RD53 Pixel Readout Chips in 65 nm CMOS for ATLAS and CMS Phase2 upgrades

Luca Pacher

on behalf of the RD53 Collaboration

XII Front-End Electronics Workshop

Torino - Jun 12, 2023

Outline

- The RD53 collaboration
- ASICs developments
- Review of design requirements
- Selected implementation details
- Verification strategy
- Summary and outlook

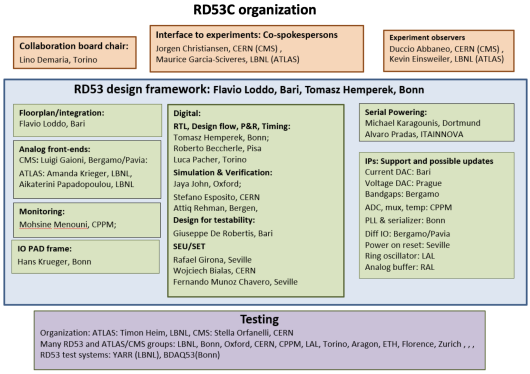
The RD53 collaboration on 65 nm CMOS technology

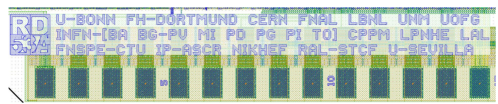
Initial mandate

- joint effort between ATLAS and CMS established back in 2013 to develop new pixel readout chips for HL-LHC pixel detectors upgrades
- **24 partecipating institutes (Europe + USA)**
- characterization of **chosen 65 nm CMOS technology** in radiation environment
- design of a **rad-hard IP library** (analog front-ends, A/D and D/A converters, CDR/PLL, high-speed serializers, RX/TX, ShuntLDO)
- design and characterization of **half-size pixel chip demonstrator** with design variations (**RD53A**)

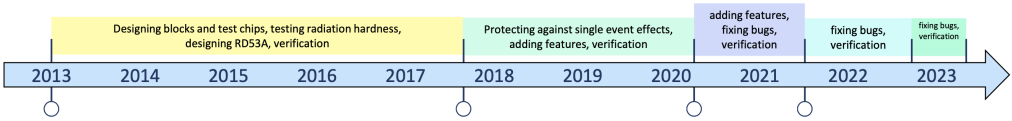
In 2018 the LHCC extended the collaboration for 3 years

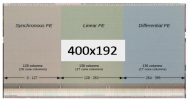
- design of **full-size pre-production** (**RD53B**) and production (**RD53C**) pixel chips for ATLAS and CMS experiments
- ATLAS and CMS chips are two "instances" of a **common ASIC design infrastructure** with different size for mechanical integration and Analog Front-Ends according to specific requirements by the experiments
- request to LHCC to further extend the collaboration by 3 years to ensure necessary support for testing and to experiments





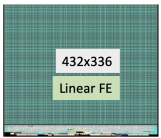
RD53 ASICs development





submitted in August 2017

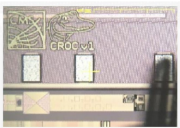
submitted in March 2020



submitted in June 2021

RD53A

RD53B-ATLAS (ItkPix_V1)



RD53B-CMS (CROC_V1)

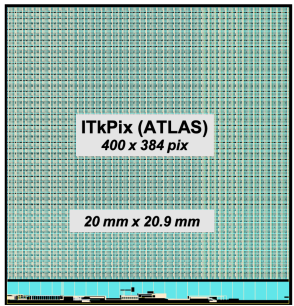
Design strategy

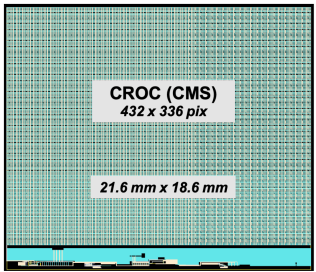
Digital-on-top (DoT) approach:

- common digital-centric design and verification framework referred to as RD53B
- RTL code and digital implementation flows parameterized in terms of "experiment flavour"
- ATLAS and CMS chips are two "instances" of the same common design
- different chip-size for mechanical integration and Analog Front-End (AFE) according to specific requirements of the experiments
- common digital synthesized logic
- unused experiment-specific additional features simply turned off by configuration

```
52
53  `ifdef ATLAS_CHIP                   // ATLAS final chip
54      `define LHC_EXPERIMENT  "ATLAS"
55      `define CORE_COLUMNS     50     // 50 x 8 = 400 pixel columns
56      `define CORE_ROWS        48     // 48 x 8 = 384 pixel rows
57
58  `elsif CMS_CHIP                     // CMS final chip
59      `define LHC_EXPERIMENT  "CMS"
60      `define CORE_COLUMNS     54     // 54 x 8 = 432 pixel columns
61      `define CORE_ROWS        42     // 42 x 8 = 336 pixel rows
62
63  `endif
```

```
1
2   setenv RD53_GIT_ROOT ${PWD}
3
4   if ( ! ${?RD53_SOS_ROOT} ) then
5      setenv RD53_SOS_ROOT ${RD53_GIT_ROOT}/cds/RD53_DM
6   endif
7
8   setenv CDS_PROJECT  ${RD53_SOS_ROOT}
9
10  #setenv LHC_EXPERIMENT  ATLAS
11  setenv LHC_EXPERIMENT CMS
12
13  #setenv RD53_VERSION RD53B
14  setenv RD53_VERSION  RD53C
```





Final **RD53C-ATLAS (ITkPix_V2)** production chip submitted on Apr 2023 !

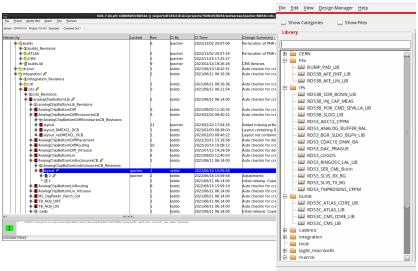
Design requirements and chip specifications

Technology	65nm CMOS
Pixel size	50 µm x 50 µm
ATLAS (CMS) pixel rows/columns	384 (336) / 400 (432)
Detector capacitance	< 100 fF (200fF for edge pixels)
Detector leakage	< 10 nA (20nA for edge pixels)
In-time threshold	< 1200 e-
Noise hits	$< 10^{-6}$
Hit rate	< 3 GHz/cm^2 (75 kHz avg. pixel hit rate)
Trigger rate	≤ 4 MHz (trigger only, without readout)
Trigger latency (L0)	≤ 12.5 µs (programmable)
Readout latency (L1)	≤ 25 µs
Manual readout rate	≤ 1 MHz
Hit loss (in-pixel pile-up + other sources)	≤ 2%
Charge resolution (Time over Threshold)	4 bits ToT (also 6-to-4 dual slope mapping)
Readout data rate	1-4 links @ 1.28Gbits/s = max 5.12 Gbits/s
Radiation tolerance	≥ 500 Mrad, 1 10^{16} 1Mev n.eq/cm^2 (at -15°C)
SEU	< 0.05 /hr/chip at 1.5GHz/cm^2 particle flux
Power consumption at max hit/trigger rate	< 1W/cm^2
Temperature range	-40°C ÷ 40°C

Collaborative design across different institutes

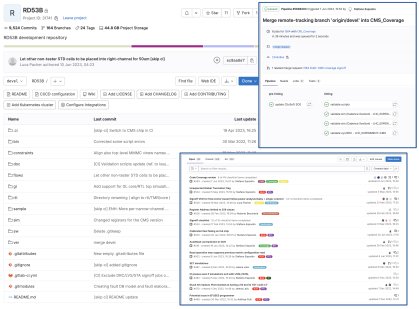
ClioSoft repository

- analog-centric repository
- all NDA-protected items (libraries, models, DRC/LVS rules etc.)
- analog and mixed/signals IP blocks
- top-level hard-macros
- additional analog-driven place-and-route input data (analog pre-routes and main power-grid, DNW enclosures etc.)
- central repository hosted by CERN IT and managed by CERN/MIC

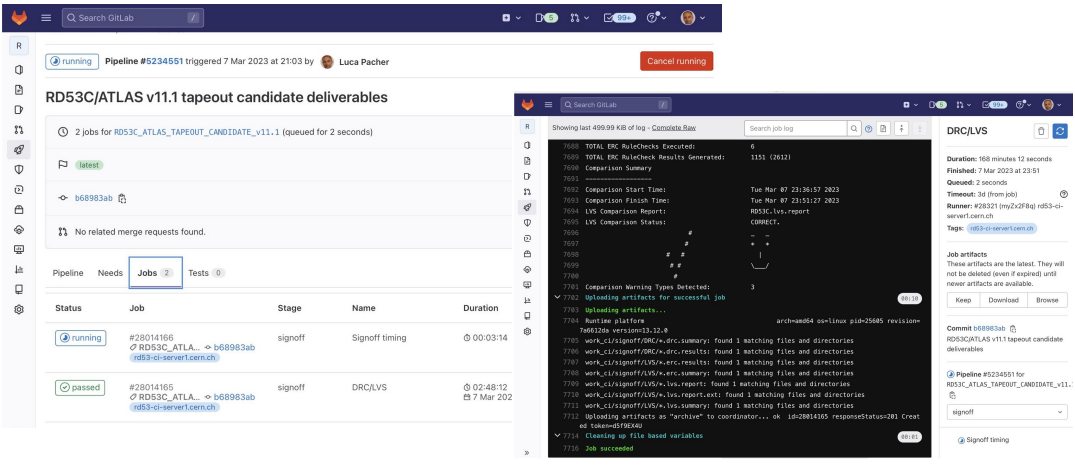


GitLab repository

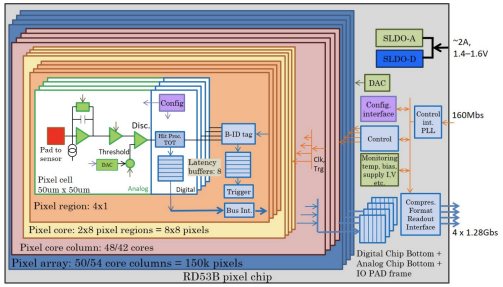
- digital-centric repository
- code versioning (RTL, constraints, flow scripts)
- UVM-based simulation and verification environment (SVE)
- issues/milestones/wiki + continuous-integration
- true software-like approach to track all design and verification activities
- service hosted by CERN IT



Continuous Integration (CI)



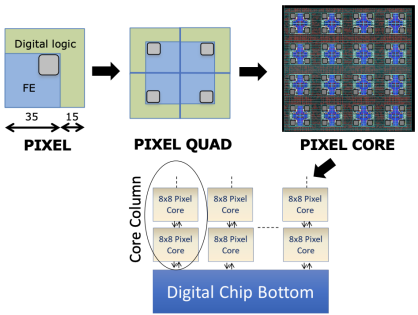
HL-LHC pixel ASIC functional overview "in a nutshell"

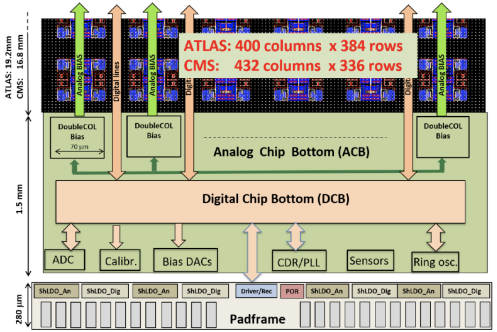


In short, extremely complex pixel readout ASIC designed to have a **minimal I/O interface** with the outside world with as fast as possible data links:

- input/output currents for **serial powering**
- one single 160 Mb/s differential serial input link used for **all chip operations** (synch/PLL lock and CDR, trigger/configuration commands, etc.)
- up to 4x 1.28 Gb/s CML output links for data readout (hit data + service data) compatible with LpGBT
- no external clock, no external reset (see later)

Logical design hierarchy





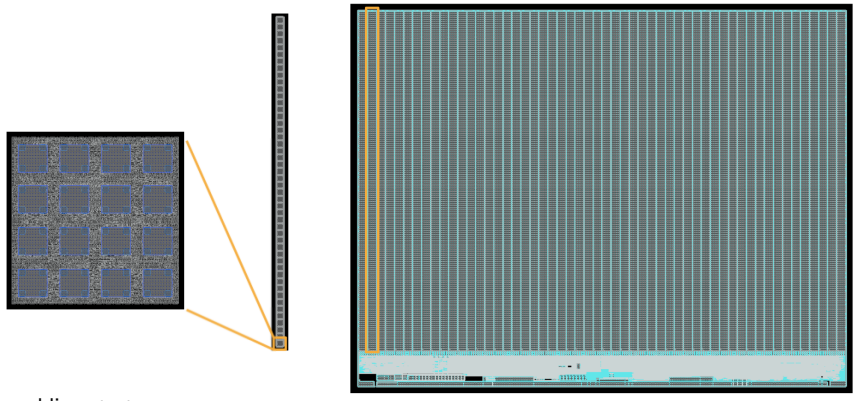
Pixel array

- 50 μm $\times$ 50 μm pixel size, approx. 50% analog and 50% digital
- digital logic synthesized for 8x8 pixels to form a **pixel core**
- all cores are identical $\rightarrow$ efficient hierarchical verification and implementation
- one digital readout block for each PixelCoreColumn in the chip periphery

Chip periphery

- **Analog Chip Bottom (ACB)**: contains all analog and mixed-signal blocks (bias DACs, CDR/PLL, ADC etc.)
- **Digital Chip Bottom (DCB)**: synthesized logic containing communication to/from the chip, readout and global configiration
- Common **padframe**: complex macro containing all I/O blocks with ESD protection and distributed Shunt-LDOs for serial powering

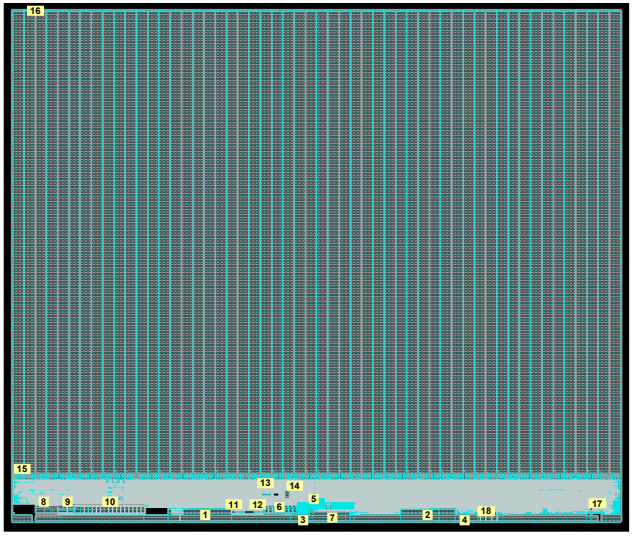
Floorplan and physical design hierarchy



Chip assembling strategy:

- assemble "pixel core" (8x8 pixels) → OA abstract/Liberty/SPEF
- step-and-repeat cores to form a "column of pixel cores" (core-column) → OA abstract/Liberty/SPEF
- chip-periphery/top-level flat (historical choice)
- assemble top-level (synthesized logic + hard-macros + complex padframe macro)

Top-level layout with analog and M/S blocks

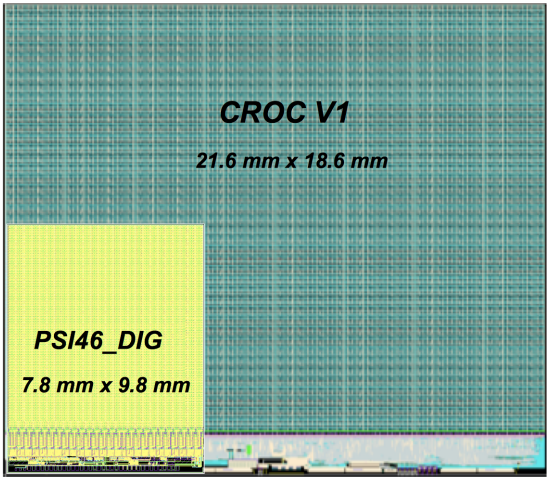


1. ShuLDO-A (master)
2. ShuLDO-D (master)
3. 160 Mb/s serial input LVDS RX
4. 4x DataMerge (DM) LVDS RX
5. PLL/CDR/DM
6. PLL biasing DACs
7. 4x 1.28 Gb/s serializers + CML drivers
8. monitoring block (V-MUX + I-MUX + 12-bit SAR ADC)
9. calibration block + CAP measurement circuit
10. Front-End specific biasing DACs
11. temp/rad sensor (ShuLDO-A hotspot)
12. temp/rad sensor (ShuLDO-D hotspot)
13. temp/rad sensor (middle of the chip periphery)
14. ring oscillators (monitor STD cell speed RAD degradation)
15. bottom-array temp sensor
16. top-array temp sensor
17. 32-bit E-FUSEs block (foundry IP)
18. 4x GPO LVDS TX

The Moore's law from another perspective...

Some "old" reference numbers from current CMS pixel detector readout chip (**PSI46**) specs...

- CMOS 0.25 μm (RHBD)
- 100 μm $\times$ 150 μm pixel size
- 52 $\times$ 80 pixels
- 7.9 mm $\times$ 9.8 mm
- 3 ke^- nominal threshold
- originally full-analog readout at 40 MHz (then moved to "digital" for Phase-I)



Analog and M/S blocks summary

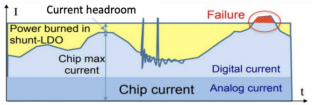
Block	Description
Analogue front end	The ATLAS chip versions use a differential front end. The CMS chip versions use a linear front end.
Shunt LDO	Enables start-up and serial powering. Constant input current shared between chips, modules on serial chains. 1 LDO for digital power, 1 LDO for analogue power.
Clock & Data Recovery (CDR)/PLL	Recovers a 160MHz clock and command/trigger stream. The PLL generates internal clocks: 160 MHz, 64 MHz, 640 MHz and 1.28 GHz.
Bias circuit	Provides biases to the pixel array. Based on bandgap references.
Calibration circuit	Injects hits into the pixel array, to calibrate its response.
Monitoring block	Digitises analogue quantities using a voltage mux, current mux and 12-bit ADC
Temperature and Radiation sensors	Temperature sensors: polysilicon resistors. Radiation sensors: based on PMOS devices with a linear variation in voltage in the dose range 10 - 1000 Mrad.
LVDS pads/drivers	Pads and drivers for differential inputs/outputs

Selected implementation details

Serial powering

Both ATLAS and CMS will adopt the innovative **serial-powering scheme**

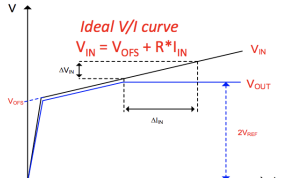
- based on complex **shunt-LDO regulators** inside the chip
- one regulator for each power domain (1x analog + 1x digital)
- **constant input current** shared among more chips (2-4) on the same module (less cables)
- modules are in serial chains: "recycle" current from one module to another
- in case of chip failure, its current can be absorbed by the other chips of the module
- not sensitive to voltage drops
- on-chip regulated supply voltages, low noise
- radiation hardness ($>$ 500 Mrad) silicon proven in RD53A and next prototypes



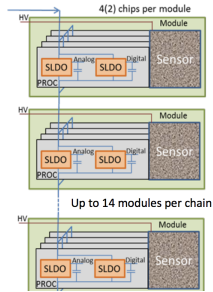
Example current consumption of one readout chip

Protections:

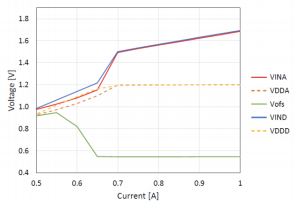
- *Over-voltage protection*: V_{IN} clamped to 2 V
- *Under-shunt protection*: V_{OUT} decreased in case shunt current goes below a certain threshold (due to excess load current)



- V/I curve parameters (Voffset, slope) defined by external resistors
- V_{OUT} tunable by chip configuration



ITkPixV1.1 measurement at room T, pre-rad



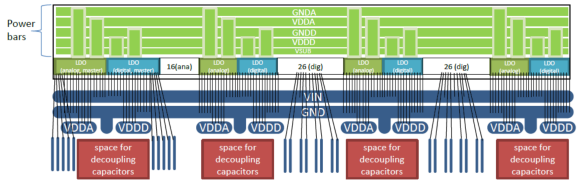
Serial powering (cont'd)

Two power domains in the chip

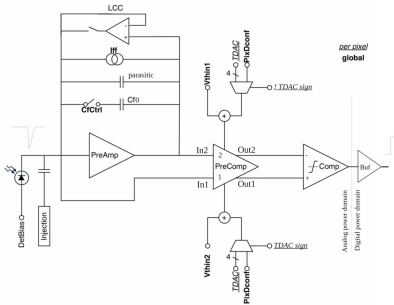
- two ShuntLDOs regulators
- analog (VDDA, GNDA)
- digital (VDDD, GNDD)

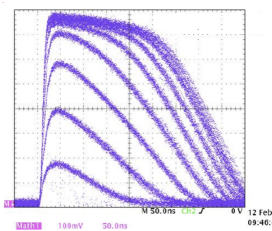
Very complex padframe design !

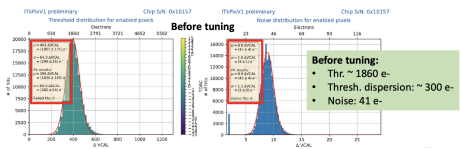
- ShuntLDOs are located in the padframe and distributed along the full chip width
- 4x blocks with sets of 5 Wire Bonds per power net
- power lines come only from the bottom
- GNDA-GNDD-VSUB connected together off-chip
- carefull design of I/O ESD protection structures, startup bandgap reference, over-current protections etc.
- standalone complex macro used for top-level integration

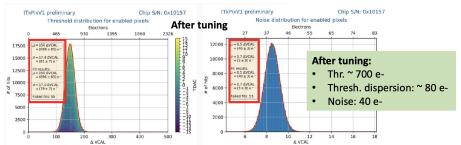


Differential analog front-end (ATLAS)



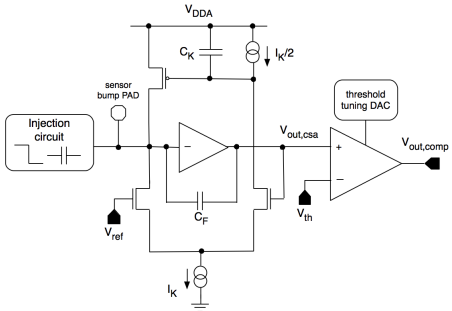




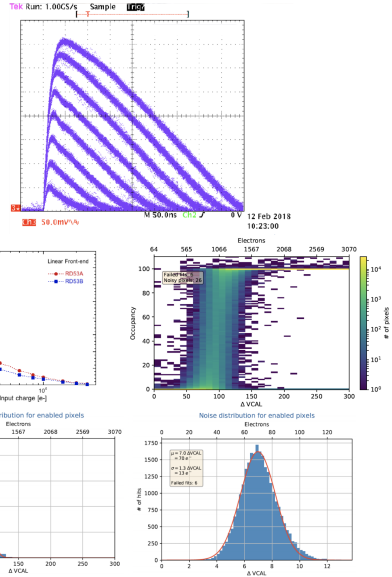


- single-stage **charge-sensitive amplifier (CSA)**
- continuous-reset integrator with tunable feedback current (global setting)
- dedicated leakage-current compensation circuit
- DC-coupled **differential pre-comparator** stage
- per-pixel **threshold adjustment** with local 4+1 bit DAC
- fully-differential comparator

Linear analog front-end (CMS)



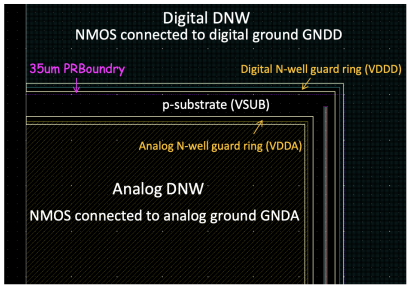
- single-stage **charge-sensitive amplifier (CSA)**
- **Krummenacher feedback** for triangular pulse-shaping and leakage-current compensation
- low-power asynchronous fast comparator for hit discrimination
- 10-bit DAC for global threshold
- per-pixel **threshold adjustment** with local 5-bit DAC

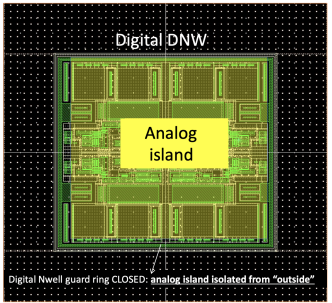


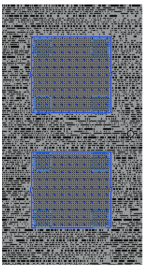
Pixel substrate floorplan

Isolation strategy between analog and digital domains adopted in all RD53 pixel chips:

- analog front-end placed into its own "analog" Deep N-Well (DNW) with local ground GNDA
- digital logic placed into independent "digital" DNW with local ground GNDD
- global p-substrate biased by dedicated VSUB bias line
- PMOS devices of the analog pixel facing the digital boundary
- 2x2 pixels mirrored to form an "analog quad" $\rightarrow$ promoted as "macro" to be used into digital flow



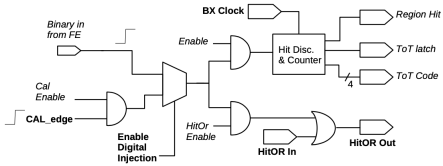




Pixel digital architecture

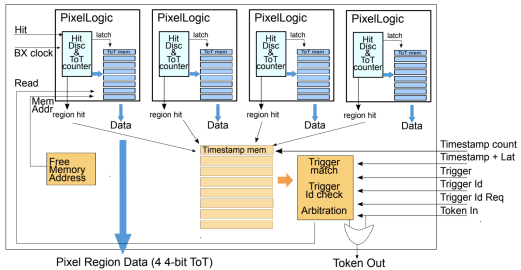
Per-pixel digital logic

- selectable synchronous or asynchronous hit-sampling mode (CMS-only feature)
- charge measurement by means of the **Time-over-Threshold (ToT) technique**
- hits stored as **Time-over-Threshold (ToT)** values + BX time-stamp
- per-pixel 6-bit ToT counters, then mapped to 4-bits only
- selectable ToT dual-slope 6-to-4 ToT mapping for **charge-compression** in case of elongated clusters
- selectable ToT clock counting (40 MHz or dual-edge 80 MHz)
- digital injection



Region-based digital logic

- pixels readout organized into 1 ($r\phi$) x 4 (z) pixel regions (FEI4-like)
- per-pixel 4x8 ToT memory slots
- timestamp memory shared among 4x pixels in the same pixel-region
- **trigger-matching** performed at pixel level with programmable 9-bit trigger latency (max. 12.5 μs)
- token-based readout of hits
- **hit-OR** path + 11-bit **Precision ToT (PTOT)** counting at 640 MHz in the chip periphery for precise sensors characterization



Data-flow architecture

Command, control and timing

- one single 160 Mb/s differential control link with a custom developed protocol driving up to 15 chips (4-bit addressing + broadcast)
- CDR/PLL recovers Data and Clock (see later)

Readout via high-speed CML serial links (1-4 x 1.28 Gb/s)

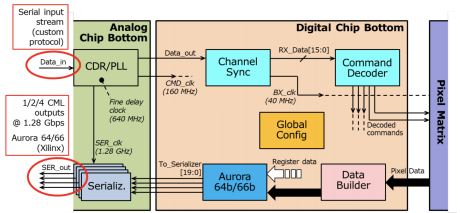
- token-based readout of hits
- hit data read in parallel for each core-column as soon as a valid trigger is received
- multiple levels of data processing, event building, data buffering and formatting
- service data collection and formatting (e.g. ADC motoring data, configuration readback values)
- Xilinx Aurora 64/66 encoding (on-chip Aurora transmitter)
- building an Aurora frame based data stream to be send to high-speed serializers

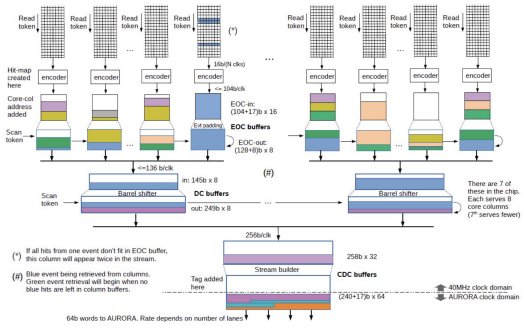
Multi-chip Data-Merging (DM) for low-rate outer layers

- one chip can be configured as "primary" to aggregate serial data coming from one or more "secondary" chips and merge them with its own output
- allows to reduce the number of links to read out data in the outer layers

Design For Testability (DFT)

- entire chip-periphery covered by scan-chain (> 99% ATPG coverage)
- industry-standard methodology adopted to speed-up Wafer Level Testing (WLT) for massive production





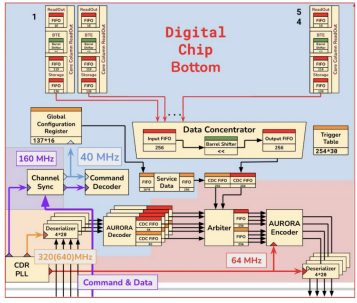
Clock and reset strategy

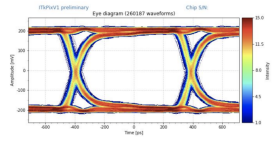
Clock generation and distribution

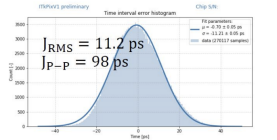
- CDR/PLL used to recover 160 MHz clock and command stream from the transitions on the input serial control link
- generated internal clocks: 160 MHz (channel-synch + DataMerge), 64 MHz (Aurora 64b/66b encoding), 640 MHz (PTOT) and 1.28 GHz (serializers + fine-delays)
- 40 MHz master clock generated by simple clock-division into ChannelSynchronizer (160MHz/4), used my most of chip-periphery and into pixel array
- 4 different clock domains overall running into chip-periphery: standard Clock Domain Crossing (CDC) techniques adopted to deal with multiple clocks (also checked with formal tools)
- automatically-inserted clock-gating used to reduce power consumption

Reset scheme

- **NO reset pin** (nor Power-On Reset) implemented in the chip to minimize the risk of false resets due to SEEs (therefore NO reset line triplication required)
- **synchronous reset** adopted everywhere (with the only exception in the per-pixel asynchronous hit-sampling logic)
- Global and Pixel configuration have **hardcoded default values** selected by a multiplexer at power-up and to switch the muxes to use register values "magic numbers" need to be written into two key registers
- individual blocks have independent reset signals that can be sent using the **GlobalPulse** command
- an additional fast-command called **Clear** can be used to fully and only resets the datapath and control state machines inside the chip
- self-reset of the CDR/PLL if no **Sync** commands are not sent for $\sim 1\ \mu s$







SEU/SET mitigation strategy

SEU protection (chip-periphery)

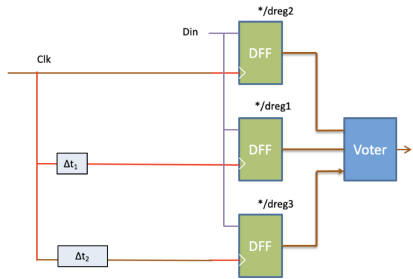
- approx. 100 Hz expected SEU rate per chip for the innermost layer
- protect vital circuitry and data as far as feasible within space limits (but... cannot protect everything!)
- priorities: protect configuration data, state machines, memory pointers
- semi-custom flow developed by RD53 (no TMRgen flow by CERN)
- TMR automatically inserted during synthesis based on register name (all FlipFlops are triplicated unless marked as `_notmr` in RTL)
- placement spacing constraint between FlipFlops then applied during place-and-route
- after TMR, low-enough SEU rate to allow **continuos re-programming** of global configuration registers and pixel configuration registers

SEU protection (pixel level)

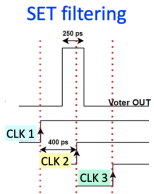
- not enough space available to TMR all FlipFlops in the 8x8 pixel core
- TMR performed only on a limited set of per-pixel configuration latches, then rely on continuos reconfiguration
- extensive SEU/SET simulations to validate the design at both RTL and GL

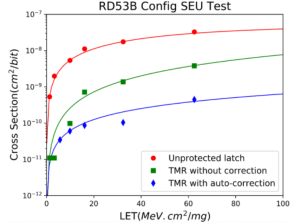
SET mitigation

- NO triplication performed for the majority-voting logic!
- SET filtering achieved by means of automatically-inserted extra delays close to clock pins to locally spread the clock between FlipFlops
- design target: 0.4 ns clock-spread in typical corner
- usage of synchronous reset (almost) everywhere

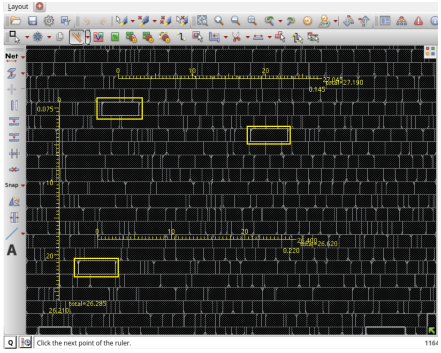


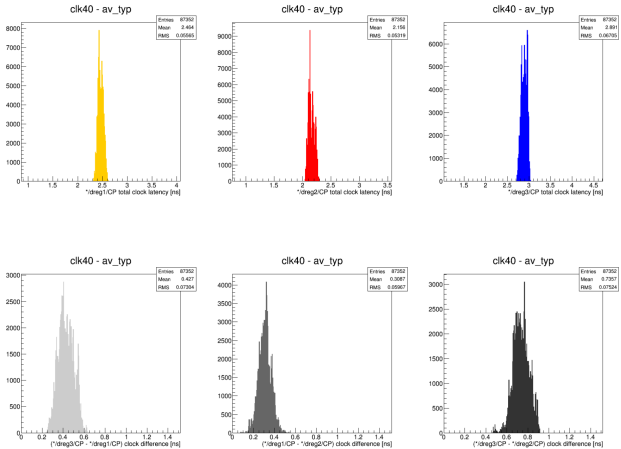
SEU cross-section comparison from heavy-ion testing





SEU/SET mitigation strategy (cont'd)





> 15 μm placement distance guaranteed for all triplicated sequential elements in the design!

Design verification

Digital verification

Verification standard adopted in RD53 is UVM

- Universal Verification Methodology (IEEE Std. 1800.2-2020)
- industry-like approch, well-supported by tool vendors

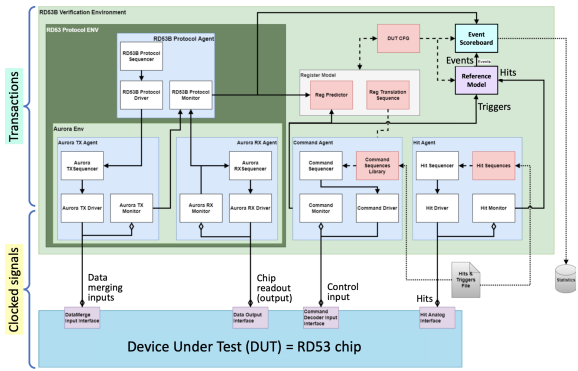
RD53 verification framework

- use of UVM concepts → implement design-specific UVM Verification Components (UVC)
- **Command** UVC: send commands according to protocol
- **Aurora** UVC: receive data, send data for DataMerging
- **Hit** UVC: send random hits or from distributions
- **Reference model**: predicts events, receive hits and triggers, read registers mirror
- **Scoreboard**: receive predictions from the Reference Model, receive data readouts from the DUT

Metric Driven Verification (MDV)

- simulate until the desired coverage is reached
- verification goal is always 100%
- **Code Coverage**
- **Functional Coverage**
- use **constrained randomization** to test all feasible combinations of configuration and inputs
- randomize sets of tests and learn which set is optimal (fewest runs) to reach coverage goals

RD53 verification framework (UVM)

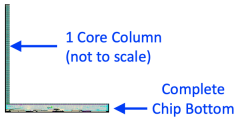


Devices under tests

Slower to simulate

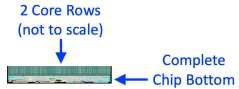
Slower to simulate

RTL Single Core Column
8 x 384 pixels
Obtain: majority of coverage



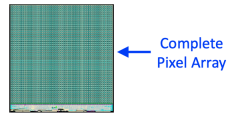
Gate Level Single Core Column
8 x 384 pixels
Obtain: additional timing checks

RTL "Baby Chip"
400 x 16 pixels
Exercise all columns



Gate Level "Baby Chip"
400 x 16 pixels
Exercise all columns

RTL Full Chip
400 x 384 pixels
Exercise filling of FIFOs



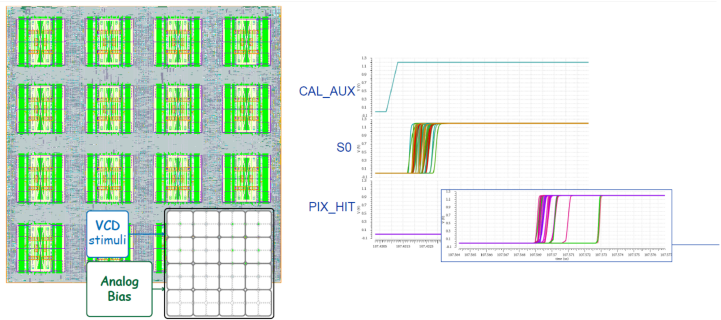
Gate Level Full Chip
400 x 384 pixels
Exercise filling of FIFOs

RTL = Register Transfer Level, before any placement or routing takes place, faster to simulate
Gate Level = after placement and routing, includes parasitics, slower to simulate

Mixed-signal verification

The verification framework is also used to generate digital stimuli (VCD) to drive all the analog and M/S blocks in pure-SPICE simulations:

- create custom digital test (directed test) to stimulate analog and M/S block digital pins
- dump digital waveforms into VCD
- import VCD into analog design environment
- apply analog bias signals
- run full-SPICE or Fast-SPICE simulations as required



Conclusions

- The readout chips for the ATLAS and CMS HL-LHC pixel detectors are being developed by the **RD53 Collaboration** on 65 nm CMOS technology since 2013
- Very big collaboration (24 institutions from Europe and USA) lasting for almost 10 years ...
- ... and (extremely) rare example of engineers and physicists of two experiments working together !
- Industry-like **collaborative design** across different institutes around the world
- The first half-size prototype **RD53A** has been crucial to initially explore design variations and to understand how to cope with the complexity of implementing such large and complex pixel chips
- Development of a common design and verification framework to implement final chips as two different "instances" of a common design having different sizes and different Analog Front-End
- Full-size **pre-production** chips **RD53B-ATLAS (ITkPixV1)** and **RD53B-CMS (CROC_V1)** implementing all functionalities requested to operate into real experiments have been submitted though 2020 and 2021
- All measurements up to now indicate that the chip is generally working fine, with some bugs that have been fixed for final submissions
- Verification is crucial in all its different aspects: functional, SEE, analog and M/S, power
- The ATLAS **final** production chip **ITkPixV2** has been just submitted (wafers expected back from the foundry in few weeks)
- The submission of the final CMS chip is foreseen for end of July

Thank you

for your attention !