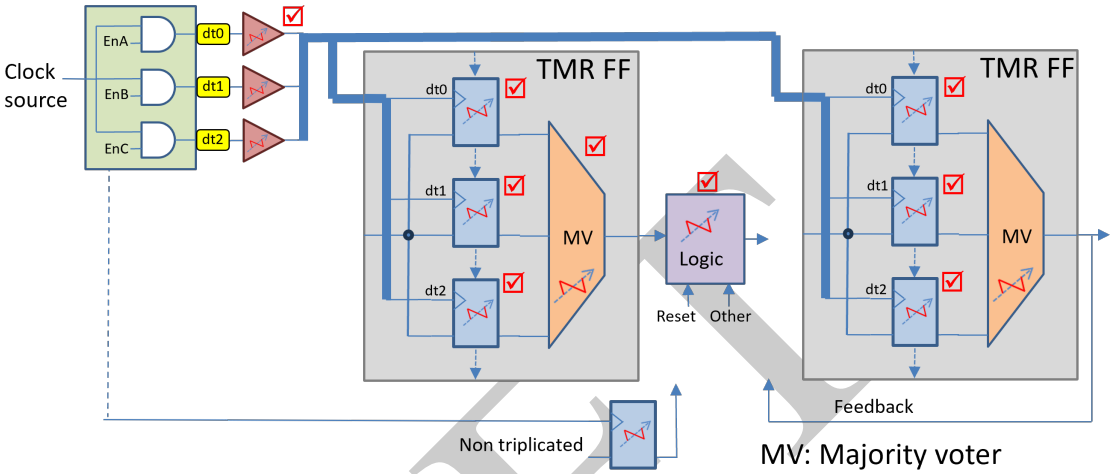


Figure 68. SEU/SET protection used for critical storage nodes in digital chip bottom. Critical flip-flops made with TMR. SETs in logic, MVs (Majority Voter) and clock drivers filtered by triplicated clocks with centralized clock skews (dt0, dt1, dt2) followed by MVs.

the characteristics of the design and how best to integrate this into the chip design flow. The selective TMR protection has been made at gate level, after RTL logic synthesis. Based on register names, with a specific name extension, single Flip-Flops (FF) have been replaced with triplicated flip-flops with Majority Voting (MV). Triplicated clocks are introduced in the design for TMR protected FFs in appropriate clock domains (40 MHz, 64 MHz, 160 MHz, 640 MHz, 1.28 GHz). SET filtering for TMR FFs is obtained with a time skew between the triplicated clocks such that short SET glitches will only be seen by a single TMR FF and then filtered by the TMR majority voter as indicated in figure 68 [47]. This approach does not require triplication of TMR voters and combinatorial logic as it prevents SET glitches to propagate to multiple TMR nodes. Triplicated clock skew was for the RD53B chips set to 300 ps, based on SET glitch width measured with a dedicated test chip [48]. It has been measured that the used partial TMR protection with triplicated skewed clocks has reduce the effective SEU cross section by a factor of 400. The triplicated clock skewing has in final RD53C chips been increased to 400 ps to further diminish the SEU and SET cross-section (see section 10). Quick and efficient production testing of the implemented TMR protection can be done by disabled one by one the triplicated clocks, and check that the chip continues to work correctly.

Extensive SEU/SET tests have been made of the RD53B chips [45][46]. Occurrences of relatively long readout link dropouts, as shown in figure 69, were seen in ion beam tests. It was confirmed in dedicated laser injection tests to be caused by short SET glitches in the biasing generating circuit, being extended to multi microsecond long biasing shifts to the PLL. The cause was confirmed with detailed circuit simulations with a dedicated analog simulation setup for SET/SEU sensitivity analysis. Biasing circuit topology changes have been implemented, based on detailed SET simulations at transistor level, to resolve this in final production chips. A critical issue with the chip event readout getting stuck was also identified and resolved based on extensive SEU verification simulations (see section 11).

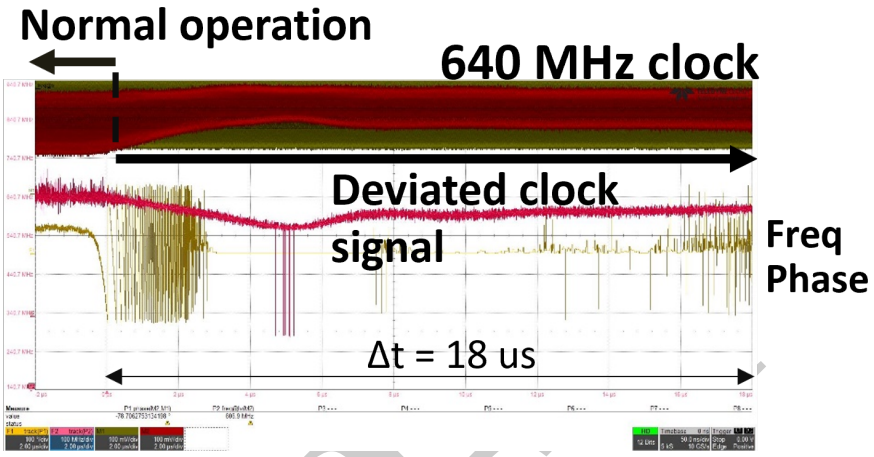


Figure 69. Measured link dropouts of RD53B chip in ion test beam. A sudden PLL frequency jump occurs at the time of the induced SET and it takes 18 us for the biasing and the PLL control loop to recover from this. On the upper trace it can be noticed that the signal amplitude of the serial output is also affected, as the serial link driver is also affected by the SET in the biasing circuit. It can be noticed that the driver output recovers much slower than the PLL, as the PLL control loop actively compensates for the long/slow biasing shift.

The critical (and sensitive) PLL has been implemented in full custom layout with triplicated counters. During its normal operation it uses a simple bang-bang phase detector, where occasional SEUs and SETs in the phase detector can only introduce very small jitter (few ps).

Recent ion and proton beam tests with the RD53C chip have confirmed that final production chips have significantly lower SEU and SET sensitivity. Link dropouts are not observed any more. When actively processing high hit and trigger rates, the RD53C chip is seen to have a HEH (High Energy Hadron) cross section a factor ~30 better than the RD53B. The effective HEH cross section for event readout getting stuck has recently been measured to be lower than $3\,10^{-13}\,\mathrm{cm}^{-2}$. In inner pixel layers this will correspond to a pixel chip running for an average period of ~1 hour between having readout issues. This is considered acceptable for a small number of inner layer pixel chips in such a hostile radiation environment. Regular system level fast buffer clear commands can be issued at rates as high as several Hz, without significant system dead-time, when having a DAQ system that can handle this appropriately. Further system level studies are ongoing to determine how the DAQ and control systems of the experiments can handle this efficiently.

An unexpected small number of Single Event Latchups (SEL) were observed in the digital part of the RD53B chip at an increased supply voltage of 1.3 V in a dedicated ion test at highest Linear Energy Loss (LET) and 45° incidence angle with an equivalent LET_{eff} of 88 MeV.cm^2/mg. This has not been seen before in the used 65 nm CMOS technology and came as a surprise as a digital library with substrate and well taps in each gate has specifically been used to avoid possible latchup issues. It has been verified that events with so high LET_{eff} will not occur in practice in the HL-LHC environment (e.g silicon recoils from nuclear reactions). It has also been verified that SEL is not seen in the ion beam when the digital logic is powered at its nominal voltage of 1.2 V

(it is known that SEL is very voltage dependent at low power supply voltages). If such a latchup would exceptionally occur in final systems, it is not expected to cause permanent chip damage in a serially powered system, driven by a constant and limited current.

10 Implementation

RD53 chips are implemented in a 65 nm CMOS technology with the maximum allowed metal stack consisting of 7 thin, 1 thick and 1 ultra-thick metal layers, and an additional top redistribution layer also used for power distribution where appropriate. The general floor plan is as shown in figure 9, and consists of the large pixel matrix of 150 k 50×50 μm^2 pixels and the Digital Chip Bottom (DCB), the Analog Chip Bottom (ACB) and the IO pad frame with SLDO power regulators.

The pixel array is assembled from 8×8 pixel cores including sixteen analog islands of 2×2 front-ends embedded in a sea of digital logic as shown in figure 70. Analog and digital circuits are implemented in separate triple wells to assure best possible noise isolation. Pixel cores have embedded power, analog and digital signal routing to make pixel core columns from abutment of pixel cores, and from this build the complete pixel array. Pixel cores have built-in digital signal buffers and skew compensation for time critical signals (clock and calibration injection) to guarantee a max time skew across the array of 1 ns (~2 ns after 1 Grad). Skew compensation is implemented in the pixel cores as shown on the right of figure 70 with a configurable delay before the local clock distribution network. The configurable skew compensation delay is driven by the pixel core address (defined by location) to get well aligned pixel clocks for different process, temperature and voltage corners as show in figure 71. It can be noticed that the skew compensation delay is effectively adjusted for each four pixel cores. Pixel core logic has been synthesized with appropriate conservative constraints to build a functional pixel core column without timing constraints violations. A timing model of the pixel core has been extracted by the Cadence Liberty tool to assure accurate pixel array timing used for full chip assembly and verification. The pixel core has also been extensively simulated at analog level to assure the best possible verification of AFEs together with the digital pixel logic (see section 11).

Logic synthesis, place & route and timing optimization of the DCB, with pre-placed analog blocks in the ACB, has been performed with the pixel core timing model with dedicated conservative process, voltage and temperature corners to get the required TID radiation tolerance. As part of the design flow, triplication of critical FFs is performed at the end of logic synthesis with dedicated triplicated clocks with strict timing constraints (and time skew as mentioned in section 9). Placement is enforced to keep a minimum spacing between TMRed FFs to prevent correlated multi-bit SEU upsets to disturb correct function of the chip. A histogram of final TMR FF distance is shown in figure 72, where it can be seen to have complied to the minimum distance constraint of 15 μm. 15 μm distance has by the HEP electronics community been seen to be give good assurance that multi bit flips will be unlikely. Figure 73 shows that an average TMR clock skewing of 400 ps has been obtained in the final RD53C chips (was 300 ps in RD53B chips). For a small number of FFs the effective local clock skew between the three TMR FFs are just below 200 ps (typical case process) which is considered acceptable. It was attempted to get the place and route tools to

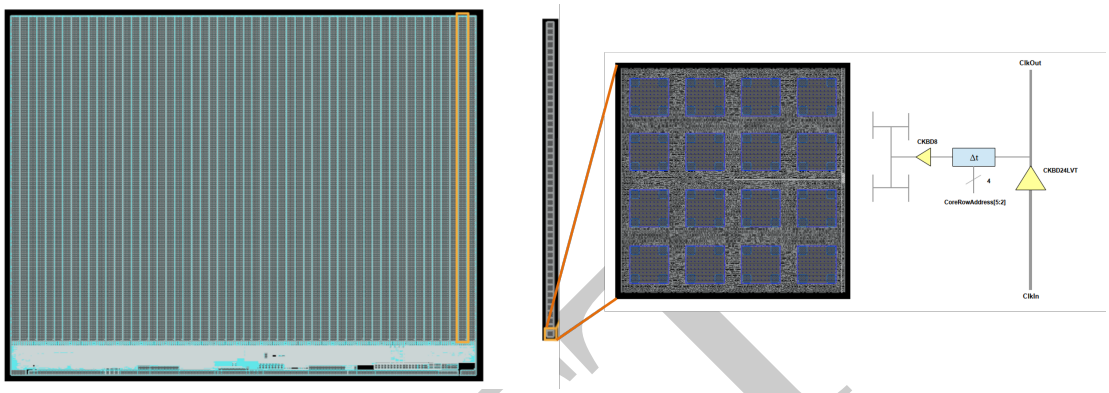


Figure 70. Pixel chip implementation with physical hierarchy used to build the pixel array. The pixel array is assembled from pixel core columns made from pixel cores with 8×8 pixels. Clock skew compensation along the pixel core column is built into the pixel cores, with its local delay determined by position along the column.

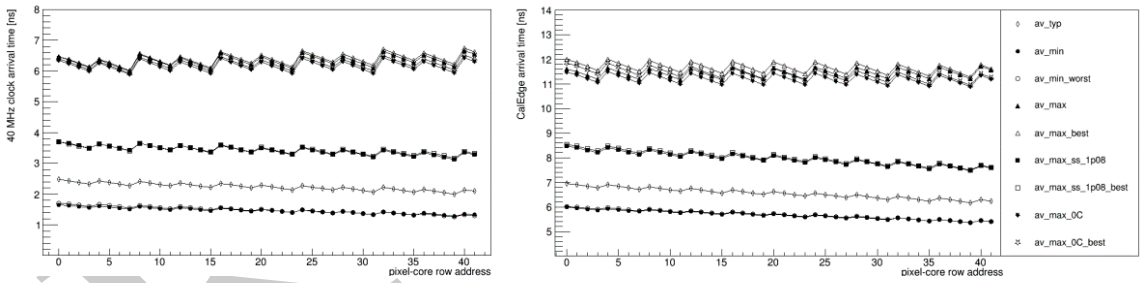


Figure 71. Time skew along pixel column for clock (left) and pixel charge injection (right) for different process corners.

reduce the tails of the clock skew distribution but convergence was not obtained after several days of running and the tools eventually crashed.

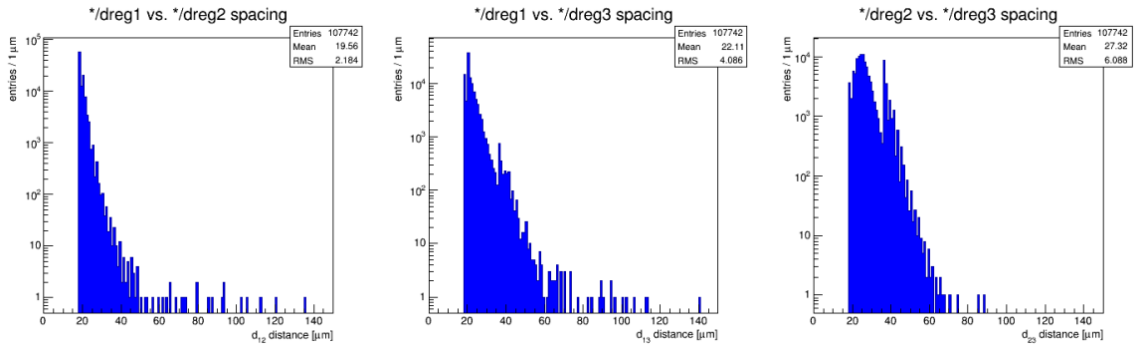


Figure 72. Distributions of distance between pairs of triplicated TMR flip-flops in final RD53C chip.

Power distribution in such a large complex mixed signal chip is critical and has been verified with dedicated Voltus power simulations as shown in figure 74. Detailed gate level simulations are required to drive dynamic power distribution verification but the RD53 chip is so large and complex

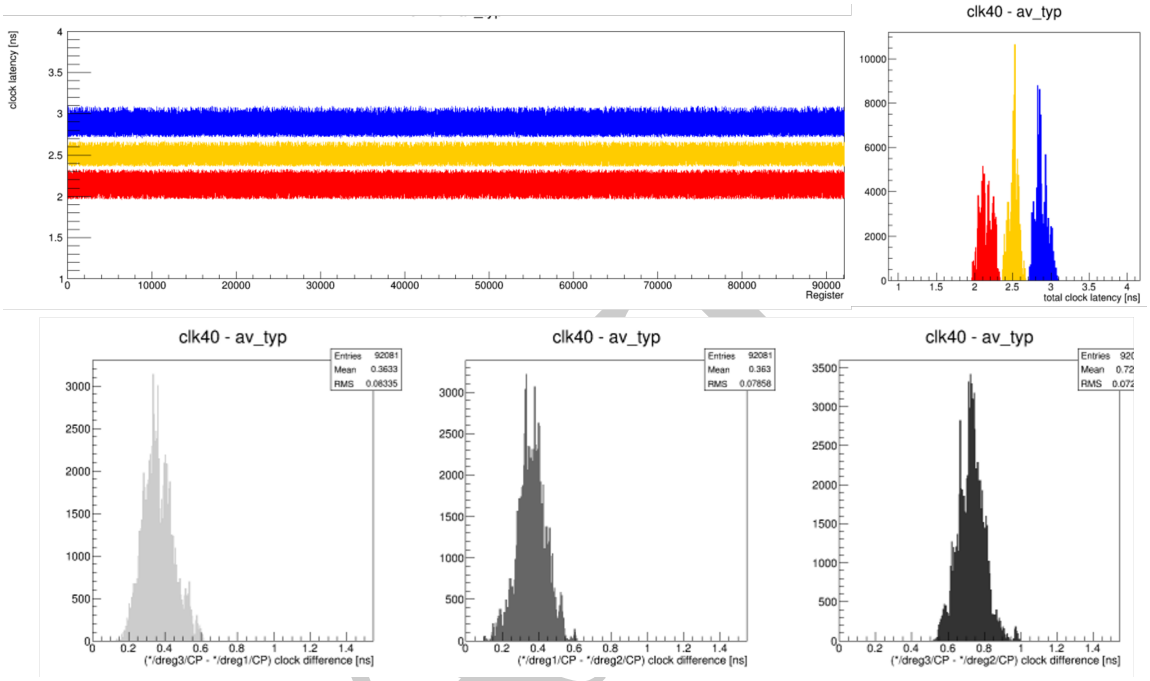


Figure 73. TMR clock skew in the DCB of the final RD53C chip at typical process. Upper: Clock phase of the three local triplicated clocks (red, yellow and blue), across 92 k TMR instances in the DCB. Lower: Histograms of relative phase difference between clock pairs: clk3 - clk1, clk1 - clk2, clk3 - clk2.

that the available tools for this crashed. Dynamic power verification has therefore been made as a combination of a detailed single core column simulation and full chip verification using a simplified core column. This is seen to be compatible with the observed RD53B ground (and VDD) voltage drop measurements as shown in figure 60. All metal layers have been used to get the best possible power distribution, so it is in practice not possible to improve this. The full sized pixel chips with this passive on-chip power distribution have demonstrated that they meet all requirements.

The wire-bonding pad-frame, with 100 μm pitch, is identical for all RD53B and RD53C generation chips, enabling the use of common testing infrastructure consisting of single chip test cards and wafer probing cards. A large majority of the wire-bonding pads are used to supply power to the chip and have low inductance connections to external decoupling capacitors, as shown in figure 75.

The final RD53 chip implementations contain 660 M transistors, 56 M standard cells and 12 M memory elements. 2.1 M memory elements are used to implement 700 k TMR protected bits, of which 85 % are pixel configuration bits and 15 % are used in the DCB. Overall for the complete design 7 % of logical bits have TMR protection.

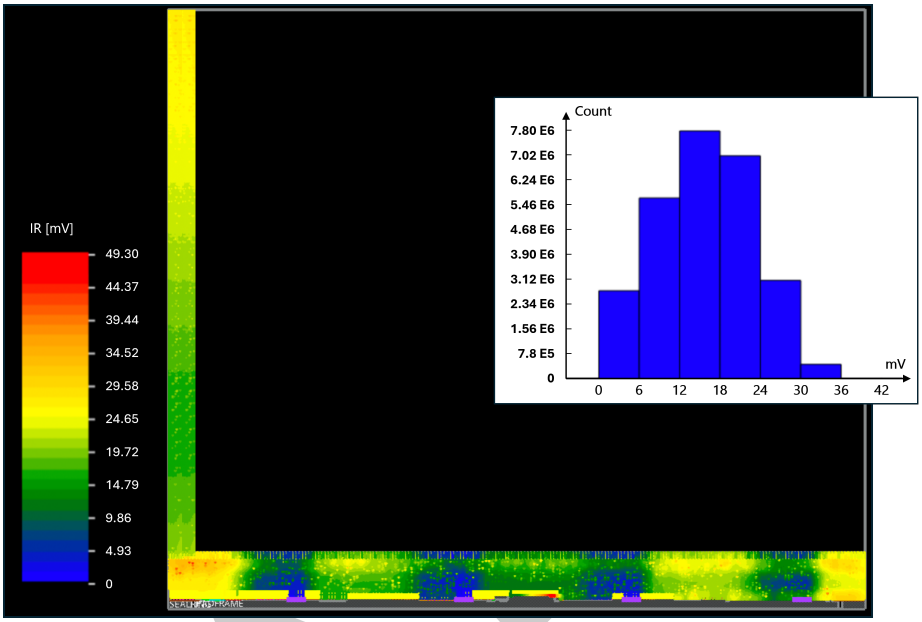


Figure 74. Dynamic Voltus IR drop analysis along pixel array column and across chip. Color coding shows voltage drop in on-chip power and ground distribution network along a single pixel core column on the left (with inserted histogram) and across the chip with the distributed SLDO power regulators. A power - ground voltage drop of up to 40 mV along a pixel core column has in AFE simulations been seen to be acceptable.

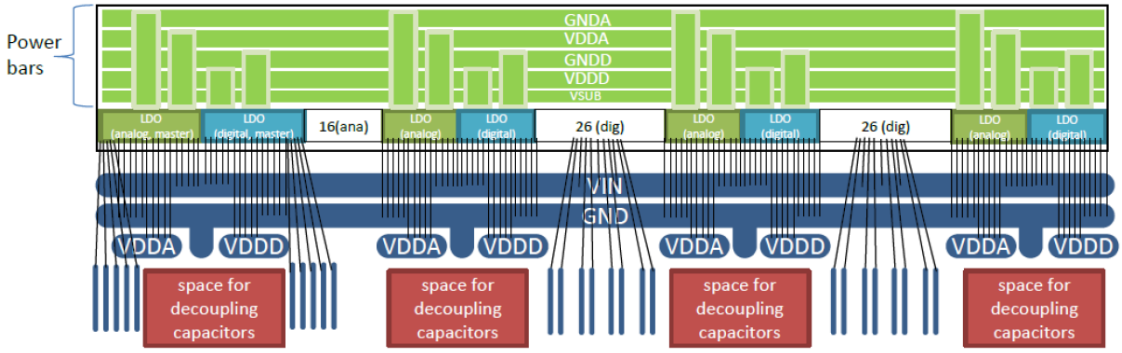


Figure 75. Wire-bonding of RD53 chip with large majority of wire-bonding used for input powering and low inductance connections to decoupling caps.

11 Verification

Such a large complex mixed signal chip, for use in an extremely hostile radiation environment, requires major efforts to perform verification of architecture, performance, functions, mixed signal behavior and SEU/SET tolerance. This must be done over different process, radiation, temperature and voltage conditions, followed by thorough testing and characterization of silicon. Verification has required a large effort and requires specific verification expertise and appropriate verification tools. This is the well known increasing verification challenge in today's chip industry, and RD53 chips have the additional complication of radiation effects (TID + SEU/SET).

An initial System Verilog chip simulation and verification framework was implemented [50][51] to develop and optimize an appropriate architecture with the required performance. This was extensively used to gradually develop a fine grained architecture and optimize RTL (Register Transfer Level) code to insure that the design fits within the available area and with acceptable power consumption. A particular critical part of this optimization is fitting sufficient trigger latency buffering in the pixels. Two alternative architectures were evaluated, verified and tested down to gate level implementation [13][12]. A highly optimized region based time tagged latency buffering scheme with 8 buffer locations was chosen for the final chips and qualified for the highest hit rates (and SEU), as shown from simulations in figure 43. Fitting one buffer location less gives unacceptable hit losses and an additional buffer location simply does not fit in the high density pixel array. This vital simulation and verification framework was gradually extended and extensively used over several years to get to a highly optimized design. It was also used for initial SEU simulations to determine an appropriate design approach, with only a small part of the design using TMR protection for area and power consumption reasons.

For the final RD53C generation chips it was initially planned to expand the existing simulation and verification framework. This was in practice found to be difficult, as it had not been optimized for this, and because the original designers of this framework were no longer available. A new UVM (Universal Verification Methodology) functional verification framework was implemented as shown in figure 76, by a new collaboration member with the necessary expertise in functional verification. This has been a considerable effort over several years that has identified several delicate and critical design issues, but also overlooked a few minor non-critical issues.

The RD53 data flow includes multiple clock domain crossings (40 MHz, 64 MHz, 160 MHz, 640 MHz, 1,28 GHz) and uses Clock Domain Crossing (CDC) FIFOs that have been specifically verified with formal verification tools and with detailed gate level simulations for possible meta-stability issues.

RD53 pixel chips have many delicate and complex mixed signal functions with ~50 % of chip area used for analog functions (AFEs, SLDO powering, biasing, DAC, ADC, monitoring sensors, PLL). This has required extensive dedicated analog and mixed signal verifications to be made. Analog blocks have been developed and submitted in small test chips that have been extensively tested. Digital simulation models of analog blocks have been developed and verified to enable realistic full chip digital simulations and verifications to be made. It was attempted to run mixed signal simulations (analog simulated by fast Spice together with digital at RTL or gate level) but this was found to be too slow and memory consuming for such a large and complex design. An alternative, and very effective, mixed signal verification scheme was used. Full digital chip simulations (at both RTL and gate level) were run and digital interface signals driving the analog blocks were stored as VCD (Value Change Dump) waveform files. These were then used as stimuli for detailed spice simulations of analog blocks to verify their correct function and correspondence with the digital simulation models used for full chip verification. This has ensured that interface issues between digital and analog have been identified and corrected. An example of results from such a digitally driven mixed signal verification simulation of an 8×8 pixel core is shown in figure 77.

Verification of appropriate SEU and SET protection, as described in section 9, is a delicate and

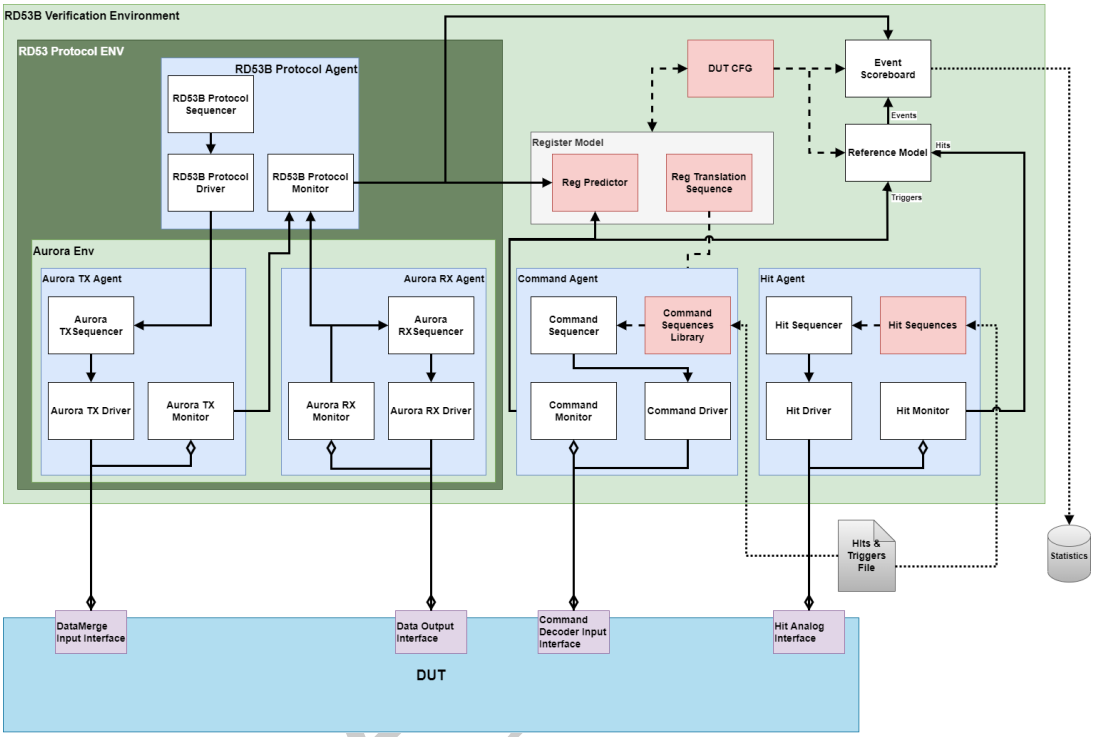


Figure 76. UVM based functional verification framework used for final production chips. UVM uses standardized signal generators (e.g. pixel hit generators) and output interfaces to be verified (e.g. readout). The functional verification is based on a high level reference model (e.g. for triggered hits being readout). Such a parameterized verification framework is run across a large set (several thousands) randomized chip configurations with a large number (millions) of hits and triggers. Final full chip verification typically takes several weeks of continuous simulations on a cluster of high performance workstations.

critical task. Partial SEU/SET protection requires extensive SEU/SET simulations to verify that the chip can not get stuck or lose system synchronization. It must be verified that the non protected parts of the chip can not induce unacceptable hit and event losses. Critical cases, of non TMR protected nodes, have been identified with extensive SEU simulations at RTL level (fast simulations) and have enabled final designs to have significantly improved SEU tolerance. Implementation of TMR registers have been verified with gate level simulations (slow). SEU/SET issues seen in RD53B tests with ion and protons beams [45], have only been possible to identify and resolve in RD53C chips with extensive SEU/SET simulations to identify the complex cause of these. An example of this was a very hard to find issue with a SEU vulnerability in the unprotected pixel array logic when transferring triggered hits out of the pixel array. In this specific case a core column readout controller could get stuck if a SEU occurred in a specific time window. The cause of this was hard to identify as only visible in simulations in a rare specific case when the SEU occur together with a specific hit pattern in a narrow time window and only detectable in chip readout output after multiple events have been processed. It was not possible to fit TMR in the relevant part of the pixel array logic, but a minor change to the readout controller prevents the chip to get stuck.

SET sensitivity has been verified with gate level simulations to confirm that short SET glitches

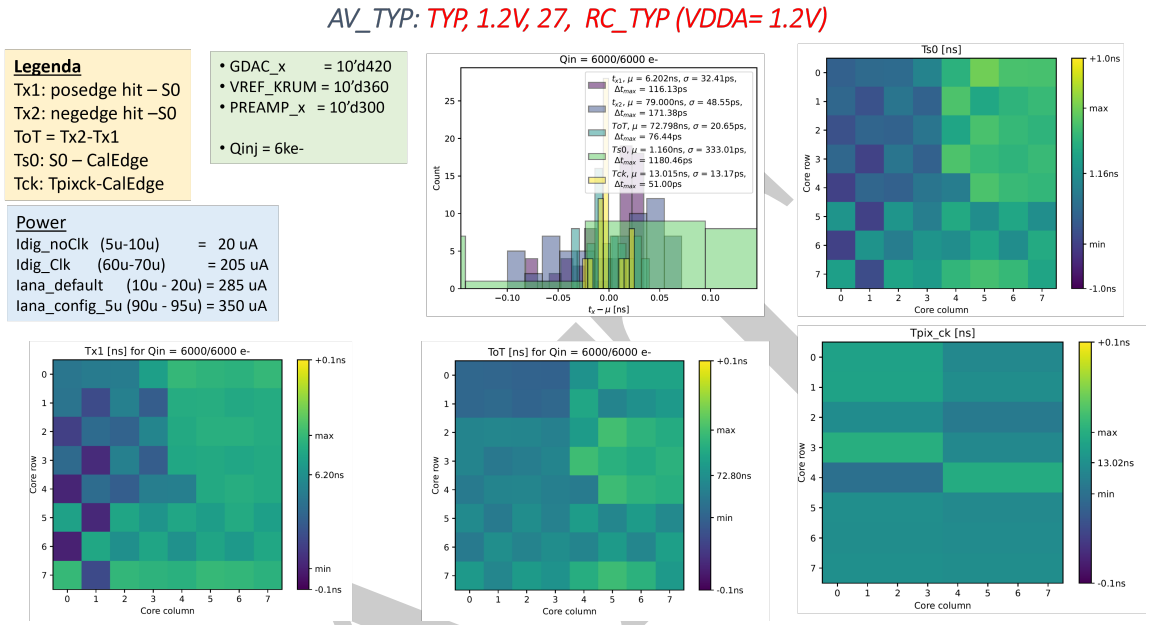


Figure 77. Summary sheet of detailed 8×8 pixel core mixed signal simulation/verification results for typical process conditions. Summary sheet shows AFE biasing together with observed pixel power consumption and timing spread across pixels (TOA, TOT, sampling clock, and injection pulse).

are filtered (and that large SET pulses disturb chip functionality as expected). SET simulations at gate level are extremely time consuming and were therefore focused on known critical parts of the chip. Several SEU/SET vulnerability issues have been identified and fixed and final chips have significantly improved SEU/SET immunity. Two simulation cases are shown in figure 78 where it can be seen that the chip readout gets blocked from an SEU and in the improved case where readout continues as normal. With the implemented and simulated improvements, it can be estimated that inner layer chips will run reliably for hours. If a chip gets stuck or out of synchronization, it has been verified in simulations that a dedicated fast clear command can recover normal functionality with short dead-time. The effective dead-time of the chip from a fast clear command is determined by the clearing of the hits in the latency buffers and clearing of pending readout data in readout and processing FIFOs. At system level, the fast clear command can be used systematically at regular intervals if needed, with very small dead time from the pixel chip, but the DAQ and control system must be prepared and optimized for this.

The general progression of RTL code together with the initial simulation framework and final verification framework is shown in figure 79 with number of reported and fixed bugs for the different chip submissions. Bugs found are from both testing of previous generation chips and issues found with extended verification simulations. It must be mentioned that the majority of issues have been found and corrected with simulations in the verification frameworks, whereas more delicate mixed signal issues have been found in chip testing.

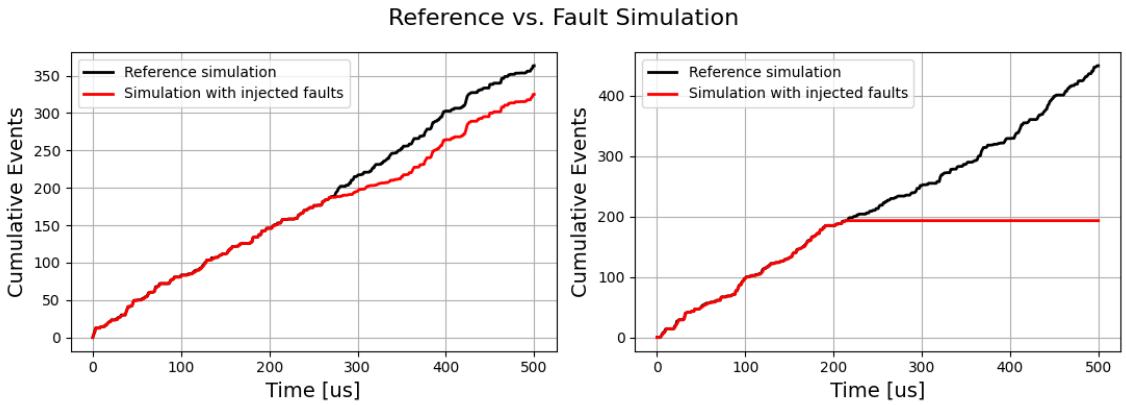


Figure 78. SEU simulation comparing reference simulation (black) with simulation with SEU/SET injections (red) with a SEU injection acceleration of a factor 1 million compared to real SEU conditions. Left: Simulation with SEU upsets provoking loss of hits/events for a limited time, but self recovers. Right: SEU simulation where chip gets stuck because of single SEU in pixel array provoking pixel core column readout to get stuck, until issuing fast buffer clear.

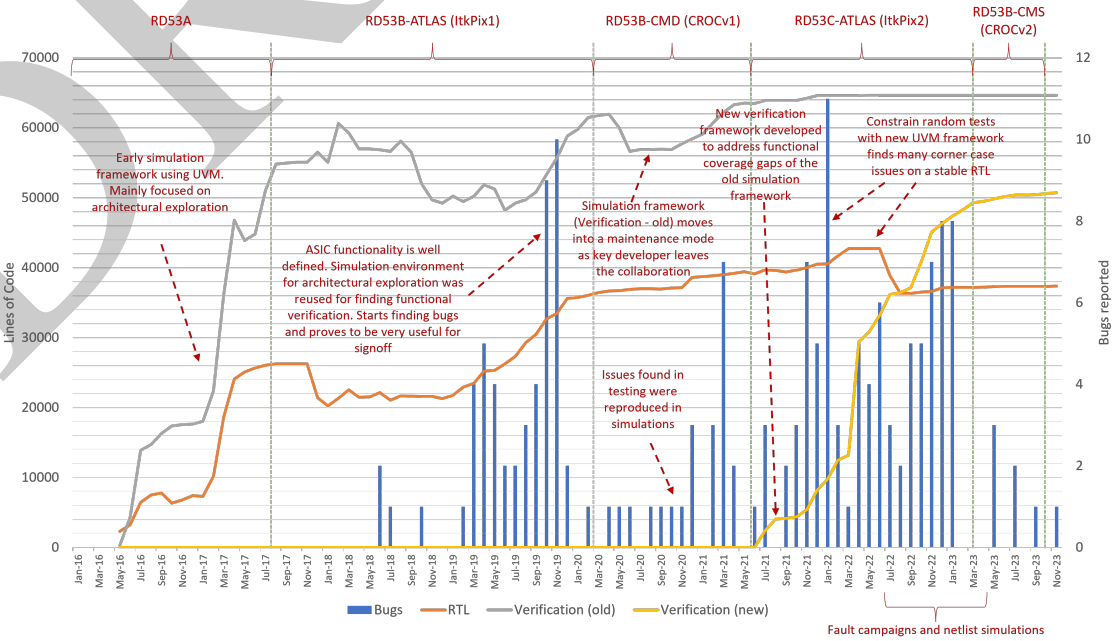


Figure 79. History of RTL design code, simulation and verification code together with reported issues/bugs over different chip generations.

12 Test and characterization

To assure appropriate testability of such large sized mixed signals chips, DFT (Design For Testability) has been implemented in the DCB with full scan path test capability. For the pixel array it has not been possible to fit a scan path. The DCB scan chain includes specific test ports to the pixel core columns to be capable of controlling and observing these if required. Characterization and production tests have shown that the pixel array can be sufficiently tested using parallel analog and digital hit injections with triggered readout. Specific and critical analog and digital signals can be observed via configurable analog and digital multiplexers to dedicated analog and digital test pins.

Several test and DAQ systems (BDAQ53 [52], YARR [53], FELIX [54], CMS-DAQ [55]) have been developed in RD53 and in ATLAS and CMS for chip characterization, wafer level testing, module testing, and system integration tests. They have been used to produce the measurement results presented in this paper. Initial test systems have been verified against chip RTL code with Cocotb [57] integration with virtual digital hit injection to debug test system firmware and software before physical chips became available (e.g. using a RD53B logo hit mask as shown in figure 80).

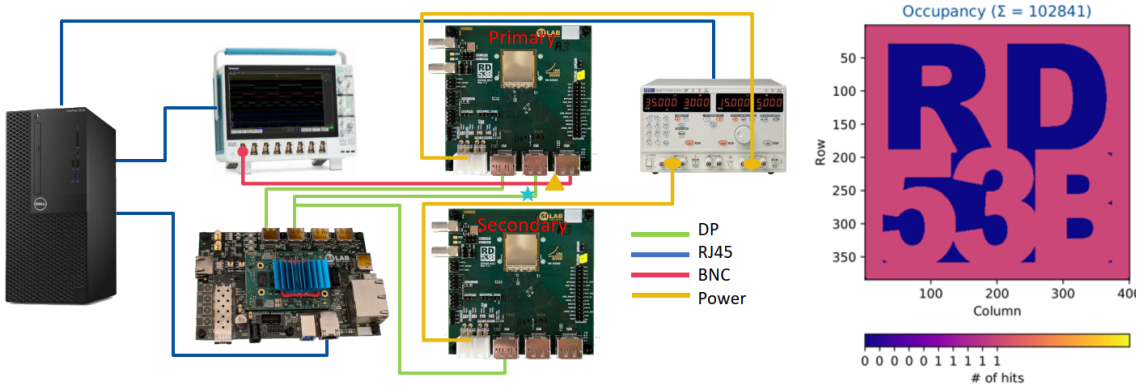


Figure 80. BDAQ53 test and characterization system [52], based on a compact custom FPGA card (lower left), for single and multi chip testing (two single chip test cards shown in the middle labled Primary and Secondary). The same test setup is used with a dedicated needle card for wafer level production testing. Right: Example of virtual RD53 logo hit mask used for software and firmware verification before having silicon chips.

Wafer level testing routines have been developed in the two experiments [58][59] to perform production testing of chips needed in the final systems (30 k chips for CMS + 60 k chips for ATLAS = 90 k chips = ~1000 wafers). These routines extract more than 50 different performance criteria measurements covering digital, analog and powering with examples shown in figure 81 and 82. Under the assumption that a few pixels per chip can be allowed to be non functional (will be disabled), the overall yield has been seen to be as good as ~90 % in 50 pre-production wafers as shown in figure 83. A breakdown of typical chip rejects per wafer, for different tests, is shown in figure 84. Extracted performance and calibration parameters are stored in appropriate databases to be used for configuration and cross checking with pixel module and system tests. Individual chips are traced during dicing, handling and mounting on pixel modules with a chip ID burned in E-fuses during wafer probing. This chip ID can though not be guaranteed to be readable after irradiation

(~100 Mrad). A complete test and characterization of a wafer with 131 chips takes 8 - 16 hours, giving a wafer throughput of 1-2 wafers per day per probing station. Wafer probing facilities have been set up in institutes of each experiment to test all production chips within a year.

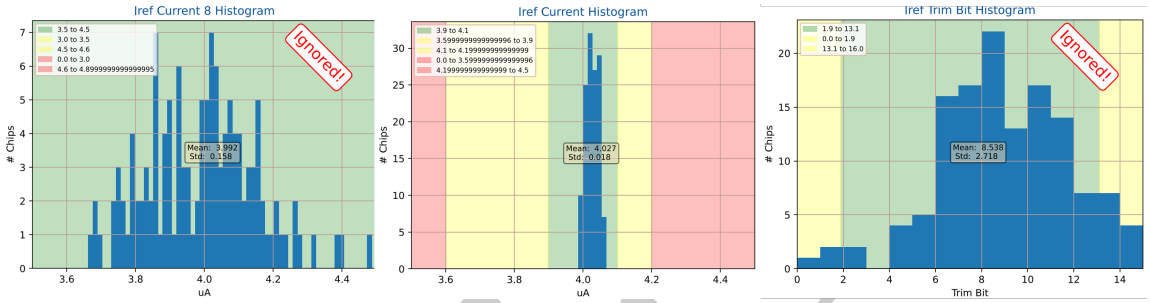


Figure 81. Example of wafer testing selecting criteria for measured reference current (Iref) and its tuning. Left: untuned reference current, Middle: tuned reference current, Right: Iref tuning distribution. Green is acceptance criteria for pixel module production, yellow is acceptance for prototype pixel modules and red is chip reject.

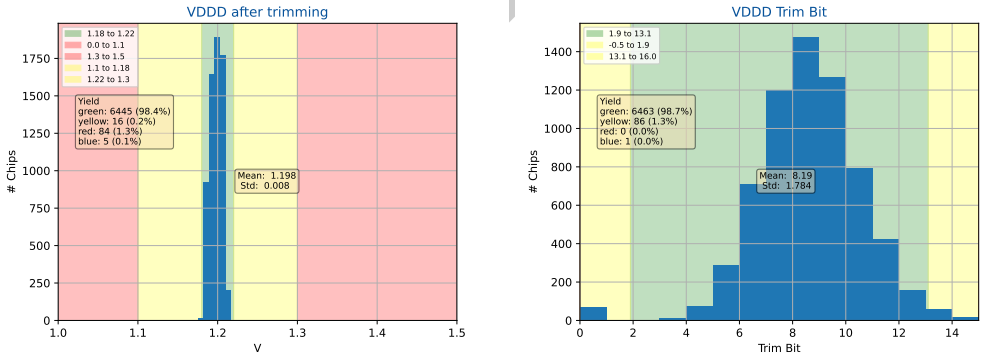


Figure 82. Wafer level testing of tuned SLDO output voltage (VDDD). Right: Used VDDD trimming bits to obtain narrow VDDD voltage distribution among chips on 50 wafers.

Pixel module tests with different bump-bonded sensors (planar, 3D, 50×50 μm^2, 25×100 μm^2) and different module configurations (single, dual, quad chips) have been performed by the ATLAS and CMS pixel detector groups in test beams and with radioactive sources with fully satisfactory results. Pixel module integration issues related to bump-bonding yield and micro cracks from thinned chip dicing, close to the active chip circuits, have been encountered. This will be resolved with improved chip dicing procedures (e.g. laser grooving followed by saw dicing), improved procedures for bump bonding and improved pixel module production and quality assurance procedures (e.g. gluing with radiation hard glues).

A measured gamma ray spectrum of an Am-241 source is shown in figure 85. Finally a clear beam spot can be seen in figure 86 from a triggered proton beam test together with a X-ray tomography of a quad pixel module, based on detected pixel hits in the module itself.

System integration tests with serial powering and concurrent readout, as shown in figure 6 have

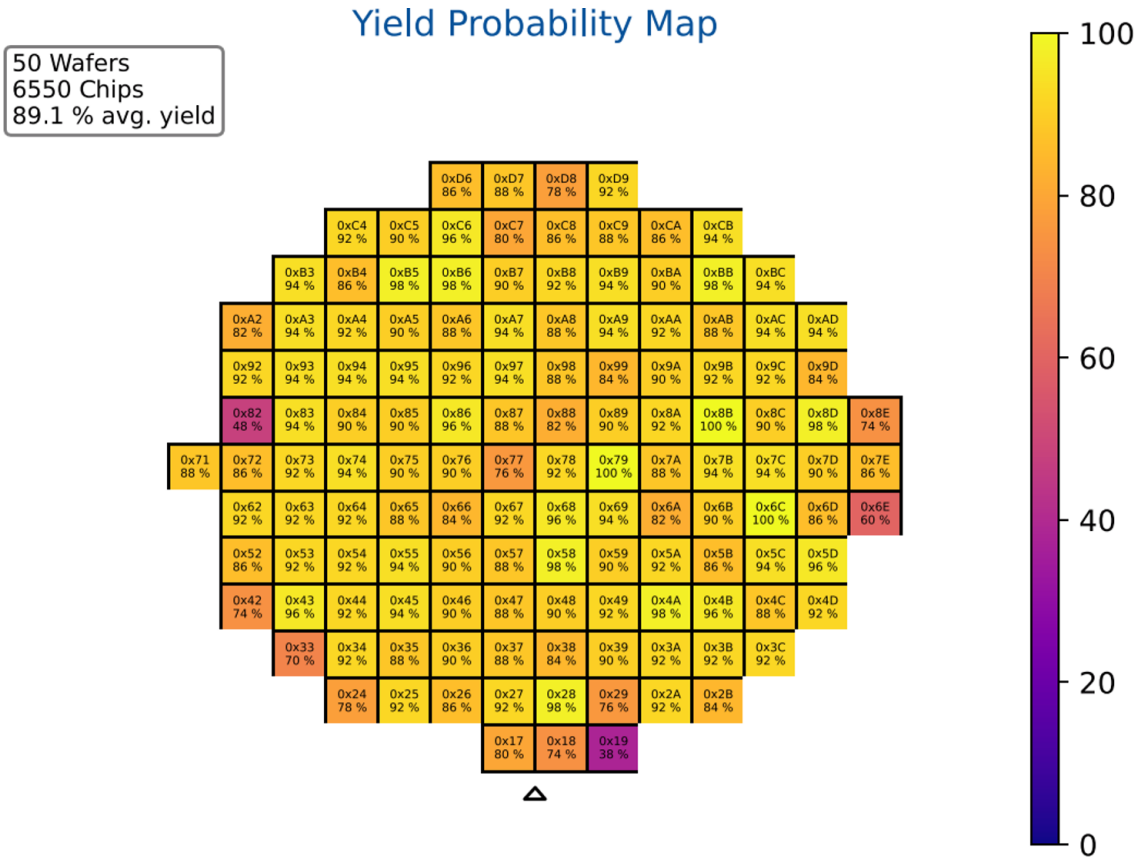


Figure 83. RD53C-ATLAS wafer yield map for 50 wafers.

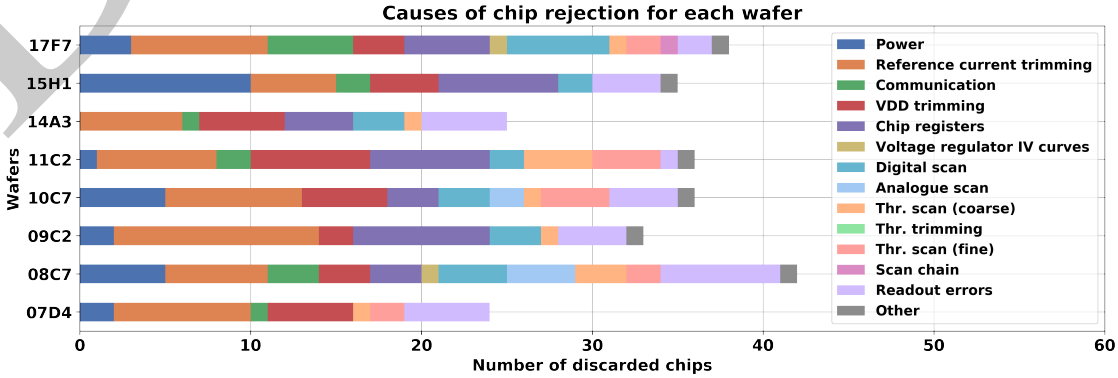


Figure 84. RD53B-CMS chip rejects for 8 wafers for different tests. To be noted that failing chips are often rejected by multiple tests, so typically only 10-20 chip rejects per wafer. Certain test failures also excludes other tests to be performed (e.g. power supply short).

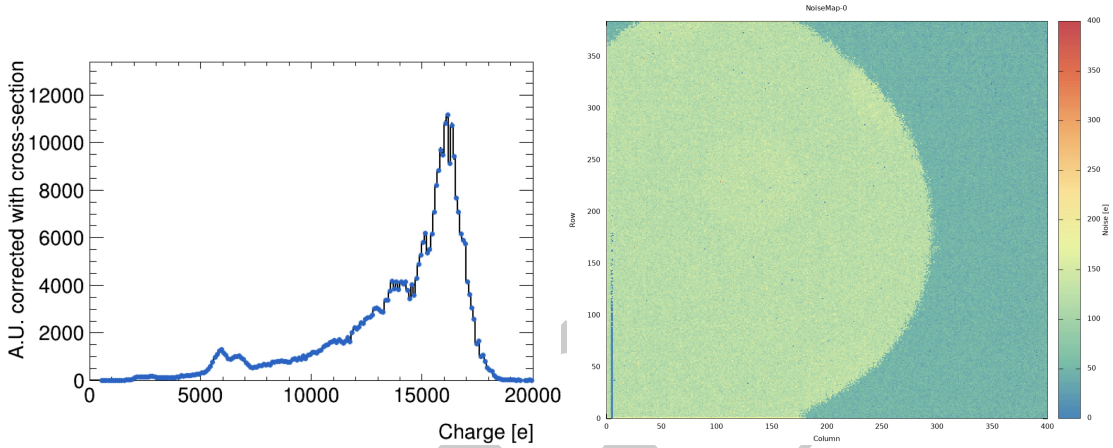


Figure 85. Am-241 gamma source test of RD53B-ATLAS partially bump-bonded single chip module. Left: Spectrum measured with the high precision TOT option and threshold of 1000 e. Right: Hit map of partially bump-bonded module with Sintef 3D sensor. Unconnected pixel bumps clearly seen as low hit count pixels (blue) with a few noise hits.

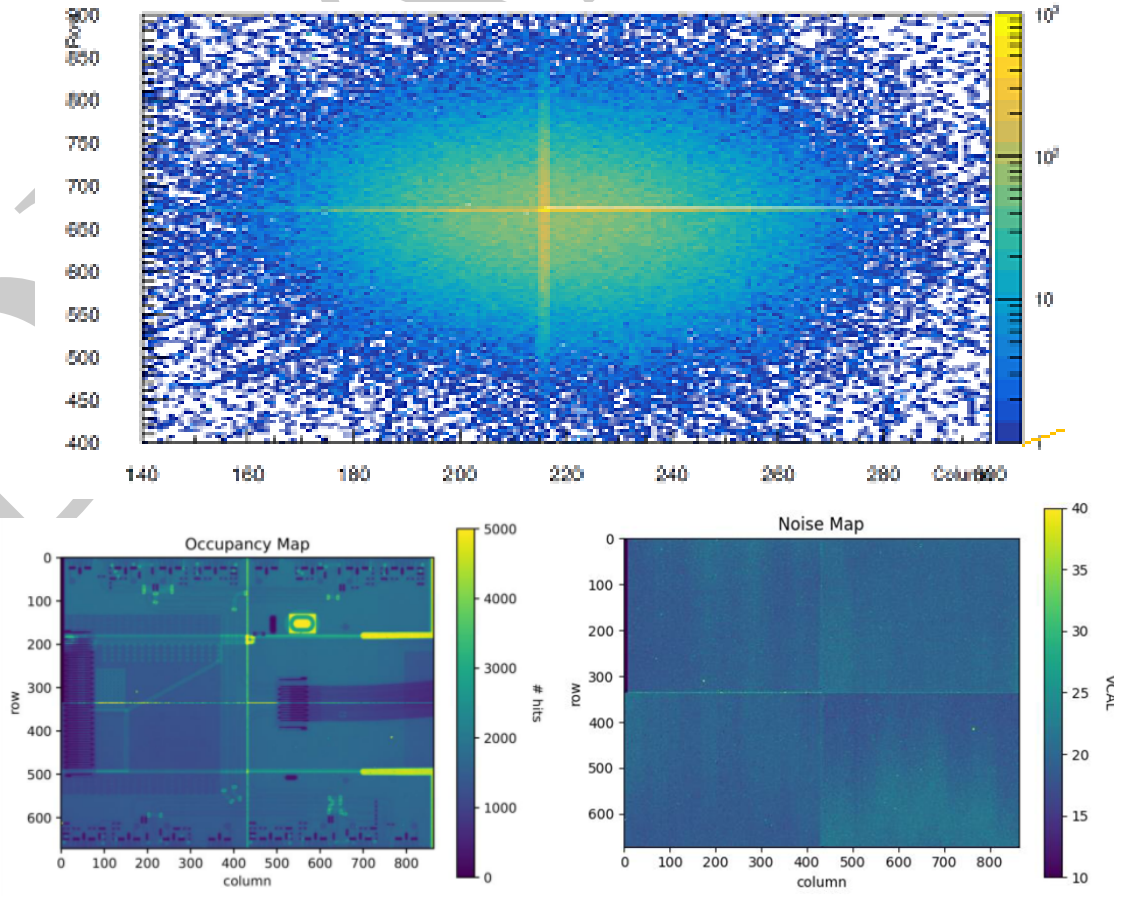


Figure 86. RD53B-CMS quad pixel module tests. Upper: Triggered proton beam profile with large pixels between chips clearly visible. Lower left: X-ray tomography with X-ray hit count as registered by quad chip pixel module. Lower right: 2D noise map of same quad chip pixel module.

demonstrated fully satisfactory chip, module and system performance for use in the pixel detectors of the two experiments. Extensive pixel module and system tests will continue to be made in the coming year in the two experiments with RD53C production chips.

13 Conclusions

The RD53 collaboration has over the last 10 years successfully developed two large complex mixed signal hybrid pixel detector chips for use in the ATLAS and CMS HL-LHC upgrades. It has been a major challenge to assure required lifetime (TID) and reliability (SEU/SET) for such an unprecedented hostile radiation environment a few cm from the interaction points at the heart of the ATLAS and CMS experiments. A novel serial powering concept has been developed for the on-chip power regulator that has been qualified and verified at system level for low noise use with up to 64 pixel chips in a serial power chain, giving major material budget reductions in the pixel detectors. Flexible control and readout interfaces enable the pixel chips to be employed efficiently across pixel detector systems with highly varying hit and readout rates. Final production chip versions have recently been submitted and are currently under thorough verification and testing at pixel module and system level in the two experiments, so the production of tens of thousands of pixel modules can get started for their integration into the upgraded ATLAS and CMS pixel detectors.

It has been a major effort for a large number of collaborators (students, post-docs, physicists and chip design engineers) across 24 institutes to get to this point after 10 years of extensive R&D. Many expected challenges and unexpected problems have gradually been resolved by a collective effort, that constantly had to be adapted to an evolving design team with regular departures of experienced team members. It has been an additional challenge to handle two slightly different chip versions. For future pixel chips of increased performance and complexity with significantly increased IC technology costs, it is recommended to develop common chips to use efficiently available HEP (High Energy Physics) chip design resources. The RD53 design team has worked very well across the two experiments. It has also been highly beneficial for the two pixel detector communities to have an open information flow on chip, module and system issues and sharing appropriate solutions.

Acknowledgments

We would like to thank and acknowledge our colleagues in the ATLAS and CMS pixel projects for their help, patience and extensive work getting to final production chips. A large number of people have been involved in defining appropriate chip specifications and make extensive chip and system tests with serially powered pixel modules with different pixel sensors.

We would also like to thank the CERN micro electronics group for their extensive technology support and handling communications with IMEC and the foundry for chip prototyping and production, as supported by the EU Europractice chip design program.

The solid and long term support from participating RD53 institutes has been critical for us to reach a successful end of 10 years of challenging R&D for these particularly difficult detector applications in an unprecedented harsh radiation environment. Funding has been provided by the following

agencies: CERN; MEYS CR (Czech Republic); CEA and CNRS/IN2P3 (France); HGF and MPG (Germany); GSRI, Greece; INFN (Italy); NWO (Netherlands); RCN (Norway); MCIN/AEI and PCTI (Spain); Swiss Funding Agencies (Switzerland); STFC (United Kingdom); Department of Energy (USA).

References

[1] RD53 collaboration web pages: http://cern.ch/RD53

[2] ATLAS-TDR-030, Technical Design Report for the ATLAS Inner Tracker Pixel Detector: https://cds.cern.ch/record/2285585

[3] CMS-TDR-11, CMS Technical Design Report for the Pixel Detector Upgrade: https://cds.cern.ch/record/1481838

[4] LPGBT, Low Power GigaBit Transfer link chip: https://lpgbt.web.cern.ch/lpgbt/

[5] The FE-I4 pixel readout integrated circuit. M. Garcia-Sciveres et al. NIMA, Volume 636, issue 1, p155-159. https://doi.org/10.1016/j.nima.2010.04.101

[6] Evaluation of planar silicon pixel sensors with the RD53A readout chip for the Phase-2 Upgrade of the CMS Inner Tracker. W. Adam et al. 2023 JINST 18 P11015. https://doi.org/10.1088/1748-0221/18/11/P11015

[7] CERN-LHCC-2012-022, Letter of Intent for the Phase-II Upgrade of the ATLAS Experiment: https://cds.cern.ch/record/1502664

[8] CERN-LHCC-2015-010, Technical Proposal for the Phase-II Upgrade of the CMS Detector: https://cds.cern.ch/record/2020886

[9] RD53B users guide, CERN-RD53-PUB-21-001. https://cds.cern.ch/record/2754251

[10] RD53B Manuals for ATLAS and CMS. CERN-RD53-PUB-19-002. https://cds.cern.ch/record/2665301

[11] RD53C manuals for ATLAS and CMS. CERN-RD53-PUB-24-001 https://cds.cern.ch/record/2890222

[12] Design and optimisation of low power hybrid pixel array logic for the extreme hit and trigger rates of the Large Hadron Collider upgrade. S. Marconi. PHD thesis Perugia. https://cds.cern.ch/record/2630094

[13] Ultra high-density hybrid pixel sensors for the detection of charge particles. A. Paterno PHD thesis, Torino university https://iris.polito.it/handle/11583/2743337

[14] Pixel detectors with local intelligence: An IC designer point of view. F. Krummenacher. NIMA, vol. 305, 1991, pp. 527-532. https://doi.org/10.1016/0168-9002(91)90152-G

[15] Optimal use of charge information for the HL-LHC pixel detector readout. Y. Chen at al. NIMA, Volume 902, 11 September 2018, P 197-210. https://doi.org/10.1016/j.nima.2018.01.091

[16] Comparative evaluation of analogue front-end designs for the CMS Inner Tracker at the High Luminosity LHC. N. Emriskova at al. JINST 16 (2021) P12014. https://doi.org/10.1088/1748-0221/16/12/P12014

[17] Development of calibration techniques and performance analysis of the CMS Inner Tracker for the High Luminosity phase of LHC. Master thesis. L. Damenti. https://cds.cern.ch/record/2893555

[18] Optimization of the 65-nm CMOS Linear Front-End Circuit for the CMS Pixel Readout at the HL-LHC. L. Gaioni et al. IEEE Transactions on Nuclear Science Volume 68, Issue 11, Pages 2682 - 26921, November 2021. https://doi.org/10.1109/TNS.2021.3117666

[19] CMS analog front-end: simulations and measurements. L. Gaioni, CERN cds document, CERN-RD53-PUB-20-002. https://cds.cern.ch/record/2746420

[20] Xilinx, Aurora 64B/66B Protocol Specification, SP011 (v1.3) October 1, 2014 https://docs.xilinx.com/v/u/en-US/aurora_64b66b_protocol_spec_sp011

[21] A Clock and Data Recovery Circuit for the ALTAS/CMS HL-LHC Pixel Front End Chip in 65 nm CMOS Technology. P. Rymaszewski et al. TWEPP 2019 proceedings. Proceedings of science V 370. https://doi.org/10.22323/1.370.0046

[22] Design and characterization of pixel IC electronics and sensors for new pixel detector generations. P. Rymaszewski PHD thesis, Bonn University. https://bonndoc.ulb.uni-bonn.de/xmlui/handle/20.500.11811/9695

[23] The design and test results of A Giga-Bit Cable Receiver (GBCR) for the ATLAS Inner Tracker Pixel Detector. L. Zhang et al. 2023 JINST 18 C03005 https://doi.org/10.1088/1748-0221/18/03/C03005

[24] The Opto-electrical conversion system for the data transmission chain of the ATLAS ITk Pixel detector upgrade for the HL-LHC. L. Franconi et al., 2022 J. Phys.: Conf. Ser. 2374 012105. https://doi.org/10.1088/1742-6596/2374/1/012105

[25] LM431AdjustablePrecisionZenerShuntRegulator, Texas instruments. https://www.ti.com/lit/gpn/LM431

[26] An integrated Shunt-LDO regulator for serial powered systems. M. Karagounis et al. ESSCIRC 2009 proceedings, https://doi.org/10.1109/ESSCIRC.2009.5325974

[27] Stabilization and Protection of the Shunt-LDO regulator for the HL-LHC pixel detector upgrades, M. Karagounis et al., Proceedings of Science, Vol. 370, Sept. 2019 https://doi.org/10.22323/1.370.0067

[28] A Rad-Hard Bandgap Voltage Reference for High Energy Physics Experiments, G. Traversi et al., Lecture Notes in Electrical Engineering, Vol. 627, Sept. 2019. https://doi.org/10.1007/978-3-030-37277-4_3

[29] Characterization and verification of the Shunt-LDO regulator and its protection circuits for serial powering of the ATLAS and CMS pixel detectors, M. Karagounis et al., Journal of Physics: Conference Series, Vol. 2374 (12022), May 2021. https://doi.org/10.1088/1742-6596/2374/1/012071

[30] An Integrated Shunt-LDO regulator for serial powered systems, M. Karagounis et al., Proceedings of the IEEE European Solid-State Circuits Conference, Athens, Greece, p. 276 -279, 2009. https://doi.org/10.1109/ESSCIRC.2009.5325974

[31] Concept, realization and characterization of serially powered pixel modules (Serial Powering), D. B. Ta et al., NIMA Volume 565, Issue 1, 1 September 2006, Pages 113-118. https://doi.org/10.1016/j.nima.2006.04.074

[32] A serial powering scheme for the ATLAS pixel detector at sLHC. L Gonella et al 2010 JINST 5 C12002. https://doi.org/10.1088/1748-0221/5/12/C12002

[33] Evaluation of a Serial Powering Scheme and its Building Blocks for the ATLAS ITk Pixel Detector. F. Hinterkeuser PHD thesis, Bonn University. https://bonndoc.ulb.uni-bonn.de/xmlui/bitstream/handle/20.500.11811/10392/6851.pdf?sequence=2&isAllowed=y

[34] Characterisation of the first digital modules with RD53B-CMS readout chips for the Phase-2 Upgrade of the CMS Inner Tracker. M. Missiroli et al 2023 JINST 18 C01027. https://doi.org/10.1088/1748-0221/18/01/C01027

[35] RD53A chip susceptibility to electromagnetic conducted noise. A. Prada et al. TWEPP2019, Proceedings of Science Volume 370. https://doi.org/10.22323/1.370.0064

[36] The RD53 chip monitoring system, M. Menouni,CERN cds document, CERN-RD53-NOTE-23-001. https://cds.cern.ch/record/2851168

[37] Split-SAR ADCs: Improved Linearity With Power and Speed Optimization. Y. Zhu, et al. IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 22, no. 02, pp. 372-383, 2014. https://doi.org/10.1109/TVLSI.2013.2242501

[38] A 12b 50MS/s 2.1mW SAR ADC with redundancy and digital background calibration. A. H. Chang, H. -S. Lee and D. Boning. 2013 Proceedings of the ESSCIRC (ESSCIRC), Bucharest, Romania, 2013, pp. 109-112, doi: 10.1109/ESSCIRC.2013.6649084. https://doi.org/10.1109/ESSCIRC.2013.6649084

[39] Radiation-induced short channel (RISCE) and narrow channel (RINCE) effects in 65 and 130 nm MOSFETs, F. Faccio et al. IEEE Trans. Nucl. Sci., vol. 62, no. 6, pp. 2933–2940, Dec. 2015. https://doi.org/10.1109/TNS.2015.2492778

[40] Dose-Rate Sensitivity of 65-nm MOSFETs Exposed to Ultrahigh Doses, G. Borghello et al. IEEE Transactions on Nuclear Science, Volume 65, Issue 8, August 2018, https://doi.org/10.1109/TNS.2018.2828142

[41] 1-Grad total dose evaluation of 65 nm CMOS technology for the HL-LHC upgrades. M. Menouni et al. 2015 JINST 10 C05009. https://doi.org/10.1088/1748-0221/10/05/C05009

[42] Characterization of radiation effects in 65 nm digital circuits with the DRAD digital radiation test chip. L.M. Jara Casas et al. 2017 JINST 12 C02039. https://doi.org/10.1088/1748-0221/12/02/C02039

[43] DRAD results obtained during irradiation campaigns. L. Jara Casas, CERN cds document, CERN-RD53-PUB-20-001. https://cds.cern.ch/record/2725573

[44] Single Event Effects mitigation with TMRG tool. S. Kulis. 2017 JINST 12 C02039. https://doi.org/10.1088/1748-0221/12/01/C01082

[45] Single event effects on the RD53B pixel chip digital logic and on-chip CDR. J. Lalic et al. 2022 JINST 17. https://doi.org/10.1088/1748-0221/17/05/C05001

[46] Single event effects testing of the RD53B chip. M. Menouni et al. 2022 J. Phys.: Conf. Ser. 2374. https://doi.org/10.1088/1742-6596/2374/1/012084

[47] Heavy-Ions induced SEE effects measurements for the STRURED ASIC. G. De Robertis et al. Nuclear Physics B (Proc. Suppl.) 215 (2011) 333–336. https://doi.org/10.1016/j.nuclphysbps.2011.04.046

[48] Characterization of Soft Error Rate Against Memory Elements Spacing and Clock Skew in a Logic with Triple Modular Redundancy in a 65nm Process. S. Miryala et al. TWEPP2018 conference proceedings. Proceedings of science TWEPP2018 029. https://doi.org/10.22323/1.343.0029

[49] Measurements of the radiation damage to the ITkPixV1 chip in X-ray irradiations. M. Mironova et al. NIMA Volume 1039, 11 September 2022, 166947. https:https://doi.org/10.1016/j.nima.2022.166947

[50] Reusable SystemVerilog-UVM design framework with constrained stimuli modeling for High Energy Physics applications. S. Marconi at al. 2015 IEEE International Symposium on Systems Engineering (ISSE). https://doi.org/10.1109/SysEng.2015.7302788

[51] The RD53 collaboration's SystemVerilog-UVM simulation framework and its general applicability to design of advanced pixel readout chips. S. Marconi at al. JINST Volume 9, Oct 2014. https://doi.org/10.1088/1748-0221/9/10/P10005

[52] BDAQ53, a versatile pixel detector readout and test system for the ATLAS and CMS HL-LHC upgrades. M. Daas et al. NIMA, Volume 986, 11 January 2021, 164721 https://doi.org/10.1016/j.nima.2020.164721

[53] YARR - A PCIe based Readout Concept for Current and Future ATLAS Pixel Modules. T. Heim et al. J. Phys.: Conf. Ser. 898 032053. http://dx.doi.org/10.1088/1742-6596/898/3/032053

[54] FELIX: the Detector Interface for the ATLAS Experiment at CERN. A. Paramonov et al. CHEP 2021 https://doi.org/10.1051/epjconf/202125104006

[55] The CMS Inner Tracker DAQ system for the High Luminosity upgrade of LHC from single-chip testing, to large-scale assembly qualification, D Mauro et al. CHEP2023, 26th International Conference on Computing in High Energy & Nuclear Physics. https://cds.cern.ch/record/2869307/files/CR2023_113.pdf

[56] Hybrid Pixel Readout Chip Verification, Characterization and Wafer Level Testing for the ATLAS-ITK Upgrade at the HL-LHC. M. Stanke PHD thesis. Bonn university. https://bonndoc.ulb.uni-bonn.de/xmlui/bitstream/handle/20.500.11811/10923/7123.pdf?sequence=2&isAllowed=y

[57] Cocotb open source co-simulation in Python. https://www.cocotb.org/

[58] RD53B Wafer Testing for the ATLAS ITk Pixel Detector. M. Stanke et al. 2022 J. Phys.: Conf. Ser. 2374 012087. https://doi.org/10.1088/1742-6596/2374/1/012087

[59] Wafer-level testing of the readout chip of the CMS Inner Tracker for HL-LHC. M. Grippo et al. NIMA V1044, 1 December 2022, 167496. https://doi.org/10.1016/j.nima.2022.167496