VERTEX 2025
33rd International Workshop on Vertex Detectors

Readout and testing procedures to characterise the CMS inner tracker pixel detector for HL-LHC

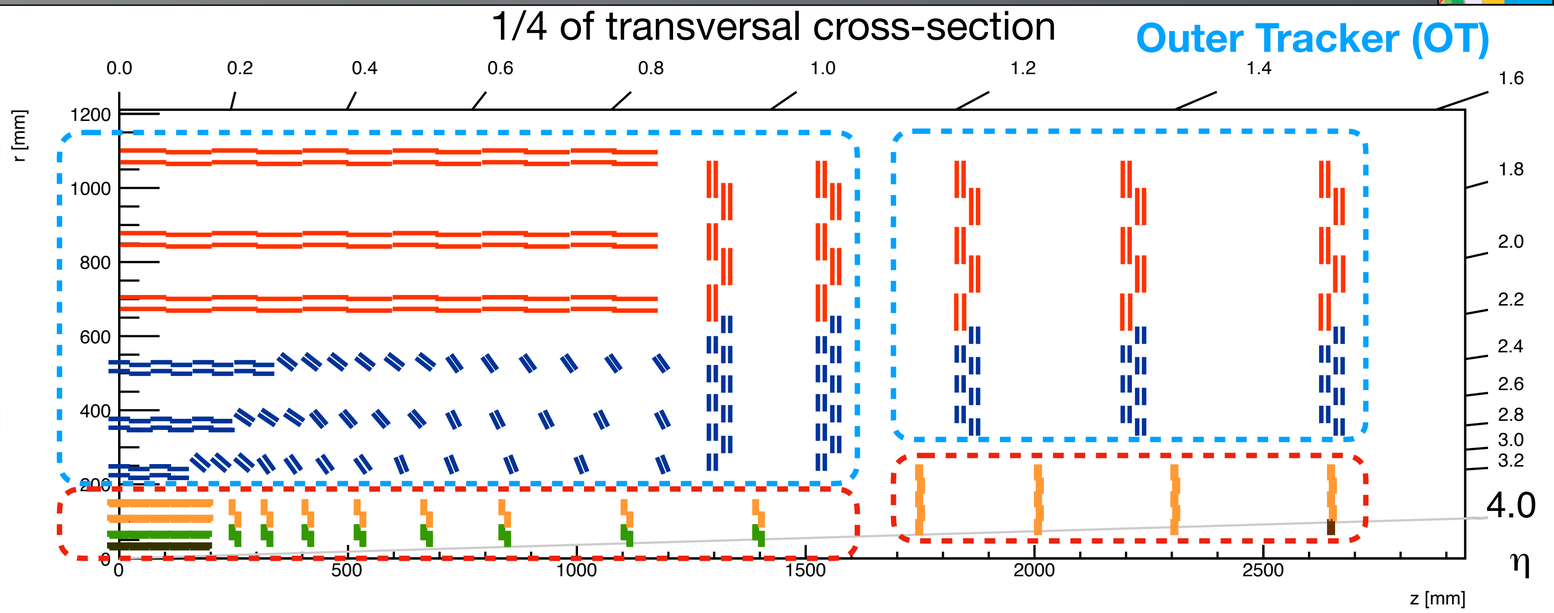
Mauro Dinardo, Università degli Studi di Milano Bicocca and INFN

Introduction to the CMS Inner Tracker

For the **High Luminosity LHC (~2030)** the instantaneous luminosity will be increased to improve discovery potential of LHC experiments

To preserve physics object performance despite average number of interactions per crossing up to 200 **experiments need a significant upgrade**

1/4 of transversal cross-section

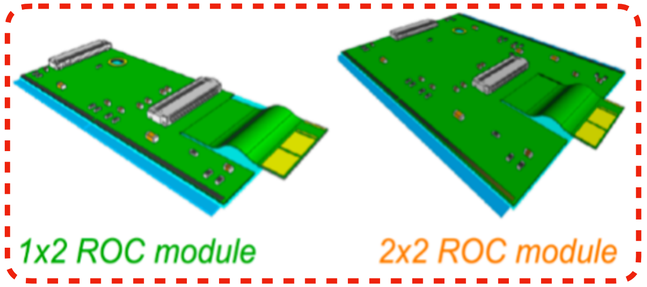


Design constraints for the Inner Tracker (IT)

- Total n. pixels **~2 billion**
- **Trigger rate ~750 kHz ➜** high speed throughput
- **Hit rate ~3.5 GHz/cm²** (innermost layer) and **Trigger latency 12.5 µs ➜** buffer depth
- Unprecedented **radiation level** (~2x10^{16} n_{eq}/cm², **1 Grad**, for innermost layer, over 10 years)

Inner Tracker Sensors

- Pixel size: **25x100 µm²**
- Active thickness: 150 µm
- **3D** for barrel layer-1 & **Planar** for the rest

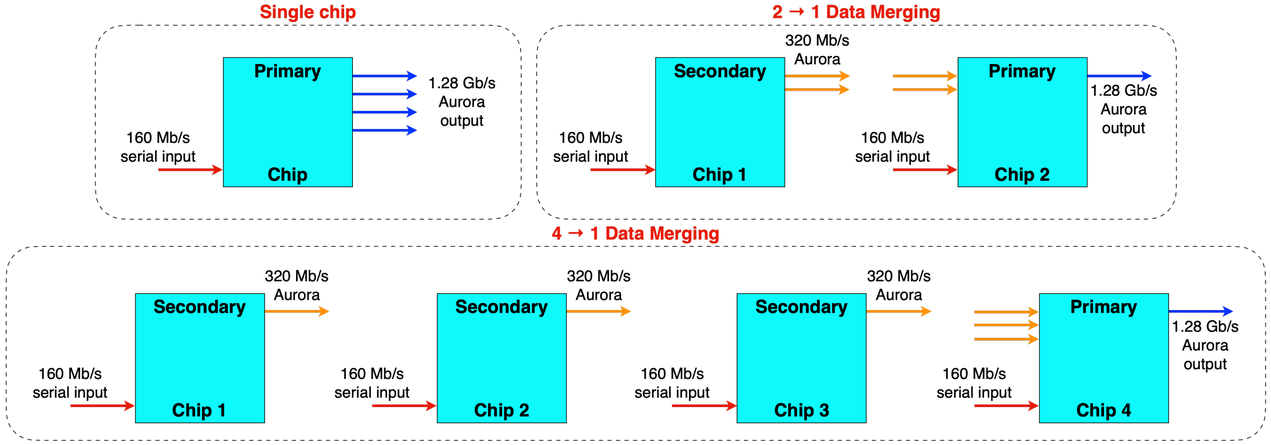


Inner Tracker ReadOut Chip (CROC)

- **Zero suppression**
- **Time-over-Threshold** hit charge measurement with **4-bits ADC**
- Global threshold with **per pixel 5-bits threshold trimming**, dual-slope gain
- Module flavours: **1x2 ROCs** and **2x2 ROCs** for a total of **3900 modules**
- **336x432** = 145152 **pixels**, zero suppression, 50x50 µm²

InnerTracker ReadOut Chip (CROC)

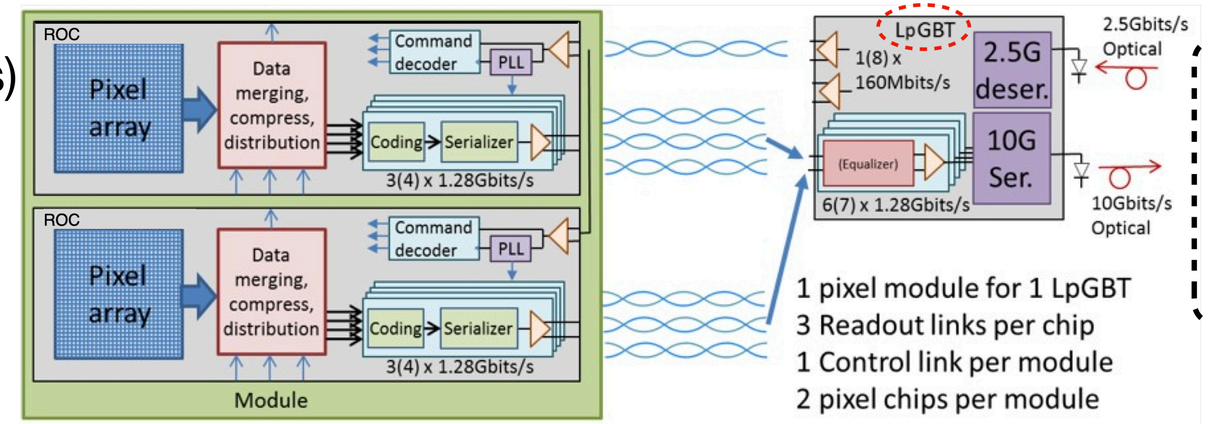
- Clock, trigger, commands to ROC **@160 Mbps**
- Multi-lane readout up to **5.12 Gbps** (1.28 Gbps x 4 lanes)
- **Data merging** capabilities:



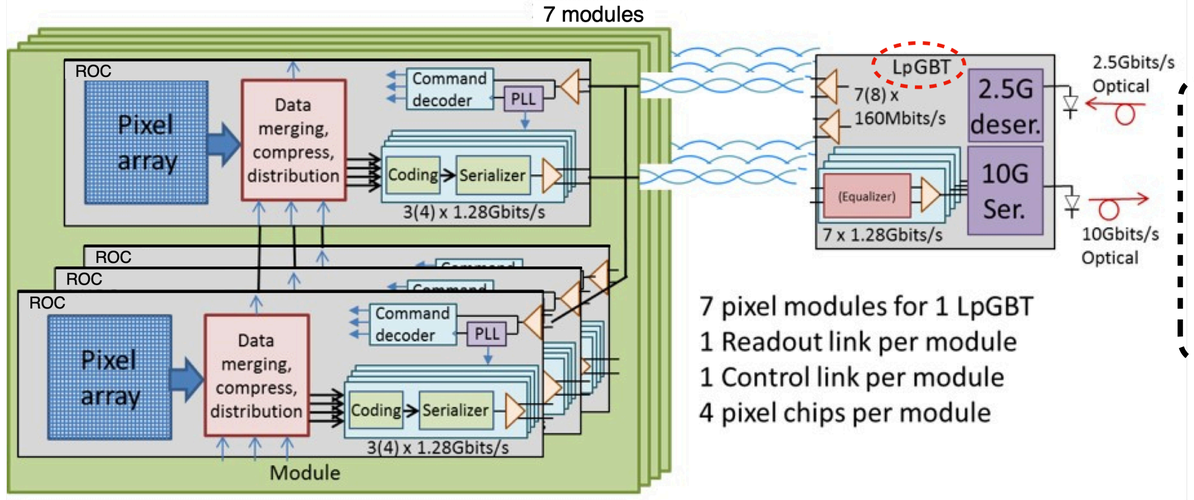
Inner Tracker optical readout

- Data from modules are sent via twisted pairs to **Low**power **G**iga**B**it**T**ransceiver (**LpGBT**) chip
- **LpGBT sends** (**receives**) data to (from) a custom **System-on-Chip** board, a.k.a. Data, Trigger and Control - **DTC** via optical fiber at **10 Gbps** (**2.5 Gbps**)

Highest occupancy link to LpGBT

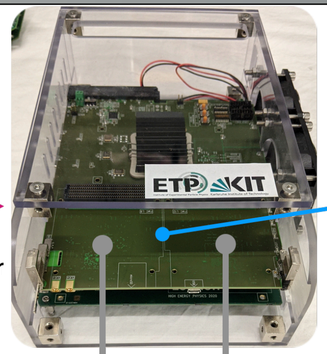


Lowest occupancy link to LpGBT

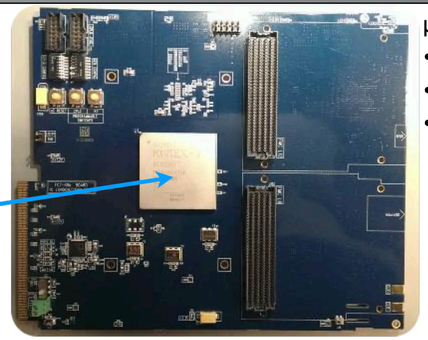


DAQ for module testing

IPbus
Control/Monitor
DAQ

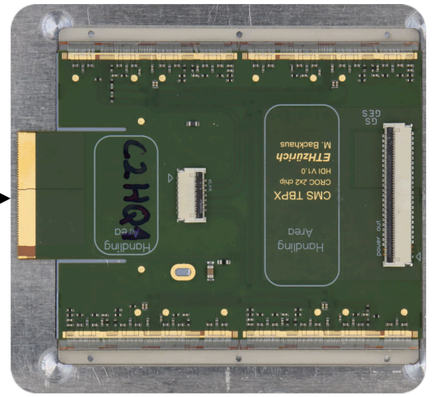


Nano crate with
µDTC board

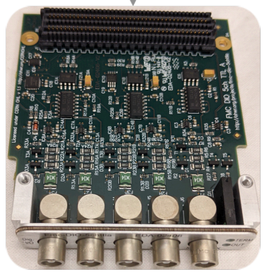


µDTC board:
- Xilinx Kintex 7 FPGA
- two FMC connectors
- DDR3 RAM 4 Gbit

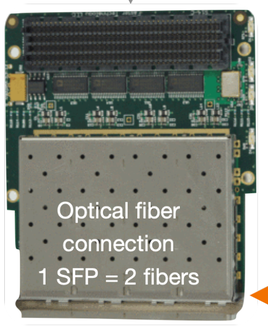
Module: **RD53B chip x 4**



Twisted pair
cable connection

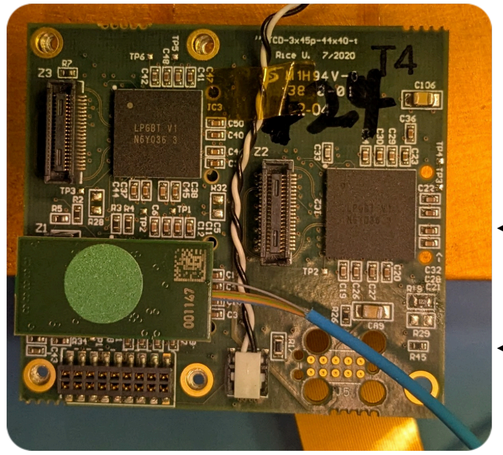


Digital I/O FMC board:
external trigger, clock,
busy signal



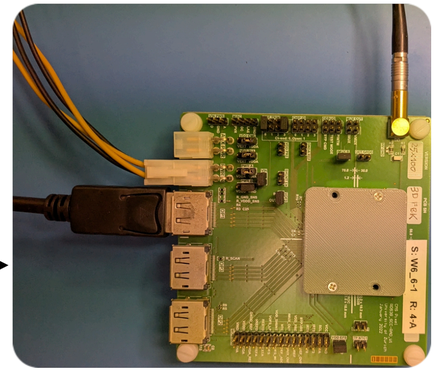
Optical readout FMC board:
up to 8 SFP connectors

Portcard:
- Low-power GigaBit Transceiver chip
- VTRx+ electro-optical converter



Display Port
connection

Single Chip Card:
RD53B chip



DAQ developed to

- Assist readout chip and sensor **development** on **test bench** and at **beam test** experiments
- Assist users in **characterising modules** during **production**

DAQ components

- Software: **Ph2-ACF**
- Hardware: **µDTC**
- Firmware

Phase-2 Acquisition and Control Framework - **Ph2-ACF**

- Designed to be user-friendly
- Based on a minimal hardware
- Conduct stress tests on modules ➜ Ensure meet all requirements

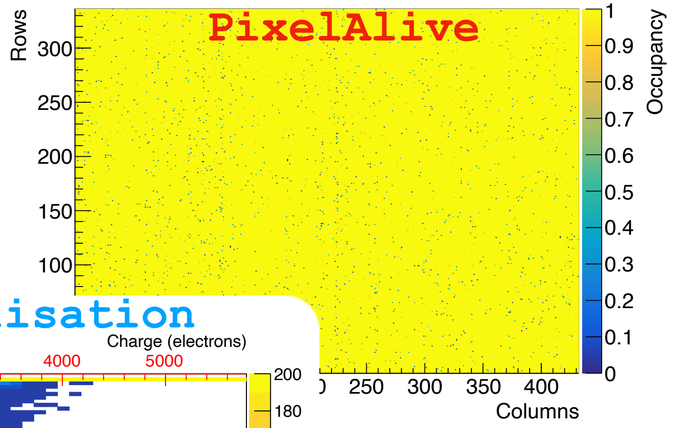
E.g. **burn-in test:** multiple thermal cycles to guarantee the needed production throughput (10 cycles on 8 modules simultaneously in less than 12 hours)

Available tests and scans

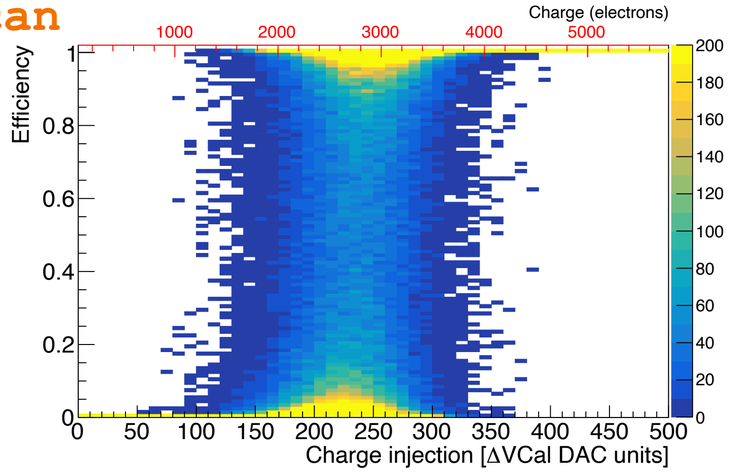
Available scans/calibrations are:

- **PixelAlive**
- **SCurve** (efficiency vs injected charge scan)
- **Threshold equalisation**
- **Injection delay scan**
- **Gain scan**
- Noise scan
- Threshold adjustment
- Bit Error Rate test
- Gain optimisation
- Threshold minimisation
- Latency scan
- Clock delay scan
- Data read back optimisation
- Chip internal voltage tuning
- Generic DAC-DAC scan
- VTRx scan
- LpGBT eye diagram
- Physics (for data talking in beam test experiments)

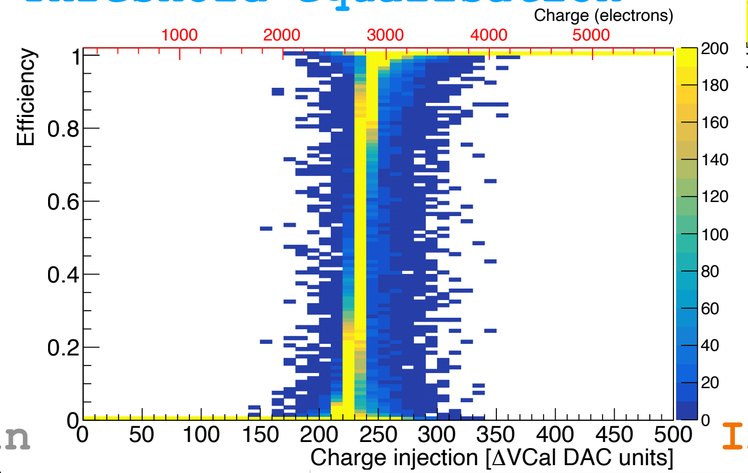
PixelAlive



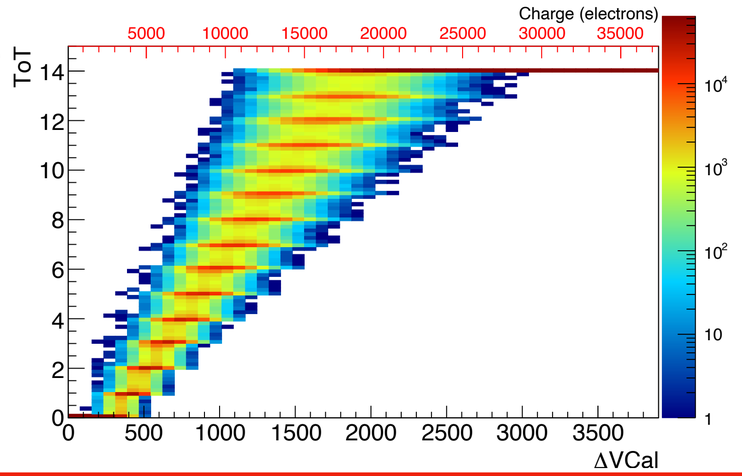
SCurve



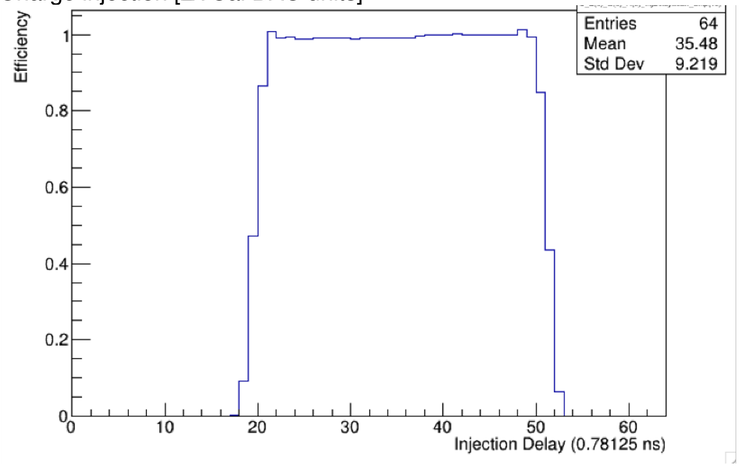
Threshold equalisation



Gain



Injection delay

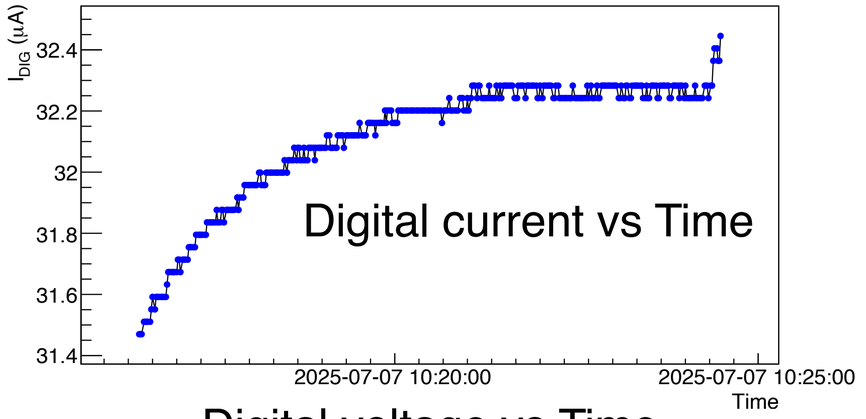


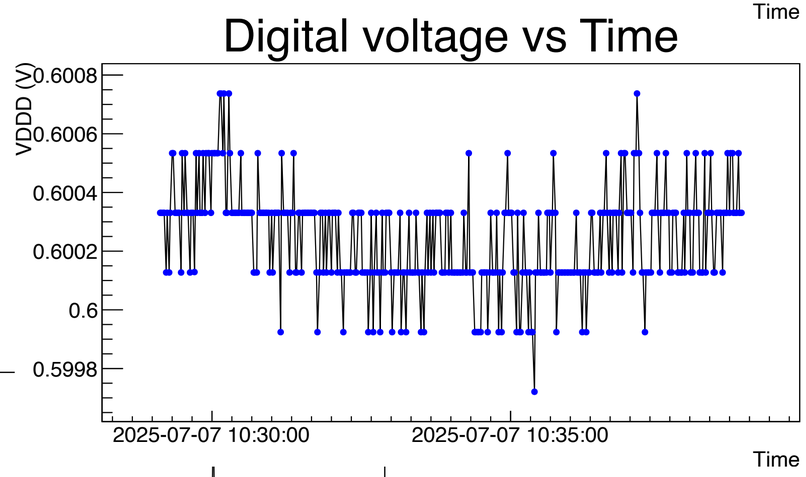
Monitor front-end electronics

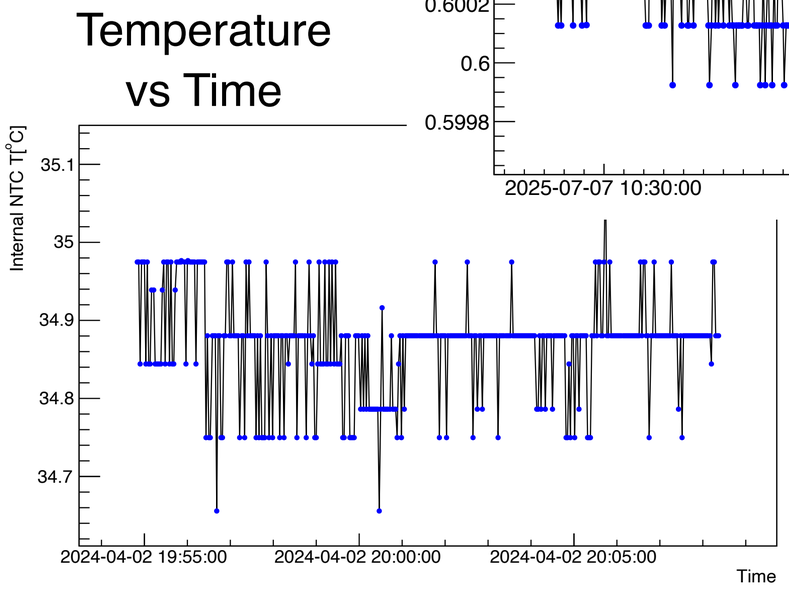
- Voltages, Currents, and Temperatures defined in `XML` configuration file
- Data are stored in `ROOT` files vs Time
- Asynchronous with respect to other tasks, e.g. calibration scans

```
<MonitoringSettings>
  <Monitoring type="RD53B" enable="1">
    <MonitoringSleepTime> 1000 </MonitoringSleepTime>
    <MonitoringElement device="RD53"   register="VINA"                enable="1"/>
    <MonitoringElement device="RD53"   register="VDDA"                enable="1"/>
    <MonitoringElement device="RD53"   register="ANA_IN_CURR"         enable="1"/>
    <MonitoringElement device="RD53"   register="VIND"                enable="1"/>
    <MonitoringElement device="RD53"   register="VDDD"                enable="1"/>
    <MonitoringElement device="RD53"   register="DIG_IN_CURR"         enable="1"/>
    <MonitoringElement device="RD53"   register="Iref"                enable="1"/>
    <MonitoringElement device="RD53"   register="POLY_TEMPSENS_TOP"   enable="1"/>
    <MonitoringElement device="RD53"   register="POLY_TEMPSENS_BOTTOM" enable="1"/>
    <MonitoringElement device="RD53"   register="TEMPSENS_ANA_SLDO"   enable="1"/>
    <MonitoringElement device="RD53"   register="TEMPSENS_DIG_SLDO"   enable="1"/>
    <MonitoringElement device="RD53"   register="TEMPSENS_CENTER"     enable="1"/>
    <MonitoringElement device="RD53"   register="INTERNAL_NTC"        enable="1"/>

    <MonitoringElement device="LpGBT" register="TEMP"                 enable="0"/>
    <MonitoringElement device="LpGBT" register="VDD"                  enable="0"/>
    <MonitoringElement device="LpGBT" register="PUSMStatus"           enable="0"/>
  </Monitoring>
</MonitoringSettings>
```







User interfaces: GUI and Dirigent

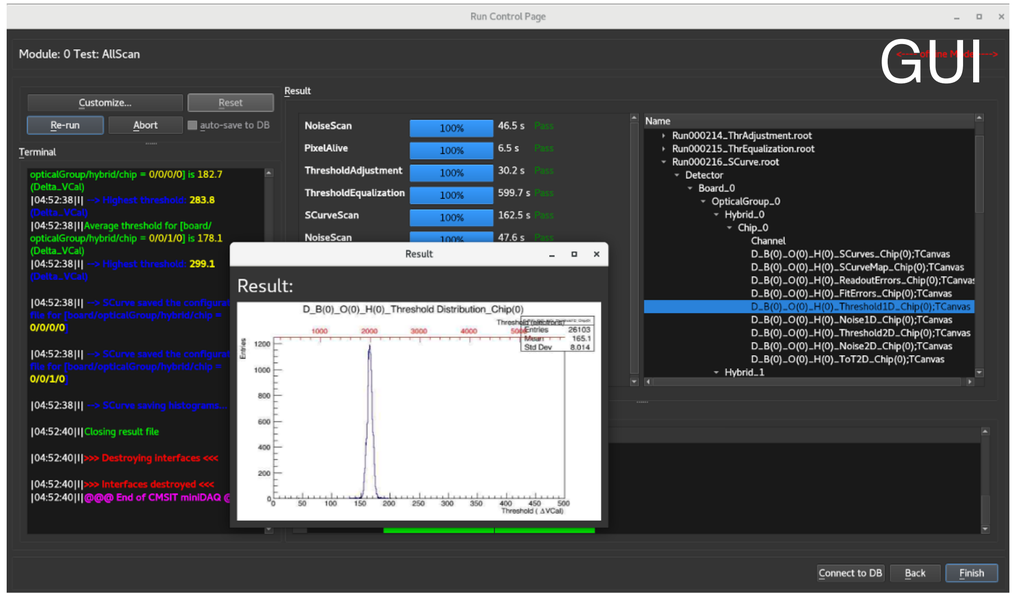
GUI/Dirigent: reliable, consistent (based on `JSON` files), and easy-to-use interfaces for module testing

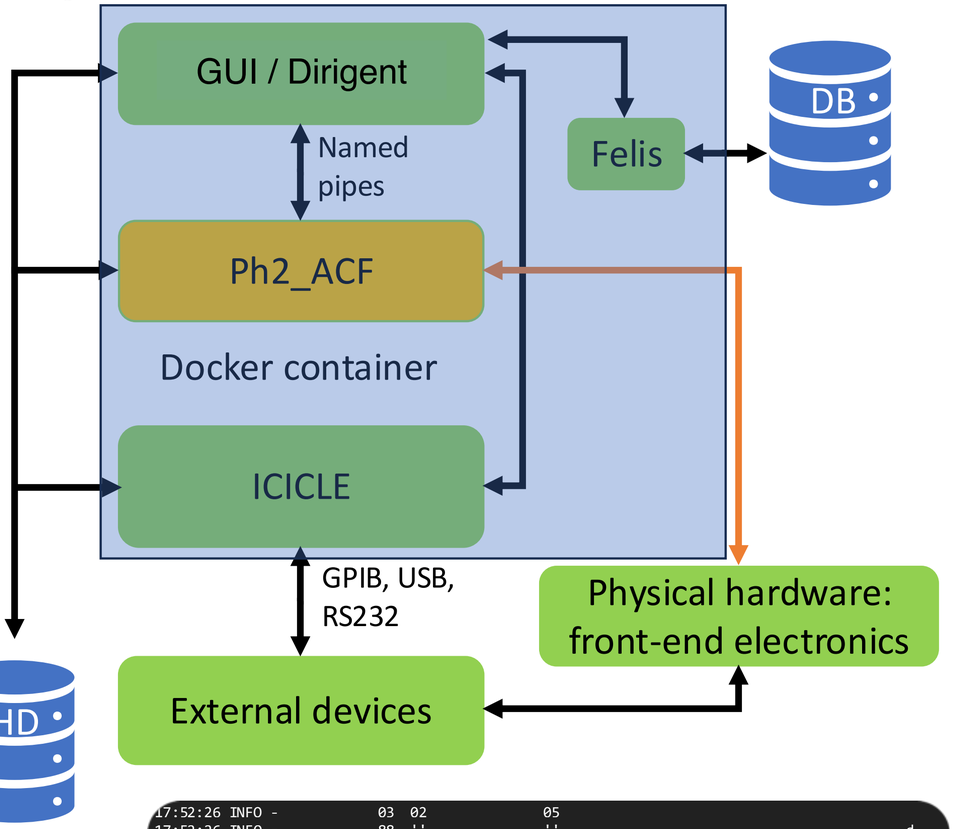
They orchestrate:

- Power supplies and interlocks (“`ICICLE`”)
- **`Ph2-ACF`**, its config. files, and `ROOT` output files
- Grading software and upload to DB (“`Felis`”)

Other main features:

- Run multiple µDTC/`Ph2-ACF` in parallel
- Disable any module that is problematic during test





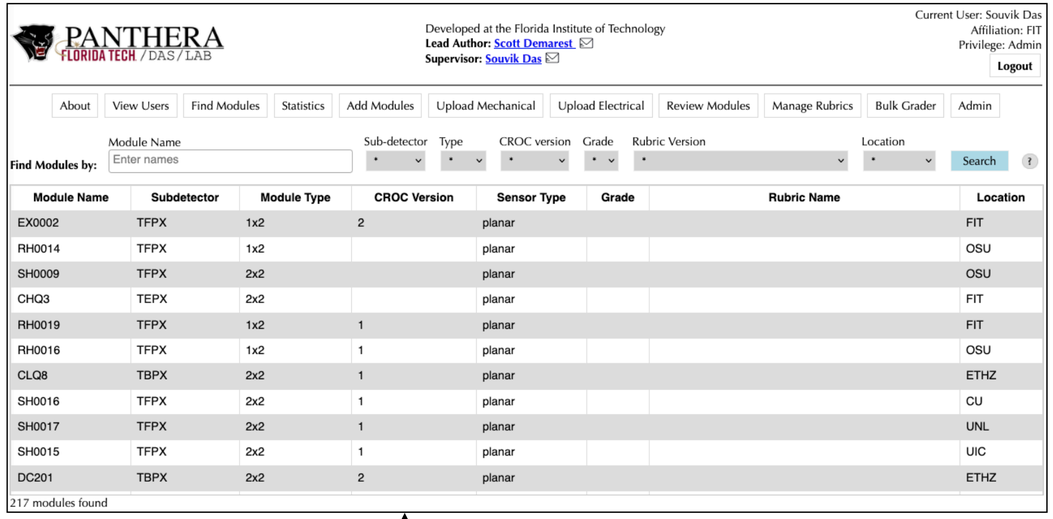
Based on command line ➜ Allows for more debugging capabilities, giving more flexibility to the testing centres

```
17:52:26 INFO -           03  02            05                                                  ,d
17:52:26 INFO -           88  ''            ''                                                  88
17:52:26 INFO -           88                                                                    88
17:52:26 INFO -    ,a6cb9b,88  88  8b,d2622a,  88   ,aaac5b,d8   ,a5842ba,  8b,d9a9ca,  M8841MM
17:52:26 INFO -   a8'    `Y88  88  88P'   'Y8  88  a8'     `Y88  a8P_____88  88P'    ``8a   88
17:52:26 INFO -   8b       88  88  88          88  8b       88  8PP''''''   88       88   88
17:52:26 INFO -   '8a,   ,d88  88  88          88  '8a,   ,d88  '8b,   ,aa  88       88   88,
17:52:26 INFO -    `'23f9P'Y8  88  88          88   `'ea04P'Y8   `'e69e8''  88       88   'Y888
17:52:26 INFO -                                    aa,    ,88
17:52:26 INFO -                                     '7f42dP'                                   master
17:54:59 INFO dirigent_cli - version 3.2.5-13-g54baf67
17:54:59 INFO dirigent_cli - Ph2 ACF v6-08
17:54:59 INFO dirigent_cli - ICICLE 2.0.0-121-g2622aac
17:54:59 INFO dirigent_cli - Felis v1.1.0
17:54:59 INFO calibrate_module - ===== parameter values as configured: =====
17:54:59 INFO calibrate_module - ========================================
17:54:59 INFO performable_scans - N_modules = 2
17:54:59 INFO performable_scans - N_ROCs = [4,4]
17:55:00 INFO performable_scans - === 1/1: physics ===
17:55:35 INFO calibration_context - Checking logs/CMSITminiDAQ000000_physics.log for aborts...
17:55:35 INFO calibration_context - No Ph2 ACF abort detected.
```

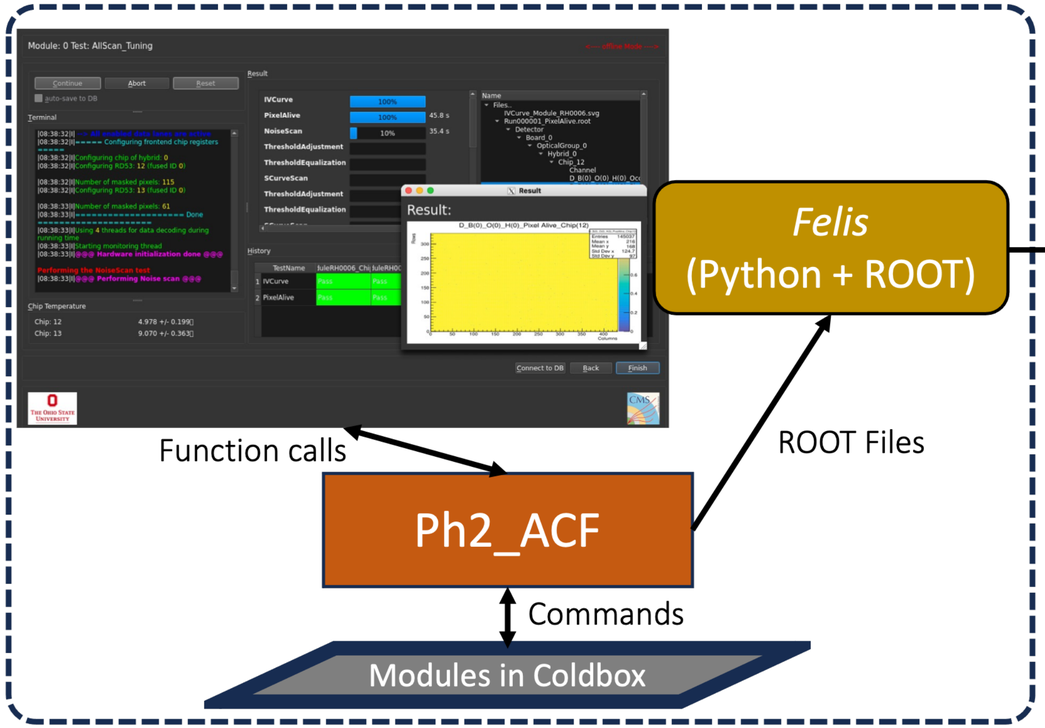
Grading software - Felis/Panthera

`Panthera` (the big cat @ panthera.fit.edu)

- Web-accessible Database & Grader
- Tests entered by GUI/Dirigent, signed by Shifter
- Shift Leader Grades modules by appropriate Rubric
- Query-based search for modules to reveal their history
- Maintains history of multiple Sequences of Tests per module (Plots, `ROOT` files, test conditions)



At Modules Testing & Assembly sites



Felis (Python + ROOT)

`ROOT` files & key numbers over `HTTP`

Web query & rendering

SQL queries

Panthera (PHP)

SQL DB (RAID 6)

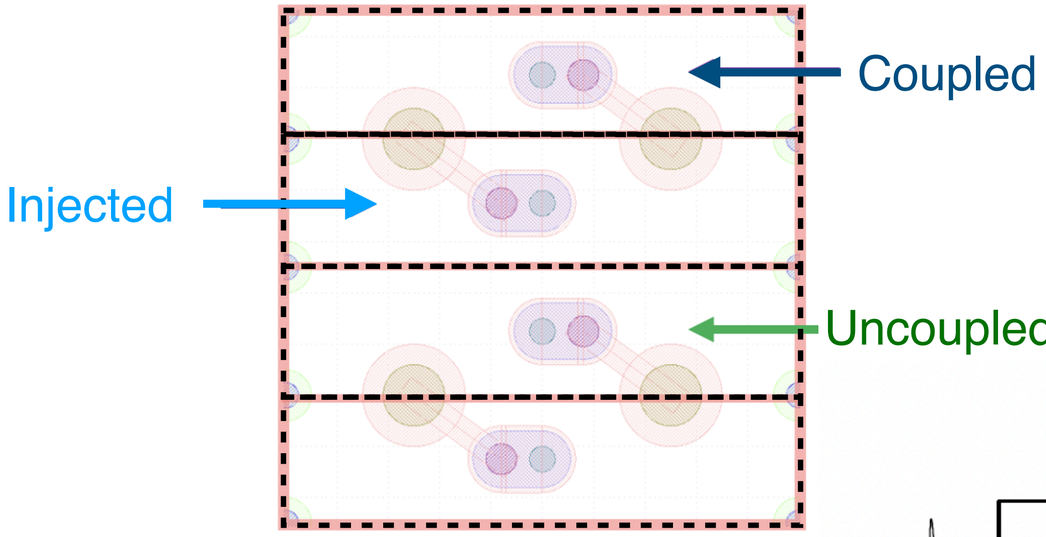
At Florida Tech server

`Felis` (the little cat)

- **`Panthera`** is complemented by **`Felis`**
- Sits at Testing Sites, packaged within GUI/Dirigent
- Analyses Test Results, extracts key numbers for grading and communicates them to `Panthera`
- Creates informative plots

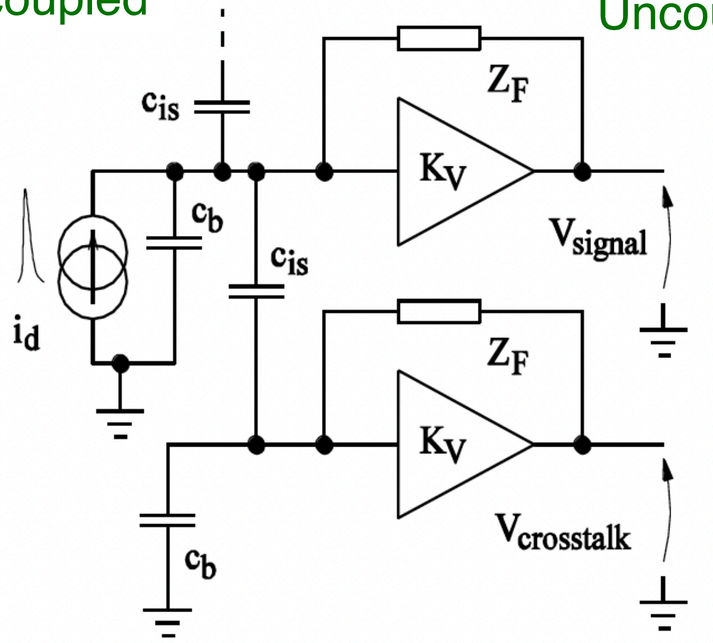
Cross-talk

Pixel cells are capacitively coupled ➜ Cross-talk different in Planar & 3D sensors

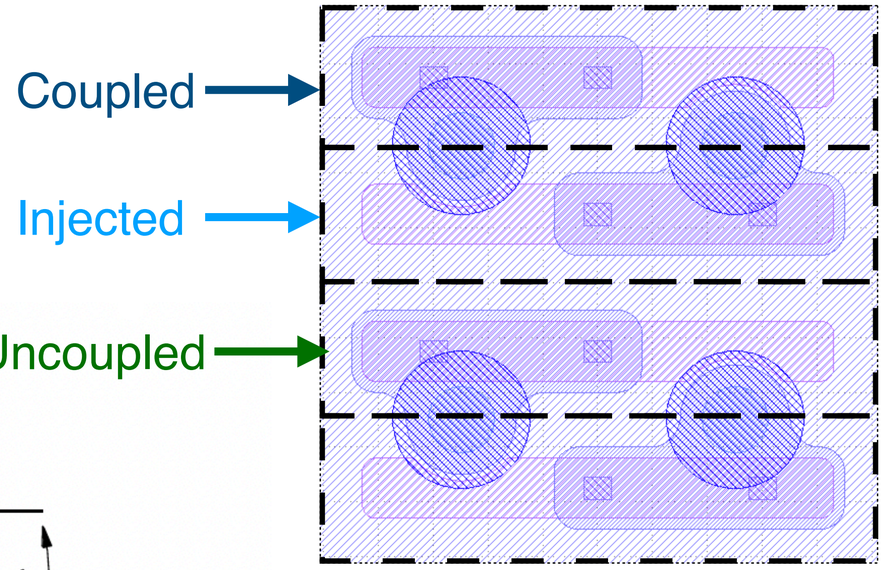


3D sensor 4 cells

Coupling capacitance:
~10 fF (both for Coupled and Uncoupled)



Equivalent circuit



Planar sensor 4 cells

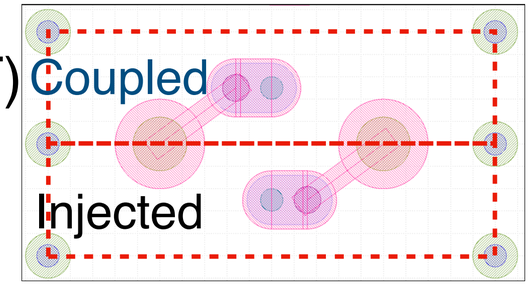
Coupling capacitance:
- **~14 fF** for Coupled
- **~6.5 fF** for Uncoupled

Jan Kaplon, CNRS School of microelectronics, Frejus, 2011

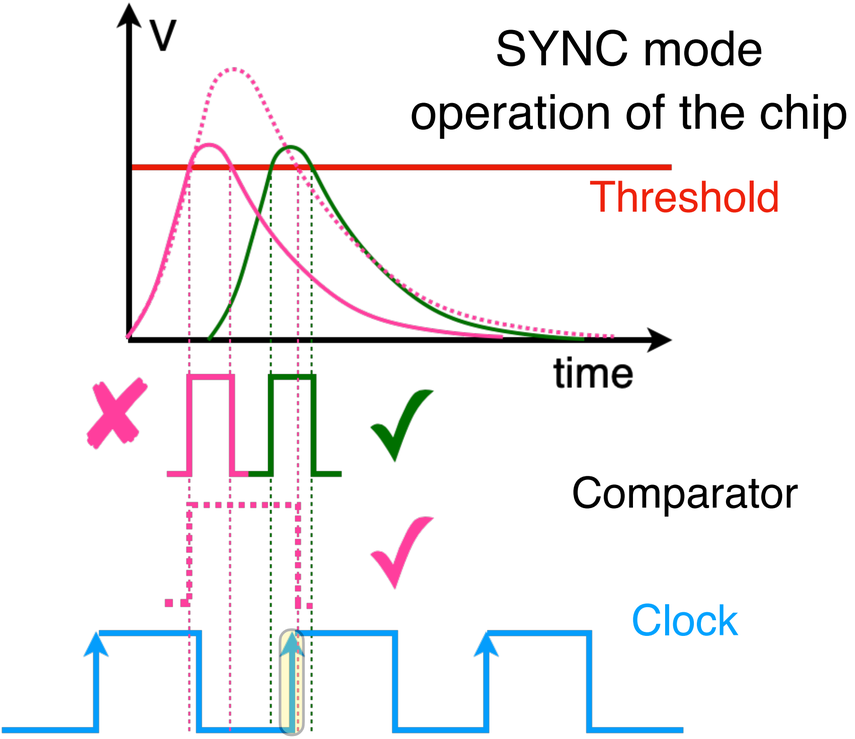
X-talk (**XT**) evaluated by means of **SCurves** ➜ Measured on **Injected** pixel and on **Adjacent** (Coupled and Uncoupled) ones

- **Q_1:** measured threshold of **Injected** pixel ➜ Actual threshold = $Q_1 \cdot (1-XT)$
- **Q_2:** measured threshold of **Adjacent** pixel ➜ Actual threshold = $Q_2 \cdot XT$

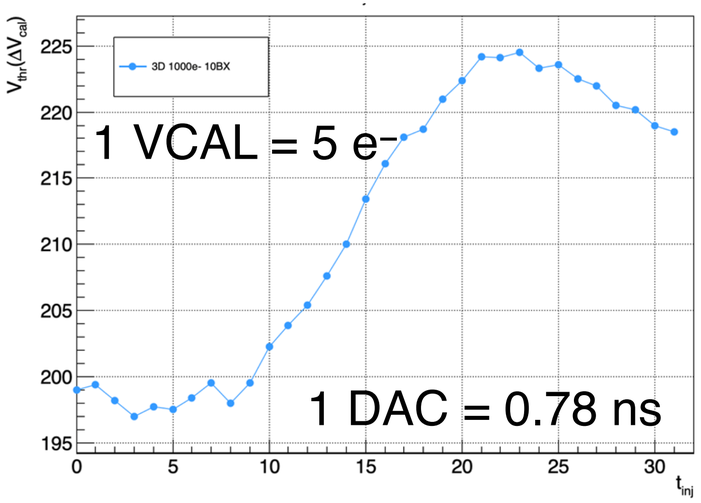
At first order all pixels have same threshold ➜ **$Q_1 \cdot (1-XT) = Q_2 \cdot XT$**



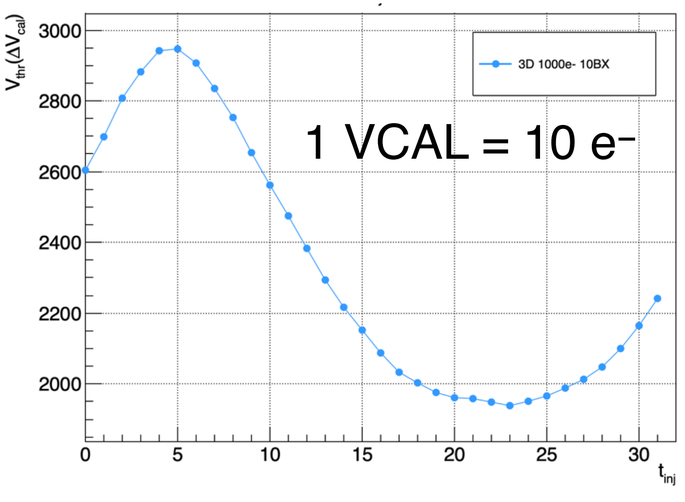
$$XT = \frac{r}{1-r}$$ where r = Q_1/Q_2, and it is evaluated as a function of **injection time**



3D sensor - **Injected**
Threshold vs Injection time

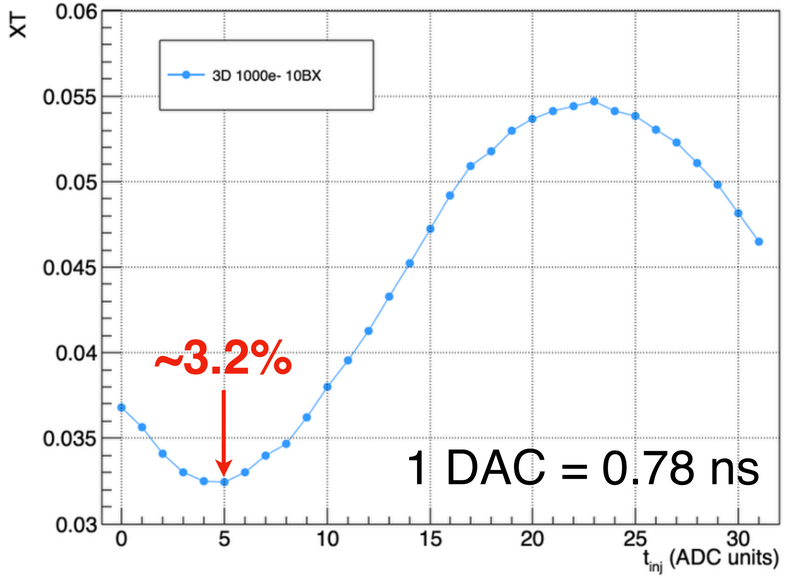


3D sensor - **Coupled**
Threshold vs Injection time

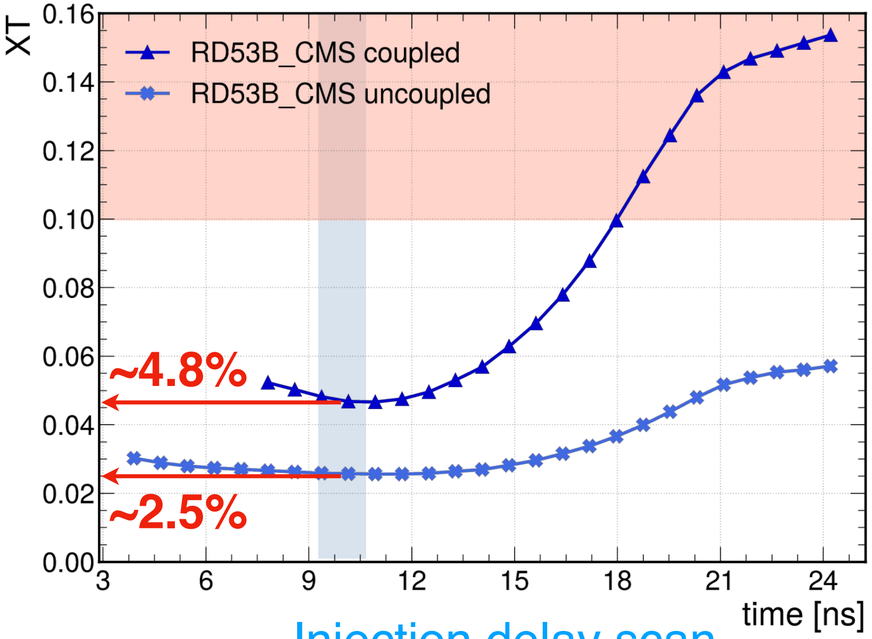


X-talk results for 3D and Planar sensors

XT 3D sensor



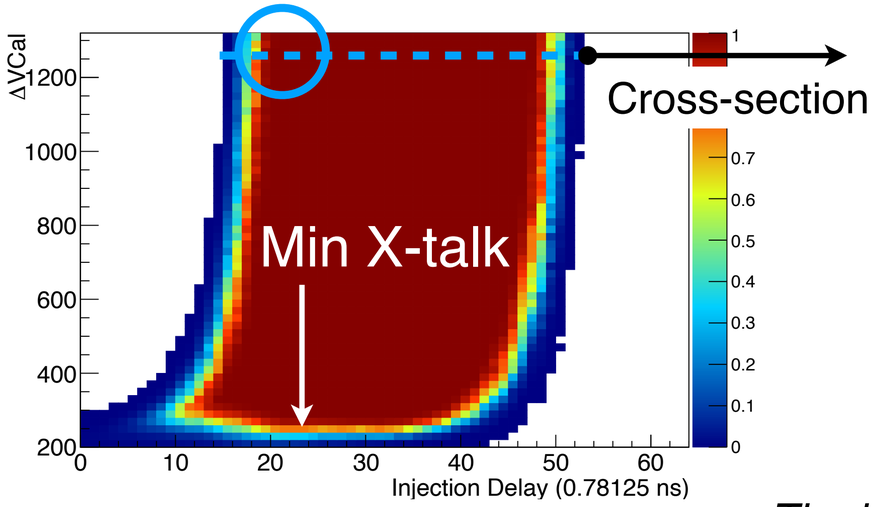
XT Planar sensor



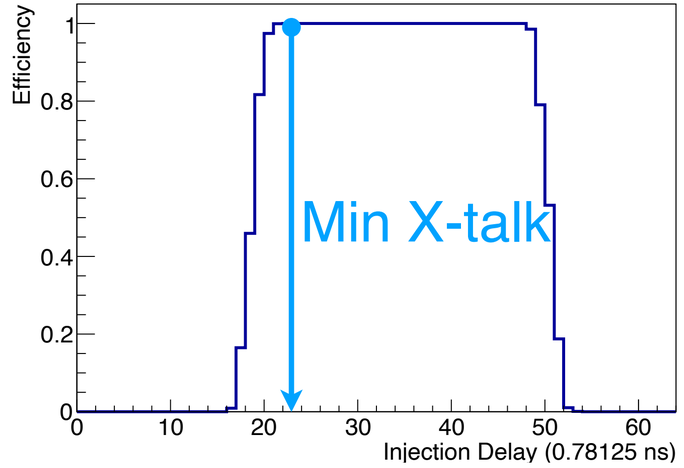
Injection delay scan

How to guess right injection time (minimising X-talk) without need of running time consuming `SCurves`?

➜ **Injection delay scan**



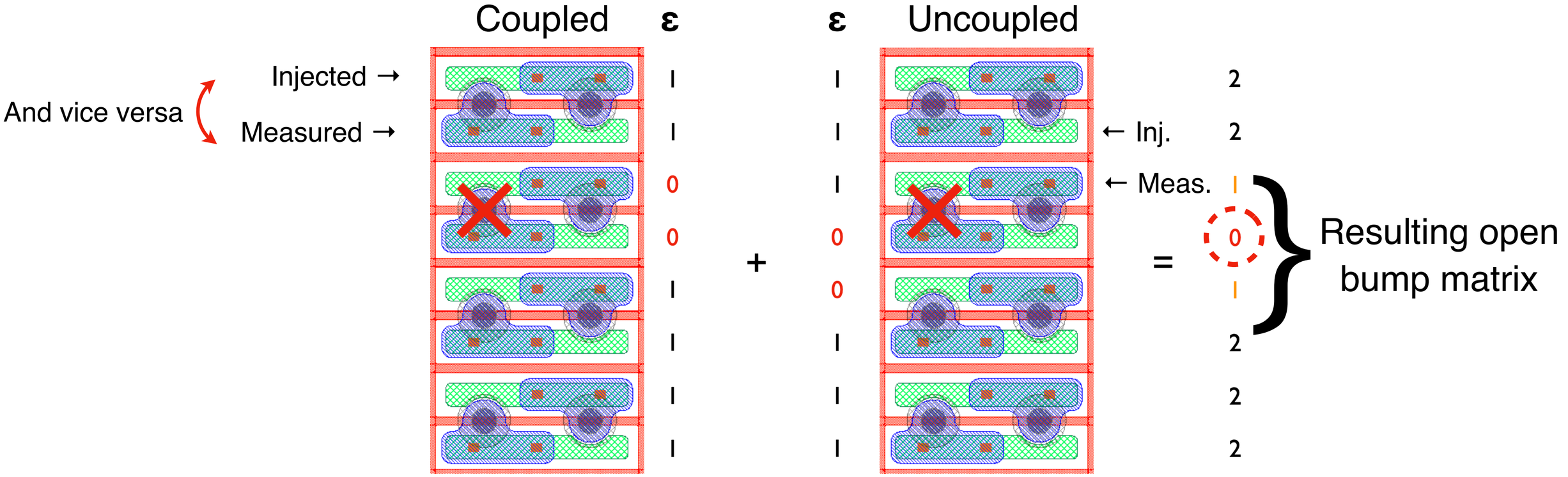
Cross-section



The bottom two plots are just for illustration

X-talk turned out to be **useful** to **detect open bumps**: neighbouring pixels “sense” the induced charge only if the bump is properly connected

- The **procedure** is **fast** because it requires **two runs of `PixelAlive`** ➜ Inject 100 times and measure efficiency per pixel
- **Two runs are needed** to disentangle the ambiguity due to whether the injected or the adjacent pixel has the missing bump



Hit detection efficiency

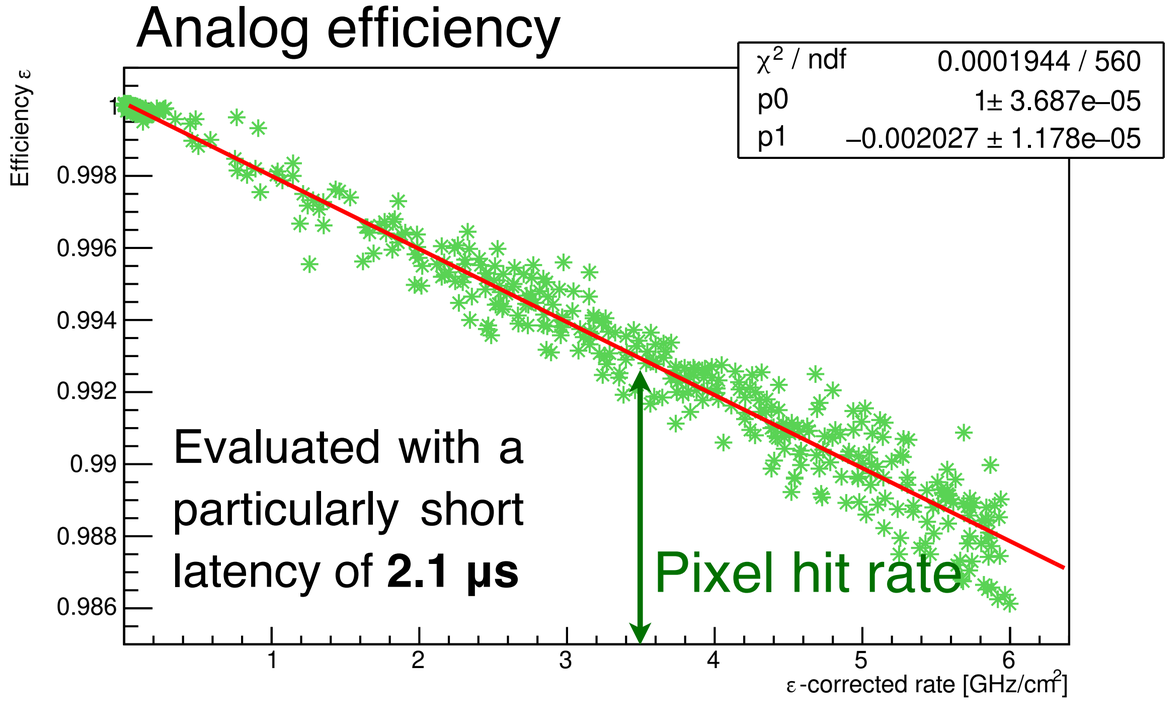
Hit rate test

Two main mechanisms responsible for the degradation of the frontend chip hit detection efficiency: “**analog**” and “**buffering**” inefficiencies

- **Evaluation:** high hit-rate generated by X-ray machine, superimposed to hits from internal injection circuit, whose timing and number is very well controlled
- **Estimation:** fraction of injections successfully detected by the chip

$$\epsilon = \frac{n_{\text{detected injections}}}{n_{\text{total injections}}}$$

To register two separate hits within the same pixel the first hit must return below threshold before the second comes in ➜ If this doesn’t happen ➜ **Analog inefficiency** ➜ Mitigated by adjusting the return to baseline of the Time-over-Threshold (ToT) charge measurement

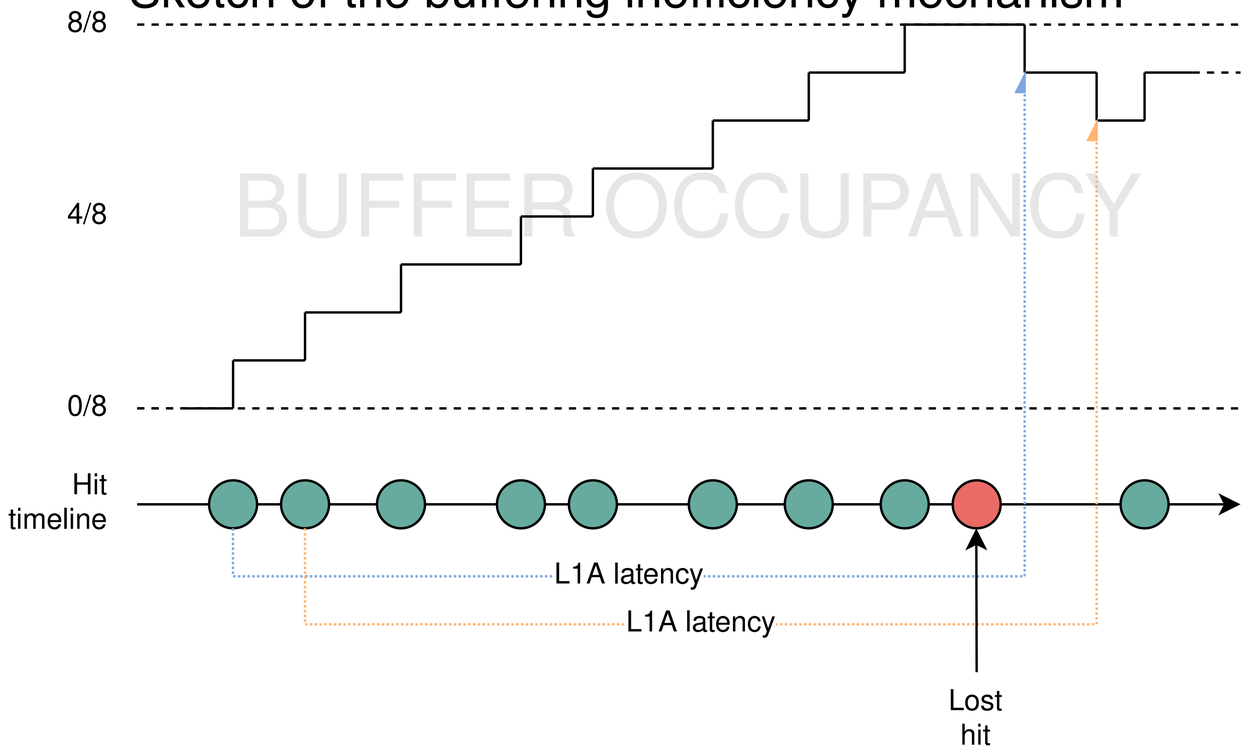


Fabio Luongo, Doctoral thesis, University of Turin, Italy (2025)

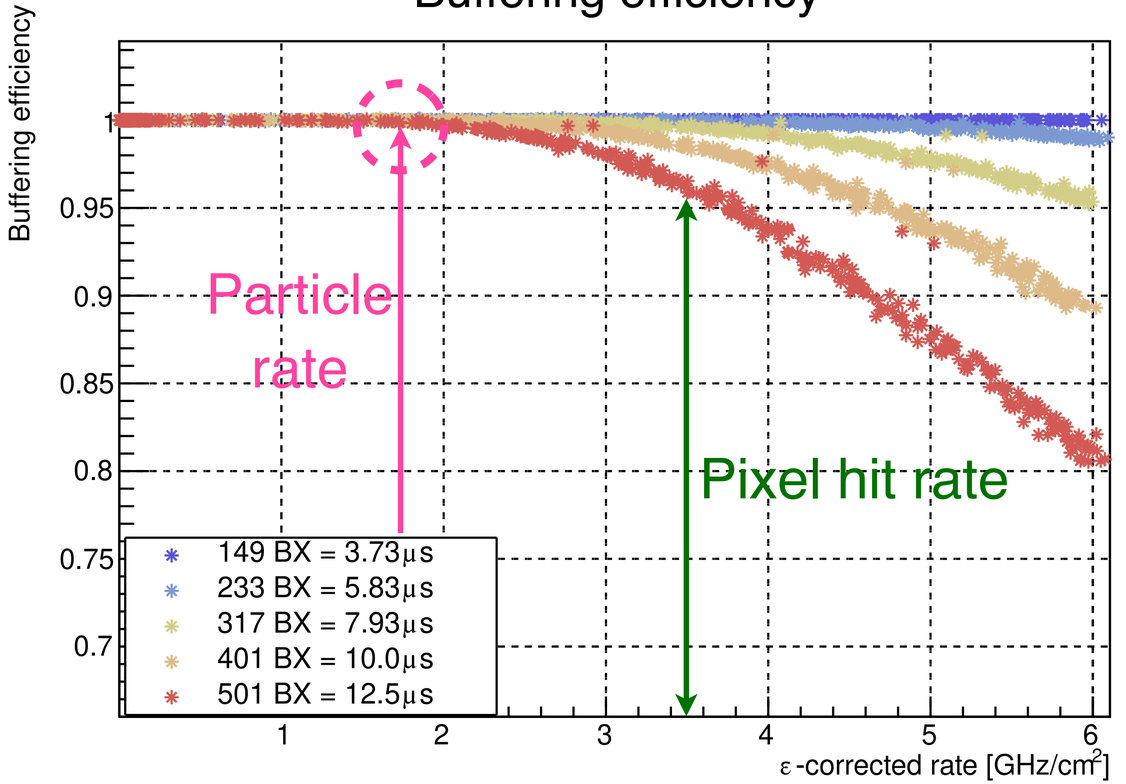
Hit rate test

The trigger for a particular bunch crossing is delayed by some fixed, Level-1-Accept (L1A) latency ➜ Hit data from pixels must be temporarily stored inside buffers for an equally long amount of time, after which they can be safely discarded ➜ Any hit detected while their buffer is full will be lost ➜ **Buffering inefficiency**

Sketch of the buffering inefficiency mechanism



Buffering efficiency



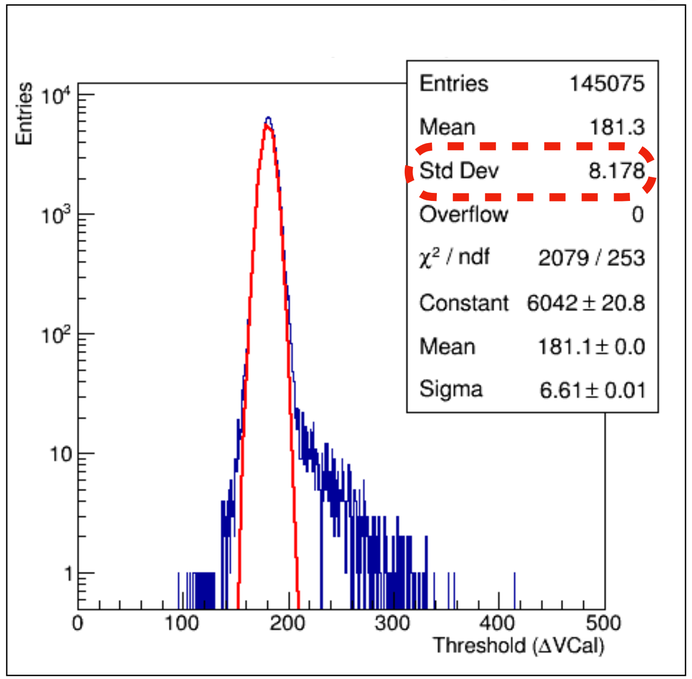
Fabio Luongo, Doctoral thesis, University of Turin, Italy (2025)

Typically particles generate clusters of hits ➜ The actual particle rate is 1.78 GHz/cm²

How to run cold

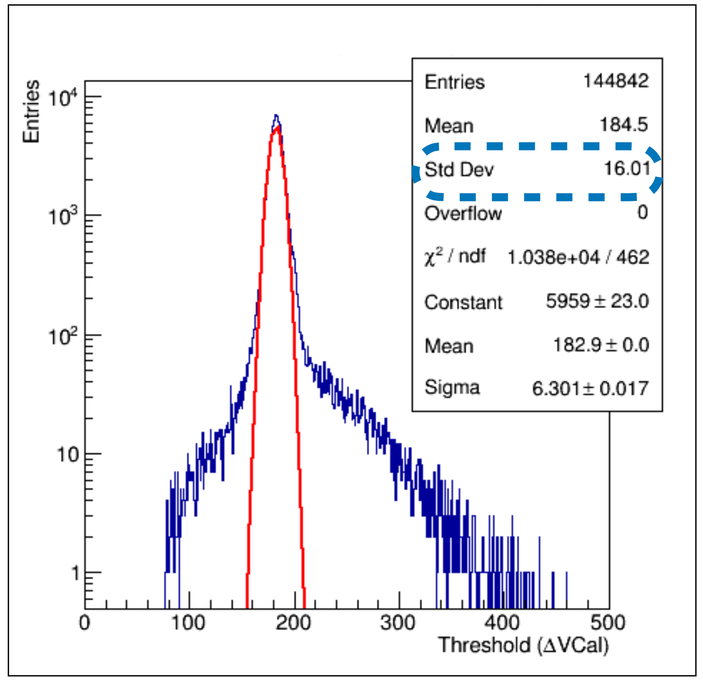
Running cold

T = **30°C** (T_{CO2} = 15°C) **Adj. range = 140**

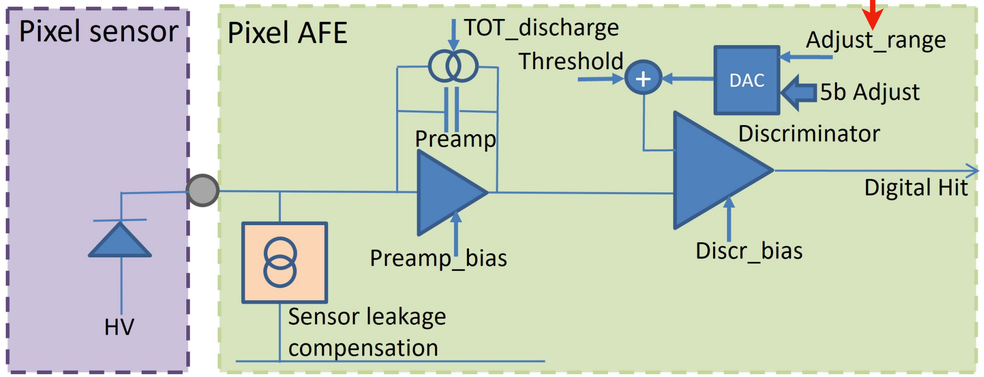


Threshold distributions after threshold equalisation

T = **–14°C** (T_{CO2} = –31°C) **Adj. range = 140**



The readout chip has 5-bits threshold trimming per pixel, together with a global register to **Adjust their range (LDAC)**

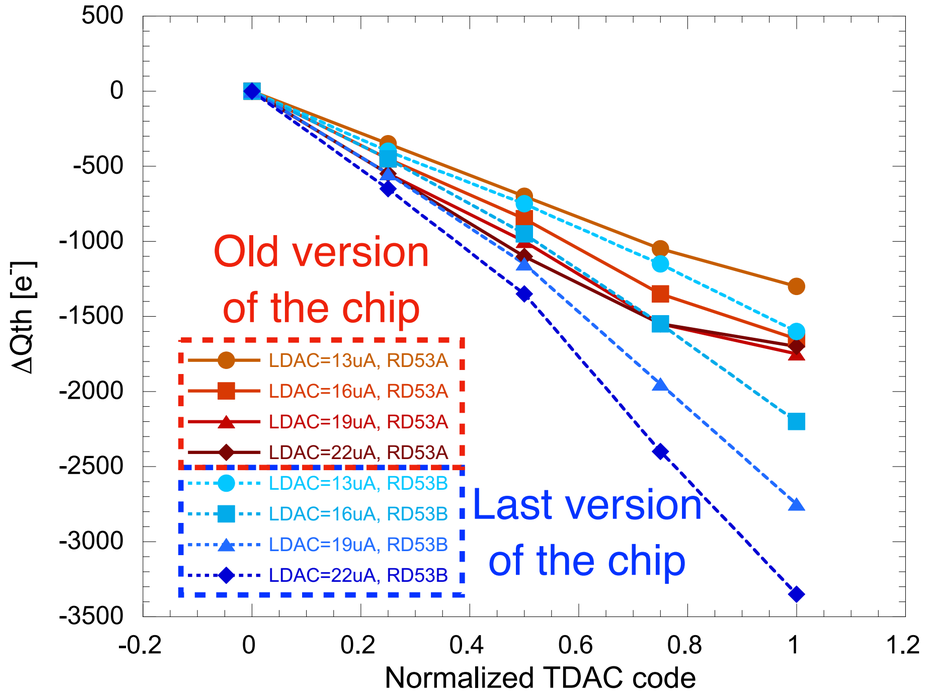


JINST (2025) 20 P03024

Threshold dispersion and **Mean threshold** for various readout chips (x-axis) at different

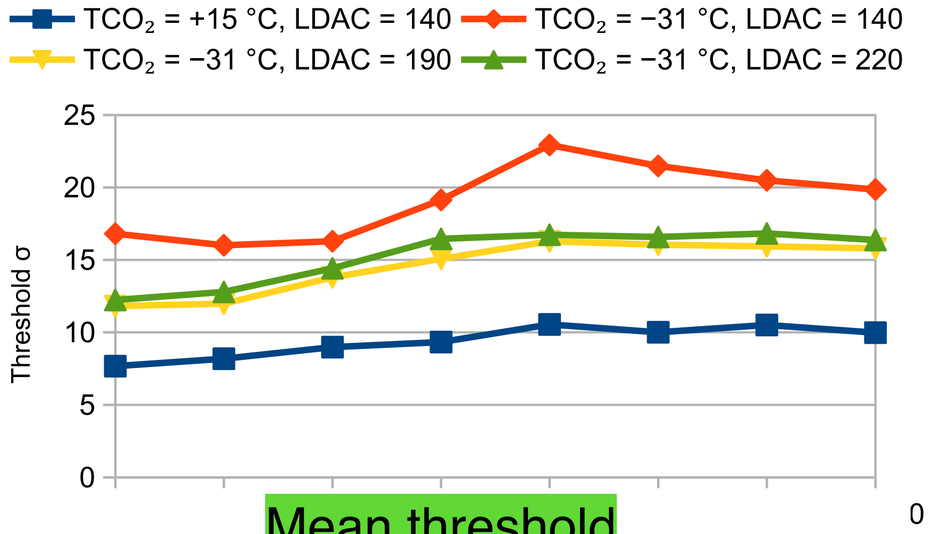
- Temperatures
- Adj. range (LDAC) values

Threshold variation vs Per pixel trim bits (TDAC) for different **Adj. range** (**LDAC**) values

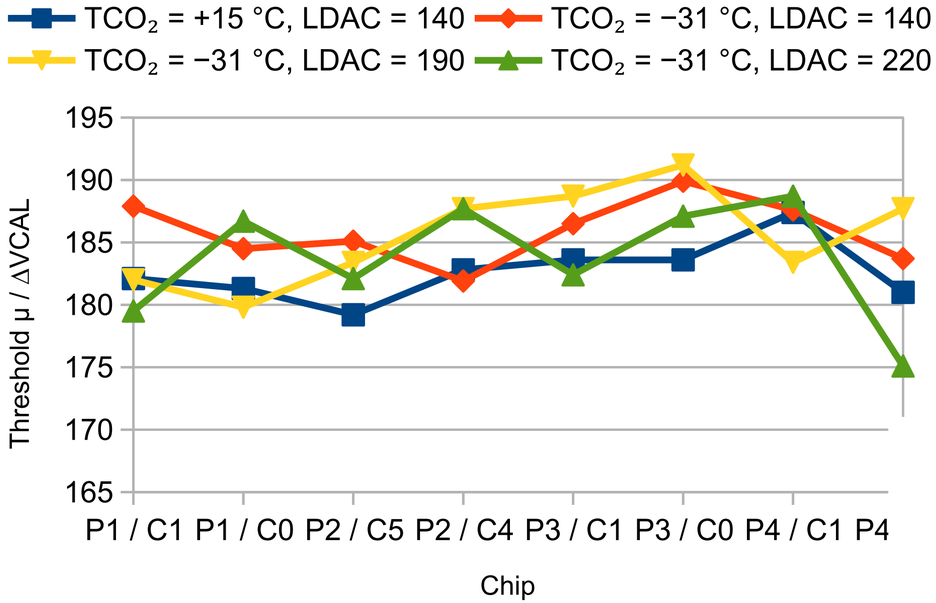


IEEE TNS, vol. 68, no. 11, pp. 2682-2692 (2021)

Threshold dispersion



Mean threshold



How to run an irradiated module

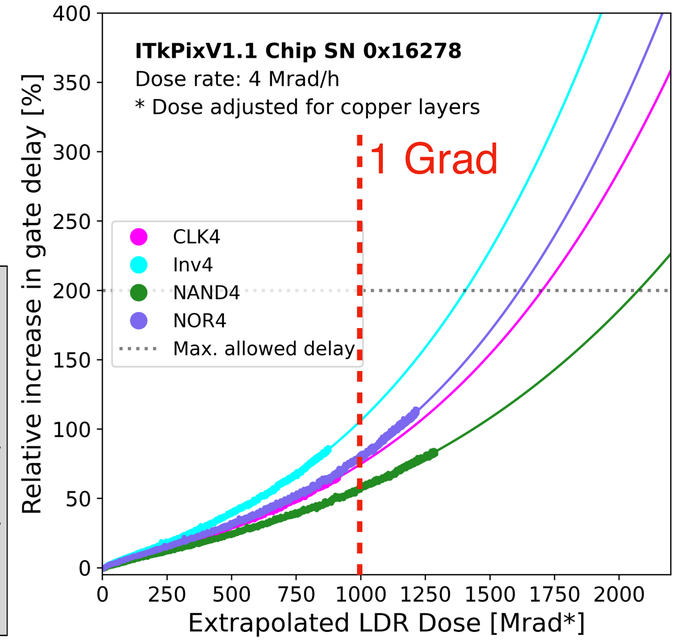
How to run an irradiated module

The **Inner Tracker readout chip** is made in 65 nm CMOS technology, radiation tolerant up to 100 Mrad ➜ Further precautions were adopted to make it more **rad. hard up to 1 Grad**

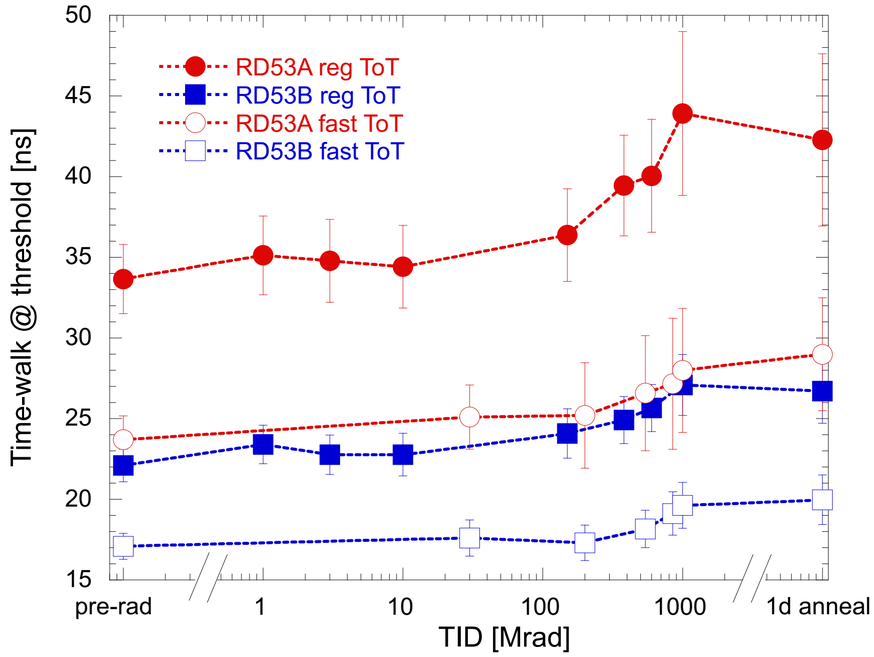
- Wider and longer transistors
- Triple Modular Redundancy
- Limit annealing periods ➜ Run cold (below –10°C)
- Homogenous Copper vias thickness of current mirrors

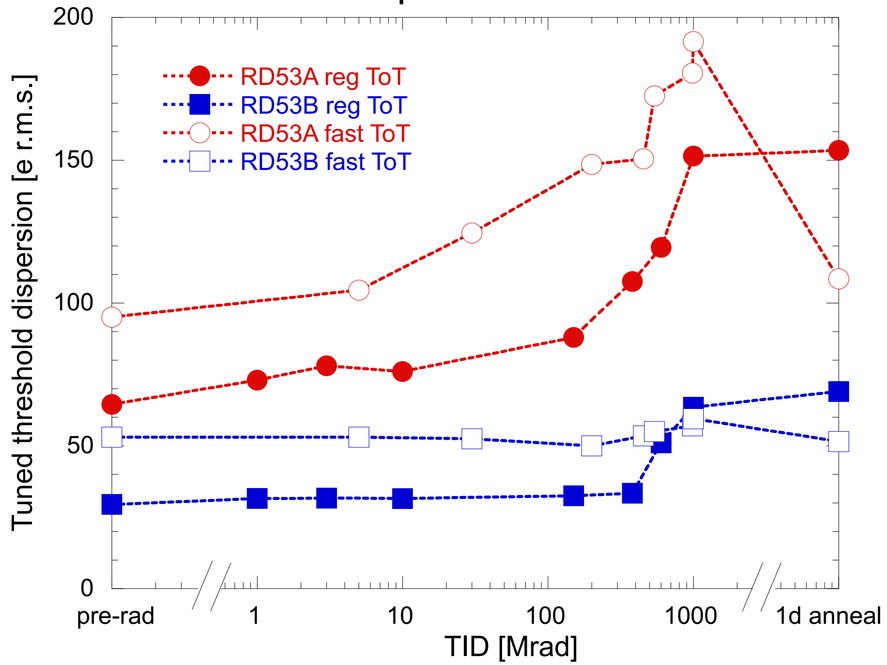
- **RD53A:** old version of the chip
- **RD53B:** last version of the chip
- Regular ToT: 6000 e⁻ corresponds to 130 ns ToT
- Fast ToT: 6000 e⁻ corresponds to 60 ns ToT

JINST (2025) 20 P03024



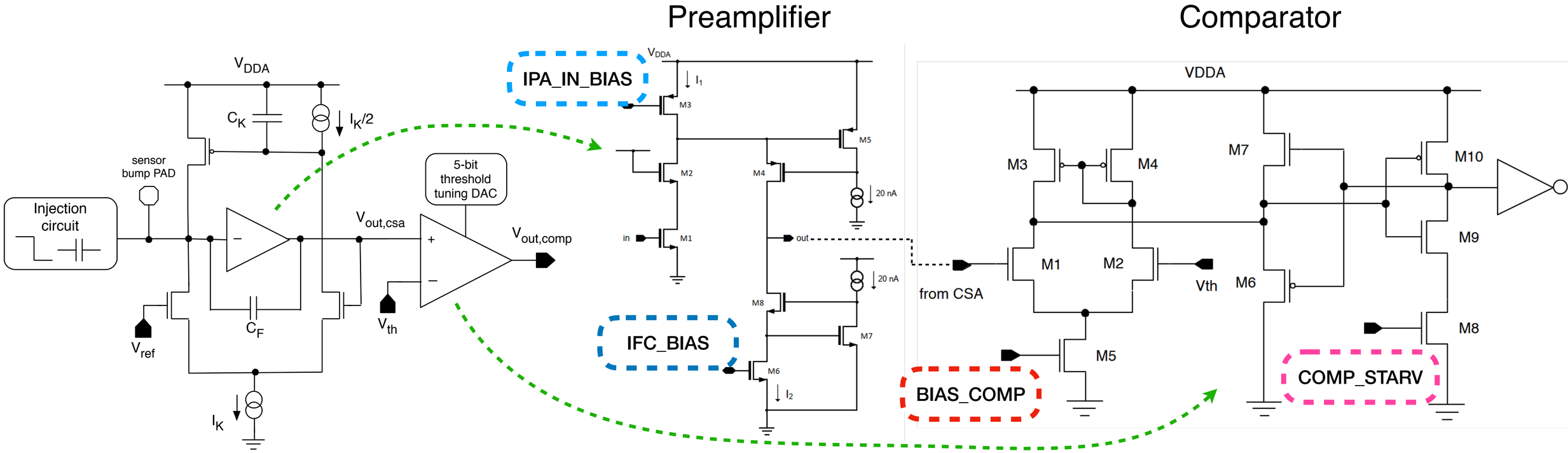
IEEE TNS, vol. 68, no. 11, pp. 2682-2692 (2021)





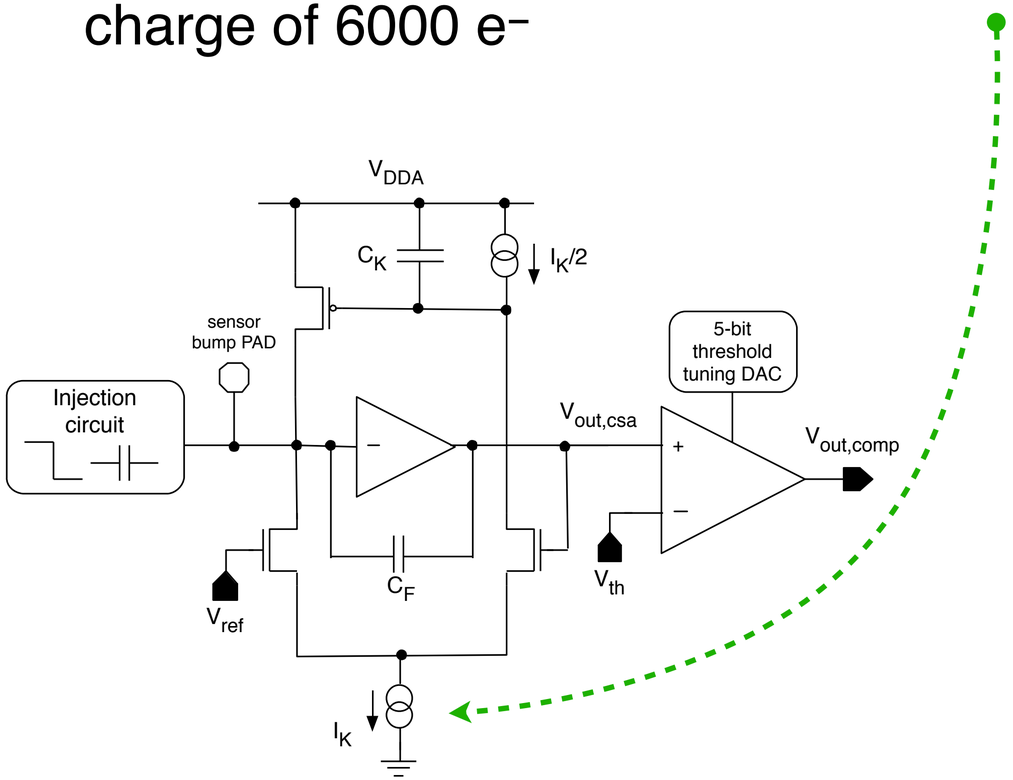
Optimal tuning procedure for irradiated chips

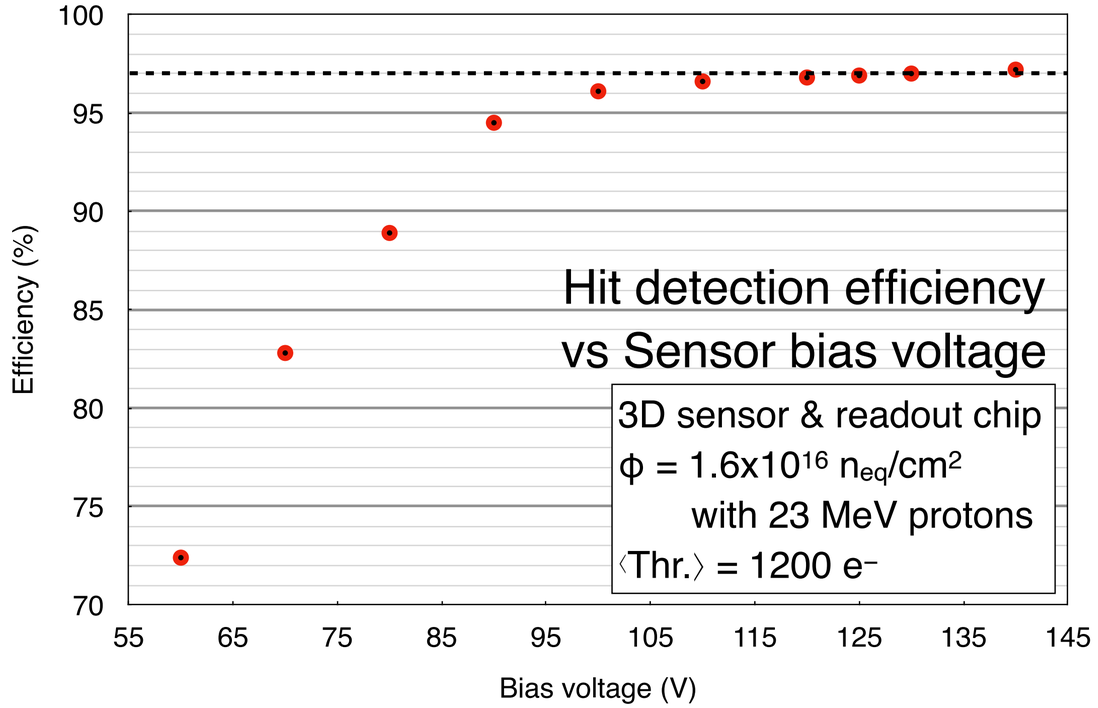
1. Set threshold to a high value (4500 e^-) and all threshold trim bits to 0
2. Set **BIAS_COMP** = 0 ➜ Measure analog current ➜ **Tune BIAS_COMP** to optimal current consumption
3. Set **IPA_IN_BIAS** = 0 ➜ Measure analog current ➜ **Tune IPA_IN_BIAS** to optimal current consumption
4. Set all all threshold trim bits to 16 ➜ **Tune LDAC** to optimal current consumption



5. **Rescale** `IFC_BIAS` and `COMP_STARV` to restore the appropriate biasing of the secondary branches of the pre-amplifier and the comparator:
   - `IFC_BIAS*` = `IFC_BIAS` x `BIAS_COMP*` / `BIAS_COMP`
   - `COMP_STARV*` = `COMP_STARV` x `BIAS_COMP*` / `BIAS_COMP`

   `BIAS_COMP*` is the value at step 2. while and `BIAS_COMP` is the initial value, before the tuning
6. **Tune I_k** to ⟨ToT⟩ = 5.3 (for slow discharge) or ⟨ToT⟩ = 2 (for fast discharge) injecting a charge of 6000 e^-

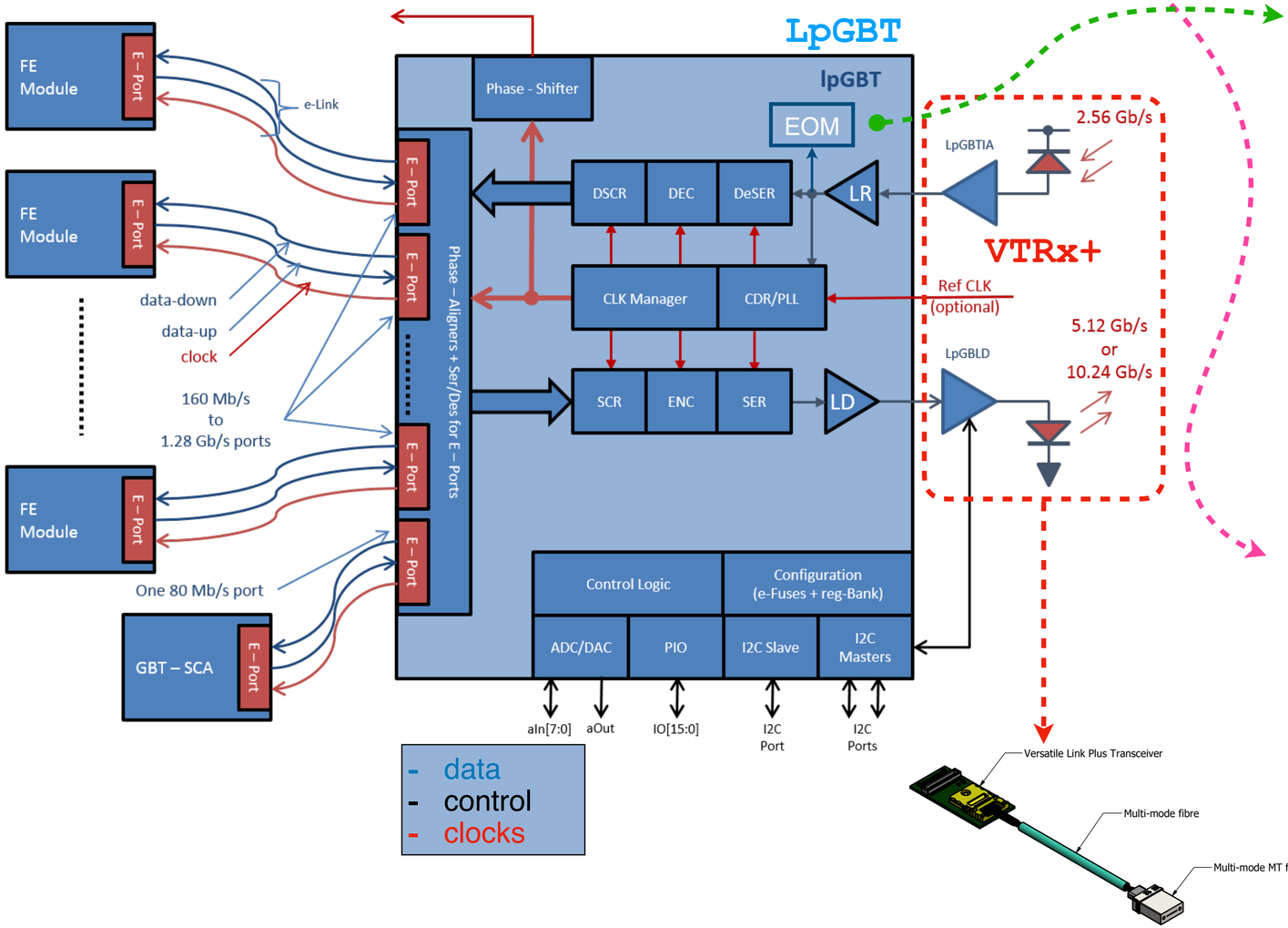


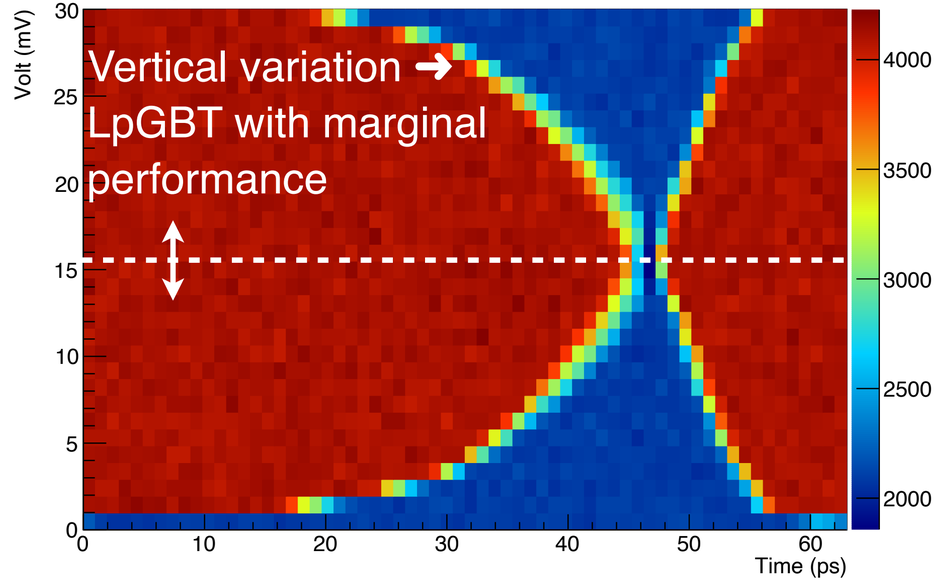


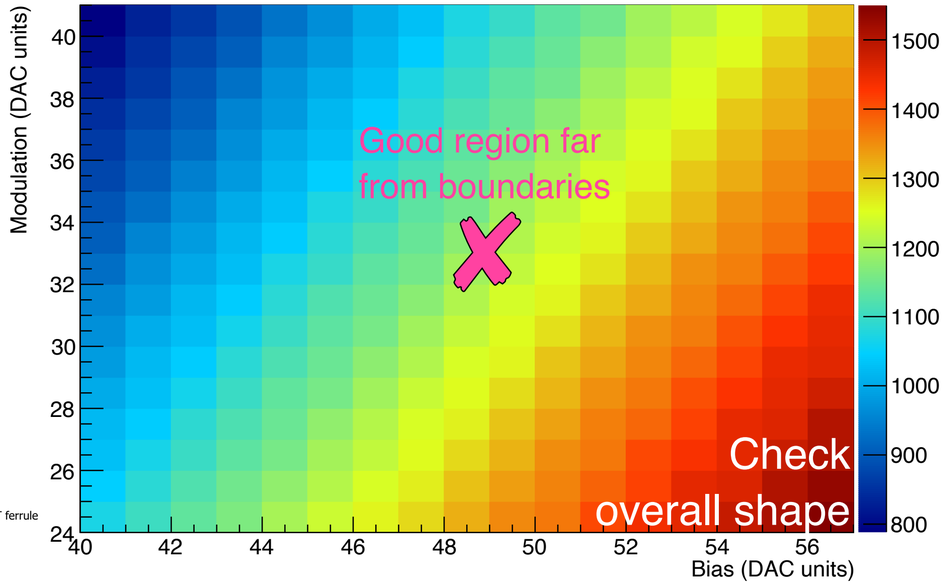
Optical connectivity

The performance of the optical connectivity can be monitored by means of two plots

- **LpGBT** eye opening monitor (EOM) ➜ **Down-link**
- **VTRx+** Bias vs Modulation plot ➜ **Up-link**







Summary

The CMS **Inner Tracker** detector was designed to face the **challenges** imposed by HL-LHC

- Limited power budget: 1 W/cm^2
- Radiation hardness up to 2x10^{16} n_{eq}/cm^2 and 1 Grad
- Trigger rate: 750 kHz
- Hit rate: 3.5 GHz/cm^2
- Latency: 12.5 μs

…

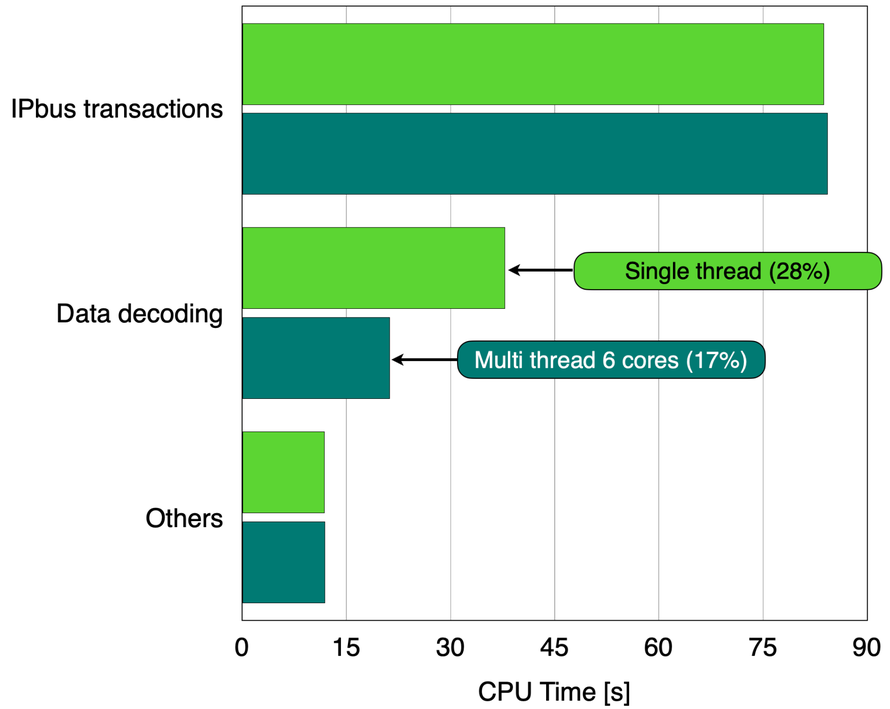
➜ A complex **readout chip** was finally implemented that **meets all requirements**

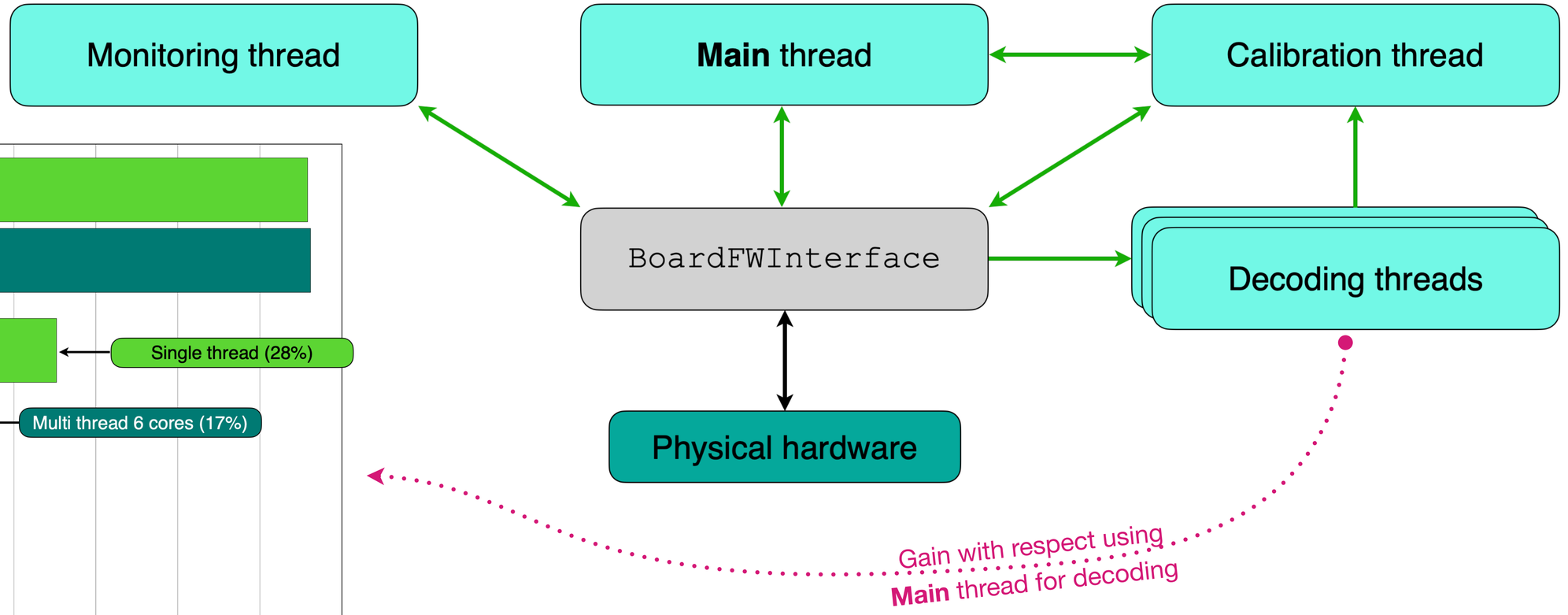
We are almost in the **production phase** where all the aspects of the pixel detector are going to be evaluated, optimised, and eventually rated through an **ad-hoc readout system**

We have here emphasised a few key features:

- X-talk
- Hit detection efficiency
- How to run cold
- How to run an irradiated module
- Optical connectivity checkout

Backup slides





Example (*noise measurement*) of running-time gain thanks to data decoding parallelisation:

- **decoding** is **second** most **time-consuming** operation, we **gained ~10%** on a 6 core CPU
- **IPbus transaction** is **first** most **time-consuming** operation, which can only be improved once we move to the SoC

- **Calibration thread:** runs a calibration while listening to commands like stop, pause, etc…
- **Decoding thread:** bit-unpacking and interpretation are intrinsically parallelizable tasks because data are logically and physically separated by triggers
- **Monitoring thread:** monitor frontend hardware asynchronously with respect to other tasks

Optimal tuning procedure for irradiated chips

1. Set threshold to 4500 e⁻ and all TDAC (threshold trim bits) to 0
2. Set **BIAS_COMP** = 0 and measure the analog current $\mathbf{I^{(0)}_A}$ ➜ **Tune BIAS_COMP** to the value resulting in a total current consumption:
   $I_A^{(1)} = I_A^{(0)} + 319\ \mathrm{mA}$ ➜ **319 mA** corresponds to 2.2 µA per pixel
3. Set **IPA_IN_BIAS** = 0 and measure the analog current $\mathbf{I^{(2)}_A}$ ➜ **Tune IPA_IN_BIAS** to the value resulting in a total current consumption:
   $I_A^{(3)} = I_A^{(2)} + 435\ \mathrm{mA}$ ➜ **435 mA** corresponds to 3 µA per pixel
4. Set all TDAC = 16 and **tune LDAC** to the value resulting in a total current consumption:
   $I_A^{(4)} = I_A^{(3)} + 58\ \mathrm{mA}$ ➜ **58 mA** corresponds to 400 nA per pixel
5. **Rescale IFC_BIAS** and **COMP_STARV** to restore the appropriate biasing of the secondary branches of the pre-amplifier and the comparator:
   - **IFC_BIAS*** = **IFC_BIAS** x **BIAS_COMP*** / **BIAS_COMP**
   - **COMP_STARV*** = **COMP_STARV** x **BIAS_COMP*** / **BIAS_COMP**

   BIAS_COMP* is the value at step 2. while and **BIAS_COMP** is the initial value, before the tuning
6. **Tune I_k** to an ⟨ToT⟩ = 5.3 (for slow discharge) or ⟨ToT⟩ = 2 (for fast discharge) injecting a charge of 6000 e⁻