

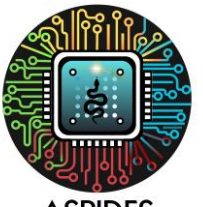
Ultra-fast photon counting with monolithic digital SiPMs in a 110 nm CMOS technology

FEE 2026, XIII FRONT-END ELECTRONICS

Tommaso Floris on behalf of the APSIDES collaboration

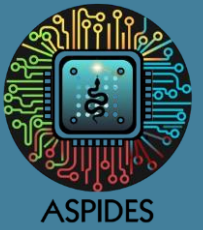


UNIVERSITÀ
DI PAVIA



ASPIDES

Outline



Applications

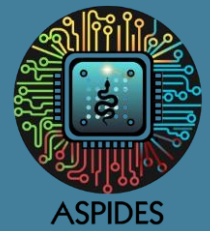
Features

Architecture overview

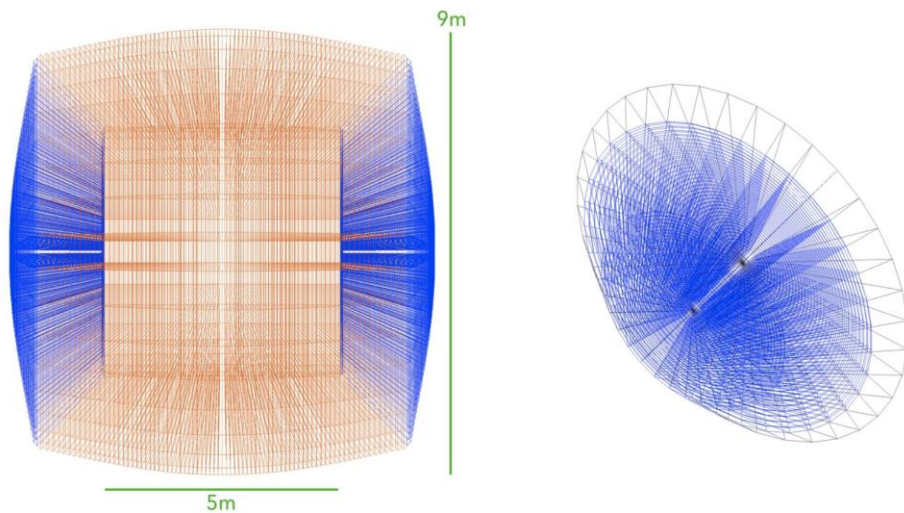
Device operation

Conclusion

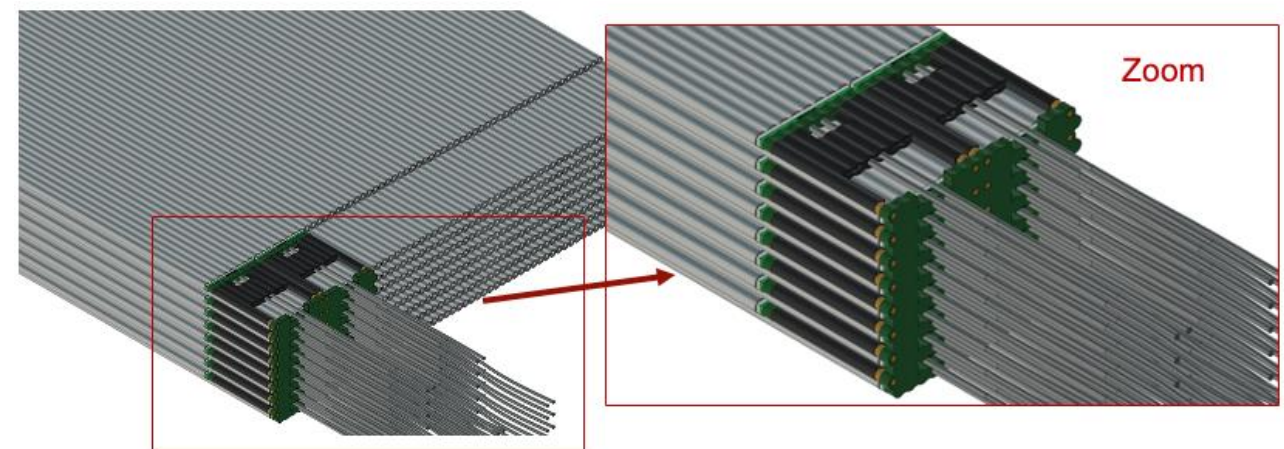
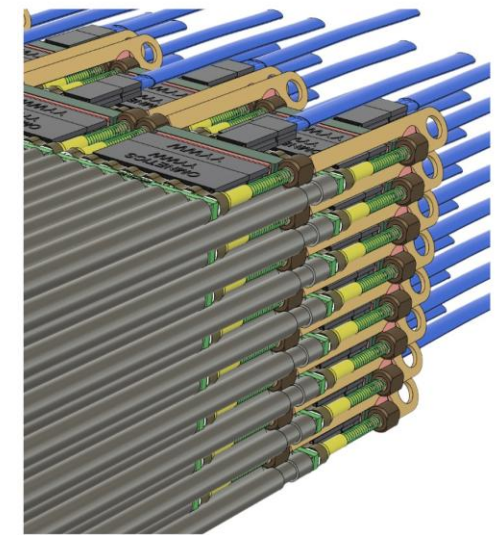
Application: dual-readout calorimeter



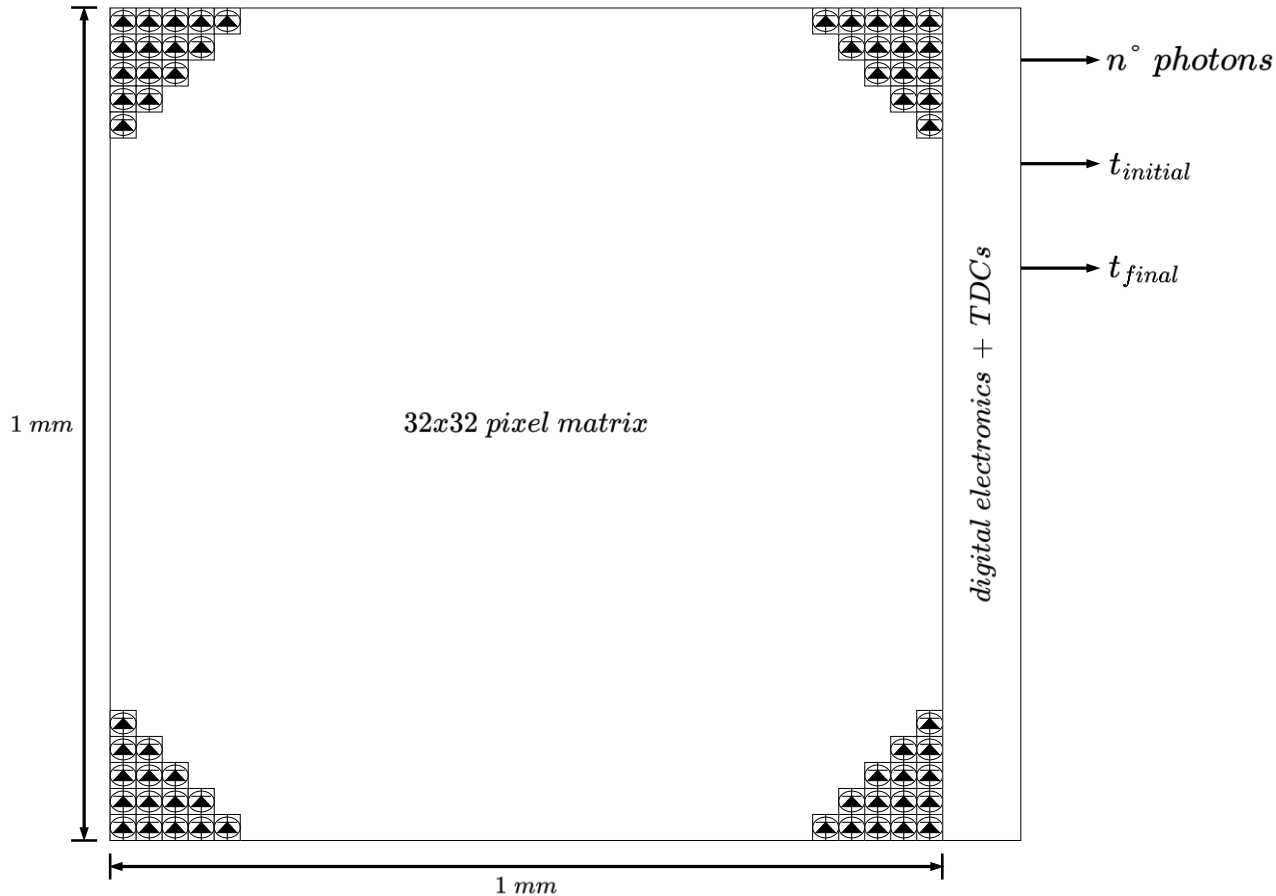
- Almost **75 millions** of 2 mm outer diameter stainless steel tubes.
- In each tube there is a 1 mm diameter fiber connected to a SiPM.
- Signals from 8-SiPMs grouped to reduce the number of channels to be read out.



dual-readout calorimeter for IDEA

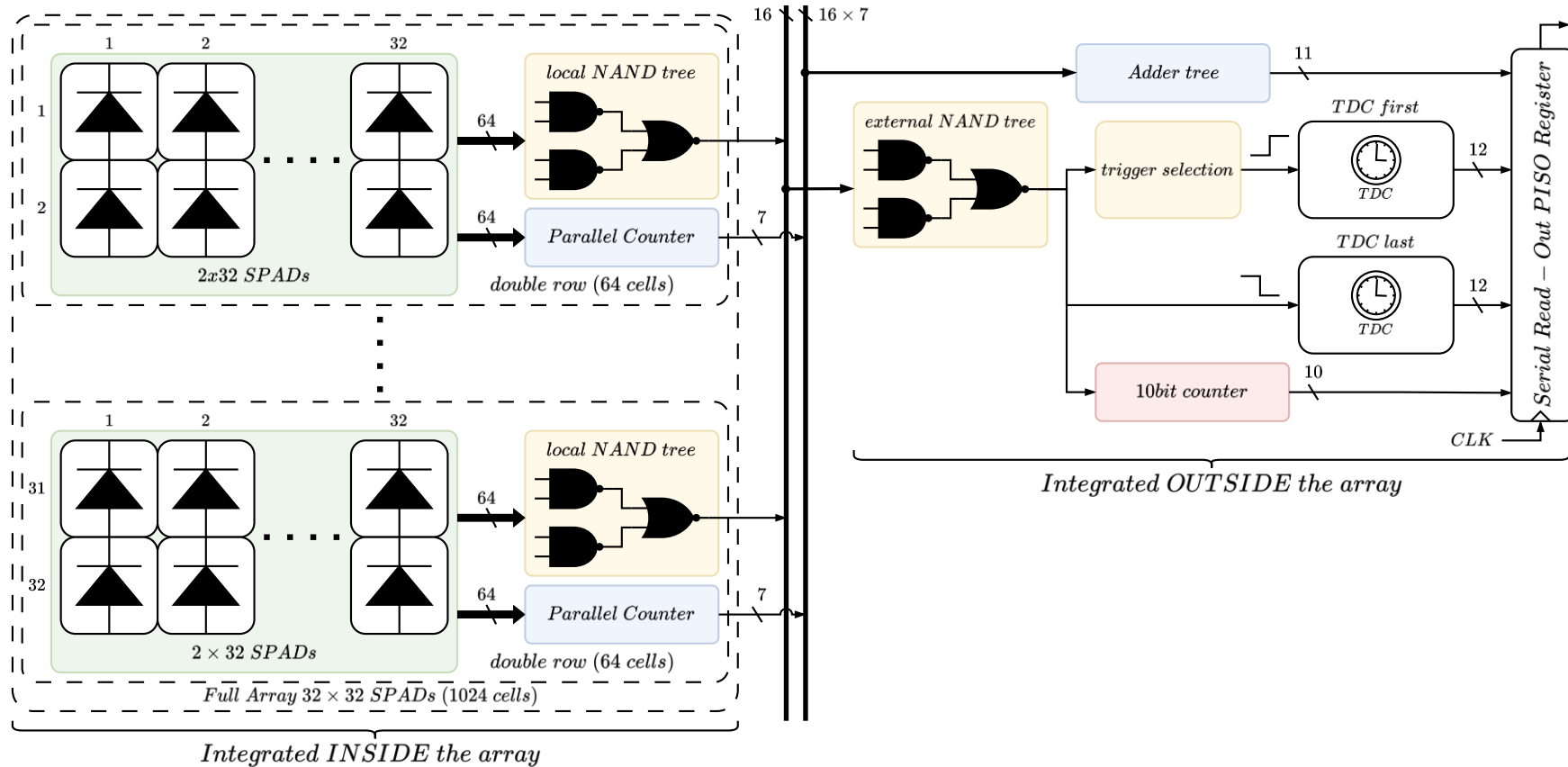
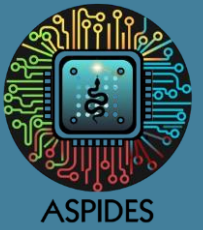


Digital SiPM Features



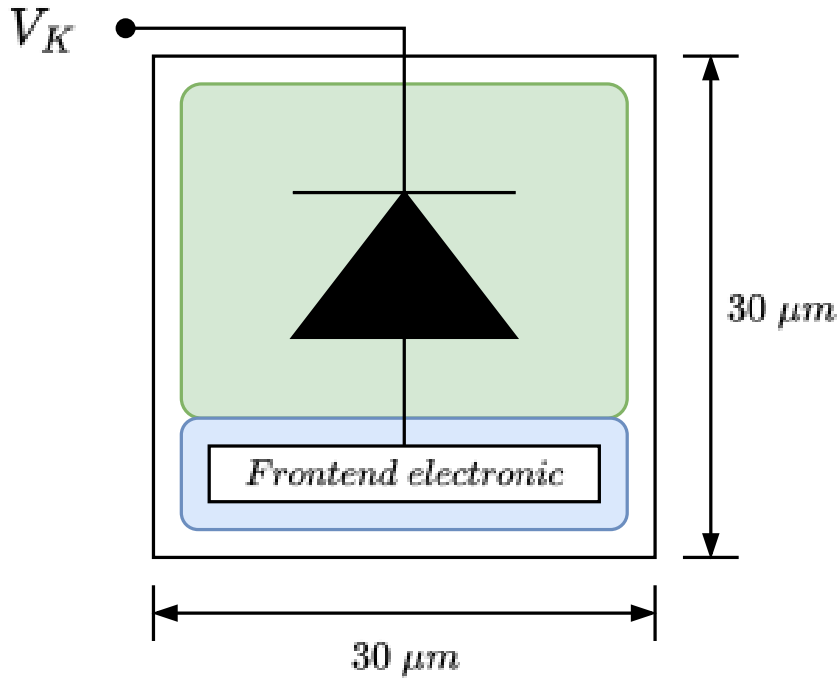
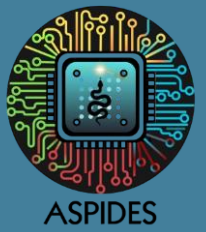
- **Size:** 1024 pixels, 1 mm²
- **Technology:** 110 nm SPAD-compatible CMOS
- **Array Functions:**
 - Hit pixels counting
 - Measurement of Time of Arrival (ToA) for **first** and **last** photons with a sub-100ps resolution
- **Fill Factor Optimization:**
 - Move part of the electronics outside the pixel matrix
 - Trade-off between **in-pixel integration** and **FF maximization**
- **Performance Goals:**
 - **Fast counting** capability
 - Accurate **time stamping** of first and last detected photons
- **Readout System:**
 - On-board readout electronics
 - Compact and efficient serial data transmission

Architecture

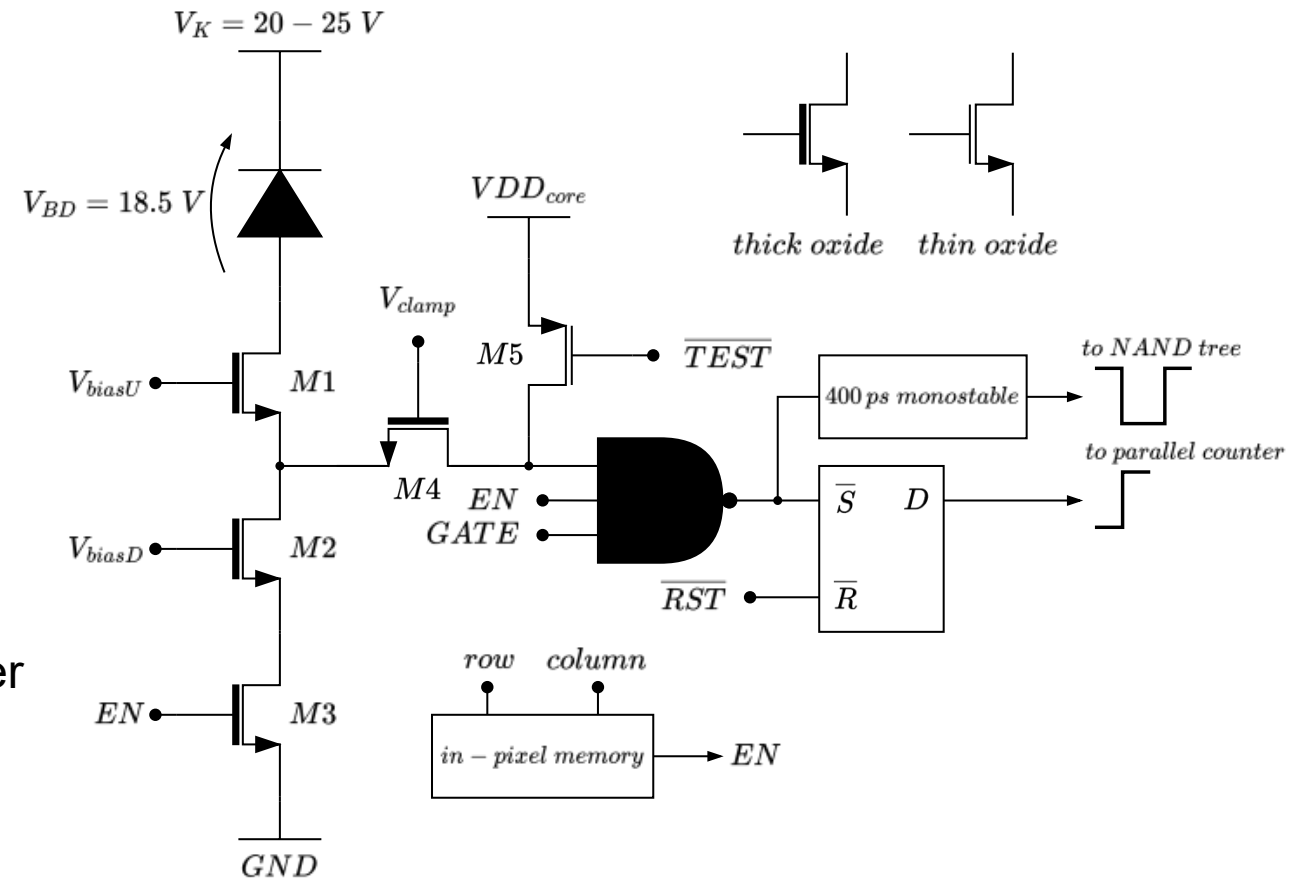


The core of the architecture is a 1024 SPAD array organized in 32 rows and 32 columns.

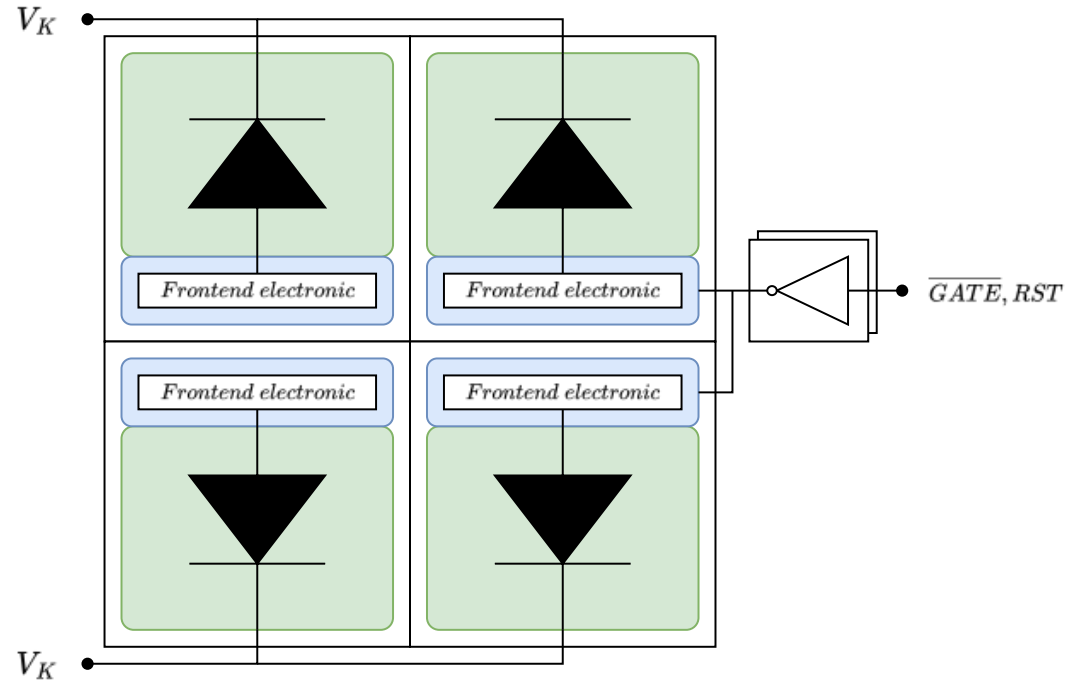
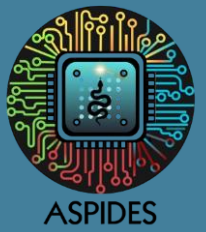
Pixel



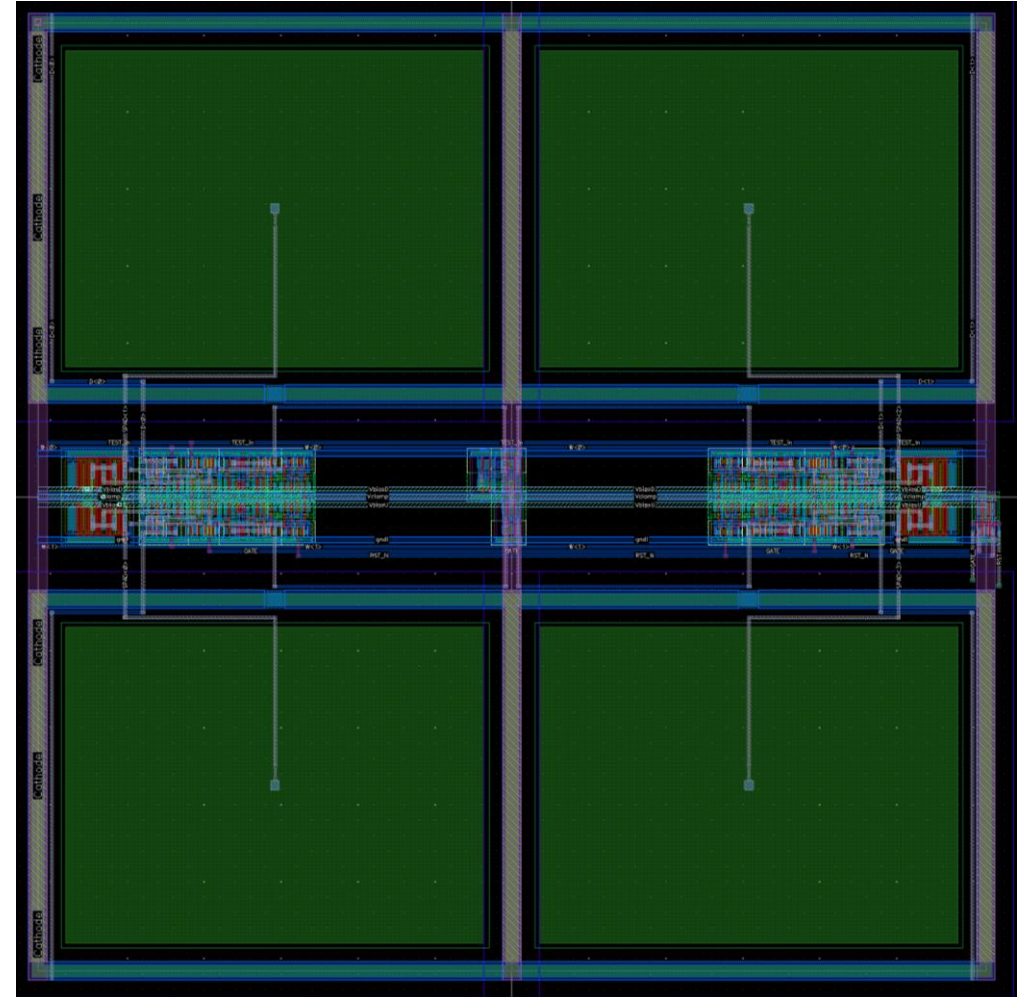
- Cascode in the quenching network to sustain higher over-voltage.
- Possibility to disable the single pixel.
- Gate signal to control simultaneously all the pixels.



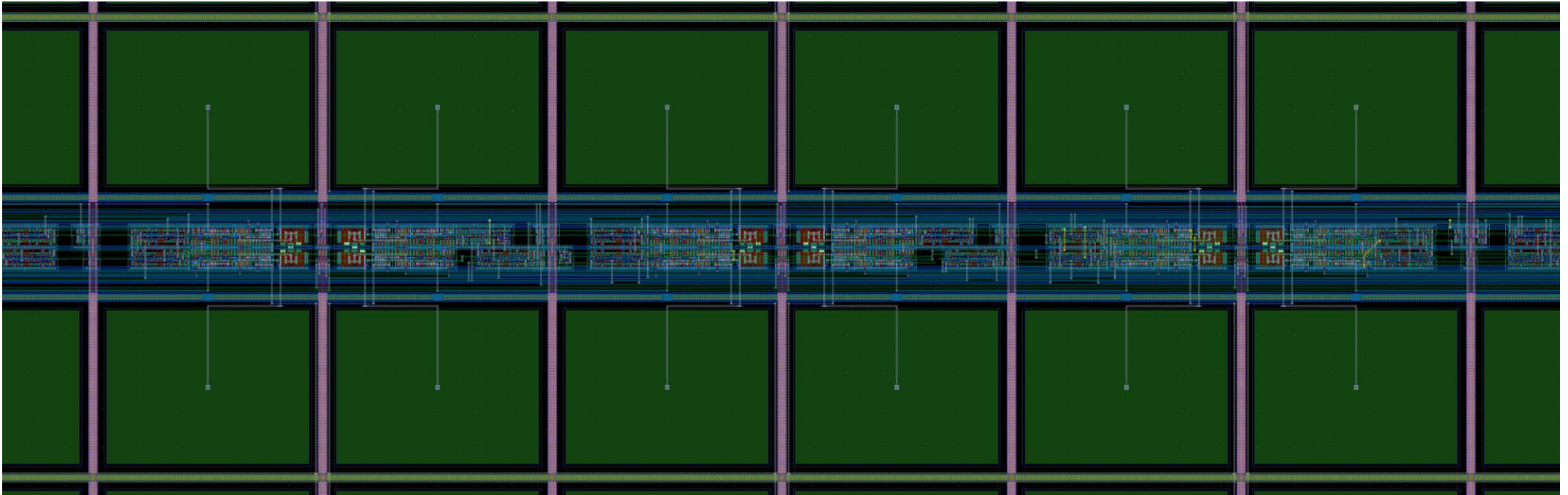
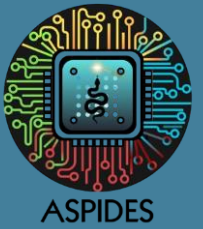
Cluster



- Four pixels **sharing well and bias networks** to increase the FF (~50%).

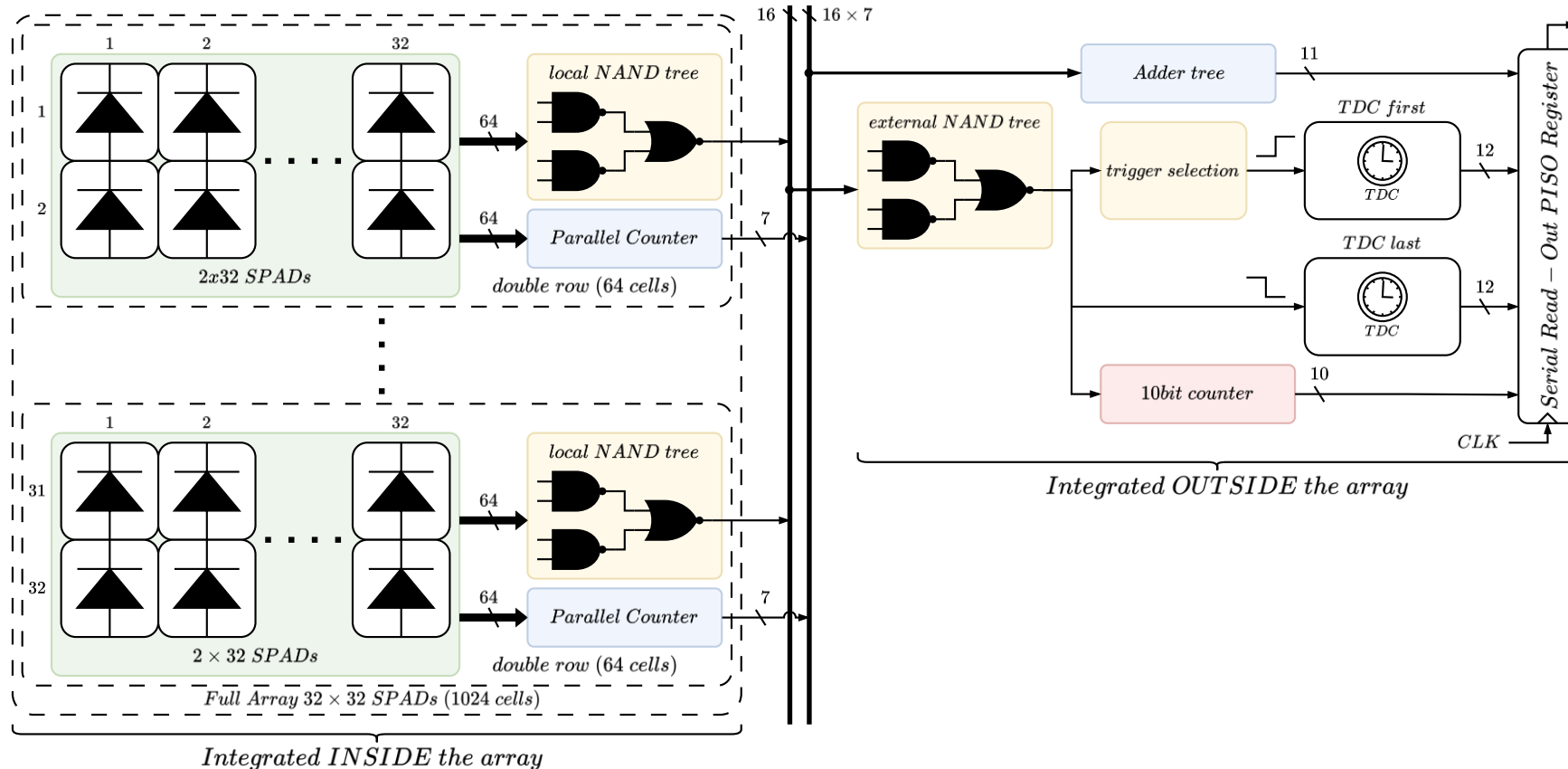
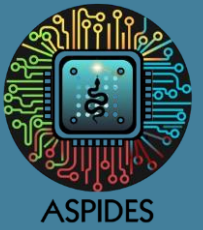


Double Row



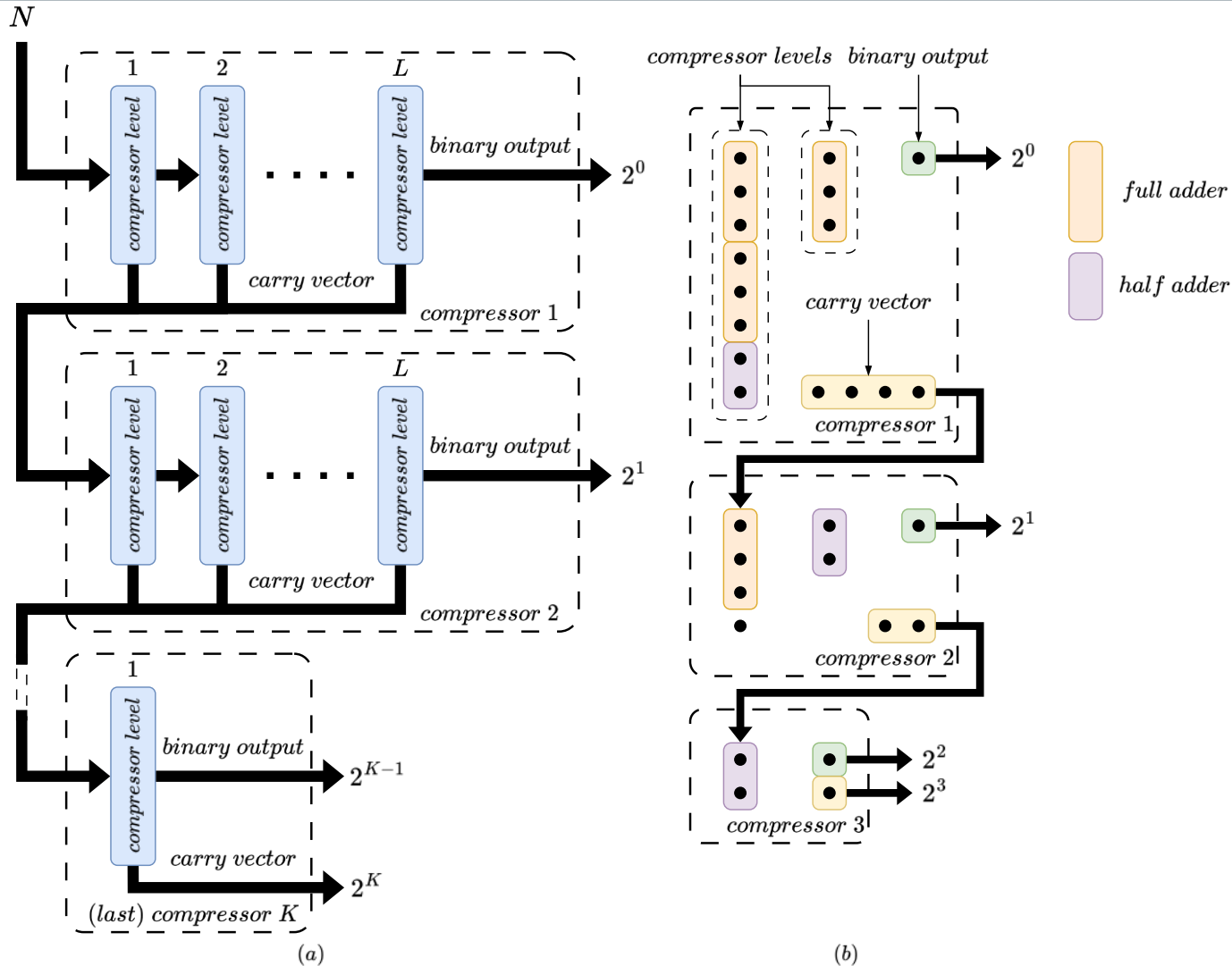
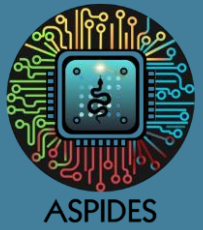
- 16 clusters compose a single double row.
- Each double row behaves like an **independent block** that generate a time and an energy output.
- 16 double rows are **stacked** to compose the full array.
- The outputs of each double row are combined outside the array.
- The front-end electronics of a pixel in a double row can only communicate with pixels in the **same double row**.

Architecture



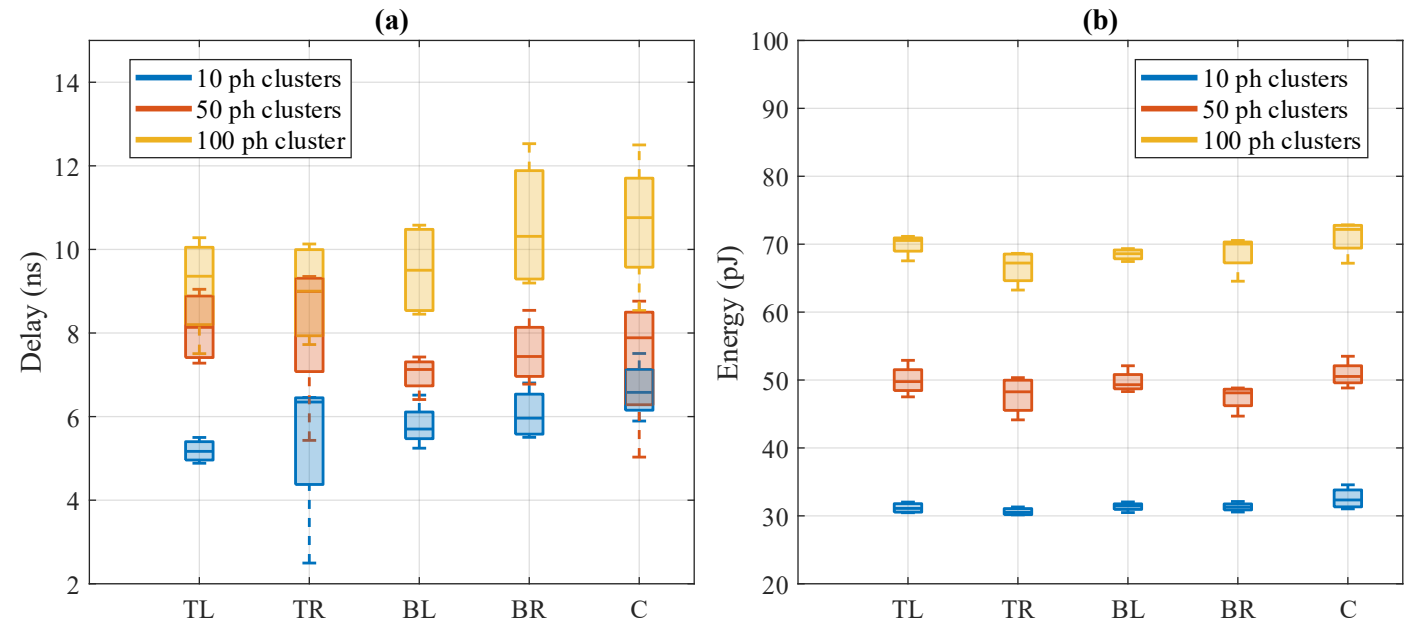
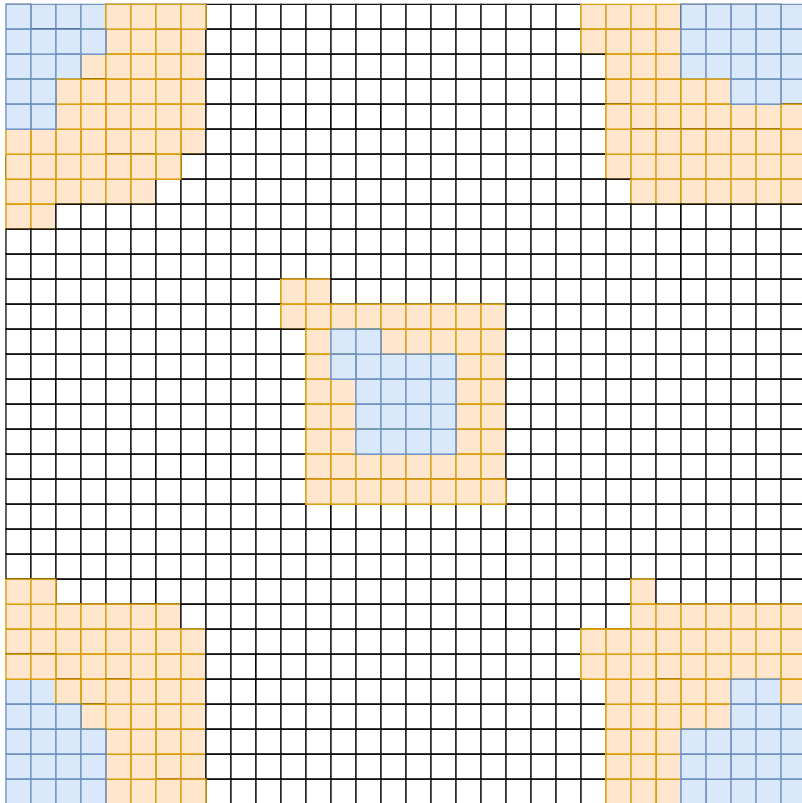
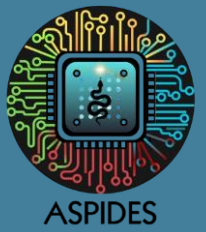
Parallel counter architecture implemented to asynchronously count the **number of fired pixel**.

Parallel Counter



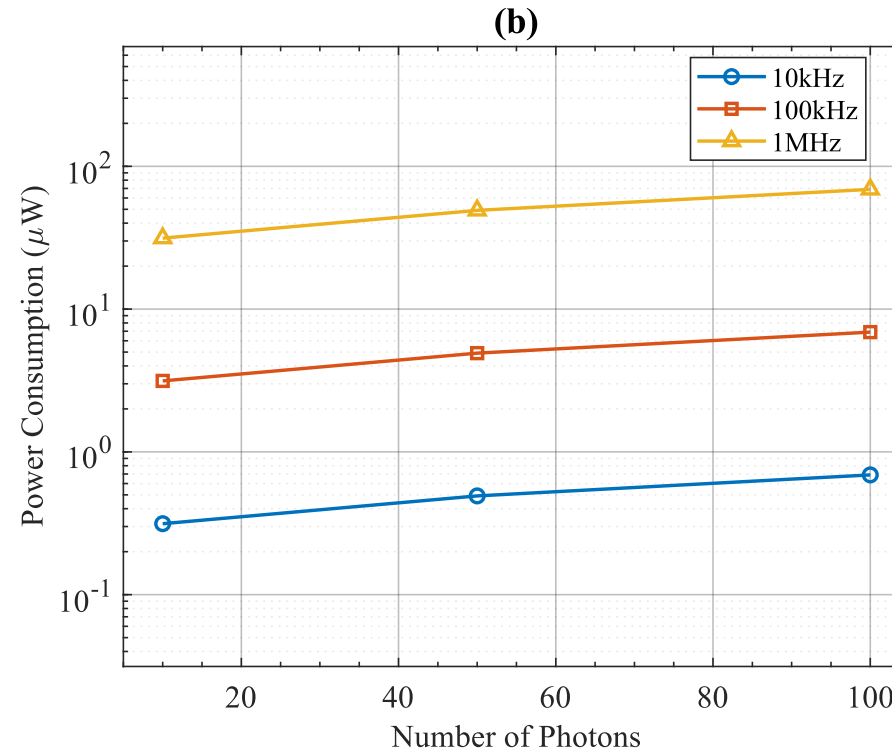
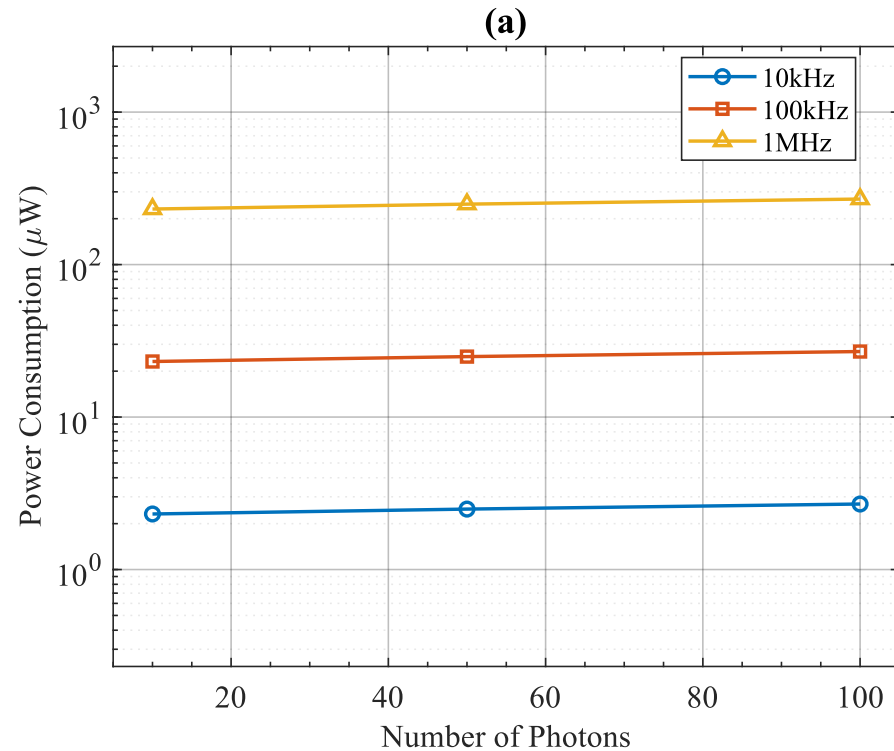
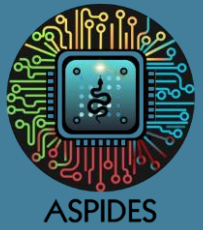
- The 64 outputs of a double row are processed by the **local parallel counter** to generate a 7-bit word.
- This operation is performed by progressively reducing the number of bits using different compressor.
- Each compressor is internally divided into several compression level composed of **full and half adder**.
- The structure of the parallel counter is recursively generated from a **Verilog code** based on the number of input bits.

Parallel Counter



Several beam cluster of different size (10, 50 and 100 photons) are fed to the array: the delay of the parallel counter and the energy consumption was extracted.

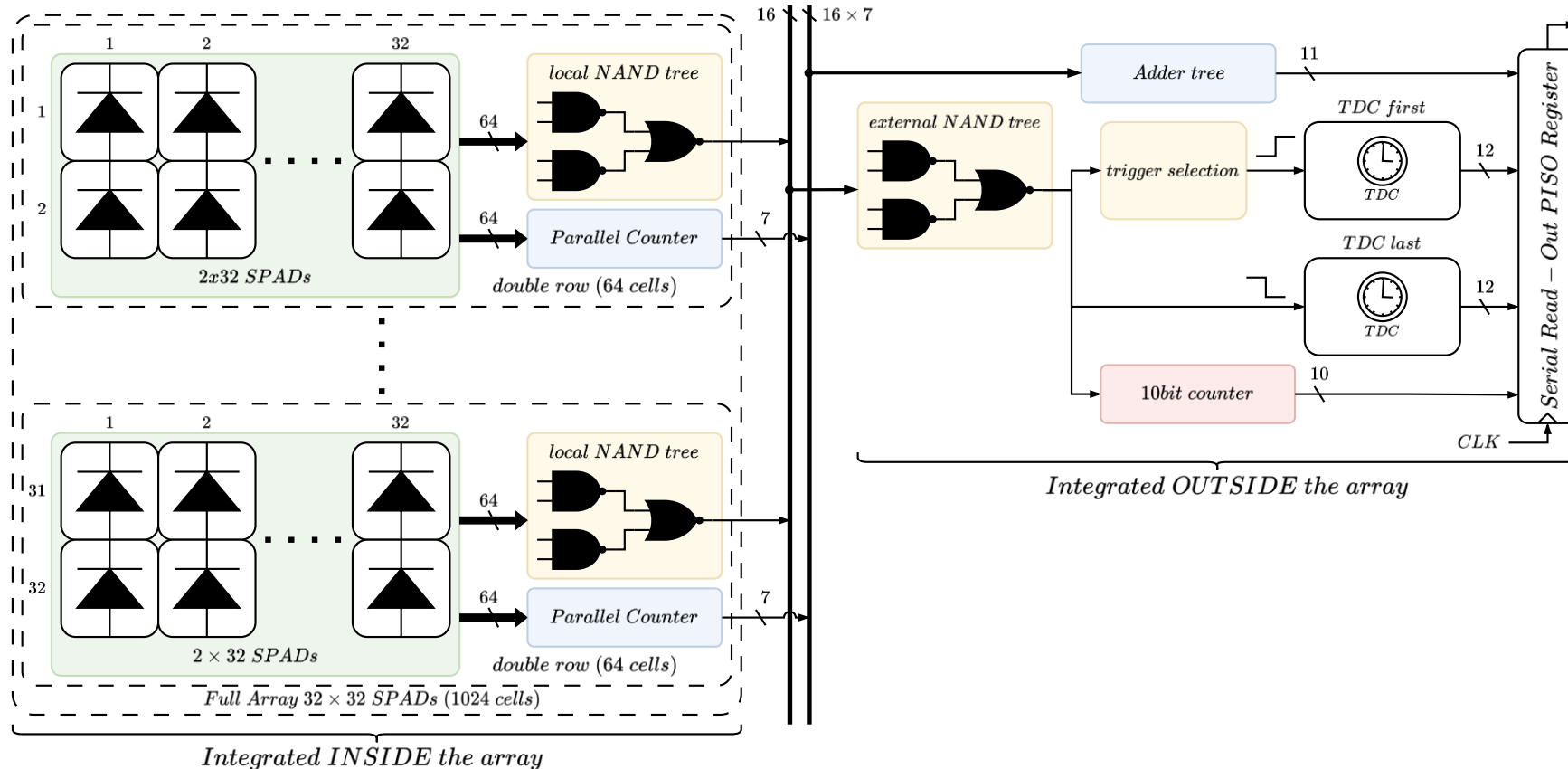
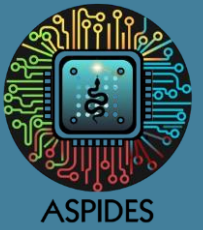
dSiPM Power Consumption



Power consumption estimation based on the simulation result on the array and on the TDCs:

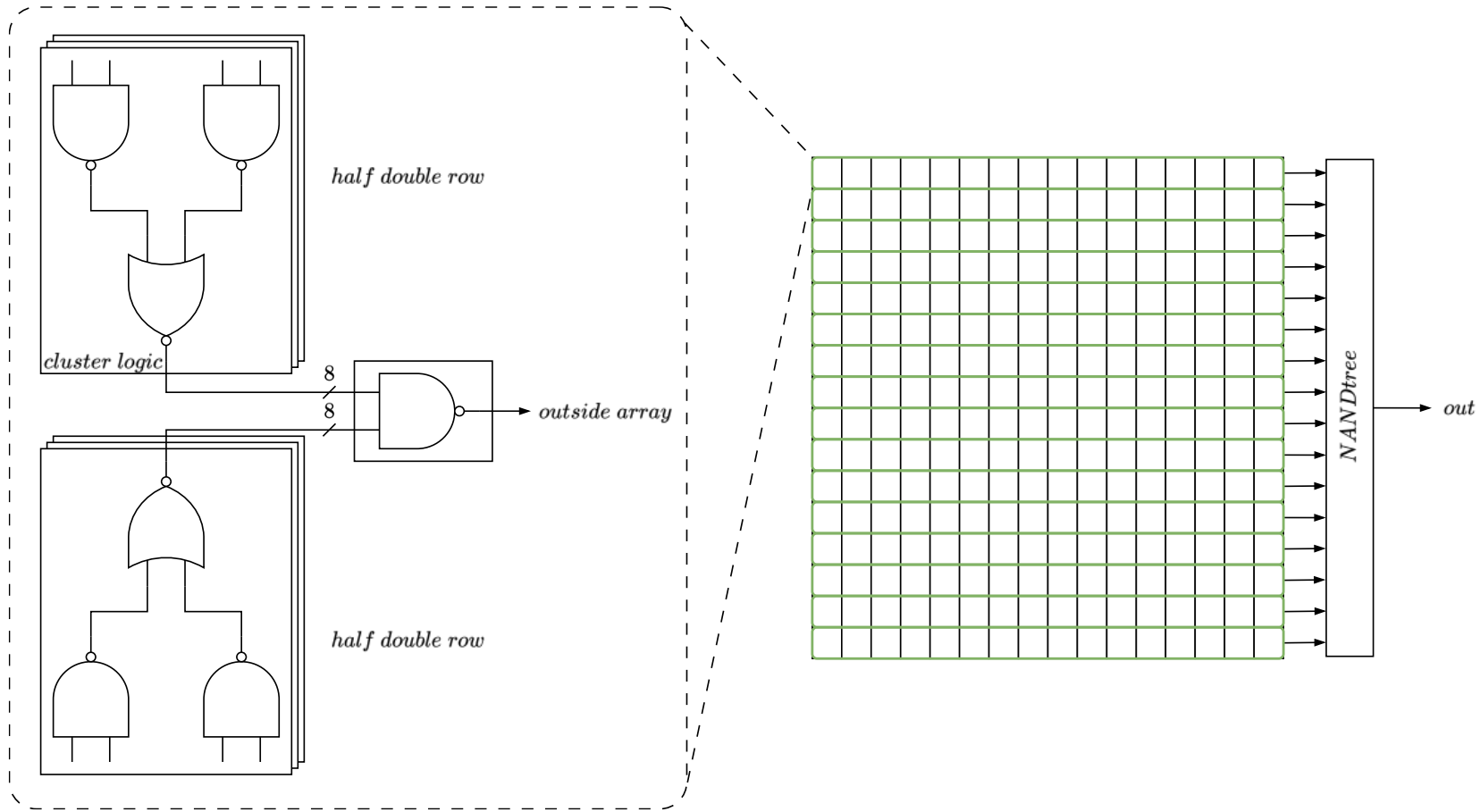
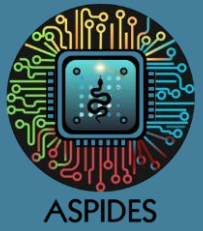
- a) Considering also the two TDCs converting a time interval equal to half their FSR (~200 ns).
- b) Only the power consumption of the array to produce the information about the number of fired photons .

Architecture



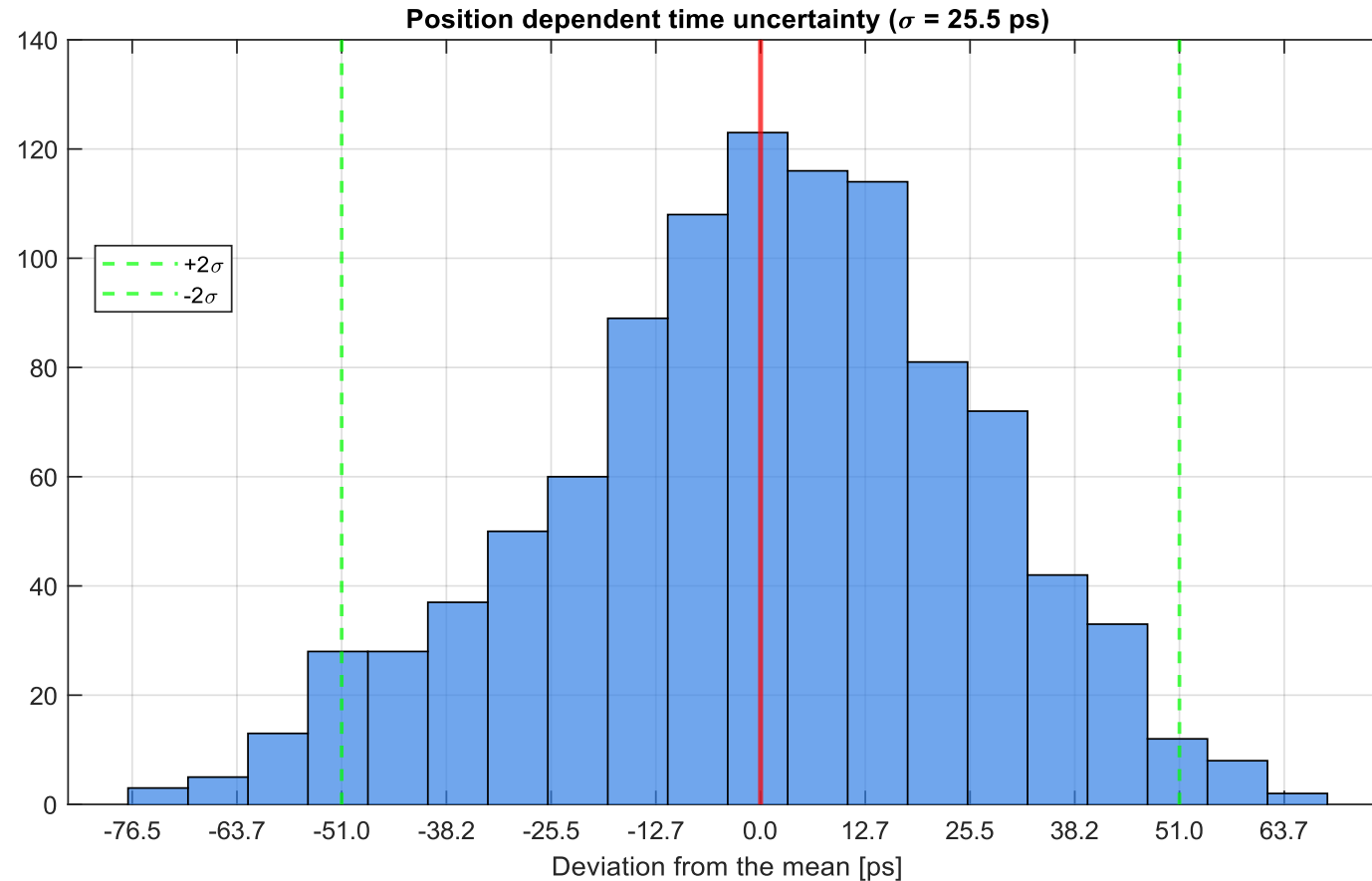
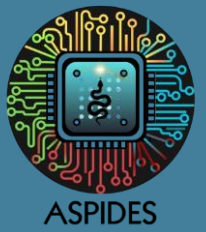
A logic gate tree **equivalent to a NAND gate** is used to detect the first and last events on the array by the timing of its **edges**.

NAND Tree



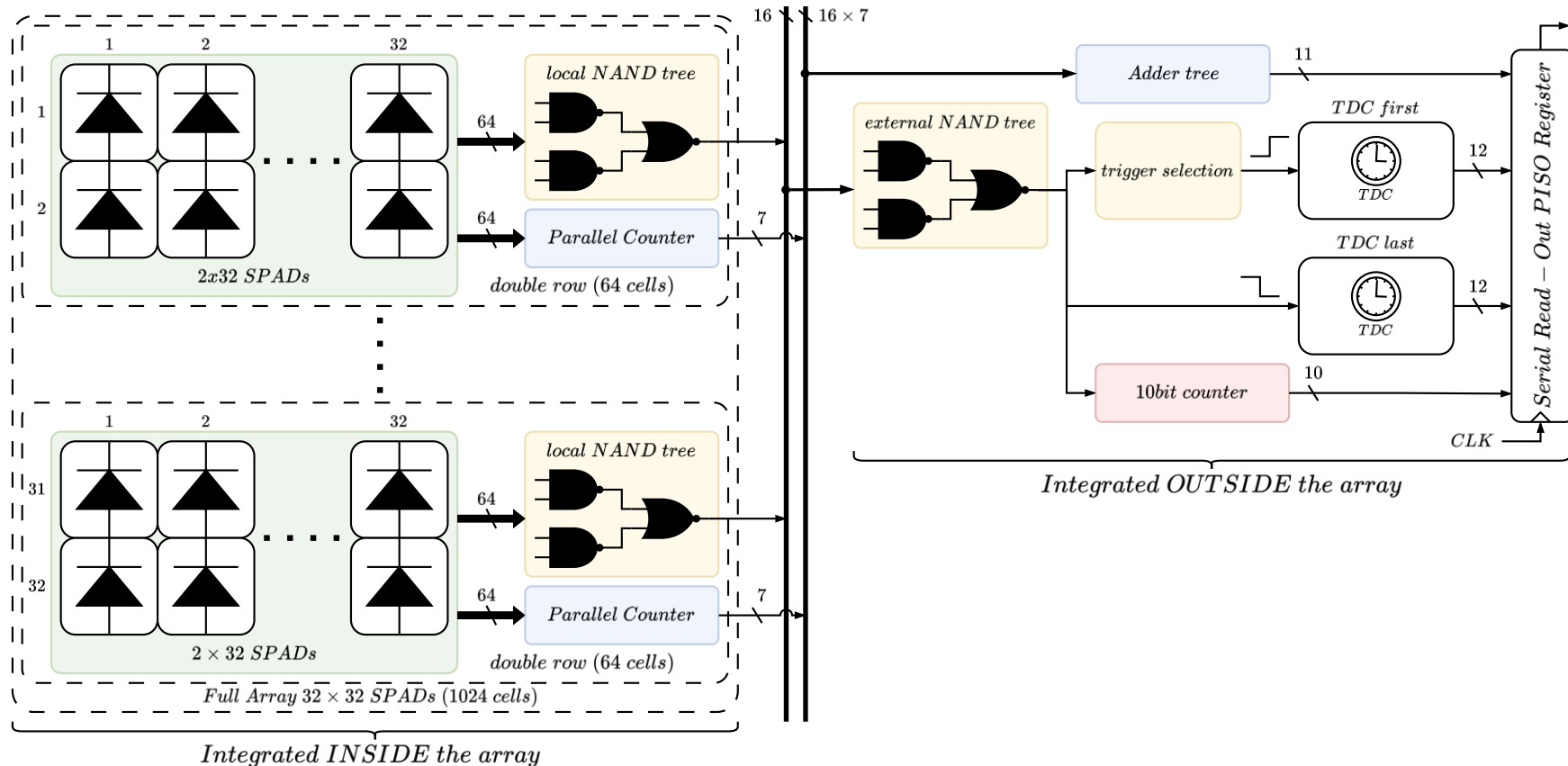
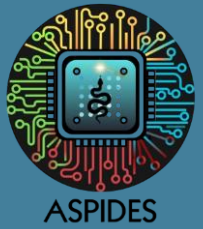
Each cluster generates 1 bit, all the bits in a double row are **combined to provide a single bit** at the double row output. All the 16 double row outputs are processed by the external NAND tree.

NAND Tree



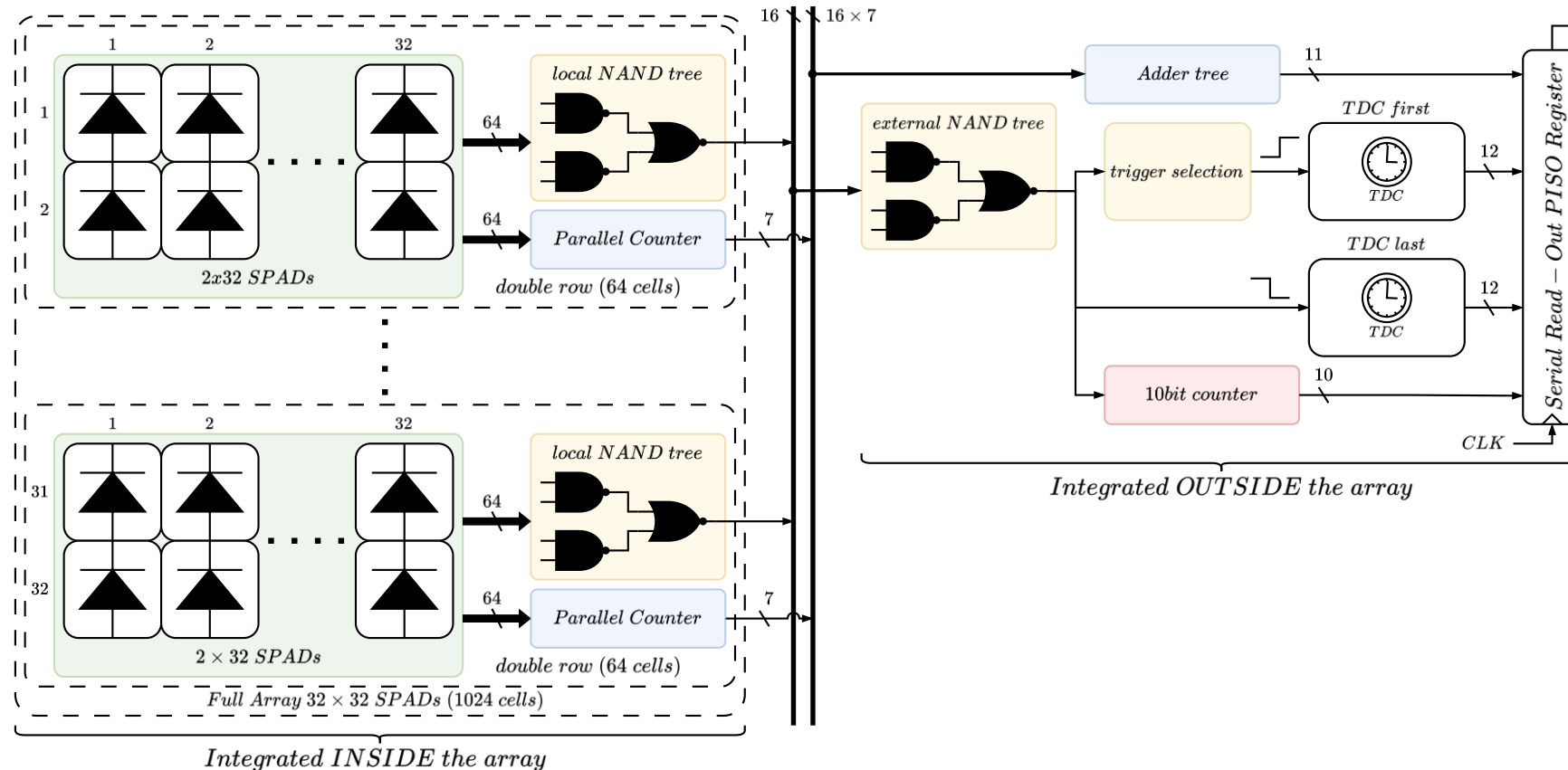
Distribution of the rising edge of the NAND tree output depending on the position of the triggered pixel.

Architecture



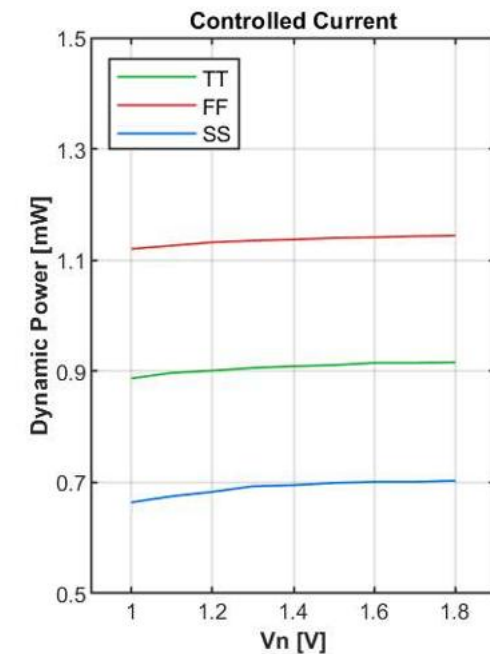
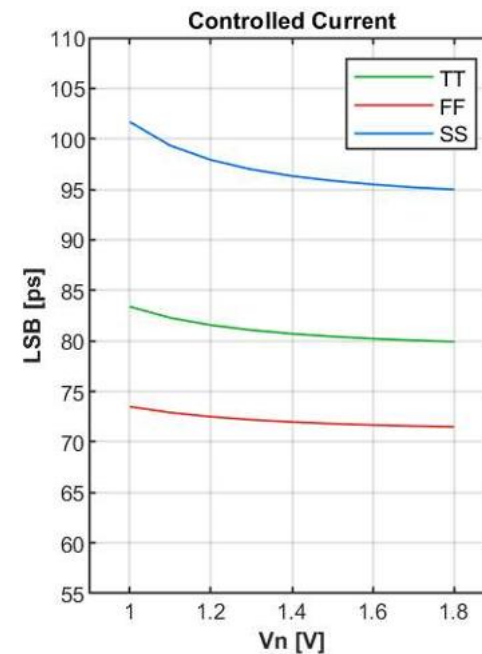
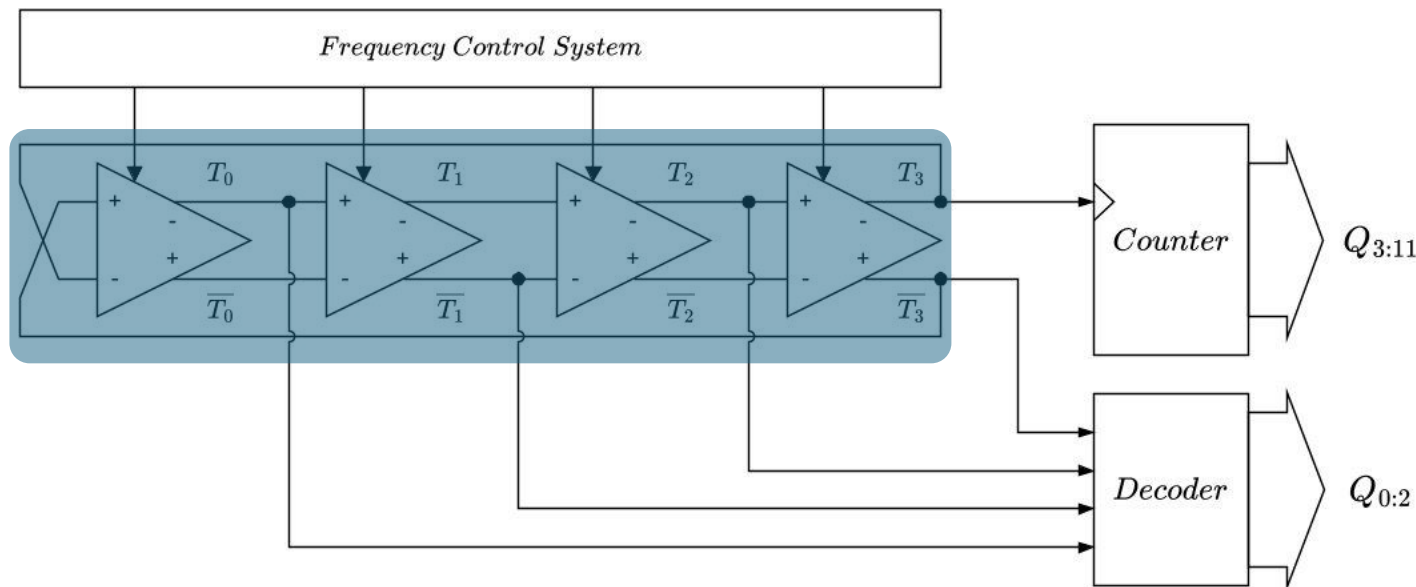
A 10-bit counter is integrated for **characterization purposes**, counts the number of rising edges of the NAND tree output signal.

Architecture



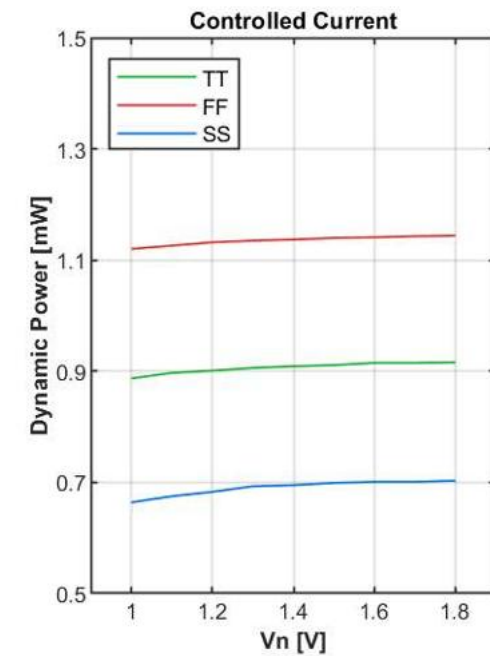
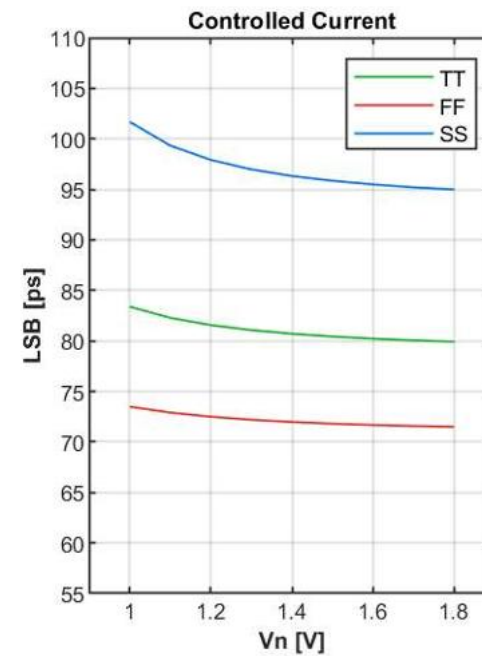
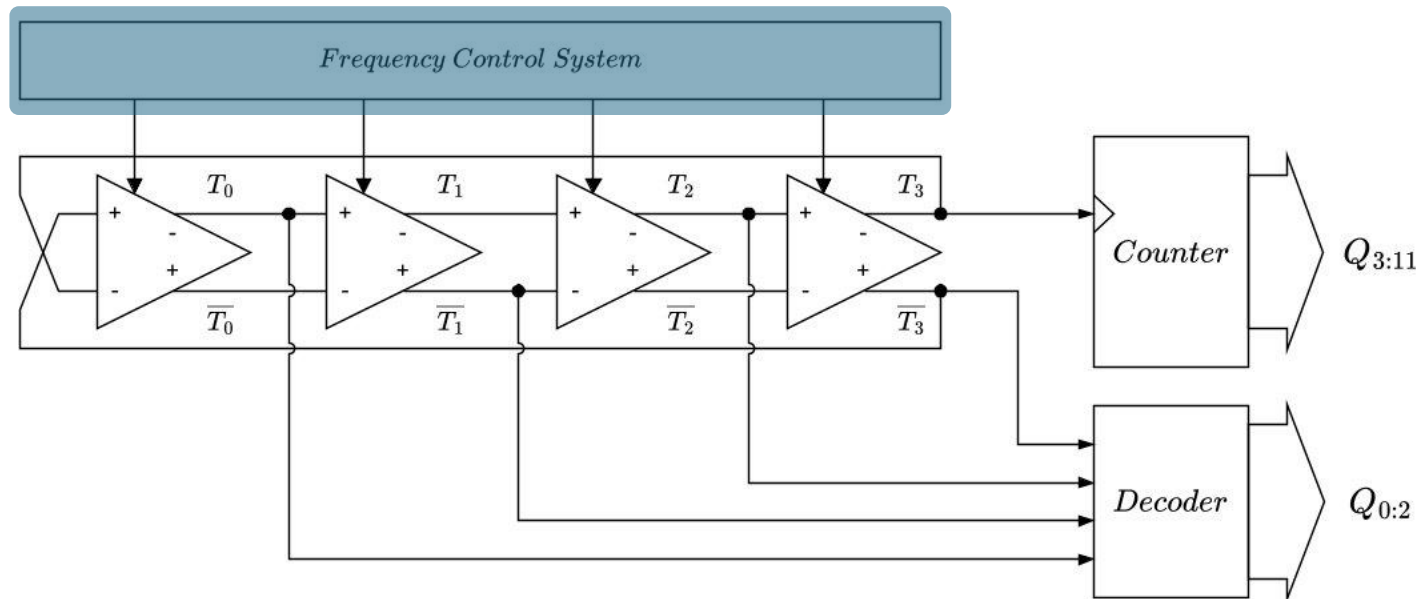
Two TDCs with sub-100ps resolution used for time-stamping: one for the first event and the other for the last one.

Time-to-Digital Converter Architecture



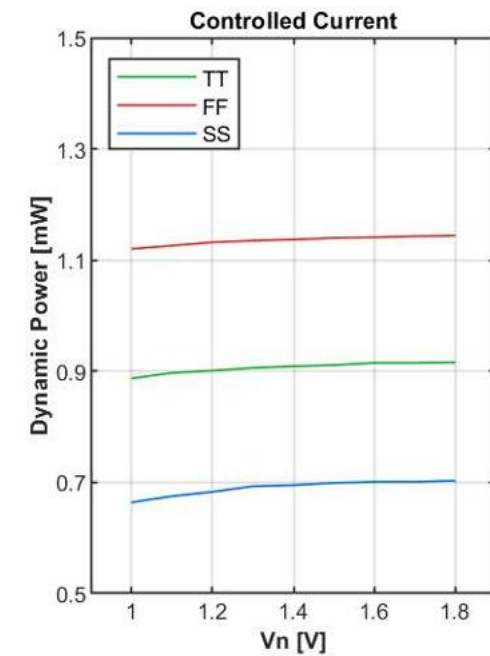
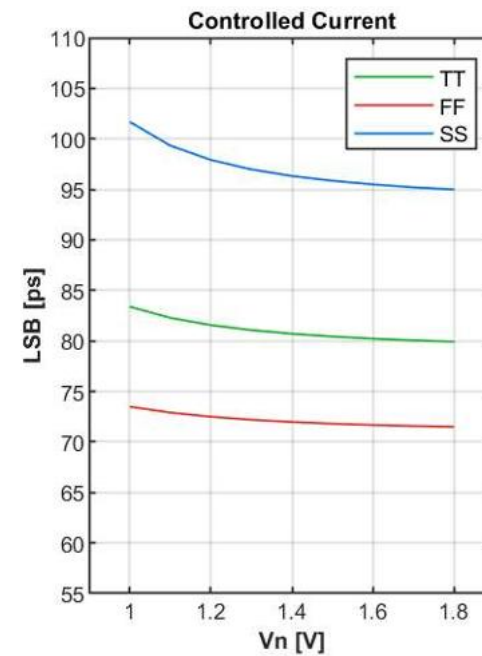
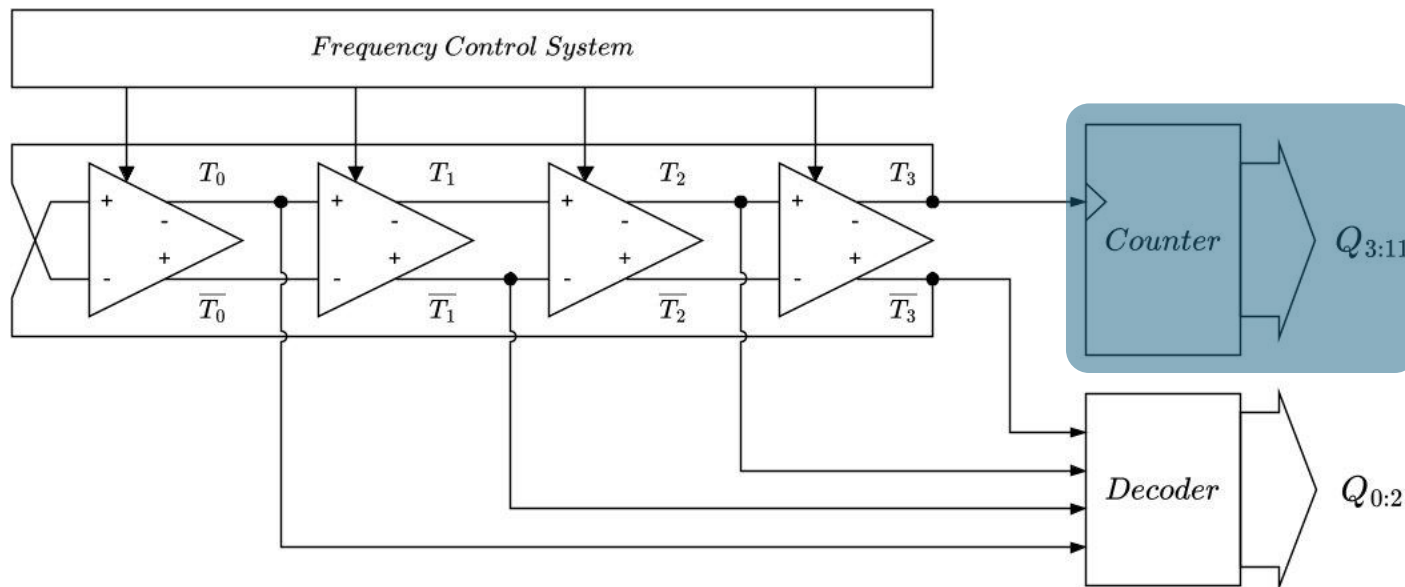
Based on a pseudo-differential ring oscillator with four stages. Resolution below 100ps and 1mW of power consumption.

Time-to-Digital Converter Architecture



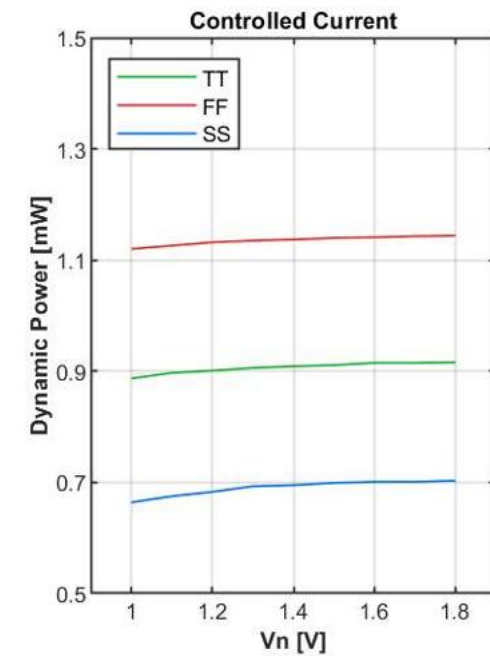
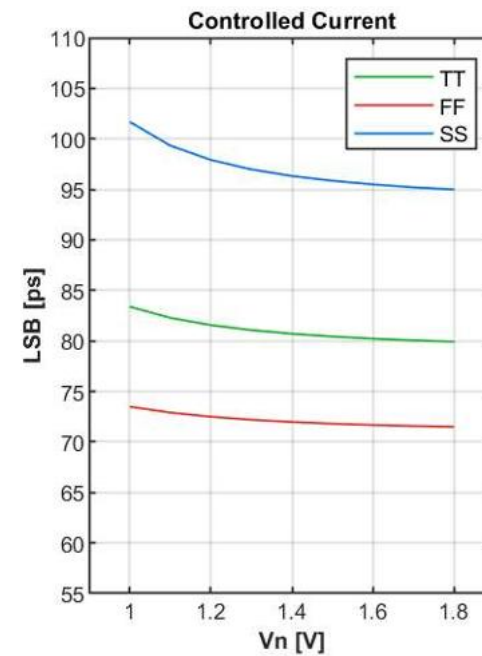
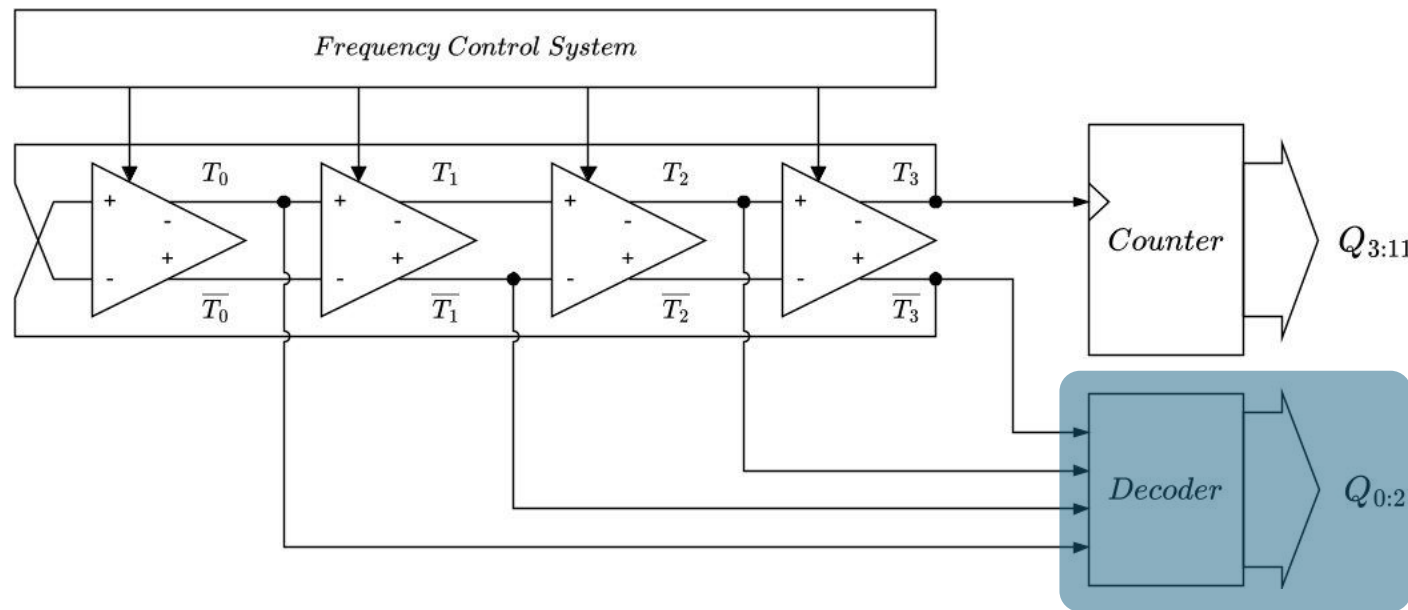
Using an external voltage is possible to **control the frequency** and, therefore, the resolution.

Time-to-Digital Converter Architecture



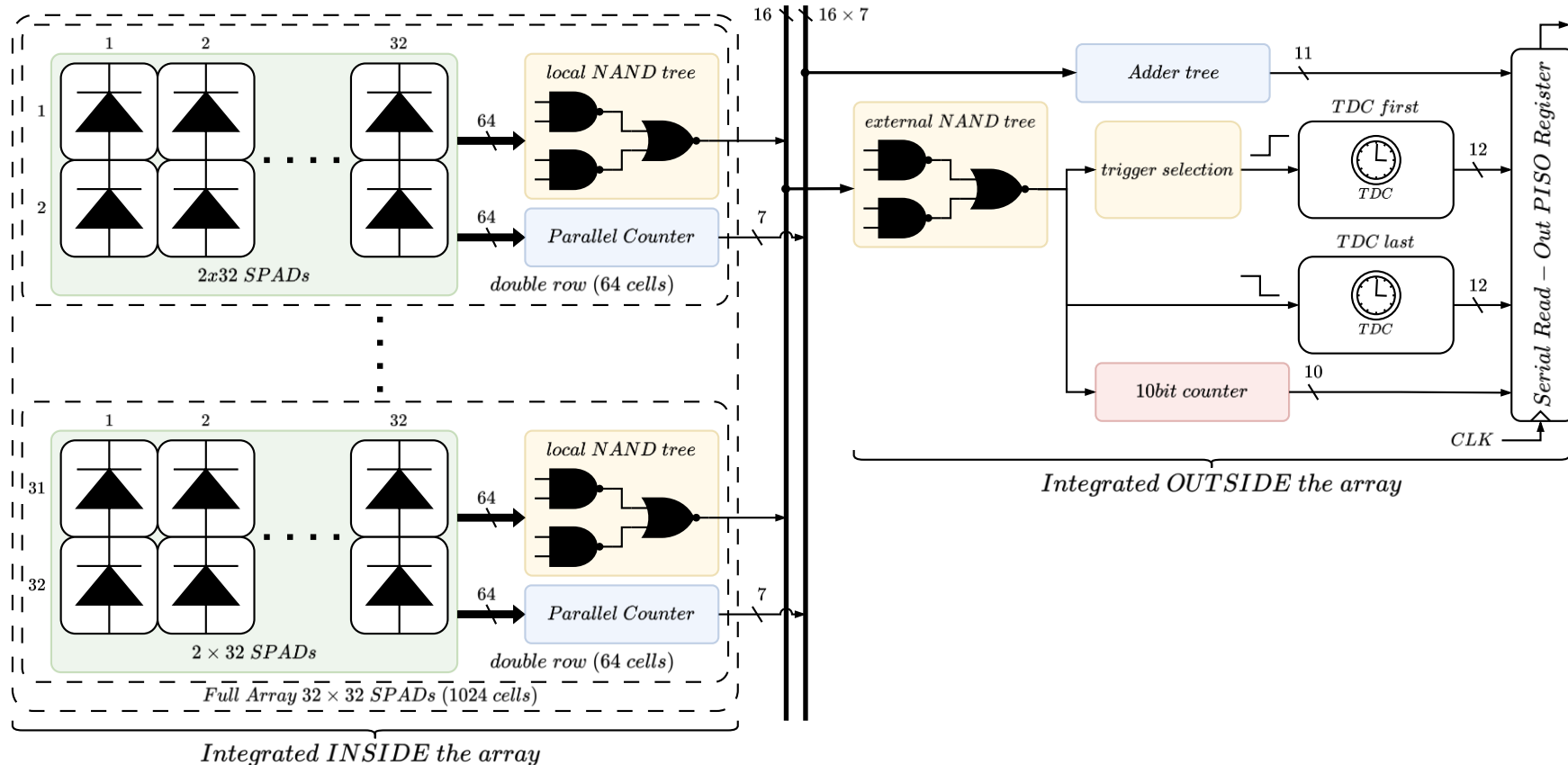
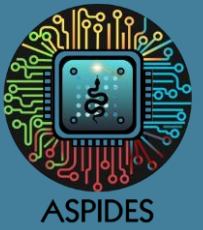
The *MSBs* are generated with a **counter** that guarantees large full-scale range and linearity.

Time-to-Digital Converter Architecture



The *LSBs* are generated by a **decoder** that reads the internal node state once the oscillation is stopped.

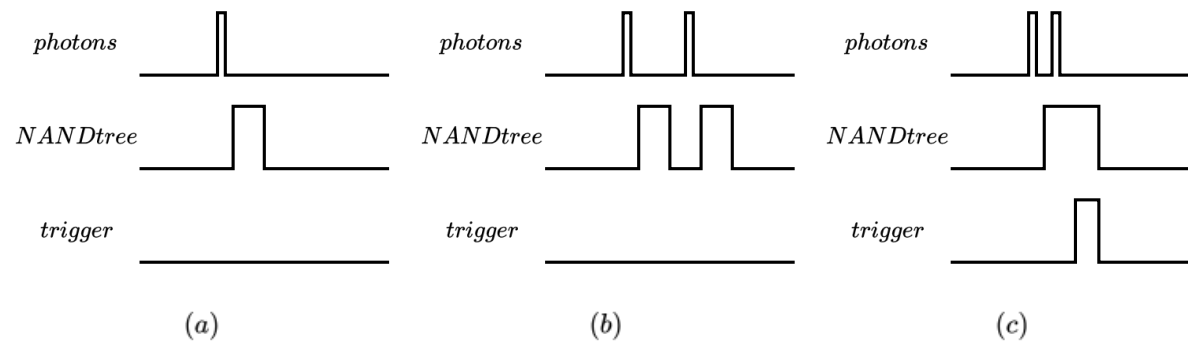
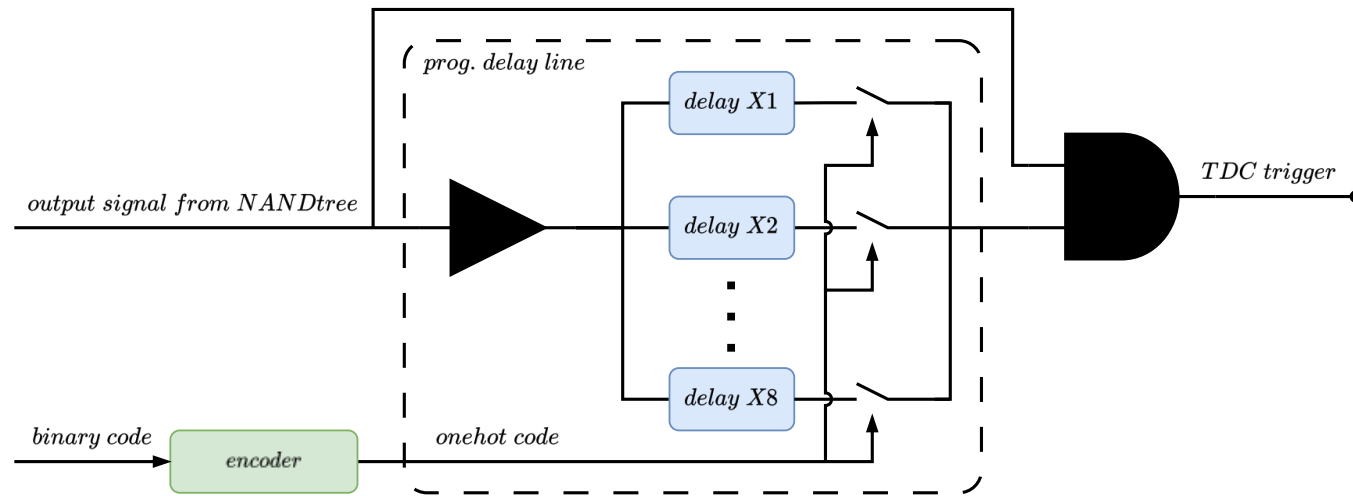
Architecture



Three selectable trigger options for TDC first, only one for TDC last.

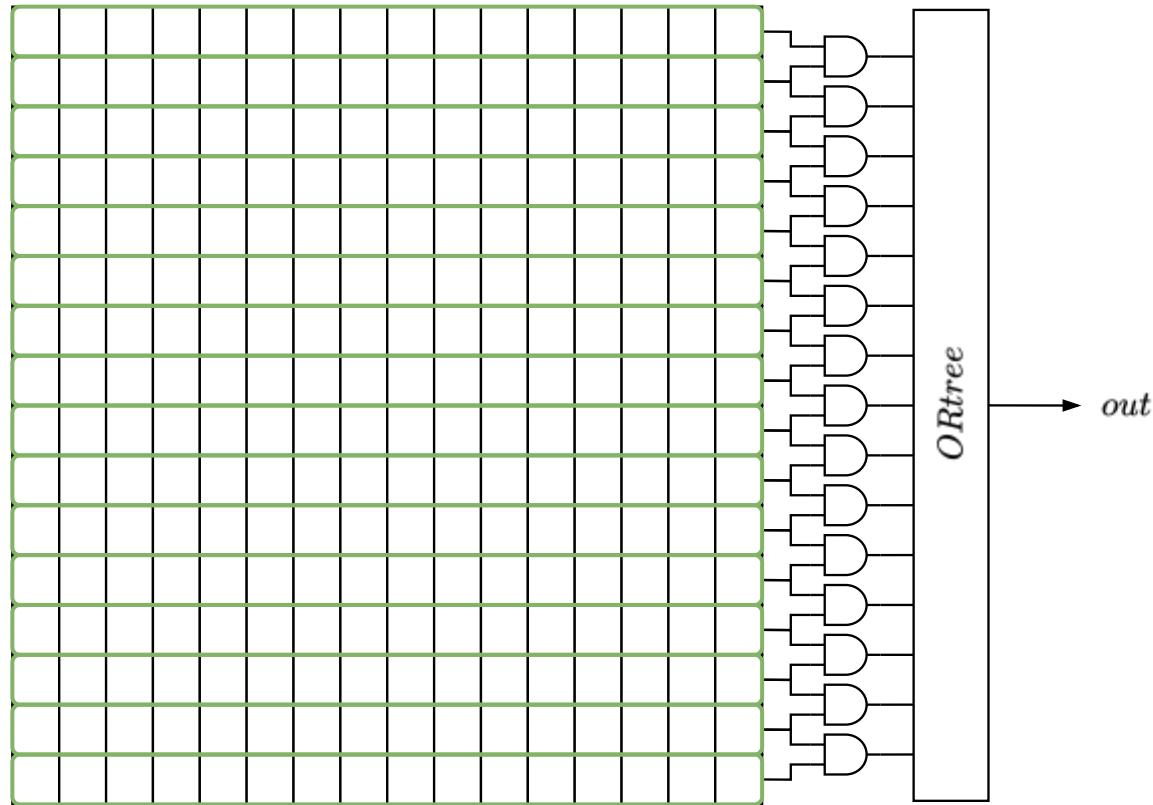
TDC last will be triggered only by the falling edge of the output of the NAND tree only if the TDC first has been triggered

Time-over-Threshold Detection

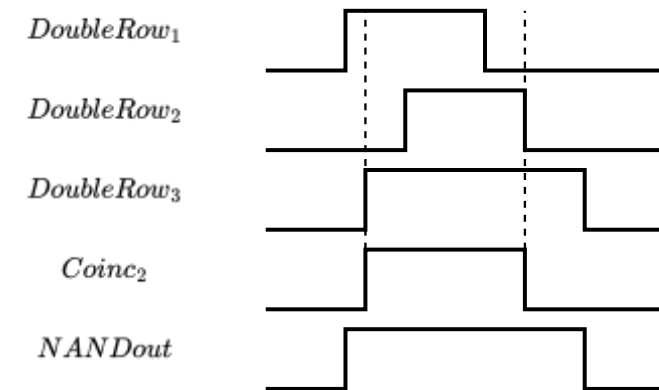


- Eight different cells, each one with a different delay, the **tuning range** spans from 600ps to 1.4ns with a step of 100ps.
- If the *NAND tree* output remains high for a time **greater than the chosen delay**, then *TDC first* will be triggered.
- The **decoder** is used to convert the 3 configuration bits in OneHot code to select the desired delay.

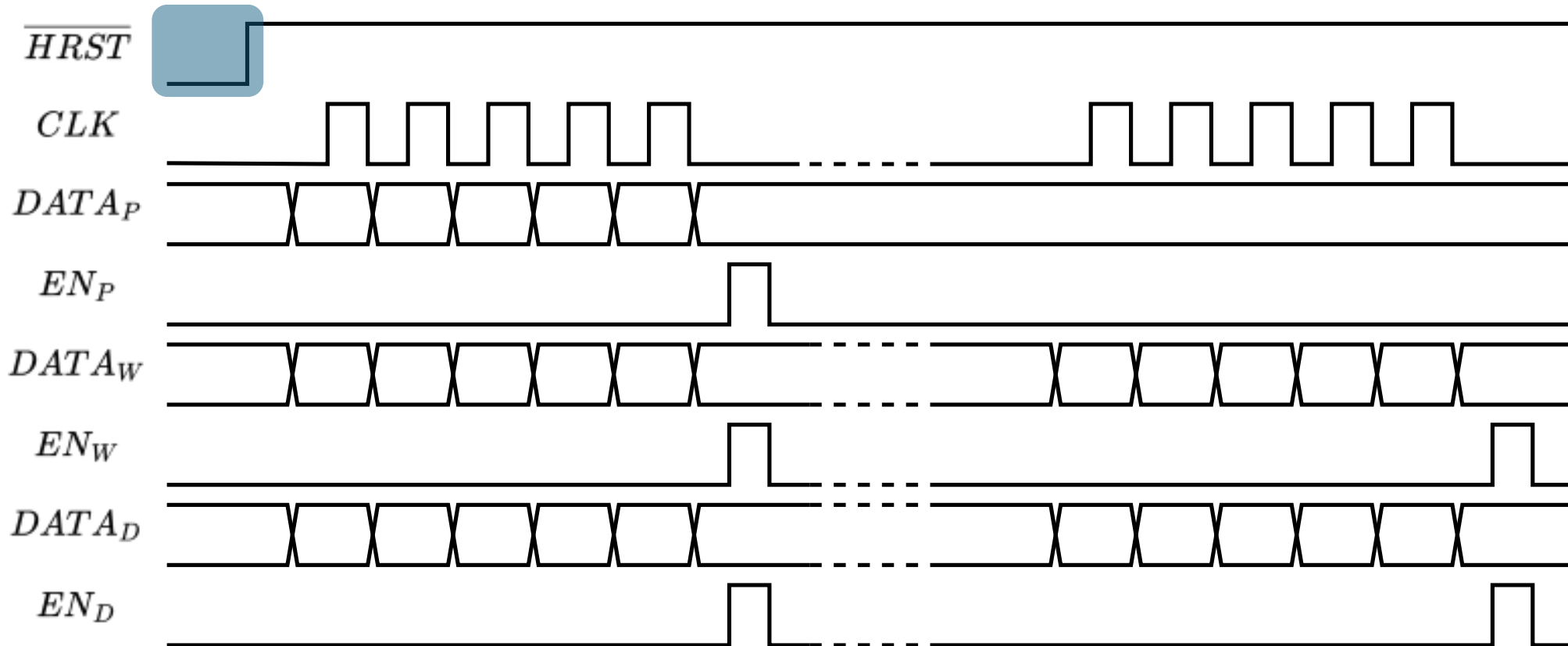
Coincidence Detection



The output of the **local NAND tree** for each double row is compared with the adjacent ones and only if an events occurs in **two/three** (selectable by configuration) of them then *TDC first* will be triggered.

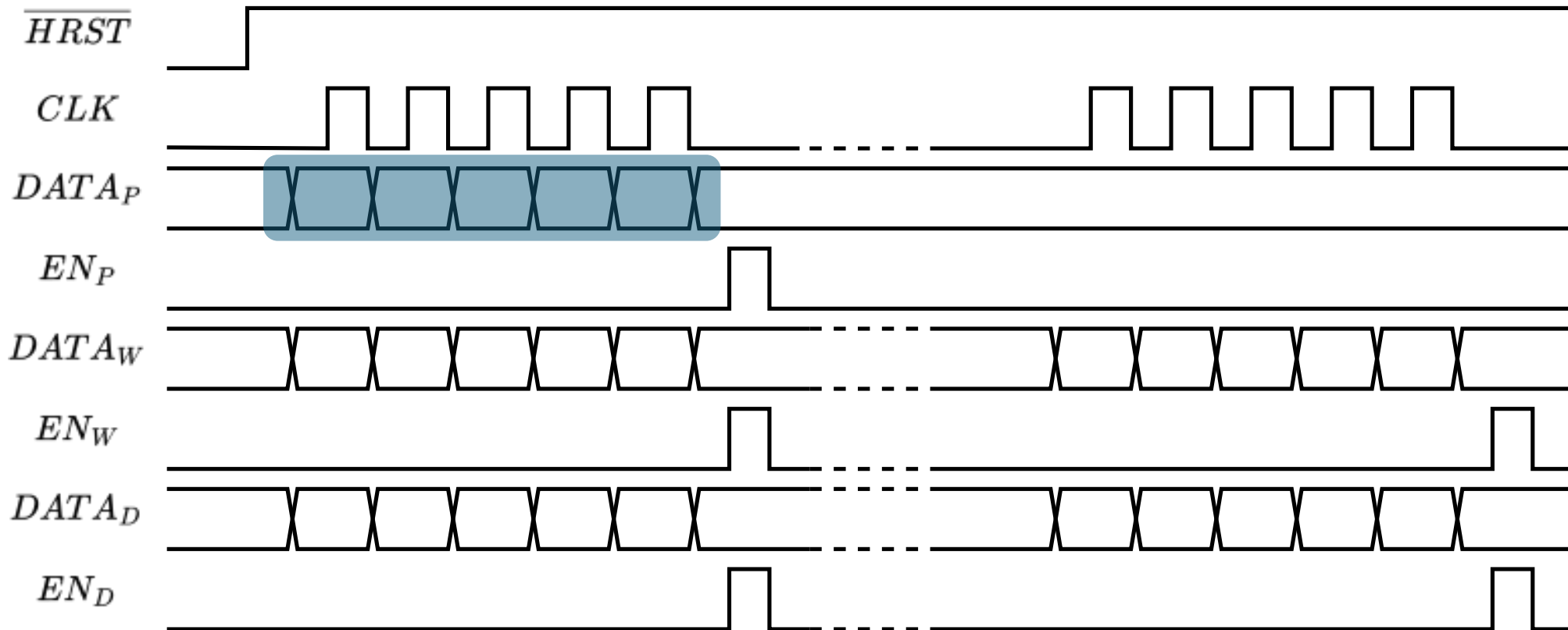


Programming Phase



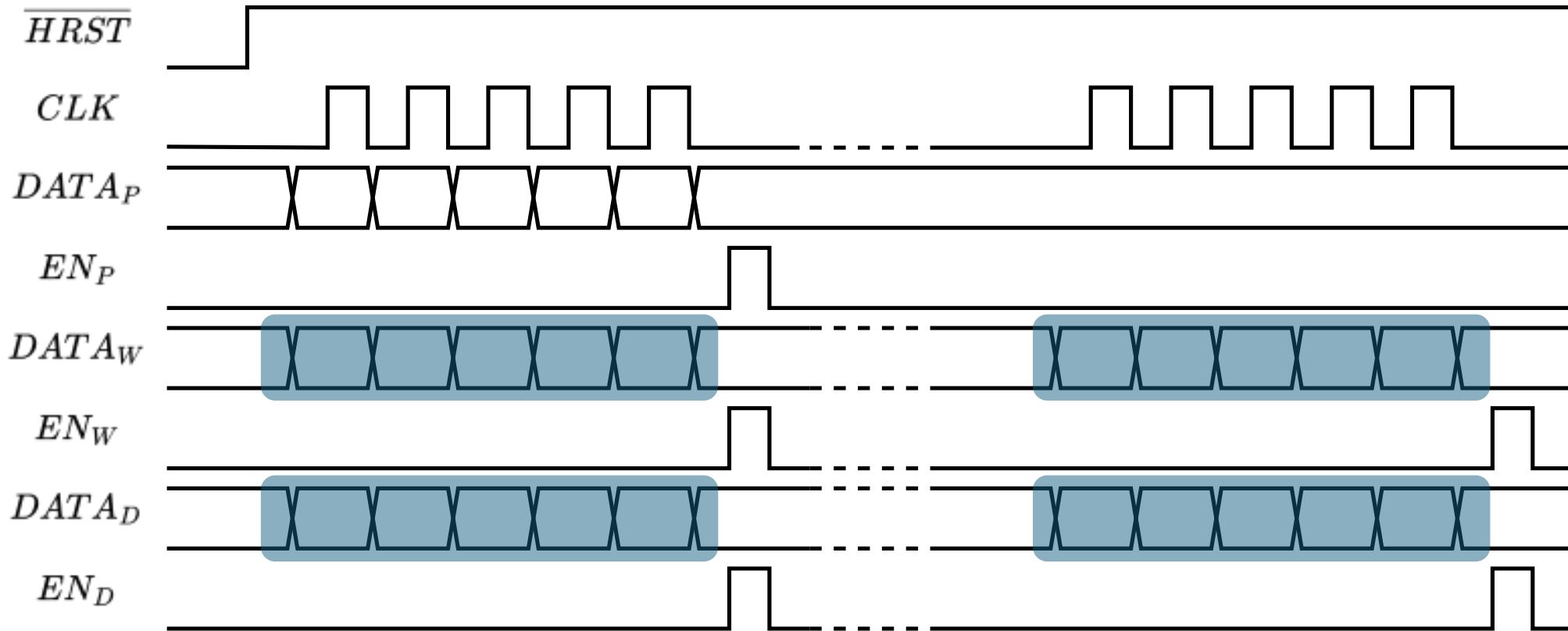
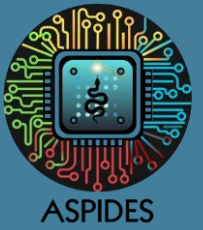
The *Hard Reset* clear all the configuration and programming register.

Programming Phase



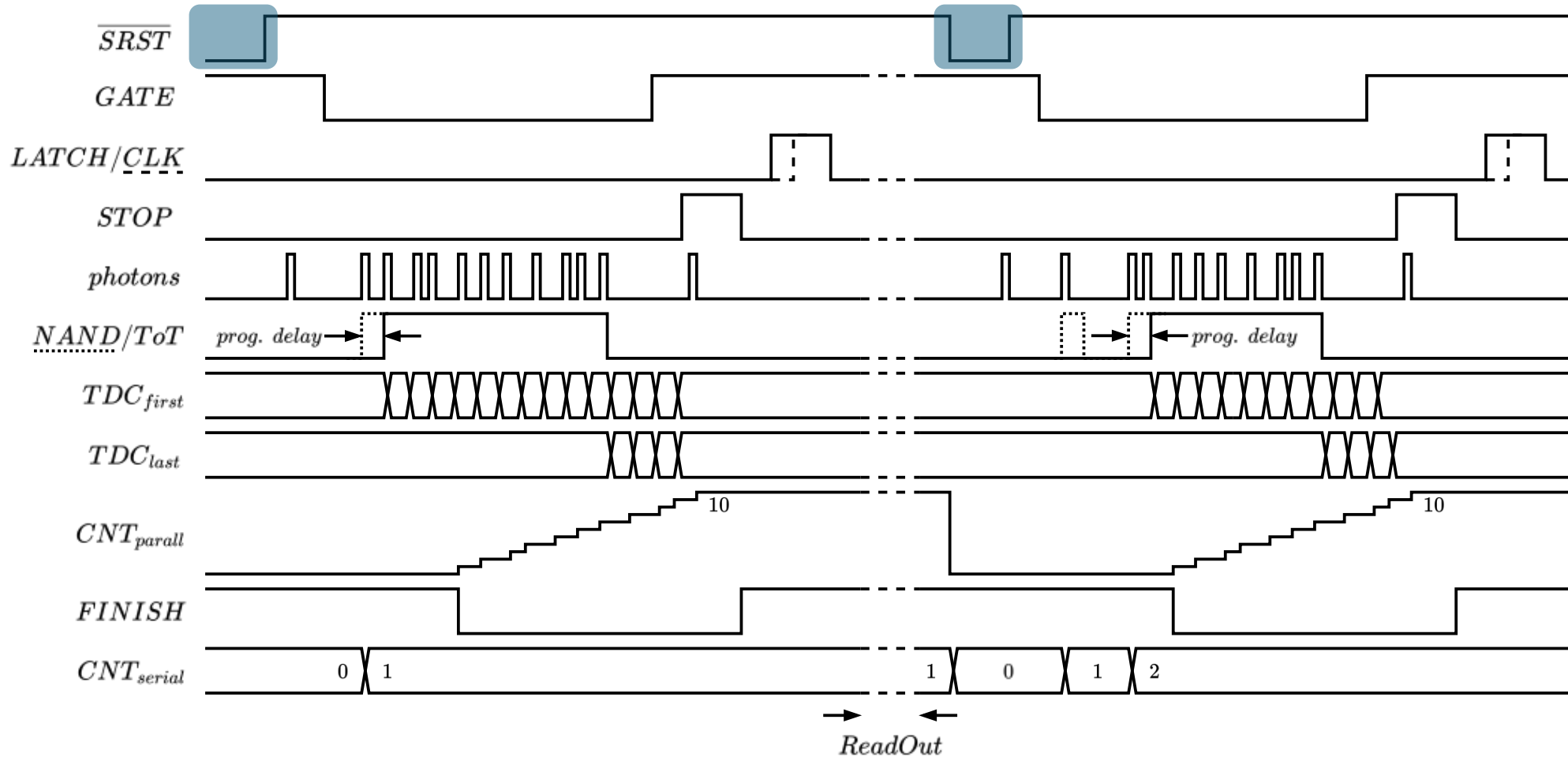
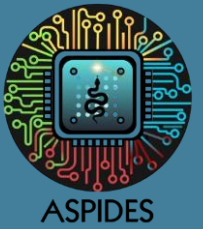
The configuration register requires only one data stream to be programmed.

Programming Phase



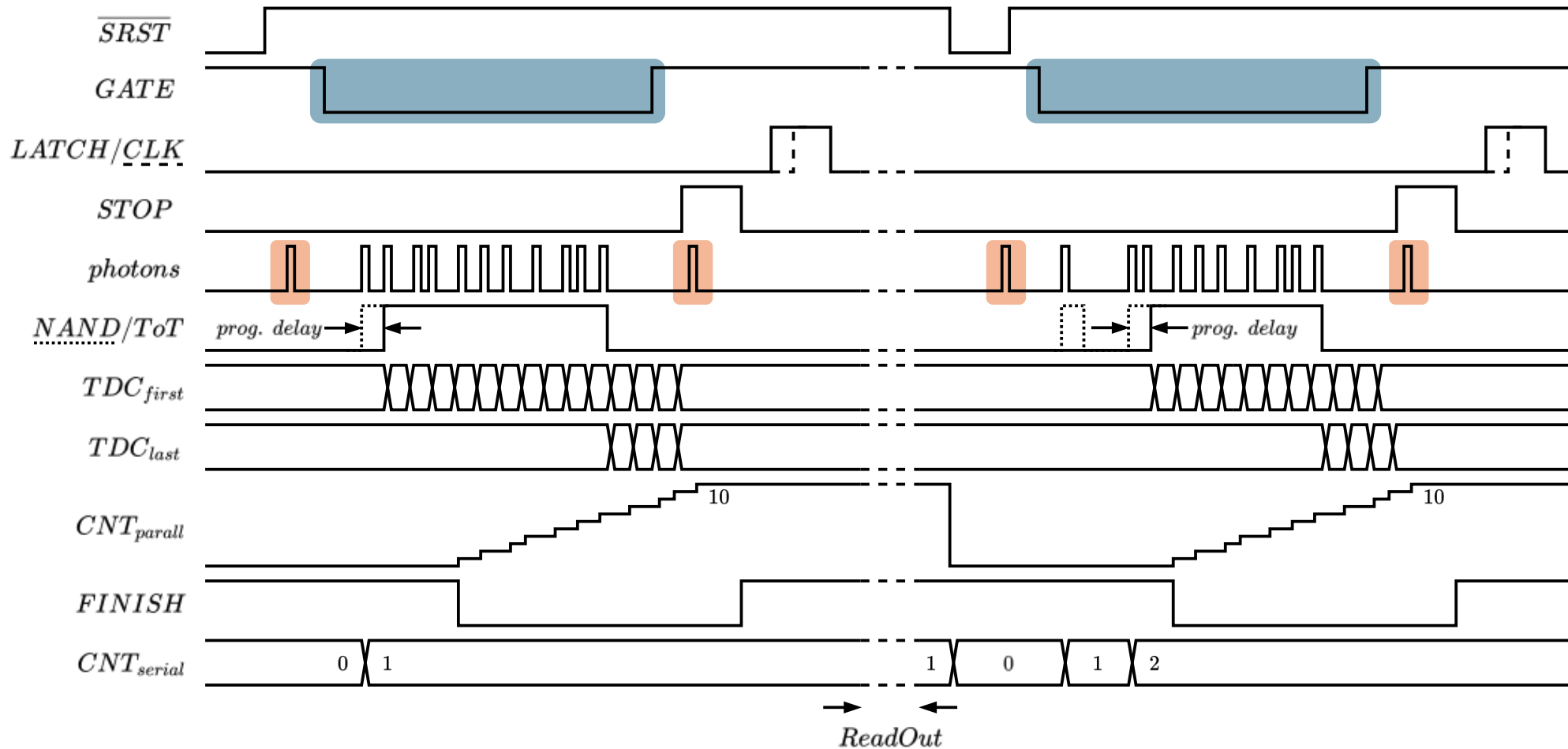
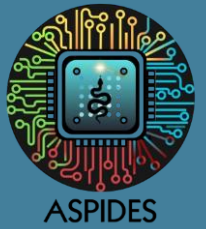
The register for pixel enabling may require more than one data stream depending on the desired configuration.

Measurement Phase



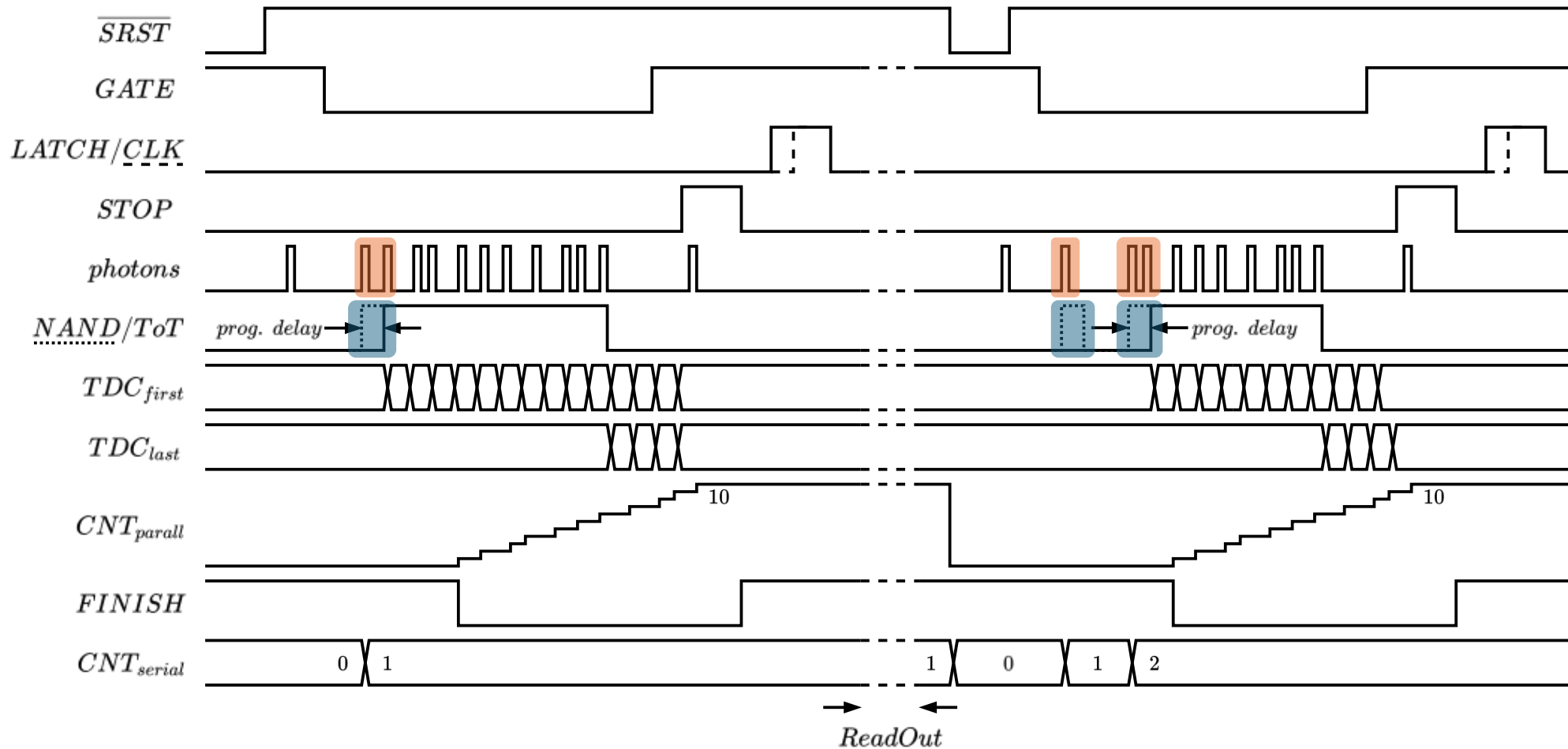
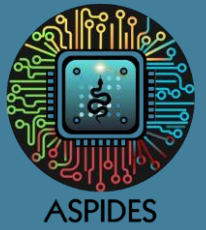
The *Soft Reset* initialize the array for a new measurement, reset the in-pixel latch, the readout register and the TDCs.

Measurement Phase



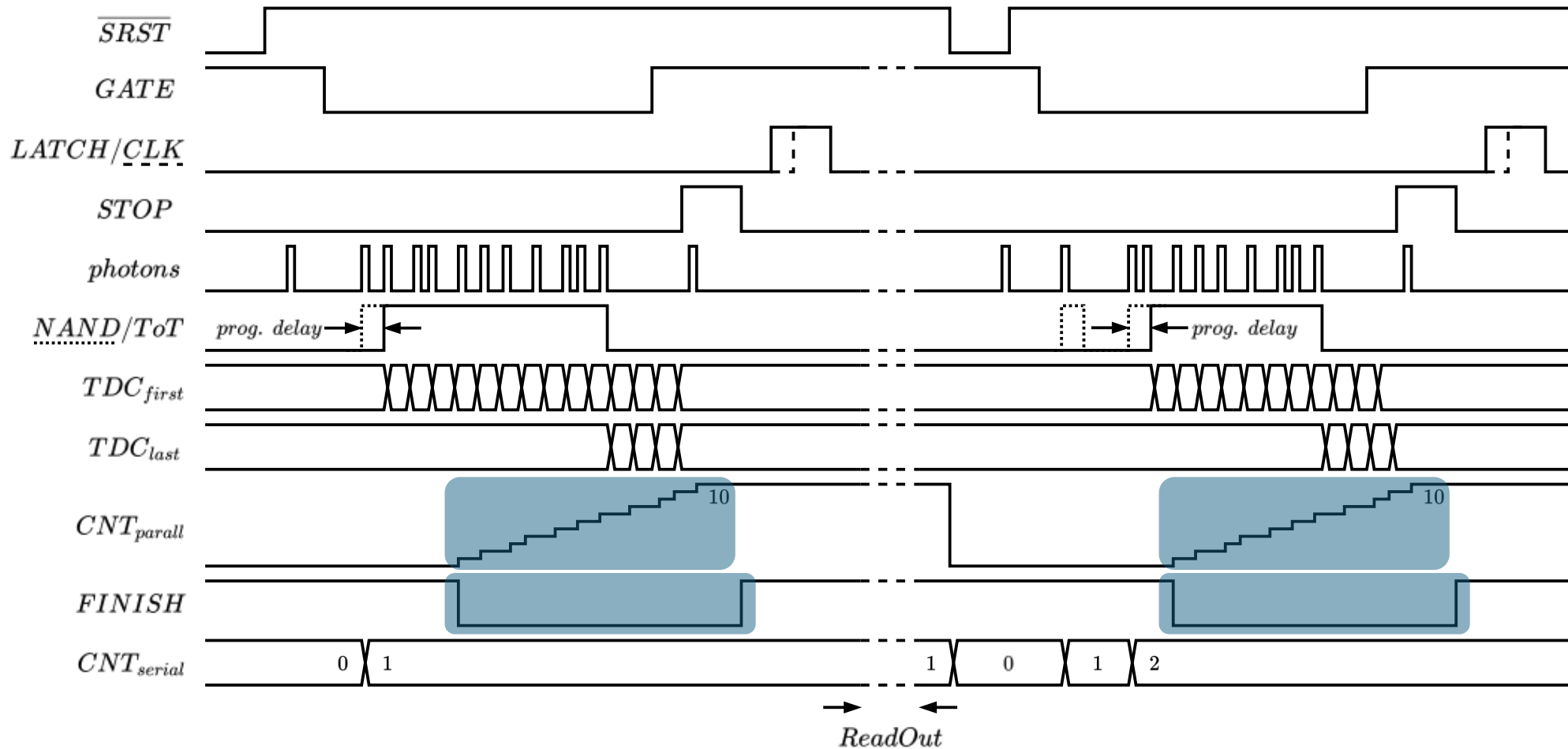
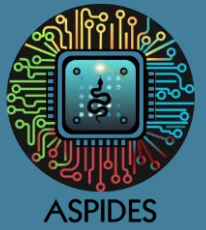
The *Gate* signal defines the window for acquisition, events outside the window will do not affect the measurement.

Measurement Phase



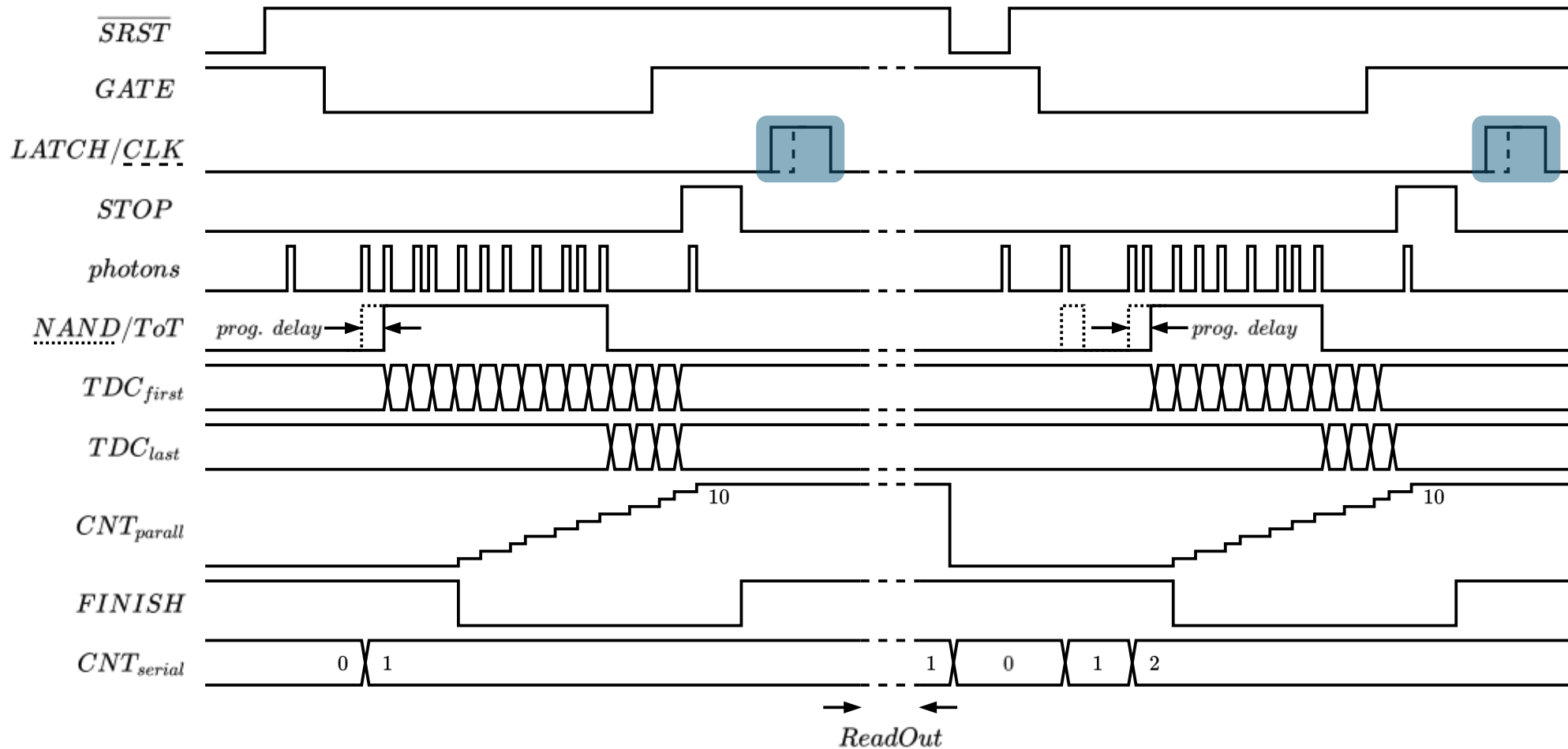
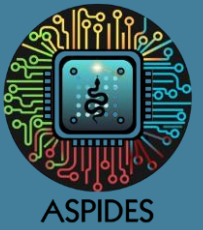
If the Time-over-Threshold trigger method is enabled, at least **two photons** are required in a programmable time window to trigger the *TDC_{first}*.

Measurement Phase



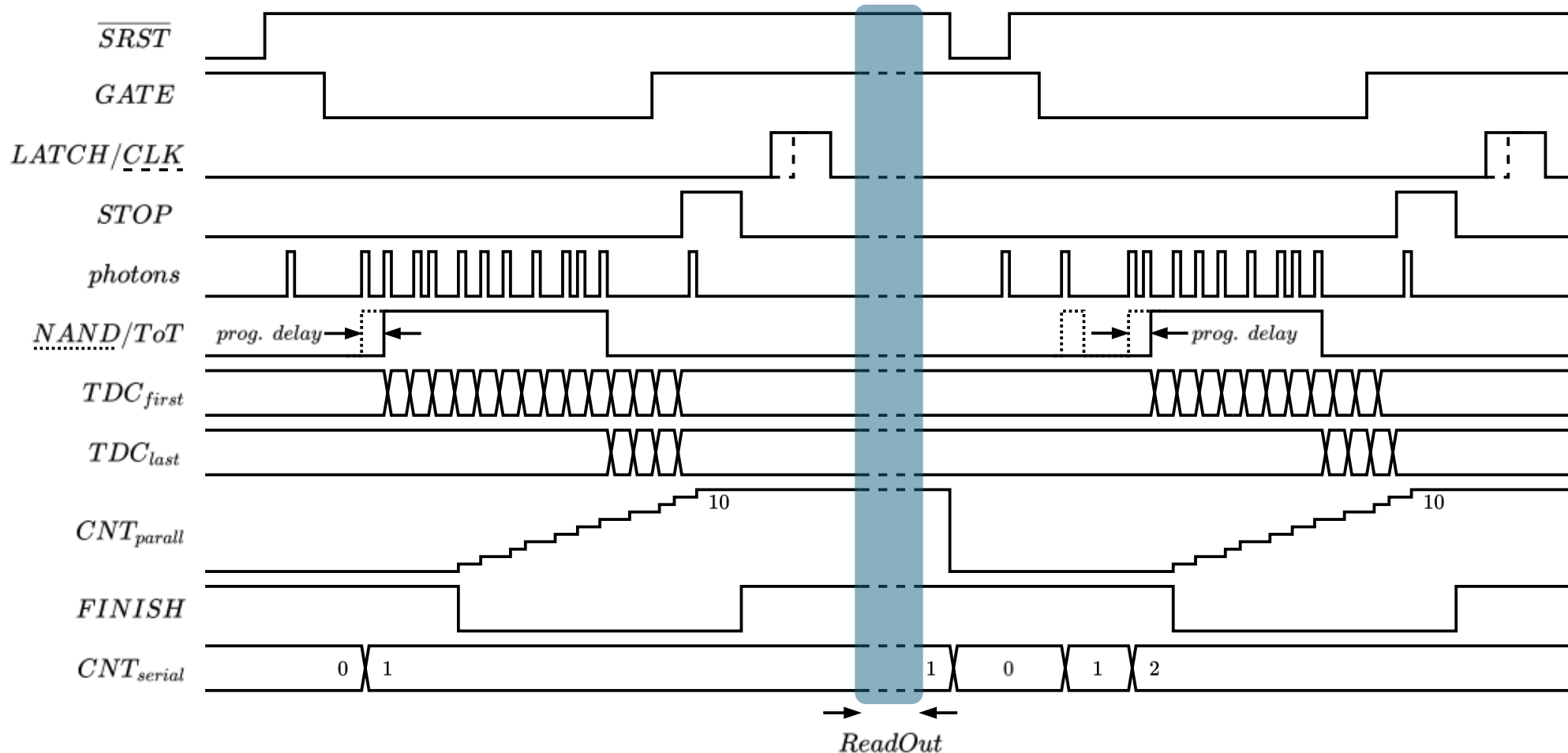
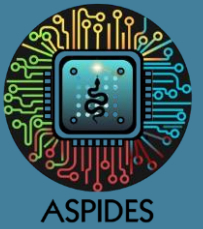
Finish signal goes to 0 while the output of the **parallel counter changes** its output.

Measurement Phase



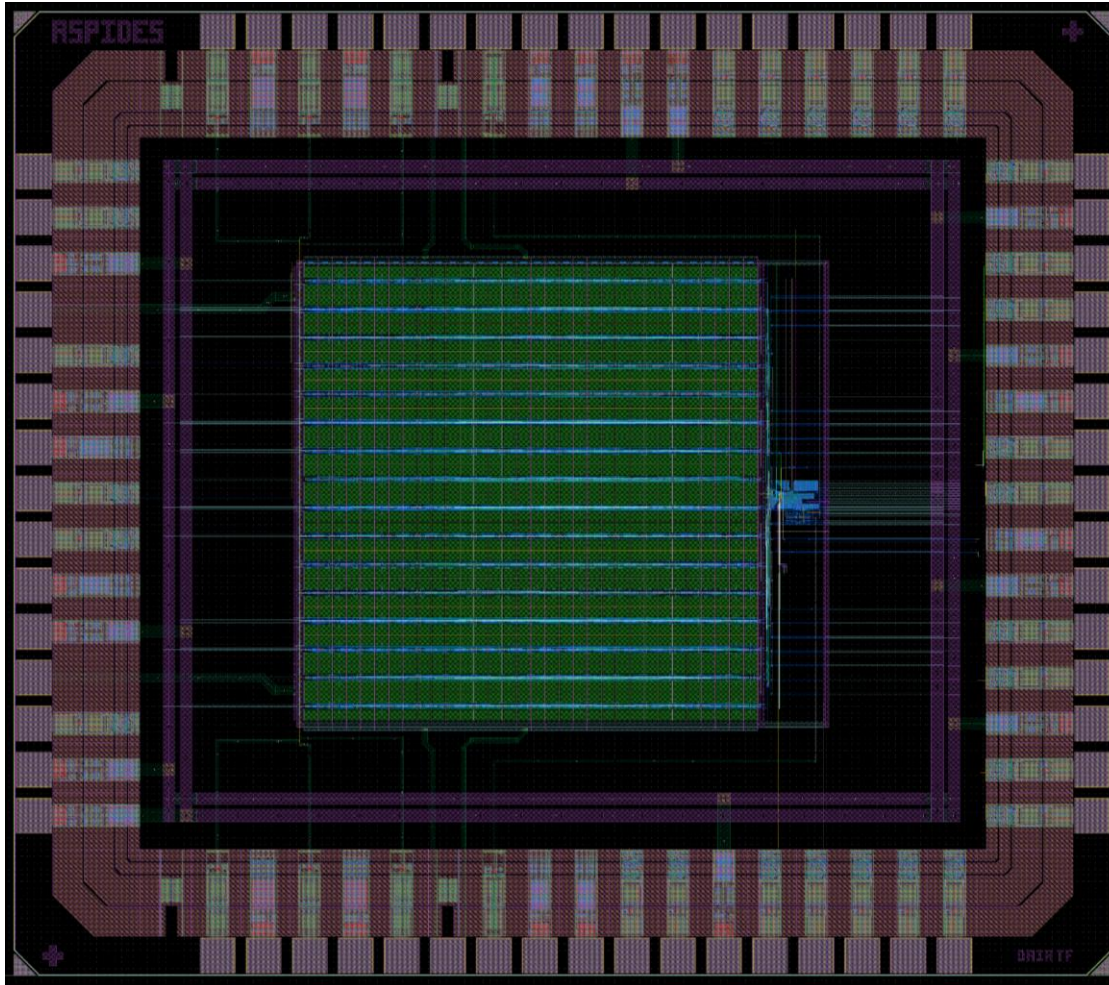
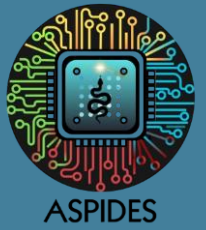
Latch and *Clock* signals are used to sample the data in the **output register for readout**.

Measurement Phase



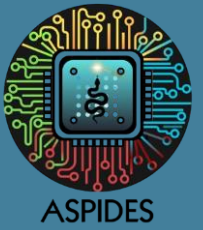
After each measurement, the data can be sent outside the chip for data processing or/and saving.

ASPIDES Prototype Chip



- 5 mm²
- 66 total pads
 - 28 digital pads
 - 14 outputs
 - 14 inputs
 - 36 analog pads
 - 8 bias voltages
 - 24 power supply
 - 4 SPAD bias
- 2 breakers

Conclusion



- Digital SiPM in 110nm CMOS is under development
 - **Photon counting** over three decades in less than 20ns by means of a parallel architecture.
 - ToA and ToT measurements with **100ps resolution** of the **first** and **last** detected events.
 - Different thresholding options available for **noise mitigation**.
 - Several **functions are configurable and tunable** to optimize performance.
- Next steps
 - A **prototype** has been **submitted** and will be **characterized** to validate the design – MPW run started in December 2025.
 - Based on the **characterization results**, a demonstrator will be designed and fabricated, integrating 8 dSiPMs in a single module.

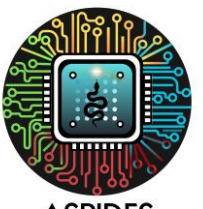


Thank you for your attention!

FEE 2026, XIII FRONT-END ELECTRONICS



UNIVERSITÀ
DI PAVIA



ASPIDES